The Normal Distribution

- Properties of the Normal Distribution
- Shapes of Normal Distributions
- Standard (Z) Scores
- The Standard Normal Distribution
- Calculate areas within the normal distribution

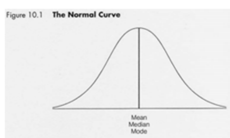
<u>Review</u>:
So far we have been examining a variety of distributions

- Frequency distributions
  - (negatively and positively skewed)
- Central tendencies
  - (mean, median, mode)
- Variance and Standard deviation

<u>Coming Up</u>:
The next type of distribution is referred to as the normal curve.

Properties of the Normal Distributions

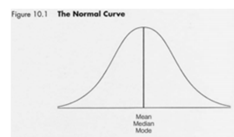
- Used with **linear** variables
- A bell-shaped and symmetrical theoretical (i.e., perfect) distribution as compared to an empirical (actual, real) distribution
- with the mean, the median, and the mode all coinciding at its peak and
- with frequencies gradually decreasing at both ends of the curve.



What do we mean when we say the normal distribution is a “theoretical” distribution and not an empirical distribution? What is an empirical distribution? A theoretical distribution?

Normal Distributions

- **Normal Distribution**
- is a theoretical ideal distribution. Real-life empirical distributions rarely, if ever, match this model perfectly.
- However, many things in life do approximate the normal distribution, and are said to be “normally distributed”
  - For example, grades, SAT scores, **others?**



The shrinking middle class may get most of the news, but there’s another middle out there that’s also in dire straits: middle-skill jobs.

By Mark Fadden
Special to the Star-Telegram

Remember the bell curve? That big, beautiful curved line on a graph that told us that, in terms of measuring human performance, there was a low number of low performers, a high number of average performers, and a low number of high performers.

For all intents and purposes, the bell curve seems about right, doesn't it? If we remember back to our school days, a few kids in every class were tapped for the “Gifted and Talented” program and a few others were held back at the end of the year and needed to, as my mother used to say, be baked in the oven a little bit longer to come out perfect. But the majority of us glided right along, just like we should of, with passing grades of varying degrees right on schedule until it was time to claim our diplomas.

The bell curve is so dependable, it's been applied to almost everything in terms of human performance to predict the outcome: sports, music, even happiness. A few will soar with wings like eagles, others will fail and drop like stones, but most of us will be able to play a few sports decently, be good enough to at least sing in the shower or in the church choir, and be happy with our lives most of the time.

In the last decade or so, there's been a worrisome trend going on that the bell curve has a hard time explaining. It's the shrinking middle class. Whereas the bell curve used to go from fairly low, to very high to fairly low again to form a classic bell shape, not only is the middle portion getting slimmer, it's also getting short-

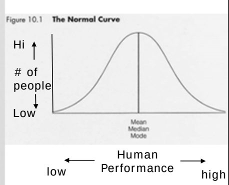
The upper class puts a significant portion of their income in investments. Stocks and real estate portfolios don't usually include iPhones and Beats headphones.. At these two ends of the spectrum the money just isn't being spent like it is in the middle. So, wherever you are on the economic spectrum, a healthy middle class bodes well for us all. And one of the biggest parts of the waning middle class has to do with the growing gap in middle-skill jobs.

Where's the beef?

Middle-skill jobs can be defined as those that require more than a high school diploma but less than a four-year college degree. Think firefighters, advanced manufacturing workers, computer support specialists, and radiology technicians. Currently in the US, approximately 70 million people work in middle-skills jobs, representing almost half of the entire labor force.

The gap between the number of qualified middle-skill workers available and middle-skills jobs that need to be filled has been caused by technology and innovation. The skills learned in high school simply aren't good enough to operate the more technology advanced equipment or understand the math and technology-centric operations in today's middle-skill jobs.

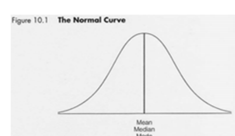
In a 2014 survey, the management consulting firm Accenture found that 69 percent of the approximately 800 human resources executives who were polled said that middle-skill talent shortages “regularly affect their performance.”



Normal Distributions

One reason we study the normal distribution is because characteristics (i.e., properties) of the normal distribution can be applied to empirical distributions (i.e. real-life variables) that are approximately normally distributed.

In this chapter, we will learn various characteristics (i.e. properties) of the normal distribution that can help us when we are examining real data that is approximately normal.



Scores "Normally Distributed?"

Table 10.1 Final Grades in Social Statistics of 1,200 Students (1983-1993)					
Midpoint Score	Frequency Bar Chart	Freq.	Cum. Freq. (below)	%	Cum % (below)
40	*	4	4	0.33	0.33
50	*******	78	82	6.5	6.83
60	***************	275	357	22.92	29.75
70	**********************	483	840	40.25	70
80	***************	274	1114	22.83	92.83
90	*******	81	1195	6.75	99.58
100	*	5	1200	0.42	100

- Is this distribution normal?
- To answer this question there are two things to initially examine: (1) look at the shape illustrated by the bar chart, and (2) calculate the mean, median, and mode.

Scores "Normally Distributed?"

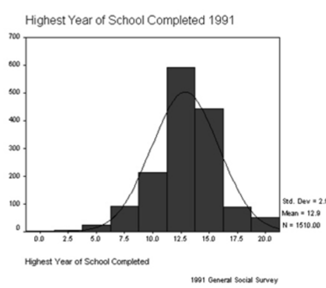
Table 10.1 Final Grades in Social Statistics of 1,200 Students (1983-1993)					
Midpoint Score	Frequency Bar Chart	Freq.	Cum. Freq. (below)	%	Cum % (below)
40	*	4	4	0.33	0.33
50	*******	78	82	6.5	6.83
60	***************	275	357	22.92	29.75
70	**********************	483	840	40.25	70
80	***************	274	1114	22.83	92.83
90	*******	81	1195	6.75	99.58
100	*	5	1200	0.42	100

- The **Mean** = 70.07
- The **Median** = 70
- The **Mode** = 70

Scores Normally Distributed!

- The **Mean** = 70.07
- The **Median** = 70
- The **Mode** = 70
- Since all three are essentially equal, and the bar graph appears to be normally distributed, we can conclude these data are ***normally distributed***.
- Also, since the median is approximately equal to the mean, we know that the distribution is **symmetrical** (equal on both sides of the mid point, that is, mirror images on each half).

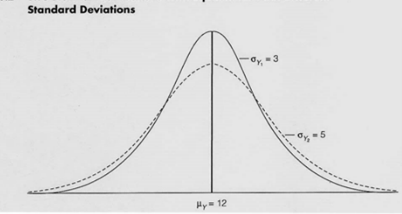
The Shape of a Normal Distribution



Notice the shape of the normal curve in this graph. Some normal distributions are tall and thin, while others are short and wide. All normal distributions, though, are taller in the middle and symmetrical (What do we mean by symmetrical?).

Different Shapes of the Normal Distribution

Figure 10.2 **Two Normal Distributions with Equal Means but Different Standard Deviations**



Notice that the **standard deviation** determines the relative width of the distribution; the larger the standard deviation, the wider the curve (because the larger the SD, the further the cases are, on average, from the mean.

Since the Standard Deviation reflects the width of the curve, lets review what the standard deviation is:

A measure of the population that reflects the average distance of the cases (or average deviation of the cases) from the mean.

A measure of variation for interval-ratio variables; it is equal to:

$$s = \sqrt{s_Y^2} = \sqrt{\frac{\sum (Y - \overline{Y})^2}{N - 1}}$$

We can use the **normal curve** and **standard deviation** to determine (or estimate) the percentage of cases that are close to the mean as well as the percentage that are far from the mean. Consider our example of a mean grade of 70.00 and SD of 10.00.

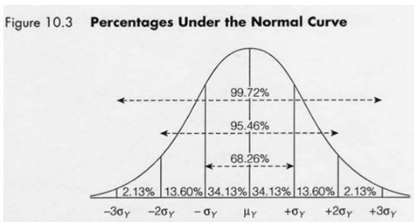


Figure 10.3 **Percentages Under the Normal Curve**

The **normal curve** allows us to estimate what percentage of the cases fall between any two points.

For example, what percentage of students scored between 70 and 80 on the statistics tests (mean=70; SD=10).
Between 70 and 60? Between 60 and 80?
Between 70 and 90? Between 70 and 50?

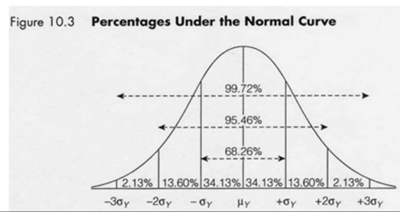


Figure 10.3 **Percentages Under the Normal Curve**

The **normal curve** also allows us to predict what percentage of the cases fall above or below a specific value or "raw score" (or in this case grade).

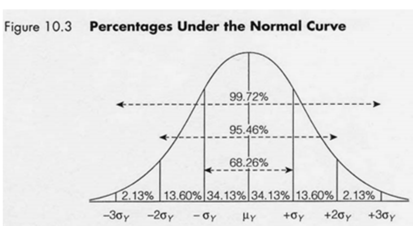


Figure 10.3 **Percentages Under the Normal Curve**

For example:
--what percentage of the students fell above a grade of 70 (mean=70; SD=10)?
--what percentage fell above a grade of 80?
--what percentage fell above a grade of 90?
--what percentage fell below a grade of 60?

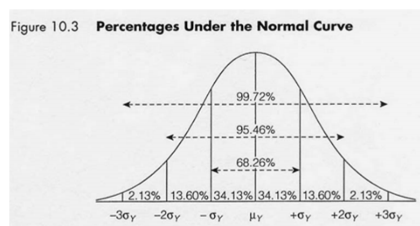


Figure 10.3 **Percentages Under the Normal Curve**

In-Class Assignment:

Finding a grade where a proportion of students fall above or below the grade.

Using the data presented in Table 10.1:

$$\sqrt{\frac{\sum(Y-\bar{Y})^2}{N-1}}$$

(1) what is the mean?
(2) the standard deviation?
(3) roughly 84% of the students received a grade of ______ or lower?
(4) roughly 16% of the students scored ______ or lower?
(5) 95% of the students scored ______ or higher?

Standard (Z) Score

The number of standard deviations that a given score is above or below the mean.

In our example, if "Jill" had a Z score (standard deviation) of 1, where would she fall on the normal cure?

What percentage of the students scored less than Jill?

If Mary had a Z score of -3 where would she fall on the normal curve and what percentage of students scored less than her?

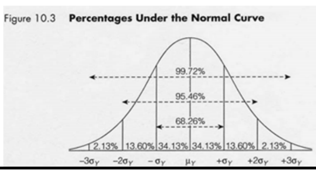
Standard (Z) Scores

What is the advantage of knowing the Z score for a particular case?

You will know whether they are close to the mean or far above or below the mean.

And

You will know the percentage of cases that fall higher and lower than a particular case.



Calculating the Z score: we can take a particular student grade, say 85, and subtract from it the mean for all students and then divide this number by the SD. This would provide us with the "standard score" for this particular grade.

$$Z = \frac{Y - \overline{Y}}{S_y}$$

Example where a person has received a performance score from her/his boss: How do we calculate a Z score for this person?

Standard Score = $\frac{\text{performance score – mean}}{\text{standard deviation}}$

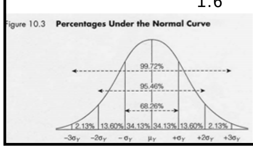
Standard (Z) Score

$$Z = \frac{Y - \overline{Y}}{S_y}$$

Example from previous slide. We'll pretend we've calculated the mean for all students and found it to be 80 and we've calculated the SD for all students and found it to be 5. We want to determine the Z score for a student who received an 88.

Z (Standard Score) = $\frac{\text{grade – mean}}{\text{standard deviation}}$

1.6 = $\frac{88 - 80}{5}$



Steps for determining how unusual a raw score is:

1. Convert the raw score to a Z score
2. Find the Z score on a normal curve displaying Z scores

or

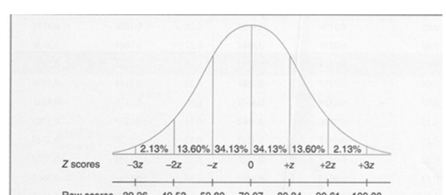
we can use the "Standard Normal Table" (page 480-483) to determine where the Z score falls on the normal curve. It displays the percentage of cases that fall between the mean and the Z score of interest (column B on table). You add 50% to this (all the cases on the other side of the mean) to determine the percent of cases that are below or above the Z score of interest.

For example:

a Z score of 1.96 = a raw score of .50 + .475

or .975

Standard Normal Curve



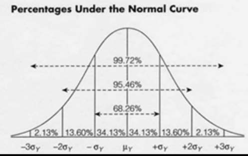
How is the "standard" normal curve different from the normal curve?

The normal curve is in raw scores while the standard normal curve shows standard scores.

In Class Problem:

Laura hopes to obtain a scholarship from TWU's Physical Therapy doctorate program. She scored a 92 on the physical therapy entry exam. What percentage of the applicants scored lower than Laura? Among those who took the exam, the mean was 78 and the standard deviation was 4. Laura needs her score to be better than 98% of the applicants in order to obtain a scholarship. Will she be offered a scholarship?

Figure 10.3 Percentages Under the Normal Curve



$$\text{Standard Score} = \frac{\text{grade – mean}}{\text{standard deviation}}$$

$$Z = \frac{Y - \overline{Y}}{S_y}$$

Using the Computer to Compute Z Scores:

1. Open a data set
2. Click Analyze, then descriptive statistics, then descriptives
3. Double click on a variable (grade point average) to move it to the variables box.
4. Click Save standardized values as variables in the Descriptives dialog box
5. Click OK.
6. Examine your data file. Notice that a new variable has been created. Each student has received a z score for the variable. A student with a large z score will have a score further from the mean score.

凩訝

(see you later)