Homework 4

Name: Shashank Singh[1]
36-705 Intermediate Statistics
Due: Thursday, October 2, 2014

1. (a) Let $\boxed{T := (X_{(1)}, X_{(n)}),}$ where $X_{(i)}$ denotes the i^{th} order statistic. Note that

$$p(x^n; \theta) = \prod_{i=1}^{n} 1_{\{x_i \in (\theta, \theta+1)\}} = 1_{\{\theta < x_{(1)}\}} 1_{\{x_{(n)} < \theta+1\}}$$

(where 1_A denotes the indicator function of and event A), which is a function only of T and θ. Hence T is sufficient. Suppose S is a sufficient statistic. Then, there exist functions h_S and g_S such that $p(x^n; \theta) = h(x^n)g(S; \theta)$. As long as $h(x^n) > 0$ (which happens almost surely)

$$x_{(1)} = \sup\{\theta \in \mathbb{R} : g(S; \theta) > 0\} \quad \text{and} \quad x_{(n)} = 1 + \inf\{\theta \in \mathbb{R} : g(S; \theta) > 0\}.$$

Hence T is a function of S, and so T is minimal. ■

(b) Since $T := (X_{(1)}, X_{(n)})$ is a minimal sufficient statistic and T is clearly not a function of X_3, X_3 is not sufficient. ■

2. Since

$$\mathbb{E}[\hat{\lambda}] = \frac{1}{n} \sum_{i=1}^{n} \mathbb{E}[X_i] = \frac{1}{n} \sum_{i=1}^{n} \lambda = \lambda,$$

$\boxed{\mathsf{bias}(\hat{\lambda}) = 0.}$

$$\mathsf{se}(\hat{\lambda}) = \sqrt{\frac{1}{n^2} \sum_{i=1}^{n} \mathbb{V}[X_i]} = \sqrt{\frac{1}{n^2} \sum_{i=1}^{n} \lambda} = \boxed{\sqrt{\frac{\lambda}{n}}.}$$

Hence, $\mathsf{bias}(\hat{\lambda}) = \mathsf{bias}^2(\hat{\lambda}) + \mathsf{se}^2(\hat{\lambda}) = \boxed{\lambda/n.}$

[1] sss1@andrew.cmu.edu

3. (a) $\mathbb{E}[X_i] = \frac{a+b}{2}$ and

$$\mathbb{E}[X_i^2] = \mathbb{V}[X_i] + \mathbb{E}^2[X_i] = \frac{1}{12}(b-a)^2 + \frac{1}{4}(a+b)^2 = \frac{1}{3}(a^2 + ab + b^2).$$

Solving for a and b gives

$$a = \mathbb{E}[X_i] - \sqrt{3(\mathbb{E}[X^2] - \mathbb{E}^2[X])} \quad \text{and} \quad b = \mathbb{E}[X_i] + \sqrt{3(\mathbb{E}[X^2] - \mathbb{E}^2[X])}.$$

Hence, the method of moments estimators for a and b are

$$\boxed{\tilde{a} = \overline{X} - \sqrt{3(\overline{X^2} - \overline{X}^2)}} \quad \text{and} \quad \boxed{\tilde{b} = \mathbb{E}[X_i] + \sqrt{3(\overline{X^2} - \overline{X}^2)}.}$$

(b) Since

$$p(X_1, \ldots, X_n|a, b) = (b-a)^{-n} 1_{a<X_1,\ldots,X_n<b},$$

the likelihood is maximized when $b - a$ is minimized, subject to $a < X_{(1)}$ and $X_{(n)} < b$. Hence, the MLEs of a and b are $\boxed{\hat{a} = X_{(1)}}$ and $\boxed{\hat{b} = X_{(n)}.}$

(c) Since $\tau = (a+b)/2$, the MLE of τ is

$$\hat{\tau} = \frac{\hat{a} + \hat{b}}{2} = \frac{X_{(1)} + X_{(n)}}{2},$$

where $\hat{a}$ and $\hat{b}$ denote the MLEs of a and b, respectively.

4. (a) Since the normal is its own conjugate prior,

$$\mu|X^n \sim \mathcal{N}\left(\frac{b^2}{b^2 + \sigma^2/n}\overline{X} + \frac{\sigma^2/n}{b^2 + \sigma^2/n}a, \frac{b^2\sigma^2}{\sigma^2 + b^2 n}\right)$$

(this is a pretty standard result, but is really algebraically messy). Hence,

$$\hat{\mu} = \mathbb{E}[\mu|X^n] = \frac{b^2}{b^2 + \sigma^2/n}\overline{X} + \frac{\sigma^2/n}{b^2 + \sigma^2/n}a.$$

(b) Under squared error loss, the risk is $R(\mu, \hat{\mu}) = \mathbb{E}^2[\hat{\mu} - \mu] + \mathbb{V}[\hat{\mu}]$, where

$$\mathbb{E}[\hat{\mu} - \mu] = \frac{b^2}{b^2 + \sigma/n}\mathbb{E}[\overline{X}] + \frac{\sigma^2/n}{b^2 + \sigma}a - \mu = \left(\frac{b^2}{b^2 + \sigma/n} - 1\right)\mu + \frac{\sigma^2/n}{b^2 + \sigma}a$$

and

$$\mathbb{V}[\hat{\mu}] = \frac{b^2}{b^2 + \sigma/n}\mathbb{V}[\overline{X}] = \frac{b^2\sigma^2/n}{b^2 + \sigma/n}.$$

(c) As can be seen from part (b), $R(\mu, \hat{\mu}) \geq \mathbb{E}^2[\mu - \hat{\mu}] \to \infty$ as $\mu \to \infty$.

(d) Note that, for the Bayes estimator, the posterior risk is

$$r(\hat{\mu}|X^n) = \mathbb{V}[\mu|X^n] = \frac{b^2\sigma^2}{\sigma^2 + nb^2}.$$

Since this does not depend on X^n,

$$B_\pi(\hat{\mu}) = \int_{\mathbb{R}^n} r(\hat{\mu}|X^n)m(X^n)\,dX^n = \boxed{\frac{b^2\sigma^2}{\sigma^2 + nb^2}.}$$

5. (a) Define $S := \sum_{i=1}^n X_i$. Then,

$$\pi(p|X^n) = \frac{\pi(X^n|p)\pi(p)}{\pi(X^n)} \propto \left(p^S(1-p)^{n-S}\right)\left(p^{\alpha-1}(1-p)^{\beta-1}\right) = p^{S+\alpha-1}(1-p)^{n-S+\beta-1},$$

which is proportional to the pdf of Beta$(S+\alpha, n-S+\beta)$, so that $p|X^n \sim$ Beta$(S+\alpha, n-S+\beta)$. Hence, the Bayes estimator is

$$\hat{p} = \mathbb{E}[p|X^n] = \frac{S+\alpha}{S+\alpha+n-S+\beta} = \boxed{\frac{S+\alpha}{n+\alpha+\beta}.}$$

(b) Under squared error loss, the risk is $R(p,\hat{p}) = \mathbb{E}^2[\hat{p}-p] + \mathbb{V}[\hat{p}]$, where

$$\mathbb{E}[\hat{p}-p] = \frac{\mathbb{E}[S]+\alpha}{n+\alpha+\beta} - p = \frac{pn+\alpha-pn-p\alpha-p\beta}{n+\alpha+\beta} = \frac{(1-p)\alpha-p\beta}{n+\alpha+\beta}$$

and

$$\mathbb{V}[\hat{p}] = \frac{\mathbb{V}[S]}{(n+\alpha+\beta)^2} = \frac{np(1-p)}{(n+\alpha+\beta)^2}$$

(since $S \sim$ Binomial(n,p)).

(c) The Bayes risk is

$$B_\pi(\hat{p}) = \int_0^1 R(p,\hat{p})\pi(p)\,dp = \frac{\int_0^1(((1-p)\alpha-p\beta)^2+np(1-p))p^{\alpha-1}(1-p)^{\beta-1}\,dp}{(n+\alpha+\beta)^2},$$

which looks like an awfully nasty integral.

(d) Setting $\boxed{\alpha = \beta = \sqrt{n}/2,}$

$$R(p,\hat{p}) = \frac{((1-p)\alpha-p\beta)^2+np(1-p)}{(n+\alpha+\beta)^2} = \frac{(n/4)(1-4p+4p^2)+np-np^2}{(n+\alpha+\beta)^2} = \frac{n}{4(n+\sqrt{n})^2},$$

which does not depend on p and is hence minimax optimal.