UC Berkeley
Department of Statistics

STAT 210A: Introduction to Mathematical Statistics

Midterm Solutions
Fall 2007

Problem 1.1
(15 points) Suppose that $X_1, \ldots, X_n$ are i.i.d. discrete random variables with the PMF

$$p(x;\theta) \quad = \quad \begin{cases} \theta & \text{for } x = -1 \\ (1-\theta)^2\theta^x & \text{for } x = 0, 1, 2, \ldots, \end{cases}$$

where $0 < \theta < 1$.

(a) 4 pts Find the Cramér-Rao lower bound for unbiased estimators of θ based on $X_1, \ldots, X_n$.

(b) 4 pts Compute the maximum likelihood estimate of θ based on $(X_1, \ldots, X_n)$.

(c) 3 pts Show that $(X_1, \ldots, X_n)$ can be written as a 2-dimensional exponential family in terms of $T_1(x) = \sum_{i=1}^n \mathbb{I}[x_i = -1]$ and $T_2(x) = \sum_{i=1}^n x_i \mathbb{I}[x_i \geq 0]$.

(d) 4 pts Is (T_1, T_2) complete? How does this relate to our result from class on completeness in full rank exponential families?

Solution 1.1: Define the statistics $T_1 = \sum_{i=1}^n \mathbb{I}(x_i = -1)$ and $T_2 = \sum_{i=1}^n x_i \mathbb{I}(x_i \geq 0)$. We then have

$$\begin{aligned} \prod_{i=1}^n p(x_i;\theta) &= \prod_{i=1}^n \left\{ \theta^{\mathbb{I}(x_i=-1)} \left((1-\theta)^2\theta^{x_i}\right)^{\mathbb{I}(x_i \geq 0)} \right\} \\ &= \exp\left\{T_1\left[\log\theta - 2\log(1-\theta)\right] + T_2\log\theta + 2n\log(1-\theta)\right\} \end{aligned}$$

(a) From the given representation as an exponential family, we can compute the log likelihood and its derivatives:

$$\begin{aligned} \ell(\theta) &= T_1(\log\theta - 2\log(1-\theta)) + T_2\log\theta + 2n\log(1-\theta) \\ \frac{\partial \ell}{\partial\theta} &= T_1\left(\frac{1}{\theta} + \frac{2}{1-\theta}\right) + \frac{T_2}{\theta} - \frac{2n}{1-\theta} \\ \frac{\partial^2\ell}{\partial\theta^2} &= -\frac{T_1+T_2}{\theta^2} - \frac{2(n-T_2)}{(1-\theta)^2}. \end{aligned}$$

We also have $\mathbb{E}(T_1) = n\mathbb{P}(X_1 = -1) = n\theta$ and

$$\mathbb{E}(T_2) \;=\; n\mathbb{E}(X_i\mathbf{1}(X_i \geq 0)) \;=\; n(1-\theta)^2\sum_{x=0}^{\infty} x\theta^x = n\theta.$$

Consequently, we have

$$I(\theta) \quad = \quad -\mathbb{E}\Big[\frac{\partial^2\ell}{\partial\theta^2}\Big] \;=\; \frac{2n}{\theta(1-\theta)},$$

so that the Cramer-Rao bound is $\text{var}(\delta) \geq \theta(1-\theta)/2n$.

(b) From our calculation of the second derivative of ℓ, we see that $\frac{\partial^2 l}{\partial \theta^2} \leq 0$. Consequently, the MLE is given by a zero point of the gradient, or equivalently

$$\hat{\theta}^{MLE} = \frac{T_1 + T_2}{2n - T_1 + T_2}.$$

(Note that if $X_i = 0$ for all $i = 1, \ldots, n$, then $\hat{\theta}^{MLE} = 0$, and if $X_i = -1$ for all $i = 1, \ldots, n$, then $\hat{\theta}^{MLE} = 1$.)

(c) Our calculations prior to (a) show this fact.

(d) Defining $g(T_1, T_2) = T_1 - T_2 \neq 0$, we see that $\mathbb{E}[g(T_1, T_2)] = 0$ for all θ, so that (T_1, T_2) is not complete. The result from class does not apply because the family is not full rank. (Indeed, if we set $\eta_1 = \log \theta - 2\log(1-\theta)$ and $\eta_2 = \log\theta$, then these parameters satisfy the constraint $\eta_1 = \eta_2 - 2\log(1 - e^{\eta_2})$.)

Problem 1.2
(18 points) Let $(X_1, \ldots, X_n)$ be an i.i.d. sample of Bernoulli random variables with $\mathbb{P}(X_i = 1) = \theta$. For parts (a) and (b), assume that $\theta \in (0, \frac{1}{2})$.

(a) 2 pts Specify the maximum likelihood estimate of $g(\theta) = \theta\,(1-\theta)$ based on $(X_1, \ldots, X_n)$.

(b) 3 pts Compute the Cramér-Rao bound for unbiased estimation of $g(\theta)$ under squared error loss.

Now assume that $\theta \in (0, 1)$, and consider estimation of $g(\theta)$ under squared error loss.

(c) 4 pts Find the UMVUE of $g(\theta)$.

(d) 4 pts Suppose that Θ is modeled as random, with a Beta(a, b) prior. Obtain the Bayes estimator of $g(\theta)$. Is this estimator unbiased? Is it consistent as $n \to +\infty$?

(e) 5 pts Consider the improper density $\lambda(\theta) = [\theta(1-\theta)]^{-1}$ on $(0, 1)$. Give necessary and sufficient conditions on $(X_1, \ldots, X_n)$ for the posterior distribution over Θ to be proper. When the posterior is proper, find the generalized Bayes estimator associated with this improper prior.

Note: The Beta(a, b) distribution has density $\lambda(\theta) = \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)}\theta^{a-1}(1-\theta)^{b-1}$ for $\theta \in (0, 1)$. Its mean is $\mathbb{E}[\Theta] = a/(a+b)$.

Solution 1.2: Define the statistic $T = \sum_{i=1}^n x_i$. We then have

$$\prod_{i=1}^n p(x_i; \theta) \;=\; \prod_{i=1}^n \theta^{x_i}(1-\theta)^{1-x_i} \;=\; \exp\left\{T \log \frac{\theta}{1-\theta} + n \log(1-\theta)\right\}. \tag{1}$$

(a) Since $g(\theta)$ is invertible on $[0, \frac{1}{2}]$, the MLE $\widehat{g(\theta)}$ is equal to $g(\widehat{\theta})$, where $\widehat{\theta}$ is the MLE of $\theta \in [0, \frac{1}{2}]$. We compute

$$\begin{aligned}
\ell(\theta) &= T \log \frac{\theta}{1-\theta} + n \log(1-\theta) \\
\frac{\partial \ell}{\partial \theta} &= T\left(\frac{1}{\theta} + \frac{1}{1-\theta}\right) - \frac{n}{1-\theta} = \frac{T}{\theta(1-\theta)} - \frac{n}{1-\theta} \\
\frac{\partial^2 \ell}{\partial \theta^2} &= -\frac{T}{\theta^2} - \frac{n-T}{(1-\theta)^2}.
\end{aligned}$$

Consequently, the MLE of θ restricted to $[0, \frac{1}{2}]$ is

$$\widehat{\theta} \;=\; \begin{cases} \frac{1}{n}\sum_{i=1}^n X_i & \text{if } \frac{1}{n}\sum_{i=1}^n X_i \leq \frac{1}{2} \\ \frac{1}{2} & \text{otherwise.} \end{cases}$$

(b) From our calculations in (a), we have $I_n(\theta) = \mathbb{E}\left(\frac{T}{\theta^2} + \frac{n-T}{(1-\theta)^2}\right) = \frac{n}{\theta(1-\theta)}$ and $g'(\theta) = 1 - 2\theta$. Thus, the Cramer-Rao bound for unbiased estimators of $g(\theta)$ is $\dfrac{\theta(1-\theta)(1-2\theta)^2}{n}$.

(c) Since this is a full rank exponential family, $T = \sum_{i=1}^{n} \mathbb{I}(x_i = 1)$ is a complete and sufficient statistics for θ. Therefore, a function $f(T)$ that is unbiased for $g(\theta)$ is a UMVUE. We compute

$$\begin{aligned}
\mathbb{E}\left[\frac{T\,(n-T)}{n^2}\right] &= \frac{1}{n^2}\left[n\mathbb{E}[T] - \mathbb{E}[T^2]\right] \\
&= \frac{1}{n^2}\left[n^2\theta - \mathrm{var}(T) - (\mathbb{E}[T])^2\right] \\
&= \frac{1}{n^2}\left[n^2\theta - n\theta(1-\theta) - n^2\theta^2\right] \\
&= \frac{n-1}{n}\theta(1-\theta).
\end{aligned}$$

Thus, $f(T) = \frac{T(n-T)}{n(n-1)}$ is the UMVUE of $g(\theta)$.

(d) The posterior distribution of θ given $X_1, \ldots, X_n$ is Beta$(T+a, n-T+b)$. The Bayes estimate of $g(\theta)$ under squared error loss is

$$\mathbb{E}(\Theta(1-\Theta)|X) = \mathbb{E}[\Theta] - \mathbb{E}[\Theta^2].$$

From the given formula, we have $\mathbb{E}[\Theta|X] = \frac{T+a}{n+a+b}$. We compute

$$\begin{aligned}
\mathbb{E}[\Theta^2|X] &= \frac{\Gamma(a+b+n)}{\Gamma(T+a)\Gamma(n-T+b)}\int \theta^2\theta^{T+a-1}(1-\theta)^{n-T+b-1}d\theta \\
&= \frac{\Gamma(a+b+n)}{\Gamma(T+a)\Gamma(n-T+b)}\int \theta^{T+a+1}(1-\theta)^{n-T+b-1}d\theta \\
&= \frac{\Gamma(a+b+n)}{\Gamma(T+a)\Gamma(n-T+b)}\frac{\Gamma(T+a+2)\Gamma(n-T+b)}{\Gamma(n+a+b+2)} \\
&= \frac{(T+a+2)(T+a+1)}{(n+a+b+2)(n+a+b+1)}.
\end{aligned}$$

Overall, we have

$$\mathbb{E}[g(\Theta) \mid X] = \frac{(T+a)(n-T+b)}{(n+a+b)(n+a+b+1)}.$$

From the calculation in part (c), we see that this estimate is biased. However, since $T/n \xrightarrow{p} \theta$ as n goes to infinity, it is consistent.

(e) The posterior distribution of θ given $X_1, \ldots, X_n$ is proportional to $\theta^{T-1}(1-\theta)^{n-T-1}$ Note that

$$\int_0^\infty \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)}x^{\alpha-1}(1-x)^{\beta-1}dx < \infty \quad\Leftrightarrow\quad \alpha > 0,\ \beta > 0$$

Thus, necessary and sufficient conditions for the posterior to be proper are $T > 0$ and $T < n$. *i.e.* $1 \leq T \leq n-1$. Since the posterior is like part (d) with $a = b = 0$, the generalized Bayes estimator is $\frac{T(n-T)}{n(n+1)}$.

Problem 1.3

(17 pts) For some $n > 2$, let $(X_1, X_2, \ldots, X_n)$ be an i.i.d. sample from the uniform distribution on $[\theta - \Delta, \theta + \Delta]$, where $\theta \in \mathbb{R}$ and $\Delta > 0$.

(a) 3 pts Suppose that $\Delta = \frac{1}{2}$ is known, but θ is unknown. Find a minimal sufficient statistic T for the family parameterized by θ. Is T complete for the family? Justify your answer.

(b) 3 pts Suppose that $\theta = 0$ is known, but $\Delta > 0$ is unknown. Find a minimal sufficient statistic S for the family parameterized by Δ. Is S complete for the family? Justify your answer.

For the remaining parts, suppose that both (θ, Δ) are unknown.

(c) 3 pts Let $X_{(i)}$ denote the i^{th} order statistic. Show that $U = (X_{(1)}, X_{(n)})$ is complete and sufficient for the family parameterized by (θ, Δ).

(d) 5 pts Find UMVUEs of θ and Δ.

(e) 3 pts Find a UMVUE of θ/Δ.

Solution 1.3:

(a) Note that $\prod_{i=1}^n p(x_i; \theta) = \mathbb{I}\left(x_{(n)} - \frac{1}{2} \leq \theta\right) \cdot \mathbb{I}\left(\theta \leq x_{(1)} + \frac{1}{2}\right)$, so that $(X_{(1)}, X_{(n)})$ is sufficient by the factorization criterion. By likelihood ratio criterion, it is also minimal sufficient. However, it is not complete since $\mathbb{E}[X_{(n)} - X_{(1)}] = 1 - \frac{2}{n+1}$ for all θ, but $X_{(n)} - X_{(1)} \neq 0$.

(b) In this case, we have $\prod_{i=1}^n p(x_i; \Delta) = \left(\frac{1}{2\Delta}\right)^n \mathbb{I}\left(x_{(n)} \leq \Delta\right) \cdot \mathbb{I}\left(\Delta \geq x_{(1)}\right)$. Thus $(X_{(1)}, X_{(n)})$ is sufficient by the factorization criterion, and minimal sufficient by the likelihood ratio criterion. However, $(X_{(1)}, X_{(n)})$ is not complete since we have $\mathbb{E}[X_1 + X_{(n)}] = 0$ for all Δ (by symmetry), yet $X_{(1)} + X_{(n)} \neq 0$.

(c) In this case, the distribution can be factored as

$$\prod_{i=1}^n p(x_i; \theta, \Delta) = \left(\frac{1}{2\Delta}\right)^n \mathbb{I}\left(x_{(n)} \leq \theta + \Delta\right) \, \mathbb{I}\left(-\Delta + \theta \geq x_{(1)}\right),$$

so that $(X_{(1)}, X_{(n)})$ is again sufficient by the factorization criterion. We claim that it is also complete: indeed, suppose that g is some function of $(X_{(1)}, X_{(n)})$ such that $\mathbb{E}[g(X_{(1)}, X_{(n)})] = 0$ for all (θ, Δ). We have

$$\mathbb{P}[X_{(1)} \geq t, X_{(n)} \leq s] \quad = \quad \prod_{i=1}^n \mathbb{P}[t \leq X_i \leq s] \quad = \quad \frac{(s-t)^n}{(2\Delta)^n} \mathbb{I}[t \geq \theta - \Delta] \mathbb{I}[s \leq \theta + \Delta].$$

Consequently, their joint density is

$$p(x_{(1)}, x_{(n)}) \quad = \quad \frac{n(n-1)}{(2\Delta)^n} (x_{(n)} - x_{(1)})^{n-2}$$

for $(x_{(1)}, x_{(n)}) \in (\theta - \Delta, x_{(n)}) \times (\theta - \Delta, \theta + \Delta)$. With this, we have

$$\begin{aligned}
& \mathbb{E}\left[g(X_{(1)}, X_{(n)})\right] = 0 \\
\Leftrightarrow\ & \int_{\theta-\Delta}^{\theta+\Delta}\int_{\theta-\Delta}^{x_{(n)}} g(x_{(1)}, x_{(n)}) n(n-1)(x_{(n)} - x_{(1)})^{n-2}(2\Delta)^{-n} dx_{(1)} dx_{(n)} = 0,\ \forall\ \theta \in \mathbb{R},\ \forall \Delta > 0 \\
\Leftrightarrow\ & \int_{\theta-\Delta}^{\theta+\Delta}\int_{0}^{x_{(n)}-\theta+\Delta} g(x_{(n)} - t, x_{(n)}) t^{n-2} dt dx_{(n)} = 0,\ \forall\ \theta \in \mathbb{R},\ \forall \Delta > 0 \\
\Leftrightarrow\ & \int_{\eta}^{\eta+2\Delta}\int_{0}^{x_{(n)}-\eta} g(x_{(n)} - t, x_{(n)}) t^{n-2} dt dx_{(n)} = 0,\ \forall\ \eta = \theta - \Delta \in \mathbb{R},\ \forall \Delta > 0 \\
\Leftrightarrow\ & \int_{0}^{x_{(n)}-\eta} g(x_{(n)} - t, x_{(n)}) t^{n-2} dt dx_{(n)} = 0,\ \forall\ \eta = \theta - \Delta \in \mathbb{R}, \\
\Leftrightarrow\ & g(x, y) = 0\ \forall\ x, y,
\end{aligned}$$

which implies that $(X_{(1)}, X_{(n)})$ is complete.

(d) Since $(X_{(1)}, X_{(n)})$ is complete and sufficient, we can obtain UMVUEs as unbiased functions of these quantities. To simplify calculation, define $Y_i = \frac{X_i - \theta}{2\Delta}$. Since each Y_i is uniform on $[-\frac{1}{2}, \frac{1}{2}]$, we have $\mathbb{E}\left[Y_{(1)} + Y_{(n)}\right] = 0$ and $\mathbb{E}\left[Y_{(n)} - Y_{(1)}\right] = \dfrac{n-1}{n+1}$. This implies that $\mathbb{E}\left[\frac{X_{(1)}+X_{(n)}}{2}\right] = \theta$ and $\mathbb{E}\left[\dfrac{X_{(n)} - X_{(1)}}{2}\dfrac{n+1}{n-1}\right] = \Delta$.

Therefore, we conclude that $\delta_1(X_{(1)}, X_{(n)}) = \frac{X_{(1)}+X_{(n)}}{2}$ is the UMVUE of θ, and $\delta_2(X_{(1)}, X_{(n)}) = \frac{X_{(n)}-X_{(1)}}{2}\frac{n+1}{n-1}$ is the UMVUE of Δ.

(e) Using the joint density of $(X_{(1)}, X_{(n)})$, we compute

$$\begin{aligned}
\mathbb{E}\left[\frac{1}{Y_{(n)} - Y_{(1)}}\right] &= \int_{-1/2}^{1/2}\int_{-1/2}^{x_{(n)}} \frac{1}{x_{(n)} - x_{(1)}} n(n-1)(x_{(n)} - x_{(1)})^{n-2} dx_{(1)} dx_{(n)} \\
&= \int_{-1/2}^{1/2}\int_{-1/2}^{x_{(n)}+1/2} n(n-1) t^{n-3} dt dx_{(n)} \\
&= \frac{n}{n-2}
\end{aligned}$$

Due to symmetry, we have $\mathbb{E}\left[\frac{Y_{(1)}+Y_{(n)}}{Y_{(n)}-Y_{(1)}}\right] = 0$. Therefore, we have

$$\begin{aligned}
\mathbb{E}\left[\frac{n-2}{n}\frac{2}{X_{(n)} - X_{(1)}}\right] &= \frac{1}{\Delta}, \quad \text{and} \\
\mathbb{E}\left[\frac{X_{(1)} + X_{(n)}}{X_{(n)} - X_{(1)}}\right] &= \theta\mathbb{E}\left[\frac{2}{X_{(n)} - X_{(1)}}\right] = \frac{\theta}{\Delta}\frac{n}{n-2}.
\end{aligned}$$

Overall, we conclude that $\delta_3(X_{(1)}, X_{(n)}) = \dfrac{n-2}{n} \cdot \dfrac{X_{(1)} + X_{(n)}}{X_{(n)} - X_{(1)}}$ is the UMVUE of $\dfrac{\theta}{\Delta}$.