Probability and Statistics

Solution sheet 13

Solution 13.1 Using the Neyman-Pearson test with the hypothesis

$$H_0 : X \sim f(x)dx,$$
$$H_1 : X \sim f(x-1)dx.$$

We have that the likelihood ratio is given by

$$L = \frac{f(x-1)}{f(x)}.$$

then in the case of the normal variable $L = e^{x-\frac{1}{2}}$, and we have rejection when $L > c$, i.e.$x > \ln c + \frac{1}{2}$. Then the rejection set is of the form (a, ∞).

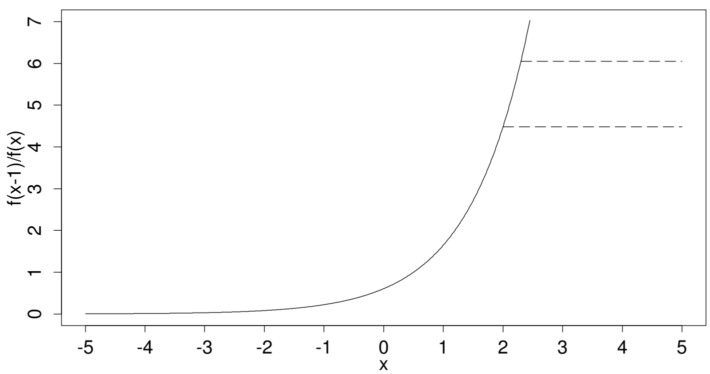

Figure 1: Rejections sets of the normal case.

In the case of the Cauchy random variable the likelihood ration is given by $L = \frac{x^2+1}{x^2-2x+2}$, then we have an interesting behavior as you can see in figure 2. If you put $c = 1$ then you will have a non bounded interval, but if you put $c > 1$ you will have a bounded interval.

This happens because the Cauchy distribution is heavy tailed.

Solution 13.2 We have to construct the test between the hypothesis:

$$H_0 : F^{-1}(0.5) = m,$$
$$H_1 : F^{-1}(0.5) \neq m.$$

We will use the statistic $T_{n,m} = \sum_{i=1}^{m} \mathbb{1}_{\{X_n \leq m\}}$ and the test is going to be given by

$$\phi(x) = 1 \Leftrightarrow \left| T_{n,m} - \frac{n}{2} \right| > c(n, \alpha),$$

where $x = (x_i)_{i=1}^n$, n is the size of the experiment and α is the level of the test. We have that, under H_0, $T_{n,m}$ follows a binomial law $Bin(n, \frac{1}{2})$. Then if we define $k = \frac{n}{2} - c(n, \alpha)$, we should have, thanks to the symmetry of the Binomial coefficients:

$$\sum_{j=0}^{k-1} \binom{n}{j} 0.5^n \leq \frac{\alpha}{2} < \sum_{j=0}^{k} \binom{n}{j} 0.5^n,$$

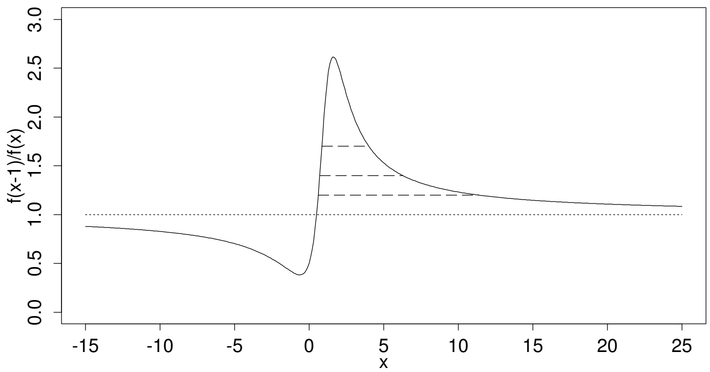

Figure 2: Rejections sets of the cauchy case.

and $\frac{n}{2} + c(n, \alpha) = n - k$. Take $C \subseteq \mathbb{R}^{n+1}$

$$C := \left\{ (x_1, ..., x_n, m) : k \leq \sum_{i=1}^{n} \mathbb{1}_{\{x_i \geq m\}} \leq n - k \right\},$$

then we would like $B((x_1, .., x_n)) = \{m : (x_1, ..., x_n, m) \in C\}$ to be a confidence interval of level α of $F(\frac{1}{2})$. Note that

$$m \in [x_{(j)}, x_{(j+1)}) \Leftrightarrow \sum_{i=1}^{n} \mathbb{1}_{\{x_i > m\}} = j$$

then $B((x_1, ..., x_n)) = [x_{(k)}, x_{(n-k)}]$. Thanks to the central limit theorem

$$\begin{aligned}\mathbb{P}_m \left(k \leq \sum_{i=1}^{n} \mathbb{1}_{\{x_i > m\}} \leq n - k \right) &= \mathbb{P}_m \left(\frac{2}{\sqrt{n}} \left(k - \frac{n}{2} \right) \leq \frac{2}{\sqrt{n}} \left(\sum_{i=1}^{n} \mathbb{1}_{\{x_i > m\}} - \frac{n}{2} \right) \leq \frac{2}{\sqrt{n}} \left(\frac{n}{2} - k \right) \right) \\ &\approx \phi \left(\frac{2}{\sqrt{n}} \left(\frac{n}{2} - k \right) \right) - \phi \left(\frac{2}{\sqrt{n}} \left(k - \frac{n}{2} \right) \right) \\ &= 2\phi \left(\frac{2}{\sqrt{n}} \left(\frac{n}{2} - k \right) \right) - 1.\end{aligned}$$

We want it to be bigger than 0.95, so $k \approx \frac{n}{2} - \frac{1.96}{2}\sqrt{n} \approx \lfloor \frac{n}{2} \rfloor - \sqrt{n}$. Then

$$B((x_1, ..., x_n)) = \left[x_{(\lfloor \frac{n}{2} \rfloor - \sqrt{n})}, x_{(\lfloor \frac{n}{2} \rfloor + \sqrt{n})} \right],$$

is a confidence interval with 95% of confidence level. Then using the same notation as in Theorem 6.4 we have that

$$\begin{aligned}C &:= \{(x, B(x))\} \\ &= \left\{ (x, m) : m \in \left[x_{(\lfloor \frac{n}{2} \rfloor - \sqrt{n})}, x_{(\lfloor \frac{n}{2} \rfloor + \sqrt{n})} \right] \right\}.\end{aligned}$$

Then $A(m) = \{x \in \mathbb{R}^n : \theta \in \left[x_{(\lfloor \frac{n}{2} \rfloor - \sqrt{n})}, x_{(\lfloor \frac{n}{2} \rfloor + \sqrt{n})} \right]\}$, so thanks to the Theorem 6.4 a test to the level 0.95 is given by $\varphi_m(x) = \mathbb{1}_{\{x_{(\lfloor \frac{n}{2} \rfloor - \sqrt{n})} \leq m \leq x_{(\lfloor \frac{n}{2} \rfloor + \sqrt{n})}\}}$.

Solution 13.3

(a) $\left[\overline{X} - \frac{S_n}{\sqrt{n}} t_{n-1,1-\frac{\alpha}{2}},\ \overline{X} + \frac{S_n}{\sqrt{n}} t_{n-1,1-\frac{\alpha}{2}}\right]$.

(b) The right-endpoint clearly diverges to ∞.

For the left-endpoint we set $a := \frac{1}{n}\sum_{i=2}^{n} x_i$, and with the hint we obtain

$$\begin{aligned}
S_n^2 &= \frac{1}{n-1}\sum_{i=1}^{n}(x_i-\overline{x})^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i-a)^2 - \underbrace{\frac{n}{n-1}(a-\overline{x})^2}_{=\frac{1}{n(n-1)}x_1^2} \\
&= \underbrace{\frac{1}{n-1}\sum_{i=2}(x_i-a)^2}_{=:b} + \frac{1}{n-1}(x_1-a)^2 - \frac{1}{n(n-1)}x_1^2 \\
&= b + \frac{1}{n-1}(x_1^2 - 2x_1 a + a^2) - \frac{1}{n(n-1)}x_1^2 \\
&= x_1^2 \underbrace{\left(\frac{1}{n-1} - \frac{1}{n(n-1)} - \frac{2a}{x_1}\frac{1}{n-1} + \frac{1}{n-1}\frac{a^2}{x_1^2} + \frac{b}{x_1^2}\right)}_{=:f(x_1)\xrightarrow[x_1\to\infty]{}\frac{1}{n-1}-\frac{1}{n(n-1)}=\frac{1}{n}}.
\end{aligned}$$

Moreover, it holds $\overline{x} = \frac{1}{n}x_1 + a = x_1 \underbrace{\left(\frac{1}{n} + \frac{a}{x_1}\right)}_{=:g(x_1)\xrightarrow[x_1\to\infty]{}\frac{1}{n}}$. We obtain,

$$\begin{aligned}
\overline{x} - \frac{S_n}{\sqrt{n}} t_{n-1,1-\frac{\alpha}{2}} &= \overline{x}\left(1 - \frac{S_n}{\overline{x}\sqrt{n}} t_{n-1,1-\frac{\alpha}{2}}\right) = \overline{x}\left(1 - \frac{x_1\sqrt{f(x_1)}}{x_1 g(x_1)\sqrt{n}} t_{n-1,1-\frac{\alpha}{2}}\right) \\
&= \overline{x}\underbrace{\left(1 - \frac{\sqrt{f(x_1)}}{g(x_1)\sqrt{n}} t_{n-1,1-\frac{\alpha}{2}}\right)}_{\xrightarrow[x_1\to\infty]{} 1-t_{n-1,1-\frac{\alpha}{2}}}.
\end{aligned}$$

For all the levels used in practice, the t-Quantile is strictly larger than 1, i.e. the left-endpoint converges to $-\infty$ The confidence interval does not give any information anymore, every value is plausible for $\mathbb{E}[X]$.

Solution 13.4

(a) This is not a paired sample.

(b) The model is given by $X_1, \ldots, X_{10}$ iid $\sim \mathcal{N}(\mu_X, \sigma^2)$ und $Y_1, \ldots, Y_{10}$ iid $\sim \mathcal{N}(\mu_Y, \sigma^2)$ where μ_X, μ_Y and σ are unknown. The null and alternative hypothesis are given by

$$H_0 : \mu_X = \mu_Y \quad \text{und} \quad H_A : \mu_X \neq \mu_Y.$$

The statistic is

$$T := \frac{\bar{X} - \bar{Y}}{S_{\text{pool}}\sqrt{2/10}}$$

and under H_0, it is t-distributed with 18 degrees of freedom. With a level of 5%, the null hypothesis will be rejected when $|T| > 2.101$. From the data we obtain $\bar{x} = 10.693, \bar{y} = 6.75$ and $S_{pool} = 4.255$, so $T = 2.0723$ i.e. H_0 is not rejected.

Solution 13.5

(a) We know that $T^{(1)}_{2n+1} \sim N\left(\mu, \frac{\sigma^2}{2n+1}\right)$, then

$$\begin{aligned}
&\mathbb{P}\left(|T^{(1)}_{2n+1} - \mu| \leq c^{(1)}_n\right) = 0.95\\
\Rightarrow &\mathbb{P}\left(\frac{|T^{(1)}_{2n+1} - \mu|}{\sigma}\sqrt{2n+1} \leq \frac{c^{(1)}_n}{\sigma}\sqrt{2n+1}\right) = 0.95\\
\Rightarrow &c^{(1)}_n = \frac{\sigma}{\sqrt{2n+1}}\Phi^{-1}(0.975) \approx 1.96\frac{\sigma}{\sqrt{2n+1}}.
\end{aligned}$$

For the second estimator, define $\tilde{X}_k := X_k - \mu \sim N(0,\sigma)$ and $\tilde{X}_{(k)} = (\tilde{X})_{(k)}$, then $F^{-1}\left(\frac{1}{2}\right) = 0$. Thanks to the example 4.6 of the Skript,we know that:

$$\mathbb{P}\left(\sqrt{2n+1}\tilde{X}_{(n+1)} \leq x\right) \to \Phi(2F'(0)x),$$

where in our case $F'(0) = \frac{1}{\sqrt{2\pi}\sigma}$. Then,

$$\begin{aligned}
\mathbb{P}\left(|T^{(2)}_n - \mu| \leq x\right) &= \mathbb{P}\left(\sqrt{2n+1}\tilde{X}_{(n+1)} \leq \sqrt{2n+1}x)\right) + \mathbb{P}\left(\sqrt{2n+1}\tilde{X}_{(n+1)} \geq -\sqrt{2n+1}x\right)\\
&\approx 2\Phi\left(\frac{\sqrt{2}}{\sqrt{\pi}\sigma}\sqrt{2n+1}x\right) - 1,
\end{aligned}$$

then if we take $c^{(2)}_n := \Phi^{-1}(0.975)\frac{\sqrt{\pi}}{\sqrt{2}\sqrt{2n+1}}\sigma$ we have what we wanted.

(b) Taking $q = \frac{\pi}{2}$ we have that:

$$\frac{c^{(2)}_{qn}}{c^1_n} \approx \frac{\sqrt{\pi}}{\sqrt{2}}\frac{\sqrt{2n+1}}{\sqrt{\pi n+1}} \to 1.$$

The parameter q represents how many more data I have to take with the estimator 2 to get the same order of error bounds than for the one of experiment 1.

Solution 13.6

(a) Using the fact that $I(\beta_0, \beta_1) \to +\infty$ as $\|(\beta_0, \beta_1)\| \to \infty$ (which is true since the x_is are not all the same), the infimum of I is in approximated in some compact set. Since I is continuous, the infimum of I is a minimum. We can look for critical points $(\hat{\beta}_0, \hat{\beta}_1)$:

$$\begin{aligned}
\partial_{\beta_0} I(\hat{\beta}_0, \hat{\beta}_1) &= 2\sum_{i=1}^n \hat{\beta}_0 + \hat{\beta}_1 x_i - y_i = 0\\
\partial_{\beta_1} I(\hat{\beta}_0, \hat{\beta}_1) &= 2\sum_{i=1}^n x_i(\hat{\beta}_0 + \hat{\beta}_1 x_i - y_i) = 0.
\end{aligned}$$

We solve the above system:

$$n\hat{\beta}_0 + n\bar{x}\hat{\beta}_1 = n\bar{y} \quad \Rightarrow \quad \hat{\beta}_0 = \bar{y} - \bar{x}\hat{\beta}_1,$$

and

$$
\begin{aligned}
& n\bar{x}\hat{\beta}_0 + \left(\sum_{i=1}^n x_i^2\right)\hat{\beta}_1 = \sum_{i=1}^n x_i y_i \\
\Rightarrow\quad & n\bar{x}(\bar{y} - \bar{x}\hat{\beta}_1) + \left(\sum_{i=1}^n x_i^2\right)\hat{\beta}_1 = \sum_{i=1}^n x_i y_i \\
\Rightarrow\quad & \left(\sum_{i=1}^n x_i^2 - n\bar{x}^2\right)\hat{\beta}_1 = \sum_{i=1}^n x_i y_i - n\bar{x}\bar{y} \\
\Rightarrow\quad & \left(\sum_{i=1}^n (x_i - \bar{x})^2\right)\hat{\beta}_1 = \sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y}) \\
\Rightarrow\quad & \hat{\beta}_1 = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sum_{i=1}^n (x_i - \bar{x})^2}.
\end{aligned}
$$

(b) We apply the above formula to find $\hat{\beta}_1 = 40.89$ and $\hat{\beta}_0 = 0.55$.

Solution 13.7

(a) If we calculate the $k+1$ partial derivatives $\partial Q/\partial\beta_0, \cdots, \partial Q/\partial\beta_k$, and we set each of these derivatives equal to 0, we obtain the following $k+1$ linear equations involving $k+1$ unknown values $\beta_0, \cdots, \beta_k$:

$$
\begin{aligned}
\hat{\beta}_0 n + \hat{\beta}_1 \sum_{i=1}^n x_i + \cdots + \hat{\beta}_k \sum_{i=1}^n x_i^k &= \sum_{i=1}^n y_i, \\
\hat{\beta}_0 \sum_{i=1}^n x_i + \hat{\beta}_1 \sum_{i=1}^n x_i^2 + \cdots + \hat{\beta}_k \sum_{i=1}^n x_i^{k+1} &= \sum_{i=1}^n x_i y_i, \\
&\vdots \\
\hat{\beta}_0 \sum_{i=1}^n x_i^k + \hat{\beta}_1 \sum_{i=1}^n x_i^{k+1} + \cdots + \hat{\beta}_k \sum_{i=1}^n x_i^{2k} &= \sum_{i=1}^n x_i^k y_i.
\end{aligned}
$$

As before, if these equations have a unique solution, that solution provides the minimum value for Q. A necessary and sufficient condition for a unique solution is that the determinant of the $(k+1) \times (k+1)$ matrix formed by the coefficients of $\hat{\beta}_0, \cdots, \hat{\beta}_k$ above is not zero.

(b) In this example, it is found that the equations are

$$
\begin{aligned}
& 10\beta_0 + 23.3\beta_1 + 90.37\beta_2 = 8.1, \\
& 23.3\beta_0 + 90.37\beta_1 + 401.0\beta_2 = 43.59, \\
& 90.37\beta_0 + 401.0\beta_1 + 1892.7\beta_2 = 204.55.
\end{aligned}
$$

The unique solution is

$$
\hat{\beta}_0 = -0.744, \quad \hat{\beta}_1 = 0.616, \quad \hat{\beta}_2 = 0.013.
$$

Hence the least squares parabola is

$$
y = -0.744 + 0.616x + 0.013x^2.
$$

Solution 13.8

(a) Just note that the coordinate i of (**??**) is given by

$$y_i = (X\beta)_i + \epsilon_i = \sum_{k=1}^{n} X_{ik}\beta_k + \epsilon_i = x_i \cdot \beta + \epsilon_i.$$

(b) We know that for the normal linear regression model is $\bar{\beta} := ((X^tX)^{-1}X)Y$, so it's a linear model. Let's compute its expected value

$$\begin{aligned}\mathbb{E}\left(\bar{\beta}\right) &= \mathbb{E}\left(((X^tX)^{-1}X^t)Y\right)\\ &= \mathbb{E}\left((X^tX)^{-1}X^t(X\beta+\epsilon)\right)\\ &= \beta + \mathbb{E}\left(\epsilon\right) = \beta,\end{aligned}$$

Then $\bar{\beta}$ is unbiased.

(c) We just have to compute

$$\begin{aligned}\mathrm{Var}(\bar{\beta}) &= \mathrm{Var}(((X^tX)^{-1}X^t)Y)\\ &= ((X^tX)^{-1}X^t)\mathrm{Var}(Y)((X^tX)^{-1}X^t)^t\\ &= \sigma^2(X^tX)^{-1}.\end{aligned}$$

(d) We just have to compute its expected value:

$$\begin{aligned}\mathbb{E}\left(\tilde{\beta}\right) &= \mathbb{E}\left(\bar{\beta} + CY\right)\\ &= \beta + C\mathbb{E}\left(X\beta+\epsilon\right)\\ &= (I + CX)\beta,\end{aligned}$$

given its expected value should be β for all $\beta \in \mathbb{R}^n$, then we have that $CX = 0$.

(e) We have to compute the covariance matrix of $\tilde{\beta}$

$$\begin{aligned}\mathrm{Var}(\tilde{\beta}) &= \mathrm{Var}(Dy) = D\mathrm{Var}(y)D^t = \sigma^2 DD^t\\ &= \sigma^2((X^tX)^{-1}X^t + C)(X(X^tX)^{-1} + C^t)\\ &= \sigma^2((X^tX)^{-1}X^tX(X^tX)^{-1} + (X^tX)^{-1}X^tC^t + CX(X^tX)^{-1} + CC^t)\\ &= \sigma^2(X^tX)^{-1} + \sigma^2(X^tX)^{-1}(\underbrace{CX}_{0})^t + \sigma^2\underbrace{CX}_{0}(X^tX)^{-1} + \sigma^2CC^t\\ &= \underbrace{\sigma^2(X^tX)^{-1}}_{\mathrm{Var}(\hat{\beta})} + \sigma^2CC^t.\end{aligned}$$

To finish note that σ^2CC^t is a positive semidefinitive matrix.