Module 18 self-assessment

Question 1

Use the t-distribution table from the EM2B exam appendix or the wikipedia page below to estimate the following values if T is a random variable drawn from a t-distribution with ν degrees of freedom.

`https://en.wikipedia.org/wiki/Student%27s_t-distribution`

(a) (i) $\mathbb{P}(T > 1.6)$ with $\nu = 3$. (ii) $\mathbb{P}(T < 1.6)$ with $\nu = 3$. (iii) $\mathbb{P}(-1.68 < T < 1.68)$ with $\nu = 29$. (iv) $\mathbb{P}(-1.6 < T < 1.6)$ with $\nu = 49$ (b) (i) The right critical value for probability $\alpha = 0.05$ when $\nu = 8$. (ii) The two-sided rejection region with probability $\alpha = 0.2$ for $\nu = 16$. (iii) Find the range for the middle 50% of probability with $\nu = 20$.

Solution:

For this exercise we will use the table 11.8 of the book. Recall that the t distribution is symmetric around $t = 0$ and it has different shape for a given number of degrees of freedom. Moreover, the tables give critical values, i.e. points on the horizontal axis and the areas under the graph to the right of those values.

(a) (i) From the table in the book, 3rd row for $\nu = 3$ we see that $\mathbb{P}(T > 1.638) = 0.1$ so we can estimate by extrapolation $\mathbb{P}(T > 1.6) \approx 0.11$.

(ii) Using the answer in (i) $\mathbb{P}(T < 1.6) \approx 1 - 0.11 = 0.89$.

(iii) Here we want the area of the t-distribution graph for $\nu = 29$ between points $t_1 = -1.68$ and $t_2 = 1.68$. The relevant section of the table 11.8 is the row before the last. By the symmetry of the graph wrt $t = 0$ it is easy to see that the area to the right of the critical value t_2 is the same as the area to the left of t_1. As the table gives critical values then the sought area is simply 1 minus twice the area to the right of $t = 1.68$. The second column in that row of the table gives $\mathbb{P}(T > 1.699) = 0.05$ so we can extrapolate to $\mathbb{P}(T > 1.680) \approx 0.051$ and therefore $\mathbb{P}(-1.68 < T < 1.68) = 1 - 2 \cdot 0.051 = 0.898$.

(iv) In this part we need the area above the t-distribution with $\nu = 49$ between $t_1 = -1.6$ and $t_2 = 1.6$. Although there are other tables freely available on the web that relate probabilities and critical values for this distribution, the one in fig. 11.8 does not include this particular numbers. We thus utilise the info on the last row for $\nu \to \infty$. Arguing similar to part (iii) the sought probability is 1 minus twice the probability for a critical value $t = 1.6$. The table gives $\mathbb{P}(T > 1.645) = 0.05$ and $\mathbb{P}(T > 1.282) = 0.10$ so from a rather crude approximation (as the graph isn't close to being linear over $1.282 \leq t \leq 1.645$) we have $\mathbb{P}(T > 1.6) \approx 0.053$. So the sought probability is $\mathbb{P}(-1.6 < T < 1.6) = 1 - 2 \cdot 0.053 = 0.894$.

(b)

(i) Straight from the table, 8th row, 2nd column $z = 1.860$.

(ii) Here we look for points t_1 and t_2 symmetrically positioned wrt $t = 0$ on the t-distribution with $\nu = 16$ so that the area to the left of t_1 and right of t_2 sums up to $\alpha = 0.2$. The 2-sided rejection region imposes that the two areas above each equal to

$\alpha = 0.1$, and in this case the region is $(-\infty, t_1] \cup [t_2, +\infty)$. So in the 16th row of the table (for $\nu = 16$ we look for t_2 such that $\mathbb{P}(T > t_2) = 0.1$ which we find in the first column as $t_2 = 1.337$. Due to symmetry $t_1 = -t_2$ so the rejection region we want is $(-\infty, -1.337] \cup [1.337, +\infty)$.

(iii) Here we want the values t_1 and t_2 on the symmetric $t-$distribution with $\nu = 20$ such that the area under the graph between these points is 0.5. In effect the area outside this region splits equally between the area to the right of the critical value t_2 and the left of the quantile t_1. As the table does not provide any data for $\alpha > 0.10$ and we need to find those for $\alpha = 0.25$, then we can look this up on a different table on the web. One I found has $t_2 = 0.687$ for $\alpha = 0.25$. In effect, the range is $-0.687 \leq t \leq 0.687$.

Question 2

Consider data $\{2.5, 7.3, 5.5, 8.5, 11.5, 9.7\}$ from a normal distribution whose mean μ and variance σ^2 are unknown. Construct the confidence intervals (i) 99%, (ii) 95% and (iii) 80% for μ.

Solution:

We compute the sample mean

$$\bar{x} = \frac{1}{6}\big(2.5 + 7.3 + 5.5 + 8.5 + 11.5 + 9.7\big) = 7.5,$$

and sample variance

$$S_n^2 = \frac{1}{5}\sum_{i=1}^{6}(x_i - \bar{x})^2 = 10.176.$$

As we do not know the variance and have only few data, the appropriate test is a two-tailed T-test at (i) $\alpha = 0.01$, (ii) $\alpha = 0.05$ and (iii) $\alpha = 0.20$ significance levels, from a student's t distribution with 5 degrees of freedom (number of samples - 1). Converting from α to $\alpha/2$ for the 2-tail areas we look for the critical values at the respective $t_{\alpha/2}$ which can be found to be

$$t_{0.005} = 4.032, \quad t_{0.025} = 2.571, \quad \text{and} \quad t_{0.10} = 1.476.$$

In effect, since $\widehat{\mathrm{SE}} = S_n/\sqrt{6} = 1.3023$ the intervals for the mean are:

(i)

$$\mu \in \big[\bar{x} - t_{0.005}\widehat{\mathrm{SE}}, \bar{x} + t_{0.005}\widehat{\mathrm{SE}}\big] = [2.2491, 12.751],$$

(ii)

$$\mu \in \big[\bar{x} - t_{0.025}\widehat{\mathrm{SE}}, \bar{x} + t_{0.025}\widehat{\mathrm{SE}}\big] = [4.1518, 10.848],$$

(iii)

$$\mu \in \big[\bar{x} - t_{0.025}\widehat{\mathrm{SE}}, \bar{x} + t_{0.025}\widehat{\mathrm{SE}}\big] = [5.5778, 9.4222].$$