

5 Random Variables

Example 5.1. Experiment: Rolling a pair of dice

$$\begin{aligned}\Omega &= \text{all possible outcomes of rolling a pair of dice} \\ X &= \text{the sum of the two dice} \\ Y &= \text{the product of the two dice} \\ Z &= \text{the minimum of the two dice}\end{aligned}$$

X , Y , and Z are functions.

1. The domain of X is $\underline{1 \leq X \leq 6}$. The range of X is $\underline{12 - 2 = 10}$
2. The domain of Y contains 36 elements. The range of Y is $\underline{1 \leq Y \leq 36}$

Definition 5.2. A *random variable* is a function defined on an outcome space Ω that assigns each outcome a real number. That is, if X is the name of the random variable and s is a possible outcome in the sample space, then $X(s)$ is a real number. A *discrete random variable* has an outcome space that is finite or countably infinite (like the integers).

Example 5.3. Experiment: Flip a fair coin three times
Given heads is 1 and tails is 0, some examples of random variables are the following

$$\begin{aligned}A &= \text{the sum of the three coin flips} \\ B &= \text{the product of the three coin flips} \\ C &= \text{the minimum of the three coin flips}\end{aligned}$$

Some examples of events in the other hand are the following

$$\begin{aligned}D &= \text{getting one heads} \\ E &= \text{getting all heads} \\ F &= \text{getting all tails}\end{aligned}$$

Definition 5.4. Let X be a random variable with finitely many range values. The **mass function** of X is the function that gives the probability of each possible value of X . The mass function is

$$p(x) := \mathbb{P}(X = x).$$

To represent the mass function, you can plot it, list it as a table, or write it as a function.

Example 5.5. Teams play a game. Each team gets 36 tokens to distribute between the numbers 2 through 12. Once all teams have chosen distributions, the host will roll two dice. If the sum of the rolls is a number where the team has placed a token, they may remove one token from that number. The winning team is the first team to remove all 36 of their tokens. What is the best way to distribute the 36 tokens?

Solution. The best way to distribute the tokens are by placing 18 tokens on 2 and 18 tokens on 12. Given X is the random variable equal to the sum of the two dice rolls and $p(x) = \mathbb{P}(X = x)$,

$$\begin{array}{lll} p(2) = p(12) = \frac{1}{36} & p(3) = p(11) = \frac{2}{36} & p(4) = p(10) = \frac{3}{36} \\ p(5) = p(9) = \frac{4}{36} & p(6) = p(8) = \frac{5}{36} & p(7) = \frac{6}{36} \end{array}$$

Meaning 2 and 12 are the best numbers to bet on as they have the least likelihood of being rolled and thus the best way to maximize keeping the team's tokens. \square

Definition 5.6. A *probability distribution* is a function that describes the probabilities of all events in an outcome space. A probability mass function completely describes a distribution.

Notation 5.7. Let X be a random variable and B an event in the *range* of X (so B is some set of real numbers).

$$\mathbb{P}(X \in B) = \sum_{x \in B} \mathbb{P}(X = x)$$

The symbol \in means “is an element of.”

Example 5.8. Let X be the number of heads out of 100 (fair) coin flips. Let event $A = \{2, 3\}$. Compute $\mathbb{P}(X \in A)$.

Solution.

$$\begin{aligned} \mathbb{P}(X \in A) &= \sum_{x \in A} \mathbb{P}(X = x) \\ &= \mathbb{P}(X = 2) + \mathbb{P}(X = 3) \\ &= \frac{1}{2^{100}} \left(\binom{100}{2} + \binom{100}{3} \right) = \frac{1}{2^{100}} (4950 + 161700) = \frac{166650}{2^{100}} \end{aligned}$$

\square

5.1 Functions of Random Variables

Definition 5.9. Two random variables are *independent* if

$$\mathbb{P}(X = x, Y = y) = \mathbb{P}(X = x)\mathbb{P}(Y = y) \quad \text{for all } x \text{ and } y$$

Or equivalently

$$\mathbb{P}(X = x|Y = y) = \mathbb{P}(X = x) \quad \text{for all } x \text{ and } y$$

Example 5.10. Experiment: Rolling a pair of dice

An example of independent random variables are the following

A = the value of the first die

B = the value of the second die

An example of dependent random variables on the other hand are the following

C = the number of even dice rolls

D = the number of odd dice rolls

Remark 5.11. Let X be a random variable. We can create a new random variable Y as a function of X . In this case, X and Y are often dependent. We could define functions like $Y = X^2$, $Y = |X - 50|$, or $Y = 13X - X^2 + \ln(X + 1)$. Each of these is a new random variable.

Example 5.12. Let X be the number of heads out of 100 coin tosses.

If we define Y to be $100 - X$, we find the distribution of Y to be the same as X given the fact $\binom{n}{k} = \binom{n}{n-k}$ or specifically $\binom{100}{k} = \binom{100}{100-k}$.

$$\begin{aligned}\mathbb{P}(X = k) &= \mathbb{P}(100 - X = 100 - k) \\ &= \mathbb{P}(Y = 100 - k) = \frac{1}{2^{100}} \binom{100}{k} \\ &= \mathbb{P}(Y = k) = \frac{1}{2^{100}} \binom{100}{100 - k}\end{aligned}$$

Theorem 5.13. Let X be a discrete random variable, and $Y = f(X)$. The probability mass function of Y is

$$\mathbb{P}(Y = y) = \sum_{x \text{ with } f(x)=y} \mathbb{P}(X = x)$$

Proof of thm 5.13.

$$\begin{aligned}\mathbb{P}(Y = y) &= \mathbb{P}(f(X) = y) \\ &= \sum_{x:f(x)=y} \mathbb{P}(X = x)\end{aligned}$$

□

Remark 5.14. Two random variables X and Y are equal if both random variables output the same value for each element of the outcome space. Suppose you roll two dice. X = number on the first die. Y = number on the second die. Do X and Y have the same distribution? Are X and Y equal? As the random variables X and Y are dependent on the number rolled on a die, they have the same distribution, specifically $\frac{1}{6}$ for all possible values for X and Y . However, as they are dependent on different dices, X and Y are not equal on all outcomes s of the outcome space Ω .

5.2 Joint Distributions

Definition 5.15. The *joint distribution* of two random variables gives the probability of all possible pairs of outcomes. So, the joint distribution of X and Y is a list of (or description of) the probabilities $\mathbb{P}(X = x \text{ and } Y = y)$ for all possible pairs (x, y) . It is often helpful to express the joint distribution as a table/grid.

Example 5.16. Several examples of joint distributions

1. Two dice are rolled. X is the value of the first die, Y is the value of the second. What is the joint distribution of X and Y ?

	1	2	3	4	5	6
1	1/6	1/6	1/6	1/6	1/6	1/6
2	1/6	1/6	1/6	1/6	1/6	1/6
3	1/6	1/6	1/6	1/6	1/6	1/6
4	1/6	1/6	1/6	1/6	1/6	1/6
5	1/6	1/6	1/6	1/6	1/6	1/6
6	1/6	1/6	1/6	1/6	1/6	1/6

2. Two coins are flipped. Given heads is 1 and tails is 0, X is the sum of the two flips, Y is the value of the second flip. The joint distribution of X and Y is

	0	1	2
0	1/4	1/4	0/4
1	0/4	1/4	1/4

Definition 5.17. The *marginal distribution* of a random variable X , is the probability mass function of X . The mass function is called the marginal distribution of X when it is calculated from the joint distribution of X with another random variable (because the calculation can be done by summing rows/columns and writing the results in the margins of the joint distribution table).

Example 5.18. By taking the sum of the columns and rows respectively from the above joint distribution table, we can determine the marginal distribution of the random variables X and Y .

	0	1	2	Total
0	1/4	1/4	0/4	2/4
1	0/4	1/4	1/4	2/4
Total	1/4	2/4	1/4	4/4

$$\begin{aligned}
P(X = 0) &= \frac{1}{4} \\
P(X = 1) &= \frac{2}{4} = \frac{1}{2} \\
P(X = 2) &= \frac{1}{4} \\
P(Y = 0) &= \frac{2}{4} = \frac{1}{2} \\
P(Y = 1) &= \frac{2}{4} = \frac{1}{2}
\end{aligned}$$

Theorem 5.19. (*Marginal distribution*)

$$\mathbb{P}(X = x) = \sum_{\text{all } y} \mathbb{P}(X = x, Y = y)$$

Proof of theorem 5.19. Summing all the $P(X = x, Y = y)$ for all values of y is equivalent to taking the sum of a column where $X = x$ in a marginal distribution table. \square

Definition 5.20. Several random variables are mutually (collectively) independent if

$$\mathbb{P}(X_1 = x_1, X_2 = x_2, \dots, X_n = x_n) = \prod_{i=1}^n \mathbb{P}(X = x_i) \quad \text{for all } x_i$$

Example 5.21. Let

$$\begin{aligned}
X &= \text{even numbers} \\
Y &= \#s \text{ above } 2
\end{aligned}$$

The random variables above are mutually independent.

Theorem 5.22. (*Consequences of independence*)

1. Any subgroup of independent variables are also independent
2. Functions of independent random variables are also independent
Example: If X_1 and X_2 are independent, then $2X_1$ and $X_2 + 3$ are also independent
3. Functions of collections disjoint collections of independent random variables are also independent
Example: If X_1, X_2, X_3, X_4 are independent, then $\min(X_1, X_2)$ and $\max(X_3, X_4)$ are also independent.