|  |
| --- |
| **Mathematics for Data Science 2023/2024** |

**Week 3- Standard Normal Distribution**

**Solution**

**Answer the following questions**

Q1. In the given figures, we have three normal curves.

In the first curve the mean is 0 and the standard deviation is 1.

The second curve has the same mean, 0, but a standard deviation of 2.

Can you see what the mean and standard deviation are for the third curve?

Diagram

Description automatically generated

*Solution μ = 1 and σ = 1.*

Q2. The figure given below shows two normal distribution curves representing the time taken to prepare personal (“S”) and business (“A”) income tax returns:

(a) Which has the larger mean?

(b) Which has the larger standard deviation?

A picture containing text, sky

Description automatically generated

Solution: a. The mean of distribution “S” is about 2.5 hours, while the mean of distribution

“A” is about 5.5 hours. Therefore distribution “A” has the larger mean.

b. The normal distribution “A” is flatter or more spread out so has the larger standard

deviation.

Q3. Let X = number of children in a household in Sydney. Suppose we take random samples of size 2 from the above parent population, that is we randomly select 2 households at a time and count the number of children in each household. The diagrams in Figure 14 represent the outcomes from six samples. So, for example, in sample 1, one household had no children the other had 2 children.

Calendar

Description automatically generated

1. Find the sample mean in each case and mark it on the diagram.
2. Draw the distribution of X for these six samples. Use the same scale on the axis

as above.

1. Based on your data above what do you guess is the mean number of children in an Australian household? That is, estimate μX from the data. What could you do to improve your estimate?
2. The sample means are 1, 2, 1.5, 1.5, 1, 1.5
3. Scatter chart

   Description automatically generated
4. We can see from a. that the values of X jump around from sample to sample.

Our best estimate of μX is μX. We estimate μX as 1.42.

To improve the estimate, increase the number of samples and the sample size

*Q4.Chart, diagram, histogram

Description automatically generated*Study carefully the diagram of the normal curve given in Figure 16 and then complete the table using the percentages given.

Solution:

*Table

Description automatically generated*

Q5. What is the area between the mean and 0.85 standard deviations below the mean (ie

between z scores of −0.85 and 0)?

Solution: Now, because the normal curve is symmetrical, the area we want is equal to the area

under the curve between 0 and +0.85. We look up that area in our table.

A picture containing text

Description automatically generated

The required area is 0.3023 or 30.23%.

Q6. What is the area between z scores of 0.33 and 1.33?

A picture containing histogram

Description automatically generated

Q7. Find the areas corresponding to the following intervals, expressing your answers as

decimals and then percentages. Show each result on a diagram of the normal curve.

Area for z scores:

a. below a z score of +0.85;

b. above a z score of +2.75;

c. below a z score of −1.03;

d. between z scores of +1.58 and +2.35;

e. between z scores of −2.80 and −2.50;

f. between z scores of −1.55 and +1.55;

g. between the mean and z = +2.33;

h. between the mean and 1.47 standard deviations above the mean;

i. between z = −0.58 and z = 0;

j. between the mean and 2.55 standard deviations below the mean

Solution:

*a*. Area = 0.8023 or 80.23%

b. Area = 0.0030 or 0.3%

c. Area = 0.1515 or 15.15%

d. Area = 0.0477 or 4.77%

e. Area = 0.0036 or 0.36%

f. Area = 0.8788 or 87.88%

g. Area = 0.4901 or 49.01%

h. Area = 0.4292 or 42.92%

i. Area = 0.2190 or 21.9%

j. Area = 0.4946 or 49.46%.

Q8. Let X be scores on a computer skills test with μX = 100 and σX = 10. Assume the

scores follow a normal distribution.

a. Find the number of standard deviations above or below the mean of each of the

following scores on the computer test: 95, 110, 130.

b. Use a diagram to find the raw scores equivalent to the following z scores: 0, −1, −2, 1, 2.

c. What is the z score for a raw score of 118.4?

Solution:

A picture containing box and whisker chart

Description automatically generated