Econ 2740: Economic Statistics
Fall, 2019. (Chapter 8)

Probability Density Functions (PDFs)…

A ***continuous random variable*** is one that can assume an **uncountable** number of values.

➔We cannot list the possible values because there is an uncountably infinite number of them.

➔For example, we cannot even attempt to count all the real numbers between zero and one.

➔Because there is an uncountably infinite number of values, the probability of each individual value is 0.

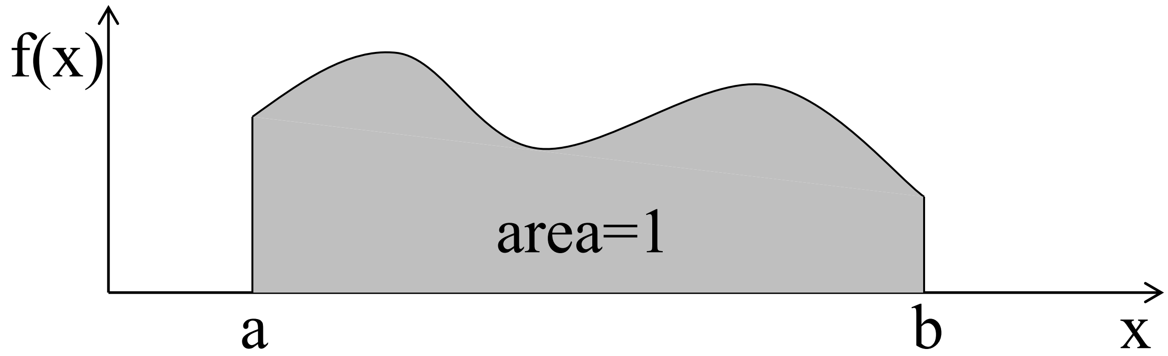
➔But we can assign a probability to a range of values

➔The probability density function (PDF) helps us do this …

Probability Density Function…

A function f(x) is called a ***probability density function (PDF)*** over the range $\mathbf{a} \le \mathbf{x} \le \mathbf{b}$ if it meets the following requirements:

1) $f(x) \ge 0$ for all **x** between **a** and **b**, and



2) The total area under the curve between **a** and **b** is 1.0

Use of PDF to Calculate Probabilities

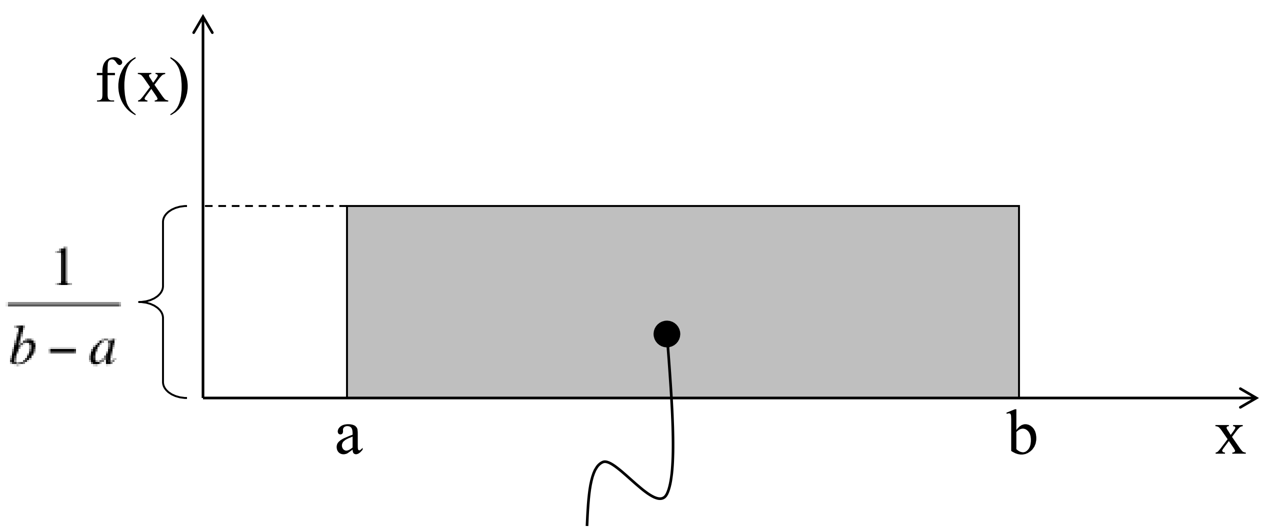
- We use to calculate probabilities
- The area underneath the PDF corresponds to the probability.
- E.g. if X is a continuous random variable and we want to calculate $P(3<X<5)$ then…
- We simply find the area under the PDF for X between 3 and 5.

Uniform Distribution…

Consider the ***uniform probability distribution*** (sometimes called the *rectangular probability distribution*).

It is described by the function:

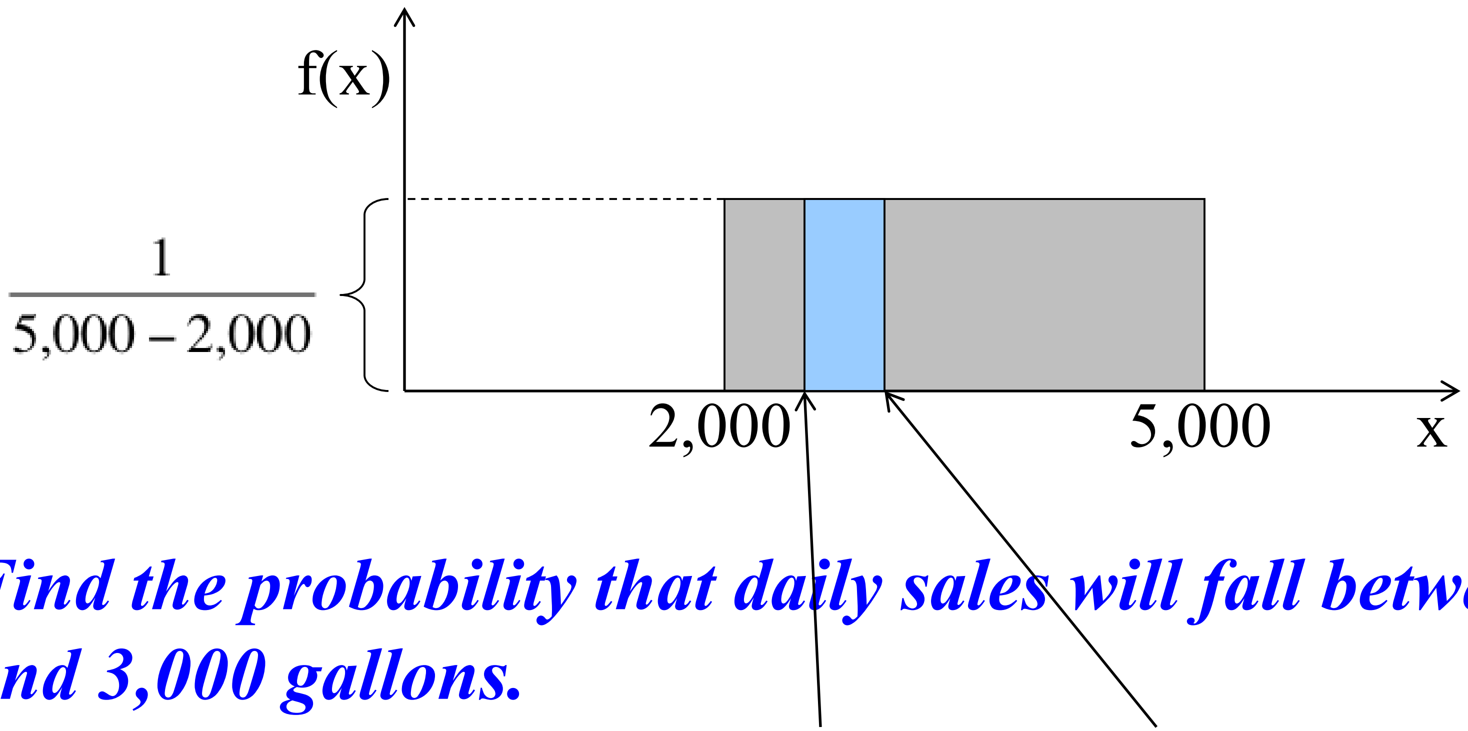
$$f(x) = \frac{1}{b-a}, \quad where \quad a \le x \le b$$



$$\text{area} = \text{width} \times \text{height} = (b-a) \times \frac{1}{b-a} = 1$$

Example 8.1(a)…

The amount of gasoline sold daily at a service station is uniformly distributed with a minimum of 2,000 gallons and a maximum of 5,000 gallons.

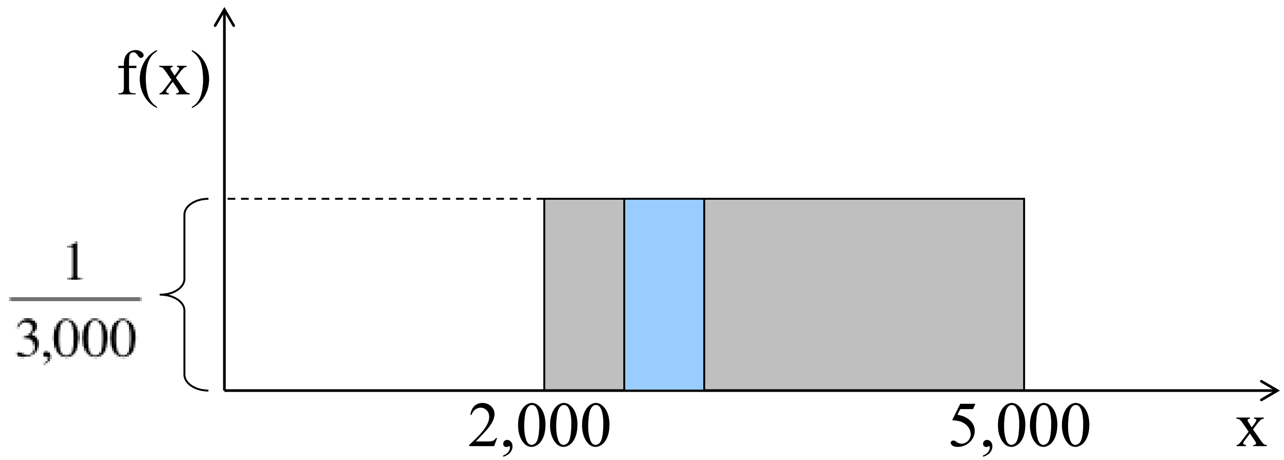


Find the probability that daily sales will fall between 2,500 and 3,000 gallons.

Algebraically: what is **P(2,500 ≤ X ≤ 3,000)** ?

Example 8.1(a)…

$$\mathbf{P(2{,}500 \le X \le 3{,}000)} = (3{,}000 - 2{,}500) \times \frac{1}{3{,}000} = .1667$$

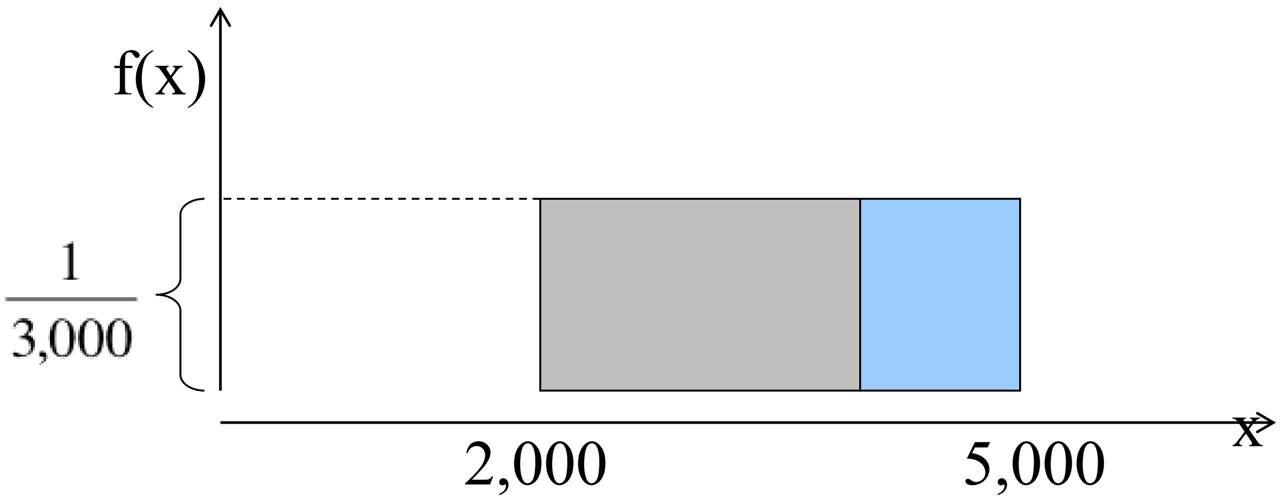


"there is about a 17% chance that between 2,500 and 3,000 gallons of gas will be sold on a given day"

Example 8.1(b)…

What is the probability that the service station will sell <u>at least</u> 4,000 gallons?

Algebraically: what is **P(X ≥ 4,000)** ?



$$P(X \geq 4,000) = (5,000 - 4,000)\frac{1}{5,000 - 2,000}$$

$$= \frac{1,000}{3,000} = \frac{1}{3}$$

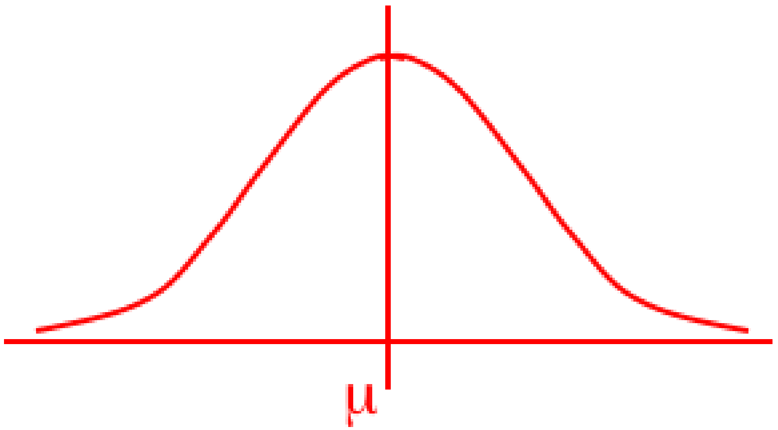
PDFs that range over all values

- Probability density functions don't have to be constrained to lie between two points a and b.
- They can range from -¥ to +¥
- But the area underneath them still has to "add up" (integrate) to one.
- The normal or bell-shaped distribution is a good example of a PDF that ranges from -¥ to +¥

The Normal Distribution…

The probability density function of a ***normal random variable*** is given by:

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



It looks like this:

Bell shaped,

Symmetrical around the mean

μ

The Normal Distribution…

Important things to note:

The normal distribution is fully defined by two parameters: its standard deviation and mean

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

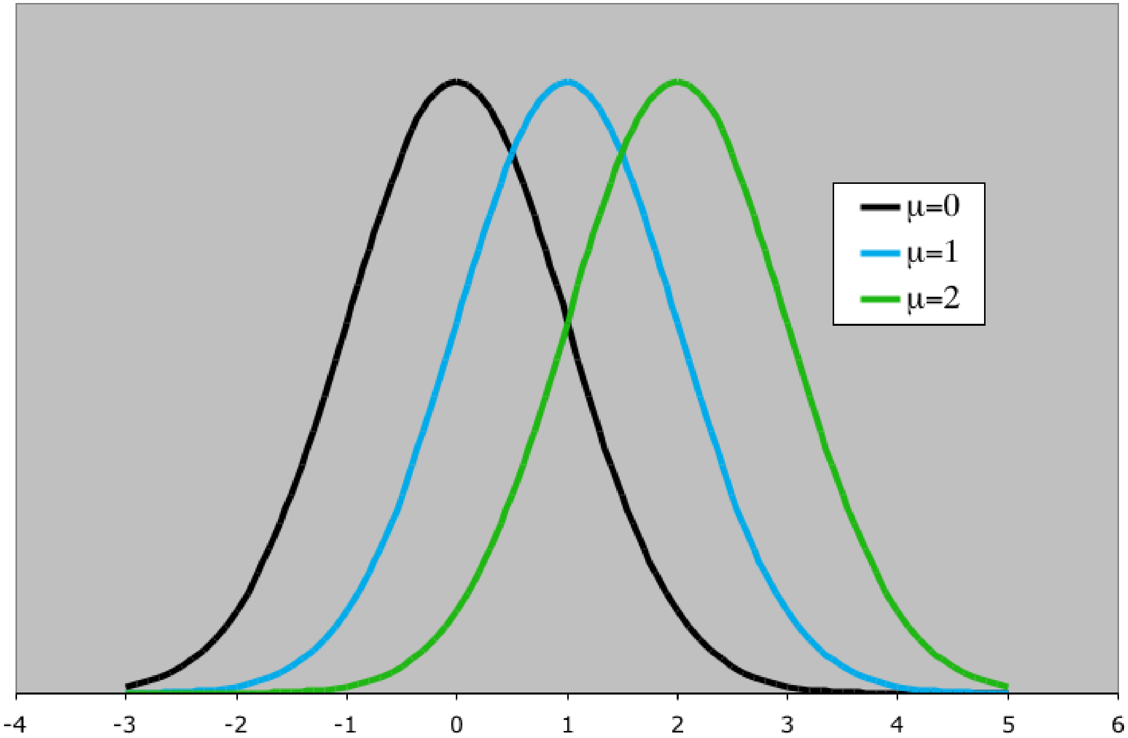
The normal distribution is bell shaped and symmetrical about the mean

Unlike the range of the uniform distribution ($a \leq x \leq b$)
Normal distributions ***range from minus infinity to plus infinity***

Normal Distribution…

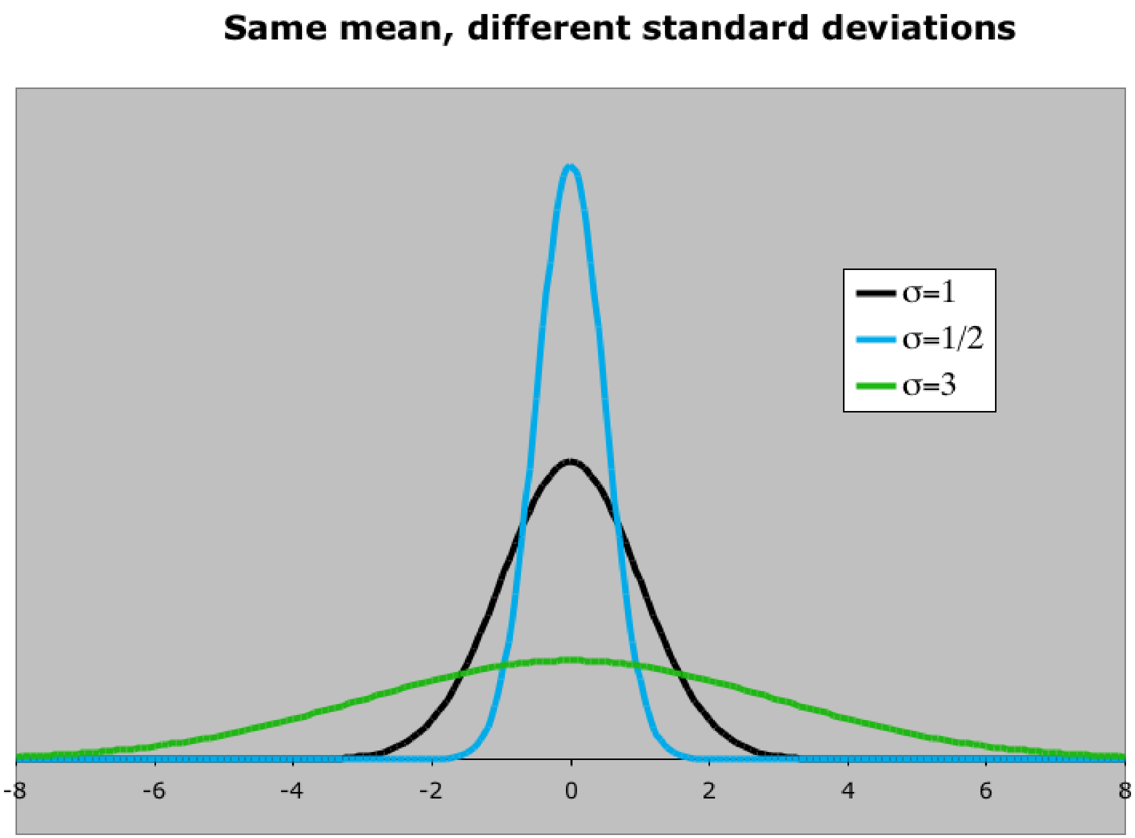
The normal distribution is described by two parameters: its mean μ and its standard deviation σ. Increasing the mean *shifts* the curve to the right…

Same variance, different means



Normal Distribution…

The normal distribution is described by two parameters: its mean μ and its standard deviation σ. Increasing the standard deviation "*flattens*" the curve…



Shorthand Notation

- The normal PDF has a complex formula
- But depends on only m and s or s^2
- So, its convenient to use the shorthand notation:

$$X \sim N\left(\mu, \sigma^2\right)$$

to indicate that X is a normally distributed random variable with population mean m and population variance s^2

ON BLACKBOARD: SOME SIMPLE EXAMPLES

Special Property of Normal Distribution

- Linear Combinations of Normal Random Variables are Still Normal Random Variables
- If X and Y are normal Random variables then so are the following linear functions of X and Y:
  - X+5
  - 3X
  - X+Y
  - 3X+4Y+18
- But, the following non-linear functions are no-longer Normal:
  - X^2
  - XY

Manipulating Normal Distributions

- If we use the laws of expectation and variance we can make new normal R.V's from linear combinations of existing random variables

E.g. if $X \sim N(3,4)$ and $Y \sim N(5,9)$

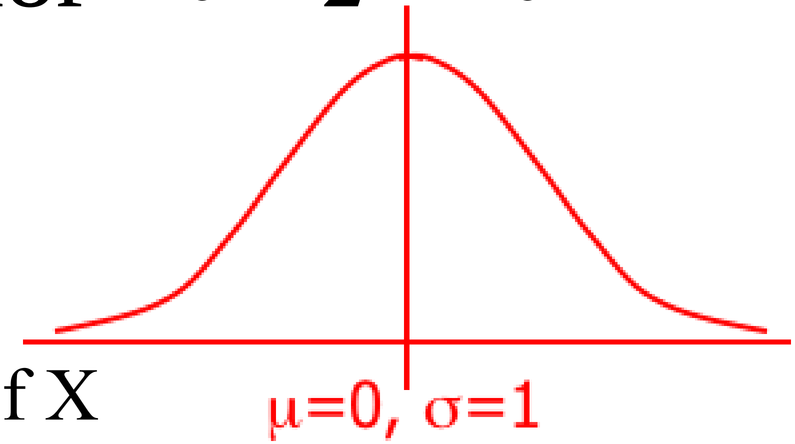
and $COV(X,Y) = 0$ then $X+Y \sim N(?,?)$

ON BLACKBOARD: MORE EXAMPLES

Standard Normal Distribution…

A normal distribution whose mean is zero and standard deviation is one is called the ***standard normal distribution***.

$$f(z) = \frac{1}{1\sqrt{2p}} e^{-\frac{1}{2}\left(\frac{z-0}{1}\right)^2} \quad \text{for } -\infty < z < \infty$$



By Convention: Refer to

Standard Normal using Z instead of X

IN SHORT HAND NOTATION: $Z \sim N(0,1)$

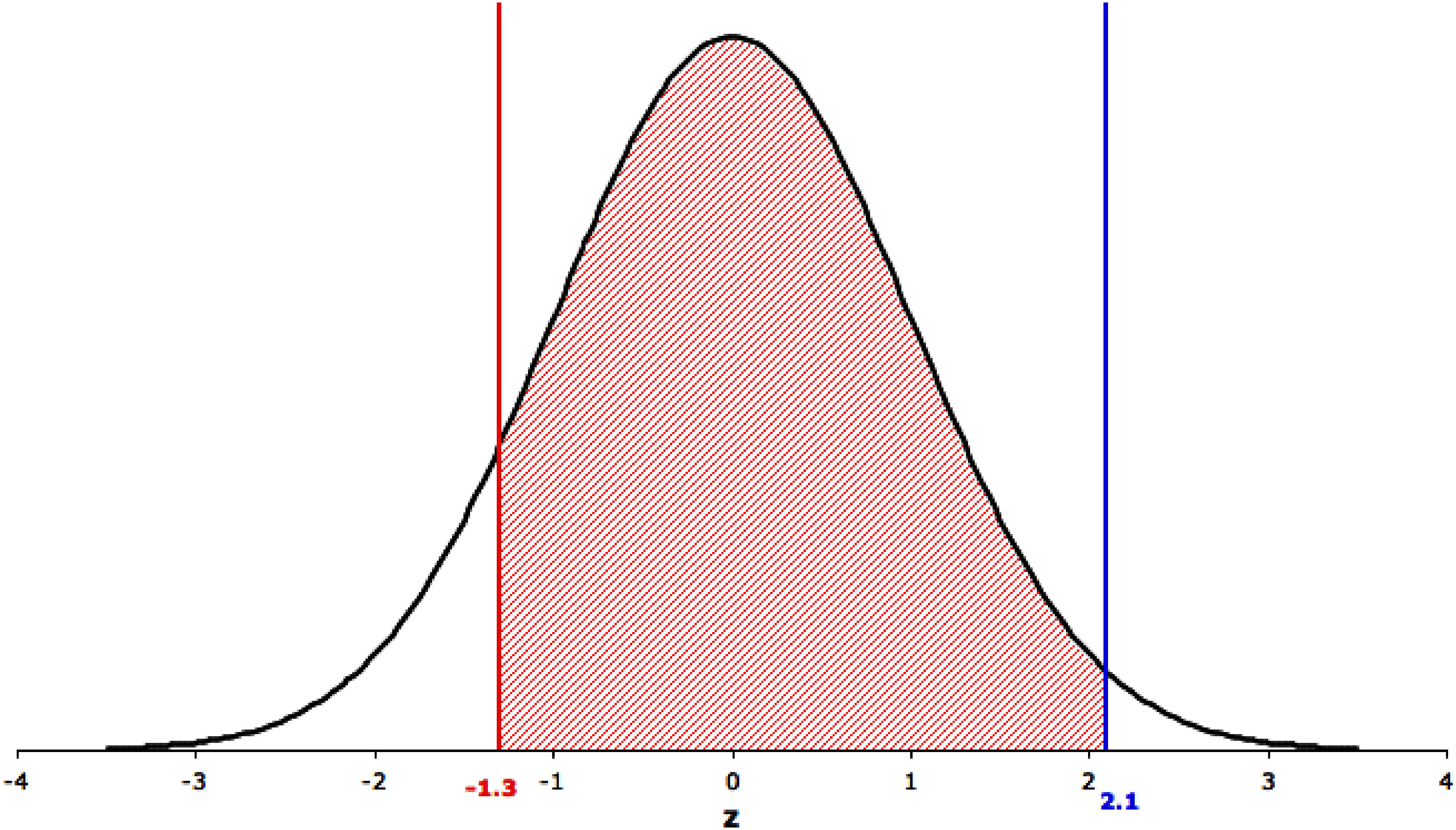
Using the Standard Normal Table

- What is $P(0<Z<1)$?
- In theory: Area under PDF between 0 and 1
- But, PDF of Z is NOT simple!
- So, statisticians have produced a table to help us answer this question.

PUT STANDARD NORMAL TABLE ON PROJECTOR

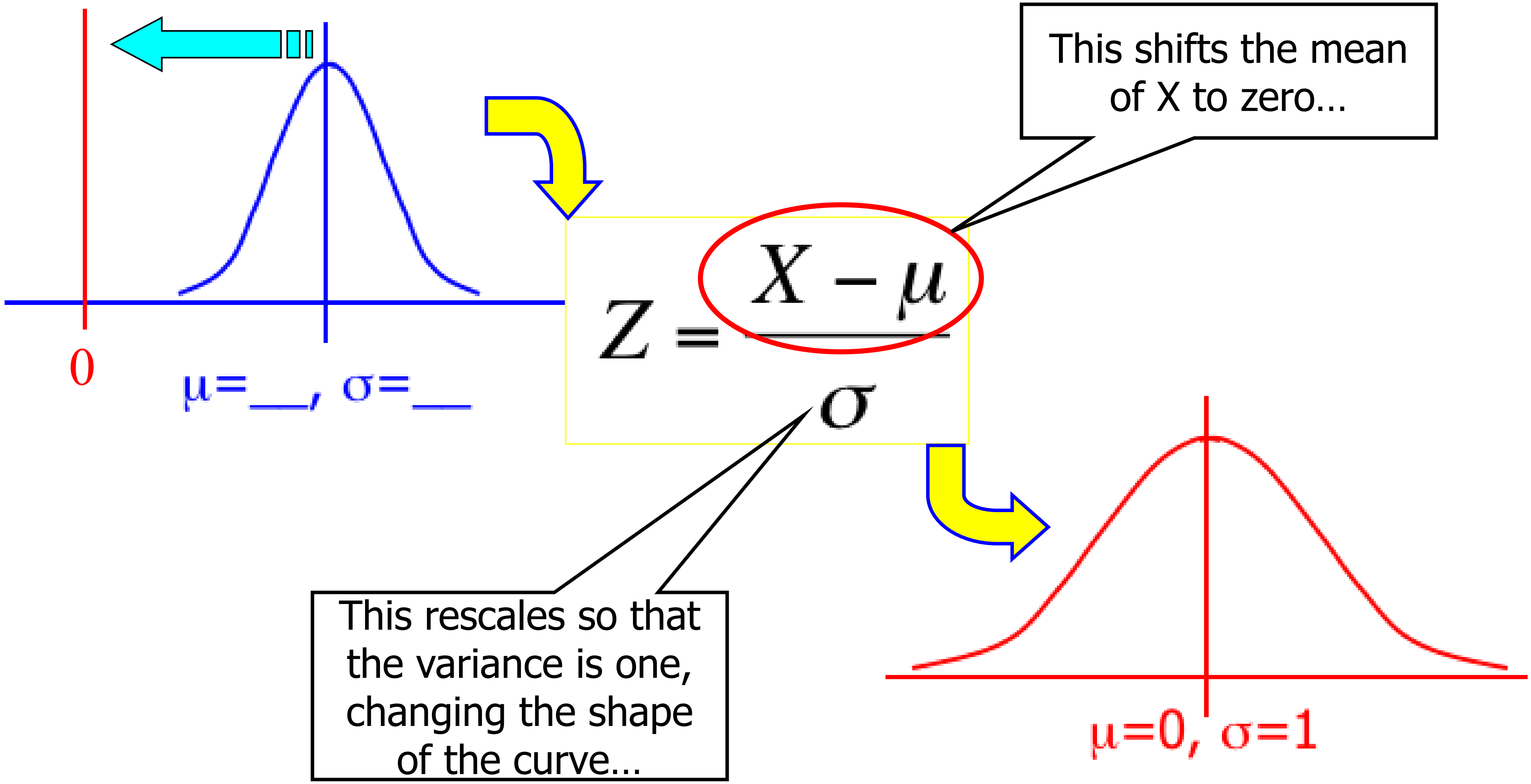
In pictures …

$P(-1.30 < Z < 2.10) = .8853$



Converting to Standard Normal…

We can use the following function to convert any normal random variable X to a **standard** normal random variable Z



In Shorthand Notation

$$X \sim N\left(\mu,\sigma^2\right)$$

$$X-\mu \sim N\left(0,\sigma^2\right)$$

$$Z=\frac{X-\mu}{\sigma} \sim \frac{1}{\sigma} N\left(0,\sigma^2\right)$$

$$=N\left(0,\frac{\sigma^2}{\sigma^2}\right)=N\left(0,1\right)$$

Looking up Normal Probabilities

What if X is normal, but not standard normal. How do we look up probabilities?

- The probabilities will depend on the population mean and variance
- Statisticians could not provide a table for every single possible combination of the mean and variance
- So instead, we convert to a standard normal and then use the standard normal table to look up the probabilities:
  - Step 1: convert to Z
  - Step 2: Look up the probability in Z-table

Now the Details …

Suppose that $X \sim N(\mu, \sigma^2)$.

What is P(X>a) where 'a' is some constant.

$$P(X>a)=P\left(\frac{X-\mu}{\sigma} > \frac{a-\mu}{\sigma}\right)$$

$$= P\left(Z > \frac{a-\mu}{\sigma}\right)$$

Look up this probability in standard normal table

ON BLACKBOARD: SEVERAL EXAMPLES

Example 8.2…

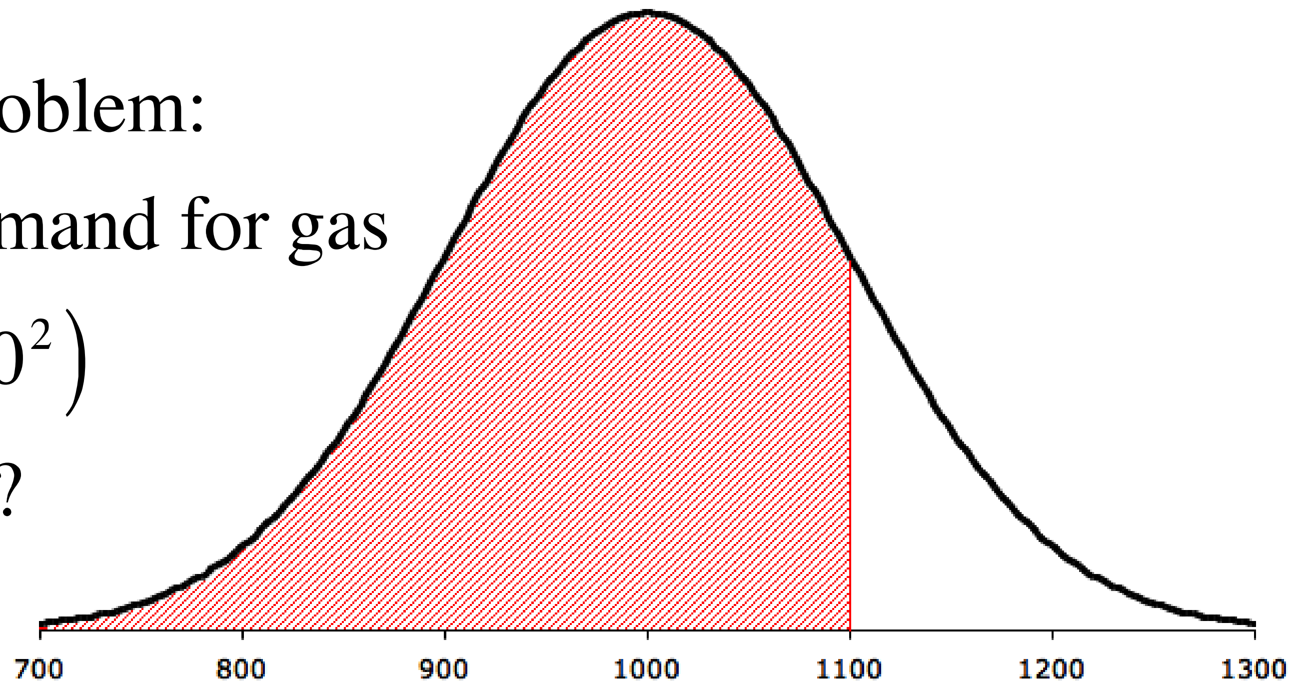
At a gas station the daily demand for regular gasoline is normally distributed with a mean of 1,000 gallons and a standard deviation of 100 gallons. There is exactly 1,100 gallons of regular gasoline in storage. The manager would like to know the probability that he will have enough regular gasoline to satisfy today's demands.

Translate the problem:

Let X be the demand for gas

$$X \sim N(1000, 100^2)$$

$$P(X < 1{,}100) = ?$$



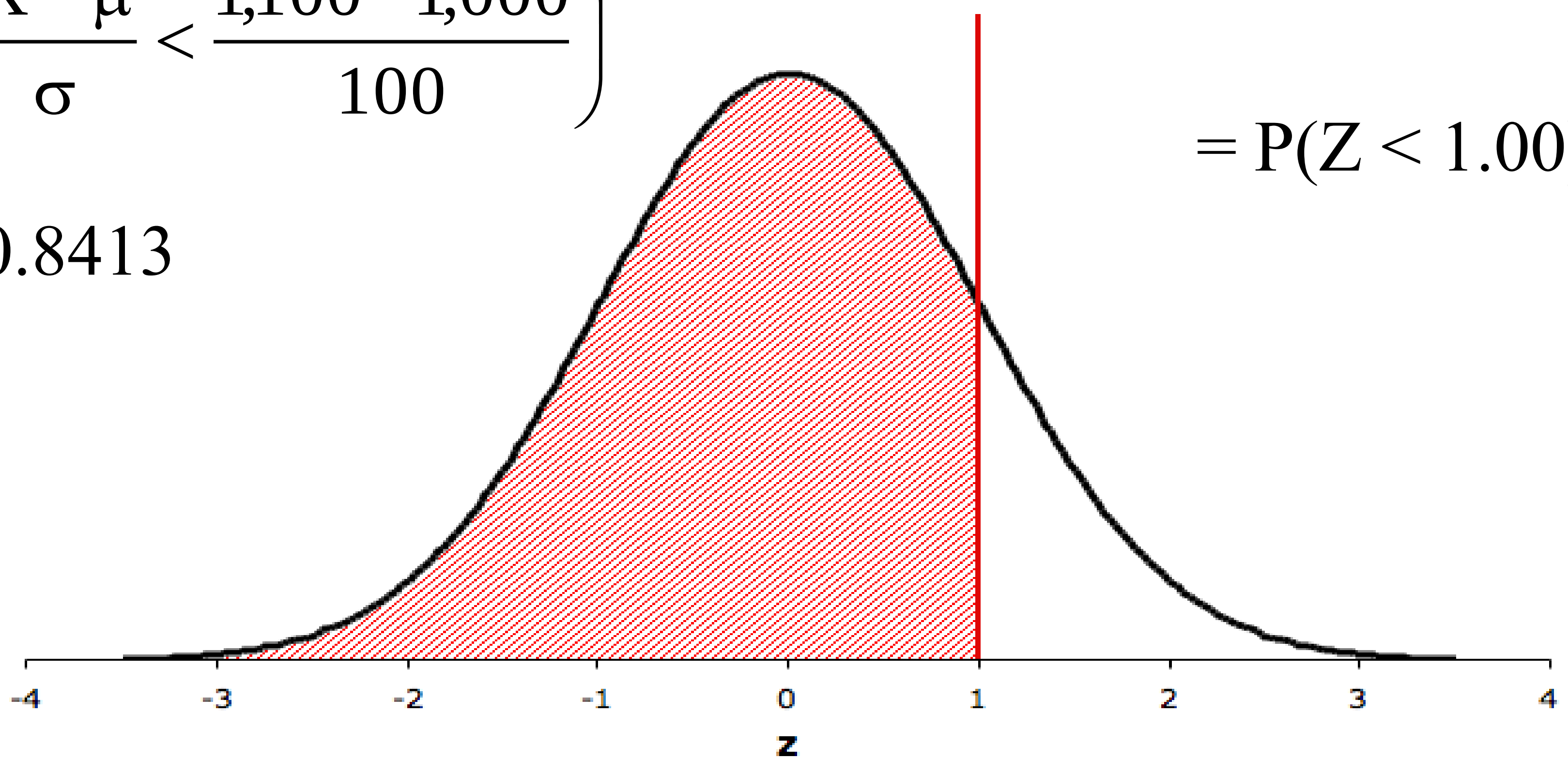
Example 8.2…

The first step is to standardize X. However, if we perform any operations on X we must perform the same operations on 1,100. Thus,

$$P\left(\frac{X-\mu}{\sigma} < \frac{1{,}100-1{,}000}{100}\right)$$

$$= P(Z < 1.00)$$

$$= 0.8413$$



APPLICATIONS IN FINANCE: Measuring Risk

Why risk is often measured by the variance and standard deviation?

Example 8.3

Consider an investment whose return is normally distributed with a mean of 10% and a standard deviation of 5%.

a. Determine the probability of losing money.

b. Find the probability of losing money when the standard deviation is equal to 10%.

Example 8.2

a The investment loses money when the return is negative. Thus we wish to determine

$$P(X < 0) = P\left(\frac{X - \mu}{\sigma} < \frac{0 - 10}{5}\right) = P(Z < -2.00)$$

From Table 3 we find $P(Z < -2.00) = .0228$

Therefore the probability of losing money is .0228

Example 8.2

b. If we increase the standard deviation to 10% the probability of suffering a loss becomes

$$P(X < 0) = P\left(\frac{X - \mu}{\sigma} < \frac{0 - 10}{10}\right)$$

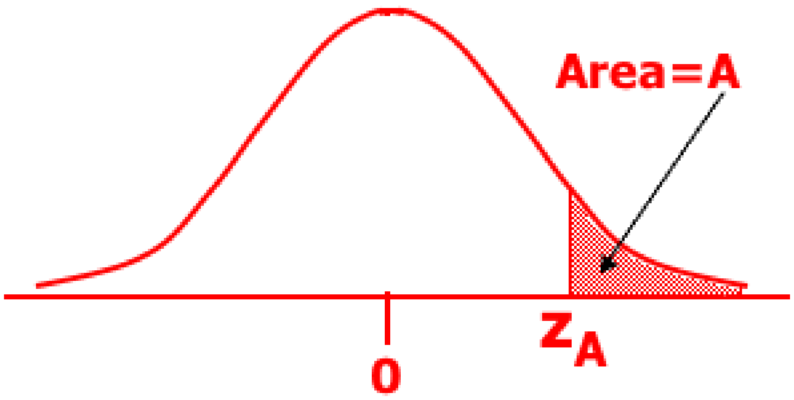
$$= P(Z < -1.00)$$

$$= .1587$$

Finding Values of Z…

Often we' re asked to find some value of Z for a given probability, i.e. given an area (A) under the curve, what is the corresponding value of z (z_A) on the horizontal axis that gives us this area? That is:

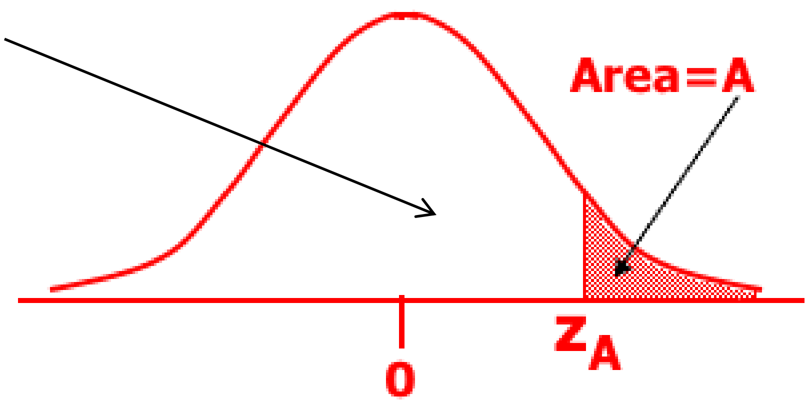
$$P(Z > z_A) = A$$



Finding Values of Z…

What value of z corresponds to an area under the curve of 2.5%? That is, what is $z_{.025}$?

(1 – A) = (1–.025) = .9750



If you do a "reverse look-up" on Table 3 for .9750, you will get the corresponding $z_A = 1.96$

Since $P(z > 1.96) = .025$, we say: $\mathbf{z_{.025} = 1.96}$

Exponential Distribution…

Another important continuous distribution is the ***exponential distribution*** which has this probability density function:

$$f(x) = \lambda e^{-\lambda x}, \quad x \geq 0$$

Note that $x \geq 0$. Time (for example) is a non-negative quantity; the exponential distribution is often used for time related phenomena such as the length of time between phone calls or between parts arriving at an assembly station.

We do not have time to discuss this distribution in detail, but you should know in which circumstances the exponential distribution is useful.

Other Continuous Distributions…

Three other important continuous distributions which will be used extensively in later sections are introduced in chapter 8:

Student ***t*** Distribution,

Chi-Squared Distribution, and

F Distribution.

We will defer further discussion of these until we use them later.