MD ABUL HAYAT
University of Arkansas

mahayat@uark.edu

HOMEWORK 3 : STAT 5113 (STATISTICAL INFERENCE)

Question A1: Find Fisher information for a sample from a $Gam(\alpha, \beta)$ distribution, and use it to determine the asymptotic distribution of the ML estimator of the unknown 2-dimensional parameter (α, β).

Answer: Suppose, we collect $X_1, \cdots, X_n$ independent samples from a $Gam(\alpha, \beta)$ distribution, then the pdf of X_i is

$$f_{X_i}(x_i) = \frac{\beta^\alpha}{\Gamma(\alpha)} e^{-\beta x_i} x_i^{\alpha-1}$$

$$\text{Likelihood function,} \quad L(\alpha,\beta) = \Big[\frac{\beta^\alpha}{\Gamma(\alpha)}\Big]^n e^{-\beta \sum_{i=1}^n x_i} \Big[\prod_{i=1}^n x_i\Big]^{\alpha-1}$$

$$\text{log-likelihood,} \quad l(\alpha,\beta) = \ln L(\alpha,\beta) = n\alpha \ln\beta - n\ln\Gamma(\alpha) - \beta\sum_{i=1}^n x_i + (\alpha-1)\Big(\sum_{i=1}^n \ln x_i\Big)$$

$$\frac{\partial l}{\partial \alpha} = n\ln\beta - n\frac{\Gamma'(\alpha)}{\Gamma(\alpha)} + \sum_{i=1}^n \ln x_i = n\ln\beta - n\psi(\alpha) + \sum_{i=1}^n \ln x_i$$

$$\text{where} \quad \psi(\alpha) = \frac{\Gamma'(\alpha)}{\Gamma(\alpha)} \quad \text{is Digamma function}$$

$$\frac{\partial^2 l}{\partial \alpha^2} = -n\psi'(\alpha) \Rightarrow \mathbb{E}\Big[-\frac{\partial^2 l}{\partial \alpha^2}\Big] = n\psi'(\alpha)$$

$$\frac{\partial l}{\partial \beta} = \frac{n\alpha}{\beta} - \sum_{i=1}^n x_i$$

$$\frac{\partial^2 l}{\partial \beta^2} = -\frac{n\alpha}{\beta^2} \Rightarrow \mathbb{E}\Big[-\frac{\partial^2 l}{\partial \beta^2}\Big] = \frac{n\alpha}{\beta^2}$$

$$\frac{\partial^2 l}{\partial \alpha \partial \beta} = \frac{n}{\beta} \Rightarrow \mathbb{E}\Big[-\frac{\partial^2 l}{\partial \alpha \partial \beta}\Big] = -\frac{n}{\beta}$$

$$\text{Fisher Information Matrix,} \quad I(\alpha,\beta) = \begin{bmatrix} n\psi'(\alpha) & -\frac{n}{\beta} \\ -\frac{n}{\beta} & \frac{n\alpha}{\beta^2} \end{bmatrix}$$

$$[I(\alpha,\beta)]^{-1} = \frac{1}{\frac{n^2}{\beta^2}[\alpha\psi'(\alpha)-1]} \begin{bmatrix} \frac{n\alpha}{\beta^2} & \frac{n}{\beta} \\ \frac{n}{\beta} & n\psi'(\alpha) \end{bmatrix} = \frac{1}{\frac{n}{\beta^2}[\alpha\psi'(\alpha)-1]} \begin{bmatrix} \frac{\alpha}{\beta^2} & \frac{1}{\beta} \\ \frac{1}{\beta} & \psi'(\alpha) \end{bmatrix}$$

Using diagonal elements of $[I(\alpha,\beta)]^{-1}$, ML estimator of the parameter α and β are

$$\hat{\alpha} \quad \dot{\sim} \quad N\Big(\alpha, \frac{\alpha}{n[\alpha\psi'(\alpha)-1]}\Big)$$

$$\hat{\beta} \quad \dot{\sim} \quad N\Big(\beta, \frac{\beta^2\psi'(\alpha)}{n[\alpha\psi'(\alpha)-1]}\Big)$$

Question A2: Find Fisher information for a Binomial distribution with probability parameter $p \in (0,1)$

Answer: Here, if we take one observation of Binomial distribution from $X \sim Bin(n,p)$ then

$$\text{likelihood,} \quad L(p) = f(x;p) = \binom{n}{x} p^x (1-p)^{n-x}$$

$$\text{log-likelihood,} \quad l(p) = \ln L(p) = \ln\binom{n}{x} + x\ln p + (n-x)\ln(1-p)$$

$$\frac{\partial l}{\partial p} = \frac{x}{p} - \frac{n-x}{1-p}$$

$$\frac{\partial^2 l}{\partial p^2} = -\frac{x}{p^2} - \frac{n-x}{(1-p)^2}$$

$$\text{Fisher information,} \quad I(p) = \mathbb{E}_p\Big[-\frac{\partial^2 l}{\partial p^2}\Big]$$

$$I(p) = \mathbb{E}_p\Big[\frac{x}{p^2} + \frac{n-x}{(1-p)^2}\Big] = \frac{\mathbb{E}(x)}{p^2} + \frac{\mathbb{E}(n-x)}{(1-p)^2} = \frac{np}{p^2} + \frac{n-np}{(1-p)^2} = \frac{n}{p} + \frac{n}{1-p} = \frac{n}{p(1-p)}$$

(for Binomial random variable $\mathbb{E}(x) = np$). Fisher information $I(p) = \frac{n}{p(1-p)}$

Question B:
Answer: The proportion of zinc can be modeled as a random variable X having the following distribution

$$f(x;\theta) = \theta(1-x)^{\theta-1}, \quad 0 < x < 1$$

$$\text{likelihood,} \quad L(\theta) = \theta^n\Big[\prod_{i=1}^n (1-x_i)\Big]^{\theta-1}$$

$$\text{log-likelihood,} \quad l(\theta) = \ln L(\theta) = n\ln\theta + (\theta-1)\sum_{i=1}^n \ln(1-x_i)$$

$$l'(\theta) = \frac{\partial l}{\partial\theta} = \frac{n}{\theta} + \sum_{i=1}^n \ln(1-x_i)$$

$$l''(\theta) = \frac{\partial l^2}{\partial\theta^2} = -\frac{n}{\theta^2}$$

$$\text{Fisher Information,} \quad I(\theta) = \mathbb{E}(-l''(\theta)) = \frac{n}{\theta^2}$$

$$[I(\theta)]^{-1} = \frac{\theta^2}{n}$$

$$l'(\theta) = 0 \Rightarrow \text{MLE} \quad \hat{\theta} = \frac{n}{-\sum_{i=1}^n \ln(1-x_i)}$$

$$\hat{\theta} \quad \dot{\sim} \quad N(\theta, \frac{\theta^2}{n})$$

$$\text{SE}_\theta(\hat{\theta}) = \sqrt{\frac{\theta^2}{n}}$$

$$\widehat{\text{SE}_\theta(\hat{\theta})} = \sqrt{\frac{\hat{\theta}^2}{n}}$$

Code in R:

```
d = c(1.2,2.7,5.2,0.9,4.8,4.0,4.2,0.8,1.6,1.0,1.6,5.1,4.6,2.7,1.2,1.8,4.8,1.9)/100
nume = length(d)
deno = -sum(log(1-d))
theta_hat = nume/deno
SE_estimate = sqrt(theta_hat^2/length(d))
```

Results:

```
> theta_hat
[1] 35.25505
> SE_estimate
[1] 8.309694
```

(b) the average zinc content of the alloy is $\mathbb{E}(X)$

$$\mathbb{E}(X) = \int_0^1 x\theta(1-x)^{\theta-1}dx = \frac{1}{1+\theta}$$

$$\widehat{\mathbb{E}(X)} = \frac{1}{1+\hat{\theta}} \quad \text{where,} \quad \hat{\theta} \quad \dot{\sim} \quad N(\theta, \frac{\theta^2}{n})$$

from the delta rule, if $\hat{\theta}$ is MLE of θ, then for any differentiable function of θ namely $\psi(\theta)$

$$\psi(\hat{\theta}) \quad \dot{\sim} \quad N\Big(\psi(\theta), \frac{\sigma^2[\psi'(\theta)]^2}{n}\Big)$$

$$\text{where one sample ayamptotic variance of} \quad \hat{\theta} = \sigma^2 = \theta^2$$

$$\text{Let,} \quad \psi(\theta) = \mathbb{E}(X) = \frac{1}{1+\theta}$$

$$\psi'(\theta) = -\frac{1}{(1+\theta)^2}$$

$$\mathbb{E}(X) \quad \dot{\sim} \quad N\Big(\frac{1}{1+\theta}, \frac{\theta^2}{n(1+\theta)^4}\Big)$$

$$\text{SE} = \frac{\theta}{\sqrt{n}(1+\theta)^2}$$

$$\text{hence the estimate,} \quad \widehat{\text{SE}} = \frac{\hat{\theta}}{\sqrt{n}(1+\hat{\theta})^2}$$

Code in R:

```
average_zinc = 1/(1+theta_hat)
SE_average_zinc = theta_hat/(sqrt(length(d))*(1+theta_hat)^2 )
```

Results:

```
> average_zinc*100 # in percentage
[1] 2.758237
> SE_average_zinc
[1] 0.006321907
```
