

# Indice

1. La parità, la gara e la stagione nel tiro con l'arco
2. Sommario
3. Il giorno che si decide in 40 secondi
4. 1. Che cosa succede quando un match finisce cinque pari
5. 2. Che cosa deve decidere una regola di spareggio
6. 3. Un ripasso: la nuvola e la sua ampiezza
7. 4. Dal punteggio all'ampiezza, e ritorno
8. 5. Come si calcola un match che non si è mai giocato
9. 6. Il paradosso della parità
10. 7. Perché la freccia non perde nulla, e perché questo non basta
11. 8. Quanto si può guadagnare, al massimo
12. 9. I quattro tetti
13. 10. Quattordici modi di decidere, uno per uno
14. 11. I pesi che gli anelli meriterebbero
15. 12. Misurare le distanze: la strada che nessuno ha provato
16. 13. La cascata, e perché sembrava la risposta
17. 14. Il ribaltamento: quando contare i dieci indica l'arciere sbagliato
18. 15. La spiegazione, che è aritmetica e non statistica
19. 16. Dove il ribaltamento vale, e dove smette di valere
20. 17. La proposta, e perché funziona
21. 18. Come si scriverebbe l'articolo
22. 19. Quanto spettacolo si vuole comprare
23. 20. Se gli arcieri non fossero come li abbiamo descritti
24. 21. Una polizza, non una regola
25. 22. Le cinque regole che restano in piedi
26. 23. Il tabellone conta più della regola — ma quanto
27. 24. Che cosa portarsi a casa dalla Parte prima
28. 25. La gara è un oggetto solo
29. 26. Quanto sbaglia la qualificazione
30. 27. Il pedaggio del piazzamento
31. 28. Sette modi di costruire un tabellone
32. 29. La doppia eliminazione: giusta in teoria
33. 30. Il ripescaggio ha una dose
34. 31. Da dove far venire il piazzamento
35. 32. Cinque modi di giocare 15 frecce
36. 33. Che cosa vuol dire spettacolo, in modo che si possa contare
37. 34. La regola che nessun formato può violare
38. 35. Come si compongono i pezzi

39. 36. Il modello precisione  
40. 37. Il modello spettacolo  
41. 38. Il modello bilanciato  
42. 39. Il modello minimo: migliorare senza cambiare niente di visibile  
43. 40. I quattro a confronto  
44. 41. Che cosa portarsi a casa dalla Parte seconda  
45. 42. Il problema opposto  
46. 43. Che cosa vuol dire il migliore dell'anno  
47. 44. Quanto una stagione riesce a distinguere  
48. 45. Perché un punteggio non si confronta con un altro  
49. 46. La procedura, passo per passo  
50. 47. Un esempio percorso per intero  
51. 48. Il problema che nessuna quantità di dati risolve  
52. 49. Chi entra in classifica, e quando il premio è individuale  
53. 50. Che cosa conta davvero in un calendario  
54. 51. Quanti dati servono  
55. 52. Il confronto finale: quanto vale una stagione  
56. 53. Che cosa portarsi a casa dalla Parte terza  
57. 54. Il filo delle tre parti  
58. 55. Che cosa il modello non vede  
59. 56. Che cosa cambia in ciascuna conclusione  
60. 57. I limiti  
61. 58. Come un risultato è scaduto senza diventare falso  
62. 59. Quello che avevo previsto, e quello che mi ha smentito  
63. 60. Che cosa direi a chi deve decidere  
64. Principi  
65. Note di tecnica  
66. E, sopra tutti, il limite che nessuno di loro vede  
67. Il modello  
68. La regola ottimale sulle coordinate  
69. I pesi sugli anelli  
70. La parità a set  
71. Il rovesciamento del conteggio dei dieci  
72. La graduatoria  
73. Convenzioni

# Quando nessuno ha vinto

Matteo Campagna · [matteocampagna.com](http://matteocampagna.com)

Sommario

*Il giorno che si decide in 40 secondi*

## **PARTE PRIMA — LA PARITÀ**

- 1. Che cosa succede quando un match finisce cinque pari*
- 2. Che cosa deve decidere una regola di spareggio*
- 3. Un ripasso: la nuvola e la sua ampiezza*
- 4. Dal punteggio all'ampiezza, e ritorno*
- 5. Come si calcola un match che non si è mai giocato*
- 6. Il paradosso della parità*
- 7. Perché la freccia non perde nulla, e perché questo non basta*
- 8. Quanto si può guadagnare, al massimo*
- 9. I quattro tetti*
- 10. Quattordici modi di decidere, uno per uno*
- 11. I pesi che gli anelli meriterebbero*
- 12. Misurare le distanze: la strada che nessuno ha provato*
- 13. La cascata, e perché sembrava la risposta*
- 14. Il ribaltamento: quando contare i dieci indica l'arciere sbagliato*
- 15. La spiegazione, che è aritmetica e non statistica*
- 16. Dove il ribaltamento vale, e dove smette di valere*
- 17. La proposta, e perché funziona*
- 18. Come si scriverebbe l'articolo*
- 19. Quanto spettacolo si vuole comprare*
- 20. Se gli arcieri non fossero come li abbiamo descritti*
- 21. Una polizza, non una regola*
- 22. Le cinque regole che restano in piedi*
- 23. Il tabellone conta più della regola — ma quanto*
- 24. Che cosa portarsi a casa dalla Parte prima*

## **PARTE SECONDA — LA GARA**

- 25. La gara è un oggetto solo*
- 26. Quanto sbaglia la qualificazione*
- 27. Il pedaggio del piazzamento*
- 28. Sette modi di costruire un tabellone*
- 29. La doppia eliminazione: giusta in teoria*
- 30. Il ripescaggio ha una dose*
- 31. Da dove far venire il piazzamento*
- 32. Cinque modi di giocare 15 frecce*
- 33. Che cosa vuol dire spettacolo, in modo che si possa contare*

- 34. *La regola che nessun formato può violare*
- 35. *Come si compongono i pezzi*
- 36. *Il modello precisione*
- 37. *Il modello spettacolo*
- 38. *Il modello bilanciato*
- 39. *Il modello minimo: migliorare senza cambiare niente di visibile*
- 40. *I quattro a confronto*
- 41. *Che cosa portarsi a casa dalla Parte seconda*

## **PARTE TERZA — L'ARCIERE DELL'ANNO**

- 42. *Il problema opposto*
- 43. *Che cosa vuol dire il migliore dell'anno*
- 44. *Quanto una stagione riesce a distinguere*
- 45. *Perché un punteggio non si confronta con un altro*
- 46. *La procedura, passo per passo*
- 47. *Un esempio percorso per intero*
- 48. *Il problema che nessuna quantità di dati risolve*
- 49. *Chi entra in classifica, e quando il premio è individuale*
- 50. *Che cosa conta davvero in un calendario*
- 51. *Quanti dati servono*
- 52. *Il confronto finale: quanto vale una stagione*
- 53. *Che cosa portarsi a casa dalla Parte terza*

## **APPENDICE B — QUANDO LA FORMA CONTA**

### **CHIUSURA**

- 54. *Il filo delle tre parti*
- 55. *Che cosa il modello non vede*
- 56. *Che cosa cambia in ciascuna conclusione*
- 57. *I limiti*
- 58. *Come un risultato è scaduto senza diventare falso*
- 59. *Quello che avevo previsto, e quello che mi ha smentito*
- 60. *Che cosa direi a chi deve decidere*

## **NOTA SU DATI, FONTI E METODO**

## **APPENDICE — VENTUNO COSE IMPARATE SBAGLIANDO**

*Principi*

*Note di tecnica*

*E, sopra tutti, il limite che nessuno di loro vede*

## **APPENDICE MATEMATICA**

*Il modello*

*La regola ottimale sulle coordinate*

*I pesi sugli anelli*

*La parità a set*

*Il rovesciamento del conteggio dei dieci*

*La graduatoria*

*Convenzioni*

#### COME CITARE QUESTO LAVORO

Campagna, M. (2026). Quando nessuno ha vinto: la parità, la gara e la stagione nel tiro con l'arco. Versione 1.1, 11 ottobre 2026. Ricerca indipendente. <https://doi.org/10.5281/zenodo.23262110>

## La parità, la gara e la stagione nel tiro con l'arco

### Sommario

Nel tiro con l'arco olimpico un match a set può finire cinque a cinque, e a quel punto il regolamento chiede una freccia sola per decidere. Questo volume misura quanto quella freccia (e ogni altra regola di spareggio ammissibile) riesca a indicare l'arciere più bravo, e poi allarga la domanda al torneo e alla stagione. Il modello è un arciere che tira una nuvola gaussiana di ampiezza nota; da lì i match si calcolano per convoluzione esatta, i tornei per simulazione appaiata, e la graduatoria di stagione con una stima congiunta di livelli e difficoltà. I risultati principali: la regola in vigore si colloca nel terzo inferiore delle alternative esaminate, e nove delle quattordici considerate la superano sulla coppia di riferimento; leggere il totale del match prima della freccia vale 0,3 punti, meno di quanto valga registrare le distanze, e conserva quel vantaggio anche quando lo spareggio è tirato peggio del resto del match; il piazzamento conta più della struttura del tabellone, e la graduatoria di stagione usata come piazzamento è l'intervento più efficace fra quelli che non richiedono di tirare una freccia in più; e le domande su cui il dibattito si concentra (la lunghezza dei set, l'ampiezza del ripescaggio) spostano i risultati di frazioni di match su sessantatré.

### Il giorno che si decide in 40 secondi

L'ultima freccia del quinto set è a terra, i giudici hanno finito di contare, e sul tabellone compare il numero che nessuno dei due voleva vedere: cinque a cinque. Non c'è un sesto set da giocare, perché il regolamento ne prevede cinque e cinque ne sono stati giocati. Accade allora una cosa che in nessun altro momento della gara accade: i due arcieri tornano in linea con una freccia sola in faretra, e quella freccia decide il match.

Dura 40 secondi. Ci finiscono dentro il turno, la medaglia, e talvolta il lavoro di un quadriennio.

Chiunque frequenti questo sport ha sentito le due obiezioni che seguono, e sono sempre quelle. Da un lato c'è chi sostiene che 40 secondi non possano annullare due ore di gara, e che affidare a un colpo solo una decisione del genere sia un sorteggio con qualche passaggio in più. Dall'altro c'è chi risponde che ti-

rare bene sotto quella pressione è precisamente il mestiere, e che il più forte lo dimostra proprio lì, dove tutto pesa di più.

Sono argomenti seri entrambi, e si fronteggiano da quando il formato a set è stato introdotto. Quello che non ho trovato, in nessuna delle due parti, è un numero.

Eppure una regola di spareggio non chiede di essere giudicata: chiede di essere misurata. Serve a rispondere a una domanda precisa, quale dei due arcieri è il più bravo, e a un attrezzo del genere si possono porre domande che a un'opinione non si possono porre. Non se sia giusto, ma ogni quante volte sbagli. E soprattutto, la domanda che al dibattito manca del tutto: se il margine per fare meglio sia grande abbastanza da valere la discussione che gli si dedica.

### **Perché il libro non si ferma alla parità**

Questo volume nasce da lì, e poi si allarga da solo. Una volta misurato quanto valga la regola di parità viene naturale chiedersi quanto valga tutto il resto: il formato del match, la struttura del tabellone, il round di qualificazione che decide chi incontra chi. E una volta fatto quel conto viene naturale la terza domanda, che è quella che di solito si pone per prima e che invece va posta per ultima: chi è, a fine anno, l'arciere più preciso, e con quale diritto lo si dice.

Sono tre domande che sembrano diverse e sono la stessa a tre scale. Un istante, una giornata, una stagione. In tutte e tre si tratta di stabilire quanta informazione ci sia davvero nei dati che abbiamo, e quanta ne butti via il modo in cui li leggiamo.

### **Le risposte, subito**

Non c'è ragione di far aspettare il lettore, quindi metto qui in apertura quello che le pagine seguenti dimostrano.

La regola di spareggio in vigore è battuta da 9 alternative su 14 fra quelle che ho misurato, e la migliore fra quelle adottabili domani sta in una riga di regolamento: vince chi ha totalizzato più punti nel match, e solo se anche i totali sono pari si tira una freccia. Il criterio che la tradizione considera più naturale, contare i dieci, indica l'arciere sbagliato più spesso di quanto indichi quello giusto ogni volta che lo si consulta dopo una parità, ed è l'unico modo in cui lo si consulta. Non per un difetto della regola, ma per una ragione aritmetica che si dimostra in due tabelle e che, una volta vista, è ovvia.

Tutto questo però conta poco, ed è la seconda cosa che queste pagine misurano. In una gara da 64 arcieri si giocano 63 match: passare dalla regola in vigore alla migliore possibile cambia il vincitore in meno di un match, e cambiare il formato con cui si giocano le 15 frecce lo cambia in un terzo di match. Detto in modo che si veda: bisogna guardare tre gare intere prima che quel cambio di formato sposti un solo arciere da una parte all'altra del tabellone.

Quello che conta davvero è da dove viene il piazzamento con cui si costruisce il tabellone. Oggi il titolo va al più bravo del campo 9,7 volte su cento, cioè una gara su dieci. Usando la classifica di stagione invece del punteggio del giorno, e ricomponendo gli accoppiamenti a ogni turno, si passa a 13,7 volte su cento: una gara su sette. Su cento tornei sono quattro titoli in più assegnati alla persona giusta, e non costano una freccia, un minuto o uno spareggio a nessuno.

Chi non volesse cambiare nemmeno quello ha comunque una strada, ed è il quarto dei modelli di gara che queste pagine descrivono: si possono guadagnare 1,1 punti, cioè un titolo in più ogni novanta gare,

senza toccare niente di ciò che si vede (stessa qualificazione, stesso tabellone, stessa durata) cambiando soltanto come si leggono i risultati che la gara produce già. Un titolo in più ogni novanta gare, senza che nessuno sugli spalti si accorga di niente. Sono due interventi, e tutti e due chiedono di registrare dove ogni freccia sia finita dentro l'anello: ordinare la qualificazione per distanza dal centro, che da solo vale 0,5, e risolvere le parità sulla stessa misura, che da solo ne vale 0,7. Insieme danno 1,1, meno della loro somma, perché in parte comprano la stessa cosa.

Chi non voglia procurarsi strumentazione ha un'alternativa più economica sulla sola parità, che vale 0,33 punti sul titolo e costa una riga di regolamento: leggere il totale del match prima della freccia di spareggio. È appunto un'alternativa, perché le due regole di parità non possono valere sullo stesso match.

E nemmeno tutto questo basta, il che è la cosa che riordina la discussione. In un torneo a eliminazione diretta il titolo va al migliore del campo 9,7 volte su cento oggi, 14,2 adottando tutto il meglio proponibile con le attrezzature esistenti, 14,6 se si misurassero le distanze. Anche facendo ogni cosa nel modo migliore, quindi, in sei gare su sette il titolo va a qualcuno che non è il più bravo del campo. Una stagione di gare, letta come si deve, individua il migliore tre volte su quattro: la differenza fra le due cifre è il vero risultato di questo libro. Il torneo non sbaglia: gli si chiedono due cose insieme, e nessuna delle due gli riesce del tutto perché sono in conflitto.

### **Come vanno letti i numeri che seguono**

Il modello lascia fuori tre cose reali: la pressione del momento, il vento, il peso della posta in gioco. Le ho misurate a parte, e il capitolo 55 fa quel conto per tutte e tre le parti insieme, che è il posto giusto perché il modello è uno solo. La domanda che ho posto a ciascuna non è se sposti i numeri, che li sposta di sicuro, ma in che direzione sposti ciascuna conclusione.

La risposta è che quasi sempre le aggravano invece di indebolirle. Se qui trovo che una regola distingue poco, sul campo distinguerà ancora meno; se trovo che il titolo va al migliore una gara su dieci, in una giornata di vento ci va più di rado; se trovo che il piazzamento di qualificazione sbaglia di otto posti, con il vento ne sbaglia di più. I numeri di queste pagine sono dunque un riferimento sotto le assunzioni dichiarate, e per le perturbazioni che ho modellato (code pesanti, gruppo scenterato, gruppo ovale, deriva dentro il match, vento, pressione) le misure si riducono. Non ne segue che il campo reale debba stare sotto: una simulazione non osserva il campo reale, e qualità correlate con la bravura potrebbero spingere nell'altro verso. Fuori da quella lista non lo dimostro, e dove il verso non è provato lo dico.

Le eccezioni che ho trovato sono due, e le dichiaro qui perché sono le volte in cui un fattore escluso lavora contro una mia conclusione. La prima: il vento attenua il ribaltamento del conteggio dei dieci di cui parla il capitolo 14, e non lo annulla (resta sotto la monetina anche con vento forte) ma lo rende meno marcato, e al capitolo 22 riporto di quanto. La seconda: il vento riduce anche quanto spesso i match finiscano cinque pari, e quindi quante volte la regola di parità venga chiamata in causa, il che attenua l'effetto di ogni intervento su quella regola senza annullarlo.

### **Come è organizzato questo saggio**

Tre parti, e ciascuna finisce con un capitolo che riassume che cosa portarsi a casa: chi voglia le conclusioni senza le derivazioni può leggere quei tre e la chiusura. La Parte prima costruisce anche gli strumenti che servono alle altre due, quindi i suoi primi cinque capitoli valgono per tutto il volume. La chiu-

sura contiene, oltre alle raccomandazioni, il conto dei fattori che il modello ignora e l'elenco di ciò che questo lavoro non può dire.

# PARTE PRIMA — LA PARITÀ

---

## 1. Che cosa succede quando un match finisce cinque pari

Conviene partire dal principio e descrivere per intero come si arrivi a quel cinque a cinque. Chi non frequenta i campi di gara conosce la scena di riflesso, e chi li frequenta la dà per talmente ovvia da non guardarla più, che è di solito la condizione in cui si sbaglia.

Nel recurve individuale la sfida è al meglio di 5 set, e ciascun set vale 3 frecce a testa. Si sommano i punti delle 3 frecce, e chi ha fatto di più incassa due punti-set; se i totali coincidono, uno ciascuno; chi ha fatto di meno, niente. Vince il match chi arriva per primo a 6 punti-set. La somma di tutte le frecce non conta: contano i set. Un arciero può tirare quindici punti in più dell'avversario lungo tutto il match e perderlo lo stesso, se quei quindici punti se li è presi tutti in un set solo, che vince largamente, e negli altri quattro perde di misura.

Il formato esiste per una ragione precisa. Il sistema precedente sommava tutte le frecce e assegnava la vittoria al totale più alto, e produceva match in cui un arciero prendeva il largo al terzo turno e il resto era una formalità: nessuna tensione, nessuna rimonta possibile, e un pubblico che se ne andava prima della fine. Con i set, ogni 3 frecce si riparte da zero, e un set perso di dieci punti costa esattamente quanto un set perso di uno. La rimonta resta aperta fino all'ultima freccia.

Il prezzo di questa scelta è aritmetico e non lo si può aggirare. Cinque set distribuiscono dieci punti, due per set; se dopo 5 set nessuno dei due ha raggiunto sei, l'unica ripartizione possibile è cinque a cinque. E a quel punto non c'è modo di proseguire, perché i set previsti sono cinque e cinque ne sono stati giocati.

### Che cosa dice il regolamento

A quel punto il regolamento chiede una freccia a testa: vince chi fa il punteggio più alto, e se il punteggio è pari decide la freccia più vicina al centro. Se i giudici non riescono a distinguerle, si torna in linea e si ripete. Nel modello concentrico che uso qui le due letture coincidono per le frecce che segnano punti, perché a punteggio più alto corrisponde distanza minore, e restano distinte solo quando una delle due esce dalla zona di punteggio.

Il sorteggio non è previsto in alcun passaggio. Ho verificato entrambe le cose alla fonte, sul regolamento della federazione internazionale, nella versione pubblica, e lo dico perché in un lavoro precedente avevo ricostruito a memoria una clausola simile e mi ero sbagliato.

Nelle finali internazionali si tira alternati, un arciero per volta con 20 secondi a disposizione ciascuno: uno spareggio sono quindi due tiri e poco più di 40 secondi di cronometro, ai quali si aggiunge il tempo che i giudici impiegano a stabilire quale delle due frecce sia più interna. Non c'è nulla, in questo sport, che si decida più in fretta, e poche cose pesano altrettanto.

C'è un terzo articolo, che sembra un dettaglio e non lo è: nei match alternati il meglio classificato della qualificazione sceglie l'ordine del primo set (non è detto che scelga di tirare per primo) e allo spareggio

parte chi aveva tirato per primo in quel set. Ne segue una conseguenza su cui tornerò alla fine, e che riguarda non tanto la regola quanto la possibilità stessa di studiarla.

### **Quanto spesso capita**

È il primo numero di queste pagine, e conviene fissarlo perché tutto il resto ci si appoggia. Fra due arcieri di livello simile (dieci punti di differenza su 720, che nel tiro con l'arco d'élite è una distanza reale ma non enorme) un match su sei finisce cinque pari.

In una gara da 64 arcieri si giocano 63 match, quindi in un pomeriggio si tirano circa dieci spareggi. Chi segue una gara internazionale dalla prima eliminatoria alla finale vede quindi il regolamento chiedere quella freccia una decina di volte, e ogni volta manda a casa qualcuno. Non è un caso da manuale: è una scena che chiunque frequenti questo sport ha visto decine di volte.

---

## **2. Che cosa deve decidere una regola di spareggio**

Contare gli errori richiede però di aver deciso prima che cosa sia un errore, e la decisione va presa qui, prima di guardare qualunque risultato. La trappola è di quelle che si vedono soltanto dopo esserci caduti: se stabilissimo che il più bravo è chi ha vinto, allora nessuna regola potrebbe mai sbagliare, vincerebbe sempre il più bravo per definizione, e la domanda che ci siamo posti si scioglierebbe fra le mani.

Il criterio che adottato è quindi uno solo, e non ne aggiungo altri.

*Il più bravo è chi, mettendosi in linea, si attende in media il punteggio più alto.*

Chiunque alleni obietterà subito che manca tutto il resto, la tenuta quando il match si stringe, la lettura del vento, la capacità di produrre il tiro che serve nel momento in cui serve, e l'obiezione è fondata. Quelle qualità restano fuori, e restano fuori proprio là dove contano di più, perché non esiste tiro più gravato di quello di spareggio.

In cambio della povertà, però, il criterio ha una qualità che le altre non hanno: si può controllare. Il punteggio atteso è una grandezza che si rileva, che si mette a confronto e che si può seguire di stagione in stagione, e due persone che discutano possono trovarsi d'accordo su quel numero. Le altre doti esistono eccome, ma nessuno dispone di un metro su cui si sia d'accordo per misurarle: assumerle come definizione vorrebbe dire calibrare uno strumento servendosi di un secondo strumento a sua volta non calibrato.

Non è una scelta neutra, e va detto. Se si crede che un torneo debba premiare la tenuta nervosa più della precisione, allora questo lavoro misura la cosa sbagliata, e la risposta corretta non è correggerne i numeri ma respingerne la definizione, sapendo però che si sta respingendo la sola qualità che questo lavoro ha imparato a misurare.

C'è anche una ragione più fine per aver scelto il punteggio atteso e non, per esempio, la larghezza del gruppo. Le due coincidono finché l'arciere tira centrato, e divergono appena non lo fa: un arciere con un gruppo strettissimo ma spostato di due centimetri dal centro ha la nuvola più piccola e il punteggio peggiore, e in quel caso il criterio giusto è il secondo, perché una federazione premia il merito e non la geometria.

### **Come leggere i numeri di questo libro**

Tutte le cifre che seguono rispondono alla stessa domanda: *ogni quante volte la regola manda avanti l'arciere più bravo*. Quando scriverò che una regola vale 57, intenderò che su cento match finiti in parità ne indirizza correttamente cinquantasette e sbaglia negli altri quarantatré.

Gli estremi della scala vanno fissati adesso, perché ogni numero che verrà va letto dentro quelli. In basso c'è il *50 su cento*: è quello che si ottiene lanciando una moneta senza guardare il bersaglio. Va detta bene una cosa che si sente ripetere: se si sa che una regola indovina meno di una volta su due, applicarla al contrario la riporta sopra, ma sapere il verso è precisamente ciò che non si ha in anticipo, e il capitolo 14 mostra un criterio plausibile che, nella situazione in cui lo si consulta, sta sotto la moneta senza che nessuno se ne fosse accorto. In alto c'è il *cento su cento*, che con 15 frecce non si raggiunge mai, ma serve come riferimento.

C'è però un secondo modo di dire la stessa cosa, e serve perché il primo nasconde un difetto. Una regola può indovinare dieci volte in più di un'altra su una situazione che capita una volta all'anno, e allora quelle dieci volte non cambiano niente. Il secondo modo tiene conto anche di quanto spesso l'occasione si presenti, e lo esprime contando i match di una gara intera: 64 arcieri, 63 match giocati.

Quando dirò che una regola vale un match e mezzo, intenderò questo: giocando una gara completa, ci sono un match e mezzo in cui il vincitore è diverso da quello che uscirebbe tirando una monetina su ogni parità. Metà volta non esiste, naturalmente, e va letta come una media su più gare: un match e mezzo a gara significa tre match ogni due gare, cioè tre arcieri che passano il turno per merito invece che per caso, con tutte le conseguenze che un cambio di match ha sul proseguimento dello stesso.

Le due letture non si convertono l'una nell'altra, e questa è la ragione per cui le riporto entrambe. Spostare il vincitore di un match qualsiasi non è la stessa cosa che spostare quello della finale: se il match che cambia è un ottavo, cambia chi va ai quarti; se è la finale, cambia chi vince la gara.

Quando parlerò della gara nel suo complesso userò invece una terza formulazione, ancora più semplice: ogni quante gare il titolo va davvero al più bravo del campo. Oggi, come vedremo nella Parte seconda, è una gara su dieci.

---

### 3. Un ripasso: la nuvola e la sua ampiezza

Da qui in avanti serve un attrezzo solo, ed è questo: la macchia di frecce che un arciere lascia sul bersaglio si può riassumere in un numero singolo.

Il perché è meno esotico di quanto suoni. Prendiamo un arciere che non abbia difetti sistematici, che non tiri cioè abitualmente basso, o a sinistra, o largo in una direzione riconoscibile: i suoi errori si distribuiscono allora in modo indifferente attorno al centro, tanto a destra quanto a sinistra, tanto in alto quanto in basso, e con la stessa intensità in ogni verso. La macchia che ne risulta è tonda, e ha il centro dove l'arciere mirava.

Stando così le cose, sapere in che direzione una freccia si sia allontanata non aiuta a riconoscere chi l'abbia tirata: sinistra e destra dicono la stessa cosa, e ciò che distingue un arciere da un altro è soltanto la distanza tipica a cui le frecce si fermano dal centro. Quel *quanto lontano* è un numero solo: lo si chiama ampiezza della nuvola, si misura in centimetri, e in statistica si indica con la lettera greca sigma. Nuvola stretta, ampiezza piccola, arciere fortissimo; nuvola larga, ampiezza grande, arciere che sbaglia di più.

Conviene fissare subito la scala con qualcosa che si possa guardare. Un arciere la cui nuvola misura quattro centimetri e mezzo mette due frecce su tre dentro un cerchio largo come un piattino da caffè, poco più di tredici centimetri. Il numero da solo non lo suggerisce, e va tenuto a mente ogni volta che se ne legge uno: la macchia è sempre più larga di quanto l'ampiezza faccia pensare.

## **Il bersaglio**

Dall'altra parte c'è il bersaglio, e anche quello è semplice. La faccia usata a 70 metri è larga 122 centimetri ed è divisa in dieci corone circolari tutte della stessa larghezza, 6,1 centimetri ciascuna: la più interna vale dieci, la successiva nove, fino all'ultima che vale uno. Dentro il dieci c'è un cerchietto più piccolo, di 3,05 centimetri di raggio, chiamato X, che vale dieci punti come il resto della corona, ma si conta a parte e serve proprio a rompere le parità.

Mettendo insieme le due cose si ottiene tutto: sapere quanto è larga la macchia significa sapere quanto spesso una freccia finisca in ciascuna corona, e sapere quello significa sapere che punteggio ci si debba attendere. La relazione fra le due si scrive in una riga, ed è l'unica formula che compare nel corpo di queste pagine perché è la sola di cui serva vedere la faccia:

$$F(r) = 1 - e^{(-r^2 / 2\sigma^2)}$$

Si legge un pezzo per volta.  $F(r)$  è la probabilità che una freccia cada entro una distanza  $r$  dal centro: vale zero se  $r$  è zero, perché nessuna freccia colpisce il punto esatto, e tende a uno per  $r$  grande, perché allargando abbastanza il cerchio prima o poi la freccia ci sta dentro di sicuro. Fra i due estremi la curva sale, e  $\sigma$  governa quanto in fretta.

Il modo concreto di leggerla è questo: un arciere da 700 punti piazza dentro il dieci (i sei centimetri centrali) settantadue frecce su cento, mentre uno da 630 ne piazza ventidue. Su una volée da 36 frecce sono ventisei gialli contro otto, e la differenza si vede da bordo campo senza contare niente. I dettagli sono in appendice, formula S1, e la traduzione dall'ampiezza al punteggio è la S2; da qui in avanti la formula non serve più, servono i numeri che produce.

La nuvola e gli anelli — sessanta frecce per ciascun livello, faccia da 122 cm, corone da 6,1 cm

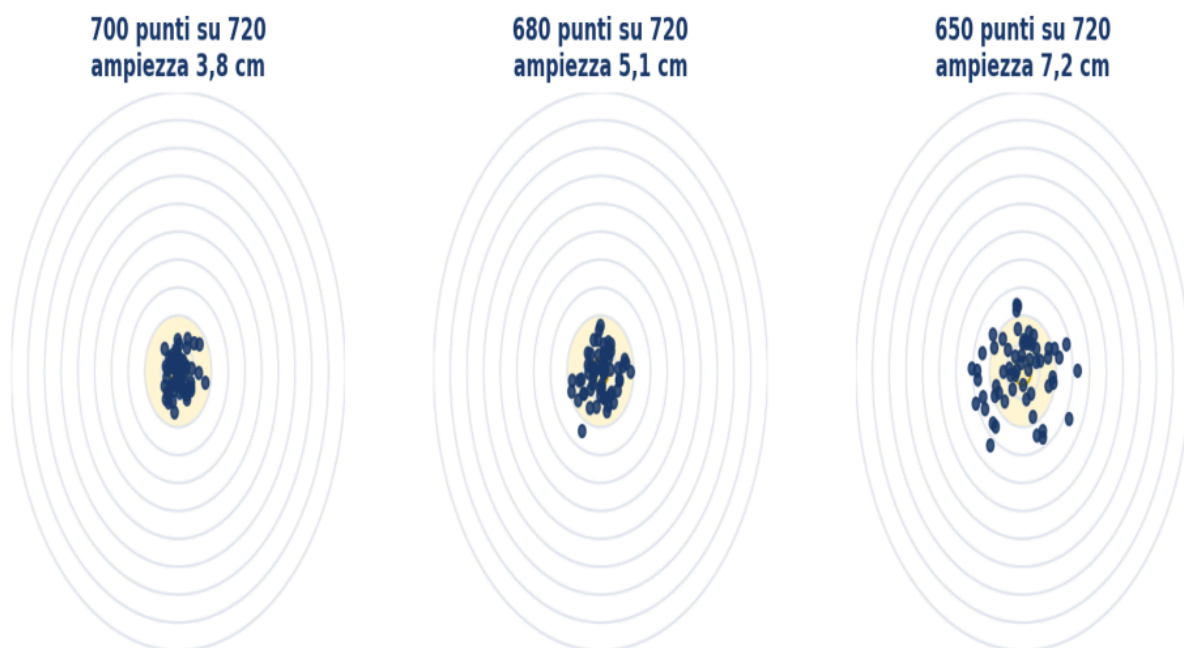


**Figura 1 — La nuvola e gli anelli.** Sessanta frecce simulate per ciascuno di tre livelli di bravura — tre arcieri da 700, 680 e 650 punti su 720 — sulla faccia da 122 cm. Le corone hanno larghezza 6,1 cm, l’anello X raggio 3,05 cm. Rappresentazione schematica originale a fini didattici. Le proporzioni non vengono da un dataset e non sono una misura: sono scelte per rendere leggibile la geometria del bersaglio, e nessuna conclusione del volume vi si appoggia.

#### 4. Dal punteggio all’ampiezza, e ritorno

La relazione si percorre in entrambe le direzioni, ed è così che la useremo: da un punteggio si risale all’ampiezza della nuvola, e da quella si calcola tutto il resto. Il round di qualificazione a 70 metri prevede 72 frecce, quindi un massimo di 720 punti.

| Punteggio atteso su 720 | Ampiezza della nuvola | Chi è, all’incirca             |
|-------------------------|-----------------------|--------------------------------|
| 700                     | 3,78 cm               | fra i primi al mondo           |
| 690                     | 4,46 cm               | finalista internazionale       |
| 680                     | 5,14 cm               | quartifinalista internazionale |
| 670                     | 5,81 cm               | nazionale di buon livello      |
| 660                     | 6,49 cm               | ottimo regionale               |
| 645                     | 7,50 cm               | buon agonista                  |
| 630                     | 8,52 cm               | agonista                       |

*(Ampiezza corrispondente a ciascun punteggio medio su 72 frecce a 70 metri, ottenuta invertendo numericamente la relazione del capitolo precedente. La colonna di destra è indicativa e serve solo a dare la scala.)*

Questa tabella contiene già una sorpresa, e vale la pena fermarsi. Dieci punti su 720 sembrano poco, l'1,4 per cento del massimo, e corrispondono invece a una differenza di ampiezza del 18 % se si parte da 700 e del 15 % se si parte da 690. Il divario in ampiezza dipende cioè dal livello a cui lo si misura, e per questo ogni volta dirò da dove parto.

Si vede meglio mettendo i due gruppi uno accanto all'altro sul bersaglio. Il finalista da 690 raccoglie due frecce su tre in un cerchio largo tredici centimetri; al quartfinalista da 680 ne serve uno largo quindici, cioè due centimetri in più. Su un bersaglio è una differenza che un allenatore riconosce a colpo d'occhio, e statisticamente è un divario enorme: due arcieri distanti dieci punti non sono due arcieri simili, sono due arcieri che un buon strumento distingue quasi sempre. Quando più avanti dirò che una coppia è vicina intenderò due o tre punti, non dieci.

### **Perché ogni numero porta con sé il suo livello**

C'è un'avvertenza che vale per tutto il resto, e che è la ragione per cui ogni numero di queste pagine porta con sé il livello a cui si riferisce. La traduzione da punti ad ampiezza dipende dal livello: un punto vale l'1,81 % di ampiezza partendo da 700, l'1,52 % partendo da 690, l'1,32 % partendo da 680. Lo stesso punto, in cima alla classifica, pesa di più, perché la nuvola è già piccola e toglierne un punto richiede di allargarla di una frazione maggiore.

Concretamente, questo vuol dire che guadagnare un punto a 700 è più difficile che guadagnarlo a 680, e chi allena lo sa senza bisogno di formule. Senza quell'indicazione lo stesso numero cambia del 40 per cento, e chi lo rileggesse otterrebbe un'altra cosa.

### **La coppia di riferimento**

Molti numeri dipendono poi da quali due arcieri si stiano confrontando, quindi ne fisso una coppia sola e la userò ovunque possibile: **un arciere da 690 contro uno da 680**. Dieci punti di differenza, cioè la distanza fra un finalista e un quartfinalista in una gara internazionale, una coppia realistica.

Le sue due nuvole misurano 4,46 e 5,14 centimetri, il loro rapporto è 1,15, e il match fra i due finisce cinque pari il 16,8 per cento delle volte: quasi un match su sei, che in un tabellone da 64 fa 10,6 match su 63. Ogni volta che un numero comparirà senza altra indicazione sarà calcolato su questa coppia.

---

## **5. Come si calcola un match che non si è mai giocato**

Arriviamo al passaggio che di regola nessuno racconta, e che invece è la parte più onesta di questo mestiere. Se quel match non è mai stato disputato, su che base si afferma quante volte una regola ne indichi il vincitore giusto?

La via che si presenta subito è quella della simulazione: si istruisce un calcolatore a fingere di tirare, gli si fanno disputare 10 milioni di incontri, e si contano gli esiti. È legittima, in molti campi è l'unica disponibile, e ha un difetto: il risultato porta con sé un margine di incertezza, perché 10 milioni di prove restano un numero finito. Quando la differenza da misurare è di tre match su mille, e in queste pagine capiterà spesso, il rumore diventa il problema principale. Dove ho potuto, quindi, non ho simulato: ho contato.

L'idea si spiega senza matematica, ed è la stessa con cui si calcola la probabilità che due dadi diano sette: non si tirano i dadi, si contano i casi. Supponiamo di voler sapere con che frequenza la freccia di A si fermi più vicina al centro di quella di B. Non c'è bisogno di tirare: sappiamo, per ciascuna distanza dal centro, con che probabilità A ci arrivi, e sappiamo la stessa cosa di B; allora per ogni possibile distanza di A si calcola quanto sia probabile che B cada più lontano, e si sommano tutti i casi pesando ciascuno per la sua probabilità.

Un set è la somma di 3 frecce, e se so quanto spesso una freccia vale dieci, quanto spesso nove e così via, posso calcolare quanto spesso 3 frecce sommino a 28, a 27, a 26: basta elencare tutti i modi di ottenere ciascun totale e sommarne le probabilità. Da lì si ricava quanto spesso A vinca un set, quanto spesso pareggi, quanto spesso perda. E un match è una successione di set, che si percorre come una scacchiera con pochi pezzi: si comincia dall'esito finale, che è noto, e si torna indietro un set alla volta calcolando con che probabilità ci si trovi in ciascuna situazione intermedia.

C'è un dettaglio che semplifica parecchio i conti, e non è ovvio: nessuno dei 51 percorsi che portano al cinque pari tocca mai sei punti prima della fine, perché con al massimo due vittorie il punteggio più alto raggiungibile prima dell'ultimo set è cinque. La regola che ferma il match a sei non morde quindi mai sulla fetta che ci interessa.

Questo modo di procedere ha una conseguenza che conviene mettere in chiaro, ed è la ragione per cui i numeri che seguono si possono attaccare con profitto: sono riproducibili all'ultima cifra. Chi rifacesse i conti domani, con la stessa specifica, non otterrebbe risultati simili ai miei, otterrebbe i miei.

## 6. Il paradosso della parità

Prima di guardare che cosa succeda al cinque pari conviene guardare che cosa succederebbe senza, ed è il modo più rapido per capire dove stia davvero la difficoltà. Immaginiamo per un momento che non esistano regolamenti, né set, né spareggi: soltanto due arcieri, un certo numero di frecce a testa, e la libertà di misurare la distanza esatta di ciascuna dal centro. Chi ha il gruppo più stretto?

| Divario di ampiezza                      | 3 frecce a testa | 6           | 15          | 30          | 72          |
|------------------------------------------|------------------|-------------|-------------|-------------|-------------|
| 1 %                                      | 50,9             | 51,3        | 52,2        | 53,1        | 54,7        |
| 2 %                                      | 51,9             | 52,7        | 54,3        | 56,1        | 59,4        |
| 5 %                                      | 54,6             | 56,6        | 60,4        | 64,7        | 72,1        |
| 10 %                                     | 58,9             | 62,7        | 69,8        | 76,9        | 87,3        |
| <b>15,2 % — la coppia di riferimento</b> | <b>63,0</b>      | <b>68,4</b> | <b>77,8</b> | <b>86,1</b> | <b>95,4</b> |
| 20 %                                     | 66,5             | 73,1        | 83,8        | 92,0        | 98,5        |

*(Quante volte su cento la lettura ottimale indica l'arciere più bravo. Le colonne contano le frecce di ciascun arciera, non il totale delle due: la colonna da 15 è quindi il match come si gioca oggi, e quella da 72 è un round di qualificazione. Su tutti i match e non soltanto su quelli in parità; calcolo esatto in forma chiusa, formula S5.)*

La riga in grassetto è la nostra coppia di riferimento, e dice una cosa che vale la pena rileggere due volte. Guardando le 15 frecce che ciascuno tira in un match, e leggendole nel modo migliore possibile, si riconosce il più bravo 78 volte su cento: su cento match sbaglia ventidue volte, cioè poco più di una su cin-

que. Bastano sei frecce a testa per arrivare a 68, che è già molto sopra la monetina, e con le 72 di un round di qualificazione si arriva a 95, cioè si sbaglia una volta ogni venti.

La conclusione, quindi, è che senza il filtro della parità il problema sarebbe quasi facile: un match contiene informazione abbondante per distinguere due arcieri che stiano a dieci punti l'uno dall'altro. Tutta la difficoltà di cui parliamo non nasce dunque dalla scarsità dei dati, e questo è il primo fatto sorprendente di queste pagine, perché è esattamente il contrario di quello che si sente dire al campo.

### Da dove nasce, allora

Quando un match finisce cinque pari non stiamo guardando due arcieri qualunque, ma due arcieri selezionati dall'aver prodotto punteggi indistinguibili per 5 set consecutivi. Il cinque pari non è un evento neutro che capita: è una lente che ci consegna esattamente i casi in cui i dati non hanno saputo decidere: è un po' come chiedersi quanto sia affidabile una bilancia guardando soltanto le volte in cui ha segnato lo stesso peso per due oggetti diversi, che per costruzione sono le volte in cui ha detto meno.

L'effetto si misura, e conviene misurarlo sulla coppia di riferimento perché è quella che useremo per tutto il resto.

| Regola                            | Su tutti i match | Sui soli match cinque pari | Quanto ha perso      |
|-----------------------------------|------------------|----------------------------|----------------------|
| Somma dei quadrati delle distanze | 77,8             | 68,8                       | un terzo             |
| Distanza media semplice           | 76,9             | 67,1                       | poco più di un terzo |
| Conteggio pesato degli anelli     | 75,1             | 62,7                       | quasi la metà        |
| Punteggio totale del match        | 74,3             | 60,5                       | poco più della metà  |
| Conteggio dei dieci               | 71,1             | 53,6                       | cinque sestì         |
| Freccia di spareggio              | 57,0             | 57,0                       | nulla                |

*(Coppia di riferimento 690 contro 680. «Quanto ha perso» è la quota di margine sopra il 50 per cento che il condizionamento porta via: una regola che passi da 70 a 60 ha perso metà dei venti punti che aveva.)*

La colonna di destra è quella che racconta la storia, e conviene leggerla capendo che cosa misuri. Una regola che indovini 70 volte su cento fa venti volte meglio della monetina; se guardando soltanto i match finiti pari scende a 60, di quelle venti volte gliene restano dieci, cioè ne ha persa la metà. È in quel senso che dico quanto ha perso: non quanto sia sceso il numero, ma quanta parte del suo vantaggio sulla monetina sia svanita.

Guardata così, la tabella dice tre cose. Le regole che leggono le distanze perdono un terzo di quel vantaggio, quelle che leggono gli anelli ne perdono circa la metà, e il conteggio dei dieci ne perde cinque sestì: da 71 scende a 53,6, che vuol dire indovinare cinquantaquattro volte su cento dove una moneta ne indovinerebbe cinquanta. È il primo segnale di una cosa che il capitolo 14 mostrerà per intero.

La freccia di spareggio, invece, non perde niente. Sta a 57 prima e a 57 dopo, ed è l'unica delle sei di cui questo si possa dire.

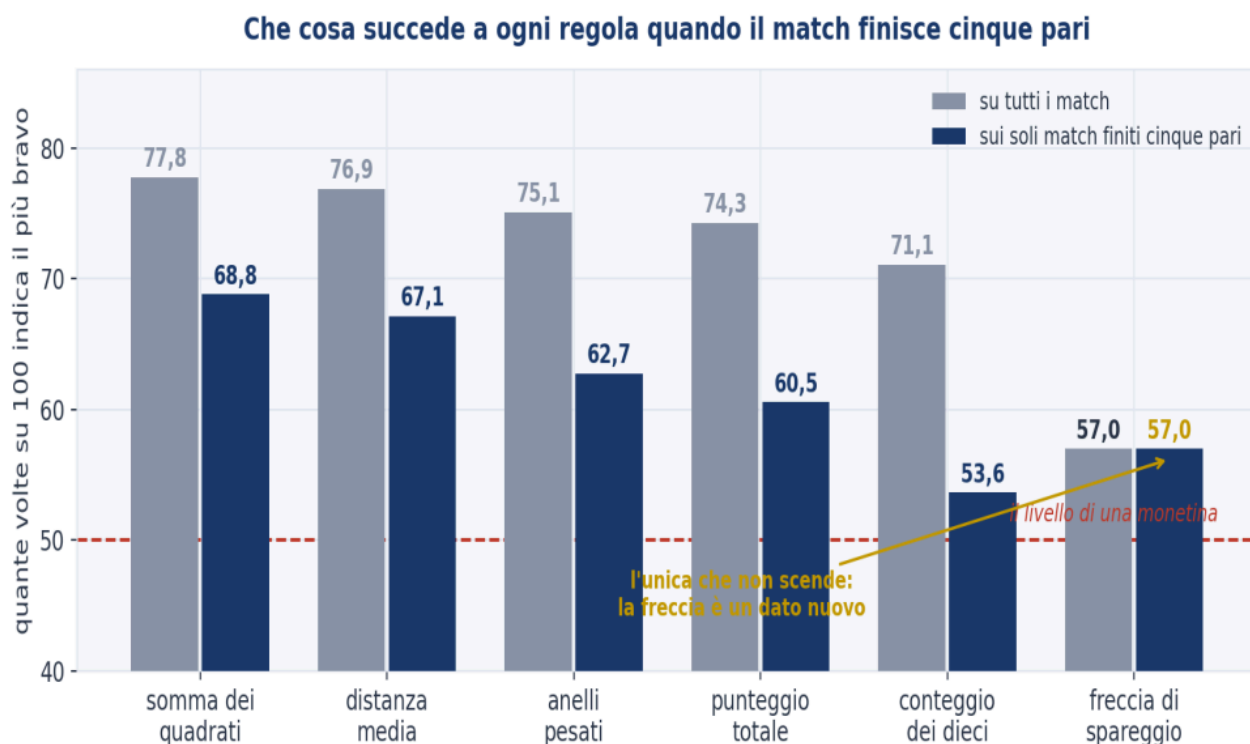


**Figura 2 — Che cosa succede a ogni regola quando il match finisce cinque pari.** Quante volte su cento ciascuna regola indica il più bravo, su tutti i match e sui soli match finiti in parità. Coppia di riferimento 690 contro 680. La freccia di spareggio è l'unica che non scende, e il conteggio dei dieci è quello che arriva più vicino alla monetina.

## 7. Perché la freccia non perde nulla, e perché questo non basta

Che la freccia conservi tutto non è un caso fortunato, ed è una di quelle cose che si dimostrano in due righe e poi non si dimenticano più. La freccia di spareggio si tira quando il match è già finito: non ha contribuito a produrre la parità, non ne fa parte, e con essa non ha nessun rapporto se non quello di venire dopo.

Chiedere a quella freccia di decidere significa dunque chiedere a un dato che non è ancora stato speso; chiedere alle quindici già tirate di rompere la parità che loro stesse hanno prodotto è una faccenda diversa, perché si sta domandando a un dato di dire una cosa che in quel momento sappiamo già non aver detto. In termini formali: guardare solo i match finiti pari non tocca ciò che con quella parità non ha rapporto, ed è la formula S4.

Sembrerebbe quindi un argomento a favore della regola in vigore, e non lo è: anzi, è l'equivoco su cui si regge buona parte della discussione. Perché la freccia conserva tutto, sì, ma di un tutto che è pochissimo. Sulla coppia di riferimento indovina 57 volte su cento e sbaglia le altre quarantatré, e non si muove di lì né prima, né dopo la parità. Le altre regole partono da 71, 74, 78, molto più in alto, e poi scendono. La domanda giusta, a questo punto, non è chi conservi di più ma chi arrivi più in alto dopo aver perso quello che deve perdere.

### Perché la freccia separa così poco i migliori

C'è un fatto sulla freccia che vale la pena raccontare, perché è insolitamente pulito e perché torna utile più avanti. La probabilità che la freccia indichi il più bravo si scrive in forma chiusa, ed è governata da una sola quantità: il rapporto fra le ampiezze delle due nuvole. Detto  $\rho$  quel rapporto, la probabilità vale  $\rho^2$  diviso  $1 + \rho^2$ , e nient'altro entra nel conto (formula S8) — non la distanza di tiro, non la larghezza degli anelli, non quanto sia teso il match. Sulla nostra coppia il rapporto è 1,15, e quella formula dà 57,0.

Questo spiega anche perché lo spareggio sia così poco discriminante fra i migliori, e conviene vederlo salendo di livello. Fra il finalista da 690 e uno da 685 il rapporto fra le nuvole scende a 1,076, e la freccia indovina 53,6 volte su cento: su venti spareggi ne sbaglia più di nove. In una finale olimpica, dove i due arcieri sono ancora più vicini, si scende sotto il 53, che vuol dire una regola distinguibile da un sorteggio soltanto guardando decine di finali.

La freccia non è imprecisa perché sia una sola: è imprecisa perché deve distinguere due nuvole che si somigliano moltissimo, e con un colpo solo non ce n'è modo. Un giudice che dovesse dire quale dei due gruppi sia più stretto guardando una freccia per parte si troverebbe nella stessa condizione, e nessuno gli chiederebbe di rispondere.

### **I due fatti su cui poggia tutta la Parte prima**

Torniamo alla tabella del capitolo precedente e leggiamone la colonna di mezzo, quella dei soli match finiti in parità, che è la sola popolazione su cui una regola di spareggio agisce.

Il primo fatto è che il punteggio totale del match indovina 60,5 volte su cento e la freccia 57. Il totale, cioè la somma delle 15 frecce, che è già scritta sul referto e che il regolamento ignora del tutto, manda avanti l'arciere giusto tre volte e mezzo in più ogni cento parità. Su una gara con dieci spareggi è un arciere ogni tre gare che passerebbe il turno per merito invece che per caso, e la somma è già lì, sul tavolo del giudice. E li batte pur lasciando irrisolti i propri pareggi, che come vedremo sono un terzo dei casi: se anche a quelli si desse una risposta, il divario si allargherebbe.

Il secondo è che il conteggio dei dieci, il criterio che chiunque proporrebbe per primo, indovina 53,6 volte su cento: fa peggio della freccia di tre volte e quattro, ed è a un soffio dalla monetina. Su cento match finiti cinque pari ne indovina cinquantatré e ne sbaglia quarantasette, mentre tirare una moneta ne indovinerebbe cinquanta, quindi tutto quello che quel criterio aggiunge, rispetto al lanciare un centesimo in aria, sono tre match e mezzo ogni cento parità. Vedremo al capitolo 14 che non si tratta di un criterio semplicemente debole.

Una precisazione sulla lettura di quei numeri, perché tornerà utile in tutto il volume. Le regole che possono pareggiare (il totale, il conteggio dei dieci) non decidono sempre, e in quei casi ho accreditato mezzo punto, cioè ho fatto come se la parità residua venisse risolta con una moneta. È una convenzione conservativa: una regola vera dovrebbe risolverli in qualche modo, e qualunque modo sensato è meglio di una moneta. Il totale a 60,5, quindi, è il suo valore da solo e non il valore della regola che si potrebbe costruirci sopra, che sarà più alto.

---

## **8. Quanto si può guadagnare, al massimo**

Prima di confrontare le regole conviene sapere quanto margine ci sia in tutto, altrimenti si scopre che una regola indovina 57 volte su cento e non si sa se sia molto o poco, perché non si sa se il massimo fos-

se 58 o 90.

Il conto si fa prendendo la regola migliore concepibile, quella che sapendo tutto sbaglia il meno possibile, e guardando quanto spostati l'esito rispetto a una monetina. Il risultato dipende da due cose che tirano in direzioni opposte: quanto spesso la parità capita, che diminuisce al crescere del divario fra i due arcieri, e quanto bene la regola sappia risolverla, che invece aumenta. Il prodotto delle due è la quantità che conta, e la chiamo guadagno (formula S10).

Conviene fermarsi un momento sull'unità, perché è il punto in cui è più facile perdersi. Il guadagno non è quante volte la regola indovini: è di quanto sposta la probabilità che a vincere il match sia il più bravo, tenendo conto di quanto spesso serve. Una regola perfetta che agisse su una situazione che non capita mai avrebbe guadagno zero, ed è la stessa ragione per cui un'ottima norma su un caso raro non cambia lo sport.

| Divario fra i due | Quanti match finiscono pari | Guadagno massimo |
|-------------------|-----------------------------|------------------|
| nullo             | 21,40 %                     | 0,00 punti       |
| 3 punti           | 20,68 %                     | 1,48 punti       |
| 6 punti           | 19,18 %                     | 2,66 punti       |
| 8 punti           | 17,87 %                     | 3,19 punti       |
| 11 punti          | 15,68 %                     | 3,66 punti       |
| 13 punti          | 14,15 %                     | 3,74 punti       |
| 17 punti          | 11,17 %                     | 3,58 punti       |
| 25 punti          | 6,31 %                      | 2,52 punti       |

*(Ancoraggio 700, parametrizzazione in punti. Guadagno in punti percentuali sulla probabilità di vincere il match, con la lettura ottimale sulle coordinate. La seconda cifra decimale è incerta di circa due centesimi.)*

La prima colonna si legge così: fra due arcieri identici circa un match su cinque finisce cinque pari, e fra due separati da venticinque punti circa uno su sedici.

La forma della colonna di destra è il primo risultato strutturale di queste pagine, e va guardata prima dei valori: parte da zero, sale, e poi torna a scendere, con il massimo attorno ai 14 punti di divario. La ragione è chiara appena la si vede. Se i due arcieri sono identici non c'è niente da distinguere, e nessuna regola può fare meglio di una moneta; se sono lontanissimi la parità quasi non si verifica, e una regola che agisce su un evento raro non sposta niente.

Detto in gare: fra due arcieri separati da venticinque punti il match finisce pari una volta su sedici, quindi in un tabellone da 64 di spareggi se ne tirano quattro invece di dieci, e su quattro occasioni anche la regola perfetta ha poco spazio per lavorare. La regola di spareggio conta soltanto in una finestra intermedia di divari, e anche lì conta poco.

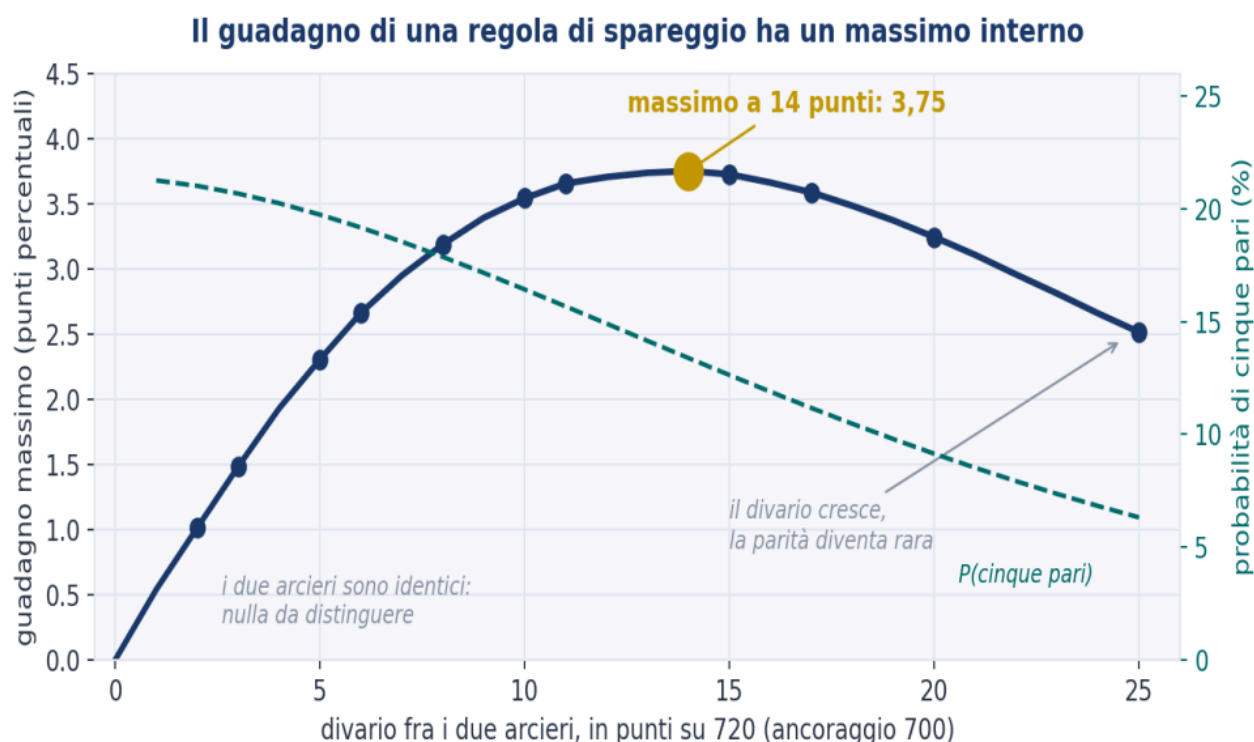


**Figura 3 — Il guadagno di una regola di spareggio ha un massimo interno.** Prodotto fra probabilità di cinque pari e capacità della regola ottimale sulle coordinate, in punti percentuali sulla probabilità di vincere il match. Ancoraggio 700, parametrizzazione in punti. La curva vale zero a divario nullo e tende a zero per divari grandi; il massimo è a 14 punti e vale 3,75. In tratteggio, la probabilità di finire cinque pari.

## 9. I quattro tetti

Quel massimo di 3,75 presuppone però una cosa che nessun regolamento prevede: poter misurare le coordinate esatte di tutte le frecce del match. Un referto di gara registra il valore dell'anello (un nove è un nove) non la posizione dentro l'anello. Quindi quel numero è un limite teorico e non un obiettivo, e il numero che conta per chi debba scrivere una regola è un altro. Ne ho calcolati quattro, uno per ciascun insieme di cose che si potrebbero guardare.

| Che cosa si lascia guardare alla regola        | Guadagno massimo | Ammesso oggi |
|------------------------------------------------|------------------|--------------|
| Coordinate del match più freccia di spareggio  | 4,00 punti       | no           |
| Coordinate del match                           | 3,75 punti       | no           |
| Anelli del match più freccia di spareggio      | 2,88 punti       | sì           |
| Anelli del match soltanto                      | 2,31 punti       | sì           |
| Regola in vigore: la sola freccia di spareggio | 1,47 punti       | in vigore    |

(Guadagno massimo sul divario, in punti percentuali di probabilità di vincere il match. Ancoraggio 700.)

La prima conclusione operativa è nella terza riga confrontata con l'ultima: la regola in vigore realizza circa metà del margine che il regolamento già consentirebbe di ottenere. Il rapporto è 1,48 su 2,88 al

punto in cui ciascuno raggiunge il proprio massimo; sulla coppia di riferimento, dove il divario è dieci punti e non tredici, la frazione realizzata scende al 46 per cento. Fra il 42 e il 53 a seconda della coppia, quindi, e la mezza misura è una media e non una costante.

La seconda conclusione si legge nella quarta riga, e nel dibattito non l'ho trovata formulata. Leggere il referto che è già sul tavolo del giudice, senza far tirare nulla a nessuno, rende il 56 per cento in più che far tirare una freccia in più a due atleti stanchi. La somma è già scritta, il giudice ce l'ha davanti, e la si sta ignorando per chiedere ai due arcieri una prestazione in più.

### **Che cosa vuol dire, su un campo di gara**

Qui il numero smette di essere astratto, e conviene ricostruire la scena per intero.

In una gara da 64 arcieri si giocano 63 match. Fra due avversari come la nostra coppia di riferimento, circa dieci finiscono pari e vanno decisi da una regola di spareggio. Una monetina ne indirizzerebbe correttamente la metà, cioè cinque: è il pavimento, e serve a leggere tutto quello che segue.

La freccia di spareggio, cioè la regola in vigore, ne indirizza correttamente cinque e tre quarti. Guadagna quindi tre quarti di arciera a gara rispetto al caso: su quattro gare, tre arcieri passano il turno per merito invece che per fortuna, e nella quarta gara la regola non ha spostato nessuno.

La migliore regola che il regolamento attuale consentirebbe ne indirizzerebbe correttamente sei e sei decimi, cioè un arciera e sei decimi in più della monetina. Rispetto a oggi il guadagno è di poco meno di un arciera a gara: su cinque gare, quattro arcieri in più passano il turno per merito. È tutto quello che c'è in palio sulla regola di parità, e vale la pena tenerlo in mente per il resto del libro.

Misurando le distanze dal centro invece degli anelli si arriverebbe a sette, cioè due arcieri in più della monetina. Nemmeno il massimo assoluto, quindi, sposta due match su 63: in una gara intera, con dieci spareggi tirati, la differenza fra il regolamento peggiore immaginabile e quello migliore possibile sono due arcieri.

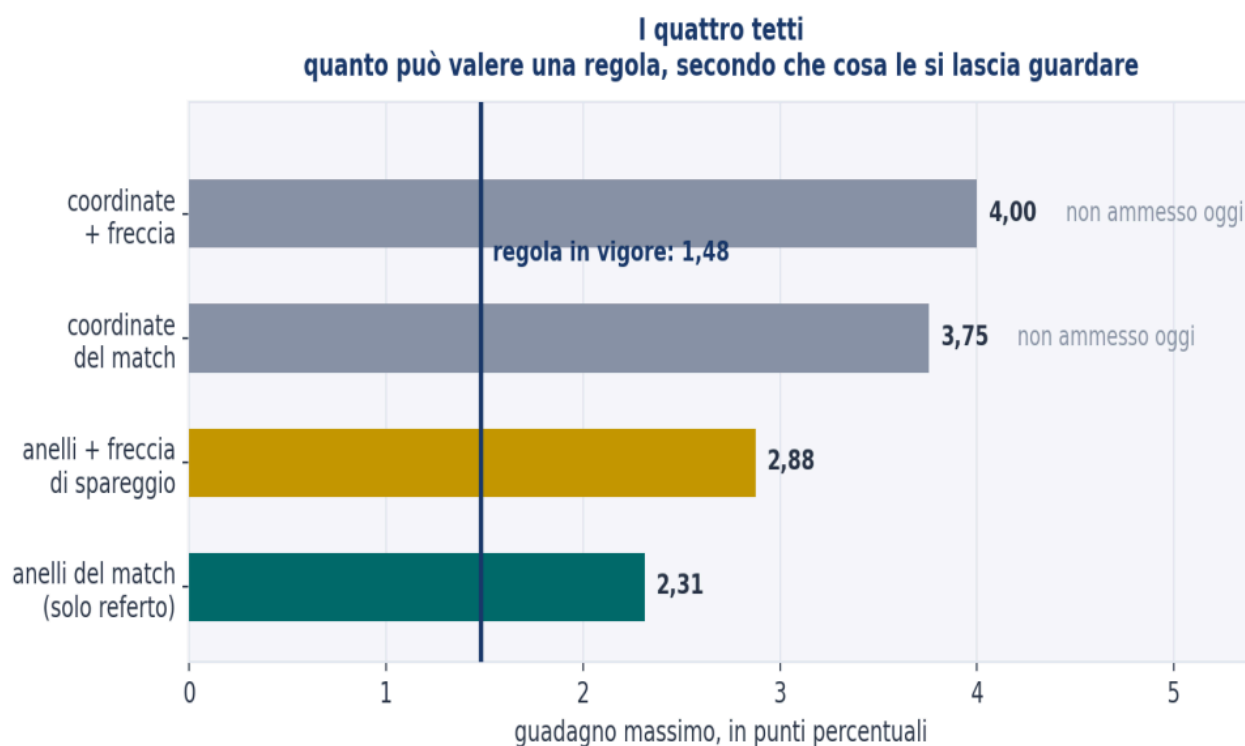


**Figura 4 — I quattro tetti.** Guadagno massimo in punti percentuali secondo che cosa si lascia guardare alla regola, con la riga del regolamento vigente evidenziata. Calcolo esatto sulla coppia che massimizza ciascun tetto, ancoraggio 700, tetto preso al massimo sul divario.

## 10. Quattordici modi di decidere, uno per uno

Fissato il tetto, la domanda diventa concreta: quali sono i modi possibili di decidere un match pari, e quanto vale ciascuno. Ne ho provati quattordici, e conviene descriverli prima di metterli in fila, perché la tabella da sola non spiega niente.

**Il sorteggio** non lo propone nessuno, ma serve da pavimento: indovina 50 volte su cento per definizione, e ogni altra regola va giudicata su quanto stia sopra questo.

**La freccia di spareggio** è la regola in vigore, e indovina 57 volte su cento. Il suo pregio è di usare un dato nuovo, non contaminato dalla parità; il suo difetto è che quel dato è uno solo, e un dato solo dice poco.

**Il conteggio dei dieci** guarda chi abbia fatto più dieci nel match, e in caso di parità chi abbia fatto più X. È il criterio che la tradizione considera più naturale (ha centrato più volte, quindi ha tirato meglio) e indovina circa 55 volte su cento, cioè meno della freccia. Il capitolo 14 mostrerà che il problema è molto più profondo di così.

**Il punteggio totale del match** somma le 15 frecce e assegna la vittoria a chi ha più punti. È la proposta che quasi tutti fanno per prima, ed è ragionevole: se il formato a set butta via l'informazione sul margine con cui i set sono stati vinti, il totale la recupera tutta. Indovina 60,5 volte su cento anche lasciando

irrisolti i propri pareggi, quindi batte già la regola in vigore; ma da solo non basta, e il capitolo 13 spiega perché.

**Il conteggio delle frecce sotto il nove** conta le frecce da otto in giù e assegna la vittoria a chi ne ha meno. L'idea è che una freccia buttata via dica molto sulla larghezza della nuvola, mentre la differenza fra un nove e un dieci dica poco. Indovina 59,3 volte su cento da solo, e 62,4 mettendo la freccia di spareggio in coda per i pareggi: è una delle proposte migliori, ed è la più facile da verificare dagli spalti, perché si tratta di contare le frecce brutte.

**Il conteggio delle frecce sotto il dieci** è la stessa idea con il taglio spostato di un anello, si contano le frecce che non sono dieci, e indovina 55,9 volte su cento con la freccia in coda. Sei volte e mezza peggio ogni cento parità rispetto alla versione precedente, per un solo anello di differenza, e la ragione di quel divario è il contenuto del capitolo 11.

**Il conteggio degli X** da solo indovina 56,3 volte su cento, sotto la regola in vigore: l'X è un criterio molto fine e molto raro, e separa poco.

**Le cascate** sono le regole che consultano più criteri in fila finché uno decide. Quella classica (prima il totale, poi il conteggio dei dieci, poi gli X, poi la freccia) indovina 59,7 volte su cento; quella rovesciata, che consulta i dieci prima del totale, 57,67. La differenza fra le due è il primo indizio del ribaltamento che vedremo fra poco.

Togliendo il conteggio dei dieci dalla cascata, cioè consultando il totale, poi gli X, poi la freccia, si sale a 62,1: meglio di entrambe le cascate, il che dice già che il conteggio dei dieci non stava aggiungendo informazione ma togliendola. E togliendo anche gli X, cioè con due soli stadi (totale e poi freccia) si sale ancora, a 62,9. La chiamerò **TOTALE+** per brevità: è la proposta di queste pagine, e il capitolo 17 la spiega per intero.

**CONCORDE** è una regola diversa dalle altre: se il totale e il conteggio dei dieci indicano lo stesso arciere vince quello, e se si contraddicono si tira. Indovina 58,9 volte su cento, e non è una regola per l'accuratezza, ma una polizza contro un'eventualità precisa, che il capitolo 21 descrive.

**Il conteggio pesato degli anelli**, invece di dare a un nove nove punti e a un otto otto, dà a ciascun anello il peso che quell'anello merita davvero per distinguere due arcieri. Indovina 62,7 volte su cento, e con la freccia in coda 64: è la più accurata di tutte quelle che leggono i soli anelli del referto. Resta sotto il tetto del capitolo 9, che vale 65,16, e la differenza non è un difetto di questa regola: il tetto guarda anche la distanza della freccia di spareggio, che il regolamento già misura, e la combina con gli anelli invece di usarla come ultima parola. Nessun giudice fa quella combinazione al tavolo, ed è per questo che resta un tetto e non una proposta. Il conteggio pesato ha però un difetto che nessun guadagno compensa, e ne parlo al capitolo 12.

**La somma dei quadrati delle distanze**, infine, misura quanto disti dal centro ogni freccia, eleva al quadrato e somma, assegnando la vittoria al totale più basso. Indovina 68,8 volte su cento, molto sopra tutte le altre. Richiede però di misurare le coordinate, che oggi non si fa, ed è l'unica strada che porterebbe davvero fuori da questo margine stretto, per cui merita un capitolo suo.

Eccole tutte insieme, nella stessa unità.

| Modalità di spareggio                         | Indovina su 100<br>parità | Arcieri spostati per<br>gara | Spareggi tirati per<br>gara |
|-----------------------------------------------|---------------------------|------------------------------|-----------------------------|
| Sorteggio                                     | 50,0                      | +0,00                        | 0,0                         |
| Conteggio dei dieci, poi X                    | 54,96                     | +0,52                        | 0,7                         |
| Frecce sotto il dieci, poi freccia            | 55,9                      | +0,63                        | 3,5                         |
| Solo il conteggio degli X, poi freccia        | 56,15                     | +0,65                        | 2,1                         |
| Freccia di spareggio (in vigore)              | 57,0                      | +0,74                        | 10,6                        |
| Cascata rovesciata: dieci, totale, X, freccia | 57,67                     | +0,81                        | 0,4                         |
| CONCORDE                                      | 58,93                     | +0,95                        | 0,8                         |
| Cascata classica: totale, dieci, X, freccia   | 59,69                     | +1,03                        | 0,4                         |
| Totale, poi X, poi freccia                    | 62,09                     | +1,28                        | 0,7                         |
| Frecce sotto il nove, poi freccia             | 62,44                     | +1,31                        | 4,7                         |
| Conteggio pesato degli anelli                 | 62,7                      | +1,34                        | 0,0                         |
| Conteggio pesato con dispersione stimata      | 62,7                      | +1,35                        | 0,0                         |
| <b>TOTALE+ (totale, poi freccia)</b>          | <b>63,0</b>               | <b>+1,37</b>                 | <b>3,64</b>                 |
| Conteggio pesato, poi freccia                 | 64,00                     | +1,48                        | 1,9                         |
| Il meglio ottenibile entro il regolamento     | 65,16                     | +1,60                        | 10,6                        |
| Somma dei quadrati delle distanze             | 68,8                      | +1,99                        | 0,0                         |

*(Coppia di riferimento 690 contro 680, tabellone da 64 arcieri con 63 match. La prima colonna ha come popolazione i match finiti cinque pari; la seconda dice quanti arcieri in più, per gara, passano il turno per merito invece che per caso; la terza quanti spareggi si tirano in una gara intera. Pareggi accreditati a metà dove la regola può pareggiare.)*

La classifica completa va letta sapendo esattamente che cosa rappresenta. L'ottimalità del conteggio pesato a parametri noti è un risultato interno al modello: i pesi sono costruiti a partire dai logaritmi dei rapporti di verosimiglianza delle probabilità d'anello sotto i due profili considerati. Una regola che incorpora direttamente il modello può quindi sfruttarne quasi tutta l'informazione disponibile, ma eredita inevitabilmente anche le conseguenze di ogni sua eventuale misspecificazione.

È il motivo per cui pubblico la graduatoria completa e, insieme, non costruisco la raccomandazione principale sulla sua parte più sofisticata. Quanto più una regola conosce il modello, tanto più può guadagnare quando il modello è corretto e tanto più può perdere quando il modello sbaglia. Il totale del match, al contrario, non richiede di conoscere la forma della distribuzione, non dipende da pesi ricavati da parametri modellistici e legge una statistica direttamente osservabile: per questo parte con meno struttura da difendere fuori dal modello.

Se dovessi formulare oggi una previsione su ciò che sopravvivrà all'incontro con dati empirici indipendenti, direi: la testa della classifica più della sua graduatoria fine. Mi aspetto che le posizioni ravvicinate si rimescolino, che alcune regole pesate paghino parte del vantaggio ottenuto dentro il modello e che TOTALE+ resti competitivo fino a poter emergere non soltanto come migliore regola a costo informati-

vo minimo, ma come migliore regola complessiva nel dominio empirico rilevante. È una previsione, non un risultato di questo volume, e la scrivo precisamente perché possa essere verificata o smentita.

Nove delle quattordici battono la regola in vigore, che quindi non sta a metà classifica ma nel terzo inferiore. Fra la peggiore e la migliore delle proposte adottabili passano circa nove decimi di arciere a gara: guardando dieci gare, nove arcieri diversi arrivano al turno successivo a seconda di quale riga si scelga.

E c'è un fatto che la tabella mostra senza spiegarlo: aggiungere stadi non aiuta. La cascata classica, con due criteri in più di TOTALE+, indovina quasi tre volte in meno ogni cento parità, e per giunta riduce gli spareggi tirati da tre e sei a quattro decimi per gara, cioè da un pomeriggio con quattro spareggi a uno con nessuno, quasi sempre. Si paga su entrambi gli assi per aggiungere un criterio, e il perché sta nel capitolo che segue.

---

## 11. I pesi che gli anelli meriterebbero

La chiave di tutta la tabella precedente è una differenza che ho lasciato in sospeso: contare le frecce sotto il nove indovina 62,4 volte su cento, contare quelle sotto il dieci 55,9. Due regole identiche tranne che per dove si mette il taglio, e sei volte e mezza di differenza ogni cento parità per un anello di spostamento.

Per capirla bisogna guardare quanto valga davvero, ai fini del distinguere due arcieri, sapere che una freccia è finita in un certo anello invece che nel successivo. Quel valore si calcola, ed è quello che chiamo il peso dell'anello.

| Anello           | Peso che meriterebbe | Punti che dà il regolamento |
|------------------|----------------------|-----------------------------|
| anello 0 — fuori | 0                    | 0                           |
| anello 1         | 1,9                  | 1                           |
| anello 2         | 3,6                  | 2                           |
| anello 3         | 5,1                  | 3                           |
| anello 4         | 6,3                  | 4                           |
| anello 5         | 7,4                  | 5                           |
| anello 6         | 8,3                  | 6                           |
| anello 7         | 9,0                  | 7                           |
| anello 8         | 9,5                  | 8                           |
| anello 9         | 9,9                  | 9                           |
| anello 10        | 10                   | 10                          |

*(Pesi ottimali per la coppia di riferimento, riscaldati sull'intervallo da 0 a 10 per confronto diretto con i punti del regolamento. La forma concava non dipende dalla coppia scelta ed è dimostrata in appendice, formula S7.)*

I pesi non sono lineari: sono concavi. Il salto da zero a uno vale 1,9 punti, quello da nove a dieci ne vale 0,15, dodici volte e mezza meno. Detto senza scale: sapere che un arciere ha buttato via una freccia e l'altro no dice sulla loro differenza dodici volte più di quanto dica sapere che uno ha fatto dieci dove l'altro ha fatto nove.

E ha senso appena lo si dice ad alta voce. Distinguere una freccia buttata via da una mediocre dice moltissimo sulla larghezza della nuvola, perché soltanto una nuvola larga produce frecce così; distinguere un nove da un dieci non dice quasi niente, perché due arcieri di livello vicino fanno entrambi molti dieci e molti nove, e il confine fra quei due anelli cade proprio dove le loro nuvole si somigliano di più.

Da qui si legge tutta la tabella del capitolo precedente in una riga. Tre regole, tre scelte diverse su quello stesso anello: il totale semplice pesa il nove-contro-dieci come un punto pieno, cioè troppo; il conteggio delle frecce basse lo pesa zero, perché butta via ogni distinzione sopra il taglio, cioè troppo poco; i pesi ottimali lo pesano 0,15. E l'ordine di merito delle regole è esattamente l'ordine di quanto ciascuna si avvicini a quel 0,15.

Confrontando le regole allo stesso stadio, ciascuna da sola, senza freccia in coda, il conteggio pesato indovina 62,7 volte su cento, il totale 60,5, il conteggio delle frecce basse 59,2 e quello dei dieci 53,6. L'ordine è quello dell'approssimazione ai pesi giusti, e il conteggio dei dieci, che pesa tutto sull'anello meno informativo, sta in fondo.

*Una regola vale quanto bene approssima i pesi giusti. E i pesi giusti dicono che una freccia buttata via conta molto più di un dieci mancato.*

Questo spiega anche perché aggiungere stadi peggiori. Ogni stadio in più è un criterio che decide su un sottoinsieme sempre più piccolo, e su quel sottoinsieme il criterio deve battere l'alternativa che verrebbe dopo. Il conteggio dei dieci, inserito dopo il totale, decide sui match in cui il totale ha pareggiato, e lì, come vedremo, non solo non batte la freccia: fa peggio del sorteggio.

---

## 12. Misurare le distanze: la strada che nessuno ha provato

Se i pesi ottimali funzionano così bene, la domanda naturale è perché non adottarli. E dietro quella ne sta una più radicale: perché non misurare direttamente le distanze, invece di leggere gli anelli. Sono due proposte diverse e conviene separarle.

### La prima: adottare il conteggio pesato

Indovina 64 volte su cento con la freccia in coda, la più accurata di tutte quelle che leggono i soli anelli del referto, e non richiede alcuna strumentazione nuova: si calcola sui numeri che il giudice ha già davanti.

Ha però un difetto che nessun guadagno compensa. I pesi dipendono da chi si ha davanti (due arcieri d'élite e due arcieri di club non hanno gli stessi pesi) e questo significa che il giudice dovrebbe calcolarli, cosa che nessun giudice fa al tavolo. Ma soprattutto significa che un arciere che perde non può ricostruire perché ha perso: dovrebbe farsi spiegare che il suo nove valeva 9,9 e l'otto dell'avversario 9,5, e credere sulla parola a chi glielo dice. In un regolamento sportivo questa non è una comodità, è la condizione perché la regola sia accettata, perché una decisione che l'atleta non può verificare è una decisione che dovrà accettare per fiducia, e la fiducia in una gara internazionale è merce scarsa.

Vale la pena dire che ho provato a salvarla. Esiste una versione che stima il livello degli arcieri dai referti stessi, senza bisogno di conoscerlo in anticipo, ma va detto come, perché la scelta decide tutto: il livello comune si ricava dal totale delle trenta frecce dei due referti presi insieme, e i pesi si calcolano su una coppia separata di uno scarto nominale dichiarato (formula S6, seconda parte).

La precisazione non è pedanteria. Stimando invece l'ampiezza di ciascun arciere dal suo solo referto la versione crolla: quando i due totali coincidono, e in un match finito cinque pari accade una volta su tre, i pesi si annullano e la regola smette di decidere, sbagliando poi una volta su dieci. Con la stima sul livello comune le due versioni hanno deciso diversamente 52 volte su 50 468 034 match in parità: una volta su un milione, cioè una volta ogni ventimila stagioni internazionali. Il problema tecnico si dissolve, quindi; quello della verificabilità resta intero.

### **La seconda: misurare le coordinate**

È più radicale, ed è l'unica proposta di queste pagine che sposti il tetto invece di limitarsi ad avvicinarvisi. Registrare per ogni freccia non solo l'anello ma la distanza esatta dal centro alza il tetto: in una gara da 63 match si passa da un arciere e otto decimi spostati rispetto al caso a due e mezzo. Il quaranta per cento in più, ed è il solo modo di superare il muro contro cui tutte le altre proposte battono.

Servirebbero bersagli con lettura elettronica del punto d'impatto, oppure un sistema di scansione delle facce a fine volée. La tecnologia esiste, sistemi di rilevamento automatico capaci di trasmettere la posizione d'impatto sono stati sperimentati in gara, anche se non dispongo di una rassegna sistematica del loro uso; la domanda vera è quanta precisione serva, e la risposta si calcola in due righe.

La posizione misurata è quella vera più un errore dello strumento, e la somma di due nuvole tonde è una nuvola tonda: misurare con un errore di ampiezza  $\varepsilon$  equivale a leggere due arcieri le cui nuvole valgono la radice di  $\sigma^2$  più  $\varepsilon^2$ . L'effetto non è che le distanze si sporcano, è che il rapporto fra le due nuvole si comprime verso uno, e con esso il segnale, cioè lo strumento fa sembrare i due arcieri più simili di quanto siano.

Sulla coppia di riferimento un errore di 3 millimetri porta il rapporto da 1,15157 a 1,15094 e quello che la regola sa distinguere da 27,80 a 27,71 su cento: tre decimi di per cento, cioè si perde un trecentesimo di quanto la regola vale. Sui soli match finiti cinque pari, che è la popolazione su cui la regola agisce, la perdita è circa tre volte maggiore, da 18,80 a 18,61, cioè l'uno per cento, perché il filtro della parità ha già tolto quasi tutta la differenza fra i due, e quel poco che resta è più fragile.

La legge serve a chi deve scrivere un capitolato: la perdita cresce col quadrato del rapporto fra errore dello strumento e ampiezza della nuvola, con un coefficiente di 0,71 costante su tutto l'intervallo utile. Il due per cento di perdita si raggiunge a 7,5 millimetri guardando tutti i match, e attorno ai 4 guardando i soli cinque pari; a un centimetro di errore si perde il tre e mezzo per cento, e a un centimetro nessuno costruirebbe uno strumento.

Detto in modo concreto: uno strumento che sbagli di mezzo millimetro in più o in meno di quanto sbaglia un arbitro che legge un anello a occhio è più che sufficiente. Non serve una precisione da laboratorio, serve una precisione da campo, e questo è il numero che dice quale.

La regola diventerebbe, letteralmente: vince chi ha la somma dei quadrati delle distanze più piccola. Il quadrato non è un capriccio, si ricava dal rapporto di verosimiglianza fra le due ipotesi, che è la formula  $S_3$ , ed è la lettura che estrae il massimo dell'informazione disponibile. Ha inoltre una proprietà che nessuna delle altre possiede: per applicarla non serve sapere niente. Qualunque sia la coppia di arcieri, la regola è sempre la stessa; il giudice non deve stimare nulla e il calcolo è una somma.

Resta la verificabilità, e non voglio nasconderla: un arciere non ricalcola a occhio una somma di quadrati. Ma può vedere i numeri, ed è una differenza sostanziale. Se il tabellone mostra 412,7 contro 428,3, la

decisione è ispezionabile in un modo in cui un peso calcolato non lo è: sono due somme di quindici numeri, ciascuno letto da uno strumento che l'atleta può contestare esattamente come oggi contesta la lettura di un anello.

Il costo non lo so stimare e non fingo di saperlo. Ma la domanda che una federazione dovrebbe porsi non è quanto costi, bensì quanto costi rispetto a quello che compra, e adesso almeno il secondo termine è misurato.

---

### **13. La cascata, e perché sembrava la risposta**

Con i pesi del capitolo 11 in mano, e le distanze del capitolo 12 messe da parte come strada da percorrere un giorno, la domanda torna a essere pratica: qual è la regola migliore fra quelle che una federazione potrebbe adottare domani mattina, senza comprare niente?

Per settimane ho creduto di averla, ed era la cascata classica: prima il totale, se pareggia il conteggio dei dieci, se pareggia ancora gli X, e solo se pareggiano tutti e tre si tira.

Il ragionamento sembra solido. Il totale indovina 60,5 volte su cento contro le 57 della freccia, ma da solo non basta perché pareggia troppo spesso: su cento match finiti cinque pari, il totale delle 15 frecce coincide in 34 casi. Detto in una gara, significa che su dieci spareggi il totale sarebbe muto in tre o quattro, e in quei casi il giudice resterebbe senza risposta. Una regola che in un terzo dei casi non decide non è una regola, e la soluzione ovvia è quella che qualunque regolamento sportivo usa in altri contesti: si mettono i criteri in fila, dal più informativo al meno, e si scende finché uno decide.

Sulla coppia di riferimento la cascata indovina 59,7 volte su cento contro le 57 della regola in vigore. Decide sempre, non richiede strumentazione, non allunga la gara, si calcola a mano su un referto in 20 secondi. Sembrava la risposta, e per settimane l'ho trattata come tale.

Poi, per una ragione che non aveva niente a che vedere con la sostanza, si è rivelata sbagliata.

---

### **14. Il ribaltamento: quando contare i dieci indica l'arciere sbagliato**

La scoperta è arrivata cambiando l'unità di misura, non cercando qualcosa.

Avevo passato settimane a esprimere ogni risultato in una valuta da analista: di quanto si sposta la probabilità che a vincere il match sia il più bravo. Poi, per rendere queste pagine leggibili a chi non fa questo mestiere, ho rifatto la stessa tabella nell'altra valuta (su cento match finiti cinque pari, quante volte passa il più bravo) e nella valuta nuova la tabella conteneva una riga impossibile: la cascata risultava sotto il punteggio totale.

Non poteva essere. Se il conteggio dei dieci fosse anche solo un po' informativo, sostituirlo alla monetina su quei match potrebbe soltanto migliorare le cose. Una riga impossibile non è una riga strana: è una riga che dice che qualcosa non torna. L'ho misurato, e questo è il risultato che ha riscritto tutta la Parte prima.

Sui match che finiscono cinque pari e in cui anche i totali sono pari, e sotto il modello di dispersione che questo volume adotta, il conteggio dei dieci indica l'arciere sbagliato. Non indica meno bene: indica quello sbagliato, più spesso di quanto indichi quello giusto. Sulla coppia di riferimento manda avanti il più bravo 44,7 volte su cento, il che vuol dire che ne manda avanti il meno bravo cinquantacinque: una moneta lanciata in aria farebbe meglio, e applicare la regola al contrario farebbe meglio ancora.

E non è un effetto locale di una coppia particolare, come mostra il calcolo esatto, non simulato, su tutta la griglia.

| Coppia  | Quanto spesso i totali pareggiano | Conteggio dei dieci | Dieci, poi X | Freccia di spareggio |
|---------|-----------------------------------|---------------------|--------------|----------------------|
| 700-695 | 46,52 %                           | 48,70               | 52,19        | 54,30                |
| 690-685 | 37,2 %                            | 47,47               | 48,81        | 53,65                |
| 690-680 | 34,4 %                            | 44,66               | 46,79        | 57,01                |
| 690-670 | 28,1 %                            | 38,97               | 41,79        | 62,94                |
| 680-660 | 24,4 %                            | 38,82               | 40,10        | 61,47                |
| 660-640 | 19,45 %                           | 40,58               | 41,02        | 59,35                |
| 650-620 | 15,2 %                            | 38,09               | 38,46        | 62,21                |
| 640-600 | 12,16 %                           | 36,57               | 36,88        | 64,39                |

*(Su cento match finiti cinque pari e con i totali in parità. Calcolo esatto, nessun errore di simulazione. Pareggi del criterio accreditati a metà.)*

Tutta la colonna sta sotto 50, dal 48,70 del caso più tenue al 36,57 del più marcato. Nella riga peggiore, contare i dieci manda a casa l'arciere più bravo in sessantatré casi su cento: chi lo applicasse sistematicamente farebbe meglio a decidere con la moneta, e meglio ancora a premiare l'arciere con meno dieci. Aggiungere gli X non ripara nulla, perché la colonna accanto resta sotto 50 quasi dappertutto.

**A totale pari, chi ha più dieci è chi ne ha avuto più bisogno · 690 contro 680, calcolo esatto**

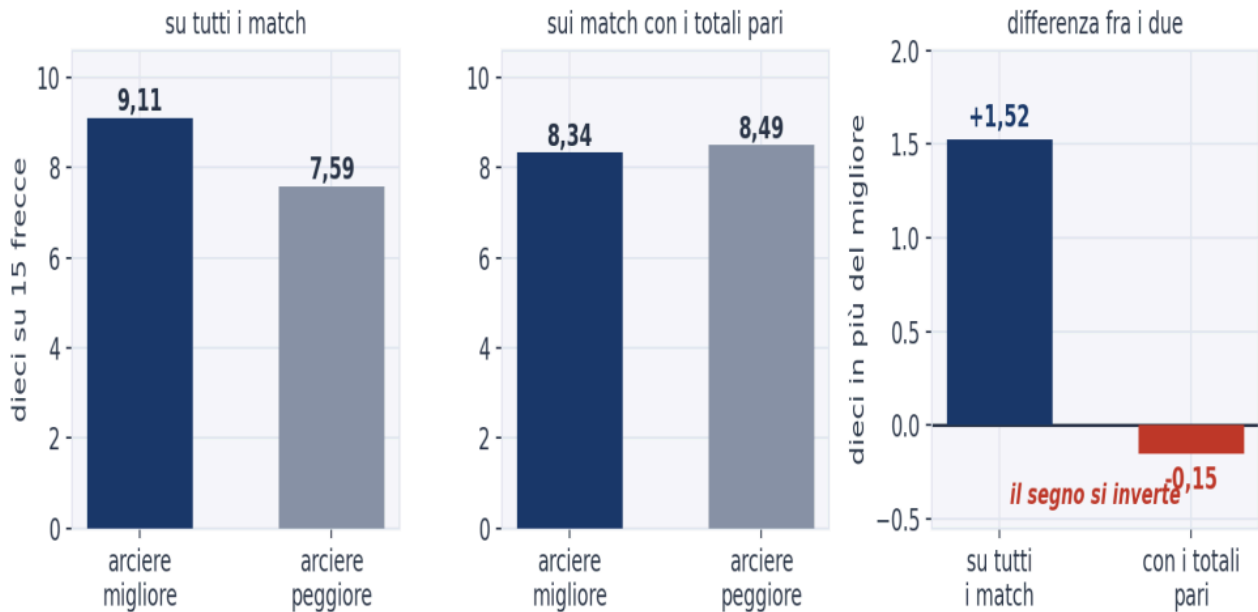


**Figura 5 — Il rovesciamento del conteggio dei dieci.** Numero medio di dieci su 15 frecce per i due arcieri, senza condizionamento e condizionando ai totali pari, con la differenza fra i due nel terzo pannello. Coppia 690 contro 680, calcolo esatto. La differenza passa da +1,52 a -0,15: il segno si inverte.

## 15. La spiegazione, che è aritmetica e non statistica

La ragione è più semplice di quanto il risultato faccia temere, e una volta vista è ovvia. Bastano due tabelle, e la sola differenza fra loro è quali match si guardano.

Guardando tutti i match che i due arcieri giocano, il quadro è quello che ci si aspetta.

|                    | Arciere migliore (690) | Arciere peggiore (680) |
|--------------------|------------------------|------------------------|
| Dieci su 15 frecce | 9,110                  | 7,588                  |
| Frecce da 8 o meno | 0,357                  | 0,894                  |

(Valori attesi, calcolo esatto sotto il modello standard.)

Il più bravo fa più dieci (nove contro sette e mezzo su quindici frecce) e butta via meno di metà delle frecce che butta via l'altro. Tutto normale, ed è quello che chiunque si aspetterebbe stando dietro la linea di tiro.

Ora guardiamo soltanto i match in cui i totali sono pari, che è la sola situazione in cui il criterio verrebbe letto; nella qualificazione è quella che resta dopo aver pareggiato anche gli X.

|                    | Arciere migliore (690) | Arciere peggiore (680) |
|--------------------|------------------------|------------------------|
| Dieci su 15 frecce | 8,343                  | 8,495                  |
| Frecce da 8 o meno | 0,443                  | 0,587                  |

*(Stessi arcieri. La sola differenza rispetto alla tabella precedente è quali match si guardano.)*

Il peggiore ne fa di più. Non è un artefatto e non è rumore, e il meccanismo si dice in una frase. Se due arcieri arrivano allo stesso totale e uno dei due ci arriva con più dieci, quello deve aver restituito quei punti da qualche altra parte, perché il totale è fisso e non si crea dal nulla: i punti mancanti si concentrano sulle frecce rimaste, che sono meno.

Un esempio rende la cosa evidente. Due arcieri chiudono il match con lo stesso totale. Il primo ci arriva con nove dieci e sei nove; il secondo ne ha fatti undici, ma per pareggiare quel totale ha dovuto piazzare da qualche parte due frecce molto basse. Chi guarda il conteggio dei dieci premia il secondo. Chi guarda il bersaglio vede che il secondo ha buttato via due frecce e il primo nessuna.

Dove finiscano di preciso i punti mancanti (su qualche freccia bassa o su molte frecce mediocri) l'aritmetica non lo impone, e referti che vanno in controtendenza si costruiscono a tavolino. Nei referti generati dal modello, la strada è quasi sempre la stessa: la correlazione fra le due differenze, misurata su quei match, vale 0,9898. Detto senza gergo, il legame è quasi ferreo: praticamente ogni volta che un arciere ha più dieci dell'altro a parità di totale, ha anche più frecce buttate via.

Ne segue che se due arcieri hanno lo stesso totale e uno ha più dieci, quell'uno è quello che ha avuto più bisogno di dieci per arrivarci: cioè quello che ne ha buttate via di più. Il che porta all'enunciato corretto, che è molto più forte di "il conteggio dei dieci è un criterio debole".

A totale fissato, il conteggio dei dieci non misura chi ha tirato meglio: misura chi ha concentrato il proprio deficit su meno frecce (formule S11 e S12). Sotto il modello standard concentrare il deficit è precisamente ciò che fa l'arciere meno bravo, e da lì discende il resto. Il conteggio dei dieci, in quella situazione, misura correttamente la cosa sbagliata.

### **La verifica che dice di quale parità si tratti**

Il meccanismo predice una cosa che si può controllare. Se il ribaltamento nasce dal vincolo sul totale, non dovrebbe dipendere dal cinque pari: dovrebbe bastare che i totali siano uguali, indipendentemente da come siano andati i set.

Ho rifatto il calcolo chiedendo soltanto che i totali fossero pari, senza chiedere che il match fosse finito cinque a cinque, e il risultato regge su tutta la griglia: 48,4 sulla coppia più tenue, 44,2 su quella di riferimento, 38,3 su quella più larga. Gli scarti rispetto alla tabella precedente stanno fra due e sette decimi, cioè il cinque pari sposta pochissimo.

È il pareggio dei totali a contare, non il cinque pari. Il che ha una conseguenza pratica per un regolamento: il ribaltamento colpirebbe il conteggio dei dieci ovunque lo si usi per rompere una parità di totale, e nel regolamento in vigore quel posto è uno solo, le parità ordinarie della classifica di qualificazione, dove si guardano prima gli X e poi i dieci. Per le parità che decidono l'accesso alle eliminatorie e per i match il regolamento prescrive invece lo spareggio, e lì il criterio non entra. Non è un difetto del formato: è un difetto del criterio.

E c'è un ultimo strato, che rende il fatto più che una curiosità tecnica. Il conteggio dei dieci non viene mai consultato in una situazione qualunque: lo si guarda dopo aver constatato una parità, cioè esattamente quando il totale ha già taciuto. Non si applica dunque un buon criterio in una situazione sfortunata: si applica un criterio nella sola situazione in cui quel criterio si inverte. Non è che la popolazione sbagliata sia capitata, è la regola stessa a crearla.

## 16. Dove il ribaltamento vale, e dove smette di valere

Un risultato del genere va delimitato, altrimenti si estende dove non regge. Ho fatto tre controlli, e tutti e tre insegnano qualcosa.

Il primo riguarda l'anello X: si comporta allo stesso modo? No, e la ragione conferma il meccanismo invece di contraddirlo. L'X è una suddivisione dentro l'anello del dieci, e un X e un dieci valgono entrambi dieci punti; passare dall'uno all'altro non richiede quindi nessuna restituzione al totale, e il vincolo aritmetico non lo tocca. Infatti il conteggio degli X, sui match con i totali pari, resta informativo sulle coppie di vertice: indovina 53,9 volte su cento a 700 contro 695, 53,1 sulla coppia di riferimento, 54,3 a 690 contro 670.

Il secondo controllo nasce da lì: se l'X è informativo, conviene usarlo come secondo criterio dopo il totale? La risposta è no, e non me l'aspettavo. Su quegli stessi match la freccia di spareggio vale più dell'X (57 contro 53) quindi inserire l'X fra il totale e la freccia peggiora invece di migliorare.

| Coppia  | Totale, poi freccia | Totale, poi X, poi freccia | Totale, poi 10, poi X, poi freccia |
|---------|---------------------|----------------------------|------------------------------------|
| 700-695 | 57,33               | 57,49                      | 56,65                              |
| 690-680 | 62,95               | 62,09                      | 59,70                              |
| 690-670 | 73,68               | 72,02                      | 68,06                              |
| 660-640 | 68,80               | 67,27                      | 65,37                              |

*(Su cento match finiti cinque pari. Ancoraggio al primo dei due punteggi. Calcolo esatto.)*

E c'è un secondo costo, meno visibile. Con l'X in mezzo il ricorso allo spareggio crolla: da una forbice fra il 19 e il 47 per cento delle parità a una fra il 5 e l'8. In una gara da 64 arcieri vuol dire passare da due o cinque spareggi tirati a meno di uno, cioè quasi sempre nessuno. Si paga accuratezza e spettacolo insieme per aggiungere uno stadio.

La conclusione generale è che quando il totale tace, una freccia nuova batte i criteri residuali che ho provato, il conteggio degli X e quello dei dieci. Che batte qualunque funzione ricavabile dal referto non lo dimostro.

Il terzo controllo chiede dove smetta di valere anche per l'X. La risposta è che l'X resta informativo finché il suo anello è grande rispetto alla nuvola: scendendo di livello la nuvola cresce, l'anello X diventa relativamente piccolo, e a un certo punto anche l'X si inverte. L'attraversamento avviene quando il raggio dell'X sta all'ampiezza della nuvola come 0,52 a uno (formula S14), che corrisponde a un livello di circa 670 punti su 720. Sopra quel livello l'X è utile, sotto si inverte come il conteggio dei dieci: la stessa regola, applicata a una gara nazionale invece che a una finale mondiale, cambia segno.

La soglia però si sposta un poco col divario fra i due arcieri, ed è onesto riportarla per intero invece di scegliere il valore che fa comodo: con 5 punti di differenza l'attraversamento è a 667 punti e a un rapporto di 0,508; con 10 punti a 669 e 0,522; con 20 punti a 673 e 0,547. Più i due arcieri sono lontani, più in alto va la soglia, e il motivo è lo stesso che governa tutto questo capitolo: l'informazione utile sta dove le due nuvole si somigliano meno.

È interessante che lo stesso rapporto governi anche un altro fenomeno di questo lavoro, cioè quanto costi leggere anelli invece di coordinate, e che le due soglie stiano nell'ordine giusto: prima la lettura per anelli comincia a costare informazione, poi comincia a invertire il segno. Se fossero venute al contrario avremmo avuto un problema.

---

## 17. La proposta, e perché funziona

Se il conteggio dei dieci va escluso, la cascata si semplifica da sé e restano due stadi soli. La regola sta in una frase.

*In caso di parità a set vince chi ha totalizzato più punti nel match. Se anche i totali sono pari, si tira una freccia per uno e vince la più vicina al centro.*

La chiamo TOTALE+ per brevità, e prima di difenderla conviene vedere quanto vale.

| Coppia  | Freccia (oggi) | Cascata | TOTALE+ | Il meglio ottenibile | Frazione realizzata |
|---------|----------------|---------|---------|----------------------|---------------------|
| 700-695 | 54,3           | 56,6    | 57,3    | 58,1                 | 90 %                |
| 690-685 | 53,6           | 55,1    | 56,7    | 57,9                 | 85 %                |
| 690-680 | 57,0           | 59,70   | 62,95   | 65,2                 | 85 %                |
| 690-670 | 62,95          | 68,06   | 73,7    | 77,4                 | 87 %                |
| 680-660 | 61,5           | 66,85   | 71,9    | 75,7                 | 86 %                |
| 660-640 | 59,4           | 65,37   | 68,80   | 72,38                | 84 %                |

*(Su cento match finiti cinque pari. Ancoraggio al primo dei due punteggi. La frazione realizzata è rispetto al meglio ottenibile entro il regolamento, contando dal pavimento del 50 per cento.)*

TOTALE+ si prende fra l'84 e il 90 per cento di tutto quello che il regolamento consente di prendere, contro il 42-53 per cento della regola in vigore. Ma il numero da solo non convince nessuno, e la parte che serve a difendere la proposta è un'altra: perché funziona.

Un match che finisce cinque pari non è un match in cui i due hanno tirato ugualmente. È un match in cui hanno vinto e perso i set nell'ordine giusto per pareggiare, cosa che può accadere anche con quindici punti di differenza sul totale. Quel totale, però, è ancora lì sul referto, e dice qualcosa. La domanda è quanto.

| Margine di totale | Quanto spesso capita | Quante volte il totale ha ragione | Quante volte ha ragione la freccia |
|-------------------|----------------------|-----------------------------------|------------------------------------|
| 0 punti — pari    | 34,4 %               | 50,48 — non decide                | 57,0                               |
| 1 punto           | 40,8 %               | 60,8                              | 57,0                               |
| 2 punti           | 17,4 %               | 71,3                              | 57,0                               |
| 3 punti           | 5,6 %                | 80,7                              | 57,0                               |
| 4 punti           | 1,5 %                | 87,8                              | 57,0                               |
| 5 punti o più     | 0,4 %                | 93,4                              | 57,0                               |

*(Coppia di riferimento 690 contro 680, popolazione dei match finiti cinque pari, 2 milioni di match simulati. La colonna di destra riporta il valore teorico: la freccia di spareggio è indipendente da come sia andato il match, quindi vale 57,0 su ogni riga per costruzione. Una simulazione la restituisce fra 56,8 e 58,5 sulle righe con pochi casi, e quelle oscillazioni sono rumore — pubblicarle suggerirebbe un effetto che non esiste.)*

Questa tabella è la spiegazione, e va letta guardando le due colonne di destra una accanto all'altra.

La colonna della freccia è costante: 57 su ogni riga. Non è un arrotondamento ma una conseguenza di che cosa la freccia sia: viene tirata dopo, non sa nulla di come sia andato il match, e la sua capacità di indicare il più bravo non può quindi dipendere dal margine con cui il totale si è chiuso. È l'argomento del capitolo 7 visto da un'altra angolazione.

La colonna del totale invece cresce. Con un punto di margine ha ragione sei volte su dieci, con due punti sette, con tre otto, con quattro quasi nove. E le due colonne si incrociano subito: già a un solo punto di margine il totale batte la freccia, e un punto è la differenza più comune di tutte, perché capita in quattro match su dieci fra quelli finiti pari.

Sommando le righe, il quadro è questo: in due casi su tre il totale dice qualcosa, e quando parla ha ragione da sei a nove volte su dieci; nel terzo caso tace, e lì la freccia vale 57 contro i 50 di una monetina. TOTALE+ usa il criterio migliore in ciascuna delle due situazioni, invece di usarne uno solo in entrambe. È tutta qui: la regola in vigore fa tirare sempre, anche quando il referto sapeva già la risposta.

### **Che cosa cambia in una gara, e in vent'anni**

In una gara da 64 arcieri la regola in vigore sposta il vincitore in tre quarti di match rispetto al caso, e fa tirare dieci frecce e sei di spareggio. TOTALE+ lo sposta in un match e quattro decimi, e di frecce di spareggio ne fa tirare tre e sei. Quasi il doppio dell'accuratezza con poco più di un terzo degli spareggi.

In un campionato che duri vent'anni sono una trentina di match il cui vincitore cambia, e centoquaranta frecce di spareggio in meno da tirare: trenta arcieri che nell'arco di vent'anni passano il turno per merito invece che per fortuna.

In titoli vinti è molto meno, e conviene dirlo prima che lo dica qualcun altro. Questa sola modifica sposta la probabilità che il titolo vada al più bravo di tre decimi e mezzo di punto per edizione, cioè meno di un titolo in vent'anni. È la ragione per cui il conto si fa sui match e non sui titoli: la regola di parità agisce spesso, ma quasi mai sulla finale, anche se può cambiarne i partecipanti.

---

## 18. Come si scriverebbe l'articolo

Una proposta che non arriva al testo dell'articolo non è una proposta, perché tutto il lavoro sta nei casi che il testo deve coprire. Ecco come lo scriverei.

### **Articolo — Risoluzione della parità a set**

1. Quando dopo 5 set il punteggio è di cinque punti-set per ciascun atleta, vince il match l'atleta che ha totalizzato il maggior numero di punti nelle 15 frecce del match.
2. Se i due atleti hanno totalizzato lo stesso numero di punti, ciascuno tira una freccia nel tempo previsto dal regolamento per una singola freccia, venti secondi in tiro alternato; trenta nel tiro simultaneo degli eventi World Ranking e annunciati, quaranta negli altri, riducibili a trenta dall'organizzatore. Vince l'atleta che realizza il punteggio più alto; a parità di punteggio, quello la cui freccia risulta più vicina al centro del bersaglio; se anche la distanza non li separa, si ripete.
3. Se le due frecce sono giudicate equidistanti, si ripete la procedura del comma 2 fino a che una decisione sia possibile.
4. L'ordine di tiro nella procedura del comma 2 è determinato per sorteggio effettuato dal giudice in presenza dei due atleti.
5. Il totale del match di cui al comma 1 è quello risultante dal referto già compilato per i 5 set, senza ulteriori verifiche. Ciascun atleta ha diritto di prenderne visione prima che la decisione sia comunicata.

Cinque commi, e ciascuno risponde a una domanda che un giudice si porrebbe.

Il primo è la regola, e non richiede alcun dato nuovo: il totale è già sul referto, perché per assegnare i punti-set il giudice ha già dovuto sommare ogni set. Il secondo conserva la freccia di spareggio dove serve davvero, cioè quando il referto non decide, e non tocca i tempi di tiro previsti oggi. Il terzo copre il caso della doppia parità con lo stesso meccanismo del regolamento attuale, senza introdurre nulla di nuovo.

Il quarto è l'unica modifica sostanziale rispetto alla prassi, e la propongo per una ragione che non riguarda l'accuratezza. Oggi allo spareggio parte chi aveva tirato per primo nel primo set, e l'ordine di quel set lo sceglie il meglio classificato della qualificazione, che può scegliere di tirare per secondo, e talvolta lo fa. Chi tira per primo dipende quindi dal piazzamento e insieme dalla preferenza di chi sceglie, e le due cose restano mescolate: chi volesse studiare se partire per primo convenga non riuscirebbe a separarle. Basterebbe registrare la scelta per rendere la domanda studiabile, e sorteggiare l'ordine costa nulla e la rende rispondibile con i dati che una federazione già raccoglie.

Il quinto protegge la verificabilità, che è la ragione per cui questa regola è preferibile ad altre più accurate: l'atleta che perde deve poter vedere il numero che lo ha eliminato.

Quanto a che cosa cambierebbe nella pratica di gara, la risposta è quasi niente, ed è il punto. Il giudice ha già il referto compilato e deve sommare due colonne di quindici numeri, o più semplicemente leggere due totali che il sistema di punteggio elettronico calcola già. In due casi su tre la decisione è più rapida di oggi, perché non richiede di far tornare due atleti in linea di tiro.

## 19. Quanto spettacolo si vuole comprare

C'è un'obiezione a TOTALE+ che va affrontata e non aggirata: fa sparire due spareggi su tre. Oggi in una gara se ne tirano 10,6, con TOTALE+ 3,6, cioè si passa da un pomeriggio con dieci di quei momenti a un pomeriggio con tre o quattro. Chi difende quella scena perché è il momento che il pubblico aspetta ha un argomento legittimo, e queste pagine non lo negano.

Ma la scelta non è secca, ed è il risultato di questo capitolo. TOTALE+ è un punto su una curva, e la curva si ottiene generalizzando la regola in una riga: si usa il totale del match solo se il margine è abbastanza ampio, altrimenti si tira. Con margine minimo di un punto si ottiene TOTALE+; alzando la soglia si mandano allo spareggio sempre più match, fino a riottenere la regola in vigore.

| Margine minimo per usare il totale | Indovina su 100 parità | Arcieri spostati per gara | Spareggi per gara | Quota di parità che va allo spareggio |
|------------------------------------|------------------------|---------------------------|-------------------|---------------------------------------|
| 1 punto (= TOTALE+)                | 63,0                   | +1,37                     | 3,6               | 34 %                                  |
| 2 punti                            | 61,4                   | +1,20                     | 8,0               | 75 %                                  |
| 3 punti                            | 58,9                   | +0,95                     | 9,8               | 93 %                                  |
| 4 punti                            | 57,7                   | +0,81                     | 10,4              | 98 %                                  |
| nessun limite (= regola in vigore) | 57,0                   | +0,75                     | 10,6              | 100 %                                 |

*(Coppia di riferimento 690 contro 680, tabellone da 64 arcieri con 63 match. Il margine è la differenza fra i due totali di 15 frecce.)*

La seconda riga è quella da guardare, confrontata con l'ultima. Con la soglia a due punti si conservano otto spareggi su dieci e si indovina comunque quattro volte e quattro decimi in più ogni cento parità rispetto a oggi. Il pubblico vede quasi la stessa gara di adesso, e nel frattempo mezzo arciere in più a gara passa il turno per merito.

Il che smonta il compromesso come veniva posto. Non è vero che accuratezza e spettacolo si escludano: la scelta secca fra leggere il referto e tirare la freccia era un artefatto dell'aver considerato soltanto regole che decidono tutto in un modo o tutto nell'altro.

Quello che ne esce, per una federazione, è una manopola con i prezzi scritti sopra. Chi voglia il massimo di accuratezza mette la soglia a uno e accetta di scendere a tre spareggi e mezzo a gara; chi voglia tenere lo spettacolo la mette a due, ne conserva otto, e paga meno di due decimi di arciere a gara, cioè un arciere ogni sei gare. Chi non voglia cambiare niente adesso sa quanto gli costa. Nel testo dell'articolo la variante si scrive cambiando tre parole al primo comma, "ha totalizzato almeno due punti in più" e adeguando il secondo.

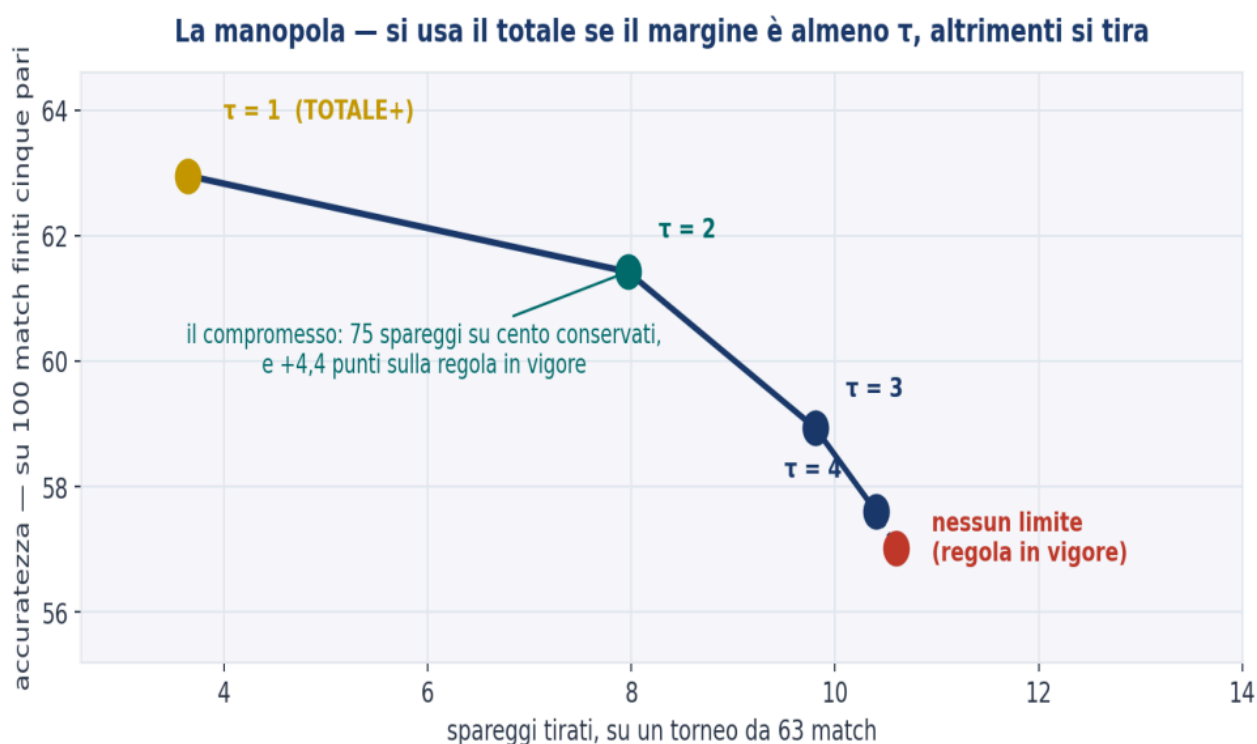


**Figura 6 — La manopola.** Accuratezza contro numero di spareggi, al variare del margine minimo  $\tau$  richiesto per usare il totale del match. Coppia di riferimento 690 contro 680, popolazione dei match finiti cinque pari; gli spareggi sono su un torneo da 63 match. L'estremo destro è la regola in vigore, che equivale a non usare mai il totale.

## 20. Se gli arcieri non fossero come li abbiamo descritti

Fin qui abbiamo assunto una cosa precisa: che le frecce si dispongano in una nuvola tonda, centrata, con le code che calano rapidamente. È il modello standard e per gli errori piccoli funziona benissimo.

Ma ogni arciere sa che non è sempre così. Esiste la freccia che non c'entra niente con le altre, quella partita storta, quella in cui si è persa la trazione, quella con la raffica arrivata al momento sbagliato. E il modello a campana è severissimo su quegli eventi: per un arciere da 690 punti colloca una freccia sotto il sette una volta su 3 milioni, cioè mai in una carriera intera. Chiunque frequenti un campo sa che non è così.

Chiamiamola coda e mettiamola nel conto. Ho verificato tutte le conclusioni precedenti sotto cinque deviazioni dal modello standard (code pesanti, gruppo scenterato, gruppo ovale, deriva durante il match, condizioni ambientali comuni) e il risultato principale è che il ribaltamento del conteggio dei dieci si ritira quando le code sono abbastanza pesanti.

Era prevedibile dal meccanismo. La coda aggiunge modi di comporre lo stesso totale (formula S13): se una freccia può finire molto lontano, il legame ferreo fra più dieci e più frecce basse si allenta, perché una sola freccia buttata via può compensare parecchi dieci. E le due cose si muovono insieme cella per cella, il che conferma che il meccanismo è quello.

| Modello generatore              | Conteggio dei dieci | Quanto è stretto il vincolo aritmetico |
|---------------------------------|---------------------|----------------------------------------|
| standard                        | 46,80               | 0,998                                  |
| coda leggera (1 % delle frecce) | 48,97               | 0,888                                  |
| coda media (3 %)                | 52,60               | 0,831                                  |
| coda marcata (6 %, molto larga) | 59,32               | 0,547                                  |

(Coppia 700-690, sui match con i totali pari. Sopra 50 il criterio torna informativo.)

Con tre frecce anomale ogni cento il criterio torna sopra la monetina, e con sei ogni cento indovina quasi sei volte su dieci: nel mondo delle code marcate contare i dieci è una regola sensata, e in quello standard è peggio di una moneta. Il verso della conclusione dipende quindi da un numero che non conosciamo.

Ma non tutte le deviazioni fanno lo stesso effetto, e questa è la parte che ha richiesto più lavoro. Il gruppo scentrato non ritira l'inversione: la approfondisce, da 46,80 a 41,76 spostando il centro del gruppo di un'ampiezza. Il gruppo ovale nemmeno, e va detto con che forma: i due semiassi stanno in rapporto 3,2 con la media geometrica preservata (la radice del prodotto delle due dispersioni resta  $\sigma$ ) così che il valore 1 riproduca esattamente il modello standard e la dispersione complessiva non cambi. Senza quel parametro il conto non si rifà: con un rapporto di 2 lo stesso calcolo dà 46,73 invece di 46,02.

Che cosa distingue allora le deviazioni che contano da quelle che non contano? Non quanta massa scenda negli anelli bassi, ma se ci scenda concentrata in poche frecce. La prova è una coppia di casi costruita apposta, con la stessa quantità di frecce basse e comportamento opposto.

| Tipo di deviazione             | Frecce da 8 o meno | Vincolo aritmetico | Conteggio dei dieci |
|--------------------------------|--------------------|--------------------|---------------------|
| molte frecce spostate di poco  | 0,2112             | 0,991              | 45,78               |
| poche frecce spostate di molto | 0,2114             | 0,715              | 49,07               |

(Coppia 700 contro 690, popolazione ristretta ai match cinque pari con i totali in parità. Le due configurazioni sono costruite per avere la stessa massa negli anelli bassi.)

Stessa massa bassa fino alla quarta cifra, comportamento opposto: è la concentrazione il meccanismo, non la quantità.

Tradotto per chiunque abbia frequentato un campo, conta la differenza fra due arcieri che tutti riconoscono. C'è quello che tira leggermente largo tutto il giorno, che chiude con parecchi otto e nessun disastro. E c'è quello che tira benissimo e ogni tanto butta via una freccia, che chiude con molti dieci e un sei. Il primo lascia il referto legato al totale, il secondo lo scioglie, ed è il secondo a rendere utile il conteggio dei dieci, il che è controintuitivo perché è anche quello che il pubblico chiamerebbe il più bravo dei due.

## 21. Una polizza, non una regola

Quale dei due mondi è il nostro? Gli arcieri d'élite tirano leggermente largo o buttano via ogni tanto una freccia?

La risposta onesta è che non lo sappiamo, e c'è una ragione strutturale per cui probabilmente non lo sapremo mai. L'informazione su quanto sia marcata la coda sta in eventi rari per definizione (le frecce but-

tate via molto lontano) e stimare quel numero con la precisione che servirebbe richiede circa diecimila frecce per atleta. Sono 140 round di qualificazione: un arciere che ne tirasse uno alla settimana impiegherebbe quasi tre anni, e nessuna quantità di gare in più cambia il fatto che l'informazione stia in eventi che si osservano di rado.

Esiste una regola che aggira il problema invece di risolverlo, e vale la pena spiegarla perché nasce da un'idea elegante. Si chiama CONCORDE, e funziona così: se il totale e il conteggio dei dieci indicano lo stesso arciere, o se uno solo dei due decide, vince quello; se sono entrambi decisivi e si contraddicono, si va allo spareggio.

Un esempio. A fa 145 con 6 dieci, B fa 143 con 7 dieci: il totale dice A, i dieci dicono B, sono in disaccordo e si va allo spareggio. Se invece A fa 145 con 7 dieci contro 143 e 6, i due criteri concordano e A vince senza tirare. E se i totali sono pari a 144 e i dieci sono 7 contro 5, decide l'unico criterio che si esprime, quindi vince chi ha fatto più dieci.

L'idea nasce da un fatto dimostrabile per esaustione: due regole che differiscano solo per l'ordine in cui consultano totale e conteggio dei dieci divergono se e solo se i due criteri sono entrambi decisivi e discordi. Fuori da quell'insieme di casi sono la stessa regola. Tutta la questione aperta vive quindi su un sottoinsieme che un giudice riconosce a colpo d'occhio dal referto, e su quel sottoinsieme non sappiamo quale dei due criteri abbia ragione. Allora non si sceglie: si tira.

| Modello generatore  | Regola in vigore | TOTALE+      | Cascata      | CONCORDE     |
|---------------------|------------------|--------------|--------------|--------------|
| standard            | 57,01            | 62,95        | 59,70        | 58,93        |
| coda leggera        | 57,01            | 61,74        | 59,32        | 59,16        |
| coda media          | 57,01            | 60,04        | 58,79        | 59,62        |
| <b>coda marcata</b> | <b>57,01</b>     | <b>53,27</b> | <b>53,41</b> | <b>59,30</b> |
| gruppo ovale        | 57,01            | 62,12        | 60,42        | 59,04        |
| gruppo scentrato    | 57,01            | 64,05        | 60,20        | 57,90        |

*(Su cento match finiti cinque pari, coppia 690 contro 680. Ogni riga è un generatore diverso; la regola in vigore non cambia per costruzione, perché la freccia è indipendente dal match.)*

La riga in grassetto è quella che conta. Sotto coda marcata TOTALE+ scende a 53,27, cioè sotto la regola in vigore, mentre CONCORDE tiene a 59,30. CONCORDE non è dunque un'alternativa da preferire: è una copertura da comprare se si teme una configurazione che non si sa accertare, e il premio della polizza si conosce: quattro volte su cento a cui si rinuncia se il mondo è quello standard.

Chi scommette sul modello standard e sbaglia non perde soltanto il guadagno: perde anche quello che aveva, perché finisce sotto la regola che voleva sostituire. Su una gara intera, vincere la scommessa vale fra quattro e sei decimi di arciere e perderla ne costa fra cinque e otto, cioè si rischia un po' più di quanto si possa vincere.

La decisione ha quindi quattro voci, e nessuna delle quattro basterebbe da sola: il guadagno è modesto, la perdita è leggermente maggiore, il ramo cattivo scende sotto la regola che si voleva sostituire, e il numero che decide quale dei due mondi sia il nostro non è accertabile.

C'è però una cosa in favore di TOTALE+ che il conto delle poste non cattura, e che restringe parecchio il rischio. La coda che lo danneggia non è una coda qualsiasi: è quella concentrata. Gruppo scentrato, gruppo ovale e deriva non gli fanno niente, anzi, sotto gruppo scentrato TOTALE+ è la migliore di tutte, con 64,05. Il ventaglio si apre su un asse solo, e l'insieme delle configurazioni pericolose è più piccolo di quanto la tabella faccia sembrare.

## 22. Le cinque regole che restano in piedi

A questo punto ci sono più regole ragionevoli e nessuna che batta le altre su tutto. Il modo corretto di presentarle non è una classifica (una classifica è una scelta di pesi travestita da risultato) ma la lista di quelle che nessun'altra batte su tutti i criteri insieme. I criteri sono sette: accuratezza attesa, accuratezza nel caso peggiore, tempo aggiunto alla gara, requisiti strumentali, decisività, verificabilità pubblica, spareggi conservati.

| Regola                        | Vince su                                                                          | Paga su                                                                                 |
|-------------------------------|-----------------------------------------------------------------------------------|-----------------------------------------------------------------------------------------|
| Conteggio pesato, poi freccia | l'accuratezza più alta, e pochi spareggi                                          | non ispezionabile, e la più fragile se il modello si muove                              |
| TOTALE+                       | la migliore fra quelle che si verificano a occhio                                 | robustezza pari alla regola in vigore                                                   |
| Manopola con soglia 2         | conserva tre quarti degli spareggi e guadagna comunque                            | meno accurata di TOTALE+                                                                |
| Regola in vigore              | il massimo di spareggi conservati, e il migliore comportamento sotto coda marcata | l'accuratezza più bassa nel modello standard e in quasi tutte le configurazioni provate |
| CONCORDE                      | il migliore caso peggiore                                                         | quasi nessuno spareggio, accuratezza attesa più bassa                                   |

*(Cinque dei sette vertici non dominati della frontiera a sette criteri, coppia di riferimento.)*

Come si sceglie fra i cinque? Non con un calcolo, e conviene dirlo chiaramente: la scelta è politica e non tecnica, perché richiede di pesare lo spettacolo contro l'accuratezza e non esiste un modo oggettivo di farlo. Quello che i conti possono fare è dire quanto costi ciascuna scelta, e adesso lo dicono.

Tre cose però il calcolo le dice, e restringono il campo.

La prima è che la regola in vigore è battuta da tutte e quattro le altre sotto il modello standard, e da CONCORDE anche sotto coda marcata. Non è però la peggiore in ogni configurazione: sotto coda marcata è la migliore delle tre che leggono il referto, e la riga della tavola va letta con quell'eccezione accanto. Se la si conservasse, la si conserverebbe per lo spettacolo e per nient'altro, e adesso quel prezzo si conosce. Va detto per completezza che esiste una configurazione in cui la regola in vigore non è la peggiore, ed è quella di cui parla il capitolo precedente: sotto coda marcata TOTALE+ scende a 53,3 e la cascata a 53,4, entrambe sotto il 57,0 della freccia. È la ragione per cui quel capitolo esiste, e per cui la scelta fra le cinque non è ovvia.

La seconda è che il conteggio pesato, che è il più accurato, non è proponibile: una decisione che l'atleta non può ricostruire non regge in una gara internazionale.

La terza è che fra TOTALE+, la manopola e CONCORDE la scelta dipende da due domande a cui una federazione può rispondere e un ricercatore no: quanto valga lo spareggio come momento di spettacolo, e

quanto si tema che gli arcieri abbiano code marcate.

C'è infine una regolarità che compare qui per la terza volta e che tornerà anche nelle altre due parti: la regola più accurata sotto il modello nominale è anche la più fragile appena il modello si muove. Il conteggio pesato stima la dispersione assumendo il modello, quindi ne eredita l'errore; le regole grossolane non lo fanno perché non assumono nulla.

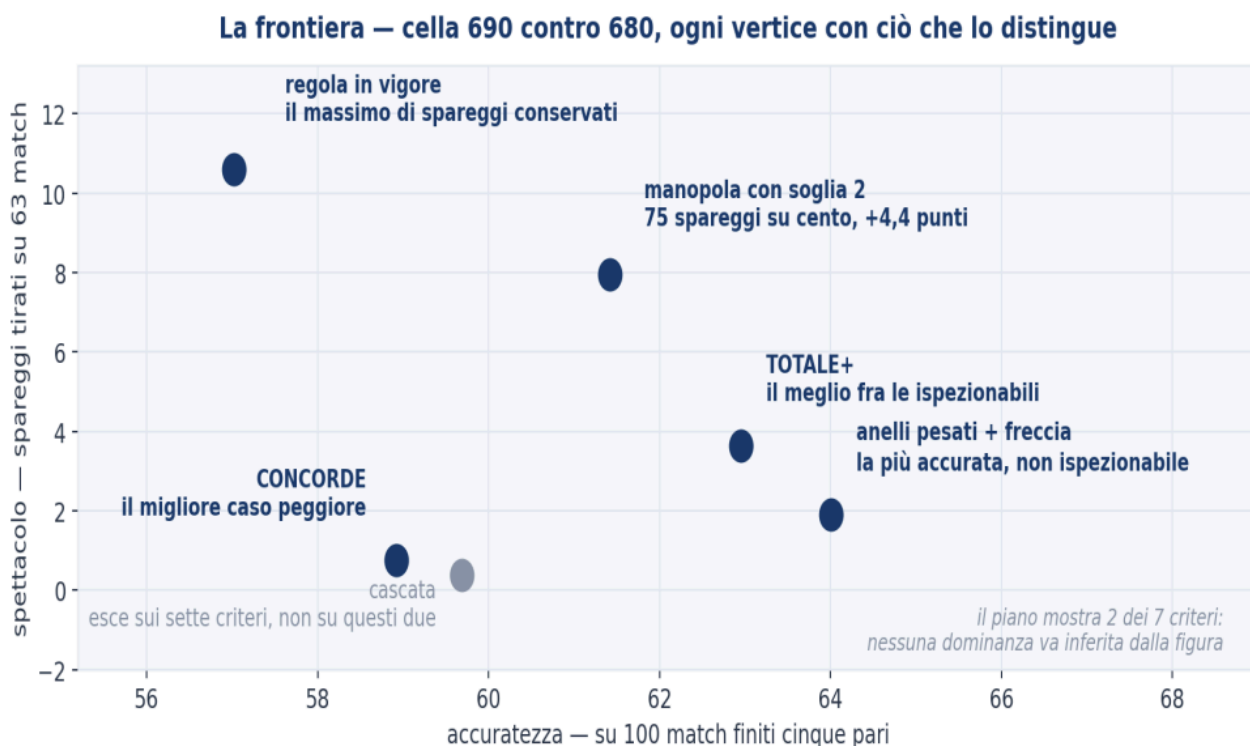


**Figura 7 — La frontiera.** Accuratezza attesa contro spareggi conservati, coppia 690 contro 680, popolazione dei match finiti cinque pari. Il piano mostra due dei sette criteri: la cascata appare qui in posizione ragionevole ed è esclusa per un asse che questa figura non mostra — la robustezza — quindi nessuna dominanza va inferita dal disegno.

## 23. Il tabellone conta più della regola — ma quanto

Tutto quello che precede riguarda un match. Una gara però è un tabellone, e prima di chiudere questa parte vale la pena mettere la regola di parità accanto alle cose con cui compete davvero.

La domanda si pone così: di quanto cambia la frequenza con cui il titolo va al migliore del campo, cambiando ciascuno dei pezzi che compongono una gara? Tutte le quantità che seguono sono misurate sulla stessa macchina, sullo stesso campo e nella stessa unità, cosa che nel dibattito su questi temi non capita quasi mai, perché questi interventi vivono in discussioni separate e nessuno li mette sullo stesso piano.

Servono prima tre parole sui termini, che torneranno per tutta la Parte seconda. Il piazzamento è l'ordine in cui gli arcieri entrano nel tabellone e serve a stabilire chi incontra chi: il primo affronta il sessantaquattresimo, il secondo il sessantatreesimo, e così via; oggi viene dalla classifica del round di qualificazione. Il riseeding è la variante in cui, invece di fissare tutto il tabellone all'inizio, dopo ogni turno si ri-

compongono gli accoppiamenti fra i sopravvissuti, così che il meglio classificato incontri sempre il peggio rimasto. La graduatoria di stagione, infine, è una stima della precisione di ciascun arciere calcolata sui punteggi di tutte le gare che ha disputato nei mesi precedenti: costruirla è il compito della Parte terza, e qui interessa soltanto come origine alternativa del piazzamento.

| Che cosa si cambia                                                              | Escursione fra la peggiore e la migliore opzione |
|---------------------------------------------------------------------------------|--------------------------------------------------|
| La regola di parità, dal sorteggio al meglio ottenibile entro il regolamento    | 1,1 punti                                        |
| — e dalla regola in vigore a quel meglio                                        | 0,4 punti                                        |
| — e fino alla lettura delle distanze, che il regolamento non prevede            | 1,5 punti                                        |
| La struttura del tabellone, fra le cinque messe a confronto in quella corsa     | 1,7 punti                                        |
| Da dove viene il piazzamento, dalla qualificazione alla graduatoria di stagione | 2,7 punti                                        |

*(Probabilità che il titolo vada al migliore del campo, in punti percentuali. Campo compatto da 64 arcieri fra 690 e 665, quarantamila tornei per le prime tre righe e dodicimila per le altre due; le ultime due misurate con il piazzamento che una qualificazione produce davvero, e prese dalle corse dell'archivio di riproduzione — non dalla tavola delle sette strutture del capitolo 28, che confronta un insieme diverso. Il capitolo 40 riporta per esteso la prima colonna.)*

Letta in gare, la tabella dice questo. Oggi il titolo va al più bravo del campo una volta ogni dieci gare. Cambiando la regola di parità dal sorteggio al meglio possibile si arriva a una volta ogni nove: un titolo in più ogni novanta gare. La scelta che una federazione ha davvero, passare dalla regola in vigore alla migliore fra quelle applicabili, vale quattro decimi di punto, cioè un titolo in più ogni duecentocinquanta gare.

Se si potessero misurare le distanze delle frecce l'escursione salirebbe a uno e mezzo, il che porta la regola di parità in mezzo alle altre due voci invece che in fondo; ma richiede strumentazione che oggi non c'è. È in ogni caso coerente con tutto quello che abbiamo visto: la regola agisce soltanto su un match su sei, e lo sposta di poco.

La struttura del tabellone vale una volta e mezza la regola, non sette volte. E qui c'è una cosa da dire, perché il numero cambia molto secondo che cosa si assuma: se si assume che il piazzamento rispecchi l'ordine vero, la stessa escursione si allarga di molto: fra le quattro strutture del capitolo 27 il titolo va al migliore nel 10,4 % dei tornei con le teste protette e nel 16,1 % con il riseeding: quasi sei punti di escursione, contro l' 1,1 della regola di parità. Ma quell'assunzione è falsa, e la Parte seconda misura di quanto.

Il piazzamento, infine, è il primo dei tre: cambiare da dove viene vale una volta e mezza la struttura e due volte e mezza la regola di parità. È il risultato che organizza la Parte seconda, e non era nel piano di questo lavoro.

C'è poi una conferma che arriva per una strada del tutto diversa e che non passa dal piazzamento. Guardando in quale configurazione ciascuna regola di parità dia il suo risultato peggiore, si trova che per sette regole su otto il caso peggiore non è il modello più ostile, ma il campo più squilibrato, quello in cui il

tabellone accoppia 16 arcieri di vertice con 48 molto più deboli. Non è il generatore a legare il caso peggiore: è il campo.

Il tabellone conta dunque più della regola di parità, e da dove viene il piazzamento conta più del tabellone. Ma tutti e tre contano poco, e il motivo si vede fissando la scala: si sta muovendo un numero che parte da una gara su dieci e non arriva a una su sei. Su un campo disperso, dove il primo e l'ultimo distano settanta punti invece di venticinque, si sale a poco più di una gara su quattro e mezzo, ma quella è una gara in cui il migliore si vede a occhio, e la misura conta meno.

Non è una critica al tiro con l'arco. Un tabellone a eliminazione diretta è fatto così, e nessuno sport che ne usi uno può fare molto meglio: è il contesto entro cui va letta ogni discussione su una regola che ne sposta mezzo punto.

---

## 24. Che cosa portarsi a casa dalla Parte prima

**Il margine esiste ed è piccolo.** In una gara da 64 arcieri si giocano 63 match, e circa dieci finiscono pari. Una monetina ne indirizzerebbe correttamente cinque; la regola in vigore cinque e tre quarti; la migliore procedura possibile sei e sei decimi. Fra quello che si fa e quello che si potrebbe fare ballano meno di un arciere a gara.

**Esiste però una regola migliore, e sta in una riga:** vince chi ha totalizzato di più, e se anche i totali sono pari si tira una freccia. Si prende l'85 per cento del margine disponibile, non costa nulla, si verifica a occhio, e fa tirare 3,6 spareggi per gara invece di 10,6.

**Funziona per una ragione che si legge in una tabella.** La freccia di spareggio indovina 57 volte su cento qualunque cosa sia successa nel match, perché è un dato nuovo che non sa nulla di quel che l'ha preceduta. Il totale ne indovina 61 con un punto di margine, 71 con due, 81 con tre. Già un punto di margine batte la freccia, e un margine c'è in due casi su tre.

**Il criterio storico è peggio di una monetina nel punto esatto in cui si applica,** finché gli impatti si distribuiscono come il modello assume, l'Appendice B mostra che solo con code abbastanza pesanti il verso si rovescia. Il conteggio dei dieci, sui match con i totali pari, indica l'arciere sbagliato, perché lì non misura la bravura ma quanto il deficit sia concentrato su poche frecce. E lo si consulta soltanto lì: nelle parità ordinarie della classifica di qualificazione, dopo la parità degli X. Nei match il regolamento in vigore non lo usa affatto, perché a cinque pari si tira la freccia, e nemmeno per le parità che decidono l'accesso alle eliminatorie, dove è previsto lo spareggio.

**Accuratezza e spettacolo non si escludono.** La manopola a due punti conserva otto spareggi su dieci e guadagna comunque quattro volte su cento rispetto alla regola in vigore.

**C'è poi una strada che porterebbe fuori da questo margine stretto,** ed è misurare le distanze invece degli anelli: porta il tetto da un arciere e sei decimi a due tondi per gara, e richiede una precisione di 3 millimetri che non è da laboratorio.

## PARTE SECONDA — LA GARA

---

## 25. La gara è un oggetto solo

La Parte prima ha misurato l'ultimo passo di una gara. Ma una gara non comincia al cinque pari: comincia con un round di qualificazione in cui tutti tirano 72 frecce e vengono messi in classifica, prosegue con un tabellone che decide chi incontra chi in base a quella classifica, e solo dopo arrivano i sei turni di eliminazione con i loro match e le loro eventuali parità.

Per settimane ho trattato questi pezzi come problemi separati, e sbagliavo. Sono un oggetto solo, e la ragione è precisa: il tabellone si costruisce sul piazzamento, e il piazzamento viene da una misura imperfetta. Chi si classifica quarto non è necessariamente il quarto più bravo, e un piazzamento sbagliato cambia chi incontra chi, quindi cambia tutto quello che viene dopo. Finché si assume che il piazzamento sia l'ordine vero, ogni struttura di tabellone appare migliore di quanto sia, ed è l'errore che questa parte corregge, cambiando un numero anche della Parte prima.

La domanda diventa allora una sola: se volessimo costruire la gara migliore possibile, come sarebbe fatta? E siccome "migliore" dipende da che cosa si voglia, la risposta è in quattro versioni. Un modello precisione, che fa vincere il più bravo il più spesso possibile. Un modello spettacolo, che massimizza quello che tiene il pubblico dentro senza però perdere accuratezza. Un modello bilanciato, che migliora l'accuratezza senza togliere niente a nessuno. E un modello minimo, per chi non voglia cambiare nemmeno quello: stessa gara identica, cambiando soltanto come si leggono i risultati. I capitoli dal 36 al 39 li descrivono per intero, dalla prima freccia all'ultima; quelli che vengono prima costruiscono i pezzi con cui sono fatti.

---

## 26. Quanto sbaglia la qualificazione

Prima di tutto il resto, il numero da cui dipende l'intera Parte seconda.

Ho simulato sessantamila round di qualificazione su un campo di 64 arcieri di bravura nota, li ho ordinati per punteggio come fa una gara vera, e ho confrontato l'ordine ottenuto con quello vero.

| Tipo di campo | Frecce | Il primo classificato è davvero il migliore | I primi otto sono gli otto migliori | Errore medio di posizione |
|---------------|--------|---------------------------------------------|-------------------------------------|---------------------------|
| compatto      | 72     | 14,2 %                                      | 0,04 %                              | 8,6 posti                 |
| compatto      | 144    | 18,9 %                                      | 0,3 %                               | 6,5                       |
| bimodale      | 72     | 24,8 %                                      | 1,9 %                               | 6,8                       |
| bimodale      | 144    | 32,1 %                                      | 5,8 %                               | 5,3                       |
| disperso      | 72     | 32,4 %                                      | 5,1 %                               | 4,1                       |
| disperso      | 144    | 41,2 %                                      | 14,7 %                              | 2,9                       |

*(64 arcieri, 60 000 qualificazioni simulate per riga, pareggi risolti a sorteggio: o l'arciere è primo o non lo è. I tre campi sono quelli dichiarati: compatto 690–665, disperso 695–625, bimodale sedici atleti 692–680 più quarantotto 655–630.)*

La prima riga è quella che conta, perché un campo di vertice è compatto per definizione: ci si arriva selezionando. E dice tre cose che vale la pena tradurre una per una.

Il primo classificato è davvero il più bravo del campo 14,2 volte su cento, cioè poco più di una volta su sette. In un circuito da sette gare l'anno, la testa di serie numero uno è l'arciere più forte presente una volta sola, e nelle altre sei è qualcun altro.

I primi otto sono davvero gli otto migliori quattro volte su diecimila. Su 120 000 gare simulate significa meno di una volta su duemila: un giudice che dirigesse una gara alla settimana per quarant'anni la vedrebbe forse una volta. Ogni volta che si parla di proteggere le teste di serie, quindi, si stanno proteggendo otto arcieri di cui in pratica nessuno è l'insieme giusto.

L'errore medio di posizione è di 8,6 posti. Chi meriterebbe il decimo posto si ritrova, in media, fra il secondo e il diciannovesimo, e chi entra nel tabellone come numero cinquanta può benissimo valere il ventesimo, e trovarsi al primo turno contro qualcuno che non avrebbe dovuto incontrare.

### **Perché sbaglia così tanto**

Non è un difetto della qualificazione, ed è importante capirlo per non trarne la conclusione sbagliata. È la stessa cosa che la Parte prima aveva già misurato, vista da un'altra angolazione.

Settantadue frecce distinguono benissimo due arcieri a dieci punti l'uno dall'altro: ci riescono 95 volte su cento, cioè sbagliano una volta su venti. Ma in un campo compatto, sessantaquattro arcieri fra 690 e 665, la distanza fra due vicini di classifica non è di dieci punti. È di quattro decimi di punto. E a quel divario le stesse 72 frecce indicano il più bravo 52,9 volte su cento, cioè quasi come una monetina.

È tutta qui la ragione per cui la qualificazione sbaglia di otto posti e mezzo: non le si chiede di separare arcieri distanti, le si chiede di ordinarne sessantaquattro che stanno tutti dentro venticinque punti. Sarebbe come pretendere che una bilancia da cucina metta in fila sessantaquattro sacchetti di farina che differiscono di un grammo l'uno dall'altro: la bilancia funziona benissimo, è la domanda a essere troppo fine per lo strumento.

E qui c'è la cosa scomoda: il caso peggiore è anche quello che conta di più.

## Quanto è affidabile il piazzamento di qualificazione – 64 arcieri, 60 000 gare, pareggi risolti

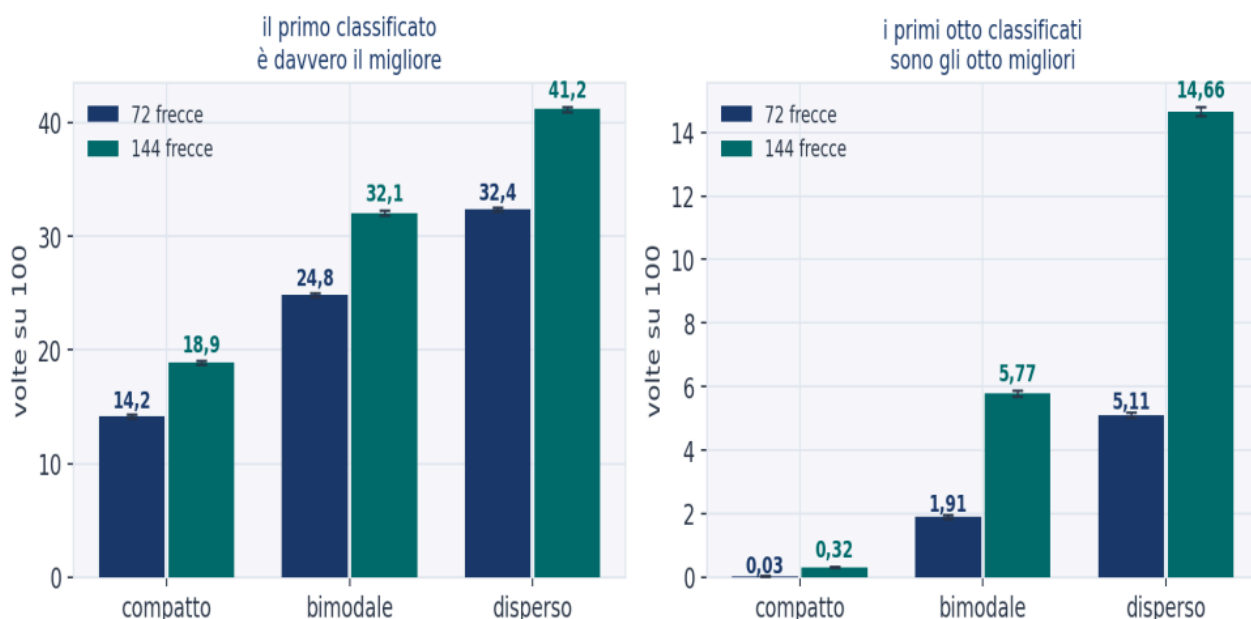


**Figura 8 — Quanto è affidabile il piazzamento di qualificazione.** Probabilità che il primo classificato sia il migliore e che i primi otto siano gli otto migliori, per tre tipi di campo e due lunghezze di qualificazione. 64 arcieri, sessantamila gare simulate per cella.

## 27. Il pedaggio del piazzamento

Se il piazzamento sbaglia, chi lo usa paga, e paga in proporzione a quanto lo usa. Chiamo pedaggio quella perdita, ed è il principio di progetto di tutta questa parte.

Si misura confrontando ogni struttura di tabellone con sé stessa: una volta assumendo che il piazzamento sia l'ordine vero, una volta usando quello che una qualificazione produce davvero. La differenza è il pedaggio, cioè quanto quella struttura perde per il solo fatto di fidarsi di una classifica imperfetta.

| Struttura               | Con piazzamento vero | Con piazzamento reale | Pedaggio     |
|-------------------------|----------------------|-----------------------|--------------|
| Tabellone standard      | 10,77 %              | 10,03 %               | +0,74 ± 0,11 |
| Teste di serie protette | 10,35 %              | 9,66 %                | +0,69 ± 0,11 |
| Ripescaggio dei quarti  | 12,46 %              | 10,83 %               | +1,63 ± 0,11 |
| Riseeding a ogni turno  | 16,11 %              | 11,29 %               | +4,82 ± 0,12 |

(Probabilità che il titolo vada al migliore. Campo compatto, centocinquantamila tornei, confronto appaiato in senso stretto: la stessa estrazione casuale decide l'esito di una data coppia in entrambe le colonne, così che a cambiare sia soltanto l'origine del piazzamento. È il modo di misurare una differenza con un errore molto minore di quello che avrebbe la stessa differenza fra due corse indipendenti: quella porterebbe un errore di circa 0,12 punti, mentre l'appaiamento la porta a 0,11-0,12 su una quantità che è la differenza fra due livelli, ciascuno con errore 0,08-0,09. Il confronto giusto è con la differenza indipendente, non con i livelli. Il ripescaggio dei turni tardivi non compare perché non è fra le strutture che il motore implementa; il capitolo seguente lo misura insieme alle altre.)

Il quadro non è un continuo, ma non è nemmeno una dicotomia: è una scala a tre gradini.

Il tabellone standard e le teste di serie protette pagano lo stesso pedaggio, sette decimi di punto ciascuno, e la differenza fra loro è più piccola dell'errore della misura, cioè non c'è. Il ripescaggio dei quarti paga più del doppio, un punto e sei. Il riseeding paga quattro e otto: da solo, più della somma degli altri tre.

Detto in titoli assegnati, su cento edizioni di un campionato il riseeding ne consegna all'arciere sbagliato quasi cinque in più di quanti ne consegnerebbe se la qualificazione dicesse la verità. Il tabellone standard, nella stessa situazione, ne sbaglia meno di uno. Il riseeding non è una struttura più fragile per sfortuna: è più fragile perché guarda la classifica a ogni turno invece che una volta sola, e ogni volta che la guarda ne eredita l'errore.

*Chi progetta un tabellone sta scegliendo, senza saperlo, quante volte fidarsi di una qualificazione.*

E ne segue una conseguenza che restringe lo spazio delle soluzioni: l'errore del primo turno non si può comprimere restando dentro la struttura. Nessuna delle quattro scende sotto sette decimi di punto, e le differenze fra le prime tre non sono distinguibili. Qualunque tabellone con teste di serie paga quel pedaggio, e per ridurlo non basta cambiare struttura, bisogna migliorare il piazzamento.

### **Una nota che vale per tutte le tabelle che seguono**

Le quantità di questa parte sono simulate e non calcolate esattamente, e capitoli diversi le misurano con corse indipendenti: la stessa quantità compare quindi con cifre leggermente diverse in tabelle diverse. Non sono valori in disaccordo, è la stessa quantità misurata più volte, come pesarsi due mattine di fila.

Gli errori vanno però letti per quello che sono. Le differenze appaiate della tavola qui sopra portano un errore di circa un decimo di punto, perché la stessa alea decide entrambe le colonne; i livelli assoluti misurati da corse indipendenti ne portano di più. Confrontare l'uno con l'altro non ha senso, e in questa tavola i livelli non hanno un errore separato. Non li ho uniformati artificialmente, perché la loro dispersione è essa stessa l'informazione che dice quanto credere alla terza cifra: nella Parte seconda vanno lette due cifre e non tre.

## Il pedaggio del piazzamento — campo compatto, 150 000 tornei appaiati

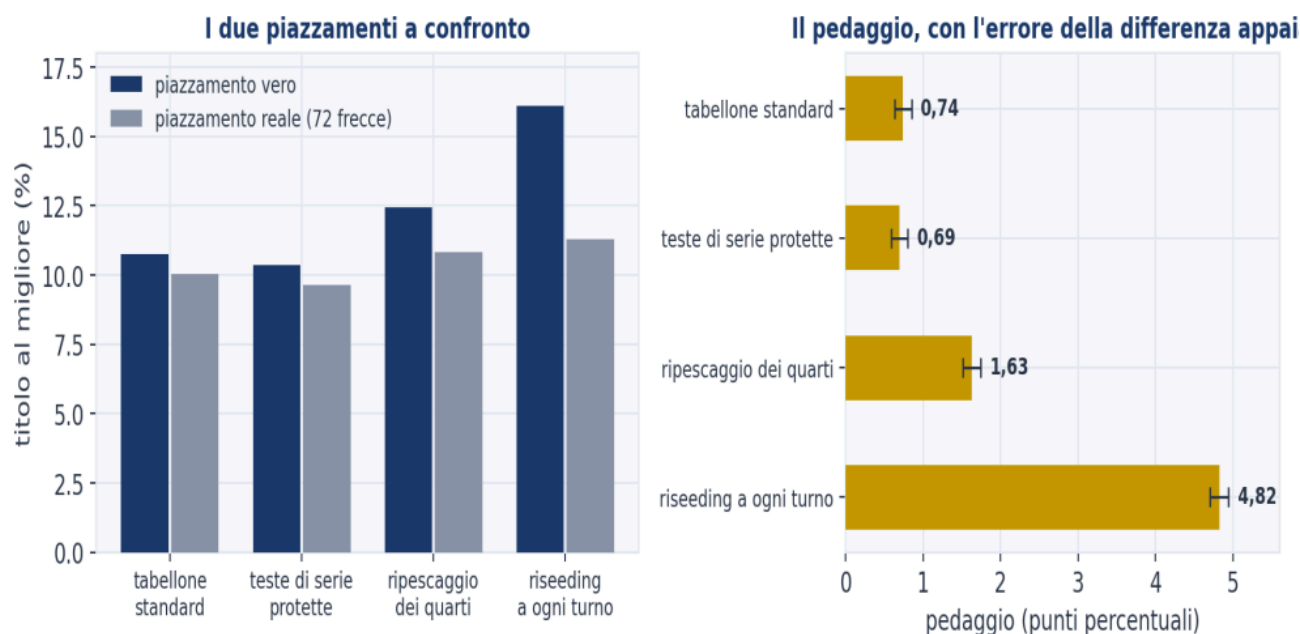


**Figura 9 — Il pedaggio del piazzamento.** Punti percentuali persi da ciascuna struttura passando dall'ordine vero a quello prodotto da una qualificazione a 72 frecce. Campo compatto, centocinquanta-mila tornei, differenze appaiate con barre a un errore standard.

## 28. Sette modi di costruire un tabellone

Ecco l'insieme completo, ciascuna struttura spiegata prima di essere misurata.

**Il tabellone standard** è quello in uso: eliminazione diretta a 64, con gli accoppiamenti fissati all'inizio secondo il piazzamento e chi perde fuori. Sessantatré match.

**Il riseeding** è la variante in cui il tabellone non è fissato: dopo ogni turno si ricompongono gli accoppiamenti fra i sopravvissuti secondo il piazzamento, così che il meglio classificato incontri sempre il peggio rimasto. L'idea è proteggere i migliori a ogni passo invece che soltanto al primo, e il numero di match non cambia.

**Il ripescaggio dei quarti** lascia il tabellone standard fino ai quarti, poi dà una seconda occasione. I quattro eliminati ai quarti giocano un mini-tabellone di tre match, due semifinali di ripescaggio e una finale, e chi ne esce affronta in uno spareggio aggiuntivo il peggio piazzato fra i quattro vincitori dei quarti; chi vince quello spareggio va in semifinale. Il torneo passa così da 63 a 67 match.

Ogni parola di quella descrizione conta, e vale la pena dire perché. La stessa struttura raccontata in modo appena diverso diventa quattro gare distinte, con risultati che vanno dal 7,3 all'11,2 per cento: dire che il ripescato affronti il peggio piazzato invece del meglio piazzato sposta il risultato di quanto sette misure sbagliate potrebbero spostarlo, cioè di una distanza che il caso non produce mai. La versione descritta qui è quella misurata, ed è anche l'unica che una federazione scriverebbe; il ripescato arriva da

una sconfitta, e farlo entrare senza giocare toglierebbe a un quartofinalista vincente un posto guadagnato sul campo.

**Il ripescaggio dei turni tardivi** è una struttura diversa, non un ripescaggio più ampio, e la differenza conta. Il tabellone principale resta quello standard, ma chi perde dagli ottavi in poi scende in un tabellone secondario dove continua a giocare a eliminazione diretta; chi perde al primo turno è invece fuori subito. Scendono quindi quindici arcieri, otto dagli ottavi, quattro dai quarti, due dalle semifinali e chi perde la finale, e il secondario, ordinato per piazzamento in qualificazione, dà il bye al meglio piazzato. Il vincitore del secondario gioca la finale contro il vincitore del principale, che ci arriva imbattuto e la può perdere. 78 match in tutto.

Vale la pena dire come è entrata in questo elenco, perché nessuno l'aveva progettata: è nata da un errore. Doveva essere una doppia eliminazione, e per un difetto nel codice raccoglieva soltanto chi perde dagli ottavi in poi (quindici arcieri) invece di tutti. L'errore è stato scoperto perché il numero diceva una cosa impossibile, che perdere il primo match desse probabilità zero di vincere il titolo, e la struttura è stata tenuta perché misura comunque una gara vera e diversa dalle altre cinque.

La stessa famiglia di funzioni ha poi prodotto un secondo difetto della stessa natura: il tabellone secondario riceveva quindici arcieri e la funzione che lo giocava, scritta per un numero di partecipanti potenza di due, ne scartava uno senza fargli tirare una freccia. Ora quella funzione rifiuta i numeri che non sa gestire, i bye vanno dichiarati, e il numero di match di ogni struttura si conta eseguendola invece di scriverlo a mano. Le due lezioni sono la stessa: un tabellone che accetta qualunque campo nasconde i propri difetti invece di segnalarli.

**Le teste di serie protette** tengono i primi otto classificati lontani fra loro fino ai quarti, con accoppiamenti sorteggiati fra i restanti.

**La doppia eliminazione**, infine, manda chi perde in un tabellone secondario invece di eliminarlo. Nella forma qui simulata la finale è però in gara unica fra il vincitore del principale, ancora imbattuto, e quello del secondario: chi viene dal principale può quindi uscire alla prima sconfitta, e per questo la chiamo doppia eliminazione a finale secca. Tutti i perdenti scendono nel secondario, compresi i trentadue del primo turno, e i due tabelloni procedono in parallelo. Centoventisei match, cioè il doppio: due giornate di gara invece di una.

**La doppia eliminazione in senso stretto** è la stessa struttura con una sola differenza, e sta tutta nella finale. Nella versione precedente il vincitore del principale arriva imbattuto e può uscire alla prima sconfitta, il che non è una doppia eliminazione: qui, se il vincitore del secondario vince la prima finale, i due hanno una sconfitta a testa e se ne gioca una seconda. Sono 126 match, o 127 quando la seconda finale si gioca, e si gioca nel 49,7 per cento dei tornei, quindi la lunghezza attesa è 126,50. È l'unica delle sette strutture che non abbia una lunghezza fissa, e per questo il costo si dichiara come intervallo e non come conteggio.

| Struttura                            | Titolo al migliore | Match giocati |
|--------------------------------------|--------------------|---------------|
| Riseeding                            | 11,56 %            | 63            |
| Ripescaggio dei quarti               | 11,06 %            | 67            |
| Tabellone standard                   | 10,26 %            | 63            |
| Ripescaggio dei turni tardivi        | 11,79 %            | 78            |
| Teste di serie protette              | 9,72 %             | 63            |
| Doppia eliminazione a finale secca   | 10,90 %            | 126           |
| Doppia eliminazione in senso stretto | 11,28 %            | 126–127       |

*(Campo compatto, piazzamento reale a 72 frecce, 25 000 tornei della stessa corsa appaiata, con seme dichiarato nel file dei valori. L'errore standard di ciascun livello è di circa due decimi di punto, quindi quello della differenza fra due righe è di circa tre: le differenze inferiori a mezzo punto vanno lette come indistinguibili. Tutte e sette le righe vengono dalla stessa corsa — stessi campi, stesse qualificazioni, stessa alea dei match, e i match ripetuti sono realizzazioni nuove come richiede una struttura a ripescaggio. La colonna dei match è un conteggio esatto della struttura, misurato a parte con semi diversi e 2 000 ripetizioni.)*

Le due strutture in cima (ripescaggio tardivo 11,79 e riseeding 11,56) non si separano fra loro, e stanno sopra il tabellone standard di più di un punto. La doppia eliminazione con finale secca sta più in basso, a 10,90, pur giocando il doppio dei match.

### **Che cos'è l'errore standard, e perché senza non si legge niente**

Ricorre in tutta questa parte, e conviene spiegarlo una volta sola.

Questi numeri vengono da simulazioni, e una simulazione dà ogni volta un risultato appena diverso, come contare quante teste escono in mille lanci di moneta: circa cinquecento, ma mai esattamente cinquecento. L'errore standard misura quanto oscilli quel «circa».

La regola pratica è semplice. Due valori distanti meno di due errori standard potrebbero benissimo essere lo stesso numero misurato due volte; a tre o più la differenza è reale. Su questa tabella l'errore standard di ogni livello vale circa due decimi, e quello della differenza fra due strutture circa tre: due strutture che distino mezzo punto vanno considerate pari, e due che ne distino uno e mezzo davvero diverse.

Ed è qui il ribaltamento che questa parte cercava. Con il piazzamento vero il riseeding staccava tutti di quattro punti; con quello reale il ripescaggio lo raggiunge, pagando un quarto del suo pedaggio.

Le due strutture arrivano allo stesso punto per vie opposte, e vale la pena dirlo perché è la cosa che si porta via da questo capitolo. Il riseeding corregge gli errori del piazzamento riordinando, quindi ne eredita il rumore cinque volte invece di una, e lo fa senza spendere un match in più. Il ripescaggio li corregge facendo giocare altri match, e un match non consulta la qualificazione: misura direttamente. Le due strade portano allo stesso risultato, ma una costa sessantatré match e l'altra fra sessantasette e settantotto, secondo quanto sia largo il ripescaggio.

Ci si aspetterebbe allora che il ripescaggio, essendo più robusto, sorpassi il riseeding quando la qualificazione peggiora. Non succede, e l'ho verificato: a 36 frecce il vantaggio del riseeding è di due decimi, a 72 di sei, a 144 di due. I tre valori portano ciascuno un errore di poco più di due decimi, quindi la sola cosa che i dati sostengono è che il vantaggio non si chiude peggiorando la qualificazione, il riseeding

parte da così tanto più in alto che non viene raggiunto. Che si allarghi migliorandola è plausibile e coerente col meccanismo, ma su questi numeri non è dimostrato.

Sui tre tipi di campo, poi, il segno non cambia mai, ma la grandezza varia di un fattore otto.

| Campo    | Riseeding | Ripescaggio | Differenza   |
|----------|-----------|-------------|--------------|
| compatto | 10,9 %    | 10,9 %      | -0,05 ± 0,17 |
| disperso | 24,4 %    | 24,5 %      | -0,16 ± 0,21 |
| bimodale | 18,3 %    | 18,9 %      | -0,57 ± 0,18 |

*(Piazzamento reale a 72 frecce, centocinquantamila tornei della stessa corsa, con seme dichiarato nel file dei valori, confronto appaiato.)*

Le due strutture non si separano quasi mai, e dove si separano è il ripescaggio ad avere la meglio. Su campo compatto e su campo disperso la differenza è dentro il rumore. Su campo bimodale il ripescaggio vince di poco più di mezzo punto con un errore di due decimi, che è una differenza che il caso non spiega.

La lettura è diversa da quella che ci si aspetterebbe. Il riseeding mette il migliore piazzato nella parte più facile del tabellone; il ripescaggio dà una seconda occasione a chi è stato eliminato da un incontro sfortunato. Su un campo dove metà degli arcieri sono indistinguibili fra loro, la seconda cosa conta più della prima.

E su un campo di vertice, che è compatto, la scelta fra le due si decide sui quattro match in più del ripescaggio e non sull'accuratezza. Proprio dove il titolo conta di più, il criterio tecnico tace e decide la logistica.

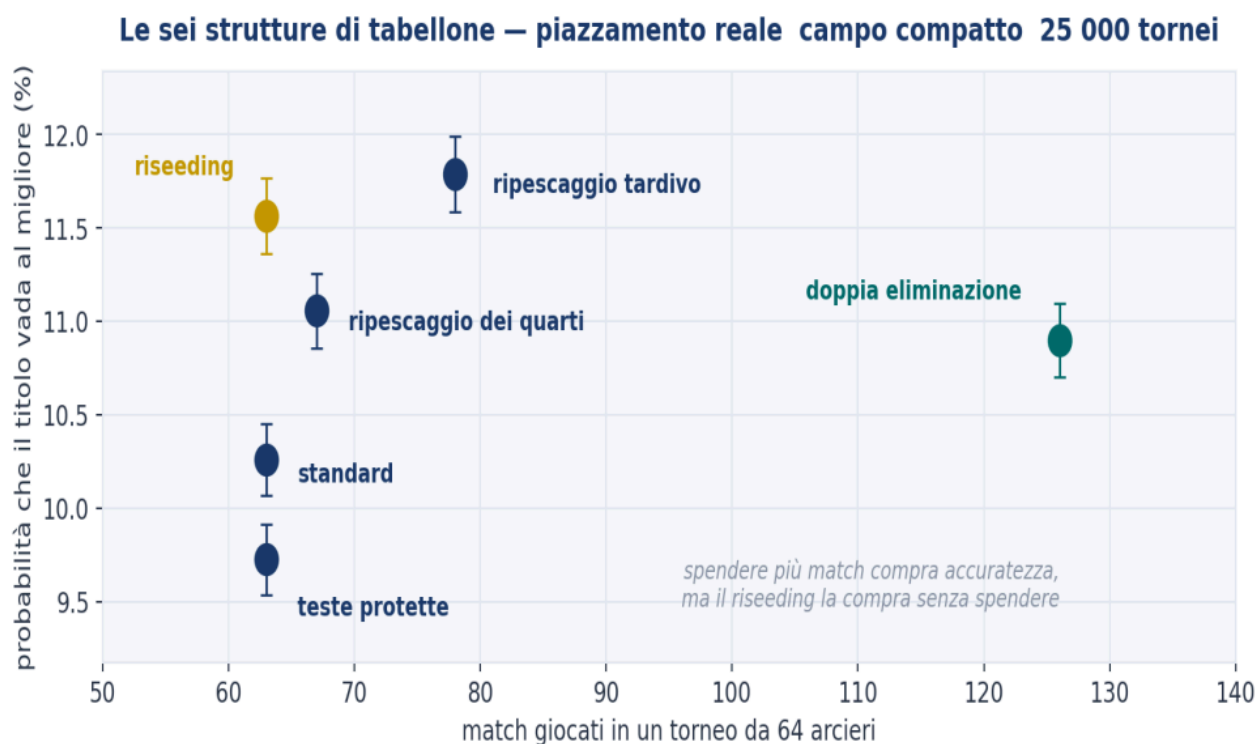


**Figura 10 — Le strutture di tabellone.** Probabilità che il titolo vada al migliore contro numero di match giocati. Campo compatto, piazzamento reale a 72 frecce, venticinquemila tornei della stessa corsa, con seme dichiarato nel file dei valori.

---

## 29. La doppia eliminazione: giusta in teoria

Merita un capitolo a sé, perché è la struttura che tutti propongono quando si lamentano dell'eliminazione diretta, e perché il verdetto è duplice.

Da un lato è strategicamente sana. La preoccupazione legittima è che un arciere possa trarre vantaggio dal perdere di proposito, per finire in un tabellone secondario più facile. Non succede, e il conto è netto: un arciere che vinca il primo match ha davanti a sé una gara in cui su cento tornei ne vince circa dodici; se lo perde apposta per scendere nel secondario, di tornei ne vince poco più di uno. Perdere di proposito costa quindi 10,8 punti di probabilità sul titolo, una differenza venti volte più grande dell'errore della misura. Nessun incentivo, e nessuna ambiguità.

Dall'altro costa il doppio. Con 64 arcieri e tutti i perdenti nel secondario si giocano 126 match invece di 63, cioè due giornate invece di una, e il risultato è 10,90 per cento.

Perché una seconda occasione aiuta, e perché aiuta meno di quanto costi? Un secondo percorso non è una seconda misura della stessa cosa: è una misura in più, ma di qualcosa di più facile. Il tabellone secondario fa giocare molti match fra arcieri deboli, e il suo vincitore arriva alla finale avendo battuto avversari mediocri; quindi il fatto che sia arrivato là dice meno di quanto direbbe lo stesso percorso nel tabellone principale. Resta però una misura vera, fatta sul campo invece che sulla qualificazione, ed è questo che spiega il guadagno: si aggiunge informazione dove il piazzamento ne aveva poca. Il prezzo è che per comprarne un punto e mezzo servono sessantatré match in più, mentre il riseeding ne compra altrettanti senza spenderne nessuno.

Il confronto vero non è però con il tabellone standard, che sta al 10,26, ma con il riseeding: 11,56 per cento con sessantatré match, cioè un risultato migliore per metà del tempo di gara. È questo che rende la doppia eliminazione a finale secca difficile da proporre: non arriva nemmeno allo stesso posto, e chiede una giornata di gara in più.

Si può obiettare che la colpa sia della finale secca e non della struttura, ed è un'obiezione misurabile: basta far valere davvero la seconda sconfitta, cioè obbligare il vincitore del secondario a battere due volte il campione del principale. Misurata sulla stessa corsa, quella variante dà 11,28 per cento contro il 10,90 della finale secca, e costa 126 match o 127 secondo che la seconda finale si giochi. Il guadagno c'è ed è di quattro decimi, quindi la finale secca costava qualcosa; ma resta sotto il riseeding, che dà 11,56 con la metà dei match. L'obiezione è fondata e non cambia la conclusione: la doppia eliminazione, in entrambe le forme, spende il doppio del tempo di gara per arrivare più in basso.

Va detto da dove viene questo numero, perché la prima edizione lo accompagnava con una cautela che nel frattempo è scaduta. Quella cautela diceva che la misura andava rifatta, perché una corsa precedente riusava l'alea del primo incontro: il vincitore del primo scontro fra due arcieri era costretto a vincere anche il rientro, e per una struttura in cui il rientro è il meccanismo stesso quel difetto è determinante.

Il motore corrente non ha più quel difetto. L'alea è legata alla coppia e al numero dell'incontro, quindi ogni rientro è una realizzazione nuova, e la funzione che decide un match si rifiuta di procedere se una coppia si incontrasse più volte di quante realizzazioni siano state estratte, invece di riusarne una. Il 10,90 riportato qui viene da quel motore, ed è riprodotto dal produttore congelato: la cautela della prima edizione va tolta e non ripetuta.

Resta invece vero il risultato sugli incentivi, calcolato a parte: perdere di proposito il primo match costa 10,8 punti di probabilità sul titolo, venti volte l'errore della misura.

È un raro caso in cui non c'è compromesso da discutere, perché la doppia eliminazione a finale secca costa il doppio e rende meno del riseeding. In un lavoro che ripete a ogni capitolo che ogni scelta è uno scambio, un caso in cui non lo è merita una riga esplicita.

Una nota tecnica per chi volesse adottarla comunque: il tabellone secondario va accoppiato per ordine di eliminazione e non per piazzamento, perché ha già perso i migliori e non c'è ordine da azzeccare.

---

### 30. Il ripescaggio ha una dose

Mettendo in fila i tre gradi di ripescaggio si vede qualcosa che nessuno aveva previsto.

| Ampiezza del ripescaggio               | Titolo al migliore | Match |
|----------------------------------------|--------------------|-------|
| nessuno (tabellone standard)           | 10,3 %             | 63    |
| solo i quarti                          | 11,1 %             | 67    |
| dagli ottavi in poi — quindici arcieri | 11,8 %             | 78    |
| tutti (doppia eliminazione)            | 10,9 %             | 126   |

*(Campo compatto, piazzamento reale a 72 frecce, centocinquantomila tornei della stessa corsa, con seme dichiarato nel file dei valori; la doppia eliminazione e il ripescaggio tardivo sono misurati a parte, con corse dedicate.)*

Aperto ai soli quarti, il ripescaggio compra otto decimi di punto sul tabellone standard. Aperto dagli ottavi in poi (i quindici arcieri che perdono dagli ottavi alla finale) ne compra uno e mezzo. Con la doppia eliminazione a finale secca soltanto sei decimi, spendendo il doppio dei match.

La curva sale e poi scende: il massimo non è agli estremi ma in mezzo, nella struttura che ripesca dagli ottavi. E la ragione sta nella finale.

Ogni match in più è una misura in più, e una misura non consulta la qualificazione: la sostituisce. Per questo il ripescaggio tardivo batte quello dei soli quarti. Ma la doppia eliminazione mette il campione del principale, arrivato imbattuto dopo sei incontri, a giocarsi tutto in una partita sola contro chi viene da un tabellone di sconfitti. Quella finale secca butta via l'informazione raccolta nel percorso, e ne butta via più di quanta i match aggiuntivi ne producano.

Non è quindi vero che spendere più match compri sempre accuratezza. È vero che ne compra fino al punto in cui la struttura non comincia a scartare quello che ha misurato.

Vale la pena dire come questa struttura sia entrata nell'insieme, perché non l'aveva progettata nessuno. È emersa da un difetto di implementazione, un ripescaggio che doveva essere una doppia eliminazione e raccoglieva soltanto chi perde dagli ottavi in poi, e si è rivelata una struttura reale, coerente, e informati-

va proprio perché sta in mezzo. Uno spazio di progetto non si esplora da solo: non avevo pensato a un ripescaggio a metà strada finché un errore non l'ha costruito.

### 31. Da dove far venire il piazzamento

Se il pedaggio del primo turno non si comprime restando dentro la struttura, l'unico modo di ridurlo è migliorare il piazzamento. Ci sono tre strade, e hanno prezzi molto diversi. Raddoppiare la qualificazione portandola a 144 frecce, il che costa mezza giornata di gara. Usare la classifica di stagione invece del punteggio del giorno, il che non costa niente perché quei dati sono già stati tirati nelle gare dei mesi precedenti. Oppure non fare nulla, che è l'opzione corrente.

Prima di misurarle serve però dire con precisione che cosa sia questa classifica di stagione, perché graduatoria è una parola che copre almeno quattro cose diverse e la scelta fra loro non è indifferente.

Non è la classifica a punti dei circuiti, quella che a ogni gara ordina i presenti e assegna 25 punti al primo, 20 al secondo, 16 al terzo e così a scendere fino a uno per il quindicesimo, zero agli altri e nulla a chi è assente. Non è la somma dei punteggi di qualificazione, e non è nemmeno la loro media. È la stima di precisione prodotta dalla procedura della Parte terza, applicata ai punteggi dei round di qualificazione delle gare precedenti: quegli stessi punteggi, corretti per la difficoltà di ciascuna gara e poi ristretti verso la media del gruppo in proporzione a quante gare l'atleta abbia disputato. Chi voglia il dettaglio lo trova al capitolo 46, dove la procedura è scritta passo per passo, e al 47, dove è percorsa su un esempio.

La distinzione conta, e si può misurare. Su 800 stagioni simulate con 64 arcieri, 10 gare e 6 presenze a testa, i quattro metodi ordinano il campo con questa fedeltà all'ordine vero.

| Su che cosa si costruisce la graduatoria        | Accordo con l'ordine vero |
|-------------------------------------------------|---------------------------|
| Stima di precisione (procedura del capitolo 46) | 0,961                     |
| Classifica a punti per piazzamento              | 0,901                     |
| Media dei punteggi di qualificazione            | 0,906                     |
| Somma dei punteggi di qualificazione            | 0,906                     |

*(64 arcieri fra 690 e 665, 10 gare da 72 frecce, 6 presenze a testa sorteggiate, effetto di gara normale con deviazione standard di 11 punti, 800 stagioni. L'accordo è la correlazione di ranghi con l'ordine vero di bravura. L'effetto di gara è il parametro che decide questi livelli, ed è dichiarato e non misurato: le differenze fra righe reggono alla riparametrizzazione, i livelli assoluti no.)*

L'accordo si legge così: vale 1 se la graduatoria mette gli arcieri esattamente nell'ordine vero, e 0 se li mette in un ordine che con la bravura non ha niente a che fare. Tutte e quattro le righe stanno molto in alto, quindi la domanda non è se funzionino ma quanto le differenze fra loro contino, e cinque centesimi di accordo, come vedremo nella tabella successiva, valgono circa un punto di probabilità sul titolo.

Media e somma coincidono, perché a presenze uguali producono lo stesso ordine, e stanno cinque centesimi sotto la procedura del capitolo 46. La classifica a punti sta sotto le altre di quasi due centesimi: normalizza per gara, quindi tiene conto meglio delle condizioni del giorno, ma perde informazione perché comprime a zero tutti i piazzamenti oltre il quindicesimo: un arciere arrivato ventesimo e uno arrivato sessantesimo contano uguale, e non lo sono.

Nel resto di questa parte, graduatoria significa sempre e soltanto la prima riga di quella tabella.

| Origine del piazzamento    | Accordo con l'ordine vero | Tabellone standard | Con riseeding |
|----------------------------|---------------------------|--------------------|---------------|
| Qualificazione, 72 frecce  | 0,813                     | 9,7 %              | 11,1 %        |
| Qualificazione, 144 frecce | 0,896                     | 10,2 %             | 12,2 %        |
| Graduatoria di 5 gare      | 0,953                     | 10,2 %             | 12,7 %        |
| Graduatoria di 10 gare     | 0,976                     | 10,7 %             | 13,5 %        |
| Graduatoria di 20 gare     | 0,988                     | 10,4 %             | 14,0 %        |

*(Campo compatto, dodicimila tornei, corsa propria di questo capitolo. L'accordo è la correlazione fra l'ordine di piazzamento e quello vero, e viene da una corsa precedente: non è prodotto da questo esperimento e va letto come riferimento qualitativo. Le graduatorie sono prodotte dalla procedura della Parte terza, non da un modello approssimato. Questa tavola serve a confrontare fra loro le origini del piazzamento e le due strutture di tabellone: i suoi livelli non vanno letti come la misura del modello bilanciato, che viene dalla corsa unica del capitolo 40 e vale 13,7 per cento.)*

Tre letture, e tutte e tre sono raccomandazioni.

**La prima.** Con il riseeding, una graduatoria costruita su cinque gare batte una qualificazione da 144 frecce, cioè una qualificazione lunga quanto due di quelle di oggi: 12,7 contro 12,2 per cento. Usa dati che sono già stati tirati, non costa una freccia, non costa un'ora, non richiede di cambiare il programma di gara. Richiede soltanto di decidere che il piazzamento venga da lì.

**La seconda, ed è la più importante.** Su un tabellone fisso, un piazzamento migliore non serve quasi a niente. Passando dalla qualificazione corta alla graduatoria di 20 gare, cioè da un accordo con l'ordine vero di 0,81 a uno di 0,99, che è quasi la perfezione, il tabellone standard passa da 9,7 a 10,4 per cento, che dentro l'errore della misura vuol dire non guadagnare nulla. Con il riseeding gli stessi identici dati portano da 11,1 a 14,0.

Detto in gare: comprando la classifica migliore possibile e lasciando il tabellone com'è, il titolo continua ad andare al più bravo una volta ogni dieci. Comprandola e cambiando anche il tabellone, ci va una volta ogni sette. La classifica migliore, da sola, è una spesa che non si vede.

Il riseeding non è dunque necessariamente la struttura più accurata (il capitolo precedente mostra che il ripescaggio le sta alla pari o sopra) ma è l'unica che converta in accuratezza un piazzamento migliore, ed è per questo che entra nel modello bilanciato.

**La terza.** I due interventi non si sommano in modo semplice, e ora posso dire di quanto. Misurati sulla stessa corsa, la qualificazione doppia da sola vale mezzo punto e il riseeding da solo uno e quattro; applicati insieme valgono due e mezzo, contro l'uno e nove che darebbero sommandosi. L'eccesso è di sei decimi con un errore di cinque: il segno dice che la coppia rende più della somma, ma la misura non lo separa da zero, e la conclusione onesta è che con questi dati non si decide né in un senso né nell'altro.

Ne segue comunque l'ordine pratico, che nessuna tabella di confronto separato poteva mostrare: prima si cambia struttura, che costa una riga di regolamento, e solo dopo si migliora il piazzamento, che costa frecce e ore.

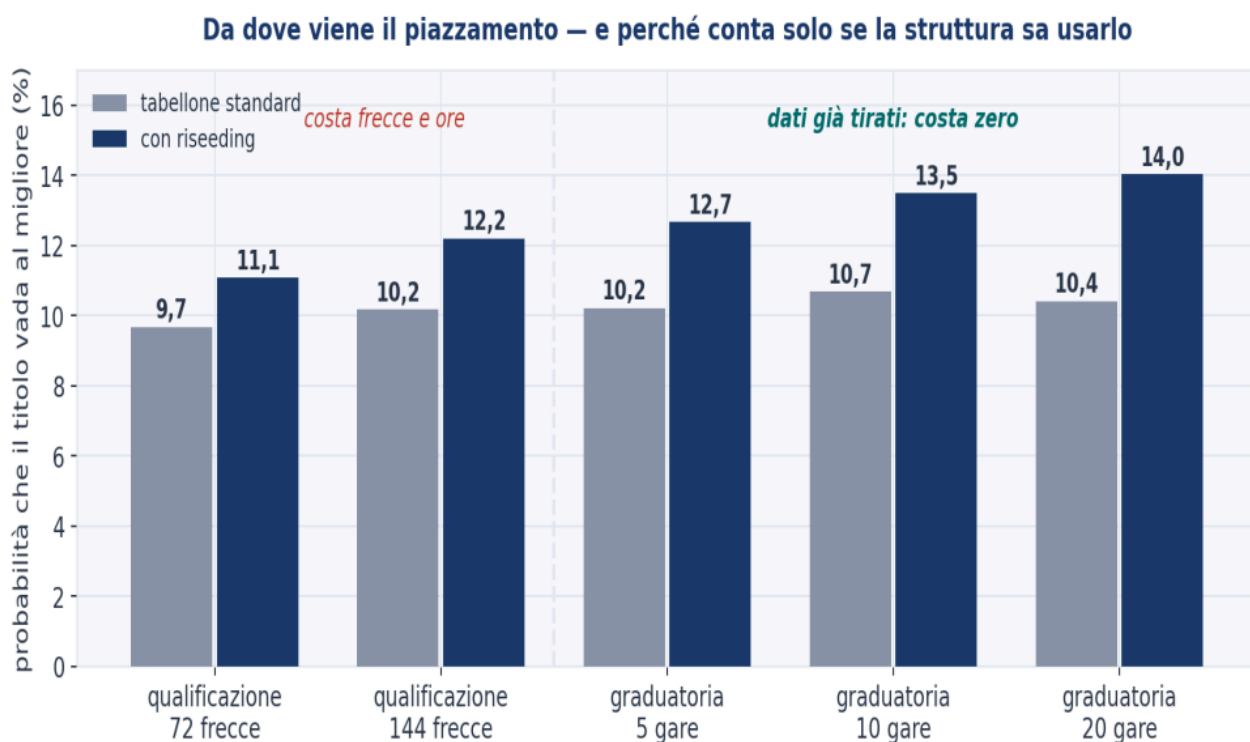


**Figura 11 — Da dove viene il piazzamento.** Probabilità che il titolo vada al migliore per cinque origini del piazzamento e due strutture di tabellone. Campo compatto, dodicimila tornei della corsa dedicata di questo capitolo. Le prime due colonne costano frecce e ore; le altre tre usano dati già disponibili.

## 32. Cinque modi di giocare 15 frecce

Un match si gioca con 15 frecce a testa, e il regolamento stabilisce come vadano raggruppate: 5 set da tre, due punti a chi vince il set, primo a sei. Ma quel raggruppamento è una scelta fra molte possibili, e la domanda di questo capitolo è se sia la scelta giusta o se, cambiando il modo di distribuire le stesse 15 frecce, si possa far vincere più spesso il più bravo.

Ne ho provati cinque. Il primo è quello attuale. Il secondo taglia i set al minimo: 15 set da una freccia sola, ciascuno vinto da chi piazza meglio quel singolo tiro, e vince il match chi arriva a sedici punti. Il terzo fa l'opposto, 3 set da cinque frecce, primo a quattro.

Il quarto e il quinto sono la stessa idea con due tarature diverse, e sono quelli su cui avevo puntato. Si gioca a set da tre frecce come oggi, ma dal quarto set in poi il match si chiude appena uno dei due conduce di almeno due punti-set; se non accade, si prosegue fino a un tetto (sette set nella prima taratura, nove nella seconda) e lì vince chi conduce, oppure, se si è pari, si applica la regola di spareggio. L'idea è che la durata si adatti all'incertezza: due arcieri che si equivalgono tirano più a lungo, un match squilibrato finisce presto. Mi sembrava la cosa più sensata di tutte.

| Formato             | Divario 6 | Divario 11 | Divario 25 | Frecce per match |
|---------------------|-----------|------------|------------|------------------|
| attuale, 5 set da 3 | 67,99     | 79,40      | 95,29      | 13,5             |
| set corti, 15 da 1  | 68,23     | 79,71      | 95,45      | 13,9             |
| set lunghi, 3 da 5  | 67,97     | 79,37      | 95,26      | 13,5             |
| variabile, tetto 7  | 67,11     | 78,04      | 94,06      | 12,9             |
| variabile, tetto 9  | 67,13     | 78,06      | 94,06      | 13,0             |

*(Su cento match, ancoraggio 700, calcolo esatto. La parità a punti-set è risolta con TOTALE+ in tutte le righe, con la concordanza calcolata sulla popolazione dei match che quel formato manda in parità e non su quella del formato in vigore — la regola dev'essere la stessa per tutti, altrimenti si confronterebbero formati e regole insieme, e con la freccia di spareggio al suo posto ogni riga scende di circa un punto senza che le distanze fra righe cambino. L'ultima colonna è calcolata a 6 punti di divario e non vale per le altre due: le frecce consumate calano al crescere del divario, perché i match equilibrati finiscono prima — a 11 punti l'attuale ne consuma 13,2 e il variabile 12,8.)*

I numeri dicono che fra i cinque non c'è quasi differenza. Fra due arcieri distanti 11 punti il formato attuale indica il più bravo 79,4 volte su cento; il migliore dei cinque, che sono i set da una freccia sola, arriva a 79,7. Tre decimi di differenza ogni cento match, che in una gara da 63 significano cambiare il vincitore di un quinto di match: bisognerebbe guardare cinque gare intere perché un solo arciere finisca dall'altra parte del tabellone. Nessuno se ne accorgerebbe in una stagione.

Il formato a lunghezza variabile, però, non si limita a non guadagnare: perde. A 11 punti di divario indica il più bravo 78,0 volte su cento contro le 79,4 dell'attuale, e in cambio fa risparmiare quattro decimi di freccia a match. Si rinuncia a un punto e quattro decimi di accuratezza per risparmiare meno di mezza freccia per match, ed è uno scambio che nessuno accetterebbe.

Ma la ragione per cui va scartato è un'altra, ed è molto più netta. Quel formato prometteva di adattare la durata all'incertezza, e non lo fa. Il tetto si raggiunge nell'1,7 per cento dei match con la taratura più stretta, e nello 0,15 con l'altra: in una gara da 63 match il tetto entra in gioco una volta sola con la prima taratura, e con la seconda una volta ogni dieci gare, cioè una volta a stagione in un circuito. In 99,2 match su cento con la prima taratura, e in 99,9 con la seconda, la lunghezza variabile non varia affatto: si comporta come un formato a lunghezza fissa con un tetto decorativo che nessuno tocca mai.

Vale la pena fermarsi su questo, perché è il tipo di cosa che i numeri sull'accuratezza non mostrano. Se avessi valutato i cinque formati soltanto sull'asse che mi interessava, avrei scartato la lunghezza variabile dicendo che costa un punto e mezzo, un'obiezione discutibile, a cui si può rispondere che il risparmio di frecce vale qualcosa. Guardando invece quante volte il tetto entri in gioco si scopre che il formato non fa la cosa per cui era stato proposto, e quella non è un'obiezione: è una constatazione che chiude la questione.

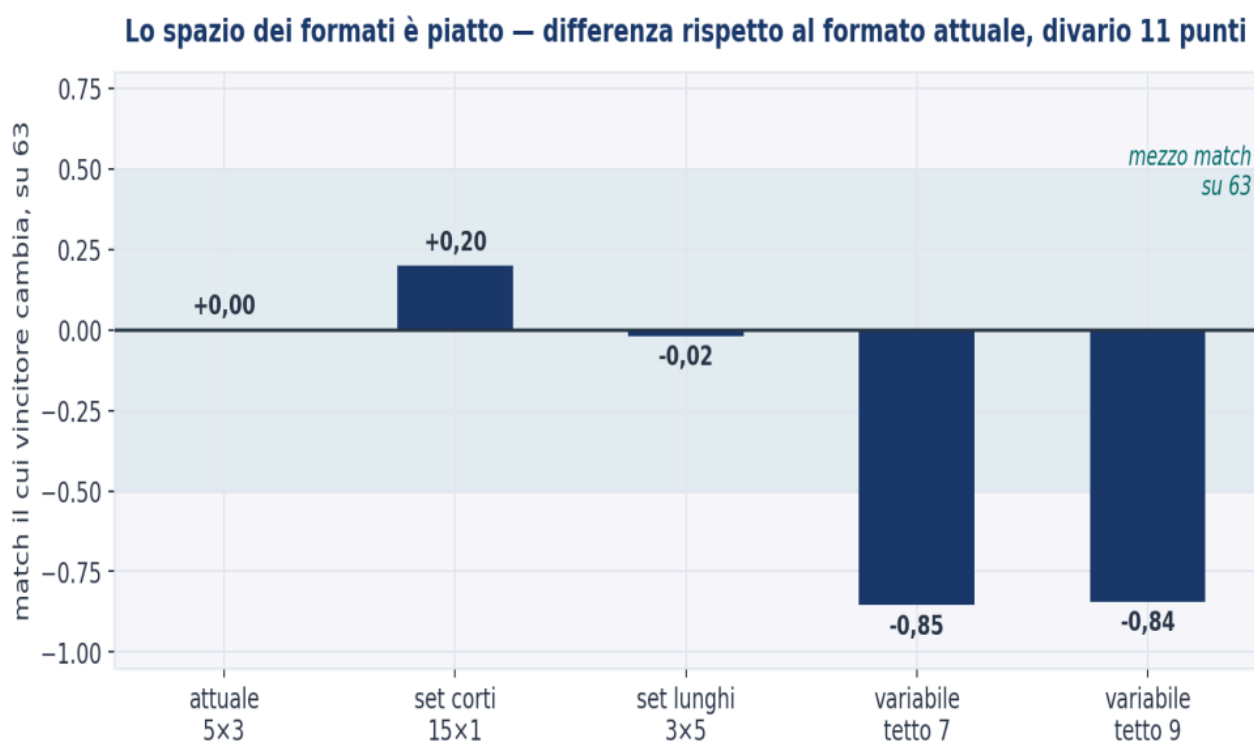


**Figura 12 — Lo spazio dei formati è piatto.** Differenza rispetto al formato attuale in match il cui vincitore cambia, su un torneo da 63. Divario 11 punti, ancoraggio 700. La banda indica mezzo match.

### Un conto aperto con la Parte prima

C'è infine una nota che chiude una questione lasciata indietro, e riguarda il formato a set in quanto tale, quello che nel gergo internazionale si chiama set system. La domanda è se quella struttura, di per sé, costi accuratezza rispetto al semplice sommare le frecce.

La risposta è che costa circa un punto ogni cento match: non dieci, non cinque. Ma con TOTALE+ come regola di parità quel costo quasi sparisce, e a undici punti di divario il set system arriva a 79,40 contro un cumulativo di 79,31, cioè lo supera.

Il costo del set system è quindi quasi tutto nella regola di spareggio e non nel formato. Con la regola in vigore si paga un punto; con TOTALE+ non si paga quasi niente e si tengono tutti i vantaggi della struttura a set. È l'argomento più forte a favore del formato attuale emerso da tutto questo lavoro, e nessuno l'aveva formulato perché nessuno aveva mai separato il costo del formato da quello della regola che lo chiude.

## 33. Che cosa vuol dire spettacolo, in modo che si possa contare

Il punto debole di ogni discussione sui formati sportivi è che lo spettacolo resta implicito: tutti sanno che cosa intendono, nessuno lo definisce, e la discussione non converge mai.

Non pretendo di misurare lo spettacolo. Propongo di misurarne cinque conseguenze, dichiarate in anticipo perché nessuno possa accusarmi di averle scelte dopo aver visto i risultati. Quanto spesso il match arrivi all'ultimo set, cioè quanto spesso la decisione non sia già presa quando l'attenzione è massima.

Quanto spesso cambi la testa dopo il primo set. Quanta incertezza resti all'inizio dell'ultimo set. Quanti set duri in media. E quante frecce si tirino in tutto.

| Formato            | Arriva all'ultimo set | Cambia la testa | Set attesi |
|--------------------|-----------------------|-----------------|------------|
| set lunghi, 3 da 5 | 63,6 %                | 9,3 %           | 2,64       |
| attuale, 5 da 3    | 52,5 %                | 15,1 %          | 4,40       |
| set corti, 15 da 1 | 34,6 %                | 17,8 %          | 13,58      |
| variabile, tetto 7 | 1,7 %                 | 16,5 %          | 4,26       |

*(Divario 11 punti, ancoraggio 700, calcolo esatto.)*

La quinta misura non compare in quella tabella perché è l'unica che non si legge dal referto: si calcola dalla distribuzione degli esiti possibili, e vale la pena riportarla a parte, sul solo formato in vigore, perché dice una cosa che le altre non dicono.

| Divario fra i due arcieri | Si arriva all'ultimo set | Incertezza residua all'ultimo set |
|---------------------------|--------------------------|-----------------------------------|
| 6 punti                   | 59,7 %                   | 1,13                              |
| 11 punti                  | 52,5 %                   | 1,07                              |
| 17 punti                  | 42,4 %                   | 0,96                              |
| 25 punti                  | 29,5 %                   | 0,81                              |

*(Formato attuale a 5 set da 3, ancoraggio 700, calcolo esatto. L'incertezza è misurata in bit: vale 1,58 quando i tre esiti — vittoria, pareggio, sconfitta — sono equiprobabili, e scende a zero quando il match è già deciso.)*

L'incertezza residua va letta come una frazione del suo massimo, e il massimo è 1,58, il valore che si avrebbe se, arrivati all'ultimo set, i tre esiti fossero ugualmente possibili e non si sapesse assolutamente nulla.

Si legga allora la riga degli undici punti di divario, che è quella su cui è costruita tutta questa parte. Fra due arcieri così distanti il match arriva all'ultimo set una volta su due, e quando ci arriva l'incertezza residua è 1,07, cioè circa due terzi del massimo. Detto in modo che si veda: chi guarda quel set ha davanti tre esiti ancora quasi tutti aperti, e non sta assistendo a una formalità. È il motivo per cui questo formato funziona come spettacolo: non solo si arriva spesso in fondo, ma quando ci si arriva non è per scaramanzia.

I ribaltamenti di testa quasi raddoppiano fra i set lunghi e quelli corti, dal 9,3 al 17,8 per cento, mentre l'accuratezza fra gli stessi due sta in un terzo di punto. E la probabilità di arrivare all'ultimo set va dal 35 al 64. La scelta del formato è quindi interamente una scelta di spettacolo, perché sull'accuratezza i formati non si distinguono: chi sostiene un formato per ragioni di equità sta argomentando su un asse piatto.

I due estremi, per giunta, hanno effetti opposti sulle due misure. I set corti massimizzano i ribaltamenti e minimizzano la probabilità di arrivare all'ultimo set; i set lunghi fanno esattamente il contrario. Sono due idee diverse di che cosa renda un match interessante (l'incertezza continua contro il finale in bilico) ed è una scelta che si può fare, ma è quella e non un'altra.

C'è poi una conclusione che riordina il compromesso prezzato nella Parte prima: TOTALE+ non cambia le misure di spettacolo. Quattro su cinque sono proprietà del formato e non toccano la regola di parità, e la quinta cambia di un decimo di freccia su tredici.

La regola di parità non è dunque un asse dello spettacolo del formato. Lo spettacolo sta nell'arrivare all'ultimo set, che accade in un match su due, e in un ribaltamento di testa in un match su sette. Lo spareggio è un momento in più al termine di un match il cui spettacolo è già stato consumato altrove.

Va detto senza esagerare, perché l'obiezione è legittima: lo spareggio ha un valore che queste cinque misure non catturano, perché è breve, è visibile, ed è il momento che finisce nei montaggi. Il calcolo dice che non è ciò che tiene il pubblico dentro il match, non che sia irrilevante per come lo sport si racconta.

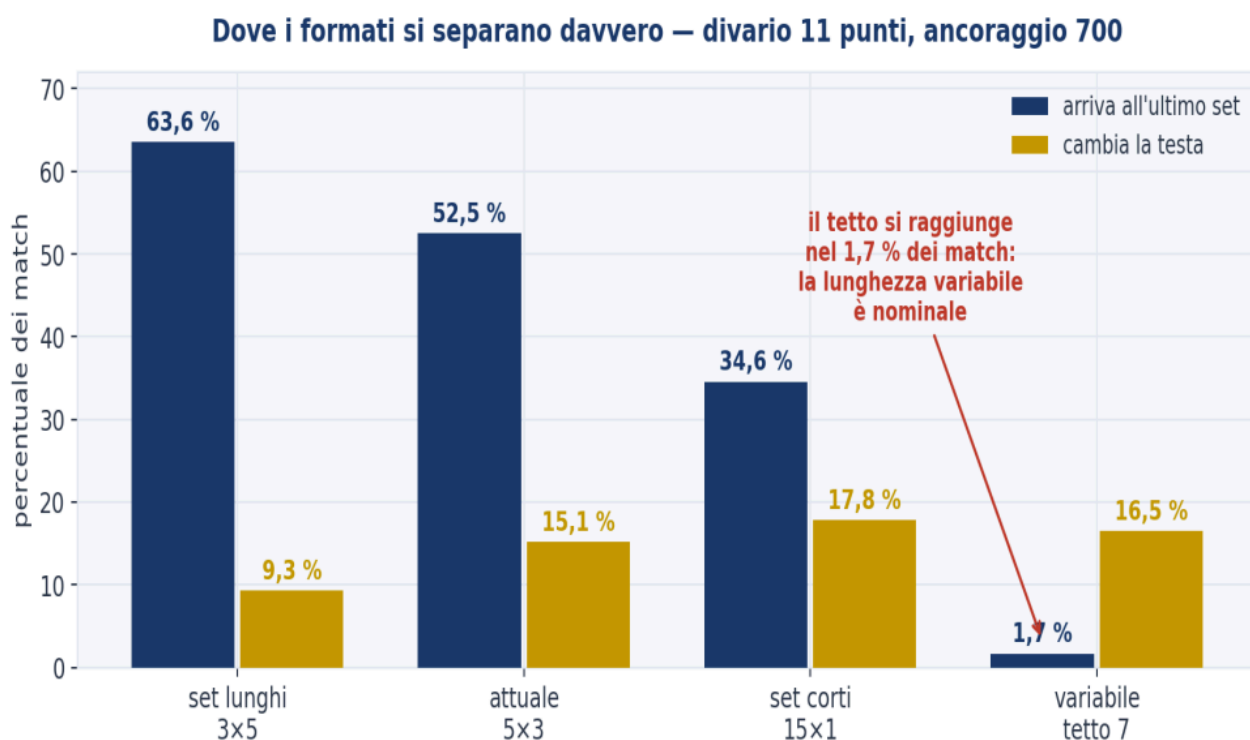


**Figura 13 — Dove i formati si separano davvero.** Probabilità che il match arrivi all'ultimo set e che cambi la testa, per quattro formati. Divario 11 punti, ancoraggio 700.

### 34. La regola che nessun formato può violare

Serve un cancello, e va messo prima di tutto il resto: un formato che cambia gli incentivi di tiro non misura più la stessa cosa.

Immaginiamo che sotto una certa regola convenga, in una certa situazione, tirare diversamente da come si tirerebbe per fare più punti possibile. Quel formato starebbe misurando un comportamento che esso stesso ha prodotto, e tutti i suoi altri numeri diventerebbero privi di senso.

Il rischio è concreto, non teorico. Un formato in cui conti il margine con cui si vince un set potrebbe rendere conveniente rischiare quando si è già avanti. Uno con un tetto potrebbe rendere conveniente trascinare il match.

La verifica si fa elencando tutti i casi, uno per uno, ed è possibile perché i modi in cui un match può andare sono in numero finito. Si elencano tutti e si controlla che migliorare un set non possa mai peggiorare la propria posizione. Per il formato attuale sono 243 sequenze; per 15 set da una freccia sono oltre 14 milioni, e si elencano lo stesso.

Il risultato è che tutti i formati e tutte le strutture passano, con zero violazioni. Il cancello quindi non discrimina, e va detto invece che presentato come un risultato.

Ma «zero violazioni» non significa niente finché non si è provato che il controllo sappia trovarle. Ho quindi costruito tre formati fatti apposta per violare, e li ho passati allo stesso controllo. Il primo dà un bonus a chi chiude il match in fretta, e il controllo trova 413 sequenze in cui conviene perdere un set per vincerne prima il seguente. Il secondo fa perdere chi totalizza esattamente sei punti-set, e ne trova 445. Il terzo (al tetto vince chi ha perso meno set) non viola, e lo riporto come controllo fallito: quel criterio non poteva violare per come è fatto, quindi un controllo rotto lo avrebbe passato ugualmente. Due controlli su tre fanno il loro lavoro; il terzo era mal costruito, e dirlo vale più che sostituirlo in silenzio.

La ragione per cui il cancello non discrimina è però interessante. Tutti i formati condividono la stessa proprietà: la soglia termina il match appena la si raggiunge, quindi fare più punti non può mai nuocere. Non è che siano ben progettati, è che condividono il tratto che rende la neutralità automatica.

L'unico caso in cui il cancello poteva scattare è il formato a lunghezza variabile, che ha due criteri di terminazione, il margine e il tetto. Con due criteri la neutralità non è più garantita dal primo. Passa, e la ragione è precisa: i due criteri sono concordi, cioè misurano la stessa cosa, quindi arrivare al tetto non premia nessuno diversamente da come lo premierebbe fermarsi prima. Un tetto è sicuro finché risolve con lo stesso criterio con cui si decide; un tetto che risolvesse sul totale delle frecce, o sul numero di set vinti invece che sui punti, introdurrebbe un secondo ordinamento in conflitto col primo e andrebbe verificato di nuovo.

---

## 35. Come si compongono i pezzi

Abbiamo quattro pezzi misurati separatamente: da dove viene il piazzamento, che struttura ha il tabellone, che formato ha il match, che regola risolve la parità. Non si possono sommare, e lo sappiamo per misura e non per precauzione: qualificazione doppia e seeding insieme danno meno della somma dei loro effetti, perché comprano in parte la stessa cosa. Un modello costruito prendendo la scelta migliore pezzo per pezzo darebbe un numero sbagliato.

Va quindi calcolato l'insieme. Ogni configurazione di gara è una quaterna (origine del piazzamento, struttura del tabellone, formato del match, regola di parità) e le combinazioni serie sono cinquantquattro: le ho calcolate tutte. Il conteggio è quello dello spazio su cui i quattro modelli sono stati ottimizzati, e non comprende la doppia eliminazione in senso stretto, che il capitolo 29 misura a parte e che nessuno dei quattro adotta.

Su quell'insieme ho però imposto un vincolo che conviene dichiarare adesso, perché decide che cosa i modelli possano contenere: ho ammesso soltanto configurazioni adottabili con le attrezzature che una gara internazionale ha già oggi. Il che esclude l'unica regola di parità che sarebbe più accurata di tutte (la somma dei quadrati delle distanze del capitolo 12) perché richiede di misurare dove ogni freccia sia finita dentro l'anello. Sistemi di rilevamento automatico capaci di trasmettere la posizione d'impatto

esistono e sono stati sperimentati in gara, ma la registrazione sistematica delle coordinate non fa parte del normale flusso di scoring, e non dispongo di un archivio di coordinate individuali su cui questo lavoro possa fare affidamento. Al capitolo 40 dico quanto costi quel vincolo, che è la domanda che un lettore attento pone a questo punto.

E siccome migliore dipende da che cosa si voglia, dall'insieme ammesso emergono tre configurazioni. Al capitolo 39 se ne aggiungerà una quarta, che non nasce da questa ottimizzazione ma da una domanda diversa (quanto si guadagni senza cambiare niente di visibile) e che per questo va letta accanto alle altre e non dentro la stessa classifica.

| Modello    | Massimizza                        | Vincolato a                                                                      |
|------------|-----------------------------------|----------------------------------------------------------------------------------|
| Precisione | probabilità che vinca il migliore | non costare più del doppio della gara attuale                                    |
| Spettacolo | numero di spareggi tirati         | non essere meno accurato della gara attuale                                      |
| Bilanciato | probabilità che vinca il migliore | conservare almeno metà degli spareggi e non costare più del 10 % di match in più |

*(I tre problemi di ottimizzazione, con i vincoli scritti prima di calcolarne le soluzioni: è la condizione perché le tre righe siano estratte e non scelte. Esiste una quarta riga, «la più accurata in assoluto senza alcun vincolo», che serve a misurare quanto si lasci sul tavolo e non è una proposta — e che, come vedremo, coincide con il modello precisione.)*

Il modello spettacolo non è quello sciatto: è vincolato anche lui, e il suo contributo è quantificare quanta accuratezza costi un punto di spettacolo. I tre capitoli che seguono descrivono le tre configurazioni per intero, dalla prima freccia all'ultima.

---

## 36. Il modello precisione

Se l'unico obiettivo fosse che vinca il più bravo, e potessimo cambiare tutto tranne la durata complessiva, la gara migliore sarebbe questa: piazzamento dalla graduatoria di stagione, tabellone con riseeding, match a 15 set da una freccia, parità risolta con TOTALE+.

Dà 14,2 volte su cento contro le 9,7 di oggi, cioè il titolo andrebbe al migliore una gara su sette invece di una su dieci. Su cento edizioni di un campionato, quattro e mezza finirebbero all'arciere giusto invece che a un altro.

Ecco come si svolgerebbe.

Non si tira nessun round di qualificazione. I 64 arcieri entrano in tabellone secondo la graduatoria di stagione, cioè la stima di precisione calcolata sui punteggi di qualificazione delle gare già disputate nei mesi precedenti, corretti per la difficoltà di ciascuna gara secondo la procedura del capitolo 46. Chi ha disputato meno del minimo richiesto entra in coda, ordinato per punteggio dell'ultima gara valida. Si elimina così la mezza giornata di qualificazione, e il tempo risparmiato è quello che paga tutto il resto.

Il primo turno accoppia il primo della graduatoria col sessantaquattresimo, il secondo col sessantatreesimo e così via: 32 match, come oggi.

Il match si gioca a 15 set da una freccia ciascuno. Ogni freccia è un set, chi la piazza meglio prende due punti, in caso di parità uno per uno, e vince chi arriva a sedici. Sembra strano a dirsi, ed è invece il formato più accurato dei cinque provati, perché moltiplica le occasioni in cui la differenza di bravura può manifestarsi invece di raggrupparle in cinque blocchi da tre. Se dopo 15 set il punteggio è quindici pari, vince chi ha totalizzato più punti nelle 15 frecce; se anche i totali sono pari, si tira una freccia con l'ordine sorteggiato.

Dal secondo turno in poi entra il riseeding, e qui sta la differenza più importante, quella che una federazione deve organizzare. Finito ogni turno, gli accoppiamenti del turno successivo non sono quelli fissati all'inizio: si riordinano i sopravvissuti secondo la graduatoria e si accoppia il primo con l'ultimo, il secondo con il penultimo, e così via.

Un esempio rende chiaro che cosa cambi. Se al primo turno escono a sorpresa il numero due e il numero tre, oggi il numero uno affronterebbe comunque l'avversario che il tabellone gli aveva assegnato mesi prima. Con il riseeding si trova davanti il sopravvissuto peggiore: il tabellone si adatta a quello che è successo, invece di restare fedele a una previsione smentita.

Quello che questo richiede in pratica è che gli accoppiamenti del turno successivo siano annunciati dopo la fine del turno precedente e non il giorno prima. Per la trasmissione è un vincolo reale, perché non si può stampare il tabellone completo in anticipo, ed è il prezzo organizzativo di questo modello.

|                              | Gara attuale     | Modello precisione |
|------------------------------|------------------|--------------------|
| Titolo al migliore           | 9,7 %            | 14,2 %             |
| Match giocati                | 63               | 63                 |
| Frecce di spareggio a torneo | 10,5             | 3,6                |
| Round di qualificazione      | 1                | 0                  |
| Accoppiamenti annunciati     | tutti all'inizio | turno per turno    |

*(Campo compatto da 64 arcieri, venticinquemila tornei della stessa corsa appaiata, con seme dichiarato nel file dei valori. La probabilità è quella che il titolo vada al migliore del campo; i match e gli spareggi sono su un torneo intero.)*

Si guadagnano quattro titoli e mezzo su cento, si risparmia mezza giornata di qualificazione, e si perdono due terzi degli spareggi. È il formato che si adotterebbe per una gara in cui il titolo conta più di tutto (un campionato del mondo, una selezione olimpica, una finale di circuito) cioè se qualcuno chiedesse conto del fatto che il migliore non vince quasi mai.

Va notato che non è più lungo della gara attuale, anzi è più corto: il vincolo sulla durata che avevo posto nell'ottimizzazione non è risultato vincolante. E c'è un modo di dirlo che vale più del vincolo stesso: questa configurazione coincide con la migliore delle cinquantaquattro combinazioni calcolate senza imporre alcun limite. Non esiste dunque una gara «costosa ma più precisa» che il vincolo abbia escluso: la frontiera si esaurisce prima, e la configurazione più accurata è già dentro il tempo che una gara occupa oggi.

---

## 37. Il modello spettacolo

Se l'obiettivo fosse il massimo di quello che tiene il pubblico dentro (spareggi, ribaltamenti, match in bilico fino all'ultimo) senza però peggiorare l'accuratezza rispetto a oggi, la gara sarebbe questa: qualifica-

zione come oggi, tabellone con ripescaggio dei quarti, formato attuale a 5 set da tre, e regola di parità in vigore. Il titolo va al migliore 10,8 volte su cento contro le 9,7 della gara attuale, e i match passano da 63 a 67.

Il vincolo è quello che rende il modello serio: non è la gara più spettacolare in assoluto, è la più spettacolare che non costi accuratezza.

Ecco come si svolgerebbe. Il round di qualificazione è a 72 frecce come oggi, senza alcun cambiamento. Dal primo turno ai quarti tutto è identico a oggi (tabellone standard, formato a 5 set da tre, parità risolta con la freccia) e questa è la parte importante, perché chi guarda non nota nessuna differenza fino ai quarti.

Dopo i quarti sta l'unica modifica, ed è quella che aggiunge spettacolo. I quattro arcieri eliminati ai quarti non tornano a casa: giocano fra loro due semifinali e una finale, tre match in tutto. Chi ne esce affronta poi in uno spareggio il peggio piazzato fra i quattro vincitori dei quarti, che è quindi l'unico a doversi difendere, mentre i tre meglio classificati restano dove sono. Quattro match in più, e il torneo passa da 63 a 67.

Perché proprio i quarti, e non prima? Perché è lì che l'eliminazione fa più danno. Un arciere forte che esce ai quarti è la perdita più costosa di tutto il tabellone (è arrivato fin lì, quindi è probabilmente forte, e il tabellone lo butta via per un match solo) mentre ripescare più indietro peggiora le cose, come il capitolo 30 ha mostrato.

|                              | Gara attuale | Modello spettacolo |
|------------------------------|--------------|--------------------|
| Titolo al migliore           | 9,7 %        | 10,8 %             |
| Match giocati                | 63           | 67                 |
| Frecce di spareggio a torneo | 10,5         | 11,2               |
| Match con seconda occasione  | 0            | 4                  |

*(Campo compatto da 64 arcieri, venticinquemila tornei della stessa corsa appaiata, con seme dichiarato nel file dei valori. I quattro match in più sono le due semifinali, la finale del ripescaggio e il match di rientro in semifinale.)*

Quattro match in più, e con essi sette decimi di freccia di spareggio in più per torneo, e una fase nuova che è di per sé un evento: quattro arcieri eliminati che si giocano il rientro, con una semifinale in palio. E il tutto senza perdere accuratezza, perché ne guadagna 1,1 punti — ripescare i quarti corregge l'errore più costoso invece di aggiungerne.

Va detto onestamente che un punto e uno sono poca cosa, e va letto nella sua unità: è la frequenza con cui il torneo finisce col migliore in cima, non il numero di match che cambiano esito. In un circuito da dieci gare l'anno significa un titolo in più ogni nove stagioni. Convertirlo in match moltiplicando per sessantatré sarebbe scorretto: la frequenza con cui vince il migliore e il numero di match il cui esito cambia sono due grandezze diverse.

Questo modello non va quindi proposto per l'accuratezza: va proposto per quello che aggiunge, e il fatto che non peggiori l'accuratezza è il suo requisito di ammissione, non il suo pregio. È il formato per una gara in cui il pubblico e la trasmissione contano più del titolo, una tappa di circuito, una finale in città, un evento pensato per essere visto. Ed è l'unica delle tre configurazioni che costi qualcosa: quattro match, cioè circa un'ora, spesi per comprare spettacolo e non accuratezza.

---

## 38. Il modello bilanciato

Se non volessimo togliere niente a nessuno (stessi match, stessi spareggi, stesso formato, stessa durata) quanto si potrebbe guadagnare? È la domanda che una federazione pone davvero, perché è l'unica che non richieda di convincere qualcuno a rinunciare a qualcosa.

La risposta è quattro titoli su cento. La gara sarebbe piazzamento dalla graduatoria di stagione, tabellone con riseeding, formato attuale a 5 set da tre, regola di parità in vigore: il titolo va al migliore 13,7 volte su cento contro le 9,7 di oggi, cioè poco meno di una gara su sette invece di una su dieci.

I 64 arcieri entrano in tabellone secondo la graduatoria di stagione, cioè la stessa stima di precisione descritta al capitolo 31, costruita sui punteggi di qualificazione delle gare precedenti e non sui piazzamenti ottenuti. Il round di qualificazione può restare, come momento di gara e per ordinare chi non abbia una graduatoria, ma non determina più il tabellone. Il match è identico a oggi, 5 set da 3 frecce, primo a sei. La parità si risolve con la freccia di spareggio, identica a oggi. E il tabellone usa il riseeding: dopo ogni turno si ricompongono gli accoppiamenti secondo la graduatoria, il meglio classificato contro il peggio rimasto.

|                              | Gara attuale | Modello bilanciato |
|------------------------------|--------------|--------------------|
| Titolo al migliore           | 9,7 %        | 13,7 %             |
| Match giocati                | 63           | 63                 |
| Frecce di spareggio a torneo | 10,5         | 10,4               |
| Formato del match            | 5 set da 3   | 5 set da 3         |
| Regola di parità             | freccia      | freccia            |
| Durata                       | una giornata | una giornata       |
| Frecce tirate                | come oggi    | come oggi          |

*(Campo compatto da 64 arcieri, venticinquemila tornei della stessa corsa appaiata, con seme dichiarato nel file dei valori. La graduatoria che compone il tabellone è quella di 10 gare, prodotta dalla procedura della Parte terza.)*

Guadagna quattro titoli su cento tenendo tutto identico: gli stessi 63 match, lo stesso formato, la stessa regola di parità, la stessa durata. Le frecce di spareggio restano circa quelle di oggi, 10,4 contro 10,5, e la differenza non viene dalla regola, che è la stessa, ma dal fatto che con un tabellone diverso si incontrano coppie diverse. Cambia soltanto da dove viene il piazzamento e come si accoppiano i sopravvissuti, due righe di regolamento, zero frecce, zero minuti.

Il perché funzioni merita di essere spiegato, perché il totale non è la somma delle parti. I due interventi agiscono sulla stessa cosa, l'errore del piazzamento, ma da lati opposti. La graduatoria di stagione riduce l'errore, perché sa chi è più bravo molto meglio di quanto lo sappia un pomeriggio di qualificazione. Il riseeding converte quella conoscenza in accoppiamenti, e senza di esso un piazzamento migliore non serve quasi a niente, come il capitolo 31 ha mostrato.

I numeri lo dicono con precisione. Applicati insieme, piazzamento dalla graduatoria e riseeding portano la gara di oggi dal 9,7 al 13,7 per cento: quattro punti, ed è la misura che conta perché viene dalla stessa corsa di tutte le altre configurazioni. Un esperimento dedicato, che serve a scomporre quel guadagno e

non a sostituirne il livello, dice che i due interventi presi da soli valgono mezzo punto e uno e uno: sommati farebbero poco più di un punto e mezzo, mentre insieme ne fanno più del doppio.

È l'unica volta in tutto questo lavoro in cui due interventi si moltiplicano invece di sommarsi, e la ragione è che nessuno dei due serve senza l'altro, una buona informazione che nessuno usa, o un meccanismo che usa un'informazione cattiva.

Le due modifiche si scrivono così.

### **Articolo — Composizione del tabellone**

1. Il tabellone di eliminazione è composto secondo la graduatoria stagionale in vigore alla data della gara, calcolata sui punteggi dei round di qualificazione delle gare del calendario ufficiale secondo la procedura di cui all'allegato tecnico. Gli atleti privi di graduatoria valida sono collocati in coda, ordinati secondo il punteggio del round di qualificazione del giorno.

2. Al termine di ciascun turno, gli accoppiamenti del turno successivo sono ricomposti secondo la graduatoria: l'atleta con la posizione migliore fra i qualificati affronta quello con la posizione peggiore, il secondo affronta il penultimo, e così via.

3. Gli accoppiamenti di ciascun turno sono resi noti entro trenta minuti dal termine del turno precedente.

Quello che richiede in pratica è poco ma va detto. Serve una graduatoria stagionale pubblicata e aggiornata, che è il prodotto della Parte terza e che una federazione deve costruire prima di poter adottare questo modello. Serve la disponibilità ad annunciare gli accoppiamenti turno per turno invece che il giorno prima, che è un vincolo per la trasmissione, anche se trenta minuti fra un turno e l'altro sono già previsti dal programma di gara. E serve una regola per chi non abbia graduatoria (un esordiente, un atleta rientrato da un infortunio, un invitato) che il primo comma risolve nel modo più semplice, collocandoli in coda ordinati per punteggio del giorno.

È l'unica delle tre proposte che non richieda di togliere niente a nessuno: non un match, non uno spareggio, non un minuto, non una freccia. Per questo è quella che consegnerei per prima.

Resta un limite che va attaccato al risultato e non messo in nota: questo modello poggia interamente sul piazzamento da graduatoria. Senza quello, cioè componendo il tabellone con la qualificazione del giorno, il guadagno scende da quattro titoli su cento a uno. Sono due raccomandazioni diverse, e la prima vale soltanto se una federazione è disposta a usare la classifica di stagione per comporre il tabellone.

---

## **39. Il modello minimo: migliorare senza cambiare niente di visibile**

I tre modelli che precedono chiedono tutti di cambiare qualcosa che il pubblico vede: gli accoppiamenti annunciati turno per turno, un ripescaggio dopo i quarti, il piazzamento che non viene più dal punteggio del giorno. Sono cambiamenti modesti, ma vanno spiegati a qualcuno, e in una federazione spiegare costa.

C'è quindi un quarto modello che merita di stare accanto agli altri tre, e risponde alla domanda più pratica di tutte: quanto si può guadagnare senza toccare nulla di ciò che si vede? Stessa qualificazione a 72 frecce, stesso tabellone a 64, stessi cinque set da tre, stessa durata, stesso ordine del giorno, cambiando soltanto come si leggono i risultati che la gara produce già.

Lo chiamo modello minimo, e non è il più accurato dei quattro: è quello che chiede meno.

Ci sono tre interventi possibili, e conviene separarli perché hanno prezzi molto diversi, e perché due di essi, le due regole di parità, sono alternativi fra loro e non possono valere sullo stesso match.

Il primo non costa niente ed è la regola di parità. Sostituire la freccia di spareggio con TOTALE+, cioè far vincere chi ha totalizzato di più nelle 15 frecce e tirare solo se anche i totali sono pari. È la proposta del capitolo 18, e non richiede nulla che non sia già sul tavolo del giudice.

Il secondo richiede strumentazione ma è invisibile: ordinare la qualificazione per distanze. Oggi la classifica di qualificazione si fa sul totale ad anelli, e i pareggi si rompono col conteggio degli X e poi dei dieci. Se si registrasse dove ogni freccia è finita dentro l'anello, la stessa gara, stesse 72 frecce, stessa mezza giornata, produrrebbe una classifica più fedele. Nulla cambia per chi guarda: gli arcieri tirano le stesse frecce e il tabellone si compone allo stesso modo, soltanto l'ordine è più giusto.

| Come si ordina la qualificazione             | Il primo è il migliore | I primi otto | Errore medio | Accordo con l'ordine vero |
|----------------------------------------------|------------------------|--------------|--------------|---------------------------|
| Totale, pareggi a sorteggio                  | 14,0 %                 | 0,04 %       | 8,62 posti   | 0,819                     |
| Totale, poi X                                | 14,2 %                 | 0,04 %       | 8,60         | 0,820                     |
| Totale, poi X, poi dieci — la regola di oggi | 14,2 %                 | 0,04 %       | 8,61         | 0,819                     |
| Somma dei quadrati delle distanze            | 15,9 %                 | 0,09 %       | 7,70         | 0,855                     |

*(64 arcieri fra 690 e 665, 72 frecce, centoventimila qualificazioni simulate. I pareggi sono rotti a sorteggio e non per ordine di classifica, altrimenti il confronto favorirebbe artificialmente i criteri che pareggiano di più. Errore standard sul primo classificato: da  $\pm 0,10$  a  $\pm 0,11$  punti.)*

Due cose vanno lette in quella tabella. Le distanze riducono l'errore medio di posizione di quasi un posto intero, da 8,61 a 7,70, e portano il primo classificato da 14,2 a 15,9 volte su cento, una differenza dodici volte più grande dell'errore della misura, cioè che il caso non spiega in alcun modo. E individuano esattamente gli otto migliori più del doppio delle volte rispetto ai criteri ad anelli, che fra loro sono indistinguibili.

I criteri di spareggio che il regolamento usa oggi, invece, non fanno quasi nulla: aggiungere il conteggio degli X vale due decimi, aggiungere anche quello dei dieci non vale niente. Qui non fanno danno come nel match, perché in classifica non c'è la parità di totale che li rovescia; semplicemente non aggiungono informazione.

**Il terzo intervento** è la regola di parità letta sulle distanze, cioè far vincere il match a chi ha la somma dei quadrati più piccola sulle 15 frecce. Richiede la stessa strumentazione del secondo.

Ecco che cosa producono, presi uno per volta e tutti insieme, su una gara che per il resto è identica a oggi.

| Ordine della qualificazione | Regola di parità     | Titolo al migliore | Guadagno          |
|-----------------------------|----------------------|--------------------|-------------------|
| anelli, X, dieci            | freccia di spareggio | 9,7 %              | — la gara di oggi |
| anelli, X, dieci            | TOTALE+              | 10,0 %             | +0,3              |
| distanze                    | freccia di spareggio | 10,3 %             | +0,5              |
| anelli, X, dieci            | distanze             | 10,4 %             | +0,7              |
| distanze                    | TOTALE+              | 10,6 %             | +0,8              |
| distanze                    | distanze             | 10,8 %             | +1,1              |

*(Tabellone standard a 64 arcieri, campo compatto fra 690 e 665, venticinquemila tornei della stessa corsa appaiata, con seme dichiarato nel file dei valori, con le matrici dei match in forma chiusa. Formato, durata e struttura identici in tutte le righe; l'unica cosa che cambia fra una riga e l'altra è come si leggono i risultati.)*

Tutte le righe di questa tavola e i cinque modelli del capitolo successivo provengono dalla stessa corsa appaiata, quindi i loro livelli sono direttamente confrontabili. La colonna da guardare è l'ultima: i guadagni sono confronti a parità di tutto il resto — stessi campi, stesse qualificazioni, stessi esiti dei match, cambiata soltanto la lettura — e su un confronto così l'errore è molto più piccolo che sui livelli.

I numeri da tenere sono quattro.

TOTALE+ da solo vale 0,33 punti e non costa nulla. È l'unico intervento di tutto questo volume che si adotti con una riga di regolamento e zero investimento.

Ordinare la qualificazione per distanze vale mezzo punto, e questo corregge quello che avevo scritto nella prima stesura. Il contrasto misurato sulla corsa decisionale vale +0,5 punti con un errore di 0,3: il segno è positivo e il valore supera l'errore di poco più di un fattore due, quindi l'effetto c'è ma non è misurato con precisione. La prima edizione riportava qui che l'ordinamento per distanze non vale niente, e quel numero veniva da un contrasto misurato su una corsa diversa: la tavola qui sopra e il produttore dicono un'altra cosa, e valgono loro.

Va detto che cosa migliora e che cosa no, perché le due domande sono diverse. Ordinare per distanze migliora il piazzamento in modo netto: mette primo il migliore 15,9 volte su cento invece di 14,2, e riduce l'errore medio da 8,6 posti a 7,7. Sul titolo l'effetto è molto più piccolo, mezzo punto, e la ragione merita di essere capita: il tabellone non consulta la classifica per sapere chi sia il migliore, la consulta per decidere chi incontra chi. Quello che conta è come distribuisce gli avversari lungo tutti e sessantatré i match, non quanto azzecca le prime posizioni. Un piazzamento molto più accurato in testa produce un guadagno modesto sui titoli.

Leggere le distanze nella parità vale 0,7 punti, contro i 0,33 di TOTALE+, perché agisce dove l'informazione è più scarsa. Il margine sulla soluzione gratuita è però stretto: se una federazione dovesse comprare strumentazione per una sola ragione sarebbe questa, ma il confronto onesto è sette decimi contro tre, in cambio di un investimento che TOTALE+ non richiede.

E i due che possono coesistere (ordinare la qualificazione per distanze e risolvere le parità sulla stessa misura) valgono insieme 1,1 punti, portando la gara di oggi da 9,7 a 10,8 volte su cento.

C'è anche una via di mezzo, per chi compri la strumentazione e non voglia toccare la regola di parità più di tanto: ordinare la qualificazione per distanze e risolvere le parità con TOTALE+ vale 0,8 punti. È

meno degli 1,1 della coppia completa e più dei 0,7 della sola parità sulle distanze, e ha il pregio di lasciare la regola di parità in una forma che si verifica a occhio. Meno della somma dei due presi da soli, 0,5 più 0,7, perché in parte comprano la stessa cosa.

Un punto scarso è poco, e va detto. Su cento tornei significa un titolo in più assegnato all'arciere giusto: in un circuito da dieci gare l'anno, una decina di stagioni per un titolo. Ma il confronto giusto non è con lo zero, è con gli altri interventi possibili — il modello bilanciato del capitolo 38 ne vale 4,0, cioè quattro volte tanto, e non richiede né strumentazione né investimenti.

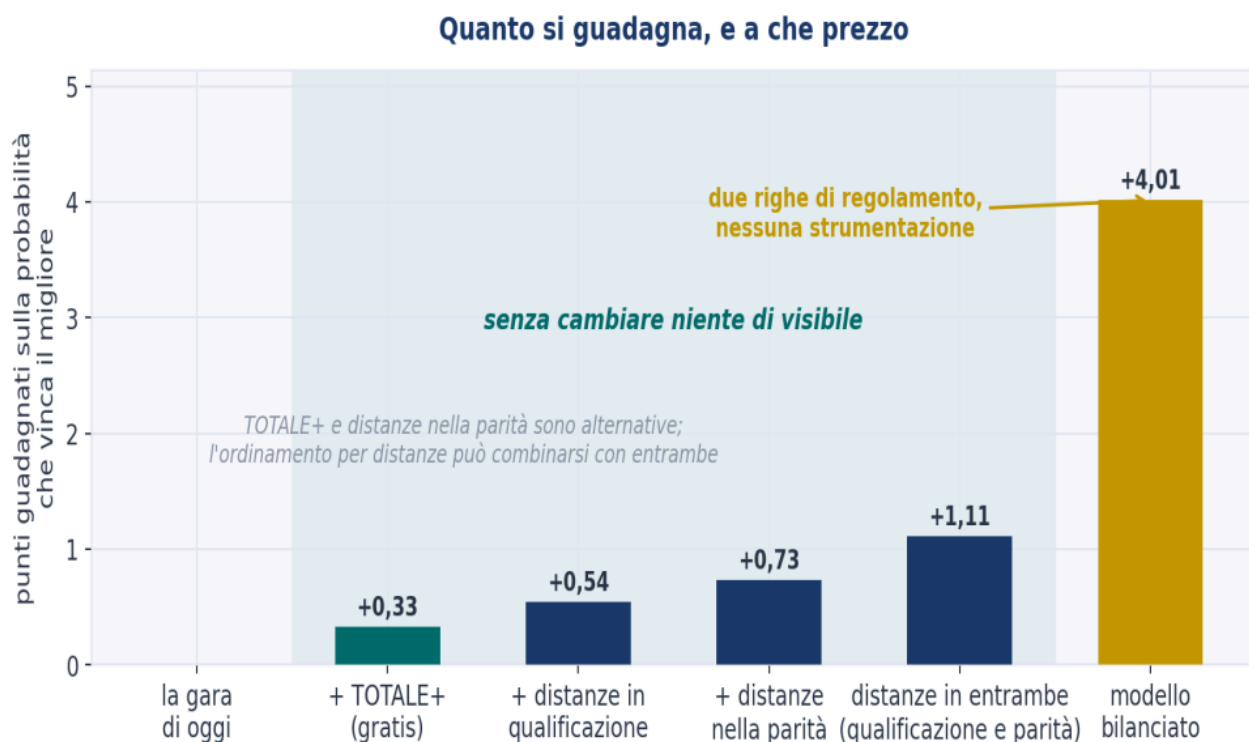


**Figura 14 — Quanto si guadagna, e a che prezzo.** Punti guadagnati sulla probabilità che il titolo vada al migliore, rispetto alla gara di oggi. Campo compatto da 64 arcieri, tutte le colonne dalla stessa corsa: venticinquemila tornei della stessa corsa appaiata, con seme dichiarato nel file dei valori. Le prime quattro non cambiano niente di ciò che si vede; l'ultima richiede due righe di regolamento e nessuna strumentazione. Gli interventi a formato invariato non si sommano al modello bilanciato: sono strade alternative. E fra loro, le due regole di parità — TOTALE+ e le distanze — sono a loro volta alternative: il modello minimo usa le distanze.

Chi non voglia cambiare niente di visibile ma possa registrare le distanze guadagna 1,1 punti facendo entrambi gli interventi. Chi non disponga della strumentazione può scegliere TOTALE+, che ne vale 0,33 e non costa nulla. Chi accetti di cambiare due righe di regolamento ne guadagna 4,0. La differenza fra le due strade non è tecnica: è quanto si è disposti a spiegare.

## 40. I quattro a confronto

Le quattro gare descritte nei capitoli precedenti nascono dalla stessa macchina e dallo stesso insieme di combinazioni: quello che le distingue è la domanda a cui rispondono. Messe una accanto all'altra, con la

gara attuale come termine di paragone, si legge in un colpo che cosa costi ciascuna e che cosa dia.

|                              | Gara attuale | Minimo        | Spettacolo    | Bilanciato    | Precisione    |
|------------------------------|--------------|---------------|---------------|---------------|---------------|
| Titolo al migliore           | 9,7 %        | 10,8 % (+1,1) | 10,8 % (+1,1) | 13,7 % (+4,0) | 14,2 % (+4,5) |
| Match giocati                | 63           | 63            | 67            | 63            | 63            |
| Frecce di spareggio a torneo | 10,5         | 0,0           | 11,2          | 10,4          | 3,6           |
| Qualificazione               | 72 frecce    | 72 frecce     | 72 frecce     | 72 frecce*    | nessuna       |
| Formato del match            | 5 set da 3   | 5 set da 3    | 5 set da 3    | 5 set da 3    | 15 set da 1   |
| Regola di parità             | freccia      | distanze      | freccia       | freccia       | TOTALE+       |
| Tabellone                    | standard     | standard      | ripescaggio   | riseeding     | riseeding     |
| Serve strumentazione         | no           | sì            | no            | no            | no            |

*(Sulla probabilità che il titolo vada al migliore, minimo e spettacolo danno lo stesso numero: due centesimi di punto li separano, contro un errore di due decimi, e questa griglia non li distingue. Si scelgono per che cosa conservano — il minimo non cambia niente di visibile, lo spettacolo tiene gli spareggi e aggiunge match — non per quanto valgono. Campo compatto da 64 arcieri. La riga delle frecce di spareggio conta quante volte in un torneo si arrivi davvero a tirarne una, non quanti match potrebbero richiederla. \\*Nel modello bilanciato la qualificazione può restare come momento di gara ma non determina il tabellone. Tutte e cinque le configurazioni vengono dalla stessa corsa appaiata.)*

Come si sceglie fra i quattro? Non con un calcolo, perché la scelta dipende da che cosa una gara debba fare, e quella è una decisione che non tocca a chi fa i conti. Ma quattro cose il calcolo le dice.

Il modello minimo completo vale 1,1 punti, cioè un titolo su cento, e richiede la registrazione delle distanze sia per ordinare la qualificazione sia per risolvere la parità: il secondo intervento vale 0,7 e il primo 0,5, e insieme meno della somma. Per chi non disponga della strumentazione esiste una variante ancora più semplice, TOTALE+, che ne vale 0,33 e richiede soltanto una modifica regolamentare. Va scelto se cambiare due righe di regolamento è più difficile che procurarsi degli strumenti, il che in alcune federazioni è vero.

Il modello precisione è il più accurato e costa meno di oggi, perché elimina la qualificazione; il suo prezzo è pagato in spareggi, che scendono da dieci e mezzo a tre e sei per torneo, e in flessibilità organizzativa.

Il modello spettacolo è il solo che costi qualcosa, quattro match, circa un'ora, e il suo guadagno di accuratezza è irrilevante, quindi va scelto per quello che aggiunge e non per quello che migliora.

E il modello bilanciato guadagna quasi quanto il precisione senza togliere niente: 4,0 punti contro 4,4, cioè l'89 per cento del guadagno massimo a costo zero.

Se dovessi consegnarne uno solo, consegnerei il bilanciato, non perché sia il migliore, ma perché è l'unico che nessuno abbia motivo di rifiutare.

## Il modello che non si può ancora costruire

A questo punto va posta la domanda più ovvia di tutte. La Parte prima ha stabilito che leggere le distanze delle 15 frecce è il modo più accurato di risolvere una parità. Perché allora nessuno dei quattro modelli lo usa?

La risposta è nel vincolo dichiarato al capitolo 35: ho ammesso soltanto configurazioni adottabili con le attrezzature esistenti. Quel vincolo però ha un prezzo, e il prezzo si misura.

| Regola di parità, a parità di tutto il resto | Titolo al migliore |
|----------------------------------------------|--------------------|
| Sorteggio                                    | 9,6 %              |
| Freccia di spareggio, la regola in vigore    | 10,0 %             |
| TOTALE+                                      | 10,4 %             |
| Somma dei quadrati delle distanze            | 10,8 %             |

*(Tabellone standard, campo compatto da 64 arcieri fra 690 e 665, 40 000 tornei della stessa corsa, seme dichiarato. Cambia soltanto la regola di parità; tutto il resto — piazzamento, tabellone, alea dei match — è tenuto fermo, così che le quattro righe siano confrontabili fra loro. È una corsa propria di questo confronto: il suo livello per la regola in vigore non va letto come la gara attuale della tavola dei cinque modelli, che viene dalla corsa unica e vale 9,7 per cento.)*

Dentro questa tavola, passare dalla freccia a TOTALE+ vale quattro decimi sul titolo, e passare da TOTALE+ alle distanze ne vale altrettanti. Sono differenze interne a questa corsa: il guadagno di TOTALE+ misurato nella corsa unica dei modelli, che è il valore che le raccomandazioni citano, vale 0,33 punti, ed è il singolo intervento sulla regola di parità che valga di più fra quelli che non richiedono strumentazione nuova; leggere le distanze ne vale 0,7, e la richiede.

Applicato al modello precisione, che con TOTALE+ arriva al 14,2 per cento, sostituire la regola con le distanze lo porta al 14,6: tre decimi e mezzo in più, con un errore di circa due decimi e mezzo su ciascuno dei due livelli. Vale la pena dire perché, perché è istruttivo. Nel formato a quindici set da una freccia il match arriva alla parità molto più di rado, e quando ci arriva i due hanno tirato quindici frecce divise in quindici set, che è una condizione più selettiva del cinque pari. La regola sulle distanze viene invocata meno spesso e decide meno bene, e il guadagno che vale otto decimi sul tabellone standard qui non si trasferisce. Estrapolarlo avrebbe promesso quindici volte quello che c'è.

Non cambia l'ordine dei quattro modelli e non cambia la raccomandazione: il bilanciato resta quello adottabile senza convincere nessuno. Ma dice una cosa che vale la pena tenere: il pezzo di gara su cui si può ancora guadagnare non è il formato né il tabellone, è la strumentazione.

Vale la pena essere precisi su quanto sia grande quel qualcosa, perché è facile leggerlo per più di quel che è. Mezzo punto è mezzo titolo ogni cento tornei, e va letto in quell'unità: non si converte in match cambiati moltiplicando per sessantatré, perché la frequenza con cui il torneo finisce col migliore in cima e il numero di match il cui esito cambia sono due grandezze diverse. È il tipo di guadagno che giustifica un investimento se quell'investimento serve anche ad altro, e la lettura elettronica del punto d'impatto servirebbe a molto altro, a partire dal contenzioso sui giudizi al bersaglio, mentre non lo giustifica da solo.

E c'è una ragione più forte per volerla, che non passa dall'accuratezza. Con le coordinate registrate, la parte di queste pagine che descrive come tirano gli arcieri smetterebbe di essere un'assunzione e diventerebbe una misura. I tornei e i regolamenti controfattuali resterebbero simulazioni, perché una gara che non si è disputata non si osserva: quello che cambierebbe è la distribuzione da cui la simulazione parte. Il caso più netto è la forma della coda: al capitolo 21 resta il numero aperto da cui dipende se TOTALE+ sia la scelta giusta o se serva una polizza, e oggi non è stimabile. Con le distanze registrate si stimerebbe in una stagione.

### I modelli di gara, contro quella attuale — campo compatto

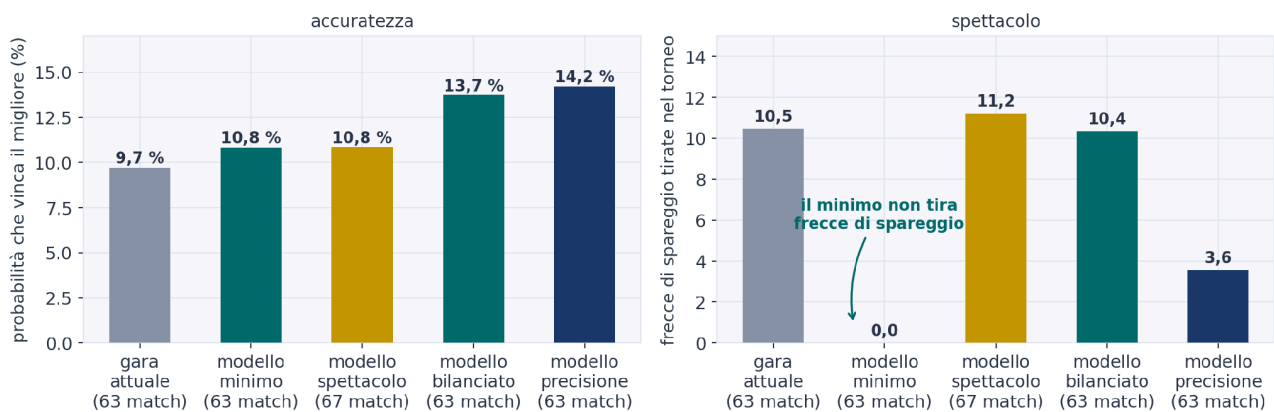


**Figura 15 — I modelli di gara.** Probabilità che vinca il migliore e frecce di spareggio tirate nel torneo, per i tre modelli e per la gara attuale. I match giocati non sono gli stessi in tutti: il ripescaggio dei quarti ne aggiunge quattro, e la figura li riporta sotto ciascun modello. Campo compatto da 64 arcieri, venticinquemila tornei della corsa unica dei modelli, con seme dichiarato nel file dei valori.

## 41. Che cosa portarsi a casa dalla Parte seconda

**Il piazzamento è molto peggiore di quanto si creda.** In un campo compatto la qualificazione a 72 frecce non indovina praticamente mai gli otto migliori, quattro volte su diecimila, e sbaglia in media di 8,6 posti.

**Ogni struttura che consulta il piazzamento ne eredita l'errore in proporzione a quante volte lo consulta.** Chi lo consulta una volta sola, cioè il tabellone standard, paga sette decimi di punto. Chi lo consulta due volte, come il ripescaggio, ne paga uno e sei. Chi lo consulta a ogni turno, cioè il ri-seeding, ne paga quattro e otto.

**La doppia eliminazione a finale secca costa il doppio dei match e rende meno del riseeding,** e la dose migliore del ripescaggio non è né zero né tutto: dagli ottavi in poi, cioè i quindici arcieri eliminati dagli ottavi alla finale, che vale 11,8 contro 11,1 dei soli quarti.

**Usare la classifica di stagione come piazzamento batte il raddoppiare la qualificazione,** e non costa niente perché quei dati esistono già.

**Il formato del match non conta.** I tre formati a frecce costanti stanno in un terzo di match a torneo, i due a lunghezza variabile fanno peggio di quasi un match, e la scelta è interamente di spettacolo.

# PARTE TERZA — L'ARCIERE DELL'ANNO

## 42. Il problema opposto

Le prime due parti hanno la stessa forma: pochissimi dati, e tetti che dicono quanto poco si possa ottenere. Quindici frecce per decidere un match, una freccia per risolvere una parità, settantadue per ordi-

nare sessantaquattro arcieri. Ogni volta la conclusione è che il margine esiste ed è stretto.

Questa parte è il problema opposto. Una stagione intera, dieci gare, settecentoventi frecce per atleta: dati in abbondanza, e la domanda diventa che cosa se ne possa ricavare. È lo stesso modello nel regime opposto, e la conclusione si rovescia: non è che l'abilità sia indistinguibile, è che una singola gara è uno strumento troppo corto.

Vale la pena quantificare subito il rovesciamento, perché è il numero che tiene insieme il volume. Nei due scenari modellati qui (un torneo da 64 arcieri, una stagione di 32) il torneo indica il migliore del campo dieci volte su cento e la stagione tre volte su quattro. Le due cifre non vengono da un confronto controllato: le popolazioni sono diverse, e il rapporto va letto come la distanza fra due scenari e non come l'efficienza relativa di due sistemi.

E questa parte non serve soltanto a incoronare qualcuno a dicembre. La Parte seconda ha mostrato che la graduatoria di stagione può essere il piazzamento su cui si costruisce un tabellone, e che in quel ruolo vale più di qualunque altro intervento. Lo strumento che segue serve quindi due volte: per il premio di fine anno e per decidere chi incontra chi a maggio.

---

## **43. Che cosa vuol dire il migliore dell'anno**

Prima della procedura, la domanda che la procedura deve risolvere. Ci sono almeno quattro risposte possibili, tutte difendibili.

La prima è chi ha vinto di più, che è il criterio dei circuiti a punti: si assegnano punti per piazzamento e si sommano. Ha il pregio di premiare chi vince quando conta, e il difetto di ereditare la casualità del tabellone, che la Parte seconda ha misurato: il titolo va al migliore una gara su dieci. Il piazzamento è quindi un segnale molto rumoroso per singola gara, non ne segue che la classifica erediti nove decimi di rumore, perché sommare dieci gare ne recupera una parte, e infatti a calendari completi la classifica a punti arriva al 72,7 per cento.

La seconda è chi ha fatto i punteggi più alti, cioè la media dei punteggi di qualificazione. Misura la precisione invece dei risultati ed è molto più stabile, ma confronta punteggi tirati in condizioni diverse, e il capitolo 45 mostra quanto costi.

La terza è chi è stato più costante, che premia chi non ha giornate storte. È una qualità reale, ma è una cosa diversa dalla precisione e va misurata a parte invece che confusa con essa.

La quarta è chi ha tirato meglio quest'anno, cioè chi ha avuto il livello medio più alto lungo la stagione. È questa che adotto, coerentemente con la definizione della Parte prima: l'arciere dell'anno è quello la cui precisione, mediata su tutta la stagione, risulta la più alta. Non chi ha vinto di più, non chi ha avuto il picco migliore; chi, tirando, si è aspettato in media il punteggio più alto.

Vale la pena distinguerla da una quinta che le somiglia e non è la stessa: chi tirerebbe meglio domani, cioè chi è più forte adesso. Sono due domande diverse, e il premio annuale è retrospettivo, si assegna per quello che si è fatto nella stagione, non per una previsione. C'è anche una ragione più stringente per non adottare la quinta qui, e la dico perché è un limite e non una preferenza: il modello di queste pagine tiene il livello fermo per tutta la stagione, e dentro quel mondo pesare di più le gare recenti può soltanto

perdere informazione. Qualunque cosa dicessi sullo stato corrente sarebbe una conseguenza dell'assunzione e non una misura. Al capitolo 57 misuro quanto quell'assunzione regga.

La ragione della scelta è la stessa del capitolo 2: è l'unica delle quattro verificabile. Si può dire con un numero quanto una procedura la indovini, perché si può simulare un mondo in cui la risposta è nota e contare quante volte la procedura ci arrivi. Con chi ha vinto di più non si può, perché la risposta è per definizione quella che si osserva.

E c'è una ragione pratica che vale il doppio: se la graduatoria deve servire da piazzamento, deve prevedere chi tirerà meglio e non registrare chi ha vinto, perché un piazzamento è una previsione e va costruito come tale.

## 44. Quanto una stagione riesce a distinguere

Prima di costruire lo strumento, il limite entro cui dovrà stare. La tabella riporta due colonne per casella: che cosa si otterrebbe leggendo le coordinate esatte, e che cosa si ottiene leggendo gli anelli, che è quello che una classifica reale ha a disposizione.

| Divario | 1 gara          | 5 gare          | 10 gare         | 20 gare         | 100 gare        |
|---------|-----------------|-----------------|-----------------|-----------------|-----------------|
|         | anelli / coord. | anelli / coord. | anelli / coord. | anelli / coord. | anelli / coord. |
| 1 punto | 57,1 / 58,5     | 65,4 / 68,5     | 71,5 / 75,2     | 78,9 / 83,2     | 96,4 / 98,4     |
| 2 punti | 64,0 / 66,5     | 78,7 / 83,0     | 87,0 / 91,1     | 94,4 / 97,2     | 100 / 100       |
| 3 punti | 70,1 / 73,6     | 88,1 / 92,1     | 95,3 / 97,7     | 99,1 / 99,8     | 100 / 100       |
| 5 punti | 80,7 / 84,9     | 97,4 / 99,0     | 99,7 / 99,9     | 100 / 100       | 100 / 100       |

*(Quante volte su cento si indica correttamente il più bravo fra due. Una gara è un round da 72 frecce. Ancoraggio 700. È la formula  $S_5$  valutata a  $n$  grande, cioè la  $S_{15}$ , ed è un limite superiore: modello noto e lettura ottimale.)*

E al contrario, quante frecce servano per arrivare a sbagliare non più di una volta su venti.

| Divario | Leggendo gli anelli    | Leggendo le coordinate | Costo del referto |
|---------|------------------------|------------------------|-------------------|
| 1 punto | 6 052 frecce (85 gare) | 4 202 (59 gare)        | 1,44 volte        |
| 2 punti | 1 542 (22 gare)        | 1 073 (15)             | 1,44              |
| 3 punti | 698 (10 gare)          | 487 (7)                | 1,43              |
| 5 punti | 261 (4 gare)           | 183 (3)                | 1,43              |

*(Minimo numero intero di frecce per cui l'affidabilità raggiunge il 95 per cento; le gare sono arrotondate per eccesso. Ancoraggio 700.)*

La riga da tenere è la terza. Una stagione di 10 gare distingue due arcieri che distino tre punti su 720 con un'affidabilità del 95,3 per cento, che è la soglia adottata qui e non una certezza: su venti coppie così ne sbaglia una: dieci gare bastano, e sono esattamente quelle che un calendario nazionale offre.

Ma la prima riga dice il limite, ed è netto. Per distinguere due arcieri che distino un punto solo servirebbero 85 gare, cioè otto stagioni e mezza a calendario pieno. Un arciere dovrebbe tirare da ventenne a ventinove anni perché la classifica sappia dire, con l'affidabilità del 95 per cento adottata qui, se è più

preciso di un suo coetaneo di un punto su settecentoventi. Non è un'impossibilità matematica: è un calendario che nessuna federazione può tirare.

Questo numero è il vincolo entro cui tutto il resto della Parte terza deve stare, e va tenuto a mente quando arriveremo alle regole di ammissione: non serve una procedura più raffinata per distinguere un punto, perché il limite non è nella procedura.

Vale la pena notare, di passaggio, che leggere gli anelli invece delle coordinate costa il 44 per cento di frecce in più, e che il rapporto è lo stesso su tutti e quattro i divari: 1,44, 1,44, 1,43, 1,43. Non è un effetto di soglia, è un tasso, e la formula S16 mostra perché sia costante. Ricompare qui identico a come compariva nella Parte prima.

---

## 45. Perché un punteggio non si confronta con un altro

Ed ecco il problema che rende non banale una classifica di stagione, e che nessuna somma di punteggi risolve.

Un 685 tirato con vento e un 685 tirato in giornata perfetta non sono la stessa prestazione. Ogni gara ha le sue condizioni, vento, luce, altitudine, temperatura, qualità del campo, e quelle condizioni agiscono su tutti i presenti insieme (formula S17).

Ne segue che due arcieri che abbiano tirato a gare diverse non sono confrontabili sui punteggi grezzi. E siccome nessuno partecipa a tutto, questa non è un'eccezione: è la regola.

| Sovrapposizione dei calendari | Nessun effetto | Effetto da 5 punti | Effetto da 11 punti |
|-------------------------------|----------------|--------------------|---------------------|
| 0 % — gare tutte diverse      | 95,3           | 84,4               | 71,7                |
| 50 %                          | 94,8           | 88,9               | 76,9                |
| 100 % — stesse gare           | 95,1           | 94,8               | 94,1                |

*(Classifica costruita sulla media dei punteggi. Divario 3 punti, 10 gare da 72 frecce, ancoraggio 700, 4 000 stagioni per cella.)*

Si legga l'ultima colonna dall'alto in basso. Con condizioni che spostano i punteggi di undici punti, come una giornata di vento forte, due arcieri che abbiano tirato alle stesse gare vengono ordinati correttamente 94 volte su cento; due che abbiano tirato a gare tutte diverse, 72. Ventidue volte su cento di differenza, per una cosa che con il tiro non c'entra niente.

Il motivo è semplice. Se due arcieri hanno tirato alle stesse gare, il vento di aprile lo hanno preso entrambi e nel confronto si cancella. Se hanno tirato a gare diverse non si cancella affatto, e diventa rumore puro che la classifica scambia per bravura.

E qui c'è il risultato che rende adottabile tutta questa parte. Rifacendo la stessa misura con la stima congiunta del capitolo 46 invece che con la media dei punteggi, la dipendenza dal calendario non si vede più: fra i disegni in cui il grafo arciere-gara resta connesso (una, due o tre gare in comune) l'affidabilità vale 64,5, 64,6 e 64,6, indistinguibili dentro un errore di quattro decimi su ciascuna. Con la media dei punteggi, nello stesso scenario, fra gli estremi ne corrono ventidue e quattro.

Il confronto con il calendario del tutto disgiunto non lo faccio, e conviene dire perché. Con zero gare in comune il grafo non è connesso in nessuna delle stagioni simulate, e con gli effetti di gara stimati insieme ai livelli la differenza di origine fra due componenti separate non sta nei dati: qualunque numero esca di là viene dalla soluzione a norma minima, cioè da un dettaglio numerico. È la stessa ragione per cui il capitolo 50 tiene fuori scala la riga a due blocchi disgiunti e per cui il capitolo 46 e l'appendice S18 lo dicono per esteso. Quello che la misura sostiene è che fra i calendari connessi la sovrapposizione smette di contare; non un rapporto fra un caso connesso e un caso che il modello non sa leggere.

Il che cambia la raccomandazione in meglio. Non «coordinate i calendari», che richiederebbe di mettere d'accordo decine di soggetti, ma: il problema del calendario si risolve nel modo di calcolare. È un intervento che si fa una volta e non richiede il consenso di nessuno.

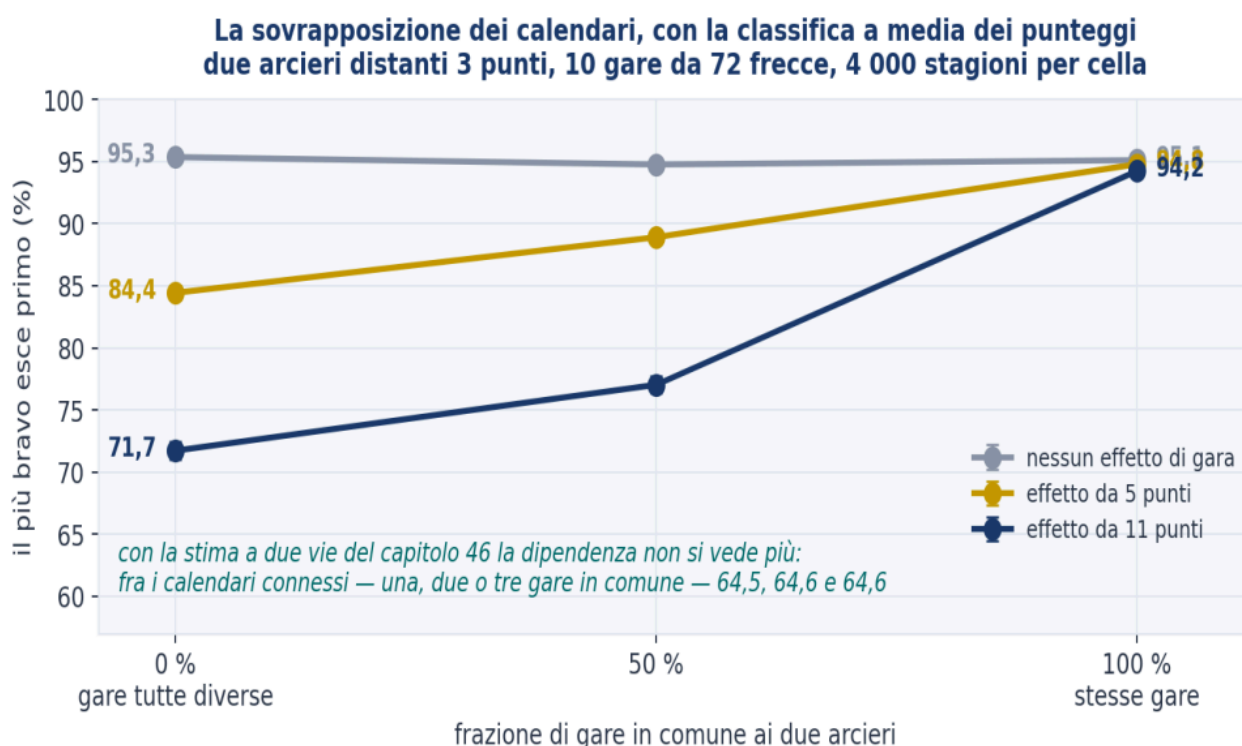


**Figura 16 — La sovrapposizione dei calendari, con la classifica a media dei punteggi.** Il più bravo esce primo, in funzione della frazione di gare condivisa e per tre ampiezze dell'effetto di gara. Due arcieri distanti 3 punti, 10 gare da 72 frecce, ancoraggio 700, 4 000 stagioni per cella. Con la stima congiunta del capitolo 46 la dipendenza non si vede più fra i calendari connessi.

## 46. La procedura, passo per passo

Ecco lo strumento, specificato in modo che chiunque possa implementarlo.

Il punto di partenza è una sola idea, e si dice in italiano prima che in simboli: il punteggio di un archiere in una gara è il suo livello, più la difficoltà di quella gara, più quello che resta. Il livello è quello che vogliamo stimare; la difficoltà è la stessa per tutti i presenti; quello che resta è la giornata buona o storta del singolo. In simboli,  $y_{ij} = \mu + a_i + e_j + \varepsilon_{ij}$ . Tutto quello che segue serve a ricavare i livelli e le difficoltà da quella.

Il primo passo raccoglie le osservazioni. Si mette in fila ogni coppia arciera-gara effettivamente disputata: chi ha tirato a sei gare su dieci contribuisce con sei righe, non con dieci. Le assenze non si riempiono e non si mediano, semplicemente non ci sono. Ed è questo che rende il resto delicato: se i calendari fossero tutti uguali quasi ogni procedura funzionerebbe, ed è perché non lo sono che il passo successivo va fatto in un modo solo.

Il secondo passo stima insieme i livelli degli arcieri e le difficoltà delle gare. Non prima gli uni e poi le altre: insieme, cercando i valori che rendono minimo lo scarto complessivo fra i punteggi previsti e quelli osservati, su tutte le righe raccolte.

Serve un'ancora, perché aggiungere una costante a tutti i livelli e sottrarla a tutte le difficoltà lascerebbe le previsioni identiche. L'ancora naturale è che le difficoltà sommino a zero, e con quella la soluzione è unica. La difficoltà di una gara diventa allora un numero di punti che si legge da solo: meno cinque significa che quella gara è costata cinque punti a chiunque vi abbia tirato.

Che i due insiemi vadano stimati insieme non è un dettaglio di implementazione, ed è il punto in cui una procedura può sbagliare senza che nessuno se ne accorga. La strada che viene in mente per prima, centrare ogni arciera sulla propria media e poi mediare gli scarti per gara, sembra fare la stessa cosa e non la fa: sottrae a ciascun arciera la media delle difficoltà del suo calendario, e se i calendari differiscono quella costante non è comune. Non è un errore che si attenua con più dati. Si costruisce senza fatica un esempio con cinque arcieri, cinque gare, nessun rumore e tutti collegati fra loro, in cui quella procedura mette davanti un arciera da 698 a uno da 700, mentre la stima congiunta recupera l'ordine esatto.

Il terzo passo restringe verso la media. La stima di chi ha tirato a due gare non va creduta quanto quella di chi ne ha fatte dieci, e il rimedio è tirare ciascuna verso la media generale di una quantità che dipende da quante gare ha alle spalle. Chi ha tirato molto resta quasi dov'è; chi ha tirato poco viene riportato verso il centro, perché il suo numero è più figlio del caso. Con presenze uguali il restringimento è lo stesso per tutti e l'ordine non cambia; con presenze diseguali fa esattamente il lavoro che ci si aspetta.

### **Quattro controlli che un'implementazione deve superare**

Non sono opzionali, perché ciascuno intercetta un guasto che lascerebbe uscire una classifica dall'aspetto perfettamente ragionevole.

Il primo verifica che tutti gli arcieri siano collegati fra loro, e va fatto prima di stimare qualunque cosa. Se il gruppo si spezza in due metà che non hanno nessuna gara in comune, la differenza fra le due non è nei dati: gli arcieri dell'una non sono confrontabili con quelli dell'altra se non aggiungendo un'ipotesi. Un'implementazione che non lo controlli produce una classifica dall'aspetto normale in cui due metà sono allineate per convenzione e non per misura, ed è l'unico controllo che intercetti quell'errore.

Il secondo prova separatamente le due manipolazioni possibili, perché un arciera può barare in due modi opposti: tirando a più gare possibile, oppure scegliendo soltanto le migliori. A presenze fissate le due coincidono, quindi un controllo solo sembra verificarle entrambe e non lo fa.

Il terzo verifica che i metodi ingenui risultino davvero manipolabili: la somma dei punteggi e la media semplice devono fallire, e se non falliscono è il controllo a essere rotto.

Il quarto confronta la stima congiunta con quella fatta in due tempi, su un caso senza rumore e con calendari sbilanciati. Devono dare risultati diversi, e la congiunta deve recuperare i livelli veri. Se coinci-

dono, l'ancora non sta agendo e la stima non è quella che si crede.

## 47. Un esempio percorso per intero

Cinque arcieri, 5 gare, presenze parziali. I punteggi sono inventati ma il meccanismo è quello vero, e alla fine confronteremo il risultato con la verità che ho usato per costruirli. Il caso è costruito attorno a un problema reale: Charlie ha saltato le due gare difficili, Bravo le ha fatte tutte.

|         | gara 1 | gara 2 | gara 3 | gara 4 | gara 5 | media | gare |
|---------|--------|--------|--------|--------|--------|-------|------|
| Alfa    | 697    | 683    | 699    | 679    | 691    | 689,8 | 5    |
| Bravo   | 691    | 677    | 691    | 679    | 682    | 684,0 | 5    |
| Charlie | 694    | —      | 687    | —      | 686    | 689,0 | 3    |
| Delta   | 685    | 666    | 684    | 671    | —      | 676,5 | 4    |
| Echo    | —      | 662    | 679    | 667    | 673    | 670,2 | 4    |

*(Esempio costruito: i punteggi sono generati da livelli veri noti e da difficoltà di gara note, che riveleremo alla fine per verificare che cosa la procedura sia riuscita a recuperare.)*

Guardando le medie grezze Charlie è secondo con 689,0, a un soffio da Alfa e cinque punti sopra Bravo. Ma Charlie ha saltato le gare due e quattro, e la domanda è se quelle due gare fossero più difficili delle altre.

Stimando insieme il livello di ogni arciere e la difficoltà di ogni gara, con l'ancora che le difficoltà sommino a zero, si ottiene questa.

|                    | gara 1 | gara 2 | gara 3 | gara 4 | gara 5 |
|--------------------|--------|--------|--------|--------|--------|
| Difficoltà stimata | +8,2   | −8,7   | +6,7   | −6,7   | +0,5   |

Le gare due e quattro erano difficili, tolgono otto e sei punti a chiunque vi abbia partecipato, e le gare uno e tre erano facili. Cioè: esattamente quello che Charlie ha evitato.

Correggendo i punteggi e restringendo si arriva alla classifica finale.

|         | Media grezza | Posizione      | Dopo la correzione | Posizione      | Livello vero |
|---------|--------------|----------------|--------------------|----------------|--------------|
| Alfa    | 689,8        | 1 <sup>a</sup> | 689,6              | 1 <sup>a</sup> | 690          |
| Charlie | 689,0        | 2 <sup>a</sup> | 683,8              | 2 <sup>a</sup> | 682          |
| Bravo   | 684,0        | 3 <sup>a</sup> | 683,9              | 3 <sup>a</sup> | 686          |
| Delta   | 676,5        | 4 <sup>a</sup> | 676,8              | 4 <sup>a</sup> | 678          |
| Echo    | 670,2        | 5 <sup>a</sup> | 672,6              | 5 <sup>a</sup> | 674          |

*(Stesso esempio. La colonna di destra è il livello vero usato per generare i punteggi, e serve soltanto a verificare il risultato: in una stagione reale non è osservabile. I valori corretti sono arrotondati a un decimale: il margine di Bravo su Charlie vale 0,15 punti, che la tabella mostra come 683,9 contro 683,8.)*

Ecco che cosa ha fatto la procedura. Sulla media grezza Charlie stava cinque punti sopra Bravo, per il solo fatto di aver saltato le due gare difficili; dopo la correzione il suo vantaggio scende a quattro decimi

di punto. La procedura ha recuperato quasi tutto il vantaggio indebito, cinque punti rubati al calendario, quattro e sei restituiti.

E adesso la parte che conta di più, che è il motivo per cui ho scelto questo esempio invece di uno che funzionasse meglio. Il margine di Bravo su Charlie è quindici centesimi di punto, e la tabella del capitolo 49 dice che con 5 gare servono tre punti di margine per riconoscere il migliore l'ottantacinque per cento delle volte. Un decimo e mezzo di punto non separa niente: la procedura sta dicendo che Charlie e Bravo sono indistinguibili, ed è la risposta corretta, perché in verità Bravo è più bravo di quattro punti e la classifica ha ancora l'ordine sbagliato.

L'esempio mostra dunque due cose insieme. La procedura fa il suo lavoro, perché toglie il vantaggio indebito di chi ha scelto le gare facili, e il quattro decimi contro i cinque punti è la misura di quanto lavoro abbia fatto. E non fa miracoli, perché con 5 gare e arcieri a quattro punti di distanza l'ordine resta incerto, ed è esattamente quello che la classifica deve dire. Non «Charlie è secondo», ma «Charlie e Bravo sono a quattro decimi, cioè non separati». Il capitolo 50 dice che cosa farne.

---

## 48. Il problema che nessuna quantità di dati risolve

Resta il difetto strutturale di ogni graduatoria sportiva, ed è quello su cui questa parte deve dire qualcosa di nuovo o non vale la pena farla: gli atleti scelgono a quali gare partecipare, e le gare che mancano non mancano a caso. Si salta la gara quando non si è in forma, si va a quella che conviene, si evita quella con il campo difficile.

Peggio, una graduatoria manipolabile scegliendo il calendario non è proponibile. Se conviene tirare poco e bene, il premio smette di misurare la precisione e comincia a misurare l'astuzia nella programmazione.

Il requisito però si può convertire in un esperimento, ed è la cosa che rende questo capitolo utile invece che una cautela. Si prende un arciere simulato, gli si dà la facoltà di scegliere le gare per arrivare più in alto possibile, si calcola la strategia migliore e si misura quanti posti guadagni.

Chi far manovrare non è indifferente, e la scelta va dichiarata. L'arciere che manipola è il sedicesimo su trentadue, cioè uno di mezza classifica, ed è la posizione da cui si guadagna di più: chi fosse già primo non avrebbe niente da manovrare, e chi fosse ultimo non arriverebbe comunque in alto. La base di confronto sono sei gare scelte a caso su dieci, e il numero che segue è il caso peggiore fra le posizioni provate, non quello tipico.

| Metodo di classifica            | Posizione senza barare | Tirando a tutte le gare | Scegliendo le migliori |
|---------------------------------|------------------------|-------------------------|------------------------|
| Somma dei punteggi              | 16,0                   | 1,00 (+15,0)            | 10,85 (+5,2)           |
| Media dei punteggi              | 16,0                   | 16,04 (+0,0)            | 10,84 (+5,2)           |
| Stima congiunta del capitolo 46 | 16,1                   | 16,03 (+0,0)            | 15,16 (+0,9)           |

*(Arciere sedicesimo su 32, stagione di 10 gare, base sei gare a caso. 600 stagioni per la prima colonna e 500 per la seconda, in confronto appaiato.)*

La prima riga è quella che dovrebbe chiudere per sempre la questione della somma dei punteggi. Un arciere di mezza classifica passa dal sedicesimo al primo posto semplicemente iscrivendosi a tutte le gare

invece che a sei: quindici posti, senza tirare una freccia meglio di prima. Chi vince quella classifica è chi si iscrive di più.

E le due manipolazioni sono opposte e speculari, il che è la cosa da capire: la somma premia chi tira di più e non premia chi salta le brutte giornate, la media fa esattamente il contrario. I tre metodi non stanno dunque su una scala di qualità, stanno su due assi, e un confronto su una manipolazione sola li ordina male, che è quello che chiunque farebbe provando la sola manipolazione che gli viene in mente.

Soltanto la stima congiunta regge entrambe. E non del tutto: nove decimi di posto, che vanno riportati e non arrotondati a zero. Un arciero che scelga bene le proprie gare guadagna meno di un posto, dove con la media ne guadagnerebbe cinque.

Il restringimento toglie l'effetto della gara ma non completamente, perché quell'effetto è stimato dagli stessi dati che l'arciero ha scelto di produrre: senza ipotesi su come sceglie, nessuna procedura che impari dai soli dati osservati può essere garantita immune a chi decide quali dati produrre. E la manipolazione che resta è proprio quella che nello sport si osserva davvero: nessuno tira male di proposito, molti saltano la gara quando non si sentono in forma.

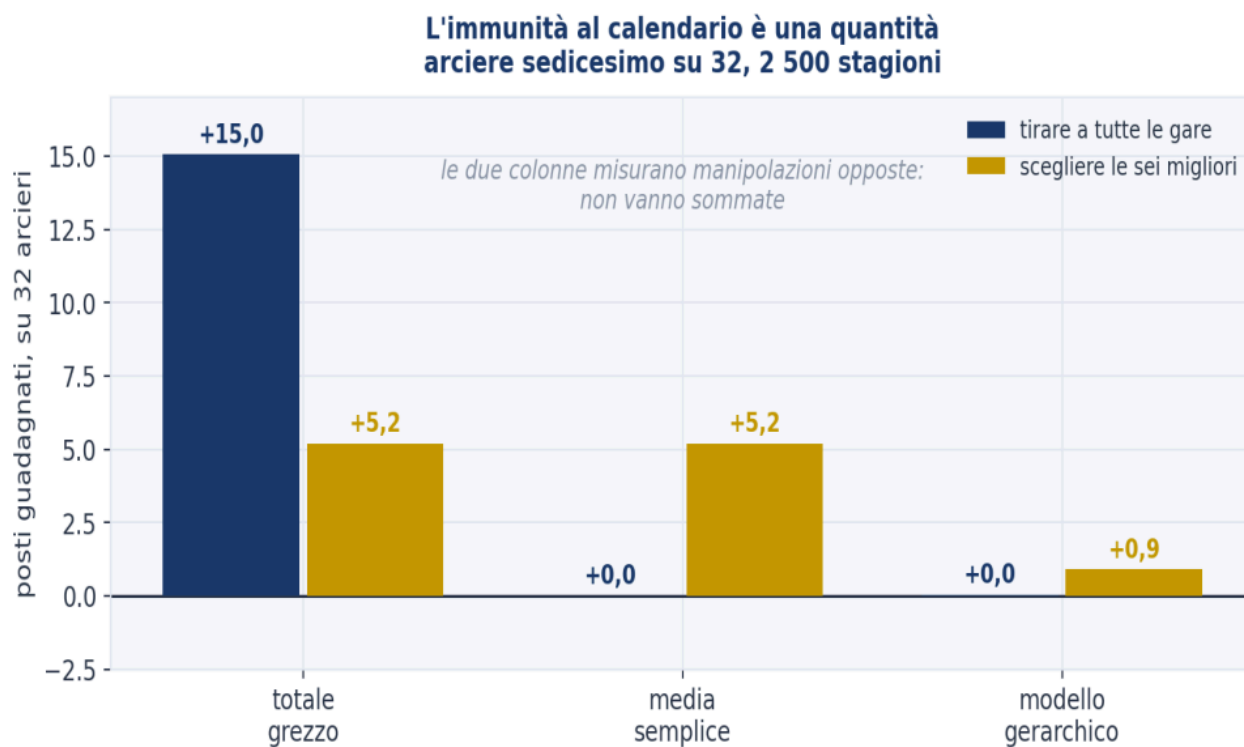


**Figura 17 — L'immunità al calendario è una quantità.** Posti guadagnati da un arciero sedicesimo su 32, in un campo di 32 atleti fra 700 e 660, che scelga strategicamente il calendario, per due manipolazioni opposte. 2 500 stagioni, in confronto appaiato. Le due colonne misurano manipolazioni diverse e non vanno sommate.

## 49. Chi entra in classifica, e quando il premio è individuale

Questa è la parte che una federazione applicherebbe per prima, e sono due regole scrivibili in un regolamento.

La prima riguarda quando un premio individuale sia difendibile. La domanda pratica non è quanto sia accurata la graduatoria in generale, ma se questo primo posto, con questo margine, regga.

| Gare disputate | Margine 1 punto       | Margine 2 punti       | Margine 3 punti        | Margine 5 punti        |
|----------------|-----------------------|-----------------------|------------------------|------------------------|
| 5              | 72 % [70–73] (60,3 %) | 79 % [77–81] (33,4 %) | 85 % [83–87] (15,8 %)  | 85 % [77–91] (1,7 %)   |
| 10             | 85 % [84–86] (58,1 %) | 91 % [90–93] (27,4 %) | 95 % [93–96] (10,1 %)  | 100 % [86–100] (0,4 %) |
| 20             | 95 % [94–95] (58,3 %) | 98 % [98–99] (23,4 %) | 100 % [98–100] (5,2 %) | 100 % [21–100] (0,0 %) |

*(Probabilità che il primo della graduatoria sia davvero il migliore del campo, dato il margine osservato sul secondo, con l'intervallo di Wilson al 95 per cento fra parentesi quadre. Fra parentesi tonde, quanto spesso quel margine si presenta. È la domanda che il premio pone: non se il primo sia migliore del secondo, ma se sia il migliore di tutti. 32 atleti fra 700 e 660, calendari completi, stima congiunta, 6 000 stagioni per riga. L'intervallo dice quanto poco significhi un cento per cento su pochi casi: con ventun stagioni sopra i cinque punti di margine vale «fra 85 e 100», con tre stagioni «fra 44 e 100».)*

Quale margine basti non lo decide la matematica: dipende da quanta certezza si voglia, e quella è una scelta della federazione. Il conto dice questo.

Con dieci gare, un punto di margine dà l'85 per cento, cioè assegnando il premio con un solo punto di vantaggio lo si sbaglia una volta su sei, e in trent'anni di premi sono quattro o cinque assegnati alla persona sbagliata. Due punti danno il 91, tre il 95. Se la soglia che si considera sufficiente è il novantacinque per cento, con dieci gare servono tre punti e non due; se è il novanta, ne bastano due; con venti gare due punti danno già il 98,5.

I valori fra parentesi tonde sono quelli che rendono la regola adottabile invece che teorica, ed è la ragione per cui li ho calcolati. A dieci gare, un margine di due punti si presenta in poco più di una stagione su quattro, e uno di tre punti in una su dieci. Alzare la soglia di certezza costa in cadenza: più si vuole essere sicuri, più raramente il premio è assegnabile.

Quindi con la soglia a due punti il premio individuale si assegna un anno su quattro, e negli altri tre la forma onesta è collettiva. Una federazione sa quale cadenza aspettarsi prima di adottare la regola, invece di scoprirlo alla terza stagione, ed è la differenza fra una regola che si può discutere e una che si subisce.

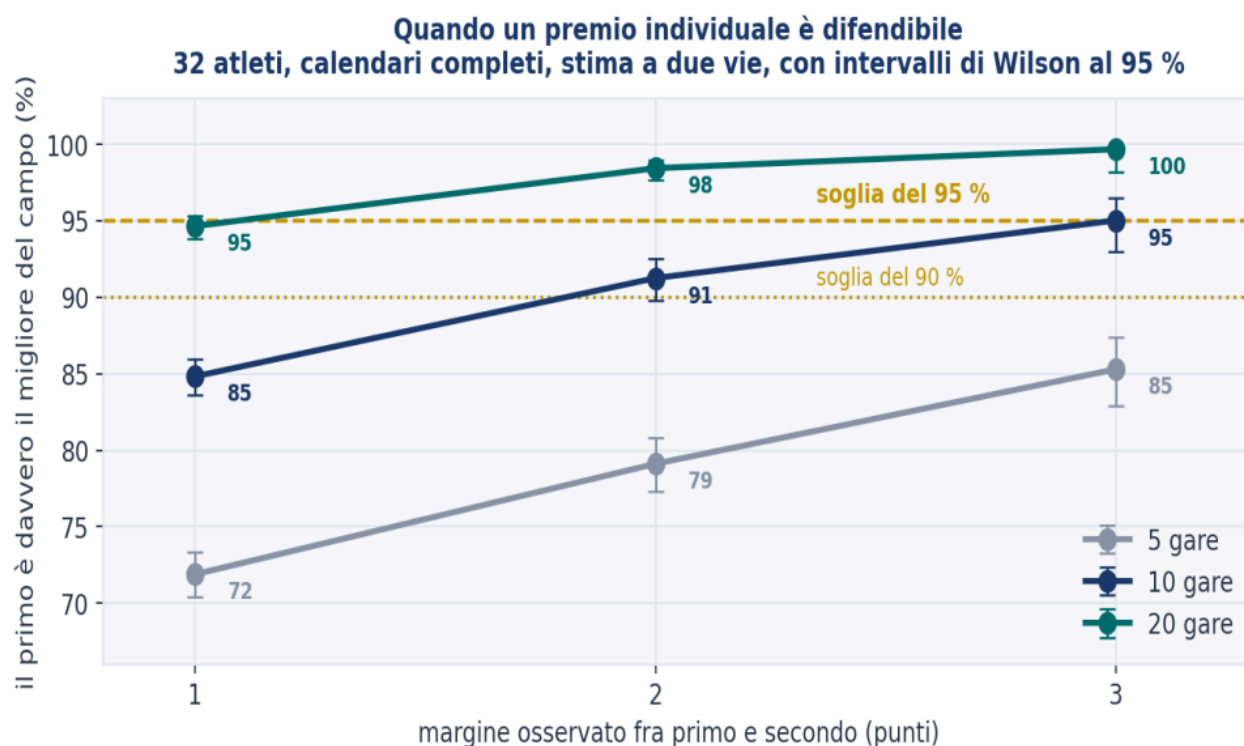


**Figura 18 — Quando un premio individuale è difendibile.** Probabilità che il primo della graduatoria sia davvero il migliore del campo, in funzione del margine osservato sul secondo e del numero di gare. 32 atleti, calendari completi, stima congiunta.

La seconda regola riguarda quante gare minime servano per entrare in classifica.

| Soglia       | Posti guadagnabili manipolando |
|--------------|--------------------------------|
| 3 gare su 10 | 1,19                           |
| 5 gare       | 0,79                           |
| 6 gare       | 0,63                           |
| 8 gare       | 0,00                           |
| 10 gare      | 0,00                           |

(Arciere sedicesimo su 32, 1 500 stagioni simulate per soglia.)

La manipolabilità cala con continuità e si azzerava quando la soglia supera le presenze tipiche, perché a quel punto saltare una gara significa uscire dalla classifica, e il costo della manipolazione diventa infinito.

Ma lo zero si compra escludendo atleti, ed è un costo che nessuna tabella di accuratezza contiene: una soglia a otto gare su dieci esclude chi si infortuna, chi ha impegni di lavoro, chi vive lontano dalle sedi di gara. Sei gare su dieci è il punto in cui la curva si è già appiattita senza pagare quel prezzo: fra sei e otto si comprano sei decimi di posto e si perde una categoria di atleti.

E vale la pena dire la cosa scomoda: una classifica perfettamente immune alla manipolazione è una classifica in cui partecipare è obbligatorio. L'immunità totale e l'accessibilità sono in conflitto per costruzio-

ne, e chi chiede la prima sta chiedendo la seconda senza saperlo.

Le due regole si scrivono così.

### **Articolo — Graduatoria stagionale e premio annuale**

1. La graduatoria stagionale è calcolata su tutte le gare del calendario ufficiale disputate nell'anno, secondo la procedura di cui all'allegato tecnico.
2. Entra in graduatoria l'atleta che abbia disputato almeno il 60 per cento delle gare del calendario, con un minimo di quattro.
3. Accanto a ciascuna posizione sono pubblicati il valore stimato, il margine sull'atleta successivo e il numero di gare disputate.
4. La federazione dichiara in anticipo la certezza richiesta per un premio individuale, e dalla tavola del capitolo 49 ricava il margine corrispondente: con dieci gare a calendario, due punti per il novanta per cento e tre per il novantacinque. Il premio di atleta dell'anno è assegnato individualmente se il margine del primo sul secondo raggiunge quella soglia; altrimenti è assegnato collettivamente agli atleti che dal primo distano meno della soglia.
5. La graduatoria è aggiornata dopo ogni gara del calendario ed è pubblicamente consultabile.

Il terzo comma è quello che rende la regola verificabile ed è la sua parte più importante: pubblicando il margine, chiunque può controllare da sé se il primo posto sia difendibile, mentre senza quel dato la regola del quarto comma sarebbe una decisione presa altrove. Il quinto serve invece alla Parte seconda, perché se la graduatoria deve comporre i tabelloni dev'essere aggiornata e pubblica prima di ogni gara.

---

## **50. Che cosa conta davvero in un calendario**

Se la procedura del capitolo 46 riduce quasi a zero la dipendenza dal calendario, resta da chiedersi se il calendario conti ancora. La risposta è sì, ma non per la ragione che ci si aspetterebbe.

I sei calendari confrontati sono costruiti così, su trentadue atleti e dodici gare con sei presenze a testa. Nei blocchi disgiunti i primi sedici tirano le gare dalla prima alla sesta e gli altri sedici dalla settima alla dodicesima, senza mai incontrarsi. Nei calendari a gare in comune una, due o tre gare sono obbligatorie per tutti e le restanti si sorteggiano dentro il proprio blocco. Nel sorteggio libero ogni atleta pesca sei gare fra dodici senza vincoli, ed è il caso realistico. Nel sorteggio bilanciato si pesca ugualmente, ma con sedici partecipanti esatti per gara — serve a separare l'effetto della struttura da quello dello squilibrio di presenze.

| Calendario            | Gare in comune, minimo | Il primo è il migliore |
|-----------------------|------------------------|------------------------|
| due blocchi disgiunti | 0                      | 63,2 %                 |
| una gara in comune    | 1                      | 64,5 %                 |
| due gare in comune    | 2                      | 64,6 %                 |
| tre gare in comune    | 3                      | 64,6 %                 |
| sorteggio bilanciato  | 0                      | 63,0 %                 |
| sorteggio libero      | 1                      | 63,8 %                 |

*(32 atleti fra 700 e 660, 12 gare da 72 frecce, 6 presenze a testa, effetto di gara normale con deviazione standard di 11 punti, 12 000 stagioni appaiate per tutti i disegni. Come per la tabella del capitolo 52, i livelli dipendono dall'ampiezza assunta per l'effetto di gara mentre il verso del confronto regge alla riparametrizzazione.)*

Va distinta una cosa dall'altra, perché sono due domande diverse, e la prima riga è un caso a parte. Con due blocchi che non condividono alcuna gara i due gruppi non sono collegati, e la differenza fra loro non è nei dati: il numero che ne esce dipende da come si chiude quell'indeterminazione, non da quanto bene la procedura funzioni. Quella riga non va messa sulla stessa scala delle altre.

Le altre cinque sono confrontabili, e dicono questo: condividere anche una sola gara vale sette decimi di punto sul sorteggio libero, dividerne due otto decimi e tre otto decimi, con un errore di cinque decimi ciascuno: le tre non sono distinguibili fra loro. Non è la struttura a blocchi in sé a comprare accuratezza.

La ragione è l'appaiamento fra i contendenti, ed è più fine di quanto sembri. In un calendario a blocchi i migliori tirano tutti alle stesse gare, quindi il confronto fra i candidati al primo posto è perfettamente appaiato: il vento e la luce si cancellano proprio nella comparazione che decide la classifica. Nel sorteggio, invece, l'atleta migliore condivide soltanto la metà delle gare con ciascun rivale, e parte del confronto passa per via indiretta.

Il controllo lo conferma sulla stessa corsa. Costringere i soli otto candidati al vertice a tirare alle stesse sei gare (lasciando gli altri ventiquattro liberi di sorteggiare) vale otto decimi di punto; appaiarne sedici 0,7; appaiarli tutti e trentadue 0,8. I tre contrasti portano un errore di cinque decimi ciascuno, quindi il verso è sempre lo stesso ma la grandezza questa corsa non la separa dallo zero, e i tre non si distinguono fra loro.

Ed è questo il criterio di progetto di un calendario, e per una federazione è una buona notizia: non serve che tutti si incontrino, serve che si incontrino i contendenti. Sei gare comuni agli otto candidati al vertice è la configurazione provata, e il resto del calendario può essere sorteggiato. Il segno del guadagno è coerente in tutte le misure, ma vale 0,8 punti con un errore di 0,6: non è separato dallo zero, e sei non è una soglia di sufficienza dimostrata. È una raccomandazione molto più economica di «tutti a tutto», perché non richiede di obbligare nessuno.

Va detto chi siano quegli otto, perché è la differenza fra una raccomandazione teorica e una applicabile. Nella prima misura sono gli otto migliori veri, che nessuna federazione conosce in anticipo: è il limite superiore per una selezione perfetta. Rifacendo l'esperimento con gli otto scelti dalla graduatoria della stagione precedente (l'informazione che una federazione ha davvero) il risultato non cambia: 64,5 con gli otto migliori veri, 64,6 con quelli scelti dalla graduatoria, una differenza di meno di un decimo, con-

tro un errore di quattro decimi su ciascuno dei due. La raccomandazione è quindi adottabile, e questo è misurato e non assunto.

Quante gare in comune bastino l'ho misurato facendone condividere agli otto da zero a sei, tenendo identico tutto il resto. Con una o due gare in comune non si misura nulla: i guadagni sono tre decimi e mezzo e tre decimi, dentro un errore di tre decimi. Con tre il segnale c'è: nove decimi, tre volte l'errore. Da lì in su la griglia non sale in modo ordinato, sei decimi con quattro gare, sette decimi con cinque, un punto e quattro con sei, e la distanza fra un punto e l'altro resta dell'ordine dell'errore. Sei è il massimo fra quelli provati, non una soglia dimostrata. I valori più elevati tendono a comparire quando aumenta il numero di gare condivise, ma questa simulazione non dimostra una crescita monotona né identifica una soglia di saturazione.

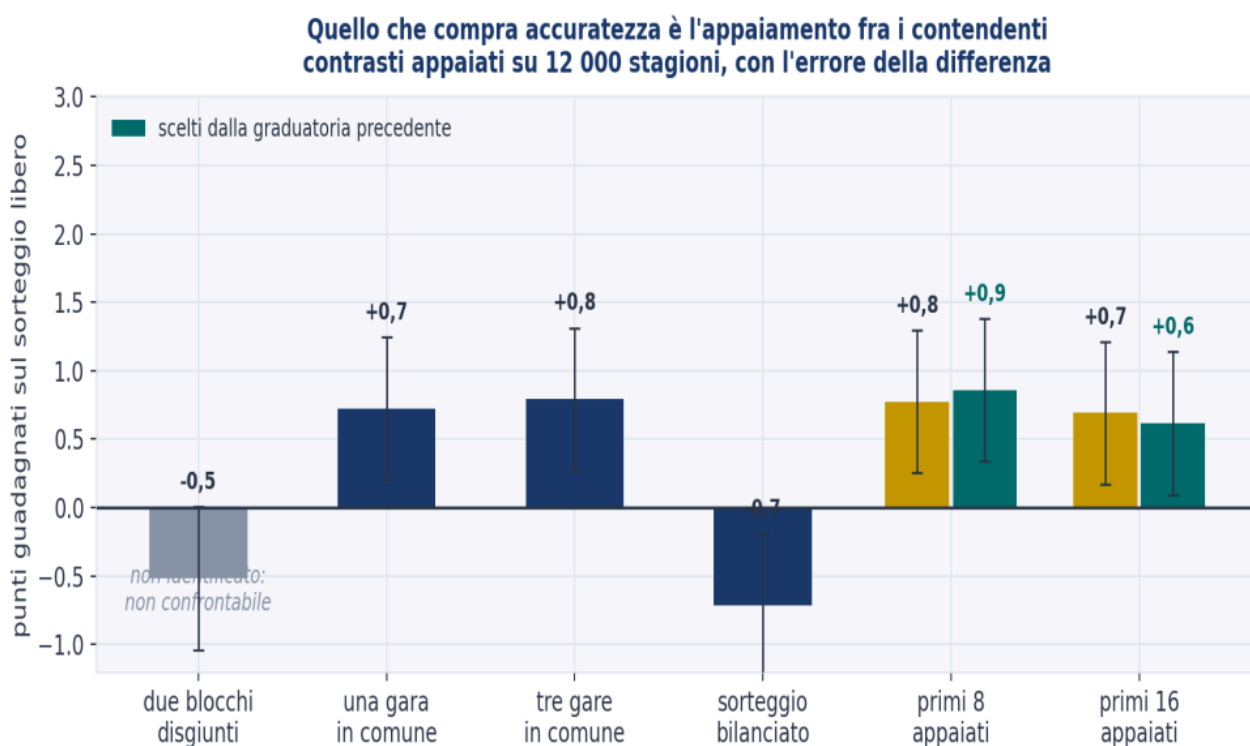


**Figura 19 — Quello che compra accuratezza è l'appaiamento fra i contendenti, non la struttura del calendario in sé.** Punti guadagnati sul sorteggio libero, misurati come contrasto appaiato stagione per stagione, con l'errore della differenza. Le barre verdi sono la selezione fatta in anticipo: i contendenti scelti dalla graduatoria dell'anno prima, cioè l'informazione che una federazione ha davvero. La prima colonna non è identificata e non va confrontata con le altre. 32 atleti, 12 gare, 6 presenze a testa, 12 000 stagioni, stima congiunta.

## 51. Quanti dati servono

Questo è il pezzo che rende lo strumento adottabile, perché una federazione deve poter sapere prima quanti dati le servano invece di scoprirlo dopo.

| Gare | 36 frecce a gara  | 72 frecce a gara  | 144 frecce a gara |
|------|-------------------|-------------------|-------------------|
|      | primo / primi tre | primo / primi tre | primo / primi tre |
| 5    | 52,0 / 34,3       | 61,5 / 50,1       | 72,8 / 67,6       |
| 10   | 62,1 / 50,6       | 71,5 / 67,2       | 82,8 / 81,8       |
| 20   | 73,0 / 68,5       | 83,7 / 82,3       | 92,2 / 91,1       |
| 40   | 83,4 / 80,6       | 92,4 / 90,5       | 97,9 / 96,8       |

*(Probabilità di collocare correttamente il primo, e di individuare correttamente i primi tre come insieme. Calendari completi, stima congiunta del capitolo 46, 32 atleti fra 700 e 660, 3 000 stagioni per cella.)*

La cella da guardare è quella del calendario tipico: dieci gare da 72 frecce, che è quello che una federazione nazionale offre. Il primo della graduatoria è davvero il migliore 71,5 volte su cento, cioè sbaglia tre stagioni su dieci. È molto meglio del titolo di un torneo, che sbaglia nove volte su dieci, e resta lontano dalla certezza.

Per superare il novanta per cento, nella griglia provata, servono venti gare da 144 frecce oppure quaranta da 72. Le due strade costano lo stesso, perché entrambe portano a 2 880 frecce per atleta contro le 720 di oggi: quattro volte l'impegno attuale, cioè quattro stagioni di oggi compresse in una. Che siano le combinazioni minime non lo dico, dico che sono le più economiche fra quelle provate.

Le due leve non sono separate da questa griglia. Raddoppiare le gare porta da 71,5 a 83,7, raddoppiare le frecce porta da 71,5 a 82,8, e le due destinazioni distano un punto con un errore di un punto. Non è una prova che valgano lo stesso: per dirlo servirebbe fissare in anticipo quanto piccola una differenza debba essere per chiamarla nulla, e verificare che l'intervallo ci stia dentro. Detto in pratica, fra venti gare da 72 frecce e dieci da 144 la simulazione non separa, quindi la scelta fra le due va fatta sul costo e sul calendario, non sull'accuratezza. Conta invece chi le tiri insieme a chi, come il capitolo precedente ha mostrato.

E c'è un fatto che vale la pena guardare: il divario fra individuare il primo e individuare il terzetto si chiude quasi del tutto al crescere dei dati.

A cinque gare da 36 frecce il primo esce giusto nel 52 per cento dei casi e il terzetto nel 34: diciotto punti di distanza. Con una stagione corta, quindi, premiare un podio è molto più difficile che premiare un vincitore, e una federazione che assegni tre premi ne sbaglia almeno uno due volte su tre. A quaranta gare da 144 frecce il primo vale 97,9 e il terzetto 96,8: un punto di distanza, che con l'errore di questa corsa non si distingue da zero.

Che le due domande costino esattamente lo stesso non lo dimostro, e la differenza fra le due affermazioni è quella fra non aver trovato una differenza e aver trovato un'uguaglianza. Il terzetto resta il problema più difficile in tutte e dodici le celle e non lo supera mai. Quello che cambia con i dati non è quale delle due domande sia più facile, ma quanto la scelta fra le due pesi: con pochi dati pesa moltissimo, con molti non pesa quasi più.

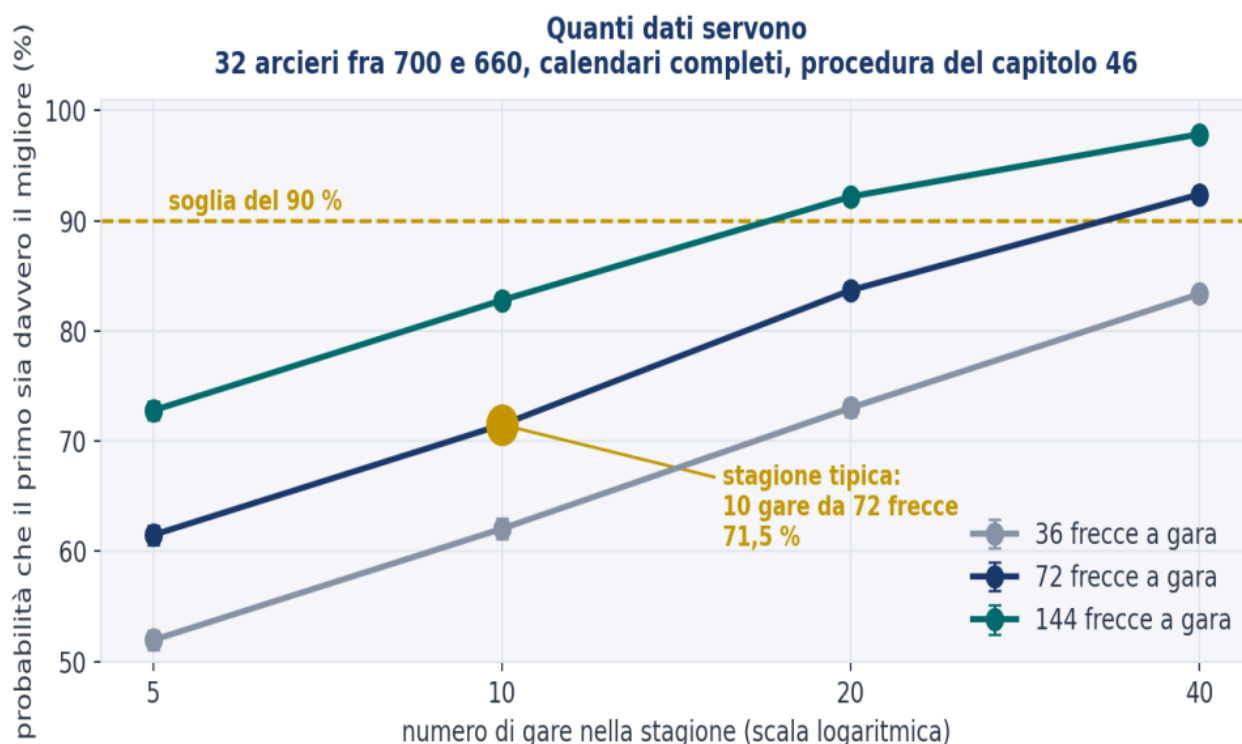


**Figura 20 — Quanti dati servono.** Probabilità di collocare correttamente il primo, per numero di gare e frecce per gara. Calendari completi, 32 atleti fra 700 e 660. Scala logaritmica; in evidenza la stagione tipica.

## 52. Il confronto finale: quanto vale una stagione

Ecco infine la quantità che si mette accanto al dieci per cento del torneo.

|                                    | Con calendari parziali | Con calendari completi |
|------------------------------------|------------------------|------------------------|
| Classifica a punti per piazzamento | 33,0 %                 | 72,7 %                 |
| Somma dei punteggi                 | 16,4 %                 | 73,0 %                 |
| Media dei punteggi                 | 37,1 %                 | 72,8 %                 |
| Stima congiunta del capitolo 46    | 58,6 %                 | 73,0 %                 |

(Probabilità che il primo della graduatoria sia davvero il migliore. 32 atleti fra 700 e 660 — quindi 1,3 punti fra vicini di classifica — 10 gare da 72 frecce, effetto di gara normale con deviazione standard di 11 punti, 4 000 stagioni. Nei calendari parziali ogni atleta tira a un numero di gare estratto fra tre e sette.)

Una nota sulla solidità di questi quattro numeri, perché è diversa da quella del resto del volume. I livelli dipendono da quanto si assuma grande l'effetto della gara, che è un valore dichiarato e non misurato: rifacendo il conto con un'assunzione un po' diversa, la prima colonna si sposta di un paio di punti e la seconda di uno. Le distanze fra le righe invece non si spostano: a calendari completi i quattro metodi restano entro 1,8 punti, e a calendari parziali la stima congiunta resta sopra la media semplice di 21,5 punti e sopra la classifica a punti di 25,6. Quello che si porta a casa da questa tabella sono le distanze, non i livelli.

**A calendari completi i quattro metodi stanno vicini**, fra 72,7 e 73,0: lo spazio fra metodi ragionevoli è largo tre decimi di punto, e la scelta del metodo quasi non decide — come non decidevano la regola di parità e il formato del match.

**A calendari parziali decide eccome**, ed è l'unico posto del volume in cui la scelta del metodo vale più della struttura. Fra il migliore e il peggiore dei quattro passano quarantadue punti.

Ma quello che separa non è la qualità del metodo: è quanto ciascuno resista al fatto che non tutti tirano alle stesse gare. La somma dei punteggi crolla al 16,4 per cento (cioè sbaglia il primo posto cinque volte su sei) perché premia chi ha tirato di più. La classifica a punti non va molto meglio, perché anch'essa accumula. La media semplice regge un po' più delle prime due perché normalizza. E la stima congiunta sta sopra tutte al 58,6 perché è l'unica che tratti la presenza come un dato del disegno invece che come merito.

E c'è una terza lettura che vale più delle altre due. Completare i calendari vale molto, e a calendario completo la scelta del metodo quasi non conta: passare da calendari parziali a completi porta la stima congiunta dal 58,6 al 73,0, cioè quattordici punti, e una volta completi i quattro metodi stanno entro tre decimi. Con calendari parziali, invece, il metodo conta moltissimo (quarantadue punti fra il migliore e il peggiore) e le due affermazioni vanno tenute distinte. Detto in stagioni: con i calendari completi la graduatoria sbaglia il primo posto due stagioni su sette, con quelli parziali quattro su dieci, e nessun raffinamento del calcolo recupera quella differenza.

Il numero da tenere è quello: con 10 gare e calendari completi il primo della graduatoria è davvero il migliore tre volte su quattro, contro una volta su dieci del titolo di un torneo com'è organizzato oggi, e una su sette con la configurazione migliore della Parte seconda, ricordando che il torneo è misurato su 64 arcieri e la stagione su 32, quindi le due cifre descrivono due scenari e non si dividono l'una per l'altra.

---

## 53. Che cosa portarsi a casa dalla Parte terza

**Una stagione ordina tre punti e non uno.** Dieci gare distinguono due arcieri separati da tre punti su 720 con un'affidabilità del 95,3 per cento; per distinguerne uno servirebbero 85 gare, cioè otto stagioni e mezza a calendario pieno.

**Un punteggio non si confronta con un altro** senza correggere per la difficoltà della gara: ignorarlo costa circa ventidue punti nello scenario provato, e la stima congiunta li annulla — fra i calendari connessi non resta dipendenza misurabile.

**L'immunità al calendario è una quantità e si compra.** Con la somma dei punteggi un arciere di mezza classifica guadagna quindici posti scegliendo dove tirare; con la procedura corretta ne guadagna meno di uno. L'immunità totale costa l'esclusione di chi non può esserci.

**Il premio è individuale sopra un margine dichiarato in anticipo e collettivo sotto.** Con dieci gare, due punti di margine danno il novanta per cento di certezza e capitano poco più di una stagione su quattro; tre punti danno il novantacinque e capitano una su dieci. Quale certezza serva lo decide chi assegna il premio, non questo conto.

**E il calendario conta per due cose.** La prima è un requisito: tutti gli arcieri devono essere collegati fra loro da gare condivise, altrimenti i livelli di due gruppi separati non sono confrontabili. La seconda,

dato quel requisito, è che i contendenti al vertice tirino alle stesse gare.

## APPENDICE B — QUANDO LA FORMA CONTA

Tutto quello che c'è in questo volume poggia su una frase che non ho mai messo alla prova. Ho descritto ogni arciere con un numero solo, quanto sono sparse le sue frecce attorno al centro, e da lì ho ricavato la probabilità di ogni anello, di ogni punteggio, di ogni match. Assumere che quel numero basti significa assumere che gli impatti si distribuiscano come una gaussiana: una nuvola simmetrica che si dirada in modo regolare allontanandosi dal centro.

È un'assunzione comoda, e c'è evidenza empirica a suo favore. Dall e Park, su oltre duecentomila impatti, trovano che la distanza dal centro segue molto da vicino la forma che una nuvola gaussiana tonda produce, e che la stessa forma vale per compound e ricurvo, uomini e donne, indoor e outdoor, dal livello intermedio all'élite.

Quello che un buon accordo sull'insieme non esclude sono le code. Chi tira sa che un arciere non sbaglia sempre allo stesso modo: la maggior parte delle frecce sta dentro un gruppo stretto, e ogni tanto ne parte una che con quel gruppo non c'entra niente — un rilascio sporco, una raffica presa male, un attimo di disattenzione al momento sbagliato. Se quelle frecce esistono e non sono rarissime, la nuvola non è più una gaussiana: è un nucleo stretto più una minoranza dispersa.

Questo capitolo chiede che cosa succede alle conclusioni del volume se il mondo è fatto così. Non sostituisce il modello: lo mette dentro una famiglia più ampia e prova a rompere ogni risultato che ho ottenuto.

### Il modello, e che cosa non afferma

Ogni impatto viene da una delle due componenti. Quasi sempre dal nucleo; ogni tanto, chiamiamo quella frequenza  $\epsilon$ , da una componente che ha la stessa forma ma è più larga di un fattore che chiamo gravità. Con  $\epsilon$  uguale a zero si ritrova esattamente il modello del volume, e non per somiglianza: la calibrazione restituisce la stessa ampiezza fino alla decima cifra, ed è una delle prove che l'apparato esegue a ogni ricostruzione (formule S22 e S23).

Un arciere non è più un numero ma un profilo di tre: quanto è stretto il nucleo, quanto spesso arriva un tiro fuori pattern, quanto è grave quando arriva.

Serve dire subito che cosa questo capitolo non fa. Non afferma che una qualunque percentuale di frecce anomale descriva gli arcieri veri. Nessuno dei valori di  $\epsilon$  che userò è stimato da referti: sono scenari. La domanda a cui rispondo è più modesta e più utile: quali conclusioni reggono su tutta una classe di distribuzioni, e quali dipendono invece dalla forma delle code. **La cosa da chiedersi non è quanto valga  $\epsilon$ : è quanto dovrebbe valere per farmi sbagliare.**

### Perché la domanda non è oziosa

Il primo conto è quello che ha reso necessario tutto il resto. Ho preso un arciere da 690 punti e l'ho ricostruito sotto ogni combinazione di frequenza e gravità, calibrando ogni volta il nucleo perché il punteggio atteso restasse 690 esatto. Tutti questi arcieri valgono lo stesso: cambia solo come ci arrivano.

Non cambia solo quello. La dispersione del punteggio di round, cioè quanto un arciere oscilla da una gara all'altra, passa da 4,58 punti nel modello gaussiano a 5,71 quando cinque frecce su cento vengono da una componente con deviazione tripla, e a 7,47 quando dieci frecce su cento vengono da una con deviazione quadrupla.

Detto in modo che si veda: nel mondo gaussiano un arciere da 690 oscilla, da una gara all'altra, di quattro punti e mezzo; nell'ultimo di sette e mezzo. È un arciere che sembra molto meno affidabile, pur avendo la stessa media.

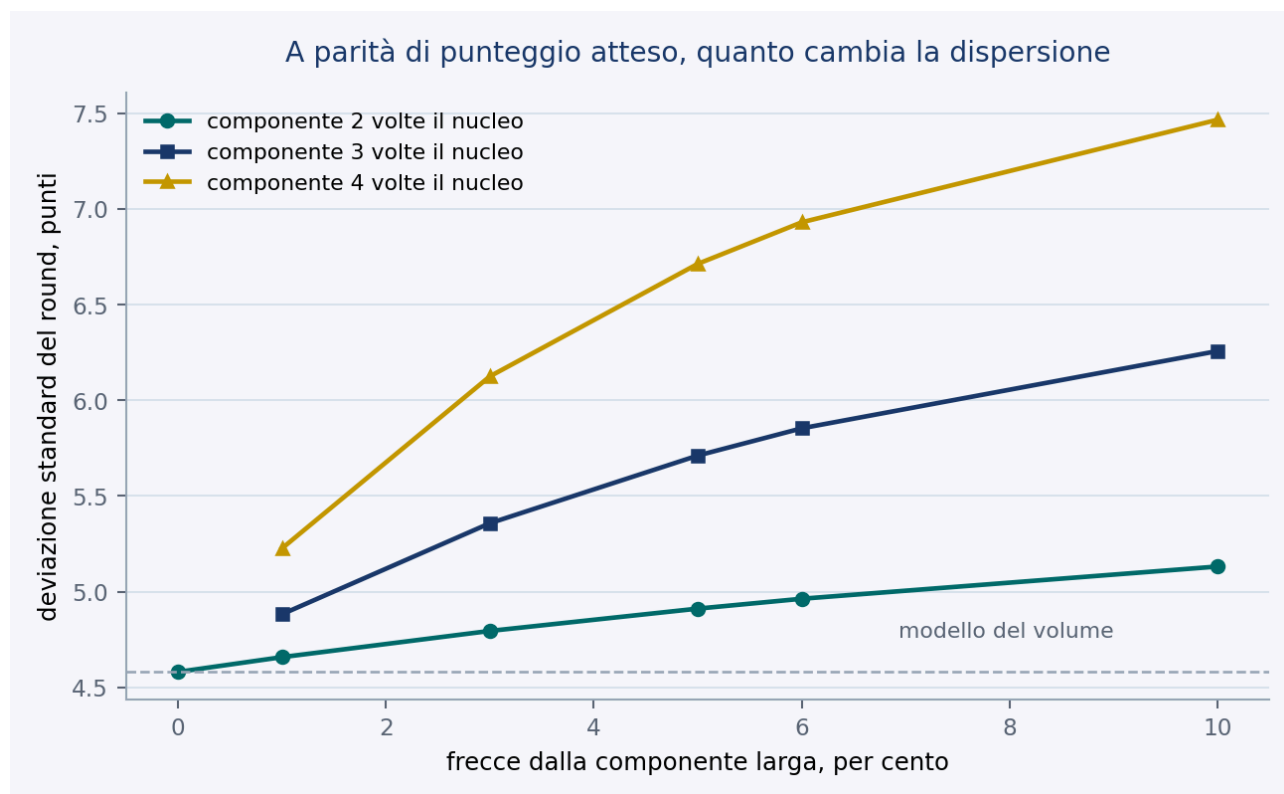


**Figura 21 — La dispersione del punteggio sotto contaminazione.** Deviazione standard del punteggio di round per un arciere da 690, al variare della frequenza e della gravità delle frecce fuori pattern. Con frequenza nulla si ritrova il modello del volume.

E la dispersione del punteggio è la grandezza su cui poggia metà di questo libro: quanto spesso due arcieri arrivino pari, l'affidabilità dello spareggio, quella del totale del match, la precisione della stima stagionale. Se cambia di oltre il sessanta per cento, non si può dare per scontato che le conclusioni restino dove sono.

### **I match reggono, e il motivo è istruttivo**

Il primo esperimento sembrava dare un risultato forte, e si è rivelato sbagliato. Vale la pena raccontarlo perché il modo in cui è caduto dice qualcosa sul problema.

Ho preso due arcieri con lo stesso punteggio atteso: il primo regolare, nucleo largo e quasi nessuna freccia anomala; il secondo più preciso ma fragile, nucleo stretto e cinque frecce su cento fuori pattern. Li ho fatti affrontare duecentomila volte sotto quattro sistemi. Il set system premiava il fragile di tre punti e mezzo, cinque volte più del totale cumulativo: sembrava che i formati pesassero diversamente precisione ordinaria e rischio di fallimento.

Poi ho appaiato anche la dispersione. Con lo stesso punteggio atteso e la stessa oscillazione da gara a gara l'effetto quasi sparisce: su tutti i sistemi provati lo scarto dal cinquanta per cento sta dentro tre decimi di punto. Per la freccia di spareggio la probabilità si calcola esattamente, senza simulare, e vale 49,7497 per cento: lo scarto dal pari è di 0,2503 punti, cioè su mille spareggi il fragile ne perde due e mezzo in più del regolare.

Su questa famiglia di distribuzioni, dunque, quasi tutto il vantaggio dell'arciere fragile veniva dalla dispersione e non dalla forma.

Sarebbe scorretto trarne che a media e oscillazione uguali le code non contino mai. Non è vero, e basta un esempio a mostrarlo. Si prenda A che fa 2 punti sette volte su dieci e 4 punti tre volte su dieci; e B che fa 0 punti una volta su dieci, 2 punti una volta su dieci e 3 punti otto volte su dieci. Hanno la stessa media, 2,6, e la stessa oscillazione — la varianza vale 0,84 per entrambi; eppure nel confronto diretto A vince 40,5 volte su cento e B 59,5. I due numeri riassuntivi non determinano chi vince, perché vincere non è una funzione lineare del punteggio.

Quello che il conto di questo capitolo mostra è che nella famiglia di miscele provata l'effetto residuo della forma è piccolo, non che sia nullo in generale.

### **Le regole di parità non reggono**

La Parte prima confronta quattordici modi di sciogliere un match finito cinque pari e conclude che nove superano quello in vigore. La domanda che conta non è se i numeri cambino, cambiano tutti, ma se cambi l'ordine, perché è l'ordine che una federazione usa per scegliere.

Due cose cambiano, e vanno tenute separate perché si misurano su match diversi. Le regole di spareggio si consultano sui match arrivati a cinque pari; il conteggio dei dieci del capitolo 14 si consulta sui match con i totali in parità, ed è lì che va misurato.

**La prima è l'ordine delle quattordici modalità.** Nel modello gaussiano vince il conteggio pesato seguito dalla freccia, che indovina 64,2 volte su cento. Basta che una freccia su cento sia fuori pattern perché passi in testa TOTALE+; con tre su cento vince CONCORDE, con cinque le frecce sotto il dieci, con dieci il conteggio dei dieci seguito dagli X. La graduatoria non ha un vertice stabile: chi scegliesse la regola migliore secondo il modello gaussiano sceglierebbe una regola diversa in ciascuno di quei mondi.

Il conteggio pesato mostra perché. È la regola costruita per estrarre il massimo dell'informazione sotto il modello standard, ed è la migliore finché quel modello è vero: passa da 64,16 a 53,45 con cinque frecce anomale su cento, e a 47,85 con dieci, cioè sotto il sorteggio. I pesi che la rendono ottimale diventano i pesi sbagliati, e una regola tarata su un modello è fragile esattamente quanto è efficiente.

**La seconda riguarda il conteggio dei dieci, ed è più grave.** Sotto la gaussiana indica l'arciere sbagliato 55,6 volte su cento, è il risultato che dà il titolo al capitolo 14 e che ho usato come argomento per mezzo volume. Con tre frecce anomale su cento indovina 52,4 volte su cento e ha già smesso di essere peggio di una moneta; con cinque vale 55,5, con dieci 63,9.

| modello               | freccia | TOTALE+ | cascata | conteggio dei dieci |
|-----------------------|---------|---------|---------|---------------------|
| gaussiano             | 56,94   | 63,06   | 59,93   | 44,45               |
| $\varepsilon = 3 \%$  | 57,09   | 59,80   | 58,84   | 52,44               |
| $\varepsilon = 5 \%$  | 56,51   | 58,62   | 58,77   | 55,54               |
| $\varepsilon = 10 \%$ | 56,08   | 54,92   | 56,75   | 63,85               |

(Tavola 21. Quanto ciascuna regola indica il più bravo, fra un arciere da 690 e uno da 680. Le prime tre sulla popolazione dei match finiti cinque pari, il conteggio dei dieci su quella dei totali in parità, che è la popolazione su cui il capitolo 14 lo esamina. Seicentomila match simulati per cella, alea appaiata fra i modelli.)

Il meccanismo è semplice una volta visto, e va nei due versi. Una freccia fuori pattern toglie molti punti al totale del match e un solo dieci al conteggio dei dieci: chi legge il totale legge anche il disastro, chi conta i dieci no. Il totale perde così otto punti passando dal modello gaussiano al dieci per cento di code; il conteggio dei dieci, sulla sua popolazione, ne guadagna diciannove.

È il meccanismo del capitolo 15 letto al contrario. Lì avevo mostrato che a totale pari avere più dieci significa aver concentrato il deficit su meno frecce, e che dunque il conteggio indica chi ha tirato peggio. Quel ragionamento vale finché il deficit si distribuisce come il modello gaussiano prevede. Se una minoranza di frecce viene da una componente larga, il deficit di chi ne ha subita una è concentrato per un'altra ragione, non perché abbia tirato disuguale, ma perché ha avuto un incidente, e allora avere più dieci torna a significare quello che l'intuizione dice: aver tirato meglio.

**L'antiinformatività del conteggio dei dieci è dunque una proprietà del modello, non del conteggio.**

Quanto sia netto il rovesciamento si legge in un numero solo: il coefficiente di concordanza fra la graduatoria delle quattordici regole nel mondo gaussiano e quella nel mondo contaminato. Vale 1 se le due graduatorie sono identiche, 0 se non hanno rapporto, e  $-1$  se una è l'esatto rovescio dell'altra. Passa da 0,76 con una freccia anomala su cento a  $-0,56$  con dieci: l'ordine non si degrada, si rovescia. Il risultato l'ho ripetuto con cinque semi indipendenti e con cinque coppie di arcieri.

Sopra quella soglia la regola che ho criticato per tutto il volume torna competitiva, e per una ragione che le va riconosciuta: è cieca al disastro, e in un mondo con disastri la cecità protegge.

Un'avvertenza sulla direzione, prima di leggere la tavola. Gli scenari contaminano la coppia d'élite non perché io creda che le frecce anomale abitino soprattutto lì: la contaminazione sta dove stanno le conclusioni. Un test di robustezza mette il difetto dove farebbe più male, e le conclusioni decisive di questa parte del volume si concentrano nel dominio d'élite, quindi è soprattutto quel dominio che andava avvelenato; provare le code soltanto sui campi minori sarebbe stato un esperimento con una posta inferiore per la domanda qui in esame.

Da questa scelta di metodo non va però dedotta una tesi sul mondo reale. Anzi, se dovessi formulare una previsione, scommetterei sull'opposto. La mia ipotesi è che una quota importante del tiro fuori pattern rifletta errori o perturbazioni di esecuzione e che l'abilità consista anche nella capacità di ridurne la frequenza: mi aspetto quindi che  $\varepsilon$  diminuisca salendo di livello, che la deviazione dal nucleo regolare non scompaia necessariamente nemmeno nell'élite, ma che sia mediamente più pronunciata nei livelli inferiori.

La compressione dei totali verso il limite di 720, invece, nasce dalla trasformazione degli impatti in un punteggio superiormente limitato e non costituisce, da sola, evidenza sulla forma della nuvola di dispersione. È una congettura, e la lascio qui apposta perché sia falsificabile. Se fosse confermata dai dati, la tavola andrebbe letta di conseguenza: scenari di contaminazione più severi per i campi minori e, per l'élite, un intorno più stretto del valore reale di  $\varepsilon$ , precisamente nella regione in cui le soglie di rottura di questo capitolo risultano più basse.

| coppia  | $\varepsilon = 3 \%$ | $\varepsilon = 5 \%$ | $\varepsilon = 10 \%$ |
|---------|----------------------|----------------------|-----------------------|
| 690-680 | 36                   | 49                   | 65                    |
| 690-685 | 33                   | 42                   | 58                    |
| 700-690 | 40                   | 44                   | 53                    |
| 680-660 | 26                   | 39                   | 73                    |
| 660-640 | 17                   | 29                   | 66                    |

*(Tavola 22. Quante coppie di regole si scambiano di posto passando dal mondo gaussiano a quello contaminato, per cinque coppie di arcieri. Fra quattordici regole le coppie possibili sono novantuno; qui si contano quelle che entrambe le corse distinguono con sicurezza.)*

La coppia d'élite è la più fragile, si rompe già con tre frecce anomale su cento, perché a quei livelli i totali sono compressi contro il massimo del round e una freccia anomala pesa relativamente di più. Le coppie di livello medio reggono più a lungo.

### Le strutture di torneo reggono

La Parte seconda conclude che il piazzamento conta più della struttura del tabellone. Qui la robustezza è netta.

| modello              | standard | riseeding | sorteggio | appaiati |
|----------------------|----------|-----------|-----------|----------|
| gaussiano            | 15,55    | 17,82     | 14,69     | 14,51    |
| $\varepsilon = 1 \%$ | 15,81    | 17,48     | 14,48     | 14,01    |
| $\varepsilon = 3 \%$ | 15,14    | 16,98     | 13,18     | 13,64    |
| $\varepsilon = 5 \%$ | 14,59    | 16,38     | 13,13     | 13,32    |

*(Tavola 23. Quanto spesso il migliore del campo vince il torneo, su cento tornei, per ciascuna struttura. Campo di 64 arcieri, dodicimila tornei per cella, alea appaiata.)*

Si legga per colonne. Tutte le strutture calano insieme di circa un punto e mezzo passando dal mondo gaussiano a quello con cinque frecce anomale su cento, e le distanze fra loro restano le stesse. Il riseeding resta primo in ogni riga, e non c'è una sola coppia di strutture che si scambi di posto.

Il motivo, nella griglia provata, è che un torneo aggrega sessantatré match e l'aggregazione livella le differenze di forma: ogni struttura perde potere discriminante nella stessa misura, perché la perdita viene dal rumore in più nei singoli match. È una regolarità osservata su questa griglia di contaminazioni, non un teorema per ogni deviazione dalla gaussiana.

La conclusione della Parte seconda è quindi la più solida delle tre. Se si dovesse scegliere un solo intervento senza sapere com'è fatto il mondo, quello sul piazzamento è l'unico che non cambia segno.

## La stagione è la parte fragile

Qui ho fatto la prova più severa che sapevo fare. Invece di cambiare insieme il modo di generare i dati e quello di analizzarli, che risponderebbe alla domanda facile, se lo strumento giusto funzioni sui dati giusti, ho generato stagioni dal mondo contaminato e le ho analizzate con la procedura del capitolo 46, quella che il volume usa davvero e che continua a credere agli errori gaussiani.

| modello                           | identifica il migliore | errore della stima | perdita               |
|-----------------------------------|------------------------|--------------------|-----------------------|
| gaussiano                         | 62,20 %                | 2,111              | —                     |
| $\varepsilon = 1 \%$              | 62,00 %                | 2,246              | trascurabile          |
| $\varepsilon = 3 \%$              | 56,93 %                | 2,485              | come 2,0 gare in meno |
| $\varepsilon = 5 \%$              | 55,97 %                | 2,638              | come 2,3 gare in meno |
| $\varepsilon = 5 \%$ , gravità 16 | 50,93 %                | 3,083              | come 4,2 gare in meno |
| $\varepsilon = 10 \%$             | 49,43 %                | 3,431              | come 4,7 gare in meno |

*(Tavola 24. La procedura del capitolo 46 applicata a stagioni generate dal modello contaminato. Trentadue arcieri, dieci gare, sei presenze a testa, tremila stagioni per riga.)*

La perdita è espressa in gare perché quella è l'unità in cui si scrive un calendario. Ho misurato quanto costa accorciare la stagione: passando da dieci gare a nove, la probabilità di identificare il migliore cala di 2,7 punti. Dire che tre frecce anomale su cento costano quanto togliere due gare è più utile che dire che costano cinque punti percentuali; una federazione sa che cosa significhi togliere due gare dal calendario, e non sa che cosa significhi perdere cinque punti.

## Quattro rimedi, tre falliti

La reazione naturale è cambiare il modo di stimare. La procedura del capitolo 46 cerca i valori che rendono minimo lo scarto complessivo, e un criterio così è notoriamente sensibile ai valori estremi: basterebbe un criterio che pesi meno gli scarti grandi.

L'ho implementato; la perdita di Huber, con la soglia al valore classico e la scala degli scarti misurata in modo che non sia essa stessa gonfiata da ciò che si vuole attenuare.

**Non funziona.** L'errore della stima robusta è peggiore di quello del metodo ordinario in tutte le configurazioni, comprese quelle contaminate.

La ragione è che le code non producono round anomali: producono round leggermente peggiori. Cinque frecce fuori pattern su settantadue spostano il totale di qualche punto, non di venti, e lo scarto di un round contaminato sta dentro la distribuzione ordinaria degli scarti. Uno stimatore robusto non ha niente da riconoscere: cerca il round mostruoso e non lo trova, perché sommare settantadue frecce ha già diluito l'incidente dentro il totale. L'aggregazione ha già fatto il lavoro che la robustezza avrebbe fatto, e quello che resta non è una coda pesante negli scarti, è oscillazione in più. Serve più informazione, non un calcolo più furbo.

Questo rende la Parte terza più solida, non meno, ma va detto con precisione che cosa la prova mostra. Mostra che questa perdita robusta, su queste configurazioni, non recupera niente: il confronto vale circa un decimo di punto di errore, sempre a sfavore della robusta e sempre a più di tre errori standard. Non mostra che nessuno stimatore possa recuperare, che è un enunciato diverso e più forte di quanto una

prova su una sola famiglia di perdite possa sostenere. La spiegazione che ho dato rende plausibile la versione forte, ma plausibile non è dimostrato.

### L'asimmetria che non viene dal tiro

Un'ultima cosa, perché è facile confondere due fenomeni diversi.

I punteggi degli arcieri d'élite sono asimmetrici: rispetto alla propria media si può perdere molto e guadagnare poco. È tentante leggerlo come prova che gli impatti non sono gaussiani. Non lo è.

Il round da settantadue frecce ha un massimo, e chi tira vicino a quel massimo ha poco spazio sopra e molto sotto. Con impatti perfettamente gaussiani e nessuna coda, l'asimmetria del punteggio cresce col livello: vale  $-0,073$  a 600 punti,  $-0,083$  a 690 e  $-0,257$  a 710. A quel livello restano tre deviazioni standard sopra la media e infinite sotto: un arciere da 710 non può fare più di 720, ma può benissimo fare 690.

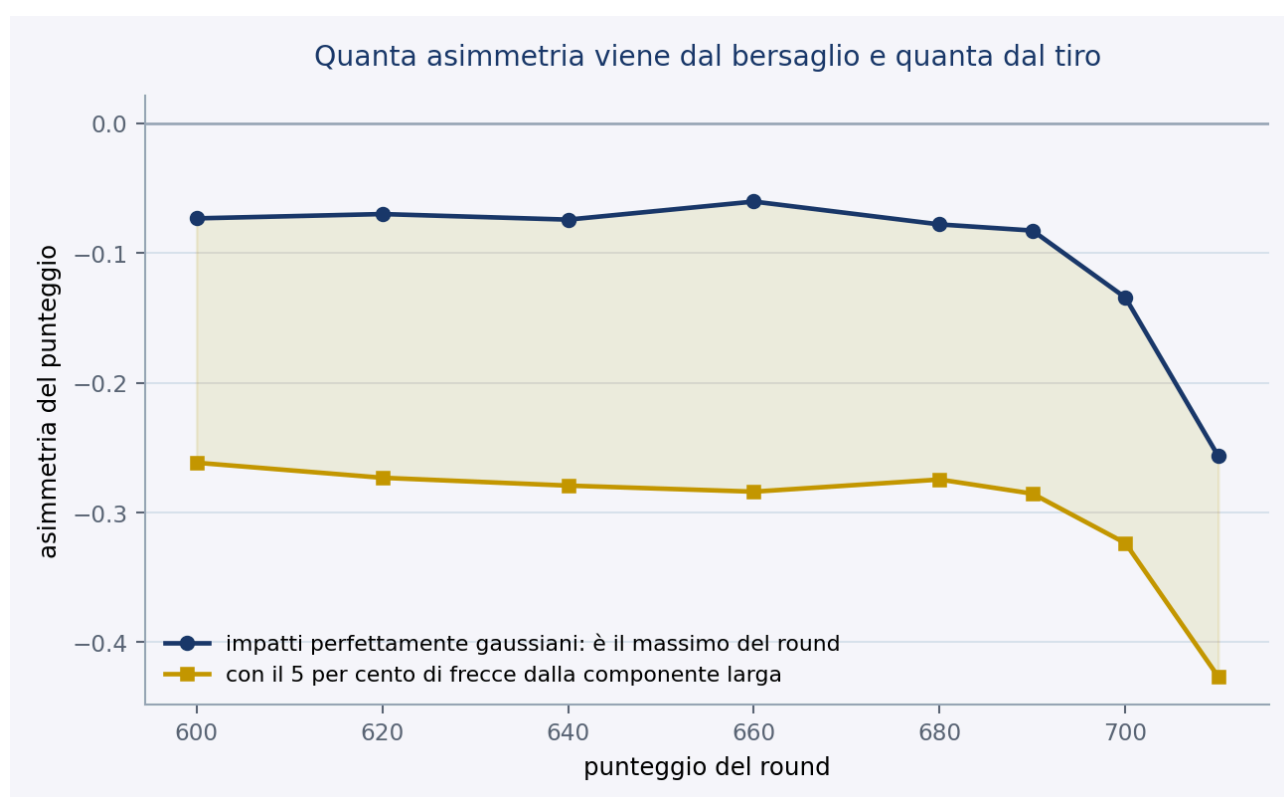


**Figura 22 — L'asimmetria che viene dal bersaglio e quella che viene dal tiro.** Asimmetria del punteggio di round in funzione del livello, su tutti i round di 72 frecce di un arciere, con impatti gaussiani e con code. La parte che cresce col livello viene dal massimo del round; quella piatta, dalle code.

Le code aggiungono un contributo di circa due decimi, ma quasi costante a tutti i livelli. La separazione è quindi netta: l'asimmetria che cresce col livello viene dal bersaglio, quella piatta dal tiro. A 690 punti il massimo del round spiega meno di un terzo dell'asimmetria osservata; a 710 ne spiega il sessanta per cento. Chi osservasse un arciere d'élite e attribuisse tutta la sua asimmetria alle code sbaglierebbe, e sbaglierebbe di più quanto più l'arciere è bravo.

### Una regola che il modello non implementa

Va dichiarata una divergenza dal regolamento, che prevede che se nello spareggio entrambi mancano l'area di punteggio si tiri di nuovo; il modello invece assegna sempre la vittoria alla freccia più vicina al centro, e quel caso non lo distingue.

Quanto costi saperlo si calcola esattamente. Su tutto il dominio del volume il valore più grande è 3,3 su diecimila, cioè un doppio miss ogni tremila spareggi, e si presenta al livello più basso con le code più pesanti. Nel dominio d'élite, dove si giocano le conclusioni, scende a due su centomila milioni con il dieci per cento di code, cioè una volta ogni cinquanta miliardi di spareggi, e a 6 su 10 elevato alla ottantaduesima con impatti gaussiani, che è un modo formale di dire mai.

Implementare la ripetizione non cambierebbe una cifra di questo capitolo. Tacere la divergenza sarebbe stato peggio che dichiararla.

### Che cosa portarsi a casa

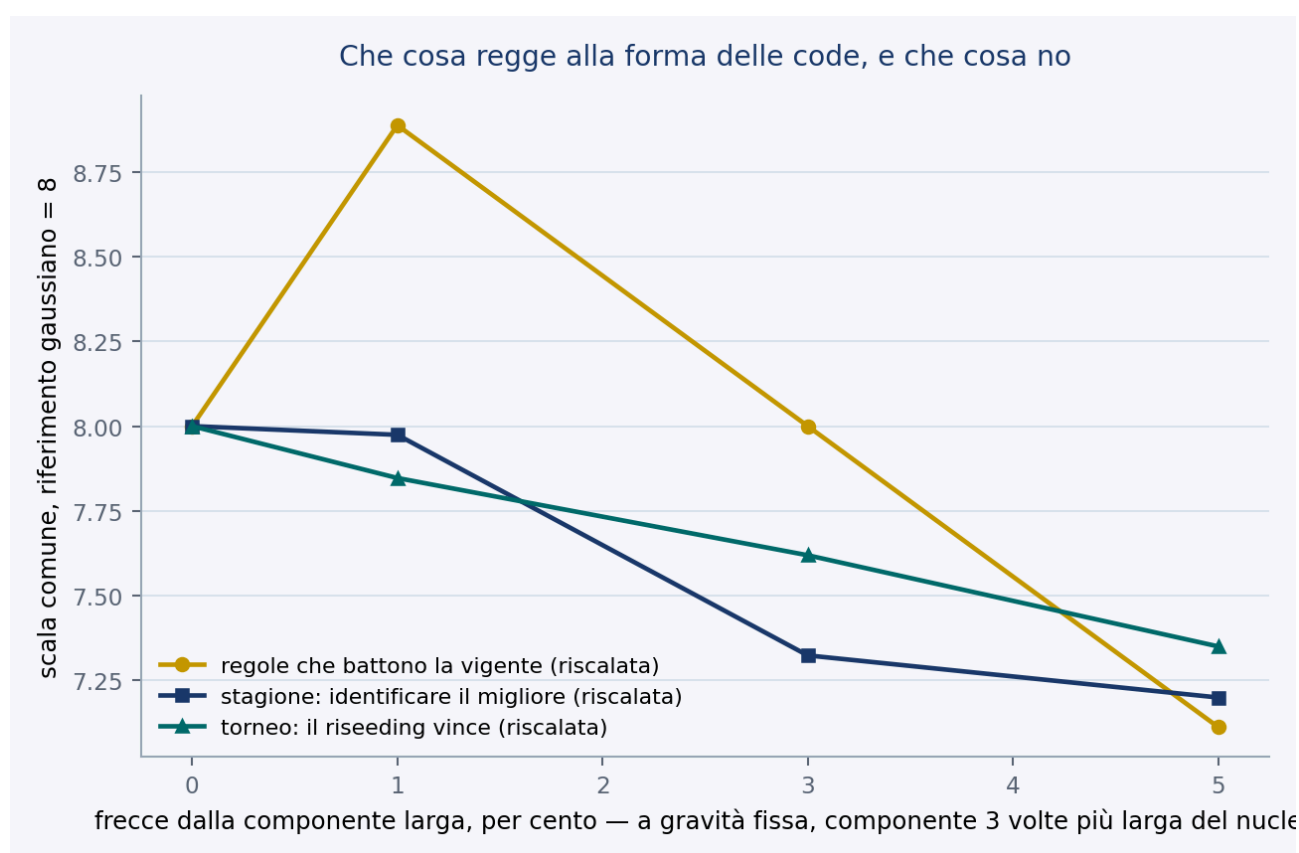


**Figura 23 — Che cosa regge e che cosa cade.** Esito di ciascuna famiglia di conclusioni del volume al crescere della frazione di frecce fuori pattern: i tornei reggono, le regole di parità si riordinano, la graduatoria stagionale perde.

Il quadro non è uniforme, ed è questo che lo rende utile.

**Le conclusioni sui match e sui tornei reggono.** Quello che quei formati misurano è soprattutto la dispersione del punteggio, e appaiate media e oscillazione l'effetto residuo della forma è di pochi decimi di punto. Non vale in generale, il controesempio di poche pagine fa lo esclude, ma vale qui.

**L'ordine delle quattordici modalità non regge nemmeno a una freccia anomala su cento:** già lì cambia la regola in testa e si scambiano di posto undici coppie di regole. Con tre su cento la con-

cordanza con l'ordine gaussiano è a zero, con dieci vale  $-0,56$ .

**La raccomandazione della Parte prima sopravvive, ma in forma più debole** di quella che avevo scritto. Non è vero che leggere il match batta sempre la sola freccia: con dieci frecce anomale su cento TOTALE+ scende a 54,92 e passa sotto lo spareggio, che vale 56,08. Restano però sei modalità che leggono il match e battono ancora la freccia, le due cascate, CONCORDE, dieci poi X, le frecce sotto il dieci, il conteggio degli X. Cambia quali, non se.

**Il risultato che dà il titolo al capitolo 14 cade prima di tutti**, e cade di parecchio: il conteggio dei dieci passa da 44,45 a 63,85, cioè diciannove punti. Smette di essere peggio di una moneta già con tre frecce anomale su cento, e con dieci indica il più bravo quasi due volte su tre. Chi volesse difendere la regola tradizionale ha qui il suo argomento, ed è onesto consegnarglielo: vale a partire dal tre per cento di frecce fuori pattern, il primo punto della griglia in cui il conteggio dei dieci supera la monetina, e nessuno sa se le frecce fuori pattern siano così tante.

**Le conclusioni sulla classifica stagionale sono le più fragili**, e qui frequenza e gravità vanno tenute distinte. Con frecce anomale tre volte più larghe del nucleo, il tre per cento costa quanto due gare di calendario e il cinque per cento poco più di due. Se le anomale sono quattro volte più larghe, il cinque per cento costa già quanto quattro gare. La perdita robusta che ho provato non ne recupera niente; se qualche altro stimatore possa farlo, questo capitolo non lo dice.

**Resta una domanda che questo capitolo non può chiudere, ed è la più importante: quanto vale e davvero.** Lo strumento per stimarlo dai referti è scritto e collaudato su una griglia che incrocia frequenza, gravità e numerosità: se la gravità viene stimata insieme alla frequenza, ritrova il valore vero entro 1,13 punti percentuali; ma se la si fissa a un valore sbagliato la frequenza viene sopravvalutata fino a 2,27 punti — e della gravità, sugli arcieri reali, questo lavoro non ha una stima empirica. Dagli anelli soli le due si confondono fra loro: servirebbero le coordinate degli impatti.

Finché quel numero non c'è, il modo onesto di leggere questo volume è con la soglia accanto a ciascuna conclusione. E la soglia più bassa, quella che dovrebbe preoccupare per prima, è quella della classifica stagionale: 2,0 % se le frecce anomale sono larghe il triplo del nucleo, 1,2 % se sono larghe il quadruplo. Non è il tre per cento, quello è soltanto il primo punto della griglia in cui l'effetto si vede.

---

## 54. Il filo delle tre parti

Le tre parti non sono tre studi accostati: sono concatenate, e il modo in cui si legano è il risultato che non avevo previsto cominciando.

La Parte terza produce una graduatoria di stagione. Quella graduatoria, usata come piazzamento, è l'ingrediente che nella Parte seconda vale più di ogni altro singolo intervento: quattro titoli in più su cento edizioni, a costo zero. E la Parte seconda contiene la Parte prima, perché la regola di parità è l'ultimo passo di una gara e non un problema a sé. La graduatoria smette quindi di essere un premio di dicembre e diventa la cosa che decide chi incontra chi a maggio.

Ogni parte, poi, fa la stessa cosa: fissa un tetto e misura la distanza fra la prassi e quel tetto. Per la parità il tetto è un arciere e sei decimi spostati per gara, e la regola in vigore ne sposta tre quarti mentre TOTALE+ ne sposta uno e quattro. Per la gara il tetto è quindici titoli su cento, e la gara attuale ne consegna

dieci. Per la graduatoria il tetto è tre quarti con calendari completi, e i calendari parziali la portano a tre quinti.

E ogni volta lo spazio si comprime, e la decisione passa a un asse che non era in discussione.

| Su che cosa si discuteva                         | Quanto separa il migliore dal peggiore | Che cosa decide invece        |
|--------------------------------------------------|----------------------------------------|-------------------------------|
| le regole di parità che battono quella in vigore | 0,4 punti fra la vigente e il meglio   | tempo, spettacolo, robustezza |
| i formati del match                              | un terzo di match su 63                | quale spettacolo si voglia    |
| le strutture di tabellone                        | 1,7 punti su una base di dieci         | da dove viene il piazzamento  |
| i metodi di graduatoria, a calendari completi    | tre decimi di punto                    | chi tira insieme a chi        |

*(Riepilogo delle quattro comparse. Le unità delle righe sono diverse fra loro — match, punti di probabilità — e non vanno confrontate in verticale: la colonna dice quanto sia stretto ciascuno spazio, non quale sia il più stretto.)*

Vale la pena mettere i tre tetti in una forma sola, perché è il modo più rapido di vedere che cosa questo lavoro abbia stabilito.

| Che cosa si decide                                | Dove sta oggi                        | Dove si può arrivare                             | Quanto resta fuori portata                   |
|---------------------------------------------------|--------------------------------------|--------------------------------------------------|----------------------------------------------|
| <b>Una parità — chi vince un match pari</b>       | 57 volte su 100                      | 63 con TOTALE+, 69 misurando le distanze         | il 31 % dei casi resta indecidibile          |
| <b>Una gara — chi vince il torneo</b>             | 9,7 volte su 100                     | 13,7 senza costi, 14,2 col modello di precisione | 86 volte su 100 il titolo non va al migliore |
| <b>Una stagione — chi è il migliore dell'anno</b> | 59 volte su 100 a calendari parziali | 75 a calendari completi                          | un quarto delle stagioni resta ambiguo       |

*(Le tre righe misurano cose diverse e non vanno confrontate in verticale: la prima è la probabilità di indicare il più bravo fra due arcieri in parità, la seconda che il titolo vada al migliore di un campo di 64, la terza che il primo di una graduatoria stagionale sia davvero il migliore di 32.)*

Letta per righe, la tabella dice quanto ciascun problema sia risolvibile. Letta per colonne dice qualcosa di più scomodo: la terza colonna è sempre molto lontana dalla quarta.

Anche facendo tutto il possibile, un torneo continua a non premiare il migliore in cinque casi su sei. E questo non perché il regolamento sia scritto male, ma perché un tabellone a eliminazione diretta fa una cosa diversa da quella che gli si chiede. Chi voglia che il titolo vada al migliore deve guardare la terza riga, non la seconda.

Quattro volte lo stesso esito, dunque: il margine si comprime e la cosa che decide non è quella che si stava discutendo. Detta così non è una conclusione sui numeri ma una descrizione di come è fatto il problema, e ha una conseguenza pratica per chi debba intervenire: **il posto dove intervenire non è quasi mai quello di cui si discute**. Non la regola di parità ma il tabellone; non il formato ma quale spettacolo si voglia; non il metodo di calcolo ma chi tira insieme a chi.

---

## 55. Che cosa il modello non vede

Le tre parti di questo volume hanno un modello solo, e quel modello ignora tre cose reali. Ogni tiro vi nasce isolato dagli altri, estratto sempre dalla stessa nuvola immutabile: non c'è tensione che monta, non c'è braccio che si stanca, non c'è vento che gira, non c'è la posta che pesa.

Il limite morde in tutte e tre le parti, ma non allo stesso modo. Nella prima morde più che altrove, perché la freccia di spareggio non è una freccia qualsiasi, è la più carica che esista in questo sport, tirata sapendo che dopo non viene più niente. Nella seconda entra dalla qualificazione, che è un pomeriggio intero durante il quale il vento può cambiare tre volte. Nella terza entra in ogni gara della stagione, ed è l'unico punto in cui il modello lo tratta esplicitamente.

Dichiarare un limite però serve a poco. Quello che serve è dire in che direzione spinga ciascuna conclusione, e dove si può, di quanto.

### La pressione

Che la freccia di spareggio sia distribuita come le quindici che l'hanno preceduta è un'ipotesi forte, e non è una questione aperta: chi ha confrontato le frecce di spareggio con quelle del match su gare vere ha trovato un calo di rendimento proprio lì. L'accentuazione nelle competizioni che contano di più è documentata, ma non per tutti allo stesso modo; nello studio che la riporta l'effetto è più marcato fra le donne, e altri lavori trovano differenze per sesso e per esperienza. Prendo quindi il calo come fatto e la sua misura come incerta, ed è la ragione per cui nel seguito lo faccio variare invece di fissarlo.

La direzione in cui quel calo spinge dipende però da come la pressione agisca, e i tre casi possibili danno tre risposte diverse.

**Se la pressione allargasse entrambe le nuvole in proporzione**, poniamo del venti per cento ciascuna, non cambierebbe assolutamente nulla. La ragione è la formula del capitolo 7: lo spareggio dipende soltanto dal rapporto fra le due ampiezze, e moltiplicandole per lo stesso fattore quel rapporto resta identico. È un caso che si esclude per costruzione, non per misura.

**Se invece la pressione aggiunge a ciascuno una quota di errore propria**, come è più naturale pensare, allora le due nuvole si assomigliano più di quanto si assomiglino gli arcieri, e ogni regola diventa meno informativa. È il caso che ho misurato, e conviene esprimerlo in una valuta comprensibile: diciamo che sotto pressione un arciere si comporta come uno più debole di un certo numero di punti. Un calo di 10 punti significa che il migliore dei due, al momento di tirare, vale 680 invece di 690.

| Calo dovuto alla pressione | I due arcieri diventano | Freccia di spareggio | TOTALE+ |
|----------------------------|-------------------------|----------------------|---------|
| nessuno                    | 690 e 680               | 57,0                 | 62,95   |
| 5 punti                    | 685 e 675               | 56,6                 | 62,8    |
| 10 punti                   | 680 e 670               | 56,15                | 62,66   |
| 20 punti                   | 670 e 660               | 55,5                 | 62,4    |
| 30 punti                   | 660 e 650               | 54,9                 | 62,24   |
| 50 punti                   | 640 e 630               | 54,1                 | 61,96   |

*(Coppia di riferimento, popolazione dei match finiti cinque pari. Calcolo esatto: il calo si applica soltanto alla freccia di spareggio e non alle 15 del match, perché è quella la freccia sotto pressione. La riga senza calo coincide con i valori del capitolo 17, come deve.)*

Il risultato è più solido di quanto ci si aspetterebbe. Un calo di 30 punti è enorme, è la distanza che separa un finalista internazionale da un buon arciere regionale, e costa alla freccia di spareggio due punti scarsi, portandola da 57,0 a 54,9. Perché la freccia scenda a 52, cioè quasi al livello di una monetina, servirebbe un calo di 130 punti: vorrebbe dire che un arciere da 690 tira, in quel momento, come uno da 560, cioè come un principiante. Nei lavori sul tiro con l'arco che ho consultato non si trovano cali di quell'ordine.

E c'è un secondo fatto, che è quello utile alla decisione: **nel modello, un calo di rendimento sulla sola freccia di spareggio allarga il vantaggio di TOTALE+ invece di ridurlo**. Senza pressione lo scarto fra le due regole è di 5,9 volte su cento; con un calo di 30 punti sale a 7,4.

La ragione è semplice: il totale del match è già stato scritto e la pressione non lo tocca più, mentre la freccia di spareggio va ancora tirata. Più la pressione morde, più conviene leggere il referto invece di aggiungere un tiro.

Resta un terzo caso, e non è misurabile. Se esistesse davvero una qualità separata, la freddezza, il colpo sotto pressione, allora lo spareggio la misurerebbe, e misurerebbe bene qualcosa di diverso dalla precisione: non uno strumento impreciso, ma uno strumento puntato altrove. Il modello non può distinguere questo caso dal precedente, e chiunque affermi il primo ha l'onere di misurarlo.

## Il vento

Qui la direzione è la stessa, ma l'entità è molto maggiore, e questo è il fattore che il volume sottovaluta di più.

Il vento aggiunge a entrambi gli arcieri un errore che non dipende da quanto siano bravi: soffia allo stesso modo sulle frecce dell'uno e su quelle dell'altro. La conseguenza è che le due nuvole si allargano di una quantità comune, e quindi si assomigliano.

| Ampiezza dell'errore da vento | Nuvola del migliore | Nuvola del peggiore | Rapporto | Freccia di spareggio |
|-------------------------------|---------------------|---------------------|----------|----------------------|
| nessuno                       | 4,46 cm             | 5,14 cm             | 1,152    | 57,0                 |
| 2 cm                          | 4,89                | 5,51                | 1,128    | 55,88                |
| 4 cm                          | 5,99                | 6,51                | 1,087    | 54,1                 |
| 6 cm                          | 7,48                | 7,90                | 1,056    | 52,7                 |
| 10 cm                         | 10,95               | 11,24               | 1,027    | 51,3                 |

*(Coppia di riferimento. L'errore da vento è modellato come una componente casuale comune che si somma a quella di ciascun arciere.)*

La colonna delle nuvole dice tutto da sola. Con sei centimetri di errore da vento il gruppo del finalista da 690 diventa più largo di quello che il quartifinalista da 680 aveva in giornata calma: 7,48 contro 5,14. Le due nuvole restano diverse, ma il loro rapporto scende da 1,152 a 1,056; cioè i due arcieri, sul bersaglio, diventano quasi indistinguibili.

E infatti lo spareggio scende da 57,0 a 52,7, cioè perde più della metà del suo margine sulla monetina. Sei centimetri sono una giornata ventosa ma non eccezionale.

L'effetto colpisce tutte le regole insieme, non soltanto la freccia.

| Ampiezza dell'errore da vento | Probabilità di cinque pari | Freccia di spareggio | TOTALE+ | Divario |
|-------------------------------|----------------------------|----------------------|---------|---------|
| nessuno                       | 16,8 %                     | 57,01                | 62,95   | 5,94    |
| 3 cm                          | 17,5 %                     | 55,05                | 59,85   | 4,80    |
| 6 cm                          | 14,3 %                     | 52,74                | 55,73   | 2,98    |
| 10 cm                         | 14,9 %                     | 51,32                | 52,90   | 1,57    |

*(Coppia di riferimento, popolazione dei match finiti cinque pari. Calcolo esatto: il vento aggiunge a entrambi un errore comune, quindi si allargano tutte e sedici le frecce e non solo la sedicesima.)*

In una giornata di vento forte tutte le regole convergono verso la monetina, e la scelta fra loro smette di contare. TOTALE+ conserva un vantaggio su tutta la scala, ma quel vantaggio si assottiglia da sei volte su cento a una e mezza.

Ne segue una conseguenza pratica che nessuno formula mai e che vale la pena scrivere: è la giornata di vento, non la regola di parità, a decidere quanto una gara sia una lotteria. Nello scenario di rumore da vento qui modellato, l'effetto è molto più grande di quello di qualunque regola di parità. Il modello però aggiunge una componente casuale comune e non simula meteorologia, sedi o orari: se una scelta di sede riduca davvero quel rumore, e di quanto, questo lavoro non lo misura.

### La posta in gioco

È il fattore su cui ho meno da dire, ed è onesto dirlo.

Non c'è un modo di metterla nel modello che non sia arbitrario. Che la posta agisca attraverso la pressione è una lettura plausibile e non una relazione che io misuri: quello che ho misurato è come cambiano i numeri se la freccia di spareggio è tirata peggio delle altre, non perché lo sia.

Quello che si può osservare è che le due cose non sono indipendenti: una finale per l'oro ha insieme la posta massima e, tipicamente, i due arcieri più vicini fra loro. Ed è la combinazione peggiore, perché il primo fattore erode la precisione e il secondo riduce il divario da misurare.

Con la coppia più stretta che una finale possa presentare, due arcieri a 5 punti di distanza, la freccia di spareggio scende a 53,6 su cento, e con un calo da pressione di 20 punti arriva a 52,9. Il che dice una cosa che vale la pena tenere: lo spareggio è meno affidabile proprio nel momento in cui viene usato per decidere di più.

---

## 56. Che cosa cambia in ciascuna conclusione

Dichiarare che i tre fattori spingono verso il basso non basta, perché il volume non contiene un numero solo: contiene conclusioni diverse, e conviene chiedersi per ciascuna in che direzione venga spostata. Ho fatto il conto prendendo il vento come fattore di riferimento, perché è il più severo dei tre.

| Conclusione del volume                         | Senza vento | Con 6 cm di vento | Effetto |
|------------------------------------------------|-------------|-------------------|---------|
| Il titolo va al migliore                       | 9,7 %       | scende            | aggrava |
| Il piazzamento sbaglia di 8,6 posti            | 8,6         | sale              | aggrava |
| La regola in vigore indica il più bravo        | 57,01       | 52,74             | aggrava |
| Il margine disponibile per una regola migliore | 5,94 punti  | 2,98              | aggrava |
| Il conteggio dei dieci è anti-informativo      | 44,66       | 47,27             | attenua |
| Quanto spesso il totale pareggia               | 34,4 %      | 19,5 %            | attenua |

(Coppia di riferimento, 2 milioni di match per riga. «Aggrava» significa che il fattore escluso rende la conclusione più marcata di come qui riportata, «attenua» il contrario.)

Quattro conclusioni su sei vengono aggravate, e sono quelle portanti. La regola in vigore, che in giornata calma indica il più bravo 57 volte su cento, nello scenario con sei centimetri di errore da vento ne indica 53. Il numero che il volume riporta è quindi quello del caso senza vento, e ogni scenario di vento provato lo abbassa; che cosa faccia il campo reale, queste simulazioni non lo osservano. Lo stesso vale per il titolo che va al migliore e per l'errore del piazzamento.

Le due eccezioni vanno però dette con la stessa chiarezza, perché sono le sole in cui un fattore escluso lavora contro una mia conclusione.

**La prima è il ribaltamento del conteggio dei dieci:** il vento lo attenua, da 44,66 a 47,27. Il meccanismo è quello del capitolo 20, il vento aggiunge dispersione a tutte le frecce, il che allenta il legame fra dieci e frecce basse. Non lo annulla, perché il criterio resta sotto la monetina anche con dieci centimetri di vento, dove vale 48,84. Ma chi volesse difendere il conteggio dei dieci ha qui il suo unico argomento, ed è onesto consegnarglielo.

**La seconda è più tecnica:** con il vento il totale pareggia meno spesso, dal 34 al 20 per cento, perché le due somme si allontanano. Il che riduce il numero di casi in cui TOTALE+ deve ricorrere alla freccia, e quindi, paradossalmente, in una giornata ventosa TOTALE+ si comporta più spesso come regola pura e meno come regola a due stadi.

Il verso resta però quello sotto tutti e tre i fattori di questo capitolo: TOTALE+ sta davanti alla freccia di spareggio con vento e senza, sotto pressione e senza, e sotto pressione le sta davanti di più. Quello che il vento erode è il margine assoluto, da 5,94 punti a 2,98 con sei centimetri di vento, e a 1,57 con dieci, non l'ordine.

Vale la pena essere precisi su che cosa questo significhi e che cosa non significhi. Pressione, vento e posta in gioco non ribaltano il confronto fra le due regole. La coda marcata del capitolo 21 sì, e resta l'unica configurazione provata in tutto questo lavoro in cui TOTALE+ finisce sotto la regola in vigore. I tre fattori di questo capitolo non aggiungono quindi un rischio nuovo alla proposta: aggravano il problema comune a tutte le regole, che è quanto poco margine ci sia in tutto.

## 57. I limiti

Prima delle raccomandazioni, i confini di questo lavoro. Sono cinque, e il quinto è quello promesso al primo capitolo.

**Il primo è che la bravura, qui, è una quantità sola:** il punteggio atteso, e nient'altro. Un atleta però è anche altro, la capacità di reggere su una giornata intera, il modo in cui legge il vento, la prontezza nel rimettersi in piedi dopo un turno andato male. Nessuna di queste cose entra nei calcoli. Se si crede che un torneo debba premiare quelle qualità, questo lavoro misura la cosa sbagliata, e la risposta corretta non è correggerne i numeri ma respingerne la definizione.

**Il secondo è che ciò che cambio è l'aritmetica e non il comportamento.** Nel confrontare due regole tengo ferme le frecce già tirate e sostituisco soltanto il criterio con cui le si legge; ma con un regolamento diverso è verosimile che gli atleti si comporterebbero in un altro modo, e forse già dal primo set. È il limite di ogni ragionamento controfattuale su un regolamento sportivo, e l'unica cosa che lo scioglierebbe è una gara realmente disputata sotto la regola nuova, che per definizione non c'è ancora.

**Il terzo riguarda quanto si possa concludere da un incontro solo,** e la risposta è: quasi nulla. Chi fosse il più forte quel pomeriggio non è una grandezza osservabile. Accumulando gare in numero sufficiente si arriva a stimare quale dei due raggruppamenti più stretto e con che margine d'errore; sulla singola giornata, invece, la certezza non si raggiunge, e nessuna conclusione di queste pagine va letta come un giudizio su una gara realmente avvenuta.

**Il quarto è che la forma della coda non è stimata.** Ho misurato quanto varrebbe conoscerla e la ragione strutturale per cui probabilmente non si saprà, ma non equivale a saperlo: è il numero che decide se TOTALE+ sia la scelta giusta o se serva CONCORDE, ed è aperto.

**Il quinto, ed è quello annunciato all'inizio, è che l'effetto dell'ordine di tiro allo spareggio non è misurabile con i dati che esistono oggi.** Il regolamento stabilisce che allo spareggio parte chi ha tirato per primo nel primo set, e l'ordine di quel set lo sceglie il meglio classificato della qualificazione, che può scegliere di tirare per secondo, e talvolta lo fa. Se un vantaggio dell'ordine esistesse si confonderebbe con un vantaggio di bravura, perché nei dati disponibili le due cose compaiono quasi sempre insieme.

E ha una conseguenza pratica che costa niente: basterebbe sorteggiare l'ordine di tiro allo spareggio perché la domanda diventasse rispondibile con i dati che una federazione già raccoglie. È il quarto comma dell'articolo del capitolo 18, l'unica modifica di queste pagine che non serva a migliorare una regola ma a rendere misurabile una cosa che oggi non lo è.

## **Il verso dei limiti**

C'è una conseguenza generale che vale la pena ripetere, perché è il modo giusto di leggere l'intero volume.

Quattro dei cinque limiti appena elencati, tutti tranne il terzo, che riguarda che cosa si possa osservare e non quanto si distingua, hanno lo stesso verso. La bravura ridotta a una quantità sola, il comportamento tenuto fermo, la coda non stimata e l'ordine di tiro non misurabile spingono ciascuno verso una capacità di distinguere minore, non maggiore. Lo stesso vale per i tre fattori del capitolo 55, con l'unica eccezione lì dichiarata.

I numeri di queste pagine sono dunque un riferimento sotto le assunzioni dichiarate, e per le deviazioni che ho modellato le misure si riducono; quanto, dipende dal vento della giornata più che da qualunque altra cosa. Che il campo reale stia sotto è un'estrapolazione, non una misura.

Il che ha una conseguenza sul modo di contestare questo lavoro, e la scrivo perché mi sembra la cosa più utile da lasciare a chi voglia farlo. **Chi obietta che il modello è troppo semplice avrebbe ragione, e obbietterebbe a favore delle mie conclusioni.** Il margine che qui misuro è il massimo disponibile sotto le assunzioni dichiarate, e le complicazioni che ho provato lo riducono.

Non ho provato che lo riduca qualunque complicazione: qualità correlate con la bravura, leggere il vento, reggere la pressione, potrebbero anche aumentare la separazione fra arcieri. Chi volesse smentire questo volume dovrebbe quindi mostrare non che ho trascurato qualcosa (l'ho fatto, e l'ho elencato) ma che qualcosa di trascurato spinge nella direzione opposta. È una cosa che si può fare, ed è un lavoro diverso da quello di elencare le semplificazioni.

C'è infine un limite che riguarda l'apparato di verifica e non il contenuto. Tutti i controlli descritti nella nota sul metodo verificano i calcoli, e nessuno verifica il significato attribuito a un numero: in una revisione finale sui file mai riletti sono emersi quattro difetti di quel tipo, tutti numeri esatti con l'etichetta sbagliata. Chi legga l'apparato senza questo limite ne inferisce una garanzia che non c'è.

---

## 58. Come un risultato è scaduto senza diventare falso

Vale la pena raccontare per esteso il caso più istruttivo di questo lavoro, perché è un modo di sbagliare che nessun controllo intercetta e che in un progetto lungo è probabilmente il più comune.

La Parte prima è stata chiusa e archiviata prima che la Parte seconda esistesse, e fra i suoi risultati c'era il confronto fra il peso della struttura del tabellone e quello della regola di parità: il rapporto misurato era di circa cinque a uno.

Quel confronto misurava però l'escursione fra strutture assumendo che il piazzamento fosse l'ordine vero, e nessuno lo aveva scritto come ipotesi, perché a nessuno era venuto in mente che potesse essere falsa. Il tabellone si costruisce sui piazzamenti, e i piazzamenti sono quello che sono.

Poi la Parte seconda ha misurato quanto siano affidabili, e la risposta è stata che in un campo compatto la qualificazione non indovina praticamente mai gli otto migliori. E le strutture che guadagnavano di più erano proprio quelle che consultano il piazzamento più volte: con l'ordine reale l'escursione fra strutture scende da sei punti e mezzo a meno di due.

Che la struttura conti più della regola resta vero, e su due derivazioni indipendenti di cui la seconda non passa dal piazzamento. Quello che cambia è il fattore: **il tabellone è la cosa che conta di più, e conta poco anche lui.**

E c'è una quantità nuova che il numero vecchio conteneva senza saperlo. La differenza fra i due è il prezzo dell'ignoranza sul piazzamento: rispetto al tabellone standard il riseeding varrebbe cinque punti se sapessimo davvero chi è più bravo, e ne vale uno e uno con la qualificazione che abbiamo. Quei quattro punti tondi sono ciò che si compra migliorando il piazzamento. Detto altrimenti: **quattro titoli su cento edizioni sono già oggi sul tavolo, e non li prende nessuno perché il tabellone li cerca dove non ci sono.** È la ragione per cui la Parte seconda finisce dove finisce.

La forma dell'errore vale più dell'errore. Il numero non era sbagliato: era corretto sui casi su cui era stato calcolato, e quei casi assumevano una cosa che nessuno aveva mai misurato. È un risultato che scade perché l'insieme su cui era definito cambia sotto di lui, e sono i casi che nessun controllo intercetta, per-

ché nel punto in cui accadono non succede niente. Nessun test fallisce, nessun numero diventa strano, e il risultato continua a essere ripetuto finché qualcuno non misura l'ipotesi che nessuno aveva scritto.

---

## 59. Quello che avevo previsto, e quello che mi ha smentito

Prima di calcolare avevo dichiarato per iscritto una serie di aspettative, e vale la pena riportare come siano andate perché è l'unica garanzia che i risultati non siano ricostruzioni di ciò che speravo di trovare.

**Erano sbagliate queste.** Che le coppie con differenze apprezzabili non arrivino al cinque pari. Che la cascata fosse la risposta. Che il set system costasse parecchio. Che il formato a lunghezza variabile fosse il candidato migliore. Che l'anti-informatività del conteggio dei dieci si ritirasse sotto qualunque deviazione dal modello. Che il valore di conoscere la coda restasse sotto la soglia. Che il ripescaggio, essendo più robusto, sorpassasse il riseeding con una qualificazione peggiore. Che su un tabellone fisso migliorare il piazzamento oltre un certo punto non convenisse. Che la connessione di un calendario contasse.

**Erano azzeccate queste.** Che la struttura contasse più della regola. Che il tabellone standard soffrisse meno del riseeding un piazzamento sbagliato. Che l'appaiamento fra i primi recuperasse l'intero divario fra calendari, e qui la misura più fine ha corretto la previsione in un punto, perché estendere l'appaiamento oltre i primi otto non produce un beneficio che questi dati sappiano risolvere.

Era invece azzeccata solo a metà quella sull'anello X: essendo una suddivisione interna al dieci sfugge davvero al vincolo aritmetico sulle coppie strette, ma al crescere del divario ci ricade, e sotto i quaranta punti di distanza scende anche lui sotto la metà.

C'è una regolarità in questo bilancio: **le intuizioni strutturali hanno funzionato quasi sempre, quelle sulle grandezze quasi mai.** Sapevo dove guardare, non sapevo quanto grande fosse quello che avrei trovato. E c'è una formulazione più precisa, arrivata dopo tre errori consecutivi dello stesso tipo: capire un meccanismo autorizza a prevedere in che verso spinga, non chi vinca il confronto, perché descrive una forza e non il suo bilancio contro tutte le altre.

Una nota per chi volesse contestare l'argomento: un alto tasso di previsioni smentite è prova di misura genuina soltanto se le previsioni erano dichiarate prima e abbastanza specifiche da poter essere sbagliate. È l'averle scritte in anticipo a rendere valido l'argomento, non il conteggio degli errori, e nei punti in cui la posta era alta sono state messe per iscritto in un documento datato prima di eseguire il calcolo.

---

## 60. Che cosa direi a chi deve decidere

Prima di tutto in una tabella, perché chi deve decidere ha bisogno di vedere insieme che cosa costi e che cosa renda ciascun intervento.

| Intervento                                     | Che cosa cambia              | Che cosa costa                            | Quanto vale                  |
|------------------------------------------------|------------------------------|-------------------------------------------|------------------------------|
| Tabellone dalla graduatoria, con riseeding     | due righe di regolamento     | niente                                    | +4,0 ± 0,3                   |
| Ripescaggio dai quarti                         | 4 match in più               | un'ora di gara                            | +1,1 ± 0,2                   |
| Regola di parità sulle distanze                | strumentazione               | niente di visibile                        | +0,7 ± 0,1                   |
| Ordinare la qualificazione per distanze        | strumentazione               | niente di visibile                        | +0,5 ± 0,3                   |
| TOTALE+ sulla parità                           | una riga di regolamento      | da 10,6 a 3,6 spareggi                    | +0,33 ± 0,09                 |
| Manopola a due punti                           | una riga, con soglia         | da 10,6 a 8,0 spareggi                    | +4,4 sulla parità            |
| Formato del match                              | qualunque cambiamento        | variabile                                 | ±0,3 di match a gara         |
| Premio individuale sopra un margine dichiarato | una riga di regolamento      | tre premi su quattro diventano collettivi | non misurabile in punti      |
| Appaiare le gare dei primi otto                | sei gare in comune agli otto | niente                                    | +0,8 ± 0,6 sulla graduatoria |

*(Le prime cinque righe vengono tutte dallo stesso file di valori, p2\_modelli.json: stessa corsa appaiata, venticinquemila tornei, seme dichiarato, campo compatto da 64 arcieri, e sono guadagni sulla probabilità che il titolo vada al migliore rispetto alla gara di oggi, che vale 9,712. L'errore riportato è quello del contrasto appaiato. La manopola e il formato del match sono misurati sulla popolazione dei match in parità e non sul titolo, quindi non si confrontano con le prime cinque; l'ultima riga è sulla probabilità che il primo di una graduatoria stagionale sia il migliore di 32 e non si somma alle altre. Le prime due modifiche possono coesistere, ma il loro effetto congiunto non è stato stimato: i due guadagni non vanno sommati. Quelli sulle distanze sono alternativi alle prime due se si sceglie la strada a formato invariato.)*

Due cose vanno lette in quella tabella prima delle altre. **Il primo intervento vale più di dodici volte il secondo e non costa nulla di più, ed è quello di cui non si discute mai.** E l'ultimo agisce su una domanda diversa, non su chi vince una gara, ma su chi risulta il migliore dell'anno, quindi non va sommato agli altri.

Poi, in ordine di quanto valgono, le nove raccomandazioni per esteso.

**Il primo intervento non riguarda la regola di parità e non costa niente:** comporre il tabellone dalla graduatoria di stagione invece che dal punteggio del giorno, e ricomporre gli accoppiamenti a ogni turno. Il titolo passa dall'andare al migliore una gara su dieci all'andarci quasi una su sette, senza togliere un match, uno spareggio, un minuto o una freccia. È il modello bilanciato del capitolo 38, e sono due righe di regolamento.

**Il secondo è la regola di parità, ed è ancora una riga:** vince chi ha totalizzato di più, e se anche i totali sono pari si tira una freccia. Si prende l'85 per cento del margine disponibile contro il 50 di oggi, si verifica a occhio, e fa tirare 3,6 spareggi per gara invece di 10,6.

**Per chi non voglia cambiare nemmeno quelle due righe c'è una terza strada,** che non tocca niente di ciò che si vede: stessa qualificazione, stesso tabellone, stessa durata, e si cambia soltanto come si leggono i risultati. Vale 1,1 punti, cioè un titolo su cento, e sono due interventi: ordinare la qualificazione per distanza dal centro vale 0,5, risolvere le parità sulla stessa misura vale 0,7, e insieme danno meno della somma perché in parte comprano la stessa cosa. Entrambi richiedono di registrare dove ogni freccia sia finita dentro l'anello, cosa che nessuno vedrebbe dagli spalti. Chi non possa permettersi

quella strumentazione ha una sola alternativa, e riguarda la parità: leggere il totale del match prima della freccia, che vale 0,3 punti e non costa nulla. È un'alternativa e non un'aggiunta, le due regole di parità non possono valere sullo stesso match.

**Il criterio storico va tolto e non riordinato.** Contare i dieci dopo una parità di punteggio (e nel regolamento in vigore è il criterio delle parità ordinarie di qualificazione, non dei cutoff né dei match) indica l'arciere sbagliato sotto il modello adottato qui, perché a totale fissato quella misura dice chi ha concentrato il proprio deficit e non chi ha tirato meglio. È l'unica raccomandazione del volume che l'Appendice B mostra dipendere dalla forma delle code. Vale anche per la classifica di qualificazione, dove il conteggio degli X e quello dei dieci non fanno danno ma non aggiungono nulla: tre decimi il primo, zero il secondo.

**Se lo spareggio va difeso, si può difendere a un prezzo noto:** la manopola a due punti conserva otto spareggi su dieci e guadagna comunque quattro volte su cento rispetto alla regola in vigore. Non è vero che accuratezza e spettacolo si escludano.

**Sul formato non c'è niente di meglio da proporre,** e la cosa utile è saperlo: i tre formati a frecce costanti stanno in un terzo di match su 63, e la scelta è su quale spettacolo si voglia (ribaltamenti o finali in bilico) non su quale sia più equo.

**Sulla classifica annuale, due regole scrivibili subito.** Premio individuale soltanto sopra un margine dichiarato in anticipo, che con dieci gare vale due punti se la certezza richiesta è il novanta per cento e tre se è il novantacinque; nel primo caso il premio è assegnabile poco più di una stagione su quattro, nel secondo una su dieci, e nelle altre la forma onesta è collettiva. Quale sia la certezza richiesta è una decisione della federazione, non un risultato di questo lavoro: qui ci sono i numeri per prenderla. E soglia di ammissione al 60 per cento delle gare e non più in alto, perché l'immunità totale si compra escludendo chi non può esserci.

**C'è poi una strada che porterebbe fuori dal margine stretto** in cui tutto il resto si dibatte, ed è misurare le distanze invece degli anelli: porta il tetto della regola di parità da un arciere e otto decimi a due e mezzo per gara, e richiede una precisione di 3 millimetri che non è da laboratorio.

**E infine la cosa che non è una raccomandazione ma riordina la discussione.** Con calendari completi una stagione individua il migliore tre volte su quattro; un torneo, com'è organizzato oggi, lo fa una volta su dieci, e una su sette anche adottando tutto il meglio che queste pagine propongono.

Il problema non è che il torneo sbagli: è che gli si chiedono due cose insieme. Chi vuole sapere chi è il migliore ha già uno strumento, e costa dieci volte di più. Chi vuole uno spettacolo ne ha un altro, e per quello funziona benissimo. Chiederli allo stesso evento è la ragione per cui nessuno dei due riesce del tutto.

## NOTA SU DATI, FONTI E METODO

**Sorgenti regolamentari.** Il regolamento citato è quello della federazione internazionale, Book 3 nella versione del 13 marzo 2026. E i tempi per singola freccia: venti secondi in alternato, trenta nel simultaneo degli eventi World Ranking e annunciati, quaranta negli altri e riducibili a trenta dall'organizzatore.

**Natura dei risultati.** Questo è uno studio di modello, non uno studio empirico. Descrive il comportamento delle regole, dei formati e delle procedure in astratto, e non pretende di dire che cosa sia successo

in una gara precisa. Nessun dato di gara reale è stato usato. La calibrazione parte da punteggi scelti come rappresentativi di un campo di vertice, 690 in giù su 720, e sono un'assunzione dichiarata del modello, non una misura di una popolazione: il resto è derivato da lì.

**Come sono stati ottenuti i numeri.** Dove esiste una formula chiusa si usa quella. Dove i casi possibili sono in numero finito si elencano tutti invece di campionarne alcuni. Dove serve simulare, si usano gli stessi match per tutte le regole confrontate, così che le differenze abbiano rumore molto minore dei livelli. Sono calcolati esattamente: la probabilità del cinque pari, la struttura del tabellone, la forma dello spareggio, i tetti per insieme informativo, il costo del set system e le misure di spettacolo. Sono simulate: la sovrapposizione dei calendari e la manipolabilità, con il numero di ripetizioni scelto perché l'errore stia sotto la cifra riportata.

### **Come sono stati controllati**

Tre cose, in ordine di quanto valgono.

**La prima è l'accordo fra metodo esatto e metodo simulato**, dove entrambi sono disponibili. Esclude insieme l'errore di modello e quello di implementazione, mentre due simulazioni che concordano misurano soltanto il loro rumore comune. Il ribaltamento del capitolo 15, per esempio, è stato ricalcolato per via esatta oltre che per simulazione, e le due strade danno lo stesso segno su tutta la griglia.

**La seconda è l'apparato di verifica** costruito attorno a questo volume, e di cui dico sotto.

**La terza è che ogni controllo è stato collaudato contro il guasto che deve intercettare.** Un controllo che non ha mai fallito non è un controllo: è una decorazione, e il collaudo consiste nel guastare deliberatamente un valore per volta e verificare che qualcuno se ne accorga.

### **L'apparato di verifica**

L'apparato che riesegue i numeri del volume è fatto di più programmi, e li descrivo qui uno per uno.

``motore.py`` contiene le funzioni di calcolo del match e del torneo, ciascuna col nome della formula dell'appendice a cui corrisponde, e tiene i campi di gara come parametro esterno invece che come ipotesi scritta nel codice. La stima congiunta della graduatoria di stagione sta in ``stagione.py``, ed è l'unico percorso per la Parte terza; rifiuta di rispondere se gli arcieri non sono tutti collegati fra loro, perché fra gruppi separati la differenza di origine non è nei dati.

``catalogo.py`` dichiara ogni grandezza prima di calcolarla — che cos'è, su quale popolazione, in quale unità, rispetto a quale livello di riferimento — e la produce dal motore.

La seconda implementazione indipendente non importa il motore principale: ricava l'ampiezza della nuvola per tentativi successivi invece che dalla formula chiusa, e gioca ogni match freccia per freccia invece di consultare una tabella precalcolata. Il suo verificatore confronta sette quantità — due risultati di match, due condizionati e tre strutture di torneo — e nella versione finale tutte rientrano nella soglia dichiarata. È una validazione forte e limitata a quelle sette: non vale per ciò che quello script non esegue.

La generazione precedente dell'apparato è tenuta separata, perché sulle strutture con ripescaggio condiziona il difetto che il motore ha poi corretto — un solo numero casuale per coppia di arcieri, che costringeva il vincitore del primo incontro a vincere anche il rientro. **Due implementazioni che condividono un difetto non si validano a vicenda: concordano perché sbagliano insieme.**

`estrattore.py` legge il testo del volume e pretende che ogni numero stampato abbia una voce nel catalogo: 2 134 quantità, tutte coperte. Se ne trova una senza voce, fallisce.

Va detto con precisione che cosa questo dimostri e che cosa no. La copertura riconosce un numero per compatibilità numerica con una voce del catalogo, e in un volume da duemila quantità una cifra sbagliata trova quasi sempre un omonimo. Esclude quindi che un numero sia inventato, non che sia quello giusto per quella cella.

La difesa contro la cella sbagliata è il legame esplicito, e accanto a esso sta un registro. `registro.py` scrive `REGISTRO.md`: per ogni grandezza che il volume usa per raccomandare qualcosa, che cosa misura, su quale popolazione, rispetto a quale riferimento, da quale produttore e da quale corsa — ripetizioni e seme — il valore, il suo errore, e quante frasi o celle vi sono ancorate. Le grandezze decisionali sono ventotto e tutte hanno almeno un legame. Il registro elenca comunque quelle che non ne avessero, perché è su quelle che una divergenza resterebbe leggibile invece di diventare rossa. `vincoli.py` ancora 88 celle decisionali e 56 frasi a una voce precisa del catalogo — identificando la cella prima di guardarne il valore — e fallisce se la cifra stampata non è quella che il produttore calcola. Fuori da quelle celle la garanzia è più debole: la maggior parte delle quantità è legata a una voce del proprio capitolo, le altre sono riprese di grandezze prodotte altrove.

`manifest.py` scrive la catena di provenienza delle figure — quale funzione produce ciascuna, da quali file legge, e se il disegno incorporato qui è quello che il codice corrente produce — e fallisce se una figura ha più di un generatore, che è il modo in cui una versione vecchia sopravvive accanto a una nuova.

`collaudo.py` guasta i valori uno alla volta e verifica che vengano intercettati. Ogni prova guasta il volume in un modo diverso — una cella che stampa un valore diverso, un produttore che cambia mentre la pagina resta indietro, una figura che non coincide col proprio generatore, un file che cambia dopo il congelamento, un conteggio dichiarato male, il segno di un guadagno invertito — e ciascuna pretende non solo che il comando fallisca, ma che fallisca per la ragione giusta. Scrivendole ho scoperto che alcuni guasti passavano indisturbati, fra cui gonfiare il guadagno del modello bilanciato nella tavola delle raccomandazioni.

Una cifra contestata si mette alla prova nello stesso modo: si cambia il suo valore nel catalogo, si riesegue e si guarda se il controllo lo intercetta. Se non lo intercetta, la tolleranza è troppo larga, ed è un difetto da correggere.

### **Una verifica di tipo diverso: riscrivere leggendo solo la prosa**

Oltre a ricalcolare i numeri ho fatto una cosa che nessun controllo automatico può fare: ho riscritto le procedure leggendo soltanto la loro descrizione in queste pagine, senza guardare il codice, e ho confrontato i risultati.

Le sette regole di parità tornano tutte entro un quarto di punto. Il tabellone standard e il riseeding, la cascata, CONCORDE e tre dei cinque formati tornano entro l'errore di simulazione. I primi due passi della procedura della Parte terza tornano alla cifra.

Una sola quantità divergeva, il formato a 15 set da una freccia, e la ricostruzione aveva ragione. Vale la pena raccontarlo per intero, perché è il difetto più profondo che questo volume abbia contenuto.

La struttura del match coincideva alla sesta cifra: entrambe le implementazioni davano la stessa probabilità di arrivare a quindici pari. La differenza stava tutta nel modo di risolverla. La causa era questa: quanto valesse TOTALE+ veniva calcolato una volta sola, sui match che finiscono cinque pari nel formato in vigore, e poi riusato per tutti e cinque i formati. Ma i match che un formato manda in parità non sono gli stessi che ci manda un altro, e su match diversi la stessa regola vale diversamente. La correzione è stata calcolare ciascun formato sui match che quel formato manda davvero allo spareggio.

Ricalcolata sui match giusti, la riga dei set corti scende di due decimi a tutti e tre i divari, quella dei set lunghi sale di due centesimi, i due formati variabili non si muovono, e il formato in vigore torna identico alla quarta cifra, perché era quello su cui il valore era stato calcolato.

Lo stesso valore entra anche nella simulazione dei tornei da cui esce il modello precisione, che è l'unico dei quattro a usare il formato a set corti. Quel modello è stato quindi ricalcolato con il metodo corretto, e il valore che ne esce, 14,22 per cento, è quello che il volume riporta ovunque: la correzione è già dentro, non è una revisione successiva.

Vale la pena dire perché la correzione sia così piccola, perché è un errore che ho commesso io prima di calcolarla. Stimandola dalla tabella dei formati, dove la riga scende di due decimi, avevo previsto un calo intorno al 14,8. Ma quella tabella è a sei, undici e venticinque punti di divario, mentre in un campo compatto due vicini di classifica distano quattro decimi di punto, e a quel divario i due valori quasi coincidono. La correzione vera è di tre centesimi, non di due decimi: dieci volte meno di quanto la stima suggerisse. È la ragione per cui il numero va calcolato e non estrapolato, e questa volta la lezione l'ho imparata sul mio stesso errore.

La conseguenza per il lettore è che il vantaggio del formato migliore sul formato attuale si dimezza, da mezzo punto a tre decimi, il che rafforza la conclusione del capitolo invece di indebolirla: lo spazio dei formati è ancora più piatto di quanto avessi scritto.

### **Che cosa l'apparato ha già corretto**

Vale la pena raccontarlo, perché è il modo in cui questo apparato si è guadagnato la fiducia che gli chiedo.

Ricostruendo indipendentemente il ripescaggio dei quarti ho ottenuto 7,8 per cento contro il 10,9 pubblicato. Non era un errore di calcolo di nessuna delle due parti: era che la frase con cui il volume descriveva quella struttura ammetteva quattro gare diverse, con risultati fra il 7,3 e l'11,2 per cento. La descrizione è stata riscritta perché una sola lettura resti possibile.

Nello stesso confronto sono emersi il numero di match di quella struttura, che è 67 e non 66; l'accuratezza del ripescaggio tardivo nella tabella delle sei strutture, che è 11,7 e non 10,60; e l'ampiezza dell'effetto di gara della Parte terza, che è di undici punti e non di otto.

Una tabella non era riproducibile ed è stata rifatta. Il confronto fra piazzamento vero e piazzamento reale del capitolo 27 veniva da una corsa non conservata come codice, ed è stato ricalcolato dal motore su centocinquantomila tornei. I livelli si sono spostati di circa un punto, mentre la conclusione del capitolo non si è mossa: il riseeding paga un pedaggio quattro volte quello delle altre strutture.

### **Quanto del volume è coperto, e in che modo**

Questa è la domanda che chiunque dovrebbe pormi, e la risposta onesta non è tutto.

Nel testo di questa seconda edizione compaiono 2 134 quantità sostanziali, e il catalogo le dichiara tutte: 983 sono legate a una voce del capitolo in cui compaiono e 1 151 coincidono con una grandezza prodotta altrove, che il testo richiama. Il controllo di copertura è stato rieseguito sul manoscritto riscritto, non ereditato dalla prima edizione, e non lascia scoperto nessun numero.

Va detto con precisione che cosa questo dimostri e che cosa no, perché è il limite che questa edizione ha reso visibile. La copertura dimostra che ogni numero stampato viene dal motore; non dimostra che la frase attorno a quel numero gli attribuisca il significato giusto. Una revisione ostile su questa edizione ha trovato più casi in cui la prosa nuova diceva più del dato, un modello con tre valori sotto la stessa etichetta, una conclusione fondata su una misura che il testo stesso dichiarava da rifare, un intervento dato per nullo che il produttore misura positivo. Tutti corretti, e nessuno di quei difetti sarebbe stato intercettato da un controllo numerico.

Ogni tabella del volume è verificabile, e una parte è anche rigenerabile da capo.

### **Che cosa i controlli non vedono**

Due cose, ed è onesto elencarle perché chi legge la descrizione dell'apparato ne inferirebbe una garanzia più ampia di quella che c'è.

**La prima è il significato attribuito a un numero.** Nessun controllo verifica che l'etichetta sopra una colonna sia quella giusta. In una revisione finale su file mai riletti sono emersi quattro difetti, e nessuno dei quattro era un numero sbagliato: erano numeri esatti con l'etichetta sbagliata. Una revisione successiva, condotta con un catalogo che dichiara popolazione, unità e riferimento prima del calcolo, ne ha trovati altri della stessa famiglia — un valore calcolato contando come primi posti anche quelli condivisi, due celle che erano i valori esatti di altre due coppie, e una tavola dichiarata esatta che non lo era.

È la classe di difetto più insidiosa, e il catalogo la riduce senza chiuderla. La procedura che la chiude non è tecnica: è rileggere ogni didascalia a distanza di tempo, dire con parole proprie che cosa misura la colonna, e confrontare quell'enunciato con quello che c'è scritto. Chi ha prodotto il numero lo rilegge sapendo già che cosa dovrebbe dire, ed è la ragione per cui quella rilettura va fatta con la voce del catalogo accanto e non a memoria.

### **Precisione dichiarata per parte**

La **Parte prima** è in forma esatta quasi ovunque, e le sue cifre reggono fino alla seconda decimale.

La **Parte seconda** è simulata, e le corse non hanno tutte la stessa lunghezza: dodicimila tornei per i confronti fra origini del piazzamento, venticinquemila per la corsa comune dei quattro modelli, quarantamila per il confronto fra regole di parità, centocinquantamila per i pedagogi appaiati. Il numero di ripetizioni e il seme di ciascuna stanno nella didascalia della tavola e nel registro delle grandezze decisionali. L'errore sulle probabilità di vittoria va da uno a quattro decimi di punto secondo la corsa, quindi vanno lette due cifre significative e le differenze inferiori a mezzo punto sono assenti, salvo dove riporto un confronto a parità di tutto il resto con il suo errore.

La **Parte terza** mescola le due: i tetti sono in forma chiusa, la sovrapposizione dei calendari e la manipolabilità sono simulate.

**Confronti a parità di tutto il resto.** Dove confronto due alternative uso gli stessi tornei simulati per entrambe, stesso campo, stesso piazzamento, stessi esiti dei match, così che la differenza abbia un ru-

more molto minore dei livelli. È il motivo per cui posso dichiarare significativo un divario di mezzo punto fra due strutture mentre i livelli assoluti oscillano di altrettanto.

**Riproducibilità.** Il seme del generatore casuale è fissato nel codice, quindi anche i valori simulati si riproducono esattamente. Cambiando il seme si ottengono numeri diversi entro la tolleranza dichiarata, ed è quello il modo giusto di verificarli: due esecuzioni con lo stesso seme misurano soltanto che il codice è deterministico. Non fornisco il codice di simulazione con questo lavoro. Per utilizzarlo, possiamo parlarne.

### **Limiti dichiarati**

Sette limiti, e valgono per tutto il volume: primo, secondo, terzo, quarto, quinto, sesto e settimo, nell'ordine in cui seguono.

**Primo.** La bravura è una quantità sola, il punteggio atteso. Un arciere è anche la sua tenuta sul lungo, la gestione del vento, la capacità di riprendersi da un turno storto: niente di tutto questo compare nei conti.

**Secondo.** Nel confronto fra regole cambia il modo di leggere i dati, non il comportamento di chi li produce. Un regolamento diverso indurrebbe verosimilmente scelte di tiro diverse, e il conto invece tiene ferme le frecce.

**Terzo.** Da un singolo incontro non si ricava quale dei due contendenti fosse il più forte: quella grandezza non è osservabile a quella scala.

**Quarto.** La forma della coda non è stimata. Ho misurato quanto varrebbe conoscerla, e la ragione strutturale per cui probabilmente non si saprà (l'informazione sta in eventi rari per definizione) ma non equivale a saperlo.

**Quinto.** Il livello di un arciere sta fermo per tutta la stagione. È l'ipotesi che rende lecito mediare dieci gare in una stima sola, e ho misurato quanto regga facendo camminare il livello di un passo casuale fra una gara e l'altra.

Per la domanda che adotto, chi ha tirato meglio quest'anno, la stima piatta resta la migliore a tutte le velocità di deriva provate: passa dal 73,2 per cento di primi posti giusti all'83,2 con tre punti di passo, perché la deriva allarga le distanze fra arcieri e rende il primo più riconoscibile. Per l'altra domanda, chi è più forte all'ultima gara, la stima piatta cede: cala dal 73,2 al 56,1, e una stima che pesi di più le gare recenti la supera oltre un punto di deriva per gara, cioè oltre tre punti accumulati in una stagione da dieci.

Sotto quella soglia la stagione si può leggere come se il livello fosse fermo. Sopra, chi volesse usare la graduatoria come piazzamento (che è una domanda sul domani e non sull'anno passato) dovrebbe pesare le gare recenti, dimezzando il peso ogni due o tre gare. È la ragione per cui distinguo i due usi della stessa graduatoria invece di trattarli come uno solo.

**Sesto.** L'effetto dell'ordine di tiro allo spareggio non è misurabile. Un eventuale vantaggio dell'ordine è inseparabile da un vantaggio di bravura, perché l'ordine lo sceglie il meglio classificato. Sorteggiare l'ordine lo renderebbe misurabile, e nessuno lo fa.

**Settimo.** Il doppio errore allo spareggio non è nel modello. Il regolamento prevede che, se entrambi mancano la zona di punteggio, entrambi tirino un'altra freccia; il motore decide sempre sulla freccia più

vicina al centro. Ai livelli di queste pagine due frecce fuori bersaglio nello stesso spareggio capitano due volte ogni cinquanta miliardi anche col dieci per cento di code, e nessun numero del volume dipende da come verrebbero trattate.

## APPENDICE — VENTUNO COSE IMPARATE SBAGLIANDO

Un lavoro lungo produce, oltre ai risultati, un elenco di errori. Li pubblico perché sono la parte più trasferibile, e perché quasi nessuno li pubblica, non per reticenza, ma perché quasi nessuno tiene il verbale mentre lavora.

Sono divisi in due categorie, e la distinzione è utile: un principio dice quando hai sbagliato, una nota dice come fare.

---

### Principi

**Uno — Scegliere quali casi si guardano abbassa il massimo raggiungibile, non solo la stima.** Il cinque pari toglie alle regole che leggono il match fra il nove e il sedici per cento del loro vantaggio, prima ancora che si scelga come leggerle. E una regola su chi può entrare in classifica restringe allo stesso modo il tetto di una graduatoria.

**Due — Ogni controllo va collaudato contro il guasto che deve intercettare.** Un controllo mai visto fallire non è un controllo, è una decorazione. È il principio con più occorrenze di tutti, e ne ha quattro forme distinte: il silenzio, quando non spara mai; il rumore, quando spara a caso; il bersaglio sbagliato, quando spara per la ragione sbagliata; lo scollegamento, quando spara e nessuno ascolta. Un controllo che verificava la forma di una curva con una tolleranza più larga dell'effetto vero ha certificato per settimane una proprietà falsa.

**Tre — Due forze in verso opposto producono un massimo in mezzo.** Quando due effetti agiscono in direzioni contrarie sullo stesso asse, la risposta sale e poi scende, e il segno può invertirsi due volte. Il guadagno di una regola di spareggio, la perdita da lettura per anelli, l'inversione del conteggio dei dieci in funzione della coda: tre volte la stessa forma.

**Quattro — Un limite superiore non è una stima.** Autorizza a dire che il problema è limitato, non quanto valga. Un limite usato come se fosse una misura si è rivelato sbagliato di otto ordini di grandezza, cioè cento milioni di volte.

**Cinque — La quantità che decide si stima con molti meno dati dei parametri che la spiega.** Duecentosedici frecce per prendere la decisione, diecimila per spiegarla.

**Sei — Dove i casi sono in numero finito, si elencano tutti invece di campionarne alcuni.** Vale per il calcolo e per la verifica: una dimostrazione con un numero finito di casi va resa eseguibile, non illustrata con un campione. Con l'avvertenza che stimare quanto costi l'elenco richiede la stessa attenzione alle unità che serve altrove; contare le frecce dove servivano le coppie ha fatto sembrare trattabile ciò che non lo era.

**Sette — Verificare richiede indipendenza di metodo, non di seme.** Due simulazioni che concordano misurano il loro rumore comune. Un metodo esatto e uno simulato che concordano escludono insieme errore di modello ed errore di implementazione.

**Otto — Due alternative si confrontano solo se sono confrontabili:** stesso stadio finale, stesso insieme di informazioni. Confrontare una regola troncata con una intera favorisce la seconda anche quando ogni singolo numero è corretto.

**Nove — Un criterio definito su una famiglia non si valuta su un membro, e viceversa.** «Entro tre errori standard» ripetuto su trentuno celle non è un controllo al 99,7 per cento: sono trentun controlli, e su trentuno tentativi qualcosa passa per caso.

**Dieci — Una quantità va riportata nell'unità in cui la decisione si prende.** Nella valuta sbagliata è vera e illeggibile. «Un fattore otto sui minuti» è vero e fuorviante quando in unità di gara significa nove minuti in meno; «due virgola ottantasette punti percentuali» è vero e inservibile quando significa un arciere e otto decimi su sessantatré match.

**Undici — I casi su cui un numero è calcolato devono essere quelli su cui la decisione agisce.** Il caso più costoso di questo lavoro è di quella forma: l'affidabilità della qualificazione era calcolata contando come primo posto anche quello condiviso, mentre una decisione un pareggio lo rompe. Su cento gare fa 16,5 contro 14,2, e la seconda è quella che conta.

Sui casi sbagliati il numero è vero e non pertinente, e a differenza dell'unità sbagliata non si converte: si ricalcola. È il difetto più insidioso, perché una valuta sbagliata si vede guardando il numero e una popolazione sbagliata no.

**Dodici — L'assenza di sintomi va spiegata, non accettata.** Prima di concludere che qualcosa funziona va stabilito che qualcuno stesse guardando. I due difetti peggiori del lavoro si nascondevano a vicenda, ciascuno impedendo di vedere l'altro: un'impronta mai scritta e una pulizia che cancellava i file prima che qualcuno li leggesse.

**Tredici — Prima di scrivere «finché non si sa», stimare quanto costi saperlo.** Una cautela in nota autorizza a scrivere nel testo l'affermazione che la cautela dichiarava non stabilita. Nel caso che ha prodotto questo principio, la misura che chiudeva la questione costava un minuto ed era già scritta. Il principio regge in tre versi: stimare quanto costa sapere prima di dire che non si sa; stimare quanto vale sapere prima di spendere per saperlo; stimare quanto costa applicare ciò che già si sa prima di limitarsi ad annotarlo.

**Quattordici — Una correzione applicata a un caso non si estende agli altri della stessa classe,** e sapere che la classe esiste non basta a proteggerli. Il rimedio è cercare la classe quando si scopre il caso, non annotare il caso.

**Quindici — Capire un meccanismo autorizza a prevedere in che verso spinga, non chi vinca il confronto.** Descrive una forza, non il suo bilancio contro tutte le altre. Tre previsioni consecutive di questo lavoro erano fondate su meccanismi corretti e portavano a conclusioni sbagliate.

**Sedici — Un sostituto costruito per somigliare all'oggetto vero su una statistica non è l'oggetto vero.** Un ordinamento costruito per avere una certa somiglianza con l'ordine vero e uno prodotto da una stagione simulata avevano la stessa somiglianza e strutture d'errore diverse — abbastanza da

produrre un massimo in mezzo che non esiste. Costruirlo per somigliare dà l'impressione di aver controllato.

**Diciassette — Un sospetto che spiega bene i dati non è per questo vero**, e la differenza la fa un controllo che possa smentirlo invece di uno che possa confermarlo. Un sospetto plausibile, coerente con un risultato già stabilito, e sbagliato è la configurazione più pericolosa: la coerenza con quel che già si sa è precisamente ciò che fa smettere di controllare.

---

## Note di tecnica

**Diciotto — Una trasformazione automatica applicata al testo rompe numeri che nessun controllo sui calcoli può vedere.** Convertendo i numeri da lettere a cifre ho spezzato una dozzina di valori: 0,9898 è diventato «0,9 otto nove otto», 692 è diventato «690due», un intervallo del 19-47 per cento è diventato «diciannove-47». Tutte le verifiche numeriche continuavano a passare, perché i calcoli erano giusti e il testo era rotto — sono due oggetti diversi, e l'apparato ne controllava uno solo.

**Diciannove — L'inventario va fatto sui numeri pubblicati, non sui controlli scritti.** Costruendo l'apparato di verifica ho scoperto che sette gruppi di valori del volume non erano coperti da nessun controllo: i quattro tetti, le quattordici modalità, i momenti del ribaltamento, l'esempio percorso della Parte terza. Non fallivano, non venivano proprio guardati, ed è così che un refuso è sopravvissuto a tre revisioni.

La domanda giusta non è «tutti i controlli passano», che è la domanda facile, ma «quale controllo si accorgerebbe se questo numero fosse sbagliato», posta un numero per volta.

E non si mantiene a mano ciò che si può derivare. Cinque applicazioni portanti: il perimetro dei controlli definito sulla cartella e non sull'elenco dei produttori; l'analisi della struttura del codice invece della ricerca testuale; l'impronta ricavata dai moduli invece che da una lista; la prosa generata dai numeri invece che scritta accanto; e lo stato di un documento scritto nel documento invece che nel messaggio che lo accompagna. Un controllo che guarda un posto trova anche ciò che non stava cercando; un elenco può solo mancare di una voce.

**Venti — Una frase che ammette più letture è un difetto che nessun controllo sui numeri può vedere.** La descrizione di una struttura di gara conteneva due ambiguità indipendenti (chi cedesse il posto, e se lo cedesse giocando o no) e quelle due ambiguità facevano quattro gare diverse, con risultati fra il 7,3 e l'11,2 per cento. Tutte le verifiche numeriche continuavano a passare, perché il numero pubblicato era giusto: era la frase a descriverne un altro.

Il rimedio non è rileggere con attenzione, perché chi rilegge sa già che cosa intendeva: è contare le letture possibili di ogni frase che descriva una procedura, e riscriverla finché non ne resta una.

**Ventuno — Due implementazioni indipendenti trovano quello che nessuna verifica interna trova.** L'ambiguità della voce precedente non è emersa da un controllo, ma dal fatto che una seconda implementazione, scritta leggendo soltanto la descrizione in prosa e senza guardare la prima, ha costruito una gara diversa e ha ottenuto un numero distante tre punti e mezzo.

Nessun apparato di verifica costruito attorno a un'implementazione sola avrebbe potuto accorgersene, perché ogni controllo avrebbe confermato che il codice fa quello che il codice fa.

La stessa cosa è successa una seconda volta, e vale la pena dirlo perché conferma la regola invece di illustrarla soltanto: una seconda implementazione riscritta dalla specifica corrente ha mostrato che l'affidabilità della regola sulle distanze veniva presa su tutti i match invece che sui soli match finiti pari, con una sovrastima di nove punti. Trentuno controlli interni erano verdi, perché il motore era coerente con sé stesso.

**La ridondanza vera non è ricalcolare due volte: è che due persone leggano la stessa descrizione e costruiscano.**

---

## E, sopra tutti, il limite che nessuno di loro vede

Tutti controllano i calcoli. Nessuno controlla il significato attribuito ai risultati.

È il limite dell'apparato, e va dichiarato con la stessa evidenza dei suoi successi, perché chi vede l'apparato senza il suo limite ne inferisce una garanzia che non c'è.

# APPENDICE MATEMATICA

Questa appendice contiene le formule usate nel volume, con il minimo di derivazione necessario a rifare i conti. Le formule sono numerate da S1 e sono richiamate nel testo. Chi non la legge non perde nessun risultato: qui non ce ne sono di nuovi, solo le ragioni di quelli già dati.

---

## Il modello

**S1 — Distribuzione radiale.** Se le due coordinate d'impatto sono indipendenti, di media nulla e uguale dispersione  $\sigma$ , allora la distanza dal centro  $R$  ha funzione di ripartizione

$$F(r) = P(R \leq r) = 1 - \exp(-r^2 / 2\sigma^2), r \geq 0$$

ed è la distribuzione di Rayleigh di parametro  $\sigma$ . La densità è  $f(r) = (r/\sigma^2) \cdot \exp(-r^2/2\sigma^2)$ . Equivalentemente,  $R^2/\sigma^2$  segue una chi quadrato a due gradi di libertà.

**S2 — Punteggio di una freccia.** Con anelli di larghezza  $W$  e faccia a dieci corone, il punteggio è  $s(r) = \max(0, 11 - \lceil r/W \rceil)$  con tetto a 10. La probabilità dell'anello di valore  $v$  vale

$$p_v(\sigma) = F(a_{v-1}) - F(a_v), \text{ con } a_v = (10 - v) \cdot W$$

e il punteggio atteso su  $n$  frecce è  $n \cdot \sum_v v \cdot p_v(\sigma)$ . L'inversione numerica di questa relazione — dato il punteggio atteso, trovare  $\sigma$  — produce la tabella di calibrazione del capitolo 4. Per la faccia da 122 cm,  $W = 6,1$  cm e il raggio dell'anello  $X$  vale 3,05 cm.

Nota di stabilità numerica. Calcolando  $p_v$  come differenza di due valori di  $F$  si perde ogni cifra significativa negli anelli bassi: per un arciere da 690 punti i tre anelli più esterni risultano esattamente zero, perché  $F$  satura a uno in aritmetica a precisione finita già a  $r \approx 36$  cm per  $\sigma \approx 3,8$  cm. La forma corretta è in scala logaritmica:  $\log p_v = -a_v^2/2\sigma^2 + \log(1 - \exp(-(a_{v-1}^2 - a_v^2)/2\sigma^2))$ . La somma delle probabilità resta uno anche con la forma instabile, quindi il controllo sulla normalizzazione non intercetta questo difetto.

---

## La regola ottimale sulle coordinate

**S3 — Rapporto di verosimiglianza.** Siano note le due dispersioni come insieme non ordinato  $\{\sigma_b, \sigma_p\}$ , con  $\sigma_b < \sigma_p$ , e si debba decidere quale arciero abbia quale. Il logaritmo del rapporto fra le due assegnazioni possibili vale

$$\log \Lambda = \frac{1}{2} \cdot (1/\sigma_p^2 - 1/\sigma_b^2) \cdot (S_A - S_B), \text{ con } S_j = \sum_i r_{\{j,i\}}^2$$

Poiché il primo fattore è negativo per ogni coppia, la regola di Bayes assegna la dispersione minore a chi ha la somma dei quadrati dei raggi minore, e il segno non dipende dai valori delle due dispersioni: la regola è ottimale uniformemente. Ne segue che è ottimale per qualunque distribuzione a priori simmetrica sulle due assegnazioni, con perdita zero-uno, ed è minimax nella stessa classe: con perdita asimmetrica o con priori sbilanciati la soglia del rapporto di verosimiglianza si sposta e l'ottimalità va riesaminata.

**S4 — Invarianza al condizionamento.** L'evento «cinque pari» è invariante allo scambio delle etichette dei due arcieri, quindi  $P(5-5 \mid H_1) = P(5-5 \mid H_2)$  e le due costanti di normalizzazione si elidono nel rapporto. La regola di S3 resta di Bayes anche condizionatamente alla parità.

**S5 — Forma chiusa non condizionata.** Poiché  $S_j/\sigma_j^2 \sim \chi^2_{\{2n\}}$ , il confronto fra le due somme è un rapporto di chi quadrato scalate, dunque

$$C_{\text{unc}}(n, \rho) = P(F_{\{2n, 2n\}} < \rho^2), \text{ con } \rho = \sigma_p/\sigma_b$$

È la tabella del capitolo 6. Per  $n = 1$  la ripartizione di  $F(2,2)$  vale  $x/(1+x)$ , da cui il caso dello spareggio a una freccia (S8).

---

## I pesi sugli anelli

**S6 — Regola ottimale sui soli anelli.** Se si osservano solo i conteggi per anello, la regola di Bayes è un conteggio pesato con pesi

$$w_v = \log[ p_v(\sigma_b) / p_v(\sigma_p) ]$$

e la decisione confronta  $\sum_v n_v w_v$  fra i due arcieri. I pesi dipendono dai valori delle due dispersioni, non solo dal loro ordine. Se le dispersioni non sono note si stimano, e il modo di stimarle decide tutto: il livello comune  $\hat{o}$  si ricava dal totale delle trenta frecce dei due referti presi insieme, invertendo il punteggio atteso, e i pesi si calcolano sulla coppia  $\{\hat{o}(1-\delta), \hat{o}(1+\delta)\}$  con  $\delta$  separazione nominale dichiarata. È la forma dei pesi a dipendere dal livello, non dallo scarto, ed è la ragione per cui la versione stimata decide come quella a dispersioni note salvo circa una volta su un milione — 52 volte su 50 468 034 match in parità. Stimando invece la dispersione di ciascun arciero dal suo solo referto la proprietà crolla: quando i due totali coincidono i pesi si annullano e la regola smette di decidere, con un disaccordo che sale al dieci per cento.

**S7 — Concavità dei pesi.** Ponendo  $t = 1/2\sigma^2$ ,  $u_k = (kW)^2$ ,  $d_k = u_k - u_{\{k-1\}} = W^2(2k-1)$  e  $h(y) = \log(1 - e^{-y})$ , la differenza seconda dei pesi rispetto all'indice d'anello si decompone in

$$\Delta^2 w = -2W^2(t_b - t_p) + [G_k(t_b) - G_k(t_p)], \text{ con } G_k(t) = h(d_{\{k-1\}}t) - 2h(d_k t) + h(d_{\{k+1\}}t)$$

Il primo termine è strettamente negativo. Riscaldando in  $v = W^2 t$  il fattore  $W^2$  si semplifica e la condizione di concavità diventa un numero puro per anello:  $\sup_v Q_k(v) < 2$ , con  $Q_k(v) = [\psi((2k-3)v) - 2\psi((2k-1)v) + \psi((2k+1)v)]/v$  e  $\psi(y) = y/(e^y - 1)$ . Per Taylor con resto,  $Q_k(v) = 4v\psi''(\xi)$  con  $\xi \geq (2k-3)v$ , da cui  $Q_k \leq 4C/(2k-3)$  con

$$C = \sup_{y>0} y \cdot \psi''(y) = 0,23154848, \text{ raggiunto a } y = 2,3469$$

Serve  $C < 1/2$ : il margine è 2,159. La costante non dipende da  $\sigma$ , dalla coppia, dalla larghezza d'anello né dal formato: la concavità è globale. Il caso di bordo dell'anello esterno si chiude separatamente osservando che  $\psi'' > 0$  implica  $|\psi'| \leq 1/2$ .

Nota di stabilità. La forma chiusa  $\psi''(y) = ey[(y-2)e^y + y + 2]/(ey - 1)^3$  è numericamente inaffidabile sotto  $y \approx 10^{-4}$  (cancellazione al numeratore) e sopra  $y \approx 700$  (traboccamento). La verifica di positività va condotta su  $\psi(y) = y/\expm1(y)$  con differenze finite, o su una forma equivalente riscalata.

## La parità a set

**S8 — Spareggio a una freccia.** Per due variabili di Rayleigh indipendenti di parametri  $a$  e  $b$ ,

$$P(R_A < R_B) = b^2 / (a^2 + b^2)$$

che con  $b = \rho a$  dà  $\rho^2/(1+\rho^2)$ . Per misture di Rayleigh la forma si estende a  $\sum_p \sum_q w_p w_q \cdot b_q^2/(a_p^2 + b_q^2)$ . Coincide con la forma di S5 a  $n = 1$ .

**S9 — Probabilità di finire cinque pari.** Detti  $p_w, p_t, p_l$  le probabilità che un set sia vinto, pareggiato, perso dall'arciere A, la parità a cinque richiede  $a$  vittorie,  $t$  pareggi,  $b$  sconfitte con  $2a + t = 5$  e  $a + t + b = 5$ , da cui  $b = a$  e  $a \in \{0,1,2\}$ . I tre termini hanno molteplicità multinomiale 1, 20, 30, per un totale di 51 cammini. Inoltre, con  $a \leq 2$  il punteggio massimo raggiungibile prima dell'ultimo set è 5: la soglia di arresto a sei non è mai vincolante sulla fetta del cinque pari, quindi passeggiata libera e catena assorbente coincidono lì. Ne segue

$$P(5-5) = p_t^5 + 20 \cdot p_w p_t^3 p_l + 30 \cdot p_w^2 p_t p_l^2$$

**S10 — Guadagno di una regola di spareggio.** Definito  $C(\text{regola}, \Delta)$  come la probabilità che la regola indichi l'arciere migliore condizionatamente alla parità, il guadagno rispetto al sorteggio vale

$$G(\Delta) = P(5-5 | \Delta) \cdot [C(\text{regola}, \Delta) - 1/2]$$

ed è la quantità riportata in punti percentuali sulla probabilità di vincere il match.  $G$  si annulla a  $\Delta = 0$  (dove  $C = 1/2$  per simmetria) e per  $\Delta \rightarrow \infty$  (dove  $P(5-5) \rightarrow 0$ ), e ha quindi un massimo interno.

## Il rovesciamento del conteggio dei dieci

**S11 — L'enunciato.** Sia  $T = \sum_i s_i$  il totale delle quindici frecce e  $D = \#\{i : s_i = 10\}$  il conteggio dei dieci. Condizionatamente a  $T$  uguale per i due arcieri,  $D$  si comporta come una misura di dispersione e non di posizione: a totale fissato entrambi gli arcieri hanno perso lo stesso deficit rispetto al massimo, e chi ha più dieci lo concentra su meno frecce. Posto  $K = 150 - T$  il deficit dal massimo, esso si distribuisce sulle  $15 - D$  frecce che non sono dieci, e il deficit medio su quelle vale  $K / (15 - D)$ : cresce con  $D$ . Il

vincolo è sul peso che ciascuna freccia non massima deve portare, non sul numero delle frecce basse né sulla varianza, che a totale fissato possono muoversi nei due sensi. Su questa popolazione la corrispondenza con le frecce basse è stretta, ma non è un vincolo: a totale 140 si danno dieci dieci con una sola freccia sotto il nove, e nove dieci con due — quanto stretta lo dice la correlazione qui sotto.

Formalmente, per un vettore di conteggi  $(n_0, \dots, n_{10})$  con  $\sum n_v = 15$  e  $\sum v \cdot n_v = T$  fissati, la coppia  $(n_{10}, \sum_{v \leq 8} n_v)$  non è vincolata: aumentare  $n_{10}$  di uno richiede di sottrarre punti altrove, e il modo che minimizza il numero di frecce coinvolte è abbassarne una di molti anelli — ma non è l'unico, e le configurazioni che vanno in controtendenza si costruiscono a tavolino, come il controesempio a totale 140 mostra. Quella che segue non è dunque una necessità combinatoria: è una proprietà del modello, e la correlazione qui sotto ne misura la forza.

Il vincolo si misura con la correlazione di Pearson fra la differenza di dieci e la differenza di frecce da otto o meno, calcolata sulla popolazione dei match con i due totali uguali e messa insieme su tutti i totali, senza centrare dentro ciascuno. Vale 0,9898 sulla cella di riferimento e decresce col livello: 0,9978 sulla coppia 700-690, 0,9488 sulla 680-660. A livelli bassi esistono più modi di comporre lo stesso totale, quindi il vincolo aritmetico si allenta e la corrispondenza fra dieci e frecce basse smette di essere quasi deterministica.

**S12 — Conseguenza.** Sotto il modello standard, dispersione maggiore corrisponde a bravura minore. A parità di totale il conteggio dei dieci indica quindi l'arciere meno bravo più spesso di quanto indichi il migliore, e la misura vale più della deduzione: singoli referti in controtendenza si costruiscono con quindici frecce e totale centoquaranta. Le misure esatte danno da 36,6 a 48,7 su cento su tutta la griglia esaminata: sotto cinquanta ovunque.

**S13 — Dominio.** Il rovesciamento si ritira quando la coda della distribuzione è marcata e concentrata, perché la coda aggiunge modi di comporre uno stesso totale e allenta il vincolo. Non si ritira sotto decentramento, anisotropia o deriva. Il discriminante non è la massa negli anelli bassi ma la sua concentrazione, e due configurazioni della griglia lo mostrano: la coda al 3 % su tre anelli porta una massa sotto il nove di 0,332 e la coda scentrata di 0,385 — quasi la stessa — ma le correlazioni fra dieci e frecce basse valgono 0,831 e 0,991, e i comportamenti si separano. La prima ritira l'inversione portando il conteggio dei dieci a 52,6, cioè sopra la monetina; la seconda la approfondisce a 41,8. Stessa massa bassa, esito opposto: quello che decide è come quella massa sia distribuita. La coppia costruita apposta per isolare il meccanismo lo mostra a massa esattamente uguale: due miscele in cui una frazione  $\varepsilon$  delle frecce viene da una nuvola  $c$  volte più larga, con  $\varepsilon = 0,40$  e  $c = 1,2$  nella prima — molte frecce spostate di poco — ed  $\varepsilon = 0,010$  e  $c = 6,0$  nella seconda, poche frecce spostate di molto. Le masse sotto il nove valgono 0,2112 e 0,2114, indistinguibili; il vincolo aritmetico vale 0,991 e 0,715; e il conteggio dei dieci vale 45,78 nella prima e 49,07 nella seconda.

**S14 — L'anello X.** L'anello X è una suddivisione interna all'anello dei dieci: passare da un dieci a un X non altera il totale, quindi il conteggio degli X non è vincolato dalla somma nello stesso modo. Resta perciò informativo condizionatamente ai totali pari, finché il raggio dell'X è grande rispetto alla dispersione. L'attraversamento dipende dal divario fra i due arcieri e non è un valore unico: vale  $r_X/\sigma = 0,508$  con 5 punti di divario, 0,522 con 10 e 0,547 con 20 — corrispondenti a livelli di 667, 669 e 673 punti su 720.

---

## La graduatoria

**S15 — Tetto a scala di stagione.** È  $S_5$  valutata a  $n$  grande. Per divari piccoli,  $\rho \approx 1 + \delta$  con  $\delta$  piccolo, e la probabilità cresce come la ripartizione di una normale in  $\sqrt{n} \cdot \delta$ . Da qui la scala: il numero di frecce necessario per una data affidabilità cresce come  $1/\delta^2$ .

**S16 — Costo del referto.** Il rapporto fra le frecce necessarie leggendo anelli e leggendo coordinate risulta costante a 1,43–1,44 su tutti i divari esaminati. In punti percentuali la stessa quantità ha un massimo interno, perché l'affidabilità è limitata fra  $1/2$  e 1 mentre il conteggio di frecce non lo è.

**S17 — Effetto evento.** Modellato come un numero di punti comune a tutti i partecipanti a una gara, con deviazione standard  $\sigma_E$ , sulla stessa scala su cui lo stimatore lo cerca. L'effetto è una favorevolezza: un  $e_j$  positivo è una gara facile e alza il punteggio di tutti, come nel modello. Applicarne uno di  $e$  punti a un arciere di livello  $L$  significa farlo tirare con l'ampiezza che corrisponde a  $L + e$ , così che valga  $e$  punti attesi per chiunque non sia vicino al massimo del round: la mappa da ampiezza a punteggio atteso non è lineare, e un fattore comune sulla dispersione varrebbe punti diversi per arcieri di livello diverso. La scala di riferimento è il round da 72 frecce: sui round da 36 e da 144 della Parte terza l'effetto è riscalato in proporzione alla lunghezza, e produce allora la stessa ampiezza di tiro su tutte e tre — è la ragione per cui i numeri di quei formati si confrontano fra loro. Per due arcieri che condividano la gara lo spostamento è lo stesso e si cancella nel confronto; per arcieri su gare diverse sono estrazioni indipendenti e si sommano al rumore. Da cui il comportamento misurato: sovrapposizione piena annulla quasi del tutto l'effetto, sovrapposizione nulla lo lascia intero.

**S18 — Connessione del disegno.** Con effetti evento fissi, il confronto fra due atleti che non condividano alcuna gara, né direttamente né attraverso una catena di atleti intermedi, non è stimabile: l'efficienza del disegno è zero. Con effetti casuali il confronto è possibile ma degradato con continuità. La scelta fra i due modelli non è tecnica ma sostanziale. Qui la stima è a effetti fissi, ai minimi quadrati, con le difficoltà vincolate a somma nulla e la connessione pretesa: se il grafo si spezza, la funzione non risponde. Un modello a effetti casuali risponderebbe anche allora, ma quel numero verrebbe dalla distribuzione assunta sugli eventi più che dalle gare osservate.

**S19 — Piazzamento rumoroso e propagazione.** Il tabellone si costruisce sull'ordine osservato dopo la qualificazione, non su quello vero. Detto  $\Pi$  l'ordine osservato e  $\Pi$  *quello vero*, ogni struttura è una funzione che consulta  $\Pi$  un numero  $k$  di volte:  $k = 1$  per il tabellone fisso, che lo consulta una volta sola per comporre il tabellone iniziale, e  $k = 6$  per il *reseeding*, che lo riconsulta prima di ciascuno dei sei turni. La perdita rispetto al caso  $\Pi = \Pi$  cresce con  $k$ , ed è misurata a  $+1,24 \pm 0,47$  per  $k = 1$  e  $+5,83 \pm 0,53$  per  $k = 6$  su campo compatto.

**S20 — Stima della graduatoria di stagione.** Il punteggio dell'arciere  $i$  nella gara  $j$  è  $y_{ij} = \mu + a_i + e_j + \varepsilon_{ij}$ , con  $a_i$  il livello dell'arciere ed  $e_j$  la difficoltà della gara in punti. I due insiemi di effetti si stimano insieme per minimi quadrati sulle sole coppie arciere-gara effettivamente disputate, con il vincolo  $\sum_j e_j = 0$  che rende la soluzione unica: aggiungere una costante a tutte le  $a_i$  e sottrarla a tutte le  $e_j$  lascerebbe infatti le previsioni identiche. Sopra la stima si applica il restringimento verso la media, con peso  $n_i \tau^2 / (n_i \tau^2 + \sigma^2)$ , dove  $\sigma^2$  è la varianza dei residui e  $\tau^2$  si ottiene togliendo alla varianza delle stime grezze la componente d'errore. È un'approssimazione, e conviene dirlo: il peso esatto dipende dalla varianza della stima del singolo atleta, che in un disegno sbilanciato non è funzione del solo numero

di presenze ma dell'intera matrice di disegno — conta anche con chi si è condiviso la gara. Uso  $n_i$  perché è la forma che si rifà a mano; per un restringimento esatto servirebbe la diagonale della matrice di covarianza dei coefficienti. Stimare in due passaggi — centrare ciascun atleta sulla propria media e poi mediare i residui per gara — non è equivalente quando le presenze sono sbilanciate: sottrae a ogni atleta la media degli effetti del suo calendario, e con calendari diversi quella costante non è comune. Non è un errore che si attenua con più dati: esiste un caso con cinque arcieri, cinque gare, nessun rumore e grafo di partecipazione connesso in cui la procedura sequenziale mette un arciere da 698 davanti a uno da 700, mentre la stima congiunta recupera l'ordine esatto. Il vincolo identifica le differenze fra livelli soltanto entro una componente connessa del grafo arciere-gara: fra componenti disgiunte l'origine non è nei dati, e il restringimento verso la media generale è un'ipotesi che le allinea, non una misura.

**S20b — Il peso del restringimento.** Detta  $\mu$  la media delle stime grezze del gruppo,  $m_i$  la stima grezza dell'atleta  $i$  e  $n_i$  il numero di gare a cui ha partecipato, la stima ristretta vale

$$\hat{m}_i = \mu + w_i \cdot (m_i - \mu), \text{ con } w_i = n_i \cdot \tau^2 / (n_i \cdot \tau^2 + \sigma^2)$$

dove  $\tau^2$  è la varianza fra i livelli veri degli atleti e  $\sigma^2$  la varianza del punteggio di una gara attorno al proprio livello. Entrambe si stimano dai dati:  $\sigma^2$  dalla dispersione degli scarti residui dopo la correzione per difficoltà,  $\tau^2$  dalla varianza delle stime grezze meno  $\sigma^2$  diviso il numero medio di gare. Il peso cresce col numero di gare e tende a uno quando i dati abbondano; con  $\tau$  molto maggiore di  $\sigma$  il restringimento è lieve anche per pochi dati, ed è il caso dell'esempio del capitolo 47, dove tutti gli atleti hanno tirato a tre gare o più e i pesi stanno tutti sopra 0,97. Il restringimento morde invece quando un atleta ha una o due gare sole, ed è lì che serve.

**S21 — Appaiamento dei contendenti.** La precisione del confronto fra i due candidati al primo posto dipende dalla frazione di gare che condividono: se tirano nelle stesse gare, la difficoltà comune si cancella nella loro differenza — esattamente finché nessuno dei due è così vicino al massimo del round da non poter salire di tutto l'effetto, perché il generatore lì tronca e lo spostamento realizzato non è più lo stesso per entrambi. Va tenuta distinta dalla connessione globale del disegno, che è una questione di identificabilità e non di precisione: senza connessione le origini di due componenti non sono nei dati, e con la stima a due vie un calendario a blocchi disgiunti non è migliore di un sorteggio. Appaiare i primi otto recupera gran parte del divario fra un calendario strutturato e uno sorteggiato, e alla precisione della simulazione corrente estendere l'appaiamento oltre gli otto non mostra un beneficio addizionale risolvibile: il recupero vale +0,8 con otto appaiati, +0,7 con sedici e +0,8 con tutti e trentadue, misurato come contrasto appaiato su dodicimila stagioni in cui cambia soltanto il calendario, con un errore di mezzo punto su ciascuno. Resta lo stesso se gli otto sono scelti dalla graduatoria della stagione precedente invece che dall'ordine vero: il contrasto vale +0,9 invece di +0,8, dentro lo stesso errore. Nella griglia simulata, estendere l'appaiamento oltre i primi otto non produce un beneficio addizionale chiaramente risolvibile alla precisione corrente. Otto è quindi un punto operativo ragionevole, non una soglia di saturazione dimostrata. Il criterio di progetto di un calendario resta lo stesso: non che tutti si incontrino, ma che si incontrino i contendenti.

---

## Convenzioni

Arrotondamento del numero di frecce. Tutte le tabelle di inversione — «quante frecce per una data affidabilità» — riportano il minimo intero  $n$  tale che l'affidabilità sia maggiore o uguale al bersaglio. Chi ri-

calcolasse troncando l'inversione continua otterrebbe valori inferiori di uno, ai quali l'affidabilità sta ancora sotto il bersaglio.

Pareggi. Dove una regola può pareggiare, i pareggi sono accreditati a metà — che equivale in media a risolverli con un sorteggio equo. La frequenza di pareggio è comunque riportata separatamente per ogni regola, perché una regola che vale 58 su cento pareggiando quattro volte su dieci non è la stessa cosa di una che vale 58 senza pareggiare mai.

Ancoraggio. Ogni divario espresso in punti è ancorato a un livello dichiarato, perché la traduzione fra punti e dispersione non è lineare: un punto vale l'1,81 per cento di dispersione a 700, l'1,32 per cento a 680. La cella di riferimento della Parte prima è 690 contro 680.

Popolazione. Ogni numero dichiara se sia condizionato alla parità a set o no. Le due popolazioni danno risposte molto diverse — è il contenuto del capitolo 7 — e confonderle è l'errore più facile del volume.