50.034 - Introduction to Probability and Statistics

Week 2 – Lecture 4

January–May Term, 2019

SINGAPORE UNIVERSITY OF TECHNOLOGY AND DESIGN

Lecture 4 Assigned Readings: 3.1, 3.2

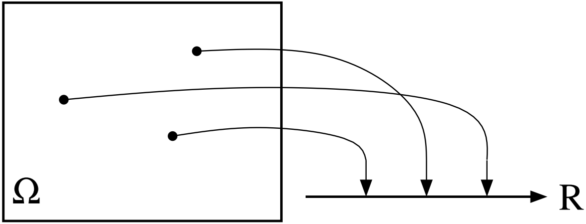
Outline

- Random variable (R.V.)
- Discrete and continuous R.V.
- Probability distribution
- Probability mass function (pmf)
- Probability density function (pdf)

Random variable (R.V.)

Definition: For a given sample space Ω of some experiment, a random variable is any rule that associates a real number with each outcome in Ω.

In other words, a random variable is a real-valued function whose domain is the sample space.



(A function is called real-valued if its range is contained in the set of real numbers.)

Discrete and continuous random variables

A random variable X is called discrete if X can take only a finite number k of different values $x_1, \dots, x_k$, or, at most, an infinite sequence of different values $x_1, x_2, x_3, \dots$

A random variable X is called continuous if the following two conditions hold:

- The set of all possible values for X is either a single interval on the real line (possibly the entire real line), or a union of disjoint intervals on the real line.
- No possible value has positive probability, that is, $\Pr(X = x) = 0$ for any possible value x.

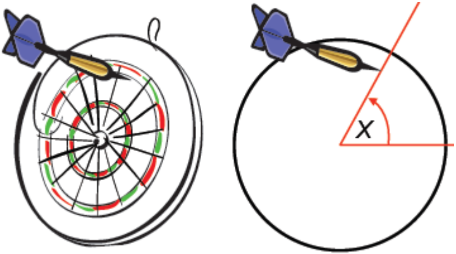
Examples of discrete random variables

A coin is tossed three times. Let X be the number of heads in the experiment. X is a **discrete random variable**; it has four possible values: 0, 1, 2 and 3.

Consider an experiment that consists of tossing a coin until the first head occurs. Let Y be the number of tosses. Y is also a **discrete random variable**; the possible values form an infinite sequence: $1, 2, 3, \ldots$.

Example of continuous random variable

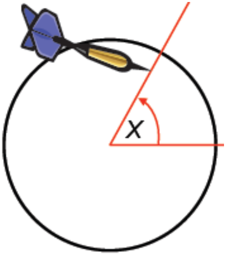
Throw a dart and record the angle X (in degrees) that the dart makes with a horizontal line in the anti-clockwise direction.



X is a **continuous random variable**; it can take on any value in the interval $[0, 360)$.

Question

What is the probability that $X = 30$?
(i.e. what is the probability that the dart makes an angle of 30°?)



Answer: It depends on how we throw. Suppose we throw the dart randomly; then there are uncountably many values that X can take. The probability that X is exactly 30 is zero.

Question

Let X be a random variable that take values from $[0, 1] \bigcup \{2\}$. The distribution of X is described as follows: it can be equally likely to fall into any point in [0,1], and

$$\Pr(X = 2) = 0.5$$

Which of the following statement is correct?
(A) X is a discrete random variable.
(B) X is a continuous random variable.
(C) X is neither discrete nor continuous.

Answer: (C).

Probability distribution of R.V.

Let X be a random variable defined on the sample space Ω.

- Thus, $X(\omega)$ is a real number for every outcome $\omega \in \Omega$.

Let C be a subset of the real line. We write $\{X \in C\}$ to mean the set $\{\omega \in \Omega : X(\omega) \in C\}$ of all outcomes whose X-value is in C.

- Recall: An event is a subset of outcomes contained in Ω.
- So $\{X \in C\}$ is an event.
- Every event is assigned a probability, so it makes sense to consider the probability of $\{X \in C\}$:

$$\Pr(X \in C) = \Pr(\{\omega \in \Omega : X(\omega) \in C\}).$$

Definition: The probability distribution of X is the collection of all probabilities of the form $\Pr(X \in C)$, for all sets C of real numbers. (This definition is for *any* R.V., not just discrete or continuous R.V.'s.)

More on probability distributions

Interpretation: For any set C of real numbers, the probability distribution of X gives the probability $\Pr(X \in C)$ of how likely the random variable X takes on values in C.

Remarks on notation:

- If the context is clear, we simply say "distribution of X" to mean "probability distribution of X".
- We usually write $\Pr(X = x)$ to mean $\Pr(X \in \{x\})$.

There are other ways to represent the same information given by the probability distribution of a random variable:

- probability mass function (only for discrete R.V.)
- probability density function (only for continuous R.V.)
- cumulative distribution function (for *any* R.V.)

Probability mass function (pmf) for discrete R.V.

Let X be a **discrete** R.V. defined on the sample space Ω.

The probability mass function (pmf) of X is a function $p(x)$, defined on every real number, such that

$$p(x) = \Pr(X = x) = \Pr(\{\omega \in \Omega : X(\omega) = x\}).$$

- In other words, $p(x)$ is the probability of the event $\{X \in \{x\}\}$, i.e. the event consisting of all outcomes whose X-value is x.
- The pmf of X completely describes the distribution of X, since
$$\Pr(X \in C) = \sum_{x \in C} p(x).$$

Note on terminology: In the textbook, "probability function" is used instead of "probability mass function". Either term is okay.

- Be careful: Whether you use pmf or "probability function", it makes sense only for **discrete** random variables!

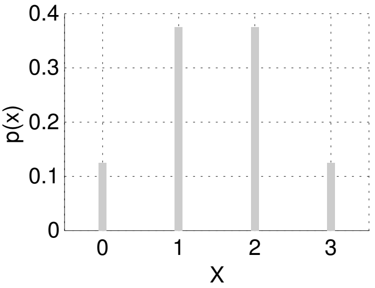
Example 1

Toss a fair coin 3 times. The sample space is

$$\Omega = \{\text{HHH, HHT, HTH, HTT, THH, THT, TTH, TTT}\}.$$

Let X be the number of heads in the experiment. What is the pmf of X? Plot it as a function of all possible values for X. **Answer:** Let $p(x)$ be the pmf of X. Then:

$$p(0) = \tfrac{1}{8},\ p(1) = \tfrac{3}{8},\ p(2) = \tfrac{3}{8},\ p(3) = \tfrac{1}{8}.$$



Bernoulli distribution

A random variable X is called a Bernoulli random variable if it takes only two values 0 and 1.

- The pmf of X is given by $\Pr(X = 1) = p$, $\Pr(X = 0) = 1 - p$ for some $0 \leq p \leq 1$.
- We say that X is the Bernoulli R.V. with parameter p.
- A Bernoulli distribution is the distribution of a Bernoulli R.V.

Since the pmf depends only on the value of p, we often write it as $p(x; p)$ rather than just $p(x)$:

$$p(x; p) = \begin{cases} 1 - p, & \text{if } x = 0; \\ p, & \text{if } x = 1; \\ 0, & \text{otherwise.} \end{cases}$$

Note: p is a parameter. Each different value of p between 0 and 1 gives a different Bernoulli distribution.

Bernoulli process and binomial R.V.

Consider an experiment with n repeated trials, such that:

- The trials are independent.
- Each trial has only two outcomes: 1 (success) and 0 (failure).
- The success rate of the trials is the same (denoted by some probability p).

Such an experiment is called a Bernoulli process.
Each trial of the experiment is called a Bernoulli trial.

The number of successful trials in a Bernoulli process is a discrete random variable called a binomial random variable.
A binomial distribution is the distribution of a binomial R.V.

Binomial distribution

The pmf of a binomial R.V. is given by

$$p(x;n,p) = \begin{cases} \binom{n}{x} p^x (1-p)^{n-x}, & \text{if } x \in \{0, 1, ..., n\}; \\ 0, & \text{otherwise}; \end{cases}$$

where n is the number of trials, and p is the success rate of each trial.

Recall: $\binom{n}{k}$ is a binomial coefficient, computed by $\binom{n}{k} = \frac{n!}{k!(n-k)!}$.

By the **binomial theorem** (binomial expansion of $(p + (1-p))^n$),

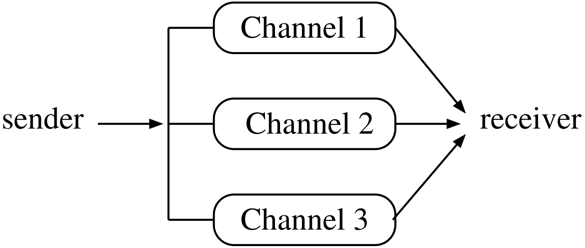
$$\sum_{x=0}^{n} p(x;n,p) = \sum_{x=0}^{n} \binom{n}{x} p^x (1-p)^{n-x} = 1.$$

The name "binomial R.V." comes from the binomial theorem.

Example 2: Communication Channels

A message is sent via three independent channels in parallel. It is received if any of the channels transmits the message successfully. The transmission success rate (probability) of each channel is 0.4. What is the probability that the message is received?

(Hint: Use the binomial distribution.)



Method 1

Consider transmission over each channel as a Bernoulli trial with success rate 0.4. Then, the experiment is a Bernoulli process with three trials. Let X be the number of channels that transmits the message successfully.

X is a binomial R.V., whose pmf is $p(x; n, p)$, where $n = 3$ and $p = 0.4$.

$$\begin{aligned}
\Pr(\text{received}) &= \Pr(X = 1) + \Pr(X = 2) + \Pr(X = 3) \\
&= \tbinom{3}{1} 0.4^1 (1 - 0.4)^2 + \tbinom{3}{2} 0.4^2 (1 - 0.4)^1 \\
&\quad + \tbinom{3}{3} 0.4^3 (1 - 0.4)^0 \\
&= 0.784
\end{aligned}$$

Method 2

Consider transmission over each channel as a Bernoulli trial with success rate 0.4. Then, the experiment is a Bernoulli process with three trials. Let X be the number of channels that transmits the message successfully.

X is a binomial R.V., whose pmf is $p(x; n, p)$, where $n = 3$ and $p = 0.4$.

$$\begin{aligned}
\Pr(\text{received}) &= 1 - \Pr(\text{all channels fail}) \\
&= 1 - \Pr(X = 0) \\
&= 1 - \tbinom{3}{0} 0.4^0 (1 - 0.4)^{3-0} \\
&= 1 - 0.6^3 \\
&= 0.784
\end{aligned}$$

Continuous random variable

Recall: A random variable X is continuous if:

- The set of all possible values for X is either a single interval on the real line (possibly the entire real line), or a union of disjoint intervals on the real line.
- No possible value has positive probability, that is, $\Pr(X = x) = 0$ for any possible value x.

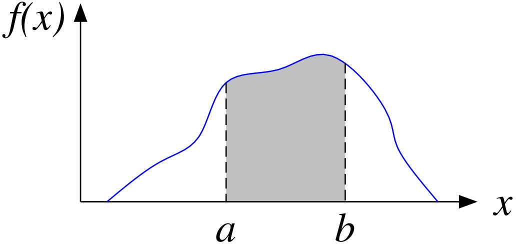
The probability density function (pdf) of a continuous R.V. X is a function $f(x)$, such that for any two numbers a and b with $a \leq b$,

$$P(a \leq X \leq b) = \int_a^b f(x)\,dx$$

Note: The pdf of X completely describes the distribution of X.

Probability distribution of R.V.

The probability that X takes on a value in the interval $[a, b]$ is the area above this interval and under the graph of the pdf.



The graph of $f(x)$ is referred to as the density curve.

Probability distribution of R.V.

For a function $f(x)$ to be a valid pdf of some continuous R.V., it must satisfy the following two conditions:

- $f(x) \geq 0$ for all x. (Density cannot be negative.)
- $\int_{-\infty}^{\infty} f(x)\,dx = 1$. (Area under the curve of $f(x)$ is 1.)

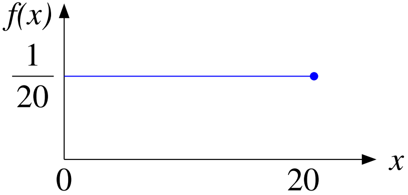
Boundary issue: As the probability of any value is 0, the following expressions have the same probability.

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

Example 3

Let X be the time (in minutes) that a person has to wait until the next bus arrives. Suppose this continuous R.V. X has the following pdf:

$$f(x) = \begin{cases} \frac{1}{20}, & \text{if } 0 \le x \le 20; \\ 0, & \text{otherwise.} \end{cases}$$



Verify that $\int_{-\infty}^{\infty} f(x)\, dx = 1$, and find the probability that a person has to wait between 5 and 15 minutes.

Example 3

We can verify that

$$\int_{-\infty}^{\infty} f(x)\,dx = \int_{0}^{20} \frac{1}{20}\,dx = 1$$

The probability that a person has to wait between 5 and 15 minutes is

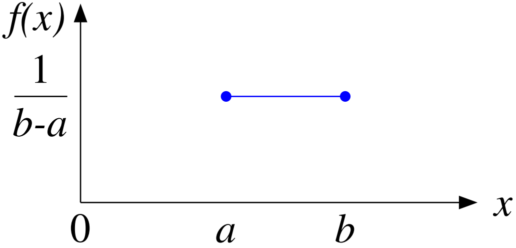
$$\Pr(5 \le X \le 15) = \int_{5}^{15} \frac{1}{20}\,dx = 0.5$$

Uniform distribution

A continuous R.V. X is said to have a uniform distribution on the interval $[a, b]$ if the pdf of X is

$$f(x; a, b) = \begin{cases} \dfrac{1}{b-a}, & \text{if } a \leq x \leq b; \\ 0, & \text{otherwise.} \end{cases}$$

Note that a and b are parameters of the pdf.



Question

Let X be a continuous R.V. that has a uniform distribution, such that its pdf is

$$f(x) = \begin{cases} \frac{1}{c}, & \text{if } 0 \le x \le 10; \\ 0, & \text{otherwise.} \end{cases}$$

What is the value of c?

Solution: We know that $\int_{-\infty}^{\infty} f(x)\,dx$ should equal 1, i.e.

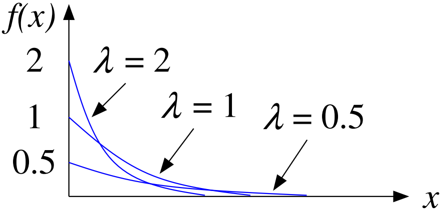
$$\int_0^{10} \frac{1}{c}\,dx = \frac{10}{c} = 1.$$

Therefore, $c = 10$.

Exponential distribution

A continuous R.V. X is said to have an exponential distribution with parameter λ ($\lambda > 0$) if the pdf of X is

$$f(x;\lambda) = \begin{cases} \lambda e^{-\lambda x}, & \text{if } x \geq 0; \\ 0, & \text{otherwise.} \end{cases}$$



(More on exponential distributions in Lecture 6..)

Example 4

Let X_1, X_2 be random variables.
Suppose X_1 has the exponential distribution with parameter λ_1.
Suppose X_2 has the exponential distribution with parameter λ_2.
Assume that $\lambda_1 > \lambda_2$. Then for any number $a > 0$, which of the following statements is true? Show your reasoning.

(A) $\Pr(X_1 \le a) < \Pr(X_2 \le a)$,

(B) $\Pr(X_1 \le a) = \Pr(X_2 \le a)$,

(C) $\Pr(X_1 \le a) > \Pr(X_2 \le a)$,

Answer: (C). Note: $\Pr(X_i \le a) = \int_0^a \lambda_i e^{-\lambda_i x}\, dx = 1 - e^{-\lambda_i a}$.
Since $1 - e^{-\lambda a}$ (as a function of λ) increases with λ, we infer:

$$\Pr(X_1 \le a) = 1 - e^{-\lambda_1 a} > 1 - e^{-\lambda_2 a} = \Pr(X_2 \le a).$$

Summary

- Random variable (R.V.)
- Discrete and continuous R.V.
- Probability distribution
- Probability mass function (pmf)
- Probability density function (pdf)

Reminder:
There is **mini-quiz 1** (15mins) next week during Cohort Class.

- Tested on materials from Lecture 1 up to and including Slide 7 ("Mean and variance of binomial R.V.") of Lecture 6.