Random Variables

Statistics 110

Summer 2006

Random Variables

A **Random Variable** (RV) is a response of a random phenomenon which is numeric.

Examples:

1. Roll a die twice and record the sum. (Discrete RV)

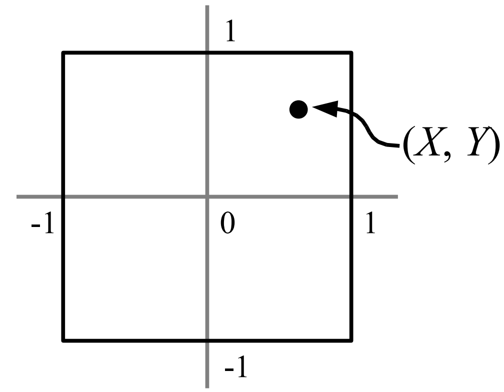
2. The coin flip example where the response X is the flip number of the first tail. (Discrete, countable RV)

3. Count α particles emitted from a low intensity radioactive source.

   Let $T_i =$ waiting time between the $i - 1^{th}$ and the i^{th} emissions. (Continuous RV)

   Let $Y_T = \#$ of emissions in time interval $[0, T]$. (Discrete RV)

4. Throw a dart at a square target and record the X and Y coordinates of where the dart hits. $Z = (X, Y)$ is a random vector. (Both are continuous RVs)



- Discrete: The values taken come from a discrete set. There may be a finite number of possible outcomes (e.g. 11 as for the sum of two dice), or countable (infinite, as for the flip that the first tail occurs).

- Continuous: The values taken come from an interval (possibly infinite). In the dart example $X \in [-1, 1]$ (finite), whereas in the radiation example $T_i \in [0, \infty)$ (infinite).

Note that RVs are usually indicated by capital letters, usually from the end of the alphabet (X, Y, Z etc). Letters from the beginning of the alphabet are left to denote events.

Instead of working with events, it is often easier to work with RVs. In fact for any event $A \subset \Omega$, we can define an indicator RV

$$I_A = \begin{cases} 1 & \text{If } A \text{ occurs} \\ 0 & \text{Otherwise} \end{cases}$$

From this definition, we get

$$\begin{aligned} I_{A^c} &= 1 - I_A \\ I_{A \cap B} &= I_A I_B \\ I_{A \cup B} &= I_A + I_B - I_A I_B \end{aligned}$$

Instead of using logic and set operations when working with events, we use algebra and arithmetic operations on RVs

For something to be a random variable, arithmetic operations have to make sense. I would not classify the following as a RV

$$X = \begin{cases} 1 & \text{if Fred} \\ 2 & \text{if Ethel} \\ 3 & \text{if Ricky} \\ 4 & \text{if Lucy} \end{cases}$$

What does $1 + 4$ (i.e. Fred + Lucy) mean here.

A RV X is a function from Ω (domain) to the real numbers. The range of the function (the values it takes) will, obviously, depend on the problem

Example: Flip a biased coin 3 times and let X be the number of heads. Assume the flips are independent with $P[H] = p$ and $P[T] = 1 - p$.

ω	$X(\omega)$	$P[\omega]$
HHH	3	p^3
HHT	2	$p^2(1-p)$
HTH	2	$p^2(1-p)$
THH	2	$p^2(1-p)$
HTT	1	$p(1-p)^2$
HTT	1	$p(1-p)^2$
HTT	1	$p(1-p)^2$
TTT	0	$(1-p)^3$

For this example, the range of $X = \{0, 1, 2, 3\}$.

The probability measure on the original sample space, Ω, gives the probability measure for X.

For the above example, the probability measure is defined by

$$P[X = i] = \binom{3}{i} p^i (1-p)^{3-i}$$

This happens to be the probability mass function for X.

Definition: Assume that the discrete RV X takes the values $\{x_1, x_2, \ldots, x_n\}$ (n possibly ∞). The **Probability Mass Function** (PMF) of the discrete RV X is a function on the range of X that gives the probability for each possible value of X:

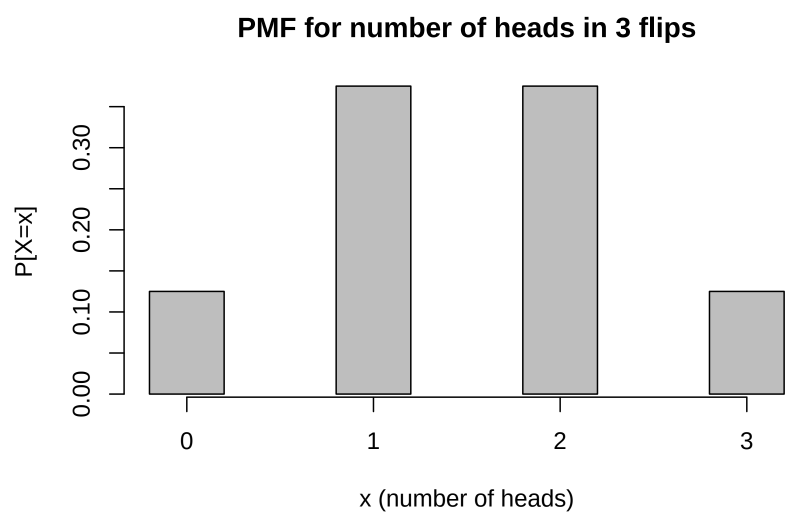
$$\begin{aligned} p_X(x_i) &= P[X = x_i] \\ &= \sum_{\omega : X(\omega) = x_i} P[\omega] \end{aligned}$$

Any valid PMF $p(x)$ must satisfy

- $p_X(x_i) \geq 0$

- $\sum_i p_X(x_i) = 1$

In the coin flipping example, assume that $p = 0.5$ (i.e. the coin is fair). Then the probability histogram of the PMF looks like



Definition: The **Cumulative Distribution Function** (CDF) of a RV X is a function on the range of X that gives the probability of being less than or equal x

$$F(x) = P[X \leq x]$$

The CDF is a nondecreasing function satisfying

$$\lim_{x \to -\infty} F(x) = 0 \quad \text{and} \quad \lim_{x \to \infty} F(x) = 1$$

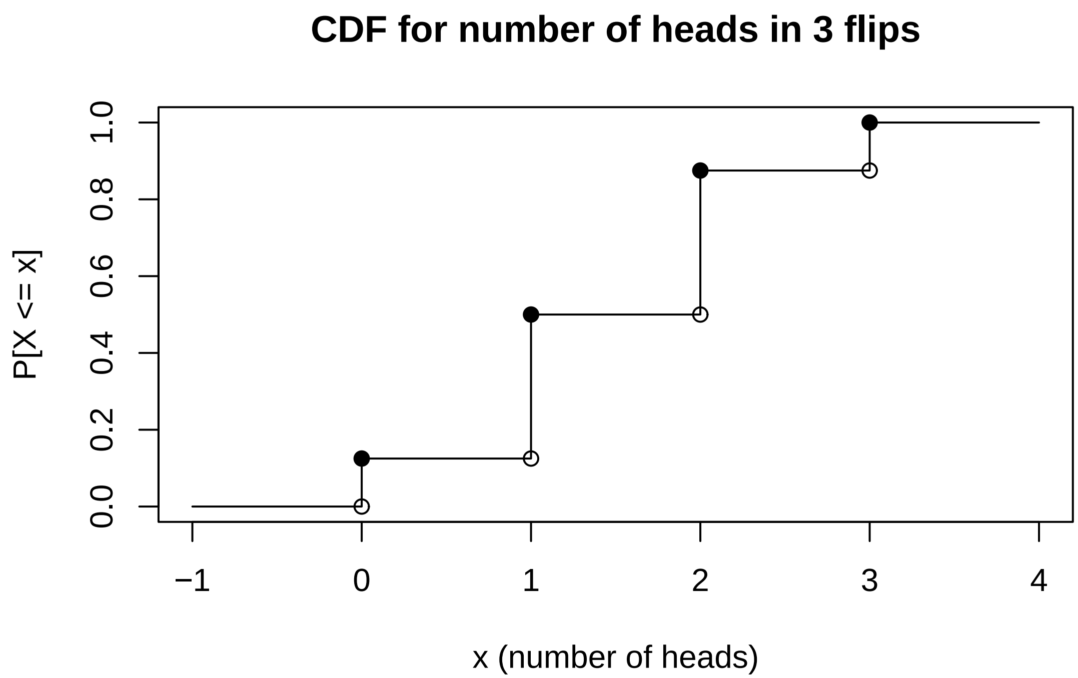
The PMF and the CDF are closely related as (for discrete RVs)

$$F(x) = \sum_{x_i : x_i \leq x} p(x_i)$$

Also, assuming that $x_1 < x_2 < \ldots$

$$p(x_i) = F(x_i) - F(x_{i-1})$$

The CDF for the coin flipping example looks like



For a discrete RV, the CDF is a step function, with jumps of $p(x_i)$ at each x_i. Also it is right continuous. For the coin flipping example

$$\lim_{x \to 0-} F(x) = 0 \text{ and } \lim_{x \to 0+} F(x) = p(0)$$

For discrete RVs, independence is easily defined. Let X and Y be two discrete RVs taking values $x_1, x_2, \ldots$ and $y_1, y_2, \ldots$ respectively. Then X and Y are said to be independent if for all i and j

$$P[X = x_i, Y = y_j] = P[X = x_i]P[Y = y_j]$$

(You don't need to check all possible events involving X and Y.)

For three or more RVs, the obvious extension is the case.

Expected Values

Definition: The **Expected Value** of a RV X (with PMF $p(x)$) is

$$E[X] = \sum_x xp(x)$$

assuming that $\sum_i |x_i|p(x_i) < \infty$. This is a technical point, which if ignored, can lead to paradoxes.

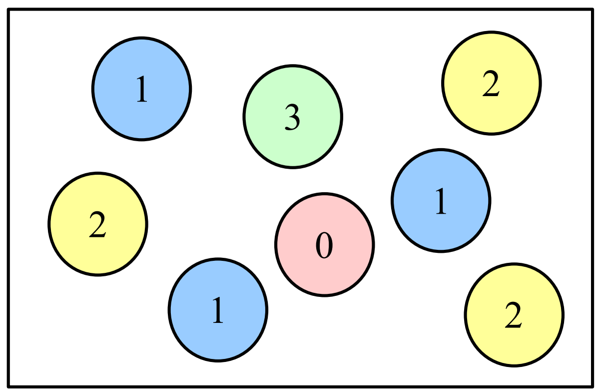
There are well defined RV that don't have a finite expection. For example, see the St. Petersberg Paradox (Example D, page 113).

- $E[X]$ can be thought of as an average

  Example: Flipping 3 coins

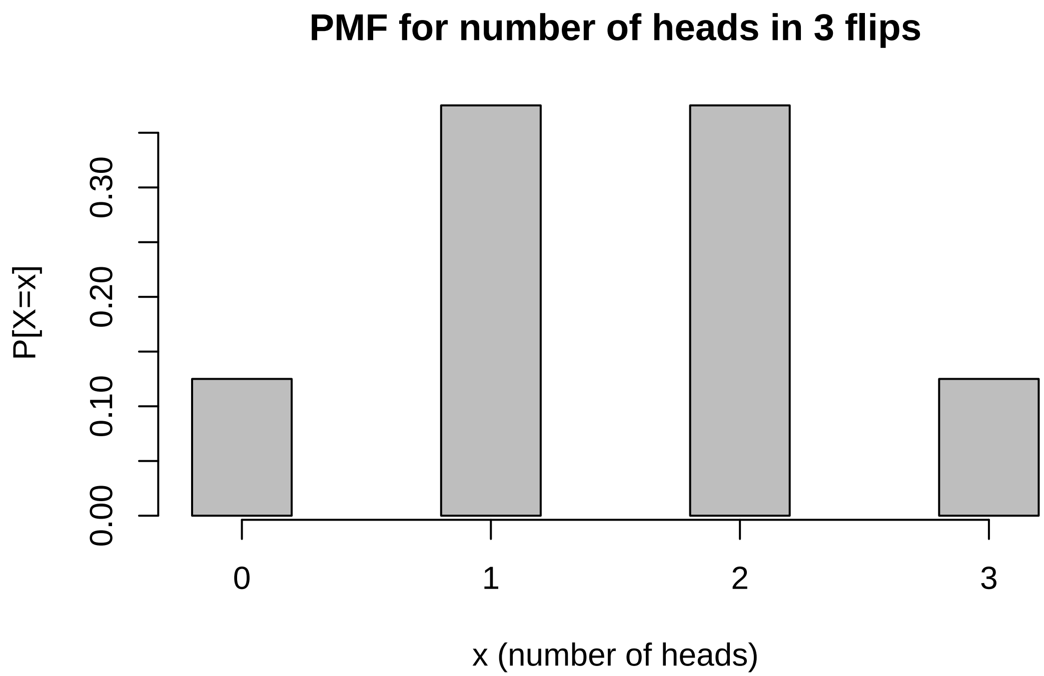
x	0	1	2	3
$p(x)$	$\frac{1}{8}$	$\frac{3}{8}$	$\frac{3}{8}$	$\frac{1}{8}$

$$E[X] = 0 \times \frac{1}{8} + 1 \times \frac{3}{8} + 2 \times \frac{3}{8} + 3 \times \frac{1}{8} = 1.5$$



$E[X]$ is the average value of the balls in this urn.

$E[X]$ serves as a “central” or “typical” value of the distribution.


PMF for number of heads in 3 flips
P[X=x]
0.00
0.10
0.20
0.30
0
1
2
3
x (number of heads)

- You bet an amount θ on each draw (with replacement) from the urn for X. The payoff in each draw is the value drawn (or $X(\omega)$ for the sample point drawn). What is the bet θ^* that makes this a fair game? [In the sense that in the long run, the winning or loss will become smaller and smaller relative to the total amount of your bets.]

  Answer: $\theta^* = E[X]$

  Justification: To come (Law of Large Numbers)

Theorem. *If $Y = g(X)$, then*

$$E[Y] = \sum_x g(x) p_X(x)$$

Proof. Since Y is also a RV is also has its own PMF.

$$\begin{aligned} p_Y(y) &= P[\{x : g(x) = y\}] \\ &= \sum_{\{x:g(x)=y\}} p_X(x) \end{aligned}$$

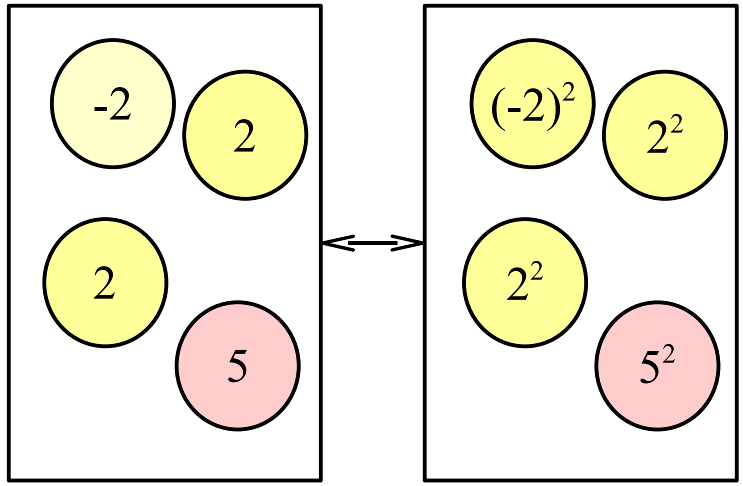
Therefore

$$
\begin{aligned}
E[Y] &= \sum_y y p_Y(y) \\
&= \sum_y y \sum_{\{x : g(x) = y\}} p_X(x) \\
&= \sum_y \sum_{\{x : g(x) = y\}} g(x) p_X(x) \\
&= \sum_x g(x) p_X(x)
\end{aligned}
$$

□

So instead of figuring out the PMF for the transformed RV, we can directly calculate the expected value.

Example: Let $Y = X^2$



x	-2	2	5
$p_X(x)$	$\frac{1}{4}$	$\frac{1}{2}$	$\frac{1}{4}$

y	4	25
$p_Y(y)$	$\frac{3}{4}$	$\frac{1}{4}$

Example: Let $X = I_A$ for some event A. Then

$$E[X] = 1P[A] + 0P[A^c] = P[A]$$

$$E[X^2] = 1P[A] + 0P[A^c] = P[A]$$

In general $E[g(X)] \neq g(E[X])$

- Expected value of a linear function

  If a and b are constants, then

$$E[a + bX] = a + bE[X]$$

  Proof. Simple algebra $\square$

- Expected value of a sum of 2 RVs

$$E[Y + Z] = E[Y] + E[Z]$$

  Note that this only defined if Y and Z are defined on a common sample space.

Proof. Both Y and Z are RVs on a common sample space Ω. Then

$$\begin{aligned} E[Y+Z] &= \sum_{\omega} (Y+Z)(\omega)P[\omega] \\ &= \sum_{\omega} (Y(\omega)+Z(\omega))P[\omega] \\ &= \sum_{\omega} Y(\omega)P[\omega] + \sum_{\omega} Z(\omega)P[\omega] \\ &= E[Y] + E[Z] \end{aligned}$$

□

For practical reasons, the results of this may not make sense, even if it is well defined. For example, let $Y =$ number of coins in pocket and $Z =$ time to get to Science Center from home, when students in the class are sampled.