SAN JOSÉ STATE UNIVERSITY

College of Social Work
S. W. 242
Spring 2008
Edward Cohen

Week 5: February 22, 2008 (Some of this will be covered in Week 6 as well)

The Normal Curve; *z* Scores; Area Under the Curve (Preliminary material for hypothesis testing)

- **Comments on Assignment 1**
- **Intro to statistical inference**
- **Intro to SPSS**
- **Final Paper preparation – bring your Introduction and research question(s) or hypothesis next week, or notes from them**

"Statistics means never having to say you're certain!"
Anon.

- Normal curve
- Normal distribution
- Skewed distribution
- *z* score
- "area under the curve"
- Statistical inference
- Null hypothesis
- Alternative hypothesis
- *p* value

I. Comments on Assignment 1

A. Adjust language for your population. Children require different language than adults. Use less technical description of research questions, e.g. avoid "The effects of …" *This is true for the Informed Consent as well—you don't have to use your exact study title.* Also avoid using words in scripts or questions like "child abuse," "mental illness," "trauma" and other words that might put people off or bias their response (except, see #4 below). In general for any population, avoid technical jargon and acronyms in scripts and questions.

B. Regarding "ethnicity"—most of the time it's best just to ask someone "What ethnicity do you identify with most?" as an open ended question. Then, later you can always categorize the responses as a nominal variable. People tend to not like being categorized in boxes, and the end result of either way you ask the question is the same.

C. In your final paper, the Informed Consent will be added as an attachment. You will be referring to it in your Methods section.

D. Make clearer the distinction between the quantitative and qualitative components. You can say "mixed methods" but you still have to differentiate research questions, design, sampling, interviewing, and analysis methods for both types of components.

E. Qualitative studies require many fewer participants in the sample than quantitative designs.

F. With some sensitive subjects, ask questions in a less personalized nature. For example, when asking youth about their emotional problems and social interactions, instead of asking "Do you think your emotional problems affect your social interactions?", ask "Do you think that people's emotional problems affect how they get along with others, or affect what their relationships are like?" In this way you approach your respondents as experts rather than as subjects, and you are more likely to get better information.

G. For your Final Paper you are required to have consistent font size throughout—even for attachments if feasible.

II. Review from last week: ***fascinating*** information about the standard deviation:

Look at this output from last week's lecture:

Comparison of Means, ScWk 240 Exams, Fall 2008

	N	Range	Mean	**Std. Deviation**	Variance
Test_1	50	4.00	14.2549	**1.29373**	1.674
Test_2	50	4.00	14.6800	**.84370**	.712

1. For those of you who like looking at formulas (you know who you are), the formula for the variance is

$$\frac{\Sigma(x_i - \bar{x})^2}{N - 1}$$

where $\Sigma(x_i - \bar{x})^2$ is "the sum of the squared differences of each individual x from the mean", or the sum of squares. *N* is the total sample. The variance is the average amount of dispersion from the mean

2. The standard deviation (the square root of the variance) is in the same units as the original measure (test score points) whereas the variance is in squared units, $(\text{test score points})^2$

3. For the SD of Test_1, we say that "one standard deviation is equivalent to 1.29 points". In Test_2, "one standard deviation is equivalent to .84 of a test score point."

***If your sample's test scores have a mean of 10 and a standard deviation of 2 (i.e. a SD is equal to 2 score points), what would be the test score for someone two standard deviations above the mean? How far from the mean (in standard deviations) would a score of 4 be?

4. A *population* can have a mean and standard deviation of a measure.

   a) Many psychological tests or other measurement instruments, when tested and validated, have been administered enough to know the population parameters (means and standard deviations), which allows you to compare your particular client's or research participant's score to the population. Note—there can be different means and SDs for women, men, children, ethnic groups, etc.

b) In much of our social sciences research, however, we usually don't know the population parameters and so we have to *infer* them from our sample statistics.

III. Normal Distributions

A. A <u>normal distribution</u> (or normal curve) is a bell-shaped curve (a type of frequency polygon) generated from the plotting of values from *interval* and *ratio* level variables. Some properties are:

1. Bell-shaped

2. Symmetrical

3. Mode, median, and mean all occur at the highest point and in the center of the distribution

4. The curve goes to infinity and never touches the x-axis. This accounts for extreme +,- values

B. What does curve of the frequency polygon tell us about the distribution of the variable?

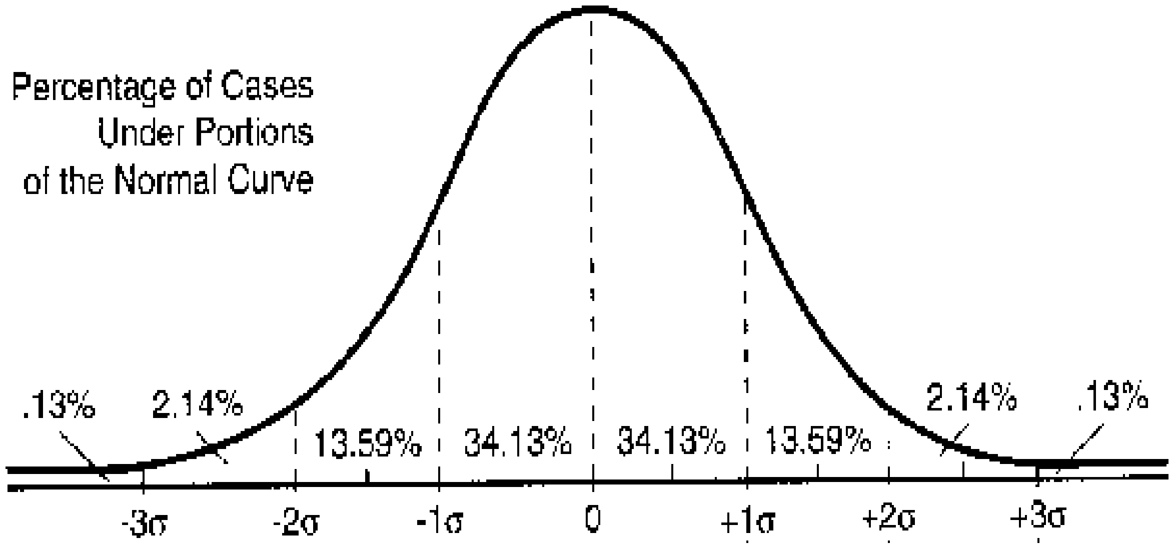
1. Whether the variable is normally distributed, also known as symmetric (see below), or **skewed**

a) Positively skewed—the right tail is longer than the left one

b) Negatively skewed—the left tail is longer than the right one

2. For statistical inference testing we are interested in what proportion of values occur in any given distance from the mean

C. The normal curve can be divided into standard deviation units:



1. The omega (σ) symbol stands for standard deviation. With continuous level data that are "normally distributed" the curve is symmetrical--the zero point in the middle is also where the mean and median are

2. Fun facts about the normal curve

   a) Approximately 68% of the data observations (34.13% + 34.13%) will be within one standard deviation on either side of the mean

   b) Approximately 95% (13.59% + 34.13% + 34.13% + 13.59%) will be within two standard deviations on either side of the mean.

***What is the percentile of a score falling 2 standard deviations above the mean?
***What is the percentile of a score falling exactly at the mean?

   c) Only a small percentage of observations (less than 5%, or less than .05) will fall in the tails.

D. If a dataset is normally distributed in the population, then we can "standardize" the original raw data into **z scores**.

1. Calculating a z score: (raw score – mean) divided by the SD, or for you formula buffs:

$$\frac{x_i \quad \bar{x}}{SD}$$

   a) For a mean of 10, with a SD of 2, what's the z score of a raw score of 12?

   (12 - 10) / 2 = 2 / 2 = 1

   b) For a mean of 10, with a SD of 2, what's the z score of a raw score of 13?

   (13 - 10) / 2 = 3 / 2 = 1.5

2. Once we have a *z* score, then we can find out the *area under the curve* between a raw score and a mean, or the percentile of the raw score.

   a) For a z score of 1, the area between the score and the mean is 34.13% (also see figure, and see Table 4.3 in Weinbach).

   b) For a z score of 1.5, the area under the curve between the score and the mean is 43.32%, or [(34.13%) + (13.59%/2)]. Also, see Table 4.3 in Weinbach.

3. And, knowing the area between the score and the mean, we also know the percentile of that score.

   a) For a z score of 1, the percentile of the raw score is (34.13% + 50%) = 84.13%, or the (area between the score and the mean, plus the rest of the area to the left of the mean). So, a raw score of 12, with a mean of 10 and SD of 2 is at approximately the 84^{th} percentile.

   b) For a z score of 1.5, the percentile of the raw score is (43.32% + 50%) = 93.32%, or the (area between the score and the mean, plus the rest of the area to the left of the mean). So, a raw score of 13, with a mean of 10 and SD of 2 is at approximately the 93rd percentile.

***With this same example (mean=10, SD=2) what percentage of scores would fall below a score of 14?

4. Why is this important? Several datasets of different scales and units of measures can be transformed into z scores and compared to each other (and their percentiles can be compared).

5. ***Understanding the area under the normal curve is important to understand hypothesis testing, in general***

IV. Statistical Significance and Hypothesis Testing—What can we *infer* from the sample about the population?

A. **Inference** – using a statistic (that is, a summary of the sample) to draw tentative conclusions about the population from which the sample was drawn, and to make a probability statement about our confidence in those conclusions

- This inference relies on
  1. Some statistic that shows summary measures our variables or relationships between the variables (Descriptive statistics, Chi Square, Correlation, ANOVA, linear regression, etc.)
  2. A measures of variability of that statistic, usually based on the standard deviation for that statistic
  3. A quantified probability of our confidence (using area under the curve), showing the likelihood that our statistic represents the population

B. What can influence this confidence?

1. Rival hypotheses (remember that from 240?)

2. Measurement problems (remember measurement error?)

These two influences are the result of research design. *No statistical testing can take the place of good research design!*

3. **Sampling bias, or sampling error** -- Chance contends that no matter how well a research study is designed, any relationship between variables within the sample is just a "fluke" or random occurrence. Sampling error is the natural tendency of any sample to differ from the population from which it was drawn. Sampling error is especially likely to occur in small samples, even when rival hypotheses and measurement error are taken care of.

***Why would sampling error be more likely to occur in small samples? (Hint—think about variance and standard deviation.)

4. *A statistic will quantify the extent of sampling error.* It's up to the researcher to interpret whether rival hypotheses or measurement error are likely culprits, based on limitations of the research design.

5. **Preview to next week's topic:**

a) Null Hypothesis: *There is no relationship between the two variables in the population*

- Treatment does not result in improved outcomes
- Poverty has no relationship to extent of emergency room usage

b) Alternative Hypothesis (or "research hypothesis"—the one you worked so hard on last semester)*There is a relationship between the variables in the population*

- Treatment results in improved outcomes
- Poverty is related to the extent of emergency room usage

c) What's a "p value" ($p \leq .05$)? "We can say with a degree of certainty that the relationship found between these variables can only happen *by chance* 5% of the time." Or, the Null Hypothesis can be rejected.