4.3 The Normal Distribution

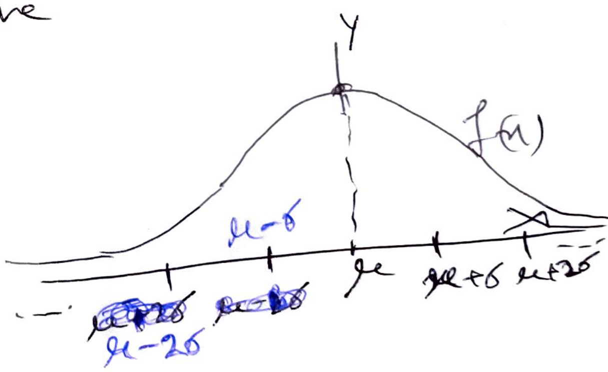
The pdf of normally distributed r.v. $X \sim N(\mu, \sigma^2)$ with mean μ and variance σ^2 is defined by

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

$$-\infty < x < \infty,$$
$$-\infty < \mu < \infty$$
$$\sigma > 0$$

* pdf of normal distribution is known as called normal curve.
* Normal distribution is also known as Gaussian distribution and the normal curve is called Gaussian curve or pdf.

Property of the normal curve



(a) The normal curve is symmetrical about y-axis.

(b) The normal curve is unimodal (mean, mode & median are same)

(c) $P[X \le \mu] = P[X \ge \mu] = 0.5$

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

(e) x-axis is assymptotic line of NC $y = f(x)$

Q. If $X \sim N(\mu, \sigma^2)$ is a normal r.v. ~~and~~ $Y = aX + b$, then prove that $a > 0$

Proof

The pdf of X is the normal curve def^n by

$$f_X(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}, \quad -\infty < x < \infty,\ -\infty < \mu < \infty,\ \sigma > 0$$

where $0 \leq f_X(x) \leq 1$.

Now to find pdf of Y:

Let $f_Y(y)$ be the pdf of r.v. $Y = aX + b$ and $F_Y(y)$ be the PDF of Y, then we have

$$f_Y(y) = F_Y'(y)$$

$$= \left[F_X\left(\frac{y-b}{a}\right)\right]'$$

$$= F_X'\left(\frac{y-b}{a}\right)\left(\frac{y-b}{a}\right)'$$

$$= f_X\left(\frac{y-b}{a}\right) \cdot \frac{1}{a}$$

$$= \frac{1}{a} f_X\left(\frac{y-b}{a}\right)$$

$$\boxed{\begin{aligned} F_Y(y) &= P[Y \leq y] \\ &= P[aX + b \leq y] \\ &= P\left[X \leq \frac{y-b}{a}\right] \\ &= F_X\left(\frac{y-b}{a}\right) \end{aligned}}$$

Since $f_Y(y)$ is a density function, so $f_Y(y) > 0$ iff $a > 0$

proved

Normal distribution fun

Let $X \sim N(\mu, \sigma^2)$ be the normal r.v with mean μ & variance σ^2, then the PDF of of X is called normal distribution function or Gaussian distribution function and is defined by

$$F_X(x) = P[X \le x] = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{1}{2}\left(\frac{u-\mu}{\sigma}\right)^2} du$$

$$= F_X(x; \mu, \sigma)$$

* Analysing of the normal curve with two changing parameters μ & σ is very critical, so to study the property of $F_X(x)$, we have to fix either μ or σ. The alternative approach is to remove the parameters with the help of another variable, i.e. r.v.

$$Z = \frac{X-\mu}{\sigma}$$

Standard Normal distribution

The normal distribution is

$$F_X(x) = P[X \le x]$$

Let $Z = \frac{X-\mu}{\sigma}$, then $X = \mu + \sigma Z$

Now $F_X(x) = P[X \le x] = P[\mu + \sigma Z \le x]$

$$= P\left[Z \le \frac{x-\mu}{\sigma}\right]$$

$$= F_Z\left(\frac{x-\mu}{\sigma}\right) = \phi\left(\frac{x-\mu}{\sigma}\right)$$

Mean of Z

$$E(Z) = E\left(\frac{X-\mu}{\sigma}\right) = \frac{1}{\sigma}E(X) - \frac{\mu}{\sigma}$$

$$= \frac{\mu}{\sigma} - \frac{\mu}{\sigma} = 0$$

variance of Z

$$V(Z) = V\left(\frac{X-\mu}{\sigma}\right) = V\left(\frac{1}{\sigma}X - \frac{\mu}{\sigma}\right)$$

$$= \frac{1}{\sigma^2}V(X) = 1 \quad (\because V(X) = \sigma^2)$$

The r.v. $Z = \frac{X-\mu}{\sigma}$ is called standard normal r.v. with mean $\mu_Z = 0$ and standard deviation $\sigma_Z = 1$, i.e. $Z \sim N(0,1)$ and $\varphi(z)$ is the standard normal distribution function.

pdf of Z

We have $F_X(x) = \varphi\left(\frac{x-\mu}{\sigma}\right)$

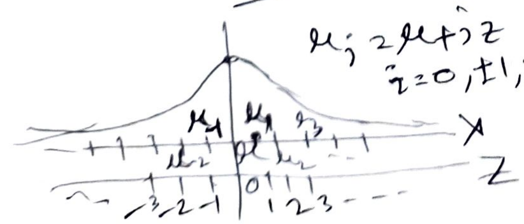
$= \varphi(z)$ where $z = \frac{x-\mu}{\sigma}$

Let $f_Z(z)$ be the pdf of standard r.v. Z.

then we have

$$\varphi'(z) = f_Z(z)$$

~~Now~~ Diff. w.r.t. z $\quad \varphi'(z) = (F_X(x))' = (F_X(\mu+\sigma z))'$



$$= f_X(\mu+\sigma z)\cdot\sigma$$

$$= \sigma \cdot \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{\mu+\sigma z-\mu}{\sigma}\right)^2}$$

$$\Rightarrow f_Z(z) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}z^2}, \quad -\infty < z < \infty$$

$$\Rightarrow f_Z(z) = \frac{1}{\sqrt{2\pi}} e^{-z^2/2}, \quad -\infty < z < \infty \text{ is the Std NC.}$$

<u>Note</u> $F_X(x) = P[X \le x] = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{1}{2}\left(\frac{u-\mu}{\sigma}\right)^2}$, $x < \infty$

$\Phi(z) = \Phi_Z(z) = P[Z \le z] = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{z} e^{-u^2/2} du$, $z < \infty$

Q: Prove that $\Phi(0) = 0.5$

<u>Proof</u>

$$\Phi(0) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{0} e^{-z^2/2} dz$$

$$= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{0} e^{-u} \cdot \frac{1}{\sqrt{2}} u^{-1/2} du$$

$\frac{z^2}{2} = u$

$\Rightarrow z\,dz = du$

$\Rightarrow dz = \frac{du}{z}$

$= \frac{1}{\sqrt{2u}} du$

$= \frac{1}{\sqrt{2}} u^{-1/2} du$

$z: -\infty \to 0$

$\Rightarrow u: -\infty \to 0$

$$= \frac{1}{2\sqrt{\pi}} \int_{-\infty}^{0} e^{-u} u^{-1/2} du$$

$$= 1 - \frac{1}{2\sqrt{\pi}} \int_{0}^{\infty} u^{-1/2} e^{-u} du$$

$$= 1 - \frac{1}{2\sqrt{\pi}} \int_{0}^{\infty} u^{\frac{1}{2}-1} e^{-u} du$$

$$= 1 - \frac{1}{2\sqrt{\pi}} \Gamma\left(\tfrac{1}{2}\right)$$

$\because \Gamma(n) = \int_0^\infty x^{n-1} e^{-x} dx$

$$= 1 - \frac{1}{2\sqrt{\pi}} \cdot \sqrt{\pi} = 1 - \frac{1}{2} = 0.5$$

<u>proved</u>

Q. Prove that $\varphi(-z) = 1 - \varphi(z)$

Proof We have

$$\varphi(z) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{z} e^{-u^2/2}\, du$$

Replacing z by $-z$, we get

$$\varphi(-z) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{-z} e^{-u^2/2}\, du$$

$$= \frac{1}{\sqrt{2\pi}}\int_{\infty}^{z} e^{-u^2/2}\,(-du)$$

Replacing u by $-u$

$$= \frac{1}{\sqrt{2\pi}}\int_{z}^{\infty} e^{-u^2/2}\, du$$

$$= 1 - \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{z} e^{-u^2/2}\, du$$

$$= 1 - \varphi(z)$$ proved

Q. Prove that $\varphi(0) = 0.5$ by the above result.

proof We have

$$\varphi(-z) = 1 - \varphi(z)$$

Taking $z = 0$, we get

$$\varphi(0) = 1 - \varphi(0)$$

$$\Rightarrow \varphi(0) = \frac{1}{2} = 0.5$$

proved

Note (a) $P[a < Z < b] = \varphi(b) - \varphi(a)$

and

$$P[a < X < b] = F_X(b) - F_X(a)$$

$$= \varphi\left(\frac{b-\mu}{\sigma}\right) - \varphi\left(\frac{a-\mu}{\sigma}\right)$$

(b) $P[Z \le z] = \varphi(z) = F_X(\mu + \sigma z)$

$$= P[X \le \mu + \sigma z]$$

where $Z \sim N(0,1)$

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

$\because F_X(x) = \varphi\left(\frac{x-\mu}{\sigma}\right)$

Replacing x by $\mu + \sigma z$, we get

$$F_X(\mu + \sigma z) = \varphi\left(\frac{\mu + \sigma z - \mu}{\sigma}\right) = \varphi(z)$$

$\Rightarrow P[X \le \mu + \sigma z] = P[Z \le z]$

(c) Since $X = \mu + \sigma Z$, so

(100p)th percentile of X

$= \mu + (100p)^{th}$ percentile of $Z \times \sigma$

(d) Z_α is $100(1-\alpha)^{th}$ percentile of Z, i.e. $1-\alpha = \varphi(z)$ for some $z \in Z$

(e) A discrete r.v. X is approximated to normal r.v. $X^* \sim N(\mu_X, \sigma_X^2)$ with 0.5 error correction

$P[X \le x]$

$\sim P[X^* \le x + 0.5]$

$\simeq \varphi\left(Z \le \frac{x + 0.5 - \mu_X}{\sigma}\right)$

$Z = \frac{X^* - \mu_X}{\sigma_X}$

$P[X \ge x] \sim P[X^* \ge x + 0.5] \simeq \varphi\left(Z \ge \frac{x + 0.5 - \mu_X}{\sigma_X}\right)$

* If $X \sim Bin(n, p)$ is a binomial r.v. with no. of trials n and prob. of success p, then

$$b(x; n, p) = \binom{n}{x} p^{x} q^{n-x}, \quad 0 \le x \le n$$
$$p + q = 1$$

& $$B(x; n, p) = P[X \le x] = \sum_{y \le x} \binom{n}{y} p^{y} q^{n-y}$$

We have $\mu_X = np$ and $\sigma^2 = npq$

$$\Rightarrow \sigma = \sqrt{npq}$$

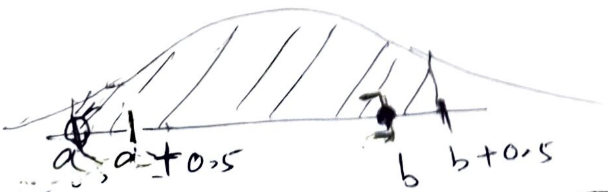
$$F_X(x) = P[X \le x] = B(x; n, p)$$
$$\simeq \phi\left(\frac{x + 0.5 - np}{\sqrt{npq}}\right)$$

and $$P[X > x] = 1 - P[X \le x]$$
$$= 1 - B(x; n, p)$$
$$\simeq 1 - \phi\left(\frac{x + 0.5 - np}{\sqrt{npq}}\right)$$

Limit Thm of De Moivre of Laplace

* If $X \sim B(n, p)$ is the binomial r.v. with $np \ge 10$ & $nq \ge 10$ then $X \sim Z$ with 0.5 error correction mean $\mu_X = np$ & SD $\sigma_X = \sqrt{npq}$,

then $$P[a < X \le b] \simeq F_X(b) - F_X(a)$$
$$= \phi\left(\frac{b + 0.5 - np}{\sqrt{npq}}\right) - \phi\left(\frac{a + 0.5 - np}{\sqrt{npq}}\right)$$



Note: $P[a < X \le b] \simeq P[a + 0.5 < X < b + 0.5]$

Ex let $X \sim B(50, 0.25)$, then find

ⓐ $P[X \leq 10]$

ⓑ $P[5 \leq X \leq 15]$

using normal distribution if condition satisfied.

Sol^n Given $X \sim B(50, 0.5)$, i.e. $n = 50$, $p = 0.25 \Rightarrow q = 0.75$

Since $np = 50 \times 0.25 = 12.5 \geq 10$

and $nq = 50 \times 0.75 = 37.5 \geq 10$,

so $X \sim X^*$ with 0.5 error correction

where $Z \approx \frac{X^* - \mu_X}{\sigma_X}$, $X^* \sim N(\mu_X, \sigma_X^2)$

We have $\mu_X = np = 12.5$

$\sigma_X^2 = npq = 50 \times 0.25 \times 0.75$

$= 9.375$

$\Rightarrow \sigma_X = 3.0619$

ⓐ $P[X \leq 10] = F_X(10) \approx \varphi\left(\frac{10 + 0.5 - 12.5}{3.0619}\right)$

$= \varphi(-0.6532)$

$= 1 - \varphi(0.6532)$

ⓑ $P[5 \leq X \leq 15] = F_X(15) - F_X(4)$

$\approx \varphi\left(\frac{15 + 0.5 - 12.5}{3.0619}\right) - \varphi\left(\frac{4 + 0.5 - 12.5}{3.0619}\right)$

$= \varphi(0.9798) - \varphi(-2.6128)$

$= \varphi(0.9798) - (1 - \varphi(2.6128))$

$= 0.8364 - 1 + 0.9955 \approx 0.8319$

$\boxed{\begin{array}{l} P[a \leq X \leq b] \\ = F_X(b) - F_X(a-1) \\ \text{for } X = \text{discrete} \end{array}}$