

Appunti Lab

Francesco Sermi

Indice

Capitolo 1

Introduzione _____ Pagina 2

- 1.1 L'errore massimo e le sue limitazioni 2
- 1.2 Precisione vs accuratezza 3

Capitolo 2

Probabilità _____ Pagina 4

- 2.1 Definizione operativa della probabilità 6
 - Definizione combinatoriale — 6 • Definizione frequentista — 6 • Definizione soggettivista della probabilità — 7
- 2.2 Elementi di calcolo combinatorio 7
- 2.3 Probabilità condizionata 10
- 2.4 Teorema di Bayes 11
- 2.5 Variabili casuali e funzioni di distribuzione 12

Capitolo 1

Introduzione

Il concetto di *incertezza* e di *misura* gioca un ruolo centrale nelle scienze sperimentali siccome noi **non** siamo mai in grado di misurare una grandezza fisica con un'accuratezza infinita, dunque il nostro risultato manca di una parte sostanziosa del proprio contenuto se non vi è assegnata una stima dell'incertezza compiuta nella misura effettuata.

Il primo sforzo nella storia delle scienze verso un sistema standard di unità di misure è stato costituito, nella Francia del XVIII secolo, del sistema metrico decimale. Una domanda molto interessante che il lettore si potrebbe chiedere sarebbe la seguente *"Come possiamo definire un'unità di misura?"* La risposta più semplice, ma non banale, è tramite dei campioni inalterabili e riproducibili: si riporta proprio la realizzazione del metro campione e del kilogrammo campione realizzati, sebbene subiscano una contaminazione, ogni anno, pari a $1\mu\text{g}$.

Oggi è stato definito il Sistema internazionale delle misure che si basa sul fissare 7 unità di base da cui è possibile esprimere tutte le altre come combinazione delle altre (si rimanda la lettura delle dispense di Baldini per usufruire dei suoi grafici e/o tabelle).

1.1 L'errore massimo e le sue limitazioni

Si osserva che nella sua formulazione più elementare il concetto di *errore massimo* è legato alla domanda *"Qual è il più piccolo intervallo che contiene con certezza il valore numerico della quantità che sto misurando?"*. In altre parole, il risultato della misura di una generica grandezza fisica x risulta essere pari a:

$$x = \hat{x} \pm \Delta x \quad (1.1)$$

intendo come Δx l'errore massimo, ovvero stiamo dicendo che l'intervallo $[\hat{x} - \Delta x; \hat{x} + \Delta x]$ è il più piccolo intervallo possibile che ci dà la certezza di includere il valore incognito x . Definiamo adesso una serie di termini comodi per esprimere al meglio i concetti successivi:

- x è il valore, incognito (siccome non lo conosciamo mai appieno), della grandezza che vogliamo misurare e che chiameremo *misurando*;
- \hat{x} è la migliore stima di x che possiamo fornire a partire dei dati a nostra disposizione (e che chiameremo *valore centrale* o *migliore stima* della nostra misura);
- Δx è l'incertezza di misura e, in questo **caso**, coincide con l'**errore massimo**

Tuttavia, nel caso di misure ripetute, giungiamo ad un assurdo: infatti, se una misura fluttua, da un punto di vista operativo, non possiamo escludere che una nuova misura della grandezza non fornisca un valore al di fuori dell'intervallo iniziale di incertezza e, ove questo accade, siamo costretti ad allargare tale intervallo, giungendo, dunque, ad un evidente paradosso, siccome acquisire nuove informazioni può solo peggiorare (o lasciare invariato, se otteniamo delle misure che fanno ancora parte di $[x - \Delta x; x + \Delta x]$) il nostro stato di conoscenza, almeno determinando l'incertezza come errore massimo.

Questo è il motivo per cui non utilizziamo quasi mai il concetto di errore massimo, ma utilizzeremo il concetto di errore statistico.

$$x = \hat{x} \pm \sigma_x \quad (1.2)$$

La 1.2 ha un significato diverso da 1.1 prima siccome questa definisce un intervallo che non ci dà la *certezza* ma solo la probabilità, ben definita, di contenere il valore del misurando.

Adesso definiamo un altro modo molto utile per quantificare quanto è grosso, rispetto alla misura, l'errore che noi commettiamo:

Definizione 1.1: Errore percentuale

L'errore relativo è il rapporto tra l'incertezza della misura e il suo valore centrale:

$$e_{\%} = \frac{\sigma_x}{|\hat{x}|} \quad (1.3)$$

1.2 Precisione vs accuratezza

E' molto sottile la distinzione fra la precisione di uno strumento e la sua accuratezza: con il primo termine si indica l'accordo tra il valore misurato dallo strumento e quello effettivo del misurando, mentre con il secondo si intende il grado di consistenza fra i risultati di misure successive della stessa quantità nelle medesime condizioni

Capitolo 2

Probabilità

Grazie al matematico Kolmogorov abbiamo la prima costruzione rigorosa della teoria della probabilità, in una struttura che, sostanzialmente, sopravvive ancora oggi nei manuali moderni.

La struttura di base su cui si fonda la struttura assiomatica della probabilità parte dal definire lo spazio campionario

Definizione 2.1: Spazio campionario Ω

Lo spazio campionario Ω è l'insieme (numerabile) di tutte le possibili realizzazioni elementari di un dato fenomeno e spazio degli eventi \mathcal{F} l'insieme di tutti i sottoinsiemi di Ω tale che:

$$\#\mathcal{F} = 2^{\#\Omega}$$

L'idea di Kolmogorov è quella di definire la probabilità direttamente sullo spazio degli eventi-cioè possiamo assegnare una probabilità non solo ad ogni elemento dello spazio campionario, ma anche ad uno qualsiasi dei suoi sottoinsiemi. Definiamo adesso il concetto di probabilità:

Definizione 2.2: Probabilità

Definiamo probabilità una misura P su \mathcal{F} che associ univocamente ad ogni elemento E di \mathcal{F} un numero reale $P(E)$ che soddisfa le seguenti tre proprietà (o assiomi di Kolmogorov):

- ① $0 \leq P(E) \leq 1 \forall E \in \mathcal{F}$
- ② $P(\Omega) = 1$
- ③ $P(E_1 \cup E_2) = P(E_1) + P(E_2)$ se $E_1 \cap E_2 = \emptyset$

Possiamo utilizzare il terzo assioma di Kolmogorov all'unione numerabili di eventi *disgiunti* (ovvero eventi per cui la loro intersezione è nulla)

Corollario 2.1 Unione numerabili di eventi disgiunti

$$P\left(\bigcup_i^n E_i\right) = \sum_i^n P(E_i) \text{ se } E_1 = E_2 = \dots = E_n = \emptyset$$

Dimostrazione: Si procede per induzione su n . Per $n = 1$ è banale, siccome:

$$P\left(\bigcup_i^1 E_i\right) = P(E_1) = \sum_i^1 P(E_i) = P(E_1)$$

Adesso mostriamo $n \implies n + 1$:

$$\sum_i^{n+1} P(E_i) = P(E_1) + \dots + P(E_{n+1}) = (\text{ip. induttiva}) P\left(\bigcup_i^n E_i\right) + P(E_{n+1}) = P\left(\sum_i^{n+1} E_i\right)$$

Siccome possiamo vedere nell'ultimo passaggio la somma fra due eventi che sono fra loro disgiunti, ovvero fra l'evento $E_1 \cup E_2 \cdots \cup E_n$ e l'evento E_{n+1} e per ipotesi sappiamo che $E_i \cap E_j = \emptyset \forall i, j$.
La dimostrazione è dunque conclusa. ☺

Corollario 2.2 Probabilità complementare

Dato un evento E e detto \bar{E} il suo complementare in Ω , si ha che:

$$P(\bar{E}) = 1 - P(E)$$

Dimostrazione: Sapendo che $\bar{E} \cap E = \emptyset$ ma $\bar{E} \cup E = \Omega$:

$$1 = P(\Omega) = P(\bar{E} + E) = P(\bar{E}) + P(E) \implies P(\bar{E}) = 1 - P(E)$$

☺

Corollario 2.3 Probabilità dell'insieme nullo

$$P(\emptyset) = 0$$

Dimostrazione:

$$P(\Omega) = P(\Omega \cup \emptyset)$$

ma siccome $\Omega \cap \emptyset = \emptyset$ allora

$$P(\Omega \cup \emptyset) = P(\Omega) + P(\emptyset) = P(\Omega) \implies 1 + P(\emptyset) = 1 \implies P(\emptyset) = 0$$

La dimostrazione è dunque conclusa. ☺

Corollario 2.4 Limitatezza della probabilità di un sottoinsieme

$$\text{Se } E_1 \subset E_2 \implies P(E_1) \leq P(E_2)$$

Dimostrazione: Se $E_1 \subset E_2 \implies E_2 = E_1 + (E_2 \setminus E_1)$ ma siccome $E_1 \cap (E_2 \setminus E_1) = \emptyset$:

$$P(E_2) = P(E_1) + P(E_2 \setminus E_1) \implies P(E_1) \leq P(E_2)$$

☺

Teorema 2.1 Addizione delle probabilità

Dati due eventi E_1 ed E_2 , si ha che:

$$P(E_1 \cup E_2) = P(E_1) + P(E_2) - P(E_1 \cap E_2) \quad (2.1)$$

Dimostrazione: Possiamo scrivere gli insiemi E_1 ed E_2 come unione di eventi disgiunti, infatti:

$$E_2 = E_2 \cap \Omega = E_2 \cap (E_1 \cup \bar{E}_1) = (E_2 \cap E_1) \cup (E_2 \cap \bar{E}_1)$$

ma siccome $(E_2 \cap E_1) \cap (E_2 \cap \bar{E}_1) = \emptyset$ (altrimenti si giungerebbe ad un assurdo, visto che se $E_2 \cap E_1$ avesse degli elementi in comune con $E_2 \cap \bar{E}_1$ implicherebbe, visto che l'intersezione fra insiemi è un'operazione che gode di proprietà associativa e commutativa, che $E_1 \cap \bar{E}_1 \cap E_2 \cap E_2 \neq \emptyset$ ma $E_1 \cap \bar{E}_1 = \emptyset$, dunque assurdo). Tornando alla dimostrazione:

$$P(E_2) = P((E_2 \cap E_1) \cup (E_2 \cap \bar{E}_1)) = P(E_2 \cap E_1) + P(E_2 \cap \bar{E}_1) \quad (2.2)$$

d'altra parte abbiamo che

$$E_1 \cup E_2 = (E_1 \cup E_2) \cap \Omega = (E_1 \cup E_2) \cap (E_1 \cup \bar{E}_1) = (E_1 \cap E_1) \cup (E_1 \cap \bar{E}_1) \cup (E_2 \cap E_1) \cup (E_2 \cap \bar{E}_1)$$

$= E_1 \cup (E_2 \cap E_1) \cup (E_2 \cap \bar{E}_1)$ ma si osserva che il termine $E_1 \cup (E_2 \cap E_1)$ è ridondante, siccome $E_1 \cup (E_2 \cap E_1) = E_1$, dunque

$$E_1 \cup E_2 = E_1 \cup (E_2 \cap \bar{E}_1) \implies P(E_1 \cup E_2) = P(E_1) + P(E_2 \cap \bar{E}_1)$$

e combinando le due relazioni ottenute

$$\begin{cases} P(E_2) = P(E_2 \cap E_1) + P(E_2 \cap \bar{E}_1) \\ P(E_1 \cup E_2) = P(E_1) + P(E_2 \cap \bar{E}_1) \end{cases}$$

dunque

$$P(E_1 \cup E_2) = P(E_1) + P(E_2) - P(E_1 \cap E_2)$$

☺

2.1 Definizione operativa della probabilità

Il lettore più attento si sarà però accorto di come la definizione che abbiamo dato di probabilità non ci dà alcun modo con cui calcolare la probabilità, ma piuttosto una serie di assiomi da cui possiamo ricavare una serie di proprietà utili di cui gode la probabilità.

2.1.1 Definizione combinatoriale

Nella sua definizione combinatoriale, definiamo la probabilità di un evento E con quanto segue

Definizione 2.3: Definizione combinatoriale della probabilità

La probabilità di un evento E coincide con il rapporto tra il numero di casi favorevoli n e il numero di casi possibili N , a condizione che questi siano tutti equiprobabili

$$P(E) = \frac{n}{N} \quad (2.3)$$

Osservazione:-

Si osserva che questa definizione di probabilità rispetta i tre assiomi di Kolmogorov: il primo assioma discende dalla ovvia condizione per cui $0 \leq n \leq N$ ed il secondo fatto deriva dal fatto che se $n = N \implies P(E) = 1$. Per quanto riguarda la terza condizione, si osserva che se E_1 e E_2 sono due eventi disgiunti con rispettivamente n_1 e n_2 casi favorevoli, allora:

$$P(E_1 \cup E_2) = \frac{n_1 + n_2}{N} = \frac{n_1}{N} + \frac{n_2}{N} = P(E_1) + P(E_2)$$

dunque anche il terzo assioma è rispettato

Tuttavia questa definizione operativa di probabilità possiede un grande problema: nella definizione è compiuto infatti un ragionamento circolare, siccome richiediamo l'equiprobabilità dei casi nella definizione stessa di probabilità.

2.1.2 Definizione frequentista

Quando è possibile ripetere un esperimento in condizioni controllate, è possibile definire la probabilità di un evento E come il limite della frequenza relativa all'evento stesso quando il numero di ripetizioni N dell'esperimento tende all'infinito. Possiamo quindi pensare di effettuare un esperimento un numero N arbitrariamente grande di volte, contare le n volte in cui è avvenuto l'evento E e definire la probabilità $P(E)$ dell'evento come il limite del rapporto $\frac{n}{N}$

Definizione 2.4: Definizione frequentista della probabilità

La probabilità di un evento E si definisce come

$$P(E) = \lim_{N \rightarrow +\infty} \frac{n}{N} \quad (2.4)$$

dove il limite va inteso nei termini della convergenza statistica, ovvero

$$\forall \epsilon, \delta > 0 \exists \tilde{N} > 0 : \forall N, N > \tilde{N} \implies P\left(\left|\frac{n}{N} - P(E)\right| \geq \delta\right) \leq \epsilon \quad (2.5)$$

Osservazione:-

Il senso di questo limite, che va inteso come limite in senso statistico piuttosto che nel senso usuale dell'analisi matematica, è il fatto che non è possibile garantire a priori l'esistenza di un numero N di ripetizioni del nostro esperimento che mi permetta di affermare con certezza che la differenza tra la frequenza registrata $\frac{n}{N} - P(E)$ sia minore di una certa quantità ϵ : infatti, se effettuiamo due diverse serie di N ripetizioni dell'esperimento otterrò frequenze relative $\frac{n}{N}$ diverse. Quello che possiamo però dire è il fatto che se N è abbastanza grande allora posso rendere piccola a piacere la probabilità che $\frac{n}{N}$ si discosti da $P(E)$ di un valore prefissato δ .

Esempio 2.1 (Lancio di un dado)

Se lanciamo N volte un dado equo a sei facce e registriamo (al crescere di N) il numero n di volte in cui esce, ad esempio, il numero 3, per N molto grande il rapporto $\frac{n}{N}$ tenderà a $P(3) = \frac{1}{6}$ (nel senso della convergenza statistica)

2.1.3 Definizione soggettivista della probabilità

Definizione 2.5: Definizione soggettivista

La probabilità di un evento E si identifica con la misura del grado di fiducia che un individuo attribuisce al verificarsi di E , sulla base dell'informazione a sua disposizione

Osservazione:-

Il termine "soggettivo" si riferisce al fatto che persone diverse, sulla base di differenti informazioni, assoceranno, in generale, una probabilità diversa allo stesso evento e, proprio per questo fatto, si dice che questa definizione è *soggettiva*.

Alla fine, sebbene questa definizione lasci inizialmente sbigottiti siccome si perde quell'oggettività che, in un certo senso, assicuravano le altre due definizioni, questa definizione rispetto di fatto come noi operiamo nella vita di tutti i giorni. All'interno della scuola soggettivista vi sono diversi approcci distinti per derivare le regole fondamentali della probabilità in un modo logicamente consistente (sebbene, con questo *approccio*, gli assiomi non sono tali, ma regole che si ricavano da un principio più formale): il più popolare di questi approcci è il principio della **scommessa coerente**, il quale afferma che *una volta assegnata la probabilità ad un evento dovremo essere disposti ad accettare scommesse sul verificarsi dell'evento stesso con un rapporto tra puntata e vincita determinato dalla probabilità stessa*¹

2.2 Elementi di calcolo combinatorio

La combinatoria è quella branca della matematica che si occupa del *contare*, dunque è strettamente connessa alla probabilità.

¹Il senso è che se diciamo che due eventi sono equiprobabili allora dobbiamo essere pronti ad accettare scommesse 1:1, ovvero che una *scommessa* è coerente se e solo se non determina **a priori** una perdita per il banco o per lo scommettitore, dunque è equivalente anche se i ruoli fossero scambiati

Introduciamo il fattoriale di un numero nella seguente maniera

Definizione 2.6: Fattoriale

Dato un numero $n \in \mathbb{N}$, indichiamo con $n!$

$$n! = \prod_{k=1}^n k = n(n-1) \cdots 1 \quad (2.6)$$

Se ci pensiamo bene, la funzione fattoriale non è altro il numero di *permutazioni*, ovvero il numero di modi in cui si possono disporre n elementi se non possiamo ripeterli e **conta** l'ordine con cui questi elementi vengono disposti: supponiamo infatti di avere 20 oggetti e di volerli disporre all'interno di un cassetto, noi abbiamo ben $20!$ modi possibili per metterli, siccome per il primo "posto" del cassetto possiamo metterci uno dei 20, nel secondo "posto" 19 oggetti e così via; ottenendo ben $20!$ fattoriale di possibilità.

Le permutazioni possono essere viste come un caso particolare delle disposizioni, ovvero i modi con cui è possibile disporre n oggetti in k posti: per esempio, riprendendo l'esempio di prima, se noi volessimo prendere da 20 oggetti 5, presi a caso da questi 20, all'interno di un cassetto in maniera tale che conti l'ordine con cui li mettiamo, si osserva che nel primo *posto* del cassetto abbiamo 20 oggetti disponibili, nel secondo 19, nel terzo 18, nel quarto 17 e nell'ultimo 16; dunque affermare che i modi totali sono $\frac{20!}{15!} = 20 \cdot 19 \cdot 18 \cdot 17 \cdot 16$. Definiamo quindi una disposizione come

Definizione 2.7: Disposizione di n elementi e di ordine k

Definiamo una disposizione di n elementi e di ordine k come il numero di modi con cui è possibile disporre n elementi in k slot, pari a

$$D_{n,k} = \frac{n!}{(n-k)!} \quad (2.7)$$

Tornando alla funzione fattoriale, è comoda l'approssimazione di Stirling (di cui non daremo una dimostrazione) per cui:

$$n! = \sqrt{2\pi n} \left(\frac{n}{e}\right)^n \quad (2.8)$$

Il numero di modi con cui è possibile scegliere k elementi non ordinati è dato dal coefficiente binomiale n su k

Definizione 2.8: Coefficiente binomiale n su k

Il coefficiente binomiale n su k è definito come

$$\binom{n}{k} = \frac{n!}{k!(n-k)!} \quad (2.9)$$

Corollario 2.5

$$\binom{n}{k} = \binom{n}{n-k}$$

Dimostrazione: Banalmente, si osserva che

$$\binom{n}{k} = \frac{n!}{k!(n-k)!} = \frac{n!}{(n-k)!(n-(n-k))!} = \binom{n}{n-k}$$

☺

Il coefficiente binomiale è strettamente connesso al triangolo di Pascal (o di Newton) e può essere anche caratterizzato attraverso la formula della potenza del binomio (o teorema binomiale). Per farlo però ci servirà il seguente lemma:

Lemna 2.2.1 Proprietà fondamentale del coefficiente binomiale

$$\binom{n+1}{k+1} = \binom{n}{k} + \binom{n}{k+1}$$

Dimostrazione:

$$\begin{aligned} \binom{n}{k+1} + \binom{n}{k} &= \frac{n!}{(k+1)!(n-k-1)!} + \frac{n!}{k!(n-k)!} = \frac{n!}{(k+1)k!(n-k-1)!} + \frac{n!}{k!(n-k)(n-k-1)!} \\ &= \frac{n!}{k!(n-k-1)!} \left(\frac{1}{k+1} + \frac{1}{n-k} \right) = \frac{n!}{k!(n-k-1)!} \left(\frac{n-k+k+1}{(k+1)(n-k)} \right) = \frac{n!}{k!(n-k-1)!} \cdot \frac{(n+1)}{(k+1)(n-k)} \\ &= \frac{(n+1)!}{(k+1)!(n-k)!} = \binom{n+1}{k+1} \end{aligned}$$

⊕

Teorema 2.2 Teorema binomiale

Lo sviluppo della potenza n -esima del binomio $x_1 + x_2$ è pari a

$$(x_1 + x_2)^n = \sum_{k=0}^n \binom{n}{k} x_1^{n-k} x_2^k \quad (2.10)$$

Dimostrazione: si può procedere per induzione. Innanzitutto si osserva che, per $n = 0$, si ha che

$$(x_1 + x_2)^0 = 1 = \sum_{k=0}^0 \binom{0}{k} x_1^{0-k} x_2^k = \binom{0}{0} x_1^0 x_2^0 = 1$$

dunque per $n = 0$ l'ipotesi è verificata. Mostriamo che $n \implies n+1$:

$$(x_1 + x_2)^{n+1} = (x_1 + x_2)^n (x_1 + x_2) = \sum_{k=0}^n \binom{n}{k} x_1^{n-k} x_2^k \cdot (x_1 + x_2) = \sum_{k=0}^n \binom{n}{k} x_1^{n+1-k} x_2^k + \sum_{k=0}^n \binom{n}{k} x_1^{n-k} x_2^{k+1}$$

Da qua soffermiamoci sul primo termine: si osserva che

$$\sum_{k=0}^n \binom{n}{k} x_1^{n+1-k} x_2^k = \binom{n}{0} x_1^{n+1} + \sum_{k=1}^n \binom{n}{k} x_1^{n+1-k} x_2^k = \binom{n}{0} x_1^{n+1} + \sum_{k=0}^{n-1} \binom{n}{k+1} x_1^{n+1-k-1} x_2^{k+1} = x_1^{n+1} + \sum_{k=0}^{n-1} \binom{n}{k+1} x_1^n x_2^{k+1}$$

in cui si è semplicemente effettuato un cambio di variabile di variabile $k' = k+1$ (ma siccome gli indici sono muti abbiamo semplicemente indicato k' sempre con k). Adesso andiamo al secondo termine:

$$\sum_{k=0}^n \binom{n}{k} x_1^{n-k} x_2^{k+1} = \sum_{k=0}^{n-1} \binom{n}{k} x_1^{n-k} x_2^{k+1} + \binom{n}{n} x_1^{n-n} x_2^{n+1} = \sum_{k=0}^{n-1} \binom{n}{k} x_1^{n-k} x_2^{k+1} + x_2^{n+1}$$

Tornando dunque alla relazione iniziale, si deve avere che

$$(x_1 + x_2)^{n+1} = x_1^{n+1} + \sum_{k=0}^{n-1} \binom{n}{k+1} x_1^n x_2^{k+1} + \sum_{k=0}^{n-1} \binom{n}{k} x_1^{n-k} x_2^{k+1} + x_2^{n+1} = x_1^{n+1} + \sum_{k=0}^{n-1} \left[\binom{n}{k} + \binom{n}{k+1} \right] x_1^{n-k} x_2^{k+1} + x_2^{n+1}$$

dunque, siccome per proprietà dei coefficienti binomiali si ha che $\binom{n+1}{k+1} = \binom{n}{k} + \binom{n}{k+1}$, allora

$$(x_1 + x_2)^{n+1} = x_1^{n+1} + \sum_{k=0}^{n-1} \binom{n+1}{k+1} x_1^{n-k} x_2^{k+1} + x_2^{n+1} = x_1^{n+1} + \sum_{k=1}^n \binom{n+1}{k} x_1^{n+1-k} x_2^k + x_2^{n+1} = \sum_{k=0}^{n+1} \binom{n+1}{k} x_1^{n+1-k} x_2^k$$

Dunque la tesi è stata dimostrata

⊕

Il coefficiente binomiale è comodo per calcolarsi le potenze del 2 siccome se $x_1 = x_2 = 1$ si ha che

$$2^n = \sum_{k=0}^n \binom{n}{k} 1^{n-k} 1^k = \sum_{k=0}^n \binom{n}{k}$$

Il coefficiente binomiale può anche essere generalizzato supponendo di voler dividere un insieme di n elementi in m sottoinsiemi disgiunti, ciascuno con un numero k_i di elementi, la cui unione costituisca l'insieme di partenza, dunque si deve avere che $\sum_{i=1}^m k_i = n$. Il numero di modi con cui si può effettuare tale suddivisione prende il nome di **coefficiente multinomiale** n su k_1, \dots, k_m e si basa sull'idea che abbiamo n su k_1 modi per scegliere il primo sottoinsieme, $n - k_1$ su k_2 modi per scegliere il secondo sottoinsieme e così via, dunque:

$$\binom{n}{k_1, k_2, \dots, k_m} = \binom{n}{k_1} \binom{n-k_1}{k_2} + \dots + \binom{n-k_1-\dots-k_{m-1}}{k_m}$$

Sempre come per il coefficiente binomiale, possiamo caratterizzare il coefficiente multinomiale come lo sviluppo per la potenza n di un numero arbitrario di monomi nella seguente maniera:

$$(x_1 + x_2 + \dots + x_m)^n = \sum_{\{(k_1, k_2, \dots, k_m): \sum_{i=1}^m k_i = n\}} \binom{n}{k_1, \dots, k_m} x_1^{k_1} x_2^{k_2} \dots x_m^{k_m}$$

2.3 Probabilità condizionata

Tutto ciò che è stato mostrato nella sezione precedente può essere comodo per calcolare la probabilità di determinati eventi² e consiglio caldamente di guardare gli esempi proposti da Baldini nel suo libro, su cui non mi soffermerò. Diamo delle definizioni

Definizione 2.9: Probabilità condizionata

Dati due eventi E_1 ed E_2 , con $P(E_2) \neq 0$, definiamo la *probabilità condizionata* $P(E_1|E_2)$ di E_1 dato E_2 (cioè la probabilità che si verifichi E_1 nel caso in cui sappiamo già che si verifica E_2) come

$$P(E_1|E_2) = \frac{P(E_1 \cap E_2)}{P(E_2)} \quad (2.11)$$

Osservazione:-

La probabilità condizionata può essere vista come una sorta di "misura" dell'intersezione $E_1 \cap E_2$ pesata a E_2 . E' possibile inoltre dimostrare che la probabilità condizionata soddisfa tutti gli assiomi della probabilità siccome ogni evento E può essere visto come una probabilità condizionata alla probabilità dello spazio campionario $P(E|\Omega)$

L'utilità della probabilità condizionata è il fatto che possiamo calcolare la probabilità che due eventi E_1 ed E_2 si verifichino contemporaneamente nella seguente maniera:

$$P(E_1 \cap E_2) = P(E_1)P(E_2|E_1) = P(E_2)P(E_1|E_2)$$

A questo punto diamo la definizione di *eventi indipendenti*:

Definizione 2.10: Eventi indipendenti

Si dice che due eventi indipendenti E_1 ed E_2 sono indipendenti se il fatto che sia verificato E_2 non influenza la probabilità che si verifichi E_1 , ovvero se

$$P(E_1|E_2) = P(E_1) \text{ e } P(E_2|E_1) = P(E_2) \quad (2.12)$$

e per la definizione di probabilità condizionata possiamo riscrivere che

$$P(E_1 \cap E_2) = P(E_1)P(E_2) \quad (2.13)$$

²usufruendo della definizione combinatoriale della probabilità

Osservazione:-

Non bisogna confondere il concetto di *eventi indipendenti* con il concetto di *eventi disgiunti* e, dunque, incompatibili: si rimanda all'esempio di Baldini sul libro, in cui si osserva proprio il classico esempio del mazzo da carte in cui viene inserito un jolly. Prima dell'aggiunta del jolly i due eventi $E_1 = \text{"pesco un re"}$ ed $E_2 = \text{"pesco una carta di cuori"}$ erano indipendenti pure avendo un'intersezione non nulla, dopo l'aggiunta del jolly non sono più indipendenti ma non sono nemmeno incompatibili

2.4 Teorema di Bayes

La naturale conseguenza di tutto ciò che abbiamo detto sulla probabilità condizionata è il *teorema di Bayes*, che lega tra loro la probabilità condizionata $P(E_1|E_2)$ con la probabilità $P(E_2|E_1)$

Teorema 2.3 Teorema di Bayes

La probabilità condizionata di E_1 dato E_2 è pari a

$$P(E_1|E_2) = \frac{P(E_2|E_1)P(E_1)}{P(E_2)} \quad (2.14)$$

dove $P(E_2|E_1)$ è la probabilità condizionata di E_2 dato E_1

Dimostrazione: Dalla definizione di probabilità condizionata $P(E_1|E_2)$ di E_1 dato E_2 sappiamo che

$$P(E_1|E_2) = \frac{P(E_1 \cap E_2)}{P(E_2)}$$

con $P(E_2) \neq 0$, ma noi sappiamo pure che $P(E_1 \cap E_2) = P(E_2|E_1)P(E_1)$, dunque si ottiene la tesi:

$$P(E_1|E_2) = \frac{P(E_2|E_1)P(E_1)}{P(E_2)}$$

La dimostrazione è dunque conclusa ⊕

Sfruttando inoltre il fatto che $E_1 \cup \bar{E}_1 = \Omega$ allora $(E_1 \cap E_2) \cup (\bar{E}_1 \cap E_2) = E_2 \cap (E_1 \cup \bar{E}_1) = E_2$, dunque:

$$P(E_2) = P(E_1 \cap E_2) + P(\bar{E}_1 \cap E_2)$$

(si ricorda che $(E_1 \cup E_2) \cup (\bar{E}_1 \cup E_2) = \Omega$ siccome se, per assurdo, $\exists c \in (E_1 \cap E_2) \cap (\bar{E}_1 \cap E_2) \neq \emptyset \implies \{c \in E_1 \cap E_2 \wedge c \in \bar{E}_1 \cap E_2\}$ ma ciò implica che $c \in E_1 \wedge c \in \bar{E}_1$ che è un assurdo; dunque possiamo usare il terzo assioma di Kolmogorov per stimare la probabilità di E_2). Ciò implica che

$$P(E_1|E_2) = \frac{P(E_2|E_1)P(E_1)}{P(E_1 \cap E_2) + P(\bar{E}_1 \cap E_2)} \quad (2.15)$$

Più in generale, possiamo dire che se si dispone di un partizionamento dell'insieme $\{E_i\}$ (tali che $E_i \cap E_j = \emptyset \forall i, j, i \neq j$ e $\bigcup_i E_i = \Omega$) allora possiamo generalizzare, per induzione su i , che $P(E_2) = \sum_i P(E_2|E_i)P(E_i)$, dunque:

$$P(E_1|E_2) = \frac{P(E_2|E_1)P(E_1)}{\sum_i P(E_2|E_i)P(E_i)} \quad (2.16)$$

Questo teorema, sebbene risulti a prima vista banale, in realtà è molto importante siccome lega la probabilità diretta al problema di probabilità inversa. Si rimanda sempre al libro di Baldini per degli esempi per comprendere meglio il teorema di Bayes, sebbene sia utile osservare una cosa: mentre lo statista, oppure il matematico, studia la probabilità in astratto, ovvero cerca di calcolare la probabilità di un determinato processo "contando"; il fisico, di professione, cerca di inferire sul calcolo della probabilità attraverso delle raccolte dati, dunque tramite queste relazioni possiamo eventualmente ricondurci alla probabilità inversa, tuttavia cambia l'approccio "metodologico".

2.5 Variabili casuali e funzioni di distribuzione

Iniziamo questa sezione introducendo il concetto di variabile casuale

Definizione 2.11: Variabile casuale

Una *variabile casuale* o *variabile aleatoria* è una variabile che rappresenta la realizzazione numerica di un processo causale, per cui il suo valore è soggetto a fluttuazioni casuali e non è noto a priori. Si distinguono in

- discrete, cioè variabili che possono assumere un numero finito o numerabile di valori;
- continue, cioè variabili che possono assumere tutti i valori compresi in un intervallo.

Esempio 2.2

- L'uscita del lancio di un dado a sei facce è una variabile casuale discreta che può assumere esattamente sei valori: $1, 2, 3, 4, 5, 6$
- Il tempo necessario per arrivare da casa al luogo di lavoro è un esempio di variabile casuale continua