**Università degli Studi di Cassino**

**Esercitazione di Statistica 2 dell'1.02.2007**

**Dott.ssa Simona Balzano**

## Esercizio 1

Il risultato di un test attitudinale ha una media pari a 220 con uno scarto quadratico medio di 20.
Effettuando il test su un campione di 100 soggetti determinare:

a) la probabilità che il punteggio medio sia maggiore di 225;
b) la probabilità che il punteggio medio sia compreso tra 216 e 223;
c) il punteggio medio minimo registrato nel 30% più abile del campione.

## Soluzione

Poiché la numerosità campionaria è elevata, si può sfruttare il teorema del limite centrale, ossia trattare il punteggio medio $\bar{x}$ come una v.c. normale con parametri $\mu_{\bar{x}} = \mu_X$ e $\sigma^2_{\bar{x}} = \frac{\sigma^2_X}{n}$

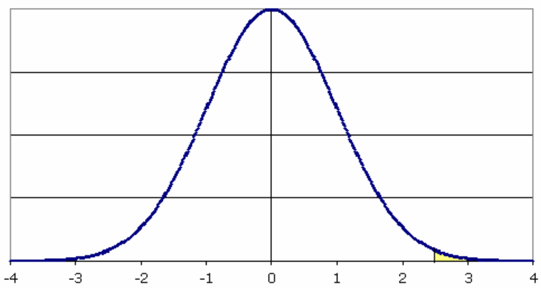
$$\bar{x} \sim N\left(\mu_X; \frac{\sigma^2_X}{n}\right)$$

ossia:

$$\bar{x} \sim N(220; 4)$$

**a)** P (X > 225)

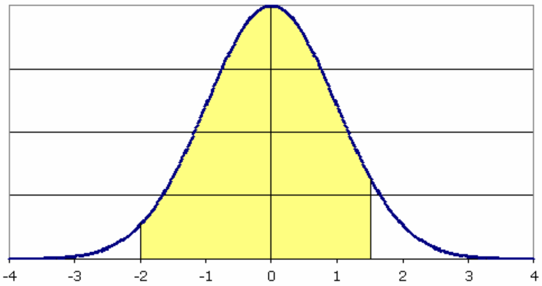
$$z_{225} = \frac{225 - 220}{2} = 2,5$$




P (X > 225) = P (Z > 2,5) = 1 – F(2,5) = 1 – 0,9938 = **0,0062**

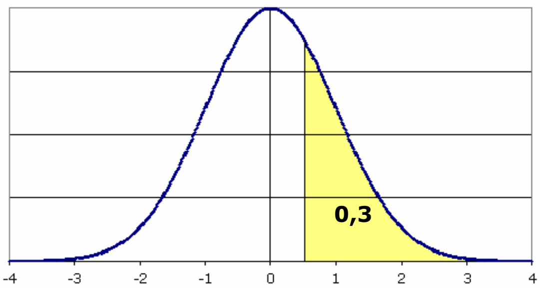
**b)** P (216 ≤ X ≤ 223)

$$z_{216} = \frac{216 - 220}{2} = -2\,; \quad z_{223} = \frac{223 - 220}{2} = 1{,}5$$




P (216 ≤ X ≤ 223) = P (-2 ≤ Z ≤ 1,5) = F(1,5) – [1 – F(2)] = 0,9332 – (1 – 0,9772) = **0,9104**

**c)** $\hat{\hat{x}} : P\left(X > \hat{\hat{x}}\right) = 0{,}3$




$\hat{z} : F\left(\hat{z}\right) = 0{,}7$

$\hat{z} = 0{,}52$ F(0,52) = 0,6985

$\hat{\hat{x}} = \hat{z} \cdot \sigma_{\bar{x}} + \mu_{\bar{x}} = 0{,}52 \cdot 2 + 220 =$ **221,4**

**Esercizio 2**

I voti riportati da N = 5 studenti ad un esame sono: 26, 27, 28, 29, 30 (popolazione). Si verifichi la relazione tra i parametri ($\mu$ e $\sigma^2$) della popolazione e quelli della distribuzione campionaria relativa a tutti i possibili campioni di ampiezza n = 2, nelle tre ipotesi di estrazione:

a) con ripetizione
b) senza ripetizione;
c) in blocco.

**Soluzione**

Prima di tutto conviene determinare $\mu$ e $\sigma^2$:

$$\mu = \frac{26+27+28+29+30}{5} = 28$$

$$\sigma^2 = \frac{(26-28)^2 + (27-28)^2 + \cdots + (30-28)^2}{5} = 2$$

**a)**
I campioni con ripetizione di ampiezza n = 2 estraibili **con ripetizione** da una popolazione di n = 5 unità sono $N^n = 5^2 = 25$ e sono elencati di seguito:

| campione | media camp. | Campione | media camp. | campione | media camp. | campione | media camp. | campione | media camp. |
|---|---|---|---|---|---|---|---|---|---|
| 26; 26 | 26 | 27; 26 | 26,5 | 28; 26 | 27 | 29; 26 | 27,5 | 30; 26 | 28 |
| 26; 27 | 26,5 | 27; 27 | 27 | 28; 27 | 27,5 | 29; 27 | 28 | 30; 27 | 28,5 |
| 26; 28 | 27 | 27; 28 | 27,5 | 28; 28 | 28 | 29; 28 | 28,5 | 30; 28 | 29 |
| 26; 29 | 27,5 | 27; 29 | 28 | 28; 29 | 28,5 | 29; 29 | 29 | 30; 29 | 29,5 |
| 26; 30 | 28 | 27; 30 | 28,5 | 28; 30 | 29 | 29; 30 | 29,5 | 30; 30 | 30 |

La distribuzione di frequenza della media campionaria è:

| $\bar{x}_i$ | $n_i$ |
|---|---|
| 26 | 1 |
| 26,5 | 2 |
| 27 | 3 |
| 27,5 | 4 |
| 28 | 5 |
| 28,5 | 4 |
| 29 | 3 |
| 29,5 | 2 |
| 30 | 1 |
| | **25** |

La variabile casuale $\bar{x}$ è così definita:

| $\bar{x}_i$ | $P(\bar{x}_i)$ |
|---|---|
| 26 | 0,04 |
| 26,5 | 0,08 |
| 27 | 0,12 |
| 27,5 | 0,16 |
| 28 | 0,2 |
| 28,5 | 0,16 |
| 29 | 0,12 |
| 29,5 | 0,08 |
| 30 | 0,04 |
| | **1** |

Il valore atteso della v.c. $\bar{x}_i$ è:

$$E\left(\bar{x}\right) = \sum_i \bar{x}_i \cdot P\left(\bar{x}_i\right) = \left(26 \cdot 0,04\right) + \left(26,5 \cdot 0,08\right) + \cdots + \left(30 \cdot 0,04\right) = 28 = \mu_X$$

Mentre la varianza di $\bar{x}_i$ è:

$$\sigma^2_{\bar{x}} = E\left[\bar{x} - E\left(\bar{x}\right)\right] = \sum_i \left[\bar{x}_i - E\left(\bar{x}_i\right)\right]^2 \cdot P\left(\bar{x}_i\right) =$$

$$= \left(26 - 28\right)^2 \cdot 0,04 + \left(26,5 - 28\right)^2 \cdot 0,08 + \cdots + \left(30 - 28\right)^2 \cdot 0,04 = 1 = \frac{\sigma^2_X}{n}$$

**b)**

Nell'ipotesi di estrazione senza ripetizione il numero di possibili campioni è

$$N \times (N - 1) = 5 \times 4 = 20$$

(dai 25 del caso precedente bisogna eliminare quelli in cui i valori sono uguali, cioè [26; 26], [27; 27], ecc.), quindi:

| campione | media camp. | Campione | media camp. | campione | media camp. | campione | media camp. | campione | media camp. |
|---|---|---|---|---|---|---|---|---|---|
| 26; 27 | 26,5 | 27; 26 | 26,5 | 28; 26 | 27 | 29; 26 | 27,5 | 30; 26 | 28 |
| 26; 28 | 27 | 27; 28 | 27,5 | 28; 27 | 27,5 | 29; 27 | 28 | 30; 27 | 28,5 |
| 26; 29 | 27,5 | 27; 29 | 28 | 28; 29 | 28,5 | 29; 28 | 28,5 | 30; 28 | 29 |
| 26; 30 | 28 | 27; 30 | 28,5 | 28; 30 | 29 | 29; 30 | 29,5 | 30; 29 | 29,5 |

La distribuzione di frequenza della media campionaria è:

| $\bar{x}_i$ | $n_i$ |
|---|---|
| 26,5 | 2 |
| 27 | 2 |
| 27,5 | 4 |
| 28 | 4 |
| 28,5 | 4 |
| 29 | 2 |
| 29,5 | 2 |
| | **20** |

La variabile casuale $\bar{x}$ è così definita:

| $\bar{x}_i$ | $P(\bar{x}_i)$ |
|---|---|
| 26,5 | 0,1 |
| 27 | 0,1 |
| 27,5 | 0,2 |
| 28 | 0,2 |
| 28,5 | 0,2 |
| 29 | 0,1 |
| 29,5 | 0,1 |
| | **1** |

Il valore atteso della v.c. $\bar{x}_i$ è:

$$E\left(\bar{x}\right) = \sum_i \bar{x}_i \cdot P\left(\bar{x}_i\right) = \left(26,5 \cdot 0,1\right) + \left(27 \cdot 0,1\right) + \cdots + \left(29,5 \cdot 0,1\right) = 28 = \mu_X$$

Mentre la varianza di $\bar{x}_i$ è:

$$\sigma^2_{\bar{x}} = E\left[\bar{x} - E\left(\bar{x}\right)\right] = \sum_i \left[\bar{x}_i - E\left(\bar{x}_i\right)\right]^2 \cdot P\left(\bar{x}_i\right) =$$

$$= \left(26,5 - 28\right)^2 \cdot 0,1 + \left(27 - 28\right)^2 \cdot 0,1 + \cdots + \left(29,5 - 28\right)^2 \cdot 0,1 = 0,75 =$$

**c)**

Nel caso di estrazione in blocco, bisogna escludere i campioni costituiti dagli stessi elementi posti in ordine inverso. Il numero di campioni estraibili è:

$$\binom{N}{n} = \binom{5}{2} = \frac{5!}{2!(5-2)!} = \frac{5 \cdot 4 \cdot 3 \cdot 2 \cdot 1}{2 \cdot 1 \cdot 3 \cdot 2 \cdot 1} = 10$$

| campione | media camp. | campione | media camp. | campione | media camp. | campione | media camp. |
|---|---|---|---|---|---|---|---|
| 26; 27 | 26,5 | 27; 28 | 27,5 | 28; 29 | 28,5 | 29; 30 | 29,5 |
| 26; 28 | 27 | 27; 29 | 28 | 28; 30 | 29 | | |
| 26; 29 | 27,5 | 27; 30 | 28,5 | | | | |
| 26; 30 | 28 | | | | | | |

La distribuzione di frequenza della media campionaria è:

| $\bar{x}_i$ | $n_i$ |
|---|---|
| 26,5 | 1 |
| 27 | 1 |
| 27,5 | 2 |
| 28 | 2 |
| 28,5 | 2 |
| 29 | 1 |
| 29,5 | 1 |
| | **10** |

La variabile casuale $\bar{x}$ è così definita:

| $\bar{x}_i$ | $P(\bar{x}_i)$ |
|---|---|
| 26,5 | 0,1 |
| 27 | 0,1 |
| 27,5 | 0,2 |
| 28 | 0,2 |
| 28,5 | 0,2 |
| 29 | 0,1 |
| 29,5 | 0,1 |
| | **1** |

Il valore atteso della v.c. $\bar{x}_i$ è:

$$E(\bar{x}) = \sum_i \bar{x}_i \cdot P(\bar{x}_i) = (26,5 \cdot 0,1) + (27 \cdot 0,1) + \cdots + (29,5 \cdot 0,1) = 28 = \mu_X$$

mentre la varianza di $\bar{x}_i$ è:

$$\sigma_{\bar{x}}^2 = E\left[\bar{x} - E\left(\bar{x}\right)\right] = \sum_i \left[\bar{x}_i - E\left(\bar{x}_i\right)\right]^2 \cdot P\left(\bar{x}_i\right) =$$

$$= \left(26,5 - 28\right)^2 \cdot 0,1 + \left(27 - 28\right)^2 \cdot 0,1 + \cdots + \left(29,5 - 28\right)^2 \cdot 0,1 = 0,75 = \frac{\sigma_x^2}{n} \cdot \frac{N - n}{N - 1}$$

**Esercizio 3**

In un ristorante il conto (per persona) viene definito come somma di 2 componenti: un fisso pari a 20 € più una parte variabile pari all'equivalente in euro del 2% del peso del cibo consumato.
Se la quantità di cibo consumata ha una media pari a 800 grammi con uno scarto quadratico medio di 400, qual è la probabilità che 100 clienti paghino a testa in media:

a) fino a 35 €;
b) più di 38 €;
c) più di 40 €;
d) tra 36 e 39 €.

Inoltre:

e) Determinare la probabilità che un solo cliente paghi più di 40 euro

**Soluzione**

Se chiamiamo Y la v.c. prezzo pro-capite e X la quantità di cibo consumata, la variabile casuale cui dobbiamo fare riferimento per il calcolo delle probabilità richieste è il prezzo **medio** pro-capite $\bar{y}$ pagato in un insieme di 30 clienti.

Y è determinabile attraverso la trasformazione lineare di X in cui $\alpha = 20$ e $\beta = 0{,}02$:

$$Y = 20 + 0{,}02\ X.$$

Essendo note media $\mu_X$ e scarto quadratico medio $\sigma_X$ di X, è possibile determinare la media e la varianza di Y come:

$$\mu_Y = \alpha + \beta\mu_X = 20 + 0{,}02 \times 800 = 36$$

$$\sigma_Y^2 = \beta^2\sigma_X^2 = 0{,}02^2 \times 400^2 = 64$$

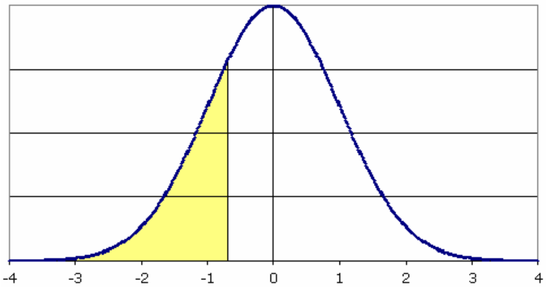
Pur non conoscendo la distribuzione di Y, per il teorema del limite centrale sappiamo che $\bar{y}$ si distribuisce normalmente, con media $\mu_Y$ e varianza $\frac{\sigma_Y^2}{n} = \frac{64}{30} = 2{,}13$ , quindi:

$$\bar{y} \sim N(36; 2{,}13)$$

Ora è possibile ricavare le probabilità cercate.

**a)** $P(\bar{y} \leq 35)$

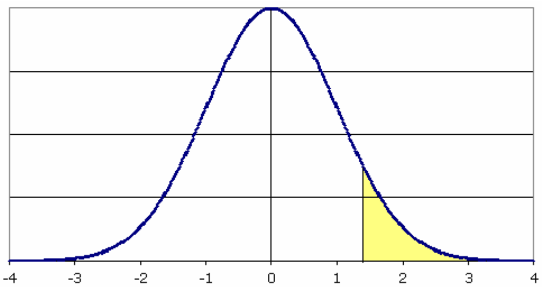
$$z_{35} = \frac{35-36}{8/\sqrt{30}} = -0,68$$




$P(\bar{y} \leq 35) = P(Z \leq -0{,}68) = 1 - F(0{,}68) = 1 - 0{,}7517 =$ **0,2483**

**b)** $P(\bar{y} > 38)$

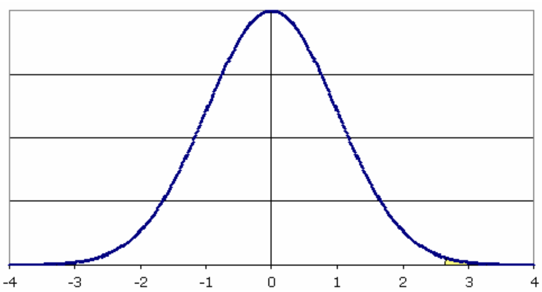
$$z_{38} = \frac{38-36}{8/\sqrt{30}} = 1,37$$




$P(\bar{y} > 38) = P(Z > 1{,}37) = 1 - F(1{,}37) = 1 - 0{,}9147 =$ **0,0853**

**c)** $P(\bar{y} > 40)$

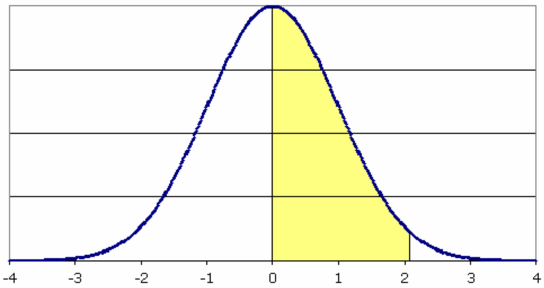
$$z_{40} = \frac{40-36}{8/\sqrt{30}} = 2,74$$



$P(\bar{y} > 40) = P(Z > 2{,}74) = 1 - F(2{,}74) = 1 - 0{,}9969 =$ **0,0031**

**d)** $P(36 \le \bar{y} \le 39)$

$$z_{39} = \frac{39-36}{8/\sqrt{30}} = 2{,}05$$




$P(36 \le \bar{y} \le 39) = P(0 \le Z \le 2{,}05) = F(2{,}05) - 0{,}5 = 0{,}9798 - 0{,}5 =$ **0,4798**

**e)**

Poiché non è nota la distribuzione della v.c. Y non è possibile determinare la probabilità che essa sia compresa in un qualsiasi intervallo.

**Nota**
Se Y si distribuisse normalmente, allora potremmo dire:

$$Y \sim N(36;64)$$

Quindi:

$$z_{40} = \frac{40-36}{8} = 0{,}5$$

$P(Y > 40) = P(Z > 0{,}5) = 1 - F(0{,}5) =$

Si noti che questo valore è minore di $P(\bar{y} > 40)$, a causa del fatto che la distribuzione della media campionaria è sempre meno variabile di quella della variabile casuale originaria.