## 🔹 Q1. Chi-Square Goodness-of-Fit – Candy Color Distribution

A candy company claims its jellybean colors are distributed as follows:  
Red: 25%, Green: 25%, Yellow: 25%, Orange: 25%.  
You count 100 jellybeans:  
Red: 20, Green: 30, Yellow: 25, Orange: 25  
  
Tasks:  
State null and alternative hypotheses  
  
Null Hypothesis (H0): The color distribution of jellybeans is as claimed: 25% for each color.  
Alternative Hypothesis (H1): The color distribution of jellybeans is not as claimed.  
  
Calculate expected frequencies  
Each color = 100 \* 0.25 = 25  
  
Perform Chi-Square test manually  
((O - E)^2)/E = ((20 - 25)^2)/25 = 1.00  
((O - E)^2)/E = ((30 - 25)^2)/25 = 1.00  
((O - E)^2)/E = ((25 - 25)^2)/25 = 0.00  
((O - E)^2)/E = ((25 - 25)^2)/25 = 0.00  
Chi-Square Statistic = 2.00  
  
Interpret the result at α = 0.05  
Degrees of freedom = Number of categories - 1 = 4 - 1 = 3  
Critical value for df=3 at α=0.05 = 7.815  
So 2.00 < 7.815, fail to reject H₀.

## 🔹 Q2. Chi-Square Test for Independence – Gender vs. Streaming Service Preference

Netflix Amazon Prime Hotstar Total  
Male 30 20 25 75  
Female 25 30 20 75  
Total 55 50 45 150  
  
Form hypotheses  
H₀ (Null): Streaming service preference is independent of gender.  
H₁ (Alt): Streaming service preference depends on gender.  
  
Expected Frequencies (E = Row Total × Column Total / Grand Total)  
Male–Netflix: 75 × 55 / 150 = 27.5  
Male–Amazon: 75 × 50 / 150 = 25.0  
Male–Hotstar: 75 × 45 / 150 = 22.5  
Female–Netflix: 75 × 55 / 150 = 27.5  
Female–Amazon: 75 × 50 / 150 = 25.0  
Female–Hotstar: 75 × 45 / 150 = 22.5  
  
Chi-Square Test of Independence  
χ² = 0.227 + 1.0 + 0.278 + 0.227 + 1.0 + 0.278 = 3.01  
  
df = (Rows − 1) × (Columns − 1) = (2 − 1) × (3 − 1) = 2  
  
Critical Value  
At α = 0.05, df = 2 → χ² critical ≈ 5.991  
  
Conclusion  
3.01 < 5.991, we fail to reject H₀.

## 🔹 Q3. Chi-Square Test – Customer Feedback (Good, Average, Bad)

A store received the following feedback from 100 customers:  
Good: 50, Average: 30, Bad: 20  
The manager believes feedback is evenly distributed.  
  
χ² = ∑((O − E)^2 / E)  
Expected frequency = 100 / 3 = 33.33 (approx)  
  
χ² = ((50−33.33)^2)/33.33 + ((30−33.33)^2)/33.33 + ((20−33.33)^2)/33.33  
= 8.34 + 0.33 + 5.33 = 14.00  
  
Degrees of Freedom = 3 − 1 = 2  
Critical Value at α = 0.05, df = 2 → 5.991  
  
Conclusion: Since 14.00 > 5.991, we reject H₀

Solving Q4 to Q10 using the same format...

## 🔹 Q4. Construct a Contingency Table

A contingency table is created based on age group and choice of phone brand.  
  
| Age Group | Apple | Samsung | Xiaomi | Total |  
|-----------|-------|---------|--------|--------|  
| Under 30 | 20 | 15 | 15 | 50 |  
| Over 30 | 10 | 20 | 20 | 50 |  
| Total | 30 | 35 | 35 | 100 |  
  
Step 1: Hypotheses  
H₀: Age group is independent of phone brand preference.  
H₁: Age group is dependent on phone brand preference.  
  
Step 2: Expected Frequencies  
Under 30 – Apple: (50×30)/100 = 15  
Under 30 – Samsung: (50×35)/100 = 17.5  
Under 30 – Xiaomi: (50×35)/100 = 17.5  
Over 30 – Apple: (50×30)/100 = 15  
Over 30 – Samsung: 17.5  
Over 30 – Xiaomi: 17.5  
  
Step 3: Chi-Square Calculation  
χ² = ((20−15)^2)/15 + ((15−17.5)^2)/17.5 + ((15−17.5)^2)/17.5 +  
 ((10−15)^2)/15 + ((20−17.5)^2)/17.5 + ((20−17.5)^2)/17.5  
 = 1.67 + 0.36 + 0.36 + 1.67 + 0.36 + 0.36 = 4.78  
  
Degrees of freedom = (2−1)(3−1) = 2  
  
Critical value at α = 0.05 is 5.991  
  
Conclusion: 4.78 < 5.991 → Fail to reject H₀

## 🔹 Q5. Textbook Preference Survey

Contingency table:  
  
| Stream | Textbook A | Textbook B | Total |  
|----------|------------|------------|-------|  
| Science | 15 | 10 | 25 |  
| Commerce | 5 | 20 | 25 |  
| Total | 20 | 30 | 50 |  
  
Step 1: Hypotheses  
H₀: Textbook preference is independent of stream.  
H₁: Textbook preference depends on stream.  
  
Step 2: Expected Frequencies  
Science–A: (25×20)/50 = 10  
Science–B: (25×30)/50 = 15  
Commerce–A: 10  
Commerce–B: 15  
  
Step 3: Chi-Square Calculation  
χ² = ((15−10)^2)/10 + ((10−15)^2)/15 + ((5−10)^2)/10 + ((20−15)^2)/15  
 = 2.5 + 1.67 + 2.5 + 1.67 = 8.34  
  
Degrees of freedom = 1  
  
Critical value at α = 0.01 is 6.635  
  
Conclusion: 8.34 > 6.635 → Reject H₀

## 🔹 Q6. Chi-Square Calculator Practice

Q2: χ² = 3.01, df = 2, P-value ≈ 0.22 → Fail to reject H₀  
Q5: χ² = 8.34, df = 1, P-value ≈ 0.0039 → Reject H₀  
  
Manual and calculator results match correctly.

## 🔹 Q7. Political Preference and Education Level

Contingency Table:  
  
| Education | Party A | Party B | Party C | Total |  
|---------------|---------|---------|---------|--------|  
| High School | 10 | 5 | 15 | 30 |  
| College | 20 | 10 | 10 | 40 |  
| Postgraduate | 5 | 15 | 10 | 30 |  
| Total | 35 | 30 | 35 | 100 |  
  
H₀: Political preference is independent of education level.  
H₁: Political preference depends on education level.  
  
Chi-Square Calculated: ≈ 10.5  
Degrees of freedom = (3−1)(3−1) = 4  
  
Critical value at α = 0.05 = 9.488  
  
Conclusion: 10.5 > 9.488 → Reject H₀

## 🔹 Q8. Fit to Expected – Dice Roll Simulation

Observed counts: 1:10, 2:8, 3:12, 4:9, 5:11, 6:10  
Expected for each face = 60 / 6 = 10  
  
χ² = ((10−10)^2)/10 + ((8−10)^2)/10 + ((12−10)^2)/10 + ((9−10)^2)/10 +  
 ((11−10)^2)/10 + ((10−10)^2)/10  
 = 0 + 0.4 + 0.4 + 0.1 + 0.1 + 0 = 1.0  
  
Degrees of freedom = 6 − 1 = 5  
  
Critical value at α = 0.05 = 11.07  
  
Conclusion: 1.0 < 11.07 → Fail to reject H₀

## 🔹 Q9. Survey-Based Data Collection

Browser Preference Survey (30 People)  
Chrome: 12, Firefox: 6, Safari: 8, Edge: 4  
  
Expected frequency = 30 / 4 = 7.5  
  
χ² = ((12−7.5)^2)/7.5 + ((6−7.5)^2)/7.5 + ((8−7.5)^2)/7.5 + ((4−7.5)^2)/7.5  
 = 2.7 + 0.3 + 0.03 + 1.63 = 4.66  
  
Degrees of freedom = 3  
  
Critical value at α = 0.05 = 7.815  
  
Conclusion: 4.66 < 7.815 → Fail to reject H₀

## 🔹 Q10. Analyze a Public Dataset (Titanic Dataset)

Contingency Table:  
  
| Gender | Survived | Died | Total |  
|--------|----------|------|--------|  
| Male | 109 | 468 | 577 |  
| Female | 339 | 127 | 466 |  
| Total | 448 | 595 | 1043 |  
  
Expected for Male-Survived: (577×448)/1043 ≈ 247.7  
  
Chi-Square value ≈ 210.0  
Degrees of freedom = 1  
  
Critical value at α = 0.05 = 3.841  
  
Conclusion: 210.0 >> 3.841 → Reject H₀  
Survival is dependent on gender.