Lecture 6 : Distributions of Continuous Random Variables

In this chapter we will be more focused with collecting and collating information, rather than delving into conceptual issues. Let us note that we shall list in this chapter some important distributions associated with continuous random variables. We are not going to make detailed computation of mean and variances or mgf of each and every distribution that we list here. But we will of course do so for ~~all~~ some of the most well known ones, like the exponential distribution, Gamma distribution and above all the jewel of statistics, the normal distribution.

Lets start with a simple one first, the uniform distribution. This is a distribution whose density function remains constant over an interval

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

It is clear that f_X is a density function. Observe that $f_X(x) \geq 0 \ \forall x$ and

$$\int_{-\infty}^{\infty} f_X(x)\,dx = \int_a^b \frac{1}{b-a}\,dx = 1.$$

Further $E(X) = \displaystyle\int_a^b x \frac{1}{b-a}\,dx = \frac{1}{b-a}\left(\frac{b^2}{2} - \frac{a^2}{2}\right) = \frac{b+a}{2}$

while $\mathrm{Var}(X) = \dfrac{(b-a)^2}{12}$ (Compute it yourself)

The mgf for the uniform distribution is given as

$$m_X(t) = E\left[e^{tx}\right] = \frac{e^{bt} - e^{at}}{(b-a)\,t}$$

Before we discuss the normal distribution, let us ~~also~~ mention few important distributions which are used in statistics.

Section 1: The Gamma Distribution and Exponential ~~distri~~ distribution

The Gamma distribution depends on the Gamma function

The Gamma Function is given as

$$\Gamma(z) = \int_0^\infty e^{-x} x^{z-1}\, dx, \quad z>0$$

If $z \in \mathbb{N}$ then

$$\Gamma(z) = (z-1)!$$

Further if $z = 1/2$ then it is well-known that $\Gamma(1/2) = \sqrt{\pi}$

The Gamma function, is an improper integral. For more details see any good book on real analysis or advanced calculus. In fact

$$\boxed{\Gamma(z+1) = z\,\Gamma(z)}$$

Just for fun let us see why $\Gamma(1/2) = \sqrt{\pi}$. Note that

$$\Gamma(1/2) = \int_0^\infty \frac{e^{-x}}{\sqrt{x}}\, dx. \quad \text{Set } x = t^2, \text{ i.e. } dx = 2t\,dt$$

$$\therefore \quad \Gamma(1/2) = \int_0^\infty \frac{e^{-t^2}}{t}\, 2t\,dt = 2\int_0^\infty e^{-t^2}\, dt$$

There is a famous formula in calculus, i.e

$$\int_0^\infty e^{-t^2}\, dt = \frac{\sqrt{\pi}}{2}.$$

Thus $\Gamma(\frac{1}{2}) = \sqrt{\pi}$.

A random variable X is said to follow the Gamma distribution if it ~~is~~ has the density

$$\boxed{f_X(x) = \frac{\lambda}{\Gamma(r)} (\lambda x)^{r-1} e^{-\lambda x}}, \quad x \geq 0 \quad (\lambda > 0, r > 0)$$

Here r and λ is the parameter. We say that $X \sim \text{Gamma}(r, \lambda)$.

What happens if $r = 1$. Then

$$f_X(x) = \frac{\lambda}{\Gamma(1)} (\lambda x)^0 e^{-\lambda x}, \quad x \geq 0, \lambda > 0$$

$$\therefore \boxed{f_X(x) = \lambda e^{-\lambda x}} \longrightarrow (*)$$

Note that $\Gamma(1) = 1$, since $0! = 1$ (By convention)

The expression in $(*)$ is also a p.d.f which is called exponential distribution.

The exponential distribution has the following amazing property called "Memorylessness". Let us see why it is called so.

Let $X \sim \exp(\lambda)$, i.e. X is a random variable with exponential distribution with parameter λ. Then

$$P[X > a+b \mid X > a] = P[X > b], \quad a > 0, b > 0$$

It does not keep the information $X > a$, used in the conditioning.

$$P[X > a+b \mid X > a] = \frac{P[X > a+b \ \& \ X > a]}{P[X > a]}$$

$$= \frac{P[X > a+b]}{P[X > a]}$$

$$P[X > a+b] = 1 - P[X \leq a+b]$$

$$= 1 - \int_0^{a+b} \lambda e^{-\lambda x}\, dx$$

$$= 1 - \left[\lambda \int_0^{a+b} e^{-\lambda x}\, dx \right]$$

$$= 1 - \left[\lambda \left[\frac{e^{-\lambda x}}{-\lambda} \right]_0^{a+b} \right]$$

$$= 1 - \cancel{\lambda} \left[-e^{-\lambda x} \right]_0^{a+b}$$

$$= 1 - \left[-e^{\lambda (a+b)} + 1 \right]$$

$$= 1 + e^{-\lambda (a+b)} - 1$$

$$= e^{-\lambda (a+b)}$$

Similarly $P[X > a] = e^{-\lambda a}$. Thus

$$P[X > a+b \mid X > a] = \frac{e^{-\lambda (a+b)}}{e^{-\lambda a}} = e^{-\lambda b} = P[X > b].$$

The idea of memoryless-ness is used in reliability theory. For example let X denote the random variable which represents the life-time ~~on~~ of a component in a device counted for example in hours. Then

$$P[X > a+b \mid X > a]$$

seeks to find the probability, that given that the component has functioned for more than a hours, what is the probability that it will work for more than additional b hours, i.e. it will work for a+b hours. Memory less-ness tells us that this probability is same as the probability that the component has worked more than b hours.

For seeing the application of exponential distribution see for example the book "Probability and Statistics, with reliability, queuing and computer science applications" by Kishore Trivedi (Wiley 2002).

We ask the reader to compute the mean and variance of an exponential distribution. The answers are

$$E[X] = \lambda \quad \text{and} \quad Var(X) = \frac{1}{\lambda^2}$$

For the Gamma distribution, we shall compute the mean and variance using mgf technique. Of course we have

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

$$= \int_0^\infty e^{tx} \frac{\lambda^r}{\Gamma(r)} e^{-\lambda x} x^{r-1} dx$$

$$= \int_0^\infty \frac{\lambda^r}{\Gamma(r)} e^{-(\lambda-t)x} x^{r-1} dx$$

$$= \frac{\lambda^r}{\cancel{(\lambda-t)^r}} \int_0^\infty \frac{1}{\Gamma(r)} e^{-(\lambda-t)x} x^{r-1} dx$$

$$= \lambda^r \int_0^\infty \frac{1}{\Gamma(r)} x^{r-1} e^{-(\lambda-t)x} dx$$

Set $(\lambda-t)x = y \Rightarrow dy = (\lambda-t)dx$

$$\therefore \quad m_X(t) = \frac{\lambda^r}{(\lambda-t)^r} \int_0^\infty \frac{(\lambda-t)^r}{\Gamma(r)} x^{r-1} e^{-(\lambda-t)x} dx \qquad (\text{Assume } \lambda > t)$$

$$= \frac{\lambda^r}{(\lambda-t)^r} \frac{1}{\Gamma(r)} \int_0^\infty y^{r-1} e^{-y} dy$$

$$= \frac{\lambda^r}{(\lambda-t)^r} \frac{1}{\Gamma(r)} \Gamma(r) = \frac{\lambda^r}{(\lambda-t)^r}$$

$$\therefore \quad m'_X(t) = r\lambda^r (\lambda-t)^{-r-1} \Rightarrow \boxed{E(X) = m'_X(0) = \frac{r}{\lambda}}$$

To compute $E[X^2]$, we need to compute $m_X''(0)$. Observe that

$$m_X''(t) = r(r+1)\lambda^r(\lambda-t)^{-r-2}$$

Hence

$$m_X''(0) = \frac{r(r+1)}{\lambda^2}$$

$$\therefore \boxed{Var(X) = \frac{r(r+1)}{\lambda^2} - \frac{r^2}{\lambda^2} = \frac{r}{\lambda^2}}$$

The Gamma distribution plays a key role in ~~quea~~ ~~queing~~ queuing theory.

Section 2: The Beta Distribution

The ~~notion~~ notion of Beta distribution depends on the idea of ~~B~~ Beta-functions an important class of functions in mathematical analysis, which is given as

$$B(a,b) = \int_0^1 x^{a-1}(1-x)^{b-1}\,dx, \qquad a>0,\ b>0.$$

To begin ~~wt~~ with we will first explore the relation between Beta and Gamma function. The ~~reati~~ relation is

$$\cancel{B(a,}$$

$$B(a,b) = \frac{\Gamma(a)\Gamma(b)}{\Gamma(a+b)}$$

(see any book on advanced calculus for a proof. You need double integrals for proving this). The Beta distribution has the pdf

$$f_X(x) = \frac{1}{B(a,b)}\, x^{a-1}(1-x)^{b-1}; \quad \cancel{x} \; 0<x<1$$

with $a>0$ & $b>0$.

It is of course simple to show that f_X is a p.d.f

We say $X \sim \text{Beta}(a,b)$. If $a=1$ & $b=1$, then the Beta distribution is just the uniform distribution. The distribution function or cdf can be in fact very compactly repented

$$F_X(x) = \begin{cases} 0, & x \le 0 \\ \int_0^x \frac{1}{B(a,b)} z^{a-1}(1-z)^{b-1}\,dz, & x \in (0,1) \\ 1, & x \ge 1. \end{cases}$$

Let us compute mean and variance of the beta distribution. We present here the approach given in the book "Introduction to the Theory of Statistics" by Mood, Graybill and Boes, Mc Graw Hill, 1974 Third Edn. We first compute

$$E[X^k] = \frac{1}{B(a,b)} \int_0^1 x^k x^{a-1} (1-x)^{b-1}\,dx$$

$$= \frac{1}{B(a,b)} \int_0^1 x^{k+a-1} (1-x)^{b-1}\,dx$$

$$= \frac{B(k+a,b)}{B(a,b)} = \cancel{\frac{\Gamma(k+a+b)}{\Gamma(k+a)\Gamma(b)}} \cdot \frac{\cancel{\Gamma(a+b)}}{\Gamma(a)\Gamma(b)}$$

$$= \frac{\Gamma(k+a)\Gamma(b)}{\Gamma(k+a+b)} \cdot \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)}$$

$$\therefore\ E[X^k] = \frac{\Gamma(k+a)\Gamma(a+b)}{\Gamma(k+a+b)\Gamma(a)}$$

$\therefore$ For $k=1$

$$E[X] = \frac{\Gamma(a+1)\Gamma(a+b)}{\Gamma(a)\Gamma(a+b+1)}$$

$$= \frac{a\Gamma(a)\Gamma(a+b)}{\Gamma(a).(a+b)\Gamma(a+b)}$$

$$= \frac{a}{a+b}$$

Now we will set $k=2$

$$\therefore \quad E[X^2] = \frac{\Gamma(a+2)\,\Gamma(a+b)}{\Gamma(2+a+b)\,\Gamma(a)}$$

$$= \frac{(a+1)\,\Gamma(a+1)\,\Gamma(a+b)}{(a+b+1)\,\Gamma(a+b+1)\,\Gamma(a)}$$

$$= \frac{a\,(a+1)\,\Gamma(a)\,\Gamma(a+b)}{(a+b)\,(a+b+1)\,\Gamma(a+b)\,\Gamma(a)}$$

$$= \frac{a\,(a+1)}{(a+b)(a+b+1)}$$

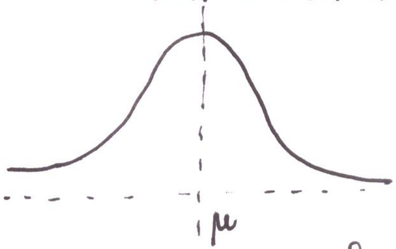
$$\mathrm{Var}(X) = \frac{a(a+1)}{(a+b)(a+b+1)} - \left(\frac{a}{a+b}\right)^2$$

$$= \frac{ab}{(a+b)^2\,(a+b+1)}$$

So we now see the advantage of computing $E[X^n]$

Section 3: Normal distribution

Normal distribution is the poster-boy of statistics. Even many members of the general public have heard about it and will immediately recognize the bell-shaped curve, which you see on the left. On a more formal level we say that a continuous random variable X has a normal distribution if it has a p.d.f



$$f_X(x) = \frac{1}{\sigma\sqrt{2\pi}}\, e^{-\frac{(x-\mu)^2}{2\sigma^2}}, \qquad x \in \mathbb{R}$$

Here μ and σ are the parameters, and we say ~~or rather~~ that X follows a normal distribution with ~~mean μ and~~ parameters μ, σ^2 and symbolically $X \sim N(\mu, \sigma^2)$.

It will turn out that μ and σ^2 are the ~~ra~~ mean and variance of a normal random variable. We shall call a continuous random variable, normal random variable, if it follows the normal distribution. Mind you, ~~mathematics~~ mathematicians do not like the term normal distribution. They want to use the term Gaussian distribution, after Karl, Federich Gauss, the great mathematician who introduced the idea ~~whi~~ of the normal curve ~~white~~ while tabulating astronomical data. Gauss was trying to measure the radius of the moon. As he repeated his measurement he got a different value (even if the difference was very less). When he plotted the values of his measurement he found that they seem to fall on a bell-shaped curve which the statisticians called as normal curve. The normal curve is also called an error curve as errors in measurements lie on the bell-shaped curve. We will say more on this later.

First let us show that f_X is a p.d.f. The fact that $f_X(x) \geq 0$ is obvious. Now observe that

$$\int_{-\infty}^{\infty} f_X(x)\,dx = \int_{-\infty}^{\infty} \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}\,dx$$

Let $\frac{x-\mu}{\sqrt{2}\,\sigma} = t \Rightarrow dx = \sqrt{2}\,\sigma\,dt$. Hence

$$\int_{-\infty}^{\infty} f_X(x)\,dx = \cancel{\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} e^{-}} \;\frac{1}{\sqrt{\pi}} \int_{-\infty}^{\infty} e^{-t^2}\,dt$$

$$= \frac{\sqrt{\pi}}{\sqrt{\pi}} = 1. \quad \left(\because \int_{-\infty}^{\infty} e^{-t^2}\,dt = 1, \text{ famous result in analysis}\right)$$

Of course that shows that f_X is indeed a p.d.f.

Suppose you do not know ~~to~~ the fact that $\int_{-\infty}^{\infty} e^{-t^2}\,dt = \sqrt{\pi}$.

Then in that case we can show that $\int_{-\infty}^{\infty} f_X(x)\,dx$; using the techniques double integral. This attempt is made in Mood, Graybill Boes, [Though I believe they tacitly use the above formula].

Theorem 6·1 : If $X \sim N(\mu, \sigma^2)$, Then

$$E[X] = \mu, \quad Var[X] = \sigma^2$$

and

$$m_X(t) = e^{\mu t + \frac{1}{2}\sigma^2 t^2}$$

Proof: You can of course directly compute the mean and the variance directly. However we will compute the $m_X(t)$ and through that compute μ & σ^2 through that.

$$\begin{aligned}
m_X(t) &= E\left[e^{tX}\right] \\
&= E\left[e^{tX - t\mu + t\mu}\right] \\
&= e^{t\mu} E\left[e^{t(X-\mu)}\right] \\
&\cancel{= e^{t\mu} \int_{-\infty}^{\infty} \frac{1}{\sigma\sqrt{2\pi}}} \\
&= e^{t\mu} \int_{-\infty}^{\infty} \frac{1}{\sigma\sqrt{2\pi}} e^{t(x-\mu)} e^{-\frac{1}{2}\frac{(x-\mu)^2}{\sigma}} dx \\
&= e^{t\mu} \int_{-\infty}^{\infty} \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2\sigma^2}\left[(x-\mu)^2 - 2\sigma^2 t^2 (x-\mu)\right]} dx
\end{aligned}$$

Now

$$\begin{aligned}
(x-\mu)^2 - 2\sigma^2 t^2 (x-\mu) &= (x-\mu)^2 - 2\sigma^2 t(x-\mu) + \sigma^4 t^2 - \sigma^4 t^2 \\
&= (x-\mu-\sigma^2 t)^2 - \sigma^4 t^2
\end{aligned}$$

$$\begin{aligned}
\therefore \quad m_X(t) &= e^{t\mu} e^{\frac{\sigma^2 t^2}{2}} \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{\left[(x-\mu-\sigma^2 t)^2\right]}{2\sigma^2}} dx \\
&= e^{t\mu + \frac{\sigma^2 t^2}{2}}
\end{aligned}$$

$$\because \quad \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{(x-(\mu+\sigma^2 t))^2}{2\sigma^2}} dx = 1 \quad \text{(Think why!!)}.$$

$$\therefore \quad m_X(t) = e^{t\mu + \frac{1}{2}\sigma^2 t^2}$$

$$m'_X(t) = (\mu + \sigma^2 t)\, e^{t\mu + \frac{1}{2}\sigma^2 t^2}$$

$$\boxed{E(X) = m'_X(0) = \mu}$$

$$m''_X(t) = \sigma^2 \left(e^{t\mu + \frac{1}{2}\sigma^2 t^2}\right) + (\mu + \sigma t)(\mu + \sigma t)\, e^{t\mu + \frac{1}{2}\sigma^2 t^2}$$

$$m''_X(0) = \sigma^2 + \mu^2 = E(X^2)$$

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

$$= \sigma^2 + \mu^2 - \mu^2$$

$$\boxed{Var(X) = \sigma^2}$$

If $\mu = 0$ and $\sigma^2 = 1$, ~~We~~ we say X follows a standard normal distribution, and its pdf is

$$\cancel{f_X(x) = \frac{1}{\sqrt{2\pi}} \int}$$

$$\boxed{f_X(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2}}$$

In fact if we ~~write th~~ consider $X \sim N(\mu, \sigma^2)$, then the

random variable $Z = \dfrac{X - \mu}{\sigma} \sim N(0,1)$.

The distribution function of a standard normal variable has a symbol $\Phi(x)$, ie

$$\Phi(x) = P(X \leq x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{1}{2}x^2}\, dx.$$

Now consider any $X \sim N(\mu, \sigma^2)$

then
$$P(X \leq x) = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{1}{2\sigma^2}(x-\mu)^2}\, dx$$

Now observe
$$P(X \leq x) = P\left(\frac{X-\mu}{\sigma} \leq \frac{x-\mu}{\sigma}\right) = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{1}{2\sigma^2}(x-\mu)^2}\, dx$$

Now $\frac{X-\mu}{\sigma} \sim N(0,1)$. Thus if $X \sim N(\mu, \sigma^2)$

Then
$$\cancel{F_X(x) = \Phi\left(\frac{x-\mu}{\sigma}\right)}$$

$$\boxed{F_X(x) = \Phi\left(\frac{x-\mu}{\sigma}\right)}$$

The values of the function Φ is tabulated and these tables are available.

$$\therefore \quad P(a \leq X \leq b) = F_X(b) - F_X(a).$$

$$\cancel{= \Phi\left(\frac{x-b}{\ }\right.}$$

$$\therefore \quad \boxed{P(a \leq X \leq b) = \Phi\left(\frac{b-\mu}{\sigma}\right) - \Phi\left(\frac{a-\mu}{\sigma}\right)}$$

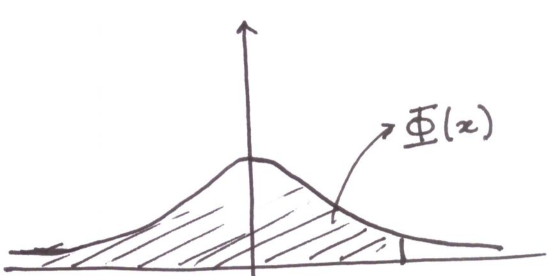


Fig A:

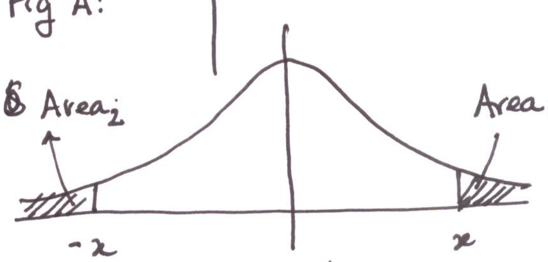


Fig B: Area 1 = Area 2 by symmetry.

For a more historical aside, let us note the De Moivre also had independently discovered the normal curve as a distribution of errors in measurement. Now observe that

$$\Phi(x) = P(X \leq x) = 1 - P(X > x)$$
$$= 1 - \int_{x}^{+\infty} f_X(x)\, dx$$

Now using Fig B we have
$$P(X > x) = P(X \leq -x)$$

Thus

$$\Phi(x) = 1 - P(X \le -x)$$ [Alaways keep in mind that $P(X=x)=0$]

$$= 1 - \Phi(-x)$$

$$\boxed{\Phi(x) = 1 - \Phi(-x)}$$

The binomial distribution which we know as a discrete distribution actually links up ~~beat~~ beautifully with normal distribution when the number of trials is very large. ~~and thus links up beautifully with the normal distrib~~

Thus the binomial distribution is a link between the discrete and the continuous. The following result asserts this

Theorem 8.1: Let $X \sim B(n,p)$, then

$$Z = \frac{X - np}{\sqrt{np(1-p)}}$$

approaches the standard normal distribution with mean zero and variance 1 as $n \to \infty$.

Proof: (A sketch only) How do we detect whether a given r.v. tends to some other distribution. Here we use the following approach. We show that as $n \to \infty$, the mgf of the binomial random variable will tend to the mgf of standard normal. We know from an assigment problem, that

- $m_{\frac{X+a}{b}}(t) = e^{\frac{a}{b}t} m_X\left(\frac{t}{b}\right)$ [See practice problem set No 2. Problem]

Since $X \sim B(n,p)$, we have

$$m_X(t) = (pe^t + (1-p))^n,$$ $1 - p = q$ (in the notes)

$\therefore$ For Z we have

$$m_Z(t) = e^{-\frac{np}{\sigma}t}\left[pe^{t/\sigma} + (1-p)\right]^n$$

where $\sigma = \sqrt{np(1-p)}$.

Let us set $np = \mu$. Then

$$m_Z(t) = e^{-\frac{\mu t}{\sigma}} \left[1 + p\left(e^{\frac{t}{\sigma}} - 1\right)\right]^n$$

In order to proceed to the limit we need to delink n & p since we will consider $n \to \infty$.

$$\cancel{\ln_Z t}$$

$$\therefore \quad \ln\left(m_Z(t)\right) = -\frac{\mu t}{\sigma} + n \ln\left[1 + p\left(e^{t/\sigma} - 1\right)\right]$$

$$= -\frac{\mu t}{\sigma} + n \ln\left[1 + p\left(\frac{t}{\sigma} + \frac{t^2}{2!\sigma^2} + \frac{1}{3!}\frac{t^3}{\sigma^3} + \cdots\right)\right]$$

Remember the series $\ln(1+x) = x - \frac{1}{2}x^2 + \frac{1}{3}x^3 - \frac{1}{4}x^4 + \cdots$ converges $|x| < 1$. Assuming that such conditions hold we have (Think why)

$$\ln\left(m_Z(t)\right) = -\frac{\mu t}{\sigma} + np\left[\frac{t}{\sigma} + \frac{1}{2!}\frac{t^2}{\sigma^2} + \frac{1}{3!}\frac{t^3}{\sigma^3} + \cdots\right]$$

$$\cancel{= -\frac{\mu t}{\sigma}} - \frac{np^2}{2}\left[\frac{t}{\sigma} + \frac{1}{2!}\frac{t^2}{\sigma^2} + \frac{1}{3!}\frac{t^3}{\sigma^3} + \cdots\right]^2$$

$$+ \frac{np^3}{3}\left[\frac{t}{\sigma} + \frac{1}{2!}\frac{t^2}{\sigma^2} + \frac{1}{3!}\frac{t^3}{\sigma^3} + \cdots\right] - \cdots$$

Now by ~~collection~~ collecting the powers of t we obtain

$$\cancel{\ln M_Z}$$

$$\ln\left(m_Z(t)\right) = \left(\frac{-\mu}{\sigma} + \frac{np}{\sigma}\right)t + \left(\frac{np}{2\sigma^2} - \frac{np^2}{2\sigma^2}\right)t^2$$

$$+ \left(\frac{np}{6\sigma^3} - \frac{np^2}{2\sigma^3} + \frac{np^3}{3\sigma^3}\right)t^3 + \cdots$$

$$= \frac{1}{\sigma^2}\left(\frac{np - np^2}{2}\right)t^2 + \left(\frac{np}{6\sigma^3} - \frac{np^2}{2\sigma^3} + \frac{np^3}{3\sigma^3}\right)t^3 + \cdots$$

$$= \frac{1}{\sigma^2} \cdot \frac{\sigma^2}{2} t^2 + \frac{n}{\sigma^3}\left[\frac{p - 3p^2 + 2p^3}{6}\right] t^3$$

$\therefore$ ~~ln~~

$$\therefore \ln(m_Z(t)) = \frac{1}{2}t^2 + \frac{n}{\sigma^3}\left[\frac{p-3p^2+2p^3}{6}\right]t^3 + \cdots\cdots$$

$$= \frac{1}{2}t^2 + \frac{n}{(\sqrt{np(1-p)})^3}\left[\frac{p-3p^2+2p^3}{6}\right]t^3 + \cdots\cdots$$

$$= \frac{1}{2}t^2 + \frac{n}{n^{3/2}\,p^{3/2}(1-p)^{3/2}}\left[\frac{p-3p^2+2p^3}{6}\right]t^3 + \cdots\cdots$$

So $$\ln m_Z(t) = \frac{1}{2}t^2 + \frac{1}{\sqrt{n}\,(p^{3/2}(1-p)^{3/2})}\left[\frac{p-3p^2+2p^3}{6}\right]t^3 + \cdots$$

We can now safely argue, that for $r > 2$ the coefficient of t^r will go to zero, as $n \to \infty$. Hence

$$\ln m_Z(t) \to \frac{1}{2}t^2 \text{ as } n \to \infty$$

as $\frac{1}{2}t^2$ is the mgf of the standard normal random variable.

~~What we~~ Theorem 8.1, is a special case of the central limit theorem we will learn later on in the course.

[The sketch of the proof given here is from the book:

John. E. Freund's Mathematical Statistics by Miller & Miller]

Pearson : 2014.]

—x—