CORSO DI STATISTICA (parte 2) - ESERCITAZIONE 2

Dott.ssa Antonella Costanzo

a.costanzo@unicas.it

Esercizio 1. La variabile Uniforme Continua

Data una scheda telefonica da 5 euro di cui non si sa se sia mai stata usata e nel caso sia stata usata non si conosce l'ammontare ancora disponibile, \`e ragionevole ipotizzare per tale ammontare X una distribuzione di probabilità uniforme (continua).

a) Calcolare valore atteso e varianza di X
b) Devo fare una telefonata interurbana il cui costo sarà di 2 euro. Quale \`e la probabilità che la scheda telefonica sia sufficiente per fare la telefonata?

Sol.

Indichiamo con X la variabile casuale "disponibilità residua nella scheda" X ∼ U(a=0, b=5).

Se X è una variabile casuale uniforme continua:

distribuzione di probabilità

$$f(x) = \begin{cases} \frac{1}{b-a} \; se \; x \in [a,b] \\ \\ 0 \; altrimenti \end{cases}$$

Momenti

$E(X) = \frac{a+b}{2}$ $\qquad Var(X) = \frac{(a-b)^2}{12}$

Funzione di ripartizione

$$F(X) = P(X \leq x) = \frac{x-a}{b-a}$$

a) $E(X) = \frac{a+b}{2} = \frac{0+5}{2} = 2.5$ $\quad Var(X) = \frac{(a-b)^2}{12} = \frac{(0-5)^2}{12} = 2.083$

b) La probabilità richiesta `e P(X ≥ 2). Usando la formula per la funzione di ripartizione di una variabile casuale uniforme, tale probabilità risulta:

$$P(X \geq 2) = 1 - P(X \leq 2) = 1 - \frac{2-0}{5-0} = 0.6$$

Esercizio 2. Variabile casuale esponenziale

Il tempo di attesa dell'arrivo tra veicoli in un incrocio particolare segue una distribuzione esponenziale con una media di 12 secondi.

a) Qual è la probabilità che il tempo di arrivo tra i veicoli è di 15 secondi o meno?

b) Qual è la probabilità che il tempo di arrivo tra i veicoli sia esattamente di 6 secondi?

c). Qual è la probabilità che il tempo di arrivo tra i veicoli è tra 6 e 10 secondi?

d). Qual è la probabilità che passino più di 30 secondi tra gli arrivi dei veicoli?

Sol.

$$X \sim Exp\left(\frac{1}{12}\right)$$

a) $P(X \leq 15) = 1 - e^{-\frac{15}{12}}$

b) $P(X = 6) = \emptyset$

c) $P(6 \leq X \leq 10) = F(10) - F(6) = 1 - e^{-\frac{10}{12}} - [1 - e^{-\frac{6}{12}}]$

d)$P(X \geq 30) = 1 - P(X \leq 30) = 1 - [1 - e^{-\frac{30}{12}}]$

Esercizio 3. Le proprietà della v.c. Normale

Siano X1, X2, and X3 variabili casuali normali i.i.d. con medie e varianze rispettivamente pari a:

	μ	σ^2
X_1	10	1
X_2	20	2
X_3	30	3

Definire la distribuzione della variabile $Y = X_1 + X_2 + X_3$, calcolare media varianza e deviazione standard di Y

Sol.

Ricordando la proprietà riproduttiva della normale per cui "la somma di n variabili casuali i.i.d che si distribuiscono secondo una legge normale è ancora una variabile casuale normale" possiamo scrivere:

$$Y \sim N(60, 6)$$

La media della somma (Y) è pari alla somma delle medie delle singole variabili, e per l'ipotesi di indipendenza, la varianza di Y è la somma delle singole varianze.

Esercizio 4. Combinazioni lineari di v.c. Normali

Un analista di costi deve stimare il costo unitario di un prodotto per il prossimo anno. Egli osserva che per la realizzazione della singola unità di prodotto sono necessarie 12 ore di lavoro e 5.8 grammi di materie prime. Attualmente ogni unità di prodotto ha un costo fisso pari a 184,50 dollari. L'analista stima che il costo di un'ora di lavoro anno prossimo sarà normalmente distribuito con un valore atteso di $ 45,75 e una deviazione standard di 1,80 $ (varianza 3.24); il costo delle materie prime anch'esso sarà normalmente distribuito con un valore atteso di 62,35 $ ed una deviazione standard di $ 2.52 (varianza 6.3504). Trovare la distribuzione del costo unitario del prodotto. Calcolare inoltre valore atteso e varianza.

Sol.

Indichiamo con L il costo del lavoro, con M il costo delle materie prime impiegate per la realizzazione e C sono i costi fissi. Il costo unitario del prodotto è Q.

La funzione di costo la possiamo quindi esprimere come segue:

$$Q = 12L + 5.8M + C$$

$L \sim N(47.75, 3.24)$

$M \sim N(62.35, 6.3504)$

Siccome ipotizziamo che il costo del lavoro non influenza il costo delle materie prime possiamo considerare queste due variabili indipendenti. A questo punto anche Q sarà normalmente distribuito con media e varianza pari rispettivamente:

$$E(Q) = 12(45.75) + 5.8(62.35) + 184.50 = 1095.13$$

$$Var(Q) = 12^2(3.24) + 5.8^2(6.3504) = 680.19$$

Quindi: $Q \sim N(1095.13, 680.19)$

Il risultato deriva dall'applicazione della proprietà delle combinazioni lineari di v.c. Normali:

Siano $X_1, X_2, \dots, X_k$ *i.i.d. e distribuite normalmente, allora* $Y = a_1X_1 + a_2X_2 + \cdots + a_kX_k$ *è tale che Y si distribuisce anch'essa come una Normale con media e varianza pari rispettivamente a:*

$$E(Y) = a_1E(X_1) + a_2E(X_2) + \cdots + a_kE(X_k)$$
$$Var(Y) = a_1^2Var(X_1) + a_2^2Var(X_2) + \cdots + a_k^2Var(X_k)$$

Esercizio 5. Applicazioni v.c. Normale

Su un campione di atleti sono state condotte delle analisi in laboratorio, ed è stata riscontrato che la quantità di proteina A presente nel sangue ha una media di 35 microgrammi e deviazione standard σ uguale 5. Quale è la probabilità di trovare:

a) individui con valori superiori a 40;

b) individui con valori inferiori a 40;

c) individui con valori inferiori a 25;

d) individui con valori compresi tra 40 e 50;

e) individui con valori tra 30 e 40.

Sol.

a) Indichiamo con $Z = \frac{X-\mu}{\sigma} = \frac{40-35}{5} = 1$ la v.c. Normale standardizzata

$P(X \geq 40) \equiv P(Z \geq 1) = 1 - P(Z \leq 1) = 1 - 0.8413 = 0.1517$

b) $P(X \leq 40) \equiv P(Z \leq 1) = 0.8413$

c) Indichiamo con $Z = \frac{X-\mu}{\sigma} = \frac{25-35}{5} = -2$

$P(X \leq 25) \equiv P(Z \leq -2) = 1 - P(Z \leq 2) = 1 - 0.9772 = 0.0288$

d) Indichiamo con $z_1 = \frac{X_1-\mu}{\sigma} = \frac{40-35}{5} = 1$ e $z_2 = \frac{X_2-\mu}{\sigma} = \frac{50-35}{5} = 3$

$P(40 \leq X \leq 50) \equiv P(3 \leq Z \leq 1) = P(Z \leq 3) - P(Z \leq 1) = 0.9987 - 0.8413 = 0.1574$

e) con $z_1 = \frac{X_1-\mu}{\sigma} = \frac{30-35}{5} = -1$ e $z_2 = \frac{X_2-\mu}{\sigma} = \frac{40-35}{5} = 1$

$P(30 \leq X \leq 40) \equiv P(-1 \leq Z \leq 1) = P(Z \leq 1) - P(Z \leq -1) = P(Z \leq 1) - [1 - P(Z \leq 1)] =$

$$= 0.8413 - [1 - 0.8413] = 0.6826$$

E' possibile utilizzare le tabelle anche nel modo inverso; cioè leggere su di esse la probabilità e ricavare il valore di Z corrispondente. Ad esempio:
Determinare il valore di proteina A presente nel sangue dal solo 5% degli atleti (novantacinquesimo percentile)

Sol.

Considerando la tavola della funzione di ripartizione della v.c. Normale standardizzata occorre individuare quel valore z che lascia a sinistra il 95% dell'area in modo da isolare il restante 5%. La proporzione 0.95 si trova tra 0. 9495(Z=1.64) e 0.9505 (Z=1.65). La loro media è Z=1,645. Essa indica che occorre spostarsi a destra della media (35) di una quantità pari a 1,645 volte la deviazione standard. In particolare: $X = \mu + 1.645\sigma$

$X = 35 + 1.645(5) = 35 + 8.225 = 43.225$

Esercizio 6.

La durata di una batteria si distribuisce secondo una variabile aleatoria normale con media pari a 50 ore e deviazione standard pari a 5 ore. Si calcoli:

(a) la probabilità che una batteria duri dalle 42 alle 52 ore;
(b) il valore in ore tale da garantire la durata del 40% delle batterie

Sol.

a) $X \sim N(50{,}25)$

$P(42 \leq X \leq 52) = ?$

Standardizzazione:

$$z_1 = \frac{x_1 - \mu}{\sigma} = \frac{42 - 50}{5} = -1.60$$
$$z_2 = \frac{x_2 - \mu}{\sigma} = \frac{52 - 50}{5} = 0.40$$

Calcolo:

$$P(z_1 \leq Z \leq z_2) = P(-1.60 \leq Z \leq 0.40) = P(Z \leq 0.40) - P(Z \leq -1.60) = 0.6554 - [1 - 0.9452]$$
$$= 0.6006$$

b) Occorre individuare il valore di z tale che $P(Z \leq z) = 0.40$. Dapprima individuo il valore z in corrispondenza dell'area richiesta che risulta pari, approssimando, a $z_{0.40}$=0.25.
Per la simmetria della curva normale allora il valore di $z_{0.40}$ è pari a -0.25.

Sapendo che:

$$z = \frac{x - \mu}{\sigma} = -0.25$$

Tramite la formula inversa di standardizzazione otteniamo:

$z = \frac{x-50}{5} = 0.25$ da cui $x = 50 + 5(-0.25) = 48.75$