Probabilità e Statistica

UniVR - Dipartimento di Informatica

Fabio Irimie

2° Semestre 2023/2024

Indice

1 Cos'è la probabilità e la statistica?

La statistica è una scienza che si occupa di raccogliere, organizzare, analizzare e interpretare i dati. Nella statistica si cerca di estrapolare informazioni da esperimenti **aleatori** (esperimenti che non si possono ripetere esattamente allo stesso modo) e di prendere decisioni basate su queste informazioni. Ogni esperimento aleatorio ha bisogno di un **modello probabilistico** che ne descriva le caratteristiche principali.

1.1 Popolazione, variabili e campione

- **Popolazione**: tutti i possibili oggetti di un'indagine statistica
- **Individuo**: un singolo oggetto della popolazione
- **Variabile**: una qualsiasi caratteristica di un individuo della popolazione soggetta a possibili variazioni da individuo a individuo; è l'oggetto di interesse in uno studio
- **Range della variabile**: R_x è l'insieme di tutti i possibili valori che la variabile x può assumere
- **Campione**: un sottoinsieme rappresentativo della popolazione composto dalle variabili relative ad un sottoinsieme di individui
- **Realizzazione del campione di dimension** n: (post esperimento) le osservazioni del campione:

$$\underline{x} = (\tilde{x}_1, \ldots, \tilde{x}_n)$$

- **Range dei dati**: $\mathcal{R}_{\underline{x}}$ i valori che la variabile può assumere tra il minimo e il massimo

1.2 Parametro e Stima

- **Parametro**:una misura che descrive una proprietà dell'intera popolazione
- **Stima**: una misura che descrive una proprietà del campione e che fornisce informazioni sul parametro

1.3 Variabili

Le variabili possono essere di diverso tipo:

- **Variabili qualitative nominali**:
  - **Ordinali**: possono essere ordinate
  - **Non ordinali**: non possono essere ordinate

  I valori che assumono si definiscono anche **modalità**
- **Variabili quantitative**: Sono valori numerici e si distinguono in:
  - **Aleatorie continue**: derivano da processi di misura e assumono i loro range (valori che possono assumere). Sono sottoinsiemi reali
  - **Aleatorie discrete**: derivano da processi di conteggio e assumono valori interi

2 Statistica descrittiva

Consiste nella raccolta, organizzazione, rappresentazione e analisi dei **dati**.

2.1 Strumenti di sintesi

2.1.1 Tabelle di frequenza

Sono tabelle di frequenze di individui con una certa caratteristica o aventi una caratteristica appartenente ad un certo intervallo.

- **Frequenza assoluta**: conteggio del numero di individui
- **Frequenza relativa**: percentuale del numero di individui
- **Frequenza cumulativa**: conteggio o percentuale del numero di individui fino ad un certo punto

2.1.2 Distribuzioni

Sono rappresentazioni del modo in cui diverse **modalità** si distribuiscono tra gli individui di una popolazione.

- **Caso discreto**: f: valore variabile $\to$ frequenza relativa
- **Caso continuo o numerabile**: f: intervallo di valori variabile $\to$ frequenza relativa

2.1.3 Distribuzioni cumulative

Sono distribuzioni che rappresentano la frequenza cumulativa di una variabile. Possono essere:

- **Caso discreto**: f: valore variabile $\to$ frequenza cumulaiva relativa
- **Caso continuo o numerabile**: f: intervallo $\to$ frequenza cumulativa relativa

2.1.4 Grafici

Sono rappresentazioni grafiche delle distribuzioni. Possono essere:

- **Istogrammi**: è costituito da rettangoli, insistenti sulle classi della partizione, attigui le cui aree sono confrontabili con le probabilità.

  $$\text{area rettangolo} \quad i = h_i \cdot |\pi_i| \approx P_X(\pi) \approx f_i$$

  $$h_i = \frac{f_i}{|\pi_i|} \quad \text{per ogni } i \in I$$

  L'area del rettangolo che insiste sulla classe π_i della partizione è pari alla frequenza relativa della classe, quindi l'area torale è 1.

- **Diagrammi a barre**: rappresentano le frequenze di una variabile. Le barre sono separate e la loro altezza è proporzionale alla frequenza

- **Diagrammi a torta**: rappresentano le frequenze relative di una variabile
- **Boxplot**: rappresentano le frequenze di una variabile
- **Poligono di frequenza (ogiva)**: è un grafico a linee continue che ha sull'asse delle ordinate le frequenze cumulative. Questo tipo di grafici è il più comune per rappresentare le frequenze cumulative.

3 Frequenze

Siano $\underline{x} = (\tilde{x}_1, \ldots, \tilde{x}_n)$ una realizzazione del campione di dimensione n e $\mathcal{R}_{\underline{x}}$ il range dei dati. Si dice **partizione** di $\mathcal{R}_{\underline{x}}$:

$$\pi = \{\pi_i\}_{i \in I}$$

La **classe i-esima** è l'elemento i-esimo della partizione

3.1 Frequenze campionarie

3.1.1 Frequenza assoluta

Si dice **frequenza assoluta** n_i per ogni $i \in I$ il numero di osservazioni che appartengono a π_i, cioè:

$$n_i = card(\tilde{x}_j \in \pi_i, \quad j = 1, \ldots, n) \quad \text{(cardinalità)}$$

$$0 \leq n_i \leq n, \text{ per ogni } i \in I \quad e \quad \sum_{i \in I} n_i = n$$

3.1.2 Frequenza relativa

Si dice **frequenza relativa** f_i per ogni $i \in I$ la percentuale delle osservazioni che appartengono a π_i, cioè:

$$f_i = \frac{n_i}{n}$$

$$0 \leq f_i \leq 1, \text{ per ogni } i \in I \quad e \quad \sum_{i \in I} f_i = 1$$

3.2 Frequenze cumulative

3.2.1 Frequenza cumulativa assoluta

Si dice **frequenza cumulativa assoluta** N_i il numero di osservazioni che appartengono alle classi π_h, con $h \leq i$, cioè:

$$N_i = \sum_{h=1}^{i} n_h$$

$$0 \leq N_i \leq n, \text{ per ogni } i \in I \quad e \quad N_i \leq N_j, \ i < j$$

3.2.2 Frequenza cumulativa relativa

Si dice **frequenza cumulativa relativa** F_i della i-esima classe la somma delle frequenze relative delle classi π_h, con $h \leq i$, cioè:

$$F_i = \sum_{h=1}^{i} f_h = \frac{1}{n} N_i = \frac{1}{n} N_{i-1} + f_i$$

$$0 \leq F_i \leq 1, \text{ per ogni } i \in I \quad e \quad F_i \leq F_j,\ i < j$$

4 Statistica descrittiva

4.1 Indici statistici

Sono misure quantitative che fornicono informazioni sulla distribuzione di una certa caratteristica.

4.1.1 Indici di posizione o centralità

Forniscono informazioni del valore attorno al quale si posizionano i dati. Consentono di valutare l'ordine di grandezza della variabile aleatoria e aiutano a "localizzare" la distribuzione. Sono espressi nella stessa unità di misura della variabile.

Sia $\underline{x} = (\tilde{x_1}, \dots, \tilde{x_n})$ un campione di dimensione n.

- **Media campionaria**: è il valore medio dei dati (baricentro dei dati):

$$\overline{x} = \frac{1}{n} \sum_{j=1}^{n} \tilde{x}_j$$

- **Moda campionaria**: m , valore che si ripete più frequentemente. Ci possono essere più valori modali.

  Sia $\underline{y} = (y_1, \dots, y_n)$ il campione ordinato ($y_i \in \{\tilde{x_1}, \dots, \tilde{x_n}\}$ e $y_i \leq y_{i+1}$)

- **Mediana campionaria**:M: è il valore centrale del campione, una volta ordinato.

$$M = \begin{cases} y_{\frac{n+1}{2}} & \text{se } n \text{ è dispari} \\ \frac{1}{2}(y_{\frac{n}{2}} + y_{\frac{n}{2}+1}) & \text{se } n \text{ è pari} \end{cases}$$

4.1.2 Indici di dispersione

Forniscono informazioni su quanto i dati si disperdono attorno ad un valore centrale. Sono:

- **Range**: differenza tra il massimo e il minimo valore:

$$r = \max_{j \in \{1,\dots,n\}} \tilde{x}_j - \min_{j \in \{1,\dots,n\}} \tilde{x}_j$$

- **Scarto Quadratico Medio campionario**: misura la dispersione dei dati attorno alla media

$$s'^2 = \frac{1}{n}\sum_{j=1}^{n}(\tilde{x_j} - \bar{x})^2$$

- **Varianza campionaria**: misura la dispersione dei dati attorno alla media

$$s^2 = \frac{1}{n-1}\sum_{j=1}^{n}(\tilde{x_j} - \bar{x})^2$$

- **Deviazione standard campionaria**: misura la distanza dei dati attorno alla media

$$s = \sqrt{s^2} = \sqrt{\frac{1}{n-1}\sum_{j=1}^{n}(\tilde{x_j} - \bar{x})^2}$$

  Per interpretare la deviazione standard si possono definire **valori usuali** di una variabile i valori del campione compresi tra:

  - **Minimo valore "usuale"**: media campionaria - 2 deviazioni standard
  - **Massimo valore "usuale"**: media campionaria + 2 deviazioni standard

4.1.3 Indici di forma

Sia $\underline{x} = (\tilde{x}_1, \ldots, \tilde{x}_n)$ un campione di dimensione n.

- **Asimmetria / Skewness**: misura la simmetria della distribuzione

$$\gamma_1 = \frac{1}{n-1}\sum_{j=1}^{n}\left(\frac{\tilde{x}_j - \bar{x}}{s}\right)^3$$

  - $\gamma_1 > 0$: distribuzione asimmetrica a destra (con coda più lunga a destra)
  - $\gamma_1 < 0$: distribuzione asimmetrica a sinistra (con coda più lunga a sinistra)
  - $\gamma_1 = 0$: distribuzione simmetrica

- **Curtosi**: misura la "appuntitura" della distribuzione

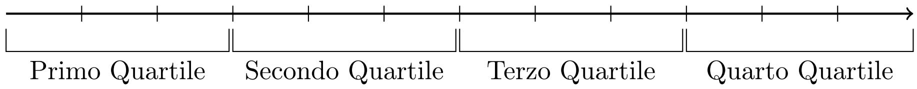
$$\gamma_2 = \frac{1}{n-1}\sum_{j=1}^{n}\left(\frac{\tilde{x}_j - \bar{x}}{s}\right)^4$$

  - $\gamma_2 = 3$: curtosi della normale standard, (variabile di riferimento)
  - $\gamma_2 > 3$: ci sono meno valori agli estremi di quanto aspettato, e di conseguenza si ha una minore dispersione dei dati. In tal caso la distribuzione risulta abbastanza appuntita
  - $\gamma_2 < 3$: ci sono più valori agli estremi di quanto aspettato, e di conseguenza si ha una maggiore dispersione dei dati. In tal caso la distribuzione risulta piatta

4.1.4 Indici di posizione relativi

Rappresentano indici di posizione, ma non centrali, bensì indici di posizionamento relativo.

- **Percentili**: Se p è un numero tra 0 e 100, il **percentile di ordine p** (o p-esimo percentile, se p è intero) è il dato che delimita il primo $p\%$ dei dati (ordinati) dai rimanenti dati.
- **Quartili**: Valori che separano i dati in quattro parti, una volta ordinati.



$$\underline{x} = (\tilde{x_1}, \dots, \tilde{x_n}) \quad \text{campione di dimensione } n$$
$$\underline{y} = (y_1, \dots, y_n) \quad \text{campione ordinato}$$

Il primo quartile è il valore che separa il 25% inferiore dal 75% superiore dei dati.

$$Q_1 = \begin{cases} \frac{y_{\frac{n}{4}} + y_{\frac{n}{4}+1}}{2} & \frac{n}{4} \text{ intero} \\ y_{\lceil \frac{n}{4} \rceil} & \frac{n}{4} \text{ non intero} \end{cases}$$

Il secondo quartile è il 50-esimo percentile, ovvero la mediana. È il valore che separa il 50% inferiore dal 50% superiore dei dati.

$$Q_2 = M = \begin{cases} \frac{y_{\frac{n}{2}} + y_{\frac{n}{2}+1}}{2} & \frac{n}{2} \text{ intero} \\ y_{\lceil \frac{n}{2} \rceil} & \frac{n}{2} \text{ non intero} \end{cases}$$

Il terzo quartil è il 75-esimo percentile, ovvero il valore che separa il 75% inferiore dal 25% superiore dei dati.

$$Q_3 = \begin{cases} \frac{y_{\frac{3n}{4}} + y_{\frac{3n}{4}+1}}{2} & \frac{3n}{4} \text{ intero} \\ y_{\lceil \frac{3n}{4} \rceil} & \frac{3n}{4} \text{ non intero} \end{cases}$$

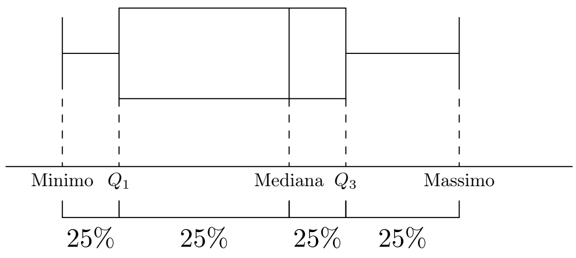
Lo scarto (o distanza interquartile) è la differenza tra il terzo e il primo quartile:

$$IR = Q_3 - Q_1$$

4.1.5 Box-Plot

Fornisce informazioni sulla forma della distribuzione:

Box-Plot



4.2 Outliers

Definizione 4.1
Gli ***Outliers*** *sono valori estremi, insolitamente grandi o piccoli, rispetto al resto dei dati. La loro presenza potrebbe distorcere i risultati dell'analisi, e richiede pertanto un'analisi più accurata.*

$$x \leq Q_1 - 1.5 \cdot IR \quad oppure \quad x \geq Q_3 + 1.5 \cdot IR$$

4.2.1 Outliers deboli

Si dicono outliers deboli:

$$Q_1 - 3 \cdot IR < x \leq Q_1 - 1.5 \cdot IR$$

$$\text{oppure}$$

$$Q_3 + 1.5 \cdot IR < x \leq Q_3 + 3 \cdot IR$$

4.2.2 Outliers forti

Si dicono outliers forti:

$$x \leq Q_1 - 3 \cdot IR$$

$$\text{oppure}$$

$$x \geq Q_3 + 3 \cdot IR$$

5 Statistica descrittiva bivariata

La statistica descrittiva bivariata si occupa di studiare la relazione tra due variabili.

5.1 Relazione tra 2 variabili

- **Correlazione**: Associazione **lineare** tra 2 variabili. La forza dell'associazione è data dal **coefficiente di correlazione**.
- **Regressione**: dipendenza di una variabile (dipendente) da un'altra variabile (indipendente)

5.1.1 Correlazione

Sia $(\underline{x}, \underline{y}) = ((\tilde{x}_1, \tilde{y}_1), \ldots (\tilde{x}_n, \tilde{y}_n))$ un campione di dimensione n di due misure x ed y, con medie campionarie $\bar{x}$ e $\bar{y}$, deviazioni standard campionarie (s_x, s_y).

$$\bar{x} = \frac{1}{n} \sum_{i=1}^{n} \tilde{x}_i$$

$$\bar{y} = \frac{1}{n} \sum_{i=1}^{n} \tilde{y}_i$$

$$s_x = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(\tilde{x}_i - \bar{x})^2}$$

$$s_y = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(\tilde{y}_i - \bar{y})^2}$$

Il coefficiente di correlazione campionario è definito come:

$$\rho_n \triangleq \frac{\sum_{i=1}^{n}(\tilde{x}_i - \bar{x})(\tilde{y}_i - \bar{y})}{\sqrt{\sum_{i=1}^{n}(\tilde{x}_i - \bar{x})^2 \sum_{i=1}^{n}(\tilde{y}_i - \bar{y})^2}}$$

Il risultato sarà un numero compreso tra -1 e 1:

$$|\rho_n| \leq 1$$

Questo indice misura il grado di dipendenza lineare tra le due variabili.

$\lvert\rho_n\rvert$	Grado di correlazione tra $\underline{x}$ e $\underline{y}$
$\rho_n = -1$	massima correlazione lineare inversa
$-1 < \rho_n < 0$	correlazione inversa
$\rho_n = 0$	assenza di correlazione
$0 < \rho_n < 1$	correlazione diretta
$\rho_n = 1$	massima correlazione lineare diretta

Sono indici qualitativi:

ρ_n	Grado di correlazione tra $\underline{x}$ e $\underline{y}$
$\lvert\rho_n\rvert \leq 0.5$	scarsa correlazione
$0.5 < \lvert\rho_n\rvert \leq 0.75$	correlazione moderata
$0.75 < \lvert\rho_n\rvert \leq 0.9$	correlazione buona
$\lvert\rho_n\rvert > 0.9$	correlazione molto buona

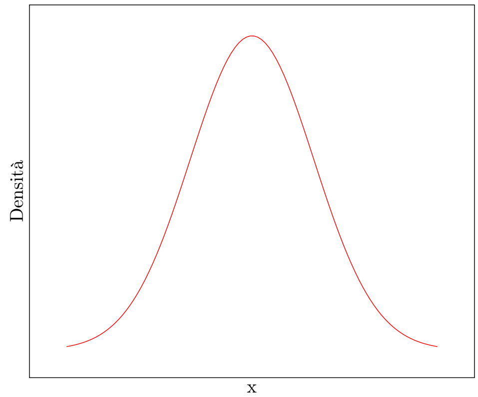
5.1.2 Regressione

La regressione lineare è un modello matematico che cerca di esprimere una variabile. Per ipotesi riteniamo che due variabili siano legate da una relazione del tipo $y = g(x)$

1. I dati accoppiati (x, y) costituiscono un campione di dati quantitativi
2. Dallo scatter plot possiamo ipotizzare che nella **popolazione** ci sia una relazione lineare del tipo:

   $$y_i = \beta_0 + \beta_1 x_i + \varepsilon_i$$

   dove ε_i è l'errore casuale, con distribuzione a campana



3. Cerchiamo di individuare l'equazione della **curba di regressione relativa del campione**:
$$\hat{y}_i = a + bx_i$$

5.1.3 Determinazione dei coefficienti della retta di regressione

L'obiettivo è quello di determinare i coefficienti a e b in modo ottimale, affinchè la retta di regressione $\hat{y}_i = a + bx_i$ sia il più possibile vicina ai punti (x_i, y_i) del campione.

Si determina quindi l'equazione generica della curva interpolante stimando i parametri in modo da rendere **minima** la distanza al quadrato dei punti osservati dalla curva.
$$\sum_{i=1}^{n}(y_i - \hat{y}_i)^2 = \sum_{i=1}^{n}(y_i - a - bx_i)^2$$

Equazioni normali:

$$\begin{cases} \sum_{i=1}^{n} y_i = na + b\sum_{i=1}^{n} x_i \\ \sum_{i=1}^{n} x_i y_i = a\sum_{i=1}^{n} x_i + b\sum_{i=1}^{n} x_i^2 \end{cases}$$

5.2 Riassunto

- Dato **un campione**: abbiamo determinato una **stima** di alcuni parametri (media, deviazione standard, varianza, quartili, ...), una stima della distribuzione (frequqnze relative) con grafici (istogramma [frequenza relativa], diagrammi [area = frequenza relativa], boxplot [quartili, outliers])

- Dati **due campioni**: abbiamo determinato una **stima** di alcuni parametri (media, deviazione standard, varianza, quartili, ...) ed una stima della distribuzione (frequenze relative) con grafici (scatter plot, retta di regressione, coefficiente di correlazione) e abbiamo fatto un **confronto**.

  Abbiamo determinato una stima della **correlazione** e la retta di regressione lineare.
  $$\rho_n = \text{coeff. di correlazione} \quad \rho_n \approx 1$$

Per capire se le informazioni tratte dal campione sono statisticamente significative si fa riferimento alla **statistica inferenziale**. Ma bisogna essere ingrado di parlare di probabilità e di distribuzioni teoriche (modelli probabilistici).

6 Probabilità

La probabilità di un evento $A \in \mathcal{A}$ rappresenta una misura di quanto ci si aspetta che si verifichi l'evento A.

Calcolare le probabilità non significa "prevedere il futuro", ma trovare come distribuire un maggiore o minore **grado di fiducia** tra i vari possibili modi in cui si potrà presentare un certo fenomeno aleatorio.

Definizioni utili 6.1
*L'**ipotesi dei modelli** è lo spazio dei campioni **finito*** $\Leftrightarrow card(\Omega) = n < \infty$
Eventi equiprobabili:

$$P(\omega_i) = P(\omega_j), \quad i, j \in \{1, \dots, n\}$$

La probabilità di un evento $A \in \mathcal{A}$ si calcola come:

$$P(A) = \frac{\text{casi favorevoli ad } A}{\text{casi possibili}} = \frac{card(A)}{card(\Omega)}$$

6.1 Esperimenti aleatori

Un fenomeno **casuale**, o aleatorio, è un fenomeno **osservabile**, ma non prevedibile. Cioè conoscendo i dati iniziali e le leggi, non possiamo prevederne il risultato. Ciò che invece possiamo conoscere è l'insieme di tutti i possibili risultati.

- **Fenomeno deterministico**: Dati + Leggi = Conoscenza
- **Fenomeno non deterministico**: Dati + Leggi = Non Conoscenza

Alcuni esempi di esperimenti sono:

- Consideriamo tre figli di una stessa coppia. Controlliamo il sesso dei tre.
- Lancio un dado. Controllo il numero che esce.
- Lancio 2 dadi. Controllo i numeri che escono.
- Considero i piselli so che possono avere il baccello verde o giallo e il fiore bianco o viola. Ne estraggo uno a caso. Che caratteristiche ha?
- Sono ad un call center. Conto il numero di telefonate che arrivano in un intervallo di tempo
- Misuro all'altezza di un uomo di 40 anni italiano

6.2 Spazio campionario ed eventi

È l'insieme di tutti i possibili risultati di un esperimento casuale:

$$\Omega = \{\omega_1, \omega_2, \ldots, \omega_n\}$$

Uno dei possibili risultati dell'esperimento si chiama **Evento elementare**:

$$\{\omega_i\}, \quad i = 1, \ldots, n$$

L'**Evento** è un sottoinsieme dello spazio campione $A \subset \Omega$ in cui sono contenuti alcuni dei possibili eventi elementari, quelli favorevoli all'evento considerato.

6.3 Esperimenti

6.3.1 Esperimento 1: Lancio di un dado

Prendiamo in considerazione il lancio di un dado:

$$\text{Lo spazio dei campioni è:} \quad \Omega = \{1, 2, 3, 4, 5, 6\}$$

I possibili eventi sono:

A = Il risultato del lancio è 1

B = Il risultato del lancio è dispari

C = Il risultato del lancio è maggiore di 4

D = Il risultato del lancio è dispari non maggiore di 4

E = Il risultato del lancio è pari

F = Il risultato del lancio è 7

G = Il risultato del lancio è tra 1 e 6

$$A = \{1\} \quad B = \{1, 3, 5\}$$

$$C = \{1, 2, 3, 4\} \quad D = \{1, 3, 5\} \bigcap \{1, 2, 3, 4\} = B \bigcap C = \{1, 3\}$$

$$E = \{2, 4, 6\} = \Omega \setminus B = \overline{\{1, 3, 5\}} = \overline{\tilde{B}}$$

$$F = \{7\} = \overline{\Omega} = \emptyset$$

$$G = \{1, 2, 3, 4, 5, 6\} = \Omega$$

6.3.2 Esperimento 2: Lancio di 2 dadi

Prendiamo in considerazione il lancio di 2 dadi:

$$\Omega_2 = \{1, 2, 3, 4, 5, 6\} \times \{1, 2, 3, 4, 5, 6\} = \{(1, 1), (1, 2), \ldots, (6, 6)\}$$

L'evento: A = Esce almeno un 6 è:

$$A = \{(6, 1), (6, 2), \ldots, (6, 6), (1, 6), (2, 6), \ldots, (5, 6)\}$$

6.3.3 Esperimento 3: Sesso dei nascituri

Consideriamo 3 figli di una stessa coppia. Controlliamo il sesso dei tre.

Se considero una **singola nascita** lo spazio dei campioni è:

$$\Omega = \{M, F\}$$

Quindi si hanno due possibili eventi elementari:

$$\{\text{M}\}, \{\text{F}\}$$

Se invece considero **tre nascite** lo spazio dei campioni è:

$$\Omega_3 = \{(\omega_1, \omega_2, \omega_3) \quad \omega_i \in \Omega\}$$

quindi è costituito da tutte le **terne** ordinate di maschi e femmine.

1° Figlio	2° Figlio	3° Figlio
M	M	M
M	M	F
M	F	M
M	F	F
F	M	M
F	M	F
F	F	M
F	F	F

Ogni terna rappresenta un **evento elementare**.

6.3.4 Caratteristiche degli esperimenti 1-3

- Lo spazio dei campioni è finito
- Gli eventi sono tutte le parti di Ω, cioè tutti i possibili sottoinsiemi di Ω

6.3.5 Esperimento 4: Tempo di attesa

Sono ad un call center e conto il numero di telefonate che arrivano in un intervallo di tempo.
Lo spazio dei campioni è:

$$\Omega = \{0, 1, 2, 3, \dots\}$$

Caratteristiche:

- Lo spazio dei campioni è **infinito numerabile**
- Gli eventi sono tutte le parti di Ω, cioè tutti i possibili sottoinsiemi di Ω

6.3.6 Esperimento 5: Misure

Misuro l'altezza di un uomo di 40 anni italiano.
Lo spazio dei campioni è:

$$\Omega \subseteq \mathbb{R}$$

Caratteristiche:

- Lo spazio dei campioni è un sottoinsieme di $\mathbb{R}$, quindi è **infinito non numerabile**
- Gli eventi sono tutti i sottointervalli di $\mathbb{R}$, le loro unioni e le loro intersezioni

6.3.7 Tipi di esperimenti

Gli esperimenti possono essere di diversi tipi:

- **Misure di conteggio**
- **Misure continue**

6.3.8 Tipi di eventi

- **Evento aleatorio**:
  È un sottoinsieme dello spazio campionario, cioè $A \subset \Omega$, ad esempio il lancio di un dado un evento aleatorio potrebbe essere: "esce un numero pari"
- **Evento elementare**:
  È un evento che contiene un solo elemento, cioè $A = \{\omega\}$, ad esempio il lancio di un dado ha come eventi elementari: "esce 1", "esce 2", "esce 3", "esce 4", "esce 5", "esce 6"
- **Eventi complementari**:
  Sono eventi che si escludono a vicenda, ad esempio nel lancio di un dado: E $=$ "esce un numero pari" e $\overline{E}$ $=$ "esce un numero dispari" sono eventi complementari
- **Eventi incompatibili**:
  Sono eventi che non possono verificarsi contemporaneamente, ad esempio nel lancio di un dado: E $=$ "esce un numero pari" e F $=$ "esce il numero 6" sono eventi incompatibili
- **Eventi compatibili**:
  Sono eventi che possono verificarsi contemporaneamente, ad esempio nel lancio di un dado: E $=$ "esce un numero pari" e F $=$ "esce il numero 2" sono eventi compatibili
- **Evento certo**:
  È un evento che si verifica sempre, cioè $A = \Omega$, ad esempio il lancio di un dato ha sempre un risultato certo.
- **Evento impossibile**:
  È un evento che non si verifica mai, cioè $A = \emptyset$, ad esempio il lancio di un dato non può avere come risultato 7.

6.4 Spazio campionario e insieme degli eventi

Definizione 6.1
Lo ***spazio dei campioni*** Ω *è l'insieme di tutti i possibili esiti (risultati). La cardinalità di uno spazio dei campioni può esssere finita, infinita numerabile e infinita non numerabile.*

$$\Omega = \{\omega_1, \omega_2, \ldots, \omega_n\} \quad \textit{oppure} \quad \Omega \subseteq \mathbb{R}$$

Definizione 6.2
*L'****insieme degli eventi*** $\mathcal{A}$ *è un insieme* ***finito*** *di parti di* Ω *tali che sia un'algebra, cioè tale che:*

A_1. $\Omega \in \mathcal{A}$

A_2. *Unione di eventi è un evento*

$$A, B \in \mathcal{A} \Rightarrow A \cup B \in \mathcal{A}$$

A_3. *se* $A, B \in \mathcal{A}$, *allora* $A \setminus B \in \mathcal{A}$

$$\Downarrow$$

$$A \in \mathcal{A} \Rightarrow A^c = \Omega \setminus A \in \mathcal{A}$$

L'insieme degli eventi rappresenta tutti gli eventi che ci ***interessati*** *rispetto all'esperimento preso in considerazione, e che* ***ben descrivono l'esperimento*** *stesso.*

Definizione 6.3
σ*-algebra* $\mathcal{F}$ *è un insieme qualsiasi* $\mathcal{F}$ *di parti di* Ω *tali che:*

A_1. $\Omega \in \mathcal{F}$

A_2^σ. *sia* $\{A_n\}_{n\in\mathbb{N}}$ *con* $A_n \in \mathcal{F}$, *allora* $\bigcup_{n\in\mathbb{N}} A_n \in \mathcal{F}$

A_3^σ. *se* $A, B \in \mathcal{F}$, *allora* $A \setminus B \in \mathcal{F}$ *Diremo* ***evento*** *ogni sottoinsieme* $A \in \mathcal{F}$.

$$\Downarrow$$

$$A \in \mathcal{F} \Rightarrow A^c = \Omega \setminus A \in \mathcal{F}$$

6.4.1 Esempi

Esempio 6.1
Lancio il dado e controllo che numero esce

$$\mathcal{A} = \mathcal{P}(\Omega) =$$

$$= \Big\{\{1\},\{2\},\{3\},\{4\},\{5\},\{6\},$$
$$\{1,2\},\{1,3\},\{1,4\},\{1,5\},\{1,6\},$$
$$\{2,3\},\{2,4\},\{2,5\},\{2,6\},$$
$$\dots$$
$$\{1,2,3\},\{1,2,4\},\{1,2,5\},\{1,2,6\},\dots$$
$$\{1,2,3,4\},\{1,2,3,5\},\{1,2,3,6\},\dots$$
$$\{1,2,3,4,5\},\{1,2,3,4,6\},\dots$$
$$\{1,2,3,4,5,6\} = \Omega, \emptyset\Big\}$$

6.5 Probabilità degli esperimenti 1-2

6.5.1 Esperimento 1: Lancio di un dado

$$\Omega_1 = \{1,2,3,4,5,6\}$$
$$A = \text{ esce almeno un 6}$$
$$P(\{i\}) = \frac{\text{casi favorevoli}}{\text{casi possibili}} = \frac{card(\{i\})}{card(\Omega)} = \frac{1}{6}, \quad i = 1,\dots,6$$

6.5.2 Esperimento 2: Lancio di 2 dadi

$$\Omega_2 = \{1,2,3,4,5,6\} \times \{1,2,3,4,5,6\} = \{(1,1),(1,2),\dots,(6,6)\}$$
$$A = \text{ esce almeno un 6}$$
$$P(A) = \frac{\text{casi favorevoli ad } A}{\text{casi possibili}} = \frac{card(A)}{card(\Omega)} = \frac{11}{36}$$

6.6 Definizione frequentista di probabilità

Definizioni utili 6.2
L'ipotesi dei modelli deve essere ripetibile all'esperimento, quindi bisogna avere tante prove ripetute (nelle stesse condizioni) ed indipendenti

La probabilità di un evento $A \in \mathcal{A}$, fatte n prove:

$$P(A) = \frac{\text{numero di occorrenze di } A}{n} = f_n(A)$$

Si basa sulla **legge empirica del caso** che sintetizza una regolarità osservabile sperimentalmente.

6.7 Definizione soggettiva di probabilità

È la misura del grado di fiducia che un individuo **coerente** assegna al verificarsi di un dato evento in base alle sue **conoscenze**

Probabilità di un evento $A \in \mathcal{A}$:

$$P(A) = \frac{\text{posta}}{\text{vincita}} = \frac{P}{V}$$

In breve, "se ci credo, pago"

7 Assiomi di Kolmogorov

L'impostazione assiomatica permette a Kolmogorov di non esplicitare esattamente come valutare la probabilità (lasciando quindi la libertà di seguire l'approccio più adatto al caso in esame), ma di limitarsi solo a indicare quali sono le regole formali che una misura di probabilità deve soddisfare per poter essere dichiarata tale.

7.1 Assiomi

7.1.1 Caso finito

Definizione 7.1

$$(\Omega, \mathcal{A}, P)$$

P_1. $P(\Omega) = 1$

P_2. *sia* $A, B \in \mathcal{A}$ *disgiunti, t.c*

$$A \cap B = \emptyset$$

allora

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

(additività finita)

7.1.2 Caso generale

Definizione 7.2

$$(\Omega, \mathcal{A}, P)$$

P_1. $P(\Omega) = 1$

P_2^σ. *sia* $\{A_n\}_n, A_n \in \mathcal{F}$ *disgiunti t.c.*

$$A_i \cap A_j = \emptyset, \quad i \neq j$$

allora

$$P\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} P(A_i)$$

(σ-additività)

8 Modelli probabilistici

8.1 Regola per l'unione e l'intersezione di eventi

Sia $(\Omega, \mathcal{A}, P)$ uno spazio di probabilità. Siano $A, B \in \mathcal{A}$ due eventi. Allora:

- **L'unione di** A **e** B:

$$P(A \cup B) = \begin{cases} P(A) + P(B) & \text{se } A \text{ e } B \text{ sono disgiunti} \\ P(A) + P(B) - P(A \cap B) & \text{caso generale} \end{cases}$$

- **L'intersezione di** A **e** B:

$$P(A \cap B) = \begin{cases} P(A)P(B) & \text{se } A \text{ e } B \text{ sono indipendenti} \\ P(A|B)P(B) = P(A)P(B|A) & \text{caso generale} \end{cases}$$

8.2 Modello equiprobabile

L'equiprobabilità è un modello in cui tutti gli eventi elementari hanno la **stessa probabilità** di verificarsi. Ciò implica che lo spazio dei campioni deve essere **finito**.

Definizione 8.1 (Proprietà della probabilità)

P_1. $P(\Omega) = 1$

P_2. *Se* $A, B \in \mathcal{P}(\Omega)$ *tale che:*

$$A \cap B = \emptyset, \text{ allora}$$

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

$$\Downarrow$$

Definizione 8.2 (Probabilità uniforme)

$$1 = P(\Omega) = P(\{1,2,3,4,5,6\}) = P(\cup_i \{\omega_i\}) =$$
$$= \sum_i P(\{\omega_i\}) = card(\Omega) \cdot P(\{\omega_i\})$$

da cui:

$$P(\{\omega_i\}) = \frac{1}{card(\Omega)}, \text{ per ogni } \omega_i \in \Omega$$

Definizione 8.3 (Modello equiprobabile o uniforme)
È lo spazio di probabilità $(\Omega, \mathcal{A}, P)$ tale che:

M_1. *Ω è finito, cioè la cardinalità di Ω ($card(\Omega) \in \mathbb{N}$) è tale che:*

$$card(\Omega) < \infty$$

M_2. *$\mathcal{A} = \mathcal{P}(\Omega)$ è l'insieme delle parti di Ω*

M_3. *per ogni $\omega \in \Omega$*

$$P(\{\omega\}) = costante$$

8.3 Spazio di probabilità

8.3.1 Caso finito

Definizione 8.4 (Spazio di probabilità)
Dato lo spazio dei campioni $\Omega = \{\omega_1, \ldots, \omega_n\}$ e l'insieme degli eventi $\mathcal{A}$, la probabilità P è definita come:

P_1. $P(\Omega) = 1$

P_2. *Se $A, B \in \mathcal{A}$, disgiunti, tale che:*

$$A \cap B = \emptyset, \text{ allora}$$

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

(additività finita)

8.3.2 Caso generale

Definizione 8.5 (Spazio di probabilità)
Dato lo spazio dei campioni $\Omega = \{\omega_1, \ldots, \omega_n\}$*, oppure* $\Omega \subseteq \mathbb{R}$ *e la* σ*-algebra* $\mathcal{F}$*, la probabilità* P *è definita come:*

P_1. $P(\Omega) = 1$

P_2^σ. *sia* $\{A_n\}_n, A_n \in \mathcal{F}$*, disgiunti, tale che:*

$$A_i \cap A_j = \emptyset, \quad \textit{per } i \neq j; \textit{ allora}$$

$$P\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} P(A_i)$$

(σ-additività)

8.4 Indipendenza di eventi

Sia $(\Omega, \mathcal{A}, P)$ uno spazio di probabilità. Siano $A, B \in \mathcal{A}$ due eventi. Allora A e B si dicono **indipendenti** se:

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

In breve due eventi si dicono indipendenti se l'occorrenza di uno dei due non influenza la probabilità di occorrenza dell'altro.

8.4.1 Proposizione

Siano $A, B \in \mathcal{A}$ due eventi indipendenti. Allora:

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

L'intersezione:

$$P(A \cap B) = \begin{cases} P(A)P(B) & \text{se } A \text{ e } B \text{ sono indipendenti} \\ P(A|B)P(B) = P(A)P(B|A) & \text{caso generale} \end{cases}$$

8.5 Probabilità condizionata

Sia $(\Omega, \mathcal{A}, P)$ uno spazio di probabilità **uniforme**. Siano $A, B \in \mathcal{A}$ due eventi. Prendiamo per ipotesi che $P(B) > 0$, quindi B è accaduto, cioè B diventa un evento certo $\Rightarrow$ $prob(B) = 1$. Allora la relazione tra la probabilità P e la probabilità P_B sullo spazio ristretto a B è:

$$P_B(A) = \frac{card_E(A)}{card(\Omega_B)} = \frac{card(A \cap B)}{card(B)} = \frac{card(A \cap B)}{card(B)} \cdot \frac{card(\Omega)}{card(\Omega)} = \frac{P(A \cap B)}{P(B)}$$

Quindi:

Definizione 8.6
La ***Probabilità condizionata a*** B *per ogni* $A \in \Omega$ *è:*

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

8.5.1 Esempio

Esempio 8.1
Consideriamo l'esperimento che consiste nel lancio di un dado equo. Consideriamo i due eventi:

$$A = \text{"esce un numero"} > 3 = \{4,5,6\}$$

$$B = \text{"esce un numero pari"} = \{2,4,6\}$$

Si lanci il dado una volta:

B1. *Calcolare la probabilità dell'evento* A

B2. *Lanciato il dado una persona guarda il risultato e afferma che è uscito un numero pari, cioè è accaduto l'evento* B. *Sapendo questa informazione, come diventa la probabilità dell'evento* A*?*

Il dado è equo, quindi si utilizza il modello equiprobabile.

B1. Calcolare $P(A)$ *:*
Lo spazio dei campioni è:

$$\Omega = \{1,2,3,4,5,6\}$$

La probabilità di A *è:*

$$P(\{\omega\}) = \frac{1}{card(\Omega)} = \frac{1}{6}, \quad \omega \in \Omega$$

$$P(A) = \frac{card(A)}{card(\Omega)} = \frac{3}{6} = \frac{1}{2}$$

B2. Calcolare $P(A|B)$ *: probabilità dell'evento* A *sapendo che sia occorso* B
Lo spazio dei campioni è:

$$\Omega_B = \{2,4,6\}$$

La probabilità di A *sapendo che sia occorso* B *è:*

$$P_B(\{\omega\}) = \frac{1}{card(\Omega_B)} = \frac{1}{3}, \quad \omega \in \Omega_B$$

$$P_B(A) = \frac{card(A)}{card(\Omega_B)} = \frac{2}{3}$$

Condizionare ad un evento significa costruire un nuovo spazio di probabilità ristretto all'evento condizionante:

$$\textit{Spazio dei campioni: } \Omega \rightarrow \Omega_B = B$$

$$\textit{Misura di probabilità: } P \rightarrow P_B : P_B(\Omega_B) = \frac{card(\Omega_B)}{card(\Omega_B)} = 1 = P(\Omega)$$

La relazione tra P e P_B:

$$P_B(A) = \frac{card_B(A)}{card(\Omega_B)} = \frac{card(A \cap B)}{card(B)} = \frac{card(A \cap B)}{card(\Omega)} \frac{card(\Omega)}{card(B)} = \frac{P(A \cap B)}{P(B)}$$

8.5.2 Proposizione

$P_B(\cdot) = P(\cdot|B)$ è una misura di **probabilità** su $(\Omega, \mathcal{A})$ tale che:

$$P_B(A) = P(A|B) = \frac{P(A \cap B)}{P(B)} \quad A \cap B \neq \emptyset$$

$$P_B(A) = P(A|B) = \frac{P(A \cap B)}{P(B)} = 0 \quad A \cap B = \emptyset$$

$$P_B(\Omega) = P(\Omega|B) = \frac{P(\Omega \cap B)}{P(B)} = 1$$

$$P_B(B) = P(B|B) = \frac{P(B \cap B)}{P(B)} = 1$$

L'ultima proprietà è coerente con l'ipotesi che B è occorso e quindi è un evento certo.

8.5.3 Come determinare la probabilità condizionata

Sia $(\Omega, \mathcal{A}, P)$ uno spazio di probabilità **uniforme** e sia $B \subset \Omega$ tale che $P(B) > 0$. Condizionare ad un evento B significa costruire un nuovo spazio di probabilità ristretto all'evento condizionante B: $(\Omega_B, \mathcal{F}_B, P_B)$

$$\text{Spazio dei campioni: } \Omega \rightarrow \Omega_B = B$$

$$\text{Misura di probabilità: } P \rightarrow P_B : P_B(B) = 1$$

Si utilizza lo spazio di partenza $(\Omega, \mathcal{A}, P)$ su cui definisco la misura condizionata, per ogni $A \subseteq \Omega$:

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

8.6 Probabilità a Priori e a Posteriori (Formula di Bayes)

Sia $(\Omega, \mathcal{A}, P)$ uno spazio di probabilità. Si dice **partizione di** Ω un insieme di eventi $\{E_j\}_{1 \leq j \leq n}$ tali che:

1. sono disgiunti:

$$i, j \in I \quad \text{tale che} \quad i \neq j,\ E_i \cap E_j = \emptyset$$

2. per ogni $i \in I$:

$$P(E_i) > 0$$

3. ricoprono tutto lo spazio:

$$\bigcup_{j=1}^{n} E_j = \Omega$$

8.6.1 Teorema delle probabilità totali

Sia $(\Omega, \mathcal{A}, P)$ uno spazio di probabilità e sia $\{E_j\}_{1 \leq j \leq n}$ una partizione di Ω. Sia $B \in \mathcal{A}$ un evento. Allora:

$$P(B) = \sum_{j=1}^{n} P(B|E_j)P(E_j)$$

8.6.2 Dimostrazione

$$B = B \cap \Omega = B \cap \bigcup_{j=1}^{n} E_j = \bigcup_{j=1}^{n} (B \cap E_j)$$

$$P(B) = P\left(\bigcup_{j=1}^{n} (B \cap E_j)\right) \sum_{j=1}^{n} P(B \cap E_j) = \sum_{j=1}^{n} P(B|E_j)P(E_j)$$

8.6.3 Teorema di Bayes

Definizione 8.7
Sia $\{E_i\}_{i \in I}$ una partizione di Ω (finita o numerabile) e sia $P(A) > 0$. Allora per un generico elemento della partizione E_n, con $n \in I$, si ha:

$$P(E_n|A) = \frac{P(A|E_n)P(E_n)}{\sum_{i \in I} P(A|E_i)P(E_i)}$$

8.6.4 Dimostrazione

$$P(E_n|A) = \frac{P(A \cap E_n)}{P(A)} = \frac{P(A|E_n)P(E_n)}{P(A)} = \frac{P(A|E_n)P(E_n)}{\sum_{i \in I} P(A|E_i)P(E_i)}$$

8.7 Probabilità a priori e a posteriori

8.7.1 Probabilità a priori

$P(A)$: è una probabilità da determinare senza altre informazioni. Non si hanno informaizoni sufficienti per determinarla.

8.7.2 Probabilità a posteriori

$P(A|E_n)$: si hanno strumenti ed informazioni sufficienti per determinare queste probabilità. Se si conoscono delle altre informazioni, quindi se si restringe lo spazio campionario, si riescono a determinare.

8.8 Esercizi

8.8.1 Esercizio 1

Una popolazione si compone per un 40% di fumatori (F) e per il restante 60% di non fumatori (N). Si sa che il 25% dei fumatori e il 7% dei non fumatori ha una malattia respiratoria cronica (M) .

1. Calcolare la probabilità che un individuo scelto a caso sia effetto dalla malattia respiratoria
2. Se l'individuo scelto è affetto dalla malattia, calcolare la probabilità che sia un fumatore.

$$\Omega = \{F, F^c\} \quad \text{partizione}$$
$$M = \text{"Malattia respiratoria"}$$
$$P(M|F) = 0.25 \quad P(M|F^c) = 0.07$$
$$P(M) = P(M|F)P(F) + P(M|F^c)P(F^c) = 0.25 \cdot 0.4 + 0.07 \cdot 0.6 = 0.14$$
$$P(F|M) \overset{\text{Bayes}}{=} \frac{P(M|F)P(F)}{P(M)} = \frac{0.25 \cdot 0.4}{0.14} = 0.70$$

9 Variabili aleatorie

Sia $(\Omega, \mathcal{F}, P)$ uno spazio di probabilità che modellizza un certo esperimento. Una **variabile aleatoria** è una funzione che associa un numero ad ogni esito possibile dell'esperimento. Formalmente:

$$X : \Omega \to \mathbb{R}$$

Sia $\mathcal{B}$ l'insieme degli eventi su $\mathbb{R}$ (intervalli, unione ed intersezione di intervalli), per ogni $A \subset \mathbb{R}$:

$$X^{-1}(A) \in \mathcal{F}$$

è l'inversa di X su A.

Le variabili aleatorie possono essere:

- **Discrete**: se assumono un numero finito o numerabile di valori. Ad esempio:
$$X \rightsquigarrow \{0, 1, 2, 3, 4, 5, 6\}$$
- **Continue**: se assumono valori in un intervallo reale. Ad esempio:
$$X \rightsquigarrow \mathbb{R}$$

9.1 Distribuzione di una variabile aleatoria

Sia $(\Omega, \mathcal{F}, P)$ uno spazio di probabilità che modella un certo esperimento e sia $X : \Omega \to \mathbb{R}$ una variabile aleatoria. Sia $\mathcal{B}$ l'insieme degli eventi su $\mathbb{R}$ (intervalli, unione ed intersezione di intervalli).

La probabilità di un evento nella funzione X prendendo in considerazione il lancio di un dado si indica come:

$$P(1) = P(X = 1) = P(\{\omega \in \Omega : X(\omega) = 1\}) = \frac{1}{6}$$

L'insieme $\{P(1), P(2), \dots, P(6)\}$ rappresenta le densità di probabilità, chiamato anche **distribuzione di** X

9.2 Variabili aleatorie discrete

Sia $(\Omega, \mathcal{B}, P)$ uno spazio di probabillità, si definisce **variabile aleatoria discreta** :

$$X : \Omega \to R_X = \{x_1, x_2, \dots, x_n\}$$

con $card(R_X)$ al più numerabile.

Definizione 9.1
Distribuzione o legge di X

$$P_X(\{x_i\}) = P(X^{-1}(\{x_i\})) = P(X = x_i) \quad \text{per ogni } x_i,\ i = 1, \dots, n$$

La legge di una variabile aleatoria discreta è dunque caratterizzata da una ***funzione di probabilità***

Esempio 9.1
Prendiamo in considerazione il lancio di 2 dadi: lancio i 2 dadi e ne sommo i valori.

$$\Omega = \{1, 2, 3, 4, 5, 6\} \times \{1, 2, 3, 4, 5, 6\}$$

$$\mathcal{F} = \mathcal{P}(\Omega)$$

$$P(\omega) = \frac{1}{36}$$

$$X : \Omega \to \{2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12\} \subset \mathbb{N}$$

$$(x, y) \to x + y$$

La legge di X è:

x_i	$X^{-1}(x_i)$	$P_X(x_i)$
2	$\{(1,1)\}$	1/36
3	$\{(1,2),(2,1)\}$	2/36
4	$\{(1,3),(2,2),(3,1)\}$	3/36
5	$\{(1,4),(2,3),(3,2),(4,1)\}$	4/36
6	$\{(1,5),(2,4),(3,3),(4,2),(5,1)\}$	5/36
7	$\{(1,6),(2,5),(3,4),(4,3),(5,2),(6,1)\}$	6/36
8	$\{(2,6),(3,5),(4,4),(5,3),(6,2)\}$	5/36
9	$\{(3,6),(4,5),(5,4),(6,3)\}$	4/36
10	$\{(4,6),(5,5),(6,4)\}$	3/36
11	$\{(5,6),(6,5)\}$	2/36
12	$\{(6,6)\}$	1/36

La rappresentazione grafica della distribuzione di X è:

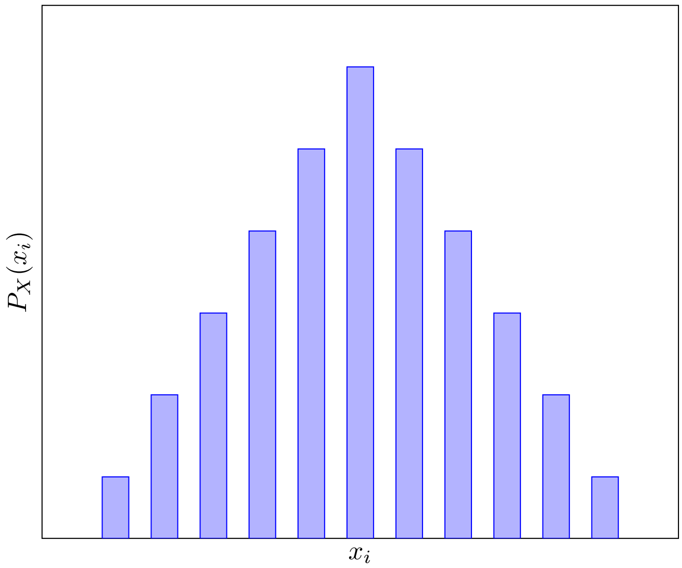


Figura 1: Distribuzione di X

cioè una distrubuzione triangolare.

Definizione 9.2
*Il **Valore atteso** $\mathbb{E}[X]$ è definito come:*

$$\mathbb{E}[X] = \sum_{i=1}^{n} x_i P_X(x_i)$$

Si tratta di una media pesata dei valori assunti dalla variabile aleatoria con i pesi dati dalle probabilità che la variabile assuma quei valori.

Definizione 9.3
*La **Varianza** di X è definita come:*

$$Var(X) = \mathbb{E}[(X - \mathbb{E}[X])^2] = \sum_{i=1}^{n}(x_i - \mathbb{E}[X])^2 P_X(x_i)$$

Esempio 9.2
Prendiamo in considerazione il lancio di una moneta:

$$\Omega = \{T, C\} \quad \mathcal{F} = \mathcal{P}(\Omega)$$

Si vuole calcolare swe esce testa, quindi X è la variabile aleatoria che rappresenta il lancio della moneta:

$$X : \Omega \to \{0, 1\} \quad t.c. \; X(T) = 1, \quad X(C) = 0$$

*Questa variabile viene chiamata **variabile aleatoria di Bernoulli** ed è una variabile **dicotomica** (solo 2 valori).*

- ***Moneta equa:***
$$P_1(\{T\}) = P_1(\{C\}) = \frac{1}{2}$$
$$\Downarrow$$
$$P_X(1) = P_1(X = 1) = P_1(\{T\}) = \frac{1}{2}$$
$$P_X(0) = \frac{1}{2}$$

- ***Moneta non equa:***
$$P_2(T) = \frac{1}{3} \quad P_2(C) = \frac{2}{3}$$
$$P_X(1) = P_2(X = 1) = P_2(\{T\}) = \frac{1}{3}$$
$$P_X(0) = 1 - P_X(1) = \frac{2}{3}$$

10 Schema di Bernoulli

10.1 Prove dicotomiche ripetute ed indipendenti

TODO

10.2 Conteggio del successo in ciascuna prova

Definizione 10.1 (MOLTO IMPORTANTE)

$$X \sim \mathcal{B}(p) \quad X \text{ è una variabile di Bernoulli}$$

$$p \equiv \textit{probabilità di successo}$$

$$\mathbb{E}[X] = p$$

$$Var(X) = p(1-p)$$

10.3 Conteggio del successo in cuascuna delle n volte

Si ripete l'esperimento dicotomico n volte in condizioni di indipendenza (lancio la moneta n volte). Chiamo $X_1, X_2, \ldots, X_n$ i risultati degli n lanci.

$$X_i \sim \mathcal{B}(p) \quad \text{per ogni } i = 1, \ldots, n$$

Si è interessati all'evento:

$$A = \text{"numero di successi ottenuti negli } n \text{ lanci"} = \text{"numero di teste"}$$

La variabile che conta il numero di successi è:

$$X = X_1 + X_2 + \ldots + X_n \equiv \text{"conta il numero di successi"}$$

$$X \sim \mathcal{B}(n,p) \quad \textbf{variabile binomiale}$$

La variabile binomiale conta il numero di successi con:

$$n \equiv \text{numero di prove}$$

$$p \equiv \text{probabilità di successo in ogni prova}$$

I valori che potrà assumere la variabile binomiale sono:

$$X \rightsquigarrow \{0, 1, 2, \ldots, n\}$$

10.3.1 Modello per il conteggio del successo

Esito prova i-esima:

$$X_1 \sim \mathcal{B}(p) \quad P_X(x_i) = P_i^x (1-p)^{1-x_i}$$

perchè:

$$P_x(0) = P^0(1-p)^1 = (1-p) \quad \text{"insuccesso"}$$
$$P_x(1) = P^1(1-p)^0 = p \quad \text{"successo"}$$

Esito delle n prove: Vettore aleatorio $X_1, X_2, \ldots, X_n$ che assume valori:

$$(x_1, x_2, \ldots, x_n) \quad x_j \in \{0, 1\}$$

Prove indipendenti (Distribuzione della variabile binomiale):

$$
\begin{aligned}
P(X_1 = x_1, X_2 = x_2, \ldots, X_n = x_n) &= P(X_1 = x_1)P(X_2 = x_2)\ldots P(X_n = x_n) \\
&= P^{x_1}(1-p)^{1-x_1} P^{x_2}(1-p)^{1-x_2} \ldots P^{x_n}(1-p)^{1-x_n} \\
&= p^{\sum_{i=1}^{n} x_i}(1-p)^{n-\sum_{i=1}^{n} x_i}
\end{aligned}
$$

dove $\sum_{i=1}^{n} x_i = k$ è il numero di successi ottenuti nei n lanci.

$$P_X(X = n) \binom{n}{k} = \frac{n!}{k!(n-k)!} \quad \textbf{coefficiente binomiale}$$

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

è la probabilità di ottenere k successi in n prove.