0.1 Variabili aleatorie / Random Variable

Una variabile aleatoria X è una funzione che, dato un qualunque numero, calcola la probabilità di ottenere quel numero. La variabile può essere discreta o continua: nel caso discreto vuol dire che la funzione può prendere solo valori interi (in $\mathbb{N}$, per esempio la variabile che descrive il lancio del dado è discreta, perchè i possibili valori che può assumere sono $\{1, 2.3..6\}$) invece le variabili continue sono quelle che prendono valori in $\mathbb{R}$ (per esempio, una variabile uniforme che pesca un numero a caso tra $[0, 1]$ è continua, perchè posso ottenere tutti i valori reali, compreo $\sqrt{2}$. Una variabile aleatoria è definita tramite la sua funzione di densità, che rappresenta la distribuzione di probabilità di tutti i possibili punti. Il valor medio di una variabile aleatoria (expected value) è il valore che mi aspetto di ottenere aritmeticamente dopo una serie di osservazioni sulla variabile (questo valore non per forza deve essere ottenuto, per esempio il valor medio del lancio di un dado è 3.5, che chiaramente non posso ottenere lanciando un dado) La varianza è il valor medio dei valori che si discostano dalla media (poi in formule sarà più chiaro), ovvero rappresenta come si discostano i valori dal valor medio (se, in un lancio casuale, ottengo i seguenti numeri $(0, 10, 5)$, $(5, 6, 4)$ la media è chiaramente 5 per tutti e due ma la varianza sarà più alta nella prima rilevazione che nella seconda, perchè i valori si discostano molto dalla media.). In un investimento finanziario, farò prevalere una variabile aleatoria con una varianza molto piccola, perchè vuol dire che tutti i possibili guadagni/perdite sono vicini al valor medio e quindi ”limito” il rischio. Vediamo ora quali variabili aleatorie avete studiato. L'importante per voi sarà ricordare le formule che determinano il valor medio, varianza e media delle variabili aleatorie (non dovete calcolarle ogni volta) e capire la consegna del testo. Dividiamo il prossimo paragrafo in variabili aleatorie discrete e quelle continue.

Variabili discrete

Il valor medio di una variabile discreta X si calcola nel seguente modo:

$$E[X] = \sum_{x=0}^{\inf} xP(X = x) \tag{1}$$

dove con $P(X = x)$ intendiamo la probabilità che la variabile X prenda il valore x. La varianza invece è definita nel seguente modo:

$$Var[X] = E[(X - E[X])^2] = E[X^2] - E^2[X] \tag{2}$$

.

Per i conti che dovrete fare voi, conoscendo la variabile aleatoria con cui state lavorando, saprete già quanto vale la varianza e il valor medio. Da cui potete calcolarvi agevolmente:

$$E[X^2] = Var[X] + E^2[X]$$

$$E[cX] = cE[X], \quad x \in \mathbb{R}$$

Nel raro caso questo non sia possibile, potete calcolare il valor medio di una qualsiasi funzione della variabile aleatoria con la seguente formula:

$$E[g(X)] = \sum_{x=0}^{\inf} g(x)P(X = x) \tag{3}$$

Vediamo alcuni esempi di variabili discrete che avete studiato.

Poisson

La variabile aleatoria di Posson si rappresenta con $X \approx Po(\lambda)$ rappresenta la probabilità di eventi che avvengono in maniera successiva e indipendente, supposto che mediamente se ne verificano λ in un determinato lasso di tempo. Per esempio, una Poisson può rappresentare la frequenza di passaggio di auto in una strada, e se questa variabile è una $Po(5)$ vuol dire che mi aspetto mediamente un passaggio di 5 maccchine all'ora (o al minuto, dipende dallo spazio in cui si lavora).

La funzione di densità di questa variabile è la seguente:

$$f(x) = \exp(-\lambda)\frac{\lambda^x}{x!}$$

dove per esempio, f(2) indica la probabilità che avvengano due avvenimenti in media. Dunque, è abbastanza immediato il fatto che, se $X \approx Po(\lambda)$:

$$E[X] = \lambda$$
$$Var[X] = \lambda$$
$$E[X^2] = \lambda + \lambda^2$$

Vale inoltre che se abbiamo n variabili aleatorie Poisson (λ), allora la variabile somma è ancora una Poisson ($n\lambda$)

Vedremo nelle applicazioni come usare queste formule e come queste formule descrivono completamente la distribuzione della variabile.

Geometrica

La variabile geometrica rappresenta la probabiità di primo successo in un qualsiasi gioco aleatorio. Indichiamo con $X \approx T(p)$ dove p indica la probabilità di vincere. Per esempio $P(T = 5)$ calcola la probabilità che alla quinta estrazione ottengo il primo successo. Vale la seguente formula:

$$P(T = k) = (1 - p)^{k-1}p$$

e ragionandoci dovrebbe essere chiara (è la probabilità che perdo le prime $k-1$ volte ma vinco alla $k - esima$. Le formule che legano valor medio e varianza sono:

$$E[T] = \frac{1-p}{p} - 1$$
$$Var[T] = \frac{1-p}{p^2}$$

$$E[T^2] = \frac{1-p}{p^2} + (\frac{1-p}{p} - 1)^2$$

Nelle applicazioni la Geometrica compare poco, però è utile averla sottomano.

Variabili Continue

Quello che le differenzia da quelle discrete, come già detto, è il fatto che possono prendere non solo valori interi, ma anche valori reali. Non ha nessun signifcato l'espressione $P(X = \sqrt{2})$ se $X \approx Po(2)$, perchè la Poisson può assumere solo valori in $\mathbb{N}$. Invece nel caso continuo la cosa ha senso, quindi le formule di modificano facendo un approssimazione delle serie con l'integrale (cioè la struttura è sempre la stessa ma si passa dalla serie all'integrale). Avremmo dunque che, se X è una variabile continua che prende valori in $[a, b]$ un intervallo di $\mathbb{R}$ (per esempio [0,1])

$$E[X] = \int_a^b xP(X = x)dx$$

$$E[X^2] = \int_a^b x^2 P(X = x)dx$$

$$E[g(X)] = \int_a^b g(x)P(X = x)dx$$

$$Var[X] = E[X^2] - E^2[X]$$

Gli integrali che possono capitare sono solo integrali di polinomi e al massimo di esponenziali; come piccolo promemoria bisogna ricordarsi che:

$$\int x^a dx = \frac{x^{a+1}}{a+1}$$

$$\int e^{f(x)} f'(x)dx = e^{f(x)}$$

Vediamo in dettaglio le più applicate negli esercizi.

Variabili uniformi

Le variabili uniformi, raprresentate da $X \approx R(a, b)$ sono le più usate e rappresenta la probabilità di ottenere un numero a caso compreso tra a e b. Dunque se $X \approx R(0, 5)$ allora la variabile simula l'estrazione a caso di un numero tra 0 e 5 , ma comprende tutti i valori reali, dunque anche $2\sqrt{2}, \exp(1)$ etc. Vale che la formula di densità è la seguente:

$$f(x) = \frac{1}{b-a}$$

chiaramente la probabilità di indovinare un numero è data da 1 diviso la lunghezza dell'intervallo (provare a farsi qualche disegno per capire meglio) Valgono inoltre le seguenti formule su varianza e valore medio.

$$E[X] = \frac{a+b}{2}$$
$$Var[X] = \frac{(b-a)^2}{12}$$
$$E[X^2] = \frac{(b-a)^2}{12} + \frac{a^2+b^2+2ab}{4}$$

Esponenziali

La variabile esponenziale esrime la durata della vita di un fenomeno aleatorio, che è indipendente rispetto al suo passato. Questa proprietà è abbastanza importante in probabilità e si può riassumere nella seguente frase: "se so dove mi trovo ora, il futuro è indipendente dal passato". Data un esponenziale $X \approx Exp(\lambda)$, abbiamo la seguente funzione di densità:

$$f(x) = \lambda \exp(-\lambda)$$

e le formule che servono nelle applicazioni sono:

$$E[X] = \frac{1}{\lambda}$$
$$Var[X] = \frac{1}{\lambda^2}$$
$$E[X^2] = \frac{2}{\lambda^2}$$
$$Mediana[X] = \frac{\log 2}{\lambda}$$

Un altra formula importante è l'applicazione della funzione di distribuzione, ovvero $P(X \leq x)$. Questa è la probabilità di ottenere valori più piccoli di x. Se nel caso discreto è poco importante, perchè assumendo solo valori interi si avrà $P(X \leq x) = P(X = x-1) + P(X = x-2).... + P(X = 0)$ nel caso continuo questo non va bene perchè assume inifiniti valori. Vale dunque che, se $X \approx Exp(\lambda)$

$$P(X \leq x) = 1 - \exp(-\lambda x)$$