Chapter Six Notes:

Spring 2023

New Problem (just slightly different from Chapter 4, number 7):

The random variables W_1 and W_2 are independent. Their expected values are $E(W_1) = \mu_1$ and $E(W_2) = \mu_2$. Their variances are $\text{var}(W_1) = \sigma_1^2 < \infty$, and $\text{var}(W_2) = \sigma_2^2 < \infty$. Find $E(W_1 - W_2)$ and $\text{var}(W_1 - W_2)$.

Solution:

First, $E(W_1 - W_2) = E(W_1) - E(W_2) = \mu_1 - \mu_2$.

Second, $\text{var}(W_1 - W_2) = E\{[(W_1 - W_2) - E(W_1 - W_2)]^2\}$

$$\text{var}(W_1 - W_2) = E\{[(W_1 - W_2) - (\mu_1 - \mu_2)]^2\}$$

$$\text{var}(W_1 - W_2) = E\{[(W_1 - \mu_1) - (W_2 - \mu_2)]^2\}$$

$$\text{var}(W_1 - W_2) = E\{[(W_1 - \mu_1)^2 + (W_2 - \mu_2)^2 - 2(W_1 - \mu_1)(W_2 - \mu_2)]\}$$

$$\text{var}(W_1 - W_2) = E[(W_1 - \mu_1)^2] + E[(W_2 - \mu_2)^2] - 2E[(W_1 - \mu_1)(W_2 - \mu_2)]$$

$$\text{var}(W_1 - W_2) = \text{var}(W_1) + \text{var}(W_2) - 2\,\text{cov}(W_1, W_2)$$

$$\text{var}(W_1 - W_2) = \sigma_1^2 + \sigma_2^2 - 2 \bullet 0$$

$$\text{var}(W_1 - W_2) = \sigma_1^2 + \sigma_2^2$$

Two Independent Sample Test

Let $X_1, X_2, \ldots, X_n$ be a random sample of size *n* from the random variable *X*, which is $N(\mu_X, \sigma_X^2)$. For example, *X* could be the response of a participant in a clinical trial to a new medicine. Let $B_1, B_2, \ldots, B_m$ be a random sample of size *m* from the random variable *B*, which is $N(\mu_B, \sigma_B^2)$. Continuing the example, B could be the response of a participant in a clinical trial to the best available medicine. The two samples are independent of each other. The context of this discussion is that there is random assignment of the participants in the study to the two groups. This has the effect of roughly balancing the two groups with respect to each and every variable other than the treatment assigned. If there is a significant difference between the two groups, that difference is either a chance event or the causal result of a difference in the treatments.

Back to the probability theory of the analysis, $\bar{X}_n$ is $N(\mu_X, \frac{\sigma_X^2}{n})$, and $\bar{B}_m$ is $N(\mu_B, \frac{\sigma_B^2}{m})$ The two sample averages are independent. We seek to use this data to test $H_0 : E(X - B) = 0$ against the alternative hypothesis $H_1 : E(X - B) \neq 0$ at the α level of significance. Our test statistic is $TS = \bar{X}_n - \bar{B}_m$.

Distribution of the Test Statistic

The distribution of $TS = \bar{X}_n - \bar{B}_m$ is normal. In the case that the random variable *X* is not normal (or the random variable *B* is not normal), then the CLT implies that the distribution of *TS* is approximately normal. From the problem at the start of class,

$E(TS) = E(\bar{X}_n - \bar{B}_m) = E(\bar{X}_n) - E(\bar{B}_m) = \mu_X - \mu_B$, and

$\mathrm{var}(TS) = \mathrm{var}(\bar{X}_n - \bar{B}_m) = \mathrm{var}(\bar{X}_n) + \mathrm{var}(\bar{B}_m) = \frac{\sigma_X^2}{n} + \frac{\sigma_B^2}{m}$.

Null Distribution of the Test Statistic

Since $H_0 : E(X - B) = 0$ which is equivalent to $H_0 : \mu_X = \mu_B$, $E_0(TS) = E_0(\bar{X}_n - \bar{B}_m) = \mu_X - \mu_B = 0$. R. A. Fisher argued that the null hypothesis should be that the random variable *X* has exactly the same distribution as the random variable *B*. That is, not only does $\mu_X = \mu_B$ under the null, but also $\sigma_X^2 = \sigma_B^2 = \sigma^2$. In fact, any probabilistic parameter of random variable has the same value for *X* and *B*. Since it is very unusual for a treatment to have a measurable effect, this conception of the null hypothesis is widely used. Using this assumption,

$\mathrm{var}_0(TS) = \mathrm{var}_0(\bar{X}_n - \bar{B}_m) = \frac{\sigma^2}{n} + \frac{\sigma^2}{m} = \sigma^2(\frac{1}{n} + \frac{1}{m})$.

Test when variance known

As in Chapter 5, we test this null hypothesis by putting TS in standard score form:

$Z=\dfrac{\overline{X}_n-\overline{B}_m-0}{\sqrt{\sigma^2(\frac{1}{n}+\frac{1}{m})}}$. If $\alpha=0.10$, we reject H_0 when $|Z|\geq 1.645$. If $\alpha=0.05$, we reject H_0 when $|Z|\geq 1.960$. If $\alpha=0.01$, we reject H_0 when $|Z|\geq 2.576$.

Test when variances unknown but equal

Just as in Chapter 5, we use an estimate of σ^2 and stretch the critical values an amount determined by the degrees of freedom of our estimate. There are a number of estimates of σ^2. For example, $S_X^2=\dfrac{\sum_{i=1}^{n}(X_i-\overline{X}_n)^2}{n-1}$ and $S_B^2=\dfrac{\sum_{i=1}^{n}(B_i-\overline{B}_m)^2}{m-1}$ are unbiased estimates of σ^2, with $n-1$ and $m-1$ degrees of freedom respectively. That is, $E(S_X^2)=E(S_B^2)=\sigma^2$. We use both of these estimates. Let $S_P^2=\dfrac{(n-1)S_X^2+(m-1)S_B^2}{n+m-2}$.

This estimator has $n+m-2$ degrees of freedom. Then our studentized statistic is

$T_{n+m-2}=\dfrac{\overline{X}_n-\overline{B}_m-0}{\sqrt{S_P^2(\frac{1}{n}+\frac{1}{m})}}$. If $\alpha=0.10$, we reject H_0 when $|T_{n+m-2}|\geq t_{1.645,n+m-2}$. Similarly, if $\alpha=0.05$, we reject H_0 when $|T_{n+m-2}|\geq t_{1.960,n+m-2}$. If $\alpha=0.01$, we reject H_0 when $|T_{n+m-2}|\geq t_{2.576,n+m-2}$.

Example Problem: from examination 1, Fall 2016, 2B

A research team took a random sample of 3 observations from a normally distributed random variable Y and observed that $\bar{y}_3=37.4$ and $s_Y^2=42.6$, where $\bar{y}_3$ was the average of the three observations sampled from Y and s_Y^2 was the unbiased estimate of $\mathrm{var}(Y)$ (i.e., the divisor in the variance was $n-1$). A second research team took a random sample of 5 observations from a normally distributed random variable X and observed that $\bar{x}_5=50.6$ and $s_X^2=48.1$, where $\bar{x}_5$ was the average of the five observations sampled from X and s_X^2 was the unbiased estimate of $\mathrm{var}(X)$ (i.e., the divisor in the variance was $n-1$). Test the null hypothesis $H_0: E(X)=E(Y)$ against the alternative $H_1: E(X)\neq E(Y)$ at the 0.10, 0.05, and 0.01 levels of significance using the pooled variance t-test. This problem is worth 40 points.

Solution: $s_P^2 = \frac{(5-1)48.1 + (3-1)42.6}{5+3-2} = 46.27$ on 6 degrees of freedom. The t stretches of 1.645, 1.960, and 2.576 for 6 degrees of freedom are 1.943, 2.447, and 3.707 respectively. The test statistic is $t_{n+m-2} = \frac{37.4 - 50.6 - 0}{\sqrt{46.27(\frac{1}{5} + \frac{1}{3})}} = -2.637$. Reject the null hypothesis at the 0.10 and 0.05. Accept the null hypothesis at the 0.01 level.

Alternate Problem:

A research team took a random sample of 3 observations from a normally distributed random variable *Y* and observed that $\bar{y}_3 = 37.4$ and $s_Y^2 = 42.6$, where $\bar{y}_3$ was the average of the three observations sampled from *Y* and s_Y^2 was the unbiased estimate of $\mathrm{var}(Y)$ (i.e., the divisor in the variance was $n-1$). A second research team took a random sample of 5 observations from a normally distributed random variable *X* and observed that $\bar{x}_5 = 50.6$ and $s_X^2 = 48.1$, where $\bar{x}_5$ was the average of the five observations sampled from *X* and s_X^2 was the unbiased estimate of $\mathrm{var}(X)$ (i.e., the divisor in the variance was $n-1$). Find the 99% confidence interval for $E(X) - E(Y)$.

Solution: The template for the 99% confidence interval for $E(X) - E(Y)$ is $\bar{x}_n - \bar{y}_m \pm t_{2.576,n+m-2} s_P \sqrt{\frac{1}{n} + \frac{1}{m}}$. That is, the 99% confidence interval for $E(X) - E(Y)$ is $50.6 - 37.4 \pm 3.707\sqrt{46.27}\sqrt{\frac{1}{5} + \frac{1}{3}} = 13.2 \pm 3.707 \bullet 6.802 \bullet 0.7303 = 13.2 \pm 18.4$. That is, the 99% confidence interval for $E(X) - E(Y)$ is the interval between -5.2 and 31.6. Since 0 is in the 99% confidence interval, we should accept $H_0 : E(X) = E(Y)$ against $H_1 : E(X) \neq E(Y)$ at the 0.01 level of significance.

Test when variances unknown and unequal

This procedure is called the unequal variance t-test or unequal variance confidence interval. As before, let $X_1, X_2, \ldots, X_n$ be a random sample of size *n* from the random variable *X*, which is $N(\mu_X, \sigma_X^2)$. Let $B_1, B_2, \ldots, B_m$ be a random sample of size *m* from the random variable *B*, which is $N(\mu_B, \sigma_B^2)$. Here, $\sigma_X^2 \neq \sigma_B^2$. Then, $\mathrm{var}(\overline{X}_n - \overline{B}_m) = \dfrac{\sigma_X^2}{n} + \dfrac{\sigma_B^2}{m}$. The standard score form of the test statistic is then

$Z = \dfrac{\overline{X}_n - \overline{B}_m - 0}{\sqrt{\dfrac{\sigma_X^2}{n} + \dfrac{\sigma_B^2}{m}}}$. The studentized standard score form of test statistic would be

$T_? = \dfrac{\overline{X}_n - \overline{B}_m - 0}{\sqrt{\dfrac{S_X^2}{n} + \dfrac{S_B^2}{m}}}$. The problem is that the probability theory calculations that are the basis of the degrees of freedom in the pooled variance t-test are not valid. There is, however, a complicated formula (called Satterthwaite's formula) that produces an approximate degrees of freedom. I will not ask you to calculate this by hand in an examination. Fortunately, the statistical programs that you will use will calculate the unequal variance t-test and unequal variance confidence interval for you. Typically, the equal variance t-test and unequal variance t-tests have essentially equal p-values when the assumption of equal variance appears reasonable. When the assumption does not appear reasonable, the one should use the unequal variance calculations. My practice is to report the unequal variance results as calculated by a reputable statistics program.

Type II Error Rate and Power Calculations

The definition of the Type II error rate is $\beta = \Pr_1\{\text{Accept } H_0\}$. The power of a statistical test is defined to be $\text{Power} = 1 - \beta$. This is the probability that the null hypothesis is correctly rejected. A large Type II error rate indicates a study that is "underpowered." The calculation of β is just a normal probability calculation. The specification of the normal distribution is based on the alternative specified in the problem. Typically, the values of the variances are assumed known.

Example Problem: from Chapter 6 Study Guide, Problem 3

In a clinical trial, 50 patients suffering from an illness will be randomly assigned to one of two groups so that 25 receive an experimental treatment and 25 receive the best available treatment. The random variable *X* is the response of a patient to the experimental medicine, and the random variable *B* is the response of a patient to the best currently available treatment. The random variables *X* and *B* are normally distributed with $\sigma_X = \sigma_B = 500$ under both the null and alternative distributions. The null hypothesis to be tested is that $E(X) - E(B) = 0$ against the alternative that $E(X) - E(B) > 0$ at the 0.01 level of significance. What is the probability of a Type II error for the test of the null hypothesis when $E(X) - E(B) = 500$?

Solution: The standard score form of the TS is $Z = \dfrac{\overline{X}_n - \overline{B}_m - 0}{\sqrt{\sigma^2(\frac{1}{n} + \frac{1}{m})}}$, and $H_0 : E(X - B) = 0$ is rejected when $Z \geq 2.326$, remembering that the problem asks for a one-sided test at level of significance 0.01. Using the TS directly, $H_0 : E(X - B) = 0$ is rejected when $\overline{X}_{25} - \overline{B}_{25} \geq 0 + 2.326\sqrt{\dfrac{500^2}{25} + \dfrac{500^2}{25}} = 0 + 2.326 \bullet 141.42 = 328.95$. For the alternative specified in the problem $\overline{X}_{25} - \overline{B}_{25}$ is $N(500, 141.42^2)$. Then

$$\beta = \Pr_1\{\text{Accept } H_0\} = \Pr_1\{\overline{X}_{25} - \overline{B}_{25} < 328.95\} = \Pr\{\frac{\overline{X}_{25} - \overline{B}_{25} - E_1(\overline{X}_{25} - \overline{B}_{25})}{\sigma_1(\overline{X}_{25} - \overline{B}_{25})} < \frac{328.95 - 500}{141.42}\}.$$

That is, $\beta = \Pr\{Z < \dfrac{328.95 - 500}{141.42} = -1.210\} = \Phi(-1.210) = 0.113$.

Sample size for two sample test:

The bad news is that the mathematics of sample size calculations is relatively complex. The good news is that one can solve a wide range of sample size problems once one knows how to solve this one. The argument is essentially the same.

Problem 6 in Chapter 6 Study Guide

> In a clinical trial, *2J* patients suffering from an illness will be randomly assigned to one of two groups so that *J* will receive an experimental treatment and *J* will receive the best available treatment. The random variable *X* is the response of a patient to the experimental medicine, and the random variable *B* is the response of a patient to the best currently available treatment. The random variables *X* and *B* are normally distributed. The null hypothesis to be tested is that $E(X) - E(B) = 0$ against the alternative that $E(X) - E(B) > 0$ at the $\alpha,\ \alpha \leq 0.5,$ level of significance. When the null hypothesis is true,

$\operatorname{var}(X)=\operatorname{var}(B)=\sigma_0^2$. When the alternative hypothesis is true, $\operatorname{var}(B)=\sigma_0^2$, but $\operatorname{var}(X)=\sigma_1^2>\sigma_0^2$. What is the number J in each group that would have to be taken so that the probability of a Type II error for the test of the null hypothesis specified in the common section is $\beta,\ \beta\le 0.5$, when $E(X)-E(B)=\Delta>0$?

Solution: The test statistic is $TS=\bar{X}_J-\bar{B}_J$, and TS is $N(E(X)-E(B),\frac{\operatorname{var}(X)}{J}+\frac{\operatorname{var}(B)}{J})$. The null distribution of TS is then $N(0,\frac{\sigma_0^2}{J}+\frac{\sigma_0^2}{J})$. Hence, we reject $H_0:E(X)-E(B)=0$ against $H_1:E(X)-E(B)>0$ at the α level of significance when $TS\ge 0+|z_\alpha|\sqrt{\frac{\sigma_0^2}{J}+\frac{\sigma_0^2}{J}}$. When $E(X)-E(B)=\Delta>0$ and $\operatorname{var}(X)=\sigma_1^2>\sigma_0^2$.and $\operatorname{var}(B)=\sigma_0^2$, the (alternative) distribution of TS is then $N(\Delta,\frac{\sigma_0^2}{J}+\frac{\sigma_1^2}{J})$. Then, the probability of a Type II error is $\beta=\Pr_1\{\text{Accept } H_0\}=\Pr_1\{TS<0+|z_\alpha|\sqrt{\frac{2\sigma_0^2}{J}}\}$. That is,

$$\beta=\Pr_1\{TS<0+|z_\alpha|\sqrt{\frac{2\sigma_0^2}{J}}\}=\Pr\{Z=\frac{TS-\Delta}{\sigma_1(\bar{X}_J-\bar{B}_J)}<\frac{0+|z_\alpha|\sqrt{\frac{2\sigma_0^2}{J}}-\Delta}{\sqrt{\frac{\sigma_0^2}{J}+\frac{\sigma_1^2}{J}}}\}.$$ Since $\beta=\Pr_1\{\text{Accept } H_0\}\le 0.5$, it is true that $\beta=\Pr\{Z<-|z_\beta|\}$. We now have two equations:

$$\beta=\Pr\{Z<\frac{0+|z_\alpha|\sqrt{\frac{2\sigma_0^2}{J}}-\Delta}{\sqrt{\frac{\sigma_0^2}{J}+\frac{\sigma_1^2}{J}}}\},$$ and

$\beta=\Pr\{Z<-|z_\beta|\}$. The problem is to choose J so that the probability of a Type II error is a specified value. That is, we should choose J so that the right hand sides of the two equations are equal: $\frac{0+|z_\alpha|\sqrt{\frac{2\sigma_0^2}{J}}-\Delta}{\sqrt{\frac{\sigma_0^2}{J}+\frac{\sigma_1^2}{J}}}=-|z_\beta|$. We have to solve for J in the equation above. This reduces to: $0+|z_\alpha|\sqrt{\frac{2\sigma_0^2}{J}}-\Delta=-|z_\beta|\sqrt{\frac{\sigma_0^2}{J}+\frac{\sigma_1^2}{J}}$. That is, $|z_\alpha|\sqrt{\frac{2\sigma_0^2}{J}}+|z_\beta|\sqrt{\frac{\sigma_0^2}{J}+\frac{\sigma_1^2}{J}}=\Delta$.

Next, solve for J to get $\sqrt{J} = \frac{|z_\alpha|\sqrt{2\sigma_0^2} + |z_\beta|\sqrt{\sigma_0^2+\sigma_1^2}}{\Delta}$. Since J has to be an integer, we increase J to the next integer value.

The scenario can be realistic. Suppose that there are two ratio scale random variables. The one with the greater mean typically has a greater variance. For example, the Poisson and the chi-square distributions have this property. This greater variance requires a somewhat greater sample size in study design.

This calculation assumes that there is no attrition of subjects. Typically, study attrition is large. An attrition rate less than 15% at a follow-up three or more years later is a very low attrition rate. Accounting for attrition is its own modeling effort, which can be very difficult to do well.

Problem 4 in Chapter 6 Study Guide

In a clinical trial, *2J* patients suffering from an illness will be randomly assigned to one of two groups so that *J* will receive an experimental treatment and *J* will receive the best available treatment. The random variable *X* is the response of a patient to the experimental medicine, and the random variable *B* is the response of a patient to the best currently available treatment. The random variables *X* and *B* are normally distributed and have $\sigma_X = \sigma_B = 500$ under both the null and alternative distributions. The null hypothesis to be tested is that $E(X) - E(B) = 0$ against the alternative that $E(X) - E(B) > 0$ at the 0.005 level of significance. What is the number *J* in each group that would have to be taken so that the probability of a Type II error for the test of the null hypothesis specified in the common section is 0.01 when $E(X) - E(B) = 250$?

Solution: For this specification, $\alpha = .005$, so that $|z_\alpha| = 2.576$. Also, $\beta = .01$, so that $|z_\beta| = 2.326$. With regard to variances, $\sigma_0^2 = \sigma_1^2 = 500^2$. Finally, $\Delta = 250$. Then, the design equation is $\sqrt{J} = \frac{2.576\sqrt{2\bullet 500^2} + 2.326\sqrt{2\bullet 500^2}}{250} = \frac{3466.24}{250} = 13.865 = \sqrt{192.24}$.
That is, there should be at least 193 in each group.

The magnitude of the difference in the expected values is 250, which is half of the assumed standard deviation of each group. This is called a one-half standard deviation effect. To detect a difference equal to 1/2 of the standard deviation, researchers need about 200 in each group. This is 400 total observations. One needs about 50 observations per group to detect a one standard deviation effect. That is, fewer observations are needed to detect a larger effect size.

Calculating Sample Size for One-Sample Tests

These calculations use the same approach as the two sample problem.

The problem is: A research team is conducting research about Y, which is normally distributed with standard deviation σ. They wish to test the null hypothesis $H_0 : E(Y) = \mu_0$ with standard deviation σ_0 at level of significance α against the alternative hypothesis that $H_1 : E(Y) > \mu_0$. How many observations n are necessary so that the probability of a Type II error is β when $E(Y) = \mu_0 + \Delta$ and $\sigma = \sigma_1$?

Solution: The test statistic is $TS = \bar{Y}_n$, and TS is $N(E(Y), \frac{\text{var}(Y)}{n})$. The null distribution of TS is then $N(\mu_0, \frac{\sigma_0^2}{n})$. Hence, we reject $H_0 : E(Y) = \mu_0$ against $H_1 : E(Y) > \mu_0$ at the α level of significance when $TS = \bar{Y}_n \geq \mu_0 + |z_\alpha| \sqrt{\frac{\sigma_0^2}{n}}$. When $E(Y) = \Delta > 0$ and $\text{var}(Y) = \sigma_1^2$, the (alternative) distribution of TS is then $N(\mu_0 + \Delta, \frac{\sigma_1^2}{n})$. Then, the probability of a Type II error is $\beta = \Pr_1\{\text{Accept } H_0\} = \Pr_1\{TS < \mu_0 + |z_\alpha| \sqrt{\frac{\sigma_0^2}{n}}\}$. That is,

$$\beta = \Pr_1\{TS < \mu_0 + |z_\alpha| \sqrt{\frac{\sigma_0^2}{n}}\} = \Pr\{Z = \frac{TS - (\mu_0 + \Delta)}{\sigma_1(\bar{Y}_n)} < \frac{\mu_0 + |z_\alpha| \sqrt{\frac{\sigma_0^2}{n}} - (\mu_0 + \Delta)}{\sqrt{\frac{\sigma_1^2}{n}}}\}.$$ Since

$\beta = \Pr_1\{\text{Accept } H_0\} \leq 0.5$, it is true that $\beta = \Pr\{Z < -|z_\beta|\}$. We now have two equations:

$$\beta = \Pr\{Z < \frac{\mu_0 + |z_\alpha|\sqrt{\frac{\sigma_0^2}{n}} - (\mu_0 + \Delta)}{\sqrt{\frac{\sigma_1^2}{n}}}\}$$, and

$\beta = \Pr\{Z < -|z_\beta|\}$. The problem is to choose *n* so that the probability of a Type II error is a specified value. That is, we should choose *n* so that the right hand sides of the two equations are equal: $\frac{\mu_0 + |z_\alpha|\sqrt{\frac{\sigma_0^2}{n}} - (\mu_0 + \Delta)}{\sqrt{\frac{\sigma_1^2}{n}}} = -|z_\beta|$. We have to solve for *n* in the equation above. This reduces to: $0 + |z_\alpha|\sqrt{\frac{\sigma_0^2}{n}} - \Delta = -|z_\beta|\sqrt{\frac{\sigma_1^2}{n}}$. That is, $|z_\alpha|\sqrt{\frac{\sigma_0^2}{n}} + |z_\beta|\sqrt{\frac{\sigma_1^2}{n}} = \Delta$ Next, solve for *n* to get $\sqrt{n} = \frac{|z_\alpha|\sqrt{\sigma_0^2} + |z_\beta|\sqrt{\sigma_1^2}}{\Delta}$. Since *n* has to be an integer, we increase *n* to the next integer value.

Problem 3, Study Guide Chapter 5

A research team is conducting research about Y, a student's score on a national examination, which is normally distributed with standard deviation $\sigma = 150$. They wish to test the null hypothesis $H_0 : E(Y) = 1000$ at level of significance 0.005 against the alternative hypothesis that $H_1 : E(Y) > 1000$. How many observations n are necessary so that the probability of a Type II error is 0.01 when $E(Y) = 1050$ and $\sigma = 150$?

Solution: For this specification, $\alpha = .005$, so that $|z_\alpha| = 2.576$. Also, $\beta = .01$, so that $|z_\beta| = 2.326$. With regard to variances, $\sigma_0^2 = \sigma_1^2 = 150^2$. Finally, $\Delta = 50$. Then, the design equation is $\sqrt{n} \ge \frac{2.576\sqrt{150^2} + 2.326\sqrt{150^2}}{50} = \frac{735.3}{50} = 14.706 = \sqrt{216.27}$.
That is, the sample size should be at least 217.

When you work on a problem like this, rephrase the question in more general terms. For example, the question asks how large a sample size is necessary to detect a difference of 1/3 of a standard deviation; note the 50=1050-1000=150/3. For this alpha and beta, one needs at least 217 observations, ignoring the non-response issue.

Permutation Tests

Suppose that we are comparing two groups *A* and *B*, with a random sample of 2 observations from the *A* group and an independent random sample of 3 observations for the *B* group. The *A* group observations are {1,2}, and the *B* group observations are {3, 4, 5}, with an observed t-value of -3.0.

The concept is that the null hypothesis essentially is that the two groups being compared are identical so that the group labels are arbitrary. In that event, each permutation is equally likely. The two independent sample t-test is calculated for each value as in the table below.

Table 1

Permutations and t-test values

A group permutation	B group permutation	t-value
1, 2	3, 4, 5	-3.0
1, 3	2, 4, 5	-1.22
1, 4	2, 3, 5	-0.52
1, 5	2, 3, 4	0.00
2, 3	1, 4, 5	-0.52
2, 4	1, 3, 5	0.00
2, 5	1, 3, 4	0.52
3, 4	1, 2, 5	0.52
3, 5	1, 2, 4	1.22
4, 5	1, 2, 3	3.00

The permutation distribution of the t-test values for this example is the set:

{-3.00, -1.22, -0.52, -0.52, 0.00, 0.00, 0.52, 0.52, 1.22, 3.00}. The two-sided permutation test p-value is equal to the number of permutations that have an absolute value of t greater than or equal to the absolute value of the t-test for the observed samples. Here, there are two permutations with absolute value of 3.0 or more. Then the two sided permutation test p-value is 2/10=0.2, which is not a small value. This example is just to illustrate the concept. In practice, the number of permutations is very large. For example, if there are *n* observations in the *A* sample and *m* observations in the *B* sample, then there are $\binom{n+m}{n}$ permutations.

Researchers typically take a random sample of permutations to calculate an estimated permutation test p-value.

Pitman's Findings

E.J.G. Pitman studied the permutation distribution of the two independent sample t-test results. He proved that the first four moments of the permutation distribution are asymptotically the same as the first four moments of the Student's t distribution with $n+m-2$ degrees of freedom. Consequently, the p-value from Student's t-test is a good approximation to the permutation test p-value.

Paired t-test

The most common application of the paired t-test is a comparison of the post-training score of a participant in a study with the same participant's pre-training score. The idea of the paired t-test is to calculate the difference (here post score-pre score) for each participant. This data is used in the one-sample t-test of Chapter 5 to test the null hypothesis that the expected post training score is equal to the expected pre training score.

Chapter 6 Study Guide, Problem 5

A research time wished to estimate the reduction of the density of contaminant in a liquid due to filtering the liquid. They filtered four samples, called A, B, C, and D. Find the 99% confidence interval for the expected reduction in the density of contaminant using the data in the table below:

Sample	Density of Contaminant before Filtering	Density of Contaminant after Filtering	Difference	Deviation of Difference	Deviation Squared
A	132	87	45	-4.75	22.5625
B	205	163	42	-7.75	60.0625
C	81	35	46	-3.75	14.0625
D	423	357	66	16.25	264.0625

Solution: The four differences of *D*=before filtering – after filtering are: 132-87=45, 42, 46, and 66. Then $\bar{d}_4 = \frac{45+42+46+66}{4} = 49.75$ and $s_D^2 = 120.25$ on 3 degrees of freedom. The 99% confidence interval for the expected reduction in density is from 17.7 to 81.8.

End of Chapter 6 Lecture Notes