Lesson 100: The Normal Distribution

HL Math - Santowski

Objectives

- Introduce the Normal Distribution
- Properties of the Standard Normal Distribution
- Introduce the Central Limit Theorem

Normal Distributions

- A random variable **X** with mean μ and standard deviation σ is normally distributed if its probability density function is given by

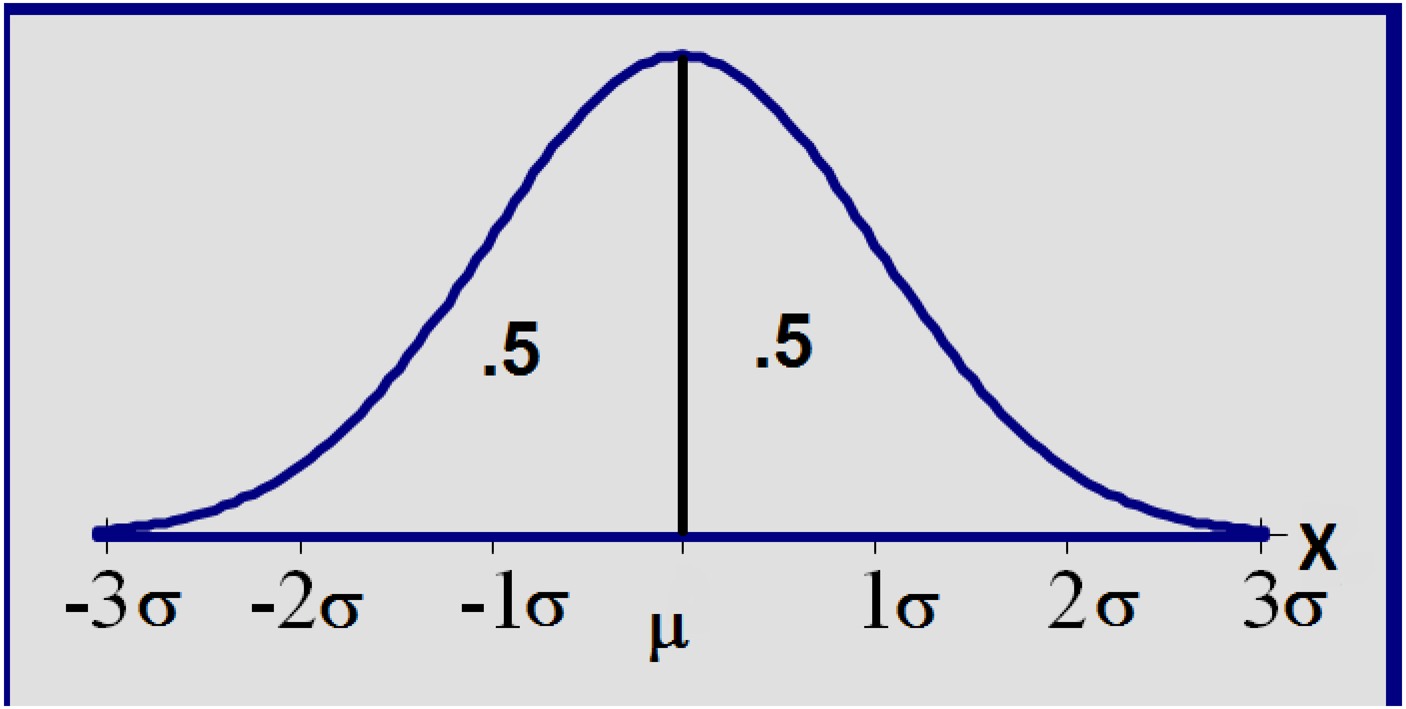
$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-(1/2)\left(\frac{x-\mu}{\sigma}\right)^2} \qquad -\infty \le x \le \infty$$

$$\text{where } \pi = 3.14159\ldots \text{ and } e = 2.71828\ldots$$

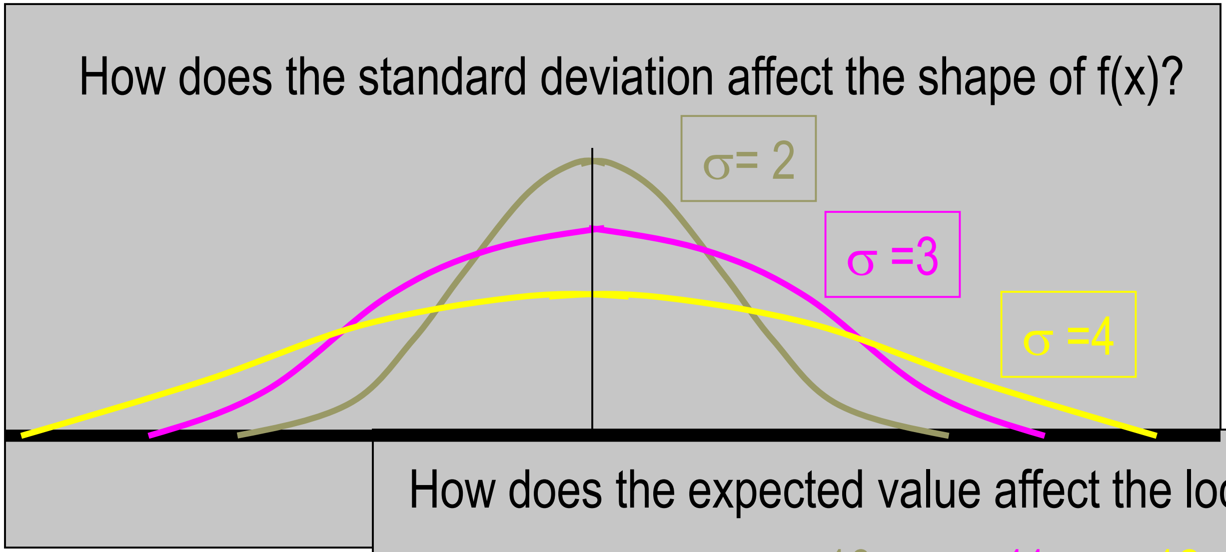
Normal Probability Distributions

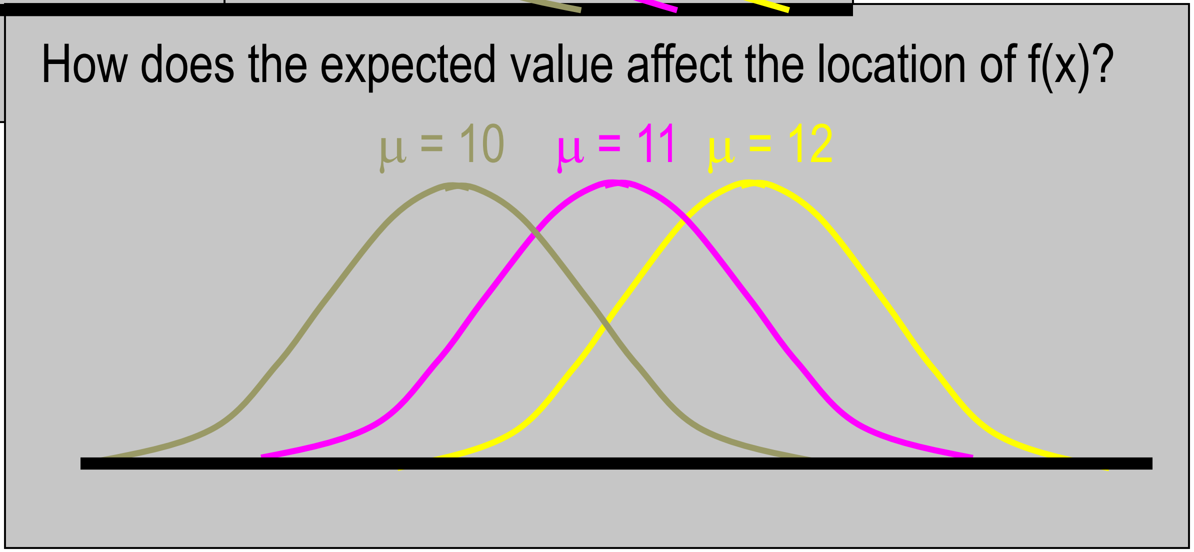
- The expected value (also called the mean) E(X) (or μ) can be any number
- The standard deviation σ can be any nonnegative number
- The total area under every normal curve is 1
- There are infinitely many normal distributions

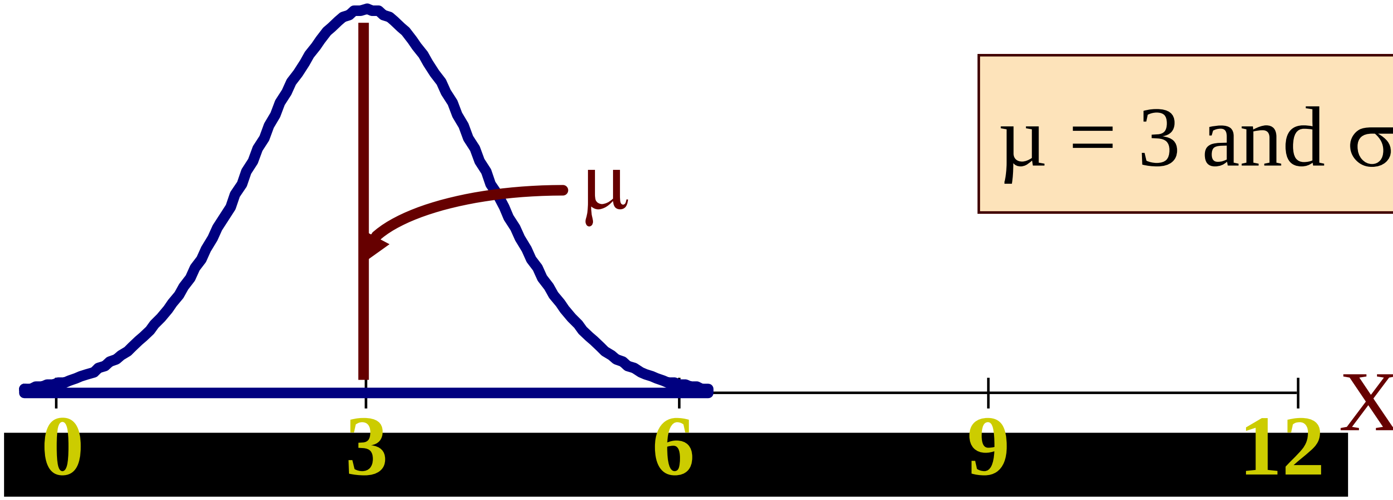
Total area =1; symmetric around μ



The effects of μ and σ





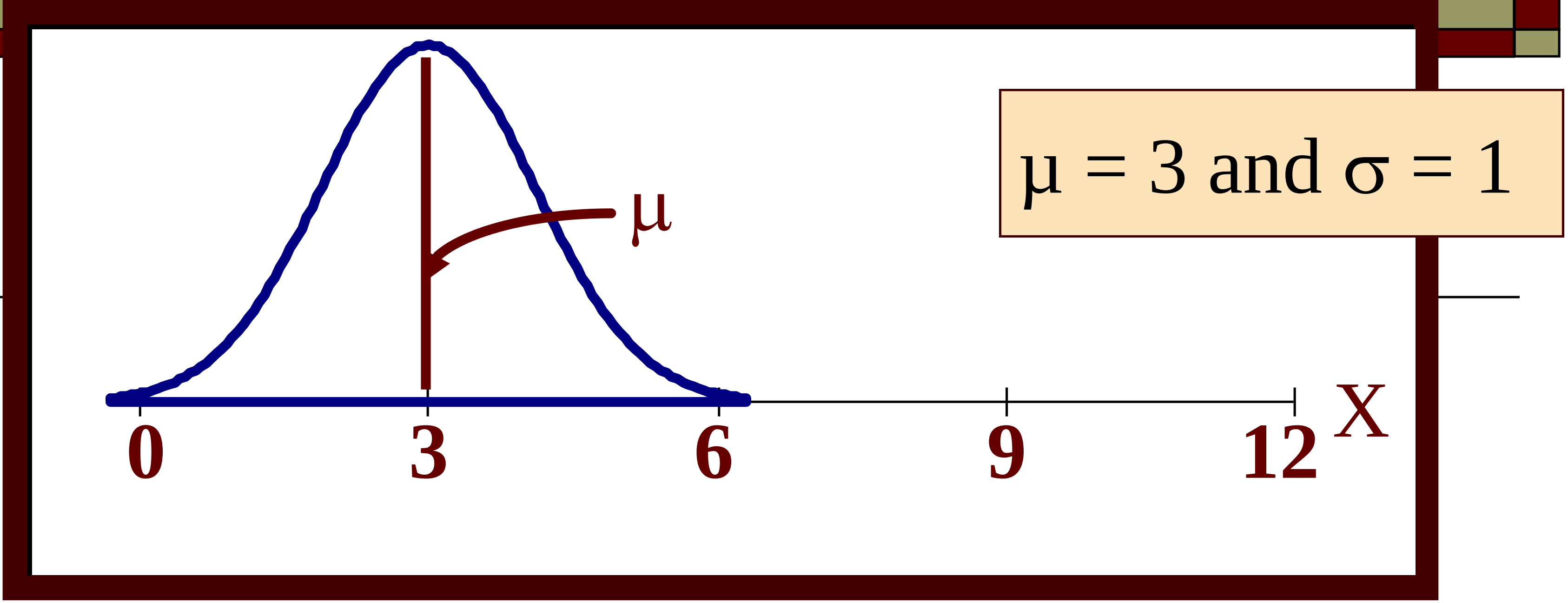


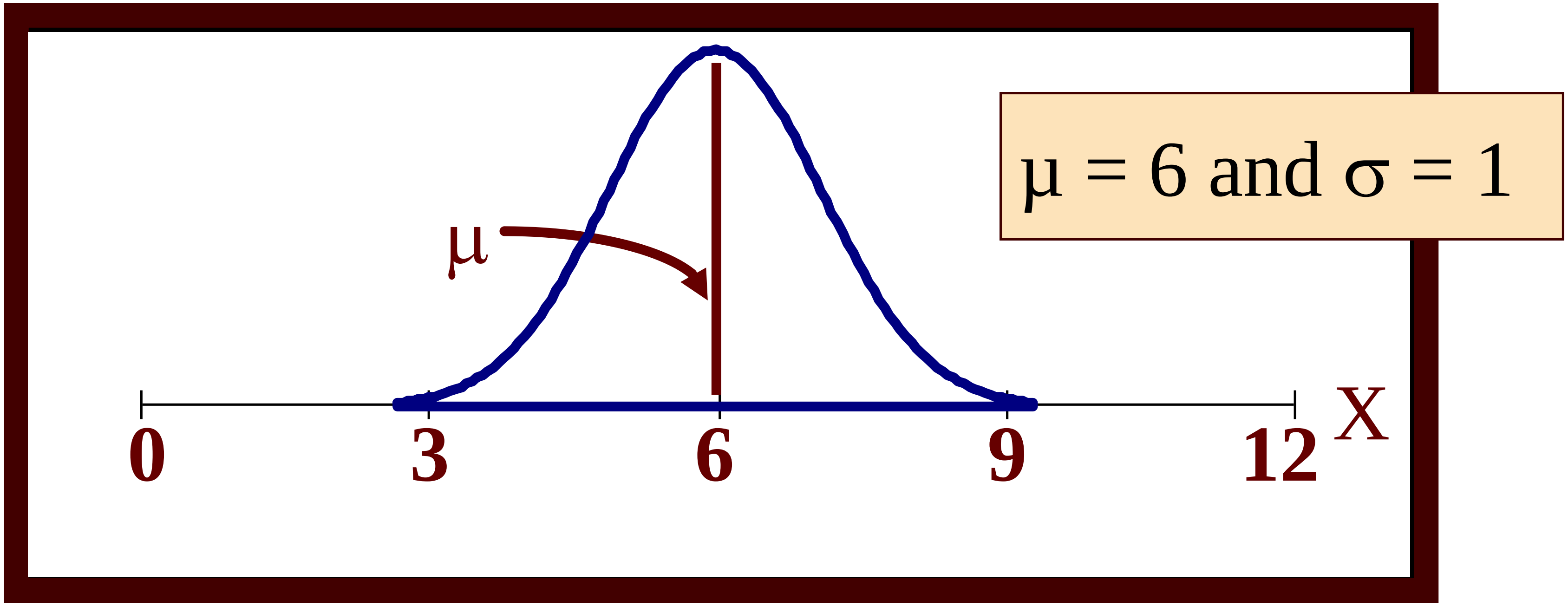
$\mu = 3$ and $\sigma = 1$

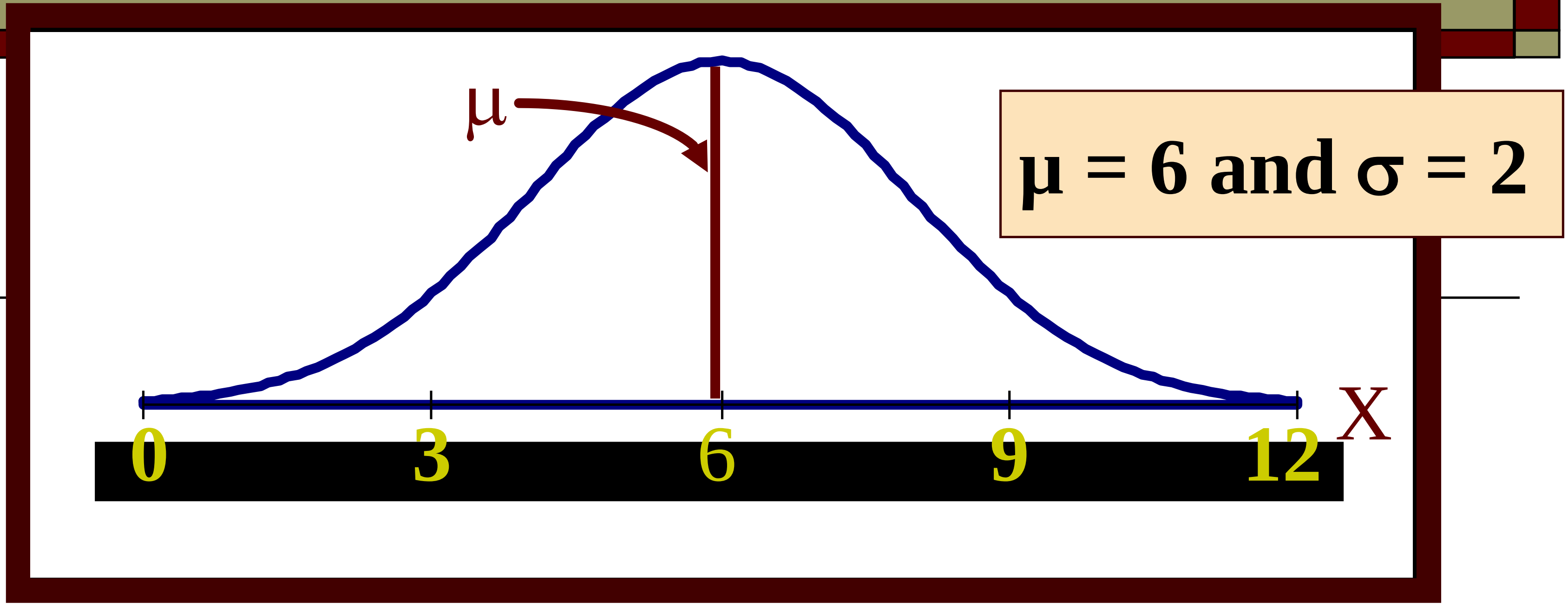
A family of bell-shaped curves that differ only in their means and standard deviations.

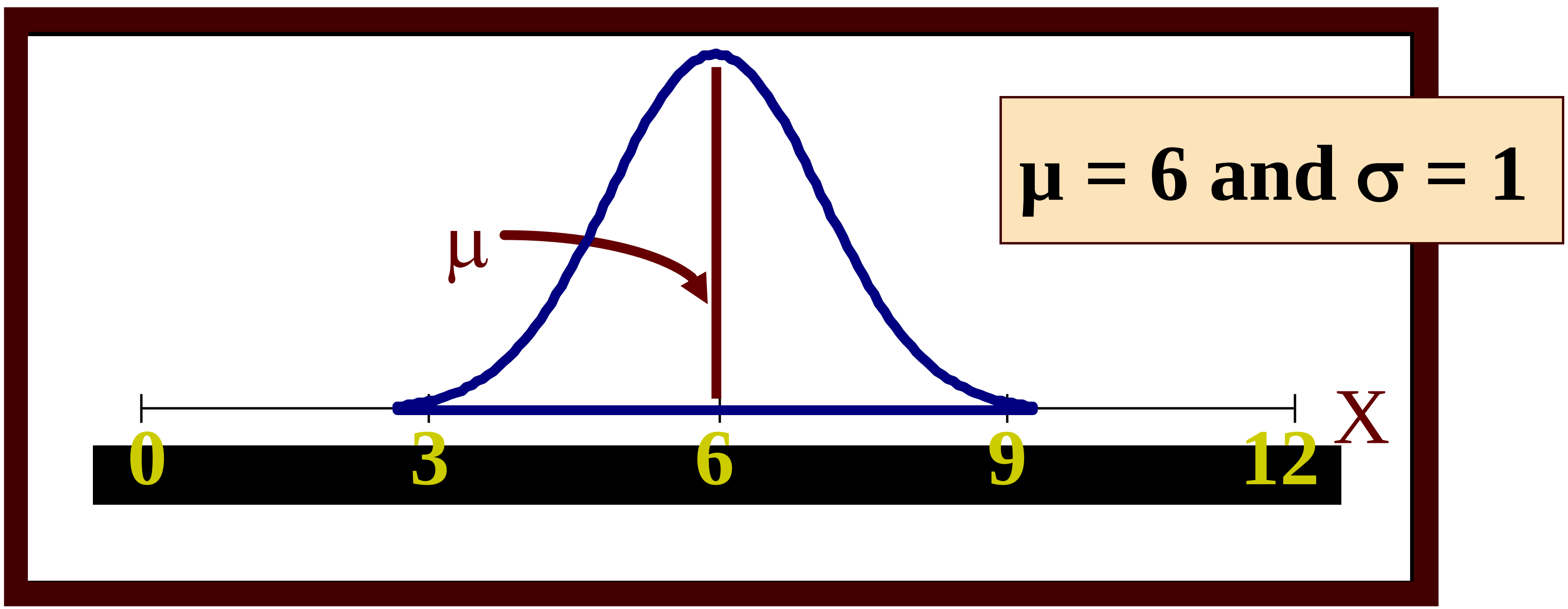
μ = the mean of the distribution

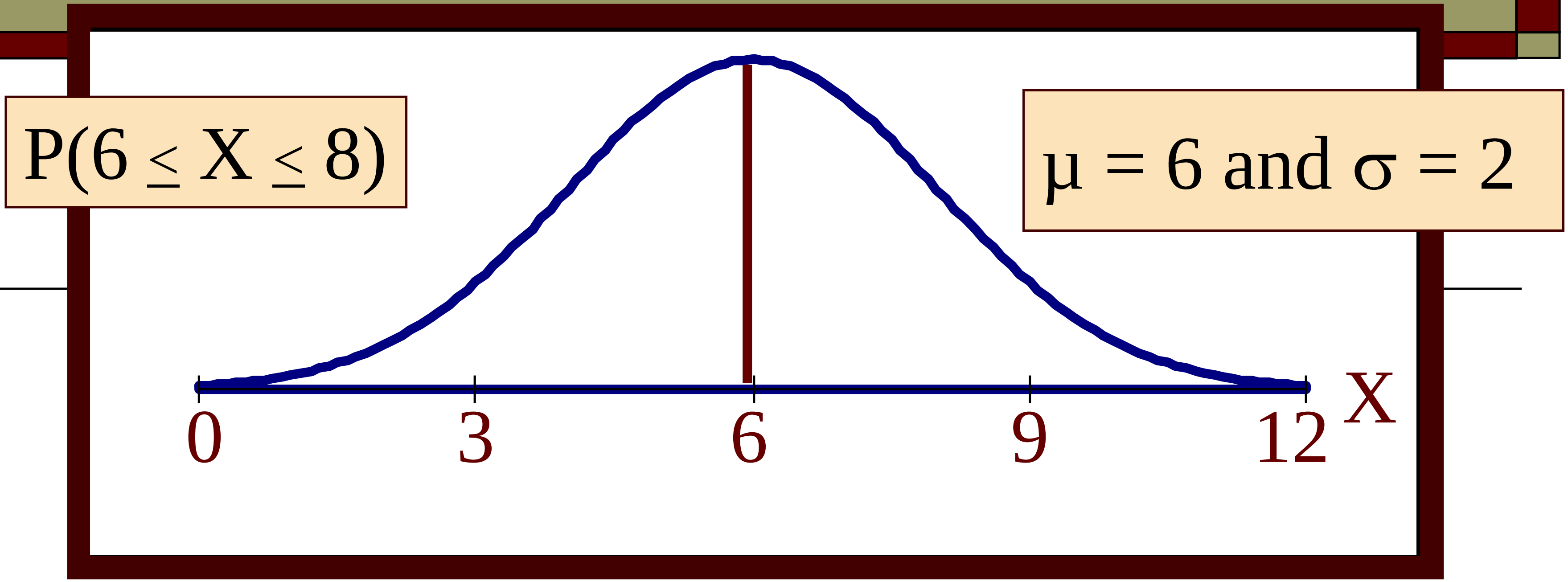
σ = the standard deviation


μ = 3 and σ = 1
μ
0
3
6
9
12
X


μ = 6 and σ = 1
μ
0
3
6
9
12
X


μ
μ = 6 and σ = 2
0
3
6
9
12
X


μ
μ = 6 and σ = 1
0
3
6
9
12
X

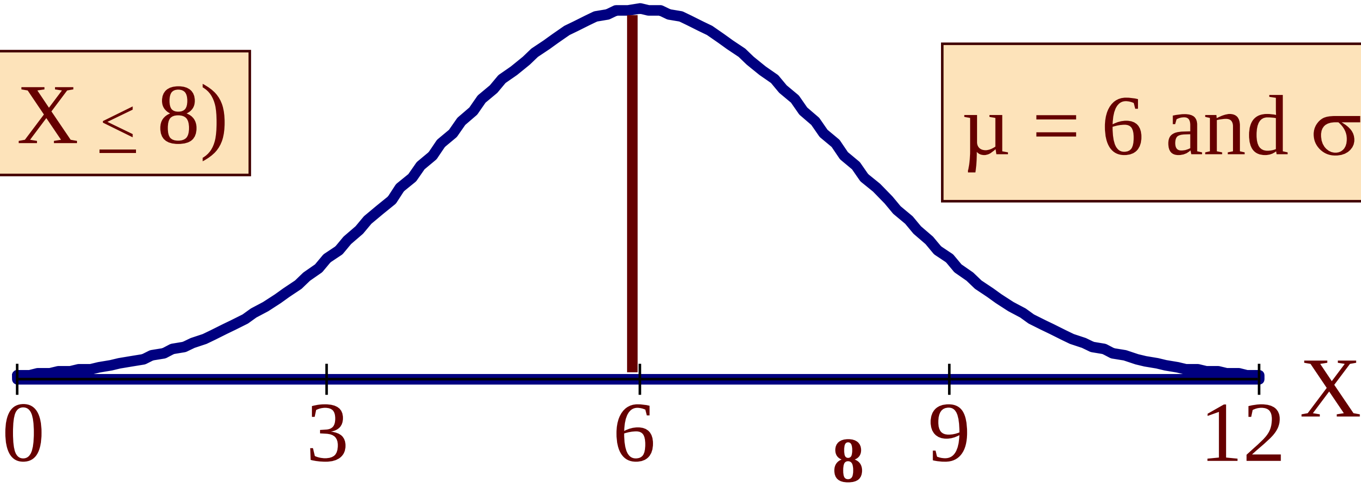


Probability = area under the density curve

P($a \le X \le b$ = area under the density curve between a and b .

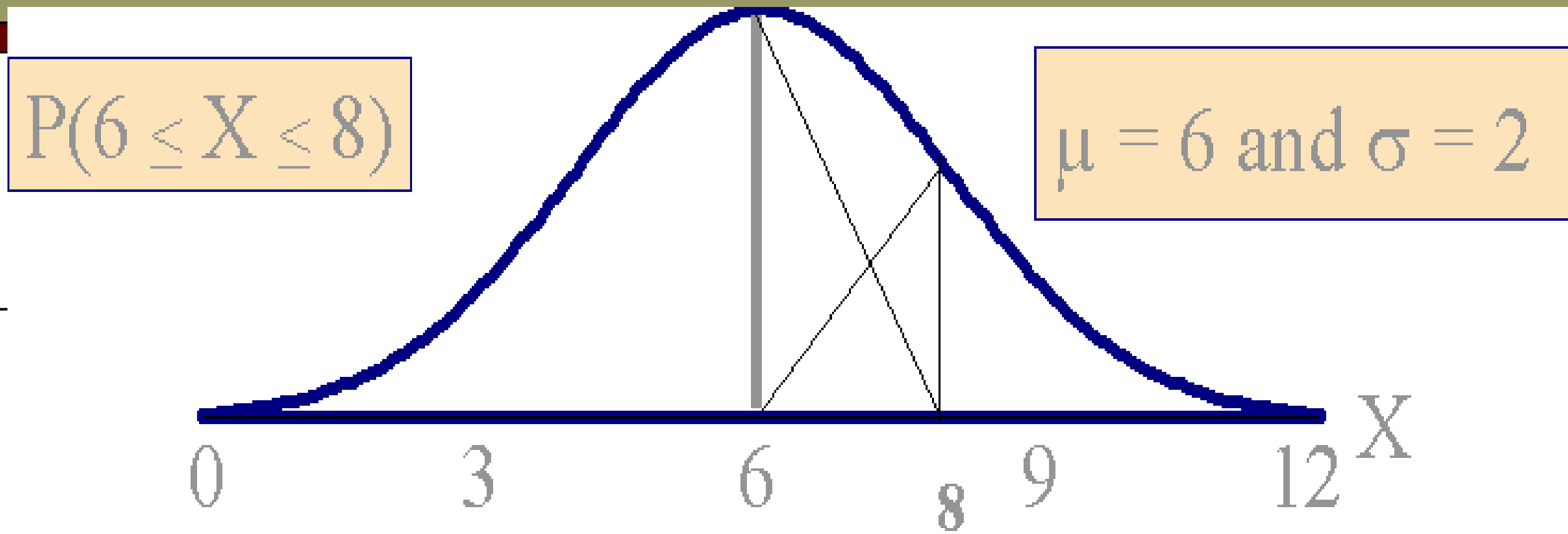
$P(6 \le X \le 8)$

$\mu = 6$ and $\sigma = 2$



Probability = area under the density curve

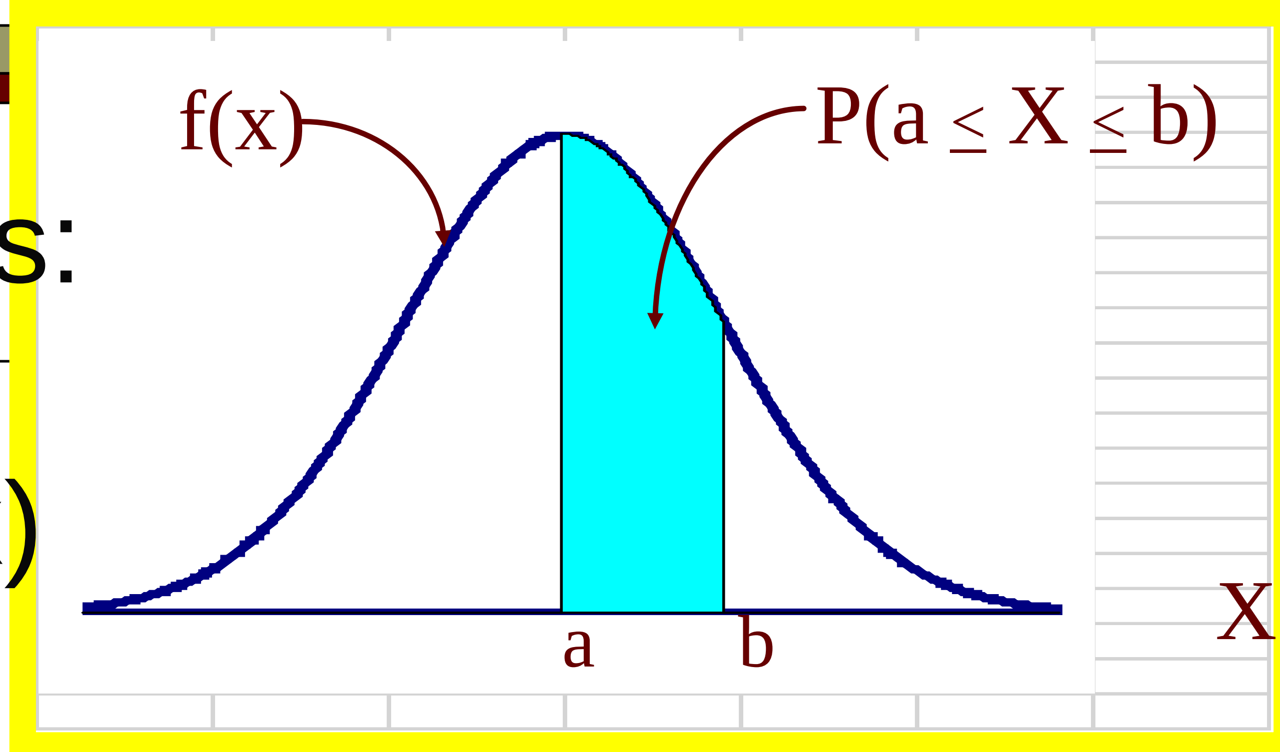
$P(\overset{6}{\cancel{a}} \le X \le \overset{8}{\cancel{b}})$ = area under the density curve between $\underset{6}{\cancel{a}}$ and $\underset{8}{\cancel{b}}$



Probability = area under the density curve

$P(\overset{6}{\cancel{a}} \le X \le \overset{8}{\cancel{b}})$ = area under the density curve between $\underset{6}{\cancel{a}}$ and $\underset{8}{\cancel{b}}$

Probabilities: area under graph of f(x)



$P(a \le X \le b)$ = area under the density curve between a and b.

$P(X=a) = 0$

$P(a \le x \le b) = P(a < x < b)$

$$P(a \le X \le b) = \int_a^b f(x)dx$$

The Normal Distribution: as mathematical function (pdf)

$$f(x) = \frac{1}{\boxed{\sigma}\sqrt{2\pi}} \cdot e^{-\frac{1}{2}(\frac{x-\boxed{\mu}}{\sigma})^2}$$

This is a bell shaped curve with different centers and spreads depending on μ and σ

<u>Note constants:</u>

π=3.14159

e=2.71828

The Normal PDF

It's a probability function, so no matter what the values of μ and σ, must integrate to 1!

$$\int_{-\infty}^{+\infty} \frac{1}{\sigma\sqrt{2\pi}} \cdot e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} dx = 1$$

Normal distribution is defined by its mean and standard dev.

E(X)=μ = $\int_{-\infty}^{+\infty} x \frac{1}{\sigma\sqrt{2\pi}} \cdot e^{-\frac{1}{2}(\frac{x-\mu}{\sigma})^2} dx$

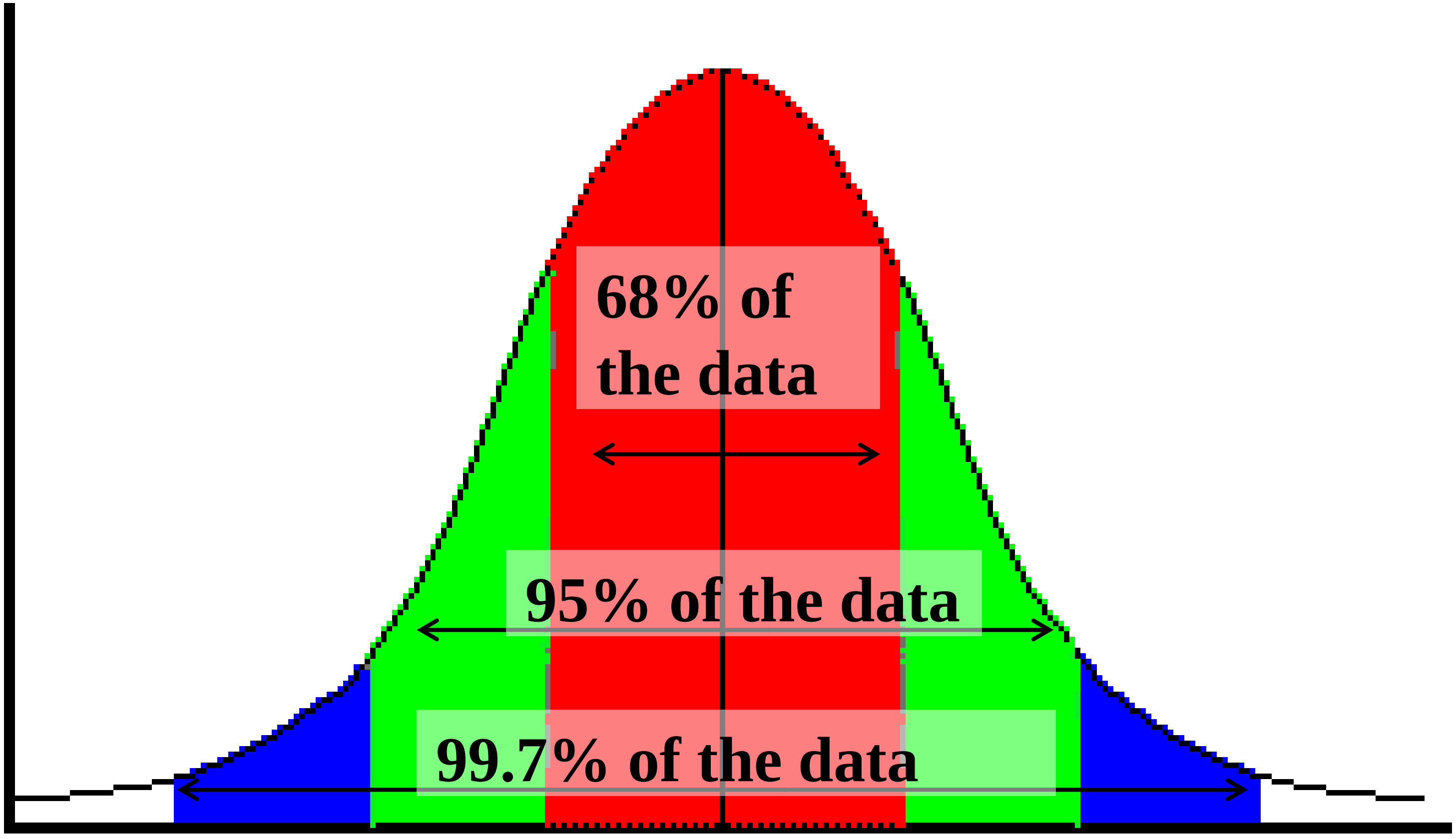
Var(X)=σ² = $\int_{-\infty}^{+\infty} x^2 \frac{1}{\sigma\sqrt{2\pi}} \cdot e^{-\frac{1}{2}(\frac{x-\mu}{\sigma})^2} dx) - \mu^2$

Standard Deviation(X)=σ

**The beauty of the normal curve:

No matter what μ and σ are, the area between $\mu-\sigma$ and $\mu+\sigma$ is about 68%; the area between $\mu-2\sigma$ and $\mu+2\sigma$ is about 95%; and the area between $\mu-3\sigma$ and $\mu+3\sigma$ is about 99.7%. Almost all values fall within 3 standard deviations.

68-95-99.7 Rule



68-95-99.7 Rule in Math terms…

$$\int_{\mu-\sigma}^{\mu+\sigma} \frac{1}{\sigma\sqrt{2\pi}} \bullet e^{-\frac{1}{2}(\frac{x-\mu}{\sigma})^2} dx = .68$$

$$\int_{\mu-2\sigma}^{\mu+2\sigma} \frac{1}{\sigma\sqrt{2\pi}} \bullet e^{-\frac{1}{2}(\frac{x-\mu}{\sigma})^2} dx = .95$$

$$\int_{\mu-3\sigma}^{\mu+3\sigma} \frac{1}{\sigma\sqrt{2\pi}} \bullet e^{-\frac{1}{2}(\frac{x-\mu}{\sigma})^2} dx = .997$$

Standardizing

- Suppose $X \sim N(\mu, \sigma)$
- Form a new random variable by subtracting the mean μ from X and dividing by the standard deviation σ:

$$(X-\mu)/\sigma$$

- This process is called <u>standardizing</u> the random variable X.

Standardizing (cont.)

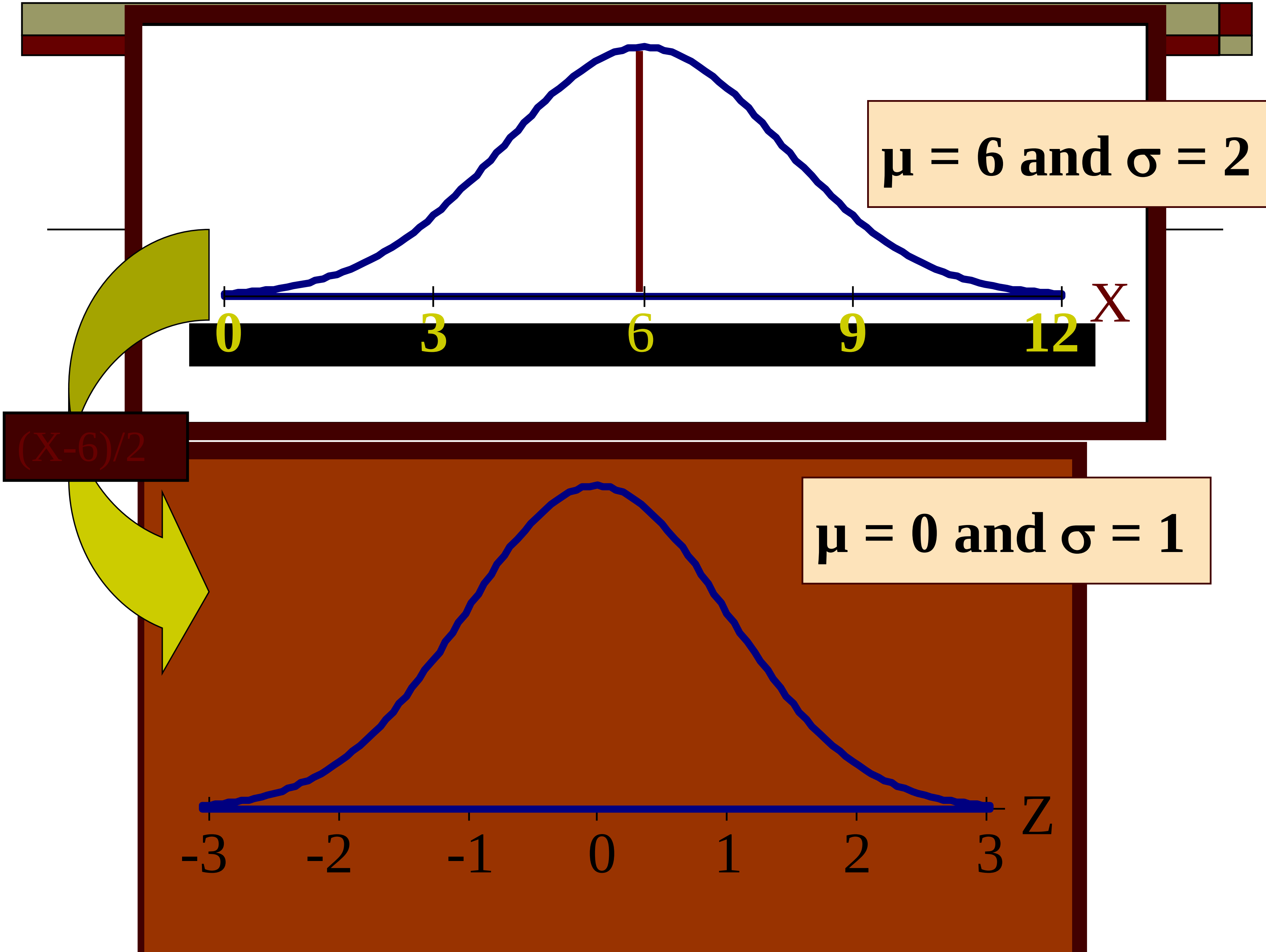
- $(X-\mu)/\sigma$ is also a normal random variable; we will denote it by Z:

$$Z = (X-\mu)/\sigma$$

- Z has mean 0 and standard deviation 1: $E(Z) = \mu = 0$; $SD(Z) = \sigma = 1$.

$$Z \sim N(0, 1)$$

- The probability distribution of Z is called the <u>standard normal distribution</u>.


μ = 6 and σ = 2
0
3
6
9
12
X
(X-6)/2
μ = 0 and σ = 1
-3
-2
-1
0
1
2
3
Z

Pdf of a standard normal rv

- A normal random variable x has the following pdf:

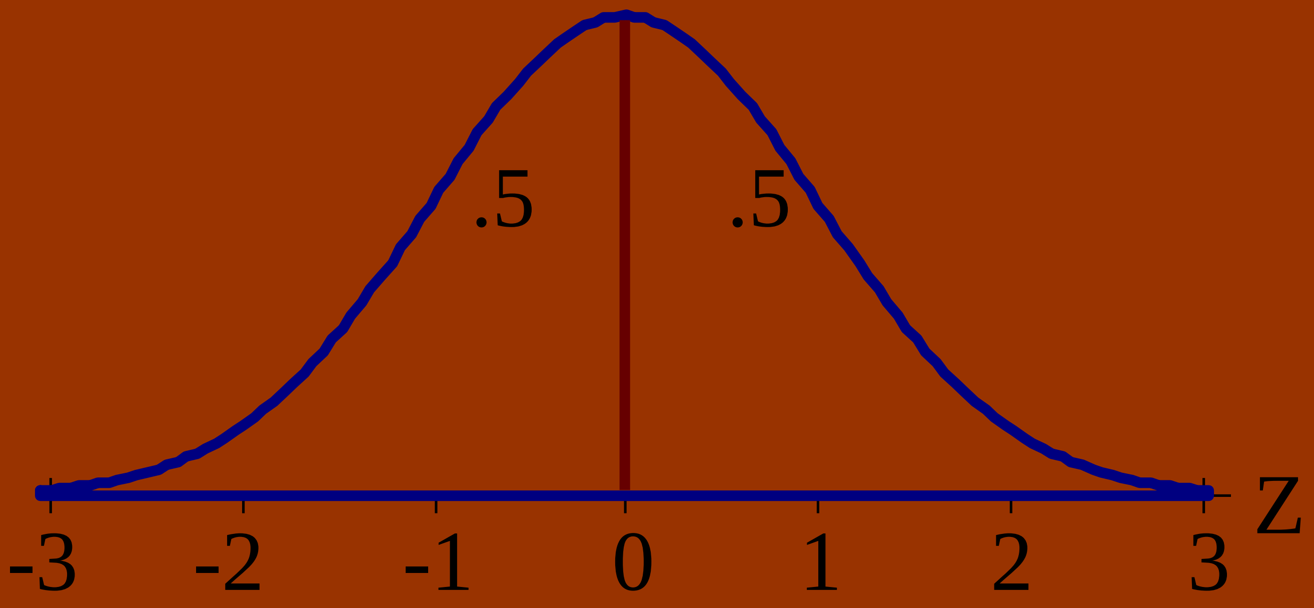
$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left[\frac{(x-\mu)^2}{\sigma^2}\right]}, -\infty < x < \infty$$

$Z \sim N(0,1)$ *substitute* 0 *for* μ *and* 1 *for* σ

pdf for the standard normal rv becomes

$$\phi(z) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}z^2}, \ -\infty < z < \infty$$

Standard Normal Distribution



Z = standard normal random variable

$\mu = 0$ and $\sigma = 1$

Important Properties of Z

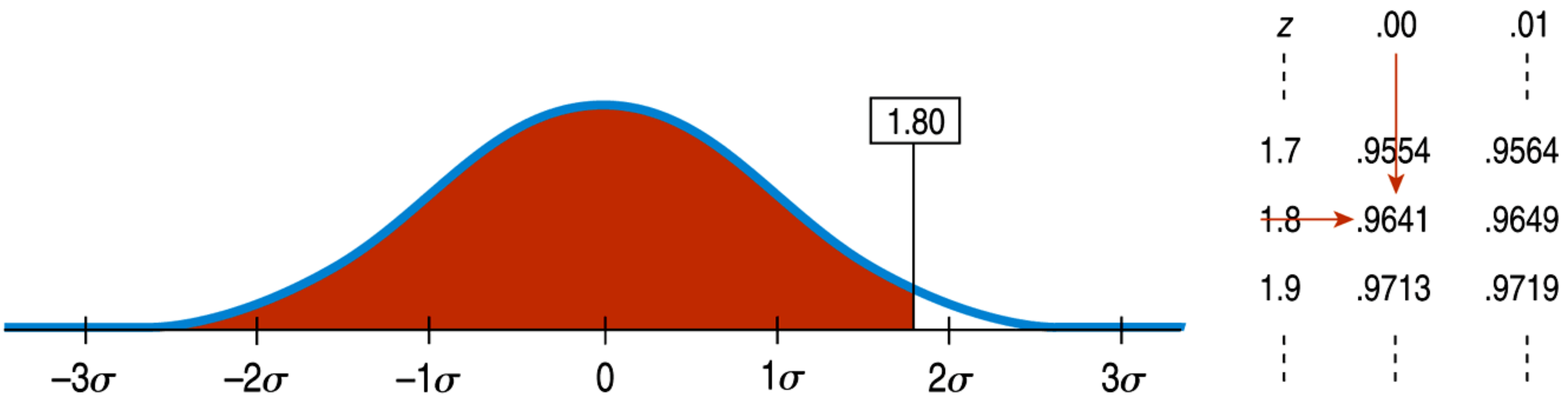
#1. The standard normal curve is symmetric around the mean 0

#2. The total area under the curve is 1;

so (from #1) the area to the left of 0 is 1/2, and the area to the right of 0 is 1/2

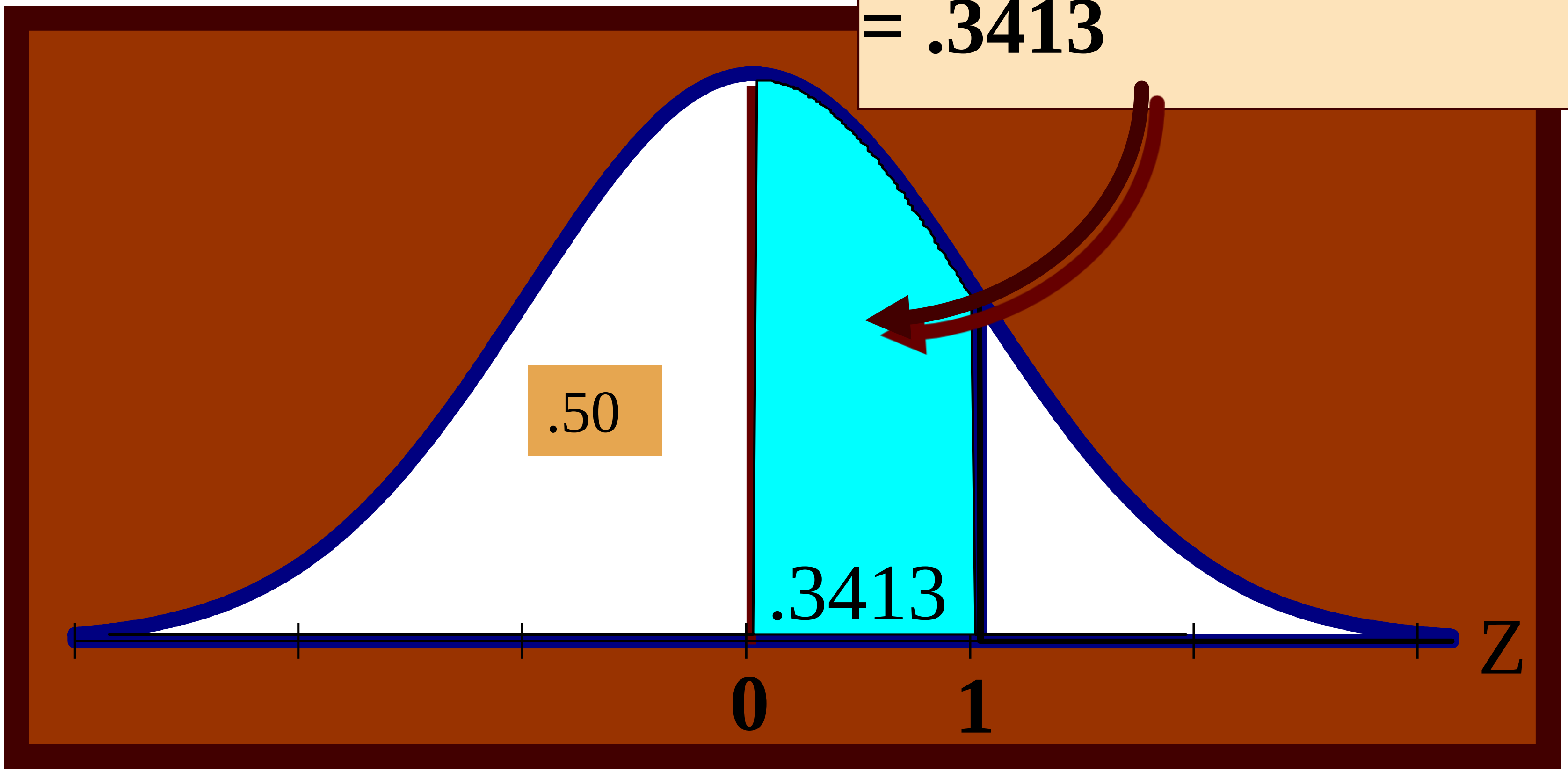
Finding Normal Percentiles by Hand (cont.)

- Table Z is the standard Normal table. We have to convert our data to z-scores before using the table.
- The figure shows us how to find the area to the left when we have a z-score of 1.80:



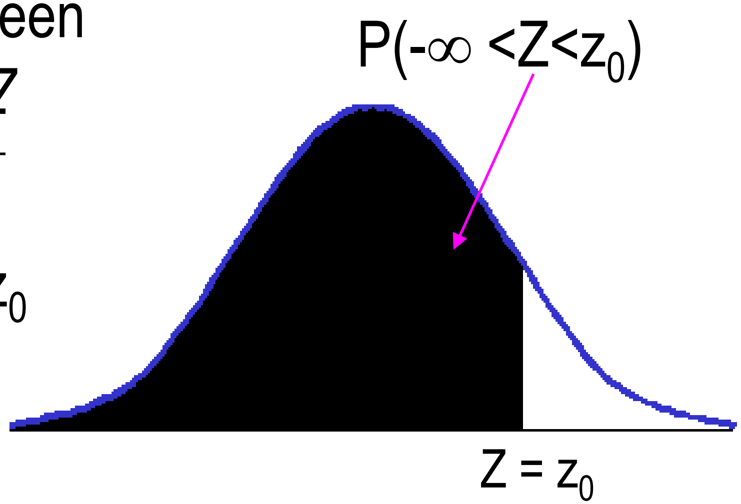
Areas Under the Z Curve: Using the Table

P(0 ≤ Z ≤ 1) = .8413 - .5 = .3413



Standard normal probabilities have been calculated and are provided in table *Z*

The tabulated probabilities correspond to the area between Z= -∞ and some z_0



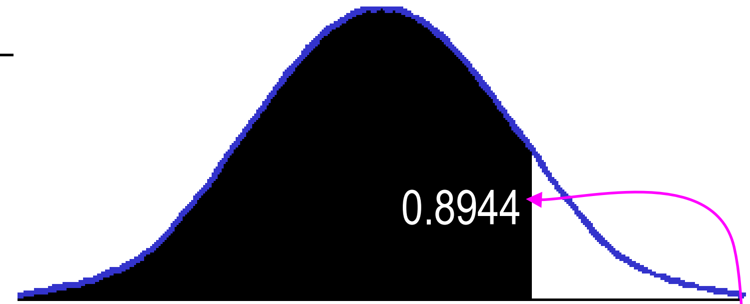
z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	**0.5000**	**0.5040**	**0.5080**	**0.5120**	**0.5160**	**0.5199**	**0.5239**	**0.5279**	**0.5319**	**0.5359**
0.1	**0.5398**	**0.5438**	**0.5478**	**0.5517**	**0.5557**	**0.5596**	**0.5636**	**0.5675**	**0.5714**	**0.5753**
0.2	**0.5793**	**0.5832**	**0.5871**	**0.5910**	**0.5948**	**0.5987**	**0.6026**	**0.6064**	**0.6103**	**0.6141**
⋮			⋮			⋮				⋮
1	**0.8413**	**0.8438**	**0.8461**	**0.8485**	**0.8508**	**0.8531**	**0.8554**	**0.8577**	**0.8599**	**0.8621**
⋮			⋮			⋮				⋮
1.2	**0.8849**	**0.8869**	**0.8888**	**0.8907**	**0.8925**	**0.8944**	**0.8962**	**0.8980**	**0.8997**	**0.9015**
⋮			⋮			⋮				⋮

Example – continued X~N(60, 8)

$$P(X < 70) = P\left(\frac{X-60}{8} < \frac{70-60}{8}\right)$$

$$= P(z < 1.25)$$

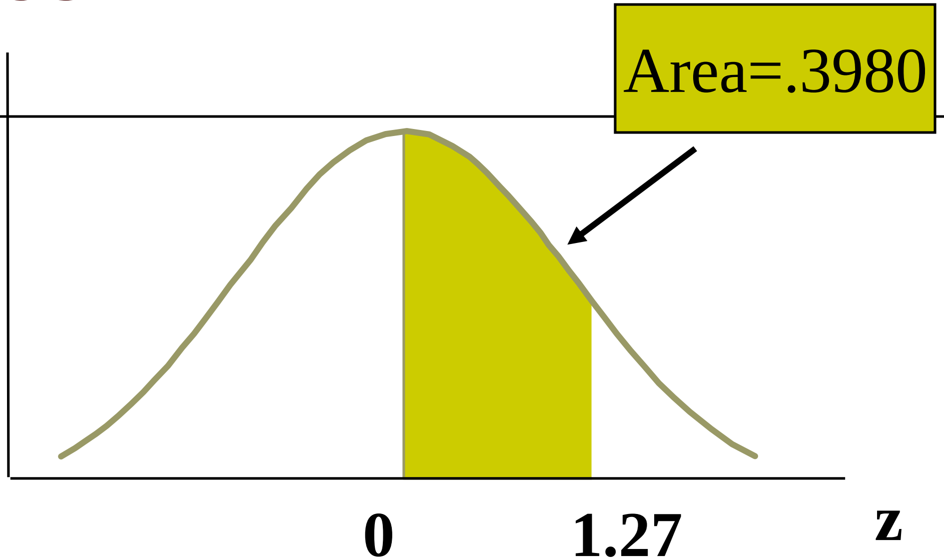
$P(z < 1.25) = 0.8944$



In this example $z_0 = 1.25$

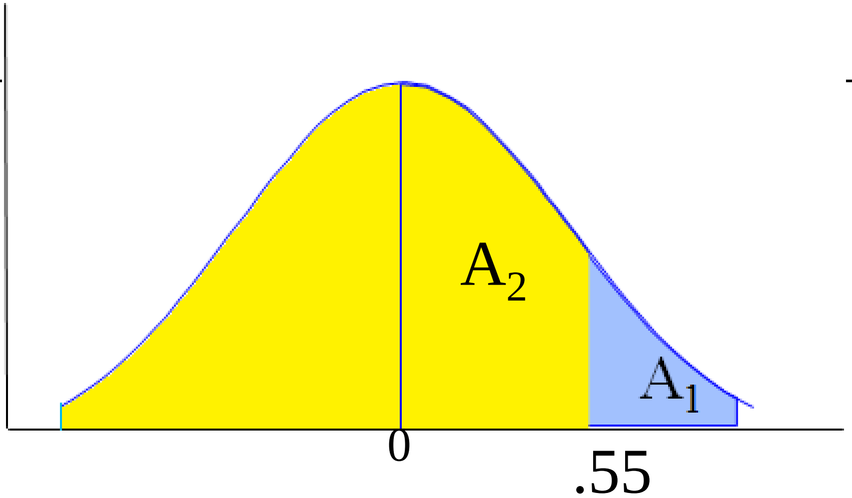
z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	**0.5000**	**0.5040**	**0.5080**	**0.5120**	**0.5160**	**0.5199**	**0.5239**	**0.5279**	**0.5319**	**0.5359**
0.1	**0.5398**	**0.5438**	**0.5478**	**0.5517**	**0.5557**	**0.5596**	**0.5636**	**0.5675**	**0.5714**	**0.5753**
0.2	**0.5793**	**0.5832**	**0.5871**	**0.5910**	**0.5948**	**0.5987**	**0.6026**	**0.6064**	**0.6103**	**0.6141**
⋮			⋮			⋮				⋮
1	**0.8413**	**0.8438**	**0.8461**	**0.8485**	**0.8508**	**0.8531**	**0.8554**	**0.8577**	**0.8599**	**0.8621**
⋮			⋮			⋮				⋮
1.2	**0.8849**	**0.8869**	**0.8888**	**0.8907**	**0.8925**	**0.8944**	**0.8962**	**0.8980**	**0.8997**	**0.9015**
⋮			⋮			⋮				⋮

Examples



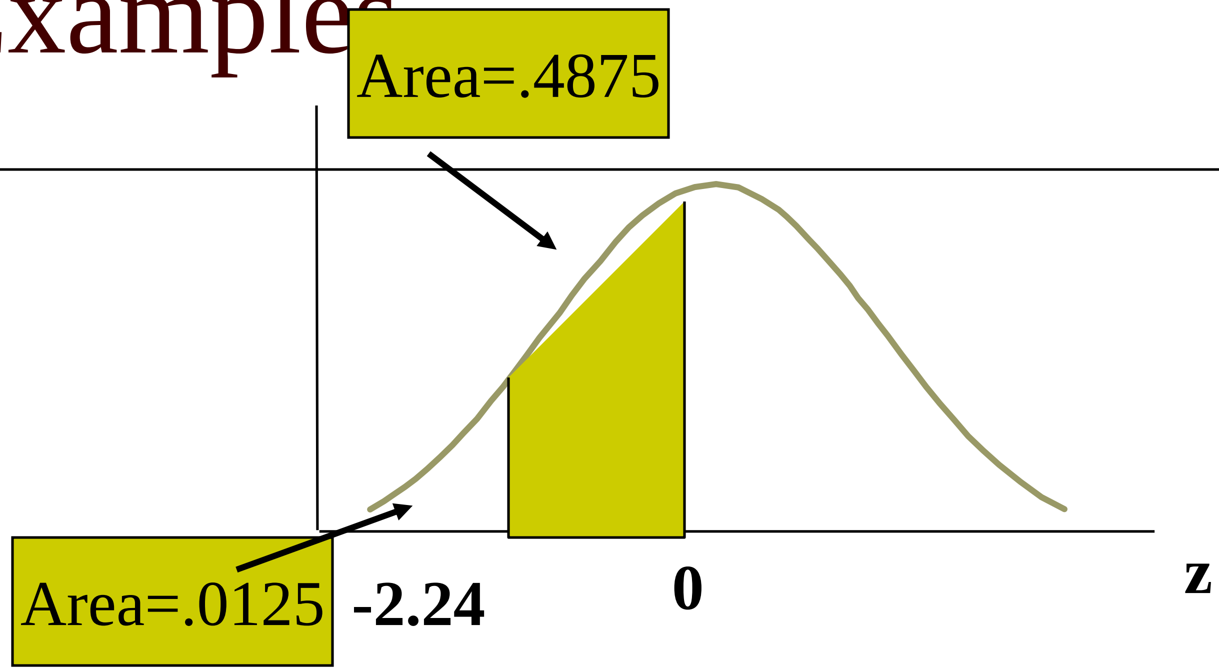
- $P(0 \le z \le 1.27) =$ **.8980-.5=.3980**

Examples



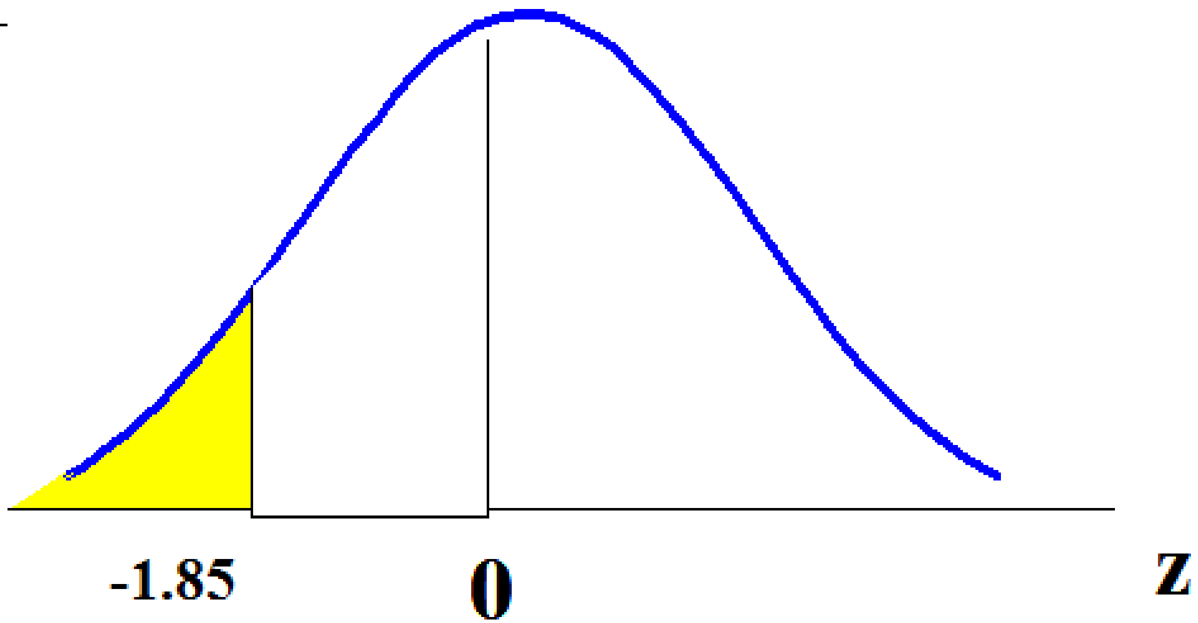
$$
\begin{aligned}
P(Z \geq .55) &= A_1 \\
&= 1 - A_2 \\
&= 1 - .7088 \\
&= .2912
\end{aligned}
$$

Examples



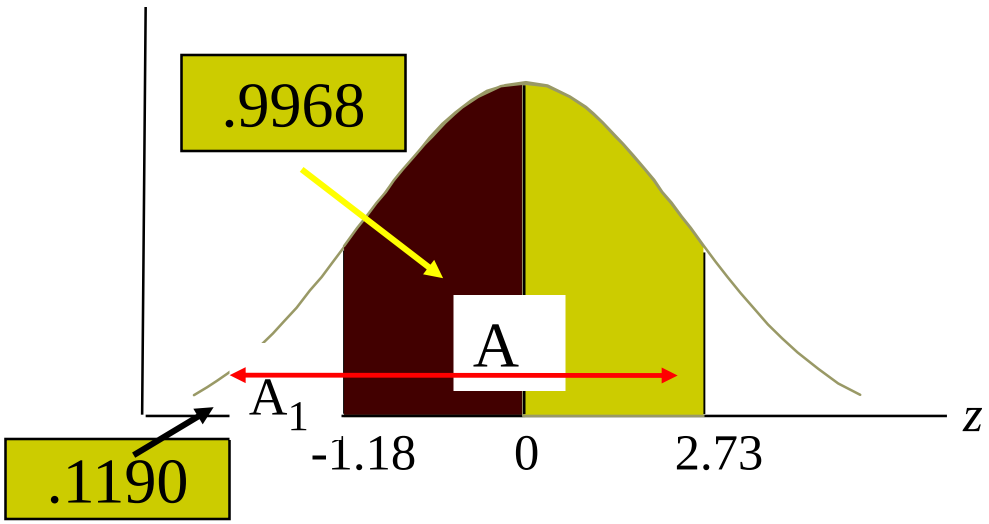
- $P(-2.24 \le z \le 0) =$ **.5 - .0125 = .4875**

Examples

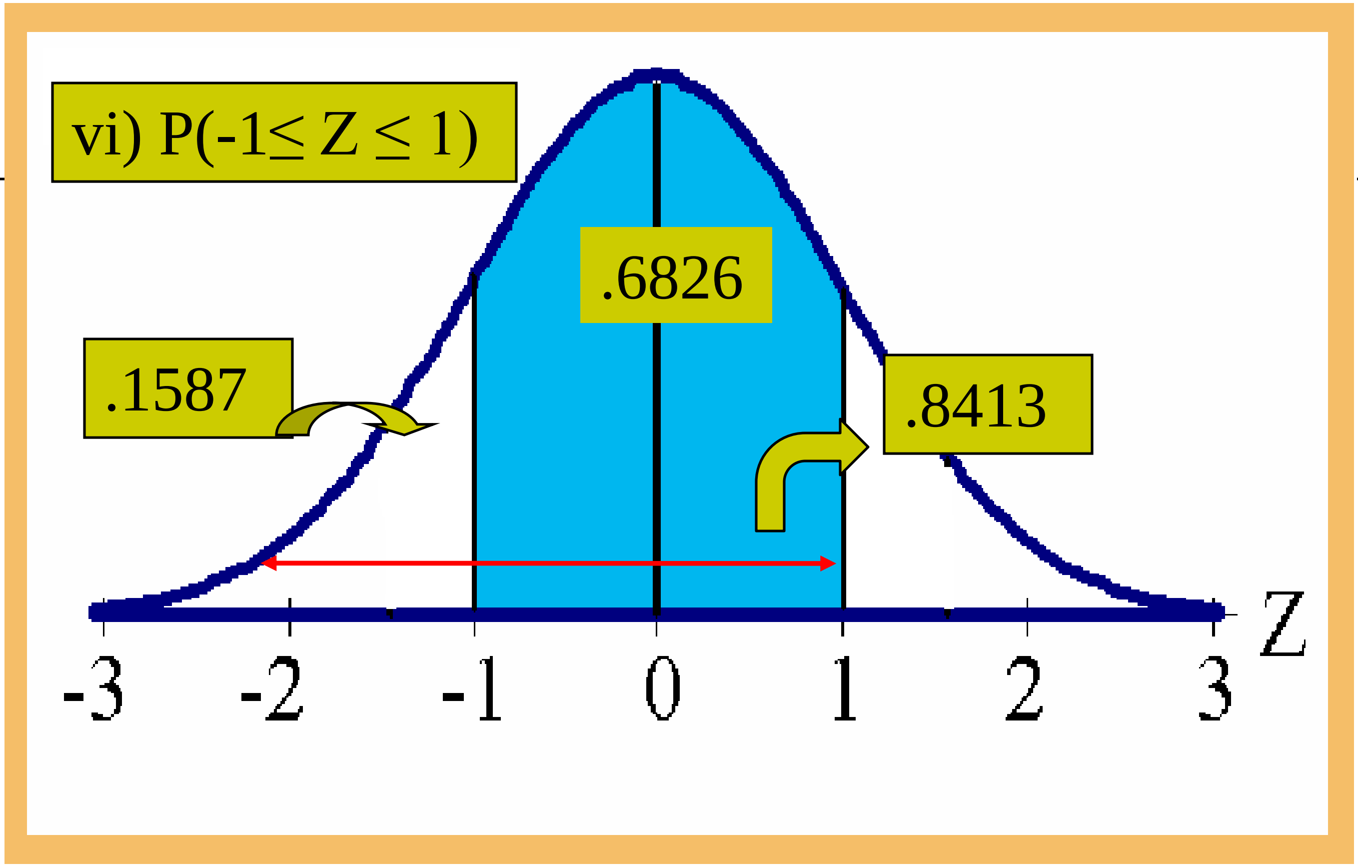


$P(z \leq -1.85) = .0322$

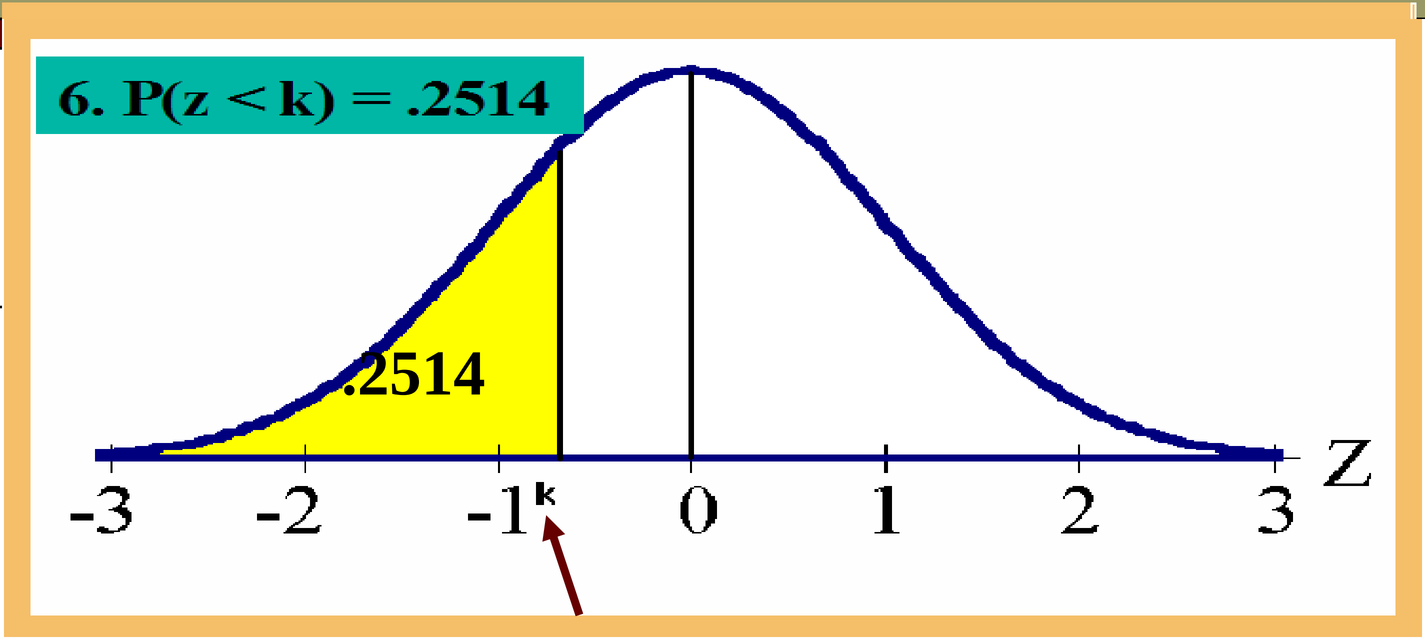
Examples (cont.)



- $P(-1.18 \le z \le 2.73)$ $= A - A_1$
- $= .9968 - .1190$
- $= .8778$



$$P(-1 \leq Z \leq 1) = .8413 - .1587 = .6826$$



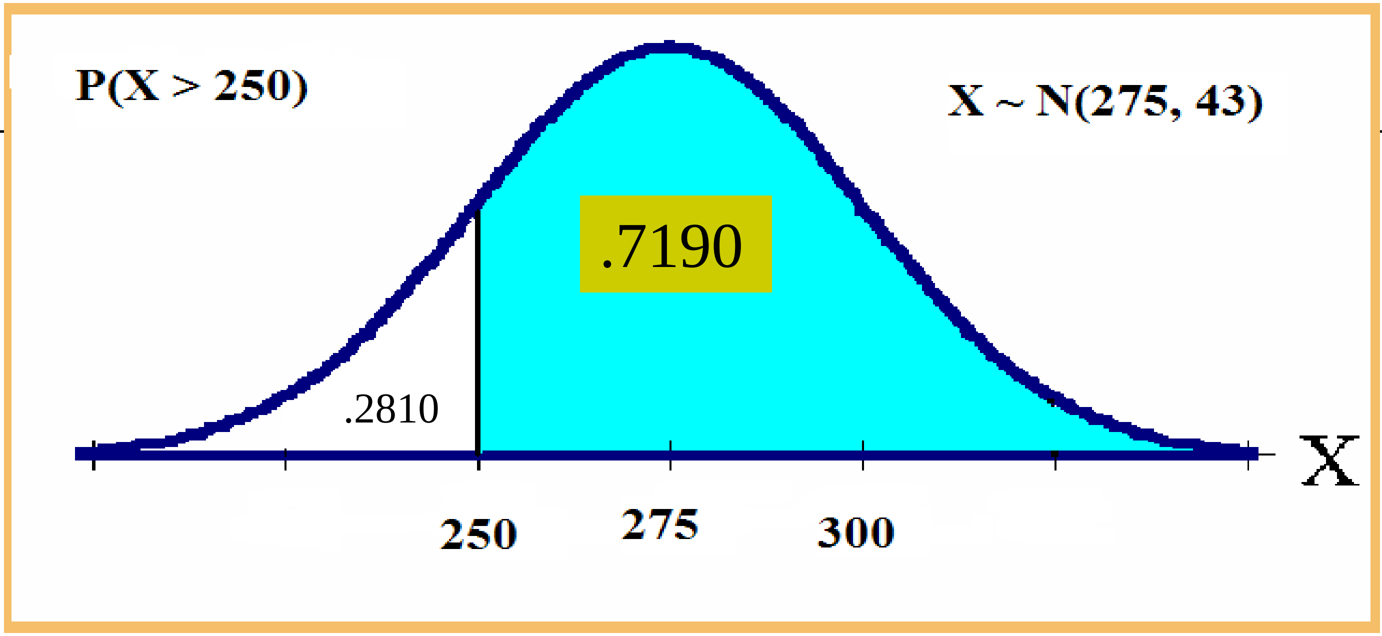
-.67

Is k positive or negative?

Direction of inequality; magnitude of probability

Look up .2514 in body of table; corresponding entry is -.67

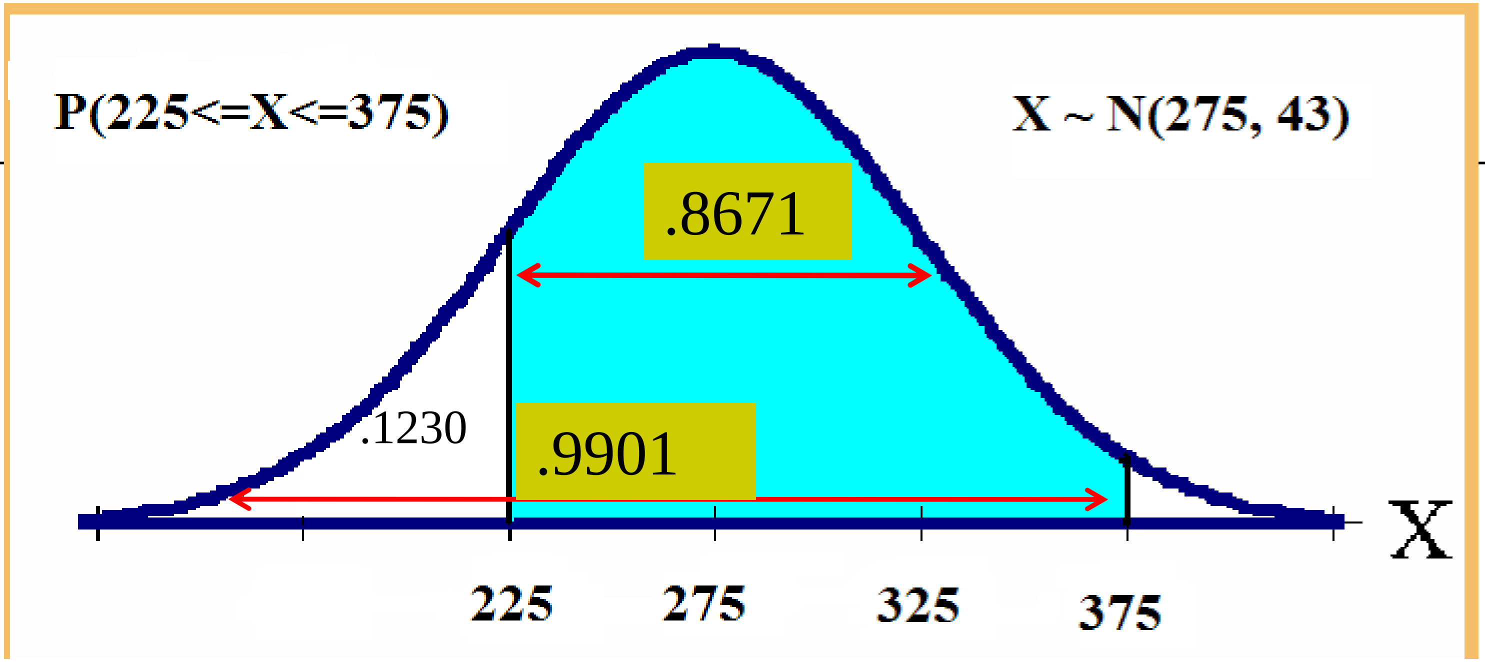
Examples (cont.) viii)



$$P(X > 250) = P(Z > \frac{250 - 275}{43})$$

$$P(Z > \frac{-25}{43}) = P(Z > -.58) = 1 - .2810 = .7190$$

Examples (cont.) ix)

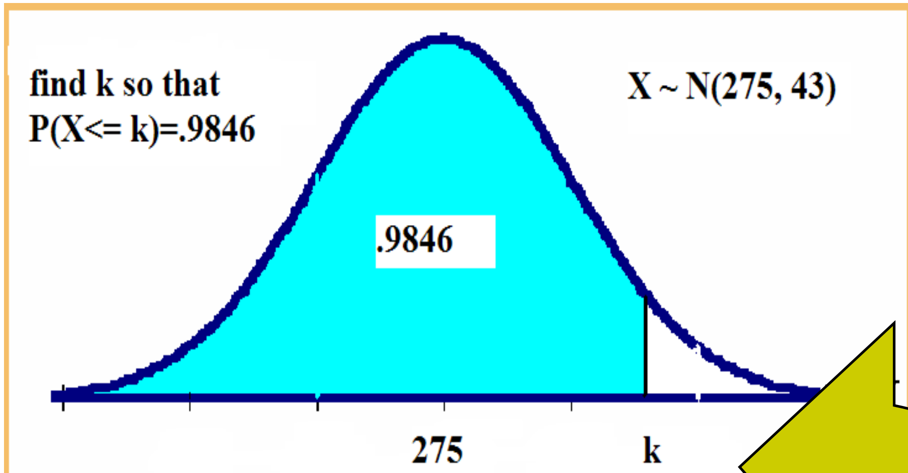


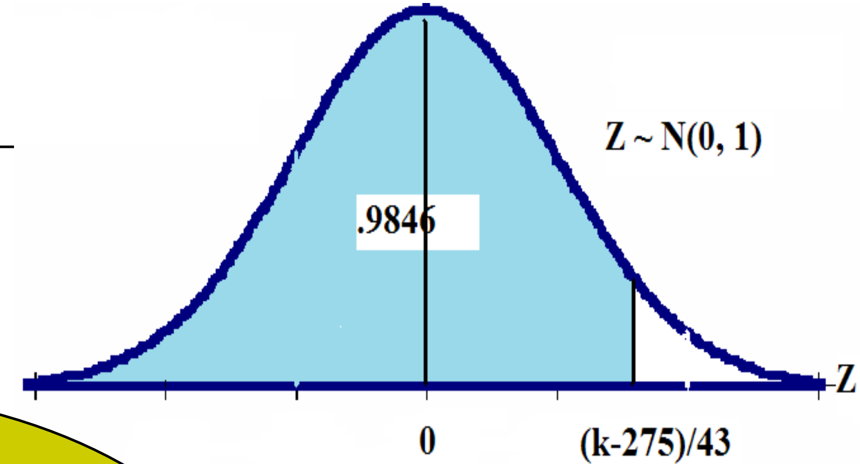
$$ix)\ P(225 \le x \le 375)$$

$$= P\left(\frac{225-275}{43} \le \frac{x-275}{43} \le \frac{375-275}{43}\right)$$

$$= P(-1.16 \le z \le 2.33) = .9901 - .1230 = .8671$$

X~N(275, 43) find k so that P(x<k)=.9846





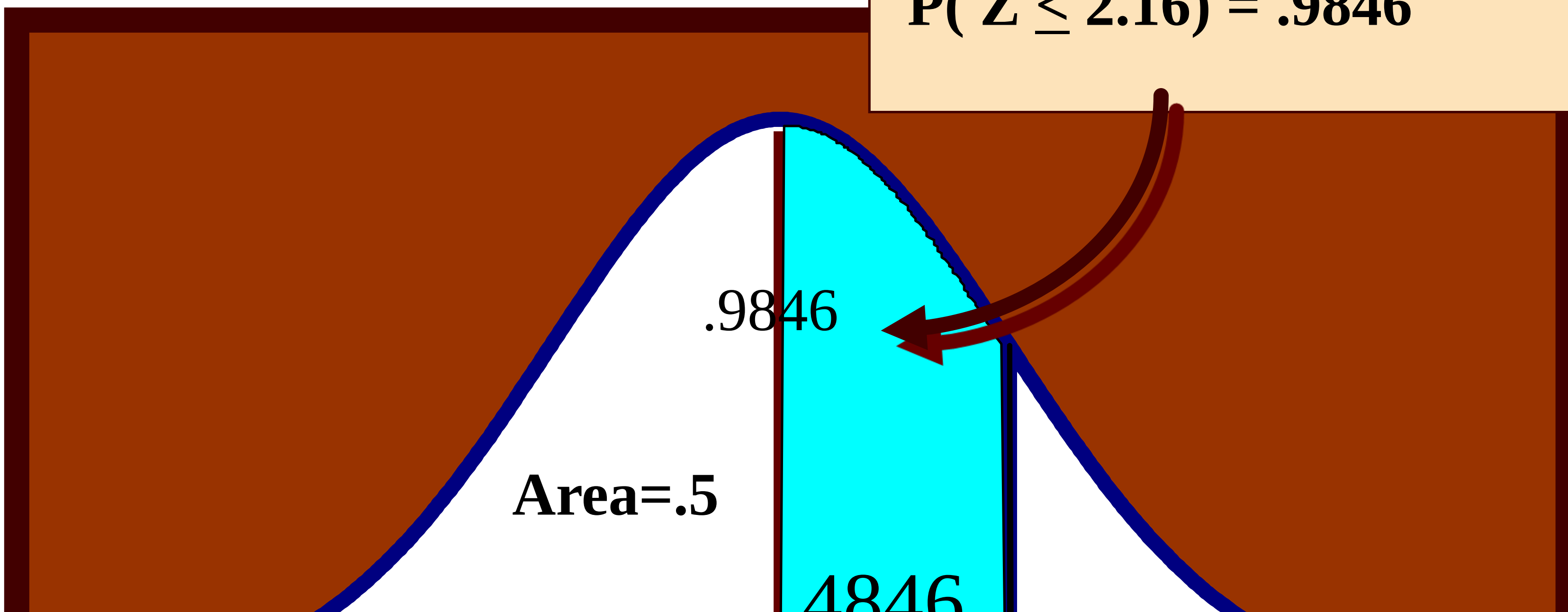
$$.9846 = P(x \le k) = P\left(\frac{x-275}{43} \le \frac{k-275}{43}\right)$$

$$= P\left(z \le \frac{k-275}{43}\right)$$

$$\Rightarrow \frac{k-275}{43} = 2.16 \text{ (from standard norm}$$

$$\Rightarrow k = 2.16(43)+275 = 367.88$$

$P(Z \leq 2.16) = .9846$



Example

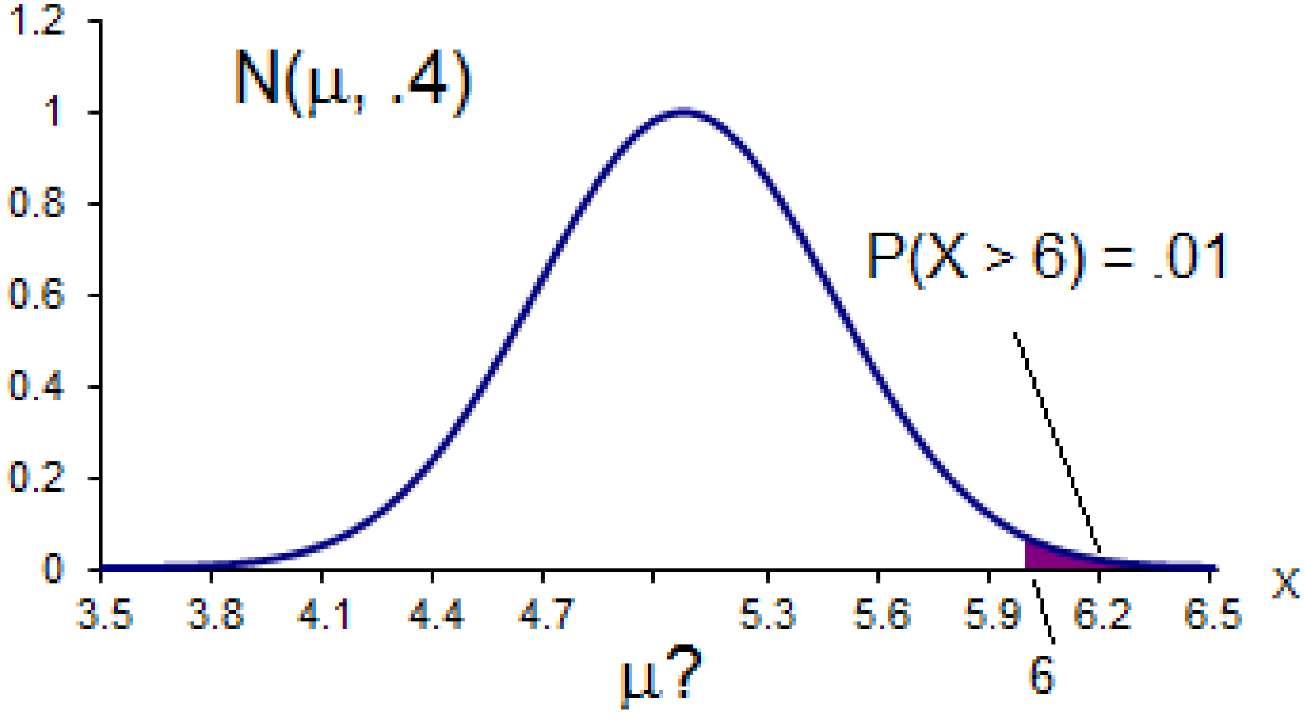
- Regulate blue dye for mixing paint; machine can be set to discharge an average of μ ml./can of paint.
- Amount discharged: $N(\mu, .4$ ml). If more than 6 ml. discharged into paint can, shade of blue is unacceptable.
- Determine the setting μ so that only 1% of the cans of paint will be unacceptable

Solution

$$X = \text{amount of dye discharged into can}$$

$$X \sim N(\mu, .4); \text{ determine } \mu \text{ so that}$$

$$P(X > 6) = .01$$



Solution (cont.)

X =amount of dye discharged into can

X ~N(μ, .4); determine μ so that

$$P(X > 6) = .01$$

$$.01 = P(x > 6) = P\left(\frac{x-\mu}{.4} > \frac{6-\mu}{.4}\right) = P\left(z > \frac{6-\mu}{.4}\right)$$

$$\Rightarrow \frac{6-\mu}{.4} = 2.33 \text{(from standard normal table)}$$

$$\Rightarrow \mu = 6\text{-}2.33(.4) = 5.068$$