UC Berkeley

Department of Statistics

STAT 210A: Introduction to Mathematical Statistics

Problem Set 1- Solutions

Issued: September 2

Due: September 9

Problem 1.1

Due to independence, the joint distribution function of $(X_1, ..., X_n)$ could be written as:

$$
\begin{aligned}
P(x_1, ..., x_n) &= \prod_i e^{-\lambda_i} \frac{\lambda_i^{x_i}}{x_i!} \mathbf{1}(x_i \in N) \\
&= (\prod_i \frac{\mathbf{1}(x_i \in N)}{x_i!}) \exp\{\sum_i x_i \ln \lambda_i - \sum_i \lambda_i\} \\
&= (\prod_i \frac{\mathbf{1}(x_i \in N)}{x_i!}) \exp\{\alpha \sum_i x_i + \beta \sum_i x_i t_i - \sum_i e^{\alpha+\beta t_i}\}[\text{plug in definition of } \lambda_i].
\end{aligned}
$$

It is belongs to a two-parameter exponential family, where $T_1 = \sum_i X_i$ and $T_2 = \sum_i t_i X_i$.

Problem 1.2

(a) To show Y has a density wrt μ, we need to show P_Y is absolutely continuous wrt to μ which is to say for any set A, $\mu(A) = 0$ leads to $P_\theta(Y \in A) = 0$. By definition, $P_\theta(Y \in A) = P_\theta(X \in A|X \in S) = P_\theta(X \in A \cap S)/P_\theta(X \in S) \leq P_\theta(X \in A)/P_\theta(X \in S)$, the last quantity equals 0 because $P_\theta(X)$ is absolutely continuous to μ. Therefore $P_\theta(Y \in A) = 0$ if $\mu(A) = 0$, then the RadonNikodym theorem tells us that Y has a density wrt μ. By similar argument as above the density should be:

$$q_\theta(y) = \frac{1}{\int_S P_\theta(x)dx} h(y) \exp\{\sum_i \eta_i(\theta) T_i(y) - B(\theta)\} \mathbf{1}(y \in S).$$

(b) Notice that

$$q_\theta(y) = h(y)\mathbf{1}(y \in S) \exp\{\sum_i \eta_i(\theta) T_i(y) - \big(B(\theta) - \log \int_S P_\theta(x)dx\big)\},\ \theta \in \Omega$$

which indicates $\{q_\theta\}$ forms an exponential family.

Problem 1.3

(a) $\forall \varepsilon > 0$, $\mathbb{P}\{|M_n - \theta| \geq \varepsilon\} = \mathbb{P}\{M_n \leq \theta - \varepsilon\} = \left(\dfrac{\theta - \varepsilon}{\theta}\right)^n \to 0$ as $n \to \infty$.

(b) For $x \in [0, \theta]$, $F(x) = \mathbb{P}(M_n \leq x) = \left(\dfrac{x}{\theta}\right)^n$.

Thus, density of M_n: $f(x) = \dfrac{n}{\theta}\left(\dfrac{x}{\theta}\right)^{n-1} \mathbf{1}(0 \leq x \leq \theta)$

$Bias = \theta - \mathbb{E}(M_n) = \theta - \displaystyle\int_0^\theta \frac{n}{\theta^n} x^n dx = \frac{1}{n+1}\theta$

$Variance = \mathbb{E}(M_n^2) - \mathbb{E}(M_n)^2 = \displaystyle\int_0^\theta \frac{n}{\theta^n} x^{n+1} dx - \left(\frac{n}{n+1}\theta\right)^2 = \left(\frac{n}{n+2} - \frac{n^2}{(n+1)^2}\right)\theta^2.$

(c) $Risk = Bias^2 + Variance = \left(\dfrac{n}{n+2} - \dfrac{n-1}{n+1}\right)\theta^2 = \dfrac{2\theta^2}{(n+1)(n+2)}$

Problem 1.4

Define

$$A(\eta) = \log \int x^2 \exp\{-\eta x\} d\mu(x) = \log\left(\sum_{i=1} i^2 \exp\{-\eta i\}\right)$$

the last equality is because μ is counting measure. By the sum of geometric series, if $\eta > 0$, $\sum_{i=1} \exp\{-\eta i\} = 1/(e^\eta - 1)$. Taking derivative of both sides yields $\sum_{i=1} i^2 \exp\{-\eta i\} = e^\eta(e^\eta + 1)/(e^\eta - 1)^3$ which is finite. And it is easy to see that when $\eta \leq 0$, $A(\eta) > \infty$. So the natural parameter space is $\{\eta, \eta > 0\}$ and $\log A(\eta) = \log\left(e^\eta(e^\eta + 1)/(e^\eta - 1)^3\right)$.

Theorem 2.4 ensures the existence of moments, and from the remark below theorem 2.4 we know that $A(\eta)' = -EX$ and accordingly $A(\eta)'' = Var(X)$. So $EX = (4e^\eta + e^{2\eta} + 1)/(e^{2\eta} - 1)$ and $Var(X) = 4e^\eta(e^\eta + e^{2\eta} + 1)/(e^{2\eta} - 1)^2$.

Problem 1.5

(a) Not identifiable. Pick $\theta_1 = (\mu_1, \mu_2, \ldots, \mu_p, \nu, \sigma)$ and $\theta_2 = (\mu_1 + \nu, \mu_2 + \nu, \ldots, \mu_p + \nu, 0, \sigma)$ and notice that they result in the same distribution;

(b) Not identifiable. Pick $\theta = (1, 1)$ and $\theta_2 = (2, 2)$ and notice they give rise to the same distribution;

(c) Identifiable.(except the case $\sigma^2 = 0$) For $\theta_1 = (\sigma_1, \alpha_1)$ we have $\text{var}(X_i) = \sigma_1^2$ and $\text{cov}(X_i, X_{i-1}) = \alpha_1\sigma_1^2$. Suppose $\mathbb{P}_{\theta_1} = \mathbb{P}_{\theta_2}$. We have $\sigma_2^2 = \text{var}_{\theta_2}(X_i) = \text{var}_{\theta_1}(X_i) = \sigma_1^2$, hence $\sigma_1 = \sigma_2$. Furthermore, $\alpha_2\sigma_2^2 = \text{cov}_{\theta_2}(X_i, X_{i-1}) = \text{var}_{\theta_1}(X_i, X_{i-1}) = \alpha_1\sigma_1^2$, yielding $\alpha_1 = \alpha_2$. Hence $\mathbb{P}_{\theta_1} = \mathbb{P}_{\theta_2} \Rightarrow \theta_1 = \theta_2$ proving identifiability.

Problem 1.6

From the distribution of $\mathbf{X}$, we have:

$$\mathbb{P}_\theta(\mathbf{X}=\mathbf{x}) = \exp\left[\left(\sum_{i=1}^n x_i\right)\log(\theta) - n\theta\right]\frac{1}{x_1!x_2!\cdots x_n!}$$

(a) We have:

$$\mathbb{P}_\theta\left(\mathbf{X}=\mathbf{x}|T(\mathbf{X})=S\right) = \frac{\mathbb{P}_\theta\left(\mathbf{X}=\mathbf{x} \text{ and } T(\mathbf{X})=S\right)}{\mathbb{P}_\theta(T\left(\mathbf{X}\right)=S)}$$

From the properties of the Poisson distribution: $T(\mathbf{X}) = \sum_{i=1}^n X_i \sim \text{Poisson}(n\theta)$:

$$\mathbb{P}_\theta(T\left(\mathbf{X}\right)=S) = \exp\left[S\log(n\theta) - n\theta\right]\frac{1}{S!}$$

Furthermore:

$$\begin{aligned}
\mathbb{P}_\theta\left(\mathbf{X}=\mathbf{x} \text{ and } T(\mathbf{X})=S\right) &= \mathbb{I}(\sum_{i=1}^n x_i = S)\exp\left[\left(\sum_{i=1}^n x_i\right)\log(\theta) - n\theta\right]\frac{1}{x_1!x_2!\cdots x_n!} \\
&= \mathbb{I}(\sum_{i=1}^n x_i = S)\exp\left[S\log(\theta) - n\theta\right]\frac{1}{x_1!x_2!\cdots x_n!}
\end{aligned}$$

which yields:

$$\begin{aligned}
\mathbb{P}_\theta\left(\mathbf{X}=\mathbf{x}|T(\mathbf{X})=S\right) &= \exp\left[-S\log(n)\right]\frac{S!}{x_1!x_2!\cdots x_n!}\cdot\mathbb{I}(\sum_{i=1}^n x_i = S) \\
&= \left(\frac{1}{n}\right)^S\frac{S!}{x_1!x_2!\cdots x_n!}\cdot\mathbb{I}(\sum_{i=1}^n x_i = S)
\end{aligned}$$

which does not involve θ proving sufficiency of T.

(b) Define:

$$\begin{aligned}
g(S,\theta) &= \exp\left[S\log(\theta) - n\theta\right] \\
h(x) &= \frac{1}{x_1!x_2!\cdots x_n!}
\end{aligned}$$

It is then possible to write $\mathbb{P}_\theta(\mathbf{X}=\mathbf{x})$ as: $\mathbb{P}_\theta(\mathbf{X}=\mathbf{x}) = g\left(T(\mathbf{X}),\theta\right)h(x)$ so sufficiency of T follows from the factorization theorem.