

Math 214 Statistics Fall 13 Sample Midterm 2

In class

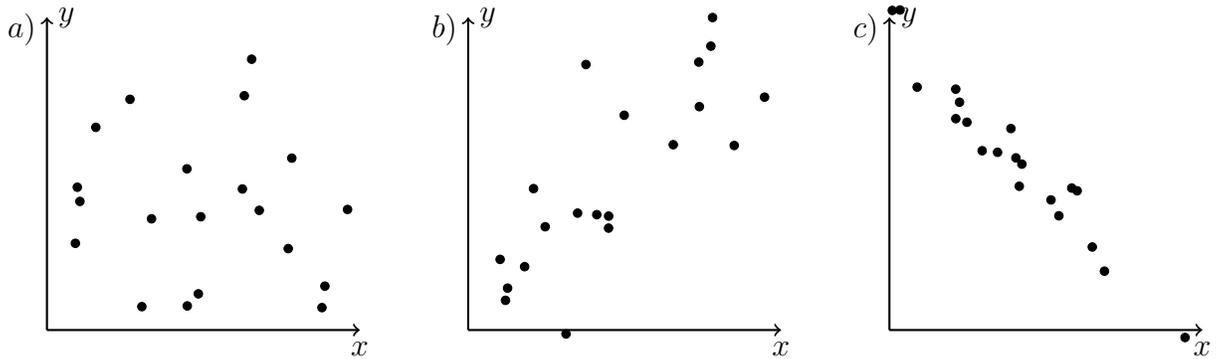
- (1) A newspaper conducts a survey to see who will vote for candidate Smith for mayor on Staten Island. A random sample of 1200 voters is taken, and 630 of them claim they will vote for Smith.
 - (a) Find an 80% confidence interval for the proportion of voters on Staten Island who will vote for Smith.
 - (b) Is this evidence that Smith will get more than 50% of the votes during the election? Test at the 10% significance level.

- (2) A survey in Brooklyn of 1000 votes finds that 475 of them will vote for Smith. Using the information from question 1, is this significant evidence that different proportions of voters in Staten Island and Brooklyn favour Smith?

- (3) 100 college students were asked what size pizza they liked, and what were their favourite toppings.

	Pepperoni	Veggie	Cheese	Total
Small	18	11	6	35
Medium	14	11	7	32
Large	3	9	21	33
Total	35	31	34	100

- (a) State the null hypothesis H_0 for a chi-square test based on this data.
 - (b) Assuming H_0 , what is the expected count for a large veggie pizza?
 - (c) What is the contribution to the chi-square statistic for the large veggie cell?
 - (d) What are the degrees of freedom df for the chi-square statistic?
 - (e) Suppose the chi-square statistic for this data is 22.8. What is the P -value (approximately)?
 - (f) At the 5% confidence level, what is the conclusion? What does it mean in this case?
-
- (4) Guess the correlation coefficients for the following data.



- (5) State whether the following questions are true or false:
- The least squares regression line is the line that makes the square of the correlation in the data as small as possible.
 - The least squares regression line is the line that makes the sums of the squares of the vertical distances from the data points to the line as small as possible.
 - The least squares regression line is the line that splits the data in half, with half the data above, and half below.
 - The least squares regression line passes through the point (\bar{x}, \bar{y}) , the means of the predictive and response variables, respectively.
- (6) Researchers studied a sample of 300 adults and found a strong negative correlation between the amount of vitamin X present in the bloodstream of an individual and the number of pounds overweight.
- State whether the following questions are true or false:
- This is quite strong evidence that large amounts of vitamin X cause you to not gain weight.
 - If the amount of vitamin X and the number of pounds overweight were plotted on a scatterplot for each individual in this study, then the points would lie close to a line with negative slope.
 - If a larger sample had been studied, then the correlation would be stronger.
 - If one of the data points has a negative residual, then it lies below the data line.

In the lab

- (1) The following table gives data about your instructor's internet usage and time spent in NY for the months of July to November.

	Jul	Aug	Sep	Oct	Nov
Internet usage in GB	41	10	59	52	22
Days present in NY	28	5	24	28	18

- What is the average number of days spent in NY during those months?
- What is the standard deviation of the number of GB used?
- Draw a graph of internet usage against days in NY.
- What is the correlation coefficient r for this data?
- What is the least squares regression line?
- What is the 90% confidence interval for the slope?
- If your instructor spends 25 days in NY in December, how much internet usage would you expect?