Università degli Studi di Cassino

Esercitazione di Statistica 2 dell'8.02.2007

Simona Balzano

Esercizio 1
Uno studente supera una prova con probabilità pari a 0,6. Considerando un campione di ampiezza 10, determinare la probabilità che:

a) tutti gli studenti tranne 1 superino la prova;
b) 2 studenti superino la prova;
c) non meno di 8 studenti superino la prova;
d) non più di 8 studenti superino la prova;
e) superino la prova da 2 a 6 studenti;

Soluzione

La v.c. di riferimento è X = "numero di studenti che superano la prova su n = 10", che segue la distribuzione Binomiale con parametri n = 10 e p = 0,6.
La probabilità di un valore X = x si ottiene come:

$$P(x) = \binom{n}{x} p^x (1-p)^{n-x}$$

che esprime la probabilità di x successi in n prove.

a) P (X = 9) = P(9)

$$P(9) = \binom{10}{9} \times 0,6^9 \times (1-0,6)^{10-9} = \frac{10!}{9! \times (10-9)!} \times 0,6^9 \times (1-0,6)^1 = 10 \times 0,0101 \times 0,4 = \mathbf{0,0403}$$

b) P (X = 2) = P(2)

$$P(2) = \binom{10}{2} \times 0,6^2 \times (1-0,6)^{10-2} = \frac{10!}{2! \times (10-2)!} \times 0,6^2 \times (1-0,6)^8 = 45 \times 0,36 \times 0,0006 = \mathbf{0,0106}$$

c) P (X ≥ 8) = P(8) + P(9) + P(10)

$$P(8) = \binom{10}{8} \times 0,6^8 \times (1-0,6)^{10-8} = \frac{10!}{8! \times (10-8)!} \times 0,6^8 \times (1-0,6)^2 = 45 \times 0,0167 \times 0,16 = 0,12$$

$$P(9) = 0,0403$$

$$P(10) = \binom{10}{10} \times 0,6^{10} \times (1-0,6)^{10-10} = 1 \times 0,6^{10} \times (1-0,6)^0 = 0,006$$

P (X ≥ 8) = P(8) + P(9) + P(10) = 0,0027 + 0,0403 + 0,006 = **0,049**

d) P (X ≤ 8) = 1 − [P(9) + P(10)] = 1 − [0,0403 + 0,006] = **0,9537**

e) P (2 ≤ X ≤ 6) = $\sum_{i=2}^{6} P(X = i)$

P(2) = 0,0106

$$P(3) = \binom{10}{3} \times 0,6^3 \times (1-0,6)^{10-3} = \frac{10!}{3! \times (10-3)!} \times 0,6^3 \times (1-0,6)^7 = 120 \times 0,216 \times 0,0016 = 0,0425$$

$$P(4) = \binom{10}{4} \times 0,6^4 \times (1-0,6)^{10-4} = \frac{10!}{4! \times (10-4)!} \times 0,6^4 \times (1-0,6)^6 = 210 \times 0,1296 \times 0,004 = 0,111$$

$$P(5) = \binom{10}{5} \times 0,6^5 \times (1-0,6)^{10-5} = \frac{10!}{5! \times (10-5)!} \times 0,6^5 \times (1-0,6)^5 = 252 \times 0,078 \times 0,01 = 0,2$$

$$P(6) = \binom{10}{6} \times 0,6^6 \times (1-0,6)^{10-6} = \frac{10!}{6! \times (10-6)!} \times 0,6^6 \times (1-0,6)^6 = 210 \times 0,047 \times 0,026 = 0,25$$

P (2 ≤ X ≤ 6) = P(2) + P(3) + P(4) + P(5) + P(6) =
= 0,0106 + 0,0425 + 0,111 + 0,2 + 0,25 = **0,6141**

Esercizio 2

La probabilità che in un paese del Nord-Europa nevichi in un giorno invernale è 0,3. determinare la probabilità che durante l'inverno:

a) non nevichi in una settimana;
b) nevichi 2 giorni in una settimana;
c) nevichi più di 3 giorni in una settimana;
d) nevichi al massimo 5 giorni in un mese;
e) nevichi da 10 a 15 giorni in un mese.
f) nevichi 10 giorni in un mese;

Soluzione
La v.c. di riferimento è X = "numero di giorni di neve", che segue la distribuzione Binomiale con parametri p = 0,3.
La probabilità di un valore X = x si ottiene come:

$$P(x) = \binom{n}{x} p^x (1-p)^{n-x}$$

che esprime la probabilità di x successi in n prove.

a) P(0); n = 7; p = 0,3

$$P(0) = \binom{7}{0} \times 0,3^0 \times (1-0,3)^{7-0} = 1 \times 1 \times 0,7^7 = \mathbf{0,0823}$$

b) P(2); n = 7; p = 0,3

$$P(2) = \binom{7}{2} \times 0,3^2 \times (1-0,3)^{7-2} = \frac{7!}{2!(7-2)!} \times 0,09 \times 0,7^5 = 21 \times 0,09 \times 0,168 = \mathbf{0,3176}$$

c) P(X ≥ 3); n = 7; p = 0,3

$$P(X > 3) = \sum_{i=4}^{7} P(X=i) = 1 - [P(0) + P(1) + P(2) + P(3)]$$

P(0) = 0,0823

$$P(1) = \binom{7}{1} \times 0,3^1 \times (1-0,3)^{7-1} = \frac{7!}{1!(7-1)!} \times 0,3 \times 0,7^6 = 1 \times 0,3 \times 0,1176 = 0,247$$

P(2) = 0,3176

$$P(1) = \binom{7}{3} \times 0,3^3 \times (1-0,3)^{7-3} = \frac{7!}{3!(7-3)!} \times 0,3^3 \times 0,7^4 = 35 \times 0,027 \times 0,24 = 0,227$$

P(X ≥ 3) = 0,0823 + 0,247 + 0,3176 = **0,6469**

d) $P(X \leq 5)$; $n = 31$ (o 30); $p = 0,3$

n è sufficientemente elevato da poter considerare l'approssimazione Normale ad una Binomiale di parametri

$$\mu = n \times p \qquad e \qquad \sigma^2 = n \times p \times (1-p).$$

Per cui, essendo:

$n \times p = 31 \times 0,3 = 9,3$

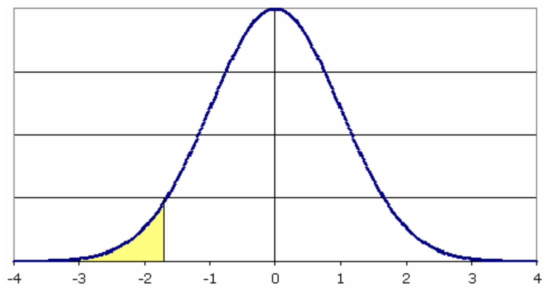
e

$\sigma^2 = n \times p \times (1-p) = 31 \times 0,3 \times 0,7 = 6,51$

si può dedurre:

$$X \sim N(9,3; 6,51)$$

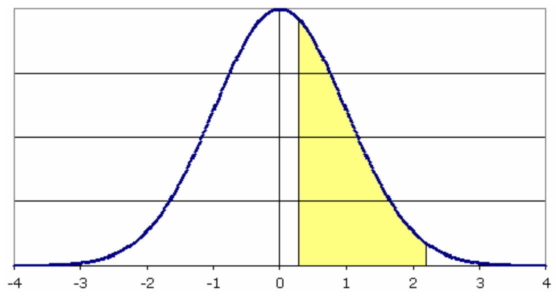
$$z_5 = \frac{5 - 9,3}{\sqrt{6,51}} = -1,68$$



$P(X \leq 5) = P(Z \leq -1,68) = 1 - F(1,68) = 1 - 0,9535 =$ **0,0465**

e) $P(10 \leq X \leq 15)$; $n = 31$; $p = 0,3$.

$$z_{10} = \frac{10 - 9,3}{\sqrt{6,51}} = 0,27; \qquad z_{15} = \frac{15 - 9,3}{\sqrt{6,51}} = 2,23$$



$P(10 \leq X \leq 15) = P(0,27 \leq Z \leq 2,23) = F(2,23) - F(0,27)$
$= 0,9871 - 0,6064 =$ **0,3807**

f) P(X = 10); n = 31; p = 0,3

In questo caso, nonostante n sia elevato, non è possibile utilizzare l'approssimazione normale perché non è possibile determinare la probabilità che una v.c. continua assuma un preciso valore.
Bisogna, quindi, utilizzare la formula della binomiale, per cui:

$$P(10) = \binom{31}{10} \times 0,3^{10} \times (1-0,3)^{31-10} = \frac{31!}{10!(31-10)!} \times 0,3^{10} \times 0,7^{21} =$$

$$= 44.352.165 \times 0,0000059 \times 0,00056 = \mathbf{0,146}$$

Esercizio 3

Una certa auto prodotta da una certa casa percorre mediamente 15 chilometri con 1 litro di carburante, con uno scarto quadratico medio di 5 chilometri.
Determinare le seguenti probabilità:

a) che 50 auto percorrano mediamente al massimo 16 Km con 1 litro;
b) che 50 auto percorrano mediamente da 13 a 17 Km con 1 litro;
c) che 16 auto percorrano mediamente da 13 a 17 km con 1 litro.

Soluzione

La v.c. descritta è X = "numero di chilometri percorsi con 1 litro", che ha media μ_X = 15 e scarto quadratico medio σ_X = 5, di cui non è nota la distribuzione.
Le probabilità richieste sono invece relative al consumo medio, per cui la v.c. di riferimento è la media campionaria $\bar{x}$, che avrà $\mu_{\bar{x}} = \mu_X = 15$ e scarto quadratico medio $\sigma_{\bar{x}} = \frac{\sigma_X}{\sqrt{n}} = \frac{5}{\sqrt{n}}$, che assumerà quindi i valori:

$\sigma_{\bar{x}} = \frac{5}{\sqrt{50}} = 0,707$ nel campione di ampiezza n = 50;

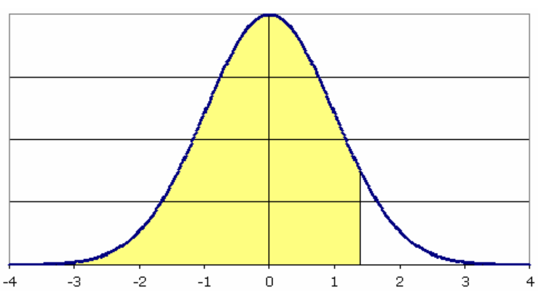
$\sigma_{\bar{x}} = \frac{5}{\sqrt{16}} = 1,25$ nel campione di ampiezza n = 16.

a) $P(\bar{x} \leq 16)$

n è sufficientemente elevato per poter utilizzare il teorema del limite centrale e affermare che la v.c. media campionaria $\bar{x}$ si distribuisce normalmente:

$$\bar{x} \sim N(15; 0,5)$$

$$z_{16} = \frac{16-15}{\sqrt{0,5}} = 1,4142$$

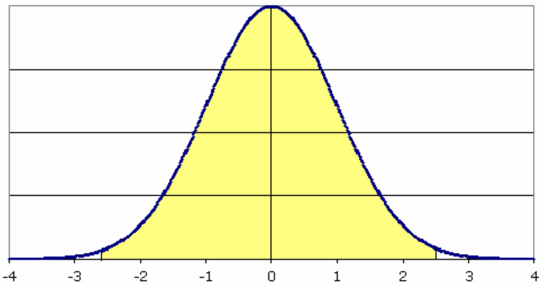


$P(\bar{x} \leq 16) = P(Z \leq 1,41) = F(1,41) =$ **0,9207**

b) $P(13 \le \bar{x} \le 17)$

Anche in questo caso si fa riferimento alla distribuzione normale del punto a).

$$z_{13} = \frac{13-15}{\sqrt{0,5}} = -2,82\,; \qquad z_{17} = \frac{17-15}{\sqrt{0,5}} = 2,82$$



$P(13 \le \bar{x} \le 17) = P(-2,82 \le Z \le 2,82) = F(2,82) - [1 - F(2,82)]$
$= 0,9976 - (1 - 0,9976) =$ **0,9952**

c) $P(13 \le \bar{x} \le 17)$

In questo caso n = 16 non è sufficientemente elevato per applicare il teorema del limite centrale.
Affermare che $13 \le \bar{x} \le 17$ equivale ad affermare che in valore assoluto lo scarto tra il valore che assume $\bar{x}$ e la sua media ($\mu_{\bar{x}} = 15$) è al massimo par a 2, ossia:

$$\left|\bar{x} - \mu_{\bar{x}}\right| \le 2$$

Si può, dunque, ricorrere alla disuguaglianza di Cebicev:

$$P\left(\left|X - \mu\right| \le k\sigma\right) \ge 1 - \frac{1}{k^2}\,,$$

che non da un valore preciso di probabilità ma una soglia minima al di sotto della quale la probabilità non scende.

In altre parole, possiamo determinare la probabilità minima per l'intervallo massimo $\left|\bar{x} - \mu_{\bar{x}}\right| \le 2$.

Tale probabilità minima è pari a $1 - \frac{1}{k^2}$.

Sapendo che $k\sigma = 2$, deriva che:

$$k = \frac{2}{\sigma} = \frac{2}{1,25} = 1,6\,.$$

Quindi:

$$P\left(\left|\bar{x} - \mu_{\bar{x}}\right| \le 2\right) \ge 1 - \frac{1}{1,6^2}$$

cioè

$$P\left(\left|\overline{x}-15\right|\leq 2\right)\geq 0,609$$

Che può essere scritto come:

$P\left(13\leq \overline{x}\leq 17\right)\geq \mathbf{0,609}$

Esercizio 4

Una v.c. X ha media μ = 40 e una varianza σ^2 = 22. Determinare la probabilità che $\overline{x}$ sia compresa tra 39 e 41,5:

a) in un campione di ampiezza n = 100;
b) in un campione di ampiezza n = 25.

Soluzione

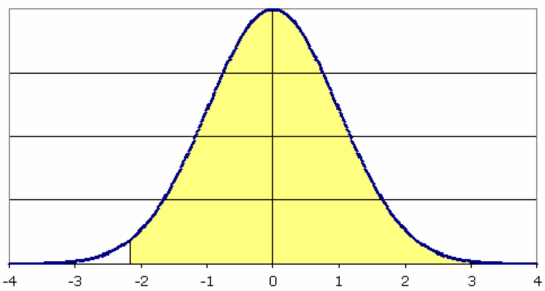
La v.c. di riferimento è la media campionaria $\overline{x}$.

a) P (39 ≤ $\overline{x}$ ≤ 41,5); n = 100

Per il teorema del limite centrale $\overline{x}$ si distribuisce normalmente con media $\mu_{\overline{x}} = 40$ e varianza $\sigma^2_{\overline{x}} = \frac{\sigma^2_X}{n} = \frac{22}{100} = 0,22$:

$$\overline{x} \sim N(40;0,36)$$

$$z_{39} = \frac{39-40}{\sqrt{0,22}} = -2,13 \qquad z_{41,5} = \frac{41,5-40}{\sqrt{0,22}} = 3,2$$



P (39 ≤ $\overline{x}$ ≤ 41,5) = P (-2,13 ≤ z ≤ 3,2) = F(3,2) – [1 – F(2,13)] =
= 0,9993 – (1 – 0,9834) = **0,9827**

b) P (39 ≤ $\overline{x}$ ≤ 41,5); n = 25

Non è possibile ricorrere al teorema del limite centrale perché n è piccolo. Si può, quindi, utilizzare la disuguaglianza di Cebicev per determinare il valore minimo della probabilità richiesta.
La v.c. media campionaria $\overline{x}$ segue una distribuzione non nota ed ha media $\mu_{\overline{x}} = 40$ e varianza:

$$\sigma^2_{\overline{x}} = \frac{\sigma^2_X}{n} = \frac{22}{25} = 0,88 \text{ (quindi } \sigma_{\overline{x}} = 0,94\text{)}$$

C'è però una difficoltà: la disuguaglianza di Cebicev fornisce il valore minimo della probabilità associata ad un intervallo simmetrico intorno alla media, mentre l'intervallo [39; 41,5] non è simmetrico intorno alla media $\mu_{\bar{x}} = 40$.

Ipotizzando la simmetria per la distribuzione incognita di $\bar{x}$, possiamo suddividere il problema in due parti, considerando le probabilità minime associate ai due intervalli simmetrici [39; 41] e [38,5; 41,5], dividerle per 2 e sommare i due risultati.
In altre parole, dobbiamo determinare i valori minimi di P (39 ≤ $\bar{x}$ ≤ 41) e P (38,5 ≤ $\bar{x}$ ≤ 41,5) con la disuguaglianza di Cebicev e dividere tali minimi per due, per ottenere le probabilità dei semi-intervalli [39; 40] e [40; 41,5], per poi calcolare la probabilità richiesta come probabilità dell'unione dei due semi-intervalli:

$$P(39 \le \bar{x} \le 41,5) = P(39 \le \bar{x} \le 40) + P(40 \le \bar{x} \le 41,5)$$

Per l'intervallo [39; 41] la disuguaglianza di Cebicev ci dice che:

$$P\left(\left|\bar{x} - \mu_{\bar{x}}\right| \le 1\right) \ge 1 - \frac{1}{k^2}$$

Ossia $k\sigma_{\bar{x}} = 1$, quindi:

$$k = \frac{1}{\sigma_{\bar{x}}} = \frac{1}{0,94} = 1,07$$

Da cui:

$$P\left(\left|\bar{x} - \mu_{\bar{x}}\right| \le 1\right) \ge 1 - \frac{1}{1,07^2}$$

$$P(39 \le \bar{x} \le 41) \ge 0,12$$

Per il semi-intervallo [39; 40], ipotizzando la simmetria della distribuzione incognita di $\bar{x}$:

$$P\left(39 \le \bar{x} \le 40\right) \ge \frac{0,12}{2}$$

$$P\left(39 \le \bar{x} \le 40\right) \ge 0,06$$

Per l'intervallo [38,5; 41,5] la disuguaglianza di Cebicev ci dice che:

$$P\left(\left|\bar{x} - \mu_{\bar{x}}\right| \le 2\right) \ge 1 - \frac{1}{k^2}$$

Ossia $k\sigma_{\bar{x}} = 2$, quindi:

$$k = \frac{2}{\sigma_{\bar{x}}} = \frac{2}{0,94} = 2,14$$

Da cui:

$$P\left(\left|\overline{x} - \mu_{\overline{x}}\right| \le 2\right) \ge 1 - \frac{1}{2,14^2}$$

$$P(38,5 \le \overline{x} \le 41,5) \ge 0,78$$

Per il semi-intervallo [40; 41,5], ipotizzando la simmetria della distribuzione incognita di $\overline{x}$:

$$P\left(40 \le \overline{x} \le 41,5\right) \ge \frac{0,78}{2}$$

$$P\left(40 \le \overline{x} \le 41,5\right) \ge 0,39$$

Considerando, infine, l'unione dei 2 semi-intervalli:

$$P\left(39 \le \overline{x} \le 41,5\right) = P\left(39 \le \overline{x} \le 40\right) + P\left(40 \le \overline{x} \le 41,5\right)$$

$$P\left(39 \le \overline{x} \le 41,5\right) \ge 0,06 + 0,39$$

$$P\left(39 \le \overline{x} \le 41,5\right) \ge \mathbf{0,45}$$