# Chapter III
# Special Distributions

## 1 Binomial Distribution (James Bernoulli)

Consider a random experiment, which is so defined that it has only two possible outcomes which we call *success* and *failure*. Let $p$ be the probability of a success and $q$ = 1-$p$ be the probability of failure. Let the experiment be repeated $n$ times. Also assume that the trails are independent and the probability of success $p$ remains unaltered by trail to trail. Then obviously the number of success in $n$ trails is a random variable. Let $X$ denote the number of successes in $n$ independent repetitions of this experiment and let $f(x)$ be the p.d.f. of $X$.

If the $x$ repetitions in which success occurs are specified, the probability of success in these $x$ repetitions and failures in the remaining $n-x$ repetitions is $p^x q^{n-x}$ by the multiplication theorem the repetitions being independent. The $x$ trials in which success occurs may be specified in $\binom{n}{x}$ mutually exclusive ways. So the event $X = x$ can occur in $\binom{n}{x}$ mutually exclusive ways and probability of each is $p^x q^{n-x}$. So by addition theorem, the required probability is $\binom{n}{x} p^x q^{n-x}$.

i.e.

$$f(x) = \binom{n}{x} p^x q^{n-x}, x = 0, 1, \ldots, n, 0 \leq p \leq 1.$$

This distribution is called binomial distribution.

**Definition**:- A discrete random variable $X$ is said to follow a binomial distribution with parameters $n$ and $p$ if its probability density function is given by

$$f(x) = \binom{n}{x} p^x q^{n-x}, x = 0, 1, \ldots, n, 0 \leq p \leq 1, p+q=1.$$

**Assumption of Binomial distribution can be applied**

1. The random experiment has two outcomes, which can be called "success" and "failure"
2. Probability for success in a single trails remains constant from trail to trail of the experiment.
3. The experiment is repeated a finite number of times
4. The trails are independent.

**Properties**

1. Mean = E($X$) = $np$ and variance = V($X$) = $npq$.
2. If $(n+1)p$ is integer the binomial distribution has two mode one at $(n+1)p - 1$ and the other at $(n+1)p$. If $(n+1)p$ is not an integer then the integer part of $(n+1)p$ is the mode.
3. If $p = q = ½$ the distribution is symmetrical and when $p \neq q$ the distribution is a skewed distribution.
4. If $X_i$ ($i$=1, 2, …, $k$) are independent binomial random variate with parameters ($n_i$, $p$). Then $S_k = \sum_{i=1}^{k} X_i$ has a binomial distribution with parameters ($\sum_{i=1}^{k} n_i$, $p$). (This property is called additive or reproductive property).
5. If $n$ is very large and if neither $p$ and $q$ is too close to zero binomial distribution may be approximated by normal distribution.

| | |
|---|---|
| Mean: | $np$ |
| Variance: | $np(1-p)$ |
| Mode: | The largest integer $\leq (n+1)p$ |
| Mean Deviation: | $2n\binom{n-1}{m}p^{m+1}(1-p)^{n-m}$, where $m$ denotes the largest integer $\leq np$. [Kamat 1965] |
| Coefficient of Variation: | $\sqrt{\frac{1-p}{np}}$ |
| Coefficient of Skewness: | $\frac{1-2p}{\sqrt{np(1-p)}}$ |
| Coefficient of Kurtosis: | $3 - \frac{6}{n} + \frac{1}{np(1-p)}$ |

## 2 Normal Distribution

A continuous random variable $X$ with p.d.f.

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left\{\frac{-(x-\mu)^2}{2\sigma^2}\right\}, \; -\infty<x<\infty$$

is said to follow normal distribution with parameters ($\mu$, $\sigma$).

**Properties**

1. The normal curve is bell shaped, unimodal, symmetric about the mean and approaches to the *x* axis at $\pm\infty$.
2. The mean, mode and median coincide at $x = \mu$ and $\sigma$ is the standard deviation.
3. The odd central moments are zero.
4. The curve is mesokurtic
5. It has points of inflexion at $x = \mu \pm \sigma$.
6. The mean deviation about the mean is $\frac{4}{5}\sigma$
7. The quartile deviation is $\frac{2}{3}\sigma$
8. The interval $(\mu - 2\sigma, \mu + 2\sigma.)$ contains 95% of the area and $(\mu - 3\sigma, \mu + 3\sigma.)$ contains 99% of the area under the curve.
9. If $X$ has normal distribution with parameter $(\mu, \sigma)$ and $a$ and $b$ are any two real numbers then the random variable $Y = aX + b$ is also a normal distribution with parameters $(a\mu + b, a\sigma)$.
10. If $X$ and Y are independent normal vairate with mean and SD $(\mu_1, \sigma_1)$ and $(\mu_2, \sigma_2)$ and $a$ and $b$ are any two real numbers then the random variable $Z = aX + bY$ is also a normal distribution with mean and SD $(a\mu_1 + b\mu_2, \sqrt{a^2\sigma_1^2 + b^2\sigma_2^2})$.

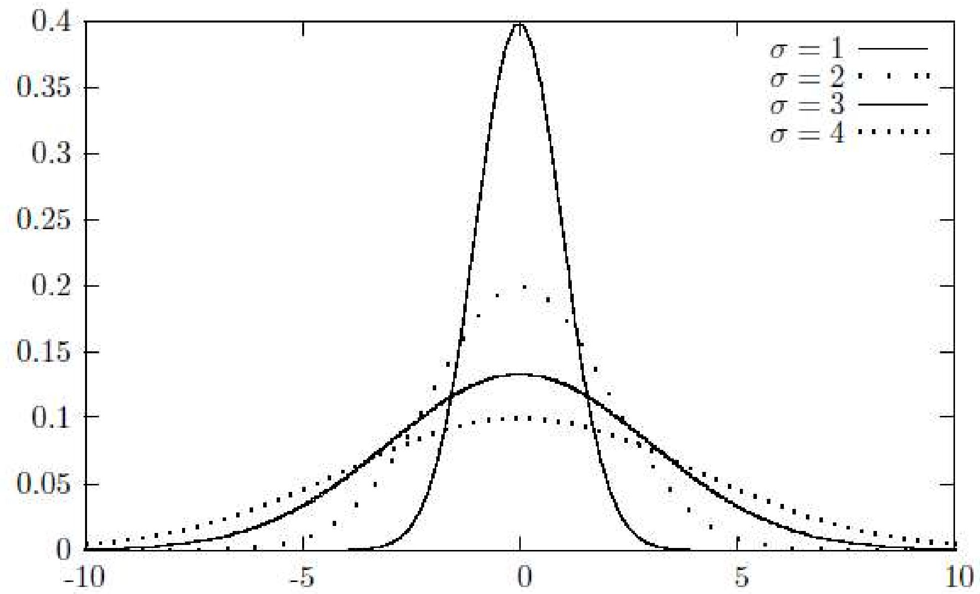


**Figure 10.2** Normal pdfs with $\mu = 0$

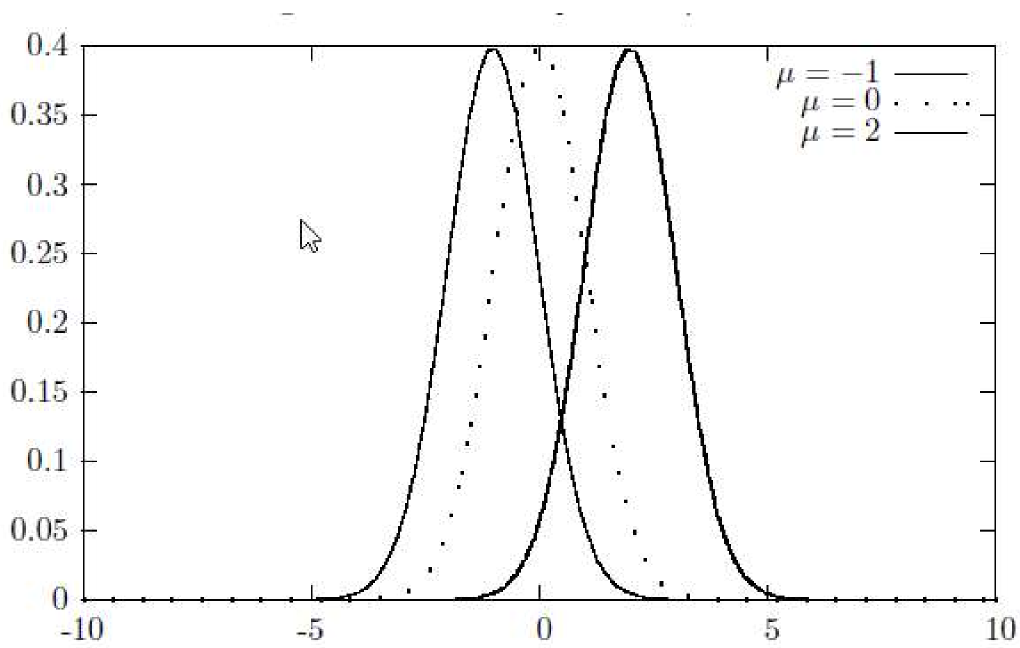


**Figure 10.3** Normal pdfs with $\sigma = 1$

| | |
|---|---|
| Mean: | $\mu$ |
| Variance: | $\sigma^2$ |
| Coefficient of Variation: | $\sigma/\mu$ |
| Median: | $\mu$ |
| Mean Deviation: | $\sqrt{\frac{2\sigma^2}{\pi}}$ |
| Coefficient Skewness: | 0 |
| Coefficient of Kurtosis: | 3 |

**Standard Normal Distribution**

A normal distribution with mean zero and variance one is called a standard normal distribution. That is the pdf of the standard normal distribution is

$$f(x) = \frac{1}{\sqrt{2\pi}} \exp\left\{\frac{-x^2}{2}\right\}, \ -\infty<x<\infty$$

**Note: If $X$ is a normal ($\mu, \sigma$) then $Z = \frac{x-\mu}{\sigma}$ has standard normal distribution.**