Name: Chaoran Li
Student ID: (UTD ID) 2021489307; (NET ID) cxl190012
Course Number: CS 6364.002

# Homework 3 Writeup Part

First five questions will be chosen from the ten exercises in the PDF document entitled "MLE questions.pdf". You can choose any 5 from the 10 exercises.

Q1: (1) A sample of six observations, ($X_1$ = 0.4, $X_2$ = 0.7, $X_3$ = 0.9, $X_4$ = 0.6, $X_5$ = 0.5, $X_6$ = 0.7) is collected from a continuous distribution with the density function

$$f(x) = \Theta x^{\Theta-1}, with\ 0 < x < 1.$$

Estimate Θ by the method of maximum likelihood.

Solution:

$$lnL = ln\prod_{i=1}^{N} Prob(X_i:\Theta) = \sum_{i=1}^{N} \ln\big(Prob(X_i:\Theta)\big)$$

$$= \sum_{i=1}^{N} \ln(\Theta x_i^{\Theta-1}) = \sum_{i=1}^{N} \ln\Theta + \sum_{i=1}^{N} \ln(x_i^{\Theta-1}) = Nln\Theta + \sum_{i=1}^{N} \ln(x_i^{\Theta-1})$$

$$\frac{d}{d\Theta} lnL = \frac{N}{\Theta} + \sum_{i=1}^{N} \frac{1}{x_i^{\Theta-1}} x_i^{\Theta-1} \ln(x_i) = \frac{N}{\Theta} + \sum_{i=1}^{N} \ln(x_i)$$

$$= \frac{6}{\Theta} + (\ln(0.4) + \ln(0.7) + \ln(0.9) + \ln(0.6) + \ln(0.5) + \ln(0.7))$$

$$= \frac{6}{\Theta} - 2.9390 = 0$$

$$\Theta = 2.0415$$

$$\frac{d^2}{d^2\Theta} lnL = -\frac{6}{\Theta^2} + 0 < 0$$

$$\therefore MLE(\Theta) = 2.0415$$

Q2: (2) Suppose that $X_1$, …, $X_n$ from a random sample from a uniform distribution on the interval (0, Θ), where of the parameter Θ > 0 but is unknown. Estimate Θ by the method of maximum likelihood.

Solution:

PDF for uniform distribution:

$$f(x) = \begin{cases} \frac{1}{\Theta}, \Theta \in (0, \Theta) \\ 0, otherwise \end{cases}$$

$$lnL = ln\prod_{i=1}^{N} Prob(X_i:\Theta) = \sum_{i=1}^{N} \ln\big(Prob(X_i:\Theta)\big)$$

$$= \sum_{i=1}^{N} \frac{1}{\Theta} = Nln\frac{1}{\Theta}$$

$$\frac{d}{d\Theta} lnL = N\frac{1}{\Theta}^{-1} - \frac{1}{\Theta^2} = -\frac{N}{\Theta} = -\frac{6}{\Theta} < 0$$

No critical point and lnL takes the maximum when Θ takes the minimum.

$$\frac{d^2}{d^2\Theta} lnL = \frac{6}{\Theta^2} > 0$$

$$\therefore MLE(\Theta) = \mathrm{MAX}\{X_1, \dots, X_n\}$$

Q3: (3) Let $X_1$, …, $X_n$ be a random sample from an exponential distribution with the density function

$$f(x) = \frac{1}{\beta} e^{-\frac{x}{\beta}}, with\ 0 < x < \infty.$$

Estimate β by the method of maximum likelihood.
Solution:
From f(x) can we get: β≠0

$$lnL = ln\prod_{i=1}^{N} Prob(X_i:\beta) = \sum_{i=1}^{N} \ln\big(Prob(X_i:\beta)\big)$$

$$= \sum_{i=1}^{N} \ln\left(\frac{1}{\beta}e^{-\frac{x_i}{\beta}}\right) = \sum_{i=1}^{N} \ln\frac{1}{\beta} + \sum_{i=1}^{N} \ln\left(e^{-\frac{x_i}{\beta}}\right) = -\sum_{i=1}^{N} \ln\beta + \sum_{i=1}^{N} -\frac{x_i}{\beta} = -N\ln\beta - \frac{1}{\beta}\sum_{i=1}^{N} x_i$$

$$\frac{d}{d\beta} lnL = -\frac{N}{\beta} + \frac{\sum_{i=1}^{N} x_i}{\beta^2} = 0$$

$$\beta_1 = \frac{\sum_{i=1}^{N} x_i}{N}$$

$$\frac{d^2}{d^2\beta} lnL = \frac{N}{\beta^2} - \frac{2\sum_{i=1}^{N} x_i}{\beta^3} = 0$$

$$\beta_2 = \frac{2\sum_{i=1}^{N} x_i}{N} > \beta_1$$

$$\therefore \frac{d^2}{d^2\beta} lnL < 0, when\ \beta < \beta_2$$

$$\therefore MLE(\beta) = \frac{\sum_{i=1}^{N} x_i}{N} = AVG\{X_1, \dots, X_n\}$$

Q4: (4) Gamma distribution has a density function as

$$f(x) = \frac{1}{\Gamma(\alpha)} \lambda^{\alpha} x^{\alpha-1} e^{-\lambda x}, with\ 0 \le x < \infty.$$

Suppose the parameter α is known, please find the MLE of ƛ based on a random sample $X_1, \dots, X_n$.
Solution:

$$lnL = ln\prod_{i=1}^{N} Prob(X_i:\lambda) = \sum_{i=1}^{N} \ln\big(Prob(X_i:\lambda)\big)$$

$$= \sum_{i=1}^{N} \ln\left(\frac{1}{\Gamma(\alpha)}\lambda^{\alpha} x_i^{\alpha-1} e^{-\lambda x_i}\right) = \sum_{i=1}^{N} \ln\frac{1}{\Gamma(\alpha)} + \sum_{i=1}^{N} \ln(\lambda^{\alpha}) + \sum_{i=1}^{N} \ln(x_i^{\alpha-1}) + \sum_{i=1}^{N} \ln\big(e^{-\lambda x_i}\big)$$

$$= -\sum_{i=1}^{N} \ln\Gamma(\alpha) + \alpha\sum_{i=1}^{N} \ln\lambda + (\alpha-1)\sum_{i=1}^{N} \ln x_i + \sum_{i=1}^{N} -\lambda x_i$$

$$= -\sum_{i=1}^{N} \ln\Gamma(\alpha) + \alpha\sum_{i=1}^{N} \ln\lambda + (\alpha-1)\sum_{i=1}^{N} \ln x_i - \lambda\sum_{i=1}^{N} x_i$$

$$\frac{d}{d\lambda} lnL = 0 + \alpha\sum_{i=1}^{N}\frac{1}{\lambda} + 0 - \sum_{i=1}^{N} x_i = \frac{\alpha N}{\lambda} - \sum_{i=1}^{N} x_i = 0$$

$$\lambda = \frac{\alpha N}{\sum_{i=1}^{N} x_i}$$

$$\frac{d^2}{d^2\lambda} lnL = -\frac{\alpha N}{\lambda^2} < 0$$

$$\therefore MLE(\lambda) = \frac{\alpha N}{\sum_{i=1}^{N} x_i}$$

Q5: (5) Suppose that $X_1, \dots, X_n$ from a random sample from a distribution for which the density function is as follows:

$$f(x) = \frac{1}{2}e^{-|x-\Theta|}, with -\infty < x < \infty.$$

Also suppose that the value of Θ is unknown (-∞ < x < ∞). Find the MLE of Θ.

Solution:

$$lnL = ln\prod_{i=1}^{N} Prob(X_i : \Theta) = \sum_{i=1}^{N} \ln\big(Prob(X_i : \Theta)\big)$$

$$= \sum_{i=1}^{N} \ln\left(\frac{1}{2}e^{-|x_i-\Theta|}\right) = \sum_{i=1}^{N} \ln\left(\frac{1}{2}\right) + \sum_{i=1}^{N} \ln\left(e^{-|x_i-\Theta|}\right) = Nln\left(\frac{1}{2}\right) + \sum_{i=1}^{N} -|x_i - \Theta| = Nln\left(\frac{1}{2}\right) + \sum_{i=1}^{N} -\sqrt{(x_i - \Theta)^2}$$

Assume $x_i \neq \Theta$.

$$\frac{d}{d\Theta}lnL = 0 + \sum_{i=1}^{N} -\frac{1}{2\sqrt{(x_i - \Theta)^2}}2(x_i - \Theta)(-1) = \sum_{i=1}^{N} \frac{(x_i - \Theta)}{\sqrt{(x_i - \Theta)^2}} = 0$$

$\Theta$ will be 0 when the number of $x_i < \Theta$ is equal to $x_i > \Theta$

$$\frac{d^2}{d^2\Theta}lnL \text{ is hard to discuss, because } \frac{d}{d\Theta}lnL \text{ is not continuous.}$$

$$\therefore MLE(\Theta) = Median\{X_1, \dots, X_n\}$$

Let's use MED to represent $Median\{X_1, \dots, X_n\}$ here.

When n is even, assume $X_p <$ MED $< X_{p+1}$, then MLE($\Theta$) can be any element of $(X_p, X_{p+1})$.

Next five questions will be chosen from the exercises in Chapter 2 of PRML. You can choose from the first 30 exercises and 2 from the last 30 exercises in Chapter 2 of PRML, excluding the exercises with the label "www".

Q6: (2.2) The form of the Bernoulli distribution given by

$$Bern(x|\mu) = \mu^x(1-\mu)^{1-x}\text{ , (2.2)}$$

is not symmetric between the two values of x. In some situations, it will be more convenient to use an equivalent formulation for which x ∈ {-1, 1}, in which case the distribution can be written

$$p(x|\mu) = \left(\frac{1-\mu}{2}\right)^{(1-x)/2}\left(\frac{1+\mu}{2}\right)^{(1+x)/2}$$

where μ ∈ {-1, 1}. Show that the distribution is normalized, and evaluate its mean, variance, and entropy.

Solution:

$$p(-1|\mu) = \left(\frac{1-\mu}{2}\right)^{(1+1)/2}\left(\frac{1+\mu}{2}\right)^{(1-1)/2} = \frac{1-\mu}{2}$$

$$p(1|\mu) = \left(\frac{1-\mu}{2}\right)^{(1-1)/2}\left(\frac{1+\mu}{2}\right)^{(1+1)/2} = \frac{1+\mu}{2}$$

$$\therefore p(-1|\mu) + p(1|\mu) = \frac{1-\mu}{2} + \frac{1+\mu}{2} = 1$$

The distribution is normalized.

Mean =

$$E(x) = \sum_{x\in\{-1,1\}} xp(x|\mu) = (-1) * \frac{1-\mu}{2} + (1) * \frac{1+\mu}{2} = \mu$$

Variance =

$$E\left(\big(x - E(x)\big)^2\right) = \sum_{x\in\{-1,1\}} (x-\mu)^2 p(x|\mu)$$

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

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

Entropy =

$$H(x) = -E\big(\log(p(x))\big) = -\sum_{x\in\{-1,1\}} p(x|\mu)\log\big(p(x|\mu)\big)$$

$$= -\frac{1-\mu}{2}\log\left(\frac{1-\mu}{2}\right) - \frac{1+\mu}{2}\log\left(\frac{1+\mu}{2}\right)$$

Q7: (2.4) Show that the mean of the binomial distribution is given by

$$E[m] = \sum_{m=0}^{N} mBin(m|N,\mu) = N\mu, (2.11).$$

To do this, differentiate both sides of the normalization condition

$$\sum_{m=0}^{N} \binom{N}{m} \mu^m (1-\mu)^{N-m} = 1, (2.264)$$

with respect to μ and then rearrange to obtain an expression for the mean of n. Similarly, by differentiating (2.264) twice with respect to μ and making use of the result (2.11) for the mean of the binomial distribution prove the result

$$var[m] = \sum_{m=0}^{N} (m - E[m])^2 Bin(m|N,\mu) = N\mu(1-\mu), (2.12)$$

for the variance of the binomial.

Solution:

$$var[m] = \sum_{m=0}^{N} (m - E[m])^2 Bin(m|N,\mu)$$

$$= \sum_{m=0}^{N} (m - N\mu)^2 \binom{N}{m} \mu^m (1-\mu)^{N-m}$$

$$= \sum_{m=0}^{N} (m^2 - 2mN\mu + N^2\mu^2) \binom{N}{m} \mu^m (1-\mu)^{N-m}$$

$$= \sum_{m=0}^{N} m^2 \binom{N}{m} \mu^m (1-\mu)^{N-m} - 2N\mu \sum_{m=0}^{N} m \binom{N}{m} \mu^m (1-\mu)^{N-m} + N^2\mu^2 \sum_{m=0}^{N} \binom{N}{m} \mu^m (1-\mu)^{N-m}$$

$$= \sum_{m=0}^{N} m^2 \binom{N}{m} \mu^m (1-\mu)^{N-m} - 2N\mu * N\mu + N^2\mu^2 * 1$$

$$= \sum_{m=0}^{N} m^2 \binom{N}{m} \mu^m (1-\mu)^{N-m} - N^2\mu^2$$

$$From\ factors\ of\ Binomial\ Coefficient, we\ have: m\binom{N}{m} = N\binom{N-1}{m-1}$$

$$= \sum_{m=0}^{N} mN \binom{N-1}{m-1} \mu^m (1-\mu)^{N-m} - N^2\mu^2$$

$$= N \sum_{m=0}^{N} m \binom{N-1}{m-1} \mu * \mu^{m-1} (1-\mu)^{(N-1)-(m-1)} - N^2\mu^2$$

$$= N\mu \sum_{m=1}^{N} m \binom{N-1}{m-1} \mu^{m-1} (1-\mu)^{(N-1)-(m-1)} - N^2\mu^2$$

$$= N\mu \sum_{l=m-1=0}^{M=N-1} (l+1) \binom{M}{l} \mu^l (1-\mu)^{M-l} - N^2\mu^2$$

$$= N\mu \sum_{l=0}^{M} l \binom{M}{l} \mu^l (1-\mu)^{M-l} + N\mu \sum_{l=0}^{M} \binom{M}{l} \mu^l (1-\mu)^{M-l} - N^2\mu^2$$

$$= N\mu \sum_{l=0}^{M} l \binom{M}{l} \mu^l (1-\mu)^{M-l} + N\mu * 1 - N^2\mu^2$$

$$= N\mu \sum_{l=0}^{M} M \binom{M-1}{l-1} \mu^{l}(1-\mu)^{M-l} + N\mu * 1 - N^2\mu^2$$
$$= N\mu M \sum_{l=0}^{M} \binom{M-1}{l-1} \mu * \mu^{l-1}(1-\mu)^{(M-1)-(l-1)} + N\mu - N^2\mu^2$$
$$= N(N-1)\mu^2 \sum_{l=0}^{M} \binom{M-1}{l-1} \mu^{l-1}(1-\mu)^{(M-1)-(l-1)} + N\mu - N^2\mu^2$$
$$= N(N-1)\mu^2 * 1 + N\mu - N^2\mu^2$$
$$= N^2\mu^2 - N\mu^2 + N\mu - N^2\mu^2$$
$$= N\mu(1-\mu)$$
Proved.

Q8: (2.6) Make use of the result

$$\int_0^1 \mu^{a-1}(1-\mu)^{b-1} d\mu = \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)}, (2.265)$$

to show that the mean, variance, and mode of the beta distribution

$$Beta(\mu|a,b) = \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} \mu^{a-1}(1-\mu)^{b-1}, (2.13)$$

are given respectively by

$$E[\mu] = \frac{a}{a+b}$$
$$var[\mu] = \frac{ab}{(a+b)^2(a+b+1)}$$
$$mode[\mu] = \frac{a-1}{a+b-2}$$

Solution:

$$E[\mu] = \int_0^1 \mu Beta(\mu|a,b) d\mu$$
$$= \int_0^1 \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} \mu^{a}(1-\mu)^{b-1} d\mu$$
$$= \int_0^1 \frac{\Gamma(a+1)\Gamma(a+b)}{\Gamma(a+1+b)\Gamma(a)} \frac{\Gamma(a+1+b)}{\Gamma(a+1)\Gamma(b)} \mu^{a}(1-\mu)^{b-1} d\mu$$
$$= \frac{\Gamma(a+1)}{\Gamma(a)} \frac{\Gamma(a+b)}{\Gamma(a+b+1)} \frac{\Gamma(a+1+b)}{\Gamma(a+1)\Gamma(b)} \int_0^1 \mu^{a}(1-\mu)^{b-1} d\mu$$

For $\mathbf{\Gamma}(x+1) = x\mathbf{\Gamma}(x)$

$$E[\mu] = a * \frac{1}{a+b} * \frac{\Gamma(a+1+b)}{\Gamma(a+1)\Gamma(b)} * \frac{\Gamma(a+1)\Gamma(b)}{\Gamma(a+1+b)} = \frac{a}{a+b}$$

$$var[\mu] = E[\mu^2] - (E[\mu])^2 = \int_0^1 \mu^2 Beta(\mu|a,b) d\mu - \frac{a^2}{(a+b)^2}$$
$$= \int_0^1 \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} \mu^{a+1}(1-\mu)^{b-1} d\mu - \frac{a^2}{(a+b)^2}$$
$$= \int_0^1 \frac{\Gamma(a+2)\Gamma(a+b)}{\Gamma(a+2+b)\Gamma(a)} \frac{\Gamma(a+2+b)}{\Gamma(a+2)\Gamma(b)} \mu^{a+1}(1-\mu)^{b-1} d\mu - \frac{a^2}{(a+b)^2}$$
$$= \frac{\Gamma(a+2)}{\Gamma(a)} \frac{\Gamma(a+b)}{\Gamma(a+b+2)} \frac{\Gamma(a+2+b)}{\Gamma(a+2)\Gamma(b)} \int_0^1 \mu^{a+1}(1-\mu)^{b-1} d\mu - \frac{a^2}{(a+b)^2}$$
$$= a(a+1) * \frac{1}{(a+b)(a+b+1)} \frac{\Gamma(a+2+b)}{\Gamma(a+2)\Gamma(b)} \frac{\Gamma(a+2)\Gamma(b)}{\Gamma(a+2+b)} - \frac{a^2}{(a+b)^2}$$
$$= \frac{a(a+1)}{(a+b)(a+b+1)} - \frac{a^2}{(a+b)^2}$$

$$= \frac{a\big((a+1)(a+b) - a(a+b+1)\big)}{(a+b)^2(a+b+1)}$$
$$= \frac{a(a^2+ab+a+b-a^2-ab-a)}{(a+b)^2(a+b+1)}$$
$$= \frac{ab}{(a+b)^2(a+b+1)}$$

$$\frac{d}{d\mu} Beta(\mu|a,b)$$
$$= \frac{d}{d\mu}\left(\frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)}\mu^{a-1}(1-\mu)^{b-1}\right)$$
$$= \frac{\boldsymbol{\Gamma}(a+b)}{\boldsymbol{\Gamma}(a)\boldsymbol{\Gamma}(b)}\frac{d}{d\mu}(\mu^{a-1}(1-\mu)^{b-1})$$
$$= \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)}\big((a-1)\mu^{a-2}(1-\mu)^{b-1} - \mu^{a-1}(b-1)(1-\mu)^{b-2}\big)$$
$$= \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)}\mu^{a-2}(1-\mu)^{b-2}\big((a-1)(1-\mu) - \mu(b-1)\big) = 0$$
$$\mu = \frac{a-1}{a+b-2}$$
$$\frac{d^2}{d^2\mu} Beta(\mu|a,b) = \frac{d}{d\mu}\left(\frac{d}{d\mu} Beta(\mu|a,b)\right) < 0 \; when \; a > 1 \; and \; b > 1$$
$$\therefore mode[\mu] = \frac{a-1}{a+b-2}$$

Q9: (2.41) Use the definition of the gamma function

$$\Gamma(x) = \int_0^\infty u^{x-1}e^{-u}du\,, (1.141)$$

to show that the gamma distribution

$$Gam(\lambda|a,b) = \frac{1}{\Gamma(a)}b^a\lambda^{a-1}e^{-b\lambda}, (2.146)$$

is normalized.
Solution:
$$\int_0^\infty Gam(\lambda|a,b)d\lambda$$
$$= \int_0^\infty \frac{1}{\Gamma(a)}b^a\lambda^{a-1}e^{-b\lambda}d\lambda$$
$$= \frac{1}{\Gamma(a)}\int_0^\infty (b\lambda)^{a-1}e^{-b\lambda}bd\lambda$$
$$= \frac{1}{\Gamma(a)}\Gamma(a) = 1$$
Gamma distribution is normalized.

Q10: (2.42) Evaluate the mean, variance, and mode of the gamma distribution

$$Gam(\lambda|a,b) = \frac{1}{\Gamma(a)}b^a\lambda^{a-1}e^{-b\lambda}, (2.146).$$

Solution
$$E[\lambda] = \int_0^\infty \lambda Gam(\lambda|a,b)d\lambda$$
$$= \int_0^\infty \frac{1}{\Gamma(a)}b^a\lambda^a e^{-b\lambda}d\lambda$$

$$= \frac{b^a}{\boldsymbol{\Gamma}(a)} \int_0^\infty ƛ^a e^{-bƛ} dƛ$$
$$= \frac{b^a}{\Gamma(a)} \int_0^\infty \left(\frac{bƛ}{b}\right)^a e^{-bƛ} \frac{d(bƛ)}{b}$$
$$= \frac{b^a}{\Gamma(a)} * \frac{1}{b^{a+1}} \int_0^\infty (bƛ)^a e^{-bƛ} d(bƛ)$$
$$= \frac{1}{b\Gamma(a)} \Gamma(a+1)$$
$$= \frac{a}{b}$$
$$var[ƛ] = E[ƛ^2] - (E[ƛ])^2$$
$$= \int_0^\infty ƛ^2 Gam(ƛ|a,b) dƛ - \left(\frac{a}{b}\right)^2$$
$$= \int_0^\infty \frac{1}{\Gamma(a)} b^a ƛ^{a+1} e^{-bƛ} dƛ - \left(\frac{a}{b}\right)^2$$
$$= \frac{b^a}{\Gamma(a)} \int_0^\infty ƛ^{a+1} e^{-bƛ} dƛ - \left(\frac{a}{b}\right)^2$$
$$= \frac{b^a}{\Gamma(a) b^{a+2}} \int_0^\infty (bƛ)^{a+1} e^{-bƛ} d(bƛ) - \left(\frac{a}{b}\right)^2$$
$$= \frac{1}{b^2 \Gamma(a)} \Gamma(a+2) - \left(\frac{a}{b}\right)^2$$
$$= \frac{a(a+1)}{b^2} - \left(\frac{a}{b}\right)^2$$
$$= \frac{a^2 + a - a^2}{b^2} = \frac{a}{b^2}$$

$$\frac{d}{dƛ} Gam(ƛ|a,b)$$
$$= \frac{d}{dƛ} \left(\frac{1}{\Gamma(a)} b^a ƛ^{a-1} e^{-bƛ}\right)$$
$$= \frac{b^a}{\Gamma(a)} \left((a-1)ƛ^{a-2} e^{-bƛ} + ƛ^{a-1}(-b) e^{-bƛ}\right)$$
$$= \frac{b^a}{\Gamma(a)} ƛ^{a-2} e^{-bƛ} \left((a-1) + ƛ(-b)\right) = 0$$
$$ƛ = \frac{a-1}{b}$$
$$\frac{d^2}{d^2ƛ} Gam(ƛ|a,b) = \frac{d}{dƛ}\left(\frac{d}{dƛ} Gam(ƛ|a,b)\right) < 0 \; when \frac{a}{b} \geq 1$$
$$\therefore mode[ƛ] = \frac{a-1}{b}$$