Contents

1 Introduzione alla probabilità

1.1 Glossario

- Probabilità $\rightarrow$ *la materia che studia i sistemi non deterministici*
  - Frequestista $\rightarrow$ *probabilità assegnata sulla base di più esperimenti ripetuti nelle stesse condizioni*
  - Soggettivista $\rightarrow$ *non esiste un valore oggettivo ma ci si basa sulla fiducia e sull'incertezza che l'individuo ha riguardo l'occorrenza di un certo evento*
- Sistemi non deterministici $\rightarrow$ *conoscendo i dati iniziali non possiamo determinare i dati finali*
- Incertezza degli eventi $\rightarrow$ *la varianza degli eventi che possono succede*
- Rumore $\rightarrow$ *possiamo misurare un evento solo approssimatamente*
- Statistica $\rightarrow$ studio dei sistemi deterministici
- Varianza $\rightarrow$ *dispersione dei dati attorno al valore centrale / media / valore*

- Confidenza → *intervallo che rappresenta una stima dei valori medi*

- Frequenza

  - Frequenza assoluta → Numero di volte che si verifica un evento

  - Frequenza relativa → Rapporto tra frequenza assoluta e il numero di prove/dati

- Dataset → numero di dati a disposizione $\boldsymbol{D_n = \{x_1 \dots x_n\}}$

- Principio di enumerazione → Passare solo una volta da ogni elemento della raccolta

- Spazio esiti (s o Ω) → Tutti i possibili esiti di un evento → $\boldsymbol{Dado = \{1 \dots 6\}}$

- Spazio eventi (e) → Tutti i possibili risultati di un esperimento → $\boldsymbol{Dado = \{1 || 2\}}$ ← che esca $\mathbf{1}$ oppure $\mathbf{2}$

- Statistica descrittiva → la branca della statistica che usa tecniche per descrivere le caratteristiche dei dati in un esperimento (tramite grafici ecc.)

- Assioma → Tre assiomi fondamentali su cui si poggia la teoria del calcolo delle probabilità

- 1' Assioma $\rightarrow$ La probabilità di E è un numero reale **non negativo**
  $\mathbb{P}(E) \in \mathbb{R}, \mathbb{P}(E) \geq 0 \ \forall E \subseteq \Omega$ | $0 \leq P(E) \leq 1$

- 2' Assioma $\rightarrow$ Allo spazio degli esiti è sempre associato ad **1**
  $\mathbb{P}(s) = 1$

- 3' Assioma $\rightarrow$ Per ogni coppia di eventi incompatibili $E_1, E_2 \subseteq \Omega$
  la probabilità di $E_1 \cup E_2$ è uguale alla **somma della loro probabilità**
  $\mathbb{P}(E_1 \cup E_2) = \mathbb{P}(E_1) + \mathbb{P}(E_2)$

1.2 Moda e Mediana

1.2.1 Moda

Definizione: La moda è il valore che presenta la **massima frequenza** all'interno del dataset

Formula generica:

$$Moda \rightarrow v_i : f_i = max f_i \begin{cases} \text{un solo valore} & \textbf{Moda} \\ \text{più di un valore} & \textbf{Valori modali} \end{cases}$$

1.2.2 Mediana

Definizione: La mediana è il **valore centrale** all'interno del dataset (dati ordinati in ordine crescente/decresente)

Formula generica:

$$Mediana = \begin{cases} \text{n pari} & \frac{x_{\frac{n}{2}} + x_{\frac{n}{2}+1}}{2} \\ \text{n dispari} & x_{[\frac{n+1}{2}]} \leftarrow \text{Intero superiore (Ceil)} \end{cases}$$

Esempio:

$$D_n = \textbf{\{ 28, 34, 51, 19, 62, 43, 29, 38, 45, 26, 49, 33 \}}$$

Per la mediana è necessario ordinare i dati in ordine crescente:

$$D_n = \textbf{\{ 19, 26, 28, 29, 33, 34, 38, 43, 45, 49, 51, 62 \}}$$

Mediana:

$$\frac{x_{\frac{12}{2}} + x_{\frac{12}{2}+1}}{2} = \frac{x_6 + x_7}{2} = \frac{34+38}{2} = \frac{72}{2} = \mathbf{36}$$

Nota: quando si trova ad esempio x_6 bisogna andare a sostituire il valore con la posizione di x

1.3 Media, Varianza e Deviazione Standard Campionaria

1.3.1 Media Campionaria

Definizione: La media campionaria è la **media** degli elementi di un campione.

Formula generica:

$$\overline{X} = \frac{1}{n} \sum_{i=1}^{n} X_i$$

1.3.2 Varianza Campionaria

Definizione: La varianza campionaria è la **dispersione** degli elementi di un campione.

Formula generica:

$$S^2 = \frac{\sum_{i=1}^{n} (x_i - \overline{x})^2}{n-1}$$

1.3.3 Deviazione Standard

Definizione: La deviazione standard è la **media** degli scarti degli elementi di un campiona.

Formula generica:

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

Esempio: (Varianza, Media e Dev. St.) $D_n =$ **{ 3, 4, 6, 7, 10 }**

$$\text{Media del campione: } \overline{X} = \frac{\mathbf{(3+4+6+7+10)}}{\mathbf{5}} = \frac{30}{5} = 6$$

$$\text{Varianza campionaria: } S^2 = \frac{\mathbf{[(-3)^2+(-2)^2+0^2+1^2+4^2]}}{\mathbf{4}} = 7.5$$

$$\text{Dev. St. campionaria: } S = \sqrt{S^2} = \sqrt{7.5}$$

1.4 Percentile

Definizione: Il percentile è un indicatore che serve ad **indicare il valore minimo** sotto al quade ricade una **determinata percentuale** degli altri elementi sotto osservazione.

$$Valore \begin{cases} \geq & \text{k \% dati} \\ \leq & \text{100 - k \% dati} \end{cases}$$

Dove il secondo quartile è sempre uguale alla **mediana**

1. Prima cosa da fare è ordinare i valori in ordine crescente

2. Si calcola il prodotto $k = np$

   - *p* indica il percentile in decimale es. $p = 0.25$

   - *n* indica il numero di dati presenti nel dataset

- se k è **intero** il valore si ottiene tramite la media dell'elemento k-esimo e del (k+1)

- se k **non è intero** il valore si ottiene arrotondando per *eccesso* al primo intero e si sceglie come percentile la posizione nel dataset del valore trovato

Esempio: $D_n =$ **{ 1, 1, 2, 3, 4, 5, 6, 7, 8, 9 }** $n = 10$
Calcolo **25esimo** percentile (1' quartile) (k è intero, **arrotondiamo** per eccesso)

$$k = 10 \cdot 0.25 = 2.5 \approx \text{pos } \mathbf{3} = \text{valore: } \mathbf{2}$$

Calcolo **50esimo** percentile (2' quartile) = mediana

$$k = 10 \cdot 0.50 = 5 \longrightarrow \frac{4+5}{2} = \text{valore: } \mathbf{4.5}$$

Calcolo **75esimo** percentile (3' quartile)

$$k = 10 \cdot 0.75 = 7.5 \approx \text{pos } \mathbf{8} = \text{valore: } \mathbf{7}$$

1.5 Disugaglianza di Chebyshev

Definizione: Dice quanti dati di un campione cadono all'interno di un intervallo con centro la **media** $\forall k \geq 1 : k \in \mathbb{R}$

$$(\overline{x} - k_s, \overline{x} + k_s) \longrightarrow \boldsymbol{S_k : [i : 1 \leq i \leq n, |x_i - \overline{x} < k_s|]}$$

Formula generica:

$$\frac{\#S_k}{n} \geq 1 - \frac{n-1}{nk^2} > 1 - \frac{1}{k^2}$$

S_k è il numero di elementi dell'insieme S_k
n è il numero di dati nel dataset
k^2 è l'incognita da trovare

Generalizzando:

$$|x - \overline{x}| < 5 \longrightarrow 68\%$$

$$|x - \overline{x}| < 25 \longrightarrow 95\%$$

$$|x - \overline{x}| < 35 \longrightarrow 99.7\%$$

Esempio: $D_n =$ **{2, 1, 4, 5, 3, 1, 4, 3, 2, 1}**
(trovare l'intervallo che ricade nel 30%)
La media è *2.6*
S_k **= 3** $\longrightarrow \mathbf{100 : 10 = 30 : \boldsymbol{x}} \longrightarrow \frac{10 \cdot 30}{100} = 3$
$\frac{3}{10} > 1 - \frac{1}{k^2}$
$k^2 = \frac{7}{10}$
$-\sqrt{\frac{7}{10}} < k < \sqrt{\frac{7}{10}}$
$(2.6 - \sqrt{\frac{7}{10}}, 2.6 + \sqrt{\frac{7}{10}})$

1.6 Insieme di dati Bivariati

Definizione: è lo studio della relazione di due variabili.

Formula generica:

$$D_n : \{(X_1, Y_1)(X_2, Y_2) \dots (X_n, Y_n)\}$$

1.6.1 Coefficiente di correlazione campionario

Definizione: utilizzato per capire se esiste un legame **lineare** tra due serie di dati.

Formula generica:

$$r = \frac{\sum_{i=1}^{n}(x_i - \overline{x})(y_i - \overline{y})}{(n-1)s_x s_y} = \frac{\sum_{i=1}^{n}(x_i - \overline{x})(y_i - \overleftarrow{y}\,)}{\sqrt{\sum_{i=1}^{n}(x_i - \overline{x})^2 \sum_{i=1}^{n}(y_i - \overline{y})^2}}$$

1.7 Permutazioni, Combinazioni e Disposizioni

1.7.1 Permutazioni

Definizione: Modi possibili per sistemare **n** oggetti (**0! = 1**)

$$n! = n \cdot (n-1) \dots (n \cdot (n-1))$$

Esempio: Fattoriale di 6

$$6! = 6 \cdot (6-1) \cdot (6-2) \dots (6-5) = 720$$

1.7.2 Combinazioni

Definizione: Modi di disporre **k** elementi scelti da **n** elementi (l'ordine **non conta**)

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

Esempio: in una classe di **26** alunni si devono eleggere **2** rappresentanti

$$C_{n,k} = \frac{n!}{k! \cdot (n-k)!}$$

Sostituiamo **n** con *26* (numero di alunni) e **k** con *2* (numeri di rappresentanti)

$$C_{26,2} = \frac{26!}{2! \cdot (26-2)!} = \frac{26!}{2! \cdot 24!} = \frac{25 \cdot 26}{2} = \mathbf{325}$$

è possibile anche semplificare i fattoriali come in questo caso

1.7.3 Disposizioni

Definizione: Modi di disporre **k** elementi scelti da **n** elementi (l'ordine **conta**)

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

Esempio: Quante parole is possono ottenere usando 4 **diverse** lettere da *youmath*
In questo caso dobbiamo contare le **disposizioni** senza ripetizione di **classe 4 di 7**

$$D_{7,4} = \frac{7!}{(7-4)!} = \frac{7!}{3!} = \frac{5040}{6} = \mathbf{840}$$

1.8 Probabilità condizionata

Definizione: è la probabilità che succeda un evento **E** dato un evento **F**

$$P(E|F) = \frac{P(E|F)}{P(F)}$$

Esempio: **3** scatole con contenuto nascosto dove in una è presente il premio

$$P(\text{Vincita}) = \frac{1}{3}$$

$$P(\text{Vincita } \mid \text{ 1' pacco contiene un gatto}) = \frac{1}{2}$$

$$P(\text{Vincita } \mid \text{ 1' pacco } \textbf{NON} \text{ contiene un gatto}) = 0$$

1.8.1 Teorema di Bayes

Formula generica:

$$P(F_j|E) = \frac{P(F_j \cap E)}{P(E)} = \frac{P(E|F_j) \cdot P(F_j)}{\sum_{i=1}^{p} P(E|F_i) \cdot P(F_i)}$$

Probablità di F_j sapendo che si sia verificato l'evento **E**

Esempio: Un oggetto si è rotto, e si suppone che possa essere rotto a causa di un componente qualsiasi dei 5 elementi che lo compongono, con uguale probabilità. Per $i = 1, 2, \ldots, 5$, sia $1 - \alpha_i$ la probabilità di diagnosticare che il componente i-esimo sia rotto. Qual è la probabilità che l'oggetto sia rotto a causa di ognuno di 5 componenti se il componente 1 è stato diagnosticato funzionante?

Risoluzione :
Per R_i: (probabilità che il componente *i-esimo* sia rotto)
Per E: (l'evento che conosciamo (in questo caso che il *componente 1* funziona))
$1-\alpha_i$ è la probabilità che il componente *i-esimo* sia rotto

Per R_1:

$$P(R_1 \mid E) = \frac{P(E \mid R_1) P(R_1)}{\sum_{i=1}^{5} P(E \mid R_i) P(R_i)}$$

1.9 Operazioni e proprietà tra eventi

Definizione: Prendiamo come esempio **E** ed **F** come eventi

- E $\cup$ F $\longleftarrow$ Unione
- E $\cap$ F $\longleftarrow$ Intersezione
- E $\subset$ F $\mid$ E $\subseteq$ F $\longleftarrow$ Contenuto
- E $\supset$ F $\mid$ E $\supseteq$ F $\longleftarrow$ Contiene
- E^c $\longleftarrow$ Complemento

Le seguenti operazioni possono essere combinate tra di loro: formando cosi le proprietà che seguono:

- $E \cup (F \cup G) = (E \cup F) \cup G \longrightarrow$ Associativa unione

- $(E \cup F) \cap G = (E \cap G) \cup (F \cup G) \longrightarrow$ Distributiva intersezione

- $E \cap (F \cap G) = (E \cap F) \cap G \longrightarrow$ Associativa intersezione

- $(E \cap F) \cup G = (E \cup G) \cap (F \cup G) \longrightarrow$ Distributiva unione

- $(E \cup F)^c = \frac{E^c \cap F^c}{(E \cap F)^c} = E^c \cup F^c$

2 Variabile aleatorie

Definizione: La variabile aleatoria è una variabile che può assumere **valori diversi** in dipendenza da *qualche esperimento casuale*.

$$X \begin{cases} Discrete & \text{Solo } \textbf{valori finiti} \\ Continue & \text{Possono assumere } \textbf{range illimitati} \end{cases}$$

2.1 Funzione di ripartizione (Tutte le variabili)

Definizione: La Probabilità che la variabile aleatoria **X** assuma un valore minore o uguale ad x

Formula generica: $F(x) = P(X \leq x)$

- F = funzione di ripartizione

- X = variabile aleatoria

- x = variabile normale

Esempio :

$P(a < X \leq b)$

$P(X \leq b) = P(X \leq a) + P(a < X \leq b)$

$P(a < X \leq b) = P(X \leq b) - P(X \leq a) = F(b) - F(a)$

2.2 Funzione di massa (Variabili discrete)

Formula generica: $p(a) = P(X = a)$

se si ha la funzione di ripartizione è possibile ottenere la funzione di massa perche:

$$X \leq a = \cup X_i$$

Formula generica:

$$F(x) = P(X \leq a) = \sum_{x \leq a} p(x_i)$$

Esempio: variabile aleatoria X che può assumere valori **1**, **2** o **3**
Dato che p(1) + p(2) + p(3) = **1**

Se: $$p(1) = \frac{1}{2} \qquad\qquad p(2) = \frac{1}{3}$$

Allora:

$$p(3) = \frac{1}{6}$$

La funzione di ripartizione F di X è data da:

$$F(a) = \begin{cases} 0 & a < 1 \\ \frac{1}{2} & 1 \leq a < 2 \\ \frac{5}{6} & 2 \leq a < 3 \\ 1 & 3 \leq a \end{cases}$$

2.3 Funzione della densità di probabilità (Variabili continue)

Formula generica:

$$P(X \in B) = \int_B f(x)\,dx$$

$$P(X \in (-\infty, +\infty)) = \int_{-\infty}^{+\infty} f(x)\,dx = 1$$

integrando $-\infty$ a $+\infty$ la probabilità che avvenga x è per forza 1 perche andiamo ad includere tutti i valori di $\mathbb{R}$

Se abbiamo che **B = [a, b]** $\longrightarrow P(a \leq X \leq b) = \int_a^b f(x)\,dx$
Se abbiamo che **B = [a]** $\longrightarrow P(X = a) = \int_a^a f(x)\,dx = 0$

Relazione che lega la funzione di ripartizione **F** alla densità **f**:

$$F(x) = P(X \leq x) = \int_{-\infty}^{x} f(x)\, dx$$

Derivando entrambi i membri otteniamo che:

$$\frac{d}{da} F(a) = f(a)$$

Esempio: Sia assegnata una variabile aleatoria X con densità data da:

$$f(x) = \begin{cases} C(4x - 2x^2) & 0 < x < 2 \\ 0 & \text{altrimenti} \end{cases}$$

(a) quanto vale C? **(b)** quanto vale P(X > 1)?

(a) siccome f è una densita allora:

$$\begin{aligned} 1 &= C \int_0^2 (4x - 2x^2)\, dx \\ &= C[2x^2 - \frac{2x^3}{3}]\Big|_{x=0}^{x=2} = C \cdot \frac{8}{3} \\ &= C = \frac{3}{8} \end{aligned}$$

(b) conoscendo ora la densità f possiamo trovare la P(X > 1):

$$P(X > 1) = \int_1^{\infty} f(x)\, dx = \frac{3}{8} \int_1^2 (4x - 2x^2)\, dx = \frac{1}{2}$$

3 Funzioni a due variabili

Questo tipo di funzioni ci sono utili quando l'utilizzo di una sola variabile è impossibile poichè *l'oggetto in questione è basato sulla relazione di due variabili aleatorie*

3.1 Funzione di ripartizione congiunta

Definizione: Funzione di ripartizione a due variabili aleatorie X e Y

Formula generica:

$$F(x, y) = P(X \leq x, Y \leq y)$$

Se vogliamo trovare solamente la funzione di ripartizione di una singola variabile aleatoria:

$$\begin{aligned} F_X(x) &= P(X \leq x) \\ &= P(X \leq x, Y \leq \infty) \\ &= F(x, \infty) \end{aligned}$$

Applicabile anche alla $F_y(y)$:

$$F_Y(y) = F(\infty, y)$$

3.2 Funzione di massa congiunta

Definizione: Probabilita che accadano due eventi (**X** e **Y**) nello stesso istante.

Formula generica:

$$p(x_i, y_j) = P(X = x_i, Y = y_j)$$

Se vogliamo trovare solamente la funzione di massa di una singola variabile aleatoria:

$$\begin{aligned} p_X(x_i) &:= P(X = x_i) \\ &= P(\bigcup_j \{X = x_i, Y = y_j\}) \\ &= \sum_j P(X = x_i, Y = y_j) \\ &= \sum_j p(x_i, y_j) \end{aligned}$$

Applicabile anche alla p_Y:

$$p_Y(y_j) = \sum_i p(x_i, y_j)$$

$$\sum_x \sum_y p(x, y) = 1$$

3.3 Funzione densità congiunta

Definizione: Due variabili aleatorie X e Y sono *congiuntamente continue* se esiste un funzione non negativa f(x,y) definita per tutti gli x e gli y

Formula generica:

$$P((X,Y) \in C) = \int \int_{(x,y) \in C} f(x,y)\, dx\, dy$$

se A e B sono sottoinsiemi qualsiasi di $\mathbb{R}$ e C:= A x B

$$C := (x,y) \in \mathbb{R}^2 : x \in A, y \in B$$

Possiamo riscrivere la funzione di ripartizione congiunta di X e Y come segue:

$$\begin{aligned} F(a,b) &= P(X \leq a, Y \leq b) \\ &= P(X \in a, Y \in b) \\ &= \int_B \int_A f(x,y)\, dx\, dy \\ &= \int_{-\infty}^{a} \int_{-\infty}^{b} f(x,y)\, dx\, dy \end{aligned}$$

Esempio: Siano X e Y due variabili aleatorie congiuntamente continue con densità di probabilità data da:

$$f(x,y) = \begin{cases} 2e^{-x}e^{-2y} & x > 0, y > 0 \\ 0 & altrimenti \end{cases}$$

Si calcolino **(a)** P(X > 1, Y < 1)

$$\begin{aligned}
P(X > 1, Y < 1) &= \int_0^1 \int_1^\infty 2e^{-x}e^{-2y}\,dx\,dy \\
&= \int_0^1 2e^{-2y}(\int_1^\infty e^{-x}\,dx)\,dy \\
&= \int_0^1 2e^{-2y}\{-e^{-x}\}|_{x=1}^{\infty}\,dy \\
&= e^{-1}\int_0^1 2e^{-2y}\,dy \\
&= e^{-1}(1-e^{-2})
\end{aligned}$$

In questo caso si è integrato prima in una variabile e poi nell'altra

3.4 Variabili aleatorie indipendenti

3.4.1 X,Y indipendenti

Definizione: Un evento su una variabile non influenza l'altra.

Formula generica: Se soddisfano questa richiesta le variabili si dicono *indipendenti*

$$P(X \in A, Y \in B) = P(X \in A)P(Y \in B)$$

Usando gli assiomi della probabilità è possibile dimostrare che la definizione di sopra è equivalente a:

$$P(X \leq a, Y \leq b) = P(X \leq a)P(Y \leq b)$$

$$\forall a, b \in \mathbb{R}$$

Ovvero che la funzione di ripartizione congiunta sia il prodotto delle marginali:

$$F(a, b) = F_X(a)F_Y(b)$$

Funzione di massa:

$$p(a, b) = p_X(a)p_Y(b)$$

$$P(X = a, Y = b) = P(X = a)P(Y = b)$$

Dimostrazione:

$$\begin{aligned} P(X \in A, Y \in B) &= \sum_{x \in A} \sum_{y \in B} p(x, y) \\ &= \sum_{x \in A} \sum_{y \in B} p_X(x)p_Y(y) \\ &= \sum_{x \in A} p_X(x) \sum_{y \in B} p_Y(y) \\ &= P(X \in A)P(Y \in B) \end{aligned}$$

Funzione di densità:

$$f(x, y) = f_X(x)f_Y(y)$$

$$\forall x, y \in \mathbb{R}$$

Esempio con variabili indipendenti continue e con stessa funzione di densità:

$$f_X(t) = f_Y(t) = \begin{cases} e^{-t} & t > 0 \\ 0 & altrimenti \end{cases}$$

Quale è la densità di probabilità della variabile aleatoria data dal rapporto X/Y

$$\begin{aligned}
F_{X|Y}(a) &= P(X|Y \le a) \\
&= \int_{(x,y)} \int_{x \le ay} f(x,y)\,dx\,dy \\
&= \int_{(x,y)} \int_{x \le ay} f(x)f(y)\,dx\,dy \\
&= \int_0^\infty \int_0^{ay} e^{-x} f(x)f(y)\,dx\,dy \\
&= \int_0^\infty e^{-y} (\int_0^{ay} e^{-x})\,dy \\
&= \int_0^\infty e^{-y}(1 - e^{-ay})\,dy \\
&= [-e^{-y} + \frac{e^{-(a+1)y}}{(a+1)}]\Big|_0^\infty \\
&= 1 - \frac{1}{a+1}
\end{aligned}$$

La funzione di densità si ricava infine **derivando** la funzione di ripartizione

$$f_{X|Y}(a) = \frac{d}{da}(1 - \frac{1}{a+1}) = \frac{1}{(a+1)^2} a > 0$$

3.5 Distribuzioni condizionate

Definizione: La distribuzione condizionata di Y dato X è la probabilità di X quando è conosciuto il valore assunto da X.
A ogni distribuzione condizionata è associato un valore atteso condizionato e una varianza condizionata

Formula generica: $P(E|F) = \frac{P(E \cap F)}{P(F)}$

3.6 funzione di massa condizionata (Discrete)

Formula generica:

$$
\begin{aligned}
p_{X|Y} = P(X|Y) = P(X = x, Y = y) \\
= \frac{P(X = x, Y = y)}{P(Y = y)} \\
= \frac{p(X, Y)}{p_Y(x, y) > 0}
\end{aligned}
$$

$$\forall x, \forall y \text{ con } p_Y(y) > 0$$

Se y non è un valore possibile di Y, ovvero se P(Y = y) = 0
la quantità $\boldsymbol{p_{X|Y}(x|y)}$ non è definita:

Esempio: Siano X e Y due variabili aleatorie discrete con funzione di massa congiunta p dato che:

$$p(0, 0) = 0.4 \qquad p(0, 1) = 0.2 \qquad p(1, 0) = 0.1 \qquad p(1, 1) = 0.3$$

Calcolare la massa di X condizionata da Y = 1

$$P(Y=1) = \sum_x p(x,1) = p(0,1) + p(1,1) = 0.5$$

Quindi:

$$P(X=0|Y=1) = \frac{p(0,1)}{P(Y=1)} = \frac{2}{5}$$

$$P(X=1|Y=1) = \frac{p(1,1)}{P(Y=1)} = \frac{3}{5}$$

Se X e Y sono variabili congiuntamente continue, non è possibile utilizzare la definizione di distribuzione condizionata valida per quelle discrete, infatti sappiamo che P(Y = y) = 0 per tutti i valori di y

3.7 funzione di densità condizionata (Continue)

Formula generica:

$$f_{X|Y}(x|y) = \frac{f(x,y)}{f_Y(y)}$$

Se X e Y sono congiuntamente continue e A è un sottoinsieme di numeri reali per ogni y si può definire:

$$P(X \in A|Y=y) := \int_A f_{X|Y}(x|y)\,dx$$

Notiamo che X e Y sono indipendenti allora:

$$f_{X|Y}(x,y) = f_X(x) \qquad\qquad P(X \in A|Y=y) = P(X \in A)$$

Esempio: è data la seguente densità congiunta di X e Y

$$f(x,y) = \begin{cases} \frac{12}{5}x(2-x-y) & 0 < x < 1, 0 < y < 1 \\ 0 & altrimenti \end{cases}$$

Si calcoli la densità condizionata di X rispetto a Y = y per 0 < y < 1.
Se questi due numeri sono compresi tra 0 e 1 abbiamo che:

$$\begin{aligned} f_{X|Y}(x,y) := & \frac{f(x,y)}{f_Y(y)} \\ &= \frac{f(x,y)}{\int_{-\infty}^{\infty} f(x',y)\,dx'} \\ &= \frac{x(2-x-y)}{\int_0^1 x'(2-x'-y)\,dx'} \\ &= \frac{x(2-x-y)}{\frac{2}{3}-\frac{y}{2}} \\ &= \frac{6x(2-x-y)}{4-3y} \end{aligned}$$

4 Valore atteso

Definizione: Rappresenta la media pesata dei valori di una variabile aleatoria

4.1 Funzione di massa (Discrete)

$$\mathbb{E}[X] := \sum_i x_i P(X = x_i)$$

Si può dire quindi che il valore atteso è anche detto *media* di X oppure *aspettazione*

Esempio semplice: Se X è una variabile aleatoria con funzione di massa

$$p(0) = \frac{1}{2} = p(1)$$

Allora:

$$\mathbb{E}[X] = 0 \cdot \frac{1}{2} + 1 \cdot \frac{1}{2} = \frac{0+1}{2} = \frac{1}{2}$$

Esempio dado fair 6 facce $P(x_i = i) = \frac{1}{6}$

$$\mathbb{E}[X] = \sum_{i=1}^{6} i \frac{1}{6} = \frac{1}{6} \sum_{i=1}^{6} i = \frac{21}{6} = \frac{7}{2} = 3.5$$

Oppure:

$$\mathbb{E}[X] := 1 \cdot \frac{1}{6} + 2 \cdot \frac{1}{6} + 3 \cdot \frac{1}{6} + 4 \cdot \frac{1}{6} + 5 \cdot \frac{1}{6} + 6 \cdot \frac{1}{6} = \frac{7}{2} = 3.5$$

dove il risultato è la media dei valori che X può assumere

Se N è molto grande allora $N_i \approx N_p(x_i)$

$$\sum_{i}^{n} x_i p(x_i) \approx \sum_{i}^{n} x_i \frac{n_i}{n}$$

4.2 Funzione di densità (Continue)

Formula generica:

$$\mathbb{E}[X] := \int_{-\infty}^{+\infty} x f(x) \, dx$$

Esempio: Siamo in attesa di una comunicazione che deve arrivare dopo le ore 17.
a partire dalle 17 è una variabile aleatoria con funzione di densità data da:

$$f(x) = \begin{cases} \frac{1}{1.5} & \text{se } 0 < x < 1.5 \\ 0 & \text{altrimenti} \end{cases}$$

Il valore atteso del tempo che trascorre tra le 17 e il momento di arrivo della comunicazione è quindi:

$$\mathbb{E}[X] = \int_{0}^{1.5} \frac{x}{1.5} \, dx = 0.75$$

4.3 Valore atteso di una funzione

Definizione: è possibile calcolare il valore atteso di una funzione g(X) notando che essa stessa è una variabile aleatoria quindi si applicano le stesse proprietà, come segue:

Variabile discreta:

$$\mathbb{E}[g(X)] = \sum_i g(x_i)p(x_i)$$

Variabile Continua:

$$\mathbb{E}[g(x)] = \int_{-\infty}^{+\infty} g(x)f(x)\,dx$$

Esempio (discrete): quanto vale il valore atteso del quadrato di una variabile X con le seguenti funzioni di massa?

p(0) = 0.2   p(1) = 0.5   p(2) = 0.3

Se poniamo $Y := X^2$ questa diventa una variabile che può assumere i valori 0^2, 1^2, 2^2

$$p_Y(0) := P(Y = 0^2) = 0.2$$
$$p_Y(1) := P(Y = 1^2) = 0.5$$
$$p_Y(4) := P(Y = 2^2) = 0.3$$

Quindi:

$$\mathbb{E}[X^2] = \mathbb{E}[Y] = 0 \cdot 0.2 + 1 \cdot 0.5 + 4 \cdot 0.3 = 1.7$$

Oppure (utilizzando la proposizione delle variabili discrete)

$$\mathbb{E}[X^2] = 0^2 \cdot 0.2 + 1^2 \cdot 0.5 + 2^2 \cdot 0.3 = 1.7$$

Esempio (continue): Il tempo – in ore – necessario per localizzare un guasto nell'impianto elettrico di una fabbrica è una variabile aleatoria X con funzione di densità

$$f(x) = \begin{cases} 1 & 0 < x < 1 \\ 0 & \text{altrimenti} \end{cases}$$

Se il danno economico provocato da una interruzione di x ore è x^3, qual è il valore atteso di questo costo?

Applicando la proposizione della variabile continua possiamo ottenere quanto segue:

$$\mathbb{E}[X^3] = \int_0^1 x^3 \, dx = \frac{1}{4}$$

4.4 Dimostrazioni

Sia per discreto che per continuo si applicano le seguenti proprietà:

$$\mathbb{E}[aX + b] = a\mathbb{E}[X] + b$$

Se proviamo a ponere a = 0 scopriamo che:

$$\mathbb{E}[b] = b$$

Se proviamo a ponere b = 0 scopriamo che:

$$\mathbb{E}[aX] = a\mathbb{E}[X]$$

Ovvero, il valore atteso di un fattore costante moltiplicato per una variabile aleatoria, è pari alla costante per il valore atteso della variabile aleatoria.

Per caso discreto:

$$\begin{aligned}
\mathbb{E}[aX + b] &= \sum_x (ax + b)p(x) \\
&= a \sum_x xp(x) + b \sum_x p(x) \\
&= a\mathbb{E}[X] + b
\end{aligned}$$

Per caso continuo:

$$\begin{aligned}\mathbb{E}[aX+b] &= \int_{-\infty}^{+\infty} (ax+b)f(x)\,dx \\ &= a\int_{-\infty}^{+\infty} xf(x)\,dx + b\int_{-\infty}^{+\infty} f(x)\,dx \\ &= a\mathbb{E}[X]+b\end{aligned}$$

4.5 Momenti N-esimi nel valore atteso

Definizione: se n = 1,2 ... n, la quantità $\mathbb{E}[X^n]$ se esiste viene detta *momento n-esimo* della variabile aleatoria X.

è possibile applicare le formule di prima, come segue:

$$\mathbb{E}[X^n] = \begin{cases} \sum_x x^n p(x) & \text{se X è discreta} \\ \int_{-\infty}^{+\infty} x^n f(x)\,dx & \text{se X è continua} \end{cases}$$

4.6 Valore atteso della somma di due variabili

Definizione: è possibile applicare le formule viste sopra anche quando abbiamo due variabili aleatorie

se in questo caso $\mathbb{E}[\boldsymbol{g(X,Y)}]$ esiste allora:

$$\mathbb{E}[g(X,Y)] = \begin{cases} \sum_x \sum_y g(x,y)p(x,y) & \text{Se discreto} \\ \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} g(x,y)f(x,y)\,dx\,dy & \text{Se continuo} \end{cases}$$

se $g(X, Y)$ come $\boldsymbol{g = X + Y}$ allora

$$\mathbb{E}[X+Y] = \mathbb{E}[X] + \mathbb{E}[Y]$$

Dimostrazione caso discreto:

$$\begin{aligned}
\mathbb{E}[X+Y] &= \sum_x \sum_y (x+y)p(x,y) \\
&= \sum_x x \cdot [\sum_j p(x_i, y_j)] + \sum_x y \cdot [\sum_i p(x_i, y_j)] \\
&= \sum_x x p_X(x) + \sum_y y p_Y(y) \\
&= \mathbb{E}[X] + \mathbb{E}[Y]
\end{aligned}$$

Dimostrazione caso continuo:

$$\begin{aligned}
\mathbb{E}[X+Y] &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (x+y) f(x,y)\, dx\, dy \\
&= \int_{-\infty}^{\infty} x [\int_{-\infty}^{\infty} f(x,y)\, dy]\, dx + \int_{-\infty}^{+\infty} y [\int_{-\infty}^{+\infty} f(x,y)\, dx]\, dy \\
&= \int_{-\infty}^{\infty} x f_X(x)\, dx + \int_{-\infty}^{\infty} y f_Y(y)\, dy \\
&= \mathbb{E}[X] + \mathbb{E}[Y]
\end{aligned}$$

è possibile applicare la ricorsione per il numero di variabili aleatori

$$\begin{aligned}\mathbb{E}[X+Y+Z] &= \mathbb{E}[(X+Y)+Z] \\ &= \mathbb{E}[X+Y]+\mathbb{E}[Z] \\ &= \mathbb{E}[X]+\mathbb{E}[Y]+\mathbb{E}[Z]\end{aligned}$$

In generale per ogni n

$$\mathbb{E}[X_1+X_2+\ldots+X_n] = \mathbb{E}[X_1]+\mathbb{E}[X_2]+\ldots\mathbb{E}[X_n]$$

Esempio: 2 dadi a 6 facce

$$\begin{aligned}\mathbb{E}[X+Y] &= \mathbb{E}[X]+\mathbb{E}[Y] \\ &= \sum_{i=1}^{6} x_i p(x_i) + \sum_{i=1}^{6} y_i p(y_i) \\ &= \sum_{i=1}^{6} x_i \frac{1}{6} + \sum_{i=1}^{6} y_i \frac{1}{6} \\ &= \frac{7}{2}+\frac{7}{2} = 7\end{aligned}$$

Dove **7** è il valore atteso della somma dei due dadi.

Se vogliamo predire il valore di X possiamo scegliere un numero che sarà ugual ad X. L'errore che commeteremo sarà di $\boldsymbol{(X-c)^2}$

Se $\boldsymbol{c = \mathbb{E}[X]}$ l'errore sarà minimizzato $\boldsymbol{\mu := \mathbb{E}[X]}$

$$\mathbb{E}[(X-c)^2] \geq \mathbb{E}[(X-\mu)^2]$$

5 Varianza

Definizione: Indica di quanto i dati si discostano dalla media al quadrato

$\mu = \mathbb{E}[X] \leftarrow$ Primo momento $\qquad$ $\mathbb{E}[X^2] \leftarrow$ Momento secondo

Formula generica:

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

Generalizzazione:

$$\begin{aligned}
Var(X) &= \mathbb{E}[(X-\mu)^2] \\
&= \mathbb{E}[X^2 - 2\mu \cdot X + \mu^2] \\
&= \mathbb{E}[X^2] - 2\mu \cdot \mathbb{E}[X] + \mu^2 \\
&= \mathbb{E}[X^2] - \mu^2 \longrightarrow \mathbb{E}[X^2] - \mathbb{E}[X]^2
\end{aligned}$$

Esempio: Varianza di un dado

$$\begin{aligned}
\mathbb{E}[X^2] &= \sum_{1}^{6} i^2 P(X=i) \\
&= 1^2 \cdot \frac{1}{6} + 2^2 \cdot \frac{1}{6} + 3^2 \cdot \frac{1}{6} + 4^2 \cdot \frac{1}{6} + 5^2 \cdot \frac{1}{6} + 6^2 \cdot \frac{1}{6} \\
&= \frac{91}{6}
\end{aligned}$$

Sapendo che $\mathbb{E}[X] = \frac{7}{2}$

$$Var(X) = \mathbb{E}[X^2] - \mathbb{E}[X]^2 = \frac{91}{6} - (\frac{7}{2})^2 = \frac{35}{12}$$

5.1 Costanti reali nella varianza

Una utile identità che riguarda la varianza è la seguente (per ogni coppia di costanti reali a e b)

$$Var(aX + b) = a^2 \cdot Var(X)$$

Per dimostrare ciò ricordiamoci sempre di $\boldsymbol{\mu := \mathbb{E}[X]}$

Dimostrazione:

$$\begin{aligned} Var(aX + b) :&= \mathbb{E}[(aX + b - \mathbb{E}[aX + b])^2] \\ &= \mathbb{E}[(aX + b - a\mu - b)^2] \\ &= \mathbb{E}[a^2(X - \mu)^2] \\ &= a^2\mathbb{E}[(X - \mu)^2] \\ &= a^2 Var(X) \end{aligned}$$

Se sostituiamo i valori di **a** e **b** troviamo che:
SE $a = 0 \longrightarrow Var(b) = 0 \longrightarrow$ le costanti hanno varianza **nulla**
SE $a = 1 \longrightarrow Var(X+b) = Var(X) \longrightarrow$ sommando una const. non cambia la varianza
SE $b = 0 \longrightarrow Var(aX) = a^2 \cdot Var(X)$

5.2 Deviazione Standard

Definizione: Indica di quanto dei dati si **discotastano dalla media** (non al quadrato)

Formula generica:

$$S = \sqrt{Var(X)}$$

$$Var(X+X) = Var(2 \cdot X) = 4 \cdot Var(X)$$

Se X è indipendente allora:

$$Var(X) + Var(X) = Var(X+X)$$

6 Covarianza

Definizione: Misura la **variazione** tra due variabili aleatorie associate tra di loro *è la media del prodotto degli scarti*

Formula generica:

$$\boldsymbol{Cov(X, Y) := \mathbb{E}[(X - \mu_X)(Y - \mu_Y)]}$$

Dove:
$\mu_x = \mathbb{E}[X]$
$\mu_y = \mathbb{E}[Y]$

La covarianza può essere negativa, positiva o nulla

Positivo $\longrightarrow$ Le due variabili crescono o decrescono insieme
Negativo $\longrightarrow$ Quando una variabile cresce l'altra decresce
Nullo $\longrightarrow$ Le due variabili sono indipendenti

è presente una formula alternativa **più semplice** *(si trova espandendo il prodotto al secondo membro)*

$$\begin{aligned} Cov(X, Y) &= \mathbb{E}[XY - \mu_X Y - \mu_Y X + \mu_X \mu_Y] \\ &= \mathbb{E}[XY] - \mu_X \mathbb{E}[Y] - \mu_Y \mathbb{E}[X] + \mu_X \mu_Y \\ &= \mathbb{E}[XY] - \mu_X \mu_Y - \mu_X \mu_Y + \mu_X \mu_Y \\ &= \mathbb{E}[XY] - \mathbb{E}[X]\mathbb{E}[Y] \end{aligned}$$

6.1 Proprietà della covarianza

$Cov(X, Y) = Cov(Y, X) \longleftarrow$ Simmetria

$Cov(X,X) = Var(X) \longleftarrow$ Generalizzazione della varianza

$Cov(X+Y,Z) = Cov(X,Z) + Cov(Y,Z)$

$Cov(X+Y,Z+W) = Cov(X,Z) + Cov(X,W) + Cov(Y,Z) + Cov(Y,W)$

$Var(X+Y) = Var(X) + Var(Y) + 2Cov(X,Y)$

- Se $X_1 \dots X_n$ e Y sono variabili aleatorie qualsiasi allora:

$$Cov(\sum_{i=1}^{n} X_i, Y) = \sum_{i=1}^{n} Cov(X_i, Y)$$

- Se $X_1 \dots X_n$ e $Y_1 \dots Y_m$ sono variabili aleatorie qualsiasi allora:

$$Cov(\sum_{i=1}^{n} X_i, \sum_{j=1}^{m} Y_j) = \sum_{i=1}^{n}\sum_{j=1}^{m} Cov(X_i, Y_j)$$

6.2 Variabili indipendenti

Definizione: Se X e Y sono variabili aleatorie **indipendenti:**

$$\mathbb{E}[XY] = \mathbb{E}[X]\mathbb{E}[Y]$$

Questo implica che:

$$Cov(X,Y) = 0$$

$$Var(\sum_{i=1}^{n} X_i) = \sum_{i=1}^{n} Var(X_i)$$

Esempio: varianza della somma di 10 lanci indipendenti di un dado
Denotiamo con X_i il punteggio del dado *i-esimo*, sappiamo che:

$$\begin{aligned} Var(\sum_{i=1}^{10} X_i) &= \sum_{i=1}^{10} Var(X_i) \\ &= 10 \cdot \frac{35}{12} \\ &= \frac{175}{6} \end{aligned}$$

Esempio: Un sistema composto di n componenti distinti si dice in parallelo se funziona fino a che almeno uno dei componenti funziona
Sia dato un sistema di questo tipo, per il quale, per $i = \{1, 2, \ldots, n\}$ il componente *i-esimo* funziona - indipendentemente da tutti gli altri - con probabilità $p_{\mathrm{i.}}$.
Qual è la probabilità che **l'intero sistema funzioni**?
Denotiamo con A_i l'evento che il componente i funzioni. Allora:

$$\begin{aligned} P(\text{ il sistema funziona }) &= 1 - P(\text{ il sistema non funziona }) \\ &= 1 - P(\text{ nessun componente funziona }) \\ &= 1 - P\left(A_1^c \cap A_2^C \cap \cdots \cap A_{\mathrm{n}}^c\right) \\ &= 1 - \prod_{i=1}^{n} (1 - P_i) \qquad \text{per } I^*i \text{ indipendenza} \end{aligned}$$

6.3 Coefficiente di correlazione lineare

Definizione: numero puro che tiene conto della deviazione standard di X e Y

Formula generica:

$$Corr(X,Y) := \frac{Cov(X,Y)}{\sqrt{Var(X) \cdot Var(Y)}}$$

La correlazione può assumere valori compresi tra **-1** e **1**

-1 $\longrightarrow$ Le due variabili sono inversamente proporzionali
0 $\longrightarrow$ Le due variabili sono indipendenti
1 $\longrightarrow$ Le due variabili sono crescono o decrescono con lo stesso rapporto

7 Funzione generatrice dei momenti

Definizione: Funzione che ci permette di calcolare i momenti della distribuzione.

Formula generica:

$$\phi(t) = \mathbb{E}[e^{tX}]$$

Dove X è una variabile aleatoria e t è un parametro reale

$$\phi(t) := \mathbb{E}[e^{tX}] = \begin{cases} \sum\limits_{x} e^{tx} p(x) & \text{se X discreta} \\ \int_{-\infty}^{+\infty} e^{tx} f(x)\, dx & \text{se X continua} \end{cases}$$

Derivando la funzione si ottengono i momenti:

$$\phi'(t) = \frac{d}{dt}\mathbb{E}[e^{tX}] = \mathbb{E}[\frac{d}{dt}e^{tX}] = \mathbb{E}[Xe^{tX}] \longrightarrow \phi'(0) = \mathbb{E}[X]$$

Analogamente:

$$\phi''(t) = \frac{d^2}{dt^2}\mathbb{E}[e^{tX}] = \mathbb{E}[\frac{d^2}{dt^2}e^{tX}] = \mathbb{E}[X^2 e^{tX} \longrightarrow \phi''(0) = \mathbb{E}[X^2]]$$

Generalizzando:

$$\phi^{(n)}(0) = \mathbb{E}[X^n]$$

Media: $\mu_x = \phi'(0)$ Varianza: $\sigma_x^2 = \phi''(0) = \{\phi(0)'\}'$

Ipotizziamo: se X e Y sono indipendenti con ϕ_X e ϕ_Y e se ϕ_{X+Y} è la funzione generatrice dei momenti di X + Y allora:

$$\phi_{X+Y}(t) = \phi_X(t)\phi_Y(t)$$

Concludiamo che:

$$\begin{aligned}
\phi_{X+Y}(t) :&= \mathbb{E}[e^{t(X+Y)}] \\
&= \mathbb{E}[e^{tX}e^{tY}] \\
&= \mathbb{E}[e^{tX}]\mathbb{E}[e^{tY}] \\
&= \phi_X(t)\phi_Y(t)
\end{aligned}$$

7.1 Disuguaglianza di Markov

Definizione: Ci permette di sapere la probabilità che una variabile assuma valori molto grandi
Serve per calcolare che una variabile aleatoria assuma un minimo di "a" $\longrightarrow a \in \mathbb{R}$

Solo per variabili positive: $X \in (0, +\infty)$

Formula generica:

$$P(X \geq a) \leq \frac{\mathbb{E}[X]}{a}$$

Dimostrazione:

$$\begin{aligned}
\mathbb{E}[X] &:= \int_0^{+\infty} xf(x)\,dx \\
&= \int_0^a xf(x)\,dx + \int_a^{+\infty} xf(x)\,dx \\
&\geq \int_a^{+\infty} xf(x)\,dx \\
&\geq \int_a^{+\infty} af(x)\,dx \\
&= a\int_a^{+\infty} f(x)\,dx \\
&= \boldsymbol{aP(X \geq a)}
\end{aligned}$$

7.2 Disuguaglianza di Chebyshev

Definizione: Ci permette di sapere la probabilità che una variabile si discosti dalla media per più di un certo numero di deviazioni standard.

$$\text{Se } X \text{ var aleatoria} \begin{cases} \mu & \text{Media} \\ \sigma^2 & \text{Varianza} \end{cases}$$

Per ogni $\mathbf{r} > \mathbf{0} \longrightarrow$ valore che indica il discostamento dalla media

Formula generica:

$$P(|X - \mu| \geq r) \leq \frac{\sigma^2}{r^2}$$

Formula generica: per calcolare la probabilità di X nell'intervallo

$$P(a \leq X \leq b) = P(|X - \mu| \leq r) \geq 1 - \frac{\sigma^2}{r^2}$$

Dimostrazione: Dimostriamo che:

$$\{|X - \mu| \geq r\} \qquad\qquad \{(X - \mu)^2 \geq r^2\}$$

Questi due eventi coincidono e quindi sono **equiprobabili**
Sapendo per certo che $\boldsymbol{(X - \mu)^2}$ è *non negativa*
Possiamo applicare **Markov** con $a = r^2$ ottenendo:

$$\begin{aligned} P(|X - \mu| \geq r) &= P((X - \mu)^2 \geq r^2) \\ &\leq \frac{\mathbb{E}[(X - \mu)^2]}{r^2} = \frac{\sigma^2}{r^2} \end{aligned}$$

La disuguaglianza di **Markov** e di **Chebyshev** servono per ottenere le stime di probabilità di eventi rari di variabili cui conosciamo solo la **media** e la **varianza**.

Postilla: in caso di *distribuzione nota* non c'è bisogno di utilizzare una di queste disuguaglianze.

Esempio: I numeri di pezzi prodotti in una settimana è una X di **50**
(a) Cosa si può dire sulla probabilità che la produzione superi i 75 pezzi?

(b) Se è nota anche la varianza pari a **25** cosa si può dire sulla probabilità che la produzione sia compresa tra i *40* e i *60* pezzi?

Risoluzione :
(a) per la disuglianza di *Markov*

$$P(X \geq 75) \leq \frac{\mathbb{E}[X]}{75} = \frac{50}{75} = \frac{2}{3}$$

(b) Applicando la disuguaglianza di *Chebyshev*
Prima di tutto dobbiamo ricavarci $\boldsymbol{r}$, essa è la distanza della media tra *40* e *60*
$\frac{40+60}{2} = 50 \longrightarrow 60 - 50 = \mathbf{10} | 50 - 40 = \mathbf{10}$

$$P(|X - 50| \geq 10) \leq \frac{25}{10^2} = \frac{\mathbf{1}}{\mathbf{4}}$$

Quindi:

$$P(40 \leq X \leq 60) = P(|X - 50| \leq 10) \geq 1 - \frac{1}{4} = \frac{\mathbf{3}}{\mathbf{4}}$$

Perciò la probabilità che la produzione sia compresa tra *40* e i *60* pezzi è almeno del **75%**

8 Legge debole dei grandi numeri

Definizione: Dice che la probabilità che la differenza tra la media campionaria e il valore atteso superi una determinata soglia diventa sempre più piccola all'aumentare del numero di osservazioni

Definizione: Sia $X_1, X_2 \dots X_n$ una successione di variabili aleatorie tutte con la media $\mathbb{E}[\boldsymbol{X_i}] \boldsymbol{=:} \boldsymbol{\mu}$ allora per ogni $\epsilon > 0$

Formula generica:

$$P(|\frac{X_1 + \dots + X_n}{n} - \mu| > \epsilon) \longrightarrow 0$$

$$\text{quando n} \longrightarrow \infty$$

Dimostrazione: Proveremo a dimostrare con l'ipotesi che le X_i hanno varianza finita σ^2 abbiamo che:

$\mathbb{E}[\frac{X_1+\dots+X_n}{n}] = \mu$ $\qquad$ $Var(\frac{X_1+\dots+X_n}{n}) = \frac{\sigma^2}{n}$

La seconda si può trovare in questo modo:

$$\begin{aligned} Var(\frac{X_1 + \dots + X_n}{n}) &= \frac{1}{n^2} Var(X_1 + \dots + X_n) \\ &= \frac{Var(X_1) + \dots + Var(X_n)}{n^2} \\ &= \frac{n\sigma^2}{n^2} = \frac{\sigma^2}{n} \end{aligned}$$

Segue allora dalla disuguaglianza di *Chebyshev* applicata alla variabile aleatoria $(X_1 + \ldots + X_n)/n$ che:

$$P(|\frac{X_1 + \ldots + X_n}{n} - \mu| > \epsilon) \leq \frac{\sigma^2}{n\epsilon^2}$$

Esempio: Supponiamo di ripetere in successione *molte copie indipendenti* di un esperimento ponendo:

$$X_i := \begin{cases} 1 & \text{se E si realizza nell'esperimento } \textit{i-esimo} \\ 0 & \text{se E } \textbf{non} \text{ si realizza nell'esperimento } \textit{i-esimo} \end{cases}$$

La sommatoria $X_1 + X_2 + \ldots + X_n$ rappresenta il numero di prove *tra le prime n* poichè:

$$\mathbb{E}[X_i] = P(X_i = 1) = P(E)$$

si deduce che la frazione delle n prove nelle quali si realizza E, tende (nel senso della legge debole dei grandi numeri) alla probabilità $P(E)$.

9 Modelli di variabili aleatorie

Definizione: Quelle che studieremo ora (porca madonna) sono dei modelli di variabili aleatorie caratterizzate dal fatto che vengono utilizzati da una vasta generalità dei campi applicativi nei quali compaiono e soprattutto usate in natura.

9.1 Bernoulli

Definizione: Una variabile X si dice *bernoulliana* se può essere solo **0** e **1**

Formula generica:

$$P(X = 0) = 1 - p$$

$$P(X = 1) = p$$

Dove con p intendiamo un valore che dovrà essere $\boldsymbol{0 \leq p \leq 1}$
Il suo valore atteso è dato da:

$$\mathbb{E}[X] := 1 \cdot P(X = 1) + 0 \cdot P(X = 0) = p$$

9.2 Binomiali

Definizione: Ipotizziamo che dobbiamo realizzare n ripetizioni di un esperimento Se X è il numero totale di successi e *n* il numero di ripetizioni di un esperimento si dice che abbiamo una *variabile aleatoria binomiale* di parametri (n, p).

La sua funzione di massa è data da:

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

$$i = 0, 1 \ldots, n$$

Dove (ricordiamo) che il coefficiente binomiale è:

$$\binom{n}{i} := \frac{n!}{i!(n-1)!}$$

Spiegazione: Per spiegare le equazioni di sopra dobbiamo fissare una *sequenza di esiti* con i successi e $n-1$ fallimenti.
La probabilità che si verichi questa sequenza è appunto $\boldsymbol{p^i(1-p)^{n-1}}$
Si continua quindi con il contare le sequenze di esiti con questa caratteristica $\binom{n}{i}$
Ad esempio, concludendo, per $n = 5$ e $i = 2$ ci sono $\binom{5}{2} = 10$ scelte possibili.

(s, s, f, f, f)	(s, f, s, f, f)	(s, f, f, s, f)	(s, f, f, f, s)	(f, s, s, f, f)
(f, s, f, s, f)	(f, s, f, f, s)	(f, f, s, s, f)	(f, f, s, f, s)	(f, f, f, s, s)

Se prendiamo in esempio l'esito (f, s, f, s, f) vediamo che i **successi** si sono verificati nelle prove numero *2* e numero *4*.
Ricordiamoci che la *somma delle probabilità* è pari a **1** tramite questa dimostrazione:

Dimostrazione:

$$\sum_i P(X = i) = \sum_{i=0}^{n} \binom{n}{i} p^i (1-p)^{n-1} = [p + (1-p)^n] = 1$$

Esempio: Se X è il numero di pezzi difettosi in 10 dischetti con X di parametri (10, 0.1) quanto è la probabilità che ne vengano ritornati esattamente **una** se ne vengono comprate **3**?

La probabilità che una scatola sia ritornata è pari a:

$$\begin{aligned} P(X > 1) &= 1 - P(X = 0) - P(X = 1) \\ &= 1 - \binom{10}{0} \cdot 0.01^0 \cdot 0.99^{10} - \binom{10}{1} \cdot 0.01^1 \cdot 0.99^9 \approx 0.0043 \end{aligned}$$

Continuo: Ogni scatola viene resa con probabilità *0.43%*
Acquistandone quindi 3 scatole otteniamo una variabile di parametri (3, 0.0043) quindi:

$$\binom{3}{1} \cdot 0.0043^1 \cdot 0.9957^2 \approx 0.013$$

9.2.1 Valore atteso e varianza di Binomiali

Definizione: La varianza di variabili aleatorie binomiali può essere vista come *somma di bernoulliane*.
Quindi se X è binomiale di parametri (n, p) si può scrivere nel seguente modo:

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

Dove X_i è una funzione indicatrice del successo dell'*i-esimo* esperimento:

$$X_i := \begin{cases} 1 & \text{se la prova } \textit{i-esima} \text{ ha successo} \\ 0 & \text{altrimenti} \end{cases}$$

Per X tramite le proprietà di *media* e *varianza* otteniamo che:

$$\mathbb{E}[X] = \sum_{i=1}^{n} \mathbb{E}[X_i] = np$$

$$Var(X) = \sum_{i=1}^{n} Var(X_i) = np(1-p)$$

9.2.2 Funzione di massa e di ripartizione di Binomiali

Definizione: Supponiamo che X sia binomiale sempre di parametri *(n, p)* possiamo calcolare la sua **funzione di ripartizione**

$$P(X \leq i) = \sum_{k=0}^{i} \binom{n}{k} p^k (1-p)^{n-k} = \sum_{k=0}^{i} P(X=k)$$

$$i = 0, 1 \ldots n$$

e la sua funzione di massa:

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

è presente una relazione tra $\boldsymbol{P(X=k+1)}$ e $\boldsymbol{P(K=k)}$:

$$\begin{aligned} P(X=k+1) &= \frac{p}{1-p} \frac{n-k}{k+1} P(X=k) \\ &= \frac{p}{1-p} \frac{n-k}{k+1} \\ &= \frac{n!}{(n-k)!k!} P^{k+1} (1-p)^{n-(k+1)} \\ &= \binom{n}{k+1} P^{k+1} (1-p)^{n-(k+1)} \end{aligned}$$

Esempio: Sia X una variabile aleatoria di parametri $n = 6$ e $p = 0.4$
Iniziando da $P(X = 0) = 0.6^6$ e applicando una ricorsione troviamo che:

$$P(X = 0) = 0.06^6 \approx 0.0467$$

$$P(X = 1) = \frac{4}{6} \cdot \frac{6}{1} \cdot P(X = 0) \approx 0.1866$$

$$P(X = 2) = \frac{4}{6} \cdot \frac{5}{2} \cdot P(X = 1) \approx 0.3110$$

$$P(X = 3) = \frac{4}{6} \cdot \frac{4}{3} \cdot P(X = 2) \approx 0.2765$$

$$P(X = 4) = \frac{4}{6} \cdot \frac{3}{4} \cdot P(X = 3) \approx 0.1382$$

$$P(X = 5) = \frac{4}{6} \cdot \frac{2}{5} \cdot P(X = 4) \approx 0.0369$$

$$P(X = 6) = \frac{4}{6} \cdot \frac{1}{6} \cdot P(X = 5) \approx 0.0041$$

9.3 Poisson

Definizione: Una variabile aleatoria X che assume valori $\boldsymbol{X \in \{1, 2, \ldots n\}}$ viene detta *poissoniana* di parametro $\boldsymbol{\lambda > 0}$
Se la sua *funzione di massa* è data da:

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

$$i = 0, 1, 2, n$$

La funzione sopra è chiaro che rappresenta una funzione di massa accettabile, difatti:

$$\sum_{i=0}^{\infty} P(X=i) = e^{-\lambda} \sum_{i=0}^{\infty} \frac{\lambda^i}{i!} = e^{-\lambda} e^{\lambda} = 1 \leftarrow \text{sviluppo in serie}$$

Per determinare la **media** e la **varianza** dobbiamo prima calcolare la sua *funzione generatrice dei momenti*:

$$\begin{aligned}
\phi(t) :&= \mathbb{E}[e^{tX}] \\
&= \sum_{i=0}^{\infty} e^{ti} P(X=i) \\
&= e^{-\lambda} \sum_{i=0}^{\infty} e^{ti} \frac{\lambda^i}{i!} \\
&= e^{-\lambda} e^{\lambda e^t} = exp\{\lambda(e^t - 1)\}
\end{aligned}$$

Derivando troviamo che:

$$\phi'(t) = \lambda e^t exp\{\lambda(e^t - 1)\}$$

$$\phi''(t) = (\lambda e^t)^2 exp\{\lambda(e^t - 1)\} + \lambda e^t exp\{\lambda(e^t - 1)\}$$

Se valutiamo le due funzioni con il parametro t = 0 ottieniamo che il $\mathbb{E}[X]$ e la $Var(X)$ **coincidono**:

$$\mathbb{E}[X] = \phi'(0) = \lambda$$

$$Var(X) = \phi''(0) - \mathbb{E}[X]^2 = \lambda^2 + \lambda - \lambda^2 = \lambda$$

La Poissoniana può essere usata come approsimazione di una binomiale di parametri (n, p) quando *n* è molto **grande** e *p* è molto **piccolo**.
Per la dimostrazione poniamo $\lambda = np$:

$$
\begin{aligned}
P(X = i) &= \frac{n!}{(n-i)!i!} p^i (1-p)^{n-i} \\
&= \frac{n(n-1)\dots(n-i+1)}{i!} \binom{\lambda}{n}^i (1 - \frac{\lambda}{n})^{n-1} \\
&= \frac{n(n-1)\dots(n-i+1)}{n^i} \cdot \frac{\lambda^i}{i!} \cdot \frac{(1-\lambda/n)^n}{(1-\lambda/n)^i} \\
&= P(X = i) \approx \frac{\lambda^i}{i!} e^{-\lambda}
\end{aligned}
$$

Possiamo dire che l'approssimazione di poisson si può usare per:

- Il numero di persone all'interno di una categoria di persone, che superano i **100** anni di età.
- La quantità di numeri di telefono errati che vengono composti in una giornata.
- Il numero di clienti che entrano in un ufficio postale in un giorno.

Esempio: Se il numero medio di incidenti in un autostrada sia pari a **3**, quanto è la probabilità che la prossima settimana ci sia almeno un incidente?

(se denotiamo il numero di incidenti con X il numero di questi sarà *approsimativamente* distribuito con Poisson di media 3):

$$\begin{aligned} P(X \geq 1) &= 1 - P(X = 0) \\ &= 1 - \frac{3^0}{0!}e^{-3} \\ &= 1 - e^{-3} \approx 0.9502 \end{aligned}$$

La distribuzione di Poisson è *riproducibile*, quindi la somma di due poissoniane è sempre una poissoniana.
Dimostrabile assegnando ai parametri X_1 e X_2 con parametri λ_1 λ_2
e calcolandone la **funzione generatrice dei momenti** della loro somma:

Dimostrazione:

$$\begin{aligned} \phi_{X_1+X_2}(t) &= \phi_{X_1}(t)\phi_{X_2}(t) \\ &= exp\{\lambda_1(e^t - 1)\}exp\{\lambda_2(e^t - 1)\} \\ &= exp\{(\lambda_1\lambda_2)(e^t - 1)\} \end{aligned}$$

Consideriamo N eventi in modo che $N = N_1 + N_2$ con probabilità p e $1 - p$ rispettivamente
Si può calcolare la *funzione di massa* di N_1 e N_2

$$\begin{aligned} P(N_1 = n, N_2 = m) &= P(N_1 = n, N = n + m) \\ &= P(N_1 = n|N = n + m)P(N = n + m) \\ &= P(N_1 = n|N = n + m)\frac{\lambda^{n+m}}{(n + m)!}e^{-\lambda} \end{aligned}$$

Condizionando al fatto che $n + m$ eventi ciascuno ha probabilità p
si scopre che ci siano esattamente *n* eventi di tipo *1*, quindi una binomiale di

parametri (n + m,p)
Quindi otteniamo che:

$$
\begin{aligned}
P(N_1 = n, N_2 = m) &= \binom{n+m}{n} p^n (1-p)^m \frac{\lambda^{n+m}}{(n+m)!} e^{-\lambda} \\
&= \frac{(n+m)!}{n!m!} p^n (1-p)^m \frac{\lambda^{n+m}}{(n+m)!} e^{-\lambda} \\
&= \frac{(\lambda p)^n}{n!} e^{-\lambda p} \frac{(\lambda(1-p))^m}{m!} e^{-\lambda(1-p)}
\end{aligned}
$$

è possibile ora calcolare le **distribuzioni marginali** di N_1 e N_2:

$$
\begin{aligned}
P(N_1 = n) &= \sum_{m=0}^{\infty} P(N_1 = n, N_2 = m) \\
&= \frac{(\lambda p)^n}{n!} e^{-\lambda p} \sum_{m=0}^{\infty} \frac{(\lambda(1-p))^m}{m!} e^{-\lambda(1-p)} \\
&= \frac{(\lambda p)^n}{n!} e^{-\lambda p}
\end{aligned}
$$

e analogamente:

$$
P(N_2 = m) = \sum_{n=0}^{\infty} P(N_1 = n, N_2 = m) = \frac{(\lambda(1-p))^m}{m!} e^{-\lambda(1-p)}
$$

Da queste equazioni segue che N_1 e N_2 sono variabili con distribuzione di Poisson di media λp e $\lambda(1-p)$ rispettivamente.

Definizione: Se N eventi sono classificati in 1,2, ..., r con probabilità p_1,p_2,...p_r (con la loro somma = 1)
allora la quantità totale di eventi sono *variabili di Poisson indipendenti* di medie λp_1,λp_2,...,λp_r

9.3.1 Calcolo esplicito della distribuzione di Poisson

Definizione: Se X è una variabile aleatoria di Poisson di media λ allora:

$$\frac{P(X=i+1)}{P(X=i)} = \frac{\lambda^{i+1}e^{-\lambda}}{(i+1)!}\frac{i!}{\lambda^i e^{-\lambda}} = \frac{\lambda}{i+1}$$

Tramite questa formula possiamo ottenere anche:

$$P(X=i+1) = \frac{\lambda}{i+1}P(X=i)$$

9.4 Ipergeometriche

Definizione: Una variabile aleatoria X che ha come *massa di probabilità* si dice *ipergeometrica* di parametri N, M e n.

Introduzione: Una scatola contiene N batterie *accettabili* e M *difettose*. se si estraggono senza rimessa
e in maniera casuale *n* batterie, con **pari probabilità** a ciascuno degli $\binom{N+M}{n}$ sottoinsiemi.

Formula generica:

$$P(X=i) = \frac{\binom{N}{i}\binom{M}{n-i}}{\binom{N+M}{n}}$$

$$i = 0, 1 \ldots n$$

Esempio: prendiamo a caso *6* componenti da una cassa di *20*. un sistema funziona solamente se tra i 6 componenti non ci siano più di 2 componenti guasti. Se nella cassa ci sono **15** componenti buoni e **5** guasti, quanto è la probabilità che il sistema funzioni?
- Se indichiamo con X il numero di componenti funzionanti tra i 6 estratti, *X è ipergeometrica* di parametri *15, 5 e 6*, quindi:

$$
\begin{aligned}
P(X \geq 4) &= \sum_{i=4}^{6} P(X = i) \\
&= \frac{\binom{15}{4}\binom{5}{2} + \binom{15}{5}\binom{5}{1} + \binom{15}{6}\binom{5}{0}}{\binom{20}{6}} \approx 0.8687
\end{aligned}
$$

9.4.1 Media e varianza delle ipergeometriche

Per determinare la media e la varianza Estrazione solo una volta:

$$
X_i := \begin{cases} 1 & \text{se la } \textit{i-esima} \text{ batteria estratta è accettabile} \\ 0 & \text{altrimenti} \end{cases}
$$

Quindi:

$$
P(X_i = 1) = \frac{N}{N+M}
$$

Dimostrazione:

$$
\begin{aligned}
P(X_i = 1, X_j = 1) &= P(X_j = 1 | X_i = 1) P(X_i = 1) \\
&= \frac{N-1}{N+M-1} \cdot \frac{N}{N+M}
\end{aligned}
$$

Ciascuna delle X_i è una bernoulliana quindi:

$$\mathbb{E}[X_i] = P(X_i = 1) = \frac{N}{N+M}$$

Utilizziamo il fatto che la X è la somme delle X_i per ottenere la **media**

$$\mathbb{E}[X] = \mathbb{E}[\sum_{i=1}^{n} X_i] = \sum_{i=1}^{n} \mathbb{E}[X_i] = n\frac{N}{N+M+}$$

Riprendendo il discorso di prima la formula della **varianza** è la seguente:

$$Var(X_i) = P(X_i = 1)P(X_i = 0) = \frac{NM}{(N+M)^2}$$

Utilizziamo la formula per il calcolo della varianza della somma di variabili aleatorie:

$$Var(X) = \sum_{i=1}^{n} Var(X_i) + 2\sum_{j=2}^{n}\sum_{i<j} Cov(X_i, X_j)$$

$$Var(X) = np(1-p)[1 - \frac{n-1}{N+M-1}]$$

9.4.2 Relazioni tra Binomiali ed Ipergeometriche

Definizione: per $N, M \longrightarrow \infty$ binomiale = ipergeometrica

$$\text{Se} \begin{cases} N, M & \text{Grande} \\ n & \text{Piccolo} \end{cases}$$

Binomiale $\approx$ Ipergeometrica

Differenze la differenza principale tra i due modelli di variabili sta nel caso in esame
se l'estrazione o l'evento **non influenza** la probabilità dell'evento successivo (quindi quando la *probabilità* è uguale per ogni esperimento) allora si usa la **binomiale**.

Se però la probabilità **cambia** dopo ogni esperimento allora si usa una **ipergeometrica**

Nei casi però in cui gli *elementi estratti* sono pochi rispetto *all'insieme totale* una ipergeometrica si può **approssimare** con una binomiale.

Binomiale	**Ipergeometrica**	**Entrambe**
Se lanciamo 10 volte una *moneta* la binomiale rappresenta *il numero di volte che esce testa*	Se abbiamo 10 biglie *6* rosse *4* nere Estrarre 3 biglie e contare quelle rosse.	Estrarre 10 biglietti vincenti da 100 (20 vincenti e 80 perdenti)

Nel caso di entrambe si può approssimare la ipergeometrica con una binomiale con prob. **0.2** e **10** estrazioni totali

9.5 Uniformi

9.5.1 Continue

Definizione: Una variabile aleatoria continua si dice *uniforme* sull'intervallo $[\alpha, \beta]$ se ha la seguente funzione di densità:

$$f(x) = \begin{cases} \frac{1}{\beta-\alpha} & \text{se } \alpha \leq x \leq \beta \\ 0 & \text{altrimenti} \end{cases}$$

Si nota che il grafico di una densità soddisfa le condizioni per essere una *densità di probabilità*:

$$\int_{-\infty}^{\infty} f(x)\,dx = \frac{1}{\beta-\alpha}\int_{\alpha}^{\beta} dx = 1$$

Se $[\boldsymbol{a}, \boldsymbol{b}] \subseteq [\boldsymbol{\alpha}, \boldsymbol{\beta}]$ possiamo ricavare la sua *funzione di ripartizione*:

$$P(a < X < b) = \frac{1}{\beta-\alpha}\int_{a}^{b} dx = \frac{b-a}{\beta-\alpha}$$

è possibile anche ricavare la *media di una variabile aleatoria X* su $[\alpha, \beta]$

$$\begin{aligned}\mathbb{E}[X] :&= \int_{\alpha}^{\beta} \frac{x\,dx}{\beta-\alpha}\\ &= \frac{\beta^2-\alpha^2}{2(\beta-\alpha)}\\ &= \frac{(\beta-\alpha)(\beta+\alpha)}{2(\beta-\alpha)} = \frac{\alpha+\beta}{2}\end{aligned}$$

e la varianza (se abbiamo il **momento secondo**):

$$\begin{aligned}Var(X) &= \frac{\alpha^2+\alpha\beta+\beta^2}{3} - (\frac{\alpha+\beta}{2})^2\\ &= \frac{\alpha^2-2\alpha\beta+\beta^2}{12}\\ &= \frac{(\beta-\alpha)^2}{12}\end{aligned}$$

9.5.2 Discrete

Definizione: se p:

$$p(x) = \begin{cases} \frac{1}{\beta-\alpha+1} & \text{se } \alpha \leq x \leq \beta \quad n = \beta - \alpha + 1 \\ 0 & \text{altrimenti} \end{cases}$$

Possiamo ricavare anche il **valore atteso**:

$$\mathbb{E}[X] = \frac{\beta + \alpha}{2}$$

E la sua **varianza**:

$$Var(X) = \mathbb{E}[X^2] - \mathbb{E}[X]^2 = \frac{(\beta - \alpha + 1)^2 - 1}{12} = \frac{n^2 - 1}{12}$$

9.6 Normali o Gaussiane

Definizione: Una variabile aleatoria X si dice *normale* o *gaussiana*
Di parametri μ e σ^2.
Se X ha funzione di densità data da:

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} exp\{-\frac{(x-\mu)^2}{2\sigma^2}\}$$

$$\forall x \in \mathbb{R}$$

La *funzione generatrice dei momenti* di una gaussiana (parametri μ, σ^2) si può dedurre da questa equazione:

$$\begin{aligned}
\phi(t) :&= \mathbb{E}[e^{tX}] \\
&= \frac{1}{\sqrt{2\pi}\sigma}\int_{-\infty}^{\infty} e^{tX} exp\{-\frac{(x-\mu)^2}{2\sigma^2}\}\,dx \\
&= \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} e^{t(\sigma y+\mu)} e^{\frac{-y^2}{2}}\,dy \\
&= \frac{e^{\mu t}}{\sqrt{2\pi}}\int_{-\infty}^{\infty} exp\{\frac{2\sigma t y - y^2}{2}\}\,dy \\
&= exp\{\mu t + \frac{\sigma^2 t^2}{2}\}\int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} exp\{-\frac{(y-\sigma t)^2}{2}\}\,dy \\
&= exp\{\mu t + \frac{\sigma^2 t^2}{2}\}
\end{aligned}$$

Se deriviamo tutto sto mappazzone otteniamo le seguenti derivate:

$$\begin{aligned}
\phi'(t) &= \left(\mu + \sigma^2 t\right)\exp\left\{\mu t + \frac{\sigma^2 t^2}{2}\right\} \\
\phi''(t) &= \left[\sigma^2 + \left(\mu + \sigma^2 t\right)^2\right]\exp\left\{\mu t + \frac{\sigma^2 t^2}{2}\right\}
\end{aligned} \tag{1}$$

Come ci ricordiamo dalle seguenti funzione generatrici di momenti possiamo ricavarci il *valore atteso* e la *varianza* (in questo caso) di una gaussiana

$$\mathbb{E}[X] = \phi'(0) = \mu$$

$$\mathbb{E}[X^2] = \phi''(0) = \sigma^2\mu^2$$

$$Var(X) = \mathbb{E}[X^2] - \mathbb{E}[X]^2 = \sigma^2$$

Cosi sappiamo che μ σ^2 rappresentano rispettivamente la *media* e la *varianza*

La trasformazione lineare di X (val. al. normale) è a sua volta una **gaussiana**:

$$\text{Per } X \backsim \mathcal{N} \longrightarrow Y = \alpha X + \beta$$

$$\alpha, \beta \text{ costanti e } \alpha \neq 0$$

Y viene detta variabile aleatoria *normale* con media $\boldsymbol{\alpha\mu + \beta}$ e varianza $\alpha^2\sigma^2$
Se $X \backsim \mathcal{N}(\mu, \sigma^2)$ allora:

$$Z := \frac{X - \mu}{\sigma}$$

variabile aletoria *normale* con media 0 e varianza 1 (anche detta **normale standard**)
La sua funzione di ripartizione (indicata con Φ) ha la seguente formula:

$$\Phi(x) := \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-y^2/2} dy, \quad \forall x \in \mathbb{R}$$

$$\text{è uguale a dire } P(X \leq x)$$

Se vogliamo trovare invece $P(X \leq b)$ (se e solo se:)

$$\frac{X - \mu}{\sigma} < \frac{b - \mu}{\sigma}$$

Formula generica: Cosi da avere:

$$\begin{aligned} P(X < b) &= P(\frac{X - \mu}{\sigma} < \frac{b - \mu}{\sigma}) \\ &= P(Z < \frac{b - \mu}{\sigma}) \\ &=: \boldsymbol{\Phi(\frac{b - \mu}{\sigma})} \end{aligned}$$

Formula generica: Con queste due equazioni possiamo fare lo stesso per $\boldsymbol{a < b}$:

$$\begin{aligned}
P(a < X < b) &= P(\frac{a-\mu}{\sigma} < \frac{X-\mu}{\sigma} < \frac{b-\mu}{\sigma}) \\
&= P(\frac{a-\mu}{\sigma} < Z < \frac{b-\mu}{\sigma}) \\
&= P(Z < \frac{b-\mu}{\sigma}) - P(Z < \frac{a-\mu}{\sigma}) \\
&=: \boldsymbol{\Phi(\frac{b-\mu}{\sigma}) - \Phi(\frac{a-\mu}{\sigma})}
\end{aligned}$$

In tutti i casi siamo arrivati sempre ad un $\Phi(x)$, per calcolare il valore effettivo c'è bisogno della tabella che segue qua sotto

Figure 1: Tabella di Φ

z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
0.0	.5000	.5040	.5080	.5120	.5160	.5199	.5239	.5279	.5319	.5359
0.1	.5398	.5438	.5478	.5517	.5557	.5596	.5636	.5675	.5714	.5753
0.2	.5793	.5832	.5871	.5910	.5948	.5987	.6026	.6064	.6103	.6141
0.3	.6179	.6217	.6255	.6293	.6331	.6368	.6406	.6443	.6480	.6517
0.4	.6554	.6591	.6628	.6664	.6700	.6736	.6772	.6808	.6844	.6879
0.5	.6915	.6950	.6985	.7019	.7054	.7088	.7123	.7157	.7190	.7224
0.6	.7257	.7291	.7324	.7357	.7389	.7422	.7454	.7486	.7517	.7549
0.7	.7580	.7611	.7642	.7673	.7704	.7734	.7764	.7794	.7823	.7852
0.8	.7881	.7910	.7939	.7967	.7995	.8023	.8051	.8078	.8106	.8133
0.9	.8159	.8186	.8212	.8238	.8264	.8289	.8315	.8340	.8365	.8389
1.0	.8413	.8438	.8461	.8485	.8508	.8531	.8554	.8577	.8599	.8621
1.1	.8643	.8665	.8686	.8708	.8729	.8749	.8770	.8790	.8810	.8830
1.2	.8849	.8869	.8888	.8907	.8925	.8944	.8962	.8980	.8997	.9015
1.3	.9032	.9049	.9066	.9082	.9099	.9115	.9131	.9147	.9162	.9177
1.4	.9192	.9207	.9222	.9236	.9251	.9265	.9279	.9292	.9306	.9319
1.5	.9332	.9345	.9357	.9370	.9382	.9394	.9406	.9418	.9429	.9441
1.6	.9452	.9463	.9474	.9484	.9495	.9505	.9515	.9525	.9535	.9545
1.7	.9554	.9564	.9573	.9582	.9591	.9599	.9608	.9616	.9625	.9633
1.8	.9641	.9649	.9656	.9664	.9671	.9678	.9686	.9693	.9699	.9706
1.9	.9713	.9719	.9726	.9732	.9738	.9744	.9750	.9756	.9761	.9767
2.0	.9772	.9778	.9783	.9788	.9793	.9798	.9803	.9808	.9812	.9817
2.1	.9821	.9826	.9830	.9834	.9838	.9842	.9846	.9850	.9854	.9857
2.2	.9861	.9864	.9868	.9871	.9875	.9878	.9881	.9884	.9887	.9890
2.3	.9893	.9896	.9898	.9901	.9904	.9906	.9909	.9911	.9913	.9916
2.4	.9918	.9920	.9922	.9925	.9927	.9929	.9931	.9932	.9934	.9936
2.5	.9938	.9940	.9941	.9943	.9945	.9946	.9948	.9949	.9951	.9952
2.6	.9953	.9955	.9956	.9957	.9959	.9960	.9961	.9962	.9963	.9964
2.7	.9965	.9966	.9967	.9968	.9969	.9970	.9971	.9972	.9973	.9974
2.8	.9974	.9975	.9976	.9977	.9977	.9978	.9979	.9979	.9980	.9981
2.9	.9981	.9982	.9982	.9983	.9984	.9984	.9985	.9985	.9986	.9986
3.0	.9987	.9987	.9987	.9988	.9988	.9989	.9989	.9989	.9990	.9990
3.1	.9990	.9991	.9991	.9991	.9992	.9992	.9992	.9992	.9993	.9993
3.2	.9993	.9993	.9994	.9994	.9994	.9994	.9994	.9995	.9995	.9995
3.3	.9995	.9995	.9995	.9996	.9996	.9996	.9996	.9996	.9996	.9997
3.4	.9997	.9997	.9997	.9997	.9997	.9997	.9997	.9997	.9997	.9998
3.5	.9998	.9998	.9998	.9998	.9998	.9998	.9998	.9998	.9998	.9998
3.6	.9998	.9998	.9999	.9999	.9999	.9999	.9999	.9999	.9999	.9999

Esempio: per trovare un valore se dobbiamo trovare $\mathbf{\Phi(1.77)}$ cerco:

1.7 nelle *righe*

0.07 nelle *colonne*

$\Phi(-x)$ è possibile trovare $\Phi(-x)$ usando la *simmetria della distribuzione* rispetto a 0.

$$\begin{aligned}\Phi(-x) &= P(Z < -x)\\ &= P(Z > x)\\ &= 1 - P(Z < x) = 1 - \Phi(x)\end{aligned}$$

Esempio:

$$P(Z < -1) = \Phi(-1) = 1 - \Phi(1) \approx 1 - 0.8413 \approx 0.1587$$

Esempio: Sia X una variabile aleatoria normale media:$\mu = 3$, varianza:$\sigma^2 = 16$ Si trovino **(a)** P(X < 11); **(b)** P(X > -1) **(c)** P(2 < X < 7).

(a)   Poniamo prima di tutto $Z := (X - \mu)/\sigma$

$$\begin{aligned}P(X < 11) &= P(\frac{X-3}{4} < \frac{11-3}{4})\\ &= P(Z < 2)\\ &= \Phi(2) \approx 0.9972\end{aligned}$$

(b)   stesso ragionamento per b | (P > -1) |

$$\begin{aligned}P(X > 1) &= P(\frac{X-3}{4} < \frac{-1-3}{4})\\ &= P(Z > -1)\\ &= P(Z < 1)\\ &= \Phi(1) \approx 0.8413\end{aligned}$$

(c) stesso ragionamento per c | P(2 < X < 7) |

$$\begin{aligned} P(2 < X < 7) &= P(\frac{2-3}{4} < \frac{X-3}{4} < \frac{7-3}{4}) \\ &= P(-1/4 < Z < 1) \\ &= \Phi(1) - \Phi(-0.25) \\ &= \Phi(1) - 1 + \Phi(0.25) \approx 0.4400 \end{aligned}$$

Riproducibilità della distribuzione normale: Dove:
$X_1, X_2 \dots X_n$ sono *aleatorie normali e indipendenti*, X_i ha media μ_i e varianza σ_i^2

La sua funzione generatrice di $\sum_{i=1}^n X_i$ è data da:

$$\begin{aligned} \phi(t) &= E\left[\exp\left\{tX_1 + tX_2 + \ldots + tX_n\right\}\right] \\ &= E\left[e^{tX_1} e^{tX_2} \ldots e^{tX_n}\right] \\ &= \prod_{i=1}^n E\left[e^{tX_i}\right] \\ &= \prod_{i=1}^n \exp\left\{\mu_i t + \frac{\sigma_i^2 t^2}{2}\right\} \\ &= \exp\left\{\bar{\mu} t + \frac{\bar{\sigma}^2 t^2}{2}\right\} \longrightarrow \mathcal{N}(\overline{\mu}, \overline{\sigma}^2) \end{aligned} \tag{2}$$

Dove:

$$\overline{\mu} := \sum_{i=1}^n \mu_i \qquad\qquad \overline{\sigma}^2 := \sum_{i=1}^n \sigma_i^2$$

Semplificazione: Per ogni $\alpha \in (0,1)$ definiamo $\boldsymbol{z_a}$ in modo che:

$$P(Z > z_a) = 1 - \Phi(Z_a) = \alpha$$

Spieghiamo meglio se no non ci capiamo un cazzo.
Definiamo $z_a := \Phi^{-1}(1-\alpha)$ in modo che la probabilità che una *normale standard* assuma un z_a esattamente ad α

Esempio :

$$1 - \Phi(1.645) \approx 0.05 \qquad 1 - \Phi(1.96) \approx 0.025 \qquad 1 - \Phi(2.33) \approx 0.01$$

Diventano uguali a:

$$z_{0.05} \approx (1.645) \qquad z_{0.025} \approx (1.96) \qquad z_{0.01} \approx (2.33)$$

9.7 Esponenziali

Definizione: Una variabile aleatoria continua la cui funzione di densità è data da

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

per $\boldsymbol{\lambda > 0}$ si dice(sesso) ***esponenziale*** con parametro/intensità λ

Definizione: L' *esponenziale* rappresenta la durata di vita di un fenomeno.

Postilla: La $\boldsymbol{\lambda}$ rappresenta *il tasso di decadimento* della probabilità.
Ovvero la **velocità** con cui la probabilità *diminuisce* al cresce del tempo.
Più è grande λ più velocemente la probabilità diminuisce

La sua *funzione di ripartizione* è data da:

$$\begin{aligned} F(x) &= P(X \leq x) \\ &= \int_0^x \lambda e^{-\lambda y}\, dy \\ &= 1 - e^{-\lambda x} \quad x \geq 0 \end{aligned}$$

Come per gli altri modelli possiamo trovare la sua *funzione generatrice dei momenti* e di conseguenza i momenti e la varianza.

$$\begin{aligned} \phi(t) &:= E\left[e^{tX}\right] \\ &= \int_0^\infty e^{tx} \lambda e^{-\lambda x} dx \\ &= \lambda \int_0^\infty e^{-(\lambda - t)x} dx \\ &= \frac{\lambda}{\lambda - t}, \quad t < \lambda \end{aligned} \tag{3}$$

Derivando ϕ otteniamo $\phi'(t)$ e $\phi''(t)$:

$$\phi'(t) = \frac{\lambda}{(\lambda - t)^2}$$

$$\phi''(t) = \frac{2\lambda}{(\lambda - t)^3}$$

Ottenendo in questo modo i soliti valori attesi e la varianza:

$$\mathbb{E}[X] = \phi'(0) = \frac{1}{\lambda}$$

$$\mathbb{E}[X^2] = \phi''(0) = \frac{2}{\lambda^2}$$

$$Var(X) = \mathbb{E}[X^2] - \mathbb{E}[X]^2 = \frac{1}{\lambda^2}$$

Per una variabile aleatoria esponenziale λ è il *reciproco* del valore atteso e la varianza è il *quadrato* di quest'ultimo.

Definizione: La proprietà centrale della distribuzione esponenziale è la sua **assenza di memoria**

Spiegazione: spieghiamo meglio quello scritto prima.
La seguente proprietà ci dice che la probabilità che un evento che si verifichi in un certo lasso di tempo **non dipende** dal tempo trascorso fino a quel momento ma solo dal tempo trascorso a partire da quel momento.

In termini di formula riferendoci ad una variabile aleatoria X intendiamo che:

$$P(X > s + t | X > t) = P(X > s) \quad \forall s, t \geq 0$$

Esempio: il numero di miglia percorse da una macchina prima che la batteria si scarichi è di media *10.000* miglia
Se una persona fa un viaggio di *5.000* miglia
Quale è la probabilità che lo porti a termine senza dover sostituire la batteria? e se la distribuzione non è esponenziale?
- ricordandoci la proprietà *di assenza di memoria della distribuzione esponenziale* il tempo di vita residuo è esponenziale
con intensità $\boldsymbol{\lambda = 1/10}$ e quindi la probabilità cercata è:

$$\begin{aligned} P(\text{vita residua} > 5) &= 1 - F(5) \\ &= e^{-5\lambda} \\ &= e^{-0.5} \approx 0.607 \end{aligned}$$

Se non avessimo saputo che la distribuzione è esponenziale, la probabilità sarebbe stata da questa equazione:

$$\begin{aligned} P(\text{vita residua} > 5) &= P(\text{vita totale} > t + 5 | \text{vita totale} > t) \\ &= \frac{1 - F(t+5)}{1 - F(t)} \end{aligned}$$

Postilla: t è il numero di miglia della batteria fino al momento del viaggio
Quindi senza l'informazione che la nostra distribuzione è esponenziale avremmo **bisogno di ulteriori informazioni**.

Proprietà con condizione in assenza di memoria:

$$\frac{P(X > s + t, X > t)}{P(X > t)} = P(X > s)$$

e quindi anche a:

$$P(X > s + t) = P(X > s)P(X > t)$$

Dimostrazione:

$$P(X > x) = e^{-\lambda x} \to e^{-\lambda(s+t)} = e^{-\lambda s} e^{-\lambda t}$$

Proposizione: se abbiamo $X_1, X_2, \ldots X_n$ *indipendenti* di parametri $\lambda_1, \lambda_2, \ldots \lambda_n$
La variabile aleatoria:
$Y := min(X_1, X_2, \ldots, X_n)$ è **esponenziale** di parametro $\sum_{i=1}^{n} \lambda_i$

Spiegazione: Basta dimostrare che $\boldsymbol{P(Y \leq x) = 1 - exp\{-x \sum_{i=1}^{n} \lambda_i\}}$
quindi che $P(Y > x) = exp\{-x \sum_{i=1}^{n} \lambda_i\}$
e ora la vera dimostrazione che tanto è inutile diomerda.

Dimostrazione:

$$\begin{aligned} P(Y > x) &= P\left(\min\left(X_1, X_2, \ldots, X_n\right) > x\right) \\ &= P\left(X_1 > x, X_2 > x, \ldots, X_n > x\right) \\ &= \prod_{i=1}^{n} P\left(X_i > x\right) \qquad \text{per l'indipendenza} \\ &= \prod_{i=1}^{n} \left(1 - F_{X_i}(x)\right) \\ &= \prod_{i=1}^{n} e^{-\lambda_i x} \\ &= e^{-x \sum_{i=1}^{n} \lambda_i} \end{aligned} \tag{4}$$

9.8 Processi stocastici (Poisson)

Definizione: Famiglia di variabili aleatorie parametrizzate da un indice (in questo caso **t**)

Definizione: Consideriamo una serie di eventi instantanei che avvengono però a intervalli di tempo **random**
Sia $N(t)$ il numero di quanti eventi se ne sono verificati nell'intervallo $[0, t]$
$N(t)$ viene detto **processo di Poisson** di intensità $\lambda, \lambda > 0$

Condizioni:

1. $N(0) = 0 \longrightarrow$ si iniziano a contare gli eventi dal **tempo 0**

2. Il numero degli eventi che hanno luogo in intervalli di tempo *disgiunti* sono **indipendenti**. $\rightarrow$ *indipendenza degli incrementi* | il numero di eventi fino al tempo t -> $N(t)$ è **indipendente** dal numero di eventi tra il tempo t e il tempo $t + s$

3. La distribuzione del numero degli eventi in un dato intervallo di tempo dipende dalla **lunghezza** dell'intervallo $\rightarrow$ *stazionarietà degli incrementi* | la *distribuzione* di $N(t + s) - N(t)$ è **la stessa** per tutti i valori di t

4. $\lim_{h \to 0} \frac{P(N(h) = 1)}{h} = \lambda \rightarrow$ Per un intervallo di tempo *molto piccolo* c'è una probabiltà di λ_h che si **verifica un solo evento**

5. $\lim_{h \to 0} \frac{P(N(h) \geq 2)}{h} = 0 \rightarrow$ Per un intervallo di tempo *molto piccolo* c'è una probabiltà **nulla** che se ne verifichino due o più.

Con queste ipotesi qua di sopra è possibile dimostrare che *il numero di eventi* che si verificano in un qualsiasi intervallo di tempo t è una *variabile aleatoria di Poisson* di media λ_t.

Se n è grande:

$$P(N(t) = k) \approx P(k \text{ sottointervalli con 1 evento, n-k con 0 eventi})$$

Sempre per n grande, la condizione **4** e le condizioni **4** e **5** insieme implicano che:

$$P(1 \text{ evento in un sottointervallo fissato}) \approx \frac{\lambda_t}{n}$$

$$P(0 \text{ eventi in un sottointervallo fissato}) \approx 1 - \frac{\lambda_t}{n}$$

Utilizzando l'indipendenza della condizione 2 (*indipendenza degli incrementi*) il numero totale di eventi è assimilabile ad una variabile aleatoria **binomiale**.

$$P(k \text{ sotto intervalli con 1 evento, n - k con eventi}) \approx \binom{n}{k}(\frac{\lambda_t}{n})^k(1 - \frac{\lambda_t}{n})^{n-k}$$

Se n tende all'infinito può essere *approssimata con Poisson* media λ_t

$$P(N(t) = k) \approx \frac{(\lambda t)^k}{k!}e^{-\lambda t}$$

Proposizione Siano $X_1, X_2, \cdots X_n$ intervalli di tempo che intercorrono rispettivamente dal 1' al 2' al 3' ecc.

Esempio: $X_1 = 5$ e $X_2 = 8$ il primo evento avviene all'istante *5* e il secondo all'istante *13* (5+8)
Vogliamo determinare la distribuzioni delle X_i (ricordando che l'evento $\{X_1 > t\}$ si verifica se nell'intervallo [0, t] *non si sono realizzati eventi*) quindi:

$$P(X_1 > t) = P(N(t) = 0) = e^{\lambda t}$$

Questo significa che:

$$F_{X_1}(t) := P(X_1 \leq t) = 1 - e^{-\lambda t}$$

X_i è una variabile aleatoria *esponenziale* di intensità λ
Per trovare X_2 si noti che qualunque valore s assuma la variabile aleatoria X_1 è data da:

$$\begin{aligned} P(X_2 > t | X_1 = s) &= P(0 \text{ eventi in} (s, s+t | X_1 = s)) \\ &= P(0 \text{ eventi in} (s, s+t)) \quad \text{per la condizione } \mathbf{2} \\ &= e^{-\lambda t} \end{aligned}$$

Questo prova che la variabile aleatoria X_1 è **esponenziale**
e X_2 è esponenziale di intensità λ e **indipendente** da X_1

Proposizione: Le X_i sono tutte *variabili esponenziali*
quindi i tempi che separano gli eventi di Poisson di intensità λ
sono una *successioni di esponenziali indipendenti*

9.9 Gamma

Definizione: Una variabile aleatoria *continua* si dice distribuzione di *tipo gamma* di parametri (α, λ) con $\alpha > 0$ e $\lambda > 0$ la sua funzione di intensità è data da:

$$f(x) = \begin{cases} \frac{\lambda^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\lambda x} & \text{se } x > 0 \\ 0 & \text{se } x \leq 0 \end{cases}$$

dove con $\boldsymbol{\Gamma}$ indichiamo la funzione *gamma di Eulero*, definita in modo da normalizzare l'integrale di f come segue:

$$\begin{aligned} \Gamma(\alpha) &:= \int_0^\infty \lambda^\alpha x^{\alpha-1} e^{-\lambda x} dx \\ &= \int_0^\infty y^{\alpha-1} e^{-y} dy \quad \text{ponendo } y = \lambda x \end{aligned} \tag{5}$$

è possibile **integrare** per parti, se $\boldsymbol{\alpha > 1}$ possiamo scrivere:

$$\begin{aligned} \int_0^\infty y^{\alpha-1} e^{-y} dy &= - y^{\alpha-1} e^{-y} \Big|_{y=0}^{\infty} + \int_0^\infty (\alpha-1) y^{\alpha-2} e^{-y} dy \\ &= (\alpha - 1) \int_0^\infty y^{\alpha-2} e^{-y} dy \end{aligned} \tag{6}$$

Dove il termine $-y^{a-1}e^{-y}\Big|_{y=0}^{\infty}$ è **nullo** perche $\boldsymbol{\alpha > 1}$ implica che $\lim_{y\to 0} y^{\alpha-1} = 0$
Abbiamo dimostrato quindi che:

$$\Gamma(\alpha) = (\alpha - 1)\Gamma(\alpha - 1)$$

La seguente proprietà ci permette di calcolare *per induzione* il valore che Γ assume

$$\Gamma(1) = \int_0^\infty e^{-y}\, dy = 1$$

e anche per una $n \geq 1$:

$$\begin{aligned}
\Gamma(n) &= (n-1)\Gamma(n-1) \\
&= (n-1)(n-2)\Gamma(n-2) \\
&\dots \\
&= (n-1)!\Gamma(1)
\end{aligned}$$

Da cui possiamo dedurre che $\boldsymbol{\Gamma(n) = (n-1)!}$

Possiamo ovviamente ottenere la *funzione generatrice dei momenti* dalla formula:

$$\begin{aligned}
\phi(t) &:= E\left[e^{tX}\right] \\
&= \int_0^\infty e^{tx}\frac{\lambda^\alpha}{\Gamma(\alpha)}x^{\alpha-1}e^{-\lambda x}dx \\
&= \frac{\lambda^\alpha}{\Gamma(\alpha)}\int_0^\infty x^{\alpha-1}e^{-(\lambda-t)x}dx \\
&= \left(\frac{\lambda}{\lambda-t}\right)^\alpha \frac{1}{\Gamma(\alpha)}\int_0^\infty y^{\alpha-1}e^{-y}dy \\
&= \left(\frac{\lambda}{\lambda-t}\right)^\alpha
\end{aligned} \tag{7}$$

Ora deriviamo per ottenere $\phi'(t)$ e $\phi''(t)$:

$$\phi'(t) = \frac{\alpha\lambda^\alpha}{(\lambda-t)^{\alpha+1}}$$

$$\phi''(t) = \frac{\alpha(\alpha+1)\lambda^\alpha}{(\lambda-t)^{\alpha+2}}$$

Ricordiamoci che è possibile ottenere dai momenti il *valore atteso* e la *varianza*:

$$\mathbb{E}[X] = \phi'(0) = \frac{\alpha}{\lambda}$$

$$\mathbb{E}[X^2] = \phi''(0) = \frac{\alpha(\alpha+1)}{\lambda^2}$$

$$Var(X) = \mathbb{E}[X^2] - \mathbb{E}[X]^2 = \frac{\alpha}{\lambda^2}$$

Riproducibilità: Come altre distribuzioni fissando λ possiamo renderle **riproducibili**

Se $X_1 e X_2$ sono variabili aleatorie gamma indipendenti, parametri (α_1, λ) e (α_2, λ) possiamo calcolare la funzione generatrice:

$$\begin{aligned}
\phi(t) &= \mathbb{E}[e^{t(X_1+X_2)}] \\
&= \mathbb{E}[e^{tX_1} e^{tX_2}] \\
&= \mathbb{E}[e^{tX_1}]\mathbb{E}[e^{tX_2}] \\
&= (\frac{\lambda}{\lambda - t})^{\alpha_1} (\frac{\lambda}{\lambda - t})^{\alpha_2} \\
&= (\frac{\lambda}{\lambda - t})^{\alpha_1+\alpha_2}
\end{aligned}$$

L'enunciato segue quindi che ϕ determina la *distribuzione*.

è possibile ovviamente generalizzare alla **somma di due variabili aleatorie**

Proposizione gamma: Se $X_1, i = 1, 2, \ldots, n$ sono variabili indipendenti con parametri gamma $(\alpha_1 + \alpha_2, \lambda)$ allora:

$$\sum_{i=1}^{n} X_i$$

è una gamma di parametri

$$\sum_{i=1}^{n} \alpha_i, \lambda$$

Proposizione esponenziali: Se $X_1, i = 1, 2, \dots, n$ sono variabili aleatorie *esponenziali* di densità λ allora è una gamma di parametri (n, λ):

$$\sum_{i=1}^{n} X_i$$

10 Distribuzioni che derivano da quella normale

10.1 Chi-quadro

Definizione: Se $Z_1, Z_2, \dots Z_n$ sono variabili aleatorie normali standard e indipendenti, la somma dei loro quadrati è:

$$X := Z_1^2 + Z_1^2 + \cdots + Z_n^2$$

Definizione: Viene definita una distribuzione *CHI-QUADRO* quando abbiamo bisogno di valutare se una *differenza* tra più insiemi di dati è statisticamenente significativo, quindi per fare confronti, e la definiamo così:

$$X \sim \chi_n^2 \quad \chi = \text{chi-quadro}$$

Riproducibilità: La distribuzione è *riproducibile* dove X_1 e X_2 sono indipendenti con n_1 n_2 gradi di libertà
Per la distribuzione normale standard, se X è una *chi-quadro* con n gradi di liberta e α $(0 \leq \alpha \leq 1)$ definiamo la quantità $\chi^2_{\alpha,n}$ tramite queste equazione:

$$P(X \geq \chi^2_{\alpha,n}) = \alpha$$

Esempio: Si determini $P(X \leq 30)$ quando X è una aleatoria chi-quadro con **26** gradi di libertà (dal software):

$$P(X \leq 30) \approx 0.7325$$

Esempio 2: Si trova vale $\chi^2_{0.05,15}$ tra le tabelle

$$\chi^2_{0.05,15} \approx 24.996$$

Figure 2: Tabella di $\chi^2_{\alpha,n}$

n \ α	0,001	0,002	0,005	0,01	0,02	0,05	0,1	0,2	0,5	0,75	0,8	0,9	0,95	0,98	0,99	0,995	0,998	0,999
1	0,000	0,000	0,000	0,000	0,001	0,004	0,016	0,064	0,455	1,323	1,642	2,706	3,841	5,412	6,635	7,879	9,550	10,828
2	0,002	0,004	0,010	0,020	0,040	0,103	0,211	0,446	1,386	2,773	3,219	4,605	5,991	7,824	9,210	10,597	12,429	13,816
3	0,024	0,039	0,072	0,115	0,185	0,352	0,584	1,005	2,366	4,108	4,642	6,251	7,815	9,837	11,345	12,838	14,796	16,266
4	0,091	0,129	0,207	0,297	0,429	0,711	1,064	1,649	3,357	5,385	5,989	7,779	9,488	11,668	13,277	14,860	16,924	18,467
5	0,210	0,280	0,412	0,554	0,752	1,145	1,610	2,343	4,351	6,626	7,289	9,236	11,070	13,388	15,086	16,750	18,907	20,515
6	0,381	0,486	0,676	0,872	1,134	1,635	2,204	3,070	5,348	7,841	8,558	10,645	12,592	15,033	16,812	18,548	20,791	22,458
7	0,598	0,741	0,989	1,239	1,564	2,167	2,833	3,822	6,346	9,037	9,803	12,017	14,067	16,622	18,475	20,278	22,601	24,322
8	0,857	1,038	1,344	1,646	2,032	2,733	3,490	4,594	7,344	10,219	11,030	13,362	15,507	18,168	20,090	21,955	24,352	26,124
9	1,152	1,370	1,735	2,088	2,532	3,325	4,168	5,380	8,343	11,389	12,242	14,684	16,919	19,679	21,666	23,589	26,056	27,877
10	1,479	1,734	2,156	2,558	3,059	3,940	4,865	6,179	9,342	12,549	13,442	15,987	18,307	21,161	23,209	25,188	27,722	29,588
11	1,834	2,126	2,603	3,053	3,609	4,575	5,578	6,989	10,341	13,701	14,631	17,275	19,675	22,618	24,725	26,757	29,354	31,264
12	2,214	2,543	3,074	3,571	4,178	5,226	6,304	7,807	11,340	14,845	15,812	18,549	21,026	24,054	26,217	28,300	30,957	32,909
13	2,617	2,982	3,565	4,107	4,765	5,892	7,042	8,634	12,340	15,984	16,985	19,812	22,362	25,472	27,688	29,819	32,535	34,528
14	3,041	3,440	4,075	4,660	5,368	6,571	7,790	9,467	13,339	17,117	18,151	21,064	23,685	26,873	29,141	31,319	34,091	36,123
15	3,483	3,916	4,601	5,229	5,985	7,261	8,547	10,307	14,339	18,245	19,311	22,307	24,996	28,259	30,578	32,801	35,628	37,697
16	3,942	4,408	5,142	5,812	6,614	7,962	9,312	11,152	15,338	19,369	20,465	23,542	26,296	29,633	32,000	34,267	37,146	39,252
17	4,416	4,915	5,697	6,408	7,255	8,672	10,085	12,002	16,338	20,489	21,615	24,769	27,587	30,995	33,409	35,718	38,648	40,790
18	4,905	5,436	6,265	7,015	7,906	9,390	10,865	12,857	17,338	21,605	22,760	25,989	28,869	32,346	34,805	37,156	40,136	42,312
19	5,407	5,969	6,844	7,633	8,567	10,117	11,651	13,716	18,338	22,718	23,900	27,204	30,144	33,687	36,191	38,582	41,610	43,820
20	5,921	6,514	7,434	8,260	9,237	10,851	12,443	14,578	19,337	23,828	25,038	28,412	31,410	35,020	37,566	39,997	43,072	45,315
21	6,447	7,070	8,034	8,897	9,915	11,591	13,240	15,445	20,337	24,935	26,171	29,615	32,671	36,343	38,932	41,401	44,522	46,797
22	6,983	7,636	8,643	9,542	10,600	12,338	14,041	16,314	21,337	26,039	27,301	30,813	33,924	37,659	40,289	42,796	45,962	48,268
23	7,529	8,212	9,260	10,196	11,293	13,091	14,848	17,187	22,337	27,141	28,429	32,007	35,172	38,968	41,638	44,181	47,391	49,728
24	8,085	8,796	9,886	10,856	11,992	13,848	15,659	18,062	23,337	28,241	29,553	33,196	36,415	40,270	42,980	45,559	48,812	51,179
25	8,649	9,389	10,520	11,524	12,697	14,611	16,473	18,940	24,337	29,339	30,675	34,382	37,652	41,566	44,314	46,928	50,223	52,620
26	9,222	9,989	11,160	12,198	13,409	15,379	17,292	19,820	25,336	30,435	31,795	35,563	38,885	42,856	45,642	48,290	51,627	54,052
27	9,803	10,597	11,808	12,879	14,125	16,151	18,114	20,703	26,336	31,528	32,912	36,741	40,113	44,140	46,963	49,645	53,023	55,476
28	10,391	11,212	12,461	13,565	14,847	16,928	18,939	21,588	27,336	32,620	34,027	37,916	41,337	45,419	48,278	50,993	54,411	56,892
29	10,986	11,833	13,121	14,256	15,574	17,708	19,768	22,475	28,336	33,711	35,139	39,087	42,557	46,693	49,588	52,336	55,792	58,301
30	11,588	12,461	13,787	14,953	16,306	18,493	20,599	23,364	29,336	34,800	36,250	40,256	43,773	47,962	50,892	53,672	57,167	59,703
35	14,688	15,686	17,192	18,509	20,027	22,465	24,797	27,836	34,336	40,223	41,778	46,059	49,802	54,244	57,342	60,275	63,955	66,619
40	17,916	19,032	20,707	22,164	23,838	26,509	29,051	32,345	39,335	45,616	47,269	51,805	55,758	60,436	63,691	66,766	70,618	73,402
45	21,251	22,477	24,311	25,901	27,720	30,612	33,350	36,884	44,335	50,985	52,729	57,505	61,656	66,555	69,957	73,166	77,179	80,077
50	24,674	26,006	27,991	29,707	31,664	34,764	37,689	41,449	49,335	56,334	58,164	63,167	67,505	72,613	76,154	79,490	83,657	86,661

Dove: $n =$ *gradi di libertà*

se abbiamo

$\boldsymbol{n = 10}$

$\boldsymbol{\alpha = 0.05}$

cerco:

10 nelle *righe*

0.05 nelle *colonne*

Trovo subito che $\chi^2 = 18.307$

10.2 Distribuzione T

Definizione: Se Z e C_n sono variabili indipendenti, con Z *normale standard* e C_n *chi-quadro* con n gradi di libertà la sua variabile aleatoria T_n è definita:

$$T_n := \frac{Z}{\sqrt{C_n/n}}$$

in questo caso si dice di avere una **distribuzione t** con n gradi di libertà, che denotiamo così:

$$T_n \sim t_n$$

$$C_n \sim X_n^2 \longrightarrow \frac{C_n}{n} = \frac{Z_1^2 + \ldots + Z_n^2}{n}$$

Se applichiamo *la legge dei grandi numeri* otteniamo che se n è grande C_n/n sarà molto vicina a $\mathbb{E}[Z_1^2] = 1$

Dimostrazione di valore atteso e varianza di T_n che sono dati da:

$$\mathbb{E}[T_n] = 0 \quad n \geq 2$$

$$Var(T_n) = \frac{n}{n-2} \quad n \geq 3$$

Se T_n è una t con n gradi di libertà e $\alpha \in (0,1)$ definiamo la quantità $t_{\alpha,n}$ in questo modo:

$$P(T_n \geq t_{\alpha,n}) = \alpha$$

è possibile **applicare la simmetria**:

$$\begin{aligned}
\alpha &= P(-T_n \geq t_{\alpha,n}) \\
&= P(T_n \leq -t_{\alpha,n}) \\
&= 1 - P(T_n > -T_{\alpha,n}) \\
&= P(T_n \geq t_{\alpha,n}) = 1 - \alpha
\end{aligned}$$

Da tutto questo otteniamo qunidi che:

$$-t_{\alpha,n} = t_{1-\alpha,n}$$

10.3 Distribuzione F

Formula generica:

$$F_{n,m} := \frac{C_n/n}{C_m/m}$$

Definizione: Se C_n e C_m sono aleatorie indipendenti, chi-quadro con n e m gradi di libertà
si dice di avere una *distribuzione F* con n e m gradi di libertà

Per ogni $\alpha \in (0,1)$ possiamo definire la quantità $F_{\alpha,n,m}$ in modo:

$$P(F_{n,m} > F_{\alpha,n,m}) = \alpha$$

Se vogliamo trovare una $\alpha > 0.5$ possiamo ottenerla in questo modo:

$$\begin{aligned}\alpha &= P(\frac{C_n/n}{C_m/m} > F_{\alpha,n,m}) \\ &= P(\frac{C_m/m}{C_n/n} < \frac{1}{F_{\alpha,n,m}}) \\ &= 1 - P(\frac{C_m/m}{C_n/n} \geq \frac{1}{F_{\alpha,n,m}})\end{aligned}$$

Facciamola più semplive vah:

$$P(\frac{C_m/m}{C_n/n} > \frac{1}{F_{\alpha,n,m}}) = 1 - \alpha$$

Per trovare invece $F_{1-\alpha,n,m}$ dobbiamo fare così:

$$P(\frac{C_m/m}{C_n/n} > \frac{1}{F_{\alpha,n,m}}) = 1 - \alpha$$

Osservando le due equazioni possiamo notare che:

$$\frac{1}{F_{\alpha,n,m}} = F_{1-\alpha,n,m}$$

Boh vabbe

Esempio: Determiniamo $P(F_{6,14} \leq 1.5)$
Guardando il software si ottiene che la soluzione è **0.752**

10.4 Distribuzione logistica

Definizione: Una variabile **continua** si dice avere una *distribuzione logistica* di parametri $(\mu, \upsilon > 0)$ se la funzione di ripartizione è data da:

$$F(x) = \frac{e^{(x-\mu)/\upsilon}}{1 + e^{(x-\mu)/\upsilon}} \quad \forall x \in \mathbb{R}$$

Se deriviamo $F(x) = 1 - 1/(1 + e^{(x-\mu)/\upsilon})$ troviamo la sua *densità di probabilita:*

$$f(x) = \frac{e^{(x-\mu)/\upsilon}}{\upsilon(1 + e^{(x-\mu)/\upsilon})^2} \quad \forall x \in \mathbb{R}$$

possiamo ricavare la *media* di una logistica **X** di parametri (μ, υ) possiamo procedere in questo modo:

$$\begin{aligned}
\mathbb{E}[X] &= \mu + \upsilon \mathbb{E}[\frac{X-\mu}{\upsilon}] \\
&= \mu + \upsilon \int_{-\infty}^{\infty} \frac{x-\mu}{\upsilon} \cdot \frac{e^{(x-\mu)/\upsilon}}{\upsilon(1+e^{(x-\mu)/\upsilon})^2}\, dx \\
&= \mu + \upsilon \int_{-\infty}^{\infty} \frac{ye^y}{(1+e^y)^2}\, dy \quad \text{ponendo } y = \frac{x-\mu}{\upsilon} \\
&= \mu
\end{aligned}$$

è possibile ottenere anche la *varianza* in questo modo:

$$Var(x) = \frac{\sigma^2}{3}$$

Postilla: In conclusione μ è la **media della distribuzione**
il parametro υ è invece detto *dispersione*.
Una variabile aleatoria logistica con *media* 0 e *dispersione* 1 è detta **logistica standard**.