

Statistics 515 Practice Exam 2 Answers

Part I: Answer seven of the following eight questions. If you complete more than seven, I will grade only the first seven. Five points each.

1)

$$P(X \geq 105) = P(X \geq 104.5) = P\left(\frac{X - np}{\sqrt{np(1-p)}} \geq \frac{104.5 - 100}{\sqrt{200(.5)(.5)}}\right) \approx P(Z \geq 0.6363961) = .5 - .2389 = .2611$$

2) Define what is meant by "Type I error". **The error of rejecting a true null hypothesis.**

3) Define what is meant by "the p-value (or observed significance level) of a test".

The probability of observing a test statistic at least as extreme as the one observed if the null hypothesis is true.

4) A variety of methods can be used to check if a sample came from a population that follows a normal distribution. Which type of graphical display was designed specifically to help address this problem?

Q-Q plot or Normal Probability Plot.

5) If you increase the sample size, you will be **more** likely to reject a false H_0 in a test of hypotheses. If you increase α , you will be **more** likely to reject a false H_0 in a test of hypotheses.

6) A sample of size 20 results in $\bar{x} = 10.0$ and $s = 2.4$. Assume that the necessary assumptions are met and construct a 95% CI for μ .

$$\bar{x} \pm t_{\alpha/2} \frac{s}{\sqrt{n}} = 10.0 \pm 2.093 \frac{2.4}{\sqrt{20}} = 10.0 \pm 1.1 = (8.9, 11.1)$$

7) The hypothesis test and confidence intervals about a single variance are conducted using the χ^2 distribution. The hypothesis test to compare to variances is conducted using the **F** distribution. In both cases the test is **not very robust**.

8) At $\alpha=0.10$ we **reject** the null hypothesis that $\mu_N=\mu_S$. At $\alpha=0.01$ we **accept** the null hypothesis that $\mu_N=\mu_S$.

		T-Tests			
Variabl e val ue	Method Pool ed	Vari ances Equal	DF 20	t Val ue 1.76	Pr > t 0.0932

Part II: Answer every part of the next two problems. Read each problem carefully, and show your work for full credit. Twenty points each.

1) A produce broker will deal with apples from an orchard only if he is quite certain that they are larger than 2.5 inches in diameter on average. To test this he randomly takes a sample of size 12 from the orchard's stock.

A) State the appropriate null and alternate hypotheses for testing if the broker will deal with the orchard's apples. **$H_0: \mu=2.5$ (don't buy) vs. $H_A: \mu>2.5$ (buy) where μ is the true average diameter of the orchard's apples.**

B) A sample of size 12 is acquired and it has a mean of 2.758 and a standard deviation of 0.3942. Test the hypothesis in A at an $\alpha=0.01$ level and state what the broker should do in this case.

$$t = \frac{\bar{x} - \mu}{s/\sqrt{n}} = \frac{2.758 - 2.5}{0.3942/\sqrt{12}} = 2.267 \quad \text{Compare this to 2.718 and fail to reject the null hypothesis. The broker should NOT buy}$$

C) Besides the sample being randomly chosen, what other assumption(s) are required to trust the test in part B? If possible, check that the assumption(s) hold. **The population of apple weights at the orchard must be normally distributed. We cannot check this without having the data.**

2) A candidate for political office wants to determine if there is a difference in his popularity between men and women. To test this he collects a sample of 250 men and 250 women and records how many of them plan on voting for him in the upcoming election.

A) State the appropriate null and alternate hypothesis for determining whether the candidate differs in popularity between men and women.

$H_0: p_m=p_f$ (equal popularity) vs. $H_A: p_m \neq p_f$ (different popularity) where p_m is the percentage of men in the population who would vote for the candidate and p_f is the percentage of women.

B) Of those sampled, 105 of the men and 128 of the women plan on voting for the candidate. Report the p-value for the test of hypothesis in A.

$$\hat{p}_m = \frac{105}{250} = 0.420 \quad \hat{p}_f = \frac{128}{250} = 0.512 \quad \bar{p} = \frac{n_m \hat{p}_m + n_f \hat{p}_f}{n_m + n_f} = \frac{105 + 128}{250 + 250} = 0.466$$

$$z = \frac{(\hat{p}_m - \hat{p}_f) - (p_m - p_f)}{\sqrt{\frac{\bar{p}(1-\bar{p})}{n_m} + \frac{\bar{p}(1-\bar{p})}{n_f}}} = \frac{(0.420 - 0.512) - 0}{\sqrt{\frac{0.466(1-0.466)}{250} + \frac{0.466(1-0.466)}{250}}} = \frac{-0.092}{0.0446} = -2.06$$

Comparing this to a z table we get that the area below -2.06 is 0.5 - 0.4803 = 0.0197, doubling this we get the p-value is 0.0394.

C) Besides the sample being randomly chosen, what other assumption(s) are required to trust the test in part B? If possible, check that the assumption(s) hold.

The sample size needs to be large enough.

$$n_m \hat{p}_m = 105, n_m(1 - \hat{p}_m) = 145, n_f \hat{p}_f = 128, n_f(1 - \hat{p}_f) = 122$$

are all greater than 5.