**Practice RAT for Week 05, Thursday**

**Answers at the end**

|  |  |
| --- | --- |
| 1 | For normal distributions, what is the approximate z-score associated with the 2nd percentile? |
|  |
| (a) | 2 |
| (b) | 1 |
| (c) | -1 |
| (d) | -2 |

|  |  |
| --- | --- |
| 2 | What is the output of the following R code? (First choose an answer. Only then test it in R.)  pnorm(2) |
|  |
| (a) | 0.95 |
| (b) | 0.9772499 |
| (c) | NaN  Warning message:  In pnorm(2) : Invalid quantile |
| (d) | Error in pnorm(2) : argument "mean" is missing, with no default |

|  |  |
| --- | --- |
| 3 | The US Environmental Protection Agency (EPA) fuel economy estimates for automobile models tested recently predicted a mean of 10.5 kilometres per litre and a standard deviation of 2.6 kilometres per litre for expressway driving. Assuming that a Normal model can be applied.  What percentage of automobiles should get more than 13.1 kilometres per litre? |
|  |
| (a) | 2.5% |
| (b) | 5.0% |
| (c) | 16.0% |
| (d) | 32.0% |

|  |  |
| --- | --- |
| 4 | The World Health Organization measures child development by comparing the weights of children who are the same height and the same gender. In 2009, weights for all 80 cm tall girls in the reference population had a mean of 10.2 kg and standard deviation of 0.8 kg.  Assuming that the distribution follows the Normal model, how can R calculate the fraction of 80-cm girls with weight between 10.0 kg and 10.5 kg? |
|  |
| (a) | rnorm(10.5, mean = 10.2, sd = 0.8) -  rnorm(10.0, mean = 10.2, sd = 0.8) |
| (b) | pnorm(10.5, mean = 10.2, sd = 0.8) -  pnorm(10.0, mean = 10.2, sd = 0.8) |
| (c) | dnorm(10.5, mean = 10.2, sd = 0.8) -  dnorm(10.0, mean = 10.2, sd = 0.8) |
| (d) | qnorm(10.5, mean = 10.2, sd = 0.8) -  qnorm(10.0, mean = 10.2, sd = 0.8) |

|  |  |
| --- | --- |
| 5 | Corey has 4929 songs in his computer’s music library. The lengths of the songs have a mean of 242.2 seconds and standard deviation of 114.51 seconds. A Normal probability plot of the song lengths look like this:    How does the distribution differ from a Normal model? |
|  |
| (a) | The distribution is nearly uniform between 0 and 1400 seconds. |
| (b) | The distribution is right-skewed. |
| (c) | The distribution is symmetric, but there are more data points near the mean than in the Normal model. |
| (d) | The distribution is left-skewed. |

**Answers:**

**1d.** Approximately 95% of values fall within 2 standard deviations of the mean, so the corresponding percentiles are 2.5 and 97.5. (More exactly, they are 2.3 and 97.7.) The knee-jerk response would be (a), which is actually almost right; it's just on the wrong side of the distribution. 68% of values fall within 1 standard deviation of the mean, with corresponding percentiles of 16 and 84.

**2b.** pnorm(2) calculates the fraction of the data that is expected to have a z-score < 2 in the normal model . This fraction is the 95% from the 68-95-99.7 rule *plus* the 2.5% of the data that have a *z*-score less than -2.

**3c.** The percentage of automobiles that should get more than 13.1 kilometres per litre is also the percentage that would get more than one standard deviation above the mean. Assuming a Normal model can be applied and using the 68-95-99.7 rule, we know that 68% of all observations lie within one standard deviation of the mean and 32% of observations lie outside of one standard deviation. However, we are only interested in those that are more than one standard deviation above the mean (i.e. half of that 32% of observations) and so the answer is 16.0%.

**4b.** There is a similar example (“What proportion of SAT scores falls between 450 and 600?”) with an explanation in the R prep material.

**5b.** Compare the plot with Figure 5.9 in our textbook and read the accompanying text on page 145 for an explanation.