Unit 14: The Normal Distribution and its Applications

14.1 Normal distribution

Review of continuous random variable, for examples:

(i) heights and weights of adults

(ii) length and width of leaves of the same species

(iii) distance jumped by a C grade boy in several times

(iv) distance jumped by a group of C grade boys

(v) actual weights of rice in 5 kg bags sold in supermarkets

For large samples, say 1,000 or even 10,000 observations, the distribution of the above examples will look bell-shaped and approximately follow the normal distribution with probability function:

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

where μ is the mean and σ^2 is the variance of the distribution. The notation $N(\mu, \sigma^2)$ should be mentioned.

The standard normal distribution is a particular case of the normal distribution with $\mu = 0$, $\sigma = 1$. The probability function is :

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

14.2 Normal curve and standard normal curve

The diagrams below show the changes in the normal distribution curves as μ varies and as σ varies:

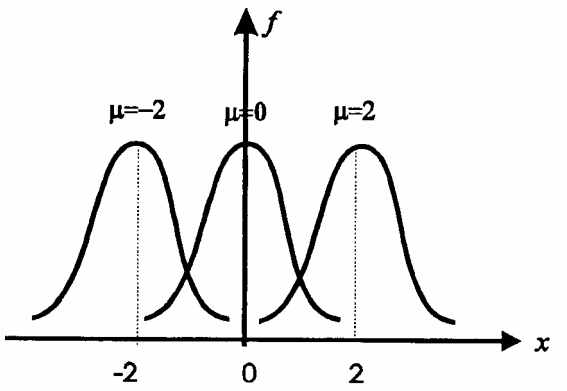

Fig. 14.2(i) Normal distribution with σ fixed ($\sigma = 1$)

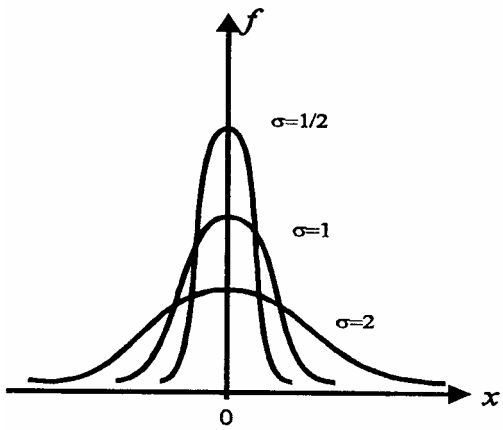

Fig. 14.2(ii) Normal d1stribution with μ fixed ($\mu = 0$)

Important properties of the normal curve:

(i) the curve is bell-shaped and symmetrical about the mean;

(ii) the mean, the median, mode are all equal;

(iii) the spread of the curve is determined by the value of σ;

(iv) the area under the curve is 1.

Standard normal curve

All normal distribution can be reduced to the standard normal distribution by using the transformation $Z = \dfrac{X - \mu}{\sigma}$ where $X \sim N(\mu, \sigma^2)$ and $Z \sim N(0, 1)$.

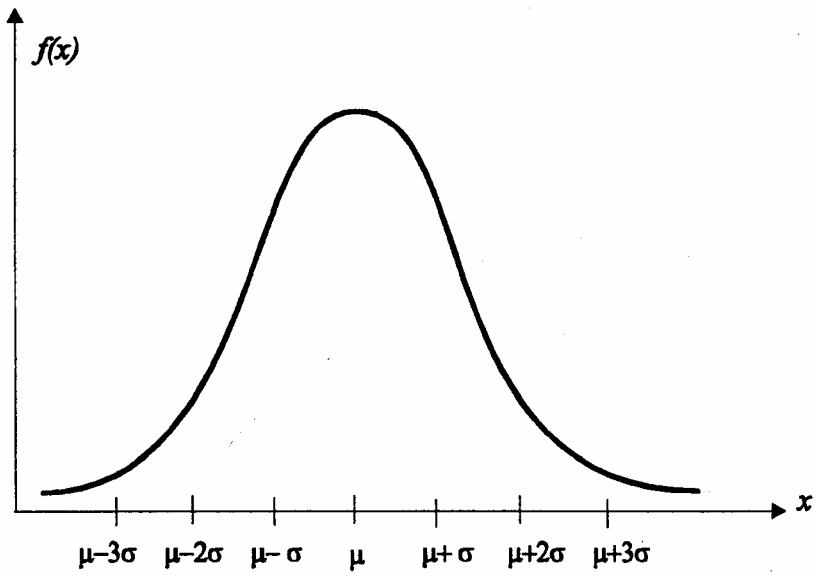

Fig. 14.2 (iii) Normal Curve

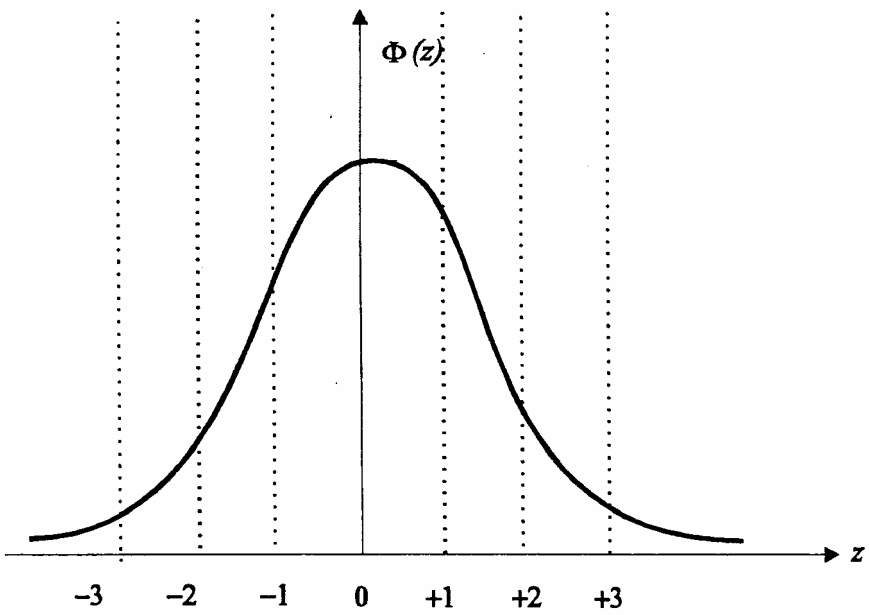

Fig. 14.2 (iv) Standard Normal Curve

14.3 Normal Table

Students are expected to know the values of the area under the standard normal curve given by

$$A(z) = P(0 < Z < z)$$
$$= \int_0^z \frac{1}{\sqrt{2\pi}} e^{\frac{-t^2}{2}} \, dt$$

which are tabulated in the normal table.

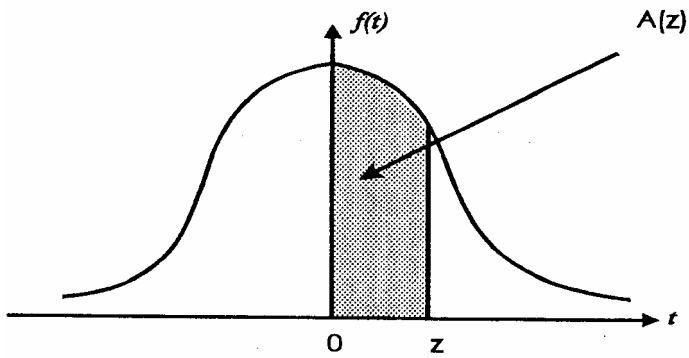

Fig. 14.3 (i)

For Z in N (0, 1), students are taught to use the table to find values like $P(Z > a)$, $P(Z < a)$, $P(Z \le b)$ and $P(a \le Z \le b)$ to solve practical problems. However, they should be reminded to transform those distributions in $N(\mu, \sigma^2)$ to $N(0, 1)$ before referring to the table.

Given $P(Z > a) = \alpha$ or $P(Z < b) = \beta$, where α is a known value. Students are expected to know how to find out the values of a and b.

Use of standard normal variable Z to find probabilities of the $N(\mu, \sigma^2)$ are expected.

Besides, students are expected to know how to work out the following problems. Given that $X \sim N(\mu, \sigma^2)$,

(i) if $\sigma^2 = 16$ and $P(X > 120) = 0.05$, find the value of μ;

(ii) if $\mu = 20$ and $P(15 \le X \le 25) = 0.99$, find the value of σ;

(iii) if $P(X > -10) = 0.72$ and $P(X < 0) = 0.90$, find the values of μ and σ.

14.4 Applications of normal distribution

<u>Example 1</u>

If the time a student stays in a classroom follows the normal distribution with $\mu = 6$ h and $\sigma = 0.1$ h, what is the probability that he stays in a classroom for less than 5 hours?

<u>Example 2</u>

A factory produces TV sets by process A or by process B. Assuming that the life-time distribution of the sets is normal. The sets produced by process A have mean life-time of

86,000 h with standard deviation 8,000 h, whereas those produced by process B have mean lifetime of 75,000 h with standard deviation 6,000 h Sets with a life-time of less than 30,000 hare regarded as definitely bad. Which of the two processes will produce more acceptable sets?

Example 3

A factory produces apple juice contained in a bottle of 1.5 litres. However, due to random fluctuations in the automatic bottling machine, the actual volume per bottle varies according to a normal distribution. It is observed that 10% of bottles are under 1.45 litres whereas 5% contain more than 1.55 litres. Calculate the mean and standard deviation of the volume distribution.

Example 4

Monthly food expenditure for families of two in Hong Kong is on the average $6,000 and has a standard deviation $300. Assuming that the monthly food expenditure are normally distributed.

(i) What is the probability that monthly food expenditure are less than $5,500?

(ii) What is the probability that monthly food expenditure are between $5,500 and $7,000 ?

(iii) What is the probability that monthly food expenditure are less than $5,500 or greater than $7,000 ?

(iv) Determine the values of lower quartile Q_1 and upper quartile Q_3 .

(v) Write a report regarding your findings.