Calcolo delle probabilità e statistica (04642)

Prof. Fabrizio Caselli - fabrizio.caselli@unibo.it A.A. 2012/2013
Appunti di Nicola Gentili – nicola.gentili2@studio.unibo.it

Sommario

Statistica descrittiva

La statistica descrittiva ha il compito di presentare i dati raccolti in una certa indagine in modo sintetico, comunicativo e rappresentativo.

Definizioni

Si definisce **POPOLAZIONE**, un insieme di elementi oggetto dell'indagine statistica. Ogni elemento della popolazione è un'**UNITÀ** o **UNITÀ STATISTICA**.

La proprietà che l'indagine statistica vuole studiare è una funzione che ha come dominio la popolazione, ed è detta **CARATTERE**.

I caratteri posso essere:

- Qualitativi (ad esempio caratteristiche quali il colore, la razza, il gruppo sanguigno, …)
- Quantitativi (rappresentano una misura del carattere)
  - Discreti (ad esempio l'età, …)
  - Continui (ad esempio l'altezza, il peso, …)

Entrambi possono essere rappresentati da numeri, ma solo per i caratteri quantitativi hanno senso determinate operazioni (ad esempio il calcolo della media).

Sia S un sottoinsieme infinito di $\mathbb{R}$, allora:

i. Se S è unione di intervalli (eventualmente illimitati) allora si dice che S ha **CARDINALITÀ DEL CONTINUO**
Es: $[0,1]$; $[0,+\infty)$; $[-1,1) \cup [2, 5]$; $\mathbb{R}$
ii. Se S si può esprimere nella forma $S = \{ x_1, x_2, \dots \}$ allora si dice che S ha **CARDINALITÀ NUMERABILE**
Es: $\mathbb{N}$; { primi }; $\mathbb{Z}$; $\{ 1/n, n = 1, 2, \dots \}$; $\mathbb{Q}$

Un insieme finito o numerabile si dice discreto.

Se S è finito, allora può essere espresso nella forma: $S = \{ x_1, x_2, \dots, x_n \}$

Se S è numerabile, allora può essere espresso nella forma: $S = \{ x_1, x_2, \dots, x_n, \dots \} = \{ x_1, x_2, \dots \}$

I valori che i caratteri qualitativi o quantitativi discreti possono assumere (codominio della funzione carattere) si dicono **MODALITÀ**.

CARATTERE : POPOLAZIONE → MODALITÀ

Esempi:

Popolazione	Carattere	Modalità	
{ studenti corso di laurea }	Materia preferita	{ materie }	qualitativo
	Voto di maturità	{ voti }	quantitativo
{ corsi di lauera }	Rapporto tra laureati e iscritti	[0, 1]	quantitativo
{ numeri interi }	La prima lettera del nome	{ lettere }	qualitativo
{ possibili lanci di una moneta }	Esito (testa o croce	{ testa, croce }	qualitativo
	Tempi di arresto	$\mathbb{Q}$	quantitativo

Frequenze relative

Considerando un **CAMPIONE** (cioè un insieme di dati) $x_1, \ldots, x_n$, (contenente n osservazione, ovvero avente cardinalità n) proveniente dall'osservazione di un carattere avente un numero finito di modalità (1, 2, …, K), si chiama **FREQUENZA ASSOLUTA** o **EFFETTIVO** della k-esima modalità il numero:

$$N_k = \#\{j; x_j = k\},\ \ k = 1, \ldots, K$$

L'effettivo N_k è la cardinalità dell'insieme degli individui del campione che assumono il valore k (ovvero il numero di osservazioni che assumono il valore k).

Si dice **FREQUENZA RELATIVA** della k-esima modalità, i numero:

$$p_k = \frac{N_k}{n},\ \ k = 1, \ldots, K$$

Essa è la proporzione tra il numero di osservazioni che assumono il valore k e il numero di individui del campione.

Si ha che:

$$N_1 + N_2 + \ldots + N_k = n \qquad e \qquad p_1 + p_2 + \ldots + p_k = 1$$

<u>Se il carattere è di tipo quantitativo continuo</u>, non è possibile praticare queste procedure, si può però suddividere l'intervallo [m,M] dei valori assunti dalle osservazioni in un insieme di sottointervalli $I_1, \ldots, I_K$.

Ogni sottointervallo è detto **CLASSE**. Ogni classe è descritta dai suoi confini. L'**AMPIEZZA** di una classe è la differenza tra il confine superiore e il confine inferiore. Il **VALORE MEDIO DI UNA CLASSE** è la media dei suoi confini.

Allora:

$$N_k = \#\{j; x_j = I_k\} \qquad e \qquad p_k = \frac{N_k}{n},\ \ k = 1, \ldots, K$$

In questo caso N_k e p_k sono gli effettivi e le frequenze relative dell'intervallo I_k.

Istogrammi

Uno strumento utilizzato dalla statistica descrittiva per rappresentare i risultati di un'indagine è costituito dagli istogrammi.

- Gli istogrammi sono grafici costituiti da tanti rettangoli quanti sono le modalità o le classi modali.
- La superficie di ogni rettangolo è proporzionale alla frequenza della modalità o della classe.
- La larghezza è costante nel caso delle modalità, dipende invece dall'ampiezza nel caso di un'indagine su classi, in quest'ultimo caso l'altezza del rettangolo è una misura della concentrazione o densità delle classi.

Esempio

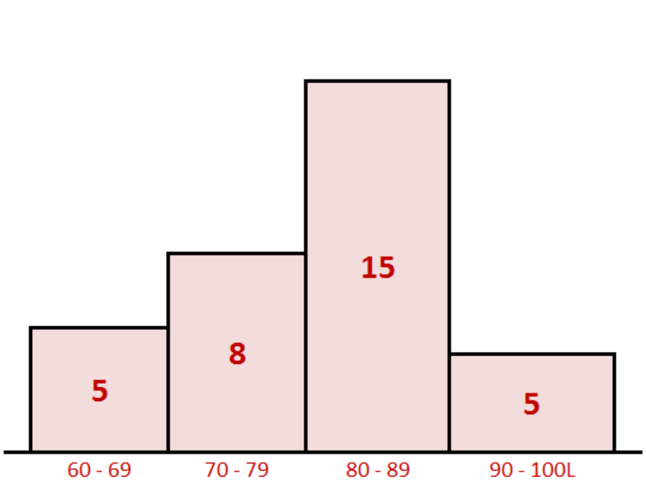
Campione: *studenti di un corso di laurea*

Carattere: *voto di maturità*

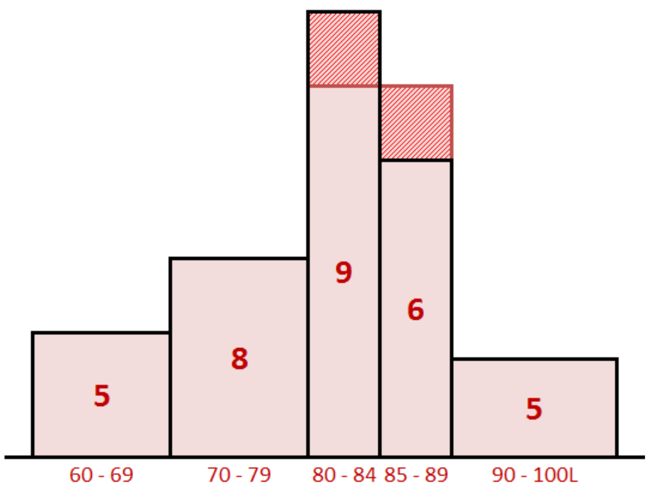
Classi	Ampiezza	Classi (con correzione di continuità)	Ampiezza	Valore medio	Frequenza assoluta	Frequenza relativa	Istogramma A
60 – 69	9	59.5 – 69.5	10	64.5	5	5/33	
70 – 79	9	69.5 – 79.5	10	74.5	8	8/33	
80 – 89	9	79.5 – 89.5	10	84.5	15	15/33	
90 – 100L	11	89.5 – 101.5	12	95.5	5	5/33	

Classi	Ampiezza	Classi (con correzione di continuità)	Ampiezza	Valore medio	Frequenza assoluta	Frequenza relativa	Istogramma B
60 – 69	9	59.5 – 69.5	10	64.5	5	5/33	
70 – 79	9	69.5 – 79.5	10	74.5	8	8/33	
80 – 84	4	79.5 – 84.5	5	82	9	9/33	
85 – 90	5	84.5 – 89.5	5	87	6	6/33	
90 – 100L	11	89.5 – 101.5	12	95.5	5	5/33	

Istogramma A



Istogramma B



Nel caso di caratteri continui, inoltre, si deve considerare che la scelta delle classi I_k e delle loro ampiezze, può determinare un istogramma troppo grossolano o troppo dettagliato evidenziando variazioni poco interessanti. Il problema di determinare quale sia la migliore ampiezza non è banale, senza entrare in particolari dettagli si indica la regola di Sturges per cui, per n osservazioni, suggerisce un numero di classi uguale a $1 + \frac{\log n}{\log 2}$.

Indici di posizione e di dispersione

Gli **INDICI DI POSIZIONE** sono singoli valori che sintetizzano la distribuzione di un certo carattere.

Moda, classe modale

Si chiama **MODA** (nel caso di caratteri qualitativi o discreti) o **CLASSE MODALE**, il carattere o la classe I_k che ha frequenza più alta. Possono esistere più mode o classi modali, si dice che un carattere può essere unimodale, bimodale o trimodale, quando si presentano più di 3 mode si dice che non è definita alcuna moda.

È possibile valutare la moda per i caratteri sia di tipo quantitativo che qualitativo.

Media

Se il carattere è quantitativo, è possibile calcolare la media.
Definito un carattere x su un campione che assume le modalità x_1, x_2, ..., x_n, si dice **MEDIA** (o **MEDIA CAMPIONARIA** o **MEDIA ARITMETICA**) il valore:

$$\bar{x} = \bar{x}_n = \frac{1}{n}(x_1 + x_2 + \ldots + x_n) = \frac{1}{n}\sum_{i=1}^{n} x_i$$

La media è invariante per cambio di unità di misura, ovvero calcolare la media e poi cambiare unità di misura equivale a calcolare la media dei valori convertiti nell'unità di misura. Infatti, se a, b ∈ ℝ, e si pone $y_i = ax_i + b$, allora:

$$\bar{y} = \frac{1}{n}\sum_{i=1}^{n} y_i = \frac{1}{n}\sum_{i=1}^{n}(ax_i + b) = a\left(\frac{1}{n}\sum_{i=1}^{n} x_i\right) + \frac{1}{n}\sum_{i=1}^{n} b = a\bar{x} + b$$

Esempio:
Considerati due caratteri x e y definiti sullo stesso campione, supponendo che $y_i = 2x_i - 3$, si verifichi che la stessa relazione vale anche per le medie dei due caratteri, ovvero: $\bar{y} = 2\bar{x} - 3$.

$$\bar{y} = \frac{1}{n}\sum_{i=1}^{n} y_i = \frac{1}{n}\sum_{i=1}^{n}(2x_i + 3) = 2\left(\frac{1}{n}\sum_{i=1}^{n} x_i\right) + \frac{1}{n}\sum_{i=1}^{n} 3 = 2\bar{x} + 3$$

Esempio:
Siano 50, 55, 53, 44, 43, 48, 50 le temperature registrate a New York in una settimana, misurate in gradi Fahrenheit. Qual è stata la temperatura media in gradi Celsius? Se x indica la temperatura in gradi Fahrenheit e y in Celsius su ha:

$$y = (x - 32)\frac{100}{180}$$

È possibile calcolare la media in gradi Farhenheit e fare un'unica conversione in gradi Celsius.

Inoltre è possibile effettuare un'altra conversione per semplificare il calcolo della media, ponendo ad esempio:

$$z = (x - 50) \qquad x = z + 50$$

$$\bar{z} = \frac{1}{7}(5 + 3 - 6 - 7 - 2) = -\frac{7}{7} \qquad \bar{x} = \bar{z} + 50 = \frac{350 - 7}{7} = \frac{343}{7} = 49$$

$$\bar{y} = (49 - 32)\frac{100}{180} = 17\frac{5}{9} = 9.44\ °C$$

Esempio

Calcolare la media di un carattere x che assume i seguenti valori: 180, 190, 150, 210, 250, 160, 140, 140, 210, 220

Per semplificare i calcoli è possibile effettuare un cambio di variabile: $y = (x - 200) / 10$

In questo modo la media del carattere y è la media dei numeri: -2, -1, -5, 1, 5, -4, -6, -6, 1, 2

Calcolare la media di questi 10 numeri è più semplice, infatti è possibile eliminare dal calcolo i valori -2, -1, -5, 5, 1, 2; sommare i restanti: $-4-6-6+1=-15$ e dividere per 10 per ottenere la media di y: $-1{,}5$

A questo punto è possibile effettuare il cambio di variabile "all'indietro" per ottenere la media di x:

$$x = 10y + 200 \quad da\ cui\ \ \bar{x} = -15 + 200 = 185$$

Se il carattere ha un numero finito di modalità $z_1, \dots, z_k$ e $p_1, \dots, p_k$ sono le frequenze relative del campione allora:

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i = \frac{1}{n}(np_1z_1 + \dots + np_kz_k) = \sum_{i=1}^{k} p_i z_i$$

Esempio:

Dato il seguente campione in cui *i* è il numero di figli, N_i è il numero di famiglie aventi *i* figli, p_i è la frequenza relativa di tali famiglie:

i	0	1	2	3	4	5	6	7	8	9	10	Totale
N_i	35	239	1058	2714	4868	6363	5674	3567	1542	397	43	26500
p_i	0.0013	0.0090	0.04	0.1024	0.1841	0.2401	0.2141	0.1358	0.0582	0.015	0.0016	1

è possibile calcolare la media semplicemente sommando le frequenze relative:

$$\bar{x} = \frac{1}{26500}(35*0 + 239*1 + 1058*2 + 2714*3 + 4868*4 + 6363*5 + 5674*6 + 3567*7 + 1542*8 + 397*9 + 43*10) = 5.17$$

$$\bar{x} = \begin{pmatrix} 0*0.0013 + 1*0.0090 + 2*0.04 + 3*0.1024 + 4*0.1841 + 5*0.2401 + \\ +6*0.2141 + 7*0.1358 + 8*0.0582 + 9*0.015 + +10*0.0016 \end{pmatrix} = 5.17$$

La **MEDIA PONDERATA** si utilizza quando si vuole assegnare un peso diverso ai diversi valori ottenuti nel campionamento.

$$\bar{x} = \frac{1}{\sum_{i=1}^{n} p_i}\sum_{i=1}^{n} p_i x_i \quad dove\ p_i\ \text{è}\ il\ peso\ della\ modalit\text{à}\ x_i$$

Esempio:

Calcolare la media ponderata dei voti di uno studente (valori ottenuti dal campionamento):

Corso	CFU	Voto
Logica Matematica	6	19
Programmazione	12	28
Architetture	12	24
Matematica discreta	6	20
Algoritmi	12	27
Analisi	12	28
Calcolo	6	22
Paradigmi	9	24

$$\bar{x} = \frac{6\cdot 19 + 12\cdot 28 + 12\cdot 24 + 6\cdot 20 + 12\cdot 27 + 12\cdot 28 + 6\cdot 22 + 9\cdot 24}{75} = \frac{1866}{75} = 24.88$$

Supponendo di avere due campioni $x = \{x_1, \dots, x_n\}$ e $y = \{y_1, \dots, y_m\}$ aventi come medie $\bar{x}$ e $\bar{y}$, rispettivamente. È possibile calcolare la media $\bar{w}$ del campione w ottenuto unendo i due campioni di origine utilizzandone le sole medie (e non i singoli elementi):

$$\bar{w} = \frac{1}{(n+m)} \sum_{i=1}^{n+m} w_i = \frac{1}{n+m}\left(\sum_{i=1}^{n} x_i + \sum_{i=1}^{m} y_i\right) = \frac{1}{n+m}(n\bar{x} + m\bar{y}) = \frac{n}{n+m}\bar{x} + \frac{m}{n+m}\bar{y} = \frac{n\bar{x} + m\bar{y}}{n+m}$$

Varianza

Si chiama **VARIANZA** del campione, la quantità:

$$\sigma^2 = \sigma_x^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^2$$

La varianza è un indice di dispersione e misura se i valori del campione sono più o meno lontani dal loro "baricentro": valori grandi indicano che vi è molta dispersione e ci sono campioni lontani dalla media.

Esempio:
i campioni -1, 0, 1 e -100, 50, 50, hanno entrambi media 0, calcolando la varianza si ottiene:

$$\sigma_x^2 = \frac{1}{3}(-1^2 + 1^2) = \frac{2}{3} \qquad \sigma_y^2 = \frac{1}{3}(-100^2 + 50^2 + 50^2) = \frac{10000 + 2500 + 2500}{3} = 5000$$

La media $\bar{x}$ gode di una importante proprietà: è il numero "a" per cui la somma dei quadrati delle distanze di x_i da a è minima.

In generale sia ha:

$$a \to \frac{1}{n}\sum_{i=1}^{n}(x_i - a)^2 \qquad (misura\ di\ quanto\ i\ valori\ x_i sono\ distanti\ da\ a)$$

Il minimo di questa funzione si ha per a = $\bar{x}$, infatti:

$$0 = \left(\frac{1}{n}\sum_{i=1}^{n}(x_i - a)^2\right)' = -\frac{2}{n}\sum_{i=1}^{n}(x_i - a) = -2\left(\frac{1}{n}\sum_{i=1}^{n} x_i - \frac{1}{n}\sum_{i=1}^{n} a\right) = -2(\bar{x} - a) \;\Rightarrow\; \bar{x} = a$$

Proprietà della varianza

$Se\ b \in \mathbb{R}\ e\ y_i = x_i + b, allora:\ \ \sigma_x^2 = \sigma_y^2\ \ infatti:$

$$\sigma_x^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i + b - \bar{x} - b)^2 = \frac{1}{n}\sum_{i=1}^{n}(y_i - \bar{y})^2 = \sigma_y^2$$

$Se\ a \in \mathbb{R}\ e\ y_i = ax_i, allora:\ \ a^2\sigma_x^2 = \sigma_y^2\ \ infatti:$

$$\sigma_y^2 = \frac{1}{n}\sum_{i=1}^{n}(y_i - \bar{y})^2 = \frac{1}{n}\sum_{i=1}^{n}(ax_i - a\bar{x})^2 = \frac{1}{n}\sum_{i=1}^{n}(a(x_i - \bar{x}))^2 = \frac{1}{n}\sum_{i=1}^{n}a^2(x_i - \bar{x})^2 = a^2\frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^2 = a^2\sigma_x^2$$

La radice quadrata σ della varianza è detta **SCARTO QUADRATICO** o **DEVIAZIONE STANDARD**.

Ancora: se il carattere ha un numero finito di modalità $z_1, \dots, z_k$ e $p_1, \dots, p_k$ sono le frequenze relative del campione allora:

$$\sigma^2 = \sum_{i=1}^{k} p_i(z_i - \bar{x})^2$$

Per il calcolo della varianza è possibile individuare una formulazione alternativa più utile:

$$\sigma^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i^2 - 2x_i\bar{x} + \bar{x}^2) = \frac{1}{n}\sum_{i=1}^{n}x_i^2 + \frac{1}{n}\sum_{i=1}^{n}\bar{x}^2 - 2\bar{x}\frac{1}{n}\sum_{i=1}^{n}x_i = \frac{1}{n}\sum_{i=1}^{n}x_i^2 + \bar{x}^2 - 2\bar{x}\bar{x} = \frac{1}{n}\sum_{i=1}^{n}x_i^2 - \bar{x}^2$$

Pertanto, la varianza è uguale alla media dei quadrati meno il quadrato della media.

Mediana e altri quantili

I quantili di ordine n, sono valori che dividono il campione in n gruppi equinumerosi.

Dato un campione x_1, …, x_n, indicando con $x_{(1)}$, …, $x_{(n)}$ i dati del campione ordinati in maniera crescente: $x_{(1)} \leq x_{(2)} \leq \ldots \leq x_{(n)}$, si chiama **MEDIANA** il più piccolo di questi valori che lasci alla sua destra almeno metà delle osservazioni. La mediana divide il campione in due gruppi equinumerosi.

Se n è dispari, la mediana è il valore in posizione (n + 1) / 2.

Se n è pari la mediana è definita come la media tra il valore $x_{(n/2)}$ e $x_{(n/2+1)}$.

Esempio:

Dato il campione: 2 5 3 3 7 8 4 2 2 1 1 1 2 1 3
è possibile individuare la mediana ordinando i valori: 1 1 1 1 2 2 2 2 3 3 3 4 5 7 8
la mediana è: 2
Infatti divide il campione in due gruppi da 7 elementi: $\boxed{1\ 1\ 1\ 1\ 2\ 2\ 2}$ $\boxed{3\ 3\ 3\ 4\ 5\ 7\ 8}$

Dato il campione: 2 5 3 3 7 8 4 2 2 1 1 1 2 1 3 6
è possibile individuare la mediana ordinando i valori: 1 1 1 1 2 2 2 2 3 3 3 4 5 6 7 8
la mediana è: 2,5
Infatti divide il campione in due gruppi da 8 elementi: $\boxed{1\ 1\ 1\ 1\ 2\ 2\ 2\ 2}$ | $\boxed{3\ 3\ 3\ 4\ 5\ 6\ 7\ 8}$

La mediana è un indice di centralità che può essere più utile e rappresentativo della media in certi casi, ad esempio con campioni asimettrici.

L'intervallo $[x_{(1)}, x_{(n)}]$ si chiama **RANGE** e rappresenta l'ampiezza del campione e rappresenta un indice di dispersione, facile da calcolare ma altrettanto sensibile agli errori.

Oltre alla mediana, altri quantili che possono essere utilizzati sono:

- Quartili: dividono il campione in quattro gruppi;
- Decili: dividono il campione in dieci gruppi;
- Percentili: dividono il campione in cento gruppi.

I **QUARTILI** di un campione sono 3: $q_{1/4}$, $q_{2/4}$ e $q_{3/4}$. Il quartile $q_{2/4}$ è la mediana.

La distanza tra il primo e il terzo quartile: $q_{3/4} - q_{1/4}$ è detta **SCARTO INTERQUARTILE** ed è un indice di dispersione più appropriato rispetto al range.

Media e mediana di classi

Nel caso in cui si conoscano le frequenze di un campione relative a classi modali definite, e non relativamente ai singoli valori del campione, è possibile calcolare:

- La media come media ponderata dei valori centrali delle classi, utilizzando come pesi le frequenze delle classi stesse
- La mediana in modo che sia quel valore che divide l'area dell'istogramma in 2 parti aventi la stessa superficie.

Esempio

In una stazione meteorologica in Siberia è stato registrato lo spessore del ghiaccio in centimetri in 21 date ottenendo i seguenti valori: 53, 60, 65, 66, 78, 77, 83, 88, 90, 92, 102, 105, 106, 111, 112, 108, 110, 106, 114, 111, 110.
La media del campione è, circa, **92.71**
La mediana, ottenuta ordinando il campione, è il valore dell'undicesima misurazione: 102
53, 60, 65, 66, 77, 78, 83, 88, 90, 92, **102**, 105, 106, 106, 108, 110, 110, 111, 111, 112, 114.

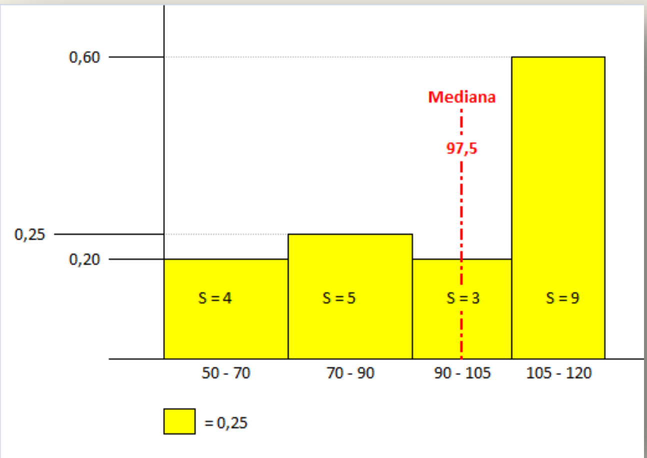
Ipotizzando di avere avuto il campione suddiviso in 4 classi:

Classe	Frequenza
50 – 70	4
70 – 90	5
90 – 105	3
105 - 120	9

Si può calcolare la media come media ponderata dei valori centrali delle classi:

$$\frac{60\cdot 4+80\cdot 5+97.5\cdot 3+112.5\cdot 9}{21}=\frac{240+400+292.5+1012.5}{21}=\frac{1945}{21}=\mathbf{92.62}\ \approx \mathbf{92.71}$$

Per calcolare la mediana è utile disegnare l'istogramma e lavorare sulle superfici.



- La prima classe ha superficie 4, ampiezza 20, da cui: altezza = 4 / 20 = 0,20.
- La seconda classe ha superficie 5, ampiezza 20, da cui: altezza = 5 / 20 = 0,25.
- La terza classe ha superficie 3, ampiezza 15, da cui: altezza = 3 / 15 = 0,20
- La quarta classe ha superficie 9, ampiezza 15, da cui: altezza = 9 / 15 = 060

La superficie totale è 21. La mediana deve dividere la superficie in due parti uguali da 10,5. In questo caso è sufficiente suddividere in 2 parti uguali la superficie della terza classe per ottenere due parti di grafico di pari superficie, pertanto la mediana si trova nel punto centrale della classe 90-105, ovvero **97.5** (piuttosto differente dalla mediana, 102, calcolata sul campione di misurazioni iniziali).

Altre medie

Talvolta, in presenza di un campione, può essere opportuno trasformare i dati in possesso per analizzare il campione e calcolare medie più significative.

In particolare, se A e B sono due intervalli (eventualmente illimitati), i dati si trovano in A e f: A→B è una trasformazione invertibile, allora la media associata alla funzione f è:

$$\bar{x}_f = f^{-1}\left(\frac{1}{n}\sum_{i=1}^{n} f(x_i)\right)$$

Questa operazione consiste nel trasformare i dati tramite la funzione f, calcolarne la media, e applicare l'inversa di f.

Alcuni esempi importanti sono:

- La **MEDIA GEOMETRICA**
  Tipicamente utilizzata per i tassi di interesse e i tassi d'inflazione, è espressa coma la radice n-esima del prodotti dei valori.

  $$f(x) = \log x \quad \Rightarrow \quad f^{-1}(y) = e^y \quad \Rightarrow \quad \bar{x}_G = \bar{x}_{log} = \left(\frac{1}{n}\sum_{i=1}^{n} \log x_i\right)^{1/n} = \left(\prod_{i=1}^{n} x_i\right)^{1/n}$$

  Esempio: supponendo di avere delle azioni in borsa per 4 anni che hanno registrato le seguenti variazioni annue: +10%, -20%, +24%, -14%; qual è stata la variazione media?
  $y_i = x_i + 1 \quad \Rightarrow \quad y_1 = 1.1;\ y_2 = 0.80;\ y_3 = 1.24;\ y_4 = 0.86$
  $\bar{y}_{log} = (1.1 \cdot 0.80 \cdot 1.24 \cdot 0.86)^{1/4} \cong 0.9842 \quad \Rightarrow \quad \bar{x}_{log} \cong 0.9384 - 1 = -0.0158 = -1.58\%$

- La **MEDIA QUADRATICA**
  È espressa come radice quadrata della media dei quadrati dei valori.

  $$f(x) = x^2 \quad \Rightarrow \quad f^{-1}(y) = \sqrt{y} \quad \Rightarrow \quad \bar{x}_Q = \sqrt{\frac{1}{n}\sum_{i=1}^{n} x_i^2}$$

  Esempio: supponendo di voler sostituire 3 tubi di diametro 1, 2, e 3 centimetri con 3 tubi uguali in modo che la portata complessiva rimanga invariata, quale diametro devono avere i 3 tubi?

  $$\bar{x}_Q = \sqrt{\frac{1}{3}(1^2 + 2^2 + 3^2)} = \sqrt{\frac{1+4+9}{3}} = \sqrt{\frac{14}{3}} = \sqrt{4.\overline{6}} \cong 2.16$$

- La **MEDIA ARMONICA**
  È espressa come l'inverso della media degli inversi.

  $$f(x) = \frac{1}{x} \quad \Rightarrow \quad f^{-1}(y) = \frac{1}{y} \quad \Rightarrow \quad \bar{x}_A = \frac{1}{\frac{1}{n}\sum_{i=1}^{n} \frac{1}{x_i}} = \frac{n}{\sum_{i=1}^{n} \frac{1}{x_i}}$$

  Esempio: supponendo di fare una passeggiata in montagna andando, all'andata, alla velocità di 3 km/h e al ritorno alla velocità di 7 km/h. Qual è la velocità media?

  $$\bar{x}_A = \frac{2}{\frac{1}{3} + \frac{1}{7}} = \frac{2}{\frac{10}{21}} = \frac{21}{5} = 4.2\ km/h$$

Correlazione

Per ogni elemento di un campione potrebbero essere studiati diversi caratteri. Può essere utile / necessario valutare il legame tra i caratteri misurati.

Correlare due caratteri significa studiare i possibile legami (correlazioni) tra i due caratteri definiti sullo stesso campione.

Supponendo che i due caratteri siano x e y, e il campione sia costituito da n unità statistiche, si possono "accoppiare" i valori dei caratteri delle stesse unità statistiche: $(x_1, y_1), (x_2, y_2), \ldots, (x_n, y_n)$.

Per stimare la correlazione tra i due caratteri è possibile utilizzare un approccio grafico utilizzando un diagramma a dispersione nel quale:

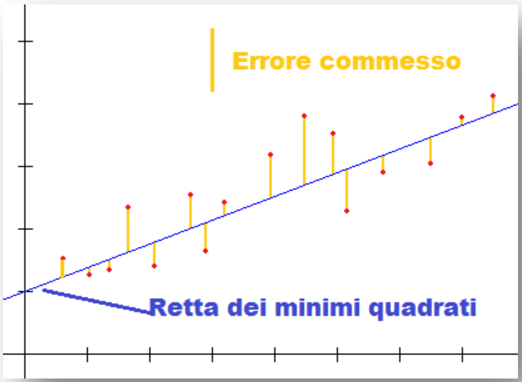
- i valori dei due caratteri x e y si sviluppano rispettivamente lungo l'asse delle ascisse e delle ordinate,
- le coppie delle misurazioni sono indicate con dei punti in corrispondenza delle coordinate da esse rappresentate.

Osservando il grafico potrebbe essere possibile intuire il tipo di correlazione (effettuare una valutazione qualitativa della correlazione): se i punti si dispongono lungo una retta si ha una **correlazione lineare**, se si dispongono lungo una parabola si ha una **dipendenza quadratica** tra i due caratteri.

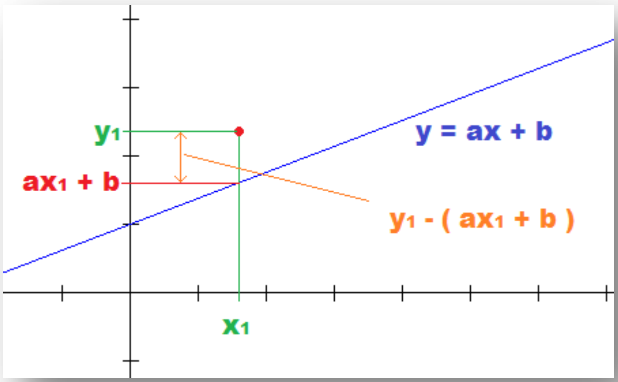
Inoltre, se all'aumentare del valore dei punti del carattere x si rileva un aumento del valore dei punti del carattere y, si ha una **correlazione positiva**, viceversa si ha una **correlazione negativa**.

Approfondendo lo studio della correlazione, allo scopo di passare da una valutazione qualitativa ad una quantitativa più precisa, si procede cercando di definire l'equazione della **retta ai minimi quadrati**.

La RETTA AI MINIMI QUADRATI è la retta che rende minimi i quadrati degli errori commessi utilizzando l'approssimazione data dalla retta stessa.



Ricordando che l'equazione di una retta generica può essere scritta come y = ax + b, è possibile osservare che la misura dell'errore per ogni misurazione corrisponde alla differenza tra y_i e $ax_i + b$.



La somma dei quadrati delle distanze è:

$$S(a,b) = \frac{1}{n}\sum_{i=1}^{n}(y_i - ax_i - b)^2$$

<u>**La retta ai minimi quadrati è la retta di equazione y = ax + b che rende minimo il valore di S(a, b).**</u>

Allo scopo di individuare i valori della retta si effettuano i calcoli necessari.

$$S(a,b) = \frac{1}{n}\sum_{i=1}^{n}(y_i - ax_i - b)^2 =$$

si aggiungono e sottraggono opportuni valori

$$= \frac{1}{n}\sum_{i=1}^{n}\left((y_i - \bar{y}) + (-ax_i + a\bar{x}) + (-b + \bar{y} - a\bar{x})\right)^2 =$$

il termine da elevare al quadrato può essere valutato come un trinomio: $(a + b + c)^2 = a^2 + b^2 + c^2 + 2ab + 2ac + 2bc$

$$= \frac{1}{n}\sum_{i=1}^{n}\begin{pmatrix}(y_i - \bar{y})^2 + (-ax_i + a\bar{x})^2 + (-b + \bar{y} - a\bar{x})^2 + 2(y_i - \bar{y})(-ax_i + a\bar{x}) + \\ +2(y_i - \bar{y})(-b + \bar{y} - a\bar{x}) + 2(-ax_i + a\bar{x})(-b + \bar{y} - a\bar{x})\end{pmatrix} =$$

si divide la sommatoria

$$= \frac{1}{n}\sum_{i=1}^{n}(y_i - \bar{y})^2 + \frac{1}{n}\sum_{i=1}^{n}(-ax_i + a\bar{x})^2 + \frac{1}{n}\sum_{i=1}^{n}(-b + \bar{y} - a\bar{x})^2 + \frac{1}{n}\sum_{i=1}^{n}2(y_i - \bar{y})(-ax_i + a\bar{x}) +$$

$$+ \frac{1}{n}\sum_{i=1}^{n}2(y_i - \bar{y})(-b + \bar{y} - a\bar{x}) + \frac{1}{n}\sum_{i=1}^{n}2(-ax_i + a\bar{x})(-b + \bar{y} - a\bar{x}) =$$

le ultime due sommatorie sono costituite da un valore costante ($2(-b + \bar{y} - a\bar{x})$) e un termine che è la somma di tutti i valori del carattere meno la media moltiplicata per il numero di valori, tale termine assume valore 0, pertanto può essere semplificato

$$= \frac{1}{n}\sum_{i=1}^{n}(y_i - \bar{y})^2 + \frac{1}{n}\sum_{i=1}^{n}(a^2(x_i - \bar{x})^2) + \frac{1}{n}\sum_{i=1}^{n}(-b + \bar{y} - a\bar{x})^2 + \frac{1}{n}\sum_{i=1}^{n}2a^2(y_i - \bar{y})(x_i - \bar{x}) =$$

ora, il primo termine è la varianza di y, il secondo è la varianza di x per a^2, il terzo termine è una costante (è pertanto possibile eliminare la sommatoria e il termine 1/n), il quarto termine rappresenta una misura chiamata **COVARIANZA** (moltiplicata per $2a^2$)

Dati due caratteri x e y, la covarianza è definita similmente alla varianza di un solo termine, come:

$$\boldsymbol{\sigma_{x,y} = \frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}$$

si ha allora:

$$S(a,b) = \sigma_y^2 + a^2\sigma_x^2 + (-b + \bar{y} - a\bar{x})^2 - 2a\sigma_{x,y}$$

Si ricorda che si vuole individuare i valori di a e b in modo tale che il valore S(a, b) sia minimo, per fare questo è necessario che:

i) $(-b+\bar{y}-a\bar{x})^2$ sia nullo, da cui $\bar{y}=a\bar{x}+b$ (ovvero la retta ai minimi quartili passa per $(\bar{x},\bar{y})$)

ii) sia minimizzato il valore di $\sigma_x^2 a^2 - 2\sigma_{x,y}a$; il valore di a che minimizza questa funzione è
$$a=-\frac{-2\sigma_{x,y}}{2\sigma_x^2}=\frac{\sigma_{x,y}}{\sigma_x^2}$$
(si ricordi infatti che il minimo di una funzione del tipo ax+bx² è x = -a/2b)

In definitiva, per individuare i termini dell'equazione della retta ai minimi quadrati servono:

1) la media del carattere x: $\bar{x}$
2) la media del carattere y: $\bar{y}$
3) la covarianza dei caratteri x e y: $\sigma_{x,y}$
4) la varianza del carattere x: σ_x^2

una volta determinati questi valori è possibile calcolare:
$$a=\frac{\sigma_{x,y}}{\sigma_x^2} \quad e \quad b=\bar{y}-a\bar{x}$$

Si ricorda che è possibile calcolare la varianza utilizzando la seguente formula:
$$\boldsymbol{\sigma_x^2=\overline{x^2}-\bar{x}^2}$$

Analogamente, per la covarianza, è possibile individuare una formula semplificata:
$$\begin{aligned}\sigma_{x,y}&=\frac{1}{n}\sum_{i=1}^{n}(x_i-\bar{x})(y_i-\bar{y})=\frac{1}{n}\sum_{i=1}^{n}(x_iy_i-x_i\bar{y}-\bar{x}y_i+\overline{xy})=\\&=\frac{1}{n}\sum_{i=1}^{n}(x_iy_i)+\frac{1}{n}\sum_{i=1}^{n}(-x_i\bar{y})+\frac{1}{n}\sum_{i=1}^{n}(-\bar{x}y_i)+\frac{1}{n}\sum_{i=1}^{n}(\overline{xy})=\\&=\frac{1}{n}\sum_{i=1}^{n}(x_iy_i)-\bar{y}\frac{1}{n}\sum_{i=1}^{n}(x_i)-\bar{x}\frac{1}{n}\sum_{i=1}^{n}(y_i)+\overline{xy}=\\&=\overline{xy}-\overline{xy}-\overline{xy}+\overline{xy}=\overline{xy}-\bar{x}\,\bar{y}\end{aligned}$$

La covarianza può essere positiva o negativa, il coefficiente a della retta ai minimi quadrati ha segno concorde con la covarianza. In termini di correlazione:

- se $\sigma_{x,y}>0$ si ha correlazione positiva
- se $\sigma_{x,y}<0$ si ha correlazione negativa

Si definisce **indice di correlazione** il valore
$$\rho_{x,y}=\frac{\sigma_{x,y}}{\sigma_x\sigma_y}$$

e si ha che:
- l'indice di correlazione ha segno concorde con la covarianza
- l'indice di correlazione è invariante per cambiamenti di scala
- l'indice di correlazione è sempre compreso tra -1 e 1
- se l'indice di correlazione è uguale a -1 o 1, allora i punti sono allineati
- se l'indice di correlazione è uguale a 0, allora i due caratteri sono incorrelati (non esiste un legame lineare).

Esempio

Si calcoli l'equazione della retta ai minimi quadrati nel caso di un campione di 7 studenti, per correlare le ore di studio al voto ottenuto all'esame (in tabella si riportano ore e voti).

Campione Studente	Ore x	Voti y	Ore^2 x^2	Voti^2 y^2	Ore · Voti $x \cdot y$
1	10	10	100	100	100
2	60	28	3.600	784	1.680
3	50	30	2.500	900	1.500
4	40	25	1.600	625	1.000
5	40	20	1.600	400	800
6	30	21	900	441	630
7	25	18	625	324	450
Somma	**255**	**152**	**10.925**	**3.574**	**6.160**
Media	**36,43**	**21,71**	**1.560,71**	**510,57**	**880,00**
	1) $\bar{x}$	2) $\bar{y}$	$\overline{x^2}$	$\overline{y^2}$	$\overline{xy}$

3) $\sigma_{x,y} = \overline{xy} - \bar{x}\,\bar{y} = 880{,}00 - 36{,}43 \cdot 21{,}71 = 880{,}00 - 791{,}02 = 88{,}98$

4) $\sigma_x^2 = \overline{x^2} - \bar{x}^2 = 1.560{,}71 - 36{,}43^2 = 1.560{,}71 - 1.327{,}04 = 233{,}67$

Quindi:

$$a = \frac{\sigma_{x,y}}{\sigma_x^2} = \frac{88{,}98}{233{,}67} = 0{,}38 \quad e \quad b = \bar{y} - a\bar{x} = 21{,}71 - 0{,}38 \cdot 36{,}43 = 7{,}84$$

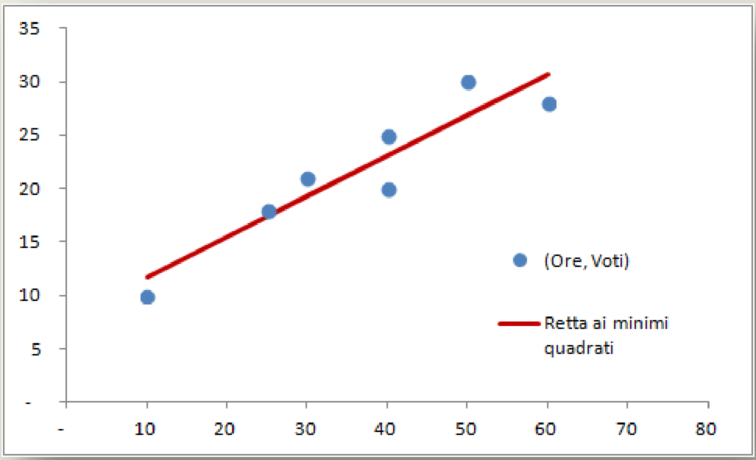
La retta ai minimi quadrati ha coefficiente angolare a = 0,38 > 0.

$$\sigma_y^2 = \overline{y^2} - \bar{y}^2 = 510{,}57 - 21{,}71^2 = 510{,}57 - 471{,}51 = 39{,}06$$

L'indice di correlazione è:

$$\rho_{x,y} = \frac{\sigma_{x,y}}{\sigma_x \sigma_y} = \frac{88{,}98}{\sqrt[2]{233{,}67}\sqrt[2]{39{,}06}} = \frac{88{,}98}{15{,}29 \cdot 6{,}25} = \frac{88{,}98}{95{,}54} = 0{,}93$$

In definitiva vi è una correlazioni positiva e l'indice di correlazione è alto, ergo: gli studenti più studiosi prendono voti più alti.



Nota: i calcoli sono stati effettuati senza approssimare i singoli valori intermedi.

Esempio

Si calcoli l'equazione della retta ai minimi quadrati nel caso di un campione di 7 misurazioni di concentrazione dell'ozono nell'aria e irraggiamento solare, per correlarli fra essi e si valuti quale valore di irraggiamento ci si può aspettare in una giornata con concentrazione di ozono pari a 20.

Campione Misurazioni	**Concentrazione ozono** x	**Irragiamento solare** y	**Concentrazione ozono ²** x^2	**Irragiamento solare ²** y^2	**Concentrazione ozono · Irragiamento solare** x · y
1	41	190	1.681	36.100	7.790
2	16	256	256	65.536	4.096
3	30	322	900	103.684	9.660
4	45	252	2.025	63.504	11.340
5	21	191	441	36.481	4.011
6	32	236	1.024	55.696	7.552
7	10	264	100	69.696	2.640
Somma	**195**	**1.711**	**6.427**	**430.697**	**47.089**
Media	**27,86**	**244,43**	**918,14**	**61.528,14**	**6.727,00**
	1) $\bar{x}$	2) $\bar{y}$	$\overline{x^2}$	$\overline{y^2}$	$\overline{xy}$

3) $\sigma_{x,y} = \overline{xy} - \bar{x}\,\bar{y} = 6.727{,}00 - 244{,}43 \cdot 27{,}86 = 6.727{,}00 - 6.809{,}08 = -82{,}08$

4) $\sigma_x^2 = \overline{x^2} - \bar{x}^2 = 918{,}14 - 27{,}86^2 = 918{,}14 - 776{,}02 = 142{,}12$

Quindi:

$$a = \frac{\sigma_{x,y}}{\sigma_x^2} = \frac{-82{,}08}{142{,}12} = -0{,}58 \quad e \quad b = \bar{y} - a\bar{x} = 244{,}43 - (-0{,}58 \cdot 27{,}86) = 260{,}52$$

La retta ai minimi quadrati ha coefficiente angolare a = – 0,58 < 0.

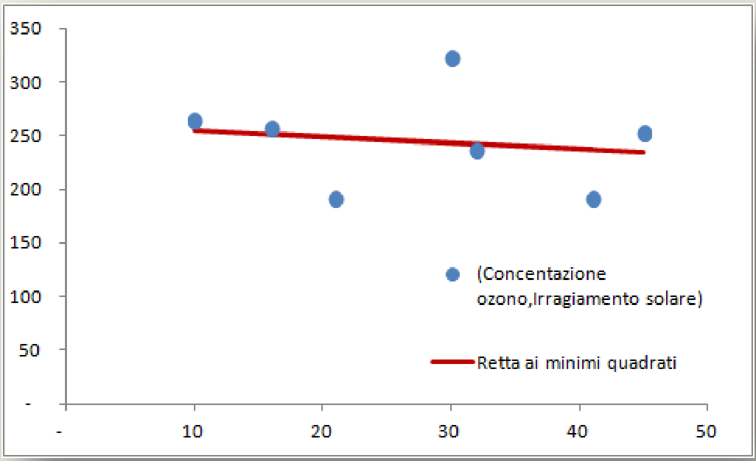
$$\sigma_y^2 = \overline{y^2} - \bar{y}^2 = 61.528{,}14 - 244{,}43^2 = 61.528{,}14 - 59.745{,}33 = 1.782{,}82$$

L'indice di correlazione è:

$$\rho_{x,y} = \frac{\sigma_{x,y}}{\sigma_x \sigma_y} = \frac{-82{,}08}{\sqrt[2]{142{,}12}\sqrt[2]{1.782{,}82}} = \frac{-82{,}08}{11{,}92 \cdot 42{,}22} = \frac{-82{,}08}{503{,}37} = -0{,}16$$

In definitiva vi è una correlazioni negativa e l'indice di correlazione è basso, quindi vi è una correlazione debole.

Per un valore di concentrazione di ozono pari a 20 si può ipotizzare un valore di irraggiamento solare pari a 248,97.



Nota: i calcoli sono stati effettuati senza approssimare i singoli valori intermedi.

Esempio (esercizio 5.3)

In 25 scatole di 100 viti si sono contati il numero di pezzi difettosi ottenendo i seguenti dati:
1; 4; 3; 1; 3; 2; 2; 1; 2; 5; 3; 0; 1; 4; 3; 7; 1; 3; 1; 7; 2; 1; 2; 4; 8.

Determinare le modalità, le frequenze, la media (2.88), i quartili (1,2,4), la moda, la varianza campionaria (4.53), la differenza interquartile e il range.

Popolazione: scatole
Carattere: pezzi difettosi

Modalità:	0	1	2	3	4	5	7	8
Frequenze:	1	7	5	5	3	1	2	1
Frequenze relative:	0.05	0.33	0.24	0.24	0.14	0.05	0.14	

Media: (1 · 0 + 1 · 7 + 2 · 5 + 3 · 5 + 4 · 3 + 5 · 1 + 7 · 2 + 8 · 1) / 25 = 71 / 25 ≅ 2.84

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25
Mediana:	0	1	1	1	1	1	1	1	2	2	2	2	$\boxed{2}$	3	3	3	3	3	4	4	4	5	7	7	8
Quartili:	0	1	1	1	1	1	$\boxed{1}$	1	2	2	2	2	$\boxed{2}$	3	3	3	3	3	$\boxed{4}$	4	4	5	7	7	8

Scarto interquartile: 3
Range: 8
Moda: 1
Varianza: (307 / 25) – 2.84^2 = 12.28 – 8.07 = 4.21

Esempio (esercizio 6.1)

Sono state recensite 512 famiglie di 6 figli. Per ognuna è stato osservato il numero di figlie femmine. I dati sono riportati nella seguente tabella.

Numero di figlie	Frequenza
0	23
1	64
2	131
3	123
4	107
5	48
6	16

Qual è il numero medio di figlie? Calcolare la varianza con entrambe le formule.

$$\bar{x} = \frac{0 \cdot 23 + 1 \cdot 64 + 2 \cdot 131 + 3 \cdot 123 + 4 \cdot 107 + 5 \cdot 48 + 6 \cdot 16}{512} = \frac{1459}{512} \cong 2.85$$

$$\sigma_x^2 = \frac{23(0-2.85)^2 + 64(1-2.85)^2 + 131(2-2.85)^2 + 123(3-2.85)^2 + 107(4-2.85)^2 + 48(5-2.85)^2 + 16(6-2.85)^2}{512} = \frac{1025.42}{512} \cong 2.00$$

$$\sigma_x^2 = \frac{23(0)^2 + 64(1)^2 + 131(2)^2 + 123(3)^2 + 107(4)^2 + 48(5)^2 + 16(6)^2}{512} - 2.85^2 = 10.12 - 8.12 \cong 2.00$$

Esempio (esercizio 6.2)

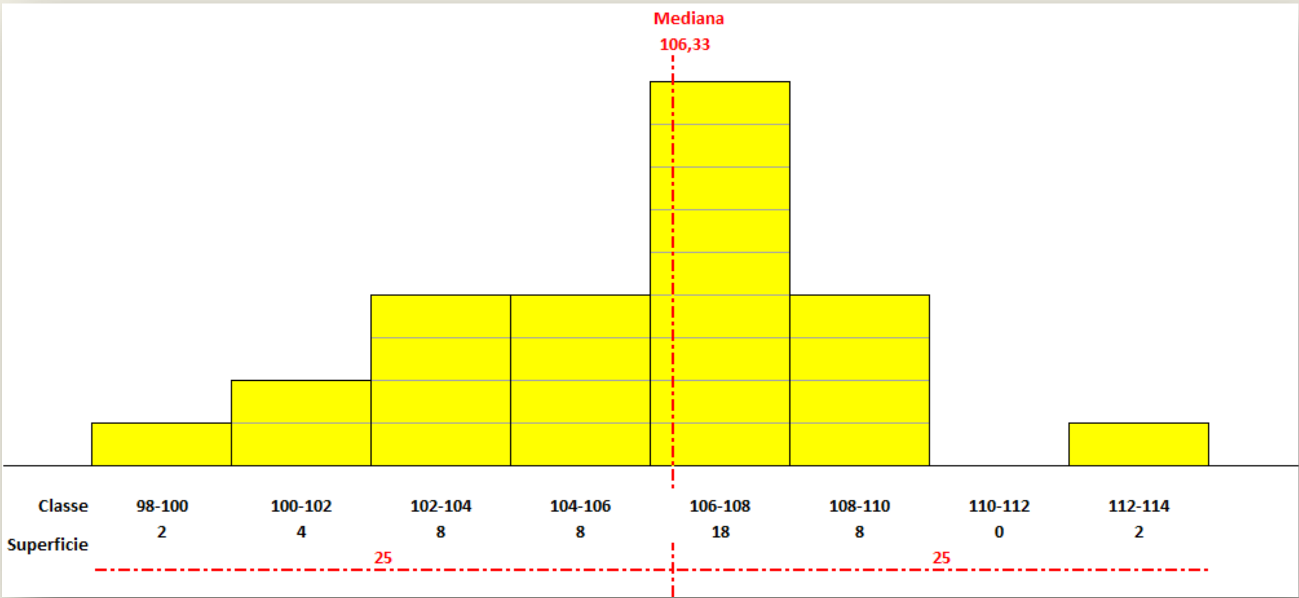
Il peso di 50 giocatori di football americano è rappresentato in questa tabella.

Peso	Frequenza
98-100	2
100-102	4
102-104	8
104-106	8
106-108	18
108-110	8
110-112	0
112-114	2

Determinare la media, la varianza e la mediana dei dati.

$$\bar{x} = \frac{99 \cdot 2 + 101 \cdot 4 + 103 \cdot 8 + 105 \cdot 8 + 107 \cdot 18 + 109 \cdot 8 + 111 \cdot 0 + 113 \cdot 2}{2 + 4 + 8 + 8 + 18 + 8 + 0 + 2} = \frac{5290}{50} = 105.8$$

$$\sigma_x^2 = \frac{99^2 \cdot 2 + 101^2 \cdot 4 + 103^2 \cdot 8 + 105^2 \cdot 8 + 107^2 \cdot 18 + 109^2 \cdot 8 + 111^2 \cdot 0 + 113^2 \cdot 2}{2 + 4 + 8 + 8 + 18 + 8 + 0 + 2} - 105.8^2 = 11202.92 - 11193.64 = 9.28$$



Mediana: 106.33

Esempio (esercizio 6.4)

Si hanno 3 statuette in rame della Tour Eiffel di altezza, rispettivamente, 5, 10 e 20 centimetri.
Volendo fonderle per ottenere tre statuette uguali, quanto dovranno essere alte per utilizzare tutto il rame a disposizione?

$$f(x) = x^3 \ ; \ f^{-1}(y) = y^{\frac{1}{3}} \quad \bar{x} = \left(\frac{5^3 + 10^3 + 20^3}{3}\right)^{\frac{1}{3}} \cong 14.49$$

Probabilità

Studia la possibilità che certi fenomeni accadano. Si costruisce “a priori” un modello esatto da cui trarre conclusioni in modo rigoroso.

Fenomeni deterministici e causali

Un qualunque accadimento che porta ad un certo risultato, appartenente ad un determinato insieme, è detto **FENOMENO**.

Se il risultato può essere previsto o calcolato con esattezza prima che si verifichi, si dice che il **FENOMENO È DETERMINISTICO**. Viceversa, se il risultato non può essere previsto, si dice che il **FENOMENO È CASUALE (O ALEATORIO)**.

Lanciando un oggetto da una determinata altezza, è possibile calcolare l’istante di impatto a terra. Il lancio è un fenomeno deterministico.

Lanciando un dado a sei facce non è invece possibile determinare a priori quale numero uscirà (fenomeno casuale).

L’insieme di tutti i possibili risultati di un fenomeno si indica solitamente con Ω. In generale, un sottoinsieme di Ω è detto **EVENTO** (non tutti i sottoinsiemi potranno essere chiamati eventi).

Sugli eventi è possibile considerare le comuni operazioni di intersezione, unione e complementazione.

La **VALUTAZIONE DI PROBABILITÀ** è una funzione che associa ad ogni evento un numero tanto più grande quanto più si ritiene che l’evento stesso possa accadere; si denota con P(E) la valutazione di probabilità dell’evento E.

Spazi di probabilità

Sia dato un insieme Ω di possibili risultati di un fenomeno aleatorio, gli eventi sono sottoinsiemi di Ω dei quali si vorrà e si potrà calcolare la probabilità. Non ha senso calcolare la probabilità di un sottoinsieme di Ω che non sia un evento.

Se E_1, E_2, … (finiti o infiniti) sono eventi, si dice che formano una **SUCCESSIONE DI EVENTI**. In generale se si specifica un ultimo elemento: E_1, E_2, …, E_n si impone che la successione sia finita, viceversa si impone che la successione sia infinita. In certi casi una successione finita può essere vista come una successione infinita (aggiungendo, ad esempio, infinite volte l’insieme vuoto).

$$\bigcup_{i=1}^{n} E_i = \bigcup_{i=1}^{\infty} E_i \qquad \bigcap_{i=1}^{n} E_i = \bigcap_{i=1}^{\infty} E_i$$

Definizione: famiglia coerente di eventi

Sia $\mathcal{A}$ una collezione di sottoinsiemi di Ω. Si dice che $\mathcal{A}$ è una **FAMIGLIA COERENTE DI EVENTI** se:

i) $\emptyset, \Omega \in \mathcal{A}$
ii) $E^C \in \mathcal{A}$ se $E \in \mathcal{A}$
iii) Unioni ed intersezioni di successioni di elementi di $\mathcal{A}$ appartengono sempre ad $\mathcal{A}$

$$\bigcup_{i=1}^{n} E_i \in \mathcal{A} \quad e \quad \bigcap_{i=1}^{n} E_i \in \mathcal{A}$$

Esempio

Sia Ω = { 1, 2, 3, 4, 5, 6 }, si considerino le seguenti collezioni di sottoinsiemi:

$\mathcal{A}_1$ = { ∅, { 1, 2, 3, 4, 5, 6 }, { 1, 3 }, { 2, 4, 5, 6 } }	è una famiglia coerente di eventi
$\mathcal{A}_2$ = { ∅, { 1, 2, 3, 4, 5, 6 }, { 1, 3 }, { 2, 4, 5, 6 }, { 1, 3, 5 }, { 2, 4, 6 } }	non è una famiglia coerente di eventi Infatti, ad esempio, { 1, 3 } ∪ { 2, 4, 6 } $\notin \mathcal{A}_2$

Una famiglia coerente di eventi è solitamente detta **σ–algebra**.

Data una σ–algebra $\mathcal{A}$, una probabilità su $\mathcal{A}$ è una funzione $P: \mathcal{A} \rightarrow \mathbb{R} \geq 0$ che soddisfa le seguenti proprietà:

$$(1)\ P(\emptyset) = 0 \quad e \quad P(\Omega) = 1$$
$$(2)\ P\left(\bigcup_{i=1}^{\infty} E_i\right) = \sum_{i=1}^{\infty} P(E_i) \quad dove\ E_i\ \text{è}\ una\ successione\ di\ eventi\ a\ due\ a\ due\ disgiunti.$$

Nel caso di unioni non numerabili, la proprietà (2) non è richiesta. È facile che si verifichino casi in cui, se $\Omega = \mathbb{R}$, allora $P(\{ x \}) = 0$ per ogni $x \in \mathbb{R}$.

Definizione: spazio di probabilità

Uno spazio di probabilità è una terna (Ω, $\mathcal{A}$, P) dove:

- Ω è un insieme;
- $\mathcal{A}$ è una famiglia coerente di eventi;
- P è una probabilità su $\mathcal{A}$.

Lo studio di un fenomeno aleatorio si suddivide in 2 parti:

a) la prima consiste nel determinare uno spazio di probabilità che permette di studiare il fenomeno, essa non ha un'unica soluzione e non è possibile determinare in maniera oggettiva se una soluzione sia migliore di un'altra *(questa parte dipende dalla sensibilità, dall'esperienza e dall'interesse dello studioso); questa fase è detta* ***modellizzazione****;*
b) la seconda fase consiste invece nello studio del modello (cioè dello spazio definito nella prima fase), non è soggetta a discrezionalità o ambiguità e i risultati sono univocamente determinati; essa è matematica.

La probabilità uniforme

Dato un insieme finito Ω i cui risultati siano tutti equiprobabili. In questo caso si sceglie $\mathcal{A}$ come l'insieme di tutti i possibili sottoinsiemi di Ω. Si può facilmente dedurre che $P(A) = \frac{|A|}{|\Omega|}$. Il calcolo delle probabilità di un evento si riduce quindi ad un problema di combinatoria, più precisamente al calcolo delle cardinalità di certi insiemi finiti.

$$Si\ ha\ che: \forall\ \omega, \omega' \in \Omega\ \ P(\{\omega\}) = P(\{\omega'\}); \quad \Omega = \{\omega_1, \omega_2, \dots, \omega_n\}\ ovvero\ \Omega = \bigcup_{i=1}^{n} \{\omega_i\}, allora:$$

$$1 = P(\Omega) = P\left(\bigcup_{i=1}^{n} \{\omega_i\}\right) = \sum_{i=1}^{n} P(\{\omega_i\}) \ \ da\ cui: P(\{\omega_i\}) = \frac{1}{n}\ \forall i = 1, 2, \dots, n \ \ perché\ \ P(\{\omega_i\})\ \text{è}\ costante\ \forall i$$

$$Se\ A\ \text{è}\ un\ insieme\ di\ eventi\ favorevoli, allora\ P(A) = \frac{|A|}{|\Omega|}$$

Esempi

1) Qual è la probabilità di lanciare un dado ed ottenere un numero pari? ½
2) Qual è la probabilità di indovinare un terno al lotto? $\frac{1}{\binom{90}{3}} = \frac{1}{\frac{90*89*88}{3*2}} = \frac{1}{117480}$
3) Costruire uno spazio di probabilità (non uniforme) sull'insieme Ω = { 1, 2, a, b }
(Ω = { 1, 2, a, b } , $\mathcal{A}$ = { ∅, { 1, 2 }, { a, b }, { 1, 2, a, b }, { 1 }, { 2, a, b } } , P(A) = 1 se 1 ∈ A; 0 altrimenti }
4) Costruire uno spazio di probabilità sull'insieme ℕ dei numeri naturali
(Ω = ℕ , $\mathcal{A}$ = { ∅, ℕ, { 2x : x ∈ ℕ }, { 2x + 1 : x ∈ ℕ } } , P($\mathcal{A}$) }

Problemi delle σ-algebre

Alcuni esempi aiutano a capire alcuni problemi legati alla cattiva modellizzazione o posizione di problemi:

i) Il fenomeno "scelta di un numero a caso" nell'insieme dei numeri naturali: se si volesse calcolare la probabilità di scegliere un determinato numero si noterebbe immediatamente che l'insieme costituito da quel numero, sarebbe un evento; per uniformità tutti i numeri sono eventi ed hanno la stessa probabilità di verificarsi. Allora: se la probabilità è 0 per tutti, allora P(Ω) sarebbe uguale a 0, viceversa se la probabilità degli eventi fosse maggiore di zero, si avrebbe P(Ω) = ∞. In entrambi i casi si contraddice la definizione di probabilità. **Per modellizzare correttamente insiemi Ω infiniti è utile considerare eventi definiti/generati da progressioni aritmetiche**: ad esempio i multipli di 2 (i numeri pari) o i multipli di 3 e si potrebbe immaginare che questi abbiano probabilità 1/2 o 1/3 rispettivamente; tuttavia cambiando l'ordine dei numeri naturali, si potrebbero fare valutazioni diverse del fenomeno (ad esempio, scrivendo prima tutti i numeri pari, si potrebbe ipotizzare una probabilità pari a 1), pertanto **si vuole che la probabilità non dipenda dall'ordine di rappresentazione degli elementi dell'insieme infinito**.

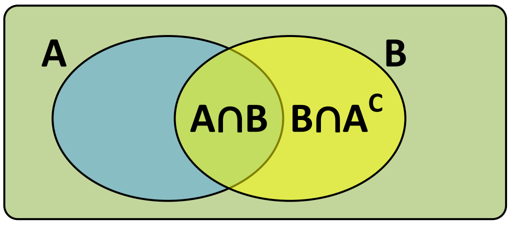
ii) Se Ω è un insieme continuo, non è possibile prendere tutti i numero e la σ-algebra su esso è troppo complessa da esaminare.

Proprietà

$Dati\ due\ eventi\ qualunque\ A\ e\ B, si\ ha\ B = (B \cap A) \cup (B \cap A^C)\ dove\ l'unione\ è\ disgiunta. Ne\ segue:$

$$\boldsymbol{P(B) = P(B \cap A) + P\left(B \cap A^C\right)}$$

(unione disgiunta)



Esempio

Fenomeno: due scatole contengono rispettivamente 2 palline rosse ed 1 blu la prima scatola, e 2 palline rosse e 2 blu la seconda. Qual è la probabilità di estrarre una pallina rossa?

$\Omega = \{\text{insieme delle 7 palline}\}$ $conviene\ considerare\ ognuna\ delle\ palline\ come\ singolo\ risultato.$

$\mathrm{R} = \{\text{insieme delle 4 palline rosse}\}$ $\mathrm{B} = \{\text{insieme delle 3 palline blu}\}$

$\mathrm{A} = \{\text{insieme delle 3 palline della prima scatola}\}$ $\mathrm{A^C} = \{\text{insieme delle 4 palline della seconda scatola}\}$

$$P(R) = P(R \cap A) \cup P(R \cap A^C) = \frac{2}{3} \cdot \frac{1}{2} + \frac{2}{4} \cdot \frac{1}{2} = \frac{1}{3} + \frac{1}{4} = \frac{7}{12}$$

$In\ generale, se\ A_1, A_2, \ldots, A_k\ è\ una\ partizione\ di\ \Omega\ si\ ha\ B = \bigcup_{i=1}^{k} (B \cap A_i)\ \ da\ cui:$

$$\boldsymbol{P(B) = P(B \cap A_1) \cup P(B \cap A_2) \cup \ldots \cup P(B \cap A_k) = \sum_{i=1}^{k} (B \cap A_i)}$$

$Se\ B = \Omega\ \mathrm{e}\ \mathrm{A} \neq \Omega\ \mathrm{e}\ \mathrm{A} \neq \emptyset, allora\ \Omega = (\Omega \cap A) \cup (\Omega \cap A^C) = \mathrm{A} \cup A^C\ \ da\ cui\ P(\Omega) = P(A) + P(A^C)\ da\ cui$
$\boldsymbol{P(A) = 1 - P\left(A^C\right)}$

Ovvero si ha che la probabilità che un evento si verifichi è uguale a 1 meno la probabilità che non si verifichi.

$\boldsymbol{Se\ A \subseteq B\ allora\ P(A) \leq P(B)}, infatti:$
$A \cap B = A, da\ cui: P(B) = P(B \cap A) + P(B \cap A^C) = P(A) + P(B \cap A^C) \geq P(A)$

$\boldsymbol{Se\ A \subset B\ allora\ P(A) \leq P(B)}, infatti:$
$A \cap B = A, da\ cui: P(B) = P(B \cap A) + P(B \cap A^C) = P(A) + P(B \cap A^C) \geq P(A)$

$Ricordando\ le\ leggi\ di\ De\ Morgan\ \left((A \cup B)^C = (A^C \cap B^C)\right)\ si\ può\ ottenere$
$\boldsymbol{P(\cup A_i) = 1 - P\left(\cap A_i^C\right)}$

Esempio
Fenomeno: si lanciano 2 dadi, qual è la probabilità di ottenere almeno un 6?

$\Omega = \{\text{tutte le possibili coppie di valori ottenute}\} \quad \Rightarrow \quad |\Omega| = 6 \cdot 6 = 36$

$A = primo\ dado\ \text{è}\ 6 \qquad B = primo\ dado\ \text{è}\ 6$

$P(A \cup B) = P(\Omega) - \mathrm{P}\left(\mathrm{A^C} \cap \mathrm{B^C}\right) = 1 - \frac{25}{36} = \frac{11}{36}$

$Ricordando\ ancora\ che\ A \cup B = A \cup (B \setminus A), dove\ la\ seconda\ unione\ \text{è}\ disgiunta, si\ ha\ che:$
$\boldsymbol{P(A \cup B)} = P(A) + P(B \setminus A) = P(A) + P(B \cap A^C) = \boldsymbol{P(A) + P(B) - P(A \cap B)}$

Esempio
Fenomeno: si lanciano 2 dadi, qual è la probabilità di ottenere almeno un 6?

$\Omega = \{\text{tutte le possibili coppie di valori ottenute}\} \quad \Rightarrow \quad |\Omega| = 6 \cdot 6 = 36$

$A = primo\ dado\ \text{è}\ 6 \qquad B = primo\ dado\ \text{è}\ 6$

$$P(A \cup B) = P(\mathrm{A}) + \mathrm{P(B)} - \mathrm{P(A \cap B)} = \frac{1}{6} + \frac{1}{6} - \frac{1}{36} = \frac{12-1}{36} = \frac{11}{36}$$

$Nel\ caso\ di\ 3\ eventi\ si\ ha:$
$\boldsymbol{i)}\ \boldsymbol{P(A \cup B \cup C) = 1 - P\left(A^C \cap B^C \cap C^C\right)}$
$\boldsymbol{ii)}\ \boldsymbol{P(A \cup B \cup C) = P(A) + P(B) + P(C) - P(A \cap B) - P(A \cap C) - P(B \cap C) + P(A \cap B \cap C)}$

Probabilità condizionale

Dato uno spazio di probabilità (Ω, $\mathcal{A}$, P), sia A un evento con P(A) > 0, si definisce un nuovo spazio di probabilità (Ω, $\mathcal{A}$, P') dove:

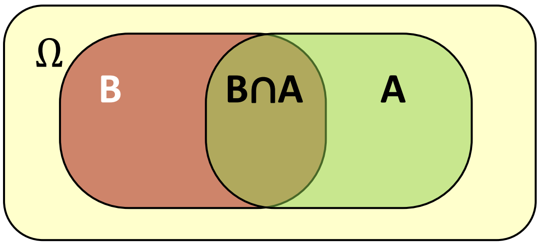
$$P'(B) = \frac{P(B \cap A)}{P(A)}$$

Si verifica che P' sia effettivamente un probabilità:

$i)\ P'(\emptyset) = \frac{P(\emptyset \cap A)}{P(A)} = \frac{P(\emptyset)}{P(A)} = 0$
$ii)\ P'(\Omega) = \frac{P(\Omega \cap A)}{P(A)} = \frac{P(A)}{P(A)} = 1$
$iii)\ siano\ B_1, B_2, \dots eventi\ disgiunti, allora\ P'\left(\bigcup_i B_i\right) = \frac{P\left(A \cap (\cup_i B_i)\right)}{P(A)} = \sum_i \frac{P(A \cap B_i)}{P(A)} = \sum_i P'(B_i)$

La probabilità condizionale rappresenta la probabilità che si verifichi l'evento B sapendo che il risultato è in A, si indica con P(B | A).

Tramite la rappresentazione grafica degli insiemi è possibile comprendere la probabilità condizionale in termini di rapporti tra le superfici.



$$P(A) = \frac{S(A)}{S(\Omega)}$$

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

Formula delle probabilità totali

$$Si\ osservi\ che: P(A|B) = \frac{P(B \cap A)}{P(B)} \quad => P(B \cap A) = P(B)P(A|B)$$

$$Allora: P(A) = P(A \cap B_1) + P(A \cap B_2) + \ldots = P(B_1)P(A|B_1) + P(B_2)P(A|B_2) + \cdots$$

Formula di Bayes

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

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

Eventi indipendenti

Due eventi A e B si dicono INDIPENDENTI se $\boldsymbol{P(A \cap B) = P(A) \cdot P(B)}$.
Analogamente, si definiscono n eventi indipendenti.

Osservando questa caratteristica, in relazione alla probabilità condizionale, si può notare che, per due eventi A e B indipendenti:

$$\boldsymbol{def \quad P(A \cap B) = P(A) \cdot P(B)}$$

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

Calcolo combinatorio - Richiami

- Dati due insiemi A e B di cardinalità n e m, si ha che l'insieme delle coppie ordinate A x B ha cardinalità n·m.

- DISPOSIZIONI: tutte le possibili k-uple ordinate di elementi distinti di un insieme A di cardinalità n.

  $$D_k^n = n(n-1)\ldots(n-k+1) = n(n-1)\ldots(n-k+1)\frac{(n-k)\ldots 1}{(n-k)\ldots 1} = \frac{n!}{(n-k)!}$$

- PERMUTAZIONI: tutte le possibili disposizioni di tutti gli elementi dell'insieme A di cardinalità n.

  $$D_n^n = P_n = n!$$

- COMBINAZIONI: tutti i possibili sottoinsiemi (quindi non ordinati) di cardinalità k dell'insieme A di cardinalità n.

  $$C_k^n = \frac{D_k^n}{k!} = \binom{n}{k} \quad (tutte\ le\ disposizioni\ diviso\ il\ numero\ di\ modi\ di\ ordinarle)$$

Esempio
Fenomeno: due scatole contengono rispettivamente 2 palline rosse ed 1 blu la prima scatola, e 2 palline rosse e 2 blu la seconda. Qual è la probabilità di estrarre una pallina rossa?

$\Omega = \{\text{insieme delle 7 palline}\}$ *conviene considerare ognuna delle palline come singolo risultato.*

$\mathrm{R} = \{\text{insieme delle 4 palline rosse}\}$ $R \cap A = \{palline\ rosse\ della\ prima\ scatola\}$

$\mathrm{A} = \{\text{prima scatola}\}$ $\mathrm{A}^{\mathrm{C}} = \{\text{seconda scatola}\}$

$$P(R) = P(R \cap A) + P(R \cap A^C) = P(A)P(R|A) + P(A^C)P(R|A^C) = \frac{1}{2}\frac{2}{3} + \frac{1}{2}\frac{2}{4} = \frac{1}{3} + \frac{1}{4} = \frac{4+3}{12} = 7/12$$

Esempio
Si studiano gli effetti del fumo su una popolazione, si ha che:
- il 40% delle persone della popolazione sono fumatori, di questi, il 25% ha una malattia respiratoria
- il 60% sono non fumatori, di questi, il 7% ha una malattia respiratoria.

1) Presa una persona a caso, qual è la probabilità che soffra di una malattia respiratoria?
2) Qual è la probabilità che un malato sia fumatore?

$F = \{fumatori\}$ $M = \{malati\}$
$P(F) = 0.4$ $P(F^C) = 0.6$ $P(M|F) = 0.25$ $P(M|F^C) = 0.07$
$$P(M) = P(F)P(M|F) + P(F^C)P(M|F^C) = 0.4 \cdot 0.25 + 0.6 \cdot 0.07 = 0.142 \qquad (1)$$
$$P(F|\ M) = P(M|F)\frac{P(F)}{P(M)} = 0.25 \cdot \frac{0.4}{0.142} \cong 0.704 \qquad (2)$$

Esempio
Un'urna contiene un certo numero b di palline bianche ed un certo numero r di palline rosse.
Si estraggono 3 palline, senza rimpiazzarle.
i) Qual è la probabilità che la prima pallina sia bianca?
ii) Qual è la probabilità che la seconda pallina sia bianca?
iii) Qual è la probabilità che la terza pallina sia bianca?

B_i = la i-esima pallina estratta è bianca R_i = la i-esima pallina estratta è rossa
Ω = insieme di tutte le terne ordinate di palline

$$i)\ P(B_1) = \frac{b}{b+r}$$

$$ii)\ P(B_2) = P(B_2 \cap B_1) + P(B_2 \cap R_1) = P(B_1)P(B_2|B_1) + P(R_1)P(B_2|R_1) =$$
$$= \frac{b}{b+r} \cdot \frac{b-1}{b+r-1} + \frac{r}{b+r} \cdot \frac{b}{b+r-1} = \frac{b^2 - b + rb}{(b+r)(b+r-1)} = \frac{b(b+r-1)}{(b+r)(b+r-1)} = \frac{b}{b+r} = P(B_1)$$

Esempio dei 3 comodini

Ci sono 3 comodini con 2 cassetti ciascuno.

- Il comodino 1 contiene una moneta d'oro in ogni cassetto.
- Il comodino 2 contiene una moneta d'oro in un cassetto ed una d'argento nell'altro.
- Il comodino 3 contiene una moneta d'argento in ogni cassetto.

Si apre un cassetto e si prende una moneta.

Sapendo che la moneta presa è d'oro, qual è la probabilità che anche l'altra sia d'oro?

$Siano: O = \{ l'insieme\ delle\ 3\ monete\ d'oro \} \quad A = \{ l'insieme\ delle\ monete\ d'argento\}$

$C_i = \{ l'insieme\ delle\ monete\ del\ comodino\ i \}$

$\boldsymbol{La\ probabilit\grave{a}\ cercata\ \grave{e}\ P(C_1|O)}$

$Usando\ la\ formual\ di\ Bayes\ si\ ha: P(C_1|O) = \dfrac{P(O|C_1)P(C_1)}{P(O)} = \dfrac{1 \cdot 1/3}{1/2} = \dfrac{1}{3} \cdot 2 = \dfrac{2}{3}$

Esempio

Ci sono 3 urne, l'urna i-esima contiene 1 pallina rossa e i bianche.

Sapendo che si è estratta una pallina bianca, qual è la probabilità di averla estratta dalla 3° urna?

$\Omega = \{insieme\ di\ tutte\ le\ palline\} \qquad U_i = \{palline\ dell'urna\ \ i-esima\}$

$B = \{insieme\ delle\ palline\ bianche\} \quad R = \{insieme\ delle\ palline\ rosse\}$

$\boldsymbol{La\ probabilit\grave{a}\ cercata\ \grave{e}\ P(U_3|B)}$

$Usando\ la\ formual\ di\ Bayes\ si\ ha: P(U_3|B) = \dfrac{P(B|U_3)P(U_3)}{P(B)}$

$P(B|U_3) = \dfrac{3}{4} \qquad P(U_3) = \dfrac{1}{3}$

$P(B) = P(B \cap U_1) + (B \cap U_2) + (B \cap U_3) = P(U_1)P(B|U_1) + P(U_2)P(B|U_2) + P(U_3)P(B|U_3) =$

$= \dfrac{1}{3} \cdot \dfrac{1}{2} + \dfrac{1}{3} \cdot \dfrac{2}{3} + \dfrac{1}{3} \cdot \dfrac{3}{4} = \dfrac{1}{6} + \dfrac{2}{9} + \dfrac{1}{4} = \dfrac{23}{36}$

$\boldsymbol{P(U_3|B) = \dfrac{3/4 \cdot 1/3}{23/36} = \dfrac{1}{4} * \dfrac{36}{23} = \dfrac{9}{23}}$

Esempio

Un'urna contiene 3 palline bianche e 5 palline rosse.

Si estraggono 2 palline con rimpiazzo.

i) Qual è la probabilità di estrarre 2 palline dello stesso colore?

ii) Qual è la probabilità di estrarre almeno 1 pallina bianca?

i) Le possibilità consistono nell'estrazione di 2 palline bianche o di 2 palline rosse.

$Sia\ B_1\ se\ la\ prima\ pallina\ estratta\ \grave{e}\ bianca, B_2\ se\ la\ seconda\ estratta\ \grave{e}\ bianca;$

$R_1\ se\ la\ prima\ estratta\ \grave{e}\ rossa\ , R_2 se\ la\ seconda\ estratta\ \grave{e}\ rossa.$

$P(B_1) = \dfrac{3}{8} \qquad P(B_2) = \dfrac{3}{8} \qquad P(B_1 \cap B_2) = \dfrac{3}{8} \cdot \dfrac{3}{8} = \dfrac{9}{64}$

$P(R_1) = \dfrac{5}{8} \qquad P(R_2) = \dfrac{5}{8} \qquad P(R_1 \cap R_2) = \dfrac{5}{8} \cdot \dfrac{5}{8} = \dfrac{25}{64}$

$P(B_1 \cap B_2) \cup P(R_1 \cap R_2) = \dfrac{9}{64} + \dfrac{25}{64} = \dfrac{34}{64} = \dfrac{17}{32}$

ii) $P(\Omega) - P(R_1 \cap R_2) = 1 - \frac{25}{64} = \frac{39}{64}$

Esempio

Date 10 urne numerate da 1 a 10. Ognuna contiene 4 palline rosse e i palline bianche.
Vengono estratte 2 palline senza rimpiazzo da un'urna scelta a caso.

i) Qual è la probabilità che le due palline siano di colori diversi
ii) Sapendo di aver estratto una pallina bianca e una pallina rossa, qual è l'urna con la maggior probabilità di esser stata scelta?

$\Omega = \{coppie\ non\ ordinate\ di\ palline\ della\ stessa\ urna\}$
$U_i = \{coppie\ dell'urna\ i-esima\}$
$BR = \{coppie\ di\ palline\ di\ colore\ diverso\}$

i) $\boldsymbol{la\ probabilit}$à $\boldsymbol{cercata}$ è $\boldsymbol{P(BR)}$

$Per\ ogni\ urna, si\ hanno\ complessivamente$: $\binom{4+i}{2}$ $modi\ di\ scegliere\ una\ coppia\ qualsiasi$;
$le\ coppie\ di\ tipo\ BR\ possono\ essere\ formate\ scegliendo\ una\ delle\ i\ palline\ bianche\ e\ una\ delle\ 4\ rosse: 4i.$
$Utilizzando\ la\ formula\ delle\ probabilit$à $totali\ si\ ha$:

$$P(BR)=\sum_{i=1}^{10}P(U_i)\big(P(BR|U_i)\big)=\sum_{i=1}^{10}\frac{1}{10}\left(\frac{4i}{\binom{4+i}{2}}\right)=\frac{4}{10}\sum_{i=1}^{10}\left(\frac{i}{\binom{4+i}{2}}\right)=$$
$$=\frac{2}{5}\left(\frac{1}{\binom{5}{2}}+\frac{2}{\binom{6}{2}}+\frac{3}{\binom{7}{2}}+\frac{4}{\binom{8}{2}}+\frac{5}{\binom{9}{2}}+\frac{6}{\binom{10}{2}}+\frac{7}{\binom{11}{2}}+\frac{8}{\binom{12}{2}}+\frac{9}{\binom{13}{2}}+\frac{10}{\binom{14}{2}}\right)=$$
$$=\frac{2}{5}\left(\frac{1}{10}+\frac{2}{15}+\frac{3}{21}+\frac{4}{28}+\frac{5}{36}+\frac{6}{45}+\frac{7}{55}+\frac{8}{66}+\frac{9}{78}+\frac{10}{91}\right)=0.5060$$

ii) $\boldsymbol{La\ probabilit}$**à** $\boldsymbol{cercata}$ **è** $\boldsymbol{P(U_i|BR)}$

$$Usando\ la\ formula\ di\ Bayes\ si\ ha:\ P(U_i|BR)=\frac{P(BR|U_i)P(U_i)}{P(BR)}=\frac{\frac{4i}{\binom{4+i}{2}}\cdot\frac{1}{10}}{0.5060}$$

Effettuando i calcoli si ottengono le seguenti probabilità:

$P(U_1|BR)=0.079$
$P(U_2|BR)=0.105$
$P(U_3|BR)=0.113$ $\quad Massima\ probabilit$à
$P(U_4|BR)=0.113$ $\quad Massima\ probabilit$à
$P(U_5|BR)=0.110$
$P(U_6|BR)=0.105$
$P(U_7|BR)=0.101$
$P(U_8|BR)=0.096$
$P(U_9|BR)=0.091$
$P(U_{10}|BR)=0.087$

Un metodo alternativo per risolvere il problema consiste nel confrontare la probabilità di estrarre 2 palline diverse dall'urna i-esima rispetto all'urna successiva.

$$\frac{P(U_i|BR)}{P(U_{i+1}|BR)}=\frac{\frac{4i}{\binom{4+i}{2}}\cdot\frac{1}{10}}{0.5060}\cdot\frac{0.5060}{\frac{4(i+1)}{\binom{4+i+1}{2}}\cdot\frac{1}{10}}=\frac{4i}{\binom{4+i}{2}}\cdot\frac{\binom{4+i+1}{2}}{4(i+1)}=\frac{i}{\frac{(4+i)(4+i-1)}{2}}\cdot\frac{\frac{(4+i+1)(4+i)}{2}}{(i+1)}=$$
$$=\frac{i(4+i+1)}{(3+i)(i+1)}=\frac{i^2+4i+i}{i^2+4i+3}$$

$$Si\ osserva\ che:\begin{cases} se\ i<3, \frac{P(U_i|BR)}{P(U_{i+1}|BR)}<1\ , quindi\ le\ urne\ 1\ e\ 2\ hanno\ probabilt\text{à}\ inferiore\ all'urna\ 3 \\ se\ i=3, \frac{P(U_i|BR)}{P(U_{i+1}|BR)}=1, quindi\ le\ probabilit\text{à}\ dell'urna\ 3\ e\ dell'urna\ 4\ sono\ uguali \\ se\ i>3, \frac{P(U_i|BR)}{P(U_{i+1}|BR)}>1, quindi\ le\ urne\ succesive\ alla\ 4\ hanno\ probabilit\text{à}\ inferiori\ all'urna\ 4 \end{cases}$$

Esempio
In una partita a POKER con 4 giocatori (32 carte dal 7 all'asso), è più probabile ottenere un poker servito o un colore servito?

$$\Omega = \{combinazioni\ di\ 5\ carte\} \qquad |\Omega| = \binom{32}{5} = \frac{32\cdot 31\cdot 30\cdot 29\cdot 28}{5\cdot 4\cdot 3\cdot 2\cdot 1} = 8\cdot 31\cdot 29\cdot 28$$

$I\ modi\ di\ ottenere\ un\ poker\ servito\ sono\ 8\cdot 28 = 224$
$(esistono\ 8\ tipi\ di\ poker, dal\ 7\ all'asso, fissate\ le\ 4\ carte\ del\ poker, si\ hanno\ 28\ modi\ per\ scegliere\ la\ quinta\ carta)$

$$I\ modi\ di\ ottenere\ un\ colore\ servito\ sono\ 4\cdot\binom{8}{5} = 4\cdot\frac{8\cdot 7\cdot 6\cdot 5\cdot 4}{5\cdot 4\cdot 3\cdot 2} = 4\cdot 8\cdot 7 = 224$$

$$Quindi\text{:}\ \ P(poker) = \frac{8\cdot 28}{8\cdot 31\cdot 29\cdot 28} = \frac{1}{31\cdot 29} \quad = \quad P(colore) = \frac{4\cdot 8\cdot 7}{8\cdot 31\cdot 29\cdot 28} = \frac{1}{31\cdot 29}$$

Nota: in realtà la possibilità del colore è leggermente inferiore perché esistono una serie di combinazioni di scala reale che, nell'esempio, sono assimilate al colore.

Variabili aleatorie discrete

Anziché calcolare le probabilità di un determinato evento, può essere spesso necessario valutare la probabilità dell'effetto di uno di determinati eventi (se ad esempio si valuta una serie di scommesse, potrebbe essere più interessante valutare le possibilità di ottenere determinate vincite complessive, piuttosto che valutare le singole combinazioni di vittorie / perdite, è più interessante sapere la probabilità di vincere più di quanto si è puntato alla fine di tutte le scommesse).

Variabile aleatoria

$Dato\ uno\ spazio\ di\ probabilità\ (\Omega, \mathcal{A}, P), una\ variabile\ aleatoria\ è\ un'applicazione X\colon \Omega \to \mathbb{R}\ tale\ che\colon$
$\forall\, t \in \mathbb{R} \quad l'insieme\ \{\omega \in \Omega\colon X(\omega) \le t\} \in \mathcal{A}$

Si osservi che se è valida la relazione $\le$, sono anche valide tutte le altre ($>$, $<$, $=$):

$\{\omega \in \Omega\colon X(\omega) > t\}\ è\ un\ evento\ complementare\ a\ \{\omega \in \Omega\colon X(\omega) \le t\}, e\ per\ definizione\ esso\ appartiene\ ad\ \mathcal{A}$

$\{\omega \in \Omega\colon X(\omega) = t\} = \{\omega \in \Omega\colon X(\omega) \le t\} \cap \{\omega \in \Omega\colon X(\omega) \ge t\}$

$$\{\omega \in \Omega\colon X(\omega) < t\}\ è\ l'evento\ \bigcup_{i=1}^{n}\left\{\omega \in \Omega\colon X(\omega) \le t - \frac{1}{n}\right\}$$

Esempio
Quale delle seguenti è una variabile aleatoria?

$$\Omega = \mathbb{Z} \qquad \mathcal{A} = \{\Omega, \emptyset, 2\mathbb{Z}, 2\mathbb{Z}+1\}$$

i) $X\colon \Omega \to \mathbb{R} \ \ X(n) = n \quad (identità)$
$Verifica\colon l'insieme\ \{\omega \in \Omega\colon X(\omega) \le k\} \in \mathcal{A} \ \ (k\ costante)?$ **NO**
$\{\omega \in \Omega\colon X(\omega) \le k\} = \{n \in \mathbb{Z}\colon n \le k\} \notin \mathcal{A} \quad (ad\ esempio\ k = 10, l'insieme\ \{10, 9, \dots, -\infty\} \notin \mathcal{A}$

ii) $Y\colon \Omega \to \mathbb{R} \ \ Y(n) = (-1)^n$
$Verifica\colon l'insieme\ \{\omega \in \Omega\colon X(\omega) \le k\} \in \mathcal{A} \ \ (k\ costante)?$ **SI**

$$Y(n) = \begin{cases} -1 & se\ n\ dispari \\ 1 & se\ n\ pari \end{cases} \quad quindi\colon \{\omega \in \Omega\colon X(\omega) \le k\} = \begin{cases} \mathbb{Z} & se \quad k \ge 1 \\ 2\mathbb{Z}+1 & se \quad -1 \le k < 1 \\ \emptyset & se\ k < -1 \end{cases}$$

$Per\ esprimere\ la\ variabile\ aleatoria\ discreta\ \{\omega \in \Omega\colon X(\omega) \le t\}\ si\ userà\ la\ notazione\ (X \le t)$

Variabili aleatorie discrete

Se X è una variabile aleatoria che assume valori finiti o numerabili, allora essa è una **VARIABILE ALEATORIA DISCRETA**.

Densità

$La\ funzione\ p_x: \mathbb{R} \to \mathbb{R}\ data\ da\ p(x_i) = P(X = x_i)\ si\ dice\ \boldsymbol{densit}\text{à}\ della\ variabile\ aleatoria\ discreta\ X.$

Si osserva che la densità soddisfa le seguenti proprietà:

1) $p(x) \neq 0 \quad per\ una\ quantit\text{à}\ finita\ o\ numerabile\ di\ valori$
2) $p(x) \geq 0 \quad per\ ogni\ x \in \mathbb{R}\ \ (essendo\ definita\ come\ probabilit\text{à}\ non\ pu\text{ò}\ assumere\ valori\ negativi)$
3) $\sum_{x \in \mathbb{R}} p(x) = 1$

La funzione densità che soddisfa le proprietà sopra indicate si dice **DENSITÀ DISCRETA**.

Densità uniforme

Una variabile aleatoria discreta X che assume tutti i valori con la stessa probabilità, si dice **DENSITÀ UNIFORME**.

$$Sia\ A = \{x_1, x_2, \dots, x_n\}\ l'insieme\ dei\ valori\ assunti\ dalla\ variabile\ aleatoria\ discreta\ X: p(x) = \begin{cases} \frac{1}{n}\ se\ x \in A \\ 0\ altrimenti \end{cases}$$

Esempio

$\Omega = \{1,2,3,4,5,6\} \qquad X(\omega) = \cos(\pi\omega)$

$$allora: p(\omega) = \begin{cases} P(Y = 1) = P(\{2,4,6\}) = {}^1/_2 \\ P(Y = -1) = P(\{1,3,5\}) = {}^1/_2 \\ 0\ \ altrimenti \end{cases}$$

Esempio

Si effettuano 3 lanci di una moneta. Sia X numero di volte che esce testa. Calcolare la densità di X.

$$\Omega = \{possibili\ risultati\ di\ 3\ lanci: 2^3 = 8\} \quad X(\omega) = \begin{cases} 0\ se\ \ \omega = \{(C,C,C)\} \\ 1\ se\ \ \omega = \{(C,C,T),(C,T,C),(T,C,C)\} \\ 2\ se\ \ \omega = \{(C,T,T),(T,T,C),(T,C,T)\} \\ 3\ se\ \ \omega = \{(T,T,T)\} \end{cases}$$

$$Allora\ p(x) = \begin{cases} {}^1/_8\ se\ \ X(\omega) = 0 \\ {}^1/_8\ se\ \ X(\omega) = 3 \\ {}^3/_8\ se\ \ X(\omega) = 1 \\ {}^3/_8\ se\ \ X(\omega) = 2 \\ 0\ altrimenti \end{cases} \qquad \frac{1}{8} + \frac{1}{8} + \frac{3}{8} + \frac{3}{8} = 1$$

Esempio

Si entra al casinò con 15 €. Si puntano per 3 volte, 5€ alla roulette. Siano X gli € con i quali si esce dal casinò. Calcolare la densità di X.

IPOTESI 1: 36 numeri, 18 vincenti, 18 perdenti.

$\Omega = \{(VVV),(VVP),(VPV),(PVV),(VPP),(PVP),(PPV),(PPP)\}$

A = {0, 10, 20, 30}

P(VITTORIA)	0,5
P(PERDITA)	0,5

ω	P(ω)	x = X(omega)	p(X=x)
VVV	0,125	30	p(30)=0,125
VVP	0,125	20	
VPV	0,125	20	p(20)=0,375
PVV	0,125	20	
VPP	0,125	10	
PVP	0,125	10	p(10)=0,375
PPV	0,125	10	
PPP	0,125	0	p(0)=0,125

IPOTESI 2: 37 numeri, 18 vincenti, 19 perdenti.

$\Omega = \{(VVV),(VVP),(VPV),(PVV),(VPP),(PVP),(PPV),(PPP)\}$

A = {0, 10, 20, 30}

P(VITTORIA)	0,486
P(PERDITA)	0,514

ω	P(ω)	x = X(omega)	p(X=x)
VVV	0,115	30	p(30)=0,115
VVP	0,122	20	
VPV	0,122	20	p(20)=0,365
PVV	0,122	20	
VPP	0,128	10	
PVP	0,128	10	p(10)=0,385
PPV	0,128	10	
PPP	0,135	0	p(0)=0,135

$$Densità: \quad p_{x(k)} = \begin{cases} \left(\frac{19}{37}\right)^3 & se\ k = 0 \\ \left(\frac{19}{37}\right)^2 \left(\frac{18}{37}\right) & se\ k = 1 \\ \left(\frac{19}{37}\right)\left(\frac{18}{37}\right)^2 & se\ k = 2 \\ \left(\frac{18}{37}\right)^3 & se\ k = 3 \\ 0 & altrimenti \end{cases}$$

Funzione caratteristica e Densità di Bernoulli

Dato un insieme $A \subseteq \Omega$, la **FUNZIONE CARATTERISTICA** associata χ_A è la funzione che, ad ogni elemento $\omega \in \Omega$, associa il valore 1 se $\omega \in A$, 0 se $\omega \notin A$.

$$\chi_A(\omega) = \begin{cases} 1 & se\ \omega \in A \\ 0 & se\ \omega \notin A \end{cases}$$

Se A è un evento, allora si può verificare che χ_A è una variabile aleatoria discreta, infatti:

$$\{\chi_A \leq t\} = \begin{cases} \emptyset & se \quad t < 0 \\ A^C & se \quad 0 \leq t < 1 \\ \Omega & se \quad t > 1 \end{cases} \qquad \Rightarrow \qquad \forall t \in \mathbb{R}\ \{\omega : \chi_A(\omega) \leq t\} \in \mathcal{A}$$

La densità della funzione caratteristica è

$$p_{\chi_A}(k) = \begin{cases} P(A) & se\ \ k = 1 \\ 1 - P(A) & se\ k = 0 \\ 0 & altrimenti \end{cases}$$

Una variabile aleatoria X che assume solo valori 0 e 1 si dice **VARIABILE DI BERNOULLI**.

Se la probabilità che la variabile di Bernoulli assuma il valore 1 è p, si scrive: $X \sim B(1, p)$.

In definitiva, per ogni evento A, la variabile χ_A è una variabile di Bernoulli: $\chi_A \sim B\big(1, P(A)\big)$.

Densità Binomiale

Dato un fenomeno aleatorio e un determinato evento su di esso, uno **SCHEMA SUCCESSO-INSUCCESSO** consiste nel ripetere un certo numero (finito o numerabile) di prove e valutare il verificarsi (successo) o meno (insuccesso) dell'evento studiato, per calcolare il numero di successi ottenuti nelle varie prove.

Ad esempio uno schema di successo-insuccesso valuta quante volte esce testa (successo) in una serie ripetuta di lanci di monete.

Se le prove sono indipendenti l'una dall'altra (es. il lancio della moneta) si parla di **SCHEMA SUCCESSO-INSUCCESSO CON RIPETIZIONE**.

La probabilità di ottenere una determinata sequenza di successi e insuccessi, nel caso di prove indipendenti, dipende solo dal numero di essi e non dall'ordine, se p è la probabilità del successo, allora la probabilità di avere una sequenza lunga n definita con k successi e n-k insuccessi è $p^k(1-p)^{n-k}$.

Per generare tutte le possibili sequenze con k successi e n-k insuccessi è necessario scegliere in quali posizioni inserire i k successi, questo può essere fatto in $\binom{n}{k}$ modi diversi.

Considerando una variabile aleatoria X definita come il numero di successi in n tentativi, si ha:

$$p_X(k) = \begin{cases} \binom{n}{k} p^k (1-p)^{n-k} & se\ k = 0, 1, 2, \dots, n \\ 0 & altrimenti \end{cases}$$

In questo caso si scrive $X \sim B(n, p)$; X è detta **VARIABILE BINOMIALE** e p_X è la sua **DENSITÀ BINOMIALE**.

Densità Ipergeometrica

Quando lo schema di successo-insuccesso è senza ripetizioni (ad esempio se si vuole considerare la probabilità che estraendo n palline, senza rimpiazzo, da un'urna contenente b palline bianche e r palline rosse, se ne estraggano esattamente k bianche), lo spazio Ω è dato da tutte le possibili combinazioni di n eventi tra i possibili b successi e r insuccessi e si vuole valutare la probabilità che lo schema successo-insuccesso conti esattamente k successi.

$$I\ possibili\ k\ successi\ possono\ essere\ scelti\ tra\ i\ b\ \text{"disponibili"}\ in\ \binom{b}{k} modi.$$

$$Viceversa\ gli\ (n-k)\ insuccessi\ potranno\ essere\ scelti\ tra\ gli\ r\ in\ \binom{r}{n-k} modi.$$

$$Gli\ eventi\ che\ fanno\ parte\ dello\ schema\ di\ succeso-insuccesso\ sono\ \binom{b+r}{n}.$$

Allora la variabile X, che rappresenta la probabilità di ottenere b successi e r insuccessi da un insieme di n eventi disponibili è una **VARIABILE IPERGEOMETRICA**: $X \sim H(n, b, r)$.

La sua **DENSITÀ IPERGEOMETRICA** è data da:

$$p_X(k) = \begin{cases} \dfrac{\binom{b}{k}\binom{r}{n-k}}{\binom{b+r}{n}} & se \quad k \in \{0,1,\dots,\min(n.,b)\} \\ 0 & altrimenti \end{cases}$$

Esempio

In un hard disk sono memorizzati 3000 file in 30 cartelle, ciascuna contenente 100 file.
Si esegue un programma che deve utilizzare 28 file.
Sapendo che una delle 30 cartelle non è installata, qual è la probabilità che il programma funzioni correttamente nei seguenti casi?

1. Il programma utilizza 28 file differenti
2. Il programma utilizza 28 file eventualmente anche ripetuti.

La probabilità che un file sia in una delle cartelle installate è pari a $\frac{29}{30}$.

1. La variabile che descrive questo schema è $X \sim B\left(28, \frac{29}{30}\right)$ e

Se il programma deve utilizzare 28 file differenti, la probabilità che essi siano installati è pari a

$$P(X=28) = \left(\frac{29}{30}\right)^{28} = 38{,}70\%$$

2. La variabile che descrive questo schema è $Y \sim H(28, 2900, 100)$

Se il programma deve utilizzare 28 file non necessariamente differenti, la probabilità che essi siano installati è pari a

$$P(Y=28) = \frac{\binom{2900}{28}\binom{100}{0}}{\binom{3000}{28}} = \frac{\binom{2900}{28}}{\binom{3000}{28}} = \frac{2900_{28}}{28!} \cdot \frac{28!}{3000_{28}} = 38{,}53\%$$

Esempio

Un urna contiene 2 palline bianche ed 1 rossa. Si estraggono senza rimpiazzo finché non si sceglie la pallina rossa. Considerando la variabile aleatoria X = numero di estrazioni necessarie per trovare la pallina rossa, si ha che X è una variabile finita che può assumere solo i valori 1, 2 o 3. La sua densità è data da:

$$p_X(k) = \begin{cases} \frac{1}{3} & se\ k = 1,2,3 \\ 0 & altrimenti \end{cases}$$

Se si rimpiazzano le palline ad ogni estrazione, la variabile aleatoria X diventa numerabile e si ha:

$$p_X(k) = \begin{cases} \left(\frac{2}{3}\right)^k \left(\frac{1}{3}\right) & se\ k = 1,2,3,\dots \\ 0 & altrimenti \end{cases}$$

Esempio

I bulloni prodotti da una fabbrica sono difettosi con una probabilità del 20% e vengono commercializzati in confezioni da 3. Qual è la probabilità che in una confezione ci sia al massimo un bullone difettoso?

$$X \sim B(3{,}0.8) \Rightarrow \quad p_X(k) = \begin{cases} 0.8^3 = 0.512 & se\ k = 0 \\ 3 \cdot 0.8^2 \cdot 0.2^1 = 0.384 & se\ k = 1 \\ 3 \cdot 0.8^1 \cdot 0.2^2 = 0.096 & se\ k = 2 \\ 0.2^3 = 0.008 & se\ k = 3 \\ 0 & altrimenti \end{cases} \quad \Rightarrow \quad p_X(0) + p_X(1) = 0.512 + 0.384 = 0.896$$

Esempio

Un’urna contiene 8 palline rosse e 2 bianche. Estraendone 3 senza rimpiazzo, qual è la probabilità di estrarne al massimo una bianca?

$$X \sim H(3{,}2{,}8) \Rightarrow \quad p_X(k) = \begin{cases} \frac{\binom{2}{k}\binom{8}{3-k}}{\binom{10}{3}} & se\ k = 1,2,\dots \\ 0 & altrimenti \end{cases} \quad \Rightarrow$$

$$p_X(0) + p_X(1) = \frac{\binom{2}{0}\binom{8}{3}}{\binom{10}{3}} + \frac{\binom{2}{1}\binom{8}{2}}{\binom{10}{3}} = \frac{\binom{8}{3} + 2\binom{8}{2}}{\binom{10}{3}} = \frac{8 \cdot 7 + 8 \cdot 7}{10 \cdot 3 \cdot 4} = \frac{112}{120} = 0.933$$

Esempio

Una compagnia aerea vende 28 biglietti per un volo di 24 posti. Sapendo che ogni passeggero ha probabilità di non presentarsi pari al 30%, qual è la probabilità che qualcuno non trovi posto?
Quanti biglietti si possono vendere ammettendo una probabilità del 10% che qualcuno non trovi posto?

Si può pensare come ad uno schema di successo insuccesso per 28 tentativi, in cui si vogliono contare le persone che parteciperanno. Tutti hanno la stessa probabilità pertanto si può vedere come ad uno schema di successo insuccesso di 2 eventi, con ripetizione.

$$X \sim B\left(28, \left(\frac{7}{10}\right)\right) \quad P(X > 24) = P(X = 25) + P(X = 26) + P(X = 27) + P(X = 28) =$$

$$= \binom{28}{25}\left(\frac{7}{10}\right)^{25}\left(\frac{3}{10}\right)^3 + \binom{28}{26}\left(\frac{7}{10}\right)^{26}\left(\frac{3}{10}\right)^2 + \binom{28}{27}\left(\frac{7}{10}\right)^{27}\left(\frac{3}{10}\right)^1 + \binom{28}{28}\left(\frac{7}{10}\right)^{28}\left(\frac{3}{10}\right)^0 =$$

$$= 0.01186 + 0.00319 + 0.00055 + 0.00005 = 0.1565$$

Densità Geometrica

Si consideri uno schema successo-insuccesso con prove indipendenti (con ripetizioni) ognuna delle quali ha la probabilità p di dare successo. Esaminando la variabile T = tempo di primo successo (indicando con I_n e S_n rispettivamente i casi di insuccesso e insuccesso al tentativo n) si ha che:

$$P(T=k)=P(I_1\cap I_2\cap\ldots\cap I_{k-1}\cap S_k)=P(I_1)P(I_2)\ldots P(I_{k-1})P(S_k)=(1-p)^{k-1}p$$

o anche

$$P(T=k)=P(X_1=0)P(X_2=0)\ldots P(X_{k-1}=0)P(X_k=1)$$
$$dove\ X_i\ \text{è}\ la\ variabile\ B(1,p)\ data\ da\ X_i=1\ se\ l'i-esimo\ tentativo\ da\ successo, X_i=0\ altrimenti$$

Allora p_T è la **DENSITÀ GEOMETRICA MODIFICATA** e si scrive $T\sim\tilde{G}(p)$:

$$p_T(k)=\begin{cases}p(1-p)^{k-1} & se\ k=1,2,3,\ldots\\ 0 & altrimenti\end{cases}$$

La **DENSITÀ GEOMETRICA (STANDARD)** si scrive $X\sim G(p)$ ed è data da:

$$p(k)=\begin{cases}p(1-p)^{k} & se\ k=0,1,2,\ldots\\ 0 & altrimenti\end{cases}$$

Si può verificare che si tratta di una densità in quanto

1. $p(k)\geq 0$
2. $p(k)\neq 0 \quad per\ una\ quantità\ numerabile\ di\ valori$
3. $\sum_k p(k)=1$

Per verificare la condizione 3, si deve ricordare la serie geometrica:

$$\sum_{k=0}^{\infty}a^k=\lim_{n\to\infty}(1+a^1+a^2+\cdots a^n)=\lim_{n\to\infty}\frac{(1+a^1+a^2+\cdots a^n)(1-a)}{(1-a)}=$$
$$=\lim_{n\to\infty}\frac{(1+a^1+a^2+\cdots a^n-a^1-a^2-\ldots-a^{n+1})}{(1-a)}=\lim_{n\to\infty}\frac{1-a^{n+1}}{1-a}=\frac{1}{1-a}\ se\ 0\leq a<1$$

Nel caso della serie geometrica si ha che:

$$\sum_{k=0}^{\infty}(1-p)^k p=p\sum_{k=0}^{\infty}(1-p)^k \ \Rightarrow\ a=(1-p)\quad 0\leq(1-p)<1$$

$$\sum_{k=0}^{\infty}a^k=\frac{1}{1-a}\ \Rightarrow\ p\sum_{k=0}^{\infty}(1-p)^k=p\frac{1}{1-1-p}=p\frac{1}{p}=1$$

Se $T\sim\tilde{G}(p)$ è una variabile di densità geometrica modificata, allora $T-1\sim\tilde{G}(p)$ è una variabile geometrica standard.

Una variabile geometrica X gode della proprietà di **MANCANZA DI MEMORIA** infatti si può osservare che:

$$P(X \geq k) = \sum_{i=k}^{\infty} p(1-p)^i = (1-p)^k \underbrace{p\sum_{i=0}^{\infty}(1-p)^i}_{=1} = (1-p)^k$$

Si ha anche che:

$$P(X = k+m | X \geq k) = \frac{P(X = k+m)}{P(X \geq k)} = \frac{(1-p)^{k+m-1}p}{(1-p)^{k-1}} = p(1-p)^m = P(X = m)$$

Questo significa che la probabilità che si verifichi un successo dopo m prove, è indipendente dal numero di prove fatte prima di iniziare a contare le m prove.

Si può anche dimostrare per induzione:

caso base: $P(X = 0) = p$

ipotesi induttiva: $P(x = i) = p(1-p)^i \; per \; ogni \; i < k$

allora:

$$P(X \geq k) = 1 - P(X < k) = 1 - \sum_{i=0}^{k-1} p(1-p)^i = 1 - p\frac{\left(1-(1-p)^k\right)}{1-(1-p)} = (1-p)^k$$

Quindi

$$P(X = k | X \geq k) = \frac{P(X = k)}{P(X \geq k)} = P(X = 0)$$

Si conclude che:

$$P(X = k) = P(X = 0)P(X \geq k) = p(1-p)^k$$

Esempio

Si lanciano due dadi, uno rosso e l'altro blu. Si ripetono i lanci (simultanei) fino ad ottenere un 6 con entrambi i dadi.
Si hanno le seguenti variabili:
X : primo lancio in cui il dado rosso è 6
Y : primo lancio in cui il dado blu è 6
Z : numero di lanci per ottenere due 6

i) determinare le densità di X, Y, Z
ii) Qual è la probabilità che il dado rosso dia un 6 prima del dado blu?

i)

$$X \sim \tilde{G}\left(\frac{1}{6}\right) \quad p_X(k) = \begin{cases} \left(\frac{5}{6}\right)^{k-1} \left(\frac{1}{6}\right) & se\ k = 1,2,3,\dots \\ 0 & altrimenti \end{cases}$$

$$Y \sim \tilde{G}\left(\frac{1}{6}\right) \quad p_Y(k) = \begin{cases} \left(\frac{5}{6}\right)^{k-1} \left(\frac{1}{6}\right) & se\ k = 1,2,3,\dots \\ 0 & altrimenti \end{cases}$$

$$Z \sim \tilde{G}\left(\frac{1}{36}\right) \quad p_Z(k) = \begin{cases} \left(\frac{35}{36}\right)^{k-1} \left(\frac{1}{36}\right) = \frac{35^{k-1}}{36^k} & se\ k = 1,2,3,\dots \\ 0 & altrimenti \end{cases}$$

ii)

Si può ragionare in questo modo: siccome X e Y sono variabili equivalenti, sarà vero che $P(X < Y) = P(X > Y)$ cioè che la probabilità che esca un 6 sul dado rosso prima che su quello blu è uguale a quella che esca un 6 sul dado blu prima che su quello rosso. Inoltre è ovvio che: $P(X < Y) + P(X > Y) + P(X = Y) = 1$.
Allora $P(X < Y) = \frac{1-P(X=Y)}{2}$

La probabilità di ottenere due 6 allo stesso lancio, ma mai un 6 nei lanci precedenti è

$$P(X = Y) = \sum_{k=0}^{\infty} \left(\frac{25}{36}\right)^k \left(\frac{1}{36}\right) = \frac{1}{36}\left(\frac{1}{1-\frac{25}{36}}\right) = \frac{1}{36}\frac{36}{11} = \frac{1}{11}$$

Pertanto si ha $P(X < Y) = \frac{1-P(X=Y)}{2} = \frac{10}{11} * \frac{1}{2} = \frac{5}{11}$

Esempio

Si hanno 4 urne con 2 palline numerate ciascuna.
L'urna 1 contiene due palline con il numero 1.
L'urna 2 contiene una pallina con il numero 1 e una con il numero 2.
L'urna 3 contiene una pallina con il numero 1 e una con il numero 3.
L'urna 4 contiene una pallina con il numero 1 e una con il numero 4.

Si estrae una pallina di ciascuna urna. Si considerano 2 variabili aleatorie:
X : somma delle palline estratte
Y : nr. di palline estratte con il numero 1

Calcolare:

i) La densità di X
ii) La densità di Y
iii) La probabilità di aver estratto la pallina 4 sapendo che la somma è 7.

Le possibili estrazioni sono:

Urna 1	Urna 2	Urna 3	Urna 4	Somma	Nr. 1	p	
1	1	1	4	7	3	1/8	
1	1	1	1	4	4	1/8	
1	1	3	4	9	2	1/8	
1	1	3	1	6	3	1/8	
1	2	1	4	8	2	1/8	
1	2	1	1	5	3	1/8	
1	2	3	4	10	1	1/8	
1	2	3	1	7	2	1/8	

$$i)\quad p_X(k) = \begin{cases} \frac{1}{8} & se\ k = 4,5,6,7,8,10 \\ \frac{2}{8} & se\ k = 7 \\ 0 & altrimenti \end{cases}$$

$$ii)\quad Y \sim B\left(3, \frac{1}{2}\right) \qquad p_Y(k) = \begin{cases} \binom{3}{k}\left(\frac{1}{2}\right)^k\left(\frac{1}{2}\right)^{3-k} & se\ k = 0,1,2,3 \\ 0 & altrimenti \end{cases}$$

$$p_Y(0) = \binom{3}{0}\left(\frac{1}{2}\right)^0\left(\frac{1}{2}\right)^3 = \frac{1}{8}$$
$$p_Y(1) = \binom{3}{1}\left(\frac{1}{2}\right)^1\left(\frac{1}{2}\right)^2 = 3 \cdot \frac{1}{2} \cdot \frac{1}{4} = \frac{3}{8}$$
$$p_Y(2) = \binom{3}{2}\left(\frac{1}{2}\right)^2\left(\frac{1}{2}\right)^1 = 3 \cdot \frac{1}{4} \cdot \frac{1}{2} = \frac{3}{8}$$
$$p_Y(3) = \binom{3}{3}\left(\frac{1}{2}\right)^3\left(\frac{1}{2}\right)^0 = \frac{1}{8}$$

La variabile Y rappresenta uno schema di successo-insuccesso con ripetizioni sulle 3 urne che contengono una pallina con numero 1 e una pallina con numero diverso da 1. Sulla prima urna, la probabilità di estrarre un 1 è 1.

$$iii)\quad P(P4|S7) = \frac{P(P4 \cap S7)}{P(S7)} = \frac{\frac{1}{8}}{\frac{2}{8}} = \frac{1}{2}$$

Indicando con P4 l'evento "estrazione della pallina 4" e con S7 l'evento "somma delle palline = 7".

Densità di Poisson

La **DENSITÀ DI POISSON** è data da:

$$p(k) = \begin{cases} e^{-\lambda}\dfrac{\lambda^k}{k!} & se\ k = 0,1,2,\dots \\ 0 & altrimenti \end{cases}$$

dove λ è un parametro strettamente positivo.

Se X è una **VARIABILE ALEATORIA DI POISSON** di parametro λ allora si scrive: $X \sim P(\lambda)$.

Ricordando che $\left(1+\frac{x}{n}\right)^n \to e^x$ si vuole verificare che una variabile aleatoria di Poisson approssima, sotto certe ipotesi, una variabile binomiale $X \sim B\left(n, \frac{\lambda}{n}\right)$, infatti in tal caso si ha:

$$p(k) = \binom{n}{k}\left(\frac{\lambda}{n}\right)^k\left(1-\frac{\lambda}{n}\right)^{n-k} = \frac{n!}{k!\,(n-k)!}\frac{\lambda^k}{n^k}\left(1-\frac{\lambda}{n}\right)^{n-k} =$$

$$= \frac{\lambda^k}{k!}\frac{n(n-1)(n-2)\dots(n-k+1)}{n^k}\left(1-\frac{\lambda}{n}\right)^n\left(1-\frac{\lambda}{n}\right)^{-k} \xrightarrow[n\to\infty]{} \frac{\lambda^k}{k!}e^{-\lambda}$$

Pertanto, una variabile B(n,p) con n molto grande e p molto piccolo può essere approssimata con una variabile di Poisson P(nλ). In statistica si adotta l'approssimazione della distribuzione binomiale tramite la distribuzione di Poisson quando n > 20 e p < 1/20, o preferibilmente quando n > 100 e np < 10.

Esempio
In un'azienda, il numero di guasti al mese che si verificano ha una densità di Poisson. La probabilità che in un mese non vi siano guasti è data da e^{-1}.

$$X \sim P(1)$$

a) Qual è la probabilità che nel prossimo mese si verifichi più di un guasto?
È possibile individuare questa probabilità quale complementare della probabilità che nel prossimo mese non si verifichino guasti:

$$P(X>1) = 1 - P(X \le 1) = 1 - P(X=0) - P(x=1) = 1 - e^{-1}\frac{1^0}{0!} - e^{-1}\frac{1^1}{1!} = 1 - e^{-1} - e^{-1} = 0.264$$

b) Qual è la probabilità che nel prossimo mese si verifichino esattamente 2 guasti?

$$P(X=2) = e^{-1}\frac{1^2}{2!} = 0.184$$

c) Qual è la probabilità che nei prossimi 6 mesi si verifichino almeno 3 guasti?
Se e^{-1} è la probabilità che in un mese non vi siano guasti, considerando i guasti da un mese all'altro come eventi indipendenti, si può considerare che la probabilità che in 6 mesi non vi siano guasti sia pari a $(e^{-1})^6 = e^{-6}$
Si considera allora una nuova variabile di Poisson di parametro $\lambda = 6$ $Y \sim P(-6)$

$$P(Y \ge 3) = 1 - P(Y<3) = 1 - P(Y=0) - P(Y=1) - P(Y=2) =$$

$$= 1 - e^{-6}\frac{6^0}{0!} - e^{-6}\frac{6^1}{1!} - e^{-6}\frac{6^2}{2!} = 1 - e^{-6}(1+6+18) = 1 - 25e^{-6} = 0.938$$

d) Qual è la probabilità che esattamente in 2 dei prossimi 6 mesi non ci siano guasti?
Si tratta di uno schema di successo insuccesso in 6 prove indipendenti con p = e^{-1}.

$$Z \sim B(6, e^{-1})$$

La probabilità cercata è data da: $P(Z=2) = P_Z(2) = \binom{6}{2}(e^{-1})^2(1-e^{-1})^4 = 0.324$

Esempio

Un gruppo di amici decide di fare 2 tipi di scherzi telefonici. Ha a disposizione un elenco telefonico e decide di fare scherzi agli abitanti dei paesi A, B e C.

Il paese A ha il doppio degli abitanti dei paesi B e C (#A = #B + #C)

Si sceglie un numero dall'elenco telefonico:

- se il numero appartiene al paese A, si fa lo scherzo di tipo 1
- se il numero appartiene al paese B, si fa lo scherzo di tipo 2
- se il numero appartiene al paese C, si fa, indifferentemente, o lo scherzo di tipo 1 o lo scherzo di tipo 2.

X = #volte in cui si è fatto lo scherzo di tipo 1

NA = #volte in cui si è chiamato un numero del paese A.

a) Qual è la probabilità che il primo scherzo fatto sia di tipo 1?

Sia S1 l'evento scherzo di tipo 1.

La sua probabilità è:

$$P(S1) = P(S1 \cap A) + P(S1 \cap B) + P(S1 \cap C) = P(A)P(S1|A) + P(B)P(S1|B) + P(C)P(S1|C) =$$
$$= \frac{1}{2} \cdot 1 + \frac{1}{4} \cdot 0 + \frac{1}{4} \cdot \frac{1}{2} = \frac{1}{2} + \frac{1}{8} = \frac{5}{8}$$

b) Qual è la densità di X?

$$X \sim B\left(50, \frac{5}{8}\right) \quad P_X(k) = \begin{cases} \binom{50}{k}\left(\frac{5}{8}\right)^k \left(\frac{3}{8}\right)^{n-k} & k = 0,1,\dots .50 \\ 0 & altrimenti \end{cases}$$

c) Qual è la densità di NA?

$$NA \sim B\left(50, \frac{1}{2}\right) \quad P_{NA}(k) = \begin{cases} \binom{50}{k}\left(\frac{1}{2}\right)^k \left(\frac{1}{2}\right)^{n-k} = \binom{50}{k}\frac{1}{2^{50}} & k = 0,1,\dots,50 \\ 0 & altrimenti \end{cases}$$

Esempio

Un album è composto da 100 figurine. Un bambino ne ha già 60.

Acquistano un pacchetto che contiene 6 figurine (tutte diverse), qual è la probabilità di trovare almeno 4 figurine non doppie?

Ogni pacchetto rappresenta un possibile sottoinsieme delle 100 figurine complessive disponibili.

$$X \sim H(6,40,60)$$

Esistono quindi $\binom{100}{6}$ possibili combinazioni diverse di figurine per riempire un pacchetto.

I pacchetti che hanno esattamente 6 figurine nuove (non doppie) sono i possibili sottoinsiemi di 6 elementi delle 40 figurine non in possesso: $\binom{40}{6}$.

I pacchetti che hanno esattamente 5 figurine nuove (non doppie) sono i possibili sottoinsiemi di 6 elementi delle 40 figurine non in possesso: $\binom{40}{5}\binom{60}{1}$.

I pacchetti che hanno esattamente 4 figurine nuove (non doppie) sono i possibili sottoinsiemi di 6 elementi delle 40 figurine non in possesso: $\binom{40}{4}\binom{60}{2}$.

La probabilità di trovare almeno 4 figurine nuove è quindi:

$$P(X \geq 4) = P(X = 4) + P(X = 5) + P(X = 6) = \frac{\binom{40}{6} + \binom{40}{5}\binom{60}{1} + \binom{40}{4}\binom{60}{2}}{\binom{100}{6}} = 0.172$$

Variabili congiunte

Risulta spesso interessante considerare più di una variabile aleatoria su uno stesso spazio di probabilità.

Se $X_1, X_2, \dots, X_m$ sono variabili aleatorie discrete, allora si può definire la **VARIABILE ALEATORIA M-DIMENSIONALE** X:

$$X = (X_1, X_2, \dots, X_m): \Omega \to \mathbb{R}$$

La variabile X assume al massimo una quantità numerabile di valori (se le X_i sono discrete), infatti:

$$\{X = \underline{x}\} = \{X_1 = x_1\} \cap \{X_2 = x_2\} \cap \dots \cap \{X_m = x_m\}$$

$\{X = \underline{x}\}$ è un evento per ogni $\underline{x} \in \mathbb{R}^m$.

La densità di una variabile m-dimensionale è data dalla funzione $p: \mathbb{R}^m \to \mathbb{R}_{\geq 0}$ data da $p(\underline{x}) = P(X = \underline{x})$.

Questa si dice **DENSITÀ CONGIUNTA** delle variabili $X_1, X_2, \dots, X_m$.

Esempio

Da una scatola con 2 palline rosse, 2 bianche, 2 verdi, si estraggono 2 palline e si considerino le variabili
X = palline rosse estratte Y = palline bianche estratte.
Determinare la densità congiunta di X ed Y.

$$X \sim H(2,2,4) \quad P_X(k) = \begin{cases} \dfrac{\binom{2}{k}\binom{4}{2-k}}{\binom{6}{2}} & se\ k = 0,1,2 \\ 0 & altrimenti \end{cases} \qquad Y \sim H(2,2,4) \quad P_Y(k) = \begin{cases} \dfrac{\binom{2}{k}\binom{4}{2-k}}{\binom{6}{2}} & se\ k = 0,1,2 \\ 0 & altrimenti \end{cases}$$

k	j	Densità congiunta	
0	0	$\frac{1}{\binom{6}{2}} = \frac{1}{15}$	Estrazione di due palline verdi
0	1	$\frac{4}{15}$	Estrazione di una pallina bianca ed una pallina verde
0	2	$\frac{1}{15}$	Estrazione di due palline bianche
1	0	$\frac{4}{15}$	Estrazione di una pallina rossa e una pallina verde
1	1	$\frac{4}{15}$	Estrazione di una pallina rossa e una pallina biancha
2	0	$\frac{1}{15}$	Estrazione di 2 palline rosse
1	2	0	Questi eventi, sono possibili per le singole densità ma non per la densità congiunta
2	1	0	
2	2	0	
		1	Totale

Quando si studia una densità congiunta, le densità relative alle variabili aleatorie $X_1, X_2, \dots, X_m$ si dicono **DENSITÀ MARGINALI** e si indicano, solitamente, con $p_1, p_2, \dots, p_m$ e sono univocamente determinate dalla densità congiunta.

$Esempio\ (m = 2)$:

$$p_1(z) = P(X_1 = z) = P\left(\bigcup_{x_2 \in \mathbb{R}} (X_1 = z, X_2 = x_2)\right) = \sum_{x_2 \in \mathbb{R}} P(X_1 = z, X_2 = x_2) = \sum_{x_2 \in \mathbb{R}} P(z, x_2)$$

Non è vero il contrario: le densità marginali non definiscono la densità congiunta.

Variabili aleatorie indipendenti

Le n variabili aleatorie $X_1, X_2, \dots, X_n$ si dicono indipendenti se gli eventi $\{X_1 \in A_1\}, \{X_2 \in A_2\}, \dots, \{X_n \in A_n\}$ sono indipendenti per ogni scelta degli insiemi $A_i \subseteq \mathbb{R}$, vale quindi la seguente relazione:

$$P(X_1 \in A_1, X_2 \in A_2, \dots, X_n \in A_n) = P(X_1 \in A_1)\, P(X_2 \in A_2) \dots P(X_n \in A_n)$$

$$p(\underline{x}) = p_1(x_1) p_2(x_2) \dots p_n(x_n)$$

Lemma

Siano X e Y due variabili aleatorie discrete, esse sono indipendenti se e solo se $p(x,y) = p_X(x) p_Y(y)$ $\forall x, y \in \mathbb{R}$

Dimostrazione

DA FARE

Funzione di ripartizione

Data una qualunque variabile aleatoria X (non necessariamente discreta) la funzione

$$F_X: \mathbb{R} \to \mathbb{R} \qquad F_X(t) = P(X \le t)$$

è detta **FUNZIONE DI RIPARTIZIONE**.

Conoscere la funzione di ripartizione permette di conoscere la densità di una variabile aleatoria:

$$\underset{\textit{densità}}{p_X(t)} = P(X = t) = P(X \le t) - P\big(X \le (t-1)\big) = \underset{\textit{funzioni di ripartizione}}{F_X(t) - F_X(t-1)}$$

Quindi

$$p_X(k) = F_X(k) - F_X(k-1) \quad e \quad F_X(k) = p_X(k) + F_X(k-1) \;\; da\ cui \quad F_X(k) = \sum_{i=0}^{\infty} p_X(k-i)$$

Esempio

$Se\ T\sim\tilde{G}(p)$ *è il tempo di primo successo a prove senza ripetizioni, allora si ha che*
$F_T(k) = P(T \leq k) = 1 - P(T > k) = 1 - (1-p)^k$

La funzione di ripartizione è utile quando si vuole calcolare il minimo e il massimo di variabili.

Esempio

Si lanciano una moneta e un dado simultaneamente fino a quando non si ottiene almeno un 6 e almeno una testa.
Z = #lanci
Qual è la densità di Z?

Si definiscono:

T = tempo di primo successo del lancio del dado $P(T=k) = (1-p)^{k-1}p = \left(\frac{5}{6}\right)^{k-1}\left(\frac{1}{6}\right)$

S = tempo di primo successo del lancio della moneta $P(S=k) = (1-p)^{k-1}p = \left(\frac{1}{2}\right)^{k-1}\left(\frac{1}{2}\right) = \left(\frac{1}{2}\right)^{k}$

Z = Max (S, T) è la variabile che rappresenta il numero di lanci richiesti

$$F_Z(k) = P(Z \leq k) = P(\max(S,T) \leq k) = P(S \leq k, T \leq k)$$

sfruttando il fatto che S e T sono indipendenti

$$P(S \leq k, T \leq k) = P(S \leq k)P(T \leq k) = F_S(k)F_T(k)$$

sfruttando quanto individuato nell'esempio precedente: $F_T(k) = P(T \leq k) = 1 - P(T > k) = 1 - (1-p)^k$

$$F_S(k)F_T(k) = \left(1 - \left(1 - \frac{1}{2}\right)^k\right)\left(1 - \left(1 - \frac{1}{6}\right)^k\right) = \left(1 - \left(\frac{1}{2}\right)^k\right)\left(1 - \left(\frac{5}{6}\right)^k\right) = 1 - \left(\frac{1}{2}\right)^k - \left(\frac{5}{6}\right)^k + \left(\frac{1}{2}\right)^k\left(\frac{5}{6}\right)^k =$$
$$= 1 - \left(\frac{1}{2}\right)^k - \left(\frac{5}{6}\right)^k + \left(\frac{5}{12}\right)^k$$

La densità di Z è:

$$p_Z(k) = F_Z(k) - F_Z(k-1) = 1 - \left(\frac{1}{2}\right)^k - \left(\frac{5}{6}\right)^k + \left(\frac{5}{12}\right)^k - \left(1 - \left(\frac{1}{2}\right)^{k-1} - \left(\frac{5}{6}\right)^{k-1} + \left(\frac{5}{12}\right)^{k-1}\right) =$$
$$= \left(\frac{1}{2}\right)^{k-1}\left(1 - \frac{1}{2}\right) + \left(\frac{5}{6}\right)^{k-1}\left(1 - \frac{5}{6}\right) + \left(\frac{5}{12}\right)^{k-1}\left(\frac{5}{12} - 1\right) = \left(\frac{1}{2}\right)^k + \left(\frac{5}{6}\right)^{k-1}\left(\frac{1}{6}\right) + \left(\frac{5}{12}\right)^{k-1}\left(\frac{7}{12}\right)$$

W = Min (S, T) rappresenta il caso in cui si ottiene o un 6 o una testa

$$1 - F_W(k) = P(W > k) = P(\min(S,T) > k) = P(S > k, T > k) = P(S > k)P(T > k) = \left(\frac{1}{2}\right)^k\left(\frac{5}{6}\right)^k = \left(\frac{5}{12}\right)^k$$

$Allora$: $F_W(k) = 1 - \left(\frac{5}{12}\right)^k$

Che è la funzione di ripartizione di una variabile geometrica modificata di parametro 7/12: $W\sim\tilde{G}\left(\frac{7}{12}\right)$

Più brevemente è possibile calcolare la probabilità dell'evento considerando che:

$p^S = \frac{1}{2}$ $\quad p^T = \frac{5}{6}$ $\quad allora\ \ p^s p^t = \frac{1}{2}\frac{5}{6} = \frac{5}{12}$ $\quad$ *è la* p^w *di non ottenere né testa né 6*

$quindi\ 1 - p^w = 1 - \frac{5}{12} = \frac{7}{12}$ $\quad$ *è la probabilità di ottenere o un 6 o testa*

ovvero, in uno schema di successo – insuccesso, il parametro della v. geometrica modificata.

Cambi di variabile

$Se\ \Phi: \mathbb{R}^m \to \mathbb{R} \quad e \quad \underline{x}\ m-dimensionale \quad allora\ \Phi(\underline{x})$ è $1-dimensionale\ e\ si\ ha\ che:$

$$P_{\Phi(\underline{x})}(k) = \sum_{\underline{x} \in \Phi^{-1}(k)} p_{\underline{x}}(\underline{x})$$

Esempio

Sia X una variabile aleatoria di densità

$$p_X(k) = \begin{cases} 1/_4 & se\ k = 1 \\ 1/_3 & se\ k = -1 \\ 1/_4 & se\ k = 2 \\ 1/_6 & se\ k = -3 \\ 0 & altrimenti \end{cases} \qquad definire\ la\ densità\ di\ \Phi(x) = x^2$$

$\Phi(x)\ assume\ i\ seguenti\ valori: \Phi(1) = 1^2 = 1;\ \ \Phi(-1) = (-1)^2 = 1;\ \ \Phi(2) = 2^2 = 4;\ \ \Phi(-3) = (-3)^2 = 9$

$$p_{\Phi(x)}(k) = \begin{cases} \frac{1}{4} + \frac{1}{3} = \frac{7}{12} & se\ k = 1 \\ \frac{1}{4} & se\ k = 4 \\ \frac{1}{6} & se\ k = 9 \\ 0 & altrimenti \end{cases}$$

Esempio

Sia X una variabile aleatoria di densità

$$p_{X_1}(k) = \begin{cases} 1/_4 & se\ k = 1 \\ 1/_3 & se\ k = -1 \\ 1/_4 & se\ k = 2 \\ 1/_6 & se\ k = -3 \\ 0 & altrimenti \end{cases} \qquad p_{X_2}(k) = \begin{cases} 1/_3 & se\ k = 2 \\ 2/_3 & se\ k = 0 \\ 0 & altrimenti \end{cases} \qquad definire\ la\ densità\ di\ \Phi(x_1, x_2) = x_1 + x_2$$

$\Phi(x_1, x_2)\ assume\ i\ seguenti\ valori:$

$\Phi(1,2) = 3;\ \ \Phi(-1,2) = 1;\ \ \Phi(2,2) = 4;\ \ \Phi(-3,2) = -1;$

$\Phi(1,0) = 1;\ \ \Phi(-1,0) = -1;\ \ \Phi(2,0) = 2;\ \ \Phi(-3,0) = -3$

$$p_{\Phi(x_1,x_2)}(k) = \begin{cases} \frac{1}{6} \cdot \frac{2}{3} = \frac{1}{9} & 4/36 & se\ k = -3 & \Phi(-3,0) \\ \frac{1}{6} \cdot \frac{1}{3} + \frac{1}{3} \cdot \frac{2}{3} = \frac{1}{18} + \frac{2}{9} = \frac{5}{18} & 10/36 & se\ k = -1 & \Phi(-3,2)\ e\ \Phi(-1,0) \\ \frac{1}{3} \cdot \frac{1}{3} + \frac{1}{4} \cdot \frac{2}{3} = \frac{1}{9} + \frac{1}{6} = \frac{5}{18} & 10/36 & se\ k = 1 & \Phi(-1,2)\ e\ \Phi(1,0) \\ \frac{1}{4} \cdot \frac{2}{3} = \frac{2}{12} & 6/36 & se\ k = 2 & \Phi(2,0) \\ \frac{1}{4} \cdot \frac{1}{3} = \frac{1}{12} & 3/36 & se\ k = 3 & \Phi(1,2) \\ \frac{1}{4} \cdot \frac{1}{3} = \frac{1}{12} & 3/36 & se\ k = 4 & \Phi(2,2) \\ 0 & & altrimenti & \end{cases}$$

Esempio
$Siano\ X\sim B(n,p)\quad e\quad Y\sim B(m,p), qual\ è\ la\ densità\ di\ Z = X + Y\quad se\ X\ e\ Y\ sono\ indipendenti?$

X descrive uno schema di successo-insuccesso a n prove indipendenti
Y descrive uno schema di successo-insuccesso a m prove indipendenti

Essendo indipendenti i due eventi, si può pensare a Z come ad uno schema di successo-insuccesso a m+n prove indipendenti: $Z\sim B(n+m,p)$

Esempio
$Siano\ X\sim P(\lambda)\quad e\quad Y\sim P(\mu), qual\ è\ la\ densità\ di\ Z = X + Y?$

Si osserva che:
$i)\ X\sim P(\lambda)\ è\ un'approssimazione\ di\ X'\sim B\left(n,\frac{\lambda}{n}\right)\quad quindi\ X\approx B\left(n,\frac{\lambda}{n}\right)$
$ii)\ Y\sim P(\mu)\ è\ un'approssimazione\ di\ Y'\sim B\left(m,\frac{\mu}{m}\right)\quad quindi\ Y\approx B\left(m,\frac{\mu}{m}\right)$
$iii)\ \frac{\mu}{\lambda}\ è\ un\ numero\ reale\ che\ può\ essere\ approssimato\ con\ un\ numero\ razionale: \frac{\mu}{\lambda}\cong\frac{a}{b}\quad a,b\in\mathbb{N}$
Allora, scegliendo $m = 10^6\cdot a$ e $n = 10^6\cdot b$ si ha:
$$\frac{\mu}{\lambda}\cong\frac{a}{b}=\frac{10^6\cdot a}{10^6\cdot b}=\frac{m}{n}\quad da\ cui:\ \frac{\mu}{\lambda}\cong\frac{m}{n}\quad e\ quindi\ \frac{\mu}{m}\cong\frac{\lambda}{n}\cong p\quad (\lambda = np\quad e\quad \mu = mp)$$
Allora:
$$X\approx B(n,p)\quad Y\approx B(m,p)\quad da\ cui\ Z = X+Y\approx B(n+m,p)\approx P\big((n+m)\cdot p\big) = P(np+mp) = P(\lambda+\mu)$$
Quindi: $\boldsymbol{X\sim P(\lambda) + Y\sim P(\mu) = Z\sim P(\lambda+\mu)}$

Esempio
Un call center lavora per 2 aziende diverse. Siano:
X = #telefonate ricevute per l'azienda 1 $X\sim P(2)$
Y = #telefonate ricevute per l'azienda 2 $Y\sim P(3)$

Stabilire la probabilità che in un giorno si ricevano meno di 4 telefonate.
Z = #telefonate ricevute per l'azienda 1 o per l'azienda 2 $Y\sim P(5)$

La probabilità di ricevere meno di 4 telefonate è
$$P(Z\le 3) = P(Z=0)+P(Z=1)+P(Z=2)+P(Z=3) = e^{-5}+5^{e^{-5}}+\frac{25}{2}e^{-5}+\frac{125}{6}e^{-5}=\frac{118}{3}e^{-5}=0.265$$

Esempio

Un'urna contiene 11 monete che danno testa con probabilità $0, \frac{1}{10}, \frac{2}{10}, \frac{3}{10}, \frac{4}{10}, \frac{5}{10}, \frac{6}{10}, \frac{7}{10}, \frac{8}{10}, \frac{9}{10}, \frac{10}{10}$

Si estrae una moneta e la si lancia 2 volte.

$$X_i = \begin{cases} 1 \; se \; il \; lancio \; i \; da \; testa \\ 0 \; se \; il \; lancio \; i \; da \; croce \end{cases}$$

a) Determinare la densità di X_1 e X_2
b) Stabilire se X_1 e X_2 sono indipendenti
c) Avendo ottenuto 2 volte testa, determinare la probabilità di aver estratto la moneta da $\frac{9}{10}$ o quella da $\frac{10}{10}$

a)

$$p(X_1) \sim B(1,p) \;\rightarrow\; p_{X_1}(k) = \begin{cases} \binom{1}{0} p^0 (1-p)^1 & se \; k = 0 \\ \binom{1}{1} p^1 (1-p)^0 & se \; k = 1 \\ 0 & altrimenti \end{cases} \quad dove \; p \; dipende \; dalla \; moneta, \;\; p(X_2) = p(X_1)$$

Si indica con M_i l'evento "estrazione della moneta i". La densità di X_1 è data, utilizzando la formula delle probabilità totali, da

$$\boldsymbol{p(X_1) = p(X_1 \cap M_0) + p(X_1 \cap M_2) + \cdots + (X_1 \cap M_{10}) = \sum_{i=0}^{10} p(M_i) \cdot p(X_1|M_i) = \sum_{i=0}^{10} \frac{1}{11} \cdot \left(\frac{i}{10}\right)^2 = \frac{1}{11} \sum_{i=0}^{10} \left(\frac{i}{10}\right) = \frac{1}{11} \cdot \frac{55}{10} = \frac{5}{10} = \frac{1}{2}}$$

b)

Intuitivamente, si può immaginare che i due eventi siano dipendenti dal fatto che se si ottiene testa dal primo lancio, è verosimile pensare che si sia scelta una moneta che da testa con maggiore probabilità e quindi, anche il secondo lancio ha maggiori probabilità di dare testa.

$$\boldsymbol{p(X_1 = 1, X_2 = 1) = P(M_0)P(X_1 = 1, X_2 = 1) + P(M_1)P(X_1 = 1, X_2 = 1) + \cdots + P(M_{10})P(X_1 = 1, X_2 = 1) = \frac{1}{11} \sum_{i=0}^{10} \left(\frac{i}{10}\right)^2 =}$$

$$\boldsymbol{= \frac{1}{11} \cdot \frac{385}{100} = \frac{1}{11} \cdot \frac{77}{20} = \frac{7}{20}}$$

$\boldsymbol{Dato\ che\ p(X_1 = 1, X_2 = 1). se\ i\ due\ eventi\ fossero\ indipendenti, sarebbe\ p(X_1 = 1) \cdot p(X_1 = 1) = \frac{1}{2} \cdot \frac{1}{2} = \frac{1}{4} \neq \frac{7}{20}}$
$\boldsymbol{allora\ i\ due\ eventi\ sono\ dipendenti}.$

c)

Sia TT l'evento "si ottengono 2 teste":

$$\boldsymbol{P(M_9|TT) + P(M_{10}|TT) =_{BAYES} = \frac{P(TT|M_9) \cdot P(M_9)}{P(TT)} + \frac{P(TT|M_{10}) \cdot P(M_{10})}{P(TT)} = \frac{\frac{9}{10} \cdot \frac{9}{10} \cdot \frac{1}{11}}{\frac{7}{20}} + \frac{\frac{10}{10} \cdot \frac{10}{10} \cdot \frac{1}{11}}{\frac{7}{20}} = \frac{81}{1100} \cdot \frac{20}{7} + \frac{1}{11} \cdot \frac{20}{7} =}$$

$$\boldsymbol{= \frac{81}{55} \cdot \frac{1}{7} + \frac{20}{77} = \frac{81}{385} + \frac{20}{77} = \frac{81 + 100}{385} = \frac{181}{385} = 0,4701}$$

Media (valore atteso o speranza matematica)

Si consideri una variabile $X \sim B\left(2, \frac{1}{3}\right)$, essa può assumere i valori 0, 1 o 2, in particolare si ha:

$$p_X(0) = \left(\frac{2}{3}\right)^2 = \frac{4}{9}$$
$$p_X(1) = \binom{2}{1}\left(\frac{1}{3}\right)\left(\frac{2}{3}\right) = \frac{4}{9}$$
$$p_X(2) = \left(\frac{1}{3}\right)^2 = \frac{1}{9}$$

La **MEDIA** di questa variabile è data da: $0 \cdot \frac{4}{9} + 1 * \frac{4}{9} + 2 * \frac{1}{9} = \frac{2}{3}$

Definizione di media

Sia X una variabile aleatoria discreta di densità $p(x)$. La media è data da: $E[X] = \sum_{x \in \mathbb{R}} x\, p(x)$

Osservazioni

- In generale, fare una sommatoria per $x \in \mathbb{R}$ non ha senso, ma siccome $p(x)$ è una variabile aleatoria discreta, essa assumerà valori diversi da 0 per una quantità numerabile di valori
- Se una sommatoria di questo tipo non converge, allora la proprietà commutativa non è valida ed è possibile riordinare i termini per ottenere un qualsiasi valore
- Poiché $\sum_{x \in \mathbb{R}} x\, p(x)$ sia una media, essa deve convergere assolutamente
- Nel caso in cui i termini della sommatoria siano finiti e nel caso in cui assumono solo valori non negativi, questi problemi non sussistono, pertanto tutti i tipi di variabili viste non presentano questi problemi.

Esempio
$Sia\ X \sim B(1,p) \Rightarrow E[X] = 0 \cdot p_X(0) + 1 \cdot p_X(1) = 0 \cdot (1-p) + 1 \cdot p = p$
$Sia\ Y \sim B(2,p) \Rightarrow E[Y] = 0 \cdot p_Y(0) + 1 \cdot p_Y(1) + 2p_Y(2) = 0 \cdot (1-p)^2 + 2 \cdot (1-p)p + 2p^2 = 2p - 2p^2 + 2p^2 = 2p$

In generale si ha:
$\boldsymbol{X \sim B(n,p)} \ \Rightarrow \ \boldsymbol{E[X] = np}$

Teorema

Sia $X = (X_1, X_2, \dots, X_m)$ una variabile aleatoria $m-$dimensionale e sia $\Phi: \mathbb{R}^m \to \mathbb{R}$.

Ponendo $Z = \Phi(X)$ si ha che: $\boldsymbol{E[Z] = \sum_{\underline{x} \in \mathbb{R}^m} \Phi(\underline{x}) p(\underline{x})}$ *dove $p(\underline{x})$ è la densità di X.*

Dimostrazione

$$E[Z] = \sum_{z \in \mathbb{R}} z\, p_Z(z) = \sum_{z \in \mathbb{R}} z \sum_{\underline{x} \in \Phi^{-1}(z)} p(\underline{x}) = \sum_{\underline{x} \in \Phi^{-1}(z)} \Phi(\underline{x}) \sum_{\underline{x} \in \Phi^{-1}(z)} p(\underline{x}) = \sum_{\underline{x} \in \mathbb{R}^m} \Phi(\underline{x}) p(\underline{x})$$

La media è un'applicazione lineare

Siano X, Y variabili aleatorie definite sullo stesso spazio di probabilità e sia $c \in \mathbb{R}$, allora:

$$\boldsymbol{E[X+Y] = E[X] + E[Y]} \qquad \boldsymbol{e} \qquad \boldsymbol{E[cX] = cE[X]}$$

Dimostrazione

$$\boldsymbol{E[X+Y]} = \sum_{(x,y)\in\mathbb{R}^2} (x+y)p_{X,Y}(x,y) = \sum_{(x,y)\in\mathbb{R}^2} (x)p_{X,Y}(x,y) + \sum_{(x,y)\in\mathbb{R}^2} (y)p_{X,Y}(x,y) =$$

$$= \sum_{x\in\mathbb{R}} x \sum_{y\in\mathbb{R}} p_{X,Y}(x,y) + \sum_{y\in\mathbb{R}} y \sum_{x\in\mathbb{R}} p_{X,Y}(x,y) = \sum_{x\in\mathbb{R}} x p_X(x) + \sum_{y\in\mathbb{R}} y p_Y(y) = \boldsymbol{E[X] + E[Y]}$$

Utilizzando la formula delle densità marginali: $\sum_{y\in\mathbb{R}} p_{X,Y}(x,y) = p_X(x)$ *e* $\sum_{x\in\mathbb{R}} p_{X,Y}(x,y) = p_Y(y)$

$$\boldsymbol{E[cX]} = \sum_{x\in\mathbb{R}} (cx)p_X(x) = c\sum_{x\in\mathbb{R}} (x)p_X(x) = \boldsymbol{cE[X]}$$

Media di variabili binomiali

È ora possibile dimostrare che:

$\boldsymbol{X \sim B(n,p) \Rightarrow E[X] = np}$

$X \sim B(n,p)$ rappresenta la densità in uno schema di successo insuccesso in n tentativi, con probabilità p di successo, con ripetizione.

$$X = X_1 + X_2 + \ldots + X_n \quad dove\ X_i \sim B(1,p)\ \forall i \quad allora \quad X_i = \begin{cases} 1\ se\ successo\ all'i-esimo\ tentativo \\ 0\ altrimenti \end{cases}$$

$$E[X_i] = p \quad \forall i$$

$$Quindi:\ E[X] = \sum_{i=1\ldots n} E[X_i] = E[X_1] + E[X_2] + \ldots + E[X_n] = p + p + \ldots + p = np$$

Media di variabili ipergeometriche

$$Sia\ X \sim H(n,b,r) \quad allora\ X_i = \begin{cases} 1\ se\ successo\ all'i-esimo\ tentativo \\ 0\ altrimenti \end{cases} \quad quindi\ X_i \sim B\left(1, \frac{b}{b+r}\right)$$

$$pertanto\ \boldsymbol{E[X] = \frac{nb}{b+r}}$$

Media di variabili di Poisson

$$Sia\ X \sim P(\lambda) \approx B\left(n, \frac{\lambda}{n}\right) \sim Y \ \ con\ n\ grande$$

$$allora, E[Y] = n \cdot \frac{\lambda}{n} = \lambda \quad da\ cui \quad \boldsymbol{E[X] = \lambda}$$

$$Lo\ dimostra\ anche\ il\ calcolo\ esplicito: \boldsymbol{E[X]} = \sum_{k=0}^{\infty} k e^{-\lambda}\frac{\lambda^k}{k!} = e^{-\lambda}\sum_{k=0}^{\infty} \lambda \frac{\lambda^{k-1}}{(k-1)!} = e^{-\lambda}\lambda\sum_{h=0}^{\infty}\frac{\lambda^h}{h!} = \lambda$$

Esempio

Si invitano a casa 10 amici e si offre il gelato.

Si hanno a disposizione 6 vaschette con gusti diversi, i primi 3 gusti sono scelti da 2 su 9 gli ultimi e da 1 su 9.

Ogni amico può scegliere un solo gusto.

$$X_i = \begin{cases} 1 \; se \; qualcuno \; sceglie \; il \; gusto \; i \\ 0 \; se \; nessuno \; sceglie \; il \; gusto \; i \end{cases}$$

Mediamente, quante vaschette di gelato si devono aprire?

$Sia\ X = \#vaschette\ da\ aprire \qquad X = X_1 + X_2 + X_3 + X_4 + X_5 + X_6$

$E[X] = E[X_1] + E[X_2] + E[X_3] + E[X_4] + E[X_5] + E[X_6]$

la probabilità che un amico scelga il gusto i, è il complementare della probabilità che nessuno lo scelga:

$$P(X_1 \geq 1) = P(X_2 \geq 1) = P(X_3 \geq 1) = 1 - P(X_1 = 0) = 1 - \left(\frac{7}{9}\right)^{10} \cong 0.92$$

$$P(X_4 \geq 1) = P(X_5 \geq 1) = P(X_6 \geq 1) = 1 - P(X_4 = 0) = 1 - \left(\frac{8}{9}\right)^{10} \cong 0.69$$

$E[X] \cong 0.92 + 0.92 + 0.92 + 0.69 + 0.69 + 0.69 \cong 4.83 \cong 5\ vaschette$

Esempio

Sia (X,Y) una variabile aleatoria bidimensionale

$$p(x,y) = \begin{cases} \frac{1}{4} & se\ (x,y) = (1,2), (-1,2) \\ \frac{1}{2} & se\ (x,y) = (0,0) \\ 0 & altrimenti \end{cases}$$

Calcolare E[X], E[Y], E[X·Y] e stabilire se X e Y sono indipendenti.

$p_X(1) = p(1,2) = \frac{1}{4} \qquad p_X(-1) = p(-1,2) = \frac{1}{4} \qquad p_X(0) = p(0,0) = \frac{1}{2} \qquad p_X(k) = \begin{cases} \frac{1}{4} & se\ k = 1 \\ \frac{1}{4} & se\ k = -1 \\ \frac{1}{2} & se\ k = 0 \\ 0 & altrimenti \end{cases} \qquad E[X] = 1 \cdot \frac{1}{4} + (-1) \cdot \frac{1}{4} + 0 \cdot \frac{1}{2} = 0$

$p_Y(2) = p(1,2) + p(-1,2) = \frac{1}{2} \qquad p_Y(0) = p(0,0) = \frac{1}{2} \qquad p_Y(k) = \begin{cases} \frac{1}{2} & se\ k = 2 \\ \frac{1}{2} & se\ k = 0 \\ 0 & altrimenti \end{cases} \qquad E[Y] = 2 \cdot \frac{1}{2} + 0 \cdot \frac{1}{2} = 1$

$p(1,2) = \frac{1}{4} \neq p_X(1)p_Y(2) = \frac{1}{4} \cdot \frac{1}{2} = \frac{1}{8} \quad p(-1,2) = \frac{1}{4} \neq p_X(-1)p_Y(2) = \frac{1}{4} \cdot \frac{1}{2} = \frac{1}{8} \quad p(0,0) = \frac{1}{2} \neq p_X(0)p_Y(0) = \frac{1}{2} \cdot \frac{1}{2} = \frac{1}{4}$

I due eventi sono dipendenti.

	$\boldsymbol{p_Y(0) = \frac{1}{2}}$	$\boldsymbol{p_Y(2) = \frac{1}{2}}$	
$\boldsymbol{p_X(1) = \frac{1}{4}}$	$\boldsymbol{p_{XY}(1 \cdot 0) = \frac{1}{8}}$	$\boldsymbol{p_{XY}(1 \cdot 2) = \frac{1}{8}}$	$p_{XY}(0) = \frac{1}{8} + \frac{1}{8} + \frac{1}{4} + \frac{1}{4} = \frac{6}{8} \quad p_{XY}(2) = \frac{1}{8} \quad p_{XY}(-2) = \frac{1}{8}$
$\boldsymbol{p_X(-1) = \frac{1}{4}}$	$\boldsymbol{p_{XY}(-1 \cdot 0) = \frac{1}{8}}$	$\boldsymbol{p_{XY}(-1 \cdot 2) = \frac{1}{8}}$	$E[X \cdot Y] = 0 \cdot \frac{6}{8} + 2 \cdot \frac{1}{8} + (-2) \cdot \frac{1}{8} = 0 = E[X]E[Y]$
$\boldsymbol{p_X(0) = \frac{1}{2}}$	$\boldsymbol{p_{XY}(0 \cdot 0) = \frac{1}{4}}$	$\boldsymbol{p_{XY}(0 \cdot 2) = \frac{1}{4}}$	

Media di variabili binomiali

Se X_1 e X_2 sono due variabili aleatorie di densità $B(1,p)$ entrambe (possono assumere solo valori 0 e 1), la variabile X_1X_2 prodotto delle due variabili può assumere ancora solo i valori 0 e 1, è quindi $B(1,p')$ dove $p' = P(X_1X_2 = 1)$, allora:

$$p' = P(X_1X_2 = 1) = P(X_1 = 1, X_2 = 1) = P(X_1 = 1)P(X_2 = 1) = p^2 \;\; \Rightarrow \;\; E[X_1X_2] = p' = E[X_1]E[X_2]$$

Se X_1 e X_2 sono invece variabili aleatorie che rappresentano uno schema di successo − insuccesso senza rimpiazzo, esse hanno densità $B\left(1, \frac{b}{b+r}\right)$, il loro prodotto è ancora una variabile di Bernoulli $B(1,q)$:

$$q = P(X_1X_2 = 1) = P(X_1 = 1, X_2 = 1) = \frac{b}{b+r}\frac{b-1}{b+r-1}$$
$$\Rightarrow \;\; E[X_1X_2] = \frac{b}{b+r}\frac{b-1}{b+r-1} \neq \left(\frac{b}{b+r}\right)^2 = E[X_1]E[X_2]$$

Proposizione

Se X ed Y sono variabili aleatorie indipendenti allora $E[XY] = E[X]E[Y]$

Dimostrazione

$$E[XY] = \sum_{(x,y)\in\mathbb{R}^2} xy\, p(x,y) = \sum_{(x,y)\in\mathbb{R}^2} xy\, p_X(x)p_Y(y) = \sum_{x\in\mathbb{R}} x\, p_X(x) \sum_{y\in\mathbb{R}} y\, p_Y(y) = E[X]E[Y]$$

Non è vero il contrario: è possibile che, se X e Y sono dipendenti, si abbia comunque E[XY] = E[X]E[Y].

Esempio

Si hanno 112 dadi di cui 56 truccati. Quelli truccati danno 1 con probabilità 1/2 mentre gli altri 5 risultati hanno tutti probabilità 1/10. Si prende un dado a caso e lo si lancia.

a) Qual è la probabilità che esca 3?

$\boldsymbol{Sia\ X = risultato\ del\ lancio.}$

$\boldsymbol{Sia\ A\ l'evento\ estrazione\ di\ un\ dado\ truccato, per\ la\ formual\ delle\ probabilità\ totali\ si\ ha\ che:}$

$$\boldsymbol{P(X=3) = P(X=3\cap A) + P(X=3\cap B) = P(A)P(X=3|A) + P(B)P(X=3|B) = \frac{1}{2}\cdot\frac{1}{10} + \frac{1}{2}\cdot\frac{1}{6} = \frac{1}{20} + \frac{1}{12} = \frac{3+5}{60} = \frac{2}{15}}$$

b) Qual è il risultato medio

$$\boldsymbol{P(X=2) = P(X=3) = P(X=4) = P(X=5) = P(X=6) = \frac{2}{15}}$$
$$\boldsymbol{P(X=1) = P(X=1\cap A) + P(X=1\cap B) = P(A)P(X=1|A) + P(B)P(X=1|B) = \frac{1}{2}\cdot\frac{1}{2} + \frac{1}{2}\cdot\frac{1}{6} = \frac{1}{4} + \frac{1}{12} = \frac{3+1}{12} = \frac{1}{3}}$$
$$\boldsymbol{E[X] = 1\cdot\frac{1}{3} + 2\cdot\frac{2}{15} + 3\cdot\frac{2}{15} + 4\cdot\frac{2}{15} + 5\cdot\frac{2}{15} + 6\cdot\frac{2}{15} = \frac{1}{3} + \frac{4}{15} + \frac{6}{15} + \frac{8}{15} + \frac{10}{15} + \frac{12}{15} = \frac{1}{3} + \frac{40}{15} = \frac{45}{15} = 3}$$

c) Si consideri un nuovo fenomeno aleatorio: si sceglie un dado a caso e lo si lancia 2 volte. Si consideri la variabile bidimensionale (X,Y) data dai risultati dei 2 lanci e sia p(x,y) la corrispondente densità congiunta. Determinare p(2,3)

$$\boldsymbol{p(2,3) = P(X=2\cap Y=3) = P(A)P(X=2\cap Y=3|A) + P(B)P(X=2\cap Y=3|B) =}$$
$$\boldsymbol{= \frac{1}{2}\cdot\frac{1}{10}\cdot\frac{1}{10} + \frac{1}{2}\cdot\frac{1}{6}\cdot\frac{1}{6} = \frac{1}{200} + \frac{1}{72} = \frac{9+25}{1800} = \frac{17}{900} \cong 0.0189}$$

d) Stabilire se le due variabili sono indipendenti

Intuitivamente le due variabili non sono indipendenti in quanto se si ottiene un numero diverso da 1 al primo lancio, è più probabile che si sia scelto un dado non truccato.

$$\boldsymbol{0.0189 \cong P(X=2\cap Y=3) \neq P(X=2)P(Y=3) = \frac{2}{15}\cdot\frac{2}{15} = \frac{4}{225} \cong 0.0178}$$

e) Determinare E[X+Y] $\quad \boldsymbol{E[X+Y] = E[X] + E[Y] = 3 + 3 = 6}$

Media di una variabile geometrica modificata

Lanciando un dado, ci si aspetta che per ottenere un 6 siano necessari, mediamente, 6 lanci.

Proposizione

$$Se\ X \sim \tilde{G}(p) \quad allora \quad E[X] = \frac{1}{p}$$

Dimostrazione

$$Si\ osservi\ che, data\ f(x) = \sum_{k=1}^{\infty} x^k = \frac{x}{1-x}, calcolando\ la\ sua\ derivata\ si\ ha\ f'(x) = \sum_{k=1}^{\infty} kx^{k-1} = \frac{1}{(1-x)^2}$$

$$\boldsymbol{E[x] = \sum_{k=1}^{\infty} k\, p(1-p)^{k-1} = p \sum_{k=1}^{\infty} k(1-p)^{k-1} = p\, f'(1-p) = p \frac{1}{\left(1-(1-p)\right)^2} = p\frac{1}{p^2} = 1/p}$$

Esempio

Due dadi equilibrati vengono lanciati, separatamente, più volte. Sia X il numero di lanci necessari ad ottenere 1 con il primo dado e Y il numero di lanci necessari ad ottenere 5 o 6 con il secondo.

a) Determinare la densità di X e Y.

$\boldsymbol{X \sim \tilde{G}\left(\frac{1}{6}\right) \quad Y \sim \tilde{G}\left(\frac{2}{6}\right)}$

b) Determinare la media di X e Y

$\boldsymbol{E[X] = 6 \quad E[Y] = 3}$

c) Calcolare la densità di Z = max (X, Y), qual è la sua media?

***Supponendo X e Y indipendenti, si calcola la funzione di ripartizione di Z*:**

$$\boldsymbol{F_Z = P(X \le k)\, P(Y \le k) = \left(1 - P(X > k)\right)\left(1 - P(Y > k)\right) = \left(1 - \left(1 - \frac{1}{6}\right)^k\right)\left(1 - \left(1 - \frac{1}{3}\right)^k\right) = \left(1 - \left(\frac{5}{6}\right)^k\right)\left(1 - \left(\frac{2}{3}\right)^k\right)}$$

Media?

Varianza, covarianza e indice di correlazione

Se X è una variabile aleatoria discreta, si definisce la **VARIANZA** di X ricordando la definizione di varianza di un carattere in statistica descrittiva:

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

Se X è una variabile di Bernoulli $X \sim B(1, p)$ (allora $X = X^2$), si ha:

$$Var(X) = \sigma_X^2 = E[X^2] - E[X]^2 = E[X] - E[X]^2 = p - p^2 = p(1 - p)$$

Lemma

Date due variabili aleatorie si ha: $\boldsymbol{Var(X + Y) = Var(X) + Var(Y) + 2Cov(X, Y)}$
Dove $2Cov(X, Y) = E[XY] - E[X]E[Y]$.

Se X e Y sono indipendenti si ha: $\boldsymbol{Var(X + Y) = Var(X) + Var(Y)}$

Dimostrazione

$$Var(X + Y) = E[(X + Y)^2] - (E[X + Y])^2 = E[(X^2 + Y^2 + 2XY)] - (E[X] + E[Y])^2 =$$

$$= E[X^2] + E[Y^2] + 2E[XY] - E[X]^2 - E[Y]^2 - 2E[X]E[Y] = Var(X) + Var(Y) + 2Cov(X, Y)$$

$Sia\ X \sim B(n, p)\ \ allora\ \ Var(X) = np(1 - p)$

$Sia\ X \sim P(\lambda), il\ parametro\ \lambda$ è $il\ limite\ di\ una\ successione\ di\ variabili\ B\left(n, \frac{\lambda}{n}\right)\ aventi\ varianza\ n\frac{\lambda}{n}\left(1 - \frac{\lambda}{n}\right)$ $allora\ \ Var(X) = \lambda.\ Il\ parametro\ \lambda$ è $allora\ sia\ la\ media\ che\ la\ varianza\ della\ variabile\ X.$

Come per i caratteri si definisce il **COEFFICIENTE DI CORRELAZIONE** come:

$$\rho_{X,Y} = \frac{Cov(X, Y)}{\sigma_X \sigma_Y}$$

dove $\sigma_X = \sqrt{Var\ (X)}$ è lo **SCARTO QUADRATICO MEDIO**.

Esempio

In uno schema successo-insuccesso senza rimpiazzo, si considerino le variabili di Bernoulli X_1 e X_2 relative alle prime 2 estrazioni. Determinare il coefficiente di correlazione.

Sapendo che X_1 e X_2 sono variabili di Bernoulli $B\left(1, \frac{b}{b+r}\right)$, si ha che

$$\sigma_{X_1} = \sigma_{X_2} = \sqrt{Var(X_1)} = \sqrt{\frac{b}{b + r} - \left(\frac{b}{b + r}\right)^2} = \sqrt{\frac{b^2 + br - b^2}{(b + r)^2}} = \sqrt{\frac{br}{(b + r)^2}}$$

$$\boldsymbol{\sigma_{X_1}\sigma_{X_2} = \frac{br}{(b + r)^2}}$$

Per determinare la covarianza è necessario calcolare la media di X_1X_2: è una variabile di Bernoulli di parametro $\frac{b}{b+r} \cdot \frac{b-1}{b+r-1}$ che è anche la media, quindi:

$$\boldsymbol{Cov(X_1, X_2)} = E[X_1X_2] - E[X_1]E[X_2] = \frac{b}{b + r} \cdot \frac{b - 1}{b + r - 1} - \frac{b}{b + r} \cdot \frac{b}{b + r} = \frac{b^2 - b}{(b + r)(b + r - 1)} - \frac{b^2}{(b + r)^2} =$$

$$= \frac{(b^2 - b)(b + r) - b^2(b + r - 1)}{(b + r)^2(b + r - 1)} = \frac{b^3 - b^2 + b^2r - br - b^3 - b^2r + b^2}{(b + r)^2(b + r - 1)} = \boldsymbol{\frac{-br}{(b + r)^2(b + r - 1)}}$$

$$\boldsymbol{\rho_{X_1,X_2} = \frac{Cov(X_1, X_2)}{\sigma_{X_1}\sigma_{X_2}} = \frac{\frac{-br}{(b + r)^2(b + r - 1)}}{\frac{br}{(b + r)^2}} = \frac{-br}{(b + r)^2(b + r - 1)} \cdot \frac{(b + r)^2}{br} = -\frac{1}{b + r - 1}}$$

Esempio

Si lancia un dado per 4 volte. Sia X la variabile che conta quante volte è uscito il 6 e Y la variabile che conta il numero di volte in cui è uscito un numero dispari seguito da un numero pari:

a) Determinare le densità marginali di X e Y

$X \sim B\left(4, \frac{1}{6}\right)$

$Sia\ Y_i = \begin{cases} 1\ se\ il\ lancio\ i\ è\ dispari\ e\ il\ lancio\ i+1\ è\ pari \\ 0\ altrimenti \end{cases}$

$Allora\ Y = Y_1 + Y_2 + Y_3.\quad Y\ può\ assumere\ solo\ valori\ compresi\ tra\ 0\ e\ 2, in\ particolare:$

$P(Y=2) = \frac{1}{16}\ in\ quanto\ vi\ è\ una\ sola\ combinazione\ (su\ 2^4 = 16\ possibili)\ con\ due\ sequenze\ DP: DPDP.$

$P(Y=1) = P(Y_2 = 1) + P(Y_1 = 1)P(Y_3 = 0) + P(Y_1 = 0)P(Y_3 = 1) = \frac{1}{4} + \frac{1}{4}\cdot\frac{3}{4} + \frac{3}{4}\cdot\frac{1}{4} = \frac{4+3+3}{16} = \frac{10}{16} = \frac{5}{8}$

$P(Y=0) = 1 - P(Y=2) - P(Y=1) = 1 - \frac{1}{16} - \frac{10}{16} = \frac{5}{16}$

b) Detta $p_{X,Y}$ la densità congiunta di X e Y, determinare $p_{X,Y}$ (3,1)

$p_{X,Y}(3,1)\ è\ la\ probabilità\ che\ esca\ 3\ volte\ il\ numero\ 6\ e\ che\ esca\ una\ sequenza\ DP.$

$Vi\ sono\ e\ possibili\ combinazioni: D666, 6D66, 66D6.\ \ p_{X,Y}(3,1) = 3\cdot\frac{1}{2}\cdot\frac{1}{6^3} = \frac{3}{432} = \frac{1}{144}$

c) Le variabili X e Y sono indipendenti?

$p_{X,Y}(3,1) = \frac{1}{144} \sim p_X(3)p_Y(1) = \binom{4}{3}\frac{5}{6}\frac{1}{6^3}\ \frac{5}{8} = \frac{5}{324}\ \frac{5}{8} = \frac{25}{2592}$

d) Determinare la varianza di X e di Y.

$E[X^2] - E[X]^2 = np(1-p) = 4\frac{1}{6}\frac{5}{6} = \frac{20}{36} = \frac{5}{9}\qquad ???\ negli\ appunti\ del\ docente\ c'è\ un\ risultato\ diverso$

$E[Y] = E[Y_1] + E[Y_2] + E[Y_3] = 2\cdot\frac{1}{16} + 1\cdot\frac{5}{8} + 0\cdot\frac{5}{16} = \frac{12}{16} = \frac{3}{4}\qquad E[Y]^2 = \frac{9}{16}$

$E[Y^2] = 2^2\frac{1}{16} + 1^2\frac{5}{8} + 0^2\frac{5}{16} = \frac{4}{16} + \frac{10}{16} = \frac{14}{16} = \frac{7}{8}$

$Var(Y) = E[Y^2] - E[Y]^2 = \frac{7}{8} - \frac{9}{16} = \frac{14-9}{16} = \frac{5}{16}$

Disuguaglianza di Chebyshev

$Sia\ X\ una\ variabile\ aleatoria\ discreta, allora\ preso\ un\ qualunque\ \eta > 0\ si\ ha: P(|X - E[X]| > \eta) \leq \frac{Var(X)}{\eta^2}$

Dimostrazione

$Si\ ricordi\ la\ funzione\ caratteristica\ di\ A: \chi_A(\omega) = \begin{cases} 1\ se\ \omega \in A \\ 0\ se\ \omega \notin A \end{cases}\quad e\ che\ \chi_A \sim B(1, P(A)).\quad Si\ ha\ E[\chi_A] = P(A).$

$Confrontando\ le\ due\ variabili\ aleatorie\ Y = \eta^2\chi_A\ \ e\ \ Z = (X - E[X])^2\ si\ osserva\ che:$

$1)\ Z(\omega) \geq Y(\omega)\ per\ ogni\ \omega \in \Omega, infatti\ se\ Y(\omega) = \eta^2\ \ Z(\omega) > \eta^2\ per\ definizione\ di\ A$

$2)\ se\ \omega \notin A, allora\ Y(\omega) = 0\ e\ Z(\omega) \geq 0\ \ infatti\ Z\ è\ un\ quadrato, quindi\ non\ è\ mai\ negativo$

$Allora\ E[Z] \geq E[Y].\ \ Ma\ E[Z] = Var(X)\ \ mentre\ E[Y] = \eta^2 P(|X - E[X]| > \eta).$

$Quindi\ E[Z] = Var(X) \geq E[Y] = \eta^2 P(|X - E[X]| > \eta)\quad da\ cui\quad P(|X - E[X]| > \eta) \leq \frac{Var(X)}{\eta^2}$

Legge dei grandi numeri

Se si lancia una moneta (non truccata) *n* volte ottenendo *k* volte testa, è difficile che k/n sia esattamente uguale a ½ anche se, intuitivamente, se il numero di lanci è molto elevato, il rapporto k/n non dovrebbe essere di molto differente da ½ . Si tratta di una sequenza di variabili X_1, X_2, …, X_n di densità B(1, ½) e si vuole valutare la variabile:

$$\bar{X}_n = \frac{1}{n}(X_1 + X_2 + \ldots + X_n)$$

Definizione

$Se\ X_1, X_2, \ldots,$ *è una successione di variabili aleatorie, si dice che esse convergono in probabilità alla variabile aleatoria* X *se, per ogni* η *fissato, si ha:*

$$\lim_{n\to+\infty} P(|X_n - X| > \eta) = 0$$

Definizione: LA LEGGE DEI GRANDI NUMERI

cia $X_n, n \in \mathbb{N}$ *una successione di variabili aleatorie indipendenti ed aventi tutte la stessa densità (con media* μ *e varianza* σ^2*) allora, posto* $\bar{X}_n = \frac{1}{n}(X_1 + X_2 + \cdots + X_n)$ *la successione* $\bar{X}_n$ *converge in probabilità alla variabile aleatoria costante* μ.

Dimostrazione

Si osservi che $E[\bar{X}_n] = \mu$ *e* $Var(\bar{X}_n) = \frac{\sigma^2}{n}$*, applicando la disuguaglianza di Chebyshev alla variabile* $\bar{X}_n$:

$$P(|\bar{X}_n - \mu| > \eta) \le \frac{Var(\bar{X}_n)}{\eta^2} = \frac{\sigma^2}{n\eta^2} \to 0$$

Legge dei grandi numeri per la varianza

Sotto le stesse ipotesi definite per la legge rei grandi numeri, si ha che se:

$$\bar{\sigma}_n^2 = \frac{1}{n}\sum_{i=1}^{n}(X_i - \bar{X}_n)^2$$

allora la successione $\bar{\sigma}_n^2$ *converge in probabilità alla variabile costante* σ^2

Dimostrazione

Per semplicità di notazione si pone $Y_i = X_i^2$ *per ogni i e si definisce* $\bar{Y}_n$*come per la leggedei grandi numeri; per la varianza vale la formula:* $\bar{\sigma}_n^2 = \bar{Y}_n - \bar{X}_n^2$
cioé la varianza è la media dei quadrati meno il quadrato della media.
Si ha che $\bar{Y}_n$ *converge in probabilià a* $E[Y_1] = E[X_1^2]$ *mentre* $\bar{X}_n^2$ *converge in probabilità a* $E[X_1]^2$.
In conclusione si ha che $\bar{\sigma}_n^2$ *converge a* $E[X_1^2] - E[X_1]^2 = \sigma^2$.

Osservazione

Considerando il problema di un'indagine statistica in cui X_i rappresenta il valore assunto da un carattere sull'i-esimo elemento di un campione di ne elementi, allora X_n non è altro che la media del campione (media campionaria).

Variabili aleatorie continue

Una variabile aleatoria $X: \Omega \to \mathbb{R}$ si dice continua se assume una quantità non numerabile di valori.

Ricordando la funzione di ripartizione $F(x) = P(X \leq x)$ si osserva che è una funzione di variabili reali a valori reali e si può notare che:

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

Lemma

Se la funzione di ripartizione F di una variabile aleatoria X è continua, allora $P(X = 0) = 0$ per ogni $x \in \mathbb{R}$.

Dimostrazione

$$P(X = x) \leq P\left(x - \frac{1}{n} < X \leq x\right) = F(x) - F\left(x - \frac{1}{n}\right) \to_{n \to \infty} = 0$$

Si ha come conseguenza che tutti gli eventi $\{X < x\}\ e\ \{X \leq x\}$ hanno la stessa probabilità. Non ha senso, per le variabili continue, utilizzare la densità definita per le variabili discrete.

Una funzione $f: \mathbb{R} \to \mathbb{R}$ integrabile si dice densità per X se: $P(a \leq X \leq b) = \int_a^b f(t)dt$

Si ha anche che: $P(-\infty \leq X \leq +\infty) = \int_{-\infty}^{+\infty} f(t)dt$

Se $f: \mathbb{R} \to \mathbb{R}$ è la densità di una variabile aleatoria X allora:

1) $f(t) \geq 0 \qquad \forall t \in \mathbb{R}$
2) $\int_{-\infty}^{+\infty} f(t)dt = 1$

Se f soddisfa queste due condizioni si dice **DENSITÀ CONTINUA**.

Nota: se una variabile X ammette densità allora la sua funzione di ripartizione è continua.

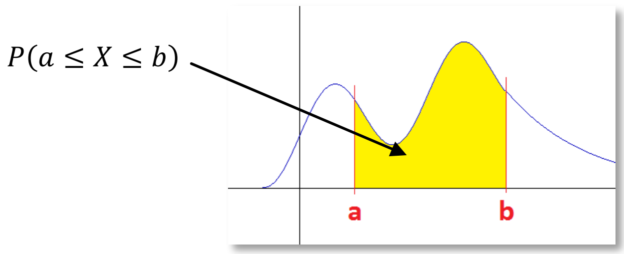
Conoscendo la densità di f è possibile calcolare la funzione di ripartizione F calcolando un integrale:

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

Viceversa, ricordando il teorema fondamentale del calcolo integrale che se F è derivabile e $F'(x) = f(x)$ allora:

$$P(x \in [a, b]) = P(a \leq X \leq b) = F(b) - F(a) = \int_a^b f(t)dt$$

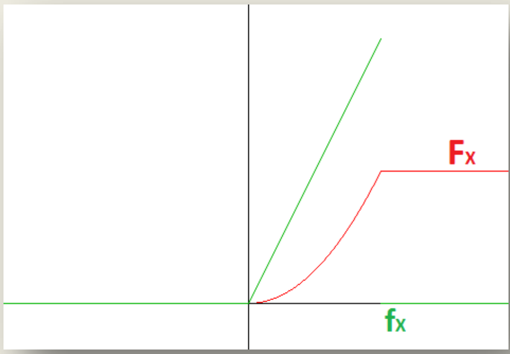
si osserva che se una variabile continua X ha funzione di ripartizione $F(x)$ derivabile (tranne, al più, in un numero finito di punti) allora $f(x) = F'(x)$ è la densità di X.



Prof. Fabrizio Caselli - fabrizio.caselli@unibo.it
Appunti di Nicola Gentili – nicola.gentili2@studio.unibo.it

Esempio

$$F_X(t) = \begin{cases} 0 & se\ t \leq 0 \\ t^2 & se\ 0 < t < 1 \\ 1 & se\ t \geq 0 \end{cases}$$



Si definisce la densità:

$$f_X(t) = \begin{cases} 0 & se\ t \leq 0 \\ 2t & se\ 0 < t < 1 \\ 0 & se\ t \geq 0 \end{cases}$$

F_X è $una\ probabilità, quindi$ è $sempre\ 0 \leq F_X \leq 1.$
$F_X(t) = P(X \leq t) \quad \Rightarrow \quad 0 \leq F_X(t) \leq 1 \quad \forall t \in \mathbb{R}.$
F_X è $monotona\ non\ descrescente$: $F_X(a) \geq F_X(b)\ \ con\ \ a \geq b.$
$\lim_{t \to +\infty} F_X\ (t) = 1 \qquad \lim_{t \to -\infty} F_X\ (t) = 0$

Media e Varianza di variabili continue

Nel caso discreto la media è definita come:

$$E[X] = \sum_{x \in \mathbb{R}} x\ p(x)$$

Nel caso continuo si ha che la media è definita come:

$$E[X] = \int_{\mathbb{R}} x\ f(x)\ dx$$

La media è lineare anche nel caso di variabili continue:

$$E[aX + bY] = aE[X] + bE[Y]$$

$$Var(x) = E[(X - E[X]^2)] = E[X^2] - E[X]^2$$

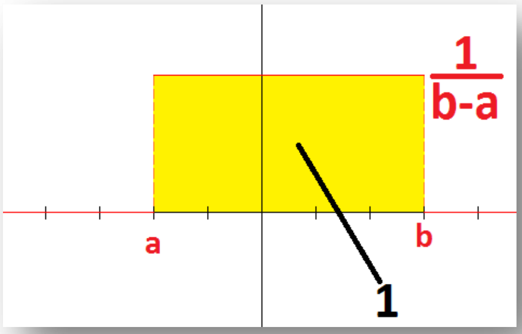
$$Var(aX) = a^2 Var(X)$$

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

Densità uniformi

La variabile X è uniforme, nell'intervallo $[a,b]$, se f_X è costante in $[a,b]$ e nulla fuori:

$$f_x = \begin{cases} \frac{1}{b-a} & se\ x \in [a,b] \\ 0 & se\ x \notin [a,b] \end{cases}$$



$$\boldsymbol{E[X]} = \int_{\mathbb{R}} x\, f(x)dx = \int_{\mathbb{R}} x \frac{1}{b-a} dx = \int_a^b x \frac{1}{b-a} dx = \frac{1}{b-a}\left[\frac{x^2}{2}\right]_a^b =$$

$$= \left(\frac{1}{b-a}\right)\frac{1}{2}(b^2-a^2) = \left(\frac{1}{b-a}\right)\frac{1}{2}(b+a)(b-a) = \boldsymbol{\frac{b+a}{2}}$$

$$F_X(t) = \int_{-\infty}^{t} f(x)dx = \begin{cases} 0 & se\ t < a \\ \int_{-\infty}^{t} \frac{1}{b-a} dx & se\ a \le t \le b \\ 1 & se\ t > b \end{cases}$$

Per calcolare la varianza: $Var(X) = E[X^2] - E[X]^2$

$$\boldsymbol{E[X^2]} = \int_{-\infty}^{+\infty} x^2\, f(x)dx = \int_a^b x^2\, f(x)dx = \int_a^b x^2 \frac{1}{b-a} dx = \frac{1}{b-a}\left[\frac{x^3}{3}\right]_a^b = \frac{1}{b-a}\frac{(b^3-a^3)}{3} = \boldsymbol{\frac{a^2+ab+b^2}{3}}$$

$$\boldsymbol{Var(X)} = \frac{a^2+ab+b^2}{3} - \left(\frac{b+a}{2}\right)^2 = \frac{a^2+ab+b^2}{3} - \frac{b^2+2ab+b^2}{4} = \frac{4a^2+4ab+4b^2-3a^2-6ab-3b^2}{12} =$$

$$= \frac{a^2-2ab+b^2}{12} = \boldsymbol{\frac{(b-a)^2}{12}}$$

Densità esponenziale

La funzione $f(t) = \begin{cases} \lambda e^{-\lambda t} & se\ t \ge 0 \\ 0 & altrimenti \end{cases}$ con λ parametro reale positivo, è una densità continua:

1) $f(t) \ge 0 \quad \forall t \in \mathbb{R}$

2) $\int_{-\infty}^{+\infty} f(t)dt = \int_0^{+\infty} \lambda e^{-\lambda t} dt = \left[-\lambda \frac{1}{\lambda} e^{-\lambda t}\right]_0^{+\infty} = 0 + 1$

Una variabile X che ha una tale densità è detta **DENSITÀ ESPONENZIALE** di parametro λ e si indica con: $X \sim Exp(\lambda)$.

Una variabile esponenziale soddisfa la proprietà di mancanza di memoria:

$$\boldsymbol{P}(\boldsymbol{X} > t+s | \boldsymbol{X} > t) = \frac{\boldsymbol{P}(\boldsymbol{X} > t+s)}{\boldsymbol{P}(\boldsymbol{X} > t)} = \frac{e^{-\lambda(t+s)}}{e^{-\lambda t}} = e^{-\lambda s} = P(X > s)$$

Cambi di variabile (x^2, ax + b)

Se X è una variabile aleatoria continua e $\phi: \mathbb{R} \to \mathbb{R}$, non è detto che $\phi(X)$ sia a sua volta una variabile aleatoria. Se ϕ è continua (tranne al più in un numero finito di punti), allora $\phi(X)$ è una variabile aleatoria.

Se X ha densità f, si può calcolare la densità di X^2, infatti, sia G la funzione di ripartizione di X^2, allora:

$$G(t) = P(X^2 \leq t) = P(-\sqrt{t} \leq X \leq \sqrt{t}) = F_X(\sqrt{t}) - F_X(-\sqrt{t})$$

Se G è derivabile allora, la densità g di X^2 è:

$$g(t) = G'(t) = -\frac{1}{2\sqrt{t}}\left(f\left(\sqrt{t} - f(-\sqrt{t})\right)\right)$$

Se X ha densità f, si può calcolare la densità di a X + b (con a ≠ 0):

$$G(t) = P(aX + b \leq t) = P\left(X \leq \frac{t-b}{a}\right) = F_X\left(\frac{t-b}{a}\right)$$

Se G è derivabile allora, la densità g di a X + b è:

$$g(t) = G'(t) = \frac{1}{a} f\left(\frac{t-b}{a}\right) \quad se \;\; a > 0, \quad in\ generale: g(t) = \frac{1}{|a|} f\left(\frac{t-b}{a}\right)$$

Indipendenza

$X\ e\ Y\ sono\ indipedenti\ se\ \forall\ coppia\ di\ intervalli\ I\ e\ J\ si\ ha\ che\colon\ P(X \in I, Y \in J) = P(X \in I)P(Y \in J)$

$Se\ I\ e\ J\ sono\ illimitati\ a\ sinistra, allora\colon P(X \leq t, Y \leq s) = P(X \leq t)P(X \leq s) = F_X(t)F_Y(s)$

Esempio:
Se $X \sim Exp(\lambda)$ e $Y \sim Exp(\mu)$ sono variabili indipendenti, la densità della variabile $Z = \max(X, Y)$ può essere calcolato come:

$$F_Z(t) = P(X \leq t)P(Y \leq t) = F_X(t)F_Y(t) = (1 - e^{-\lambda t})(1 - e^{-\mu t})$$

Derivando si ha:

$$F'_Z(t) = f_Z(t) = \lambda e^{-\lambda t}(1 - e^{-\lambda t}) + \mu e^{-\mu t}(1 - e^{-\mu t})$$

$\min(X, Y) = W \sim Exp(\lambda + \mu)$

Esempio:
Sia X una variabile continua uniforme nell'intervallo [0,1] e Y la variabile (discreta) data da
$Y \;=\; \lfloor 5X \rfloor \;\; \lfloor x \rfloor\colon parte\ intera$
Determinare la densità di Y.

$La\ densità\ di\ X\ è\ f_X(t) = \begin{cases} 1\ se\ t \in [0,1] \\ 0\ altrimenti \end{cases} \quad la\ densità\ di\ Y\ è\ data\ da\colon f_Y(t) = \frac{1}{5} f\left(\frac{t}{5}\right) = \begin{cases} \frac{1}{5} & se\ \ t \in [0,1] \\ 0 & altrimenti \end{cases}$

Esempio:
Una certa pianta ha tempo di vita esponenziale di media 8 anni (e quindi parametro 1/8). Si acquistano 2 piante, per quanto tempo mediamente vivrà almeno una delle due piante?

$$X \sim Exp\left(\frac{1}{8}\right)\colon tempo\ di\ vita\ medio\ di\ una\ pianta \quad \min(X_1, X_2) = W \sim Exp\left(\frac{1}{8} + \frac{1}{8}\right) = W \sim Exp\left(\frac{1}{4}\right) \to E[W]$$

Esempio:

Si consideri la funzione $f: \mathbb{R} \to \mathbb{R}$ $\quad f(t) = \begin{cases} e^{-2t} & se\ t \geq 0 \\ e^{2t} & se\ t \leq 0 \end{cases}$

1) Verificare che f è la densità di una variabile aleatoria X

$$a)\ f(t) \geq 0 \quad \forall\, t \in \mathbb{R}$$

$$b)\ \int_{-\infty}^{+\infty} f(t)dt = 2\int_0^{+\infty} e^{-2t}dt = 1$$

2) Determinare la funzione di ripartizione di X

$$se\ t < 0 \qquad F_X(t) = \int_{-\infty}^{t} e^{2x}dx = \left[\frac{1}{2}e^{2x}\right]_{-\infty}^{t} = \frac{1}{2}e^{2t}$$

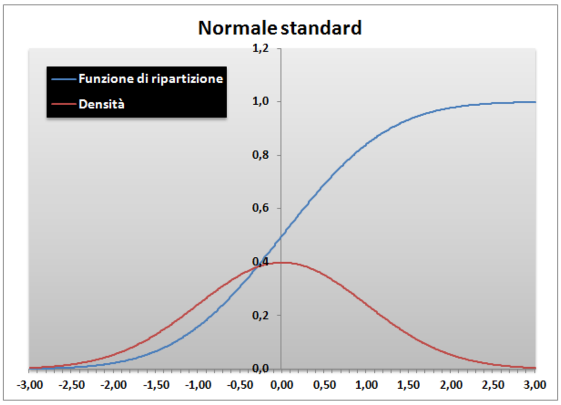
$$se\ t > 0 \qquad F_X(t) = \int_{-\infty}^{t} e^{-2x}dx = \int_{-\infty}^{0} e^{2x}dx + \int_0^t e^{-2x}dx = 1 - \frac{1}{2}e^{-2t}$$

3) Determinare la media di X

4) Determinare la probabilità che X > 1

$$P(X > 1) = 1 - P(X \leq 1) = 1 - F_X(1) = 1 - \left(1 - \frac{1}{2}e^{-2}\right) = \frac{1}{2}e^{-2} = 6{,}7\%$$

$Se\ \phi(x): \mathbb{R} \to \mathbb{R}\ allora\ E[\phi(x)] = \int_{\mathbb{R}} \phi(x)\, f(x)\, dx$



Densità normali

Sapendo che

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

allora

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

è una densità che viene detta **NORMALE STANDARD** o **GAUSSIANA STANDARD** e si indica con $X \sim N(0,1)$.
Per indicare una densità normale standard si utilizza il simbolo ζ_0.
Siano σ > 0 e μ due parametri reali e $\zeta_0 \sim N(0,1)$, considerando la variabile $\zeta = \sigma\zeta_0 + \mu$ si ha che essa avrà densità

$$f_\zeta(y) = \frac{1}{\sigma} f\zeta_0\left(\frac{y-\mu}{\sigma}\right) = \frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{(y-\mu)^2}{2\sigma^2}}$$

In questo caso si scrive: $\zeta \sim N(\mu, \sigma^2)$.

Quando si ha a che fare con una variabile $\zeta \sim N(\mu, \sigma^2)$ è utile effettuare il cambio di variabile: $\zeta_0 = \frac{\zeta - \mu}{\sigma}$ dove ζ_0 è una normale standard. In questo modo è possibile effettuare i calcoli con la variabile ζ_0 e tradurre i risultati nella variabile ζ.

Per indicare la funzione di ripartizione di una variabile standard, si usa spesso $\Phi(x)$. I valore di questa funzione sono riportati in apposite tabelle e si osserva che per valori di x < -3 essa assume valori molto prossimi a 0 e per valori di x > 3 essa assume valori molto prossimi a 1. È pertanto sufficiente calcolare i valori di $\Phi(x)$ solo tra -3 e 3.

Grazie alla simmetria della densità normale, si ha inoltre che $\Phi(-x) = 1 - \Phi(x)$.

$$E[\zeta_0] = \int_{-\infty}^{+\infty} x \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx = 0 \qquad lo\ si\ deriva\ dal\ fatto\ che\ la\ funzione\ è\ dispari.$$

$$E[\zeta_0^2] = \int_{-\infty}^{+\infty} x^2 \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx = \int_{-\infty}^{+\infty} x \;\; x \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx = (*)$$

Integrazione per parti: $\int f'(x)g(x)dx = [f(x)g(x)] - \int f(x)g'(x)dx \rightarrow \; f = x \quad f' = 1 \quad g = -\frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} \quad g' = x\frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$

$$(*) = \left[-x\frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}\right]_{-\infty}^{+\infty} - \int_{-\infty}^{+\infty} -\frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx = 0 - (-1) = 1$$

$$Var(\zeta_0) = E[\zeta_0^2] - E[\zeta_0]^2 = 1 - 0 = 1$$

$Sfruttando\ il\ fatto\ che\ se\ \zeta \sim N(\mu, \sigma^2)\ \ allora\ \ \zeta = \mu + \sigma\zeta_0\ \ e\ la\ linearità\ della\ media, si\ calcolano$:

$$E[\zeta] = E[\mu + \sigma\zeta_0] = E[\mu] + E[\sigma\zeta_0] = E[\mu] + \sigma E[\zeta_0] = \mu + \sigma 0 = \mu$$

$Var(\zeta) = Var(\mu + \sigma\zeta_0)$ si consideri che μ è costante e certamente μ e $\sigma\zeta_0$ sono indipendenti, allora:

$$Var(\zeta) = Var(\mu + \sigma\zeta_0) = Var(\mu) + Var(\sigma\zeta_0) = 0 + \sigma^2 Var(\zeta_0) = \sigma^2 1 = \sigma^2$$

Teorema del limite centrale

Considerando una sequenza di variabili uniformi nell'intervallo [-1, 1] tutte indipendenti tra loro: X_1, X_2, …

La somma di questa sequenza di variabili, approssima sempre meglio la forma a parabola della densità uniforme, all'aumentare dei termini sommati.

Questa proprietà vale anche per le variabili discrete.

Siano: $S_n = X_1 + X_2 + \ldots + X_n$ e $\bar{X}_n = \frac{1}{n}(S_n)$ se le variabili X_i hanno tutte la stessa densità di media μ e varianza σ^2 e sono indipendenti allora:

$$E[S_n] = n\mu \qquad Var(S_n) = n\sigma^2$$

$$E[\bar{X}_n] = \frac{1}{n}E[X_1 + X_2 + \cdots + X_N] = \frac{1}{n}(\mu + \mu + \ldots + \mu) = \frac{1}{n}n\mu = \mu \qquad Var(\bar{X}_n) = \frac{\sigma^2}{n}$$

Teorema del limite centrale

Sia $X_1, X_2, \ldots$ una successione di variabili aleatorie indipendenti aventi tutte la stessa densità (discreta o continua) di media μ e varianza σ^2; allora per n abbastanza grande si ha, con buona approssimazione:

$$S_n \sim N(n\mu, n\sigma^2) \qquad \bar{X}_n \sim N\left(n, \frac{\sigma^2}{n}\right)$$

Esempio:

Sia $f(x) = \begin{cases} e^{-x} & se\ x \geq 0 \\ 0 & se\ x \leq 0 \end{cases}$

1) <u>f è una densità?</u>

È necessario verificare che siano verificate le due proprietà delle densità:

$a)\ f(x) \geq 0\ \forall x \in \mathbb{R}$ ✓

$b)\ \int_{-\infty}^{+\infty} f(x)dx = \int_{-\infty}^{0} 0dx + \int_{0}^{+\infty} e^{-x}dx = 0 + [-e^{-x}]_0^{+\infty} = 0 - (-1) = 1$ ✓

f è una densità.

2) <u>Detta X una variabile di densità $f(x)$ calcolare $F(x)$</u>

$$F_X(t) = \begin{cases} \int_0^t f(x)dx = \int_0^t e^{-x}dx = [-e^{-x}]_0^t = 1 - e^{-t} \\ 0 \quad se\ t \leq 0 \end{cases}$$

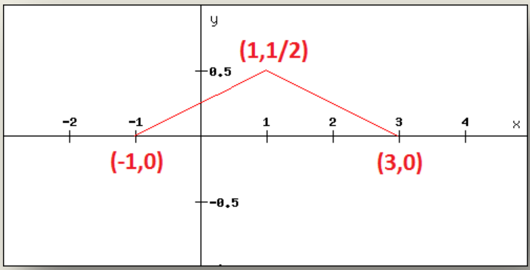
3) <u>Detta X una variabile di densità $f(x)$ calcolare $E[X]$</u>

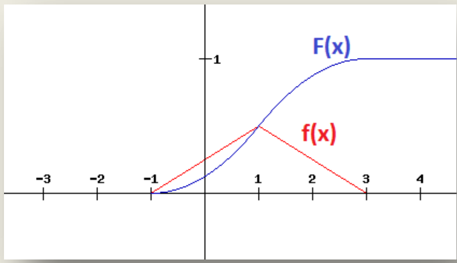
$$E[X] = \int_{-\infty}^{+\infty} x\, f(x)dx = \int_0^{+\infty} x\, e^{-x}dx = (*)$$

Integrazione per parti: $\int f'(x)g(x)dx = [f(x)g(x)] - \int f(x)g'(x)dx \rightarrow \quad f = -e^{-x} \quad f' = e^{-x} \quad g = x \quad g' = 1$

$$(*) = [-xe^{-x}]_0^{+\infty} - \int_0^{+\infty} -e^{-x}dx = 0 + [-e^{-x}]_0^{+\infty} = 0 - (-1) = 1$$

Esempio:
Sia f definita graficamente:





1) <u>Scrivere la formula di f</u>

Si ricorda l'equazione della retta: $y = mx + q \rightarrow \quad m = \dfrac{f(x_2) - f(x_1)}{x_2 - x_1} \quad q = y - mx$

$$f \text{ è definita in quattro parti: } f(x)\begin{cases} se\ x < -1 & \Rightarrow f(x) = 0 \\ se\ -1 < x < 1 \Rightarrow m = \dfrac{\frac{1}{2} - 0}{1-(-1)} = \dfrac{1}{4};\, q = \dfrac{1}{2} - \dfrac{1}{4}\cdot 1 = \dfrac{1}{4} & \Rightarrow f(x) = \dfrac{x+1}{4} \\ se\ 1 < x < 3 \quad \Rightarrow m = \dfrac{0 - \frac{1}{2}}{3-1} = -\dfrac{1}{4};\, q = \dfrac{1}{2} + \dfrac{1}{4}\cdot 1 = \dfrac{3}{4} & \Rightarrow f(x) = \dfrac{3-x}{4} \\ se\ x > 3 & \Rightarrow f(x) = 0 \end{cases}$$

2) <u>Calcolare F</u>

Anche F deve essere valutata in 4 parti:

$$F_X(t) = \begin{cases} se\ t < -1 \quad \Rightarrow F_X(t) = 0 \\ se\ -1 < t < 1 \Rightarrow F_X(t) = \displaystyle\int_{-1}^{t} \frac{x+1}{4}dx = \frac{1}{4}\int_{-1}^{t} x dx + \frac{1}{4}\int_{-1}^{t} 1 dx = \frac{1}{4}\left(\left[\frac{x^2}{2}\right]_{-1}^{t} + [x]_{-1}^{t}\right) = \frac{1}{4}\left(\frac{t^2}{2} - \frac{1}{2} + t + 1\right) = \\ \quad = \dfrac{1}{4}\left(\dfrac{t^2 - 1 + 2t + 2}{2}\right) = \dfrac{t^2 + 2t + 1}{8} \qquad si\ ha\ che\ F_X(1) = \dfrac{1^2 + 2 + 1}{8} = \dfrac{1}{2} \\ se\ 1 < t < 3 \quad \Rightarrow F_X(t) = \dfrac{1}{2} + \displaystyle\int_{1}^{t} \frac{3-x}{4}dx = \frac{3}{4}\int_{1}^{t} 1dx - \frac{1}{4}\int_{1}^{t} x dx = \frac{3}{4}[x]_1^t - \frac{1}{4}\left[\frac{x^2}{2}\right]_1^t = \frac{3t-3}{4} - \frac{1}{4}\frac{t^2-1}{2} = \\ \quad = \dfrac{1}{2} + \dfrac{-t^2 + 6t - 5}{8} \\ se\ t > 3 \quad \Rightarrow F_X(t) = 1 \end{cases}$$

3) <u>Sia X una variabile di densità f calcolare la densità di X^2 e $2X+1$</u>

Si vuole calcolare $f_{X^2}(t)$, *conviene allora calcolare la funzione di ripartizione e poi derivarla:*

$F_{X^2}(t) = P(X^2 \le t) = P(-\sqrt{t} \le X \le \sqrt{t}) = F_X(\sqrt{t}) - F_X(-\sqrt{t})$

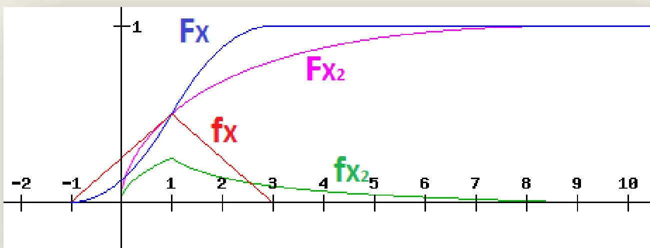
A questo punto possono verificarsi diversi casi a seconda del valore di t, infatti F_X è definita diversamente nei diversi intervalli in cui può trovarsi t; si noti che t può assumere solo valori positivi;

$se\ 0 \le t \le 1 \qquad (\sqrt{t} \in [0,1], -\sqrt{t} \in [-1,0]) \Rightarrow F_{X^2}(t) = \dfrac{t + 2\sqrt{t} + 1}{8} - \dfrac{t - 2\sqrt{t} + 1}{8} = \dfrac{4\sqrt{t}}{8} = \dfrac{\sqrt{t}}{2}$

$se\ 1 < t \le 9 \qquad (\sqrt{t} \in [1,3], -\sqrt{t} \in [-3,-1]) \Rightarrow F_{X^2}(t) = \dfrac{-t + 6\sqrt{t} - 5}{8} - 0 = \dfrac{-t + 6\sqrt{t} - 5}{8}$

$se\ t > 9 \qquad \Rightarrow F_{X^2}(t) = 1$

$$f_{X^2}(t) = \begin{cases} se\ t < 0 & 0 \\ se\ 0 \le t \le 1 & \left(\dfrac{\sqrt{t}}{2}\right)' = \dfrac{1}{4\sqrt{t}} \\ se\ 1 < t \le 9 & \left(\dfrac{-t + 6\sqrt{t} - 5}{8}\right)' = \dfrac{3}{8\sqrt{t}} - \dfrac{1}{8} \\ se\ t > 9 & (1)' = 0 \end{cases}$$

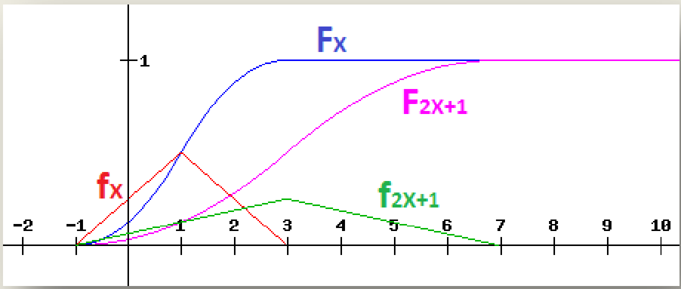


Anche per calcolare $f_{2X+1}(t)$, conviene calcolare la funzione di ripartizione e poi derivarla:

$$F_{2X+1}(t) = P(2X+1 \leq t) = P\left(X \leq \frac{t-1}{2}\right) = F_X\left(\frac{t-1}{2}\right)$$

$$\text{Allora } F_{2X+1}(t) = \begin{cases} se\ t < -1 & \Rightarrow F_{2X+1}(t) = F_X\left(\frac{t-1}{2}\right) = 0 \\ se\ -1 < t < 3 & \Rightarrow F_{2X+1}(t) = F_X\left(\frac{t-1}{2}\right) = \frac{\left(\frac{t-1}{2}\right)^2 + 2\left(\frac{t-1}{2}\right) + 1}{8} = \frac{t^2+2t+1}{32} \\ se\ 3 < t < 7 & \Rightarrow F_{2X+1}(t) = F_X\left(\frac{t-1}{2}\right) = \frac{1}{2} + \frac{-\left(\frac{t-1}{2}\right)^2 + 6\left(\frac{t-1}{2}\right) - 5}{8} = \frac{1}{2} + \frac{-t^2+14t-33}{32} \\ se\ t > 7 & \Rightarrow F_{2X+1}(t) = F_X\left(\frac{t-1}{2}\right) = 1 \end{cases}$$

$$\text{da cui } f_{2X+1}(t) = F'_{2X+1}(t) = \begin{cases} se\ t < -1 & \Rightarrow f_{2X+1}(t) = 0 \\ se\ -1 < t < 3 & \Rightarrow f_{2X+1}(t) = \left(\frac{t^2+2t+1}{32}\right)' = \frac{t+1}{16} \\ se\ 3 < t < 7 & \Rightarrow f_{2X+1}(t) = \left(\frac{-t^2+14t-33}{32}\right)' = -\frac{t-7}{16} \\ se\ t > 7 & \Rightarrow f_{2X+1}(t) = 0 \end{cases}$$



Correzione di continuità

Se X è una variabile aleatoria discreta approssimata dalla variabile normale, per ottenere risultati più accurati è preferibile effettuare una **CORREZIONE DI CONTINUITÀ**, ovvero modificare i valori interi di analisi considerandoli come intervalli di ampiezza 1 centrati sul valore.

Se ad esempio X assume solo valori interi, $X = 5$ dovrà essere sostituito con $4.5 < x < 5.5$;

oppure, $2 < X \leq 5$ dovrà essere sostituito con $1.5 < x < 5.5$;

o ancora, $X > 4$ sarà sostituito da $X > 4.5$.

Esempio (biglietti aerei):

Esempio (divise, taglie):

Studi sperimentali mostrano che l'altezza di un uomo adulto ha densità normale $N(\mu,\sigma^2)$ con $\mu = 173cm$ e $\sigma = 9cm$.

Una divisa per reclute è disponibile in 5 taglie:

Taglia	Altezza
XS	h < 162
S	162 < h < 170
M	170 < h < 180
L	180 < h < 190
XL	h > 190

Dovendo vestire 500 nuove reclute, quante divise di ogni taglia si dovrebbero acquistare?

Inviduato tramite apposita tabella

$P(\zeta < 162)$ *è la probabilità che un uomo adulto abbia un'altezza inferiore a 162 cm.*

Si effettua allora il cambio di variabile, ricordando che $\zeta_0 = \frac{\zeta-\mu}{\sigma}$:

$$P(\zeta < 162) = P\left(\frac{\zeta-173}{9} < \frac{162-173}{9}\right) = P(\zeta_0 < -1.22) = \Phi(-1.22) = 1-\Phi(1.22) = 1-0.88877 = 0.11123$$

$$P(162 < \zeta < 170) = P\left(\frac{162-173}{9} < \frac{\zeta-173}{9} < \frac{170-173}{9}\right) = P(-1.22 < \zeta_0 < -0.33) = \Phi(-0.33)-\Phi(-1.22) = \Phi(1.22)-\Phi(0.33) = 0.88877-0.62930 = 0.25947$$

$$P(170 < \zeta < 180) = P\left(\frac{170-173}{9} < \frac{\zeta-173}{9} < \frac{180-173}{9}\right) = P(-0.33 < \zeta_0 < 0.78) = \Phi(0.78)-\Phi(-0.33) = \Phi(0.78)-\big(1-\Phi(0.33)\big) = \Phi(0.78)-1+\Phi(0.33) = 0.78230-1+0.62930 = 0.41160$$

$$P(180 < \zeta < 190) = P\left(\frac{180-173}{9} < \frac{\zeta-173}{9} < \frac{190-173}{9}\right) = P(0.78 < \zeta_0 < 1.89) = \Phi(1.89)-\Phi(0.78) = 0.97062-0.7823 = 0.18832$$

$$P(\zeta > 190) = P\left(\frac{\zeta-173}{9} > \frac{190-173}{9}\right) = P(\zeta_0 > 1.89) = 1-\Phi(1.89) = 1-0.97062 = 0.02938$$

Taglia	Altezza	Probabilità	Nr.divise	
XS	h < 162	0.11123	55	(0.11123 x 500)
S	162 < h < 170	0.25947	130	(0.25947 x 500)
M	170 < h < 180	0.41160	206	(0.41160 x 500)
L	180 < h < 190	0.18832	94	(0.18832 x 500)
XL	h > 190	0.02938	15	(0.02938 x 500)

Esempio (lampadine):

Una lampadina ha tempo di vita definito da densità esponenziale di media 10 giorni.

Qual è la probabilità che 40 lampadine bastino per un anno?

Siano $X_1, X_2, \dots, X_4$ *i tempi di vita delle lampadine.* $X_i \sim Exp(10)$ *per ogni* $i = 1,\dots,40$

$S_{40} = (X_1 + X_2 + \cdots + X_{40})$

Si vuole calcolare $P(X_1 + X_2 + \cdots + X_{40} > 365) = P(S_{40} > 365)$

$\left(\boldsymbol{se\ X \sim Exp(\lambda)\quad allora\quad E[X] = \frac{1}{\lambda}\quad e\quad Var(x) = \frac{1}{\lambda^2}}\right)$

$$Allora: E[X_i] = \mu = 10 \quad e \quad Var(X_i) = \sigma^2 = \frac{1}{\frac{1}{10^2}} = 100 \qquad per\ ogni\ i = 1,\dots,40$$

Per il teorema del limite centrale si ha che: $S_{40} \sim N(40 \cdot 10, 40 \cdot 100) = N(400, 4000) \qquad S_{40} = \zeta$

$$\boldsymbol{P(S_{40} \geq 365) = P(\zeta \geq 365) = P(400 + \sqrt{4000}\,\zeta_0 \geq 365) = P\left(\zeta_0 \geq \frac{365-400}{\sqrt{4000}}\right) = P(\zeta_0 \geq -0.55) = P(\zeta_0 < 0.55) = \Phi(0.55) = 0.70884}$$

Esempio (100 lanci di dado):
Lanciando un dado 100 volte, qual è la probabilità che la somma dei lanci si maggiore o uguale a 400?

$X \sim B\left(1, \frac{1}{6}\right)$ *rappresenta il lancio del dado*

$$E[X] = \sum_{i=1}^{6} \frac{i}{6} = \frac{21}{6} \qquad Var(X) = E[X^2] - E[X]^2 = \sum_{i=1}^{6} \frac{i^2}{6} - \left(\frac{21}{6}\right)^2 = \frac{91}{6} - \frac{441}{36} = \frac{546 - 441}{36} = \frac{105}{36} = \frac{35}{12}$$

$$Allora\ si\ ha\ che\ S_{100} \sim N\left(\frac{21}{6} \cdot 100, \frac{35}{12} \cdot 100\right) = N\left(\frac{2100}{6}, \frac{3500}{12}\right) = N\left(350, \frac{3500}{12}\right)$$

$$Quindi\ P(S_{100} > 400) = P\left(350 + \sqrt{\frac{3500}{12}}\zeta_0 > 400\right) = P\left(\zeta_0 > \frac{400 - 350}{\sqrt{\frac{3500}{12}}}\right) = P\left(\zeta_0 > \frac{50}{17{,}0789}\right)$$

$$= P(\zeta_0 > 2.93) = 1 - 0.99825 = 0.00175$$

I due parametri $N\left(350, \frac{3500}{12}\right)$ *esprimono come la maggior parte dei casi si concentri attorno al valore* 350 *con uno scarto di* $\pm\sqrt{\frac{3500}{12}} = \pm 17,$ *pertanto la maggior parte dei valori ottenibili è compresa tra* 333 *e* 367, *ciò motiva la scarsa probabilità di ottenere un valore superiore a* 400.

Esempio (esame 17/09/2012):

Statistica inferenziale

Ha il compito di dedurre proprietà di una popolazione conoscendo i dati relativi ad un campione.

Intervallo di confidenza

Data una certa popolazione relativamente alla quale si vuole studiare un certo carattere x, se il carattere è quantitativo si può considerare una variabile aleatoria X data da x su un elemento scelto a caso dalla popolazione.

Se si sceglie un campione a caso di n elementi, si può considerare la variabile aleatoria X su tutti gli elementi del campione: $X_1, X_2, \dots, X_n$

La media campionaria si ottiene come

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

Conoscendo media $\mu = E[X_i]$ e $\sigma^2 = Var(X_i)$, per il **teorema centrale del limite** è possibile rispondere a domande del tipo "qual è la probabilità che $\bar{X}_n$ sia compreso tra due valori?: $P(a \le \bar{X}_n \le b)$?"

I valori di μ e σ^2 non sono però noti a priori, si vuole allora stimare la media di una variabile aleatoria sulla base di un campione.

$$Dato\ 0 < c < 1\ \ sia\ z_c\ \ tale\ che\ P(-z_c < \zeta_0 < z_c) \qquad da\ cui\ \ \ \phi(z_c) - \phi(-z_c) = c$$

$$ovvero\ \left(ricordando\ che\ \phi(-z_c) = 1 - \phi(z_c)\right)\ \ \phi(z_c) - \left(1 - \phi(z_c)\right) = 2\phi(z_c) - 1 = c$$

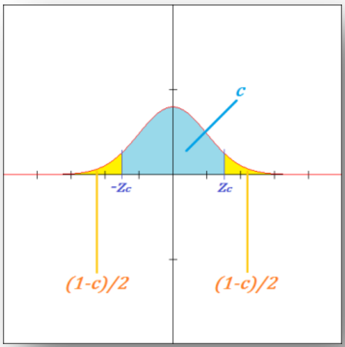
$$quindi\ \ \phi(z_c) = \frac{c+1}{2} = \frac{c}{2} + \frac{1}{2}$$

Ad esempio:

$$se\ c = 0.80 \quad \phi(z_c) = \frac{0.80+1}{2} = 0.90 \qquad z_{0.8} = 1.28$$

$$se\ c = 0.90 \quad \phi(z_c) = \frac{0.90+1}{2} = 0.95 \qquad z_{0.9} = 1.64$$

$$se\ c = 0.95 \quad \phi(z_c) = \frac{0.95+1}{2} = 0.975 \qquad z_{0.95} = 1.96$$



$$Allora\ si\ ha\ che:\ \ \frac{\bar{X}_n - \mu}{\frac{\sigma}{\sqrt{n}}} \sim N(0,1)\ \ da\ cui:$$

$$P\left(-z_c < \frac{\bar{X}_n - \mu}{\frac{\sigma}{\sqrt{n}}} < z_c\right) = c \ \Rightarrow\ P\left(-z_c\frac{\sigma}{\sqrt{n}} < \bar{X}_n - \mu < z_c\frac{\sigma}{\sqrt{n}}\right) \ \Rightarrow\ P\left(-z_c\frac{\sigma}{\sqrt{n}} - \bar{X}_n < -\mu < z_c\frac{\sigma}{\sqrt{n}} - \bar{X}_n\right) = c$$

$$\Rightarrow\ \ P\left(\bar{X}_n - z_c\frac{\sigma}{\sqrt{n}} < \mu < \bar{X}_n + z_c\frac{\sigma}{\sqrt{n}}\right) = c$$

Questa espressione rappresenta la probabilità (c) per cui la media reale (μ) è compresa in un determinato intervallo.

Non si conosce però ancora lo scarto quadratico medio.

Richiamando la **Legge dei Grandi Numeri** si ha che: $\bar{\sigma}_n^2 = \frac{1}{n}\sum_{i=1}^{n}(X_i - \bar{X}_n)^2 \to \sigma^2$ $\bar{\sigma}_n^2$ è detta VARIANZA CAMPIONARIA.

Prof. Fabrizio Caselli - fabrizio.caselli@unibo.it

Appunti di Nicola Gentili – nicola.gentili2@studio.unibo.it

Sostituendo allora σ con $\bar{\sigma}_n$ si ottiene la confidenza:

$$\Rightarrow \quad Conf\left(\bar{X}_n - z_c \frac{\bar{\sigma}_n}{\sqrt{n}} < \mu < \bar{X}_n + z_c \frac{\sigma\bar{\sigma}_n}{\sqrt{n}}\right) = c$$

Si ha dunque che la media della carattere per la popolazione è inclusa nell'intervallo $[\bar{X}_n - z_c \frac{\bar{\sigma}_n}{\sqrt{n}}, \bar{X}_n + z_c \frac{\sigma\bar{\sigma}_n}{\sqrt{n}}]$ che è detto intervallo di confidenza c per μ.

Esempio (albicocche):

Esempio (binomali):

Esempio (exit poll):

Test statistici

(inferire è sinonimo di dedurre)