Final

MATH 281B

March 25, 2015

Problem 1.

a. It is straight forward to see that

$$\pi(\lambda|X) \propto f(x|\theta)\pi(\theta) \propto \lambda e^{-(X+\mu)\lambda}$$

This is a Gamma distribution with parameter $(2, (X + \mu)^{-1})$. You may have another parametrization of Gamma, which gives you a different second parameter.

Pay attention that this is **not an exponential distribution**. Therefore this is not an example of conjugate prior.

b. The reason to use the posterior mean as the Bayes estimator is that posterior mean minimizes the Bayes risk under **squared loss**. To see this,

$$R_B = \mathbb{E}_\Lambda \mathbb{E}_{X|\lambda}(\delta - \lambda)^2 = \mathbb{E}_X \mathbb{E}_{\lambda|X}(\delta - \lambda)^2,$$

where the second equation is realized by Fubini's theorem.

A simple expansion can show that $\mathbb{E}_{\lambda|X}(\delta - \lambda)^2$ is minimized when $\delta = \mathbb{E}_{\lambda|X} \lambda = \mathbb{E}(\lambda|X)$.

Problem 2.

a. One example comes from question 1, where we have posterior distribution $\text{Gamma}(2, (X+\mu)^{-1})$. The prior has variance μ^{-2}, and the posterior variance is $2(X + \mu)^{-2}$. Therefore as long as

$$\mu^{-2} < 2(X + \mu)^{-2}$$

or

$$X < (\sqrt{2} - 1)\mu,$$

we have a larger posterior variance.

Another example comes from Poisson distribution with conjugate Gamma prior. Suppose Gamma prior is equipped with parameter (k, θ) and we have one observation. It is known that the posterior is $\text{Gamma}(k + X, \theta/(\theta + 1))$. If we have posterior variance larger than the prior variance, we need

$$k\theta^2 < (k + X)(\theta)^2(\theta + 1)^{-2}$$

or

$$X > k\theta(\theta + 2).$$

Other examples may be possible. However, you need X to satisfy certain criteria to make the posterior variance large. Any example leading to unconditional larger posterior variance is wrong.

b. By the formula

$$\mathrm{Var}\,(\lambda) = \mathbb{E}(\mathrm{Var}\,(\lambda|X)) + \mathrm{Var}\,(\mathbb{E}(\lambda|X)) \geq \mathbb{E}(\mathrm{Var}\,(\lambda|X)),$$

we can tell that larger posterior variance can not happen with probability 1, or the inequality will be reversed.

Problem 3.

a. Delta method is as follows. Suppose that some X_n satisfies

$$\sqrt{n}(X_n - \theta) \xrightarrow{L} N(0, \sigma^2).$$

Then for some function ϕ which is differentiable at θ, with nonzero first derivative, we have

$$\sqrt{n}(\phi(X_n) - \phi(\theta)) \xrightarrow{L} N(0, \phi'(\theta)^2\sigma^2).$$

b. Using central limit theorem on the MLE estimator, we can obtain

$$\sqrt{n}(\hat{p} - p) \xrightarrow{L} N(0, p(1-p)).$$

The plug-in idea shows that the MLE for λ is $\hat{\lambda} = \log(\hat{p}/(1-\hat{p}))$. Consider the function $\phi(x) = \log(x/(1-x))$, we have $\phi'(x) = 1/x(1-x)$. Therefore

$$\sqrt{n}(\phi(\hat{p}) - \phi(p)) \xrightarrow{L} N(0, p(1-p)\phi'(p)^2).$$

or

$$\sqrt{n}(\hat{\lambda} - \lambda) \xrightarrow{L} N(0, \frac{1}{p(1-p)}).$$

Problem 4.

a. The proof of Neyman-Pearson Lemma can be found at Wikipedia.
`http://en.wikipedia.org/wiki/Neyman%E2%80%93Pearson_lemma`
Another version can be found here. Credit goes to Professor J. Hart of TAMU.
`http://www.stat.tamu.edu/~hart/611/neyman.proof.pdf`

b. Consider a set of test function $\{\phi_\theta(X)\}$, corresponding to the null $X \sim P_\theta$ under the same level α, then the set $\theta : \phi_\theta \neq 1$ is a $1-\alpha$ confidence set for θ.

Reversely, the null $X \sim P_\theta$ can not be rejected if a confidence set of level $1-\alpha$ covers θ, under level α.

c. By the Neyman-Pearson lemma, the test function takes the form

$$\phi(X) = \mathbb{1}\{\frac{f_1(X)}{f_0(X)} > k\}$$

for some certain k, where f_0 and f_1 are density functions under null and alternative. Under the setup of this question, the test function is

$$\phi(X) = \mathbb{1}\{\frac{e^X}{\mathbb{1}_{[0,1]}(X)} > k\}.$$

Clearly we reject the null when $X > 1$. For $X < 1$, since $\mathbb{E}(\phi(X)) = \alpha$ and ϕ is monotone decreasing when X increases in $[0, 1]$, we know that ϕ is equivalent to $\mathbb{1}\{X < k'\}$, where $k' = \alpha$.

To sum up, we reject H_0 when $X < \alpha$ or $X > 1$.

For the second part, by using the alternative form of Neyman-Pearson:

$$\phi(X) = \mathbb{1}\{\frac{\sup_{\theta \in H_0} f_\theta(X)}{\sup_{\theta \in H_0 \cap H_1} f_\theta(X)} < k\}.$$

In this certain situation, we see that

$$\sup_{\theta \in H_0} f_\theta(X) = 1/X$$

and

$$\sup_{\theta \in H_0 \cap H_1} f_\theta(X) = \max(1/X, 1/eX) = 1/X.$$

(Taking derivative to the parameter and solving for the maximum will give you the answer.)

Therefore

$$\frac{\sup_{\theta \in H_0} f_\theta(X)}{\sup_{\theta \in H_0 \cap H_1} f_\theta(X)} = 1,$$

which means that the only way out is $\phi = \alpha$, which is a randomized guessing. This example tells us that only one observation is not enough to distinguish two families of distributions, even if the families are totally different.