Exam 2 Solutions

Statistics 401

November 3, 2006

TRUE/FALSE: (3 pts each) For each of the following, circle T or F. You need not justify your answer.

1. **F** The following is a valid pair of statistical hypotheses: $\text{H}_0\colon \pi > 0.6$ versus $\text{H}_a\colon \pi = 0.7$.
2. **F** The following is a valid pair of statistical hypotheses: $\text{H}_0\colon \overline{X} = 4$ versus $\text{H}_a\colon \overline{X} > 4$.
3. **F** The larger the p-value, the stronger the evidence that H_0 is true.
4. **F** Rejecting H_0 implies H_a is true. (Rejecting H_0 only SUGGESTS that H_a is true)
5. **F** A Type II Error is made if a false null hypothesis is rejected.
6. **F** The standard deviation of the population distribution decreases as the sample size n increases.
7. **F** The probability of a Type I Error decreases as the sample size increases.
8. **T** $\overline{X}_1 - \overline{X}_2$ is an unbiased estimator of the population mean $\mu_1 - \mu_2$.
9. **F** A t distribution has skinnier tails than the standard normal distribution.
10. **T** The Power of a test increases as α increases.

MULTIPLE CHOICE: (3 pts each) Circle the single best answer.

11. When concerned about normality of the sampling distribution of $\overline{X}$, one should consider transforming a sample if

    E. The sample size is small and the data does not come from a normal distribution.

12. If a 95% confidence interval for μ is (452, 470), then the probability is ______ that the true mean μ falls in the interval (452, 470).

    E. Either 0 or 1, but cannot determine which one is true.

13. Consider statements I, II, III, and IV. Which pair of statements is **TRUE**? Given the sample standard deviation s remains constant, the confidence interval for μ will become

I.	Wider as the sample size increases.
II.	Narrower as the sample size increases.
III.	Wider as the confidence level increases.
IV.	Narrower as the confidence level increases.

D. II. and III.

14. If the population distribution for some variable X is N(μ, σ), then the sample mean $\overline{X}$ of from a small simple random sample is distributed as _____________. Assume the population size is *small* relative to the sample size.

B. N(μ, $\frac{\sigma}{\sqrt{n}}$)

15. For a random sample of 9 women, the average resting pulse rate is $\bar{x} = 76$ beats per minute and the sample standard deviation is $s = 5$ beats per minute. The *standard error* of the sample mean is

C. 1.667

SHORT ANSWER AND COMPUTATION: Show all work to receive full credit! Write legibly.

16. (4 pts) For a random sample of 9 women, the average resting pulse rate is $\bar{x} = 76$ beats per minute and the sample standard deviation is $s = 5$ beats per minute. A 95% confidence interval for μ is (72.6, 79.4). Interpret this CI in **terms of the problem**.

I am 95% confident that the average resting pulse rate for women is between 72.6 and 79.4 beats per minute.

17. (4 pts) State the Central Limit Theorem.

For large simple random simples, $\overline{X} \dot\sim N\left(\mu_x, \frac{\sigma_x}{\sqrt{n}}\right)$.

18. (4 pts) A real estate agent would like to estimate the proportion of offices in a large city that are currently occupied. How large of a sample is required to ensure that the estimate will be accurate to within 0.1 with 99% confidence?

$n = \frac{1}{2}\left(1 - \frac{1}{2}\right)\left(\frac{2.576}{.1}\right)^2 \approx 165.89$. Thus, a sample size of 166 is recommended.

19. (4 pts) A study released in September of 2006 found that 70% of the people who worked at ground zero in New York City after the terrorist attacks on September 11, 2001 suffer severe respiratory problems. Suppose that $\pi = 0.7$. If a random sample of $n = 40$ workers at ground zero are selected, give the sampling distribution of p, the sample proportion out of these 40 who have severe respiratory problems. Be sure to give the mean and standard deviation of p.

20. Suppose that a population of $N = 3$ female wolves have each had a litter of pups of size X=2, 3 and 6 pups respectively. An investigator plans on taking a random sample of $n = 2$ wolves (without replacement) from these three and then computing the largest value in the sample. Denote the largest value of the sample by M. For example, if the sample of wolves has $X = 2, 3$ then the largest value in this sample is $M = 3$.

    (a) (4 pts) Construct the sampling distribution of M, sampling without replacement. There are 3 different samples of size 2!

Sample	M	$P(M)$
$\{2, 3\}$	3	$\frac{1}{3}$
$\{2, 6\}$	6	$\frac{1}{3}$
$\{3, 6\}$	6	$\frac{1}{3}$

    (b) (3 pts) Calculate the mean of M, μ_M.
    $\mu_M = \sum_i x_i P(x_i) = 3\left(\frac{1}{3}\right) + 6\left(\frac{2}{3}\right) = 5.$

    (c) (3 pts) Is M unbiased for the largest value in the population? Justify your answer.
    The largest value in the population is 6 and $\mu_M = 5$. Since $\mu_M = 5 \neq 6$, then M is a biased statistic.

21. An industrial plant claims to discharge no more than 1000 gallons of waste water per hour, on the average, into a neighboring lake. The Environmental Protection Agency decides to monitor the plant, in case this limit is being exceeded. Based on a random sample of size $n = 44$ they find $\bar{x} = 1021$ gallons and $s = 200$ gallons.

(a) (2 pts) What type of study is this?

Observational Study

(b) (4 pts) State the null and alternative hypotheses.

$\text{H}_0 : \mu = 1000$ versus $\text{H}_a : \mu > 1000$.

(c) (4pts) What assumptions do you need to check before conducting a hypothesis test? Are these assumptions satisfied?

Since the sample size is $n = 44 > 30$, then we may invoke the Central Limit Theorem and assume that $\overline{X}$ is normally distributed.

(d) Conduct the hypothesis test.

i. (3 pts) Calculate the value of the test statistic.
$t = \frac{1021-1000}{\frac{200}{\sqrt{44}}} \approx .6965$.

ii. (2 pts) Give the distribution of the test statistic assuming that the null hypothesis is true.
$T \sim t(43)$

iii. (2 pts) Give the p-value.
The p-value is $P(T > .6965) = .25$ which we approximated with .1.

iv. (2 pts) Make a decision. Use a significance level of $\alpha = .05$.
Since the p-value$= .25 > .05$, then FAIL TO REJECT the null hypothesis.

v. (4 pts) Make a conclusion in terms of the problem.
The evidence fails to suggest that the mean volume of waste water discharged into the lake is larger than 1000 gallons.

(e) (4 pts) Explain what a Type I Error is in terms of this problem.

A Type I Error is finding that the industrial plant of is discharging more than 1000 gallons into the lake when it actually is not.

(f) (2 pts) For this problem, what is the probability that a Type I Error occurs?

$P(\text{Type I Error}) = \alpha = .05$.