POLI 443
Applied Political Research

Session 2 – Areas Under the Normal Curve

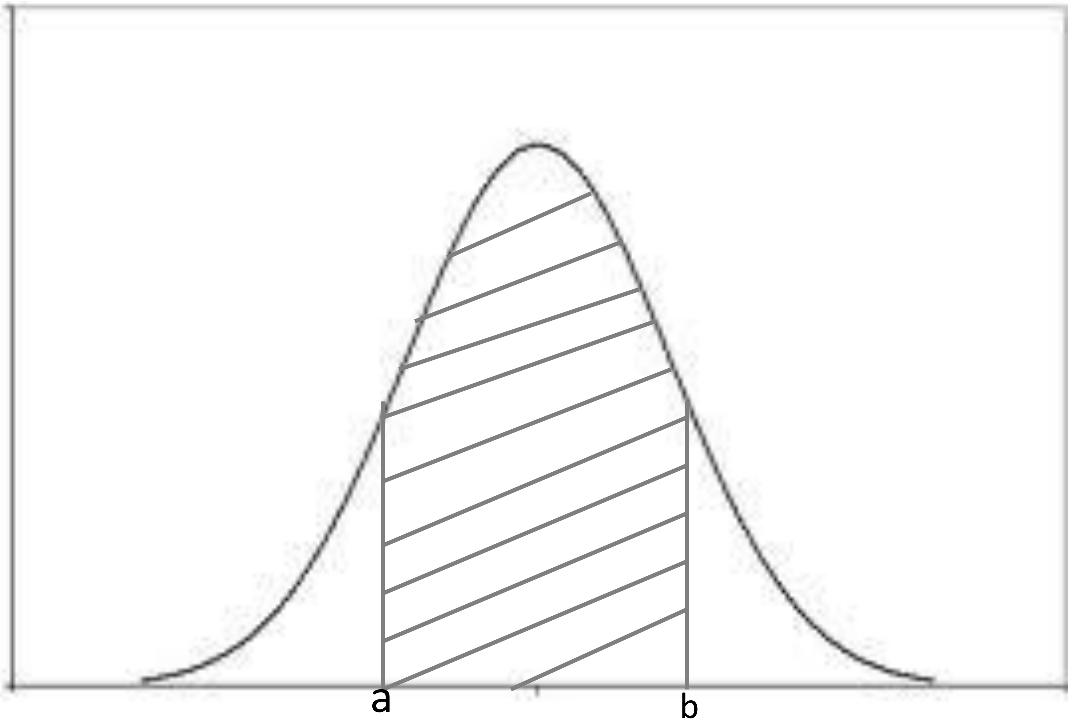
Lecturer: Prof. A. Essuman-Johnson, Dept. of Political Science
Contact Information: aessuman-johnson@ug.edu.gh

UNIVERSITY OF GHANA
College of Education
School of Continuing and Distance Education
2014/2015 – 2016/2017

Areas under the Normal Curve

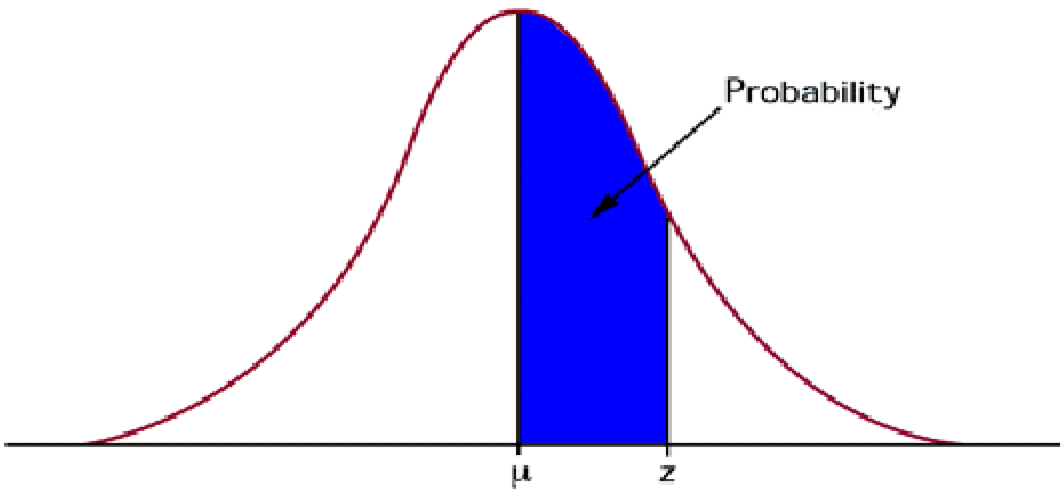
- To make inferences about a population we need to know the probability of a sample. We need therefore to be able to calculate the probability that a random variable with a normal probability distribution will fall within a specified interval e.g. if X is a random variable which is normally distributed with mean (μ) and standard deviation (s) we might need to know the probability that X lies in an interval between points a and b.

Fig 4



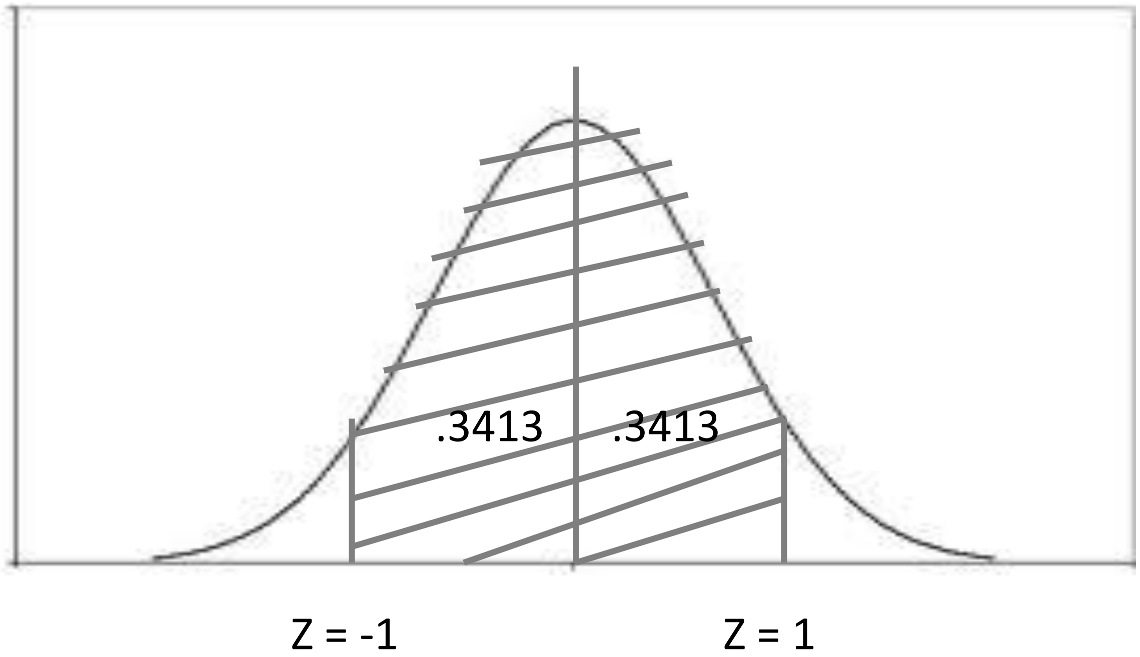
- The above shaded portion of the curve corresponds to the probability that X will lie between a and b. The desired area can be obtained by using a table of areas under the normal curve (ref. to normal table). The tables give areas under the normal curve between the mean and a measurement which is Z standard deviations to the right of the mean (μ).

Fig 5



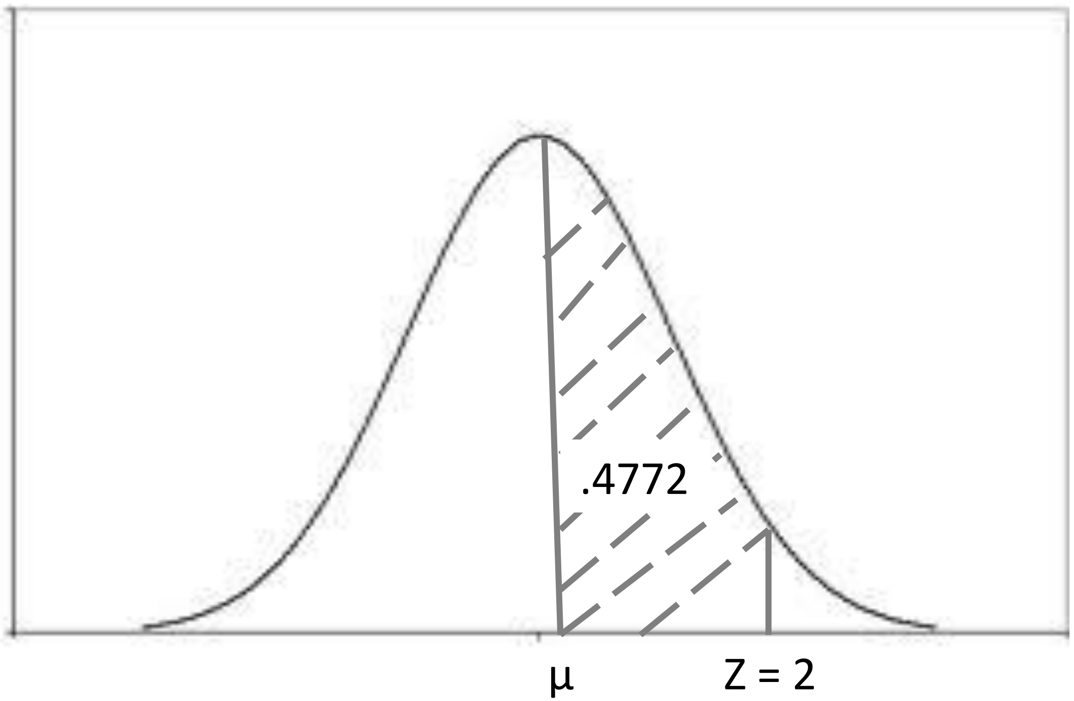
- Areas to the left of the mean need not be tabulated because the normal curve is symmetrical about the mean. Thus the area between the mean and a point z (sd) to the right is the same as the area between the mean a similar point z(sd) to the left of the mean (refer to normal distribution tables and work some examples with students) e.g. the area 1 sd to the right of the mean (i.e. $Z = 1$) is 0.3413. But the area within 1 sd of the mean (i.e. $Z = +/- 1$) = 0.6826 or approximately 68%.

Fig 6



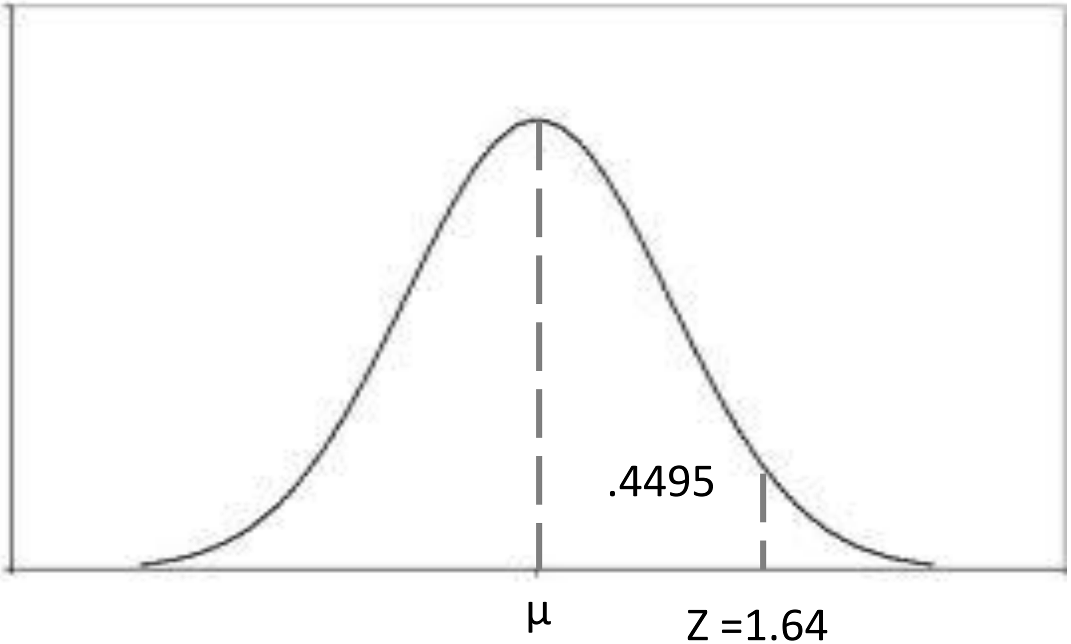
- Similarly the area between the mean and a point Z = 2 sd to the right of the mean is shown in the fig. below:

Fig 7



The area within approximately 2 sd of the mean = Z = +/- 2 = 2(0.4772) = 0.9544 or approximately 95%. Suppose we wish to find the area corresponding to Z = 1.64

Fig 8



)/s = 11-8/2 = 1.5

To determine how many standard deviations (s) a measurement X lies from the mean we use the standard score

$$Z = X - (\mu)/(\delta) \quad \text{or} \quad X - (\overline{x})/s$$

To calculate the area under the normal curve between the mean (mu) and a specified value X to the right of the mean, we first find Z i.e. the number of standard deviations that X lies from (mu) (standardize X) using the expression:

$$Z = \overline{x} - (\mu)/s$$

We then refer to the normal distribution table and obtain the entry corresponding to the calculated value of Z. This entry is the desired area or probability under the normal curve between (mu) and the value of X. Social scientists refer to Z as the Z score

Illustration 1

Suppose that X is a normally distributed random variable with mean 8 and standard deviation of 2. Find the probability that X lies between 8 and 11 i.e. what fraction of the total area lies under the curve between 8 and 11?

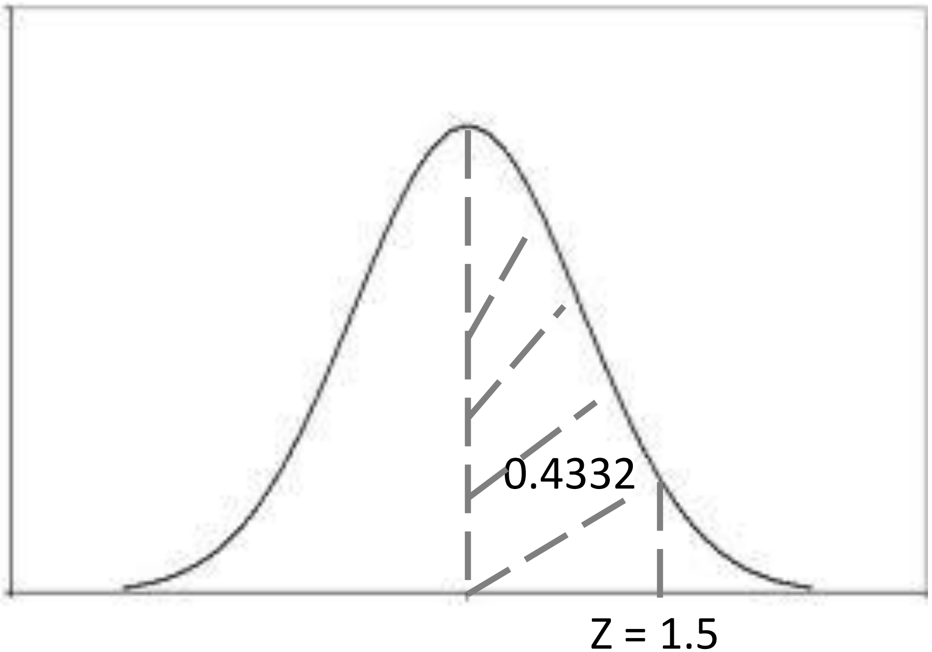
Solution

Compute the Z value for 11

From N distribution tables find the area under N at the Z value

$Z = X - (\overline{x})/s = 11\text{-}8/2 = 1.5$

Fig 9



Required area is the shaded area. From N Distribution tables Z = 1.5 = 0.4332. Therefore, the probability that X lies between 8 & 11 is 0.4332.

Illustration 2

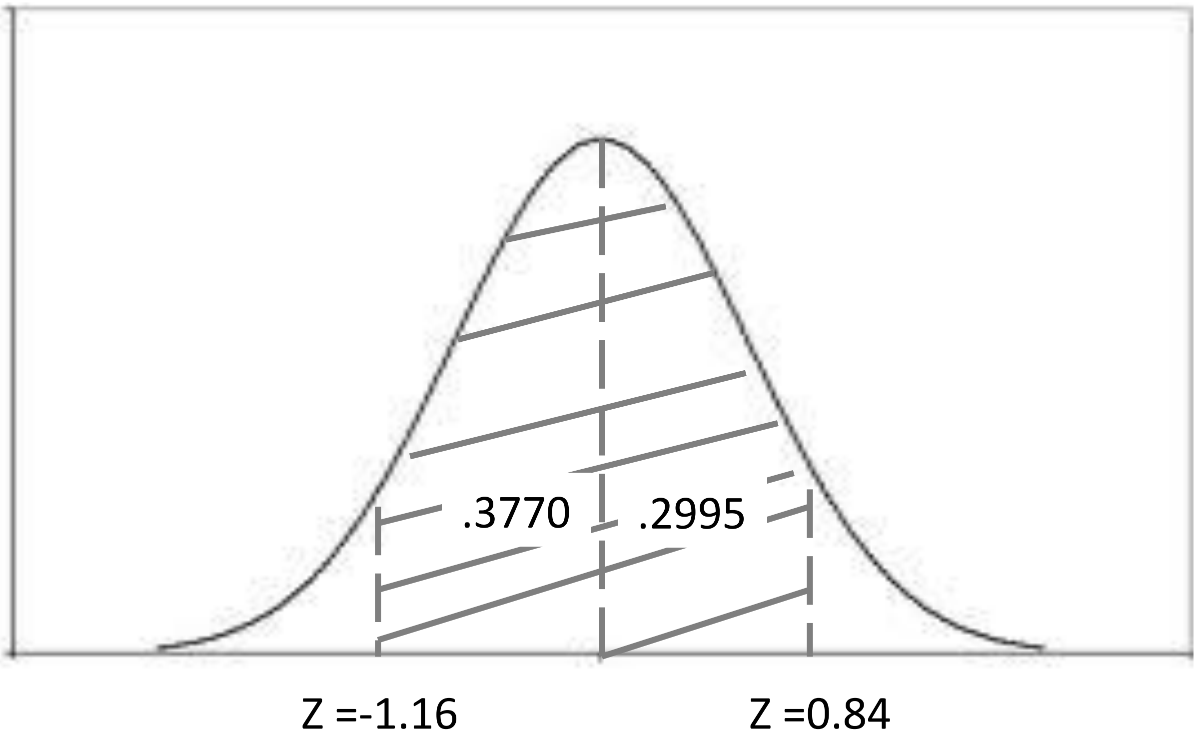
In a sample of 1000 students, the mean score of an examination was 14.4 with standard deviation of 2.5. Assuming the distribution is normal (a) How many students scored between 12 & 16 (b) How many scored above 18 and below 8 (c) what are the chances that any student selected at random will score above 15?

Solution

(a) Students scoring between 12 and 16 will actually score any mark from 11.5 to 16.5

$$Z(11.5) = X - (\bar{x})/s = 11.5 - 14.4/2.5 = -2.9/2.5 = -1.16$$

$$Z(16.5) = X - (\bar{x})/s = 16.5 - 14.4/2.5 = 2.1/2.5 = 0.84$$



Required area = Area between Z = -1.16 & Z = 0.84

= Area between Z = -1.16 & Z = 0 + (Z = 0 & Z = 0.84)

= 0.3770 + 0.2995

= 0.6765

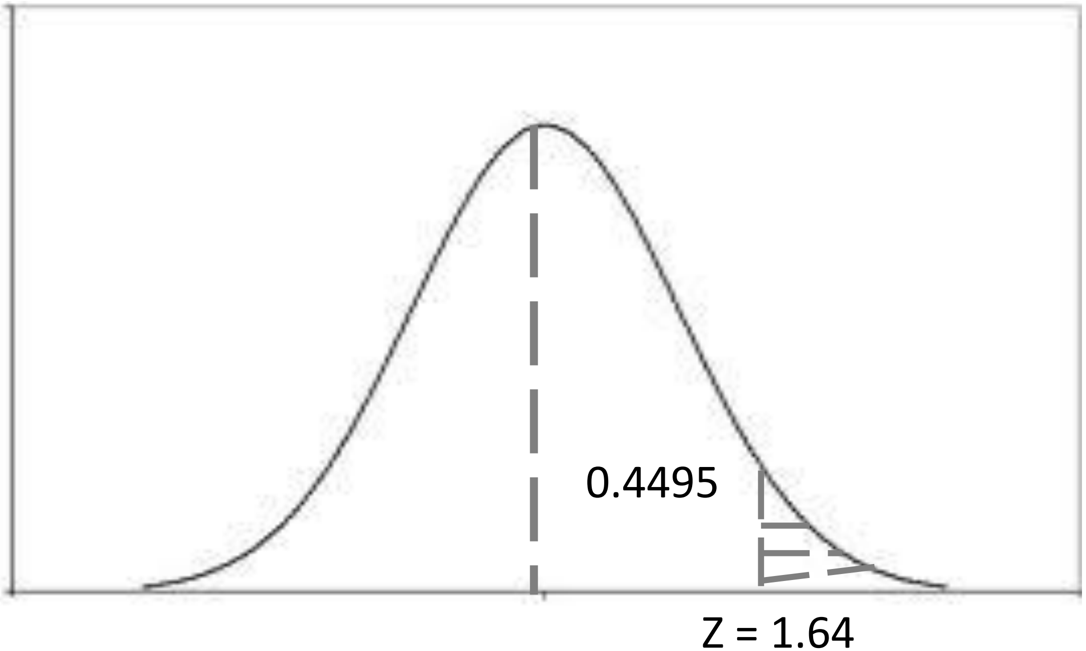
Therefore the number of students scoring between 12 & 16 = 0.6765 of 1000 = 677

Students scoring above 18 would at least score 18.5 and those scoring below 8 would score at least 7.5

$Z (18.5) = X – ((\overline{x})/s = 18.5 – 14.4/2.5 = 4.1/2.5 = 1.64$

The proportion of students scoring above 18 is shown in the diagram below:

Fig 11



Area shaded: From normal distribution table

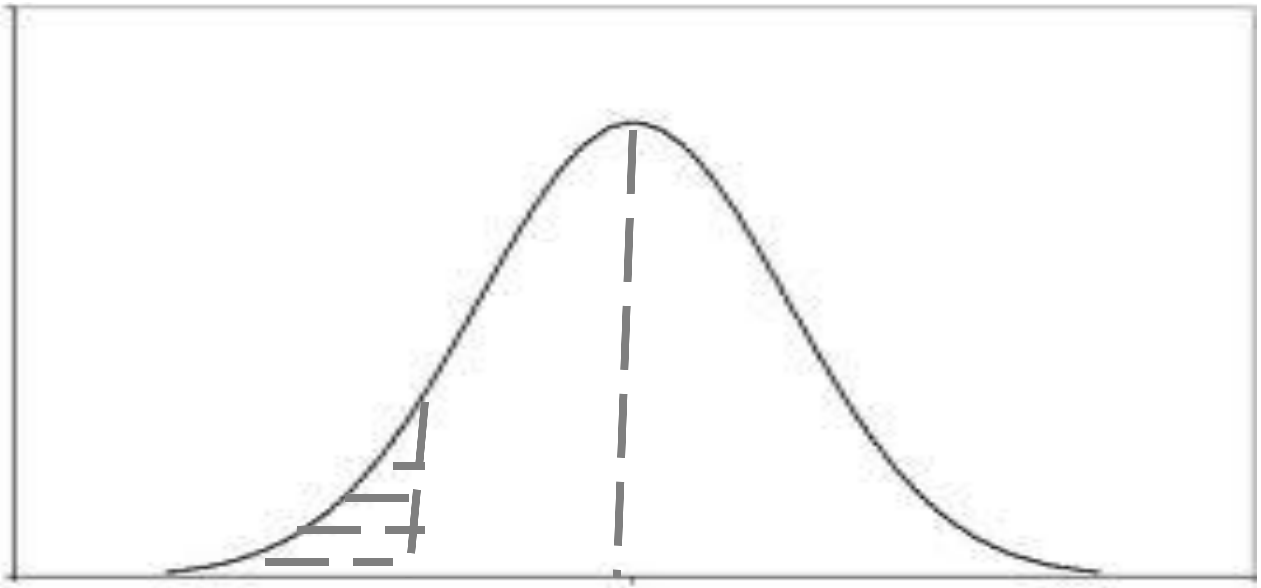
Z = 1.64 = 0.4495

The required area = 0.5000 – 0.4495 = 0.0505 = 5.05%

Therefore the number of students scoring above 18 = 0.0505 of 1000 = 50.5 = 51

Z (7.5) = X – ($\bar{x}$)/s = 7.5 – 14.4/2.5 = -6.9/2.5 = - 2.76

Fig 12



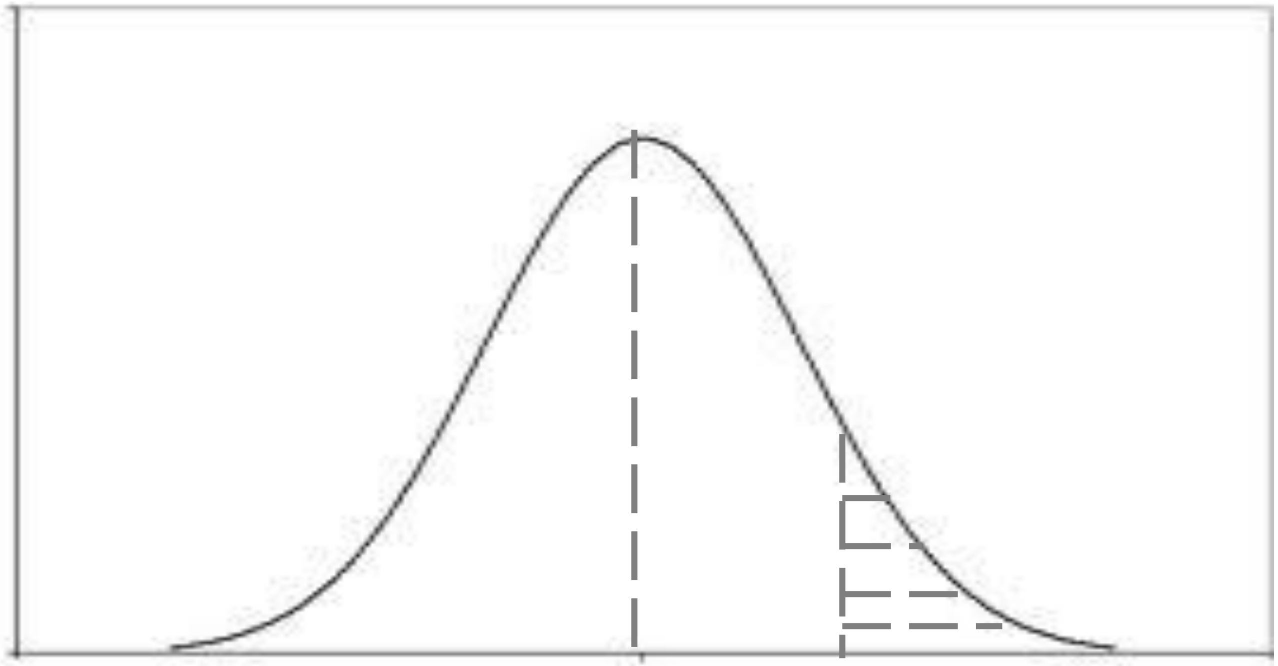
Z = 2.76

From Normal tables Z = 2.76 = 0.4971. The shaded area = 0.5000 – 0.4971 = 0.0029 = 29%. Therefore the number of students scoring below 8 = 0.29 of 1000 =2.9 = 3

(C)The upper limit of 15 is 15.5

Z (15.5) = X – ($\overline{x}$)/s = 15.5 – 14.4/2.5 = 1.1/2.5 = 0.44

Fig 13



Z = 0.44

From normal tables Z = 0.44 = 0.1700. The required area therefore is 0,5000 – 0,1700 = 0.3300. Therefore the chances that a student selected at random will score above 15 is 0.33.

THANK YOU