Definizione di variabile casuale

In alcuni settori della scienza e della tecnica intervengono delle grandezze variabili i cui valori numerici dipendono dalla probabilità che si verifichino certi eventi casuali. Tali grandezze prendono il nome di variabili casuali e possono essere di due tipi: discrete e continue.

Una variabile casuale si dice discreta se può assumere un numero finito o un'infinità numerabile di valori, si dice continua se può assumere tutti i valori appartenenti ad un intervallo o all'unione di più intervalli.

Si dice variabile casuale discreta o variabile casuale aleatoria una variabile **X** che può assumere i valori $x_1, x_2, x_3, \cdots\cdots, x_n$ al verificarsi di un sistema completo di eventi casuali $E_1, E_2, E_3, \cdots\cdots, E_n$ le cui probabilità sono rispettivamente uguali a $p_1, p_2, p_3, \cdots\cdots, p_n$. Essendo gli eventi $E_1, E_2, E_3, \cdots\cdots, E_n$ incompatibili e complementari vuol dire che il verificarsi di uno di essi esclude il verificarsi contemporaneo di qualsiasi altro evento fermo restando che uno di essi deve necessariamente verificarsi. Ne segue che l'evento $E_1 \cup E_2 \cup E_3 \cup \cdots\cdots \cup E_n = \Omega$ è l'evento certo per cui possiamo scrivere:

$$p\left(E_1 \cup E_2 \cup E_3 \cup \cdots\cdots \cup E_n\right) = p\left(E_1\right) + p\left(E_2\right) + p\left(E_3\right) + \cdots\cdots + p\left(E_n\right) = p_1 + p_2 + p_3 + \cdots\cdots + p_n = 1$$

Di solito, le variabili casuali sono indicate con le lettere maiuscole **X**, **Y**, dell'alfabeto latino, mentre i valori che esse assumono [1] sono indicati con le corrispondenti lettere minuscole dell'alfabeto latino x_i , y_i .

[1] Sono le **modalità** o i valori della variabile casuale

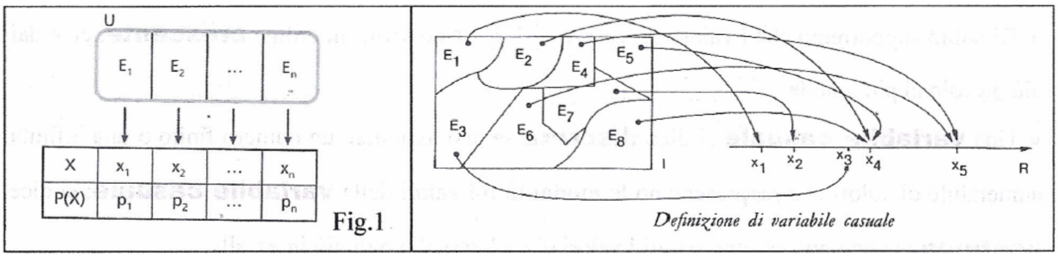


Fig.1 *Definizione di variabile casuale*

Nella figura 1 ad ogni evento E_i corrisponde una modalità (valore) x_i della **variabile casuale X** e la probabilità p_i di verificarsi. Gli insiemi (eventi aleatori) $E_1, E_2, \cdots, E_n$ sono **disgiunti** e la loro unione è l'insieme universo U, quindi tali insiemi costituiscono una **partizione** di U.

Dallo schema della seconda tabella si possono dedurre le due seguenti tabelle:

Eventi	Probabilità
E_1	p_1
E_2	p_2
E_3	p_3
E_4	p_4
E_5	p_5
E_6	p_6
E_7	p_7
E_8	p_8
Totale	1

$\longrightarrow$

Valori della v.c. X	Probabilità
x_1	p_1
x_2	p_7
x_3	p_3
x_4	$p_2 + p_5 + p_8$
x_5	$p_4 + p_6$
Totale	1

Dal punto di vista formale per l'evento E_1 scriveremo $P(E_1) = P(X = x_1) = p_1$ in quanto l'affermazione che "si verifica l'evento E_1 con probabilità $P(E_1)$" è equivalente all'affermazione che "la variabile casuale X assume il valore x_1 con probabilità $P(E_1)$"

Sostanzialmente una variabile casuale discreta **X** è un insieme discreto di coppie (x_i, p_i), dove p_i rappresenta la probabilità che la variabile casuale discreta **X** assuma la modalità x_i.

Spesso una variabile casuale **X** è rappresentata mediante una tabella del tipo:

$$X = \begin{bmatrix} x_1 & x_2 & \cdots & x_n \\ p_1 & p_2 & \cdots & p_n \end{bmatrix} \quad \text{o del tipo} \quad X = \begin{cases} x_1 & x_2 & \cdots & x_i & \cdots & x_n \\ p_1 & p_2 & \cdots & p_i & \cdots & p_n \end{cases}$$

o anche, in maniera ancora più sintetica : $X = [x_i, p(x_i)]$ con $\sum_{i=1}^{n} p(x_i) = 1$ e $p(x_i)$ rappresenta la probabilità che la variabile casuale discreta **X** assuma il valore x_i (ovvero la modalità x_i).

La prima riga rappresenta l'insieme di tutte le modalità che la variabile casuale discreta **X** può assumere; la seconda riga rappresenta la sua funzione di probabilità o la sua legge di distribuzione di probabilità o, più semplicemente , la sua distribuzione di probabilità.

Precisazioni

- Una variabile casuale discreta **X** è definita quando conosciamo l'insieme dei valori x_i che essa assume, nonché l'insieme delle probabilità p_i corrispondenti
- i valori x_i che la variabile casuale discreta **X** assume prendono il nome di modalità della variabile casuale discreta **X**
- l'insieme delle probabilità p_i costituisce la distribuzione di probabilità della variabile casuale discreta **X**, detta anche funzione di probabilità o funzione di massa o funzione di densità e spesso viene indicata col simbolo $f(x)$, cioè: $\mathbf{f(x)=P(X=x)}$
- Di solito supporremo che i valori $x_1, x_2, x_3, \cdots\cdots, x_n$ siano scritti in ordine crescente, cioè dal più piccolo al più grande
- Una variabile casuale **X** si dice discreta se può assumere un numero finito o una infinità numerabile di valori che rappresentano le modalità della variabile casuale; si dice continua quando può assumere tutti i valori di un intervallo o di più intervalli

Esempi di variabili casuali discrete

Si lanciano in aria tre monete e si considera la variabile casuale $\mathbf{X = numero\ di\ croci}$. Costruire la tabella della variabile casuale

Lo spazio degli eventi elementari è il seguente:

$$\Omega = \{TTT, TTC, TCT, CTT, TCC, CTC, CCT, CCC\}$$ [2]

in quanto gli eventi elementari sono i seguenti:

$$e_1 = T \cap T \cap T = TTT \ , \ e_2 = T \cap T \cap C = TTC \ , \ e_3 = T \cap C \cap T = TCT$$

$$e_4 = C \cap T \cap T = CTT \ , \ e_5 = T \cap C \cap C = TCC \ , \ e_6 = C \cap T \cap C = CTC$$

$$e_7 = C \cap C \cap T = CCT \ , \ e_8 = C \cap C \cap C = CCC$$

La probabilità di ogni evento elementare è $\frac{1}{8}$, cioè:

$$p(e_1) = p(e_2) = p(e_3) = p(e_4) = p(e_5) = p(e_6) = p(e_7) = p(e_8) = \frac{1}{8}$$

La variabile casuale **X** può assumere i valori **0, 1, 2, 3** (cioè possiamo avere 0 croci, una croce, due croci, tre croci) con le probabilità di seguito indicate:

$$X = \begin{bmatrix} x_1 & x_2 & x_3 & x_4 \\ 0 & 1 & 2 & 3 \\ \frac{1}{8} & \frac{3}{8} & \frac{3}{8} & \frac{1}{8} \\ p_1 & p_2 & p_3 & p_4 \end{bmatrix}$$

0, 1, 2, 3 rappresentano le modalità della variabile casuale **X.**

Se vogliamo calcolare la probabilità che la variabile casuale **X** assuma il valore 2, cioè se vogliamo calcolare la probabilità che in un lancio si abbiano due croci dovremo procedere come segue:

$$p(X=2) = p(e_5 \cup e_6 \cup e_7) = p(e_5) + p(e_6) + p(e_7) = \frac{1}{8} + \frac{1}{8} + \frac{1}{8} = \frac{3}{8}$$

[2] Nel lancio di tre monete diverse si hanno 8 diversi casi possibili , corrispondenti alle disposizioni con ripetizione di 2 elementi di classe 3 :

Una tabella più completa della precedente può essere la seguente:

Modalità x_i **della v.c X**	**Eventi corrispondenti** E_i	**Probabilità** p_i **di ciascun evento**
$x_1 = 0$	$E_1 = e_1$	$p_1 = \frac{1}{8}$
$x_2 = 1$	$E_2 = e_5 \cup e_6 \cup e_7$	$p_2 = \frac{3}{8}$
$x_3 = 2$	$E_3 = e_2 \cup e_3 \cup e_4$	$p_3 = \frac{3}{8}$
$x_4 = 3$	$E_4 = e_8$	$p_4 = \frac{1}{8}$
$e_1 = TTT$, $e_2 = TTC$, $e_3 = TCT$ $e_4 = CTT$, $e_5 = TCC$, $e_6 = CTC$ $e_7 = CCT$, $e_7 = CCC$		

Calcolare la probabilità che nel lancio delle tre monete si ottenga più di una croce.

$$p(X>1) = p[(X=2)\cup(X=3)] = p(X=2)+p(X=3) = \frac{3}{8}+\frac{1}{8} = \frac{1}{2}$$

Ovviamente la somma di tutte le probabilità è uguale ad 1 , come prevede la teoria .

Caratteristiche numeriche delle variabili casuali discrete

La legge di distribuzione di probabilità di una variabile casuale X descrive completamente la variabile casuale X.Nella pratica, spesso capita di essere interessati alla conoscenza di alcuni parametri numerici che mettono in evidenza alcune proprietà delle variabili casuali, quali la tendenza centrale, la variabilità, l’interdipendenza. Spesso è sufficiente conoscere soltanto alcuni parametri numerici che caratterizzano i tratti essenziali della variabile casuale in esame. Per esempio, un valore medio attorno al quale si raggruppano i valori possibili della variabile casuale; un numero qualsiasi che caratterizza la dispersione di questi valori attorno al valore medio. Si tratta di valori caratteristici o sintetici che forniscono un'immagine riassuntiva sufficiente per gli scopi che ci prefiggiamo.

Le caratteristiche numeriche utilizzate nel calcolo della probabilità sono numerose, ma noi ci limiteremo a considerarne alcune fra le più importanti quali la speranza matematica, la varianza, i momenti, lo scarto quadratico medio (o deviazione standard). Come esistono un valore medio, una varianza e dei momenti per una variabile statistica così esistono un valore medio, una varianza e dei momenti per una variabile casuale **X**.

Definizione di speranza matematica

Una delle caratteristiche numeriche di una variabile casuale **X** è la speranza matematica detta anche valore medio [3]

Sia data la variabile casuale **X** le cui modalità $x_1, x_2, x_3, \cdots\cdots, x_n$ abbiano, rispettivamente, probabilità $p_1, p_2, p_3, \cdots\cdots, p_n$. Si chiama speranza matematica della variabile casuale **X** la somma dei prodotti delle modalità di questa variabile per le rispettive probabilità.

$$\mathbf{E(X)=M(X)=\mu(X)=x_1 \times p_1 + x_2 \times p_2 + x_3 \times p_3 + \cdots + x_n \times p_n = \sum_{i=1}^{n} x_i\, p_i}$$

Ricordando che $p_1 + p_2 + p_3 + \cdots\cdots + p_n = 1$ possiamo scrivere:

$$\mathbf{E(X)=M(X)=\frac{x_1 \cdot p_1 + x_2 \cdot p_2 + x_3 \cdot p_3 + \cdots + x_n \times p_n}{p_1 + p_2 + p_3 + \cdots + p_n} = \frac{\sum_{i=1}^{n} x_i\, p_i}{\sum_{i=1}^{n} p_i = 1} = \sum_{i=1}^{n} x_i\, p_i}$$

In tal modo si osserva che la speranza matematica di una variabile casuale **X** rappresenta la media aritmetica ponderata delle modalità $x_1, x_2, x_3, \cdots\cdots, x_n$ della variabile casuale **X** per i pesi $p_1, p_2, p_3, \cdots\cdots, p_n$.

---

[3] O **valore atteso** o **valore sperato** o **misura di tendenza centrale** o **media aritmetica**

La speranza matematica di una **variabile casuale X** ci fornisce un valore singolo che rappresenta tutti i valori della X e per mesto motivo è chiamato anche una **misura di tendenza centrale**.

Varianza

La **varianza** è un particolare **parametro numerico** di **dispersione** della **variabile casuale X**.

Essa è definita in maniera del tutto analoga alla varianza di una variabile statistica.

Definizione di varianza: Si chiama **varianza** della **variabile casuale X** la speranza matematica del quadrato dello scarto $(x - \mu)$ tra i valori x della **variabile casuale X** ed il suo valore medio μ

In simboli abbiamo: $\mathbf{Var(X) = \sigma^2 = V(X) = E\left[(X-\mu)^2\right] = \sum_{i=1}^{n} (x_i - \mu)^2 \cdot p(x_i)}$

Formula pratica per il calcolo della varianza

La **varianza**, che per definizione è il valore medio dello scarto al quadrato, può essere calcolata in base alla seguente regola pratica:

$$\mathbf{Var(X) = E(X^2) - [E(X)]^2 = \sum_{i=1}^{n} x_i^2 \cdot p(x_i) - \left[\sum_{i=1}^{n} x_i \cdot p(x_i)\right]^2}$$

cioè la varianza di una **variabile casuale X** è uguale alla differenza fra la media dei quadrati della **variabile casuale X** ed il quadrato del valore medio della **variabile casuale X**.

$\mathbf{E(X^2) = M(X^2)}$ = **speranza matematica** o **media aritmetica** o **valore medio** della **variabile casuale X** $Y = X^2$

$\mathbf{E(X) = M(X)}$ = **speranza matematica** della **variabile casuale X**

La varianza gode delle seguenti proprietà :

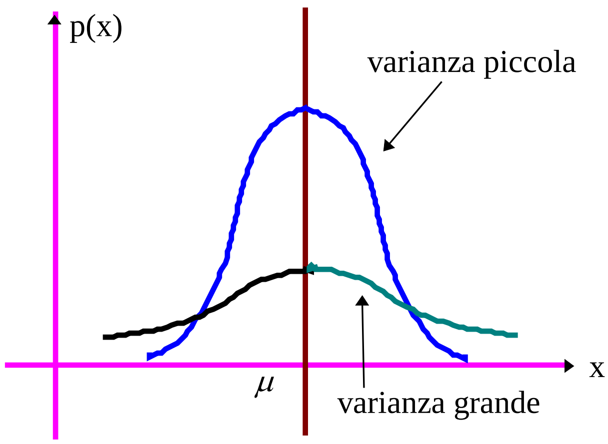
- $Y = aX \Rightarrow Var(Y) = a^2 \cdot Var(X)$ cioè $\mathbf{Var(a \cdot X) = a^2 \cdot Var(X)}$
- $Y = X + b \Rightarrow Var(Y) = Var(X)$ cioè : $\mathbf{Var(X + b) = Var(X)}$
- $Y = aX + b \Rightarrow Var(Y) = a^2 \cdot Var(X)$ cioè $\mathbf{Var(a \cdot X + b) = a^2 \cdot Var(X)}$

la varianza σ^2 (o la deviazione standard σ) è una misura della dispersione dei valori della variabile casuale **X** attorno al valore medio μ.

Se i valori di probabilità delle modalità della variabile casuale **X** sono concentrati vicino alla media μ , la varianza è piccola , mentre la varianza è grande quando i valori sono dispersi lontani dalla media .

Per questo motivo, quanto minore è la varianza (che è un valore costante) di una variabile casuale **X**, tanto più i valori di probabilità $p_1, p_2, p_3, \cdots\cdots, p_n$ sono concentrati intorno al valore medio μ, come mostrato schematicamente in figura.

Nel caso particolare $Var(X) = 0$ non c'è dispersione dei valori della variabile casuale **X** intorno al valore medio . In questo caso la variabile casuale **X** è una costante che coincide col valore medio: **X = μ**



Scarto quadratico medio o deviazione standard

Se si è scelto il valore medio μ come indicatore di localizzazione è normale scegliere la varianza come indicatore della dispersione. Tuttavia la varianza è una quantità di secondo grado in **x** per cui spesso si preferisce usale la deviazione standard (o lo scarto quadratico medio) definita come la radice quadrata aritmetica della varianza . In simboli abbiamo: $\sigma(X)=\sqrt{Var(X)}=\sqrt{\sum_{i=1}^{n}(x_i-\mu)^2\cdot p(x_i)}$

Se X è una variabile casuale con media μ e varianza σ^2 allora la trasformata:

$$Y=\frac{X-\mu}{\sigma} \qquad [\ X=\sigma Y+\mu\ ]$$

è detta variabile casuale standardizzata della variabile casuale X, cioè una variabile casuale standardizzata è una nuova variabile casuale che gode di particolari proprietà.

La variabile casuale standardizzata Y gode delle seguenti proprietà:

a) è una grandezza adimensionata b) $E(Y) = 0$ c) $Var(Y) = 1$ cioè la sua varianza vale **1**

La varianza è un ottimo criterio per stabilire la dispersione in una distribuzione di probabilità che abbia valore medio μ finito ed univocamente definito. Tuttavia non sempre è possibile usare il valore medio come parametro di localizzazione. In questi casi si può ricorrere al altri parametri di dispersione.

Distribuzione di probabilità

Data una variabile casuale X, con valori $x_1, x_2, x_3, \cdots\cdots, x_n$, la successione delle probabilità $p_1, p_2, p_3, \cdots\cdots, p_n$ ad essi associata si chiama distribuzione di probabilità della variabile casuale X. Essa può essere rappresentata con una tabella come la seguente dove sono indicate le modalità della X e le corrispondenti probabilità.

X	$x_1, x_2, x_3, \cdots\cdots, x_n$
p	$p_1, p_2, p_3, \cdots\cdots, p_n$

La legge di distribuzione di probabilità è detta anche funzione di probabilità o funzione di massa o funzione di densità della variabile casuale **X** e spesso viene indicata col simbolo $f(x)$, cioè: $\mathbf{f(x)=P(X=x)}$.

Funzione di ripartizione

La variabile casuale **X** può essere descritta oltre che dalla funzione di massa $f(x)$ anche da un'altra funzione detta funzione di ripartizione (o funzione di probabilità cumulata o funzione di distribuzione cumulativa) indicata col simbolo $\mathbf{F(x)}$.

Sia **X** una variabile casuale che può assumere le modalità $x_1, x_2, x_3, \cdots\cdots, x_n$ disposte in ordine crescente e con distribuzione di probabilità $p_1, p_2, p_3, \cdots\cdots, p_n$.

Definizione: Si chiama funzione di ripartizione della variabile casuale **X**, una funzione $\mathbf{F(x)}$ della variabile reale $\mathbf{x}$ che è uguale, per ogni valore della $\mathbf{x}$, alla probabilità che **X** assuma un valore minore o uguale ad $\mathbf{x}$, cioè:

$$\mathbf{F(x_k)=P(X\le x_k)=\sum_{i=1}^{k} p(x_i)=p(x_1)+p(x_2)+p(x_3)+L,+p(x_k)} \quad \text{dove } \mathbf{p(x_i)=P(X=x_i)}$$

Dalla definizione data segue che:

- $x < x_1 \Rightarrow F(x) = 0$
- $x_1 \le x < x_2 \Rightarrow F(x) = p_1 = P(x \le x_1)$
- $x_2 \le x < x_3 \Rightarrow F(x) = p_1 + p_2 = P(x \le x_2)$
- $x_3 \le x < x_4 \Rightarrow F(x) = p_1 + p_2 + p_3 = P(x \le x_3)$
- $x_k \le x < x_{k+1} \Rightarrow F(x) = p_1 + p_2 + p_3 + \cdots + p_k = P(x \le x_k)$
- $x \ge x_n \Rightarrow F(x) = p_1 + p_2 + p_3 + \cdots\cdots + p_n = 1$

$\mathbf{X}$	x_1	x_2	x_3	$\cdots$	x_n
$\mathbf{P(X)}$	p_1	p_2	p_3	$\cdots$	p_n
$\mathbf{F(X)}$	p_1	p_1+p_2	$p_1+p_2+p_3$	$\cdots$	$p_1+p_2+p_3+\cdots+p_n$

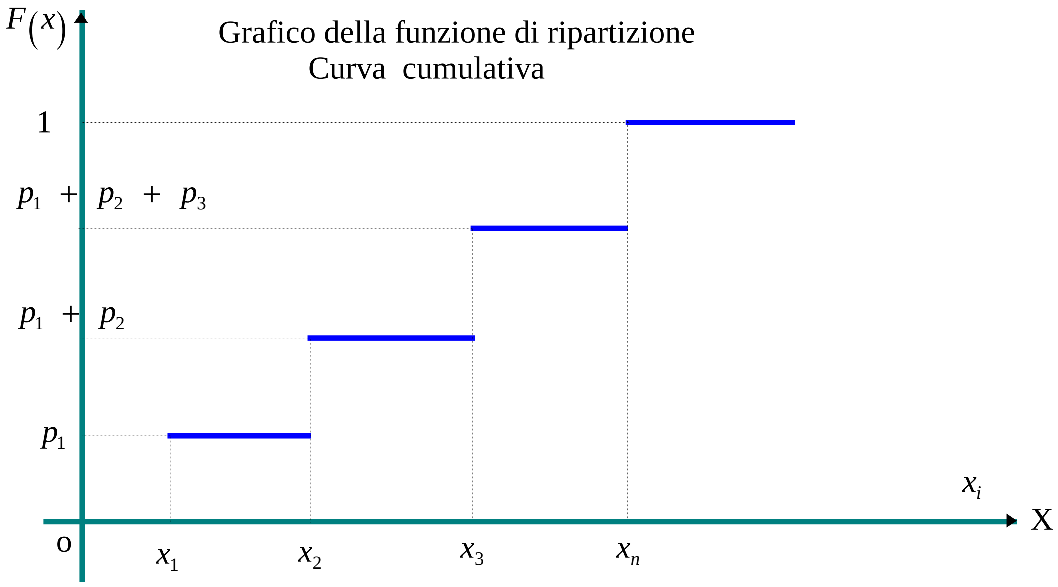
Teorema: Se $F(x)$ è la **funzione di ripartizione** della **variabile casuale X**, allora: $x_i \leq x < x_{i+1} \quad \Rightarrow \quad F(x) = p_1 + p_2 + p_3 + \cdots + p_i = P(X \leq x_i)$

Per la **variabile casuale discreta X**, avente la seguente funzione di probabilità $X = \begin{cases} x_1 & x_2 & \cdots & x_i & \cdots & x_n \\ p_1 & p_2 & \cdots & p_i & \cdots & p_n \end{cases}$

la **funzione di ripartizione** è la seguente:

$$F(x) = \begin{cases} 0 & se & x < x_1 \\ p_1 & se & x_1 \leq x < x_2 \\ p_1 + p_2 & se & x_2 \leq x < x_3 \\ p_1 + p_2 + p_3 & se & x_3 \leq x < x_4 \\ \cdots\cdots\cdots\cdots & \cdots & \cdots\cdots\cdots\cdots \\ p_1 + p_2 + \cdots + p_i & se & x_i \leq x < x_{i+1} \\ \cdots\cdots\cdots\cdots & \cdots & \cdots\cdots\cdots\cdots \\ 1 & se & x \geq x_n \end{cases}$$

Il grafico della funzione di ripartizione è quello riportato in figura e prende il nome di **curva cumulativa**.



A causa del suo grafico, la **funzione di ripartizione** è detta **funzione a gradini**.

- E' importante rilevare che la conoscenza della funzione di ripartizione di una variabile casuale X ci consente di ricavare, mediante semplici differenze, la probabilità dei singoli eventi elementari, cioè:

$$p(X=x_k)=p(x_{k-1}<X\leq x_k)=F(x_k)-F(x_{k-1})$$

Mediante la funzione di ripartizione possiamo calcolare la probabilità che la variabile casuale X assuma un valore x appartenente ad un intervallo di estremi a,b:

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

La funzione di densità (o di massa) di una variabile casuale X concettualmente rappresenta la probabilità che la variabile casuale X assuma un ben determinato valore.

La funzione di ripartizione di una variabile casuale X rappresenta la probabilità che la variabile casuale X assuma un valore ≤ ad x:

$F(x)=p(X\leq x)$ e quindi la probabilità che la variabile casuale X assuma un valore maggiore di X è: $1-F(x)$

Il problema delle prove ripetute e la variabile casuale con distribuzione binomiale

Definiamo variabile casuale binomiale una variabile aleatoria con due soli esiti possibili. Un problema di grande importanza pratica è quello delle prove ripetute tutte nelle stesse condizioni. Si tratta di un problema che porta alla costruzione della variabile casuale con distribuzione binomiale.

Si dice esperimento di *B*ernoulli una sequenza di n prove con le seguenti caratteristiche:

1) ogni prova è un esperimento che può avere soltanto due esiti possibili, uno lo chiameremo successo (si verifica l'evento **E**) ed ha probabilità **p**, l'altro lo chiameremo insuccesso (si verifica l'evento contrario $\overline{\mathbf{E}}$ cioè non si verifica l'evento **E**) ed ha probabilità **q = 1 - p**

2) il risultato di ciascuna prova è indipendente dai risultati delle prove precedenti. [4]

3) la probabilità **p** di successo e quindi la probabilità **q = 1 - p** di insuccesso sono costanti in ciascuna prova.

Il problema delle prove ripetute si presenta quando consideriamo un evento **E** che ha probabilità **p** costante di verificarsi in qualsiasi prova effettuata e vogliamo calcolare la probabilità che ha tale evento (detto anche successo) di verificarsi **h** su **n** prove effettuate, con $0 \le h \le n$.

Consideriamo un evento **E** ripetibile quanto si vuole e supponiamo di fare con esso **n** prove, tutte nelle stesse condizioni, per cui la probabilità che l'evento **E** si realizzi sia uguale in ogni prova effettuata. Indichiamo con **p** la probabilità che in ciascuna prova l'evento **E** si verifichi e con **q = 1 - p** la probabilità che l'evento **E** non si verifichi (insuccesso) e quindi si verifichi l'evento contrario $\overline{\mathbf{E}}$.

Vogliamo determinare la probabilità $p_{n,h}(E)$ che nelle **n** prove effettuate l'evento **E** si verifichi **h** volte.

Possiamo distinguere due casi:

a) **è fissato l'ordine delle prove in cui l'evento E deve verificarsi**

b) **non è fissato l'ordine delle prove in cui l'evento E deve verificarsi**

---

[4] Questo significa che se si tratta dell'estrazione di una pallina da un'urna, dopo l'estrazione la pallina va rimessa nell'urna, altrimenti la probabilità dell'evento (che abbiamo chiamato successo) non sarebbe costante.

a) In questo caso le n prove sono eventi semplici, fra loro compatibili ed indipendenti in quanto la probabilità p è costante.

Pertanto, fissato l'ordine delle **h** prove in cui l'evento **E** deve realizzarsi, ed indicando con **q** la probabilità dell'evento contrario, per il teorema della probabilità composta abbiamo:

$$\mathbf{p_{n,h}(E)} = \overbrace{\mathbf{p\cdot p\cdot p\cdots p}}^{\text{h volte}}\cdot\overbrace{\mathbf{q\cdot q\cdot q\cdots q}}^{\text{n - h volte}}=\mathbf{p^h\cdot q^{n-h}}$$

Un dado viene lanciato 10 volte. Calcolare la probabilità che ha la faccia contraddistinta dal numero 3, di presentarsi nel primo, nel secondo, nel sesto, nel settimo lancio.

$p=\dfrac{1}{6}$ = probabilità che ha la faccia contraddistinta dal numero **3** di presentarsi in un lancio

$q=\dfrac{5}{6}$ = probabilità che ha la faccia contraddistinta dal numero **3** di non presentarsi in un lancio

L'evento considerato $\mathbf{E_{1,2,6,7}}$ si verifica quando si ottiene il punto **3** nel primo (evento E_1) e nel secondo (evento E_2), si ottiene un punto diverso dal **3** nel terzo (evento E_3), nel quarto (evento E_4), nel quinto (evento E_5), si ottiene il punto **3** nel sesto (evento E_6) e nel settimo (evento E_7) lancio, si ottiene un punto diverso dal **3** nell'ottavo (evento E_8), nel nono (evento E_9) e decimo (evento E_{10}) lancio.

$$p(E)=p(E_1)\cdot p(E_2)\cdots p(E_{10})=\frac{1}{6}\cdot\frac{1}{6}\cdot\frac{5}{6}\cdot\frac{5}{6}\cdot\frac{5}{6}\cdot\frac{1}{6}\cdot\frac{1}{6}\cdot\frac{5}{6}\cdot\frac{5}{6}\cdot\frac{5}{6}=\left(\frac{1}{6}\right)^4\cdot\left(\frac{5}{6}\right)^6=0,026=p^4q^6$$

$$n=10 \qquad h=4 \qquad n-h=10-4=6$$

$\mathbf{E_{1,2,6,7}}$ = probabilità che il numero **3** si presenti nel primo, secondo, sesto, settimo lancio.

E' evidente che i seguenti eventi elementari hanno tutti la stessa probabilità, cioè:

$$p\left(\mathbf{E_{1,2,6,7}}\right)=p\left(\mathbf{E}_{2,5,8,10}\right)=p\left(\mathbf{E}_{7,9,5,3}\right)=\cdots=p^4q^6=0,026$$

Cosa succede quando non è fissato l'ordine delle prove? Il fatto che l'evento **E** si verifichi **h** implica che l'evento contrario $\overline{\mathbf{E}}$ si presenta **n-h** volte.

Valgono le seguenti considerazioni:

- l'evento "**E** si verifica **h** volte su **n** prove" può presentarsi con $\binom{\mathbf{n}}{\mathbf{h}}$ modalità (configurazioni)
- la probabilità di ogni singola modalità, per il teorema della probabilità composta, è:

$$\overbrace{p\cdot p\cdot p\cdots p}^{\text{h volte}}\cdot\overbrace{q\cdot q\cdot q\cdots q}^{\text{n - h volte}}=p^h\cdot q^{n-h}$$

- essendo le $\binom{\mathbf{n}}{\mathbf{h}}$ modalità (configurazioni) diverse fra loro e a due a due incompatibili, per il teorema della probabilità totale, la probabilità dell'evento "**E** si verifica **h** volte su **n** prove", è:

$$p_{n,h}(E)=\underbrace{p^h\cdot q^{n-h}+p^h\cdot q^{n-h}+\cdots+p^h\cdot q^{n-h}}_{\binom{n}{h}}=\binom{n}{h}\cdot p^h\cdot q^{n-h}$$

Se, come abbiamo detto in precedenza, non è fissato l'ordine delle prove, l'evento complesso può presentarsi in tanti modi quante sono le combinazioni di **n** elementi (prove effettuate) di classe **h** (numero delle prove favorevoli o successi). Tali modi di presentarsi sono eventi complessi tra loro incompatibili in quanto uno di essi esclude il verificarsi degli altri. La probabilità di ciascuno di tali eventi è data dalla relazione precedente: $\mathbf{p_{n,h}(E)=p^h\cdot q^{n-h}}$ in quanto ciascuno di essi differisce per l'ordine in cui le **h** prove favorevoli (**h** successi) all'evento elementare **E** considerato si sono presentate rispetto alle **n** prove effettuate.

La probabilità dell'evento richiesto sarà la somma delle singole probabilità, tante volte quante sono le combinazioni semplici delle **n** prove effettuate rispetto alle **h** prove favorevoli. In simboli abbiamo:

$$p_{n,h}(E) = \binom{n}{h} \cdot p^h \cdot q^{n-h} = \frac{n!}{h!(n-h)!} \cdot p^h \cdot q^{n-h} = p(E,n,h) = p_{E,n,h}$$

$p_{n,h}(E)$ = probabilità che l'evento **E** si verifichi **h** volte in **n** prove = probabilità che in **n** prove si abbiano **h** successi ed **n-h** insuccessi.

Chiariamo meglio quanto detto con un esempio particolare. "Si lancia una moneta 5 (n) volte. Si chiede di calcolare la probabilità che esca **testa** 2 (h) volte."

$p = \frac{2}{5}$ = probabilità che esca **testa** in una singola prova

$q = 1 - p = 1 - \frac{2}{5} = \frac{3}{5}$ = probabilità che in una singola prova non esca **testa**

Innanzitutto osserviamo che $TT\bar{T}\bar{T}\bar{T}$ $[1]$ è una (ma non la sola) configurazione (modalità) dell'evento E = la testa esce due volte in 5 lanci

$$p(E) = p(TT\bar{T}\bar{T}\bar{T}) = p(T) \cdot p(T) \cdot p(\bar{T}) \cdot p(\bar{T}) \cdot p(\bar{T}) = p^2 \cdot (1-q)^3 \quad [2]$$

Nel caso di h successi in n prove abbiamo: $p(E) = p^h \cdot (1-p)^{n-h}$

Ma la configurazione (modalità) $[1]$ non è la sola che risolve la nostra richiesta perché, anche la configurazione (modalità) $\bar{T}TT\bar{T}\bar{T}$ va altrettanto bene e si presenta con la stessa probabilità $[2]$.

Ne consegue la necessità di accertare quante sono le configurazioni (modalità) equiprobabili che soddisfano il quesito proposto. Esse sono tante quante le permutazioni con ripetizione di 5 elementi di cui 2 (successi) uguali tra loro e 3 (insuccessi) pure uguali tra loro ma distinti dai precedenti.

Ricordando che $P_n^{(\alpha,\beta)} = \frac{n!}{\alpha!\beta!}$ = permutazioni di n elementi dei quali α uguali tra loro e β uguali tra loro ma diversi da α con $\alpha + \beta = n$

$$\binom{n}{k} = C_{n,k} = \frac{n!}{k!(n-k)!} = \text{combinazioni semplici di } n \text{ elementi di classe } k$$

Ritornando all'esempio precedente possiamo scrivere: $P_5^{(2,3)} = \frac{5!}{2!(5-2)!} = \binom{5}{2} = C_{5,2}$

Nel caso di h successi in n prove le configurazioni (modalità) sono: $\frac{n!}{k!(n-k)!} = \binom{n}{k} = C_{n,k}$

Il quesito proposto richiede di determinare la probabilità che si presentino due teste (2 successi) in 5 prove. Abbiamo:

$$p(X=2) = \underbrace{p(TT\bar{T}\bar{T}\bar{T}) \cdot p(T\bar{T}T\bar{T}\bar{T}) \cdots p(\bar{T}\bar{T}\bar{T}TT)}_{\binom{5}{2}\text{ addendi}} = \binom{5}{2} p^2 (1-p)^3$$

$$p(X=2) = \binom{5}{2} p^2 (1-p)^3 = \binom{5}{2} \cdot \left(\frac{2}{5}\right)^2 \cdot \left(\frac{3}{5}\right)^3 = 10 \cdot \frac{4}{25} \cdot \frac{27}{125} = \frac{216}{625} = 0{,}3456$$

Nel caso generale, la probabilità di avere h successi (h teste) in n prove (n lanci) è:

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

Un dado viene lanciato 4 volte. Qual è la probabilità che la faccia contrassegnata col numero 6 si presenti 3 volte?

$p = \frac{1}{6}$, $q = \frac{5}{6}$, $\binom{4}{3} = \frac{4 \cdot 3 \cdot 2}{3 \cdot 2} = 4$ $\quad \mathbf{p_{4,3}(E) = \binom{4}{3} \cdot \left(\frac{1}{6}\right)^3 \cdot \left(\frac{5}{6}\right)^1 = 4 \cdot \left(\frac{1}{6}\right)^3 \cdot \left(\frac{5}{6}\right)^1 = \frac{5}{324} = 0.015}$

$\mathbf{E}$ = la faccia del dado segnata col numero 6 esce 3 volte in 4 lanci

$\mathbf{E_1}$ = la faccia del dado segnata col numero 6 esce al primo lancio

$\mathbf{E_2}$ = la faccia del dado segnata col numero 6 esce al secondo lancio

$\mathbf{E_3}$ = la faccia del dado segnata col numero 6 esce al terzo lancio

$\mathbf{E_4}$ = la faccia del dado segnata col numero 6 esce al quarto lancio

Gli eventi $B_1 = E_1 \cap E_2 \cap E_3 \cap \overline{E}_4 = 6 \cap 6 \cap 6 \cap \overline{6}$, $B_2 = E_1 \cap E_2 \cap \overline{E}_3 \cap E_4 = 6 \cap 6 \cap \overline{6} \cap 6$, $B_3 = E_1 \cap \overline{E}_2 \cap E_3 \cap E_4 = 6 \cap \overline{6} \cap 6 \cap 6$, $B_4 = \overline{E}_1 \cap E_2 \cap E_3 \cap E_4 = \overline{6} \cap 6 \cap 6 \cap 6$ hanno tutti la stessa probabilità $p = \left(\frac{1}{6}\right)^3 \cdot \frac{5}{6}$ di verificarsi.

$$\mathbf{p(B_1) = p(E_1) \cdot p(E_2) \cdot p(E_3) \cdot p(\overline{E}_4) = \frac{1}{6} \cdot \frac{1}{6} \cdot \frac{1}{6} \cdot \frac{5}{6} = \left(\frac{1}{6}\right)^3 \cdot \frac{5}{6}}$$

Il numero di eventi B_i coincide col numero che esprime le combinazioni di 4 elementi di classe 3. $\mathbf{E = B_1 \cup B_2 \cup b_3 \cup B_4}$

$$\boxed{p(E) = p(B_1 \cup B_2 \cup b_3 \cup B_4) = p(B_1) + p(B_2) + p(B_3) + p(B_4) = \binom{4}{3} \cdot p(B_1) = \binom{4}{3} \cdot \left(\frac{1}{6}\right)^3 \cdot \frac{5}{6} = 0{,}015}$$

Calcolare la probabilità che su 12 lanci di una moneta si ottengano 8 volte croce.

E = la croce della moneta esce 8 volte su 12 lanci

Si tratta di un esperimento di Bernoulli nel quale il successo coincide con l'evento esce croce e l'insuccesso coincide con l'evento esce testa.

Risulta: $n = 12$, $h = 8$, $p = q = \frac{1}{2}$

$$p_{12,8}(E) = \binom{12}{8} \cdot \left(\frac{1}{2}\right)^8 \cdot \left(\frac{1}{2}\right)^4 = \binom{12}{8} \cdot \left(\frac{1}{2}\right)^{12} = \frac{12 \cdot 11 \cdot 10 \cdot 9 \cdot 1}{4 \cdot 3 \cdot 2 \cdot 2^{12}} = 0{,}1208$$

Calcolare la probabilità che estraendo per 5 volte una carta da un mazzo di 40 (reinserendo ogni volta la carta estratta e rimescolando bene le carte) si ottengano: a) tre figure b) almeno tre figure c) almeno una figura d) al massimo una figura

Se non reintroducessimo la carta nel mazzo, l'esperimento non sarebbe di Bernoulli in quanto la probabilità di estrarre una figura cambierebbe ad ogni successiva estrazione. Nel caso nostro si tratta di un esperimento di Bernoulli in cui il successo coincide con l'estrazione di una figura e l' insuccesso con l'estrazione di una carta diversa da una figura.

$$p=\frac{12}{40}=\frac{3}{10}\ ,\ q=1-p=\frac{7}{10}$$

a) **E** = la carta estratta è una figura

$p_{5,3}(E)=\binom{5}{3}\cdot\left(\frac{3}{10}\right)^3\cdot\left(\frac{7}{10}\right)^2=\frac{1323}{10000}=0{,}1323$ = probabilità che la figura esca 3 volte in 5 estrazioni

b) La probabilità di ottenere almeno 3 figure coincide con la somma delle probabilità di ottenere 3, 4, 5 figure. $p = p_{5,3} + p_{5,4} + p_{5,5}$

$$\mathbf{p}=\binom{5}{3}\cdot\left(\frac{3}{10}\right)^3\cdot\left(\frac{7}{10}\right)^2+\binom{5}{4}\cdot\left(\frac{3}{10}\right)^4\cdot\left(\frac{7}{10}\right)+\binom{5}{5}\cdot\left(\frac{3}{10}\right)^5=\frac{1323}{10000}+\frac{567}{20000}+\frac{243}{100000}=0{,}16308$$

c) Per rispondere alla terza domanda potremmo procedere come per la domanda 2 calcolando: $\mathbf{p} = \mathbf{p}_{5,1}+\mathbf{p}_{5,2}+\mathbf{p}_{5,3}+\mathbf{p}_{5,4}+\mathbf{p}_{5,5}$

Però è molto più rapido considerare la probabilità dell'evento contrario, ossia la probabilità di non ottenere alcuna figura:

$$\mathbf{p}_{5,0}=\binom{5}{0}\cdot\left(\frac{3}{10}\right)^0\cdot\left(\frac{7}{10}\right)^5=1\cdot 1\cdot\frac{16807}{100000}=0{,}16807$$

La probabilità di ottenere almeno una figura è: $\mathbf{p}=1-\mathbf{p}_{5,0}=1-0{,}16807=0{,}83193$

d) $$\mathbf{p}=\mathbf{p}_{5,0}+\mathbf{p}_{5,1}=1\cdot 1\cdot\frac{16807}{100000}+\binom{5}{1}\cdot\left(\frac{3}{10}\right)^1\cdot\left(\frac{7}{10}\right)^5=\frac{16807}{100000}\text{ » }0{,}16807$$

Distribuzione binomiale o distribuzione di Bernoulli

Indichiamo con **X** la variabile casuale binomiale che rappresenta il numero **h** di successi in **n** prove.

Definizione di variabile casuale binomiale: Si dice che una variabile casuale discreta $\mathbf{X}$, con valori $x = 0, 1, 2, 3, \cdots, n$, ha una distribuzione di probabilità binomiale di parametri $\mathbf{n}$ e $\mathbf{h}$ se la sua funzione di massa è data da:

$$\mathbf{f(x) = P(X=h) = p_{n,h}(X) = p_n(X=h) = \binom{n}{h} \cdot p^h \mathbb{R} q^{n-h} = \frac{n!}{h!(n-h)!} \cdot p^h \cdot q^{n-h} = p(X,n,h) = p_{X,n,h}}$$

dove si hanno $\mathbf{h}$ successi ed $\mathbf{n-h}$ insuccessi, e $\mathbf{p}$ rappresenta la probabilità che si abbia un successo (si verifica un singolo evento $\mathbf{E}$) e $\mathbf{q = 1-p}$ rappresenta la probabilità che si abbia un insuccesso (si verifica un singolo insuccesso cioè si verifica l'evento contrario $\overline{\mathbf{E}}$).

$\mathbf{p(X,n,h) = p_{X,n,h}}$ = probabilità che l'evento $\mathbf{X}$, in $\mathbf{n}$ prove, si verifichi $\mathbf{h}$ volte.

Una variabile casuale con distribuzione binomiale descrive il numero di volte $\mathbf{h}$ che si può verificare un evento aleatorio $\mathbf{E}$ di probabilità $\mathbf{p}$ su $\mathbf{n}$ prove.

In forma tabellare una variabile casuale binomiale $\mathbf{X}$ assume la seguente forma:

$$X = \begin{cases} \begin{array}{ccccccc} x_1 & x_2 & x_3 & \cdots & x_h & \cdots & x_n \\ 0 & 1 & 2 & \cdots & h & \cdots & n \\ p(x_1) = p_{n,0} & p(x_2) = p_{n,1} & p(x_3) = p_{n,2} & \cdots & p(x_h) = p_{n,h} & \cdots & p(x_n) = p_{n,n} \\ q^n & \binom{n}{1} p\, q^{n-1} & \binom{n}{2} p^2 q^{n-2} & \cdots & \binom{n}{h} p^h q^{n-h} & \cdots & p^n \end{array} \end{cases}$$

Spesso, in un esperimento con prove ripetute, ognuna delle quali ha la probabilità $\mathbf{p}$ di successo, la probabilità di ottenere $\mathbf{h}$ successi si indica con la scrittura:

$$P(n,h,p)$$

Ad esempio la scrittura $P\left(12,4,\frac{2}{7}\right)$ significa che si sono effettuate 12 prove ottenendo 4 successi sapendo che è uguale a $\frac{2}{7}$ la probabilità di avere un successo (che cioè si verifichi l'evento **E**.

La funzione di ripartizione $F(X)$ della variabile casuale binomiale **X** è:

$$\mathbf{F(X) = P(X \le h) = p(X=0) + p(X=1) + p(X=2) + \cdots + p(X=h) = \sum_{k=0}^{k=h} \binom{n}{k} \cdot p^k \cdot q^{n-k}}$$

La probabilità che in **n** prove il numero di successi sia compreso tra x_3 ed x_5 ci viene data da: $P_n(x_3 \le X \le x_5) = p_n(X = x_3) + p_n(X = x_4) + p_n(X = x_5)$

$P_n(E = h)$ =probabilità che in **n** prove l'evento **E** si presenti **h** volte

$$P_n(E = h) = \binom{n}{h} \cdot p^h \cdot q^{n-h} = \frac{n!}{h!(n-h)!} \cdot p^h \cdot q^{n-h}$$

$P_n(E \le h)$ = probabilità che in **n** prove l'evento **E** si presenti al massimo **h** volte

$$P_n(E \le h) = p_n(E = 1) + p_n(E = 2) + \cdots + p_n(E = h)$$

$P_n(h \le E \le n)$ = probabilità che in **n** prove l'evento **E** si presenti al almeno **h** volte

$$P_n(h \le E \le n) = p_n(E = h) + p_n(E = h+1) + \cdots + p_n(E = n)$$

B = in **n** prove l'evento **E** si presenti al almeno **h** volte

$\overline{\mathbf{B}}$ = in **n** prove l'evento **E** si presenti al massimo **h-1** volte

$$\mathbf{p(B) + p(\overline{B}) = 1} \qquad \mathbf{p(B) = 1 - p(\overline{B})}$$

Gli eventi **B** e $\overline{\mathbf{B}}$ sono uno opposto dell'altro.

Le principali proprietà della distribuzione binomiale sono:

Media	$\mu = M(X) = np$
Varianza	$\sigma^2 = npq = np(1-p)$
Deviazione standard o scarto quadratico medio	$\sigma = \sqrt{npq} = \sqrt{np(1-p)}$

Si lanci **6** volte una moneta non truccata. Sia <<esce croce>> il successo. Calcolare: a) la probabilità che si presentino **2** croci b) la probabilità di ottenere almeno **4** croci c) la probabilità che l'evento croce non si presenti mai (cioè che tutti gli eventi siano insuccessi).

a) Calcoliamo la probabilità che si presentino **2** croci.

E = nel lancio di una moneta esce croce

Abbiamo: $n = 6$, $p(\mathbf{E}) = q(\overline{\mathbf{E}}) = \frac{1}{2}$, $h = 2$ $\quad p_6(E=2) = \binom{6}{2} \cdot \left(\frac{1}{2}\right)^2 \cdot \left(\frac{1}{2}\right)^4 = \frac{15}{64}$

$p_6(E=2)$ = probabilità che lanciando **6** volte una moneta la croce esca **2** volte

b) Calcoliamo la probabilità di ottenere almeno **4** croci

$P_6(4 \le E \le 6) = p_6(4 \le E \le 6)$ = probabilità che lanciando **6** volte una moneta la croce esca almeno **2** volte

$$P_6(4 \le E \le 6) = p_6(E=4) + p_6(E=5) + p_6(X=6) = \binom{6}{4} \cdot \left(\frac{1}{2}\right)^4 \cdot \left(\frac{1}{2}\right)^2 + \binom{6}{5} \cdot \left(\frac{1}{2}\right)^5 \cdot \left(\frac{1}{2}\right) + \binom{6}{6} \cdot \left(\frac{1}{2}\right)^6$$

$$p_6(4 \le E \le 6) = \frac{15}{64} + \frac{6}{64} + \frac{1}{64} = \frac{11}{32}$$

c) Calcoliamo la probabilità che l'evento croce non si presenti mai

$$P_6(E=0) = p_6(E=0) = q^n = q^6 = \left(\frac{1}{2}\right)^6 = \frac{1}{64}$$

e quindi, la probabilità di ottenere nei 6 lanci almeno una croce è data da:

$$P_6(1 \le E \le 6) = 1 - P_6(E=0) = 1 - q^6 = 1 - \frac{1}{64} = \frac{63}{64}$$

Un'urna contiene 7 palline rosse e 3 nere. Una prova consiste nell'estrarre, una di seguito all'altra, tre palline. Si considera successo il caso in cui le 3 palline estratte siano rosse. Determinare, nell'ipotesi che la prova venga ripetuta 96 volte, i parametri μ , σ^2 , □ quando:

a) ogni prova prevede la reimmissione della pallina estratta

b) la reimmissione non è prevista

a) $\mathbf{E_1}$ = la prima pallina estratta è rossa $\mathbf{E_2}$ = la seconda pallina estratta è rossa

$\mathbf{E_3}$ = la terza pallina estratta è rossa

$\mathbf{E}=\mathbf{E_1}\cap\mathbf{E_2}\cap\mathbf{E_3}$ = tutte e tre le palline estratte sono rosse $p(E_1)=p(E_2)=p(E_3)=\frac{7}{10}$

$p(\mathbf{E_1}\cap\mathbf{E_2}\cap\mathbf{E_3})=p(E_1)\cdot p(E_2)\cdot p(E_3)=\frac{7}{10}\cdot\frac{7}{10}\cdot\frac{7}{10}=\left(\frac{7}{10}\right)^3$ = probabilità di successo in ogni prova

$q=1-p=1-\left(\frac{7}{10}\right)^3$ = probabilità di insuccesso in ogni prova

$$\mu=n\,p=96\cdot\left(\frac{7}{10}\right)^3=\mu=n\,p=96\cdot\left(\frac{7}{10}\right)^3=\frac{4116}{125}=32{,}93$$

$$\sigma^2=npq=\frac{4116}{125}\cdot\left[1-\left(\frac{7}{10}\right)^3\right]=\frac{676053}{31250}=21{,}63 \qquad \sigma=\sqrt{n\,p\,q}=\sqrt{21{,}63}=4{,}65$$

b) Se ad ogni prova la reinserimento non è previsto abbiamo:

$$p(E_1)=\frac{7}{10}\ ,\ p(E_2)=\frac{6}{9}\ ,\ p(E_3)=\frac{5}{8}$$

$$p(\mathbf{E_1}\cap\mathbf{E_2}\cap\mathbf{E_3})=p(E_1)\cdot p(E_2)\cdot p(E_3)=\frac{7}{10}\cdot\frac{6}{9}\cdot\frac{5}{8}=\frac{210}{720}=\frac{7}{24}=0{,}29$$

$$q=1-p=1-\frac{210}{720}=\frac{510}{720}=\frac{17}{24}=0{,}71$$

$$\mu=n\,p=96\cdot\frac{210}{720}=28 \qquad \sigma^2=npq=28\cdot\frac{410}{720}=19{,}83 \qquad \sigma=\sqrt{n\,p\,q}=\sqrt{19{,}83}=4{,}45$$

Da un mazzo di 40 carte si effettuano 4 estrazioni con rimessa della carta nel mazzo. Descrivere la variabile casuale X = numero di assi.

Si tratta di una variabile casuale che segue la legge di distribuzione binomiale avente come parametri $n=4$, $p=\frac{4}{40}=\frac{1}{10}$, $q=1-p=\frac{9}{10}$

$p(\mathbf{X}=0)=\left(\frac{9}{10}\right)^4=0,66 \qquad p(\mathbf{X}=1)=\binom{4}{1}\left(\frac{1}{10}\right)\left(\frac{9}{10}\right)^3=0,29$

$p(\mathbf{X}=2)=\binom{4}{2}\left(\frac{1}{10}\right)^2\left(\frac{9}{10}\right)^2=0,05 \qquad p(\mathbf{X}=3)=\binom{4}{3}\left(\frac{1}{10}\right)^3\left(\frac{9}{10}\right)^1=0,0036$

$$X=\begin{cases} 0 & 1 & 2 & 3 & 4 \\ \left(\frac{9}{10}\right)^4=0,66 & \binom{4}{1}\frac{1}{10}\left(\frac{9}{10}\right)^3=0,29 & \binom{4}{2}\left(\frac{1}{10}\right)^2\left(\frac{9}{10}\right)^2=0,05 & \binom{4}{3}\left(\frac{1}{10}\right)^3\frac{9}{10}=0,0036 & \left(\frac{1}{10}\right)^4=0,0001 \end{cases}$$

$$X=\begin{cases} 0 & 1 & 2 & 3 & 4 \\ 0,66 & 0,29 & 0,05 & 0,0036 & 0,0001 \end{cases}$$

Si supponga di lanciare 3 monete. Si vuole sapere qual è la probabilità di avere: (1) due teste (2) al massimo una testa

Lanciare tre monete equivale a lanciare una moneta tre volte; siamo nel caso delle prove ripetute, cioè nella distribuzione binomiale.

(1) $\mathbf{E}$ = nel lancio di una moneta esce testa.

$n=3,\ p(\mathbf{E})=q(\overline{\mathbf{E}})=\frac{1}{2},\ h=2 \quad n-h=1$

$$p_3(\mathbf{E}=2)=\binom{3}{2}\cdot\left(\frac{1}{2}\right)^2\cdot\left(\frac{1}{2}\right)=\frac{3!}{2!\cdot 1!}\left(\frac{1}{2}\right)^3=3\cdot\frac{1}{8}=\frac{3}{8}=0,375$$

(2) $p_3(\mathbf{E}\le 1)=p_3(\mathbf{E}=0)+p_3(\mathbf{E}=1)=\binom{3}{0}\cdot\left(\frac{1}{2}\right)^0\cdot\left(\frac{1}{2}\right)^3+\binom{3}{1}\cdot\left(\frac{1}{2}\right)^1\cdot\left(\frac{1}{2}\right)^2=\left(\frac{1}{3}\right)^3+\frac{3!}{2!\cdot 1!}\left(\frac{1}{2}\right)^3\cdot 7$

$$p_3(\mathbf{E}\le 1)=\frac{1}{8}+\frac{3}{8}=\frac{1}{2}=0,5$$

Un'urna contiene 12 palline delle quali 4 sono bianche ed 8 sono nere. Si vuole conoscere la probabilità che estraendo 5 palline (cioè facendo 5 prove) 4 siano bianche ed 1 sia nera. Si sa che dopo ogni estrazione la pallina estratta viene rimessa nell'urna per cui le successive estrazioni sono effettuate sempre nelle stesse condizioni.

$\mathbf{E}$ =si estrae una pallina bianca $\overline{\mathbf{E}}$ = si estrae una pallina nera, cioè non si estrae una pallina bianca $n=5$, $h=4$ $n-h=1$ $p(\mathbf{E})=\frac{4}{12}=\frac{1}{3}$ $q(\overline{\mathbf{E}})=\frac{8}{12}=\frac{2}{3}$

La probabilità che estraendo 5 palline 4 siano bianche ed una nera è:

$$p_5(\mathbf{E}=\mathbf{4})=\binom{5}{4}\cdot\left(\frac{1}{3}\right)^4\cdot\left(\frac{2}{3}\right)^1=5\cdot\frac{1}{3^5}=0,04115$$

Lanciando una coppia di dadi cinque volte qual è la probabilità che si ottenga un punteggio totale maggiore di sette almeno due volte?

evento A = in un lancio di due dadi otteniamo un punteggio totale maggiore di **7**

I casi possibili sono **36**, i casi favorevoli sono **15**

6	7	**8**	**9**	**10**	**11**	**12**
5	6	7	**8**	**9**	**10**	**11**
4	5	6	7	**8**	**9**	**10**
3	4	5	6	7	**8**	**9**
2	3	4	5	6	7	**8**
1	2	3	4	5	6	7
	1	2	3	4	5	6

$$p(A)=\frac{15}{36}=\frac{5}{12} \qquad p(\overline{A})=1-p(A)=1-\frac{5}{12}=\frac{7}{12}$$

Si tratta di una distribuzione binomiale: Indicando con n il numero di prove, con k il numero di "successi", con p la probabilità del "successo" e con q la probabilità dell'insuccesso abbiamo: $p_{n,k}(A)=P_n(A=k)=\binom{n}{k}\cdot p^k\cdot q^{n-k}$

La probabilità che che l'evento A si verifichi $k=2$ volte in $n=5$ lanci è:

$$p_{5,2}(A)=\binom{5}{2}\cdot\left(\frac{5}{12}\right)^2\cdot\left(\frac{7}{12}\right)^3$$

B= l'evento A si verifica in cinque lanci almeno due volte

$$p(B)=p_5(A=2)+p_5(A=3)+p_5(A=4)+p_5(A=5)$$

Però è molto più rapido considerare la probabilità dell'evento contrario, ossia la probabilità che l'evento A non si verifichi due volte, cioè al massimo si verifica una volta:

$$p(\overline{B})=P_5(A=0)+P_5(A=1)=\binom{5}{0}\left(\frac{5}{12}\right)^0\left(\frac{7}{12}\right)^{5-0}+\binom{5}{1}\left(\frac{5}{12}\right)^1\left(\frac{7}{12}\right)^{5-1}=\binom{5}{0}\left(\frac{5}{12}\right)^0\left(\frac{7}{12}\right)^5+\binom{5}{1}\left(\frac{5}{12}\right)^1\left(\frac{7}{12}\right)^4$$

Adesso possiamo calcolare la probabilità che l'evento A si verifichi almeno due volte

$$p(B)=1-p(\overline{B})=1-p_5(A=0)-p_5(A=1)=1-\binom{5}{0}\left(\frac{5}{12}\right)^0\left(\frac{7}{12}\right)^5-\binom{5}{1}\left(\frac{5}{12}\right)^1\left(\frac{7}{12}\right)^4$$

$$p(B)=1-\left(\frac{7}{12}\right)^5-5\cdot\left(\frac{5}{12}\right)\cdot\left(\frac{7}{12}\right)^4=1-\frac{16807}{248832}-5\cdot\left(\frac{5}{12}\right)\cdot\frac{2401}{20736}$$

$$p(B)=1-\frac{16807}{248832}-\frac{60025}{248832}=\frac{5375}{7776}\approx 0{,}691\approx 69\%$$

Una prova scritta di matematica è strutturata come test a risposta multipla: vi sono 9 domande, ognuna con 4 possibili risposte, tra le quali una sola è esatta. Per ottenere la sufficienza occorre rispondere esattamente ad almeno 6 domande. Uno studente impreparato indica le risposte a caso: qual è la probabilità che lo studente ottenga la sufficienza?

E = l'alunno raggiunge la sufficienza

Se la risposta a ciascuna domanda è scelta a caso, la probabilità che sia stata scelta la risposta è $p=\frac{1}{4}$. Anche in questo caso siamo in presenza di prove ripetute, e dunque il numero di risposte esatte è una variabile casuale **X** a distribuzione binomiale, con:

$n=9 \qquad p=\frac{1}{4} \qquad q=1-p=\frac{3}{4}$

Lo studente ottiene la sufficienza se $\mathbf{X}=6 \vee \mathbf{X}=7 \vee \mathbf{X}=8 \vee \mathbf{X}=9$

L'evento **E** è l'unione di **4** eventi disgiunti e la sua probabilità sarà la somma di tali probabilità, ossia: $p(\mathbf{E})=p_9(X=6)+p_9(X=7)+p_9(X=8)+p_9(X=9)$

$p_9(X=6)=\binom{9}{6}\cdot\left(\frac{1}{4}\right)^6\cdot\left(\frac{3}{4}\right)^3=0,008651 \qquad p_9(X=7)=\binom{9}{7}\cdot\left(\frac{1}{4}\right)^7\cdot\left(\frac{3}{4}\right)^2=0,001235$

$p_9(X=8)=\binom{9}{8}\cdot\left(\frac{1}{4}\right)^8\cdot\left(\frac{3}{4}\right)^1=0,000103 \qquad p_9(X=9)=\binom{9}{9}\cdot\left(\frac{1}{4}\right)^9\cdot\left(\frac{3}{4}\right)^0=0,0000038$

$$p(\mathbf{E})=0,008651+0,001235+0,000103+0,0000038=0,009994\approx 1\%$$

Distribuzione di Poisson

Definizione: Sia **X** una variabile casuale discreta che può assumere i valori $0,1,2,3,\cdots,n$ in modo che la distribuzione di probabilità di **X** sia: $\mathbf{P}(\mathbf{X}=\mathbf{x})=\frac{\boldsymbol{\lambda}^{\mathbf{x}}\mathbf{e}^{-\boldsymbol{\lambda}}}{\mathbf{x}!}$ con $x=0,1,2,3,\cdots,n$ dove λ rappresenta una costante positiva assegnata. Tale distribuzione di probabilità è detta distribuzione di Poisson di parametro λ e la variabile casuale **X** $X=\begin{cases} x: & 0 & 1 & 2 & \cdots & n \\ p: & e^{-\lambda} & \frac{\lambda}{1!}e^{-\lambda} & \frac{\lambda^2}{2!}e^{-\lambda} & \cdots & \frac{\lambda^x}{x!}e^{-\lambda}\end{cases}$ si chiama variabile casuale di Poisson di parametro λ.

Si può dimostrare che il parametro λ di una distribuzione di Poisson coincide col valore medio e la varianza della variabile casuale X, cioè risulta:

$$\lambda = M(X) = var(X) = \sigma^2$$

Questo significa che il parametro λ ha il significato di media e di varianza della variabile casuale di Poisson (detta anche legge degli eventi rari o legge dei piccoli numeri) è una variabile casuale discreta che può assumere una infinità numerabile di valori. Essa ha un grande interesse applicativo, poiché si presta a rappresentare il numero di manifestazioni di un dato fenomeno in un intervallo di tempo.

La distribuzione di Poisson è una buona approssimazione della distribuzione binomiale quando n è grande e p è piccolo. Sotto queste ipotesi risulta: $\lambda = np$ dove n il numero di modalità (valori)che può assumere la variabile casuale X e p rappresenta la probabilità dell'evento.

Esempio: Un'urna contiene 30 palline numerate da 1 a 30. Determinare il numero di volte in cui può uscire, effettuando 4 estrazioni nelle stesse condizioni, la pallina col numero 1.

In questo caso possiamo utilizzare la distribuzione di Poisson in quanto è un valore la probabilità dell'evento esce il numero 1. $P(X=1) = \frac{1}{30}$

$n = 4$ in quanto si effettuano 4 estrazioni $\lambda = np = 4 \cdot \frac{1}{30} = \frac{2}{15}$

In questo caso la distribuzione di Poisson diventa: $P(X = x) = \frac{\lambda^x e^{-\lambda}}{x!} = \frac{\left(\frac{2}{15}\right)^x}{x!} \times e^{-\frac{2}{15}}$

$$P(X=1)=\frac{\left(\frac{2}{15}\right)^1}{1!}\cdot e^{-\frac{2}{15}}=\frac{2}{15}\cdot e^{-\frac{2}{15}}\approx 0{,}35 \qquad \lambda=\frac{2}{15}$$

La distribuzione di Poisson può essere considerata una approssimazione della distribuzione binomiale, approssimazione tanto più precisa e nello stesso tempo più utile quando il parametro **n** è grande e risulta disagevole il calcolo dei coefficienti binomiali. Quando nella distribuzione binomiale **n** è grande mentre la probabilità **p** è vicina allo zero così che $q=1-p$ è vicino ad **1**, l'evento è detto raro. Nella pratica si considera raro un evento se il numero **n** delle prove sarà maggiore di **50** mentre **np** sarà minore di **5**. In questi casi la distribuzione binomiale è approssimata dalla distribuzione di Poisson ponendo $\lambda=np$.

Per tale motivo tale distribuzione trova applicazione in tutte quelle situazioni che possono essere ricondotte al problema delle prove ripetute, laddove il numero delle prove **n** sia molto alto. Si osservi infine che, mentre la distribuzione binomiale dipende da due parametri (n,p), la distribuzione di Poisson dipende solo dal parametro λ, che rappresenta la media. E' questo l'unico dato richiesto per potere applicare la distribuzione di Poisson: in particolare, non è necessario conoscere il numero **n** delle prove.

Problema N°1: **All'arrivo in un aeroporto, i passeggeri passano la dogana alla media di 2 ogni 30 secondi. Ammettendo che il numero di passeggeri che attraversano la dogana in un dato intervallo di tempo segua la legge della distribuzione di Poisson, determinare la probabilità che non più di 2 passeggeri abbiano attraversato la dogana, sempre in un periodo di 30 secondi.**

$$P(X\le 2)=P(X=0)+P(X=1)+P(X=2) \quad \text{con} \quad \lambda=2$$

Gli eventi sono 0,1,2 passeggeri hanno attraversato la dogana

$P(x=0)=\frac{2^0\cdot e^{-2}}{0!}=0{,}13534 \quad P(x=1)=\frac{2^1\cdot e^{-2}}{1!}=0{,}27067 \quad P(x=2)=\frac{2^2\cdot e^{-2}}{2!}=0{,}27067$

$$P(X\le 2)=0{,}13534+0{,}27067+0{,}27067=0{,}67668$$

Problema N°2: **Una macchina produce pezzi difettosi con una probabilità $p=0{,}006$. Calcoliamo la probabilità che su $n=500$ pezzi:**

(a) nessun pezzo risulti difettoso **(b)** risultino difettosi 3 pezzi **(c)** risultino difettosi più di 5 pezzi

La variabile casuale relativa al numero di pezzi difettosi ha valore medio:

$\lambda=\mu=n\cdot=500\cdot 0{,}006=3$ e distribuzione di probabilità $P(X=x)=\frac{\lambda^x e^{-3}}{x!}$

(a) la probabilità che nessun pezzo risulti difettoso è: $P(X=0)=\frac{3^0 e^{-3}}{0!}\simeq 0{,}04979$

(b) la probabilità che risultino difettosi 3 pezzi è: $P(X=3)=\frac{3^3 e^{-3}}{3!}\simeq 0{,}22404$

(c) Per calcolare la probabilità che risultino difettosi più di 5 pezzi bisogna procedere come segue.

Mi calcolo la probabilità che non vi siano meno di 5 pezzi difettosi, cioè: $P(X\le 5)$

La probabilità che risultino difettosi più di 5 pezzi si calcola applicando la seguente formula:

$$P(X>5)=1-P(X\le 5)$$

$$P(X\le 5)=P(X=0)+P(X=1)+P(X=2)+P(X=3)+P(X=4)+P(X=5)$$

$$P(X\le 5)=0{,}049787+0{,}149361+0{,}224042+0{,}224042+0{,}168031+0{,}100819=0{,}916082$$

Problema N°3: **Ad uno sportello bancario arrivano in media 30 persone all'ora. Calcolare la probabilità che in 5 minuti: (a) arrivino 4 persone (b) arrivino meno di 3 persone.**

Per calcolare il valore del parametro positivo λ della distribuzione di Poisson bisogna applicare la seguente proporzione: $30:60 = y:5$ in quanto $1\,\text{ora} = 60\,\text{minuti}$ $\quad y = \lambda = \dfrac{30 \cdot 5}{60} = 2{,}5$

Questo significa che ad uno sportello bancario arrivano in media $2{,}5$ persone ogni 5 minuti

Il parametro λ può essere determinato calcolando la probabilità che una sola persona si presenti ogni minuto. Bisogna applicare la seguente formula: $p = \dfrac{30}{60} = \dfrac{1}{2} = 0{,}5 \quad \lambda = 5 \cdot 0{,}5 = 2{,}5$

(a) la probabilità che arrivino 4 persone è: $P(X=4) = \dfrac{2{,}5^4 e^{-\lambda}}{4!} = 0{,}133602$

(b) la probabilità che arrivino meno di 3 persone è:

$$P(X \leq 3) = P(X = 0) + P(X = 1) + P(X = 2) + P(X = 3) + P(X = 4)$$

$$P(X \leq 3) = 0{,}082085 + 0{,}205212 + 0{,}256516 + 0{,}213763 + 0{,}133602 = 0{,}891178$$

Problema N°4: **In una stazione di servizio autostradale arrivano in media 3 automobili ogni minuto. Calcolare la probabilità che: (a) che in un minuto arrivi una sola automobile (b) che in un minuto arrivino meno di 4 automobili (c) che in un minuto almeno una sola automobile (d) che in due minuti consecutivi arrivi una sola automobile (e) che in due minuti consecutivi arrivino due automobili**

- Il problema posto può essere ricondotto a quello delle prove ripetute: ogni automobilista può essere considerato una "**prova**" che, nel corso del minuto considerato può recarsi alla stazione di servizio (**successo**) o non recarvisi (**insuccesso**). Poiché il numero delle prove, ossia degli automobilisti, è incognito ma comunque molto grande, si può applicare la distribuzione di Poisson con $\lambda = 3$.

(a) la probabilità che in un minuto arrivi una sola automobile è: $P(X=1)=\dfrac{3^1 e^{-3}}{3!}=0{,}149$

(b) la probabilità che in un minuto arrivino meno di 4 automobili.

Dobbiamo calcolare la probabilità che in un dato minuto arrivino $0,1,2,3,4$ automobili. Applicando il teorema della probabilità totale e tenendo presente che tali eventi sono incompatibili, otteniamo:

$P(X\le 4)=P(X=0)+P(X=1)+P(X=2)+P(X=3)$ $\quad P(X\le 4)=\dfrac{3^0 e^{-3}}{0!}+\dfrac{3^1 e^{-3}}{1!}+\dfrac{3^2 e^{-3}}{2!}+\dfrac{3^3 e^{-3}}{3!}=0{,}647$

(c) la probabilità che in un minuto almeno una sola automobile

Conviene utilizzare il teorema della probabilità contraria: l'evento di cui si chiede la probabilità è l'evento contrario dell'evento "**arrivano 0 automobili**", che ha la seguente probabilità:

$$P(X=0)=\frac{3^0 e^{-3}}{0!}=e^{-3}$$

Perciò la probabilità che in un minuto almeno una sola automobile è: $P(X\ge 1)=1-e^{-3}=0{,}9502$

(d) la probabilità che in due minuti consecutivi arrivi una sola automobile

L'evento considerato è l'unione di due eventi incompatibili: "**nel primo minuto arriva una sola automobile e nel secondo minuto non ne arriva nessuna**" oppure "**nel primo minuto non arriva nessuna automobile e nel secondo minuto ne arriva una**". Ciascuno di tali due eventi può a sua volta essere considerato come l'intersezione di due eventi indipendenti. La probabilità richiesta è:

$$p(E_1\cup E_2)=p(E_1)+P(E_2)=P(X=1)\cdot P(X=0)+P(X=0)\cdot P(X=1)=2P(X=0)\cdot P(X=1)$$

$$p(E_1\cup E_2)=2P(X=1)\cdot P(X=0)=2\cdot\frac{3^0 e^{-3}}{0!}\cdot\frac{3^1 e^{-3}}{1!}=6\cdot e^{-3}=0{,}0148$$

(e) la probabilità che in due minuti consecutivi arrivino due automobili

Ragionando come nel caso precedente otteniamo:

$$p(E_1 \cup E_2 \cup E_3) = p(E_1) + P(E_2) + P(E_3) = P(X=0) \cdot P(X=2) + P(X=1) \cdot P(X=1) + P(X=2) \cdot P(X=0)$$

$$p(E_1 \cup E_2 \cup E_3) = 2P(X=0) \cdot P(X=2) + P^2(X=1) = 2\frac{3^0 e^{-3}}{0!}\frac{3^2 e^{-3}}{2!} + \left(\frac{3^1 e^{-3}}{1!}\right)^2 = 18e^{-6} = 0{,}0446$$

I quesiti **(d)** e **(e)** possono essere risolti anche considerando la variabile casuale che indica il numero di automobili che arrivano alla stazione in due minuti. Tale variabile casuale ha una distribuzione di Poisson con media $\lambda = 6$. Infatti, se in un minuto arrivano in media 2 due automobili, in due minuti arrivano in media 6 automobili. $\lambda = 6$.

Nel quesito **(d)** abbiamo: $P(X=1) = \frac{6^1 \cdot e^{-6}}{1!} = 6 \cdot e^{-6} = 0{,}0148$

Nel quesito **(e)** abbiamo: $P(X=2) = \frac{6^2 \cdot e^{-6}}{2!} = \frac{6^2}{2!} \cdot e^{-6} = 18 \cdot e^{-6} = 0{,}0446$

Simulazione del 22 Aprile 2015 Quesito 10

In una stazione ferroviaria, fra le 8 e le 10 del mattino, arrivano in media ogni 20 minuti 2 treni. Determinare la probabilità che in 20 minuti: (a) non arrivi alcun treno (b) ne arrivi uno solo (c) ne arrivino al massimo 4.

Per calcolare le probabilitàrichieste dobbiamo per prima cosa capire come modellizzare gli eventi, cioè trovare la variabile casuale che descrive gli arrivi dei treni. La variabile casuale che descrive il numero di successi (in questo caso il numero di treni che arrivano) in un determinato tempo è la variabile di Poisson. Per definire tale variabile abbiamo bisogno di un parametro che descrive la media dei successi che vine inbdicata col simbolo λ. Per la risoluzione del problema possiamo applicare in tutti e tre i casi la distribuzione di probabilità di Poisson secondo cui la probabilità che l'evento si verifichi è data dalla relazione: $P(X=x) = \frac{\lambda^x}{x!} \cdot e^{-\lambda}$

Sappiamo che in media ogni 20 minuti arrivano 2 treni, quindi la media è $\lambda = 2$ per cui la formula precedente diventa: $P(X=x) = \frac{2^x}{x!} \cdot e^{-2}$

(a) $x=0 \quad \Rightarrow \quad P(X=0)=\frac{2^0}{0!}\cdot e^{-2}=e^{-2}\approx 0{,}135=13{,}5\%$

(b) $x=1 \quad \Rightarrow \quad P(X=1)=\frac{2^1}{1!}\cdot e^{-2}=\frac{2}{e^2}\approx 0{,}271=27{,}1\%$

(c) $x\le 4 \quad \Rightarrow \quad P(X\le 4)=P(X=0)+P(X=1)+P(X=2)+P(X=3)+P(X=4)$

$$P(X\le 4)=\frac{2^0}{0!}\cdot e^{-2}+\frac{2^1}{1!}\cdot e^{-2}+\frac{2^2}{2!}\cdot e^{-2}+\frac{2^3}{3!}\cdot e^{-2}+\frac{2^4}{4!}\cdot e^{-2}$$

$$P(X\le 4)=\frac{1}{e^2}+\frac{2}{e^2}+\frac{2}{e^2}+\frac{4}{3e^2}+\frac{2}{3e^2}=\frac{1}{e^2}\left(1+2+2+\frac{4}{3}+\frac{2}{3}\right)=\frac{1}{e^2}\left(1+2+2+\frac{6}{3}\right)$$

$$P(X\le 4)=\frac{1}{e^2}(1+2+2+2)=\frac{7}{e^2}\approx 0{,}947=94{,}7\%$$