

ANOVA

Case	Normality	Homoscedasticity	ANOVA Test
1	Verified	Verified	Parametric Homoscedastic ANOVA
2	Not Verified	Verified	Non Parametric Homoscedastic ANOVA
3	Verified	Not Verified	Parametric Heteroscedastic ANOVA
4	Not Verified	Not Verified	Non Parametric Heteroscedastic ANOVA

- caso2: normalità non verificata, faccio il kruskal-wallis test, non si applica il classico F-test test non-parametrico, si riferisce al fatto che i dati non appartengono ad alcuna distribuzione particolare
- caso3: non verificata l'omoscedasticità, faccio il welsch test
- caso4: test di kruskal-wallis, quindi lo faccio a prescindere dall'omoscedasticità

Web Server (DOE):

somma dei quadrati → importanza

shapiro-wilk: quando ho dei dubbi col test visivo, mi permette di verificare la normalità (ottengo il p-value in output)

residui → stima continua → attiva stimatori legacy

stimatori legacy → stima continua → stima normale

stima normale → bontà di adattamento

devo verificare l'omoscedasticità: arancione → omoscedasticità rigettata → F-test non applicabile; quindi, applico welsch

Welch:

- nel primo caso p-value basso, significatività accettata
- nel secondo caso il p-value è alto quindi il fattore cache non è significativo

caso 3 su JMP: analizza y rispetto a x → analisi ad una via y rispetto fattore → varianze uguali (in questa analisi trascuriamo l'interazione, mi interessa solo capire se hanno varianze omogenee)

caso2/4 su JMP: analizza y rispetto a x → analisi ad una via Y rispetto fattore → test non parametrico → tesi di kruskal-wallis → devo vedere il test a una variabile

esercizio: DOE

y: response time

fattori: tipo di pagina, intensità (CTT o request rate)

dato che ci limitiamo a 2 fattori, possiamo fare diverse opzioni:

- design con più livelli, req basso req alto e 4 tipi di pagine, oppure 3 tipi di pagine e diversi req rate

faremo un full factorial design

usable capacity è il massimo da non eccedere

2 livelli: 25 e 75% della usable capacity (per esempio)

A) importanza dei fattori, faccio la divisione con la variazione totale

B) verifico normalità, omoscedasticità e poi capisco quale test applicare in base alla tabella

posso limitare la durata del test, posso fare anche 1-2 minuti

Web Server (Workload Characterization):

tabella di contingenza