Hw 5

7. Suppose that X follows a geometric distribution,

$$P(X=k)=p(1-p)^{k-1}$$

① Asymptotic variance of MLE

② Fisher information of X and asymptotic sampling dist. of MLE

Sol. ※ By hw 4.7(b), we have MLE of p be $\hat{p}=\frac{1}{\bar{X}_n}$

We first get Fisher info of X

let $f(x)=p(1-p)^{x-1}$

$$I(p)=-E\left[\frac{\partial^2 \ln f(x)}{\partial p^2}\right]=-E\left[\frac{\partial^2(\ln p+(X-1)\ln(1-p))}{\partial p^2}\right]$$

$$=-E\left[-\frac{1}{p^2}+\frac{(1-X)}{(1-p)^2}\right]=\frac{1}{p^2}+\frac{-1}{(1-p)^2}+\frac{1}{p(1-p)^2}$$

$$=\frac{(1-p)^2-p^2+p}{p^2(1-p)^2}=\frac{1}{p^2(1-p)}$$

using $E[X]=\frac{1}{p}$

we have fisher information be $nI(p)=\frac{n}{p^2(1-p)}$

By Thm 6.6 (CH8 P39), we have

$$\sqrt{nI(p)}\,(\hat{p}-p)\sim N(0,1)$$

$\Rightarrow \hat{p}\sim N\left(p,\frac{p^2(1-p)}{n}\right)$, the asymptotic variance of MLE is $\frac{p^2(1-p)}{n}$

16. Consider an i.i.d. sample of random variables with density function

$$f(x|\sigma) = \frac{1}{2\sigma}\exp\left(-\frac{|x|}{\sigma}\right)$$

① Asymptotic variance of MLE

② Fisher information of X and asymptotic sampling dist. of MLE

Sol. By hw 4. 16 (b), we have MLE of σ be $\hat{\sigma} = \frac{\sum_{i=1}^{n}|X_i|}{n}$

First get fisher information of X

$$I(\sigma) = -E\left[\frac{\partial^2 \ln f(X|\sigma)}{\partial \sigma^2}\right]$$

$$= -E\left[\frac{1}{\sigma^2} - \frac{2|X|}{\sigma^3}\right]$$

$$= -\frac{1}{\sigma^2} + \frac{2}{\sigma^3}E[|X|]$$

$$E[|X|] = \int_{-\infty}^{\infty} |x| \frac{1}{2\sigma}\exp\left(-\frac{|x|}{\sigma}\right) dx$$

$$= \sigma\int_0^{\infty} \frac{|x|}{\sigma}\exp\left(-\frac{|x|}{\sigma}\right) d\frac{x}{\sigma}$$

$$= \sigma\left[-\int_0^{\infty} \frac{|x|}{\sigma}\, d\exp\left(-\frac{|x|}{\sigma}\right)\right]$$

$$= \sigma\left[-\frac{|x|}{\sigma}\exp\left(\frac{-|x|}{\sigma}\right)\Big|_0^{\infty} + \int \exp\left(\frac{-|x|}{\sigma}\right) d\frac{|x|}{\sigma}\right]$$

$$= \left[\underbrace{-x\exp\left(\frac{-x}{\sigma}\right)}_{①} - \sigma\exp\left(\frac{-x}{\sigma}\right)\right]\Big|_0^{\infty}$$

$= \delta$

① $\lim_{x\to\infty} x\exp\left(\frac{-x}{\delta}\right) = \lim_{x\to\infty} \frac{x}{\exp\left(\frac{x}{\delta}\right)} \quad \left(\frac{\infty}{\infty}\right)$

$\overset{L'hos}{=} \lim_{x\to\infty} \frac{1}{\frac{1}{\delta}\exp\left(\frac{x}{\delta}\right)} = 0$

we have $I(\delta) = -\frac{1}{\delta^2} + \frac{2}{\delta^3}\cdot\delta = \frac{1}{\delta^2}$

Fisher information be $nI(\delta) = \frac{n}{\delta^2}$

By Thm 5.6, we have

$\sqrt{nI(\delta)}\,(\hat{\delta} - \delta) \sim N(0,1)$

$\Rightarrow \hat{\delta} \sim N\Big(\delta, \underbrace{\frac{\delta^2}{n}}\Big)$ III

↳ asym. var of MLE

18. Suppose that $X_1, X_2, \ldots, X_n$ are i.i.d. random variables on the interval $[0, 1]$ with the density function

$$f(x|\alpha) = \frac{\Gamma(3\alpha)}{\Gamma(\alpha)\Gamma(2\alpha)} x^{\alpha-1}(1-x)^{2\alpha-1}$$

where $\alpha > 0$ is a parameter to be estimated from the sample. It can be shown that

$$E(X) = \frac{1}{3}$$

$$\text{Var}(X) = \frac{2}{9(3\alpha+1)}$$

① Asymptotic variance of MLE

② Fisher information of X and asymptotic sampling dist. of MLE

Sol. ※ From Hw4 18(b), there is no close form for α's MLE $\hat{\alpha}$

We first get fisher information

$I(\alpha) = -E[\underline{\frac{d^2}{d\alpha^2} \ln f(x|\alpha)}]$

$$\frac{d^2}{d\alpha^2} \ln f(x|\alpha) = \frac{d^2}{d\alpha^2}\Big(\ln \Gamma(3\alpha) - \ln\Gamma(\alpha) - \ln\Gamma(2\alpha) + (\alpha-1)\ln x + (2\alpha-1)\ln(1-x)\Big)$$

(using digamma function $\psi(\cdot)$ with $\psi(\alpha) = \frac{d}{d\alpha}\ln\Gamma(\alpha)$)

$$= \frac{d}{d\alpha}\Big(3\psi(3\alpha) - \psi(\alpha) - 2\psi(2\alpha) + \ln x + 2\ln(1-x)\Big)$$

$$= 3\psi'(3\alpha) - \psi'(\alpha) - 2\psi'(2\alpha)$$

we have $I(\alpha) = -3\psi'(3\alpha) + \psi'(\alpha) + 2\psi'(2\alpha)$

Fisher information be $nI(\alpha) = -3n\psi'(3\alpha) + n\psi'(\alpha) + 2n\psi'(2\alpha)$

by thm 6.6, we have

$$\sqrt{nI(\alpha)}\,(\hat{\alpha} - \alpha) \sim N(0,1)$$

$$\Rightarrow \hat{\alpha} \sim N\Big(\alpha, \underbrace{\frac{1}{n(-3\psi'(3\alpha) + \psi'(\alpha) + 2\psi'(2\alpha))}}_{\text{Variance of MLE}}\Big) \quad \blacksquare$$

53. Let $X_1, \ldots, X_n$ be i.i.d. uniform on $[0, \theta]$.

a. Find the method of moments estimate of θ and its mean and variance.

b. Find the mle of θ.

d. Find a modification of the mle that renders it unbiased.

(a) $\mu_1 = E(X) = \frac{\theta}{2} \Rightarrow \hat{\mu}_1 = \bar{X} \Rightarrow \frac{\hat{\theta}}{2} = \bar{X} \quad \therefore \hat{\theta}_{MME} = 2\bar{X} = \frac{2}{n}\sum_{i=1}^{n} X_i$

$$E(\hat{\theta}_{MME}) = 2E(\bar{X}) = 2 \cdot \frac{\theta}{2} = \theta$$

$$Var(\hat{\theta}_{MME}) = 4Var(\bar{X}) = 4\frac{1}{n^2}\sum_{i=1}^{n} Var(X_i) = \frac{4}{n^2} \cdot n \cdot \frac{\theta^2}{12} = \frac{\theta^2}{3n}$$

(b) $L(\theta|\underset{\sim}{x}) = \prod_{i=1}^{n} f_{X_i}(x_i;\theta) = \frac{1}{\theta^n} I_{[X_{(n)},\infty)}(\theta)$

$\because L(\theta|\underset{\sim}{x})$ is decreasing function of θ

$\therefore \hat{\theta}_{MLE} = X_{(n)}$

(d) $f_{X_{(n)}}(x) = \frac{n!}{(n-1)!}\left(\frac{x}{\theta}\right)^{n-1}\frac{1}{\theta} = \frac{nx^{n-1}}{\theta^n}, \quad 0 \le x \le \theta$

$$E(X_{(n)}) = \int_0^\theta x\,\frac{nx^{n-1}}{\theta^n}\,dx = \frac{n}{\theta^n}\int_0^\theta x^n\,dx = \frac{n}{\theta^n}\,\frac{\theta^{n+1}}{n+1} = \frac{n}{n+1}\theta$$

$$\Rightarrow E\left(\frac{n+1}{n}X_{(n)}\right) = \theta$$

Therefore $\hat{\theta} = \frac{n+1}{n}X_{(n)}$ is an unbiased estimate of θ.

58. If gene frequencies are in equilibrium, the genotypes AA, Aa, and aa occur with probabilities $(1-\theta)^2$, $2\theta(1-\theta)$, and θ^2, respectively. Plato et al. (1964) published the following data on haptoglobin type in a sample of 190 people:

Haptoglobin Type		
Hp1-1	Hp1-2	Hp2-2
10	68	112

a. Find the mle of θ.
b. Find the asymptotic variance of the mle.

(a) Let n_1 be the number of people with AA genotype.
Let n_2 be the number of people with Aa genotype.

$(n_1, n_2) \sim \text{Trinomial}\left(n=190,\ p_1=(1-\theta)^2,\ p_2=2\theta(1-\theta)\right)$

$$L(\theta|\underset{\sim}{x}) = \frac{n!}{n_1!\,n_2!\,(n-n_1-n_2)!}(1-\theta)^{2n_1}(2\theta(1-\theta))^{n_2}\,\theta^{2(n-n_1-n_2)}$$
$$= \frac{n!\cdot 2^{n_2}}{n_1!\,n_2!\,(n-n_1-n_2)!}(1-\theta)^{2n_1+n_2}\,\theta^{2(n-n_1-n_2)+n_2}$$

$$\Rightarrow \ell(\theta|\underset{\sim}{x}) = \ln\left(\frac{n!\cdot 2^{n_2}}{n_1!\,n_2!\,(n-n_1-n_2)!}\right) + (2n_1+n_2)\ln(1-\theta) + (2(n-n_1-n_2)+n_2)\ln(\theta)$$

$$\frac{\partial}{\partial\theta}\ell(\theta|\underset{\sim}{x}) = \frac{-2n_1-n_2}{1-\theta} + \frac{2(n-n_1-n_2)+n_2}{\theta} = 0$$
$$\Rightarrow \frac{-2n_1\theta - n_2\theta + 2(n-n_1-n_2)+n_2 - 2(n-n_1-n_2)\theta - n_2\theta}{\theta(1-\theta)} = 0$$
$$\Rightarrow \frac{-2n\theta + 2(n-n_1-n_2)+n_2}{\theta(1-\theta)} = 0$$
$$\Rightarrow \hat{\theta}_{MLE} = \frac{2n-2n_1-n_2}{2n}$$

$\because$ $n = n_1+n_2+n_3$, $n_3 \equiv$ The number of people with aa genotype.

$\therefore$ $\hat{\theta}_{MLE}$ can be expressed in other forms.

eg., subsituting $n_2 = n-n_1-n_3$ into the MLE formula $\Longrightarrow \hat{\theta}_{MLE} = \dfrac{n-n_1+n_3}{2n}$

subsituting $n_1 = n-n_2-n_3$ into the MLE formula $\Longrightarrow \hat{\theta}_{MLE} = \dfrac{n_2+2n_3}{2n}$

$$\frac{\partial^2}{\partial\theta^2}\ell(\theta|\underset{\sim}{x}) = \frac{-2n_1-n_2}{(1-\theta)^2} - \frac{2(n-n_1-n_2)+n_2}{\theta^2} < 0,\quad \forall\,\theta\in(0,1)$$

MLE of θ is $\hat{\theta}_{MLE} = \dfrac{292}{380} = \dfrac{73}{95} \approx 0.7684$

(b) By trinomial distribution, $E(n_1) = n(1-\theta)^2$, $E(n_2) = 2n\theta(1-\theta)$, $E(n_3) = n\theta^2$

$$
\begin{aligned}
I_{\underset{\sim}{X}}(\theta) &= -E\left(\frac{\partial^2}{\partial\theta^2}\ell(\theta|\underset{\sim}{x})\right) \\
&= E\left(\frac{2n_1+n_2}{(1-\theta)^2} + \frac{2(n-n_1-n_2)+n_2}{\theta^2}\right) \\
&= \frac{2n(1-\theta)^2+2n\theta(1-\theta)}{(1-\theta)^2} + \frac{2n\theta^2+2n\theta(1-\theta)}{\theta^2} \\
&= \frac{2n\left(\theta(1-\theta)+\theta^2+(1-\theta)^2+\theta(1-\theta)\right)}{\theta(1-\theta)} \\
&= \frac{2n}{\theta(1-\theta)}
\end{aligned}
$$

Asymptotic variance of $\hat{\theta}_{MLE}$ is $Var(\hat{\theta}_{MLE}) = \dfrac{1}{I_{\underset{\sim}{X}}(\theta)} = \dfrac{\theta(1-\theta)}{2n}$,

and estimated value of $Var(\hat{\theta}_{MLE})$ is

$$\widehat{Var(\hat{\theta}_{MLE})} = \frac{1}{I_{\underset{\sim}{X}}(0.7684)} = \frac{0.7684\times(1-0.7684)}{2\times 190} \approx 0.00047$$

By LN Ch8 P.39 Theorem 6.6, we have

$$\sqrt{I_{\underset{\sim}{X}}(\theta)}\left(\hat{\theta}_{MLE}-\theta\right) \longrightarrow N(0,1) \Longrightarrow \hat{\theta}_{MLE} \longrightarrow N\left(\theta, \frac{1}{I_{\underset{\sim}{X}}(\theta)} = \frac{\theta(1-\theta)}{2n}\right)$$