Module 2: Probability

Lecture 5: Discrete Random Variables

Overview

Table 1: Supplementary Reading

Topics	**Relevant Ch**
Discrete Random Variables	OIS: 3.4
Bernoulli, Binomial Distribution	OIS: 4.3
Poisson Distribution	OIS: 4.5

RbE = R by Example

OIS = OpenIntro Statistics

Discrete Random Variables

Random Variables

- So far we have found probabilities of events $A, B, C, ...$ that correspond to a subset of outcomes in the sample space.
- As we move towards the concept of a probability distribution, we first need to understand the concept of a random variable.
- A *random variable* is a numeric quantity whose value depends on the outcome of a random event
- In essence, a random variable is a *function* that allows us to convert experimental outcomes to a numerical representation of those outcomes.

RV example

If the experiment is tossing a coin 3 times then we could have a random variable X representing the number of heads obtained. Then we ask questions like: what is $P(X = 2)$?

A = seeing exactly 2 heads

$$S = \{HHH, HHT, HTH, THH, TTH, THT, HTT, TTT\}$$

X = f

count the "H"s

3, 2, 2, 2, 1, 1, 1, 0

X possible value range 0, 1, 2, 3

Hence, we have defined a *function* over the sample space S.

Random variables

- The previous exercise presents an example of a **discrete random variable**.
- There are two types of random variables:
  - *Discrete random variables* often take only integer values
    - a variable whose value is obtained by counting
    - Example: Number of credit hours, Difference in number of credit hours this term vs last
  - *Continuous random variables* take real (decimal) values
    - a variable whose value is obtained by measuring
    - Example: Cost of books this term, Difference in cost of books this term vs last
- For now, we focus on discrete random variables.

Discrete Random variable

Definition 1 (random variable (RV))

A random variable X is a function that maps the sample space S to the real line; that is, for each element $\omega \in S$, $X(\omega) \in \mathbb{R}$.

$S = \{ \omega_1, \omega_2, \omega_3, \omega_4 \ldots \}$

- Often the outcomes of a sample space are already numbers on the real line. e.x. rolling a die counts the dots $= X$ $S = \{1, 2, 3, 4, 5, 6\}$ $\downarrow 1, 2, 3, 4, 5, 6$
- We denote random variables by capital letters (typically from the end of the alphabet)
- The real number associated with outcome ω is a realization still Random (or value) of $X(\omega)$ and is denoted by the corresponding lower case letter x. observed, no more randomness, experiment is complete.
- Herein we will use X in place of $X(\omega)$ $\cancel{X = \{ x_1, x_2, x_3, \ldots x \}}$

Discrete Probability Distribution

Definition 2

The set of all possible realizations is called the ***support*** *or* ***range*** *of* X

Definition 3

If a discrete random variable X *can take values* $x_1, x_2, \ldots, x_n$ *with probabilities* $p_1, p_2, \ldots, p_n$ *such that*

1. $p_1 + p_2 + \cdots + p_n = 1$, *and*
2. $p_i \geq 0$ *for all* i,

then this defines a ***discrete probability distribution*** *or* *probability mass function* *for* X. (pmf)

Discrete Probability Distribution

Toss two fair coins. Let X represent the number of tails obtained.

What is the discrete probability distribution for X. (pmf)

counts the tails

$S = \{ \underset{\omega_1}{HH}, \underset{\omega_2}{HT}, \underset{\omega_3}{TH}, \underset{\omega_4}{TT} \}$

$X(\omega) = \quad \underset{x_1}{0} \quad \underbrace{1 \quad 1}_{x_2} \quad \underset{x_3}{2}$

range (AKA support) of X

range

PMF version

x	x_1	x_2	...	x_n
$p(X=x)$	p_1	p_2	....	p_n

$p_1 + p_2 + \dots + p_n = 1$

x	0	1	2
$P(X=x)$	$1/4$	$2/4$	$1/4$

$\frac{1}{4} + \frac{2}{4} + \frac{1}{4} = \frac{4}{4}$

PMF

$p_1 = P(X=0) = P(HH) = 1/4$

$p_2 = P(X=1) = P(HT \cup TH) = P(HT) + P(TH)$

$= \frac{1}{4} + \frac{1}{4} = \frac{2}{4}$

$p_3 = P(X=2) = P(TT) = \frac{1}{4}$

Discrete Probability Distribution

Suppose we roll a pair of fair dice, one green and one red. If we are interested in the number of dots facing up on two dice, we can define the sample space and present it by an array;

$$S = \left\{ \begin{array}{cccccc} (1,1) & (1,2) & (1,3) & (1,4) & (1,5) & (1,6) \\ (2,1) & (2,2) & (2,3) & (2,4) & (2,5) & (2,6) \\ (3,1) & (3,2) & (3,3) & (3,4) & (3,5) & (3,6) \\ (4,1) & (4,2) & (4,3) & (4,4) & (4,5) & (4,6) \\ (5,1) & (5,2) & (5,3) & (5,4) & (5,5) & (5,6) \\ (6,1) & (6,2) & (6,3) & (6,4) & (6,5) & (6,6) \end{array} \right\}$$

Let X be a random variable to represent the sum of two dies.

The realized values of X for the entire sample space would be:

$$x = \begin{Bmatrix} 2 & 3 & 4 & 5 & 6 & 7 \\ 3 & 4 & 5 & 6 & 7 & 8 \\ 4 & 5 & 6 & 7 & 8 & 9 \\ 5 & 6 & 7 & 8 & 9 & 10 \\ 6 & 7 & 8 & 9 & 10 & 11 \\ 7 & 8 & 9 & 10 & 11 & 12 \end{Bmatrix}$$

Hence the distribution can be written

x	2	3	4	5	6	7	8	9	10	11	12
$P(X = x)$	$\frac{1}{36}$	$\frac{2}{36}$	$\frac{3}{36}$	$\frac{4}{36}$	$\frac{5}{36}$	$\frac{6}{36}$	$\frac{5}{36}$	$\frac{4}{36}$	$\frac{3}{36}$	$\frac{2}{36}$	$\frac{1}{36}$

Clicker question

Clicker 1

Does the following represent a valid probability mass function?

x	2	3	4	5	6	7
$P(X = x)$	$\frac{2}{10}$	$\frac{2}{10}$	$\frac{2}{10}$	$\frac{2}{10}$	$\frac{2}{10}$	$\frac{2}{10}$

(a) Yes

(b) No

(c) not enough information to say

Clicker question

Clicker 1

Does the following represent a valid probability mass function?

x	2	3	4	5	6	7
$P(X = x)$	$\frac{2}{10}$	$\frac{2}{10}$	$\frac{2}{10}$	$\frac{2}{10}$	$\frac{2}{10}$	$\frac{2}{10}$

$= \frac{12}{10}$ ✗

(a) Yes

(b) No

(c) not enough information to say

the probabilities don't add up to 1!

Clicker question

Clicker 2

Does the following represent a valid probability mass function?

x	-10	-8	0	6	7	10
$P(X = x)$	$\frac{2}{10}$	$\frac{1}{10}$	$\frac{2}{10}$	$\frac{2}{10}$	$\frac{2}{10}$	$\frac{1}{10}$

(a) Yes

(b) No

(c) not enough information to say

Clicker question

Clicker 2

Does the following represent a valid probability mass function?

x	-10	-8	0	6	7	10
$P(X = x)$	$\frac{2}{10}$	$\frac{1}{10}$	$\frac{2}{10}$	$\frac{2}{10}$	$\frac{2}{10}$	$\frac{1}{10}$

(a) Yes

(b) No

(c) not enough information to say

We can often represent the probability mass function (pmf) as a function in the more classical sense. Returning to the example from slide 8 we could represent the $P(X)$ by the function:

PMF version

$$\boxed{P(X=x) = \frac{6-|x-7|}{36} \quad \text{for } x = 2, 3, \ldots, 12}$$

$$P(X=4) = \frac{6-|4-7|}{36} = \frac{3}{36}$$

Probability mass function (pmf)

Definition 4 (probability mass function (pmf))

more common

If X is a discrete random variable, the function given by $f(x) = P(X = x)$ for each x within the range (or support) of X is called the probability distribution or probability mass function of X.*

pmf

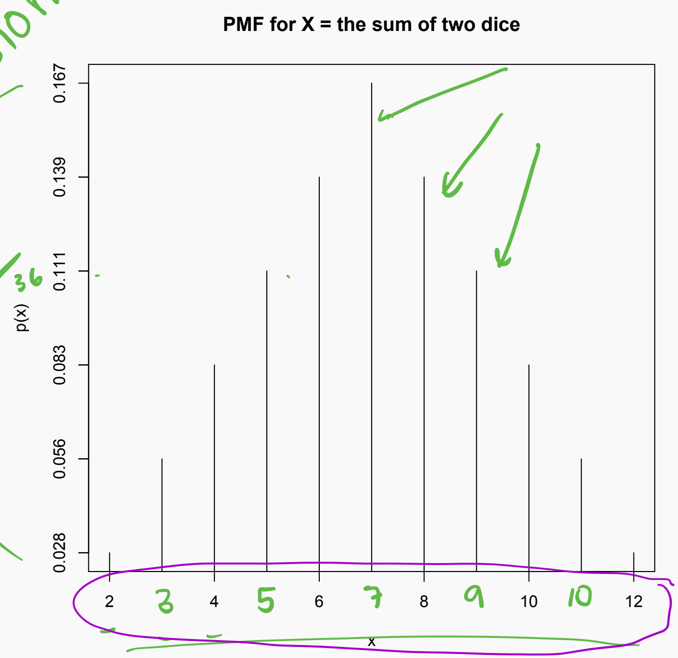
- The pmf may take the form of a table, a function, or a plot.
- The key components of all of these pmf forms is that they have:
    1. The possible values that random variable X can take on (the support/range)
    2. The probabilities, $p(x)$, for each value in the support.
    3. The $\sum p(x) = 1$

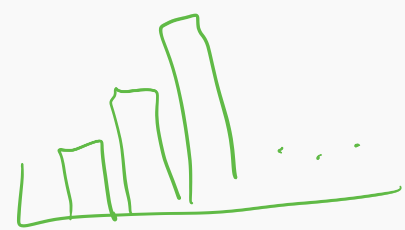
Visual representation of a pmf

- One perfectly valid way of representing a pmf is using a line graph.
- For example, the pmf in Ex. 8 could be represented in the following way.

pmf version 3.

1





Cumulative distribution function (cdf)

$x = \boxed{2, 3, 4,} \ldots 12$

- For some fixed value x, we often wish to compute the probability that the observed value of X will be at most x.
- For example, returning to Example 8, what is the probability that the sum of two dice is less than or equal to 4?

$$P(X \leq 4) = P(X=2 \cup X=3 \cup X=4)$$

$$= P(X=2) + P(X=3) + P(X=4)$$

$$= f(2) + f(3) + f(4)$$

$$= \frac{1}{36} + \frac{2}{36} + \frac{3}{36} = \frac{6}{36} = \frac{1}{6}$$

Cumulative distribution function (cdf)

Definition 5

If X is a discrete random variable, and $f(t)$ is the value of the ~~probability distribution~~ of X at t, then the function $f(x) = P(X = x)$

pmf

$$F(x) = P(X \leq x) = \sum_{t \leq x} f(t), \quad \text{for } -\infty < x < \infty$$

is the cumulative distribution function (cdf), or simply distribution function of X. In other words $F(x)$ accumulates the probabilities of $X \leq x$

Returning to Example 8

$$F(4) = P(X \leq 4) = \sum_{t \leq 4} f(t) = f(2) + f(3) + f(4) = 1/36$$

Cumulative distribution function (cdf)

Theorem 1

The values $F(x)$ of a RV X satisfy the conditions

(1) $\lim_{x\to-\infty} F(x) = 0,$ $P(X \leq x)$

(2) $\lim_{x\to\infty} F(x) = 1.$

(3) *If $a < b$, then $F(a) \leq F(b)$ for any real numbers a and b.*

Returning to Example 8

$F(12) =$ $\sum_{t \leq 12} f(t) = f(2) + f(3) + \ldots + f(12) = 1$

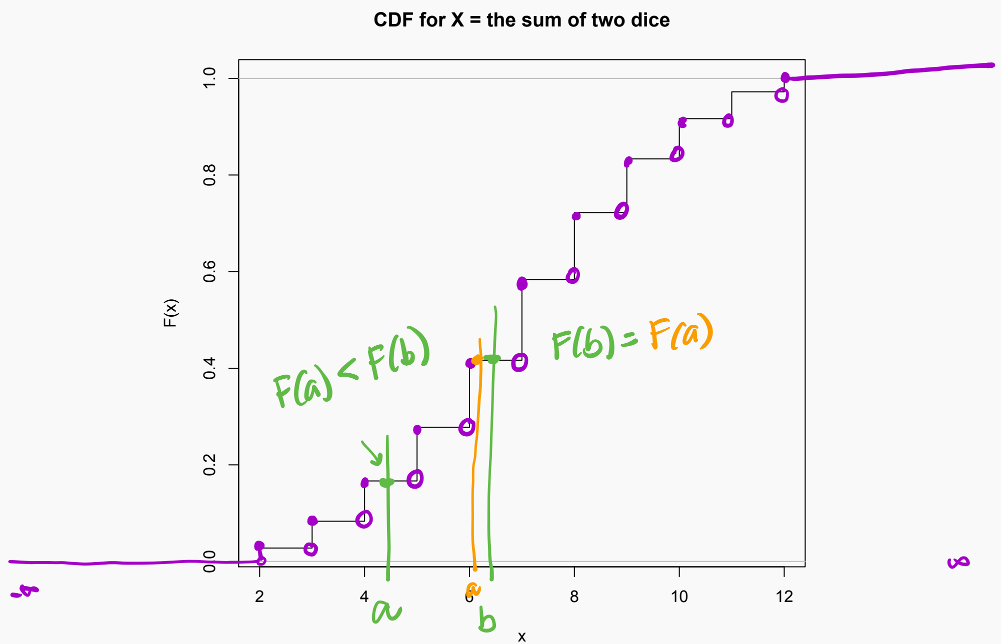
$F(14) = \sum_{t \leq 14} f(t) =$

$3 < 4 \rightarrow F(3) \leq F(4)$

Visual representation of a pmf

The CDF of a discrete RV can be represented in a **step function**.

The graph of $F(x)$ will have a jump at every possible value of X and will be flat between possible values. For example, the cdf from Ex. 8 could be represented in the following way.



CDF example

Example 1

A store carries flash drives with either 1, 2, 4, 8, or 16 GB of memory. The accompanying table gives the ~~distribution~~ probability mass function of Y = the amount of memory in a purchased drive:

instance/observed y_1

y	1	2	4	8	16
$P(Y = y)$	0.05	0.10	0.35	0.40	0.10

random variable

p_1

$=$

$P(Y = y_1)$

Find the cdf for Y.

CDF example

$F(1) = P(Y \leq 1) = P(Y=1) = 0.05$

$F(2) = P(Y \leq 2) = p_1 + p_2 = 0.05 + 0.10$

$F(4) = P(Y \leq 4) = P(Y=1) + P(Y=2) + P(Y=4)$

$= 0.05 + 0.10 + 0.35$

$F(8) = \cancel{P(Y \leq 8) =} \ldots$

$F(16) = P(Y \leq 16) = 1$

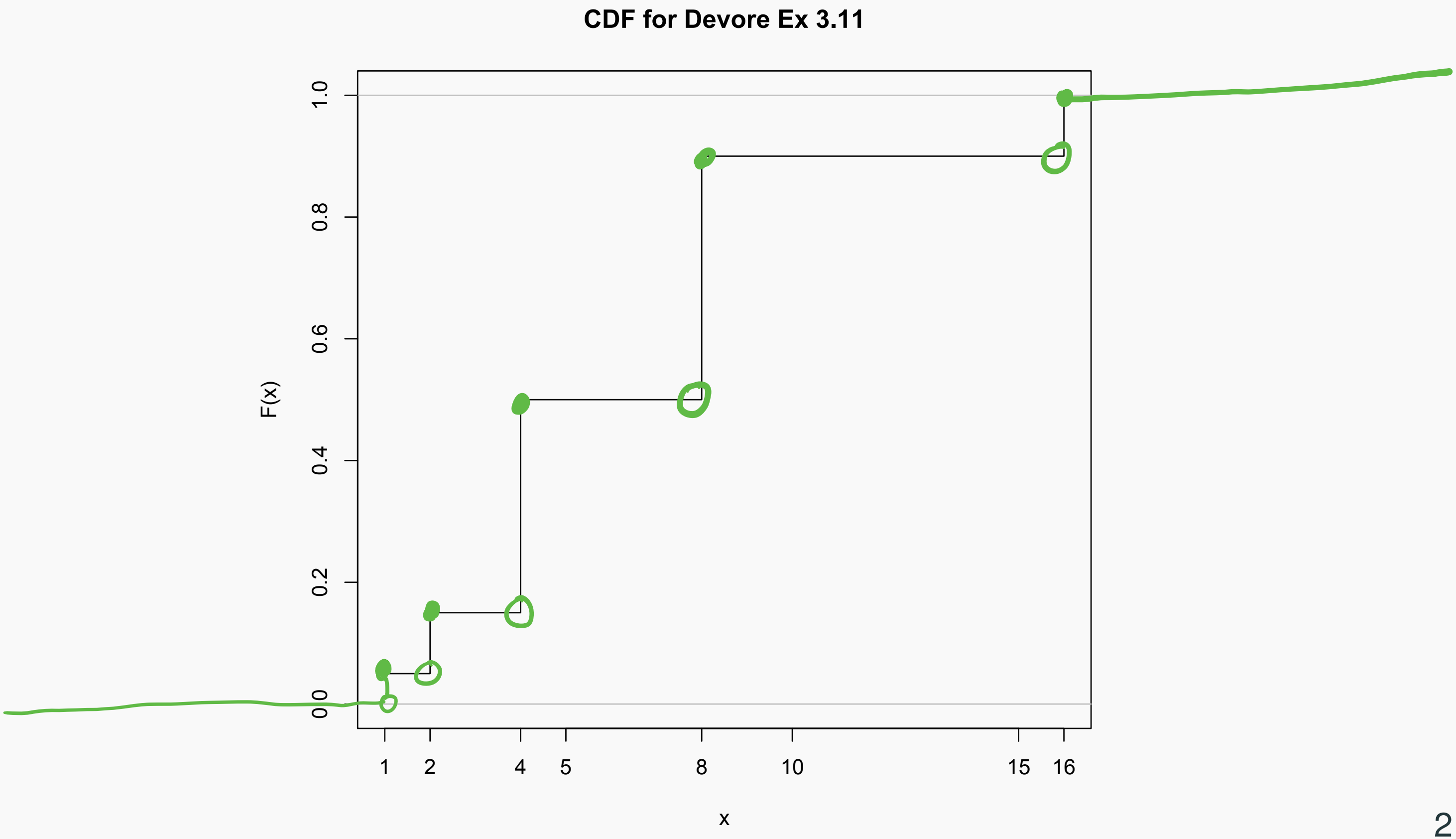
version 1

$$F(y) = \begin{cases} 0 & y < 1 \\ 0.05 & 1 \leq y < 2 \\ 0.15 & 2 \leq y < 4 \\ 0.5 & 4 \leq y < 8 \\ 0.9 & 8 \leq y < 16 \\ 1 & y \geq 16 \end{cases}$$

CDF example

A graph of this cdf is shown below

version 2



pmf from cdf

- So far our examples have derived the cdf from the pmf.
- We can reverse this process to obtain the pmf from the cdf in the event that the cdf is the only information we have available.

Theorem 2

If the range of a random variable X consists of the values $x_1 < x_2 < \cdots < x_n$, then $f(x_1) = F(x_1)$ and

$$f(x_i) = F(x_i) - F(x_{i-1}), \qquad \text{for } i = 2, 3, \ldots, n.$$

Probabilities using CDF

Returning to Example 1, find the probability that the amount of memory in a purchased drive is 2 GB.

$x_1 < x_2 < x_3 < \boxed{x_4} < x_5$

$1 < 2 < 4 < \boxed{8} < 16$

$f(2) = P(X=2) = F(2) - F(1)$

$f(x_2) = F(x_2) - F(x_1)$

$= \cancel{f(1)} + f(2) - \cancel{f(1)} = f(2)$

$= 0.15 - 0.05 = 0.10$

Probability that a purchased drive is 8 GB

$f(8) = F(8) - F(4)$

$f(x_4) = F(x_4) - F(x_3)$

$= f(8) + \cancel{f(4)} + \cancel{f(2)} + \cancel{f(1)} - [\cancel{f(4)} + \cancel{f(2)} + \cancel{f(1)}] = f(8)$

$F(Y \leq 8) \to f(Y=8)$

$- F(Y \leq 7)$

$\underbrace{}_{F(Y \leq 4)}$

Probabilities using CDF

More generally, the probability that X falls in a specified interval is easily obtained from the cdf.

$x_1 < x_2 < x_3 < x_4 < x_5$

$1 < 2 < 4 < 8 < 16$

For Example 1 the $P(2 \leq Y \leq 8) = ?$

$$= \underbrace{[f(1) + f(2) + f(4) + f(8)]}_{} - f(1)$$

$$= F(8) - F(1)$$

from CDF

$$= 0.9 - 0.05 = 0.85$$

from pmf

$0.10 + 0.35 + 0.40 = 0.85$

Probabilities using CDF

Proposition 1

For any two numbers a and b with $a \leq b$,

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

where $F(a-)$ represents the maximum of $F(x)$ values to the left of a.

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

$$\begin{array}{ccccccccc} \boxed{x_1} & < & \boxed{x_2} & < & x_3 & < & \boxed{x_4} & < & x_5 \\ 1 & < & 2 & < & 4 & < & 8 & < & 16 \\ a- & & a & & & & b & & \end{array}$$

$$P(\underset{a}{2} \leq X \leq \underset{b}{8}) = F(8) - F(1)$$

$$P(2 \leq X \leq 8) \neq F(8) - F(2) = f(8) + f(4) + \cancel{f(2)} + \cancel{f(1)} - [\cancel{f(2)} + \cancel{f(1)}]$$

$$= P(4 \leq X \leq 8)$$

$$= P(2 < X \leq 8)$$

Another view at pmfs

- It is often helpful to think of a pmf as a mathematical model for a discrete population.
- Consider selecting a random student at UBCO among the 9000 registered students. Let $X =$ the number of courses for which that student is registered. Suppose the pmf is given by:

x	1	2	3	4	5	6	7
$P(X = x)$	0.01	0.03	0.13	0.25	0.39	0.17	0.02

- If we think of the 900 individuals to each have his or her own X value; then we could view $p(x)$ as the *proportion* of the population with each X value. i.e 25% of students at UBCO are enrolled in 4 courses.

Some Special Discrete Probability Distributions

Introduction/preview

- Focusing still on discrete random variables, we will introduce some special probability distributions that are often used in many applications.
- These are so commonly used that we have given them names (Bernoulli, Poisson, binomial, etc.).
- These special distributions have well defined pmfs, indexed by constants called parameters.
- Distributions that have the same name and general form of pmf, but with different values of the parameters, are said to be in the same family

Definition 6 (Bernoulli RV)

Any random variable whose only possible values are 0 and 1 is called a Bernoulli random variable.

Definition 7 (Bernoulli trial)

A Bernoulli trial is an experiment with only two possible outcomes: success and failure. "1" "0" *If $X \sim$ Bernoulli(p), then*

- $P($*success*$) = P(X = 1) = p$, *and*
- $P($*failure*$) = P(X = 0) = 1 - p$.

- Ex 1: tossing a coin and seeing if it lands "heads" (getting a head is success and getting a tail is failure)
- Ex 2: Rolling a die and seeing if it lands on 6 (success = rolling a 6, failure = rolling a 1,2,3,4, or 5) trial.

Definition 8 (Bernoulli distribution)

A random variable X has a Bernoulli distribution, denoted by $X \sim$ Bernoulli(p), if and only if it has probability mass function: (pmf)

$$f(x) = p^x (1-p)^{1-x} \qquad \text{for } x = 0, 1$$

where $0 \leq p = P(X = 1) \leq 1$.

If $X = 1$ if a coin lands heads, and $X = 0$ if the coin lands tails, $X \sim Bernoulli(p = 0.5)$. What is the probability that the coin lands tails?

(P(X=0)

$f(0) = p^0 (1-p)^{1-0}$

$= (0.5)^0 (1-0.5)^1 = 0.5$

Parameters of a Probability Distribution

- On the 28 slide, I gave two examples of a Binomial trial:
  - $X_1 \rightarrow$ Ex 1: tossing a coin and seeing if it lands "heads"
  - $X_2 \rightarrow$ Ex 2: Rolling a die and seeing if it lands on 6 trial.
- Both of these random variables have a pmf as outlined in Definition 8, however, Ex. 1 will use $\boxed{p = 0.5}$ while Ex. 2 will use $\boxed{p = 1/6}$
- We refer to p as a parameter of a Bernoulli distribution. Moreover, this parameter is restricted to the values $0 \leq p \leq 1$
- Sometimes, you may see the pmf written as $f(x; p)$ (rather than $f(x)$) to imply that this distribution depends on the particular value of p. $f(x; p)$ $f(x)$
- If X follows a Bernoulli distribution, we usually write: $X \sim Bernoulli(p)$. X follows a Bernoulli distribution with $p = 0.5$

Definition 9 (parameter)

Suppose $f(x)$ depends on a quantity that can be assigned any one of a number of possible values, with each different value determining a different probability distribution. Such a quantity is called a <u>*parameter*</u> *of the distribution. The collection of all probability distributions for different values of the parameter is called a* <u>*family*</u> *of probability distributions.*

same family

$$\begin{cases} X_1 \sim \text{Bernoulli}(p = 0.5) \\ X_2 \sim \text{Bernoulli}(p = 1/6) \end{cases}$$

differ by parameter value.

Binomial Distribution

- The Bernoulli distribution is about the experiment with one single *trial*.
- Now suppose we conduct multiple independent Bernoulli trials and define X to count the number of successes.
- A sequence of Bernoulli trials (also referred to as a Binomial experiment) conform to the following list of requirements:
  1. the experiment consists of a sequence of n trials, where n is fixed in advance.
  2. there are only two possible outcomes for each trial (success or failure)
  3. the trials are independent (the outcome one trial does not influence the outcome of another).
  4. the probability of "success" on each trial doesn't change (constant p).

Definition 10 (Binomial Distribution)

*The **Binomial Probability Distribution** comes about as the result of n independent Bernoulli trials, each trial having success probability p. The probability of obtaining x successes in these n trials is given by*

parameters

n choose x $\quad {}^nC_x$

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

for $x = 0, 1, 2, \ldots, n$ where $0 \leq p \leq 1$ and n *is a non-negative integer.*

x=0 if all tails

x=1 1 head and 9 tails... x=n → all coins landed heads.

As the notation implies, the binomial distribution has 2 parameters: p, the probability of success, and n the number of trials. We often express this as $X \sim \text{Binom}(n, p)$

$P(X=2$ $\qquad HHTT = 0.5\times0.5\times0.5\times0.5 = (p)^2(1-p)^2$

Looking at this formula from an intuitive viewpoint, it makes sense

1. there are ${}^nC_x$ ways of getting x successes from n trials,
2. p^x is the probability of success, x times and $\qquad P(X=x)$
3. $(1-p)^{n-x}$ is the probability of failure, $n-x$ times.

$$f(x;n,p) = P(X=x) = \binom{n}{x} p^x (1-p)^{n-x}$$

$$\underbrace{SSS\ldots}_{x}\ \underbrace{FFF\ldots}_{n-x} \leftarrow \underbrace{p\cdot p\cdot\ldots p}_{x}\ \underbrace{(1-p)(1-p)\cdot\ldots(1-p)}_{n-x} = \underbrace{p^x(1-p)^{n-x}}$$

$$\frac{n!}{x!(n-x)!} := \binom{n}{x}$$

$$P(\underline{SSS\ldots SFF\ldots F}) \leftarrow P(S\cap S\cap S\ldots\cap F\ldots) = p^x(1-p)^{n-x}$$

$$+\ P(FFF\ldots FSS\ldots S) \leftarrow$$

$$+\ P(\sim\sim\sim) \leftarrow$$

A fair coin is tossed six times. What is the probability of obtaining exactly four heads? $X \sim \text{Binomial}(n=6, p=0.5)$

"Success" = H

X = counting the # of successes

$$P(X=4) = \binom{n}{x} p^x (1-p)^{n-x}$$

$$= \binom{6}{4} (0.5)^4 (1-0.5)^{6-4}$$

$$= 0.234$$

HHHHTT
HHHTHT
⋮
TTHHHH

$\binom{6}{4}$

$$P(\text{obtain 4 heads}) = P(HHHHTT) + P(HHHTHT) + \dots + P(TTHHHH)$$

$$= p^4(1-p)^2 + p^4(1-p)^2 + \dots + p^4(1-p)^2$$

$$= \binom{6}{4} p^4(1-p)^2$$

Looking at this formula from an intuitive viewpoint, it makes sense

1. there are ${}^6C_4$ ways of getting 4 successes in 6 trials,
2. $\left(\frac{1}{2}\right)^4$ is the probability of success, 4 times and
3. $\left(\frac{1}{2}\right)^2$ is the probability of failure, 2 times.

Exercises to Try at home

Exercise 1 A fair die is rolled 10 times.

a) What is the probability that it shows the number 6 exactly 5
b) What is the probability that it does not show the number 1 at all?
c) What is the probability that it shows the number 4 less than 2 times?

Exercise 2 A couple decide to keep having children until they have a boy, then they stop.

a) What is the probability that the couple have 5 children? (Hint: Geometric distribution)
b) Eight couples take this approach. What is the probability that more than 2 of these couples have 5 children?

Sampling With Replacement

- Sampling such that each unit is replaced before the next sample is drawn is called **sampling with replacement**.

- The binomial distribution will be applicable in cases where we sample with replacement (since p will not change)

- What if we don't sample with replacement, such as in a lottery or a raffle?

Sampling Without Replacement

- Sampling such that each unit is not replaced before the next sample is drawn is called **sampling without replacement**.
- The binomial distribution will not be applicable (since the p will no longer remain constant)
- Consider a situation where we are looking for defects in units.

- Suppose that there are N units in total and K of these are defective. If we sample n of these units then the probability that x are defective is given by

$$P(X = x) = \frac{{}^{K}C_{x}\,{}^{N-K}C_{n-x}}{{}^{N}C_{n}} = \frac{\binom{K}{x}\binom{N-K}{n-x}}{\binom{N}{n}}.$$

X = # defects.

for $max\{0, n - (N - K)\} \leq x \leq min\{n, K\}$

- The numerator is the particular outcome that we 'want' and the denominator is the total number of possible outcomes.

(A) $N=10$, $K=7$ defective, ($N-K=3$ not defective) $n=4$

$max\{0, 4-(10-7)\} \leq x \leq min(4,7)$

$1 \leq x \leq 4$

(B) $N=10$ $K=3$ defective ($N-K=7$ not defective) $n=4$

$max\{0, 4-(10-3)\} \leq x \leq min\{4,3\}$

$0 \leq x \leq 3$

Hypergeometric Distribution

Definition 11 N

A random variable X has a *hypergeometric distribution*, *denoted by $X \sim$ hypergeometric$\underbrace{(n, N, K)}_{\text{parameters}}$, if and only if its probability distribution given by:*

pmf

$$f(x; N, K, n) = \frac{{}^{K}C_{x}\ {}^{N-K}C_{n-x}}{{}^{N}C_{n}} = \frac{\binom{K}{x}\binom{N-K}{n-x}}{\binom{N}{n}}.$$

for $max\{0, n-(N-K)\} \leq x \leq min\{n, K\}$, where $N \in \{0, 1, 2, \ldots\}$, $K \in \{0, 1, 2, \ldots, N\}$, and $n \in \{0, 1, 2, \ldots, N\}$

This distribution has three parameters: N the number of objects, K the number of "special" objects, n, the number of draws.

Lottery Example

A lottery involves drawing 6 numbers from 1–49, without replacement. Each player chooses 6 numbers and prizes are awarded for matching 4, 5 or 6 of the numbers that are drawn. What is the probability of matching 4 numbers?

X = count the # of winning numbers we have on our ticket.

$X \sim \text{hypergeometric}(\underline{n=6}, N=49, K=6)$

$$P(X=\overset{x}{4}) = \frac{\binom{K}{x}\binom{N-K}{n-x}}{\binom{N}{n}} = \frac{\binom{6}{4}\binom{49-6}{6-4}}{\binom{49}{6}} = 0.00097$$

Exercises II

Exercise 3 An athlete conceals two performance enhancing tablets in a bottle containing eight vitamin pills that look similar. If a drug surveillance scheme is in operation that involves randomly sampling three of these pills, what is the probability that cheating will be detected? Assume that the analysis of the tablets is not error prone.

Exercise 4 Eight cards are chosen at random from a well-shuffled pack. What is the probability of obtaining three spades, two hearts, two diamonds and one club?

The Poisson Distribution

The Poisson distribution is often used to model counts. Usually, the observation process is considered to be taking place over continuous intervals of time or space. For example:

- the number of accidents in an intersection during a given time interval;
- the number of customers that arrive to a service line during a given time period;
- the number of the incoming calls per hour of a given phone line in day time;
- the number of whales in a circle of radius r.
- the number of fish caught in an hour of ice fishing.

The Poisson Distribution

X counts the number of events occurring in a fixed interval, when events occur randomly and independently with constant rate.

Definition 12

Let $\boxed{X}$ be a random variable that follows a ***Poisson distribution*** *with parameter λ. Then,*

$$P(X = x) = f(x; \lambda) = e^{-\lambda} \frac{\lambda^x}{x!}, \qquad x = 0, 1, 2, \ldots$$

X counts the number of events occurring in a fixed interval, when events occur randomly and independently with constant rate.

Definition 12

Let X be a random variable that follows a ***Poisson distribution*** *with parameter λ. Then,*

$$P(X=x) = \boxed{f(x;\lambda) = e^{-\lambda}\frac{\lambda^x}{x!},} \quad x = 0, 1, 2, \ldots$$

where $\lambda > 0$.

$$P(X=x) = \frac{e^{-\lambda}\lambda^x}{x!}$$

$$\sum_{x=0}^{5} \frac{e^{-\lambda}\lambda^x}{x!}$$

This distribution has a single parameter, λ.

$$P(X=0) + P(X=1) + \ldots + P(X=5)$$

$$P(X \le 5) = e^{-7.5} \cdot 7.5^0/0! + e^{-7.5}\, 7.5^1/1! + \ldots + e^{-7.5}\, 7.5^5/5! =$$

The Poisson Distribution: some comments

- λ is the average rate, or the number of events per time period/space interval (it is often called the rate parameter).
- As we will see in an upcoming example, it is imperative that we make sure that the average rate is consistent with the time interval in the question!
  - Generally, if r is the average number that occurs in a single unit time, and X counts the number of events in the interval $[0, t]$ then $X \sim \text{Possion}(\lambda = rt)$
- A *Poisson Process* is something that conforms to the following:
  1. Events are independent of each other.
  2. λ is constant.
  3. Two events cannot occur at the same time.

Poisson Example

A certain kind of sheet metal has, on average, five defects per 10 square feet. What is the probability that a 15-square-foot sheet of the metal will have at least 6 defects?

X = count the # of defects in $15ft^2$

$X \sim \text{Poisson}(\lambda = 7.5/15ft^2)$

$P(A) = 1 - P(A^c)$

$$\frac{5}{10ft^2} = \frac{\lambda}{15ft^2} \Rightarrow \lambda = \frac{15(5)}{10} = 7.5/15ft^2$$

$x = 0, 1, 2, 3, 4, 5, 6, 7$

① $P(X \geq 6) = 1 - P(X < 6)$

$= 1 - \underbrace{P(X \leq 5)}_{\text{cdf}}$ ppois(5, 7.5)

CDF $P(X \leq x)$

②

③ $= 1 - [P(X=0) + P(X=1) + \dots + \underbrace{P(X=5)}_{\text{pmf}}]$

④ $= 1 - \left[\sum_{x=0}^{5} \frac{7.5^x e^{-7.5}}{x!}\right] = 1 - \left[\frac{7.5^0 e^{-7.5}}{0!} + \frac{7.5^1 e^{-7.5}}{1!} + \dots \frac{7.5^5 e^{-7.5}}{5!}\right]$

$= 1 - e^{-7.5}\left(1 + 7.5 + \dots + \frac{7.5^5}{5!}\right)$

$= 0.7585$

Poison Example II

Customers arrive at an average rate of 3.7/hour. If the store opens at 8:00 what is the probability that there are at least two arrivals by 8:45?

X = counting the # of customers arriving in a $\boxed{\text{45min interval}}$

$X \sim \text{Poisson}(\lambda = 2.775)$

$$\frac{3.7}{60\text{min}} = \frac{\lambda^?}{45\text{min}} \Rightarrow \lambda = \frac{3.7(45)}{60} = 2.775/45\text{min}$$

$x = \boxed{0, 1}, 2, 3, 4, \ldots$

$$P(X \geq 2) = 1 - P(X < 2)$$
$$= 1 - P(X \leq 1)$$
$$= 1 - [P(X=0) + P(X=1)]$$
$$= 1 - \left[\frac{e^{-2.775}\, 2.775^0}{0!} + \frac{e^{-2.775}\, 2.775^1}{1!}\right]$$
$$= 1 - e^{-2.775}(1 + 2.775) \approx 0.7646$$

Poison Example II

In reference to Example 47, what is the probability that their second customer arrives between 8:20 and 8:45?

I'll put up solution as Bonus in Appendix

(more complicated than I'll ask you in test scenarios).

Exercises III

Exercise 5 Coliform bacteria are randomly distributed in river water at an average concentration of 1 per 25cc of water. What is the probability of finding more than two bacteria in a sample of 10cc of river water?

Exercise 6 Each morning, after opening her e-mail account, Lucy has to discard, on average, ten spam messages. If the number of spam messages may be described by a Poisson distribution, what is the probability that on any given morning Lucy will receive less than four spam messages?