ALMA MATER STUDIORUM · UNIVERSITÀ DI BOLOGNA

FACOLTÀ DI SCIENZE MATEMATICHE, FISICHE E NATURALI
Corso di Laurea in Matematica

VALUTAZIONI COERENTI DI PROBABILITÀ

Tesi di Laurea in Probabilità e Statistica Matematica

Relatore:
Chiar.mo Prof.
MASSIMO CAMPANINO

Presentata da:
ANNA CAVALLARI

Seconda Sessione
Anno Accademico 2011-2012

La probabilità non esiste

Bruno de Finetti

Introduzione

Fin dal suo apparire, il calcolo delle probabilità ha creato discussioni sul proprio significato. I primi problemi trattati erano essenzialmente riconducibili al lancio di dadi, più in generale relativi a giochi d'azzardo in cui è presente una simmetria tra i casi possibili. In queste situazioni la valutazione delle probabilità che un evento si verifichi si può definire come il rapporto tra il numero di casi favorevoli e il numero di casi possibili. Una più matura riflessione mostra che nella migliore delle ipotesi una tale definizione è circolare, in quanto presuppone che i casi possibili siano equiprobabili. Inoltre questa definizione necessita della preesistenza di un concetto di probabilità; non è però accettabile che una definizione si basi sullo stesso concetto che dovrebbe descrivere.

Gli sviluppi del calcolo delle probabilità, e le sue applicazioni ad esempio in campo economico, dove il requisito di equiprobabilità viene meno, hanno fatto vacillare la definizione classica, senza però trovare motivate risposte per rimpiazzarla con un'altra migliore definizione.

Lo studio dei fenomeni fisici, in cui è possibile effettuare un numero qualsiasi di prove, ha portato a una definizione frequentista della probabilità. In base ad essa la probabilità di un evento è data dalla frequenza asintotica dei risultati positivi ottenibile in una successione di prove effettuate nelle stesse condizioni.

De Finetti e Savage contestano l'eccessiva ristrettezza derivante dalla ripetibilità delle prove, e propongono una definizione soggettivistica della probabilità, secondo la quale, la probabilità di un evento è il grado di fiducia (variabile da soggetto a soggetto) riposta nel verificarsi del fatto stesso.

Infine Kolmogorov ha introdotto una teoria assiomatica delle probabilità.

Proprio dall'impostazione soggettivista di Bruno de Finetti deriva l'aforisma, nonchè dedica della tesi: "*La probabilità non esiste*"; cioè la probabilità esprime il punto di vista di un osservatore e quindi non ha una propria autonoma esistenza.

Nella vita quotidiana ci troviamo spesso in situazioni d'incertezza e nelle frasi abitudinarie a volte compaiono espressioni che sottointendono all'incertezza: "forse", "credo", "è probabile", ecc., tali espressioni occupano una parte marginale, mentre il fulcro del discorso è riconducibile alla certezza. Questo appena descritto è solitamente il nostro modo di ragionare, naturalmente, ragionando si può cadere in errore, per ridurre tale rischio al ragionamento si può affiancare la logica.
Nell'oggetto di studio della presente tesi tale aiuto è chiamato logica del certo, accompagnato dalla coerenza.
Le conclusioni del nostro ragionamento logico possono essere o certe o impossibili o possibili, la terza eventualità è in una posizione intermedia e negativa: rispecchia la nostra ignoranza sul fatto, sottoposto al nostro ragionamento, nel senso che, in base alle nostre conoscenze non siamo in grado di stabilire se è vero o falso. L'ambito della possibilità è quindi l'ambito su cui si estende la nostra incertezza.
Per superare l'incertezza e il campo delle possibilità bisogna introdurre la previsione, facendo molta attenzione sul suo significato, con previsione non si intende predizione, non si vuole indovinare nulla, bensì la previsione consiste nel valutare, giudicare e stimare tutte le alternative possibili ed associare ad ognuna una probabilità.

Nella presente tesi dopo brevi cenni storici e dopo aver definito i quattro diversi concetti di probabilità qui introdotti, mi soffermerò sulla concezione soggettivista di Bruno de Finetti e grazie al principio di coerenza definirò gli assiomi e le definizioni di base della probabilità, introducendo prima le proprietà della previsione; in particolare seguendo l'impostazione del sopracitato matematico spianerò il terreno fino a giungere al *Teorema fondamentale per le probabilità*, alla sua dimostrazione e ad una semplice applicazione del teorema stesso. Concluderò con un breve accenno sulle valutazioni coerenti di probabilità e previsione nel caso di un'infinità numerabile di valori possibili.

Indice

Capitolo 1

Cenni storici

I primi studi che portarono a concetti legati alla probabilità risalgono a metà del XVI secolo in *Liber de ludo aleae* di Girolamo Cardano (scritto nel 1526, ma pubblicato un secolo e mezzo dopo, nel 1663) e in *Sulla scoperta dei dadi* di Galileo Galilei (pubblicato del 1656). In particolare, Galileo spiegò come mai, lanciando tre dadi, la probabilità di uscita delle somme 10 e 11 sia più probabile dell'uscita 9 e del 12, nonostante che entrambi i risultati si ottengano da un uguale numero di combinazioni, cioè tali somme si ottengono dallo stesso numero di modi, uscite, in cui si possono "combinare", unire, i dadi.
Il problema della ripartizione della posta in gioco nel caso che un gioco d'azzardo debba essere interrotto, venne affrontato da Luca Pacioli nella sua *Summa de arithmetica, geometria, proportioni et proportionalita* (pubblicata nel 1494) e successivamente da Tartaglia, per poi essere risolto da Pascal e Fermat.

La nascita del concetto moderno di probabilità viene attribuita a Blaise Pascal (1623-1662) e Pierre de Fermat (1601-1665). Il Cavalier de Méré (un accanito giocatore passato alla storia per questo) aveva calcolato che ottenere almeno un 6 in 4 lanci di un dado non truccato era equivalente a ottenere almeno un doppio 6 in 24 lanci, sempre di un dado non truccato. Tuttavia, giocando secondo tale convinzione, invece di vincere perdeva e scrisse a Pascal lamentando che la matematica falliva di fronte all'evidenza empirica. Da ciò scaturì una corrispondenza tra Pascal e Fermat in cui cominciò a delinearsi il concetto di probabilità nell'accezione frequentistica.

Pascal annunciò nel 1654 all'Accademia di Parigi che stava lavorando sul problema della ripartizione della messa in gioco. In una lettera del 29 luglio dello stesso anno a Fermat propose la soluzione del problema, affrontato con il metodo per ricorrenza, mentre Fermat utilizzava metodi basati sulle combinazioni.

Nel 1657 Christiaan Huygens (1629-1695) scrisse un *Libellus de ratiociniis in ludo aleae*, il primo trattato sul calcolo delle probabilità, nel quale introduceva il concetto di valore atteso.
I suoi lavori influenzarono tra l'altro Pierre de Montmort (1678-1719), che scrisse nel 1708 un *Essai d'analyse sur le jeux de hasard*, ma anche Jakob Bernoulli e Abraham de Moivre.

Nel 1713 viene pubblicato postumo *Ars conjectandi* di Jakob Bernoulli, dove veniva dimostrato il teorema che porta il suo nome, noto anche come legge dei grandi numeri. Successivamente, de Moivre pervenne ad una prima formulazione, poi generalizzata da Pierre Simon Laplace (1749-1827), del Teorema centrale del limite. La teoria delle probabilità raggiunse così basi matematicamente solide e, con esse, il rango di nuova disciplina.
In essa esercita un ruolo centrale il rapporto tra casi favorevoli e casi possibili e la probabilità è un numero intrinsecamente legato ad un evento. Negli anni centrali del XX secolo, tuttavia, prima Bruno de Finetti e poi Leonard Jimmie Savage hanno elaborato una concezione soggettivistica della probabilità, secondo cui essa è il grado di fiducia che una persona ha nel verificarsi dell'evento.

Nello stesso periodo, Andrey Nikolaevich Kolmogorov ha dato inizio alla moderna teoria assiomatica (*Grundbegriffe der Wahrscheinlichkeitsrechnung*, 1933), ispirandosi alla teoria della misura. Si è così affermata una teoria della probabilità puramente matematica.

Capitolo 2

Definizioni di probabilità

In probabilità si considera un fenomeno osservabile esclusivamente dal punto di vista della possibilità o meno del suo verificarsi, prescindendo dalla sua natura. Tra due estremi, detti *evento certo* (ottenere un numero compreso tra 1 e 6 lanciando un dado) ed *evento impossibile* (ottenere 1 come somma dal lancio di due dadi), si collocano eventi più o meno probabili.

Si possono distinguere le seguenti definizioni di probabilità:

- definizione classica,
- definizione frequentista,
- definizione soggettivistica,
- "definizione assiomatica".

Secondo la prima definizione di probabilità, per questo detta **classica**, la probabilità di un evento è *il rapporto tra il numero dei casi favorevoli all'evento e il numero dei casi possibili, purchè questi ultimi siano tutti equiprobabili.* Questa definizione è spesso attribuita a Pierre Simon Laplace e quindi anche identificata come definizione classica di Laplace.
La definizione classica consente di calcolare effettivamente la probabilità in molte situazioni. Inoltre, è una definizione operativa e fornisce quindi un metodo per il calcolo. Presenta tuttavia diversi aspetti negativi non irrilevanti:

- dal punto di vista formale, è una definizione circolare: richiede che i casi possiedano tutti la medesima probabilità, che è però ciò che si vuole definire;

- non definisce la probabilità in caso di eventi non equiprobabili;

- presuppone un numero finito di risultati possibili e di conseguenza non è utilizzabile nel continuo.

Per superare tali difficoltà, Richard von Mises (1883-1953) propose di definire la probabilità di un evento come *il limite cui tende la frequenza relativa dell'evento al crescere del numero degli esperimenti.* Questa è la definizione **frequentista** e si applica ad esperimenti casuali i cui eventi elementari non siano necessariamente ritenuti ugualmente possibili, ma assume che l'esperimento sia ripetibile più volte, idealmente infinite, sotto le stesse condizioni.

De Finetti e Savage hanno proposto una definizione di probabilità, detta **soggettiva**, applicabile ad esperimenti casuali i cui eventi elementari non siano necessariamente ritenuti ugualmente possibili e che non siano necessariamente ripetibili più volte sotto le stesse condizioni: *la probabilità di un evento è il prezzo che un individuo ritiene equo pagare per ricevere 1 se l'evento si verifica, 0 se l'evento non si verifica.*
Al fine di rendere concretamente applicabile la definizione, si aggiunge un criterio di coerenza: *le probabilità degli eventi devono essere attribuite in modo tale che non sia possibile ottenere una vincita o una perdita certa.*
La definizione soggettiva consente quindi di valutare la probabilità di eventi anche quando gli eventi elementari non sono equiprobabili e quando l'esperimento non può essere ripetuto. Rimane fondata, tuttavia, sull'opinione di singoli individui, che potrebbero presentare diverse propensioni al rischio. Basta pensare che molti sarebbero disposti a giocare 1 euro per vincerne 1.000, ma pochi giocherebbero un milione di euro per vincerne un miliardo.

L'impostazione **assiomatica** della probabilità venne proposta da Andrey Nikolaevich Kolmogorov nel 1933 in *Grundbegriffe der Wahrscheinlichkei-*

tsrechnung (*Concetti fondamentali del calcolo delle probabilità*), sviluppando la ricerca che era ormai cristallizzata sul dibattito fra quanti consideravano la probabilità come limite di frequenze relative e quanti cercavano un fondamento logico della stessa.
Va notato che la "definizione assiomatica" non è una vera e propria definizione operativa e non fornisce indicazioni su come calcolare la probabilità.
Il nome deriva dal procedimento di "assiomatizzazione" quindi nell'individuare i concetti primitivi, da questi nell'individuare i postulati da cui si passa a dimostrare i teoremi.

2.1 Concezione oggettiva - concezione soggettiva: breve confronto

Consideriamo le affermazioni, suggerite da de Finetti [1], spesso fatte da oggettivisti alle quali contrapponiamo quelle di un soggettivista per capire meglio la distinzione tra queste due impostazioni di "pensiero" probabilistico.
`Oggettivista`: due eventi dello stesso tipo in identiche condizioni sono "uguali" e hanno quindi necessariamente la stessa probabilità.
`Soggettivista`: due eventi distinti sono sempre diversi per infinite circostanze, sono ugualmente probabili per un individuo se egli li giudica tali.
`Oggettivista`: due eventi sono indipendenti se il verificarsi dell'uno non influisce sulla probabilità dell'altro.
`Soggettivista`: per definizione, due eventi per un individuo sono indipendenti se la conoscenza dell'esito di uno non gli fa modificare la valutazione di probabilità per l'altro.
`Oggettivista`: supponiamo per ipotesi che questi eventi siano ugualmente probabili e indipendenti.
`Soggettivista`: non ha senso considerare come "ipotesi" qualcosa che non è un'affermazione oggettiva; un'affermazione sulle probabilità o è la valutazione di probabilità o non è nulla.

Quando un soggettivista esprime una valutazione di probabilità è per lui chiara e rigorosa, ha delle motivazioni che possono essere condivise o meno, giudicate più o meno ragionevoli; le contestazioni altrui possono interessare ma non modificano nulla. In contrapposizione un oggettivista si accontenta di schematizzazioni semplificate di casi, trascura ogni conoscenza e opera in modo grossolano.
Il soggettivista poi si mantiene entro i confini del realismo, non si discosta dal caso di interesse o al massimo tratta casi immediatamente vicini; contrariamente l'oggettivista alla realtà sostituisce modelli schematizzati e anzichè soffermarsi sul caso specifico di interesse "fugge in avanti" e si occupa di moltissimi casi, o addirittura infiniti.

Capitolo 3

Eventi ed enti aleatori

L'oggetto fondamentale di studio del calcolo delle probabilità sono gli enti aleatori, in particolare i numeri aleatori, questi ultimi insieme agli eventi costituiscono gli strumenti che si trovano alla base dello studio dei problemi dell'incertezza. Considereremo le nozioni di:

- eventi aleatori,
- numeri aleatori,
- enti aleatori,
- funzioni aleatorie.

Con aleatorio intendiamo semplicemente qualcosa di sconosciuto, di non noto, pertanto incerto ma di per sè ben determinato.

La definizione di questi quattro concetti, qui introdotti, segue l'impostazione data da de Finetti [1].

3.1 Numeri aleatori

Si tratta di numeri ben definiti, ma non necessariamente conosciuti; ad esempio un risultato di un determinato esperimento, una quotazione azionaria in un determinato istante; sono tutti valori ben definiti e determinati, ma

possono essere sconosciuti o perchè si riferiscono al futuro o perchè non si hanno i mezzi per conoscerli.
Indicheremo i numeri aleatori con le lettere maiuscole X (o Y ecc.), e con $I(X)$ l'insieme dei valori possibili. Un numero aleatorio X si dice:

- superiormente limitato se l'insieme dei valori possibili $I(X)$ è superiormente limitato ($\sup I(X) < +\infty$);
- inferiormente limitato se l'insieme dei valori possibili $I(X)$ è inferiormente limitato ($\inf I(X) > -\infty$);
- limitato se l'insieme dei valori possibili $I(X)$ è sia superiormente che inferiormente limitato ($\sup I(X) < +\infty$, $\inf I(X) > -\infty$);
- illimitato se l'insieme dei valori possibili $I(X)$ non è nè superiormente nè inferiormente limitato ($\sup I(X) = +\infty$, $\inf I(X) = -\infty$).

Con i numeri aleatori si possono effettuare le operazioni aritmetiche usuali. Inoltre definiamo le seguenti operazioni logiche [3]:

1. $X \vee Y := \max(X, Y)$
2. $X \wedge Y := \min(X, Y)$
3. $\widetilde{X} := 1 - X$

Tali operazioni hanno le seguenti proprietà:

1. proprietà distributiva

$$X \vee (Y \wedge Z) = (X \vee Y) \wedge (X \vee Z)$$

$$X \wedge (Y \vee Z) = (X \wedge Y) \vee (X \wedge Z)$$

2. proprietà associativa

$$X \vee (Y \vee Z) = (X \vee Y) \vee Z$$

$$X \wedge (Y \wedge Z) = (X \wedge Y) \wedge Z$$

3. proprietà commutativa

$$X \vee Y = Y \vee X$$

$$X \wedge Y = Y \wedge X$$

4. proprietà connesse alla negazione $\widetilde{}$

$$\widetilde{\widetilde{X}} = X$$

$$(X \vee Y)^{\sim} = \widetilde{X} \wedge \widetilde{Y}$$

$$(X \wedge Y)^{\sim} = \widetilde{X} \vee \widetilde{Y}$$

3.2 Eventi aleatori

Un caso particolare di numero aleatorio è dato dagli eventi. Un evento ammette solo due valori: vero o falso, per convenzione li poniamo rispettivamente uguali a 1 (evento vero) oppure a 0 (evento falso).
Un evento si indica ancora con le lettere maiuscole preferibilmente E, F, ecc.; un evento E è un numero aleatorio tale che $I(E) \subset \{0, 1\}$.
Dati due eventi E e F, $E \vee F$ si dice somma logica e $E \wedge F$ prodotto logico.
Si verifica che:

1. $E \vee F = E + F - EF$

2. $E \wedge F = EF$

Dato un evento E, si definisce complementare di E (o negazione di E) l'evento $\widetilde{E} = 1 - E$.

3.3 Enti aleatori

Riferendoci alla visione già introdotta con i numeri aleatori proviamo a precisarla e ad estenderla.

Nel caso di un numero aleatorio X possiamo indicare come spazio delle alternative, $\mathcal{S}$, una retta, asse x, e su di esso l'insieme, $\mathcal{Q}$, dei soli valori (punti) possibili. Se consideriamo due numeri aleatori, X e Y, possiamo pensare il piano cartesiano con coordinate x e y come spazio $\mathcal{S}$, nel quale avremo un insieme $\mathcal{Q}$ di punti possibili; lo stesso possiamo dirlo nello spazio ordinario nel caso di tre numeri aleatori X, Y e Z e via dicendo nel caso di n numeri aleatori.

Indipendentemente dal significato geometrico possiamo immaginare per ogni ente aleatorio uno spazio astratto $\mathcal{S}$ di tutte le alternative possibili; nel caso di strutture lineari possiamo considerare vettori aleatori o matrici aleatorie, nel caso bidimensionale insiemi aleatori o addirittura curve aleatorie (per esempio la traiettoria del percorso di una mosca) ed infine insiemi aleatori superficiali (per esempio la parte della superficie terrestre in ombra in un dato istante).

3.4 Funzioni aleatorie

Una funzione aleatoria, indichiamola con $Y(t)$, supponendo per comodità che la variabile t sia il tempo, è una funzione nota e ben definita ma il suo andamento non è necessariamente conosciuto da ogni individuo. Se la funzione è nota a meno di alcuni parametri, per esempio $Y(t) = A\cos(Bt + C)$ con A, B e C aleatori la cosa è banale perchè si riconduce allo spazio dei parametri.

Generalmente parlando di funzione aleatoria ci si riferisce al caso in cui l'incertezza sussiste istante per istante: noti i valori della $Y(t)$ in un numero comunque grande ma finito di istanti $t = t_1, t_2, ..., t_n$, sarà ancora incerto il valore in un diverso istante t.

Capitolo 4

Relazioni fra eventi

Una proposizione E rimane sempre una proposizione anche se diciamo $E = 1$ o $\widetilde{E} = 0$ oppure se affermiamo "E è vera" o "non E è falsa". Se alla proposizione (o all'avento) E anteponiamo il simbolo $\vdash$ intendiamo dire che è sicuramente vera: quindi $\vdash E$ diviene l'*asserzione* che "E è certo"; naturalmente $\vdash \widetilde{E}$ è l'asserzione che "E è impossibile".

Usiamo la notazione $E \subset F$ se $\vdash E \leq F$, questa scrittura significa "E implica F". Dicendo che un evento E *implica* l'evento F, oppure che E è *contenuto* in F, affermiamo che E non può verificarsi se non si verifica anche F, ossia che è impossibile $E\widetilde{F}$.
Inoltre $\vdash E = F$ si scrive in modo equivalente come $E \equiv F$ se $E \subset F$ e $F \subset E$ cioè E è identico a F, ossia E e F sono certamente entrambi veri o entrambi falsi: *uguaglianza di E e F.*

Definizione 4.1. Si definiscono le seguenti proprietà:

- **incompatibilità**: E, F si dicono incompatibili se $\vdash EF = 0$;
- **esaustività**: $E_1, ..., E_n$ si dicono esaustivi se $\vdash E_1 + ... + E_n \geq 1$;
- **partizione**: $E_1, ..., E_n$ si dicono una partizione se $\vdash E_1 + ... + E_n = 1$ (esaustivi e incompatibili).

Dicendo che due eventi, E e F, sono *incompatibili*, affermiamo che è impossibile si verifichino entrambi, ossia che è impossibile EF. Dicendo che

n eventi $E_1, ..., E_n$ sono incompatibili affermiamo che lo sono a due a due: $\vdash E_iE_j = 0$ per $i \neq j$, ossia che non può verificarsene più di uno.

Dicendo che due eventi, E e F sono *esaustivi*, affermiamo che è impossibile non se ne verifichi nessuno, ossia che è impossibile $\widetilde{E}\widetilde{F}$. Altra forma per affermare l'esaustività è: $\vdash E + F \geq 1$. Possiamo estendere la definizione di esaustività al caso di n eventi: dicendo che essi sono esaustivi affermiamo che almeno uno tra essi deve verificarsi.

Diremo *partizione* una famiglia di eventi incompatibili ed esaustivi, dei quali cioè è certo che deve verificarsi uno e uno solo: la coesistenza dell'incompatibilità e dell'esaustività significa infatti $\vdash E_1 + ... + E_n = 1$. Una partizione può essere finita o infinita.
Dati $E_1, ..., E_n$ eventi, per costruire una partizione a partire dagli n eventi si utilizza il metodo dei costituenti.
Consideriamo i 2^n prodotti $E'_1E'_2...E'_n$ dove ciascun E'_i può essere uguale a E_i o al suo complementare $\widetilde{E}_i$. Alcuni di questi 2^n prodotti possono risultare impossibili e non vanno considerati; quelli che, essendo possibili, rimangono, in numero $s \leq 2^n$, si dicono i *costituenti* $C_1, ..., C_s$ della partizione determinata da $E_1, ..., E_n$. Sono possibili tutti i costituenti solo quando gli E_i sono logicamente indipendenti. I costituenti possibili costituiscono una partizione. Infatti

$$1 = (E_1 + \widetilde{E}_1)...(E_n + \widetilde{E}_n) = \sum_{C\ costituente} C$$

dalla somma possono essere esclusi tutti i costituenti impossibili.
Se $E_1, ..., E_n$ sono una partizione, allora i costituenti possibili sono:

$$E_1\widetilde{E}_2...\widetilde{E}_n \qquad \widetilde{E}_1E_2\widetilde{E}_3...\widetilde{E}_n \qquad ... \qquad \widetilde{E}_1...\widetilde{E}_{n-1}E_n$$

che si possono identificare con gli eventi stessi.

4.1 Dipendenza e indipendenza logica di eventi

Diciamo che n eventi tutti necessariamente possibili sono *logicamente indipendenti* quando danno luogo a 2^n costituenti possibili, questo significa che ciascuno di tali eventi rimane incerto anche dopo la conoscenza dell'esito degli altri.
Supponiamo che uno dei prodotti sia impossibile, consideriamo ad esempio il prodotto $E_1 E_2 ... E_n$; E_1 è possibile, $E_1 E_2$ può esserlo o no, $E_1 E_2 E_3$ può esserlo o no, e così via, se uno di questi prodotti è impossibile lo sono ovviamente tutti i successivi. Sia impossibile il prodotto $E_1 E_2 E_3 E_4$ questo significa che è possibile si verifichino gli eventi E_1, E_2 e E_3 e che E_4 non può essere vero. Gli eventi in tal caso, ossia quando il numero dei costituenti possibili $s < 2^n$, si dicono *logicamente dipendenti*.

Consideriamo un evento E e la sua dipendenza da $E_1, ..., E_n$, si distinguono più casi: se E rimane incerto dopo la conoscenza dell'esito degli E_i significa che è logicamente indipendente, se E risulta determinato dopo la conoscenza dell'esito degli E_i significa che è logicamente dipendente, infine vi è il caso intermedio e cioè che E sia incerto o no a seconda dell'esito degli E_i in questo caso è detto logicamente semidipendente.

Siano $C_1, ..., C_s$ i costituenti determinati dagli E_i. Ogni evento E si può scrivere nel seguente modo: $E = EC_1 + EC_2 + ... + EC_s$. Consideriamo uno qualunque degli addendi, sia EC_h, si avranno tre casi distinti: $EC_h = C_h$, oppure $EC_h = 0$, oppure $0 \subset EC_h \subset C_h$.
I risultati possibili degli E_i corrispondono al realizzarsi di uno qualunque dei costituenti C_h, a seconda che C_h appartenga al I, al II o al III tipo, E risulta rispettivamente *certo*, *impossibile* o *incerto*. Quindi possiamo classificare i costituenti nel seguente modo:

- **I tipo** $C_h \subset E$,
- **II tipo** $C_h \subset \widetilde{E}$,

- **III tipo** altrimenti.

Quindi diciamo che l'evento E è:

- *logicamente dipendente* se non esistono costituenti del III tipo, ossia se tutti i costituenti di $E_1, ..., E_n$ sono del I o del II tipo,
- *logicamente indipendente* se tutti i costituenti sono del III tipo,
- *logicamente semidipendente* se esistono costituenti del III tipo e del I tipo o/e del II tipo.

Capitolo 5

Rappresentazioni in forma lineare

L'insieme $\mathcal{Q}$ dei punti possibili di due numeri aleatori X e Y è costituito dalle coppie dei punti cartesiani (x, y) del piano che la coppia (X, Y) può assumere come valore, dove il piano rappresenta lo spazio $\mathcal{S}$ delle alternative; tutto ciò si può generalizzare a tre punti o più, passando quindi ad uno spazio di dimensione qualunque.

Seguendo sempre l'impostazione di de Finetti [1], consideriamo $\mathcal{S}$ come un sottoinsieme di uno *spazio lineare affine* che chiamiamo *ambiente lineare* e lo indichiamo con $\mathcal{A}$; $\mathcal{S}$ è una varietà meno estesa rispetto ad $\mathcal{A}$ che contiene $\mathcal{Q}$; ad esempio se $\mathcal{A}$ è lo spazio ordinario ed X, Y e Z sono legati dall'equazione $X^2 + Y^2 + Z^2 = R^2$ $\mathcal{S}$ è la superficie sferica su cui si trovano i punti possibili $\mathcal{Q}$.
I numeri aleatori linearmente rappresentati in un ambiente $\mathcal{A}$ costituiscono un sistema lineare, che indichiamo con $\mathcal{L}$, e che è duale di $\mathcal{A}$.

Consideriamo gli eventi $E_1, E_2, ..., E_n$, l'ambiente lineare $\mathcal{A}$ è lo spazio vettoriale affine a n dimensioni, con sistema di coordinate $x_1, x_2, ..., x_n$ sulle quali saranno considerati i valori dei numeri aleatori $X_1, X_2, ..., X_n$; in questo caso i numeri aleatori sono gli eventi $E_1, E_2, ..., E_n$ cioè ai quali associamo i soli valori 0 ed 1, e quindi l'insieme $\mathcal{Q}$ dei punti possibili è formato al

massimo dai 2^n punti di coordinate 0 o 1 e altrimenti da una parte di essi; questi corrispondono agli s ($s \leq 2^n$) costituenti.
È comodo pensare a tali punti come i vertici di un cubo (o ipercubo).

Il sistema lineare $\mathcal{L}$, delle combinazioni lineari di $E_1, E_2, ..., E_n$, è formato dai numeri aleatori $X = u_1E_1 + u_2E_2 + ... + u_nE_n$, interpretabili come il guadagno (positivo o negativo) dell'importo u_1 se si verifica E_1, più u_2 se si verifica E_2, e così via; le X possono avere al più tanti valori possibili distinti quanti sono i costituenti, e questo caso si ha quando i corrispondenti punti possibili si trovano su iperpiani distinti $\sum_i u_i x_i = cost.$

Per $n = 3$ indichiamo con x, y, z le coordinate cartesiane dell'ambiente lineare $\mathcal{A}$ e con u, v, w quelle del sistema duale $\mathcal{L}$, cioè se $X = uE_1 + vE_2 + wE_3$ si ottiene che il valore di X è $ux + vy + wz$ quando E_1 assume il valore x, E_2 il valore y ed E_3 il valore z. Essendo E_1, E_2, E_3 eventi, i valori di x, y, z possono solo essere 0 o 1, e considerati i vertici del cubo Q i corrispondenti valori che assume il numero aleatorio X sono raccolti nella seguente tabella:

Vertici del cubo	Valori di X
$(0, 0, 0)$	0
$(0, 0, 1)$	w
$(0, 1, 0)$	v
$(0, 1, 1)$	$v + w$
$(1, 0, 0)$	u
$(1, 0, 1)$	$u + w$
$(1, 1, 0)$	$u + v$
$(1, 1, 1)$	$u + v + w$

In particolare se $u = v = w = 1$ il numero di successi è 0 in un caso, 1 in tre casi, 2 in tre casi, 3 in un caso $(1 + 3 + 3 + 1 = 8)$.

Finora abbiamo considerato solo combinazioni lineari omogenee

$$X = u_1E_1 + u_2E_2 + ... + u_nE_n$$

ora partiamo dall'ipotesi che $X_i = E_i$ e introduciamo anche quelle complete consideriamo quindi la costante u_0 e X_0 che rappresenta il numero aleatorio che vale certamente uno ($X_0 \equiv 1$); quindi l'addendo u_0X_0 vale u_0 senza alterare la formula $X = \sum_i u_iX_i$, va solo tenuto presente che c'è una variabile fittizia, x_0, in più.

Abbiamo quindi le combinazioni lineari di n numeri aleatori X_i,

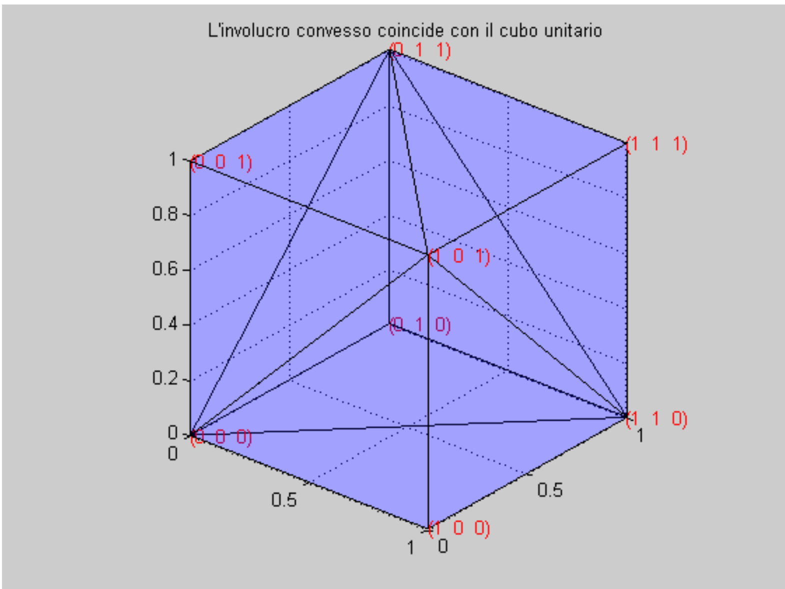
$$X = \sum_i u_iX_i$$

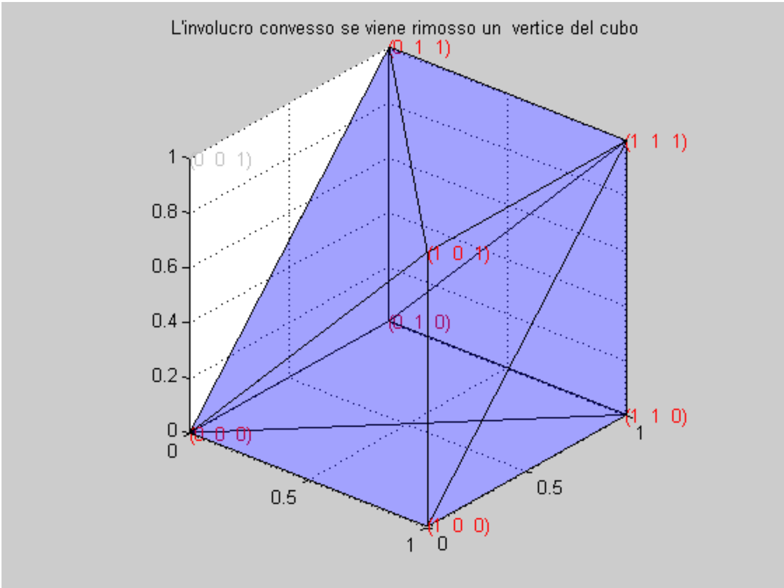
X si dice linearmente dipendente dalle X_i, può darsi però che già le X_i date siano linearmente dipendenti, cioè che tra le combinazioni lineari si trovi qualcuna identicamente nulla e quindi almeno una delle X_i è combinazione lineare delle altre e non deve essere considerata. Geometricamente significa che l'insieme dei punti possibili $\mathcal{Q}$ appartiene ad un sottospazio lineare $\mathcal{A}'$ di $\mathcal{A}$ ed allora è sufficiente considerare $\mathcal{A}'$.

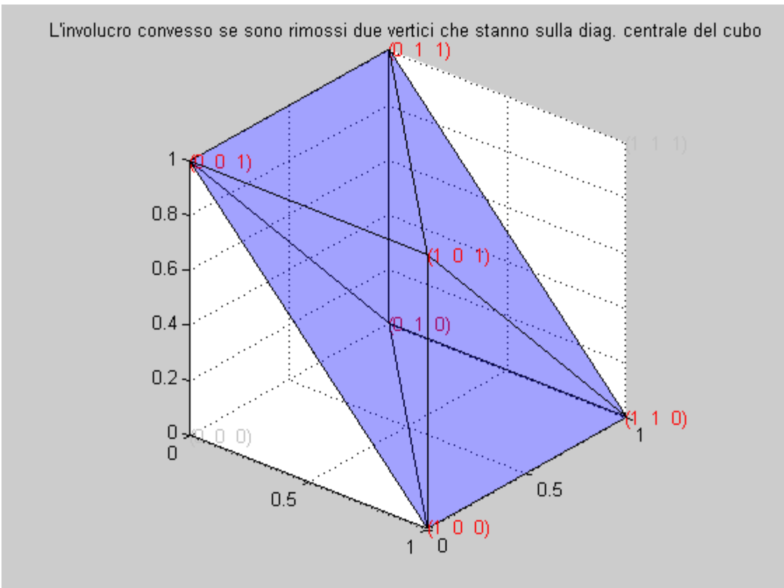
Consideriamo la somma $\sum_i u_ix_i$ che rappresenta una funzione lineare sia di X (ossia delle sue componenti u_i) e sia di Q (ossia delle sue coordinate x_i) quindi possiamo indicarla sia con $X(Q)$ pensandola come il valore di un dato X al variare di Q, sia con $Q(X)$ pensandola come il valore assegnato dal risultato Q ai diversi X.
Pensiamo al baricentro P di due punti Q_1 e Q_2 con "masse" q_1 e q_2 ($q_1+q_2 = 1$) per le proprietà della meccanica abbiamo che ogni funzione lineare X assume in P il valore $X(P) = q_1X(Q_1) + q_2X(Q_2)$ questa uguaglianza può essere estesa a qualsiasi numero di punti.
Il baricentro può essere allora un qualunque punto appartenente all'**involucro convesso** dei punti Q_h considerati. In matematica si definisce involucro convesso di un qualsiasi insieme I l'intersezione di tutti gli insiemi convessi che contengono I. Poichè l'intersezione di insiemi convessi è a sua volta convessa, una definizione alternativa di involucro convesso è "il più piccolo insieme convesso contenente I". Intuitivamente, l'involucro convesso di un insieme di punti del piano è la forma che assumerebbe un elastico allargato in modo da contenere tutti i punti e poi lasciato libero di restringersi: un poligono che ha alcuni di quei punti come vertici e li contiene tutti. Nel calcolo delle probabilità è fondamentale considerare l'involucro convesso determinato dai

punti possibili, cioè dai $Q \in \mathcal{Q}$, ossia l'involucro convesso di $\mathcal{Q}$.
Dualmente l'involucro convesso di $\mathcal{Q}$ è anche l'intersezione di tutti i semispazi contenenti $\mathcal{Q}$; in altre parole, se un punto P appartiene all'involucro convesso $K(I)$ di un insieme I, esso si trova dalla stessa parte di I rispetto ad ogni iperpiano che non taglia l'insieme I, mentre se non vi appartiene esiste un iperpiano che lo separa da I. Quindi ogni funzione lineare nonnegativa su I lo è anche su tutto $K(I)$ e tale proprietà non vale per nessuno punto non appartenente a $K(I)$.

Il caso del cubo sopracitato è un esempio semplice ma significativo di un involucro convesso, considerando tutti gli otto vertici o una parte di essi.
L'involucro convesso generato dagli otto vertici, da sette vertici e da sei vertici sono mostrati rispettivamente nelle figure seguenti [6]:




L'involucro convesso se viene rimosso un vertice del cubo
(0 1 1)
(1 1 1)
(0 0 1)
(1 0 1)
(0 1 0)
(1 1 0)
(0 0 0)
(1 0 0)
1
0.8
0.6
0.4
0.2
0
0
0.5
1
0
0.5
1


L'involucro convesso se sono rimossi due vertici che stanno sulla diag. centrale del cubo
(0 1 1)
(1 1 1)
(0 0 1)
(1 0 1)
(0 1 0)
(1 1 0)
(0 0 0)
(1 0 0)
1
0.8
0.6
0.4
0.2
0
0
0.5
1
0
0.5
1

Capitolo 6

Coerenza e probabilità

Relativamente alla nozione di *probabilità di un evento*, diversamente da altri autori che richiedono agli eventi di soddisfare proprietà particolari, come la simmetria o la ripetibilità, proprietà che limitano eccessivamente la possibilità di eseguire valutazioni probabilistiche nella maggior parte dei casi di interesse, Bruno de Finetti definisce $\mathbf{P}(E)$ come espressione numerica del grado di fiducia che un individuo ripone nel verificarsi dell'evento E in base alle sue informazioni ed opinioni. Si tratta evidentemente di una nozione *soggettiva* perchè, con riferimento ad un dato evento E, la probabilità $\mathbf{P}(E)$ dipende dall'individuo che la valuta, ma, per uno stesso individuo, dipende dal suo stato di conoscenza su E, che generalmente varia con il tempo.

Basandosi su un'interpretazione della nozione di probabilità in termini del grado di fiducia che un dato individuo ha, in base alle sue convizioni e conoscenze, sull'avverarsi di un dato evento, de Finetti pone a fondamento del calcolo delle probabilità il *principio di coerenza.* In questo modo la teoria quantitativa della probabilità soggettiva viene identificata con la teoria matematica delle valutazioni coerenti di probabilità che soddisfano il principio di coerenza.

Questo principio viene presentato da de Finetti tramite due formulazioni alternative, tra loro equivalenti, che si basano su schemi concettuali diversi: il *metodo della scommessa* e il *metodo della penalità.*

Questi due criteri consistono:

- in uno schema di decisione che viene sottoposto ad un individuo in modo tale che egli attribuisca un valore ad un evento, tale valore per definizione si chiama previsione,

- in una condizione di coerenza che permette di distinguere se le previsioni date dall'individuo costituiscono un insieme coerente ed accettabile oppure contraddittorio.

Il **metodo della scommessa** consiste nel considerare un numero aleatorio X, l'individuo sceglie un valore $\overline{x}$ e dopo tale scelta si impegna ad accettare qualsiasi scommessa con guadagno

$$c(X - \overline{x})$$

dove $c \in \mathbb{R}$ è un coefficiente di proporzionalità scelto da chi ci propone la scommessa.
Quindi per definizione abbiamo $\mathbf{P}(X) = \overline{x}$, previsione di X secondo l'opinione dell'individuo. Inoltre si suppone che l'individuo non accetti una scommessa che dia certamente una perdita, pertanto un insieme di previsioni si dice coerente se tra le combinazioni di scommesse accettate nessuna dà guadagni sicuramente negativi. Quindi per il criterio di coerenza non si può scegliere $\overline{x}$ in modo che ci sia una perdita certa.

Il **metodo della penalità** suppone che l'individuo subirà una penalizzazione L proporzionale al quadrato della differenza tra X ed un valore $\overline{x}$ che egli presceglie a tal fine come meglio crede:

$$L = \left(\frac{X - \overline{x}}{k}\right)^2$$

con k arbitrariamente fissato.
Quindi si ha $\mathbf{P}(X)$, previsione di X, uguale al valore $\overline{x}$ prescelto. Inoltre si suppone che l'individuo non preferisca un valore se può sceglierne un altro che dà una penalizzazione certamente minore, pertanto l'insieme di previsioni

si dice coerente se non ne potevano essere scelte altre in modo da rendere certamente minore la penalizzazione. Quindi per il criterio di coerenza non deve esistere un valore $\overline{x'}$ tale che la corrispondente penalità sia sicuramente minore.

Si verifica l'equivalenza dei due criteri. Sia $\overline{x}$ la previsione di X definita in base al metodo della scommessa e sia $\overline{x'}$ quella definita in base al metodo della penalità, ciò significa rispettivamente che:

- per quanto riguarda il primo criterio, il guadagno aleatorio X è equivalente al guadagno certo $\overline{x}$ quindi è preferibile ad ogni $x < \overline{x}$ e non ad un $x > \overline{x}$;
- nel secondo criterio il guadagno, nonchè la penalizzazione $-(X - \overline{x'})^2$ è preferibile ad ogni altro $-(X - x)^2$ con $x \neq \overline{x'}$, ossia il guadagno $G = (X - x)^2 - (X - \overline{x'})^2$ è preferito a 0.

Consideriamo le penalizzazioni corrispondenti a due valori qualunque di x, siano $x = a$ e $x = b$, e indichiamo con $c = \frac{a+b}{2}$ il valore centrale dell'intervallo (a, b). La scelta di a è preferibile a quella di b se è preferibile a 0 il guadagno $G = (X - b)^2 - (X - a)^2$, ossia, sviluppando i calcoli, se è preferibile a 0

$$G = (X^2-2bX+b^2)-(X^2-2aX+a^2) = 2(a-b)X-(a^2-b^2) = 2(a-b)(X-c)$$

La preferibilità di G a 0 significa $\mathbf{P}(G) > 0$, in base al primo metodo risulta $\mathbf{P}(G) = 2(a - b)(\overline{x} - c)$, espressione che è positiva se $a > b$ ed $\overline{x} > c$, o viceversa, se $a < b$ ed $\overline{x} < c$, ossia se $\overline{x}$ cade nell'intervallo tra c (punto medio dell'intervallo (a, b)) ed a, ossia se a è più vicino ad $\overline{x}$ che non a b.

Consideriamo ora l'insieme $\mathcal{P}$ delle previsioni coerenti e diamo ai nostri due criteri un'interpretazione geometrica come nell'impostazione di de Finetti [1]. Dati n numeri aleatori $X_1, X_2, ..., X_n$ nell'ambiente lineare $\mathcal{A}$ fissiamo in tale ambiente, cioè nello spazio a n dimensioni delle coordinate $x_1, x_2, ..., x_n$, gli n valori $\overline{x_1}, \overline{x_2}, ..., \overline{x_n}$ che corrispondono alle rispettive previsioni $\mathbf{P}(X_1), \mathbf{P}(X_2), ..., \mathbf{P}(X_n)$. Le condizioni di coerenza dicono che l'insieme $\mathcal{P}$ delle previsioni $\mathbf{P}$ coerenti è l'involucro convesso chiuso dell'insieme $\mathcal{Q}$ dei

punti Q possibili.
Nel metodo della scommessa la condizione necessaria e sufficiente per la coerenza si può esprimere dicendo che ogni equazione lineare tra numeri aleatori X_i

$$c_1X_1 + c_2X_2 + ... + c_nX_n = c$$

deve essere rispettata per le rispettive previsioni $\mathbf{P}(X_i)$:

$$c_1\mathbf{P}(X_1) + c_2\mathbf{P}(X_2) + ... + c_n\mathbf{P}(X_n) = c$$

Geometricamente un punto rappresenta una previsione coerente se e solo se nessun iperpiano lo separa dall'insieme $\mathcal{Q}$ dei punti possibili, e ciò caratterizza i punti dell'involucro convesso.
Per quanto riguarda il metodo della penalizzazione introduciamo nell'ambiente lineare $\mathcal{A}$ una metrica $\rho^2 = \sum_i (x_i/k_i)^2$, ponendo:

penalizzazione $= L = (P - Q)^2 =$ quadrato della distanza tra il punto-previsione P e il punto-risultato Q, secondo detta metrica

La condizione necessaria e sufficiente per la coerenza richiede, geometricamente, che P non può essere spostato in modo da ridurne la distanza da tutti i punti Q, questo caratterizza i punti dell'involucro convesso.

Dalla condizione di coerenza e quindi dall'ambito della teoria delle valutazioni coerenti di probabilità si ottengono le proprietà della previsione e, come teoremi, gli *assiomi* delle probabilità.

6.1 Proprietà della previsione

Dal principio di coerenza si ottiene che la previsione $\mathbf{P}$ gode:

1. dell'additività $\mathbf{P}(X + Y) = \mathbf{P}(X) + \mathbf{P}(Y)$;

2. della convessità cioè $\mathbf{P}(X)$ non deve essere inferiore all'estremo inferiore dei valori possibili per X nè superiore all'estremo superiore dei valori possibili per X;

3. della linearità per un numero finito qualunque di addendi, questa proprietà discende dall'additività e dalla convessità.

Otteniamo quindi la seguente proposizione [3]:

Proposizione 6.1.1. *La previsione ha le seguenti proprietà:*

1. ***monotonia:*** $\inf I(X) \leq \mathbf{P}(X) \leq \sup I(X)$

2. ***linearità:*** *se* $X = a_1X_1 + a_2X_2 + ... + a_nX_n$ *allora*

$$\mathbf{P}(X) = a_1\mathbf{P}(X_1) + a_2\mathbf{P}(X_2) + ... + a_n\mathbf{P}(X_n)$$

dove ricordiamo che $I(X)$ è l'insieme dei valori possibili del numero aleatorio X.

Dimostrazione. Per dimostrare la monotonia utilizziamo il metodo della scommessa e il guadagno $c(X - \overline{x})$. La previsione $\overline{x}$ deve essere tale che non si possa scegliere c in modo tale che si abbia un guadagno certo o una perdita certa. Se fosse $\overline{x} < \inf I(X)$, allora per $c < 0$:

$$\vdash c(X - \overline{x}) < 0$$

Se invece fosse $\overline{x} > \sup I(X)$, allora per $c > 0$:

$$\vdash c(X - \overline{x}) < 0$$

Ne segue che

$$\inf I(X) \leq \mathbf{P}(X) \leq \sup I(X)$$

dove naturalmente $\mathbf{P}(X) = \overline{x}$.

Dimostriamo ora la linearità, consideriamo il numero aleatorio $Z = X+Y$. Posto $\overline{x} = \mathbf{P}(X)$, $\overline{y} = \mathbf{P}(Y)$ e $\overline{z} = \mathbf{P}(Z)$, utilizzando sempre il metodo della scommessa sia G il guadagno pari a:

$$G = c_1(X-\overline{x})+c_2(Y-\overline{y})+c_3(Z-\overline{z}) = (c_1+c_3)X+(c_2+c_3)Y-c_1\overline{x}-c_2\overline{y}-c_3\overline{z}$$

Scegliendo c_1, c_2 e c_3 in modo tale da annullare la parte aleatoria, ovvero ponendo $c_1 = c_2 = -c_3$, si ottiene il guadagno complessivo:

$$G = c_3(\overline{x} + \overline{y} - \overline{z})$$

Per evitare che si possa scegliere c_3 in modo che $\vdash G < 0$, dovrà essere $\overline{x} + \overline{y} - \overline{z} = 0$, ovvero $\overline{z} = \overline{x} + \overline{y}$.
Per ogni punto P, la proiezione ortogonale P' di P sul piano $z = x + y$ ha distanza minore di P da ogni punto (X, Y, Z) possibile. In base al principio di coerenza dovrà essere $P = P'$, ovvero P deve appartenere al piano $z = x+y$. Ne segue che $\overline{z} = \overline{x} + \overline{y}$.
Analogamente, per $Z = aX$, $a \in \mathbb{R}$, si ottiene $\overline{z} = a\overline{x}$.
In generale, se $X = a_1X_1 + a_2X_2 + ... + a_nX_n$, allora

$$\mathbf{P}(X) = a_1\mathbf{P}(X_1) + a_2\mathbf{P}(X_2) + ... + a_n\mathbf{P}(X_n)$$

□

6.2 Probabilità di eventi

Quando si parla di un evento E, la previsione $\mathbf{P}(E)$ si chiama anche *probabilità* di E. Le proprietà delle probabilità di eventi non sono che particolarizzazioni delle proprietà delle previsioni di numeri aleatori.

Dalla proprietà di monotonia, segue che:

1. $0 \leq \mathbf{P}(E) \leq 1$,
2. $E \equiv 0 \Rightarrow \mathbf{P}(E) = 0$,
3. $E \equiv 1 \Rightarrow \mathbf{P}(E) = 1$.

Se $E \equiv 1$, E si dice *evento certo*, se $E \equiv 0$ si dice *evento impossibile.*

Consideriamo inoltre:

somma logica: $\mathbf{P}(E_1 \vee E_2) = \mathbf{P}(E_1 + E_2 - E_1E_2) \leq \mathbf{P}(E_1 + E_2)$, *somma aritmetica*: $\mathbf{P}(E_1 + E_2) = \mathbf{P}(E_1) + \mathbf{P}(E_2)$.

Dati i due eventi E_1 ed E_2 somma logica e somma aritmetica coincidono se i due eventi sono incompatibili, lo stesso vale per un numero finito qualunque di eventi incompatibili: $E = E_1 \vee E_2 \vee ... \vee E_n = E_1 + E_2 + ... + E_n$ e quindi

$$\mathbf{P}(E) = \mathbf{P}(E_1) + \mathbf{P}(E_2) + ... + \mathbf{P}(E_n)$$

Nel caso di una partizione finita si ha

$$\vdash E_1 + E_2 + ... + E_n = 1 \Rightarrow \sum_i \mathbf{P}(E_i) = 1$$

quindi in una partizione le probabilità devono dare come somma 1.
Ad esempio se consideriamo due eventi complementari E ed $\widetilde{E}$, cioè nel caso di una partizione con due eventi, risulta $\mathbf{P}(E) + \mathbf{P}(\widetilde{E}) = 1$, ossia $\mathbf{P}(\widetilde{E}) = 1 - \mathbf{P}(E)$. Da questo esempio discende il fatto che le probabilità di eventi complementari devono essere complementari.

Per eventi qualunque, prescindendo cioè dall'ipotesi di incompatibilità, abbiamo

$$E = E_1 \vee E_2 \vee ... \vee E_n = 1 \wedge (E_1 + E_2 + ... + E_n) \leq E_1 + E_2 + ... + E_n$$

e quindi

$$\mathbf{P}(E) \leq \mathbf{P}(E_1) + \mathbf{P}(E_2) + ... + \mathbf{P}(E_n)$$

cioè la probabilità dell'evento somma deve essere minore o uguale della somma delle probabilità. Nel caso di eventi compatibili, oltre la precedente disuguaglianza, nulla si può dire di più per $\mathbf{P}(E)$. Introducendo altri elementi abbiamo qualche informazione in più.
Mediante i costituenti, basta considerare il solo costituente

$$C = \widetilde{E}_1 \widetilde{E}_2 ... \widetilde{E}_n$$

perchè $E = \widetilde{C}$ quindi otteniamo

$$\mathbf{P}(E) = 1 - \mathbf{P}(C)$$

Se consideriamo il prodotto degli E_i (due a due, tre a tre, ...), grazie allo sviluppo,

$$E = \sum_i E_i - \sum_{i,j} E_i E_j + \sum_{i,j,k} E_i E_j E_k - ... \pm E_1 E_2 ... E_n$$

otteniamo che la probabilità dell'evento somma deve essere sempre

$$\mathbf{P}(E) = \sum_i \mathbf{P}(E_i) - \sum_{i,j} \mathbf{P}(E_i E_j) + \sum_{i,j,k} \mathbf{P}(E_i E_j E_k) - ... \pm \mathbf{P}(E_1 E_2 ... E_n)$$

Quest'ultima espressione è lineare nelle probabilità dei prodotti. Riportiamo i casi particolari di due e di tre eventi:

$$\mathbf{P}(E_1 \vee E_2) = \mathbf{P}(E_1) + \mathbf{P}(E_2) - \mathbf{P}(E_1 E_2)$$

$$\mathbf{P}(E_1 \vee E_2 \vee E_3) =$$
$$\mathbf{P}(E_1) + \mathbf{P}(E_2) + \mathbf{P}(E_3) - \mathbf{P}(E_1 E_2) - \mathbf{P}(E_1 E_3) - \mathbf{P}(E_2 E_3) + \mathbf{P}(E_1 E_2 E_3)$$

Una formula analoga serve per esprimere la probabilità che dati n eventi se ne verifichino esattamente h, non importa quali. Il verificarsi di $E_1 E_2 ... E_h$ è:

$$E_1 E_2 ... E_h (1 - E_{h+1})(1 - E_{h+2}) ... (1 - E_n) =$$
$$E_1 E_2 ... E_h - \textstyle\sum_i E_1 E_2 ... E_h E_{h+i} + \sum_{i,j} E_1 E_2 ... E_h E_{h+i} E_{h+j} - ... \pm E_1 E_2 ... E_n$$

Capitolo 7

Dipendenza lineare

Seguendo l'impostazione di de Finetti [1] consideriamo tre eventi logicamente indipendenti E_1, E_2, E_3 ed il sistema cartesiano di riferimento (x, y, z) cui sovrapponiamo l'ambiente lineare $\mathcal{A}$ ed il sistema lineare $\mathcal{L}$. Gli otto vertici del cubo unitario

$$(0,0,0) \quad (1,0,0) \quad (0,1,0) \quad (0,0,1) \quad (0,1,1) \quad (1,0,1) \quad (1,1,0) \quad (1,1,1)$$

pensati come punti di $\mathcal{A}$ rappresentano i costituenti Q_i, punti possibili di $\mathcal{Q}$:

$$Q_0 = \widetilde{E}_1\widetilde{E}_2\widetilde{E}_3 \quad Q_1 = E_1\widetilde{E}_2\widetilde{E}_3 \quad Q_2 = \widetilde{E}_1E_2\widetilde{E}_3 \quad Q_3 = \widetilde{E}_1\widetilde{E}_2E_3$$

$$Q_0' = E_1E_2E_3 \quad Q_1' = \widetilde{E}_1E_2E_3 \quad Q_2' = E_1\widetilde{E}_2E_3 \quad Q_3' = E_1E_2\widetilde{E}_3$$

(dove agli zeri corrispondono negazioni, agli uno affermazioni); pensati come punti di $\mathcal{L}$ rappresentano i numeri aleatori

$$0 \qquad E_1 \qquad E_2 \qquad E_3 \qquad E_1+E_2+E_3 \qquad E_2+E_3 \qquad E_1+E_3 \qquad E_1+E_2$$

(dove compaiono gli addendi corrispondenti agli uno).

Consideriamo il punto generico di $\mathcal{A}$, (x, y, z), questo significa che E_1 assume il valore x, E_2 il valore y ed E_3 il valore z, quindi ciò può valere come previsione cioè $\mathbf{P}(E_1) = x$, $\mathbf{P}(E_2) = y$ e $\mathbf{P}(E_3) = z$, ossia (x, y, z) rappresenta la previsione $\mathbf{P}$ che attribuisce ad E_1, E_2, E_3 le probabilità $(p_1, p_2, p_3) = (x, y, z)$ e che è anche esprimibile come baricentro dei punti Q_i con opportuni

pesi (masse) q_i.
Pensato come punto di $\mathcal{L}$, (x, y, z) rappresenta il numero aleatorio $X = uE_1 + vE_2 + wE_3$ di coefficienti $(u, v, w) = (x, y, z)$. Poichè $\mathbf{P}(X) = up_1 + vp_2 + wp_3 = ux + vy + wz$, $\mathbf{P}(X)$ può essere interpretato come prodotto interno dei vettori duali $\mathbf{P}$ di $\mathcal{A}$ ed X di $\mathcal{L}$.

Ora vediamo i vari casi che si ottengono precisando le ipotesi sugli E_i, ossia stabilendo quali tra gli otto prodotti sono effettivamente costituenti possibili. Se esistono tutti gli otto costituenti, $\mathcal{P}$ è tutto il cubo.

Se gli E_i costituiscono una partizione, i costituenti sono tre:

$$Q_1 = (1, 0, 0),\ Q_2 = (0, 1, 0) \text{ e } Q_3 = (0, 0, 1)$$

sappiamo che le probabilità p_i possono essere tre numeri qualunque non negativi di somma pari a 1, ossia le $P = (x, y, z)$ ammissibili appartengono al piano $x + y + z = 1$, e più precisamente al triangolo avente per vertici i tre punti possibili Q_1, Q_2, Q_3 e sono esprimibili univocamente come baricentro

$$P = q_1Q_1 + q_2Q_2 + q_3Q_3$$

di tali punti con pesi $q_1 = x$, $q_2 = y$, $q_3 = z$. Tale triangolo costituisce lo spazio $\mathcal{P}$ delle previsioni ammissibili, ed è l'involucro convesso dell'insieme $\mathcal{Q}$ delle eventualità possibili che si riduce ai tre vertici detti.
Pensandolo in $\mathcal{L}$, possiamo dire che il punto $(1, 1, 1)$ rappresenta il numero aleatorio certamente pari a 1, dato che $E_1 + E_2 + E_3 = 1$.

Se gli E_i sono incompatibili (ma non esaustivi) i costituenti sono quattro:

$$Q_0 = (0, 0, 0),\ Q_1 = (1, 0, 0),\ Q_2 = (0, 1, 0) \text{ e } Q_3 = (0, 0, 1)$$

in questo caso vale la relazione $x + y + z \leq 1$; si ha ancora P esprimibile come baricentro

$$P = q_0Q_0 + q_1Q_1 + q_2Q_2 + q_3Q_3$$

dei Q con pesi $q_0 = 1 - x - y - z$, $q_1 = x$, $q_2 = y$, $q_3 = z$; lo spazio $\mathcal{P}$ è il tetraedro di vertici Q_0, Q_1, Q_2 e Q_3.

Siano E_1 ed E_2 logicamente indipendenti ed E_3 ne sia il prodotto:

$$E_3 = E_1E_2$$

i costituenti sono i seguenti:

$$Q_0 = (0,0,0),\ Q_1 = (1,0,0),\ Q_2 = (0,1,0) \text{ e } Q'_0 = (1,1,1)$$

i primi tre sono sul piano $z = 0$, gli ultimi tre su $z = x + y - 1$, gli altri due gruppi di tre su $z = y$ e su $z = x$; lo spazio $\mathcal{P}$ è quindi il tetraedro $z \leq 0$, $z \leq x + y - 1$, $z \leq x$, $z \leq y$, ossia:

$$[\max(0, x + y - 1) =] \quad 0 \vee (x + y - 1) \leq z \leq x \wedge y \quad [= \min(x, y)]$$

Anche qui P è esprimibile come baricentro

$$P = q_0 Q_0 + q_1 Q_1 + q_2 Q_2 + q'_0 Q'_0$$

dei Q con pesi $q_0 = 1 - x - y + z$, $q_1 = x - z$, $q_2 = y - z$, $q'_0 = z$.

Siano E_1 ed E_2 logicamente indipendenti ed E_3 l'evento somma:

$$E_3 = E_1 \vee E_2$$

poichè l'evento somma è pari a $E_1 + E_2 - E_1 E_2$ il seguente caso si riconduce al precedente. I costituenti sono i seguenti:

$$Q_0 = (0,0,0),\ Q'_1 = (0,1,1),\ Q'_2 = (1,0,1) \text{ e } Q'_0 = (1,1,1)$$

le disuguaglianze del tetraedro $\mathcal{P}$ aventi quei quattro vertici sono:

$$[\max(x, y) =] \quad x \vee y \leq z \leq 1 \wedge (x + y) \quad [= \min(1, x + y)]$$

i pesi per ottenere $P = (x, y, z)$ come baricentro dai Q sono $q_0 = 1 - z$, $q'_1 = z - y$, $q'_2 = z - x$, $q'_0 = x + y - z$.

Nei casi precedenti ogni P è esprimibile come baricentro dei Q_i con pesi q_i univocamente determinati e questo si verifica solo e soltanto quando i Q_i sono linearmente indipendenti.

Se affermiamo che E_1, E_2, E_3 sono esaustivi, i costituenti sono sette:

$$Q_1 = (1,0,0),\ Q_2 = (0,1,0),\ Q_3 = (0,0,1),\ Q'_1 = (0,1,1),\ Q'_2 = (1,0,1),$$
$$Q'_3 = (1,1,0) \text{ e } Q'_0 = (1,1,1)$$

mancando il vertice $Q_0 = (0, 0, 0)$ del cubo, l'involucro convesso $\mathcal{P}$ è il cubo stesso amputato del tetraedro determinato da tale vertice e dai tre adiacenti, cioè la parte del cubo $0 \leq x,\ y,\ z \leq 1$ soddisfacente la disuguaglianza $x + y + z \geq 1$. Ogni suo punto P si può esprimere, ma in infiniti modi (a meno che non coincida con un vertice o non appartenga ad uno spigolo o a una faccia triangolare), come baricentro di punti Q: basta considerare pesi q non negativi, di somma pari a 1, e tali che

$$q_1 + q'_2 + q'_3 + q'_0 = x, \qquad q_2 + q'_1 + q'_3 + q'_0 = y, \qquad q_3 + q'_1 + q'_2 + q'_0 = z$$

Se escludiamo entrambi i costituenti estremi, cioè sia $Q_0 = (0, 0, 0)$ sia $Q'_0 = (1, 1, 1)$, ne rimangono sei. Il cubo viene amputato di due tetraedri opposti, e rimane la parte $\mathcal{P}$ delimitata dalla doppia disuguaglianza $1 \leq x + y + z \leq 2$.

Casi analoghi sono:

- $E_3 \subset E_1E_2$ i costituenti sono cinque:

$$Q_0 = (0, 0, 0),\ Q_1 = (1, 0, 0),\ Q_2 = (0, 1, 0),\ Q'_3 = (1, 1, 0) \text{ e } Q'_0 = (1, 1, 1)$$

- $E_3 \subset (E_1 = E_2)$ i costituenti sono sei:

$$Q_0 = (0, 0, 0),\ Q_1 = (1, 0, 0),\ Q_2 = (0, 1, 0),\ Q_3 = (0, 0, 1),\ Q'_3 = (1, 1, 0) \text{ e } Q'_0 = (1, 1, 1)$$

- $E_3 \equiv (E_1 = E_2)$ i costituenti sono quattro:

$$Q_1 = (1, 0, 0),\ Q_2 = (0, 1, 0),\ Q_3 = (0, 0, 1) \text{ e } Q'_0 = (1, 1, 1)$$

  I primi tre sono sul piano $x + y + z = 1$, gli ultimi tre su $x = y + z - 1$, gli altri due gruppi di tre su $z = x + y - 1$ e su $y = x + z - 1$. Lo spazio $\mathcal{P}$ è quindi il tetraedro delimitato dalle seguenti disuguaglianze: $z \leq 1 - x - y$, $z \leq x - y + 1$, $z \leq x + y - 1$ e $z \leq y - x + 1$.
  Anche in questo caso P è esprimibile come baricentro dei Q

$$P = q_1Q_1 + q_2Q_2 + q_3Q_3 + q'_0Q'_0$$

con pesi rispettivamente $q_1 = 1 + x - y - z$, $q_2 = 1 - x + y - z$, $q_3 = 1 - x - y + z$ e $q'_0 = x + y + z - 1$.

Tutto quanto detto si può estendere in una dimensione qualunque. Se $E_1, E_2, ..., E_n$ sono eventi logicamente indipendenti, avremo 2^n costituenti Q_i, i vertici dell'ipercubo unitario, ossia punti dell'ambiente lineare $\mathcal{A}$ $(x_1, x_2, ..., x_n)$, con gli $x_i = 0$ o 1. Le previsioni $\mathbf{P}$ ammissibili sono quelle del cubo $\mathcal{P}$, involucro convesso dell'insieme $\mathcal{Q}$ costituito dai vertici. Il sistema lineare $\mathcal{L}$ è formato dai numeri aleatori $X = u_1E_1 + u_2E_2 + ... + u_nE_n$ linearmente dipendenti dagli eventi E_i.

Ogni caso particolare differisce da quello precedente per il fatto che una parte dei costituenti va esclusa: anzichè considerare 2^n costituenti ne avremo $s < 2^n$. Essi determineranno uno spazio lineare di dimensione d. Se $d < n$, gli n eventi E_i sono linearmente dipendenti, quindi eliminando gli E_i superflui si ottiene l'involucro convesso $\mathcal{P}$, un poliedro a d dimensioni, avente per vertici i Q che formano $\mathcal{Q}$.
Valutate le probabilità $\mathbf{P}(E_i)$ degli eventi dati, la $\mathbf{P}$ risulta determinata per tutti e soli i numeri aleatori X dipendenti linearmente dagli E_i, ossia appartenenti a $\mathcal{L}$; in particolare la probabilità di un evento E è determinata se e soltanto se E è un tale X. È interessante il caso in cui il numero aleatorio X possa assumere solo i valori 0 e 1, cioè nel caso di un evento; vedremo che, se E non è linearmente dipendente dagli E_i, si potrà soltanto dire che $p' \leq \mathbf{P}(E) \leq p''$ con $p' = \sup \mathbf{P}(X)$ per gli X di $\mathcal{L}$ certamente $\leq E$, e $p'' = \inf \mathbf{P}(X)$ per gli X di $\mathcal{L}$ certamente $\geq E$.

7.1 Teorema fondamentale per le probabilità

Un problema importante è quello dell'estensione di una valutazione coerente fatta sugli eventi di un certo insieme ad una valutazione ancora coerente sugli eventi di un soprainsieme arbitrario del primo. De Finetti dimostra che tale estensione si può sempre effettuare e, con riferimento ad un numero finito di eventi, stabilisce un risultato che nella letteratura viene indicato come *teorema fondamentale delle probabilità.* Un enunciato semplificato del teorema è il seguente.

Teorema 7.1.1. *Considerato un arbitrario insieme finito di n eventi E_i e assegnate, in modo coerente, le loro probabilità $\mathbf{P}(E_i)$, quella $\mathbf{P}(E_{n+1})$ di un altro evento E_{n+1}*

1. *risulta determinata univocamente se E_{n+1} dipende linearmente dagli n eventi E_i;*
2. *altrimenti può assumere un qualunque valore di un intervallo chiuso $[p', p'']$, i cui estremi dipendono dalle probabilità già assegnate $\mathbf{P}(E_i)$.*

Quindi dopo aver valutato le probabilità $p_1, p_2, ..., p_n$ dei rispettivi n eventi $E_1, E_2, ..., E_n$, assegnare la probabilità p_{n+1} all'ulteriore evento E_{n+1}, rispettando la condizione che la valutazione $(p_1, p_2, ..., p_n, p_{n+1})$, sulla famiglia $\{E_1, E_2, ..., E_n, E_{n+1}\}$ sia coerente si utilizza il risultato di De Finetti [1] che qui enunciamo in modo completo.

Teorema 7.1.2. ***Teorema fondamentale per le probabilità***. *Date le probabilità $\mathbf{P}(E_i)$ $(i = 1, 2, ..., n)$ di un numero finito di eventi, la probabilità di un altro evento E, $\mathbf{P}(E)$:*

1. *risulta determinata (qualunque sia $\mathbf{P}$) se E è linearmente dipendente dagli E_i;*
2. *altrimenti può esserle attribuito, coerentemente, un qualunque valore di un intervallo chiuso, $p' \leq \mathbf{P}(E) \leq p''$ (che potrà spesso dare una limitazione illusoria se $p' = 0$ e $p'' = 1$, o anche, in casi limiti per $\mathbf{P}$ particolari, dare un risultato determinato $p = p' = p''$).*

Precisamente, p' è l'estremo superiore, $\sup \mathbf{P}(X)$, delle valutazioni per difetto $\mathbf{P}(X)$ date da numeri aleatori X di $\mathcal{L}$ (cioè linearmente dipendenti dagli E_i) tali che certamente $X \leq E$. Se E non è logicamente dipendente dagli E_i, si *noti che $X \leq E$ si può più espressivamente sostituire con $X \leq E'$ ove E' è il massimo evento logicamente dipendente dagli E_i contenuto in E. Lo stesso dicesi per p'' (mutando* sup *in* inf*, e dovunque il senso delle disuguaglianze, massimo in minimo, E' in E'', ecc.).*

Osserviamo che se i due eventi E' ed E'' sono rispettivamente il primo somma di costiuenti del I tipo ed il secondo somma di costituenti del I e del III tipo risulta che $E' \subseteq E \subseteq E''$, ed E' è il massimo evento logicamente dipendente da $E_1, E_2, ..., E_n$ certamente contenuto in E, E'' è il minimo evento logicamente dipendente da $E_1, E_2, ..., E_n$ che certamente contiene E.

Dimostrazione. Detti $Q_1, Q_2, ..., Q_s$ i costituenti relativi ad $E_1, E_2, ..., E_n$, ed essendo E logicamente (ma non linearmente) dipendente dagli E_i, l'ambiente lineare $\mathcal{A}'$ ottenuto con l'aggiunta di E (ossia di una nuova coordinata x alle precedenti $x_1, x_2, ..., x_n$) ha i medesimi costituenti Q_h ma disposti sui vertici di un ipercubo ad $n+1$ dimensioni anzichè n: ogni $Q = (x_1, x_2, ..., x_n)$ viene lasciato al suo posto (su $x = 0$) o spostato sull'S_n parallelo ($x = 1$), diventando $(x_1, x_2, ..., x_n, 0)$ o $(x_1, x_2, ..., x_n, 1)$, a seconda che Q è contenuto in $\widetilde{E}$ o in E. L'involucro convesso $\mathcal{P}'$ in S_{n+1} (in $\mathcal{A}'$) ha come proiezione sull'S_n precedente ($\mathcal{A}$) il precedente $\mathcal{P}$; per ogni $\mathbf{P}$ ivi ammissibile (di coordinate $p_i = \mathbf{P}(E_i)$) le ammissibili estensioni in $\mathcal{A}'$ sono i punti $\mathbf{P}'$ che si proiettano su $\mathbf{P}$ e appartengono a $\mathcal{P}'$, ossia appartengono al segmento $p' \leq x \leq p''$ intersezione della retta $(p_1, p_2, ..., p_n, x)$ con $\mathcal{P}'$. I punti estremi ($x = p'$, $x = p''$) sono sul contorno di $\mathcal{P}'$ cioè su uno degli iperpiani (ad n dimensioni) che ne costituiscono le facce; sia esso $\sum u_i x_i + ux = c$, ossia sia soddisfatta su di esso la relazione $\sum u_i E_i + uE = c$ ossia la $E = (c - \sum u_i E_i)/u$: la $X \in \mathcal{L}$ a secondo membro realizza la condizione asserita e fornisce $p' = \mathbf{P}(X)$. Così per p''. □

Il teorema fondamentale per le probabilità permette di procedere anche in modo infinito ad attribuire a tutti gli eventi aleatori desiderabili, uno dopo l'altro, probabilità e previsioni coerenti con le precedenti. Vediamo ora alcune applicazioni.

Consideriamo il numero di successi

$$Y = E_1 + E_2 + ... + E_n$$

il caso più semplice è quello in cui il numero di successi è noto, sia $Y = y$, ossia è nota la *frequenza* $\frac{Y}{n} = \frac{y}{n}$; estendiamo questo esempio considerando i limiti entro cui Y è compreso tra dati valori estremi y' ed y'': $y' \leq Y \leq y''$ e quindi la frequenza deve rimanere tra $\frac{y'}{n}$ ed $\frac{y''}{n}$.
Nel primo caso, quello più semplice, la somma delle $\mathbf{P}(E_i)$, cioè $\mathbf{P}(Y)$ deve dare y, nel secondo un valore compreso tra y' e y''; quindi dividendo per n, le probabilità $\mathbf{P}(E_i)$ devono essere tali che la loro media aritmetica coincida con la frequenza nota $\frac{y}{n}$, oppure cada tra gli estremi $\frac{y'}{n}$ ed $\frac{y''}{n}$.
Per raggiungere lo scopo consideriamo l'ambiente lineare $\mathcal{A}^*$ generato dagli s costituenti Q_h oppure generato dagli E_i e dai loro prodotti. Supponiamo gli E_i logicamente indipendenti, e quindi $s = 2^n$, vediamo il caso particolare di tre eventi.
Siano E_1, E_2 ed E_3 eventi logicamente indipendenti ed $F = E_1E_2$, $G = E_1E_3$, $H = E_2E_3$ ed $E = E_1E_2E_3$ i loro prodotti, questi sette eventi sono linearmente indipendenti dato che fra i $2^3 = 8$ costituenti esiste una sola relazione lineare (somma = 1); tra essi sussitono però delle disuguaglianze ad esempio $E_1 \geq F \geq E$, come è ovvio pensando che dei $2^7 = 128$ vertici del cubo a sette dimensioni sono possibili solo gli otto corrispondenti ai costituenti relativi ad E_1, E_2 ed E_3.
Elenchiamo i costituenti con le coordinate nell'ambiente $\mathcal{A}^*$ e le rispettive espressioni lineari nello spazio duale $\mathcal{L}^*$:

$$E_1E_2E_3FGHE = (1,1,1,1,1,1,1) = E$$

$$E_1E_2\widetilde{E}_3F\widetilde{G}\widetilde{H}\widetilde{E} = (1,1,0,1,0,0,0) = F - E$$

$$E_1\widetilde{E}_2E_3\widetilde{F}G\widetilde{H}\widetilde{E} = (1,0,1,0,1,0,0) = G - E$$

$$\widetilde{E}_1E_2E_3\widetilde{F}\widetilde{G}H\widetilde{E} = (0,1,1,0,0,1,0) = H - E$$

$$E_1\widetilde{E}_2\widetilde{E}_3\widetilde{F}\widetilde{G}\widetilde{H}\widetilde{E} = (1,0,0,0,0,0,0) = E_1 - F - G + E$$

$$\widetilde{E}_1 E_2 \widetilde{E}_3 \widetilde{F}\widetilde{G}\widetilde{H}\widetilde{E} = (0,1,0,0,0,0,0) = E_2 - F - H + E$$

$$\widetilde{E}_1 \widetilde{E}_2 E_3 \widetilde{F}\widetilde{G}\widetilde{H}\widetilde{E} = (0,0,1,0,0,0,0) = E_3 - G - H + E$$

$$\widetilde{E}_1 \widetilde{E}_2 \widetilde{E}_3 \widetilde{F}\widetilde{G}\widetilde{H}\widetilde{E} = (0,0,0,0,0,0,0) = 1 - E_1 - E_2 - E_3 + F + G + H - E$$

Per la coerenza occorre e basta che le probabilità dei costituenti siano non negative (la somma è automaticamente pari a uno), e sono pertanto necessarie e sufficienti le seguenti disuguaglianze:

$$\mathbf{P}(E) \geq 0$$

$$\mathbf{P}(F), \mathbf{P}(G), \mathbf{P}(H) \geq \mathbf{P}(E)$$

$$\mathbf{P}(E_1) \geq \mathbf{P}(F) + \mathbf{P}(G) - \mathbf{P}(E)$$

$$\mathbf{P}(E_2) \geq \mathbf{P}(F) + \mathbf{P}(H) - \mathbf{P}(E)$$

$$\mathbf{P}(E_3) \geq \mathbf{P}(G) + \mathbf{P}(H) - \mathbf{P}(E)$$

$$(\mathbf{P}(E_1) + \mathbf{P}(E_2) + \mathbf{P}(E_3)) - (\mathbf{P}(F) + \mathbf{P}(G) + \mathbf{P}(H)) + \mathbf{P}(E) \leq 1$$

7.1.1 Un semplice esempio

Un imputato innocente deve essere giudicato da una giuria composta da tre giurati il cui verdetto finale è raggiunto a maggioranza. I tre giurati A, B e C assumono la loro decisione indipendentemente.
Calcolare la probabilità che l'imputato sia assolto.

Siano E_A, E_B ed E_C gli eventi ai quali attribuiamo il seguente significato:

$$E_A = \text{“A assolve l'imputato”}$$
$$E_B = \text{“B assolve l'imputato”}$$
$$E_C = \text{“C assolve l'imputato”}$$

Se consideriamo i costituenti logici associati agli eventi E_A, E_B ed E_C vediamo che l'evento:

$$E = \text{“l'imputato é assolto da A, B e C”}$$

è somma logica dei costituenti $E_A E_B E_C$, $E_A E_B \widetilde{E}_C$, $E_A \widetilde{E}_B E_C$ ed $\widetilde{E}_A E_B E_C$.

7.2 Numeri aleatori con infiniti valori possibili

Consideriamo un'infinità numerabile di valori possibili x_h con $h = 1, 2, ...$ ad essi corrispondono le probabilità p_h positive o nulle con

$$\sum_h p_h = 1 - p^* \leq 1 \qquad 0 \leq p^* \leq 1$$

Per ogni intervallo o insieme I, conoscendo solo i valori possibili $x_h \in I$ e le rispettive p_h, possiamo dire che dato $X \in I$, se I contiene un insieme finito di punti

$$\mathbf{P}(X) = \sum_h p_h x_h$$

altrimenti, se ne contiene infiniti possiamo affermare solamente che

$$\sum_h p_h x_h \leq \mathbf{P}(X) \leq \sum_h p_h x_h + p^*$$

Se x è punto di accumulazione per gli x_h abbiamo delle probabilità *aderenti* (non nulle). Le probabilità aderenti si distinguono in aderenti a sinistra e aderenti a destra, queste si definiscono rispettivamente come il limite di $\mathbf{P}(x - \varepsilon < X < x)$ e di $\mathbf{P}(x < X < x + \varepsilon)$ per $\varepsilon \to 0$. Le probabilità aderenti non possono superare p^* neppure se considerate complessivamente, quindi le probabilità aderenti potrebbero avere o probabilità $< p^*$, o non esistere e in questo caso la loro probabilità sarebbe nulla; pur valendo per p^* la seguente doppia disuguaglianza $0 \leq p^* \leq 1$.

Ci chiediamo cosa possiamo affermare riguardo alla previsione $\mathbf{P}(X)$ conoscendo i valori possibili x_h e le rispettive probabilità p_h, cioè seguendo l'impostazione soggettivista [1] quali vincoli vengono imposti dalla conoscenza degli x_h e dalla valutazione delle p_h alla quale vogliamo rimanere coerenti per il calcolo della previsione di X.

Prima di studiare il caso di un numero aleatorio X *illimitato* vediamo quello di X *limitato*, prendiamo il minimo e il massimo dei punti di accumulazione degli x_h, indichiamoli rispettivamente con x' e x'', avremo quindi la seguente catena di disuguaglianze

$$-\infty < \inf X \leq x' \leq x'' \leq \sup X < +\infty$$

Proposizione 7.2.1. *Se* $p^* = 0$ *risulta univocamente* $\mathbf{P}(X) = \sum_h p_h x_h$ *(come nel caso finito); al di fuori di tale caso particolare possiamo solo dire che*

$$\sum_h p_h x_h + p^* x' \leq \mathbf{P}(X) \leq \sum_h p_h x_h + p^* x''$$

e $\mathbf{P}(X)$ *risulta univocamente determinato se* $p^* = 0$ *o se e solo se* $x' = x''$*, ossia se gli* x_h *hanno un unico punto di accumulazione, che è quindi un limite al quale convergono.*

Dimostrazione. Fissato $\varepsilon > 0$ prendiamo N abbastanza grande al fine che risulti $\sum p_h < \varepsilon$ con $h \geq N$, e poniamo $X = X_1 + X_2 + X_3$ con

$$X_1 = X = x_h \text{ se } h < N, \text{ altrimenti } X_1 = 0$$
$$X_2 = X = x_h \text{ se } h \geq N \text{ e } x_h < x' - \varepsilon \text{ oppure } x_h > x'' + \varepsilon, \text{ altrimenti } X_2 = 0$$
$$X_3 = X = x_h \text{ se } h \geq N \text{ e } x' - \varepsilon \leq x_h \leq x'' + \varepsilon, \text{ altrimenti } X_3 = 0$$

Abbiamo

$$\mathbf{P}(X_1) = \textstyle\sum_h p_h x_h \ (h < N) \to \sum_h p_h x_n \ (x_h \text{ limitate})$$
$$\varepsilon \inf X \leq \mathbf{P}(X_2) \leq \varepsilon \sup X$$

perchè i valori possibili tra $\inf X$ e $x' - \varepsilon$ e tra $x'' + \varepsilon$ e $\sup X$ sono al più in numero finito, e la probabilità complessiva di quelli tra essi con $h \geq N$ è la somma di un numero finito dei p_h per i quali la somma della serie è $< \varepsilon$; e infine

$$p^*(x' - \varepsilon) \leq \mathbf{P}(X_3) \leq p^*(x'' + \varepsilon)$$

Tutto ciò vale per ogni ε e quindi per $\varepsilon \to 0$ si ha la tesi. □

Passiamo ora al caso di X *illimitato*, consideriamo dapprima il caso di X illimitato unilateralmente e precisamente di illimitatezza superiore, il caso generale deriverà come corollario.
Sia $X \geq 0$ ossia $\inf X \geq 0$, nel caso generale basterà porre $X = X_1 - X_2$ con

$X_1 = 0 \vee X$ e $X_2 = |0 \wedge X|$.
Inoltre supponiamo non esistano punti d'accumulazione al finito, potremo quindi supporre le x_h in ordine crescente, e tendenti a $+\infty$ al divergere di h.
In tali condizioni, posto

$$P_n = \sum_{h=1}^{n} p_h \qquad P = \lim P_n \qquad p^* = 1 - P$$

$$S_n = \sum_{h=1}^{n} p_h x_h \qquad S = \lim S_n$$

abbiamo

$$P_n = \mathbf{P}(X \leq x_n) \qquad 1 - P_n = \mathbf{P}(X > x_n)$$

p^* = massa aderente a sinistra di $+\infty$, oppure in $x = +\infty$;

$$S_n = \mathbf{P}\big\{X(X \leq x_n)\big\} \qquad S_n + x_n(1 - P_n) = \mathbf{P}(X \wedge x_n)$$

previsione di x "troncato" a x_n.Poichè ogni X troncata è sempre $\leq X$, si ha

$$\begin{gathered}\mathbf{P}(X) \geq S_n + x_n(1 - P_n) \text{ per qualsiasi } n \text{, e quindi} \\ \mathbf{P}(X) \geq S + x_n(1 - P) = S + x_n p^* \text{ per qualsiasi } n\end{gathered}$$

Da qui possiamo dire subito che $\mathbf{P}(X) = \infty$, se $S = \infty$ (la serie $\sum p_h x_h$ diverge) oppure se $p^* \neq 0$ (esiste una probabilità collocata o aderente in $+\infty$).
Nel caso opposto, cioè se $p^* = 0$ e S finito ($\sum p_h = 1$ e $\sum p_h x_h$ convergente) è ammissibile per $\mathbf{P}(X)$ la valutazione

$$\mathbf{P}(X) = S = \sum_{h=0}^{\infty} p_h x_h$$

o qualunque valore superiore, $+\infty$ incluso.

Ciò si dimostra *per continuità*. Poniamo $X'_n = X(X \leq x_n)$ (X amputata), con $p'_h = p_h$ per $h \leq n$, $p'_h = 0$ per $h > n$ e $p'_0 = \sum p_h(h > n) = \mathbf{P}(X'_n = 0)$; al crescere di n tutte le $p_h = \mathbf{P}(X'_n = h)$ tendono a p_h; ma $\mathbf{P}(X'_n) = S_n \to S$. Poniamo poi $X''_n = X'_n + a_n(X > n)$, cioè X''_n coincide con X fino ad x_n, ma, quando lo supera, viene sostituito con a_n; e con a_n indichiamo il primo degli

x_h per cui $x_h p_0 \geq n$. Il valore a_n dà un contributo $\geq n$, quindi senz'altro $\mathbf{P}(X''_n) \geq n \to \infty$.
Schematizziamo la conclusione raggiunta nei due casi:

$$p^* > 0 \qquad \mathbf{P}(X) = +\infty$$

$$p^* = 0 \begin{cases} S = +\infty & \mathbf{P}(X) = +\infty \\ S =< +\infty & S \leq \mathbf{P}(X) \leq +\infty \end{cases}$$

Se X è bilateralmente illimitato $\mathbf{P}(X)$ è completamente indeterminato. Consideriamo la valutazione *privilegiata* consistente nel prendere sia per la parte positiva $0 \vee X$ che per quella negativa $0 \wedge X$ la previsione minima (in valore assoluto) ammissibile, indichiamola con $\widehat{\mathbf{P}}$ e ponendo in generale

$$\widehat{\mathbf{P}}(X) = \widehat{\mathbf{P}}(0 \vee X) - \widehat{\mathbf{P}}(|0 \wedge X|)$$

o brevemente

$$\widehat{S} = S^+ + S^-$$

Questa previsione asintotica risulta:

$$\widehat{S} = S^+ + S^- \begin{cases} \textit{finita}, \text{ se lo sono } S^+ \text{ ed } S^-; \\ \textit{infinita}, \text{ se lo è una: } +\infty \text{ se } S^+ = +\infty; \ -\infty \text{ se } S^- = -\infty; \\ \textit{non definita}, \text{ se entrambe sono infinite.} \end{cases}$$

7.2.1 La proprietà di continuità

La proprietà dice che la coerenza si conserva in un passaggio al limite [1]. La proprietà non vale nel caso dell'additività completa. Essa è molto utile per dimostrazioni di ammissibilità, come quella precedente sul caso di previsione di un X illimitato superiormente.

Teorema 7.2.2. *Siano $\mathbf{P}_n(E)$ delle valutazioni di probabilità (coerenti) definite per un medesimo campo di eventi $\mathcal{E}$, e poniamo $\mathbf{P}(E) = \lim \mathbf{P}_n(E)$ quando esiste (e sia $\mathcal{E}' \subseteq \mathcal{E}$ l'insieme degli E per cui il limite esiste). In tale campo la $\mathbf{P}(E)$ costituisce anch'essa una valutazione di probabilità (coerente).*

Dimostrazione. Le condizioni di coerenza sono espresse da equazioni (o disequazioni) lineari implicanti un numero finito di elementi (eventi, o numeri aleatori); nel passaggio al limite si conservano. □

Bibliografia

[1] B. de Finetti, *Teoria delle probabilità*, Giulio Einaudi, Torino, 1970.

[2] B. de Finetti a cura di P. Monari e D. Cocchi, *Probabilità e induzione*, Clueb, Bologna, maggio 1993.

[3] F. Biagini, M. Campanino, *Elementi di probabilità e statistica*, Springer, Milano, 2006.

[4] http://www.brunodefinetti.it/

[5] http://it.wikipedia.org/wiki/Probabilità

[6] http://www.unipa.it/sanfilippo/pub/sigad/altro/treeventi/

[7] http://www.dm.unibo.it/∼campanin/htdocs/diddist/1.html

Ringraziamenti

Il primo ringraziamento desidero rivolgerlo al Professor Massimo Campanino per la grande disponibilità, la pazienza, i consigli e l'aiuto che mi ha dato durante tutta la stesura della tesi.

Un grazie speciale alla mia famiglia che mi è sempre stata vicina, soprattutto nei momenti più duri, che mi ha sostenuto e fatto arrivare fino a questo punto.