Continuous Variables and the Normal Distribution

Prof. Wells

STA 209, 4/12/23

Outline

In this lecture, we will. . .

Outline

In this lecture, we will. . .

- Define and explore continuous random variables
- Investigate properties of the Normal Distribution

Section 1

Continuous Random Variables

The Distribution of a Continuous Variable

- A continuous random variable is one that any value in an interval of real numbers.
  Example: the temperature Noyce 2401 at 4pm on 4/12

The Distribution of a Continuous Variable

- A continuous random variable is one that any value in an interval of real numbers.
  Example: the temperature Noyce 2401 at 4pm on 4/12
- Previously, we described the distribution of *discrete* random variables by listing the probabilities of taking each possible values.

The Distribution of a Continuous Variable

- A continuous random variable is one that any value in an interval of real numbers.
  Example: the temperature Noyce 2401 at 4pm on 4/12
- Previously, we described the distribution of *discrete* random variables by listing the probabilities of taking each possible values.
  - But for continuous variables, there are too many possible values to provide a meaningful probability for each.

The Distribution of a Continuous Variable

- A continuous random variable is one that any value in an interval of real numbers.
  Example: the temperature Noyce 2401 at 4pm on 4/12
- Previously, we described the distribution of *discrete* random variables by listing the probabilities of taking each possible values.
  - But for continuous variables, there are too many possible values to provide a meaningful probability for each.
- Instead, we describe the probabilities that continuous variables are in certain ranges of values, specified by a **density curves**

The Distribution of a Continuous Variable

- A continuous random variable is one that any value in an interval of real numbers. *Example: the temperature Noyce 2401 at 4pm on 4/12*
- Previously, we described the distribution of *discrete* random variables by listing the probabilities of taking each possible values.
  - But for continuous variables, there are too many possible values to provide a meaningful probability for each.
- Instead, we describe the probabilities that continuous variables are in certain ranges of values, specified by a **density curves**
- The **density curve** for a continuous random variable X is the function f so that...

The Distribution of a Continuous Variable

- A continuous random variable is one that any value in an interval of real numbers.
  Example: the temperature Noyce 2401 at 4pm on 4/12
- Previously, we described the distribution of *discrete* random variables by listing the probabilities of taking each possible values.
  - But for continuous variables, there are too many possible values to provide a meaningful probability for each.
- Instead, we describe the probabilities that continuous variables are in certain ranges of values, specified by a **density curves**
- The **density curve** for a continuous random variable X is the function f so that...
  - The values of the function are always non-negative

The Distribution of a Continuous Variable

- A continuous random variable is one that any value in an interval of real numbers. *Example: the temperature Noyce 2401 at 4pm on 4/12*
- Previously, we described the distribution of *discrete* random variables by listing the probabilities of taking each possible values.
  - But for continuous variables, there are too many possible values to provide a meaningful probability for each.
- Instead, we describe the probabilities that continuous variables are in certain ranges of values, specified by a **density curves**
- The **density curve** for a continuous random variable X is the function f so that...
  - The values of the function are always non-negative
  - The total area under the function is 1

The Distribution of a Continuous Variable

- A continuous random variable is one that any value in an interval of real numbers.
  Example: the temperature Noyce 2401 at 4pm on 4/12
- Previously, we described the distribution of *discrete* random variables by listing the probabilities of taking each possible values.
  - But for continuous variables, there are too many possible values to provide a meaningful probability for each.
- Instead, we describe the probabilities that continuous variables are in certain ranges of values, specified by a **density curves**
- The **density curve** for a continuous random variable X is the function f so that...
  - The values of the function are always non-negative
  - The total area under the function is 1
  - The area under the density curve on an interval is the probability that the variable is in that interval.

The Distribution of a Continuous Variable

- A continuous random variable is one that any value in an interval of real numbers.
  Example: the temperature Noyce 2401 at 4pm on 4/12
- Previously, we described the distribution of *discrete* random variables by listing the probabilities of taking each possible values.
  - But for continuous variables, there are too many possible values to provide a meaningful probability for each.
- Instead, we describe the probabilities that continuous variables are in certain ranges of values, specified by a **density curves**
- The **density curve** for a continuous random variable X is the function f so that...
  - The values of the function are always non-negative
  - The total area under the function is 1
  - The area under the density curve on an interval is the probability that the variable is in that interval.
- Using the language of calculus,

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

Density Curve

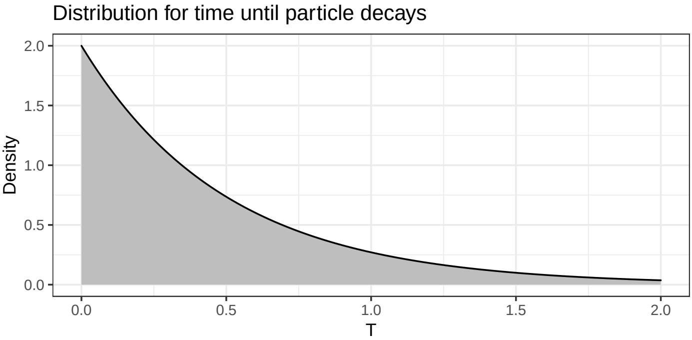
- The density curve for the number of seconds T until a radioactive particle decays is:

$$f(t) = 2e^{-2t} \qquad \text{for } t \geq 0$$

Density Curve

- The density curve for the number of seconds T until a radioactive particle decays is:

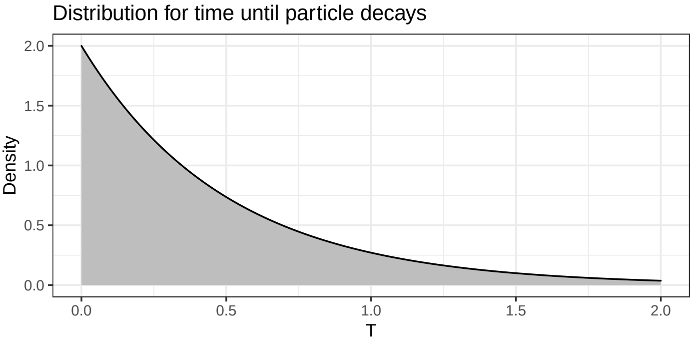
$$f(t) = 2e^{-2t} \qquad \text{for } t \geq 0$$



Density Curve

- The density curve for the number of seconds T until a radioactive particle decays is:

$$f(t) = 2e^{-2t} \qquad \text{for } t \geq 0$$

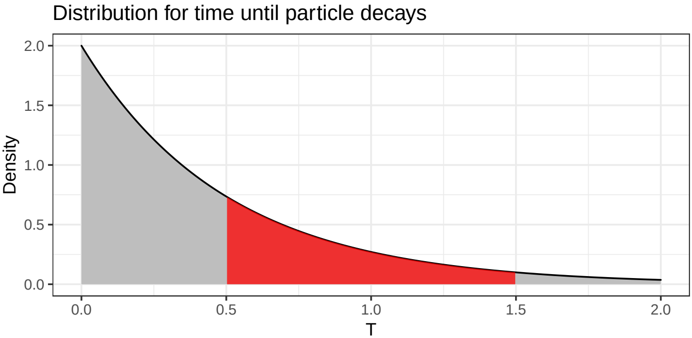


- The probability that it takes between 0.5 and 1.5 seconds to decay is the area under the curve between 0.5 and 1.5. $P(0.5 < T < 1.5) = \int_{0.5}^{1.5} 2e^{-2t}\, dt =$

Density Curve

- The density curve for the number of seconds T until a radioactive particle decays is:

$$f(t) = 2e^{-2t} \qquad \text{for } t \geq 0$$



- The probability that it takes between 0.5 and 1.5 seconds to decay is the area under the curve between 0.5 and 1.5. $P(0.5 < T < 1.5) = \int_{0.5}^{1.5} 2e^{-2t}\, dt = 0.34$

Mean and Standard Deviation for Continuous Variables

- Just as with discrete random variables, we can define the mean, variance and standard deviations of continuous variables.

Mean and Standard Deviation for Continuous Variables

- Just as with discrete random variables, we can define the mean, variance and standard deviations of continuous variables.
    - But we cannot use the same definition as before (the sum of values, weighted by the probability of each value)

Mean and Standard Deviation for Continuous Variables

- Just as with discrete random variables, we can define the mean, variance and standard deviations of continuous variables.
  - But we cannot use the same definition as before (the sum of values, weighted by the probability of each value)
  - Note that for any real number c, $P(X = c) = 0$. (Why?)

Mean and Standard Deviation for Continuous Variables

- Just as with discrete random variables, we can define the mean, variance and standard deviations of continuous variables.
  - But we cannot use the same definition as before (the sum of values, weighted by the probability of each value)
  - Note that for any real number c, $P(X = c) = 0$. (Why?)
- Instead, we use the *integral* from calculus to define the mean and variance:

$$E[X] = \int xf(x)\,dx \qquad \mathrm{Var}(X) = \int (x-\mu)^2 f(x)\,dx \qquad \mathrm{SD}(X) = \sqrt{\mathrm{Var}(X)}$$

Mean and Standard Deviation for Continuous Variables

- Just as with discrete random variables, we can define the mean, variance and standard deviations of continuous variables.
  - But we cannot use the same definition as before (the sum of values, weighted by the probability of each value)
  - Note that for any real number c, $P(X = c) = 0$. (Why?)
- Instead, we use the *integral* from calculus to define the mean and variance:

$$E[X] = \int xf(x)\,dx \qquad \mathrm{Var}(X) = \int (x-\mu)^2 f(x)\,dx \qquad \mathrm{SD}(X) = \sqrt{\mathrm{Var}(X)}$$

  - These **integrals** are tools to meaningfully average infinitely many values (but we won't compute any integrals in this class)

Mean and Standard Deviation for Continuous Variables

- Just as with discrete random variables, we can define the mean, variance and standard deviations of continuous variables.
  - But we cannot use the same definition as before (the sum of values, weighted by the probability of each value)
  - Note that for any real number c, $P(X = c) = 0$. (Why?)
- Instead, we use the *integral* from calculus to define the mean and variance:

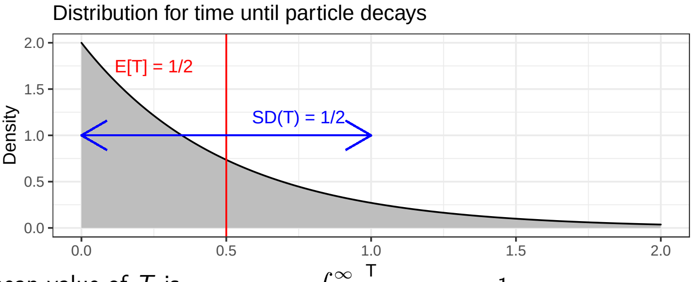
$$E[X] = \int xf(x)\,dx \qquad \mathrm{Var}(X) = \int (x-\mu)^2 f(x)\,dx \qquad \mathrm{SD}(X) = \sqrt{\mathrm{Var}(X)}$$

  - These **integrals** are tools to meaningfully average infinitely many values (but we won't compute any integrals in this class)
- As with discrete variables, the mean of a continuous variables represents its typical value. The standard deviation represents the typical size of deviations from the mean.

Density Curve

- The density curve for the number of seconds T until a radioactive particle decays is:

$$f(t) = 2e^{-2t} \qquad \text{for } t \geq 0$$



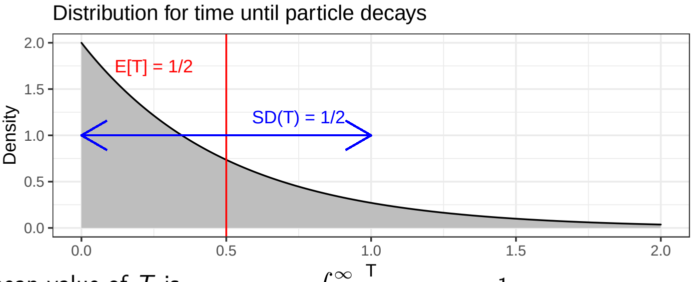
- The mean value of T is

$$E[T] = \int_0^\infty t \cdot 2e^{-2t}\, dt = \frac{1}{2}$$

Density Curve

- The density curve for the number of seconds T until a radioactive particle decays is:

$$f(t) = 2e^{-2t} \qquad \text{for } t \geq 0$$



- The mean value of T is
$$E[T] = \int_0^\infty t \cdot 2e^{-2t}\, dt = \frac{1}{2}$$

- The variance and standard deviation of T are
$$\mathrm{Var}(T) = \int_0^\infty \left(t - \frac{1}{2}\right)^2 \cdot 2e^{-2t}\, dt = \frac{1}{4} \qquad \mathrm{SD}(T) = \sqrt{\mathrm{Var}(T)} = \frac{1}{2}$$

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

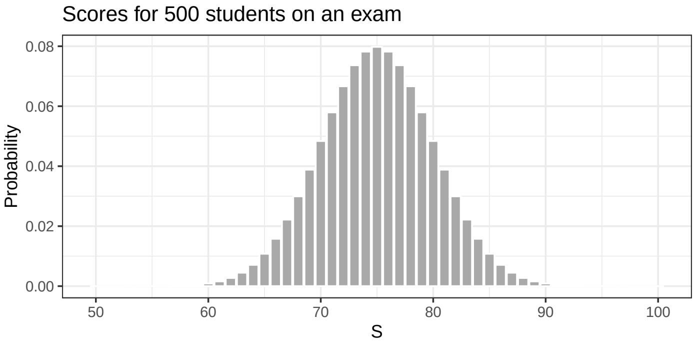
If a discrete variable takes a large number of values which are close together, we can often approximate it using a continuous variable.

- 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.

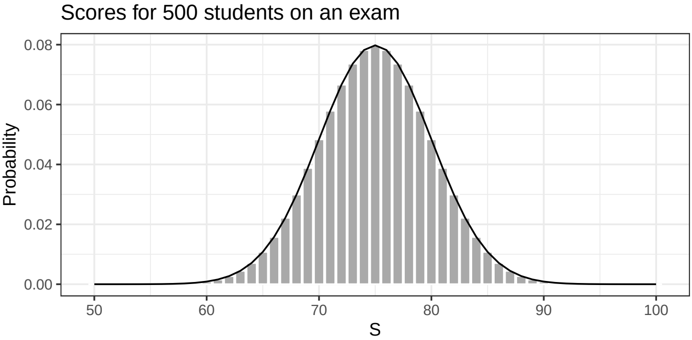
- Suppose 500 students take a standardized exam, with mean 75 points. The distribution for the score S of a randomly chosen student is:



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:



Section 2

The Normal Distribution

The Normal Distribution

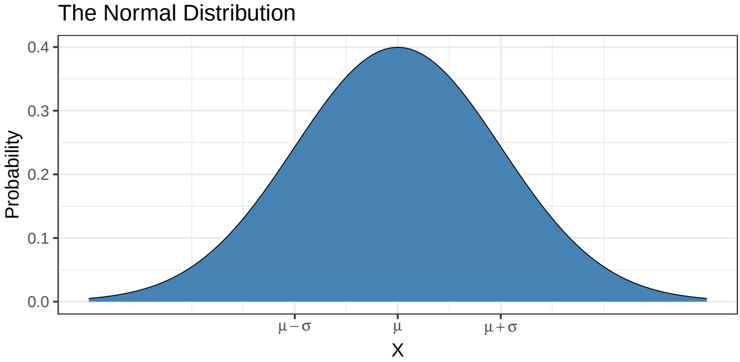
- The general Normal density curve with mean μ and standard deviation σ is given by the formula

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

The Normal Distribution

- The general Normal density curve with mean μ and standard deviation σ is given by the formula

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



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

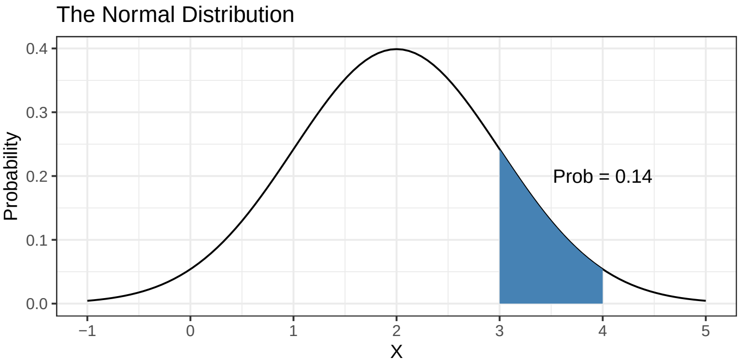
Recall that for a random variable which has a **continuous** distribution, we find probabilities by looking at areas under the density curve.

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
```

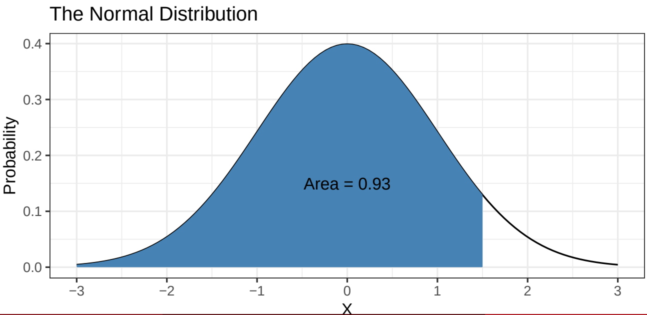
Calculating Normal Probabilities in R

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

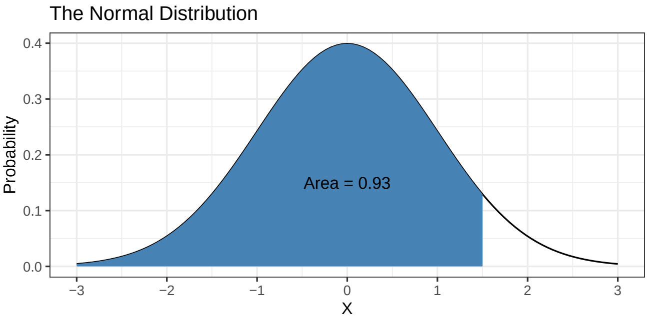
The pnorm function lets us compute cumulative areas (i.e. all area to the left of a given value). But how do we compute the area between two values?

- 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

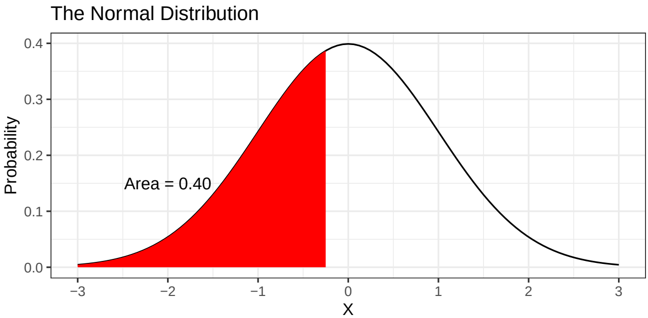
The `pnorm` function lets us compute cumulative areas (i.e. all area to the left of a given value). But how do we compute the area between two values?

- 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

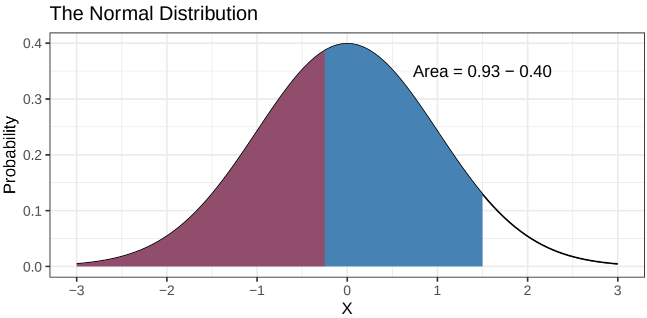
The `pnorm` function lets us compute cumulative areas (i.e. all area to the left of a given value). But how do we compute the area between two values?

- 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

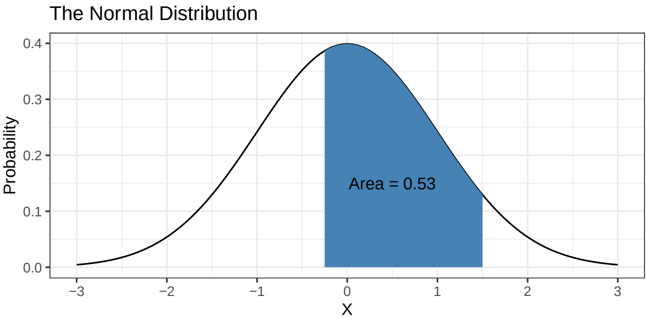
The pnorm function lets us compute cumulative areas (i.e. all area to the left of a given value). But how do we compute the area between two values?

- 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.

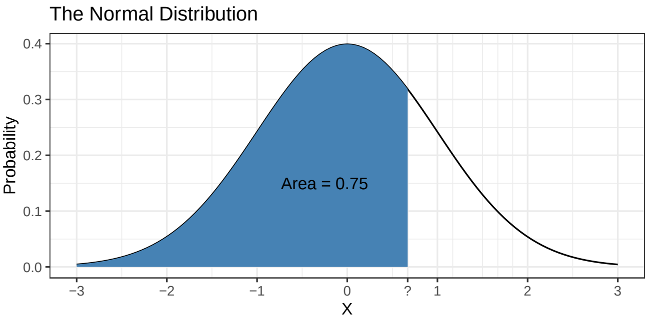
```
pnorm(q =1.5 , mean =0 , sd =1 ) - pnorm(q = -.25 , mean =0 , sd =1 )
```

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



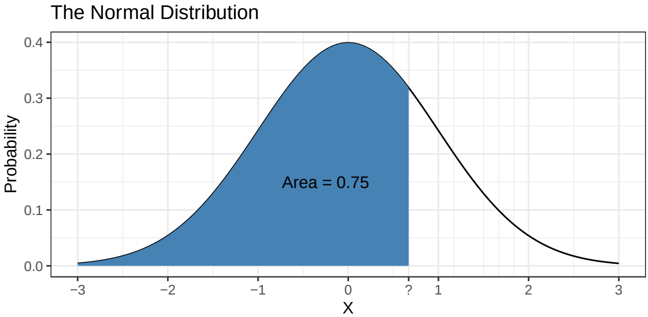
Finding Quantiles

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



Finding Quantiles

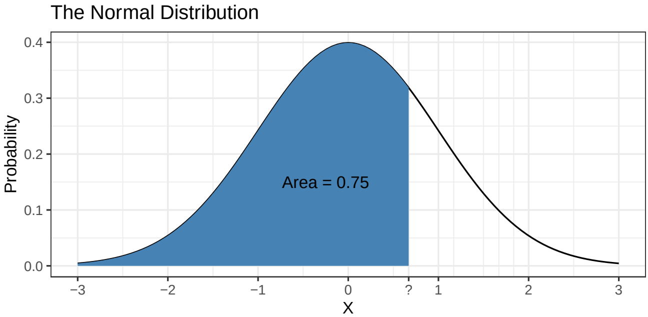
Suppose we instead have the opposite problem: We want to FIND the value of X with a given cumulative area.



- 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.

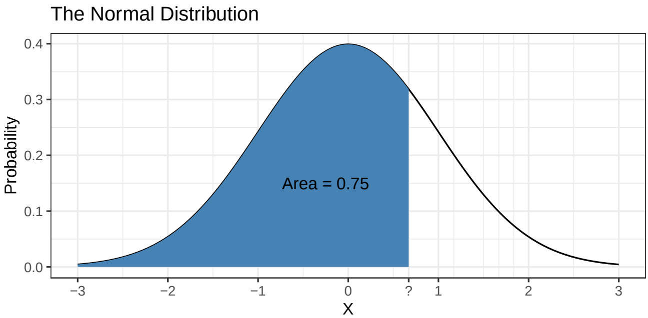


- That is, we want to find the .75 quantile (i.e. the 75th percentile)

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)

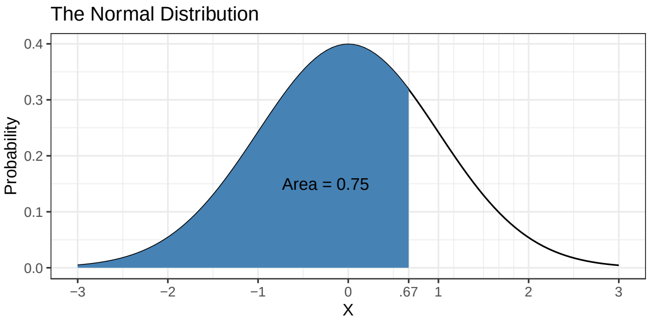
R has a built-in function for that too! `qnorm(p =... , mean =... , sd =... )`

```
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 )
```

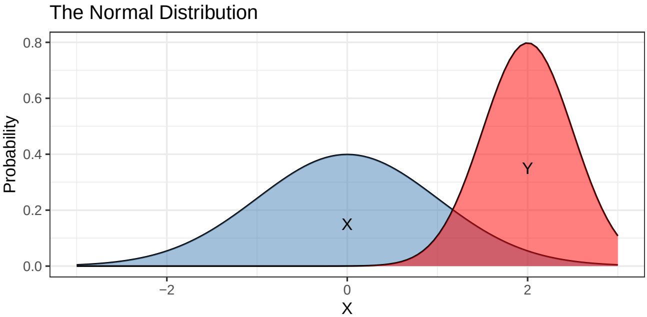
```
## [1] 0.6744898
```

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 $\sigma = .25$.

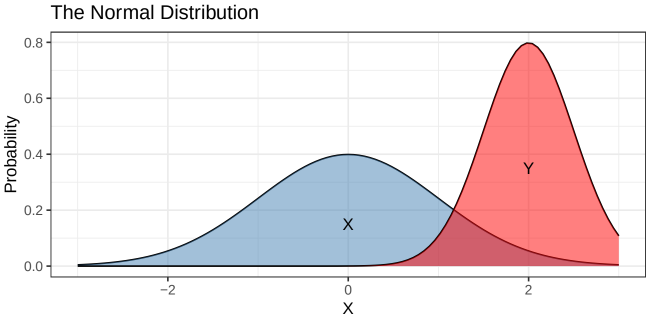
Scale and Translation Invariance

- 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

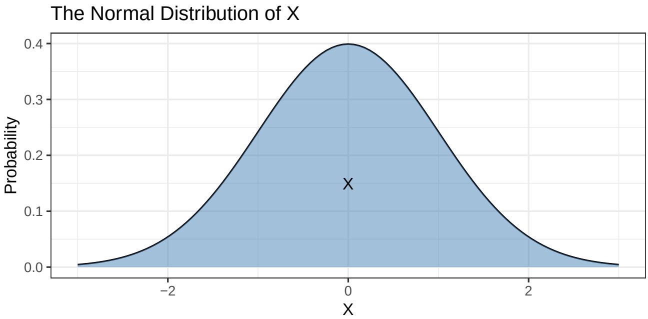
- Consider a Normal variable X with $\mu = 0$ and $\sigma = 1$, and another Normal variable Y with mean $\mu = 2$ and $\sigma = .25$.



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

Scale and Translation Invariance

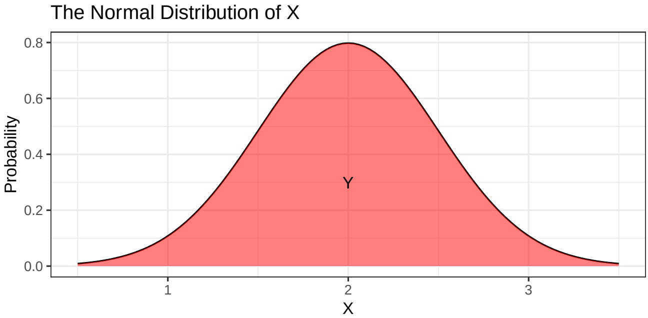
- Consider a Normal variable X with $\mu = 0$ and $\sigma = 1$, and another Normal variable Y with mean $\mu = 2$ and $\sigma = .25$.



- 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 μ and standard deviation σ. 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 μ and standard deviation σ. 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 μ and standard deviation σ. 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 μ and standard deviation σ. 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 μ and σ)