Completely Randomized Designs with One Factor II

Sean Hellingman ©

Design for Data Science (ADSC2030)

shellingman@tru.ca

Winter 2025

Topics

Introduction

- **Completely Randomized Designs** (CRD) are used when there is only one factor under study and the experimental units are homogeneous.
- Once the model is estimated, we need to check the validity of our inferences.
- There is a set of assumptions that must be satisfied in order for our model to be valid.
  - We will cover some techniques to fix common problems.
- *Note: Example code continues from part I*

Model Assumptions

- There are two main assumptions required for the validity of the analysis based on the linear model:

  1. Constancy of the variance of the experimental error, σ^2, across all levels of the treatment factor (constant variance).

  2. The experimental errors follow a normal distribution.

Constant Variance

- Examine a scatterplot of the residuals VS the factor levels to see if the variability is approximately the same across the factor levels.
- Plot the model residuals VS the means of the response for each factor to see if the variability increases when the mean level of the response increases.
- We can use the `ncvTest()` similar to a (Breusch–Pagan test) in R with the null hypothesis being a constant variance.

Normality

- Examine the histogram of the *standardized residuals* for approximate normality.
- Q-Q plot of the *standardized residuals*.
  - In R we can also use: `plot(model, 2)`
- We can use a Shapiro–Wilk test on the *standardized residuals*.
  - Shapiro–Wilk in R: `shapiro.test(standardized.residuals)`

Using R

- par(mfrow = c(2,2)) *Examine simultaneously*
- plot(anova.model, which=5) *standardized residuals versus the factor levels*
- plot(anova.model, which=1) *plot of residuals versus the cell means or fitted values*
- plot(anova.model, which=2) *Q-Q*
- plot(residuals(anova.model) $\sim$ Exp.Unit, main="Residuals vs Exp. Unit", data=data)

  abline(h = 0, lty = 2) *plot of residuals versus experimental unit numbers*

Example 1

- Use the techniques outlined in the previous slides to see if the assumptions hold for the bread example.

- Comment on your findings.

Heterogeneity of Variances

- One of the most common violations is the assumption of a constant variance.
  - Usually occurs when the relationship is non-linear.
  - As the mean increases so does the variance.

- One common solution is to transform the data prior to analysis.

Box-Cox Power Transformation I

- We can examine the scatterplot to help determine if we need a variance stabilizing transformation.

- In 1964 Box and Cox proposed a series of transformations that usually work well ($Y = y^{\lambda}$).
  - If the variance increases with the mean, choose $\lambda < 1$.
  - If the variance decreases as the mean increases, choose $\lambda > 1$

Box-Cox Power Transformation II

Relationship Between σ and μ	λ	Transformation
$\sigma \propto \mu^2$	-1	Reciprocal
$\sigma \propto \mu^{3/2}$	$-\frac{1}{2}$	Square Root of Reciprocal
$\sigma \propto \mu$	0	Log
$\sigma \propto \mu^{1/2}$	$\frac{1}{2}$	Square Root

- It is very common that $\sigma \propto \mu$.
- **In practice, the value of λ that minimizes the error sum of squares would be the most appropriate.**

Box-Cox Transformation in R

- We can use the *MASS* package to obtain the *best* λ.
- In R:
  - `bc <- boxcox(model)`
  - `lambda <- bc$x[which.max(bc$y)]`
    To transform your data:
  - `transform.df <- transform(df, variable = variable^(lambda))`
- Then you can conduct the ANOVA analysis again.

Example 2

- Use the *MASS* package to Box-Cox power transformation on the ANOVA we conducted in Example 4.
- What is the value of λ?
- In R, perform an ANOVA test to determine if there is any significant differences in the means of the response for each of the three factors in the bread example.
- Conduct the diagnostics from Example 5 on the new model.
  - Do the assumptions hold?

Box-Cox Power Transformation III

- When variance heterogeneity is very pronounced, the Box-Cox transformation can greatly increase the sensitivity in detecting treatment effects.
- The Box-Cox transformation can usually help with the normality assumption.
- *If one of the simpler transformations (slide 11) falls within the confidence interval, you may select that transformation.*

Distribution Transformations

- The distribution assumption for the model for the CRD is assumed to be normal.
- It may be that we *know* the data follows some other distribution (Binomial, Poisson, or Lognormal).
  - Then the variance would no longer be constant (Binomial $\sigma = \sqrt{np(1-p)}$)
- If the distribution is known, we can make an appropriate transformation to stabilize the variance.

Response Distribution-Based Transformations

Response Distribution	Variance in Terms of mean μ	Transformation $f(y)$
Binomial	$\frac{\mu(1-\mu)}{n}$	$\sin^{-1}\sqrt{y/n}(radians)$
Poisson	μ	$\sqrt{y}$ or $\sqrt{y+\frac{1}{2}}$
Lognormal	$c\mu^2$	$\log(y)$

- *Just to make you aware that these techniques exist.*

Weighted Least Squares I

- If the variance is not constant and is **not** related to the factor-level means we can use weighted least squares.
- **Weighted least squares** a generalization of ordinary least squares and linear regression in which knowledge of the unequal variance of observations (heteroscedasticity) is incorporated into the regression.
- A weight is assigned to each observation based on the variability.

$$W_{ii} = \frac{1}{\sigma_i^2}$$

- *Essentially, the inverse of the variance within each treatment level.*

Weighted Least Squares in R

- We can create a vector of the weights and a new model in R:
  - ```
    with(data,{ vars <- tapply(response, factor, var)
        weights <- rep(1/vars, each = r)
        model <- lm(response ~ factor, weights = weights,
           data = data)
        anova(model)
        })
    ```
- `with()` applies all statements to the data data frame.
- Calculate the variance of each group and use its inverse at the weights.
- Estimate a new linear model.
- Estimate a new ANOVA table.
    ```

Example 3

- Estimate a weighted regression model using the data from the bread example.
- How does the ANOVA table change?

Weighted Least Squares II

- The weights are essentially, the inverse of the variance within each treatment level.
- *The linear algebra required to arrive at the estimates is not covered in this course.*
- We are still trying to minimize a distance, just now each observation is weighted by variability of the factor level it belongs to.

Number of Replicates

- Replicates are needed to estimate the variance of the error in order to use the ANOVA F-test.
- The power of the test will increase as the number of replicates increases (could be expensive).
- Determining the number of replicates needed to achieve a desired power of a test falls into the category of *error control*.
- **The experimenter must select a value of practical difference (Δ).**
  - Usually obtained through the previous/pilot studies.

Power of a Test

- $1 - \beta$ is called the **power of the test** which is the probability of correctly rejecting the null hypothesis when it is indeed false.
- The power of a test depends on the true value of the population mean, the level of confidence used, and the sample size.
- Assuming that H_1 is true, there will be some overlap in the sampling distributions of μ_0 and μ_1.
- This means that the test statistic may fall into the *acceptance* region even when H_1 is true.

Power of a Test Graphics I

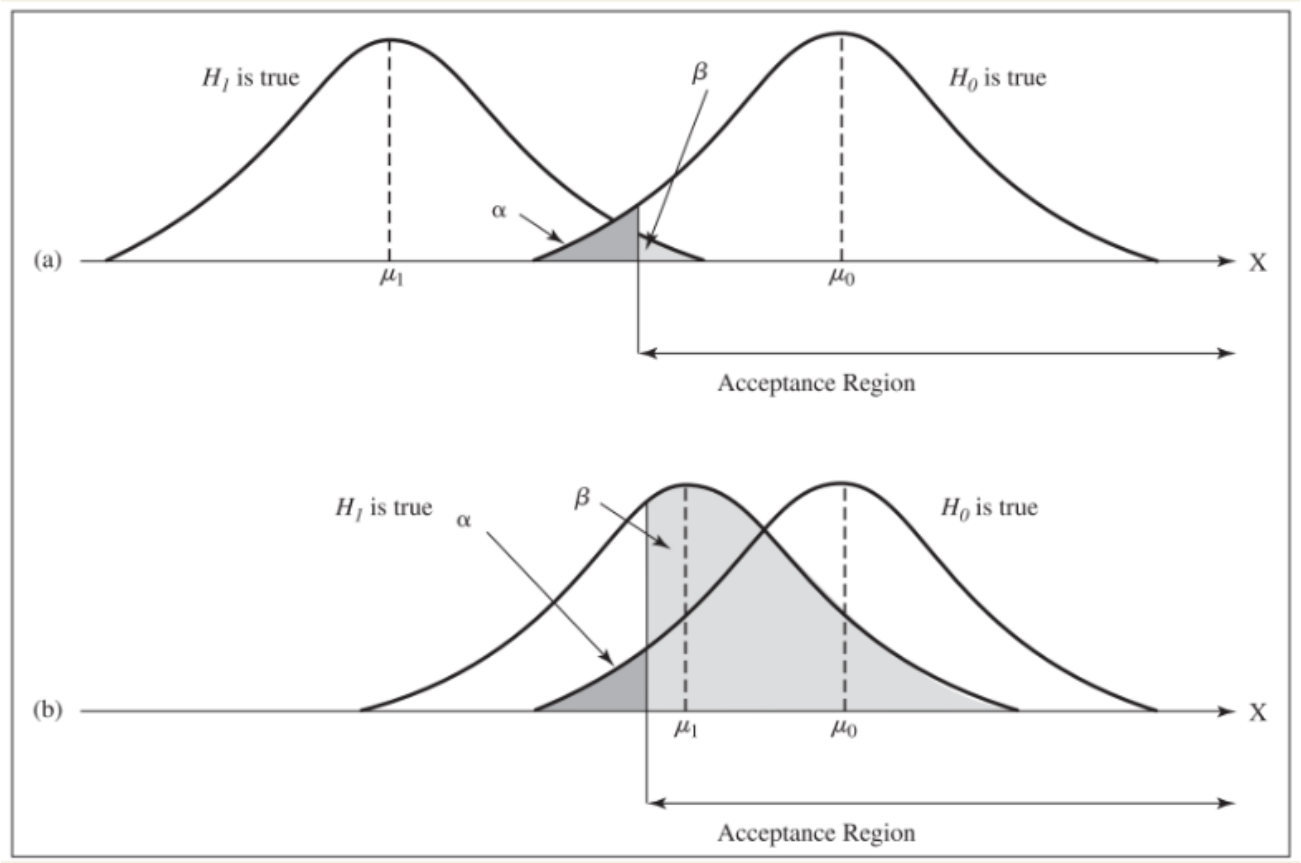

Figure: Source: (2)

Power of a Test Graphics II

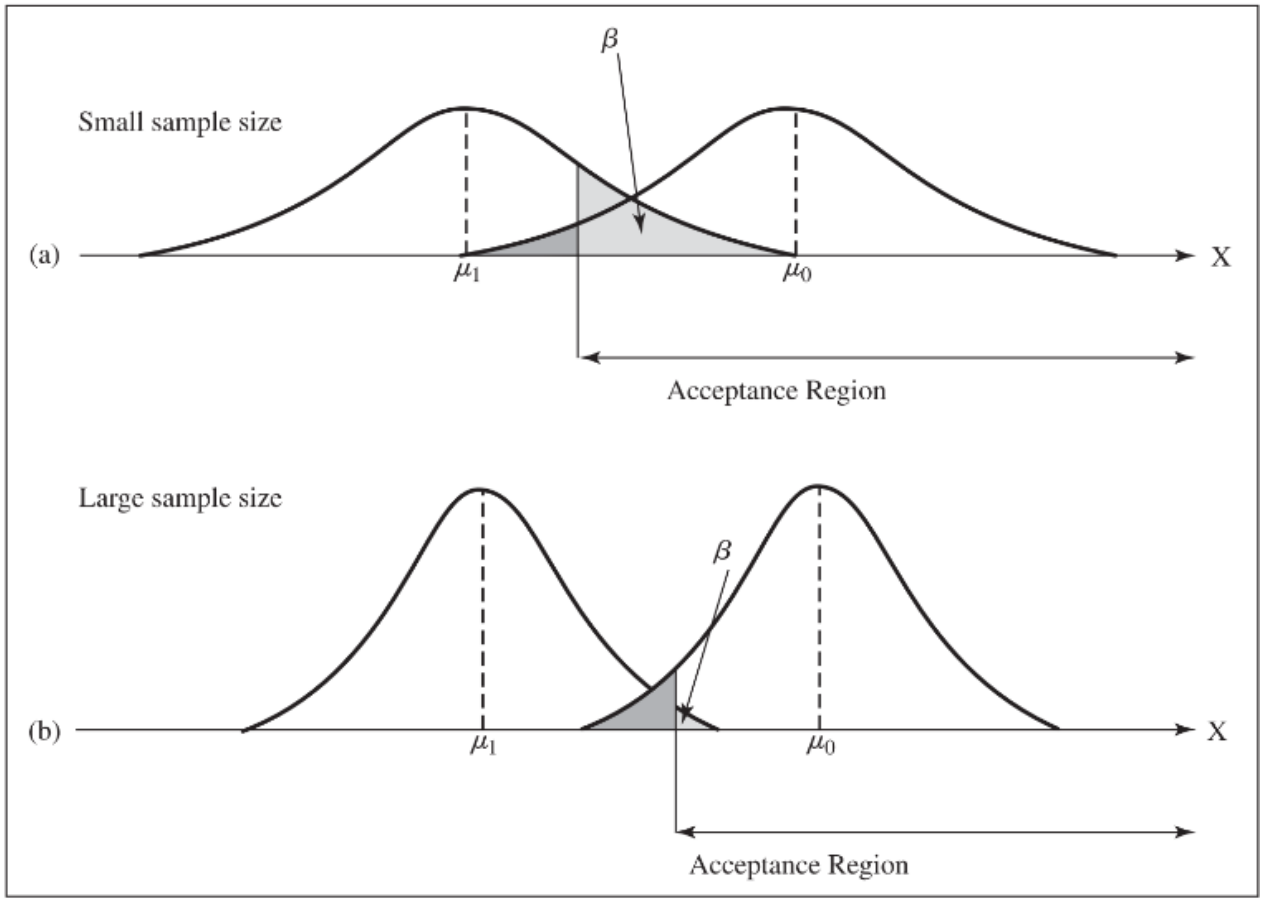

Figure: Source: (2)

Number of Replicates in R

- We can use the `Fpower()` function from the *daewr* package:
  - `library(daewr)`
  - `rmin <- s # smallest number of replicates considered`
  - `rmax <- m # largest number of replicates considered`
  - `alpha <-rep(`α`, rmax - rmin +1) #` ($\alpha = 0.05$)
  - `sigma <- sqrt(sigma_squared) # from pilot study`
  - `nlev <- t # number of factor levels`
  - `nreps <-rmin:rmax`
  - `Delta <- delta # practical difference`
  - `power <-Fpower1(alpha,nlev,nreps,Delta,sigma)`
  - `power # results`
- Rule of thumb: *The number of replicates that result in a power between 0.80 and 0.90 is usually sufficient for most designs.*

Example 4

- Use the following information to determine how many replicates to use for the bread example:
  - Assume the variance of the experimental error was obtained from the pilot experiment: $\hat{\sigma}^2 = 2.1$
  - Fewest replicates under consideration: 2
  - Most replicates under consideration: 6
  - Selected practical difference: $\Delta = 3$
  - Significance level: $\alpha = 0.05$

After F-test

- When we reject the null hypothesis of the ANOVA, we can assume that *at least* one mean is significantly different from another.
- This does not imply that all of the means are significantly different, nor which means are different.
- Further investigation into the differences can be conducted:
  1. The investigator will have pre-planned comparisons they would like to make.
  2. The investigator has no idea what differences to look for.

Pre-planned Comparisons I

- We may examine specific hypotheses about the response mean based on different factor levels (at specific points).

- For example (assuming 4 treatment factors):
  - H_0: $\mu_1 = \frac{1}{3}(\mu_2 + \mu_3 + \mu_4)$
  - H_0: $\mu_2 = \mu_3$
  - H_0: $\mu_3 = \mu_4$

Pre-planned Comparisons in R

- We may use the *daewr* and *gmodels* packages in R to test for specific (pre-planned) differences (at specific points).

- In R:
  - `model <- aov(response ~ factor, data = data) # Model`
  - `con <- matrix(c(1, -1/3, -1/3, -1/3, 0, 1, -1, 0, 0, 0, 1, -1 ), 4, 3) # row matrix of comparisons`
  - `L <- t(con) # transpose the matrix`
  - `rownames(L) <- c("Comparison.1", "Comparison.2", "Comparison.3") # Label the comparisons`
  - `options(digits=3) # round to 3 digits`
  - `fit.contrast(model, "factor", L) # Test the hypotheses`

- *Note: You will have to change the con matrix for your own comparisons these values come from the example on the previous slide.*

Pre-planned Comparisons II

- When factors are quantitative (numeric), pre-planned comparisons can be done to examine the significance of linear or higher order polynomial trends in the response.

- In R:
  - `contrasts(data$factor) <- contr.poly(order) # Set the order of the polynomial (`$t-1$`)`
  - `contrasts(data$factor) # Examine the contrasts`
  - `model <- aov( response ~ factor, data)`
  - `summary.lm(model) # significance of trends`

Example 5

1. Test the following (pre-planned) hypotheses from the bread example:
   - H_0: $\mu_{45} = \frac{1}{2}(\mu_{35} + \mu_{40})$
   - H_0: $\mu_{35} = \mu_{40}$

2. Test the significance of a linear trend and a second-degree polynomial ($t - 1 = 2$) trend in the bread example.

Unplanned Comparisons

- If the comparisons are not pre-planned, we might want to make comparisons based on the means found in the data.
  - This implies that we make all possible comparisons.

- **The possibility of making a Type I error increases when we conduct all possible comparisons.**
  - Tukey's HSD (or honestly significant difference) method adjusts the critical region to account for this possibility.
  - Can also use the Newman-Keuls method.

Tukey's HSD in R

- In R we can make the pairwise comparisons:
  - `model <- aov(response ~ factor, data=data)`
  - `model.tukey <- TukeyHSD(model,ordered = TRUE) # All pairs`
  - `model.tukey` # Print results

Newman-Keuls method in R

- In R we can make the pairwise comparisons:
  - `library(agricolae)`
  - `model <- aov(response ~ factor, data=data)`
  - `compare <- SNK.test(model,"factor",alpha = 0.05)`
  - `compare` # Print results
- In the table generated by `compare$groups` groups with the same indicator are not significantly different.

Example 6

- Use both the Tukey's HSD and Newman-Keuls methods to identify any differences in the individual effects of the rise times on the height of the bread dough.

Compare all Means to a Control or the *Best*

- In some experiments we want to compare our results to a default or *control* level.
- We can use a method developed by Dunnett (1955) that can control the experiment-wise type I error rate in R:
  - `library(multcomp)`
  - `results.dun <- glht(model,linfct = mcp(treat = "Dunnett"), alternative = "greater")`
  - `summary(results.dun)`
- Compares the mean at each factor level to a control.
- *Control is set alphabetically in R.*

`glht()` Comments

- We usually have an idea of the difference that we would like to test.
  - Therefore we usually look at one-tailed tests.
- To change the nature of the comparison in R:
  - `alternative = "less" # all means are less`
  - `alternative = "two.sided" # all means are different`
- Change the factor levels to compare with the *best*.

Example 11

- Use the `glht()` function in R to determine if the average heights of the 40 and 45 minute rise-time factors are significantly larger than the reference of 35 minutes.

CRD Selection

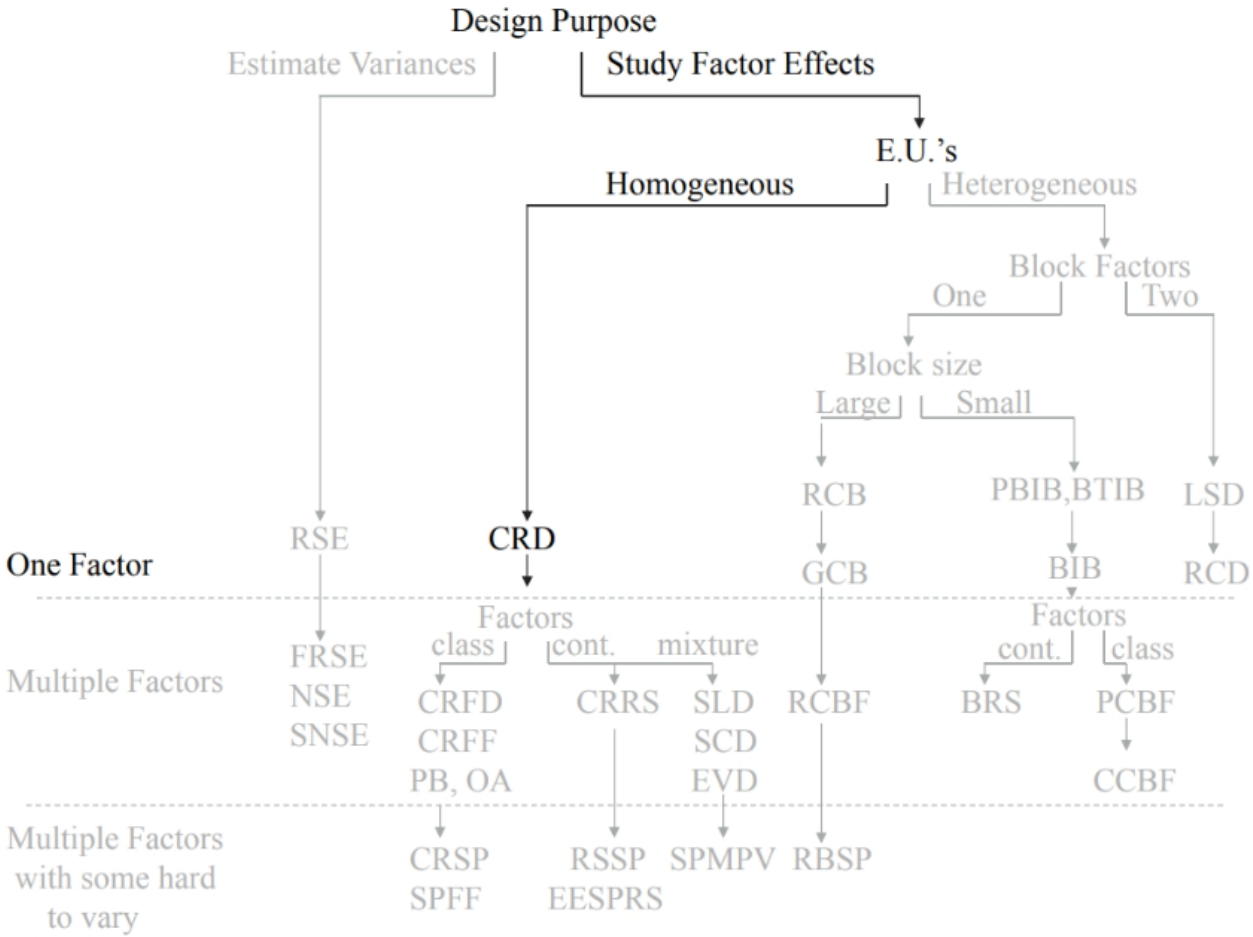

Figure: Source: (1)

Exercise 1

- Use R to construct a CRD with one factor for the following situation:
  - Set your seed to 2030.
  - Want to design an experiment to determine the tensile strength of different cotton/synthetic blends.
    - Four cotton percentages (25%, 30%, 35%, and 40%)
    - Five replications (r) for each cotton percentage

Exercise 2

- Assume the following results vector (from Exercise 1):
  - Results = (14,19,25,8,7,25,22,10,10,18,18,19,10,11,19,18,15,23,12,11)
- Estimate the appropriate model to estimate the effects of the cotton content on the tensile strength response variable.
- Use the `fit.contrast()` function to estimate the average difference in the means for the second and fourth levels of the treatment factor.

Exercise 3

- Take some time to apply the techniques that we have learned to test various assumptions about your results from Exercise 1 and Exercise 2.

References & Resources

1. Lawson, J. (2014). *Design and Analysis of Experiments with R (Vol. 115)*. CRC press.

2. Evans, J. R., Olson, D. L., & Olson, D. L. (2007). *Statistics, data analysis, and decision modeling.* Upper Saddle River, NJ: Pearson/Prentice Hall.

- fit.contrast
- Box-Cox Transformations
- Matrix Transpose
- Tukey's HSD
- Newman-Keuls
- Generalized Linear Hypothesis Test (glht)