ST 511, Exp Stat Bio Sci I
SSI 2002, Exam II
Part A, Questions 1, 2, 3
June 14, 2002

Last_Name__________________________
First_Name________________________
Student_ID_________________________

GROUND RULES

• This is an open notes, open book exam. You may use a calculator.
• Show all of your relevant work on the exam. Scratch paper is available at the front of the class.
• Be sure to put your name on both parts of this exam.
GROUND RULES

- This is an open notes, open book exam. You may use a calculator.
- Show all of your relevant work on the exam. Scratch paper is available at the front of the class.
- Be sure to put your name on both parts of this exam.

Do not write below this line.

Question 4: ______ × 4 points = ______
Question 5: ______ × 4 points = ______
Total points = ______
1. (20 points)
A popular flyfishing magazine published data from a comparative study of two flyrods (same length and line weight, but from different manufacturers). Six fishermen fished with a rod of each type for two hours, and assigned an overall performance rating to each rod on a scale from 1 to 100 (1=worst, 100=best). The ratings data appear in the table below.

<table>
<thead>
<tr>
<th>Fisherman</th>
<th>Rod A</th>
<th>Rod B</th>
<th>Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>73</td>
<td>72</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>84</td>
<td>82</td>
<td>2</td>
</tr>
<tr>
<td>3</td>
<td>86</td>
<td>83</td>
<td>3</td>
</tr>
<tr>
<td>4</td>
<td>85</td>
<td>78</td>
<td>7</td>
</tr>
<tr>
<td>5</td>
<td>85</td>
<td>77</td>
<td>8</td>
</tr>
<tr>
<td>6</td>
<td>95</td>
<td>86</td>
<td>9</td>
</tr>
</tbody>
</table>

Do the data provided evidence of a difference in mean performance rating between the two rods?

Answer

The appropriate method of analysis is a two-sample, paired-data, \( t \) test. The relevant hypothesis is \( H_0 : \mu_{\text{Rod A}} = \mu_{\text{Rod B}} \), or in terms of differences, \( H_0 : \mu_D = 0 \) where by definition \( \mu_D = \mu_{\text{Rod A}} - \mu_{\text{Rod B}} \). The alternative is two-sided, \( H_A : \mu_{\text{Rod A}} \neq \mu_{\text{Rod B}} \), or in terms of differences \( H_0 : \mu_D \neq 0 \).

Here’s the relevant row from the \( t \) table ...


\[
\begin{array}{ccccccccc}
\alpha & 0.200 & 0.150 & 0.100 & 0.050 & 0.025 & 0.010 & 0.005 & 0.001 \\
t(5, \alpha) & 0.9195 & 1.1557 & 1.4759 & 2.0150 & 2.5706 & 3.3648 & 4.0319 & 5.8924 \\
\end{array}
\]

Because \(|t_{\text{calc}}| = 3.5960 > t(5, \alpha/2) = 2.5706\), the null hypothesis of no difference in mean yields is rejected. That is, there is evidence at the \( \alpha = 0.0500 \) level of significance that Rod A and Rod B have different mean response.

EOA
2. (20 points) The North Carolina Wildlife Resources Commission stocks trout in selected streams in the North Carolina mountains. Brook, brown and rainbow trout are stocked in the proportions .4, .2 and .4 respectively. That is 40% of the trout stocked are brook trout, 20% brown trout, and 40% are rainbow. Assuming that the three types of trout are equally catchable, the probability of catching a brook, brown and rainbow trout would be .4, .2, and .4 respectively. This spring I fished a stocked stream in the mountains and in two days of fishing I caught a total of 73 trout, of which 34 were brook trout, 6 were brown trout, and 33 were rainbows. Are the numbers of trout I caught consistent with the hypothesis that the three types of trout are equally catchable?

<table>
<thead>
<tr>
<th>Type of Trout</th>
<th>Brook</th>
<th>Brown</th>
<th>Rainbow</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Numbers caught</td>
<td>34</td>
<td>6</td>
<td>33</td>
<td>73</td>
</tr>
</tbody>
</table>

Answer

The problem requires testing a completely specified set of probabilities (proportions) using a Chi-squared goodness-of-fit statistic. The null hypothesis category probabilities are

<table>
<thead>
<tr>
<th>Type of Trout</th>
<th>Brook</th>
<th>Brown</th>
<th>Rainbow</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Probabilities $\theta_i$</td>
<td>0.400</td>
<td>0.200</td>
<td>0.400</td>
<td>1</td>
</tr>
</tbody>
</table>

Under the null hypothesis probability distribution, the expected frequency for the $i^{th}$ category is estimated using the formula

$$\hat{f}_i = f \cdot \theta_i.$$ 

For example,

$$\hat{f}_1 = f \cdot \theta_1 = 73 \times 0.400 = 29.200.$$ 

Here’s the table of expected frequencies...

<table>
<thead>
<tr>
<th>Type of Trout</th>
<th>Brook</th>
<th>Brown</th>
<th>Rainbow</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Expected Numbers caught</td>
<td>29.20</td>
<td>14.60</td>
<td>29.20</td>
<td>73</td>
</tr>
</tbody>
</table>

...and the test statistic for testing equality to the hypothesized probabilities,

$$\chi^2_{calc} = \sum_{cells} \frac{(observed - expected)^2}{expected} = \frac{(34 - 29.20)^2}{29.20} + \frac{(6 - 14.60)^2}{14.60} + \cdots + \frac{(33 - 29.20)^2}{29.20} = 6.3493.$$ 

Here’s the relevant row from the $\chi^2$ tables.

$$\begin{array}{cccccccc}
\alpha & 0.200 & 0.150 & 0.100 & 0.050 & 0.025 & 0.010 & 0.005 & 0.001 \\
\end{array}$$ 

Because $5.9915 \leq \chi^2_{calc} < 7.3778$ we know that $0.025 < p \leq 0.050$, where $p$ is the attained level of significance of the test (the $p$-value).
3. (20 points) Length restrictions on “keepers” are an integral component of fisheries management regulations. (“Keepers” are fish that fisherman are allowed to keep, all other fish must be returned to the water.)

Suppose that the regulations for bass caught in Lake Endofssone state that a bass may be kept only if it is between 8 and 16 inches in length. In other words all bass caught with lengths less than 8 inches or greater than 16 inches must be returned to the lake (alive!).

Assume that the lengths of bass in Lake Endofssone are normally distributed with a mean of 12 inches and a standard deviation of 4 inches.

A fisherman catches 25 bass from the lake, of which only 5 are keepers. (Hint: the fisherman thinks 25 is a large number of fish.)

(a) Are the fisherman’s data consistent with what is assumed about the distribution of lengths of bass in Lake Endofssone?

**Answer**

This calls for a hypothesis test. But first a preliminary calculation is required. Let $W$ denote the length of a bass in Lake of Endofssone. Based on the information given, the probability of a keeper from Lake Endofssone is

$$\Pr(\text{Keeper}) = \Pr(8 < W < 16) = \Pr((8 - 12)/4 < Z < (16 - 12)/4) = \Pr(-1 < Z < 1) = 0.6826.$$ 

Now let $Y$ denote the number of keepers in a sample of $n = 25$ bass. Then $Y$ has Binomial($25, \pi$) distribution. The question calls for testing $H_0 : \pi = 0.6826$ versus $H_A : \pi \neq 0.6826$ based on the observed catch of 5 keepers in 25 bass caught. The point estimate of $\pi$ is $\hat{\pi} = 5/25 = 0.2$. The relevant test statistic is

$$Z_{calc} = \frac{0.2 - 0.6826}{\sqrt{0.6826(1 - 0.6826)/25}} = -5.18,$$

which is so large (in absolute value) that the null hypothesis is rejected at any of the usual levels of significance.

EOA

(b) Using the fisherman’s data, construct a 95% confidence interval for the proportion of bass in Lake Endofssone that are “keepers.”

**Answer**

The point estimate of $\pi$ is $\hat{\pi} = 5/25 = 0.2$. The required confidence interval is

$$0.2 \pm 1.96\sqrt{\frac{0.2(1 - 0.2)}{25}}$$

$$0.2 \pm 1.96(0.08)$$

$$0.2 \pm 0.1568$$

$$0.0432, 0.3568$$

EOA
4. (20 points) We studied the following 10 methods in Chapters 5 and 6:

Box 4.2 Inferences about $\mu$;
Box 4.3 Inferences about $\mu_1 - \mu_2$ with independent samples and $\sigma_1^2 = \sigma_2^2$;
Box 4.4 Inferences about $\mu_1 - \mu_2$ with independent samples and $\sigma_1^2 \neq \sigma_2^2$;
Box 4.5 Inferences about $\mu_1 - \mu_2$ with paired observations;
Box 4.6 Inferences about $\sigma^2$;
Box 4.7 Inferences about the ratio $\sigma_1^2/\sigma_2^2$;
Box 6.1 Large-sample inferences about a population proportion $\pi$;
Box 6.3 Inferences about a completely specified set of proportions;
Box 6.5 Large-sample inferences about the difference between two proportions $\pi_1 - \pi_2$;
Box 6.8 Large-sample test of the equality of distribution of two categorical populations.

Read each of the five problem descriptions (a) . . . (e) below and on the following pages, and decide what the most appropriate method of analysis is. Do not do the analysis! I simply want to know what method is most appropriate.

(a) The data below are from a study designed to compare no-till and conventional till methods. The response variable is crop yield in the growing season after preparation with either conventional or no-till methods. The experimental design used two fields (one prepared using no-till methods the other using conventional tilling methods) from each of six farms. The fields were planted with a crop and yields were measured.

<table>
<thead>
<tr>
<th>Farm</th>
<th>Tilling Method</th>
<th>Conventional</th>
<th>No-Till</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>11</td>
<td>12</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>13</td>
<td>10</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>33</td>
<td>29</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>21</td>
<td>16</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>17</td>
<td>11</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>24</td>
<td>17</td>
<td></td>
</tr>
</tbody>
</table>

Do the data indicate a difference in the mean crop yield of fields prepared with conventional and no-till methods? Circle the answer corresponding to the most appropriate method for this problem.

[Box 4.2] [Box 4.3] [Box 4.4] [Box 4.5] [Box 4.6] [Box 4.7] [Box 6.1] [Box 6.3] [Box 6.5] [Box 6.8]

Answer

The appropriate method of analysis is a two-sample, paired-data, $t$ test. The relevant hypothesis is $H_0 : \mu_{\text{CONVENTIONAL}} = \mu_{\text{NO-TILL}}$, or in terms of differences, $H_0 : \mu_D = 0$ where by definition $\mu_D = \mu_{\text{CONVENTIONAL}} - \mu_{\text{NO-TILL}}$. The alternative is two-sided, $H_A : \mu_{\text{CONVENTIONAL}} \neq \mu_{\text{NO-TILL}}$, or in terms of differences $H_0 : \mu_D \neq 0$. 
### Tilling Method

<table>
<thead>
<tr>
<th>Farm</th>
<th>Conventional</th>
<th>No-Till</th>
<th>Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>11</td>
<td>12</td>
<td>-1</td>
</tr>
<tr>
<td>2</td>
<td>13</td>
<td>10</td>
<td>3</td>
</tr>
<tr>
<td>3</td>
<td>33</td>
<td>29</td>
<td>4</td>
</tr>
<tr>
<td>4</td>
<td>21</td>
<td>16</td>
<td>5</td>
</tr>
<tr>
<td>5</td>
<td>17</td>
<td>11</td>
<td>6</td>
</tr>
<tr>
<td>6</td>
<td>24</td>
<td>17</td>
<td>7</td>
</tr>
</tbody>
</table>

\[
\sum D_i = 24.000000 \\
\bar{D} = 4.0000000 \\
\sum (D_i - \bar{D})^2 = 40.000000 \\
\begin{align*}
   s_D &= 2.8284271 \\
   t_{calc} &= \left(\sqrt{6}\right)\bar{D}/s_D \\
   &= 3.4641016
\end{align*}
\]

Here’s the relevant row from the *t* table ...

<table>
<thead>
<tr>
<th>(\alpha)</th>
<th>0.200</th>
<th>0.150</th>
<th>0.100</th>
<th>0.050</th>
<th>0.025</th>
<th>0.010</th>
<th>0.005</th>
<th>0.001</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>t</em>(5, (\alpha))</td>
<td>0.9195</td>
<td>1.1557</td>
<td>1.4759</td>
<td>2.0150</td>
<td>2.5706</td>
<td>3.3648</td>
<td>4.0319</td>
<td>5.8924</td>
</tr>
</tbody>
</table>

Because \(|t_{calc}| = 3.4641 > t(5, \alpha/2) = 2.5706\), the null hypothesis of no difference in mean yields is rejected. That is, there is evidence at the \(\alpha = 0.0500\) level of significance that Conventional and No-Till have different mean response.

---

**EOA**
(b) In the late 1800s H.P. Bowditch measured the heights of several thousand Massachusetts school children. His data for \( n = 1600 \) 11-year-old boys had a mean of 52.9 inches and a standard deviation of 6.2 inches. Do these data provide evidence that the mean height of 11-year-old boys in the 1800s exceeded 52.6 inches?

Circle the answer corresponding to the most appropriate method for this problem.

[Box 4.2] [Box 4.3] [Box 4.4] [Box 4.5] [Box 4.6]
[Box 4.7] [Box 6.1] [Box 6.3] [Box 6.5] [Box 6.8]

**Answer**

The hypotheses to be tested are:

\[ H_0 : \mu = 52.6 \quad \text{versus} \quad H_A : \mu > 52.6. \]

The standard error of the sample mean is

\[ \hat{\sigma}_Y = \frac{s}{\sqrt{n}} = \frac{6.2}{\sqrt{1600}} = 0.155, \]

and so the relevant test statistic is

\[ t_{calc} = \frac{\bar{Y} - \mu_0}{\hat{\sigma}_Y} = \frac{52.9 - 52.6}{0.155} = 1.94. \]

This level-0.05 test rejects \( H_0 \) when \( t_{calc} \) exceeds \( t(1599, 0.05) \approx z(0.05) = 1.65 \). Thus \( H_0 \) is rejected in favor of \( H_A \); that is, there is evidence at the 0.05 level of significance that the mean height of 11-year-old boys in the 1800s exceeded 52.6 inches.

EOA

(c) The Federal Tobacco Commission (FTC) studied the nicotine content of two brands of cigarettes, Cough Lites and Choke Lights, and recorded the following summary statistics.

<table>
<thead>
<tr>
<th>Sample Size</th>
<th>Mean (mg.)</th>
<th>Standard Deviation (mg.)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cough Lites</td>
<td>( n_1 = 16 )</td>
<td>( \bar{Y}_1 = 0.6 )</td>
</tr>
<tr>
<td>Choke Lights</td>
<td>( n_2 = 16 )</td>
<td>( \bar{Y}_2 = 0.9 )</td>
</tr>
</tbody>
</table>

The FTC has determined that variability in nicotine content of the two brands is similar and would now like to determine whether the data support the claim that the mean nicotine content of Cough Lites is less than that of Choke Lights?

Circle the answer corresponding to the most appropriate method for this problem.
Answer

First calculate the pooled estimate of variance

\[
\begin{align*}
    s_{\text{Pooled}}^2 &= \frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2} \\
    &= \frac{(15)0.04 + (15)0.09}{30} \\
    &= 0.065.
\end{align*}
\]

The indicated test statistic is

\[
\begin{align*}
    t_{\text{calc}} &= \frac{0.6 - 0.9}{s_{\text{Pooled}} \sqrt{1/16 + 1/16}} \\
    &= \frac{-0.3}{0.255\sqrt{1/16 + 1/16}} \\
    &= -0.3 \times \frac{10}{\sqrt{2.55}} \\
    &= -0.3 \times \frac{10}{1.60} \\
    &= -1.886.
\end{align*}
\]

The hypotheses to be tested are \( H_0 : \mu_1 - \mu_2 = 0 \) versus \( H_A : \mu_1 - \mu_2 < 0 \) and the test rejects \( H_0 \) in favor of \( H_A \) when \( t_{\text{calc}} < -t(30, 0.05) = -1.6973 \). So there is not sufficient evidence to conclude that the nicotine content of Cough Lites is less than that of Choke Lights at the \( \alpha = 0.05 \) level of significance.
(d) The toxicant potassium cyanate was applied to two vials of fertilized trout fish eggs. For one vial the toxicant was applied immediately after fertilization. For the other vial the eggs were allowed to water-harden for six hours after fertilization before the toxicant was applied. The total numbers of eggs and the numbers of living eggs 19 days after fertilization were counted in each vial. There were 463 non-water-hardened eggs of which 274 were living; there were 431 water-hardened eggs of which 394 were living. Do these data provide evidence supporting the conclusion that the survival probability of water-hardened eggs is greater than that of non water-hardened eggs?

Circle the answer corresponding to the most appropriate method for this problem.

[Box 4.2] [Box 4.3] [Box 4.4] [Box 4.5] [Box 4.6]
[Box 4.7] [Box 6.1] [Box 6.3] [Box 6.5] [Box 6.8]

Answer

Define \( \pi_1 \) = the probability of survival of a non-water-hardened egg, and \( \pi_2 \) = the probability of survival of a water-hardened egg. The hypotheses to be tested are \( H_0 : \pi_1 - \pi_2 = 0 \) versus \( H_0 : \pi_1 - \pi_2 < 0 \). Let \( n_1 = 463, Y_1 = 274, n_2 = 431, Y_2 = 394, \hat{\pi}_1 = 274/463 = 0.59179266, \hat{\pi}_2 = 394/431 = 0.91415313 \). The relevant test statistics is

\[
Z_{calc} = \frac{\hat{\pi}_1 - \hat{\pi}_2}{\sqrt{\hat{\pi}_1(1 - \hat{\pi}_1)/n_1 + \hat{\pi}_2(1 - \hat{\pi}_2)/n_2}}
\]

\[
= \frac{0.59179266 - 0.91415313}{\sqrt{0.59179266(1 - 0.59179266)/463 + 0.91415313(1 - 0.91415313)/431}}
\]

\[= -12.150799.\]

The null hypothesis is rejected in favor of the alternative when \( Z_{calc} < -1.645 \). Thus there is evidence to conclude that water-hardened eggs have higher survival probability at the \( \alpha = .05 \) level of significance.

EOA

(e) Data collected in order to obtain an interval estimate of the relative variability in heights of 11-year-old male and female children were summarized and appear in table below.

<table>
<thead>
<tr>
<th></th>
<th>Sample Size</th>
<th>Mean (in.)</th>
<th>Standard Deviation (in.)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Male</td>
<td>( n_1 = 16 )</td>
<td>( \bar{Y}_1 = 54 )</td>
<td>( s_1 = 8 )</td>
</tr>
<tr>
<td>Female</td>
<td>( n_2 = 16 )</td>
<td>( \bar{Y}_2 = 51 )</td>
<td>( s_2 = 4 )</td>
</tr>
</tbody>
</table>

Circle the answer corresponding to the most appropriate method for this problem.
Answer

The phrase “interval estimate of the relative variability” says it all. The appropriate methods are given in Box 4.7. The problem is to calculate a confidence interval for $\frac{\sigma_1^2}{\sigma_2^2}$ where $\sigma_1^2$ ($\sigma_2^2$) is the population variance of heights of 11-year-old male (female) children.

EOA
5. (20 points) Some of the statements below make sense (they are correct, reasonable, consistent) and
some do not make sense (they are incorrect, unreasonable, inconsistent). Indicate which statements
make sense and which do not (circle the appropriate answer).

(a) “We found a statistically significant difference between treatment and control groups (p-value < .0003).”

consistent/correct inconsistent/incorrect

Answer
Consistent.

(b) “The p-value of the test was large (p-value > .6552) indicating a high degree of confidence in the results.”

consistent/correct inconsistent/incorrect

Answer
Inconsistent. A p-value is not a measure of confidence.

(c) “The 95% confidence interval for the difference in treatment means is (2.34, 4.73), indicating no
difference between mean treatment response.”

consistent/correct inconsistent/incorrect

Answer
Inconsistent. The 95% confidence does not contain 0, indicating that the treatment means are
significantly different from 0.

(d) “The p-value of a hypothesis test is a measure of the strength of evidence against the null hypo-
thesis in the direction of the alternative hypothesis.”

consistent/correct inconsistent/incorrect

Answer
Consistent.

(e) “… the value of the test statistic is $t_{calc} = 0.37$ (with 72 degrees of freedom) with a p-value < .0001 suggesting that …”

consistent/correct inconsistent/incorrect

Answer
Inconsistent. A $t_{calc}$ test statistic near 0 will have a p-value close to 0.5.