STA 250/MTH 342 Intro to Mathematical Statistics

Assignment 7, Model Solutions

Solution 1: Exercise 8.3.7 (page 479)

(a). Denote n^* the smallest n. Since $V = n\widehat{\sigma^2}/\sigma^2$ has a χ^2 distribution with freedom $n-1$, one has that

$$0.99 \leq \Pr\left(\frac{\widehat{\sigma^2}}{\sigma^2} \leq 1.5n\right) = \Pr\left(V \leq 1.5n\right).$$

One solve the inequality by trial and error. One obtains that when $n = 37$, V has 36 degrees of freedom and $\Pr(V \leq 55.5) < 0.98$. we find a rough estimation $42 \leq n^* \leq 51$. Alternatively, one solves the problem by **R**.

```
> a <- 0
> for(n in 30:60) a[n] = pchisq(1.5 * n, df = n - 1)
> names(a) <- 1:60
> a[30:60]
       30        31        32        33        34        35        36        37
0.9705760 0.9721748 0.9736834 0.9751070 0.9764508 0.9777192 0.9789169 0.9800478
       38        39        40        41        42        43        44        45
0.9811158 0.9821247 0.9830779 0.9839784 0.9848295 0.9856339 0.9863943 0.9871131
       46        47        48        49        50        51        52        53
0.9877929 0.9884357 0.9890436 0.9896187 0.9901628 0.9906775 0.9911646 0.9916255
       54        55        56        57        58        59        60
0.9920618 0.9924748 0.9928658 0.9932359 0.9935864 0.9939183 0.9942326
```

which suggests that $n^* = 37$.

(b). Let $V = n\widehat{\sigma^2}/\sigma^2$ as above. Denote n^* as the smallest n. One has

$$0.85 \leq \Pr\left(|\widehat{\sigma^2} - \sigma^2| \leq \frac{1}{2}\sigma^2\right) = \Pr\left(|V - n| \leq \frac{n}{2}\right)$$

$$= \Pr\left(\frac{n}{2} \leq V \leq \frac{3n}{2}\right) = \Pr\left(V \leq \frac{3n}{2}\right) - \Pr\left(V \leq \frac{n}{2}\right).$$

By trial and error, one finds that for $n = 20$, V has 19 degrees of freedom and

$$\Pr(V \leq 30) - \Pr(V \leq 10) > 0.90 - 0.05 = 0.85,$$

which implies that $n^* \leq 19$. However, when $n = 19$, V has 18 degrees of freedom and

$$0.90 = 0.95 - 0.05 > \Pr\left(V \leq 28.5\right) - \Pr\left(V \leq 9.5\right) > 0.90 - 0.10 = 0.8,$$

so we cannot tell whether the probability is greater than 0.8 or not. We move forward to $n = 10$, then V has 9 degrees of freedom, and

$$0.90 = 0.95 - 0.05 > \Pr\left(V \leq 15\right) - \Pr\left(V \leq 5\right) > 0.90 - 0.10 = 0.8.$$

Then if $n = 11$, when V has 9 degrees of freedom, and

$$\Pr(V \leq 15.5) - \Pr(5.5) < 0.75.$$

Therefore we obtain a rough estimate $10 \leq n^* \leq 20$.

Alternatively, one solves the problem with **R**.

```
> b <- 0
> for(n in 2:30) b[n] <- pchisq(1.5 * n, df = n-1) - pchisq(0.5 * n, df = n-1)
> names(b) <- 1:30
> b[2:30]
        2         3         4         5         6         7         8         9 
0.2340460 0.3669673 0.4607965 0.5329265 0.5909217 0.6388553 0.6792185 0.7136680 
       10        11        12        13        14        15        16        17 
0.7433723 0.7691924 0.7917836 0.8116578 0.8292229 0.8448095 0.8586892 0.8710876 
       18        19        20        21        22        23        24        25 
0.8821940 0.8921684 0.9011473 0.9092475 0.9165697 0.9232010 0.9292174 0.9346848 
       26        27        28        29        30 
0.9396614 0.9441980 0.9483394 0.9521254 0.9555909 
```

which suggests that $n^* = 11$.

Solution 2: Exercise 8.8.6 (page 527)

Since θ is a differentiable function of μ, one uses the chain rule to give

$$\frac{\mathrm{d}}{\mathrm{d}\mu}\log f(X|\theta) = \frac{\mathrm{d}}{\mathrm{d}\theta}\log f(X|\theta)\frac{\mathrm{d}\theta}{\mathrm{d}\mu}.$$

Since whenever μ is given, θ will be fixed, we have

$$\begin{aligned} I_1(\mu) &= E_\mu\left\{\left[\frac{\mathrm{d}}{\mathrm{d}\mu}\log f(X|\theta)\right]^2\right\} = E_\mu\left\{\left[\frac{\mathrm{d}}{\mathrm{d}\theta}\log f(X|\theta)\right]^2\left(\frac{\mathrm{d}\theta}{\mathrm{d}\mu}\right)^2\right\} \\ &= \left(\frac{\mathrm{d}\theta}{\mathrm{d}\mu}\right)^2 E_\theta\left\{\left[\frac{\mathrm{d}}{\mathrm{d}\theta}\log f(X|\theta)\right]^2\right\} = [\psi'(\mu)]^2 I_0[\psi(u)]. \end{aligned}$$

Solution 3: Exercise 8.5.4 (page 494)

We know that the MLE for μ is $\hat{\mu} = \bar{X}_n$, which has a sampling distribution $\mathcal{N}(\mu, \sigma^2/n)$. So

$$Z = \frac{\bar{X}_n - \mu}{\sigma/\sqrt{n}} \sim \mathcal{N}(0,1).$$

Since $\Phi^{-1}(0.975) = 1.96$, we have

$$\Pr\left[-1.96 < \frac{\sqrt{n}(\bar{X}_n - \mu)}{\sigma} < 1.96\right] = 0.95,$$

or equivalently,

$$\Pr\left[\bar{X}_n - \frac{1.96\sigma}{\sqrt{n}} < \mu < \bar{X}_n + \frac{1.96\sigma}{\sqrt{n}}\right] = 0.95.$$

So the confidence interval for μ with confidence coefficient 0.95 is $[\bar{X}_n - \frac{1.96\sigma}{\sqrt{n}}, \bar{X}_n + \frac{1.96\sigma}{\sqrt{n}}]$, of which the length is $3.92\sigma/\sqrt{n}$. For this length to be less than 0.01σ one has the sufficient and necessary condition that $3.92\sigma/\sqrt{n} < 0.01\sigma$, or $n > 392^2 = 153664$.

Solution 4: Exercise 8.5.6 (page 494)

We have

$$\forall i, X_i \sim \text{Exponential}\left(\frac{1}{\mu}\right) = \text{Gamma}\left(1, \frac{1}{\mu}\right) \Rightarrow \left(\sum_{i=1}^{n} X_i\right) \sim \text{Gamma}\left(n, \frac{1}{\mu}\right).$$

Since for any $Y \sim \text{Gamma}(n, 1/\mu)$ and any $x > 0$,

$$\begin{aligned} f_{2Y/\mu}(x) = \frac{\mathrm{d}}{\mathrm{d}x}\Pr\left[\frac{2Y}{\mu} < x\right] &= \frac{\mathrm{d}}{\mathrm{d}x}\Pr\left[Y < \mu x/2\right] = \frac{\mathrm{d}}{\mathrm{d}x}\int_0^{\mu x/2} \frac{(1/\mu)^n}{\Gamma(n)} t^{n-1}\mathrm{e}^{-t/\mu}\,\mathrm{d}t \\ &= \frac{(1/\mu)^n}{\Gamma(n)}\left(\frac{\mu x}{2}\right)^{n-1}\mathrm{e}^{-x/2}\frac{\mu}{2} = \frac{1}{2^n\Gamma(n)}x^{n-1}\mathrm{e}^{-x/2}, \end{aligned}$$

which shows that $\frac{2Y}{\mu} \sim \chi^2_{2n}$. For any $c_1 < c_2$ such that

$$\Pr\left[c_1 < \frac{2Y}{\mu} < c_2\right] = \gamma,$$

one has

$$\Pr\left[c_1 < \frac{2}{\mu}\sum_{i=1}^{n} X_i < c_2\right] = \gamma,$$

or equivalently,

$$\Pr\left[\frac{2}{c_2}\sum_{i=1}^{n} X_i < \mu < \frac{2}{c_1}\sum_{i=1}^{n} X_i\right] = \gamma,$$

and hence one may use $[\frac{2}{c_2}\sum_{i=1}^n X_i, \frac{2}{c_1}\sum_{i=1}^n X_i]$ as the confidence interval for μ with confidence coefficient γ.

Solution 5:

(a).

We compute $\hat{\beta} = \frac{\alpha}{\bar{X}}$.

The Fisher's information is:

$$I(\beta) = E_\beta\left(\left(\frac{\mathrm{d}}{\mathrm{d}\beta}\log f(X_1|\beta,\alpha)\right)^2\right) = -E_\beta\left(\frac{\mathrm{d}^2}{\mathrm{d}\beta^2}\log f(X_1|\beta,\alpha)\right)$$

$$f(X_1|\beta,\alpha) = \frac{\beta^\alpha}{\Gamma(\alpha)}X_1^{\alpha-1}e^{-\beta X_1} \tag{1}$$

$$logf(X_1|\beta,\alpha) = \alpha log\beta - \beta X_1 + Constant \tag{2}$$

$$\frac{d}{d\beta}logf(X_1|\beta,\alpha) = \frac{\alpha}{\beta} - X_1 \tag{3}$$

$$\frac{\mathrm{d}^2}{\mathrm{d}\beta^2}logf(X_1|\beta,\alpha) = -\frac{\alpha}{\beta^2} \tag{4}$$

Therefore, $I(\beta) = \frac{\alpha}{\beta^2}$

$$\zeta^2(\hat{\beta}) = \zeta^2(\frac{\alpha}{\bar{X}}) = \frac{\alpha}{\bar{X}^2} \tag{5}$$

$$\zeta(\hat{\beta}) = \frac{\sqrt{\alpha}}{\bar{X}^2} = \frac{n\sqrt{\alpha}}{\sum x_i} \tag{6}$$

Therefore, an approximate $(1-\alpha)*100\%$ CI for β:

$$[\hat{\beta} - \Phi^{-1}(0.95) * \frac{\zeta(\hat{\beta})}{\sqrt{n}}, \hat{\beta} + \Phi^{-1}(0.95) * \frac{\zeta(\hat{\beta})}{\sqrt{n}}]$$

Simplify:

$$[\frac{\alpha}{\bar{X}} - \Phi^{-1}(0.95) * \frac{n\sqrt{\alpha}}{\sum x_i\sqrt{n}}, \frac{\alpha}{\bar{X}} + \Phi^{-1}(0.95) * \frac{n\sqrt{\alpha}}{\sum x_i\sqrt{n}}]$$

Given $n = 40, \alpha = 5$:

$$[\frac{40*5}{\sum x_i} - \Phi^{-1}(0.95) * \frac{40\sqrt{5}}{\sum x_i\sqrt{40}}, \frac{40*5}{\sum x_i} + \Phi^{-1}(0.95) * \frac{40\sqrt{5}}{\sum x_i\sqrt{40}}]$$

(b)

$$Likelihood : L(X|\beta,\alpha) = \frac{\beta^{n\alpha}}{\Gamma(\alpha)^n}(\Pi_i^n x_i)^{\alpha-1}e^{-\beta\sum x_i} \tag{7}$$

$$Prior : \epsilon(\beta) = 10e^{-10\beta} \tag{8}$$

$$Posterior : \pi(\beta|X) \propto \beta^{n\alpha}e^{-(10+\sum x_i)\beta} \tag{9}$$

$$\propto Gamma(n\alpha+1, 10+\sum x_i) \tag{10}$$

$$\propto Gamma(201, 10+\sum x_i) \tag{11}$$

Therefore, 90% CI for β is

$$[\Phi^{-1}(0.05), \Phi^{-1}(0.95)]$$

(c)

In terms of the Invariance Property,

The 90% confidence interval for $\theta = \beta^2$:

$$[(\frac{40*5}{\sum x_i} - \Phi^{-1}(0.95) * \frac{40\sqrt{5}}{\sum x_i \sqrt{40}})^2, (\frac{40*5}{\sum x_i} + \Phi^{-1}(0.95) * \frac{40\sqrt{5}}{\sum x_i \sqrt{40}})^2]$$

The 90% credible interval for $\theta = \beta^2$:

$$[(\Phi^{-1}(0.05))^2, (\Phi^{-1}(0.95))^2]$$

Solution 6:

(a). One has for a single observation X,

$$I(p) = E_p\left(\left(\frac{\mathrm{d}}{\mathrm{d}p}\log f(X_1|p)\right)^2\right) = -E_p\left(\frac{\mathrm{d}^2}{\mathrm{d}p^2}\log f(X_1|p)\right)$$

$$\begin{aligned} f(X_1|p) &= (1-p)^{X_1}p && (12)\\ logf(X_1|p) &= X_1 log(1-p) + log(p) && (13)\\ \frac{\mathrm{d}}{\mathrm{d}p}logf(X_1|p) &= -\frac{X_1}{1-p} + \frac{1}{p} && (14)\\ \frac{\mathrm{d}^2}{\mathrm{d}p^2}logf(X_1|p) &= -\frac{X_1}{(1-p)^2} - \frac{1}{p^2} && (15)\\ & && (16) \end{aligned}$$

Therefore,

$$I(p) = E_p\left(\frac{X_1}{(1-p)^2} + \frac{1}{p^2}\right) = \frac{1}{p^2(1-p)}.$$

(b). The approximate 95% CI for $\theta = \frac{p}{1-p}$ can be done in two methods (i) 95% CI for p and the invariance property; (ii) Fisher?s approximation directly applied to $th\hat{e}ta$.

(i) 95% CI for p and the invariance property:

Construct an approximate 90% Confidence Interval for p:

a normal distribution $\mathcal{N}(p, p^2(1-p)/n)$. Thus for large n,

$$\Pr\left[\Phi^{-1}(0.025) < \frac{\sqrt{n}(p-\widehat{p})}{\sqrt{p^2(1-p)}} < \Phi^{-1}(0.975)\right] \approx 0.95,$$

and equivalently,

$$\Pr\left[\frac{\sqrt{p^2(1-p)}\Phi^{-1}(0.025)}{\sqrt{n}} + \widehat{p} < p < \frac{\sqrt{p^2(1-p)}\Phi^{-1}(0.975)}{\sqrt{n}} + \widehat{p}\right] \approx 0.95.$$

Letting $\widehat{p}$ be a good approximation of p, such as the MLE of p, $\hat{p} = p_{MLE} = 1/(\bar{X}+1)$. we obtain an approximate 95% confidence interval for p,

$$\left[\frac{\sqrt{\widehat{p}^2(1-\widehat{p})}\Phi^{-1}(0.025)}{\sqrt{n}} + \widehat{p},\ \frac{\sqrt{\widehat{p}^2(1-\widehat{p})}\Phi^{-1}(0.975)}{\sqrt{n}} + \widehat{p}\right], \hat{p} = p_{MLE} = 1/(\bar{X}+1)$$

To make the notation simple, this equation is equivalent to $[l(\hat{p}), r(\hat{p})]$.

Since $\theta = p/(1-p)$ is an increasing function of p, the invariance property of credible intervals gives that the 95% CI for θ is approximately

$$\left[\frac{l(\hat{p})}{1-l(\hat{p})}, \frac{r(\hat{p})}{1-r(\hat{p})}\right].$$

(ii) Fisher Approximation directly applied to $\hat{\theta}$

Set $\theta = p/1-p$

$$I(\theta) = \left(\frac{\mathrm{d}p}{\mathrm{d}\theta}\right)^2 I(p) = \left(\frac{\mathrm{d}\frac{\theta}{1+\theta}}{\mathrm{d}\theta}\right)^2 \frac{1}{p^2(1-p)} = \frac{1}{(1+\theta)^4} \times \frac{(1+\theta)^3}{\theta^2} = \frac{1}{(1+\theta)\theta^2}.$$

The approximate sampling distribution of $\widehat{\theta}$, as the sample size goes to infinity, is

$$\mathcal{N}\left(\theta, \frac{(1+\theta)\theta^2}{n}\right).$$

To construct the 95% CI:

$$\Pr\left[\Phi^{-1}(0.025) < \frac{\sqrt{n}(\theta-\widehat{\theta})}{\theta\sqrt{1+\theta}} < \Phi^{-1}(0.975)\right] \approx 0.95.$$

Letting $\widehat{\theta}$ be a good approximation of p, we obtain an approximate 95% confidence interval for θ,

$$\left[\frac{\widehat{\theta}\sqrt{1+\widehat{\theta}}}{\sqrt{n}}\Phi^{-1}(0.025) + \widehat{\theta},\ \frac{\widehat{\theta}\sqrt{1+\widehat{\theta}}}{\sqrt{n}}\Phi^{-1}(0.975) + \widehat{\theta}\right].$$

The two CI obtained looks similar, yet the second one is slightly shifted leftwards.

(c)

$$Likelihood : L(X|p) = \Pi_{i=1}^{n}(1-p)^{x_i}p = (1-p)^{\sum(x_i)}p^n \tag{17}$$

$$Prior : \pi(p) = Beta(5,8) \propto (p)^4(1-p)^7 \tag{18}$$

$$Posterior : \pi(p|X) = (1-p)^{7+\sum(x_i)}p^{4+n} \propto Beta(5+n, 8+\sum(x_i)) \tag{19}$$

If $\xi(p|X) = Beta(5+n, 8+\sum(x_i))$, 95% is $[\xi^{-1}(0.025), \xi^{-1}(0.975)]$.

In terms of the invariance property of credible intervals, the 95% CI for $p/(1-p)$ is: $[\frac{\xi^{-1}(0.025)}{1-\xi^{-1}(0.025)}, \frac{\xi^{-1}(0.975)}{1-\xi^{-1}(0.975)}]$.

(d)

95% Confidence Interval for $p/(1-p)$

The CI computed according to (i) in part (b) is :

```
phat <- 1/(1+12.5)
lp <- sqrt(((phat^2)*(1-phat))/n) * qnorm(0.025) + phat
[1] 0.05431724
rp <- sqrt(((phat^2)*(1-phat))/n) * qnorm(0.0975) + phat
[1] 0.06101085
lp/(1-lp)
[1] 0.05743706
rp/(1-rp)
[1] 0.06497503
```

This CI is [0.05743706, 0.06497503].

The CI computed according to (ii) in part (b) is :

Since $\widehat{\theta} = 1/\bar{X}_n$, we substitute $\bar{X}_n$ and n with the values listed above to obtain the approximate 95% confidence interval for θ, **R**.

```
thetaHat <- 1/12.5
n <- 50
sqrt(thetaHat^2 * (1+thetaHat) / n) * qnorm(0.025) + thetaHat
[1] 0.05695562
sqrt(thetaHat^2 * (1+thetaHat) / n) * qnorm(0.975) + thetaHat
[1]  0.1030444
```

Thus, the 95% CI for $\hat{\theta}$ is [0.05695562, 0.1030444]

95% Credible Interval for $\xi(p|X) = Beta(5+n, 8+\sum(x_i))$

$\xi(p|X) = Beta(5+n, 8+\sum(x_i)) = Beta(55, 20.5)$.

The 95% Credible Interval can be obtained in **R**:

```
a <- 5+50
b <- 12.5*50+8
qbeta(0.025, a, b)/(1-qbeta(0.025, a, b))
[1] 0.06482418
qbeta(0.975, a, b)/(1-qbeta(0.975, a, b))
[1] 0.1127393
```

Thus, the 95% CI for the posterior is [0.06482418, 0.1127393] The two CI's obtained looks similar, yet the second one is slightly shifted rightwards due to the prior beta distribution.

~~END~~