Capitolo 1

Teoria della probabilità

1.1 Definizione assiomatica della probabilità

1.1.1 Introduzione al concetto di probabilità

Una teoria che formalizzi il concetto di probabilità con particolare riferimento alle sue declinazioni scientifiche può essere sviluppata invocando numerosi sistemi logici. Tuttavia, una descrizione assiomatica coadiuvata dall'apparato deduttivo tipico di ogni disciplina analitica rappresenta, indubbiamente, il procedimento più intuitivo ed esauriente che si possa produrre.
Senza scendere in argomentazioni di carattere epistemologico, che esulano, peraltro, dagli ambiti di questo Corso, occorre preliminarmente individuare il dominio di validità della teoria che si intende derivare. Esso sarà costituito da un modello dei fenomeni affatto comuni al mondo empirico che faccia astrazione di alcune categorie dotate di quelle proprietà enumerative che ne permettano una rappresentazione insiemistica. Attraverso l'introduzione di un opportuno funzionale ogni insieme sarà, quindi, associato ad un intervallo di numeri reali a cui applicare un'operazione di misura probabilistica. Estendendo a questa particolare metrica il classico concetto di funzione, si renderà possibile l'applicazione dei risultati derivati dall'Analisi Matematica.
I concetti primitivi strettamente necessari per affrontare una definizione assiomatica della probabilità sono riportati di seguito.

Definizione 1.1. *Si definisce* esperimento aleatorio *qualsiasi fenomeno empirico il cui esito non sia esattamente predicibile a priori, ma di cui si conoscano i possibili risultati.*

Un esempio di esperimento aleatorio è costituito dal lancio di un dado il cui esito è, appunto, l'uscita di una delle facce del dado.

Definizione 1.2. *Si definisce* spazio dei risultati *l'insieme* $S = \{\zeta_1, \zeta_2, \zeta_3, \dots\}$*, dove* ζ_i *sono i possibili risultati dell'esperimento.*

Definizione 1.3. *Si definisce* evento *un qualunque sottoinsieme di S. Un evento si verifica quando un esperimento aleatorio genera almeno un risultato tra quelli che lo caratterizzano. L'insieme di tutti i possibili eventi associati ad un esperimento aleatorio si indica come* spazio degli eventi S_e. *In particolare, si individuano:*

- *$\emptyset$ (insieme vuoto) è detto evento impossibile (perché non contiene nessun risultato).*
- *S è detto evento certo (perché contiene tutti i risultati).*

Si osservi che se la cardinalità dello spazio dei risultati è uguale ad N, allora la cardinalità dello spazio degli eventi è pari a 2^N.

Definizione 1.4. *Due eventi si dicono* incompatibili *se il verificarsi di uno esclude il verificarsi dell'altro. Da un punto di vista insiemistico, ciò significa che la loro intersezione è l'evento impossibile.*

Avvalendosi di questa terminologia è possibile definire un funzionale:

$$\mathrm{Prob}\big\{\big\} : S \to \mathbb{R}$$

che associa ad ogni evento A la sua probabilità $\mathrm{Prob}\big\{A\big\}$ ottemperando i seguenti assiomi:

Assioma 1.1. *La probabilità è un numero non negativo:*

$$\mathrm{Prob}\big\{A\big\} \geq 0.$$

Assioma 1.2. *L'evento certo S ha probabilità 1:*

$$\mathrm{Prob}\big\{S\big\} = 1.$$

Assioma 1.3. *La probabilità dell'unione di due eventi incompatibili A e B è la somma delle rispettive probabilità:*

$$A \cap B = \emptyset \quad \Longrightarrow \quad \mathrm{Prob}\big\{A \cup B\big\} = \mathrm{Prob}\big\{A\big\} + \mathrm{Prob}\big\{B\big\}.$$

Dai tre assiomi si ricavano alcune proprietà notevoli della probabilità:

1. L'evento impossibile ha probabilità 0:

   $$\mathrm{Prob}\big\{\emptyset\big\} = 0$$

   Dimostrazione.

   $$\mathrm{Prob}\big\{A\big\} = \mathrm{Prob}\big\{A \cup \emptyset\big\} \stackrel{1.3}{=} \mathrm{Prob}\big\{A\big\} + \mathrm{Prob}\big\{\emptyset\big\} \quad \Longrightarrow \quad \mathrm{Prob}\big\{\emptyset\big\} = 0$$

   □

2. La probabilità di un evento è un numero non superiore ad 1:

$$\text{Prob}\{A\} \leq 1$$

Dimostrazione.

$$\begin{aligned}\text{Prob}\{S\} = \text{Prob}\{A \cup \overline{A}\} &\stackrel{1.3}{=} \text{Prob}\{A\} + \text{Prob}\{\overline{A}\} \\ \stackrel{1.2}{\Longrightarrow} \quad \text{Prob}\{A\} &= 1 - \text{Prob}\{\overline{A}\} \stackrel{1.1}{\leq} 1\end{aligned}$$

□

3. Dati due eventi qualunque A e B, la probabilità della loro unione è:

$$\text{Prob}\{A \cup B\} = \text{Prob}\{A\} + \text{Prob}\{B\} - \text{Prob}\{A \cap B\}$$

Dimostrazione.

$$\begin{aligned}\text{Prob}\{A \cup B\} &= \text{Prob}\{A \cup \{B \setminus \{A \cap B\}\}\} \\ &\stackrel{1.3}{=} \text{Prob}\{A\} + \text{Prob}\{B\} - \text{Prob}\{A \cap B\}\end{aligned}$$

□

La relazione precedente può essere generalizzata a quella che viene spesso indicata come *proprietà di subadditività numerabile dello spazio degli eventi rispetto alla misura probabilistica* che si riassume sinteticamente come segue:

$$\text{Prob}\{\bigcup_j A_j\} \leq \sum_j \text{Prob}\{Aj\}$$

4. Dati due eventi A e B, se A è contenuto in B allora la probabilità di A è non superiore a quella di B:

$$A \subseteq B \quad \Longrightarrow \quad \text{Prob}\{A\} \leq \text{Prob}\{B\}$$

Dimostrazione.

$$\begin{aligned}\text{Prob}\{B\} &= \text{Prob}\{A \cup \{B \setminus A\}\} \\ &\stackrel{1.3}{=} \text{Prob}\{A\} + \text{Prob}\{B \setminus A\} \\ &\stackrel{1.1}{\Longrightarrow} \quad \text{Prob}\{A\} \leq \text{Prob}\{B\}\end{aligned}$$

□

1.1.2 Eventi condizionati

Definizione 1.5. *Dati due eventi A ed M con $M \neq \emptyset$, si definisce* probabilità condizionata *di A rispetto ad M, la probabilità che si verifichi l'evento A qualora si sia verificato anche l'evento M. Tale probabilità si indica con:*

$$\mathrm{Prob}\{A \mid M\} \doteq \frac{\mathrm{Prob}\{A \cap M\}}{\mathrm{Prob}\{M\}} \tag{1.1}$$

Alcune proprietà delle probabilità condizionate sono:

1.
$$A \subseteq M \quad \Longrightarrow \quad \mathrm{Prob}\{A \mid M\} \geq \mathrm{Prob}\{A\}$$

Dimostrazione.

$$\mathrm{Prob}\{A \cap M\} = \mathrm{Prob}\{A\} \quad \Longrightarrow \quad \mathrm{Prob}\{A \mid M\} = \frac{\mathrm{Prob}\{A\}}{\mathrm{Prob}\{M\}} \geq \mathrm{Prob}\{A\}$$

□

2.
$$M \subseteq A \quad \Longrightarrow \quad \mathrm{Prob}\{A \mid M\} = 1$$

Dimostrazione.

$$\mathrm{Prob}\{A \cap M\} = \mathrm{Prob}\{M\}$$

□

Definizione 1.6. *Un insieme di eventi $\{A_i\}$ costituiscono una* partizione *dello spazio dei risultati S se:*

- $A_i \cap A_j = \emptyset, \ \forall i, j : i \neq j$
- $\bigcup_i A_i = S$

Introdotti questi concetti, si possono dedurre i seguenti risultati:

Teorema 1.1 (della probabilità totale)**.** *Se $\{A_i\}$ costituiscono una partizione dello spazio dei risultati S, e B è un evento arbitrario, allora la probabilità di quest'ultimo può essere espressa come:*

$$\mathrm{Prob}\{B\} = \sum_i \mathrm{Prob}\{B \mid A_i\} \, \mathrm{Prob}\{A_i\}$$

Dimostrazione.

$$B = \bigcup_i \{B \cap A_i\} \quad \Longrightarrow \quad \text{Prob}\big\{B\big\} \overset{1.3}{=} \sum_i \text{Prob}\big\{B \cap A_i\big\}$$
$$\doteq \sum_i \text{Prob}\big\{B \,\big|\, A_i\big\}\,\text{Prob}\big\{A_i\big\}$$

□

Si osservi che una partizione qualunque dello spazio dei risultati S forma una *base* per sviluppare in serie un generico evento.

Teorema 1.2 (di Bayes)**.** *Se $\{A_i\}$ formano una partizione dello spazio dei risultati S, e B è un evento arbitrario, allora le probabilità degli eventi A_i condizionate all'evento B possono essere espresse, alla luce del risultato precedente, come:*

$$\text{Prob}\big\{A_i \,\big|\, B\big\} = \text{Prob}\big\{B \,\big|\, A_i\big\}\frac{\text{Prob}\big\{A_i\big\}}{\text{Prob}\big\{B\big\}} = \frac{\text{Prob}\big\{B \,\big|\, A_i\big\}\,\text{Prob}\big\{A_i\big\}}{\sum_i \text{Prob}\big\{B \,\big|\, A_i\big\}\,\text{Prob}\big\{A_i\big\}}$$

1.1.3 Eventi indipendenti

Definizione 1.7. *Due eventi A e B si dicono* indipendenti *se:*

$$\text{Prob}\big\{A \cap B\big\} = \text{Prob}\big\{A\big\}\,\text{Prob}\big\{B\big\}. \tag{1.2}$$

Intuitivamente ciò equivale ad affermare che il verificarsi dell'uno non è legato in alcun modo al verificarsi dell'altro.

Alcune proprietà significative della probabilità associata a due eventi indipendenti sono:

1. Se A e B sono indipendenti allora:

   $$\text{Prob}\big\{A \,\big|\, B\big\} = \text{Prob}\big\{A\big\} \quad \text{e} \quad \text{Prob}\big\{B \,\big|\, A\big\} = \text{Prob}\big\{B\big\}$$

   Il precedente risultato costituisce una giustificazione più rigorosa della proprietà di indipendenza tra due eventi enunciata dalla (1.2).

2. Se A e B sono indipendenti allora lo sono anche gli eventi $\overline{A}$ e B, gli eventi A e $\overline{B}$, e gli eventi $\overline{A}$ e $\overline{B}$.

Generalizzando il concetto, si può affermare che tre eventi A_1, A_2, A_3 risultano tra loro indipendenti se sono indipendenti a due a due e vale anche che:

$$\text{Prob}\big\{A_1 \cap A_2 \cap A_3\big\} = \text{Prob}\big\{A_1\big\}\,\text{Prob}\big\{A_2\big\}\,\text{Prob}\big\{A_3\big\}. \tag{1.3}$$

Procedendo, quindi, per induzione, n eventi $A_1, A_2, \ldots, A_n$ si dicono indipendenti se sono indipendenti a gruppi di $n-1$ e risulta che:

$$\text{Prob}\Big\{\bigcap_{i=1}^{n} A_i\Big\} = \prod_{i=1}^{n} \text{Prob}\big\{A_i\big\} \tag{1.4}$$

1.2 Variabili aleatorie

Definizione 1.8. *Sia dato un esperimento aleatorio con spazio dei risultati S, si definisce* variabile aleatoria reale *una funzione:*

$$\mathbf{x} : S \to \mathbb{R}$$

che associa ad ogni risultato $\zeta \in S$ il numero reale $\mathbf{x}(\zeta)$ e soddisfa le seguenti due condizioni:

1. *l'insieme $\{\mathbf{x} \leq x\}$ è un evento $\forall x \in \mathbb{R}$;*
2. *le probabilità degli eventi $\{\mathbf{x} = \infty\}$ e $\{\mathbf{x} = -\infty\}$ sono nulle:*

$$\text{Prob}\big\{\mathbf{x} = \infty\big\} = \text{Prob}\big\{\mathbf{x} = -\infty\big\} = 0.$$

Si osservi come, attraverso l'introduzione del concetto di variabile aleatoria, si passi da una rappresentazione insiemistica della probabilità ad una descrizione eminentemente analitica con tutti i vantaggi del caso. Ricorrendo, infatti ai risultati matematici derivati in altri ambiti scientifici, è possibile caratterizzare completamente l'esperimento a cui essa è associata.
In generale, una variabile aleatoria $\mathbf{z}$ può essere anche a valori complessi e, in questo caso, è costituita da due variabili aleatorie reali $\mathbf{x}$ e $\mathbf{y}$ tali che:

$$\mathbf{z} = \mathbf{x} + j\mathbf{y}.$$

1.2.1 Funzione di distribuzione di probabilità

Definizione 1.9. *Data una variabile aleatoria $\mathbf{x}$ si definisce* funzione di distribuzione (cumulative distribution function, CDF) *la seguente funzione:*

$$F_{\mathbf{x}}(x) \doteq \text{Prob}\big\{\mathbf{x} \leq x\big\},\ \forall x \in (-\infty, +\infty)\,. \tag{1.5}$$

La funzione di distribuzione gode delle seguenti proprietà:

1. La funzione di distribuzione assume i seguenti valori agli estremi del dominio:

$$F_{\mathbf{x}}(-\infty) = 0 \qquad F_{\mathbf{x}}(\infty) = 1.$$

2. La funzione di distribuzione è monotona non decrescente:

$$\forall x_1, x_2 \in \mathbb{R} : x_1 \leq x_2 \quad \Longrightarrow \quad F_{\mathbf{x}}(x_1) \leq F_{\mathbf{x}}(x_2).$$

3. Se la funzione di distribuzione si annulla in un punto, allora è nulla anche in tutti i punti alla sua sinistra:
$$\textit{se } \exists x_0 : F_{\mathbf{x}}(x_0) = 0 \quad \Longrightarrow \quad F_{\mathbf{x}}(x) = 0 \quad \forall x \leq x_0.$$
4. La probabilità che $\mathbf{x}$ assuma valori a destra di un punto è esprimibile come:
$$\text{Prob}\big\{\mathbf{x} > x\big\} = 1 - F_{\mathbf{x}}(x)$$
5. La funzione di distribuzione è continua a destra:
$$F_{\mathbf{x}}(x^+) = F_{\mathbf{x}}(x) \quad \forall x \in \mathbb{R}.$$
6. La probabilità che $\mathbf{x}$ assuma valori in un intervallo finito aperto a sinistra è esprimibile come:
$$\text{Prob}\big\{x_1 < \mathbf{x} \leq x_2\big\} = F_{\mathbf{x}}(x_2) - F_{\mathbf{x}}(x_1).$$
7. La probabilità che $\mathbf{x}$ assuma un preciso valore puntuale è esprimibile come:
$$\text{Prob}\big\{\mathbf{x} = x\big\} = F_{\mathbf{x}}(x) - F_{\mathbf{x}}(x^-).$$
8. La probabilità che $\mathbf{x}$ assuma valori in un intervallo chiuso è esprimibile come:
$$\text{Prob}\big\{x_1 \leq \mathbf{x} \leq x_2\big\} = F_{\mathbf{x}}(x_2) - F_{\mathbf{x}}(x_1^-).$$

1.2.2 Variabili aleatorie continue, discrete e miste

Una variabile aleatoria si dice *discreta*, *continua* o *mista* in base alle caratteristiche della sua funzione di distribuzione. In particolare la variabile aleatoria $\mathbf{x}$ si dice:

Discreta se la sua funzione di distribuzione è costante a tratti. In questo caso la variabile aleatoria $\mathbf{x}$ può assumere al più un'infinità numerabile di valori $\{x_i,\ i = 1, 2, \dots\}$ che costituiscono, appunto, i punti di discontinuità con probabilità date da:
$$\text{Prob}\big\{\mathbf{x} = x_i\big\} = F_{\mathbf{x}}(x_i) - F_{\mathbf{x}}(x_i^-).$$

Continua se la sua funzione di distribuzione è continua, ovvero:
$$F_{\mathbf{x}}(x^-) = F_{\mathbf{x}}(x^+) = F_{\mathbf{x}}(x) \quad \forall x.$$
In questo caso la variabile aleatoria può assumere una infinità non numerabile di valori ognuno dei quali ha probabilità nulla.

Mista se la funzione di distribuzione non è né completamente continua né costante a tratti. In questo caso la variabile aleatoria può assumere una infinità non numerabile di valori ognuno dei quali con probabilità nulla, ed un insieme numerabile di valori $\{x_i,\ i = 1, 2, \dots\}$ che costituiscono i punti di discontinuità con probabilità date da:
$$\text{Prob}\big\{\mathbf{x} = x_i\big\} = F_{\mathbf{x}}(x_i) - F_{\mathbf{x}}(x_i^-)$$

1.2.3 Funzione di densità di probabilità

Definizione 1.10. *Data una variabile aleatoria* $\mathbf{x}$ *con funzione di distribuzione* $F_{\mathbf{x}}(x)$*, si definisce* densità di probabilità (probability density function, pdf)*la sua derivata prima:*

$$f_{\mathbf{x}}(x) \doteq \frac{d}{dx} F_{\mathbf{x}}(x) \tag{1.6}$$

Questa definizione è valida per qualunque tipo di variabile aleatoria. Tuttavia, nel caso di variabili aleatorie continue la densità di probabilità è una funzione ordinaria continua a tratti, per variabili discrete essa assume la forma seguente:

$$f_{\mathbf{x}}(x) = \sum_i \delta(x - x_i) \,\mathrm{Prob}\{\mathbf{x} = x_i\} \tag{1.7}$$

mentre per variabili aleatorie miste è costituita da un termine ordinario e da una sommatoria di funzioni impulsive (Dirac comb).
Le proprietà più importanti della funzione densità di probabilità sono le seguenti:

1. La densità di probabilità è una funzione non negativa:

   $$f_{\mathbf{x}}(x) \geq 0$$

   Dimostrazione.

   $F_{\mathbf{x}}(x)$ *è infatti non decrescente.*

   □

2. La funzione di distribuzione è legata alla densità di probabilità dalla relazione:

   $$F_{\mathbf{x}}(x) = \int_{-\infty}^{x} f_{\mathbf{x}}(\alpha) d\alpha$$

   Dimostrazione.

   E' la tesi del Teorema Fondamentale del Calcolo Integrale.

   □

3. La densità di probabilità ha area unitaria:

   $$\int_{-\infty}^{+\infty} f_{\mathbf{x}}(x) dx = 1$$

Dimostrazione.

$$F_{\mathbf{x}}(+\infty) \doteq 1.$$

□

4. La densità di probabilità si annulla a $+\infty$ e $-\infty$:

$$f_{\mathbf{x}}(-\infty) = f_{\mathbf{x}}(+\infty) = 0$$

Dimostrazione.

Si tratta di una condizione necessaria per l'integrabilità di $f_{\mathbf{x}}(x)$.

□

5. La probabilità che la variabile aleatoria $\mathbf{x}$ stia in un intervallo è data da:

$$\text{Prob}\{x_1 < \mathbf{x} \leq x_2\} = \int_{x_1}^{x_2} f_{\mathbf{x}}(x)dx$$

Dimostrazione.

L'operatore integrale smussa le eventuali discontinuità di prima specie di $f_{\mathbf{x}}(x)$.

□

Sebbene la funzione densità di probabilità così definita abbia validità generale, normalmente nel caso di variabili aleatorie discrete si preferisce utilizzare direttamente la funzione di probabilità:

$$\text{Prob}\{\mathbf{x} = x_i\}$$

per cui si può affermare che:

1. La funzione di distribuzione è legata alla densità di probabilità dalla relazione:

$$F_{\mathbf{x}}(x) = \sum_{i:\, x_i \leq x} \text{Prob}\{\mathbf{x} = x_i\}$$

2. La probabilità che la variabile aleatoria $\mathbf{x}$ stia in un intervallo è data da:

$$\text{Prob}\{a < \mathbf{x} \leq b\} = \sum_{i:\, x_i \in (a,b]} \text{Prob}\{\mathbf{x} = x_i\}$$

1.2.4 Funzione di variabili aleatorie

Data la variabile aleatoria $\mathbf{x}$ con pdf $f_{\mathbf{x}}(x)$ e CDF $F_{\mathbf{x}}(x)$ e la funzione reale $g(x)$, la grandezza:

$$\mathbf{y} = g(\mathbf{x})$$

è una nuova variabile aleatoria caratterizzata da una funzione di distribuzione in generale diversa da quella originale.
L'espressione anlaitica della funzione di distribuzione della variabile aleatoria $\mathbf{y}$ può essere ricavata dalla seguente considerazione:

$$\begin{aligned} F_{\mathbf{y}}(y) &= \text{Prob}\big\{\mathbf{y} \leq y\big\} = \text{Prob}\big\{g(\mathbf{x}) \leq y\big\} = \\ &= \text{Prob}\big\{\mathbf{x} \in A_y\big\} = \int_{A_y} f_{\mathbf{x}}(x)dx \end{aligned} \tag{1.8}$$

dove l'insieme A_y è:

$$A_y = \{x : g(x) \leq y\} \tag{1.9}$$

La densità di probabilità si può ricavare derivando l'espressione precedente oppure ricorrendo alla seguente relazione:

$$f_{\mathbf{y}}(y) = \sum_i \frac{f_{\mathbf{x}}(x_i)}{|g'(x_i)|} \tag{1.10}$$

dove gli x_i sono le soluzioni dell'equazione $y = g(x)$.

Esempio 1.1. *Data le variabile aleatoria* $\mathbf{x}$ *con densità di probabilità* $f_{\mathbf{x}}(x)$ *calcolare la densità di probabilità della variabile aleatoria:*

$$\mathbf{y} = \mathbf{x}^2$$

Soluzione. La funzione di distribuzione può essere ottenuta considerando due casi:

$y < 0$**:** La variabile aleatoria $\mathbf{x}^2$ non può essere negativa quindi:

$$F_{\mathbf{y}}(y) = \text{Prob}\big\{\mathbf{x}^2 \leq y\big\} = 0$$

$y \geq 0$**:** L'insieme A_y è $A_y = \{x : -\sqrt{y} \leq x \leq \sqrt{y}\}$ quindi:

$$\begin{aligned} F_{\mathbf{y}}(y) &= \text{Prob}\big\{\mathbf{x}^2 \leq y\big\} = \text{Prob}\big\{-\sqrt{y} \leq \mathbf{x} \leq \sqrt{y}\big\} = \\ &= \int_{-\sqrt{y}}^{\sqrt{y}} f_{\mathbf{x}}(x)dx = F_{\mathbf{x}}(\sqrt{y}) - F_{\mathbf{x}}(-\sqrt{y}) \end{aligned}$$

La densità di probabilità si ottiene come derivata della precedente oppure può essere calcolata direttamente dalla (1.10) considerando che le soluzioni della equazione $y = x^2$ sono $\{x_1 = -\sqrt{y}, x_2 = \sqrt{y}\}$ quando $x \geq 0$, quindi:

$$f_{\mathbf{y}}(y) = \begin{cases} 0 & y < 0 \\ \frac{f_{\mathbf{x}}(-\sqrt{y})}{2\sqrt{y}} + \frac{f_{\mathbf{x}}(\sqrt{y})}{2\sqrt{y}} & y \geq 0 \end{cases}$$

□

1.2.5 Momenti statistici di una variabile aleatoria

Definizione 1.11. *Data la variabile aleatoria* $\mathbf{x}$*, si definisce* valore atteso *o* media statistica *di tale v.a.:*

$$\mathrm{E}\{\mathbf{x}\} \doteq \int_{-\infty}^{+\infty} x f_{\mathbf{x}}(x) dx. \tag{1.11}$$

Nel caso particolare di una v.a. discreta, contraddistinta da una pdf impulsiva, può anche essere scritto come:

$$\mathrm{E}\{\mathbf{x}\} = \sum_i x_i \,\mathrm{Prob}\big\{\mathbf{x} = x_i\big\}. \tag{1.12}$$

Il valore medio $\mathrm{E}\{\mathbf{x}\}$ è il valore attorno al quale si distribuisce la variabile aleatoria $\mathbf{x}$. In proposito la legge dei grandi numeri afferma che se un esperimento aleatorio viene ripetuto n volte in modo indipendente la media aritmetica degli n valori della variabile aleatoria tende al valore atteso al crescere di n.
Una proprietà interessante del valore atteso stabilisce che, il valore atteso della variabile aleatoria $\mathbf{y} = g(\mathbf{x})$, dato da:

$$\mathrm{E}\{\mathbf{y}\} = \int_{-\infty}^{\infty} y f_{\mathbf{y}}(y) dy$$

può essere calcolato senza conoscere la $f_{\mathbf{y}}(y)$ come:

$$\mathrm{E}\{g(\mathbf{x})\} = \int_{-\infty}^{\infty} g(x) f_{\mathbf{x}}(x) dx. \tag{1.13}$$

Questa proprietà consente di calcolare immediatamente i parametri di una variabile aleatoria che stanno per essere introdotti.

Definizione 1.12. *Si definisce* varianza *di una variabile aleatoria* $\mathbf{x}$*, la quantità:*

$$\mathrm{Var}\{\mathbf{x}\} \doteq \mathrm{E}\left\{(\mathbf{x} - \mathrm{E}\{\mathbf{x}\})^2\right\} = \int_{-\infty}^{\infty} (x - \mathrm{E}\{\mathbf{x}\})^2 f_{\mathbf{x}}(x) dx. \tag{1.14}$$

Tale parametro dà una misura di quanto la variabile aleatoria si disperde attorno al suo valore medio.

Definizione 1.13. *Si definisce* momento di ordine k *di una variabile aleatoria la grandezza:*

$$\mathrm{E}\left\{\mathbf{x}^k\right\} \doteq \int_{-\infty}^{\infty} x^k f_{\mathbf{x}}(x)dx. \tag{1.15}$$

La conoscenza dei valori di tutti i momenti di ordine k, per $k = 1, 2, 3, \ldots$ è sufficiente per caratterizzare completamente una variabile aleatoria, prescindendo dalla conoscenza di $f_{\mathbf{x}}(x)$ e $F_{\mathbf{x}}(x)$.
Rivestono una particolare importanza il momento di ordine 1, che coincide con il valore atteso, ed il momento di ordine 2, detto *valore quadratico medio*, che è legato alla varianza dalla seguente relazione:

$$\mathrm{Var}\left\{\mathbf{x}\right\} = \mathrm{E}\left\{\mathbf{x}^2\right\} - \left(\mathrm{E}\left\{\mathbf{x}\right\}\right)^2 \tag{1.16}$$

considerando, infatti, che:

$$\mathrm{Var}\left\{\mathbf{x}\right\} dot{=} \mathrm{E}\left\{(\mathbf{x} - \mathrm{E}\left\{\mathbf{x}\right\})^2\right\} = \mathrm{E}\left\{(\mathbf{x})^2 - 2\mathbf{x}\,\mathrm{E}\left\{\mathbf{x}\right\} + (\mathrm{E}\left\{\mathbf{x}\right\})^2\right\} = \mathrm{E}\left\{\mathbf{x}^2\right\} - (\mathrm{E}\left\{\mathbf{x}\right\})^2 \tag{1.17}$$

1.2.6 Funzioni di distribuzione e densità di probabilità congiunte

In generale un esperimento aleatorio può coinvolgere più di una grandezza osservabile, ovvero più di una variabile aleatoria. Pertanto non può essere caratterizzato eaurientemente soltanto dalle funzioni di distribuzione delle singole variabili (dette *distribuzioni marginali*), perché occorre tenere conto delle possibili interrelazioni che possono essere descritte introducendo il concetto di *funzione di distribuzione congiunta.*

Definizione 1.14. *Data una sequenza di n variabili aleatorie $\mathbf{x}_1, \ldots, \mathbf{x}_n$ si definisce* funzione di distribuzione congiunta *la probabilità:*

$$F_{\mathbf{x}_1,\ldots,\mathbf{x}_n}(x_1, \ldots, x_n) \doteq \mathrm{Prob}\left\{\mathbf{x}_1 \leq x_1, \ldots, \mathbf{x}_n \leq x_n\right\} \tag{1.18}$$

e funzione di densità di probabilità congiunta *la sua derivata rispetto a tutte le variabili:*

$$f_{\mathbf{x}_1,\ldots,\mathbf{x}_n}(x_1, \ldots, x_n) \doteq \frac{\partial^n}{\partial x_1 \ldots \partial x_n} F_{\mathbf{x}_1,\ldots,\mathbf{x}_n}(x_1, \ldots, x_n) \tag{1.19}$$

Alcune proprietà di queste funzioni sono:

1. La funzione di distribuzione si ricava dalla densità di probabilità dalla relazione:

$$F_{\mathbf{x}_1,\ldots,\mathbf{x}_n}(x_1, \ldots, x_n) = \int_{-\infty}^{x_1} \cdots \int_{-\infty}^{x_n} f_{\mathbf{x}_1,\ldots,\mathbf{x}_n}(t_1, \ldots, t_n) dt_1 \ldots dt_n \tag{1.20}$$

   che si può dimostrare integrando la pdf congiunta sull'ipercubo ottenuto dal prodotto cartesiano dei rispettivi domini ed applicando il teorema di Fubini.

2. Le distribuzioni e densità di probabilità marginali del sottoinsieme di variabili $\mathbf{x}_1, \ldots, \mathbf{x}_k$ con $k < n$ si ricavano da:

$$F_{\mathbf{x}_1,\ldots,\mathbf{x}_k}(x_1, \ldots, x_k) = F_{\mathbf{x}_1,\ldots,\mathbf{x}_n}(x_1, \ldots, x_k, +\infty, \ldots, +\infty) \tag{1.21}$$

   e:

$$f_{\mathbf{x}_1,\ldots,\mathbf{x}_k}(x_1, \ldots, x_k) = \int_{-\infty}^{+\infty} \cdots \int_{-\infty}^{+\infty} f_{\mathbf{x}_1,\ldots,\mathbf{x}_n}(x_1, \ldots, x_k, t_{k+1}, \ldots, t_n) dt_{k+1} \ldots dt_n \tag{1.22}$$

1.2.7 Variabili aleatorie statisticamente indipendenti

Premesse queste definizioni, è possibile esprimere il concetto di indipendenza statistica tra n eventi mediante le rispettive variabili aleatorie associate. In particolare, si dice che n variabili aleatorie sono indipendenti se la loro funzione di distribuzione (o densità di probabilità) congiunta è data da prodotto delle singole funzioni di distribuzione (o densità di probabilità) marginali, ovvero:

$$F_{\mathbf{x}_1,\ldots,\mathbf{x}_n}(x_1, \ldots, x_n) = \prod_{i=1}^{n} F_{\mathbf{x}_i}(x_i) \tag{1.23}$$

$$f_{\mathbf{x}_1,\ldots,\mathbf{x}_n}(x_1, \ldots, x_n) = \prod_{i=1}^{n} f_{\mathbf{x}_i}(x_i) \tag{1.24}$$

1.2.8 Variabili aleatorie condizionate

Definizione 1.15. *Data la variabile aleatoria* $\mathbf{x}$ *e l'evento generico* M*, si definisce* funzione di distribuzione condizionata *di* $\mathbf{x}$ *rispetto ad* M*, la funzione:*

$$F_{\mathbf{x}}(x \,|\, M) \doteq \mathrm{Prob}\big\{\mathbf{x} \leq x \,\big|\, M\big\} = \frac{\mathrm{Prob}\big\{\mathbf{x} \leq x, M\big\}}{\mathrm{Prob}\big\{M\big\}} \tag{1.25}$$

mentre la densità di probabilità è:

$$f_{\mathbf{x}}(x \,|\, M) \doteq \frac{d}{dx} F_{\mathbf{x}}(x \,|\, M) \tag{1.26}$$

Ovviamente l'evento M può essere espresso mediante una variabile aleatoria che può essere anche la $\mathbf{x}$ stessa. Quindi, se ad esempio $M = \{\mathbf{y} \leq y\}$, si ottiene:

$$F_{\mathbf{x}}(x \,|\, \mathbf{y} \leq y) = \mathrm{Prob}\big\{\mathbf{x} \leq x \,\big|\, \mathbf{y} \leq y\big\} = \frac{F_{\mathbf{x},\mathbf{y}}(x, y)}{F_{\mathbf{y}}(y)} \tag{1.27}$$

Un caso particolare è quello in cui $M = \{\mathbf{y} = y\}$ in questo caso, infatti, se la variabile aleatoria $\mathbf{y}$ è continua si ha che $\text{Prob}\big\{\mathbf{y} = y\big\} = 0$ ed occorre utilizzare un passaggio al limite per rendere consistente la definizione.
Si definisce anzitutto:

$$M = \{y \leq \mathbf{y} \leq y + \Delta y\}$$

e si ottiene la seguente funzione di distribuzione:

$$\begin{aligned} F_{\mathbf{x}}(x \,|\, y \leq \mathbf{y} \leq y + \Delta y) &= \frac{\text{Prob}\big\{\mathbf{x} \leq x, y \leq \mathbf{y} \leq y + \Delta y\big\}}{\text{Prob}\big\{y \leq \mathbf{y} \leq y + \Delta y\big\}} = \\ &= \frac{\displaystyle\int_{y}^{y+\Delta y} \int_{-\infty}^{x} f_{\mathbf{xy}}(\alpha, \beta) d\alpha \, d\beta}{F_{\mathbf{y}}(y + \Delta y) - F_{\mathbf{y}}(y)}. \end{aligned} \tag{1.28}$$

Derivando rispetto a x si ottiene:

$$f_{\mathbf{x}}(x \,|\, y \leq \mathbf{y} \leq y + \Delta y) = \frac{\displaystyle\int_{y}^{y+\Delta y} f_{\mathbf{xy}}(x, \beta) d\beta}{F_{\mathbf{y}}(y + \Delta y) - F_{\mathbf{y}}(y)} \tag{1.29}$$

calcolando il limite per $\Delta y \to 0$ si ottiene:

$$\begin{aligned} f_{\mathbf{x}}(x \,|\, \mathbf{y} = y) &= \lim_{\Delta y \to 0} \frac{\displaystyle\int_{y}^{y+\Delta y} f_{\mathbf{xy}}(x, \beta) d\beta}{F_{\mathbf{y}}(y + \Delta y) - F_{\mathbf{y}}(y)} = \\ &= \lim_{\Delta y \to 0} \frac{\Delta y}{F_{\mathbf{y}}(y + \Delta y) - F_{\mathbf{y}}(y)} \cdot \frac{\displaystyle\int_{0}^{y+\Delta y} f_{\mathbf{xy}}(x, \beta) d\beta - \int_{0}^{y} f_{\mathbf{xy}}(x, \beta) d\beta}{\Delta y} = \\ &= \frac{f_{\mathbf{xy}}(x, y)}{f_{\mathbf{y}}(y)} \end{aligned} \tag{1.30}$$

Nel caso che non si introducano ambiguità, la precedente distribuzione condizionata può essere scritta più semplicemente come $f_{\mathbf{x}}(x \,|\, y)$. Ed in questo caso si hanno:

$$f_{\mathbf{x}}(x \,|\, y) = \frac{f_{\mathbf{xy}}(x, y)}{f_{\mathbf{y}}(y)} \qquad f_{\mathbf{y}}(y \,|\, x) = \frac{f_{\mathbf{xy}}(x, y)}{f_{\mathbf{x}}(x)} \tag{1.31}$$

Anche per le distribuzioni condizionate valgono il teorema della probabilità totale:

$$f_{\mathbf{x}}(x) = \int_{-\infty}^{+\infty} f_{\mathbf{x}}(x \,|\, y) f_{\mathbf{y}}(y) dy \tag{1.32}$$

ed il teorema di Bayes:

$$f_{\mathbf{x}}(x \,|\, \mathbf{y} = y) = \frac{f_{\mathbf{y}}(y \,|\, x) f_{\mathbf{x}}(x)}{f_{\mathbf{y}}(y)} = \frac{f_{\mathbf{y}}(y \,|\, x) f_{\mathbf{x}}(x)}{\displaystyle\int_{-\infty}^{+\infty} f_{\mathbf{x}}(x \,|\, y) f_{\mathbf{y}}(y) dy}. \tag{1.33}$$

1.2.9 Valore atteso di variabili aleatorie condizionate

Per essere più generali possibile, si consideri una variabile aleatoria $\mathbf{x}$, un evento M ed una funzione $g(x)$. Applicando quanto detto in precedenza, si ottiene che il valore atteso di $\mathbf{y} = g(\mathbf{x})$ condizionato al verificarsi dell'evento M è:

$$\mathrm{E}\left\{g(\mathbf{x}) \,\middle|\, M\right\} = \int_{-\infty}^{+\infty} g(x) f_{\mathbf{x}}(x \,|\, M) dx \tag{1.34}$$

e come casi particolari si hanno il valore medio:

$$\mathrm{E}\left\{x \,\middle|\, M\right\} = \int_{-\infty}^{+\infty} x f_{\mathbf{x}}(x \,|\, M) dx \tag{1.35}$$

e la varianza:

$$\mathrm{Var}\left\{x \,\middle|\, M\right\} = \int_{-\infty}^{+\infty} (x - \mathrm{E}\left\{x \,\middle|\, M\right\})^2 f_{\mathbf{x}}(x \,|\, M) dx \tag{1.36}$$

Ancora una volta un caso di particolare interesse è quello in cui $M = \{\mathbf{y} = y\}$ infatti, in questo caso si ottiene:

$$\mathrm{E}\left\{\mathbf{x} \,\middle|\, y\right\} = \int_{-\infty}^{+\infty} x f_{\mathbf{x}}(x \,|\, y) dx \tag{1.37}$$

Al variare di y, questa è una funzione:

$$\varphi(y) = \mathrm{E}\left\{\mathbf{x} \,\middle|\, y\right\} \tag{1.38}$$

per cui se applicata alla variabile aleatoria $\mathbf{y}$ fornisce una nuova variabile aleatoria $\varphi(\mathbf{y})$ il cui valore medio è il valore medio (non condizionato) di $\mathbf{x}$, infatti:

$$\begin{aligned}
\mathrm{E}\left\{\mathrm{E}\left\{\mathbf{x} \,\middle|\, \mathbf{y}\right\}\right\} = \mathrm{E}\left\{\varphi(\mathbf{y})\right\} &= \int_{-\infty}^{+\infty} \varphi(y) f_{\mathbf{y}}(y) dy = \\
&= \int_{-\infty}^{+\infty} f_{\mathbf{y}}(y) \int_{-\infty}^{+\infty} x f_{\mathbf{x}}(x \,|\, y) dx dy = \\
&= \int_{-\infty}^{+\infty} x \int_{-\infty}^{+\infty} f_{\mathbf{x}}(x \,|\, y) f_{\mathbf{y}}(y) dy dx = \\
&= \int_{-\infty}^{+\infty} x f_{\mathbf{x}}(x) dx = \\
&= \mathrm{E}\left\{\mathbf{x}\right\}
\end{aligned} \tag{1.39}$$

Questo concetto può essere generalizzato al caso in cui si abbia una funzione di due variabili aleatorie $g(\mathbf{x}, \mathbf{y})$ ottenendo:

$$\begin{aligned}
\mathrm{E}\left\{\mathrm{E}\left\{g(\mathbf{x},\mathbf{y}) \,\middle|\, \mathbf{y}\right\}\right\} &= \int_{-\infty}^{+\infty} f_{\mathbf{y}}(y) \int_{-\infty}^{+\infty} g(x,y) f_{\mathbf{x}}(x \,|\, y) dx dy = \\
&= \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} g(x,y) f_{\mathbf{xy}}(x,y) dx dy = \\
&= \mathrm{E}\left\{g(\mathbf{x},\mathbf{y})\right\}
\end{aligned} \tag{1.40}$$

Esempio 1.2. *Date le variabili aleatorie $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n$ indipendenti e identicamente distribuite (i.i.d.) e la variabile aleatoria $\mathbf{n}$ a valori interi indipendente dalle $\mathbf{x}_i$, ma con distribuzione arbitraria, si calcoli il valore atteso della nuova variabile aleatoria:*

$$\mathbf{y} = \sum_{i=1}^{\mathbf{n}} \mathbf{x}_i.$$

Soluzione. Il calcolo di valore medio e varianza di $\mathbf{y}$ si può effettuare condizionandosi al valore di $\mathbf{n}$:

$$\mathrm{E}\left\{\mathbf{y} \,\middle|\, n\right\} = \mathrm{E}\left\{\textstyle\sum_{i=1}^{\mathbf{n}} \mathbf{x}_i \,\middle|\, n\right\} = \sum_{i=1}^{n} \mathrm{E}\left\{\mathbf{x}_i\right\} = n\,\mathrm{E}\left\{\mathbf{x}_i\right\} \tag{1.41}$$

quindi:

$$\begin{aligned}
\mathrm{E}\left\{\mathbf{y}\right\} &= \mathrm{E}\left\{\mathrm{E}\left\{\mathbf{y} \,\middle|\, n\right\}\right\} = \sum_{n=0}^{\infty} \mathrm{E}\left\{\mathbf{y} \,\middle|\, n\right\} \mathrm{Prob}\left\{\mathbf{n} = n\right\} = \\
&= \sum_{n=0}^{\infty} n\,\mathrm{E}\left\{\mathbf{x}_i\right\} \mathrm{Prob}\left\{\mathbf{n} = n\right\} = \\
&= \mathrm{E}\left\{\mathbf{x}_i\right\} \mathrm{E}\left\{\mathbf{n}\right\}
\end{aligned} \tag{1.42}$$

per calcolare il valore quadratico medio invece:

$$\begin{aligned}
\mathrm{E}\left\{\mathbf{y}^2 \,\middle|\, n\right\} &= \mathrm{E}\left\{(\textstyle\sum_{i=1}^{\mathbf{n}} \mathbf{x}_i)^2 \,\middle|\, n\right\} = \\
&= \mathrm{E}\left\{\textstyle\sum_{i=1}^{\mathbf{n}} \mathbf{x}_i^2 + \sum_{i=1}^{\mathbf{n}} \sum_{j=1;\, j\neq i}^{\mathbf{n}} \mathbf{x}_i \mathbf{x}_j \,\middle|\, n\right\} = \\
&= n\,\mathrm{E}\left\{\mathbf{x}_i^2\right\} + (n^2 - n)(\mathrm{E}\left\{\mathbf{x}_i\right\})^2
\end{aligned} \tag{1.43}$$

da cui:

$$\begin{aligned}
\mathrm{E}\left\{\mathbf{y}^2\right\} &= \mathrm{E}\left\{\mathrm{E}\left\{\mathbf{y}^2 \,\middle|\, \mathbf{n}\right\}\right\} = \sum_{n=0}^{\infty} \mathrm{E}\left\{\mathbf{y}^2 \,\middle|\, n\right\} \mathrm{Prob}\left\{\mathbf{n} = n\right\} = \\
&= \sum_{n=0}^{\infty} \left(n\,\mathrm{E}\left\{\mathbf{x}_i^2\right\} + (n^2 - n)(\mathrm{E}\left\{\mathbf{x}_i\right\})^2\right) \mathrm{Prob}\left\{\mathbf{n} = n\right\} = \\
&= \mathrm{E}\left\{\mathbf{n}\right\} \mathrm{E}\left\{\mathbf{x}_i^2\right\} + \left(\mathrm{E}\left\{\mathbf{n}^2\right\} - \mathrm{E}\left\{\mathbf{n}\right\}\right) (\mathrm{E}\left\{\mathbf{x}_i\right\})^2
\end{aligned} \tag{1.44}$$

infine la varianza vale:

$$\begin{aligned}\mathrm{Var}\{\mathbf{y}\} &= \mathrm{E}\{\mathbf{y}^2\} - (\mathrm{E}\{\mathbf{y}\})^2 = \\ &= \mathrm{E}\{\mathbf{n}\}\,\mathrm{E}\{\mathbf{x}_i^2\} + \left(\mathrm{E}\{\mathbf{n}^2\} - \mathrm{E}\{\mathbf{n}\}\right)(\mathrm{E}\{\mathbf{x}_i\})^2 - (\mathrm{E}\{\mathbf{x}_i\}\,\mathrm{E}\{\mathbf{n}\})^2 = \\ &= \mathrm{E}\{\mathbf{n}\}\,\mathrm{Var}\{\mathbf{x}_i\} + (\mathrm{E}\{\mathbf{x}_i\})^2\,\mathrm{Var}\{\mathbf{n}\}\,.\end{aligned} \tag{1.45}$$

□

1.2.10 Funzione generatrice dei momenti per variabili aleatorie discrete (trasformata Z)

Definizione 1.16. *Data la variabile aleatoria discreta ed a valori interi* $\mathbf{x}$, *si definisce* funzione generatrice dei momenti *la funzione complessa:*

$$X(z) \doteq \mathrm{E}\{z^{\mathbf{x}}\} = \sum_k z^k\,\mathrm{Prob}\{\mathbf{x}=k\}\ ,\ z \in \mathbb{C}. \tag{1.46}$$

La funzione generatrice dei momenti è, in pratica, la trasformata Z della funzione di probabilità, a meno del segno dell'esponente.
La funzione generatrice dei momenti di una qualunque distribuzione discreta gode delle seguenti proprietà:

1. Assume valore unitario per $z = 1$:

$$X(1) = \sum_k \mathrm{Prob}\{\mathbf{x}=k\} = 1 \tag{1.47}$$

2. è analitica all'interno della cerchio unitario (per $|z| \le 1$):

$$|X(z)| = \left|\sum_k z^k\,\mathrm{Prob}\{\mathbf{x}=k\}\right| \le \sum_k \left|z^k\right|\mathrm{Prob}\{\mathbf{x}=k\} \le 1 \tag{1.48}$$

3. Il valore medio può essere ricavato dalla derivata prima della funzione generatrice:

$$X'(z) = \sum_k k z^{k-1}\,\mathrm{Prob}\{\mathbf{x}=k\} \tag{1.49}$$

   da cui:

$$X'(1) = \sum_k k\,\mathrm{Prob}\{\mathbf{x}=k\} = \mathrm{E}\{\mathbf{x}\} \tag{1.50}$$

4. Il valore quadratico medio e la varianza possono essere ricavati dalle derivate prima e seconda della funzione generatrice:

$$X''(z) = \sum_k k(k-1)z^{k-2}\,\mathrm{Prob}\big\{\mathbf{x} = k\big\} \tag{1.51}$$

da cui:

$$\begin{aligned} X''(1) &= \sum_k k^2\,\mathrm{Prob}\big\{\mathbf{x} = k\big\} - \sum_k k\,\mathrm{Prob}\big\{\mathbf{x} = k\big\} = \\ &= \mathrm{E}\left\{\mathbf{x}^2\right\} - \mathrm{E}\left\{\mathbf{x}\right\} \end{aligned} \tag{1.52}$$

quindi:

$$\mathrm{E}\left\{\mathbf{x}^2\right\} = X''(1) + X'(1) \tag{1.53}$$

Per la varianza invece:

$$\begin{aligned} \mathrm{Var}\left\{\mathbf{x}\right\} &= \mathrm{E}\left\{\mathbf{x}^2\right\} - (\mathrm{E}\left\{\mathbf{x}\right\})^2 = \\ &= X''(1) + X'(1) - (X'(1))^2 \end{aligned} \tag{1.54}$$

5. Date le n variabili aleatorie indipendenti $\mathbf{x}_1, \ldots, \mathbf{x}_n$ con funzioni generatrici $X_1(z), \ldots, X_n(z)$, allora la variabile aleatoria:

$$\mathbf{y} = \sum_{k=1}^{n} \mathbf{x}_k$$

ha per funzione generatrice:

$$Y(z) = \prod_{k=1}^{n} X_k(z) \tag{1.55}$$

infatti:

$$Y(z) = \mathrm{E}\left\{z^{\sum_{k=1}^{n}\mathbf{x}_k}\right\} = \mathrm{E}\left\{\prod_{k=1}^{n} z^{\mathbf{x}_k}\right\} = \prod_{k=1}^{n}\mathrm{E}\left\{z^{\mathbf{x}_k}\right\} = \prod_{k=1}^{n} X_k(z) \tag{1.56}$$

6. Siano $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_k$ variabili aleatorie i.i.d. con funzioni generatrici $X(z)$, e sia $\mathbf{k}$ una variabile aleatoria indipendente dalle precedenti e con funzione generatrice $K(z)$. Allora la variabile aleatoria:

$$\mathbf{y} = \sum_{i=1}^{\mathbf{k}} \mathbf{x}_i$$

ha per funzione generatrice:

$$Y(z) = K(X(z)) \tag{1.57}$$

infatti:

$$\begin{aligned}
Y(z) &= \mathrm{E}\left\{z^{\mathbf{y}}\right\} = \mathrm{E}\left\{z^{\sum_{i=1}^{\mathbf{k}} \mathbf{x}_i}\right\} = \\
&= \sum_k \mathrm{E}\left\{z^{\sum_{i=1}^{k} \mathbf{x}_i} \,\middle|\, \mathbf{k}=k\right\} \mathrm{Prob}\left\{\mathbf{k}=k\right\} = \\
&= \sum_j (X(z))^k \,\mathrm{Prob}\left\{\mathbf{k}=k\right\} = \\
&= K(X(z))
\end{aligned} \tag{1.58}$$

1.2.11 Funzione generatrice dei momenti per variabili aleatorie continue (trasformata di Laplace)

Data la variabile aleatoria continua a valori non negativi $\mathbf{x}$, con funzione di densità di probabilità $f_{\mathbf{x}}(x)$ si definisce trasformata di Laplace la funzione complessa:

$$\beta_{\mathbf{x}}(s) = \mathrm{E}\left\{e^{-s\mathbf{x}}\right\} = \int_0^{\infty} f_{\mathbf{x}}(x) e^{-sx} dx \ , \ s \in \mathbb{C}. \tag{1.59}$$

La Trasformata di Laplace ha proprietà simili a quelle della funzione generatrice:

1. La trasformata di Laplace ha valore unitario in $s = 0$:

$$\beta_{\mathbf{x}}(0) = \int_0^{\infty} f_{\mathbf{x}}(x) dx = 1 \tag{1.60}$$

2. La trasformata di Laplace è una funzione analitica per $\mathrm{Re}(s) \geq 0$:

$$\begin{aligned}
|\beta_{\mathbf{x}}(s)| &= \left| \int_0^{\infty} f_{\mathbf{x}}(x) e^{-sx} dx \right| \leq \\
&\leq \int_0^{\infty} f_{\mathbf{x}}(x) \left| e^{-sx} \right| dx \leq \\
&\leq \int_0^{\infty} f_{\mathbf{x}}(x) dx = 1
\end{aligned} \tag{1.61}$$

3. Dalle derivate della trasformata di Laplace calcolate per $s = 0$ è possibile ricavare tutti i momenti della variabile aleatoria mediante la relazione:

$$\mathrm{E}\left\{\mathbf{x}^k\right\} = (-1)^k \beta_{\mathbf{x}}^{(k)}(0) \tag{1.62}$$

infatti:

$$
\begin{aligned}
\beta'_{\mathbf{x}}(s) &= \int_0^\infty -x f_{\mathbf{x}}(x) e^{-sx} dx \\
\beta''_{\mathbf{x}}(s) &= \int_0^\infty (-x)^2 f_{\mathbf{x}}(x) e^{-sx} dx \\
&\vdots \\
\beta^{(k)}_{\mathbf{x}}(s) &= \int_0^\infty (-x)^k f_{\mathbf{x}}(x) e^{-sx} dx
\end{aligned}
$$

quindi per $s = 0$:

$$\beta^{(k)}_{\mathbf{x}}(0) = \int_0^\infty (-x)^k f_{\mathbf{x}}(x) dx = (-1)^k \,\mathrm{E}\left\{\mathbf{x}^k\right\} \tag{1.63}$$

In particolare il valore medio, il valore quadratico medio e la varianza sono:

$$\mathrm{E}\{\mathbf{x}\} = -\beta'_{\mathbf{x}}(0) \tag{1.64}$$

$$\mathrm{E}\left\{\mathbf{x}^2\right\} = \beta''_{\mathbf{x}}(0) \tag{1.65}$$

$$\mathrm{Var}\{\mathbf{x}\} = \beta''_{\mathbf{x}}(0) - (\beta'_{\mathbf{x}}(0))^2 \tag{1.66}$$

4. Date le n variabili aleatorie indipendenti $\mathbf{x}_1, \ldots, \mathbf{x}_n$ con trasformate di Laplace $\beta_{\mathbf{x}_1}(s), \ldots, \beta_{\mathbf{x}_n}(s)$, allora la variabile aleatoria:

$$\mathbf{y} = \sum_{k=1}^{n} \mathbf{x}_k$$

ha per trasformata di Laplace:

$$\beta_{\mathbf{y}}(s) = \prod_{k=1}^{n} \beta_{\mathbf{x}_k}(s) \tag{1.67}$$

infatti:

$$
\begin{aligned}
\beta_{\mathbf{y}}(s) &= \mathrm{E}\left\{e^{-s\sum_{k=1}^n \mathbf{x}_k}\right\} = \mathrm{E}\left\{\prod_{k=1}^{n} e^{-s\mathbf{x}_k}\right\} = \\
&= \prod_{k=1}^{n} \mathrm{E}\left\{e^{-s\mathbf{x}_k}\right\} = \prod_{k=1}^{n} \beta_{\mathbf{x}_k}(s)
\end{aligned} \tag{1.68}
$$

5. Siano $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_k$ variabili aleatorie i.i.d. con trasformata di Laplace $\beta_{\mathbf{x}}(s)$, e sia $\mathbf{k}$ una variabile aleatoria indipendente dalle precedenti e con funzione generatrice $K(z)$. Allora la variabile aleatoria:

$$\mathbf{y} = \sum_{i=1}^{\mathbf{k}} \mathbf{x}_i$$

ha per trasformata di Laplace:

$$\beta_{\mathbf{y}}(s) = K(\beta_{\mathbf{x}}(s)) \tag{1.69}$$

infatti:

$$\begin{aligned} \beta_{\mathbf{y}}(s) &= \mathrm{E}\left\{e^{-s\mathbf{y}}\right\} = \mathrm{E}\left\{e^{-s\sum_{i=1}^{\mathbf{k}} \mathbf{x}_i}\right\} = \\ &= \sum_k \mathrm{E}\left\{e^{-s\sum_{i=1}^{k} \mathbf{x}_i} \,\middle|\, \mathbf{k} = k\right\} \mathrm{Prob}\big\{\mathbf{k} = k\big\} = \\ &= \sum_k (\beta_{\mathbf{x}}(s))^k \,\mathrm{Prob}\big\{\mathbf{k} = k\big\} = \\ &= K(\beta_{\mathbf{x}}(s)) \end{aligned} \tag{1.70}$$

1.2.12 Funzione caratteristica

Data la variabile aleatoria $\mathbf{x}$ con densità di probabilità $f_{\mathbf{x}}(x)$, si definisce funzione caratteristica l'integrale:

$$\Phi_{\mathbf{x}}(\omega) = \mathrm{E}\left\{e^{j\omega\mathbf{x}}\right\} = \int_{-\infty}^{\infty} f_{\mathbf{x}}(x) e^{j\omega x} dx \tag{1.71}$$

che in pratica non è altro che la trasformata di Fourier (a parte il segno della variabile ω) della densità di probabilità. Per variabili aleatorie discrete la funzione caratteristica diventa:

$$\Phi_{\mathbf{x}}(\omega) = \sum_i e^{j\omega x_i} \,\mathrm{Prob}\big\{\mathbf{x} = x_i\big\} \tag{1.72}$$

Le proprietà più rilevanti della funzione caratteristica sono:

1. La funzione caratteristica assume il valore 1 nell'origine:

$$\Phi_{\mathbf{x}}(0) = \int_{-\infty}^{\infty} f_{\mathbf{x}}(x) dx = 1 \tag{1.73}$$

2. La funzione caratteristica ha modulo massimo nell'origine:

$$|\Phi_{\mathbf{x}}(\omega)| = \left| \int_{-\infty}^{\infty} f_{\mathbf{x}}(x) e^{j\omega x} dx \right| \leq \int_{-\infty}^{\infty} f_{\mathbf{x}}(x) = 1 \tag{1.74}$$

3. I momenti della variabile aleatoria possono essere calcolati dalle derivate della funzione caratteristica calcolata in 0 con la relazione:

$$\mathrm{E}\left\{\mathbf{x}^k\right\} = (-j)^k \Phi_{\mathbf{x}}^{(k)}(0) \tag{1.75}$$

infatti:

$$\begin{aligned}
\Phi'_{\mathbf{x}}(\omega) &= \int_{-\infty}^{\infty} j x f_{\mathbf{x}}(x) e^{j\omega x} dx \\
\Phi''_{\mathbf{x}}(\omega) &= -\int_{-\infty}^{\infty} x^2 f_{\mathbf{x}}(x) e^{j\omega x} dx \\
&\vdots \\
\Phi^{(k)}_{\mathbf{x}}(\omega) &= \int_{-\infty}^{\infty} j^k x^k f_{\mathbf{x}}(x) e^{j\omega x} dx
\end{aligned} \tag{1.76}$$

quindi per $\omega = 0$

$$\Phi^{(k)}_{\mathbf{x}}(0) = j^k \int_{-\infty}^{\infty} x^k f_{\mathbf{x}}(x) dx \tag{1.77}$$

4. Date le n variabili aleatorie $\mathbf{x}_1, \ldots, \mathbf{x}_n$ indipendenti e con funzioni caratteristiche $\Phi_{\mathbf{x}_1}(\omega), \ldots, \Phi_{\mathbf{x}_n}(\omega)$, allora la variabile aleatoria:

$$\mathbf{y} = \sum_{k=1}^{n} \mathbf{x}_k$$

ha per funzione caratteristica:

$$\Phi_{\mathbf{y}}(\omega) = \prod_{k=1}^{n} \Phi_{\mathbf{x}_k}(\omega) \tag{1.78}$$

infatti:

$$\begin{aligned}
\Phi_{\mathbf{y}}(\omega) &= \mathrm{E}\left\{e^{j\omega \sum_{k=1}^{n} \mathbf{x}_k}\right\} = \mathrm{E}\left\{\prod_{k=1}^{n} e^{j\omega \mathbf{x}_k}\right\} = \\
&= \prod_{k=1}^{n} \mathrm{E}\left\{e^{j\omega \mathbf{x}_k}\right\} = \prod_{k=1}^{n} \Phi_{\mathbf{x}_k}(\omega)
\end{aligned} \tag{1.79}$$

5. Siano $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_k$ variabili aleatorie i.i.d. con funzione caratteristica $\Phi_{\mathbf{x}}(\omega)$, e sia $\mathbf{k}$ una variabile aleatoria indipendente dalle precedenti e con funzione generatrice $K(z)$. Allora la variabile aleatoria:

$$\mathbf{y} = \sum_{i=1}^{\mathbf{k}} \mathbf{x}_i$$

ha come funzione caratteristica:

$$\beta_{\mathbf{y}}(s) = K(\Phi_{\mathbf{x}}(\omega)) \tag{1.80}$$

La dimostrazione é completamente analoga a quella relativa alla trasformata di Laplace di una somma di k v.a. i.i.d. ed é, pertanto, omessa.

1.3 Distribuzioni notevoli

In questo paragrafo sono elencate le principali distribuzioni delle variabili aleatorie continue e discrete, ed i relativi parametri.

1.3.1 Distribuzione Uniforme

Una variabile aleatoria continua $\mathbf{x}$ si dice uniformemente distribuita nell'intervallo $[a, b]$ se la sua funzione di densità di probabilità è:

$$f_{\mathbf{x}}(x) = \begin{cases} \frac{1}{b-a} & a \leq x \leq b \\ 0 & \text{altrimenti} \end{cases} \tag{1.81}$$

e la funzione di distribuzione risulta quindi:

$$F_{\mathbf{x}}(x) = \begin{cases} 0 & x < a \\ \frac{x-a}{b-a} & a \leq x \leq b \\ 1 & x > b \end{cases} \tag{1.82}$$

I parametri della distribuzione uniforme sono:

$$\mathrm{E}\{\mathbf{x}\} = \int_a^b \frac{x}{b-a} dx = \frac{a+b}{2} \tag{1.83}$$

$$\mathrm{E}\left\{\mathbf{x}^2\right\} = \int_a^b \frac{x^2}{b-a} dx = \frac{a^2+b^2+ab}{3} \tag{1.84}$$

$$\mathrm{Var}\{\mathbf{x}\} = \mathrm{E}\left\{\mathbf{x}^2\right\} - (\mathrm{E}\{\mathbf{x}\})^2 = \frac{(a-b)^2}{12} \tag{1.85}$$

La trasformata di Laplace invece vale:

$$\beta_{\mathbf{x}}(s) = \int_a^b \frac{e^{-sx}}{b-a} dx = \frac{e^{-as} - e^{-bs}}{s(b-a)} \tag{1.86}$$

Una variabile aleatoria discreta $\mathbf{x}$ che assuma i valori $\{a, a+1, a+2, \ldots, a+N-1\}$ con a intero, si dice uniformemente distribuita, se la sua funzione di probabilità è:

$$\text{Prob}\big\{\mathbf{x} = k\big\} = \begin{cases} \frac{1}{N} & a \leq k \leq a+N-1 \\ 0 & \text{altrimenti} \end{cases} \tag{1.87}$$

I parametri della distribuzione uniforme sono:

$$\text{E}\left\{\mathbf{x}\right\} = \sum_{k=a}^{a+N-1} \frac{k}{N} = \sum_{k=0}^{N-1} \frac{a+k}{N} = a + \frac{(N-1)}{2} \tag{1.88}$$

$$\begin{aligned} \text{E}\left\{\mathbf{x}^2\right\} &= \sum_{k=a}^{a+N-1} \frac{k^2}{N} = \sum_{k=0}^{N-1} \frac{(a+k)^2}{N} = \\ &= a^2 + 2a\frac{N-1}{2} + \frac{(2N-1)(N-1)}{6} \end{aligned} \tag{1.89}$$

$$\text{Var}\left\{\mathbf{x}\right\} = \text{E}\left\{\mathbf{x}^2\right\} - (\text{E}\left\{\mathbf{x}\right\})^2 = \frac{N^2-1}{12} \tag{1.90}$$

La funzione generatrice invece vale:

$$X(z) = \sum_{k=a}^{a+N-1} \frac{z^k}{N} = \frac{z^a(z^N-1)}{N(z-1)} \tag{1.91}$$

1.3.2 Distribuzione Esponenziale

Una variabile aleatoria continua $\mathbf{x}$ segue una distribuzione esponenziale se la sua densità di probabilità è:

$$f_{\mathbf{x}}(x) = \begin{cases} \lambda e^{-\lambda x} & x \geq 0 \\ 0 & \text{altrimenti} \end{cases} \tag{1.92}$$

e la funzione di distribuzione risulta quindi:

$$F_{\mathbf{x}}(x) = \begin{cases} 1 - e^{-\lambda x} & x \geq 0 \\ 0 & \text{altrimenti} \end{cases} \tag{1.93}$$

I parametri della distribuzione esponenziale sono:

$$\begin{aligned} \text{E}\left\{\mathbf{x}\right\} &= \int_0^\infty x\lambda e^{-\lambda x} dx = \\ &= \left[-xe^{-\lambda x}\right]_0^\infty + \int_0^\infty e^{-\lambda x} dx = \\ &= \left[-xe^{-\lambda x}\right]_0^\infty - \left[\frac{e^{-\lambda x}}{\lambda}\right]_0^\infty = \\ &= \frac{1}{\lambda} \end{aligned} \tag{1.94}$$

$$\begin{aligned}
\mathrm{E}\left\{\mathbf{x}^2\right\} &= \int_0^\infty x^2\lambda e^{-\lambda x}dx = \\
&= \left[-x^2e^{-\lambda x}\right]_0^\infty + \int_0^\infty 2xe^{-\lambda x}dx = \\
&= \left[-x^2e^{-\lambda x}\right]_0^\infty + \left[-\frac{2xe^{-\lambda x}}{\lambda}\right]_0^\infty - \left[\frac{2e^{-\lambda x}}{\lambda^2}\right]_0^\infty = \\
&= \frac{2}{\lambda^2}
\end{aligned} \tag{1.95}$$

$$\mathrm{Var}\left\{\mathbf{x}\right\} = \mathrm{E}\left\{\mathbf{x}^2\right\} - \left(\mathrm{E}\left\{\mathbf{x}\right\}\right)^2 = \frac{1}{\lambda^2} \tag{1.96}$$

La trasformata di Laplace, invece, vale:

$$\begin{aligned}
\beta_{\mathbf{x}}(s) &= \int_0^\infty \lambda e^{-\lambda x}e^{-sx}dx = \\
&= \left[\frac{\lambda}{s+\lambda}e^{-(s+\lambda)x}\right]_0^\infty = \\
&= \frac{\lambda}{s+\lambda}
\end{aligned} \tag{1.97}$$

La proprietà più importante della distribuzione esponenziale è l'*assenza di memoria*, cioè, se il tempo di attesa di un certo evento è una variabile aleatoria esponenziale $\mathbf{t}$ con media $\frac{1}{\lambda}$ allora, supponendo che sia già trascorso un tempo τ, il tempo residuo $\mathbf{x} = \mathbf{t} - \tau$ prima del verificarsi dell'evento è ancora una variabile aleatoria esponenziale con la stessa media:

$$\begin{aligned}
F_{\mathbf{x}}(x \,\big|\, \mathbf{t} > \tau) &= \mathrm{Prob}\left\{\mathbf{x} \le x \,\big|\, \mathbf{t} > \tau\right\} = \\
&= \frac{\mathrm{Prob}\left\{\mathbf{t} - \tau \le x, \mathbf{t} > \tau\right\}}{\mathrm{Prob}\left\{\mathbf{t} > \tau\right\}} = \\
&= \frac{\mathrm{Prob}\left\{\tau < \mathbf{t} \le x + \tau\right\}}{\mathrm{Prob}\left\{\mathbf{t} > \tau\right\}} = \\
&= \frac{1 - e^{-\lambda(x+\tau)} - (1 - e^{-\lambda\tau})}{1 - (1 - e^{-\lambda\tau})} = \\
&= \frac{e^{-\lambda\tau} - e^{-\lambda(x+\tau)}}{e^{-\lambda\tau}} = \\
&= 1 - e^{-\lambda x}
\end{aligned} \tag{1.98}$$

1.3.3 Distribuzione Gaussiana

Una variabile aleatoria continua $\mathbf{x}$ segue una distribuzione gaussiana con media μ e varianza σ^2 se la sua funzione di densità di probabilità è:

$$f_{\mathbf{x}}(x) = \frac{1}{\sqrt{2\pi\sigma^2}}e^{-\frac{(x-\mu)^2}{2\sigma^2}} \tag{1.99}$$

Non è possibile derivare in modo semplice i parametri di questa distribuzione, ma essi possono essere dimostrati a posteriori. Per la media basta considerare:

$$\begin{aligned}
&\int_{-\infty}^{+\infty} x\frac{1}{\sqrt{2\pi\sigma^2}}e^{-\frac{(x-\mu)^2}{2\sigma^2}}dx - \mu + \mu = \\
&= \int_{-\infty}^{+\infty} x\frac{1}{\sqrt{2\pi\sigma^2}}e^{-\frac{(x-\mu)^2}{2\sigma^2}}dx - \mu\int_{-\infty}^{+\infty}\frac{1}{\sqrt{2\pi\sigma^2}}e^{-\frac{(x-\mu)^2}{2\sigma^2}}dx + \mu = \\
&= \int_{-\infty}^{+\infty}\frac{x-\mu}{\sqrt{2\pi\sigma^2}}e^{-\frac{(x-\mu)^2}{2\sigma^2}}dx + \mu = \\
&= \int_{-\infty}^{+\infty}\frac{z}{\sqrt{2\pi\sigma^2}}e^{-\frac{z^2}{2\sigma^2}}dx + \mu = \mu
\end{aligned}$$

Per la varianza invece si considera la seguente uguaglianza:

$$\int_{-\infty}^{+\infty}\frac{1}{\sqrt{2\pi\sigma^2}}e^{-\frac{(x-\mu)^2}{2\sigma^2}}dx = 1$$

si moltiplica per $\sqrt{2\pi\sigma^2}$:

$$\int_{-\infty}^{+\infty} e^{-\frac{(x-\mu)^2}{2\sigma^2}}dx = \sqrt{2\pi\sigma^2}$$

si deriva rispetto a σ:

$$\int_{-\infty}^{+\infty} -\frac{(x-\mu)^2}{2}\cdot\left(-\frac{2}{\sigma^2}\right)e^{-\frac{(x-\mu)^2}{2\sigma^3}}dx = \sqrt{2\pi}$$

quindi moltiplicando per σ^2 e dividendo per $\sqrt{2\pi}$:

$$\int_{-\infty}^{+\infty}(x-\mu)^2\frac{1}{\sqrt{2\pi\sigma^2}}e^{-\frac{(x-\mu)^2}{2\sigma^2}}dx = \sigma^2$$

1.3.4 Distribuzione di Bernoulli

Una variabile aleatoria discreta $\mathbf{x}$ segue una distribuzione di Bernoulli se può assumere solo i valori 0 e 1 e la sua funzione di probabilità è:

$$\text{Prob}\{\mathbf{x} = k\} = \begin{cases} 1-p & k=0 \\ p & k=1 \end{cases} \tag{1.100}$$

I parametri della distribuzione di Bernoulli sono:

$$\text{E}\{\mathbf{x}\} = 0\cdot(1-p) + 1\cdot p = p \tag{1.101}$$

$$\mathrm{E}\left\{\mathbf{x}^2\right\} = 0^2 \cdot (1-p) + 1^2 \cdot p = p \tag{1.102}$$

$$\mathrm{Var}\left\{\mathbf{x}\right\} = \mathrm{E}\left\{\mathbf{x}^2\right\} - (\mathrm{E}\left\{\mathbf{x}\right\})^2 = p(1-p) \tag{1.103}$$

La funzione generatrice invece vale:

$$X(z) = (1-p)z^0 + pz = 1 + p(z-1) \tag{1.104}$$

1.3.5 Distribuzioni Geometrica e Geometrica modificata

Una variabile aleatoria discreta $\mathbf{x}$ segue una distribuzione geometrica se la sua funzione di probabilità è:

$$\mathrm{Prob}\left\{\mathbf{x} = k\right\} = p^k(1-p) \qquad 0 \le p \le 1; \quad k = 0, 1, 2, \ldots \tag{1.105}$$

fisicamente rappresenta il numero di prove indipendenti consecutive in cui si verifica un evento bernoulliano con probabilità p. I parametri della distribuzione geometrica sono:

$$\begin{aligned}
\mathrm{E}\left\{\mathbf{x}\right\} &= \sum_{k=0}^{\infty} kp^k(1-p) = p(1-p)\sum_{k=0}^{\infty} kp^{k-1} = \\
&= p(1-p)\frac{d}{dp}\sum_{k=0}^{\infty} p^k = p(1-p)\frac{d}{dp}\frac{1}{1-p} = \\
&= \frac{p}{1-p}
\end{aligned} \tag{1.106}$$

$$\begin{aligned}
\mathrm{E}\left\{\mathbf{x}^2\right\} &= \sum_{k=0}^{\infty} k^2p^k(1-p) = p(1-p)\sum_{k=0}^{\infty} k^2p^{k-1} = \\
&= p(1-p)\frac{d}{dp}\sum_{k=0}^{\infty} kp^k = p(1-p)\frac{d}{dp}\left(p\frac{d}{dp}\frac{1}{1-p}\right) = \\
&= \frac{p(1+p)}{(1-p)^2}
\end{aligned} \tag{1.107}$$

$$\mathrm{Var}\left\{\mathbf{x}\right\} = \mathrm{E}\left\{\mathbf{x}^2\right\} - (\mathrm{E}\left\{\mathbf{x}\right\})^2 = \frac{p}{(1-p)^2} \tag{1.108}$$

La funzione generatrice invece vale:

$$X(z) = \sum_{k=0}^{\infty} z^kp^k(1-p) = \frac{1-p}{1-pz} \tag{1.109}$$

Una variante della distribuzione geometrica la si ottiene considerando la variabile aleatoria che esprime il numero di prove indipendenti necessarie per il varificarsi del primo successo in un esperimento bernoulliano con probabilità $1-p$ (compresa quella con esito positivo):

$$\mathrm{Prob}\left\{\mathbf{x} = k\right\} = p(1-p)^{k-1} \qquad 0 \le p \le 1; \quad k = 1, 2, \ldots \tag{1.110}$$

I parametri della distribuzione geometrica modificata sono:

$$\begin{aligned} \mathrm{E}\{\mathbf{x}\} &= \sum_{k=1}^{\infty} kp(1-p)^{k-1} = p\sum_{k=0}^{\infty} k(1-p)^{k-1} = \\ &= -p\frac{d}{dp}\sum_{k=0}^{\infty}(1-p)^k = -p\frac{d}{dp}\frac{1}{p} = \\ &= \frac{1}{p} \end{aligned} \tag{1.111}$$

$$\begin{aligned} \mathrm{E}\left\{\mathbf{x}^2\right\} &= \sum_{k=1}^{\infty} k^2p(1-p)^{k-1} = p\sum_{k=0}^{\infty} k^2(1-p)^{k-1} = \\ &= p\frac{d}{dp}\sum_{k=0}^{\infty} -k(1-p)^k = p\frac{d}{dp}\left((1-p)\frac{d}{dp}\frac{1}{p}\right) = \\ &= \frac{2-p}{p^2} \end{aligned} \tag{1.112}$$

$$\mathrm{Var}\{\mathbf{x}\} = \mathrm{E}\left\{\mathbf{x}^2\right\} - (\mathrm{E}\{\mathbf{x}\})^2 = \frac{1-p}{p^2} \tag{1.113}$$

La funzione generatrice, invece, vale:

$$\begin{aligned} X(z) &= \sum_{k=1}^{\infty} z^k p(1-p)^{k-1} = pz\sum_{k=1}^{\infty}(z(1-p))^{k-1} = \\ &= \frac{pz}{1-z(1-p)} \end{aligned} \tag{1.114}$$

1.3.6 Distribuzione Binomiale

La distribuzione binomiale è quella che si ottiene considerando come variabile aleatoria $\mathbf{x}$ il numero di successi in n ripetizioni indipendenti di un esperimento bernoulliano con probabilità di successo p. La relativa funzione di probabilità è:

$$\mathrm{Prob}\big\{\mathbf{x} = k\big\} = \binom{n}{k} p^k (1-p)^{n-k} \qquad \begin{array}{l} 0 \leq p \leq 1 \\ k = 1, 2, \ldots, n \end{array}$$

I parametri della distribuzione binomiale sono:

$$
\begin{aligned}
\mathrm{E}\{\mathbf{x}\} &= \sum_{k=0}^{n} k \binom{n}{k} p^k (1-p)^{n-k} = \\
&= np \sum_{k=1}^{n} \frac{(n-1)!}{(k-1)!(n-k)!} p^{k-1} (1-p)^{n-k} = \\
&= np \sum_{k=0}^{n-1} \frac{(n-1)!}{k!(n-1-k)!} p^{k} (1-p)^{n-1-k} = \\
&= np(p+(1-p))^{n-1} = \\
&= np
\end{aligned}
\tag{1.116}
$$

$$
\begin{aligned}
\mathrm{E}\left\{\mathbf{x}^2\right\} &= \sum_{k=0}^{n} k^2 \sum_{d=d_{free}^{H}}^{\infty} nkp^k (1-p)^{n-k} = \\
&= np \sum_{k=1}^{n} (1+(k-1)) \frac{(n-1)!}{(k-1)!(n-k)!} p^{k-1} (1-p)^{n-k} = \\
&= np + np \sum_{k=0}^{n-1} k \frac{(n-1)!}{k!(n-k)!} p^{k} (1-p)^{n-k} = \\
&= np + np((n-1)p) \\
&= np + n(n-1)p^2
\end{aligned}
\tag{1.117}
$$

$$
\mathrm{Var}\{\mathbf{x}\} = \mathrm{E}\left\{\mathbf{x}^2\right\} - (\mathrm{E}\{\mathbf{x}\})^2 = np(1-p) \tag{1.118}
$$

La funzione generatrice invece vale:

$$
\begin{aligned}
X(z) &= \sum_{k=1}^{\infty} z^k \binom{n}{k} p^k (1-p)^{n-k} = \\
&= \sum_{k=1}^{\infty} \binom{n}{k} (pz)^k (1-p)^{n-k} = (pz+(1-p))^n = \\
&= (p(z-1)+1)^n
\end{aligned}
\tag{1.119}
$$

1.3.7 Distribuzione di Poisson

Si consideri l'intervallo temporale $[-\frac{T}{2}, \frac{T}{2}]$ nel quale siano distribuiti casualmente k punti (ad esempio gli istanti di arrivo di altrettante persone in una coda). Sia inoltre $\mathbf{x}$ la variabile aleatoria che esprime il numero di punti che si trovano in un intervallo di lunghezza τ

comunque posizionato all'interno di $[-\frac{T}{2}, \frac{T}{2}]$. La variabile aleatoria $\mathbf{x}$ ha una distribuzione binomiale:

$$\text{Prob}\{\mathbf{x} = k\} = \binom{n}{k} p^k (1-p)^{n-k} \tag{1.120}$$

dove:

$$p = \frac{\tau}{T} \tag{1.121}$$

allargando l'intervallo temporale T ed il numero di punti n in modo che il numero di punti per unità di tempo (detto *tasso di arrivi*) rimanga costante pari a:

$$\lambda = \frac{n}{T} \tag{1.122}$$

si ottiene $p = \frac{\tau\lambda}{n}$, quindi per $n \to \infty$:

$$\begin{aligned}
\text{Prob}\{\mathbf{x} = k\} &= \lim_{n\to\infty} \binom{n}{k} p^k (1-p)^{n-k} = \\
&= \lim_{n\to\infty} \binom{n}{k} \left(\frac{\lambda\tau}{n}\right)^k \left(1 - \frac{\lambda\tau}{n}\right)^{n-k} = \\
&= \lim_{n\to\infty} \frac{(\lambda\tau)^k}{k!} \left(1 - \frac{\lambda\tau}{n}\right)^n \frac{n!}{(n-k)!(n-\lambda\tau)^k} = \\
&= \lim_{n\to\infty} \frac{(\lambda\tau)^k}{k!} \left(1 - \frac{\lambda\tau}{n}\right)^n \frac{n(n-1)\dots(n-k+1)}{(n-\lambda\tau)^k} = \\
&= \frac{(\lambda\tau)^k}{k!} e^{-\lambda\tau}
\end{aligned} \tag{1.123}$$

questa è la distribuzione di Poisson.
Una variabile aleatoria discreta $\mathbf{x}$ segue quindi una distribuzione di Poisson se la sua funzione di probabilità è:

$$\text{Prob}\{\mathbf{x} = k\} = \frac{(\lambda\tau)^k}{k!} e^{-\lambda\tau} \qquad \lambda > 0,\ \tau > 0 \tag{1.124}$$

I parametri della distribuzione di Poisson sono:

$$\text{E}\{\mathbf{x}\} = \sum_{k=0}^{\infty} k \frac{(\lambda\tau)^k}{k!} e^{-\lambda\tau} = \lambda\tau e^{-\lambda\tau} \sum_{k=1}^{\infty} \frac{(\lambda\tau)^{k-1}}{(k-1)!} = \lambda\tau \tag{1.125}$$

$$\begin{aligned}
\mathrm{E}\left\{\mathbf{x}^2\right\} &= \sum_{k=0}^{\infty} k^2 \frac{(\lambda\tau)^k}{k!} e^{-\lambda\tau} = \lambda\tau e^{-\lambda\tau} \sum_{k=1}^{\infty} k \frac{(\lambda\tau)^{k-1}}{(k-1)!} = \\
&= \lambda\tau e^{-\lambda\tau} \sum_{k=1}^{\infty} (1 + (k-1)) \frac{(\lambda\tau)^{k-1}}{(k-1)!} = \\
&= \lambda\tau e^{-\lambda\tau} \left(e^{\lambda\tau} + \sum_{k=1}^{\infty} (k-1) \frac{(\lambda\tau)^{k-1}}{(k-1)!} \right) = \\
&= \lambda\tau e^{-\lambda\tau} \left(e^{\lambda\tau} + \lambda\tau e^{\lambda\tau} \right) = \\
&= \lambda\tau(1 + \lambda\tau)
\end{aligned} \tag{1.126}$$

$$\mathrm{Var}\left\{\mathbf{x}\right\} = \mathrm{E}\left\{\mathbf{x}^2\right\} - \left(\mathrm{E}\left\{\mathbf{x}\right\}\right)^2 = \lambda\tau \tag{1.127}$$

La funzione generatrice invece vale:

$$X(z) = \sum_{k=0}^{\infty} z^k \frac{(\lambda\tau)^k}{k!} e^{-\lambda\tau} = e^{-\lambda\tau} e^{\lambda\tau z} = e^{\lambda\tau(z-1)} \tag{1.128}$$

La distribuzione di Poisson ha alcune proprietà importanti:

1. Indicando con $\mathbf{x}$ la somma di due variabili poissoniane indipendenti $\mathbf{x}_1$ e $\mathbf{x}_2$, $\mathbf{x}$ è ancora una variabile aleatoria di Poisson con valore medio pari alla somma dei valori medi di $\mathbf{x}_1$ e $\mathbf{x}_2$. Infatti, calcolando le funzioni generatrici si ottiene:

$$\begin{aligned}
X_1(z) &= e^{\lambda_1\tau(z-1)} \\
X_2(z) &= e^{\lambda_2\tau(z-1)} \\
X(z) &= X_1(z)X_2(z) = e^{(\lambda\tau_1 + \lambda_2\tau)(z-1)}
\end{aligned} \tag{1.129}$$

2. Se $\mathbf{t}_i$ con $i \in \mathbb{Z}$ sono le posizioni dei punti sull'asse temporale distribuiti secondo Poisson, allora la distanza:

$$\boldsymbol{\tau}_i = \mathbf{t}_{i+1} - \mathbf{t}_i$$

fra due di tali punti è una variabile aleatoria con distribuzione esponenziale con media $\frac{1}{\lambda}$:

$$\begin{aligned}
F_{\boldsymbol{\tau}_i}(t) &= \mathrm{Prob}\left\{\boldsymbol{\tau}_i \le t\right\} = \\
&= \mathrm{Prob}\left\{\text{almeno un arrivo in } (t_i, t_i + t)\right\} = \\
&= 1 - \mathrm{Prob}\left\{\text{nessun un arrivo in } (t_i, t_i + t)\right\} = \\
&= 1 - e^{-\lambda t}
\end{aligned} \tag{1.130}$$

1.3.8 Distribuzione di Pareto

Una variabile aleatoria continua a valori positivi $\mathbf{x}$ segue una distribuzione di Pareto se la relativa densitá di probabilitá assume la seguente espressione analitica:

$$f_{\mathbf{x}}(x) = \frac{\alpha k^{\alpha}}{x^{\alpha+1}}, \ x \geq k \tag{1.131}$$

dove k rappresenta il valore minimo assunto dalla v.a. $\mathbf{x}$ e viene designato come *location*, mentre il parametro α, indicato come *shape*, indica come decresce la pdf stessa. Si noti che l'andamento è tipicamente iperbolico e, pertanto, si potranno presentare dei valori elevati in ampiezza con probabilità di occorrenza non trascurabili come accade, invece, nel caso di pdf memoryless (sostanzialmente la pdf esponenziale).
Per quanto attiene la CDF si ha che:

$$F_{\mathbf{x}}(x) = \begin{cases} 0, \ x < k \\ 1 - \left(\frac{k}{x}\right)^{\alpha} \ , \ x \geq k \end{cases} \tag{1.132}$$