



CORSO SIGA

“STATISTICA DI BASE PER IL MIGLIORAMENTO GENETICO TRADIZIONALE ED AVANZATO”

Salsomaggiore Terme, 29 Settembre-3 Ottobre 2014

Obiettivi formativi: *le lezioni sono rivolte a ricercatori in formazione che intendano acquisire i principali aspetti dell'applicazione della biometria ai problemi di genetica applicata. Il corso non intende affrontare nozioni computazionali o di carattere formale, ma è principalmente rivolto alla comprensione delle nozioni di base essenziali per una corretta pianificazione degli esperimenti e della conseguente interpretazione dei dati sperimentali.*

Il corso è articolato in lezioni al mattino e in seminari applicativi nel pomeriggio.

PROGRAMMA ED OBIETTIVI SPECIFICI

Lunedì 29 settembre 2014 (dalle ore 14:00 alle ore 17:00)

Obiettivi: Presentazione dei partecipanti. Verifica del livello di conoscenza pregressa e delle attese sulla base di una semplice discussione con il docente.

Chiarimento sulle finalità del corso: Poco spazio ai calcoli, ma maggiore enfasi sulla consapevolezza dei significati e delle attese delle procedure statistiche, oggi facilmente risolte dal calcolo automatico. Il rischio del procedere per ricette, perdendo la possibilità di introdurre innovazione.

Argomenti: Quello che sappiamo o dovremmo sapere: le statistiche elementari, l'approccio frequentista, le statistiche e le loro attese. La variabilità fenotipica e la sua descrizione (uni variata e multivariata). La variabile casuale e richiamo delle principali distribuzioni di probabilità e di densità di probabilità.

L'inferenza statistica. Il problema delle assunzioni. Massima verosimiglianza e minimi quadrati. Gli intervalli di confidenza. Come strutturiamo una tabella di risultati descrittivi.

Martedì 30 settembre (dalle ore 9:00 alle ore 13:00)

Obiettivi: Comprendere l'importanza della pianificazione dell'esperimento anche ai fini dell'inferenza statistica. Il significato di ipotesi e la sua verifica in presenza di assunzioni e in caso di loro assenza.

Argomenti: Inferenza statistica e verifica delle ipotesi. Errori di I e II tipo. Illustrazione di ipotesi semplici e complesse. La pianificazione dell'esperimento. Le repliche e la precisione del dato. La

fenotipizzazione. Ruolo della pianificazione per una inferenza efficace. Alcuni casi classici di verifica delle ipotesi: la regressione, l'analisi della varianza ad un criterio di classificazione ecc.

Martedì 30 settembre (dalle ore 14:00 alle ore 17:00)

Prof. Marco Acutis (Università di Milano).

Approfondimento ANOVA

Mercoledì 1 ottobre (dalle ore 9:00 alle ore 13:00)

Obiettivi: Il modello lineare come approccio unificato nell'analisi dei dati sperimentali. Comprensione dell'importanza della matrice del disegno e della pianificazione delle fonti di variazione.

Argomenti: Il modello lineare. Significato e sua interpretazione. Stima dei parametri. Modello fisso e modello casuale e loro significato nell'analisi genetica. Il modello lineare nell'ANOVA e nella regressione multipla e curvilinea. La matrice del disegno, modelli vincolati. Cenni al modello lineare generalizzato, logistico ecc. Applicazioni in Genetica: strutturazione in famiglie, ereditabilità, breeding value, regressione genitori figli, localizzazione di QTL, ecc

Mercoledì 1 ottobre (dalle ore 14:00 alle ore 17:00)

Dott.ssa Elisabetta Frascaroli (Università di Bologna)

Approfondimento ed esempi applicativi (stima di parametri genetici, QTL ecc.)

Giovedì 2 ottobre (dalle ore 9:00 alle ore 13:00)

Obiettivi: Verso la "high dimension biology": la comprensione dell'analisi bayesiana e l'uso di distribuzioni generate dai dati sperimentali.

Argomenti: I metodi di ricampionamento: bootstrap e jackknife. Introduzione all'analisi bayesiana. Il teorema di Bayes: dalla massima verosimiglianza alla stima a posteriori della distribuzione dei parametri. La verifica delle ipotesi nell'analisi bayesiana. Confronto con l'approccio frequenti sta.

Giovedì 2 ottobre (dalle ore 14:00 alle ore 17:00)

Prof. Federico Mattia Stefanini (Università di Firenze)

Approfondimento ed esempi applicativi

Venerdì 3 ottobre (dalle ore 09:00 alle ore 13:00)

Obiettivi: Discutere i vantaggi offerti dall'analisi multidimensionale dei dati. Comprendere le principali caratteristiche dell'analisi multivariata.

Argomenti: I metodi univariati sono un sottoinsieme dell'analisi multivariata dei dati. Distribuzione di densità multi normale. La matrice dei dati e l'analisi "per righe" e "per colonne". Elenco delle principali tecniche di analisi e relative connessioni.

Discussione finale e distribuzione degli attestati.

Informazioni generali

E' opportuno che i partecipanti siano muniti di proprio elaboratore portatile. Nella sede didattica è garantita la possibilità di connessione Wi-fi. Non è richiesta la conoscenza di uno specifico software applicativo, anche se alcuni esempi e/o applicazioni prevedono l'uso del linguaggio R.

Testo di riferimento. Federico M. Stefanini Introduzione alla Statistica applicata con esempi in R. Pearson Education. (ISBN 9788871923567)

Altri testi per consultazione reperibili in rete verranno comunicati ai partecipanti prima dell'inizio del corso.