

Supponiamo ora che la variabile casuale discreta X debba assumere uno dei valori x_1, x_2, \dots, x_k e che le probabilità associate a quei valori siano $f(x_1), f(x_2), \dots, f(x_k)$.

Allora il valore atteso della variabile X è definito dalla formula

$$(1) E[X] = \sum_{i=1}^k x_i f(x_i) \quad \text{"E" sta per "Expected value" = "Valore atteso"}$$

Nell'esempio appena visto la variabile X è la somma vinta i cui valori possibili sono

$x_1 = 0, x_2 = 1, x_3 = 2, x_4 = 3$ e le probabilità corrispondenti a quei valori sono

$$f(x_1) = \frac{1}{8}, f(x_2) = \frac{3}{8}, f(x_3) = \frac{3}{8}, f(x_4) = \frac{1}{8}$$

Ora supponiamo che il gioco considerato venga modificato in modo da vincere la somma $g(x_i)$ invece di x_i quando si ottiene il valore x_i , ad esempio per $g(x)$ si potrebbe scegliere la funzione $g(x) = x^2$, il che significa che si vincerebbe in dollari il quadrato di teste uscite, allora il valore atteso nel gioco così modificato sarebbe dato dalla formula

$$(2) E[g(X)] = \sum_{i=1}^k g(x_i) f(x_i)$$

Nelle formule (1) e (2) si assume che ci sia soltanto un numero finito di possibili valori della variabile casuale X , per limare questa limitazione si introduce la seguente definizione più generale applicabile a qualsiasi variabile casuale discreta.

Definizione

Il valore atteso della funzione $g(X)$ della variabile casuale discreta X , la cui funzione di densità è f , è dato da

$$(3) E[g(X)] = \sum_{i=1}^{\infty} g(x_i) f(x_i)$$

Si intende in questa definizione che x_1, x_2, \dots sono i possibili valori di X , e $f(x_1), f(x_2), \dots$ sono le probabilità corrispondenti, il valore atteso della variabile casuale X viene chiamato di solito **la media della variabile casuale** o **la media della distribuzione della variabile casuale X** .

Se la variabile X può assumere soltanto un numero finito, ad esempio k , di possibili valori, l'indice superiore nella somma al secondo membro della (3) sarà naturalmente k anziché ∞ .

Quando la variabile X possiede un numero infinito di possibili valori con probabilità positive è necessario assumere che la serie infinita al secondo membro della (3) converga.

Momenti

I momenti della distribuzione di una variabile casuale discreta X si possono definire in termini di valori attesi come segue:

Definizione

Definiamo il momento d'ordine k dall'origine della distribuzione della variabile casuale discreta X , la cui funzione di densità è f ed è data da

$$(4) \mu_k' = E[X^k] = \sum_{i=1}^{\infty} x_i^k f(x_i)$$

Il momento d'ordine k di una distribuzione è anche comunemente chiamato il momento d'ordine k della variabile casuale X avente si intende quella distribuzione.



Così si può parlare di μ'_k come del momento d'ordine k della variabile X o come del momento d'ordine k della distribuzione X .

Il momento del primo ordine $\mu'_1 = E[X]$ ricorre così spesso che ad esso si assegna il simbolo particolare μ . Poiché anche i momenti dalla media sono usati ampiamente, essi pure necessitano di essere definiti.

Definizione

Il momento d'ordine k dalla media della distribuzione della variabile casuale discreta X , la cui densità è f , è dato da

$$(5) \mu_k = E[(X - \mu)^k] = \sum_{i=1}^{\infty} (x_i - \mu)^k f(x_i)$$

I momenti di una distribuzione aiutano molto a descrivere la distribuzione quando la funzione di densità non è disponibile. Una gran parte dei problemi che implicano i momenti sarà interessata soltanto ai primi due momenti, perché essi spesso sono sufficienti a descrivere due utili proprietà della distribuzione e cioè dove la distribuzione è centrata e in che grado la distribuzione è concentrata attorno a questo centro.

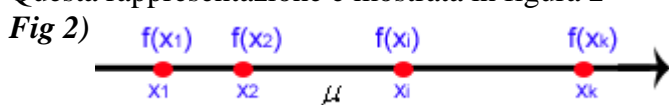
Il momento del primo ordine dall'origine μ si usa per determinare dove la distribuzione è centrata e il momento del secondo ordine dalla media μ_2 si usa per determinare il grado di concentrazione della distribuzione attorno alla media μ .

Poiché anche **il momento del secondo ordine** dalla media si usa molto spesso a questo proposito esso si indica con il simbolo particolare $\mu_2 \rightarrow \sigma^2$ e **viene chiamato varianza della distribuzione**.

La radice quadrata positiva della varianza cioè σ **viene chiamata scarto quadratico medio o deviazione standard della distribuzione**. Esso si impiega al posto della varianza quando si vuole una misura della concentrazione nelle stesse unità di misura della variabile casuale.

Per esaminare in che modo μ e σ^2 aiutano a descrivere una distribuzione di probabilità discreta le probabilità $f(x_1), f(x_2), \dots, f(x_k)$ associate ai valori x_1, x_2, \dots, x_k della variabile casuale X siano rappresentate geometricamente come masse puntiformi la cui somma è uguale a 1 localizzate nei punti x_1, x_2, \dots, x_k sull'asse x .

Questa rappresentazione è mostrata in figura 2



Il momento del primo ordine dall'origine μ allora, per come è definito, rappresenta il centro di gravità di queste masse puntiformi e perciò può servire come una misura di dove è localizzata o centrata la distribuzione.

Il momento del secondo ordine dalla media, cioè la varianza σ^2 tenderà ad essere piccolo se la maggior parte delle masse di probabilità sono concentrate vicino alla media μ e non vi sono masse localizzate molto lontano da μ . Tanto più disperse sono le masse attorno alla media tanto più grande è verosimile che diventi σ^2 , così la varianza di una distribuzione può servire come una misura del grado in cui la distribuzione è concentrata attorno alla media. E' facile tuttavia dare esempi di distribuzioni per cui la varianza è una misura insufficiente della concentrazione della distribuzione attorno alla sua media. Ciò ostante per la maggior parte delle distribuzioni più comuni essa assolve il suo compito in modo soddisfacente.



Come esempio in cui la varianza serve a misurare la concentrazione, consideriamo le due distribuzioni di probabilità mostrate nelle figure seguenti, in cui le probabilità sono rappresentate da segmenti verticali:

Fig. 3)

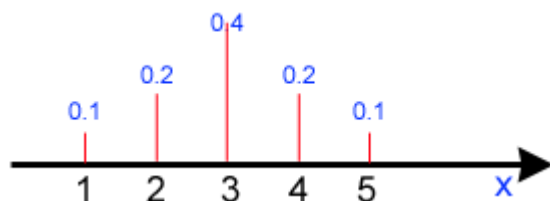
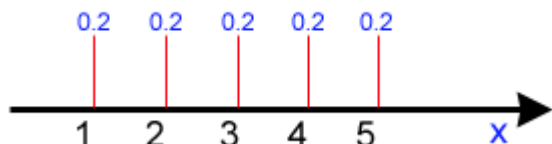


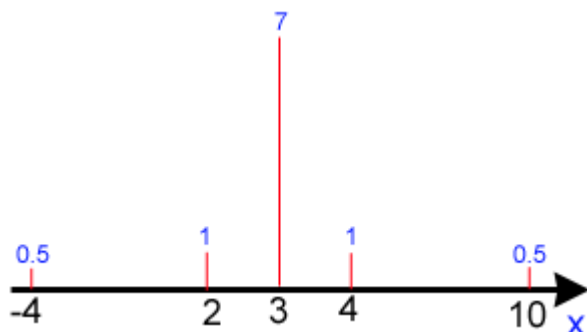
Fig. 4)



Per entrambe le distribuzioni applicando la formula che definisce i momenti dall'origine, cioè la (4) per $k=1$, si ottiene che $\mu = 3$. Calcoli basati sui momenti della media, cioè la formula (5) per $k=2$, si ottiene rispettivamente $\sigma^2 = 1,2$ e $\sigma^2 = 2,0$ mostrando quindi che la prima distribuzione è più altamente concentrata attorno alla sua media rispetto alla seconda.

Come esempio, invece, in cui il valore della varianza non sembra indicare la natura della concentrazione della distribuzione, attorno alla sua media consideriamo la distribuzione mostrata in figura 5

Fig. 5)



I calcoli danno ancora $\mu = 3$ e $\sigma^2 = 5,1$, valore quest'ultimo che confrontato con quelli ottenuti per le fig. 3, 4 indicherebbe scarsa concentrazione della distribuzione attorno alla sua media. Mentre, invece, la maggior parte della distribuzione è concentrata attorno alla media stessa. Si può notare quindi che ponendo una massa di probabilità molto piccola sufficientemente lontano dal centro della distribuzione, il valore della varianza si può rendere tanto grande quanto si vuole. Nonostante le limitazioni della varianza come misura della concentrazione di una distribuzione attorno alla sua media come mostrato nell'esempio di figura 3, la media e la varianza sono dimostrate quantità molto utili nel trattare le distribuzioni di probabilità più comunemente usate.

