Department of Mathematics
College of William & Mary

MATH 451/551

Chapter 4. Common Discrete Distributions

4.2 Binomial Distribution

Bernoulli

GuanNan Wang
gwang01@wm.edu

n Bernoulli trial (indep)
X = # success among n trials

Binomial Distribution

- When n Bernoulli trials are conducted, each with an identical probability of success p, the experiment is known as a **binomial** random experiment.
- A binomial random experiment satisfies the following criteria.
  1. The random experiment consists of n identical trials
  2. There are two possible outcomes for each trial
  3. The trials are mutually independent
  4. The probability of success on each trial is identical
- $X \sim Binomial(n, p)$ models the number of successes in n mutually independent Bernoulli trials, each with probability of success p, where n is a positive integer. $X \sim Bin(n,p)$
- **Support of X:** $\mathcal{A} = \{0, 1, 2, \ldots, n\}$.
- **PMF:** $f(x) = \binom{n}{x} p^x (1-p)^{n-x}, \; x = 0, 1, 2, \ldots, n$ for some positive integer n and $0 < p < 1$ is a $Binomial(n, p)$ random variable.

Mean

$X \sim Bin(n,p) \Rightarrow \mathcal{X} = \{0,1,2,\dots,n\} \Rightarrow X$ is discrete.

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

$$E(X) = \sum_{\mathcal{X}} x f(x) = \sum_{x=0}^{n} x f(x) = \sum_{x=1}^{n} x f(x) = \sum_{x=1}^{n} \cancel{x} \frac{n!}{\underset{(x-1)!}{\cancel{x!}}(n-x)!} p^x (1-p)^{n-x}$$

$$\underset{y=x-1}{=} \sum_{y=0}^{n-1} \frac{n!}{y!(n-(y+1))!} p^y \cdot p (1-p)^{n-(y+1)}$$

$$= \sum_{y=0}^{n-1} \frac{(n-1)!\, n}{y!((n-1)+\cancel{1}-(y+\cancel{1}))!} p^y p (1-p)^{(n-1)+1-(y+1)}$$

$$= \sum_{y=0}^{n-1} \frac{(n-1)!\, np}{y!((n-1)-y)!} p^y (1-p)^{n-1-y} \underset{m=n-1}{=} \sum_{y=0}^{m} \frac{m!}{y!(m-y)!} p^y (1-p)^{m-y}\, np$$

$$= np \sum_{y=0}^{m} \frac{m!}{y!(m-y)!} p^y (1-p)^{m-y} = np \cdot 1 = np$$

Variance

$$\frac{x^2}{\cancel{x(x-1)}(x-2)\cdots 1}$$

$$V(X) = E(X^2) - \{E(X)\}^2 = \underbrace{E(X^2)} - n^2p^2$$

$$= E(X(X-1) + X) - n^2p^2 = E(X(X-1)) + E(X) - n^2p^2 = \underbrace{E(X(X-1))}_{?} + np - n^2p^2$$

$$E\{X(X-1)\} = \underbrace{\sum_{x=0}^{n} x(x-1) \frac{n!}{\cancel{x!}(n-x)!} p^x (1-p)^{n-x}}_{(x-2)!}$$

$$= \sum_{x=2}^{n} \frac{(n-2)!\, n(n-1)}{\underbrace{(x-2)}!\{(n-2)-(x-2)\}!} p^{x-2} p^2 (1-p)^{\{(n-2)-(x-2)\}}$$

$$\overset{y=x-2}{\underset{m=n-2}{=}} \underbrace{\sum_{y=0}^{m} \frac{m!}{y!(m-y)!} p^y (1-p)^{m-y}}_{1} \underbrace{n(n-1)p^2} = n(n-1)p^2$$

$$V(X) = n(n-1)p^2 + np - n^2p^2 = np\{\overset{np-p}{\underbrace{(n-1)p}} + 1 - \cancel{np}\} = np(1-p)$$

X_i = ith Bernoulli Trial.

Bernoulli Trial. $\begin{cases} p & X=1 \\ (1-p) & X=0 \end{cases}$

Y = # success among n trial

$E(X_i) = p \qquad V(X_i) = p(1-p)$

$n \Downarrow$

Bin $Y = X_1 + X_2 + \cdots + X_n$

$\Rightarrow E(Y) = E(X_1) + E(X_2) + \cdots + E(X_n) = p + p + \cdots + p = np.$

$V(Y) = V(X_1 + X_2 + \cdots + X_n) \overset{\text{indep}}{=\!=} V(X_1) + V(X_2) + \cdots + V(X_n)$

$= np(1-p)$

MGF

$X_i \overset{iid}{\sim} Ber(p)$

indep

identically

$Y = X_1 + X_2 + \cdots + X_n \sim Bin(n,p)$

$$M_Y(t) = E(e^{tY}) = E\left(e^{t(X_1+X_2+\cdots+X_n)}\right)$$

$$= E\left(e^{tX_1} e^{tX_2} \cdots e^{tX_n}\right)$$

$$\underset{\text{indep}}{=} E(e^{tX_1})\, E(e^{tX_2}) \cdots E(e^{tX_n})$$

$$= M_{X_1}(t)\, M_{X_2}(t) \cdots M_{X_n}(t)$$

$$\underset{\text{identical}}{=} \{M_{X_i}(t)\}^n = (1-p+pe^t)^n, \quad -\infty < t < \infty$$

Skewness

$$E\left\{\left(\frac{X-\mu}{\delta}\right)^3\right\} = \frac{1-2p}{\sqrt{np(1-p)}}$$

l

Kurtosis

$$E\left\{\left(\frac{X-\mu}{\delta}\right)^4\right\} = 3 + \frac{1-6p(1-p)}{np(1-p)}$$

Example 1

If Emma takes 3 free throws (she is a 70% free throw shooter) and X is the number that she makes, find $f(x)$, μ, σ^2, and $P(X = 2)$.

①. $n=3$ identical Bernoulli trials
②. make $p=70\%$, fail $1-p=30\%$
③ indep.

$\Rightarrow X \sim Bin(n=3, p=70\%)$

$f(x) = \binom{3}{x} 0.7^{x} 0.3^{3-x}, \quad x = 0,1,2,3$

$\mu = np = 3 \times 0.7 = 2.1$

$\sigma^2 = np(1-p) = 3 \times 0.7 \times 0.3 = 0.63$

$P(X=2) = \binom{3}{2} 0.7^2 0.3^1 = 0.441$

Example 2

A dozen eggs contain 3 defectives. If a sample of 5 is taken with replacement, find the probability that

1. exactly 2 of the eggs sampled are defective
2. 2 or fewer of the eggs sampled are defective

Let X = # defective eggs in the sample

① $n=5$ Bernoulli trial

② $\begin{cases} \text{defective egg} \\ \text{good egg} \end{cases}$

③. indep

④. $p = \frac{3}{12}$

$X \sim Bin(n=5,\ p=\frac{1}{4})$

1. $P(X=2) = \binom{5}{2}\left(\frac{1}{4}\right)^2\left(\frac{3}{4}\right)^{5-2}$

2. $P(X \le 2) = P(X=2) + P(X=1) + P(X=0)$

$= \sum_{x=0}^{2} \binom{5}{2}\left(\frac{1}{4}\right)^x\left(\frac{3}{4}\right)^{5-x}$

Example 3

An airplane has 100 seats. The airline "overbooks" a flight (sells more tickets than available seats) in order to maximize their profit. Assume that each ticket holder's decision to show up for a flight is an independent Bernoulli trial with a probability of showing up for the flight of 0.92. If the airline profit is \$10 for each seat sold and the airline loses \$40 for each "bumped" passenger, what is the expected profit if 103 seats are sold?

Let X = # ticket holders show up.

① $X_i \overset{iid}{\sim} Ber(0.92)$

② $X = \sum X_i \sim Bin(103, 0.92) \Rightarrow f(x) = \binom{103}{x} 0.92^{x} \, 0.08^{103-x}, \; x = 0, 1, 2, \dots, 103$

Let Y = # profit

$$Y = \begin{cases} 10 * 103 & 0 \le x \le 100 \\ 10 * 103 - 40(103 - x) & 100 < x \le 103 \end{cases}$$

$$E(Y) = 10 \times 103 \times \sum_{x=0}^{100} f(x) + \{10 \times 103 - 40(103 - x)\} \sum_{x=101}^{103} f(x)$$

$$= 1029.548$$

Type I error $= P(\text{reject } H_0 \mid H_0 \text{ is true})$

Example 4

Type II error $= P(\text{fail to reject } H_0 \mid H_0 \text{ is false})$

A company produces biased coins that come up heads when flipped with probability 0.7. You are not sure whether you have one of these biased coins or whether you have a fair coin, so you devise the following experiment: (1) Flip the coin 100 times, (2) If there are 62 or more heads conclude that the coin is biased, otherwise, conclude that the coin is fair.

$X = \#$ heads among 100 tosses

		True Coin Status	
	?	Coin fair	Coin biased
Experiment Conclusion	Coin fair	✓ 0.9895	✗ 0.0340
	Coin biased	✗ 0.0105	✓ 0.9660

Fair: $X \sim Bin(n=100, p=0.5)$

$P(X \le 62) = \sum_{x=0}^{62} \binom{100}{x} 0.5^x (1-0.5)^{100-x}$

$P(X > 62) = 1 - P(X \le 62)$

Biased: $X \sim Bin(n=100, p=0.7)$

$P(X \le 62) = \sum_{x=0}^{62} \binom{100}{x} 0.7^x (0.3)^{100-x}$

$P(X > 62) = 1 - P(X \le 62)$

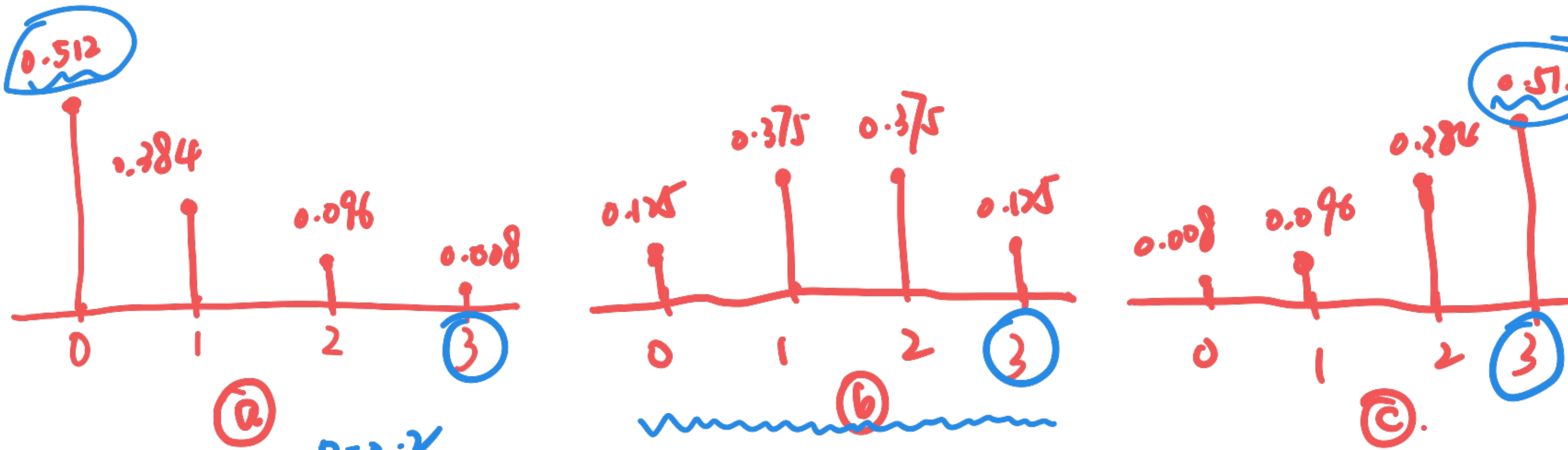
R Functions

R Functions

$P(X = x) = ?$

Function	Returned Value
dbinom(x, n, p)	calculates the probability mass function $f(x)$
pbinom(x, n, p)	calculates the cumulative ~~mass~~ density function $F(x)$
qbinom(u, n, p)	calculates the percentile (quantile) $F^{-1}(u)$
rbinom(m, n, p)	generates m random variates

$f(x)$ $\quad$ $Bin(n, p)$



$p = 0.2$

$n = 3 \quad p = \frac{1}{2}$

$p = 0.8$

$S = \{0, 1, 2, \cdots, n\}$

$p = 0.8$

$p = 0.2$

$p = 0.5$

THANK YOU!