Lecture - 14

Probability & Statistics.

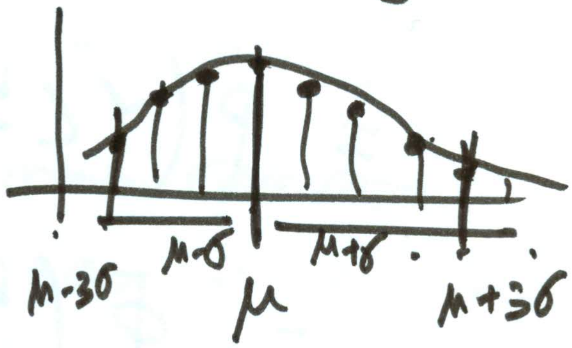
'Normal' (random variable) distribution

Continuous random variable whose pdf

$$P(X=x) = f(x) = \frac{1}{\sqrt{2\pi}\,\sigma} e^{-(x-\mu)^2/2\sigma^2}, \quad -\infty < x < \infty$$

$\sigma > 0$ — $f(x)$ becomes a density fn.

Discovered in 1733 — in order to approximate binomial distribution when n is large.



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

$$N(0,1) \sim Y = \frac{X-\mu}{\sigma} = -\frac{\mu}{\sigma} + \frac{1}{\sigma}X$$

$$\equiv a + bX$$

unit normal distn.

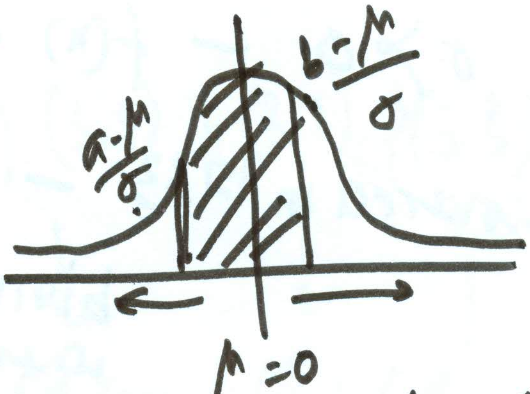
standard normal dist.

$E(Y) = 0$, $var(Y) = 1$

$$P(a < X < b)$$

$$= P\left(\frac{a-\mu}{\sigma} < \frac{X-\mu}{\sigma} < \frac{b-\mu}{\sigma}\right)$$

$$= P\left(\frac{a-\mu}{\sigma} < Y < \frac{b-\mu}{\sigma}\right)$$

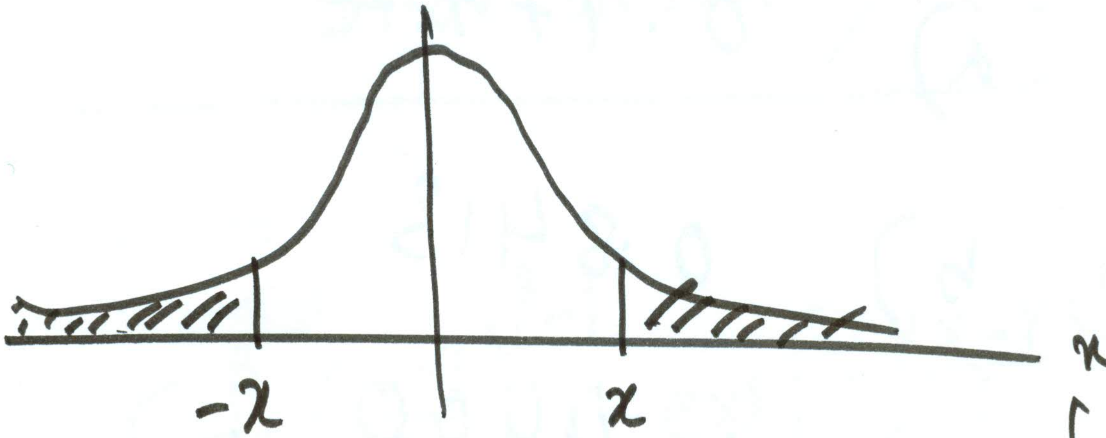


$$= P\left(Y < \frac{b-\mu}{\sigma}\right) - P\left(Y < \frac{a-\mu}{\sigma}\right)$$

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

where Φ denotes the CDF of $Y \sim N(0,1)$

But Φ is not known still now.



$$\Phi(x) = P(X < x) = \int_{-\infty}^{x} f(t)\,dt$$

$$\Phi(-x) = P(X \le -x)$$

$$= 1 - P(X < x)$$

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

<u>Prob.</u> If X is a normal random variable with mean $\mu = 3$ and variance $\sigma^2 = 16$.

$X \sim N(\mu, 16)$

Find a) $P(X < 11)$

b) $P(X > -1)$

c) $P(2 < X < 7)$

Ans.

a) 0.9772

b) 0.8413

c) .4400.

$P(X > -1)$

$= P\left(\frac{X-3}{4} > \frac{-1-3}{4}\right)$

$= P(Y > -1)$

$= P(Y < 1)$

$= 0.8413$

Weibull Distribution / RV.

PDF is

$$f(x) = \begin{cases} \frac{\beta}{\delta}\left(\frac{x-\gamma}{\delta}\right)^{\beta-1} \exp\left[-\left(\frac{x-\gamma}{\delta}\right)^{\beta}\right], & x \geq \gamma \\ 0, & \text{otherwise.} \end{cases}$$

$-\infty < \gamma < \infty \mapsto$ called the location parameter

$\delta > 0 \mapsto$ scale parameter

$\beta > 0 \mapsto$ the shape parameter

(Reliability engineering)

— $E(X) = \gamma + \delta\,\Gamma\left(1+\frac{1}{\beta}\right)$

$Var(X) = \delta^2 \left\{ \Gamma\left(1+\frac{2}{\beta}\right) - \left(\Gamma\left(1+\frac{1}{\beta}\right)\right)^2 \right\}$

The CDF $F(x) = 1 - \exp\left[-\left(\frac{x-\gamma}{\delta}\right)^{\beta}\right]$

$x \geq \gamma$

The CDF of Weibull distn. can be obtained in <u>excel</u> fn.

WEIBULL(x, β, γ, TRUE)

Function of random variable.

$$Y = a + bX$$

$$Y = g(X)$$

$$E[Y] = \int_{-\infty}^{\infty} g(x)\, f(x)\, dx.$$

Theorem is concerned with PDF of a function of a continuous r.v.

<u>Theorem</u>. If X is a cont. r.v. with pdf f_X s.t. $f_X > 0$ for $x \in (a, b)$ and if $y = H(x)$ is a continuous fn strictly increasing or strictly decreasing fn of x, then the random variable $Y = H(X)$ has density fn

$$f_Y(y) = f_X(x) \left| \frac{dx}{dy} \right|$$

with $x = H^{-1}(y)$ expressed in terms of y. (If H is increasing then $f_Y(y) > 0$ if $H(a) < y < H(b)$; and if H is decreasing then $f_Y(b) > 0$ if $H(b) < b < H(a)$.)

<u>Pr8B</u>. Let $f_X(x) = \begin{cases} x/8, & 0 \le x \le 4 \\ 0, & \text{otherwise.} \end{cases}$

and $H(x) = 2x + 8$ which is (cont. & strictly increasing)

Then find $f_Y(y)$ using the theorem <u>mentioned before</u>.

$$f_Y(y) = \begin{cases} \frac{y}{32} - \frac{1}{4}, & 8 \le y \le 16 \\ 0, & \text{otherwise} \end{cases}$$

$f_Y(y) = f_X(x) \left|\frac{dx}{dy}\right|$

$= \frac{y-8}{16} \times \frac{1}{2}$

$= \frac{y}{32} - \frac{1}{4}$

$y = 2x + 8$

$\Rightarrow x = \frac{y-8}{2}$

> 0 then $\Rightarrow \frac{dx}{dy} = \frac{1}{2}$

$0 < \frac{y-8}{2} < 4$

$\Rightarrow 8 < y < 16$

Jointly distributed random variables.

In an experiment into the possible causes of cancer, we might be interested in relationship between the average # of cigarettes smoked daily and the age at which an individual contracts cancer.

$$F(x,y) = P(X \le x, Y \le y)$$

$$F_X(x) = P(X \le x)$$
$$= P(X \le x, Y \le \infty)$$
$$= F(x, \infty)$$

lly, $F_Y(y) = F(\infty, y)$

~~Let~~ ~~An~~

Let X & Y be discrete random variables.

$$\boxed{P(X = x_i, Y = y_j) = f(x_i, y_j)}$$

Suppose $f(x_i, y_j)$ is known,

Q. How to find $P(X = x_i)$.

$$\{X = x_i\} = \bigcup_j \{X = x_i, Y = y_j\}$$

$$P\{X = x_i\} = P\left(\bigcup_j \{X = x_i, Y = y_j\}\right)$$

$$= \sum_j P\{X = x_i, Y = y_j\}$$

$$= \sum_j P(x_i, y_j)$$

Hly $P(Y = y_j) = \sum_i P(x_i, y_j)$

Note: The reverse is not true.

Prob. Suppose that 3 batteries are randomly chosen from a group of 3 new, 4 used but still working, and 5 defective batteries. If we let X & Y denote respectively the # of 'new' and 'used but still working' batteries that are chosen, then the joint PMF of X & Y is

$$\phi(i,j) = P(X=i,\ Y=j)$$

$$= \frac{\binom{3}{i}\binom{4}{j}\binom{5}{3-i-j}}{\binom{12}{3}}$$

$i \backslash j$	0	1	2	3	$p(X=x_i)$
0	$p(0,0)$	$p(0,1)$	$p(0,2)$	$p(0,3)$	→ row sum
1	$p(1,0)$	$p(1,1)$	.	.	→ row sum
2	.				
	.				
3	$p(3,0)$	$p(3,1)$	$p(3,2)$	$p(3,3)$	→ row sum
$p(Y=j)$	↓ col. sum	↓ col. sum	↓ col. sum.	↓ col. sum.	