Chi-Square Goodness-of-Fit Test

Statistics Worksheet · Grade 11–12 / AP Statistics

Name: ________________________________ Date: ___________________

Learning Objectives

- State the null and alternative hypotheses for a chi-square goodness-of-fit test
- Calculate expected counts and verify the conditions for the chi-square goodness-of-fit test
- Compute the chi-square test statistic and interpret the results in context

Problems

1. A spinner has 4 equal sections labeled A, B, C, and D. If the spinner is fair, what is the expected proportion for each section? A spinner is spun 200 times. Fill in the expected count for each section in the table below.

Section	Expected Proportion	Expected Count (n = 200)
A	0.25	
B	0.25	
C	0.25	
D	0.25	

2. For a chi-square goodness-of-fit test, state the null and alternative hypotheses for the following scenario: A die manufacturer claims their six-sided die is fair, meaning each face should appear with equal probability of 1/6.

$$H_0 : p_1 = p_2 = p_3 = p_4 = p_5 = p_6 = \frac{1}{6}$$

3. Using the car dealership problem from the video, the manager surveyed 150 customers. The expected proportions based on last year are given. Calculate the expected count for each car color.

Color	Last Year %	Expected Count (n = 150)
Red	30%	
Yellow	20%	

Scan to watch

Color	Last Year %	Expected Count (n = 150)
Green	10%	
Blue	10%	
White	30%	

4. One condition for a chi-square goodness-of-fit test is that all expected counts must be greater than 5. A researcher surveys 40 people about their favorite season and expects 25% for each of the four seasons. Do the conditions for the chi-square goodness-of-fit test hold? Explain.

$$E_i = 40 \times 0.25 = 10$$

5. Using the car dealership data from the video, compute the chi-square contribution for the Red car color only. The observed count for Red is 50 and the expected count is 45.

$$\frac{(O - E)^2}{E} = \frac{(50 - 45)^2}{45}$$

6. Using the car dealership survey data from the video, compute the chi-square contribution for each color and find the total chi-square test statistic. Observed counts: Red = 50, Yellow = 35, Green = 30, Blue = 10, White = 25. Expected counts: Red = 45, Yellow = 30, Green = 15, Blue = 15, White = 45.

Color	O	E	$(O - E)^2$	$(O - E)^2 / E$
Red	50	45		
Yellow	35	30		
Green	30	15		
Blue	10	15		
White	25	45		
Total	150	150	—	

7. Determine the degrees of freedom for the car dealership chi-square goodness-of-fit test, which has 5 color categories. Then use the chi-square critical value table to find the critical value at a significance level of 0.05.

Scan to watch

$$df = k - 1$$

8. A market researcher claims that customers choose among four phone brands in the proportions: Brand A = 40%, Brand B = 30%, Brand C = 20%, Brand D = 10%. A sample of 300 customers gives: Brand A = 110, Brand B = 100, Brand C = 65, Brand D = 25. Compute the expected counts, verify the conditions, and calculate the chi-square test statistic.

Brand	O	Expected %	E	(O − E)² / E
A	110	40%		
B	100	30%		
C	65	20%		
D	25	10%		
Total	300	100%		

9. Using the phone brand problem from Problem 8, the chi-square test statistic is approximately 3.194 with 3 degrees of freedom. At a significance level of 0.05, the critical value is 7.815. State the decision and write a conclusion in context.

$$\chi^2 \approx 3.194 < \chi^2_{critical} = 7.815$$

10. A genetics researcher expects offspring to follow Mendel's 9:3:3:1 ratio across four phenotypes from a total of 320 offspring. The observed counts are: Phenotype 1 = 180, Phenotype 2 = 55, Phenotype 3 = 58, Phenotype 4 = 27. Carry out a full chi-square goodness-of-fit test at the 0.05 significance level: state hypotheses, compute expected counts, verify conditions, calculate the test statistic, find degrees of freedom and the critical value, then state your conclusion.

Phenotype	Ratio	O	E	(O − E)² / E
1	9/16	180		
2	3/16	55		
3	3/16	58		
4	1/16	27		
Total	16/16	320		

Scan to watch

Chi-Square Goodness-of-Fit Test — Answer Key

Statistics Worksheet · Grade 11–12 / AP Statistics

Answer Key

1. Answer: 50 for each section

Section	Expected Proportion	Expected Count (n = 200)
A	0.25	50
B	0.25	50
C	0.25	50
D	0.25	50

- Expected Count = $n \times p_i = 200 \times 0.25 = 50$
- Each section has an expected count of 50

2. Answer: H0: All six faces appear with equal probability 1/6. Ha: At least one face appears with a probability different from 1/6.

- The null hypothesis states that the distribution matches the claimed distribution: each face has probability 1/6.
- The alternative hypothesis states that at least one proportion differs from 1/6, meaning the die is not fair.

3. Answer: Red: 45, Yellow: 30, Green: 15, Blue: 15, White: 45

Color	Last Year %	Expected Count (n = 150)
Red	30%	45
Yellow	20%	30
Green	10%	15
Blue	10%	15
White	30%	45

- Red: $150 \times 0.30 = 45$
- Yellow: $150 \times 0.20 = 30$
- Green: $150 \times 0.10 = 15$
- Blue: $150 \times 0.10 = 15$
- White: $150 \times 0.30 = 45$

4. Answer: Yes. Each expected count = 10, which is greater than 5, so the condition is satisfied.

- Expected count for each season = $40 \times 0.25 = 10$
- Since all expected counts (10) are greater than 5, the condition is met and we can proceed with the chi-square test.

Scan to watch

5. Answer: Approximately 0.556

- O = 50, E = 45
- $(O - E)^2 = (50 - 45)^2 = 5^2 = 25$
- $25 \div 45 \approx 0.556$

6. Answer: Chi-square test statistic ≈ 23.19

Color	O	E	$(O - E)^2$	$(O - E)^2 / E$
Red	50	45	25	0.556
Yellow	35	30	25	0.833
Green	30	15	225	15.000
Blue	10	15	25	1.667
White	25	45	400	8.889
Total	150	150	—	26.945

- Red: $(50–45)^2/45 = 25/45 \approx 0.556$
- Yellow: $(35–30)^2/30 = 25/30 \approx 0.833$
- Green: $(30–15)^2/15 = 225/15 = 15.000$
- Blue: $(10–15)^2/15 = 25/15 \approx 1.667$
- White: $(25–45)^2/45 = 400/45 \approx 8.889$
- $\chi^2 = 0.556 + 0.833 + 15.000 + 1.667 + 8.889 \approx 26.945$

7. Answer: df = 4; critical value = 9.488

- Degrees of freedom = $k - 1 = 5 - 1 = 4$
- Using the chi-square table with df = 4 and $\alpha = 0.05$, the critical value is 9.488.

8. Answer: E: A=120, B=90, C=60, D=30. All > 5, conditions met. $\chi^2 \approx 4.097$

Brand	O	Expected %	E	$(O - E)^2 / E$
A	110	40%	120	0.833
B	100	30%	90	1.111
C	65	20%	60	0.417
D	25	10%	30	0.833
Total	300	100%	300	3.194

- Expected counts: A = 300×0.40 = 120; B = 300×0.30 = 90; C = 300×0.20 = 60; D = 300×0.10 = 30
- All expected counts > 5: condition satisfied.
- A: $(110–120)^2/120 = 100/120 \approx 0.833$
- B: $(100–90)^2/90 = 100/90 \approx 1.111$
- C: $(65–60)^2/60 = 25/60 \approx 0.417$
- D: $(25–30)^2/30 = 25/30 \approx 0.833$

Scan to watch

• $\chi^2 \approx 0.833 + 1.111 + 0.417 + 0.833 = 3.194$

9. Answer: Fail to reject H0. There is not enough evidence to conclude that the distribution of customer brand preferences differs from the claimed proportions.

• Compare test statistic to critical value: $3.194 < 7.815$
• Since the test statistic does not exceed the critical value, we fail to reject H0.
• Conclusion: At the 0.05 significance level, there is insufficient evidence that the actual brand preferences differ from the researcher's claimed distribution.

10. Answer: $\chi^2 \approx 1.817$; df = 3; critical value = 7.815; Fail to reject H0. Data fits the 9:3:3:1 ratio.

Phenotype	Ratio	O	E	$(O - E)^2 / E$
1	9/16	180	180	0.000
2	3/16	55	60	0.417
3	3/16	58	60	0.067
4	1/16	27	20	2.450
Total	16/16	320	320	2.934

• H0: Offspring follow the 9:3:3:1 Mendelian ratio. Ha: At least one proportion differs from the expected ratio.
• Expected counts: P1 = 320×(9/16) = 180; P2 = 320×(3/16) = 60; P3 = 320×(3/16) = 60; P4 = 320×(1/16) = 20
• All expected counts > 5: conditions are satisfied.
• P1: $(180–180)^2/180 = 0.000$
• P2: $(55–60)^2/60 = 25/60 \approx 0.417$
• P3: $(58–60)^2/60 = 4/60 \approx 0.067$
• P4: $(27–20)^2/20 = 49/20 = 2.450$
• $\chi^2 \approx 0 + 0.417 + 0.067 + 2.450 = 2.934$
• df = 4 – 1 = 3; critical value at $\alpha = 0.05$ is 7.815
• Since $2.934 < 7.815$, fail to reject H0. The data is consistent with Mendel's 9:3:3:1 ratio.

Scan to watch