IGP 304 (Spring 2006) Homework 2 Keys:

1. Because serum cholesterol is related to age and sex, some investigators prefer to express it in terms of z-scores. If X is raw serum cholesterol, then $z = (X - \mu)/\sigma$, where μ is the mean and σ is the standard deviation of serum cholesterol for a given age-sex group. Suppose z is regarded as a standard normal distribution. Also suppose a person is regarded as having high cholesterol if $z > 2.0$ and borderline cholesterol if $1.5 < z < 2.0$. [Make sure you can get the answers from both tables and Stata, although you don't need to turn in Stata output.]
    - What is $\Pr(z < 0.5)$?
    - What is $\Pr(z > 0.5)$?
    - What is $\Pr(-1.0 < z < 1.5)$?

$\Pr(z < 0.5) = 0.69$; $\Pr(z > 0.5) = 0.31$; $\Pr(-1.0 < z < 1.5) = 0.77$. Stata command `di normal(x)` should give $\Pr(z < x)$.

    - What proportion of people have high cholesterol?
    - What proportion of people have borderline cholesterol?

$\Pr(z > 2.0) = 0.023$ or 2.3% of people have high cholesterol; $\Pr(1.5 < z < 2.0) = 0.044$ or 4.4% of people have borderline cholesterol.

2. Beta-carotene is hypothesized to prevent macular degeneration (an important eye disease in the elderly). A dietary survey was undertaken to measure beta-carotene intake in the typical American diet. Assume the distribution of ln(beta-carotene intake) is normal, with mean 8.34 and standard deviation 1.00. [Units are in ln(IU). ln = natural logarithm.]
    - What percentage of people have dietary beta-carotene intake below 2000 IU? (Note that ln(2000)=7.60.)

$\Pr(IU < 2000) = \Pr(\ln(IU) < 7.60) = \Pr[z < (7.60 - 8.34)/1.00] = \Pr(z < -0.74) = 0.23$. Thus, 23% of people have dietary beta-carotene intake below 2000 IU.

    - What percentage of people have dietary beta-carotene intake below 1000 IU? (Note that ln(1000)=6.91.)

$\Pr(IU < 1000) = \Pr(\ln(IU) < 6.91) = \Pr[z < (6.91 - 8.34)/1.00] = \Pr(z < -1.43) = 0.076$. Thus, 7.6% of people have dietary beta-carotene intake below 1000 IU.

    - Some studies suggest beta-carotene intake over 10,000 IU may protect against macular degeneration. What percentage of people have a dietary intake of at least 10,000 IU?

Because $\ln(10000) = 9.21$, $\Pr(IU > 10000) = \Pr(\ln(IU) > 9.21) = \Pr[z > (9.21 - 8.34)/1.00] = \Pr(z > 0.87) = 0.19$. Thus, 19% of people have dietary beta-carotene intake at least 10,000 IU.

- Suppose each person took a beta-carotene supplement pill of dosage 5000 IU in addition to his or her normal diet. Assume the resulting distribution of ln(beta-carotene intake) is normally distributed, with mean 9.12 and standard deviation 1.00. What percentage of people would have an intake from diet and supplements of at least 10,000 IU?

Pr(IU > 10000) = Pr(ln(IU) > 9.21) = Pr[z > (9.21 – 9.12)/1.00] = Pr(z > 0.09) = 0.46. Thus, 46% of people would have an intake from diet and supplements of at least 10,000 IU.

3. Much discussion has appeared in the medical literature in recent years on the role of diet in the development of heart disease. The serum-cholesterol levels of a group of people who eat a primarily macrobiotic diet are measured. Among 24 of them, aged 20-39, the mean cholesterol level was found to be 175 mg/dL with a standard deviation of 35 mg/dL.
   - If the mean cholesterol level in the general population in this age group is 230 mg/dL and the distribution is assumed to be normal, then test the hypothesis that the group of people on a macrobiotic diet have cholesterol levels different from those of the general population.

Let μ be the average cholesterol level in the group of people on a macrobiotic diet and $\mu_0 = 230$ be the mean cholesterol level in the general population in this age group. The null hypothesis is H_0: $\mu = \mu_0 = 230$ and the alternative hypothesis is H_a: $\mu \neq 230$. We will carry out a one-sample t-test. The test statistic is $(\bar{x} - \mu_0)/(s/\sqrt{n}) = (175 - 230)/(35/\sqrt{24}) = -7.698$. Comparing this statistic with the t-distribution with 23 degrees of freedom yields a two-sided p-value 8×10^{-8} (obtained by Stata command `di 2*ttail(23, 7.698)`). Because this p-value is very small, there is very strong evidence to reject the null hypothesis.

   - Compute a 95% confidence interval for the true mean cholesterol level in this group.

The 97.5% percentile of the t-distribution with 23 degrees of freedom is $t' = 2.0687$ (obtain by Stata command `di invttail(23, .025)`). Thus, the 95% confidence interval for the true mean cholesterol level in this group is, using formula $\bar{x} \pm t' \times s/\sqrt{n}$, from $175 - 2.0687\times35/\sqrt{24} = 160.2$ to $175 + 2.0687\times35/\sqrt{24} = 189.8$. We are 95% confident that the true mean cholesterol level μ in this group is $160.2 \le \mu \le 189.8$.

   - What type of complementary information is provided by the hypothesis test and confidence interval in this case?

With the same significance level, say α, there is a one-to-one correspondence between hypothesis testing and confidence interval: We reject the null hypothesis H_0: $\mu = \mu_0$ at level α if and only if the (1– α)% confidence interval doesn't contain the null value μ_0. However, confidence intervals tend to provide more information. In this problem, hypothesis testing tells us we have very strong evidence against the null hypothesis that the mean cholesterol level in the group is 230 mg/dL, while the confidence interval tells us the range of values that the true mean cholesterol level in the group will likely fall into.

4. One method for assessing the bioavailability of a drug is to note its concentration in blood and/or urine samples at certain periods of time after giving the drug. Suppose we want to compare the concentrations of two types of aspirin (types A and B) in urine specimens taken from the same person, 1 hour after he or she has taken the drug. Hence, a specific dosage of either type A or type B aspirin is given at one time and the 1-hour urine concentration is measured. One week later, after the first aspirin has presumably been cleared from the system, the same dosage of the other aspirin is given to the same person and the 1-hour urine concentration is noted. Because the order of giving the drugs may affect the results, a table of random numbers is used to decide which of the two types of aspirin to give first. This experiment is performed on 10 people; the results are given in this table. Suppose we want to test the hypothesis that the mean concentrations of the two drugs are the same in urine specimens.
    - What are the appropriate hypotheses?

The null hypothesis is that the mean concentrations of the two drugs are the same in urine specimens; the alternative hypothesis is the mean concentrations are different. Let μ_A and μ_B be the mean concentrations for drugs A and B. Then the null hypothesis is H_0: $\mu_A - \mu_B = 0$ and the alternative hypothesis is H_a: $\mu_A - \mu_B \neq 0$. (You also can write H_0: $\mu_A = \mu_B$ vs. H_a: $\mu_A \neq \mu_B$.)

    - What are the appropriate procedures to test these hypotheses? Conduct the tests mentioned.

A paired t-test is appropriate. There are 10 differences, with mean 3.6 and standard deviation 3.098. Thus the test statistic is $3.6/(3.098/\sqrt{10}) = 3.675$, with 9 degrees of freedom. The two-sided p-value is 0.0051. Because this p-value is very small, we have strong evidence to reject the null hypothesis.

    - What is the best single-number estimate of the mean difference in concentrations between the two drugs? (A single-number estimate is also called a point estimate.)

The best point estimate for the mean difference in concentrations between the two drugs is 3.6.

    - What is a 95% CI for the mean difference? (A CI is also called an interval estimate.)

The 95% CI for the mean difference is $3.6 \pm t' \times 3.098/\sqrt{10}$, where $t' = 2.262$ is the 97.5% percentile of t_9. Thus the 95% CI is from 1.38 to 5.82. In other words, we have 95% confidence that the true mean difference in concentrations between the two drugs is between 1.38 and 5.82.

    - Suppose a significance level of .05 is used in your test. What is the relationship between the decision reached with the test procedure and the nature of the confidence interval?

See my answer to the last question of Problem 3. In short, we reject the null hypothesis at level 0.05 if and only if zero is not inside the 95% CI.

5. True or false? If false, state (1) why it is wrong and (2) the correct results/statements/conclusions that is beyond just grammatically negating a false statement.
    - In a test to compare the birth weights of children born to 14 smokers with those of children born to 15 non-smokers, the p-value was 0.0064. So the probability of the null hypothesis of no difference is 0.0064.

False. The p-value measures the level of departure of data from the null hypothesis. Assuming the null hypothesis is true, it is calculated as the probability of observing data with the same level of departure as or more extreme than the data in hand. It cannot be interpreted as the probability of the null hypothesis. A correct statement should be: If there is no birth weight difference between children born to smokers and those born to non-smokers, the probability of observing data with at least the level of difference in our data would be 0.0064, less than one in a hundred. Thus, it is likely there is a difference.

    - In a clinical trial to compare the effect of a drug with that of placebo, both the drug and placebo effects were measured for every subject (assuming carry-over effect is ignorable). We can calculate the average of the drug effects, the average of the placebo effects, and the average of per-subject drug-placebo effect differences. Since the difference of the first two averages is the same as the third average, we can carry out either a two-sample t-test or a paired t-test to give us the same results.

False. Even though these two ways of calculating average difference lead to the same estimate, a two-sample t-test and a paired t-test won't give the same results. The two-sample t-test ignores the information inherent in the data that two numbers are observed on the same subject, while the paired t-test takes this information into consideration. As a result, the two-sample t-test is not as powerful to detect the drug effect as the paired t-test. A correct statement should be: We should carry out a paired t-test to test for the effect difference between drug and placebo.

    - For a sample, we want to know its variation or how the data spread out. There are two ways of capturing variation: standard deviation and inter-quartile range. The two measures are similar numerically and are effectively interchangeable.

False. Although both concepts reflect variation in the data, they are not numerically the same. Strictly speaking, they also are not directly comparable. For data generated from a normal distribution with mean μ and variance σ^2, the SD will be around σ, while the IQR will be around 1.35σ, starting from about $\mu - 0.67\sigma$ to about $\mu + 0.67\sigma$. (The numbers are obtained by Stata command `di invnormal(.75)` and `di invnormal(.25)`.) In addition, SD is sensitive to extreme values while IQR is not. A correct statement should be: These two measures are not similar numerically and thus are not interchangeable.