Economics 103 – Statistics for Economists

Francis J. DiTraglia

University of Pennsylvania

Lecture # 11

Today

1. Discrete RVs – Part IV
2. Continuous RVs – Part I

Recall from Last Time:

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

Linearity of Expectation, Again

Holds for Continuous RVs as well, but different proof.

In general, $E[g(X, Y)] \neq g(E[X], E[Y])$. The key exception is when g is a linear function:

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

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

Proof of Linearity of Expectation for Discrete RVs

$$
\begin{aligned}
E[aX + bY + c] &= \sum_x \sum_y (ax + by + c)p(x, y) \\
&=
\end{aligned}
$$

Proof of Linearity of Expectation for Discrete RVs

$$
\begin{aligned}
E[aX + bY + c] &= \sum_x \sum_y (ax + by + c) p(x, y) \\
&= \sum_x \sum_y [axp(x, y) + byp(x, y) + cp(x, y)] \\
&=
\end{aligned}
$$

Proof of Linearity of Expectation for Discrete RVs

$$
\begin{aligned}
E[aX+bY+c] &= \sum_x \sum_y (ax+by+c)p(x,y) \\
&= \sum_x \sum_y [axp(x,y)+byp(x,y)+cp(x,y)] \\
&= a\sum_x \sum_y xp(x,y) + b\sum_y \sum_x yp(x,y) + c\sum_y \sum_x p(x,y) \\
&=
\end{aligned}
$$

Proof of Linearity of Expectation for Discrete RVs

$$
\begin{aligned}
E[aX + bY + c] &= \sum_x \sum_y (ax + by + c)p(x, y) \\
&= \sum_x \sum_y [axp(x, y) + byp(x, y) + cp(x, y)] \\
&= {\color{red}a} \sum_x \sum_y xp(x, y) + {\color{red}b} \sum_y \sum_x yp(x, y) + {\color{red}c} \sum_y \sum_x p(x, y) \\
&= a \sum_x \sum_y xp(x, y) + b \sum_y \sum_x yp(x, y) + {\color{red}c} \\
&=
\end{aligned}
$$

Proof of Linearity of Expectation for Discrete RVs

$$
\begin{aligned}
E[aX + bY + c] &= \sum_x \sum_y (ax + by + c)p(x, y) \\
&= \sum_x \sum_y [axp(x, y) + byp(x, y) + cp(x, y)] \\
&= a \sum_x \sum_y xp(x, y) + b \sum_y \sum_x yp(x, y) + c \sum_y \sum_x p(x, y) \\
&= a \sum_x \sum_y xp(x, y) + b \sum_y \sum_x yp(x, y) + c \\
&= a \sum_x x \left(\sum_y p(x, y) \right) + b \sum_y y \left(\sum_x p(x, y) \right) + c \\
&=
\end{aligned}
$$

Proof of Linearity of Expectation for Discrete RVs

$$
\begin{aligned}
E[aX + bY + c] &= \sum_x \sum_y (ax + by + c)p(x, y) \\
&= \sum_x \sum_y [axp(x, y) + byp(x, y) + cp(x, y)] \\
&= a \sum_x \sum_y xp(x, y) + b \sum_y \sum_x yp(x, y) + c \sum_y \sum_x p(x, y) \\
&= a \sum_x \sum_y xp(x, y) + b \sum_y \sum_x yp(x, y) + c \\
&= a \sum_x x \left(\sum_y p(x, y) \right) + b \sum_y y \left(\sum_x p(x, y) \right) + c \\
&= a \sum_x xp(x) + b \sum_y yp(y) + c \\
&=
\end{aligned}
$$

Proof of Linearity of Expectation for Discrete RVs

$$
\begin{aligned}
E[aX + bY + c] &= \sum_x \sum_y (ax + by + c)p(x, y) \\
&= \sum_x \sum_y [axp(x, y) + byp(x, y) + cp(x, y)] \\
&= a \sum_x \sum_y xp(x, y) + b \sum_y \sum_x yp(x, y) + c \sum_y \sum_x p(x, y) \\
&= a \sum_x \sum_y xp(x, y) + b \sum_y \sum_x yp(x, y) + c \\
&= a \sum_x x \left(\sum_y p(x, y) \right) + b \sum_y y \left(\sum_x p(x, y) \right) + c \\
&= a \sum_x x p(x) + b \sum_y y p(y) + c \\
&= aE[X] + bE[Y] + c
\end{aligned}
$$

Application: Shortcut Formula for Variance

By the Linearity of Expectation,

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

Application: Shortcut Formula for Variance

By the Linearity of Expectation,

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

Application: Shortcut Formula for Variance

By the Linearity of Expectation,

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

Application: Shortcut Formula for Variance

By the Linearity of Expectation,

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

Application: Shortcut Formula for Variance

By the Linearity of Expectation,

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

We saw in a previous lecture that it's typically much easier to calculate variances using the shortcut formula.

Another Application: Shortcut Formula for Covariance

Similar to Shortcut for Variance: in fact $Var(X) = Cov(X, X)$

$$
\begin{aligned}
Cov(X, Y) &= E[(X - \mu_X)(Y - \mu_Y)] \\
&= E[XY - \mu_X Y - \mu_Y X + \mu_X \mu_Y] \\
&\vdots \\
&= E[XY] - E[X]E[Y]
\end{aligned}
$$

You'll fill in the details for homework...

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

Joint pmf equals product of marginal pmfs (see last lecture): Knowing the 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.)

Binomial(n, p) Random Variable

Definition

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

Parameters

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

Support

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

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).

Which of these is the PMF of a Binomial(n, p) RV?

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

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

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

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

(e) $p(x) = p^n(1-p)^x$

Which of these is the PMF of a Binomial(n, p) RV?

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

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

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

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

(e) $p(x) = p^n(1-p)^x$

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

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:

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

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:

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

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:

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

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:

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

Extremely Important:

Variance of Sum $\neq$ Sum of Variances!

Variance of a Sum

$$Var(aX + bY) = E\left[\{(aX + bY) - E[aX + bY]\}^2\right]$$

Variance of a Sum

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

Variance of a Sum

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

Variance of a Sum

$$
\begin{aligned}
Var(aX + bY) &= E\left[\{(aX + bY) - E[aX + bY]\}^2\right] \\
&= E\left[\{a(X - \mu_X) + b(Y - \mu_Y)\}^2\right] \\
&= E\left[a^2(X - \mu_X)^2 + b^2(Y - \mu_Y)^2 + 2ab(X - \mu_X)(Y - \mu_Y)\right] \\
&= a^2E[(X - \mu_X)^2] + b^2E[(Y - \mu_Y)^2] + 2abE[(X - \mu_X)(Y - \mu_Y)]
\end{aligned}
$$

Variance of a Sum

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

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

We showed last time that if X and Y are independent, $Cov(X, Y) = 0$. Hence, independence implies

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

This is also true for more than two 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

It is always true that

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

Variance

It is not true in general that

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

but it is true in the special case where $X_1, \ldots X_n$ are independent.

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

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

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

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

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

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

Continuous RVs – Part I

What Changes?

1. Probability Density Functions replace Probability Mass Functions (aka Probability Distributions)
2. Integrals Replace Sums

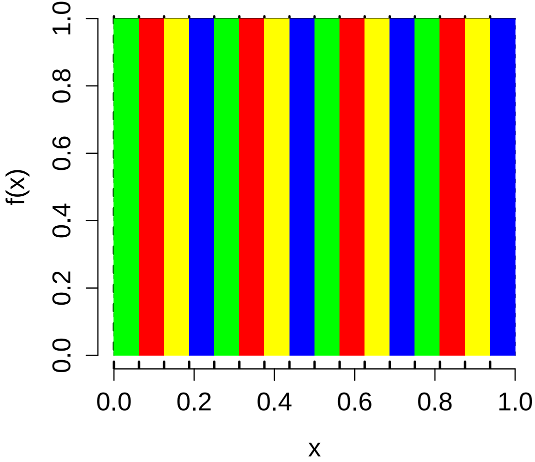
Everything Else is Essentially Unchanged!

What is the probability of “Yellow?”

What is the probability of "Right Hand Blue?"

What is the probability that the spinner lands in any *particular* place?

From Twister to Density – Probability as *Area*



Continuous Random Variables

For continuous RVs, probability is a matter of finding the area of *intervals*. Individual *points* have *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, $P(X = x) = 0 \neq f(x)$!

Properties of PDFs

1. $\int_{-\infty}^{\infty} f(x)\, dx = 1$
2. $f(x) \geq 0$ for all x
3. $f(x)$ is *not* a probability and can be greater than one!
4. $P(X \leq x_0) = F(x_0) = \int_{-\infty}^{x_0} f(x)\, dx$

We'll start with the simplest possible example: the Uniform$(0, 1)$ RV.

Uniform$(0, 1)$ Random Variable

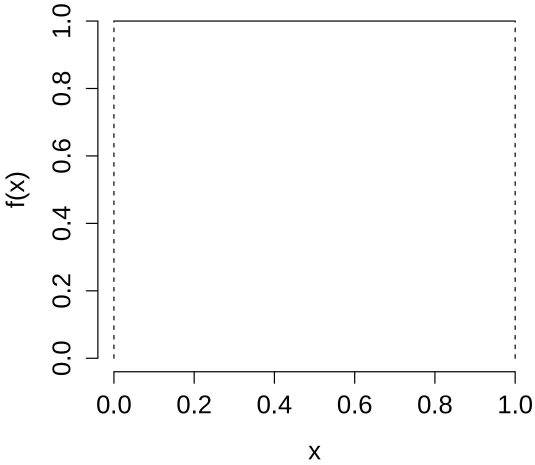
$X \sim \text{Uniform}(0, 1)$

We say that X follows a Uniform(0,1) distribution, if it is equally likely to take on *any value* in the range $[0, 1]$ and never takes on a value outside this range.

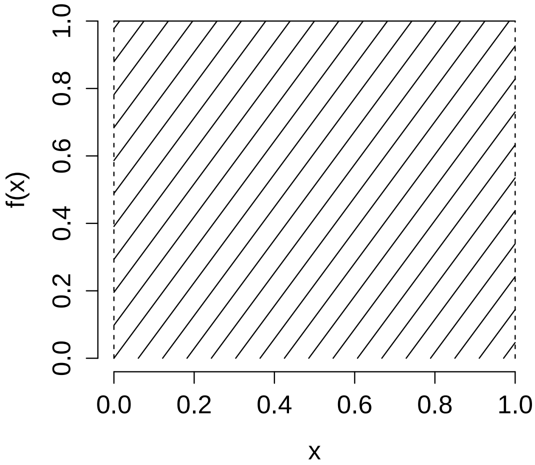
Uniform PDF

$f(x) = 1$ for $0 \leq x \leq 1$, zero elsewhere.

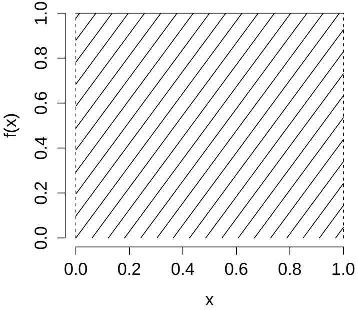
Uniform$(0, 1)$ PDF



What is the area of the shaded region?

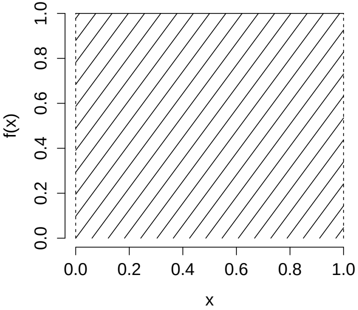


What is the area of the shaded region?



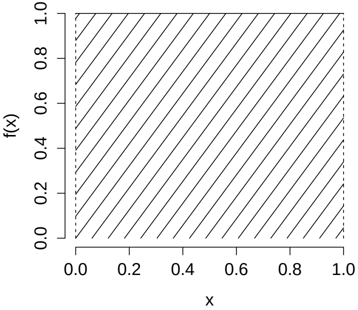
$$\int_{-\infty}^{\infty} f(x)\, dx =$$

What is the area of the shaded region?



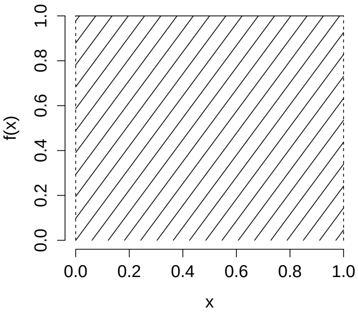
$$\int_{-\infty}^{\infty} f(x)\, dx = \int_0^1 1\, dx =$$

What is the area of the shaded region?



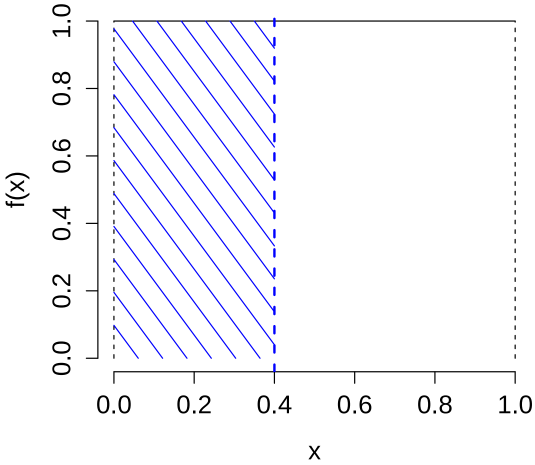
$$\int_{-\infty}^{\infty} f(x)\, dx = \int_0^1 1\, dx = \left. x \right|_0^1 =$$

What is the area of the shaded region?

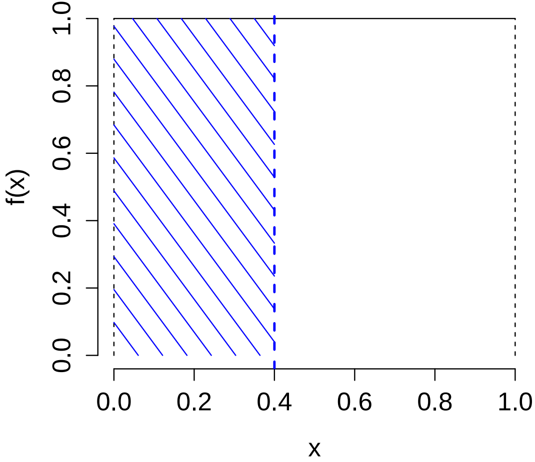


$$\int_{-\infty}^{\infty} f(x)\ dx = \int_{0}^{1} 1\ dx = \left. x \right|_{0}^{1} = 1 - 0 = 1$$

What is the area of the shaded region?



$$F(0.4) = P(X \leq 0.4) = 0.4$$



Relationship between PDF and CDF

Integrate the pdf to get the CDF

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

Differentiate the CDF to get the pdf

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

This is just the Fundamental Theorem of Calculus.

Example: Uniform$(0, 1)$ RV

Integrate the pdf, $f(x) = 1$, to get the CDF

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

Example: Uniform(0, 1) RV

Integrate the pdf, $f(x) = 1$, to get the CDF

$$F(x_0) = \int_{-\infty}^{x_0} f(x)\, dx = \int_0^{x_0} 1\, dx =$$

Example: Uniform$(0, 1)$ RV

Integrate the pdf, $f(x) = 1$, to get the CDF

$$F(x_0) = \int_{-\infty}^{x_0} f(x)\, dx = \int_0^{x_0} 1\, dx = \left. x \right|_0^{x_0} =$$

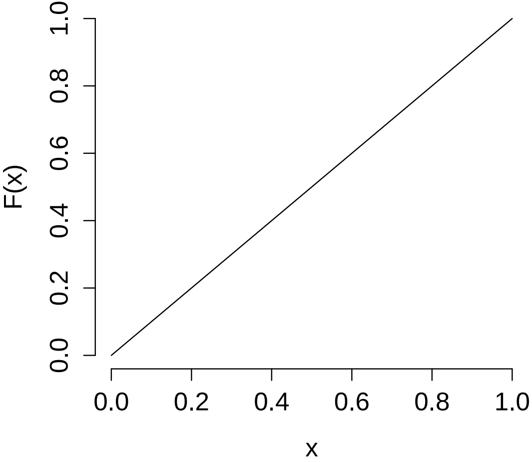
Example: Uniform$(0, 1)$ RV

Integrate the pdf, $f(x) = 1$, to get the CDF

$$F(x_0) = \int_{-\infty}^{x_0} f(x)\, dx = \int_0^{x_0} 1\, dx = \left. x \right|_0^{x_0} = x_0 - 0 = x_0$$

$$F(x_0) = \begin{cases} 0, x_0 < 0 \\ x_0, 0 \leq x_0 \leq 1 \\ 1, x_0 > 1 \end{cases}$$

Uniform(0, 1) CDF



Example: Uniform$(0, 1)$ RV

Differentiate the CDF, $F(x_0) = x_0$, to get the pdf

$$\frac{d}{dx}F(x) = 1 = f(x)$$