Lecture No.7

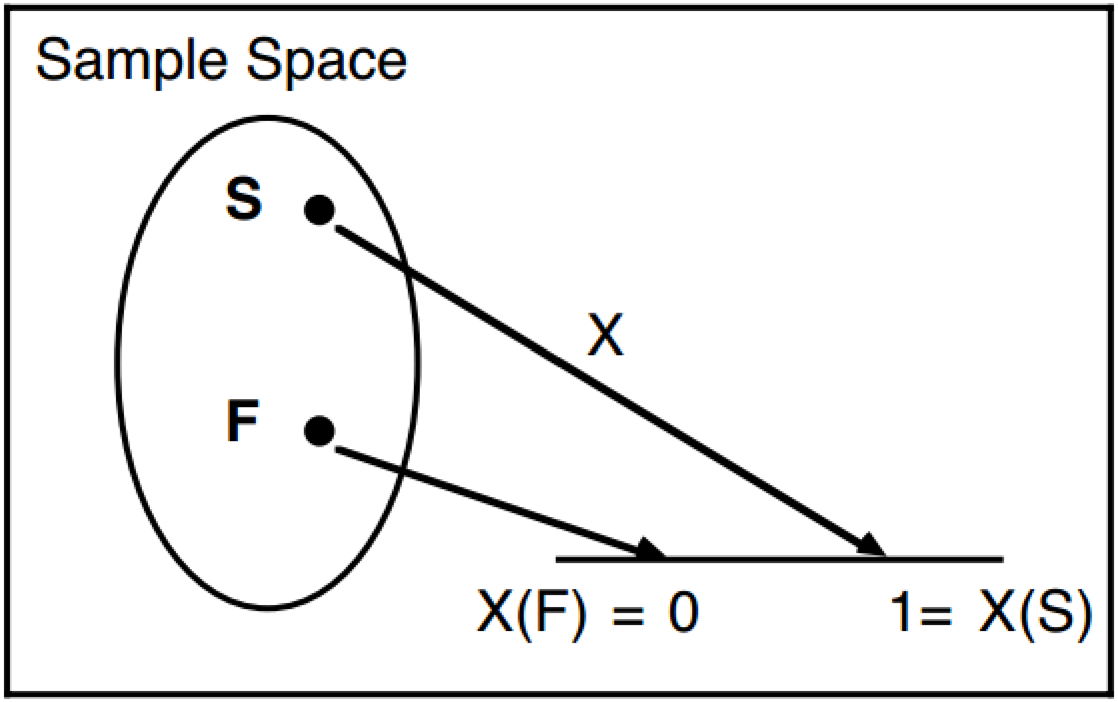
SPECIAL DISCRETE DISTRIBUTIONS

Azeem Iqbal

University of Engineering and Technology, Lahore
(Faisalabad Campus)

Bernoulli Distribution

- A **Bernoulli trial** is a random experiment in which there are precisely two possible outcomes, which we conveniently call ‘failure’ (F) and ‘success’ (S).



Bernoulli Distribution

- The probability density function of this random variable is

$$p(1) = P(X = 1) = p$$

$$p(0) = P(X = 0) = 1 - p$$

- Where p denotes the probability of success. Hence

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

We denote the Bernoulli random variable by writing $X \sim BER(p)$

Bernoulli Distribution

What is the probability of getting a score of not less than 5 in a throw of a six-sided die?

Answer: Although there are six possible scores $\{1, 2, 3, 4, 5, 6\}$, we are grouping them into two sets, namely $\{1, 2, 3, 4\}$ and $\{5, 6\}$. Any score in $\{1, 2, 3, 4\}$ is a failure and any score in $\{5, 6\}$ is a success. Thus, this is a Bernoulli trial with

$$P(X = 0) = P(\text{failure}) = \frac{4}{6} \qquad \text{and} \qquad P(X = 1) = P(\text{success}) = \frac{2}{6}.$$

Hence, the probability of getting a score of not less than 5 in a throw of a six-sided die is $\frac{2}{6}$.

Expected Value of Bernoulli Distribution

$$\mu_X = \sum_{x=0}^{1} x \cdot p_X(x)$$

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

$$= \sum_{x=0}^{1} x \cdot p^x(1-p)^{1-x}$$

$$= (0)p^0(1-p)^{1-0} + (1)p(1-p)^{1-1}$$

$$= p$$

Variance of Bernoulli Distribution

$$\sigma_X^2 = \sum_{x=0}^{1} (x - \mu_X)^2 \cdot p_X(x)$$

$$= \sum_{x=0}^{1} (x - p)^2 \, p^x (1 - p)^{1-x}$$

$$\mu_X = p$$

$$= p^2(1 - p) + p(1 - p)^2$$

$$= p(1 - p)[p + (1 - p)] \quad = p(1 - p)$$

Binomial Distribution

- Consider a fixed number n of mutually independent Bernoulli trails.
- Suppose these trials have same probability of success, say p.
- A random variable $\boldsymbol{X}$ is called a **binomial random variable** if it represents the total number of successes in n independent Bernoulli trials.
- To find the probability density function of $\boldsymbol{X}$ we have to find the probability of $\boldsymbol{x}$ successes in $\boldsymbol{n}$ independent trails.

Binomial Distribution

- If we have x successes in $\boldsymbol{n}$ trails, then the probability of each n-tuple with x successes and $\boldsymbol{n} - \boldsymbol{x}$ failures is

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

- We will denote a binomial random variable with parameters $\boldsymbol{p}$ and $\boldsymbol{n}$ as $\boldsymbol{X \sim BIN(n, p)}$.

Practice Problem

On a five-question multiple-choice test there are five possible answers, of which one is correct. If a student guesses randomly and independently, what is the probability that she is correct only on questions 1 and 4?

Here the probability of success is $p = 1/5$, and thus $1 - p = 4/5$.
Therefore, the probability that she is correct on questions 1 and 4 is

$$
\begin{aligned}
P(\text{correct on questions 1 and 4}) &= p^2(1-p)^3 \\
&= \left(\frac{1}{5}\right)^2 \left(\frac{4}{5}\right)^3 \\
&= \frac{64}{5^5} = 0.02048.
\end{aligned}
$$

Practice Problem

On a five-question multiple-choice test there are five possible answers, of which one is correct. If a student guesses randomly and independently, what is the probability that she is correct only on two questions?

Here the probability of success is $\boldsymbol{p = 1/5}$, and thus $\boldsymbol{1 - p = 4/5}$.

There are $\binom{5}{2}$ different ways she can be correct on two questions.

Therefore, the probability that she is correct on two questions is

$$
\begin{aligned}
P(\text{correct on two questions}) &= \binom{5}{2} p^2 (1-p)^3 \\
&= 10 \left(\frac{1}{5}\right)^2 \left(\frac{4}{5}\right)^3 \\
&= \frac{640}{5^5} = 0.2048.
\end{aligned}
$$

Expected Value and Variance of Binomial RV

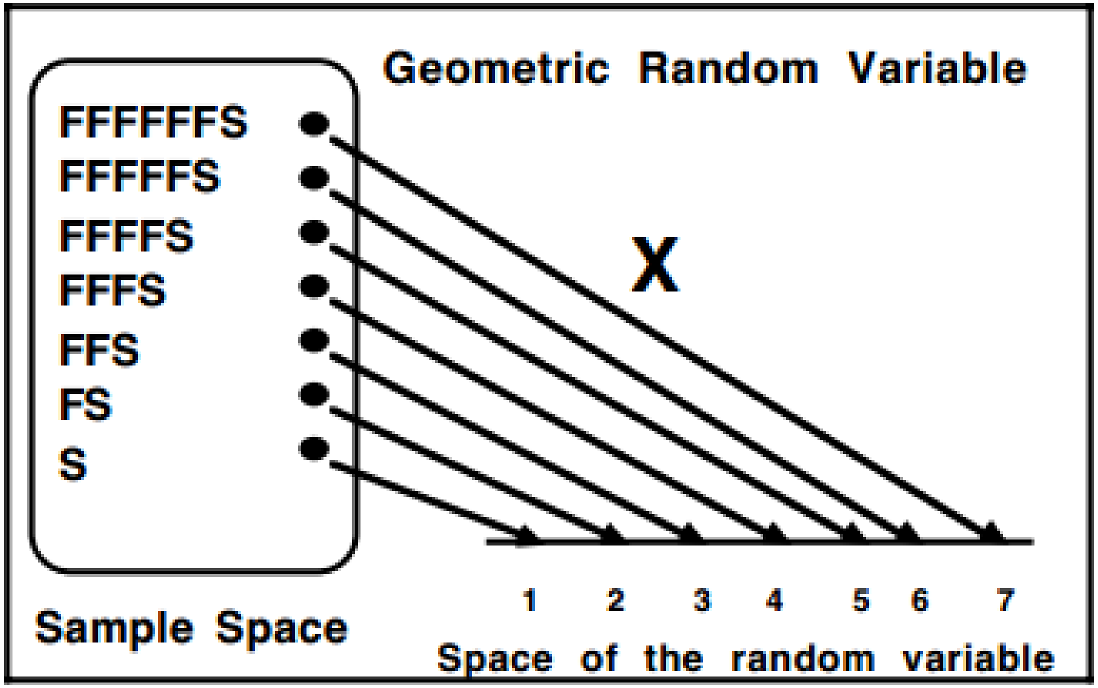
- If $\boldsymbol{X}$ is binomial random variable with parameters $\boldsymbol{p}$ and $\boldsymbol{n}$, then the mean and variance are respectively given by

$$E[X] = np$$

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

Geometric Distribution

- Let **X** denote the trial number on which the first success occurs.



Geometric Distribution

- Hence the probability that the first success occurs on x^{th} trial is given by.

$$p(x) = p(X = x) = (1-p)^{x-1} \cdot p$$

- Hence, the probability density function of X is

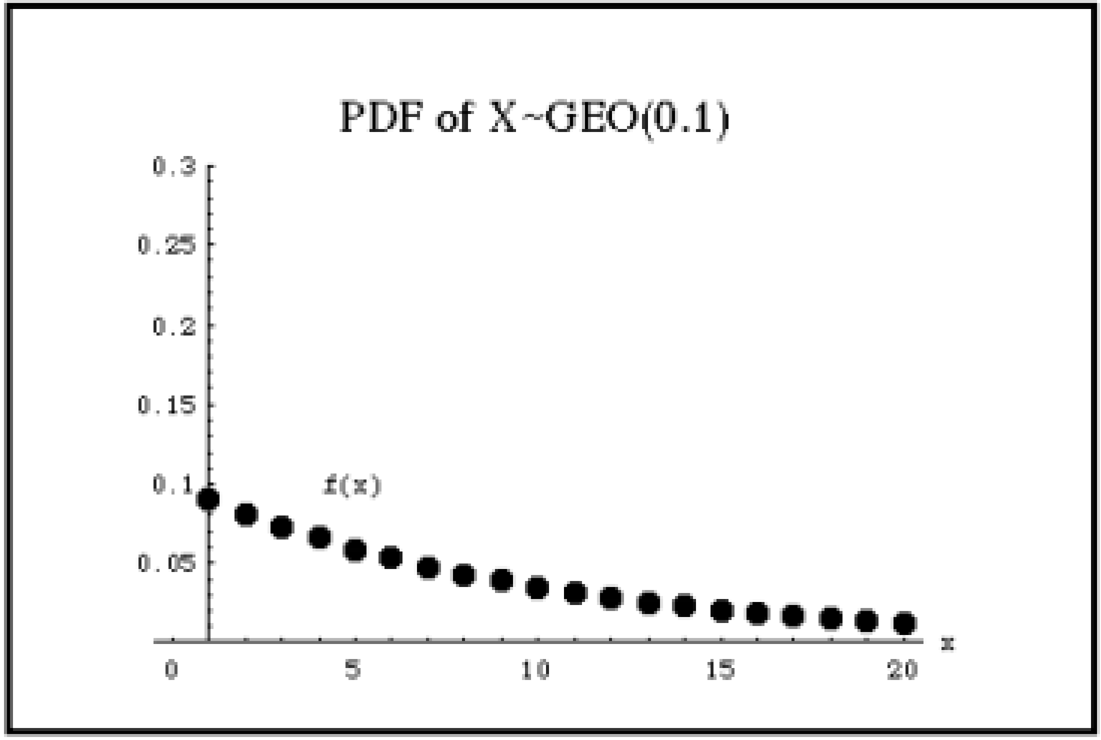
$$p(x) = (1-p)^{x-1} \cdot p \quad , x = 1, 2, 3 \dots , \infty$$

- where p denotes the probability of success in a single Bernoulli trial.

Geometric Distribution

- If X has a geometric distribution we denote it as

$X \sim GEO(p)$.



Geometric Distribution

- Is the real valued function $\boldsymbol{p}(\boldsymbol{x})$ is valid pmf defined by

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

- For a pmf to be valid $\sum_{x=1}^{\infty} p(x) = 1$

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

$$= p \sum_{y=0}^{\infty} (1-p)^{y} , \qquad where\ y = x - 1$$

$$= p \frac{1}{1-(1-p)}$$

$$= 1$$

Practice Problem

The probability that a machine produces a defective item is 0.02. Each item is checked as it is produced. Assuming that these are independent trials, what is the probability that at least 100 items must be checked to find one that is defective?

Solution

- Let X denote the trial number on which the first defective item is observed. We want to find

$$\begin{aligned}
P(X \geq 100) &= \sum_{x=100}^{\infty} f(x) \\
&= \sum_{x=100}^{\infty} (1-p)^{x-1}\, p \\
&= (1-p)^{99} \sum_{y=0}^{\infty} (1-p)^{y}\, p \\
&= (1-p)^{99} \\
&= (0.98)^{99} = 0.1353.
\end{aligned}$$

- Hence the probability that at least 100 items must be checked to find one that is defective is 0.1353.

Expected Value and Variance of Geometric RV

- If $\boldsymbol{X}$ is Geometric RV variable with parameters $\boldsymbol{p}$, then the mean and variance are respectively given by

$$E[X] = \boldsymbol{\mu_X} = \frac{1}{p}$$

$$Var[X] = \boldsymbol{\sigma_X^2} = (1-p)/\mathrm{p}^2$$

Poisson Distribution

Poisson Distribution

- Experiments yielding numerical values of a random variable X, the number of outcomes occurring during a given time interval or in a specified region, are called **Poisson experiments**.

- The given time interval may be of any length, such as a minute, a day, a week, a month, or even a year.

Poisson Distribution

- Hence a Poisson experiment can generate observations for the random variable X representing the

  - Number of telephone calls per hour received by an office.
  - Number of days school is closed due to snow during the winter.
  - Number of postponed games due to rain during a baseball season.

Poisson Distribution

- **Poisson Distribution** is an important discrete distribution which is widely used for modeling many real life situations. First, we define this distribution and then we present some of its important properties.
- **Definition**: A random variable X is said to have a Poisson distribution if its probability density function is given by

$$p(x) = \frac{e^{-\lambda}\lambda^{x}}{x!}$$

Poisson Distribution

- Where $0<\lambda<\infty$ is a parameter. We denote such a random variable by $X \sim POI(\lambda)$.

- If $X \sim POI(\lambda)$, then

$$E[X] = \mu_X = \lambda$$

$$Var[X] = \sigma_X^2 = \lambda$$

Practice Problem

The number of hits at a Web site in any time interval is a Poisson random variable. A particular site has on average $\lambda = 2$ hits per second.

- What is the probability that there are no hits in an interval of 0.25 seconds?

- What is the probability that there are no more than two hits in an interval of one second?

Practice Problem

- What is the probability that there are no hits in an interval of 0.25 seconds?

In an interval of **0.25 seconds**, the number of hits H is a Poisson random variable with

$$\boldsymbol{\lambda = (2\,hits/s) \times (0.25\,s) = 0.5\,hits.}$$

The PMF of H is

$$P_H(h) = \frac{0.5^h e^{-0.5}}{h!}, \qquad h = 0,1,2,\ldots.$$

Practice Problem

The probability of no hits is

$$P_H(0) = \frac{0.5^0 e^{-0.5}}{0!} = 0.607$$

Practice Problem

- What is the probability that there are no more than two hits in an interval of one second?

In an interval of **1 seconds**, the number of hits J is a Poisson random variable with

$$\boldsymbol{\lambda = (2\ hits/s) \times (1\ s) = 2 hits.}$$

The PMF of J is

$$P_J(j) = \frac{2^j e^{-2}}{j!}, \qquad j = 0,1,2, \dots .$$

Practice Problem

- What is the probability that there are no more than two hits in an interval of one second?

$$P\,[J \le 2] = P\,[J = 0] + P\,[J = 1] + P\,[J = 2]$$
$$= P_J\,(0) + P_J\,(1) + P_J\,(2)$$
$$= e - 2 + 2^1 e^{-2}/1! + 2^2 e^{-2}/2! = 0.677$$

Homework Problem

Suppose that the number of inquiries arriving at a certain interactive system follows a Poisson distribution with arrival rate of 12 inquiries per minute.

Find the probability of 10 inquiries arriving

a) in a 1-minute interval; **(Answer=0.1048)**
b) in a 3-minute interval. **(Answer=0.000002337)**
c) What is the expectation and the variance of the number of arrivals during each of these intervals? **(Answer= 12 and 36)**

a) Let X be the number of inquiries in a 1-min interval. Then X is a Poisson random variable with the parameter 12.

$$P(X = 10) = \frac{e^{-12} \cdot 12^{10}}{10!} = 0.1048$$

b) Let X be the number of inquiries in a 3-min interval. Then X is a Poisson random variable with the parameter 12*3=36.

$$P(X = 10) = \frac{e^{-36} \cdot 36^{10}}{10!} = 0.0000002337$$

c) What is the expectation and the variance of the number of arrivals during each of these intervals?

c. For Poisson distribution with parameter λ, we have $E(X) = Var(X) = \lambda$.

Therefore, for a 1-min interval, $E(X) = Var(X) = 12$, and for a 3-min interval, $E(X) = Var(X) = 36$.