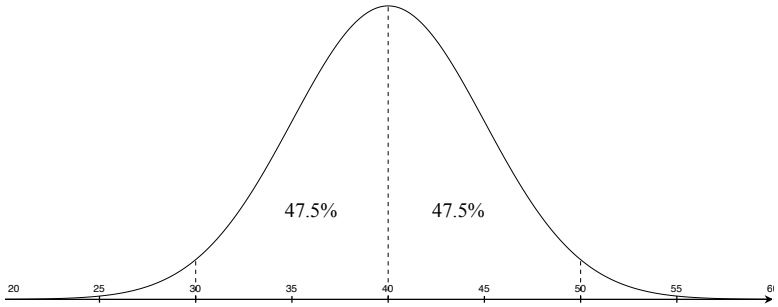
UCSC AMS 5 **Quiz 2 – Solutions** **Summer Session I, 2016**

- For full credit, ***show your work and explain your reasoning*** on every question.
- Please write clearly — if we can't read it, we won't give you credit for it.
- Please do ***not*** write your answers between the lines of the questions.

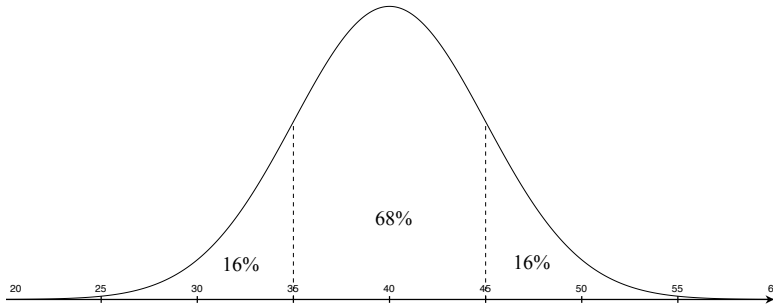
1. (3 pts) A certain set of data has average $\bar{x} = 40$ and standard deviation $SD_x = 5$. Assuming that the data has an approximately normal distribution, what percentage of the data lies between 30 and 40?

About 95% *of the data lies within 2 SD from the average,* 40*. I.e., about* 95% *of the data lies between* 30 *and* 50 *because* $10 = 2SD$ *in this case. The normal distribution is symmetric around the average, so about half of this data lies between* 30 *and* 40 *and half lies between* 40 *and* 50*. I.e., about* $47.5\% = 95\%/2$ *of the data lies between* 30 *and* 40*. See figure below*



2. (4 pts) What percentage of the data from question 1. lies above 45?

For approximately normally distributed data, about 68% *lies between one SD below to one SD above average. So in this case, about* 68% *of the data lies between* 35 *and* 45*. This means that about* $32\% = 100\% - 68\%$ *of the data lies either above* 45 *or below* 35*, and since the distribution is (approximately) symmetric, about half lies above* 45*, so about* $16\% = 32\%/2$ *of the data lies above* 45*. See figure below.*



3. (3 pts) A different set of data has average $\bar{y} = 18$, standard deviation $SD_y = 7$ and the biggest value in the data is $y_{max} = 24$. Does this data have an (approximately) normal distribution? Explain why or why not briefly, but precisely.

This data does ***not*** have an (even approximately) normal distribution, because there is ***no*** data more than 1 SD above average (the largest value is 24, which is less than 1 SD above average). If it were approximately normal, then about 16% of the data would lie more than 1 SD above average (like in problem 2).

4. (3 pts) Data from the 2005 *Current Population Survey* was used to summarize the relationship between age and educational level (measured in years of education) of U.S. women age 25 and above with the following statistics.

$$\begin{aligned} \text{average age} &\approx 50 \text{ years}, \quad SD_{age} \approx 16 \text{ years} \\ \text{average ed. level} &\approx 13.2 \text{ years}, \quad SD_{ed} \approx 3 \text{ years}, \quad r \approx -0.2 \end{aligned}$$

Does this mean that as women age, they become less educated? If not, what accounts for the negative correlation?

No, women do not 'lose' years of education as they age. The survey data was collected in one year, not over a period of many years. This means that the older women in the survey were born years before the younger women in the survey. This explains the negative correlation between age and education — women born in the 1940s or 1950s were much less likely to go to college, for example, than women born in the 1960s and 1970s (and later).

5. In 2005, the *Educational Testing Service* observed the average Math SAT score for each of the 51 states (including D.C.) as well as the percentage of high school seniors in the state who took the test. The correlation between these two variables was -0.84.

(a) (4 pts) True or False: test scores tend to be lower in states where a higher percentage of students take the test. If true — explain why. If false — explain the negative correlation.

True: *the negative correlation does say that the higher the percentage of students that take the test in a state, the lower the average score tends to be in that state. The simplest explanation is that in states where a large proportion of students take the SAT, there is more variation in the students' math abilities, so there are more weak students whose scores bring the average down. If the percentage of students taking the test is small in a particular state, then it is likely that the student who do take the test are the stronger students, for the most part, so the average score in these states is higher.*

(b) (3 pts) In Connecticut the average Math SAT score in 2005 was 517, and in Iowa it was 608. Do the data show that schools in Iowa do a better job teaching math than schools in Connecticut? Explain your answer.

It is possibly true that Iowa schools are better at teaching math than Connecticut schools, but ***the data does not show this.*** *In fact the data offers another possible explanation for the lower average score in Connecticut: perhaps the percentage of students taking the test in Connecticut is higher than in Iowa.*