

## Analisi della Varianza (ad un criterio)

**ANOVA**  
(Analysis Of Variance)

## ANALISI DELLA VARIANZA - 2

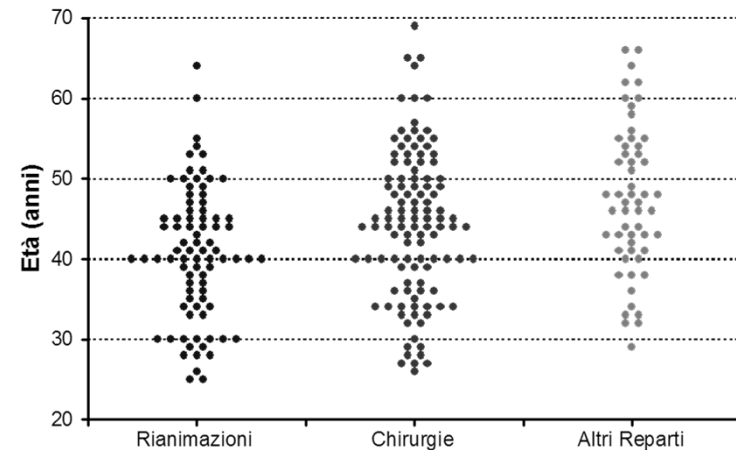
Oltre ad una media generale,  $\bar{x}$ , abbiamo k medie, una per ognuno dei singoli gruppi,  $\bar{x}_1, \bar{x}_2, \bar{x}_3, \dots, \bar{x}_k$ .

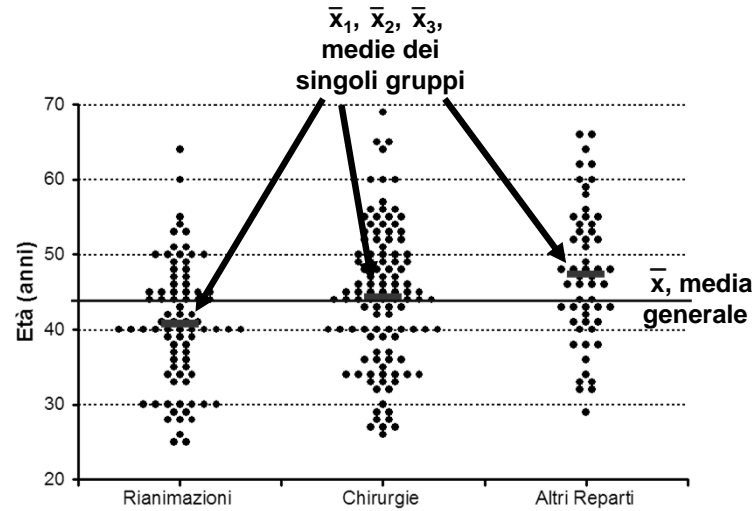
## ANALISI DELLA VARIANZA - 1

Abbiamo k gruppi, con un numero variabile di unità statistiche. Nella notazione classica, ogni unità statistica viene individuata da due numeri in posizione pedice: il primo indica il gruppo di appartenenza, e il secondo indica la posizione del soggetto all'interno del gruppo.

gruppo 1	gruppo 2	gruppo 3	.....	gruppo k
$x_{11}$	$x_{21}$	$x_{31}$	.....	$x_{k1}$
$x_{12}$	$x_{22}$	$x_{32}$	.....	$x_{k2}$
$x_{13}$	$x_{23}$	$x_{33}$	.....	$x_{k3}$
$x_{14}$	$x_{24}$	$x_{34}$	.....	$x_{k4}$
$x_{15}$	$x_{25}$	$x_{35}$	.....	$x_{k5}$
$x_{16}$	$x_{26}$	$x_{36}$	.....	$x_{k6}$
$x_{17}$	$x_{27}$	$x_{37}$	.....	$x_{k7}$
$x_{18}$		$x_{38}$	.....	$x_{k8}$
		$x_{39}$	.....	$x_{k9}$
$\bar{x}_1$	$\bar{x}_2$	$\bar{x}_3$	.....	$\bar{x}_k$

Età dei medici operanti nelle Rianimazioni, Chirurgie ed altri Reparti di un'Azienda Ospedaliera del Veneto





## ANALISI DELLA VARIANZA - 3

Ipotesi  $\begin{cases} H_0: \mu_1 = \mu_2 = \mu_3 = \dots = \mu_0 \\ H_1: \text{almeno una media differisce dalle altre} \end{cases}$

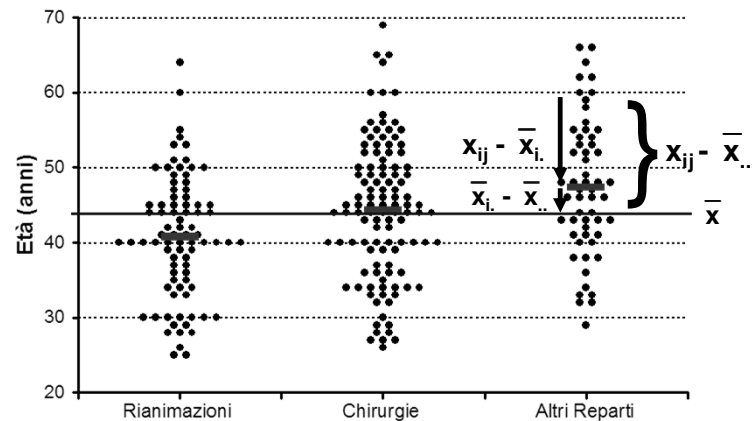
Per rispondere a questa domanda, possiamo fare tante t di Student, confrontando tutte le possibili coppie di medie?

NO, perché altrimenti avremmo un'inflazione (aumento abnorme) di  $\alpha$  (alfa), probabilità di errore del I tipo.

E' meglio quindi ricorrere ad un test globale, che confronti fra di loro tutti i gruppi:

**l'analisi della varianza.**

## SCOMPOSIZIONE DELLA DEVIANZA nell'Analisi della Varianza - 1



$x_{ij} - \bar{x}_{..}$  = scarto di una singola osservazione (valore *jesimo* del gruppo *iesimo*) dalla media generale

$\bar{x}_i - \bar{x}_{..}$  = scarto della media del gruppo *iesimo* dalla media generale

$x_{ij} - \bar{x}_i$  = scarto di una singola osservazione (valore *jesimo* del gruppo *iesimo*) dalla media del gruppo *iesimo*

## SCOMPOSIZIONE DELLA DEVIANZA nell'Analisi della Varianza - 2

Per una singola osservazione:

$$\begin{array}{ccc}
 \text{Variabilità} & & \text{Variabilità} \\
 \text{totale} & & \text{entro gruppi} \\
 (x_{ij} - \bar{x}_{..}) & = & (\bar{x}_{i.} - \bar{x}_{..}) + (x_{ij} - \bar{x}_{i.}) \\
 & & \text{Variabilità fra} \\
 & & \text{gruppi}
 \end{array}$$

Si può dimostrare che, per tutte le osservazioni:

$$\begin{array}{ccc}
 \text{Devianza} & & \text{Devianza} \\
 \text{totale, SST} & & \text{entro gruppi} \\
 \sum_{i,j} (x_{ij} - \bar{x}_{..})^2 & = & \sum_{i,j} (\bar{x}_{i.} - \bar{x}_{..})^2 + \sum_{i,j} (x_{ij} - \bar{x}_{i.})^2 \\
 & & \text{Devianza tra} \\
 & & \text{gruppi}
 \end{array}$$

I° STEP: definire l'ipotesi da verificare

$$\begin{cases}
 H_0: \mu_0 = \mu_1 = \dots = \mu_k \\
 H_1: \mu_i \neq \mu_j
 \end{cases}$$

almeno una media  
diversa dalle altre

II° STEP: definire la statistica test

III° STEP: calcolare il p-value

$$F = \frac{MS_{tra}}{MS_{entro}} \stackrel{H_0}{\sim} F_{(k-1), (N-k)}$$

distribuzione F di Snedecor  
con (k-1) gradi di libertà  
al numeratore e (N-k) gradi di  
libertà al denominatore

Se valgono le assunzioni di normalità e  
omoschedasticità

## Tavola ANOVA ad un criterio di classificazione

Fonte di variabilità	SS	df	MS
Tra gruppi	$\sum_{ij} (\bar{x}_{i.} - \bar{x}_{..})^2$	k - 1	$SS_{tra} / (k - 1)$
Entro gruppo	$\sum_{ij} (x_{ij} - \bar{x}_{i.})^2$	N - k	$SS_{entro} / (N - k)$
Totale	$\sum_{ij} (x_{ij} - \bar{x}_{..})^2$	N - 1	$SS_{tot} / (N - 1)$

VARIANZA DELLE MEDIE

↑

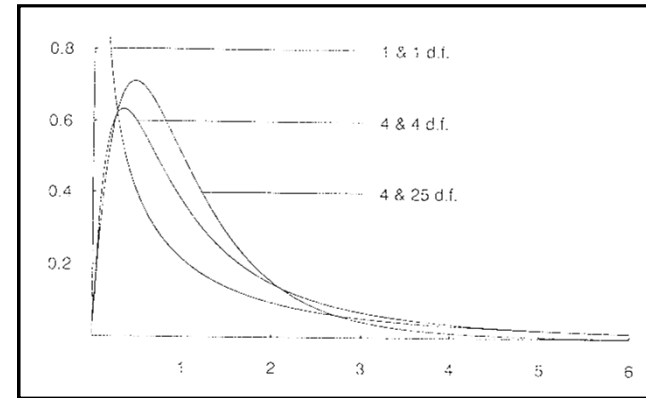
VARIANZA POOLED

Scomposizione della devianza campionaria per fonti di variabilità:

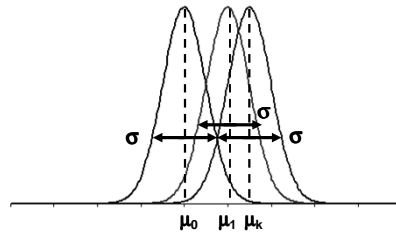
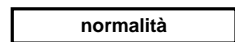
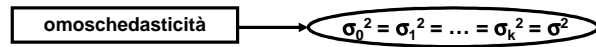
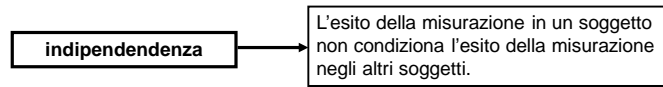
- Differenze tra gruppi:  $SS_{tra}$
- Errore campionario:  $SS_{entro}$

## Distribuzione F di Snedecor

$$F \sim F_{v_1, v_2}$$



## ASSUNZIONI SU CUI SI BASA L'ANALISI DELLA VARIANZA:



L'età dei medici (n=251) è significativamente diversa nelle Rianimazioni, nelle Chirurgie e negli altri Reparti?

fonte di variabilità	gradi di libertà	devianza	varianza	test F (significatività)
TRA gruppi	2	1546,10	773,05	9,779
ENTRO gruppi	248	19604,73	79,05	(P<0,001)
TOTALE	250	21150,84		

Si rifiuta l'ipotesi nulla: l'età differisce significativamente tra i vari Reparti.