Example 2.7

An insurance company classifies people as "accident-prone" and "none-accident prone". For a fixed year, the probability of an accident prone person has an accident is 0.4, and the probability of an non-accident prone-person has an accident is 0.2.
The population is estimated to be 30 percent accident prone.

$$A = \{\text{policy holder has an accident}\}$$

$$B = \{\text{policy holder is accident prone}\}.$$

We are given that

$$P(B) = 0.3$$
$$P(A \mid B) = 0.4$$
$$P(A \mid \bar{B}) = 0.2.$$

(a) What is the probability of a new policy holder will have an accident?

By the Law of total probability:

$$P(A) = P(A|B)\,P(B) + P(A|\bar{B})\,P(\bar{B})$$

$$= 0.3 \times 0.4 + 0.2 \times 0.7$$

$$= 0.12 + 0.14 = 0.26$$

(b) Suppose the policy holder does have an accident, what is the probability that he/she was accident-prone

$$P(B|A) = \frac{P(A|B)\,P(B)}{P(A|B)\,P(B) + P(A|\bar{B})\,P(\bar{B})}$$

$$= \frac{0.4 \times 0.3}{0.4 \times 0.3 + 0.2 \times 0.7} = 0.46$$

2.6 Random variable

Random variable.

A random variable Y is a variable whose value is determined by chance; The distribution of a random variable consists of two parts. [We keep talking how to assign probs to events.

1, ~~an~~ the set of all possible values of Y (called support.

2 a function that describes how to assign ~~probability~~ probabilities to events involving Y.

Notation

We denote random variables by upper case letters towards the end of the alphabet, e.g. W, X, Y, Z.

A possible variable [realization] of Y

is denoted by the lower case version y.

In words, $P(Y=y)$ is read

" The probability of the random variable

Y equals the value y"

[$P(X=3) = 0.5$.

The probability of seeing 3

stormtroopers on your way back h~~omework~~

home is 0.5.]

<u>discrete random variable.</u>

random variable { discrete

continuous.

If ~~they~~ the random variable Y can assume only a finite (or countable) number of values.

We call Y a discrete random variable.

If it makes more sense to envision Y as taking values in an interval of numbers, we call Y a continuous random variable.

[Think about the possible values / support of Y]

Example 2.8

Classify the following random variables as discrete or ~~contino~~ continuous and specify the support of each random variable.

W = number of unbroken eggs in a randomly selected carton.

X : length of time between accidents at a factory

Y : whether or not you will pass the class

Z : number of aircraft arriving at CAE.

The random variable W is discrete.

$W : w \{ w : \cancel{=} \; w = 0, 1, \cancel{2}, \ldots, 12 \}$

$\downarrow$

realizations W.

so here is lower case letter

Thes random variable X is continuous.

$\{ x : x > 0 \}$.

[① time cannot be negative

② It can be very large

Hence, there is no point putting

an upper limit].

[Note that the different

of continous variable &

discrete variable is whether

we can think them as COUNTABLE,

both can be <u>infinite</u>]

Y: The random variable Y is

discrete It can assume values

in $\{y : y = 0, 1\}$.

where I've arbitrarily labeled

1 for passing and 0 for failing.

Remark:

Random variable that take exactly 2 values are called binary.

[fraud, not fraud]

[Republican, Democrat]

[Buy, NOT Buy] → Amazon,

[Win not Win]

The random variable Z is discrete. It can assume values in

$$\{ z : z = 0, 1, 2, \cdots \}$$

I've allowed for the probability of a very large number of aircraft coming.

Chapter 3.

3.1 Discrete distributions.

$\boxed{\text{pmf}}$ suppose that Y is a discrete random variable. The function

$$p_Y(y) = P(Y = y)$$

is called the probability mass function for Y.

$$p_Y(y) = P(Y = y)$$

The pmf $p_Y(y)$ is a function that assigns probabilities to each possible value of Y.

Properties: A pmf $p_Y(y)$ for a discrete random variable Y satisfies the following.

1. $0 < p_Y(y) < 1$ for all possible values of y

2. The sum of the probabilities, taken over all possible values of Y must equal 1, i.e.

$$\sum_{\text{all } y} p_Y(y) = 1$$

E.g. 3.1

A company has 6 telephone lines. Let Y denote the # lines used at a specific time. Suppose the pmf of Y is given by

$y =$ 0 1 2 3 4 5 6

$p_Y(y) = 0.1 \quad 0.15 \quad 0.20 \quad 0.25 \quad 0.20 \quad 0.06 \quad 0.04.$

[note that we imply $p_Y(y) = 0$,

if y is not equal to 0, 1, ... 6]

① What's the probability of exactly 2 lines are in use?

$$p_Y(2) = P(Y=2) = 0.20$$

② what is the probability that <u>at most</u> two lines are in use.

$$P(Y \le 2) = P(Y=0) + P(Y=1) + P(Y=2)$$

$$= p_Y(0) + p_Y(1) + p_Y(2)$$

$$= 0.10 + 0.15 + 0.20 = 0.45.$$

③ What is the probability ~~of~~ that at least five lines are in use?

$$P(Y \geq 5) = P(Y = 5) + P(Y = 6)$$

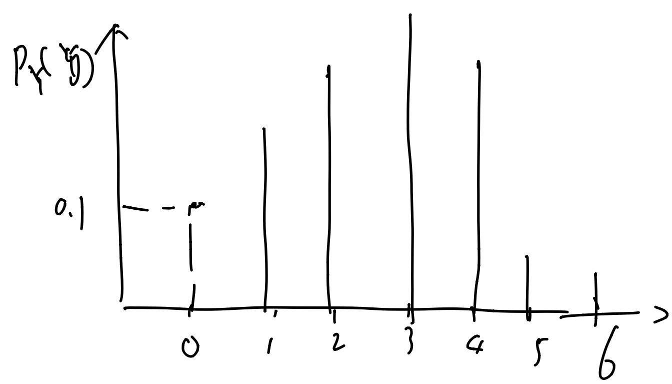
$$= p_Y(5) + p_Y(6) = 0.06 + 0.04$$

$$= 0.10$$

Visualization of this p

Display of pmf function.

The height of the bar is equal to $p_Y(y)$
$= P_Y(Y = y)$



Cumulative distribution of Y. [CDF]

$$F_Y(y) = P(Y \leq y)$$

~~$P(Y \leq y)$~~

$$P(Y \leq 2) = P(Y=2) + P(Y=0) + P(Y=1)$$

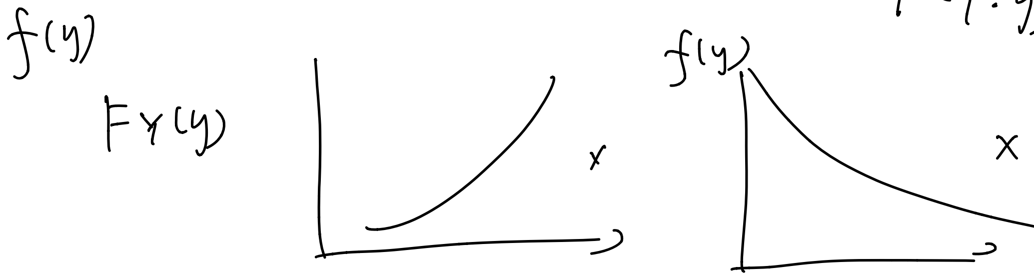
properties

① The cdf $F_Y(y)$ is a non-decreasing function

② $0 \leq F_Y(y) \leq 1$; (this thing is a probability)

③ The cdf $F_Y(y)$ is this example (when y is discrete) takes a step at each possible value of Y and stays constant otherwise.

- The height of each step at a particular y is equal to $p_Y(y) = P(Y=y)$



→ Think ~~about~~ of a shape of it?.

$f(2) = F_Y(2) = P(Y \le 2) = 0.45$

$P(Y \le 2.1) = 0.2.$

$P(Y \le \underline{\underline{6}}) =$

$p(X \le 1) = 0.2$

$p(X \le 1.2) = 0.2$

Because You cannot give birth 0.2.

$p(X \le 2) = 0.4$