# Statistics

## Factorial ANOVA and Repeated Measures ANOVA

March 3, 2010

Jobayer Hossain, Ph.D. & Tim Bunnell, Ph.D.
*Nemours Bioinformatics Core Facility*

# Class Objectives -- You will Learn:

- Analysis of variance (ANOVA) for Factorial Experiment
- Analysis of variance (ANOVA) for Repeated Measures
- Data analysis in SPSS for Factorial and Repeated Measures ANOVA

# ANOVA of factorial experiment (Two or more factors)

- Factorial experiment:
  - The effects of the two or more factors including their interactions are investigated simultaneously.
  - For example, consider two factors A and B. Then total variation of the response will be split into variation for A, variation for B, variation for their interaction AB, and variation due to error.

# ANOVA of factorial experiment (Two or more factors)

- Model with two factors (A, B) and their interactions:

$$y_{ijk} = \mu + \alpha_i + \beta_j + (\alpha\beta)_{ij} + e_{ijk}$$

$\mu$ is the general mean

$\alpha_i$ is the effect of $ith$ level of the factor A

$\beta_j$ is the effect of $jth$ level of the factor B

$(\alpha\beta)_{ij}$ is the interaction effect of $ith$ level A and $jth$ level of B

$e_{ijk}$ is the error

- Assumptions: The same as in One-way ANOVA.

# ANOVA of factorial experiment (Two or more factors)

- Null Hypotheses:
- $H_{oa}$: Means of all groups of the factor A are equal.
- $H_{ob}$: Means of all groups of the factor B are equal.
- $H_{oab}$:$(\alpha\beta)_{ij} = 0$, i. e. two factors A and B are independent

# ANOVA of factorial experiment (Two or more factors)

- ANOVA for two factors A and B with their interaction AB.

| Sources of Variation | Sum of Squares | df | Mean Sum of Squares | F-Ratio |
|---|---|---|---|---|
| Main Effect A | SSA | k-1 | MSA=SSA/k-1 | MSA/MSE |
| Main Effect B | SSB | P-1 | MSB=SSB/p-1 | MSB/MSE |
| Interaction Effect AB | SSAB | (k-1)(p-1) | MSAB=SSAB/ (k-1)(p-1) | MSAB/MSE |
| Error | SSE | kp(r-1) | MSE=SSE/ kp(r-1) | |
| Total | SST | Kpr-1 | | |

# ANOVA of factorial experiment (Two or more factors) - SPSS Demonstration

Statistics -> General Linear Model-> Select dependent variable (e.g. change from PLUC), factors (Treatment group and gender) and covariate( PLUC_pre)

# Repeated Measures

- The term **repeated measures** refers to data sets with multiple measurements of a response variable on the same experimental unit or subject.
- In repeated measurements designs, we are often concerned with two types of variability:
  - **Between-subjects** - Variability associated with different groups of subjects who are treated differently (equivalent to between groups effects in one-way ANOVA)
  - **Within-subjects** - Variability associated with measurements made on an individual subject.

# Repeated Measures

- Examples of Repeated Measures designs:

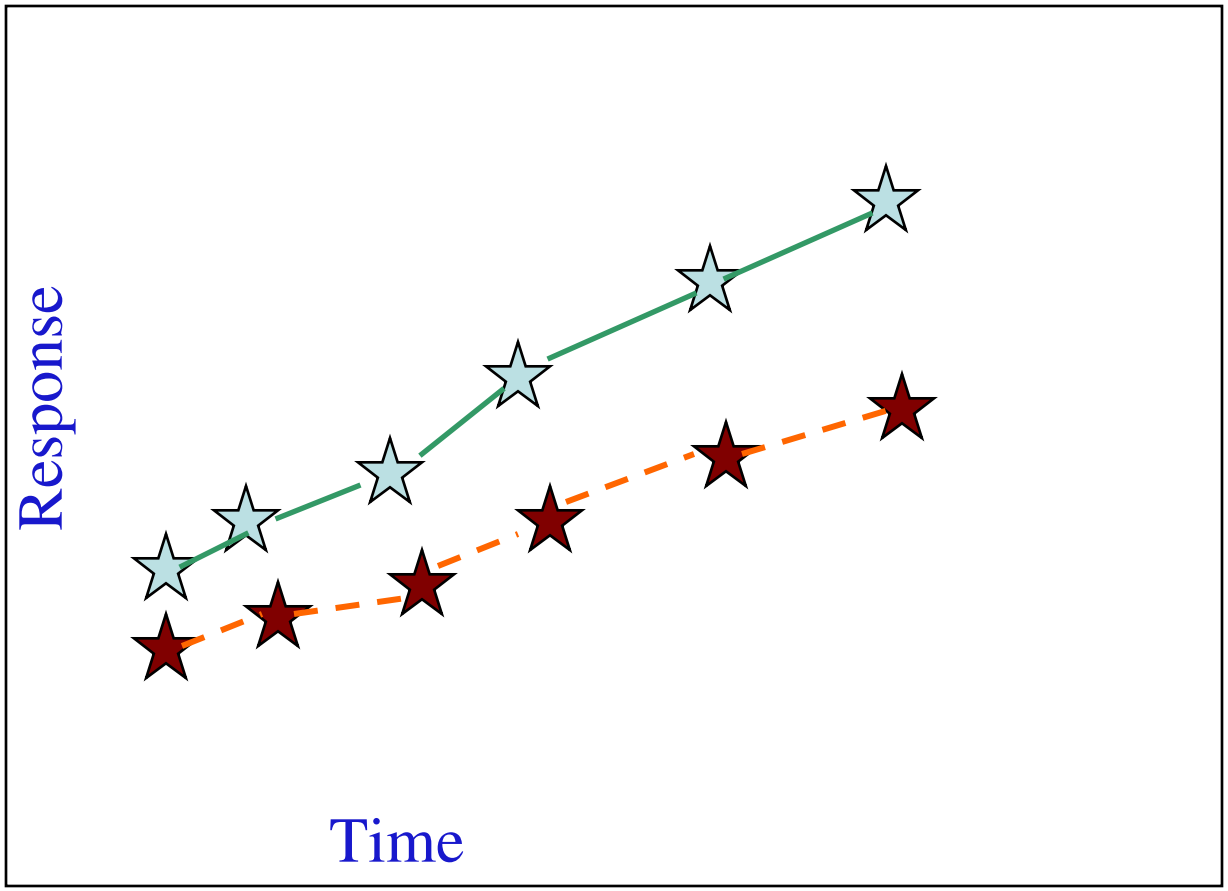
  A. Two groups of subjects treated with different drugs for whom responses are measured at six-hour increments for 24 hours. Here, DRUG treatment is the between-subjects factor and TIME is the within-subjects factor.

  B. In a weight loss study, 100 overwieght persons are to be given the same diet and their weights measured at the end of each week for 12 weeks to assess the weight loss time. Here subjects are observed repeatedly to assess the effect of diet.

# Repeated Measures

- When measures are made over time as in example A we want to assess:
  - how treatment effects differ at different times (i.e. the treatment by time interaction).
  - how the dependent measure changes over time independent of treatment (i.e. the main effect of time)
  - how treatments differ independent of time (i.e., the main effect of treatment)
- Repeated measures require special treatment because:
  - Observations made on the same subject are not independent of each other.
  - Adjacent observations in time are likely to be more correlated than non-adjacent observations

# Repeated Measures



# Repeated Measures

- Methods of repeated measures ANOVA
  - Univariate - Uses a single outcome measure.
  - Multivariate - Uses multiple outcome measures.
  - Mixed Model Analysis - One or more factors (other than subject) are random effects.
- We will discuss only univariate approach

# Repeated Measures

- Assumptions:
  - Subjects are independent.
  - The repeated observations for each subject follows a multivariate normal distribution
  - The correlation between any pair of within subjects levels are equal. This assumption is known as **sphericity**.

# Repeated Measures

- Test for Sphericity:
  - Mauchley's test
- Violation of sphericity assumption leads to inflated F statistics and hence inflated type I error.
- Three common corrections for violation of sphericity:
  - Greenhouse-Geisser correction
  - Huynh-Feldt correction
  - Lower Bound correction
- All these three methods adjust the degrees of freedom using a correction factor called Epsilon.
- Epsilon lies between 1/k-1 to 1, where k is the number of levels in the within subject factor.

# Repeated Measures: SPSS demonstration

- Analyze -> general Linear Model-> Repeated Measures -> define within subject factor name (e.g. time), number of levels ( i.e, number of times of dependent measurements e.g. we have two measurements of PLUC, so # of level is 2), click on Add, then select within subject variables chronologically (e.g. PLUC_pre and PLUC_post), select between subject factor (s) (e.g. grp), and select covariate (s) e.g. baseline age. Then select other output options.

# Thank you