Stat 512
Homework key 8

December 12, 2015

REGULAR PROBLEMS

1. Let $Y_1, \dots, Y_n$ be independent identically distributed random variables according to a Rayleigh distribution with parameter $\sigma^2 > 0$ and probability density function

$$f_Y(y|\sigma^2) = \frac{y}{\sigma^2} e^{-\dfrac{y^2}{2\sigma^2}} 1_{(0,\infty)}(x).$$

   (a) Derive the cumulative distribution function F_Y and the cdf F_W for $W = Y^2$.

   (b) Derive $E[Y]$.

   (c) Derive $Var(Y)$.

   (d) Find a method of moments estimator (MME) for σ^2 using the first moment. Find its asymptotic distribution.

   (e) Find a method of moments estimator (MME) for σ^2 using the second moment. Find its asymptotic distribution.

   (f) Find a maximum likelihood estimator (MLE) for σ^2. Find its asymptotic distribution.

   (g) Which of the above estimators is to be preferred? Why?

Ans:

(a)

$$\begin{aligned} F_Y(x) &= \int_0^x \frac{y}{\sigma^2} e^{-\dfrac{y^2}{2\sigma^2}} dy \\ &= \int_0^{x^2/2\sigma^2} e^{-z} dz \\ &= 1 - e^{-x^2/2\sigma^2}. \end{aligned}$$

We have used the change of variable technique, $z = \dfrac{y^2}{2\sigma^2}$ to evaluate the integral.

Let $W = Y^2$.

$$\begin{aligned} f_W(w) &= \frac{f(\sqrt{w})}{2\sqrt{w}} \\ &= \frac{e^{-w/2\sigma^2}}{2\sigma^2} \end{aligned}$$

Hence, $W \sim \mathcal{E}(2\sigma^2)$ (Mean parametrization).

(b)

$$EY = \int_0^\infty \frac{y^2}{\sigma^2} e^{-\dfrac{y^2}{2\sigma^2}} dy.$$

We will use the change of variable $z = \dfrac{y^2}{2\sigma^2}$ again to get

$$\begin{aligned} EY &= \sqrt{2}\sigma \int_0^\infty \sqrt{z} e^{-z} dz \\ &= \sqrt{2}\sigma\Gamma(3/2) \int_0^\infty \frac{1^{3/2} z^{3/2-1} e^{-z}}{\Gamma(3/2)} dz. \end{aligned}$$

Now the integrand is the PDF of $Gamma(3/2, 1)$. Hence it will integrate to 1. Since $\Gamma(3/2) = \dfrac{\sqrt{\pi}}{2}$, we will be left with

$$EY = \sigma\sqrt{\frac{\pi}{2}}.$$

(c) First we will find EY^2.

$$\begin{aligned} EY^2 &= \int_0^\infty \frac{y^3}{\sigma^2} e^{-\frac{y^2}{2\sigma^2}} dy \\ &= 2\sigma^2 \int_0^\infty z e^{-z} dz \\ &= 2\sigma^2 \Gamma(2) \int_0^\infty \frac{z^{2-1}e^{-z}}{\Gamma(2)} dz \\ &= 2\sigma^2 \end{aligned}$$

Hence,

$$Var(Y) = EY^2 - (EY)^2 = 2\sigma^2 - \sigma^2\frac{\pi}{2} = \frac{4-\pi}{2}\sigma^2$$

(d) Equating the first sample moment with the first population moment we get,

$$\begin{aligned} \sqrt{\frac{\pi}{2}}\sigma &= \bar{Y}. \\ \Rightarrow \hat{\sigma}^2 &= \frac{2\bar{Y}^2}{\pi} \end{aligned}$$

We will use the CLT and Delta method to find the asymptotic distribution.

$$\sqrt{n}\left(\bar{Y} - \sqrt{\frac{\pi}{2}}\sigma\right) \to_d \mathcal{N}\left(0, \frac{4-\pi}{2}\sigma^2\right)$$

Take the monotonic transformation $g(x) = x^2$. We get

$$\sqrt{n}\left(\bar{Y}^2 - \frac{\pi}{2}\sigma^2\right) \to_d \mathcal{N}\left(0, \pi(4-\pi)\sigma^4\right)$$

Multiplying the random variable by $2/\pi$,

$$\sqrt{n}\left(\frac{2}{\pi}\bar{Y}^2 - \sigma^2\right) \to_d \mathcal{N}\left(0, \frac{4}{\pi}(4-\pi)\sigma^4\right).$$

Therefore,

$$\sqrt{n}\left(\hat{\sigma}^2 - \sigma^2\right) \to_d \mathcal{N}\left(0, \frac{4}{\pi}(4-\pi)\sigma^4\right). \tag{1}$$

(e) Equating the second raw moments from the sample and the population we get that,

$$\begin{aligned} \sum_{i=1}^n \frac{Y_i^2}{n} &= EY^2 = 2\sigma^2 \\ \Rightarrow \hat{\sigma}^2 &= \frac{\sum_{i=1}^n Y_i^2}{2n} \end{aligned}$$

Since $W_i = Y_i^2$-s are i.i.d. exponential random variables with mean $2\sigma^2$, and hence variance $4\sigma^4$, applying CLT we get,

$$\begin{aligned} \sqrt{n}\left(\sum_{i=1}^n \frac{Y_i^2}{n} - 2\sigma^2\right) &\to_d \mathcal{N}\left(0, 4\sigma^4\right) \\ \sqrt{n}\left(\sum_{i=1}^n \frac{Y_i^2}{2n} - \sigma^2\right) &\to_d \mathcal{N}\left(0, \sigma^4\right). \end{aligned}$$

Therefore,

$$\sqrt{n}\left(\hat{\sigma}^2 - \sigma^2\right) \to_d \mathcal{N}\left(0, \sigma^4\right). \tag{2}$$

(f)

$$\mathcal{L}(\sigma^2|y) = \prod_{i=1}^n \frac{y_i e^{-y_i^2/2\sigma^2}}{\sigma^2}$$

$$\Rightarrow l(\sigma^2|y) = \sum_{i=1}^n \log y_i - \sum_{i=1}^n \frac{y_i}{2\sigma^2} - n\log\sigma^2.$$

d.w.r.t. σ^2 and setting the derivative to 0 we get,

$$0 = \frac{\partial l}{\partial \sigma^2} = \frac{\sum_{i=1}^n y_i^2}{2\sigma^4} - \frac{n}{\sigma^2} = 0$$

$$\Rightarrow \frac{\sum_{i=1}^n y_i^2}{2\sigma^2} = n$$

$$\Rightarrow \hat{\sigma}^2 = \frac{\sum_{i=1}^n y_i^2}{2n}$$

Hence this $\hat{\sigma}^2$ is same as the previous one. Therefore from (2) we get that its asymptotic distribution is

$$\sqrt{n}\left(\hat{\sigma}^2 - \sigma^2\right) \to_d \mathcal{N}\left(0, \sigma^4\right). \tag{3}$$

(g) From (1), (2), (3) we find that the asymptotic variances of the MME estimators from the 1st, 2nd moments and the MLE are $\frac{4}{\pi}(4-\pi)\sigma^4$, σ^4, σ^4 respectively. Now $\pi < 3.15$. Therefore, $\frac{4}{\pi}(4-\pi) > 1$. Hence the MLE and the MME estimator from the second moment (both are the same estimator) have lower asymptotic variance. All the estimators are unbiased. Therefore, the MLE or the MME estimator from the second moment, i.e. $\frac{\sum_{i=1}^n Y_i^2}{2n}$ must be preferred.

2. Let $Y_1, \ldots, Y_n$ be independent identically distributed random variables according to a one parameter exponential family with parameter θ. For each of the following distributions, find an MME $\tilde{\theta}$ of θ and its asymptotic distribution, as well as the MLE $\hat{\theta}$ of θ and its asymptotic distribution.

(a) Bernoulli distribution: $Y_i \sim \mathcal{B}(1, \theta)$.

Ans: We have that $E[Y_i] = \theta$, so the natural MME is just the sample mean

$$\tilde{\theta} = \bar{Y}_n.$$

The MLE is the solution of the score equation, given by

$$\sum_{i=1}^n \frac{\partial}{\partial\theta} \log f_Y(y_i|\theta) = 0.$$

Here, we have that

$$\begin{aligned}
\sum_{i=1}^n \frac{\partial}{\partial\theta} \log f_Y(y_i|\theta) &= \sum_{i=1}^n \frac{\partial}{\partial\theta} \log[\theta^{y_i}(1-\theta)^{1-y_i}] \\
&= \sum_{i=1}^n \frac{\partial}{\partial\theta}[y_i \log\theta + (1-y_i)\log(1-\theta)] \\
&= \sum_{i=1}^n \left(\frac{y_i}{\theta} - \frac{1-y_i}{1-\theta}\right) \\
&= \sum_{i=1}^n \frac{1}{\theta(1-\theta)}(y_i - \theta) = 0.
\end{aligned}$$

The solution to the score equation is seen to be $\hat{\theta} = \bar{Y}_n$, so that the MME and MLE are equivalent. Invoking the CLT, the asymptotic sampling distribution for either estimator is

$$\sqrt{n}(\bar{Y}_n - \theta) \to_d \mathcal{N}(0, Var[Y_i] = \theta(1-\theta)).$$

(b) Poisson distribution: $Y_i \sim \mathcal{P}(\theta)$.

Ans: We have that $E[Y_i] = \theta$, so the natural MME is just the sample mean

$$\tilde{\theta} = \bar{Y}_n.$$

The MLE is the solution of the score equation, given by

$$\sum_{i=1}^{n} \frac{\partial}{\partial \theta} \log f_Y(y_i|\theta) = 0.$$

Here, we have that

$$\begin{aligned}
\sum_{i=1}^{n} \frac{\partial}{\partial \theta} \log f_Y(y_i|\theta) &= \sum_{i=1}^{n} \frac{\partial}{\partial \theta} \log[(y_i!)^{-1} e^{-\theta} \theta^{y_i}] \\
&= \sum_{i=1}^{n} \frac{\partial}{\partial \theta}[-\log(y_i!) - \theta + y_i \log \theta] \\
&= \sum_{i=1}^{n} (-1 + y_i/\theta) \\
&= \sum_{i=1}^{n} \frac{1}{\theta}(y_i - \theta) = 0.
\end{aligned}$$

The solution to the score equation is seen to be $\hat{\theta} = \bar{Y}_n$, so that the MME and MLE are equivalent. Invoking the CLT, the asymptotic sampling distribution for either estimator is

$$\sqrt{n}(\bar{Y}_n - \theta) \to_d \mathcal{N}(0, Var[Y_i] = \theta).$$

(c) Exponential distribution: $Y_i \sim \theta$ with $F_Y(y) = (1 - e^{-y/\theta})\mathbf{1}_{(0,\infty)}(y)$.

Ans: We have that $E[Y_i] = \theta$, so the natural MME is just the sample mean

$$\tilde{\theta} = \bar{Y}_n.$$

The MLE is the solution of the score equation, given by

$$\sum_{i=1}^{n} \frac{\partial}{\partial \theta} \log f_Y(y_i|\theta) = 0.$$

Here, we have that

$$\begin{aligned}
\sum_{i=1}^{n} \frac{\partial}{\partial \theta} \log f_Y(y_i|\theta) &= \sum_{i=1}^{n} \frac{\partial}{\partial \theta} \log[\theta^{-1} e^{-y_i/\theta}] \\
&= \sum_{i=1}^{n} \frac{\partial}{\partial \theta}[-\log \theta - y_i/\theta] \\
&= \sum_{i=1}^{n} \left(-\frac{1}{\theta} + \frac{y_i}{\theta^2}\right) \\
&= \sum_{i=1}^{n} \frac{1}{\theta^2}(y_i - \theta) = 0.
\end{aligned}$$

The solution to the score equation is seen to be $\hat{\theta} = \bar{Y}_n$, so that the MME and MLE are equivalent. Invoking the CLT, the asymptotic sampling distribution for either estimator is

$$\sqrt{n}(\bar{Y}_n - \theta) \to_d \mathcal{N}(0, Var[Y_i] = \theta^2).$$

Note this pattern: for estimation of the mean in each of these distributions, the MLE solved an estimating equation of the form

$$\sum_{i=1}^{n} \frac{1}{Var[Y_i]}(y_i - E[Y_i]) = 0,$$

where $E[Y_i] = \theta$ and $Var[Y_i] = h(\theta)$. Question 5 below expands on this result.

3. For each of the estimators found in problems 1 and 2, there was a mean-variance relationship: The variance of the asymptotic distribution of our estimator involved the unknown parameter:

$$\sqrt{n}(T_n(\overrightarrow{Y}) - \theta) \to_d \mathcal{N}(0, V(\theta)).$$

This sometimes creates a problem when trying to find confidence intervals for θ based on the asymptotic distribution. One approach that is sometimes used to use the delta method to find a "variance stabilizing transformation" $g_{F,\hat{\theta}}(t)$ for each distribution function F and estimator $\hat{\theta}$ such that

$$\sqrt{n}(g(T_n) - g(\theta)) \to_d \mathcal{N}(0, V^*),$$

where V^*) is independent of θ.

(a) Find a variance stabilizing transformation for the estimator found in problem 1d.

(b) Find a variance stabilizing transformation for the estimator found in problem 1e.

(c) Find a variance stabilizing transformation for the estimator found in problem 1f.

(d) Find a variance stabilizing transformation for the estimator found in problem 2a.

(e) Find a variance stabilizing transformation for the estimator found in problem 2b.

(f) Find a variance stabilizing transformation for the estimator found in problem 2c.
Note: The use of these variance stabilizing transformations will at times help us in the following manner:

- *We use the resulting asymptotic normal distribution to construct an approximate distribution for* $g(T_n)$*:*

$$g(T_n) \sim \mathcal{N}\,\mathrm{lg}(\theta), \frac{V^*}{n}\Big)$$

- *Based on observation* $T_n = t$*, we construct a confidence interval for* $g(\theta)$ *using the usual approach for a normal distribution:*

$$100(1-\alpha)\%CI\ \ for\ g(\theta) = (L, U) = \left(g(t) - z_{1-\alpha/2}\sqrt{\frac{V^*}{n}}, g(t) + z_{1-\alpha/2}\sqrt{\frac{V^*}{n}}\right).$$

- *Then for invertible* g*, we compute*

$$100(1-\alpha)\%CI\ \ for\ \theta = (g^{-1}(L), g^{-1}(U)).$$

Ans:

(a) We have to find a g such that the asymptotic variance of $g(T_n)$, i.e. $V(\sigma^2)g'(\sigma^2)^2 = c$ where c is some constant. Now from (1) we find that $V(\sigma^2) = \frac{4}{\pi}(4-\pi)\sigma^4$. Therefore we need $g'(\sigma^2)^2\sigma^4 = c$. Let us take $c = 1$. Hence,

$$\begin{aligned} g'(\sigma^2)^2 &= \frac{1}{\sigma^4} \\ \therefore g'(\sigma^2) &= \frac{1}{\sigma^2} \\ \therefore g(\sigma^2) &= \log(\sigma^2) \end{aligned}$$

With this transformation,

$$\sqrt{n}(\log\hat{\sigma}^2 - \log\sigma^2) \to_d \mathcal{N}\left(0, \frac{4}{\pi}(4-\pi)\right)$$

where $\hat{\sigma}^2 = \dfrac{2\bar{Y}^2}{\pi}$.

(b) From (2) we find that $V(\sigma^2) = \sigma^4$. Therefore we need $g'(\sigma^2)^2\sigma^4 = c$. So the g in problem 3.(a) will work. Hence, $g = \log$. With this transformation,

$$\sqrt{n}(\log\hat{\sigma}^2 - \log\sigma^2) \to_d \mathcal{N}\Big(0, 1\Big)$$

where $\hat{\sigma}^2 = \dfrac{\sum_{i=1}^n Y_i^2}{2n}$.

(c) The estimator of 1f is same as 1e. Therefore the transformation will again be the log transformation and

$$\sqrt{n}(\log \hat{\sigma}^2 - \log \sigma^2) \to_d \mathcal{N}\Big(0, 1\Big)$$

where $\hat{\sigma}^2 = \dfrac{\sum_{i=1}^n Y_i^2}{2n}$.

(d) The estimator found in 2a is $\bar{Y}$ where

$$\sqrt{n}(\bar{Y} - \theta) \to_d \mathcal{N}(0, \theta(1-\theta)).$$

We need to find a function g such that $g'(\theta)^2 = \dfrac{c}{\theta(1-\theta)}$. Let us take $c = 1$.

$$g'(\theta) = \frac{1}{\sqrt{\theta(1-\theta)}}$$
$$g(\theta) = \int \frac{d\theta}{\sqrt{\theta(1-\theta)}}.$$

Take change of variable $\theta = \dfrac{1+w}{2}$ to get

$$g(\theta) = \int \frac{dw}{\sqrt{1-w^2}} = \arcsin w + d = \arcsin(2\theta - 1) + d$$

where d is a constant for integration. Hence, we can use $g(\theta) = \arcsin(2\theta - 1)$. Applying Delta method we derive,

$$\sqrt{n}(\arcsin(2\bar{Y} - 1) - \arcsin(2\theta - 1)) \to_d \mathcal{N}(0, 1).$$

(e) The estimator in 2b is $\bar{Y}$ where

$$\sqrt{n}(\bar{Y} - \theta) \to_d \mathcal{N}(0, \theta)$$

So, $g'(\theta) = \dfrac{c}{\sqrt{\theta}}$. Let us take $c = 1$. Integrating $g'(\theta)$ we get $g(\theta) = 2\sqrt{\theta}$. Taking this transformation,

$$\sqrt{n}(2\sqrt{\bar{Y}} - 2\sqrt{\theta}) \to_d \mathcal{N}(0, 1).$$

(f) The estimator in 2c is $\bar{Y}$ where

$$\sqrt{n}(\bar{Y} - \theta) \to_d \mathcal{N}(0, \theta^2).$$

So, $g'(\theta) = \dfrac{c}{\theta}$. Let us take $c = 1$. Integrating $g'(\theta)$ we get $g(\theta) = \log(\theta)$. Taking this transformation we get,

$$\sqrt{n}(\log(\bar{Y}) - \log(\theta)) \to_d \mathcal{N}(0, 1).$$

4. Let $X_1, X_2, \dots$ be a sequence of i.i.d. random variables having mean $\mu > 0$ and variance $\sigma^2 > 0$, and let $Y_1, Y_2, \dots$ be a sequence of i.i.d. random variables having mean $\nu > 0$ and variance $\tau^2 > 0$. Further suppose that X_i and Y_j are independent if $i \neq j$, but that the correlation between X_i and Y_j is ρ if $i = j$. For notational convenience, denote $\vec{X}_n = (X_1, \dots, X_n)$ and $\vec{Y}_n = (Y_1, \dots, Y_n)$.

(a) Find a method of moments estimator $\hat{\theta}_n = \hat{\theta}_n(\vec{X}_n, \vec{Y}_n)$ for $\theta = \mu - \nu$ and derive its asymptotic distribution. Is your estimator unbiased? Consistent?

Ans: We first note that we can consider our sample as i.i.d. observations of the random vector

$$\begin{pmatrix} X_i \\ Y_i \end{pmatrix} \sim \left(\vec{\alpha} = \begin{pmatrix} \mu \\ \nu \end{pmatrix}, \Sigma = \begin{pmatrix} \sigma^2 & \rho\sigma\tau \\ \rho\sigma\tau & \tau^2 \end{pmatrix}\right).$$

An MME in terms of this random vector for its mean is given by

$$\hat{\vec{\alpha}} = \begin{pmatrix} \bar{X}_n \\ \bar{Y}_n \end{pmatrix}.$$

As a consequence of the multivariate CLT, we have that

$$\sqrt{n}(\hat{\vec{\alpha}} - \vec{\alpha}) \to_d \mathcal{N}_2(0, \Sigma).$$

We now take the function $g(u,v) = u-v$, having gradient $\nabla g(u,v) = (1,-1)^T$ and let $\hat{\theta} = g(\bar{X}_n, \bar{Y}_n)$. Application of the delta method yields

$$\sqrt{n}(g(\hat{\vec{\alpha}}) - g(\vec{\alpha}) \to_d \mathcal{N}_1\left(0, \nabla g(u,v)^T \Sigma \nabla g(u,v)\right).$$

This reduces to the asymptotic sampling distribution,

$$\sqrt{n}((\bar{X}_n - \bar{Y}_n) - (\mu - \nu)) \to_d \mathcal{N}_1(0, \sigma^2 + \tau^2 - 2\rho\sigma\tau).$$

Note that this estimator and its asymptotic sampling distribution could also be derived by considering the variable $W_i = X_i - Y_i$ (we might call this a paired difference). We have that $E[W_i] = \mu - \nu$ and $Var[W_i] = \sigma^2 + \tau^2 - 2\rho\sigma\tau$, so for MME given by $\bar{W}_n = \bar{X}_n - \bar{Y}_n$, the univariate CLT yields

$$\sqrt{n}(\bar{W}_n - (\mu - \nu)) \to_d \mathcal{N}_1(0, \sigma^2 + \tau^2 - 2\rho\sigma\tau).$$

Regardless of which route was used to derive it, the consistency of $\hat{\theta}$ follows from the WLLN, and its unbiasedness follows from linearity of expectations. Parts (b) and (c) below demonstrate that the two methods are not always equivalent. Note in particular that the contrast (transformation) $g(u,v) = u - v$ in this problem was linear, while in (b) and (c), the contrast was nonlinear.

(b) Find a method of moments estimator $\hat{\psi}_n = \hat{\psi}_n(\vec{X}_n, \vec{Y}_n)$ for $\psi = \mu/\nu$ and derive its asymptotic distribution. Is your estimator unbiased? Consistent?

Ans: Recall from (a) that considering our sample as i.i.d. random vectors with mean $\vec{\alpha}$, that

$$\sqrt{n}(\hat{\vec{\alpha}} - \vec{\alpha}) \to_d \mathcal{N}_2(0, \Sigma).$$

We now take the function $g(u,v) = u/v$, having gradient $\nabla g(u,v) = (1/v, -u/v^2)^T$ and let $\hat{\psi} = g(\bar{X}_n, \bar{Y}_n)$. Application of the delta method yields

$$\sqrt{n}(g(\hat{\vec{\alpha}}) - g(\vec{\alpha}) \to_d \mathcal{N}_1\left(0, \nabla g(u,v)^T \Sigma \nabla g(u,v)\right).$$

This reduces to the asymptotic sampling distribution,

$$\sqrt{n}(\hat{\psi} - \psi) = \sqrt{n}((\bar{X}_n/\bar{Y}_n) - (\mu/\nu)) \to_d \mathcal{N}_1\left(0, \frac{\mu^2\tau^2}{\nu^2} + \frac{\sigma^2}{\nu^2} - \frac{2\mu\rho\sigma\tau}{\nu^3}\right) = \mathcal{N}_1\left(0, \psi^2\tau^2 + \nu^{-2}\sigma(\sigma - 2\psi\tau)\right).$$

The Mann-Wald (continuous mapping theorem) yields the consistency of $\hat{\psi}$ for ψ. It is not generally the case that $\hat{\psi}$ will be unbiased. The counterexample for (c) demonstrates this.

(c) Define $W_i = X_i/Y_i$. Let $\gamma = E[W_i]$. Show that a MME estimator $\hat{\gamma}_n = \hat{\gamma}_n(\vec{X}_n, \vec{Y}_n)$ of γ is in general a biased, inconsistent estimator for ψ.

Ans: Let $X_i \sim \mathcal{B}(1,p)$ and $Y_i \sim \mathcal{B}(1,p) + 1$ be independent random variables ($\rho = 0$). Then $EX_i = \mu = p > 0$ and $EY_i = \nu = 1 + p > 0$. We have that

$$\psi = \frac{\mu}{\nu} = \frac{p}{1+p}.$$

Now, for $W_i = X_i/Y_i$ we have

$$\gamma = EW_i = E\frac{X_i}{Y_i} = 0 + \frac{1}{2}p^2 + p(1-p) = p(1 - p/2).$$

In order for $\gamma = \psi$, we find that p must be either 0 or 1. For instance, if $p = 1/2$, $\gamma = 1/3 \neq 3/8 = \psi$. For this example, we have that $\bar{W}_n$ is unbiased and consistent for γ but clearly not for ψ.

MORE INVOLVED PROBLEMS

5. (This problem draws heavily on problems 3 and 4 on Homework 3.) Let $\vec{X} = (X_1, \ldots, X_n)$ be a random vector in which the X_i are independently distributed with a one exponential family distribution having canonical parameter η and density (probability mass function) of the form

$$f_X(x|\eta) = h(x)\exp[T(X)\eta - A(\eta)].$$

Assume that we may twice interchange integration of f_X with respect to x and differentiation of f_X with respect to η and that $A(\eta)$ is twice differentiable with invertible derivatives.

(a) Find the asymptotic distribution of the maximum likelihood estimate $\hat{\eta}$ of η.

Ans: We defined the efficient score in Homework 3, Problem 4, as

$$U(X) = \frac{\partial}{\partial \eta} \log f_X(X|\eta).$$

This is directly related to the derivative of the log likelihood, $\dot{\ell}$, for a sample $\vec{X}$ by

$$\dot{\ell}(\eta|\vec{X}) = \sum_{i=1}^{n} U(x_i).$$

For the one parameter exponential family under consideration, we have (denoting $A'(x)$ by $\dot{A}(x)$) that

$$U(X) = T(X) - \dot{A}(\eta).$$

Hence an MLE for $\dot{A}(\eta)$ is given by solving,

$$\sum_{i=1}^{n}[T(X_i) - \dot{A}(\eta)] = 0.$$

This yields that the MLE for $\dot{A}(\eta)$ is just the mean $\bar{T}_n$. From Homework 3, Problem 4, we see that $E[U(X)] = 0$. This implies that

$$E[T(X)] = \dot{A}(\eta).$$

An additional result was that $Var[U(X)] = -E\left[\frac{\partial}{\partial \eta} U(X)\right]$. This implies that

$$Var[T(X)] = Var[T(X) - \dot{A}(\eta)] = Var[U(X)] = \ddot{A}(\eta).$$

These results, along with the CLT, yield the asymptotic sampling distribution

$$\sqrt{n}(\bar{T}_n - \dot{A}(\eta)) \to_d \mathcal{N}(0, \ddot{A}(\eta)).$$

Since $\dot{A}$ is invertible by assumption, we can apply the invariance of the MLE to arrive at the conclusion that the MLE, $\hat{\eta} = \dot{A}^{-1}(\bar{T}_n)$. We conclude from the delta method that the asymptotic sampling distribution of $\hat{\eta}$ is

$$\sqrt{n}(\hat{\eta} - \eta) \to_d \mathcal{N}\left(0, \frac{\ddot{A}(\eta)}{\ddot{A}(\eta)^2}\right) = \mathcal{N}\left(0, \frac{1}{\ddot{A}(\eta)}\right).$$

(b) Find the asymptotic distribution of the MLE $\hat{\theta}$ of $\theta = E[X]$ if $\eta = g(\theta)$.

Ans: We assume that g is differentiable with an invertible derivative. Given this assumption, we have by the invariance of MLEs that

$$\hat{\theta} = g^{-1}(\hat{\eta}).$$

To calculate the asymptotic distribution, we appeal again to the delta method.

$$\sqrt{n}(\hat{\theta} - \theta) \to_d \mathcal{N}\left(0, \frac{1}{\ddot{A}(\eta)\dot{g}(g^{-1}(\eta))^2}\right) = \mathcal{N}\left(0, \frac{1}{\ddot{A}(g(\theta))\dot{g}(\theta)^2}\right).$$