18.443 Problem Set 4 Spring 2015
Statistics for Applications
Due Date: 3/13/2015
prior to 3:00pm

Problems from John A. Rice, Third Edition. [$Chapter.Section.Problem$]

1. Problem 8.10.5

   X is a descrete random varialbe with

   $$P(X = 1) = \theta \text{ and } P(X = 2) = 1 - \theta.$$

   Three independent ovservations of X are made:

   $$x_1 = 1, x_2 = 2, x = 2.$$

   (a). Find the method of moments estimate of θ

   $$\mu_1 = E[X] = 1 \times (\theta) + 2 \times (1 - \theta) = 2 - \theta.$$

   Setting $\hat{\mu}_1 = \overline{X} = \mu_1$ and solving for θ gives

   $$\hat{\theta} = 2 - \hat{\mu}_1 = 2 - \overline{X} = 2 - 5/3 = 1/3.$$

   (b). What is the likelihood function?

   $$\begin{array}{rcl} Lik(\theta) & = & f(x_1 \mid \theta) \times f(x_2 \mid \theta) \times f(x_3 \mid \theta) \\ & = & \theta \times (1 - \theta) \times (1 - \theta) = \theta(1 - \theta)^2. \end{array}$$

   (c).The mle for θ is found by solving

   $$\begin{array}{rcl} 0 = \frac{\partial}{\partial\theta}[\log\left(Lik(\theta)\right)] & = & \frac{\partial}{\partial\theta}[\ln(\theta) + 2\ln(1 - \theta)] \\ & = & \frac{1}{\theta} - \frac{2}{1-\theta} \\ \Longrightarrow \hat{\theta} & = & 1/3. \end{array}$$

   (d). The uniform distribution has density

   $$\pi(\theta) = 1,\ 0 < \theta < 1$$

   The posterior distribution has density

   $$\begin{array}{rcl} \pi(\theta \mid x) & \propto & \pi(\theta) \times Lik(\theta) \\ & = & \theta(1 - \theta)^2 \times 1 \end{array}$$

   By inspection, the posterior density must be that of a $Beta(a, b)$ distribution with

   $$a = 2 \text{ and } b = 3.$$

2. Problem 8.10.7

   Suppose that X follows a geometric distribution

   $$P(X = x \mid p) = p(1-p)^{x-1},\ x = 1, 2\ldots$$

   Assume an i.i.d. sample of size n.

   (a). Find the method of moments estimate of p.

   The first moment of X is

   $$\mu_1 \;=\; \sum_{k=1}^{\infty} kp(1-p)^{k-1} = 1/p$$

   (See Appendex A, page A1.)

   Solving $\hat{\mu}_1 = \overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i = \mu_1$ gives

   $$\hat{p} = \frac{1}{\overline{X}}.$$

   (b). Find the mle of p. The mle of a sample $x_1, \ldots, x_n$ minimizes

   $$Lik(p) = \prod_{i=1}^{n}[p(1-p)^{x_i-1}] = p^n(1-p)^{(\sum_{i=1}^{n} X_i)-n}$$

   The mle is the same as the mle for a sample of size $(\sum_{i=1}^{n} x_i)$ from a $Bernoulli(p)$ distribution with n successes:

   $\hat{p} \;=\; \frac{n}{\sum_{i=1}^{n} x_i - n + n} \;=\; 1/\overline{X}$ (same as the method-of-moments estimate)

   (c) The asympotic variance of the mle is

   $$Var(\hat{p}_{MLE}) \approx \frac{1}{nI(p)} = \frac{p^2(1-p)}{n}$$

   where

   $$\begin{aligned} I(p) &= E[-\frac{\partial^2 log[P(x|p)]}{\partial p^2}] \\ &= E[-\frac{\partial^2}{\partial p^2}[\ln(p) + (x-1)\ln(1-p)] \\ &= E[-[-\frac{1}{p^2} - \frac{(x-1)}{(1-p)^2}]] \\ &= \frac{1}{p^2} + \frac{E[X]-1}{(1-p)^2} \\ &= \frac{1}{p^2} + \frac{(1-p)}{p(1-p)^2} \\ &= \frac{1}{p^2} + \frac{1}{p(1-p)} \\ &= \frac{1}{p^2(1-p)} \end{aligned}$$

   (d). Let p have a uniform distribution on $[0, 1]$. The posterior distribution of p has density

   $$\begin{aligned} \pi(p \mid x_1, \ldots, x_n) &\propto Lik(p) \times \pi(\theta) \\ &= p^n(1-p)^{(\sum_{i=1}^{n} x_i)-n} \times 1 \\ &= p^{a^*-1}(1-p)^{b^*-1} \end{aligned}$$

   which can be recognized as a $Beta(a^*, b^*)$ distribution with

$a^* = (n+1)$ and $b^* = (\sum_{i=1}^n x_i) - n + 1$

The mean of the posterior distribution is the mean of the $Beta(a^*, b^*)$ distribution which is

$$\begin{aligned} E[p \mid x_1, \ldots, x_n] &= \frac{a^*}{a^* + b^*} \\ &= \frac{n+1}{n+1+(\sum_{i=1}^n x_i) - n + 1} \\ &= \frac{n+1}{\sum_{i=1}^n x_i + 2} \end{aligned}$$

(Note: the book solution is only for $n = 1$ and $x_1 = k$)

3. Problem 8.10.32

The R code in *Problem_8_10_32.html* provides numerical solutions to (a), (b), and (c).

(a). Reasonable guesses for the mean μ and variance σ^2 are the mle for μ and the unbaised sample variance for σ^2.

$\hat{\mu} = \overline{X}$.

$\hat{\sigma}^2 = \frac{1}{n-1} \sum_{i=1}^n (X_i - \overline{X})^2$

(Alternates are reasonable if suggested, e.g., the mle for the variance.)

(b). Confidence intervals for μ :

Let $CI.level$ ($\times 100\%$) be the confidence level of the interval.

Set $\alpha = 1 - CI.level$ and define $t^* = t(\alpha/2, df = (n-1))$, the upper $\alpha/2$ quantile of a t distribution with $n-1$ degrees of freedom. Then the confidence interval for μ is

$$\{\mu : \hat{\mu} - t^* \sqrt{s^2}/\sqrt{n} < \mu < \hat{\mu} + t^* \sqrt{s^2}/\sqrt{n}\}.$$

Confidence intervals for σ^2 :

Let $CI.level$ ($\times 100\%$) be the confidence level of the interval.

Set $\alpha = 1 - CI.level$ and define $\chi^*_{Upper} = \chi(\alpha/2, df = (n-1))$, the upper $\alpha/2$ quantile of a $chi - square$ random variable with $(n-1)$ degrees of freedom, and $\chi^*_{Lower} = \chi(1 - \alpha/2, df = (n-1))$, the lower $\alpha/2$ quantile of the same distribution. The confidence interval is:

$$\{\sigma^2 : \hat{\sigma}^2 \times \frac{n-1}{\chi_{Upper}} < \sigma^2 < \hat{\sigma}^2 \times \frac{n-1}{\chi_{Lower}}\}$$

(c). To compute confidence intervals for σ rather than σ^2, just take the square roots of the corresponding confidence intervals for $sigma^2$.

(d). To halve the confidence interval for μ, the sample size should increase about 4 times. The standard deviation of the sample mean is $\sigma/\sqrt{n}$ so this is halved when the sample size is increased by a factor of 4. This argument does not account for the fact that the critical value of the t distribution of a given $\alpha/2$ level is smaller the higher the degrees of freedom. So, the factor of 4 is an upper bound on the the sample size increase required. Also, the length of the confidence interval for μ is random. The previous comments apply on average, but in any sample, the confidence interval will have random length and could be arbitrarily large, albeit with very small probability.

4. Problem 8.10.63

   Suppose that 100 items are sampled from a manufacturing process and 3 are found to be defective. Let $X = 3$ be the outcome of a $Binomial(n = 100, \theta)$ random variable.

   The likelihood of the data is

   $$Lik(\theta) = \begin{pmatrix} 100 \\ 3 \end{pmatrix} \theta^3(1-\theta)^{97}$$

   Consider a prior distribution for θ which is $Beta(a, b)$, (with $a > 0, b > 0$).

   $$\pi(\theta) = \frac{1}{\beta(a,b)}\theta^{a-1}(1-\theta)^{b-1}.$$

   The posterior distribution has density:

   $$\begin{aligned} \pi(\theta \mid x) &\propto Lik(\theta) \times \pi(\theta) \\ &= [\begin{pmatrix} 100 \\ 3 \end{pmatrix}]\theta^3(1-\theta)^{97}] \times [\tfrac{1}{\beta(a,b)}\theta^{a-1}(1-\theta)^{b-1}] \\ &\propto \theta^{a+3-1}(1-\theta)^{97+b-1} \end{aligned}$$

   Which normalizes to a $Beta(a^*, b^*)$ distribution where

   $a^* = a + 3$, and $b^* = b + 97$.

   The mean of the posterior distribution is

   $$E[\theta \mid X] = a^*/(a^* + b^*).$$

   For $a = b = 1$ we have $E[\theta \mid X] = 4/102$ and

   For $a.5$, and $b = 5$ we have $E[\theta \mid X] = 3.5/105.5$.

   The R code in *Problem_8_10_63.html* plots the prior/posterior densities and computes the posterior means.