VARIABILI ALEATORIE DISCRETE E DISTRIBUZIONI DI PROBABILITA'

Def. Si dice **variabile aleatoria discreta** X una variabile che può assumere valori $X_1, X_2, ...X_n$ corrispondenti ad eventi aleatori $E_1, E_2, ...E_n$ non impossibili, che si escludono a vicenda e tali che sicuramente uno di essi si verifichi.

Ogni valore X_i che la variabile può assumere è correlato alla probabilità p_i dell'evento corrispondente E_i, cioè la probabilità che la variabile X assuma il valore X_i.

Def. Data una variabile aleatoria discreta X, con valori $X_1, X_2, ...X_n$, la successione delle probabilità $p_1, p_2, ...p_n$ ad essi associate si chiama **distribuzione di probabilità** della variabile X.

Esempio

Un'urna contiene 6 palline verdi e 4 gialle. Si estraggono 3 palline senza reimmissione. Si considera l'evento "uscita di una pallina verde". Esso può avvenire da 0 a 3 volte. Il numero di volte in cui esce una pallina verde è la variabile aleatoria X. Esso è una variabile casuale teorica perché ad esso possiamo associare un valore di probabilità calcolato teoricamente. Si può costruire una tabella che rappresenta la distribuzione di probabilità della variabile X.

X	*0*	*1*	*2*	*3*
P(X)	*1/30*	*3/10*	*1/2*	*1/6*

(verificare per esercizio la correttezza delle probabilità calcolate)

<u>Si noti che, poiché gli eventi si escludono a vicenda e uno di essi si deve sicuramente verificare, la somma delle probabilità è 1.</u> (condizione di normalizzazione)

Def. Si dice **funzione di ripartizione** di una variabile aleatoria X la funzione F(X) che fornisce la probabilità che X non assuma valore superiore ad un valore prefissato Xi.
Pertanto il suo valore si ottiene sommando le probabilità relative ai valori minori o uguali a Xi.

X	*X1*	*X2*	*X3…*	*Xn*
P(X)	*P1*	*P2*	*P3…*	*Pn*
F(X)	*P1*	*P1+P2*	*P1+P2+P3*	*P1+P2+…Pn*

Ritornando all'esempio precedente:

X	*0*	*1*	*2*	*3*
P(X)	*1/30*	*3/10*	*1/2*	*1/6*
F(X)	*1/30*	*1/3*	*5/6*	*1*

I valori di F(X) rappresentano la probabilità che la pallina verde esca "al massimo 0 volte", al massimo 1 volta", " al massimo 2 volte", "al massimo 3 volte".

GIOCHI ALEATORI

Una delle prime questioni considerate dai matematici agli albori della nascita della teoria della probabilità fu la possibilità di calcolare la probabilità di vincita e quindi la "convenienza" o meno per una persona di partecipare ad un gioco d'azzardo, considerata la somma investita e l'eventuale guadagno.
Tale convenienza è espressa dalla grandezza detta "**speranza matematica**".

Def. Si chiama "speranza matematica" M(X) della variabile X la somma dei prodotti della variabile X per le rispettive probabilità:

$$M(X) = X_1 \cdot p_1 + X_2 \cdot p_2 + ... + X_n \cdot p_n$$

Se M(X) = 0 si dice che il gioco è equo; se M(X) > 0 il gioco è favorevole (per il giocatore), se M(X) < 0 il gioco è sfavorevole.

ESEMPIO
In un gioco si lancia un dado. Se esce il numero 1, il giocatore vince 7 euro, se esce un numero pari 4 euro, in caso contrario deve pagare 2 euro.
Moltiplicando i valori delle vincite (positivi) e delle perdite (negativi) per le rispettive probabilità si ottiene un valore che rappresenta la vincita che il giocatore avrebbe potuto sperare di realizzare mediamente.

X	*7*	*4*	*-2*
P(X)	*1/6*	*1/2*	*1/3*

M(X) = 7/6 + 2 -2/3 = 5/2 = 2,5

In questo caso il gioco è favorevole al giocatore.
Se vogliamo calcolare quanto deve essere l'ammontare della posta pagata dal giocatore in caso di perdita, affinché il gioco sia equo, poniamo:

M(X) = 7/6 + 2 – P 1/3 = 0. Si trova P = 9,5. La tabella nel caso di gioco equo diventa pertanto:

X	*7*	*4*	*- 9,5*
P(X)	*1/6*	*1/2*	*1/3*

Se si organizza il gioco diversamente, si può chiedere una posta da pagare al giocatore per poter partecipare al gioco gestito da un "banco", prima che il gioco inizi.
Poiché il giocatore vince 7 euro se esce 1 e la posta relativa al gioco equo è 9,5, il giocatore incasserà una vincita lorda di 16,5 euro in caso esca il numero 1; poiché vince 4 euro se esce un numero pari, la vincita lorda in questo caso sarà 13,5.
(16,5 - 9,5 = 7, vincita netta; 13,5 – 9,5 = 4, vincita netta)
<u>La posta P che il giocatore deve pagare affinché il gioco sia equo è la somma dei prodotti delle vincite lorde per le rispettive probabilità di vincita:</u>

$$P = 16,5 \cdot \frac{1}{6} + 13,5 \cdot \frac{1}{2} = 9,5$$

La tabella è:

X	*16,5*	*13,5*	*0*
P(X)	*1/6*	*1/2*	*1/3*

<u>La posta P che il giocatore deve pagare, affinchè il gioco sia equo, è la speranza matematica della variabile "vincita lorda"</u>.

VALORI CARATTERIZZANTI UNA VARIABILE ALEATORIA DISCRETA

Def. Si chiama "**speranza matematica o valor medio**" M(X) della variabile X la somma dei prodotti della variabile X per le rispettive probabilità:

$$M(X) = X_1 \cdot p_1 + X_2 \cdot p_2 + ... + X_n \cdot p_n$$

Si noti che l'espressione equivale a:

$$M(X) = \frac{X_1 \cdot p_1 + X_2 \cdot p_2 + ... + X_n \cdot p_n}{p_1 + p_2 + ... p_n}$$

cioè alla media ponderata dei valori di X, essendo la somma delle probabilità uguale a 1.

<u>Teorema</u>
La speranza matematica di una variabile aleatoria costante è la costante stessa.

<u>Teorema</u>
Se S è una variabile aleatoria discreta e a è una costante, allora: M (aX) = a M(X).

(La probabilità di X coincide con la probabilità di aX)

<u>Teorema</u>
La speranza matematica della somma S = X + Y di due variabili aleatorie è pari alla somma delle speranze matematiche.

M (X + Y) = M(X) + M(Y)

<u>Teorema</u>
La speranza matematica del prodotto P = XY di due variabili aleatorie indipendenti è pari al prodotto delle due speranze matematiche.

M (XY) = M(X) M(Y)

Può accadere che due distribuzioni abbiano lo stesso valor medio, ma i singoli valori Xi abbiano distanze più o meno grandi dal valore medio, cioè la distribuzione presenti una maggiore o minore dispersione rispetto al valor medio. Si introduce quindi una nuova grandezza correlata a tali differenze.

Def. Si dice **deviazione o scarto o variabile aleatoria centrata** la variabile differenza : X – Mx.

Le probabilità p (X) = p (X –Mx) essendo Mx costante.

Si noti che: M (X –Mx) = M(X) – M (Mx) = Mx – Mx = 0.

<u>La speranza matematica di una variabile aleatoria centrata è nulla.</u>

Sostituire ad una variabile X la deviazione X - Mx equivale quindi a una traslazione di sistema di riferimento che porti il valor medio nell'origine degli assi.

Def. Si chiama **varianza** della variabile aleatoria X la speranza matematica del quadrato della deviazione X – Mx.

$$V(X) = M\left[(X - M_x)^2\right] \qquad \text{oppure} \qquad V(X) = \sum_1^n (X_i - M_x)^2 \cdot p_i$$

Si hanno le seguenti proprietà della varianza.
Siano X, Y variabili aleatorie e k una costante, allora:

$V(kX) = k^2 \cdot V(X)$

$V(X + k) = V(X)$

$V(X + Y) = V(X) + V(Y)$, se le variabili X, Y sono indipendenti.

Def. Si definisce **scarto quadratico medio o deviazione standard** della variabile aleatoria X, la radice quadrata della varianza.

$$\sigma(X) = \sqrt{M\left[(X - M_x)^2\right]} = \sqrt{\sum_1^n (X_i - M_x)^2 \cdot p_i}$$

<u>Teorema</u>

$$\boxed{V(X) = \sigma^2 = M(X^2) - (Mx)^2}$$

Il valor medio e la varianza hanno un'importante significato fisico: se associamo ai punti di ascissa $X_1, X_2, ...X_n$ le masse $p_1, p_2, \ldots p_n$, allora si può dimostrare che M rappresenta il centro di gravità del sistema, V il momento di inerzia.

DISTRIBUZIONI DI PROBABILITA' CLASSICHE

Distribuzione binomiale o di Bernoulli

Tale distribuzione riguarda il problema delle prove ripetute. Si ripete n volte una certa prova nelle stesse condizioni; ogni prova è indipendente dalle altre. Se p è la probabilità costante dell'evento A "successo", q = 1 – p è la probabilità dell'evento B "insuccesso". Al termine delle n prove, l'evento A potrà aver avuto luogo da 0 ad n volte e l'evento B avrà quindi avuto luogo da n a 0 volte. La legge di distribuzione è fornita dal teorema di Bernoulli delle prove ripetute.

<u>Teorema</u>

La probabilità che su n prove ripetute l'evento A si verifichi x ($x \le n$) volte è data dalla funzione di probabilità:

$$f(x) = P(x) = \binom{n}{x} p^x q^{n-x}$$

Riportiamo senza dimostrazione il seguente

Teorema

Se X è una variabile aleatoria bernoulliana, di ordine n e parametro p:

- il suo valor medio è: $M(X) = np$
- la sua varianza è: $V(X) = npq$
- lo scarto quadratico medio è: $\sigma = \sqrt{npq} = \sqrt{np(1-p)}$

Distribuzione multinomiale

Si ha una generalizzazione della distribuzione binomiale quando ogni prova dà luogo a più di due alternative, conducendo ad eventi $A_1, A_2, ...A_n$ indipendenti con probabilità associate $p_1, p_2, ...p_n$, costanti da prova a prova, e tali che la loro somma sia 1. Per determinare la probabilità che gli eventi $A_1, A_2, ...A_n$ si presentino $X_1, X_2, ...X_n$ volte su n prove, si fa ricorso al

Teorema

La probabilità dell'evento composto dato dal verificarsi di X_1 eventi A_1, X_2 eventi A_2 ecc. è data dalla legge

$$p = \frac{N!}{X_1!X_2!...X_N!} p_1^{X_1} p_2^{X_2} ... p_N^{X_N}$$

(Ex. Un'urna contiene 10 palline di cui 3 rosse, 2 bianche, 5 verdi. Calcolare la probabilità che estraendo successivamente 5 palline, con reimmissione, ne escano 2 rosse, 1 bianca e 2 verdi)

Distribuzione geometrica

Considerata una prova ripetuta molte volte nelle stesse condizioni, il numero di prove necessarie affinchè si verifichi per la prima volta l'evento A "successo" è una variabile aleatoria X detta "tempo di attesa dell'evento A" e la sua distribuzione si dice "distribuzione geometrica".

Si noti che:
l'evento A si verifica alla prima prova con probabilità p;
l'evento A si verifica alla seconda prova con probabilità pq;
l'evento A si verifica alla terza prova con probabilità pq^2 …
Quindi la probabilità che l'evento A si verifichi alla x-esima prova sarà:

$$p(x) = pq^{x-1}$$

E' facile notare che il rapporto $\frac{p(x+1)}{p(x)} = q$, per questo motivo la distribuzione viene detta geometrica.
Essendo q<1 si avrà che le probabilità sono sempre decrescenti: $p(x+1) < p(x)$ e, dopo un certo termine, divengono trascurabili.

Enunciamo i teoremi:

- $\sum_{1}^{\infty} p(x) = \sum_{1}^{\infty} p \cdot q^{x-1} = 1$ condizione di normalizzazione

- La speranza matematica della variabile X geometrica è $M(X) = \frac{1}{p}$

- La varianza della variabile X geometrica è : $V(X) = \frac{q}{p^2} = \frac{1-p}{p^2}$

<u>Teorema</u>
Data una variabile di legge geometrica X di parametro p, la variabile aleatoria X condizionata dall'evento X >k , è una variabile aleatoria di parametro p.

Ciò significa che la variabile "tempo di attesa del primo successo", anche se condizionata dal non avverarsi dell'evento per k volte, mantiene la stessa probabilità. La variabile X di legge geometrica è <u>senza memoria</u>. E' quindi erroneo il ragionamento di certi giocatori del lotto che scommettono su numeri che hanno accumulato un ritardo. La probabilità di uscita alla prossima estrazione è sempre la stessa.

Distribuzione ipergeometrica

Si consideri questa volta una serie di prove che siano effettuate contemporaneamente o in modo successivo e dipendente (ad esempio l'estrazione di palline da un'urna senza reimmissione).
Se si hanno N elementi di cui K con la caratteristica di poter verificare l'evento A (ad esempio, se A = estrazione di una pallina bianca, K = numero delle palline bianche), allora si può verificare facilmente che la probabilità di ottenere x successi (estrazione di x palline bianche) su n prove ripetute, pensata come rapporto tra casi favorevoli e casi possibili, è data dalla legge:

$$p(x) = \frac{\binom{K}{x}\binom{N-K}{n-x}}{\binom{N}{n}}$$

Tale funzione rappresenta la distribuzione di probabilità ipergeometrica di una variabile casuale discreta.

Si può dimostrare che per le variabili con distribuzione ipergeometrica valgono le seguenti:

$M(X) = np$ dove $p = \frac{K}{N}$, $q = 1 - p$

$$V(X) = npq\frac{N-n}{N-1}$$

$$\sigma(X) = \sqrt{npq\frac{N-n}{N-1}}$$

Distribuzione di Poisson

Sia X una variabile aleatoria discreta che può assumere i valori 0,1,2,3…e sia λ una costante assegnata. Si dice distribuzione di Poisson di parametro λ la distribuzione regolata dalla legge di probabilità:

$$p(x) = \frac{\lambda^x e^{-\lambda}}{x!}$$

Tale distribuzione si manifesta in molti fenomeni naturali, come numero di chiamate telefoniche ad un centralino, errori di stampa in un libro, particelle emesse da una sostanza radioattiva, ecc.

Teorema

Per una variabile aleatoria di Poisson si ha che:

$$M(x) = \lambda$$
$$V(x) = \sigma^2 = \lambda$$
$$\sigma = \sqrt{\lambda}$$

Relazione tra le distribuzioni di Bernoulli e di Poisson

La distribuzione di Poisson può esser vista come una buona approssimazione della distribuzione binomiale per piccoli valori di x, purchè p sia molto piccolo (eventi rari) e $\lambda = np$.

Teorema

$$\lim_{n \to +\infty} \binom{n}{x} p^x q^{n-x} = \frac{\lambda}{x!} e^{-\lambda}$$

Per valori elevati di n si può sostituire la distribuzione binomiale con quella di Poisson se $p<0,1$.
(Ad esempio se $n>50$ ed $np<5$)

Teorema di Bienaymé-Cebicev

Supponiamo di conoscere sperimentalmente i valori assunti da una variabile X senza conoscerne la distribuzione di probabilità. Vogliamo determinare qual è la probabilità che i valori X differiscano in valore assoluto dal valor medio almeno di un certo $\varepsilon > 0$.
A tale questione risponde il teorema seguente.

Teorema

Data una variabile casuale X con scarto quadratico medio σ, la probabilità che i valori assunti da X differiscano in valore assoluto dal suo valore medio M di almeno $\varepsilon > 0$ (con $\varepsilon > \sigma$) è data da:

$$P(|X - M| > \varepsilon) \leq \frac{\sigma^2}{\varepsilon^2}$$

Dimostrazione

Sappiamo che $\sigma^2 = V(X) = \sum_1^n (X_i - M_x)^2 \cdot p_i$.

Se trascuriamo tutti gli scarti minori in valore assoluto di ε, che supponiamo siano i primi m, avremo:

$\sigma^2 \geq \sum_{i=m+1}^{n} (X_i - M_x)^2 \cdot p_i$, quindi $\sigma^2 \geq \sum_{i=m+1}^{n} \varepsilon^2 \cdot p_i$

e infine $\sigma^2 \geq \varepsilon^2 \cdot \sum_{i=m+1}^{n} p_i$, dove la sommatoria rappresenta la probabilità cercata:

$P(|X - M| > \varepsilon) \leq \frac{\sigma^2}{\varepsilon^2}$ c.v.d.

Si può allora calcolare la probabilità dell'evento contrario, cioè che i valori assunti da X differiscano in modulo dal valor medio per un fattore minore di ε.

$$P(|X - M| < \varepsilon) > 1 - \frac{\sigma^2}{\varepsilon^2}$$

Utilizzando tale relazione si può determinare un intervallo di valori che una variabile X può assumere con una data probabilità.

Il teorema di Bienaymé-Cebicev permette di superare la legge empirica del caso con una formulazione astratta nota come "**Legge dei grandi numeri o teorema di Bernoulli**".

"Se un evento ha probabilità costante p di verificarsi ad ogni prova, la probabilità che il valore della frequenza relativa differisca in valore assoluto dalla probabilità per meno di un arbitrario ε > 0, piccolo a piacere, tende ad 1, cioè alla certezza, quando il numero delle prove tende all'infinito".

Dimostrazione (caso di una distribuzione binomiale)

Se *n* è il numero di prove, la variabile *X/n* rappresenta la frequenza del verificarsi dell'evento associato alla variabile *X*.

$$M\left(\frac{X}{n}\right) = \frac{1}{n} M(X) = \frac{1}{n} np = p$$

$$V\left(\frac{X}{n}\right) = \frac{1}{n^2} npq = \frac{pq}{n}$$

Applicando il teorema di Cebicev

$$P\left(\left|\frac{X}{n} - p\right| < \varepsilon\right) > 1 - \frac{pq}{n \cdot \varepsilon^2}$$

Poiché $\lim_{n \to \infty} \frac{pq}{n \cdot \varepsilon^2} = 0$, la probabilità che la differenza tra la frequenza e la probabilità teorica p dell'evento sia minore di un ε fissato tende ad 1, cioè alla certezza.