

# Lecture 6

## Bernoulli and Binomial Distributions

---

STAT 330 - Iowa State University

1 / 18

## Discrete Distributions

### Common distributions for discrete random variables

- Bernoulli distribution

$$X \sim \text{Bern}(p)$$

- Binomial distribution

$$X \sim \text{Bin}(n, p)$$

- Geometric distribution

$$X \sim \text{Geo}(p)$$

- Poisson distribution

$$X \sim \text{Pois}(\lambda)$$

We will also discuss *joint distributions* for 2 or more discrete random variables

2 / 18

## Bernoulli Distribution

---

### Bernoulli Distribution

*Bernoulli Experiment:* Random experiment with only 2 outcomes:

- Success (S)
- Failure (F)

where  $P(\text{Success}) = P(S) = p$  for  $p \in [0, 1]$

Example 1: (Bernoulli experiments):

1. Flip a coin:  $S = \text{heads}$ ,  $F = \text{tails}$
2. Watch stock prices:  $S = \text{increase}$ ,  $F = \text{decrease}$
3. Cancer screening:  $S = \text{cancer}$ ,  $F = \text{no cancer}$

## Working with Bernoulli Random Variable

Suppose we have a Bernoulli experiment (only 2 outcomes: “success”, “failure”).

We obtain the outcome “success” with probability  $p$

When random variable  $X$  follows a *Bernoulli Distribution*, we write

$$X \sim \text{Bern}(p)$$

- Define a random variable  $X$

$$X = \begin{cases} 1 & \text{Success (S)} \\ 0 & \text{Failure (F)} \end{cases}$$

4 / 18

## Bernoulli Random Variable Cont.

- Probability Mass Function (pmf)

1.  $\text{Im}(X) = \{0, 1\}$
2.  $P(X = 1) = P(S) = p$   
 $P(X = 0) = P(F) = 1 - p$

The pmf can be written in tabular form:

$x$	0	1
$p_X(x)$	$1 - p$	$p$

The pmf can be written as a function:

$$p_X(x) = \begin{cases} p^x(1-p)^{1-x} & x \in \{0, 1\} \\ 0 & \text{otherwise} \end{cases}$$

Typically, we use the above functional form to describe the *probability mass function (pmf)* of Bernoulli random variable.

5 / 18

## Bernoulli Random Variable Cont.

- Cumulative distribution function (cdf)

$$F_X(t) = P(X \leq t) = \begin{cases} 0 & t < 0 \\ 1 - p & 0 \leq t < 1 \\ 1 & t \geq 1 \end{cases}$$

- Expected Value:  $E(X) = p$

$$E(X) = \sum_{x \in \{0,1\}} xP(X = x) = 0(1 - p) + 1(p) = p$$

- Variance:  $Var(X) = p(1 - p)$

6 / 18

## Binomial Distribution

---

## Binomial Distribution

Set up: Conduct multiple trials of *identical* and *independent* Bernoulli experiments

- Each trial is independent of the other trials
- $P(\text{Success}) = p$  for each trial

We are interested in the number of success after  $n$  trials. The random variable  $X$  is

$$X = \text{"\# of successes in } n \text{ trials"}$$

This random variable  $X$  follows a *Binomial Distribution*

$$X \sim \text{Bin}(n, p)$$

where  $n$  is the number of trials, and  $p$  is the probability of success for each trial.

7 / 18

## Binomial Distribution Cont.

Example 2: Flip a coin 10 times, and record the number of heads.

Success = "heads";  $P(\text{Success}) = p = 0.5$

- Define the random variable  $X$

$$X = \text{"\# of heads in } n = 10 \text{ trials"}$$

- The distribution of  $X$  is ...

$$X \sim \text{Bin}(10, 0.5)$$

8 / 18

## Derivation of Binomial pmf

- Probability Mass Function (pmf)

1.  $Im(X) = \{0, 1, 2, 3, 4, \dots, n\}$
2.  $P(X = x) = ?$

Recall  $P(\text{Success}) = P(S) = p$ ,  $P(\text{Failure}) = P(F) = 1 - p$

Case:  $X = 0$      $\underline{F} \ \underline{F} \ \underline{F} \ \dots \ \underline{F}$

$$P(X = 0) = (1 - p)^n$$

Case:  $X = 1$

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

Case:  $X = 2$

$$P(X = 2) = \binom{n}{2} p^2 (1 - p)^{n-2}$$

9 / 18

## Binomial Random Variables

In general, the *probability mass function (pmf)* of a Binomial R.V can be written as:

$$p_X(x) = \begin{cases} \binom{n}{x} p^x (1 - p)^{n-x} & x = 0, 1, 2, \dots, n \\ 0 & \text{otherwise} \end{cases}$$

- Cumulative distribution function (cdf)

$$F_X(t) = P(X \leq t) = \sum_{x=0}^{\lfloor t \rfloor} \binom{n}{x} p^x (1 - p)^{n-x}$$

(Add up the pmfs to obtain the cdf)

- Expected Value:  $E(X) = np$
- Variance:  $Var(X) = np(1 - p)$

10 / 18

## IID Random Variables

---

### Properties of IID Random Variables

*Independent and identically distributed (iid)* random variables have properties that simplify calculations

Suppose  $Y_1, \dots, Y_n$  are iid random variables

- Since they are *identically* distributed,
  - $E(Y_1) = E(Y_2) = \dots = E(Y_n)$   
 $\rightarrow E(\sum Y_i) = \sum E(Y_i) = nE(Y_1)$
  - $Var(Y_1) = Var(Y_2) = \dots = Var(Y_n)$
- Since they are also *independent*,
  - $\rightarrow Var(\sum Y_i) = \sum Var(Y_i) = nVar(Y_1)$

## Working with IID Random Variables

A Binomial random variable,  $X$ , is the sum of  $n$  *independent and identically distributed (iid)* Bernoulli random variables,  $Y_i$ .

Let  $Y_i$  be a sequence of iid Bernoulli R.V. For  $i = 1, \dots, n$ ,

$$Y_i \stackrel{iid}{\sim} \text{Bern}(p)$$

with  $E(Y_i) = p$  and  $\text{Var}(Y_i) = p(1 - p)$ . Then,

$$X = \sum_{i=1}^n Y_i \sim \text{Bin}(n, p)$$

Then, we obtain  $E(X)$  and  $\text{Var}(X)$  using properties of iid R.V.s

$$E(X) = nE(Y_1) = np$$

$$\text{Var}(X) = n\text{Var}(Y_1) = np(1 - p)$$

12 / 18

## Examples

---



## Binomial Distribution Examples

Example 3: A box contains 15 components that each have a defective rate of 5%. What is the probability that ...

1. exactly 2 out of 15 components are defective?
2. at most 2 components are defective?
3. more than 3 components are defective?
4. more than 1 but less than 4 components are defective?

### How to approach solving these types of problems?

1. Define the random variable
2. Determine the R.V's distribution (and values for the parameters)
3. Use appropriate pmf/cdf/ $E(X)$ / $Var(X)$  formulas to solve

13 / 18

## Binomial Distribution Examples Cont.

Define the R.V:  $X = \#$  defective out of  $n = 15$  components

State the Distribution of X:  $X \sim Bin(15, 0.05)$

$n = 15, p = 0.05$

1. What is the probability that exactly 2 out of 15 components are defective?  
 $P(X = 2) =$

14 / 18

## Binomial Distribution Examples Cont.

2. What is the probability that at most 2 components are defective?

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

15 / 18

## How to Use Binomial CDF Table (Appendix A)

Suppose we have random variable  $X \sim \text{Bin}(n = 15, p = 0.05)$ .

$$P(X \leq 2) = ?$$

- Find the  $n = 15$  sub-table
- $P(X \leq 2)$  is found inside the table corresponding to  $p = 0.05$  (column) and  $x = 2$  (row).

P(X ≤ 2) = 0.9637998					
n=15	p=0.01	0.05	0.1	0.15	1/6
x=0	0.8600584	0.4632912	0.2058911	0.08735422	0.06490547
1	0.9903702	0.8290475	0.5490430	0.31858598	0.25962189
2	0.9995842	0.9637998	0.8159389	0.60422520	0.53222487
3	0.9999875	0.9945327	0.9444444	0.82265520	0.76848078
4	0.9999997	0.9993853	0.9872795	0.93829461	0.91023433
5	1.0000000	0.9999472	0.9977503	0.98318991	0.97260589
6	1.0000000	0.9999965	0.9996894	0.99639441	0.99339642
7	1.0000000	0.9999999	0.9999999	0.9999999	0.9999999

16 / 18

## Binomial Distribution Examples Cont.

3. What is the probability that more than 3 components are defective?

17 / 18

## Binomial Distribution Examples Cont.

4. What is the probability that more than 1 but less than 4 components are defective?

18 / 18