STAT 201 Chapter 6

Distribution

Random Variable

- We know variable
- **Random Variable**: a numerical measurement of the outcome of a random phenomena

- Capital letter refer to the random variable
- Lower case letters refer to specific realization

Random Variable: Example

- We roll a 6 sided dice once
- **X**: The number we get from one roll. X is random variable because it is random, we don't know the result before rolling
- x=5, is a realization – a concrete observation

Discrete Random Variable

- The possible outcomes must be countable
- We have a valid discrete probability distribution if
  - All the probabilities are valid
    - $0 \leq P(x) \leq 1 \quad for\ all\ x$
  - We've accounted for all possible outcomes
    - $\sum P(x) = 1$

Discrete Distribution

- **Probability Distribution**: a summary of all possible outcomes of a random phenomena along with their probabilities

- A way to understand randomness

Discrete Distribution: Example 1

- Suppose X is the number we get from rolling dice once

X = Number of rolling	P(X) = Probability
1	1/6
2	1/6
3	1/6
4	1/6
5	1/6
6	1/6

- All probabilities are between 0 and 1
- $\sum P(x) = \frac{1}{6} + \frac{1}{6} + \frac{1}{6} + \frac{1}{6} + \frac{1}{6} + \frac{1}{6} = 1$

Discrete Distribution: Example 2

- Suppose X is the number of red lights on your way to campus and assume there are only 3 traffic lights

X = Number of lights	P(X) = Probability
0	.10
1	.10
2	.10
3	.40

- All probabilities are between 0 and 1; but summation of probabilities isn't 1. <u>Not a valid discrete distribution!</u>

Discrete Distribution: Example 2

- Suppose X is the number of red lights on your way to campus and assume there are only 3 traffic lights

X = Number of lights	P(X) = Probability
0	.10
1	.20
2	.30
3	.40

- All probabilities are between 0 and 1; summation of probabilities is 1. <u>A valid discrete distribution!</u>

The mean of a discrete distribution

- The **Expectation (Mean)** of a probability distribution represents the average of a large number of observed values. [in the long run].

- We denote this with the Greek letter as below
  - $\mu_X = E(X) = Expected\ value\ of\ X = \sum[X * P(X)]$

Discrete Distribution: Example

- Suppose X is the number of red lights on your way to campus and assume there are only 3 traffic lights

X = Number of lights	P(X) = Probability	X*P(X)
0	0.1	0*0.1=0
1	0.2	1*0.2=0.2
2	0.3	2*0.3=0.6
3	0.4	3*0.4=1.2

- $\mu_X = E(X) = \sum X * P(X) = 0 + 0.2 + 0.6 + 1.2 = 2$
- On average, we expect to see 2 red lights on your way to campus

Discrete Distribution: Example

- What if we change the distribution?

X = Number of lights	P(X) = Probability	X*P(X)
0	0.4	0*0.4=0
1	0.3	1*0.3=0.3
2	0.2	2*0.2=0.4
3	0.1	3*0.1=0.3

- $\mu_X = E(X) = \sum X * P(X) = 0 + 0.3 + 0.4 + 0.3 = 1$
- On average, we expect to see 1 red lights on your way to campus

Discrete Distribution: Example

- What if we change the distribution again?

X = Number of lights	P(x) = Probability	x*P(x)
0	0.2	0*0.2=0
1	0.2	1*0.2=0.2
2	0.2	2*0.2=0.4
3	0.2	3*0.2=0.6

- Is it a valid distribution? No!
- What’s its expectation? It has no expectation!

Let’s Apply This to Categorical Variables: Binomial

- We look at a categorical variable with **two categories**
  - Success v.s. Failure, or 1 v.s. 0

X		P(X)
Success (denoted as 1)	This is what we’re interested in, even if it isn’t particularly successful in the sense of the English word	p = Probability of a ‘success’
Failure (denoted as 0)	This is the other case – what we aren’t interested in ,even if it isn’t particularly a failure in the sense of the English word	q = Probability of a ‘failure’ = 1- p

The Binomial Experiment: 3 restrictions

- It consists of n trials with binary (two) output
  (They are denoted 1 or 0, or success and failure)
- The probability of success on each trial is the same
  (The trials are **identical)**
- The outcome of one trial does not affect the outcome of another trial
  (The trials are **independent)**

The Binomial Distribution: Notation

- n = the number of trials
- p = the probability of success for any given trial
- q = the probability of failure for any given trial
    - q=1-p
- X = {the number of successes in n trials}
- X is the random variable, n and p are parameters

Binomial Distribution: Formula

- $P(x) = {}_nC_x\, p^x q^{n-x} = \frac{n!}{x!(n-x)!} p^x q^{n-x}$

- Recall: n! = n*(n-1)*(n-2)*…*2*1
  - Examples
    - 5! = 5*4*3*2*1=120
    - 0!=1
    - 5!/3!= 5*4

Binomial Distribution: Example 1

- Two friends, say Tom and Jerry, find a bottle of apple juice. Both of them want to have it.
- They decide to make the decision based on a best two-out-of three coin toss.

- Let’s say Tom chooses heads.

Binomial Distribution: Example 1

- A fair copper coin is flipped three times
  - Let X bet the number of heads that occur
  - $n = 3$
  - $p = .50$
  - $q = 1 - p = 1 - .50 = .50$
- Trials are identical since we flip the same coin
- Trials are independent as the outcome of one trial doesn’t affect another

Binomial Distribution: Example 1

- A fair copper coin is flipped three times
  - Let X bet the number of heads that occur
  - n = 3, p = .50, q =.50
  - Find the probability that there are exactly 2 heads

$$\mathrm{P}(\mathrm{x}=2) = P(x) = \frac{n!}{x!\,(n-x)!}p^x q^{n-x} = \frac{3!}{2!\,(3-2)!}.5^2\,.5^{3-2} = \frac{3!}{2!*1!}.\,5^2.\,5^1$$

$$= \frac{3*2*1}{(2*1)*(1)}(.25)(.5)$$

$$= .375$$

Binomial Distribution: Example 1

- A fair copper coin is flipped three times
  - Let X bet the number of heads that occur
  - n = 3, p = .50, q =.50
  - Find the probability that there are exactly 3 heads

$$\mathrm{P}(\mathrm{x}=3)=P(x)=\frac{n!}{x!\,(n-x)!}p^{x}q^{n-x}=\frac{3!}{3!\,(3-3)!}.5^{3}\,.5^{3-3}=\frac{3!}{3!*0!}.\,5^{3}.\,5^{0}$$

$$=\frac{3*2*1}{3*2*1}(.125)(1)$$

$$=.125$$

Binomial Distribution: Example 1

- A fair copper coin is flipped three times
  - Let X bet the number of heads that occur
  - n = 3, p = .50, q =.50
  - Find the probability that Tom wins (there are 2 or more heads)

$$\mathrm{P}(\mathrm{x} \geq 2) = P(x = 2) + P(x = 3) = .375 + .125 = .5$$

OR

$$\mathrm{P}(\mathrm{x} \geq 2) = 1 - P(x < 2) = 1 - (P(x = 1) + P(x = 0))$$

Binomial Distribution: Example 1

- A fair copper coin is flipped three times
  - Let X bet the number of heads that occur
  - n = 3, p = .50, q =.50
  - Find the probability that Tom loses(there are less than 2 heads)

$\mathrm{P}(\mathrm{x} < 2) = 1 - \mathrm{P}(\mathrm{x} \geq 2) = 1 - \left(P(x = 2) + P(x = 3)\right) = 1 - .5 = .5$

OR

$P(x < 2) = P(x = 1) + P(x = 0)$

Binomial Distribution: Example 1

- The probability that Tom wins is 0.5;
- The probability that Tom loses is 0.5;
- We see that this is a fair game – each of them has a 50% chance of winning

- <u>Why not just flip the coin once? (Hint: Variability)</u>

Binomial Experiment - Example 2

- A survey tells us that the probability that a USC student think himself or herself is attractive is 0.8; Consider a class of 48 students.
- $n = 48$, $p = 0.8$, $q = 1 – 0.8 = 0.2$
- Trials are independent – one student’s decision doesn’t affect the others
- Let’s assume students are identical!

Binomial Experiment - Example 2

- n = 48, p = 0.8, q = 1 – 0.8 = 0.2
- The probability that exactly half of the 48 students think they are attractive is
- $\mathrm{P}(\mathrm{x} = 24) = P(x) = \frac{n!}{x!(n-x)!} p^x q^{n-x} = \frac{48!}{24!(48-24)!} .8^{24} .2^{48-24} = \frac{48!}{24!24!} .8^{24} .2^{24} = .00000255$
- This is an **almost impossible** event – we expect the probability that half of the class to think they are attractive is almost 0%.

Binomial Experiment - Example 2

- n = 48, p = 0.8, q = 1 – 0.8 = 0.2
- The probability that at least half of the 48 students think they are attractive is
  $\mathrm{P}(\mathrm{x} \geq 24) = P(x = 24) + P(x = 25) + \ldots P(x = 48)$
  $= .99999922$
- This is an **almost certain** event – we expect the probability at least half of the class to think they are attractive is 99.99%.

Binomial Experiment - Example 2

- n = 48, p = 0.8, q = 1 – 0.8 = 0.2
- The probability that at least 1 student think they are attractive is

$$\mathrm{P}(\mathrm{x} \geq 1) = P(x = 1) + P(x = 2) + \ldots P(x = 48)$$
$$= .99999999999999$$

- This is an **almost certain** event – we expect the probability at least half of the class to think they are attractive is 99.99%.

Mean and Variance For A Binomial

- $\boldsymbol{Expectation\ (Mean)}\ =\ n * p$
- $\boldsymbol{Standard\ Deviation} = \sqrt{npq}$

Binomial Experiment - Example

- A fair coin is flipped four times
  - Let X be the number of heads that occur
  - n = 4
  - p = .50
  - q = 1 - p = 1 - .50 =.50
- $\boldsymbol{Mean} = n * p = 4 * .50$ =2
- So, on average we expect 2 heads among all total 4 flips

Binomial Experiment - Example

- A fair coin is flipped four times
  - Let X be the number of heads that occur
  - n = 4
  - p = .50
  - q = 1 - p = 1 - .50 =.50
- $\boldsymbol{Standard\ Deviation} = \sqrt{4 * .50 * .50} = 1$

Binomial Experiment – Example 2

- Considering a class of 48 students in which the probability that they think themselves are attractive is 0.8
- n = 48, p = 0.8, q = 1 – p = 0.2
- $\boldsymbol{Mean} = n * p = 48 * 0.8 = 38.4$
- So, on average, we expect 38 to 39 students in the class who think themselves are attractive

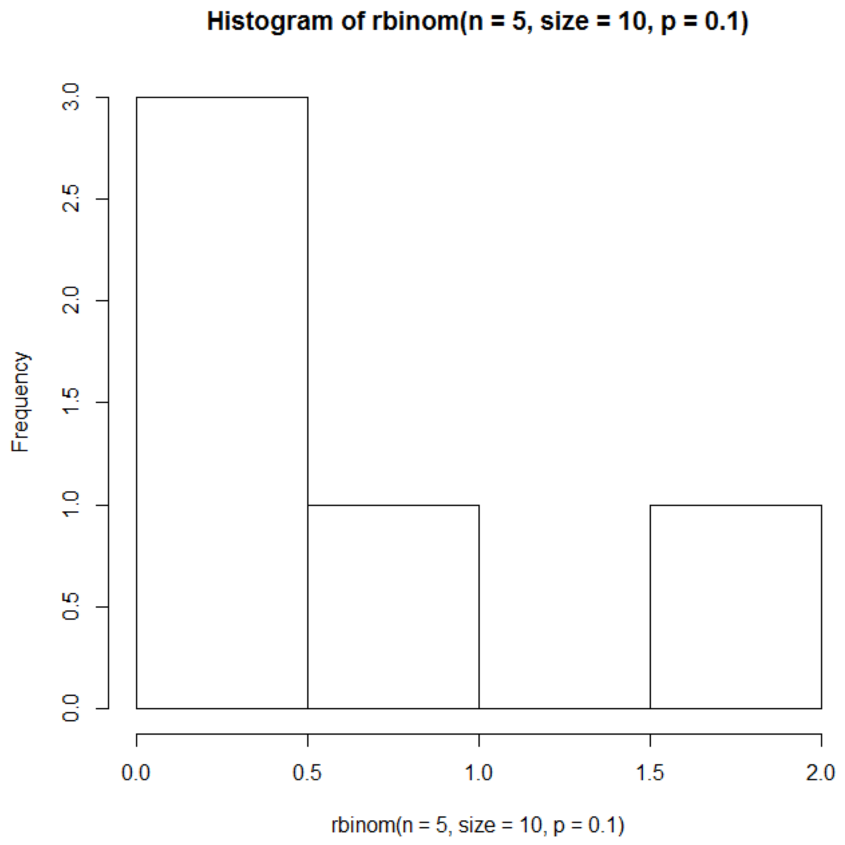
Binomial Experiment – Example 2

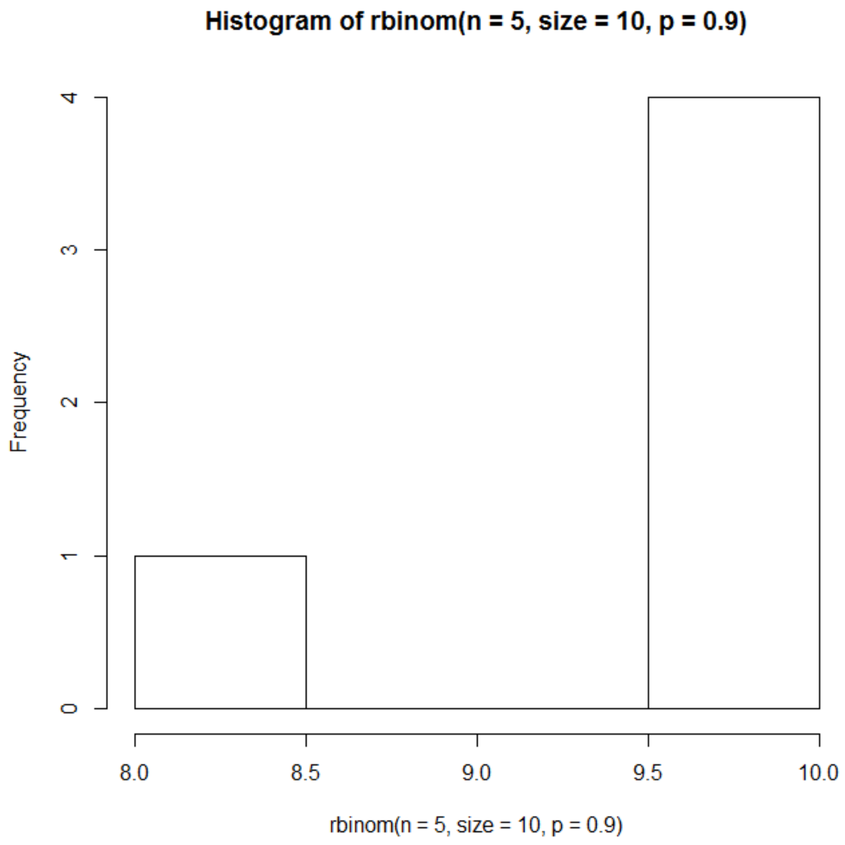
- Considering a class of 48 students in which the probability that they think themselves are attractive is 0.8
- n = 48, p = 0.8, q = 1 – p = 0.2
- $\boldsymbol{Standard\ Deviation} = \sqrt{48 * 0.8 * 0.2}$
  $= 2.77$

Shape of a binomial

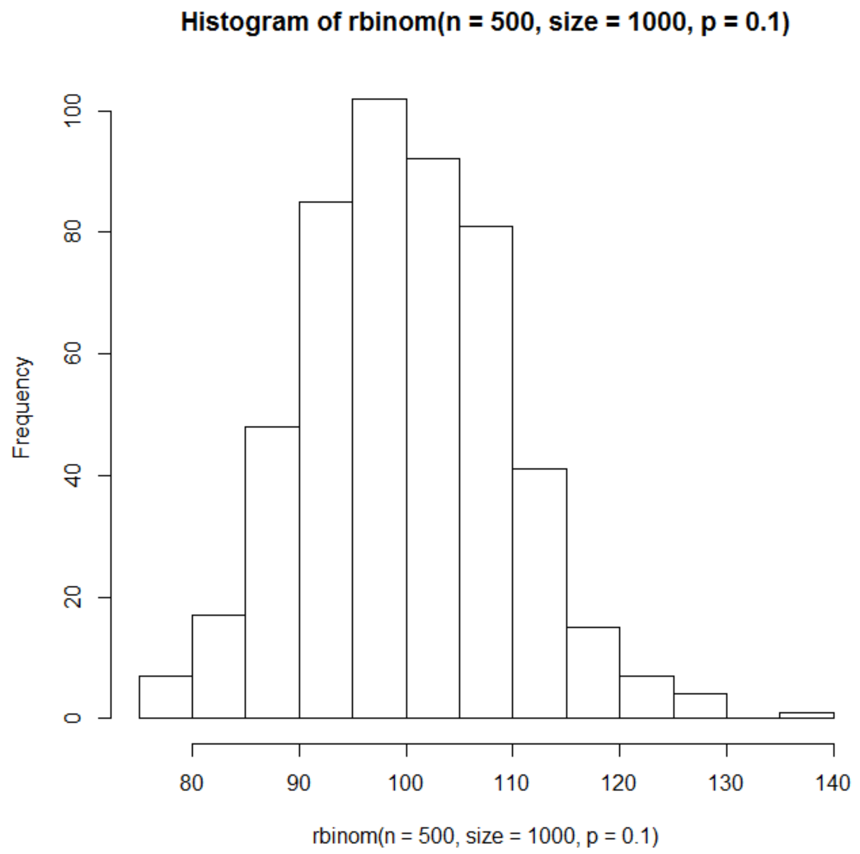
- In small sample size and $p < 0.5$, the distribution is right skewed;
- In small sample size and $p > 0.5$, the distribution is left skewed;
- For any p, as the sample size increases to be large enough, the probability distribution becomes bell shaped
- Large enough? $n * p \geq 15 \ and \ \mathrm{n} * (1 - p) \geq 15$

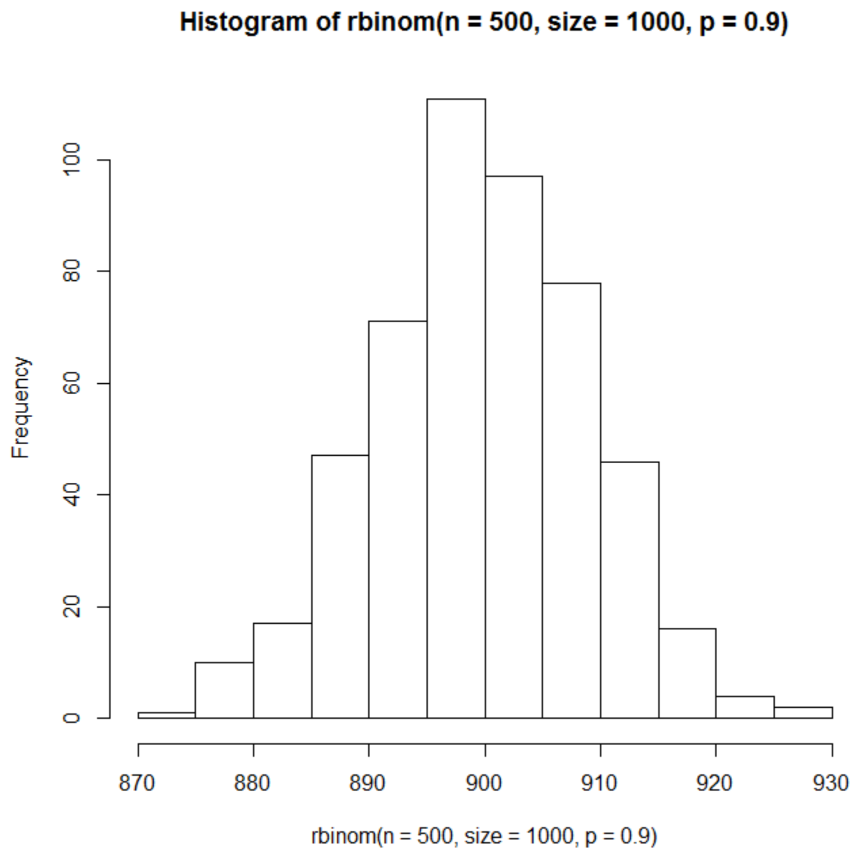
Shape of A Binomial – Small Size





Shape of A Binomial – Large Size

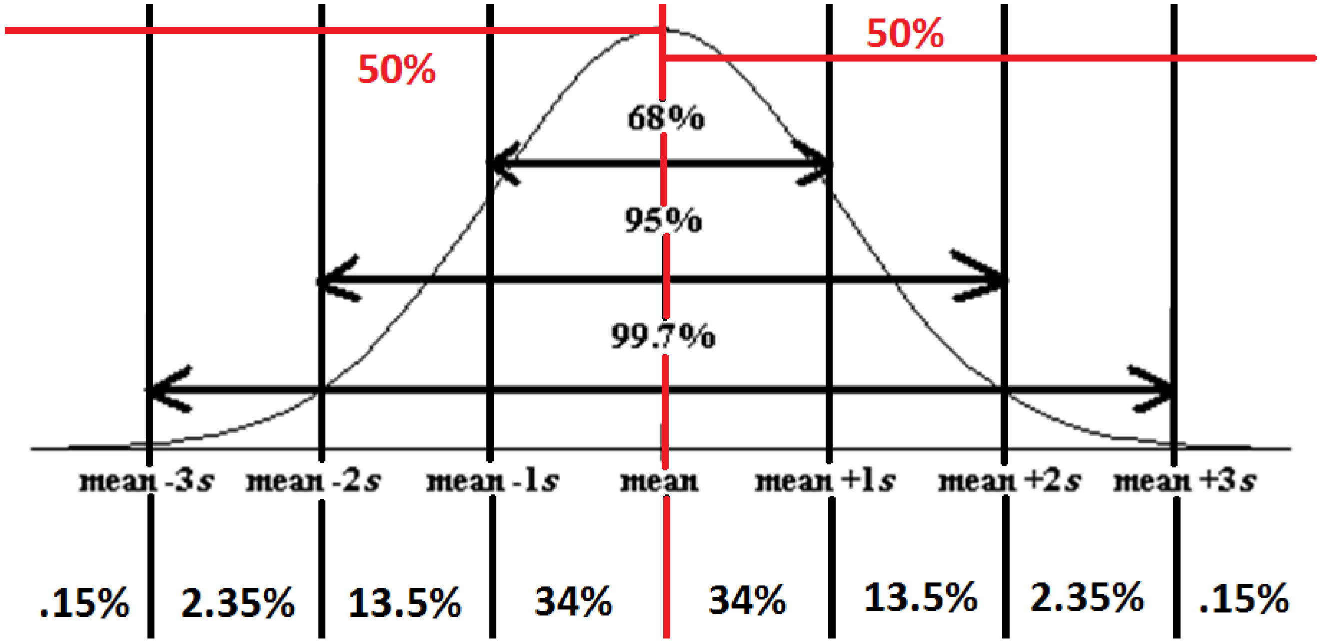




Continuous Variables

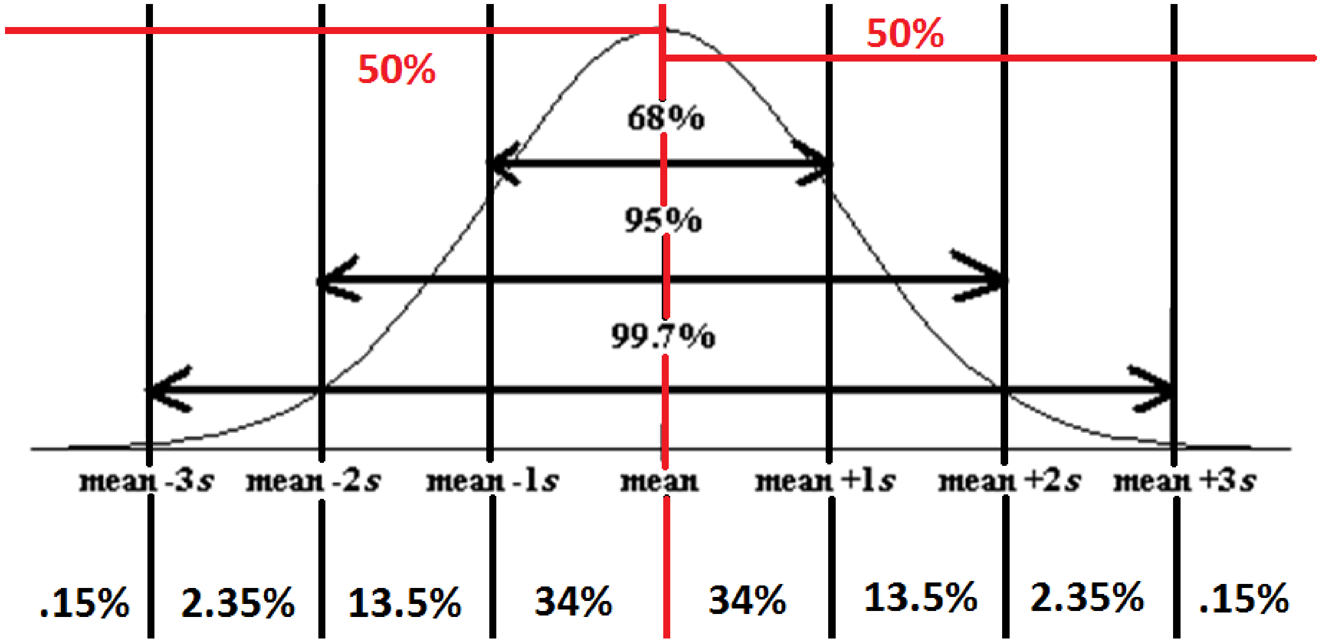
- The possible values for a continuous random variable form an interval
- There are infinitely many numbers in any interval, so the probability at any point is 0, and we look at the probability of intervals
- Each interval has a probability between 0 and 1
- The interval containing all possible values has probability equal to 1

Remember This?



- The total probability is 1 (100%)
- We'll use Z-scores to find the probability of other intervals not covered by the Empirical Rule

Remember This?



- The area under the graph of a density function over an interval represents the probability of observing a value of the random variable in the interval

Continuous Distributions

- A continuous probability distribution has some sort of curve associated with it

- Normal distribution: symmetric, bell shaped
- We use Z-scores, the number of standard deviations from the mean to look up the probabilities

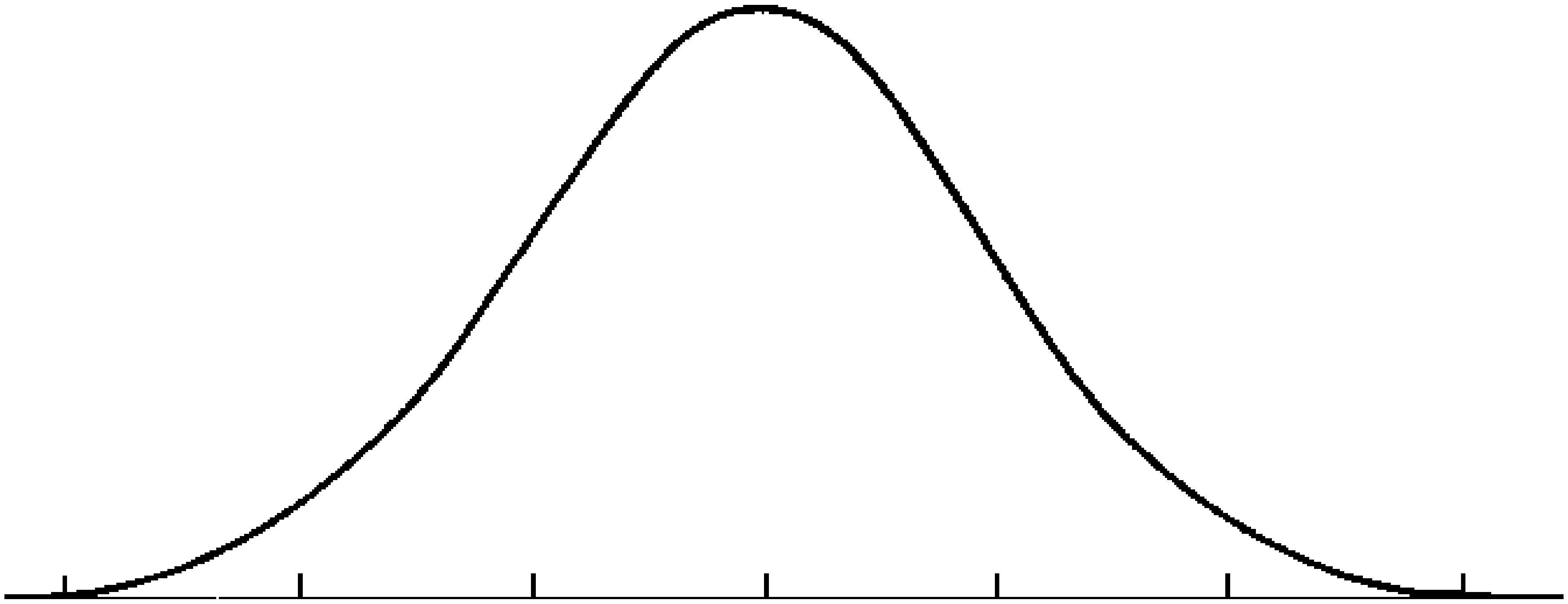
Vocabulary

- A **model** is an equation, table, or graph used to describe the reality

- The **normal curve** is a model used to describe a continuous random variable that is said to be normally distributed (symmetric about the mean, bell shaped)

Normal Curve

- The normal curve is Bell shaped, symmetric about the mean, follows the empirical rule

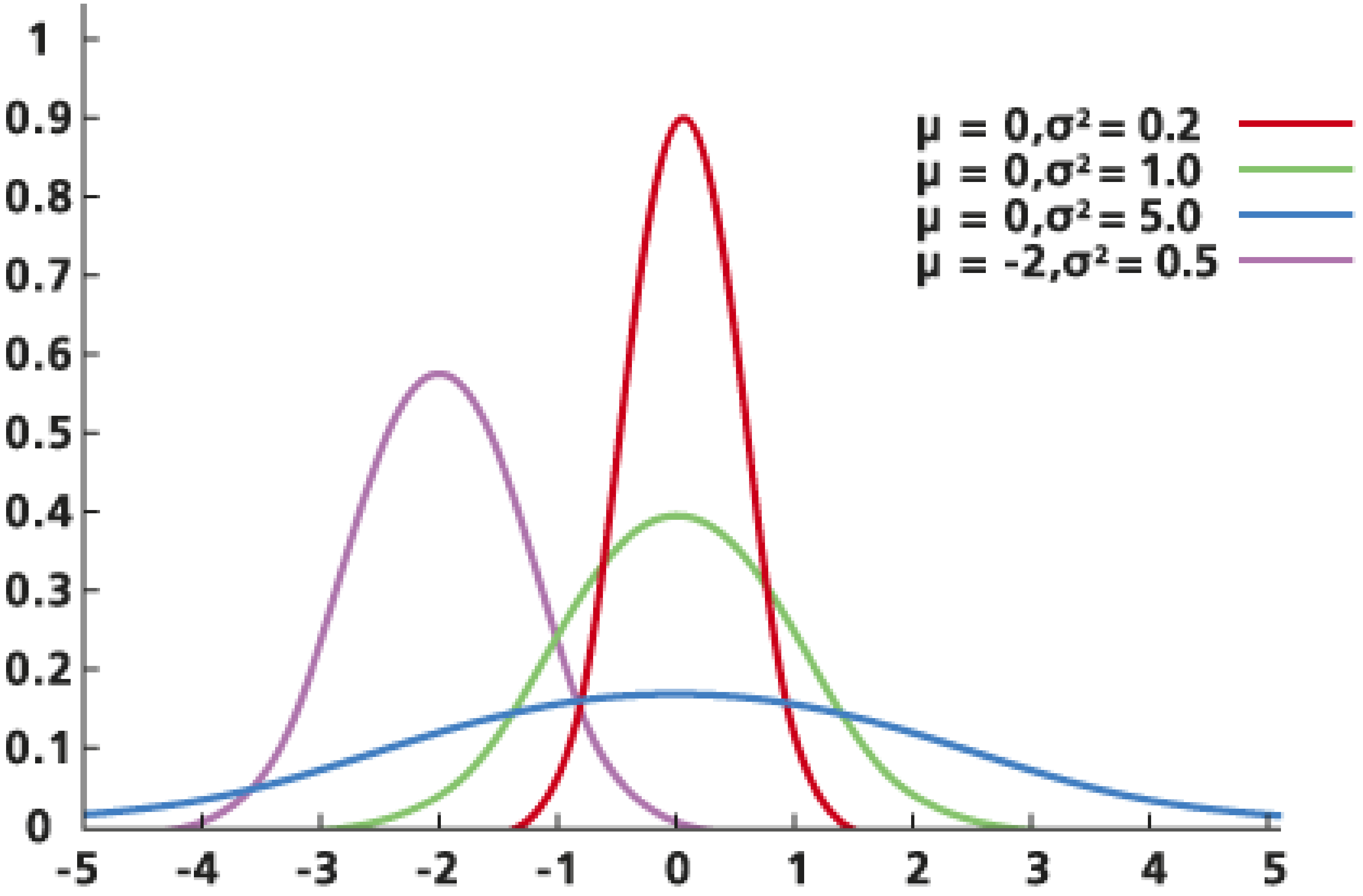


Normal Curve

- Decrease the mean, normal curve will shift to the left
- Increase the mean, normal curve will shift to the right

- Decrease the standard deviation, normal curve will get narrower
- Increase the standard deviation, normal curve will get fatter

Normal Curve



Summary of Properties of a Normal Curve

- 1. Symmetric around the mean
- 2. Highest point is at the mean=median=mode
- 3. Inflection points at $\mu \pm \sigma$
- 4. Total area under the curve is 1 (what about <> mean?)
- 5. The graph approaches 0 as we go out to either side
- 6. The Empirical Rule applies

Area Under a Normal Curve

- The area under a normal curve across any interval of values represents:

- 1. The **proportion** of the population with the characteristic described
- 2. The **probability** that a **randomly selected individual** from the population will have the characteristic

Calculating Probabilities

- We have the Empirical Rule to find probabilities for certain points within 1, 2 or 3 standard deviations from the mean, but what about the other points?
- It turns out that calculating these probabilities directly is very difficult and involves lots of calculus without any trick.

Z score

- The trick is to transform out x to z
- The Z-score represents the number of standard deviations from the mean

- $z = \frac{observation - mean}{standard\ deviation} = \frac{x - \mu}{\sigma}$

Calculating Probabilities

- So, in terms of z we have the Empirical Rule to find probabilities between points where z={-3,-2,-1,0,1,2,3}
  - 68% of the data lies between -1 and 1
  - 95% of the data lies between -2 and 2
  - 99.7% of the data lies between -3 and 3

Calculating Probabilities

- To figure out probabilities for points between other values, we need to look into a table someone made for us. (too hard to calculate directly!)
- Chart:

http://www.stat.ufl.edu/~athienit/Tables/Ztable.pdf

Calculating Probabilities

z	B .00	.01	.02	.03	.04	.05	.06	.07	.08	.09
0.0	.5000	.5040	.5080	.5120	.5160	.5199	.5239	.5279	.5319	.5359
0.1	.5398	.5438	.5478	.5517	.5557	.5596	.5636	.5675	.5714	.5753
0.2	.5793	.5832	.5871	.5910	.5948	.5987	.6026	.6064	.6103	.6141
0.3	.6179	C .6217	.6255	.6293	.6331	.6368	.6406	.6443	.6480	.6517
A 0.4	.6554	.6591	.6628	.6664	.6700	.6736	.6772	.6808	.6844	.6879
0.5	.6915	.6950	.6985	.7019	.7054	.7088	.7123	.7157	.7190	.7224
0.6	.7257	.7291	.7324	.7357	.7389	.7422	.7454	.7486	.7517	.7549
0.7	.7580	.7611	.7642	.7673	.7704	.7734	.7764	.7794	.7823	.7852

- A and B tell us that the z-score is 0.40
  - A gives us the ones place and the tenths place (**0.4**0)
  - B gives us the hundredths place (0.4**0**)

Calculating Probabilities

z	.00 (B)	.01	.02	.03	.04	.05	.06	.07	.08	.09
0.0	.5000	.5040	.5080	.5120	.5160	.5199	.5239	.5279	.5319	.5359
0.1	.5398	.5438	.5478	.5517	.5557	.5596	.5636	.5675	.5714	.5753
0.2	.5793	.5832	.5871	.5910	.5948	.5987	.6026	.6064	.6103	.6141
0.3	.6179	.6217	.6255	.6293	.6331	.6368	.6406	.6443	.6480	.6517
A 0.4	.6554 C	.6591	.6628	.6664	.6700	.6736	.6772	.6808	.6844	.6879
0.5	.6915	.6950	.6985	.7019	.7054	.7088	.7123	.7157	.7190	.7224
0.6	.7257	.7291	.7324	.7357	.7389	.7422	.7454	.7486	.7517	.7549
0.7	.7580	.7611	.7642	.7673	.7704	.7734	.7764	.7794	.7823	.7852

- C tells us that the probability that we see an observation with a z-score of 0.40 or less is .6554
- The cross-hairs created when we look right of A and down from B gives us the less-than probability for that Z-score

Calculating Probabilities

z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
0.0	.5000	.5040	.5080	.5120	.5160	.5199	.5239	.5279	.5319	.5359
0.1	.5398	.5438	.5478	.5517	.5557	.5596	.5636	.5675	.5714	.5753
0.2	.5793	.5832	.5871	.5910	.5948	.5987	.6026	.6064	.6103	.6141
0.3	.6179	.6217	.6255	.6293	.6331	.6368	.6406	.6443	.6480	.6517
0.4	.6554	.6591	.6628	.6664	.6700	.6736	.6772	.6808	.6844	.6879
0.5	.6915	.6950	.6985	.7019	.7054	.7088	.7123	.7157	.7190	.7224
0.6	.7257	.7291	.7324	.7357	.7389	.7422	.7454	.7486	.7517	.7549
0.7	.7580	.7611	.7642	.7673	.7704	.7734	.7764	.7794	.7823	.7852

- A and B tell us that the z-score is 0.25
  - A gives us the ones place and the tenths place (**0.2**5)
  - B gives us the hundredths place (0.2**5**)

Calculating Probabilities

z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
0.0	.5000	.5040	.5080	.5120	.5160	.5199	.5239	.5279	.5319	.5359
0.1	.5398	.5438	.5478	.5517	.5557	.5596	.5636	.5675	.5714	.5753
0.2	.5793	.5832	.5871	.5910	.5948	.5987	.6026	.6064	.6103	.6141
0.3	.6179	.6217	.6255	.6293	.6331	.6368	.6406	.6443	.6480	.6517
0.4	.6554	.6591	.6628	.6664	.6700	.6736	.6772	.6808	.6844	.6879
0.5	.6915	.6950	.6985	.7019	.7054	.7088	.7123	.7157	.7190	.7224
0.6	.7257	.7291	.7324	.7357	.7389	.7422	.7454	.7486	.7517	.7549
0.7	.7580	.7611	.7642	.7673	.7704	.7734	.7764	.7794	.7823	.7852

- C tells us that the probability that we see an observation with a z-score of 0.25 or less is .5987
- The cross-hairs created when we look right of A and down from B gives us the less-than probability for that Z-score

Z-Table

- $z = \frac{x-\mu}{\sigma}$
- We can then find $P(Z < z) = P(Z < \frac{x-\mu}{\sigma})$ in the z table
    - We can only look up P(Z<z) so we often have to rewrite our probabilities to look like that using rules like complements and fitting pieces
    - Z follows a **standard normal distribution**
        - Mean=0
        - Standard Deviation=1

Z-Table: Show the Empirical Rule

- The Empirical Rule states that 68% of the data lies between $\mu - \sigma$ and $\mu + \sigma$

- Let's pretend that we didn't know the Empirical Rule and find this probability using the z-table

Z-Table: Show the Empirical Rule

- **Step 1:** Make sure the data you're talking about is normally distributed
- This will be given in the problem; if not, you can look at a histogram of the data to see whether or not the histogram is symmetric and bell-shaped

Z-Table: Show the Empirical Rule

- **Step 2:** Find the z-scores

$$z_{\mu-\sigma} = \frac{(\mu \quad - \sigma \quad) - \mu}{\sigma} = -1$$

$$z_{\mu+\sigma} = \frac{(\mu \quad + \sigma \quad) - \mu}{\sigma} = 1$$

Z-Table: Show the Empirical Rule

- **Step 3:** Find the percentiles by finding the crosshairs in the z-table

$$P(Z < 1) = .8413$$

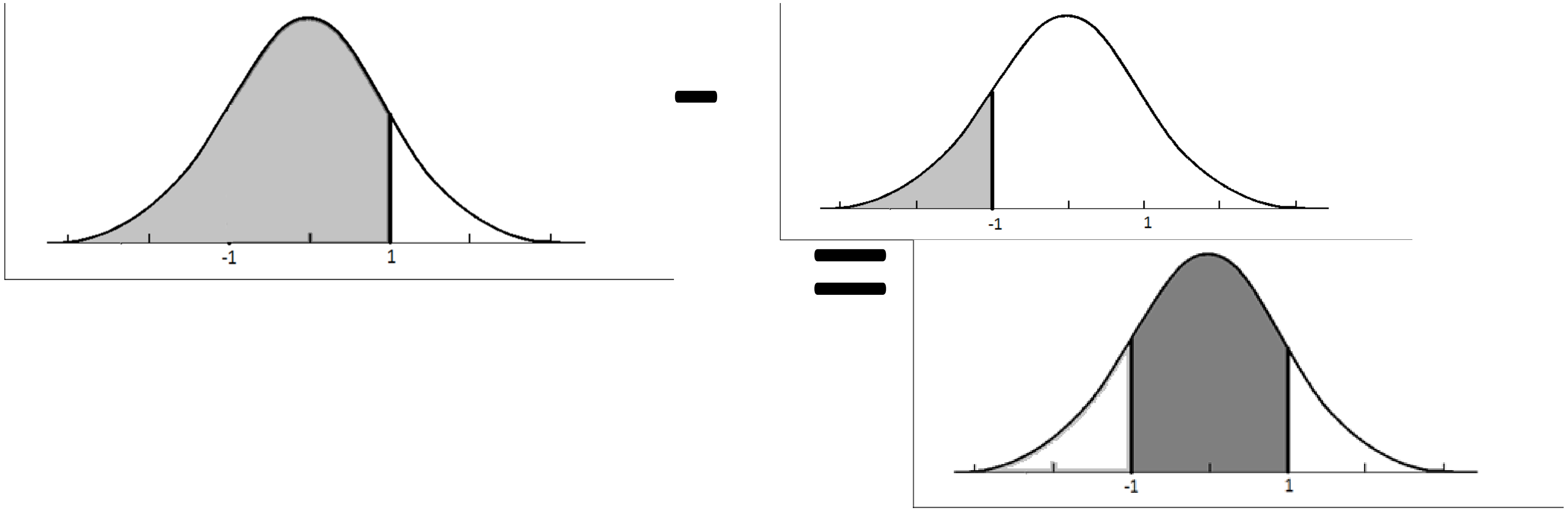
$$P(Z < -1) = .1587$$

Z-Table: Show the Empirical Rule

- **Step 4:** We can write the following by ‘fitting pieces’

$$P(\mu - \sigma < X < \mu + \sigma) = P(-1 < Z < 1)$$
$$= P(Z < 1) - P(Z < -1)$$

- The dark grey shows the difference is what we want.



Z-Table: Show the Empirical Rule

- **Step 5**: Plug in the percentiles from part 2 into the equation formed in art 3.

$$\mathrm{P}(\mu - \sigma < X < \mu + \sigma) = P(-1 < Z < 1)$$
$$= P(Z < 1) - P(Z < -1)$$
$$= .8413 - .1587 = .6826$$

Z-Table: Example 1

- Suppose you have a giant pet dragon, and you want to figure out how big he is compared to other dragons

Z-Table: Example 1

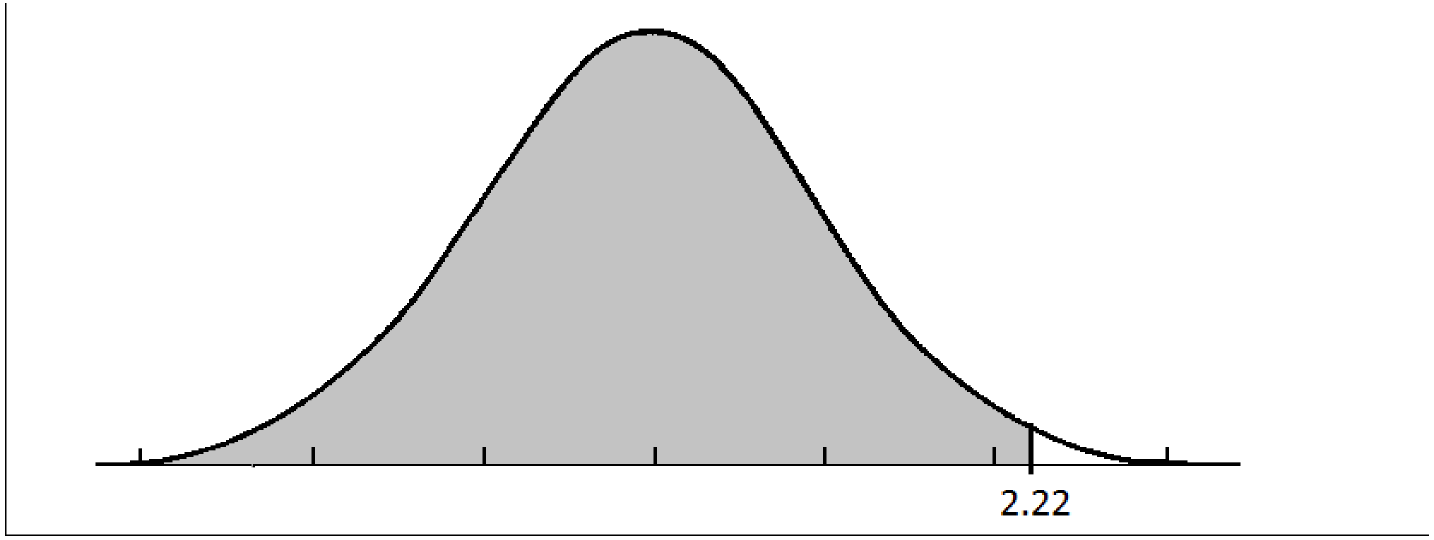
- You bring your dragon to the vet and the vet tells you that your pet dragon is 2.22 standard deviations above the mean in wingspan length, which is normally distributed, with a measurement of 397 cm.

Z-Table: Example 1

- This is equivalent to saying that your pet dragon's wingspan of x=397 gives him a z-score of 2.22
- Now that we know x=397 gives z=2.22, we can find the percentile of dragons that are smaller than yours!

Z-Table: Example 1

- Look up 2.22 in the z-table, we find P(Z<2.22) = .9868
- We now know that your dragon's wingspan is greater than 98.68% of all pet dragons!



Z-Table: Example 1

- What if we want to know the percentage of dragons whose wingspan is greater than yours?
- Here we use complement rule because:

$$P(Z>2.22)=1-P(Z<2.22) = .0132$$

- Now, we know that only 1.32% of dragons are bigger than yours!

Z-Table: Example 1

- This example could have played out very differently depending on what your vet told you.
- Your vet can say that your pet dragon has a longer wingspan than 98.68% of dragon
- Now, you only know the percentile of your dragon is 98.68, but what if we want to know your dragon's actual wingspan?

Z-Table: Example 1

- Suppose you did some research and find that the pet dragon's wingspan has mean 200 cm and standard deviation 88.738 cm

Z-Table: Example 1

- 1st step: we need to find out which z-score gives us a percentile of 98.68 by checking z-table. It is gives us z=2.22!
- 2nd step: Set up z-score equation $z = \frac{x-\mu}{\sigma}$ and solve
  - $2.22 = \frac{x-200}{88.\overline{738}}$
  - $2.22 * 88.738 = x - 200$
  - $x = 2.22 * 88.738 + 200 = 397$ cm
  - Note: this matches the measurement the vet gave in the first scenario

Z-Table: Example 2

- Now, suppose we want to know the temperature of the fire your pet dragon breaths!
- Before you leave the vet, the vet tells you that your pet dragon is 1.48 standard deviations below the mean in fire temperature with a measurement of 400 Celsius degrees (752 Fahrenheit!)

Z-Table: Example 2

- This is equivalent to say your dragon's fire temperature of x=400 gives him a z-score of -1.48

- Now that we know x=400 gives z=-1.48 we can find the percentage of dragons who's fire is cooler than yours

Z-Table: Example 2

- Look up -1.48 in the z-table, it gives us that P(Z<-1.48) = .0694

- Only 6.94% of dragons breath fire is cooler than yours – looks that yours is pretty “safe” ☺

Z-Table: Example 3

- Your giant pet dragon is adamant about getting "elective surgery" on his tail simply because it's not long enough and he is worried about finding a girl friend.
- You fly with him to Korea to express his desire with a Korean vet and the vet tells him that he doesn't need the surgery because his tail is about the average in length – its between 0.45 standard deviations of the mean tail length.

Z-Table: Example 3

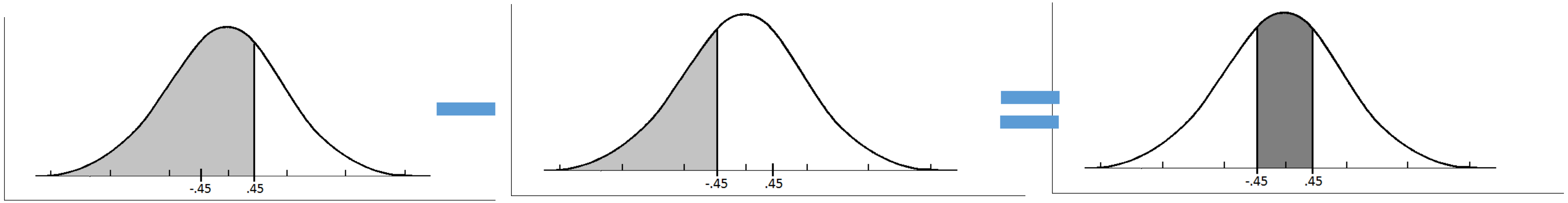
- You make it home with your disappointed dragon and you want to make him feel better by analyzing the statistics the Korean vet gave.
- After all, maybe the Korean vet was right – what percentage of dragons have a comparable tail size?

Z-Table: Example 3

- Within 0.45 standard deviations of the mean tell us that we have to worry about two z-scores:

  $z = -0.45$ and $z = 0.45$

- Let's consider the puzzle pieces we can fit here



Z-Table: Example 3

- Within 0.45 standard deviations of the mean
  1) Find the probability of a z-score less than 0.45
     - P(Z<0.45)=0.6736
  2) 2) Find the probability of a z-score less than -0.45
     - P(Z<-0.45)=0.3264

  3) This gives us P(-0.45<Z<0.45) = 0.6736-0.3264 = 0.3472
     - This means that your dragon has a similar tail to 34.72% of dragons

Z-Table: Example 3

- Even though your dragon now knows he has a similar tail length to 34.72% of other pet dragons, he still doesn’t feel confident about himself!
- Your dragon argues that the group of comparable dragons is too small by explaining 65.28% of dragons aren’t comparable
- Let’s check his math

Z-Table: Example 3

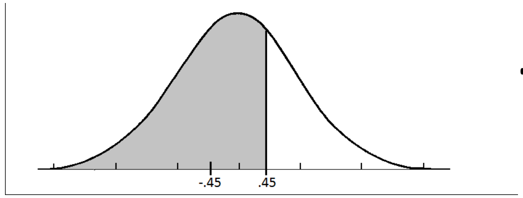
- Not having comparable tail means being outside of 0.45 standard deviations of the mean

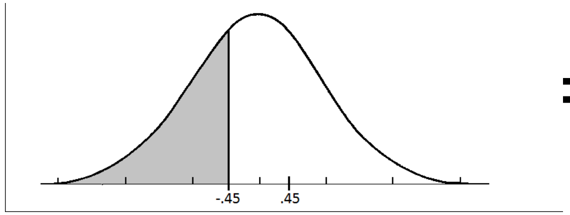
  1) Find what is within 0.45 standard deviations of the mean → $P(-0.45<Z<0.45)$

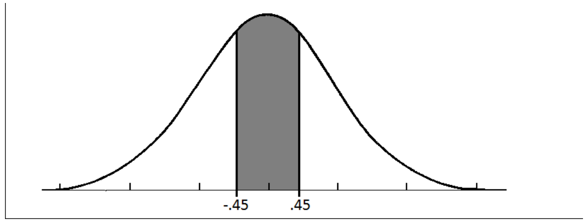
  2) Use complement rule → $1-P(-0.45<Z<0.45)$

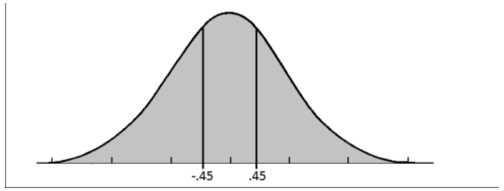
Z-Table: Example 3

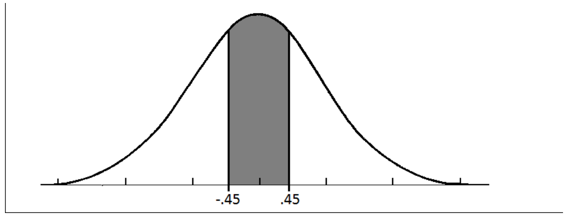
- Outside of .45 standard deviations of the mean

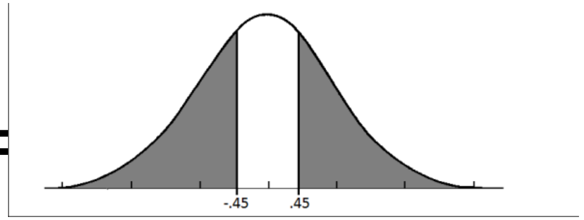
- 

-


=


- 

-


=


Z-Table: Example 3

- Outside of 0.45 standard deviations of the mean

  1) Find what is within 0.45 standard deviations of the mean
  → $P(-0.45<Z<0.45) = P(Z<0.45) – P(Z<-0.45) = 0.3472$

  2) Use complement rule

  → $1-P(-0.45<Z<0.45) = 1 – 0.3472 = 0.6528$

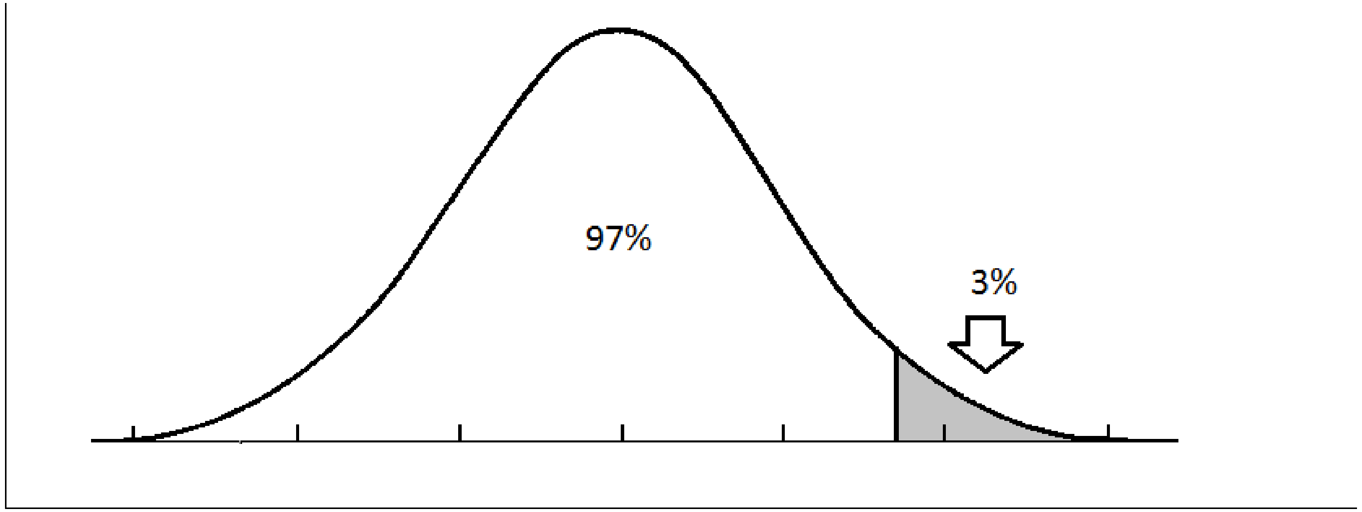
- Your dragon was right, but you explain that not all dragons outsides of his group have bigger tail – actually, half of them have smaller tails!

Z-Table: Example 4

- Your dragon isn’t receptive to any of the statistics and explain. ☹ It doesn’t matter what other dragon’s tails look like. He isn’t comfortable and wouldn’t be happy until he has a tail in the top 3% of dragons!
- Finally, you decide to fly back to Korea with your dragon and let him have the elective surgery to make him happy - which may cost $80,000, your family’s annual income!

Z-Table: Example 4

- How long does your dragon's tail have to be to be in the top 3%?
- Find the z-score of the observation where 3% of the data lies about it.



Z-Table: Example 4

- We can only find percentile in the z-table, so we need to use complement rule to figure out how much of the data lies below it. (1 – 0.03) = 0.97
- We need to find the z-score that gives us 0.97, which is z = 1.88

Z-Table: Example 4

- With z = 1.88 we can find the x observation that coincides with it, so we can tell the Korean vet how long the tail the dragon needs.
- To find the length from the z-score you did some research to find that dragon tail length is normally distributed with mean 244 cm and standard deviation 25 cm
  - $\mu = 244cm \qquad \sigma = 25$

Z-Table: Example 4

- Set up the z-score equation $z = \frac{x-\mu}{\sigma}$ and solve
  - $1.88 = \frac{x-244}{25}$
  - $1.88 * 25 = x - 244$
  - $x = 1.88 * 25 + 244 = 291$ cm
- To be in the top 3%, your dragon's tail should be 291 cm long (around 3 m!)

Z-Table: Example 4

- You dragon is very happy about his “new” tail.
- However, here is the new problem. He is too big to fit the airplane back to US. He has to stay in Korean in the rest of his life, competing with all those pet dragons with “elective surgery”!!

Z-Table: Word Problem 1

- Wii Fit showed that average calories burned in 30 minutes is 111
- Assume a bell-shaped distribution with a standard deviation 20
- Find the z-score of someone that burns 130 calories

$$z = \frac{observation - mean}{st.dev} = \frac{x - \mu}{\sigma} = \frac{130 - 111}{20} = 0.95$$

Z-Table: Word Problem 1

- What is his percentile?
  - It is **NOT** 0.95 or 95TH percentile!!!!
  - Look up 0.95 in the z table and that will give you the percentile
- $P(Z<0.95)=0.8289$ → 82.89th percentile

Z-Table: Word Problem 1

- How many calories would a man have to burn to be at the 85th percentile?
  1) Find the Z that gives the 85th percentile

     0.8508 is the closest value we can find in the z-table and

     z=1.04 is the z-value that corresponds to it.
  2) Set up and solve the following equation
     - $1.04 = \frac{x - 111}{20}$
     - x= $(1.04 * 20) + 111 = 131.8$
  3) Someone would have to burn 131.8 calories to be at the 85th percentile.

Z-Table: Word Problem 2

- Some students apply the university of SAT score and some with ACT score.

- Tom scored 550 on the SAT and Jerry scored 22.5 on the ACT – who has a better score?

Z-Table: Word Problem 2

- SAT score is symmetric, bell-shaped with mean score of 500 with a standard deviation of 100

- ACT score is symmetric, bell-shaped with mean score of 21 with a standard deviation of 4.1

Z-Table: Word Problem 2

- Tom scored 550 on the SAT and Jerry scored 22.5 on the ACT – who has a better score?

$$z_{Tom} = \frac{550-500}{100} = 0.5 \quad \rightarrow \text{P(Z<0.5)=0.6915}$$

$$z_{Jerry} = \frac{22.5-21}{4.1} = 0.37 \quad \rightarrow \text{P(Z<0.37)=0.6443}$$

- We can see that Tom's score means he did better than 69.15% of students and Jerry did better than 64.43% of students. Thus, Tom did better!