Bernoulli Trial

Definition: Bernoulli trial is a random experiment whose outcomes are classified as one of the two categories, (S, F) or (Success, Failure) or (1, 0)

\[ X(\text{Success}) = 1, \quad X(\text{Failure}) = 0 \]

Example:
• Tossing a coin, observing Head or Tail
• Observing patient’s status Died or Survived.

Bernoulli Probability Distribution

Example: (Tossing a balanced coin)

\[ P(X=1) = f(1) = p = 0.5 \]
\[ P(X=0) = f(0) = 1 - p = 0.5 \]

Bernoulli Distribution

\[ X \sim \text{Bernoulli}(p) \]

The expected value of \( X \) is \( \mu = p \)
The variance of \( X \) is \( \sigma^2 = p(1 - p) \)
The standard deviation of \( X \) is \( \sigma = \sqrt{p(1 - p)} \)

Binomial Experiment

A random experiment involving a sequence of independent and identical Bernoulli trials.

Example:
• Toss a coin ten times, and observing Head turns up.
• Roll a die 3 times, and observing a 6 turns up or not.
• In a random sample of 5 from a large population, and observing subjects’ disease status. (Almost binomial)
Discrete Distributions

Binomial Probability Model

A model to find the probability of having $x$ number successes in a sequence of $n$ independent and identical Bernoulli trials. (Random sample of size $n$ from Bernoulli distribution.)

Binomial Probability Distribution $b(n, p)$

In a binomial experiment involving $n$ independent and identical Bernoulli trials each with probability of success $p$, the probability of having $x$ successes can be calculated with the binomial probability mass function, and it is, for $x = 0, 1, \ldots, n$,

$$f(x) = P(X = x) = \frac{n!}{x!(n-x)!} \cdot p^x \cdot (1 - p)^{n-x}$$

Example:

A balanced die is rolled three times (or three balanced dice are rolled), what is the probability to see two 6’s?

Identify $n = 3$, $p = \frac{1}{6}$, $x = 2$

$$f(x) = \frac{3!}{2!(3-2)!} \left(\frac{1}{6}\right)^2 \left(\frac{5}{6}\right)^{1}$$

$$= \frac{3!}{2!(3-2)!} \left(\frac{1}{6}\right)^2 \left(\frac{5}{6}\right)^{1}$$

$$= \frac{3}{2} \cdot \frac{1}{6} \cdot \frac{5}{6}$$

$$= .069$$

Example:

If 10% of the population in a community have a certain disease, what is the probability that 4 people in a random sample of 5 people from this community has the disease? (Assume binomial experiment.)

Identify $n = 5$, $p = .10$, $x = 4$

$$f(x) = \frac{5!}{4!(5-4)!} \cdot (.10)^4 \cdot (1-.10)^{5-4}$$

$$= \frac{5!}{4! 1!} \cdot (.1)^4 \cdot (.9)^{1}$$

$$= .0004$$

Binomial Probability

Example: In the previous problem, what is the probability that 4 or more people have the disease?

Identify $n = 5$, $p = .10$, $x = 4$

$$P(X \geq 4) = P(X = 4) + P(X = 5) = \frac{5!}{4! \cdot 1!} \cdot (.1)^4 \cdot (.9)^1 + \frac{5!}{5! \cdot 0!} \cdot (.1)^0 \cdot (.9)^5$$

$$= .00045 + .00001 = .00046$$

$$= .0004$$  (What is this number telling us?)

Factorial

$n! = 1 \cdot 2 \cdot 3 \cdot \ldots \cdot n$

$0! = 1$

Example: $3! = 1 \cdot 2 \cdot 3 = 6$

Example:

$$\binom{5}{2} = \frac{5!}{2!(5-2)!} = \frac{5 \cdot 4 \cdot 3 \cdot 2 \cdot 1}{3 \cdot 2 \cdot 1 \cdot 2 \cdot 1} = 10$$

Binomial Probability

Example: A balanced die is rolled three times (or three balanced dice are rolled), what is the probability to see two 6’s?

Identify $n = 3$, $p = \frac{1}{6}$, $x = 2$

$$f(x) = \frac{3!}{2!(3-2)!} \cdot \left(\frac{1}{6}\right)^{1} \cdot \left(\frac{5}{6}\right)^{2}$$

$$= \frac{3!}{2!(3-2)!} \cdot \left(\frac{1}{6}\right)^{1} \cdot \left(\frac{5}{6}\right)^{2}$$

$$= 3 \cdot \left(\frac{1}{6}\right)^{1} \cdot \left(\frac{5}{6}\right)^{2}$$

$$= .069$$

Example: In the previous problem, what is the probability that 4 or more people have the disease?

Identify $n = 5$, $p = .10$, $x = 4$

$$P(X \geq 4) = P(X = 4) + P(X = 5) = \frac{5!}{4! \cdot 1!} \cdot (.1)^4 \cdot (.9)^1 + \frac{5!}{5! \cdot 0!} \cdot (.1)^0 \cdot (.9)^5$$

$$= .00045 + .00001 = .00046$$

$$= .0004$$  (What is this number telling us?)
Discrete Distributions

Parameters of Binomial Distribution

Parameters of the distribution:
- Mean of the distribution, $\mu = np$
- Variance of the distribution, $\sigma^2 = np(1 - p)$
- Standard deviation, $\sigma = \sqrt{np(1 - p)}$

Binomial Distribution

$n = 5, p = .10$

- $f(0) = .5905$
- $f(1) = .3280$
- $f(2) = .0729$
- $f(3) = .0081$
- $f(4) = .0004$
- $f(5) = .00001$

Cumulative Distribution Function

The cumulative distribution function (c.d.f. or distribution function, d.f.) of a discrete random variable is defined as

$$F(x) = P(X \leq x)$$

Distribution function for binomial distribution

$$F(x) = P(X \leq x) = \sum_{k=0}^{x} \binom{n}{k} \cdot p^k \cdot (1 - p)^{n-k}$$

Use of Table (p.677)

Binomial table presents cumulative probabilities

$n = 5, p = .10$

- $P(X \leq 1) = .9185$
- $P(X \leq 2) = .9914$
- $P(X = 2) = P(X \leq 2) - P(X \leq 1) = .0729$

Hypergeometric Distribution

(Sampling without replacement)

As $N$ increases, $p = \frac{N}{N}$

Binomial Distribution

(Sampling with replacement)
Chapter 2 Discrete Distributions

2.5 The Moment Generating Function

Moment Generating Function

The moment generating function (m.g.f.) of a discrete random variable \( X \), with a space \( S \), where it exists, is given by

\[
M_X(t) = E(e^{tx}) = \sum_{x \in S} e^{tx} f(x)
\]

for a positive value \( h \), such that \(-h < t < h\).

\( M_X(0) = 1 \).

Why generating function?

The Maclaurin’s Series expansion of \( e^{tx} \) is

\[
e^{tx} = 1 + tx + \frac{t^2 x^2}{2!} + \frac{t^3 x^3}{3!} + \ldots + \frac{t^r x^r}{r!} + \ldots
\]

If two random variables have the same m.g.f., they must have the same distribution.

\[
M_X(t) = e^{\theta h} f(h_1) + e^{\theta h} f(h_2) + e^{\theta h} f(h_3) + \ldots
\]

Where \( f(h) = g(h), \quad \forall \ i = 1, 2, 3, \ldots \)
Discrete Distributions

Example
If the random variable $X$ has the m.g.f.

$$M(t) = e^{t \left( \frac{3}{6} + e^{2t} \left( \frac{2}{6} \right) + e^{3t} \left( \frac{1}{6} \right) \right)}$$

then the probability distribution would be

$$P(X=1) = \frac{3}{6}, \quad P(X=2) = \frac{2}{6}, \quad P(X=3) = \frac{1}{6}$$

and the p.m.f. of $X$ would be

$$f(x) = \frac{4-x}{6}, \quad x = 1, 2, 3.$$ 

Example
If the moments of the random variable $X$ is

$$E[X^r] = 0.8, \quad r = 1, 2, 3, \ldots,$$

then the m.g.f. of $X$ would be

$$M(t) = 1 + \sum_{r=1}^{\infty} E[X^r] \left( \frac{t^r}{r!} \right) = 1 + 0.8 \sum_{r=1}^{\infty} \left( \frac{t^r}{r!} \right)$$

$$= 0.2 + 0.8 \sum_{r=1}^{\infty} \left( \frac{t^r}{r!} \right) = 0.2e^{0.8t} + 0.8e^{0.8t}$$

Thus the p.m.f. of $X$ would be

$$P(X=0) = 0.2, \quad P(X=1) = 0.8.$$

Derivatives of m.g.f.

**Theorem:**
If $X$ is a random variable that has a m.g.f. $M(t)$, for $-h < t < h$, and $h > 0$, if the derivatives of all orders exist at $t = 0$ for $M(t)$, then

$$M^{(r)}(0) = E[X^r]$$

$$M_{x}(t) = E(e^{tX}) = 1 + E[X]t + E[X^2] \frac{t^2}{2!} + \ldots + E[X^r] \frac{t^r}{r!} + \ldots$$

Some Special Moments

$$\mu = M'(0)$$

$$\sigma^2 = M''(0) - [M'(0)]^2$$

Binomial Distribution

The p.m.f. of the binomial distribution is

$$f(x) = \binom{n}{x} p^x (1-p)^{n-x}, \quad x = 1, 2, 3, \ldots, n.$$ 

So, the m.g.f. of the binomial distribution is

$$M_x(t) = E(e^{tX}) = \sum_{x=0}^{n} e^{tx} \binom{n}{x} p^x (1-p)^{n-x}$$

$$= \sum_{x=0}^{n} \binom{n}{x} (pe^t)^x (1-p)^{n-x}$$

$$= [(1-p) + pe^t]^n$$

$$M(t) = (1-p + pe^t)^n$$

$$M'(t) = n[(1-p) + pe^t]^{n-1} (pe^t)$$

$$M''(t) = n(n-1)[(1-p) + pe^t]^{n-2} (pe^t)^2$$

$$+ n[(1-p) + pe^t]^{n-1} (pe^t)$$

Mean and variance of Binomial Distribution

$$\mu = E[X] = M'(0) = np$$

$$\sigma^2 = E[X^2] - [E(X)]^2 = M''(0) - [M'(0)]^2$$

$$= n(n-1)p^2 + np - (np)^2 = np(1-p)$$
Discrete Distributions

It is reported that 10% of the apple from the Apple Farm is bad. If apple is randomly selected from this farm one after another, what is the probability that the 10th apple selected is the 2nd bad apple selected?

Negative Binomial Distribution

Let random variable $X$ be the number of trials needed to observe the $r$-th success in a sequence of Bernoulli trials. The random variable $X$ has a Negative Binomial Distribution with p.m.f.

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

The probability of having the first $r-1$ successes in first $x-1$ trials.

Geometric Distribution

If $r = 1$ in the negative binomial distribution, then the random variable $X$ has a geometric distribution with p.m.f.

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

Sum of Geometric Series

$$\sum_{x=0}^{\infty} ar^x = \sum_{x=0}^{\infty} ar^{x-1} = \frac{a}{1-r}, \quad \text{for } |r| < 1$$

$$\sum_{x=1}^{\infty} f(x) = \sum_{x=1}^{\infty} p(1-p)^{x-1} = \frac{p}{1-(1-p)} = 1$$

Special Result for Distribution Function:

$$F(x) = P(X \leq x) = 1 - (1-p)^x$$
Discrete Distributions

Example
It is reported that 10% of the apple from the Apple Farm is bad. If apple is randomly selected from this farm one after another, what is the probability that the 10th apple selected is the 2nd bad apple selected?

Negative Binomial Distribution: \( p = .1, r = 2, x = 10 \)

\[
f(x) = \binom{x-1}{r-1} p^r (1-p)^{x-r}, \quad \text{for } x = r, r+1, \ldots
\]

\[
f(10) = \binom{10-1}{2-1} .1^2 (1-.1)^{10-2} = 0.0387
\]

Example
It is reported that 10% of the apple from the Apple Farm is bad. If apple is randomly selected from this farm one after another, what is the probability that the 3rd apple selected is the 1st bad apple selected?

Geometric Distribution: \( p = .1, x = 3 \)

\[
f(x) = p(1-p)^{x-1}, \quad x = 1, 2, 3, \ldots
\]

\[
f(3) = .1(1-.1)^{3-1} = 0.081
\]

Negative Binomial Distribution (in R)
Let random variable \( X \) be the number of failures needed to observe the \( r \)-th success in a sequence of Bernoulli trials. The random variable \( X \) has a Negative Binomial Distribution with \( p.m.f. \)

\[
f(x) = \binom{x+r-1}{r-1} p^r (1-p)^{x-r}, \quad \text{for } x = 0, 1, 2, \ldots
\]

The probability of having the first \( x \) failures in first \( x+r-1 \) trials.

\[
f(x) = \binom{x+r-1}{r-1} p^r (1-p)^{x-r}
\]

\[
f(x) = \binom{x+r-1}{r-1} .1^r (1-.1)^{x-1}
\]

\[
\text{Chance of success at the last trial.}
\]

Geometric Distribution (In R)
If \( r = 1 \) in the negative binomial distribution, then the random variable \( X \) has a geometric distribution with \( p.m.f. \),

\[
f(x) = p(1-p)^{x}, \quad x = 0, 1, 2, 3, \ldots
\]

where \( x \) is number of failures before observing the 1st success.

Chapter 2 Discrete Distributions

2.6 The Poisson Distribution
Discrete Distributions

Examples of Poisson Process

• Number of people visiting to the emergency room for treatment per hour.
• Number of customers coming to the Arby’s to buy sandwich per ten minutes.

Poisson Distribution

Let $X$ be a random variable takes on the number of occurrences of some event of interest over a given interval from a Poisson process, and the $\lambda$ is the mean of the distribution, then the probability of observing $x$ occurrences is, for $x = 0, 1, 2, ...$

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

It can be used to approximate Binomial prob., for large $n$.

$e = 2.71828...$

Poisson Process

• The probability that a single event occurs within an interval is proportional to the length of the interval.
• Within a single interval, an infinite number of occurrences is possible.
• The events occur independently both within the same interval and between consecutive non-overlapping intervals.

Poisson Model

$$\mu = \lambda$$

$$\sigma^2 = \lambda$$

$$M(t) = e^{\mu(e^\lambda - 1)}$$

Poisson Probability

If on average there are 4 people catch flu in a given week in a community during a certain season, what is the probability of observing 2 people catch flu in this community in a given week period during the season? (Assume the number of people catching flu in a given period of time follow a Poisson Process.)

$$\lambda = 4$$

$$x = 2$$

$$P(X = 2) = .1465 = \frac{e^{-4} 4^2}{2!}$$

Poisson Probability

If on average there are 4 people catch flu in a given week in a community during a certain season, what is the probability of observing 2 people catch flu in this community in a given two weeks period during the season? (Assume the number of people catching flu in a given period of time follow a Poisson Process.)

$$\lambda = 4 \times 2 = 8$$

$$x = 2$$

$$P(X = 2) = .011 = \frac{e^{-8} 8^2}{2!}$$
Discrete Distributions

Poisson Probability

If on average there are 4 people catch flu in a given week in a community during a certain season, what is the probability of observing less than 2 people catch flu in this community in a given week period during the season? (Assume the number of people catching flu in a given period of time follow a Poisson Process.)

\[
\lambda = 4 \\
x = 0 \text{ and } 1 \\
P(X < 2) = P(X = 0) + P(X = 1) \\
= (e^{-4}4^0)/0! + (e^{-4}4^1)/1! = .0916
\]

Discrete Probability Models

- **Binomial**: Number of successes in a binomial experiment. (There is a sample taken.)
- **Poisson**: Number of successes in a given time period or in a given unit space. (No sample taken.)

Poisson approximation to Binomial Probability:

\[
(n \geq 20 \text{ and } p \leq .05 \text{ or } n \geq 100 \text{ and } p \leq .10)
\]

\[
\lambda = np, \quad \frac{(np)^x e^{-np}}{x!} \approx \binom{n}{x} p^x (1-p)^{n-x}
\]