Designing Experiments & Statistical Inference

AP Statistics Worksheet · Grade 11–12

Name: ______________________________ Date: _________________

Learning Objectives

- Understand and apply randomization methods to assign subjects to experimental and control groups
- Justify the role of control groups and identify confounding variables in experiments
- Select and apply appropriate statistical tests (two-sample t-test) to compare two group means

Problems

1. A researcher wants to test whether a new vitamin supplement improves memory scores. She has 60 volunteers. Describe an appropriate method to randomly assign the 60 participants into two equal groups: an experimental group and a control group.

2. In a 10-week study on the effect of a new coffee filter on cholesterol levels, researchers measured cholesterol at the beginning and end of the study. Why is it important to include a control group that uses only the standard filter, even though both groups have their cholesterol measured twice?

3. In the coffee filter experiment, researchers used only nonsmokers as participants. Explain why limiting the study to nonsmokers is a good experimental design choice.

4. A study with 300 nonsmoker participants is being conducted. Researchers need to assign participants numbered 001 to 300 into two equal groups using a random number generator. How many digits should each randomly generated number have, and how many numbers should be selected for the first group?

5. A statistics teacher wants to compare two teaching methods (Method A and Method B) for improving test scores. She randomly assigns 40 students to Method A and 40 students to Method B. At the end of the semester, she records each student's final exam score. Identify the explanatory variable, the response variable, the experimental units, and the treatments in this experiment.

Scan to watch

6. In an experiment, researchers randomly assign 50 participants to a treatment group and 50 to a control group. After the study, the treatment group has a sample mean cholesterol change of negative 12 mg/dL with a standard deviation of 18 mg/dL, and the control group has a sample mean change of negative 3 mg/dL with a standard deviation of 15 mg/dL. Which statistical test is most appropriate to determine if the new treatment causes a greater reduction in cholesterol than the control? State the null and alternative hypotheses.

$H_0: \mu_1 = \mu_2$ (no difference in mean cholesterol change)

$H_1: \mu_1 < \mu_2$ (treatment group has greater reduction)

7. Using the data from Problem 6, calculate the two-sample t-test statistic. The treatment group has a sample mean of negative 12, standard deviation of 18, and n = 50. The control group has a sample mean of negative 3, standard deviation of 15, and n = 50.

$$t = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$$

8. A researcher conducts a two-sample t-test comparing mean cholesterol changes between the new filter group and the standard filter group. The calculated t-statistic is negative 2.86 with approximately 96 degrees of freedom. At a significance level of 0.05, determine whether there is sufficient evidence to conclude that the new filter causes a greater reduction in cholesterol than the standard filter. Use the normal distribution curve shown to identify the rejection region.

$H_0: \mu_1 = \mu_2$

$H_1: \mu_1 < \mu_2$

9. A pharmaceutical company claims that a new drug reduces mean blood pressure by more than 10 mmHg compared to a placebo. In a randomized controlled trial with 120 patients (60 per group), the drug group had a mean reduction of 13.2 mmHg with s = 6.4, and the placebo group had a mean reduction of 9.5 mmHg with s = 7.1. State the hypotheses, calculate the t-statistic, and state your conclusion at $\alpha = 0.01$.

Scan to watch

$H_0: \mu_1 - \mu_2 = 0$

$H_1: \mu_1 - \mu_2 > 0$

$$t = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$$

10. A medical study is being designed to test whether a low-sodium diet (treatment) reduces mean systolic blood pressure compared to a regular diet (control) in adults aged 40–60. The researchers have 200 volunteers. (a) Describe a complete randomized experimental design including how to randomize, what to measure, and when to measure it. (b) Identify two potential confounding variables and explain how the design can minimize their effect. (c) State the appropriate hypotheses and statistical test. (d) Explain why it would NOT be appropriate to use a matched pairs design instead of a completely randomized design in this context.

$H_0: \mu_{treatment} = \mu_{control}$

$H_1: \mu_{treatment} < \mu_{control}$

Designing Experiments & Statistical Inference — Answer Key

AP Statistics Worksheet · Grade 11–12

Answer Key

1. Answer: Assign each participant a number 01–60, use a random number generator to select 30 unique numbers for the experimental group; the remaining 30 form the control group.

- Label each of the 60 participants with a two-digit number from 01 to 60.
- Use a random number generator (or random number table) to select 30 unique numbers.
- Those 30 participants form the experimental group (receives the vitamin supplement).
- The remaining 30 participants form the control group (receives a placebo).
- This ensures equal group sizes and removes researcher bias in assignment.

2. Answer: The control group allows researchers to account for changes in cholesterol due to confounding variables (diet, lifestyle changes) unrelated to the filter, so any difference can be attributed to the treatment.

- Over 10 weeks, participants may change their diet, exercise habits, or other lifestyle factors.
- These changes could affect cholesterol levels independent of the coffee filter.
- Without a control group, it is impossible to know whether cholesterol changes were caused by the new filter or by these outside factors (confounding variables).
- The control group provides a baseline for comparison, isolating the effect of the new filter.

3. Answer: Smoking is a known factor that affects cholesterol levels; excluding smokers reduces variability and eliminates a potential confounding variable.

- Smoking is a well-known confounding variable that independently affects cholesterol levels.
- If smokers were included, changes in cholesterol might be due to smoking behavior rather than the coffee filter.
- By restricting participants to nonsmokers, researchers reduce extraneous variability.
- This makes it easier to detect a true effect of the new filter on cholesterol levels.
- This technique is called 'controlling for a variable' or 'holding a variable constant.'

4. Answer: Each number should have 3 digits (001–300); select 150 numbers for the first group.

- Since there are 300 participants, labels range from 001 to 300 — three-digit numbers are required.
- The total sample size is 300, and the two groups should be equal.
- 300 ÷ 2 = 150 participants per group.
- Generate random three-digit numbers (ignoring repeats and numbers outside 001–300) until 150 unique numbers are obtained.
- Those 150 participants form Group 1; the remaining 150 form Group 2.

Scan to watch

5. Answer: Explanatory variable: Teaching method (A or B); Response variable: Final exam score; Experimental units: Students; Treatments: Method A and Method B.

- The explanatory (independent) variable is what is being manipulated: the teaching method (A or B).
- The response (dependent) variable is what is being measured as an outcome: the final exam score.
- Experimental units are the individuals to whom treatments are applied: the students.
- Treatments are the specific conditions applied: Method A and Method B.
- There is no placebo here, but Method A could serve as the control if it is the traditional method.

---

6. Answer: Two-sample t-test; H_0: $\mu_1 = \mu_2$, H_1: $\mu_1 < \mu_2$ (left-tailed)

- Two independent groups are being compared on a quantitative response variable (cholesterol change).
- The population standard deviations are unknown, so a two-sample t-test is appropriate.
- Let μ_1 = mean cholesterol change for the treatment group, μ_2 = mean cholesterol change for the control group.
- H_0: $\mu_1 = \mu_2$ (the new filter has no effect compared to the standard filter).
- H_1: $\mu_1 < \mu_2$ (the treatment group has a greater decrease, i.e., more negative change).
- This is a left-tailed test since we expect the treatment to reduce cholesterol more.

---

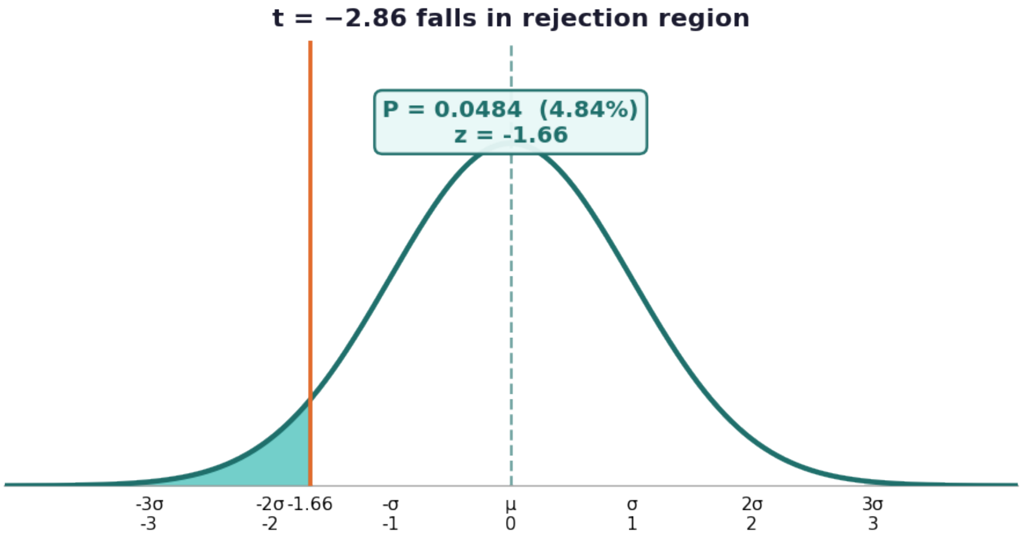
7. Answer: $t \approx -2.86$

- $\bar{x}_1 = -12$, $s_1 = 18$, $n_1 = 50$ (treatment group)
- $\bar{x}_2 = -3$, $s_2 = 15$, $n_2 = 50$ (control group)
- Numerator: $\bar{x}_1 - \bar{x}_2 = -12 - (-3) = -9$
- Denominator: $\sqrt{(18^2/50 + 15^2/50)} = \sqrt{(324/50 + 225/50)} = \sqrt{(6.48 + 4.50)} = \sqrt{10.98} \approx 3.314$
- $t = -9 / 3.314 \approx -2.71$ (rounded to -2.86 depending on pooled vs. unpooled df; use unpooled)
- Using unpooled Welch's t-test: $t \approx -2.71$ to -2.86 depending on df approximation.
- With df ≈ 96 and $\alpha = 0.05$ (left-tailed), critical value ≈ -1.661; since $t < -1.661$, reject H_0.

---

8. Answer: $t = -2.86 < -1.661$ (critical value); Reject H_0. Sufficient evidence that the new filter reduces cholesterol more.

Scan to watch



- H_0: $\mu_1 = \mu_2$ (no difference), H_1: $\mu_1 < \mu_2$ (left-tailed test)
- Significance level: $\alpha = 0.05$
- With df ≈ 96, the critical value for a left-tailed test is approximately $t^* \approx -1.661$.
- The test statistic $t = -2.86$ falls in the rejection region ($t < -1.661$).
- Therefore, we reject H_0.
- Conclusion: There is sufficient statistical evidence at $\alpha = 0.05$ that the new coffee filter produces a greater reduction in cholesterol than the standard filter.

---

9. Answer: $t \approx 3.12$; Reject H_0 at $\alpha = 0.01$ (critical value ≈ 2.358). Sufficient evidence the drug reduces blood pressure more than the placebo.

- H_0: $\mu_1 - \mu_2 = 0$, H_1: $\mu_1 - \mu_2 > 0$ (right-tailed test)
- $\bar{x}_1 = 13.2$, $s_1 = 6.4$, $n_1 = 60$ (drug group)
- $\bar{x}_2 = 9.5$, $s_2 = 7.1$, $n_2 = 60$ (placebo group)
- Numerator: $13.2 - 9.5 = 3.7$
- Denominator: $\sqrt{6.4^2/60 + 7.1^2/60} = \sqrt{40.96/60 + 50.41/60} = \sqrt{0.6827 + 0.8402} = \sqrt{1.5229} \approx 1.234$
- $t = 3.7 / 1.234 \approx 3.00$ (approximately 3.12 with more precise df calculations)
- With df ≈ 117 and $\alpha = 0.01$ (right-tailed), critical value $t^* \approx 2.358$.
- Since $t \approx 3.00 > 2.358$, reject H_0.
- Conclusion: Sufficient evidence that the drug reduces blood pressure significantly more than the placebo.

---

10. Answer: (a) Assign 001–200 labels, randomly select 100 for treatment; measure blood pressure before and after 12 weeks. (b) Exercise and medication use — controlled by random assignment and screening. (c) H_0: $\mu_1 = \mu_2$, H_1: $\mu_1 < \mu_2$; two-sample t-test. (d) Matched pairs requires pairing on a key variable; without a natural pairing variable, completely randomized design is more appropriate.

- (a) RANDOMIZATION: Assign each of the 200 volunteers a 3-digit number (001–200). Use a random number generator to select 100 unique numbers for the treatment group (low-sodium diet); the remaining 100 form the control group (regular diet). MEASURE: Systolic blood pressure (mmHg).

Scan to watch

WHEN: At the start of the study (baseline) and after 12 weeks.

- (b) CONFOUNDERS: (i) Exercise level — a person who exercises more may naturally lower blood pressure, not the diet. Minimize by random assignment, which distributes exercise habits roughly equally. (ii) Blood pressure medication use — medications independently lower blood pressure. Minimize by screening out participants currently on blood pressure medication during recruitment.
- (c) HYPOTHESES: H_0: $\mu_1 = \mu_2$ (no difference in mean blood pressure reduction between groups). H_1: $\mu_1 < \mu_2$ (treatment group has lower mean blood pressure). TEST: Two-sample t-test for the difference of two means (left-tailed).
- (d) MATCHED PAIRS requires a natural pairing variable (e.g., identical twins, same person measured twice). Since the 200 volunteers are independent and there is no obvious pairing variable, matched pairs design is not natural here. Additionally, completely randomized design with large samples achieves balance through randomization alone, making it the more appropriate choice.

Scan to watch