Study material ①

Paper – Probability & Statistics (MA2011)

Section 4.3 : Normal distribution.

pdf is given by

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

$$\sigma > 0 \text{ and } -\infty < \mu < \infty.$$

If $\mu = 0$ & $\sigma = 1$, then the above distribution is called standard normal distribution.

It's density function is denoted by $f(z)$, defined by,

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

Now, probability $P(a \le Z \le b)$ is given in a table, called standard normal table. From that table we can find very easily using the following symmetric curve.

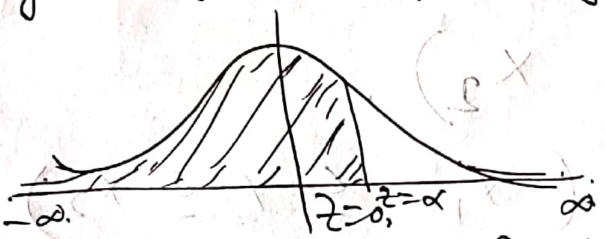

Table may be provided in different way. Some table may provide probability upto $z = \alpha$, when $\alpha \geq 0$ i.e., $P(Z \le \alpha)$ is given in table. Since $\alpha \geq 0$, naturally every value given in the table must be $\geq .5$.

Now if you need $P(Z \le -\alpha)$, then first find $P(Z \geq \alpha)$.

$$P(Z \geq \alpha) = 1 - P(Z \le \alpha).$$

Again from symmetry you see,

$$P(Z \le -\alpha) = P(Z \geq \alpha) = 1 - P(Z \le \alpha).$$

Thus you find $P(Z \le -\alpha)$.

Suppose you need to find $P(-\alpha \leq Z \leq \alpha)$. Then find in the following way:

$$P(0 \leq Z \leq \alpha) = P(Z \leq \alpha) - 0.5$$
$$= P(-\alpha \leq Z \leq 0)$$
(From symmetry)

$$\therefore P(-\alpha \leq Z \leq \alpha) = 2P(Z \leq \alpha) - 1.$$

This is obtained from Table.

Thus we find this value.

Now we connect the standard normal distribution to general normal distribution (μ, σ).

Suppose, we have a problem dealing with a normal distribution where mean $= \mu$ and standard deviation $= \sigma$ and we need $P(X_1 \leq X \leq X_2)$.

Now observe $P(X_1 \leq X \leq X_2)$

$$= P\left(\frac{X_1 - \mu}{\sigma} \leq \frac{X - \mu}{\sigma} \leq \frac{X_2 - \mu}{\sigma}\right)$$

$$= P(Z_1 \leq Z \leq Z_2)$$

[Just we renamed our events]

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

Now, we know that if $Z = \frac{X-\mu}{\sigma}$, then

$$E(Z) = 0, \quad \sigma_Z = 1.$$

Hence, for this transformation, Z becomes standard normal distribution and correspondingly Z_1 and Z_2 as found above, we note down. Then using standard normal table we find $P(z_1 < Z < z_2)$.

Thus we find $P(x_1 \leq X \leq x_2)$ with the help of standard normal table.

The normal distribution and discrete population

The normal distribution is often used as an approximation to the distribution of values in a discrete population. In such situations, extra care should be taken to ensure that probabilities are computed in an accurate manner.

The method is like this:

To find $P(a \leq X \leq b)$ in a discrete distribution:

Actually we will find $P(a - 0.5 \leq X \leq b + 0.5)$ to have more accurate value of the prob. by normal distribution.

Application:

Approximation of Binomial distribution to normal distribution.

When $n \geq 20$, we can prove that

$P(a \leq X \leq b)$ obtained using normal distribution with same mean $\mu = np$ and same standard deviation $\sigma_x = \sqrt{np(1-p)}$ is approximately equal to the original discrete distribution. Now more accurate value will be obtained if we consider

$$P(a - 0.5 < X < b + 0.5).$$

[explanation given in your book, page no. 161, Fig 4.25]

Now using standard normal table, we find

$$P(a \le X \le b)$$

$$\simeq P(a-0.5 \le X \le b+0.5)$$

$$\simeq P\left(\frac{a-0.5-np}{\sqrt{np(1-p)}} \le \frac{X-np}{\sqrt{np(1-p)}} \le \frac{b+0.5-np}{\sqrt{np(1-p)}}\right)$$

$$= P(z_1 \le Z \le z_2), \text{ with}$$

$$Z = \frac{X-np}{\sqrt{np(1-p)}},$$ which is standard normal variable.

(From hard copy) Now we start solving problems.

(Soft copy also) Prob. 28 & 29 we have discussed in the class room. Now prob. No. 31.

31 (From hard copy) (Soft copy 37) (a) Here $z = \dfrac{x-\mu}{\sigma} = \dfrac{105-104}{5} = 0.2$ | Given, $\mu = 104$, $\sigma = 5$, $x = 105$.

Now $P(X < 105) = P(Z < 0.2) = 0.5793$

(obtained from standard normal table)

Since probability at any particular point in a continuous distribution is always zero, therefore

$P(X \le 105)$ is also 0.5793.

Also $P(X = 105) = 0$.

(b) x is 1 standard deviation from the mean.

$$x = \mu \pm \sigma.$$

$$z = \frac{x-\mu}{\sigma} = \frac{\mu \pm \sigma - \mu}{\sigma} = \pm 1.$$

Now the required probability here is $P(|x| \geq \mu \pm \sigma)$.

$$P(|x| > \mu+\sigma) = P(x > \mu+\sigma)$$
$$= P\left(\frac{x-\mu}{\sigma} > 1\right)$$
$$= P(z > 1).$$

$$P(|x| > \mu-\sigma) = P(\cancel{\sigma}\, x \cancel{\geq} < \mu-\sigma)$$
$$= P(z < -1).$$

$$\therefore P(|x| \geq \mu \pm \sigma) = P(z < -1 \text{ or } z > 1)$$
$$= 2P(z < -1)$$
$$= 2 \times 0.1587 = 0.3174.$$

Note: The prob. is not dependent on the values of μ and σ.

(c) The most extreme 0.1% of chloride concentration values are the lowest 0.05% of the chloride values and the heighest 0.05% of the chloride values.

$P(z > z_1) = 0.0005$ [0.005% in prob.^{ty} $= 0.0005$]

$\therefore z_1 = 3.29.$

Also, $P(z < -z_1) = 0.0005.$

Now $z_1 = \frac{x_1 - \mu}{\sigma} \Rightarrow x_1 = \mu + z_1 \sigma.$

$\Rightarrow x_1 = 104 + 5 \times 3.29 = 120.45.$

$\therefore P(x > 120.45) = 0.0005$.

Similarly

$$\frac{x_2 - \mu}{\sigma} = -3.29$$

$\Rightarrow x_2 = 104 - 5 \times 3.29 = 87.55$.

Thus the most extreme 0.1 % of chloride concentration values are below 87.55 mmol/L and above 120.45 mmol/L.

33 (a) $\mu = 46.8$, $\sigma = 1.75$.

[Hard copy & soft copy] $P(x \le 50) = P\left(Z \le \frac{50 - 46.8}{1.75}\right) = P(Z \le 1.83)$
$= 0.9664$.

(b) $P(X \geq 48) = P\left(Z \geq \frac{48 - 46.8}{1.75}\right) = P(Z \geq 0.69) =$

$= 1 - \Phi(0.69) = 1 - 0.7549 = 0.2451$.

(c) $|x| \geq \mu \pm 1.5\sigma$

Now $P(|X| \geq \mu \pm 1.5\sigma)$

$= P(Z \le -1.5 \text{ or } Z \geq 1.5)$

Now $P(Z \geq 1.5) = 1 - P(Z \le 1.5)$

$= 1 - 0.9332 = P(Z < -1.5)$

$\therefore P(-1.5 < Z < 1.5) = 0.9332 - (1 - 0.9332)$

$= 1.8664 - 1 = 0.8664$.

[mind that here the words are "at most".

So $-1.5 < Z < 1.5$]

[In problem 31, the words were "by more than"].

[Suggested exercise problems:

34 – 57.]

41 (both Hard copy & soft copy).

$X \rightarrow$ opening altitude of the parachute.

$\therefore$ $\mu = 200$, $\sigma = 30$. (Given)

~~Let~~ It is given that equipment damage will occur if parachute opens below 100 m.

Hence the probability of equipment damage occuring can also be represented as $P(X < 100)$.

~~standard~~

$$X < 100$$

$$\Rightarrow \frac{X - 200}{30} < \frac{100 - 200}{30} \quad \text{[events are same transformed]}$$

$$\Rightarrow Z < -3.33$$

$$\therefore P(X < 100) = P(Z < -3.33) = 0.0004$$

Hence, P(parachute suffers damage) = 0.0004.

$\therefore$ P(parachute doesn't suffer damage) = 1 − 0.0004
= 0.9996.

Let E be the event that there is equipment damage to the payload of at least one of five independently dropped parachutes.

We first find P(none out of 5 cases suffers ~~do~~ damage)

= P(all out of 5 cases doesn't suffer damage)

$$= (0.9996)^5 = 0.998$$

$\therefore$ using $P(\bar{E}) = 1 - P(E)$ = 1 − 0.998 = 0.002

47. (Hard copy)

55. (Soft copy)

$X \rightarrow$ No. of ~~A randomly selected~~ drivers wears a seat belt.

A random sample of 500 drivers is selected.

So $p = 3/4$ [As given 75% drivers wear seat belt].

$n = 500$.

$q = 1 - p = 1/4$. $\therefore np = 500 \times \frac{3}{4} = 375 > 10$.

$\therefore \sigma^2 = npq = 500 \times \frac{3}{4} \times \frac{1}{4} = \frac{1500}{16} = \frac{375}{4}$

$\therefore \sigma = \sqrt{\frac{375}{4}}$

Now we use normal approximation to binomial distribution.

$P(360 \le X \le 400)$

$= P(359.5 \le X < 400.5)$

$= P(-1.60 < Z < 2.63)$

$= P(Z < 2.63) - P(Z < -1.60)$

$= 0.9957 - 0.0548$ [from standard normal table]

$= 0.9409$

$Z_1 = \frac{x_1 - \mu}{\sigma}$

$= \frac{359.5 - 375}{\sqrt{\frac{375}{4}}}$

$\simeq -1.60$

$Z_2 = \frac{x_2 - \mu}{\sigma}$

$= \frac{400.5 - 375}{\sqrt{\frac{375}{4}}}$

$\simeq 2.63$

2nd part:

$P(X < 400) = ?$

The z-score is the value (using the continuity correction) decreased by the mean np and divided by the standard deviation $\sqrt{npq} = \sqrt{np(1-p)} = \sqrt{\frac{375}{4}}$.

$$z = \frac{x - np}{\sqrt{np(1-p)}} = \frac{399.5 - 375}{\sqrt{\frac{375}{4}}} \approx 2.53$$

$\therefore P(X < 400) = P(X < 399.5) = P(Z < 2.53)$

$= 0.9943$.

Solve some problems from your book.

— x —