# MATH3911 Assignment 1

Victor Chan

This assignment is my own work. I have read and understood the University Rules in respect to Student Academic Misconduct.

1. (a) Now $X_1, ..., X_n$ are i.i.d. Poisson ($\lambda$) random variables, so $\mathrm{E}_\lambda[X] = \lambda$ and their common probability function is

$$f(x;\lambda) = e^{-\lambda}\frac{\lambda^x}{x!} \qquad \text{for x } \in \{0,1,2,...\},$$

Then the likelihood function is

$$\begin{aligned} L(\mathbf{x},\lambda) &= \prod_{i=1}^{n} f(x_i;\lambda) \qquad \text{as } X_1,...,X_n \text{ are i.i.d} \\ &= e^{-n\lambda}\frac{\lambda^{\sum_{i=1}^{n} x_i}}{\prod_{i=1}^{n}(x_i)!} \end{aligned}$$

and the score function is

$$\begin{aligned} V(\mathbf{X},\lambda) &= \frac{\partial}{\partial\lambda}\log L(\mathbf{X},\lambda) \\ &= \frac{\partial}{\partial\lambda}\left(-n\lambda + \log\lambda\sum_{i=1}^{n} X_i - \sum_{i=1}^{n}\log(X_i)!\right) \\ &= -n + \frac{1}{\lambda}\sum_{i=1}^{n} X_i, \end{aligned}$$

so the Fisher information about $\lambda$ in $\mathbf{X}$ is

$$\begin{aligned} I_{\mathbf{X}}(\lambda) &= \mathrm{E}_\lambda\left[-\frac{\partial}{\partial\lambda}V(\mathbf{X},\lambda)\right] \\ &= \mathrm{E}_\lambda\left[\frac{1}{\lambda^2}\sum_{i=1}^{n} X_i\right] \\ &= \frac{1}{\lambda^2}\sum_{i=1}^{n}\mathrm{E}_\lambda[X_i] \\ &= \frac{n\lambda}{\lambda^2} = \frac{n}{\lambda}. \end{aligned}$$

Hence the Cramer-Rao lower bound for the variance of an unbiased estimator of $\tau(\lambda) = \lambda e^{-\lambda}$ is

$$\begin{aligned} \frac{\{\frac{\partial}{\partial\lambda}\tau(\lambda)\}^2}{I_{\mathbf{X}}(\lambda)} &= \frac{[e^{-\lambda}(1-\lambda)]^2}{n/\lambda} \\ &= \frac{\lambda e^{-2\lambda}(1-\lambda)^2}{n}. \end{aligned}$$

Suppose that $T(\mathbf{X})$ is an unbiased estimator of $\tau(\lambda)$ which attains this variance bound. Then the score function $V(\mathbf{X},\lambda)$ can be expressed as $k_n(\lambda)[T(\mathbf{X})-\tau(\lambda)]$, where $k_n(\lambda)$ is independent of $\mathbf{X}$. Now

$$\begin{aligned} V(\mathbf{X},\lambda) &= \frac{1}{\lambda}\sum_{i=1}^{n} X_i - n \\ &= \frac{n}{\lambda}(\bar{X}-\lambda) \\ &= \frac{ne^{\lambda}}{\lambda}(\bar{X}e^{-\lambda}-\lambda e^{-\lambda}) \\ &= \frac{ne^{\lambda}}{\lambda}[\bar{X}e^{-\lambda}-\tau(\lambda)] \end{aligned}$$

which implies that $T(\mathbf{X})=\bar{X}e^{-\lambda}$, but this is not an estimator as it depends on $\lambda$. Hence there is no unbiased estimator of $\tau(\lambda)$ which attains the Cramer-Rao lower bound.

(b) Consider the estimator $W(\mathbf{X}) = I_{\{X_1=1\}}(\mathbf{X})$. Now

$$\mathrm{E}_\lambda[W(\mathbf{X})] = \Pr_\lambda[X_1=1] = \lambda e^{-\lambda} = \tau(\lambda)$$

so $W(\mathbf{X})$ is an unbiased estimator of $\tau(\lambda)$.

Also

$$f(x;\lambda) = \frac{e^{-\lambda}}{x!}e^{x\log\lambda} = a(\lambda)b(x)e^{c(\lambda)d(x)}$$

where $a(\lambda)=e^{-\lambda}, b(x)=\frac{1}{x!}, c(\lambda)=\log\lambda$ and $d(x)=x$. Then the family of distributions $\{f(x;\lambda)\,|\,\lambda>0\}$ is a one-paramter exponential family, so the statistic $T(\mathbf{X})=\sum_{i=1}^n d(X_i)=\sum_{i=1}^n X_i$ is complete and sufficient.

Now $T(\mathbf{X})$ ~Poisson$(n\lambda)$ and $\sum_{i=2}^n X_i$ ~Poisson$((n-1)\lambda)$, so

$$\begin{aligned} \mathrm{E}\left[W(\mathbf{X})\,|\,T(\mathbf{X})=t\right] &= \Pr[X_1=1\,|\,T(\mathbf{X})=t] \\ &= \frac{\Pr_\lambda[X_1=1,\sum_{i=1}^n X_i=t]}{\Pr_\lambda[T(\mathbf{X})=t]} \\ &= \frac{\Pr_\lambda[X_1=1,\sum_{i=2}^n X_i=t-1]}{\Pr_\lambda[T(\mathbf{X})=t]} \\ &= \frac{\Pr_\lambda[X_1=1]\Pr_\lambda[\sum_{i=2}^n X_i=t-1]}{e^{-n\lambda}(n\lambda)^t(t!)^{-1}} \\ &\quad \text{as } X_1,\ldots,X_n \text{ are independent} \\ &= \frac{[e^{-\lambda}\lambda][e^{-\lambda(n-1)}(\lambda(n-1))^{t-1}((t-1)!)^{-1}]}{e^{-n\lambda}(n\lambda)^t(t!)^{-1}} \\ &= \frac{t!\,e^{-n\lambda}\,\lambda^t\,(n-1)^{t-1}}{(t-1)!\,e^{-n\lambda}\,\lambda^t\,n^t} \\ &= \frac{t}{n-1}\left(\frac{n-1}{n}\right)^t = \frac{t}{n-1}\left(1-\frac{1}{n}\right)^t. \end{aligned}$$

Then by the Lehmann-Scheffe Theorem, the UMVUE is

$$\tilde{\tau}(\mathbf{X}) = \mathrm{E}\left[W(\mathbf{X}) \,|\, T(\mathbf{X})\right] = \frac{T(\mathbf{X})}{n-1}\left(1-\frac{1}{n}\right)^{T(\mathbf{X})} = \frac{\sum_{i=1}^n X_i}{n-1}\left(1-\frac{1}{n}\right)^{\sum_{i=1}^n X_i}.$$

(c) Let the MLE of $\lambda$ be $\hat{\lambda}(\mathbf{X})$. Then

$$\begin{aligned} 0 &= V(\mathbf{X}, \hat{\lambda}(\mathbf{X})) \\ &= \frac{\sum_{i=1}^n X_i}{\hat{\lambda}(\mathbf{X})} - n \\ &= n\left(\frac{\bar{X}}{\hat{\lambda}(\mathbf{X})} - 1\right) \\ 0 &= \frac{\bar{X}}{\hat{\lambda}(\mathbf{X})} - 1 \end{aligned}$$

so $\hat{\lambda}(\mathbf{X}) = \bar{X}$.

Let the MLE of $\tau(\lambda)$ be $\hat{\tau}(\mathbf{X})$. By the invariance property of MLEs,

$$\hat{\tau}(\mathbf{X}) = \tau\left(\hat{\lambda}(\mathbf{X})\right) = \tau(\bar{X}) = \bar{X}e^{-\bar{X}}.$$

As the MLE is asymptotically normal, unbiased and efficient, the asymptotic distribution of $\sqrt{n}\,(\hat{\tau}(\mathbf{X}) - \tau(\lambda))$ is $N\left(0,\, \lambda e^{-2\lambda}(1-\lambda)^2\right)$, where

$$\lambda e^{-2\lambda}(1-\lambda^2) = \frac{\left[\frac{\partial}{\partial\lambda}\tau(\lambda)\right]^2}{I_{X_1}(\lambda)} = \frac{\left[\frac{\partial}{\partial\lambda}\tau(\lambda)\right]^2}{\frac{1}{n}I_{\mathbf{X}}(\lambda)}.$$

(d) For the given data, $n = 25$ and $\sum_{i=1}^n x_i = 76$, so $\bar{x} = 3.04$. Then the UMVUE is $\tilde{\tau}(\mathbf{x}) = \frac{76}{24}\left(1-\frac{1}{25}\right)^{76} = 0.14230$ and the MLE is $\hat{\tau}(\mathbf{x}) = 3.04(e^{-3.04}) = 0.14542$. The two numerical values are very close to one another, which would be expected as

$$\begin{aligned} \lim_{n\to\infty} \tilde{\tau}(\mathbf{X}) &= \lim_{n\to\infty} \frac{n\bar{X}}{n-1}\left(1-\frac{1}{n}\right)^{n\bar{X}} \\ &= \bar{X}\left[\lim_{n\to\infty}\frac{n}{n-1}\right]\left[\lim_{n\to\infty}\left(1-\frac{1}{n}\right)^{n\bar{X}}\right] \\ &= \bar{X}e^{-\bar{X}} \\ &= \hat{\tau}(\mathbf{X}), \end{aligned}$$

which shows that the UMVUE approaches the MLE asymptotically.

(e) The asymptotic distribution of $\sqrt{n}\,(\hat{\tau}(\mathbf{X}) - \tau(\lambda))$ is $N\left(0,\, \lambda e^{-2\lambda}(1-\lambda)^2\right)$, so the distribution of $\hat{\tau}(\mathbf{X})$ is approximately $N\left(\tau(\lambda),\, \frac{1}{n}\lambda e^{-2\lambda}(1-\lambda)^2\right)$. Then the standard error of $\hat{\tau}(\mathbf{X})$ is

$$\begin{aligned} \hat{\mathrm{se}}\left[\hat{\tau}(\mathbf{X})\right] &\approx \sqrt{\frac{1}{n}\hat{\lambda}(\mathbf{x})\, e^{-2\hat{\lambda}(\mathbf{x})}\, (1-\hat{\lambda}(\mathbf{x}))^2} \\ &= \sqrt{\frac{1}{n}\bar{x}e^{-2\bar{x}}(1-\bar{x})^2} \\ &= \sqrt{\frac{1}{25}(3.04)e^{-2(3.04)}(1-3.04)^2} \\ &= 0.03403 \end{aligned}$$

so an asymptotic 90% confidence interval for $\tau(\lambda)$ is

$$\begin{aligned}
&\left(\hat{\tau}(\mathbf{x}) - \Phi^{-1}(0.95)\hat{\mathrm{se}}\left[\hat{\tau}(\mathbf{X})\right],\ \hat{\tau}(\mathbf{x}) + \Phi^{-1}(0.95)\hat{\mathrm{se}}\left[\hat{\tau}(\mathbf{X})\right]\right) \\
= &\ (0.14542 - 1.645(0.03403),\ 0.14542 + 1.645(0.03403)) \\
= &\ (0.08944,\ 0.20139),
\end{aligned}$$

where $\Phi$ is the standard normal distribution function.

2. Let $g$ be a function from $\{1, 2, 3, 4\}$ to $\mathbb{R}$. Suppose that $X$ follows Distribution I and for all $\theta \in \Theta = (0, 0.1)$, $\mathrm{E}_\theta[g(X)] = 0$.

Then for all $\theta \in (0, 0.1)$,

$$\begin{aligned}
0 &= \sum_{x=1}^{4} g(x) \Pr_\theta[X = x] \\
&= \theta\, g(1) + (\theta - \theta^4)\, g(2) + (2\theta^3 - \theta^4)\, g(3) + (1 + 2\theta^4 - 2\theta^3 - 2\theta)\, g(4) \\
&= g(4) + \theta\,[g(1) + g(2) - 2g(4)] + \theta^3[2g(3) - 2g(4)] + \theta^4[2g(4) - g(2)],
\end{aligned}$$

so on equating coefficients,

$$\begin{aligned}
g(4) &= 0 \qquad (1) \\
g(1) + g(2) - 2g(4) &= 0 \qquad (2) \\
2g(3) - 2g(4) &= 0 \qquad (3) \\
2g(4) - g(2) &= 0 \qquad (4).
\end{aligned}$$

Now from equation (1), $g(4) = 0$, so from equation (3), $g(3) = g(4) = 0$, and from equation (4), $g(2) = 2g(4) = 0$. Then from equation (2), $g(1) = 2g(4) - g(2) = 0$, so for all $x \in \{1, 2, 3, 4\}$, $g(x) = 0$. Then $\Pr_\theta[g(X) = 0] = 1$, for all $\theta \in (0,\ 0.1)$, so this family of distributions is complete.

Now suppose that $X$ follows Distribution II and let $g : \{1, 2, 3, 4\} \to \mathbb{R}$ be defined by $g(1) = 1$, $g(2) = -1$, $g(3) = g(4) = 0$. Then for all $\theta \in \Theta = (0, 0.1)$,

$$\begin{aligned}
\mathrm{E}_\theta[g(X)] &= \sum_{x=1}^{4} g(x) \Pr_\theta[X = x] \\
&= 2\theta^2\, g(1) + 2\theta^2\, g(2) + (2\theta^3 - \theta^4)\, g(3) + (1 + \theta^4 - 4\theta^2 - 2\theta^3)\, g(4) \\
&= 2\theta^2 \times 1 + 2\theta^2 \times (-1) + (2\theta^3 - \theta^4) \times 0 + (1 + \theta^4 - 4\theta^2 - 2\theta^3) \times 0 \\
&= 2\theta^2 - 2\theta^2 = 0
\end{aligned}$$

but

$$\begin{aligned}
\Pr_\theta[g(X) = 0] &= \Pr_\theta[X = 3] + \Pr_\theta[X = 4] \\
&= 1 - \Pr_\theta[X = 1] + \Pr_\theta[X = 2] \\
&= 1 - 4\theta^2 < 1
\end{aligned}$$

as $\theta > 0$. Hence this family of distributions is not complete.

3. (a) Now $X_1, ..., X_n$ are i.i.d. random variables with common density $f(x;\theta) = e^{\theta - x}\, I_{(-\infty,x]}(\theta) = \begin{cases} e^{\theta-x} & \text{, if x} \geq \theta \\ 0 & \text{, otherwise.} \end{cases}$

Then the likelihood function is

$$\begin{aligned} L(\mathbf{x},\theta) &= \prod_{i=1}^{n} f(x_i;\theta) \qquad\qquad \text{as } X_1,...,X_n \text{ are i.i.d.} \\ &= \prod_{i=1}^{n} e^{\theta - x_i}\, I_{(-\infty,x_i]}(\theta) \\ &= \exp(n\theta - \sum_{i=1}^{n} x_i)\, I_{(-\infty,x_{(1)}]}(\theta). \end{aligned}$$

Let $\psi(\theta)$ be the ratio of the likelihoods for data vectors $\mathbf{x}$ and $\mathbf{y}$. Then

$$\psi(\theta) = \frac{L(\mathbf{x},\theta)}{L(\mathbf{y},\theta)} = \frac{\exp(n\theta - \sum_{i=1}^{n} x_i)\, I_{(-\infty,x_{(1)}]}(\theta)}{\exp(n\theta - \sum_{i=1}^{n} y_i)\, I_{(-\infty,y_{(1)}]}(\theta)} = \exp(\sum_{i=1}^{n} y_i - \sum_{i=1}^{n} x_i)\frac{I_{(-\infty,x_{(1)}]}(\theta)}{I_{(-\infty,y_{(1)}]}(\theta)}.$$

Now when $x_{(1)} < y_{(1)}$, we have

$$\psi(\theta) = \begin{cases} \exp(\sum_{i=1}^{n} y_i - \sum_{i=1}^{n} x_i) & , \theta \leq x_{(1)} \\ 0 & , x_{(1)} < \theta \leq y_{(1)} \\ \text{undefined} & , \theta > y_{(1)} \end{cases}$$

and when $y_{(1)} < x_{(1)}$, we have

$$\psi(\theta) = \begin{cases} \exp(\sum_{i=1}^{n} y_i - \sum_{i=1}^{n} x_i) & , \theta \leq y_{(1)} \\ \infty & , y_{(1)} < \theta \leq x_{(1)} \\ \text{undefined} & , \theta > x_{(1)} \end{cases}$$

so when $x_{(1)} \neq y_{(1)}$, $\psi(\theta)$ is not constant with respect to $\theta$ (where it is defined). When $x_{(1)} = y_{(1)}$, we have

$$\psi(\theta) = \begin{cases} \exp(\sum_{i=1}^{n} y_i - \sum_{i=1}^{n} x_i) & , \theta \leq x_{(1)} = y_{(1)} \\ \text{undefined} & , \theta > x_{(1)} = y_{(1)} \end{cases}$$

so $\psi(\theta)$ is constant with respect to $\theta$ (where it is defined). Hence, by the method of Lehmann and Scheffe, $T(\mathbf{X}) = X_{(1)}$ is a minimal sufficient statistic for $\theta$.

(b) Let $F(x;\theta)$ be the common distribution function for $X_1, ..., X_n$. Then for $x \geq \theta$,

$$\begin{aligned} F(x;\theta) &= \int_{\theta}^{x} f(u;\theta)\, du \\ &= \int_{\theta}^{x} e^{\theta - u} du \\ &= e^{\theta}[-e^{-u}]_{u=\theta}^{x} \\ &= e^{\theta}(e^{-\theta} - e^{-x}) \\ &= 1 - e^{\theta - x} \end{aligned}$$

so

$$
\begin{aligned}
F_{X_{(1)}}(x;\theta) &= 1-\Pr_\theta[X_{(1)}>x] \\
&= 1-\Pr_\theta[X_1>x,...,X_n>x] \\
&= 1-(\Pr_\theta[X_1>x])^n \qquad\qquad \text{as } X_1,...,X_n \text{ are i.i.d.} \\
&= 1-[1-F(x;\theta)]^n \\
&= 1-[e^{\theta-x}]^n \\
&= 1-e^{n\theta}(e^{-nx})
\end{aligned}
$$

and hence

$$f_{X_{(1)}}(x;\theta)=\frac{d}{dx}F_{X_{(1)}}(x;\theta)=ne^{n\theta}(e^{-nx})=ne^{n(\theta-x)}.$$

Also for $x<\theta$, $f_{X_{(1)}}(x;\theta)=F_{X_{(1)}}(x;\theta)=0$. Hence

$$
\begin{aligned}
\mathrm{E}_\theta\left[X_{(1)}\right] &= \int_\theta^\infty x\, f_{X_{(1)}}(x;\theta)\,dx \\
&= e^{n\theta}\int_\theta^\infty xne^{-nx}\,dx \\
&= e^{n\theta}\left\{[-xe^{-nx}]_\theta^\infty+\int_\theta^\infty e^{-nx}dx\right\} \\
&= e^{n\theta}\left\{\theta e^{-n\theta}-\frac{1}{n}[e^{-nx}]_\theta^\infty\right\} \\
&= \theta+e^{n\theta}\left(\frac{1}{n}e^{-n\theta}\right) \\
&= \theta+\frac{1}{n}.
\end{aligned}
$$

(c) Let $g$ be a function from $[\theta,\infty)$ to $\mathbb{R}$ and suppose that $\mathrm{E}_\theta[g(X_{(1)})]=0$ for all $\theta\in\mathbb{R}$. Then

$$
\begin{aligned}
0 &= \int_\theta^\infty g(x)\, f_{X_{(1)}}(x;\theta)\,dx \\
&= ne^{n\theta}\int_\theta^\infty g(x)e^{-nx}dx \\
0 &= \int_\theta^\infty g(x)e^{-nx}dx.
\end{aligned}
$$

On differentiating both sides, $-g(\theta)e^{-n\theta}=0$ and so $g(\theta)=0$ for all $\theta\in\mathbb{R}$, as $e^{-n\theta}>0$. Then $g(x)=0$ for all $x\in[\theta,\infty)$, so $\Pr_\theta[g(X_{(1)})=0]=1$ for all $\theta\in\mathbb{R}$, and hence $X_{(1)}$ is complete. From part a), it is also sufficient for $\theta$.

Consider the estimator $W(\mathbf{X})=X_{(1)}-\frac{1}{n}$. Now $\mathrm{E}_\theta[W(\mathbf{X})]=\mathrm{E}_\theta[X_{(1)}]-\frac{1}{n}=\theta$, so $W(\mathbf{X})$ is an unbiased estimator of $\theta$. Then by the Lehmann-Scheffe Theorem, the UMVUE of $\theta$ is $\mathrm{E}\left[W(\mathbf{X})|X_{(1)}\right]=\mathrm{E}\left[X_{(1)}|X_{(1)}\right]-\frac{1}{n}=X_{(1)}-\frac{1}{n}$.

4. (a) Now $X_1,...,X_n$ are i.i.d NegativeBinomial$(k,\theta)$ random variables with common probability function

$$f(x;\theta)=\binom{k-1+x}{x}\theta^k(1-\theta)^x \qquad \text{for } x\in\{0,1,...\},$$

so the likelihood function is

$$\begin{aligned}
L(\mathbf{x},\theta) &= \prod_{i=1}^{n} f(x_i;\theta) \qquad \text{as } X_1,...,X_n \text{ are i.i.d.} \\
&= \left[\prod_{i=1}^{n} \binom{k-1+x_i}{x_i}\right] \theta^{nk}(1-\theta)^{\sum_{i=1}^{n} x_i}.
\end{aligned}$$

Then the ratio of the likelihoods for data vectors $\mathbf{x}$ and $\mathbf{y}$ is

$$\begin{aligned}
\frac{L(\mathbf{x},\theta)}{L(\mathbf{y},\theta)} &= \frac{\left[\prod_{i=1}^{n} \binom{k-1+x_i}{x_i}\right] \theta^{nk}(1-\theta)^{\sum_{i=1}^{n} x_i}}{\left[\prod_{i=1}^{n} \binom{k-1+y_i}{y_i}\right] \theta^{nk}(1-\theta)^{\sum_{i=1}^{n} y_i}} \\
&= \frac{\left[\prod_{i=1}^{n} \binom{k-1+x_i}{x_i}\right]}{\left[\prod_{i=1}^{n} \binom{k-1+y_i}{y_i}\right]}(1-\theta)^{\sum_{i=1}^{n} x_i-\sum_{i=1}^{n} y_i}
\end{aligned}$$

which is independent of $\theta$ if and only if $\sum_{i=1}^{n} x_i = \sum_{i=1}^{n} y_i$. Hence, by the method of Lehmann and Scheffe, $T(\mathbf{X}) = \sum_{i=1}^{n} X_i$ is a minimal sufficient statistic for $\theta$.

(b) The expected value of $X_1$ is

$$\begin{aligned}
\mathrm{E}_\theta[X_1] &= \sum_{x=0}^{\infty} x\, f(x;\theta) \\
&= \sum_{x=1}^{\infty} x \binom{k-1+x}{x} \theta^k (1-\theta)^x \\
&= \theta^k \sum_{x=1}^{\infty} x \frac{(k-1+x)!}{x!(k-1)!}(1-\theta)^x \\
&= \theta^k \sum_{x=1}^{\infty} \frac{(k-1+x)!}{(x-1)!(k-1)!}(1-\theta)^x \\
&= \theta^k \sum_{x=0}^{\infty} \frac{(k+x)!}{x!(k-1)!}(1-\theta)^{x+1} \\
&= \frac{k}{\theta}(1-\theta) \sum_{x=0}^{\infty} \frac{(k+x)!}{x!k!}\theta^{k+1}(1-\theta)^x \\
&= \frac{k(1-\theta)}{\theta} \sum_{x=0}^{\infty} \frac{(l-1+x)!}{x!(l-1)!}\theta^{l}(1-\theta)^x \qquad \text{where } l = k+1 \\
&= \frac{k(1-\theta)}{\theta} \sum_{x=0}^{\infty} f_Y(x;\theta) \qquad \text{where } Y \sim \text{NegativeBinomial}(l,\theta) \\
&= \frac{k(1-\theta)}{\theta}.
\end{aligned}$$

As $X_1,...,X_n$ are i.i.d., this is also the expected value of $X_2,...,X_n$.

(c) Now the score function is

$$\begin{aligned} V(\mathbf{X};\theta) &= \frac{\partial}{\partial\theta}\left[\log L(\mathbf{X};\theta)\right] \\ &= \frac{\partial}{\partial\theta}\left[\sum_{i=1}^{n}\log\binom{k-1+X_i}{X_i} + nk\log\theta + \log(1-\theta)\sum_{i=1}^{n}X_i\right] \\ &= \frac{nk}{\theta} - \frac{\sum_{i=1}^{n}X_i}{1-\theta} \end{aligned}$$

so the Fisher information about $\theta$ in $\mathbf{X}$ is

$$\begin{aligned} I_{\mathbf{X}}(\theta) &= \mathrm{E}_\theta\left[-\frac{\partial}{\partial\theta}V(\mathbf{X};\theta)\right] \\ &= \mathrm{E}_\theta\left[\frac{nk}{\theta^2} + \frac{\sum_{i=1}^{n}X_i}{(1-\theta)^2}\right] \\ &= \frac{nk}{\theta^2} + \frac{\sum_{i=1}^{n}\mathrm{E}_\theta[X_i]}{(1-\theta)^2} \\ &= \frac{nk}{\theta^2} + \frac{nk(1-\theta)/\theta}{(1-\theta)^2} \\ &= nk\left[\frac{1}{\theta^2} + \frac{1}{\theta(1-\theta)}\right] \\ &= nk\left[\frac{(1-\theta)+\theta}{\theta^2(1-\theta)}\right] \\ &= \frac{nk}{\theta^2(1-\theta)}. \end{aligned}$$

As $T(\mathbf{X})$ is a sufficient statistic for $\theta$, from part a), we have $I_{T(\mathbf{X})}(\theta) = I_{\mathbf{X}}(\theta) = \frac{nk}{\theta^2(1-\theta)}$.

(d) Now the score function can be factorised into

$$\begin{aligned} V(\mathbf{X};\theta) &= \frac{nk}{\theta} - \frac{\sum_{i=1}^{n}X_i}{1-\theta} \\ &= -\frac{nk}{1-\theta}\left[\frac{1}{nk}\sum_{i=1}^{n}X_i - \frac{1-\theta}{\theta}\right] \\ &= K_n(\theta)\left[T(\mathbf{X}) - \tau(\theta)\right] \end{aligned}$$

where $K_n(\theta) = -\frac{nk}{1-\theta}$ is independent of $\mathbf{X}$. Hence the UMVUE of $\tau(\theta) = \frac{1-\theta}{\theta}$ is $T(\mathbf{X}) = \frac{1}{nk}\sum_{i=1}^{n}X_i = \frac{\bar{X}}{k}$, and the variance of the UMVUE attains the Cramer-Rao lower bound.

(e) Let the common m.g.f. of $X_1, ..., X_n$ be $m_X(t)$. Now

$$\begin{aligned}
m_X(t) &= \mathrm{E}_\theta[e^{Xt}] \\
&= \sum_{x=0}^{\infty} e^{xt} f(x;\theta) \\
&= \sum_{x=0}^{\infty} e^{xt} \binom{k-1+x}{x} \theta^k (1-\theta)^x \\
&= \theta^k \sum_{x=0}^{\infty} \binom{k-1+x}{x} \left(e^t(1-\theta)\right)^x \\
&= \theta^k \sum_{x=0}^{\infty} \binom{k-1+x}{x} (1-p)^x \qquad \text{where } p = 1 - e^t(1-\theta) \\
&= \frac{\theta^k}{p^k} \sum_{x=0}^{\infty} \binom{k-1+x}{x} p^k (1-p)^x \\
&= \frac{\theta^k}{p^k} \sum_{x=0}^{\infty} f_U(x) \qquad \text{where } U \sim \text{NegativeBinomial}(k, p) \\
&= \left(\frac{\theta}{p}\right)^k \\
&= \left(\frac{\theta}{1-e^t(1-\theta)}\right)^k.
\end{aligned}$$

Let $T(\mathbf{X}) = \sum_{i=1}^n X_i$. Then the m.g.f. of $T(\mathbf{X})$ is

$$\begin{aligned}
m_{T(\mathbf{X})}(t) &= \prod_{i=1}^n m_X(t) \qquad \text{as } X_1, ..., X_n \text{ are i.i.d.} \\
&= \prod_{i=1}^n \left(\frac{\theta}{1-e^t(1-\theta)}\right)^k \\
&= \left(\frac{\theta}{1-e^t(1-\theta)}\right)^{nk}
\end{aligned}$$

so $T(\mathbf{X}) \sim$NegativeBinomial$(nk, \theta)$, as an m.g.f. uniquely characterises a distribution. Then the probability function of $T(\mathbf{X})$ is

$$f_T(t) = \binom{nk-1+t}{t} \theta^{nk} (1-\theta)^t \qquad \text{for } t \in \{0, 1, ...\}.$$

Similarly, $\sum_{i=2}^n X_i \sim$NegativeBinomial$((n-1)k, \theta)$.
Now let $g$ be a function from $\{0, 1, ...\}$ to $\mathbb{R}$ and suppose that $\mathrm{E}_\theta\left[g\left(T(\mathbf{X})\right)\right] = 0$

for all $\theta \in (0,1)$. Then

$$\begin{aligned}
0 &= \sum_{t=0}^{\infty} g(t)\, f_T(t) \\
&= \sum_{t=0}^{\infty} g(t) \binom{nk-1+t}{t} \theta^{nk}(1-\theta)^t \\
&= \theta^{nk} \sum_{t=0}^{\infty} g(t) \binom{nk-1+t}{t} (1-\theta)^t \\
0 &= \sum_{t=0}^{\infty} g(t) \binom{nk-1+t}{t} (1-\theta)^t,
\end{aligned}$$

where the right hand side is a power series in $(1-\theta)$. Then for all $t \in \{0,1,...\}$, $g(t)\binom{nk-1+t}{t} = 0$, so $g(t) = 0$ as $\binom{nk-1+t}{t} \geq 1$. Hence $\Pr_\theta[T(\mathbf{X}) = 0] = 1$ for all $\theta \in (0,1)$, so $T(\mathbf{X})$ is complete.

Now consider the estimator $W(\mathbf{X}) = I_{\{X_1=0\}}(\mathbf{X})$. We have $\mathrm{E}_\theta\,[W(\mathbf{X})] = \Pr_\theta[X_1 = 0] = f(0;\theta) = \theta^k$, so $W(\mathbf{X})$ is an unbiased estimator of $\eta(\theta) = \theta^k$. Now

$$\begin{aligned}
\mathrm{E}\,[W(\mathbf{X}) \,|\, T(\mathbf{X}) = t] &= \Pr[X_1 = 0 \,|\, T(\mathbf{X}) = t] \\
&= \frac{\Pr_\theta[X_1 = 0, \sum_{i=1}^n X_i = t]}{\Pr_\theta[T(\mathbf{X}) = t]} \\
&= \frac{\Pr_\theta[X_1 = 0, \sum_{i=2}^n X_i = t]}{\binom{nk-1+t}{t}\theta^{nk}(1-\theta)^t} \\
&= \frac{\Pr_\theta[X_1 = 0]\Pr_\theta[\sum_{i=2}^n X_i = t]}{\binom{nk-1+t}{t}\theta^{nk}(1-\theta)^t} \qquad \text{as } X_1, ..., X_n \text{ are independent} \\
&= \frac{\theta^k \binom{nk-k-1+t}{t}\theta^{k(n-1)}(1-\theta)^t}{\binom{nk-1+t}{t}\theta^{nk}(1-\theta)^t} \\
&= \frac{\binom{nk-k-1+t}{t}}{\binom{nk-1+t}{t}} \\
&= \frac{(nk-k-1+t)!}{t!(nk-k-1)!}\,\frac{t!(nk-1)!}{(nk-1+t)!} \\
&= \frac{(nk-k-1+t)!(nk-1)!}{(nk-1+t)!(nk-k-1)!}
\end{aligned}$$

so by the Lehmann-Scheffe Theorem, the UMVUE of $\eta(\theta) = \theta^k$ is

$$\mathrm{E}\,[W(\mathbf{X}) \,|\, T(\mathbf{X})] = \frac{(nk-k-1+T(\mathbf{X}))!(nk-1)!}{(nk-1+T(\mathbf{X}))!(nk-k-1)!} = \frac{(nk-k-1+\sum_{i=1}^n X_i)!(nk-1)!}{(nk-1+\sum_{i=1}^n X_i)!(nk-k-1)!}.$$