Tutorial on Normal Distribution
(one-dimesional normal)

Dear students,

We try to solve some problems related to normal distribution.

Recall that if $X \sim N(\mu, \sigma^2)$ then

$$f_X(x) = \frac{1}{\sqrt{2\pi}\,\sigma} e^{-\frac{1}{2\sigma^2}(x-\mu)^2}, \quad \begin{array}{l} -\infty < x < \infty \\ -\infty < \mu < \infty \\ 0 < \sigma < \infty. \end{array}$$

$\mu = E(X) \rightarrow$ mean value of normal dist^n

$\sigma^2 \rightarrow V(X) \rightarrow$ variance of normal dist^n.

What is the CDF of a $N(\mu, \sigma^2)$ dist^n

It is $\boxed{F_X(x) = \Phi\left(\frac{x-\mu}{\sigma}\right)}$ —— (1)

but then what does the function $\Phi(\cdot)$ represent.

Let us recall some properties of this func^n.

For that you need to discuss

"Properties of $N(0,1)$ distribution,

that is standard normal dist^n".

So let $Z \sim N(0,1)$ then

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

the cdf is

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

$$= \Phi(z)$$

So we are able to specify Φ function of Equation 1. The Φ func^n is the cdf of a standard normal distribution defined as

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

look at this nice representation of Equation 2 when $z > 0$

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

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

$$= \Phi(0) + P(0 < Z \le z) \quad \text{--- (3)}$$

what is the value of $\Phi(0)$. see next page.

$$\boxed{\begin{aligned}&\Phi(-z) = 1 - \Phi(z) \quad \forall\ -\infty < z < \infty \\ &\Phi(0) = 1/2 .\end{aligned}}$$

So Equation 3 becomes

$$\boxed{\Phi(z) = P(Z \le z) = 0.5 + P(0 < Z \le z)}$$ here z is positive.

Now a simple Conclusion from Equation 1 is that "CDF of any $N(\mu, \sigma^2)$ distribution can be written using standard normal CDF."

So if you want to compute probability for any $N(\mu, \sigma^2)$ you need to take help of $N(0,1)$ CDF. For example, if you want to compute $P(a < X \le b)$ where $X \sim N(\mu, \sigma^2)$ then proceed as follows: $\left[X \sim N(\mu, \sigma^2) \Rightarrow \frac{X-\mu}{\sigma} \sim N(0,1) \right]$

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

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

$$= \Phi\left(\frac{b-\mu}{\sigma}\right) - \Phi\left(\frac{a-\mu}{b}\right), \quad \text{--- } 4$$

Formula: If Y is a rv then $P(a < Y \le b) = F_Y(b) - F_Y(a)$

Can you guess what does equation 4 say??

It says that $P(a < X \leq b)$ where $X \sim N(\mu, \sigma^2)$

is given by $\underbrace{\Phi\left(\frac{b-\mu}{\sigma}\right)} - \underbrace{\Phi\left(\frac{a-\mu}{\sigma}\right)}$.

Both are standard normal CDF quantity.

$$\boxed{\text{So to compute probabilities for any } N(\mu, \sigma^2) \text{ You have use standard normal CDF } \Phi(z).}$$

Next we discuss a very nice example ~~real~~ related to standard normal CDF.

Ex: If $\Phi(z)$ is the CDF of a standard normal $N(0,1)$ random variable Z the compute following

(i) for $0 \leq a < b$, $P(a < Z \leq b)$

(ii) for $a \leq b < b$, $P(a < Z \leq b)$

(iii) $a \leq b < 0$, $P(a < Z \leq b)$

(iv) $a > 0$, $P(-a < Z \leq a)$

(v) $P(0 < Z < 1.5)$, $P(-2.33 < Z < 2.33)$
$P(Z \leq 1.69)$, $P(Z \geq -1.25)$

(vi) Compute 10^{th}, 25^{th}, 45^{th}, 65^{th}, 75^{th} percentile of a $N(1,4)$ distribution.

Solution: (i) $0 \le a < b$ given then

$$P(a < Z \le b) = \Phi(b) - \Phi(a)$$

(ii) $a \le 0 < b$ given. then

$$P(a < Z \le b) = \Phi(b) - \Phi(a) = \Phi(b) - (1 - \Phi(-a))$$
$$= \Phi(b) + \Phi(-a) - 1$$

(iii) $a \le b < 0$ given, then

$$P(a < Z \le b) = \Phi(b) - \Phi(a)$$
$$= 1 - \Phi(-b) - (1 - \Phi(-a))$$
$$= \Phi(-a) - \Phi(-b)$$

(iv) $a > 0$ given, $P(-a < Z \le a) = \Phi(a) - \Phi(-a)$

$$= \Phi(a) - (1 - \Phi(a))$$
$$= 2\Phi(a) - 1$$

[Note: these four parts cover events of all types for which you want to compute probability]

(V) Now we have to use table.

$$P(0 < Z < 1.5) = 0.4332$$

$P(-2.33 < Z \le 2.33)$

$= 2\,\Phi(2.33) - 1$

$= 2\left[\Phi(0) + P(0 < Z \le 2.33)\right] - 1$

$= 2(0.5 + 0.4901) - 1$

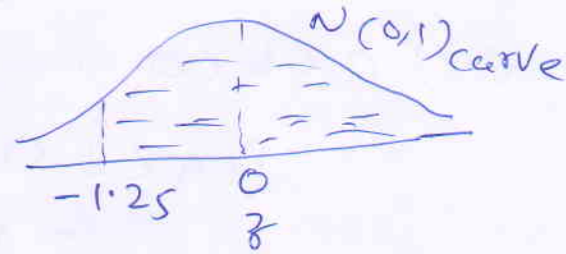
$= 0.9802$

$P(Z \le 1.69) = \Phi(1.69)$

$= \Phi(0) + P(0 < Z \le 1.69)$

$= 0.5 + 0.4545$

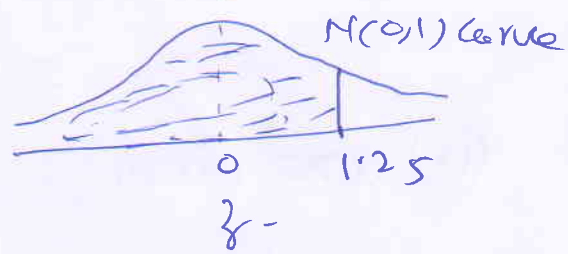
$= 0.9545$



$P(Z \geq -1.25)$

$= P(Z \le 1.25)$

(∵ standard normal symmetric about 0).



$\Phi(1.25) = \Phi(0) + P(0 < Z < 1.25)$

$= 0.5 + 0.3944$

$= 0.8944$

(till now whatever probabilities we have computed in part (v), all are for $N(0,1)$ distribution).

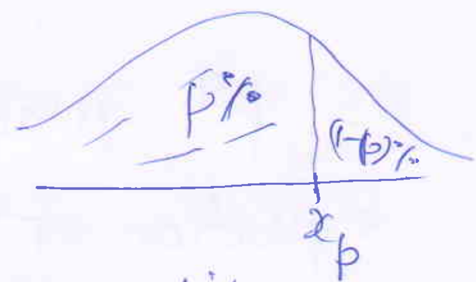
Let us try to work out part (vi). Before that one query 'Do you remember how to compute quantiles of a continuous prob-distribution'

Please note that

'for a continuous random variable X, its pth $(0<p<1)$ quantile is given by the equation

$$P(X \le x_p) = p$$

x_p: pth quantile



So for continuous case pth quantile x_p is the solution to the following equation

$$\boxed{F_X(x_p) = p}$$

In particular if X is $N(0,1)$ then corresponding pth quantile is obtained as

$$\boxed{\Phi(z_p) = p}$$

here z_p is the pth quantile of $N(0,1)$ distribution.

Let us now solve part (vi).

Also note that if $X \sim N(\mu, \sigma^2)$ then to find its pth quantile proceed as follows:

First write the CDF of X as

$$P(X \leq x) = P\left(\frac{X-\mu}{\sigma} \leq \frac{x-\mu}{\sigma}\right)$$

$$= P\left(Z \leq \frac{x-\mu}{\sigma}\right) = \Phi\left(\frac{x-\mu}{\sigma}\right)$$

So pth quantile x_p is the solution of equation

$$\boxed{\Phi\left(\frac{x_p-\mu}{\sigma}\right) = p} \quad \text{——— } (*)$$

pth quantile x_p.

(vi) $X \sim N(1, 4)$. CDF of X is

$$F_X(x) = \Phi\left(\frac{x-1}{2}\right) \quad \text{——— } (**)$$

Let us find 75th percentile. So in this case $p = 0.75$. From Equation (**) we have to solve $\Phi\left(\frac{x_{0.75}-1}{2}\right) = 0.75$ ——— (***)

[how to get $x_{0.75}$]

write (***) as

$$\Phi\left(\frac{x_{0.75}-1}{2}\right) = 0.75 = 0.5+0.25$$

now look at prob 0.25 in the normal table and find corresponding z. [$\frac{x_{0.75}-1}{2}$ will be that value of z].

From table $z = \cancel{0.0987}\ 0.68$

$$\therefore \frac{x_{0.75}-1}{2} = \cancel{0.0987}\ 0.68$$

$$\Rightarrow x_{0.75} = \cancel{2\times 0.0987+1}$$
$$= 0.68\times 2+1 = 1.36+1 = 2.36.$$

So 75th percentile of $N(1,4)$ dist^n is 2.36.
(that is $P(X \le 2.36) = 0.75$).

Let us compute 25th percentile of the $N(1,4)$ distribution. Following Equation (***) we now need to find $x_{0.25}$ such that

$$\Phi\left(\frac{x_{0.25}-1}{2}\right) = 0.25$$

we rewrite this as

$$1 - \Phi\left(\frac{1 - x_{0.25}}{2}\right) = 0.25 \quad \left[\begin{array}{l}\text{using} \\ \Phi(z) = 1 - \Phi(-z)\end{array}\right]$$

$$\Rightarrow \Phi\left(\frac{1 - x_{0.25}}{2}\right) = 0.75 = 0.5 + .25$$

like previous part

$$\frac{1 - x_{0.25}}{2} = 0.68$$

$$\Rightarrow x_{0.25} = -2 \times 0.68 + 1 = -1.36 + 1$$

$$= -0.36$$

So 25th percentile of $N(1,4)$ distribution is -0.36. [That is $P(X \leq -0.36) = 0.25$]

Next Find 45th percentile:

Here defining equation is

$$\Phi\left(\frac{x_{0.45} - 1}{2}\right) = 0.45$$

$$\Rightarrow 1 - \Phi\left(\frac{1 - x_{0.45}}{2}\right) = 0.45$$

$$\Rightarrow \Phi\left(\frac{1 - x_{0.45}}{2}\right) = 0.55 = 0.5 + 0.05$$

From table we get $\frac{1 - x_{0.45}}{2} = 0.13$

$$1 - x_{0.45} = 2 \times 0.13 = 0.26$$

$$\Rightarrow x_{0.45} = 1 - 0.26 = 0.74$$

$\therefore$ 45th percentile of $N(1,4)$ dist^n is 0.74.

Ex: Let heights X of 800 students are $N(66,25)$.
Find the number of students with heights
(i) between 65 and 70 inches
(ii) greater than or equal to 72 inches (try yourself).

Sol^n: $X \sim N(66,5)$

(i) $P(65 < X \le 70) = P\left(\frac{65-66}{5} < \frac{X-66}{5} \le \frac{70-66}{5}\right)$

$$= P\left(-\tfrac{1}{5} < Z \le \tfrac{4}{5}\right) = P(-0.2 < Z \le 0.8)$$

$$= \Phi(0.8) - \Phi(-0.2)$$

$$= \Phi(0.8) - (1 - \Phi(0.2))$$

$$= \Phi(0.8) + \Phi(0.2) - 1$$

$$= [\Phi(0) + P(0 < Z \le 0.8) + \Phi(0) + P(0 < Z < 0.2)] - 1$$

$$= [0.5 + 0.2881 + 0.5 + 0.0793] - 1$$

$= 0.3674$. So total no. of students is $= 800 \times 0.3674 = 294$.

Ex: A soft drink machine is regulated so that it discharges an average of 200 ml per cup. let amount of drink is normally distributed with standard deviation 15 ml.

(i) what ~~cup of~~ fractions of cup will contain more than 224 ml.

(ii) what is prob. that a cup contains between 191 and 209 ml.

(iii) Below what value do we get the smallest 25% of the drinks.

Ex: steel rods are manufactured to be 3 inches in diameter but they are acceptable if they are inside the limits 2.99 to 3.01 inches. It is observed that 5% are rejected as oversize and 5% are rejected as undersize Assuming that diameter are normally distributed find the mean and standard deviation of the distribution. Hence evaluate what would be the proportion of rejects if the permissible limits are widened to 2.985 inches and 3.015 inches