

Statistics 1040, Sections 003 & 004, Midterm 1 (200 Points)

February 14, 2003

Your Name: _____

From: Stat 1040, Fall 1999, Final Test, December 17, 1999, Question 19, 16
& Stat 1040, Fall 2002, Section 006, Midterm I, Question 3

Question 1: Observational Studies/Controlled Experiments (40 Points)

As part of a study on exercise and health, a group of 1,000 people was followed for 5 years. At the beginning of the the study, the researchers asked each person whether they exercised regularly or not. At the end of the study, the researchers measured several health-related variables, and in doing so, they noticed that the death rate for the exercise group was lower than for the no-exercise group.

1. (15 Points) Is the study described in this article an observational study or a controlled experiment? Circle your answer and explain!

Workbook:

They were asked what they did - they weren't told what to do.

2. (5 Points) When we follow people over a longer period of time, as in the study described in this article, we call such a study a Longitudinal study.

3. (20 Points) Does the result of this study necessarily imply that if people who do not exercise start to exercise regularly, they will live longer, on average, than if they do not? Explain!

Workbook:

No - this result merely shows that the type of people who choose to exercise differ from the type who do not.

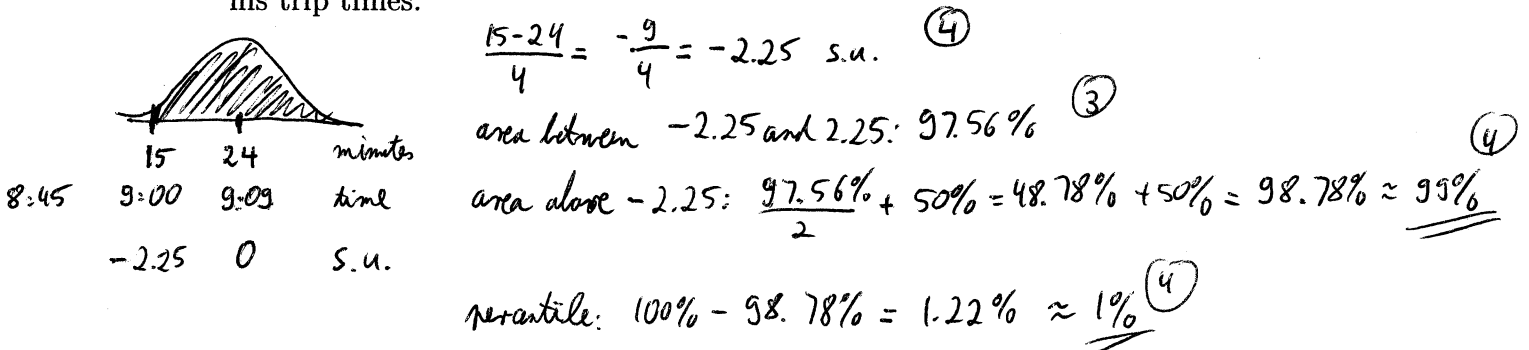
This is an observational study and there could be many confounding factors. For example, people who exercise might smoke less and it might be smoking that's harmful not exercise that's beneficial. Taking a smoker & making them exercise might have no impact.

Question 2: Normal Distribution (45 Points)

A lawyer commutes daily from his suburban home to his midtown office. On the average, one way of his trip (to work) takes 24 minutes with a standard deviation of 4 minutes. Assume that the histogram of his trip times closely follows the normal curve.

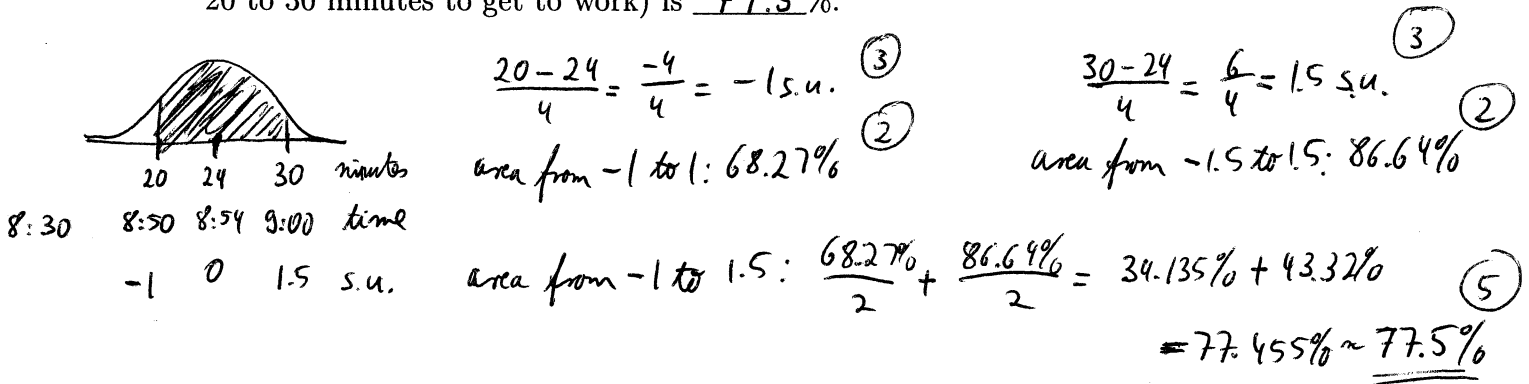
Fill the blanks in the statements below and show all the work needed to obtain the answer:

1. (15 Points) Assume the office opens at 9:00am and he leaves his house at 8:45am. The percentage of times he will be late for work (i.e., he needs more than 15 minutes to get to work) is 99%. This means 15 minutes is the 1stth percentile of his trip times.

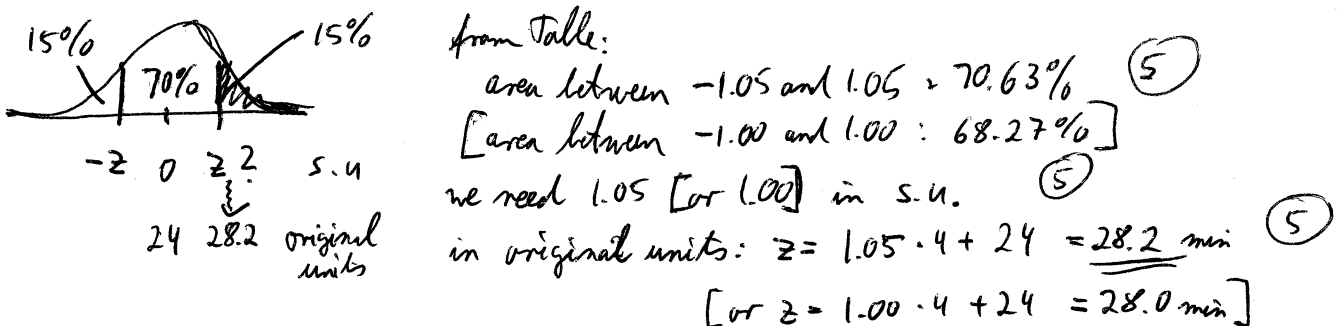


(3) if graph only (in any part)

2. (15 Points) Not satisfied with the result above, he decides to leave his house at 8:30am already. Coffee is served at his office from 8:50am to 9:00am. The proportion of times he arrives at the office during the time coffee is served (i.e., he needs between 20 to 30 minutes to get to work) is 77.5%.



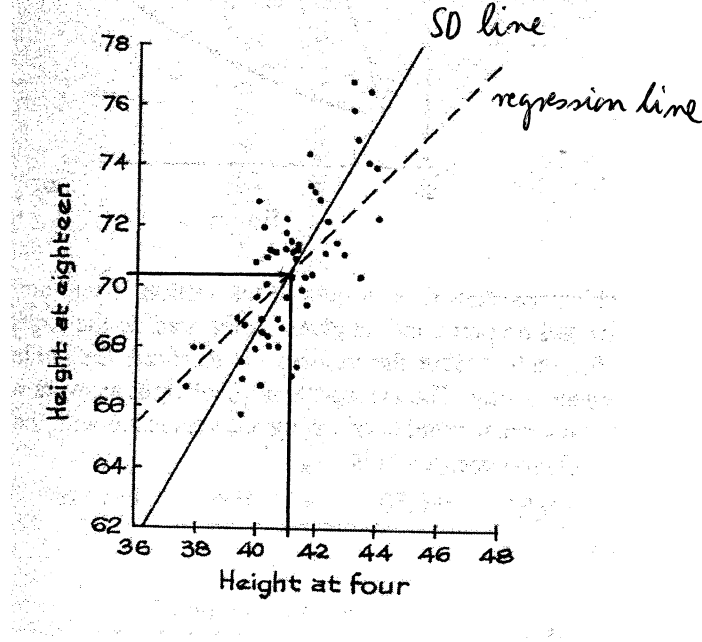
3. (15 Points) In 15% of his trips, he needs more than 28.2 minutes to get to work (i.e., the slowest 15% of his trip times are longer than this).



From: FPP, Chapter 8, page 137, Review Exercise 8 a, c, d

Question 3: Scatter Diagram & Correlation (20 Points)

A longitudinal study of human growth has been under way since 1929, at the Berkley Institute of Human Development. The scatter diagram below shows the heights of 64 boys, measured at ages 4 and 18.



1. (5 Points) The average height at age 4 is around

- 38 inches 42 inches 44 inches 66 inches 68 inches 71 inches

(2) SD line & regression line go through the point of averages; 42 inches is the closest choice

2. (5 Points) The average height at age 18 is around

- 38 inches 42 inches 44 inches 66 inches 68 inches 71 inches

(2) SD line & regression line go through the point of averages; 71 inches is the closest choice

3. (5 Points) The correlation coefficient is around

- 0.50 0.80 0.95

(2) the football is not extremely narrow, but also not too wide; 0.80 is the closest choice

4. (5 Points) Which is the SD line - solid or dashed?

(2) the SD line is steeper than the regression line

Explain your answers!

Question 4: Regression (45 Points)

A statistical analysis was made of the midterm and final scores in a large course with following results:

X : average midterm score ≈ 50 SD ≈ 25
 Y : average final score ≈ 55 SD ≈ 15
correlation $r \approx 0.60$

1. (15 Points) Predict the final score for a student whose midterm score was 80.

$$s_{u_x} = \frac{80-50}{25} = \frac{30}{25} = 1.2 \quad (5) \quad [s_{u_x} = \frac{x - \text{avg}_x}{SD_x}]$$

$$s_{u_y} = 0.6 \cdot 1.2 = 0.72 \quad (5) \quad [s_{u_y} = r \cdot s_{u_x}]$$

$$Y = 0.72 \cdot 15 + 55 = \underline{\underline{65.8}} \quad (5) \quad [Y = s_{u_y} \cdot SD_y + \text{avg}_y]$$

2. (15 Points) Predict the final score for a student whose midterm score was 15.

$$s_{u_x} = \frac{15-50}{25} = \frac{-35}{25} = -1.4 \quad (5)$$

$$s_{u_y} = 0.6 \cdot (-1.4) = -0.84 \quad (5)$$

$$Y = -0.84 \cdot 15 + 55 = \underline{\underline{42.4}} \quad (5)$$

3. (15 Points) Apparently, students who had very good midterm scores did not study enough later on and, consequently, did not do so well on the final. On the other hand, students with poor midterm scores worked hard during the rest of the semester and considerably improved their performance on the final. Do you agree with this statement? Why or why not? Explain!

Nothing unexpected happened - this is just the regression effect. (7)

The statement above is incorrect and an example for the regression fallacy. (8)

-3 if "regression effect" not stated,
but correct explanation

Question 5: Change of Scale (40 Points)

From the subjects in a health survey, the following data were collected:

average height = 68 inches SD = 2.5 inches
average blood pressure = 120 mm SD = 15 mm
correlation $r = -0.2$

You want to provide a summary of these results to a friend in Europe and report heights in centimeters (instead of inches) to make it easier for your friend to interpret the results. Recall that 1 inch = 2.54 centimeters.

1. (10 Points) The average height (in centimeters) is: $2.54 \cdot 68 = 172.72$ centimeters
2. (10 Points) The SD [for the height] (in centimeters) is: $2.54 \cdot 2.5 = 6.35$ centimeters
3. (10 Points) The correlation r is: -0.2 (this does not change when we change units!)

4. (10 Points) Circle the correct answer: r is measured in

- centimeters
- inches
- mm
- inches · mm
- centimeters · mm
- inches · centimeters
- none of these – r is a unitless number

Question 6: Valentine's Day Question (10 Points)

Mark the correct box:

(10 Points):

YES, I would like to get 10 points for free for taking the Valentine's Day Midterm.

(0 Points):

No, thank you, I don't need any additional points.