# ST102 Outline solutions to Exercise 16

1. For a 98% confidence interval, we need the lower and upper percentile values from the $\chi^2_{n-1} = \chi^2_{17}$ distribution. These are, respectively, $\chi^2_{0.99,\,17} = 6.408$ and $\chi^2_{0.01,\,17} = 33.409$. Hence we obtain:

$$\left(\frac{(n-1)s^2}{\chi^2_{\alpha/2,\,n-1}}, \frac{(n-1)s^2}{\chi^2_{1-\alpha/2,\,n-1}}\right) = \left(\frac{17 \times 27.64}{33.409}, \frac{17 \times 27.64}{6.408}\right) = (14.06, 73.33).$$

   Note that this is a very wide confidence interval due to (i.) a high level of confidence (98%), and (ii.) a small sample size ($n = 18$).

2. Using Murdoch and Barnes' *Statistical Tables*, we obtain the following:

   (a) If $P(Z < k) = 0.0681$, then $P(Z > -k) = 0.0681$, hence $k = -1.49$.

   (b) $k = P(Z < -1.72) = P(Z > 1.72) = 0.0427$.

   (c) $k = P(-0.23 < Z < 2.04) = P(Z < 2.04) - P(Z \leq -0.23) = (1 - P(Z \geq 2.04)) - P(Z \geq 0.23) = (1 - 0.02068) - 0.4090 = 0.57032$.

   (d) If $P(X_6 < k) = 0.20$, then $k = \chi^2_{0.80,\,6} = 3.070$.

   (e) If $P(X_{26} > k) = 0.025$, then $k = \chi^2_{0.025,\,26} = 41.923$.

   (f) $k = P(18.939 < X_{28} < 50.993) = P(X_{28} \geq 18.939) - P(X_{28} \geq 50.993) = 0.90 - 0.005 = 0.895$.

   (g) If $P(T_7 > k) = 0.99$, then $P(T_7 > -k) = 0.01$, hence $-k = t_{0.01,\,7} = 2.998$ and so $k = -2.998$.

   (h) $k = P(T_{19} < 2.861) = 1 - P(T_{19} \geq 2.861) = 1 - 0.005 = 0.995$.

   (i) Since $P(|T_{60}| < k) = 0.98$, $P(T_{60} > k) = (1 - 0.98)/2 = 0.01$. Hence $k = t_{0.01,\,60} = 2.390$.

   Note that it is important that you are familiar with Murdoch and Barnes' *Statistical Tables* which contains a lot of information. In the summer examination, you will be provided with a copy of these tables.

3. Let $T = X_1 + X_2 + \cdots + X_{12}$. Therefore, $T \sim \text{Bin}(12, \pi)$. We use $T$ as the test statistic. With the given sample, we observe $t = 3$. We now determine which are the more extreme values of $T$ if $\text{H}_0$ is true.

   Under $\text{H}_0$, $\text{E}(T) = n\pi_0 = 6$. Hence 9 is as extreme as 3, and the more extreme values are:

$$0, \quad 1, \quad 2, \quad 10, \quad 11 \quad \text{and} \quad 12.$$

Therefore, the $p$-value is:

$$\left(\sum_{i=0}^{3} + \sum_{i=9}^{12}\right) P_{\mathrm{H}_0}(T = i) = \left(\sum_{i=0}^{3} + \sum_{i=9}^{12}\right) \frac{12!}{i!\,(12-i)!}\,(0.50)^i (1-0.50)^{12-i}$$

$$= 2 \times (0.50)^{12} \sum_{i=0}^{3} \frac{12!}{i!\,(12-i)!}$$

$$= 2 \times (0.50)^{12} \times \left(1 + 12 + \frac{12 \times 11}{2!} + \frac{12 \times 11 \times 10}{3!}\right)$$

$$= 0.1460.$$

Since $\alpha = 0.10 < 0.1460$, we do not reject the null hypothesis of a fair coin at the 10% significance level. The observed data are consistent with the null hypothesis of a fair coin.

4. When $\sigma^2$ is unknown, we use the test statistic $T = \sqrt{n}(\bar{X} - 9)/S$. Under $\mathrm{H}_0$, $T \sim t_{15}$. With $\alpha = 0.05$, we reject $\mathrm{H}_0$ if:

   (a) $t > t_{0.05,\,15} = 1.753$, against $\mathrm{H}_1 : \mu > 9$.

   (b) $t < -t_{0.05,\,15} = -1.753$, against $\mathrm{H}_1 : \mu < 9$.

   (c) $|t| > t_{0.025,\,15} = 2.131$, against $\mathrm{H}_1 : \mu \neq 9$.

   For the given sample, $t = 2.02$. Hence we reject $\mathrm{H}_0$ against the alternative $\mathrm{H}_1 : \mu > 9$, but we will not reject $\mathrm{H}_0$ against the two other alternative hypotheses.

   When $\sigma^2$ is known, we use the test statistic $T = \sqrt{n}(\bar{X} - 9)/\sigma$. Now under $\mathrm{H}_0$, $T \sim N(0,1)$. With $\alpha = 0.05$, we reject $\mathrm{H}_0$ if:

   (a) $t > z_{0.05} = 1.645$, against $\mathrm{H}_1 : \mu > 9$.

   (b) $t < -z_{0.05} = -1.645$, against $\mathrm{H}_1 : \mu < 9$.

   (c) $|t| > z_{0.025} = 1.96$, against $\mathrm{H}_1 : \mu \neq 9$.

   For the given sample, $t = 2.02$. Hence we reject $\mathrm{H}_0$ against the alternative $\mathrm{H}_1 : \mu > 9$ and $\mathrm{H}_1 : \mu \neq 9$, but we will not reject $\mathrm{H}_0$ against $\mathrm{H}_1 : \mu < 9$.

   With $\sigma^2$ known, we should be able to perform inference better simply because we have more information about the population. More precisely, for the given significance level, we require less extreme values to reject $\mathrm{H}_0$. Put another way, the $p$-value of the test is reduced when $\sigma^2$ is given. Therefore, the risk of rejecting $\mathrm{H}_0$ is also reduced.

5. Let $T$ be the number of car accidents per week such that $T \sim \text{Poisson}(\lambda)$. We are interested in testing:

   $$\mathrm{H}_0 : \lambda = 8 \quad \text{vs.} \quad \mathrm{H}_1 : \lambda < 8.$$

   Under $\mathrm{H}_0$, then $T \sim \text{Poisson}(8)$, and we observe $t = 3$. Hence the $p$-value is:

   $$P(T \leq 3) = \sum_{t=0}^{3} \frac{\mathrm{e}^{-8} 8^t}{t!} = \mathrm{e}^{-8}\left(1 + 8 + \frac{8^2}{2!} + \frac{8^3}{3!}\right) = 0.0424.$$

   Since $0.0424 < 0.05$, we reject $\mathrm{H}_0$ and conclude that there is evidence to suggest that the accident rate has dropped.