LESSON 10: MEASURES OF RELATIVE POSITION

How Does a Value Compare (“Stack Up”) to the Others?

PART A: 100 TEST SCORES

Example 1 (100 Test Scores)

Here is a list of 100 test scores in a class. You took that test and scored 50 points. How did you do relative to the class?

70  40  60  93  86  87  13  86  66  80  
71  65  95  75  69  47  72  100  39  99  
90  63  81  88  82  51  89  58  46  95  
50  15  81  52  47  76  90  70  89  44  
46  53  46  52  37  38  54  64  70  40  
94  60  63  56  92  92  63  86  56  87  
56  51  83  79  91  87  53  97  84  49  
91  81  69  82  77  91  71  90  51  92  
68  88  38  84  54  47  50  14  48  29  
61  97  79  36  52  93  46  92  28  38  

§

We want standardized measures that will work for practically all populations involving quantitative data.
PART B: z SCORES

The formula below will convert (or “transform”) the original data set of $x$-values into a new data set of $z$ scores, which will tell us how many standard deviations above (or below) the mean the $x$-values are.

Formula for $z$ Scores

$$z = \frac{x - \text{mean}}{\text{SD}}$$

(Idea: $\Rightarrow$ new mean = 0
\Rightarrow new SD = 1)

• Subtracting the mean from all of the original $x$-values recenters the data set so that the new mean will be 0; we obtain the deviations from the mean.

• Dividing by the SD rescales the data set so that the new SD will be 1.

Notation

Population (Size $N$) $z = \frac{x - \mu}{\sigma}$

Sample (Size $n$) $z = \frac{x - x}{s}$

Round off $z$ scores to two decimal places.

$z$ scores have no units, because we divide by the SD. We can use $z$ scores in many different applications where different units are involved. $z$ scores are standardized measures of relative position.

• If we are studying heights, for example, we may use inches, feet, meters, etc. and still obtain the same $z$ scores.
Signs of $z$ Scores; Interpretations

- If $z > 0$, then the $x$-value is $z$ SDs above the mean.
- If $z = 0$, then the $x$-value is at the mean.
- If $z < 0$, then the $x$-value is $|z|$ SDs below the mean.

Example 2 ($z$ Scores: 100 Test Scores)

In Example 1, you scored 50 points on that test. The mean on the test was about 66.9 points, and the SD was about 21.4 points. What was your $z$ score? Interpret that result.

§ Solution

\[
 z = \frac{x - \text{mean}}{\text{SD}} \quad \left(\text{that is, } \frac{x - \mu}{\sigma} \text{ for population data}\right)
\]

\[
 \approx \frac{50 - 66.9}{21.4} \\
\approx -0.79
\]

Your $z$ score was about $-0.79$, which means that you scored about 0.79 standard deviations below the mean. §

The “Two SD” Rule for Unusual $z$ Scores

$z$ scores that are either less than $-2$ or greater than 2 (that is, $z < -2$ or $z > 2$) are considered “unusual.” The corresponding $x$-values are also considered unusual.

- In other words, data values that are more than 2 SDs from the mean are considered “unusual.”
Think About It: How would you feel if your instructor tells you that your z score for the test you just took was −2.52?

**Example 3 (Interval of Usual Values: Revisiting Lesson 9, Example 4)**

In Lesson 9, Example 4, the test scores had mean $\mu = 50$ points and SD $\sigma = 10$ points. The **interval of usual values** was $(\mu - 2\sigma, \mu + 2\sigma)$, which was (30 points, 70 points). What are the z scores for 30 points, 50 points, and 70 points?

§ **Solution**

- 30 points is 2 SDs **below** the mean, so its $z = -2$.
- 50 points is **at** the mean, so its $z = 0$.
- 70 points is 2 SDs **above** the mean, so its $z = 2$.

"Realistic Range" $\approx 4\sigma = 40$
Example 4 (Interval of Usual Values: 100 Test Scores; Revisiting Examples 1, 2)

In Example 2, the mean on the test was about 66.9 points, and the SD was about 21.4 points. Find the interval of usual values.

§ Solution

The interval of usual values is given by: \( \mu - 2\sigma, \mu + 2\sigma \).

\[
\begin{align*}
\mu - 2\sigma & \approx 66.9 - 2(21.4) = 24.1 \text{ points} \\
\mu + 2\sigma & \approx 66.9 + 2(21.4) = 109.7 \text{ points}
\end{align*}
\]

If the maximum possible score on the test is 100 points, then 109.7 points might not make sense. We could “truncate” and give the interval of usual values as \([24.1 \text{ points, 100 points}]\).

§

Think About It: If you scored 50 points on the test, would that score be unusually low? Note: Your \( z \) score was about \(-0.79\).

Think About It: If you scored 13 points on the test, would that score be unusually low? Note: Your \( z \) score was about \(-2.52\).
Example 5 (Comparing z Scores from Different Populations)

Two statistics classes take different tests. The mean and SD for Class 1 are: $\mu_1 = 55$ points, $\sigma_1 = 5$ points.
The mean and SD for Class 2 are: $\mu_2 = 60$ points, $\sigma_2 = 10$ points.

In which class is a score of 70 points more impressive?

§ Solution

Find the corresponding $z$ scores for 70 points in the two classes. Use:

$$ z = \frac{x - \text{mean}}{\text{SD}} $$

Class 1:

$$ z = \frac{70 - 55}{5} = 3.00 $$

Class 2:

$$ z = \frac{70 - 60}{10} = 1.00 $$

A score of 70 points is more impressive in Class 1; in fact, it is “unusually” high in that class.

Note: If we assume normality in both classes, we have:
PART C: QUANTILES

We will look at three types of quantiles, which help us describe the relative position of data values.

**Percentiles**

Think: Pennies.

- That is, think about how a dollar can be broken up into 100 pennies and how these pennies can be separated.

*Example 6 (Percentiles)*

Because 92% of the scores lie below 85 points, we can say that the 92nd percentile of the data is 85 points: $P_{92} = 85$ points.

- (Some ties may be permitted.)

**Deciles**

Think: Dimes.

\[1\text{st decile} = D_1 = P_{10} \quad \text{(Think: 1 dime is worth 10 pennies.)}\]

\[2\text{nd decile} = D_2 = P_{20}\]

\[\vdots\]

\[9\text{th decile} = D_9 = P_{90}\]
**Quartiles**

Think: Quarters.

- 1\(^{st}\) quartile = \(Q_1 = P_{25}\) \(\text{ (Think: 1 quarter is worth 25 pennies.)}\)
- 2\(^{nd}\) quartile = \(Q_2 = P_{50}\) \(\approx \text{Median, usually; rounding issues may arise})\)
- 3\(^{rd}\) quartile = \(Q_3 = P_{75}\)

The Five-Number Summary of a quantitative data set consists of the \textbf{Min} (the lowest observed value), \(Q_1\), the \textbf{Median}, \(Q_3\), and the \textbf{Max} (the highest observed value).

![Five-Number Summary Diagram]

**PART D: BOXPLOTS**

The \textbf{Five-Number Summary} can be represented graphically as a boxplot (or box-and-whisker plot).

![Boxplot Diagram]

Boxplots can help us compare different populations, such as men and women.
**Example 7 (Boxplot: 100 Test Scores; Revisiting Examples 1, 2)**

Below is a boxplot for our 100 test scores from Examples 1 and 2.

![Boxplot]

We see that:

- Min = 13 points
- $Q_1 = 50$ points
- Median = 69 points
- $Q_3 = 87$ points
- Max = 100 points

**Example 8 (Comparing Boxplots)**

Think About It: **Compare** the two boxplots below; the **same scale** is used for both. The first boxplot is from Example 7. The second boxplot is for another large class that took the same test.

![Boxplot]

- Which boxplot suggests a **symmetric** distribution?
- Which boxplot suggests possible **skewness** to the left?
PART E: IQR

We can now define the Interquartile Range (IQR), the fourth measure of variation that was promised in Lesson 9.

\[ \text{IQR} = Q_3 - Q_1 \]

- In a boxplot, this is the length of the two sides of the box that are parallel to the axis.
- This is the range of the “middle 50%” of the data set.

Example 9 (IQR: Revisiting Example 8)

In Example 8, the IQRs are:

- IQR = \( Q_3 - Q_1 = 87 - 50 = 37 \) points for the first boxplot.
- IQR = \( Q_3 - Q_1 = 30 - 20 = 10 \) points for the second boxplot.

The range is \( \text{Max} - \text{Min} = 100 - 13 \) points = 87 points for the first boxplot.

FOOTNOTES (OPTIONAL)

#1) Quantiles. Quantiles can be thought of as numerical break points that divide quantitative data into approximately equally-populated intervals.

#2) Calculating quantiles. There are many ways to calculate percentiles, quartiles, etc. Rounding issues can come into play. Sometimes, linear interpolation is used (for example, when the median is taken to be the average of two middle values in sorted data).

#3) Outliers. Outliers are sometimes defined as values that lie more than 1.5 IQR away from the “box” (that is, farther than 1.5 IQR lower than \( Q_1 \) or farther than 1.5 IQR higher than \( Q_3 \)). An outlier may be indicated by an asterisk (*) in a boxplot, with the whiskers extending from the box in such a way that they ignore outliers. For example, the score of 90 points below is an outlier.