Stats Review
Chapter 3

Note:

This review is composed of questions similar to those found in the chapter review and/or chapter test. This review is meant to highlight basic concepts from the course. It does not cover all concepts presented by your instructor. Refer back to your notes, unit objectives, handouts, etc. to further prepare for your exam.

The questions are displayed on one slide followed by the answers are displayed in red on the next.

This review is available in alternate formats upon request.

Find the Mean, Median, Mode, and Range

Data Set: 71, 74, 67, 64, 72, 71, 65, 66, 69,70

Find the Mean, Median, Mode, and Range

Data Set: 71, 74, 67, 64, 72, 71, 65, 66, 69,70

Mean: add up all the numbers and divide by the amount of numbers

$$\frac{71+74+\ 67+\ 64+72+71+65+66+69+70}{10} = \mathbf{68.9}$$

Median: Arrange numbers from smallest to largest then find the middle number

64, 65,66, 67,69,70,71, 71, 72, 74

The middle number is between 69 and 70. Finding the mean of these two numbers will give us the median which is **69.5**

Mode: The most repeated number (can have none, one, or more than one)

The mode is **71**.

Range: The highest number – the smallest number

74-64 = **10**

Find the Five-Number Summary and construct a box plot.

Data Set: 64, 65, 66, 67, 69, 70, 71, 71, 72, 74

The data set is the same as previous problem but in order

Find the Five-Number Summary and construct a box plot.

Data Set: 64, 65, 66, 67, 69, 70, 71, 71, 72, 74

The data set is the same as previous problem but in order

Min: the smallest number **64**

Q1: Find the median of the numbers from the minimum to the median (do not include the median)

These numbers are 64, 65, 66, 67, 69. The median of these numbers are 66. So Q1 is **66.**

Median: the middle number **69.5**

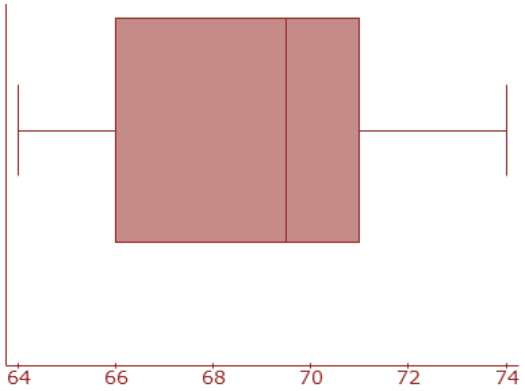
Q3: Find the median of the numbers from the median (not included) to the maximum.

These numbers are 70, 71, 71, 72, 74. The median of these numbers are 71. So Q1 is **71.**

Max: the largest number **74**

The five-number summary is

64 66 69.5 71 74



Find the upper and lower Fences. Are there any outliers?

Data Set: 71, 74, 67, 64, 72, 71, 65, 66, 69, 70

Find the upper and lower fences. Are there any outliers?

Data Set: 71, 74, 67, 64, 72, 71, 65, 66, 69, 70

1) Find the IQR. IQR=Q3-Q1=71-66=5
2) Multiply the IQR by 1.5: 5*1.5=7.5
3) **Upper Fence**: Add 7.5 to Q3; 7.5+71= **78.5**
4) **Lower Fence**: Subtract 7.5 from Q1; 66-7.5= **58.5**

Any numbers outside this range would be considered outliers.

There are no numbers outside this range so there is no outlier.

For the following data set find the standard deviation as if the data was a) from a sample b) from a population. The mean is 15.8.

Data Set: 20,19,11,15,14

For the following data set find the standard deviation as if the data was a) from a sample b) from a population. The mean is 15.8.

Data Set: 20,19,11,15,14

Steps 1-3 are the same for sample and population standard deviation

	Step 1	Step 2
Data, x_i	$x_i - \mu$	$(x_i - \mu)^2$
20	20-15.8=4.2	$(4.2)^2 = 17.64$
19	3.2	10.24
11	-4.8	23.04
15	-.8	.64
14	-1.8	3.24

Step 3: Add numbers from step 2 =54.8

a) Sample Standard Deviation

Step 4: Divide by n-1= $\frac{54.8}{5-1} = 13.7$

Step 5: Take the square root $\sqrt{13.7} = \mathbf{3.7}$

b) Population Standard Deviation

Step 4: Divide by n= $\frac{54.8}{5} = 10.96$

Step 5: Take the square root $\sqrt{10.96} = \mathbf{3.3}$

Relationship Between Mean, Median and Shape

Is the distribution of the data skewed left, skewed right, or symmetric. Match the shape to the relationship between mean and median.

Mean>Median Mean=Median Mean<Median

Relationship Between Mean, Median and Shape

Is the distribution of the data skewed left, skewed right, or symmetric.
Match the shape to the relationship between mean and median.

Mean>Median Mean=Median Mean<Median

Skewed Left

Mean<Median

Symmetric

Mean=Median

Skewed right

Mean>Median

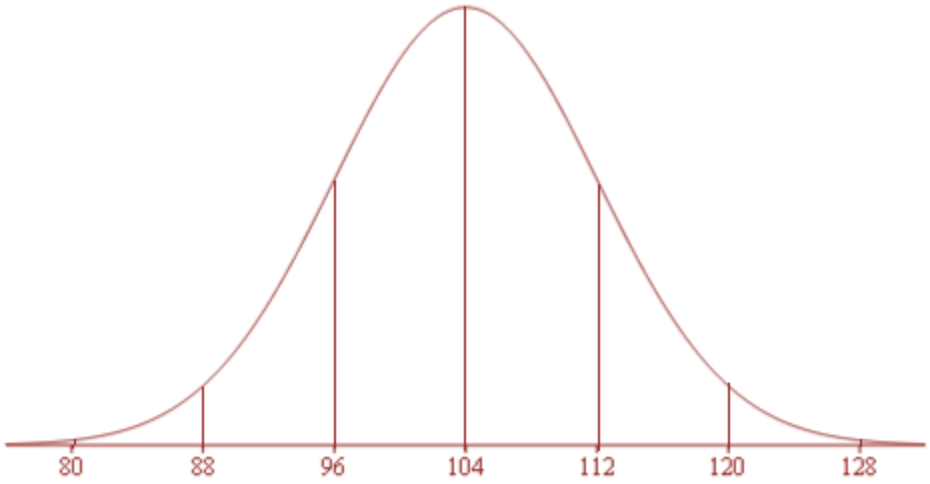
If data is skewed, median is more representative of the typical observation.

Empirical Rule to Find the Probability

At a tennis tournament, a statistician keeps track of every serve. She reported that the mean serve speed of a particular player was 104 mph and the standard deviation of the serve speeds was 8 mph. Assume that the statistician also gave us the information that the distribution of the serve speeds was bell shaped. Using the Empirical Rule, what proportion of the player's serves are expected to be between 112 mph and 120 mph?

Empirical Rule to Find the Probability

At a tennis tournament, a statistician keeps track of every serve. She reported that the mean serve speed of a particular player was 104 mph and the standard deviation of the serve speeds was 8 mph. Assume that the statistician also gave us the information that the distribution of the serve speeds was bell shaped. Using the Empirical Rule, what proportion of the player's serves are expected to be between 112 mph and 120 mph?



Start by drawing the bell curve
Compare this to the empirical rule curve on page 149 or find the percent as described on page 149. We can see that the probability is **13.5% or .135.**

Chebyshev's Inequality

Using the information from the previous slide and Chebyshev's Inequality

a) determine the minimum percentage of tennis serves within 2.5 hours from the mean.

b) determine the minimum percentage of tennis serves with a speed of 92 and 116?

Chebyshev's Inequality

Using the information from the previous slide and Chebyshev's Inequality

a) Determine the minimum percentage of tennis serves within 2.5 standard deviations from the mean.

b) Determine the minimum percentage of tennis serves with a speed of 92 and 116?

Chebyshev's Inequality: At least $\left(1-\frac{1}{k^2}\right)100\%$ of the observations lie within k standard deviations of the mean.

a) $\left(1-\frac{1}{2.5^2}\right)100\%$= **84%** of the observations lie within 2.5 standard deviations of the mean.

b) First find how many standard deviations 92 and 116 is away from the mean.

Both are within 1.5 standard deviations away from the mean ($\frac{92-104}{8}=-1.5$ and $\frac{116-104}{8}=1.5$) so k=1.5

$\left(1-\frac{1}{1.5^2}\right)100\%$= **55.5%** of the observations lie within 92 and 116 mph.

Comparisons

Test scores for Mr. Nelson's stats class have a mean of 64 and a standard deviation of 7.4. Test scores for Ms. Jackson's chemistry class have a mean of 68 and standard deviation of 4.9. Suppose a student gets a 72 on the stats test and a 73 on the chemistry test. On which test did the student preform better relative to the other students in each class?

Comparisons

Test scores for Mr. Nelson's stats class has a mean of 64 and a standard deviation of 7.4. Test scores for Ms. Jackson's chemistry class has a mean of 68 and standard deviation of 4.9. Suppose a student gets a 72 on the stats test and a 73 on the chemistry test. On which test did the student preform better relative to the other students in each class?

1) Calculate z for both classes $z = \frac{x-\mu}{\sigma}$

Stats: $z = \frac{72-64}{7.4} = 1.08$ Chemistry: $z = \frac{73-68}{4.9} = 1.02$

2) The test with the higher z-score is where the student preformed better relative to the other students in the class.

Stats is the test where the student preformed better relative to the other students in the class.

Find the Mean

Days	Frequency
1-2	2
3-4	21
5-6	20
7-8	10
9-10	30

Find the Mean

Days	Frequency
1-2	2
3-4	21
5-6	20
7-8	10
9-10	30

	Step 1↓		Step 2↓
Days	**Class Midpoint, x_i**	**Frequency, f_i**	**$x_i f_i$**
1-2	$\frac{1+2}{2} = 1.5$	2	1.5*2=3
3-4	3.5	21	73.5
5-6	5.5	20	110
7-8	7.5	10	75
9-10	9.5	30	285

Step 3: Find the sum of the $x_i f_i$ column: 3+73.3+110+75+285=546.5

Step 4: Find the sum of the frequency column: 83

Step 5: Divide the number from step 3 by the number from step 4: 546.5/83=6.6

The mean is 6.6.

Find the Sample Standard Deviation

Data From Previous Problem

Days	Frequency
1-2	2
3-4	21
5-6	20
7-8	10
9-10	30

Find the Sample Standard Deviation

Data From Previous Problem

Step 1: Find the mean (see previous slide). Mean=6.6

Step 2: Find the midpoint-mean: $x_i - \bar{x}$

Step 3: Find $(x_i - \bar{x})^2 f_i$ i.e. square each number from step 2 and multiply by the frequency.

				Step 2↓	**Step 3↓**
Days	**Class Midpoint, x_i**	**Frequency, f_i**	$\bar{x}$	$x_i - \bar{x}$	$(x_i - \bar{x})^2 f_i$
1-2	$\frac{1+2}{2} = 1.5$	2	6.6	1.5-6.6=-5.1	$(-5.1)^2 * 2 = 52.02$
3-4	3.5	21	6.6	-3.1	201.81
5-6	5.5	20	6.6	-1.1	24.2
7-8	7.5	10	6.6	.9	8.1
9-10	9.5	30	6.6	2.9	252.3

Step 3: Find the total of the $(x_i - \bar{x})^2 f_i$ column. Total =538.43

Step 4: Find the total of the frequency column. Total =83. Take this number and subtract 1. 83-1=82

Step 5: Take the number from Step 3 and divide by the number from step 4.

538.43/82=6.56622

Step 6: Take the square root of the number from step 5. $\sqrt{6.56622} = \mathbf{2.6}$