Jack Rodenni
STAT 630
Assignment 8
11/5/21

1) Chp 6 Exercise 6.2.4:

If $(x_1, \ldots, x_n)$ is a sample from a poisson (θ) distribution, where $\theta \in (0,\infty)$, then

(a) determine the mle of θ.

- $L(\theta \mid x_1,\ldots,x_n) = \prod_{i=1}^{n} f_\theta(x_i) = \prod_{i=1}^{n} \frac{\theta^{x_i}}{x_i!} e^{-\theta} = \frac{e^{-n\theta}\, \theta^{\sum_{i=1}^{n} x_i}}{\prod_i x_i!}$

$$\ell(\theta \mid x_1,\ldots,x_n) = \ln\left[\frac{e^{-n\theta}\,\theta^{\sum_i x_i}}{\prod_i x_i!}\right] = \ln\left(e^{-n\theta}\,\theta^{\sum_i x_i}\right) - \ln\left(\prod_i x_i!\right)$$

$$= \ln\left(e^{-n\theta}\right) + \ln\left(\theta^{\sum x_i}\right) - \ln\left(\prod x_i!\right)$$

$$\ell(\theta \mid x_1,\ldots,x_n) = -n\theta + \sum x_i \ln(\theta) - \ln\left(\prod_i x_i!\right)$$

- $S(\theta \mid s) = \frac{\partial \ell(\theta \mid s)}{\partial \theta} = -n + \frac{\sum x_i}{\theta}$
- $S(\theta \mid s) = 0 = -n + \sum x_i/\theta \iff n = \frac{\sum x_i}{\theta} \iff \boxed{\hat{\theta} = \frac{\sum x_i}{n}}$

(b) Show that $\sum_{i=1}^{n} x_i$ is sufficient for θ.

NOTE: Theorem 6.1.1 (Factorization Theorem). If the density (or probability function) for a model factors as $f_\theta(s) = h(s)\, g_\theta(T(s))$, where g_θ and h are nonnegative, then T is a sufficient statistic.

$$f_\theta(x_1,\ldots,x_n) = \prod_{i=1}^{n} \frac{\theta^{x_i}}{x_i!} e^{-\theta} = \frac{e^{-n\theta}\,\theta^{\sum x_i}}{\prod x_i!} = \frac{e^{-n\theta}\,\theta^{T}}{\prod x_i!} \quad \text{where } T = \sum x_i$$

- Let $h(s) = \frac{1}{\prod x_i!}$, $g_\theta(T(s)) = e^{-n\theta}\,\theta^{T}$

Thus

$$f_\theta(x_1,\ldots,x_n) = h(s)\, g_\theta(T(s)) \Rightarrow T = \sum x_i \text{ is a sufficient statistic}$$

(c) Evaluate the bias, variance & MSE of the maximum likelihood estimator.

- Bias $= E[\hat{\theta}] - \theta \Rightarrow E[\hat{\theta}] = E\left[\frac{\sum x_i}{n}\right] = \frac{n\theta}{n}$

$$\text{Bias} = E[\hat{\theta}] - \theta = \frac{n\theta}{n} - \theta = 0.$$

- $\text{var}(\hat{\theta}) = \text{var}\left(\frac{x_1 + \cdots + x_n}{n}\right) = \frac{1}{n^2}\text{var}(x_1 \ldots x_n) = \frac{n\theta}{n^2} = \frac{\theta}{n}$

- $\text{MSE} = \text{Var}_\theta(T) + \left[\text{Bias}_\theta(T)\right]^2 = \frac{\theta}{n} + 0^2 = \frac{\theta}{n}$

1.) (Cont'd)

(d) What is the MLE for θ^2? Is it unbiased? If not, what is its bias?

[Note: Invariance property says that if $\hat{\theta}$ is the MLE of θ and $\psi(\theta)$ is a one-to-one function, then $\psi(\hat{\theta})$ is the MLE of $\psi(\theta)$.]

- By Invariance property:

$$\boxed{MLE(\theta^2) = \left(\frac{\sum x_i}{n}\right)^2 = \bar{X}^2}$$

- Is it unbiased:

$$Bias(\theta^2) = E[\theta^2] - \theta^2 = \theta^2 - \theta^2 = 0.$$

(e.) What is the MLE for $P(x_i = 0)$?

$$P(X_i = x) = \frac{\theta^x e^{-\theta}}{x!} \Rightarrow P(X_i = 0) = \frac{\hat{\theta}^0 e^{-\hat{\theta}}}{0!} = \boxed{e^{-\frac{1}{n}\sum x_i} = P(X_i = 0)}$$

2.) Chp 6 Exercise 6.2.7

If $(x_1, \dots x_n)$ is a sample from a Beta$(\alpha, 1)$ distribution (see problem 2.4.24), where $\alpha > 0$ is unknown, then determine the MLE of α.

(a) $$f_{\alpha,\beta}(x_1, \dots, x_n) = \prod_{i=1}^{n} \frac{1}{B(\alpha,\beta)} x^{\alpha-1}(1-x)^{\beta-1} \overset{\beta=1}{=} \prod \frac{\Gamma(\alpha+1)}{\Gamma(\alpha)\Gamma(1)} x^{\alpha-1}(1-x)^0$$

$$L(\alpha \mid x_1 \dots x_n) = \prod_{i=1}^{n} \frac{\alpha\Gamma(\alpha)}{\Gamma(\alpha)} x_i^{\alpha-1} = \prod_{i=1}^{n} \alpha x_i^{\alpha-1} = \alpha^n \prod_{i=1}^{n} x_i^{\alpha-1}$$

$$\ell(\alpha \mid x_1, \dots, x_n) = \ln(\alpha^n \Sigma x_i^{\alpha-1}) = n\ln(\alpha) + \sum_{i=1}^{n} \ln(x_i^{\alpha-1})$$
$$= n\ln(\alpha) + \sum_{i=1}^{n} \ln(x_i)(\alpha-1)$$

$$\frac{\partial \ell}{\partial \alpha} = \frac{n}{\alpha} + \Sigma \ln(x_i) = 0 \iff \boxed{\hat{\alpha} = \frac{-n}{\Sigma \ln(x_i)}}$$

(b) Provide a statistic (reduced from the sample itself) that is sufficient for α

Theorem 6.1.1 (Factorization Theorem): If the density (or probability function) for a model factors as $f_\theta(s) = h(s)\, g_\theta(T(s))$ where g_θ & h are nonnegative, then T is a sufficient statistic.

$$f_\theta(s) = \prod_{i=1}^{n} \frac{1}{B(\alpha,1)} x_i^{\alpha-1} = \prod_{i=1}^{n} \alpha x_i^{\alpha-1} = \alpha^n \left(\prod_{i=1}^{n} x_i\right)^{\alpha-1}$$

Then $T = \prod_{i=1}^{n} x_i$ is a sufficient statistic

(c) What is the MLE for $Var(x_i)$

By the invariance property:

Plugging in our MLE $\hat{\alpha}$ & $\beta = 1$

$$Var(x_i) = \frac{\alpha\beta}{(\alpha+\beta)^2(\alpha+\beta+1)} = \frac{-\frac{n}{\Sigma \ln(x_i)}}{\left(-\frac{n}{\Sigma x_i}+1\right)^2\left(-\frac{n}{\Sigma x_i}+2\right)} = \boxed{\frac{\frac{n}{\Sigma \ln(x_i)}}{\left(-\frac{n}{\Sigma \ln(x_i)}+1\right)^2\left(\frac{n}{\Sigma \ln(x_i)}-2\right)}}$$

2.) Chp 6 Exercise 6.2.7:

(d) Compute $E[X_i]$ and use this to obtain a MOM estimator for α.

$E[X_i] = \frac{\alpha}{\alpha+1} \Leftrightarrow \alpha E[X_i] + E[X_i] = \alpha$

$E[X_i] = \alpha - \alpha E[X_i]$

$\boxed{\frac{E[X_i]}{1-E[X_i]} = \alpha}$

* Write down likelihood function later.

3.) $W_1, \dots, W_n$ are iid from the distribution w/ pdf $f(w) = \frac{3w^2}{\beta^3} I_{[0,\beta]}(w)$?

(a) Write down the likelihood function and use it to find the MLE for β. Careful - note the support of f; it may help to first sketch what the likelihood function looks like.

Note: $f(w) = \frac{3w^2}{\beta^3} I_{[0,\beta]}(w)$ is a decreasing function for β for $\beta > 0$

Also, the constraints on the w_i's imply that $w_i \le \beta \ \forall i$.

Thus, (by slide 22 of Chp 6 notes) we know that to maximize a decreasing function, we take the smallest allowable value of β to be the MLE.

$\boxed{\hat{\beta} = \max(w_1, \dots, w_n) = W_{(n)}}$

$\boxed{* L(\beta | w_1, \dots, w_n) = (3^n \prod_{i=1}^{n} w_i^2) / \beta^{3n} \; I_{[0,\beta]}(w_i)}$

(b) Find Method of Moments estimator for β. Is it unbiased?

$E[w] = \frac{1}{\beta^3} \int_0^\beta 3w^3 \, dw = 3\frac{1}{\beta^3} \left[\frac{w^4}{4}\right]_{w=0}^{w=\beta} = \frac{3}{4}\beta$

$\frac{3}{4}\beta = E[\hat{\beta}] \Leftrightarrow E[\hat{\beta}] = 4/3\,\beta \Leftrightarrow \boxed{\hat{\beta} = 4/3\,\bar{\beta}}$

2

4.) <u>Chp 6 Exercise 6.2.12</u>: If $(x_1, \dots, x_n)$ is a sample from an $N(\mu_0, \sigma^2)$ distribution, where $\sigma^2 > 0$ is unknown and μ_0 is known, then determine the MLE of σ^2. How does the MLE differ from the plug-in MLE of σ^2 computed using the location-scale normal model?

$$L(\sigma^2 | x_1, \dots, x_n) = \prod_{i=1}^{n} \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x_i - \mu_0}{\sigma}\right)^2} = \frac{1}{\sigma^n (2\pi)^{n/2}} e^{\sum_{i=1}^{n} -\frac{1}{2}\left(\frac{x_i - \mu_0}{\sigma}\right)^2}$$

$$\left[L(\sigma^2 | x_1, \dots, x_n) = \frac{1}{\sigma^n (2\pi)^{n/2}} e^{-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i - \mu_0)} \right]$$

$$l = \ln(L(\sigma^2 | x_1, \dots, x_n)) = \ln\left(e^{-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i - \mu_0)}\right) - \ln\left(\sigma^n (2\pi)^{n/2}\right)$$

$$= -\frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i - \mu_0) - \left(n\ln(\sigma) + \frac{n}{2}\ln(2\pi)\right)$$

$$= -\frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i - \mu_0) - n\ln(\sigma) - \frac{n}{2}\ln(2\pi)$$

$$\frac{\partial l}{\partial \sigma} = \frac{1}{\sigma^3}\sum x_i - \mu_0 - \frac{n}{\sigma} = 0 \Rightarrow \boxed{\hat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i - \mu_0)^2}$$

(b) Evaluate the bias, variance, & MSE of this estimator.

$$\cdot\, E[\hat{\sigma}^2] = E\left[\frac{1}{n}\sum_{i=1}^{n}(x_i - \mu_0)^2\right] = \frac{1}{n}E\left[\sum_{i=1}^{n} x_i^2 - n(\mu_0)^2\right] = \frac{1}{n}\left[\sum_{i=1}^{n} E[x_i^2] - nE(\mu^2)\right]$$

$$= \frac{1}{n}\left[\sum_{i=1}^{n}(\sigma^2 + \mu^2) - n\left[\frac{\sigma^2}{n} + \mu^2\right]\right]$$

$^*(var(x) = E[x^2] - E[x]^2 \Rightarrow E[x^2] = var(x) + E[x]^2$

$(var(\bar{x}) = E[\bar{x}^2] - E[\bar{x}]^2 \Rightarrow E[\bar{x}^2] = \frac{\sigma^2}{n} - \bar{x}^2)^*$

$$= \frac{1}{n}\left[n\sigma^2 + n\mu^2 - (\sigma^2 + n\mu_0^2)\right] = \frac{1}{n}(n\sigma^2 - \sigma^2) = \frac{n-1}{n}\sigma^2$$

$$\boxed{\text{Bias}(\hat{\sigma}^2) = \frac{n-1}{n}\sigma^2 - \sigma^2 = -\frac{1}{n}\sigma^2}$$

$$\cdot\, Var(\hat{\sigma}^2) = var\left(\frac{1}{n}\sum(x_i - \bar{x})^2\right) = var\left(\frac{\sigma^2}{n}\sum(x_i - \bar{x})^2/\sigma^2\right)$$

$$= \frac{\sigma^4}{n^2} var\left(\sum(x_i - \bar{x})^2/\sigma^2\right) \quad \left[\text{*NOTE: } \frac{\sum(x_i - \bar{x})^2}{\sigma^2} \sim \chi^2(n-1)\right]$$

$$\boxed{var(\hat{\sigma}^2) = \frac{2(n-1)\sigma^4}{n^2}}$$

$$MSE(\hat{\sigma}^2) = Var(\sigma^2) + (Bias(\sigma^2))^2 = \frac{2(n-1)\sigma^4}{n} + \frac{\sigma^4}{n^2} = \boxed{\frac{(2n-1)\sigma^4}{n^2}}$$

4.) Chp 6 Exercise 6.2.12 (contd)

(c) Compare its MSE to those of S^2 ; $\hat{\sigma}^2$ from Example 45 on slides 42-44 of the chapter 6 lecture notes. Which of the unbiased estimators has smallest MSE? Why is this reasonable.

- If we compare the MSEs of the two estimators, we see that the MSE of the biased estimator, the MLE $\hat{\sigma}^2$, is smaller than that of the unbiased estimator the sample variance S^2.

5.) Chp 6 Exercise 6.2.8: (NOTE this equation would need to be solved numerically; do not try and do this yourself.)

- If $(x_1, \ldots, x_n)$ is a sample from a Weibull (β) distribution (see problem 2.4.19) where $\beta > 0$ is unknown, then determine the score equation for the MLE of β.

$$f_\beta(x_i) = \beta x_i^{\beta-1} e^{-x_i^\beta} \quad \text{for } 0 < x_i < \infty$$

$$L(\beta \mid x_1, \ldots, x_n) = \prod_{i=1}^{n} \beta x_i^{\beta-1} e^{-x_i^\beta} = \beta^n \left(\prod_{i=1}^{n} x_i^{\beta-1}\right) e^{-\sum x_i^\beta}$$

$$\ell(\beta \mid x_1, \ldots, x_n) = \ln\left(\beta^n \left(\prod_{i=1}^{n} x_i^{\beta-1}\right) e^{-\sum x_i^\beta}\right)$$

$$= \ln(\beta^n) + \ln\left(\prod_{i=1}^{n} x_i^{\beta-1}\right) + \ln\left(e^{-\sum x_i^\beta}\right)$$

$$= n\ln(\beta) + (\beta-1)\sum \ln(x_i) - \sum x_i^\beta$$

$$\boxed{\frac{\partial \ell}{\partial \beta} = \frac{n}{\beta} + \sum \ln(x_i) - \sum \beta x_i^{\beta-1}}$$

6.) Chap 6 Exercise 6.2.19: Hint review the multinomial model (Example 6.15) and note that the parameter space is reduced to one dimension in this exercise (the Hardy-Weinberg model)

- The Hardy-Weinberg law in genetics says that the proportions of genotypes AA, Aa, & aa are θ^2, $2\theta(1-\theta)$ and $(1-\theta)^2$ respectively, where $\theta \in [0,1]$. & That in a sample of n from the population (small relative to the size of the population), we observe x_1, x_2, x_3 individuals of ty AA, Aa & aa respetively.

a.) what distribution do the counts (X_1, X_2, X_3) follow?

$$(X_1, X_2, X_3) \sim \text{Multinomial}(n, \theta^2, 2\theta(1-\theta), (1-\theta)^2)$$

b.) Record the likelihood function, the log likelihood function and the score function for θ.

- $$L(\theta \mid x_1, x_2, x_3) = \binom{n}{x_1 x_2 x_3}(\theta^2)^{x_1}(2\theta(1-\theta))^{x_2}((1-\theta)^2)^{x_3}$$
$$= \binom{n}{x_1 x_2 x_3}(\theta^{2x_1})(2^{x_2}\theta^{x_2}(1-\theta)^{x_2})(1-\theta)^{2x_3}$$
$$\boxed{L(\theta \mid x_1, x_2, x_3) = \binom{n}{x_1, x_2, x_3}(\theta^{2x_1+x_2})((1-\theta)^{x_2+2x_3})(2^{x_2})}$$

- $$l(\theta \mid x_1, x_2, x_3) = \ln(L(\theta \mid x_1, x_2, x_3))$$
$$= \ln\binom{n}{x_1, x_2, x_3} + \ln(\theta^{2x_1+x_2}) + \ln((1-\theta)^{x_2+2x_3}) + \ln(2^{x_2}).$$
$$\boxed{l(\theta \mid x_1, x_2, x_3) = \ln\binom{n}{x_1 x_2 x_3} + (2x_1+x_2)\ln(\theta) + (x_2+2x_3)\ln(1-\theta) + x_2\ln(2)}$$

- $$S(\theta \mid s) = \frac{\partial}{\partial\theta}\left[\ln\binom{n}{x_1 x_2 x_3} + (2x_1+x_2)\ln(\theta) + (x_2+2x_3)\ln(1-\theta) + x_2\ln(2)\right]$$
$$\boxed{S(\theta \mid s) = \frac{2x_1+x_2}{\theta} - \frac{x_2+2x_3}{1-\theta}}$$

c.) Record the form of the MLE for θ.

- $$S(\theta \mid s) = 0 \iff \frac{2x_1+x_2}{\theta} - \frac{x_2+2x_3}{1-\theta} = 0$$
$$\iff (1-\theta)(2x_1+x_2) - \theta(x_2+2x_3) = 0$$
$$\iff -\theta[(2x_1+x_2)+(x_2+2x_3)] + (2x_1+x_2) = 0$$
$$\iff \theta = \frac{2x_1+x_2}{2x_1+2x_2+2x_3} \iff \theta = \frac{2x_1+x_2}{2(x_1+x_2+x_3)}$$

[* note: $n = x_1+x_2+x_3$]

$$\iff \boxed{\hat{\theta} = \frac{2x_1+x_2}{2n}}$$

$ $X_1, \ldots, X_n$ is a random sample from an $Exp(\lambda)$ distribution. Find the estimator for λ of the form $L_a = \frac{a}{\sum X_i}$ w/ the smallest MSE. That is, Find a to minimize $E[(L_a - \lambda)^2]$, and give the minimum value. Hint: you will first need to take note of the distribution of $T = \sum_{i=1}^{n} X_i$ and use that to find $E(1/T)$; $E[\frac{1}{T^2}]$

- $E[(L_a - \lambda)^2] = E[(\frac{a}{\sum X_i} - \lambda)^2]$; letting $T = \sum_{i=1}^{n} X_i$ we get

$$E[(\tfrac{a}{T} - \lambda)^2] = E[\tfrac{a^2}{T^2} - \tfrac{2a}{T}\lambda + \lambda^2]$$
$$= a^2 E[\tfrac{1}{T^2}] - 2a\lambda E[\tfrac{1}{T}] + \lambda^2$$

- Note $T \sim \text{Gamma}(n, \lambda)$ w/ pdf $f(t) = \frac{\lambda^n}{\Gamma(n)} x^{n-1} e^{-\lambda t} I_{(0,\infty)}(t)$, $\lambda > 0$, $n > 0$

Then $E[\frac{1}{T^k}] = \int_0^\infty \frac{1}{t^k} \cdot \frac{\lambda^n}{\Gamma(n)} t^{n-1} e^{-\lambda t} \, dt$

$$E[\tfrac{1}{T^k}] = \frac{\lambda^k \Gamma(n-k)}{\Gamma(n)} \int_0^\infty \frac{\lambda^{n-k}}{\Gamma(n-k)} t^{n-k-1} e^{-\lambda t} \, dt = \frac{\lambda^k \Gamma(n-k)}{\Gamma(n)}$$

$$E[\tfrac{1}{T}] = \frac{\lambda^1 \Gamma(n-1)}{\Gamma(n)} = \frac{\lambda}{n-1} \quad ; \quad E[\tfrac{1}{T^2}] = \frac{\lambda^2 \Gamma(n-2)}{\Gamma(n)} = \frac{\lambda^2}{(n-2)(n-1)}$$

- $E[(L_a - \lambda)^2] = a^2 E[\frac{1}{T^2}] - 2a\lambda E[\frac{1}{T}] + \lambda^2 = \underbrace{a^2\left(\frac{\lambda^2}{(n-2)(n-1)}\right) - 2a\lambda\left(\frac{\lambda}{n-1}\right) + \lambda^2}$

$$= \frac{a^2\lambda^2}{(n-2)(n-1)} - \frac{2a\lambda^2}{n-1} + \lambda^2$$

- Taking the derivative of $E[(L_a - \lambda)^2]$ w.r.t. a and setting this equal to zero, we get:

$$\frac{\partial}{\partial a}\left[\frac{a^2\lambda^2}{(n-2)(n-1)} - \frac{2a\lambda^2}{n-1} + \lambda^2\right] = \frac{2a\lambda^2}{(n-2)(n-1)} - \frac{2\lambda^2}{n-1} = 0$$

$$2a\lambda^2 - 2\lambda^2(n-2) = 0 \iff 2a\lambda^2 = 2\lambda^2(n-2)$$

$$\boxed{\hat{a} = n-2}$$

$$\Rightarrow \boxed{\hat{L}_a = \frac{n-2}{\sum X_i}}$$

$$\min[MSE] = E(\hat{L}_a - \lambda)^2 = \frac{(n-2)^{\cancel{2}}\lambda^2}{\cancel{(n-2)}(n-1)} - \frac{2(n-2)\lambda^2}{(n-1)} + \lambda^2$$

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

$$= \lambda^2 - \frac{(n-2)\lambda^2}{(n-1)}$$

$$= \frac{\lambda^2(n-1) - \lambda^2(n-2)}{(n-1)}$$

$$= \frac{\lambda^2(n-1-n+2)}{n-1} = \boxed{\frac{\lambda^2}{n-1}}$$