Given a probability space \((\Omega, \mathcal{F}, P)\), \(X\) is a **real random variable on** \((\Omega, \mathcal{F}, P)\) if and only if

1. \(X : \Omega \rightarrow [-\infty, \infty]\), which means that the domain of \(X\) is \(\Omega\) and \(X\) and the range of \(X\) is a subset of \([-\infty, \infty]\).

2. \(X\) is **measurable** with respect to \(\mathcal{F}\). This means that, for every interval \(I\), the set \(\{\omega : X(\omega) \in I\}\) is an event (is in \(\mathcal{F}\)). Consequently, we can discuss \(P(X \in I) = P(\{\omega : X(\omega) \in I\})\). Here \(I\) can be \((a, b)\), \((a, b]\), \([a, b)\), or \([a, b]\), for \(\infty \leq a < b \leq \infty\).

   - When \(\Omega\) is finite or countably infinite, and \(\mathcal{F}\) is the set of all subsets of \(\Omega\), then every \(X : \Omega \rightarrow [-\infty, \infty]\) is a random variable.
   - For the spaces \(U[0, 1]\) and \(N(0, 1)\), random variables have to “nice”. Continuous and step functions are measurable.
The textbook has a definition on page 7.

The textbook requires that, for all Borel subsets $B \subset [-\infty, \infty]$,\n\[
\{ \omega \in \Omega : X(\omega) \in B \} \in \mathcal{F}
\] (1)

Ramsey’s definition is less work than Equation 1, because it requires that you check Equation 1 only for intervals.

The two definitions are equivalent, which is a nice theorem.
Examples of RVs

Consider the repeated, independent tossing of a fair coin, stopping with the first H. Let $X(\omega) = i^2$ where $i$ is position of the first H and set $X(\omega_{\infty}) = \infty$. So $X(H) = 1$, $X(TH) = 4$, $X(TTH) = 9$, $X(TTTTH) = 16$, etc. For the interval $I = (9, 36]$,

$$\{ \omega : 9 < X(\omega) \leq 36 \} = \{ TTTH, TTTTH, TTTTTTH \}$$

and $P(X \in I) = 1/16 + 1/32 + 1/64 = 7/64$.

In general, if $X$ is a random variable and $g : [-\infty, \infty] \to [-\infty, \infty]$ is “nice”, then $g(X)$ is a random variable—$\exp(X) = e^X$ is an example. In general, “nice” functions include Borel functions. Borel functions include continuous and step functions.
Definition 1.2.3 (page 9): Let \( X \) be an RV on \((\Omega, \mathcal{F}, P)\). The distribution measure \( \mu_X \) is a function \( \mu_X : \mathcal{B}^* \rightarrow [0, 1] \) where
- \( \mathcal{B}^* \) is the set of Borel subsets of \([-\infty, \infty]\)
- For all \( A \in \mathcal{B}^* \), we have

\[
\mu_X(A) = P(X \in A) = P(\{ \omega \in \Omega : X(\omega) \in A \})
\]

Theorem 1: \([[-\infty, \infty], \mathcal{B}^*, \mu_X] \) is a probability space.

Theorem 2: Let \( X \) be an RV on \((\Omega, \mathcal{F}, P)\) and \( Y \) be an RV on \((\Omega', \mathcal{F}', P')\). Suppose that for all subintervals \((a, b]\) of \([-\infty, \infty]\) we have

\[
\mu_X((a, b]) = \mu_Y((a, b])
\]

Then \( \mu_X = \mu_Y \).
The conclusion of Theorem 2 also holds if you can prove Equation 2 for all closed sub-intervals of \([-\infty, \infty]\), or for all open sub-intervals of \([-\infty, \infty]\), or for all sub-intervals of \([-\infty, \infty]\) of the form \([a, b)\).

**Theorem 3:** Let \(\Omega = [-\infty, \infty]\) and \(\mathcal{F} = \mathcal{B}^*\) (the Borel subsets of \([-\infty, \infty]\). Suppose that

1. \((\Omega, \mathcal{F}, P_1)\) and \((\Omega, \mathcal{F}, P_2)\) are probability spaces
2. For all sub-intervals of \([-\infty, \infty]\) of the form \((a, b]\) we have \(P_1((a, b]) = P_2((a, b])\).

Then \(P_1 = P_2\).

Again, Theorem 3 remains valid if one works with all the closed intervals, all the open intervals, or all the intervals of the form \([a, b)\).
Let $(\Omega, \mathcal{F}, P)$ be the $U[0, 1]$ example. Let $X(x) = x$ for all $x \in [0, 1]$. For $[a, b] \in [-\infty, \infty]$, 

$$
\mu_X([a, b]) = \begin{cases} 
  b - a & \text{when } 0 \leq a \leq b \leq 1 \\
  1 & \text{when } a \leq 0 \text{ and } b \geq 1 \\
  1 - a & \text{when } 0 \leq a \leq 1 \text{ and } b \geq 1 \\
  0 & \text{when } 1 \leq a \leq b \\
  b & \text{when } a \leq 0 \text{ and } 0 \leq b \leq 1 \\
  0 & \text{when } a \leq b \leq 0 
\end{cases}
$$
Let $Y(x) = 1 - x$ for all $x \in [0, 1]$. Then $\mu_X = \mu_Y$. This can be proved using Theorem 2 above, and checking all 6 cases on the previous slide.

Let $Z(x) = x^{1/5}$ for all $x \in [0, 1]$. For $[a, b] \subset [-\infty, \infty]$ we have

$$
\mu_Z([a, b]) = \begin{cases} 
 b^5 - a^5 & \text{when } 0 \leq a \leq b \leq 1 \\
 1 & \text{when } a \leq 0 \text{ and } b \geq 1 \\
 1 - a^5 & \text{when } 0 \leq a \leq 1 \text{ and } b \geq 1 \\
 0 & \text{when } 1 \leq a \leq b \\
 b^5 & \text{when } a \leq 0 \text{ and } 0 \leq b \leq 1 \\
 0 & \text{when } a \leq b \leq 0
\end{cases}
$$
CDF (Cumulative Distribution Function)

Definition

Given a real random variable $X$ on some $(\Omega, \mathcal{F}, P)$, for every real $x$ let

$$F(x) = P(X \leq x) = P(\{\omega : X(\omega) \leq x\})$$

This is called the Cumulative Distribution Function (CDF) for $X$. In terms of it, all interval probability questions can be answered: for real $a < b$

- $P(X \in (a, b]) = F(b) - F(a)$
- $P(X \in [a, b]) = F(b) - F(a^-)$, where $F(a^-)$ is the left limit of $F$ at $a$.
- $P(X \in [a, b)) = F(b^-) - F(a^-)$.
- $P(X \in (a, b)) = F(b^-) - F(a)$.
- $P(X = a) = F(a) - F(a^-)$.
We can extend the formulas of the previous slide to include \( b = \infty \) if we agree that \( F(+\infty) = 1 \).

If we agree that
\[
F(-\infty) = \lim_{x \downarrow -\infty} F(x)
\]
then all the formulas of the previous slide are also valid with \( a = -\infty \) except where the formula invokes \( a^- \).

- Formulas with \(-\infty-\) in them make no sense, because there is no left limit at \(-\infty\).
- With this definition \( F(-\infty) = P(X = -\infty) \).

**Nice Theorem:** \( \mu_X = \mu_Y \) if and only if \( F_X = F_Y \).
Probability Density Function (PDF) for $dx$

Definition

Given: $X$ a real random variable on some $(\Omega, \mathcal{F}, P)$ and $F$ is its CDF. Suppose that $F$ has a constant $C$ such that

$$F(x) = \int_0^x F'(t) \, dt + C,$$

for all $x$.

That is, $F$ is nice enough that it is the integral of its own derivative. Then we call $F'$ the probability density function (PDF) for $X$ (and for $F$) with respect to $dx$. Consequently,

$$P(X \in (a, b]) = F(b) - F(a) = \int_a^b F'(x) \, dx$$

Often, people describe $X$ indirectly, by simply specifying $F'(x)$. They leave it up to you to create the corresponding $\Omega$ and $\mathcal{F}$ (which often you don’t need to know in detail).
Two Earlier Examples Have PDFs for $dx$

- Let $X(x) = x$ for $x \in [0, 1]$ for $U[0, 1]$. Let
  \[
  f_X(u) = \begin{cases} 
  0 & \text{when } u \leq 0 \\
  1 & \text{when } 0 < u < 1 \\
  0 & \text{when } u \geq 1 
  \end{cases}
  \]

  Then for all $[a, b] \subset [-\infty, \infty]$,
  \[
  \mu_X([a, b]) = \int_a^b f_X(u) \, du
  \]

- Let $Y(x) = 1 - x$ for $x \in [0, 1]$ for $U[0, 1]$. Then $f_Y = f_X$.

- Let $Z(x) = x^{1/5}$ for $x \in [0, 1]$ for $U[0, 1]$. Then
  \[
  f_Z(u) = \begin{cases} 
  0 & \text{when } u \leq 0 \\
  5u^4 & \text{when } 0 < u < 1 \\
  0 & \text{when } u \geq 1 
  \end{cases}
  \]
Suppose that that random variable $X$ has a finite number of distinct values, $\{x_j\}_{j=1}^N$, or a sequence of distinct values $\{x_j\}_{j=1}^\infty$, such that

$$1 = \sum_{j=1}^n P(X = x_j) \quad \text{or} \quad 1 = \sum_{j=1}^\infty P(X = x_j)$$

Then we call $X$ a discrete real random variable. Discrete is intended to evoke pebble-like and lumpy versus “mud"-like and smooth. Such RVs are often introduced by describing all values for $P(X = x_j)$. This description of $X$ is called a probability mass function. We add with it to compute probabilities. For any event $A$, we have

$$P(A) = \sum_{x_j \in A} P(X = x_j)$$
Examples of Discrete RVs

Let $0 < p < 1$ and set

$$P(X = j) = \binom{n}{j} p^j (1 - p)^{n-j}, \quad \text{for} \ 0 \leq j \leq n$$

This is the **binomial distribution**. It counts the number of heads in $n$ independent coin tosses, where the probability of $H$ is $p$ on each toss.

Let $\lambda > 0$ and set

$$P(X = j) = \frac{e^{-\lambda} \lambda^j}{j!}, \quad \text{for} \ 0 \leq j < \infty$$

This is the **Poisson distribution** for counting the number of rare events in a given time interval. For example, the number of tropical storms within 100 miles of Oahu within one year.
More Examples of Discrete RVs

Toss a fair coin (with independent tosses) until you get a first $H$. Let $W$ be the number of the toss on which the first $H$ occurs. Let $M = W^3$. Then both $M$ and $W$ are discrete, with these probability mass functions:

$$f_W(x) = \begin{cases} 2^{-x} & \text{if } x \text{ is a positive integer} \\ 0 & \text{otherwise} \end{cases}$$

Also,

$$f_M(x) = \begin{cases} 2^{-\frac{3}{\sqrt{x}}} & \text{if } x \text{ is the cube of a positive integer} \\ 0 & \text{otherwise} \end{cases}$$
Consider this experiment. First, you toss a fair coin. If you get H, you then spin a spinner that gives (with equal probability) a real number from the interval $[1, 6]$. If you get T, you are given the number $-1$. Let $X$ be the random number that you get from this process, and think of it as a random variable.

1. (Ito, Wang) Explain whether $X$ is discrete. If $X$ is discrete, describe completely is PMF.

2. (Kim, Won) Continue with $X$ as in the previous problem. Explain whether $X$ has a PDF for $dx$.

3. (Ishibashi, Urias) Describe completely the CDF for $X$.

4. (Fu, Raza) For all intervals of the form $(a, b] \subset [-\infty, \infty]$, give the value of $\mu_X((a, b])$. 

Ramsey Random Variables
5. (Liu, Wu) Let \((Ω, ℱ, P)\) be the \(N(0, 1)\) probability space. Let \(Y(x) = x\) for all \(x \in \mathbb{R}\) (here \(Ω = \mathbb{R}\)). Let \(Z = Y^4\). Explain that \(Z\) has a PDF for \(dx\) and find that PDF.

6. (Itano, Sidy D.) Continue with \(Z\) as in the previous problem. For all intervals of the form \((a, b] \subset [−∞, ∞]\), give the value of \(\mu_Z((a, b])\).

7. (Vu, Itoga) Continue with \(Z\) as in the previous problem. Find the CDF of \(Z\).