UNIVERSITÀ IUSVE
Istituto Universitario Salesiano Venezia

**Prof. D. Marchioro**

# *Formulario di Metodologia della ricerca – Metodi quantitativi*

## Media

$$\bar{x} = \frac{\sum_{i=1}^{n} x_i * n_i}{n}$$

$\bar{x}$ = Media dei punteggi di x
$x_i$ = Punteggio grezzo/punteggio di x
$n_i$= Frequenza assoluta
$n$ = Numero totale dei casi (soggetti/osservazioni)

## Varianza campionaria

$$s^2 = \frac{\sum_{i=1}^{n}(x_i - \bar{x})^2}{n-1}$$

$s^2$ = Varianza
$\bar{x}$ = Media dei punteggi di x
$x_i$ = Punteggio grezzo/punteggio di x
$n$ = Numero totale dei casi (soggetti/osservazioni)

## Varianza di popolazione

$$\sigma^2 = \frac{\sum_{i=1}^{n}(x_i - \mu)^2}{N}$$

$\sigma^2$ = Varianza della popolazione
$\mu$ = Media della popolazione per la variabile x
$x_i$ = Punteggio grezzo/punteggio di x
$N$ = Numero totale dei casi (soggetti/osservazioni)

## Deviazione standard campionaria

$$s = \sqrt{\frac{\sum(x_i - \bar{x})^2}{n-1}}$$

$s$ = Deviazione standard
$\bar{x}$ = Media dei punteggi di x
$x_i$ = Punteggio grezzo/punteggio di x
$n$ = Numero totale dei casi (soggetti/osservazioni)

## Deviazione standard di popolazione

$$\sigma = \sqrt{\frac{\sum_{i=1}^{n}(x_i - \mu)^2}{N}}$$

$\sigma$ = Deviazione standard della popolazione
$\mu$ = Media della popolazione per la variabile x
$x_i$ = Punteggio grezzo/punteggio di x
$N$ = Numero totale dei casi (soggetti/osservazioni)

## Effect size per differenza tra due medie *(Cohen's d)*

$$d = \frac{\bar{x}_1 - \bar{x}_2}{\sigma_E}$$

$\bar{x}_1$= media del primo campione
$\bar{x}_2$= media del secondo campione
$\sigma_E$ = deviazione standard stimata della popolazione di ciascun gruppo:

$$\sigma_E = \sqrt{\frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}}$$

## *Cohen's K*

$$K = \frac{\sum A_{oss} - \sum A_{att}}{N - \sum A_{att}}$$

K = Indice K di Cohen
N = Totale delle osservazioni
$A_{oss}$ = Numero di accordi osservati $n_{ij}$
$A_{att}$= Numero di accordi attesi $n'_{ij}$ per una data cella:

$$A_{att} = \frac{Marginale\ di\ riga\ \times Marginale\ di\ colonna}{N}$$

IC= indice di concordanza:

$$IC = \frac{\sum A_{oss}}{N}$$

$IC_{caso}$= indice di accordo dovuto al caso:

$$IC_{caso} = \frac{\sum A_{att}}{N}$$

$IC_{vero}$= K corretto per il caso:

$$IC_{vero} = IC - IC_{caso}$$

## Guida per interpretazione di K e significatività di K

*(Landis & Koch, 1977)*

- $K < 0 =$ *disaccordo*
- $0 < K < .20 =$ *accordo pessimo*
- $.20 < K < .40 =$ *accordo modesto*
- $.40 < K < .60 =$ *accordo moderato*
- $.60 < K < .80 =$ *accordo buono*
- $.80 < K < 1.00 =$ *accordo eccellente*

Statistica test per il coefficiente K di Cohen:

$$z = \frac{K}{\sqrt{Var\,(K)}}$$

dove:

$$Var\,(K) = \frac{IC - IC^2}{N \cdot (1 - IC_{caso})^2}$$

## Costruzione del quadrato latino

Se le n prove sono un numero pari, è possibile creare esattamente n sequenze di condizioni che rispettino la condizione appena illustrata. Basta completare la serie:

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| 1 | 2 | $\boldsymbol{n}$ | 3 | $\boldsymbol{n-1}$ | 4 | $\boldsymbol{n-2}$ | 5 | $\boldsymbol{n-3}$ | **…** |

Per ottenere la prima sequenza, dalla quale poi si possono generare le altre semplicemente aggiungendo 1 alla cifra in ogni posizione.
Se il numero n di prove è dispari, non è possibile generare n sequenze che soddisfino le condizioni del quadrato latino bilanciato, ma è possibile costruirne 2n semplicemente generandone n con il metodo appena visto per il caso di un numero di prove pari, e poi costruendo le altre n invertendo una per una le prime.

## Distribuzione campionaria della media

Il valore atteso – media della distribuzione campionaria della media – è uguale alla media della distribuzione di probabilità della popolazione, relativamente alla quale la distribuzione campionaria è stata individuata:

$$\mathrm{E}(\bar{X}) = \mu$$

se indichiamo con $\mathrm{E}(\bar{X}) = \mu_{\bar{x}}$ la media della distribuzione campionaria della media, allora:

$$\boldsymbol{\mu_{\bar{x}} = \mu}$$

$$\boldsymbol{\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}}}$$

Quando la media campionaria segue una distribuzione normale, **possiamo standardizzarla** e la variabile casuale:

$$Z = \frac{\bar{X} - \mu}{\sigma_{\bar{X}}} = \frac{\bar{X} - \mu}{\sigma / \sqrt{n}}$$

## Intervalli di confidenza per la distribuzione delle medie campionarie

Ipotizzando di conoscere la media $\mu$ e la varianza $\sigma^2$ della popolazione, possiamo costruire l'intervallo di accettazione simmetrico:

$$\mu \pm z_{\alpha/2} \cdot \sigma_{\bar{X}}$$

$z_{\alpha/2} =$ valore z della normale standard che lascia nella coda destra la probabilità $\alpha/2$

L'intervallo casuale $L_1(X_1, X_2, \dots, X_n)$, $L_2(X_1, X_2, \dots, X_n)$ si definisce **intervallo di confidenza di livello $\boldsymbol{1-\alpha}$** per il parametro θ se contiene con probabilità $\boldsymbol{1-\alpha}$ il parametro ignoto θ della popolazione, ossia:

$$P\left[L_1\left(X_1, \dots, X_n\right) \le \theta \le L_2\left(X_1, \dots, X_n\right)\right] = 1 - \alpha$$

## Test t di Student

**Test-t per campioni indipendenti** (2 condizioni; soggetti diversi):

$$t = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$$

$d = (n_1 - 1) + (n_2 - 1)$

**Test-t per un campione** (confronto con un valore medio di riferimento):

$$t = \frac{\bar{X} - \mu}{s / \sqrt{n}}$$

$d = n - 1$

**Test-t per campioni appaiati** (disegno a misure ripetute con 2 campioni):

$$t = \frac{\bar{X}_{differenze} - 0}{s_{differenze} / \sqrt{n}}$$

dove:

$\bar{X}_{differenze}$= media delle differenze tra le due misurazioni appaiate;
$s_{differenze}$= dev. St. delle differenze tra le due misurazioni appaiate

$d = n - 1$ (se test/re-test)
$d = (n_1 - 1) + (n_2 - 1)$ (se 2 campioni appaiati – soggetti diversi)

<table>
<tr><th>Statistica test di Chi quadro</th><th>Indici di associazione binaria relativi derivanti dalla statistica di Chi quadro</th></tr>
<tr><td>

$$\chi^2 = \sum_{i=1}^{H}\sum_{j=1}^{K}\frac{\left(n_{ij} - n'_{ij}\right)^2}{n'_{ij}}$$

$\chi^2$ = Statistica test di chi quadro<br>
$n_{ij}$ = Frequenze osservate all’incrocio in corrispondenza della riga i-esima e della colonna j-esima<br>
$n'_{ij}$ = Frequenze attese all’incrocio in corrispondenza della riga i-esima e della colonna j-esima

$$n'_{ij} = \frac{Marginale\ di\ riga\ \times Marginale\ di\ colonna}{N}$$

$$d\ =\ gradi\ di\ libert\grave{a}\ =\ (r-1)\cdot(c-1)$$

dove:

$$r = numero\ di\ righe\ della\ tabella$$
$$c = numero\ di\ colonne\ della\ tabella$$

</td><td>

1) Indice di contingenza quadratica (Phi-quadro):

$$\Phi^2 = \frac{\chi^2}{N}$$

2) Indice V di Cramér:

$$V = \sqrt{\frac{\Phi^2}{\min\left[(H-1),(K-1)\right]}}$$

dove:

la dicitura $\mathbf{min}\left[(\boldsymbol{H}-\mathbf{1}),(\boldsymbol{K}-\mathbf{1})\right]$ indica il valore più basso tra (H-1) e (K-1), dove con H e K ci si riferisce, rispettivamente, al numero delle modalità del primo e del secondo carattere. Se i due caratteri presentano, rispettivamente, 4 e 3 modalità, al denominatore si dovrà inserire il valore $(3-1) = 2$.

</td></tr>
</table>