

9-A Relationships Between Two Continuous Variables: Linear Correlation

1. Estimate the correlation coefficient for the following variable pairs, or write "not applicable."

- a) car size versus miles per gallon b) zip code versus annual income c) skin tone versus hair color

2. Fill in the blanks below to calculate r for the correlation between states' average gas prices (x), in dollars, and Obama approval rating (y), in percent, around the time of the 2012 election.

| State | x | y | $x - \bar{x}$ | $y - \bar{y}$ | $(x - \bar{x})(y - \bar{y})$ |
|----------------|-------|------|---------------|---------------|------------------------------|
| Georgia | 3.863 | 44.6 | _____ | _____ | _____ |
| Maryland | 3.997 | 55.5 | _____ | _____ | _____ |
| New Hampshire | 3.854 | 38.7 | _____ | _____ | _____ |
| South Carolina | 3.737 | 40.9 | _____ | _____ | _____ |
| Wyoming | 3.609 | 30.6 | _____ | _____ | _____ |

$\bar{x} =$ _____ $\bar{y} =$ _____ $\Sigma(x - \bar{x})(y - \bar{y}) =$ _____
 $s_x =$ _____ $s_y =$ _____ $s_{xy} =$ _____ \div _____ = _____
 $r =$ _____ \div _____ = _____

3. Use a calculator to do a two-tailed r test of a correlation on the data in #2.

- a) What is the alternate hypothesis?
- b) What is the value of r ? c) What is the p value? d) Are the data statistically significant?
- e) What can be concluded?
- f) Based on the concept that correlation does not imply causation, what cannot be concluded?

4. Find the equation of the line of best fit for the data in #2.

- a) What is the equation?
- b) Based on the data, what was the best prediction for Obama's approval rating in California, where gas averaged \$4.264 per gallon?
- c) Are you interpolating or extrapolating in (b)?
- d) What does your answer to (c) mean about your prediction in (b)?

9-B Relationships Between Two Discrete Variables: Independence

5. I want to see if there is a relationship between gender and assigned seat (front, middle, back) in my math classes.

a) Fill in the totals in the chart of observed data below.

| Seating | Front | Middle | Back | Total |
|---------|-------|--------|------|-------|
| Girls | 22 | 28 | 6 | |
| Boys | 38 | 25 | 12 | |
| Total | | | | |

b) How many degrees of freedom are there? c) What is my critical value of χ^2 ?

d) Sketch the χ^2 distribution, label the critical value, and shade the critical region.

e) Fill in the expected values in the chart below and verify that the totals are the same as in the observed chart.

| Seating | Front | Middle | Back | Total |
|---------|-------|--------|------|-------|
| Girls | | | | |
| Boys | | | | |
| Total | | | | |

f) Fill in the blanks below to calculate χ^2 .

| <u>gender</u> | <u>seat row</u> | <u>O</u> | <u>E</u> | <u>O - E</u> | <u>(O - E)²</u> | <u>(O - E)² ÷ E</u> |
|---------------|-----------------|----------|----------|--------------|----------------------------|--------------------------------|
| girls | front | _____ | _____ | _____ | _____ | _____ |
| boys | front | _____ | _____ | _____ | _____ | _____ |
| girls | middle | _____ | _____ | _____ | _____ | _____ |
| boys | middle | _____ | _____ | _____ | _____ | _____ |
| girls | back | _____ | _____ | _____ | _____ | _____ |
| boys | back | _____ | _____ | _____ | _____ | _____ |
| | | | | | $\chi^2 =$ | _____ |

g) Are the data statistically significant?

h) What can be concluded?

i) Enter the data into the calculator and calculate p .

9-C Distributions: Goodness of Fit

6. As of 2016, 45% of registered California voters are Democrats and 27% are Republicans. As a class, we will test to see if we can conclude that SVHS seniors do not follow this distribution.

a) Poll this class to find the number of people planning to register as Democrats: _____, Republicans: _____, and neither: _____.

b) Fill in the blanks below to calculate χ^2 .

| Political Party | Observed | % Expected | Expected | $(O - E)^2 \div E$ |
|-----------------|----------|------------|----------|--------------------|
| _____ | _____ | _____ | _____ | _____ |
| _____ | _____ | _____ | _____ | _____ |
| _____ | _____ | _____ | _____ | _____ |

$\chi^2 =$ _____

b) How many degrees of freedom are there? c) What is the critical value of χ^2 ? d) Are the data statistically significant?

e) State Joe's conclusion, followed by χ^2 and a range for p .

7. Last semester's PreCalculus final scores were 106, 111, 114, 114, 117, 123, 123, 123, 126, 126, 131, 132, 135, 135, 136, 137, 138, 141, 144, 144, 145, 147, 150, 150, 150, 153, 153, 153, 153, 155, 156, 156, 160, 162, 162, 165, 168, 171, 171, 174, 174, 174, 180, 180, 180, 183, 186, 186, 189, 195, 198, 198, 204, 213, 216, 216, and 222. The mean was 158 and the standard deviation was 29.

a) Fill in the chart below and calculate χ^2 . Values for z and for % expected for this problem are already calculated in Classwork 6 problem #8.

| x Range | z Range | Observed (O) | % Expected | Expected (E) | $(O - E)^2 \div E$ |
|--------------------|-----------|------------------|------------|------------------|--------------------|
| $x < 100$ | _____ | _____ | _____ | _____ | _____ |
| $100 \leq x < 120$ | _____ | _____ | _____ | _____ | _____ |
| $120 \leq x < 140$ | _____ | _____ | _____ | _____ | _____ |
| $140 \leq x < 160$ | _____ | _____ | _____ | _____ | _____ |
| $160 \leq x < 180$ | _____ | _____ | _____ | _____ | _____ |
| $180 \leq x < 200$ | _____ | _____ | _____ | _____ | _____ |
| $200 \leq x < 220$ | _____ | _____ | _____ | _____ | _____ |
| $220 \leq x$ | _____ | _____ | _____ | _____ | _____ |

$\chi^2 =$ _____

b) How many degrees of freedom are there? c) What is the critical value of χ^2 ? d) Are the data statistically significant?

e) State the conclusion, followed by χ^2 and a range for p .

9-D Standard Deviations: Single Variance

8. To see if Asian Americans have a standard deviation different from $\sigma = 100$ on the SAT, Majixkai gets a sample of 30 scores and finds $\bar{x} = 525$ and $s = 129$.

a) Calculate χ^2 .

b) How many degrees of freedom are there?

c) What are the critical values?

d) Are the data statistically significant?

e) State the conclusion, followed by χ^2 and a range for p .

9-E Differences Between Standard Deviations: Two Variances

9. Jordan is testing to see if one gender has more variation in academics than the other. She uses spring semester Statistics grades as her dependent variable and the 38 boys and 31 girls taking Statistics last year at SVHS as her sample.

a) How many degrees of freedom are there for each sample?

b) Explain why you do not have enough information yet to find the critical value.

c) For the boys, $\bar{x} = 78.0\%$ and $s = 7.96\%$. For the girls, $\bar{x} = 87.8\%$ and $s = 8.19\%$. Find the critical value.

d) Calculate F .

e) Are the data statistically significant?

f) State Jordan's conclusion, followed by F and a range for p .

9-F Differences Between Means: Analysis of Variance (ANOVA)

10. For a team competition, a member from each team will estimate 4.00 seconds several times.

a) What is the null hypothesis for an ANOVA of these data?

b) Fill in the blanks to calculate SS_W .

| | | | | | | | | |
|--|-----------------------------------|---------------------|-----------------------------------|---------------------|-----------------------------------|-----------|-----------------|---------------------|
| Team: _____ | Team: _____ | Team: _____ | | | | | | |
| \bar{x} | $x - \bar{x}_1$ | $(x - \bar{x}_1)^2$ | \bar{x} | $x - \bar{x}_2$ | $(x - \bar{x}_2)^2$ | \bar{x} | $x - \bar{x}_3$ | $(x - \bar{x}_3)^2$ |
| _____ | _____ | _____ | _____ | _____ | _____ | _____ | _____ | _____ |
| _____ | _____ | _____ | _____ | _____ | _____ | _____ | _____ | _____ |
| _____ | _____ | _____ | _____ | _____ | _____ | _____ | _____ | _____ |
| _____ | _____ | _____ | _____ | _____ | _____ | _____ | _____ | _____ |
| $\bar{x}_1 =$ _____ | $\Sigma(x - \bar{x}_1)^2 =$ _____ | $\bar{x}_2 =$ _____ | $\Sigma(x - \bar{x}_2)^2 =$ _____ | $\bar{x}_3 =$ _____ | $\Sigma(x - \bar{x}_3)^2 =$ _____ | | | |
| $SS_W =$ _____ + _____ + _____ = _____ | | | | | | | | |

c) How many degrees of freedom are in the denominator?

d) Fill in the blanks to calculate SS_B .

$\bar{X} = (\text{_____} (\text{_____}) + \text{_____} (\text{_____}) + \text{_____} (\text{_____})) \div (\text{_____} + \text{_____} + \text{_____}) = \text{_____}$

| Team | n | \bar{x} | $\bar{x} - \bar{X}$ | $(\bar{x} - \bar{X})^2$ | $n(\bar{x} - \bar{X})^2$ |
|-------|-------|-----------|---------------------|-------------------------|--------------------------|
| _____ | _____ | _____ | _____ | _____ | _____ |
| _____ | _____ | _____ | _____ | _____ | _____ |
| _____ | _____ | _____ | _____ | _____ | _____ |
| | | | | | $SS_B =$ _____ |

e) How many degrees of freedom are in the numerator?

f) What is the critical value?

g) Calculate MS_B .

h) Calculate MS_W .

i) Calculate F .

j) Are the data statistically significant?

11. Use a calculator to do an ANOVA on the data in #10.

a) What is the p value?

b) State the conclusion, followed by F and a range for p .

9-G Selecting a Statistical Test

12. State a statistical test that would be appropriate for the following research questions.

a) Are Whites more likely to apply for college than minorities?

b) Does the die I use to roll players roll each number with equal frequency?

c) Does the percentage of girls in a math class influence their level of test anxiety?

d) Is there a difference in average age of football fans, soccer fans, and hockey fans?

e) Is it harder to remember the middle 10 names in a list of 20 than the other 10 names?

f) Are people's time estimates more consistent when they are alone than when others are present?

g) Is there a relationship between high school attended (SVHS, SLVHS, SCHS, HHS, SHS) and college plans (none, 2-year, 4-year)?