STAT241/251 Lecture Notes
Chapter 4 Part A

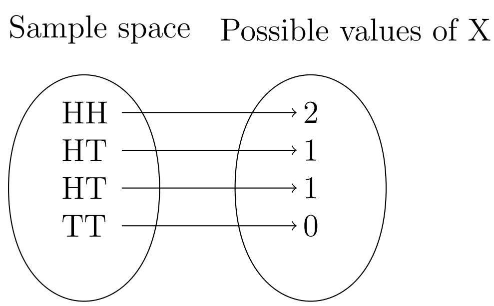
Yew-Wei Lim

Ch4

Random Variables

Random variables are customarily denoted by uppercase letter, such as X and Y. But what is a random variable? A random variable X is a function (rule) that associates a number with each outcome in the sample space.

Example: Toss a coin twice and X is the random variable that represents the number of heads.



RANDOM VARIABLES

Discrete

Set of all possible values is finite or countable. e.g. Counting number of heads in 5 tosses of a coin.

Another example:

X	0	1	2
probability*	0.1	0.4	0.5

$\star$:
Note that probabilities must add up to 1 :

$$\sum P(X = x) = 1$$

We learn more about discrete random variables in Ch 6.

Continuous

Set of outcomes are continuous (intervals).

.eg.:

$$f(x) = \begin{cases} \frac{1}{9}(4 - x^2), & -1 \leq x \leq 2, \\ 0 & \text{otherwise} \end{cases}$$

$f(x)$ is called a pdf (probability density function).

Ch4 deals mostly with continuous random variables. We'll learn the following:

(a) How to calculate probability from $f(x)$

(b) Compute cumulative density function $F(x)$

(c) Find median, Q_1, Q_3, IQR

(d) find mean (expected value)

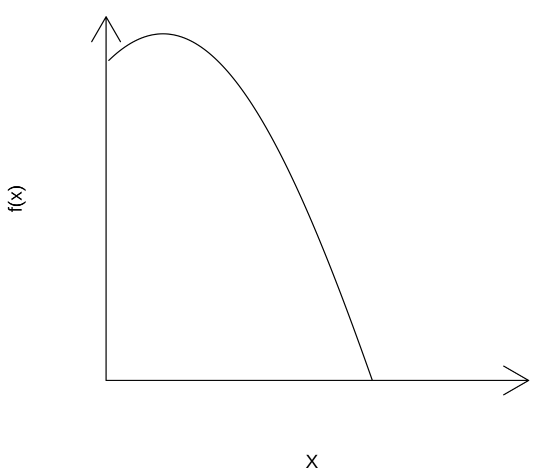
(e) find variance and standard deviation

In addition, in the last section of Ch4, we will learn

(1) how to calculate the mean and variance of the sum/difference of several independent random variable

(2) how to answer questions regarding minimum and maximum of a sequence of several independent random variables.

This is an example of the graph of a pdf.



Rules:

(a) $f(x) \geq 0$ for all x

(b) Area under curve must be 1

That is,

$$\int_{-\infty}^{\infty} f(x)dx = 1$$

(c) how to calculate probability?

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

Explain to class $P(X = x) = \int_x^x f(x)dx = 0$ [Note that we do not interpret $P(X = x) = 0$ when X is a continuous random variable as probability]

Also

$$\begin{aligned} P(a \leq X \leq b) =& P(a < X \leq b) \\ =& P(a \leq X < b) \\ =& P(a < X < b) \end{aligned}$$

Example

$$f(x) = \begin{cases} \frac{1}{c}, & 0 \leq x < 360, \\ 0 & \text{otherwise} \end{cases}$$

(a) Find the value of c

(b) Find $P(90 \leq X \leq 180)$

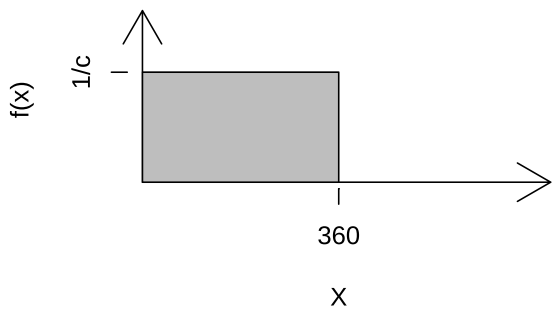
Solution:

(a) 2 ways to solve :

Method 1:

$$\begin{aligned} \int_{-\infty}^{\infty} \frac{1}{c} dx &= 1 \\ \int_{0}^{360} \frac{1}{c} dx &= 1 \\ \frac{1}{c}\Big[x\Big]_{0}^{360} &= 1 \\ \frac{360}{c} &= 1 \Rightarrow c = 360 \end{aligned}$$

Method 2:



Area under curve $= 1$, so $360 \times \frac{1}{c} = 1 \Rightarrow c = 360$

(b) Since $f(x) = \frac{1}{360}$

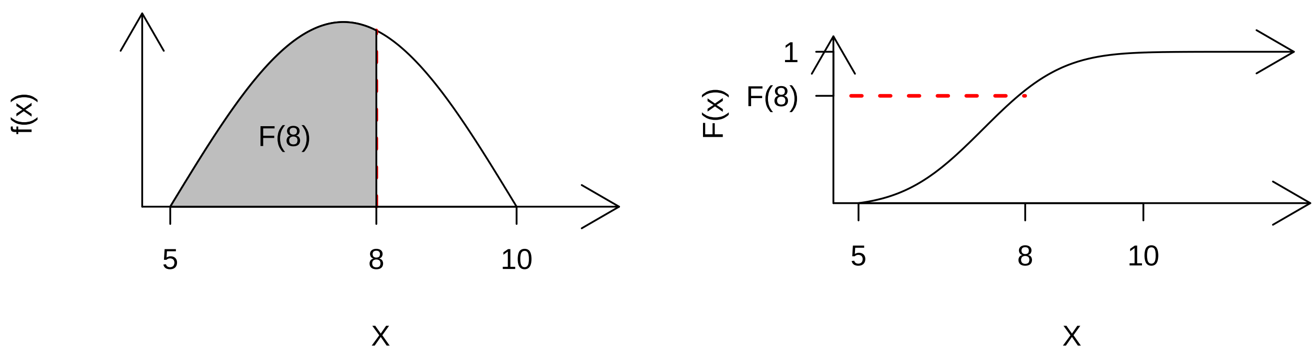
$$
\begin{aligned}
P(90 \leq X \leq 180) &= \int_{90}^{180} \frac{1}{360} dx \\
&= \frac{1}{360} \Big[x \Big]_{90}^{180} \\
&= \frac{1}{360} \times (180 - 90) \\
&= 0.25
\end{aligned}
$$

What is $F(x)$?

The cumulative distribution function (cdf) is

$$F(x) = P(X \leq x) = \int_{-\infty}^{x} f(t)dt$$

you can visualize this as:



Why learn about $F(x)$?

$F(x)$ is very useful and you can use $F(x)$ to calculate probabilities.

$$P(a \leq X \leq b) = F(b) - F(a)$$

Also

$$P(X > a) = 1 - F(a)$$

Note: learn to find $F(x)$ from $f(x)$. Also, learn to find $f(x)$ from $F(x)$.

How?

$$F'(x) = f(x).$$

See next example.

Example

$$f(x) = \begin{cases} \frac{1}{8} + \frac{3}{8}x, & 0 \le x \le 2, \\ 0 & \text{otherwise} \end{cases}$$

(a) Find $F(x)$, the cdf

(b) Use $F(x)$ to find $P(1 \le X \le 1.5)$

(c) Find $P(X > 1)$

(a)

$$\begin{aligned} F(x) &= \int_{-\infty}^{x} f(x)dx \\ &= \int_{0}^{x} (\frac{1}{8} + \frac{3}{8}y)dy \\ &= \left[\frac{y}{8} + \frac{3}{8}\frac{y^2}{2}\right]_0^x \\ &= \frac{x}{8} + \frac{3}{16}x^2 \end{aligned}$$

Hence,

$$F(x) = \begin{cases} 0, & x < 0, \\ \frac{x}{8} + \frac{3}{16}x^2, & 0 \le x \le 2, \\ 1 & x > 2 \end{cases}$$

(b)

$$\begin{aligned} P(1 \le X \le 1.5) &= F(1.5) - F(1) \\ &= (\frac{1.5}{8} + \frac{3}{16} \times 1.5^2) - (\frac{1}{8} + \frac{3}{16}) \\ &= \frac{19}{64} \\ &= 0.297 \end{aligned}$$

(c)

$$\begin{aligned} P(X > 1) =& 1 - P(X \le 1) \\ =& 1 - F(1) \\ =& 1 - (\frac{1}{8} + \frac{3}{16}) \\ =& \frac{11}{16} \\ =& 0.688 \end{aligned}$$

Extra:

From $F(x)$, we can get back $f(x)$.

$$\begin{aligned} f(x) =& F'(x) \\ =& \frac{d}{dx}(\frac{x}{8} + \frac{3}{16}x^2) \\ =& \frac{1}{8} + \frac{3}{8}x \end{aligned}$$

How to find median, Q_1, Q_3, IQR, from $f(x)$?

Steps:

To find median

(a) Find $F(x)$ first

(b) then find x such that $F(x) = 0.5.$ x is the median.

To find Q_1

(a) Find $F(x)$ first

(b) then find x such that $F(x) = 0.25.$ x is the value of Q_1.

To find Q_3

(a) Find $F(x)$ first

(b) then find x such that $F(x) = 0.75.$ x is the value of Q_3.

To find IQR

Use $IQR = Q_3 - Q_1$

Example

$$f(x) = \begin{cases} 2e^{-2x}, & x \geq 0, \\ 0 & \text{otherwise} \end{cases}$$

(a) Find median

(b) Find Q_1

(c) Find Q_3

(d) Find IQR

Solution:

(a) To find median, first find $F(x)$.

$$\begin{aligned} F(x) &= \int_0^x 2e^{-2t} dt \\ &= 2\left[-\frac{1}{2}e^{-2t}\right]_0^x \\ &= 1 - e^{-2x} \end{aligned}$$

Solve F(x)=0.5 to get median

$$\begin{aligned}
1 - e^{-2x} =& 0.5 \\
e^{-2x} =& 0.5 \\
\ln e^{-2x} =& -\ln 2 \\
-2x =& -\ln 2 \\
x =& \frac{\ln 2}{2} \\
=& 0.347
\end{aligned}$$

Thus, median = 0.347

(b) Solve F(x)=0.25 to get Q_1

$$\begin{aligned}
1 - e^{-2x} =& 0.25 \\
e^{-2x} =& 0.75 \\
-2x \ln e =& \ln 0.75 \\
x =& \frac{\ln 0.75}{-2} \\
=& 0.144
\end{aligned}$$

Thus, $Q_1 = 0.144$

(c) Solve F(x)=0.75 to get Q_3

$$\begin{aligned}
1 - e^{-2x} =& 0.75 \\
e^{-2x} =& 0.25 \\
-2x \ln e =& \ln 0.25 \\
x =& \frac{\ln 0.25}{-2} \\
=& 0.693
\end{aligned}$$

Thus, $Q_3 = 0.693$

(d) $IQR = Q_3 - Q_1 = 0.693 - 0.144 = 0.549$

How to find the mean and variance of a discrete random variable?

To find the mean $E(X)$,

$$E(X) = \sum_{x \in D} x \cdot P(x) \text{ where D is the set of possible values}$$

$$E(h(x)) = \sum_{x \in D} h(x)P(x)$$

To find variance,

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

Example:

x	4	6	8
P(x)	0.5	0.3	0.2

$$\begin{aligned}
E(X) =& (4 \times 0.5) + (6 \times 0.3) + (8 \times 0.2) \\
=& 5.4 \\
Var(X) =& E(X^2) - [\underbrace{E(X)}_{5.4}]^2 \\
E(X^2) =& 4^2 \cdot (0.5) + 6^2 \cdot (0.3) + 8^2 \cdot (0.2) \\
=& 31.6 \\
Var(X) =& E(X^2) - [E(X)]^2 \\
=& 31.6 - 5.4^2 \\
=& 2.44
\end{aligned}$$

How to find the mean and variance of a continuous random variable?

(a) Formula to get mean

$$E(X) = \int_{-\infty}^{\infty} x \cdot f(x)dx$$

In general,

$$E(g(x)) = \int_{-\infty}^{\infty} g(x) \cdot f(x)dx$$

(b) Formula for variance

$$Var(X) = E(X^2) - [E(X)]^2, \text{ where } E(X^2) = \int_{-\infty}^{\infty} x^2 \cdot f(x)dx$$

(c) Standard deviation is the square root of variance.

Example:

Find the mean and standard deviation of

$$f(x) = \begin{cases} \frac{3}{2}(1-x^2), & 0 \le x \le 1, \\ 0 & \text{otherwise} \end{cases}$$

(a)

$$\begin{aligned} E(X) &= \int_{-\infty}^{\infty} x \cdot f(x)dx \\ &= \int_{0}^{1} x \cdot \frac{3}{2}(1-x^2)dx \\ &= \frac{3}{2}\int_{0}^{1} (x - x^3)dx \\ &= \frac{3}{2}[\frac{x^2}{2} - \frac{x^4}{4}]_0^1 \\ &= \frac{3}{2}(\frac{1}{2} - \frac{1}{4}) = \frac{3}{2} \cdot \frac{1}{4} = \frac{3}{8} \end{aligned}$$

(b)

$$\begin{aligned}
Var(X) =& E(X^2) - [E(X)]^2 \\
E(X^2) =& \int_{-\infty}^{\infty} x^2 \cdot f(x)dx \\
=& \int_0^1 x^2 \cdot \frac{3}{2}(1 - x^2)dx \\
=& \frac{3}{2}\int_0^1 (x^2 - x^4)dx \\
=& \frac{3}{2}\left[\frac{x^3}{3} - \frac{x^5}{5}\right]_0^1 \\
=& \frac{3}{2}(\frac{1}{3} - \frac{1}{5}) = \frac{3}{2} \cdot \frac{2}{15} = \frac{1}{5} \\
Var(X) =& E(X^2) - [E(X)]^2 \\
=& \frac{1}{5} - (\frac{3}{8})^2 \\
=& \frac{19}{320} \\
=& 0.059
\end{aligned}$$

standard deviation of $X = \sqrt{0.059} = 0.244$.

Test your knowledge

Find $E(X)$ and $Var(X)$ of

$$f(x) = \begin{cases} 2e^{-2x}, & x \geq 0, \\ 0 & \text{otherwise} \end{cases}$$

Hint: your need to use integration by parts.

Solution: In coursenotes (page 73)