2.4 The Binomial Distribution

Bernoulli Experiment

1 0

A **Bernoulli experiment** is a random experiment where the outcome can be classified as one of two mutually exclusive ways (Heads/Tails, Pass/Fail)

A sequence of **Bernoulli trials** occurs when a Bernoulli experiment is performed several **independent** times, and the success probability, p, remains the same.

coin $p = 0.5$

Bernoulli experiment

- e.g. Flipping a fair coin. If I count the event “heads” as a success, this is a Bernoulli experiment with p=0.5.

- If I toss the coin 10 times, results correspond to 10 Bernoulli trials with p=0.5

Bernoulli Distribution

$\boxed{X \sim Bernoulli(p)}$

" is distributed as "

p prob of a success

If random variable, X, has a Bernoulli distribution:

if $x=1$,

$p^1 \cancel{(1-p)^0}$

→ $f(x) = \underline{p^x (1-p)^{1-x}},$ $\underline{x = \{0,1\}}$

— ▫ $\underline{E[X]} = \sum_{\underline{x=0}}^{1} x \underbrace{p^x (1-p)^{1-x}}_{f(x)} = 0(1-p) + 1(p) = \boxed{p}$

→ ▫ $Var[X] = \sum_{x=0}^{1}(\underline{x} - \underline{p})^2 \underbrace{p^x (1-p)^{1-x}}_{f(x)} = \boxed{p(1-p)}$

→ ▫ $SD[X] = \sqrt{p(1-p)}$

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

$X \sim Bern(p)$

Definition: Random sample

An observed sequence of n Bernoulli trials can be written as a vector of zeroes and ones, with length n. We call this a **random sample** of size n from a Bernoulli distribution.

- X_i denotes the Bernoulli random variable associated with the i^{th} trial.

$i=4$

$n=5$

X_1	X_2	X_3	X_4	X_5
1	0	0	0	1

X_4

$x_1=1$

Example: Sequence of Bernoulli Trials

1 succ

0 fail

$\boxed{p = 0.3}$

$p^x (1-p)^{1-x}$

$\cancel{0.3}\,(1-0.3)^{1-0} = 0.7$

Suppose 30% of all lottery tickets are winners. If five tickets are purchased, (0, 1, 0, 0, 1) is one possible observed outcome.

F S F F S

X_1 X_2

$\{0, 1\}$

0.3

0.7

Assuming independence, the probability of this exact outcome is (0.7)(0.3)(0.7)(0.7)(0.3) = $(0.7)^3(0.3)^2$

$0 = x_1 \qquad x_2 = 1$

p is prob of getting a 1
"success"

- X ~ Bernoulli(p=0.3)

$x = \{0, 1\}$

Binomial Distribution

Often, we are interested only in the total number of successes, but not the actual order of occurrence.

If we let X = the # of observed successes in n Bernoulli trials, then the possible values of X are $0,1,2, \ldots n$.

- For x successes, there are $n - x$ failures.
- X has a **binomial distribution.**

Binomial Distribution

X is a binomial random variable if the following are all true

1. A Bernoulli (success/fail) experiment is performed a constant number of times, n. fixed
2. The random variable, X, is the number of successes in n trials.
3. All trials are independent
4. The success probability, p, for every trial is constant. (The failure probability, $1 - p$, is also constant).

Binomial Distribution

total # success in n trials

Notation: $X \sim Binomial(n,p)$

pmf
- $f(x) = \binom{n}{x} p^x (1-p)^{n-x}, \quad x = 0,1,2,\ldots n$

- $E[X] = np$
- $Var[X] = np(1-p)$

if $n=1$: Bernoulli Dist

https://shiny.rit.albany.edu/stat/binomial/

Lottery Ticket - Binomial Example

$p = .3$

Suppose 30% of lottery tickets are winning tickets.

Let X = the number of winning tickets out of n=5 purchased. The probability of purchasing two winning tickets is

$\binom{n}{x} p^{x} (1-p)^{n-x}$

$x=2$

$$f(2) = P(X = 2) = \underbrace{\binom{5}{2} 0.3^2 (1 - 0.3)^{5-2}}_{\text{Stats}} = \underline{\qquad}$$

- Binomial is a more general case of the Bernoulli

- Bernoulli is a more specific case of Binomial
  - (specifically, $n = 1$)

2.5 Negative Binomial & Geometric Distribution

flip coin until 1st H

Geometric Distribution

Say we observe a sequence of independent Bernoulli trials until the first success occurs.

If X is the number of trials needed to observe the 1^{st} success, then X follows a **Geometric Distribution** with parameter, p.

$p = .2$ 1st success on 8th trial? $.2$

$f(8) = \underline{0.2}\,\underline{(1-0.2)^7}$

Geometric Distribution

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

$f(x;p)$

$f(x) = p(1\text{-}p)^{x\text{-}1}$, $\boxed{x = 1,2,3, \ldots}$

$\sum x \cdot f(x) = 1/p$

- $E[X] = 1/p$
- $Var[X] = \dfrac{1-p}{p^2}$

$P[X \le x] = \sum_{x=1}^{6} x \cdot p(1-p)$

$p = 0.5$ $\quad \frac{1}{0.5} = 2$

$p = 0.1$ $\quad \frac{1}{0.1} = 10$

Pop Quiz

Can you show that the geometric distribution is a valid pmf? How?

Negative Binomial Distribution

More generally, suppose we observe a sequence of independent Bernoulli trials until the r^{th} success occurs.

If X is the number of trials needed to observe the r^{th} success, then X follows a **Negative Binomial** distribution with parameters r, p.

success prob = 0.4

How many trials until I get 3rd success?

$p = .4$

$r = 3$

$\boxed{X \sim NB(3, 0.4)}$

Negative Binomial Distribution

$\cancel{\binom{8}{3}}$ $\binom{7}{2}$

$\boxed{X \sim NB(r, p)}$

$$f(x) = \binom{x-1}{r-1} p^r (1-p)^{x-r}, \quad x = r, r+1, r+2, \ldots$$

$\boxed{f(x) = \binom{x-1}{3-1} 0.4^3 (0.6)^{x-3}}$

- $\boxed{E[X] = r/p}$
- $Var[X] = \dfrac{r(1-p)}{p^2}$

$r=3$

$f(8) =$ _ _ _ _ _ _ _ _ S

2 success

?

If I have a negative binomial with r = 1,

This is a geometric distribution

Examples

2.4 - 2.5

1 A magical beer machine vending machine gives a random beer to the customer. It gives you a stout 30% of the time and an IPA 70% of the time. Let X be the number of stouts you get out of 20 beers.

$p = 0.3$

$X \sim \text{Binom}$ $\quad n = 20$

- What is the distribution of X? What is its pmf?

$X \sim \text{Binomial}(20, 0.3)$

$f(x) = \binom{20}{x} 0.3^{x} (0.7)^{20-x}$

$P[X < 7]$

- What is the probability of getting fewer than 7 stouts?

$P[X \leq 6] = P[X=0] + P[X=1] + \ldots P[X=6]$

dbinom( ) ← pmf

pbinom( ) ← cdf

$= .608$

pbinom(q, size, prob, lower.tail = TRUE, log.p = FALSE)

2 A magical beer machine vending machine randomly gives you a stout 30% of the time and an IPA 70% of the time. Thor gently smashes the machine until it gives him a stout. Let X represent the number of trials required for Thor to get his first stout.

p $(1-p) \leftarrow$ fail OR $q \leftarrow$ fail $p = 0.3$

What is the distribution of X? What is its pmf?

$X \sim Geo(0.3)$ $\rightarrow f(x) = 0.3\,(0.7)^{x-1}$

What is the probability of getting a stout on the 5th trial? ✓

pmf

$\hookrightarrow f(5) = 0.3\,(0.7)^4 = 0.07203$ dgeom(4, 0.3)

What is the probability of getting a stout within the first 5 trials?

$P[X \le 5]$ cdf pgeom( ) cdf

dgeom( ) pmf

.8319

3 A random variable X has a binomial distribution with $\mu = 6$, $\sigma^2=3.6$.

$E[X] = np$
$Var[X] = np(1-p)$

What is the distribution of X? $\boxed{X \sim Bin(n=15,\ p=0.4)}$

$3.6 = np(1-p)$
$\div$
$6 = np$

$\Rightarrow 0.6 = 1-p \Rightarrow p = 0.4$

$6 = (n)(0.4)$
$\Rightarrow n = 15$

Find $P(X = 4)$.

$f(4) = \binom{15}{4} 0.4^4 (0.6)^{11}$

Find $F(2)$. cdf $P[X \le 2] = f(0) + f(1) + f(2)$

dbinom(0, 15, 0.4) + dbinom(1, 15, 0.4) + dbinom(2, 15, 0.4) = .027
pbinom(2, 15, 0.4) = .027

4 Jacqueline hits her free throws with $p = 0.9$.

assume independence between f.t.

What is the probability that she has her first miss on the 7th free throw?

$X \sim Geom(0.1)$ $p(\text{miss}) = 0.1$ $p = 0.1$

$f(7) = 0.1(0.9)^6$

(hit all first 11)

What is the probability that she has her first miss on the 12th attempt or later?

$P[X \geq 12] = 1 - P[X \leq 11]$ 0.9^{11}

$_{29}C_2$ S S F x27

What is the probability that she has her 3rd miss on the 30th free throw?

$\boxed{X \sim NB(3, 0.1)}$

$f(30) = \left[\binom{29}{2} \underbrace{0.1^2}_{} \underbrace{0.9^{27}}_{}\right] 0.1$ or $\binom{29}{2} 0.1^3 \, 0.9^{27}$

- Probability that her first ~~miss~~ success is on 12 or later?

Fail

S S

11

12

$p = 0.1$

$f(x) = 0.1(0.9)^{x-1}$

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

$\frac{a}{1-r}$

$r = 1-p$

Jay Quellin

- Probability that first miss (success) occurs on 12 or later?

$$P[X \geq 12] = \sum_{x=12}^{\infty} p(1-p)^{x-1} = \frac{p(1-p)^{12-1}}{1-(1-p)}$$

$$= \frac{\cancel{p}(1-p)^{11}}{\cancel{p}} = (1-p)^{11} = (0.9)^{11}$$

$$= 0.9^{11} = 0.313$$

$$\boxed{1 - P[X \leq 11] = 1 - \text{pgeom}(10, 0.1)}$$