AMS 206: Classical and Bayesian Inference (Winter 2015)

Homework 1 solutions

1. Exercises 7.5.2 and 7.5.3
Solution: In exercise 7.5.2, where there is no restriction on the parameter space, the MLE for p is given by $58/70 = 0.8286$. In Exercise 7.5.3, the likelihood function is again given by $p^{58}(1-p)^{12}$, but now for p restricted in the interval $[1/2, 2/3]$. Since the likelihood function is increasing in p within this interval, the MLE is given by $2/3$.

2. Exercise 7.5.5
Solution: (a) To determine the MLE for θ, begin by writing the likelihood,

$$f_n(\boldsymbol{x} \mid \theta) = \prod_{i=1}^{n} \frac{e^{-\theta}\theta^{x_i}}{x_i!}$$

from which the log-likelihood function is given by

$$L(\theta) = -n\theta + \log(\theta)(\sum\nolimits_{i=1}^{n} x_i) - \sum\nolimits_{i=1}^{n} \log(x_i!)$$

Therefore, the derivative w.r.t. θ,

$$\frac{dL(\theta)}{d\theta} = -n + \frac{\sum_{i=1}^{n} x_i}{\theta}$$

equals 0 when $\theta = \sum_{i=1}^{n} x_i/n$, i.e., $\widehat{\theta} = \bar{x}_n$ (it is straightforward to verify that the second derivative takes negative values).
(b) Note that if $x_i = 0$ for all i, then $f_n(x \mid \theta) = e^{-n\theta}$, which is decreasing as a function of θ. Since $\theta > 0$, θ can get arbitrarily close to 0 without ever reaching it, and therefore the MLE does not exist.

3. Exercise 7.5.7
Solution: The likelihood has the form

$$f_n(\boldsymbol{x} \mid \beta) = \prod_{i=1}^{n} \beta \exp\{-\beta x_i\} = \beta^n \exp\{-\beta \sum\nolimits_{i=1}^{n} x_i\}$$

Therefore the log-likelihood function, $L(\beta) = n\log(\beta) - \beta\sum_{i=1}^{n} x_i$, and $\frac{dL(\beta)}{d\beta} = \frac{n}{\beta} - \sum_{i=1}^{n} x_i$. Setting this to 0 and solving for β, we get $\widehat{\beta} = n/(\sum_{i=1}^{n} x_i)$ (again, it is easy to check that the second derivative takes negative values).

4. Exercise 7.5.8
Solution: (a) The likelihood is $f_n(\boldsymbol{x} \mid \theta) = \prod_{i=1}^{n} \exp\{\theta - x_i\}$ if all $x_i > \theta$, and 0 otherwise. The likelihood can therefore be written as $\exp\{n\theta - \sum_{i=1}^{n} x_i\}$, if $\min\{x_1, \ldots, x_n\} > \theta$, and 0 otherwise. This is an increasing function of θ, and is maximized when θ is made as large as possible, subject to the constraint $\theta < \min(x_1, \ldots, x_n)$. Since θ can get arbitrarily close to this value but never reach it, the MLE does not exist.
(b) However, if the strict inequality $x > \theta$ in the pdf becomes $x \geq \theta$, and $x \leq \theta$ is replaced by $x < \theta$, then $\widehat{\theta} = \min\{x_1, \ldots, x_n\}$.

5. Exercise 7.5.11
Solution: The likelihood is given by $f_n(\boldsymbol{x} \mid \theta_1, \theta_2) = (\theta_2 - \theta_1)^{-n}$ for $\theta_1 \leq x_i \leq \theta_2$ for all x_i, and is equal to 0 otherwise. This boundary condition for the likelihood to be non-zero can be stated as $\theta_1 \leq \min\{x_1, \ldots, x_n\} < \max\{x_1, \ldots, x_n\} \leq \theta_2$. To maximize the likelihood, we need to make $\theta_2 - \theta_1$ as small as possible subject to the inequalities above. That is, assign to θ_2 its smallest possible value and to θ_1 its largest possible value, resulting in $\widehat{\theta_2} = \max\{x_1, \ldots, x_n\}$ and $\widehat{\theta_1} = \min\{x_1, \ldots, x_n\}$.

6. Exercise 7.5.12

Solution: The likelihood is given by the multinomial distribution (see Section 5.9 of the textbook), which is a generalization of the Binomial distribution for $k \geq 2$ categories. The multinomial likelihood is proportional to $\prod_{i=1}^{k} \theta_i^{n_i}$, where n_i represents the number of individuals of type i in the sample, and θ_i, for $i = 1, ..., k$, are the respective probabilities that satisfy $\sum_{i=1}^{k} \theta_i = 1$. Therefore, the log-likelihood for the parameter vector $\boldsymbol{\theta} = (\theta_1, ..., \theta_{k-1})$ can be written as

$$L(\boldsymbol{\theta}) = \sum_{i=1}^{k-1} n_i \log(\theta_i) + n_k \log(1 - \sum_{i=1}^{k-1} \theta_i).$$

The $k-1$ partial derivatives of this function are

$$\frac{\partial L(\boldsymbol{\theta})}{\partial \theta_l} = \frac{n_l}{\theta_l} - \frac{n_k}{\theta_k}, \quad l = 1, \ldots, k-1.$$

Setting these equations equal to 0, gives $\frac{n_l}{\theta_l} = \frac{n_k}{\theta_k}$, for $l = 1, \ldots, k-1$, and therefore,

$$\frac{n_1}{\widehat{\theta}_1} = \frac{n_2}{\widehat{\theta}_2} = \cdots = \frac{n_k}{\widehat{\theta}_k} = \frac{1}{C},$$

subject to the constraint $\sum_{i=1}^{k} \widehat{\theta}_i = 1$. We thus have $\widehat{\theta}_i = Cn_i$, for $i = 1, ..., k$, and $\sum_{i=1}^{k} \widehat{\theta}_i = 1$, which solving for C yields $C = 1/n$. Therefore, $\widehat{\theta}_i = n_i/n$, for $i = 1, ..., k$.

To verify that the critical point above provides the maximum, we need to show that the $(k-1) \times (k-1)$ Hessian matrix $H(\boldsymbol{\theta})$, with elements $\partial^2 L(\boldsymbol{\theta})/\partial\theta_i\partial\theta_j$, for $i, j = 1, ..., k-1$, is a negative definite matrix when evaluated at $\boldsymbol{\theta} = (\widehat{\theta}_1, ..., \widehat{\theta}_k) = (n_1/n, ..., n_k/n)$. Consider the case with $k = 3$ which is the simplest extension of the Binomial distribution. In this case, the Hessian matrix is a 2×2 matrix with upper diagonal element $-(n_1/\theta_1^2) - (n_3/(1-\theta_1-\theta_2)^2)$, lower diagonal element $-(n_2/\theta_2^2) - (n_3/(1-\theta_1-\theta_2)^2)$, and off-diagonal element $-(n_3/(1-\theta_1-\theta_2)^2)$. Since $-(n_1/\theta_1^2) - (n_3/(1-\theta_1-\theta_2)^2) < 0$, for all (θ_1, θ_2), and $\det(H(\theta_1, \theta_2)) > 0$, for all (θ_1, θ_2), the Hessian matrix is negative definite. The result can be extended for general $k \geq 3$; recall that one approach to check that matrix $H(\boldsymbol{\theta})$ is negative definite is by verifying that its m-th order principal minor is negative when m is odd, and is positive when m is even.

7. Exercise 7.6.2

Solution: The standard deviation of a Poisson distribution with mean θ is $\theta^{1/2}$. Assuming that at least one observed count in the random sample is different from 0, the MLE of θ was shown in exercise 7.5.5 to be $\widehat{\theta} = \bar{x}_n$. Therefore, using the invariance property for MLEs, the MLE of the standard deviation is $\bar{x}_n^{1/2}$.

8. Exercise 7.6.4

Solution: Let Y denote the lifetime of the specific type of lamp. Then, the probability of a single lamp failing within T hours is given by $p = \Pr(Y \leq T) = \int_0^T \beta \exp(-\beta y) \mathrm{d}y = 1 - \exp(-\beta T)$. Based on the problem assumptions, the number X of lamps that fail within T hours follows a binomial distribution with parameters n and p: there is a fixed number n of "trials" (testing of lamps); the trials are independent (assumption of a random sample); there are 2 possible outcomes on each trial (the lamp fails during the period of T hours or not); and p can be taken to be the same across trials (all n lamps are of the same type). The MLE for p is $\widehat{p} = X/n$, and using the invariance property of MLEs, the MLE of β can be obtained as $\widehat{\beta} = -\log(1 - (X/n))/T$.

9. Exercise 7.6.6

Solution: Subtract μ and divide by σ in both sides of the inequality included in the probability, $0.95 = \Pr(X < \theta)$, to obtain $0.95 = \Pr(Z < (\theta - \mu)/\sigma) = \Phi((\theta - \mu)/\sigma)$, where $Z \sim N(0, 1)$. Using the table for the distribution function of the standard normal distribution, we obtain $\Phi(1.645) = 0.95$, and therefore $(\theta - \mu)/\sigma = 1.645$. The MLE for θ is then obtained by solving for θ in this equation and plugging in $\widehat{\mu}$ and $\widehat{\sigma}$, that is, $\widehat{\theta} = 1.645\widehat{\sigma} + \widehat{\mu}$, where the MLEs for μ and σ are obtained in Example 7.5.6.

10. Exercise 7.6.15
Solution: Done in class.

11. Suppose that $X_1, ..., X_n$ form a random sample from a double exponential distribution (also referred to as Laplace distribution) for which the p.d.f. is given by

$$f(x \mid \mu, \sigma) = \frac{1}{2\sigma} \exp\left(-\frac{|x-\mu|}{\sigma}\right), \quad \text{for} \quad -\infty < x < \infty.$$

Here, μ is a location parameter and σ a scale parameter for the distribution, where $-\infty < \mu < \infty$ and $\sigma > 0$. Describe how the M.L.E. $\widehat{\mu}$ of μ can be obtained (there is no closed-form expression for $\widehat{\mu}$), and obtain the expression for the M.L.E. $\widehat{\sigma}$ of σ. (*Hint:* use the profile likelihood approach discussed in class in the context of Example 7.5.6.)
Solution: The log-likelihood function can be expressed as

$$L(\mu, \sigma) = -n \log(2\sigma) - \frac{1}{\sigma} \sum_{i=1}^{n} |x_i - \mu|.$$

Using the profile likelihood approach, we first find the MLE of μ under fixed σ. Directly from the form of the log-likelihood function for fixed σ, maximizing $L(\mu, \sigma)$ w.r.t. μ is equivalent to minimizing $\sum_{i=1}^{n} |x_i - \mu|$ w.r.t. μ. The key observation is that, since this function does not depend on σ, the solution to this minimization problem will provide the MLE $\widehat{\mu}$ for μ.
Next, to obtain the MLE of σ, we can maximize the profile log-likelihood for this parameter:

$$L(\widehat{\mu}, \sigma) = -n \log(2\sigma) - \frac{1}{\sigma} \sum_{i=1}^{n} |x_i - \widehat{\mu}|$$

which is straightforward using derivatives. The answer is $\widehat{\sigma} = n^{-1} \sum_{i=1}^{n} |x_i - \widehat{\mu}|$.

Note that it can shown that one solution for the MLE of μ is the sample median, though it is not the unique solution when n is even. For two different proofs of this result, refer to: Norton, R.M. (1984), "The Double Exponential Distribution: Using Calculus to Find a Maximum Likelihood Estimator", *The American Statistician*, vol. 38, pp. 135-136; and Hurley, W. J. (2009), "An Inductive Approach to Calculate the MLE for the Double Exponential Distribution", *Journal of Modern Applied Statistical Methods*, vol. 8, pp. 594-596.