2.9 The Random Effects Model

- So far we have assumed the factor levels were fixed. That is, the factor levels were set at fixed levels by the experimenter.
- In many one-factor CRDs, the a factor levels are *randomly selected from a population* of levels. For this type of experiment, the factor is called a **random factor**, and the associated effects are called **random effects**.
- In theory, for random effects, we assume the population is infinite. In practice, it is acceptable if the number of randomly selected factor levels (a) is small relative to the number of levels in the population (N). In general, we want $a/N < .10$ (or, $< 10\%$).
- The **random effects model** for a one-factor CRD is:

$$y_{ij} = \tag{5}$$

  where both τ_i and ϵ_{ij} are random variables. The model assumptions are

  - The ϵ_{ij}'s are IID $N(0, \sigma^2)$ and the τ_i's are IID $N(0, \sigma_\tau^2)$.
  - τ_i and ϵ_{ij} are independent for all i, j.

- τ_i is a random variable with $\text{Var}(\tau_i) = \sigma_\tau^2$. σ_τ^2 is the variance associated with the distribution or population of all τ_i's. We assume τ_i is independent of the random error ϵ_{ij} which has variance σ^2.
- The variances σ_τ^2 and σ^2 are called .
- The hypotheses of interest involve the variance component σ_τ^2:

  and

- If $\sigma_\tau^2 = 0$, then the random τ_i effects are identical. In this case, the variability of the $\widehat{\tau}_i$ estimates ($i = 1, 2, \ldots, a$) should be close to 0 in comparison to the MSE.
- If $\sigma_\tau^2 > 0$, then the random τ_i effects are not identical. In this case, the variability of the $\widehat{\tau}_i$ estimates ($i = 1, 2, \ldots, a$) should be large in comparison to the MSE.
- Testing hypotheses about the equality of means is inappropriate in the random effects case. Therefore, we do not perform a multiple comparison procedure to compare means.
- The ANOVA table for a random factor is the same as the ANOVA table for a fixed factor with $SS_T = SS_{trt} + SS_E$.
- To see this we need to look at the expected mean squares for the random effects model in (5).

2.9.1 Expected Mean Squares

- Theoretically, the expected values of MS_E and MS_{trt} are

$$E(MS_E) = \sigma^2 \tag{6}$$
$$E(MS_{trt}) = \sigma^2 + \tag{7}$$

  where $n_0 = \dfrac{1}{a-1}\left[N - \dfrac{\sum_{i=1}^a n_i^2}{N}\right]$.

- If all of the sample sizes are equal ($n_i = n$ for $i = 1, 2, \ldots, a$), then $n_0 = n$.
- If $H_0 : \sigma_\tau^2 = 0$ is **true**, then $E(MS_{trt}) = \sigma^2 + 0 = \sigma^2$.

If $H_0 : \sigma_\tau^2 = 0$ is **false**, then $\sigma_\tau^2 > 0$, and

$$E(MS_{trt}) \;=\; \sigma^2 + (\textit{positive quantity}) \;\longrightarrow$$

- As the variability among the τ_i's increases, the F-ratio

$$F = \frac{E(MS_{trt})}{E(MS_E)} \;=$$

  also increases.

- $F_0 = \dfrac{MS_{trt}}{MS_E}$ is an estimate of F, and $H_0 : \sigma_\tau^2 = 0$ will be rejected for large values of F_0.

- The ANOVA table for the random effects model is the same as the ANOVA for the fixed effects model. However, **the hypotheses being tested are different.**

- Remember: The hypotheses for the random effects model apply to a distribution or population (with variance component σ_τ^2) while the hypotheses for a fixed effects model apply to equality of fixed treatment effects (τ_i's) or means (μ_i's).

2.9.2 Estimation of Variance Components

- For the random effects model, we are interested in finding estimates $\widehat{\sigma}^2$ and $\widehat{\sigma}_\tau^2$ of the variance components σ^2 and σ_τ^2.

  If we replace $E(MS_E)$ with MS_E in equation (6), we get $\widehat{\sigma}^2 =$

  If we replace σ^2 with $\widehat{\sigma}^2 = MS_E$ and $E(MS_{trt})$ with MS_{trt} in equation (7), we get

$$MS_{trt} \;\approx$$

  Solving for σ_τ^2 gives

$$\widehat{\sigma}_\tau^2 \;=$$

- This estimation approach can give a negative estimate of σ_τ^2 ($\widehat{\sigma}_\tau^2 < 0$). But we know that a variance component cannot be negative. The following are 3 possible ways to handle this situation:

  1. Assume $\sigma_\tau^2 = 0$ and the negative estimate occurred due to random sampling. The problem is that using zero instead of a negative number can affect other estimates.
  2. Estimate σ_τ^2 using the REML (restricted maximum likelihood) method because it always yields a nonnegative estimate. This method will adjust other variance component estimates. REML methods are included in *SAS*.
  3. Assume the model is incorrect, and examine the problem in another way. For example, add or remove an effect from the model, and then analyze the new model.

2.9.3 A Random Effects Example

A company supplies a customer with a large number of batches of raw materials used in a chemical production process. The customer wants a high percentage of usable chemical to be produced from the raw material. The customer is concerned that there may be significant variation among the batches (which is not good for the production process). An experiment was run. Ten random batches are selected, and four random samples are taken from each batch. The response is 'percent usable chemical'. The experimental data are

Batch									
1	2	3	4	5	6	7	8	9	10
74	68	75	72	79	74	72	72	79	72
76	71	77	74	81	73	74	75	81	68
75	72	77	73	79	75	73	77	80	68
75	70	75	72	80	73	75	74	77	69

- The model is $y_{ij} = \mu + \tau_i + \epsilon_{ij}$ with

$$\epsilon_{ij} \sim IIDN(0, \sigma^2) \qquad \tau_i \sim IIDN(0, \sigma^2_\tau) \quad i = 1, 2, \ldots, 10$$

- The ANOVA table below indicates the p-value $< .0001$ for testing

$$H_0 : \sigma^2_\tau = 0 \qquad \text{and} \qquad H_1 : \sigma^2_\tau > 0.$$

  Therefore we **reject** $H_0 : \sigma^2_\tau = 0$, and conclude that there is significant variability in percent usable chemical in the population of batches.

- The estimates of the variance components are

$$\widehat{\sigma}^2 \; = \; MSE \; = \qquad\qquad \widehat{\sigma}^2_\tau \; = \; \frac{MS_{batch} - MS_E}{n} \; =$$

- In the SAS code, include a RANDOM statement to perform the F-test and output the expected mean squares. Use the VARCOMP procedure to generate estimates of the variance components.

ANOVA WITH RANDOM BATCH EFFECTS

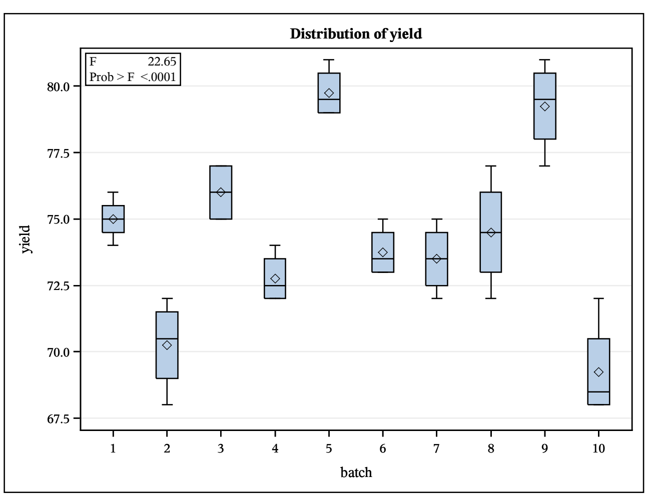
The GLM Procedure

Dependent Variable: yield

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	9	411.1000000	45.6777778	22.65	<.0001
Error	30	60.5000000	2.0166667		
Corrected Total	39	471.6000000			

R-Square	Coeff Var	Root MSE	yield Mean
0.871713	1.908728	1.420094	74.40000

Source	DF	Type III SS	Mean Square	F Value	Pr > F
batch	9	411.1000000	45.6777778	22.65	<.0001



Level of batch	N	yield Mean	yield Std Dev
1	4	75.0000000	0.81649658
2	4	70.2500000	1.70782513
3	4	76.0000000	1.15470054
4	4	72.7500000	0.95742711
5	4	79.7500000	0.95742711
6	4	73.7500000	0.95742711
7	4	73.5000000	1.29099445
8	4	74.5000000	2.08166600
9	4	79.2500000	1.70782513
10	4	69.2500000	1.89296945

SAS NOTATION: $VAR = \sigma^2$

(*)

Source	Type III Expected Mean Square
batch	Var(Error) + 4 Var(batch)

$= \sigma^2 + \sigma^2_{BATCH}$

The GLM Procedure
Tests of Hypotheses for Random Model Analysis of Variance

Dependent Variable: yield

(*) CORRECT ANOVA

Source	DF	Type III SS	Mean Square	F Value	Pr > F
batch	9	411.100000	45.677778	22.65	<.0001
Error: MS(Error)	30	60.500000	2.016667		

VERY STRONG EVIDENCE TO REJECT $H_0: \sigma^2_{BATCH} = 0$

THIS TABLE IS THE SAME AS THE ANOVA TABLE ASSUMING BATCHES WERE FIXED. THIS WILL NOT BE TRUE LATER IN THIS COURSE WITH FACTORIAL AND NESTED EFFECTS MODELS.

ANOVA WITH RANDOM BATCH EFFECTS
VARIANCE COMPONENT ESTIMATION

Variance Components Estimation Procedure

REML Iterations			
Iteration	Objective	Var(batch)	Var(Error)
0	55.4378867503	10.9152777778	2.0166666667
1	55.4378867503	10.9152777778	2.0166666667

(**)

Convergence criteria met.

REML Estimates	
Variance Component	Estimate
Var(batch)	10.91528
Var(Error)	2.01667

$\hat{\sigma}^2_{BATCH} = \dfrac{MS_{BATCH} - MS_E}{n}$

$\hat{\sigma}^2 = MS_E$

SAS code for random effects analysis

```
**************************************************;
*** A COMPLETELY RANDOMIZED ONE-FACTOR DESIGN ***;
***     WITH RANDOM BATCH EFFECTS             ***;
**************************************************;
TITLE 'ANOVA WITH RANDOM BATCH EFFECTS';
DATA in;
  DO batch = 1 TO 10;
  DO rep = 1 TO 4;
     INPUT yield @@; OUTPUT;
  END; END;
 LINES;
 74 76 75 75   68 71 72 70   75 77 77 75   72 74 73 72   79 81 79 80
 74 73 75 73   72 74 73 75   72 75 77 74   79 81 80 77   72 68 68 69
 ;
PROC GLM DATA=IN;
     CLASS batch;
     MODEL yield = batch / SS3;
     MEANS batch;
     RANDOM batch / TEST;
   [ OUTPUT OUT=diag P=pred R=resid;]
PROC VARCOMP DATA=in METHOD=REML;
     CLASS batch;
     MODEL yield = batch ;
TITLE2 'VARIANCE COMPONENT ESTIMATION';
RUN;
```

GENERATES (*) OUTPUT

← NOT NEEDED UNLESS YOU WANT TO LOOK AT VALUES FOR $\hat{y}$ AND RESIDUALS.
USE: PROC PRINT DATA=diag;

GENERATES (**) OUTPUT

2.9.4 Confidence Intervals for Variance Components

Given the normality and independence assumptions of the random effects model, we can generate various confidence intervals related to the variance components.

1. For σ^2: Because $\dfrac{SS_E}{\sigma^2} = \dfrac{(N-a)MS_E}{\sigma^2} \sim \chi^2_{N-a}$, a $100(1-\alpha)\%$ confidence interval for σ^2 is

$$\leq$$

2. For $\dfrac{\sigma^2_\tau}{\sigma^2}$: Because MS_E and MS_{trt} are independent,

$$\frac{MS_{trt}/(\sigma^2 + n_0\sigma^2_\tau)}{MS_E/\sigma^2} \quad = \quad F_0\left(\frac{\sigma^2}{\sigma^2 + n_0\sigma^2_\tau}\right) \quad \sim \quad F(a-1, N-a).$$

Let $F_L = F_{1-\alpha/2}(a-1, N-a)$ and $F_U = F_{\alpha/2}(a-1, N-a)$. Then,

$$\begin{aligned}
1-\alpha &= P\left[F_L \;\leq\; F_0\left(\frac{\sigma^2}{\sigma^2+n_0\sigma^2_\tau}\right) \;\leq\; F_U\right] \\
&= P\left[\frac{1}{F_L} \;\geq\; \frac{1}{F_0}\left(\frac{\sigma^2+n_0\sigma^2_\tau}{\sigma^2}\right) \;\geq\; \frac{1}{F_U}\right] \;=\; P\left[\frac{1}{F_U} \;\leq\; \frac{1}{F_0}\left(\frac{\sigma^2+n_0\sigma^2_\tau}{\sigma^2}\right) \;\leq\; \frac{1}{F_L}\right] \\
&= P\left[\frac{1}{F_U} \;\leq\; \frac{1}{F_0}\left(1+\frac{n_0\sigma^2_\tau}{\sigma^2}\right) \;\leq\; \frac{1}{F_L}\right] \\
&= P\left[\frac{1}{n_0}\left(\frac{F_0}{F_U}-1\right) \;\leq\; \frac{\sigma^2_\tau}{\sigma^2} \;\leq\; \frac{1}{n_0}\left(\frac{F_0}{F_L}-1\right)\right]
\end{aligned}$$

Thus, a $100(1-\alpha)\%$ confidence interval for σ^2_τ/σ^2 is (L, U) where

$$L \;= \qquad\qquad \text{and} \qquad U \;=$$

3. For $\dfrac{\sigma^2_\tau}{\sigma^2_\tau+\sigma^2}$: Note that

$$\begin{aligned}
1-\alpha &= P[\,L \leq \sigma^2_\tau/\sigma^2 \leq U\,] \;=\; P[\,1+L \leq 1+\sigma^2_\tau/\sigma^2 \leq 1+U\,] \\
&= P[\,1+L \leq \frac{\sigma^2+\sigma^2_\tau}{\sigma^2} \leq 1+U\,] \;=\; P\left[\frac{1}{1+L} \geq \frac{\sigma^2}{\sigma^2+\sigma^2_\tau} \geq \frac{1}{1+U}\right] \\
&= P\left[1-\frac{1}{1+L} \leq 1-\frac{\sigma^2}{\sigma^2+\sigma^2_\tau} \leq 1-\frac{1}{1+U}\right] \\
&= P\left[\frac{L}{1+L} \leq \frac{\sigma^2_\tau}{\sigma^2+\sigma^2_\tau} \leq \frac{U}{1+U}\right]
\end{aligned}$$

Thus, $\left(\dfrac{L}{1+L}, \dfrac{U}{1+U}\right)$ is a $100(1-\alpha)\%$ confidence interval for $\dfrac{\sigma^2_\tau}{\sigma^2_\tau+\sigma^2}$ which represents the proportion of the total variability attributable to the variability among the treatments.

4. For σ^2_τ: There is no closed form for an exact confidence interval for σ^2_τ. The following formula produces an approximate confidence interval.

(L_τ, U_τ) is an approximate $100(1-\alpha)\%$ confidence interval for σ^2_τ where

$$L_\tau \;=\; \frac{SS_{trt}\left(1-\frac{F_U}{F_0}\right)}{n_0\chi^2_{\alpha/2,a-1}} \qquad \text{and} \qquad U_\tau \;=\; \frac{SS_{trt}\left(1-\frac{F_L}{F_0}\right)}{n_0\chi^2_{1-\alpha/2,a-1}}$$

Example of 95% Confidence Intervals for Variance Components

Dependent Variable:	yield

Type 1 Analysis of Variance				
Source	DF	Sum of Squares	Mean Square	Expected Mean Square
batch	4	147.733333	36.933333	Var(Error) + 3 Var(batch)
Error	10	18.000000	1.800000	Var(Error)
Corrected Total	14	165.733333		

Type 1 Estimates			
Variance Component	Estimate	95% Confidence Limits	
Var(batch)	11.71111	3.77313	101.00296
Var(Error)	1.80000	0.87877	5.54363

- **We are 95% confident that the observed variability attributable to random (replication) error is between .8788 and 5.5436.**
- **We are 95% confident that the ratio of variance attributable to differences in batches to the variance attributable to random error is between 1.197 and 60.156.**
- **We are 95% confident that the proportion of variability of an observation attributable to differences in batches is between .5448 and .9945.**
- **We are 95% confident that the observed variability attributable to random batch-to-batch variability is between 3.773 and 101.003.**

In this example, $SS_E = 18.0$, $SS_{trt} = 147.7\overline{3}$, $F_0 = 20.51852$, and $n_0 = n = 3$, with

- $\chi^2(.975, 10) = 3.2470$ and $\chi^2(.025, 10) = 20.4832$
- $F_L = F(.975, 4, 10) = 0.11307$ and $F_U = F(.075, 4, 10) = 4.46834$
- $\chi^2(.975, 4) = 0.4844$ and $\chi^2(.025, 4) = 11.1433$

2.10 Tests for Homogeneity of Variance

- In an ANOVA, one assumption is the **homogeneity of variance (HOV)** assumption or (**equal variances assumption**. That is, in an ANOVA we assume that treatment variances are equal:

$$H_0 : \sigma_1^2 = \sigma_2^2 = \cdots = \sigma_a^2.$$

- Moderate deviations from the assumption of equal variances do not seriously affect the results in the ANOVA. Therefore, the ANOVA is robust to small deviations from the HOV assumption. We only need to be concerned about large deviations from the HOV assumption.
- Evidence of a large heterogeneity of variance (or unequal variances) problem is easy to detect in residual plots. Residual plots also provide information about patterns among the variance.
- Some researchers like to perform a hypothesis test to validate the HOV assumption. We will consider two common HOV tests: Levene's Test and the Brown-Forsythe Test.
- Both tests require replication for each of the a treatments.
- These tests are not powerful for detecting small or moderate differences in variances. This is okay because we are only concerned about large deviations from the HOV assumption.
- A third test, called Bartlett's Test, will not be covered because it is no longer popular due to its sensitivity to nonnormality.

2.10.1 Levene's Test and the Brown-Forsythe Tests

To perform Levene's Test:

1. Calculate each $z_{ij} = |e_{ij}|$ where residual $e_{ij} = y_{ij} - \overline{y}_{i\cdot}$.
2. Run an ANOVA on the set of z_{ij} values.
3. If p-value $\leq \alpha$, reject H_0 and conclude the variances are not all equal.

- Levene's Test is robust because the true significance level is very close to the nominal significance level for a large variety of distributions.
- It is not sensitive to symmetric heavy-tailed distributions (e.g., double exponential and t distributions).

To perform the Brown-Forsythe Test:

1. Calculate each $z_{ij} = |y_{ij} - \widetilde{y}_i|$ where $\widetilde{y}_i$ is the median for the i^{th} treatment.
2. Run an ANOVA on the set of z_{ij}'s.
3. If p-value $\leq \alpha$, reject H_o and conclude the variances are not all equal.

- The Brown-Forsythe Test is relatively insensitive to departures from normality.
- It is not sensitive to skewed distributions (e.g., χ^2) and extremely heavy-tailed distributions (e.g., Cauchy). In these cases, it is more robust than Levene's Test.

2.10.2 Example of Levene's Test and the Brown-Forsythe Test

A textile company has five looms that weave cloth. The company is concerned that there may be significant variability in the strengths of the cloth produces by the looms. Five random samples of cloth are taken from the cloth produced by each loom. Each sample is tested and the strength of the cloth is recorded. The data are:

Loom				
1	2	3	4	5
14.0	13.9	14.1	13.6	13.8
14.1	13.8	14.2	13.8	13.6
14.2	13.9	14.1	14.0	13.9
14.0	14.0	14.0	13.9	13.8
14.1	14.0	13.9	13.7	14.0

SAS Output for HOV Tests

The GLM Procedure

Dependent Variable: cloth

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	4	0.34160000	0.08540000	5.77	0.0030
Error	20	0.29600000	0.01480000		
Corrected Total	24	0.63760000			

R-Square	Coeff Var	Root MSE	cloth Mean
0.535759	0.872957	0.121655	13.93600

Source	DF	Type III SS	Mean Square	F Value	Pr > F
loom	4	0.34160000	0.08540000	5.77	0.0030

Level of loom	N	cloth	
		Mean	Std Dev
1	5	14.0800000	0.08366600
2	5	13.9200000	0.08366600
3	5	14.0600000	0.11401754
4	5	13.8000000	0.15811388
5	5	13.8200000	0.14832397

Levene's Test for Homogeneity of cloth Variance ANOVA of Absolute Deviations from Group Means					
Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
loom	4	0.0122	0.00304	0.67	0.6179
Error	20	0.0902	0.00451		

Brown and Forsythe's Test for Homogeneity of cloth Variance ANOVA of Absolute Deviations from Group Medians					
Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
loom	4	0.0136	0.00340	0.57	0.6897
Error	20	0.1200	0.00600		

- From the following analysis in SAS, the p-values for Levene's Test and the Brown-Forsythe are .6179 and .6897, respectively.
- We would **fail to reject** $\mathrm{H}_0 : \sigma_1^2 = \sigma_2^2 = \sigma_3^2 = \sigma_4^2 = \sigma_5^2$. Therefore, we conclude that we cannot reject the HOV assumption for the oneway ANOVA.
- And, assuming there are no serious violations of any other assumptions, we would reject
  H_0 : for the oneway ANOVA.

SAS Code for HOV Tests

```
DM 'LOG; CLEAR; OUT; CLEAR;';

ODS GRAPHICS ON;
* ODS PRINTER PDF file='C:\COURSES\ST541\HOVTEST.PDF';
OPTIONS NODATE NONUMBER;

*************************************************;
*** 5 Looms,    Response = Cloth Output,   n=5 ***;
*** Bartlett's, Brown-Forsythe, Levene's Tests ***;
*************************************************;

DATA in; INPUT loom cloth @@; CARDS;
1 14.0 1 14.1 1 14.2 1 14.0 1 14.1
2 13.9 2 13.8 2 13.9 2 14.0 2 14.0
3 14.1 3 14.2 3 14.1 3 14.0 3 13.9
4 13.6 4 13.8 4 14.0 4 13.9 4 13.7
5 13.8 5 13.6 5 13.9 5 13.8 5 14.0
;
PROC GLM DATA=in;
     CLASS loom;
     MODEL cloth = loom / ss3 ;
     MEANS loom / HOVTEST=BF;
     MEANS loom / HOVTEST=LEVENE(TYPE=ABS);
RUN;
```

Data Analysis Options When the HOV Assumption is Not Valid

- If we reject $H_0 : \sigma_1^2 = \sigma_2^2 = \cdots = \sigma_a^2$, then what options do we have to analyze the data? We will consider the following two options:
  1. Weighted least squares.
  2. Using a variance stabilizing transformation.

2.11 Weighted Least Squares

- Linear regression models (such as the models used in this course) can be fitted by the **weighted least squares (WLS)** method.
- With the WLS method, the squared deviation between the observed data value and the predicted value $(y_i - \widehat{y}_i)^2$ is multiplied by a weight w_i. This weight is proportional to the variance of y_i.
- For a one factor CRD, the WLS function is

$$W(\mu, \tau_1, \ldots, \tau_a) =$$

- To find the least squares normal equations, you simultaneously solve

$$\partial W / \partial \mu = 0 \quad \text{and} \quad \partial W / \partial \tau_i = 0 \quad \text{for} \quad i = 1, 2, \ldots, a.$$

- The WLS normal equations are:

$$\sum_{i=1}^{a} \sum_{j=1}^{n_i} w_{ij} y_{ij} = \widehat{\mu} \sum_{i=1}^{a} \sum_{j=1}^{n_i} w_{ij} + \sum_{i=1}^{a} \left(\widehat{\tau}_i \sum_{j=1}^{n_i} w_{ij} \right)$$

$$\sum_{j=1}^{n_i} w_{ij} y_{ij} = \widehat{\mu} \sum_{j=1}^{n_i} w_{ij} + \widehat{\tau}_i \sum_{j=1}^{n_i} w_{ij} \quad \text{for } i = 1, 2, \ldots, a$$

- The solution to these $(a+1)$ equations subject to one constraint (such as $\sum_{i=1}^{a} \tau_i = 0$) are the WLS estimates for $\widehat{\mu}$ and $\widehat{\tau}_i$ for $i = 1, 2, \ldots, a$.
- However, because the variances $\sigma_1^2, \sigma_2^2, \ldots, \sigma_a^2$ are unknown, we need to estimate the weight $w_{ij} = 1/\sigma_i^2$ for $j = 1, 2, \ldots, n_i$ from the data.
- For the one-factor CRD, we know the sample variance s_i^2 for treatment i is an unbiased estimate of σ_i^2 $(E(s_i^2) = \sigma_i^2)$. The estimated weight is $\widehat{w}_{ij} = 1/s_i^2$ for $j = 1, 2, \ldots, n_i$.
- SAS and Minitab will perform a WLS analysis, but you have to supply the weights.

Weighted Least Squares (WLS) Example: A company wants to test the effectiveness of a new chemical disinfectant. Five dosage levels were considered (1 through 5 grams per 100 ml). The experiment involved applying equal amounts of the disinfectant at each level to a surface that was covered with a common bacteria. The results are given below. The design was completely randomized.

Dose	%	Dose	%	Dose	%	Dose	%	Dose	%
1	5	2	13	3	12	4	17	5	22
1	1	2	13	3	16	4	13	5	30
1	3	2	6	3	9	4	16	5	27
1	5	2	7	3	18	4	19	5	32
1	2	2	11	3	16	4	26	5	32
1	6	2	4	3	7	4	15	5	43
1	1	2	14	3	14	4	23	5	29
1	3	2	12	3	13	4	27	5	26

The sample variances s_i^2 are

$s_1^2 =$ $\qquad$ $s_2^2 =$ $\qquad$ $s_3^2 =$ $\qquad$ $s_4^2 =$ $\qquad$ $s_5^2 =$

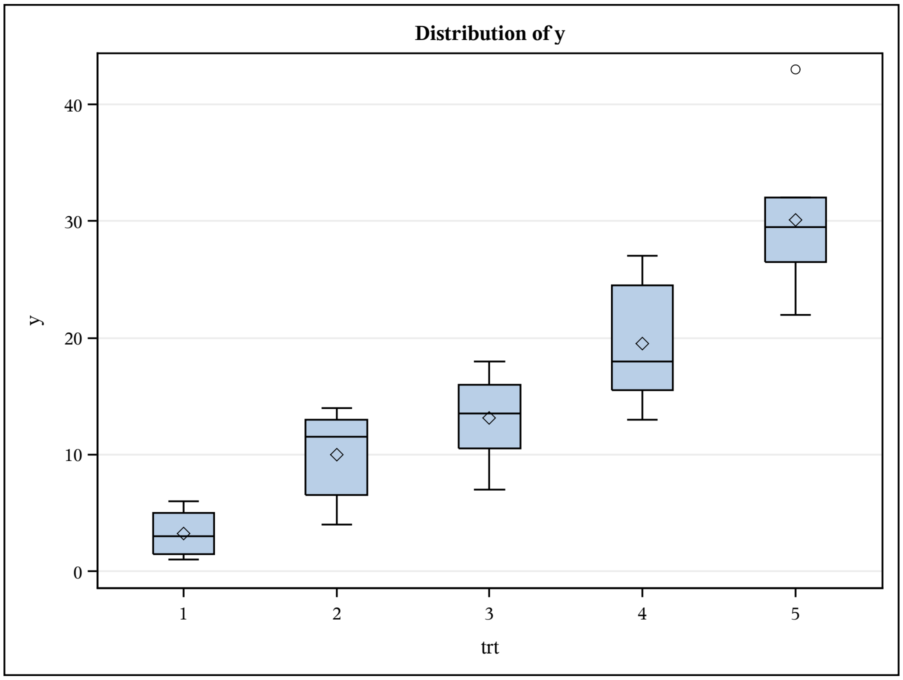
Thus, the weights $w_{ij} = 1/s_i^2$ for $j = 1, 2, \ldots, n_i$ are

$w_{1j} =$ $\qquad$ $w_{2j} =$ $\qquad$ $w_{3j} =$ $\qquad$ $w_{4j} =$ $\qquad$ $w_{5j} =$

SAS Output for WLS Example

SAMPLE VARIANCES AND WEIGHTS FOR EACH TREATMENT trt

Obs	trt	var_y	wgt
1	1	3.6429	0.27451
2	2	14.2857	0.07000
3	3	13.8393	0.07226
4	4	27.4286	0.03646
5	5	38.1250	0.02623



Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	4	207.5551273	51.8887818	51.89	<.0001
Error	35	35.0000000	1.0000000		
Corrected Total	39	242.5551273			

R-Square	Coeff Var	Root MSE	y Mean
0.855703	11.86288	1.000000	8.429653

Source	DF	Type III SS	Mean Square	F Value	Pr > F
trt	4	207.5551273	51.8887818	51.89	<.0001

Bonferroni (Dunn) t Tests for y

Note: This test controls the Type I experimentwise error rate, but it generally has a higher Type II error rate than Tukey's for all pairwise comparisons.

REJECT ALL

$H_0: \mu_i = \mu_j \quad i,j = 1,2,3,4,5$

EXCEPT FOR

$H_0: \mu_3 = \mu_4$
$H_0: \mu_2 = \mu_3$ } FAIL TO REJECT (FTR)

Alpha	0.05
Error Degrees of Freedom	35
Error Mean Square	1
Critical Value of t	2.99605

Comparisons significant at the 0.05 level are indicated by ***.

trt Comparison	Difference Between Means	Simultaneous 95% Confidence Limits		
5 - 4	10.6250	2.0487	19.2013	***
5 - 3	17.0000	9.3642	24.6358	***
5 - 2	20.1250	12.4564	27.7936	***
5 - 1	26.8750	20.0292	33.7208	***
4 - 5	-10.6250	-19.2013	-2.0487	***
4 - 3	6.3750	-0.4297	13.1797	
4 - 2	9.5000	2.6586	16.3414	***
4 - 1	16.2500	10.3455	22.1545	***
3 - 5	-17.0000	-24.6358	-9.3642	***
3 - 4	-6.3750	-13.1797	0.4297	
3 - 2	3.1250	-2.4926	8.7426	
3 - 1	9.8750	5.4460	14.3040	***
2 - 5	-20.1250	-27.7936	-12.4564	***
2 - 4	-9.5000	-16.3414	-2.6586	***
2 - 3	-3.1250	-8.7426	2.4926	
2 - 1	6.7500	2.2649	11.2351	***
1 - 5	-26.8750	-33.7208	-20.0292	***
1 - 4	-16.2500	-22.1545	-10.3455	***
1 - 3	-9.8750	-14.3040	-5.4460	***
1 - 2	-6.7500	-11.2351	-2.2649	***

FTR $H_0: \mu_3 = \mu_4$

FTR $H_0: \mu_2 = \mu_3$

```
DATA in; INPUT trt y @@;  LINES;
1  5    1  1    1  3    1  5    1  2    1  6    1  1    1  3
2 13    2 13    2  6    2  7    2 11    2  4    2 14    2 12
3 12    3 16    3  9    3 18    3 16    3  7    3 14    3 13
4 17    4 13    4 16    4 19    4 26    4 15    4 23    4 27
5 22    5 30    5 27    5 32    5 32    5 43    5 29    5 26
;
PROC SORT  DATA=in; BY trt;             <-- Sort the data by treatments.
PROC MEANS DATA=in noprint; BY trt;     <-- Calculate and save sample
     VAR y;                             <-- variances in 'wset'.
     OUTPUT OUT=wset VAR=var_y;

DATA wset; SET wset;                    <-- Calculate the weights from
     wgt = 1/var_y;                     <-- the sample variance in wset.
DROP _FREQ_ _TYPE_;

PROC PRINT DATA=wset;
TITLE 'SAMPLE VARIANCES AND WEIGHTS FOR EACH TREATMENT trt';

DATA in; MERGE in wset; BY trt;         <-- Attach the weights by treatment.

PROC GLM DATA=in;
    WEIGHT wgt;                         <-- Include the WEIGHT statement.
    CLASS trt;
    MODEL y = trt / SS3;
     MEANS trt / BON;
TITLE 'WEIGHTED LEAST SQUARES EXAMPLE WITH BONFERRONI MCP';
RUN;
```

$\rightarrow w_{ij} = \frac{1}{s_i^2}$

] OUTPUT (*) TABLE ON PAGE 51

2.12 Variance Stabilizing Transformations

- If the homogeneity of variance assumption is only moderately violated, the F-test results are only slightly affected when the design is balanced (equal n_i's). No transformation should be considered.

- If the homogeneity of variance assumption is either (i) seriously violated or (ii) moderately violated with very different n_i sample sizes (serious imbalance), then the effects on the F-test are more serious.

  - If the treatments having the larger variances have the smaller sample sizes, the true Type I error is larger than the nominal level.
  - If the treatments having the larger variances have the larger sample sizes, the true Type I error is smaller than the nominal level.

- A common approach to deal with nonconstant variance (heterogeneity of variance) is to apply a of the response that will equalize the variances across treatments. We then perform the ANOVA on the transformed data.

- Sometimes the variance of the response increases or decreases as the mean of the response increases. If this is the case, the residuals vs predicted values plot would have a funnel shape. This is when a variance stabilizing transformation may be appropriate.

- The statistical problem is to use the data to determine the form of the required transformation.

- Let μ_i be the mean for treatment i. Suppose the standard deviation of y_{ij} is proportional to a power of μ_i. That is, $\sigma_i = \theta \mu_i^\alpha$ for some α and θ. θ is called the .

- The goal is to find a transformation $y^* = y^\lambda$ such that y^* has constant or near constant variance across all treatments.

- This implies that after transforming each y_{ij} to y_{ij}^*, we no longer have a HOV problem when the ANOVA is run with the y_{ij}^* values.

- It can be shown that the variance is constant if $\lambda = 1 - \alpha$. If $\lambda \approx 0$, then we use a log transformation. We will now discuss two methods for estimating λ.

2.12.1 The Empirical Method

- If $\sigma_i = \theta \mu_i^\alpha$, then $\log(\sigma_i) = \log(\theta) + \alpha \log(\mu_i)$. A plot of $\log(\sigma_i)$ vs $\log(\mu_i)$ is linear with slope equal to α. Thus, a simple way to estimate α would be to

  1. Calculate s_i and $\overline{y}_{i\cdot}$ for treatment $i = 1, 2, \ldots, a$.
  2. Fit a regression line $\log(s_i) =$
  3. The least squares estimate of the slope $\widehat{\alpha}$ is the estimate of α.
  4. Transform each y_{ij} to $y_{ij}^* = y_{ij}^\lambda$ where $\lambda = 1 - \widehat{\alpha}$.
  5. Run the ANOVA on the y_{ij}^* values.

- Note that if $\alpha = 0$, then $\sigma_i = \theta$ for all i. Thus, the homogeneity of variance assumption is met without a transformation.

2.12.2 The Box-Cox Procedure

- Another approach is the **Box-Cox procedure** which will estimate the value of λ corresponding to the transformation y_{ij}^λ that maximizes the model R^2.

- SAS can find the Box-Cox transformation using the TRANSREG procedure.

- To find the Box-Cox transformation,
  1. For a sequence of λ values, calculate $R^2(\lambda)$. $R^2(\lambda)$ is the model R^2 value from the ANOVA on the transformed y^λ values.
  2. Select the λ that maximizes $R^2(\lambda)$ (which is equivalent to maximizing the likelihood function).
  3. Run the ANOVA on the y_{ij}^λ values.

2.12.3 Transformation Example using the Empirical and Box-Cox Methods

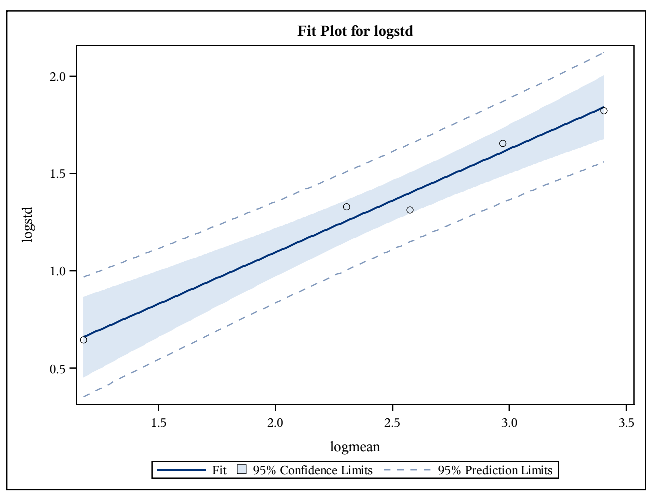
EXAMPLE: We will use the same data used in the WLS example:

Dose	%	Dose	%	Dose	%	Dose	%	Dose	%
1	5	2	13	3	12	4	17	5	22
1	1	2	13	3	16	4	13	5	30
1	3	2	6	3	9	4	16	5	27
1	5	2	7	3	18	4	19	5	32
1	2	2	11	3	16	4	26	5	32
1	6	2	4	3	7	4	15	5	43
1	1	2	14	3	14	4	23	5	29
1	3	2	12	3	13	4	27	5	26

- We will see that the recommended transformation is a square root ($\lambda = .5$) transformation. The following SAS output contains
  - The empirical method results and the Box-Cox method results.
  - The analysis of the original data. Note that the variability increases with the Dose treatment levels (from 1 to 5).
  - The analysis of the transformed (square root) data. Note that the variability is now nearly constant across the Dose treatment levels (from 1 to 5).

EMPIRICAL SELECTION OF ALPHA

Obs	mean	std	logstd	logmean
1	3.250	1.90863	0.64638	1.17865
2	10.000	3.77964	1.32963	2.30259
3	13.125	3.72012	1.31376	2.57452
4	19.500	5.23723	1.65579	2.97041
5	30.125	6.17454	1.82044	3.40536



ANOVA TO FIND EMPIRICAL SELECTION OF ALPHA

The GLM Procedure

Variable: logstd

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	1	0.79599334	0.79599334	153.30	0.0011
Error	3	0.01557765	0.00519255		
Corrected Total	4	0.81157099			

GOOD FIT OF THE LINEAR REGRESSION

R-Square	Coeff Var	Root MSE	logstd Mean
0.980806	5.325109	0.072059	1.353200

$r^2 = .98$

Source	DF	Type III SS	Mean Square	F Value	Pr > F
logmean	1	0.79599334	0.79599334	153.30	0.0011

$\hat{\beta}_0$ $\hat{\beta}_1$

Parameter	Estimate	Standard Error	t Value	Pr > |t|
Intercept	0.0347067133	0.11126036	0.31	0.7755
logmean	0.5303019549	0.04283106	12.38	0.0011

$\hat{\beta}_0 = \log \hat{\theta} = .0347$

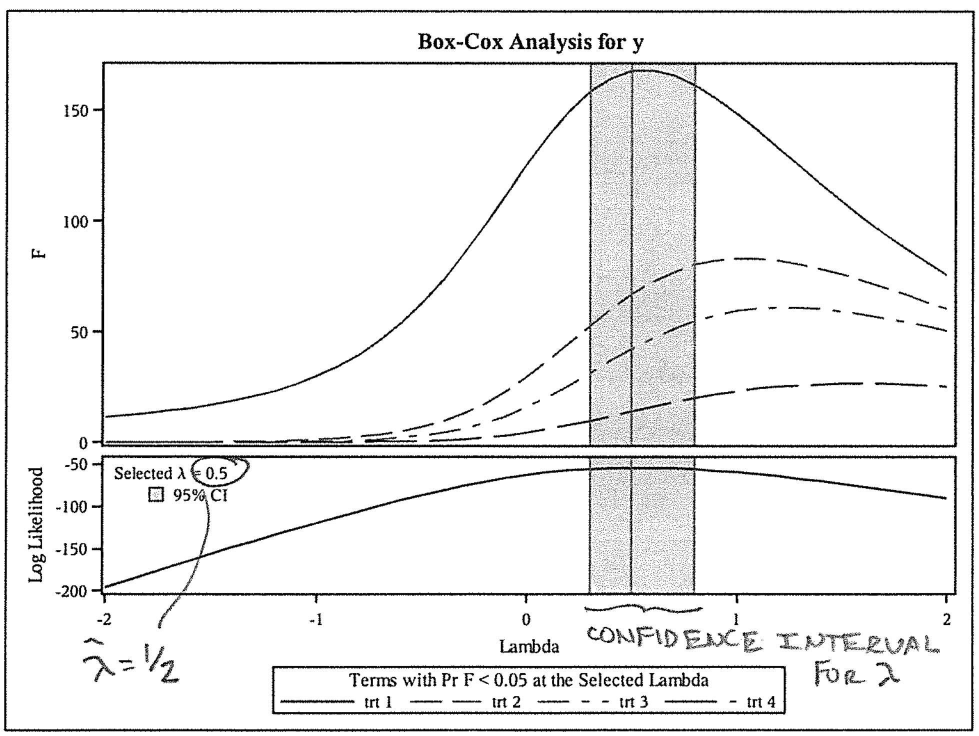
$\hat{\beta}_1 = \hat{\alpha} = .530$

$\Rightarrow \hat{\lambda} = 1 - \hat{\alpha} = .470$

Find the Box-Cox Transformation using PROC TRANSREG

The TRANSREG Procedure

$\Rightarrow$ SQUARE ROOT TRANSFORMATION IS REASONABLE ($\hat{\lambda} = .50$) USING EMPIRICAL METHOD



$\hat{\lambda} = \frac{1}{2}$ FOR THE Box-Cox TRANSFORMATION

BEFORE

Level of trt	N	y Mean	y Std Dev
1	8	3.2500000	1.90862703
2	8	10.0000000	3.77964473
3	8	13.1250000	3.72011905
4	8	19.5000000	5.23722937
5	8	30.1250000	6.17454452

(A)

AFTER

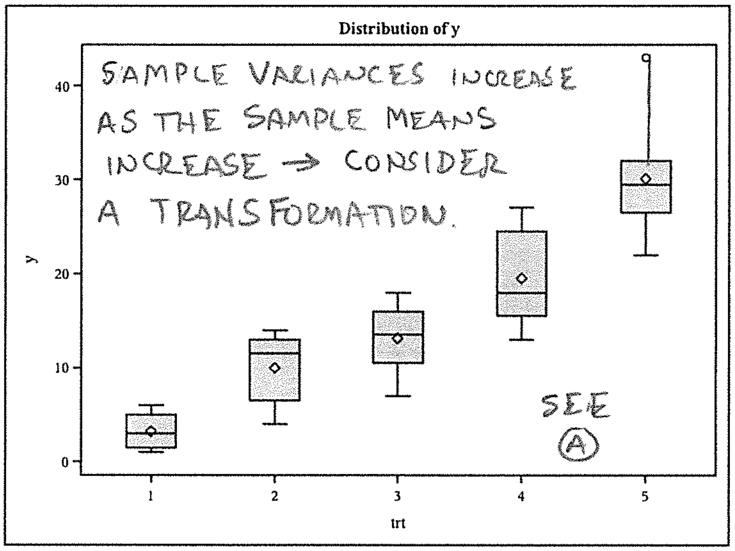
Level of trt	N	sqrty Mean	sqrty Std Dev
1	8	1.72499261	0.56000051
2	8	3.10359092	0.64827099
3	8	3.58746278	0.53995849
4	8	4.38144283	0.58842037
5	8	5.46489064	0.54507700

(B)

SIMILAR VARIABILITY AFTER TAKING A SQUARE ROOT TRANSFORMATION

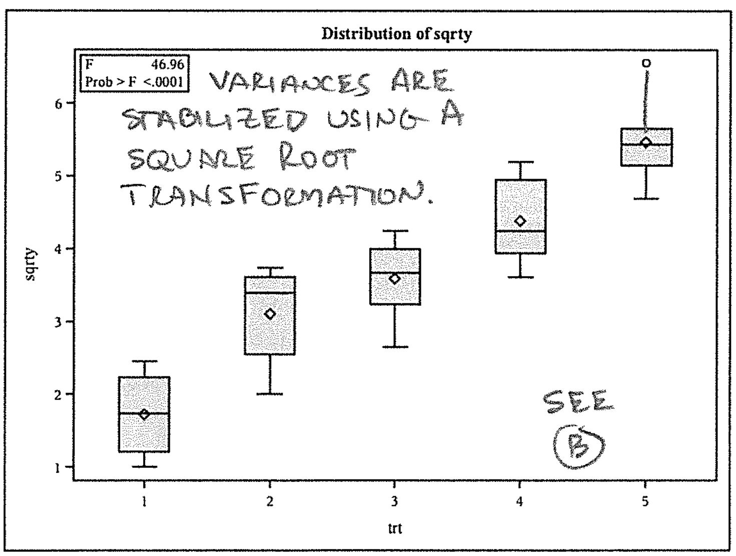
Boxplots of the Original Data (y)

PLOT 1



Boxplots of the Square Root Transformed Data $(\sqrt{y})$

PLOT 2



HOV ASSUMPTION IS NOT VIOLATED USING $\sqrt{y}$ AS THE RESPONSE.

This page contains the SAS output using the original response values. Note the funneling pattern in the *Residuals vs Predicted Value* plot. This suggests that a transformation should be considered.

ANOVA -- ORIGINAL DATA

The GLM Procedure

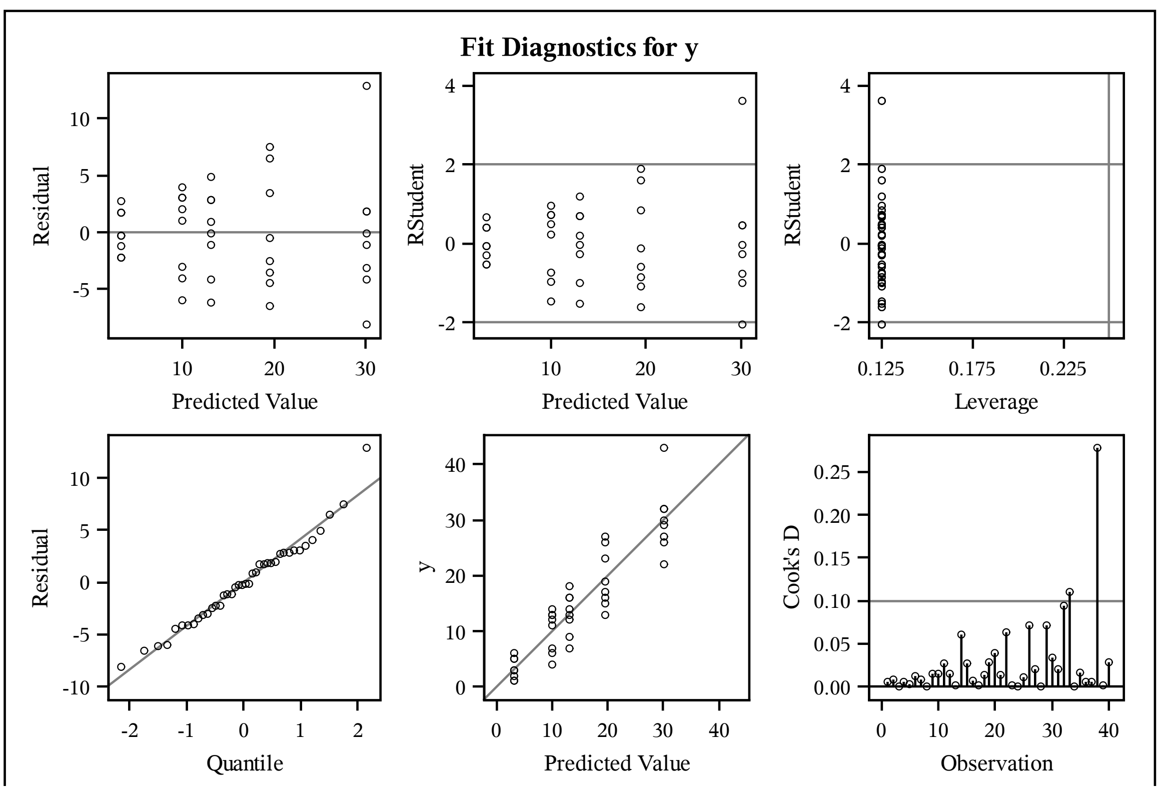
Variable: y

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	4	3323.150000	830.787500	42.68	<.0001
Error	35	681.250000	19.464286		
Corrected Total	39	4004.400000			

R-Square	Coeff Var	Root MSE	y Mean
0.829875	29.02523	4.411835	15.20000

Source	DF	Type III SS	Mean Square	F Value	Pr > F
trt	4	3323.150000	830.787500	42.68	<.0001

Dependent Variable: y



This page contains the SAS output using the square root transformation of the response values. Note that the variances are stabilized (nearly constant) in the *Residuals vs Predicted Value* plot.

ANOVA -- SQUARE ROOT TRANSFORMATION

The GLM Procedure

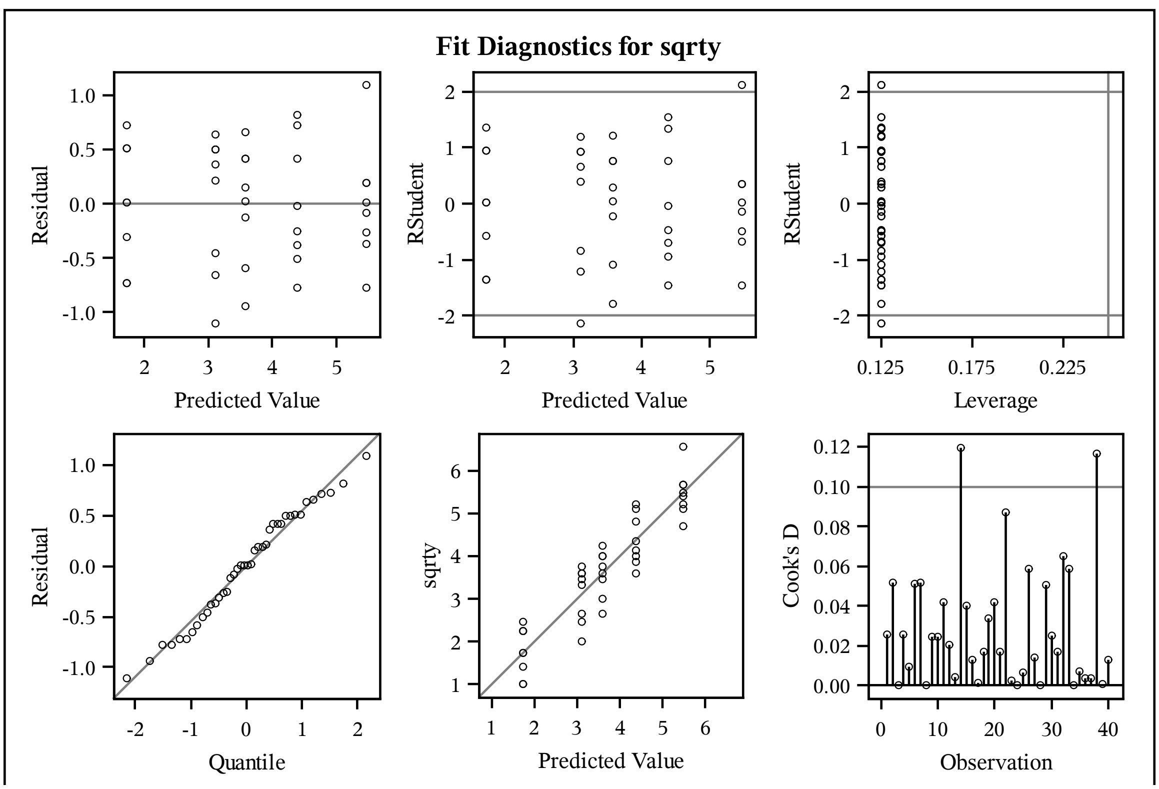
Variable: sqrty

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	4	62.69546580	15.67386645	46.96	<.0001
Error	35	11.68130939	0.33375170		
Corrected Total	39	74.37677520			

R-Square	Coeff Var	Root MSE	sqrty Mean
0.842944	15.81701	0.577712	3.652476

Source	DF	Type III SS	Mean Square	F Value	Pr > F
trt	4	62.69546580	15.67386645	46.96	<.0001

Dependent Variable: sqrty



```
DM 'LOG; CLEAR; OUT; CLEAR;';

ODS GRAPHICS ON;
ODS PRINTER PDF file='C:\COURSES\ST541\BOXCOX.PDF';
OPTIONS NODATE NONUMBER;

********************************************;
*** Variance Stabilizing Transformations ***;
********************************************;
DATA in; INPUT trt y @@;  CARDS;
1  5   1  1   1  3   1  5   1  2   1  6   1  1   1  3
2 13   2 13   2  6   2  7   2 11   2  4   2 14   2 12
3 12   3 16   3  9   3 18   3 16   3  7   3 14   3 13
4 17   4 13   4 16   4 19   4 26   4 15   4 23   4 27
5 22   5 30   5 27   5 32   5 32   5 43   5 29   5 26

**********************************************************;
*** Find the transformation using the empirical method ***;
**********************************************************;
```

REGRESSION OF $\log(s_i)$ vs $\log(\bar{y}_{i.})$

```
PROC SORT DATA=in; BY trt;
PROC MEANS DATA=in NOPRINT; BY trt;
     VAR y; OUTPUT OUT=yset MEAN=mean STD=std;
DATA yset; SET yset;
     logstd =LOG(std);   logmean=LOG(mean);
PROC PRINT DATA=yset;
     VAR mean std logstd logmean;
TITLE 'EMPIRICAL SELECTION OF ALPHA';

PROC GLM DATA=yset;
     MODEL logstd=logmean / SS3 solution;
TITLE 'ANOVA TO FIND EMPIRICAL SELECTION OF ALPHA';
```

↳ ln

EMPIRICAL METHOD TO FIND $\hat{\lambda}$

```
*** Use the output of the GLM procedure regressing the  ***;
*** log standard deviations on the log means. Apply the ***;
*** Apply the transformation to the response and        ***;
*** rerun the analysis with the transformed response.   ***;

********************************************************;
*** Find the transformation using the Box-Cox method ***;
********************************************************;

PROC TRANSREG DATA=in;
     MODEL BOXCOX(y / LAMBDA=-2 to 2 by .1) = CLASS(trt);
TITLE 'Find the Box-Cox Transformation using PROC TRANSREG';

*** Use the output of the TRANSREG procedure to find the ***;
*** the Box-Cox transformation. Apply the transformation ***;
*** to the response and rerun the analysis with the      ***;
*** transformed response                                 ***;
```

41 POSSIBLE $\hat{\lambda}$ VALUES ARE CHECKED $(-2, -1.9, \ldots, 1.9, 2.0)$

$\Rightarrow$ 41 ANOVAS ARE PERFORMED TO FIND THE LARGEST r^2 VALUE. (SEE PAGE 55)

→ BOX-COX METHOD

```
*************************************;
*** ANOVA BEFORE A TRANSFORMATION ***;
*************************************;
PROC GLM DATA=in PLOTS=(DIAGNOSTICS);
     CLASS trt;
     MODEL y= trt / SS3;
     MEANS trt ;
TITLE 'ANOVA -- ORIGINAL DATA';
```

ANOVA USING y (NO TRANSFORMATION)

```
************************************;
*** ANOVA AFTER A TRANSFORMATION ***;
************************************;

DATA in; SET in;;  sqrty = SQRT(y);

PROC GLM DATA=in PLOTS=(DIAGNOSTICS);
     CLASS trt;
     MODEL sqrty= trt / SS3;
     MEANS trt;
TITLE 'ANOVA -- SQUARE ROOT TRANSFORMATION';
RUN;
```

ANOVA USING $\sqrt{y}$ TRANSFORMATION

2.13 Simulations to Study the ANOVA HOV Assumption

- The file *simanova.r* posted on the course webpage contains R code that will do the following assessing the impact of equal ($\sigma_1 = \sigma_2 = \sigma_3$) and unequal (not all σ_i are equal) on the probability of rejecting $H_0 : \mu_1 = \mu_2 = \mu_3$.
- Various cases can be studied by varying the following:
  - The values of σ_1, σ_2, and σ_3.
  - The values of μ_1, μ_2, and μ_3.
  - The values of n_1, n_2, and n_3.
- The program will output the estimated probability of rejecting $H_0 : \mu_1 = \mu_2 = \mu_3$ assuming α levels of .01. .05, and .10 from a oneway ANOVA. That is, the proportion of samples that lead to a rejection of H_0 using $\alpha = .01..05, .10$.
- If $H_0 : \mu_1 = \mu_2 = \mu_3$ is **<u>true</u>**, these estimated probabilities represent estimates of the **Type I error** (i.e. the probability of incorrectly rejecting H_0 when H_0 it is true). Thus, the values should be close to the nominal (stated) α levels of .01, .05, and .10.
- If $H_0 : \mu_1 = \mu_2 = \mu_3$ is **<u>false</u>**, these estimated probabilities represent estimates of the **power of the F-test**.

  The **power** of a test equals the probability of correctly rejecting H_0 when H_0 is false

  $= 1 -$ **the probability of not rejecting H_0 when H_0 is false**

  $= 1 -$ **Type II error**.

Case	n_1	n_2	n_3	σ_1	σ_2	σ_3	μ_1	μ_2	μ_3	H_0
1	9	9	9	1	1	1	10	10	10	True
2	9	9	9	1	2	3	10	10	10	True
3	9	6	3	1	2	3	10	10	10	True
4	3	6	9	1	2	3	10	10	10	True
5	9	9	9	1	1	1	10	10	11	False
6	3	3	3	1	1	1	10	10	11	False
7	20	20	20	1	1	1	10	10	11	False

Case	H_0	HOV	Sample sizes	Treatment means	Rejection Probability for $\alpha = .01$	$\alpha = .05$	$\alpha = .10$
1	True	OK	$=$				
2	True	Violated	$=$				
3	True	Violated	$\neq$				
4	True	Violated	$\neq$				
5	False	OK	$=$				
6	False	OK	$=$				
7	False	OK	$=$				

R Code to perform the simulations

```
##################################################################
#  ASSESSING THE IMPACT OF EQUAL vs UNEQUAL STANDARD DEVIATIONS
##################################################################

  a    <- 3    # Enter number of treatments
  sd.1 <- 1    # Enter sigma_1  1 1 1
  sd.2 <- 1    # Enter sigma_2  1 2 2
  sd.3 <- 1    # Enter sigma_3  1 3 3
```

ENTER $\sigma_1, \sigma_2, \sigma_3$

```
  n.1 <- 9     # Enter n_1      9 9 9
  n.2 <- 9     # Enter n_2      9 9 6
  n.3 <- 9     # Enter n_3      9 9 3
```

ENTER n_1, n_2, n_3

```
  mu.1 <- 10   # Enter mu_1     10 10 10
  mu.2 <- 10   # Enter mu_2     10 10 10
  mu.3 <- 10   # Enter mu_3     10 10 10
```

ENTER μ_1, μ_2, μ_3

```
  iter <- 10000   # Enter the number of t-statistics to simulate
```

PERFORM 10,000 ANOVAS USING RANDOMLY GENERATED SAMPLES OF SIZES n_1, n_2, n_3 FROM $N(\mu_1,\sigma_1^2)$, $N(\mu_2,\sigma_2^2)$, $N(\mu_3,\sigma_3^2)$ DISTRIBUTIONS, RESPECTIVELY.

```
## Simulate F-statistics
  N <- n.1 + n.2 + n.3
  df.MSE   <- N - a
  df.MSE
  df.MStrt <- a -1
  df.MStrt

  F.stat <- numeric(iter)

  for (i in 1:iter) {
     sample1 <- rnorm(n.1, mean=mu.1, sd=sd.1)
     sample2 <- rnorm(n.2, mean=mu.2, sd=sd.2)
     sample3 <- rnorm(n.3, mean=mu.3, sd=sd.3)

     var.1 <- var(sample1)
     var.2 <- var(sample2)
     var.3 <- var(sample3)

     SS.E  <- (n.1-1)*var.1 + (n.2-1)*var.2 + (n.3-1)*var.3
     MS.E  <- SS.E/df.MSE

     all.dat  <- c(sample1,sample2,sample3)
     all.dat
     SS.TOTAL <- (N-1)*var(all.dat)

     SS.TRT   <- SS.TOTAL - SS.E
     MS.TRT   <- SS.TRT/df.MStrt

     F.stat[i] <- MS.TRT/MS.E
     }
windows()
Fmax = max(F.stat)

hist(F.stat, freq=FALSE, nclass=50, xlim=c(-.01,Fmax), ylim=c(0,.8),
     main="Histogram of F-statistics with superimposed F pdf")
curve(df(x,df.MStrt,df.MSE), add=TRUE, col=2, lwd=2)

  F.01 <- qf(.99,df.MStrt,df.MSE)
  F.05 <- qf(.95,df.MStrt,df.MSE)
  F.10 <- qf(.90,df.MStrt,df.MSE)
  F.01
  F.05
  F.10

# Simulated rejection probabilities for alpha = .01, .05, .10.
# If Ho: mu.1 = mu.2 = mu.3 is true, this is the estimated Type I error.
# If Ho is not true, this is the estimated Power = 1 - Type II error.

  reject.01  <- ifelse(F.stat >= F.01,1,0)
  pvalue.01  <- sum(reject.01)/iter
  pvalue.01

  reject.05  <- ifelse(F.stat >= F.05,1,0)
  pvalue.05  <- sum(reject.05)/iter
  pvalue.05

  reject.10  <- ifelse(F.stat >= F.10,1,0)
  pvalue.10  <- sum(reject.10)/iter
  pvalue.10
```

OUTPUT THE PROPORTION OF CASES THAT REJECT $H_0: \mu_1 = \mu_2 = \mu_3$ USING $\alpha = .01, .05, .10$.

2.14 Power and Design Size

FIND n, GIVEN POWER

- The GLMPOWER procedure in SAS can be used to
  - Determine the number of replicates needed to achieve a specified power for ANOVA F-tests.
  - Determine the power for ANOVA F-tests given a specified number of replicates. This can be used after an experiment is run to estimate the actual power for tests related to that experiment.

FIND POWER, GIVEN n

- The experimenter must specify
  - An estimate of σ.
  - Information about the mean responses (such as preliminary estimates or a minimum difference the researcher would like to detect as significant).
  - The desired power for tests if the goal is to determine the number of replicates.
  - The number of replicates if the goal is to determine an estimated power for tests.
- In the SAS output Ntotal is the total number of experimental runs. This must be divided by a to determine the number of replicates.

2.14.1 Example 1: Sample size determination and power estimation

Determine N given a nominal power level (Case 1) and determine power given N (Case 2) for a specified pattern of means or effects

- Suppose there are 5 (a) treatments, and that one mean (say μ_1) is 2 units larger than the other four means which are equal. That is:

$$\mu_2 = \mu_3 = \mu_4 = \mu_5 \quad \text{and} \quad \mu_1 = \mu_i + 2 \text{ for } i = 2,3,4,5$$

E.G. MEANS = 12 10 10 10 10

- Or, equivalently in terms of effects

$$\tau_2 = \tau_3 = \tau_4 = \tau_5 \quad \text{and} \quad \tau_1 = \tau_i + 2 \text{ for } i = 2,3,4,5$$

E.G. EFFECTS = 2 0 0 0 0

- Our prior estimate of σ is 1.25 ($\hat{\sigma}$), and the significance level is set to $\alpha = .05$ for tests.
- For <u>Case 1</u>, determine the total sample size $N = 5n$ setting the power for the ANOVA F-test at power levels $1 - \beta = .50, .80, .85, .90, .95,$ and $.99$.
- For <u>Case 2</u>, determine the power $1 - \beta$ for the ANOVA F-test when the total sample size $N =$ 10, 15, 20, 25, 30, 35, 40, 45, 50, 55, and 60.

SAS code for Case 1: Determine N for a nominal power level

```
DATA oneway;
     DO level = 1 to 5;   INPUT delta @@;  OUTPUT;   END;
LINES;
2 0 0 0 0   -> EFFECT PATTERN FOR (T1, T2, T3, T4, T5) OR FOUR MEANS ARE
;                EQUAL (M2 = M3 = M4 = M5) AND 1 MEAN IS 2 UNITS LARGER (M1)
PROC GLMPOWER DATA=oneway;
     CLASS level;
     MODEL delta = level;
     POWER
         STDDEV = 1.25   (σ̂)              <-- sigma estimate
         ALPHA  = 0.05   (α)              <-- alpha level
         NTOTAL = .                       <-- determine N
         POWER  = .50 .80 .85 .90 .95 .99;    <-- choices for power

TITLE  'Case 1: Determining sample size for detecting a 2-unit difference';
TITLE2 'for 4 equal means and 1 unequal mean (power=.50 .80 .85 .90 .95 .99)';
RUN;
```

SAS output for Case 1: Determine N for a nominal power level

```
Case 1: Determining sample size for detecting a 2-unit difference
for 4 equal means and 1 unequal mean (with power=.50 .80 .85 .90 .95 .99)

                    The GLMPOWER Procedure

          Dependent Variable                    delta
          Alpha                                  0.05
          Error Standard Deviation               1.25
```

α $\hat{\sigma}$

Computed N Total

Index	Type	Source	Nominal Power	Test DF ($a-1$)	Error DF ($a(n-1)$)	Actual Power	N Total ($an = 5n$)	$n = N/a$
1	Effect	level	0.50	4	20	0.612	25	5
2	Effect	level	0.80	4	30	0.812	35	7
3	Effect	level	0.85	4	35	0.875	40	8
4	Effect	level	0.90	4	40	0.919	45	9
5	Effect	level	0.95	4	50	0.968	55	11
6	Effect	level	0.99	4	65	0.993	70	14

FIND SMALLEST N (or n) NEEDED TO EXCEED THE NOMINAL POWER (COMPARE NOMINAL POWER TO ACTUAL POWER.)

SAS code for Case 2: Determine power for a given N

```
DATA oneway2;
     DO level = 1 to 5;   INPUT delta @@;  OUTPUT;   END;
LINES;
2 0 0 0 0
;
PROC GLMPOWER DATA=oneway2;
     CLASS level;
     MODEL delta = level;
     POWER
         STDDEV = 1.25
         ALPHA  = 0.05
         NTOTAL = 10 15 20 25 30 35 40 45 50 55 60    <-- choices for N
         POWER  = .;                                  <-- determine power

TITLE  'Case 2: Determining power for detecting a 2-unit difference';
TITLE2 'for 4 equal means and 1 unequal mean
TITLE3 '(for total sample size N=10 15 ... 55 60)';
RUN;
```

$N = 5n$

SPECIFY N

OUTPUT POWER

SAS output for Case 2: Determine power for given N

```
  Case 2: Determining power for detecting a 2-unit difference
  for 4 equal means and 1 unequal mean
  (for total sample size N=10 15 ... 55 60)

                    The GLMPOWER Procedure

          Dependent Variable                     delta
          Alpha                                   0.05
          Error Standard Deviation                1.25
```

SUPPOSE YOU WANT POWER $\geq .90$.

Computed Power

Index	Type	Source	n	N Total ($5n$)	Test DF	Error DF	Power
1	Effect	level	2	10	4	5	0.162
2	Effect	level	3	15	4	10	0.318
3	Effect	level	4	20	4	15	0.474
4	Effect	level	5	25	4	20	0.612
5	Effect	level	6	30	4	25	0.725
6	Effect	level	7	35	4	30	0.812
7	Effect	level	8	40	4	35	0.875
8	Effect	level	9	45	4	40	0.919
9	Effect	level	10	50	4	45	0.948
10	Effect	level	11	55	4	50	0.968
11	Effect	level	12	60	4	55	0.980

$< .90$

FIRST CASE WITH POWER $> .90$

$\Rightarrow n = 9$ REPLICATES PER TREATMENT

2.14.2 Example 2: Sample size determination and power estimation using estimated means

Determine total sample size N given a nominal power level (Case 3) and determine power given N (Case 4) assuming $\mu_1 = 35.6,\ \mu_2 = 33.7,\ \mu_3 = 30.2,\ \mu_4 = 28.0,$ and $\mu_5 = 25.9.$

- Suppose there are 5 treatments, and using data from a previous study, we have estimates of the treatment means: $\widehat{\mu}_1 = 35.6,\ \widehat{\mu}_2 = 33.7,\ \widehat{\mu}_3 = 30.2,\ \widehat{\mu}_4 = 28.0,$ and $\widehat{\mu}_5 = 25.9.$
- Our prior estimate of σ is 3.75, and the significance level is set to $\alpha = .05$ for tests.
- For Case 3, determine the total sample size $N = 5n$ setting the power for the ANOVA F-test at levels $1 - \beta = .80,\ .90,\ .95,$ and $.99$.
- For Case 4, determine the power $1 - \beta$ for the ANOVA F-test when the total sample size $N = 10,\ 15,\ 20,\ 25,\ 30,\ 35,\ 40,\ 45,$ and 50.

SAS Code for Case 3: Find N given a nominal power level and specified treatment means.

```
DATA oneway3;
     INPUT level $ meanest @@;
LINES;
A1 35.6     A2 33.7     A3 30.2     A4 28.0     A5 25.9
;  μ̂1          μ̂2          μ̂3          μ̂4          μ̂5
PROC GLMPOWER DATA=oneway3;
      CLASS level;
      MODEL meanest = level;
      POWER
         STDDEV = 3.75  σ̂
         ALPHA  = 0.05  α        OUTPUT N
         NTOTAL = .
         POWER  = .80 .90 .95 .99;  NOMINAL POWER
TITLE  'Case 3: Determining sample size for given power and specified means';
TITLE2 'for a oneway ANOVA with 5 treatments (power .80 .90 .95 .99)';
RUN;
```

SAS output for Case 3: Determine N for given a nominal power level assuming

$\mu_1 = 35.6,\ \mu_2 = 33.7,\ \mu_3 = 30.2,\ \mu_4 = 28.0,\ \mu_5 = 25.9$

```
Case 3: Determining sample size for given power and specified means
      for a oneway ANOVA with 5 treatments (power .80 .90 .95 .99)

                     The GLMPOWER Procedure
                    Fixed Scenario Elements

            Dependent Variable                 meanest   α
            Alpha                                 0.05
            Error Standard Deviation              3.75   σ̂

                        Computed N Total
```

Index	Type	Source	Nominal Power	Test DF	Error DF	Actual Power	N Total
1	Effect	level	0.80	4	15	0.844	20
2	Effect	level	0.90	4	20	0.941	25
3	Effect	level	0.95	4	25	0.980	30
4	Effect	level	0.99	4	30	0.994	35

COMPARE NOMINAL POWER
TO ACTUAL POWER.

SAS Code for Case 4: Determine power given N and specified treatment means.

```
DATA oneway4;
     INPUT level $ meanest @@;
LINES;
A1 35.6    A2 33.7    A3 30.2    A4 28.0    A5 25.9  ← INPUT MEANS
;
PROC GLMPOWER DATA=oneway4;
      CLASS level;
      MODEL meanest = level;
      POWER
         STDDEV = 3.75 - σ̂
         ALPHA  = 0.05 - α
         NTOTAL = 10 15 20 25 30 35 40 45 50  SPECIFY N
         POWER  = . ;                         ← OUTPUT POWER
TITLE 'Case 4: Determining power for given design size and specified means';
TITLE2 'for a oneway ANOVA with 5 treatments';
TITLE3 '(total sample size N = 10 15 ... 45 50))';
RUN;
```

SAS output for Case 4: Determine power for given N assuming

$\mu_1 = 35.6,\ \mu_2 = 33.7,\ \mu_3 = 30.2,\ \mu_4 = 28.0,\ \mu_5 = 25.9$

```
Case 4: Determining power for given design size and specified means
      for a oneway ANOVA with 5 treatments
      (total sample size N = 10 15 ... 45 50))

                    The GLMPOWER Procedure
                   Fixed Scenario Elements

          Dependent Variable                  meanest
          Alpha                                  0.05
          Error Standard Deviation               3.75
```

SUPPOSE YOU WANT POWER ≥ .95

Computed Power

Index	Type	Source	n	N Total	Test DF	Error DF	Power
1	Effect	level	2	10	4	5	0.317
2	Effect	level	3	15	4	10	0.641
3	Effect	level	4	20	4	15	0.844 < ,95
4	Effect	level	5	25	4	20	0.941
5	Effect	level	6	30	4	25	0.980
6	Effect	level	7	35	4	30	0.994
7	Effect	level	8	40	4	35	0.998
8	Effect	level	9	45	4	40	>.999
9	Effect	level	10	50	4	45	>.999

USE $n = 6$ REPLICATES TO ACHIEVE POWER ≥ .95