

Practice Final Exam Questions (2) -- Answers

Part A. Multiple Choice Questions. For each question, you are encouraged to give a reason or show work for partial credit. You must show your work or reason if the question is marked with an asterisk (*).

- Confidence intervals are useful when trying to estimate _____.
 - unknown parameters ←
 - known parameters
 - unknown statistics
 - known statistics
 - The one-sample z statistic is used instead of the one-sample t statistic when _____.
 - μ is known
 - μ is unknown
 - σ is known ←
 - σ is unknown
 - The ____ the P-value, the stronger the evidence against the null hypothesis provided by the data.
 - larger
 - smaller ←
 - (*) The test statistic for a two-sided significance test for a population mean is $z = -2.12$. What is the corresponding P-value?
 - 0.017
 - 0.034 ←
 - 0.483
 - 0.983
- Answer: $P(Z < -2.12) = 0.0170$. Since this is a two-sided test, P-value is twice as much to include the other tail. $0.0170 * 2 = 0.034$.
- The probability you reject the null hypothesis when in fact the null hypothesis is true is called _____.
 - a Type I error ←
 - a Type II error
 - the power
 - (*) A random sample of 20 observations produced a sample mean of $\bar{x} = 92.4$ and $s = 25.8$. What is the value of the standard error of \bar{x} ?
 - 4.6
 - 15.9
 - 1.3
 - 5.8 ←
 - 2.6

Answer: $SE_{\bar{x}} = \frac{25.8}{\sqrt{20}} = \frac{25.8}{4.472} = 5.8$

7. (*) The heights (in inches) of adult males in the United States are believed to be Normally distributed with mean μ . The average height of a random sample of 25 American adult males is found to be $\bar{x} = 69.72$ inches, and the standard deviation of the 25 heights is found to be $s = 4.15$. A 90% confidence interval for μ is
- 69.72 ± 1.09
 - 69.72 ± 1.37
 - 69.72 ± 1.42 ←

Answer: The t critical value for 90% CI for $df = 24$ is 1.711. So

$$\text{margin_of_error} = (1.711) \cdot \frac{4.15}{\sqrt{25}} = (1.711) \cdot (0.83) = 1.42$$

8. Suppose we were interested in determining if there were differences in the average prices among two local supermarkets. We randomly pick six items to compare at both supermarkets. Which statistical procedure would be best to use for this study?
- Matched-pairs t procedure ← if you interpreted as 'same six items'
 - One-sample t test
 - Two-sample t test ← if you interpreted as 'different six items'
 - None of the above
9. (*) Perform a one-sample t-test using the following statistics:

$$n = 5 \quad \bar{x} = 3.871 \quad s = 0.679$$

The null hypothesis is $\mu = 5.0$ is

- accepted at the 5% level; accepted at the 1% level.
- accepted at the 5% level; rejected at the 1% level.
- rejected at the 5% level; accepted at the 1% level. ←
- rejected at the 5% level; rejected at the 1% level.

Answer: Test statistic: $t = \frac{\bar{x} - \mu}{\frac{s}{\sqrt{n}}} = \frac{3.871 - 5.0}{\frac{0.679}{\sqrt{5}}} = -3.71$

Assuming a two-sided test, for $df = 4$,

- 95% CI of the t-statistic assuming H_0 is $[-2.776, 2.776]$
- 99% CI of the t-statistic assuming H_0 is $[-4.604, 4.604]$

Therefore we reject H_0 for $\alpha = 0.05$ but not for $\alpha = 0.01$.

10. (*) You buy a package of 122 Smarties and 19 of them are red. What is a 95% confidence interval for the true proportion of red Smarties?
- $(0.092, 0.220)$ ←
 - $(0.103, 0.230)$
 - $(0.085, 0.199)$

Answer: $\hat{p} = \frac{19}{122} = 0.156$. $\text{margin_of_error} = (1.96) \cdot \sqrt{\frac{(0.156)(0.844)}{122}} = (1.96) \cdot (0.033) = 0.0647$.

So the 95% CI is $[0.156 - 0.0647, 0.156 + 0.0647]$, which yields $[0.0913, 0.2207]$.

11. (*) It is thought that 12% of all students taking a particular course received a grade of A. In a sample of 155 students, it is found that 21 made an A. What is the test statistic for testing the true proportion is 12%?
- a. 0.53
 - b. 0.01
 - c. 0.577 ←

Answer: $\hat{p} = \frac{21}{155} = 0.135$. $Z = \frac{0.135 - 0.12}{\sqrt{\frac{(0.12)(0.88)}{155}}} = \frac{0.015}{0.026} = 0.577$

Part B. Short Answer Questions

1. [Textbook Exercise 5.30, 5.32] A tablet PC contains 3217 music files. The distribution of file size is highly skewed with many small file sizes. Assume that the standard deviation for this population is 3.25 megabytes (MB).

- a. What is the standard deviation of the average file size when you take an SRS of 25 files from this population?

Answer: $\sigma_{\bar{x}} = \frac{3.25}{\sqrt{25}} = 0.65$

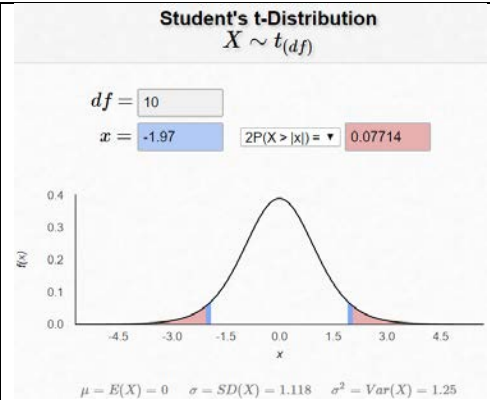
- b. Suppose you plan to take an SRS of $n = 50$ files instead. Explain why it may be reasonable to assume that the average \bar{x} is approximately Normal even though the population distribution is highly skewed.

Answer: Because of the Central Limit Theorem (CLT), we can assume the sample mean is normally distributed when the sample size is large enough. In this case, $n = 50$ is greater than 30, thus the CLT applies.

2. [Textbook Exercise 7.57] For each of the following, answer the question and give a short explanation of your reasoning.

- a. A significance test for comparing two means gave $t = -1.97$ with 10 degrees of freedom. Can you reject the null hypothesis that the μ 's are equal versus the two-sided alternative at the 5% significance level?

Answer:

 <p style="font-size: small; margin-top: 5px;">$\mu = E(X) = 0$ $\sigma = SD(X) = 1.118$ $\sigma^2 = Var(X) = 1.25$</p>	<p>Probability for $t = -1.97$ for $df = 10$ is between 0.025 and 0.05. So for two-sided test, the p-value is between 0.05 and 0.1. Therefore we fail to reject H_0 at the significance level $\alpha = 0.05$.</p> <p>The graph on the left is calculated by a software. You see the exact P-value (0.07714).</p>
---	---

- b. Answer part (a) for the one-sided alternative that the difference between means is negative.

Answer: If the alternative hypothesis was $H_a: \mu_1 - \mu_2 < 0$, we can reject the null hypothesis at alpha = 0.05 level because the p-value would be between 0.025 and 0.05.

=====

Part C. Problems

1. **One-sample t-test.** To test the hypothesis that eating fish makes one smarter, a random sample of 12 persons take a fish oil supplement for one year and then are given an IQ test. Here are the results:

116 111 101 120 99 94 106 115 107 101 110 92

Test using the following hypotheses, report the test statistic with the P-value, then summarize your conclusion.

$H_0: \mu = 100$

$H_a: \mu > 100$

Answer:

Hypotheses:

$H_0: \mu = 100$ (no effect -- eating fish does not help increase the mean IQ)

$H_a: \mu > 100$ (effect -- eating fish helps increase the mean IQ)

Test statistic:

From the data, we obtain $\bar{x} = 106$ and $s_x = 8.83$. Then we get

$$t = \frac{\bar{x} - \mu_0}{\frac{s_x}{\sqrt{n}}} = \frac{106 - 100}{\frac{8.83}{\sqrt{12}}} = \frac{6}{2.55} = 2.35.$$

P-value:

Because $n = 12$, we use the t distribution with $df = 11$ to find the probability. According to Table D, for $t = 2.35$ for $df = 11$, the probability is between 0.01 and 0.02. Since this is a one-sided (upper-tail) test, **P-value is between 0.01 and 0.02.**

Conclusion:

Since values between 0.01 and 0.02 are < 0.05 , we **reject the null hypothesis at the significance level 0.05**, and conclude that the fish oil supplement did make a significant increase in the mean IQ.

2. **Matched one-sample t-test.** The water diet requires you to drink 2 cups of water every half hour from when you get up until you go to bed but eat anything you want. Four adult volunteers agreed to test this diet. They are weighed prior to beginning the diet and 6 weeks after. Their weights in pounds are

Person	1	2	3	4	mean	s.d.
Weight before	180	125	240	150	173.75	49.56
Weight after	170	130	215	152	166.75	36.09
Difference	10	-5	25	-2	7	13.64

Conduct a one-sample t-test using the difference with the following hypotheses:

H₀: Diff = 0

H_a: Diff ≠ 0

Report the test statistic with the P-value, then summarize your conclusion.

Answer:

Hypotheses:

H₀: Diff = 0 (no difference -- there is no difference in the mean of the weight difference)

H_a: Diff ≠ 0 (difference – diet made difference in the mean of the weight difference)

Test statistic:

From the data, we know $\overline{Diff} = 7$ and $s_{Diff} = 13.64$. Then we get

$$t = \frac{\overline{Diff} - \mu_0}{\frac{s_{Diff}}{\sqrt{n}}} = \frac{7 - 0}{\frac{13.64}{\sqrt{4}}} = \frac{7}{6.82} = 1.026.$$

P-value:

Because n = 4, we use the t distribution with df = 3 to find the probability.

According to Table D, for t = 1.026 for df = 3, the probability is between 0.15 and 0.20.

Since this is a two-sided test, **P-value is between 0.30 and 0.40.**

Conclusion:

Since values between 0.30 and 0.40 are > 0.05, we fail to reject the null hypothesis at the 0.05 significance level. We do not have enough evidence to conclude that the water diet has an impact on the weight.

3. **Two-sample t procedure.** Two different alloys are being considered for making lead-free solder used in the wave soldering process for printed circuit boards. A crucial characteristic of solder is its melting point, which is known to follow a Normal distribution. A study was conducted using a random sample of 21 pieces of solder made from each of the two alloys. In each sample, the temperature at which each of the 21 pieces melted was determined. The mean and standard deviation of the sample for Alloy 1 were $\bar{x}_1 = 218.9^\circ\text{C}$ and $s_1 = 2.7^\circ\text{C}$; for Alloy 2 the results were $\bar{x}_2 = 215.5^\circ\text{C}$ and $s_2 = 3.6^\circ\text{C}$. If we were to test $H_0: \mu_1 = \mu_2$ against $H_a: \mu_1 \neq \mu_2$, what would be the value of the test statistic??

Answer:

The two-sample test statistic (for the hypothesis $\mu_1 - \mu_2 = 0$) is

$$t = \frac{(\bar{x}_1 - \bar{x}_2)}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}} = \frac{218.9 - 215.5}{\sqrt{\frac{(2.7)^2}{21} + \frac{(3.6)^2}{21}}} = \frac{3.4}{\sqrt{0.347 + 0.617}} = \frac{3.4}{\sqrt{0.964}} = \frac{3.4}{0.982} = 3.46$$

And the degrees of freedom is $21 - 1 = 20$.