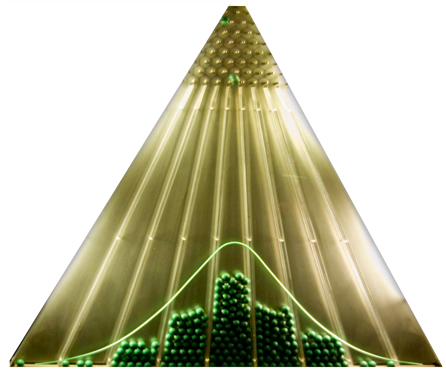
6.2 Applications of the Normal Distribution

The bean machine, a device invented by Francis Galton, can be called the first generator of normal random variables. This machine consists of a vertical board with interleaved rows of pins. Small balls are dropped from the top and then bounce randomly left or right as they hit the pins. The balls are collected into bins at the bottom and settle down into a pattern resembling the Gaussian curve.



A virtual demonstration of this box can be seen at http://www.youtube.com/watch?v=AUSKTk9ENzg

In order to utilize certain features of the normal distribution to solve statistics problems, we must transform the given distribution to the standard normal distribution in which the mean, $\mu = 0$ and $\sigma = 1$.

In some math classes, students are taught to transform the graphs of equations including shifting and stretching. Transforming a normal curve to a standard normal curve is a real world example of how the techniques of transformation are useful.

Here is the formula by which this is accomplished:

$$z = \frac{X - \mu}{\sigma}$$

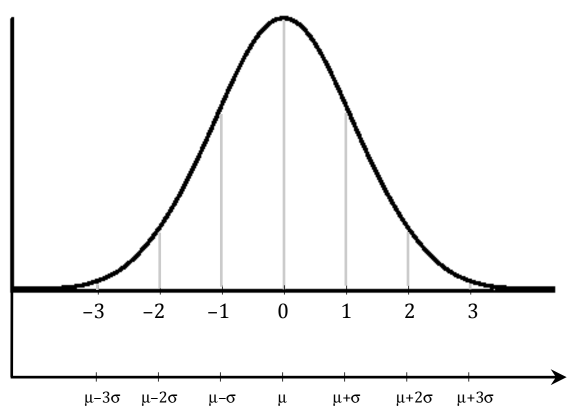
This formula shifts the curve so that the

mean is 0

and stretches or shrinks the curve so that the

standard deviation is 1.

That is, $X \sim N(0, 1)$.



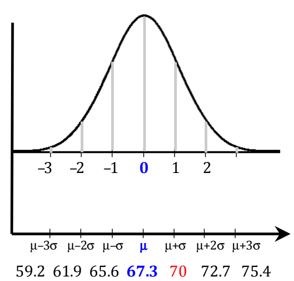
Since men's heights are approximately normally distributed, we can rework two of the Connecticut Agricultural College ROTC cadet height problems with the aid of the *z*-table.

Here is the work we did in section 6.1

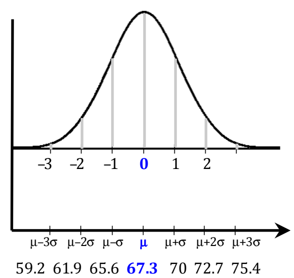
1. What is the probability that a randomly selected cadet from this group of 175 will have a height greater than 70 inches? $P(X > 70) = \frac{26}{175} \approx 0.149$	**2.** What percent of the group is shorter than 68 inches? Percent shorter than 68" $= \frac{99}{175} \approx 0.566$
3. What are the heights of the cadets in the shortest 1% of this group? Number of cadets in shortest 1% = 0.01 • 175 = 1.75 ⇒ 2 Heights of shortest 1% are 4'10" and 5'1"	**4.** What are the heights of the cadets in the tallest 5% of this group? Number of cadets in tallest 5% = 0.05 • 175 = 8.75 ⇒ 9 Heights of tallest 5% are 6', 6'1", and 6'2"

Let's rework these problems using a standardized normal curve.

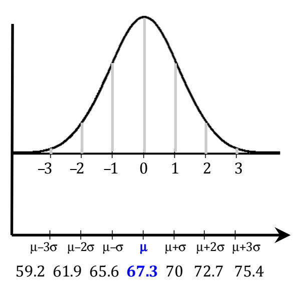
1. What is the probability that a randomly selected cadet of the group is taller than **70** inches? *The mean of this data is* ***67.3*** *inches and the standard deviation is* ***2.7*** *inches. That is* $X \sim N(\mathbf{67.3}, \mathbf{2.7})$.

Step 1	Label the *X*-axis with μ, label the *z*-axis with z_L, shade the appropriate region	**Step 3**
Identify the variable values $X_L =$ $\sigma =$ $\mu =$		Use the *z*-table or statistics calculator to find the associated area under the normal curve
Step 2		**Step 4**
Find $z_L = \frac{X_L - \mu}{\sigma}$		Find the desired probability $P(X > 70)$ $= P(z >\qquad)$ $=$

2. What percent of the group is shorter than **68** inches? $X \sim N(\mathbf{67.3}, \mathbf{2.7})$

<table>
<tr><th>Step 1</th><td rowspan="4">Label the X-axis with μ, label the z-axis with z_U, shade the appropriate region
</td><th>Step 3</th></tr>
<tr><td>Identify the variable values
$X_U =$ $\sigma =$
$\mu =$</td><td>Use the z-table or statistics calculator to find the associated area under the normal curve</td></tr>
<tr><th>Step 2</th><th>Step 4</th></tr>
<tr><td>Find $z_U = \frac{X_U - \mu}{\sigma}$</td><td>Find the desired probability
$P(X < 68)$
$= P(z <$ $)$
$=$</td></tr>
</table>

3. What are the heights of the cadets in the shortest 1% of this group? $X \sim N(\mathbf{67.3}, \mathbf{2.7})$

<table>
<tr><th>Step 1</th><td rowspan="4">Label the X-axis with μ, shade the appropriate region, label the z-axis with z_A,
</td><th>Step 3</th></tr>
<tr><td>Identify the variable values
$A =$ $\sigma =$
$\mu =$</td><td rowspan="3">Use the formula, $z_A = \frac{X - \mu}{\sigma}$ to solve for the desired height.</td></tr>
<tr><th>Step 2</th></tr>
<tr><td>Use the z-table or statistics calculator to find the z_A-value:</td></tr>
</table>

4. What is the shortest height in the tallest 5% of this group? $X \sim N(\mathbf{67.3}, \mathbf{2.7})$

Step 1

Identify the variable values

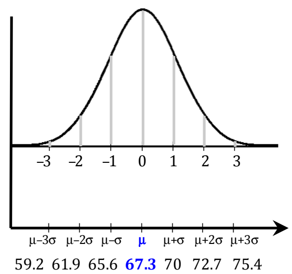
$A =$ $\sigma =$

$\mu =$

Step 2

Use the z-table or statistics calculator to find the z_A-value:

Label the X-axis with μ, shade the appropriate region, label the z-axis with z_A,



Step 3

Use the formula, $z_A = \dfrac{X - \mu}{\sigma}$ to solve for the desired height.

Additional problems:

1. Population in U.S. Jails The average daily jail population in the United States is 706,242. If the distribution is normal and the standard deviation is 52,145, find the probability that on a randomly selected day, the jail population is between 600,000 and 700,000. Note that $X \sim N(\mathbf{706{,}252}, \mathbf{52{,}145})$.

Step 1

Identify the variable values

$X_L =$ $\mu =$

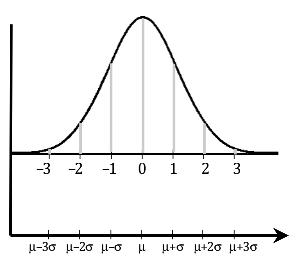
$X_U =$ $\sigma =$

Step 2

Find

$$z_L = \frac{X_L - \mu}{\sigma} \text{ and } z_U = \frac{X_U - \mu}{\sigma}$$

Label the X-axis with μ, label the z-axis with z_L, z_U, shade the appropriate region



Step 3

Use the z-table or a statistics calculator to find the associated area under the normal curve

Step 4

Find the desired probability

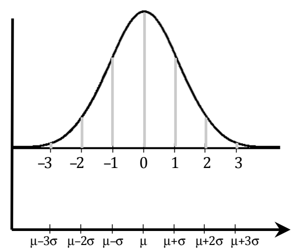
$P(600{,}000 < X < 700{,}000)$

$= P(\qquad < z < \qquad)$

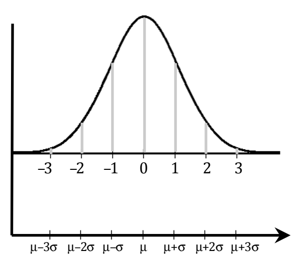
$=$

2. SAT Scores The national average SAT score (for Verbal and Math) is 1028. Assume a normal distribution with σ = 92. Note that $X \sim N(\mathbf{1028}, \mathbf{92})$.

(a) What is the 90th percentile score?

<table>
<tr><td>Step 1
Identify the variable values
$A =$
$\mu =$ $\sigma =$</td><td rowspan="2">Label the X-axis with μ, shade the appropriate region, label the z-axis with z_A,

</td><td rowspan="2">Step 3
Use the formula, $z_A = \dfrac{X - \mu}{\sigma}$ to solve for the desired score.</td></tr>
<tr><td>Step 2
Use the z-table or statistics calculator to find the z_A-value:</td></tr>
</table>

(b) What is the probability that a randomly selected score exceeds 1200?

<table>
<tr><td>Step 1
Identify the variable values
$X_L =$
$\mu =$ $\sigma =$</td><td rowspan="2">Label the X-axis with μ, label the z-axis with z_L, shade the appropriate region

</td><td>Step 3
Use the z-table or statistics calculator to find the associated area under the normal curve</td></tr>
<tr><td>Step 2
Find $z_L = \dfrac{X_L - \mu}{\sigma}$</td><td>Step 4
Find the desired probability
$P(X > 1{,}200)$
$= P(z > \qquad)$
$=$</td></tr>
</table>

3. Miles Driven Annually The mean number of miles driven per vehicle annually in the United States is 12,494. Assume the distribution is normal with a standard deviation of 1290. Note that $X \sim N(\mathbf{12{,}494}, \mathbf{1290})$. For a randomly chosen vehicle, what is the probability that the vehicle was

(a) driven more than 15,000 miles?

Step 1

Identify the variable values

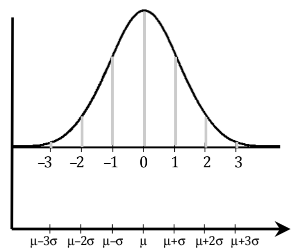
$X_L =$

$\mu =$ $\sigma =$

Step 2

Find $z_L = \dfrac{X_L - \mu}{\sigma}$

Label the X-axis with μ, label the z-axis with z_L, shade the appropriate region



Step 3

Use the z-table or statistics calculator to find the associated area under the normal curve

Step 4

Find the desired probability

$P(X > 15{,}000)$

$= P(z >$ $)$

$=$

(b) driven less than 8000 miles?

Step 1

Identify the variable values

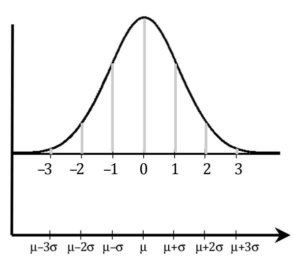
$X_U =$

$\mu =$ $\sigma =$

Step 2

Find $z_U = \dfrac{X_U - \mu}{\sigma}$

Label the X-axis with μ, label the z-axis with z_U, shade the appropriate region



Step 3

Use the z-table or statistics calculator to find the associated area under the normal curve

Step 4

Find the desired probability

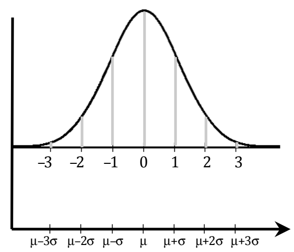
$P(X < 8{,}000)$

$= P(z <$ $)$

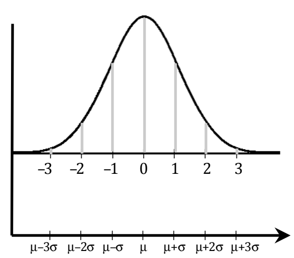
$=$

(c) Would you be comfortable with buying a vehicle for which you had been told it had been driven less than 6000 miles in the last year?

4. New Home Size A contractor has decided to build homes that will include the middle 80% of the market If the average size of homes built is 1810 square feet, find the maximum and minimum sizes of the homes the contractor should build. Assume that the data is normally distributed and the standard deviation is 92 square feet. Note that $X \sim N(\mathbf{1810}, \mathbf{92})$.

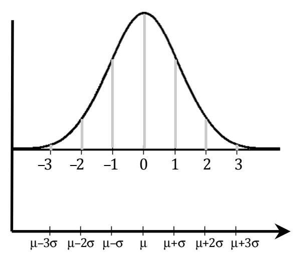
<table>
<tr><th>Step 1</th><td rowspan="4">Label the X-axis with μ, shade the appropriate region, label the z-axis with z_L, and z_U
</td><th>Step 3</th></tr>
<tr><td>Identify the variable values
$A =$
$\mu =$ $\sigma =$</td><td rowspan="3">Use the formulas,
$z_L = \dfrac{X_L - \mu}{\sigma}$ and $z_U = \dfrac{X_U - \mu}{\sigma}$
to solve for the desired minimum and maximum sizes.</td></tr>
<tr><th>Step 2</th></tr>
<tr><td>Use the z-table or statistics calculator to find the lower and upper z_A-values</td></tr>
</table>

5. Reading Improvement Program To help students improve their reading skills, a school district decides to implement a reading program. It is to be administered to the bottom 5% of the students in the district, based on the scores on a reading achievement exam. If the average score for the students in the district is 122.6, find the cutoff score that will make a student eligible for the program. The standard deviation is 18. Assume the data is normally distributed.
Note that $X \sim N(\mathbf{122.6}, \mathbf{18})$.

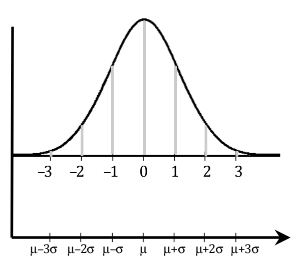
<table>
<tr><th>Step 1</th><td rowspan="4">Label the X-axis with μ, shade the appropriate region, label the z-axis with z_A,
</td><th>Step 3</th></tr>
<tr><td>Identify the variable values
$A =$
$\mu =$ $\sigma =$</td><td rowspan="3">Use the formula, $z_A = \dfrac{X - \mu}{\sigma}$ to solve for the desired cutoff score.</td></tr>
<tr><th>Step 2</th></tr>
<tr><td>Use the z-table or statistics calculator to find the z_A-value:</td></tr>
</table>

6. Cost of Personal Computers The average price of a personal computer is \$949. Assume the computer prices are approximately normally distributed and σ = \$100. Note that $X \sim N(\mathbf{949}, \mathbf{100})$.

(a) What is the probability that a randomly selected PC costs more than \$1200?

Step 1	Label the X-axis with μ, label the z-axis with z_U, shade the appropriate region	**Step 3**
Identify the variable values $X_L =$ $\mu =$ $\sigma =$		Use the z-table or statistics calculator to find the associated area under the normal curve
Step 2		**Step 4**
Find $z_L = \dfrac{X_L - \mu}{\sigma}$		Find the desired probability $P(X > 1{,}200)$ $= P(z >$ $)$ $=$

(b) The least expensive 10% of personal computers cost less than what amount?

Step 1	Label the X-axis with μ, shade the appropriate region, label the z-axis with z_A,	**Step 3**
Identify the variable values $A =$ $\mu =$ $\sigma =$		Use the formula, $z_A = \dfrac{X - \mu}{\sigma}$ to solve for the desired computer cost.
Step 2		
Use the z-table or statistics calculator to find the z_A-value:		