CONFRONTO DI PIU' MEDIE

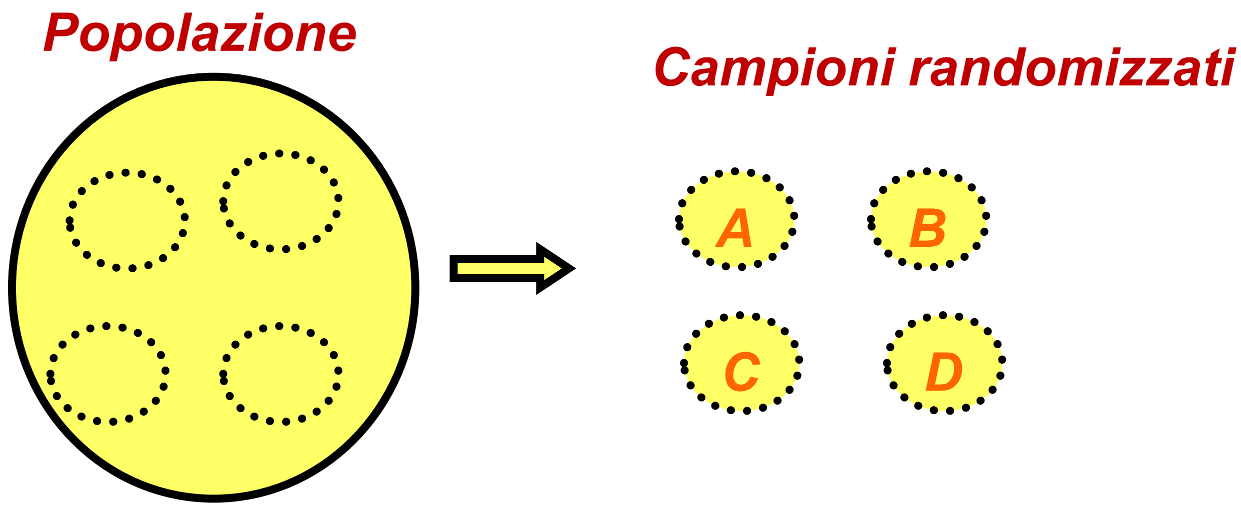
IL METODO DI

ANALISI DELLA VARIANZA

IL PROBLEMA

Supponiamo di voler studiare l'effetto di 4 diverse diete su un campione casuale di 24 cavie rilevando per ciascuna cavia il peso in gr..
Suddividiamo le 24 cavie in 4 gruppi, quanti sono i trattamenti, in modo casuale (random)

	Trattamenti			
Repliche	**A**	**B**	**C**	**D**
1	**716**	**634**	**668**	**728**
2	**761**	**676**	**723**	**786**
3	**684**	**678**	**710**	**754**
4	**792**	**682**	**762**	**789**
5	**838**	**726**	**744**	**836**
6	**774**	**694**	**732**	**842**



Dati

Si dispone del peso in grammi delle 24 cavie che sono assegnate casualmente ai 4 diversi trattamenti.

le osservazioni relative a ciascun trattamento, le possiamo indicare genericamente y_{ij}.

Con **$i=1,2,3,\ldots,k$ gruppi**

$j=1,2,3,4,\ldots,n_i$ osservazioni

Assunzioni

- Distribuzione della variabile : Gauss.
- Campioni indipendenti.
- Varianze omogenee.

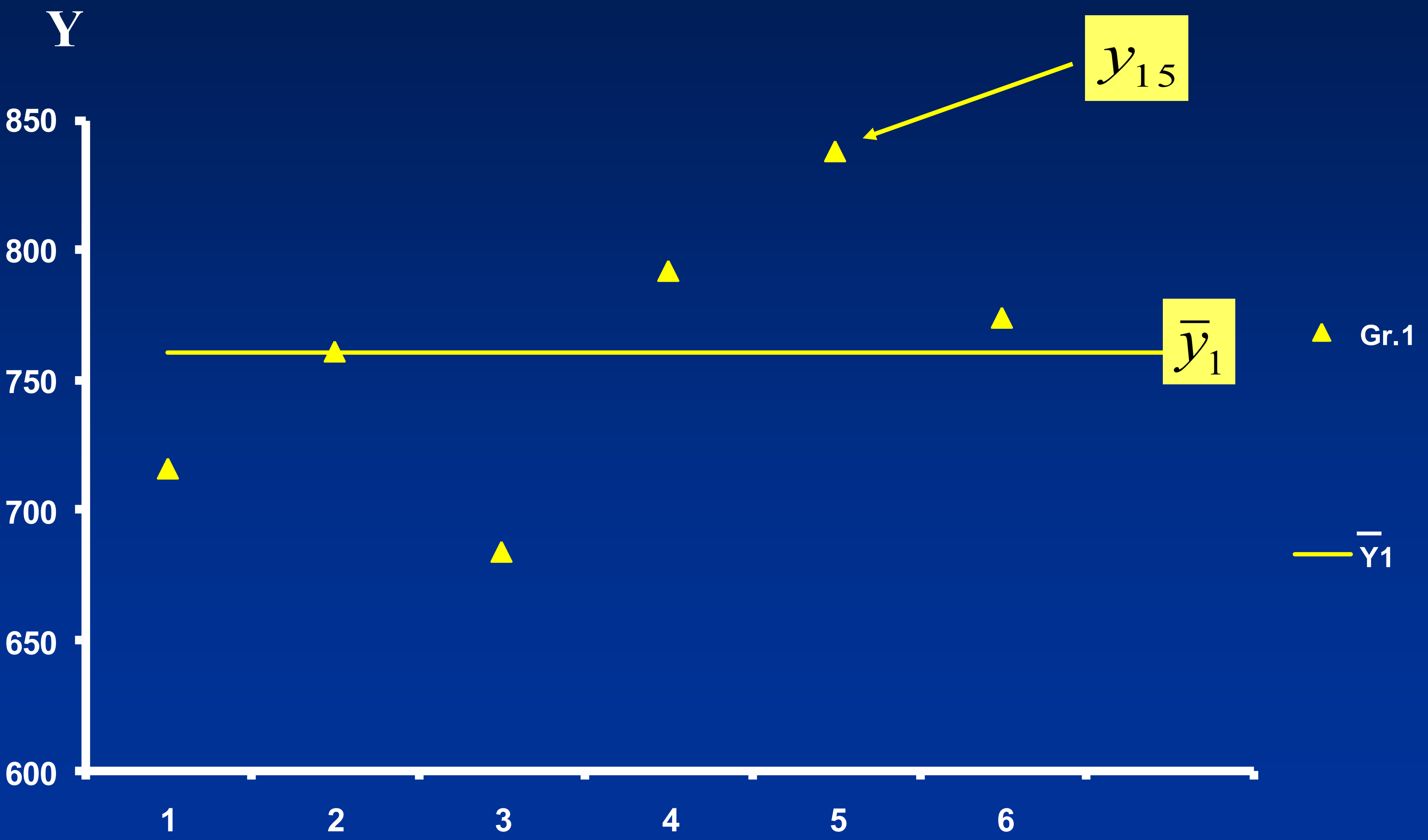
Ipotesi

$$\begin{cases} H_0: \quad \mu_a=\mu_b=\mu_c=\mu_d=\ldots=\mu_k=\mu \\ H_1: \quad \mu_r\neq\mu_s \end{cases}$$

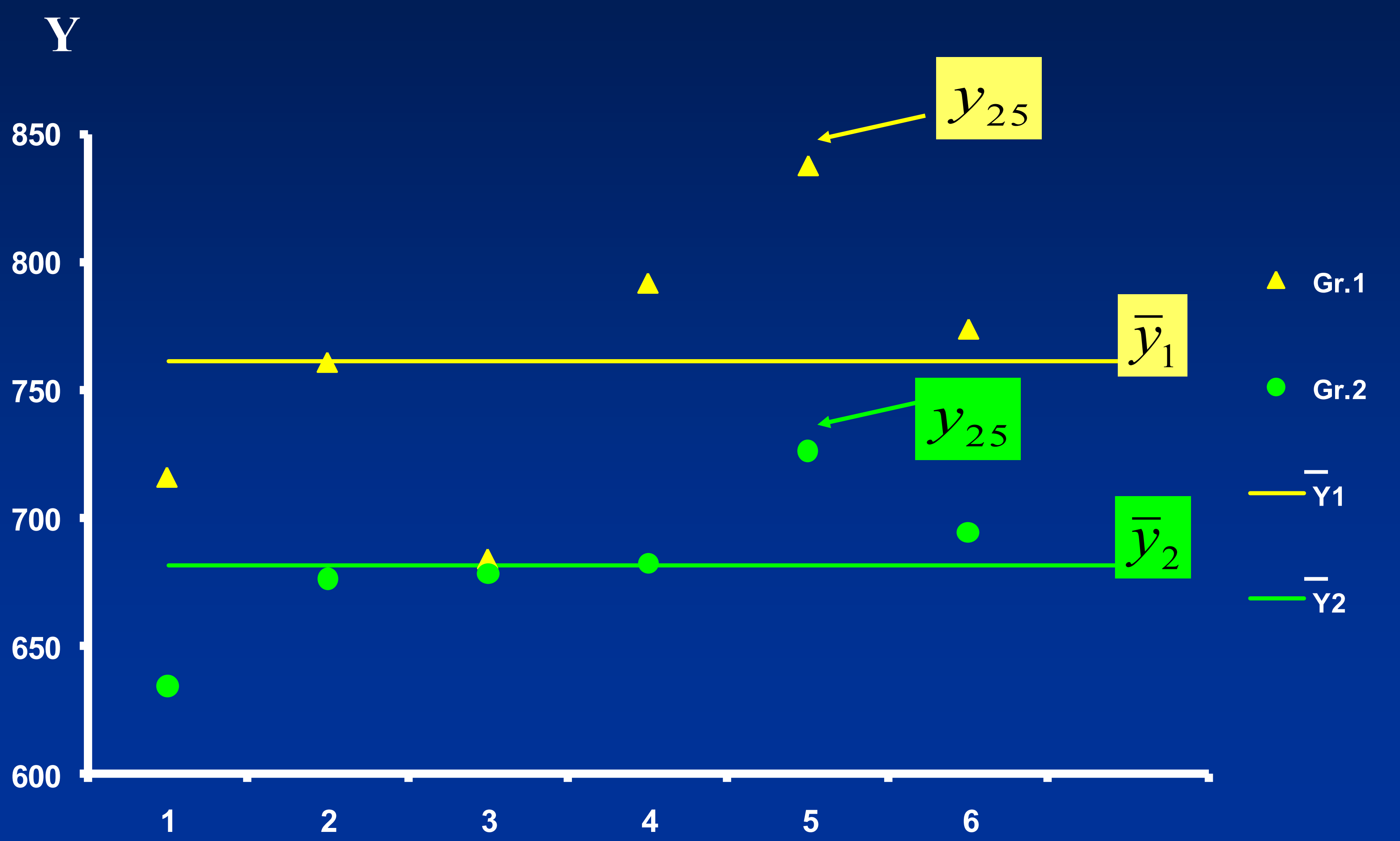
L'ipotesi da saggiare è che tutti i trattamenti siano uguali contro un ipotesi alternativa che almeno due siano diversi tra loro.

Repliche	1	2	…	i	…	k	
1	y_{11}	y_{21}	…	y_{i1}	…	y_{k1}	
2	y_{12}	y_{22}	…	y_{i2}	…	y_{k2}	
…	…	…	…		…	…	
j	y_{1j}	y_{2j}	…	y_{ij}	…	y_{kj}	
…	…	…	…	…	…	…	
n_i	y_{1n1}	y_{2n2}	…	y_{ini}	…	y_{knk}	
$\Sigma y_{ij}=T_i$	T_1	T_2	…	T_i	…	T_k	T
$\bar{y}_i$	$\bar{y}_1$	$\bar{y}_2$	…	$\bar{y}_i$	…	$\bar{y}_k$	$\bar{y}$
T^2_i/n_i	T^2_1/n_1	T^2_2/n_2	…	T^2_i/n_i	…	T^2_k/n_k	$\Sigma\, (T^2_i/n_i)$
$\Sigma y^2_{ij}=S_i$	S_1	S_2	…	S_i	…	S_k	S

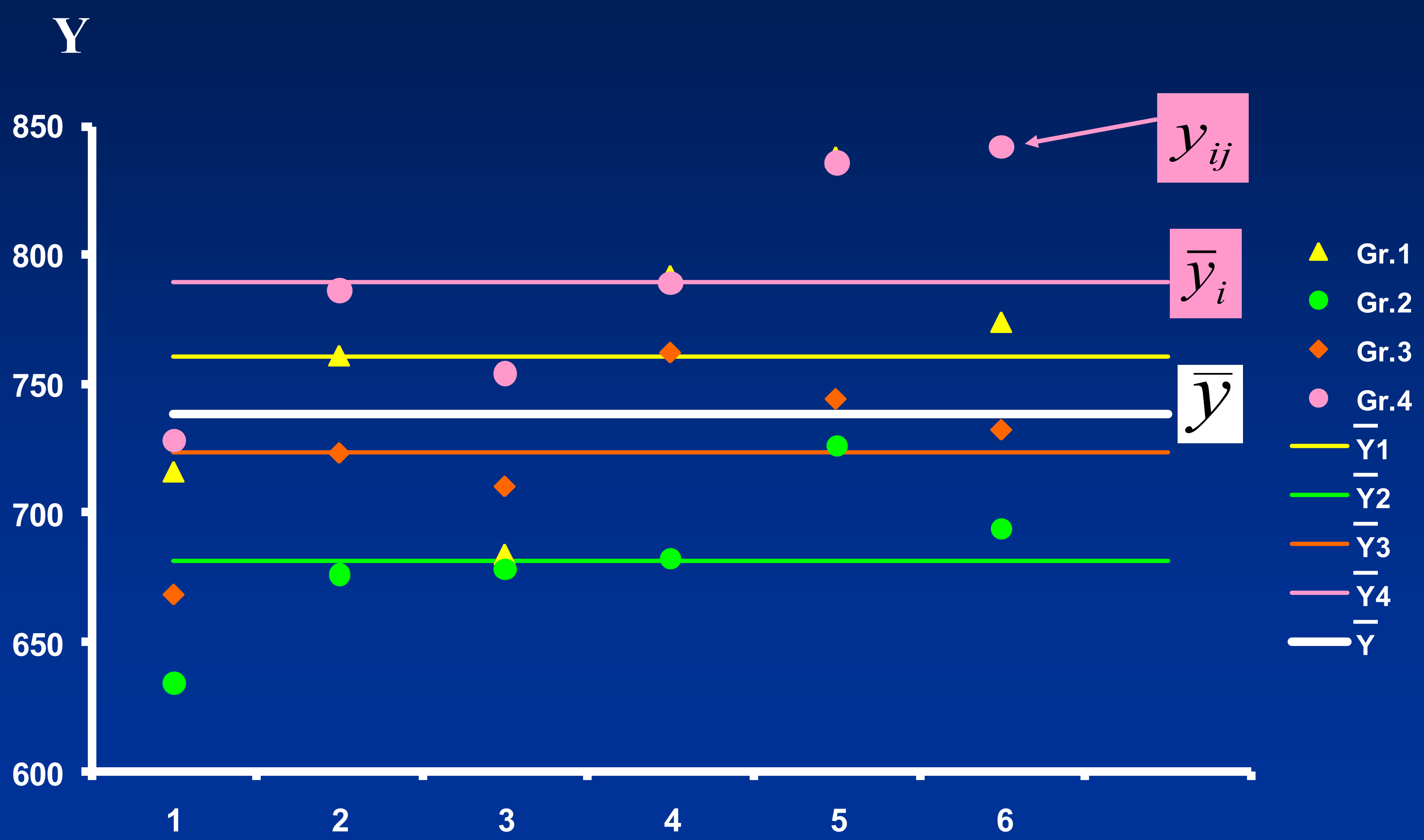
ANALISI DELLA VARIANZA



ANALISI DELLA VARIANZA

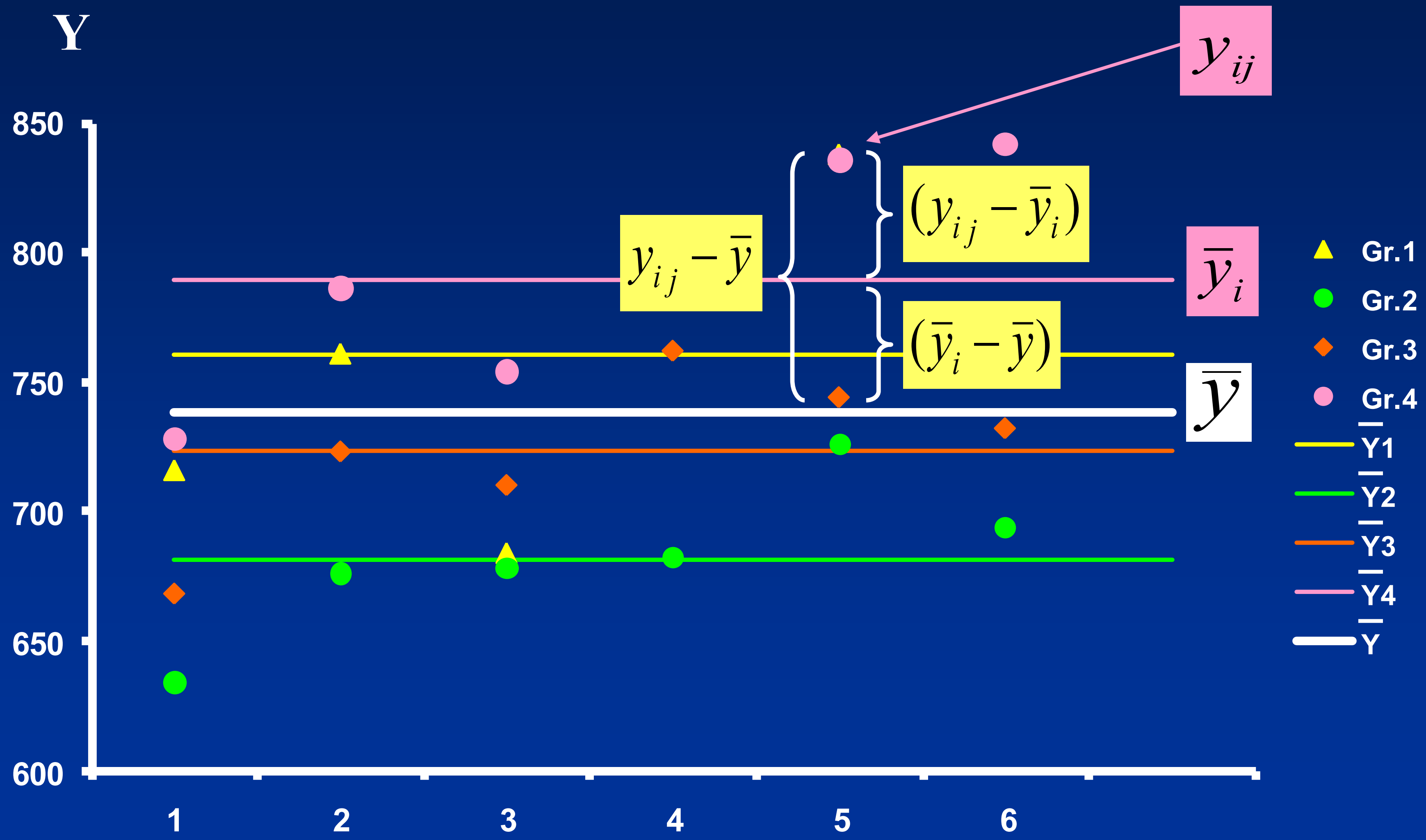


ANALISI DELLA VARIANZA



ANALISI DELLA VARIANZA

$$y_{ij} - \bar{y} = (y_{ij} - \bar{y}_i) + (\bar{y}_i - \bar{y})$$



Costruzione della Statistica test

$$\boxed{y_{ij} - \bar{y} = (y_{ij} - \bar{y}_i) + (\bar{y}_i - \bar{y})}$$

$$\boxed{\sum_{i=1}^{k}\sum_{j=1}^{n_i}\left(y_{ij}-\bar{y}\right)^2 = \sum_{i=1}^{k}\sum_{j=1}^{n_i}\left(y_{ij}-\bar{y}_i\right)^2 + \sum_{i=1}^{k}\sum_{j=1}^{n_i}\left(\bar{y}_i-\bar{y}\right)^2}$$

devianza totale = devianza entro gruppi + dev. tra gruppi.

G.l. N-1 N-k k-1

Con i=1,2,3,…,k gruppi
j=1,2,3,4,…,ni osservazioni

La statistica test consiste nel valutare quanta parte della variabilità totale è attribuibile alla differenza tra i trattamenti.

☞ **La statistica test sarà :**

$$F = \frac{\sum_{i=1}^{k}\sum_{j=1}^{n_i}\left(\bar{y}_i - \bar{y}\right)^2 \Big/ k-1}{\sum_{i=1}^{k}\sum_{j=1}^{n_i}\left(y_{ij} - \bar{y}_i\right)^2 \Big/ N-k} = \frac{\text{varianza tra gruppi (g.l. k - 1)}}{\text{varianza entro gruppi (g.l. N - k)}}$$

Distribuzione della statistica test

La distribuzione della statistica test è F-Fisher, dipende dai gradi di libertà del numeratore (k-1) e del denominatore (n-K)

Regola di decisione

Fisso α (livello di significatività, errore di I tipo) accettabilmente basso (0,05) e in corrispondenza dei gradi di libertà del numeratore e del denominatore si determina un valore tabulato che delimita la zona di rifiuto.

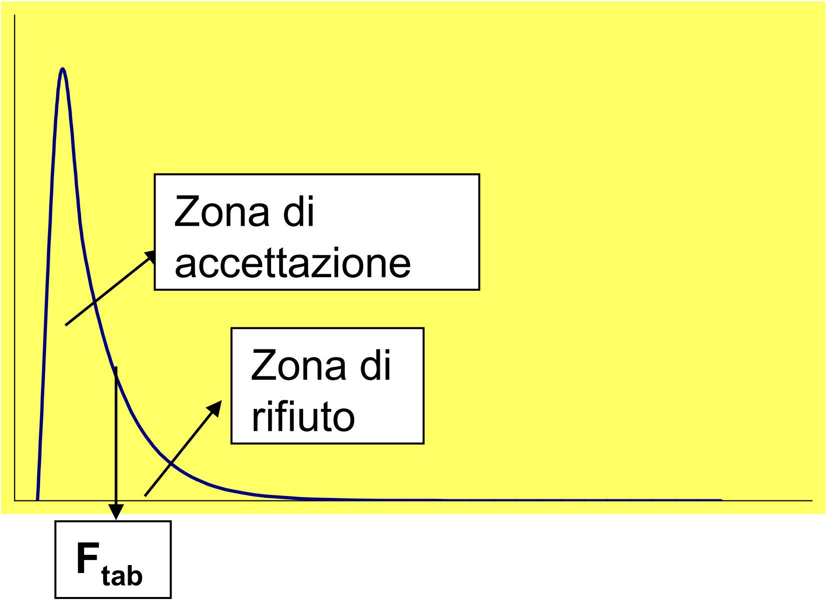


Tavola per il calcolo dell'analisi della varianza (ANOVA)

Sorgenti di variazione	Devianze	Gradi di libertà	Varianza	Fcalcolato
Trattamento	**Dev. tra gruppi**	**k-1**	**S_f=Dev. tra gruppi/k-1**	**$F=S_f$ / S_e**
Errore	**Dev. entro gruppi**	**N-k**	**S_e=Dev. entro gruppi/N-k**	
Totale	**Dev. totale**	**N-1**	**Dev. tot/N-1**	

$$SSqTOT=\sum_i\sum_j(y_{ij}-\bar{y})^2=\sum_i\sum_j y_{ij}^2-\frac{\left(\sum_i\sum_j y_{ij}\right)^2}{N}=S-\frac{T^2}{N}$$

$$SSqENTRO=\sum_i\sum_j(y_{ij}-\bar{y}_i)^2=\sum_i\sum_j y_{ij}^2-\sum_i\frac{\left(\sum_j y_{ij}\right)^2}{n_i}=S-\sum_i\frac{T_i^2}{n_i}$$

$$SSqTRA=\sum_i\sum_j(\bar{y}_i-\bar{y})^2=\sum_i\frac{\left(\sum_j y_{ij}\right)^2}{n_i}-\frac{\left(\sum_i\sum_j y_{ij}\right)^2}{N}=\sum_i\frac{T_i^2}{n_i}-\frac{T^2}{N}$$

$S = \Sigma\Sigma y_{ij}^2$ $T = \Sigma\Sigma\, y_{ij}$ $T_i = \Sigma y_{ij}$ N = num. delle osservazioni

Nel nostro esempio la tabella precedente diventa:

	Trattamento				
Repliche	**A**	**B**	**C**	**D**	
1	716	634	668	728	
2	761	676	723	786	
3	684	678	710	754	
4	792	682	762	789	
5	838	726	744	836	
6	774	694	732	842	
$\Sigma y_{ij}=T_i$	4565	4090	4339	4735	T =**17729**
$\bar{y}_i$	760,8	681,7	723,2	789,2	$\bar{y}$ =**738,7**
T^2_i/n_i	3473204	2788017	3137820	3736704	$\Sigma(T^2_i/n_i)$ =**13135745**
$S_i=\Sigma y^2_{ij}$	3488217	2792452	3143057	3746677	S =**13170403**

Applicando le formule precedenti si ottiene:

devianza totale = S – (T^2 / N)=13170403-13096560= 73843

devianza entro gruppi = S –Σ(T^2_i / n_i)=13170403-13135745=34658

devianza tra gruppi = Σ (T^2_i / n_i) - (T^2 / N)=13135745-13096560=39185

Da cui:

TAVOLA DI ANALISI DELLA VARIANZA

Sorgenti	Devianze	Gradi di libertà	Varianze	**F**
Tra gruppi	39185	3	13061,67	**7,54**
Entro gruppi	34658	20	1732,90	
Totale	73843	23		

Varianza Residua

Conclusioni

Dato che il valore calcolato F = 7,54
è maggiore del valore tabulato
per $\alpha=0{,}05$ $F_{3,20} = 3{,}10$
allora posso rifiutare H_0

⇓

Almeno due diete sono differenti tra loro.

TEST DI BARTLETT
PER L'OMOGENEITA' DELLE VARIANZE

Dati k gruppi di osservazioni, le corrispondenti varianze devono tutte risultare stime della stessa varianza incognita della popolazione, e in base a questo presupposto la varianza di errore può essere ottenuta come media ponderata delle k varianze stimate in ciascun gruppo

$$H_0: \quad \sigma_1^2 = \sigma_2^2 = \ldots\ldots = \sigma_k^2 = \sigma^2$$

contro l'alternativa che almeno uno dei valori $\boldsymbol{\sigma_i^2}$ risulti diverso dai rimanenti k-1

La statistica test

$$\frac{A}{B}=\frac{2.3026\bullet\left[\left(\sum n_i-k\right)\bullet\mathbf{lg}_{10}S^2-\left(\sum n_i-1\right)\bullet\mathbf{lg}_{10}S_i^2\right]}{1+\dfrac{\sum_i\dfrac{1}{n_i-1}-\dfrac{1}{\sum_i n_i-k}}{3(k-1)}}$$

La distribuzione della statistica test è $\chi 2$ con k-1 gradi di libertà
S^2 è la varianza residua e S^2_i è la varianza del gruppo i-esimo

Si rifiuta l'ipotesi se il valore calcolato di A/B risulta maggiore del valore tabulato $\chi 2$ con k-1 gradi di libertà

$$\frac{A}{B} = \frac{2.3026 \bullet \left[\left(\sum n_i - k\right) \bullet \mathbf{lg}_{10} S^2 - \left(\sum n_i - 1\right) \bullet \mathbf{lg}_{10} S_i^2\right]}{1 + \dfrac{\sum_i \dfrac{1}{n_i - 1} - \dfrac{1}{\sum_i n_i - k}}{3(k-1)}}$$

$$\left(\sum n_i - k\right) = 20 \qquad \mathbf{lg}_{10} S^2 = \mathbf{lg}_{10} 1632{,}9 = 3{,}21$$

$$\left(\sum n_i - 1\right) \bullet \mathbf{lg}_{10} S_i^2 = 17{,}07 + 14{,}73 + 15{,}10 + 16{,}49 = 63{,}4$$

$$1 + \frac{\sum_i \dfrac{1}{n_i - 1} - \dfrac{1}{\sum_i n_i - k}}{3(k-1)} = 1 + \frac{\sum_i \dfrac{1}{6-1} - \dfrac{1}{24-4}}{3(4-1)} = 1{,}049$$

$$\frac{A}{B} = \frac{2{,}3026 \bullet [20 \bullet 3{,}21 - 63{,}4]}{1{,}049} = 1{,}79$$

χ2 con 3 gradi di libertà = 7,81>A/B=1,79 ⇒Varianze omogenee

CONFRONTI MULTIPLI

Vanno effettuati quando l'ANOVA ha portato al rifiuto dell'ipotesi nulla

Nella valutazione della significatività bisogna stare attenti poichè questi tests possono comportare una distorsione dell'errore di I tipo così come della potenza del test

Procedura LSD di Fisher

esegue tutti i possibili test t

(Least Significant Difference = Minima differenza significativa)

Se eseguita solo nel caso in cui l'ipotesi nulla su tutte le medie viene rifiutata risulta piuttosto efficiente nel mantenere un ragionevole grado di controllo sui falsi errori.

$$\mathbf{LSD = t_{\alpha/2}\sqrt{(2\ Var\ Residua/n)}}$$

$t_{\alpha/2}$ Valore della t-Student con gradi di libertà N-k

- **Si ordinano le medie dalla più piccola alla più grande**
- **Si confronta la media più grande con la più piccola**
  - calcolando la differenza
  - confrontando la differenza ottenuta con il valore LSD

⬇

Se la differenza supera il valore LSD
si conclude che le due medie sono diverse

Si procede confrontando la più grande con la seconda più piccola e via di seguito

Nessuna coppia di medie può essere dichiarata significativamente diversa, se si trova all'interno di un'altra coppia già dichiarata non differente

$$\textbf{LSD} = t_{\alpha/2}\sqrt{(2\ \text{Var Residua}/n)} = 2.086\sqrt{(2\text{x}1732.90/6)} = 50.14$$

	Gr.B	Gr.C	Gr.A	Gr.D
$\bar{y}_i$	681.7	723.2	760.8	789.2

Dif1 = 789.2 – 681.7 = 107.5 **> 50,14** ⇨ medie differenti D-B

Dif2 = 789.2 – 723.2 = 66.0 **> 50,14** ⇨ medie differenti D-C

Dif3 = 789.2 – 760.8 = 28.4 < 50,14 ⇨ medie uguali D-A

Dif4 = 760.8 – 681.7 = 79.1 **> 50.14** ⇨ medie differenti A-B

Dif5 = 760.8 – 723.2 = 37.6 < 50.14 ⇨ medie uguali A-C

Dif6 = 723.2 – 681.7 = 41.5 < 50.14 ⇨ medie uguali C-B

Test t multiplo di Bonferroni

E' basato sulla costruzione degli intervalli di confidenza per la differenza delle medie poste a confronto

$$\left[\left(\bar{y}_r - \bar{y}_s \right) - t_{1-\frac{\alpha}{2m}, g.l.} \sqrt{\frac{2S^2_{residua}}{n}} \leq \mu_r - \mu_s \leq \left(\bar{y}_r - \bar{y}_s \right) + t_{1-\frac{\alpha}{2m}, g.l.} \sqrt{\frac{2S^2_{residua}}{n}} \right]$$

m rappresenta il numero di intervalli predeterminati per il confronto
g.l sono i gradi di libertà della varianza residua ottenuti nell'ANOVA
α è il livello di significatività stabilito per l'ANOVA

Se l'intervallo ottenuto non contiene lo zero le medie poste a confronto possono ritenersi diverse

$$\left[\left(\bar{y}_r - \bar{y}_s\right) - t_{1-\alpha/2m,g.l.}\sqrt{\frac{2S^2_{residua}}{n}} \le \mu_r - \mu_s \le \left(\bar{y}_r - \bar{y}_s\right) + t_{1-\alpha/2m,g.l.}\sqrt{\frac{2S^2_{residua}}{n}}\right]$$

$\bar{y}_i$	681.7	723.2	760.8	789.2	$\alpha/12 = 0.004$

Int 1-4 107.5 ± 2.84 x 24.03 = 107.5 ± 68.25 ⇒ **38.75 --- 175.25**

Int 2-4 66.0 ± 2.84 x 24.03 = 66.0 ± 68.25 ⇒ - 2.25 --- 134.25

Int 3-4 28.4 ± 2.84 x 24.03 = 28.4 ± 68.25 ⇒ - 39.85 --- 96.65

Int 1-3 79.1 ± 2.84 x 24.03 = 79.1 ± 68.25 ⇒ **10.85 --- 147.35**

Int 2-3 37.6 ± 2.84 x 24.03 = 37.6 ± 68.25 ⇒ - 30.65 --- 105.85

Int 1-2 41.5 ± 2.84 x 24.03 = 41.5 ± 68.25 ⇒ - 26.75 --- 109.75

Test del campo di variazione multiplo di Duncan

Il test pur essendo conservativo sui livelli di significatività non provoca una elevata perdita di potenza

✓Si pongono in ordine crescente le medie dalla più piccola alla più grande

✓Il test di Duncan tiene conto di quante medie (r) sono presenti tra quelle poste a confronto e definisce ***livello di protezione***

la quantità $(1-\alpha)^{r-1}$ per due medie separate da r passi e il livello di significatività viene approssimato dal valore $1-(1-\alpha)^{r-1}$

✓Il risultato è un diverso sistema di moltiplicatori per

$\sqrt{\textbf{Var Residua/n}}$ utili per calcolare la statistica **Critical Range**.

⬇

Se la differenza fra le medie è maggiore del Critical Range si rifiuta l'ipotesi di uguaglianza delle medie

QUANDO NON E’ POSSIBILE FARE L’ASSUNZIONE DELLA DISTRIBUZIONE GAUSSIANA SULLA VARIABILE, IL CONFRONTO DI PIU’ GRUPPI SI EFFETTUA CON IL METODO DELL’ANALISI DELLA VARIANZA NON PARAMETRICA

IL PROBLEMA

Si vuole studiare l'effetto di due farmaci sul tempo di reazione ad un certo stimolo su animali da laboratorio; il terzo gruppo è quello di controllo.

Possiamo affermare che i tre campioni si differenziano rispetto ai tempi di reazione (misurati in secondi)?

GRUPPO 1	GRUPPO 2	GRUPPO 3
17	8	2
20	7	5
40	9	4
31	8	3
35		

ANALISI DELLA VARIANZA NON PARAMETRICA

TEST DI KRUSKAL-WALLIS

Le n_1, n_2, …, n_k osservazioni provenienti da k campioni sono aggregate in un'unica serie di dati di dimensione n e messe in ordine crescente

Alle n osservazioni vengono assegnati i ranghi. Quando due o più osservazioni hanno lo stesso valore, ad ogni osservazione viene assegnata la media dei ranghi di tutte le osservazioni con lo stesso valore

I ranghi assegnati alle osservazioni in ognuno dei k gruppi vengono sommati tra loro ottenendo k somme dei ranghi

Sui nostri dati

Gr I	R1	Gr. II	R2	Gr III	R3
17	**9**	8	**6,5**	2	**1**
20	**10**	7	**5**	5	**4**
40	**13**	9	**8**	4	**3**
31	**11**	8	**6,5**	3	**2**
35	**12**				
	55		**26**		**10**

I valori in **nero** sono i dati originari

I valori in **blu** rappresentano i ranghi

I valori in **rosso** sono le somme dei ranghi

La statistica test

$$H = \frac{12}{n(n+1)} \sum_{j=1}^{k} \frac{R_j^2}{n_j} - 3(n+1)$$

Dove:

K= numero dei campioni

nj = numero di osservazioni nel j-esimo campioni

n = numero totale delle osservazioni

Rj =somma dei ranghi nel j-esimo campioni

Quando ci sono tre campioni e 5 o meno osservazioni in ogni campione, la significatività di H viene determinata usando le tavole di Kruskal-Wallis.

Quando ci sono più di 5 osservazioni in uno o più campioni, H viene confrontato con i valori tabulati del χ^2 con k-1 gradi di libertà.

Se ci sono osservazioni con il medesimo valore (ties) bisogna correggere la statistica H

$$1-\sum T \Big/ n^3 - n$$

$$T = t^3 - t$$

$$H_{corr} = \frac{H}{1-\left(\dfrac{\sum T}{n^3 - n}\right)}$$

Per il nostro insieme di dati:

$$H=\frac{12}{13(13+1)}\left[\frac{55^2}{5}+\frac{26^2}{4}+\frac{10^2}{4}\right]-3(13+1)=10.68$$

Per α=0,009
Htab=7.76

Rifiuto l'ipotesi nulla

I tempi di reazione sono diversi

Nel nostro caso poiché ci sono valori e ranghi uguali si ha:

Correzione

$$1-\sum T \Big/ n^3 - n$$

$$T = t^3 - t$$

$T=t^3-t=2^3-2=6$

Correzione = 1 - (6 / 13^3-13) = 0,9973

$$H_{corr} = \frac{H}{1-\dfrac{\sum T}{n^3-n}} = \frac{10.68}{0.9973} = 10.71$$

ANALISI DELLA VARIANZA
A
BLOCCHI RANDOMIZZATI COMPLETI

Il blocco rappresenta un insieme di dati omogenei come misure successive eseguite in tempi diversi sullo stesso individuo, animali appartenenti alla stessa nidiata,ecc.

Il modello è:

$$y_{ij} = \mu + \alpha_i + \beta_j + e_{ij}$$

Dove:

	μ	è l'effetto dovuto alla media generale
	α_i	è l'effetto dovuto al blocco
	βj	è l'effetto dovuto al trattamento o gruppo
	e_{ij}	è l'errore casuale

DATI

COLONNA o Gruppo

RIGA o BLOCCO

	1	2	…	j	…	c	TOTALE	Media R_i/c
1	y_{11}	y_{12}	…	y_{1j}	…	y_{1c}	R_1	$\bar{y}_{1.}$
2	y_{21}	y_{22}	…	y_{2i}	…	y_{1c}	R_2	$\bar{y}_{2.}$
…	…	…	…		…	…		
i	y_{i1}	y_{i2}	…	y_{ij}	…	y_{ic}	R_i	$\bar{y}_{i.}$
…	…	…	…	…	…	…		
r	y_{r1}	y_{r2}	…	y_{ri}	…	y_{rc}	R_r	$\bar{y}_{r\cdot}$
TOTALE	C_1	C_2		C_J		C_C	T	
Media C_J/r	$\bar{y}_{.1}$	$\bar{y}_{\cdot 2}$	…	$\bar{y}_{\cdot j}$	…	$\bar{y}_{\cdot c}$		$T/N= \bar{y}$

L'ipotesi nulla H_0 è formulata per verificare sia l'effetto del blocco che quello del gruppo o tempo ripetuto.

$$\begin{cases} \mathbf{H_0:} & \boldsymbol{\alpha_i = 0} & \textbf{per ogni i} \\ \mathbf{H_1:} & \boldsymbol{\alpha_i \neq 0} & \textbf{per ogni i} \end{cases}$$

$$\begin{cases} \mathbf{H_0:} & \boldsymbol{\beta_j = 0} & \textbf{per ogni j} \\ \mathbf{H_1:} & \boldsymbol{\beta_j \neq 0} & \textbf{per ogni j} \end{cases}$$

Costruzione della Statistica

$$y_{ij} - \bar{y}_{..} = (\bar{y}_{i.} - \bar{y}_{..}) + (\bar{y}_{.j} - \bar{y}_{..}) + (y_{ij} - \bar{y}_{i.} - \bar{y}_{.j} + \bar{y}_{..})$$

$$\sum (y_{ij} - \bar{y}_{..})^2 = \sum (\bar{y}_{i.} - \bar{y}_{..})^2 + \sum (\bar{y}_{.j} - \bar{y}_{..})^2 + \sum (y_{ij} - \bar{y}_{i.} - \bar{y}_{.j} + \bar{y}_{..})^2$$

Dev.TOTALE = Dev. Tra soggetti + Dev.Tra gruppi + Dev.Residua

G.l. (r x c)-1 r-1 c-1 (r-1)x(c-1)

dove N= r x c

STATISTICA TEST

$$F_{blocco} = \sum (\bar{y}_{i.} - \bar{y}_{..})^2 / \sum (y_{ij} - \bar{y}_{i.} - \bar{y}_{.j} + \bar{y}_{..})^2$$

$$F_{gruppo} = \sum (\bar{y}_{.j} - \bar{y}_{..})^2 / \sum (y_{ij} - \bar{y}_{i.} - \bar{y}_{.j} + \bar{y}_{..})^2$$

DISTRIBUZIONE DELLA STATISTICA TEST

Entrambe le statistiche hanno una distribuzione F di Fisher

F_{blocco} con gradi di libertà r-1 e (r-1)(c-1)

F_{gruppo} con gradi di libertà c-1 e (r-1)(c-1)

REGOLA DI DECISIONE

Si rifiuta l'ipotesi se la F calcolata risulta maggiore della F tabulata con gli opportuni gradi di libertà

Formule da utilizzare per i calcoli

$$Totale = \sum (y_{ij} - \bar{y}_{..})^2 = \sum y_{ij}^2 - T^2/N$$

$$Tra\ soggetti = \sum (\bar{y}_{i.} - \bar{y}_{..})^2 = \sum R_i^2/c - T^2/N$$

$$Tra\ gruppi = \sum (\bar{y}_{.j} - \bar{y}_{..})^2 = \sum C_i^2/r - T^2/N$$

$$Residua = \sum (y_{ij} - \bar{y}_{i.} - \bar{y}_{.j} + \bar{y}_{..})^2 =$$

$$= Totale - Tra\ soggetti - Tra\ gruppi$$

Tavola di Analisi della Varianza

ANOVA

Sorgenti di variazione	devianze	Gradi di libertà	varianze	F
Tra soggetti	$\Sigma(\bar{y}_{i.} - \bar{y}_{..})^2$	**r-1**	$S^2 sog = \frac{\Sigma(\bar{y}_{i.} - \bar{y}_{..})^2}{r-1}$	$\frac{S^2sog}{S^2res}$
Tra gruppi	$\Sigma(\bar{y}_{.j} - \bar{y}_{..})^2$	**c-1**	$S^2 grup = \frac{\Sigma(\bar{y}_{.j} - \bar{y}_{..})^2}{c-1}$	$\frac{S^2grup}{S^2res}$
Residua	$\Sigma(y_{ij} - \bar{y}_{i.} - \bar{y}_{.j} + \bar{y}_{..})^2$	**(r-1)(c-1)**	$S^2 res = \frac{\Sigma(y_{ij} - \bar{y}_{i.} - \bar{y}_{.j} + \bar{y}_{..})^2}{(r-1)(c-1)}$	
Totale	$\Sigma(y_{ij} - \bar{y}_{..})^2$	**rc-1** **N-1**		

TEST DI FRIEDMAN

Il test di Friedman è l'equivalente non parametrico dell'analisi della varianza a blocchi randomizzati

Se

- n = numero di righe indipendenti non omogenee (soggetti)
- k = numero di colonne, trattamenti, tempo, misure ripetute
- R_i = Somma dei ranghi per i-esima colonna .
  I ranghi devono essere assegnati per riga

La statistica test con distribuzione $\chi 2$ con k-1 gradi di libertà è:

$$\chi^2 = \left(\frac{12}{nk(k+1)}\sum_{i=1}^{k} R_i^2\right) - 3n(k+1)$$