USING Z-SCORES

0:00 / 6:57 1.0x

SPEAKER: MICHAEL J. MAHOMETA, Ph.D.

When we've been looking at distributions,

we've noticed that shape plays a pretty major role in some of our decision

making.

If the shape is skewed - in other words, it's not symmetrical

- we need to approach our question a little bit differently

by using slightly different measures of

Help

center and spread

than what we might be used in the past.

In general, the symmetrical distribution wins our favor every time.

But it's not just a symmetrical distribution.

After all, a uniform distribution or a bimodal distribution

can also be considered symmetrical.

If a distribution of data falls into a category of shape classified

as "normal" then wonderful things can happen.

But more on that later.

First, let's start with a question.

Right up until that last little bit of information, everything was fine.

Let's break down this into smaller questions.

First, "Can we compare two students from two classes

on who took an exam on the same material?"

In general, the answer is yes - as long as we were confident

that the exams were close in material and difficulty.

It would be even better if the exams were exactly the same.

Help

Then we could simply compare the actual values:

John got an 83 while Jayne got an 89, so Jayne performed better on the exam.

But what if the exams were NOT all that similar?

What if one was harder than the other in terms of material?

Could we compare John and Jayne then?

The answer is yes.

We can do it with the help of z-scores.

You see, z-scores are a way to give everyone within a single distribution -

regardless of what the values are for that distribution - the same "ruler."

And here's how: First, we start with the difference between a single score

and the mean of the distribution that that score came from.

For John, the mean of his distribution (or his class)

was 74, while for Jayne her mean was 80.

Both John and Jayne show a difference to the mean of 9 points.

But that's not the whole story: each distribution has a measure of center

and a measure of SPREAD.

In our case, because we're assuming a

Help

normal, symmetrical, distribution,

the Standard Deviation is the measure of choice for spread.

So let's use it.

We take the difference that we've already found,

and we divide by the Standard Deviation of each distribution.

John's distribution Standard Deviation is 4,

while for Jayne, her distribution Standard Deviation is 6.

So, if we solve for both z-scores, we find

that John's z-score of exam performance is 2.25,

while for Jayne, her z-score for her performance is 1.5.

So, who performed better?

Well Jayne got a HIGHER raw score, but her class

did better overall (maybe it was just an easier exam).

And more importantly, Jayne's class had a larger Standard Deviation.

It had more spread around the mean more variability to each one's score.

So on the "ruler" of performance - John was 2.25 units better than average,

while, for Jayne, she was only 1.5 units

Help

better than average.

And we call these particular units z-scores.

But that's not all.

For a limited time, we'll also throw in proportion under the curve.

That's right proportion under the curve.

Here's a typical "bell shape" curve: One peak, symmetrical, most

cases fairly close to the center or mean,

and fewer cases going out in both directions from the mean.

We know that with this idea of a symmetrical distribution,

we get to use the mean and the Standard Deviation

as our measure of center and spread.

We know that for a normal distribution that mean is right in the middle,

and the Standard Deviation occurs at the point at which the "curves" sort

of change - about here and here.

Now it turns out that the normal distribution, this wonderful curve,

has some really stable properties that allow us to find out

the percentage of cases that fall both above

or below or inside or outside particuliar z-scores.

Help

Here's the empirical rule: it's a great rough guideline

to start with when it comes to area under the curve.

Here's where the first Standard Deviation

falls (right about here - where the curve changes).

The empirical rule says that 68% of the cases in this distribution will fall

between these first Standard Deviations, of -1 and +1 for a z-score (one

Standard Deviation below and 1 Standard Deviation above the mean).

95% of the cases will fall between -2 and +2 in terms of z scores or Standard

Deviations, and 99.7% will fall between -3 and +3.

In actuality, we can find the proportion under the curve

for ANY z-score we wish.

Let's go head and take a look at John and Jayne again.

Here's Jayne's score of 89 and here is her assumed normal distribution

with a mean of 80 and a Standard Deviation of 6.

And here's her corresponding z-score of 1.5.

Using a z-table or a function in R, we can

Help

find that roughly 7% of her class

performed better than her.

While for John, on the other hand, with his specific distribution

and his z score, we can see that only 1% performed better than him.

What's great about this is that z-scores will work on any distribution scale.

So, if we wanted to see who preformed better on an exam

where John's exam was actually out of 100

and Jayne's exam was out of 75 points, we could STILL do it.

We simply convert to z-scores first.

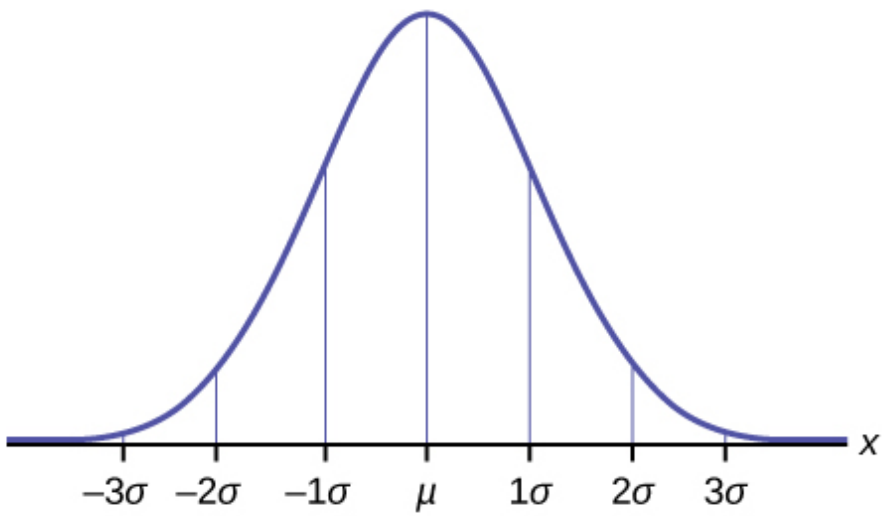
So, to answer our question of who did better on the exam,

and get the added bonus of finding the percent of distribution that

reflects that performance, we can use z-scores.

Comprehension Check

A full-term baby in the US weighs, on average, 7.5 pounds. Birth weights are normally distributed, with a standard deviation of 1.1 pounds. Sketch this distribution and use your sketch to help you answer the following questions.



Using the EMPIRICAL RULE, we can say that:

(3/3 points)

68% of all full-term babies weigh

- between 6.4 and 8.1 pounds.
- between 6.4 and 8.6 pounds. ✔
- between 5.9 and 8.6 pounds.
- between 5.9 and 8.1 pounds.

Help

95% of all full-term babies weigh

- between 5.3 and 9.7 pounds. ✔
- between 6.2 and 10.1 pounds.
- between 6.2 and 9.7 pounds.
- between 5.3 and 10.1 pounds.

99.7% of all full-term babies weigh

- between 3.7 and 10.5 pounds.
- between 3.7 and 10.8 pounds.
- between 4.2 and 10.5 pounds.
- between 4.2 and 10.8 pounds. ✔

Check | Show Answer

A distribution has a σ = 4. Find the z-score for a score that is:

(4/4 points)

4 points **above** the mean: z = ____

1

1

12 points **above** the mean: z = ____

3

3

2 points **below** the mean: z = ____

-0.5

-0.5

8 points **below** the mean: z = ____

Help

-2

-2

Help

Check Show Answer

With a height of 75 in., Lyndon Johnson was the tallest president in the 20th century. Presidents of the past century have heights with a mean of 71.5 in. and a standard deviation of 2.1 in.

With a height of 85 in., Shaquille O'Neal was the tallest player on the Miami Heat basketball team. Basketball players for the Miami Heat during 2004 to 2008 had heights with a mean of 80.0 in. and a standard deviation of 3.3 in.

(3/3 points)

What is the z-score for Lyndon Johnson? *(Round to 2 decimal places.)*

1.67

1.67

What is the z-score for Shaquille O'Neal? *(Round to 2 decimal places.)*

1.52

1.52

Help

Who is relatively taller among their respective groups: Lyndon Johnson or Shaquille O'Neal?

Lyndon Johnson

Check Show Answer

(1/1 point)

Which of the following is NOT a characteristic of a Normal Distribution?

- (selected) Half the data values are positive, and half the data values are negative. ✔
- The center value of the distribution is the mean, median and mode.
- The distribution is bell-shaped and unimodal.
- The scores are symmetrical around the mean.

Check Show Answer

Help

EdX offers interactive online classes and MOOCs from the world's best universities. Online courses from MITx, HarvardX, BerkeleyX, UTx and many other universities. Topics include biology, business, chemistry, computer science, economics, finance, electronics, engineering, food and nutrition, history, humanities, law, literature, math, medicine, music, philosophy, physics, science, statistics and more. EdX is a non-profit online initiative created by founding partners Harvard and MIT.

Terms of Service and Honor Code

Privacy Policy (Revised 4/16/2014)

About & Company Info

About

News

Contact

FAQ

edX Blog

Donate to edX

Jobs at edX

Follow Us

Twitter

Facebook

Meetup

LinkedIn

Google+