# Random Variables and the Normal Distribution

Nate Wells

Math 141, 3/31/21

## Outline

In this lecture, we will. . .

## Outline

In this lecture, we will. . .

- Define and explore Random Variables
- Investigate properties of the Normal Distribution
- Discuss the Central Limit Theorem and its role in statistics

Section 1

# Random Variables

## Definitions

A **random variable** is a numeric quantity whose value depends on the outcome of a random process.

## Definitions

A **random variable** is a numeric quantity whose value depends on the outcome of a random process.

- We use capital letters at the end of the alphabet $(W, X, Y, Z)$ to denote random variables.
  - We use lowercase letters $(w, x, y, z)$ to denote the particular values of a random variable

## Definitions

A **random variable** is a numeric quantity whose value depends on the outcome of a random process.

- We use capital letters at the end of the alphabet $(W, X, Y, Z)$ to denote random variables.
  - We use lowercase letters $(w, x, y, z)$ to denote the particular values of a random variable
- We use equation to express events associated to random variables.
  - Let $X = 5$ denotes the event "The random variable $X$ takes the value 5".

# Definitions

A **random variable** is a numeric quantity whose value depends on the outcome of a random process.

- We use capital letters at the end of the alphabet $(W, X, Y, Z)$ to denote random variables.
  - We use lowercase letters $(w, x, y, z)$ to denote the particular values of a random variable
- We use equation to express events associated to random variables.
  - Let $X = 5$ denotes the event "The random variable $X$ takes the value 5".
- Events associated to variables have probabilities of occurring.
  - $P(X = 5) = .5$ means $X$ has 50% probability of taking the value 5.

## Types of Random Variables

There are two main types of random variables:

1. **Discrete** variables can take only finitely many different values.
2. **Continuous** variables can take values equal to any real number in an interval.

# Types of Random Variables

There are two main types of random variables:

1. **Discrete** variables can take only finitely many different values.
2. **Continuous** variables can take values equal to any real number in an interval.

- Examples of discrete variables:
    - The number of credits a randomly chosen Reed student is taking.
    - The number of vegetarians in a random sample of 10 people.
    - The results of a coin flip, where 0 indicates Tails and 1 indicates Heads.

## Types of Random Variables

There are two main types of random variables:

1. **Discrete** variables can take only finitely many different values.
2. **Continuous** variables can take values equal to any real number in an interval.

- Examples of discrete variables:
  - The number of credits a randomly chosen Reed student is taking.
  - The number of vegetarians in a random sample of 10 people.
  - The results of a coin flip, where 0 indicates Tails and 1 indicates Heads.
- Examples of continuous variables:
  - The temperature of my office at a particular time of the day.
  - The amount of time it takes a radioactive particle to decay.

# Types of Random Variables

There are two main types of random variables:

1. **Discrete** variables can take only finitely many different values.
2. **Continuous** variables can take values equal to any real number in an interval.

- Examples of discrete variables:
  - The number of credits a randomly chosen Reed student is taking.
  - The number of vegetarians in a random sample of 10 people.
  - The results of a coin flip, where 0 indicates Tails and 1 indicates Heads.
- Examples of continuous variables:
  - The temperature of my office at a particular time of the day.
  - The amount of time it takes a radioactive particle to decay.
- Some discrete variables can be well-described by continuous variables:
  - The height of a random person selected from a large population.
  - The proportion of heads in a long sequence of coin flips.

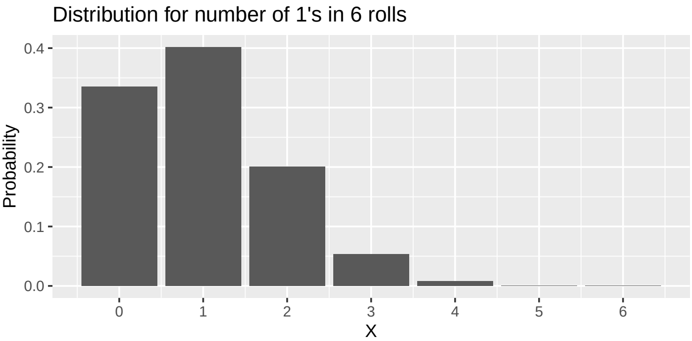
## The Distribution of a Random Variable

We often use histograms or bar charts to visualize discrete random variables.

# The Distribution of a Random Variable

We often use histograms or bar charts to visualize discrete random variables.

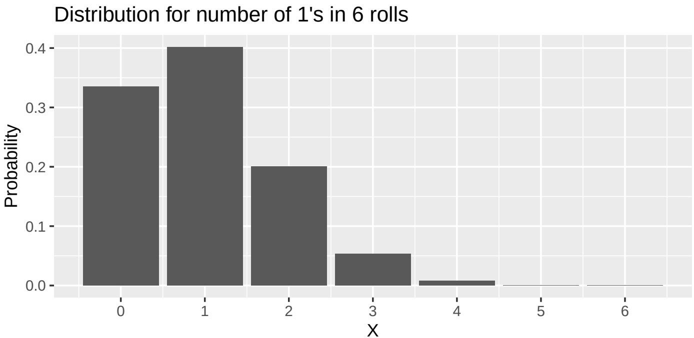
- Suppose a fair 6-sided die is rolled 6 times. Let $X$ be the number of 1s rolled. The distribution of $X$ is given by:



# The Distribution of a Random Variable

We often use histograms or bar charts to visualize discrete random variables.

- Suppose a fair 6-sided die is rolled 6 times. Let $X$ be the number of 1s rolled. The distribution of $X$ is given by:

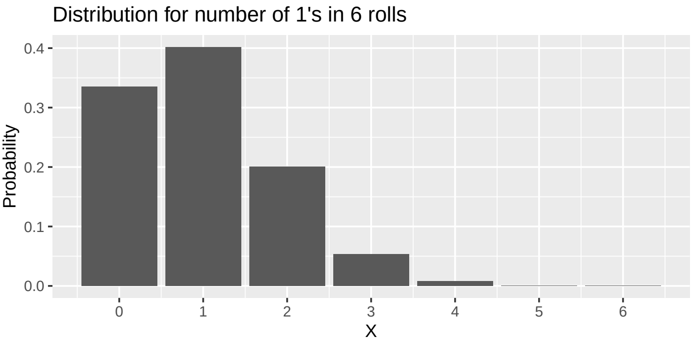



- We can use the plot to find probabilities of outcomes associated to the variable.

# The Distribution of a Random Variable

We often use histograms or bar charts to visualize discrete random variables.

- Suppose a fair 6-sided die is rolled 6 times. Let $X$ be the number of 1s rolled. The distribution of $X$ is given by:



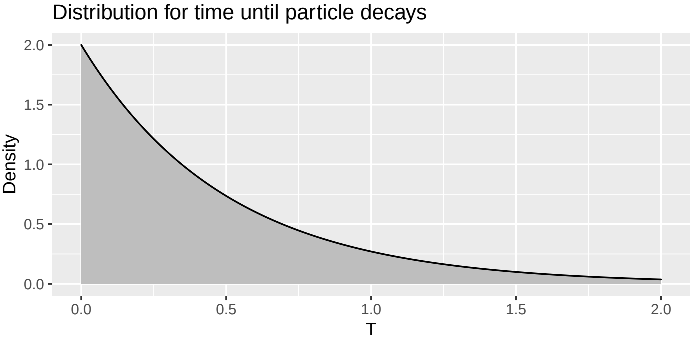

- We can use the plot to find probabilities of outcomes associated to the variable.
  - Calculate $P(X \leq 1)$. Then find $x$ so that $P(X \leq x) \geq .75$.

# The Distribution of a Continuous Variable

- We use **density plots** to visualize the distribution of a continuous variable. Areas under the plot correspond to probabilities.

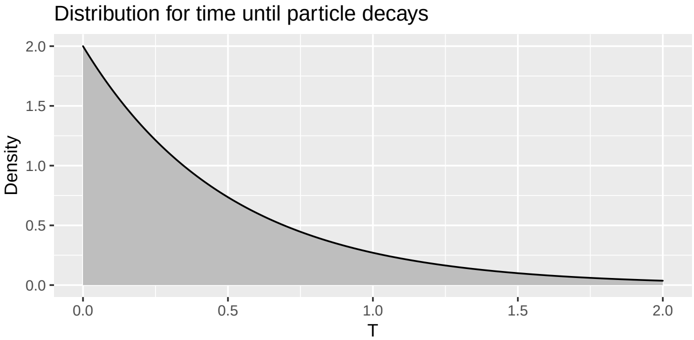
# The Distribution of a Continuous Variable

- We use **density plots** to visualize the distribution of a continuous variable. Areas under the plot correspond to probabilities.
- The distribution for the amount of time $T$ until a radioactive particle decays is given below:



# The Distribution of a Continuous Variable

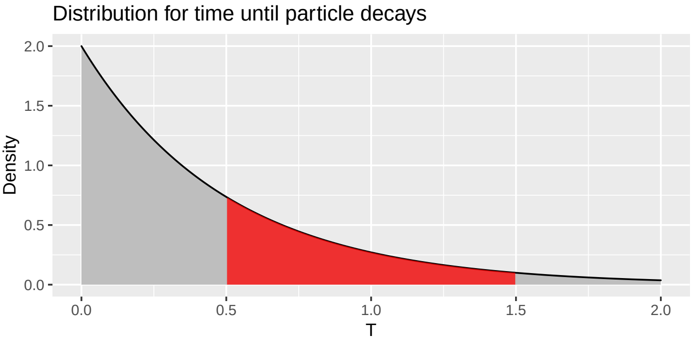
- We use **density plots** to visualize the distribution of a continuous variable. Areas under the plot correspond to probabilities.
- The distribution for the amount of time $T$ until a radioactive particle decays is given below:




- The probability that it takes between 0.5 and 1.5 units of time to decay is the area under the curve between 0.5 and 1.5. $P(0.5 < T < 1.5) =$

# The Distribution of a Continuous Variable

- We use **density plots** to visualize the distribution of a continuous variable. Areas under the plot correspond to probabilities.
- The distribution for the amount of time $T$ until a radioactive particle decays is given below:




- The probability that it takes between 0.5 and 1.5 units of time to decay is the area under the curve between 0.5 and 1.5. $P(0.5 < T < 1.5) = 0.34$

# Using Densities for Discrete Variables

If a discrete variable takes a large number of values which are close together, we can often approximate it using a continuous variable.

## Using Densities for Discrete Variables

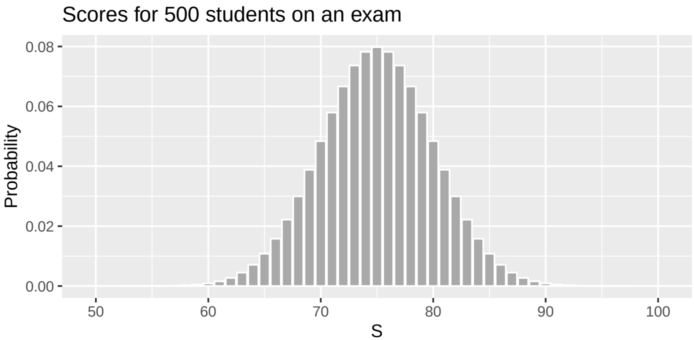
If a discrete variable takes a large number of values which are close together, we can often approximate it using a continuous variable.

- Suppose 500 students take a standardized exam, with mean 75 points. The distribution for the score $S$ of a randomly chosen student is:

# Using Densities for Discrete Variables

If a discrete variable takes a large number of values which are close together, we can often approximate it using a continuous variable.

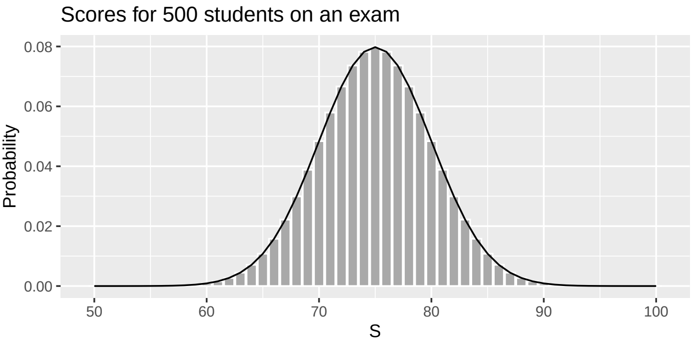
- Suppose 500 students take a standardized exam, with mean 75 points. The distribution for the score $S$ of a randomly chosen student is:



# Using Densities for Discrete Variables

If a discrete variable takes a large number of values which are close together, we can often approximate it using a continuous variable.

- Suppose 500 students take a standardized exam, with mean 75 points. The distribution for the score $S$ of a randomly chosen student is:



# Expected Value

The **expected value** (or mean) of a discrete random variable $X$ is

$$E[X] = x_1 P(X = x_1) + x_2 P(X = x_2) + \dots x_n P(X = x_n) = \sum_{i=1}^{n} x_i P(X = x_i)$$

## Expected Value

The **expected value** (or mean) of a discrete random variable $X$ is

$$E[X] = x_1 P(X = x_1) + x_2 P(X = x_2) + \ldots x_n P(X = x_n) = \sum_{i=1}^{n} x_i P(X = x_i)$$

- The expected value of $X$ is the sum of the value $X$ can take, weighted by the probability it takes those values.

# Expected Value

The **expected value** (or mean) of a discrete random variable $X$ is

$$E[X] = x_1 P(X = x_1) + x_2 P(X = x_2) + \dots x_n P(X = x_n) = \sum_{i=1}^{n} x_i P(X = x_i)$$

- The expected value of $X$ is the sum of the value $X$ can take, weighted by the probability it takes those values.
- Suppose we have a data set consisting of values $\{1, 1, 2, 2, 2, 2, 3, 4, 5, 5\}$. Let $X$ be a value chosen from this data set randomly. What is the expected value of $X$?

# Expected Value

The **expected value** (or mean) of a discrete random variable $X$ is

$$E[X] = x_1 P(X = x_1) + x_2 P(X = x_2) + \dots x_n P(X = x_n) = \sum_{i=1}^{n} x_i P(X = x_i)$$

- The expected value of $X$ is the sum of the value $X$ can take, weighted by the probability it takes those values.
- Suppose we have a data set consisting of values $\{1, 1, 2, 2, 2, 2, 3, 4, 5, 5\}$. Let $X$ be a value chosen from this data set randomly. What is the expected value of $X$?

$$\begin{aligned} E[X] =& 1P(X = 1) + 2P(X = 2) + 3P(X = 3) + 4P(X = 4) + 5P(X = 5) \\ =& 1\frac{2}{10} + 2\frac{4}{10} + 3\frac{1}{10} + 4\frac{1}{10} + 5\frac{2}{10} = \frac{27}{10} = 2.7 \end{aligned}$$

## Expected Value

The **expected value** (or mean) of a discrete random variable $X$ is

$$E[X] = x_1 P(X = x_1) + x_2 P(X = x_2) + \dots x_n P(X = x_n) = \sum_{i=1}^{n} x_i P(X = x_i)$$

- The expected value of $X$ is the sum of the value $X$ can take, weighted by the probability it takes those values.
- Suppose we have a data set consisting of values $\{1, 1, 2, 2, 2, 2, 3, 4, 5, 5\}$. Let $X$ be a value chosen from this data set randomly. What is the expected value of $X$?

$$\begin{aligned} E[X] =& 1P(X = 1) + 2P(X = 2) + 3P(X = 3) + 4P(X = 4) + 5P(X = 5) \\ =& 1\frac{2}{10} + 2\frac{4}{10} + 3\frac{1}{10} + 4\frac{1}{10} + 5\frac{2}{10} = \frac{27}{10} = 2.7 \end{aligned}$$

- But also notice that

$$\begin{aligned} E[X] =& \frac{1}{10}\left(1 \cdot 2 + 2 \cdot 4 + 3 \cdot 1 + 4 \cdot 1 + 5 \cdot 2\right) \\ =& \frac{1}{10}\left(1 + 1 + 2 + 2 + 2 + 2 + 3 + 4 + 5 + 5\right) \end{aligned}$$

## The Law of Large Numbers, again

Previously, we said that by the Law of Large numbers, the proportion of times an outcome occurs in a long sequence of trials is close to the probability for that outcome.

# The Law of Large Numbers, again

Previously, we said that by the Law of Large numbers, the proportion of times an outcome occurs in a long sequence of trials is close to the probability for that outcome.

This is a generalization:

**Theorem (The Law of Large Numbers)**

*Let $X$ be a random variable whose value depends on a random experiment. Suppose the experiment is repeated $n$ times and let $\bar{x}_n$ denote the arithmetic mean of the values of $X$ in each trial. As $n$ gets larger, the arithmetic mean $\bar{x}_n$ approaches the expected value $E[X]$ of that variable.*

## Gambler's Ruin

A roulette wheel consists of 37 wedge (18 black, 18 red, 1 green). A player may bet \$10 that a spun ball will land on a black wedge. If the ball lands on black, the player wins \$10. Otherwise, the player loses \$10.

## Gambler's Ruin

A roulette wheel consists of 37 wedge (18 black, 18 red, 1 green). A player may bet $10 that a spun ball will land on a black wedge. If the ball lands on black, the player wins $10. Otherwise, the player loses $10.

- Assuming each wedge has equal probability, what is the expected value of the bet?

- Suppose a gambler begins with $10,000. What will the gambler's fortune look like after 1000 plays?

Section 2

# The Normal Distribution

# The Normal Distribution

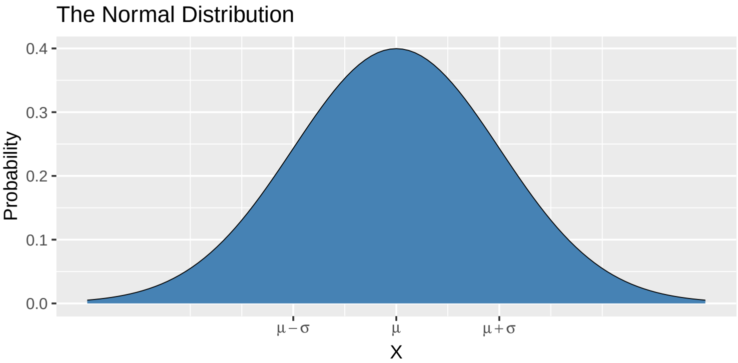
- The general Normal density curve with mean $\mu$ and standard deviation $\sigma$ is given by the formula

$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-(x-\mu)/2\sigma} \qquad \text{Don't memorize this}$$

# The Normal Distribution

- The general Normal density curve with mean $\mu$ and standard deviation $\sigma$ is given by the formula

$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-(x-\mu)/2\sigma} \qquad \text{Don't memorize this}$$



## Normal Probabilities

Recall that for a random variable which has a **continuous** distribution, we find probabilities by looking at areas under the density curve.

## Normal Probabilities

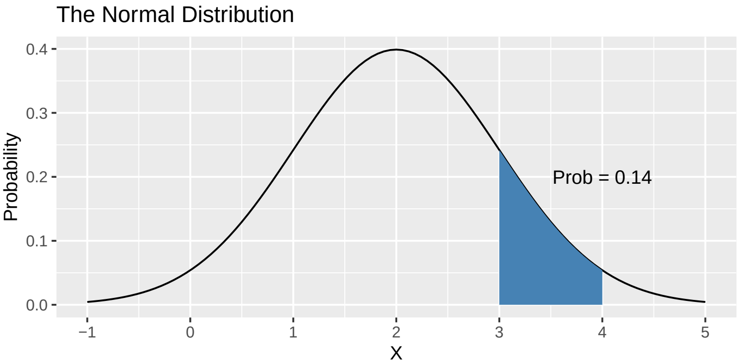
Recall that for a random variable which has a **continuous** distribution, we find probabilities by looking at areas under the density curve.

Suppose $X$ is Normally distributed with mean 2 and standard deviation 1. What is the probability that $X$ is between 3 and 4?

# Normal Probabilities

Recall that for a random variable which has a **continuous** distribution, we find probabilities by looking at areas under the density curve.

Suppose $X$ is Normally distributed with mean 2 and standard deviation 1. What is the probability that $X$ is between 3 and 4?



# Calculating Normal Probabilities in R

How do we actually find areas under the Normal density curve?

# Calculating Normal Probabilities in R

How do we actually find areas under the Normal density curve?

- R has a built-in function for computing cummulative probabilites under Normal densities: `pnorm(q =... , mean =... , sd =... )`

# Calculating Normal Probabilities in R

How do we actually find areas under the Normal density curve?

- R has a built-in function for computing cummulative probabilites under Normal densities: `pnorm(q =... , mean =... , sd =... )`
- For example, the following code computes the area left of 1.5 in the Normal distribution with mean 0 and standard deviation 1:

```
pnorm(q =1.5 , mean =0 , sd =1 )
```

```
## [1] 0.9331928
```

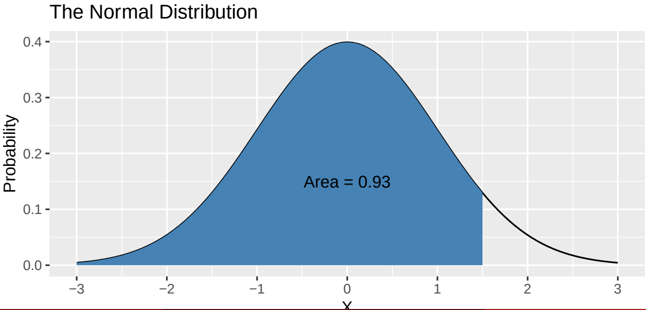
# Calculating Normal Probabilities in R

How do we actually find areas under the Normal density curve?

- R has a built-in function for computing cummulative probabilites under Normal densities: `pnorm(q =... , mean =... , sd =... )`
- For example, the following code computes the area left of 1.5 in the Normal distribution with mean 0 and standard deviation 1:

```
pnorm(q =1.5 , mean =0 , sd =1 )
```

```
## [1] 0.9331928
```



## Finding Areas of General Regions

The `pnorm` function lets us compute cumulative areas (i.e. all area to the left of a given value). But how do we compute the area between two values?

# Finding Areas of General Regions

The `pnorm` function lets us compute cumulative areas (i.e. all area to the left of a given value). But how do we compute the area between two values?

- Answer: By computing two cumulative areas and subtracting the results!

# Finding Areas of General Regions

The `pnorm` function lets us compute cumulative areas (i.e. all area to the left of a given value). But how do we compute the area between two values?

- Answer: By computing two cumulative areas and subtracting the results!

Find the area between -.25 and 1.5 under the Normal density with mean 0 and standard deviation 1.

```
pnorm(q =1.5 , mean =0 , sd =1 ) - pnorm(q = -.25 , mean =0 , sd =1 )
```

```
## [1] 0.5318991
```

# Finding Areas of General Regions

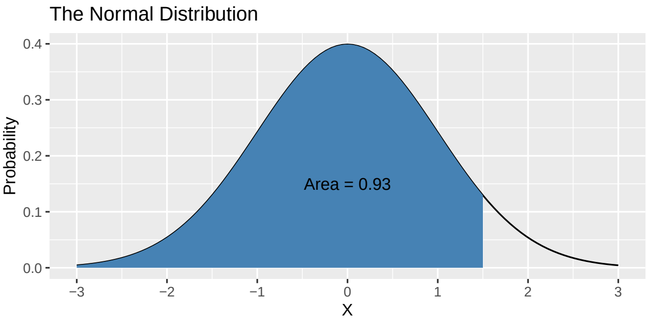
The pnorm function lets us compute cumulative areas (i.e. all area to the left of a given value). But how do we compute the area between two values?

- Answer: By computing two cumulative areas and subtracting the results!

Find the area between -.25 and 1.5 under the Normal density with mean 0 and standard deviation 1.

```
pnorm(q =1.5 , mean =0 , sd =1 ) - pnorm(q = -.25 , mean =0 , sd =1 )
```

```
## [1] 0.5318991
```



# Finding Areas of General Regions

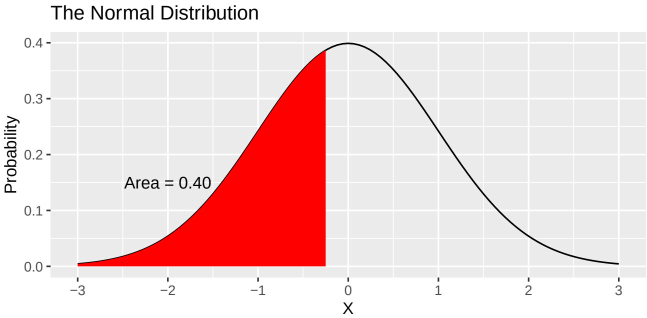
The pnorm function lets us compute cumulative areas (i.e. all area to the left of a given value). But how do we compute the area between two values?

- Answer: By computing two cumulative areas and subtracting the results!

Find the area between -.25 and 1.5 under the Normal density with mean 0 and standard deviation 1.

```
pnorm(q =1.5 , mean =0 , sd =1 ) - pnorm(q = -.25 , mean =0 , sd =1 )
```

```
## [1] 0.5318991
```



# Finding Areas of General Regions under Normal curve

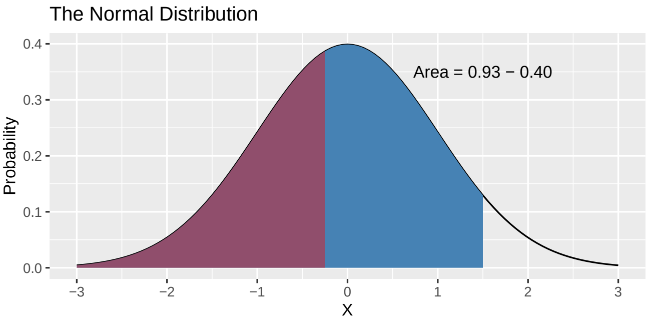
The `pnorm` function lets us compute cumulative areas (i.e. all area to the left of a given value). But how do we compute the area between two values?

- Answer: By computing two cumulative areas and subtracting the results!

Find the area between -.25 and 1.5 under the Normal density with mean 0 and standard deviation 1.

```
pnorm(q =1.5 , mean =0 , sd =1 ) - pnorm(q = -.25 , mean =0 , sd =1 )
```

```
## [1] 0.5318991
```



# Finding Areas of General Regions

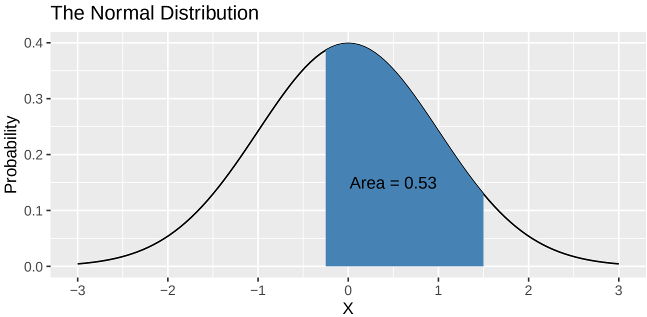
The pnorm function lets us compute cumulative areas (i.e. all area to the left of a given value). But how do we compute the area between two values?

- Answer: By computing two cumulative areas and subtracting the results!

Find the area between -.25 and 1.5 under the Normal density with mean 0 and standard deviation 1.

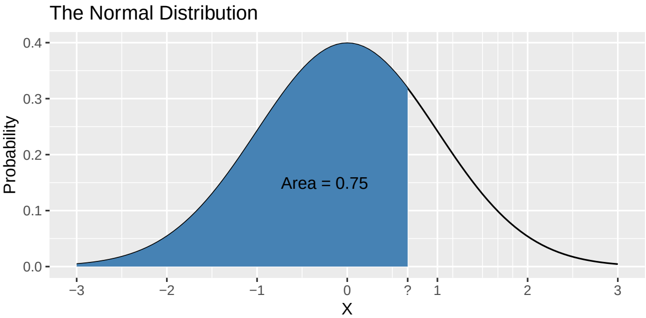
```
pnorm(q =1.5 , mean =0 , sd =1 ) - pnorm(q = -.25 , mean =0 , sd =1 )
```

```
## [1] 0.5318991
```



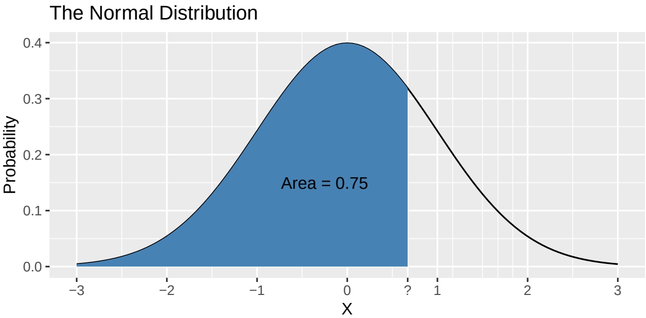
# Finding Quantiles

Suppose we instead have the opposite problem: We want to FIND the value of X with a given cumulative area.



# Finding Quantiles

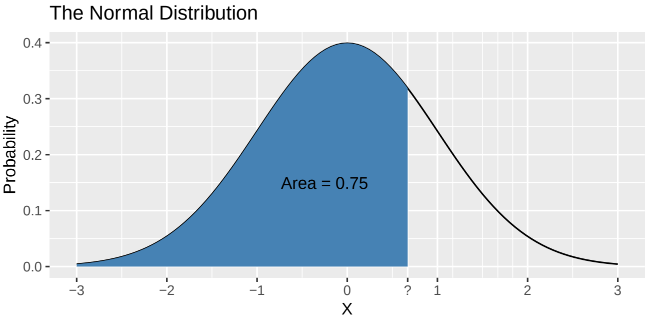
Suppose we instead have the opposite problem: We want to FIND the value of X with a given cumulative area.




- That is, we want to find the .75 quantile (i.e. the 75th percentile)

# Finding Quantiles

Suppose we instead have the opposite problem: We want to FIND the value of X with a given cumulative area.

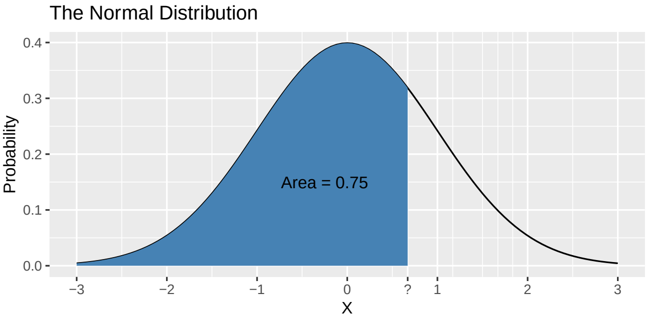



- That is, we want to find the .75 quantile (i.e. the 75th percentile)

R has a built-in function for that too! `qnorm(p =... , mean =... , sd =... )`

# Finding Quantiles

Suppose we instead have the opposite problem: We want to FIND the value of X with a given cumulative area.




- That is, we want to find the .75 quantile (i.e. the 75th percentile)

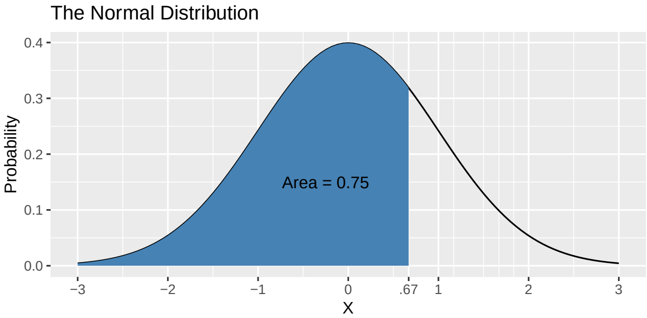
R has a built-in function for that too! `qnorm(p =... , mean =... , sd =... )`

```
qnorm(p =.75 , mean =0 , sd =1 )
```

```
## [1] 0.6744898
```

# Finding Quantiles

Suppose we instead have the opposite problem: We want to FIND the value of X with a given cumulative area.




- That is, we want to find the .75 quantile (i.e. the 75th percentile)

R has a built-in function for that too! `qnorm(p =... , mean =... , sd =... )`

```
qnorm(p =.75 , mean =0 , sd =1 )
```

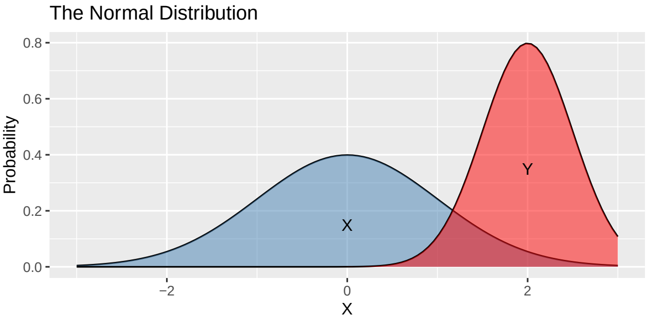
```
## [1] 0.6744898
```

## Scale and Translation Invariance

- Consider a Normal variable $X$ with $\mu = 0$ and $\sigma = 1$, and another Normal variable $Y$ with mean $\mu = 2$ and $\sigma = .25$.

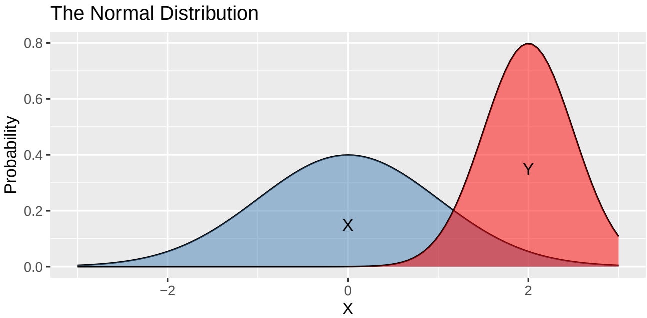
# Scale and Translation Invariance

- Consider a Normal variable $X$ with $\mu = 0$ and $\sigma = 1$, and another Normal variable $Y$ with mean $\mu = 2$ and $\sigma = .25$.



# Scale and Translation Invariance

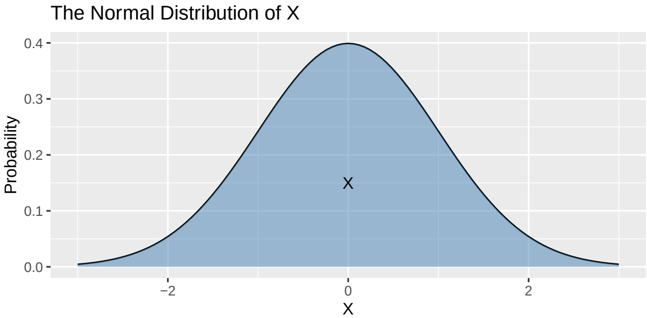
- Consider a Normal variable $X$ with $\mu = 0$ and $\sigma = 1$, and another Normal variable $Y$ with mean $\mu = 2$ and $\sigma = .25$.




- The distributions for $X$ and $Y$ have different means and different heights and widths. . .
  - But otherwise have identitical shapes!

# Scale and Translation Invariance

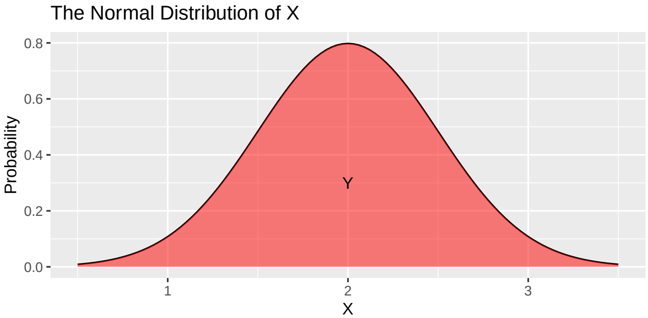
- Consider a Normal variable $X$ with $\mu = 0$ and $\sigma = 1$, and another Normal variable $Y$ with mean $\mu = 2$ and $\sigma = .25$.




- The distributions for $X$ and $Y$ have different means and different heights and widths. . .
  - But otherwise have identitical shapes!

# Scale and Translation Invariance

- Consider a Normal variable $X$ with $\mu = 0$ and $\sigma = 1$, and another Normal variable $Y$ with mean $\mu = 2$ and $\sigma = .25$.




- The distributions for $X$ and $Y$ have different means and different heights and widths. . .
  - But otherwise have identical shapes!

# Location-Scale Transformations

The previous example suggest that if we shift and rescale a Normal random variable, we should still get a Normal random variable

## Location-Scale Transformations

The previous example suggest that if we shift and rescale a Normal random variable, we should still get a Normal random variable

### Theorem

*Suppose $X$ is a Normal random variable with mean $\mu$ and standard deviation $\sigma$. Then $Z = \frac{X-\mu}{\sigma}$ is a Normal random variable with mean 0 and standard deviation 1.*

# Location-Scale Transformations

The previous example suggest that if we shift and rescale a Normal random variable, we should still get a Normal random variable

**Theorem**

*Suppose $X$ is a Normal random variable with mean $\mu$ and standard deviation $\sigma$. Then $Z = \frac{X-\mu}{\sigma}$ is a Normal random variable with mean 0 and standard deviation 1.*

The Normal variable with mean 0 and standard deviation 1 is given a special name: **the standard Normal**.

# Location-Scale Transformations

The previous example suggest that if we shift and rescale a Normal random variable, we should still get a Normal random variable

## Theorem

*Suppose $X$ is a Normal random variable with mean $\mu$ and standard deviation $\sigma$. Then $Z = \frac{X-\mu}{\sigma}$ is a Normal random variable with mean 0 and standard deviation 1.*

The Normal variable with mean 0 and standard deviation 1 is given a special name: **the standard Normal**.

The process of subtracting off the mean from a random variable and dividing by the standard deviation is called **standardizing.**

# Location-Scale Transformations

The previous example suggest that if we shift and rescale a Normal random variable, we should still get a Normal random variable

**Theorem**

*Suppose $X$ is a Normal random variable with mean $\mu$ and standard deviation $\sigma$. Then $Z = \frac{X-\mu}{\sigma}$ is a Normal random variable with mean 0 and standard deviation 1.*

The Normal variable with mean 0 and standard deviation 1 is given a special name: **the standard Normal**.

The process of subtracting off the mean from a random variable and dividing by the standard deviation is called **standardizing.**

It's often useful to standardize a variable so that we only need to consider a single density function (the *standard* Normal density) rather than many (one for each choice of $\mu$ and $\sigma$)

Section 3

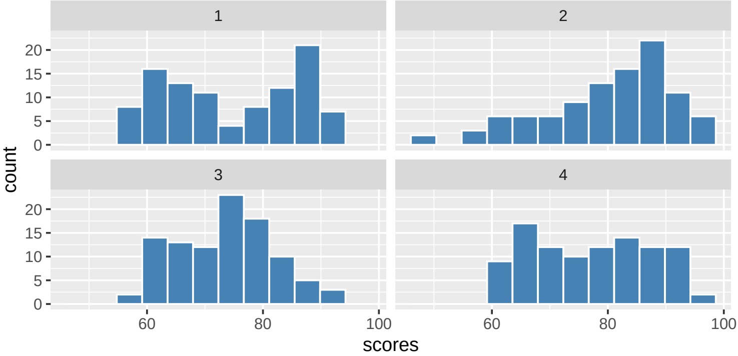
# The Central Limit Theorem

# Exam scores

Consider the following distributions for scores on a statistics exam for 4 classes of 100 students:

# Exam scores

Consider the following distributions for scores on a statistics exam for 4 classes of 100 students:



# Random Sample Means

Suppose we repeatedly take samples of 10 students from each class, and compute the average score $\bar{x}$ for each sample

## Random Sample Means

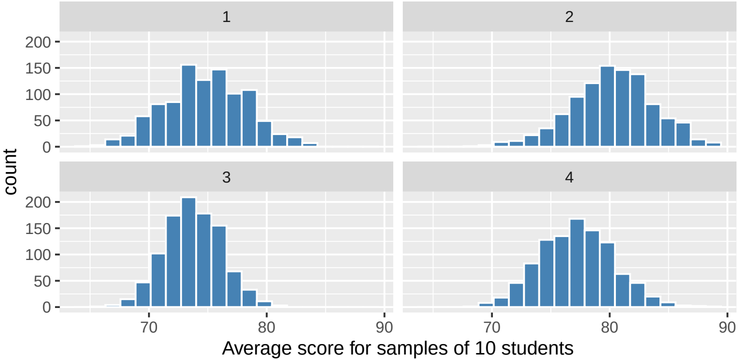
Suppose we repeatedly take samples of 10 students from each class, and compute the average score $\bar{x}$ for each sample

- What does the distribution of sample means $\bar{x}$ look like?

# Random Sample Means

Suppose we repeatedly take samples of 10 students from each class, and compute the average score $\bar{x}$ for each sample

- What does the distribution of sample means $\bar{x}$ look like?

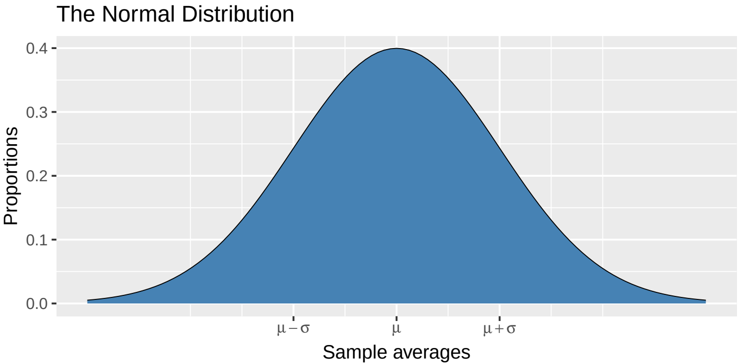


# The Normal Distribution

- In the previous example, the sampling distribution for *each* class appeared approximately Normal, regardless of the shape of the population distribution.

# The Normal Distribution

- In the previous example, the sampling distribution for *each* class appeared approximately Normal, regardless of the shape of the population distribution.

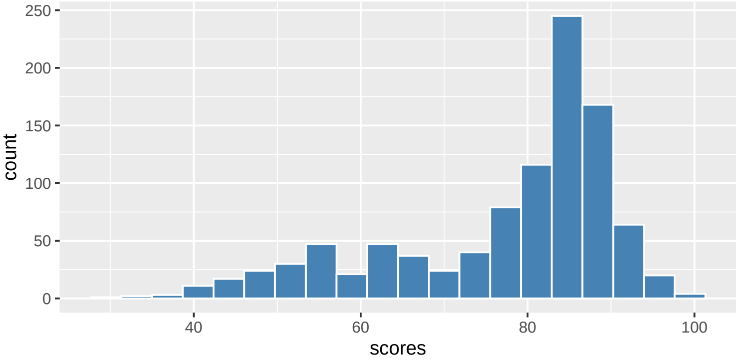


# Effect of Sample Size

Suppose we have a class of 1000 students with the following score distribution

# Effect of Sample Size

Suppose we have a class of 1000 students with the following score distribution

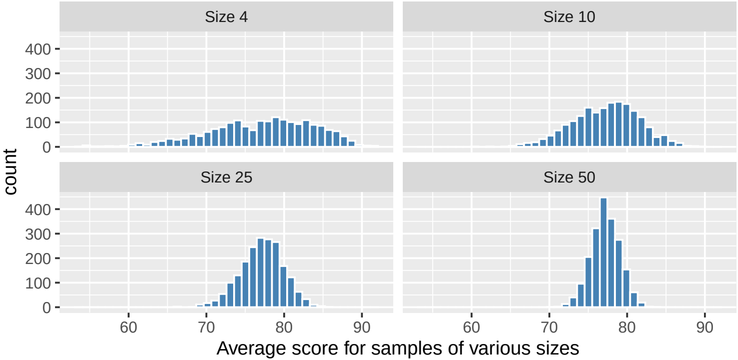


# Effect of Sample Size II

What happens to the distribution of sample means as we increase the size of each sample (keeping the number of samples drawn constant)?

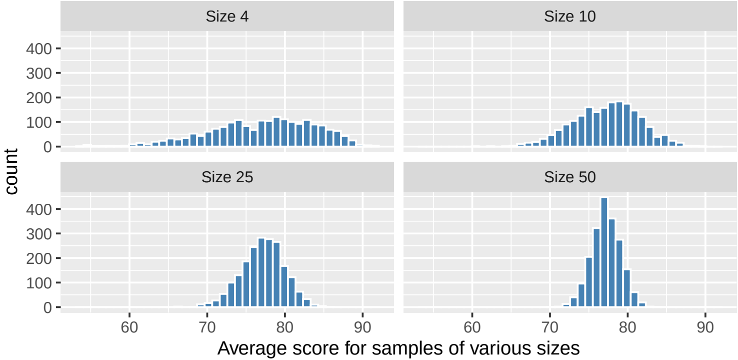
# Effect of Sample Size II

What happens to the distribution of sample means as we increase the size of each sample (keeping the number of samples drawn constant)?



# Effect of Sample Size II

What happens to the distribution of sample means as we increase the size of each sample (keeping the number of samples drawn constant)?




- As sample size increases, sampling distribution becomes **more** Normal, with **decreasing** variance

# The Central Limit Theorem

## Theorem

*Suppose an SRS of size n is drawn from a population with mean $\mu$ and standard deviation $\sigma$. When n is large, the sample mean $\bar{x}$ is approximately Normally distributed, with mean $\mu$ and standard deviation $\frac{\sigma}{\sqrt{n}}$.*

# The Central Limit Theorem

**Theorem**

*Suppose an SRS of size n is drawn from a population with mean $\mu$ and standard deviation $\sigma$. When n is large, the sample mean $\bar{x}$ is approximately Normally distributed, with mean $\mu$ and standard deviation $\frac{\sigma}{\sqrt{n}}$.*

A proof of the CLT requires more advanced techniques in probability (See Math 391). But intuitively. . .

# The Central Limit Theorem

## Theorem

*Suppose an SRS of size n is drawn from a population with mean $\mu$ and standard deviation $\sigma$. When n is large, the sample mean $\bar{x}$ is approximately Normally distributed, with mean $\mu$ and standard deviation $\frac{\sigma}{\sqrt{n}}$.*

A proof of the CLT requires more advanced techniques in probability (See Math 391). But intuitively. . .

> *A sample mean is obtained by adding together INDEPENDENT values from the population.*
>
> *In order to get a very large or very small value, nearly ALL of the independent values need to be extreme.*
>
> *To get a moderate value, many can be extreme in the opposite direction, or many can be moderate (or several variations in between).*
>
> *There are more ways to obtain moderate values in an average than to obtain extreme values*

## Implications for Statistics

- **Regardless** of the underlying population distribution, when sample size is large, the distribution of sample means is predictable, and variance in means descreases as sample size increases

## Implications for Statistics

- **Regardless** of the underylying population distribution, when sample size is large, the distribution of sample means is predictable, and variance in means descreases as sample size increases
- We can use properties of the Normal distribution to determine probabilities of obtaining extreme sample statistics

# Implications for Statistics

- **Regardless** of the underlying population distribution, when sample size is large, the distribution of sample means is predictable, and variance in means descreases as sample size increases
- We can use properties of the Normal distribution to determine probabilities of obtaining extreme sample statistics
- Statistical inference can be performed using theoretical density functions, in addition to using simulation and bootstrapping

# Implications for Statistics

- **Regardless** of the underlying population distribution, when sample size is large, the distribution of sample means is predictable, and variance in means descreases as sample size increases
- We can use properties of the Normal distribution to determine probabilities of obtaining extreme sample statistics
- Statistical inference can be performed using theoretical density functions, in addition to using simulation and bootstrapping