PS 211: Introduction to Experimental Design

Spring 2025 · Section C1

Lecture 6: The Normal Curve, Standardization, and Z Scores

Updates and reminders

- Homework 1 has been graded.
- We will post answers and grades tomorrow morning.
- We will not be returning individual homeworks, so if you would like to see or go over what you got wrong, please come to office hours.
- We will **not** be responding to individual questions about homework grades via Slack.
- For the most part, people did pretty well (~91 average).
- We graded very leniently, so even if you did well, please review the answers to make sure you understand everything.
- Please bring your questions to the review session on Tuesday.
- I posted a review sheet for Exam 1 on Slack and the course website.

Review: Variance

Why do we care about variance and standard deviations?

- Variance and standard deviations tell us how **spread out** scores are around the mean.
- They are used in many statistical tests (e.g., t-tests, ANOVA).
- They are also useful in the real world!

Review: Variance in the real world

Imagine you have the choice between two summer jobs: You can be a lifeguard or you can be a tour guide. Both jobs pay, on average, $15/hour. However, the lifeguard job has a standard deviation of $1/hour, while the tour guide job has a standard deviation of $10/hour. Which job would you choose? Why?

Imagine that you are a boating instructor and you need to order lifejackets for a group of 100 people. You know their average weight is 150 lbs. The options for lifejacket sizes range from XXS to XXL, which correspond to different weights. How would knowing the standard deviation of their weights help you decide how many of each lifejacket size to order?

Review: Variance in the real world

Now imagine that you need to estimate the standard deviation of the weights of the 100 people. You decide to randomly poll 10 of the 100 and ask their weights. How would you use the weights of the 10 people to estimate the standard deviation of the weights of all 100 people?

Answer: Calculate the standard deviation of the 10 weights (your sample) and use that as an estimate of the standard deviation of all 100 weights (your population).

Let's do it: Here are your sample weights (in lbs): $120, 130, 140, 150, 160, 170, 180, 190, 200, 210$

Step 1: Calculate the mean of the weights.

$$M = (120 + 130 + 140 + 150 + 160 + 170 + 180 + 190 + 200 + 210)/10 = 165$$

Step 2: Calculate the sum of squared deviations from the mean.

$$\sum (X_i - M)^2 = (120 - 165)^2 + (130 - 165)^2 + (140 - 165)^2 ... + (210 - 165)^2 = 8250$$

Review: Variance in the real world (continued)

Step 3: Calculate the *sample* variance.

$$Variance = \frac{\sum(X_i - M)^2}{n - 1} = \frac{8250}{10 - 1} = 916.67$$

Step 4: Calculate the *sample* standard deviation.

$$SD = \sqrt{Variance} = \sqrt{916.67} \approx 30.28$$

Now we have our *sample* standard deviation of 30.28 lbs, which we can use as an estimate of the *population* standard deviation of all 100 weights.

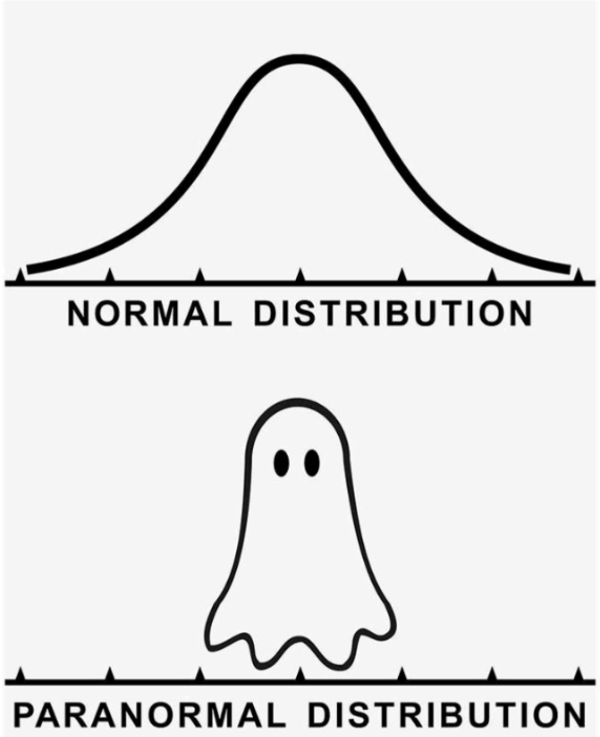
Variance is a very important concept in statistics!

It will come up again and again in this class and in future statistics classes.

Review: Normal Curves

- A **normal distribution** is:
  - Bell-shaped: Most scores cluster in the center
  - Symmetrical: Left side mirrors the right
  - Unimodal: Only one “hump”
- Ends of the curve are called **tails**

Today, we'll learn why normal distributions are so important in statistics.



Cheater Detection Using Normal Curves

- We often expect natural patterns to be normally distributed.
- Deviations from normality can indicate **cheating** or **manipulation**, or more generally that something is off.

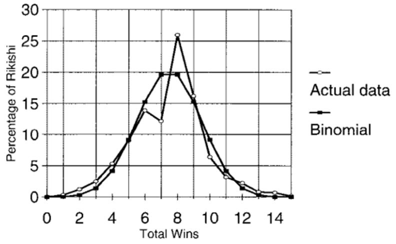


FIGURE 2. WINS IN A SUMO TOURNAMENT (ACTUAL VS. BINOMIAL)

Study of sumo wrestlers:

- 26% finished with 8 wins
- 12.2% finished with 7 wins
- Expected (if normal): ~19.6% for both outcomes

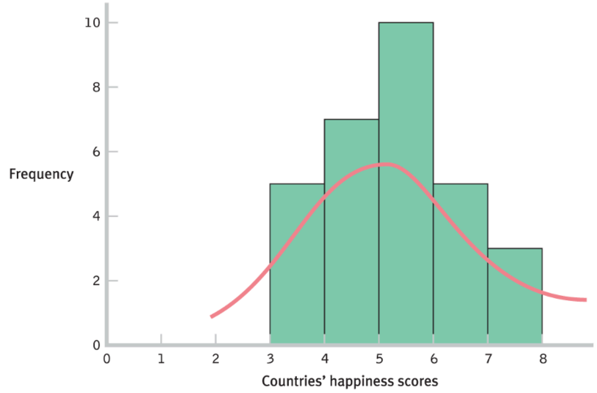
Why might the sumo results deviate from a normal distribution?

Answer: Sumo wrestlers were throwing matches to help wrestlers who had won 7 matches win their 8th (and have a winning season / advance rounds).

This is also called "anomaly detection" and is used in many fields, including fraud detection.

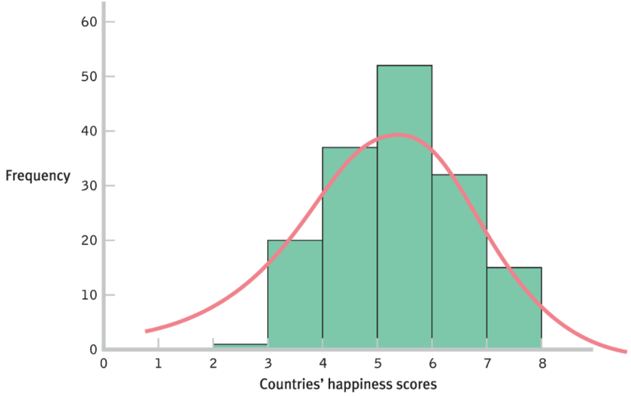
More Scores = More “Normal”

- As sample size increases, samples drawn from normally distributed populations look more normal.
- Larger samples better approximate the true population distribution.
- Small samples can look irregular.



Nolan/Heinzen, *Statistics for the Behavioral Sciences*, 5e, © 2020 Worth Publishers

Nolan/Heinzen, *Statistics for the Behavioral Sciences*, 5e, © 2020 Worth Publishers



Nolan/Heinzen, *Statistics for the Behavioral Sciences*, 5e, © 2020 Worth Publishers

Standardization: Comparing Apples and Oranges

- When data are normally distributed, we can compare scores across different distributions.
  - This can be very useful!
- We actually often do this implicitly, without realizing it or using math.
- Example:
  - If a movie critic typically gives 4-star reviews, and another typically gives 2-star reviews, who gave the better review if they both gave a movie 3 stars?
  - You might intuitively say the first critic, because 3 stars is above their average, while it's below the second critic's average.

But how much better?

- We can use math to quantify this!

Standardization: Comparing Apples and Oranges

- Standardization lets us compare scores across **different distributions**.
- We accomplish by converting raw scores into a common metric: “number of SDs from the mean.”
- This common metric is called a **z score**.
- We can then directly compare *z* scores from different distributions.

z score: How many standard deviations a score is from the mean of its distribution.

Review: Standard Deviation

- **Variance:** average squared deviation from the mean
- **Standard Deviation (SD):** square root of variance

$$SD = \sqrt{\frac{\sum(X_i - M)^2}{N}}$$

- SD tells us the **typical deviation** from the mean.

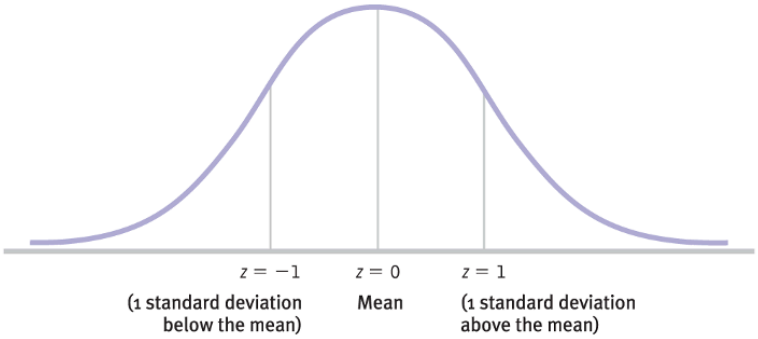
The Z Distribution

`z score: How many standard deviations a score is from the mean of its distribution.`

- A z distribution is a distribution of z scores.
- Properties:
  - Mean = 0
  - SD = 1

Why is the mean of a z distribution always 0?

Answer: The mean is **0 standard deviations from the mean**!



Nolan/Heinzen, *Statistics for the Behavioral Sciences*, 5e, © 2020 Worth Publishers

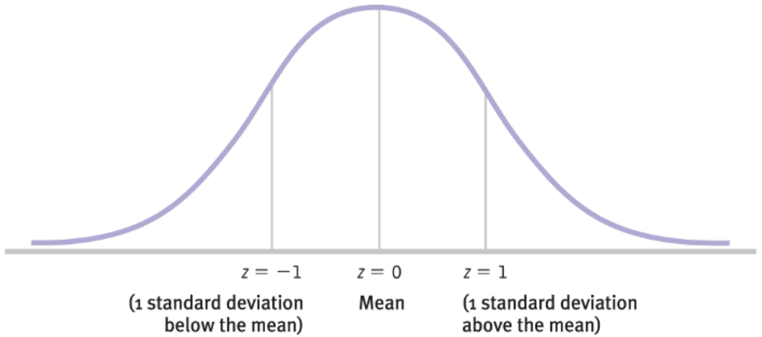
The Z Distribution (Continued)

`z score: How many standard deviations a score is from the mean of its distribution.`

- A z distribution is a distribution of z scores.
- Properties:
  - Mean = 0
  - SD = 1

Why is the SD of a z distribution always 1?

Answer: If a raw score is 1 SD above the mean, its z score is 1. If a raw score is 2 SDs above the mean, its z score is 2. And so on. Thus, the SD of the z distribution is always 1.



Nolan/Heinzen, *Statistics for the Behavioral Sciences*, 5e, © 2020 Worth Publishers

Understanding a score's relation to the mean of its distribution gives us important information.

How to Calculate a Z Score

You can sometimes calculate z scores without a calculator.

Example:

- Exam mean = 70
- SD = 10
- Your score = 80
- Your friend's score = 75
- Your enemy's score = 60

What is your z score? What is your friend's z score? What is your enemy's z score?

Answer: Your $z = 1$, friend's $z = 0.5$, enemy's $z = -1$

How to Calculate a Sample Z Score: Formula

Example: Computing a Z Score

- Mean happiness score of 157 countries = 5.382
- SD = 1.138
- Australia's score = 7.313

How many SDs above the mean is Australia's level of happiness?

$$z = \frac{7.313 - 5.382}{1.138} \approx 1.7$$

Interpretation:
Australia's happiness level is **1.7 SDs above the mean**.

Example: Computing a Z Score

- Mean happiness score of 157 countries = 5.382
- SD = 1.138
- Egypt's score = 4.362

$$z = \frac{4.362 - 5.382}{1.138} \approx -0.9$$

How many SDs below the mean is Egypt's level of happiness?

Interpretation:
Egypt's happiness level is **0.9 SDs below the mean**.

Reverse: Transforming Z Scores into Raw Scores

- You can also convert z scores back into raw scores.

Formula:

$$X = z \times SD + M$$

Example:

- France happiness: $z = 0.963$
- Mean happiness score of 157 countries = 5.382
- SD = 1.139

$$X = (0.963)(1.139) + 5.382 \approx 6.48$$

Why is this useful?

- Z scores allow comparison across different scales.

Example:

- Imagine we are comparing grades across two PS 211 classes, taught by different professors.
- Your grade: 92/100, class mean = 78.1, SD = 12.2
- Your friend's grade: 8.1/10, class mean = 6.8, SD = 0.74

Who did better, you or your friend?

- Your $z = \frac{92-78.1}{12.2} = 1.14$
- Friend's $z = \frac{8.1-6.8}{0.74} = 1.76$

Conclusion: Both of you received above average grades, but your friend did better *relative to their class*.

Comparing scores across different scales: Practice!

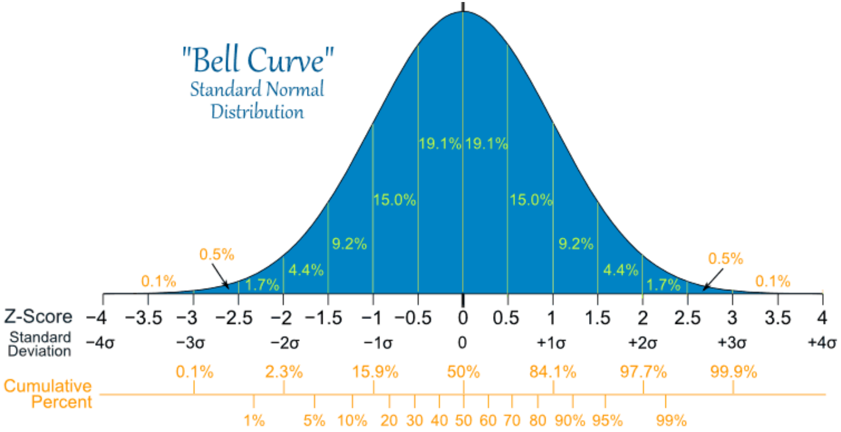
- Imagine you want to decide which of two new restaurants to go to.
- Restaurant A was reviewed by Critic A, who gave it a 7/10.
- Restaurant B was reviewed by Critic B, who gave it a 7/10.
- You have the smart idea to take a *sample* of each critic's past restaurant ratings to see how harsh or lenient they are.
- Critic A's ratings: $6, 6, 7, 8, 4$
- Critic B's ratings: $4, 5, 5, 9, 8$

Which restaurant should you go to, based on the critics' reviews? Explain your reasoning!

Comparing scores across different scales: Practice (Continued)

Z Scores, Normal Distributions, and Percentiles

- z scores allow us to determine **percentiles**.
- Percentile: percentage of scores below a given score.
- Normal distributions are standard, so we know the percentage of scores below any z score.



- 100% of scores fall below $z = +\infty$; 0% fall below $z = -\infty$.
- 50% of scores fall below $z = 0$ (the mean).
- 84% of scores fall below $z = 1$ (1 SD above the mean).
- 68% of scores fall between $z = -1$ and $z = +1$ (within 1 SD of the mean).
- **Can use a z table to find exact percentiles for any z score.**

The Central Limit Theorem

- The theorem: Any **distribution of sample means** will be approximately normal if the sample size is sufficiently large.

What is a "distribution of sample means"?

Answer: A "distribution of sample means" is the distribution of the means of multiple samples taken from a population. It shows how the sample means vary and allows us to make inferences about the population mean.

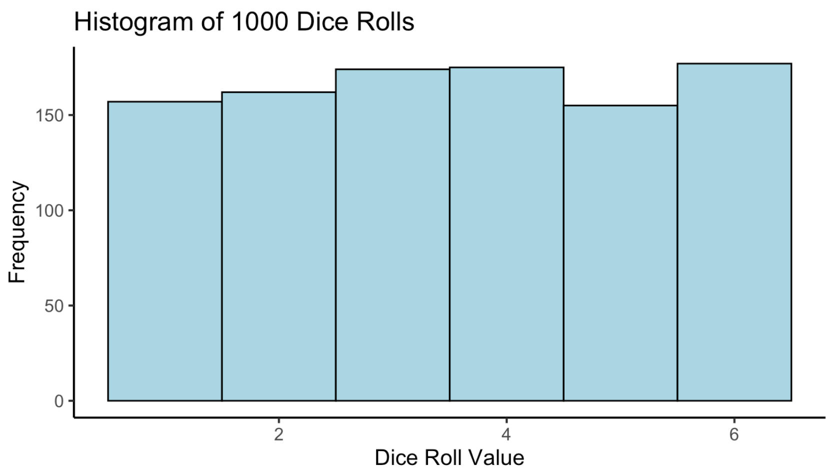
Example: We want to estimate the average height of all PS 211 students. Each class, we measure the height of 5 randomly selected students and calculate the mean height of those 5 students. We repeat this process many times, each time selecting a new random sample of 5 students and calculating the mean height. The distribution of these sample means will be approximately normal, even if the original distribution of individual heights is not normal.

Central Limit Theorem: Demo

Dice Rolls

- Imagine rolling a 6-sided die.
- Roll it once, record the result.
- Repeat many times, plot the distribution of results.

What does this distribution look like?



Answer: The distribution is uniform, with each outcome (1-6) equally likely.

Central Limit Theorem: Demo (Continued)

Distribution of Scores vs. Means

- Even if scores in a population aren't normally distributed, the distribution of sample means will be approximately normal if the sample size is large enough.
- This is the essence of the Central Limit Theorem.
- A distribution of means is **less variable** than a distribution of raw scores.
- This means it is less spread out.

Why is the distribution of means less variable than the distribution of raw scores?

Answer: The distribution of means is less variable because averaging reduces the impact of extreme values. When we take the mean of a sample, we are essentially smoothing out the variability that exists in individual scores.

Standard Error of the Mean

- The standard error (SE) is the name for the standard deviation of a distribution of sample means.

- The formula for the standard error is:

$$SE = \frac{s}{\sqrt{n}}$$

where s is the sample standard deviation and n is the sample size.

Where does this formula come from?

Answer: The formula for the standard error comes from the fact that the variability of sample means is related to the variability of individual scores and the sample size. As we increase the sample size, the standard error decreases, reflecting the increased precision of our estimate of the population mean.

We can derive it mathematically, but that is beyond the scope of this class.

Standard Error of the Mean

Calculating Standard Error

- If the SD of a distribution of individual scores = 5
- If we take samples of size $n = 25$, then the SE of the distribution of sample means is:

$$SE = \frac{5}{\sqrt{25}} = \frac{5}{5} = 1$$

What if we take samples of size $n = 9$? What if we take samples of size $n = 100$? Which will have a smaller SE? Why?

Answer:

- For $n = 9$:

$$SE = \frac{5}{\sqrt{9}} = \frac{5}{3} \approx 1.67$$

- For $n = 100$:

$$SE = \frac{5}{\sqrt{100}} = \frac{5}{10} = 0.5$$

The sample size of $n = 100$ will have a smaller SE because the standard error decreases as the sample size increases. This reflects the increased precision of our estimate of the population mean with larger samples.

Practice: Putting concepts together

Practice: Putting concepts together (Continued)

That's all for today!

See you Tuesday for the Exam 1 review session!