**Remark:**

VAR[X] is a nonlinear operator.
However, E[X] is a linear operator.
Therefore,

- $E[X+X] = E[2X]$
$= 2E[X]$

- $\sigma_{X+X}^2 = \sigma_{2X}^2$
$= 4\sigma_X^2$.

## PROBABILITY DISTRIBUTIONS For Continuous Random Variables

For a continuous random variable, we are unable to assign a probability to sample point or experimental outcome in the same manner as we did for the discrete random variable, because the range contains an uncountably infinite number of values.

ARMA 2-D LATTICE

Probability and statistics

2

## Definition: Probability Density Functions (pdf)

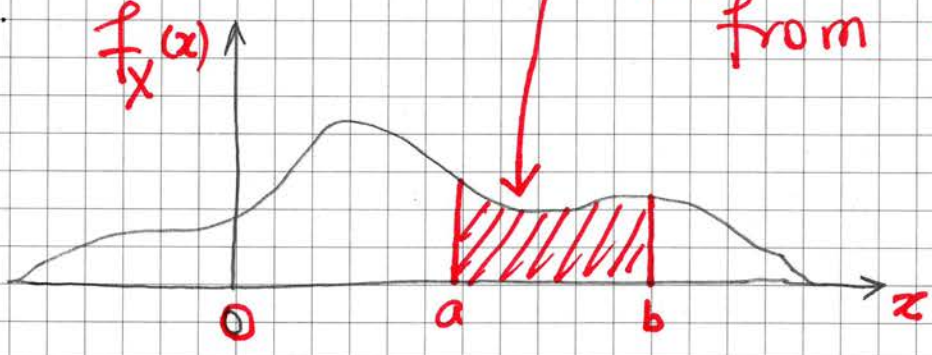
For a continuous random variable $X$, probability density function (pdf), $f_X(x)$ is a function such that

(1) $f_X(x) \geq 0$

(2) $\displaystyle\int_{-\infty}^{\infty} f_X(x)\,dx = 1$

(3) $\displaystyle Pr\{a \leq X \leq b\} = \int_a^b f_X(x)\,dx$

= Area under $f_X(x)$ from a to b.




## Warnings:

- For a notational simplicity, the sub index $X$ in the $f_X(x)$ can be dropped. Therefore, we will use $f(x)$ instead of $f_X(x)$.

## Important Remark:

- For a continuous random variable X and any value x,

$$Pr\{X=x\}=0.$$

Because every point has zero width, and the area at any point is zero.

- Since each point has zero probability, one need distinguish between inequalities such as < or ≤ for continuous random variables.
- If X is a continuous random variable, for any $x_1$ and $x_2$,

$$\begin{aligned} Pr\{x_1 \le X \le x_2\} &= Pr\{x_1 < X \le x_2\} \\ &= Pr\{x_1 \le X < x_2\} \\ &= Pr\{x_1 < X < x_2\}. \end{aligned}$$

An alternative method to describe the probability distribution of a continuous random variable is cumulative distribution function (cdf).

Defintion:

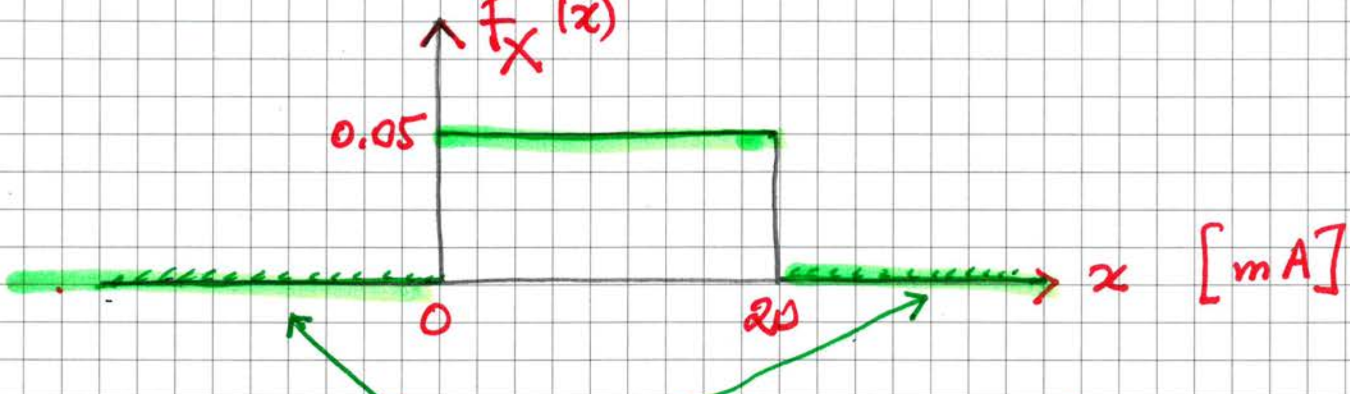
Cumulative Distribution Function (cdf) is defined as follows:

$$F_X(x) = \Pr\{X \leq x\}$$

$$= \int_{-\infty}^{x} f_X(u)\,du .$$

Example:

- Let us consider the random variable $X$ is the current measured in a copper wire in mili ampers.
- We assume that the range of $X$ is $[0, 20\,mA]$ and pdf is given as follows:




It is assumed that $f_X(x)=0$ whenever

- What is the probability that the current measurement is less than 10 mA?

$$\Pr\{X < 10\} = \int_0^{10} f_X(x)\,dx$$

$$= \int_0^{10} 0.05\,dx = 0.5$$

- Another example,

$$\Pr\{5 < X < 20\} = \int_5^{20} f_X(x)\,dx$$

$$= 0.75$$

- What is the cdf of this random variable?

We observe that the variable X consists of three expressions:

- If $x < 0$, $f_X(x) = 0$, Therefore,

$$F_X(x) = 0 \quad , \text{ for } x < 0 .$$

and

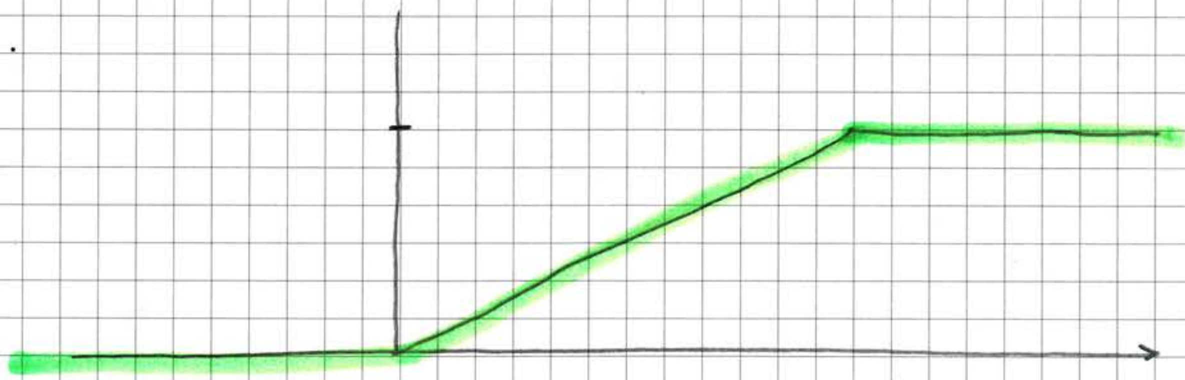
- $F_X(x) = \int_0^x f_X(u)\,du = 0.05x$ , for $0 \le x < 20$

Finally,

$$F_X(x) = \int_0^x f_X(u)\,du = 1 \quad \text{for} \quad x \geq 20 .$$

Therefore,

$$F_X(x) = \begin{cases} 0 & ; \quad x < 0 \\ 0.05x & ; \quad 0 \leq x < 20 \\ 1 & ; \quad x \geq 20 . \end{cases}$$



## Properties of the Cdf:

- From the definition of cdf, we can state the following properties:

(1) $0 \leq F_X(x) \leq 1$ , for $-\infty < x < \infty$

(2) $\lim_{x \to -\infty} F_X(x) = 0$ and $\lim_{x \to +\infty} F_X(x) = 1$ .

(3) $F_X(x_1) \leq F_X(x_2)$ for $x_1 \leq x_2$.

Remark:

- The pdf $f_X(x)$ and the cdf $F_X(x)$ are related with each other;

$$f_X(x) = \frac{d}{dx} F_X(x)$$

$$F_X(x) = \int_{-\infty}^{x} f_X(t)\,dt.$$

- For a discrete random variable $X$,

$$P_X(x_i) = F_X(x_i) - F_X(x_{i-1})$$

$$F_X(x) = \sum_{x_i \leq x} P_X(x_i).$$

We can also define the mean and the variance of a continuous random variable $X$.

## Definition:

Suppose $X$ is a continuous random variable with a pdf $f_X(x)$. The mean or expected value of $X$, denoted as $\mu_X$ or $E[X]$, is

$$\mu_X = E[X] = \int_{-\infty}^{\infty} x \, f_X(x) \, dx .$$

The variance of $X$, denoted as $VAR[X]$ or $\sigma_X^2$, is

$$\sigma_X^2 = VAR[X] = \int_{-\infty}^{\infty} (x-\mu)^2 f_X(x) \, dx = \int_{-\infty}^{\infty} x^2 f_X(x) \, dx - \mu_X^2$$

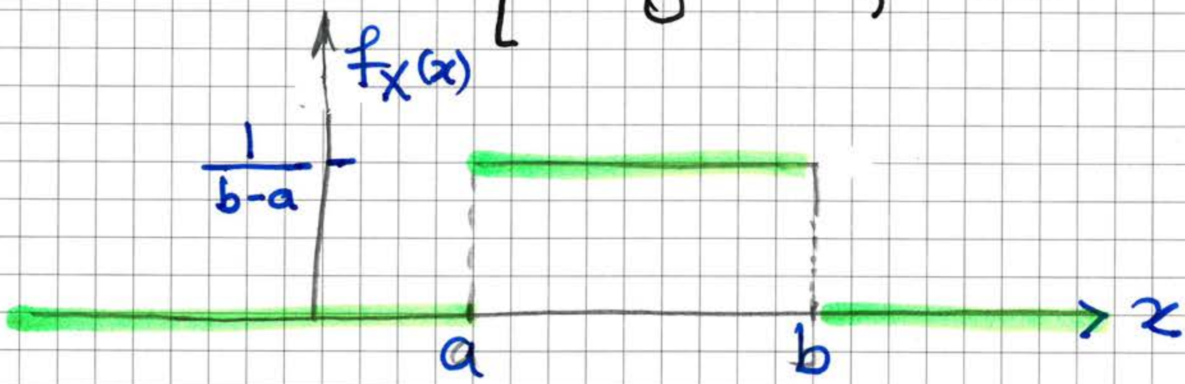
- The standard deviation of $X$ is

$$\sigma_X = +\sqrt{\sigma_X^2} .$$

Example:

Find the mean and the variance of the random variable $X$, which has a pdf function as follows:

$$f_X(x) = \begin{cases} \dfrac{1}{b-a}, & \text{for } a \le x \le b \\ 0, & \text{elsewhere} \end{cases}$$




· the mean and variance of $X$ are easily found to be

$$\mu_X = \int_a^b x f_X(x)$$

$$= \frac{1}{b-a} \int_a^b x\,dx = \frac{a+b}{2};$$

$$\sigma_X^2 = \frac{1}{b-a} \int_a^b \left(x - \frac{a+b}{2}\right)^2 dx$$

$$= \frac{(b-a)^2}{12}.$$

# CONDITIONING A CONTINUOUS RANDOM VARIABLE

## Definition:

### Conditional pdf given an event

For a random variable $X$ with pdf $f_X(x)$ and an event $A \subset S_X$ with $Pr\{A\} \geq 0$, The conditional pdf of $X$ given an event $A$ is

$$f_{X|A}(x) = \begin{cases} \dfrac{f_X(x)}{Pr\{A\}} & ; \ x \in A \\ 0 & ; \ x \notin A \ \text{(otherwise)} \end{cases}$$

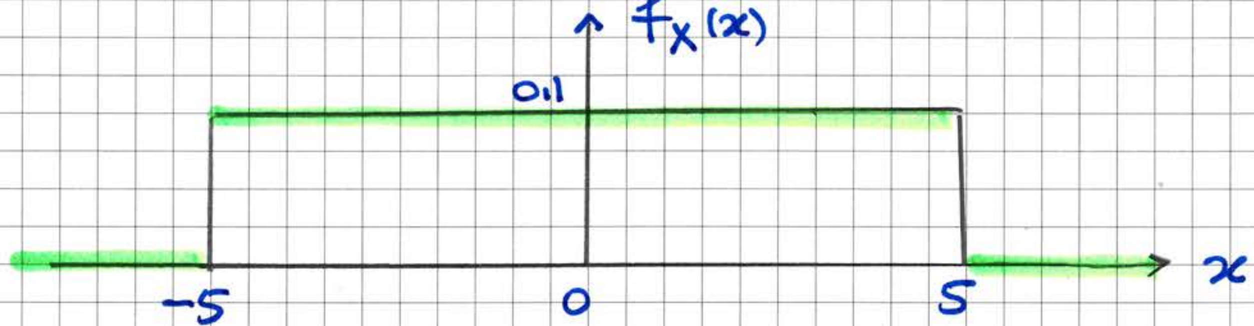
- The function $f_{X|A}(x)$ is a probability model for a new random variable related to $X$. Thus it has the same properties as any $f_X(x)$.
  - For example, the integral of the conditional pdf over all $x$ is 1.

$$\int_{-\infty}^{\infty} f_{X|A}(x)\, dx = 1.$$

## Example:

Let us assume $X$ has a uniform pdf over $[-5, 5]$

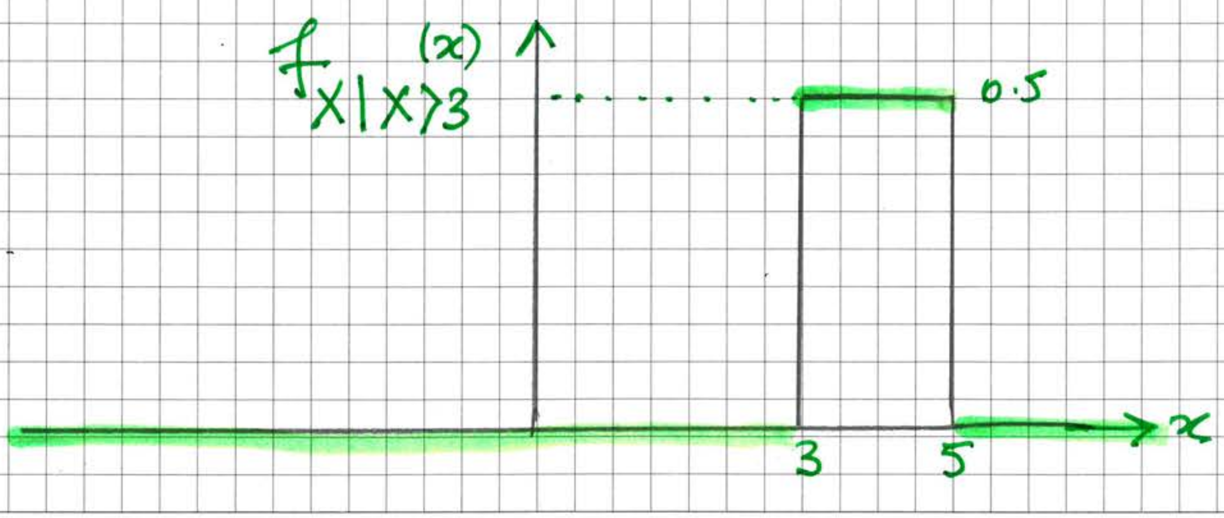
$$f_X(x) = \begin{cases} 0.1 & ; \ -5 \le x \le 5 \\ 0 & ; \ \text{otherwise} \end{cases}$$




$$Pr\{X > 3\} = \int_3^5 (0.1)\,dx = 0.2$$

and

$$f_{X|X>3}(x) = \begin{cases} \dfrac{f_X(x)}{Pr\{X>3\}} = \dfrac{0.1}{0.2} = 0.5 & ; \ 3 \le x \le 5 \\ 0 & ; \ \text{otherwise} \end{cases}$$



Remark:

The conditional pmf of a discrete random variable X can be defined similarly,

$$p_{X|A}(x_i) = \begin{cases} \dfrac{p_X(x_i)}{\Pr\{A\}} , & x_i \in A \\ 0 , & x_i \notin A . \end{cases}$$

## MOMENTS and MOMENT-GENERATING FUNCTIONS

While a probability distribution,

$$F_X(x),\ p_X(x) \text{ or } f_X(x)$$

contains a complete description of a random variable X, we are often interested in seeking a set of simple numbers that gives some of the dominat features of the random variable.

Let us define the momonts of a random variable X:

**Definition:** The expectation $E[X^k]$, when it exists, is called the n-th moment of X. It is given by

$$E[X^k] = \begin{cases} \sum_i x_i^k P_X(x_i) & \text{, for } X \text{ discrete} \\ \int_{-\infty}^{\infty} x^k f_X(x)\,dx & \text{, for } X \text{ continuous} \end{cases}$$

where k is a positive integer, $k = 1, 2, \ldots$

**Remark:**

- One of the most important moments is $E[X]$, the first moment. Using the mass analogy for the probability distribution, the first moment may be regarded as the center of mass of its distribution
- The first moment of X is called the mean, expectation value, or average value of X.
- Common notation for it is $\mu_X$ or $m_X$, or simply $\mu$ or $m$.

Now, we can define a moment generating function, $M_X(t)$ in order to compute all moments

$$E[X^k], \text{ for } k=1,2,\ldots$$

## Definition:

Given a random variable $X$, the moment-generating function $M_X(t)$ is the expected value $e^{tX}$ and expressed mathematically

$$M_X(t) = E\left[e^{tX}\right]$$

$$= \begin{cases} \int_{-\infty}^{\infty} e^{tx} f_X(x)\,dx, & X \text{ continuous} \\ \sum_k e^{tx_k} P_X(x_k), & X \text{ discrete.} \end{cases}$$

## Warning!

For certain probability distributions, the moment-generating function may not exist for all real values of $t$. However, in this lecture, we will consider the probability distributions $[F_X(x), P_X(x_i) \text{ or } f_X(x)]$ that their moment-generating function always exist.

## Example:

The probability mass function (pmf) of a Bernoulli trail is given as

$$p_X(x_i) = \begin{cases} p & , \ x=1 \\ 1-p & , \ x=0 \\ 0 & , \ \text{otherwise}. \end{cases}$$

Its moment generating function is

$$M_X(t) = E\left[e^{tX}\right]$$

$$= \sum_i e^{tx_i} p_X(x_i)$$

$$= e^{t \cdot 0} p_X(0) + e^{t \cdot 1} p_X(1)$$

$$= (1-p) + pe^t .$$

- In order to show the relationship between the moment-generating function and the moments, we can expand $e^{tX}$ as a power series in $t$, we obtain,

$$e^{tX} = 1 + tX + \frac{t^2X^2}{2!} + \frac{t^3X^3}{3!} + \cdots$$

- On taking expectations of the both side, we see that

$$M_X(t) = E\left[e^{tX}\right]$$
$$= 1 + E[X]t + E[X^2]\frac{t^2}{2!} + E[X^3]\frac{t^3}{3!} + \cdots .$$

- $M_X(t)$ is written as a power series in $t$; the coefficient of $t^k/k!$ in the expansion is the kth moment about the origin, $E[X^k]$.

- Therefore, the kth derivative of $M_X(t)$ with respect to $t$, evaluated at $t=0$, is just

$$\left.\frac{d^k}{dt^k} M_X(t)\right|_{t=0} = E\left[X^k\right] \quad \text{for } k = 1, 2, \ldots$$

gives the kth moment of $X$.

- <u>Properties of the Moment-Generating Function:</u>

- If $a$ is constant,

$$M_{X+a} = e^{at} M_X(t) \quad \text{and} \quad M_{aX}(t) = M_X(at) .$$

- If $X_1, X_2, \ldots, X_n$ are independent random variables and their moment-generating functions are $M_{X_1}(t), M_{X_2}(t), \ldots, M_{X_n}(t)$, respectively,

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

the moment generating function $M_Y(t)$ is

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

## Remark:

- Unfortunately, a random variable may not have a moment-generating function. For this reason, we define the expectation $E[e^{jtX}]$ of a random variable $X$ as the characteristic function of $X$. Denoted by $\phi_X(t)$ and always exists.

## CHARACTERISTIC FUNCTIONS

$$\phi_X(t) = E[e^{jtX}] = \begin{cases} \int_{-\infty}^{\infty} e^{jtx} f_X(x)\,dx, & X \text{ continuous} \\ \sum_i e^{jtx_i} p_X(x_i), & X \text{ discrete} \end{cases}$$

where $t$ is an arbitrary real-valued parameter and $j=\sqrt{-1}$. $\phi_X(t)$ is the expectation of a complex function and is generally complex valued.

- Since

$$|e^{jtX}| = |\cos tX + j\sin tX| = 1,$$

the sum and the integral in the above equations exist and therefore $\phi_X(t)$ always exist.

- Similar to the Fourier transform, $f_X(x)$ pdf can be obtained

$$f_X(x) = \frac{1}{2\pi}\int_{-\infty}^{\infty} \phi_X(t)\, e^{-jtx}\, dx$$

(look at the textbook).

- Furthermore, we note

$$\phi_X(0) = 1$$

$$\phi_X(-t) = \phi_X^*(t)$$

$$|\phi_X(t)| \leq 1 .$$

## Generation of Moments:

- If $t$ is changed to $jt$ at the moment-genarating function, the resulting relation is the characteristic function of $X$:

$$\phi_X(t) = M_X(jt) = E\left[e^{jtX}\right].$$

Then,

$$\phi_X(t) = \phi_X(0) + \phi_X'(0)\,t + \phi_X''(0)\frac{t^2}{2!} + \cdots$$

where the primes denotes derivatives.

and

$$E[X^k] = \frac{1}{j^k}\phi_X^{(k)}(0)$$

$$= \frac{1}{j^k}\frac{d^k}{dt^k}\phi_X(t)\Big|_{t=0}.$$

**Example:** If $X$ is exponentially distributed random variable, its pdf is given as

$$f_X(x) = \begin{cases} a e^{-ax} & , \text{ for } x \geq 0 \\ 0 & , \text{ elsewhere.} \end{cases}$$

Its characteristic function $\phi_X(t)$ is

$$\phi_X(t) = E\left[e^{jtX}\right]$$

$$= \int_0^\infty e^{jtx} (a e^{-ax}) \, dx$$

$$= a \int_0^\infty e^{-(a-jt)x} \, dx$$

$$= \frac{a}{a-jt} .$$

- The moments are

$$E[X] = \frac{1}{j} \frac{d}{dt} \left( \frac{a}{a-jt} \right) \Big|_{t=0} = \frac{1}{j} \left[ \frac{ja}{(a-jt)^2} \right] \Big|_{t=0} = \frac{1}{a} ,$$

$$E[X^2] = \frac{1}{j^2} \frac{d^2}{dt^2} \left( \frac{a}{a-jt} \right) \Big|_{t=0} = \frac{2}{a^2} ,$$

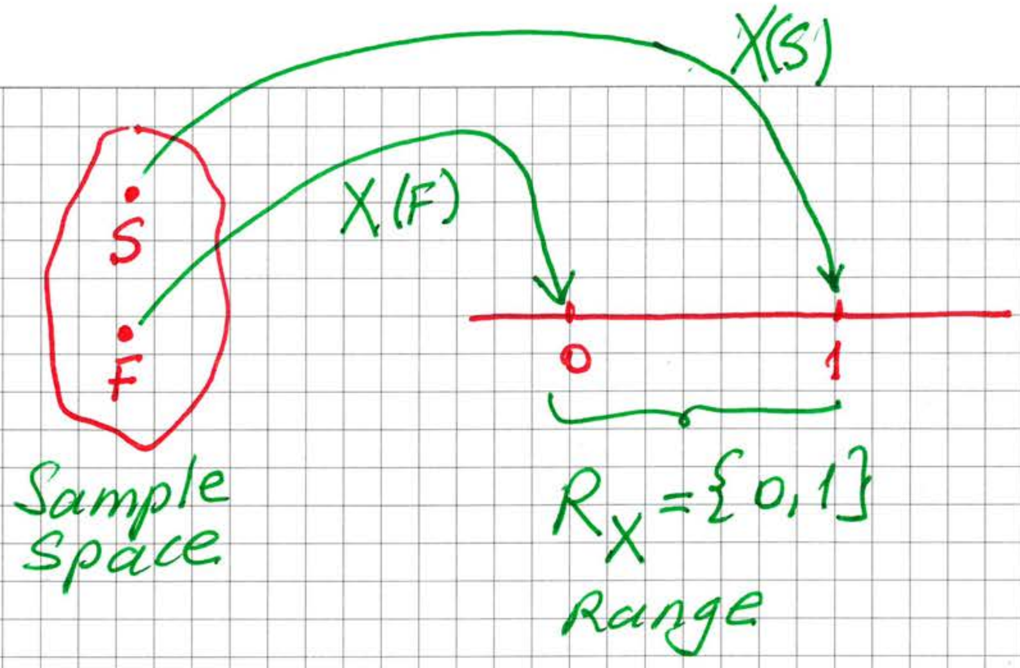
$$\sigma_X^2 = E[X^2] - E[X]^2 = \frac{2}{a^2} - \frac{1}{a^2} = \frac{1}{a^2} .$$

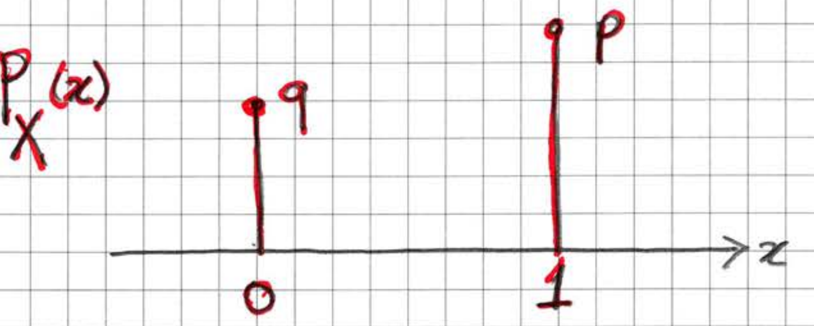
# SOME IMPORTANT DISCRETE DISTRIBUTIONS

- We will deal some distributions of discrete random variables that are important as models of scientific phenomena.
- The nature and applications of these distributions will be discussed.

## 1. Bernoulli Trails (or Experiments)

- A sequence of trails is performed so that
    - For each trail, there are only two possible outcomes, say, success or failure.
    - The probabilities of the occurrene of these outcomes remain the same throughout the trails;
    - Trails are carried out independently.
- Trails performed a under these conditions are named as Bernoulli trails.
- Let us denote event "success" by $S$, and the event "failure" by $F$. Also, let $Pr\{S\} = p$ and $pr\{F\} = q$, where $p+q=1$.




Bernoulli distribution

The probability mass function (pmf)

$p + q = 1.$

$$P_X(x) = \begin{cases} p \; ; & x = 1 \\ q \; ; & x = 0 \\ 0 \; ; & \text{otherwise} \end{cases}$$

- **Bernoulli Process**
  - Possible outcomes resulting from performing a sequence of Bernoulli trails can be represented by

    SSFFSFSSS ... FF

    FSFSSFFFS ... SF
  - The $n$ Bernoulli trails $\mathcal{E}_1, \mathcal{E}_2, \ldots, \mathcal{E}_n$ are called a Bernoulli process. Here we assume

that each trail has only two possible outcomes, Say S and F, and trails are independent.

- The probabilities of those possible outcomes are easily computed. For example, $(q \triangleq 1-p)$

$$Pr\{SSFFSF\cdots FF\} = Pr\{S\}Pr\{S\}Pr\{F\}Pr\{F\}\cdots Pr\{F\}Pr\{F\}$$

$$= ppqqpq\cdots qq.$$

- In general, the joint probability mass function is,

$$P_X(x_1, x_2, \ldots, x_n) = Pr\{x_1\}Pr\{x_2\}\cdots Pr\{x_n\}.$$

- The Mean of $X \sim$ Bernoulli $(p)$

$$\mu_X = E[X] = \sum_{i=1}^{2} x_i P_X(x_i)$$

$$= 0 \cdot q + 1 \cdot p = p.$$

- The Variance:

$$\sigma_X^2 = VAR[X] = E[(X-\mu_X)^2]$$

$$= E[(X-p)^2] = \sum_i (x_i - p)^2 P_X(x_i)$$

$$\sigma_X^2 = \sum_i x_i^2 \, p_X(x_i) - \mu_X^2$$

$$= [0^2 \cdot q + 1^2 p^2] - p^2$$

$$= p - p^2 = p(1-p) = pq.$$

## The Moment-Genering Function:

$$M_X(t) = E[e^{tX}]$$

$$= \sum e^{tx_i} \, p_X(x_i)$$

$$= e^{t \cdot 0} q + e^{t \cdot 1} p = q + pe^t.$$

## Characteristic Function:

$$\Phi_X(t) = E[e^{jtX}]$$

$$= M_X(jt)$$

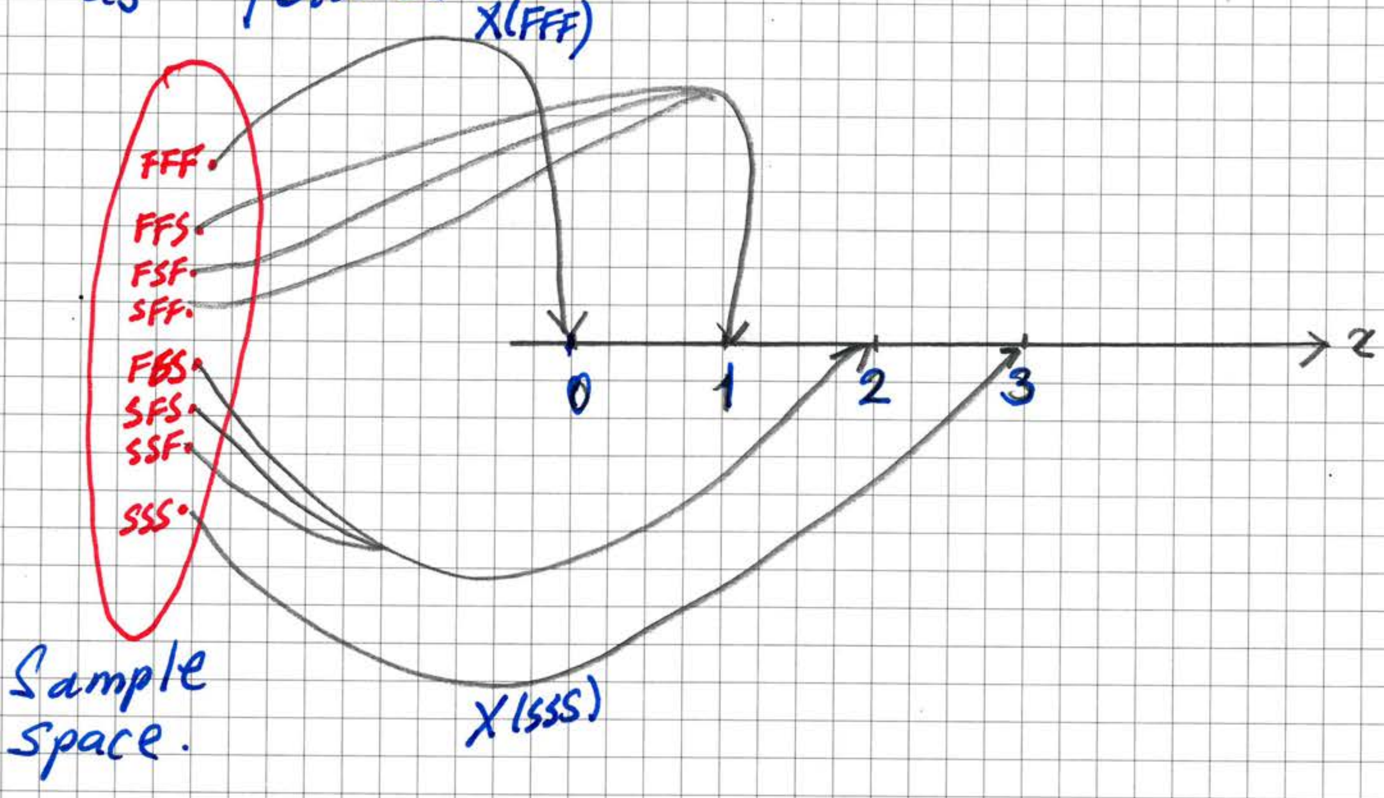
$$= q + pe^{jt}.$$

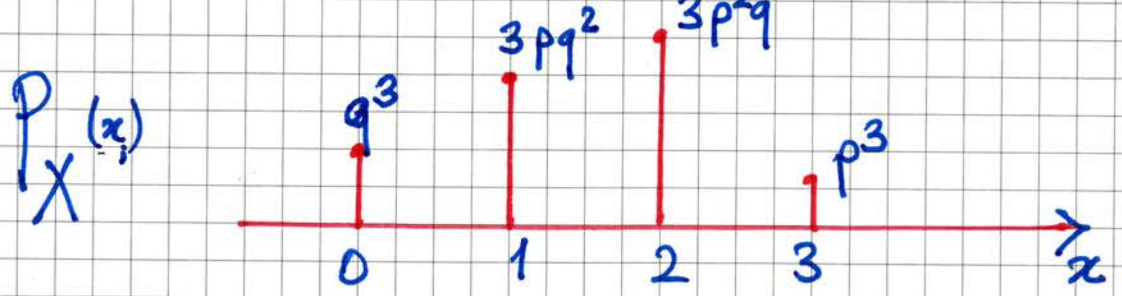
## Example:

Suppose an experiment of three Bernolli trails and the probability of success is $p$ on each trail. The random variable $X$ is given by

$$X = X_1 + X_2 + X_3$$

The distribution of $X$ is determined as follows:



# BINOMIAL DISTRIBUTION

- A random experiment consists of $n$ Bernoulli trails such that

  (1) The trails are independent

  (2) Each trail results in only two possible outcomes, labelled as "success" and "failure".

  (3) The probability of a success in each trail, denoted as $p$, remains constant.

- Let the random varible $X$ be the number of times a success occurs in $n$ trails. It is clear that

$$0 \le X \le n.$$

Question:

What is the probability that you will observe $X=r$ consecutive successes followed by $n-r$ consecutive failures?

$$Pr\{\underbrace{SS\cdots S}_{r}\underbrace{FF\cdots F}_{n-r}\} = p^r q^{n-r}$$

- The probability of such an event is

$$p^r q^{n-r}.$$

- If we know this, we can ask for the probability of exactly $X=r$ successes out of $n$ trails,

- then $p^r q^{n-r}$ must be multiplied by the number of ways these $r$ successes can occur, that is, the number of combinations of $r$ successes in $n$ trails.

$$P_r = Pr\{X=r\}$$
$$= \binom{n}{r} p^r q^{n-r} \quad , \text{for } r=0,1,\dots,n.$$

- Indeed, for $n$ trails $r$ success can be realized $\binom{n}{r}$ different forms. The binom random variable is shown as

$$X \sim B(n, p)$$

number of trails

The probability of success

## Mean and Variance:

We can express a binomial random variable as a sum of n Bernoulli random variables:

$$X = X_1 + X_2 + \cdots + X_n .$$

- Then we can take the expected value of the both sides

$$E[X] = E[X_1 + X_2 + \cdots + X_n]$$
$$= \sum_{k=1}^{n} E[X_k] = \sum_{k=1}^{n} p = np.$$

- The variance of the sum of independent random variables is the sum of the variances of the each random variable:

$$\sigma_X^2 = \sigma^2_{X_1 + X_2 + \cdots + X_n}$$
$$= \sum_{k=1}^{n} \sigma_{X_k}^2$$
$$= \sum_{k=1}^{n} p(1-p) = np(1-p).$$

# Characteristic Function of the binomial random Variable:

$$\Phi_X(t) = E\left[e^{jXt}\right]$$

$$= \sum_{k=0}^{n} e^{jtk} P_k$$

$$= \sum_{k=0}^{n} \binom{n}{k} p^k (1-p)^{n-k} e^{jtk} .$$

- The binomial expansion is

$$(a+b)^n = \sum_{k=0}^{n} \binom{n}{k} a^k b^{n-k}$$

- Therefore,

$$\Phi_X(t) = \sum_{k=0}^{n} \binom{n}{k} \underbrace{(pe^{jt})}_{a}{}^k \underbrace{(1-p)}_{b}{}^{n-k}$$

$$= \left(pe^{jt} + (1-p)\right)^n .$$

Example: [T.T. Soong]

A homeowner has just installed 20 light bulbs in a new home. Suppose that each has a probability 0.2 of functioning more than three months.

- What is the probability that at least five of these function more than three months?
- What is the average number of bulbs the homeowner has to replace in three months?

Answer:

It is reasonable to assume that the light bulbs Perform independently. If X is the number of bulbs functioning more than three months (success), it has a binomial distribution with $n = 20$ and $p = 0.2$. Namely,

$$X \sim B(20, 0.2)$$

$$P_X(k) = \binom{20}{k} 0.2^k \, 0.8^{20-k}, \quad k = 0, 1, \ldots, 20.$$

- The answer to the first question is thus given by

$$\sum_{k=5}^{20} p_X(k) = 1 - \sum_{k=0}^{4} p_X(k)$$

$$= 1 - \sum_{k=0}^{4} \binom{20}{k} 0.2^k \, 0.8^{20-k}$$

$$= 1 - (0.012 + 0.058 + 0.137 + 0.205 + 0.218)$$

$$= 0.37.$$

- The average number of replecements is

$$20 - E[X] = 20 - np$$

$$= 20 - 20(0.2) = 16.$$

## Example [T.T. Soong]

- Suppose that three telephone users use the same number.
- We are interested in estimating the probability that more than one will use it at the same time.
- If the independence of telephone habit is assumed, the probability of exactly $k$ persons requiring use of the telephone at the same time is

given by the mass function associated with the binomial distribution. In this case,

$$P_X(k) = \binom{3}{k} p^k (1-p)^{3-k}, \text{ for } k=0,1,2,3.$$

- Let it be given that, on average, a telephone user is on the phone 5 minutes per hour; an estimate of $p$ is

$$p = \frac{5}{60} = \frac{1}{12}.$$

Hence

$$P_X(k) = \binom{3}{k} \left(\frac{1}{12}\right)^k \left(\frac{11}{12}\right)^{3-k}, \text{ for } k=0,1,2,3,$$

and $X \sim B(3, 1/12)$.

- The solution to this problem is given by

$$\Pr\{X>1\} = P_X(2) + P_X(3)$$

$$= \binom{3}{2}\left(\frac{1}{12}\right)^2\left(\frac{11}{12}\right) + \binom{3}{3}\left(\frac{1}{12}\right)^3\left(\frac{11}{12}\right)^0$$

$$= \frac{11}{864} = 0.0197.$$

# GEOMETRIC DISTRIBUTION

- Another event of interest arising from Bernoulli trails is the number of trails to the first occurrence of success.

- If $X$ is used to represent this number, it is a discrete Random variable with possible values ranging from one to infinity

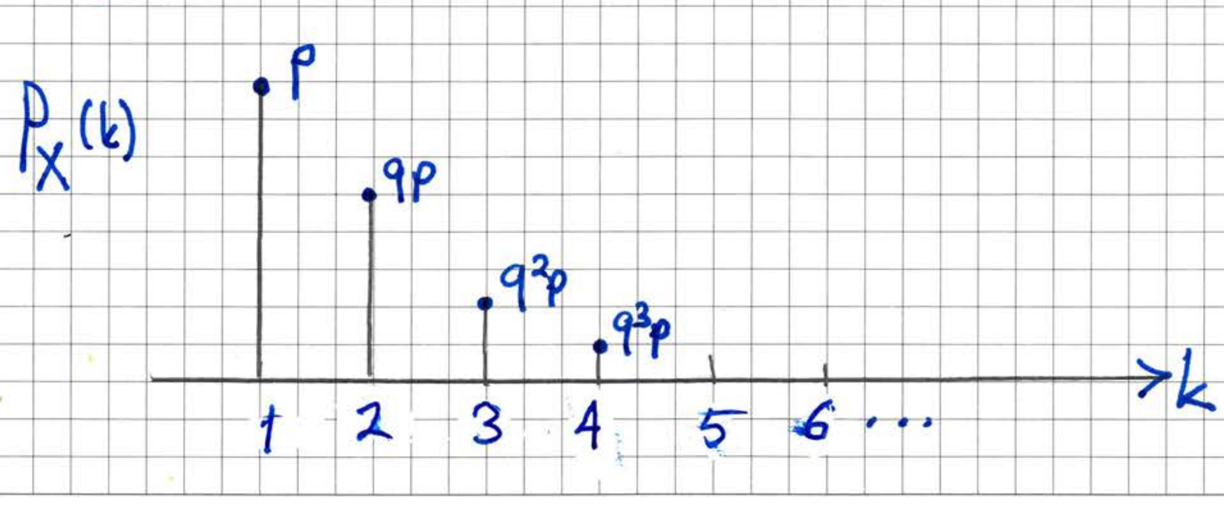
- Its pmf is easily computed to be

$$p_X(k) = Pr\{\underbrace{FF\ldots F}_{k-1}S\}$$

$$= \underbrace{Pr\{F\}Pr\{F\}\cdots Pr\{F\}}_{k-1} Pr\{S\}$$

$$= q^{k-1}p \,, \qquad k = 1, 2, \ldots$$

- This distribution is known as the geometric distribution with parameter $p$, A plot of $p_X(k)$ is given as follows:



Example: [T.T. Soong]

- A driver is looking for a parking place some distance down the street.
- There are five cars in front of the driver, each of which having a probability 0.2 of taking place.
- What is the probability that the car immediately ahead will enter parking space?

Answer:

For this problem, we have a geometric distribution and need to evaluate $p_X(k)$ for $k=5$ and $p=0.2$. Thus,

$$P_X(5) = (0.8)^4(0.2) = 0.82.$$

The Cumulative Probabity Function (cdf):

$$F_X(x) = \sum_{k=1}^{m \le x} p_X(k) = p + qp + \cdots + q^{m-1}p$$

$$= p\left(1+q+\cdots+q^{m-1}\right)$$

$$= (1-q)\left(1+q+\cdots+q^{m-1}\right)$$

$$= 1 - q^m,$$ where $m$ is the largest integer less than or equal to $x$.

# The Mean of the Geometric Random Variable:

$$\mu = E[X] = \sum_{k=1}^{\infty} k \, p_X(k)$$

$$= \sum_{k=1}^{\infty} k \, q^{k-1} p = p \sum_{k=1}^{\infty} \frac{d}{dq} q^k$$

$$= p \frac{d}{dq} \underbrace{\sum_{k=1}^{\infty} q^k}_{\text{since } q=1-p<1} = p \frac{d}{dq}\left(\frac{q}{1-q}\right) = \frac{1}{p}.$$

## Variance:

Following the same procedure, the variance has the form

$$\sigma_X^2 = E\left[(X-\mu_X)^2\right]$$

$$= E\left[\left(X-\frac{1}{p}\right)^2\right]$$

$$= \sum_{k=1}^{\infty} \left(k-\frac{1}{p}\right)^2 p \, q^{k-1} = \frac{q}{p^2}.$$

To see more details, you can read the texbook [kayran].

**Example:** [T.T. Soong]

• Assume that the probability of a sample (specimen) failing during a given experiment is 0.1.

• What is the probability that it will take more than three samples to have one surviving the experiment?

**Answer:**

Let $X$ denote the number of trails requiered for the first sample to survive.

It has than a geometric distribution with $p = 0.9$.

The desired probability is

$$\Pr\{X > 3\} = 1 - F_X(3)$$

$$= 1 - (1 - q^3)$$

$$= q^3 = 0.1^3 = 0.001.$$

**Example:** [T.T. Soong]

Let The probability of occurrence of a flood of magnitude greater than a critical magnitude in any given year be 0.01. Assuming that floods occur independently, determine $E[N]$, the average return period.

- The <u>average return period</u>, or simply <u>return period</u>, is defined as the average number of years between floods for which the magnitude is greater than the critical magnitude

<u>Answer:</u>

It is clear that $N$ is a geometric random variable with $p = 0.01$, namely

$$N \sim \text{Geometric}(0.01).$$

The return period is then

$$E[N] = \frac{1}{p} = 100 \text{ years.}$$

The critical magnitude which gives rise to $E[N] = 100$ years is often referred to as the "100-year flood".

# LACK OF MEMORY PROPERTY

- A geometric random variable has been defined as the number of trails until the fir success.

- However, because the trails are independent the count of the number of trails until the next success can be started at any trail without changing the probability distribution of the random variable.

- The geometric random variable has a lack of memory property. The system always "forgets" its failures and begins as if it were performing the first trail.

- For example,

100 bits transmitted. 6th bit is in error.

$$Pr\{X=106 \mid X>100\} = Pr\{X=6\}$$

106th bit is in error

In the transmission of bits, if 100 bits are transmitted, the probability that the first error, after bit 100, occurs on bit 106 is the probability that the next six outcomes are "OOOOOE".

- The Probability of an error remains constant for all transmissions. In this sense, the geometric distribution is said to lack any memory.

## Example:

At a fair die rolling experiment, what is the expected or average number of rolling in order to get 6?

Until we get first 6, the required number of roll is shown by the geometric random variable. Therefore,

$$X \sim Geo(1/6)$$

and

$$E[X] = \mu_X = \frac{1}{P} = \frac{1}{1/6} = 6.$$

# Poisson Random Variable

(the Distribution of rare events)

- The probability model of a Poisson random variable describes phenomena that occur randomly in time.
- While the time of each occurrence is completely random, there is a known average number of occurrence per unit time.
- The poisson model is used widely in many fields. For example,

- The arrival of information request at a World Wide Web server,
- Passenger arrivales at an airline terminal,
- The initiation of telephone calls,
- Emission of particles from a radioctive source,
- Car arrivals at an intersection,
- and many other similar pheomena.

• To describe a Poisson random variable, we call the occurrence of the phenomenon of interest an <u>arrival</u>.

• A Poisson model often specifies an average rate $\lambda$ [arrivals per second]. The pmf of X has the following form:

$$P_X(k) = \frac{\lambda^k}{k!} e^{-\lambda}, \quad k = 0, 1, 2, \ldots$$

" # of arrials per second."

$\lambda$ is the parameter of this distribution and it is the average rate [arrivals per second].

- If the time interval is T seconds, then, for this time interval, the parameter of the distribution is $\alpha = \lambda T$, and the pmf is as follows

$$P_X(k) = \frac{\alpha^k}{k!} e^{-\alpha}$$

$k = 0, 1, 2, \dots$ [# of arrivals at a time interval T seconds]

## THE POISSON APPROXIMATION TO THE BINOMIAL DISTRIBUTION

- For large $n$ ($n \to \infty$) and small $p$ ($p \to 0$) in a such a way that

$$\lambda = np$$

remains fixed. We note that $\lambda$ is the mean of $X$, which is assumed to remain constant.

- The poisson distribution may be developed as a limiting form of the binomial distribution.

- We return to the binomial distribution $X \sim Binom(n, p)$,

$$P_X(k) = \binom{n}{k} p^k (1-p)^{n-k}$$

$$= \frac{n(n-1)(n-2)\cdots(n-k+1)}{k!} p^k (1-p)^{n-k}$$

- Multiply and divide the right-hand side of this equation by $n^k$, we can write

$$\binom{n}{k} p^k (1-p)^{n-k} = \frac{\overbrace{\left(1-\frac{1}{n}\right)\left(1-\frac{2}{n}\right)\cdots\left(1-\frac{k-1}{n}\right)}^{\to 1}}{k!} \underbrace{(np)^k}_{\lambda^k} \left(1-\frac{np}{n}\right)^{n-k}$$

if $n \to \infty$, $\lambda = np$ will be fixed,

and

$$\left(1-\frac{np}{n}\right)^{n-k} = \left(1-\frac{\lambda}{n}\right)^n \Big/ \underbrace{\left(1-\frac{\lambda}{n}\right)^k}_{\to 1}$$

If $n \to \infty$, then the denomitor $\left(1-\frac{\lambda}{n}\right)^k$ will go to 1 and the numerator will approach to $e^{-\lambda}$. Indeed,

$$\lim_{n\to\infty} \left(1-\frac{\lambda}{n}\right)^n = e^{-\lambda}.$$

Therefore, when $\lambda = np$ remains fixed, at the limiting case, for

For $n \to \infty$ or $p = \frac{\lambda}{n} \to 0 \to$

$$\underbrace{\binom{n}{k} p^k (1-p)^{n-k}}_{\text{Binomial}} \xrightarrow[n \to \infty]{} \underbrace{\frac{\lambda^k}{k!} e^{-\lambda}}_{\text{Poisson}}$$

$X \sim \text{Binom}(n, p)$ $\qquad$ $X \sim \text{Pois}(\lambda = np)$

Arrivals per second.

Example [T.T. Soong]:

In oil exploration, the probability of an oil strike in the North Sea 1 in 500 drillings. What is the probability exactly 3 oil-producing wells in 1000 explorations?

Answer:

In this case, $n = 1000$, and $p = \frac{1}{500} = 0.002$, and the Poisson approximation is appropriate.

We have $\lambda = np = 1000 \cdot (0.002) = 2$, the pmf is

$$P_X(k) = \frac{2^k e^{-2}}{k!}, \quad \text{for } k = 0, 1, 2, \ldots$$

and $P_X(3) = \frac{2^3 e^{-2}}{3!} = 0.18.$

## • SPATIAL DISTRIBUTIONS

• The Poisson distribution has been derived based on arrivals developing in time. However, we can use the same argument to distribution of points in the space.

• We can consider the distribution of imperfections (flaws) in a material.

• If the number of flaws in a given volume has a Poisson distribution, time interval can be replaced by volumes

### Example:

• A good example of this application is the study carried out by Clark (1946) concerning the distribution of flying-bomb hits in one part of London during World War 2.

• The area divided into 576 small areas of 0.25 $km^2$ each. In the following table, the number $n_k^o$ denotes the number of areas with exactly $k$ hits recorded and $n_k^p$ denotes the predicted number based on a Poisson distribution, with

$$\lambda t = \frac{\text{number of total hits}}{\text{per number of areas}} = \frac{537}{576} = 0.932.$$

# Comparison of the observed and theoretical Distributions of flying-bomb hits, in London.

| $n_k$ \ $k$ | 0 | 1 | 2 | 3 | 4 | $\geq 5$ |
|---|---|---|---|---|---|---|
| $n_k^o$ | 229 | 211 | 93 | 35 | 7 | 1 |
| $n_k^p$ | 226.7 | 211.4 | 98.5 | 30.6 | 7.1 | 1.6 |

We observe an excellent agreement between the predicted and observed results.

$$P_X(k) = \frac{(0.932)^k e^{-0.932}}{k!}, \text{ for } k=0,1,2,\ldots$$

**Example:** [T.S. Soong]

- Suppose that the probability of a transistor manufactured by a certain firm being defective is 0.015.

- What is the probability that there is no defective transistor in a batch of 100?

**Answer:** Let X be the number of defective transistors in 1000. The desired probability is

$$P_X(k) = \binom{100}{k} (0.015)^k (0.985)^{100-k}$$

For $k=0$:

$$P_X(k) = \binom{100}{0} (0.015)^0 (0.985)^{100-0}$$

$$= (0.985)^{100} = 0.2206.$$

Since $n$ is large and $p$ is small in this case, the Poisson approximation is appropriate and we obtain

$\lambda = 100 \cdot 0.015 = 1.5$,

$$P_X(k) = \frac{(1.5)^k e^{-1.5}}{k!}, \text{ for } k = 0, 1, \ldots$$

For $k=0$,

$$P_X(0) = \frac{(1.5)^0 e^{-1.5}}{0!} = e^{-1.5} = 0.223.$$

which is very close to exact answer.

Remark:

In practice, the Poisson approximation is frequently used when $n > 100$, and $p < 0.1$.

Example:

The probability that a particular rivet (persin) in the wing surface of a new aircraft is defective is 0.001. There are 4000 rivets in the wing.

What is the probability that not more than six defective rivets will be installed?

$$\Pr\{X \le 6\} = \sum_{x=0}^{6} \binom{4000}{x} (0.001)^x (0.999)^{4000-x}$$

Using the Poisson approximation;

$$\lambda = 4000\,(0.001) = 4$$

$$\Pr\{X \le 6\} = \sum_{x=0}^{6} e^{-4} \frac{4^x}{x!} = 0.889.$$

Expected Value of the Poisson Distribution

$$E[X] = \sum_{k=0}^{\infty} k\, p_k = \sum_{k=0}^{\infty} k \frac{\lambda^k}{k!} e^{-\lambda}$$

$$= \lambda e^{-\lambda} \sum_{k=1}^{\infty} \frac{\lambda^{k-1}}{(k-1)!} = \lambda e^{-\lambda} \underbrace{\sum_{k=0}^{\infty} \frac{\lambda^k}{k!}}_{e^{\lambda}}$$

$$= \lambda e^{-\lambda} e^{\lambda} = \lambda.$$

- Similarly, it can be shown that

$$\sigma_X^2 = VAR[X]$$
$$= E[(X-\lambda)^2] = \lambda .$$

- For a poission random variable,

$$\mu_X = \sigma_X^2 = \lambda .$$

Remarks:

1. $X_1 \sim Pois(\lambda_1)$ and $X_2 \sim Pois(\lambda_2)$

For $\lambda_2 > \lambda_1$, we have

$$Pr\{X_2 > k\} > Pr\{X_1 > k\}, \quad k = 0, 1, \ldots .$$

It is possible to show the proof from the definition of the random variable.

2. $X_1 + X_2 \sim Pois(\lambda_1 + \lambda_2)$

its pmf is,

$$P_{X_1+X_2}(k) = \frac{(\lambda_1+\lambda_2)^k}{k!} e^{-\lambda_1-\lambda_2}, \quad k = 0, 1, \ldots .$$

The Moment-Generating Function

$$M_X(t) = E[e^{tX}] = \sum_{k=0}^{\infty} e^{tk} p_k = \sum_{k=0}^{\infty} e^{tk}\left(\frac{\lambda^k e^{-\lambda}}{k!}\right)$$
$$= e^{\lambda(e^t-1)} .$$

(more details in the textbook (Kayman)).