ST210A - Homework 4

Hoang Duong

September 30, 2014

Problem 1. Geometric Distribution

Proof. Let $g(X)$ be an unbiased estimator of θ. We need:

$$
\begin{aligned}
\forall \theta \in 0,1(), \mathbb{E}g(X) =& \theta \\
\Leftrightarrow \sum_{k=0}^{\infty} g(k)\theta(1-\theta)^k =& \theta \\
\Leftrightarrow \sum_{k=0}^{\infty} g(k)(1-\theta)^k =& 1, \forall \theta \in (0,1) \\
\Leftrightarrow \sum_{k=0}^{\infty} g(k)\lambda^k =& 1, \forall \lambda = 1-\theta \in (0,1)
\end{aligned}
$$

Thus $g(0) = 1, g(k) = 0, \forall k \geq 1$. So $g(X) = \mathbb{I}[X = 0]$.

$$
\begin{aligned}
\text{Var}(g(X)) =& \mathbb{E}g^2(X) - (\mathbb{E}g(X))^2 \\
=& \theta - \theta^2
\end{aligned}
$$

The Fisher information for θ is:

$$
\begin{aligned}
I(\theta) =& \mathbb{E}\left[-\frac{\partial^2 \log\left(\theta(1-\theta)^x\right)}{\partial\theta^2}\right] \\
=& \mathbb{E}\left[-\frac{\partial^2 (\log\theta + x\log(1-\theta))}{\partial\theta^2}\right] \\
=& \mathbb{E}\left[-\frac{\partial}{\partial\theta}\left(\frac{1}{\theta} - \frac{x}{1-\theta}\right)\right] \\
=& \mathbb{E}\left[\frac{1}{\theta^2} + \frac{x}{(1-\theta)^2}\right] \\
=& \frac{1}{\theta^2} + \frac{1}{(1-\theta)^2}\left(\frac{1}{\theta} - 1\right) \\
=& \frac{1}{\theta^2} + \frac{1}{\theta(1-\theta)} = \frac{1}{\theta^2(1-\theta)}
\end{aligned}
$$

Thus the error bound for an unbiased estimator according to Theorem 4.9 in Keener is: $\theta^2(1-\theta)$. This error bound is strictly smaller than the variance of $g(X)$ for $0 < \theta < 1$. ($g(X)$ is the only unbiased estimator). $\square$

Problem 2. Poisson with Gamma Prior

Proof. (a) We have the posterior density of θ is:

$$\begin{aligned}
p(\theta \mid x) =& \frac{p(x \mid \theta)p(\theta)}{p(x)} \\
=& \frac{1}{p(x)} \frac{\theta^x e^{-\theta}}{x!} \frac{1}{b^a \Gamma(a)} \theta^{a-1} e^{-\theta/b}
\end{aligned}$$

Thus we have for any estimator $d(X)$ of Θ:

$$\begin{aligned}
\mathbb{E}\left[L(\Theta, d(X)) \mid X = x\right] =& \int_0^\infty \frac{(\theta - d(X))^2}{\theta} p(\theta \mid x) d\theta \\
=& \int_0^\infty \frac{(\theta - d(X))^2}{\theta} \frac{1}{p(x)} \frac{\theta^x e^{-\theta}}{x!} \frac{1}{b^a \Gamma(a)} \theta^{a-1} e^{-\theta/b} d\theta \\
=& \frac{1}{p(x) b^a \Gamma(a)} \int_0^\infty (\theta - d(X))^2 \theta^{x+a-2} e^{-\theta - \theta/b} \frac{\left(\frac{b}{b+1}\right)^{x+a-1} \Gamma(x+a-1)}{\left(\frac{b}{b+1}\right)^{x+a-1} \Gamma(x+a-1)} d\theta \\
=& h(x, b, a) \int_0^\infty (\theta - d(X))^2 \frac{1}{(\frac{b}{b+1})^{x+a-1} \Gamma(x+a-1)} \theta^{x+a-2} e^{-\frac{-\theta}{(b/(b+1))}} d\theta
\end{aligned}$$

We can see that the above integral is $\mathbb{E}(\lambda - d(X))^2$ for λ follows $Gamma(x+a-1, \frac{b}{b+1})$. This expectation can be rewritten as: $\mathbb{E}(\lambda - \mu(\lambda) + \mu(\lambda) - d(X)) = \mathrm{Var}(\lambda) + \mathbb{E}(d(X) - \mu(\lambda))^2$, and is minimized when $d(X) = \mu(\lambda) = (x + a - 1)\frac{b}{b+1}$.

According to Theorem 7.1 from Keener, $d(X) = (x + a - 1)\frac{b}{b+1}$ is the Bayes estimator w.r.t the family of Gamma(a, b) prior.

(b) The estimator $\delta(X) = X$ can be obtained as $a = 1, b \to \infty$. This is exponential distribution with mean parameter λ approaching zero. $\square$

Problem 3. Uniform with log-normal prior

Proof. (a) First the conditional density of $X = (X_1, X_2, ..., X_n)^T$ is:

$$\begin{aligned}
p(X_i = x_i, \forall 1 \le i \le n \mid \Theta = \theta) =& \frac{1}{\theta^n} \mathbb{I}[0 < x_i < \theta, \forall 1 \le i \le n] \\
=& \frac{1}{\theta^n} \mathbb{I}[\min x_i > 0] \mathbb{I}[\max x_i < \theta]
\end{aligned}$$

Since $\log \theta$ is a bijective function on $\mathbb{R}^+$, conditioning on θ is the same as conditioning on $\log \theta$. Thus:

$$\begin{aligned}
p(\theta \mid X) =& \frac{1}{p(X)} p(X \mid \theta) p(\theta) \\
=& \frac{1}{p(X)} \frac{\mathbb{I}[\min x_i > 0] \mathbb{I}[\max x_i < \theta]}{\theta^n} \frac{1}{\sqrt{2\pi} \theta \sigma_0} \exp\left(-\frac{1}{2\sigma_0^2} (\ln \theta - \mu_0)^2\right)
\end{aligned}$$

Applying the change of variable density for $\lambda = g(\theta) = \log \theta$, we have:

$$\begin{aligned}
p_\lambda(\lambda \mid X) =& \frac{1}{p(X)} \frac{\mathbb{I}[\min x_i > 0] \mathbb{I}[\max x_i < \exp \lambda]}{\exp(n\lambda)} \frac{1}{\sqrt{2\pi} \exp(\lambda) \sigma_0} \exp\left(-\frac{1}{2\sigma_0^2} (\lambda - \mu_0)^2\right) \\
=& \frac{\mathbb{I}[\min x_i > 0] \mathbb{I}[\max x_i < \exp \lambda]}{\sqrt{2\pi} \sigma_0 p(X) \exp(n\lambda + \lambda)} \exp\left(-\frac{1}{2\sigma_0^2} (\lambda - \mu_0)^2\right) \\
=& \frac{\mathbb{I}[\min x_i > 0] \mathbb{I}[\max x_i < \exp \lambda]}{\sqrt{2\pi} \sigma_0 p(X)} \exp\left(-\frac{1}{2\sigma_0^2} (\lambda - \mu_0)^2 + (n+1)\lambda\right)
\end{aligned}$$

(b) Under the specified loss function, we need to find $\delta(X)$ that minimizes:

$$\begin{aligned}\mathbb{E}\left[L(\Theta,\delta(X)\mid X=x\right] =& \mathbb{P}\left[\Theta\neq\delta(X)\mid X=x\right]\\ =&1-\mathbb{P}\left[\Theta=\delta(X)\mid X=x\right]\end{aligned}$$

We can minimize the expectation by maximize the a posterior probability $\mathbb{P}\left[\Theta=\delta(X)\mid X=x\right] = \mathbb{P}\left[\log\Theta=\log(\delta(X))\right]$, but this doesn't make sense because this probability is always zero. Instead we will attempt to maximize the density function $p_{\log\Theta}$. From part (a), we have the density is maximized iff $\frac{1}{2\sigma_0^2}\left(\lambda-\mu_0\right)^2+(n+1)\lambda$ is minimized and $\lambda\geq\log\max x_i$. The quadratic expression is minimized at either $\frac{1}{\sigma_0^2}(\lambda-\mu_0)+n+1=0$, the point that make derivative equal to zero, or at the critical point $\log\max x_i$. The linear equation has solution $\lambda=\mu_0-(n+1)\sigma_0^2$. Thus the minimum is attained at $\lambda=\max(\log\max x_i,\mu_0-(n+1)\sigma_0^2)$ by the nature of quadratic function left-half truncated. This is equivalent to $\theta=\max\left(\max x_i,\exp\left(\mu_0-(n+1)\sigma_0^2\right)\right)$.

In conclusion the Beyes estimator is $\delta(X)=\max\left(\max x_i,\exp\left(\mu_0-(n+1)\sigma_0^2\right)\right)$. □

Problem 4. Beyes Estimator vs. Unbiased Estimator

Proof. Assuming that $\exists\delta(X)$ that is both unbiased and Bayes estimator. As a Beyes estimator, it minimizes:

$$\mathbb{E}\left[L(g(\Theta),\delta(X))\mid X=x\right] = \mathbb{E}\left[(g(\Theta)-\delta(X))^2\mid X=x\right],\forall x$$

As shown before any random variable X, $\mathbb{E}(X-a)^2$ is minimized w.r.t a when $a=\mathbb{E}X$. Thus the above expectation is minimized when $\delta(x)=\mathbb{E}\left[g(\Theta)\mid X=x\right]$ (a.e). But as an unbiased estimator of $g(\Theta)$, $\mathbb{E}\left[\delta(X)\mid\Theta=\theta\right]=g(\theta)$.

Thus the Bayes risk:

$$\begin{aligned}\mathbb{E}\left[g(\Theta)\delta(X)\right] =& \mathbb{E}\left[\mathbb{E}\left[g(\Theta)\delta(X)\mid X\right]\right]\\ =&\mathbb{E}\left[\delta(X)\mathbb{E}\left[g(\Theta)\mid X\right]\right]\\ =&\mathbb{E}\left[\delta^2(X)\right]\\ \mathbb{E}\left[g(\Theta)\delta(X)\right] =& \mathbb{E}\left[\mathbb{E}\left[g(\Theta)\delta(X)\mid\Theta\right]\right]\\ =&\mathbb{E}\left[g(\Theta)\mathbb{E}\left[\delta(X)\mid X\right]\right]\\ =&\mathbb{E}\left[g^2(\Theta)\right]\\ \Rightarrow\mathbb{E}\left[\left(g(\theta)-\delta(X)\right)^2\right] =& \mathbb{E}\left[g^2(\Theta)\right]+\mathbb{E}\left[\delta^2(X)\right]-2\mathbb{E}\left[g(\Theta)\delta(X)\right]\\ =&0\end{aligned}$$

Thus under quadratic loss, unbiased estimator agree with Bayes estimator implies Bayes risk is zero. On the other hand, if Bayes Risk is zero, then $g(\theta)=\delta(X)$ almost everywhere. Thus $\delta(X)$ is an unbiased estimator and also a Bayes estimator. So it is a two way relation. □

Problem 5. Moment of Normal Distribution

Proof. (a) We have:

$$
\begin{aligned}
\mathbb{E}g'(X) =& -\int \exp\left(\sum_{i=1}^{d}\theta_i T_i(x) - A(\theta)\right) h(x) g'(x) dx \\
=& -\int \exp\left(\sum_{i=1}^{d}\theta_i T_i(x) - A(\theta)\right) h(x) dg(x) \\
=& -\exp\left(\sum_{i=1}^{d}\theta_i T_i(x) - A(\theta)\right) h(x) g(x)\Big|_{-\infty}^{\infty} + \int g(x) d\exp\left(\sum_{i=1}^{d}\theta_i T_i(x) - A(\theta)\right) h(x) \\
=& 0 + \int g(x)\Bigg\{ h'(x)\exp\left(\sum_{i=1}^{d}\theta_i T_i(x) - A(\theta)\right) + \\
& + h(x)\exp\left(\sum_{i=1}^{d}\theta_i T_i(x) - A(\theta)\right)\left(\sum_{i=1}^{d}\theta_i T_i'(X)\right)\Bigg\} \\
=& \int g(x)\left[\frac{h'(x)}{h(x)} + \sum_{i=1}^{d}\theta_i T_i^{'}(x)\right]\exp\left(\sum_{i=1}^{d}\theta_i T_i(x) - A(\theta)\right) h(x) dx \\
=& \mathbb{E}\left\{\left[\frac{h'(X)}{h(X)} + \sum_{i=1}^{d}\theta_i T_i'(X)\right] g(X)\right\}
\end{aligned}
$$

(b) For normal, the density is:

$$
\begin{aligned}
f(x) =& \frac{1}{\sqrt{2\pi}\sigma}\exp\left\{-\frac{(x-\mu)^2}{2\sigma^2}\right\} \\
=& \frac{1}{\sqrt{2\pi}}\exp\left\{-\frac{1}{2\sigma^2}x^2 + \frac{\mu}{\sigma^2}x - \frac{\mu}{2\sigma^2} - \log\sigma\right\}
\end{aligned}
$$

So from (a), we have:

$$
\begin{aligned}
\mathbb{E}\left[g'(X)\right] =& -\mathbb{E}\left\{\left[\frac{h'(X)}{h(X)} + \sum_{i=1}^{d}\theta_i T_i'(X)\right] g(X)\right\} \\
=& -\mathbb{E}\left[\left(-\frac{1}{2\sigma^2}2X + \frac{\mu}{\sigma^2}\right) g(X)\right] \\
=& \frac{1}{\sigma^2}\mathbb{E}\left[(X-\mu)\, g(X)\right] \\
=& \frac{1}{\sigma^2}\left(\mathbb{E}\left[Xg(X)\right] - \mathbb{E}\left[X\right]\mathbb{E}\left[g(X)\right]\right) \\
=& \frac{1}{\sigma^2}\mathrm{Cov}(X, g(X))
\end{aligned}
$$

(c) We have:

$$
\begin{aligned}
\operatorname{Cov}(X^2, X) =& \sigma^2 \mathbb{E}\left[2X\right] \\
\Leftrightarrow \mathbb{E}X^3 - \mathbb{E}X^2\mathbb{E}X =& \sigma^2 2\mu \\
\Leftrightarrow \mathbb{E}X^3 =& 2\mu\sigma^2 + (\sigma^2 + \mu^2)\mu \\
=& 3\mu\sigma^2 + \mu^3 \\
\operatorname{Cov}(X^3, X) =& \sigma^2 \mathbb{E}\left[3X^2\right] \\
\Leftrightarrow \mathbb{E}X^4 - \mathbb{E}X^3\mathbb{E}X =& \sigma^2 3(\sigma^2 + \mu^2) \\
\Leftrightarrow \mathbb{E}X^4 =& (3\mu\sigma^2 + \mu^3)\mu + 3\sigma^4 + 3\sigma^2\mu^2 \\
=& 6\mu^2\sigma^2 + \mu^4 + 3\sigma^4
\end{aligned}
$$

$\square$