every measurement you ever make is in actuality discrete, but it's useful to regard rvs that are conceptually continuous (i.e., no limit in principle to the precision of measurement) as continuous.

\[ \text{(23 Apr 19)} \]

**Definition**

Given a (mass) discrete rvs \( Z \), the probability function (pmf or pf) of \( Z \) is the function \( f_z(z) \) that keeps track of the probability associated with \( Z \): \( f_z(z) = P(Z = z) \).

The set \( \{ z : f_z(z) > 0 \} \) is called the support of (the distribution of) \( Z \).

(CS is almost unique in using "pf" nearly everybody talks about the pmf.)
A rv $X$ that only takes on the values $\{0, 1\}$ - i.e., a binary rv - is said to have a Bernoulli distribution with parameter $p$, written Bernoulli$(p)$ -

\[ f(x) = P(X = x) = \begin{cases} p & \text{for } x = 1 \\ 1 - p & \text{for } x = 0 \\ 0 & \text{else} \end{cases} \]

Notation: $X$ follows a Bernoulli$(p)$ distribution is distributed as $X \sim$ Bernoulli$(p)$ or $(X \mid p)$.
Example. In the powerball lottery (see Problem 2) 5 white balls are drawn at random without replacement from a bin with balls numbered \{1, 2, \ldots, 69\}.

Let \(W_i\) = # of \(i\)th drawn ball, \(i = 1, 2, \ldots, 6\).

Clearly \(p(W_1 = w_1) = \left\{ \begin{array}{ll} \frac{1}{69} & \text{for } w_1 = 1, 2, \ldots, 69 \\ 0 & \text{otherwise} \end{array} \right.\)

less clearly (but true) \(W_2, \ldots, W_5\) follow the same distribution if nothing is known about the previous draws.

Definition. For any two integers \(a \leq b\), a rv \(X\) that's equally likely to be any of the values \(\{a, a+1, \ldots, b\}\) has the uniform distribution \text{Uniform} \{a, b\}. Evidently
its pdf is \[ f(y) = P(X = y) = \begin{cases} \frac{1}{b-a+1} & \text{for } y = a, \ldots, b \\ 0 & \text{else} \end{cases} \]

\[ Y \sim \text{Uniform } \{ a, b \} \iff Y \text{ chosen at random from } \{ a, a+1, \ldots, b \}. \]

**Definition**  \( n \) **trials are performed**, with each trial recorded as a success \( S \) or failure \( F \). If each trial is **independent of all the others** and the chance \( P \) of success is **constant across the trials**, then \( Y = \# \text{ of successes} \) has the **Binomial distribution**

\[ f(y) = P(Y = y) = \binom{n}{y} p^y (1-p)^{n-y} \quad \text{for } y = 0, 1, \ldots, n \]

(with parameters \( n \) and \( p \)).
In shorthand \( Y \sim \text{Binomial} \left( n, p \right) \) or \((Y \mid n, p)\).

Let \( \Theta_i = \begin{cases} 1 & \text{if trial } i \text{ is a success} \\ 0 & \text{failure} \end{cases} \)

for \( i = 1, \ldots, n \); then under these assumptions \( \Theta_i \sim \text{Bernoulli} \left( p \right) \) and all the \( \Theta_i \) are independent.

Notation \( X_i \sim f(X_i) \) means that all of the r.v.s \( X_1, X_2, \ldots \) are independent and identically distributed.

Thus with the success/failure trials, \( \Theta_i \sim \text{Bernoulli} \left( p \right) \) and

\[
\left( Y = \sum_{i=1}^{n} \Theta_i \right) \sim \text{Binomial} \left( n, p \right).
\]
This is our first example of the distribution of the sum of a bunch of IID rvs, a topic we'll examine in detail later.
Continuous random variables

Example (round-off error in computer science)

Single-precision floating-point decimal numbers carry about 7 significant digits of accuracy,

\[ 3.141592653589 \]

leading to roundoff error in the last digit, implying to study how these errors accumulate as the number of steps in a calculation increases. Since there's no reason one decimal digit would be favored over another in rounding, the uniform distribution is key to these calculations. Consider first

Uniform \{0, 0.1, \ldots, 0.9\} and then \{0, 0.1, 0.2, \ldots, 0.999\}

\[ 0 \quad \underbrace{0.1 \quad \ldots \quad 0.9} \quad 1 \]

\[ 0.1 \quad 1 \quad \ldots \quad 1 \]

\[ 0 \quad \overbrace{1 \quad \ldots \quad 1} \quad 1 \]
In the limit with more & more right

This should go to the

continuous uniform distribution

Uniform (0,1) on the unit interval.

The analogue of the discrete pt in this

continuous case is the smooth function

\[ f(x) = \begin{cases} 1 & \text{for } 0 \leq x \leq 1 \\ 0 & \text{else} \end{cases} \]

Definition

A random variable

(\(X\) has a continuous distribution) if there exists a continuous nonnegative function

\(f_X\) defined on \(\mathbb{R}\) such that for every interval \([a, b]\),

\[ P(a \leq X \leq b) = \int_a^b f_X(y)\,dy. \]
In this definition, $a$ can be $-\infty$ and $b$ can be $\infty$. 

**Definition:** If $Y$ is a continuous rv, the function $f_Y$ in the previous definition is called the probability density function (PDF) of $Y$. The set $\{y : f_Y(y) > 0\}$ is called the support of (the distribution of) $Y$. 

Clearly (a) $f_Y(y) \geq 0$ for all $y$ and (b) \( \int_{-\infty}^{\infty} f_Y(y) \, dy = 1 \). What about individual points - singletons - \{y\} on $\mathbb{R}$?

You'll recall from calculus that if $f_Y$ is continuous...
on its support, \( \int_{a}^{b} f_{\bar{Y}}(y) \, dy \) can equally well stand for \( P(\bar{a} \leq \bar{Y} \leq \bar{b}) \) or \( P(\bar{a} < \bar{Y} \leq \bar{b}) \) or \( P(\bar{a} \leq \bar{Y} < \bar{b}) \) or \( P(\bar{a} < \bar{Y} < \bar{b}) \), because (e.g.) \( \int_{a}^{\bar{a}} f_{\bar{Y}}(y) \, dy = 0 \) if \( f_{\bar{y}} \) is continuous at \( y = \bar{a} \). Thus, importantly, \( P(\bar{Y} = y) = 0 \) for all \( -\infty < y < \infty \).

Weirdly, this doesn't mean that the value \( y \) of \( \bar{Y} \) is impossible, or it does with discrete \( Y \); it just means that singletons have to have 0 probability (otherwise \( \int_{-\infty}^{\infty} f_{\bar{Y}}(y) \, dy = +\infty \) not 1).
Definition (with a and b any two real numbers satisfying $a < b$)

$Y \sim \text{Uniform}(a, b) \iff P\left(\frac{y}{b-a} \in (0, 1)\right) = \text{the length of the subinterval} \Leftarrow$ 

$f_Y(y) = \begin{cases} \frac{1}{b-a} & \text{for } a \leq y \leq b \\ 0 & \text{else} \end{cases}$

Definition (The indicator function)

The indicator function $(\text{true/false})$

for any proposition $A$ is $I(A) = \begin{cases} 1 & \text{if true} \\ 0 & \text{if false} \end{cases}$

People sometimes also write (with a set)

$I_A(y) = \begin{cases} 1 & \text{if } y \in A \\ 0 & \text{else} \end{cases}$
With this definition, $Y \sim \text{Uniform}(a, b)$.

$$
f_Y(y) = \frac{1}{b-a} I(a \leq y \leq b) = \frac{I_{[a,b]}(y)}{b-a}.
$$

Contrast $Y \sim \text{Uniform}(a, b)$, continuous, and uniform on $(a, b)$ or $\mathbb{R}$ versus $Y \sim \text{Uniform}\{a, b\}$ for $a, b$ integers with $a < b \iff Y$ discrete and uniform on $\{a, a+1, \ldots, b\}$.

Density and probability are not the same things.

Density values $f_Y(y)$ are themselves not probabilities; for example, they can easily be $>1$ and can even be $\to$, but probabilities must be $\leq 1$. 


we'll see later.  \( \text{Density function:} \quad P(a \leq Y \leq b) = \int_{a}^{b} f_{X}(y) \, dy. \)

For small \( \epsilon > 0 \) you can see from this sketch that
\[
P(a - \frac{\epsilon}{2} \leq Y \leq a + \frac{\epsilon}{2}) = \int_{a - \frac{\epsilon}{2}}^{a + \frac{\epsilon}{2}} f_{X}(y) \, dy
\]

**Example**

(\text{Triangular distribution})

\( f_{X}(y) \) line with slope \( m = \frac{c}{b-a} \)

Can a continuous \( RV \) have a pdf that looks like a triangle? Let's see what, if any, restrictions would be needed.
The line in the sketch has slope \( \frac{c}{b-a} \) and passes through the point \((a, 0)\), so the equation of the line is \( y - y_1 = m(x - x_1) \) \( \Rightarrow y = \frac{c}{b-a} \left( \frac{x-a}{x} \right) \).

Densities have to integrate to 1, so
\[
\int_a^b \frac{c}{b-a} (x-a) \, dx = 1 \Leftrightarrow c = \frac{2}{b-a}
\]

Easier way: area of a triangle is \( \frac{1}{2} \) (base)(height), so
\[
1 = \frac{1}{2} (b-a)c \quad \text{and} \quad c = \frac{2}{b-a}
\]

(25 Apr 19)