STAT5002 Weekly Independent Exercises - solution

Sheet 1 - Week 4

STAT5002

1 `pnorm` and `qnorm`

The purpose of this question is to provide you with examples of how the functions pnorm and qnorm might be used in tests and exams. Additional examples and exercises involving other R functions will be provided in the coming weeks.

The daily sales of a typical sales person at a local clothing store can be assumed to follow the normal curve, with a mean of \$720 and a standard deviation of \$50. Daniel is a salesman in the store. His current sales is at the 25th percentile. If he would like to get to the 75th percentile, how much more does he have to sell? Select the line of R code that will give the correct answer.

(A) `qnorm(0.25, 720, 50)-qnorm(0.75, 720, 50)`
(B) `pnorm(0.25, 720, 50)-pnorm(0.75, 720, 50)`
(C) `pnorm(0.75, 720, 50)-pnorm(0.25, 720, 50)`
(D) `qnorm(0.75, 720, 50)-qnorm(0.25, 720, 50)`

Answer: D

2 The 68%-95%-99.7% rule

The 68%-95%-99.7% rule provides approximations for the area under a normal curve (proportions). In this section, we will apply this rule to determine proportions and quantiles of a normal distribution used to model the heights of Australian men.

We assume: - Mean height = 178.84 cm - Standard deviation (SD) = 6.53 cm

2.1

Question: What percentage of Australian men are shorter than 198.43cm?

Answer: 99.85%

The z-score (standard unit) of 198.43cm is

$$(198.43 - 178.84)/6.53 = 3$$

By the 68%-95%-99.7% rule, approximately 99.7% of the data falls within three standard deviations of the mean. This means that the remaining 0.3% is split equally between the two tails of the normal distribution.

Thus, the proportion of Australian men taller than 198.43 cm is 0.15%. Therefore, the percentage of Australian men shorter than 198.43 cm is 1 - 0.15% = 99.85%.

2.2

Question: What percentage of Australian men are taller than 165.78cm?

Answer: 97.5%

The z-score (standard unit) of 165.78cm is

$$(165.78 - 178.84)/6.53 = -2$$

By the 68%-95%-99.7% rule, approximately 95% of the data falls within two standard deviations of the mean. This means that the remaining 5% is split equally between the two tails of the normal distribution.

Thus, the proportion of Australian men shorter than 165.78 cm is 2.5%. Therefore, the percentage is 1 - 2.5% = 97.5%.

2.3

Question: What percentage of Australian men are between 172.31cm and 191.9cm?

Answer: 81.5%

The z-scores of 172.31cm and 191.9cm are

$$(172.31 - 178.84)/6.53 = -1$$

and

$$(191.9 - 178.84)/6.53 = 2$$

respectively.

Approximately 68% of the data falls within one standard deviations of the mean. This means that the remaining 32% is split equally between the two tails of the normal distribution. Thus, the proportion of Australian men shorter than 172.31 cm is 16%.

Approximately 95% of the data falls within two standard deviations of the mean. This means that the remaining 5% is split equally between the two tails of the normal distribution. Thus, the proportion of Australian men taller than 191.9 cm is 2.5%.

What percentage of Australian men are between 172.31cm and 191.9cm is given by 1 - 16% - 2.5% = 81.5%.

2.4

Question: What is the 97.5th percentile of heights of Australian men?

Answer: 191.9cm

Approximately 95% of the data falls within two standard deviations of the mean. This means that the remaining 5% is split equally between the two tails of the normal distribution. Thus, the proportion of Australian men taller than (mean + 2SD = 191.9cm) is 2.5%. Therefore, (1-2.5%=97.5%) of Australian men are shorter than (mean + 2SD = 191.9cm), which gives the percentile.

2.5

Question: What is the 2.5th percentile of heights of Australian men?

Answer: 165.78cm

Approximately 95% of the data falls within two standard deviations of the mean. This means that the remaining 5% is split equally between the two tails of the normal distribution. Thus, the proportion of Australian men shorter than (mean - 2SD = 165.78cm) is 2.5%, which gives the percentile.

2.6

Question: Write down the interval that contains the middle 95% of the heights.

Answer: [165.78 cm,191.9 cm]

This is the interval bounded between (mean - 2SD = 165.78cm) and (mean + 2SD = 191.9cm).

3 Correlation coefficient

Question: Consider the following two tables with pairs of values.

x	y
2	4
4	1
4	5
8	3

u	v
1	9
2	3
2	11
4	*

What is the value of $*$ in the second table so that both tables have the same **correlation coefficient**?

Hint: You don't need to actually calculate the correlation coefficient.

Answer: 7

The correlation coefficient is shift and scale invariant. Since $x = 2 * u$, multiplying u with 2 does not affect the correlation. Then we can use the transformation $v = 2 * y + 1$, which gives $^* = 2 * 3 + 1 = 7$.