2014 Qualifying Exam Section 1

February 21, 2019

1 Question 1

We consider two groups of independent observations: $X_1, \dots, X_n$ are i.i.d from Unif$(0, \alpha)$ and $Y_1, \dots, Y_n$ are i.i.d from Unif$(0, \beta)$, where both α and β are unknown parameters assumed to be positive. For comparison, we are interested in inference on $\theta = \beta/\alpha$

1.a

Derive the UMVUEs for α and β and calculate their respective variances. Note that

$$f(x, y) = \alpha^{-n} \mathbf{1}_{\{x_{(n)} \leq \alpha\}} \mathbf{1}_{\{x_{(1)} \geq 0\}} \beta^{-n} \mathbf{1}_{\{y_{(n)} \leq \alpha\}} \mathbf{1}_{\{y_{(1)} \geq 0\}}$$

By the factorization theorem, we see that $(x_{(n)}, y_{(n)})$ is sufficient for (α, β).

Note that for $x \in (0, \alpha), y \in (0, \beta)$

$$\begin{aligned}
F(x, y) &= P(X_{(n)} \leq x, Y_{(n)} \leq y) \\
&= P(X_{(n)} \leq x) P(Y_{(n)} \leq y) \\
&= P(X_1 \leq x)^n P(Y_1 \leq y)^n \\
&= \alpha^{-n} x^n \beta^{-n} y^n
\end{aligned}$$

So that

$$f(x, y) = n^2 \alpha^{-n} \beta^{-n} x^{n-1} y^{n-1}$$

Suppose there is a measurable function g such that $\mathbb{E}g(X_{(n)}, Y_{(n)}) = 0$. Then we have

$$\begin{aligned}
0 &= \mathbb{E}g(X_{(n)}, Y_{(n)}) \\
\implies 0 &= \int_0^\alpha \int_0^\beta g(x,y) n^2 \alpha^{-n} \beta^{-n} x^{n-1} y^{n-1} dy\ dx \\
\implies 0 &= \int_0^\alpha \int_0^\beta g(x,y) x^{n-1} y^{n-1} dy\ dx \\
\implies \frac{\partial^2}{\partial\alpha\partial\beta} 0 &= \frac{\partial^2}{\partial\alpha\partial\beta} \int_0^\alpha \int_0^\beta g(x,y) x^{n-1} y^{n-1} dy\ dx
\end{aligned}$$

$$\implies g(\alpha, \beta)\alpha^{n-1}\beta^{n-1} = 0 \text{ a.e.}$$

$$\implies g(\alpha, \beta) = 0 \text{ a. e.}$$

Hence, $P(g(X_{(n)}, Y_{(n)}) = 0) = 1$, so $(X_{(n)}, Y_{(n)})$ are complete for (α, β), and they are also sufficient for (α, β) by the above.

We need to find a function h such that $\mathbb{E}h(X_{(n)}) = \alpha$. We have

$$\begin{aligned}
\alpha &= \mathbb{E}h(X_{(n)}) \\
&= \int_0^\alpha h(x)\alpha^{-n} n x^{n-1} dx \\
\iff \frac{\alpha^{n+1}}{n} &= \int_0^\alpha h(x) x^{n-1} dx
\end{aligned}$$

We see that in order to get the integral to come out with α^{n+1}, $h(x)$ must be some multiple of x. Since the integral if $h(x) = x$ gives us $\frac{1}{n+1}\alpha^{n+1}$, we see that $h(x) = \frac{n+1}{n}x$. Thus, by the Lehmann-Scheffe theorem, the UMVUE is $h(X_{(n)}) = \frac{n+1}{n}X_{(n)}$ since it is unbiased for α and a function of the complete sufficient statistic. Similarly, we have that the UMVUE for β is $\frac{n+1}{n}Y_{(n)}$

$$\begin{aligned}
\mathbb{E}\left(\frac{n+1}{n}X_{(n)}\right)^2 &= \frac{(n+1)^2}{n^2} \int_0^\alpha x^2 n \alpha^{-n} x^{n-1} dx \\
&= \frac{(n+1)^2}{n\alpha^n} \int_0^\alpha x^{n+1} dx \\
&= \frac{(n+1)^2}{n(n+2)\alpha^n} \alpha^{n+2} \\
&= \frac{(n+1)^2}{n(n+2)} \alpha^2
\end{aligned}$$

Hence,

$$\begin{aligned}
\operatorname{Var}\left(\frac{n+1}{n}X_{(n)}\right) &= \frac{(n+1)^2}{n(n+2)}\alpha^2 - \alpha^2 \\
&= \alpha^2\frac{n^2+2n+1-n^2-2n}{n(n+2)} \\
&= \frac{\alpha^2}{n(n+2)}
\end{aligned}$$

1.b

Calculate the MLEs for α and β, denoted as $\hat{\alpha}$ and $\hat{\beta}$ respectively. Derive the asymptotic distributions for $\hat{\alpha}$ and $\hat{\beta}$ after some normalization.

Solution:
Note that $f(x, y)$ is 0 if $y_{(n)} > \beta$ or if $x_{(n)} > \alpha$. If $y_{(n)} \leq \beta$ and $x_{(n)} \leq \alpha$, then f is strictly decreasing in α and in β. It follows that the MLEs are $\hat{\alpha} = X_{(n)}$ and $\hat{\beta} = Y_{(n)}$.

To find the asymptotic distribution, we want to find a sequence of constants k_n and c_n such that the cdf of $k_n(\hat{\alpha} - c_n)$ converges to a non-degenerate function. We have

$$\begin{aligned} P(k_n(X_{(n)} - c_n) \leq x) &= P(X_{(n)} \leq c_n + x/k_n) \\ &= \left(\frac{c_n + x/k_n}{\alpha}\right)^n \end{aligned}$$

If we take $c_n = \alpha$ and $k_n = -n$, then the probability above converges to $e^{-x/\alpha}$ as $n \to \infty$. Note that

$$\begin{aligned} P(-n(X_{(n)} - \alpha) \leq x) &= P(n(\alpha - X_{(n)}) \leq x) \\ &= P(X_{(n)} \geq \alpha - x/n) \\ &\to 1 - e^{-x/\alpha} \end{aligned}$$

And hence $n(\alpha - X_{(n)})$ converges in distribution to a $\text{Exp}(\alpha)$ random variable as $n \to \infty$

Similarly, we have $n(\beta - Y_{(n)})$ converges in distribution to an $\text{Exp}(\beta)$ random variable.

1.c

The MLE for θ is then $\hat{\theta} = \hat{\beta}/\hat{\alpha}$. Derive the asymptotic distribution of $\hat{\theta}$ after normalization. Construct an asymptotic 95% confidence interval for based on the observations.

<u>*Solution:*</u>
We need to find a sequence of constants k_n and c_n such that the cdf of $k_n(c_n - \hat{\theta})$ converges to some non-degenerate function.

$$\begin{aligned}
P(k_n(c_n - \hat{\theta}) \leq z) &= P(X_{(n)}/Y_{(n)} \geq c_n - z/k_n) \\
&= 1 - P(X_{(n)} \leq Y_{(n)}(c_n - z/k_n) \\
&= 1 - \int_0^\beta \left(\frac{y(c_n - z/k_n)}{\alpha}\right)^n n\beta^{-n}y^{n-1}dy \\
&= 1 - n\alpha^{-n}\beta^{-n}(c_n - z/k_n)^n \int_0 \beta y^{2n-1}dy \\
&= 1 - n\alpha^{-n}\beta^{-n}(c_n - z/k_n)^n \frac{1}{2n}\beta^{2n} \\
&= 1 - \frac{1}{2}\theta^{-n}(c_n - z/k_n)^n \\
&= 1 - \frac{1}{2}\left[\frac{1}{\theta}\left(c_n - \frac{z}{k_n}\right)\right]^n
\end{aligned}$$

We want the probability to converge to the CDF of some random variable. To get rid of the $1/2$, we need to be able to factor out a $2^{1/n}$ from what is inside the square brackets. Clearly, c_n should be proportional to θ in order to make the limit converge to some exponential function. Thus, if $c_n = 2^{1/n}\theta$ and $k_n = 2^{-1/n}n$, we have

$$\begin{aligned}
1 - \frac{1}{2}\left[\frac{1}{\theta}\left(c_n - \frac{z}{k_n}\right)\right]^n &= 1 - \frac{1}{2}\left[\frac{1}{\theta}\left(2^{1/n}\theta - \frac{z2^{1/n}}{n}\right)\right]^n \\
&= 1 - \left(1 - \frac{z/\theta}{n}\right)^n \\
&\to 1 - e^{z/\theta}
\end{aligned}$$

as $n \to \infty$, which is the CDF of an $\mathrm{Exp}(\theta)$ random variable. Thus, we have that the random variable

$$n2^{-1/n}(2^{1/n}\theta - \hat{\theta}) = n(\theta - 2^{-1/n}\hat{\theta}) \xrightarrow{\mathrm{d}} \mathrm{Exp}(\theta)$$

as $n \to \infty$.

To construct a 95% asymptotic confidence interval, the simplest way is to just consider one-side confidence intervals. Let $Z \sim Exp(\theta)$. Then

$$.05 = P(Z > c) = e^{-c/\theta} \implies c = -\theta \log(.05)$$

And hence

$$P(n(\theta - 2^{-1/n}\hat{\theta}) \leq -\theta \log(.05)) \to .95$$

Removing the probability, we have that a 95% asymptotic confidence interval is

$$\begin{aligned} \left\{\theta : n\theta - n2^{-1/n}\hat{\theta} \leq -\theta \log(.05)\right\} &= \left\{\theta : (n + \log(.05))\theta \leq n2^{-1/n}\hat{\theta}\right\} \\ &= \left\{\theta : \theta \leq \frac{2^{-1/n}n\hat{\theta}}{n + \log(.05)}\right\} \end{aligned}$$

1.d

We wish to test the hypothesis $H_0 : \alpha = \beta$ versus $H_a : \alpha \neq \beta$. What is the likelihood ratio test static. Derive the exact distribution of this test statistic.

<u>*Solution:*</u>
We have already found the MLEs for the unrestricted model $(\hat{\alpha}, \hat{\beta}) = (X_{(n)}, Y_{(n)})$. In the restricted model $\alpha = \beta \equiv \theta$ and the likelihood can be written as

$$\prod_{i=1}^{n}(\theta^{-1}\mathbf{1}_{0 \leq x_i \leq \theta})\prod_{i=1}^{n}(\theta^{-1}\mathbf{1}_{0 \leq y_i \leq \theta}) = \theta^{-2n}\mathbf{1}_{\min(x_{(1)}, y_{(1)}) \geq \theta}\mathbf{1}_{\max(x_{(n)}, y_{(n)}) \leq \theta}$$

The likelihood is 0 if $\max(x_{(n)}, y_{(n)}) \geq \theta$ and it is decreasing in θ otherwise. Thus, the MLE under the null hypothesis is $\tilde{\theta} = \max\{x_{(n)}, y_{(n)}\}$

The likelihood ratio statistic is

$$\begin{aligned}
\Lambda &= \frac{L(\tilde{\theta}, \tilde{\theta})}{L(\hat{\alpha}, \hat{\beta})} \\
&= \frac{(X_{(n)})^{-2n}\mathbf{1}_{\{X_{(n)} > Y_{(n)}\}} + (Y_{(n)})^{-2n}\mathbf{1}_{\{X_{(n)} < Y_{(n)}\}}}{(X_{(n)})^{-n}(Y_{(n)})^{-n}} \\
&= \left(\frac{Y_{(n)}}{X_{(n)}}\right)^n \mathbf{1}_{\{X_{(n)} > Y_{(n)}\}} + \left(\frac{X_{(n)}}{Y_{(n)}}\right)^n \mathbf{1}_{\{X_{(n)} \leq Y_{(n)}\}}
\end{aligned}$$

$$P(\Lambda \leq z) = P(\Lambda \leq z | X_{(n)} \leq Y_{(n)})P(X_{(n)} \leq Y_{(n)}) + P(\Lambda \leq z | X_{(n)} > Y_{(n)})P(X_{(n)} > Y_{(n)})$$

$$\begin{aligned}
P(X_{(n)} \leq Y_{(n)}) &= \int_0^\beta P(X_{(n)} \leq y) f_{Y_{(n)}}(y) dy \\
&= \alpha^{-n}\beta^{-n} n \int_0^\beta y^{2n-1} dy \\
&= \frac{1}{2}\alpha^{-n}\beta^{-n}\beta^{2n} \\
&= \frac{1}{2}\left(\frac{\beta}{\alpha}\right)^n
\end{aligned}$$

And analogously,

$$P(X_{(n)} > Y_{(n)}) = P(Y_{(n)} \leq X_{(n)}) = \frac{1}{2}\left(\frac{\alpha}{\beta}\right)^n$$

$$
\begin{aligned}
P(\Lambda \leq z | X_{(n)} \leq Y_{(n)}) &= P((X_{(n)}/Y_{(n)})^n \leq z) \\
&= P(X_{(n)} \leq z^{1/n} Y_{(n)}) \\
&= \int_0^\beta \left(\frac{z^{1/n} y}{\alpha} \right)^n n\beta^{-n} y^{n-1} dy \\
&= nz\alpha^{-n}\beta^{-n} \int_0^\beta y^{2n-1} dy \\
&= nz\alpha^{-n}\beta^{-n} \frac{1}{2n} \beta^{2n} \\
&= \frac{1}{2} \left(\frac{\beta}{\alpha} \right)^n z
\end{aligned}
$$

and analogously,

$$
P(\Lambda \leq z | X_{(n)} > Y_{(n)}) = \frac{1}{2} \left(\frac{\alpha}{\beta} \right)^n z
$$

Putting it all together,

$$
P(\Lambda \leq z) = \frac{z}{4} \left[\left(\frac{\alpha}{\beta} \right)^{2n} + \left(\frac{\beta}{\alpha} \right)^{2n} \right]
$$

And hence

$$
\Lambda \sim \text{Unif}\left(0, \frac{4}{(\alpha/\beta)^{2n} + (\beta/\alpha)^{2n}} \right)
$$

Under H_0, $\alpha = \beta$ and we have

$$
\Lambda \sim \text{Unif}\,(0, 2)
$$

1.e

Note that $\text{Cov}(\bar{X}_n, \bar{Y}_n) = 0$ since X and Y are independent. Thus, we have

$$\sqrt{n}\begin{pmatrix} \bar{X}_n - \alpha/2 \\ \bar{Y}_n - \alpha/2 \end{pmatrix} \overset{\text{d}}{\to} N(0, \Sigma)$$

where

$$\Sigma = \begin{pmatrix} \alpha^2/12 & 0 \\ 0 & \beta^2/12 \end{pmatrix}$$

Let $g(a, b) = a/b$. Then

$$\frac{\partial g}{\partial a} = -b/a^2|_\mu = -4\beta/\alpha^2$$

$$\frac{\partial g}{\partial \beta} = 1/a|_\mu = 2/\alpha$$

Let $\nabla g = \left(\frac{\partial g}{\partial \alpha}, \frac{\partial g}{\partial \beta}\right)'$. By the Delta Method,

$$\sqrt{n}(\bar{Y}_n/\bar{X}_n - \theta) \overset{\text{d}}{\to} N(0, \tau^2)$$

where

$$\tau^2 = \nabla g^T \Sigma \nabla g == \frac{2\theta^2}{3}$$

$$ARE(\hat{\theta}, \bar{Y}_n/\bar{X}_n) = \frac{\frac{1}{n}(\theta^2/6)}{\frac{1}{n^2}(\theta^2)} = \frac{n}{6}$$

We need to compute the other two asymptotic disributions. Let k_n be a sequence of constants. We have

$$\begin{aligned} k_n\left\{\frac{2\bar{Y}_n}{X_{(n)}} - \frac{\beta}{\alpha}\right\} &= \frac{1}{X_{(n)}}k_n\left\{2\bar{Y}_n - \beta\frac{X_{(n)}}{\alpha}\right\} \\ &= \frac{1}{X_{(n)}}k_n\left\{2\bar{Y}_n - \beta + \beta - \beta\frac{X_{(n)}}{\alpha}\right\} \\ &= \frac{2}{\bar{X}_n}k_n\left\{\bar{Y}_n - \frac{\beta}{2}\right\} + \beta k_n\left\{\frac{1}{X_{(n)}} - \frac{1}{\alpha}\right\} \end{aligned}$$

Now, by part (b), $n(\alpha - X_{(n)}) = O_p(1)$. By the Delta Method, $n(1/\alpha - 1/X_{(n)}) = O_p(1)$ Thus, $\sqrt{n}(\alpha - X_{(n)}) = O_p(n^{-1/2}) = o_p(1)$. Since $X_{(n)}$ is consistent for α, by the continuous mapping theorem, we have $\frac{2}{X_{(n)}} \xrightarrow{\text{p}} \frac{2}{\alpha}$. Thus, by Slutsky's Theorem and the Central Limit Theorem,

$$\sqrt{n}\left\{\frac{2\bar{Y}_n}{X_{(n)}} - \theta\right\} \xrightarrow{\text{d}} \frac{2}{\alpha}N(0, \beta^2/12) \stackrel{\text{d}}{=} N(0, \theta^2/3)$$

Hence,

$$ARE\left(\frac{\bar{Y}_n}{\bar{X}_n}, \frac{2\bar{Y}_n}{\hat{\alpha}}\right) = \frac{\frac{1}{n}\theta^2/3}{\frac{1}{n}\theta^2/6} = 2$$

Finally, let k_n be another arbitrary sequence of constants. Then

$$\begin{aligned}
k_n\left\{\frac{Y_{(n)}}{2\bar{X}_n} - \frac{\beta}{\alpha}\right\} &= \frac{k_n}{2\bar{X}_n}\left\{Y_{(n)} - \beta\frac{2\bar{X}_n}{\alpha}\right\} \\
&= \frac{k_n}{2\bar{X}_n}\left\{Y_{(n)} - \beta + \beta\left(1 - \frac{2\bar{X}_n}{\alpha}\right)\right\} \\
&= \frac{k_n}{2\bar{X}_n}(Y_{(n)} - \beta) + \frac{\beta}{2}k_n\left(\frac{1}{\bar{X}_n} - \frac{1}{\alpha/2}\right)
\end{aligned}$$

By part (b), $Y_{(n)} - \beta = O_p(1/n)$ and by the CLT and the Delta Method, $1/\bar{X}_n - 1/(\alpha/2 = O_p(1/\sqrt{n})$. Thus, if we choose $k_n = \sqrt{n}$, the LHS goes to 0. We need to find the asymptotic distribution of the RHS.

By the CLT, we have

$$\sqrt{n}(\bar{X}_n - \alpha/2) \xrightarrow{\text{d}} N(0, \alpha^2/12)$$

By the Delta Method,

$$\sqrt{n}(1/\bar{X}_n - 1/(\alpha/2)) \xrightarrow{\text{d}} N(0, 1/3)$$

Thus, by the continuous mapping theorem,

$$\sqrt{n}\left\{\frac{Y_{(n)}}{2\bar{X}_n} - \frac{\beta}{\alpha}\right\} \xrightarrow{\text{d}} N(0, \beta^2/12)$$

2 Problem 2

Consider a decision problem with a parameter space Θ having a finite number of values, $\theta_1, \ldots, \theta_l$, $l < \infty$.

2.a

Show that a Bayes rule d_B with respect to a prior distribution Λ on Θ having positive probabilities $\lambda_1, \ldots, \lambda_l > 0$ is admissible.

Proof:
Suppose that d_B is inadmissible. Then there exists some rule d^* such that $R(\theta, d^*) \leq R(\theta, d_B)$ for all θ with strict equality holding for at least one θ.

We have

$$\begin{aligned}
\mathcal{R}(\Lambda, d_B) &= \mathbb{E}_\Lambda R(\theta, d_B) \\
&= \sum_{i=1}^{l} \lambda_i R(\theta_i, d_B) \\
&> \sum_{i=1}^{l} \lambda_i R(\theta_i, d^*) && \text{(since } d_B \text{ is inadmissible)} \\
&\equiv \mathcal{R}(\Lambda, d^*)
\end{aligned}$$

But this contradicts the definition of d_B as a Bayes rule, which minimizes Bayes Risk. Thus, we must have that d_B is admissible.

2.b

The result in part (a) conflicts with other results for continuous parameter spaces where Bayes rules may not be admissible, eg, James-Stein estimation. In the discrete case described above, show that if $\lambda_i = 0$, some $i = 1, \ldots, l$, then the resulting Bayes rule d_B may not be admissible.

Proof:
Suppose $\lambda_i = 0$ and d_B is a Bayes rule with respect to Λ. We want to show that d_B may be inadmissible. Suppose d_B were always admissible. Then for any d and for some θ

$$R(\theta, d_B) < R(\theta, d)$$

Suppose the inequality holds for θ_i and $\lambda_i = 0$. We have

$$\begin{aligned}
\mathcal{R}(\Lambda, d_B) &= \sum_{j=1}^{l} \lambda_j R(\theta_j, d_B) \\
&= \sum_{j=1}^{i-1} \lambda_j R(\theta_j, d_B) + \lambda_i R(\theta_i, d_B) + \sum_{j=i+1}^{l} \lambda_j R(\theta_j, d_B) \\
&= \sum_{j=1}^{i-1} \lambda_j R(\theta_j, d_B) + \sum_{j=i+1}^{l} \lambda_j R(\theta_j, d_B)
\end{aligned}$$

Thus if $R(\theta_i, d_B) < r(\theta_i, d)$ for all d, and $\lambda_i = 0$ we can see that it is possible for $R(\Lambda, d_B)$ not to minimize Bayes risk, a contradiction. Thus, d_B may not be admissible.

2.c

Suppose that the frequentist risk of d_B in part (b) is finite and constant on those θ_i's having $\lambda_i > 0$. Show that this decision rule is minimax, that is, minimizes the maximum risk, on those θ_i's with $\lambda_i > 0$.

<u>*Proof:*</u>
Suppose that the first k elements of Θ have nonzero prior probabilities, (i.e., that $\lambda_1, \lambda_2, \ldots, \lambda_k > 0$ and $\lambda_{k+1} = \lambda_{k+2} = \ldots = \lambda_l = 0$). Since the frequentist risk is constant over the nonzero prior indices, we have

$$R(\theta_i, d_B) = c < \infty$$

for all $i = 1, 2, \ldots, k$ where c is a constant.

Suppose d_B is not minimax over $\theta_1, \ldots, \theta_k$. Then there exists some rule d^* such that

$$\sup_{\theta \in \{\theta_1, \ldots, \theta_k\}} R(\theta, d^*) < \sup_{\theta \in \{\theta_1, \ldots, \theta_k\}} R(\theta, d_B) = c$$

and hence, $R(\theta_i, d^*) < c$ for every $i = 1, \ldots, k$.

We have

$$\begin{aligned}
\mathcal{R}(\Lambda, d^*) &= \sum_{i=1}^{l} \lambda_k R(\theta_i, d^*) \\
&= \sum_{i=1}^{k} \lambda_i R(\theta_i, d^*) && \text{(since } \lambda_i = 0 \text{ for all } i > k.\text{)} \\
&< \sum_{i=1}^{k} \lambda_i R(\theta_i, d_B) && \text{(since } d^* \text{ is minimax and } d_B \text{ has constant risk over these } \theta_i\text{'s)} \\
&= \sum_{i=1}^{l} \lambda_i R(\theta_i, d_B) && \text{(since } \lambda_i = 0 \text{ for all } i > k\text{)} \\
&= R(\Lambda, d_B)
\end{aligned}$$

This contradicts the fact that d_B is bayes with respect to Λ. Hence, d_B is minimax.

2.d

Can anything be said about whether or not d_B in part (b) is minimax on $\theta_i, i = 1, \ldots, l$ Discuss.

Solution:

In (e), (f), and (g), consider the following classification problem. Suppose that X is an observation from the density

$$p(x|\theta) = \theta^1 I(0 < x < \theta)$$

where $I(\dot{)}$ denotes the indicator function and the parameter space is $\Theta = 1, 2, 3$. It is desired to classify X as arising from $p(x|1), p(x|2)$, or $p(x|3)$, under a 0-1 loss function (zero loss for a correct decision, a loss of one for an incorrect decision).

2.e

Find the form of the Bayes rule for this problem.

<u>*Solution:*</u>
The Bayes rule minimizes the posterior expected loss function. Let a_i denote the choice that we take action i, $i = 1, 2, 3$. We have

$$\begin{aligned}
E_{\theta|X} L(\theta, a_1) &= \mathbb{E}_{\theta|X} \{I(\theta = 2) + I(\theta = 3)\} \\
&= P(\theta = 2|X) + P(\theta = 3|X) \\
&= \frac{p(x|2)\lambda_2}{p(x)} + \frac{p(x|3)\lambda_3}{p(x)} \\
&= \frac{1}{p(x)} \left\{ \frac{\lambda_2}{2} I(0 < x < 2) + \frac{\lambda_3}{3} I(0 < x < 3) \right\}
\end{aligned}$$

And similarly, we have

$$E_{\theta|X} L(\theta, a_2) = \frac{1}{p(x)} \left\{ \lambda_1 I(0 < x < 1) + \frac{\lambda_3}{3} I(0 < x < 3) \right\}$$

$$E_{\theta|X} L(\theta, a_3) = \frac{1}{p(x)} \left\{ \lambda_1 I(0 < x < 1) + \frac{\lambda_2}{2} I(0 < x < 2) \right\}$$

Let $\phi_i(x) = P(d(x) = i)$.

<u>Case 1:</u> $0 < x < 1$.

$$\begin{aligned}
\mathbb{E}_{\theta|X} L(\theta, a_1) &= \frac{1}{p(x)} \left\{ \frac{\lambda_2}{2} + \frac{\lambda_3}{3} \right\} \\
\mathbb{E}_{\theta|X} L(\theta, a_2) &= \frac{1}{p(x)} \left\{ \lambda_1 + \frac{\lambda_3}{3} \right\} \\
\mathbb{E}_{\theta|X} L(\theta, a_3) &= \frac{1}{p(x)} \left\{ \lambda_1 + \frac{\lambda_2}{2} \right\}
\end{aligned}$$

Thus, for $0 < x < 1$,

$$\begin{aligned}\phi_1(x) = 1 &\iff \lambda_2/2 + \lambda_3/3 < \lambda_1 + \lambda_3/3 \text{ and } \lambda_2/2 + \lambda_3/3 < \lambda_1 + \lambda_2/2\\ &\iff \lambda_1 > \lambda_2/2 \text{ and } \lambda_1 > \lambda_3/3\end{aligned}$$

$$\begin{aligned}&\phi_1(x) = \gamma_1 = 1 - \phi_2(x) \iff \lambda_1 = \lambda_2/2, \lambda_1 > \lambda_3/3\\ &\phi_1(x) = \gamma_2 = 1 - \phi_3(x) \iff \lambda_1 = \lambda_3/3, \lambda_1 > \lambda_2/2\\ &\phi_1(x) = \gamma_3 \iff \lambda_1 = \lambda_2/2 = \lambda_3/3\\ &\phi_2(x) = 1 \iff \lambda_1 < \lambda_2/2 \text{ and } \lambda_2 > 2\lambda_3/3\\ &\phi_2(x) = \gamma_4 \iff \lambda_1 < \lambda_2/2 \text{ and } \lambda_2 = 2\lambda_3/3\\ &\phi_2(x) = \gamma_5 \iff \lambda_1 = \lambda_2/2 = \lambda_3/3\\ &\phi_3(x) = 1 \iff \lambda_1 < \lambda_3/3 \text{ and } \lambda_2 < 2\lambda_3/3\\ &\phi_3(x) = 1 - \gamma_3 - \gamma_5 \iff \lambda_1 = \lambda_2/2 = \lambda_3/3\end{aligned}$$

<u>Case 2:</u> $1 \le x < 2$

$$\begin{aligned}\mathbb{E}_{\theta|X} L(\theta, a_1) &= \frac{\lambda_2}{2} + \frac{\lambda_3}{3}\\ \mathbb{E}_{\theta|X} L(\theta, a_2) &= \frac{\lambda_3}{3}\\ \mathbb{E}_{\theta|X} L(\theta, a_3) &= \frac{\lambda_2}{2}\end{aligned}$$

Thus, for $1 \le x < 2$

$$\begin{aligned}&\phi_1(x) = 0\\ &\phi_2(x) = 1 \iff \lambda_2 > 2\lambda_3/3\\ &\phi_2(x) = \gamma_6 = 1 - \phi_3(x) \iff \lambda_2 = 2\lambda_3/3\\ &\phi_3(x) = 1 \iff \lambda_2 < 2\lambda_3/3\end{aligned}$$

<u>Case 3:</u> $2 \le x < 3$

$$\begin{aligned}\mathbb{E}_{\theta|X} L(\theta, a_1) &= \frac{\lambda_3}{3}\\ \mathbb{E}_{\theta|X} L(\theta, a_2) &= \frac{\lambda_3}{3}\\ \mathbb{E}_{\theta|X} L(\theta, a_3) &= 0\end{aligned}$$

Thus, for $2 \le x < 3$ $\phi_1(x) = \phi_2(x) = 0$, $\phi_3(x) = 1$

Let $\phi_i(x) = P(d(x) = i)$. Then the Bayes rule is given by

$$\phi_1(x) = I(0 < x < 1)\,[I(\lambda_1 > \lambda_2/2)I(\lambda_1 > \lambda_3/3) + \gamma_1 I(\lambda_1 = \lambda_2/2)I(\lambda_1 > \lambda_3/3) \\ +\gamma_2 I(\lambda_1 > \lambda_2/2)I(\lambda_1 = \lambda_3/3) + \gamma_3 I(\lambda_1 = \lambda_2/2)I(\lambda_1 = \lambda_3/3)]$$

$$\phi_2(x) = I(0 \le x < 1)\,[I(\lambda_1 < \lambda_2/2)I(\lambda_2 > 2\lambda_3/3) + (1-\gamma_1)I(\lambda_1 = \lambda_2/2)I(\lambda_1 > \lambda_3/3) \\ +\gamma_4 I(\lambda_1 < \lambda_2/2)I(\lambda_2 = 2\lambda_3/3) + \gamma_5 I(\lambda_1 = \lambda_2/2)I(\lambda_1 = \lambda_3/3)] \\ + I(1 \le x < 2)\,[I(\lambda_2 > 2\lambda_3/3) + \gamma_6 I(\lambda_2 = 2\lambda_3/3)]$$

$$\phi_3(x) = I(0 \le x < 1)\,[I(\lambda_1 < \lambda_3/3)I(\lambda_2 < 2\lambda_3/3) + (1-\gamma_2)I(\lambda_1 > \lambda_2/2)I(\lambda_1 = \lambda_3/3) \\ +(1-\gamma_4)I(\lambda_1 < \lambda_3/3)I(\lambda_2 = 2\lambda_3/3) + (1-\gamma_3-\gamma_5)I(\lambda_1 = \lambda_2/2)I(\lambda_1 = \lambda_3/3)] \\ + I(1 \le x < 2)\,[I(\lambda_2 > 2\lambda_3/3) + (1-\gamma_6)I(\lambda_2 = 2\lambda_3/3) + I(2 \le x < 3)]$$

2.f

Find the decision rule which minimizes the maximum risk over Θ and the corresponding least favorable prior distribution.

Solution: To find the minimax rule, we seek a Bayes rule with constant risk. Note that

$$R(\theta_i, \phi) = \sum_{j=1}^{3} L(\theta_i, a_j)\mathbb{E}\phi_j(x) = 1 - \mathbb{E}_{\theta_i}\phi_i(x)$$

$R(\theta_i, \phi)$ is constant if and only if $\mathbb{E}_{\theta_i}\phi_i(x)$ is constant. Note that $\mathbb{E}_{\theta_3}\phi_3(x) \ge P(2 \le x < 3) > 0$, so we must have $E_{\theta_1}\phi_1(x) > 0$ and thus $\phi_1(x) \ne 0$ when $0 \le x < 1$

Looking at $\phi_1(x)$, the only case where $E_{\theta_1}\phi_1(x)$ is constant can be the case $\lambda_1 = \lambda_2/2 = \lambda_3/3$. Thus, the least favorable prior satisfies $\lambda_1 = \lambda_2/2 = \lambda_3/3$. Since the priors add up to 1, we have

$$1 = \lambda_1 + 2\lambda_1 + 3\lambda_1 = 6\lambda_1 \iff \lambda_1 = 1/6$$

and hence $\lambda_2 = 2/6 = 1/3$ and $\lambda_3 = 3/6 = 1/2$. Hence the least favorable prior is $(\lambda_1, \lambda_2, \lambda_3) = (1/6, 1/3, 1/2)$.

In order for the risk to be constant, we need

$$\gamma_3 = \frac{1}{2}(\gamma_5 + \gamma_6) = \frac{1}{3}(1 - \gamma_3 - \gamma_5 + 1 - \gamma_6) \\ \iff 2\gamma_3 = \gamma_5 + \gamma_6, 3\gamma_3 = (2 - \gamma_3 - \gamma_5 - \gamma_6)$$

Hence, the minimax rule is any rule satisfying the least favorable prior above as well as the constraints on the γ's.

2.g

Looking at the posterior expected losses above, we can find the Bayes rule

<u>Case 1:</u> $0 < x < 1$

$$\begin{aligned} \phi_1(x) = 1 &\iff \lambda_1 > \lambda_2/2 \\ \phi_1(x) = \gamma_1 = 1 - \phi_2(x) &\iff \lambda_1 = \lambda_2/2 \\ \phi_2(x) = 1 &\iff \lambda_1 < \lambda_2/2 \end{aligned}$$

<u>Case 2:</u> $1 \leq x < 2$
$\phi_2(x) = 1$ and $\phi_1(x) = 0$.

<u>Case 3</u> $2 \leq x < 3$
$\phi_1(x) = \gamma_2 = 1 - \phi_2(x)$

And thus, we have

$$\phi_1(x) = I(0 < x < 1)[I(\lambda_1 > \lambda_2/2) + \gamma_1 I(\lambda_1 = \lambda_2/2)] + I(2 \leq x < 3)\gamma_2$$

$$\phi_2(x) = I(0 < x < 1)[I(\lambda_1 < \lambda_2/2)+(1-\gamma_1)I(\lambda_1 = \lambda_2/2)]+I(1 \leq x < 2)+(1-\gamma_2)I(2 < x \leq 3)$$

In order for the Bayes rule to be constant risk, we neeed $E_{\theta_i}\phi_i$ to be constant for all i. Thus, this happens iff $\lambda_1 = \lambda_2/2$ So the least favorable prior is $(1/3, 2/3, 0)$.

$\gamma_1 = 1/2(1 - \gamma_1) + 1 \implies \frac{3}{2}\gamma_1 = 3/2 \implies \gamma_1 = 1$

3 Question 3

Suppose that (X, Y) are two random variables with joint distribution

$$f(x, y|\alpha, \beta) = c(\alpha, \beta)\exp(-\alpha x - \beta y)\sum_{j=0}^{\infty}\frac{x^j y^j}{(j!\,)^2}$$

for $x > 0$, $y > 0$. Also, let (X_1, Y_1), . . . ,(X_n, Y_n) be a random sample from (X, Y), and let $\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$, $\bar{Y} = \frac{1}{n}\sum_{i=1}^{n} Y_i$.

3.a

Show that the joint distribution of (X, Y) in (1) is in the multiparameter exponential family and identify the rank, show that $c(\alpha, \beta) = \alpha\beta - 1$, and find the parameter space of (α, β).

<u>*Solution:*</u>
We can write

$$\begin{aligned} f(x, y|\alpha, \beta) &= \exp\left\{-\alpha x - \beta y - \log 1/c(\alpha, \beta) + \log\sum_{j=0}^{\infty}\frac{x^j y^j}{(j!\,)^2}\right\} \\ &= \exp\left\{T(x, y)'\theta - b(\theta) - d(x, y)\right\} \end{aligned}$$

where $T(x, y) = (-x, -y)'$, $\theta = (\alpha, \beta)'$ $b(\theta) = \log 1/c(\alpha, \beta)$, and $d(x, y) = -\log\sum_{j=1}^{\infty}\frac{x^j y^j}{(j!)^2}$. Thus, f is in the multiparameter exponential family. The rank is 2 since x, y are linearly independent and so are α, β.

We need to find the normalizing constant. Let $k(x, y)$ denote the kernel of the joint distribution. We have

$$
\begin{aligned}
\int_{\mathcal{X}}\int_{\mathcal{Y}} k(x,y)dydx &= \int_0^\infty \int_0^\infty e^{-\alpha x}e^{-\beta y}\sum_{j=0}^{\infty}\frac{x^j y^j}{(j!)^2}dydx \\
&= \int_0^\infty \int_0^\infty e^{-\alpha x}e^{-\beta y}\lim_{n\to\infty}\sum_{j=0}^{n}\frac{x^j y^j}{(j!)^2}dydx \\
&= \lim_{n\to\infty}\int_0^\infty \int_0^\infty e^{-\alpha x}e^{-\beta y}\sum_{j=0}^{n}\frac{x^j y^j}{(j!)^2}dydx && \text{(MCT)} \\
&= \lim_{n\to\infty}\sum_{j=0}^{n}\int_0^\infty \int_0^\infty e^{-\alpha x}e^{-\beta y}\sum_{j=0}^{n}\frac{x^j y^j}{(j!)^2}\ dy\ dx \\
&= \lim_{n\to\infty}\sum_{j=0}^{n}\frac{1}{(j!)^2}\int_0^\infty x^j e^{-\alpha x}\int_0^\infty y^j e^{-\beta y}\ dy\ dx \\
&= \lim_{n\to\infty}\sum_{j=0}^{n}\frac{1}{(j!)^2}\left[\Gamma(j+1)\right]^2\alpha^{-(j+1)}\beta^{-(j+1)} && \text{(provided } \alpha>0 \text{ and } \beta>0\text{)} \\
&= \sum_{j=0}^{\infty}\left(\frac{1}{\alpha\beta}\right)^{j+1} \\
&= \frac{1}{\alpha\beta}\sum_{j=0}^{\infty}\left(\frac{1}{\alpha\beta}\right)^{j} && \text{(provided } |\tfrac{1}{\alpha\beta}|<1\text{)} \\
&= \frac{1}{\alpha\beta}\left(\frac{1}{1-1/(\alpha\beta)}\right) \\
&= \frac{1}{\alpha\beta-1}
\end{aligned}
$$

Hence, we must have $c(\alpha,\beta)=\alpha\beta-1$ since the PDF must integrate to 1. The MCT result holds because each term in the series is non-negative so the series itself is non-negative and nondecreasing. Moreover, the series is convergent because factorials dominate powers. Thus, the MCT conditions are satisfied.

In order to derive the result, we made two assumptions about α and β. In particular, we assumed $\alpha,\beta>0$ and $\left|\frac{1}{\alpha\beta}\right|<1$. Combining these together, we must have $\alpha\beta>1$, i.e., we must have $\alpha>1/\beta$

Thus, the parameter space is

$$\Theta=\{(\alpha,\beta):\alpha\beta>1\}$$

So clearly α,β are linearly independent and hene the exponential family is full rank.

3.b

Derive the marginal distribution of X from (1) and show that $E(X) = \frac{\beta}{\alpha\beta - 1}$

$$
\begin{aligned}
f_X(x) &= \int_0^\infty f(x,y)\ dy \\
&= \int_0^\infty (\alpha\beta - 1)e^{-\alpha x}e^{-\beta y}\sum_{j=0}^{\infty}\frac{x^j y^j}{(j!)^2}\ dy \\
&= (\alpha\beta - 1)e^{-\alpha x}\int_0^\infty e^{-\beta y}\lim_{n\to\infty}\sum_{j=0}^{n}\frac{x^j y^j}{(j!)^2}\ dy \\
&= (\alpha\beta - 1)e^{-\alpha x}\lim_{n\to\infty}\sum_{j=0}^{n}\frac{x^j}{(j!)^2}\int_0^\infty y^j e^{-\beta y}\ dy && \text{(MCT)} \\
&= (\alpha\beta - 1)e^{-\alpha x}\sum_{j=0}^{\infty}\frac{x^j}{(j!)^2}\ \frac{\Gamma(j+1)}{\beta^{j+1}} \\
&= \frac{\alpha\beta - 1}{\beta}\ e^{-\alpha x}\sum_{j=0}^{\infty}\frac{x^j}{\beta^j j!} \\
&= \frac{\alpha\beta - 1}{\beta}e^{-\alpha x}e^{x/\beta} \\
&= (\alpha - 1/\beta)e^{-x(\alpha - 1/\beta)}
\end{aligned}
$$

Hence, $X \sim \text{Exp}(\alpha - 1/\beta)$, where we are using the rate parameterization. Thus, $E(X) = \frac{1}{\alpha - 1/\beta} = \frac{\beta}{\alpha\beta - 1}$.

3.c

From part (a), since the canonical parameter is $\theta = (\theta_1, \theta_2), = (\alpha, \beta)$ the cumulant function is given by

$$b(\theta) = b(\alpha, \beta) = \log(1/c(\alpha, \beta)) = \log S(\alpha, \beta)$$

Let $t = (t_1, t_2)'$. Then the moment generating function (mgf) is

$$\begin{aligned} M(t) &= \exp\{b(\theta + t) - b(\theta)\} \\ &= \exp\left\{\log S(\alpha + t_1, \beta + t_2) - \log(S)\right\} \\ &= S^{-1} S(\alpha + t_1, \beta + t_2) \end{aligned}$$

Thus,

$$\begin{aligned} \mathbb{E}X^j Y^k &= \left.\frac{\partial^2 M^{j+k}}{\partial t_1^j \partial t_2^k}\right|_{t=0} \\ &= S^{-1} \frac{\partial^2}{\partial t_1^j \partial t_2^k} \left[(\alpha + t_1)(\beta + t_2) - 1\right]^{-1}\Big|_{t=0} \\ &= S^{-1}(-1)^{j+k} \left[(\alpha + t_1)(\beta + t_2) - 1\right]^{-(j+k+1)} (\beta + t_2)^j (\alpha + t_1)^k \Big|_{t=0} \\ &= S^{-1}(-1)^{j+k} \alpha^k \beta^j \left[\alpha\beta - 1\right]^{-(j+k+1)} \end{aligned}$$

Now,

$$\begin{aligned} \frac{\partial S^{j+k}}{\partial \alpha^j \beta^k} &= \frac{\partial}{\partial \alpha^j \beta^k} (\alpha\beta - 1)^{-1} \\ &= (-1)^{j+k} (\alpha\beta - 1)^{-(j+k+1)} \alpha^k \beta^j \end{aligned}$$

The result follows except the -1 term is included in the partial of S.