Statistics: Tutorial Week 4 - Solutions to Practice Exercises

Practice Exercises

Exercise 1. Find the ML estimator for p_0 in the statistical model $\{\text{Bernoulli}(p) \mid p \in [0,1]\}$. In this case we have

$$g(x \mid p) = p^x(1-p)^{1-x} \qquad \text{if } x \in \{0,1\}.$$

SOLUTION. For the likelihood we have:

$$L(p \mid \boldsymbol{x}) = \prod_{i=1}^{n} p^{x_i}(1-p)^{(1-x_i)}.$$

This gives us the log-likelihood:

$$\ell(p \mid \boldsymbol{x}) = \sum_{i=1}^{n} \log\left(p^{x_i}(1-p)^{(1-x_i)}\right) = \sum_{i=1}^{n} x_i \log p + (1-x_i)\log(1-p).$$

We obtain derivatives

$$\frac{d}{dp}\ell(p \mid \boldsymbol{x}) = \sum_{i=1}^{n} \frac{x_i}{p} - \frac{1-x_i}{1-p}.$$
$$\frac{d^2}{dp^2}\ell(p \mid \boldsymbol{x}) = \sum_{i=1}^{n} -\frac{x_i}{p^2} - \frac{1-x_i}{(1-p)^2}.$$

We find the stationary point by setting the first derivative to zero and obtain

$$\tilde{p} = \overline{x}.$$

Note that this is a stationary point only if $\overline{x}$ is not equal to zero or one. If $\tilde{p}$ is not equal to zero or one, then the second derivative of the log likelihood evaluated at $\tilde{p}$ is strictly negative, so the found stationary point is a local maximum. In this case we have a unique stationary point, which is a local maximum, hence it is the global maximum. The special cases $\overline{x} = 0$ and $\overline{x} = 1$ remain. In these cases we do not have a stationary point, so the maximum must be obtained in a boundary point. Note that if $\overline{x} = 0$, then we must have $\boldsymbol{x} = (0, \ldots, 0)$ and similarly if $\overline{x} = 1$, then we must have $\boldsymbol{x} = (1, \ldots, 1)$. It follows that

$$L(p \mid (0,\ldots,0)) = \prod_{i=1}^{n} 1 - p = (1-p)^n \qquad \Rightarrow \qquad \hat{p} = 0 = \overline{x}$$
$$L(p \mid (1,\ldots,1)) = \prod_{i=1}^{n} p = p^n \qquad \Rightarrow \qquad \hat{p} = 1 = \overline{x},$$

where we used that $(1-p)^n$ is decreasing in p so that the maximum is attained at the most left point in the parameter space $[0, 1]$, while p^n is increasing in p so that the maximum is attained at the most right point in the parameter space $[0, 1]$. We conclude that

$$\hat{p}_{ML} = \overline{X}.$$

Exercise 2. Let $L(\theta \mid \boldsymbol{x})$ be the likelihood corresponding to a statistical model. Prove that $L(\theta \mid \boldsymbol{x})$ has a *unique* maximum at $\tilde{\theta} \in \Theta$ if and only if $\ell(\theta \mid x)$ has a *unique* maximum at $\tilde{\theta}$. Hint: Use that the logarithm is a monotone increasing function, that is $\log y_1 > \log y_2$ if and only if $y_1 > y_2$.

Solution. We show the "only if" part first. We have by definition of $\tilde{\theta}$ that for any different $\theta \in \Theta$

$$y_1 = L(\tilde{\theta} \mid \boldsymbol{x}) > L(\theta \mid \boldsymbol{x}) = y_2.$$

Therefore by the monotone property of the logarithm we get

$$\ell(\tilde{\theta} \mid \boldsymbol{x}) = \log y_1 > \log y_2 = \ell(\theta \mid \boldsymbol{x}).$$

Thus $\tilde{\theta}$ is the unique maximiser of $\ell(\theta \mid \boldsymbol{x})$. The "if" part follows by doing the steps the other way around. We now have by definition of $\tilde{\theta}$ that for any different $\theta \in \Theta$

$$\log y_1 = \ell(\tilde{\theta} \mid \boldsymbol{x}) > \ell(\theta \mid \boldsymbol{x}) = \log y_2.$$

By the monotone property of the logarithm we get

$$L(\tilde{\theta} \mid \boldsymbol{x}) = y_1 > y_2 = L(\theta \mid \boldsymbol{x}).$$

Thus $\tilde{\theta}$ is the unique maximiser of $L(\theta \mid \boldsymbol{x})$.

Exercise 3. Prove Lemma 4.14, which says the following: Let $\Theta \in \mathbb{R}$ be an interval and let $h : \Theta \to \mathbb{R}$ be a function that is twice differentiable. Suppose that there exists a unique stationary point $\tilde{\theta} \in \Theta$, that is $h'(\tilde{\theta}) = 0$, and that it satisfies the second derivative test $h''(\tilde{\theta}) < 0$. Then $\tilde{\theta}$ is the unique point at which the global maximum is attained. Hint: use the mean value theorem.

Solution. An *intuitive proof* can be given by looking at Figure 1. Suppose that we find the local maximum $\tilde{\theta} = 2.1$ as our unique stationary point. Assume there exists another θ^* such that $h(\theta^*) \geq h(\tilde{\theta})$, such as $\theta^* = 7$ in the figure. Then there has to be a point between $\tilde{\theta}$ and θ^* where h has to turn around to go up. However, the point at which the function turns creates a local minimum, and therefore $\tilde{\theta}$ is not the unique stationary point.

A more *formal proof* can be given by using the mean value theorem. Suppose there exists another $\theta^* \in \Theta$ such that $h(\theta^*) \geq h(\tilde{\theta})$ and assume without loss of generality that $\theta^* < \tilde{\theta}$. Then by the mean value theorem there exists a $\zeta \in (\theta^*, \tilde{\theta})$ such that

$$h(\theta^*) - h(\tilde{\theta}) = h'(\zeta)(\theta^* - \tilde{\theta}).$$

Note that $h(\theta^*) - h(\tilde{\theta}) \geq 0$ and $\theta^* - \tilde{\theta} < 0$, thus $h'(\zeta) \leq 0$. Recall that $h''(\tilde{\theta}) < 0$, so there exists a $\bar{\theta}$ slightly to the left of $\tilde{\theta}$ such that $h'(\bar{\theta}) > 0$. We conclude, by the intermediate value theorem, that there must exist a point in $(\theta^*, \tilde{\theta})$ at which h' is zero, which is a contradiction since $\tilde{\theta}$ was the unique stationary point.

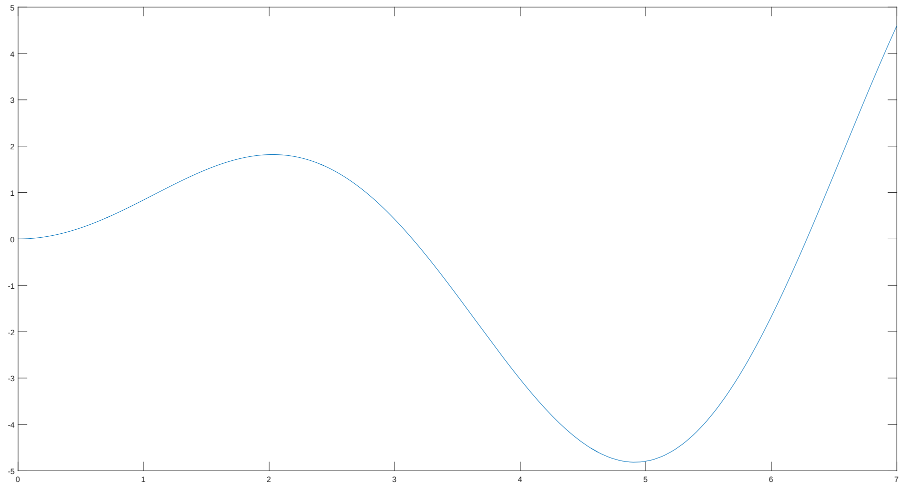

Figure 1

Exercise 4. The goal of this exercise is to combine most of what we have learned until now to tackle a complete statistics problem. An insurance company has asked us for help to predict the expected amount of money that they will have to pay next month. We have the following data:

- The number of monthly claims that had to be paid in the previous n months. We will denote this n_i for $1 \leq i \leq n$.
- The value of each individual claim. We write $c_{i,j}$, where $1 \leq j \leq n_i$, for the value of claim j in month i.

We use the following model strategy. The number of claims can be modelled nicely as a random sample from a Poisson population:

$$N_i \sim \text{Poisson}(\mu).$$

This follows by a similar intuition as for the milk store in week one, we have many independent clients with small probability to enter a claim. We model the claim size as all independent with an exponential distribution:

$$C_{ij} \sim \text{Exponential}(\lambda).$$

a. Write down a formula for the total amount that has to be paid in month i.

b. Find the expected value for the total amount that has to be paid in month i.

c. Let $\theta = (\mu, \lambda)$. Show that the joint density

$$f(n_i, c_{i1}, c_{i2}, \ldots, c_{in_i} \mid \theta)$$

of $(N_i, C_{i1}, C_{i2}, \ldots, C_{iN_i})$ is equal to

$$\left(\prod_{j=1}^{n_i} \lambda e^{-\lambda c_{ij}} \right) e^{-\mu} \frac{\mu^{n_i}}{n_i!}.$$

d. Derive the maximum likelihood estimator for the expected value of the amount that has to be paid next month.

SOLUTION.

a. We have N_i claims in month i, with values C_{ij}. The total amount to pay in month i is therefore equal to

$$\sum_{j=1}^{N_i} C_{ij}.$$

b. We calculate the expected value using the law of total expectation:

$$\mathbb{E}_\theta\left(\sum_{j=1}^{N_i} C_{ij}\right) = \mathbb{E}_\theta\left(\mathbb{E}_\theta\left(\sum_{j=1}^{N_i} C_{ij} \middle| N_i\right)\right) = \mathbb{E}_\theta\left(N_i \mathbb{E}_\theta(C_{11})\right) = \mathbb{E}_\theta\left(\frac{N_i}{\lambda}\right) = \frac{\mu}{\lambda}.$$

c. We use the conditional distribution to obtain

$$\begin{aligned} f(n_i, c_{i1}, \ldots, c_{in_i} \mid \theta) &= f(c_{i1}, \ldots, c_{in_i} \mid N_i = n_i, \theta) P_\mu(N_i = n_i) \\ &= \left(\prod_{j=1}^{n_i} \lambda e^{-\lambda c_{ij}}\right) e^{-\mu} \frac{\mu^{n_i}}{n_i!}. \end{aligned}$$

d. Using the previous question we obtain the following for the likelihood:

$$\prod_{i=1}^{n}\left(\prod_{j=1}^{n_i} \lambda e^{-\lambda c_{ij}}\right) e^{-\mu} \frac{\mu^{n_i}}{n_i!}.$$

After taking logarithms we get

$$\begin{aligned} &\sum_{i=1}^{n}\sum_{j=1}^{n_i} \log\left(\lambda e^{-\lambda c_{ij}}\right) + \sum_{i=1}^{n} \log\left(e^{-\mu}\frac{\mu^{n_i}}{n_i!}\right) \\ &\quad = \left(\sum_{i=1}^{n}\sum_{j=1}^{n_i} \log\lambda - \lambda c_{ij}\right) + \left(\sum_{i=1}^{n} -\mu + n_i \log\mu - \log n_i!\right). \end{aligned}$$

The first set of sums depends only on λ, while the second sum depends only on μ. Therefore we can find the joint ML estimator of θ by maximising these two groups separately. Finding the maximum of the second sum with respect to μ is the same as finding the ML estimator for μ in the classic Poisson setting. Therefore we immediately have

$$\hat{\mu}_{ML} = \frac{1}{n}\sum_{i=1}^{n} N_i.$$

Setting the derivative with respect to λ of the first one equal to zero gives us

$$\sum_{i=1}^{n}\sum_{j=1}^{n_i} \frac{1}{\lambda} - c_{ij} = 0 \qquad \Rightarrow \qquad \tilde{\lambda} = \frac{\sum_{i=1}^{n} n_i}{\sum_{i=1}^{n}\sum_{j=1}^{n_i} c_{ij}}$$

Note that the second derivative is equal to

$$-\frac{1}{\lambda^2}\sum_{i=1}^{n} n_i.$$

This is smaller than zero, except if all n_i are zero, which is an irrelevant scenario for us. Therefore we obtain

$$\hat{\lambda}_{ML} = \frac{\sum_{i=1}^{n} N_i}{\sum_{i=1}^{n}\sum_{j=1}^{N_i} C_{ij}}$$

We conclude by the invariance of ML estimators that the ML estimator for the expected value of the amount that has to be paid next month is equal to

$$\widehat{\left(\frac{\mu}{\lambda}\right)}_{ML} = \frac{\hat{\mu}_{ML}}{\hat{\lambda}_{ML}} = \frac{1}{n}\sum_{i=1}^{n}\sum_{j=1}^{N_i} C_{ij}.$$