CORSO DI STATISTICA (parte 2) - ESERCITAZIONE 1

Dott.ssa Antonella Costanzo

a.costanzo@unicas.it

Esercizio 1. Distribuzione di probabilità, funzione di ripartizione di una v.c. discreta

Il tasso di cambio dollaro/euro nei primi tre mesi del 2013 è caratterizzato dalla seguente distribuzione di probabilità:

X	P(X)
0.73	0.05
0.74	0.10
0.75	0.25
0.76	0.40
0.77	0.15
0.78	0.05

a) Verificare che P(x) è una distribuzione di probabilità
b) Costruire la funzione di ripartizione empirica
c) Qual è la probabilità che il tasso di cambio in un dato giorno nell'arco di questi tre mesi sarà almeno pari a 0.75?
d) Qualè la probabilità che il tasso di cambio sarà meno di 0.77?
e) Se i valori del tasso di cambio giornalieri sono tra loro indipendenti, qual è la probabilità che per 2 giorni consecutivi il tasso di cambio sia superiore a 0.75?
f) Calcolare il valore atteso e varianza di X

Sol.

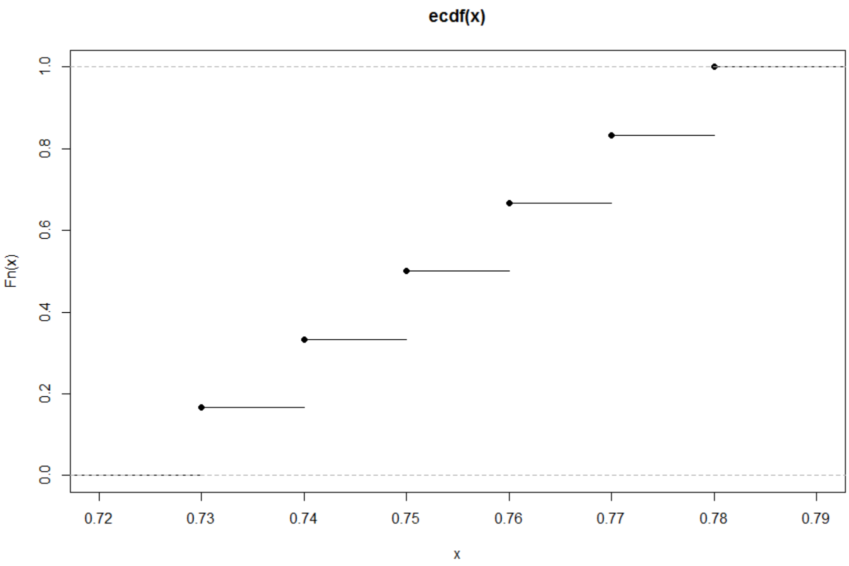
a) P(X) esprime la probabilità del verificarsi di ciascun valore di X, ossia P(X=x). Affichè una v.c. sia definita occorre che siano rispettati i postulati del calcolo della probabilità:

1. $P(X = x_i) = p_i \geq 0$

2. $\sum_i P(X = x_i) = \sum_i p_i = 1$

b) Funzione di ripartizione empirica $F(X) = P(X \leq x)$

X	X^2	P(X)	X^2P(X)	F(X)
0.73	0.5329	0.05	0.026	0.05
0.74	0.5476	0.10	0.054	0.15
0.75	0.5625	0.25	0.141	0.40
0.76	0.5776	0.40	0.231	0.80
0.77	0.5929	0.15	0.090	0.95
0.78	0.6084	0.05	0.030	1
			0.572	



c) Probabilità che il tasso di cambio nel periodo considerato sia almeno pari a 0.75?

$P(X \geq 0.75) = 1 - P(X \leq 0.75) = 1 - 0.40 = 0.60$ oppure

$P(X \geq 0.75) = P(X = 0.76) + P(X = 0.77) + P(X = 0.78) = 0.40 + 0.15 + 0.05 = 0.60$

d) probabilità che il tasso di cambio avrà un valore minore di 0.77

$P(X \leq 0.77) = 0.95$

e) Se i valori del tasso di cambio giornalieri sono tra loro indipendenti la probabilità che per 2 giorni consecutivi il tasso di cambio sia superiore a 0.75 è dato dal prodotto delle singole probabilità associate ai valori del tasso di cambio in corrispondenza delle due posizioni consecutive dopo il valore 0.75:

$P(X > 0.75) = P(X = 0.76) * P(X = 0.77) = 0.06$

f) Calcolare il valore atteso e varianza di X

$$\mu = E(X) = \sum_i xP(x)$$
$$= 0.73 * 0.05 + 0.74 * 0.10 + 0.75 * 0.25 + 0.76 * 0.40 + 0.77 * 0.15 + 0.78 * 0.05$$
$$= 0.7565$$

$$VAR(X) = E[(X - \mu)^2] = \sum_i (x_i - \mu)^2 p_i$$

Per calcolare la varianza si può utilizzare anche la seguente formula operativa:

$VAR(X) = E(X^2) - [E(X)]^2 = 0.572 - (0.7565)^2 \approx 0$

Esercizio 2. Valore atteso e varianza di trasformazioni lineari di una v.c.

La tabella seguente mostra la distribuzione di probabilità relativa al numero mensile di pezzi venduti da un impresa:

X=numero di pezzi venduti	P(x)
5000	0.2
6000	0.3
7000	0.2
8000	0.2
9000	0.1

Sapendo che la funzione di profitto di un'impresa a fronte della vendita dei suoi prodotti è $h(X) = 2X - 8000$ calcolare:

a) il valore atteso di X e di h(X)
b) la varianza e la deviazione standard di X e di h(X)

Sol.

a) Il valore atteso di X è al solito $\mu = E(X) = \sum_i xP(x) = 6700$.

Notare che h(X) è una trasformazione lineare del tipo a+bX per cui sappiamo che:

$$E(a + bX) = a + bE(X)$$

e questa regola vale $\forall X$ discreta o continua. Quindi

$$E[h(X)] = E[2X - 8000] = 2E(X) - 8000 = 2(6700) - 8000 = 5400$$

Per quanto riguarda la varianza di X la calcoliamo al solito come: $VAR(X) = E(X^2) - [E(X)]^2$

X=numero di pezzi venduti	x^2	P(x)	x^2P(x)
5000	25000000	0.2	5000000
6000	36000000	0.3	10800000
7000	49000000	0.2	9800000
8000	64000000	0.2	12800000
9000	81000000	0.1	8100000
			46500000

$E(X^2) = 46500000$

$[E(X)]^2 = 6700^2 = 44890000$

$VAR(X) = E(X^2) - [E(X)]^2 = 46500000 - 44890000 = 1610000$

Calcolare la varianza dei profitti corrisponde al calcolo della varianza per una v.c. soggetta a trasformazione lineare:

$$VAR(a + bX) = b^2 Var(X)$$

$VAR\big(h(X)\big) = VAR(2X - 8000) = 2^2(1610000) = 6440000$

$\sigma(h(X)) = \sqrt{6440000} = 2537.72$

Esercizio 3. Modelli discreti: binomiale, ipergeometrica, geometrica

In occasione della lotteria a premi vengono acquistati 6 biglietti per partecipare all'estrazione a sorte di 6 premi. Sapendo che sono stati venduti 30 biglietti in totale

a) quale modello di probabilità può essere utilizzato per descrivere la variabile casuale "numero di premi vinti" nell'ipotesi che i sei biglietti vincenti vengano estratti con reimmissione?

Sol.

E' possibile utilizzare un modello binomiale che consente di descrivere la probabilità di ottenere x successi su n possibili prove indipendenti.

X=nr.premi vinti   n=6

$p = \frac{6}{30} = \frac{1}{5} = 0.2$

quindi: $X \sim Bin(n, p)$

b) Applicare il modello di probabilità scelto per costruire la variabile casuale "numero di premi vinti"

Sol.

$X \sim Bin(n, p)$

La sua funzione di probabilità è:

$$P(X = x) = \binom{n}{x} p^x (1-p)^{n-x} = \frac{n!}{x!\,(n-x)!} p^x (1-p)^{n-x}$$

Distribuzione di probabilità di X:

$$P(X=0)=\binom{6}{0}0.2^0(1-0.2)^{6-0}=\frac{6!}{0!\,(6-0)!}0.2^0 0.8^6=\frac{6*5*4*3\ldots}{1*(6*5*4*3*\ldots)}*1*0.8^6=0.262$$

$$P(X=1)=\binom{6}{1}0.2^1(1-0.2)^{6-1}$$

$$P(X=2)=\binom{6}{2}0.2^2(1-0.2)^{6-2}$$

$$P(X=3)=\binom{6}{3}0.2^3(1-0.2)^{6-3}$$

$$P(X=4)=\binom{6}{4}0.2^4(1-0.2)^{6-4}$$

$$P(X=5)=\binom{6}{5}0.2^5(1-0.2)^{6-5}$$

$$P(X=6)=\binom{6}{6}0.2^6(1-0.2)^{6-6}$$

c) rappresentare graficamente la variabile casuale X= numero di premi vinti

d) In seguito all'estrazione, calcolare la probabilità di non vincere alcun premio, di vincere al massimo 2 premi, di vincere più di 4 premi.

Sol.

$P(X=0)$

$P(X\leq 2)=P(X=0)+P(X=1)+P(X=2)$

$P(X\geq 4)=P(X=4)+P(x=5)+P(X=6)$

oppure $1-P(X\leq 4)=1-P(X=0)-P(X=1)-P(X=2)-P(X=3)-P(X=4)$

e) Calcolare valore atteso e varianza della variabile casuale X=numero di premi vinti

Sol.

Se $X\sim Bin(n,p)$ allora $E(X)=np$ e $Var(X)=np(1-p)$

per cui:

$E(X)=6*0.2=1.2$ $\qquad$ $Var(X)=6*0.2(0.8)=0.96$

f) Supponiamo adesso di comprare 3 biglietti della lotteria per l'estrazione a sorte di 3 premi su un totale di 30 biglietti e che l'estrazione dei premi avvenga in blocco. Quale modello probabilistico può essere utilizzato per descrivere la variabile casuale X=numero di premi vinti?

Sol.

Siccome l'estrazione è in blocco, viene meno l'ipotesi di indipendenza e quindi il modello probabilistico corretto per X è un modello ipergeometrico. La differenza con il modello binomiale è che, in questo caso, la probabilità condizionata del verificarsi di un successo x si modifica in funzione delle precedenti estrazioni.

$X \sim IP(n, b, H)$

n= numero di prove

b=numero casi favorevoli

H=numero di oggetti

La corrispondente distribuzione di probabilità è:

$$P(X=x) = \frac{\binom{b}{x}\binom{H-b}{n-x}}{\binom{H}{n}} = \frac{\frac{b!}{x!\,(b-x)!}\frac{(H-b)!}{(n-x)!\,(H-b-n+x)!}}{\frac{H!}{n!\,(H-n)!}}$$

Nel nostro caso:

H=30

b=3

x=0,1,2,3

n=3

$$P(X=0) = \frac{\binom{3}{0}\binom{30-3}{3-0}}{\binom{30}{3}} = 0.720$$

$$P(X=1) = \frac{\binom{3}{1}\binom{30-3}{3-1}}{\binom{30}{3}} = 0.2593$$

$$P(X=2) = \frac{\binom{3}{2}\binom{30-3}{3-2}}{\binom{30}{3}} = 0.0125$$

$$P(X=3) = \frac{\binom{3}{3}\binom{30-3}{3-3}}{\binom{30}{3}} = 0.00024$$

g) Calcolare valore atteso e varianza della variabile casuale X=numero di premi vinti nell'ipotesi che i tre biglietti vengano estratti in blocco.

$$E(X) = n\frac{b}{H} = 3\frac{3}{30} = 0.3$$

$$Var(X) = n\frac{b}{H}\left(1 - \frac{b}{H}\right) * \left(\frac{H - n}{H - 1}\right) = 0.3 * 0.9 * 0.93 = 0.25$$

h) Consideriamo sempre il caso della lotteria a premi. Partecipiamo all'estrazione di 3 premi su un totale di 30 biglietti venduti. Viene estratto un biglietto per volta con reinserimento fino a quando non si ottiene il primo biglietto vincente. Qual è la probabilità di osservare il primo biglietto vincente dopo 3 estrazioni? E dopo almeno 4 estrazioni?

Sol.

In questo caso siamo interessati a modellare il numero di estrazioni necessarie affinchè si ottenga il primo biglietto vincente. La variabile casuale di interesse è rappresentata dal numero di prove necessarie per il primo "successo" per cui indichiamo con X=vincita alla prima estrazione.

La variabile casuale da utilizzare è un caso particolare della v.c.binomiale negativa, ovvero la v.c. geometrica, per cui:

$$X \sim Geom(p)$$

p= probabilità di successo $\frac{3}{30} = 0.1$

x=numero di prove

k=numero di successi. In tal caso k=1

La sua distribuzione di probabilità è la seguente:

$P(X = x) = p^k(1 - p)^{x-k}$ ossia $P(X = x) = p(1 - p)^{x-1}$

Quindi:

$P(X = 3) = 0.1^1(1 - 0.1)^{3-1} = 0.1 * 0.9^2 = 0.081$ probabilità che esca il primo biglietto vincente dopo 3 estrazioni

$$P(X \geq 4) = 1 - P(X \leq 4) = 1 - P(X = 1) - P(X = 2) - P(X = 3) - P(X = 4)$$

Calcoliamo le singole probabilità:

$$P(X = 1) = 0.1^1(1 - 0.1)^{1-1} = 0.1$$

$$P(X = 2) = 0.1^1(1 - 0.1)^{2-1} = 0.1 * 0.9 = 0.09$$

$P(X = 4) = 0.1^1(1 - 0.1)^{4-1} = 0.1 * 0.9^3 = 0.0729$

Avremo quindi:

$P(X \geq 4) = 1 - 0.1 - 0.09 - 0.081 - 0.0729 = 0.6561$

i) Qual è il numero medio di biglietti che bisogna estrarre prima di trovare un biglietto vincente?

$E(X) = \frac{1}{p} = \frac{1}{0.1} = 10$

Esercizio 4. Il modello di Poisson

In un negozio di abbigliamento arrivano mediamente 5 clienti ogni ora. Supponendo che gli arrivi seguano una distribuzione di Poisson, si determini:

a) La probabilità che in un'ora arrivino più di due clienti.
b) La probabilità che in 2 ore arrivino 5 clienti.
c) Il valore atteso e la varianza del numero di clienti che si presentano in una giornata (il negozio rimane aperto 8 ore al giorno).

Sol.

Con X indichiamo la variabile "numero di clienti che arrivano al negozio in un'ora". E' una variabile casuale di Poisson con parametro λ=5.

a) $P(X \geq 2) = 1 - P(X = 0) - P(X = 1) - P(X = 2) =$

$$1 - \frac{e^{-5}5^0}{0!} - \frac{e^{-5}5^1}{1!} - \frac{e^{-5}5^2}{2!} = 1 - 0.0067 - 0.0337 - 0.0842 = 0.8754$$

b) "numero di clienti che arrivano al negozio in due ore" è una variabile casuale di Poisson con parametro λ=5*2=10 (vedi condizione 2). Indichiamo con Y=numero di clienti che arrivano in negozio in due ore per cui:

$$Y \sim Poi(10)$$

$P(Y = 5) = \frac{e^{-10}10^5}{5!} = 0.0378$

c) La variabile numero di clienti che si presenta in una giornata lavorativa (8 ore) si distribuisce come una Poisson di parametro λ=8*5=40

$E(X) = Var(X) = 40$