Introduction to Statistics

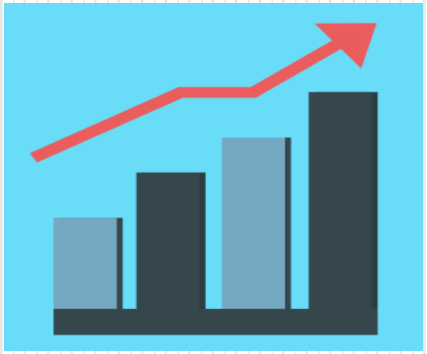
Dr. Farman Ali

Assistant Professor

DEPARTMENT OF SOFTWARE

SEJONG UNIVERSITY

Lecture-8



Course Syllabus

- Introduction to the course
- Sampling and data presentation
- Basic of probability
- Distributions
- Confidence intervals
- Hypothesis testing
- Correlation and simple linear regression
- Multiple regression

Outline

- **Mathematical Expectation of random variables**
- **Bernoulli and Binomial Distribution**
- **Poisson Distribution**

Mathematical Expectation of a Random Variable

Mathematical Expectation of a Random Variable

Let a discrete r.v. X have possible values, x1, x2, …, xn, … with corresponding probabilities f(x1), f(x2), …, f(xn), …. Such that

$$\sum f(x) = 1$$

Then the mathematical expectation or the expectation or the expected value of X, denoted by E(X), is defined as:

$$E(X) = x_1 f(x_1) + x_2 f(x_2) + \ldots + x_n f(x_n) + \ldots = \sum_{i=1}^{\infty} x_i f(x_i)$$

Mathematical Expectation of a continuous r..v. X is defined as:

$$E(X) = \int_{-\infty}^{+\infty} x\, f(x)\, dx$$

Properties of Mathematical Expectation

Properties of mathematical Expectation of a random variable:

- $E(a)=a$, where 'a' is any constant.
- $E(aX+b)=a\,E(X)+b$, where a and b both are constants
- $E(X+Y)=E(X)+E(Y)$
- $E(X-Y)=E(X)-E(Y)$
- If X and Y are independent r.v's then

  $$E(XY)=E(X).\,E(Y)$$

Mathematical Expectation: Examples

Example: What is the mathematical expectation of the number of heads when 3 fair coins are tossed?

Solution: Here S={HHH, HHT, HTH, THH, HTT, THT, TTH, TTT}

Let X= number of heads then x=0,1,2,3

Then X has the following p.d.f:

(xi)	f(xi)
0	1/8
1	3/8
2	3/8
3	1/8

Mathematical Expectation: Examples

Example: What is the mathematical expectation of the number of heads when 3 fair coins are tossed?

Solution: Here S={HHH, HHT, HTH, THH, HTT, THT, TTH, TTT}

Let X= number of heads then x=0,1,2,3

Then X has the following p.d.f:

(xi)	f(xi)	x*f(x)
0	1/8	0
1	3/8	3/8
2	3/8	6/8=3/4
3	1/8	3/8
Total		**12/8**

$$E(X) = \sum xf(x) = \frac{12}{8} = \frac{3}{2} = 1.5,$$

Note: Since E(X)=1.5 which is not an integer, so we can say that When coin is tossed a large no of time then on average, we would get 1.5 heads

Mathematical Expectation: Examples

Example: If it rains, an umbrella salesman can earn \$30 per day. If it is fair, he can lose \$6 per day. What is his expectation if the probability of rain is 0.3?

Solution: Here, P(rain)=0.3, then P(no rain)=0.7

Let X= number of dollars the salesman earns.

Then X can take values, 30 and -6 with corresponding probabilities 0.3 and 0.7 respectively. Then X has the following p.d.f:

(xi)	f(xi)	x*f(x)
30	0.3	9
-6	0.7	-4.2
Total		**4.8**

$$E(X) = \sum xf(x) = 4.8 = \$4.8\, per\ day$$

Expectation of functions of two Random Variable

If X and Y are two random variables having joint p.m.f (p.d.f) $f_{x,y}(x,y)$ then the expectation of $\varphi(x,y)$ is given by

Discrete case:

$$E[\varphi(x,y)] = \sum_{i=1}^{m}\sum_{j=1}^{n}\varphi(x_i,y_j)f_{x,y}(x_i,y_i)$$

Continuous case:

$$E[\varphi(x,y)] = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\varphi\,(x,y)f_{x,y}(x,y)dxdy$$

Expectation of a Function of Random Variable

Let H(X) be a function of the r.v. X. Then H(X) is also a r.v. and also has an expected value (as any function of a r.v. is also a r.v.).

If X is a discrete r.v. with p.d f(x) then

$$E\left[H\left(X\right)\right]=H\left(x_1\right)f\left(x_1\right)+H\left(x_2\right)f\left(x_2\right)+\ldots+H\left(x_n\right)f\left(x_n\right)=\sum_i H\left(x_i\right)f\left(x_i\right)$$

If X is a continuous r.v. with p.d.f. f(x) then

$$E\left[H\left(X\right)\right]=\int_{-\infty}^{+\infty} H\left(x_i\right)f\left(x_i\right)dx$$

Expectation of a Function of Random Variable

We have $$E\left[H\left(X\right)\right]=\sum_i H\left(x_i\right) f\left(x_i\right)$$

- If H(X)=X², then $E\left(X^2\right)=\sum_i x_i^2 f\left(x_i\right)$
- If H(X)=Xᵏ, then $E\left(X^k\right)=\sum_i x_i^k f\left(x_i\right)=\mu_k'$

This is called 'k-th moment about origin of the r.v. X.

- If $H\left(X\right)=\left(X-\mu\right)^k$, then $\mu_k=E\left(X-\mu\right)^k=\sum_i\left(x_i-\mu\right)^k f\left(x_i\right)$

This is called 'k-th moment about Mean of the r.v. X

- **Variance:** $\mu_2=\sigma^2=E\left(X-\mu\right)^2=E\left(X^2\right)-\left[E\left(X\right)\right]^2$

Mathematical Expectation: Examples

Example: Let X be a r.v. with the following probability distribution.

Calculate E(X), E(X^2) and $E(2X+1)^2$

X=x	-3	6	9
P(X=x)	1/6	1/2	1/3

Solution:

$$E(X) = \sum xf(x)$$

$$E(X^2) = \sum x^2 f(x)$$

Mathematical Expectation: Examples

Example: Let X be a r.v. with the following probability distribution.

Calculate E(X), $E(X^2)$ and Var (2X+3)

X=x	-3	-2	-1	0	1	2	3
P(X=x)	0.05	0.10	0.30	0	0.30	0.15	0.10

Solution:

$$E(X) = \sum xf(x)$$

$$E(X^2) = \sum x^2 f(x)$$

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

Mathematical Expectation: Examples

Example: Let X be a r.v. with probability distribution:

Calculate E(X), E(X^2) and Var(X)

x	f(x)
-1	0.125
0	0.5
1	0.2
2	0.05
3	0.125

Solution:

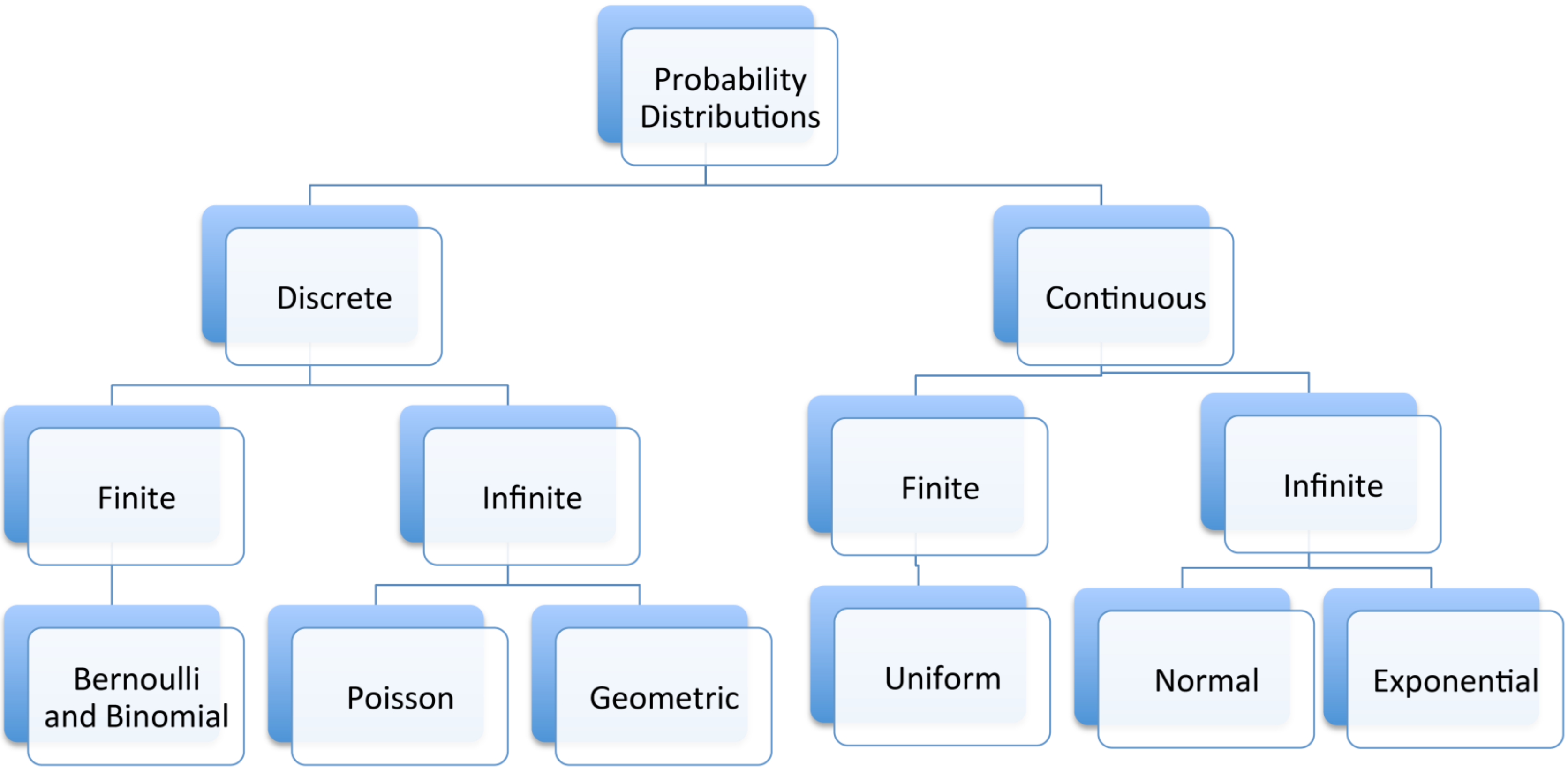
x	f(x)	x*f(x)	x^2*f(x)
-1	0.125	-0.125	0.125
0	0.5	0	0
1	0.2	0.2	0.2
2	0.05	0.1	0.2
3	0.125	0.375	1.125
Total=		**0.55**	**1.65**

$$E(X) = \sum xf(x) = 0.55$$

$$E(X^2) = \sum x^2 f(x) = 1.65$$

$\mathrm{Var}(X)=E(X^2)-[E(X)]^2=1.65-[0.55]^2=1.3475$

Some Common Distributions



Discrete Probability Distributions

Some important discrete Probability Distributions are:

- Bernoulli and Binomial Distribution

- Poisson Distribution

Bernoulli and Binomial Experiment

- Any trial of random experiment is called Bernoulli trail if it satisfy the following four properties:

- The outcome of each trial may be classified into one of two categories, conventionally called Success (S) and Failure (F). Usually the outcome of interest is called a success and the other, a failure.
- The probability of success, denoted by p, remains constant for all trials.
- The successive trials are all independent.
- The experiment is repeated a fixed number of times, say n.

The experiment having n Bernoulli trials is called **Binomial experiment**.

Binomial Probability Distribution

- When X denotes the number of successes in n trials of a binomial experiment, then it is called a **binomial random variable**.

- The probability distribution of a binomial random variable is called the **Binomial Probability Distribution**

- Let experiment repeat n-times. Then the probability of x-success in n-trail can be calculated by

$$P(X = x) = n_{C_x}\, p^x q^{n-x} = \binom{n}{x} p^x q^{n-x}, x = 0,1,2,\ldots,n$$

Where, q=1-p, is the probability of failure on each trial.

Cumulative Binomial Distribution

$$P\left(X \le r\right) = \sum_{x=0}^{r} \left\{ \binom{n}{x} p^{x} q^{n-x} \right\}$$

Binomial Distribution: Example

Example: A coin is tossed 5 times. Find the probabilities of obtaining various number of heads.

Solution: The random variable X which denotes the number of heads (successes) has a binomial probability distribution with p=1/2 and n=5.

- The possible values of X are: 0,1,2,3,4 and 5.

$$P(X=0)=\binom{5}{0}\left(\frac{1}{2}\right)^{0}\left(\frac{1}{2}\right)^{5-0}=\left(\frac{1}{2}\right)^{5}=\frac{1}{32}$$

$$P(X=1)=\binom{5}{1}\left(\frac{1}{2}\right)^{1}\left(\frac{1}{2}\right)^{5-1}=\frac{5}{32}$$

$$P(X=2)=\binom{5}{2}\left(\frac{1}{2}\right)^{2}\left(\frac{1}{2}\right)^{5-2}=\frac{10}{32}$$

$$P(X=3)=\binom{5}{3}\left(\frac{1}{2}\right)^{3}\left(\frac{1}{2}\right)^{5-3}=\frac{10}{32}$$

$$P(X=4)=\binom{5}{4}\left(\frac{1}{2}\right)^{4}\left(\frac{1}{2}\right)^{5-4}=\frac{5}{32}$$

$$P(X=5)=\binom{5}{5}\left(\frac{1}{2}\right)^{5}\left(\frac{1}{2}\right)^{5-5}=\frac{1}{32}$$

Binomial Distribution: Example

Example: Let X be a r.v. having binomial distribution with n=4 and p=1/3.
Find P(X=1), P(X=3/2), P(X<=2)

Solution: The binomial probability distribution for n=4 ad p=1/3 is:

$$P(X=x)=\binom{4}{x}\left(\frac{1}{3}\right)^{x}\left(\frac{2}{3}\right)^{4-x}, \qquad x=0,1,2,3,4$$

$$P(X=1)=\binom{4}{1}\left(\frac{1}{3}\right)^{1}\left(\frac{2}{3}\right)^{4-1}=\frac{32}{81}$$

$$P\left(X=\frac{3}{2}\right)=0$$

(b/c, a r.v. X with a binomial distribution takes only one of the integer values; 0,1,2,...,n)

Binomial Distribution: Example

Example: Let X be a r.v. having binomial distribution with n=4 and p=1/3.

Find P(X=1), P(X=3/2), P(X<=2)

Solution: The binomial probability distribution for n=4 and p=1/3 is:

$$P(X=x)=\binom{4}{x}\left(\frac{1}{3}\right)^{x}\left(\frac{2}{3}\right)^{4-x}, \qquad x=0,1,2,3,4$$

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

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

$$=\frac{16}{81}+\frac{32}{81}+\frac{24}{81}=\frac{72}{81}=\frac{8}{9}$$

Binomial Distribution: Example

Example: A coin is tossed 4-times, what is the probability to getting

1. Two heads
2. At least two heads

Solution:

Poisson Distribution

Poisson Distribution

Suppose we are given an interval (this could be time, length, area or volume) and we are interested in the finding the number of “successes" in that interval.

Assume that the interval can be divided into very small subintervals

Then r.v. X has a Poisson Distribution with parameter λ, which is given by:

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

The random variable X denotes the number of successes in the whole interval. λ is the mean number of successes in the interval. Where, e is a constant and approximately equal to 2.71828.

Poisson Distribution: Example

Example: If the r.v. X follows a Poisson distribution with mean 3.4, i.e. X~Po (3.4), Find P(X=6).

Solution: Note that we have:

$$P\left(X=x\right)=\frac{e^{-\lambda}\lambda^{x}}{x!},\qquad x=0,1,2,\ldots$$

Replacing x by 6 and λ by 3.4, we get:

$$P\left(X=6\right)=\frac{e^{-3.4}\left(3.4\right)^{6}}{6!}=0.072$$

Poisson Distribution: Example

Example: Find the Probability that at most 5 defective fuses will be found in a box of 200 fuses, if experience show that 2 percent of such fuses are defective.

Solution:

Poisson Distribution: Example

Example: The number of industrial injuries per working week in a particular factory is known to follow a Poisson distribution with mean 0.5. Find the probability that in a particular week there will be:

(i). Less than 2 accidents; **(ii).** More than 2 accidents;

Solution: Let X= The number of accidents in one week, so X~Po (0.5),

(i). Less than 2 accidents

$$P(X<2)=P(X=0)+P(X=1)$$
$$=\frac{e^{-0.5}(0.5)^0}{0!}+\frac{e^{-0.5}(0.5)^1}{1!}$$
$$=0.9098$$

(ii). More than 2 accidents

$$P(X>2)=1-P(X\le 2)$$
$$=1-\{P(X=0)+P(X=1)+P(X=2)\}$$
$$=1-\left(\frac{e^{-0.5}(0.5)^0}{0!}+\frac{e^{-0.5}(0.5)^1}{1!}+\frac{e^{-0.5}(0.5)^2}{2!}\right)$$
$$=1-0.9856=0.0144$$

Summary

- Mathematical Expectation of random variables
- Bernoulli and Binomial Distribution
- Poisson Distribution

Introduction to Statistics

!