ChatGPT

Insiemistica

Un **insieme** è una collezione di elementi distinti considerati come una singola entità [1]. Un insieme A è un **sottoinsieme** di B (si scrive $A \subseteq B$) se tutti gli elementi di A appartengono anche a B [2]. Ad esempio, se $A = \{1, 2\}$ e $B = \{1, 2, 3\}$, allora $A \subseteq B$. Se $A \subseteq B$ e $A \neq B$, si parla di **sottoinsieme proprio**.

Differenza e complementare: la differenza tra due insiemi A e B ($A - B$) è l'insieme degli elementi di A che non sono in B [3]. Similmente, dato un insieme universale U, il complemento di A in U (denotato A^c) è l'insieme di elementi in U che non appartengono ad A [4]. Ad esempio, se $U = \{1, 2, 3, 4\}$ e $A = \{1, 3\}$, allora $A^c = \{2, 4\}$. Il **campo di variazione** (o *range*) di un insieme di dati è la differenza tra il valore massimo e minimo.

Unione e intersezione: l'**unione** di due insiemi A e B, indicata $A \cup B$, contiene tutti gli elementi che stanno in A o in B (o in entrambi) [5]. La **intersezione**, $A \cap B$, contiene invece solo gli elementi comuni ad entrambi gli insiemi [6]. Inoltre, se $A \cap B = \varnothing$ (insiemi **incompatibili/disgiunti**), allora $P(A \cap B) = 0$. In generale la probabilità dell'unione è data da

$$P(A \cup B) = P(A) + P(B) - P(A \cap B)\,,$$

ovvero si sommano le probabilità dei singoli eventi e si sottrae quella dell'intersezione [7]. I due insiemi A e B sono **indipendenti** se la realizzazione di uno non influenza l'altro, ovvero se vale $P(A \cap B) = P(A)P(B)$ (equivalente a dire $P(A|B) = P(A)$) [8].

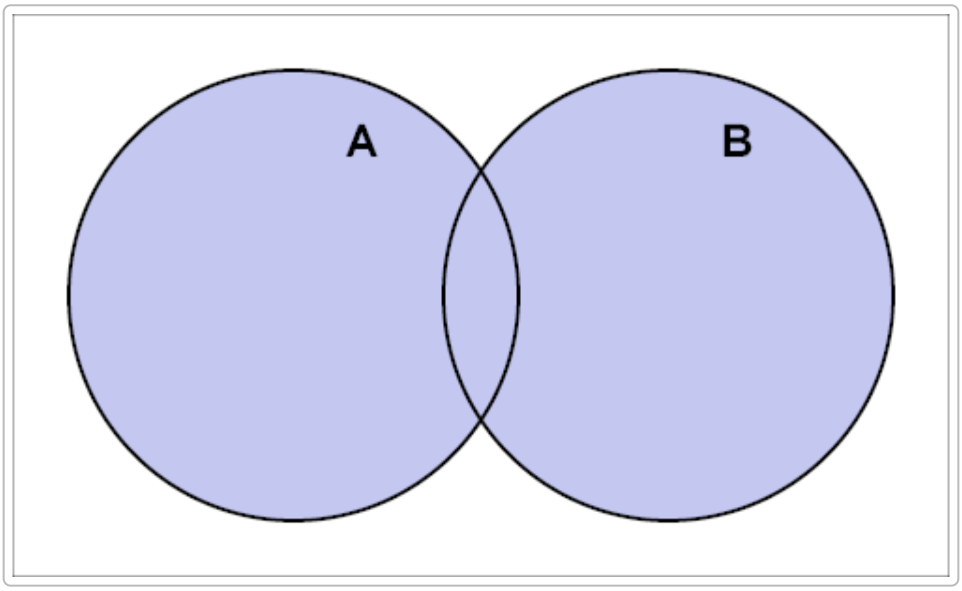

Figura – Venn diagramma dell'unione di due insiemi A e B. In un'unione i cerchi di A e B includono ogni elemento presente in almeno uno dei due insiemi, colorando entrambe le regioni.

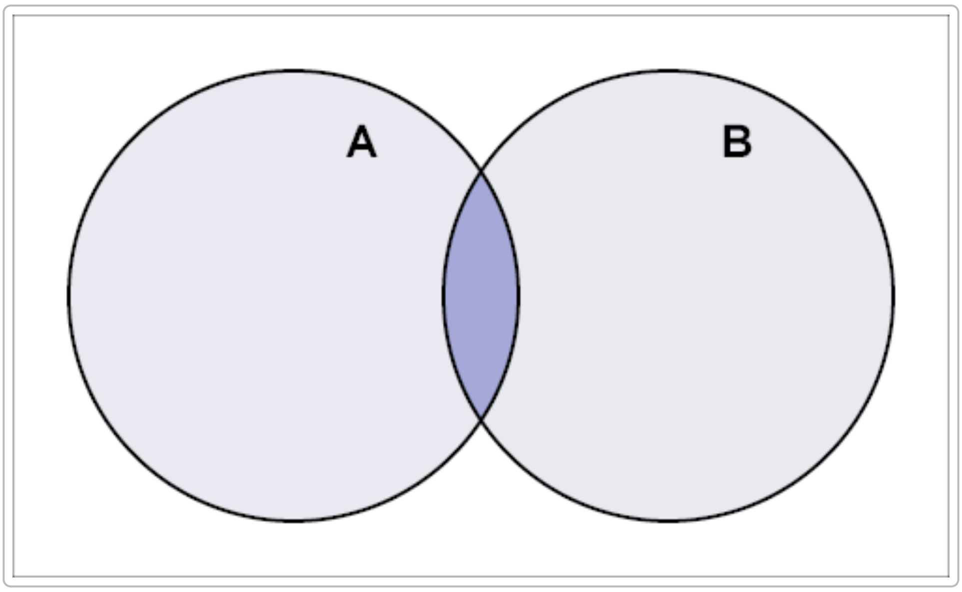

Figura – Venn diagramma dell'intersezione di due insiemi A e B. Nell'intersezione è evidenziata solo l'area comune a A e B.

Insieme delle parti e prodotto cartesiano

Dato un insieme A, si chiama **insieme delle parti** (o *power set*) $P(A)$ l'insieme che contiene *tutti* i sottoinsiemi di A [9]. Se A ha n elementi, allora $P(A)$ ha 2^n elementi [10]. Ad esempio, se $A = \{1, 2\}$, allora $P(A) = \{\varnothing, \{1\}, \{2\}, \{1, 2\}\}$. Il **prodotto cartesiano** di due insiemi A e B, denotato $A \times B$, è l'insieme di tutte le coppie ordinate (a, b) con $a \in A$ e $b \in B$ [11]. Ad esempio, se $A = \{x, y\}$ e $B = \{1, 2\}$, allora $A \times B = \{(x, 1), (x, 2), (y, 1), (y, 2)\}$.

Calcolo differenziale e integrale

Derivata

La **derivata** di una funzione $f(x)$ misura il suo tasso di variazione locale rispetto alla variabile x. Formalmente, la derivata $f'(x)$ è il limite del rapporto incrementale

$$f'(x) = \lim_{h \to 0} \frac{f(x + h) - f(x)}{h} \,.$$

Geometricamente, $f'(x)$ corrisponde alla pendenza della retta tangente al grafico di f nel punto di ascissa x [12] [13]. Fisicamente, rappresenta il tasso di variazione di una grandezza rispetto al tempo o a un'altra variabile (ad esempio, $s'(t)$ è la velocità quando $s(t)$ è la posizione) [12] [14]. Le regole fondamentali di derivazione sono:
- **Somma:** $(f + g)' = f' + g'$.
- **Differenza:** $(f - g)' = f' - g'$.
- **Costante moltiplicativa:** $(kf)' = kf'$.
- **Prodotto:** $(f\,g)' = f'\,g + f\,g'$.

- **Quoziente:** $\left(\frac{f}{g}\right)' = \frac{f'\,g - f\,g'}{g^2}$.
- **Funzione composta (regola di catena):** se $h(x) = f(g(x))$, allora $h'(x) = f'(g(x)) \cdot g'(x)$ [15].

[16] Mostrano in dettaglio queste formule: ad esempio, $(f \cdot g)' = f'\,g + f\,g'$ e $\big(f(g(x))\big)' = f'(g(x))\,g'(x)$ [16] [15]. Questi sono i *formulari principali* per calcolare le derivate di funzioni elementari (potenze, esponenziali, logaritmi, trigonometriche, etc.) usando somme, prodotti, quozienti e composizioni.

Primitiva e integrali

Una **primitiva** di $f(x)$ (o *antiderivata*) è una funzione $F(x)$ tale che $F'(x) = f(x)$ [17]. L'insieme di tutte le primitive di f è $F(x) + C$, dove C è una costante arbitraria [17] [18]. In pratica, l'**integrale indefinito** $\int f(x)\,dx$ rappresenta la famiglia di funzioni $F(x) + C$. Per funzioni elementari esistono regole di integrazione "immediate" (ad esempio $\int k\,dx = kx + C$, $\int x^n dx = x^{n+1}/(n+1) + C$ se $n \neq -1$, $\int e^{ax} dx = \frac{1}{a} e^{ax} + C$, ecc.).

Integrazione per parti: se $u(x)$ e $v(x)$ sono funzioni derivabili, vale

$$\int u'(x)v(x)\,dx = u(x)v(x) - \int u(x)v'(x)\,dx.$$

Nella forma tipica: $\int u\,dv = uv - \int v\,du$ [19]. Questa regola nasce dall'integrazione della derivata di un prodotto e permette di scambiare la derivazione su due fattori per semplificare l'integrale.

Integrazione per sostituzione (cambio di variabile): consiste nel porre $x = \varphi(t)$ per trasformare $\int f(x)\,dx$ in $\int f(\varphi(t))\,\varphi'(t)\,dt$ [20]. Questa tecnica è l'inverso della regola di derivazione di una funzione composta [21] [20]. In pratica, si sceglie $\varphi(t)$ tale che $dx = \varphi'(t)\,dt$ renda l'integrale più semplice. Per integrali definiti $\int_a^b f(x)\,dx$ con $x = \varphi(t)$ si cambia anche i limiti nel dominio di t [20].

Statistica descrittiva

Statistica univariata

Per una serie di dati **univariati**, si definiscono diversi indici di posizione e di dispersione:
- **Media aritmetica:** $\overline{x} = \frac{1}{n} \sum_{i=1}^{n} x_i$ [22], cioè la somma dei valori divisa per il numero di osservazioni.
- **Mediana:** è il valore che divide la distribuzione in due parti di uguale frequenza. In una serie ordinata, se n è dispari la mediana è il $(n+1)/2$-esimo valore; se n è pari, è la media dei due valori centrali [23] [24].
- **Moda:** è la modalità (o valore) più frequente nella serie [25]. Una distribuzione può essere unimodale, bimodale, o multimodale a seconda del numero di valori di moda.
- **Varianza:** misura la dispersione media dei dati attorno alla media. Definita come $\sigma^2 = \frac{1}{N} \sum_{i=1}^{N} (x_i - \overline{x})^2$ (per una popolazione di N osservazioni) [26].
- **Deviazione standard:** è la radice quadrata della varianza, $\sigma = \sqrt{\sigma^2}$ [27], e ha la stessa unità dei dati.
- **Campo di variazione (range):** differenza tra il valore massimo e minimo. Ad esempio, se i dati vanno da 10 a 25, il range è $25 - 10 = 15$ [28].

- **Istogramma:** è un grafico a barre dove l'area di ciascuna barra indica la frequenza (o densità) dei dati in una data classe di valori. In pratica, si suddivide l'intervallo dei dati in classi e si disegna un rettangolo per ogni classe la cui area (non l'altezza!) rappresenta la frequenza [29] .

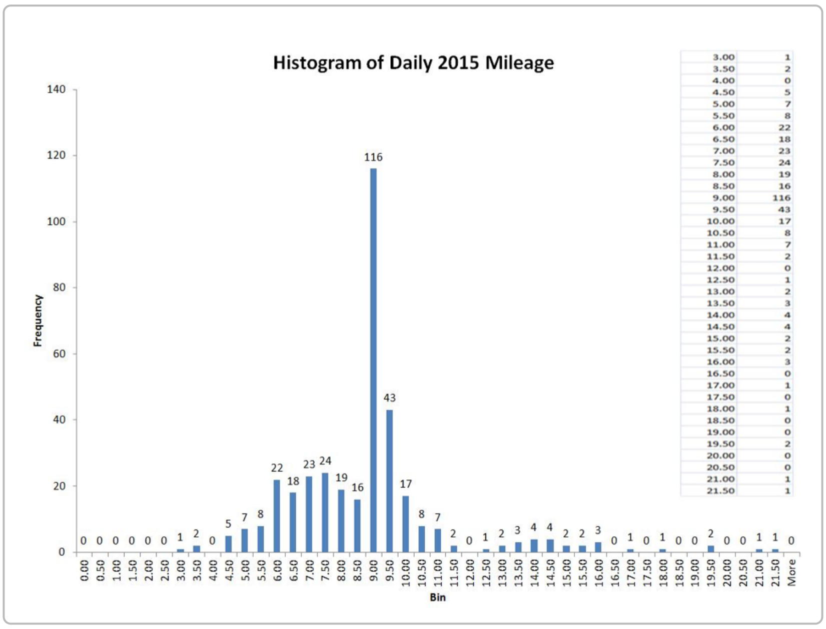

Figura – Esempio di istogramma. L'istogramma sopra mostra la distribuzione delle frequenze su classi di valori. Le aree delle barre (non solo le altezze) rappresentano le frequenze dei dati [29] .

Statistica bivariata

La **distribuzione congiunta** di due variabili (o distribuzione doppia) descrive la frequenza congiunta di coppie (X, Y). Due indici importanti sono:
- **Covarianza:** indica la tendenza di X e Y a variare insieme. Se x_i, y_i sono i dati congiunti di N osservazioni e $\bar{x}, \bar{y}$ le loro medie, la covarianza campionaria è $\mathrm{Cov}(X,Y) = \dfrac{1}{N}\displaystyle\sum_{i=1}^{N}(x_i - \bar{x})(y_i - \bar{y})$ [30] . Un valore positivo indica che X e Y aumentano insieme, negativo indica variazioni opposte, zero indica mancanza di relazione lineare.
- **Coefficiente di correlazione di Pearson:** normalizza la covarianza rendendola adimensionale. Si definisce $\rho_{XY} = \dfrac{\mathrm{Cov}(X,Y)}{\sigma_X\,\sigma_Y}$ [31] , dove σ_X, σ_Y sono le deviazioni standard di X e Y. Il valore ρ è compreso fra -1 e 1: $\rho = 1$ significa correlazione lineare positiva perfetta, $\rho = -1$ correlazione negativa perfetta, $\rho = 0$ nessuna correlazione lineare.

Retta di regressione lineare: è il modello più semplice per predire Y da X con una relazione lineare. Si assume $Y = a + bX$, dove b (pendenza) indica di quanto varia mediamente Y al variare di X, e a (intercetta) è il valore stimato di Y quando $X = 0$. I coefficienti a, b si calcolano col *metodo dei minimi quadrati*, che minimizza la somma degli scarti al quadrato $\sum_i (y_i - a - bx_i)^2$. In pratica si trova

$$b = \frac{\backslash\mathrm{Cov}(X,Y)}{\backslash\mathrm{Var}(X)}, \qquad a = \bar{Y} - b\,\bar{X},$$

dove $\backslash\mathrm{Cov}$ e $\backslash\mathrm{Var}$ sono calcolate sui dati osservati. In termini di correlazione, $b = \rho_{XY}(\sigma_Y/\sigma_X)$. La retta dei minimi quadrati passa il più vicino possibile a tutti i punti in senso quadratico.

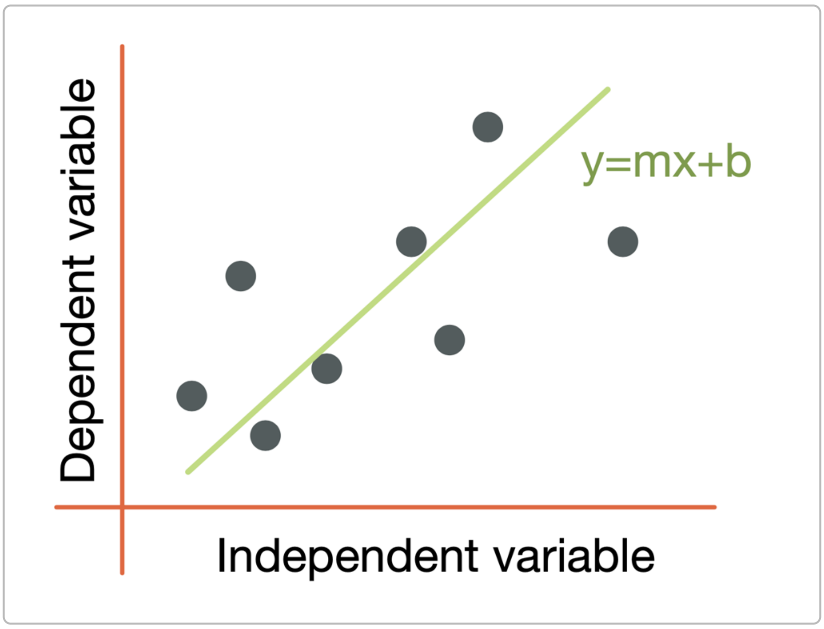

Figura – Scatter plot con retta di regressione lineare. In rosso sono riportati i punti dati (x_i, y_i) e in blu la retta di regressione $Y = a + bX$. Si osserva che la pendenza b è strettamente correlata al coefficiente ρ_{XY} della correlazione [31]: se ρ è alto (vicino a 1), i punti si dispongono intorno a una retta ascendente.

Calcolo delle probabilità

Si considerano **eventi** A,B come insiemi di esiti in uno spazio di probabilità. La probabilità $P(E)$ di un evento E soddisfa $0 \leq P(E) \leq 1$ e $P(\Omega) = 1$. Le principali interpretazioni della probabilità sono:
- **Definizione classica:** se in un esperimento ci sono N esiti equiprobabili e N_E sono favorevoli a E, allora $P(E) = N_E/N$ [32] (es. estrazione casuale da un mazzo di carte).
- **Definizione frequentista:** la probabilità di E è il limite della frequenza relativa con cui E si verifica in un numero molto grande di prove indipendenti [33].

Le **operazioni sugli eventi** riprendono le operazioni insiemistiche:
- **Unione:** $A \cup B$ (evento "A o B o entrambi"). La probabilità segue $P(A \cup B) = P(A) + P(B) - P(A \cap B)$ [7].
- **Intersezione:** $A \cap B$ (evento "A e B assieme").
- **Incompatibilità (disgiunti):** se $A \cap B = \varnothing$, allora $P(A \cap B) = 0$ e $P(A \cup B) = P(A) + P(B)$.
- **Indipendenza:** A e B sono indipendenti se $P(A \cap B) = P(A)\, P(B)$ (equivalente a dire $P(A|B) = P(A)$) [8]. In tal caso gli eventi non si influenzano a vicenda.

La **probabilità condizionata** di A rispetto a B (con $P(B) > 0$) è

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

ovvero la probabilità che A accada sapendo che B è accaduto [34].

Teorema di Bayes: in presenza di informazioni aggiuntive, permette di ribaltare le probabilità condizionate. In forma semplice (per due eventi A, B),

$$P(A \mid B) = \frac{P(B \mid A)\, P(A)}{P(B)}.$$

In generale, se $\{A_i\}$ è una partizione dello spazio e E un evento, allora

$$P(A_i \mid E) = \frac{P(E \mid A_i)\, P(A_i)}{\sum_j P(E \mid A_j)\, P(A_j)},$$

come afferma il teorema di Bayes [35]. Questo permette di aggiornare le probabilità *a priori* $P(A_i)$ ("cause") alla luce di un evento osservato E.

Variabili aleatorie discrete: sono funzioni che associano un valore numerico a ciascun esito di un esperimento su uno spazio discreto. Si descrivono tramite la **funzione di massa di probabilità** $P(X = x_i)$. Il valore atteso (o media) di X è $\mu = \sum_i x_i\, P(X = x_i)$ e la varianza è $\Var(X) = \sum_i (x_i - \mu)^2 P(X = x_i)$ (formule analoghe alla statistica descrittiva).

Esempi importanti:
- **Legge di Bernoulli $B(p)$:** $P(X = 1) = p$ (successo) e $P(X = 0) = 1 - p$ (insuccesso) [36]. Questa variabile rappresenta un esperimento con due esiti. Ha valore atteso $E[X] = p$ e varianza $\Var(X) = p(1 - p)$ [37].
- **Legge binomiale $B(n, p)$:** descrive il numero di successi in n prove indipendenti di Bernoulli(p). La probabilità che $X = k$ successi avvengano in n prove è $P(X = k) = \binom{n}{k} p^k (1 - p)^{n-k}$. Questa variabile ha media $E[X] = np$ e varianza $\Var(X) = np(1 - p)$ [38].

In generale, le variabili aleatorie discrete notevoli includono anche la geometrica, la Poisson, l'ipergeometrica, ecc., ciascuna con i propri indici caratteristici (media e varianza) dipendenti dai parametri.

Fonti: Concetti e definizioni tratti da testi di calcolo e statistica [1] [9] [12] [17] [29] [30] [34] [35] [36] [38], e immagini illustrative da Wikimedia Commons.

---

[1] [2] [10] Gli insiemi in matematica: definizione e rappresentazione
https://sapere.virgilio.it/scuola/medie/matematica-algebra/la-teoria-degli-insiemi/introduzione-concetto

[3] La differenza tra due insiemi - Andrea Minini
https://www.andreaminini.org/matematica/teoria-degli-insiemi/differenza-tra-insiemi

[4] L'insieme complemento o complementare (teoria degli insiemi) - Andrea Minini
https://www.andreaminini.org/matematica/teoria-degli-insiemi/insieme-complemento

[5] Unione (insiemistica) - Wikipedia
https://it.wikipedia.org/wiki/Unione_(insiemistica)

[6] Intersezione (insiemistica) - Wikipedia
https://it.wikipedia.org/wiki/Intersezione_(insiemistica)

[7] Teorema della probabilità totale - Wikipedia
https://it.wikipedia.org/wiki/Teorema_della_probabilit%C3%A0_totale

8 34 Probabilità condizionata - Wikipedia
https://it.wikipedia.org/wiki/Probabilit%C3%A0_condizionata

9 Insieme delle parti di un insieme - Matematicamente
https://www.matematicamente.it/appunti/insiemi/insieme-delle-parti-di-un-insieme/

11 Prodotto cartesiano - Wikipedia
https://it.wikipedia.org/wiki/Prodotto_cartesiano

12 13 14 Derivata - Wikipedia
https://it.wikipedia.org/wiki/Derivata

15 La derivata della funzione composta - Andrea Minini
https://www.andreaminini.org/matematica/calcolo-differenziale/la-derivata-della-funzione-composta

16 La derivata della somma, differenza, prodotto e quoziente di funzioni - Andrea Minini
https://www.andreaminini.org/matematica/calcolo-differenziale/la-derivata-della-somma-differenza-prodotto-e-quoziente-di-funzioni

17 18 Primitiva (matematica) - Wikipedia
https://it.wikipedia.org/wiki/Primitiva_(matematica)

19 Integrazione per parti - Wikipedia
https://it.wikipedia.org/wiki/Integrazione_per_parti

20 21 Integrazione per sostituzione - Wikipedia
https://it.wikipedia.org/wiki/Integrazione_per_sostituzione

22 Media (statistica) - Wikipedia
https://it.wikipedia.org/wiki/Media_(statistica)

23 24 Mediana (statistica) - Wikipedia
https://it.wikipedia.org/wiki/Mediana_(statistica)

25 Moda (statistica) - Wikipedia
https://it.wikipedia.org/wiki/Moda_(statistica)

26 27 Scarto quadratico medio - Wikipedia
https://it.wikipedia.org/wiki/Scarto_quadratico_medio

28 Campo di variazione - Wikipedia
https://it.wikipedia.org/wiki/Campo_di_variazione

29 Category:Histograms - Wikimedia Commons
https://commons.wikimedia.org/wiki/Category:Histograms

30 31 Covarianza (probabilità) - Wikipedia
https://it.wikipedia.org/wiki/Covarianza_(probabilit%C3%A0)

32 33 Tipologie di probabilità - Wikiversità
https://it.wikiversity.org/wiki/Tipologie_di_probabilit%C3%A0

35 Teorema di Bayes - Wikipedia
https://it.wikipedia.org/wiki/Teorema_di_Bayes

36 37 Distribuzione di Bernoulli - Wikipedia
https://it.wikipedia.org/wiki/Distribuzione_di_Bernoulli

38 Distribuzione binomiale - Wikipedia
https://it.wikipedia.org/wiki/Distribuzione_binomiale