

1) 6.1.7

$$P(\theta) = \frac{e^{-\theta} \theta^x}{x!}$$

$$a) X = \text{Poisson}(\lambda) = \frac{e^{-\lambda} \lambda^x}{x!}$$

$$\begin{aligned} L(\theta|x_1, \dots, x_n) &= \prod_{i=1}^n \frac{e^{-\theta} \theta^{x_i}}{x_i!} \\ &= \frac{e^{-\theta} \theta^x}{x!}, \text{ where } x = \sum_{i=1}^n x_i \end{aligned}$$

$$\begin{aligned} b) f_\theta &= h(s)g_\theta(T(s)) \\ f_\theta(x_1, \dots, x_n) &= \prod_{i=1}^n \frac{e^{-\theta} \theta^{x_i}}{x!} \\ &= \frac{e^{-\theta} \theta^{T(x_1, \dots, x_n)}}{T(x_1, \dots, x_n)!} 1 \\ \text{where, } T(x_1, \dots, x_n) &= \sum_{i=1}^n x_i, \text{ and} \\ g_\theta(T) &= \frac{e^{-\theta} \theta^T}{T!}, \text{ and} \\ h(x_1, \dots, h_n) &= 1 \end{aligned}$$

2) 6.1.19

$$\begin{aligned} f_\theta &= \frac{\theta^{a_0}}{\gamma(a_0)} x^{a_0-1} e^{-\theta x} \\ f_\theta &= h(x)g_\theta(T(x)) \\ &= \prod_{i=1}^n \frac{\theta^{a_0}}{\gamma(a_0)} x^{a_0-1} e^{-\theta x} 1 \\ \text{where, } h &= 1 \\ T(x_1, \dots, x_n) &= \sum_{i=1}^m x_i \\ g_\theta(T) &= \frac{\theta^{a_0}}{\gamma(a_0)} \gamma^{a_0-1} e^{-\theta \gamma} \end{aligned}$$

3) 6.2.4

$$Y = \text{Poisson}(\theta) = \frac{e^{-\theta}\theta^x}{n!}$$

$$l(\theta|x_1, \dots, x_n) = \sum_{i=1}^n \log(f_\theta(x_i))$$

$$\log(f_\theta(x)) = -\theta + x\log(\theta) + \log\left(\frac{1}{n}\right)$$

$$\frac{dl(\theta|x_1, \dots, x_n)}{d\theta} = \frac{x}{\theta} - n = 0$$

$$\theta = \frac{x}{n}$$

$$b) \text{Bias}_\theta(T) = E(T) - \psi(\theta)$$

$$= \theta - \frac{x}{n}$$

$$\text{variance} = \theta$$

$$MSE = \theta - \theta - \frac{x}{n} = \frac{x}{n}$$

4) 6.2.5

$$f = a_0 \log(\theta) + (a_0 - 1) \log(x) - \theta x + \log\left(\frac{1}{\gamma^{a_0}}\right)$$

$$f' = \frac{a_0}{\theta} - x = 0$$

$$\theta = \frac{a_0}{x}$$

$$b) \text{bias} = E_\theta(T) - \psi(\theta)$$

$$= \frac{a_0}{\theta} - \frac{a_0}{\theta} = 0$$

$$c) E(x) = a_0 \theta, \theta = \frac{a_0}{E(x)}$$

$$\hat{\theta} = \frac{a_0}{\bar{x}}$$

5) 6.2.6

a) These are independent events of successes or failures so we can use the negative binomial distribution

$$\begin{aligned}
P_x(x) &= \binom{r-1-x}{x} \theta^r (1-\theta)^x \text{ where, } r=1 \\
&= \binom{x}{x} \theta(\theta^x) = \theta^{x+1} \\
\log(\theta^{x+1}) &= (x+1)\log(\theta) \\
\frac{dl(\theta^{x+1})}{d\theta} &= \frac{x+1}{\theta} = 0 \\
b) E(x) &= \frac{(1-\theta)}{\theta}, \neq \frac{x+1}{\theta}
\end{aligned}$$

6) 6.2.8

$$\begin{aligned}
Weibull(\beta) &= \beta(1+x)^{-\beta-1} \\
l(\beta|x) &= \log(\beta) + (-\beta-1)\log(x+1) \\
\frac{dl(\beta|x)}{x\beta} &= \frac{1}{\beta} + \frac{(-\beta-1)}{1+x}
\end{aligned}$$

7) 6.2.12

a)  $mle = \hat{\sigma}^2 = \frac{n-1}{n}s^2 = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2$  any distribution with a mean  $\mu$  and variance  $\sigma^2 \dots$  They are the same

$$\begin{aligned}
L(\mu, \sigma | x_1, \dots, x_n) &= \prod_{i=1}^n \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x_i - \mu)^2}{2\sigma^2}} \\
l(\mu, \sigma) &= \log(L(\mu, \sigma | x_1, \dots, x_n)) \\
&= \frac{-n}{2} \log(2\pi) - n \log(\sigma) - \frac{1}{2\sigma^2} \sum_{i=1}^n (x_i - \mu)^2 \\
\frac{dl(\mu, \sigma)}{d\sigma} &= \frac{-n}{\sigma} + \frac{1}{\sigma^3} \sum_{i=1}^n (x_i - \mu)^2 = 0 \\
\sigma^2 &= \frac{1}{n} \sum_{i=1}^n (x_i - \hat{x})^2
\end{aligned}$$

8) 6.2.19

$$AA = \theta, Aa = 2\theta(1 - \theta), aa = (1 - \theta)^2$$

a) *Bernoulli Distribution*

$$b) L(\theta|x_1, x_2, x_3) = \theta^{x_1} 2\theta(1 - \theta)^{x_2} (1 - \theta)^{2x_3}$$

$$l(\theta|x_1, x_2, x_3) = x_1 \log(\theta) + x_2 \log(2\theta - \theta^2) + 2x_3 \log(1 - \theta)$$

$$score = \frac{x_1}{\theta} + \frac{x_2}{2\theta - \theta^2} + \frac{2x_3}{1 - \theta} = 0$$

9) 6.3.15

a) yes

b) no

10) 6.3.24

$$\psi(\theta) = aT_1 + (1 - a)T_2$$

$$Bias_{\theta}(T) = E_{\theta}(T) - \psi(\theta)$$

$$0 = E_{\theta}(T) - \psi(\theta)$$

$$\psi(\theta) = E_{\theta}(T)$$

$$b) var_{\theta}(T_1) = aT_1, var_{\theta}(T_2) = T_2(1 - a)$$

c) when  $a = .5$ , both  $T_1$  and  $T_2$  have the same weight and will minimize  $T_1 + (1 - a)T_2$  when  $T_1$  and  $T_2$  are unknown

10) additional

a)  $x_1 \dots x_n, exponential(\lambda) = \lambda e^{-\lambda x}$   $T_c = \frac{c}{\sum_{i=1}^n x_i}$ , when  $C = 1$  the estimator will have the smallest MSE because the expected mean  $exponential(\lambda) = \frac{1}{\lambda}$