



UNIVERSITÀ DEGLI STUDI DI MILANO

DIPARTIMENTO DI SCIENZE AGRARIE E AMBIENTALI
PRODUZIONE, TERRITORIO, AGROENERGIA

Metodologia Sperimentale Agronomica / Metodi Statistici per la Ricerca Ambientale

Marco Acutis

marco.acutis@unimi.it

www.acutis.it

a.a. 2018 - 2019

CdS Scienze della Produzione e Protezione delle Piante (g59)

CdS Biotecnologie Vegetali, Alimentari e Agro-Ambientali (g61)

CdS Scienze Agro-Ambientali (g57)

Lezione 07 - Sommario

- ❑ Schemi sperimentali
 - Randomizzazione completa
 - Blocchi randomizzati
 - Quadrato latino
 - Split plot



Introduzione

Schemi sperimentali più usati in Agronomia:

- ✓ randomizzazione completa
- ✓ blocchi randomizzati
- ✓ quadrato latino
- ✓ split plot
- ✓ strip splot

Schemi sperimentali - Randomizzazione completa

Definizione

Consiste

- nell'attribuire mediante sorteggio un trattamento a ciascuna unità sperimentale, non considerando la sua posizione fisica.

Si usa

- in indagini territoriali scegliendo campioni completamente casuali, laddove non sia possibile predisporre uno schema;
- talvolta in prove in ambiente controllato o in laboratorio, cioè in situazioni in cui i fattori non sperimentali sono controllati al meglio.

Non si usa

- in prove impostate in campo, poiché non offre nessun controllo della variabilità accidentale, che invece è sempre elevata.



Schemi sperimentali - Randomizzazione completa

Esempio

Dati 4 trattamenti e 3 ripetizioni per ciascuno,

- 1) tracciare $4 \times 3 = 12$ parcelle;
- 2) attribuire a ogni parcella il proprio trattamento mediante sorteggio;
- 3) eseguire l'esperimento e raccogliere i risultati;
- 4) elaborare i dati secondo la tecnica usuale di analisi della varianza a 1 via.

A1	A4	A2
A2	A3	A1
A3	A1	A2
A4	A3	A4

Qui le unità sperimentali sono rappresentate come adiacenti, ma non è affatto necessario che lo siano!



Schemi sperimentali - Blocchi randomizzati

Definizione

Consiste

- nel suddividere l'area sperimentale in **blocchi**, in modo che questi ultimi abbiano la massima omogeneità al loro interno e siano il più possibile differenziati tra loro;
- nel disporre casualmente i livelli del trattamento, in modo che in ogni blocco sia rappresentato uno e un solo livello (si parla perciò di **randomizzazione ristretta**).

Si usa

- nella gran parte delle prove di tipo manipolativo: è lo schema più usuale in prove agronomiche.

Richiede

- esperimenti bilanciati.

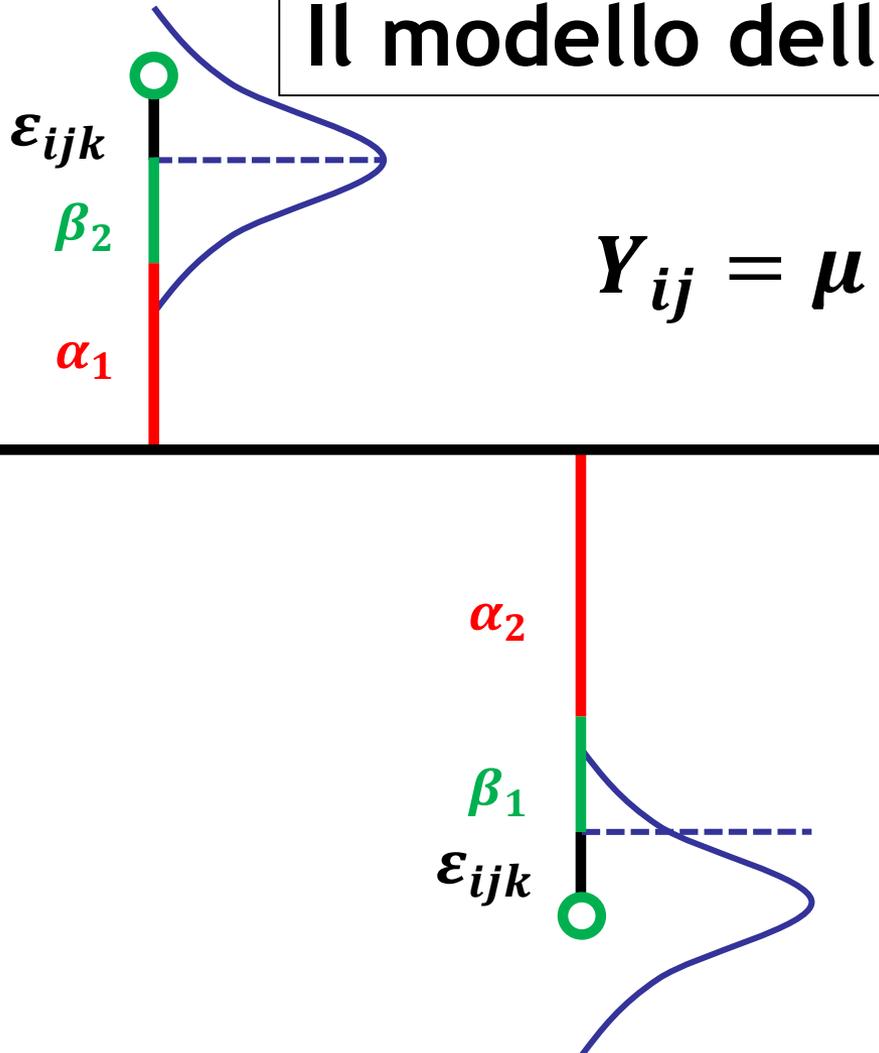
Offre

- la possibilità di controllare, almeno in parte, gli effetti dell'eterogeneità del terreno, migliorando la potenza dell'esperimento grazie all'eliminazione della variabilità tra i blocchi dall'errore sperimentale;
- la possibilità di suddividere il lavoro tra più operatori o in più giorni (1 per blocco) ed eliminarne la conseguente variabilità.



Schemi sperimentali - Blocchi randomizzati

Il modello dell'Anova a blocchi



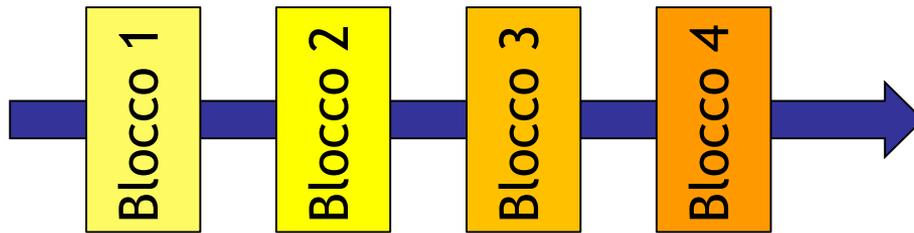
$$Y_{ij} = \mu + \alpha_i + \beta_j + \epsilon_{ij}$$

Il valore di un dato Y_{ijk} è la somma dell'effetto di uno specifico livello del 1° fattore (α_i), dell'effetto del blocco di appartenenza (β_j) e di una componente accidentale (ϵ_{ij}), tenuto conto della media generale μ . È esplicitamente esclusa l'interazione tra blocco e fattore.

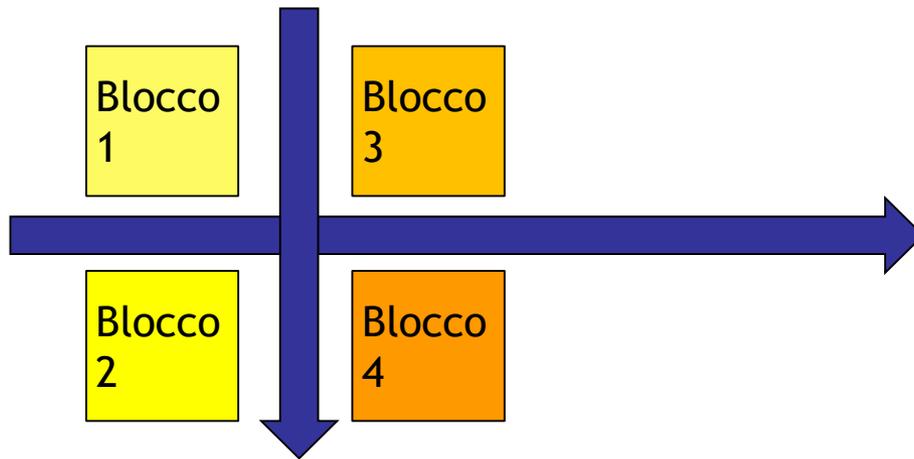
Schemi sperimentali - Blocchi randomizzati

Disposizione e forma dei blocchi

Uno dei più frequenti casi in cui si rende necessaria la suddivisione dell'area sperimentale in blocchi è quello in cui è nota la presenza di un gradiente (ad esempio di fertilità).



In caso di un solo gradiente, occorre predisporre blocchi lunghi e stretti perpendicolarmente al gradiente stesso.



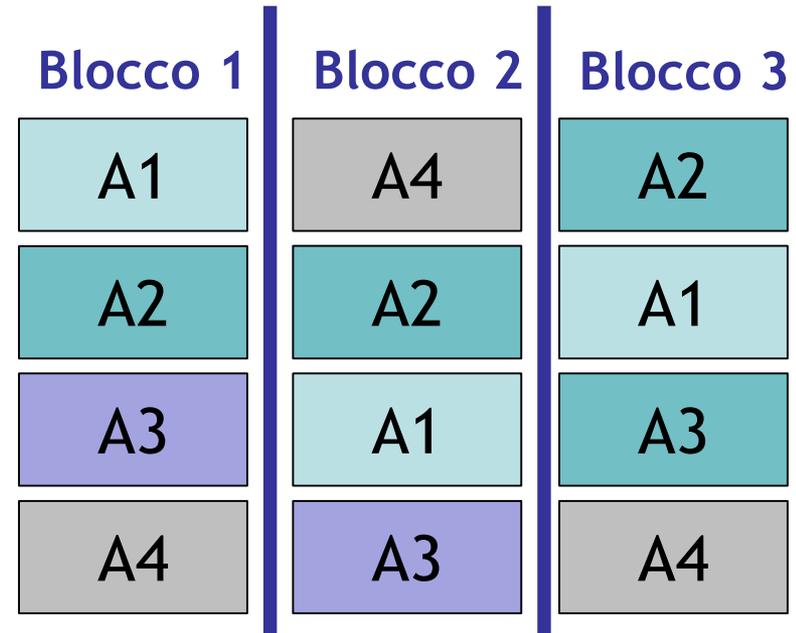
In presenza di 2 gradienti perpendicolari, occorre predisporre blocchi di forma quadrata posizionati «a quadrato». Analogamente si procede se i gradienti non sono noti.

Schemi sperimentali - Blocchi randomizzati

Esempio

Dati 4 trattamenti e 3 ripetizioni per ciascuno,

- 1) tracciare 3 blocchi e suddividerli in 4 parcelle;
- 2) attribuire, all'interno di ciascun blocco, a ogni parcella il proprio trattamento mediante sorteggio;
- 3) eseguire l'esperimento e raccogliere i risultati;
- 4) elaborare i dati, sottraendo devianza e gradi di libertà dei blocchi da quelli dell'errore di una Anova a 1 via usuale (eseguita cioè trascurando i blocchi).



Schemi sperimentali - Blocchi randomizzati

Calcoli (1/2)

Sia nb il numero dei blocchi e nt il numero dei livelli del trattamento.

✓ Per la **devianza trattamenti** procediamo come già visto nel caso dell'Anova «ordinaria»

$$dev_{TRA} = nb \sum_{i=1}^{nt} (\bar{X}_i - \bar{\bar{X}})^2.$$

✓ Per la **devianza totale** procediamo come già visto nel caso dell'Anova «ordinaria»

$$dev_{TOT} = \sum_{i=1}^{nb*nt} (X_i - \bar{\bar{X}})^2.$$

✓ Per la **devianza blocchi** procediamo considerando i blocchi come fossero trattamenti, sostituiamo cioè a ogni parcella del blocco il valore medio del blocco e infine calcoliamo la devianza di tutti questi dati

$$dev_B = nt \sum_{i=1}^{nb} (\bar{X}_i - \bar{\bar{X}})^2.$$

✓ Per la **devianza errore** procediamo per differenza

$$dev_{ERR} = dev_{TOT} - dev_{TRA} - dev_B.$$



Schemi sperimentali - Blocchi randomizzati

Calcoli (2/2)

Fonte di Variazione	Devianza	GdL	Varianza	F calcolato	$P(F)$
Totale	dev_{TOT}	$nt * nb - 1$			
Blocchi	dev_B	$nb - 1$	$\frac{dev_B}{nb - 1}$	$\frac{Var_B}{Var_{ERR}}$	
Trattamenti	dev_{TRA}	$nt - 1$	$\frac{dev_{TRA}}{nt - 1}$	$\frac{Var_{TRA}}{Var_{ERR}}$	
Errore	dev_{ERR}	$(nt - 1) * (nb - 1)$	$\frac{dev_{ERR}}{(nt - 1) * (nb - 1)}$		

Il giudizio sulla significatività dell'effetto dei trattamenti si fa in base al valore di $P(F)$. La $P(F)$ relativa ai blocchi indica se l'applicazione dello schema a blocchi è risultata efficace: se $P(F) \geq \alpha$, allora i blocchi sono inutili, mentre se $P(F) < \alpha$, allora i blocchi hanno apportato una significativa riduzione dell'errore sperimentale.

Schemi sperimentali - Blocchi randomizzati

Efficienza Relativa

Se l'effetto dei blocchi è significativo, si può valutare il parametro di **efficienza relativa** ($R.E.$) del disegno sperimentale a blocchi randomizzati (RCBD - Randomized Complete Block Design) rispetto a un analogo schema a randomizzazione completa (CRD - Completely Randomized Design):

$$R.E. = \frac{(nb - 1)Var_B + nb(nt - 1)Var_{ERR}(RCBD)}{(ntnb - 1)Var_{ERR}(CRD)}$$

N.B. La formula vale per $GdL_{ERR} > 20$, altrimenti sovrastima e perciò occorre una correzione.

$R.E.$ permette di **stimare** il **numero di ripetizioni** che dovrebbe avere un esperimento a **randomizzazione completa** per ottenere la stessa potenza dell'esperimento a blocchi iniziale.

Immaginiamo di aver un RCBD con nb blocchi e calcoliamo $Var_{ERR}(RCBD)$. Facciamo poi finta che i blocchi non ci siano, assegniamo loro il ruolo di ripetizioni e determiniamo $Var_{ERR}(CRD)$. $R.E.* nb$ rappresenta il numero di ripetizioni che dovremmo fare in un CRD per avere la stessa potenza del RCBD da cui siamo partiti.



Schemi sperimentali - Blocchi randomizzati

Requisiti

1. Normalità delle popolazioni da cui sono tratti i campioni .
2. Omogeneità delle varianze.
3. Indipendenza dei trattamenti.
4. Assenza di interazione tra trattamenti e blocchi.

N.B. I primi tre requisiti sono gli stessi dello schema sperimentale a randomizzazione completa.



Interazione Blocco x Trattamento (1/2)

Esiste un test, dovuto a Tukey, per verificare la presenza di interazione, che si basa sull'assumere che l'interazione abbia la forma particolare:

$(\alpha\beta)_{ij} = \gamma \cdot \alpha_i\beta_j$; in questo caso si ha 1 GdL per l'interazione.

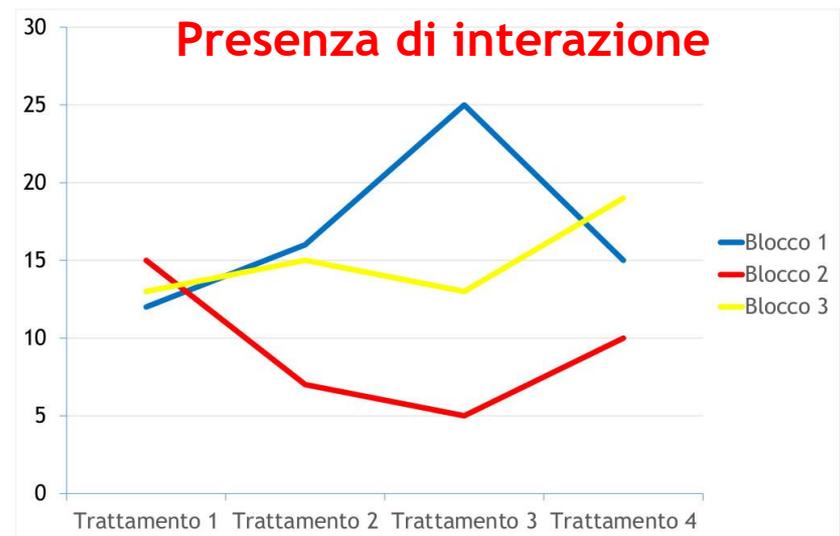
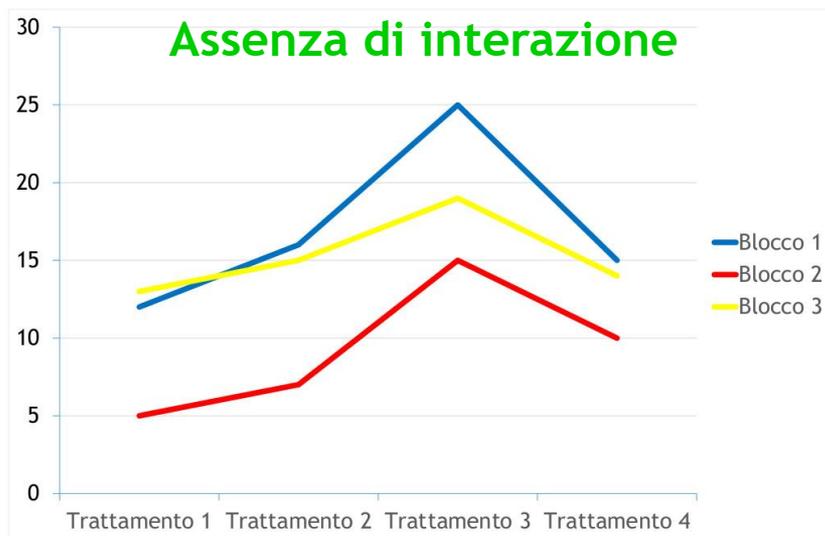
La derivazione matematica è oltre gli obbiettivi del corso, mentre l'applicazione è semplice e si trova su tutti i testi.



Schemi sperimentali - Blocchi randomizzati

Interazione Blocco x Trattamento

In alternativa è possibile procedere come già visto trattando l'Anova a 2 vie, vale a dire creando il grafico dei valori ottenuti all'interno di ciascun blocco in funzione del trattamento e osservandone l'andamento.



N.B. In caso di forte sospetto di interazione tra blocco e trattamento, il disegno sperimentale potrebbe essere scorretto e il modello dell'Anova a blocchi non potrebbe essere applicato.



Schemi sperimentali - Quadrato latino

Costruzione

I risultati di un esperimento a **quadrato latino** si possono schematizzare tramite una tabella, in cui i primi due fattori (che sono sconosciuti e rappresentano fonti di variabilità da sottrarre all'errore sperimentale) corrispondono rispettivamente alle righe e alle colonne, mentre il terzo, vale a dire il fattore oggetto di studio, è contenuto nelle celle formate dall'incrocio tra riga e colonna.

I passaggi da compiere per allestire la prova sperimentale sono:

1. individuare un'area di dimensioni adeguate;
2. distribuire il fattore oggetto di studio in modo casuale, ma ordinato in ciascuna cella, di modo che esso compaia una sola volta in ogni riga e in ogni colonna.

N.B. Si assume A PRIORI l'assenza d'interazioni tra il fattore in studio e ciascuno dei due fattori in riga e in colonna.

Modello additivo dell'analisi della varianza:

$$Y_{ijk} = \mu + \alpha_i + \beta_j + \gamma_k + \varepsilon_{ijk}$$



Schemi sperimentali - Quadrato latino

Esempio

Si intende confrontare la produttività di 5 varietà di sementi (A, B, C, D, E).

	Colonna1	Colonna2	Colonna3	Colonna4	Colonna5
Riga1	$Y_{1,1,SemC}$	$Y_{1,2,SemB}$	$Y_{1,3,SemA}$	$Y_{1,4,SemE}$	$Y_{1,5,SemD}$
Riga2	$Y_{2,1,SemB}$	$Y_{2,2,SemD}$	$Y_{2,3,SemC}$	$Y_{2,4,SemA}$	$Y_{2,5,SemE}$
Riga3	$Y_{3,1,SemA}$	$Y_{3,2,SemE}$	$Y_{3,3,SemD}$	$Y_{3,4,SemC}$	$Y_{3,5,SemB}$
Riga4	$Y_{4,1,SemD}$	$Y_{4,2,SemA}$	$Y_{4,3,SemE}$	$Y_{4,4,SemB}$	$Y_{4,5,SemC}$
Riga5	$Y_{5,1,SemE}$	$Y_{5,2,SemC}$	$Y_{5,3,SemB}$	$Y_{5,4,SemD}$	$Y_{5,5,SemA}$

- Il numero di righe e di colonne dev'essere uguale (altrimenti non parleremmo di quadrato...) e questo comporta che il numero di osservazioni necessarie per confrontare p livelli di un fattore sia pari a p^2 .
- Righe e colonne devono essere intese come suddivisioni concettuali, non necessariamente spaziali.

Schemi sperimentali - Split plot

Definizione (1/3)

Lo **Split Plot Design** (SPD), o disegno a unità suddivise, appartiene alla categoria dei disegni sperimentali multifattoriali ed è spesso utilizzato negli esperimenti condotti in contesti agricoli e industriali o in laboratorio.

In particolare, esso viene utilizzato nei casi in cui all'interno della situazione sperimentale si vogliono analizzare determinati fattori, i cui livelli difficilmente possono essere variati rispetto ad altri fattori, sia per eccessivi costi, sia per restrizioni imposte dal processo stesso.

Talvolta può risultare utile in situazioni in cui non è conveniente l'uso del disegno a blocchi randomizzati, che risulta in generale meno efficiente.



Schemi sperimentali - Split plot

Definizione (2/3)

Lo split plot si applica quando:

- ci sono almeno 2 fattori;
- si intende privilegiare la potenza relativamente all'effetto di uno dei fattori e dell'interazione;
- le condizioni operative impediscono di eseguire uno dei trattamenti singolarmente per ogni unità sperimentale (esempi classici sono dati dalle prove di lavorazione o irrigazione del suolo, dove non è possibile lavorare o irrigare correttamente solo alcuni metri quadri di terreno).



Schemi sperimentali - Split plot

Definizione (3/3)

Lo split plot prevede una sistemazione gerarchica o annidata dei trattamenti in parcelle principali e sub-parcelle. In altre parole, lo split plot si può immaginare come il risultato della sovrapposizione di due schemi a blocchi randomizzati, che vanno analizzati separatamente.

In questo modo si viene ad avere una tesi sperimentale di I ordine (es. lavorazioni) e una tesi sperimentale di II ordine (es. varietà); la prima è randomizzata sulle parcelle principali (main plots), mentre la seconda è randomizzata sulle sub-parcelle (sub-plots).



Schemi sperimentali - Split plot

Esempio (1/2)

Si consideri un esperimento in cui vi sia un **primo fattore**, la **fertilizzazione**, studiato a **4 livelli** (Fert1, Fert2, Fert3, Fert4) disposti secondo un **disegno a blocchi randomizzati**, in cui i **tre blocchi** (B1, B2, B3) sono costituiti da appezzamenti di terreno, che devono avere una determinata estensione prefissata. Si ha dunque:

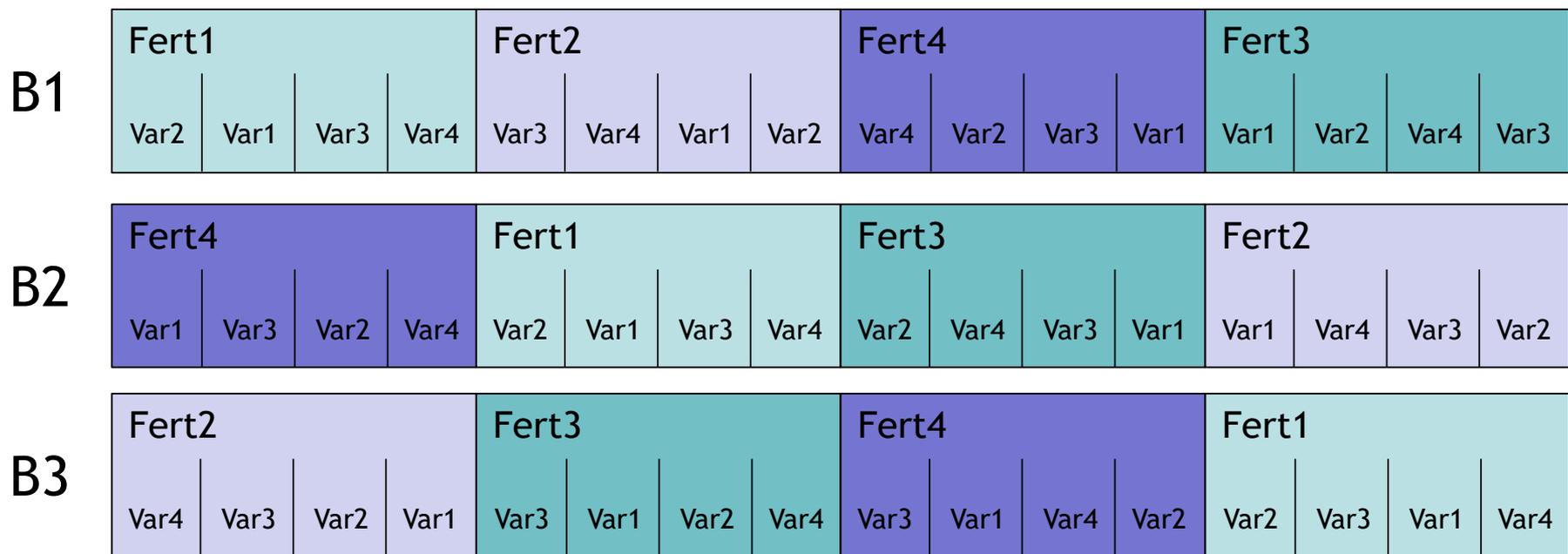
B1	Fert1	Fert2	Fert4	Fert3
B2	Fert4	Fert1	Fert3	Fert2
B3	Fert2	Fert3	Fert4	Fert1



Schemi sperimentali - Split plot

Esempio (2/2)

Si supponga inoltre di avere un **secondo fattore di interesse**, la **varietà**, analizzata a 4 livelli (Var1, Var2, Var3, Var4). Essendo necessaria, per ogni livello, una minore estensione di terreno, si ha allora la seguente struttura:



Schemi sperimentali - Split plot

Struttura

Chiamiamo **fattore di classificazione** un fattore non considerato di maggiore interesse e/o i cui livelli difficilmente possono essere fatti variare. Tali fattori vengono introdotti nell'esperimento per analizzare in che modo essi, interagendo con gli altri fattori, ne possano modificare l'azione, oppure, semplicemente, per analizzarne il comportamento.

Agli effetti principali dei fattori di classificazione non si attribuisce particolare importanza: sono introdotti nel disegno soprattutto con lo scopo di analizzare la loro relazione con gli altri fattori considerati. Per questo motivo si è disposti ad accettare una minore precisione delle stime degli effetti dei fattori di classificazione, in ragione di una migliore precisione per gli altri fattori e per le interazioni di particolare interesse.

Nella pianificazione di un disegno split plot le unità sperimentali considerate vengono suddivise in una prima fase in **whole-units (parcelle principali)**, ovvero in gruppi di unità a cui vengono **assegnati casualmente i diversi livelli dei fattori di classificazione**. Successivamente, ogni whole-unit viene suddivisa in **sub-units (sub-parcelle)**, ovvero in sottogruppi di unità per i quali il livello dei fattori di classificazione rimane fisso, ma a cui vengono **assegnati casualmente i livelli degli altri fattori**.



Schemi sperimentali - Split plot

Modello lineare

$$Y_{ijk} = \mu + \rho_i + \tau_j + \delta_{ij} + \beta_k + (\tau\beta)_{jk} + \varepsilon_{ijk}$$

$\rho_i, i = 1, \dots, r$ → Effetto dovuto al **Blocco**

$\tau_j, j = 1, \dots, a$ → Effetto dovuto al fattore di classificazione (**Fertilizzazione**)

$\beta_k, k = 1, \dots, b$ → Effetto dovuto al secondo fattore (**Varietà**)

$(\tau\beta)_{jk}$ → Effetto dovuto all'interazione (**Fertilizzazione x Varietà**)

$\delta_{ij} \sim N(0, \sigma_\delta^2)$ → Errore associato alla j-esima parcella (trattata con il j-esimo livello del fattore di classificazione) dell'i-esimo blocco

$\varepsilon_{ijk} \sim N(0, \sigma_\varepsilon^2)$ → Errore associato alle sub-parcelle

requisiti

Schemi sperimentali - Split plot

Devianze

	Fonte di Variazione	Devianza	GdL
Parcelle principali	Blocco (ρ_i)	$\frac{1}{ab} \left(\sum_{i=1}^r \left(\sum_{j=1}^a \sum_{k=1}^b Y_{ijk} \right)^2 \right) - \frac{1}{rab} \left(\sum_{i=1}^r \sum_{j=1}^a \sum_{k=1}^b Y_{ijk} \right)^2$	$r - 1$
	Fattore di classificazione (τ_j)	$\frac{1}{rb} \left(\sum_{j=1}^a \left(\sum_{i=1}^r \sum_{k=1}^b Y_{ijk} \right)^2 \right) - \frac{1}{rab} \left(\sum_{i=1}^r \sum_{j=1}^a \sum_{k=1}^b Y_{ijk} \right)^2$	$a - 1$
	Errore 1 (δ_{ij})	$\frac{1}{b} \left(\sum_{i=1}^r \sum_{j=1}^a \left(\sum_{k=1}^b Y_{ijk} \right)^2 \right) - \frac{1}{rab} \left(\sum_{i=1}^r \sum_{j=1}^a \sum_{k=1}^b Y_{ijk} \right)^2 - dev(\rho) - dev(\tau)$	$(r - 1)(a - 1)$
Sub-parcelle	Totale	$\sum_{i=1}^r \sum_{j=1}^a \sum_{k=1}^b Y_{ijk}^2 - \frac{1}{rab} \left(\sum_{i=1}^r \sum_{j=1}^a \sum_{k=1}^b Y_{ijk} \right)^2$	$rab - 1$
	Secondo fattore (β_k)	$\frac{1}{ra} \left(\sum_{k=1}^b \left(\sum_{i=1}^r \sum_{j=1}^a Y_{ijk} \right)^2 \right) - \frac{1}{rab} \left(\sum_{i=1}^r \sum_{j=1}^a \sum_{k=1}^b Y_{ijk} \right)^2$	$b - 1$
	Interazione $(\tau\beta)_{jk}$	$\frac{1}{r} \left(\sum_{j=1}^a \sum_{k=1}^b \left(\sum_{i=1}^r Y_{ijk} \right)^2 \right) - \frac{1}{rab} \left(\sum_{i=1}^r \sum_{j=1}^a \sum_{k=1}^b Y_{ijk} \right)^2 - dev(\tau) - dev(\beta)$	$(a - 1)(b - 1)$
	Errore 2 (ε_{ijk})	$dev(TOT) - dev(\rho) - dev(\tau) - dev(\delta) - dev(\beta) - dev(\tau\beta)$	$a(r - 1)(b - 1)$

Schemi sperimentali - Split plot

Tabella Anova

	Fonte di Variazione	Varianza	F
Parcelle principali	Blocco (ρ_i)	$\frac{dev(\rho)}{r - 1}$	
	Fattore di classificazione (τ_j)	$\frac{dev(\tau)}{a - 1}$	$\frac{Var(\tau)}{Var(\delta)}$
	Errore 1 (δ_{ij})	$\frac{dev(\delta)}{(r - 1)(a - 1)}$	
Sub-parcelle	Totale	$\frac{dev(TOT)}{rab - 1}$	
	Secondo fattore (β_k)	$\frac{dev(\beta)}{b - 1}$	$\frac{Var(\beta)}{Var(\varepsilon)}$
	Interazione ($\tau\beta)_{jk}$	$\frac{dev(\tau\beta)}{(a - 1)(b - 1)}$	$\frac{Var(\tau\beta)}{Var(\varepsilon)}$
	Errore 2 (ε_{ijk})	$\frac{dev(\varepsilon)}{a(r - 1)(b - 1)}$	



Schemi sperimentali - Split plot

Spiegazione intuitiva

Il trattamento nelle parcelle principali ha un valore corretto solo se è mediato sulle sub-parcelle incluse in esso.

Se ragionassimo come nel caso del fattoriale semplice, dovremmo considerare, come ripetizioni, tutte le unità che «subiscono» lo stesso livello del fattore di classificazione (cioè quello nella parcella principale). In questo modo però non si terrebbe conto del fatto che le sub-unità sono sempre aggregate per produrre la parcella principale, ovvero che la variabilità tra le parcelle principali è diversa dalla variabilità tra le sub-parcelle.

Pertanto appare necessario, con **2 varianze**, avere **2 errori**.

Alternativamente possiamo osservare che, se non possiamo calcolare l'errore basandoci sulle sub-parcelle, allora occorre assumere come errore l'interazione tra blocco e fattore di classificazione, che non è nient'altro che l'errore dello split-plot relativo alle parcelle principali.



Schemi sperimentali - Split plot

Il confronto tra medie

1. Se l'interazione non è significativa allora si può:
 - ✓ comparare le medie dei trattamenti nelle parcelle principali;
 - ✓ comparare le medie dei trattamenti nelle sub-parcelle.
2. Se l'interazione è significativa allora si può:
 - ✓ comparare le medie di trattamenti nelle sub-parcelle a un livello prefissato del trattamento nelle parcelle principali;
 - ✓ comparare le medie di trattamenti nelle parcelle principali a un livello prefissato del trattamento nelle sub-parcelle.

N.B. Se c'è interazione, non si possono comparare gli effetti semplici!



Schemi sperimentali - Split plot

Interazione non significativa

- ✓ Comparazione tra **2 medie di trattamenti nelle parcelle principali**:

la stima della varianza della differenza tra 2 medie è data da $2Var(\delta)/rb$.

rb è il prodotto del numero di ripetizioni (blocchi) per il numero di livelli del fattore nelle sub-parcelle, cioè il numero effettivo di unità sperimentali che ha dato origine alle medie in esame.

- ✓ Comparazione tra **2 medie di trattamenti nelle sub-parcelle**:

la stima della varianza della differenza tra 2 medie è data da $2Var(\varepsilon)/ra$.

ra è il prodotto del numero di ripetizioni (blocchi) per il numero di livelli del fattore nelle parcelle principali, cioè il numero effettivo di unità sperimentali che ha dato origine alle medie in esame.



Schemi sperimentali - Split plot

Interazione significativa

- ✓ Comparazione tra **2 medie del fattore nelle sub-parcelle a parità di livello del fattore nelle parcelle principali:**

la stima della varianza della differenza tra 2 medie è data da $2Var(\varepsilon)/r$.

r è il numero di ripetizioni (blocchi), cioè il numero effettivo di unità sperimentali che ha dato origine alle medie in esame.

- ✓ Comparazione tra **2 medie del fattore nelle parcelle principali a parità di livello del fattore nelle sub-parcelle:**

la stima della varianza della differenza tra 2 medie è data da $2(Var(\delta) + (b - 1)Var(\varepsilon))/rb$.



Schemi sperimentali - Strip plot

Definizione (1/2)

Possono presentarsi situazioni in cui occorra testare nello stesso esperimento due fattori, ciascuno dei quali necessita di estese unità sperimentali (ad es. irrigazione e aratura).

Se adottassimo uno schema a blocchi randomizzati per analizzare tutte le combinazioni, avremmo bisogno di parcelle enormi e ciò causerebbe un'elevata eterogeneità.

Problemi analoghi si presenterebbero nella costruzione di uno schema split plot poiché, anche assegnando all'uno il ruolo di fattore di classificazione, la natura dell'altro richiederebbe sub-parcelle molto grandi.

In casi come questi si deve utilizzare lo schema sperimentale detto **Strip Plot**.



Schemi sperimentali - Strip plot

Definizione (2/2)

Lo schema a strip plot (detto anche **split block**) si applica quando:

- ci sono almeno 2 fattori;
- si intende privilegiare la potenza relativamente all'interazione tra i fattori;
- le condizioni operative impediscono di eseguire 2 trattamenti singolarmente per ogni unità sperimentale (ad esempio per prove di lavorazione del suolo e irrigazione a scala di campo, dove per entrambi i fattori non è possibile lavorare correttamente solo alcuni metri quadri di terreno).

Lo schema potrebbe avere particolare valore nell'ambito di prove dimostrative, in quanto è possibile “vedere” i livelli dei fattori muovendosi in campo lungo linee rette.



Schemi sperimentali - Strip plot

Costruzione e considerazioni statistiche

Lo schema a strip-plot deve prevedere innanzitutto dei **blocchi**, disposti analogamente a quanto indicato per la prova a blocco randomizzato, uguale al **numero di ripetizioni**.

Nell'ambito di ciascun blocco prima si predispongono **casualmente**, una e una sola volta, i **livelli del primo fattore**, poi, **ortogonalmente** ai primi, si posizionano **casualmente** i **livelli del secondo fattore**.

Dal punto di vista dell'elaborazione statistica, i due fattori devono essere trattati separatamente, così come l'interazione. Al **fattore in riga**, al **fattore in colonna** e all'**interazione tra i due** corrisponde un errore diverso. Queste **tre varianze d'errore** vengono usate per verificare la **significatività**, **rispettivamente**, del **fattore in riga**, del **fattore in colonna** e dell'**interazione**.

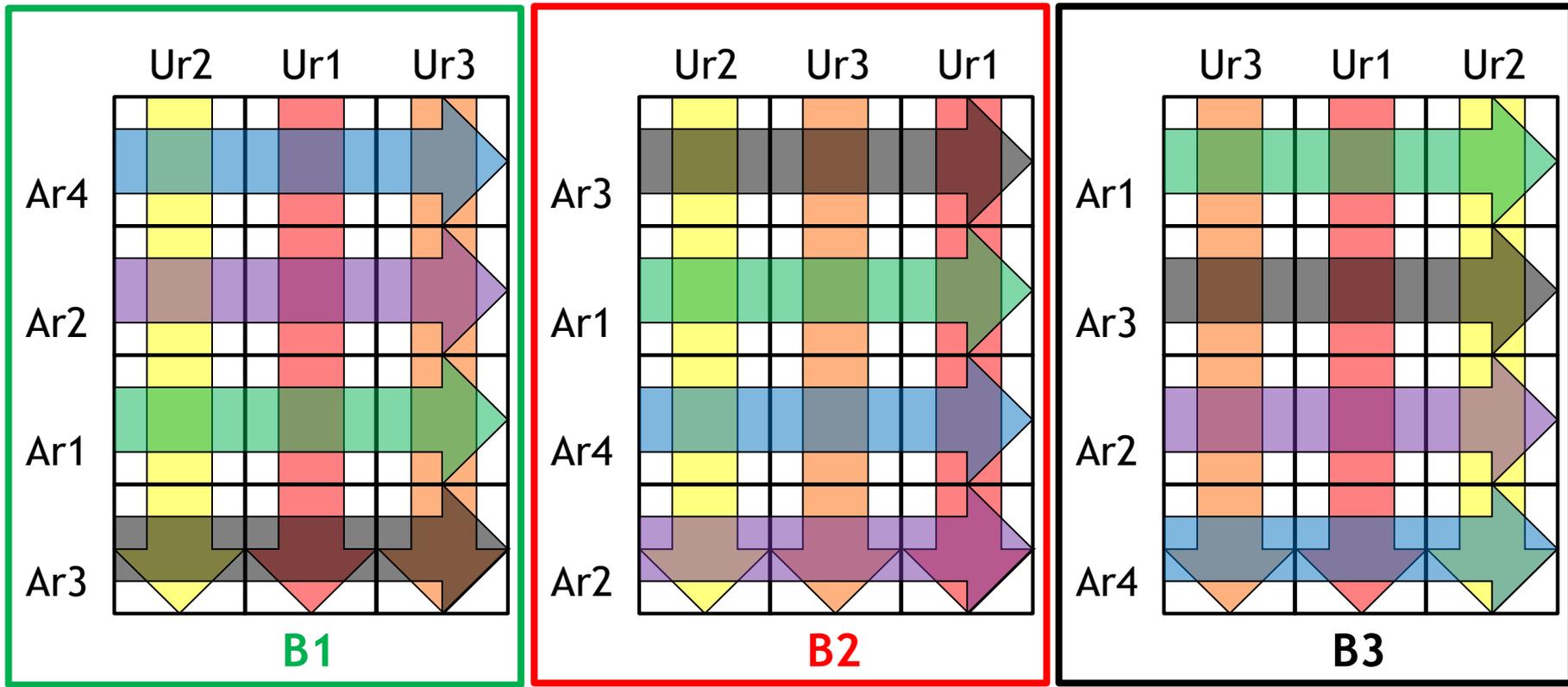
Le parcelle in riga e quelle in colonna hanno un errore sperimentale pari a quello che avrebbero se non esistessero gli "incroci", mentre per l'**analisi dell'interazione** è possibile escludere dall'errore sperimentale tutta la variabilità tra righe e colonne, ottenendo così un **incremento di potenza** rispetto allo schema fattoriale a blocchi e anche rispetto allo split plot. Tale aumento di potenza va a **scapito** della potenza che si avrebbe per gli **effetti semplici**.



Schemi sperimentali - Strip plot

Esempio

Supponiamo di voler analizzare l'effetto congiunto di 3 dosi di urea distribuita a spaglio in presemina (Ur1, Ur2, Ur3) e di 4 profondità di aratura (Ar1, Ar2, Ar3, Ar4) e di disporre di 3 blocchi (B1, B2, B3).



Schemi sperimentali - Strip plot

Modello lineare

$$Y_{ijk} = \mu + \rho_i + \alpha_j + (\rho\alpha)_{ij} + \beta_k + (\rho\beta)_{ik} + (\alpha\beta)_{jk} + \varepsilon_{ijk}$$

$\rho_i, i = 1, \dots, r$ → Effetto dovuto al **Blocco**

$\alpha_j, j = 1, \dots, a$ → Effetto dovuto al primo fattore (**Urea**)

$\beta_k, k = 1, \dots, b$ → Effetto dovuto al secondo fattore (**Aratura**)

$(\alpha\beta)_{jk}$ → Effetto dovuto all'interazione (**Urea x Aratura**)

$(\rho\alpha)_{ij} = \delta_{ij} \sim N(0, \sigma_\delta^2)$ → Errore associato alla j-esima parcella orizzontale dell'i-esimo blocco

$(\rho\alpha)_{ik} = \eta_{ik} \sim N(0, \sigma_\eta^2)$ → Errore associato alla k-esima parcella verticale dell'i-esimo blocco

$\varepsilon_{ijk} \sim N(0, \sigma_\varepsilon^2)$ → Errore associato agli «incroci»

requisiti



Schemi sperimentali - Strip plot

Devianze

Fonte di Variazione	Devianza	GdL
Blocco (ρ_i)	$ab \left(\sum_{i=1}^r (\bar{Y}_i - \bar{Y})^2 \right)$, dove \bar{Y}_i è la media dei dati relativi all'i-esimo blocco	$r - 1$
Primo fattore (α_j)	$rb \left(\sum_{j=1}^a (\bar{Y}_j - \bar{Y})^2 \right)$, dove \bar{Y}_j è la media dei dati relativi al j-esimo livello del primo fattore	$a - 1$
Errore 1 (δ_{ij})	$b \left(\sum_{i=1}^r \sum_{j=1}^a (\bar{Y}_{ij} - \bar{Y})^2 \right) - dev(\rho) - dev(\alpha)$, dove \bar{Y}_{ij} è la media dei dati relativi al j-esimo livello del primo fattore dell'i-esimo blocco	$(r - 1)(a - 1)$
Secondo fattore (β_k)	$ra \left(\sum_{k=1}^b (\bar{Y}_k - \bar{Y})^2 \right)$, dove \bar{Y}_k è la media dei dati relativi al k-esimo livello del secondo fattore	$b - 1$
Errore 2 (η_{ik})	$a \left(\sum_{i=1}^r \sum_{k=1}^b (\bar{Y}_{ik} - \bar{Y})^2 \right) - dev(\rho) - dev(\beta)$, dove \bar{Y}_{ik} è la media dei dati relativi al k-esimo livello del secondo fattore dell'i-esimo blocco	$(r - 1)(b - 1)$
Totale	$\sum_{i=1}^r \sum_{j=1}^a \sum_{k=1}^b (Y_{ijk} - \bar{Y})^2$, dove \bar{Y} è la media generale	$rab - 1$
Interazione ($\alpha\beta$) $_{jk}$	$r \left(\sum_{j=1}^a \sum_{k=1}^b (\bar{Y}_{jk} - \bar{Y})^2 \right) - dev(\alpha) - dev(\beta)$, dove \bar{Y}_{jk} è la media dei dati relativi all'incrocio tra j-esimo livello del primo fattore e k-esimo livello del secondo fattore	$(a - 1)(b - 1)$
Errore 3 (ε_{ijk})	$dev(TOT) - dev(\rho) - dev(\alpha) - dev(\delta) - dev(\beta) - dev(\eta) - dev(\alpha\beta)$	$(r - 1)(a - 1)(b - 1)$



Schemi sperimentali - Strip plot

Tabella Anova

Fonte di Variazione	Varianza	F
Blocco (ρ_i)	$\frac{dev(\rho)}{r - 1}$	
Primo fattore (α_j)	$\frac{dev(\alpha)}{a - 1}$	$\frac{Var(\alpha)}{Var(\delta)}$
Errore 1 (δ_{ij})	$\frac{dev(\delta)}{(r - 1)(a - 1)}$	
Secondo fattore (β_k)	$\frac{dev(\beta)}{b - 1}$	$\frac{Var(\beta)}{Var(\eta)}$
Errore 2 (η_{ik})	$\frac{dev(\eta)}{(r - 1)(b - 1)}$	
Totale	$\frac{dev(TOT)}{rab - 1}$	
Interazione ($\alpha\beta$) _{jk}	$\frac{dev(\alpha\beta)}{(a - 1)(b - 1)}$	$\frac{Var(\alpha\beta)}{Var(\varepsilon)}$
Errore 3 (ε_{ijk})	$\frac{dev(\varepsilon)}{(r - 1)(a - 1)(b - 1)}$	



Schemi sperimentali - Strip plot

Il confronto tra medie

1. Se l'interazione non è significativa allora si può:
 - ✓ comparare le medie dei livelli del primo fattore;
 - ✓ comparare le medie dei livelli del secondo.
2. Se l'interazione è significativa allora si può:
 - ✓ comparare le medie dei livelli del primo fattore a un livello prefissato del secondo fattore;
 - ✓ comparare le medie dei livelli del secondo fattore a un livello prefissato del primo fattore;
 - ✓ comparare le medie dei due fattori a livelli differenti.

N.B. Se c'è interazione, non si possono comparare gli effetti semplici!



Schemi sperimentali - Strip plot

Interazione non significativa

- ✓ Comparazione tra **2 medie del primo fattore**:

la stima della varianza della differenza tra 2 medie è data da $2Var(\delta)/rb$.

rb è il prodotto del numero di ripetizioni (blocchi) per il numero di livelli del secondo fattore, cioè il numero effettivo di unità sperimentali che ha dato origine alle medie in esame.

- ✓ Comparazione tra **2 medie del secondo fattore**:

la stima della varianza della differenza tra 2 medie è data da $2Var(\eta)/ra$.

ra è il prodotto del numero di ripetizioni (blocchi) per il numero di livelli del primo fattore, cioè il numero effettivo di unità sperimentali che ha dato origine alle medie in esame.



Interazione significativa

- ✓ Comparazione tra **2 medie del primo fattore a parità di livello del secondo fattore** ($Y_{i1,j}$ vs $Y_{i2,j}$):

la stima della varianza della differenza tra 2 medie è data da $2(\text{Var}(\delta) + (b - 1)\text{Var}(\varepsilon))/rb$.

- ✓ Comparazione tra **2 medie del secondo fattore a parità di livello del primo fattore** ($Y_{i,j1}$ vs $Y_{i,j2}$):

la stima della varianza della differenza tra 2 medie è data da $2(\text{Var}(\eta) + (a - 1)\text{Var}(\varepsilon))/ra$.

- ✓ Comparazione tra **le medie dei 2 fattori a livelli differenti** ($Y_{i2,j1}$ vs $Y_{i1,j2}$):

la stima della varianza della differenza tra 2 medie è data da $2(a \cdot \text{Var}(\delta) + b \cdot \text{Var}(\eta) + (a - 1)(b - 1) \cdot \text{Var}(\varepsilon))/rab$.