Stochastic processes:

A stochastic process $\{X(t), t \in T\}$ is a collection of random variables. That is, for each $t \in T$, $X(t)$ is a random variable. The index t is often interpreted as time and, as a result, we refer to $X(t)$ as the state of the process at time t.

For example, $X(t)$ might equal the total number of customers that have entered a supermarket by time t; or the number of customers in the supermarket at time t; or the total amount of sales that have been recorded in the market by time t, etc.

The set T is called the index set of the process. When T is countable set, the stochastic process is said to be a discrete-time process. If T is an open or closed interval of the real line, the stochastic process is said to be a continuous -time process. ~~process indexed by the nonnegative integers, while $\{X(t), t \geq 0\}$ is a continuous-time stochastic process indexed~~

For instance, $\{X_n, n = 0, 1, 2, 3, \ldots\}$ is a discrete-time stochastic process indexed by the nonnegative integers. while $\{X(t) : t \geq 0\}$ is a continuous-time stochastic

Scanned with CamScanner

process indexed by the nonnegative real numbers.
The state space of a stochastic process is defined as the set of all possible values that the random variables $X(t)$ can assume.
Thus, a stochastic process is a family of random variables that describes the evolution through time of some (physical) process.

In this chapter, we will define and discuss certain parametric families of univariate probability distributions that have been widely applied in a variety of decision-making situations.

A parametric family of distributions is a collection of distributions that is indexed by a quantity called "parameter". These distributions have standard names and can be derived under certain plausible conditions about the random variables involved.

In addition, they can be expressed by algebraic - functions and hence are capable of further mathematical treatment. The distributions to be presented here include both discrete and continuous distributions.

The most frequently encountered discrete distributions, among others, are Bernoulli distribution, Binomial distribution, Hypergeometric distribution, Poisson distribution, Geometric distribution, Negative binomial distribution, and multinomial distribution.

The commonly discussed continuous distributions are Normal distribution, Exponential distribution, Beta distribution, Gamma distribution, and Rectangular distribution.

Bernoulli Distribution : / Bernoulli Process :

A considerable group of random phenomena, known as Bernoulli process, named after James Bernoulli (1654–1705), follows the simplest probability distribution - involving only two possible outcomes, such as head or tail, success or failure, defective or non-defective, smoker or non-smoker, married or unmarried and like that.

To standardize the terminology describing these and many other similar processes, we would call one of the possible outcomes a "success" and the other a "failure". These names are used only to identify the outcomes and bear no connotation of success or failure in real life.

Customarily, the outcome of primary interest in a study is labeled a success (even if is a disastrous event).

In a study of prevalence of marriage dissolution or rate of electricity failure, the marital status "divorced" or "load shading" may be attributed the statistical name "success".

The events "success" and "failure" may be viewed as outcomes of experiments composed of repeatitions of independent trials. When the probabilities of the two outcomes remain unchanged from one trial to another, we name the trials "Bernoulli trials", in honor of James Bernoulli.

In all applications, it is convenient to designate the two possible outcomes of such an experiment as 0 and 1. Each of the following examples demonstrates a Bernoulli trial:

(i) Suppose a fair coin is tossed repeatedly. Let, $X_i = 1$ if a head is obtained on the ith toss and let $X_i = 0$ if a tail is obtained $(i = 1, 2, 3, \ldots)$. Then the random variables $X_1, X_2, X_3, \ldots$ form a sequence of Bernoulli trials with parameter $p = 0.5$.

(ii) Suppose that 20 percent of the item produced by a machine of a manufacturing industry are defective and that n items are selected at random and inspected. Assigning $X_i = 0$ if the ith item is non-defective and $X_i = 1$, if it is defective $(i = 1, 2, 3, \ldots)$, then the variables $X_1, X_2, \ldots$ form n Bernoulli trials with parameter $p = 0.20$.

Def^n of Bernoulli distribution :

A random variable X is said to have a Bernoulli distribution with parameter p $(0 \leq p \leq 1)$ if X can take on only the values 0 and 1 and the probabilities are

$$\left.\begin{aligned} P(X=1) &= p \\ \& \; P(X=0) &= 1-p \end{aligned}\right\} \quad \text{——} \; (1)$$

Since the distribution involves only two cases of events, it is also known as two-point probability distribution.

If we let $q = 1-p$, then the probability function of X can be written as follows:

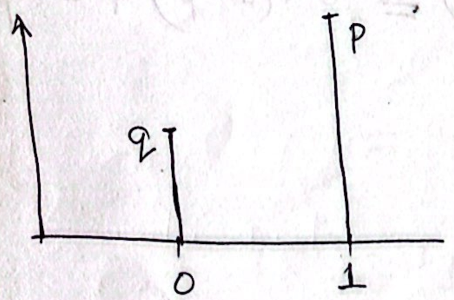
$$f(x;p) = \begin{cases} p^x q^{1-x} \; ; & \text{for } x = 0, 1 \\ 0 \; ; & \text{otherwise} \end{cases} \quad \text{——} \; (2)$$

Scanned with CamScanner

To verify that the probability function (2) actually represents the Bernoulli distribution specified by the probabilities ①, it is necessary to note that

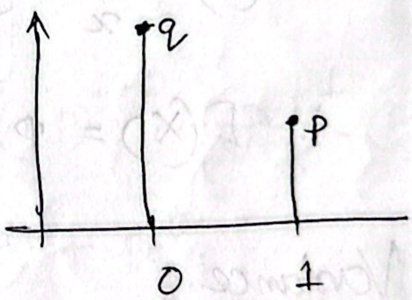
$f(1, p) = p$ and $f(0, p) = q$

The tabular representation of the random variable X is as follows:



fig(a)

$q < p$



fig(b)

$q > p$

Bernoulli functions with $q < p$ (fig-a) & $q > p$ (fig.-b)

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

Properties of Bernoulli Distribution:

Similar to other distributions, the Bernoulli distribution has its mean, variance, and other descriptive measures. These are obtained as follows:

Mean:

$$E(X) = \sum_x x f(x, p) = 0 * (1-p) + 1 * p = 0 + p = p$$

$$\therefore E(X) = p$$

Variance:

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

$$= \sum_x x^2 f(x,p) - [E(X)]^2$$

$$= 0^2 * (1-p) + 1^2 * p - (p)^2$$

$$= p - p^2$$

$$= p(1-p)$$

$$= pq \quad [\because 1-p = q]$$

$$\therefore V(X) = pq$$

Moment generating function :

$$M_X(t) = E\left(e^{tX}\right) = \sum_x e^{tx} f(x,p)$$

$$= \sum_x e^{tx} p^x (1-p)^{1-x}$$

$$= \sum_x \left(pe^t\right)^x (1-p)^{1-x}$$

$$= (pe^t)^0 (1-p)^{1-0} + (pe^t)^1 \cdot (1-p)^{1-1}$$

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

$$\therefore M_X(t) = pe^t + q$$

Moments :

Chebychev's Theorem / Chebyshev's Inequality:

If the probability or population histogram is roughly bell-shaped and the mean and variance are known, the empirical rule (i.e., $\mu \pm \sigma$, $\mu \pm 2\sigma$, $\mu \pm 3\sigma$ contain respectively atmost 68%, 95% and 99.99% of the measurements) is of great help in approximating the probabilities of certain intervals.

However in many instances, the shapes of probability histograms differ significantly from a mound shape, and the empirical rule may not yield useful approximations to the probability of interest. The following result, known as Chebychev's theorem, can be used to determine a lower bound for the probability that the random variable X of interest falls in an interval $\mu \pm k\sigma$.

Chebychev's Theorem / Chebychev's Inequality:

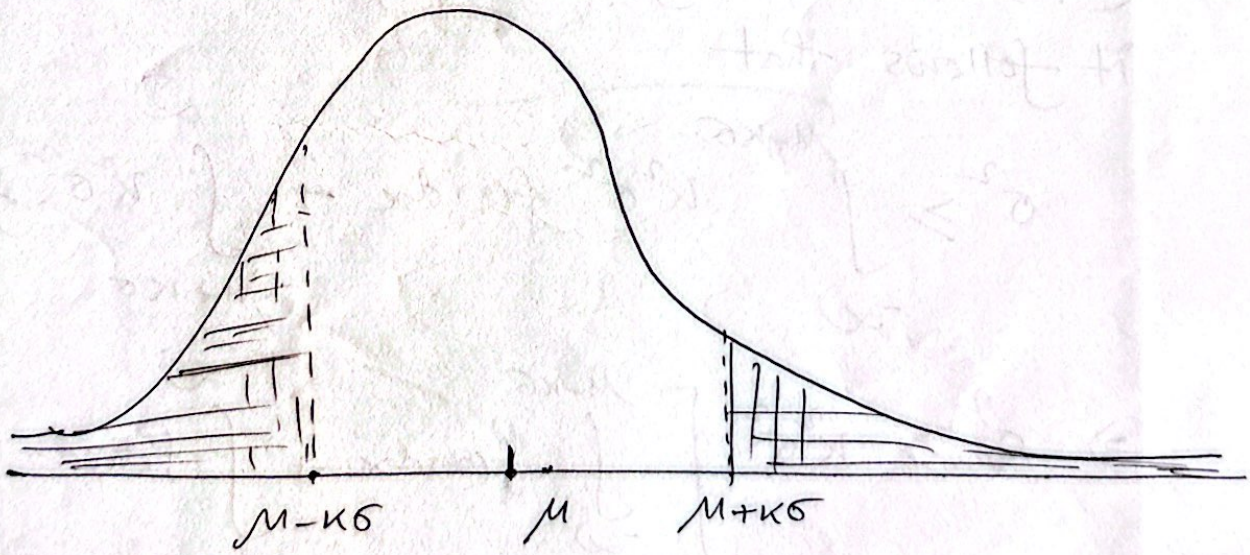
If μ and σ are the mean and the standard deviation of a random variable X, then for any positive constant k the probability is at least $1-\frac{1}{k^2}$ that X will take on a value within k standard deviations of the mean; symbolically,

$$P\left(|X-\mu|<k\sigma\right) \geq 1-\frac{1}{k^2}, \quad \sigma \neq 0$$

Proof:

By definition of variance, we can write

$$\sigma^2 = E\left[(X-\mu)^2\right] = \int_{-\infty}^{\infty} (x-\mu)^2 f(x)\,dx$$



Then dividing the integral into three parts as shown in the above figure, we get

$$\sigma^2 = \int_{-\infty}^{\mu-k\sigma} (x-\mu)^2 f(x)\,dx + \int_{\mu-k\sigma}^{\mu+k\sigma} (x-\mu)^2 f(x)\,dx$$

$$+ \int_{\mu+k\sigma}^{\infty} (x-\mu)^2 f(x)\,dx$$

Since the integrand $(x-\mu)^2 f(x)$ is nonnegative, we can form the inequality

$$\sigma^2 \geq \int_{-\infty}^{\mu-k\sigma} (x-\mu)^2 f(x)\,dx + \int_{\mu+k\sigma}^{\infty} (x-\mu)^2 f(x)\,dx$$

by deleting the second integral. Therefore, since $(x-\mu)^2 \geq k^2\sigma^2$ for $x \leq \mu - k\sigma$ or $x \geq \mu + k\sigma$ it follows that

$$\sigma^2 \geq \int_{-\infty}^{\mu-k\sigma} k^2\sigma^2 f(x)\,dx + \int_{\mu+k\sigma}^{\infty} k^2\sigma^2 \cdot f(x)\,dx$$

$$\Rightarrow \sigma^2 \geq k^2\sigma^2 \left[\int_{-\infty}^{\mu-k\sigma} f(x)\,dx + \int_{\mu+k\sigma}^{\infty} f(x)\,dx \right]$$

$$\Rightarrow \frac{1}{k^2} \geq \int_{-\infty}^{\mu-k\sigma} f(x)\,dx + \int_{\mu+k\sigma}^{\infty} f(x)\,dx \; ; \; \sigma^2 \neq 0$$

Since the sum of the two integrals on the right-hand side is the probability that X will take on a value less than or equal to $\mu - k\sigma$ or greater than or equal $\mu + k\sigma$, we have thus shown that

$$P(|X-\mu| \geq k\sigma) \leq \frac{1}{k^2}$$

$$\Rightarrow 1 - P(|X-\mu| < k\sigma) \leq \frac{1}{k^2}$$

$$\Rightarrow 1 - \frac{1}{k^2} \leq P(|X-\mu| < k\sigma)$$

$$\Rightarrow P(|X-\mu| < k\sigma) \geq 1 - \frac{1}{k^2}.$$

(Proved)

Scanned with CamScanner

Example:

If the probability density of X is given by

$$f(x) = \begin{cases} 630\, x^4 (1-x)^4 & \text{for } 0 < x < 1 \\ 0; & \text{otherwise} \end{cases}$$

find the probability that it will take on a value within two standard deviations of the mean and compare this probability with the lower bound provided by Chebychev's theorem.

Sol^n:

Mean, $\mu = E[X] = \int_0^1 x\, f(x)\, dx$

$$\Rightarrow \mu = \int_0^1 x\; 630\, x^4 (1-x)^4\, dx$$

$$= 630 \int_0^1 x^5 (1-x^4)\, dx$$

$$= \frac{1}{2}$$

Variance, $\sigma^2 = V[X] = E[X^2] - \{E[X]\}^2$ —①

Now, $E[X^2] = \int_0^1 x^2 f(x)\, dx$

Scanned with CamScanner

$$\Rightarrow E[X^2] = 630 \int_0^1 x^6 (1-x)^4 \, dx$$

$$=$$

$$\therefore \sigma^2 = \frac{1}{44}$$

$$\Rightarrow \sigma = \sqrt{\frac{1}{44}} \simeq 0.15$$

Thus, the probability that X will take on a value within two standard deviations of the mean is the probability that it will take on a value between $(\mu - k\sigma = \{0.5 - (2 * 0.15)\} = 0.20)$ and $(\mu + k\sigma = 0.5 + 0.30 = 0.80)$, that is

$$P(0.20 < X < 0.80) = \int_{0.20}^{0.80} 630x^4 (1-x)^4 \, dx$$

$$= 0.96$$

Observe that the statement "the probability is 0.96" is a much stronger statement than "the probability is at least $1 - \frac{1}{k^2} = 1 - \frac{1}{2^2} = 0.75$

Scanned with CamScanner

which is provided by Chebychev's inequ... or th^m.

Example:

The number of customers X per day at a sales counter, has been observed for a long time and found to have a mean 20 and variance 4. The probability distribution of X is not known. What can be said about the probability that tomorrow the number of customers will exceed 16 but in no case will be greater than 24?

Solution:

The problem here is to find $P(16 < X < 24)$.

Applying Chebychev's theorem

$$P(\mu - k\sigma < X < \mu + k\sigma) \geq 1 - \frac{1}{k^2}$$

For $\mu = 20$ & $\sigma = 4$ we have $k = 2$

$$P(16 < X < 24) = P(\mu - 2\sigma < X < \mu + 2\sigma)$$

$$\geq 1 - \frac{1}{2^2} = 0.75$$

$\therefore P(16 < X < 24) \geq 0.75$

The result indicates that tomorrow's customer total will be between 16 and 24 with a fairly high probability of at least 0.75 or 75%.