Università degli Studi di Cassino, Anno accademico 2004-2005
Corso di Statistica 2, Prof. M. Furno

Esercitazione del 15/02/2005
dott. Claudio Conversano

Esercizio 1
Alle elezioni amministrative del comune di Cassino del giugno 2003, i voti ottenuti dal partito XXX sono stati pari a 12.000. Il numero totale di elettori era 30.000. Si determini la probabilità che, estraendo con ripetizione un campione di 1000 elettori, almeno il 38% di essi abbia votato XXX.

SVOLGIMENTO

Si tratta di *n=1.000* estrazioni con ripetizione. I possibili risultati sono due: l'elettore ha votato per XXX (evento successo), l'elettore non ha votato per il partito XXX (evento insuccesso). La probabilità di successo in una sottoprova è 12.000/30.000 = 0.4. La probabilità richiesta può essere calcolata come:

$$P(X \geq 380) = \sum_{x=380}^{1000} \binom{1000}{x} 0{,}4^{x} (1-0{,}4)^{1000-x}$$

Il calcolo è evidentemente lungo e laborioso. Tuttavia, per n che tende ad infinito, la v.c. X tende ad una distribuzione normale $X \approx N\left(\mu = p, \sigma^2 = p(1-p)\right)$.Quando si ha un campione sufficientemente grande, come in questo caso, si può approssimare la distribuzione binomiale con la distribuzione normale.

La probabilità richiesta risulta:

$$P\left(Z \geq \frac{X - np}{\sqrt{n \cdot p \cdot (1-p)}}\right) = P\left(Z \geq \frac{380 - 1000 \cdot 0.4}{\sqrt{1000 \cdot 0.4 \cdot 0.6}}\right) = P(Z \geq -1.29) = 1 - P(Z \geq -1.29) = 0.5985$$

Esercizio 2
Sia X una variabile casuale che ha una distribuzione *t* di Student con *r* gradi di libertà. Supponendo che sia:

a) $Var(X) = 1.25$, si determinino i valori c_2, c_3, e c_4 per i quali risulta $P(|X| \geq c_2) = 0.1$, $P(|X| \geq c_3) = 0.05$, $P(|X| \leq c_4) = 0.99$.

SVOLGIMENTO

a) essendo $Var(X) = \frac{r}{r-2} = 1.25$ risulta $r = 10$.

Per determinare il valore c_2, dalla relazione $P(|X| \geq c_2) = 1 - P(-c_2 \leq X \leq c_2) = 0.01$ si ricava $P(-c_2 \leq X \leq c_2) = 0.99$, ovvero $P(X \leq c_2) = 0.995$ e quindi $P(X \geq c_2) = 0.005$. Per $r=10$ risulta $c_2 = 3.169$

Per determinare il valore c_3, dalla relazione $P(|X| \geq c_3) = 1 - P(-c_3 \leq X \leq c_3) = 0.05$ si ricava $P(-c_3 \leq X \leq c_3) = 0.95$, ovvero $P(X \leq c_3) = 0.975$ e quindi $P(X \geq c_3) = 0.025$. Per $r=10$ risulta $c_3 = 2.228$

Per determinare il valore c_4, dalla relazione $P(|X| \leq c_4) = P(-c_4 \leq X \leq c_4) = 0.99$ si ricava $P(X \leq c_4) = 0.995$ e $P(X \geq c_4) = 0.005$. Per $r=10$ risulta $c_4 = 3.169$

Esercizio 3 (campionamento – costruzione della distribuzione campionaria)
Si consideri una popolazione di 4 studenti che, nella sessione d'esame del mese di dicembre 2004 hanno sostenuto rispettivamente 1, 1, 3, 5 esami.

a) Si costruisca la distribuzione campionaria del numero mediano di esami sostenuti per campioni di ordine 3 estratti con ripetizione e si calcoli il valore atteso e la varianza.
b) Si costruisca la distribuzione campionaria del numero mediano di esami sostenuti per campioni di ordine 3 estratti senza ripetizione e si calcoli il valore atteso e la varianza.

SVOLGIMENTO

a) Si costruisca la distribuzione campionaria del numero mediano di esami sostenuti per campioni di ordine 3 estratti con ripetizione e si calcoli il valore atteso e la varianza.

Effettuando un'estrazione con ripetizione di tre unità, l'universo dei campioni è costituito da $N^n = 4^3 = 64$ elementi. L'universo dei campioni di determinazioni del carattere X (numero di esami sostenuti nella sessione…..) risulta:

$$
\begin{array}{cccccccc}
(1,1,1) & (1,1,1) & (1,3,1) & (1,5,1) & (1,1,1) & (1,1,1) & (1,3,1) & (1,5,1) \\
(1,1,3) & (1,1,3) & (1,3,3) & (1,5,3) & (1,1,5) & (1,1,5) & (1,3,5) & (1,5,5) \\
(1,1,1) & (1,1,1) & (1,3,1) & (1,5,1) & (1,1,1) & (1,1,1) & (1,3,1) & (1,5,1) \\
(1,1,3) & (1,1,3) & (1,3,3) & (1,5,3) & (1,1,5) & (1,1,5) & (1,3,5) & (1,5,5) \\
(3,1,1) & (3,1,1) & (3,3,1) & (3,5,1) & (3,1,1) & (3,1,1) & (3,3,1) & (3,5,1) \\
(3,1,3) & (3,1,3) & (3,3,3) & (3,5,3) & (3,1,5) & (3,1,5) & (3,3,5) & (3,5,5) \\
(5,1,1) & (5,1,1) & (5,3,1) & (5,5,1) & (5,1,1) & (5,1,1) & (5,3,1) & (5,5,1) \\
(5,1,3) & (5,1,3) & (5,3,3) & (5,5,3) & (5,1,5) & (5,1,5) & (5,3,5) & (5,5,5)
\end{array}
$$

Poiché ciascuna unità ha la stessa probabilità di essere estratta, ogni campione ha probabilità pari a $1/64$ di presentarsi. Le mediane dei 64 campioni sono:

$$
\begin{array}{cccccccc}
1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\
1 & 1 & 3 & 3 & 1 & 1 & 3 & 5 \\
1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\
1 & 1 & 3 & 3 & 1 & 1 & 3 & 5 \\
1 & 1 & 3 & 3 & 1 & 1 & 3 & 3 \\
3 & 3 & 3 & 3 & 3 & 3 & 3 & 5 \\
1 & 1 & 3 & 5 & 1 & 1 & 3 & 5 \\
3 & 3 & 3 & 5 & 5 & 5 & 5 & 5
\end{array}
$$

La distribuzione campionaria della mediana risulta:

$$
\begin{array}{lccc}
\bar{X}_{Me}: & 1 & 3 & 5 \\
f\left(\bar{X}_{Me}\right): & 32/64 & 22/64 & 10/64
\end{array}
$$

$$E\left(\bar{X}_{Me}\right) = 1 \cdot 32/64 + 3 \cdot 22/64 + 5 \cdot 10/64 = 2.3125$$

$Var\left(\overline{X}_{Me}\right) = 1^2 \cdot 32/64 + 3^2 \cdot 22/64 + 5^2 \cdot 10/64 - 2.3125^2 = 7.5 - 5.35 = 2.15$

b) **Si costruisca la distribuzione campionaria del numero mediano di esami sostenuti per campioni di ordine 3 estratti senza ripetizione e si calcoli il valore atteso e la varianza.**

Effettuando un'estrazione senza ripetizione di tre unità, l'universo dei campioni è costituito da $N\left(N-1\right) = 4 \cdot 3 \cdot 2 = 24$[1] elementi. Essi sono:

$$\begin{array}{cccccc} U_1U_2U_3 & U_1U_2U_4 & U_1U_3U_2 & U_1U_3U_4 & U_1U_4U_2 & U_1U_4U_3 \\ U_2U_1U_3 & U_2U_1U_4 & U_2U_3U_1 & U_2U_3U_4 & U_2U_4U_1 & U_2U_4U_3 \\ U_3U_1U_2 & U_3U_1U_4 & U_3U_2U_1 & U_3U_2U_4 & U_3U_4U_1 & U_3U_4U_2 \\ U_4U_1U_2 & U_4U_1U_3 & U_4U_2U_1 & U_4U_2U_3 & U_4U_3U_1 & U_4U_3U_2 \end{array}$$

Il corrispondente universo dei campioni di determinazioni del carattere X (numero di esami sostenuti nella sessione…..) risulta:

$$\begin{array}{cccccc} (1,1,3) & (1,1,5) & (1,3,5) & (1,3,5) & (1,5,1) & (1,5,3) \\ (1,1,3) & (1,1,5) & (1,3,1) & (1,3,5) & (1,5,1) & (1,5,3) \\ (3,1,1) & (3,1,5) & (3,1,1) & (3,1,5) & (3,5,1) & (3,5,1) \\ (5,1,1) & (5,1,3) & (5,1,1) & (5,1,3) & (5,3,1) & (5,3,1) \end{array}$$

Poiché ciascuna unità ha la stessa probabilità di essere estratta, ogni campione ha probabilità pari a 1/16 di presentarsi. Le mediane dei 24 campioni sono:

$$\begin{array}{cccccc} 1 & 1 & 3 & 3 & 1 & 3 \\ 1 & 1 & 1 & 3 & 1 & 3 \\ 1 & 3 & 1 & 3 & 3 & 3 \\ 1 & 3 & 1 & 3 & 3 & 3 \end{array}$$

La distribuzione campionaria della mediana risulta:

$$\begin{array}{lcc} \overline{X}_{Me}: & 1 & 3 \\ f\left(\overline{X}_{Me}\right): & 11/24 & 13/24 \end{array}$$

$E\left(\overline{X}_{Me}\right) = 1 \cdot 11/24 + 3 \cdot 13/24 = 2.08\overline{3}$

$Var\left(\overline{X}_{Me}\right) = 1^2 \cdot 11/24 + 3^2 \cdot 13/24 - 2.08\overline{3}^2 = 5.3\overline{3} - 4.34 = 0.9\overline{9}$

---

[1] In generale, il numero di possibili campioni di ordine *n* che possono essere estratti senza ripetizione da una popolazione di *N* elementi è: $N\left(N-1\right)\left(N-2\right)\ldots\left(N-n+1\right)$. Ad esempio, il numero di possibili campioni di ordine 4 estraibili da una popolazione di 10 elementi è: *10(10-1)(10-2)(10-3)=5040.*