Chapter 3: The Analysis of Variance (a Single Factor)

What If There Are More Than Two Factor Levels?

- The *t*-test does not directly apply
- There are lots of practical situations where there are either more than two levels of interest, or there are several factors of simultaneous interest
- The **analysis of variance** (ANOVA) is the appropriate analysis "engine" for these types of experiments
- The ANOVA was developed by Fisher in the early 1920s, and initially applied to agricultural experiments
- Used extensively today for industrial experiments

ANOVA

- A method to separate different components of variance (related to the experimental factors) from experimental data.
- Tests statistics are formed from the variance estimates.
  - Do factors have an effect?
  - The test statistics vary for different experimental designs.

ANOVA

- Different types of experiments will have a different assumed underlying mathematical models.
- The model will reflect characteristics of the experiment.
  - Fixed or random factors– Factor levels are set at particular values.
  - Nested factors
  - Randomization
- The previous information will determine how test statistics (for effects) are formed , and how estimates are computed.

Single Factor Experiments

- A generalization to the single factor, two level experiments where basic statistical inference methods (hypothesis tests and confidence intervals) were applied.
- $a > 2$ levels of the single factor are considered.
- Is there an effect of the factor?
  - Not yet identifying differences between treatments.

Example (pg. 66)

- An engineer is interested in investigating the relationship between the RF power setting and the etch rate for a tool. The objective of an experiment like this is to model the relationship between etch rate and RF power, and to specify the power setting that will give a desired target etch rate.
- The response variable is etch rate.
- She is interested in a particular gas (C2F6) and gap (0.80 cm), and wants to test four levels of RF power: 160W, 180W, 200W, and 220W. She decided to test five wafers at each level of RF power.
- The experimenter chooses 4 **levels** of RF power 160W, 180W, 200W, and 220W
- The experiment is **replicated** 5 times – runs made in random order

Example

Table 3-1 Etch Rate Data (in Å/min) from the Plasma Etching Experiment

Power (W)	Observations 1	2	3	4	5	Totals	Averages
160	575	542	530	539	570	2756	551.2
180	565	593	590	579	610	2937	587.4
200	600	651	610	637	629	3127	625.4
220	725	700	715	685	710	3535	707.0

- Does changing the power change the mean etch rate?
- Note that t-test doesn't apply here since there are more than 2 factor levels.

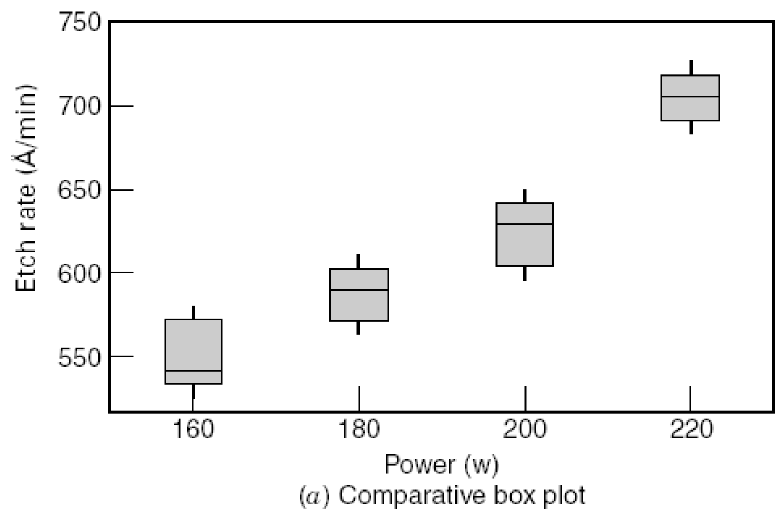


Figure 3-2 Box plots and scatter diagram of the etch rate data.

Generalization of the Example

Table 3-2 Typical Data for a Single-Factor Experiment

Treatment (level)	Observations				Totals	Averages
1	y_{11}	y_{12}	$\cdots$	y_{1n}	$y_{1.}$	$\bar{y}_{1.}$
2	y_{21}	y_{22}	$\cdots$	y_{2n}	$y_{2.}$	$\bar{y}_{2.}$
$\vdots$	$\vdots$	$\vdots$	$\cdots$	$\vdots$	$\vdots$	$\vdots$
a	y_{a1}	y_{a2}	$\cdots$	y_{an}	$y_{a.}$	$\bar{y}_{a.}$
					$y_{..}$	$\bar{y}_{..}$

- In general, there will be ***a*** **levels** of the factor, or ***a*** **treatments**, and ***n*** **replicates** of the experiment, run in **random order**…a completely randomized design **(CRD)**.
- $N = a*n$ total runs.
- We consider **fixed effects** – the factors are fixed (conclusions are applicable only to the treatments (factor levels) considered).
- Objective is to test hypotheses about the equality of the ***a*** treatment means.

The Analysis of Variance -ANOVA

- The basic single-factor ANOVA model is a linear model:

$$y_{ij} = \mu + \tau_i + \varepsilon_{ij}, \begin{cases} i = 1, 2, ..., a \\ j = 1, 2, ..., n \end{cases}$$

$\mu =$ an overall mean, $\tau_i = ith$ treatment effect,

$\varepsilon_{ij} =$ experimental error, $NID(0, \sigma^2)$

- The name “analysis of variance” stems from a partitioning of the total variability in the response variable into components that are consistent with a model for the experiment.
- The model defines how the variability will be partitioned.
- Fixed factor: a is specifically chosen by the experimenter
- Random factor: Otherwise

Models for the Data

There are several ways to write a model for the data:

$y_{ij} = \mu + \tau_i + \varepsilon_{ij}$ is called the effects model

Let $\mu_i = \mu + \tau_i$, then

$y_{ij} = \mu_i + \varepsilon_{ij}$ is called the means model

Regression models can also be employed

ANOVA

- **Total variability** is measured by the total sum of squares:

$$SS_T = \sum_{i=1}^{a}\sum_{j=1}^{n}(y_{ij} - \bar{y}_{..})^2$$

- The basic ANOVA partitioning is:

$$\sum_{i=1}^{a}\sum_{j=1}^{n}(y_{ij} - \bar{y}_{..})^2 = \sum_{i=1}^{a}\sum_{j=1}^{n}[(\bar{y}_{i.} - \bar{y}_{..}) + (y_{ij} - \bar{y}_{i.})]^2$$

$$= n\sum_{i=1}^{a}(\bar{y}_{i.} - \bar{y}_{..})^2 + \sum_{i=1}^{a}\sum_{j=1}^{n}(y_{ij} - \bar{y}_{i.})^2$$

$$SS_T = SS_{Treatments} + SS_E$$

ANOVA

$$SS_T = SS_{Treatments} + SS_E$$

- A large value of $SS_{Treatments}$ reflects large differences in treatment means.
- A small value of $SS_{Treatments}$ likely indicates no differences in treatment means.
- Formal statistical hypotheses are:

$$H_0 : \mu_1 = \mu_2 = \cdots = \mu_a$$

$$H_1 : \text{ At least one mean is different}$$

ANOVA

- While sums of squares cannot be directly compared to test the hypothesis of equal means, **mean squares** can be compared.
- A mean square is a sum of squares divided by its degrees of freedom:

$$df_{Total} = df_{Treatments} + df_{Error}$$
$$an - 1 = a - 1 + a(n-1)$$

$$MS_{Treatments} = \frac{SS_{Treatments}}{a-1}, MS_E = \frac{SS_E}{a(n-1)}$$

- If the treatment means are equal, the treatment and error mean squares will be (theoretically) equal.
- If treatment means differ, the treatment mean square will be larger than the error mean square.

Analysis of Variance Table

Table 3-3 The Analysis of Variance Table for the Single-Factor, Fixed Effects Model

Source of Variation	Sum of Squares	Degrees of Freedom	Mean Square	F_0
Between treatments	$SS_{\text{Treatments}} = n \sum_{i=1}^{a} (\bar{y}_{i.} - \bar{y}_{..})^2$	$a - 1$	$MS_{\text{Treatments}}$	$F_0 = \frac{MS_{\text{Treatments}}}{MS_E}$
Error (within treatments)	$SS_E = SS_T - SS_{\text{Treatments}}$	$N - a$	MS_E	
Total	$SS_T = \sum_{i=1}^{a} \sum_{j=1}^{n} (y_{ij} - \bar{y}_{..})^2$	$N - 1$		

- The **reference distribution** for F_0 is the $F_{a-1,\, a(n-1)}$ distribution
- **Reject** the null hypothesis (equal treatment means) if

$$F_0 > F_{\alpha, a-1, a(n-1)}$$

ANOVA

An estimator of σ^2 is

$$S_i^2 = \text{Sample variance in the } ith \text{ treatment} = \frac{\sum_{j=1}^{n}(y_{ij}-\bar{y}_{i.})^2}{n-1}$$

Pooled variance estimate from a treatments

$$= \frac{(n-1)S_1^2 + \cdots + (n-1)S_a^2}{(n-1)+\cdots+(n-1)} = \frac{\sum_{i=1}^{a}\left[\sum_{j=1}^{n}(y_{ij}-\bar{y}_{i.})^2\right]}{\sum_{i=1}^{a}(n-1)}$$

$$= ?$$

ANOVA

If the treatment means are the same then an estimator of σ^2/n is

$$\frac{S^2}{n} = \text{Sample variance of the treatments} = \frac{\sum_{i=1}^{n}(\bar{y}_{i.} - \bar{y}_{..})^2}{a-1}$$

Therefore

$$\frac{n\sum_{j=1}^{n}(\bar{y}_{i.} - \bar{y}_{..})^2}{a-1} \text{ is an estimator of } \sigma^2.$$

$$\frac{n\sum_{j=1}^{n}(\bar{y}_{i.} - \bar{y}_{..})^2}{a-1} = ?$$

ANOVA

- A more rigorous approach is to take expected values of the mean squares.

$$E(MS_E) = E\left(\frac{SS_E}{N-a}\right) = \frac{1}{N-a} E\left[\sum_{i=1}^{a}\sum_{j=1}^{n}(y_{ij} - \bar{y}_{i.})^2\right]$$

$$= \frac{1}{N-a} E\left[\sum_{i=1}^{a}\sum_{j=1}^{n} y_{ij}^2 - \frac{1}{n}\sum_{i=1}^{a} y_{i.}^2\right]$$

$$= \frac{1}{N-a} E\left[\sum_{i=1}^{a}\sum_{j=1}^{n}(\mu + \tau_i + \varepsilon_{ij})^2 - \frac{1}{n}\sum_{i=1}^{a}\left(\sum_{j=1}^{n}(\mu + \tau_i + \varepsilon_{ij})\right)^2\right]$$

$$= \sigma^2$$

ANOVA

Also

$$E(MS_{Treatments}) = \sigma^2 + \frac{n\sum_{i=1}^{a}\tau_i^2}{a-1}$$

No difference in treatment means $\Rightarrow \tau_{\text{i}} = 0 \Rightarrow E(MS_E) = E(MS_{Treatments})$.

To statistically test whether there is no difference in treatment means we need a test statistic and a sampling distribution.

ANOVA

It can be shown that $SS_{Treatments}/\sigma^2$ and SS_E/σ^2 are independent random variables

(Cochran's Theorem page 69 of the text) that have

a Chi-square distribution (sums of squared standard normal

random variables) with a-1 and N-1 degrees of freedom respectively.

What is the distribution of the ratio of two Chi-square random variables divided by their degrees of freedom?

Analysis of Variance Table

Table 3-3 The Analysis of Variance Table for the Single-Factor, Fixed Effects Model

Source of Variation	Sum of Squares	Degrees of Freedom	Mean Square	F_0
Between treatments	$SS_{\text{Treatments}} = n \sum_{i=1}^{a} (\bar{y}_{i.} - \bar{y}_{..})^2$	$a - 1$	$MS_{\text{Treatments}}$	$F_0 = \dfrac{MS_{\text{Treatments}}}{MS_E}$
Error (within treatments)	$SS_E = SS_T - SS_{\text{Treatments}}$	$N - a$	MS_E	
Total	$SS_T = \sum_{i=1}^{a} \sum_{j=1}^{n} (y_{ij} - \bar{y}_{..})^2$	$N - 1$		

- The **reference distribution** for F_0 is the $F_{a-1,\, a(n-1)}$ distribution
- **Reject** the null hypothesis (equal treatment means) if

$$F_0 > F_{\alpha, a-1, a(n-1)}$$

ANOVA for the Example

Table 3-4 ANOVA for the Plasma Etching Experiment

Source of Variation	Sum of Squares	Degrees of Freedom	Mean Square	F_0	P-Value
RF Power	66,870.55	3	22,290.18	$F_0 = 66.80$	<0.01
Error	5339.20	16	333.70		
Total	72,209.75	19			

The Reference Distribution:

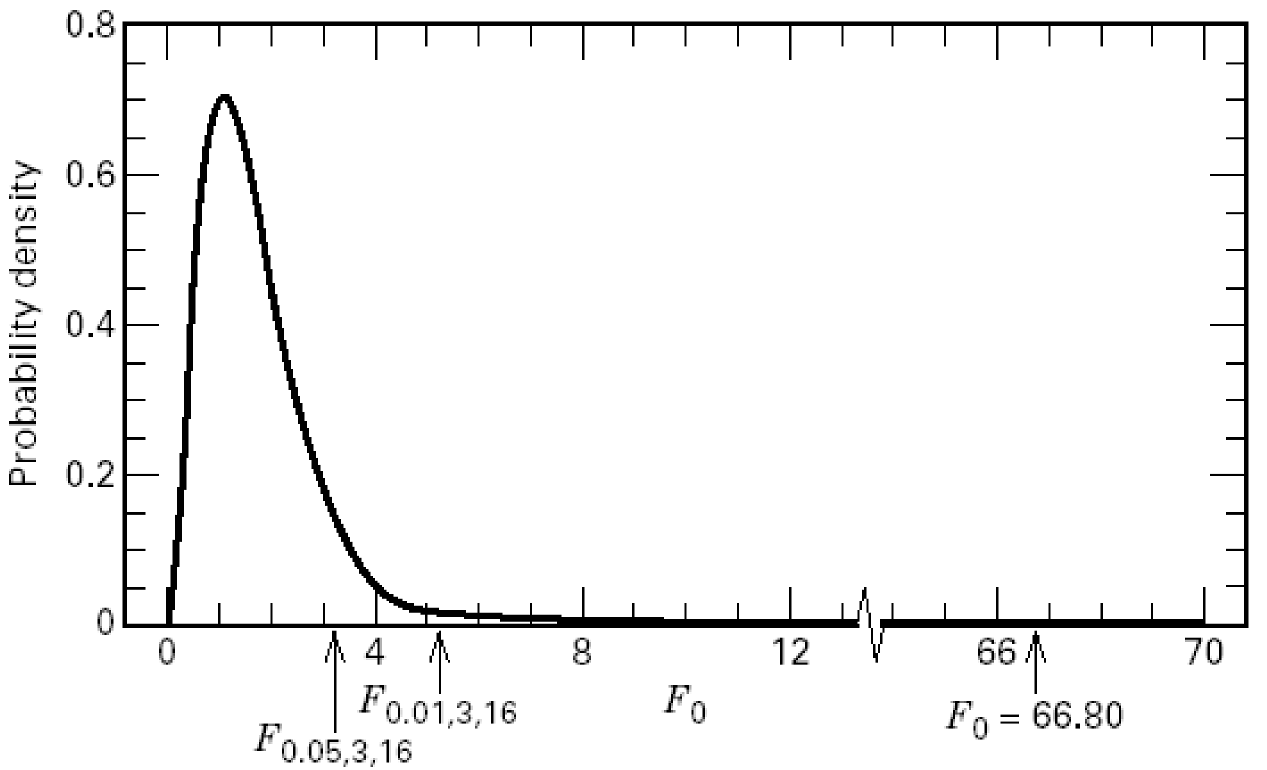


Figure 3-3 The reference distribution ($F_{3,16}$) for the test statistic F_0 in Example 3-1.

Predicted Values – Estimation of Model Parameters

$$\hat{\mu} = \bar{y}_{..}$$

$$\hat{\tau}_i = \bar{y}_{i.} - \bar{y}_{..}$$

$$\hat{\mu}_i = \hat{\mu} + \hat{\tau}_i = \bar{y}_{i.}$$

ANOVA calculations are usually done via computer

- Text exhibits sample calculations from three very popular software packages, Design-Expert, JMP and Minitab
- See activity for SAS coding
- Text discusses some of the summary statistics provided by these packages

Model Adequacy Checking in the ANOVA

- Checking assumptions
  - Normality
  - Constant variance
  - Independence
- What to do if some of these assumptions are violated.

Model Adequacy Checking in the ANOVA

- Verified by an examination of **residuals**

$$e_{ij} = y_{ij} - \hat{y}_{ij}$$
$$= y_{ij} - \bar{y}_{i.}$$

- **Residual plots** are very useful.
- **Normal probability plot** of residuals.

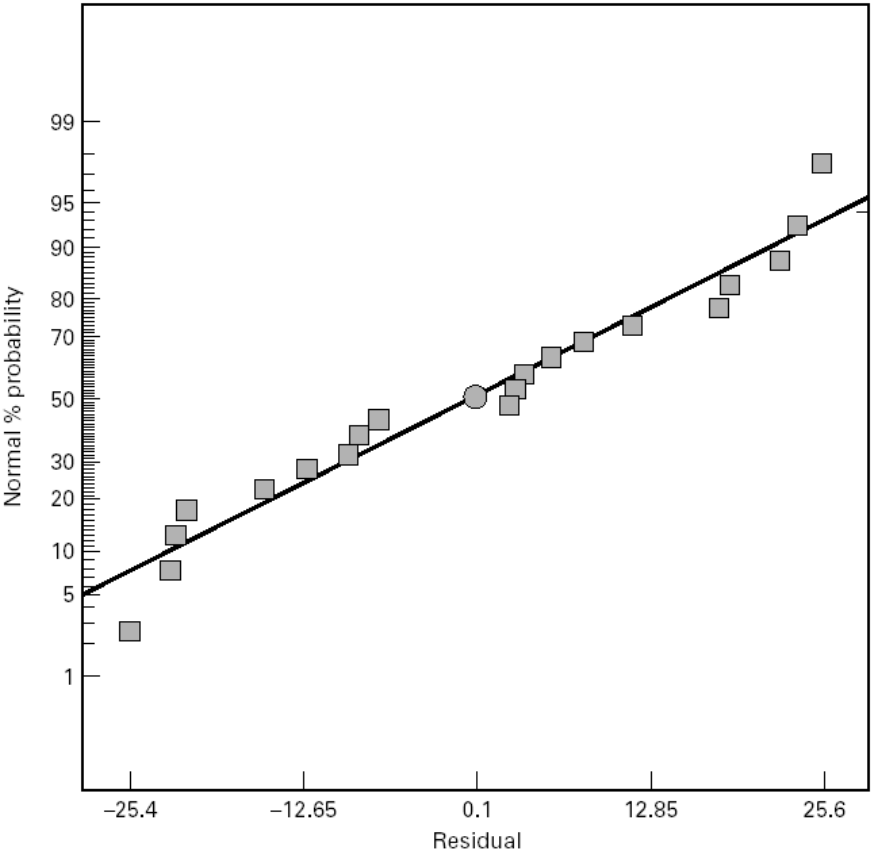


Figure 3-4 Normal probability plot of residuals for Example 3-1.

Examination of Residuals

- Standardized residuals

$$d_{ij} = \frac{e_{ij}}{\sqrt{MS_E}} \text{ should be approximately } N(0,\sigma^2)$$

Standardized residuals > 3 may be considered outliers based on standard normal probabilities.

Other Residual Plots

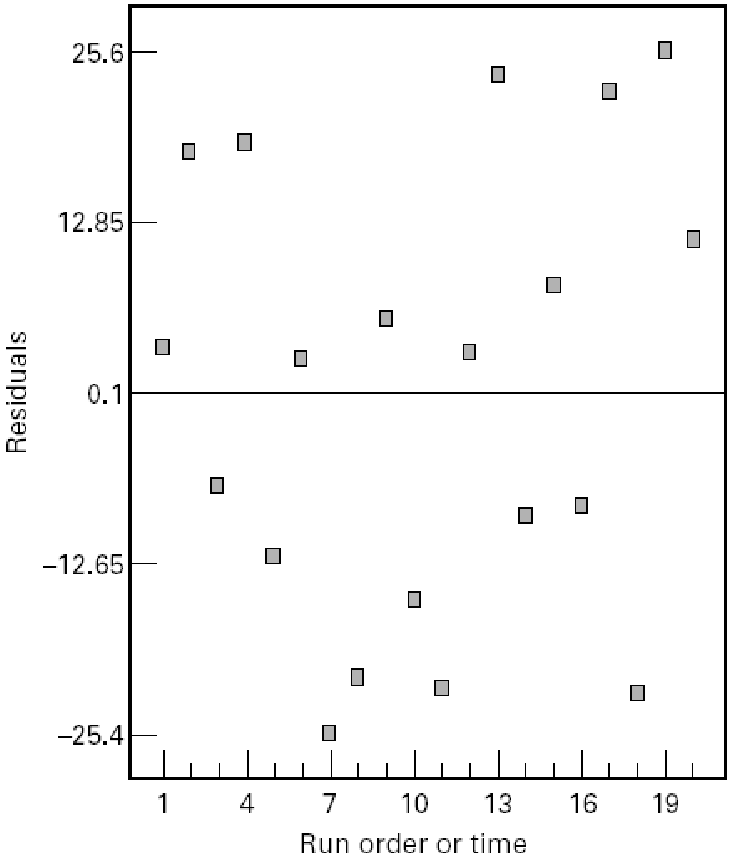


Figure 3-5 Plot of residuals versus run order or time.

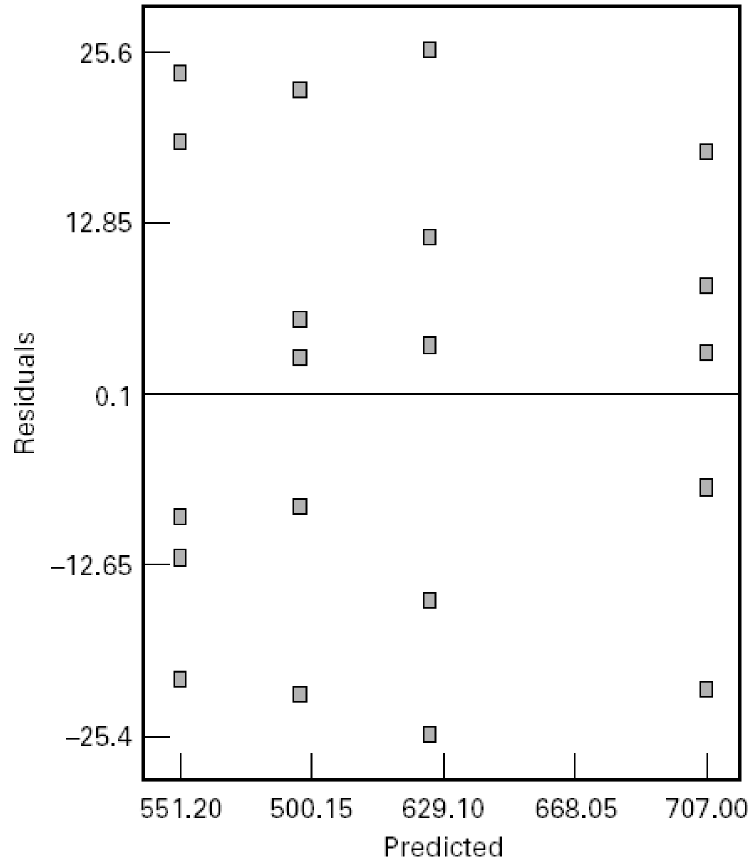


Figure 3-6 Plot of residuals versus fitted values.

Formal Test for Equality of Variance

$$H_0: \quad \sigma_1^2 = \sigma_1^2 = \ldots = \sigma_a^2$$

- Compute deviations of each observation from the treatment median.

$$d_{ij} = \left| y_{ij} - \tilde{y}_i \right| \text{ where } \tilde{y}_i = \text{treatment } i \text{ median.}$$

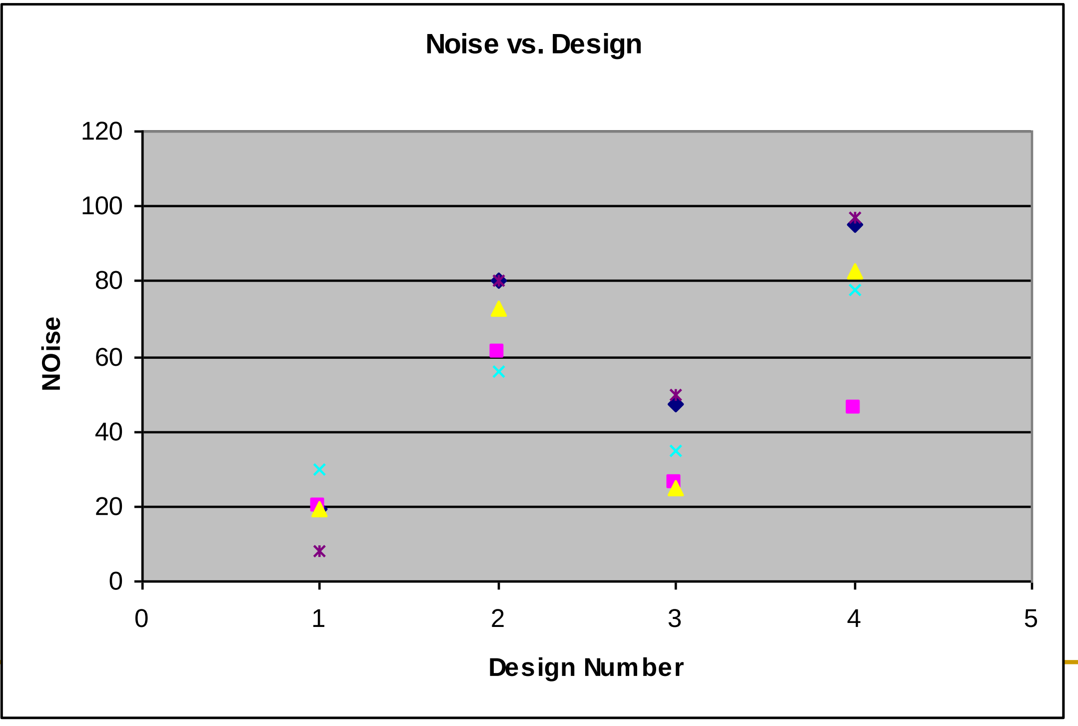
- Apply the ANOVA F-test for equality of the deviation means.
- Called the modified Levene test.

Example – Problem 3-24

Four different designs for a digital computer circuit are being studied to compare the amount of noise present.

Noise Observed from Different Circuit Designs

	Replication				
Design	**1**	**2**	**3**	**4**	**5**
D1	19	20	19	30	8
D2	80	61	73	56	80
D3	47	26	25	35	50
D4	95	46	83	78	97



Example – Problem 3-24

Anova: Single Factor

SUMMARY

Groups	*Count*	*Sum*	*Average*	*Variance*
D1	5	96	19.2	60.7
D2	5	350	70	121.5
D3	5	183	36.6	134.3
D4	5	399	79.8	420.7

ANOVA

Source of Variation	*SS*	*df*	*MS*	*F*	*P-value*	*F crit(0.05)*
Between Groups	12042	3	4014	21.77971	6.8E-06	3.2388715
Within Groups	2948.8	16	184.3			
Total	14990.8	19				

Example – Problem 3-24

- Check assumptions

Noise Observed from Different Circuit Designs

Design	Replication 1	2	3	4	5	$y_{i.}$
D1	19	20	19	30	8	19.2
D2	80	61	73	56	80	70
D3	47	26	25	35	50	36.6
D4	95	46	83	78	97	79.8

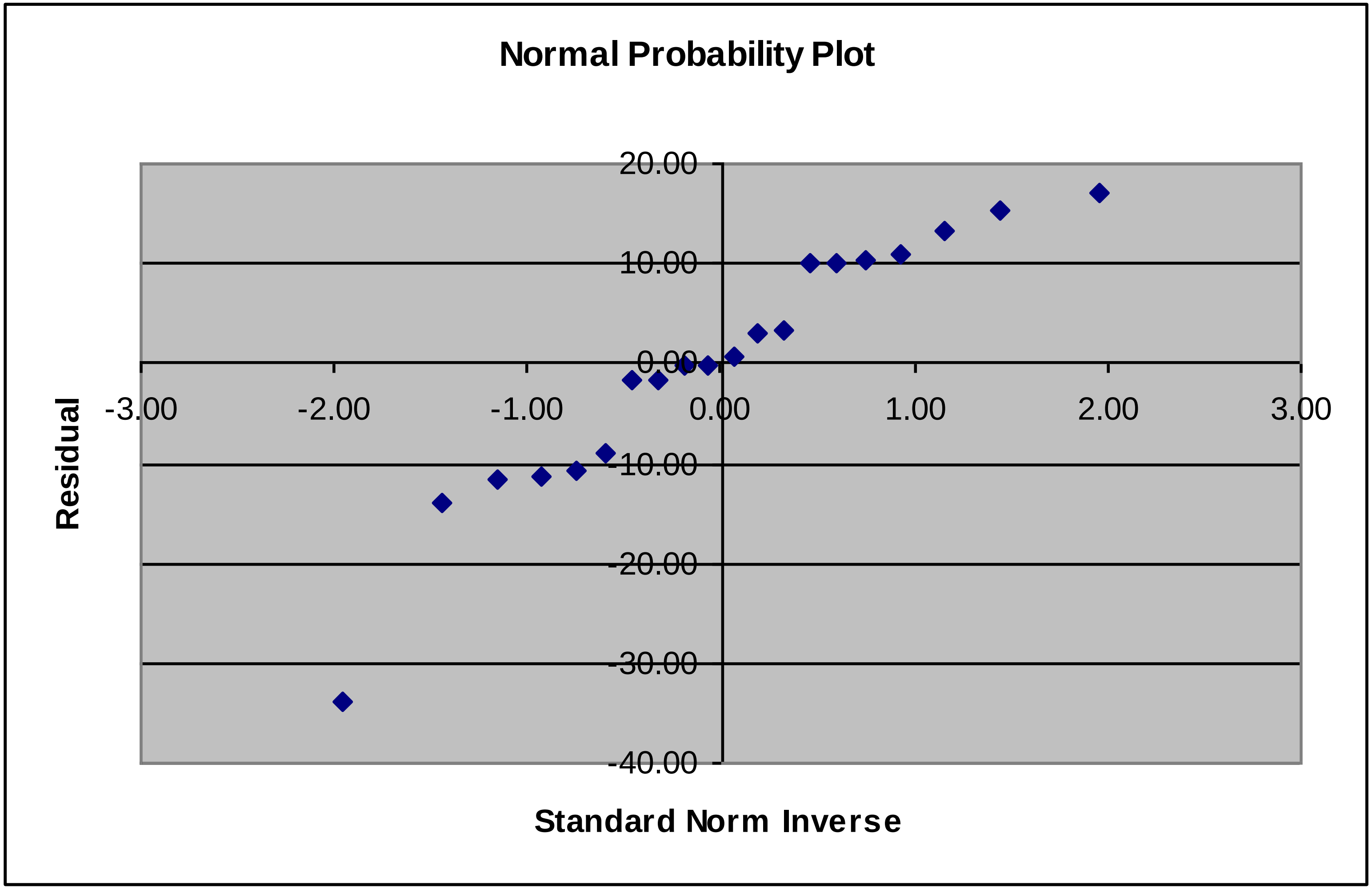
Residuals

Design	Replication 1	2	3	4	5
D1	-0.2	0.8	-0.2	10.8	-11.2
D2	10	-9	3	-14	10
D3	10.4	-10.6	-11.6	-1.6	13.4
D4	15.2	-33.8	3.2	-1.8	17.2

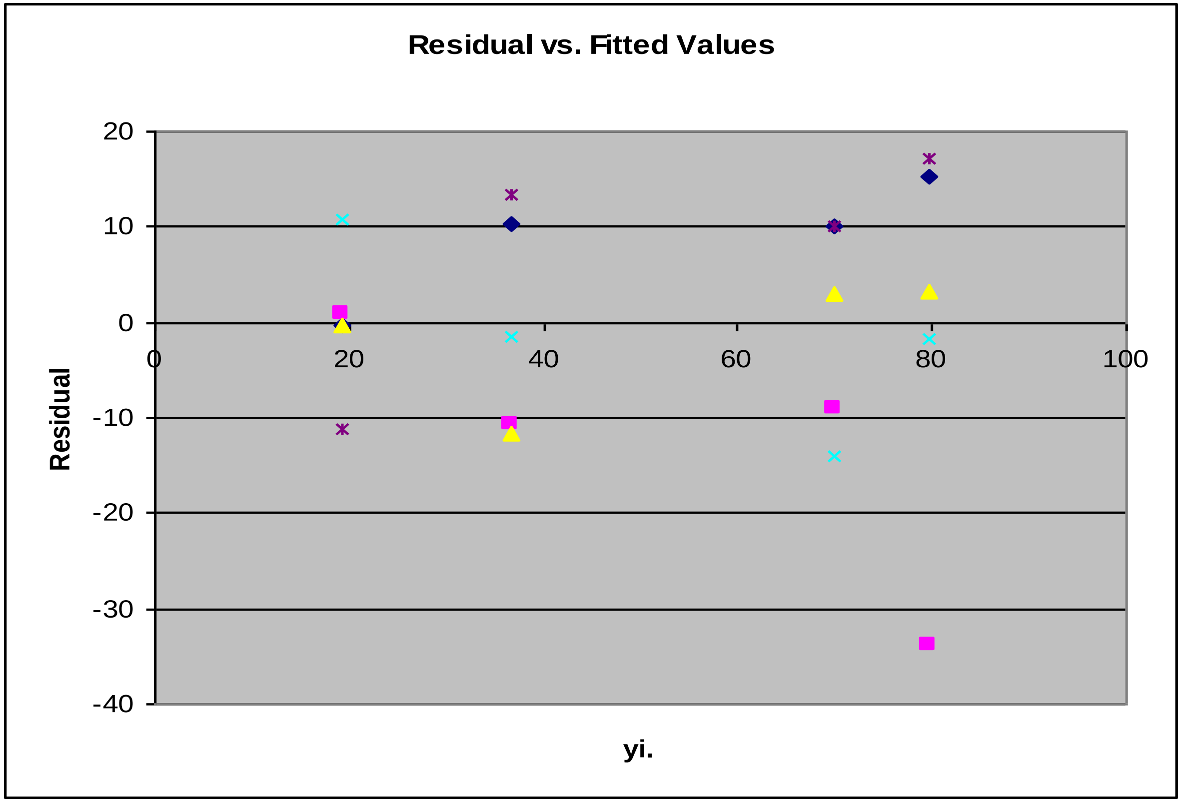
Example – Problem 3-24

j	Data	(j - 0.5)/n	Standard Norm Inverse	Sorted Data
1	-0.20	0.03	-1.96	-33.80
2	0.80	0.08	-1.44	-14.00
3	-0.20	0.13	-1.15	-11.60
4	10.80	0.18	-0.93	-11.20
5	-11.20	0.23	-0.76	-10.60
6	10.00	0.28	-0.60	-9.00
7	-9.00	0.33	-0.45	-1.80
8	3.00	0.38	-0.32	-1.60
9	-14.00	0.43	-0.19	-0.20
10	10.00	0.48	-0.06	-0.20
11	10.40	0.53	0.06	0.80
12	-10.60	0.58	0.19	3.00
13	-11.60	0.63	0.32	3.20
14	-1.60	0.68	0.45	10.00
15	13.40	0.73	0.60	10.00
16	15.20	0.78	0.76	10.40
17	-33.80	0.83	0.93	10.80
18	3.20	0.88	1.15	13.40
19	-1.80	0.93	1.44	15.20
20	17.20	0.98	1.96	17.20

Example – Problem 3-24



Example – Problem 3-24



Standardized Residuals

Design	Replication 1	2	3	4	5
D1	-0.01	0.06	-0.01	0.80	-0.82
D2	0.74	-0.66	0.22	-1.03	0.74
D3	0.77	-0.78	-0.85	-0.12	0.99
D4	1.12	-2.49	0.24	-0.13	1.27

Example – Problem 3-24

- Conduct a modified Levene test for equality of variance.

Noise Observed from Different Circuit Designs

	Replication					
Design	**1**	**2**	**3**	**4**	**5**	**Median**
D1	19	20	19	30	8	19
D2	80	61	73	56	80	73
D3	47	26	25	35	50	35
D4	95	46	83	78	97	83

Deviations for the modified Levene test

	Replication				
Design	**1**	**2**	**3**	**4**	**5**
D1	0	1	0	11	11
D2	7	12	0	17	7
D3	12	9	10	0	15
D4	12	37	0	5	14

Example – Problem 3-24

Anova: Single Factor

SUMMARY

Groups	*Count*	*Sum*	*Average*	*Variance*
D1	5	23	4.6	34.3
D2	5	43	8.6	40.3
D3	5	46	9.2	31.7
D4	5	68	13.6	202.3

ANOVA

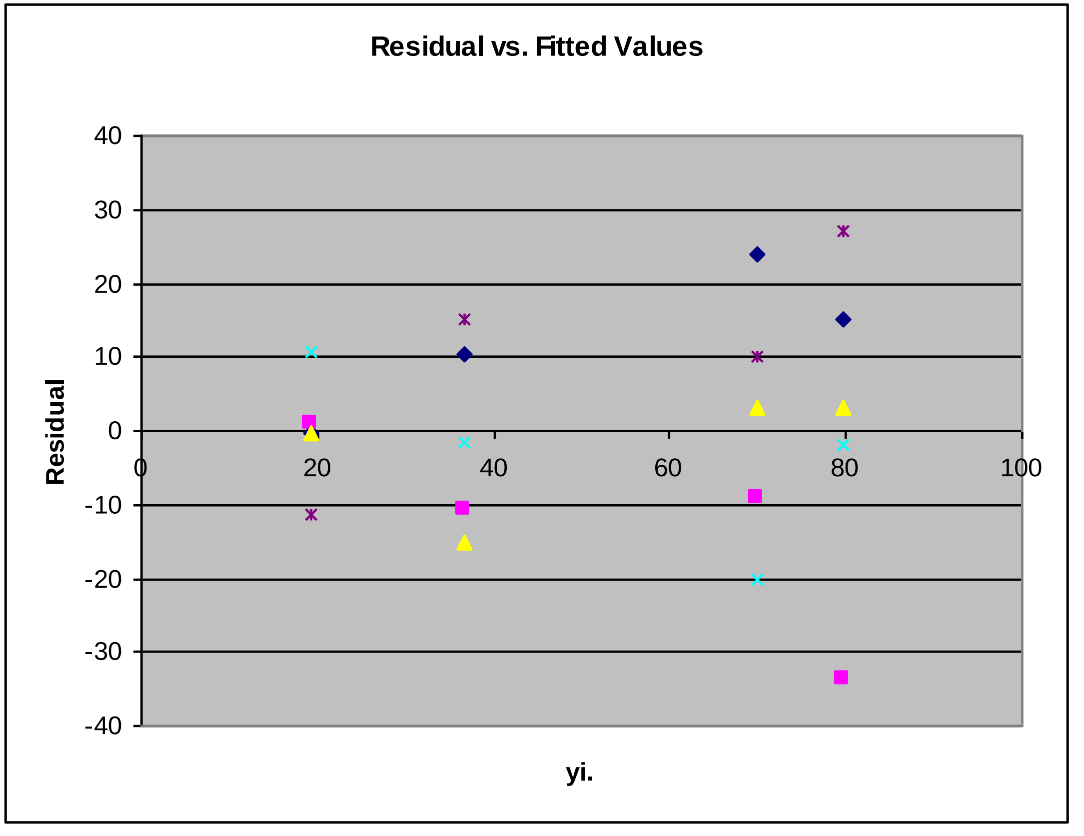
Source of Variation	*SS*	*df*	*MS*	*F*	*P-value*	*F crit(0.05)*
Between Groups	203.6	3	67.86666667	0.879672	0.472383	3.2388715
Within Groups	1234.4	16	77.15			
Total	1438	19				

What if Assumptions are Not Met?

- Unequal variances and non-normality of residuals often occur together.
  - First address unequal variance (through data transformations).
  - Assess transformed data for normality of residuals.

Example

- Variance appears to be increasing as a function of the treatment mean.



Variance Stabilizing Transformation

- Transform the raw data
  - $y^* = f(y)$.
- Apply ANOVA to the transformed data.
  - Conclusions apply to the transformed population.

Data Transformations

Transform the data to equalize variance that changes as a function of the mean.

Let Y be a random variable with $E(Y) = \mu$ and standard deviation σ_Y.

Assuming the unequal variances change as a function of the mean (e.g., the variance of the residuals increases as $y_{i.}$ increases) and the standard deviation of $Y = \sigma_Y$ is roughly proportional to a power of μ.

$$\sigma_Y \propto \mu^{\alpha}$$

Let $Y^* = Y^{\lambda}$ be a transformation of Y (with data, individual observations would be transformed $y_{ij}^* = (y_{ij})^{\lambda}$).

Then $\sigma_{Y^*} \propto \mu^{\alpha+\lambda-1} \Rightarrow$ Transform the data with $\lambda = 1 - \alpha$.

Data Transformations

Estimating α

In the *ith* treatment $\sigma_{y_i} \propto \mu_i^{\alpha} = \theta \mu_i^{\alpha}$

$\Rightarrow \ln \sigma_{y_i} = \ln \theta + \alpha \ln \mu_i$

$\Rightarrow$ Plot $\ln \sigma_{y_i}$ vs. $\alpha \ln \mu_i$ using estimates of σ_{y_i} and μ_i .

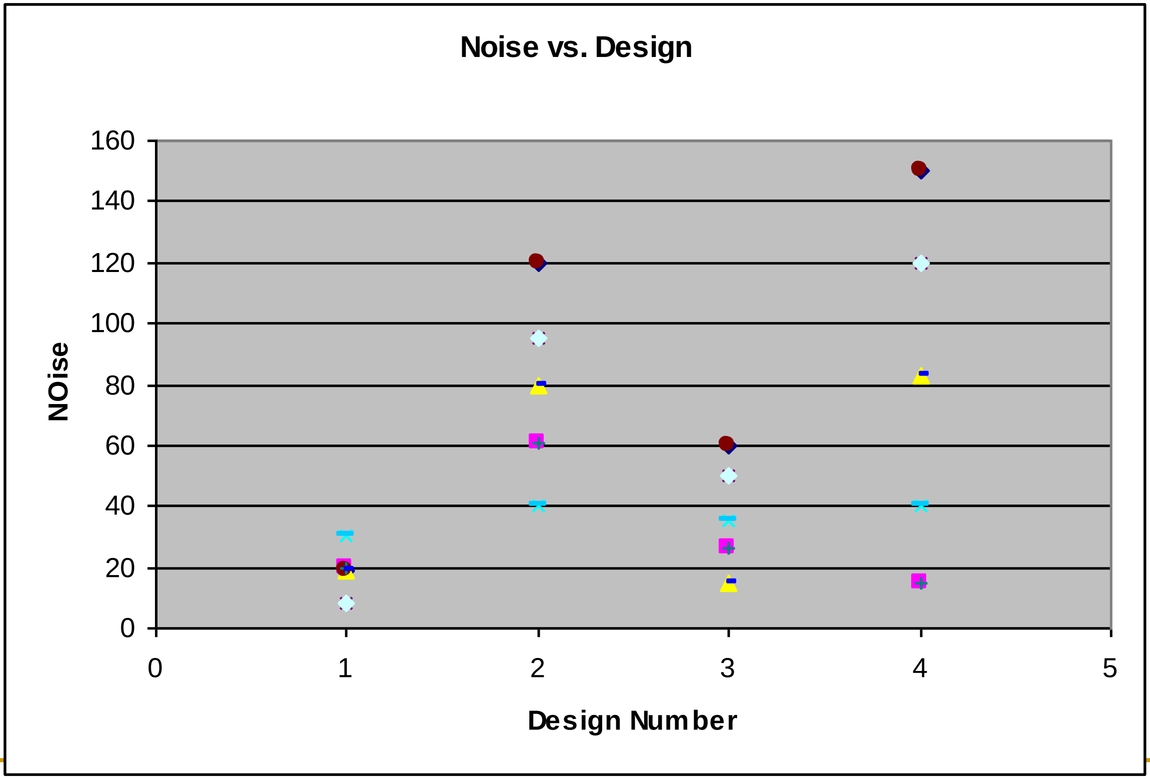
$\hat{\sigma}_{y_i}$ = Sample standard deviation of the ith treatment.

$\hat{\mu}_i$ = Sample average of the ith treatment.

Example – Problem 3-24 (Original)

Noise Observed from Different Circuit Designs

	Replication					
Design	**1**	**2**	**3**	**4**	**5**	$y_{i.}$
D1	19	20	19	30	8	19.2
D2	120	61	80	40	95	79.2
D3	60	26	15	35	50	37.2
D4	150	15	83	40	120	81.6



Example – Problem 3-24 (Original)

Anova: Single Factor

SUMMARY

Groups	Count	Sum	Average	Variance
D1	5	96	19.2	60.7
D2	5	396	79.2	945.7
D3	5	186	37.2	326.7
D4	5	408	81.6	3080.3

ANOVA

Source of Variation	SS	df	MS	F	P-value	F crit
Between Groups	14448.6	3	4816.2	4.36507	0.019936	3.2388715
Within Groups	17653.6	16	1103.35			
Total	32102.2	19				

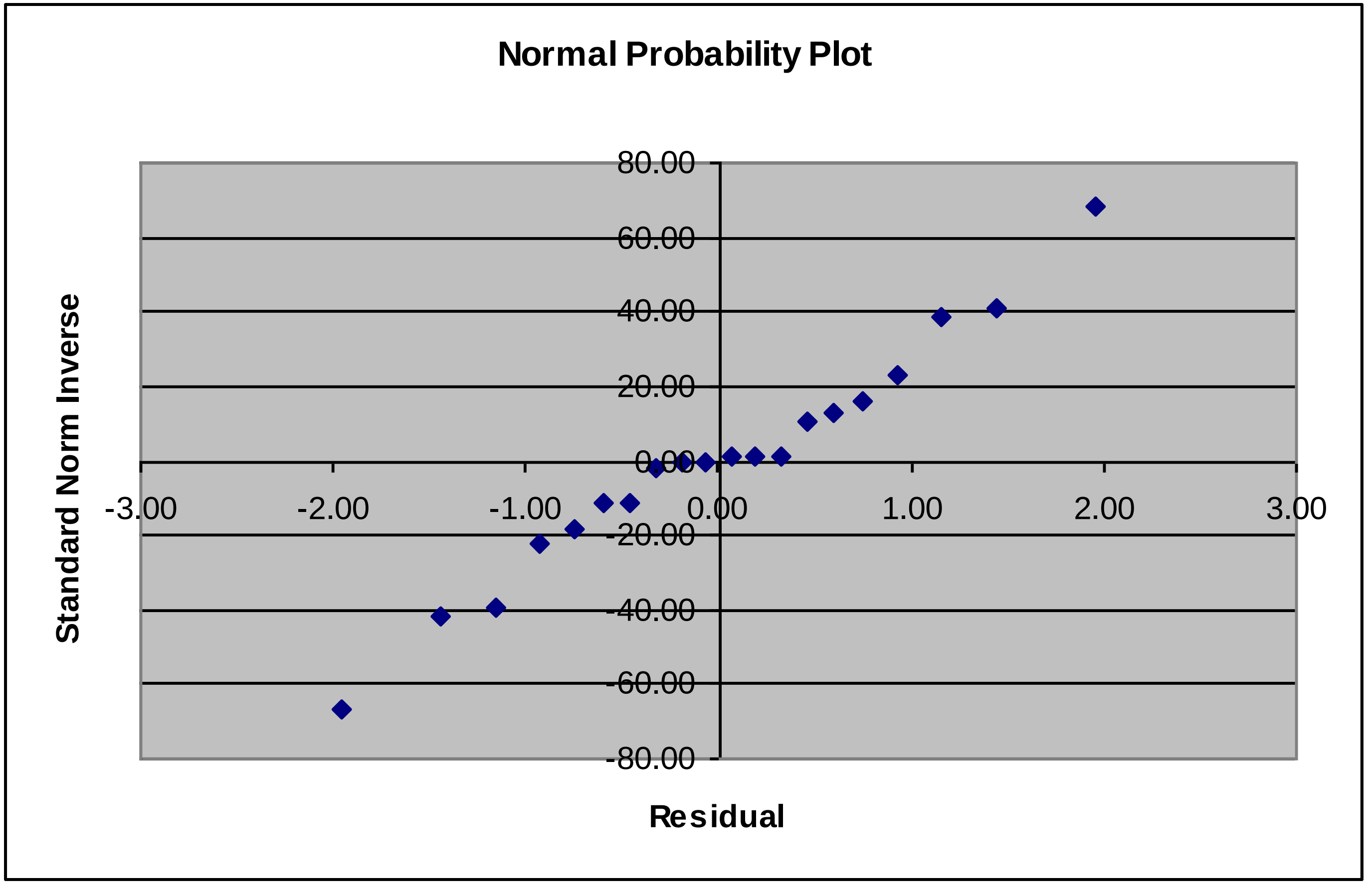
Example – Problem 3-24 (Original)

Residuals

	Replication				
$y_{i.}$	1	2	3	4	5
19.2	-0.2	0.8	-0.2	10.8	-11.2
70	40.8	-18.2	0.8	-39.2	15.8
36.6	22.8	-11.2	-22.2	-2.2	12.8
79.8	68.4	-66.6	1.4	-41.6	38.4

j	Data	(j - 0.5)/n	Standard Norm Inverse	Sorted Data
1	-0.20	0.03	-1.96	-66.60
2	0.80	0.08	-1.44	-41.60
3	-0.20	0.13	-1.15	-39.20
4	10.80	0.18	-0.93	-22.20
5	-11.20	0.23	-0.76	-18.20
6	40.80	0.28	-0.60	-11.20
7	-18.20	0.33	-0.45	-11.20
8	0.80	0.38	-0.32	-2.20
9	-39.20	0.43	-0.19	-0.20
10	15.80	0.48	-0.06	-0.20
11	22.80	0.53	0.06	0.80
12	-11.20	0.58	0.19	0.80
13	-22.20	0.63	0.32	1.40
14	-2.20	0.68	0.45	10.80
15	12.80	0.73	0.60	12.80
16	68.40	0.78	0.76	15.80
17	-66.60	0.83	0.93	22.80
18	1.40	0.88	1.15	38.40
19	-41.60	0.93	1.44	40.80
20	38.40	0.98	1.96	68.40

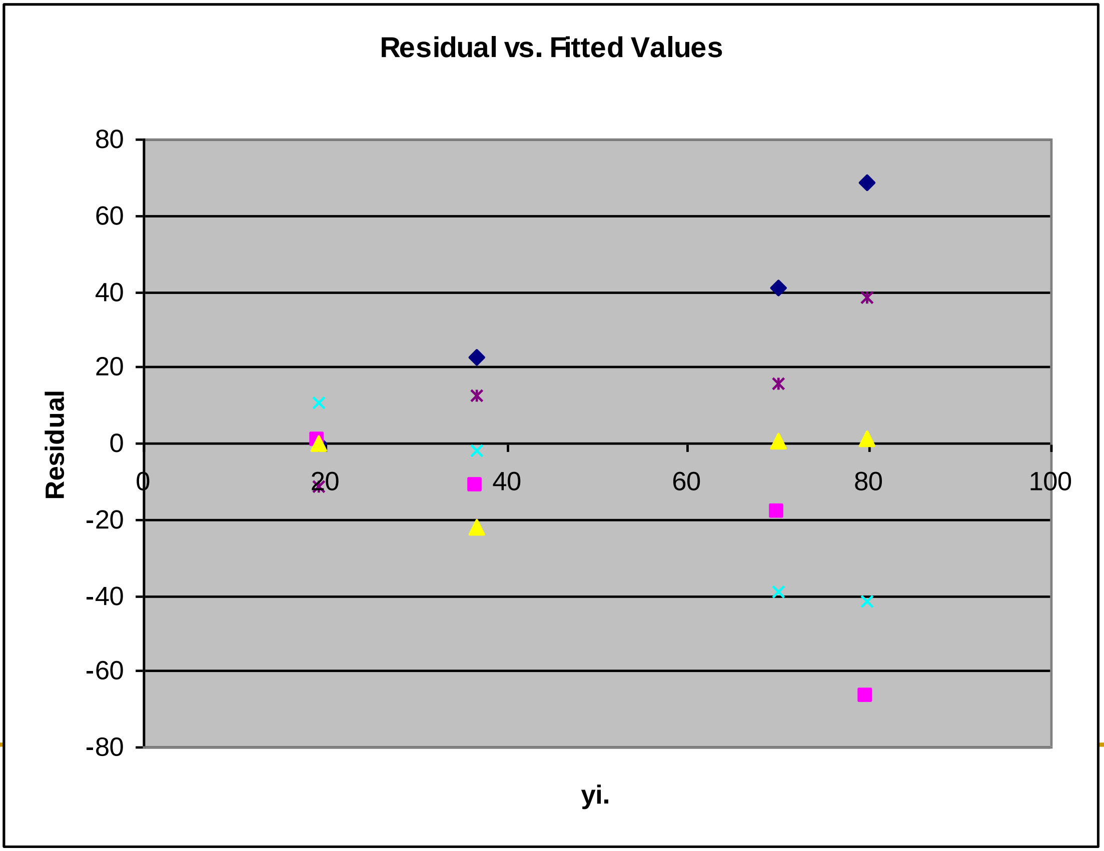
Example – Problem 3-24 (Original)



Example – Problem 3-24 (Modified 1)

Standardized Residuals

Design	Replication 1	2	3	4	5
D1	-0.01	0.02	-0.01	0.33	-0.34
D2	1.23	-0.55	0.02	-1.18	0.48
D3	0.69	-0.34	-0.67	-0.07	0.39
D4	2.06	-2.01	0.04	-1.25	1.16



Example – Problem 3-24 (Modified 1)

- Modified Levine test

Noise Observed from Different Circuit Designs

Design	Replication 1	2	3	4	5	Median
D1	19	20	19	30	8	19
D2	120	61	80	40	95	80
D3	60	26	15	35	50	35
D4	150	15	83	40	120	83

Deviations for the modified Levene test

Design	Replication 1	2	3	4	5
D1	0	1	0	11	11
D2	40	19	0	40	15
D3	25	9	20	0	15
D4	67	68	0	43	37

Example – Problem 3-24 (Modified 1)

Anova: Single Factor

SUMMARY

Groups	*Count*	*Sum*	*Average*	*Variance*
D1	5	23	4.6	34.3
D2	5	114	22.8	296.7
D3	5	69	13.8	94.7
D4	5	215	43	771.5

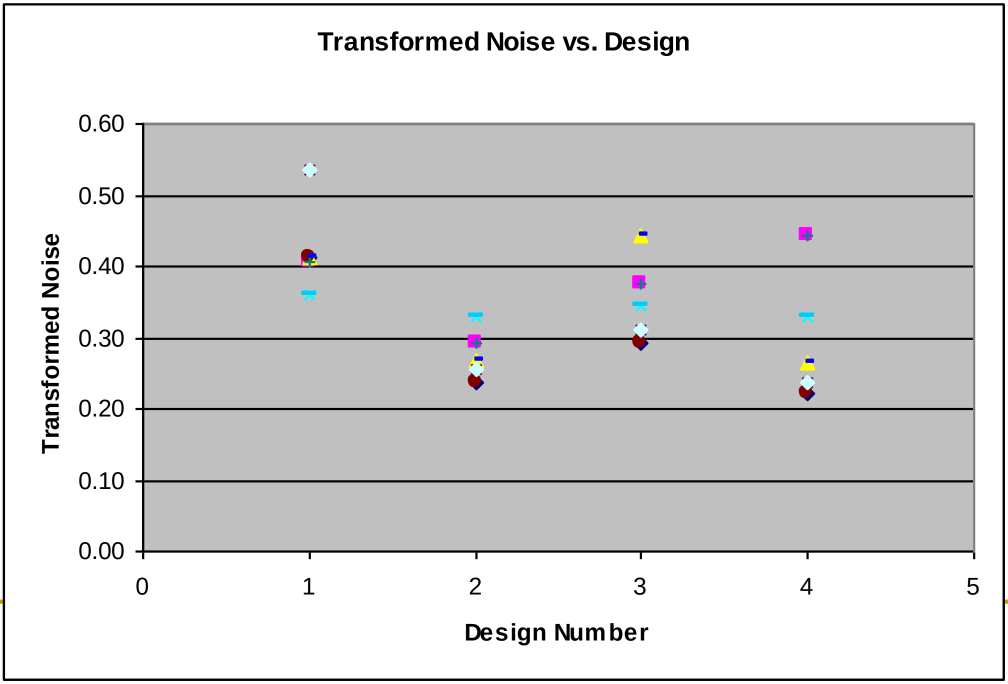
ANOVA

Source of Variation	*SS*	*df*	*MS*	*F*	*P-value*	*F crit*
Between Groups	4040.15	3	1346.716667	4.499555	0.017974	3.2388715
Within Groups	4788.8	16	299.3			
Total	8828.95	19				

Example – Problem 3-24 (Modified 2)

Transformed Noise Observed from Different Circuit Designs

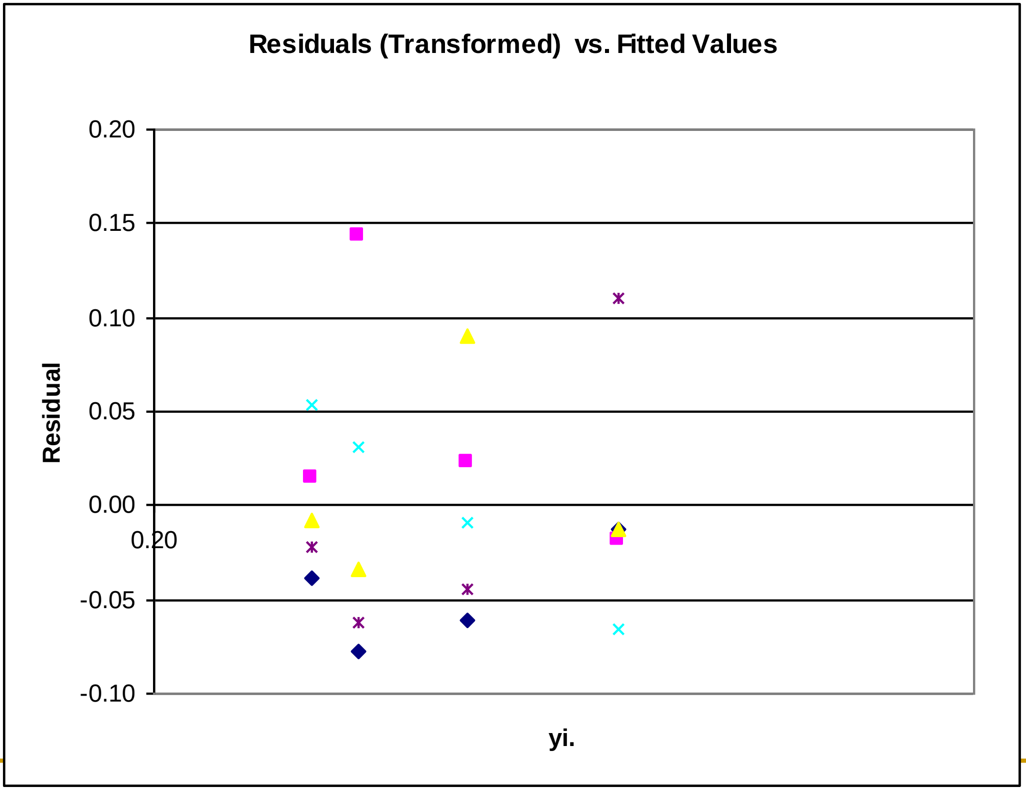
	Replication					
Design	**1**	**2**	**3**	**4**	**5**	$y_{i.}$
D1	0.41	0.41	0.41	0.36	0.54	0.43
D2	0.24	0.29	0.27	0.33	0.26	0.28
D3	0.29	0.38	0.44	0.34	0.31	0.35
D4	0.22	0.44	0.27	0.33	0.24	0.30



Example – Problem 3-24 (Modified 2)

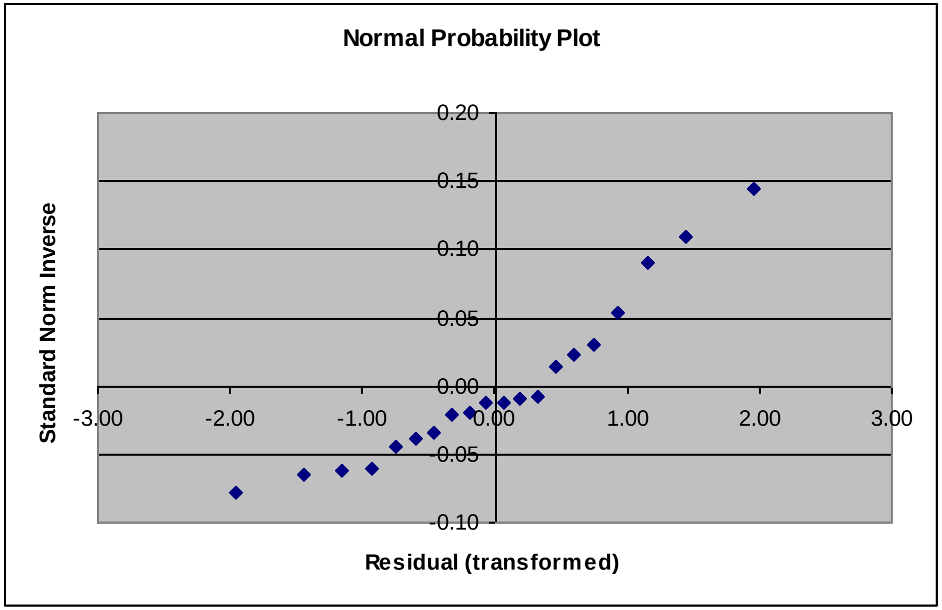
Residuals

	Replication				
$y_{i.}$	**1**	**2**	**3**	**4**	**5**
0.43	-0.01	-0.02	-0.01	-0.07	0.11
0.28	-0.04	0.01	-0.01	0.05	-0.02
0.35	-0.06	0.02	0.09	-0.01	-0.04
0.30	-0.08	0.14	-0.03	0.03	-0.06



Example – Problem 3-24 (Modified 2)

j	Data	(j - 0.5)/n	Standard Norm Inverse	Sorted Data
1	-0.01	0.03	-1.96	-0.08
2	-0.02	0.08	-1.44	-0.07
3	-0.01	0.13	-1.15	-0.06
4	-0.07	0.18	-0.93	-0.06
5	0.11	0.23	-0.76	-0.04
6	-0.04	0.28	-0.60	-0.04
7	0.01	0.33	-0.45	-0.03
8	-0.01	0.38	-0.32	-0.02
9	0.05	0.43	-0.19	-0.02
10	-0.02	0.48	-0.06	-0.01
11	-0.06	0.53	0.06	-0.01
12	0.02	0.58	0.19	-0.01
13	0.09	0.63	0.32	-0.01
14	-0.01	0.68	0.45	0.01
15	-0.04	0.73	0.60	0.02
16	-0.08	0.78	0.76	0.03
17	0.14	0.83	0.93	0.05
18	-0.03	0.88	1.15	0.09
19	0.03	0.93	1.44	0.11
20	-0.06	0.98	1.96	0.14



Example – Problem 3-24 (Modified 2)

Deviations for the modified Levene test

	Replication				
Design	**1**	**2**	**3**	**4**	**5**
D1	0.00	0.01	0.00	0.05	0.12
D2	0.03	0.02	0.00	0.06	0.01
D3	0.05	0.03	0.10	0.00	0.03
D4	0.04	0.18	0.00	0.07	0.03

Anova: Single Factor

SUMMARY

Groups	*Count*	*Sum*	*Average*	*Variance*
D1	5	0.181734	0.03634681	0.00281
D2	5	0.129099	0.025819714	0.000542
D3	5	0.218024	0.043604775	0.001327
D4	5	0.314207	0.062841478	0.004716

ANOVA

Source of Variation	*SS*	*df*	*MS*	*F*	*P-value*	*F crit*
Between Groups	0.003653	3	0.001217681	0.518494	0.675524	3.2388715
Within Groups	0.037576	16	0.002348496			
Total	0.041229	19				

Comparison Among Treatment Means

- If the null hypothesis that all treatments are equal is rejected (from the ANOVA) more analysis is needed to determine which means differ.
- The methods used are called <u>multiple comparison methods</u>.

Comparison Among Treatment Means

- Various hypotheses about the equality or inequality of treatment means can be expressed as a contrast.
- A contrast is a linear combination of parameters expressed as

$$\Gamma = \sum_{i=1}^{a} c_i \mu_i \quad \text{where} \sum_{i=1}^{a} c_i = 0.$$

Examples

- Suppose there are four treatments in a single factor experiment.

$$H_0 : \mu_3 - \mu_4 = 0$$
$$H_1 : \mu_3 - \mu_4 \neq 0 \quad \Rightarrow c_1, c_2 = 0, c_3 = 1, c_4 = -1$$

Comparison Among Treatment Means

- **Method1: t-test**

Use $\bar{y}_{i.}$ as an estimator of μ_i, and MS_E as an estimator of σ^2.

If $C = \sum_{i=1}^{a} c_i \bar{y}_{i.}$ then $Var(C) = \frac{\sigma^2}{n} \sum_{i=1}^{a} c_i^2$.

If the null hypothesis $H_0 : \sum_{i=1}^{a} c_i \mu_i = 0$ is true, then

$$\frac{\sum_{i=1}^{a} c_i \bar{y}_{i.}}{\sqrt{\frac{\sigma^2}{n} \sum_{i=1}^{a} c_i^2}} \text{ will be } \sim N(0,1).$$

Since we estimate σ^2 with MS_E, the test statistic

$$t_0 = \frac{\sum_{i=1}^{a} c_i \bar{y}_{i.}}{\sqrt{\frac{MS_E}{n} \sum_{i=1}^{a} c_i^2}} \text{ will be } \sim t_{N-a}.$$

Reject H_0 if $t_0 > t_{\alpha/2, N-a}$

Comparison Among Treatment Means

- Method 2: F-test

Result : The square of a t random variable with $N-a$ degrees of freedom has an F distribution with 1 numerator and $N-a$ denominator degrees of freedom.

Test statistic

$$F_0 = t_0^{\ 2} = \frac{\left(\sum_{i=1}^{a} c_i \bar{y}_{i.}\right)^2}{\frac{MS_E}{n}\sum_{i=1}^{a} c_i^2}$$

Reject H_0 if $F_0 > F_{\alpha,1,N-a}$

This test statistic can be expressed as a ratio of mean squares

$$F_0 = \frac{MS_C}{MSE} \text{ where } MS_C = \frac{SS_C}{1} = \frac{\left(\sum_{i=1}^{a} c_i \bar{y}_{i.}\right)^2}{\frac{1}{n}\sum_{i=1}^{a} c_i^2}$$

Comparison Among Treatment Means

- Orthogonal contrast (A special case)

Two contrasts with coefficients c_i and d_i are orthogonal contrasts if

$$\sum_{i=1}^{a} c_i d_i = 0.$$

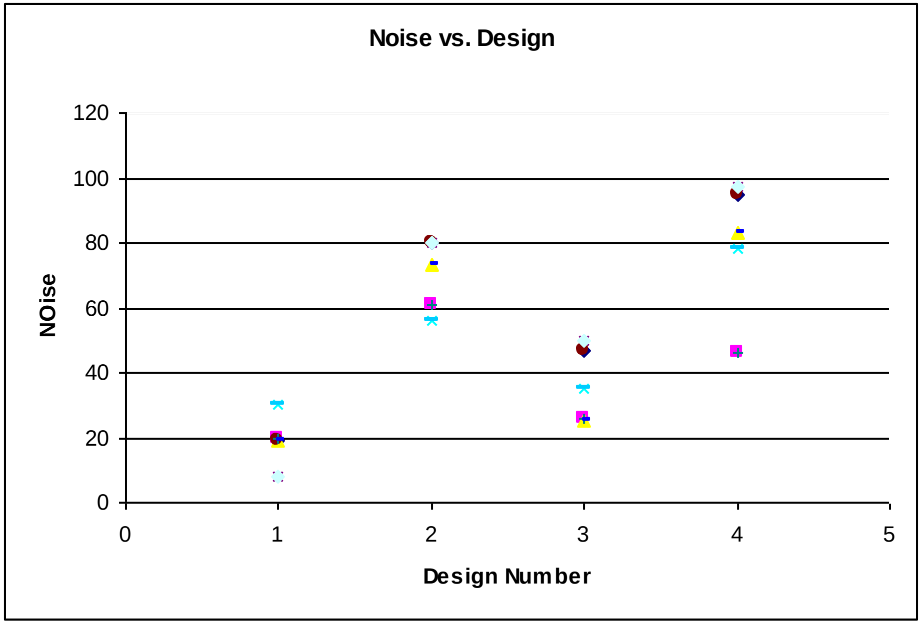
Tests perfromed on orthogonal contrasts are independent. For a treatments, a set of a-1 orthogonal contrasts will sum to $SS_{Treatments}$.

- Orthogonal contrast is useful for preplanned comparisons. Example see page 95.
- Data snooping: Scheffe's method

Example – Problem 3-24

Noise Observed from Different Circuit Designs

	Replication				
Design	**1**	**2**	**3**	**4**	**5**
D1	19	20	19	30	8
D2	80	61	73	56	80
D3	47	26	25	35	50
D4	95	46	83	78	97



Example – Problem 3-24

Anova: Single Factor

SUMMARY

Groups	*Count*	*Sum*	*Average*	*Variance*
D1	5	96	19.2	60.7
D2	5	350	70	121.5
D3	5	183	36.6	134.3
D4	5	399	79.8	420.7

ANOVA

Source of Variation	*SS*	*df*	*MS*	*F*	*P-value*	*F crit(0.05)*
Between Groups	12042	3	4014	21.77971	6.8E-06	3.2388715
Within Groups	2948.8	16	184.3			
Total	14990.8	19				

Example – Problem 3-24

- Compare noise from D1+D4 to D2+D3.

$$c_1 = 1, c_2 = -1, c_3 = -1, c_4 = 1$$

$$t_0 = \frac{\sum_{i=1}^{a} c_i \bar{y}_{i.}}{\sqrt{\frac{MS_E}{n} \sum_{i=1}^{a} c_i^2}} = \frac{19.2 - 70 - 36.6 + 79.8}{\sqrt{\frac{184.3}{5}(1+1+1+1)}} = -.626$$

$$t_{0.025,16} = 2.12 \Rightarrow \text{Do not reject.}$$

Example – Problem 3-24

- Compare noise from D1+D4 to D2+D3.

$$F_0 = t_0 = (-.626)^2 = 0.392$$

$$F_0 = \frac{MS_C}{MS_E} = \frac{SS_C/1}{MS_E} = \frac{\left(\sum_{i=1}^{a} c_i \bar{y}_{i.}\right)^2}{\frac{1}{n}\sum_{i=1}^{a} c_i^2} \div MS_E$$

$$= \frac{(19.2 - 70 - 36.6 + 79.8)^2}{4/5} \div 184.3 = 0.392$$

$$F_{0.05,1,16} = 4.49 \Rightarrow \text{Do not reject.}$$

Example – Problem 3-24

- Orthogonal contrasts

	Contrast Coefficient		
Design	**1**	**2**	**3**
D1	1	0	1
D2	0	1	-1
D3	-1	0	1
D4	0	-1	-1

Example – Problem 3-24

$$SS_{C1} = \frac{(19.2 - 36.6)^2}{2/5} = 756.9$$

$$SS_{C2} = \frac{(70 - 79.8)^2}{2/5} = 240.1$$

$$SS_{C3} = \frac{(19.2 - 70 + 36.6 - 79.8)^2}{4/5} = 11{,}045.0$$

$$SS_{C1} + SS_{C2} + SS_{C2} = 12{,}042 = SS_{Treatments}$$

Example – Problem 3-24

ANOVA

Source of Variation	*SS*	*df*	*MS*	*F*	*P-value*	*F crit(0.05)*
Treatments	12042	3	4014	21.77971	6.8E-06	3.2388715
C1	756.9	1	756.9	4.106891	0.059713	4.4939984
C2	240.1	1	240.1	1.302767	0.270504	4.4939984
C3	11045	1	11045	59.92946	8.47E-07	4.4939984
Error	2948.8	16	184.3			
Total	14990.8	19				

Scheffé's Method

- A method for hypothesis testing or forming confidence intervals for all possible contrasts among factor level means.
  - Confidence levels are applicable to infinitely many contrasts.
  - Applies to equal and unequal sample sizes.

Scheffé's Method

Consider m contrasts among factor level means

Contrast $u, 1 \le u \le m: \Gamma_u = c_{1u}\mu_1 + \cdots + c_{au}\mu_a$

The estimator of Γ_u is $C_u = c_{1u}\bar{y}_{1.} + \cdots + c_{au}\bar{y}_{a.}$

The null hypothesis is that the contrasts equal zero.

The estimator of the standard deviation of contrast u is

$$S_{C_u} = \sqrt{MS_E \sum_{i=1}^{a} \frac{c_{iu}^2}{n_i}} \text{ where } n_i = \text{\# of replications in the } i\text{th treatment.}$$

The critical value is $S_{\alpha,u} = S_{C_u}\sqrt{(a-1)F_{\alpha,a-1,N-a}}$

If $|C_u| > S_{\alpha,u}$ reject the null hypothesis

Example – Problem 3-24

- Orthogonal contrasts

	Contrast Coefficient		
Design	**1**	**2**	**3**
D1	1	0	1
D2	0	1	-1
D3	-1	0	1
D4	0	-1	-1

Noise Observed from Different Circuit Designs

	Replication					
Design	**1**	**2**	**3**	**4**	**5**	$y_{i.}$
D1	19	20	19	30	8	19.2
D2	80	61	73	56	80	70
D3	47	26	25	35	50	36.6
D4	95	46	83	78	97	79.8

Example – Problem 3-24

$$C_1 = 19.2 - 36.6 = -17.4$$

$$C_2 = 70 - 79.8 = -9.8$$

$$C_3 = 19.2 - 70 + 36.6 - 79.8 = -94$$

$$S_{C_1} = \sqrt{184.3 * (1/5 + 1/5)} = 8.59$$

$$S_{C_2} = \sqrt{184.3 * (1/5 + 1/5)} = 8.59$$

$$S_{C_3} = \sqrt{184.3 * (1/5 + 1/5 + 1/5 + 1/5)} = 12.14$$

$$S_{0.05,1} = 8.59\sqrt{3 * 3.239} = 26.78$$

$$S_{0.05,2} = 8.59\sqrt{3 * 3.239} = 26.78$$

$$S_{0.05,3} = 12.14\sqrt{3 * 3.239} = 37.84$$

Comparison of Prior Methods

- What method should be used?
  - If specific contrasts (orthogonal) are of interest use the t-test or equivalent F-test.
- Compare confidence interval half-width.
  - Difference in two means.
    - Scheffé’s method = 26.78.
    - t-confidence interval = 18.20

Comparing Pairs of Treatment Means

- **Tukey's test**

If only the pairwise comparison of treatment means is equal, specific test procedures have been developed.

The null hypothesis is:

$$H_0 : \mu_i = \mu_j \text{ for all } i \neq j.$$

Tukey's test is a procedure that controls confidence level for all pairwise comparisons.

Comparing Pairs of Treatment Means

The critical value for comparison is constructed from the distribution of the studentized range statistic:

$$q = \frac{\bar{y}_{\max} - \bar{y}_{\min}}{\sqrt{MS_E / n}}$$

The distribution of this statistic *q(a,f)* has two parameters.

a = Number of treatment means.
f = Degrees of freedom associated with MS_E.

Comparing Pairs of Treatment Means

The critical value is

$$T_\alpha = q_\alpha(a, f)\sqrt{\frac{MS_E}{n}}$$

Reject the hypothesis that two means are equal if the estimated absolute value of their difference exceeds the critical value.

Problem 3-24

- Compare all pairs of treatment means using the Tukey test. Ask about the distribution value required.

Noise Observed from Different Circuit Designs

	Replication					
Design	**1**	**2**	**3**	**4**	**5**	**$y_{i.}$**
D1	19	20	19	30	8	19.2
D2	80	61	73	56	80	70
D3	47	26	25	35	50	36.6
D4	95	46	83	78	97	79.8

Example – Problem 3-24

The critical value is computed as

$$T_{0.05} = q_{0.05}(4,16)\sqrt{MS_E / n} = 4.05\sqrt{184.3/5} = 24.59$$

$$\left|\bar{y}_{1.} - \bar{y}_{2.}\right| = \left|19.2 - 70\right| = 50.8$$
$$\left|\bar{y}_{1.} - \bar{y}_{3.}\right| = \left|19.2 - 36.6\right| = 17.4$$
$$\left|\bar{y}_{1.} - \bar{y}_{4.}\right| = \left|19.2 - 79.8\right| = 60.6$$
$$\left|\bar{y}_{2.} - \bar{y}_{3.}\right| = \left|70 - 36.6\right| = 33.4$$
$$\left|\bar{y}_{2.} - \bar{y}_{4.}\right| = \left|70 - 79.8\right| = 9.8$$
$$\left|\bar{y}_{3.} - \bar{y}_{4.}\right| = \left|36.6 - 79.8\right| = 43.2$$

Other Procedures

- Fisher Least Significant Difference (LSD) method (read page 99)
  - Comparison of all treatment mean pairs.
- Dunnett's procedure for comparison of treatment means with a control. (read page 101)

Determining Sample Size

- Recall from before the two types of error.
  - Type I
  - Type II

Determining Sample Size

- In single factor ANOVA an F-test is used to evaluate the null hypothesis.

$$H_0 : \mu_1 = \mu_2 = \cdots = \mu_a$$

$$H_1 : \text{ At least one mean is different}$$

$$\beta = 1 - P(\text{Reject } H_0 \mid H_0 \text{ is false})$$

$$= 1 - P(F_0 > F_{\alpha, a-1, N} \mid H_0 \text{ is false})$$

What is the distribution of the test statistic if H_0 is false?

If H_0 is false then $F_0 = MS_{Treatments}/MS_E$ has a noncentral F distribution.

This distribution has three parameters: numerator and denominator degrees of freedom, and a noncentrality parameter δ.

Determining Sample Size

- Operating characteristic curves are used
  - α, β, Φ, and the denominator degrees of freedom are parameters of the curve.
  - Different curves are constructed for different numerator degrees of freedom (factor levels -1).
  - Read page 105-108

$$\Phi^2 = \frac{n\sum_{i=1}^{a}\tau_i^2}{a\sigma^2}$$

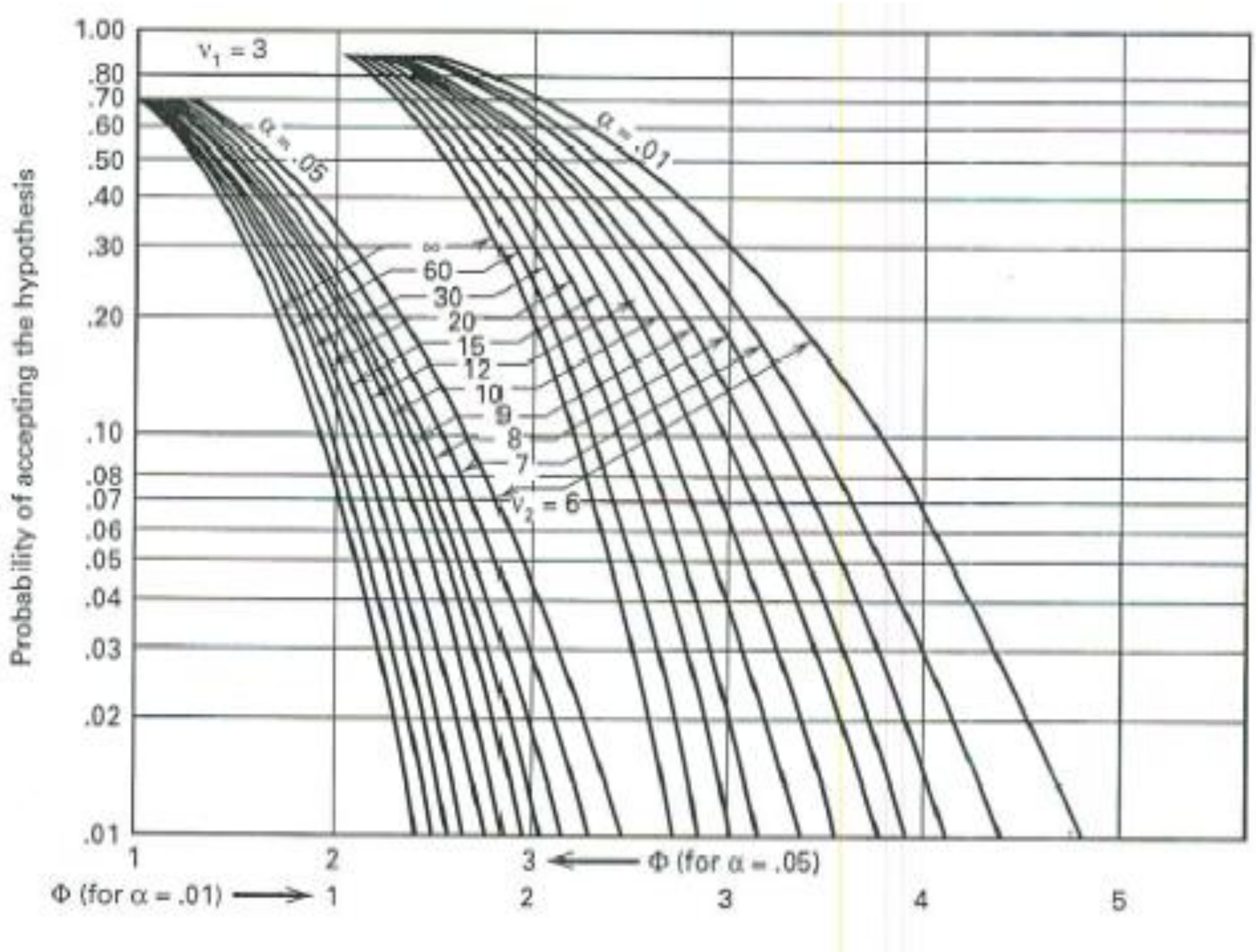
Example 3.10 (Part I)

- Consider the plasma etching experiment described before. Suppose that the experimenter is interested in rejecting the null hypothesis with a probability of at least 0.9 if the four treatment means are

$$\mu_1 = 575, \quad \mu_2 = 600, \quad \mu_3 = 650, \quad \mu_4 = 675$$

How many replications should be taken from each population to obtain a test with the required power? (use $\alpha = 0.01$ and σ=25)

Determining Sample Size



Example 3.10 (Part II)

- Consider the plasma etching experiment described before. Suppose that the experimenter is interested in rejecting the null hypothesis with a probability of at least 0.9 if any two treatment means differed by as much as 75 A/min.

How many replications should be taken from each population to obtain a test with the required power? (use $\alpha = 0.01$ and $\sigma=25$)

Example 3.10 (Part III)

- Consider the plasma etching experiment described before. Suppose that we want a 95 percent confidence interval on the difference in mean etch rate for any two power settings to be ± 30

How many replications should be taken from each population to obtain the desired accuracy? (use $\alpha = 0.01$ and σ=25)

Fixed Factors V.S. Random Factors

- Previous chapters have focused primarily on **fixed** factors
  - A specific set of factor levels is chosen for the experiment
  - Statistical inference made about this factors are confined to the specific levels studies
  - For example, if three material types are investigated as in the battery life experiment of Example 5.1, our conclusions are valid only about those specific material types.
- When factor levels are chosen at random from a larger population of potential levels, the factor is **random**
  - Sometimes, the factor levels are chosen at random from a larger population of possible levels
  - The experimenter wishes to draw conclusions about the entire population of levels, not just those that were used in the experimental design.

How to deal with random factors?

- A Single Random Factor Model
- Two-Factor Factorial with Random Factors
- Two-Factor Mixed Model
- Two-Stage Nested Design
- Split-Plot Design

A Single Random Factor Model

The linear statistical model is

$$y_{ij} = \mu + \tau_i + \epsilon_{ij} \begin{cases} i = 1, 2, \ldots, a \\ j = 1, 2, \ldots, n \end{cases} \tag{3.47}$$

where both the treatment effects τ_i and ϵ_{ij} are random variables. We will assume that the treatment effects τ_i are NID $(0, \sigma_\tau^2)$ random variables[1] and that the errors are NID $(0, \sigma^2)$, random variables, and that the τ_i and ϵ_{ij} are independent. Because τ_i is independent of ϵ_{ij}, the variance of any observation is

$$V(y_{ij}) = \sigma_\tau^2 + \sigma^2$$

↑

Variance components

The basic ANOVA sum of squares identity

$$SS_T = SS_{\text{Treatments}} + SS_E \tag{3.48}$$

is still valid. That is, we partition the total variability in the observations into a component that measures the variation between treatments ($SS_{\text{Treatments}}$) and a component that measures the variation within treatments (SS_E). Testing hypotheses about individual treatment effects is not very meaningful because they were selected randomly, we are more interested in the **population** of treatments, so we test hypotheses about the variance component σ_τ^2.

$$\begin{aligned} H_0&: \sigma_\tau^2 = 0 \\ H_1&: \sigma_\tau^2 > 0 \end{aligned} \tag{3.49}$$

If $\sigma_\tau^2 = 0$, all treatments are identical; but if $\sigma_\tau^2 = 0$, variability exists between treatments. As before, SS_E/σ^2 is distributed as chi-square with $N - a$ degrees of freedom and, under the null hypothesis, $SS_{\text{Treatments}}/\sigma^2$ is distributed as chi-square with $a - 1$ degrees of freedom. Both random variables are independent. Thus, under the null hypothesis $\sigma_\tau^2 = 0$, the ratio

$$F_0 = \frac{\dfrac{SS_{\text{Treatments}}}{a-1}}{\dfrac{SS_E}{N-a}} = \frac{MS_{\text{Treatments}}}{MS_E} \tag{3.50}$$

is distributed as F with $a - 1$ and $N - a$ degrees of freedom. However, we need to examine the expected mean squares to fully describe the test procedure.

Estimating the variance components using the ANOVA method:

$$MS_{\text{Treatments}} = \sigma^2 + n\sigma_\tau^2$$

$$MS_E = \sigma^2$$

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

$$\hat{\sigma}_\tau^2 = \frac{MS_{\text{Treatments}} - MS_E}{n}$$

EXAMPLE 3.11

A textile company weaves a fabric on a large number of looms. It would like the looms to be homogeneous so that it obtains a fabric of uniform strength. The process engineer suspects that, in addition to the usual variation in strength within samples of fabric from the same loom, there may also be significant variations in strength between looms. To investigate this, she selects four looms at random and makes four strength determinations on the fabric manufactured on each loom. This experiment is run in random order, and the data obtained are shown in Table 3.17. The ANOVA is conducted and is shown in Table 3.18. From the ANOVA, we conclude that the looms in the plant differ significantly.

■ **TABLE 3.17**
Strength Data for Example 3.11

	Observations				
Looms	**1**	**2**	**3**	**4**	$y_{i.}$
1	98	97	99	96	390
2	91	90	93	92	366
3	96	95	97	95	383
4	95	96	99	98	388
					$1527 = y_{..}$

The variance components are estimated by $\hat{\sigma}^2 = 1.90$ and

$$\hat{\sigma}_\tau^2 = \frac{29.73 - 1.90}{4} = 6.96$$

Therefore, the variance of any observation on strength is estimated by

$$\hat{\sigma}_y = \hat{\sigma}^2 + \hat{\sigma}_\tau^2 = 1.90 + 6.96 = 8.86.$$

Most of this variability is attributable to differences *between* looms.

■ **TABLE 3.18**
Analysis of Variance for the Strength Data

Source of Variation	Sum of Squares	Degrees of Freedom	Mean Square	F_0	*P*-Value
Looms	89.19	3	29.73	15.68	<0.001
Error	22.75	12	1.90		
Total	111.94	15			