Riassunto di Probabilità
Corso di Analisi Matematica 2 e Probabilità

Francesco Gori Elisa Farri

5 luglio 2024

Indice

1 Descrizione di un Esperimento Casuale

1.1 Spazi di Probabilità - nozioni base

Definiamo Spazio di Probabilità la terna $(\Omega, \mathcal{A}, \mathbb{P})$. Ogni spazio è composto da:

- Ω: è lo spazio campionario degli eventi.
- $\mathcal{A}$: è una σ-algebra di sottoinsiemi di Ω con una certa struttura.
- $\mathbb{P}$ è la misura di probabilità su Ω.

Ω, ovvero lo **spazio campionario**, rappresenta l'insieme di tutti i possibili esiti di un esperimento casuale.

Esempio - lancio di un dado:
$\Omega = \{1, 2, 3, 4, 5, 6\}$

$\mathcal{A}$, ovvero la **σ-algebra**, rappresenta l'insieme di eventi legati all'esperimento considerato, ovvero gli eventi misurabili per un dato esperimento. Ad esempio, nel caso del lancio di un dato, A potrebbe rappresentare il caso in cui il numero uscito sia pari.
Nel caso in cui Ω sia finito, $\mathcal{A}$ è l'insieme delle parti di Ω, ovvero
$\mathcal{A} = \mathbb{P}(\Omega) = \{\Omega, \emptyset, \{1\}, \{2\}, \ldots, \{6\}, \{1, 2\}, \ldots, \{1, 2, 3, 4, 5\}, \ldots\}$
Assiomi per $\mathcal{A}$ a partire da Ω:

1. $\Omega \in A$
2. se $E \in \mathcal{A} \implies E^c = \Omega \setminus E \in \mathcal{A}$
3. se $E \in \mathcal{A}$

$\mathbb{P}$, **mis. di probabilità**, assegna a ciascun evento misurabile la probabiltà che esso si verifichi.

Definizione: $\mathbb{P} : \mathcal{A} \to [0, 1]$
$E \to \mathbb{P}(E) \in [0, 1]$ t.c.

1. $\mathbb{P}(\Omega) = 1$
2. se $(E_n)_{n\geq 1} \subseteq \mathcal{A}$ disgiunti (ovvero $\forall i, j\ E_i \cap E_j = \emptyset$) $\implies \mathbb{P}(\bigcup_{n\geq 1} E_n) = \sum_{n\geq 1} \mathbb{P}(E_n)$

Esempio - lancio di un dado:
$\mathbb{P}(\omega) = 1/6$ se il dado è equo. Altrimenti $\mathbb{P}(i) \neq 1/6$, ma in ogni caso $\sum \mathbb{P}(i) = 1$.

Una misura di probabilità $\mathbb{P}$ si dice uniforme se:

$$\forall \omega \in \Omega \mathbb{P}(\omega) = p = \frac{1}{|\Omega|}$$

con ω singleton, ovvero gli eventi singolari dell'esperimento.
Se $E \in \mathcal{A}$, allora

$$\mathbb{P}(E) = \sum_{\omega \in E} \frac{1}{|\Omega|} = \frac{|E|}{|\Omega|} = \frac{casi\, favorevoli}{casi\, possibili}$$

Esempio - lancio di due dadi la cui somma sia k:
$\Omega = \{\omega = (\omega_1, \omega_2) : \omega_1, \omega_2 \in \{1, 2, 3, 4, 5, 6\}\}$
$\mathbb{P}(\text{somma} = k) = \frac{\text{numero di coppie la cui somma è } k}{36}$

1.2 Proprietà e Formule

1.2.1 Proprietà Varie

- $\forall A, B \in \mathcal{A}, B = B \cap \Omega = B \cap (A \cup A^c) = (B \cap A) \coprod (B \cap A^c)$
  Da questo segue che $\mathbb{P}(B) = \mathbb{P}(B \cap A) + \mathbb{P}(B \cap A^c)$
- se $B = \Omega$ allora $\mathbb{P}(\Omega) = \mathbb{P}(A) + \mathbb{P}(A^c) \implies \mathbb{P}(A^c) = 1 - \mathbb{P}(A)$
- se $A \subseteq B$ allora $\mathbb{P}(B) = \mathbb{P}(A) + \mathbb{P}(B \setminus A) \geq \mathbb{P}(A)$
- $\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B \cap A^c) = \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cap B)$

Esempio - lancio di 2 dadi, probabilità che almeno uno sia un 6. Possiamo adottare 3 strategie diverse:

- contiamo i casi in cui abbiamo almeno un 6, che sono $\frac{11}{36}$
- $\mathbb{P}(\text{almeno un } 6) = 1 - \mathbb{P}(\text{neanche un } 6) = 1 - \frac{5*5}{36} = \frac{11}{36}$
- $A_i = \{\omega = (\omega_1, \omega_2) : \omega_i = 6\}$
  $\mathbb{P}(\mathcal{A}_1 \cup \mathcal{A}_2) = \mathbb{P}(\mathcal{A}_1) + \mathbb{P}(\mathcal{A}_2) - \mathbb{P}(\mathcal{A}_1 \cap \mathcal{A}_2) = \frac{6}{36} + \frac{6}{36} - \frac{1}{36} = \frac{11}{36}$

1.2.2 Probabilità Condizionata e Formula di Bayes

Assumiamo che $A \in \mathcal{A}$ con $\mathbb{P}(A) > 0$, allora avremo che $\forall B \in \mathcal{A}$

$$\mathbb{P}(B|A) := \frac{\mathbb{P}(B \cap A)}{\mathbb{P}(A)}$$

e si dice probabilità di B condizionato da A. La probabilità condizionata, indicata come $\mathbb{P}(B|A)$, rappresenta la probabilità che l'evento B si verifichi dato che l'evento A è già avvenuto. In altre parole, misura la probabilità di B nel contesto in cui sappiamo che A è verificato.

Esempio - fumatori che sviluppano malattie respiratorie: Supponiamo che una popolazione sia divisa in fumatori (40%) e non fumatori (60%). Supponiamo che il 25% dei fumatori sviluppi una malattia respiratoria, mentre per i non fumatori questa percentuale sia del 7%. Scegliendo una persona casualmente, qual è la probabilità che essa sviluppi una malattia respiratoria?
$\mathbb{P}(f) = 0.4, \quad \mathbb{P}(n) = 0.6$
$\mathbb{P}(m|f) = 0.25, \quad \mathbb{P}(m|n) = 0.07$
$\mathbb{P}(m) = \mathbb{P}(m \cap f) + \mathbb{P}(m \cap n) = \mathbb{P}(m|f)\mathbb{P}(f) + \mathbb{P}(m|n)\mathbb{P}(n) = 0.25 \cdot 0.4 + 0.07 \cdot 0.6 = 0.142$

Dati $A, B \in \mathcal{A}$, con $\mathbb{P}(A), \mathbb{P}(B) > 0$, si ha che

$$\mathbb{P}(B|A) := \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(A)} = \frac{\mathbb{P}(A|B)\mathbb{P}(B)}{\mathbb{P}(A|B)\mathbb{P}(B) + \mathbb{P}(A|B^c)\mathbb{P}(B^c)}$$

Questa viene detta formula di Bayes.

1.2.3 Formula delle Probabilità Totali

Data una partizione di Ω in insiemi disgiunti $(A_n)_{n\geq 1}$ con $\mathbb{P}(A_n) > 0$ si ha che $\forall B \in \mathcal{A}$

$$\mathbb{P}(B) = \mathbb{P}(B \cap \Omega) = \mathbb{P}(B \cap \left[\bigcup_n A_n\right]) = \mathbb{P}\left(\bigcup_n (B \cap A_n)\right) = \sum_{n\geq 1} \mathbb{P}(B \cap A_n) = \sum_{n \geq 1} \mathbb{P}(B|A_n) * \mathbb{P}(A_n)$$

La formula delle probabilità totali descrive come calcolare la probabilità di un evento B considerando tutte le possibilità degli eventi della partizione A_n e le loro intersezioni con B, pesate dalle probabilità degli eventi della partizione. In breve, è una tecnica per calcolare la probabilità di B considerando tutti i possibili scenari dati dalla partizione dell'insieme campionario.

Esempio - almeno due persone con lo stesso compleanno: Siano n il numero di individui e $D = 365$ il numero di giorni in un anno. La probabilità che nessuna coppia di persone abbia lo stesso compleanno (tutte date diverse) può essere rappresentata dalla formula
$\mathbb{P}(\text{compleanni diversi}) = \frac{D}{D} * \frac{D-1}{D} * \ldots * \frac{D-n+1}{D}$

La probabilità che almeno due individui abbiamo lo stesso compleanno è data quindi dal complementare di questa probabilità, ovvero:
$\mathbb{P}(\text{almeno un compleanno in comune}) = 1 - \mathbb{P}(\text{compleanni diversi}) = 1 - \frac{D}{D} * \frac{D-1}{D} * \ldots * \frac{D-n+1}{D}$

Ipotizzando un numero di persone $n = 23$, questa probabilità viene 0.507, quindi più del 50%. Alzando il numero $n = 50$, si ottiene una probabilità uguale a 0.971, quindi quasi del 100%.

1.2.4 Indipendenza

Due eventi $A, B \in \mathcal{A}$ si dicono indipendenti se $\mathbb{P}(A \cap B) = \mathbb{P}(A) * \mathbb{P}(B)$
Questo significa che l'occorrenza o la non-occorrenza di uno degli eventi non influisce sulla probabilità dell'altro evento.

2 Variabili Aleatorie Discrete

2.1 Variabili Aleatorie Discrete - nozioni base

Definizione: dato uno spazio di probabilità $(\Omega, \mathcal{A}, \mathbb{P})$, X di dice **Variabile Aleatoria** (Reale) se è una funzione $X : \Omega \to \mathbb{R}$, $\omega \to X(\omega)$ t.c. $\forall x \in \mathbb{R} \{\omega : X(\omega) \leq x\} \in \mathcal{A}$

Definizione: si dice **Variabile Aleatoria Discreta** una v.a. per cui $Im(X) = X(\Omega) = S$, con $|S| \leq |\mathbb{N}|$, ovvero un sottoinsieme numerabile di $\mathbb{N}$.
Definiamo quindi la **Funzione di Densità di Probabilità** $p_X(x) = \mathbb{P}(X = x) \neq 0 \iff x \in Im(X)$ ano

Osservazioni:

- $E_X = \{\omega : X(\omega) \leq x\} \in \mathcal{A}$
- $E_X^c = \{\omega : X(\omega) > x\} \in \mathcal{A}$
- $\{\omega : X(\omega) = x\} = \bigcap_{n\geq 1} \left\{\omega : x - \frac{1}{n} < X(\omega) \leq x\right\} = \bigcap_{n\geq 1}(E_X \cap E^c_{x-\frac{1}{n}}) \in \mathcal{A}$

Esempio - N lanci indipendenti di una moneta che dà testa con probabilità $\mathbb{P}$, probabilità di ottenere una data stringa di risultati:
$\Omega = \{\omega(\omega_1, \ldots, \omega_N), \omega_i \in \{1, 0\}\}$ con 1=testa e 0=croce, $|\Omega| = 2^N$
$A_i := \{\omega_i \in |omega : \omega_i = 1\} = \{\text{i-esimo lancio è testa} = 1\}$
$\mathbb{P}$(ottenere una data stringa) =?
La stringa sarà della forma $\bar{\omega}(T, \ldots, T, C, \ldots, C)$ con K teste e N-K croci, l'ordine non conta.
$\mathbb{P}(\bar{\omega}) = \mathbb{P}(\bigcap_{i=1}^K A_i \cap \bigcap_{j=K+1}^N A_i^c) = \prod_{i=1}^K \mathbb{P}(A_i) \prod_{j=K+1}^N \mathbb{P}(A_j^c) = p^K(1-p)^{N-K}$

2.2 V.A. Discrete Principali

2.2.1 Binomiale

Si dice Binomiale una v.a. $X \sim Bin(N, p)$ con $N =$ numero prove e $p =$ probabilità di successo per ognuna, per cui vale $X : \Omega \to \mathbb{R}$ $X(\omega) =$ numero di successi nella stringa $\omega = \sum_{i=1}^N \omega_i$.
La sua densità è definita come

$$p_X(k) := p(X = k) = p(A_k) = \sum_{\omega \in A_k} p(\omega) = |A_k|p^k(1-p)^{N-k} = \binom{N}{k} p^k(1-p)^{N-k}$$

Rappresentandola graficamente, otterremmo un istogramma simmetrico.

2.2.2 Bernoulli

Si dice di Bernoulli una v.a. $Y \sim Ber(p) = Bin(a, p)$.
La sua densità è definita come

$$p_Y(x) = \begin{cases} p & \text{se } x = 1 \\ 1-p & \text{se } x = 0 \\ 0 & \text{altrimenti} \end{cases}$$

Può quindi avere valori compresi in $S = \{0, 1\}$

2.2.3 Geometrica

Si dice Geometrica di parametro $p \in [0,1]$, con $T \in S = \mathbb{N}$ e $k \geq 1$, la variabile aleatoria $Y \sim Geom(p)$ di densità:

$$p_T(k) = \mathbb{P}(T = k) = (1-p)^{k-1}p$$

Può essere interpretata come il tempo necessario per ottenere il primo successo in prove indipendenti iterate k volte.
Osservazioni:

- $\sum_{k \geq 1} p_T(k) = 1$
- le v.a. geometriche godono di perdita di memoria. Infatti, essendo le prove indipendenti, $\mathbb{P}(T = m + k | T > m) = \mathbb{P}(T = k)$

2.2.4 Poisson

Si dice Poisson di parametro $\mu > 0$, con $Y \in S = \mathbb{N} \cup \{0\}$, la variabile aleatoria $Y \sim Poi(\mu)$ di densità:

$$p_Y(k) = e^{-\mu}\frac{\mu^k}{k!}, k \in \mathbb{N} \cup \{0\}$$

La v.a. di Poisson può anche essere vista come: $Poi(\mu) = \lim_{N \to \infty} Bin(N, \frac{\mu}{N})$

2.2.5 Funzione Indicatrice di un Evento

Dato uno spazio di probabilità $(\Omega, \mathcal{A}, \mathbb{P})$ e dato l'insieme $E \in \mathcal{A}$, si dice Funzione Indicatrice di un Evento la funzione $\mathbb{1}_E : \Omega \to \{0,1\} \subseteq \mathbb{R}$ la cui densità vale:

$$\mathbb{1}_E(\omega) = \begin{cases} 1 & \text{se } \omega \in E \\ 0 & \text{se } \omega \in E^c \end{cases}$$

2.3 Cambiamento di Variabile Aleatoria

Sia X una v.a. con densità p_X. Posso quindi definire un cambiamento di v.a. considerando $Y := \Phi(X)$, $\Phi : \mathbb{R} \to \mathbb{R}$

Esempio - 1:
$Y = \Phi(X) = aX + b$, con $a, b \in \mathbb{R}, a \neq 0$
$p_Y(y) = \mathbb{P}(aX + b = y) = \mathbb{P}(X = \frac{y-b}{a})$

Esempio - 2:
$\Omega : \Phi(X) = x^2 = y$
$\mathbb{P}_Y(y) = \mathbb{P}(x^2 = y) = \mathbb{P}(x = \sqrt{y}) + \mathbb{P}(X = -\sqrt{y}) = p_X(\sqrt{y}) + p_X(-\sqrt{y})$

2.4 Variabili Aleatorie Multidimensionali

Si dice v.a. Multidimensionale di dimensione d il vettore
$\vec{x} = (x_1, \ldots, x_d) : \Omega \to \mathbb{R}^d$ t.c. x_i sia una v.a. distinta $\forall i \leq d$.
Possiamo quindi definirne la Densità Congiunta: $p_{\vec{X}}(\vec{x}) := \mathbb{P}(X_1 = x_1 \cap \ldots \cap X_d = x_d)$

Le prossime cose le ha dette in una lezione successiva.
Funzione di Ripartizione Congiunta: $F_{\vec{X}}(\vec{x}) := \mathbb{P}(X_1 \leq x_1 \cap \ldots \cap X_d \leq x_d) \in [0,1]$
Con $\vec{x}$ discreta avremo $F_{\vec{X}}(\vec{x}) = \sum_{\vec{y} \in \mathbb{R}^d, \vec{y}_i \leq \vec{x}_i} p_{\vec{X}}(\vec{y})$, con $p_{\vec{X}}(\vec{y}) = \mathbb{P}(X_1 = y_1 \cap \ldots \cap X_d = y_d)$

2.4.1 Densità Marginale e Indipendenza

Date due v.a. (X, Y) con densità congiunta $p_{X,Y}(x, y) = \mathbb{P}(X = x \cap Y = y)$, si dice Densità Marginale la formula
$p_X(x) = \mathbb{P}(X = x) = \sum_{y \in Im(Y)} \mathbb{P}(X = x, Y = y) = \sum_y p_{(X,Y)}(x, y)$
Le v.a. X e Y si diranno quindi indipendenti se $\forall X, Y \in \mathbb{R}^2$ si ha $p_{(X,Y)}(x, y) = p_X(x) * p_Y(y)$

2.5 Limiti della Funzione di Distribuzione

Dato uno s. di probabilità $(\Omega, \mathcal{A}, \mathbb{P})$, data una v.a. $x : \Omega \to \mathbb{R}$ t.c. $\forall x \in \mathbb{R}\, \{\omega : X(\omega) \leq X\} \in \mathcal{A}$, data $F_X(x) := \mathbb{P}(\{\omega : X(\omega) \leq x\}) \in [0, 1]$, avremo quindi:

- $\lim_{x \to +\infty} F_X(x) = \mathbb{P}(X \in \mathbb{R}) = 1$
- $\lim_{x \to -\infty} F_X(x) = 0$

2.6 Densità Discreta

Per v.a. discrete si può definire la Densità Discreta come:
$p_X(x) = \mathbb{P}(X = x) \neq 0 \iff x \in \mathcal{S}$ $F_X(x) = \mathbb{P}(X \leq x) = \sum_{y \in \mathcal{S}, y \leq x} p_X(y)$

3 Variabili Aleatorie Continue

ATTENZIONE: in questo capitolo vengono trattate parallelamente v.a. discrete e continue!

Definiamo Variabile Aleatoria (Assolutamente) Continua una X v.a. reale t.c.
$F_X(x) = \int_{-\infty}^{x} f_X(y)dy, \quad$ con $f_X(x) \geq 0$ e $\int_{\mathbb{R}} f_X(x)dx = 1$

3.1 Variabili Aleatorie Multidimensionali Assolutamente Continue

Definiamo una v.a. multidimensionale assolutamente continua come:
$\vec{X} : \Omega \to \mathbb{R}^d, \quad F_{\vec{X}}(\vec{x}) = \int_{-\infty}^{x_1} dy_1 \int \dots \int_{-\infty}^{x_d} dy_d f_{\vec{x}}(\vec{y}) \quad$ con $\vec{x} = (x_1, \dots, x_d)$
$\mathbb{P}(\vec{X} \in A_{\vec{X}}) = \int_{A_{\vec{X}}} f_{\vec{X}}(\vec{y})dy, \quad$ con $A_{\vec{X}} := \{\vec{y} = (y_1, \dots, y_d) : y_i \leq x_i, i = 1, \dots, d\}$

La densità congiunta sarà quindi $f_{\vec{X}} : \mathbb{R}^d \to \mathbb{R}^+$ tale che $\int_{\mathbb{R}^d} f_{\vec{X}}(\vec{x})\, d\vec{x} = 1$

3.2 Distribuzioni Marginali a partire dalla Congiunta

$F_{X_1}(x_1) = \mathbb{P}(X_1 \leq x_1) = \mathbb{P}(X_1 \leq x_1, X_2 \leq +\infty, \dots, X_d \leq +\infty)$

Caso Discreto:
$F_{X_1}(x_1) = \sum_{y_2,\dots,y_d} p_{\vec{X}}(x_1, y_2, \dots, y_d)$

Caso Continuo:
$F_{X_1}(x_1) = \int_{\mathbb{R}^{d-1}} dx_2 \, \dots \, dx_d \, f_{\vec{X}}(\vec{x})$

3.3 Indipendenza di Variabili Aleatorie

Due v.a. X, Y si dicono indipendenti se $F_{X,Y}(x, y) = \mathbb{P}(X \leq x \cap Y \leq y) = \mathbb{P}(X \leq x) * \mathbb{P}(Y \leq y)$

Caso Discreto:
$p_{\vec{X}}(\vec{x}) = \prod_{i=1}^{d} p_{X_i}(x_i)$

Caso Continuo:
$f_{\vec{X}}(\vec{x}) = \prod_{i=1}^{d} f_{X_i}(x_i)$

Esempio - date due V.A. X e Y reali e indipendenti, calcolare la legge di $Z = \phi(X, Y) = min(X, Y)$:
$F_Z(z) = \mathbb{P}(Z \leq z) = \mathbb{P}(min(X, Y) \leq z) = 1 - \mathbb{P}(min(X, Y) > z) = 1 - \mathbb{P}(X > z \cap Y > z) = 1 - \mathbb{P}(X > z) * \mathbb{P}(Y > z) = 1 - [1 - F_X(z)]\, [1 - F_Y(z)]$

3.4 Probabilità di un Intervallo per Variabili Aleatorie

$a < b \quad \mathbb{P}(X \in (a, b]) = F_X(b) - F_X(a)$

Caso Discreto:
$\mathbb{P}(X \in (a, b]) = \sum_{y \in S, y \in (a,b]} p_X(y)$

Caso Continuo:
$\mathbb{P}(X \in (a, b]) = \int_a^b f_X(x)\, dx$

Osservazione: nel caso continuo
$\mathbb{P}(X \in (a, b)) = \mathbb{P}(X \in (a, b]) = \mathbb{P}(X \in [a, b)) = \mathbb{P}(X \in [a, b]) = \int_a^b f_X(x)\, dx$

3.5 Probabilità di un Punto per Variabili Aleatorie

Data una v.a. continua X fissata si ha $\mathbb{P}(X = x) = \mathbb{P}(\bigcap_{n>1} \{X \in (x - \frac{1}{n}, x]\}) = \lim_{n\to\infty} \mathbb{P}(X \in (x - \frac{1}{n}, x]) = \lim_{n\to\infty} F_X(x) - F_X(x - \frac{1}{n})$

Caso Discreto:
$\mathbb{P}(X = x) = p_X(x)$

Caso Continuo:
$\mathbb{P}(X = x) = 0$

Esempi:

- $X \sim Unif(0,1)$
  $f_X(x) = \mathbb{1}\{x \in [0,1]\} \geq 0 \quad \int_{\mathbb{R}} f_X(x)\,dx = \int_0^1 dx = 1$
  $|I| = |I'|, \quad I, I' \subseteq [0,1] \quad \mathbb{P}(X \in I) = \int_I dx = |I| = |I'| = \mathbb{P}(X \in I')$
- $X \sim Unif(a,b)$
  $f_X(x) = \frac{\mathbb{1}\{x\in[a,b]\}}{b-a} \geq 0$
- $X \sim exp(\lambda)$ tempo di 1^a attesa
  $f_X(x) = \lambda e^{-\lambda x} * \mathbb{1}_{[x\geq 1]} = \lambda \int_0^\infty e^{-\lambda x}\,dx = \left[-e^{-\lambda x}\right]_0^\infty = 1$
  $\int_{-\infty}^x f_X(y)\,dy = \ldots = 1 - e^{-\lambda x}$
  In questo caso abbiamo perdita di memoria in quanto $\mathbb{P}(X > t+s | x > t) = \mathbb{P}(x > s)$

3.6 Variabile Aleatoria Gaussiana Standard

Si dice Gaussiana standard (o normale) una v.a. $X \sim Gauss$ la cui densità di probabilità è definita come:

$$f_X(x) = \frac{e^{-\frac{x^2}{2}}}{\sqrt{2\pi}}$$

Ne segue quindi che $\int_{\mathbb{R}} f_X(x) = 1$

3.7 Legge della Somma di Variabili Aleatorie

Siano X e Y v.a. e sia $\Phi(X,Y) = X + Y = Z$,
allora $F_Z(z) = \mathbb{P}(Z \leq z) = \mathbb{P}(X+Y \leq z) = \mathbb{P}((X,Y) \in A_Z)$, con $A_Z := \left\{(X,Y) \in \mathbb{R}^2 t.c. X + Y \leq z\right\}$.

Caso Continuo:
$\int_{A_Z} f_{(X,Y)}(x,y)\,dx\,dy = \int_{\mathbb{R}} dx \int_{-\infty}^{z-x} dy\, f_{(X,Y)}(x,y)$
Ponendo $u = y + x$ l'integrale diventa $\int_{\mathbb{R}} dx \int_{-\infty}^{z} du\, f_{(X,Y)}(x, u-x)$
Nel caso X e Y fossero due v.a. indipendenti, allora si ha che $f_{(X,Y)}(x,y) = f_X(x)\,f_Y(y)$, quindi l'integrale diventa $= \int_{-\infty}^z du \int_{\mathbb{R}} dx\, f_X(x) f_Y(y)$
Si ha anche che $f_Z(z) = \frac{d}{dz} F_Z(z) = \int_{\mathbb{R}} dx\, f_X(x) f_Y(z-x) = (f_X * f_Y)(z)$ ovvero la convoluzione.

Caso Discreto:
$p_Z(z) = \mathbb{P}(Z = z) = \sum_K p_X(K) p_Y(z-K)$

3.8 Media Matematica di una Variabile Aleatoria Reale X

Sapendo che $F_X(x) = \mathbb{P}(X \leq x) =$
Caso Continuo: $= \int_{-\infty}^x f_X(y)\,dy$ Caso Discreto: $\sum_{y\in Im(X),\, y\leq x} p_X(y)$
Si chiama Media (o Aspettazione, o Speranza) Matematica di una v.a. reale X la formula:
$\mathbb{E}[X] :=$ Caso Continuo: $\int_{\mathbb{R}} dx\, x\, f_X(x)$ Caso Discreto: $\sum_{K\in Im(X)} K\, p_X(K)$
Corrisponde al baricentro di massa.
Esempi:

- $T \sim Exp(\lambda), \quad f_T(t) = \lambda e^{-\lambda t} \quad \Rightarrow \quad \mathbb{E}[T] = \int_0^{+\infty} t \lambda e^{-\lambda t}\,dt = \frac{1}{\lambda}$
- $X \sim Bin(p, N) \quad \Rightarrow \quad \mathbb{E}[X] = \sum_{K=0}^N K \binom{K}{N} p^K (1-p)^{N-K} = \ldots = N\,p$
  $\Phi : \mathbb{R}^d \to \mathbb{R}, \quad \vec{X} : \Omega \to \mathbb{R}^2, \quad Z = \Phi(\vec{X})$

Teorema: $\mathbb{E}[Z] = \mathbb{E}\left[\Phi(\vec{X})\right] = \int_{\mathbb{R}^d} d\vec{x}\, \Phi(\vec{x})\, f_{\vec{X}}(x_1, \ldots, x_d)$

Teorema: date X, Y, a, b, c avremo che
$\mathbb{E}[aX + bY + c] = a\mathbb{E}[X] + b\mathbb{E}[Y] + c$
ne segue quindi che la media di una v.a. è un operatore lineare.

3.9 Covarianza tra Due Variabili Aleatorie

Si dice covarianza tra due v.a. X e Y la formula
$Cov(X,Y) := \mathbb{E}[(X - \mathbb{E}[X])(Y - \mathbb{E}[Y])] = \mathbb{E}[XY - X\mathbb{E}[Y] - \mathbb{E}[X]Y + \mathbb{E}[X]\mathbb{E}[Y]] = \ldots = \mathbb{E}[XY] - \mathbb{E}[X]\mathbb{E}[Y]$

3.10 Varianza di Una Variabile Aleatoria

Si dice varianza di una v.a. la covarianza della v.a. con se stessa, in formula:
$Var(X) = Cov(X,X) = \mathbb{E}\left[|X - \mathbb{E}[X]|^2\right]$
La varianza dice quanto una v.a. fluttua intorno alla sua media, non è un operatore lineare, infatti:

- $Var(aX) = a^2 Var(X) = Var(aX + c)$
- $Var(X + Y) = Var(X) + Var(Y) + 2Cov(X,Y) \quad$ se indipendenti $= Var(X) + Var(Y)$

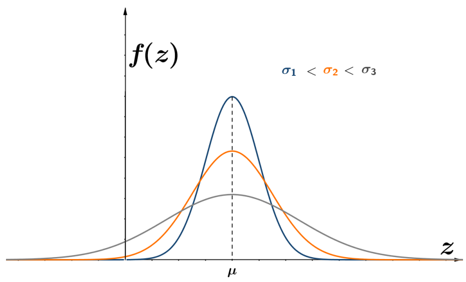
Osservazione: X, Y indip. $\Rightarrow Cov(X,Y) = 0$, invece $Cov(X,Y) = 0 \Rightarrow X, Y$ indip.

3.11 Normale Gaussiana Standard

Si dice Normale Gaussiana Stanrdard la v.a. $X \sim N(0,1)$ tale che $f_X(x) = \frac{e^{-\frac{x^2}{2}}}{\sqrt{2\pi}}$.
Notiamo anche che $\int_{\mathbb{R}} \frac{e^{-\frac{x^2}{2}}}{\sqrt{2\pi}}\, dx = 1$ e che $F_X(x) := \mathbb{P}(X \leq x) = \int_{-\infty}^{x} \frac{e^{-\frac{x^2}{2}}}{\sqrt{2\pi}}\, dx = \Phi(x)$
Essendo una funzione pari avremo anche che $\Phi(-x) = 1 - \Phi(x)$.

Esempio - Gaussiana Standard con parametro σ:

Data $X \sim N(0,1)$, data la v.a. $Y = \sigma X + \mu$, con $\mu \in \mathbb{R}, \sigma > 0$, allora varrà che $Y \sim N(\mu, \sigma^2)$.
Possiamo quindi calcolare $f_Y(y) = \frac{d}{dx}\mathbb{P}(\sigma X + \mu \leq y) = \frac{d}{dy}\mathbb{P}(X \leq \frac{y-\mu}{\sigma}) = \frac{d}{dy}F_X(\frac{y-\mu}{\sigma}) = \frac{1}{\sigma}f_X(\frac{y-\mu}{\sigma}) = \frac{e^{-\frac{(y-\mu)^2}{2\sigma^2}}}{\sqrt{2\pi\sigma^2}}$



Osservazioni:

- La media di una Gaussiana Standard è 0, e la sua varianza è 1
  $X \sim N(0,1) \quad \mathbb{E}[X] = 0 \quad Var(X) = 1$
- Se la Gaussiana fosse traslata, ovvero $Y = \sigma X + \mu$, allora avremo
  $Y \sim N(\mu, \sigma^2) \quad \mathbb{E}[Y] = \mu \quad Var(Y) = \sigma^2$

4 Legge dei Grandi Numeri (L.L.N.) e Teorema del Limite Centrale (C.L.T.)

4.1 Legge dei Grandi Numeri

Data la successione $(X_n)_{n\geq 1}$ di eventi indipendenti e con stessa distribuzione, avremo che $\mathbb{E}[X_i] = \mu < \infty$ e che $Var(X_i) = \sigma^2 < \infty$. Avremo quindi che

$$\frac{S_N}{N} = \frac{\sum_{i=1}^{N} X_1}{N} \xrightarrow[N\geq\infty]{p} \mathbb{E}\left[\frac{S_N}{N}\right] = \frac{1}{N}\sum_{i=1}^{N}\mathbb{E}\left[X_i\right] = \mu = \mathbb{E}\left[X_1\right], \text{ quindi } \frac{S_N}{N} \approx \mu$$

Questo risultato ci dice che, quando la dimensione del campione N tende a infinito, la media empirica converge alla media teorica. In formule: $\frac{S_N}{N} \approx \mu$ per $N \to \infty$.
Osservazione: la scrittura $\xrightarrow[N\geq\infty]{p}$ significa che "converge in probabilità", ovvero che
$\mathbb{P}(|X_n - X| > \varepsilon) \xrightarrow[N\to\infty]{} 0$

4.2 Teorema del Limite Centrale

Data la successione $(X_n)_{n\geq 1}$ di eventi indipendenti e con stessa distribuzione, avremo che $\mathbb{E}[X_i] = \mu < \infty$ e che $Var(X_i) = \sigma^2 < \infty$. Avremo quindi che

$$\frac{S_N - \mathbb{E}\left[S_N\right]}{\sigma\sqrt{N}} \xRightarrow{N\to\infty} N(0,1)$$

Il teorema afferma che, nel caso in cui sommiamo un elevato numero di v.a. indipendenti e identicamente distribuite, la standardizzazione della somma converge in distribuzione a una variabile Normale Gaussiana Standard. Questa convergenza avviene indipendentemente dalla distribuzione originale delle variabili casuali X_i.
Esempio: la vita di una lampadina è espressa come una v.a.
$T_i \sim exp(\lambda)$, con $\lambda = 0.1, Var[T_i] = 10, \mathbb{E}[T_i] = 10$.
Qual è la probabilità che 40 lampadine bastino per un anno?
$\mathbb{P}(S_{40} = \sum_{i=1}^{40} Ti \geq 365) =$ (per il C.L.T.) $\mathbb{P}(S_{40} - \mathbb{E}(S_40) \geq 365 - 400) = \mathbb{P}(\frac{S_{40}-\mathbb{E}(S_{40})}{\sqrt{400}} - \frac{365-400}{\sqrt{400}} \approx \mathbb{P}(N(0,1) \geq -0.55) = 1 - \Phi(-0.55)$