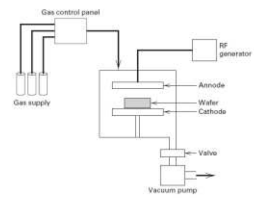
Analysis of Variance (ANOVA)

- Suppose we want to compare more than two means?

For example, suppose an engineer is investigating the relationship between the RF power setting and the etch rate for this 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.



■ FIGURE 3.1 A single-wafer plasma etching tool

If there are 2 different RF power settings (say, 160W and 180W), then a z-test or t-test is appropriate:

H_0: $\mu_1 = \mu_2$
H_1: $\mu_1 \neq \mu_2$

Comparing > 2 Means

- What if there are 3 different power settings (say, 160, 180, and 200 W)?

H_0: $\mu_1 = \mu_2$ *and* H_0: $\mu_1 = \mu_3$ *and* H_0: $\mu_2 = \mu_3$
H_1: $\mu_1 \neq \mu_2$ H_1: $\mu_1 \neq \mu_3$ H_1: $\mu_2 \neq \mu_3$

- How about 4 different settings (say, 160, 180, 200, and 220 W)?

All of the above, ***PLUS***

H_0: $\mu_1 = \mu_4$ *and* H_0: $\mu_2 = \mu_4$ *and* H_0: $\mu_3 = \mu_4$
H_1: $\mu_1 \neq \mu_4$ H_1: $\mu_2 \neq \mu_4$ H_1: $\mu_3 \neq \mu_4$

- What about 5 settings? 10?

Comparing > 2 means

- Also, suppose $\alpha = 0.05$
  - $(1 - \alpha) = P(\text{accept } H_0 \mid H_0 \text{ is true}) = 0.95$
    - 4 settings: $(0.95)^4 = 0.814$
    - 5 settings: ____________
    - 10 settings: ____________
- Instead, use Analysis of Variance (ANOVA)
  - *treatment, factor, independent variable*: that which is varied (*a* levels)
  - *observation, response, dependent variable*: the result of concern (*n* per treatment)
  - *randomization*: performing experimental runs in random order so that other factors don't influence results.

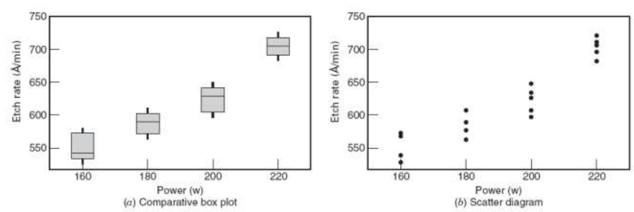
Randomizing the data collection

- Do we need to *stratify* the data collection?
  - Is it physically possible and feasible to run in random order?
  - If not, how do we manage our strata so as to avoid introducing variability that we don't intend?
    - blocking?
- Use Excel or Minitab to determine a random order for our data, e.g. … (example)

Our example (See pg. 62)

■ 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



■ FIGURE 3.2 Box plots and scatter diagram of the etch rate data

There are a couple of basic questions we'd like to answer …

- Does changing the power change the mean etch rate?
- Is there an optimum level for power?

We would like to have an objective way to answer these questions

The Analysis of Variance (Sec. 3.2, pg. 62)

■ TABLE 3.2
Typical Data for a Single-Factor Experiment

Treatment (Level)	Observations				Totals	Averages
1	y_{11}	y_{12}	...	y_{1n}	$y_{1.}$	$\bar{y}_{1.}$
2	y_{21}	y_{22}	...	y_{2n}	$y_{2.}$	$\bar{y}_{2.}$
⋮	⋮	⋮	⋮⋮⋮	⋮	⋮	⋮
a	y_{a1}	y_{a2}	...	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* = *an* total runs
- We consider the *fixed effects* case…the *random effects* case will be discussed later
- Our objective is to test hypotheses about the equality of the a treatment means

What does it mean?

- 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 basic single-factor ANOVA model is

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

$$\mu = \text{an overall mean}, \ \tau_i = i^{th} \text{ treatment effect},$$

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

Models for the data

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

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

$$\text{Let } \mu_i = \mu + \tau_i, \text{ then}$$

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

Regression models can also be employed

Variability is measured by sums of squares …

- Total *sums of squares* is partitioned as …

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

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

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

- 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

Variability is measured by sums of squares …

- Formal statistical hypotheses are:

$$H_0 : \mu_1 = \mu_2 = \cdots = \mu_a$$
$$H_1 : \text{ At least one mean is different}$$

- 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, so …

$df_{treat} = a - 1 =$ _____

$df_E = a(n - 1) =$ _____

$df_{total} = an - 1 =$ _____

Determining the Difference

- Mean Square, $MS = SS/df$

$MS_{treat} =$ ____________

$MS_E =$ ____________

- 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.
- The statistic is: $F_0 = Ms_{treat}/MS_E =$ ____________

The ANOVA is summarized in a 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

  ____

ANOVA table

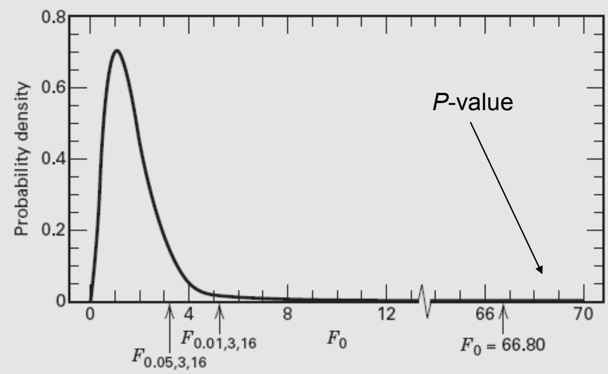
- For our example, this looks like …

■ **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			

- How do we interpret this?

The reference distribution:

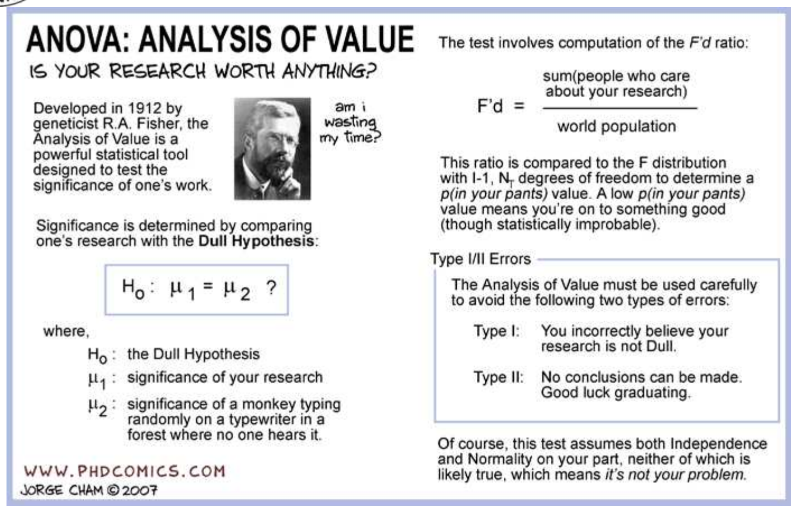


■ **FIGURE 3.3** The reference distribution ($F_{3,16}$) for the test statistic F_0 in Example 3.1

Try this on Excel and Minitab

- Excel …
- Minitab …

A (very) little humor…



Which means are different?

- Graphical methods
  - Box plots
  - Dot diagrams
  - etc.
- Numerical methods
  - Tukey's test (available on Minitab)
  - Duncan's Multiple Range test
  - Fisher's LSD

Model adequacy checking in the ANOVA

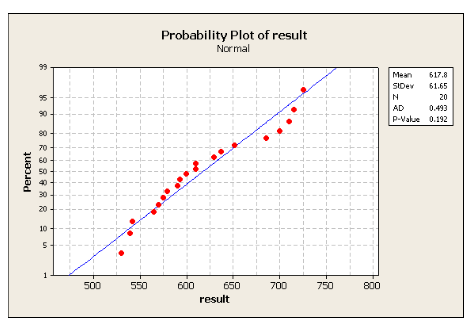
- Checking assumptions is important
  - Normality
  - Constant variance
  - Independence
  - Have we fit the right model?
- Later we will talk about what to do if some of these assumptions are violated
- For more, see section 3.4, pg. 75

Model adequacy checking in the ANOVA

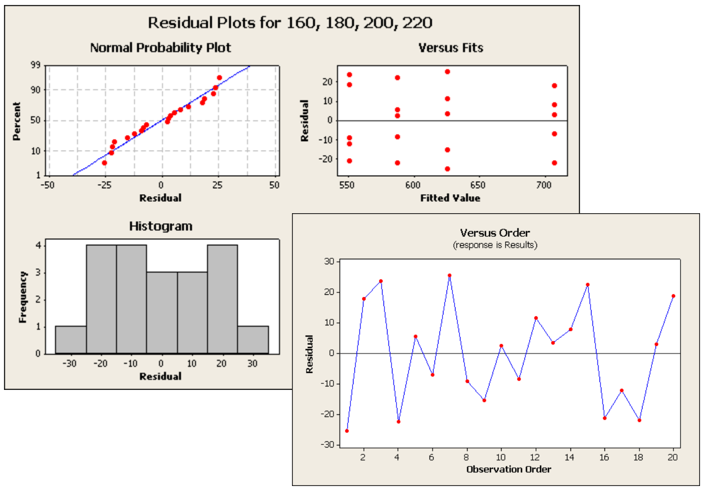
- Examination of **residuals** (see text, Sec. 3-4, pg. 75)

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

- Computer software generates the residuals
- **Residual plots** are very useful
- **Normal probability plot** of residuals



Other important residual plots

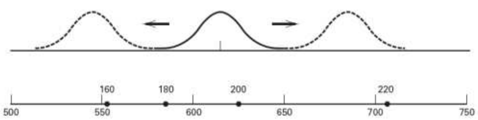


Post-ANOVA comparison of means

- The analysis of variance tests the null hypothesis of equal treatment means
  - Assume that residual analysis is satisfactory
  - If the null hypothesis is rejected, we don't know which specific means are different
- Determining which specific means differ following an ANOVA is called the multiple comparisons problem
  - There are lots of ways to do this…see text, Section 3.5, pg. 84
  - We will use a pairwise t-tests on means
    - Tukey's Method
    - Fisher's Least Significant Difference (or Fisher's LSD) Method -

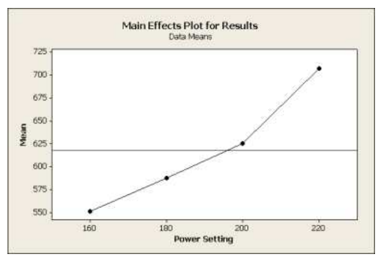
Graphical comparison of means

- From text, pg. 86

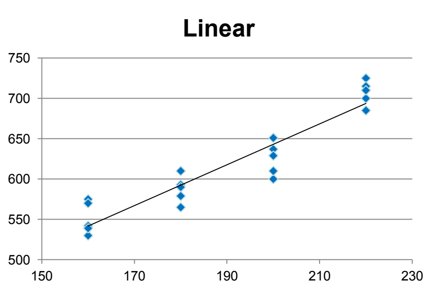


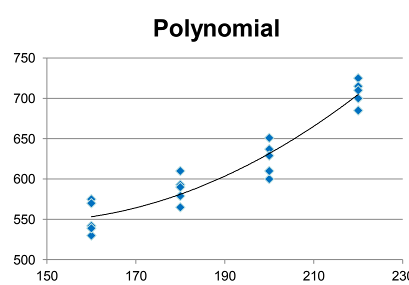
■ FIGURE 3.11 Etch rate averages from Example 3.1 in relation to a t distribution with scale factor $\sqrt{MS_E/n} = \sqrt{330.70/5} = 8.13$

- From Minitab



The regression model





$\hat{y} = 137.62 + 2.527x$

$\hat{y} = 1147.77 - 8.2555\,x + 0.028375\,x^2$

Determining the sample size

- A *FAQ* in designed experiments
- Answer depends on lots of things; including what type of experiment is being contemplated, how it will be conducted, resources, and desired *sensitivity*
- Sensitivity refers to the *difference in means* that the experimenter wishes to detect
- Generally, increasing the number of replications increases the sensitivity or it makes it easier to detect small differences in means

Determining the sample size: Fixed effects case

- Can choose the sample size to detect a specific difference in means and achieve desired values of type I and type II errors
- Type I error – reject H_0 when it is true (α)
- Type II error – fail to reject H_0 when it is false (β)
- *Power* = 1 - β
- *Operating characteristic curves* plot β against a parameter Φ where

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

OC curves to determine sample size: fixed effects case

- The *OC curves* for the fixed effects model are in the Appendix, Table V
- A very common way to use these charts is to define a difference in two means *D* of interest, then the minimum value of Φ^2 is

$$\Phi^2 = \frac{nD^2}{2a\sigma^2}$$

- Typically, we work with the ratio of *D/σ* and try values of *n* until the *desired power* is achieved
- Most statistics software packages will perform power and sample size calculations (we'll use Minitab)
- There are some other methods discussed in the text

Minitab output

Power and Sample Size

One-way ANOVA

Alpha = 0.05 Assumed standard deviation = 25 Number of Levels = 4

SS Means	Sample Size	Target Power	Actual Power	Maximum Difference
2812.5	5	0.9	0.953578	75

The sample size is for each level.

Power and Sample Size

One-way ANOVA

Alpha = 0.01 Assumed standard deviation = 25 Number of Levels = 4

SS Means	Sample Size	Target Power	Actual Power	Maximum Difference
2812.5	6	0.9	0.915384	75

The sample size is for each level.

Homework

- Solve the following using Excel or Minitab (preferred)
  - 3.5
  - 3.10
  - 3.20
  - 3.36
- Bring your solutions (in electronic form) to class on Wednesday