ST210A - HW1

Hoang Duong

September 15, 2014

Problem 1. Statistics of joint Posson distribution

Proof. Since $X_i's$ are independent, the joint distribution is:

$$\begin{aligned} p(x_1, x_2, ..., x_n) =& \prod_{i=1}^{n} \frac{\exp[x_i(\alpha + \beta t_i) - \exp(\alpha + \beta t_i)]}{x_i!} \\ =& \frac{1}{\prod_{i=1}^{n} x_i!} \exp\left[\alpha \sum_{i=1}^{n} x_i + \beta \sum_{i=1}^{n} x_i t_i - \sum_{i=1}^{n} \exp(\alpha + \beta t_i)\right] \end{aligned}$$

So the joint distributions form a two parameter exponential family α, β and the corresponding statistics are:

$$T_1 = \sum_{i=1}^{n} x_i, T_2 = \sum_{i=1}^{n} t_i x_i$$

□

Problem 2. Exponential family restricted to a region

Proof. We have the cumulative distribution of Y is:

$$\begin{aligned} \mathbb{P}_\theta[Y \le y] =& \mathbb{P}[X \le y \mid X \in S] \\ =& \frac{\mathbb{P}[X \le y \wedge X \in S]}{\mathbb{P}[X \in S]} \\ =& \frac{1}{\mathbb{P}[X \in S]} \int_{X \in S \wedge X \le y} p_\theta(X) d\mu(X) \\ =& \frac{1}{\mathbb{P}[X \in S]} \int_{X \le y} I_S(X) p_\theta(X) d\mu(X) \\ \Rightarrow q_\theta(x) =& \frac{h(x) I_S(x)}{\mathbb{P}[X \in S]} \exp\left[\sum_{i=1}^{s} \eta_i(\theta) T_i(x) - B(\theta)\right] \end{aligned}$$

For I_S indicates the indicator function for set S.

From the formula of $q_\theta(x)$ above, we can see that $\{q_\theta(x), \theta \in \Omega\}$ forms an exponential family that share everything with $\{p_\theta(x), \theta \in \Omega\}$ but the $h(x)$ function. The new $h'(x) = \frac{h(x)I_S(x)}{\mathbb{P}[X \in S]}$ satisfies the condition that it is non-negative. □

Problem 3. Consider an i.i.d sample $\{X_1, X_2, ..., X_n\} \sim U([0, \theta])$, $M_n = \max\{X_1, X_2, ..., X_n\}$

Proof. a. Prove that $M_n \xrightarrow{p} \theta$ as $n \to +\infty$.

By definition of convergence in probability, let $\epsilon > 0, \epsilon \in \mathbb{R}$ be arbitrary. If $\epsilon < \theta$, we have:

$$\begin{aligned}
\mathbb{P}[|M_n - \theta| > \epsilon] &= \mathbb{P}[\theta - M_n > \epsilon] \\
&= \mathbb{P}[M_n < \theta - \epsilon] \\
&= \mathbb{P}[\max\{X_1, X_2, ..., X_n\} < \theta - \epsilon] \\
&= \mathbb{P}[X_1 < \theta - \epsilon \wedge X_2 < \theta - \epsilon \wedge ... \wedge X_n < \theta - \epsilon] \\
&= \mathbb{P}[X_1 < \theta - \epsilon]\mathbb{P}[X_2 < \theta - \epsilon]...\mathbb{P}[X_n < \theta - \epsilon] \\
&= (\frac{\theta - \epsilon}{\theta})^n
\end{aligned}$$

Where the fifth equality holds because $X_i's$ are independent. We have:

$$\begin{aligned}
\lim_{n\to\infty} \mathbb{P}[|M_n - \theta| > \epsilon] &= \lim_{n\to\infty} (\frac{\theta - \epsilon}{\theta})^n \\
&= 0
\end{aligned}$$

If $\epsilon \geq \theta$, then:

$$\begin{aligned}
\mathbb{P}[M_n < \theta - \epsilon] &= 0 \\
\Rightarrow \lim_{n\to\infty} \mathbb{P}[M_n < \theta - \epsilon] &= 0
\end{aligned}$$

So the limit is 0 for any arbitrary $\epsilon > 0$, thus M_n converges in probability to θ.

b. Compute the Bias of M_n. We have $\forall x \in [0, \theta]$:

$$\begin{aligned}
\mathbb{P}[M_n < x] =& \mathbb{P}[X_1 < x \wedge X_2 < x \wedge ... \wedge X_n < x] \\
=& \mathbb{P}[X_1 < x]\mathbb{P}[X_2 < x]...\mathbb{P}[X_n < x] \\
=& (\frac{x}{\theta})^n
\end{aligned}$$

Thus the density for M_n is:

$$\begin{aligned}
f_\theta(M_n) =& \frac{d\mathbb{P}[M_n < x]}{dx} = \frac{n}{\theta^n} x^{n-1} \\
\Rightarrow \mathbb{E}M_n =& \int_0^\theta x f_\theta(x) dx \\
=& \int_0^\theta \frac{n}{\theta^n} x^n dx \\
=& \frac{n}{\theta^n} \frac{x^{n+1}}{n+1}\bigg|_0^\theta = \frac{n}{n+1}\theta \\
\Rightarrow \text{Bias(M}_\text{n}) =& (\mathbb{E}[M_n - \theta])^2 \\
=& \frac{1}{(n+1)^2}\theta^2
\end{aligned}$$

Compute the Variance of M_n:

$$
\begin{aligned}
\mathrm{Var}(M_n) &= \int_0^\theta (x-\mathbb{E}x)^2 f_\theta(x)dx \\
&= \int_0^\theta (x-\frac{n}{n+1}\theta)^2 \frac{n}{\theta^n}x^{n-1}dx \\
&= \int_0^\theta \{\frac{n}{\theta^n}x^{n+1} - \frac{2n^2}{(n+1)\theta^{n-1}}x^n + \frac{n^3}{(n+1)^2\theta^{n-2}}x^{n-1}\}dx \\
&= \theta^2(\frac{n}{n+2} - \frac{2n^2}{(n+1)^2} + \frac{n^2}{(n+1)^2}) \\
&= \theta^2\frac{n}{(n+1)^2(n+2)}
\end{aligned}
$$

c. The risk of M_n under quadratic loss is:

$$
\begin{aligned}
R(\theta, M_n) &= \mathrm{Bias}(M_n)^2 + \mathrm{Var}(M_n) \\
&= \theta^2(\frac{2n+2}{(n+1)^2(n+2)}) \\
&= \theta^2\frac{2}{(n+1)(n+2)}
\end{aligned}
$$

□

Problem 4. Find the natural parameter space

Proof. We need to find η such as $A(\eta) < \infty$.

$$
\begin{aligned}
A(\eta) &= \log\int \exp(-\eta x)x^2 d\mu(x) \\
&= \log\sum_{x=1}^{\infty}\exp(-\eta x)x^2
\end{aligned}
$$

Now consider the series $\sum_{x=1}^{\infty} e^{-\eta x}$, if $\eta \leq 0$ then it is obvious that the series diverges as it is at least as large as $\sum_{x=1}^{\infty} 1$. If $\eta > 0$, then the series converges according to the ratio test. So the parameter space is $\mathbb{R}^+$. For $\eta > 0$, let:

$$
\begin{aligned}
C =& \sum_{x=1}^{\infty} e^{-\eta x} \\
=& e^{-\eta}(1 + \sum_{x=1}^{\infty} e^{-\eta x}) \\
=& e^{-\eta}(1 + C) \\
\Rightarrow C =& \frac{e^{-\eta}}{1 - e^{-\eta}} = \frac{e^{\eta}}{e^{\eta} - 1} \\
\frac{dC}{d\eta} =& \sum_{x=1}^{\infty} -xe^{-\eta x} = \left(\frac{e^{\eta}}{e^{\eta} - 1}\right)' = \frac{-e^{\eta}}{(e^{\eta} - 1)^2} \\
\Rightarrow \frac{d^2C}{d\eta^2} =& \sum_{x=1}^{\infty} x^2 e^{-\eta x} = \left(\frac{-e^{\eta}}{(e^{\eta} - 1)^2}\right)' = \frac{e^{2\eta} + e^{\eta}}{(e^{\eta} - 1)^3} \\
\Rightarrow A(\eta) =& \log \frac{e^{2\eta} + e^{\eta}}{(e^{\eta} - 1)^3} \\
\Rightarrow A'(\eta) =& \frac{e^{2\eta} + 4e^{\eta} + 1}{1 - e^{2\eta}} \\
\Rightarrow A''(\eta) =& \frac{4e^{\eta}(e^{2\eta} + e^{\eta} + 1)}{(e^{2\eta} - 1)^2}
\end{aligned}
$$

So we have:

$$
\begin{aligned}
\mathbb{E}X =& -\mathbb{E}[T(X)] = -A'(\eta) = \frac{e^{2\eta} + 4e^{\eta} + 1}{e^{2\eta} - 1} \\
\mathrm{Var}[X] =& \mathrm{Var}[T(X)] = A''(\eta) = \frac{4e^{\eta}(e^{2\eta} + e^{\eta} + 1)}{(e^{2\eta} - 1)^2}
\end{aligned}
$$

□

Problem 5. Identifiable parameterization

Proof. a. This parameterization is not identificable since for

$$
\begin{aligned}
\theta_1 =& (\alpha_1, \alpha_2, ..., \alpha_p, \nu, \sigma^2) \\
\theta_2 =& (\alpha_1 + 1, \alpha_2 + 1, ..., \alpha_p + 1, \nu - 1, \sigma^2)
\end{aligned}
$$

Then $\theta_1 \neq \theta_2$, but each of X_i have the same distribution for these θ_1, θ_2 and they are independent thus they the two joint distribution are the same.

b. First we need the probability mass to sum up to 1. For that we need to divided by the constant $e^{\theta_0} + e^{\theta_1}$. And so we have the probability mass function:

$$
\begin{aligned}
p(x; \theta) =& \frac{\exp(\theta_0 x + \theta_1(1 - x))}{e^{\theta_0} + e^{\theta_1}} \\
\Rightarrow p(0; \theta) =& \frac{e^{\theta_1}}{e^{\theta_0} + e^{\theta_1}} = \frac{1}{1 + e^{\theta_0 - \theta_1}} \\
p(1; \theta) =& \frac{e^{\theta_0}}{e^{\theta_0} + e^{\theta_1}} = \frac{1}{1 + e^{\theta 1 - \theta_0}}
\end{aligned}
$$

This parameterization is not identificable since for any $\theta = (\theta_0, \theta_1), \theta' = (\theta_0', \theta_1')$ such that $\theta_0 - \theta_1 = \theta_0' - \theta_1', \theta_0 \neq \theta_0'$, we have $\theta \neq \theta'$ but the $p(x; \theta) = p'(x, \theta)$.

c. For $\theta \neq \theta'$, assume the contradiction that $\mathbb{P}_\theta = \mathbb{P}_{\theta'}$. Thus the marginal distribution of X_1 is the same as the marginal distribution for X_1'. Thus $\theta = \theta'$ contradiction. So $\mathbb{P}_\theta \neq \mathbb{P}_{\theta'}$. So this parameterization is identificable. ☐

Lemma 1. *Let $X, Y \sim Poisson(\lambda_1), Poisson(\lambda_2)$, then $X + Y \sim Poisson(\lambda_1 + \lambda_2)$.*

Proof. We have:

$$\begin{aligned}
\mathbb{P}[X+Y=n] &= \sum_{m=0}^{n} \mathbb{P}[X=m]\mathbb{P}[Y=n-m] \\
&= \sum_{m=0}^{n} e^{-\lambda_1}\frac{\lambda_1^m}{m!}e^{-\lambda_2}\frac{\lambda_2^{n-m}}{(n-m)!} \\
&= e^{-\lambda_1-\lambda_2}\sum_{m=0}^{n}\frac{1}{m!(n-m)!}\lambda_1^m\lambda_2^m \\
&= e^{-\lambda_1-\lambda_2}(\lambda_1+\lambda_2)^n\frac{1}{n!}
\end{aligned}$$

which is the probability mass function of $Poisson(\lambda_1 + \lambda_2)$.

Recursively, we have if $X_1, X_2, ..., X_n$ are independent Poisson with parameters $\lambda_1, \lambda_2, ..., \lambda_n$ then $X_1 + X_2 + ... + X_n \sim Poisson(\lambda_1 + \lambda_2 + ... + \lambda_n)$. ☐

Problem 6. Sufficient Statistics

Proof. a. From Lemma 1., we have $T \sim Poisson(n\theta)$. Now $\forall x_1, x_2, ..., x_n, t \in\in \mathbb{N}$, if $t = \sum_{i=1}^n x_i$, then:

$$\begin{aligned}
\mathbb{P}[X_1=x_1, X_2=x_2, ..., X_n=x_n \mid T=t] &= \frac{\mathbb{P}[(X_1=x_1, X_2=x_2, ..., X_n=x_n) \wedge T=t]}{\mathbb{P}[T=t]} \\
&= \frac{\mathbb{P}[X_1=x_1, X_2=x_2, ..., X_n=x_n]}{\mathbb{P}[T=t]} \\
&= \frac{\mathbb{P}[X_1=x_1]\mathbb{P}[X_2=x_2]...\mathbb{P}[X_n=x_n]}{\mathbb{P}[T=t]} \\
&= \frac{\prod_{i=1}^n \frac{1}{x_i!}\theta^{x_i}e^{-\theta}}{\frac{1}{t!}(n\theta)^t e^{-n\theta}} \\
&= \frac{(\sum x_i)!}{\prod_{i=1}^n x_i!}\frac{1}{n^t}
\end{aligned}$$

If $t \neq \sum_{i=1}^n x_i$, then $\mathbb{P}[X_1=x_1, X_2=x_2, ..., X_n=x_n \mid T=t] = 0$. So in all cases, the conditional joint distribution $\mathbb{P}[X_1, X_2, ..., X_n \mid T]$ does not depend on θ. So T is a sufficient statistics.

b. We have:

$$\begin{aligned}
\mathbb{P}[X_1=x_1, X_2=x_2, ..., X_n=x_n] &= \mathbb{P}[X_1=x_1]\mathbb{P}[X_2=x_2]...\mathbb{P}[X_n=x_n] \\
&= \prod_{i=1}^n \frac{1}{x_i!}\theta^{x_i}e^{-\theta} \\
&= \theta^{\sum_{i=1}^n x_i}e^{-n\theta}\prod_{i=1}^n\frac{1}{x_i!}
\end{aligned}$$

Let $g_\theta(T(x)) = \theta^{T(x)}e^{-n\theta}, h(x) = \prod_{i=1}^n \frac{1}{x_i!}$, then we $g_\theta \geq 0, \forall\theta > 0, h(x) \geq 0$, thus $p_\theta(x)$ satisfies the factorization condition. So TT is a sufficient statistics. ☐