Standard Normal Table
Area Under the Curve

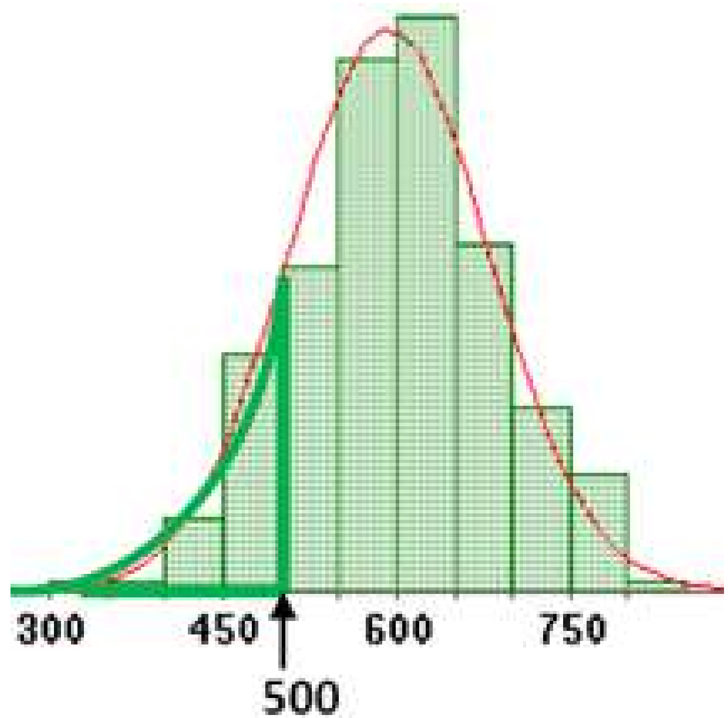
Learning Objectives

By the end of this lecture, you should be able to:

- Be able to look up a z-score on a standard Normal table and interpret the number that is found there.
- Be able to do various calculations involving areas under the density curve using the Normal table

* Review: Area under the density curve

- Recall that the area under a density curve refers to the **proportion of observations that fall inside that range**.
- In the graph shown here, the area under the curve to the left of '500' represents the proportion of people who took the SAT Math that scored below 500.



: SAT math scores for a set of 200

Another shortcut to learn

- In addition to Greek letters, statisticians sometimes employ written shortcuts to represent various distributions and their key numbers.
- For example, a shortcut to represent a Normal distribution is 'N' followed by two numbers in parentheses: N(n1, n2)
  - n1 represents the mean
  - n2 represents the standard deviation
- For our grade equivalent score, rather than say: "This distribution was approximately Normal with a mean of 7 and a standard deviation of 2.17, we would simply write: **N(7, 2.17)**.

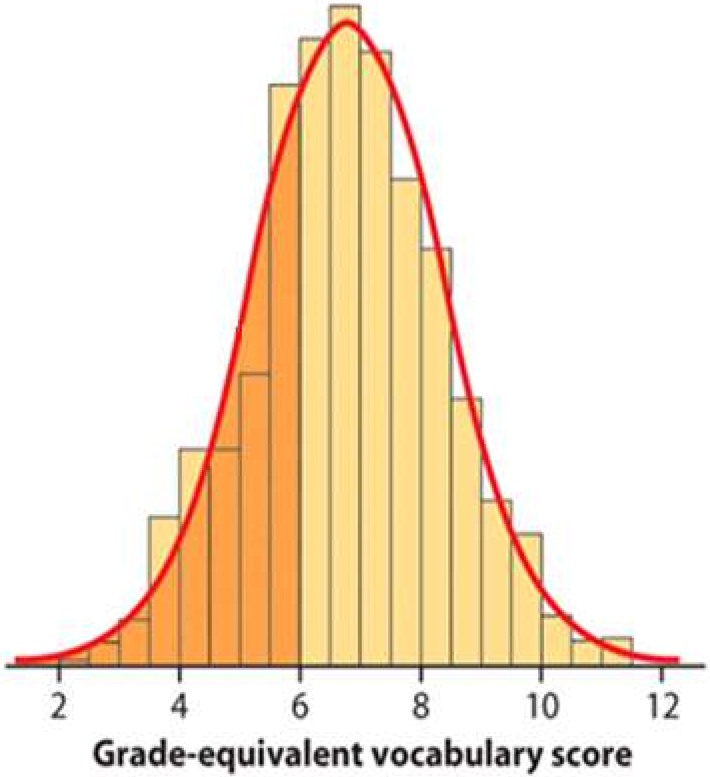
Review:
Overview of determining the area under the curve

Recall from a prior lecture:

> ***ANY value 'x' can be converted to a corresponding z-score.***

The process:

1. Convert your 'x' into a z-score
2. Look up that z-score on a **Standard Normal Table** (aka: a **z-table**)
3. The value you find on the z-table, is the area under the curve to the left of your z-score.

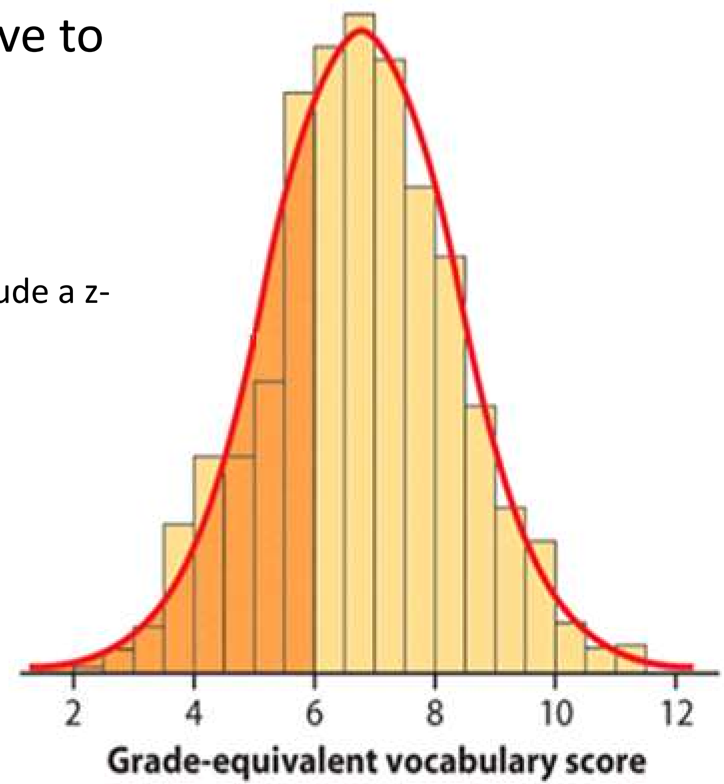


Use the “**standard Normal table**”

- Also known as the “***z-table***”
- Values found in this table tell us the area under the curve to the left of that z-score.
- So if your calculated z-score is -1.46 , the value you find on the z-table tells you the area under the curve to the left of -1.46. (Example on next slide)
- If your calculated z-score is 0.81, the value you find on the z-table tells you the area under the curve to the left of 0.81.

- This is a widely used table.
  - I guarantee that whichever textbook you purchased will include a z-table
  - There is a link to one at the top of the course web page
  - Google z-table and you’ll find hundreds

- http://lilt.ilstu.edu/dasacke/eco148/ztable.htm



Looking up -1.46 on a z-table

z	.00	.01	.02	.03	.04	.05	.06	.07
-3.4	.0003	.0003	.0003	.0003	.0003	.0003	.0003	.000
-3.3	.0005	.0005	.0005	.0004	.0004	.0004	.0004	.000
-3.2	.0007	.0007	.0006	.0006	.0006	.0006	.0006	.000
-3.1	.0010	.0009	.0009	.0009	.0008	.0008	.0008	.000
-3.0	.0013	.0013	.0013	.0012	.0012	.0011	.0011	.001
-2.9	.0019	.0018	.0018	.0017	.0016	.0016	.0015	.001
-2.8	.0026	.0025	.0024	.0023	.0023	.0022	.0021	.002
-2.7	.0035	.0034	.0033	.0032	.0031	.0030	.0029	.002
-2.6	.0047	.0045	.0044	.0043	.0041	.0040	.0039	.003
-2.5	.0062	.0060	.0059	.0057	.0055	.0054	.0052	.005
-2.4	.0082	.0080	.0078	.0075	.0073	.0071	.0069	.006
-2.3	.0107	.0104	.0102	.0099	.0096	.0094	.0091	.008
-2.2	.0139	.0136	.0132	.0129	.0125	.0122	.0119	.011
-2.1	.0179	.0174	.0170	.0166	.0162	.0158	.0154	.015
-2.0	.0228	.0222	.0217	.0212	.0207	.0202	.0197	.019
-1.9	.0287	.0281	.0274	.0268	.0262	.0256	.0250	.024
-1.8	.0359	.0351	.0344	.0336	.0329	.0322	.0314	.030
-1.7	.0446	.0436	.0427	.0418	.0409	.0401	.0392	.038
-1.6	.0548	.0537	.0526	.0516	.0505	.0495	.0485	.047
-1.5	.0668	.0655	.0643	.0630	.0618	.0606	.0594	.058
-1.4	.0808	.0793	.0778	.0764	.0749	.0735	.0721	.070
-1.3	.0968	.0951	.0934	.0918	.0901	.0885	.0869	.085
-1.2	.1151	.1131	.1112	.1093	.1075	.1056	.1038	.102
-1.1	.1357	.1335	.1314	.1292	.1271	.1251	.1230	.121
-1.0	.1587	.1562	.1539	.1515	.1492	.1469	.1446	.142

Looking up +0.81 on a z-table

z	.00	.01	.02	.03	.04	.05	.06
0.0	.5000	.5040	.5080	.5120	.5160	.5199	.523
0.1	.5398	.5438	.5478	.5517	.5557	.5596	.563
0.2	.5793	.5832	.5871	.5910	.5948	.5987	.602
0.3	.6179	.6217	.6255	.6293	.6331	.6368	.640
0.4	.6554	.6591	.6628	.6664	.6700	.6736	.677
0.5	.6915	.6950	.6985	.7019	.7054	.7088	.712
0.6	.7257	.7291	.7324	.7357	.7389	.7422	.745
0.7	.7580	.7611	.7642	.7673	.7704	.7734	.776
0.8	.7881	.7910	.7939	.7967	.7995	.8023	.805
0.9	.8159	.8186	.8212	.8238	.8264	.8289	.831
1.0	.8413	.8438	.8461	.8485	.8508	.8531	.855
1.1	.8643	.8665	.8686	.8708	.8729	.8749	.877
1.2	.8849	.8869	.8888	.8907	.8925	.8944	.896
1.3	.9032	.9049	.9066	.9082	.9099	.9115	.913
1.4	.9192	.9207	.9222	.9236	.9251	.9265	.927
1.5	.9332	.9345	.9357	.9370	.9382	.9394	.940
1.6	.9452	.9463	.9474	.9484	.9495	.9505	.951
1.7	.9554	.9564	.9573	.9582	.9591	.9599	.960
1.8	.9641	.9649	.9656	.9664	.9671	.9678	.968
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]

What do these Normal table numbers *mean*?

z	.00	.01	.02
0.0	.5000	.5040	.508
0.1	.5398	.5438	.547
0.2	.5793	.5832	.587
0.3	.6179	.6217	.625
0.4	.6554	[illegible]	.662

- On a "standard Normal table", the numbers that you see represent the proportion of observations under the curve to the left of that z-score.
  - Put another way: The area under a Normal density curve to the left of a z score.
- Recall that every 'x' has a corresponding z-score. ($z = (x-\mu) / \sigma$)
- When you look up a z-score on a Normal table, the value you find tells you what the area is under the curve to the left of your z-score. **It is a percentage** of values.
  - It initially doesn't look like a percentage though. In fact, it is what we call a 'probability'. However, a probability and a percentage are essentially the same thing, just multiply the probability by a 100.
  - So for a z-score of 0.31, your probability shows 0.6217. To get a percentage, multiply by 100: 62%.
  - **This means that the area under the curve to the left of z-0.31 is 62%.**

Example: What percentage of people scored less than 6?

Answer:

You can't tell! Recall that you need the mean and SD to determine a z-score.

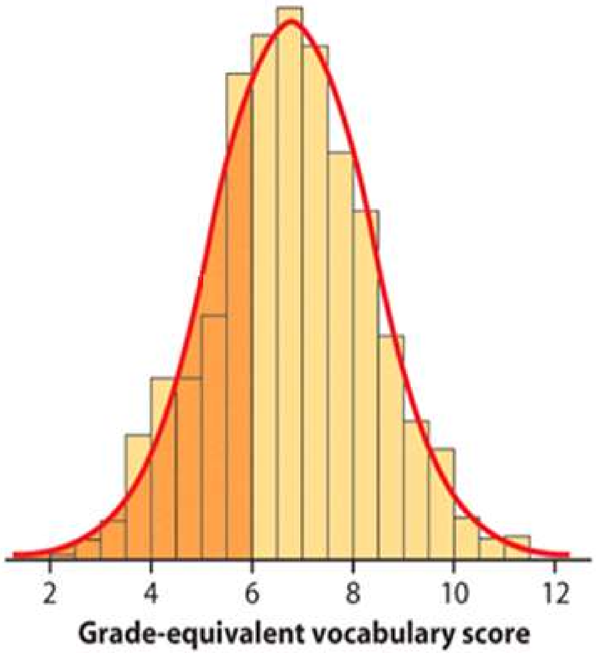
Recall: This is statistical shorthand for saying that the dataset is Normally distributed with a mean of 7 and an sd of 2.44.

Okay - suppose I give you: **N(7, 2.44)**. Now what is the z-score?

Answer: $z = (6 – 7) / 2.44$

$= \mathbf{-0.41}$

Important point: Once you have the z-score, you can – and probably should – forget the actual number! Focus only on the z-score. So, in this case, we no longer think in terms of the grade score '6'. Instead, we only think in terms of being 0.41 standard deviations below the mean.



Looking up **-0.41** on a z-table

-0.7	.2420	.2389	.2358	.2327	.2296
-0.6	.2743	.2709	.2676	.2643	.2611
-0.5	.3085	.3050	.3015	.2s981	.2946
-0.4	.3446	.3409	.3372	.3336	.3300
-0.3	.3821	.3783	.3745	.3707	.3669
-0.2	.4207	.4168	.4129	.4090	.4052
-0.1	.4602	.4562	.4522	.4483	.4443
0.0	.5000	.4960	.4920	.4880	.4840

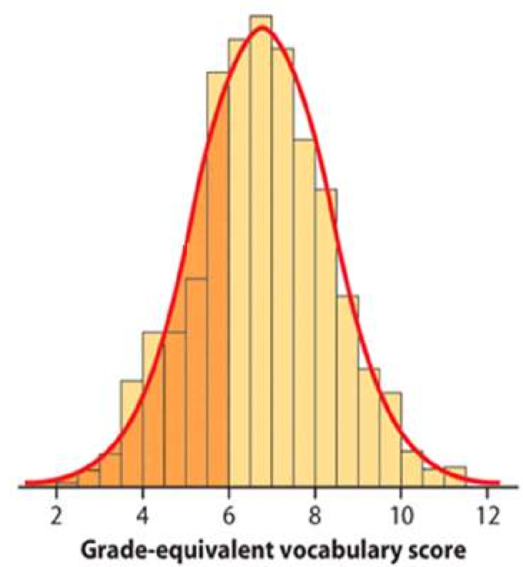
z	.00	.01	.02	.03	.04
0.0	.5000	.5040	.5080	.5120	.5160
0.1	.5398	.5438	.5478	.5517	.5557
0.2	.5793	.5832	.5871	.5910	.5948
0.3	.6179	.6217	.6255	.6293	.6331

What does this number from the z-table tell us?

-0.5	.3085	.3050	
-0.4	.3446	.3409	
-0.3	.3821	.3783	
-0.2	.4207	.4168	

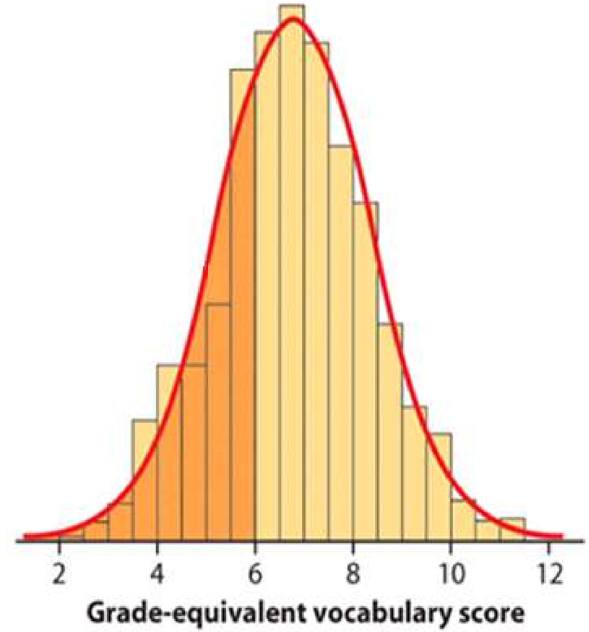
- Our table tells us that a z-score of -0.41 corresponds to a probability of **0.3409**. We multiply by 100 to get: 34.09, or **34%.**

- **Question:** Do you remember what this percentage tells us in terms of **area under the density curve**?
- **Answer:** It tells us the area to the **left** of the z-score. So the area to the left of z = -0.41 is 34%.
- **Conclusion:** 34% of people who took this exam scored less than 6. Put another way, if a person had a score of exactly 6, you would say that they were in the 34th percentile.



Summary slide of previous example:

About what percentage of students scored less than 6 on this vocabulary exam? N(7, 2.44)



1. Using the **standard deviation**, convert the value you are interested in (e.g. a Grade equivalent score of 6.0) into a **<u>z-score</u>**.
    - $z = (x - \mu) / \sigma$
    - $= (6-7) / (2.44)$
    - $= -0.41$

2. Look up your z-score on a **Standard Normal Table** (aka z-table)
    - \- 0.41 corresponds to 0.3409 or 34%

-0.5	.3085	.3050	
-0.4	.3446	.3409	
-0.3	.3821	.3783	
-0.2	.4207	.4168	

3. The value you find on the z-table, is the area under the curve to the left of your score (e.g. 6.0).
    - In other words, 34% of people scored less than 6 on this exam.

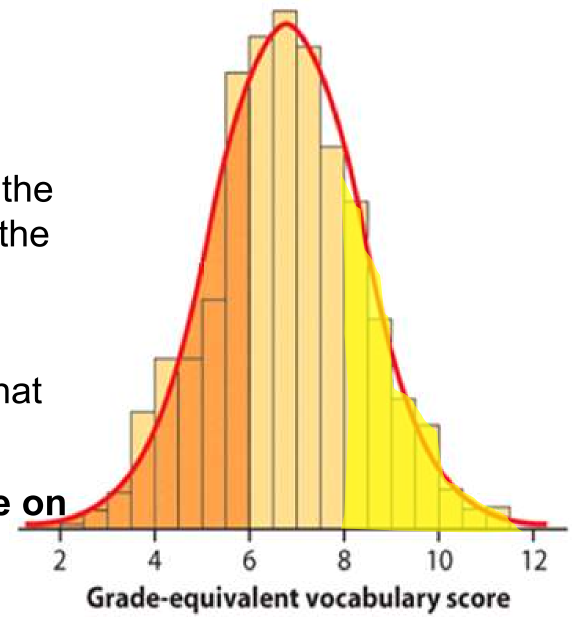
Another example:

What percentage of students scored 8.5 or more on this exam? Assume N(7, 2.17).

1. First: **PICTURE what is being asked**. You are being asked to determine the area under the curve to the right of 8.5 (shaded in yellow).
2. Find the z-score
   - $z = (x-\mu) / \sigma$
   - $= (8.5 - 7) / (2.17)$
   - $= +0.69$

7	.08	.09
79	.5319	.5359
75	.5714	.5753
64	.6103	.6141
43	.6480	.6517
08	.6844	.6879
57	.7190	.7224
86	.7517	.7549
94	.7823	.7852
78	.8106	.8133

3. Look up your z-score on a **Standard Normal Table** (aka z-table)
   - A z of 0.69 corresponds to 0.7549 or 75.5%



IMPORTANT! Remember that the value on a z-table corresponds to the area to the LEFT of the z-score. So the value of 0.7549 tells you the area under the curve to the LEFT of 8.5. However, the **question** asked for the number of students who scored MORE than 8.5.

Therefore, if about 75% of people scored less than 8.5, this means that 25% of people scored more.

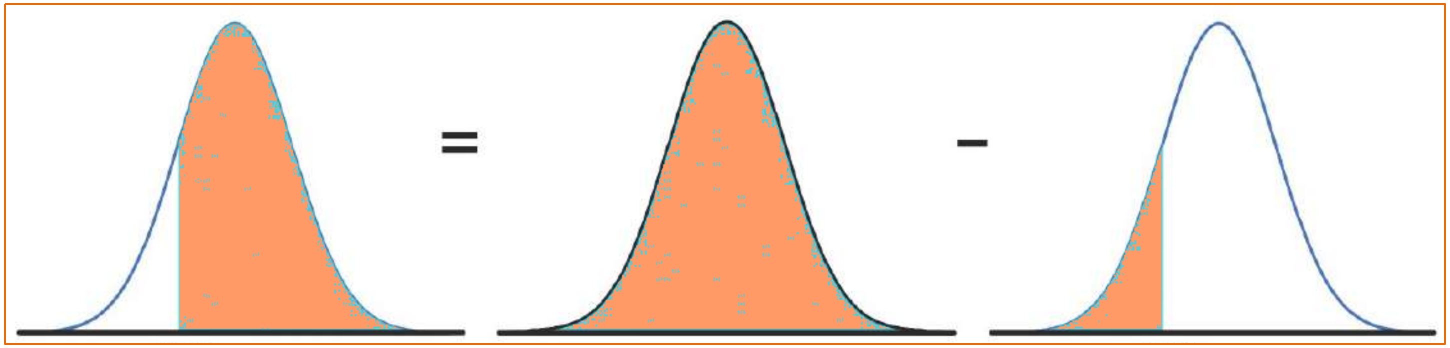
- **Answer: (100-75.5) = <u>24.5%</u> of people scored 8.5 or more on this exam.**

Values on the Normal table give you areas to the LEFT of 'z'

Remember: The Normal table gives you the area under the curve to the LEFT of z. Therefore, if you need to determine the area to the right of a z-value, simply subtract that value from 1.

That is, area to the right = (1 – area to the Left)

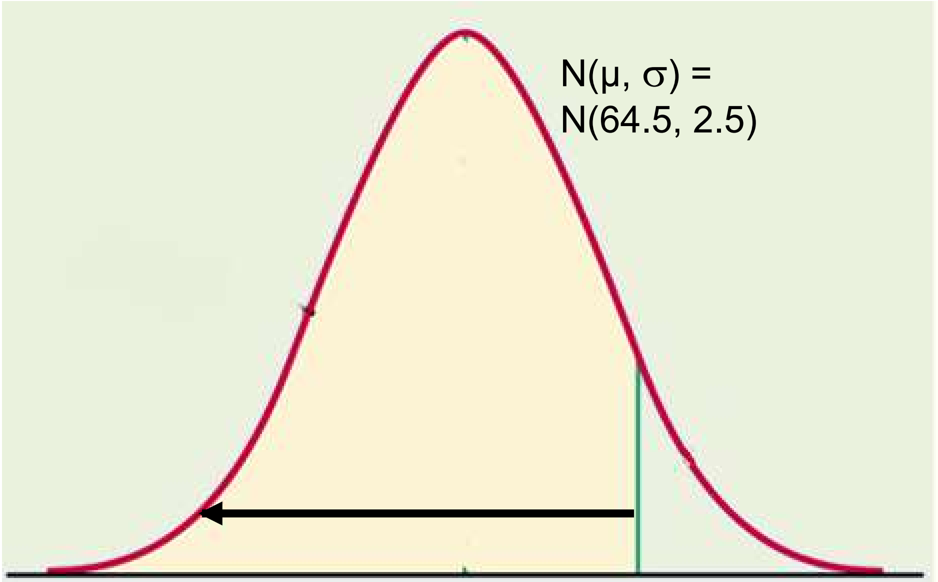
Example: A z value of -0.53 comes up on the z-table as 0.30 (or 30%). This means that the area under the curve to the left of -0.53 is 30%. Therefore, the area under the curve to the RIGHT of -0.53 is 70%



area right of z = 1 - area left of z

Ex. Women heights

Women's heights have the distribution **N(64.5,2.5)**. What percent of women are shorter than 67 inches tall (that's 5'6")?



mean μ = 64.5"
standard deviation σ = 2.5"
x (height) = 67"

Start by calculating the z-score.

$$z = \frac{(x - \mu)}{\sigma}, \quad z = \frac{(67 - 64.5)}{2.5} = \frac{2.5}{2.5} = 1 \Rightarrow \ +1 \text{ SD's above the mean}$$

Use a z-table to find the area under the curve to the left of 1.

The table tells us that the area under the Normal curve to the left of z = 1.0 is 0.84 (or 84%).

Conclusion:

84% of women are shorter than 67. (And we could also say that 16% of women are 67" or taller).

Example

- A different sample of women had their heights recorded in inches. When graphed, the distribution was found to be: **N(64.5,1.9)**. From this sample, about what percentage of women are taller than 67 inches tall (that is, 5’6”)?
- **Answer:**
- **z = (x – μ) / σ**

  = (67 – 64.5) / 1.9

  = +1.32
- On the z-table, we find that +1.32 corresponds to an area of 0.9066, or about 91%. Again, this means that a 91% of values lie to the left of 1.32. However, because we are asking for the number of people *taller* than 1.32, we are interested in the area under the curve to the RIGHT of 1.32. This corresponds to 1-91% or <u>9%</u>.

Example: The National Collegiate Athletic Association (NCAA) requires Division I athletes to score at least 820 on the combined math and verbal SAT exam to compete in their first college year. The SAT scores of 2003 were *approximately* normal with mean 1026 and standard deviation 209.

What proportion of all students would be NCAA qualifiers (SAT ≥ 820)?

$$x = 820$$

$$\mu = 1026$$

$$\sigma = 209$$

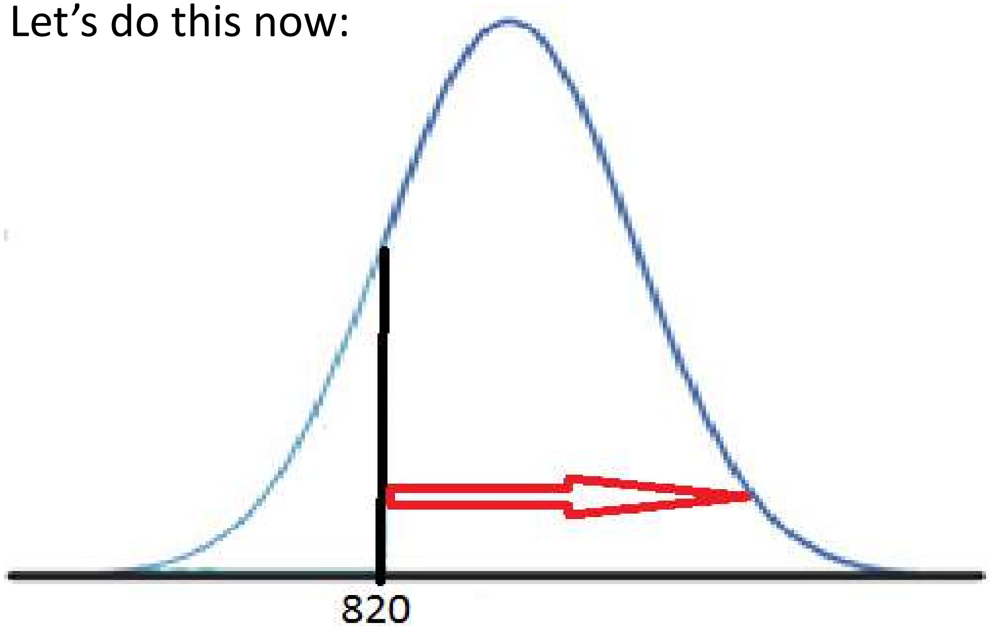
$$z = \frac{(x - \mu)}{\sigma}$$

$$z = \frac{(820 - 1026)}{209}$$

$$z = \frac{-206}{209} \approx -0.99$$

Area under
the curve to the left of
z -.99 is 0.1611
or approx. 16%.
Area to the RIGHT is about 84%.

Remember that the first step is always to get a good picture in your head of what is being asked. The WORST thing you can do is to start banging numbers into a calculator without being sure you understand what you are being asked to do!
Let's do this now:



Answer: About 84% of students would qualify.

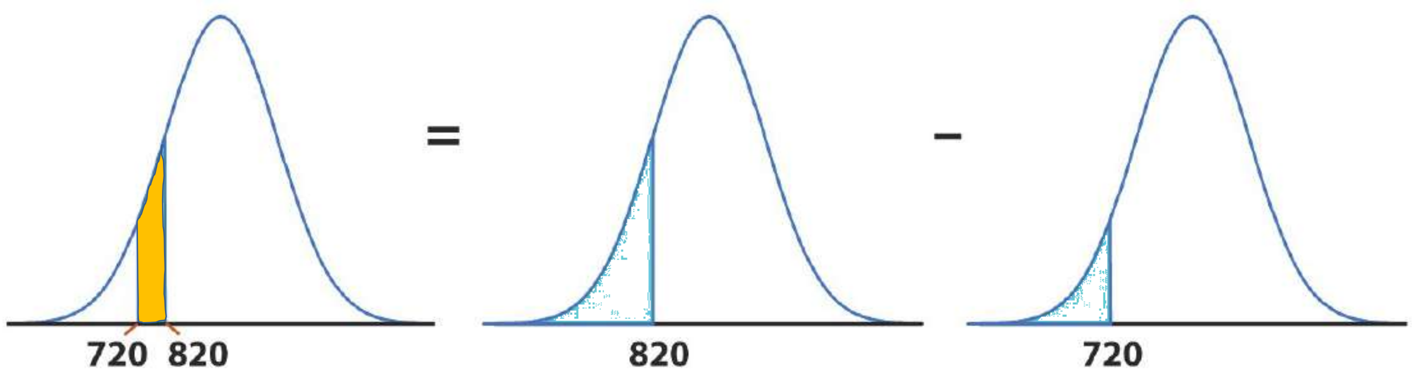
What do we mean by “approximately normal”?

- IMPORTANT: why do they say / what do they mean by “approximately normal”?
- **Answer**: It is rare to have a distribution that is perfectly normal. It is more of an ideal (much in the same way that there is, technically speaking, no such thing as the perfect circle). In appreciation of this fact, you will often find statisticians saying that a given distribution is *approximately* Normal.

The NCAA defines a “partial qualifier” eligible to practice and receive an athletic scholarship, but not to compete, with a combined SAT score of at least 720.

What proportion of all students who take the SAT would be partial qualifiers? That is, what proportion have scores between 720 and 820?

Answer: The key here is to find the area <u>between</u> 720 and 820. To do this, we calculate the area to the left of <u>each</u> and then find the difference between the two.



area between 720 and 820 **=** **area left of 820** **-** **area left of 720**

= **0.1611** **-** **0.0721**

= **about 9%**

Conclusion: About 9% of students would be considered “partial qualifiers.”

**** Example:** The SAT scores of 2003 were approximately normal with mean 1026 and standard deviation 209. One student reports that he was in the 62nd percentile. What was his score?

In this case, we have to work backwards:

1. **Look on a z-table for the 62nd percentile (or any value close to 0.62).**

z	.00	.01	.02
0.0	.5000	.5040	.5080
0.1	.5398	.5438	.5478
0.2	.5793	.5832	.5871
0.3	.6179	.6217	.6255
0.4	.6554	.6591	.6628
0.5	.6915	.6950	.6985
0.6	.7257	.7291	.7324

2. ***Find the corresponding z-score.*** *Note: In this case, a z-score of either 0.30 or 0.31 would be perfectly acceptable answers.*

3. We are familiar with the formula: $z = (x - \mu) / \sigma$. However, in this case, the missing variable is 'x'. No problem: back to high-school algebra, and we rearrange: $\mathbf{x = (z * \sigma) + \mu}$

   $x = (0.30 * 209) + 1026$

   $x =$ **1089**

Who cares about the area under the curve???

- We will be spending a lot of time looking at Normal curves throughout the course and calculating the areas under the curve. For this reason, it is very important that you do not simply learn to calculate the answer, but rather, **that you understand what it is you are concluding when you come up with a numeric “answer”**.
- When we want to draw conclusions, or make predictions, and so on, we begin with a sample. For example, if we are interested in the average height of women at our university, we may start by taking a random sample of, say, 25 women and measure their heights. From there, we graph the data and decide if it looks like a Normal distribution. If it does, then we can take this data and try to infer information about ALL women at our university.

The Worst Thing You Can Do:

- Is to keep plugging numbers into formulas until you find an answer that seems right.

- When answering these types of questions, it is VITAL that you make sure you understand the question being asked in terms of how it looks on the graph.

- In fact, truly understanding the question being asked should be your goal not just for Normal distribution / z-score type questions, but for nearly all statistical questions.