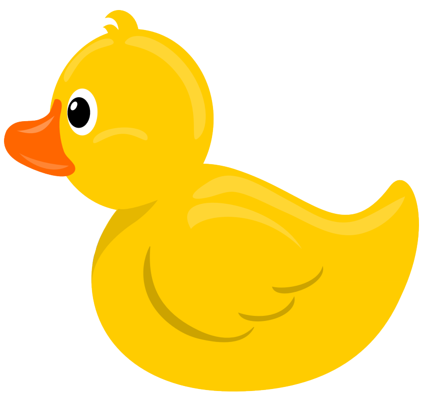
Probabilità e Statistica per l'Informatica

Ver 1.0

Falbo Andrea

A.A 2022/2023
Prof. Caravenna
LilQuacky



Indice

Capitolo 1

Statistica Descrittiva

1.1 Introduzione

Statistica Descrittiva, Statistica Inferenziale e Probabilità

La statistica è l'arte di imparare dai dati. Possiamo suddividerla in:

- **Statistica Descrittiva**: *Descrive* e riassume i dati
- **Statistica Inferenziale**: *Trae* conclusioni dai dati

Per poter trarre conclusioni dai dati, bisogna tenere conto del ruolo che gioca il caso. Definiamo quindi la **probabilità** come la descrizione matematica di eventi *casuali*.

1.2 Descrivere i dati

1.2.1 Frequenze, Istogrammi, Classi

Frequenza Assoluta e Relativa

Misuriamo una certa variabile in un campione, ottenendo un insieme di dati. Se i dati non sono distinti, ovvero *abbiamo ripetizioni*, possiamo riassumerli in una **tabella delle frequenze**. Possiamo definire dunque:

- **Frequenza Assoluta** f_i := numero di volte in cui i compare nel campione di dati.
- **Frequenza Relativa** $p_i := \frac{f_i}{N}$ = numero di volte in cui compare i rispetto al totale.

Esempio: Marta intervista i suoi N = 20 compagni di classe e chiede la squadra di calcio preferita, ottenendo le risposte:
Juve, Milan, Inter, Atalanta, Juve, Milan, Nessuna, Nessuna, Inter, Milan, Juve, Nessuna, Atalanta, Juve, Nessuna, Milan, Inter, Milan, Nessuna, Nessuna.

Squadre di Calcio		
Valori	Frequenze assolute	Frequenze relative
Juve	4	$4/20 = 0.20\%$
Milan	5	$5/20 = 0.25\%$
Inter	3	$3/20 = 0.15\%$
Nessuna	6	$6/20 = 0.30\%$
Atalanta	2	$2/20 = 0.10\%$

Istogramma

É utile rappresentare le frequenze mediante un grafico a barre detto **istogramma**.

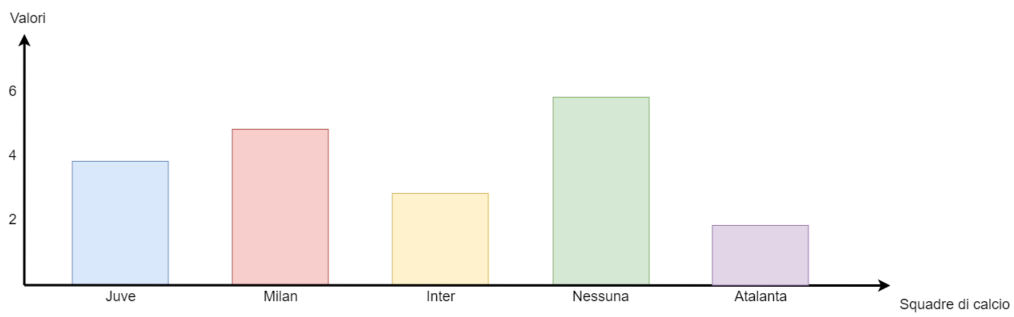


Figura 1.1: Istogramma sulle frequenze assolute delle squadre di calcio

Classi

Può capitare di avere insiemi di dati che assumono un numero elevato di valori distinti. In tal caso conviene *suddividere* i valori assunti in intervalli detti **classi**.

Esempio: Vengono misurati i livelli di colesterolo nel sangue di un insieme di N = 40 individui:

213 174 193 796 220 183 194 200 192 200 200 199 178 183 188 193 187 181 193 205 196 211 202 213 216 206 195 191 171 194 184 191 221 212 221 204 204 191 183 227

Molti valori sono distinti e dunque hanno $f_i = 1$. Scegliamo le classi:

$$[170, 180) \quad [180, 190) \quad [190, 200) \quad [200, 210) \quad [210, 220) \quad [220, 230)$$

Livelli di colesterolo		
Valori	Frequenze assolute	Frequenze relative
$[170, 180)$	3	$3/40 = 7.5\%$
$[180, 190)$	7	$7/40 = 17.5\%$
$[190, 200)$	13	$13/40 = 32.5\%$
$[200, 210)$	8	$8/40 = 20\%$
$[210, 220)$	5	$5/40 = 12.5\%$
$[220, 230)$	4	$4/40 = 10\%$

1.2.2 Dati Bivariati

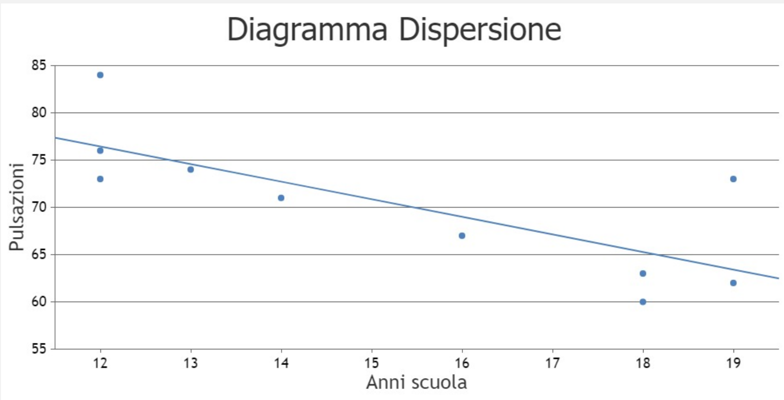
I **dati bivariati** ci permettono di mostrare *due variabili* per ogni singolo elemento dell'insieme. Per un elemento i, indichiamo con x_i la prima variabile e con y_i la seconda.

Diagramma Di Dispersione

Un **diagramma di dispersione** rappresenta i punti $(x_i,\ y_i)$, in modo da evidenziare una possibile *correlazione* tra i due valori.

Esempio: Rileviamo il numero di anni di scuola (prima variabile) e le pulsazioni a riposo (seconda variabile) in un campione di N=10 individui. I dati $(xi\ ,\ yi)$ sono

(12,73) (16,67) (13,74) (18,63) (19,73) (12,84) (18,60) (19,62) (12,76) (14,71)



Possiamo evidenziare una correlazione negativa tra i due valori.

1.3 Riassumere i dati

Una **statistica** campionaria riassume l'insieme di dati mediante una quantità numerica.

1.3.1 Indici di Posizione

Un **indice di posizione** *sintetizza la posizione* di una distribuzione sostituendo l'insieme dei dati con un unico valore tale da fornire una rappresentazione globale.

Media Campionaria

Per definire il valore medio dell'insieme dei dati, definiamo la **media campionaria** come

$$\overline{x} := \frac{x_1 + x_2 + \cdots + x_N}{N} = \frac{1}{N}\sum_{i=1}^{n} x_i$$

In generale, sapendo le frequenze assolute dei valori, possiamo scrivere la media campionaria come

$$\overline{x} = \frac{z_1 \cdot f_1 + z_2 \cdot f_2 + \cdots + z_M \cdot f_M}{N}$$

con $N = f_1 + f_2 + \cdots + f_M$, dove z_i sono i valori e f_i sono le frequenze assolute associate.

Osservazione: La media è lineare: $\overline{y} = a \cdot \overline{x} + b$

Mediana Campionaria

Un'altra misura del centro dell'insieme dei dati è la **mediana campionaria**, ovvero quel valore che, ordinati i dati, si trova in *posizione centrale*. In base al numero di dati, si calcola in due modi:

- **N dispari**: la mediana è il dato di posto $\frac{N+1}{2}$:

$$m := x_{\left(\frac{N+1}{2}\right)}$$

- **N pari**: la mediana è la media aritmetica tra il dato di posto $\frac{N}{2}$ e quello di posto $\frac{N}{2} + 1$:

$$m := \frac{x_{\left(\frac{N}{2}\right)} + x_{\left(\frac{N}{2}+1\right)}}{2}$$

Percentile Campionario

Fissiamo un numero $k \in [0, 100]$. Definiamo il **k-esimo percentile campionario** il valore t per cui:

- almeno il $k\%$ dei dati è $\leq$ t
- almeno il $(100 - k)\%$ dei dati è $\geq$ t

Quartili

I casi più importanti sono:

- **Primo Quartile** q_1: $p = \frac{1}{4}$, $k = 100p$
- **Secondo Quartile** q_2: $p = \frac{1}{2}$, $k = 100p$
- **Terzo Quartile** q_3: $p = \frac{3}{4}$, $k = 100p$

Come per la mediana, si calcola in due modi:

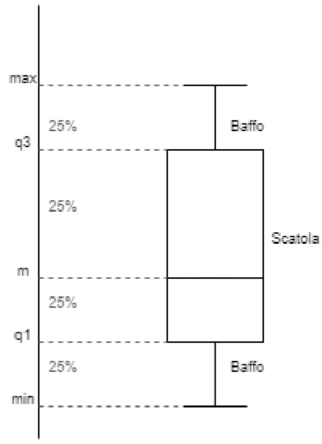
- Se $N \cdot p$ non è intero, $t = x_{(i)}$ è il dato di posizione i definito come l'intero successivo a $N \cdot p$
- Se $N \cdot p$ è intero, $t = \frac{x_{(N \cdot p)} + x_{(N \cdot p + 1)}}{2}$

Esempio: Ai 1000 abitanti di un piccolo comune viene chiesto di esprimere un giudizio su un nuovo servizio comunale, usando una scala da 0 a 4. 251 persone hanno votato 0, 260 persone hanno votato 1, 80 persone hanno votato 2, 154 persone hyanno votato 3 mentre 255 persone hanno votato 4
Vogliamo calcolare i 3 indici di posizione: Media, Mediana, Quartili.

- **Media**: $\overline{x} = \frac{0 \cdot 251 + 1 \cdot 260 + 2 \cdot 80 + 3 \cdot 154 + 4 \cdot 255}{1000} = 1,902 \simeq 1,9$
- **Mediana**: $m = \frac{x_{(500)} + x_{(501)}}{2} = \frac{1+1}{2} = 1$
- **Primo quartile**: $1000 \cdot \frac{1}{4} = 250$, $q_1 = \frac{x_2 50 + x_2 51}{2} = \frac{0+0}{2} = 0$
- **Secondo quartile**: mediana $= 1$
- **Terzo quartile**: $1000 \cdot \frac{3}{4} = 750$, $q_3 = \frac{x_{(750)} + x_{(751)}}{2} = \frac{4+4}{2} = 4$

Box Plot

Una rappresentazione grafica di mediana e quartili viene fornita dal **box plot**:



1.3.2 Indici di Dispersione

Un **indice di dispersione** descrive la *variabilità di distribuzione* quantitativa dei dati, ovvero quanto i valori presenti distano da un valore centrale.

Scarti

Gli **scarti** sono la distanza di ogni singolo elemento dal valore medio. La somma di tutti gli scarti è nulla.

$$w := (x_i - \overline{x})$$

Varianza Campionaria

Considerando gli scarti elevati al quadrato e facendone una sorta di media, otteniamo la **varianza campionaria** come:

$$s^2 := \frac{1}{N-1} \sum_{i=1}^{N} (x_i - \overline{x})^2$$

Deviazione Standard

Per ottenere una statistica omogenea ai dati, si definisce **deviazione standard** come:

$$s := \sqrt{s^2} = \sqrt{\frac{1}{N-1}\sum_{i=1}^{N}(x_i - \overline{x})^2}$$

Scarto Interquartile

Un altro indicatore di dispersione rispetto alla mediana m è lo **scarto interquartile**. Per costruzione, questo intervallo contiene almeno il 50% dei dati. Il suo valore è:

$$IQR = \Delta := q_3 - q_1$$

Osservazione: La varianza e la deviazione standard non sono lineari.

1.3.3 Correlazione

Coefficiente di Correlazione Lineare

Consideriamo un insieme di dati bivariati. Vogliamo quantificare la correlazione tra le due variabili x e y, ossia la *tendenza* per cui a valori di x grandi corrispondo valori di y grandi o piccoli. Definiamo quindi il **coefficiente di correlazione lineare**:

$$r = \frac{\sum_{i=1}^{N}(x_i - \overline{x})(y_i - \overline{y})}{(N-1)\cdot s_x \cdot s_y} \quad \text{oppure} \quad r = \frac{\sum_{i=1}^{N} x_i \cdot y_i - N \cdot \overline{x} \cdot \overline{y}}{(N-1)\cdot s_x \cdot s_y}$$

Il coefficiente di correlazione lineare può assumere valori $\in [-1, 1]$. Nello specifico diremo che:

· $|r| \gtrsim 0,7$ avremo correlazione *significativa*

· $|r| \lesssim 0,3$ avremo correlazione *debole*.

Esercizio Completo: Prendiamo un campione di N = 10 elementi e assegnamo x = numero anni di scuola e y = pulsazioni. Otteniamo i seguenti valori

(12, 73) (16, 67) (13, 74) (18, 63) (19, 73) (12, 84) (18, 60) (19, 62) (12, 76) (14, 71)

Calcoliamo $\overline{x}, m, s^2, s, r$.

· $\overline{x} = \frac{12+16+13+18+19+19+12+18+19+12+14}{10} = 15,3$

· $s_x^2 = \frac{(12+16+13+18+19+19+12+18+19+12+14)^2-10\cdot(15,3)^2}{9} \simeq 9,12$

· $s_x = \sqrt{s_x^2} \simeq 3,02$

· $\overline{y} = \frac{73+67+74+63+73+84+60+62+76+71}{10} = 70,3$

· $s_y^2 = \frac{(73+67+74+63+73+84+60+62+76+71)^2-10\cdot(70,3)^2}{9} \simeq 54,23$

· $s_y = \sqrt{s_y^2} \simeq 7,36$

· $\sum_{i=1}^{N} x_i \cdot y_i = 12 \cdot 73 + 16 \cdot 67 + 13 \cdot 74 + \cdots + 14 \cdot 71 = 10603$

· $r = \frac{10603-10\cdot 15,3\cdot 70,3}{9\cdot 3,02\cdot 7,36} \simeq -0,76$

Il valore di r conferma che x e y mostrano una significativa correlazione negativa

Capitolo 2

Spazi di Probabilità

2.1 Introduzione

Il **calcolo delle probabilità** è una teoria matematica che permette di descrivere gli *esperimenti aleatori*, ovvero fenomeni il cui esito non è prevedibile con certezza a priori.

2.2 Assiomi della Probabilità

Spazio Campionario, Evento, Probabilità

La descrizione matematica di esperimento aleatorio si descrive in 3 passi:

1. **Spazio Campionario**: Insieme Ω che contiene tutti i possibili esiti dell'esperimento.

2. **Eventi**: Affermazioni sull'esito dell'esperimento aleatorio. $A \subseteq \Omega$.

3. **Probabilità**: Funzione P che associa a ogni evento $A \subseteq \Omega$ un valore $\in [0, 1]$ che soddisfa opportune proprietà. Ci sono almeno due interpretazioni su che cos'è P(A):

    · **Soggettivista**: P(A) = prezzo equo di una scommessa che paga 1 se si verifica A, altrimenti 0.

    · **Frequentista**: P(A) = frazione asintotica di volte in cui si verifica A ripetendo l'esperimento.

In ogni caso la probabilità deve soddisfare due proprietà:

- $P(\Omega) = 1$
- Se A e B sono eventi disgiunti, cioè $A \cap B = \emptyset$, allora:

$$P(A \cup B) = P(A) + P(B)$$

Esempio: Se prendiamo in considerazione un dado a sei facce, avremo:

- Spazio Campionario: $\Omega = \{1, 2, 3, 4, 5, 6\}$
- Evento: A = esce un numero pari = {2,4,6}
- Probabilità di A= 0.5 = 50%

La coppia (Ω, P) è detta **spazio di probabilità**.

Indichiamo con $|A|$ la cardinalità di un insieme A. La **probabilità uniforme** su un insieme finito Ω si definisce come:

$$P(A) := \frac{|A|}{|\Omega|}$$

Abbiamo dunque che per ogni $w \in \Omega$, $P(w) = \frac{1}{|\Omega|}$

Proprietà di base

Fissiamo uno spazio di probabilità (Ω, P). Valgono le seguenti proprietà:

1. **Insieme Vuoto**: $P(\emptyset) = 0$
2. **Regola del complementare**: $P(A^C) = 1 - P(A)$
3. **Regola dell'addizione di probabilità**: $P(A \cup B) = P(A) + P(B) - P(A \cap B)$
4. **Monotonia**: se $A \subseteq B$ allora $P(A) \leq P(B)$

2.3 Calcolo Combinatorio

In uno spazio di probabilità uniforme, calcolare una probabilità significa contare gli elementi di un insieme. Contare è un problema non banale per insiemi grandi. Le tecniche di conteggio formano il **Calcolo combinatorio**.

Principio fondamentale: Consideriamo un esperimento costituito da due parti:

1. prima parte: n esiti possibili
2. seconda parte: m esiti possibili

Allora l'esperimento totale può avere $n \cdot m$ esiti possibili.

Esempio: Se lancio due dadi a 6 facce ho $\Omega = 6^2 = 36$ esiti possibili, se lancio tre dadi a 6 facce ho $\Omega = 6^3 = 216$ esiti possibili ecc.

2.3.1 Disposizioni con Ripetizione

Le **disposizioni con ripetizioni** sono sequenze ordinate di k elementi, *anche ripetuti*, scelti tra n possibili. Sono in numero

$$n^k$$

Esempio: Estraggo casualmente k persone: qual è la probabilità che siano nate tutte e k in primavera?
$|\Omega| = 365^k$
A = "Tutti nati in primavera " = "20 marzo - 20 giugno" = 92 giorni
$|A| = 92^k$
P(A) = $(\frac{92}{365})^k \simeq \frac{1}{4^k}$

2.3.2 Disposizioni Semplici

Le **disposizioni semplici** sono sequenze ordinate di k elementi **distinti** scelti tra n possibili. Sono in numero

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

Osservazione: Se n=k si parla di permutazioni, in questo caso sono in numero n!

Esempio: Estraggo casualmente k persone: qual è la probabilità che almeno due abbiano lo stesso compleanno?
$|\Omega| = 365^k$
A = "almeno due persone hanno lo stesso compleanno"
A^C = "tutti hanno compleanni distinti"
$P(A^C) = 365 \cdot 364 \cdots (365 - k + 1)$
$P(A) = 1 - P(A^C) = 1 - \frac{|A^C|}{|\Omega|} = 1 - \frac{365 \cdot 364 \cdots (365-k+1)}{365^k} = 1 - \frac{365}{365} \cdot \frac{364}{365} \cdots \frac{(365-k+1)}{365}$

- con K=10 persone avremo il 12% di trovare due persone con lo stesso compleanno.
- con K=23 persone avremo il 50% di trovare due persone con lo stesso compleanno.
- con K=50 persone avremo il 97% di trovare due persone con lo stesso compleanno.

2.3.3 Combinazioni

Fino ad ora abbiamo considerato l'ordine importante. Ad esempio, nel lancio di due dadi, (2,5) $\neq$ (5,2). Le combinazioni si possono ottenere dalle disposizioni semplici dimenticando l'ordine degli elementi.
Le **combinazioni** sono collezioni *non ordinate* di k elementi distinti scelti tra n possibili. Sono in numero

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

Esempio: In una mano a Poker un giocatore riceve 5 carte su un mazzo di 52. Le possibili combinazioni sono quindi

$$\binom{52}{5} = \frac{52!}{5! \cdot 47!} = 2.598.960$$

2.4 Probabilità Condizionata

Consideriamo uno spazio di probabilità (Ω, P). Consideriamo un evento $A \subseteq \Omega$ con probabilità P(A). Supponiamo di ricevere informazioni su un evento B che si è verificato.
La **probabilità condizionata** è l'aggiornamento della probabilità di A dopo un *informazione aggiuntiva*. la probabilità condizionata di A dato B si scrive come

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

Esempio: Qual è la probabilità che la somma di due dadi a 6 facce valga 4, sapendo che il primo dado vale 2?
$|\Omega| = 6^2 = 36$
A = "la somma vale 4" = (1,3) (2,2) (3,1), $|A| = 3$
B = "il primo vale 2" = (2,1), (2,2), (2,3), (2,4), (2,5), (2,6), $|B| = 6$

$$P(A|B) = \frac{A \cap B}{B} = \frac{\frac{|A \cap B|}{|\Omega|}}{\frac{|B|}{|\Omega|}} = \frac{|A \cap B|}{|B|} = \frac{1}{6} \simeq 16,7\%$$

Proprietà

· **Regola del prodotto**: $P(A \cap B) = P(B) \cdot P(A|B)$

· **Formula di disintegrazione**: $P(A) = P(A \cap B) + P(A \cap B^C)$

· **Formula delle probabilità totali**: $P(A) = P(A|B) \cdot P(B) + P(A|B^C) \cdot P(B^C)$

· **Secondo elemento fissato**: $P(*|B)$ è una probabilità: $P(A^C|B) = 1 - P(A|B)$

· **Formula di Bayes**: $P(B|A) = \frac{P(A|B) \cdot P(B)}{P(A)}$

Esempio: Per rilevare la presenza di un virus viene effettuato un test con le seguenti caratteristiche:

- Sensibilità: se il virus è presente, il test dà esito positivo al 99%
- Specificità: se il virus è assente, il test dà esito negativo al 99.7%
- Prevalenza: è noto che 4 persone su 1000 hanno il virus.

Estraendo un cittadino a caso, qual è la probabilità che l'esito sia positivo? E qual è la probabilità che abbia effettivamente il virus se l'esito è positivo?
Introduciamo gli eventi e le relative probabilità:

- A: "il test dà esito positivo"
- B: "l'individuo ha il virus"
- $P(A|B) = 0.99$
- $P(A^C|B^C) = 0.997$
- $P(B) = 0.004$

Dai seguenti valori ricaviamo che:

- $P(A|B^C) = 1 - P(A^C|B^C) = 1 - 0.997 = 0.003$
- $P(B^C) = 1 - P(B) = 1 - 0.004 = 0.996$

Ora possiamo usare la formula delle probabilità totali per trovare P(A), ovvero la probabilità che il test sia positivo:

$$P(A) = P(A|B) \cdot P(B) + P(A|B^C) \cdot P(B^C) \simeq 0,004 + 0,003 = 0.007$$

Adesso vogliamo sapere P(B — A), ovvero la probabilità che abbia effettivamente il virus se risulta positivo al test, e per farlo usiamo la Formula di Bayes:

$$P(B|A) = \frac{P(A|B) \cdot P(B)}{P(A)} = \frac{0.99 \cdot 0.004}{0.007} \simeq 57\%$$

2.5 Indipendenza di eventi

Due eventi sono **indipendenti** se al verificarsi di uno, la probabilità che si verifichi l'altro non cambia:

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

Esempio: Consideriamo
A = "il primo dado vale 2" e C = "somma = 7" e D = "somma = 4".
A = $\{(2,1),(2,2),(2,3),(2,4),(2,5),(2,6)\}$ $|A| = 6$ $P(A) = \frac{1}{6}$
C = $\{(1,6),(2,5),(3,4),(4,3),(5,2),(6,1)\}$ $|C| = 6$ $P(C) = \frac{1}{6}$
D = $\{(1,3),(2,2),(3,1)\}$ $|D| = 3$ $P(D) = \frac{1}{12}$

$A \cap C = \{(2,5)\}$ $|A \cap C| = 1$ $P(A \cap C) = \frac{1}{36}$
Dato che $P(A \cap C) = P(A) \cdot P(C)$ A e C sono indipendenti

$A \cap D = \{(2,2)\}$ $|A \cap D| = 1$ $P(A \cap C) = \frac{1}{36}$
Dato che $P(A \cap D) \neq P(A) \cdot P(D)$ A e C sono dipendenti

Osservazione: Eventi indipendenti $\neq$ eventi disgiunti! Due eventi indipendenti non possono essere disgiunti.

Estensioni

Tre eventi A,B,C si dicono indipendenti se valgono *tutte* le seguenti regole:

- $P(A \cap B \cap C) = P(A) \cdot P(B) \cdot P(C)$
- $P(A \cap B) = P(A) \cdot P(B)$
- $P(A \cap C) = P(A) \cdot P(C)$
- $P(B \cap C) = P(B) \cdot P(C)$

Capitolo 3

Variabili aleatorie

3.0.1 Introduzione

Consideriamo uno spazio di probabilità (Ω, P). Spesso non siamo interessati a tutti i dettagli dell'esito dell'esperimento, ma solo a una *quantità determinata dall'esito* dell'esperimento. Tale quantità è detta **variabile aleatoria**. Una variabile aleatoria può essere descritta come una **funzione** $\Omega \to \mathbb{R}$

Osservazione

Evento e variabile aleatoria sono diverse ma in relazione. Ogni variabile aleatoria mi permette di determinare più eventi. Sia X una variabile aleatoria e x un suo possibile valore: $\{X = x\}$ è un evento.

3.1 Variabili Aleatorie Discrete

Una variabile aleatoria X si dice **discreta** se i valori che può assumere sono un insieme finito o un insieme infinito numerabile.
Ad ogni variabile aleatoria discreta X possiamo associare un indice chiamato **Densità Discreta**:

$$P_X(x_i) := P(X = x_i)$$

Proprietà

- Se x non è un valore assunto da X, si pone $P_X(x) := 0$
- P_X è una funzione da $\mathbb{R}$ a $[0, 1]$

- $P_X(x_i) \geq 0$
- $\sum_{i \geq 1} P_X(x_i) = 1$

Osservazione

La densità discreta ci risulta utile quando ci interessa sapere la densità di un sottoinsieme:

$$P(X \in B) = \sum_{x_i \in B} P_X(x_i)$$

Esempio: Una pasticceria prepara 3 torte al giorno. Sappiamo che:

- il 20% dei giorni nessun cliente ordina una torta
- il 30% dei giorni un cliente ordina una torta
- il 35% dei giorni due clienti ordinano una torta
- il restante dei giorni tre o più clienti ordinano una torta

1. Qual è la densità discreta di X?
2. Qual è la probabilità q che il numero di torte invendute sia pari?

Sappiamo che:

- $P_X(3) = P(X = 3) = 20\%$
- $P_X(2) = P(X = 2) = 30\%$
- $P_X(1) = P(X = 1) = 35\%$

Ricaviamo che

$$P_X(0) = P(X = 0) = 1 - P(X = 3) - P(X = 2) - P(X = 1) = 15\%$$

La probabilità che il numero di torte invendute sia pari è

$$P_X(0) + P_X(2) = 45\%$$

3.1.1 Valore Medio di Variabile Aleatoria Discreta

Sia X una variabile aleatoria discreta. Si definisce **valore medio di X**:

$$E[X] := \sum_{i=1}^{N} x_i \cdot p_X(x_i)$$

Proprietà Valore Medio

Per ogni variabile aleatoria X,Y e per ogni costante c $\in \mathbb{R}$ avremo:

· *Traslazione*: E[X + c] = E[X]

· *Moltiplicazione*: E[c · X] = c · E[X]

· *Somma*: E[X + Y] = E[X] + E[Y]

· *Formula di trasferimento*: $E[f(x)] = \sum_{i=1}^{N} f(x_i) \cdot P_X(x_i)$

Osservazioni

· Il valore medio è un operatore *lineare*: $E[Z] = X + c$.

· *Momento Secondo*: $E[X^2] = \sum_{i=1}^{N}(x_i)^2 \cdot P_X(x_i)$

· Il valore medio E[X] non è necessariamente un valore x_i assunto da X

· Il valore medio E[X] è quel valore tale per cui se ripetiamo un esperimento più volte, la media di tale esperimenti si avvicinerà a E[X]. Questo fenomeno è detto **Legge dei Grandi Numeri**:

$$\frac{\sum_{i=1} x_i}{N} \simeq E[X]$$

3.1.2 Varianza e Deviazione Standard

Sia X una variabile aleatoria discreta e $\mu = E[X]$. Definiamo la **varianza** come:

$$Var[X] := E[(X - \mu)^2]$$

A livello di calcolo, risulta più semplice utlizzare la formula:

$$Var[X] = E[X^2] - E[X]^2$$

Definiano la *deviazione standard* come:

$$Sd[X] := \sqrt{Var[X]}$$

Propietà della Varianza

Per ogni variabile aleatoria X e per ogni costante c $\in \mathbb{R}$ avremo:

· Var[X + c] = Var[X]

· Var[c · X] = c^2 Var[X]

Esempio Completo: Sia X il numero di figli maschi in una famiglia con due figli. Trovare: $P_X, E[X], Var[X], Sd[X]$

Spazio Campionario

$$\Omega = \{MM, MF, FM, FF\}$$

Densità Discreta

$$P_X(2) = \frac{1}{4} \qquad P_X(1) = \frac{1}{2} \qquad P_X(0) = \frac{1}{4}$$

Valore Medio

$$E[X] = 0 \cdot P(X = 0) + 1 \cdot P(X = 1) + 2 \cdot P(X = 2) = 1 \cdot \frac{1}{2} + 2 \cdot \frac{1}{4} = 1$$

Momento Secondo

$$E[X^2] = 0^2 \cdot P(X = 0) + 1^2 \cdot P(X = 1) + 2^2 \cdot P(X = 2) = 1 \cdot \frac{1}{2} + 4 \cdot \frac{1}{4} = 1 \cdot \frac{3}{2}$$

Varianza

$$\text{Var[X]} = E[X^2] - E[X]^2 = \frac{3}{2} - 1^2 = \frac{1}{2}$$

Deviazione Standard

$$\sqrt{Var[X]} = \frac{1}{\sqrt{2}}$$

Variabili Indipendenti

Siano X e Y due variabili aleatorie. Il valore della loro somma dipende da come sono "legate":

· Consideriamo Y = X. Allora Var[Y] = Var[X], quindi

$$Var[X + Y] = Var[2X] = 4 \cdot Var[X]$$

· Consideriamo invece Y = -X. Allora

$$Var[X + Y] = Var[X - X] = Var[0] = 0$$

Gli esempi precedenti sono casi estremi di dipendenza. Definiamo invece X e Y come variabili **indipendenti** se gli eventi {X = x} e {Y = y} sono indipendenti, ovvero

$$P(X = x, Y = y) = P(X = x) \cdot P(Y = y)$$

Definizione: Se X e Y sono indipendenti allora

$$Var[X + Y] = Var[X] + Var[Y]$$

3.2 Distribuzioni Notevoli Discrete

Consideriamo una variabile aleatoria X in un certo esperimento aleatorio: $X : \Omega \to \mathbb{R}$. Possiamo calcolare P($X \in A$) per ogni $A \subseteq \mathbb{R}$. L'insieme di tali probabilità definisce la **Distribuzione della variabile aleatoria X**.

Osservazione

Per variabili aleatorie discrete, la distribuzione di X è determinata dalla densità discreta e quindi, con abuso di notazione, si può dire che la distribuzione è la densità discreta.

3.2.1 Bernoulli

Variabile aleatoria X che può assumere soltanto i valori 0 e 1. Scriveremo X ∼ Be(p). Sia p := (X = 1). Dato che

$$\sum_{i=1}^{N} P_X(x_i) = P_X(0) + P_X(1) = 1$$

si ottiene:

$$P(X = x) = \begin{cases} p & \text{se x = 1} \\ (1 - p) & \text{se x = 0} \end{cases} \tag{3.1}$$

Allora, nel caso della variabile aleatoria di Bernoulli troviamo che:

· Il **valore medio** E[X] = p.

· Il**momento secondo** E[X^2] = p

· La **varianza** Var[X] = p(1 - p)

3.2.2 Binomiale

Consideriamo un esperimento aleatorio costituito da prove ripetute ed indipendenti dove abbiamo solo due esiti. Siano n il numero di prove e p la probabilità di successo di ciascuna e X il numero di successi.
La distribuzione di X è detta **binomiale** di parametri n e p indicata con
X $\sim$ Bin(n, p). La densità discreta è data da:

$$P(X = k) = \binom{n}{k} \cdot p^k \cdot (1 - p)^{n-k}$$

dove:

- P(X = k) è esattamente il numero k di successi in n prove
- p^k è la probabilità di k successi fissati
- $(1 - p)^{n-k}$ è la probabilità di (n - k) insuccessi fissati
- $\binom{n}{k}$ scelte di quali prove hanno successo

Introduciamo le variabili:

$$X_i = \begin{cases} 1 & \text{se la i-esima prova ha successo} \\ 0 & \text{se la i-esima prova non ha successo} \end{cases} \tag{3.2}$$

Possiamo allora scrivere:

$$X = \sum_{i=1}^{N} X_i$$

Avremo che ogni $X_i \sim Be(p)$ e quindi sappiamo valore medio e varianza. Otteniamo allora:

- **Valore medio** E[X] $= n \cdot p$
- **Varianza** Var[X] $= n \cdot p \cdot (1 - p)$

Esempio: Sia X numero di figli maschi in una famiglia con 2 figli. Abbiamo $\Omega = \{MM, MF, FM, FF\}$ e avremo :

$$P_X(0) = \frac{1}{4} \qquad P_X(1) = \frac{1}{2} \qquad P_X(2) = \frac{1}{4}$$

Allora X $\sim$ Bin(2; $\frac{1}{2}$):

$$P_X(k) = \binom{2}{k} \cdot \frac{1}{2}^k \cdot \frac{1}{2}^{2-k} = \frac{2}{k! \cdot (2-k)!} \cdot \frac{1}{2^2} = \frac{1}{2} \cdot \frac{1}{k! \cdot (2-k)!}$$

Infatti, se sostituiamo k=0, ritroveremo $P_X(0) = \frac{1}{4}$, se k=1 avremo $P_X(1) = \frac{1}{2}$ e con k=3 avremo $P_X(2) = \frac{1}{4}$

3.2.3 Poisson

Una variabile aleatoria X si dice di Poisson di parametro $\lambda \in (0, \infty)$ ovvero X $\sim Pois(\lambda)$ se X(Ω) = $\mathbb{N}_0$ e

$$P(X = k) = e^{-\lambda} \cdot \frac{\lambda^k}{k!}$$

Approssimazione di Poisson: Possiamo ottenere $X \sim Pois(\lambda)$ attraverso una variabile aleatoria binomiale $Y \sim Bin(n, p)$ quando:

$$n \to \infty \qquad p \to 0 \qquad \text{allora} \quad n \cdot p = \lambda$$

Allora avremo che:

· Il **Valore medio** E[X] = λ

· La **Varianza** Var[X] = λ

Osservazione

Le variabili aleatorie di Poisson sono approssimazioni per variabili aleatorie che contano il numero di successi quando si considera una grande quantità di prove la cui probabilità di successo è piccola.

Esempio: In un ospedale nascono mediamente 2,2 bambini ogni giorno. Qual è la probablità che nessun bambino nasca in ogni giorno. E qual è la probabilità che ne nascono più di 3?
Sia X il numero di nascite in un giorno. Supponiamo $X \sim Pois(\lambda)$. Sappiamo che E[X] = λ e per ipotesi E[X] = 2,2. calcoliamo:

· P(X = 0) = $e^{-\lambda} \cdot \frac{e^0}{0!} = e^{-2,2} \simeq 11\%$

· P(X > 3) = 1 - P(X = 0) - P(X = 1) - P(X = 2) - P(X = 3) = $1 - e^{-\lambda}(1 + \lambda + \frac{\lambda^2}{2} + \frac{\lambda^3}{6}) = 18\%$

3.2.4 Geometrica

Una variabile aleatoria X si dice **Geometrica** di parametro p $\in$ (0, 1] e si scrive X $\sim$ Geo(p) se

$$P(X = k) = p \cdot (1 - p)^{k-1}$$

Osservazione

Se p è 1, allora possiamo dire che, per quanto piccola sia la probabilità, l'evento prima o poi accadrà.

Possiamo ottenere una variabile aleatoria geometrica partendo da una successione di prove ripetute dove consideriamo T l'istante del primo successo. Avremo che

$$P(T = k) = P(X_1 = 0, ..., X_{K-1} = 0, X_K = 1) = (1 - p)^{k-1} \cdot p$$

Allora troviamo:

· il **Valore medio** E[X]= $\frac{1}{p}$

· la **Varianza** Var[X] = $\frac{1-p}{p^2}$

Osservazione

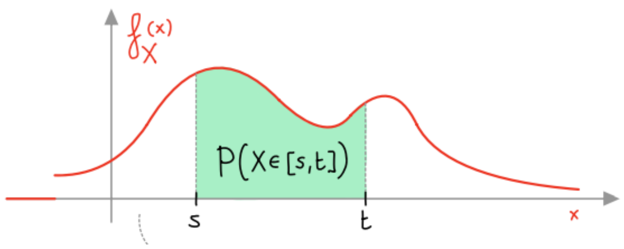
Possiamo calcolare la probabilità di coda: $\geq$ n:

$$P(T > n) = (1 - p)^n \qquad P(T \leq n) = \sum_{k=1}^{N} P_T(k) = (1 - p)^n$$

3.3 Var. Aleatorie Assolutamente Continue

Consideriamo una classe complementare di variabili aleatorie, dette **assolutamente continue**, che assumono un insieme *infinito più che numerabile di valori*, come ad esempio un intervallo in $\mathbb{R}$.
Una variabile X è assolutamente continua se la sua distribuzione è determinata da una funzione $f_X(x)$ a valori positivi detta **densità della variabile aleatoria X** nel modo seguente:

$$P(X \in [s,t]) = \int_s^t f_X(x)dx$$



Osservazione

Troviamo una analogia tra una variabile aleatoria discreta ed una assolutamente continua:

$$\text{V.A. discreta: } \sum_{x\in[s,t]} P_X(x_i) \qquad \text{V.A. assolutamente continua: } \int_s^t f_X(x)$$

Notiamo anche delle differenze importanti: se X è assolutamente continua $\forall x \in \mathbb{R} : P(X = x) = 0$. In particolare, tranne per $f_X(x) = 0$:

$$f_X(x) \neq P(X = x)$$

Densità

La **densità** di una variabile aleatoria assolutamente continua X è una funzione $f_X : \mathbb{R} \to \mathbb{R}$ integrabile tale che:

$$f_X(x) \geq 0 \quad \forall x \in \mathbb{R} \qquad\qquad \int_{-\infty}^{+\infty} f_X(x)dx = 1$$

Valore medio e varianza di v.a. assolutamente continue

Le definizioni di valore medio e varianza di v.a. assolutamente continue ricalcano quelle delle v.a. discrete:

- $E[X] = \int_{-\infty}^{+\infty} x \cdot f(x)dx$
- $Var[X] = E[X^2] - E[X]^2$
- $Sd[X] = \sqrt{Var[X]}$

Anche le proprietà definite per le variabili discrete continuano a valere:

- $E[X+c] = E[X]+c \quad E[c \cdot X] = c \cdot E[X] \quad \cdot E[X+Y] = E[X]+E[Y]$
- $Var[X+c] = Var[X] \quad Var[c \cdot X] = c^2 \cdot Var[X]$
- Se X e Y sono indipendenti: $Var[X+Y] = Var[X] + Var[Y]$

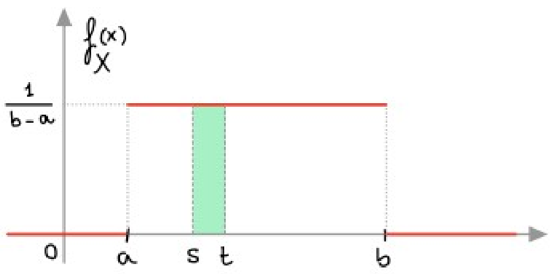
3.3.1 Uniforme Continua

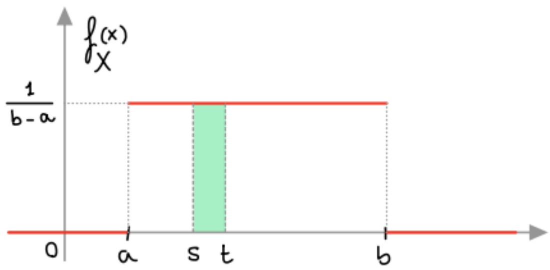
Una variabile aleatoria X è **uniforme continua in [0,1]** e si indica con $X \sim U(0,1)$ se è definita da una funzione:

$$f_X(x) = \begin{cases} c = 1 & \text{se } x \in [0,1] \\ 0 & \text{se } x \notin [0,1] \end{cases} \tag{3.3}$$

Una variabile aleatoria X è **uniforme continua in [a,b]** e si indica con $X \sim U(a,b)$ se è definita da una funzione:

$$f_X(x) = \begin{cases} \frac{1}{b-a} = 1 & \text{se } x \in [a,b] \\ 0 & \text{se } x \notin [a,b] \end{cases} \tag{3.4}$$





Dato un intervallo $[s, t] \subseteq [a, b]$ avremo che

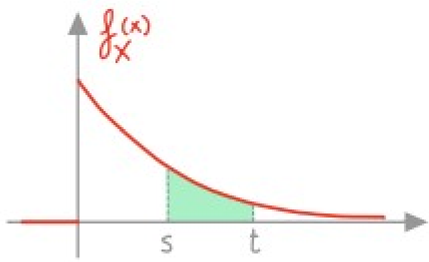
$$P(X \in [s, t]) = \int_s^t f(x)dx + \frac{1}{b-a} = \int_s^t 1\ dx = \frac{t-s}{b-a}$$

Una variabile aleatoria X con tale densità è detta esponenziale di parametro $\lambda \in (0, +\infty)$

3.3.2 Esponenziale

Una variabile aleatoria **esponenziale** è spesso utilizzata per descrivere il tempo tra gli arrivi di eventi casuali indipendenti di un processo di Poisson. Misuriamo il tempo medio di evento come τ, avremo che $\lambda = \frac{1}{\tau}$. La funzione che la descrive è

$$f_X(x) = \begin{cases} \lambda \cdot e^{-\lambda \cdot x} & \text{se } x \geq 0 \\ 0 & \text{se } x < 0 \end{cases} \tag{3.5}$$



Una variabile aleatoria X con tale densità è detta Esponenziale di parametro $\lambda \in (0, \infty)$ e si scrive X $\sim$ Exp(λ). Avremo che:

- $P(X \in [s, t]) = \int_s^t f_X(x)dx = e^{-\lambda \cdot s} - e^{-\lambda \cdot t}$
- $E[X] = \frac{1}{\lambda}$

· $Var[X] = \frac{1}{\lambda^2}$

Funzione di ripartizione

Finora abbiamo studiato le v.a. discrete e assolutamente continue. Introduciamo un nuovo oggetto per v.a. generica: **funzione di ripartizione**

$$F_X(x) := P(X \leq x)$$

· F_X è ben definita per ogni v.a.

· F_X determina la distribuzione della v.a. X:

$$P(X \in (s,t]) = F_X(t) - F_X(s)$$

· F_X è legata alla densità discreta/densità di X:

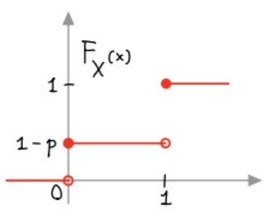
$$F_X(x) = \begin{cases} \sum_{x_i \in (-\infty, x]} P_X(x_i) & \text{se x è discreta} \\ \int_{-\infty}^{x} f_X(t)dt & \text{se x è assolutamente continua} \end{cases} \tag{3.6}$$

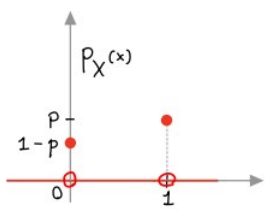
Esempio: Funzione di ripartizione per variabile discreta: Bernoulli.
Sia $X \sim Be(p)$ con $p \in (0,1)$:

$$X(\Omega) = \{0,1\} \qquad p_X(0) = 1-p \qquad p_X(1) = p$$

Allora $F_X(x) = P(X \leq x)$ vale:

$$f_X(x) = \begin{cases} 0 & \text{se } x < 0 \\ 1-p & \text{se } 0 \leq x < 1 \\ 1 & \text{se } x \geq 1 \end{cases} \tag{3.7}$$



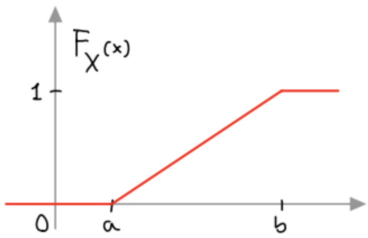


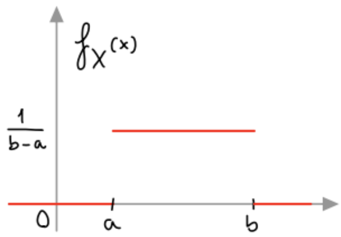
Esempio: Funzione di ripartizione per variabile uniforme continua
Sia $X \sim U(a, b)$: $X(\Omega) = [a, b]$

$$f_X(x) = \begin{cases} \frac{1}{b-a} & \text{se } a \leq x \leq b \\ 0 & \text{se } x < a \text{ o } x > b \end{cases} \tag{3.8}$$

Allora $F_X(x) = P(X \leq x)$ vale:

$$F_X(x) = \begin{cases} 0 & \text{se } x < a \\ \frac{x-a}{b-a} & \text{se } a \leq x \leq b \\ 1 & \text{se } x > b \end{cases} \tag{3.9}$$





Teorema di ripartizione di v.a. discrete

- X è una variabile aleatoria discreta $\iff$ F_X è costante a tratti
- Valori assunti x_i $\iff$ punti di discontinuità di F_X
- Densità discreta $\iff$ ampiezze dei salti
- $P_X(x_i) = F_X(x_i) - F_X(x_{\bar{i}})$ dove $x_{\bar{i}} = \lim_{t \to x_{\bar{i}}} F_X(t)$

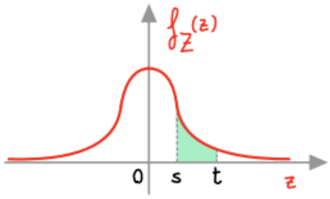
Teorema di ripartizione di v.a. assolutamente continue

- X è una variabile aleatoria assolutamente continua $\iff$ F_X è una funzione continua ed è derivabile a tratti
- Densità $F_X(x) = (F_x)'(x)$

3.4 V.A. Assolutamente Continue Normali

Abbiamo già visto due classi notevoli assolutamente continue: uniforme continua ed esponenziale. Consideriamo ora la più importante: **variabili aleatorie normali** (o gaussiane).
Una variabile aleatoria X si dice **normale standard** e si indica con
$Z \sim N(0, 1)$ se è assolutamente continua e ha densità:

$$f_Z(z) = \frac{1}{\sqrt{2 \cdot \pi}} \cdot e^{-\frac{z^2}{2}}$$



Osservazione: Se X è una variabile aleatoria normale standard, allora la variabile aleatoria $Y = a \cdot X + b$ ha una distribuzione normale con media b e deviazione standard $|a|$, dunque la sua distribuzione sarà $Y \sim N(b, a^2)$.

Una variabile aleatoria normale standard avrà:

- E[Z] = 0
- Var[Z] = 1
- P(Z $\in [s, t]) = \int_s^t f_Z(z)dz$

Purtroppo l'integrabile non è calcolabile, allora si introduce la **funzione di ripartizione di Z** indicata con:

$$\Phi(z) := F_Z(z) = P(Z \leq z) = \int_s^t f_Z(t)dt$$

Anche questa non è calcolabile esplicitamente ma i valori che può assumere sono riportati in una tabella.

Osservazione: per valori negativi si applica la formula:

$$\Phi(z) = 1 - \Phi(-z)$$

z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.5000	0.5040	0.5080	0.5120	0.5160	0.5199	0.5239	0.5279	0.5319	0.5359
0.1	0.5398	0.5438	0.5478	0.5517	0.5557	0.5596	0.5636	0.5675	0.5714	0.5753
0.2	0.5793	0.5832	0.5871	0.5910	0.5948	0.5987	0.6026	0.6064	0.6103	0.6141
0.3	0.6179	0.6217	0.6255	0.6293	0.6331	0.6368	0.6406	0.6443	0.6480	0.6517
0.4	0.6554	0.6591	0.6628	0.6664	0.6700	0.6736	0.6772	0.6808	0.6844	0.6879
0.5	0.6915	0.6950	0.6985	0.7019	0.7054	0.7088	0.7123	0.7157	0.7190	0.7224
0.6	0.7257	0.7291	0.7324	0.7357	0.7389	0.7422	0.7454	0.7486	0.7517	0.7549
0.7	0.7580	0.7611	0.7642	0.7673	0.7704	0.7734	0.7764	0.7794	0.7823	0.7852
0.8	0.7881	0.7910	0.7939	0.7967	0.7995	0.8023	0.8051	0.8078	0.8106	0.8133
0.9	0.8159	0.8186	0.8212	0.8238	0.8264	0.8289	0.8315	0.8340	0.8365	0.8389
1.0	0.8413	0.8438	0.8461	0.8485	0.8508	0.8531	0.8554	0.8577	0.8599	0.8621
1.1	0.8643	0.8665	0.8686	0.8708	0.8729	0.8749	0.8770	0.8790	0.8810	0.8830
1.2	0.8849	0.8869	0.8888	0.8907	0.8925	0.8944	0.8962	0.8980	0.8997	0.9015
1.3	0.9032	0.9049	0.9066	0.9082	0.9099	0.9115	0.9131	0.9147	0.9162	0.9177
1.4	0.9192	0.9207	0.9222	0.9236	0.9251	0.9265	0.9279	0.9292	0.9306	0.9319
1.5	0.9332	0.9345	0.9357	0.9370	0.9382	0.9394	0.9406	0.9418	0.9429	0.9441
1.6	0.9452	0.9463	0.9474	0.9484	0.9495	0.9505	0.9515	0.9525	0.9535	0.9545
1.7	0.9554	0.9564	0.9573	0.9582	0.9591	0.9599	0.9608	0.9616	0.9625	0.9633
1.8	0.9641	0.9649	0.9656	0.9664	0.9671	0.9678	0.9686	0.9693	0.9699	0.9706
1.9	0.9713	0.9719	0.9726	0.9732	0.9738	0.9744	0.9750	0.9756	0.9761	0.9767
2.0	0.9772	0.9778	0.9783	0.9788	0.9793	0.9798	0.9803	0.9808	0.9812	0.9817
2.1	0.9821	0.9826	0.9830	0.9834	0.9838	0.9842	0.9846	0.9850	0.9854	0.9857
2.2	0.9861	0.9864	0.9868	0.9871	0.9875	0.9878	0.9881	0.9884	0.9887	0.9890
2.3	0.9893	0.9896	0.9898	0.9901	0.9904	0.9906	0.9909	0.9911	0.9913	0.9916
2.4	0.9918	0.9920	0.9922	0.9925	0.9927	0.9929	0.9931	0.9932	0.9934	0.9936
2.5	0.9938	0.9940	0.9941	0.9943	0.9945	0.9946	0.9948	0.9949	0.9951	0.9952
2.6	0.9953	0.9955	0.9956	0.9957	0.9959	0.9960	0.9961	0.9962	0.9963	0.9964
2.7	0.9965	0.9966	0.9967	0.9968	0.9969	0.9970	0.9971	0.9972	0.9973	0.9974
2.8	0.9974	0.9975	0.9976	0.9977	0.9977	0.9978	0.9979	0.9979	0.9980	0.9981
2.9	0.9981	0.9982	0.9982	0.9983	0.9984	0.9984	0.9985	0.9985	0.9986	0.9986
3.0	0.9987	0.9987	0.9987	0.9988	0.9988	0.9989	0.9989	0.9989	0.9990	0.9990
3.1	0.9990	0.9991	0.9991	0.9991	0.9992	0.9992	0.9992	0.9992	0.9993	0.9993
3.2	0.9993	0.9993	0.9994	0.9994	0.9994	0.9994	0.9994	0.9995	0.9995	0.9995
3.3	0.9995	0.9995	0.9995	0.9996	0.9996	0.9996	0.9996	0.9996	0.9996	0.9997
3.4	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9998
3.5	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998
3.6	0.9998	0.9998	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999
3.7	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999
3.8	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999
3.9	0.99995	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
4.0	0.99997									
5.0	0.9999997									
6.0	0.9999999990									

Una variabile aleatoria X si dice **normale** con media μ e varianza σ^2 e si scrive con $X \sim N(\mu, \sigma^2)$ se X è assolutamente continua con:

$$f_X(x) = \frac{1}{\sqrt{2 \cdot \pi \cdot \sigma^2}} \cdot e^{-\frac{z^2}{2}}$$

Osservazione

Ci si può sempre ricondurre da una variabile normale ad una normale standard:

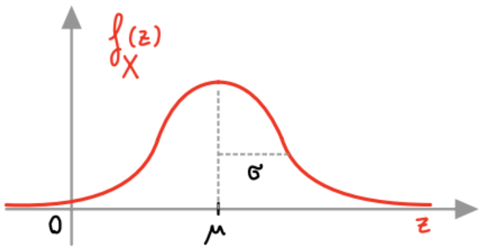


Figura 3.1: grafico a campana centrato in μ di ampiezza σ^2

· $X \sim N(\mu, \sigma^2) \to Z := \frac{X-\mu}{\sigma} \sim N(0,1)$

· $X \sim N(\mu, \sigma^2) \to E[X] = \mu \quad Var[X] = \sigma^2$

Teoremi

Se X è normale, $Y = a \cdot X + b$ è normale.
Se X e Y sono normali indipendenti allora X + Y è normale

Esempio: $X \sim N(0,1)$, $Y \sim N(0,1)$ allora

$$X + Y \sim N(0,2) \qquad X - Y = X + (-1) \cdot (Y) \sim N(0,2)$$

3.5 Vettori Aleatori

Abbiamo studiato le v.a. individualmente, ma spesso è interessante lo studio *congiunto* di v.a. relative allo stesso esperimento aleatorio.

$$(\Omega, P) = \begin{cases} X : \Omega \to \mathbb{R} & \text{variabile aleatoria} \\ Y : \Omega \to \mathbb{R} & \text{variabile aleatoria} \end{cases} \tag{3.10}$$

La coppia (X,Y) è detta **vettore aleatorio**:

$$(X, Y) : \Omega \to R^2$$

3.5.1 Vettori Aleatori Discreti

Un vettore (X, Y) si dice **discreto** se i valori che può assumere sono contenuti in un insieme finito o numerabile $\{(x_i), (y_i)\}$

Si definisce **densità discreta congiunta** come:

$$P_{(X,Y)}(x_i, Y_i) := P(X = x_i, Y = y_i)$$

La relazione tra densità discreta congiunta e la densità discreta delle singole variabili aleatorie troviamo le **densità discrete marginali**:

$$P_X(x_i) = \sum_{y_i} P_{(X,Y)}(x_i, y_i) \qquad P_Y(y_i) = \sum_{x_i} P_{(X,Y)}(x_i, y_i)$$

Il valore medio del prodotto tra due variabili aleatorie discrete X e Y è

$$E[X, Y] = \sum_{x_i} \sum_{y_i} x_i \cdot y_i \cdot P_{X,Y}(x_i, y_i)$$

Esempio: Lancio due monete dove X è prima moneta testa e Z numero totale di teste.
$X(\Omega) = 0, 1$
$Y(\Omega) = 0, 1, 2$
$\mathrm{p}_{(X,Z)}(x, z) = P(X = x, Z = z)$

x \ z	0	1	2
0	$\frac{1}{4}$	$\frac{1}{4}$	0
1	0	$\frac{1}{4}$	$\frac{1}{4}$

3.5.2 Vettori Aleatori Assolutamente Continui

Un vettore (X, Y) si dice assolutamente continuo se esiste una funzione $f_{(X,Y)}(x, y) \geq 0$ detta **densità congiunta** di X e Y tale che

$$P(X \in [s, t], Y \in [u, v]) = \int_s^t \left(\int_u^v f_{(X,Y)}(x, y) dy \right) dx$$

Si definiscono **densità marginali** come:

$$f_X(x) = \int_{-\infty}^{+\infty} f_{(X,Y)}(x, y) dy \quad f_Y(y) = \int_{-\infty}^{+\infty} f_{(X,Y)}(x, y) dx$$

Il valore medio del prodotto tra due v.a. assolutamente continue X e Y è

$$E[X \cdot Y] = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} x \cdot y \cdot f_{(X,Y)}(x, y) \; dx \; dy$$

3.5.3 Indipendenza

Consideriamo due v.a. X e Y che insieme formano un vettore aleatorio (X,Y). Si dice che X e Y sono **indipendenti** se

$$P(X \in [s,t], Y \in [u,v]) = P(X \in [s,t] \cdot Y \in [u,v])$$

Da questa formula si può ricavare che X e Y sono indipendenti se conoscendo il valore che assume Y *non cambia la distribuzione* di X:

$$P(X \in [s,t] | Y \in [u,v]) = P(X \in [s,t]$$

Osservazione: Per vettori discreti o assolutamente continui, l'indipendenza ha una riformulazione equivalente:

(X,Y) discreto, X e Y sono indipendenti sse $p_{X,Y}(x_i, y_i) = p_X(x_i) \cdot p_Y(y_i)$

(X,Y) assolut. continuo, X e Y sono indipendenti sse $f_{(X,Y)}(x,y) = f(x) \cdot f(y)$

Osservazione: Quando le v.a. non sono indipendenti, le densità marginali forniscono meno informazioni della congiunta. Quando le v.a. sono dipendenti, le densità dei singoli fornisce la stessa quantità di informazioni della congiunta.

Il valore medio del prodotto tra due variabili aleatorie indipendenti X e Y è

$$E[X \cdot Y] = E[X] \cdot E[Y]$$

3.5.4 Covarianza e Correlazione

Consideriamo due v.a. X e Y che insieme formano un vettore aleatorio (x, Y). Indichiamo i valori medi con $\mu_X = E[X] e \mu_Y = E[Y]$.
Si definisce **covarianza** di X e Y:

$$Cov[X,Y] := E[(X - \mu_X) \cdot (Y - \mu_Y)]$$

La covarianza misura il *grado di associazione* tra X e Y:
$Cov[X,Y] > 0$ quando a valori grandi di X corrispondono valori grandi di Y $Cov[X,Y] < 0$ quando a valori grandi di X corrispondono valori piccoli di Y

Osservazione: la covarianza si può riscrivere come

$$Cov[X,Y] = E[X \cdot Y] - E[X] \cdot E[Y]$$

Proprietà della covarianza

· Annullamento: $Var[X] = Cov[X, X]$

· Simmetria: $Cov[X, Y] = Cov[Y, X]$

· Costanti: $Cov[X, c] = 0$

· Bilinearità: $Cov[a \cdot X, Y] = a \cdot Cov[X, Y]$

· Bilinearità: $Cov[X + Z, Y] = Cov[X, Y] + Cov[Z, Y]$

· **Formula della somma**: $Var[X] + Var[Y] + 2 \cdot Cov[X, Y]$

Si definisce **coefficiente di correlazione lineare**:

$$\rho[X, Y] := \frac{Cov[X, Y]}{Sd[X] \cdot Sd[Y]}$$

Si tratta di una versione normalizzata della covarianza:

· $-1 \leq \rho[X, Y] \leq 1$

· $\rho = \pm 1 \longleftrightarrow Y = a \cdot X + b$

Osservazione: Se Cov[X,Y] = 0, X e Y si dicono **scorrelate**. Se abbiamo due variabili aleatorie indipendenti, allora sicuramente sono scorrelate, ma *non vale viceversa*

Esempio: Lancio due dadi regolari a 6 facce. Siano X la somma dei risultati e Y la differenza. Allora X e Y sono scorrelate:

$$\begin{aligned} Cov[X, Y] &= Cov[A + B, A - B] \\ &= Cov[A, A] - Cov[A, B] - Cov[B, A] - Cov[B, B] \\ &= Cov[A, A] - Cov[B, B] \\ &= Var[A] - Var[B] = 0 \end{aligned} \tag{3.11}$$

Tuttavia X e Y non sono indipendenti. Per esempio:

$$P(X = 12) = \frac{1}{6} \quad P(Y = 5)\frac{1}{6}$$

$$P(X = 12, Y = 5) = 0$$

$$0 \neq \frac{1}{6} \cdot \frac{1}{6}$$