VARIABILI ALEATORIE

Viene lanciata una moneta. Se esce testa vinco 100 euro, se esce croce non vinco niente. Quale è il valore della mia vincita?

Osserviamo che il valore della vincita dipende dal risultato dell'esperimento "lancio della moneta". Se esce testa, la vincita è 100, mentre se esce croce la vincita è 0. Ciò significa che la vincita non è un numero determinato, ma una funzione che assume un certo valore in corrispondenza dell'uscita testa e un altro valore in corrispondenza dell'uscita croce:

V(testa)$=$ 100

V(croce)$=$ 0,

dove V sta per vincita.

VARIABILI ALEATORIE

Possiedo 100 euro e decido di giocarli alla roulette puntandoli sul numero 7. Quale è il valore del mio capitale dopo questa operazione? Quale è il valore della mia vincita?

Naturalmente il valore del capitale non è noto a priori, ma dipende dal numero uscito alla roulette. Possiamo descrivere questa situazione con l'insieme Ω delle possibili uscite nel gioco della roulette, che sono i numeri tra 0 e 36:

$$\Omega = \{0, 1, 2, ..., 36\}$$

In caso di vittoria, e qundi se esce 7, il capitale sarà $100 \cdot 36 = 3600$. In caso di sconfitta, cioè se esce uno qualunque degli altri numeri, il capitale sarà 0. Il valore del capitale è quindi una funzione definita su Ω a valori reali che ad ogni numero diverso da 7 associa 0 e a 7 associa 3600.

Indicando tale funzione con C si ha:

$$C : \Omega \to \mathbf{R}$$

con $C(i) = 0 \quad se \quad i \neq 7$

$C(i) = 3600 \quad se \quad i = 7$

VARIABILI ALEATORIE

Viene lanciato un dado, se il numero che compare sulla faccia superiore del dado è 1 mi vengono consegnati 100 euro, altrimenti me ne vengono consegnati 10. Quanti euro mi vengono consegnati dopo il lancio del dado?

Anche questa volta il capitale che mi viene consegnato non è noto a priori ma dipende dal numero che è uscito sulla faccia superiore del dado. Questa volta

$$\Omega = \{0, 1, 2, 3, 4, 5, 6\}$$

e il capitale C vale $C(i) = 100 \quad se \quad i = 1$

$C(i) = 10 \quad se \quad i \neq 1.$

Giochiamo un euro sull'uscita di un ambo sulla ruota di Torino. Se l'ambo esce viene pagata 250 volte la posta, cioè 250 euro, se l'ambo non esce perdo semplicemente l'euro che ho puntato inizialmente. Quanto vinco?

La vincita vale -1 se l'ambo non esce e vale 249 se l'ambo esce.

Due giocatori giocano a testa o croce con una moneta equilibrata. Vince chi per primo totalizza due vincite. Quanto dura la partita?

VARIABILI ALEATORIE

Una variabile aleatoria si indica usualmente con una lettera maiuscola, per esempio C, V o X, e i suoi possibili valori (ricordiamo che abbiamo fatto l'ipotesi che Ω sia finito) con $x_1, x_2, ...x_n$.

Se x è un numero reale, si indica con :

$(X = x) = \{\omega \in \Omega \text{ tali che } X(\omega) = x\}$.

$(X \leq x) = \{\omega \in \Omega \text{ tali che } X(\omega) \leq x\}$.

Naturalmente :

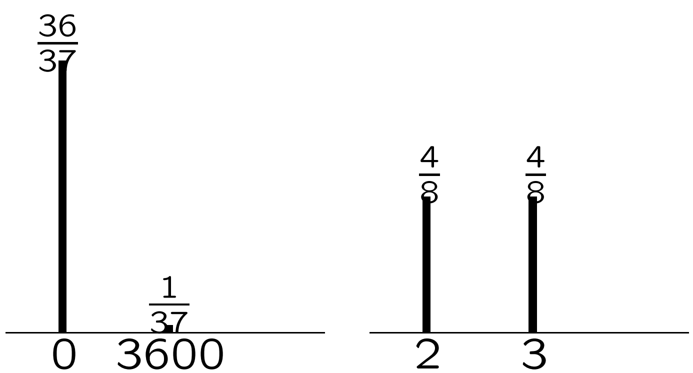
$A_1 = (X = x_1), A_2 = (X = x_2), ..., A_n = (X = x_n)\}$

sono eventi .

Indichiamo con $p_1, p_2, ..., p_n$ le rispettive probabilità. Naturalmente $p_1 + p_2 + ... + p_n = 1$, infatti $p_1 + p_2 + ... + p_n$ è la probabilità che X assuma uno dei suoi possibili valori, cioè la probabilità dell'evento certo.

VARIABILI ALEATORIE

La legge o distribuzione di una variabile aleatoria è l'insieme formato dai suoi valori $x_1,\ x_2, ...\ x_n$ **e le corrispondenti probabilità** $p_1, p_2, ...p_n$. Rappresentiamo nel modo seguente le leggi delle variabili aleatorie del primo e del quarto esempio:



VARIABILI ALEATORIE

La legge di una variabile aleatoria di fatto induce una misura di probabilità sull'insieme

$$\{x_1, x_2, ...x_n\},$$

definita come $p_X(E) = p\{\omega \in \Omega \quad t.c. \quad X(\omega) \in E\}$.

In sostanza quindi una v.a. "trasporta" la struttura di spazio di probabilità da Ω a **R**

Due variabili aleatorie che hanno la stessa legge si dicono **equidistribuite. Si dice speranza matematica o media o valore atteso** della variabile aleatoria X il numero

$$E(X) = x_1 \cdot p_1 + x_2 \cdot p_2 + ...x_n \cdot p_n.$$

Ovviamente variabili aleatorie equidistribuite hanno la stessa media. La media di una v.a. non è altro che una "media pesata" dei valori che essa ssume, dove i pesi sono le probabilità con cui un valore viene assunto. Dimostrare che se a è un numero reale

$$E(a \cdot X) = a \cdot E(x)$$

Infatti se X assume i valori $x_1, x_2, ..., x_n$ con probabilità rispettivamente $p_1, p_2, ..., p_n$, aX assume i valori $ax_1, ax_2, ..., ax_n$ con le stesse probabilità e dunque:

$$E(aX) = p_1ax_1 + p_2ax_2 + ... + p_nax_n = a(p_1x_1 + p_2x_2 + ... + p_nx_n) = aE(X)$$

VARIABILI ALEATORIE

Calcoliamo la media della variabile aleatoria del secondo esempio:

$$E(C) = 100 \cdot 1/6 + 10 \cdot 5/6 = 150/6 = 25$$

Calcoliamo la media della variabile aleatoria del terzo esempio: La probabilità di indovinare un ambo su una ruota giocando due numeri (si potrebbe calcolare) è circa $\frac{1}{400}$, dunque la media è:

$$249 \cdot 1/400 + (-1) \cdot 399/400 = -0,375$$

Osserviamo che non giocando si ha una vincita certa che vale 0 e che quindi ha media 0. Calcoliamo ora la media della variabile aleatoria del quarto esempio:

$$E(Y) = 2 \cdot 4/8 + 3 \cdot 4/8 = 5/2.$$

Osserviamo che in generale la media di una variabile aleatoria non coincide con alcuno dei valori che la variabile aleatoria assume, come accade in tutti gli esempi che abbiamo studiato. **Si dice varianza della variabile aleatoria X il numero**

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

Si può calcolare la varianza con la formula seguente:

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

La varianza misura quanto una variabile aleatoria è dispersa intorno alla sua media.

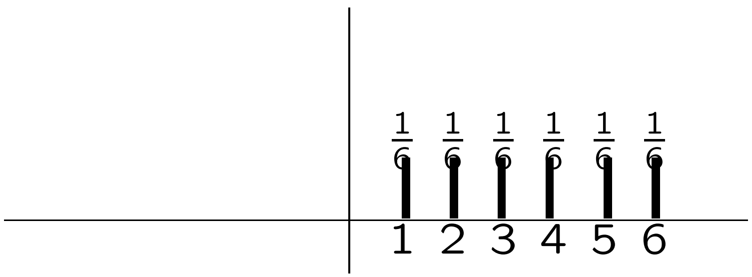
Se una variabile aleatoria è costante, la sua media è uguale al valore costante assunto dalla variabile aleatoria e la sua varianza è uguale a 0. Infatti se X è costante, assume solo un valore x_1 con probabilità 1 e dunque $E(X) = x_1$, X^2 assume solo il valore x_1^2 con probabilità 1 e dunque $V(X) = E((X - E(X)^2) = x_1^2 - x_1^2$. Calcoliamo la varianza delle variabili aleatorie del primo e del quarto esempio

$V(X) = \frac{(3600)^2}{37} - (\frac{3600}{37})^2$

$V(Y) = \frac{1}{4}$

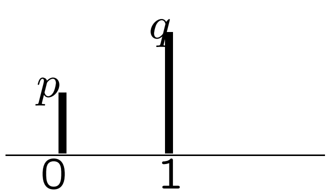
VARIABILI ALEATORIE

Lancio un dado e considero la v.a. definita su $\Omega = \{1, 2, ..., 6\}$ con $X(i) = i$, cioè X è semplicemente l'uscita del dado. Si ha



VARIABILI ALEATORIE

Lancio una moneta non necessariamente equilibrata, in cui cioè p(T)=p e p(C)=q=1-p e considero la variabile aleatoria X definita su $\Omega = \{T, C\}$ con $X(T) = 1$ e $X(C) = 0$. In questo caso



VARIABILI ALEATORIE

Distribuzioni di v.a. su spazi finiti L'esempio più semplice di distribuzione di una v.a. è la cosidetta distribuzione di Bernoulli. Abbiamo già visto questa distribuzione studiando il lancio di una moneta non necessariamente equilibriata e la variabile aleatoria X definita su $\Omega = \{T, C\}$ come $X(T) = 1$ e $X(C) = 0$, in cui X rappresenta il numero delle teste osservate dopo il lancio. Una variabile aleatoria X che assume solo i valori 0 e 1 con $p(X = 1) = p$ si dice che ha una

Distribuzione di Bernoulli di parametro p

e si indica con $B(p)$.

Siamo interessati a calcolare la media e la varianza di una variabile aleatoria con distribuzione $B(p)$. Si ha:

$$E(X) = p \ , \ V(X) = E(X^2) - (E(X))^2 = p - p^2 = p(1-p)$$

VARIABILI ALEATORIE

Siamo ora interessati a studiare altre distribuzioni di probabilità su spazi finiti Lancio una moneta non necessariamente equilibrata, in cui cioè p(T)=p e p(C)=q=1-p n volte e considero come spazio degli eventi l'insieme di tutte le possibili n-ple di uscite, per esempio nel caso di un lancio avremo:

$$\Omega = \{T, C\}$$

nel caso di due lanci

$$\Omega = \{TT, TC, CT, CC\}$$

nel caso di tre lanci

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

e così via per il caso di più lanci. Osserviamo che nel caso di n lanci lo spazio Ω contiene esattamente 2^n elementi. Questa è la classica situazione di indipendenza di eventi nel senso che l'evento "esce testa al k-esimo lancio" e "esce testa all' h-esimo lancio" se $h \neq k$ sono eventi indipendenti. (Osserviamo che l'evento "esce testa al k-esimo lancio" è rappresentato dal sottoinsieme di Ω formato da tutte le sequennze che hanno testa al k-esimo posto.) L'indipendenza ci permette di calcolare la probabilità dell'intersezione degli eventi come il prodotto delle probabilità dei singoli eventi. Per esempio nel caso dei 3 eventi la probabilità che esca la sequenza TCT è uguale al prodotto della probabilità che esca la prima volta testa per la probabilità che esca la seconda volta croce per la probabilità che esca la terza volta testa cioè $p \cdot q \cdot p = p^2 \cdot q$.

Siamo ora interessati a sapere quanto vale la probabilità che esca due volte testa e una volta croce in un ordine qualsiasi, cioè alla probabilità dell'evento:

$$\{TCT, CTT, TTC\}.$$

Per simmetria e per l'additività della misura di probabilità abbiamo:

$$p(\{TCT, CTT, TTC\}) = 3p^2 \cdot q.$$

Vogliamo ora generalizzare questo risultato, cioè stabilire in n lanci consecutivi di una moneta:

1) qual'è la probabilità che esca una ben determinata sequenza

2) qual'è la probabilità che esca esattamente k volte testa e quindi $n - k$ volte croce, in un ordine qualsiasi.

Il primo problema si risolve osservando che gli eventi corrispondenti agli n lanci sono tra loro indipendenti e quindi la probabilità dell'intersezione è uguale al prodotto delle probabilità: quindi se nella sequenza ci sono k teste e $n-k$ croci la probabilità dell'evento è $p^k \cdot q^{n-k}$

Per il secondo è necessario contare quante sono le sequenze distinte che contengono k teste e $n-k$ croci. Questo è un risultato di calcolo combinatorio: Il numero delle sequenze di k teste e n-k croci è esattamente $\binom{n}{k} = \frac{n!}{k! \cdot (n-k)!}$ dove:

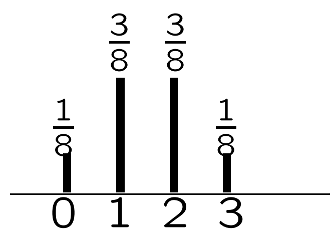
In definitiva la probabilità che escano esattamente k teste e $n-k$ croci risulta essere:

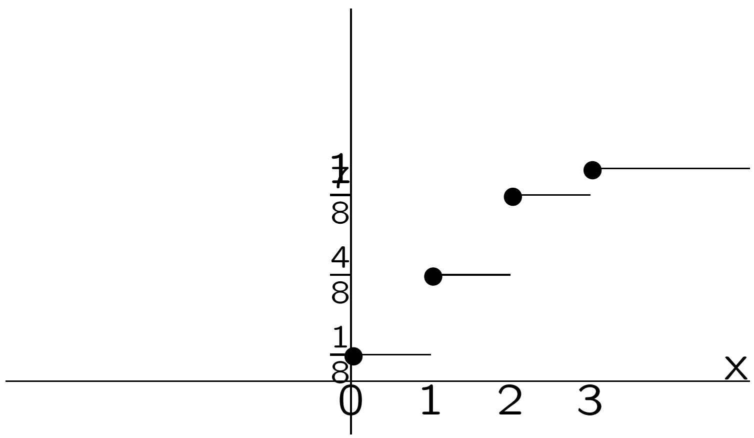
$$\frac{n!}{k! \cdot (n-k)!} p^k q^{n-k}$$

La variabile aleatoria X definita sullo spazio delle sequenze di n lanci di una moneta che associa a ogni lancio il numero delle teste uscite in quel lancio con

$$p(X=k) = \binom{n}{k} p^k q^{n-k}$$

si dice distribuzione binomiale di parametro p e si indica con $B(n,p)$.





La distribuzione di probabilità binomiale si adatta a descrivere tutte quelle situazioni in cui un esperimento viene ripetuto n volte" indipendentemente l'una dall'altra ". Facciamo alcuni esempi:

Supponiamo di avere un'urna contenente n palline di cui r bianche e n-r nere. Estraiamo una pallina e registriamo se è bianca o nera, la reimbussoliamo e procediamo a una nuova estrazione. Registriamo se la nuova estrazione ha dato una pallina bianca o nera e procediamo sempre in questo modo sino ad aver effettuato n estrazioni. La distribuzione di probabilità relativa al numero di palline bianche estratte è una binomiale con $p = \frac{r}{n}$ e $q = \frac{n-r}{n}$

La situazione analoga a quella dell'esempio precedente in cui la pallina non viene di volta in volta reimbussolata non dà luogo a una distribuzione binomiale, in quanto gli eventi corrispondenti alle successive estrazioni non sono più indipendenti.

Calcoliamo ora la media di una variabile aleatoria $X = B(n, p)$. Basta osservare che X è la somma di n variabili aleatorie bernoulliane di parametro p indipendenti, ciascuna delle quali ha media p. Poiché la media di una somma è la somma delle medie, si ha $E(X) = np$.