# Random Variables

Guy Lebanon

January 6, 2006

**Definition 1.** *A random variable (RV) $X$ is a function $X : \Omega \to \mathbb{R}$. The notation $\{X \in A\}$ represents the event $\{\omega \in \Omega : X(\omega) \in A\}$.*

Examples: a RV could be the sum of the dice when a pair of dice are rolled, the number of heads in three coin tosses, the lifetime of a flashlight battery etc. If $X$ is the sum of the dice when a pair of dice are rolled, $\{X = 3\} = \{(1, 2), (2, 1)\}$, $\{X = 4\} = \{(2, 2), (3, 1), (1, 3)\}$, $\{X < 3\} = \{(1, 1)\}$, $\{X < 2\} = \emptyset$. If $X$ is the sum of heads in three coin tosses, $\{X = 2\} = \{HHT, HTH, THH\}$, $\{X > 2\} = \{HHH\}$. The probability of sets concerning random variables such as $X \in A$ is simply $P(\{X \in A\}) = P(\{\omega \in \Omega : X(\omega) \in A\})$. For brevity, we sometime just write $P(X \in A)$ instead of $P(\{X \in A\})$.

Example: Let $X$ be the number of heads in three (ordered) fair coin tosses. The probabilities of sets concerning $X$ are

$$\begin{aligned}
P(X = 0) &= P(\{\omega \in \Omega : X(\omega) = 0) = P(\{TTT\}) = 0.5^3 = 1/8 \\
P(X = 1) &= P(\{\omega \in \Omega : X(\omega) = 1) = P(\{HTT, THT, TTH\}) = 3 \cdot 0.5^3 = 3/8 \\
P(X = 2) &= P(\{\omega \in \Omega : X(\omega) = 2) = P(\{HHT, HHT, HTH\}) = 3 \cdot 0.5^3 = 3/8 \\
P(X = 3) &= P(\{\omega \in \Omega : X(\omega) = 3) = P(\{HHH\}) = 0.5^3 = 1/8.
\end{aligned}$$

**Definition 2.** *A random variable $X$ is called discrete if there is a finite or countable set $K$ of real numbers such that $P(X \in K) = 1$. If $P(X = x) = 0$ for all $x \in \mathbb{R}$, $X$ is called a continuous RV. Note that a RV can be neither discrete nor continuous.*

If $\Omega$ is finite or countable, $X$ is always discrete (since the range of possible values of $X$ is at most countable). Similarly, if the range of $X$ is finite or countable, $X$ is discrete. In fact, we can assume without loss of generality that a discrete RV has a countable or finite range (recall that the range of a function $X : \Omega \to \mathbb{R}$ is the set $\{X(\omega) : \omega \in \Omega\}$. If $\Omega$ is continuous (for example $\mathbb{R}$), $X$ may be discrete, continuous, or neither.

Example: Consider the dart throwing experiment (under classical interpretation) with $\Omega = \{(x, y) : x, y \in \mathbb{R}, x^2 + y^2 < 1\}$ and $X(\omega) = 0$ if the dart hit the bullseye $\omega \in \{x, y \in \mathbb{R} : \sqrt{x^2 + y^2} < 0.1\}$ and 1 otherwise. $X$ is a discrete random variable because $P(X \in \{0, 1\}) = 1$. Now, consider $X$ to be the distance of the location of the dart from the center,

$$X(\omega) = X((x, y)) = \sqrt{x^2 + y^2}.$$

$X$ is a continuous random variable since for any real number $r$ the set $\{X = r\} = \{(x, y) : x, y \in \mathbb{R}, \sqrt{x^2 + y^2} = r\}$ is a circle with 2D volume (or area) zero, and hence $P(X = r) = 0$, $\forall r$. We also have, for example,

$$P(0.2 < X < 0.5) = P(\{(x, y) : x, y \in \mathbb{R}, 0.2 < \sqrt{x^2 + y^2} < 0.5\}) = \frac{\pi(0.5^2 - 0.2^2)}{\pi 1^2}.$$

**Definition 3.** *For a RV $X$, we define the probability mass function (pmf) as the function $p_X : \mathbb{R} \to \mathbb{R}$ given by $p_X(x) = P(X = x)$.*

**Definition 4.** *For a RV $X$ we define the cumulative distribution function (cdf) as the function $F_X : \mathbb{R} \to \mathbb{R}$ given by $F_X(x) = P(X \leq x)$.*

**Definition 5.** *For a RV $X$ we define the probability density function (pdf) $f_X : \mathbb{R} \to \mathbb{R}$ as the derivative of the cdf $\frac{dF_X(x)}{dx}$ where it exists, and 0 elsewhere.*

The cdf is used for both discrete and continuous RVs. The pmf is used for discrete RV only (for continuous RVs it is identically zero) and the pdf is used for continuous RV only (for discrete RVs it is identically zero). In the coin tossing example above, the pmf and cdf are

$$p_X(x) = P(X = x) = \begin{cases} 1/8 & x = 0 \\ 3/8 & x = 1 \\ 3/8 & x = 2 \\ 1/8 & x = 3 \\ 0 & \text{else} \end{cases} \qquad F_X(x) = P(X \le x) = \begin{cases} 0 & x < 0 \\ 1/8 & 0 \le x < 1 \\ 1/2 & 1 \le x < 2 \\ 7/8 & 2 \le x < 3 \\ 1 & 3 \le x \end{cases}.$$

For the dart throwing experiment and the continuous RV the cdf and pdf are

$$F_X(x) = P(X \le x) = \begin{cases} 0 & 0 < x \\ \frac{\pi x^2}{\pi 1^2} = x^2 & 0 \le x < 1 \\ 1 & 1 \le x \end{cases} \qquad f_X(x) = \begin{cases} 0 & 0 < x \\ 2x & 0 \le x < 1 \\ 0 & 1 \le x \end{cases}.$$

### Properties of pmf

The pmf $p_X$ of a discrete RV satisfies the following properties

1. $p_X(x) \ge 0$ (since $P(A) \ge 0$)
2. $\{x \in \mathbb{R} : p_X(x) \neq 0\}$ is finite or countable (by definition of discrete RV)
3. $\sum_x p_X(x) = 1$, where by (2) the summation here contains only a finite or countable number of terms:
$$1 = P(\Omega) = P(\cup_x \{X = x\}) = \sum_x P(X = x) = \sum_x p_X(x).$$

A function $p_X$ that satisfies the above properties is a valid pmf for some discrete RV.

### Properties of cdf

We use $F_X(x^-)$ to denote the limit of $F_X$ from the left $\lim_{0<\Delta\to 0} F_X(x - \Delta)$ and $F_X(x^+)$ the limit from the right $\lim_{0<\Delta\to 0} F_X(x + \Delta)$. Note that $F_X(x^-) = P(X < x)$. The cdf $F_X$ of a RV satisfies:

1. $$\lim_{x\to -\infty} F_X(x) = \lim_{x\to -\infty} P(X \le x) = P(\emptyset) = 0$$
2. $$\lim_{x\to \infty} F_X(x) = \lim_{x\to \infty} P(X \le x) = P(\Omega) = 1$$
3. $F_X$ is a non-decreasing function - that is if $a < b$, $F_X(a) \le F_X(b)$ (since $P(B) \ge P(A)$ if $A \subset B$)
4. $F_X$ is continuous from the right $F_X(x) = F(x^+)$ (without proof)

A function $F_X$ that satisfies the above properties is a valid cdf for some RV.

In addition, we have the following consequences regarding cdf: for all RV $X$ and $a < b$, $P(a < X \le b) = F_X(b) - F_X(a)$, $P(a \le X \le b) = F_X(b) - F_X(a^-)$, $P(a < X < b) = F_X(b^-) - F_X(a)$, $P(a \le X < b) = F_X(b^-) - F_X(a^-)$ and $P(X = x) = F_X(x) - F_X(x^-)$. Since $P(X = x) = F_X(x) - F_X(x^-)$ we have that $F_X$ is continuous from the left if and only if $X$ is a continuous RV (in this case for all $x$, $P(X = x) = 0$ and $F_X(x) = F_X(x^-)$). Therefore, a RV is continuous if and only if its cdf is everywhere continuous and there is no distinction between $F_X(a^-), F_X(a), F_X(a^+)$.

**Definition and properties of pdf (continuous RV)**

A pdf of a continuous RV $X$ satisfies the following

1. $f_X$ is a non-negative function (since $f_X$ is the derivative of a function $F_X$ that is non-decreasing)
2. $\int_{-\infty}^{\infty} f_X(x)dx = 1$. This may be proven by

$$\int_{-\infty}^{\infty} f_X(x)dx = \lim_{n\to\infty} \int_{-n}^{n} f_X(x)dx = \lim_{n\to\infty} P(-n \le X \le n) = \lim_{n\to\infty} (F_X(n) - F_X(-n)) = 1 - 0 = 1.$$

A function $f_X$ that satisfies the above properties is a valid pdf for some continuous RV.

The pdf has the following intuition. Since by definition

$$f_X(x) = \lim_{0<\Delta\to 0} \frac{F_X(x+\Delta) - F_X(x)}{\Delta}$$

if $\Delta$ is a small positive number

$$f_X(x)\Delta \approx F_X(x+\Delta) - F_X(x) = P(x < X < x + \Delta).$$

**Theorem 1.** *Let $X : \Omega \to \mathbb{R}$ be a RV. Then for any subset $A$ of real numbers,*

$$P(X \in A) = \begin{cases} \sum_{x\in A} p_X(x) & X \text{ is discrete} \\ \int_A f_X(x)dx & X \text{ is continuous} \end{cases} \tag{1}$$

*Proof.* We prove only the the discrete case

$$P(X \in A) = P(\cup_{x\in A}\{X = x\}) = \sum_{x\in A} P(X = x) = \sum_{x\in A} p_X(x).$$

□

Example: recall the example above of summing the number of heads in three coin tosses. As shown above pmf of X is $p_X(0) = 1/8, p_X(1) = 3/8, p_X(2) = 3/8, p_X(3) = 1/8$ and indeed $\sum_{x\in X} p_X(x) = 1$. By the theorem above $P(X > 0) = \sum_{x\in\{1,2,\ldots\}} p_X(x) = p_X(1) + p_X(2) + p_X(3) = 7/8$.

For the dart throwing example and the continuous RV, we already computed above the cdf. The pdf is $f_X(x) = 2x$ for $0 < x < 1$ and 0 otherwise. By the theorem above the probability that we hit the bullseye is

$$P(X < 0.1) = P(X \in (-\infty, 0.1)) = \int_{-\infty}^{0.1} f_X(x)dx = \int_0^{0.1} 2xdx = x^2\Big|_0^{0.1} = 0.01$$

As mentioned above a RV is completely determined by a valid cdf, as well as pmf (for discrete) or pdf (for continuous). For this reason, we will sometime supply the cdf, pmf or pdf of a RV. In this case, we can omit the sample space $\Omega$ and the original probability distribution on events $E \subset \Omega$. The probability of events concerning the RV is completely determined by the supplied cdf, pmf or pdf.