Economics 103 – Statistics for Economists

Francis J. DiTraglia

University of Pennsylvania

Lecture #9 – Discrete RVs III

Variance and Standard Deviation of a Random Variable

Binomial Random Variable

Variance and Standard Deviation of a RV

The Defs are the same for continuous RVs, but the method of calculating will differ.

Variance (Var)

$$\sigma^2 = Var(X) = E\left[(X-\mu)^2\right] = E\left[(X-E[X])^2\right]$$

Standard Deviation (SD)

$$\sigma = \sqrt{\sigma^2} = SD(X)$$

Key Point

Variance and std. dev. are *expectations of functions of a RV*

It follows that:

1. Variance and SD are constants
2. To derive facts about them you can use the facts you know about expected value

How To Calculate Variance for Discrete RV?

Remember: it's just a function of X!

Recall that $\mu = E[X] = \sum_{\text{all } x} x p(x)$

$$Var(X) = E\left[(X-\mu)^2\right] = \sum_{\text{all } x} (x-\mu)^2 p(x)$$

Shortcut Formula For Variance

This is *not* the definition, it's a shortcut for doing calculations:

$$Var(X) = E\left[(X-\mu)^2\right] = E[X^2] - (E[X])^2$$

We'll prove this in an upcoming lecture.

Example: The Shortcut Formula

Let $X \sim \text{Bernoulli}(1/2)$. Calculate $Var(X)$.

$$
\begin{aligned}
E[X] &= 0 \times 1/2 + 1 \times 1/2 = 1/2 \\
E[X^2] &= 0^2 \times 1/2 + 1^2 \times 1/2 = 1/2
\end{aligned}
$$

$$E[X^2] - (E[X])^2 = 1/2 - (1/2)^2 = 1/4$$

Variance of Bernoulli RV – via the Shortcut Formula

Step 1 – $E[X]$

$$\mu = E[X] = \sum_{x \in \{0,1\}} p(x) \cdot x = (1-p) \cdot 0 + p \cdot 1 = p$$

Step 2 – $E[X^2]$

$$E[X^2] = \sum_{x \in \{0,1\}} x^2 p(x) = 0^2(1-p) + 1^2 p = p$$

Step 3 – Combine with Shortcut Formula

$$\sigma^2 = Var[X] = E[X^2] - (E[X])^2 = p - p^2 = p(1-p)$$

Variance of a Linear Transformation

$$
\begin{aligned}
Var(a+bX) &= E\left[\{(a+bX)-E(a+bX)\}^2\right] \\
&= E\left[\{(a+bX)-(a+bE[X])\}^2\right] \\
&= E\left[(bX-bE[X])^2\right] \\
&= E[b^2(X-E[X])^2] \\
&= b^2E[(X-E[X])^2] \\
&= b^2Var(X) = b^2\sigma^2
\end{aligned}
$$

The key point here is that variance is defined in terms of expectation and expectation is linear.

Variance and SD are *NOT* Linear

$$
\begin{aligned}
Var(a + bX) &= b^2\sigma^2 \\
SD(a + bX) &= |b|\sigma
\end{aligned}
$$

These should look familiar from the related results for sample variance and std. dev. that you worked out on an earlier problem set.

Binomial Random Variable

Let $X =$ the sum of n independent Bernoulli trials, each with probability of success p. Then we say that: $X \sim \text{Binomial}(n, p)$

Parameters

$p =$ probability of "success," $n = \#$ of trials

Support

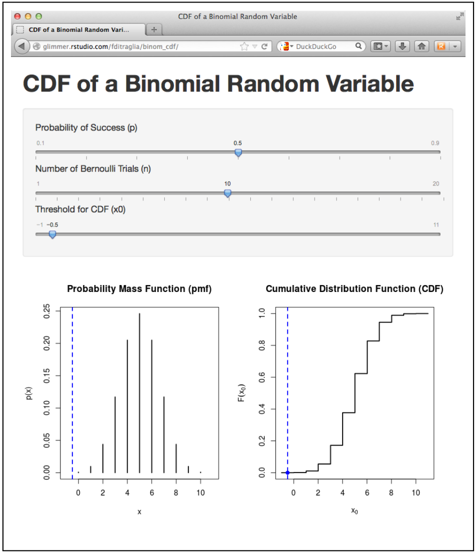
$\{0, 1, 2, \ldots, n\}$

Probability Mass Function (pmf)

$$p(x) = \binom{n}{x} p^x (1-p)^{n-x}$$

http://fditraglia.shinyapps.io/binom_cdf/

Try playing around with all three sliders. If you set the second to 1 you get a Bernoulli.



Where does the Binomial pmf come from?

Question

Suppose we flip a fair coin 3 times. What is the probability that we get exactly 2 heads?

Answer

Three basic outcomes make up this event: $\{HHT, HTH, THH\}$, each has probability $1/8 = 1/2 \times 1/2 \times 1/2$. Basic outcomes are mutually exclusive, so sum to get $3/8 = 0.375$

Where does the Binomial pmf come from?

Question

Suppose we flip an *unfair* coin 3 times, where the probability of heads is $1/3$. What is the probability that we get exactly 2 heads?

Answer

No longer true that *all* basic outcomes are equally likely, but those with exactly two heads *still are*

$$
\begin{aligned}
P(HHT) &= (1/3)^2(1-1/3) = 2/27 \\
P(THH) &= 2/27 \\
P(HTH) &= 2/27
\end{aligned}
$$

Summing gives $2/9 \approx 0.22$

Where does the Binomial pmf come from?

Starting to see a pattern?

Suppose we flip an unfair coin *4* times, where the probability of heads is $1/3$. What is the probability that we get exactly 2 heads?

HHTT	TTHH
HTHT	THTH
HTTH	THHT

Six equally likely, mutually exclusive basic outcomes make up this event:

$$\binom{4}{2}(1/3)^2(2/3)^2$$

R Commands for Binomial(n, p) RV

Probability Mass Function

`dbinom(x, size, prob)`, where `size` is n and `prob` is p

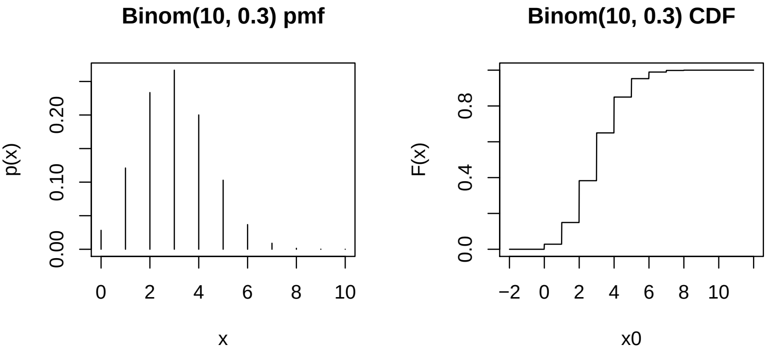
Cumulative Distribution Function

`pbinom(q, size, prob)`, where `q` is x_0, `size` is n and `prob` is p

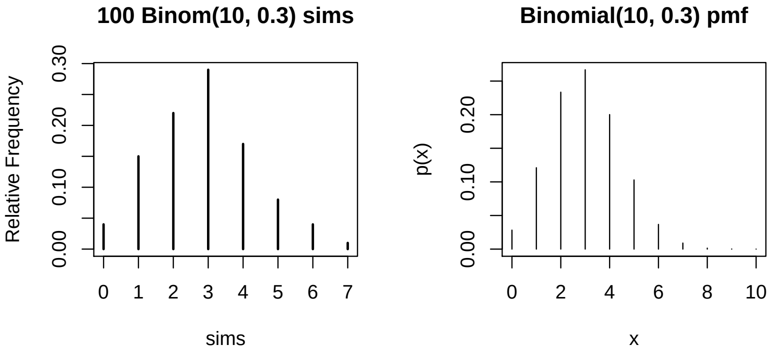
Make Random Draws

`rbinom(n, size, prob)`, where `n` is the number of draws, `size` is n and `prob` is p

```r
x <- 0:10
px <- dbinom(x, size = 10, prob = 0.3)
x0 <- seq(from = -2, to = 12, by = 0.01)
Fx <- pbinom(x0, size = 10, prob = 0.3)
par(mfrow = c(1, 2))
plot(x, px, type = 'h', ylab = 'p(x)', main = 'Binom(10, 0.3) pmf')
plot(x0, Fx, type = 'l', ylab = 'F(x)', main = 'Binom(10, 0.3) CDF')
```



```r
set.seed(5545)
sims <- rbinom(100, size = 10, prob = 0.3)
par(mfrow = c(1, 2))
rel_freq <- prop.table(table(sims))
plot(rel_freq, main = '100 Binom(10, 0.3) sims',
     ylab = 'Relative Frequency')
plot(x, px, type = 'h', ylab = 'p(x)', main = 'Binomial(10, 0.3) pmf')
```



Lecture #10 – Discrete RVs IV

Joint vs. Marginal Probability Mass Functions

Conditional Probability Mass Function & Independence

Expectation of a Function of Two Discrete RVs, Covariance

Linearity of Expectation Reprise, Properties of Binomial RV

Multiple RVs *at once* - Definition of Joint PMF

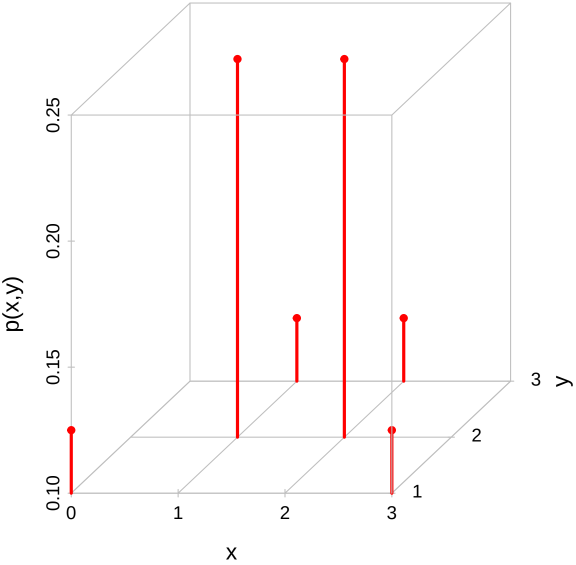
Let X and Y be discrete random variables. The joint probability mass function $p_{XY}(x, y)$ gives the probability of each pair of realizations (x, y) in the support:

$$\boxed{p_{XY}(x, y) = P(X = x \cap Y = y)}$$

Example: Joint PMF in Tabular Form

		Y		
		1	2	3
X	0	1/8	0	0
	1	0	1/4	1/8
	2	0	1/4	1/8
	3	1/8	0	0

Plot of Joint PMF



What is $p_{XY}(1,2)$?

		Y		
		1	2	3
X	0	1/8	0	0
	1	0	1/4	1/8
	2	0	1/4	1/8
	3	1/8	0	0

$$
\begin{aligned}
p_{XY}(1,2) &= P(X = 1 \cap Y = 2) = 1/4 \\
p_{XY}(2,1) &= P(X = 2 \cap Y = 1) = 0
\end{aligned}
$$

Properties of Joint PMF

1. $0 \leq p_{XY}(x, y) \leq 1$ for any pair (x, y)
2. The sum of $p_{XY}(x, y)$ over all pairs (x, y) in the support is 1:

$$\sum_x \sum_y p(x, y) = 1$$

Joint versus Marginal PMFs

Joint PMF

$$p_{XY}(x,y) = P(X = x \cap Y = y)$$

Marginal PMFs

$$p_X(x) = P(X = x)$$
$$p_Y(y) = P(Y = y)$$

You can't calculate a joint pmf from marginals alone but you *can* calculate marginals from the joint!

Marginals from Joint

$$\boxed{p_X(x) = \sum_{\text{all } y} p_{XY}(x,y)}$$

$$\boxed{p_Y(y) = \sum_{\text{all } x} p_{XY}(x,y)}$$

Why?

$$\begin{aligned} p_Y(y) &= P(Y=y) = P\left(\bigcup_{\text{all } x} \{X = x \cap Y = y\}\right) \\ &= \sum_{\text{all } x} P(X = x \cap Y = y) = \sum_{\text{all } x} p_{XY}(x,y) \end{aligned}$$

To get the marginals sum "into the margins" of the table.

		Y			
		1	2	3	
X	0	1/8	0	0	1/8
	1	0	1/4	1/8	3/8
	2	0	1/4	1/8	3/8
	3	1/8	0	0	1/8
					1

$$
\begin{aligned}
p_X(0) &= 1/8 + 0 + 0 = 1/8 \\
p_X(1) &= 0 + 1/4 + 1/8 = 3/8 \\
p_X(2) &= 0 + 1/4 + 1/8 = 3/8 \\
p_X(3) &= 1/8 + 0 + 0 = 1/8
\end{aligned}
$$

What is $p_Y(2)$?

		Y			
		1	2	3	
X	0	1/8	0	0	
	1	0	1/4	1/8	
	2	0	1/4	1/8	
	3	1/8	0	0	
		1/4	1/2	1/4	1

$$
\begin{aligned}
p_Y(1) &= 1/8 + 0 + 0 + 1/8 = 1/4 \\
p_Y(2) &= 0 + 1/4 + 1/4 + 0 = 1/2 \\
p_Y(3) &= 0 + 1/8 + 1/8 + 0 = 1/4
\end{aligned}
$$

Definition of Conditional PMF

How does the distribution of y change with x?

$$p_{Y|X}(y|x) = P(Y = y|X = x) = \frac{P(Y = y \cap X = x)}{P(X = x)} = \frac{p_{XY}(x, y)}{p_X(x)}$$

Conditional PMF of Y given $X = 2$

		Y = 1	2	3	
X	0	1/8	0	0	1/8
	1	0	1/4	1/8	3/8
	2	0	1/4	1/8	3/8
	3	1/8	0	0	1/8

$$p_{Y|X}(1|2) = \frac{p_{XY}(2,1)}{p_X(2)} = \frac{0}{3/8} = 0$$

$$p_{Y|X}(2|2) = \frac{p_{XY}(2,2)}{p_X(2)} = \frac{1/4}{3/8} = 2/3$$

$$p_{Y|X}(3|2) = \frac{p_{XY}(2,3)}{p_X(2)} = \frac{1/8}{3/8} = 1/3$$

What is $p_{X|Y}(1|2)$?

		Y			
		1	2	3	
X	0	1/8	0	0	
	1	0	1/4	1/8	
	2	0	1/4	1/8	
	3	1/8	0	0	
		1/4	1/2	1/4	

$$p_{X|Y}(1|2) = \frac{p_{XY}(1,2)}{p_Y(2)} = \frac{1/4}{1/2} = 1/2$$

Similarly:

$$p_{X|Y}(0|2) = 0, \quad p_{X|Y}(2|2) = 1/2, \quad p_{X|Y}(3|2) = 0$$

Independent RVs: Joint Equals Product of Marginals

Definition

Two discrete RVs are independent if and only if

$$p_{XY}(x,y) = p_X(x)p_Y(y)$$

for all pairs (x,y) in the support.

Equivalent Definition

$$p_{Y|X}(y|x) = p_Y(y) \text{ and } p_{X|Y}(x|y) = p_X(x)$$

for all pairs (x,y) in the support.

Are X and Y Independent?

(A = YES, B = NO)

		Y			
		1	2	3	
X	0	1/8	0	0	1/8
	1	0	1/4	1/8	3/8
	2	0	1/4	1/8	3/8
	3	1/8	0	0	1/8
		1/4	1/2	1/4	

$$\begin{aligned} p_{XY}(2,1) &= 0 \\ p_X(2) \times p_Y(1) &= (3/8) \times (1/4) \neq 0 \end{aligned}$$

Therefore X and Y are *not* independent.

Expectation of Function of Two Discrete RVs

$$\boxed{E[g(X,Y)] = \sum_x \sum_y g(x,y) p_{XY}(x,y)}$$

Some Extremely Important Examples

Same For Continuous Random Variables

Let $\mu_X = E[X], \mu_Y = E[Y]$

Covariance

$$\sigma_{XY} = Cov(X, Y) = E[(X - \mu_X)(Y - \mu_Y)]$$

Correlation

$$\rho_{XY} = Corr(X, Y) = \frac{\sigma_{XY}}{\sigma_X \sigma_Y}$$

Shortcut Formula for Covariance

Much easier for calculating:

$$\boxed{Cov(X, Y) = E[XY] - E[X]E[Y]}$$

I'll mention this again in a few slides...

Calculating $Cov(X, Y)$

		Y			
		1	2	3	
X	0	1/8	0	0	1/8
	1	0	1/4	1/8	3/8
	2	0	1/4	1/8	3/8
	3	1/8	0	0	1/8
		1/4	1/2	1/4	

$$E[X] = 3/8 + 2 \times 3/8 + 3 \times 1/8 = 3/2$$

$$E[Y] = 1/4 + 2 \times 1/2 + 3 \times 1/4 = 2$$

$$\begin{aligned} E[XY] &= 1/4 \times (2+4) + 1/8 \times (3+6+3) \\ &= 3 \end{aligned}$$

$$\begin{aligned} Cov(X, Y) &= E[XY] - E[X]E[Y] \\ &= 3 - 3/2 \times 2 = 0 \\ Corr(X, Y) &= Cov(X, Y) / [SD(X)SD(Y)] = 0 \end{aligned}$$

Hence, zero covariance (correlation) does *not* imply independence!

Zero Covariance versus Independence

While zero covariance (correlation) *does not* imply independence,
independence *does* imply zero covariance (correlation).

You will prove this in an extension problem...

Linearity of Expectation, Again

Holds for Continuous RVs as well, but different proof.

In general $E[g(X, Y)] \neq g(E[X], E[Y])$. But if g is linear, then:

$$\boxed{E[aX + bY + c] = aE[X] + bE[Y] + c}$$

where X, Y are random variables and a, b, c are constants.

You will prove this as a review problem...

Application: Proof of Shortcut Formula for Variance

By the Linearity of Expectation,

$$
\begin{aligned}
Var(X) &= E[(X-\mu)^2] = E[X^2 - 2\mu X + \mu^2] \\
&= E[X^2] - 2\mu E[X] + \mu^2 \\
&= E[X^2] - 2\mu^2 + \mu^2 \\
&= E[X^2] - \mu^2
\end{aligned}
$$

Expected Value of Sum = Sum of Expected Values

Repeatedly applying the linearity of expectation,

$$E[X_1 + X_2 + \ldots + X_n] = E[X_1] + E[X_2] + \ldots + E[X_n]$$

regardless of how the RVs $X_1, \ldots, X_n$ are related to each other. In particular it doesn't matter if they're dependent or independent.

Independent and Identically Distributed (iid) RVs

Example

$X_1, X_2, \ldots X_n \sim$ iid Bernoulli(p)

Independent

Realization of one of the RVs gives no information about the others.

Identically Distributed

Each X_i is the same kind of RV, with the same values for any parameters. (Hence same pmf, cdf, mean, variance, etc.)

Recall: Binomial(n, p) Random Variable

Definition

Sum of n independent Bernoulli RVs, each with probability of "success," i.e. 1, equal to p

Using Our New Notation

Let $X_1, X_2, \ldots, X_n \sim$ iid Bernoulli(p), $Y = X_1 + X_2 + \ldots + X_n$.
Then $Y \sim$ Binomial(n, p).

Expected Value of Binomial RV

Use the fact that a Binomial(n, p) RV is defined as the sum of n iid Bernoulli(p) Random Variables and the Linearity of Expectation:

$$
\begin{aligned}
E[Y] &= E[X_1 + X_2 + \ldots + X_n] = E[X_1] + E[X_2] + \ldots + E[X_n] \\
&= p + p + \ldots + p \\
&= np
\end{aligned}
$$

Variance of a Sum $\neq$ Sum of Variances!

$$
\begin{aligned}
Var(aX + bY) &= E\left[\{(aX + bY) - E[aX + bY]\}^2\right] \\
&\vdots \\
&= a^2 Var(X) + b^2 Var(Y) + 2abCov(X, Y)
\end{aligned}
$$

You'll fill in the missing steps as an extension problem...

Since $\sigma_{XY} = \rho\sigma_X\sigma_Y$, this is sometimes written as:

$$\boxed{Var(aX + bY) = a^2\sigma_X^2 + b^2\sigma_Y^2 + 2ab\rho\sigma_X\sigma_Y}$$

Independence $\Rightarrow$ $Var(X+Y) = Var(X) + Var(Y)$

X and Y independent $\Rightarrow Cov(X,Y) = 0$. Hence:

$$\begin{aligned} Var(X+Y) &= Var(X) + Var(Y) + 2Cov(X,Y) \\ &= Var(X) + Var(Y) \end{aligned}$$

Also true for three or more RVs

If $X_1, X_2, \ldots, X_n$ are independent, then

$$Var(X_1 + X_2 + \ldots X_n) = Var(X_1) + Var(X_2) + \ldots + Var(X_n)$$

Crucial Distinction

Expected Value

Always true that

$$E[X_1 + X_2 + \ldots + X_n] = E[X_1] + E[X_2] + \ldots + E[X_n]$$

Variance

Not true in general that

$$Var[X_1 + X_2 + \ldots + X_n] = Var[X_1] + Var[X_2] + \ldots + Var[X_n]$$

except in the special case where $X_1, \ldots X_n$ are independent (or at least uncorrelated).

Variance of Binomial Random Variable

Definition from Sequence of Bernoulli Trials

If $X_1, X_2, \ldots, X_n \sim$ iid Bernoulli(p) then

$$Y = X_1 + X_2 + \ldots + X_n \sim \text{Binomial}(n, p)$$

Using Independence

$$\begin{aligned}
Var[Y] &= Var[X_1 + X_2 + \ldots + X_n] \\
&= Var[X_1] + Var[X_2] + \ldots + Var[X_n] \\
&= p(1-p) + p(1-p) + \ldots + p(1-p) \\
&= np(1-p)
\end{aligned}$$

Lecture #11 – Continuous RVs I

Introduction: Probability as Area

Probability Density Function (PDF)

Relating the PDF to the CDF

Calculating the Probability of an Interval

Calculating Expected Value for Continuous RVs

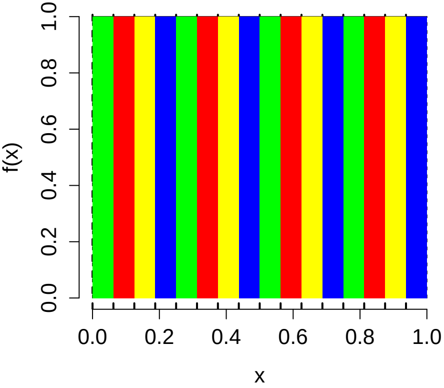
Continuous RVs – What Changes?

1. Probability Density Functions replace Probability Mass Functions
2. Integrals Replace Sums

Everything Else is Essentially Unchanged!

What is the probability of "Yellow?"

From Twister to Density – Probability as *Area*



For continuous RVs, probability is defined as *area under a curve*.

Zero area means zero probability!

Probability Density Function (PDF)

For a continuous random variable X,

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

where $f(x)$ is the *probability density function* for X.

Extremely Important

For any realization x, $P(X = x) = 0$ since $\int_a^a f(x)dx = 0$. In other words, zero area means zero probability!

For a Continuous RV, Zero Probability $\neq$ Impossible

It is *crucial* to specify the support set of a continuous RV:

- Any x outside the support set of X is *impossible*.
- Any x in the support set of X is a *possible outcome* even though $P(X = x) = 0$ for all x.

There is no way around this slightly awkward situation: it is a consequence of defining probability as the *area under a curve*.

Properties of PDFs

1. $f(x) \geq 0$ for all x in the support of X and zero otherwise.
2. $\int_{-\infty}^{\infty} f(x)\, dx = 1$

Warning: $f(x)$ is *not a probability*

Can have $f(x) > 1$ for some x as long as $\int_{-\infty}^{\infty} f(x) dx = 1$.

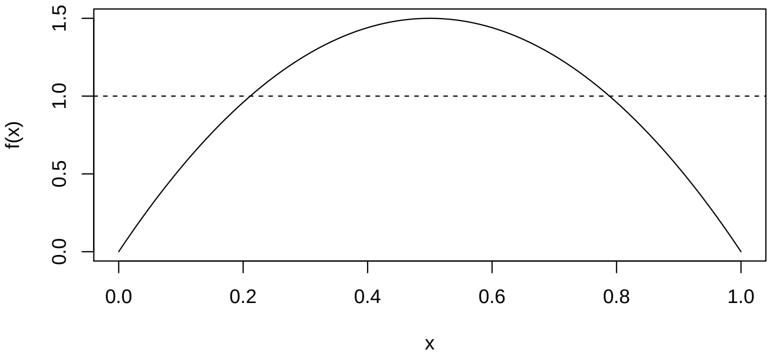
Relating the CDF to the PDF

$$F(x_0) \equiv P(X \leq x_0) = \int_{-\infty}^{x_0} f(x)\, dx$$

Example: Suppose X has Support Set $[0, 1]$

Let $f(x) = 6x(1 - x)$ for $x \in [0, 1]$ and zero otherwise.

```r
curve(6 * x * (1 - x), from = 0, to = 1, ylab = 'f(x)')
abline(h = 1, lty = 2)
```



Example: Suppose X has Support Set $[0,1]$

Let $f(x) = 6x(1-x)$ for $x \in [0,1]$ and zero otherwise.

Is f a valid PDF?

1. Is $f(x) \geq 0$ for $x \in [0,1]$ and zero otherwise?
2. Does the total area under f equal one?

$$\begin{aligned}\int_{-\infty}^{\infty} f(x)dx &= \int_0^1 6x(1-x)dx = 6\int_0^1 (x - x^2)dx \\ &= 6\left(\frac{x^2}{2} - \frac{x^3}{3}\right)\Big|_0^1 = 1\end{aligned}$$

So yes, f is a valid PDF ✓

Integrating a Function in R

```r
pdf <- function(x) {
  6 * x * (1 - x)
}

integrate(pdf, lower = 0, upper = 1)

## 1 with absolute error < 1.1e-14
```

You can use this to check your work!

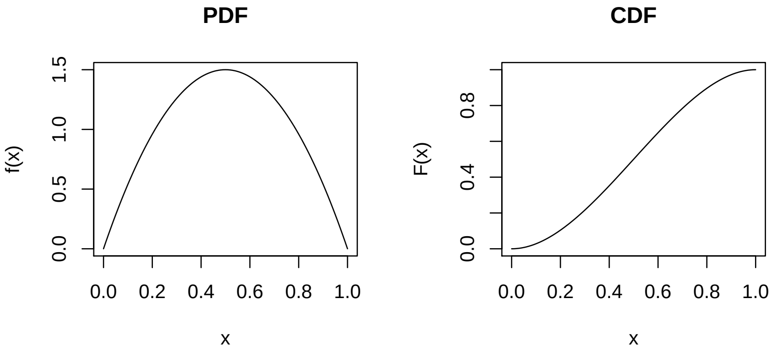
Example: $f(x) = 6x(1-x)$ for $x \in [0,1]$, zero otherwise.

What is the CDF of X?

$$F(x_0) \equiv P(X \leq x_0) = \int_{-\infty}^{x_0} f(x)\, dx = \int_0^{x_0} 6x(1-x)\, dx$$
$$= \; 6\left(\frac{x^2}{2} - \frac{x^3}{3}\right)\Bigg|_0^{x_0} = 3x_0^2 - 2x_0^3$$

$$F(x_0) = \begin{cases} 0, & x_0 < 0 \\ 3x_0^2 - 2x_0^3, & 0 \leq x_0 \leq 1 \\ 1, & x_0 > 1 \end{cases}$$

```r
par(mfrow = c(1,2))
curve(6 * x * (1 - x), from = 0, to = 1, ylab = 'f(x)', main = 'PDF')
curve(3 * x^2 - 2 * x^3, from = 0, to = 1,  ylab = 'F(x)', main = 'CDF')
```




```r
par(mfrow = c(1,1))
```

Relationship between PDF and CDF

Integrate PDF to get CDF

$$F(x_0) = P(X \leq x_0) = \int_{-\infty}^{x_0} f(x)\, dx$$

Differentiate CDF to get PDF

$$f(x) = \frac{d}{dx} F(x)$$

This is just the First Fundamental Theorem of Calculus.

Example: $f(x) = 6x(1-x)$ for $x \in [0,1]$, zero otherwise.

Differentiate CDF to get PDF

$$
\begin{aligned}
f(x) = \frac{d}{dx}F(x) &= \frac{d}{dx}\left(3x^2 - 2x^3\right) \\
&= 6x - 6x^2 \\
&= 6x(1-x)
\end{aligned}
$$

Key Idea: Probability of an Interval for a Continuous RV

$$\boxed{P(a \leq X \leq b) = \int_a^b f(x)\, dx = F(b) - F(a)}$$

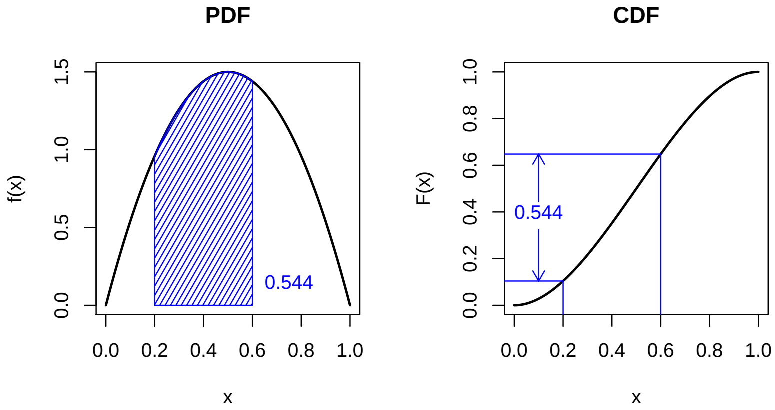
This is just the Second Fundamental Theorem of Calculus.

Example: $f(x) = 6x(1-x)$ for $x \in [0,1]$, zero otherwise.

Two equivalent ways of calculating $P(0.2 \leq X \leq 0.6)$

```
cdf <- function(x0) {
  3 * x0^2 - 2 * x0^3
}
cdf(0.6) - cdf(0.2)

## [1] 0.544

integrate(pdf, lower = 0.2, upper = 0.6)

## 0.544 with absolute error < 6e-15
```

Example: $f(x) = 6x(1-x)$ for $x \in [0,1]$, zero otherwise.



$P(0.2 \leq X \leq 0.6) = 0.544$

Expected Value for Continuous RVs

$$\boxed{E[X] = \int_{-\infty}^{\infty} x f(x)\, dx}$$

$$\boxed{E[g(X)] = \int_{-\infty}^{\infty} g(x) f(x)\, dx}$$

Integrals Replace Sums!

What about all those rules for expected value?

- The only difference between expectation for continuous versus discrete is how we do the *calculation*.
- Sum for discrete; integral for continuous.
- All *properties* of expected value continue to hold!
- Includes linearity, shortcut for variance, etc.

Variance of Continuous RV

$$\boxed{Var(X) = \int_{-\infty}^{\infty} (x - \mu)^2 f(x)\, dx}$$

where

$$\mu = E[X] = \int_{-\infty}^{\infty} x f(x)\, dx$$

Shortcut formula still holds for continuous RVs!

$$Var(X) = E[X^2] - (E[X])^2$$

Example: $f(x) = 6x(1-x)$ for $x \in [0,1]$, zero otherwise.

$$E[X] = \int_{-\infty}^{\infty} xf(x)\ dx = \int_0^1 x \cdot 6x(1-x) = 6\left(\frac{x^3}{3} - \frac{x^4}{4}\right)\Bigg|_0^1 = \frac{1}{2}$$

$$E[X^2] = \int_{-\infty}^{\infty} x^2 f(x)\ dx = \int_0^1 x^2 \cdot 6x(1-x) = 6\left(\frac{x^4}{4} - \frac{x^5}{5}\right)\Bigg|_0^1 = \frac{3}{10}$$

$$Var(X) = E[X^2] - (E[X])^2 = \frac{3}{10} - \left(\frac{1}{2}\right)^2 = 1/20$$

Complete the algebra at home and check using `integrate` in R.

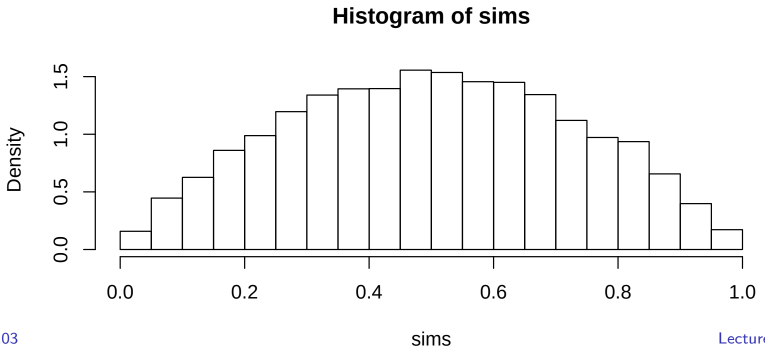
Simulating a Beta$(2, 2)$ Random Variable

Our example from above is a special case of the *Beta distribution*. The command `rbeta(n, 2, 2)` makes `n` draws for this RV. These simulations agree with our calculations from above:

```
set.seed(12345)
sims <- rbeta(10000, 2, 2)
mean(sims)

## [1] 0.5007002

var(sims)

## [1] 0.05012776
```

Simulating a Beta$(2, 2)$ Random Variable

```
mean(sims^2)

## [1] 0.3008234

hist(sims, freq = FALSE)
```



The Uniform Random Variable

Several of your review questions along with one of your extension questions will involve the so-called *Uniform Random Variable*:

Uniform(0,1) Random Variable

$f(x) = 1$ for $x \in [0, 1]$, zero otherwise.

Uniform(a,b) Random Variable

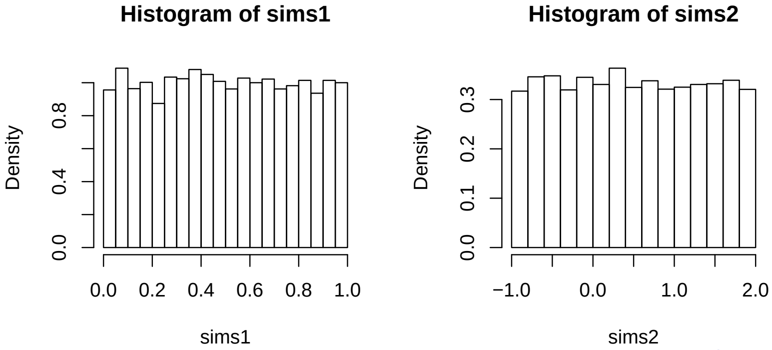
$f(x) = 1/(b - a)$ for $x \in [a, b]$, zero otherwise.

Simulating from a Uniform RV

`runif(n, a, b)` makes `n` draws from a Uniform(a, b) RV.

Simulating Uniform Random Variables

```r
sims1 <- runif(10000, 0, 1)
sims2 <- runif(10000, -1, 2)
par(mfrow = c(1, 2))
hist(sims1, freq = FALSE)
hist(sims2, freq = FALSE)
```



We don't have time to cover these in Econ 103:

Joint Density

$P(a \leq X \leq b \cap c \leq Y \leq d) = \int_c^d \int_a^b f(x,y)\, dxdy$

Marginal Densities

$f_X(x) = \int_{-\infty}^{\infty} f(x,y)\, dy, \qquad f_Y(y) = \int_{-\infty}^{\infty} f(x,y)\, dx$

Independence in Terms of Joint and Marginal Densities

$f_{XY}(x,y) = f_X(x) f_Y(y)$

Conditional Density

$f_{Y|X} = f_{XY}(x,y)/f_X(x)$

So where does that leave us?

What We've Accomplished

We've covered all the basic properties of RVs on this Handout.

Where are we headed next?

Next up is the most important RV of all: the normal RV. After that it's time to do some statistics!

How should you be studying?

If you *master* the material on RVs (both continuous and discrete) and in particular the normal RV the rest of the semester will seem easy. If you don't, you're in for a rough time...

Lecture #12 – Continuous RVs II: The Normal RV

The Standard Normal RV

Linear Combinations and the $N(\mu, \sigma^2)$ RV

Where does the Empirical Rule come from?

From $N(0, 1)$ to $N(\mu, \sigma^2)$ and Back Again

Percentiles/Quantiles for Continuous RVs

Available on Etsy, Made using R!

Figure: Standard Normal RV (PDF)

Standard Normal Random Variable: $N(0,1)$

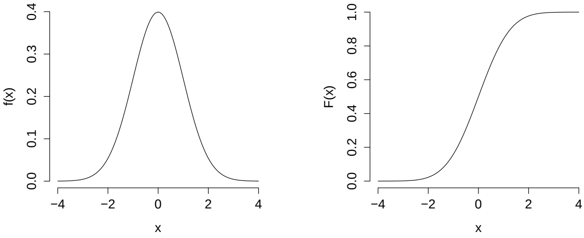
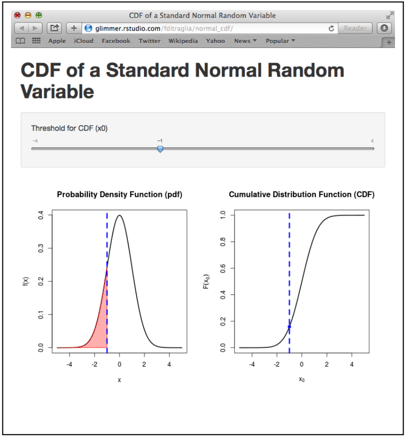


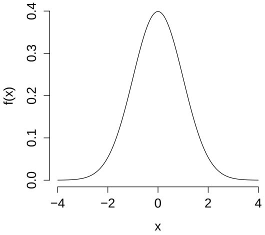
Figure: Standard Normal PDF (left) and CDF (Right)

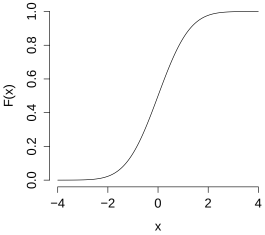
- Notation: $X \sim N(0,1)$
- Symmetric, Bell-shaped, $E[X] = 0$, $Var[X] = 1$
- Support Set $= (-\infty, \infty)$

https://fditraglia.shinyapps.io/normal_cdf/



Standard Normal Random Variable: $N(0,1)$





- There is no closed-form expression for the $N(0,1)$ CDF.
- For Econ 103, don't need to know formula for $N(0,1)$ PDF.
- You *do need* to know the R commands...

R Commands for the Standard Normal RV

`dnorm` – Standard Normal PDF

- Mnemonic: `d` = density, `norm` = normal
- Example: `dnorm(0)` gives height of $N(0,1)$ PDF at zero.

`pnorm` – Standard Normal CDF

- Mnemonic: `p` = probability, `norm` = normal
- Example: `pnorm(1)` $= P(X \leq 1)$ if $X \sim N(0,1)$.

`rnorm` – Simulate Standard Normal Draws

- Mnemonic: `r` = random, `norm` = normal.
- Example: `rnorm(10)` makes ten iid $N(0,1)$ draws.

Add a knitr frame giving examples. . .

The $N(\mu, \sigma^2)$ Random Variable

Idea

Take a linear function of the $N(0,1)$ RV.

Formal Definition

$N(\mu, \sigma^2) \equiv \mu + \sigma X$ where $X \sim N(0,1)$ and μ, σ are constants.

Properties of $N(\mu, \sigma^2)$ RV

- Parameters: Expected Value $= \mu$, Variance $= \sigma^2$
- Symmetric and bell-shaped.
- Support Set $= (-\infty, \infty)$
- $N(0,1)$ is the special case where $\mu = 0$ and $\sigma^2 = 1$.

Expected Value: μ *shifts* PDF

all of these have $\sigma = 1$

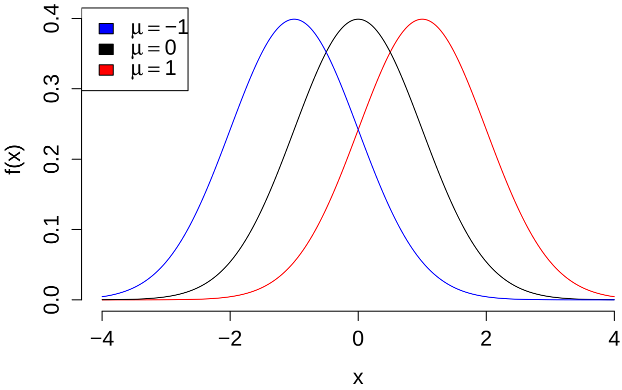


Figure: Blue $\mu = -1$, Black $\mu = 0$, Red $\mu = 1$

Standard Deviation: σ *scales* PDF

all of these have $\mu = 0$

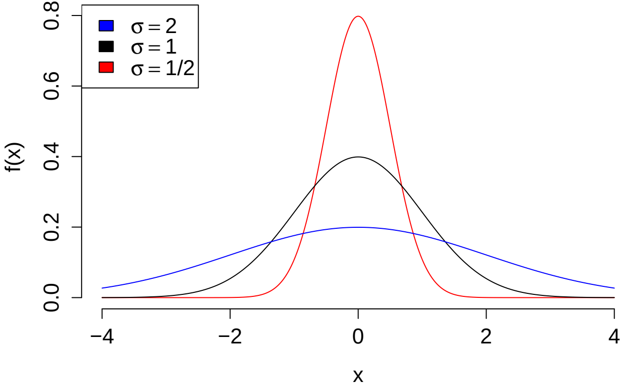


Figure: Blue $\sigma^2 = 4$, Black $\sigma^2 = 1$, Red $\sigma^2 = 1/4$

Linear Function of Normal RV is a Normal RV

Suppose that $X \sim N(\mu, \sigma^2)$. Then if a and b constants,

$$\boxed{a + bX \sim N(a + b\mu, b^2\sigma^2)}$$

Important

- For *any* RV X, $E[a + bX] = a + bE[X]$ and $Var(a + bX) = b^2 Var(X)$.
- Key point: linear transformation of normal is still normal!
- Linear transformation of Binomial is *not* Binomial!

Example

Suppose $X \sim N(\mu, \sigma^2)$ and let $Z = (X - \mu)/\sigma$. What is the distribution of Z?

(a) $N(\mu, \sigma^2)$

(b) $N(\mu, \sigma)$

(c) $N(0, \sigma^2)$

(d) $N(0, \sigma)$

(e) $N(0, 1)$

Linear Combinations of *Multiple Independent* Normals

Let $X \sim N(\mu_x, \sigma_x^2)$ independent of $Y \sim N(\mu_y, \sigma_y^2)$. Then if a, b, c are constants:

$$\boxed{aX + bY + c \sim N(a\mu_x + b\mu_y + c, a^2\sigma_x^2 + b^2\sigma_y^2)}$$

Important

- Result assumes independence
- Particular to Normal RV
- Extends to more than two Normal RVs

Suppose $X_1, X_2, \sim$ iid $N(\mu, \sigma^2)$

Let $\bar{X} = (X_1 + X_2)/2$. What is the distribution of $\bar{X}$?

(a) $N(\mu, \sigma^2/2)$

(b) $N(0, 1)$

(c) $N(\mu, \sigma^2)$

(d) $N(\mu, 2\sigma^2)$

(e) $N(2\mu, 2\sigma^2)$

Where does the Empirical Rule come from?

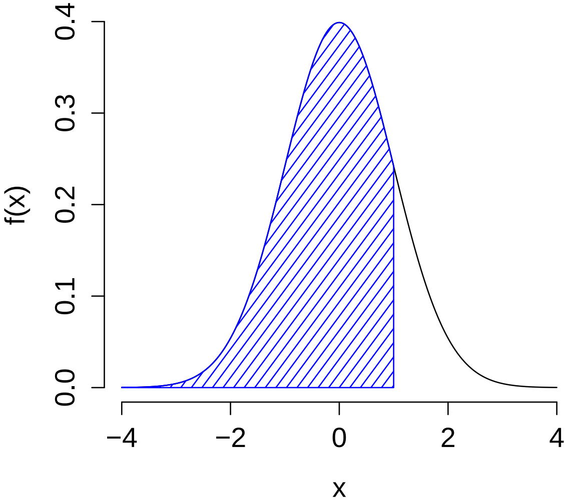
Empirical Rule

Approximately 68% of observations within $\mu \pm \sigma$

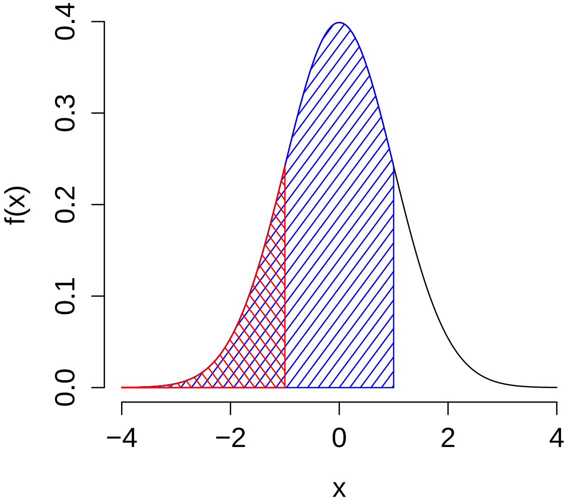
Approximately 95% of observations within $\mu \pm 2\sigma$

Nearly all observations within $\mu \pm 3\sigma$

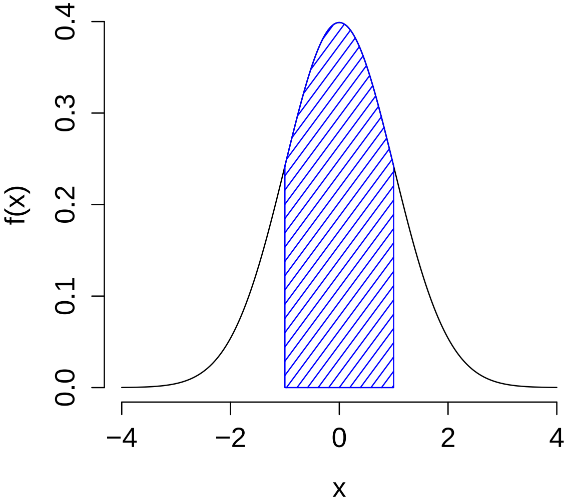
pnorm(1) ≈ 0.84



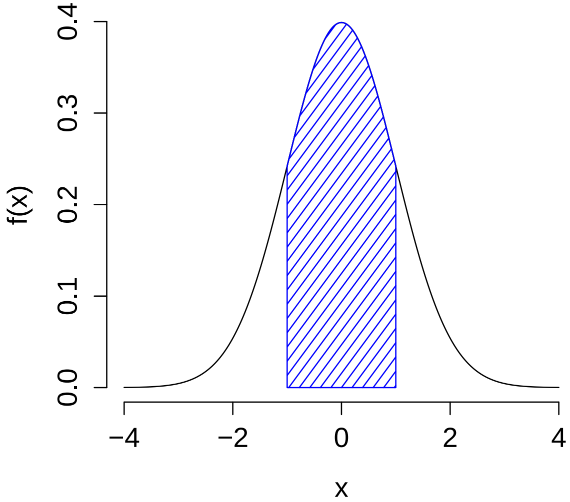
`pnorm(1) - pnorm(-1)` $\approx 0.84 - 0.16$



`pnorm(1) - pnorm(-1)` ≈ 0.68



Middle 68% of $N(0,1) \Rightarrow$ approx. $(-1,1)$



Suppose $X \sim N(0, 1)$

$$
\begin{aligned}
P(-1 \leq X \leq 1) &= \texttt{pnorm(1) - pnorm(-1)} \\
&\approx 0.683
\end{aligned}
$$

$$
\begin{aligned}
P(-2 \leq X \leq 2) &= \texttt{pnorm(2) - pnorm(-2)} \\
&\approx 0.954
\end{aligned}
$$

$$
\begin{aligned}
P(-3 \leq X \leq 3) &= \texttt{pnorm(3) - pnorm(-3)} \\
&\approx 0.997
\end{aligned}
$$

What if $X \sim N(\mu, \sigma^2)$?

$$
\begin{aligned}
P(X \leq a) &= P(X - \mu \leq a - \mu) \\
&= P\left(\frac{X-\mu}{\sigma} \leq \frac{a-\mu}{\sigma}\right) \\
&= P\left(Z \leq \frac{a-\mu}{\sigma}\right) \\
&= \texttt{pnorm}\left(\frac{a-\mu}{\sigma}\right)
\end{aligned}
$$

Where Z is a standard normal random variable, i.e. $N(0,1)$.

Probability *Above* a Threshold: $X \sim N(\mu, \sigma^2)$

$$\begin{aligned}
P(X \geq b) &= 1 - P(X \leq b) = 1 - P\left(\frac{X-\mu}{\sigma} \leq \frac{b-\mu}{\sigma}\right) \\
&= 1 - P\left(Z \leq \frac{b-\mu}{\sigma}\right) \\
&= 1 - \texttt{pnorm}((b-\mu)/\sigma)
\end{aligned}$$

Where Z is a standard normal random variable.

Suppose $X \sim N(\mu = 1, \sigma^2 = 4)$

What is $P(-1 \leq X \leq 3)$?

$$
\begin{aligned}
P(-1 \leq X \leq 4) &= P\left(-2 \leq X - 1 \leq 2\right) \\
&= P\left(-1 \leq \frac{X-1}{2} \leq 1\right) \\
&= P\left(-1 \leq Z \leq 1\right) \\
&= \texttt{pnorm(1)} - \texttt{pnorm(-1)} \\
&\approx 0.68
\end{aligned}
$$

Probability of an Interval: $X \sim N(\mu, \sigma^2)$

$$
\begin{aligned}
P(a \leq X \leq b) &= P\left(\frac{a-\mu}{\sigma} \leq \frac{X-\mu}{\sigma} \leq \frac{b-\mu}{\sigma}\right) \\
&= P\left(\frac{a-\mu}{\sigma} \leq Z \leq \frac{b-\mu}{\sigma}\right) \\
&= \texttt{pnorm}((b-\mu)/\sigma) - \texttt{pnorm}((a-\mu)/\sigma)
\end{aligned}
$$

Where Z is a standard normal random variable.

Percentiles/Quantiles for Continuous RVs

Quantile Function $Q(p)$ is the inverse of CDF $F(x_0)$

Plug in a probability p, get out the value of x_0 such that $F(x_0) = p$

$$Q(p) = F^{-1}(p)$$

In other words:

$$Q(p) = \text{the value of } x_0 \text{ such that } \int_{-\infty}^{x_0} f(x)\, dx = p$$

Inverse exists as long as $F(x_0)$ is *strictly increasing*.

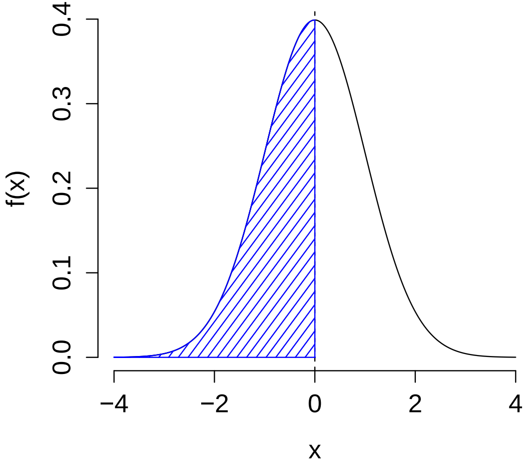
Example: Median

The median of a continuous random variable is $Q(0.5)$, i.e. the value of x_0 such that

$$\int_{-\infty}^{x_0} f(x)\, dx = 1/2$$

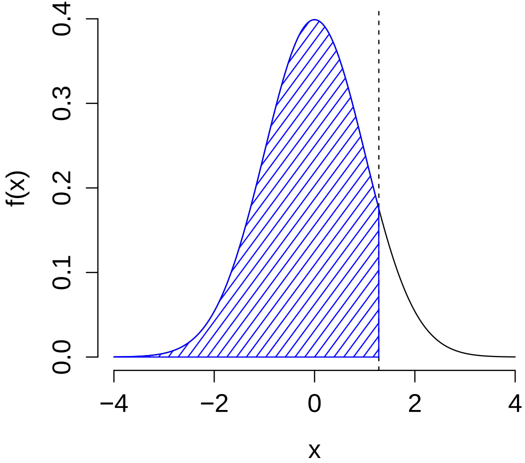
What is the median of a standard normal RV?

By symmetry, $Q(0.5) = 0$. R command: `qnorm()`



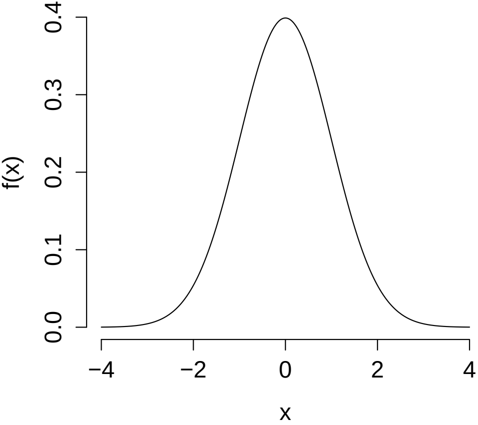
90th Percentile of a Standard Normal

`qnorm(0.9)` ≈ 1.28



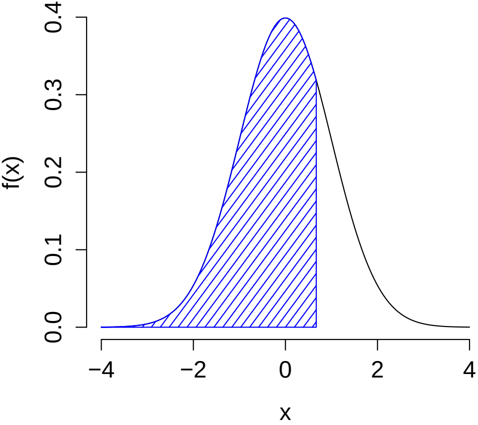
Using Quantile Function to find Symmetric Intervals

Suppose X is a standard normal RV. What is the value of c such that $P(-c \leq X \leq c) = 0.5$?



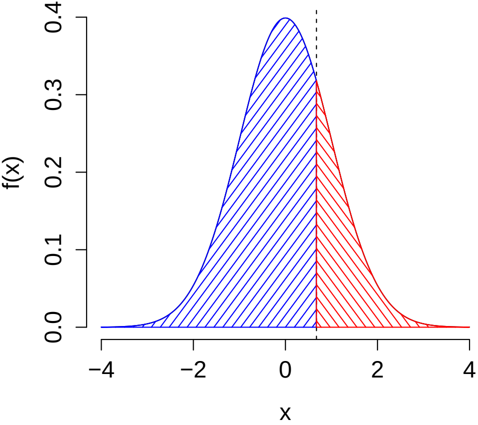
qnorm(0.75) ≈ 0.67

Suppose X is a standard normal RV. What is the value of c such that $P(-c \leq X \leq c) = 0.5$?



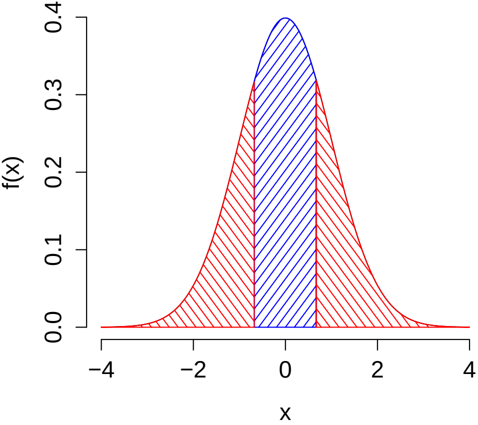
qnorm(0.75) ≈ 0.67

Suppose X is a standard normal RV. What is the value of c such that $P(-c \leq X \leq c) = 0.5$?



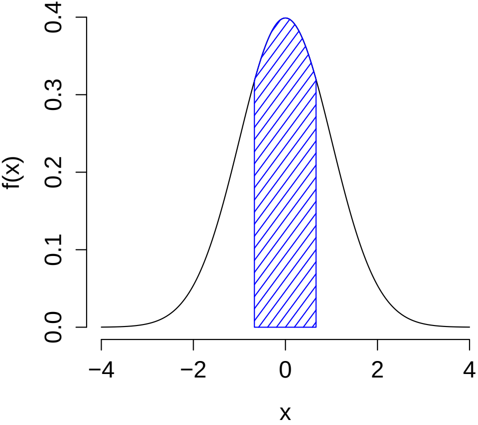
`pnorm(0.67)-pnorm(-0.67)`$\approx$?

Suppose X is a standard normal RV. What is the value of c such that $P(-c \leq X \leq c) = 0.5$?



`pnorm(0.67)-pnorm(-0.67)`≈ 0.5

Suppose X is a standard normal RV. What is the value of c such that $P(-c \leq X \leq c) = 0.5$?



95% Central Interval for Standard Normal

Suppose X is a standard normal random variable. What value of c ensures that $P(-c \leq X \leq c) \approx 0.95$?

R Commands for *Arbitrary* Normal RVs: $X \sim N(\mu, \sigma^2)$

PDF $f(x)$	`dnorm(x, mean =` μ`, sd =` σ`)`
Random Draws	`rnorm(n, mean =` μ`, sd =` σ`)`
CDF $F(x)$	`pnorm(x, mean =` μ`, sd =` σ`)`
Quantile Function $Q(p)$	`qnorm(p, mean =` μ`, sd =` σ`)`

Notice that this means you don't have to transform X to a standard normal in order to find areas under its pdf using R.

Example: $X \sim N(0, 16)$

One Way:

$$\begin{aligned}
P(X \geq 10) &= 1 - P(X \leq 10) = 1 - P(X/4 \leq 10/4) \\
&= 1 - P(Z \leq 2.5) = 1 - \texttt{pnorm(2.5)} \\
&\approx 0.006
\end{aligned}$$

An Easier Way:

$$\begin{aligned}
P(X \geq 10) &= 1 - P(X \leq 10) \\
&= 1 - \texttt{pnorm(10, mean = 0, sd = 4)} \\
&\approx 0.006
\end{aligned}$$