



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 08 - Sommario

- ❑ Schemi sperimentali a misure ripetute
 - Premessa: repliche, blocchi e soggetti
 - Definizione
 - Vantaggi
 - Dipendenza delle osservazioni
 - Sfericità



Premessa: repliche

In un disegno sperimentale a randomizzazione completa si assume in partenza che le unità sperimentali (vale a dire le repliche) rispondano tutte allo stesso modo alla variabile misurata.

Ne consegue che la **variabilità osservata** è ascrivibile soltanto al **caso** e ai **trattamenti** cui le unità sperimentali sono sottoposte.

L'errore (o caso o componente accidentale) incorpora quindi la naturale diversità esistente tra gli individui, sebbene essi per ipotesi vengano considerati «cloni» o, per l'appunto, repliche.

Premessa: blocchi

In un disegno sperimentale a blocchi randomizzati si assume che le unità sperimentali all'interno del blocco rispondano tutte allo stesso modo alla variabile misurata, ma anche che potrebbero esserci differenze tra individui appartenenti a blocchi diversi.

Ne consegue che la variabilità osservata è ascrivibile soltanto ai blocchi, al caso e ai trattamenti cui le unità sperimentali sono sottoposte.

L'errore incorpora quindi la naturale diversità esistente tra individui nell'ambito dello stesso blocco, ma non quella tra blocchi diversi.

Premessa: soggetti

In un disegno sperimentale «a misure ripetute» (sullo stesso soggetto) si assume che vi siano già in partenza differenze tra le unità sperimentali, poiché esse vengono considerate soggetti tutti diversi tra loro che potenzialmente non rispondono sempre allo stesso modo alla variabile misurata.

Ne consegue che la variabilità osservata è ascrivibile ai soggetti, al caso e ai trattamenti cui le unità sperimentali (vale a dire i soggetti stessi) sono sottoposti.

L'errore **NON** incorpora quindi la naturale diversità esistente tra gli individui.

Esempi (1/2)

Immaginiamo di avere a disposizione 1 vigneto e 5 anni di sperimentazione per valutare l'efficacia di 3 fitofarmaci nel controllo della peronospora della vite.

- ✓ **Approccio a randomizzazione completa** (Anova a 2 vie con 1 fattore fisso e 1 fattore casuale) :
 1. all'inizio di ogni anno scegliamo casualmente 180 piante (diverse ogni anno);
 2. assegniamo altrettanto casualmente uno dei tre fitofarmaci a ciascuna pianta;
 3. alla fine di ogni anno misuriamo l'indice di malattia su tutte e sole le piante selezionate.
- ✓ **Approccio a blocchi randomizzati** (Anova a blocchi a 2 vie con 1 fattore fisso e 1 fattore casuale):
 1. all'inizio della sperimentazione suddividiamo latitudinalmente il vigneto in 3 fasce (blocchi), dal momento che è probabile vi sia un gradiente di umidità;
 2. all'inizio di ogni anno suddividiamo ciascun blocco in 3 parti cui casualmente assegniamo uno dei fitofarmaci;
 3. alla fine di ogni anno scegliamo casualmente, all'interno di ogni blocco, 20 piante (diverse ogni anno) per ciascun trattamento;
 4. alla fine di ogni anno determiniamo l'indice di malattia per ciascuna combinazione calcolando la media dei dati misurati sulle 20 piante.

Esempi (2/2)

Immaginiamo di avere a disposizione 1 vigneto e 5 anni di sperimentazione per valutare l'efficacia di 3 fitofarmaci nel controllo della peronospora della vite.

✓ **Approccio a misure ripetute** :

1. all'inizio della sperimentazione scegliamo casualmente 180 piante;
2. assegniamo altrettanto casualmente uno dei tre fitofarmaci a ciascuna pianta;
3. alla fine di ogni anno misuriamo l'indice di malattia su tutte e sole le piante selezionate.

✓ **Approccio a blocchi randomizzati e misure ripetute:**

1. all'inizio della sperimentazione suddividiamo latitudinalmente il vigneto in 3 fasce (blocchi), dal momento che è probabile vi sia un gradiente di umidità;
2. all'inizio della sperimentazione suddividiamo ciascun blocco in 3 parti cui casualmente assegniamo uno dei fitofarmaci;
3. all'inizio della sperimentazione scegliamo casualmente, all'interno di ogni blocco, 20 piante per ciascun trattamento;
4. alla fine di ogni anno determiniamo l'indice di malattia per ciascuna combinazione calcolando la media dei dati misurati sulle 20 piante.

Definizione (1/2)

Uno schema sperimentale si dice dunque «a misure ripetute» quando viene fatta più di una osservazione sulla stessa unità sperimentale o soggetto .

Un caso particolarmente rappresentativo è quello delle colture pluriennali (ad esempio le arboree), dove, una volta assegnati all'inizio della prova i trattamenti alle unità sperimentali attraverso estrazione casuale, questi vengono poi ripetuti negli anni sempre sulle medesime unità.

Tutti gli schemi sperimentali visti fino a questo punto possono essere «arricchiti» con misure ripetute.

Definizione (2/2)

Nel caso di disegni sperimentali a misure ripetute, dobbiamo perciò tenere ben presente che esistono due tipi di **variabilità**:

- **TRA i soggetti (BETWEEN subjects)**, associata al fatto che diversi gruppi di soggetti ricevono diversi trattamenti (equivalente alla variabilità TRA i gruppi già vista nel caso dell'Anova a 1 via);
- **ENTRO i soggetti (WITHIN subjects)**, associata al fatto che i gruppi di soggetti che ricevono lo stesso trattamento vengono valutati più volte in condizioni differenti.

Vantaggi

Gli schemi sperimentali a misure ripetute consentono:

- ✓ di avere molte condizioni sperimentali con pochi soggetti e questo determina maggior potere statistico;
- ✓ di misurare il cambiamento nel tempo;
- ✓ di rimuovere la variabilità individuale non inerente alle variabili misurate e questo determina minore varianza di errore.

Dipendenza delle osservazioni

Fin qui abbiamo visto che tutti i test parametrici (quelli cioè che presuppongono una distribuzione normale della popolazione cui appartengono i dati) hanno come requisito fondamentale l'indipendenza delle osservazioni.

Nel caso di misure ripetute questo non può accadere, dal momento che dati provenienti dai medesimi soggetti (o unità sperimentali) sono necessariamente e naturalmente correlati. Per di più, è facilmente intuibile come il grado di dipendenza tra due misurazioni debba essere **PROBABILMENTE** più forte quanto più queste sono vicine temporalmente.

Sfericità (1/2)

L'esistenza di una relazione tra i dati osservati sul medesimo soggetto impone un'assunzione aggiuntiva che prende il nome di Sfericità. Essa consiste, in pratica, nel richiedere che il grado di dipendenza tra due misurazioni effettuate sullo stesso soggetto sia approssimativamente sempre lo stesso, quali che siano le misurazioni (vale a dire i trattamenti) e quali che siano i soggetti. L'ipotesi di sfericità quindi equivale a presupporre che tutte le condizioni in cui è stato condotto l'esperimento (ad esempio gli anni) siano correlate tra loro alla stessa «intensità» e che perciò l'effetto di tali condizioni si possa considerare invariante.

Tecnicamente parlando, il requisito di sfericità «assomiglia» all'omogeneità delle varianze perché si verifica, attraverso il test di Mauchly, controllando che sia approssimativamente uguale la varianza delle differenze tra dati ottenuti dallo stesso soggetto, ma in condizioni diverse (cioè, ad esempio, da un anno all'altro).

Sfericità (2/2)

Le conseguenze della **violazione del requisito di sfericità** sono: (i) perdita di potenza del test (cioè aumento della probabilità di commettere l'Errore del II tipo) e (ii) impossibilità di confrontare il rapporto F calcolato con il corrispondente valore tabulato.

In campo agronomico e ambientale, quando l'esperimento viene ripetuto da un'annata all'altra, è molto difficile che il requisito di sfericità venga rispettato. Inoltre ci sono moltissimi altri casi in cui lo sperimentatore già sa in partenza che, operando in condizioni diverse, le ripercussioni varieranno enormemente.

Per questo motivo, sono state elaborate le seguenti correzioni: **Greenhouse-Geisser**, **Huynh-Feldt** e **Lower-bound**. Queste consentono di coprire la maggior parte dei casi in cui la sfericità è violata, fornendo un valore del rapporto F corretto, sebbene più conservativo.