

AMS 507

Chapter 4
Random Variables

Sections 6-8

Hongshik Ahn

4.6 Bernoulli and Binomial RV's

- Bernoulli Trial

a. Each trial yields one of two outcomes: Success (S) or Failure (F)

b. $P(S) = P(X = 1) = p;$
 $P(F) = P(X = 0) = 1 - p$

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

c. Each trial is independent.

Bernoulli RV's

Let $X \sim \text{Bernoulli}(p)$. Then

$$p(x) = p^x(1 - p)^{1-x}, \quad x = 0, 1$$

$$E(X) = p$$

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

Proof

Binomial Distribution

n : a fixed number of independent Bernoulli trials

p : the probability of success in each trial

X : # successes in n trials

X is a binomial random variable: $X \sim \text{Bin}(n, p)$

If Y_1, Y_2, \dots, Y_n are independent Bernoulli(p), then

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

$$p(x) = \binom{n}{x} p^x (1-p)^{n-x}, x = 0, 1, \dots, n$$

Show that $p(x)$ is a pmf.

$$E(X) = np, \text{Var}(X) = np(1-p)$$

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

$$E(X) = np, \text{Var}(X) = np(1 - p)$$

Proof

Example 4.6.1

Let X be the number of heads in 7 independent tosses of an unbiased coin. Then $X \sim \text{Bin}(7, 1/2)$.

$$P(0 \leq X \leq 1) =$$

$$E(X) =$$

$$\text{Var}(X) =$$

Example 4.6.2

If the probability is 0.1 that a certain device fails a comprehensive safety test, what are the probabilities that among 15 of such devices,

(a) at most two will fail?

(b) at least three will fail?

Example 4.6.3

30% of the automobiles in a certain city are foreign made. Four cars are selected at random.

X : #cars sampled that are foreign made

$$P(X = 3) =$$

$$P(X \geq 3) =$$

$$P(X \leq 1) =$$

$$P(X < 2) =$$

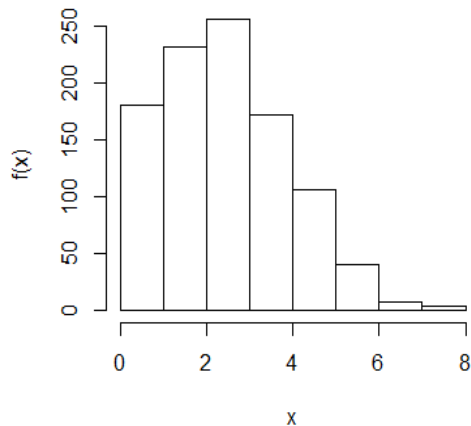
Binomial distribution

$p < 0.5$: skewed to the right

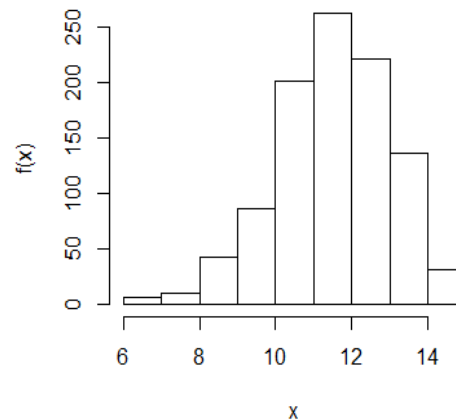
$p > 0.5$: skewed to the left

$p = 0.5$: symmetric

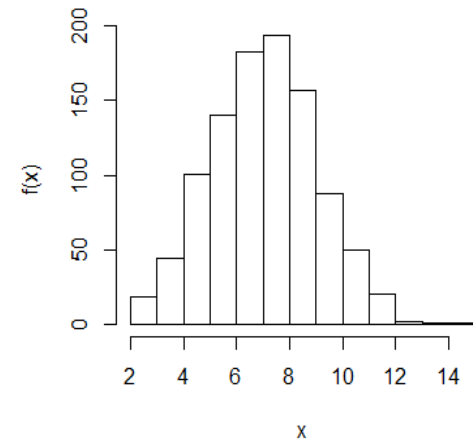
(a) Histogram of Bin(15,0.2)



(b) Histogram of Bin(15,0.8)



(c) Histogram of Bin(15,0.5)



4.7 Poisson RV's

- Properties:
 1. The probability that an event occurs in an interval is proportional to the length of the interval.
 2. Two events cannot occur at exactly the same instant.
 3. Events occur independently.
- The probability of x events occurring in a time period t for a Poisson random variable with parameter λ : $X \sim \text{Poisson}(\lambda)$

$$p(x) = P(X = x) = \frac{\lambda^x e^{-\lambda}}{x!}, x = 0, 1, \dots$$

$$E(X) = \text{Var}(X) = \lambda$$

Poisson Distribution

Show that $p(x)$ is a pmf.

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

$$E(X) = \lambda$$

$$\text{Var}(X) = \lambda$$

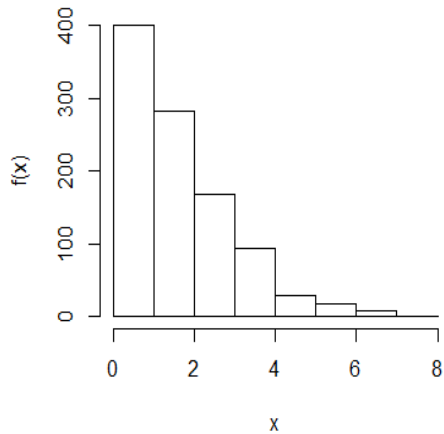
Proof

- Use of Poisson distributions, examples:
 - number of occurrences in a given time
 - distribution of bomb hits in an area
 - distribution of fish in a lake

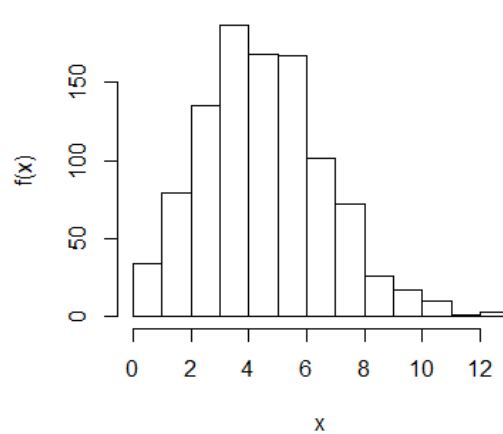
Poisson Distribution

- The Poisson distribution is highly skewed for small values of λ .
- As λ increases, the distribution becomes more symmetric.

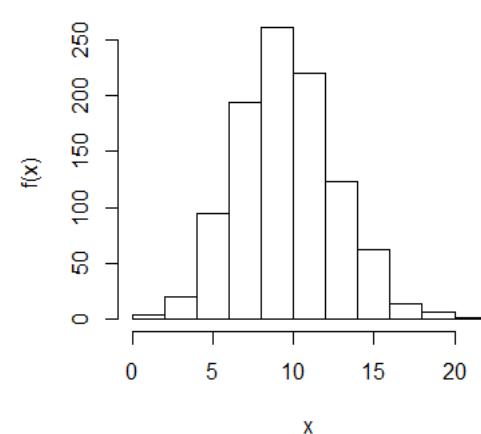
(a) Histogram of Poisson(2)



(b) Histogram of Poisson(5)



(c) Histogram of Poisson(10)



Example 4.7.1

A 911 operator handles 4 calls every 3 hours on average.

- (a) What is the probability of no calls in the next hour?

- (b) Find the probability of at most two calls in the next hour.

Example 4.7.2

On average, 12 cars pass a toll gate booth in a minute during rush hours.

(a) Probability that one car passes the booth in 3 seconds:

(b) Probability that at least two cars pass in 5 seconds:

(c) Probability that at most one car passes in 10 seconds:

Poisson Approximation to the Binomial

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

Let $n \rightarrow \infty$ and $p \rightarrow 0$ such that np remains constant (np is a moderate number). Then

X is approximately $\text{Poisson}(\lambda)$ such that $\lambda = np$.

$n \geq 20, p \leq 0.05$: acceptable

$n \geq 100, p \leq 0.01$: excellent approximation

$$X \sim \text{Bin}(n, p) \implies E(X) = np = \lambda$$

$$\text{Var}(X) = np(1 - p) = \lambda(1 - p) \rightarrow \lambda \text{ as } n \rightarrow \infty \text{ \& } p \rightarrow 0$$

Poisson Approximation to the Binomial

Alternative Approach

$$X \sim \text{Bin}(n, p), \lambda = np, n \rightarrow \infty$$

$$\begin{aligned} \Rightarrow P(X = x) &= \frac{n!}{(n-x)! x!} p^x (1-p)^{n-x} = \frac{n!}{(n-x)! x!} \left(\frac{\lambda}{n}\right)^x \left(1 - \frac{\lambda}{n}\right)^{n-x} \\ &= \frac{n(n-1) \cdots (n-x+1)}{n^x} \frac{\lambda^x}{x!} \frac{[(1 - (\lambda/n))]^n}{[1 - (\lambda/n)]^x}, \quad x = 0, 1, 2, \dots \end{aligned}$$

For $n \rightarrow \infty$,

$$\left(1 - \frac{\lambda}{n}\right)^n \approx e^{-\lambda}, \quad \frac{n(n-1) \cdots (n-x+1)}{n^x} \approx 1, \quad \left(1 - \frac{\lambda}{n}\right)^x \approx 1$$

Hence,

$$P(X = x) \approx \frac{\lambda^x e^{-\lambda}}{x!}, \quad x = 0, 1, 2, \dots$$

Poisson Approximation to the Binomial

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

$$\Rightarrow P(X = x) = \frac{\lambda^x e^{-\lambda}}{x!} = \frac{\lambda}{x} P(X = x - 1), \quad x = 1, 2, \dots$$

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

$$\Rightarrow P(Y = y) = \binom{n}{y} p^y (1 - p)^{n-y} = \frac{n - y + 1}{y} \cdot \frac{p}{1 - p} P(Y = y - 1)$$

$$= \frac{np - p(y - 1)}{y - yp} P(Y = y - 1) \approx \frac{\lambda}{y} P(Y = y - 1), \quad y = 1, 2, \dots, n$$

If $\lambda = np$ and $p \approx 0$,

$$P(Y = 0) = (1 - p)^n = \left(1 - \frac{\lambda}{n}\right)^n \rightarrow e^{-\lambda} = P(X = 0)$$

as $n \rightarrow \infty$

Example 4.7.3

A publisher of mystery novels tries to keep its books free of typos. The probability of any given page containing at least one typo is 0.003 and errors are independent from page to page. What is the approximate probability that a 500 page book has

(a) exactly 1 page with typos?

(b) at most 2 pages with typos?

4.8 Other Discrete Probability Distribution

Hypergeometric RV

- Sampling w/o replacement

- 1) Finite population with N individuals (binomial: infinite population or with replacement)
- 2) S or F . There are m successes in the population.
- 3) Sample size: n . Each subset of size n is equally likely to be chosen.

$$P(X = x) = \frac{\binom{m}{x} \binom{N-m}{n-x}}{\binom{N}{n}}$$

$$\max\{0, n - N + m\} \leq x \leq \min\{n, m\}$$
$$(\because x \leq m \text{ and } n - x \leq N - m)$$

$$E(X) = \frac{nm}{N}, \quad Var(X) = \frac{nm(N-m)(N-n)}{N^2(N-1)}$$

Example 4.8.4

A shipment of 25 CD's contains 5 that are defective. If 10 of them are randomly chosen without replacement, what is the probability that 2 of the 10 will be defective?

Hypergeometric RV

$$E(X) = \frac{nm}{N}, \quad \text{Var}(X) = \frac{nm(N-m)(N-n)}{N^2(N-1)}$$

Let $\frac{m}{N} = p$. Then

$$E(X) = np, \text{Var}(X) = np(1-p) \left(\frac{N-n}{N-1} \right)$$

If $N \gg n$, then $\text{Var}(X) \approx np(1-p)$

$\frac{N-n}{N-1}$: finite population correction factor

Binomial approximation to the hypergeometric, when

$\frac{n}{N} \leq 0.1$, and $p = \frac{m}{N}$ is not too close to either 0 or 1.

Example 4.8.5

From Example 4.8.4, for a lot of 100 CD's, 20 are defective. Find the probability that among a randomly selected sample of 10 CD's, 2 are defective, by using

- (a) The formula for the hypergeometric distribution:

- (b) The binomial approximation to the hypergeometric:

- Geometric RV: $X \sim \text{Geometric}(p)$
 1. Each trial: Success (S) or Failure (F)
 2. Independent trials
 3. $P(S) = p$
 4. Continues until the first success is observed
- Probability of getting the first success on the x -th trial:

$$p(x) = p(1 - p)^{x-1}, x = 1, 2, \dots$$
$$E(X) = \frac{1}{p}, \quad \text{Var}(X) = \frac{1 - p}{p^2}$$

Geometric RV

$$p(x) = p(1 - p)^{x-1}, x = 1, 2, \dots$$

Show $p(x)$ is a pmf.

Geometric RV

- CDF:

$$F(x) = \begin{cases} 1 - (1 - p)^x, & k \leq x < k + 1, k = 1, 2, \dots \\ 0, & x < 1 \end{cases}$$

Proof

Geometric RV

$$E(X) = \frac{1}{p}$$

Proof

Geometric RV

$$E(X) = \frac{1}{p}$$

Alternative proof

Geometric RV

$$\text{Var}(X) = \frac{1 - p}{p^2}$$

Proof

Example 4.8.1

A fair die is tossed until a certain number appears. What is the probability that the first six appears at the fifth toss?

Negative Binomial RV: $X \sim \text{NB}(r, p)$

- Independent Bernoulli Trials
- X : number of trials until r -th success
- $r = 1 \Rightarrow X \sim \text{Geometric}(p)$

$$p(x) = \binom{x-1}{r-1} p^r (1-p)^{x-r}, x = r, r+1, \dots$$

$$E(X) = \frac{r}{p}$$

$$\text{Var}(X) = \frac{r(1-p)}{p^2}$$

Negative Binomial RV

$$E(X) = \frac{r}{p}$$

Proof

Negative Binomial RV

$$\text{Var}(X) = \frac{r(1-p)}{p^2}$$

Proof

Negative Binomial RV

- Let X_1, X_2, \dots, X_n be independent with $X_i \sim \text{Geometric}(p), i = 1, 2, \dots, n$.

Then

$$Y = \sum_{i=1}^n X_i \sim \text{NB}(n, p).$$

- Let X_1, X_2, \dots, X_n be independent with $X_i \sim \text{NB}(r, p), i = 1, 2, \dots, n$.

Then

$$Y = \sum_{i=1}^n X_i \sim \text{NB}(nr, p).$$

Example 4.8.2 (Ross Example)

A pipe-smoking mathematician carries, at all times, 2 matchboxes, 1 in his left-hand pocket and 1 in his right-hand pocket. Each time he needs a match, he is equally likely to take it from either pocket. Consider the moment when the mathematician first discovers that one of his matchboxes is empty. If it is assumed that both matchboxes initially contained N matches, what is the probability that there are exactly k matches in the other box, $k = 0, 1, \dots, N$?

Answer to Example 4.8.2

Example 4.8.3 (Ross Example)

Find the expected value and the variance of the number of times one must throw a die until the outcome 1 has occurred 4 times.