Question 1

Suppose that $X_1, X_2, \ldots, X_n$ follow a $\mathrm{N}\left(\theta_1, \sigma^2\right)$ distribution where θ_1 is unknown and σ^2 is unknown, and the independent random variables $Y_1, Y_2, \ldots, Y_m$ follow a $\mathrm{N}\left(\theta_2, \sigma^2\right)$ distribution where θ_2 is unknown (note that both sets of random variables are assumed to have the same unknown variance σ^2). We further assume that each X_i is independent of each Y_j.

Use Theorem 3.4.4 in the notes to show that if we assume $\mu_1 = \mu_2$ then the statistic

$$T = \frac{\overline{X} - \overline{Y}}{S_p\sqrt{\frac{1}{n} + \frac{1}{m}}},$$

where

$$S_p^2 = \frac{1}{n+m-2}\left((n-1)S_X^2 + (m-1)S_Y^2\right),$$

$$S_X^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \overline{X})^2, \qquad S_Y^2 = \frac{1}{m-1}\sum_{j=1}^{m}(Y_j - \overline{Y})^2.$$

follows Student's t-distribution with $n+m-2$ degrees of freedom, i.e. $T \sim t_{n+m-2}$.

Question 2

Suppose that the random variables $X_1, X_2, \ldots, X_n$ are independent and identically distributed according to a normal distribution with unknown mean μ and known variance $\sigma^2 = 9$. Suppose that $\mathbf{X} = (X_1, X_2, \ldots, X_n)$ is observed as $\mathbf{x} = (x_1, x_2, \ldots, x_n)$, where

$$\sum_{i=1}^{n} x_i = 740, \qquad n = 100.$$

Using the tables in your notes, test the hypothesis that $\mu = 8$ at the significance level $\alpha = 0.05$.

Question 3

Suppose that the random variables $Y_1, Y_2, \ldots, Y_n$ are independent and identically distributed according to a normal distribution with unknown mean μ and unknown variance σ^2. Suppose that $\mathbf{Y} = (Y_1, Y_2, \ldots, Y_n)$ is observed as $\mathbf{y} = (y_1, y_2, \ldots, y_n)$, where

$$\sum_{i=1}^{n} y_i = 32, \qquad \sum_{i=1}^{n} y_i^2 = 124, \qquad n = 16.$$

(a) Test the hypothesis that $\mu = 0.9$ at the significance level $\alpha = 0.01$.

(b) If this hypothesis is not rejected at $\alpha = 0.01$, find the smallest significance level at which it is rejected.

R question

There is no R question this week, because the coursework will be released on 23 February. Make sure you have completed the R exercises in Problem Sheets 8, 9, 10 and 11.