RLP-STATSB · RLP Statistics B — Probability Models and Statistical Inference

Unit 3: Testing Claims: Significance Tests

Practice Workbook. The same problems, in the same order, as the Practice sections of the online lessons.

Name _____ Period _____

Lesson 3.1 — The Logic of a Significance Test

FLUENCY

1. Test H_0: $p = 0.4$ against H_a: $p > 0.4$, given 52 successes in a random sample of 150. Find the test statistic and the P-value.

2. Test H_0: $p = 0.25$ against H_a: $p > 0.25$, given 112 successes in a random sample of 500. Find the test statistic and the P-value.

3. Test H_0: $p = 0.2$ against H_a: $p < 0.2$, given 69 successes in a random sample of 500. Find the test statistic and the P-value.

4. Test H_0: $p = 0.3$ against H_a: $p \neq 0.3$, given 66 successes in a random sample of 300. Find the test statistic and the P-value.

5. Test H_0: $p = 0.6$ against H_a: $p > 0.6$, given 203 successes in a random sample of 300. Find the test statistic and the P-value.

6. Test H_0: $p = 0.3$ against H_a: $p \neq 0.3$, given 55 successes in a random sample of 150. Find the test statistic and the P-value.

7. Which is a correctly stated null hypothesis? (a) H_0: $\hat{p} = 0.6$ (b) H_0: $p = 0.6$ (c) H_0: $p > 0.6$

8. A test gives $P = 0.03$ at $\alpha = 0.05$. State the decision in words.

9. A test gives $P = 0.21$ at $\alpha = 0.05$. State the decision in words.

10. In a one-proportion z-test, why does the standard deviation use p_0 rather than $\hat{p}$?

APPLICATION

11. A random sample of 200 St. Louis days (2000-2024) contains 66 rainy days. A friend claims one day in four is rainy. Test H_0: $p = 0.25$ against H_a: $p \neq 0.25$ at $\alpha = 0.05$, with conditions, and state the conclusion in context.

12. A random sample of 150 Springfield days (2000-2024) contains 55 rainy days. Test whether fewer than 40% of Springfield days are rainy (H_a: $p < 0.40$), at $\alpha = 0.05$.

13. A student writes: 'P = 0.046, so there is a 4.6% chance that H_0 is true.' Rewrite the sentence correctly.

CHALLENGE

14. For the St. Louis test above, the normal approximation gives $P \approx 0.009$ and the exact binomial gives $P \approx 0.014$. When can two methods lead to different decisions at $\alpha = 0.05$, and how should a report handle that?

REVIEW

15. In a random sample of 800, 162 said yes. Construct a 95% confidence interval for the population proportion.

16. (Unit 2) A 95% interval for p is (0.14, 0.25). Would a two-sided test of H_0: $p = 0.30$ at $\alpha = 0.05$ reject?

Lesson 3.2 — t-Tests for Means and Paired Data

FLUENCY

1. A sample of 16 has $\bar{x} = 52.4$ and $s = 6$. Find t for H_0: $\mu = 50$.

2. Find the two-sided P-value for $t = 1.6$ with $df = 15$.

3. Find the two-sided P-value for $t = 2.53$ with $df = 29$.

4. Paired differences: $n = 20$, $\bar{d} = 1.1$, $s_d = 2.0$. Find t and the two-sided P-value for $H_0: \mu_d = 0$.

5. A 95% t-interval for μ is (65.9, 73.1). Will a two-sided test of $H_0: \mu = 72$ at $\alpha = 0.05$ reject?

6. What should you plot before a paired t-test on yearly differences, and why?

APPLICATION

7. A random sample of 25 Springfield daily highs (2000-2024) has $\bar{x} = 67.60°F$ and $s = 25.18°F$. A travel site claims Springfield's average daily high is 72°F. Test $H_0: \mu = 72$ against $H_a: \mu \neq 72$.

8. St. Louis minus Kansas City precipitation, paired by year, 1995-2024 (NOAA): $\bar{d} = 3.414$ in, $s_d = 9.065$ in, $n = 30$. Test whether St. Louis averages more rain than Kansas City (H_a: $\mu_d > 0$).

9. Joplin minus St. Louis annual mean temperature, paired by year, 1995-2024 (30 years with both values): $\bar{d} = 0.880°F$, $s_d = 0.780°F$. Carry out a two-sided paired t-test and give the 95% interval.

CHALLENGE

10. Explain why a two-sided test at $\alpha = 0.05$ rejects H_0: $\mu = \mu_0$ exactly when μ_0 lies outside the 95% t-interval.

REVIEW

11. (Unit 2) Find the 95% t-interval for $n = 16$, $\bar{x} = 52.4$, $s = 6$.

12. (Unit 1) Find $P(Z > 2.53)$ for a standard normal Z.

Lesson 3.3 — Errors, Power and Significance in the Real World

FLUENCY

1. A test of H_0: 'the new filter does not reduce smoke particles' rejects H_0, but in fact the filter does nothing. Which error occurred?

2. A test fails to reject H_0: 'this batch of lettuce is not contaminated', but the batch is contaminated. Which error, and why does it matter here?

3. What is the probability of a Type I error when H_0 is true and $\alpha = 0.05$?

4. Name three ways to increase the power of a test.

5. Ten independent tests at $\alpha = 0.05$, all nulls true: find P(at least one false positive).

6. True or false: a very small P-value means the effect is large.

APPLICATION

7. A two-sided one-proportion test uses H_0: $p = 0.5$ and $\alpha = 0.05$. If the true p is 0.6, find the power for n = 100, 200 and 400.

8. (Illustrative) A survey of 100,000 people in each of two regions finds 50.6% and 50.2% agreeing with a statement. The two-proportion z is about 2.53 and $P \approx 0.011$. Write a sentence that reports both statistical and practical significance.

9. A study tests 20 foods for a link to headaches and reports only that chocolate was 'significantly linked' ($P = 0.03$). Critique the report.

CHALLENGE

10. A small study (n = 12 per group) finds P = 0.30 and concludes 'the tutoring program has no effect.' Its 95% interval for the effect runs from −4 to +11 points. What should the conclusion say instead?

REVIEW

11. Test H_0: $p = 0.4$ against H_a: $p > 0.4$, given 74 successes in a random sample of 200. Find the test statistic and the P-value.

12. (Unit 2) Why is an interval for the effect more informative than a P-value alone?

Lesson 3.4 — Two-Sample Tests from Experiments and Observational Data

FLUENCY

1. Two independent samples: $\bar{x}_1 = 13.2$, $s_1 = 1.93$, $n_1 = 10$; $\bar{x}_2 = 11.0$, $s_2 = 1.83$, $n_2 = 10$. Find the two-sample t statistic.

2. Two proportions: 45 of 100 and 30 of 100. Find the pooled proportion and the z statistic for H_0: $p_1 = p_2$.

3. Random assignment, no random selection: what can the conclusion claim?

4. Random selection, no random assignment: what can the conclusion claim?

5. A two-sample t-test gives $t = 2.62$ with $df = 17.9$. Find the two-sided P-value.

6. Why use the pooled proportion in a two-proportion test but not in the interval?

APPLICATION

7. (Illustrative data, RLP Statistics A) The helicopter experiment: long blades mean 2.5063 s, short 2.3812 s, n = 8 each. A two-sample t-test gives $t \approx 2.78$, $df \approx 13.7$, two-sided $P \approx 0.0148$. The exact randomization test gave a one-sided P of about 0.0099. Compare and conclude.

8. Springfield winter days 2000-2024: rain followed 234 of 584 rainy days and 353 of 1,672 dry days. Carry out a two-proportion z-test and state the scope.

9. Joplin annual mean temperature, 1995-2024 against 1950-1979 (NOAA): means 58.77 and 57.39°F. A Welch test gives $t \approx 3.87$, $df \approx 58.0$, $P \approx 0.0003$. State the conclusion and its scope.

CHALLENGE

10. Show that for a 2×2 table the square of the two-proportion z statistic equals the chi-square statistic, using the St. Louis winter values $z \approx 8.29$ and $\chi^2 \approx 68.8$.

REVIEW

11. (Unit 2) Paired or independent: before-and-after blood pressure on the same 30 patients?

12. (RLP Statistics A) What design feature supports a causal conclusion?

Unit 3 checkpoint

1. Which hypothesis pair is correctly written?
   A. H_0: $\hat{p} = 0.5$; H_a: $\hat{p} \neq 0.5$
   B. H_0: $p = 0.5$; H_a: $p \neq 0.5$
   C. H_0: $p \neq 0.5$; H_a: $p = 0.5$
   D. H_0: $\bar{x} = 3$; H_a: $\bar{x} > 3$
2. With H_0: $p = 0.60$, $n = 150$ and $\hat{p} = 0.52$, the test statistic is $z =$ _____.
3. True or false: A P-value is the probability that the null hypothesis is true.
4. $P = 0.59$. The correct conclusion is…
   A. Accept H_0
   B. H_0 is proven
   C. Not convincing evidence against H_0
   D. Reject H_0
5. For paired data with $\bar{d} = 5.339$, $s_d = 7.204$ and $n = 30$, $t \approx$ _____.
6. True or false: A two-sided test at $\alpha = 0.05$ rejects H_0: $\mu = 65$ if 65 lies outside the 95% t-interval.
7. Rejecting a true H_0 is a…
   A. Type I error
   B. Type II error
   C. Power
   D. P-value
8. Which increases power?
   A. Smaller sample
   B. Smaller α
   C. Larger sample
   D. Smaller true effect
9. Twenty independent tests at $\alpha = 0.05$ with all nulls true: P(at least one false positive) ≈ _____.
10. True or false: With 100,000 people per group, a 0.5-point difference can be statistically significant but practically unimportant.
11. A randomized experiment gives $P = 0.018$. What can be concluded?
    A. Association only
    B. A causal effect for the participants
    C. A causal effect for everyone
    D. Nothing

12. True or false: The pooled proportion is used in the standard error of a two-proportion confidence interval.