

Random Variables

Definition: Let S be a sample space. A function $X : S \rightarrow \mathbb{R}$, that associates a real number $X(s)$ to each outcome s is called a *random variable*.

Notation: The range of the random variable is denoted R_X . That is, the R_X is the set of all possible values that X can take.

Note: We use upper-case letters (often at the end of the alphabet) to denote random variables, e.g. X, Y, Z, W, \dots

Note: As an outcome is observed and we evaluate the random variables, we obtain an observed value. The observed values are denoted with lower-case letters, e.g. x, y, z, w, \dots

Classification: Let X be a random variable with range R_X .

1. If the range R_X is a finite set or it is countable, then we say that X is a *discrete* random variable.
2. If the range R_X is an interval (finite or infinite) of real numbers, then X is a *continuous* random variable.

Example 1: Consider calls to a communication system. Let X be the waiting time between two calls. Let Y be the number of calls between 8 am and 9 am.

(a) Determine an appropriate range for these random variables.

Solution: $R_X = [0, \infty)$ and $R_Y = \{0, 1, 2, 3, \dots\}$

(b) Are these random variables discrete or continuous?

Solution: X is continuous and Y is discrete.

Events with random variables: We can construct events with random variables. Here are a few examples:

1. Let $A \subseteq \mathbb{R}$, we interpret $\{X \in A\}$ as the following event:

$$\{X \in A\} = \{s \in S : X(s) \in A\}.$$

2. Let x be a real number, we interpret $\{X = x\}$ as the following event

$$\{X = x\} = \{s \in S : X(s) = x\}.$$

3. Let x be a real number, we interpret $\{X \leq x\}$ as the following event

$$\{X \leq x\} = \{s \in S : X(s) \leq x\}.$$

Furthermore,

$$P(X \in A) = P(\{s \in S : X(s) \in A\}).$$

Discrete Random Variables

Probability Distribution : We will specify the probabilities associated with the possible values of a discrete random variable X with the following functions:

- (a) its probability mass function (pmf), denoted by p_X .
- (b) its cumulative distribution function (cdf), denoted by F_X .

Définition: Let X be a random variable with sample space S . Its *probability mass function* p_X is defined by

$$p_X(x) = P(X = x) = P(\{s \in S : X(s) = x\}), \quad \text{for } x \in R_X.$$

The *cumulative distribution function* F_X is defined by

$$F_X(x) = P(X \leq x) = P(\{s \in S : X(s) \leq x\}).$$

Remark: We can interpret a probability as a “mass” that we attribute to a number on the real line, where the total mass on the entire real line is 1. The cumulative distribution function $F_X(x)$ gives us the total amount of this mass on the interval $(-\infty, x]$.

Properties of the probability mass function :

(a) $0 \leq p_X(x) \leq 1$ for all x

(b) $\sum_{x \in R_X} p_X(x) = 1$

Example 19 : An electronic device contains two components such that each works independently of the other. The probability that the first component is defective is 0.1 and the probability that the second component is defective is 0.2. Let X be the number of defective components in the device.

(a) Give the range of X ?

Solution: $R_X = \{0, 1, 2\}$

(b) Determine the p.m.f. of X .

Solution: Let A be the event that the first component is defective and let B be the event that the second component is defective. We are told that $P(A) = 0.1$ and $P(B) = 0.2$. By independence, we have that $P(A \cap B) = P(A) \times P(B)$ and $P(A' \cap B') = P(A') \times P(B')$, so

$$p_X(0) = P(X = 0) = P((A \cup B)') = P(A' \cap B') = P(A') \times P(B')$$

$$= (1 - P(A)) \times (1 - P(B)) = (1 - 0.1) \times (1 - 0.2) = 0.72$$

$$p_X(1) = P(X = 1) = P([A \cup B] \cap [A \cap B]') = P(A) + P(B) - 2P(A \cap B) = P(A) + P(B) - 2P(A \cap B)$$

$$= P(A) + P(B) - 2P(A) \times P(B) = 0.1 + 0.2 - 2 \times 0.1 \times 0.2 = 0.26$$

$$p_X(2) = P(X = 2) = P(A \cap B) = 0.1 \times 0.2 = 0.02$$

$$p_X(x) = \begin{cases} 0.72; & x = 0 \\ 0.26; & x = 1 \\ 0.02; & x = 2 \end{cases}$$

(c) Determine the c.d.f. of X . **Solution:**

$$F_X(x) = \begin{cases} 0; & x < 0 \\ 0.72; & 0 \leq x < 1 \\ 0.98; & 1 \leq x < 2 \\ 1; & x \geq 2 \end{cases}$$

(d) What is the probability that there will be at least one defective component?

Solution: $P(X \geq 1) = 1 - P(X < 1) = 1 - p_X(0) = 1 - 0.72 = 0.28$

(e) What is the probability that there are more than 2 defective components?

Solution: $P(X > 2) = 1 - P(X \leq 2) = 1 - F_X(2) = 1 - 1 = 0$

(f) Determine $P(-0.2 < X \leq 1.5)$? **Solution:** $P(-0.2 < X \leq 1.5) = F_X(1.5) - F_X(-0.2) = 0.98 - 0 = 0.98$

Example 20 : Let X be the number of calls to a service center during a 15 minute time interval. Suppose that X has the following cumulative distribution function.

$$F_X(x) = \begin{cases} 0; & x < 0 \\ 0.1; & 0 \leq x < 1 \\ 0.35; & 1 \leq x < 2 \\ 0.55; & 2 \leq x < 3 \\ 0.65; & 3 \leq x < 4 \\ 0.95; & 4 \leq x < 5 \\ 1; & x \geq 5 \end{cases}$$

(a) Determine the range of X and give its probability mass function.

Solution: $R_X = \{0, 1, 2, 3, 4, 5\}$

$$p_X(x) = \begin{cases} 0.1; & x = 0 \\ 0.25; & x = 1 \\ 0.20; & x = 2 \\ 0.10; & x = 3 \\ 0.30; & x = 4 \\ 0.05; & x = 5 \end{cases}$$

(b) Calculate the following probabilities in two ways : (i) using p_X ; (ii) using F_X .

(i) $P(X \geq 4)$ **Solution:**

$$P(X \geq 4) = p_X(4) + p_X(5) = 0.3 + 0.05 = 0.35$$

$$P(X \geq 4) = 1 - P(X < 4) = 1 - \lim_{x \uparrow 4} F_X(x) = 1 - 0.65 = 0.35$$

(ii) $P(1 \leq X < 4)$ **Solution:**

$$P(1 \leq X < 4) = p_X(1) + p_X(2) + p_X(3) = 0.25 + .2 + 0.1 = 0.55$$

$$P(1 \leq X < 4) = \lim_{x \uparrow 4} F_X(x) - F_X(1) = 0.65 - 0.1 = 0.55$$

(iii) $P(1 \leq X \leq 4)$ **Solution:**

$$P(1 \leq X \leq 4) = p_X(1) + p_X(2) + p_X(3) + p_X(4) = 0.25 + .2 + 0.1 + 0.3 = 0.85$$

$$P(1 \leq X \leq 4) = F_X(4) - F_X(1) = 0.95 - 0.1 = 0.85$$

(iv) $P(X < 3)$ **Solution:**

$$P(X < 3) = p_X(0) + p_X(1) + p_X(2) = 0.1 + 0.25 + .2 = 0.55$$

$$P(X < 3) = \lim_{x \uparrow 3} F_X(x) = 0.55$$

(v) $P(X > 6)$ **Solution:**

$$P(X > 6) = 0$$

(since there is no probability mass for $x > 6$)

$$P(X > 6) = 1 - P(X \leq 6) = 1 - F_X(6) = 1 - 1 = 0$$

Expected Value, Mean, and Variance

Definition: Let X be a discrete random variable with range R_X and probability mass function p_X . The *expected value* of $h(X)$ is defined as

$$E[h(X)] = \sum_{x \in R_X} h(x) p_X(x).$$

Remark:

- An expected value is a weighted average.
- We say that $E[h(X)]$ is an **expected value** for the following reason: If we repeat the random experiment a large number of times, then the average of the observed values of $h(X)$ should be approximately equal to $E[h(X)]$.
- We will use the expected value to define the *mean* and *variance* of a random variable X .

Definition: Let X be a discrete random variable with range R_X and probability mass function f . Its mean (also called its expected value) is

$$\mu_X = \mu = E[X] = \sum_{x \in R_X} x f(x).$$

Its *variance* is

$$\sigma_X^2 = \sigma^2 = V(X) = E[(X - \mu)^2] = \sum_{x \in R_X} (x - \mu)^2 f(x).$$

Its *standard deviation* is

$$\sigma_X = \sigma = \sqrt{V(X)}.$$

Alternative Formula for the variance: The following is an alternative formula that can be used to compute the variance. It is more efficient compared to the definition.

$$V(X) = E[X^2] - E[X]^2$$

Remarks :

- The mean is a centre of mass of the distribution of probability masses. So the mean can be thought of as a measure of “location”.
- The variance and standard deviation can be thought of as a measure of “dispersion” or “variability”.

Example 21: Consider the random variable X in example 20. Illustrate its probability mass function with a bar plot. Next, compute the mean and standard deviation of X .

Solution:

$$E[X] = 0 \times p_X(0) + 1 \times p_X(1) + 2 \times p_X(2) + 3 \times p_X(3) + 4 \times p_X(4) + 5 \times p_X(5)$$

$$= 0 \times 0.1 + 1 \times 0.25 + 2 \times 0.2 + 3 \times 0.1 + 4 \times 0.3 + 5 \times 0.05 = 2.4$$

$$E[X^2] = 0^2 \times p_X(0) + 1^2 \times p_X(1) + 2^2 \times p_X(2) + 3^2 \times p_X(3) + 4^2 \times p_X(4) + 5^2 \times p_X(5)$$

$$= 0 \times 0.1 + 1 \times 0.25 + 4 \times 0.2 + 9 \times 0.1 + 16 \times 0.3 + 25 \times 0.05 = 8$$

so

$$V[X] = E[X^2] - (E[X])^2 = 8 - 2.4^2 = 2.24$$

$$\sigma_X = \sqrt{V[X]} = \sqrt{2.24}$$

Linear Transformations

Let Y be a linear transformation of the random variable X , i.e.,

$$Y = aX + b,$$

where a and b are real numbers. The mean and the variance of the random variable Y are, respectively,

$$E[Y] = aE[X] + b \quad \text{et} \quad V[Y] = a^2V[X].$$

Proof: We begin with the expectation:

$$\begin{aligned} E[Y] &= \sum_{x \in R_X} (ax + b) p_X(x) \\ &= a \left(\sum_{x \in R_X} x p_X(x) \right) + b \left(\sum_{x \in R_X} p_X(x) \right) \\ &= aE[X] + b, \end{aligned}$$

since $\sum_{x \in R_X} p_X(x) = 1$. Now consider the variance :

$$\begin{aligned} V[Y] &= E[(Y - \mu_Y)^2] \\ &= E[((aX + b) - (a\mu_X + b))^2] \\ &= E[a^2(X - \mu_X)^2] \\ &= a^2E[(X - \mu_X)^2] \\ &= a^2V[X] \end{aligned}$$

Example 22 : Consider the random variable X from Example 19. (Recall that X was the number of defective components)

We can calculate the mean and the standard deviation of X :

$$\mu_X = 0.3 \quad \text{and} \quad \sigma_X = 0.5$$

Suppose there is a fixed cost for the device of \$125 and that there is an additional cost of \$25 for each defective component. So the total cost is

$C = 125 + 25X$. Determine the mean and variance of the total.

Solution:

$$E[C] = E[125 + 25X] = 125 + 25E[X] = 125 + 25 \times 0.3 = 132.5$$

$$V[C] = V[125 + 25X] = 25^2V[X] = 25^2 \times \sigma_X^2 = 625 \times 0.5^2 = 156.25$$

Continuous Random Variables

Recall: The cumulative distribution function F_X is a function such that

$$F_X(x) = P(X \leq x), \quad x \in \mathbb{R}.$$

We will prove that F_X is a **non-decreasing function**. Let a and b be two real numbers such that $a < b$. We will prove that $F_X(a) \leq F_X(b)$:

Consider the events $A = \{X \leq a\}$ and $B = \{X \leq b\}$. If $s \in A$, then we have $X(s) \leq a$. But this implies that $X(s) < b$, since $a < b$. Therefore, $s \in B$. This implies that $A \subseteq B$ and that $P(A) \leq P(B)$. Therefore, we have $F_X(a) = P(X \leq a) = P(A) \leq P(B) = P(X \leq b) = F_X(b)$. Hence, $F_X(a) \leq F_X(b)$.

For a discrete random variable, the growth of F_X is caused by the accumulation of the probability masses. We can express F_X as the following sum :

$$F_X(x) = \sum_{y \in R_X: y \leq x} p_X(y).$$

Now suppose that we have a random variable that can take values in some interval, for example the time taken to accomplish a task. The task could take 1 minute, or 2 minutes, or any value between 1 and 2 minutes. It's not possible to calculate a sum over an interval, so we must replace the sum by an **integral**.

We will say that X is a continuous random variable if there exists a function f_X such that

$$F_X(x) = \int_{-\infty}^x f_X(t) dt.$$

The function f_X is called the **probability density** of X . By the fundamental theorem of calculus, we have

$$f_X(x) = \frac{d}{dx} F_X(x),$$

so f_X is the derivative of F_X . Thus, $f_X(x) = F'_X(x)$ is the rate at which we accumulate probability at x .

Interpretation of a probability as the area under a curve :

First of all, F_X is a non-decreasing function, so its derivative is non-negative. Therefore, $f_X(x) \geq 0$, for all $x \in \mathbb{R}$.

Consider the probability

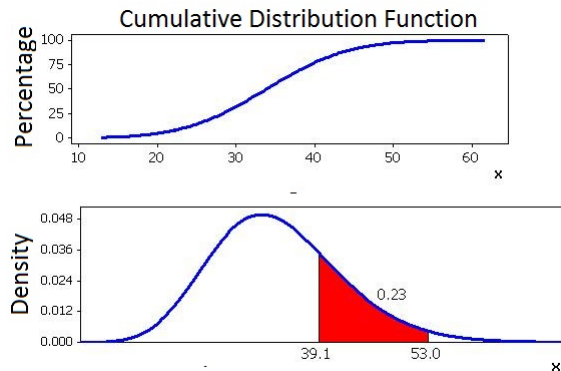
$$\begin{aligned} P(a < X \leq b) &= P(X \leq b) - P(X \leq a) = F_X(b) - F_X(a) \\ &= \int_{-\infty}^b f_X(x) dx - \int_{-\infty}^a f_X(x) dx = \int_a^b f_X(x) dx. \end{aligned}$$

So the probability that X falls in the interval $[a, b]$ is the area under the density curve f_X between $x = a$ and $x = b$.

We can also prove that $P(X = x) = 0$, for all $x \in \mathbb{R}$, when X is continuous.

So we have

$$\begin{aligned} P(a \leq X \leq b) &= P(a < X < b) = P(a \leq X < b) \\ &= P(a < X \leq b) = P(a < X \leq b) = \int_a^b f_X(x) dx. \end{aligned}$$



The above figure is an illustration of the cumulative distribution function and probability density function for a continuous random variable X . We have

$$P(39.1 < X < 53) = \int_{39.1}^{53} f_X(x) dx = 0.23$$

Example 23 : The time taken to complete a manufacturing task (in hours) can be represented by a continuous random variable X with the following density:

$$f(x) = \frac{1}{8}, \quad 0 < x < 8.$$

(a) Determine the cumulative distribution function of X .

Solution:

$$F_X(x) = \int_{-\infty}^x f(x)dx = \int_0^x \frac{1}{8}dx = \frac{1}{8}x \quad \text{if } 0 < x < 8.$$

So

$$F_X(x) = \begin{cases} 0; & x < 0 \\ \frac{1}{8}x; & 0 \leq x < 8 \\ 1; & x \geq 8 \end{cases}$$

(b) Determine the probability that the time to complete the task will be more than 2 hours but less than 7 hours.

Solution: $P(2 < X < 7) = F_X(7) - F_X(2) = \frac{7}{8} - \frac{2}{8} = \frac{5}{8}$

(c) Determine a value of x such that the task will have exactly 90% probability of being completed in more than x hours.

Solution: We want to find X such that $0.9 = P(X > x) = 1 - P(X \leq x) = 1 - F_X(x) = 1 - \frac{1}{8}x$
so $0.9 = 1 - \frac{1}{8}x$
and so $x = 0.1 \times 8 = 0.8$

Remarks:

- The **mean** of a continuous random variable is

$$\mu_X = E[X] = \int_{-\infty}^{\infty} x f_X(x) dx.$$

- The **variance** is

$$\sigma_X^2 = V[X] = E[(X - \mu_X)^2] = \int_{-\infty}^{\infty} (x - \mu_X)^2 f_X(x) dx.$$

The variance can also be calculated in the following way:

$$\sigma_X^2 = V[X] = E[X^2] - \mu_X^2 = \left[\int_{-\infty}^{\infty} x^2 f_X(x) dx \right] - \mu_X^2.$$

- The **standard deviation** is $\sigma_X = \sqrt{V[X]}$.
- **Mean and variance of a linear transformation :** Let Y be a linear transformation of the continuous random variable X , that is,

$$Y = aX + b,$$

where a and b are real numbers. The mean and the variance of Y are, respectively,

$$E[Y] = aE[X] + b \quad \text{et} \quad V[Y] = a^2 V[X].$$

Proof: The proof is the same as in the discrete case, except that we replace sums with integrals.

Example 24: Consider the random variable X of Example 23.

- (a) Determine the mean and standard deviation of X .

Solution: $E[X] = \int_{-\infty}^{\infty} xf(x)dx = \int_0^8 x \times \frac{1}{8}dx = \frac{1}{8} \times \frac{1}{2}x^2|_0^8 = 4$
 $V[X] = \int_{-\infty}^{\infty} x^2f(x)dx = \int_0^8 x^2 \times \frac{1}{8}dx = \frac{1}{8} \times \frac{1}{3}x^3|_0^8 = 21.33$
so $\sigma_X = \sqrt{V[X]} = \sqrt{21.33} = 4.62$.

Suppose that the cost (in dollars) of accomplishing the task is $C = 500 + 1500 X$. What is the expected value of the cost?

- (c) What is the standard deviation of the cost? $V[C] = V[500 + 1500 X] = 1500^2 V[X]$ so $\sigma_C = \sqrt{V[C]} = \sqrt{1500^2 V[X]} = 1500\sigma_X = 1500 \times 4.62 = 6930$

Properties of the probability density function f_X :

- (a) $f_X(x) \geq 0$ for all $x \in \mathbb{R}$ [since F_X is a non-decreasing function]
- (b) $\int_{-\infty}^{\infty} f_X(x) dx = 1$ [since the total probability mass is 1]

Example 25 : Let X be a continuous random variable with the following probability density function :

$$f(x) = c e^{-2x}, \quad x > 0$$

- (a) Determine the value of c .

Solution: c must be such that $\int_{-\infty}^{\infty} f(x) dx = 1$

so

$$1 = \int_{-\infty}^{\infty} f(x) dx = \int_0^{\infty} c e^{-2x} dx = -\frac{1}{2} c e^{-2x} \Big|_0^{\infty} = -\frac{1}{2} c \times (0 - 1)$$

so $1 = \frac{1}{2} c$ and hence $c = 2$.

- (b) Calculate $P(X \leq 1/2)$ and $P(X > 3/4)$.

Solution:

$$P(X \leq 1/2) = \int_0^{1/2} 2 e^{-2x} dx = -e^{-2x} \Big|_0^{1/2} = -[e^{-1} - 1] = 1 - e^{-1}$$

$$P(X > 3/4) = \int_{3/4}^{\infty} 2 e^{-2x} dx = -e^{-2x} \Big|_{3/4}^{\infty} = -[e^{-\infty} - e^{-6/4}] = e^{-3/2}$$