

VARIABILI ALEATORIE DISCRETE E CONTINUE

Una **VARIABILE ALEATORIA QUANTITATIVA X** può essere definita come un fenomeno le cui modalità sono espresse da numeri o, più precisamente, come una funzione definita sullo spazio degli eventi Ω che associa ad ogni elemento ω di Ω un numero reale $X(\omega)=x$.

Le variabili aleatorie quantitative sono **DISCRETE** se producono risposte numeriche che derivano da un processo di conteggio. Ad es. “Il numero dei componenti la famiglia”, “il numero delle stanze di un’abitazione”, “il numero dei biglietti dell’autobus utilizzati in un mese”, ecc.

Le variabili aleatorie quantitative sono **CONTINUE** se generano risposte che derivano da un processo di misurazione. Ad es. “l’altezza”, “il reddito”, “il fatturato”, ecc.

Ad una variabile aleatoria discreta è associata una **DISTRIBUZIONE DI PROBABILITA'**.

La **DISTRIBUZIONE DI PROBABILITA'** di una variabile aleatoria discreta è rappresentata dall'elenco delle modalità che la variabile assume, a ciascuna delle quali è associata la relativa probabilità.

Esempio 1: Nella tabella 3 è riportata la distribuzione del numero di ipoteche approvate settimanalmente da parte di una banca.

Tabella 3: Distribuzione del numero di ipoteche approvate in una settimana da parte di una banca

Numero di ipoteche approvate	Probabilità = $P(x_i)$
0	0,10
1	0,10
2	0,20
3	0,30
4	0,15
5	0,10
6	0,05

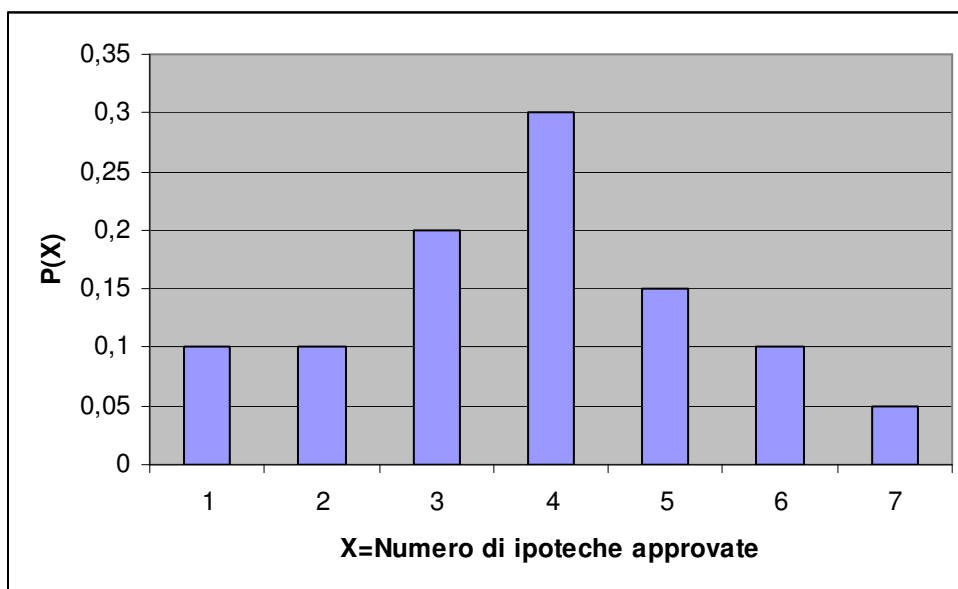
Chiamiamo la funzione $P(X = x_i) = P(x_i)$ **FUNZIONE DI PROBABILITA'**.

Poiché le modalità elencate sono mutuamente esclusive e collettivamente esaustive, le probabilità sommano ad 1:

$$P(x_i) \in [0,1] \quad \forall i$$

$$\sum_{i=1}^m P(x_i) = 1$$

La distribuzione può essere rappresentata graficamente come segue.



Un'altra funzione che descrive il comportamento di una variabile casuale è la **FUNZIONE DI RIPARTIZIONE**.

La **FUNZIONE DI RIPARTIZIONE** fornisce la probabilità che la variabile aleatoria X assuma valori \leq ad un particolare valore x_i :

$$F(X) = P(X \leq x_i) = P(x_1) + P(x_2) + \dots + P(x_i) = \sum_{k=1}^i P(x_k)$$

Quindi, nell'esempio precedente:

Tabella 4: Distribuzione del numero di ipoteche approvate in una settimana da parte di una banca e funzione di ripartizione

Numero di ipoteche approvate	P(xi)	F(xi)
0	0,10	0,10
1	0,10	0,20
2	0,20	0,40
3	0,30	0,70
4	0,15	0,85
5	0,10	0,95
6	0,05	1,00

Rappresentandola graficamente si ottiene una funzione a gradini.

Un modo per sintetizzare una distribuzione di probabilità discreta consiste nel calcolarne le principali misure di sintesi: il valore atteso e la varianza o la sua radice (lo scarto quadratico medio).

Il VALORE ATTESO di una variabile aleatoria discreta è una media ponderata delle modalità assunte dalla variabile, con coefficienti di ponderazione rappresentati dalle probabilità associate a ciascuna modalità:

$$\mu = E(X) = \sum_{i=1}^m x_i P(x_i)$$

dove

x_i = i .esima modalità della v.a. X ;

$P(x_i)$ = probabilità associata alla modalità x_i .

La VARIANZA di una variabile aleatoria discreta è la media ponderata dei quadrati degli scarti fra ciascuna modalità ed il valore atteso della variabile, dove i coefficienti di ponderazione sono rappresentati dalle probabilità associate a ciascuna modalità:

$$\sigma^2 = V(X) = E[X - E(X)]^2 = \sum_{i=1}^m [x_i - E(X)]^2 P(x_i)$$

dove

x_i = *i*.esima modalità della v.a. X ;

$P(x_i)$ = probabilità associata alla modalità x_i .

Si verifica facilmente che $V(X) = E(X^2) - \mu^2$.

Lo SCARTO QUADRATICO MEDIO è dato dalla radice della varianza:

$$\sigma = \sqrt{\sum_{i=1}^m [x_i - E(X)]^2 P(x_i)}$$

Il valore atteso, la varianza e lo scarto quadratico medio dell'esempio 1 risultano rispettivamente 2,8, 2,46 e 1,57.

Il valore atteso, la varianza e lo scarto quadratico medio soddisfano le stesse proprietà enunciate nel caso di variabili statistiche.

Passiamo ad analizzare alcuni modelli matematici in grado di rappresentare alcuni fenomeni di interesse. Nel caso delle variabili aleatorie discrete il modello matematico è costituito dalla distribuzione di probabilità.

Esempio 2: l'estrazione casuale con reinserimento di 10 palline numerate (da 1 a 10) da un'urna, con quale modello distributivo è rappresentabile?

Tabella 5: Estrazione di 10 palline numerate da un'urna

Numero	P(xi)	F(xi)
1	0,10	0,10
2	0,10	0,20
3	0,10	0,30
4	0,10	0,40
5	0,10	0,50
6	0,10	0,60
7	0,10	0,70
8	0,10	0,80
9	0,10	0,90
10	0,10	1,00

Si tratta della distribuzione UNIFORME, in cui $P(x_i) = \frac{1}{m}$ per ogni $i = 1, \dots, m$.

Altri modelli matematici sono stati sviluppati per rappresentare diversi fenomeni discreti tipici delle scienze sociali e naturali e ricorrenti in campo medico ed in economia. Uno dei più importanti è il modello BINOMIALE. Anche in questo caso la presentiamo partendo da un esempio.

Esempio 3. Il sistema informativo di una casa farmaceutica ha il compito di revisionare gli ordini di vendita. Gli ordini giudicati scorretti vengono segnalati all'attenzione della società. Si stima pari a 0,1 la probabilità che un ordine venga giudicato insoddisfacente da parte del sistema informativo. Sulla base di questa informazione la società vuole calcolare la probabilità che si abbia un certo numero di segnalazioni in un dato campione di ordini di vendita.

Per esempio, se in un giorno vengono realizzati 4 ordini di vendita, qual è la probabilità che nessuno di questi venga giudicato insoddisfacente da parte del sistema informativo?

La distribuzione binomiale è una funzione di probabilità discreta, caratterizzata da quattro essenziali proprietà.

Le osservazioni da una distribuzione binomiale possono essere ottenute con due diversi metodi di campionamento: *un campionamento da popolazione infinita senza reimmissione* oppure *un campionamento da popolazione finita con reimmissione*.

Ciascuna osservazione può essere classificata in due categorie incompatibili ed esaustive, chiamate per convenzione *successo* ed *insuccesso*.

La probabilità di ottenere un successo, p , è costante per ogni osservazione, così come la probabilità di ottenere un insuccesso, $1-p$.

Il risultato di un osservazione, successo od insuccesso, è indipendente dal risultato di qualsiasi altra.

La distribuzione binomiale è la legge della variabile aleatoria che rappresenta “il numero dei successi ottenuti in un campione di n osservazioni”.

Tornando all'esempio, un ordine viene classificato come *successo* se viene segnalata la presenza di errori. Quindi in un campione di 4 ordini il numero dei successi può essere 0, 1, 2, 3 o 4. In generale, su n osservazioni il numero dei successi è un numero compreso fra 0 ed n .

Supponiamo di osservare il seguente risultato:

1° ORDINE	2° ORDINE	3° ORDINE	4° ORDINE
segnalato	segnalato	non segnalato	segnalato

Qual è la probabilità di ottenere questa particolare sequenza di successi ed insuccessi?

Le probabilità associate alle 4 osservazioni sono:

1° ORDINE	2° ORDINE	3° ORDINE	4° ORDINE

Poiché le osservazioni sono indipendenti, la probabilità associata alla sequenza è:

$$p^3(1-p) = 0,1^3 \cdot 0,9 = 0,0009$$

Tuttavia questa rappresenta la probabilità di ottenere una sequenza con 3 successi ed un insuccesso in un campione di 4 osservazioni, ed è la stessa indipendentemente dall'ordine con cui si presentano. In quanti modi diversi possiamo ottenere 3 successi su 4 ordini?

Dobbiamo affidarci alla regola delle combinazioni.

COMBINAZIONI

Il numero di modi in cui si possono selezionare X oggetti da un campione di n unità è dato da:

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

Nel nostro esempio si ha:

$$\binom{4}{3} = \frac{4!}{3!(4-3)!} = \frac{4 \cdot 3!}{3!} = 4$$

Di conseguenza, la probabilità di osservare 3 ordini scorretti su 4 è pari a:

$$\begin{aligned} & \text{numero di possibili sequenze} \times \text{probabilità di una particolare sequenza} \\ & = 4 \times 0,0009 = 0,0036 \end{aligned}$$

Allo stesso modo possono essere derivate le probabilità degli altri 4 possibili risultati della v.a.: 0, 1, 2 e 4. Tuttavia all'aumentare di n questo calcolo diventa piuttosto laborioso e conviene far riferimento al modello rappresentato dalla distribuzione di probabilità della binomiale.

DISTRIBUZIONE BINOMIALE

$$P(X = x_i) = \frac{n!}{x_i!(n - x_i)!} \cdot p^{x_i} \cdot (1 - p)^{n - x_i}$$

dove

$P(X = x_i)$ = probabilità di ottenere x_i successi dati i parametri n e p ;

n = ampiezza campionaria;

p = probabilità di successo;

x_i = numero di successi nel campione (0, 1, 2, ..., n).

Tale modello non è altro se non la formalizzazione di quanto già derivato intuitivamente.

VALORE ATTESO DI UNA DISTRIBUZIONE BINOMIALE

$$\mu = n \cdot p$$

SCARTO QUADRATICO MEDIO DI UNA DISTRIBUZIONE BINOMIALE

$$\sigma = \sqrt{n \cdot p \cdot (1 - p)}$$

Nel nostro esempio il valore atteso e lo scarto quadratico medio della variabile che rappresenta il numero di ordini scorretti sono rispettivamente pari a 0,4 e 0,6.

VARIABILI CASUALI CONTINUE

Nel caso di fenomeni continui si può calcolare la probabilità che la v.a. assuma valori compresi in un intervallo, mentre la probabilità che assuma un particolare valore è zero. Questa caratteristica distingue i fenomeni che risultano da un processo di misurazione (come ad es. il tempo) da quelli ottenuti attraverso un processo di conteggio.

Il calcolo di probabilità, del valore atteso e dello scarto quadratico medio per un fenomeno continuo coinvolge espressioni matematiche che richiedono la conoscenza del calcolo integrale.

Una variabile casuale X continua è caratterizzata dall'esistenza di una funzione $f(x)$ continua, detta **FUNZIONE DI DENSITA'**, che gioca un ruolo analogo alla funzione di probabilità nel caso discreto, con la differenza fondamentale che $f(x)$ non è interpretabile come una probabilità.

Una funzione di densità $f(x)$ deve soddisfare le seguenti proprietà:

1. $f(x) \geq 0$ per ogni x ;

$$2. \int_{-\infty}^{+\infty} f(x) dx = 1$$

Analogamente al caso discreto, possiamo definire la **FUNZIONE DI RIPARTIZIONE**, il **VALORE ATTESO** e la **VARIANZA** di una variabile casuale continua. Dobbiamo però sostituire l'integrale alla sommatoria...

FUNZIONE DI RIPARTIZIONE

$$F(x) = P(X \leq x) =$$

a differenza di $f(x)$ è una probabilità.

VALORE ATTESO

$$E(X) =$$

VARIANZA

$$V(X) = E[X - E(X)]^2 =$$

Si verifica facilmente che $V(X) = E(X^2) - \mu^2$

Il valore atteso, la varianza e lo scarto quadratico medio soddisfano le stesse proprietà enunciate nel caso di variabili statistiche.

LA DISTRIBUZIONE NORMALE

E' sicuramente la distribuzione più importante nell'ambito di tutta la statistica. E' rappresentata da una FUNZIONE DI DENSITA' DI PROBABILITA' CONTINUE, in quanto nasce da un processo di misurazione del fenomeno di interesse.

La distribuzione normale è importante in statistica per tre motivi fondamentali:

- 1. diversi fenomeni continui sembrano seguire, almeno approssimativamente, una distribuzione normale;**
- 2. la distribuzione normale può essere utilizzata per approssimare numerose distribuzioni di probabilità discrete;**
- 3. la distribuzione normale è alla base dell'*inferenza statistica classica* in virtù del *teorema del limite centrale*.**

PROPRIETA' DELLA DISTRIBUZIONE NORMALE

- 1. ha una forma campanulare e simmetrica;**
- 2. le sue misure di posizione centrale (valore atteso, mediana, moda) coincidono;**
- 3. ha due punti di flesso in $\mu - \sigma$ e $\mu + \sigma$ (è concava nell'intervallo compreso fra questi punti e convessa altrove);**
- 4. assume valori compresi fra $-\infty$ e $+\infty$:**

$$\lim_{x \rightarrow -\infty} f(x) = \lim_{x \rightarrow +\infty} f(x) = 0.$$

La funzione di densità di probabilità della normale è data dalle seguente espressione:

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{1}{2}[(x-\mu)/\sigma]^2}$$

dove

e = costante matematica approssimata da 2,71828;

π = costante matematica approssimata da 3,14159;

μ = valore atteso della popolazione;

σ = scarto quadratico medio della popolazione;

x = valori assunti dalla variabile aleatoria, con $-\infty < X < +\infty$.

Essendo e e π delle costanti matematiche, le probabilità di una distribuzione normale dipendono soltanto dai valori assunti dai due parametri μ e σ .

L'uso delle funzione di densità vista per il calcolo di probabilità relative ad una v.a. con distribuzione normale è estremamente complesso. Per questo motivo solitamente si ricorre all'impiego di tavole in grado di fornirci le probabilità desiderate. Tuttavia, poiché il numero di combinazioni dei parametri μ e σ è infinito, conviene trasformare le osservazioni, *standardizzando* la v.a. normale, per poi ricavare le probabilità desiderate dalle tavole della distribuzione normale standardizzata.

LA STANDARDIZZAZIONE

Chiamiamo Z la variabile ottenuta sottraendo alla variabile X il suo valore atteso e rapportando il risultato allo scarto quadratico medio:

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

La v.a. Z ha valore atteso nullo e scarto quadratico medio pari a 1:

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

Esempio 4: Il responsabile dell'assemblaggio in una società che fabbrica automobili si propone di ridurre il tempo necessario. Nell'attuale processo i lavoratori sono addestrati individualmente. Dopo aver raccolto informazioni il responsabile stabilisce che il tempo di assemblaggio segue approssimativamente una distribuzione normale con valore atteso pari a 75 secondi e scarto quadratico medio pari a 6 secondi.

Supponiamo che il responsabile dell'assemblaggio voglia determinare la probabilità che un addetto scelto a caso impieghi un tempo compreso fra 75 ed 81 secondi per assemblare il pezzo. Come si procede?

1. Standardizzo la variabile.

2. Cerco il valore della probabilità nella tavola della normale standardizzata.

(Ris. 0,3413)

Altre possibili domande:

- 1. Qual è la probabilità che un addetto selezionato a caso impieghi al massimo 81 secondi per completare il lavoro? In questo caso guardo la tavola delle probabilità cumulate. (Ris. 0,8413)**
- 2. Quanto tempo deve passare perché il 10% degli addetti abbia completato l'assemblaggio? Dobbiamo individuare il valore di Z che lascia alla sua sinistra una massa (probabilità) pari a 0,1, e ricordare che da Z dobbiamo tornare a $X = \mu + Z\sigma \dots$ (Ris. $Z=-1,28$, $X=67,32$).**

La DISTRIBUZIONE T DI STUDENT nasce dalla necessità di stimare la media di una popolazione normale quando σ non è noto.

La distribuzione t di Student ha una forma molto simile alla normale standardizzata. Entrambe sono unimodali, simmetriche ed hanno media normale standardizzata. Tuttavia l'area sottesa alla distribuzione t sulle code è maggiore di quella che caratterizza la distribuzione normale, poiché σ non è noto e viene stimato attraverso S : l'incertezza sul valore di σ fa sì che i valori di t presentino una maggiore variabilità dei valori di Z .

La sua funzione di densità dipende infatti da una costante caratteristica r , intero positivo, detta “numero di gradi di libertà”. Ha valore atteso nullo e varianza pari a $\frac{r}{r-2}$.

All'aumentare dei gradi di libertà la distribuzione t converge alla distribuzione normale: all'aumentare dell'ampiezza del campione S diventa uno stimatore sempre più affidabile di σ .

Anche la funzione di densità della variabile casuale CHI-QUADRATO dipende da una costante caratteristica r detta “numero dei gradi di libertà”.

Ha la particolarità di assumere solo valori maggiori o uguali a 0. Ha una forma campanulate ma non simmetrica, che si appiattisce al crescere dei gradi di libertà.

Si dimostra che nel caso del chi-quadrato $E(X)=r$ e che $V(X)=2r$.

Questa distribuzione viene utilizzata nel caso di: costruzione di intervalli di confidenza e verifica di ipotesi su varianze; test di indipendenza fra variabili casuali; test di adattamento di distribuzioni empiriche a distribuzioni teoriche.

VARIABILI ALEATORIE BIDIMENSIONALI

Una variabile aleatoria bidimensionale (X,Y) è una funzione che associa ad ogni elemento ω di Ω una coppia di numeri reali (x,y) , dove $x = X(\omega)$ e $y = Y(\omega)$.

Se tutti gli elementi della variabile aleatoria bidimensionale sono discreti si può definire la **FUNZIONE DI PROBABILITA' CONGIUNTA** $P(x_i, y_h) = P(X = x_i, Y = y_h)$ che soddisfa le stesse proprietà viste nel caso unidimensionale.

Se tutti gli elementi della variabile aleatoria bidimensionale sono continui si può definire la **FUNZIONE DI DENSITA' CONGIUNTA** $f(x, y)$ che soddisfa le stesse proprietà viste nel caso unidimensionale.

Ciò che interessa in questo caso è la **COVARIANZA** fra le due variabili aleatorie X e Y , che si definisce come:

COVARIANZA

$$Cov(X, Y) =$$

$$\text{Si verifica facilmente che } Cov(X, Y) = E(XY) - E(X) \cdot E(Y)$$

Vediamo come si calcola nel caso discreto:

$$Cov(X,Y)=$$

E nel caso continuo:

$$Cov(X,Y)=$$

**ESEMPIO DI VARIABILE ALEATORIA DISCRETA
BIDIMENSIONALE**

Valori di Y	Valori di X		
	0	1	
0	0,1	0,2	0,3
1	0,4	0,3	0,7
	0,5	0,5	

Si calcoli la covarianza fra X e Y.

INDIPENDENZA FRA VARIABILI CASUALI

Due variabili aleatorie X e Y sono indipendenti se e solo se la funzione di probabilità (se sono discrete) o di densità (se sono continue) congiunta può essere fattorizzata come segue:

$$P(x_i, y_h) = P(x_i) \cdot P(y_h)$$

$$f(x, y) = f(x) \cdot f(y)$$

Infatti, si può notare che quando ciò è vero la covarianza si annulla:

$$Cov(X, Y) = \sum_{i=1}^r \sum_{h=1}^s (x_i - \mu_X)(y_h - \mu_Y)P(x_i, y_h)$$

=

SOMMA (E DIFFERENZA) DI VARIABILI CASUALI

Date due variabili aleatorie X e Y ne consideriamo la somma o la differenza. Vediamo come sono il loro valore atteso e la varianza di tale somma, ricordando le proprietà degli operatori media e varianza già viste.

$$E(X \pm Y) = E(X) \pm E(Y)$$

$$V(X \pm Y) = V(X) + V(Y) \pm Cov(X, Y)$$

Quindi, se X e Y sono indipendenti:

$$V(X \pm Y) = V(X) + V(Y)$$