MA2261 - DLI, Linear Statistical Models, Year 2022-2023

Solutions of exercises for feedback class 2

(Note: the exercise number refers to the workbook)

EXERCISE 2.2

i) Recalling that the distribution is valid under the constraint $y \geq b$, we calculate:

$$E[Y] = \int_b^\infty y f_Y(y) dy = \int_b^\infty y \frac{ab^a}{y^{a+1}} dy = ab^a \int_b^\infty y^{-a} dy =$$

$$= -\frac{ab^a}{a-1} \left[y^{-(a-1)} \right]_b^\infty = \frac{ab}{a-1}, \quad a > 1$$

ii) Let us calculate the likelihood function for a,

$$\mathcal{L}(a) = \prod_{i=1}^n f_Y(y_i) = \prod_{i=1}^n \frac{ab^a}{y_i^{a+1}} = a^n b^{an} \prod_{i=1}^n y_i^{-(a+1)}$$

and the log likelihood,

$$\ell(a) = \log \mathcal{L}(a) = n \log a + an \log b - (a+1) \sum_{i=1}^n \log(y_i)$$

Differentiating with respect to a and equating to zero we get,

$$\frac{d\ell(a)}{da} = \frac{n}{a} + n \log b - \sum_{i=1}^n \log(y_i) = 0$$

So that we obtain the estimate $\hat{a}$

$$\hat{a} = \frac{n}{-n \log b + \sum_{i=1}^n \log(y_i)} = \frac{n}{\sum_{i=1}^n \log\left(\frac{y_i}{b}\right)}$$

Finally, knowing that $b = 2$, we obtain

$$\hat{a} = \frac{n}{\sum_{i=1}^n \log\left(\frac{y_i}{2}\right)}$$

As $\frac{d^2\ell(a)}{da^2} = -\frac{n}{a^2} < 0$, then $\hat{a}$ above is the MLE of a.

iii) Calculating the logarithms of the given data, we have $\sum_1^{15} \log(y_i) = 16.0318$, thus

$$\hat{a} = \frac{15}{16.0318 - 15 \log 2} = 2.6621$$

EXERCISE 2.6

a) We have $H_0 : \mu = 11$ and $H_1 : \mu \neq 11$.

b) $T = \frac{\overline{X}-\mu_0}{S/\sqrt{n}} = \frac{\overline{X}-11}{S/\sqrt{25}}$ is t-distributed with d.f $n-1 = 24$.

c) We reject H_0 when $T \leq -t_{24,0.025}$ or $T \geq t_{24,0.025}$ where

$$P(T \leq -t_{24,0.025}) + P(T \geq t_{24,0.025}) = 5\%.$$

That gives $t_{24,0.025} = 2.064$. The rejection region is hence the interval $(-\infty,\ -2.064) \cup (2.064,\ +\infty)$.

d) That gives $T = \frac{\overline{x}-11}{S/\sqrt{25}} = -\frac{5}{3}$. Since the observed value of T is outside the rejection region, we accept H_0. There is hence sufficient evidence to conclude that the average absence due to illness probably has not changed.

Alternatively, the p-value is $P(T > \frac{5}{3}) + P(T < -\frac{5}{3}) = 0.1086 > 0.05$, thus, we accept H_0.

EXERCISE 2.8

We want to do a left-sided test of $H_0 : \mu = 0$ against $H_1 : \mu < 0$.

We construct the test statistic

$$T = \frac{\bar{X} - \mu_0}{S/\sqrt{n}}$$

which follows a t-distribution with d.f. $n - 1 = 8$.

And from the data, $T = \frac{-11.2-0}{9.6/\sqrt{9}} = -3.5$. When we use a 5% significance level, the critical value is then $t_{8,0.05} = 1.86$ and therefore the critical region is $(-\infty, -1.86)$. We should hence reject H_0 as $-3.5 < -1.86$. We reject H_0 and claim that the expected value is probably negative.

Alternatively, the p-value is $P(T < -3.5) = 0.004 < 0.05$, which also leads to reject H_0.