MATH10282 Introduction to Statistics
Semester 2, 2019/2020
Example Sheet 12 - Solutions

1. (i) We have that the times for the stoves of Type A are $X_{11}, \ldots, X_{1n_1} \sim N(\mu_1, \sigma_1^2)$, and the times for Type B stoves are $X_{21}, \ldots, X_{2n_2} \sim N(\mu_2, \sigma_2^2)$. An appropriate test statistic is

$$Z = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}\,.$$

Under $H_0 : \mu_1 - \mu_2 = 0$, the statistic Z is distributed as $N(0,1)$. We reject H_0 in favour of $H_1 : \mu_1 - \mu_2 > 0$ at the 5% significance level if $Z > z_{0.95} = 1.645$.

Here, the observed value of Z is

$$z = \frac{11.4 - 9.9}{\sqrt{\frac{6.25}{10} + \frac{9.0}{12}}} = 1.279\,.$$

Thus, $z < z_{0.95}$ and so we do not reject H_0 at the 5% level. There is insufficient evidence to reject the claim that the mean heating capacities of the two stoves are equal.

(ii) We would reject H_0 if

$$\bar{X}_1 - \bar{X}_2 > 1.645\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}} = 1.929\,.$$

The probability of rejecting H_0 is

$$\begin{aligned}
\mathrm{P}\left(\bar{X}_1 - \bar{X}_2 > 1.929\right) &= \mathrm{P}\left(\frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} > 1.645 - \frac{\mu_1 - \mu_2}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}\right) \\
&= 1 - \Phi\left(1.645 - \frac{2}{\sqrt{\frac{6.25}{10} + \frac{9.0}{12}}}\right) \\
&\qquad \text{since } \frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} \sim N(0,1) \\
&= 1 - \Phi(-0.061) = 0.52\,.
\end{aligned}$$

2. We denote the wear amounts for Type A tyres as $X_{11}, \ldots, X_{1n_1} \sim N(\mu_1, \sigma^2)$, and those for Type B tyres as $X_{21}, \ldots, X_{2n_2} \sim N(\mu_2, \sigma^2)$, where σ^2 denotes the common variance.

(i) An appropriate test statistic is

$$Z = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{\frac{\sigma^2}{n_1} + \frac{\sigma^2}{n_2}}}\,.$$

We reject $H_0 : \mu_1 - \mu_2 = 0$ in favour of $H_1 : \mu_1 - \mu_2 \neq 0$ at the 5% significance level if $|Z| > z_{0.95} = 1.96$.
Here, the observed value of Z is

$$z = \frac{10.24 - 9.76}{\sqrt{\frac{2\times 1.742}{25}}} = 1.286\,.$$

Thus $|z| < 1.96$ and we do not reject H_0 at the 5% significance level. We conclude that there is no evidence of a difference between the mean tyre wear for Types A and B.

(ii) H_0 is not rejected if

$$-1.96\sigma\sqrt{\frac{1}{n_1} + \frac{1}{n_2}} < \bar{X}_1 - \bar{X}_2 < 1.96\sigma\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}\,.$$

Hence the probability that we do not reject H_0 is

$$\mathrm{P}\left(-1.96\sigma\sqrt{\frac{1}{n_1} + \frac{1}{n_2}} < \bar{X}_1 - \bar{X}_2 < 1.96\sigma\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}\right)$$

$$= \mathrm{P}\left(-1.96 - \frac{(\mu_1 - \mu_2)}{\sigma\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} < \frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{\sigma\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} < 1.96 - \frac{\mu_1 - \mu_2}{\sigma\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}\right)$$

$$= \Phi\left(1.96 - \frac{\mu_1 - \mu_2}{\sigma\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}\right) - \Phi\left(-1.96 - \frac{\mu_1 - \mu_2}{\sigma\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}\right)$$

$$\text{since } \frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{\sigma\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} \sim N(0,1)$$

$$= \Phi(-0.719) - \Phi(-4.639) = 0.24.$$

3. (i) An appropriate test statistic is

$$Z = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}\,,$$

which, is approximately distributed as $N(0,1)$ under $H_0 : \mu_1 - \mu_2 = 0$ if both n_1 and n_2 are large. At the approximate 5% significance level, we reject H_0 in favour of $H_1 : \mu_1 - \mu_2 \neq 0$ if $|Z| > z_{0.975} = 1.96$.

Here, the observed value of Z is

$$z = \frac{28.2 - 30.7}{\sqrt{\frac{16.36}{33} + \frac{18.92}{37}}} = -2.49\,.$$

Thus $|z| > z_{0.975}$ and so we reject H_0 at the approximate 5% significance level in favour of the alternative hypotheses that the mean absorption times for the two drugs are unequal.

(ii) Assume $\sigma_1^2 = \sigma_2^2 = \sigma^2$. An estimator of this common variance is

$$\hat{\sigma}^2 = \frac{(n_1 - 1)S_1^2 + (n_2 - 1)S_2^2}{n_1 + n_2 - 2}\,.$$

Here, the estimate is $\hat{\sigma}^2 = \frac{32\times16.36+36\times18.92}{33+37-2} = 17.715$. The test statistic is

$$T = \frac{\bar{X}_1 - \bar{X}_2}{\hat{\sigma}\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}\,,$$

which under H_0 has a Student t distribution with $n_1 + n_2 - 2$ degrees of freedom. The observed value of T is

$$t = \frac{28.2 - 30.7}{\sqrt{17.715 \times \left(\frac{1}{33} + \frac{1}{37}\right)}} = -2.481\,.$$

Now, the exact critical value for a two-sided test at the 5% significance level is $t_{0.975} = 1.995$, i.e. the 0.975 point of a $t(68)$ distribution. The observed value of T satisfies $|t| > t_{0.975}$. Hence, at the 5% significance level we reject H_0 and conclude that $\mu_1 \neq \mu_2$.

4. Suppose that the mean and variance of the first population are μ_1 and σ_1^2, and that the mean and variance of the second population are μ_2 and σ_2^2. We are not told that the data are normally distributed.

We wish to test $H_0 : \mu_1 - \mu_2 = 0$ vs $H_1 : \mu_1 - \mu_2 \neq 0$. An appropriate test statistic is as follows:

$$Z = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}\,.$$

Under H_0, by asymptotic results Z is approximately distributed as $N(0, 1)$ provided both n_1 and n_2 are large. At the 1% significance level an appropriate rejection region is $\{|Z| > z_{0.995} = 2.576\}$.

Here, the observed value of Z is

$$z = \frac{9.8 - 11.9}{\sqrt{\frac{2.9^2}{49} + \frac{4.2^2}{64}}} = -3.14.$$

Thus $|z| > z_{0.995}$ and we reject H_0 at the 1% significance level, concluding that $\mu_1 \neq \mu_2$.

5. Let p_1 denote the proportion who believe in extraterrestrials in the population of adults who did not attend college, and let p_2 denote the proportion who believe in extraterrestrials in the population of adults who did attend college.

$$\begin{aligned} \text{Did not attend college: } & n_1 = 100, \quad r_1 = 37, \quad \hat{p}_1 = 0.37 \\ \text{Attended college: } & n_2 = 100, \quad r_2 = 47, \quad \hat{p}_2 = 0.47\,. \end{aligned}$$

We wish to test $H_0 : p_1 - p_2 = 0$ vs $H_1 : p_1 - p_2 < 0$. Under H_0, we have that $p_1 = p_2 = p$. We may estimate the common probability p by

$$\bar{p} = \frac{37 + 47}{200} = 0.42.$$

An appropriate test statistic is $\hat{p}_1 - \hat{p}_2$, which has standard error $\sqrt{p(1-p)\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}$ under H_0. The estimate of the standardized version of this test statistic is

$$Z = \frac{\hat{p}_1 - \hat{p}_2}{\sqrt{\bar{p}(1-\bar{p})\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}\,.$$

Under H_0, the statistic Z is approximately distributed as $N(0, 1)$ provided both n_1 and n_2 are large. We reject H_0 at the 5% level if $Z < -z_{0.95} = -1.645$. Here the observed value of Z is

$$z = \frac{0.37 - 0.47}{\sqrt{0.42 \times 0.58 \times \frac{2}{100}}} = -1.433\,.$$

This satisfies $Z > -z_{0.95}$ and so we do not reject H_0. There is insufficient evidence to reject the claim that the proportion believing in extraterrestrials is the same among those who did and did not attend college.