ANALYSIS OF VARIANCE: Principles of Experimental Design:

1. Difference between *observational* studies and *experiments.*

* *Observational* Study – data is recorded without interfering with the course of the events
  + Opinion Polls.
  + Ecological studies.
* *Experiments* – researchers actively intervene to control the study conditions and record the responses.
  + *Factors* – value of the predictors. Purposely altered to measure resulting change in response.
    - *Levels* – predefined values for the factors.
    - *Treatment* – specific combination of the levels of factors.
  + *Experimental Units* (EU) – entities to which the treatments are applied.

1. Common designs:
   * *One-way design*:
     + One factor, k levels
     + Each level is a treatment
   * *Multi-way factorial* design:
     + Multiple factors.
     + Treatments consist of every possible combination of the levels of the factors.
   * *Block designs* – created by the researcher to help explain variation in the response.

3. *Experimental Error* – the treatments should help explain some of the total variability in the response. Excess variability is experimental error.

* Possible sources of experimental error:
  + Natural variation among EUs.
  + Variability in measurement response.
  + Failure to include the factors that affect the response.
  + Inability to exactly reproduce experimental conditions from one EU to another.

4. *Randomization* – key component of any designed experiment. Treatments and EUs are linked by random assignment. (Application of treatments is designed randomly)

* Small studies – randomly picking an EU and assign a treatment (#1, #2, etc.)
* Large studies – computer generated.

5. *Replication* – number of independent repetitions of a treatment.

* Without replication – can’t estimate the experimental error, or the variability of the treatment effects.

6. *Balance designs* – have the same number of experimental units (EU) in every treatment. Typically data sets are not balanced, thus you should start the experiments with a design that is balanced.

a. Example. Motor Oil Containers. Plastic containers for motor oil are blow molded in a machine that has two feeders, each feeding into three molding stations (for a total of six stations). Plastic is extruded through the feeders into continuous cylinders between two halves of each mold. As the molds close, the cylinders are pinched off and air is blown into them to form the container shapes. The production engineer is concerned that the molded containers at different stations are not the same weight. To investigate whether or not this is true, a random sample of 8 containers is taken from each station.

Components:

(a) One-way design:

* Single factor (station) with 6 levels.
* Six treatments (six stations).

(b) Experimental Unit – a container.

(c) Replications –

* Eight containers per station.
* 6 stations and 8 containers per station = 48 total.

(d) Response variable: weight of container.

Example: Cloth Dyeing Experiment. The quality control department of a fabric finishing plant is studying the effect of several factors on dyeing for a cotton cloth used to manufacture shirts. Two operators (1 or 2), three cycle times (40, 50, or 60 minutes) and two temperatures (300 or 350 deg) were selected, and three small specimen of cloth were dyed under each set of conditions. The finished cloth was compared to a standard, and a numerical score was assigned.

Components

(a) 3 way factorial experiment.

(b) Three factors:

* Operator, 2 levels.
* Cycle time, 3 levels.
* Temperature, 2 levels.

(c) Treatments: 2 x 3 x 2 = 12 treatments.

(d) Replications:

* 3 replications per treatment
* 3 x 12 = 36 observations

(e) Experimental Unit – cloth.

(f) Response variable – numeric score.

Quiz. Given the following scenarios, decide whether or not the data are the result of an observational study or an experiment. If the data are the result of an experiment, identify each of the key components: response variable, experimental unit, factor(s), treatment(s) and number of replications.

1. A spare part is manufactured by a production company once a month in lots that vary in size as demand fluctuates. A random sample of ten recent production runs are examined, and the lot size and man-hours are recorded. *This is an observational study. No attempt is made to control either the man-hours or the lot size.*

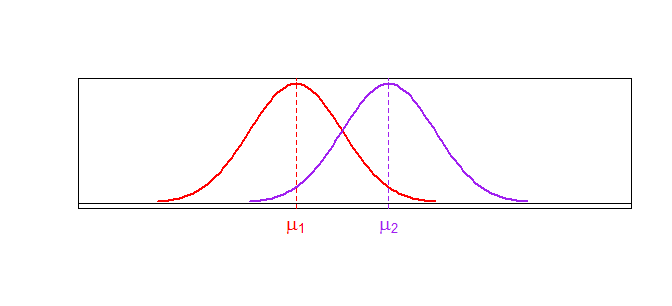
2. We wish to compare three brands of golf balls with respect to the distance they travel when hit by a mechanical driver. Five balls of brand A, five of brand B, and five of brand C are driven by the mechanical device in a random order and the distance for each ball is recorded. *This is an experiment: response variable: distance, experimental unit: golf ball, factor(s): brand, treatment(s): A, B, C and number of replications: 5.*

3. A substance used in biological and medical research is shipped by airfreight to users in cartons of 1,000 ampules. A random sample of 25 shipments was selected, and the shipping distance and the number of broken ampules were recorded. *This is an observational study. No attempt is made to control either the distance or the number of broken ampules*.

4. The data set consists of the number of pages in a manuscript and the dollar cost of correcting typographical errors in a random sample of recent orders handled by a firm specializing in technical manuscripts. *This is an observational study. The number of pages is observed (not controlled or randomly assigned).*

5. The data set contains information on 21 patients with high blood pressure, who were randomly assigned to one of three groups. One group is subjected to a very restrictive diet, another group undergoes a strict exercise program, and the third group continues their usual habits. After 6 months, the change in blood pressure is recorded. *This is an experiment: response variable: change in blood pressure, experimental unit: patient, factor(s): type of intervention, treatment(s): Diet, Exercise, or Neither and number of replications: 7.*

**ANALYSIS OF VARIANCE: Single Factor Studies.**



Test H0: μ1 = μ2 vs. Ha: μ1 ≠ μ2

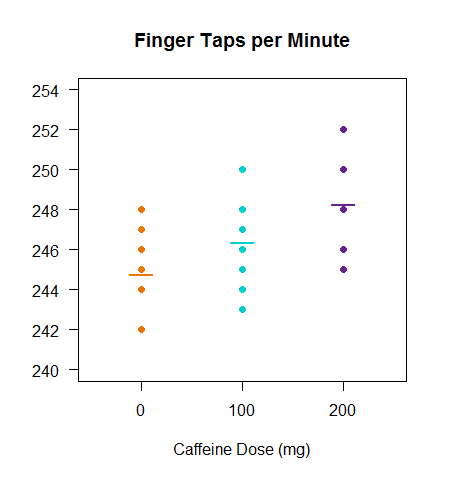
The above diagram is a two-sample t-test:

* Two groups (populations)
* Y is some measured characteristic
* For both groups, Y follows a normal distribution.
* For more than two groups?

1. Compare the Means of 2> Groups

* ANOVA Hypothesis: *comparing the mean* of more than 2 groups
  + H0: μ1 = μ2 = μ3 = . . . = μt vs. Ha: at least one mean is different.

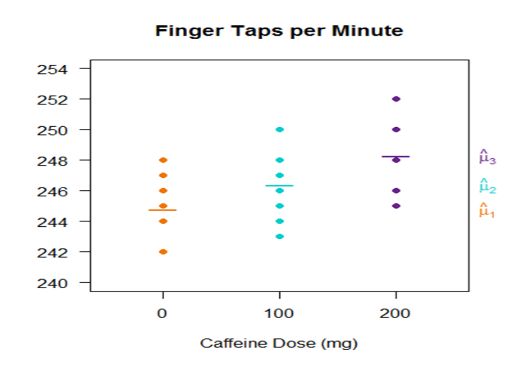
Within Group Variability



Each treatment group has a mean and variability around the mean.

If we combine the variability of these groups, we get the within group variability.

Between Group Variability



There is also variability among the means of the three groups.

This is the between group variability.



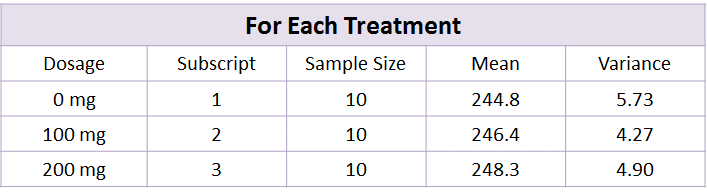
Table Layout:



observed values

sample means & variances

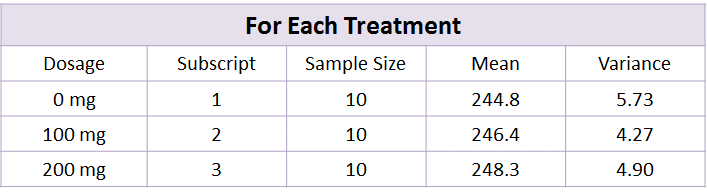
Calculate SSE





**SS Within = SSE = 134.1**

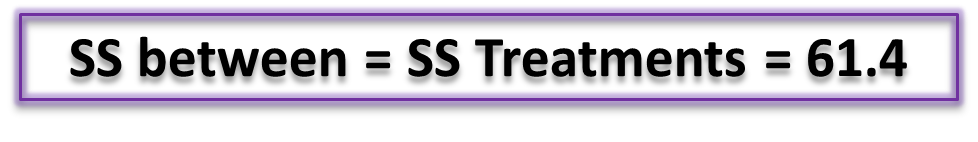
Calculate SS Treatments

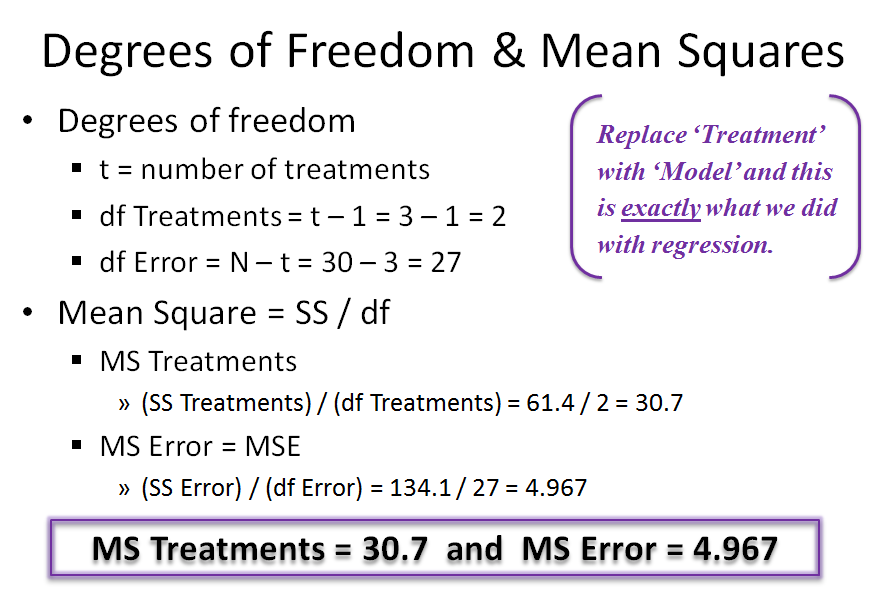


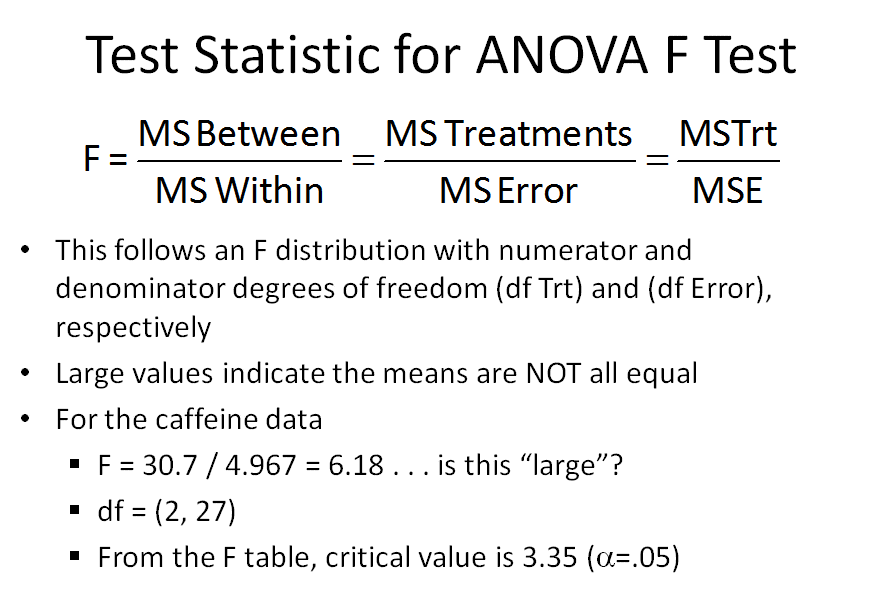


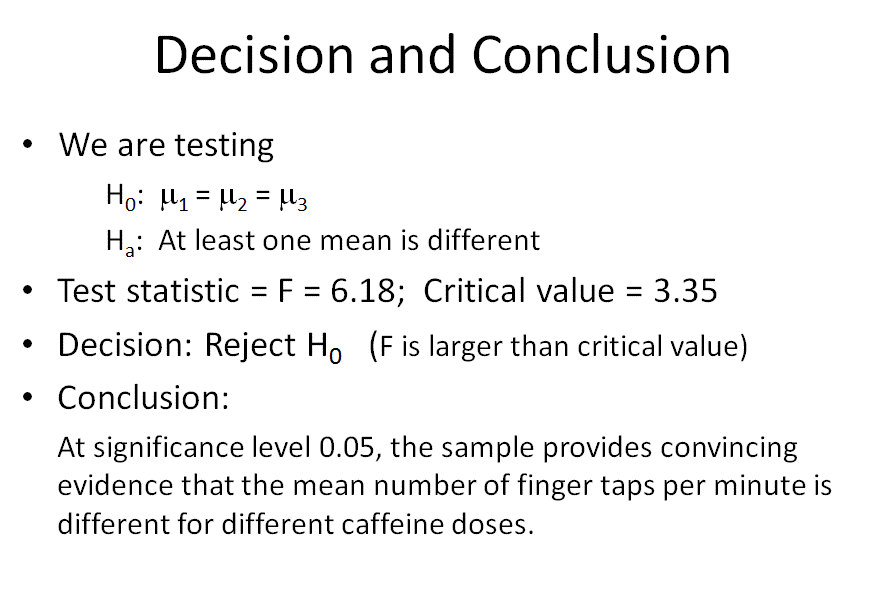
Total variation in the treatment means

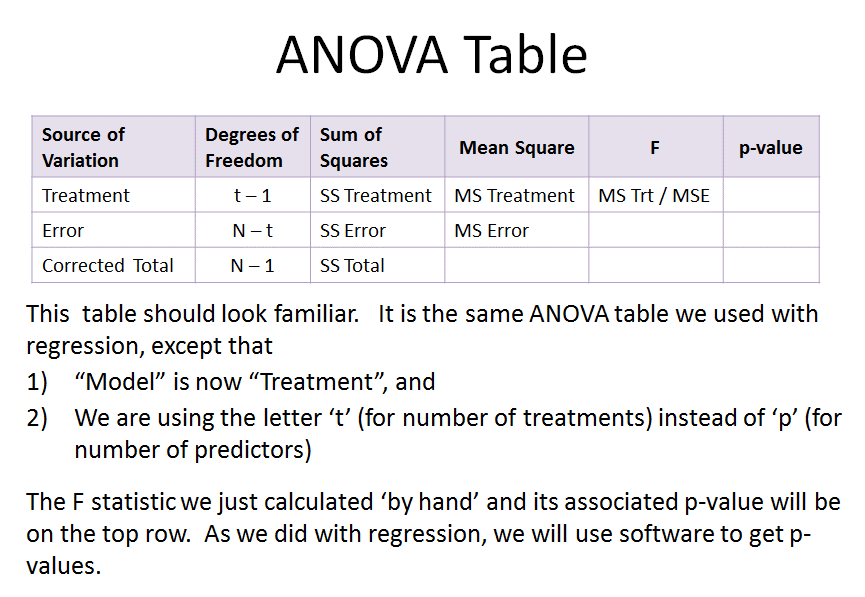












**ANALYSIS OF VARIANCE: LINEAR MODELS**

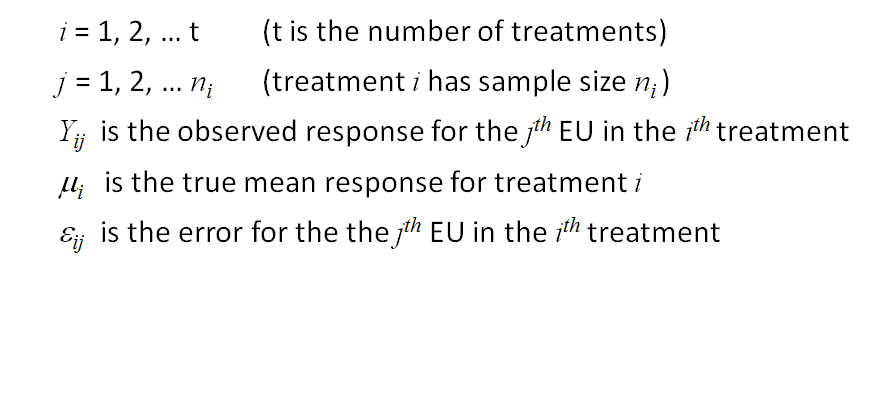
1. *Cell* means *model* – models the *mean response* for each treatment.

2. *Effects* model – models how the *mean* response for each treatment is *different* from the overall average response. This difference is called the *effect* of the treatment.

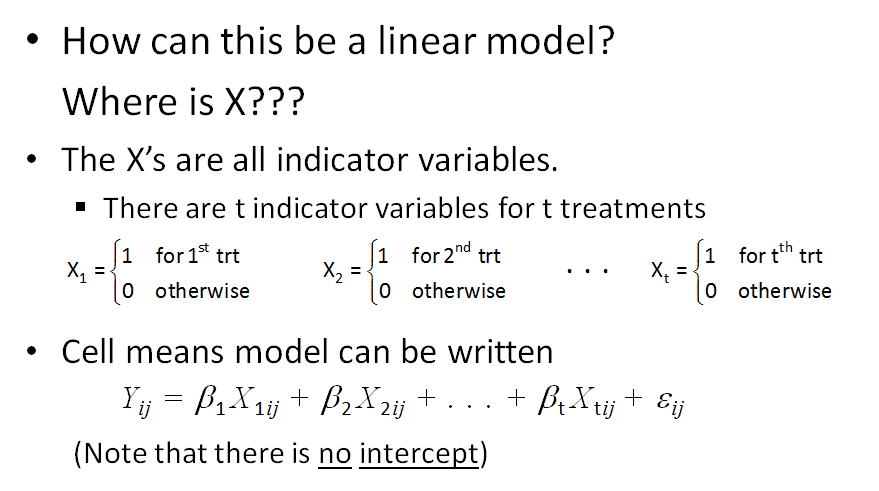
3. *Reference* treatment – a single pre-specified treatment that each treatment is compared to.

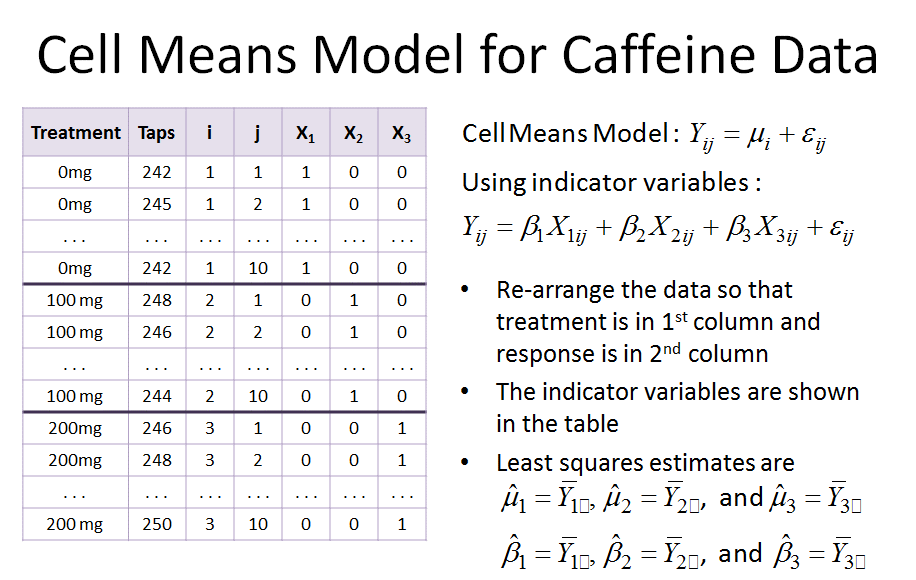
**Cell Means Model**

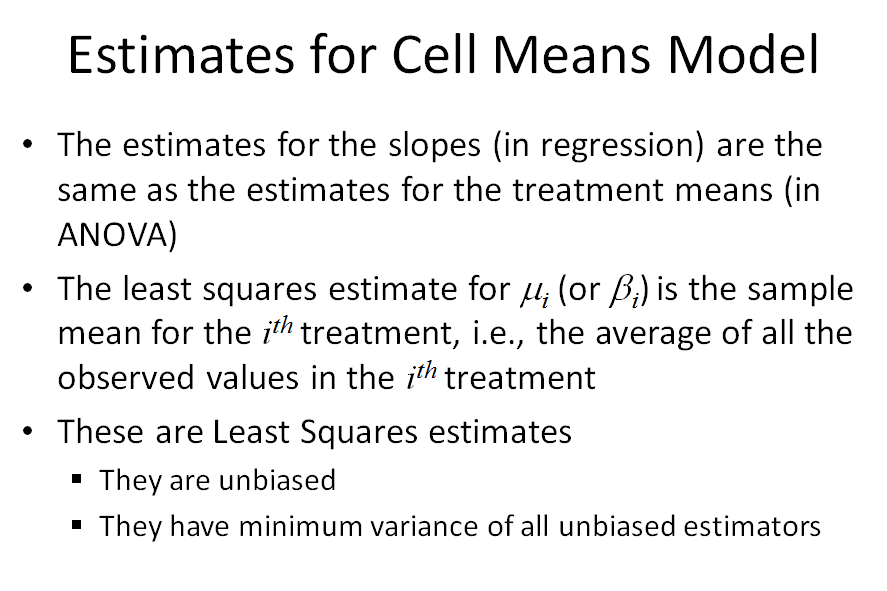


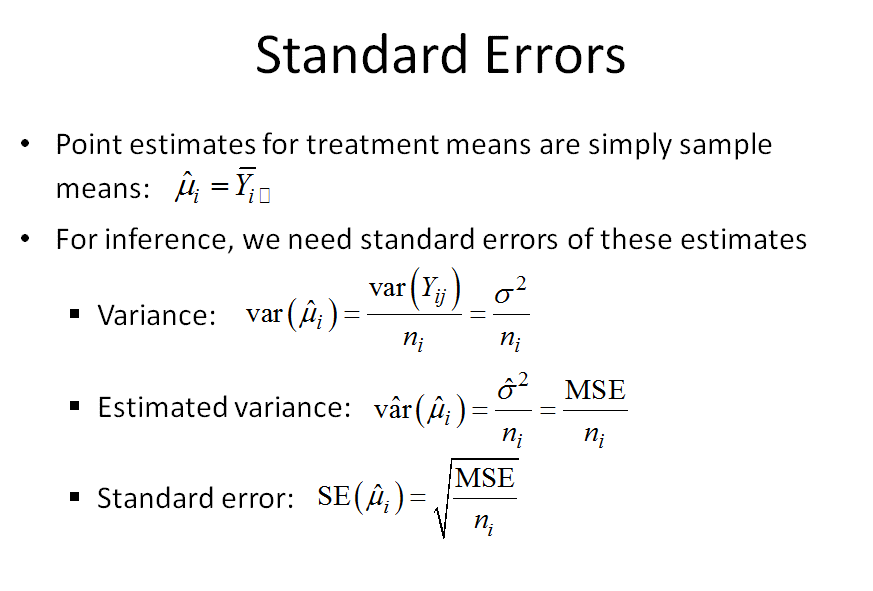


**ASSUME: ε*ij* ~ niid(0, σ2)**



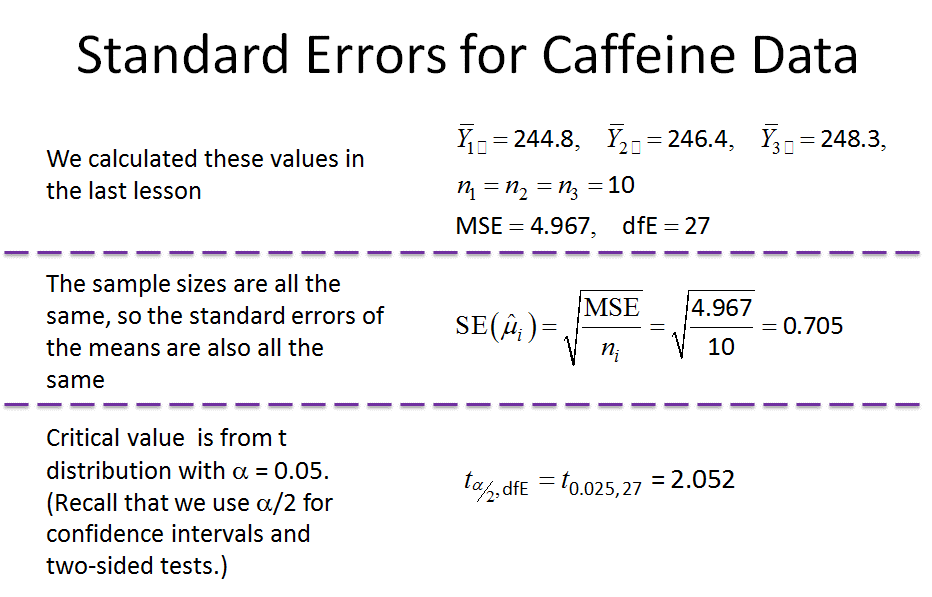






5. Comments on Standard Errors:

* Standard errors of the treatment means depend on the sample size for the treatment.
* If the treatments all have the same sample size (ie: caffeine example, *n1*=10, *n2*=10, *n3*=10)
  + Standard errors for the means will all be the same.
  + This is called “Balanced data”.
* If the treatments have different sample sizes
  + Treatment means will have different standard errors.
  + This is “unbalanced data”.

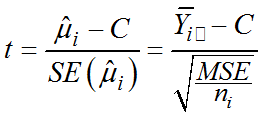


6. Confidence Intervals for Means

* Confidence Interval: (point estimate) ± (critical value) x SE
* For the caffeine data
  + Margin of error = (critical value) x SE = 2.052 x 0.705 = 1.447
  + Same margin of error for all the treatment means because the ***data are balanced.***
* 95% CI for the caffeine treatment means
  + Dose 0 mg: 244.8 ± 1.447, or (243.353, 246.247) finger taps.
  + Dose 100 mg: 246.4 ± 1.447, or (244.953, 247.847) finger taps.
  + Dose 200 mg: 248.3 ± 1.447, or (246.853, 249.747) finger taps.

7. Hypothesis Test for Means

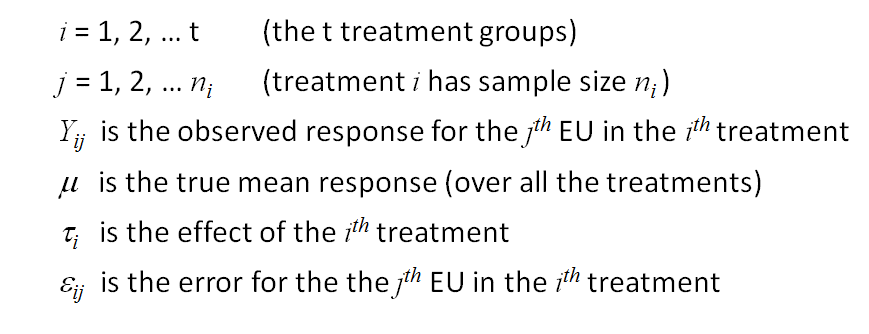
* For some constant C:
  + Test H0: μ*i* = C vs. Ha: μ*i* ≠ C
  + Test statistic:



* + Critical value is from *t* distribution, with df = df Error.
  + Reject Ho if |t| > critical value
  + Special case: C = 0
    - Then we are testing H0: μ*i* = 0 vs. Ha: μ*i* ≠ 0.

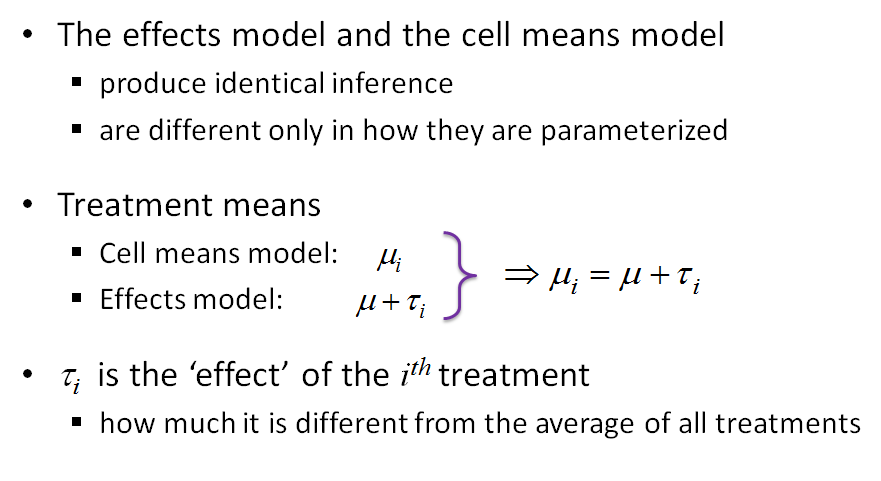
8. Effects Model:



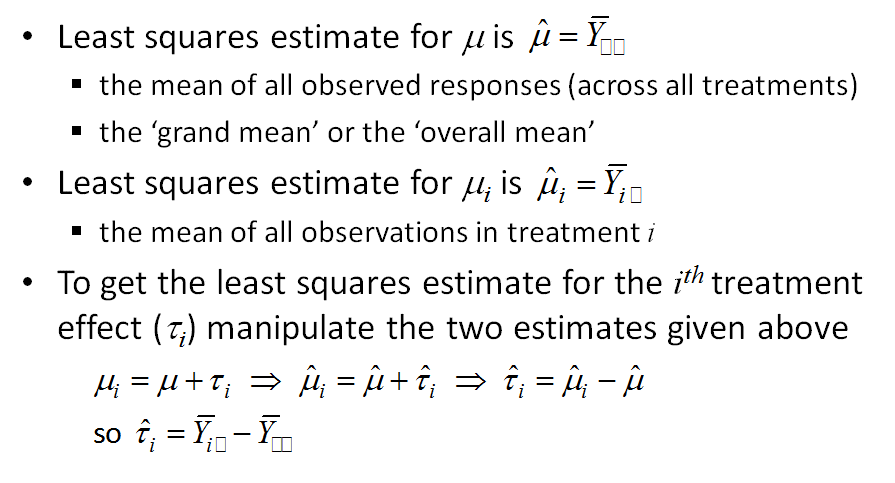




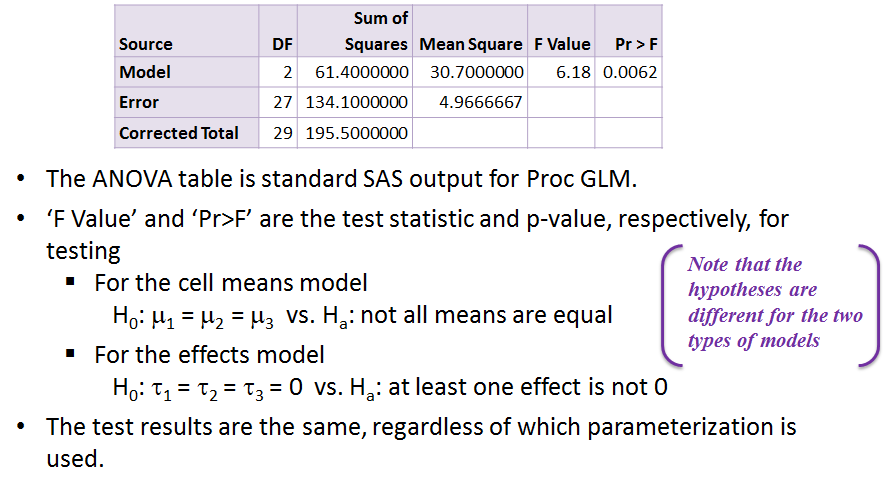
9. Parameterizations:



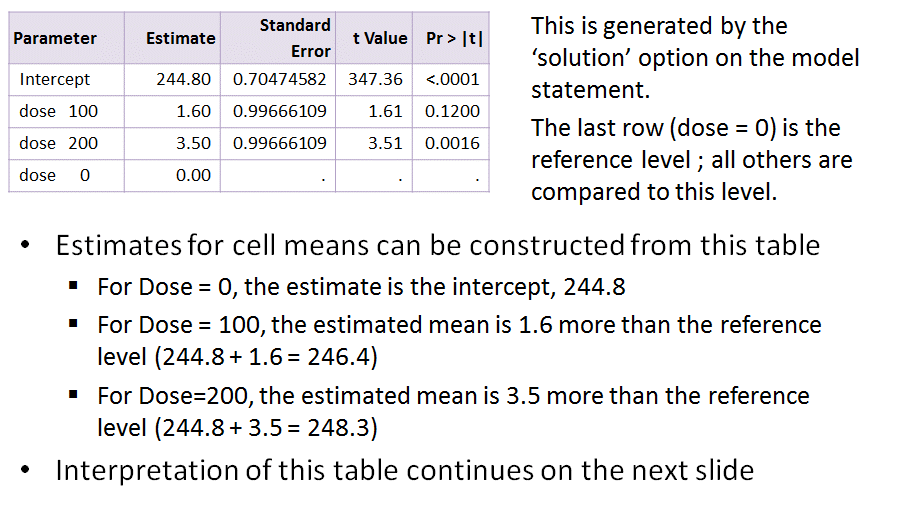
10. Estimation in Effects Model:



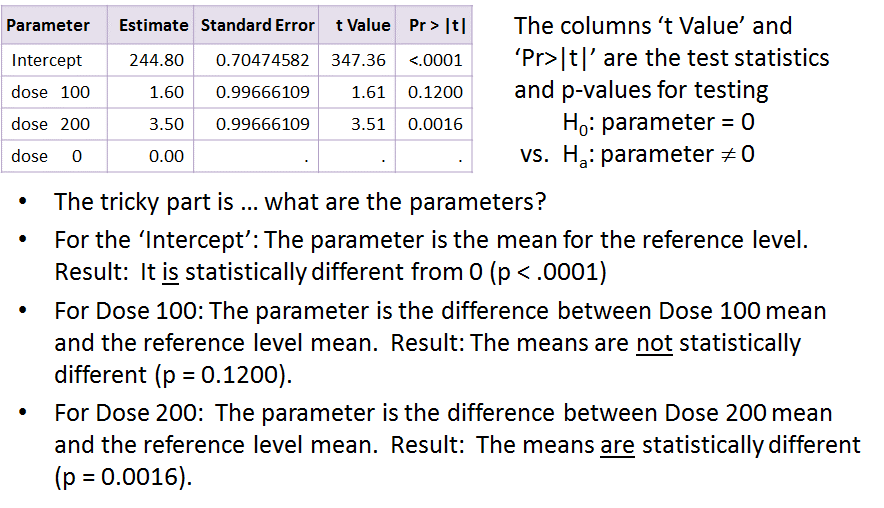
11. SAS Output: ANOVA Table



12. SAS Output: Effects Estimates



13. SAS Output: Hypothesis Tests



14. SAS Output:

