

Statistical Reasoning - Collecting and Analyzing Data Test Review 3

1. In formulating hypotheses for a statistical test of significance, the null hypothesis is often

(a) A statement that there is "no effect" or "no difference"
 (b) Proven correct
 (c) A statement that the data are all 0
 (d) 0.05
 (e) The probability of observing the data you actually obtained
2. A medical researcher is working on a new treatment for a certain type of cancer. The average survival time after diagnosis for patients on the standard treatment is two years. In an early trial, she tries the new treatment on three subjects who have an average survival time after diagnosis of four years. Although the survival time has doubled, the results are not statistically significant, even at the 0.10 significance level. The best explanation for this result is that

(a) The placebo effect is present, which limits statistical significance
 (b) The sample size is too small to determine if the observed increase cannot be reasonably attributed to chance
 (c) Although the survival time has doubled, in reality the actual increase is still two years
 (d) Subjects who survive two years are more likely to survive four years
 (e) The calculation was in error. The researchers forgot to include the sample size
3. A test of $H_0: \mu = 1.5$ vs. $H_a: \mu > 1.5$ produces a P-value of 0.0097. At an $\alpha = 0.01$ level, which of the following is an appropriate conclusion?

(a) Fail to reject H_0
 (b) Reject H_0
 (c) Cannot make any decision at all because the value of the test statistic is not available.
 (d) Accept H_a
 (e) Accept H_0
4. If we reject the null hypothesis when, in fact, it is true, we have

(a) Committed a Type I error
 (b) A computation error
 (c) Committed a Type II error
 (d) A correct decision
5. What is the null and alternative hypotheses for the question, "Do more than half of all adults think TV is less moral than society?"

$H_0: P = 50\%$ $H_a: P > 50\%$
6. A poll of 2,525 adults found that 1,970 said they support health care reform. What is the sample proportion (\hat{p}) who support health care reform?

$1970/2525 = 0.778$
7. A test of $H_0: \mu = 60$ vs. $H_a: \mu \neq 60$ produces a sample mean of $\bar{x} = 58$ and a P-value of 0.04. At an $\alpha = 0.05$ level, which of the following is an appropriate conclusion?

Reject H_0
8. A test of significance produces a P-value of 0.024. Would you Reject the null or Fail to Reject the null given the following significance levels?

1% 5% 10%

Reject / Fail to Reject Reject / Fail to Reject Reject / Fail to Reject

9. A radio talk show host is interested in the proportion p of adults in his listening area who think that the drinking age should be lowered to 18. To find this proportion, he poses the following question to his listeners: "Do you think that the drinking age should be reduced to 18 in light of the fact that 18-year-olds are eligible for military service?" He asks listeners to phone in and vote "Yes" or "no" depending upon their opinions. Of 200 people who phone in, 140 answer "Yes." If the radio host was to run a hypothesis test to test his sample data against the reported national average, what is the sample proportion (\hat{p})?

$X = 140$ $n = 200$

$140/200 = 0.70$

Type 1
10. Since testing the lifetime of a battery requires draining the battery completely, the supervisor wants to sample as few batteries as possible from each hour's production. She is considering a sample size of $n = 4$.

a. Explain a Type 1 and Type 2 error.

b. Would you recommend that the quality control supervisor use a significance level of $\alpha = 0.01$ or $\alpha = 0.10$ in future tests? Explain.

11. When the manufacturing process is working properly, NeverReady batteries have lifetimes that follow a right-skewed distribution with $\mu = 7$ hours and $\sigma = 0.5$ hours. A quality control supervisor selects a simple random sample of n batteries every hour and measures the lifetime of each. If she is convinced that the mean lifetime of all batteries produced that hour is less than 7 hours at the 5% significance level, then all those batteries are discarded.

a. State the null and alternative hypotheses for the quality control supervisor to test.

$H_0: \mu = 7$ hours
 $H_a: \mu < 7$ hours

(b) Describe a Type I and a Type II error in this situation, and explain which is more serious.
12. Fill in the following decision rules then use the p-value to arrive a conclusion.

HYPOTHESES	SIGNIFICANCE LEVEL	P-VALUE	CONCLUSION (CIRCLE ONE)
$H_0: p = 0.25$	$\alpha = 0.05$	P-value = 0.0418	Reject the null hypothesis Fail to reject the null hypothesis
$H_0: p < 0.25$	$\alpha = 0.05$	P-value = 0.002	Reject the null hypothesis Fail to reject the null hypothesis
$H_0: p \neq 0.25$	$\alpha = 0.10$	P-value = 0.0325	Reject the null hypothesis Fail to reject the null hypothesis
$H_0: p = 0.25$	$\alpha = 0.01$	P-value = 0.0213	Reject the null hypothesis Fail to reject the null hypothesis

13. A manufacturer of cold medicine claims that 60% of all adults suffer at least one cold during every winter. In a simple random sample of 200 adults, 130 adults reported that they had at least one cold last winter? Is there enough statistical evidence that the manufacturer is making false claims about colds to gain a profit?

- State: $H_0: p = 60\%$, $H_a: p \neq 60\%$
- Plan: 1 prop
- Do: $p = 0.15$ ($\alpha = 0.05$)
- Conclude: Fail to reject

$n = 200$
 $X = 130$

14. An SRS of 25 male faculty members of a large university found that 15 felt that the university was supportive of female and minority faculty. An independent SRS of 20 female faculty found that 5 felt that the university was supportive of female and minority faculty. Is there evidence, at the 10% significance level, that the proportion of male faculty members who felt the university was supportive of female and minority faculty is larger than the corresponding proportion for female faculty members?

- State: $H_0: p_m = p_{fem}$, $H_a: p_m > p_{fem}$
- Plan: 2 sample prop
- Do: $p = 0.009$
- Conclude: Reject H_0

male 15 supp 11
female 5 25
 $\alpha = 10$

15. A drug manufacturer forms tablets by compressing a granular material that contains the active ingredient and various fillers. The hardness of a sample from each batch of tablets produced is measured to control the compression process. The target value for the hardness is $\mu = 11.5$. The hardness data for a random sample of 14 tablets are

11.627	11.613	11.493	11.701	11.360	11.374	11.592
11.458	11.552	11.463	11.383	11.715	11.485	11.509

Is there significant evidence at the 5% level that the mean hardness of the tablets differs from the target value? Carry out an appropriate test to support your answer.

- State: $H_0: \mu = 11.5$, $H_a: \mu \neq 11.5$
- Plan: 1 sample T
- Do: $p = 0.46$
- Conclude: Fail to reject

16. A random sample of students from a high school were chosen to determine if their sitting pulse rate was lower than their standing pulse rate. Each student's pulse rate was measured in both positions. Can we conclude the sitting pulse rate is lower?

4	62	74	82	88	82	66	64	84	72	82	80	72	64	62
42	68	78	80	92	58	96	72	100	82	76	92	74	60	58

- State: $H_0: \mu_{sitting} = \mu_{standing}$, $H_a: \mu_{sit} < \mu_{stand}$
- Plan: 2 sample T
- Do: $p = 0.21$
- Conclude: Fail to reject

17. Do students tend to improve their SAT Mathematics (SAT-M) score the second time they take the test? Four randomly sampled students who took the test twice received the

following scores:

Student	1	2	3	4
First Score	450	520	720	600
Second Score	440	600	720	630

Given a significance level of 5%, is there enough evidence to support that students score better on their second SAT-M attempt?

- State: $H_0: \mu_{1st} = \mu_{2nd}$
 $H_a: \mu_{1st} < \mu_{2nd}$
- Plan: 2 sample T
- Do: $p = 0.39$
- Conclude: Fail to Reject