1. Consider a distribution with the probability density function

$$f(y;\mu,\phi) = \frac{1}{\Gamma(\phi)}(\phi\mu)^{\phi} y^{\phi-1}\exp(-\phi\mu y), \quad \text{for } y>0,\ \mu>0,\ \text{and } \phi>0,$$

where $\Gamma()$ is the Gamma function.

(a) Show that Y belongs to the exponential family.
Solution: Note that

$$f(y) = \frac{1}{\Gamma(\phi)}(\phi\mu)^{\phi}\exp\{(\phi-1)\log(y)-\phi\mu y\}$$

Hence, it follows exponential family with

$$\begin{aligned} A(\theta) &= \frac{1}{\Gamma(\phi)}(\phi\mu)^{\phi} \\ T &= (\log(Y), Y) \\ \zeta(\theta) &= (\phi-1,\ -\phi\mu), \\ h(y) &= 1. \end{aligned}$$

(b) Determine the value k such that it belongs to a strictly k-parameter exponential family.
Solution: It is easy to see that $\phi-1$ and $-\phi\mu$ are linearly independent. If we consider

$$c_1\log(Y)+c_2Y = c_0$$

for all y, then we must have

$$\begin{aligned} c_1\log(1)+c_2 &= c_0, \quad y=1, \\ c_1\log(e)+c_2e &= c_0, \quad y=e, \\ c_1\log\left(e^2\right)+c_2e^2 &= c_0, \quad y=e^2. \end{aligned}$$

The only solution of the linear system is $c_0=c_1=c_2=0$. Hence, $k=2$.

(c) Find $\mathrm{E}[\log(Y)]$, $\mathrm{E}[Y]$ and $\mathrm{Cov}(\log(Y), Y)$. If you need the derivative of a gamma function such as $\Gamma(x)$, just express the derivative as $\Gamma'(x)$, etc.
Solution: Define $\zeta_1=\phi-1$ and $\zeta_2=-\phi\mu<0$. We can express the distribution using natural parameters as

$$f(y) = \frac{1}{\Gamma(\zeta_1+1)}(-\zeta_2)^{\zeta_1+1}\exp\{\zeta_1\log(y)+\zeta_2 y\}$$

Then,

$$\begin{aligned} \mathrm{E}[\log(Y)] = \mathrm{E}[T_1(x)] &= -\frac{d\log C(\zeta)}{d\zeta_1} = -\frac{d-\log\Gamma(\zeta_1+1)+(\zeta_1+1)\log(-\zeta_2)}{d\zeta_1} \\ &= \frac{1}{\Gamma(\zeta_1+1)}\frac{d\Gamma(\zeta_1+1)}{d\zeta_1}-\log(-\zeta_2) \\ &= \frac{1}{\Gamma(\phi)}\frac{d\Gamma(\phi)}{d\phi}-\log(\phi\mu), \\ \mathrm{E}[Y] = \mathrm{E}[T_2(x)] &= -\frac{d\log C(\zeta)}{d\zeta_2} = -\frac{d-\log\Gamma(\zeta_1+1)+(\zeta_1+1)\log(-\zeta_2)}{d\zeta_2} \\ &= -\frac{\zeta_1+1}{\zeta_2} = \frac{1}{\mu}, \\ \mathrm{Cov}(T_1,T_2) &= -\frac{d^2\log C(\zeta)}{d\zeta_1 d\zeta_2} = -\frac{1}{\zeta_2} = \frac{1}{\phi\mu}. \end{aligned}$$

(d) Suppose that we have an i.i.d. sample Y_1, ..., Y_n from this distribution. Show that the joint distribution of $(Y_1, \cdots, Y_n)$ also belongs to an exponential family.
Solution: From Theorem 2.2, it is easy to see that

$$f(\boldsymbol{y}) = \left[\frac{1}{\Gamma(\phi)}(\phi\mu)^{\phi}\right]^n \exp\left\{(\phi-1)\sum_{i=1}^n \log(y_i) - \phi\mu\sum_{i=1}^n y_i\right\}$$

where

$$\begin{aligned} A(\theta) &= \left[\frac{1}{\Gamma(\phi)}(\phi\mu)^{\phi}\right]^n \\ T &= \left(\sum_{i=1}^n \log(y_i), \sum_{i=1}^n y_i\right) \\ \zeta(\theta) &= (\phi-1,\ -\phi\mu), \\ h(y) &= 1. \end{aligned}$$

2. (2p) Suppose that X follows a Poisson distribution with mean λ. Find the Fisher information of $\theta = \log\lambda$.
Solution: Poisson with mean λ has the pmf

$$p(x \mid \lambda) = \exp(x\log(\lambda) - \log(x!) - \lambda).$$

Define $\theta = \log(\lambda)$ as the natural parameter. Then

$$p(x \mid \theta) = \exp\left(x\theta - \log(x!) - e^{\theta}\right).$$

Differenation can pass through integration for exponential family. The second derivative of $\log p(x)$ is

$$\frac{d^2 \log p(x)}{d\theta^2} = -\exp(\theta).$$

Hence, the Fisher information is

$$-\mathrm{E}\left[\frac{d^2 \log p(x)}{d\theta^2}\right] = \exp(\theta).$$

3. (1p) Let X_1, ..., X_n be a random sample from uniform distribution on $[a, b]$. Find the sufficient statistic when $a = \theta$ is unknown but $b = 1$ is known, and when $a = \theta - 0.5$ and $b = \theta + 0.5$ with unknown θ.
Solution: When $b = 1$ is known, the joint density is

$$p(x \mid a) = \prod_{i=1}^n 1_{(a,1)}(x_i) = 1_{(a,1)}(x_{\min}).$$

By the factorization theorem, $T = X_{\min}$ with $h(x) = 1$.
When $a = \theta - 0.5$ and $b = \theta + 0.5$, the joint density is

$$p(x \mid a) = \prod_{i=1}^n 1_{(\theta-0.5,\theta+0.5)}(x_i) = 1_{(\theta-0.5,\theta+0.5)}(x_{\min})\, 1_{(\theta-0.5,\theta+0.5)}(x_{\max}).$$

By the factorization theorem, $T = (X_{\min}, X_{\max})$ with $h(x) = 1$.

4. (2p)Let X_1, ..., X_n be iid $N\left(0,\sigma^2\right)$. Find the minimal sufficient statistic for σ^2. Show also that sample variance is not minimal sufficient.
Solution: Consider the ratio

$$\frac{p(x)}{p(y)} = \frac{\prod_{i=1}^n \frac{1}{\sqrt{2\pi}\sigma}\exp\left\{-\frac{x_i^2}{2\sigma^2}\right\}}{\prod_{i=1}^n \frac{1}{\sqrt{2\pi}\sigma}\exp\left\{-\frac{y_i^2}{2\sigma^2}\right\}} = \exp\left\{\frac{\sum_{i=1}^n y_i^2 - \sum_{i=1}^n x_i^2}{2\sigma^2}\right\},$$

which does not depend on σ^2 if and only if $\sum_{i=1}^n y_i^2 - \sum_{i=1}^n x_i^2 = 0$. Hence, a minimal sufficient statistic is $\sum_{i=1}^n X_i^2$.
Also note that

$$\frac{\prod_{i=1}^n \frac{1}{\sqrt{2\pi}\sigma}\exp\left\{-\frac{x_i^2}{2\sigma^2}\right\}}{\prod_{i=1}^n \frac{1}{\sqrt{2\pi}\sigma}\exp\left\{-\frac{y_i^2}{2\sigma^2}\right\}} = \exp\left\{\frac{\left(\sum_{i=1}^n y_i^2 - n\bar{y}^2\right) - \left(\sum_{i=1}^n x_i^2 - n\bar{x}^2\right) + n\bar{y}^2 - n\bar{x}^2}{2\sigma^2}\right\}$$

Sample variances being the same is not enough to guarantee that the ratio is independent of σ^2. We also need $\bar{y}^2 = \bar{x}^2$. Hence, sample variance is not minimal sufficient.

5. Let X_1, ..., X_n be i.i.d. Poisson random variables with mean $\lambda > 0$.

   (a) Find the MLE of $\theta = \log(\lambda)$.
   Solution: The log-likelihood function is

   $$\ell(\lambda) = \sum_{i=1}^n \left[x_i \log(\lambda) - \log(x_i!) - \lambda\right].$$

   Note that

   $$\frac{\partial \ell(\lambda)}{\partial \lambda} = \sum_{i=1}^n \left[\frac{x_i}{\lambda} - 1\right],$$
   $$\frac{\partial^2 \ell(\lambda)}{\partial \lambda^2} = -\frac{\sum_{i=1}^n x_i}{\lambda^2} < 0.$$

   Hence the MLE of λ is $\hat{\lambda} = \bar{X}$. Because of the invariance property of MLE, $\hat{\theta} = \log\left(\bar{X}\right)$.

   (b) Is the MLE of θ unbiased?
   Solution: By Jensen's inequality, $E\left[\log\left(\bar{X}\right)\right] \le \log\left(E\left(\bar{X}\right)\right) = \log(\lambda) = \theta$, where the equality does not hold for the $\log()$ function that is strictly concave. Hence, the MLE of θ is not an unbiased estimator of θ.

   (c) Find lower bound in the variance for the unbiased estimators for θ.
   Solution: The Fisher information is

   $$\mathcal{I}(\lambda) = -E\left[\frac{\partial^2 \ell(\lambda)}{\partial \lambda^2}\right] = \frac{n}{\lambda}.$$

   Hence, the Cramér-Rao lower bound is

   $$\mathrm{V}(T) \ge \frac{\left[g'(\lambda)\right]^2}{\mathcal{I}(\lambda)} = \frac{\left[\frac{d\log(\lambda)}{d\lambda}\right]^2}{\mathcal{I}(\lambda)} = \frac{1}{n\lambda} = \frac{1}{n\exp(\theta)}.$$

   Strictly speaking, we need to make sure the regularity conditions hold. It is easy to see that the Poisson distributions have a common support and the parameter space $\Theta = (0,\infty) \subseteq \mathbb{R}$ is an open interval (finite or infinite). For any $\boldsymbol{x} \in \mathcal{A}$ and all $\theta \in \Theta$, the derivative

   $$\frac{\partial p(\boldsymbol{x};\theta)}{\partial \theta} = p(\boldsymbol{x};\theta)\left[\sum_{i=1}^n x_i - n\exp(\theta)\right]$$

   exists and is finite. The Poisson distribution belongs to the exponential family. Hence, we can interchange the order of differentiation and integration.

(d) Is your MLE of $\log(\lambda)$ a consistent estimator of $\log(\lambda)$?
Solution: By the law of large numbers, $\bar{X}$ is a consistent estimator of λ. Note that log is a continuous function. Hence, $\log(\bar{X})$ is also consistent by the continuous mapping theorem. In fact, as an MLE, you can also show consistency.

(e) Approximate the distribution of your MLE of $\log(\lambda)$.
Solution: The distribution of MLE can be applied by a normal distribution:

$$\sqrt{n}\left(g\left(\hat{\lambda}\right) - g(\lambda)\right) \xrightarrow{\mathcal{D}} N(0, \Sigma(\lambda)),$$

where

$$\Sigma(\lambda) = \frac{\partial g(\lambda)}{\partial \lambda^T}\left[\mathcal{I}_1(\lambda)\right]^{-1}\left(\frac{\partial g(\lambda)}{\partial \lambda^T}\right)^T.$$

Here, $g(\lambda) = \log(\lambda)$, and

$$\begin{aligned}\Sigma(\lambda) &= \left(\frac{d\log(\lambda)}{d\lambda}\right)^2\left[\mathcal{I}_1(\lambda)\right]^{-1} \\ &= \frac{1}{\lambda^2}\left[\frac{1}{\lambda}\right]^{-1} = \frac{1}{\lambda}\end{aligned}$$

(f) Show that $\bar{X}$ is the UMVUE of λ.
Solution: We have obtained above

$$\mathcal{I}(\lambda) = -\left(E\left[\frac{\partial^2 \ell(\lambda)}{\partial \lambda^2}\right]\right)^{-1} = \frac{n}{\lambda}.$$

Hence, the smallest variance is λ/n. The variance of $\bar{X}$ is also λ/n. Hence, $\bar{X}$ is the UMVUE of λ, since it is an efficient estimator of λ (attaining the C-R lower bound).

(g) Consider the estimator $\hat{\lambda} = S^2$ of λ, where S^2 is the sample variance. Show that $\mathrm{E}\left[\hat{\lambda} \mid \bar{X}\right] = \bar{X}$.
Solution: We know that $\bar{X}$ is the UMUVE of λ. Hence, $\mathrm{E}\left[\hat{\lambda} \mid \bar{X}\right] = \bar{X}$, if $\bar{X}$ is also sufficient and complete. To show that $\bar{X}$ is sufficient and complete, we express the likelihood function as

$$L(\lambda) = \exp\{-n\lambda\}\exp\left\{\sum_{i=1}^n x_i \log(\lambda)\right\} - \lambda\prod_{i=1}^n x_i!,$$

which belongs to the exponential family. It is easy to see that $\sum_{i=1}^n x_i$ is sufficient. The natural parameter is $\log(\lambda)$ with natural parameter space $\mathbb{R}$, which contains an open set. Hence, $\sum_{i=1}^n x_i$ is also complete.

6. (1p) Let X_1, ..., X_n be i.i.d. random variables from a distribution with the probability density function

$$p(x;\theta) = \frac{\log(\theta)}{\theta - 1}\theta^x, \qquad 0 < x < 1, \theta > 1.$$

Is a function of θ for which there exists an unbiased estimator whose variance attains the Cramér-Rao lower bound? If so, find it. If not, show why not.
Solution: It is easy to see that the distribution belongs to the exponential family, so we can change the order of integration and differentiation. It is also easy to see that the distributions

have a common support and the parameter space $\Theta = (1, \infty) \subseteq \mathbb{R}$ is an open interval (finite or infinite). For any $\boldsymbol{x} \in \mathcal{A}$ and all $\theta \in \Theta$, the derivative

$$\begin{aligned} \frac{\partial \log p(\boldsymbol{x}; \theta)}{\partial \theta} &= \frac{\partial n \log[\log \theta] - n \log(\theta - 1) + \sum_{i=1}^{n} x_i \log \theta}{\partial \theta} \\ &= \frac{n}{\log \theta} \frac{1}{\theta} - \frac{n}{\theta - 1} + \frac{\sum_{i=1}^{n} x_i}{\theta} \end{aligned}$$

exists and is finite.
Note that

$$\frac{\partial \log p(\boldsymbol{x}; \theta)}{\partial \theta} = \frac{n}{\theta}\left[\frac{1}{n}\sum_{i=1}^{n} x_i - \frac{\theta}{\theta - 1} + \frac{1}{\log(\theta)}\right].$$

We know that $\bar{X}$ is the efficient estimator of $\frac{\theta}{\theta-1} - \frac{1}{\log(\theta)}$.

7. (1p) Let X be a sample of iid random variables from uniform distribution on $[\theta - 0.5, \theta + 0.5]$. Show that $(X_{\min}, X_{\max})$ is not complete.
Solution: Consider the function $h(t_1, t_2) = t_1 - t_2 - \frac{n-1}{n+1}$. The cumulative distribution function of T_1 is

$$P(T_1 \leq t) = 1 - \prod_{i=1}^{n}\left[\theta + \frac{1}{2} - t\right].$$

The probability density is

$$f(t) = n\left(\theta + \frac{1}{2} - t\right)^{n-1}.$$

The expectation is $\theta + \frac{1}{2} - \frac{n}{n+1}$. The cumulative distribution function of T_2 is

$$P(T_2 \leq t) = \prod_{i=1}^{n}\left[t - \theta + \frac{1}{2}\right].$$

The probability density is

$$f(t) = n\left(t - \theta + \frac{1}{2}\right)^{n-1}.$$

The expectation is $\frac{n}{n+1} + \theta - \frac{1}{2}$.Its expectation is

$$\begin{aligned} E[h(X_{\min}, X_{\max})] &= \theta + \frac{1}{2} - \frac{n}{n+1} - \left(\frac{n}{n+1} + \theta - \frac{1}{2}\right) + \frac{n-1}{n+1} \\ &= 1 - \frac{2n}{n+1} + \frac{n-1}{n+1} = 0 \end{aligned}$$

for any θ. But $h \neq 0$. Hence, $T = (X_{\min}, X_{\max})$ is not complete.

8. (1p) Suppose that we have observed an X with probability mass function

$$p(x) = \binom{x+k-1}{k-1} \theta^k (1-\theta)^x.$$

Find the sufficient and complete statistic.
Solution: Consider the probability mass function

$$p(x) = \binom{x+k-1}{k-1} \theta^k (1-\theta)^x.$$

Note that

$$p(x) \quad = \quad \theta^k \exp\{x \log(1-\theta)\} \binom{x+k-1}{k-1},$$

which means that the distribution belongs to the exponential family. Hence, X is sufficient and complete.

9. Show that a test with $C_1 = \left\{ \frac{1}{10} \sum_{i=1}^{10} x_i > c \right\}$ is a Neyman-Pearson test for the test problem H_0: Bernoulli with probability 0.25 versus H_1: Bernoulli with probability 0.5 .
Solution: The likelihoods are

$$\begin{aligned} p_0(\boldsymbol{x}) &= (1/4)^{\sum_{i=1}^{10} X_i} (1-1/4)^{10-\sum_{i=1}^{10} X_i} \\ p_1(\boldsymbol{x}) &= (1/2)^{\sum_{i=1}^{10} X_i} (1-1/2)^{10-\sum_{i=1}^{10} X_i}. \end{aligned}$$

The Neyman-Pearson test is

$$\phi(\boldsymbol{x}) \quad = \quad \begin{cases} 1, & \text{if } p_0(\boldsymbol{x}) < k p_1(\boldsymbol{x}), \\ \gamma, & \text{if } p_0(\boldsymbol{x}) = k p_1(\boldsymbol{x}), \\ 0, & \text{if } p_0(\boldsymbol{x}) > k p_1(\boldsymbol{x}), \end{cases}$$

where

$$\begin{aligned} \frac{p_0(\boldsymbol{x})}{p_1(\boldsymbol{x})} &= \left(\frac{1}{2}\right)^{\sum_{i=1}^{10} X_i} \left(\frac{3}{2}\right)^{10-\sum_{i=1}^{10} X_i} \\ &= \left(\frac{1}{2}\right)^{10} 3^{10} - \sum_{i=1}^{10} X_i. \end{aligned}$$

The ratio is small if and only if $\sum_{i=1}^{10} X_i$ is large. Hence, the test is equivalent to

$$\phi(\boldsymbol{x}) \quad = \quad \begin{cases} 1, & \text{if } \bar{\boldsymbol{x}} > c, \\ \gamma, & \text{if } \bar{\boldsymbol{x}} = c, \\ 0, & \text{if } \bar{\boldsymbol{x}} < c. \end{cases}$$

In order to determine the value of c, we solve the equation

$$\mathrm{E}[\phi(\boldsymbol{X})] = \mathrm{P}\left(\bar{\boldsymbol{X}} > c\right) + \gamma \mathrm{P}\left(\bar{\boldsymbol{X}} = c\right) \quad = \quad \alpha.$$

10. Let P_0, P_1, P_2 be the probability distributions assigning to the integers 1, ..., 6 the following probabilities

X	1	2	3	4	5	6
P_0	0.03	0.02	0.02	0.01	0	0.92
P_1	0.06	0.05	0.08	0.02	0.01	0.78
P_2	0.09	0.05	0.12	0.01	0.02	0.71

Consider the test H_0: $P = P_0$ versus H_1: $P = P_1$ or P_2. Determine whether there exists the uniformly most powerful test with size $\alpha = 0.08$. Present the most powerful test if it exists.
Solution: In order to construct a UMP size 0.08 test for H_0: $P = P_0$ versus H_1: $P = P_1$ or P_2, the most powerful size 0.08 test for H_0: $P = P_0$ versus H_1: $P = P_1$ has to be the same as the most powerful size 0.08 test for H_0: $P = P_0$ versus H_1: $P = P_2$.
We start with the Neyman-Pearson test for H_0: $P = P_0$ versus H_1: $P = P_1$. Note that the likelihood ratio is

X	1	2	3	4	5	6
P_0	0.03	0.02	0.02	0.01	0	0.92
P_1	0.06	0.05	0.08	0.02	0.01	0.78
$\frac{P_0}{P_1}$	$\frac{1}{2}$	$\frac{2}{5}$	$\frac{1}{4}$	$\frac{1}{2}$	0	$\frac{92}{78}$

We reject the null hypothesis if the likelihood ratio is too small. The corresponding Neyman-Pearson test of size 0.08 has the rejection region to be $R = \{P_0/P_1 < 92/78\}$. The rejection region of the test is equivalent to $R = \{x : x \neq 6\}$. If it is also the rejection region for the test H_0: $P = P_0$ versus H_1: $P = P_2$, then it is the UMP size 0.08 test for H_0: $P = P_0$ versus H_1: $P = P_1$ or P_2. The likelihood ratio of P_0/P_2 is

X	1	2	3	4	5	6
P_0	0.03	0.02	0.02	0.01	0	0.92
P_2	0.09	0.05	0.12	0.01	0.02	0.71
$\frac{P_0}{P_2}$	$\frac{1}{3}$	$\frac{2}{5}$	$\frac{1}{6}$	1	0	$\frac{92}{71}$

It is easy to observe that the most powerful size 0.08 test has the rejection region $R = \{x : x \neq 6\}$, which is the same as the most powerful test for H_0: $P = P_0$ versus H_1: $P = P_1$. Therefore, the UMP test with size 0.08 has the rejection region $R = \{x : x \neq 6\}$.

11. Assume that X_1, X_2,..., X_n are i.i.d. random variables from a distribution with the probability density function

$$p(x;\theta) = \frac{x}{\theta}\exp\left(-\frac{x^2}{2\theta}\right), \quad x \geq 0,\ \theta > 0.$$

We would like to test $H_0 : \theta = \theta_0$ versus $H_1 : \theta \neq \theta_0$ with $\alpha = 0.05$, where $\theta_0 > 0$. Derive the LRT and its asymptotic distribution. You can pretend all the regularity conditions are satisfied.

Solution: The likelihood function is

$$L(\theta;\boldsymbol{x}) = \frac{1}{\theta^n}\left(\prod_{i=1}^n x_i\right)\exp\left(-\frac{\sum_{i=1}^n x_i^2}{2\theta}\right),$$

and the log-likelihood function is

$$\ell(\theta;\boldsymbol{x}) = -n\log\theta + \sum_{i=1}^n \log x_i - \frac{1}{2\theta}\sum_{i=1}^n x_i^2.$$

Note that

$$\frac{\partial \ell(\theta;\boldsymbol{x})}{\partial\theta} = -\frac{n}{\theta} + \frac{1}{2\theta^2}\sum_{i=1}^n x_i^2.$$

Hence the MLE is $\hat{\theta} = \sum X_i^2/(2n)$. Note that the MLE under H_0 is $\hat{\theta}_0 = \theta_0$. So the likelihood ratio is

$$\begin{aligned}\lambda &= \frac{L(\theta_0;\boldsymbol{x})}{L(\hat{\theta};\boldsymbol{x})} = \frac{\frac{1}{\theta_0^n}\left(\prod_{i=1}^n x_i\right)\exp\left(-\frac{\sum_{i=1}^n x_i^2}{2\theta_0}\right)}{\frac{1}{\hat{\theta}^n}\left(\prod_{i=1}^n x_i\right)\exp\left(-\frac{\sum_{i=1}^n x_i^2}{2\hat{\theta}}\right)} \\ &= \left(\frac{\sum_{i=1}^n x_i^2}{2n\theta_0}\right)^n \exp\left\{n - \sum_{i=1}^n x_i^2/2\theta_0\right\}.\end{aligned}$$

The difference in the number of parameters is 1 here, so $-2\log\lambda(\boldsymbol{X}) \xrightarrow{d} \chi_1^2$. We reject the null hypothesis if $-2\log\lambda(\boldsymbol{x}) > \chi_{1,\alpha}^2$.