

Empirical and Chebyshev's Rule Review

Empirical rule only works for bell-shaped, symmetric data.

- * *Approximately 68% of the values will lie within one standard deviation of the mean*
- * *Approximately 95% of the values will lie within two standard deviations of the mean*
- * *Approximately 99.7% of the data will lie within three standard deviations of the mean*

The Empirical rule is sometimes called the "68-95-99.7 Rule".

Chebyshev's Rule - valid for *any* data set, **it states the minimum amount of data** that will fall within **k** standard deviations of the mean.

For any number greater than 1, the proportion of data that will fall within **k** standard deviations of the mean is **at least** $1 - \frac{1}{k^2}$ where **K** is the number of standard deviations away from the mean. If we have a symmetric interval of the form: $[\mu - k\sigma, \mu + k\sigma]$, we know at least $(1 - \frac{1}{k^2})100\%$ of the data will lie inside the interval.

k	k ²	1/k ²	(1 - 1/k ²)%
2	4	.25	75%
3	9	.11	89%
4	16	.0625	93.75%

1. The 2014 Australian Open statistician reported that the mean serve speed of the best women tennis players was 100 miles per hour (mph) and the standard deviation was 10 mph.

a) Assume that the statistician also gave us the information that the distribution of serve speeds was **mound shaped and symmetric**.

What proportion of serves was between 90 mph and 130 mph? **68/2 + 99.7/2 = 83.5%**

b) If nothing is known about the shape of the distribution, **give an interval of speeds** that will contain the speeds of **at least 15/16** serves.

K = 4 100 ± 40 (60, 140)

c) Suppose that the statistician indicated that the serve speed distribution was **skewed to the right**. Which of the following values is most likely the value of the **median** serve speed?

- a. 105 mph **b. 95 mph** c. 100 mph d. 108 mph

2. Suppose the average height and standard deviation of 50 students in a class are 66 inch and 3 inch respectively.

a) If **nothing is known** about the shape of the distribution, what proportion represents the number of students **outside** the interval **from 60 to 72 inch**?

At most 25%

b) If the heights have a **mound shaped and symmetric** histogram, what proportion of the observations will be **less than 57 inch**?

0.15%

c) If the heights have a **mound shaped and symmetric** histogram, what proportion of the observations will be **less than 60 and more than 69 inch**?

2.5% + 0.15% = 2.65%

3. Solar energy is considered by many to be the energy of the future. A recent survey was taken to compare the cost of solar energy to the cost of gas or electric energy. Results of the survey revealed that the distribution of the amount of the monthly utility bill of a 3-bedroom house using gas or electric energy had a mean of \$121 and a standard deviation of \$13. **If nothing is known about the shape of the distribution, what percentage of homes will have a monthly utility bill of less than \$95?**

At most 25%

4. Suppose a light bulb manufacturer claims that the mean lifetime of its bulbs is 35 hours. Assume you have prior knowledge that the bulb lifetimes have mound shaped distribution with a standard deviation of 5 hours.

a.) If the manufacturer's claim is true, what percent of light bulbs will burn out in less than 20 hours?

50% - 99.7/2 = 0.15%

b.) Suppose you randomly select one of the bulbs and it burns out in less than 20 hours. Do you suspect the manufacturer's claim is incorrect? EXPLAIN.

Yes, since Probability = 0.15%

c.) What percentage of bulbs can be expected to burn out between 30 and 45 hours?

68/2 + 95/2 = 81.5%

5. A professor believes that if a class is allowed to work on an examination as long as desired, the times spent by the students would be approximately mound shaped with a mean of 40 minutes and a standard deviation of 6 minutes. Approximately how long should be allotted for the examination if the professor wants almost all, say 97.5%, of the class to finish?

40 + 2 x 6 = 52 min

6. Suppose the mean and SD are 74 and 10, respectively. **If we assume that the distribution is mound-shaped and symmetric, what percentage of the data will be between 54 and 84?**

95/2 + 68/2 = 81.5%

7. Suppose, at FIU the μ and σ of all students cumulative GPAs, are 2.5 and 0.5, respectively. The president of FIU wishes to graduate the top 2.5% of the students with cum laude honors and the top .15% with summa cum laude honors. Assume that distribution for the GPAs scores is mound shaped and symmetric.

a) Where should be the limits be set in terms of GPAs? **3.5, 4.0**

b) In terms of percentile scores? **$P_{97.5}$, $P_{99.85}$**

8. Time to take standardized Exam is known to have mound shaped and symmetric distribution with $\sigma = 10$ min and $P_{2.5} = 55$ min.

How much time will it take for 50% of the entire class to finish this Exam? **75 min**

How much time will it take for 97.5% of the entire class to finish this Exam? **95 min**

9. The time to complete a Test have a mound – shaped distribution with mean of 52 min and SD = 4 min.

a) What percentage of students will complete the exam in less than 48 min? **16%**

b) How much time should be allowed, if I need to be sure that 97.5% of entire class finished the test? **60 min**

c) What percentage of students will complete the exam between 44 and 64 min? **$95/2 + 99.7/2 = 97.35\%$**

10. The time to complete a Test has unknown distribution with mean of 52 min and SD = 4 min.

a) What percentage of students will complete the exam after 1 hr 8 min? **At most 1/16**

b) How much time has to be provided, if I need at least 8/9 of entire class to finish the test? **64 min**

11. You want to buy a fuel efficient hybrid, but you do not know if the new technology is reliable yet. A Consumer Report article list the average repair cost over five years of ownership as \$658 with a standard deviation of \$85. If you decide that the maximum amount of money that you can afford to spend over the five years on repairs is \$1000, is the hybrid a safe buy? Explain.

Yes, $1000 - 658 = 342$; $342 : 85 = 4$; $K = 4 \rightarrow$ At least 15/16 of all repairs will fall below \$1000

12. The average of a set of data points is 13.5, and the median of the same data set is 8.2. Is the distribution skewed? If so, is it right or left skewed?

Skewed Right

13. The average human can run at a maximum short distance speed of 22 miles per hour with a standard deviation of 2 miles per hour. A runner from Kenya was recently clocked at a speed of 28 while running the 100 yrd dash. Is this speed considered unusually fast for a human being?

Yes, $k = 3$

14. Which test result is relatively better: an 85 on a psychology exam or a 45 on an economics exam? The psychology exam has a mean of 90 and a standard deviation of 10, and the economics exam has a mean of 55 and a standard deviation of 5?

Psychology exam $Z_1 = -0.5$ $Z_2 = -2.0$