

Stat 1040 Recitation 4 Solutions

1. Leonardo da Vinci (1452-1519) theorized that if you put your arms out to the side and measured from the fingertip of one hand to the fingertip of the other, this "wingspan" distance would approximately equal your height. A group of fourth-grade students measured their height and wingspan and found

average height = 49.5 inches with an SD of 1.8 inches average wingspan = 48.9 inches with an SD of 2.1 inches.

The scatter diagram was football-shaped and the correlation coefficient was 0.8.

- (a) If one of these fourth-graders is 52 inches tall and is right on the SD line, what is their wingspan?

52" is 2.5" above the average.

That's $\frac{2.5}{1.8} = 1.4$ SDs above the average

They are 1.4 SDs above ave in wingspan (SD line)

That's $1.4 \times 2.1 = 2.94$ " above average

So their wingspan is $48.9 + 2.94 = \underline{\underline{51.8}}$ "

- (b) Predict the wingspan of a randomly chosen fourth-grader who is 52 inches tall.

52" is 2.5" above average

That's $\frac{2.5}{1.8} = 1.4$ SDs above ave

By regression, they should be around 0.8×1.4 SDs above average in wingspan. $0.8 \times 1.4 = 1.12$ SDs.

That's $1.12 \times 2.1 = 2.35$ " above average
 ≈ 2.4

So their wingspan should be $48.9 + 2.4 = \underline{\underline{51.3}}$ "

- (c) Find the rms error for your answer in part (b).

$$\sqrt{1-r^2} (2.1) = \underline{\underline{1.26}}$$

- (d) (2 points) Is it likely that the fourth-grader from part (b) would have a wingspan of 45 inches or more? Explain clearly using your answers to (b) and (c).

Yes - 95% of 4th graders who are 52" tall will be

between $51.3 + 2(1.26) = 53.8$ "

and $51.3 - 2(1.26) = 48.8$ "

2. For a random sample of 20 car models, the average weight (in pounds) was 3236, with an SD of 523. The average gas mileage (in miles per gallon) was 21.4 with an SD of 4.2. The correlation between weight and gas mileage was -0.87 . The scatter diagram was football shaped.

- (a) Ignoring weight, and assuming the histogram for gas mileage follows the normal curve, would you be surprised if someone told you that one of these cars got 27 miles per gallon? Explain your reasoning.



$$\frac{27 - 21.4}{4.2} = 1.3$$

They are only 1.3 SD's above average - this is not very unusual. (If they were more than 2 SDs above or below average, I'd be surprised)

- (b) Predict the gas mileage of a car that weighs 3500 pounds.

3500 is 264 pounds above average

That's $\frac{264}{523} = .5$ SDs above average

We expect them to be $(-.87)(.5) = -.435$ SDs above ave in mpg, that's .435 SDs below average.

.435 SDs is $(.435)(4.2) = 1.8$

so the mpg is $21.4 - 1.8 = \underline{19.6}$ mpg

- (c) Would you be surprised if someone told you that one of these cars weighing 3500 pounds got 27 miles per gallon? Explain your reasoning, using the rms error.

$$\text{RMS error} = \sqrt{1 - (-.87)^2} (4.2) = 2.07 \approx 2.1$$

95% → Most cars will be between $\underline{19.6} - 2(2.1) = 15.4$ and $19.6 + 2(2.1) = 23.8$ mpg.

So it would be surprising for such a heavy car to get such good gas mileage!

- (d) In one sentence, explain what the correlation coefficient tells you about the relationship between gas mileage and weight of cars like these.

Cars that are heavier tend to get fewer mpg.

The relationship is quite strong - if we know the weight of the car, we can predict its gas mileage quite accurately.

3. In a study on reading and math scores, tests were given to all third graders in Happy county. There were 13 schools in Happy County. The average math score and average reading score was calculated for each of these 13 schools and a scatter diagram was drawn with math score on the vertical axis and reading score on the horizontal axis. The correlation coefficient for these 13 points was .85. A newspaper reporter saw this and concluded "there is a very strong correlation between reading scores and math scores for Happy County third graders". What mistake was the reporter making? Explain briefly.

They calculated an ecological correlation because each of the 13 points was an average for a school. Ecological correlations tend to be artificially strong - if we looked at the children themselves, there would be a lot more scatter, so r would be quite a bit less than 0.85.

4. A teacher sees the positive correlation in Question 4 and concludes that "The positive correlation shows that if we focus on teaching verbal skills, math skills will improve also". What does statistics say?

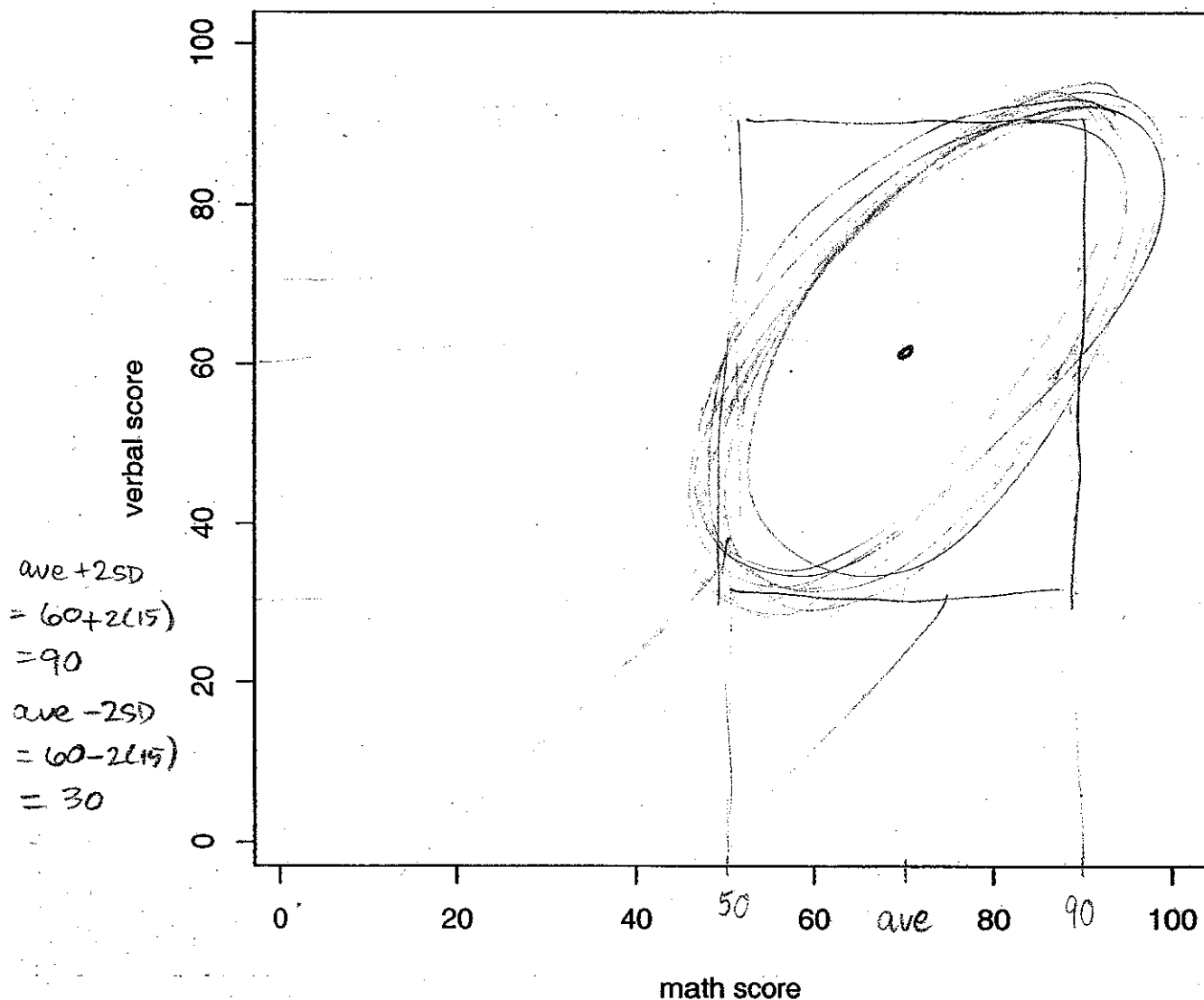
That statement is a causal statement - it's saying that if we increase verbal skills it will cause math skills to go up.

Statistics says that association is not causation - it could be that the positive correlation is an indicator of intelligence/diligence/parental involvement.

5. In one study students took a math test and a verbal test. The following summary statistics are calculated:

Average math score 70, SD = 10
Average verbal score 60, SD = 15 $r = 0.6$.

The scatter-diagram is football-shaped. Sketch the football.



ave + 2SD
= $60 + 2(15)$
= 90
ave - 2SD
= $60 - 2(15)$
= 30

math score

ave + 2SDs = $70 + 2(10) = 90$
ave - 2SDs = $70 - 2(10) = 50$

Oval should be quite fat because

$r = 0.6$.

Most of the oval should be in the box.