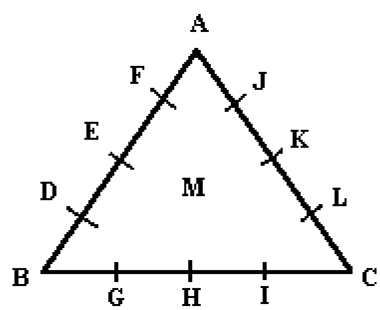
Psychology 420
Spring Semester 2000
Instructor: Dr. H. Lee

Lecture Notes:

Week 1 Lecture: See Course Syllabus for description of the course. Students will be expected to complete at the mastery level 7 assignments. If the assignments are done well, no final examination will be given. The assignments relate various analysis of variance and linear regression models to experimental design. A diagnostic test will be administered in the first class meeting to assess the student's knowledge of elementary statistics. The introductory lectures will be based upon the knowledge base of the students. Bruno's information reference testing method will be used for scoring the test. The use of this method assesses a large penalty for incorrect answers, but partial knowledge is rewarded with partial credit. Using the response triangle, test takers can indicate their degree of confidence in their chosen answers. The evaluation is formative, i.e. no grade is give.



IRT Response Triangle

Week 1: Lab: Students are introduce to the computer/statistical laboratory. The computer program used for the course will be SPSS (Statistical Package for the Social Sciences) 8.0 and 9.0. The essence of using statistical software is the data table or matrix. The data table consists of rows and columns. The rows represent the observations and the columns represent the variables. Those familiar with a teacher's roll book or spreadsheet already understand what a data table or matrix looks like. The observations are generally people and the variables are the different measures taken on each person.

When we enter SPSS for the first time, SPSS gives us an empty spreadsheet. This spreadsheet or data table looks like

	var	var	var	var	var	var
1						
2						
3						
4						
5						
6						
7						
8						
9						
10						

The observations are already numbered, but the variables remain undefined. In order to use SPSS properly, the variables need to be defined. One of the ways a user can define the variables is to use the mouse. Point the mouse to the first column with "var". Using the right button on the mouse click on the "var". From this you will get a menu. That menu will have an option "define variable." Select this and the result is another window where the variable can be named and defined. Each variable whether they are independent or dependent variables must be defined prior to their use in SPSS.

For this course students are expected to read the handout (or chapter) on Probability. Probability plays an important role in the study of statistics. One of the important concepts is the notion of a random variable. A random variable is one whose values occur at random. We don't know what value will come up, but we can still study this random process and make definite statements about occurrences. In order to do this, we study the notion of the expected value. Even though we cannot say if the toss of the coin will produce a "head" or a "tail" exactly, we can make a probability statement, such as "there is a 50% of a "head". Or if we toss a fair coin 10 times, we would expect 5 heads or 5 tails. Probabilities range from 0 to 1.00. In the study of probability, the instructor has pointed out that human decision making behavior most closely follows Bayes' Theorem and Keynes Weight of Evidence. Bayes' Theorem states that a decision is based upon the current information influence on past knowledge or probabilities. That is, the minute-by-minute decision making is based on the most current information and past information. A person may be inclined to buy a new Honda, but at the last moment a new reports indicates that they have defects, the buyer would decide not to buy. Keynes' contribution concerns the confidence humans place on the most current information. If the current information is consistent with past information, the information base is increased. If the information is contradictory, the decision-maker may decide to ignore the new information. However, in cases where there is very little past information or prior experience, the decision maker may place little confidence in the current information. Up until the Chernobyl nuclear accident, experts knew very little about what would happen with a nuclear meltdown. There were guesses, but little confidence was given to those estimates. When people venture into areas with little past information, every new bit of information would alter the information used in decision making.

Such studies using probabilities are important. When we consider statistical analysis, probability enters into the discussion. Expected values are examined in details.

Week 2: Lecture

There will be no lab this week. The results of the diagnostic test are returned to the student. Be aware that the evaluation each student received is a formative evaluation as compared to a summative evaluation. A summative evaluation gives grades and rankings. A formative evaluation gives the students feedback on their performance.

Each student in the course received a data set with the same variables. There are 72 participants. Each student received a different replication data of the same study. The data are not exactly the same for each person. However, the variables are exactly the same.

Before considering statistical analysis, research design need to be considered. In research designs, there are generally one or more independent variables and one or more dependent variables. Independent variables can be manipulated or not manipulated by the researcher. All independent variables have two or more levels (groups). The manipulated independent variable is under the direct control of the researcher. The researcher can change the levels of the independent variable and watch the change on the dependent variable. The dependent variable is the outcome measure. In psychology this is usually in the form of a score. In the data set used for all of the assignments in the class, there are three independent variables: Anxiety, Feedback Type and Grade Type. There are also four dependent variables: Scores on three papers and IQ. Feedback Type and Grade Type are manipulated or active independent variables. The researcher controls them. Certain participants receive different levels than others and that was done by design. The IV (abbreviation for independent variable) Anxiety is in this study a non-manipulated independent variable. The researcher does not have direct control over the variable. People come into the study with pre-set levels of anxiety. Anxiety can be measured and evaluated using the Galvanic Skin Response (GSR) or some type of personality test. The researcher could have manipulated anxiety by producing conditions that provoke high or low anxiety, but that was not done in this study.

The researcher tries to control the study and the variables as much as possible. By doing so, the results obtained in the study can then be attributed to the independent variables. This is called "internal validity". There are two types of research studies. One uses groups of participants and the other uses a single individual or a few participants. These are sometimes called "N=1" studies. With this type of study, there is no need for higher order statistics or inferential statistics, which we will be learning in this course. Like those studies by Skinner or Pavlov or Ebbinghaus, or in psychophysics, groups of subjects are not used. Learning curves are what interest the researcher. Our major concentration in this course will be the use of groups. The growth in the development of this area took place in the late 1920's led by R. A. Fisher.

There are two important theorems we need to consider in statistics and experimental design for groups. The first is the Central Limit Theorem. It is the Central Limit Theorem that allows us to use the normal curve in our work with statistics. Informally, the Central Limit Theorem states "regardless of the shape of the population distribution, if all samples of size n (where n is large) are drawn from the population, the sample means will be approximately normal in its distribution." First let's make a distinction between population and sample. The population is the universe of events defined by the researcher. The researcher can define a population as "all CSUN students." Or it could have been "All CSU students." It is whatever the researcher defines the universe under study. A subset of the population is called a sample. So while the population can have a size N, the sample would have a size n, where $n < N$. This powerful theorem allows us to use the normal curve in our analyses. In some situations where the researcher feels this is not tenable, nonparametric methods can be used in analyzing experimental data.

The second and equally important theorem concerns the linear combination of variables. Before an actual introduction to this, we may want to consider the data set that we will be analyzing. Remember we have three independent variables: Anxiety, Feedback and Grade. The linear combination theorem allows us to write and analyze equations such as:

Score = Anxiety + Feedback + Grade.

In general form: $l = a_1y_1 + a_2y_2 + a_3y_3 + \ldots + a_ny_n$ where

$a_1, a_2, a_3, \ldots, a_n$ are constants or weights and $y_1, y_2, y_3, \ldots, y_n$ are random variables. A linear function of normally distributed random variables will be normally distributed in repeated sampling.

We will see that this forms the basis for our analysis of variance model as well as regression models.

Week 2: Lab: None

Week 3: Lecture.

Let $y_1, y_2, y_3, \ldots, y_n$ be random variables and let $l = a_1y_1 + a_2y_2 + a_3y_3 + \ldots + a_ny_n$.
Then $E(l) = a_1\mu_1 + a_2\mu_2 + a_3\mu_3 + \ldots + a_n\mu_n = \Sigma a_i\mu_i$

and

$$V(l) = a_1^2\sigma_1^2 + a_2^2\sigma_2^2 + a_3^2\sigma_3^2 + \ldots.. + a_n^2\sigma_n^2 \\ + 2a_1a_2\sigma_{12} + 2a_1a_3\sigma_{13} + 2a_2a_3\sigma_{23} + \ldots. + 2a_{n-1}a_n\sigma_{n-1,n}$$

$$V(l) = \sum a_i^2\sigma_i^2 + 2\sum\sum a_ia_j\sigma_{ij}$$

The E(l) is an expected value. V(l) is the variance of the linear combination. The weights are estimated from data.

These equations are important in the following way. When we take into consideration all the variables we study in a research experiment, these equations are implied. These equations play a key role in experimental design. The type of design we choose to use in our study dictates the shape and form of the equations, especially V(l). These equations are used if we have a multivariate or univariate design (and statistical analyses). The traditional use of "multivariate" implies more than one dependent variable. Univariate usually mean only one dependent variable. So it is the number of dependent variables that determines whether we use univariate or multivariate statistics. In this course we will talk about univariate statistics. However, the student should be aware that the linear combination equation also applies to multivariate statistics.

With certain experimental designs, the number of terms in V(l) varies. Some terms may drop others or additional ones would be added. We will see that regression analysis uses this same set of equations. Although factor analysis is not discussed in this course, the linear model presented here also applies. When we refer to "linear", we mean the weights are of the first order. They don't need to be complex in their estimation. Some that may appear complex can be reduced to linearity if a certain transformation is applied to the data. Such transformations will be discussed later.

One of the aspects of experimental design is statistical inference. Based on sample data, we make an inference to the population. The inferential process involves hypothesis testing. Generally when we have a research question, we can form from it a research hypothesis or a set of hypotheses. Research hypotheses make take the form like

Teaching method A will yield higher scores than Method B.
Software design G will produce greater productivity than design J.
Graphical display R will yield fewer perceptual errors than display O.
The coin chosen to gamble with will produce fewer tails than heads.

When we state our hypotheses in this way, we call it the research hypothesis. However, the research hypothesis is not directly testable. Consider the hypothesis about the coin. We can translate the statement into saying $p_{Tails} < p_{Heads}$ or $p_{Heads} > .5$. How many values can pHeads take on and still satisfy the hypothesis statement? If there isn't an infinite number, one can say that there is a lot and probably uncountable. Hence, what we need to do is to state another hypothesis that is the antithesis of the research hypothesis. We would write pHeads $\leq .5$. Note that this can take on a large number of values, but, it points to one value, namely .5. So in hypothesis testing, we assume this hypothesis, called the null hypothesis, to be true and then build up evidence that it is not tenable. This then leads us to believe that the research hypothesis is true. The evidence is in the form of data collected. These empirical data are then processed using statistical methods. The result is the test statistic. However, we still need to make a decision. We need some rule to guide us as to whether or not we have enough evidence to say the null hypothesis is false. The development of such a rule is called the decision rule. A summary of the steps is given below.

Steps in a hypothesis test:

1. Statement of the null hypothesis
2. Statement of the research or alternative hypothesis
3. Computation of the test statistic
4. Develop the Decision Rule
5. Apply the Decision Rule
6. State a conclusion

In summary, the research hypothesis is generally the alternative hypothesis. The alternative hypothesis is an inexact hypothesis. This means it can take on a large number (maybe infinite) of values. It cannot be tested directly. Hypothesis testing involves setting up the null hypothesis as the contrary to the research hypothesis. It is an exact hypothesis and can be tested directly. The process is the reverse to what we normally think. We want to show the null hypothesis to be false so that this leads us to believe the alternative is true. The test statistic is computed from empirical evidence. It is used to show that the null hypothesis is not tenable, i.e., it can't happen very often in light of the data.

Types of decision errors:

Type I: Reject H_0 when H_0 is true
Type II: Not rejecting H_0 when H_0 is false.

If we wanted to determine if a coin is fair or not, we can set up the hypothesis test summarized above. Let's say we use as a decision rule if we toss a coin 100 times to reject H0 if the number of heads exceed 53. Let's say we obtain 54 heads in our experiment. Using the decision rule, we would reject H0. However, a fair coin could have produced 54 heads. When this happens, we have made a Type I error. In the other case, let's say our experiment yield exactly 50 heads. We would decision not to reject the H0. However, an unfair coin could have produced 50 heads. When this situation occurs, we have committed a Type 2 error. We try to design studies to minimize both types of errors. In psychology we generally set the probability of the Type I error to .05 or .01. There is no real supporting evidence for doing this. The probability of a Type 1 error is designated with the Greek letter α. The Type 2 error probability is designated as β. $1-\beta$ is called the power of the test and we will discuss that later when we talk about how it is related to sample size determination. As the sample size is increased with a fixed amount for α, the $1-\beta$ gets larger (this also means β is getting smaller). Sample size determination is very important for those applying for research grants. Most granting agencies want to know what the costs will be per participant. Some marketing research studies are priced to the customer on the basis of costs per participant.