

Bruno de Finetti e il Calcolo delle Probabilità

Carla Rossi
allieva di Bruno de Finetti

Giornata in onore di Bruno de Finetti
Centro “Beniamino Segre” dell’Accademia dei Lincei
Roma 30 aprile 2015



Chi sono io di Bruno de Finetti (1981)

*"Chi sono?", la prima cosa che mi sembra di dover dire come punto di partenza è che **di me stesso**, come persona qualunque, **m'importa assai di meno** che di ciò che attiene al **benessere collettivo, all'equilibrio ecologico** secondo la linea tenacemente difesa da Aurelio Peccei, **al progresso sociale e civile** secondo la linea ispirata a Lelio Basso (membro tra l'altro del tribunale Russell); linea cui vorrei che tutti mirassero per aver diritto a goderne quanto a ciascuno può ragionevolmente spettare. **Uno per tutti e tutti per uno**, senza eccessive differenze o rivalità tra individui o classi o nazioni: **rivalità utili soltanto se mirano a migliorare ovunque il benessere collettivo** anziché curarsi soltanto di quello **egoisticamente (e miopemente)** individuale o settoriale o classista.*



- *Quanto al mio modo di pensare, di prospettarmi i problemi ed esporre le mie tesi, dirò che **cerco sempre di rendere quanto più possibile chiari e semplici e "naturali" e "intuitivi"** - magari presentandoli in modi concreti e divertenti - i concetti e i ragionamenti in ogni campo e, ovviamente, soprattutto in quello della probabilità che particolarmente mi interessa, e che è, purtroppo, una delle nozioni più esposte al rischio di velleitari fraintendimenti e distorsioni e addirittura **travisamenti di ogni peggior specie**.*

Impossibile parlare di tutti i contributi

- Dato il tempo a disposizione, mi limiterò a parlare di:
 - Probabilità, incertezza e informazione;
 - Coerenza;
 - Teorema di de Finetti;
 - Critica all'additività completa.
- Dato lo stile di de Finetti i temi non possono essere separati rigidamente e le sue parole sono incomparabili.

Trascurerò importanti questioni alle quali de Finetti ha dato rilevanti contributi

- Processi stocastici, legami con *distribuzioni indefinitamente scomponibili e stabili*.
- Modelli per problemi applicativi, tra cui i lavori sulla *Genetica di Popolazioni*. Lì, oltre a introdurre la rappresentazione grafica oggi chiamata, in tutto il mondo, “*triangolo di de Finetti*”, ha utilizzato, 50 anni prima degli altri, i modelli con *generazioni sovrapposte*.
- Distribuzioni in più dimensioni, correlazioni ammissibili (talune idee riprese in considerazione vari decenni più tardi..)
- Lavori sulle *scienze attuariali e la finanza matematica*, per cui molti hanno dichiarato che avrebbe meritato il premio Nobel per l'Economia (riconoscimento conseguito da altri matematici...)

Evento (da un suo libro...AMASES)

- Un evento, o la proposizione che lo esprime, è un'entità logica capace di assumere due modalità: **vero** e **falso**; corrisponde cioè ad una questione così formulata da ammettere due sole risposte, **si** o **no**, o ancora, ad una divisione netta in due campi di tutte possibilità.
- Per esprimerli in modo **intuitivo**, e che risulterà strettamente in relazione coll'argomento del corso, potremo dire che un **evento è definito** quando, stabilendo che una **scommessa** sarà vinta se l'evento risulterà vero e perduta se risulterà falso, non si lascia **alcun adito a casi ambigui o discutibili**.

Che cos'è la probabilità: definizioni

LEZIONI SULLA PROBABILITA' di B. de Finetti Univ. Trieste 1932-33
57 pagine relative alle prime 8 lezioni

- *Dice un'antica sentenza latina, "tot capita, tot sententiae";*
- *in nessun campo essa è tanto vera quanto nella teoria delle probabilità, e fin dai principi, fin da questa stessa domanda sul significato della probabilità.*
- *Tuttavia, fra un matematico che la definisca come rapporto tra il numero di casi favorevoli e possibili,*
- *uno statistico che la interpreti come un valore più o meno ideale della frequenza,*
- *e l'uomo della strada che dica "è la sensazione che mi guida in tutta la vita",*
- *non esito a dire che la risposta migliore, più completa, più sensata, è proprio quella dell'uomo della strada.*

Definizioni vaghe e corrette o distorte

- E' un po' troppo vaga, è vero, ma è la sola che riallacci il **concetto di probabilità** a quelle che sono le sue radici profonde, intuitive, vive, da cui non possiamo pensarlo avulso senza cadere in quel ginepraio di dubbi difficoltà oscurità che fecero del calcolo delle probabilità il ramo più discusso di tutte le matematiche.
- **E' un po' troppo vaga, ma si può analizzare e precisare il suo significato sino a mostrare che basta a dedurre in modo perfettamente rigoroso tutta la teoria matematica delle probabilità.**
- In particolare, anche i due criteri di valutazione delle probabilità basati sulla numerazione di casi favorevoli e possibili e sull'osservazione di frequenze, **criteri** che hanno un'importanza grandissima, ma che non si prestano ad essere assunti come definizione, perché, come **definizioni**, avrebbero il torto di non scendere fino in fondo all'essenza dei concetti, e di abbracciare ciascuno non più di una piccola parte di problemi di probabilità.

Definizione rigorosa di probabilità

- Diremo dunque per definizione “probabilità di un evento E (per un dato individuo)” il numero $p = P(E)$ che rappresenta il valore che ha, per lui, il possesso di una lira subordinato al verificarsi di E ; si può anche dire che p rappresenta la “quota di scommessa” sull’evento E .
- Fissata ad esempio la quota in $p = 0,30$, egli accetterà indifferentemente di promettere 10, 100, 1000 lire nel caso che si verifichi l’evento E a chi gliene versa 3, 30, 300, o inversamente a pagare 3, 30, 300 lire a chi si obblighi a versargliene 10, 100, 1000, sempre nel caso considerato.

Approccio pragmatico

- de Finetti, nel passare da un concetto vago a una precisazione matematica, si preoccupa di dare una definizione operativa e non puramente formalistica.
- Il meccanismo della scommessa (invertibile) è comprensibile a tutti. Si collega palesemente all'esperienza di de Finetti nel campo delle assicurazioni, che è anche il campo (a parte i giochi d'azzardo) in cui è più antica l'applicazione economica concreta, anche se empirica, della probabilità.

Evoluzione

- Negli anni del Dopoguerra, de Finetti elaborò un ulteriore meccanismo per la valutazione delle probabilità: quello della “**penalizzazione quadratica**”. (E' possibile sia stato influenzato dai concetti di Wald, da lui citato, anche senza aderire alla sua visione delle “decisioni”).
- Questo meccanismo, ovviamente equivalente all'altro, tende a eliminare l'ipotesi di presenza di un “competitore” nella scommessa e dell'invertibilità della medesima. E' anche di più agevole implementazione operativa, tant'è che de Finetti lo sperimentò (personalmente, e con l'aiuto di “volontari”...) nel suo “**concorso pronostici calcistici**” degli anni Sessanta.
- A riprova dell'interesse definettiano alla costante concretezza e applicabilità dei concetti teorici.

Esempio importante ricordato in

Chi sono io?

- *..Col medesimo intento di imparare ad usare valutazioni di probabilità, di abituare le persone a pensare e ragionare (e conseguentemente, comportarsi) in base a valutazioni (ragionate, ma naturalmente soggettive) di probabilità, è stato ripetuto per diversi anni all'Università di Roma un esperimento di pronostici probabilistici con riferimento ai risultati delle partite del campionato di calcio.....*
- *....secondo il mio punto di vista l'esperienza era educativa perché non solo non era basata sul banale e antieducativo malvezzo del "tirare a indovinare" (come al Lotto e al Totocalcio), ma, al contrario, obbligava a indicare la probabilità (dei possibili risultati) numericamente (p_1, p_2, p_X).*

Originale da Pittsburgh (1961-1969)

Esperimento valutazioni probabilità

Sarà effettuato per le partite del girone di ritorno del campionato di calcio, Serie A, o comunque per le partite incluse nei primi 9 posti delle schede SISAL dal 5/2/61 alla fine del campionato (incluse eventuali giornate con partite internazionali o simili).

Potranno prendervi parte gli studenti del corso di Matematica attuariale, ed eventualmente altri, compatibilmente col volume del lavoro di spoglio-conteggio che sarà fatto con la calcolatrice 610 se ciò risulterà pratico (altrimenti a mano su dati arrotondati in base a una tabella).

I partecipanti potranno accordarsi per raccogliere una quota da ciascuno partecipante, per formare dei premi da distribuire in base al punteggio finale alla chiusura dell'esperimento. Di tale eventuale parte finanziaria si occuperanno essi stessi. L'Istituto ha interesse esclusivamente all'esperimento del metodo di espressione delle opinioni riguardo alle probabilità, e fornirà il punteggio finale indicando come, in base ad esso, un eventuale fondo vada ripartito fra i migliori risultati (si può fare o in proporzione al punteggio finale, oppure - per avere distacchi più significativi - in proporzione alla eccedenza sul punteggio medio, per quelli soli che lo superano: ciò dovrà esser deciso dai partecipanti).

Modalità del "pronostico"

Non si tratta propriamente di pronostici ma di valutazioni di probabilità. Ogni partecipante riempirà ogni settimana una ~~e~~ una sola schedina (usando quella del Totocalcio) indicando per ciascuna delle prime nove partite la probabilità che attribuisce ai tre risultati 1-X-2 (vittoria in casa - pareggio - vittoria esterna), espresse in % con numeri di due cifre (p.es. 46 33 21 significa prob. di "1", 46%; di "X", 33%; di "2", 21%). Indicare i tre numeri nelle tre coppie di colonne a destra in ciascuna delle tre sezioni (Figlia, che sarà restituita per ricevuta; Spoglio che rimane al raccoglitore; Matrice che va al Centro elettronico per l'elaborazione).

La somma dei tre numeri deve dare 100 (eventuali istruzioni diverse saranno date in seguito se per l'elaborazione risultasse opportuno limitare le scelte o allargarle, p.es. ammettendo solo numeri terminanti in 0 o 5, come 45 35 20 anzichè 46 33 21, oppure anche numeri con somma 99 o 98 che andrebbero automaticamente interpretati come aumentati ciascuno di $1/3$ o di $2/3$).

Ogni partecipante avrà un suo numero progressivo, e lo indicherà su ogni sezione della scheda (nel posto libero delle partite 10-13).

□ □ □ □ □ □

Per ogni partita, da zero a 10.000 punti, in base alle probabilità ~~distribuite~~ ai due risultati non verificatisi: detti m ed n i due numeri esprimenti tali probabilità in %, il punteggio sarà:

$$10.000 - (m+n)^2 + mn .$$

(Geometricamente: è il quadrato della distanza fra la valutazione prescelta e il risultato effettivo).

Roma, 2 febbraio 1961

B. de Finetti

CONCORSO

23



COMITATO OLIMPICO
NAZIONALE ITALIANO

Totoc

“AL SERVIZIO I

PARTITE DEL 26-2-1961

FIGLIA

N.	Squadra 1ª	Squadra 2ª	Concorso 23 del 26-2-61					
1	Catania	Juventus			3	0	3	5
2	Fiorentina	Padova			8	5	1	0
3	Lanerossi	Inter			2	0	3	0
4	Lazio	Udinese			6	5	2	5
5	Lecco	Bari			5	5	3	0
6	Milan	Sampdor.			8	5	1	0
7	Napoli	Bologna			4	0	2	0
8	Spal	Roma			2	0	3	5
9	Torino	Atalanta			4	0	3	5
10	Catanz.	Simm. Mon.						
11	Venezia	Messina						
12	Livorno	Lucchese						
13	Marsala	Trapani						
Riserve	1 Sambenedett. Como							
	2 Spezia Mestrina							

CONSERVATE IL TAGLIANDO FIGLIA DELLA SCHEDA VINCENTE!

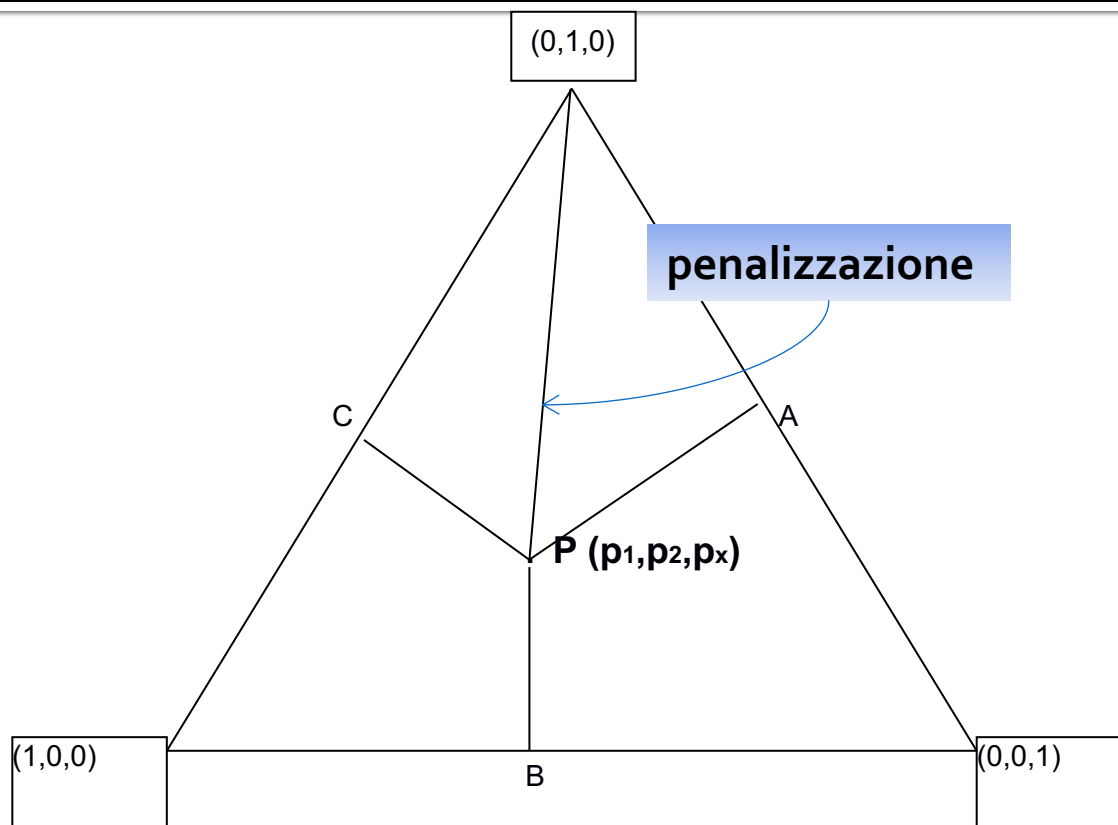
Raci

PERUGINA

SEMPRE PIÙ GRADITI

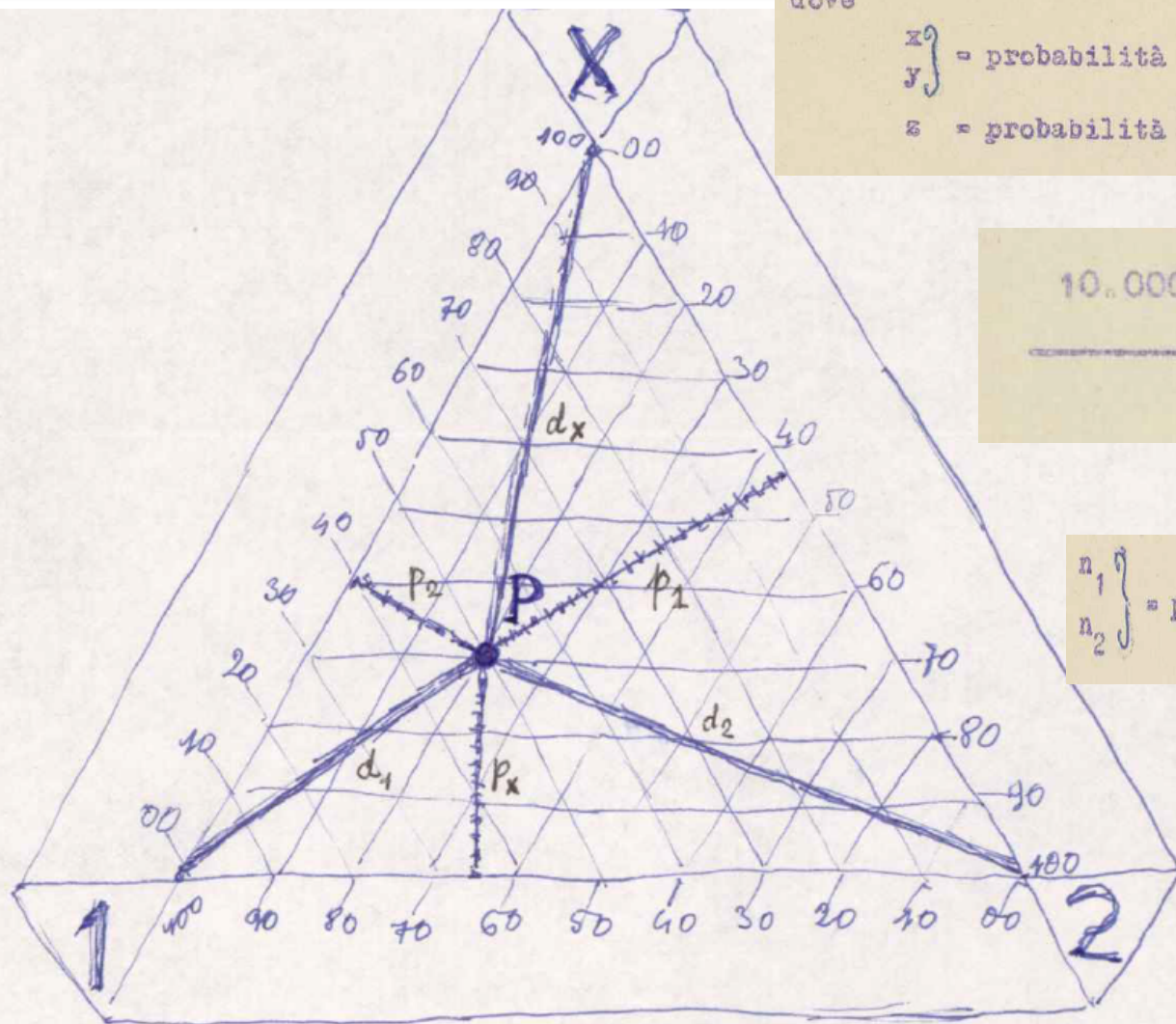
Giocata effettuata da Bruno de Finetti con il risultato delle partite segnato per il calcolo del punteggio

Rappresentazione geometrica (triangolo di de Finetti) (probabilità naturali non percentuali)



Tenendo conto che nel pronostico le probabilità si esprimono, per comodità, in forma percentuale bisogna moltiplicare le coordinate baricentriche per 100 e la penalizzazione (quadrato di una distanza) si esprime come ha comunicato ufficialmente Bruno de Finetti considerando anche che la somma delle coordinate baricentriche in % vale 100.

Rilanciato il pronostico nel 1980 rivolto ai giornalisti (triangolo disegnato da de Finetti)



dove

$\left. \begin{matrix} x \\ y \end{matrix} \right\}$ = probabilità attribuita ai segni non verificatesi
 z = probabilità attribuita al segno verificatesi.

$$\frac{10.000 - \frac{x^2 + y^2 + (100-z)^2}{2}}{100}$$

$\left. \begin{matrix} n_1 \\ n_2 \end{matrix} \right\}$ = probabilità attribuite ai segni non verificatesi

$$\frac{10.000 - (n_1^2 + n_2^2 + n_1 n_2)}{100}$$

premi

- 16°) La quota di L.2.000 verrà ripartita settimanalmente nella misura di L.90 , per le venti settimane, che andranno a formare il monte premi settimanale; il rimanente andrà a formare un monte-premi speciale
- a) monte premi settimanale: sarà ripartito in proporzione allo scarto del punteggio di ciascun partecipante, da un punteggio fisso di 540 (°) alla settimana. La liquidazione avverrà alla fine dell'anno;
- b) monte premi speciale: sarà suddiviso alla fine del concorso, sulla classifica finale dei partecipanti, calcolando gli scarti dal livello di 11.000(°) -

(°) suscettibili di opportuni ritocchi, che saranno tempestivamente comunicati.

TUTTI I CALCOLI RELATIVI AL CONCORSO PRONOSTICI VIENGONO EFFETTUATI CON LA MACCHINA CALCOLATRICE ELETTRONICA "610" I.B.M.

CONCORSO N. 18 DEL 10/12/61
1* GIORNATA

		N.	PUNTI	VINCITA L.
1	PALLOTTA	2	656.25	176
2	DI TORO	14	650.00	167
3	CORNELIO	5	639.00	151
4	ROSATI	19	628.50	134
5	TORTORICI	3	623.75	126
6	MASSARI	24	621.25	123
7	TIBERI	21	619.62	120
8	MANCINI	25	615.00	114
9	DE FINETTI	4	614.21	113
10	SIROTTI	22	613.00	111
10	DE RISIO	16	613.00	111
12	GASPARRI	12	610.54	106
13	CALI'	15	607.25	102
14	RAMBERTI	1	601.25	93
15	DIANA	27	600.97	91
15	FURST	26	600.97	91
17	D'ORAZIO	20	598.86	88
18	VARRONE	8	594.00	82
19	ROMAGNOLI	9	587.25	72
20	CARBONI	11	583.18	64
21	GAMBONI	10	582.00	64
22	RETTAROLI	23	577.50	56
23	PACI	6	574.00	52
24	BALDINI	28	570.25	46
25	DAQUANNO	18	565.50	38
26	DE FRANCESCO	17	550.25	15
27	GERMANI	13	535.75	
28	ROBERTO	7	535.28	

M = 598.83

Si parla anche di valutazione media

SCHEDINA MEDIA

ATALANTA	ROMA	39.48	35.52	24.99	X
FIorentINA	LECCO	67.19	20.24	12.56	1
INTER	CATANIA	68.25	20.69	11.05	X
PALERMO	L.R.VICENZA	49.05	32.66	18.28	1
SAMPDORIA	BOLOGNA	37.41	33.75	28.83	2
SPAL	MANTOVA	43.95	34.32	21.72	1
TORINO	MILAN	35.02	34.58	30.38	X
UDINESE	JUVENTUS	25.26	34.22	40.51	1
VENEZIA	PADOVA	45.94	34.87	19.18	X

PUNTEGGIO DELLA SCHEDA MEDIA 615.04

DALLA PROSSIMA SETTIMANA OGNI CONCORRENTE DOVRA' INDICARE SULLA PROPRIA SCHEDINA IL NUMERO D'ORDINE RIPORTATO NELLA TERZA COLONNA

Si introduce sia la media della scheda delle probabilità valutate, ma anche del **punteggio** che è un **numero aleatorio** di cui è utile parlare e allarga il campo.

Il punteggio ottenuto da de Finetti è il secondo sotto quello della schedina media.

2* GIORNATA

	RISULTATI DELLA GIORNATA	CLASSIFICA GENERALE	VINCITA DI GIORNATA L.	VINCITA TOTALE
1	DI TORO 1* 771.50	1421.50	143	310
2	FALLOTTA 6* 718.25	1374.50	110	287
3	CORNELIO 2* 728.25	1367.25	116	267
4	DE FINETTI 3* 726.47	1340.68	115	228
5	TORTORICI 13* 707.00	1330.75	103	230
6	DE RISIO 9* 716.75	1329.75	109	220
7	VARRONE 4* 724.00	1318.00	113	195
8	DIANA 11* 711.00	1311.97	105	198
9	SIROTTI 16* 695.00	1308.00	96	207
10	RAMBERTI 14* 704.00	1305.25	101	194
11	GASPARRI 17* 689.00	1299.54	92	199
12	ROMAGNOLI 12* 709.75	1297.00	105	177
13	MANCINI 19* 678.91	1293.91	86	200
14	TIBERI 22* 673.88	1293.50	83	204
15	BALDINI 5* 723.00	1293.25	113	159
16	PACI 8* 717.50	1291.50	109	161
17	ROSATI 24* 659.25	1287.75	74	209
18	DAQUANNO 6* 718.25	1283.75	110	149
19	CARBONI 15* 701.21	1283.39	99	163
20	CALI' 21* 674.00	1281.25	83	185
21	D'ORAZIO 18* 679.80	1278.66	86	176
22	MASSARI 25* 647.25	1268.50	66	190
23	DE FRANCESCO 10* 712.75	1263.00	107	123
24	RETTAROLI 20* 675.00	1252.50	83	140
25	FURST 26* 601.97	1202.94	30	131
26	GERMANI 23* 662.75	1198.50	76	76
27	GAMBONI 601.97	1183.97	--	64
28	ROBERTO 601.97	1137.25	--	--

$$M = 697.17$$

SCHEDINA MEDIA

ATALANTA	FIorentina	28.73	37.96	33.29	X
BOLOGNA	UDINESE	64.46	24.00	11.53	1
JUVENTUS	VENEZIA	58.88	28.80	12.30	1
L.R.VICENZA	SAMPDORIA	38.15	34.42	27.42	X
LECCO	INTER	23.00	33.00	44.00	2
MANTOVA	CATANIA	44.96	35.73	19.30	X
MILAN	SPAL	64.19	24.34	11.46	1
PALERMO	TORINO	37.76	35.40	26.83	1
ROMA	PADOVA	62.69	25.19	12.11	1

PUNTEGGIO DELLA SCHEDINA MEDIA 708.82

LE VINCITE DELLA SCORSA SETTIMANA NON ERANO APPROSSIMATE IN MODO CORRETTO
IN QUESTA CLASSIFICA SONO RIPORTATI I RISULTATI CORRETTI

Elenco delle squadre partecipanti al Concorso valutazioni
probabilità calcio 1967-68

ANCONA	Collegio Einaudi	Bianchi Renzo	N. codice	11
		Garofoli Gioacchino	"	12
		Rocchetti M. Laura	"	13
		Soratte	"	14
Ancona	Ist. Olivetti	Belloni Valeriano	"	16
		Ercolani Paolo	"	17
		Pettenati Paolo	"	18
		Ercolani Maurizio	"	19
ROMA	C. E. Alitalia	De Propriis Giorgio	"	21
		Mancini Luigi	"	22
		Mattei Mariano	"	23
		Pompei Tommaso	"	24
FIRENZE		Sellieri Augusto	"	26
		Biggeri Luigi	"	27
		Ricci Renzo	"	28
		Vope Ernesto	"	29
MILANO		Muzio Giovanni	"	31
		Nicola Piercarlo	"	32
		Nicola Serenella	"	33
		Quadrio Alberto	"	34
NAPOLI		Fausto Domenicantonio	"	36
		Leccisotti Mario	"	37
		Pedone Antonio	"	38
		Pica Federico	"	39
ROMA	Fac. Matematica	Beltrami Anna	"	41
		Brizio Luisa	"	42
		De Finetti Bruno	"	43
		Pala Gianfranco	"	44
ROMA	Fac. Economia	Acocella Nicola	"	46
		Miconi Bruno	"	47
		Regoli Giorgio	"	48
		Retteroli Riccardo	"	49

Torino	Brosio Giorgio	N. codice	51
	Fornengo Graziella	"	52
	Penizza Roberto	"	53
	Vercelli Sandro	"	54
Trieste	Crisma Lucio	"	56
	Depollo Arrigo	"	57
	Strudthoff Mario	"	58
	Wedlin Attilio:	"	59

Gli ultimi anni mostrano i punteggi elaborati al Calcolatore IBM di Matematica e i nomi di molti miei amici. Gli ultimi elaborati scansionati da Princeton sono del 1969.

Dal concorso si imparano tante cose importanti sulle basi della probabilità

- la probabilità è *soggettiva* (ciò non toglie che una valutazione di probabilità possa essere anche ampiamente intersoggettiva); il soggetto che la valuta ha a disposizione alcune regole di *coerenza*, contravvenendo alle quali incorre in perdite certamente superiori al minimo possibile (*assiomi* di Kolmogorov eccetto completa additività);
- ogni valutazione di probabilità dipende dall'*informazione* che si possiede; in altre parole è una “*probabilità condizionata*”, **concetto fondamentale**;
- l'acquisizione di notizie (nell'esempio: quelle dovute ai risultati delle partite già effettuate; ma anche altre) fornisce ulteriore *informazione* e permette di *modificare la valutazione delle probabilità* sugli eventi futuri; gli strumenti analitici per questi calcoli sono il teorema di Bayes e si può anche introdurre la *funzione di verosimiglianza*;



- le nuove informazioni (sulle partite effettuate, sulle squadre) influiscono non solo tutte insieme (alla fine del campionato;, ovvero quando la ricerca è conclusa), ma ognuna influisce singolarmente e dinamicamente (qui si ritrovano le basi teoriche [per l'induzione e l'approccio bayesiano alla statistica, mucchio o non mucchio](#));
- una considerazione diversa, ulteriormente istruttiva, riguarda l'ordine di acquisizione delle informazioni sui risultati delle partite; per taluno può essere influente ai fini della valutazione delle probabilità per la prossima partita; per talaltro invece può non esserlo.
- Questo secondo caso corrisponde alla situazione di “[scambiabilità](#)” di de Finetti, inizialmente chiamata [equivalenza](#), che, se verificata, agevola fortemente molte valutazioni di inferenza statistica. Infatti l'informazione globale che si possiede non dipende dall'ordine con cui si sono ottenuti le singole informazioni dello stesso tipo che la costituiscono.

Numeri aleatori o incogniti

LEZIONI SULLA PROBABILITA' di B. de Finetti Univ. Trieste 1932-33
57 pagine relative alle prime 8 lezioni

La funzione di penalizzazione permette di introdurre i **numeri aleatori**.
Si introduce anche la distribuzione di un numero aleatorio dato che si parla di medie di giocate e di perdite.

LEZIONE VI.

20. Numeri aleatori.

Passiamo dunque a trattare dei "numeri aleatori": concetto assai fecondo, esplicitamente introdotto soltanto di recente, benchè implicitamente considerato da quando esiste il calcolo delle probabilità ().

() La denominazione più usata ma meno indovinata, è "variabile casuale"; preferiamo dire "numero aleatorio" (denominazione specialmente usata dai francesi), perchè vedremo che, piuttosto che di un numero variabile, si potrebbe caso mai parlare di un numero incognito.

L'utilità di questo concetto è grande, non perchè esso costituisca qualcosa di veramente, essenzialmente nuovo, ma in quanto molti importanti problemi che si potevano porre e risolvere anche senza conoscerlo vengono presentati e chiariti in modo suggestivo e intuitivo.

Distribuzioni

- Ci sono 3 tipi di distribuzioni sulla retta, facilmente generalizzabili:
- Il primo tipo è quello di una distribuzione totalmente discontinua (salti in numero finito, $d_1 + d_2 + \dots + d_n = 1$, oppure numerabile, $d_1 + \dots + d_n + \dots = 1$;). In particolare una distribuzione statistica (con N individui, N finito) è sempre discontinua (e i salti sono al più N , tutti uguali a $1/N$ o multipli).
- Il secondo tipo è quello di una distribuzione assolutamente continua: ciò significa che $F(x)$ (la quale, essendo monotona, è derivabile quasi ovunque) è l'integrale della propria derivata $f(x)$; la $f(x)$ si dice (nello stesso senso intuitivo della meccanica) **densità** della distribuzione in x . Col linguaggio degli infinitesimi, si direbbe che $f(x)dx$ rappresenta la massa contenuta tra x ed $x + dx$.
- Il terzo tipo di distribuzioni che possono aversi, oltre quelle discontinue e continue, ha un carattere piuttosto "patologico": la massa è distribuita in modo continuo in un insieme di misura nulla.

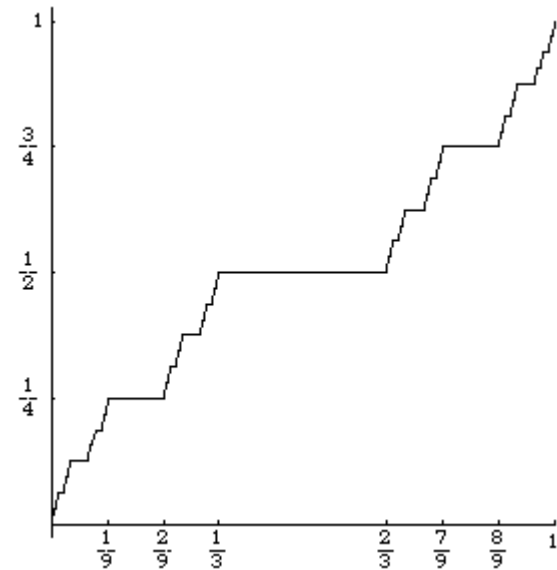
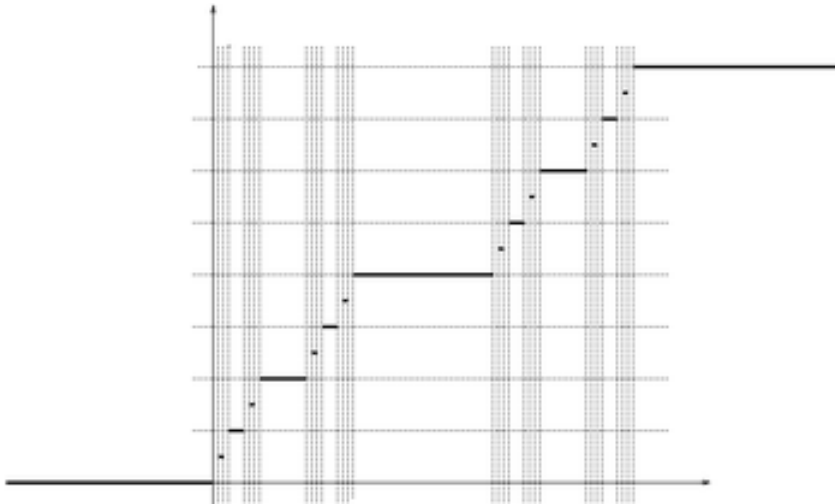


Teoria delle Probabilità del 1970

- In questo libro il terzo caso veniva esemplificato attraverso l'insieme di Cantor, che è uno dei più semplici oggetti frattali cui si possa far riferimento (anche Cantor era un precursore).
- Nonostante la sua semplicità, presenta le tipiche caratteristiche degli oggetti frattali. Esso è costruito partendo da un segmento di lunghezza unitaria cui sono iterativamente eliminati alcuni tratti.
- Sia E_0 l'intervallo $[0,1]$. Sia E_1 il set ottenuto togliendo dal centro del segmento E_0 un segmento di lunghezza $1/3$, cosicché E_1 sia composto dai due intervalli $[0, 1/3]$ e $[2/3, 1]$.
- Ripetendo la stessa operazione con questi due intervalli si ottiene il set E_2 composto dagli intervalli $[0, 1/9]$, $[2/9, 1/3]$, $[2/3, 7/9]$, $[8/9, 1]$. In generale E_k è ottenuto togliendo la parte centrale di lunghezza $1/3$ del segmento E_{k-1} ; dunque E_k è formato da 2^k intervalli ciascuno di lunghezza pari a 3^{-k} .
- Il set F è ottenuto dall'intersezione di tutti gli E_k .

Rappresentazione geometrica

Sotto sono riportate le funzioni di ripartizione ottenute in diversi casi di approssimazione. L'insieme ha misura nulla di Lebesgue sull'intervallo, ma possiede un'infinità continua di punti con probabilità totale pari a 1.



Distribuzioni come misture in generale

- la più generale distribuzione $F(x)$ si ottiene come combinazione di distribuzioni F_1, F_2, F_3 dei tre tipi indicati: $F(x) = a_1F_1(x) + a_2F_2(x) + a_3F_3(x)$ con $a_i \geq 0$, $a_1 + a_2 + a_3 = 1$.
- In genere, date delle distribuzioni definite in un medesimo spazio, una loro combinazione lineare (con coefficienti non negativi di somma = 1) dà ancora una distribuzione (che diremo senz'altro **mistura** di esse).

Verso il teorema di de Finetti

(Funzione caratteristica di un fenomeno aleatorio. In: *Memorie della R. Accademia dei Lincei*, (IV), fasc. V (1930), p.86-133.)

- Quelli che abbiamo definito col nome di «fenomeni aleatori» sono i fenomeni le cui prove sarebbero, nell'ordinaria terminologia, *indipendenti e con probabilità costante ma incognita*.
- Ora, il parlar di *probabilità incognite* è, secondo la concezione soggettiva della probabilità, cosa **priva di senso**, e in ogni caso **oscura** e **capziosa**.
- Parlare di prove *indipendenti*, nel caso che ci interessa, è per lo meno improprio, dato che in seguito all'esito delle prove precedenti la probabilità delle prove successive si modifica.

Proprietà di scambiabilità

Ad ogni modo, dall'uso che si fa ordinariamente della nozione, non certo troppo chiara e felice, di *fenomeno «a prove indipendenti e con probabilità costante ma incognita»* risulta che da essa si ritiene di poter dedurre che, se sappiamo che sono state fatte n prove e m di esse sono risultate favorevoli, tutti i modi possibili in cui le prove favorevoli e sfavorevoli si possono alternare fra loro ci appaiono ugualmente probabili.

Si deduce cioè la proprietà che definisce i nostri «fenomeni aleatori», e cioè una proprietà perfettamente sensata e significativa.

La prima formulazione data da de Finetti al suo teorema (1928-30)

- Inversamente, il Teor. I che enunceremo nel prossimo §, o, meglio, la [20] del § 10, mostrano che, ammesso *che* la solita concezione abbia senso, un fenomeno aleatorio è per l'appunto un fenomeno «a prove indipendenti con probabilità costante ma incognita p », ove la Φ si interpreti come «legge di probabilità (funzione di ripartizione) della probabilità incognita.
- Ciò che costituisce la giustificazione *formale*, ma *solamente formale*, della solita impostazione, che rimane sempre, concettualmente, per lo meno discutibile.

Teorema (versione del 1930)

$$[20] \quad \omega_h^{(n)} = \binom{n}{h} \int_0^1 \xi^h (1 - \xi)^{n-h} d\Phi$$

Teorema di “rappresentazione” di de Finetti (in forma più usuale)

Se $X_1, X_2, \dots, X_k, \dots$ formano una sequenza *infinita* di numeri aleatori scambiabili a valori 0-1 (“eventi”), allora esiste una distribuzione Q tale che la distribuzione congiunta P di n qualsiasi di essi si **rappresenta** come:

$$p(X_1, X_2, \dots, X_n) = \int_0^1 \left\{ \prod_{i=1}^n \theta^{X_i} (1 - \theta)^{1-X_i} \right\} dQ(\theta)$$

Hewitt e Savage hanno generalizzato il teorema nel 1955

Se $X_1, X_2, \dots, X_k, \dots$ formano una sequenza *infinita* di numeri aleatori scambiabili, allora esiste una distribuzione Q sullo spazio $\mathcal{F}$ delle distribuzioni su R^n tale che la distribuzione congiunta P di n qualsiasi di essi si **rappresenta** come:

$$p(X_1, X_2, \dots, X_n) = \int_{\mathcal{F}} \prod_{i=1}^n F(X_i) dQ(F)$$

Molti hanno dimostrato estensioni del teorema in varie direzioni

- Diaconis e Freedman e poi Accardi e Lu hanno esteso il teorema al caso delle **sequenze scambiabili finite**.
- Kerns e Szekely l'hanno esteso alle **sequenze scambiabili markoviane**.
- Aldous e Hoover hanno dato estensioni al caso di scambiabilità **parziale**, attraverso le nozioni di **scambiabilità separata e congiunta**.
-

Bibliografia parziale

- E. Hewitt, L.F. Savage, "Symmetric measures on Cartesian products" *Trans. Amer. Math. Soc.* , **80** (1955) pp. 470–501
- D.A. Freedman, "Invariance under mixing which generalize De Finetti's theorem: continuous time parameter" *Ann. Math. Stat.* , **33** (1962) pp. 916–923
- D.A. Freedman, "Invariance under mixing which generalize De Finetti's theorem: continuous time parameter" *Ann. Math. Stat.* , **34** (1963) pp. 1194–1216
- Diaconis P and Freedman D (1980a). "Finite exchangeable sequences." *Annals of Probability* 8(4), 745-764.
- Diaconis P and Freedman D (1980b). "De Finetti's theorem for Markov chains." *Annals of Probability*, 8(1), 115-130.

... segue Bibliografia parziale

- E. Hewitt, L.F. Savage, "Symmetric measures on Cartesian products" *Trans. Amer. Math. Soc.* , **80** (1955) pp. 470–501
- D.A. Freedman, "Invariance under mixing which generalize De Finetti's theorem: continuous time parameter" *Ann. Math. Stat.* , **33** (1962) pp. 916–923
- D.A. Freedman, "Invariance under mixing which generalize De Finetti's theorem: continuous time parameter" *Ann. Math. Stat.* , **34** (1963) pp. 1194–1216
- Diaconis P and Freedman D (1980a). "Finite exchangeable sequences." *Annals of Probability* 8(4), 745-764.
- Diaconis P and Freedman D (1980b). "De Finetti's theorem for Markov chains." *Annals of Probability*, 8(1), 115-130.

L'inferenza statistica come conseguenza del teorema

(da "Teoria delle probabilità", 1970)

- Dobbiamo fare inferenza sulla base di osservazioni $X_1, X_2, \dots, X_n$ scambiabili; (che, prima d'essere conosciute, sono *numeri aleatori scambiabili*);
- Abbiamo delle "ipotesi" relative al meccanismo di generazione delle osservazioni; denotiamo ciascuna di esse con il generico elemento θ ("parametro") appartenente ad uno spazio Θ ;
- Subordinatamente ad ogni ipotesi θ , le X_n hanno la stessa distribuzione $f(x|\theta)$ e sono stocasticamente indipendenti (ipotesi equivalente alla scambiabilità, per il teorema di de Finetti);
- Nello spazio Θ è definita una distribuzione di probabilità iniziale in forma di densità $p(\theta)$. (*Questa ipotesi è tipica dell'inferenza "bayesiana"; nella statistica classica se ne prescinde*).

Distribuzione “a posteriori” per le ipotesi

Nel caso di scambiabilità, e quindi di validità del teorema di de Finetti, si può ricavare (teorema di Bayes) che la distribuzione “a posteriori” per il parametro, e cioè per le ipotesi, è data dalla formula:

$$g(\theta \mid x_1, x_2, \dots, x_n) \propto \prod_{i=1}^n f(x_i \mid \theta) p(\theta)$$

Non è ancora pienamente soddisfacente, perché ipotizza una distribuzione di probabilità su di un *oggetto* (il parametro), che in generale *non è osservabile*. Questo non è accettabile in ambito soggettivo.

Distribuzione predittiva di future osservazioni

E' molto più interessante ed operativo, nella concezione di de Finetti, calcolare la *distribuzione predittiva* di una nuova osservazione Y della stessa natura delle X , che si scrive:

$$h(y | x_1, x_2, \dots, x_n) \propto \int_{\Theta} f(y | \theta) g(\theta | x_1, x_2, \dots, x_n) d\theta$$

Questo contesto **predittivo** definettiano è la base dello sviluppo del “prequential probability approach”, che si basa sull’idea che *we can judge the quality of an inference method by converting it into a forecasting system and assessing the empirical success of the sequence of one-step-ahead forecasts that it implies* (Dawid e Vovk, 1999), come si vede si parla di previsione nello stile di de Finetti.

Critica alla additività completa

(Teoria delle Probabilità, 1970)

- Di una valutazione di probabilità soggettiva... possiamo solo giudicare se è o non è *coerente*....
- Tale condizione di coerenza dovrà pertanto essere *la più debole, volendo che sia la più forte in quanto ad assoluta validità*.
- Essa infatti deve escludere soltanto le valutazioni assolutamente inammissibili, quelle cioè che non possiamo non giudicare contraddittorie....
- E tale condizione ...si riduce alla *additività semplice* (e *nonnegatività*).
- Renderla più restrittiva... è cosa inammissibile: farebbe erroneamente escludere valutazioni ammissibili.



- Ciò che le concezioni oggettivistiche ...generalmente **postulano**, è invece che valga l'additività completa ..e che il campo in cui la probabilità viene definita sia un'intera algebra di Boole.
- Dal punto di vista soggettivistico ciò è troppo e troppo poco: ci si può limitare a molto meno od anche andare oltre.
- Si potranno attribuire probabilità (semplicemente ma non completamente additive) a tutti e soli gli eventi che conviene far intervenire nella formulazione di un problema e nei ragionamenti per venirne a capo.

....Può un evento possibile avere probabilità nulla?....

- ... e, se SI, una riunione di eventi a probabilità nulla può avere probabilità positiva (in particolare essere l'evento certo) ?
- ...pensare e dire che eventi possibili di probabilità nulla possono aversi *se fanno parte di partizioni infinite (!)* ... è ***cosa mostruosa*** :
- se E ha probabilità= p (in particolare $=0$), seguirà ad averla sia considerandolo a sé stante, sia nella partizione dicotomica composta da E e $nonE$, sia in ogni altra in pochi o molti o infiniti eventi ottenibile frantumando $nonE$ in qualsiasi modo si voglia.

....Se in una partizione infinita si possa attribuire probabilità nulla a tutti gli eventi

- ... le diverse concezioni, che ammettono diversi tipi di additività per le probabilità di eventi, ...danno risposte diverse alla domanda.
- Per memoria:
- - *additività semplice*: postulata solo per somme di un numero finito di addendi (*concezione di de Finetti e pochi altri*).
- - *additività perfetta*: postulata per qualsiasi riunione di eventi (*concezione di Carnap e pochi altri*).
- - *additività completa*: postulata per somme di un numero finito o un'infinità numerabile di addendi (*concezione prevalente, codificata negli assiomi di Kolmogorov*).

Le risposte sono 3: prima e seconda

- Additività semplice: affermativa
- *La probabilità è semplicemente additiva. La riunione di un'infinità di eventi incompatibili di probabilità nulla può sempre avere probabilità positiva, ed anche essere l'evento certo.*
- Additività perfetta: negativa
- *La probabilità è perfettamente additiva. In ogni partizione c'è un numero finito o un'infinità numerabile di eventi con probabilità positive, di somma 1; gli altri hanno probabilità nulla sia individualmente che complessivamente.*

La terza risposta

- Additività numerabile: condizionata
- *Dipende. La risposta è NO se si tratta di una partizione numerabile, perché la probabilità è completamente additiva; la somma di un'infinità numerabile di zeri è zero.*
- *La risposta è SI se si tratta di un'infinità non numerabile, perché la probabilità non è perfettamente additiva: la somma di un'infinità non numerabile di zeri può essere positiva.*

Conclusione da Teoria delle Probabilità, 1970

- E' possibile, giunti alla fine, tentare di trarre qualche conclusione?
-
- Perché mai, si chiederà qualcuno, non rimanere ...nel "giusto mezzo" secondo l'attuale consuetudine, consistente nello spingersi fin dove l'additività completa fa funzionare tutto a meraviglia e nel fermarsi dove il miracolo cessa?



- Perché –rispondo- a mio avviso si tratta non di “giusto mezzo” bensì di “doppia stortura”.
- A mio avviso, tutto ciò che in quella impostazione si afferma, al di là di ciò per cui basta il livello Jordan-Peano-Riemann, è irrilevante agli effetti pratici e ingiustificabile sul piano teorico e concettuale.



- I due tipi, apparentemente contrastanti, di considerazioni sono intesi a provare, convergendo da direzioni opposte, questa stessa tesi: *di complicazioni si può fare a meno* (forse è la cosa più saggia), ma, se si vuol farne, *bisogna farle sul serio, in modo costruttivo anche se scomodo*.
- Forse ho torto. Ma le critiche non saranno state inutili se per confutarle qualcuno spiegherà e giustificherà in modo sensato e significativo quelle che, fino allora, non saranno che “Adhockeries for mathematical convenience”.