Lecture 4

Completely Randomized Designs: Model, Estimation, Inference

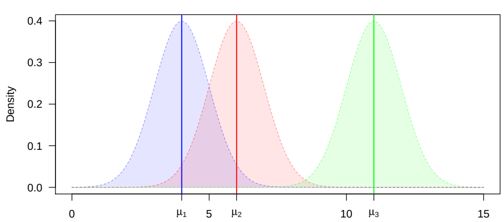
STAT 8050 Design and Analysis of Experiments
January 21, 2020

Whitney Huang
Clemson University

Notes

Means Model

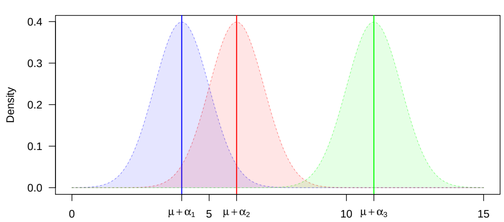
$$Y_{ij} = \mu_i + \epsilon_{ij}, \quad i = 1, \cdots, g, \quad j = 1, \cdots, n_i, \quad \epsilon_{ij} \sim \mathrm{N}(0, \sigma^2)$$



Notes

Effects Model

$$Y_{ij} = \mu + \alpha_i + \epsilon_{ij}, \quad i = 1, \cdots, g, \quad j = 1, \cdots, n_i, \quad \epsilon_{ij} \sim \mathrm{N}(0, \sigma^2)$$

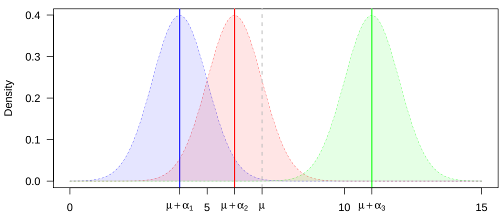


Overparameterized. Need to add a constraint so that the parameters are estimable.

Notes

Effects Model Cont'd

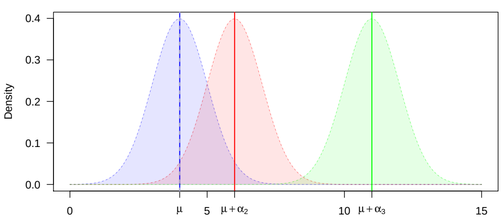
Suppose we let $\sum_{i=1}^{g} n_i\alpha_i = 0$



Notes

Effects Model Cont'd

Suppose we let $\alpha_1 = 0$



Notes

Data Layout & the Dot Notation

y_{ij} is the "observed" response for the j^{th} experimental unit to treatment i.

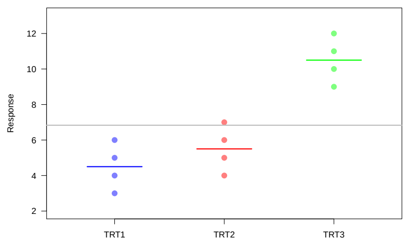
Treatment	Observations				Totals	Averages
1	y_{11}	y_{12}	$\cdots$	y_{1n_1}	$y_{1\cdot}$	$\bar{y}_{1\cdot}$
2	y_{21}	y_{22}	$\cdots$	y_{2n_2}	$y_{2\cdot}$	$\bar{y}_{2\cdot}$
$\vdots$	$\vdots$	$\vdots$	$\ldots$	$\vdots$	$\vdots$	$\vdots$
g	y_{g1}	y_{g2}	$\cdots$	y_{gn_g}	$y_{g\cdot}$	$\bar{y}_{g\cdot}$
					$y_{\cdot\cdot}$	$\bar{y}_{\cdot\cdot}$

Notes

ANOVA

Decomposition of y_{ij}: $y_{ij} = \bar{y}_{\cdot\cdot} + (\bar{y}_{i\cdot} - \bar{y}_{\cdot\cdot}) + (y_{ij} - \bar{y}_{i\cdot})$

$$\Rightarrow \underbrace{\sum_{i=1}^{g}\sum_{j=1}^{n_i}(y_{ij} - \bar{y}_{\cdot\cdot})^2}_{\text{SS}_T} = \underbrace{\sum_{i=1}^{g} n_i(\bar{y}_{i\cdot} - \bar{y}_{\cdot\cdot})^2}_{\text{SS}_{TRT}} + \underbrace{\sum_{i=1}^{g}\sum_{j=1}^{n_i}(y_{ij} - \bar{y}_{i\cdot})^2}_{\text{SS}_E}$$



Notes

ANOVA Table

Source	df	SS	MS	EMS
Treatment	$g-1$	SS_{TRT}	$\text{MS}_{TRT} = \frac{\text{SS}_{TRT}}{g-1}$	$\sigma^2 + \frac{\sum_{i=1}^{g} n_i\alpha_i^2}{g-1}$
Error	$N-g$	SS_E	$\text{MS}_E = \frac{\text{SS}_E}{N-g}$	σ^2
Total	$N-1$	SS_T		

$$\text{SS}_T = \sum_{i=1}^{g}\sum_{j=1}^{n_i}(y_{ij} - \bar{y}_{\cdot\cdot})^2 = \sum_{i=1}^{g}\sum_{j=1}^{n_i} y_{ij}^2 - \frac{y_{\cdot\cdot}^2}{N}$$

$$\text{SS}_{TRT} = \sum_{i=1}^{g} n_i(\bar{y}_{i\cdot} - \bar{y}_{\cdot\cdot})^2 = \sum_{i=1}^{g}\frac{y_{i\cdot}^2}{n_i} - \frac{y_{\cdot\cdot}^2}{N}$$

$$\text{SS}_E = \sum_{i=1}^{g}\sum_{j=1}^{n_i}(y_{ij} - \bar{y}_{i\cdot})^2 = \sum_{i=1}^{g}\sum_{j=1}^{n_i} y_{ij}^2 - \sum_{i=1}^{g}\frac{y_{i\cdot}^2}{n_i} = \text{SS}_T - \text{SS}_{TRT}$$

Notes

F-Test

Testing for treatment effects

$$H_0 : \alpha_i = 0 \quad \text{for all } i$$
$$H_a : \alpha_i \neq 0 \quad \text{for some } i$$

Test statistics: $F = \frac{\text{MS}_{TRT}}{\text{MS}_E}$. Under H_0, the test statitic follows an F-distribution with $g-1$ and $N-g$ degrees of freedom

Reject H_0 if

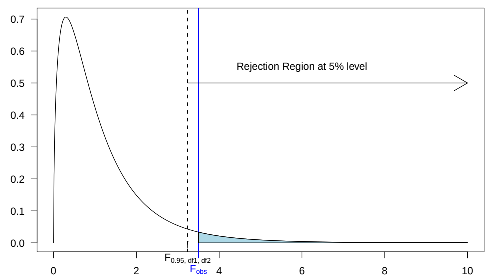
$$F_{obs} > F_{g-1,N-g;\alpha}$$

for an α-level test, $F_{g-1,N-g;\alpha}$ is the $100 \times (1-\alpha)$% percentile of a central F-distribution with $g-1$ and $N-g$ degrees of freedom.

The P-value of the F-test is the probability of obtaining F at least as extreme as F_{obs}, that is, $\mathrm{P}(F > F_{obs}) \Rightarrow$ reject H_0 if P-value $< \alpha$.

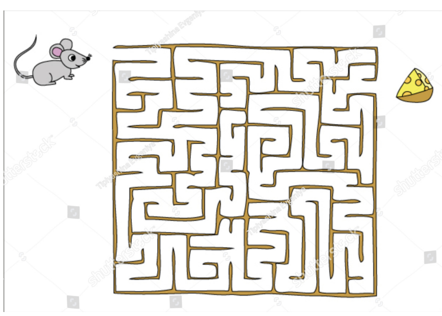
Notes

F Distribution and the F-Test



Notes

Example



Source: https://www.shutterstock.com/image-vector/find-your-way-cheese-mouse-maze-232569073

Notes

Notes