Solutions for Midterm I STA 131B, Winter 2016

Exam Rules: This exam is closed book and closed notes. Use of calculators, cell phones or other communication devices is not allowed. You must show all of your work to receive credit.

Name : ______________________ **Signature** : ______________________

ID : ______________________

1. (65 points) Suppose a random sample $X_1, \dots, X_n$ is taken from a negative binomial distribution $NB(r, \theta)$. That is, $P(X = x) = \binom{r+x-1}{x}\theta^r(1-\theta)^x$, for $x = 0, 1, 2, ...$

   Note that the mean of $NB(r, \theta)$ is $\frac{r(1-\theta)}{\theta}$ and the variance is $\frac{r(1-\theta)}{\theta^2}$.

   Assume in this problem that $r = 2$ is known but θ is unknown.

   a) Find the M.L.E. of θ.
   Answer: The likelihood function is

   $$f(x_1, \dots, x_n|\theta) = \prod_{i=1}^{n} \binom{r+x_i-1}{x_i}\theta^r(1-\theta)^{x_i} = \prod_{i=1}^{n} \binom{r+x_i-1}{x_i}\theta^{nr}(1-\theta)^{\sum_{i=1}^n x_i}.$$

   The log-likelihood function is

   $$L(\theta) = \log f(x_1, \dots, x_n|\theta) = nr\log\theta + \sum_{i=1}^{n}\log(1-\theta) + \log(\prod_{i=1}^{n}\binom{r+x_i-1}{x_i}).$$

   Let

   $$\frac{dL(\theta)}{d\theta} = \frac{nr}{\theta} - \frac{\sum_{i=1}^n x_i}{1-\theta} = \frac{nr - (nr + \sum_{i=1}^n x_i)\theta}{\theta(1-\theta)} = 0$$

   we have $\hat\theta = \dfrac{nr}{nr + \sum_{i=1}^n x_i} = \dfrac{2n}{2n + \sum_{i=1}^n x_i}$.
   Since $\dfrac{dL^2(\theta)}{d\theta^2} = -\dfrac{nr}{\theta^2} - \dfrac{\sum_{i=1}^n x_i}{(1-\theta)^2} < 0$, $\hat\theta$ is the M.L.E. of θ.

   b) Find the M.L.E. of $E(X_1)$.
   Answer: By the invariance property, M.L.E. of $E(X_1)$ is

   $$\frac{r(1-\hat\theta)}{\hat\theta} = \frac{2(1-2n/(2n+\sum_{i=1}^n x_i))}{2n/(2n+\sum_{i=1}^n x_i)} = \frac{\sum_{i=1}^n x_i}{n} = \bar{x}_n.$$

   c) Find the M.L.E. of θ, if you were told that $\theta > 0.5$. Justify your answer rigorously.
   Answer:
   i) If $\sum_{i=1}^n x_i < 2n$, then $\frac{2n}{2n+\sum_{i=1}^n} > 0.5$, so $\hat\theta$ is the M.L.E.
   ii) If $\sum_{i=1}^n x_i \geq 2n$, then $\frac{dL(\theta)}{d\theta} < 0$ for $\theta > 0.5$, it implies that $L(\theta)$ is strictly decreasing for $\theta > 0.5$. Thus, the M.L.E. does not exist.

d) Find a method of moment estimator for θ.

Answer: Since $E(X_1) = \frac{r(1-\theta)}{\theta}$, let $\frac{r(1-\theta)}{\theta} = \bar{x}_n$, we get the MoM estimator $\tilde{\theta} = r/(r + \bar{x}_n)$.

e) Find another method of moment estimator for θ.

Answer: Note that $\dfrac{E(X_1)}{\text{var}X_1} = \theta$, then the MoM estimator is $\dfrac{\bar{x}_n}{\bar{x_n^2} - \bar{x}_n^2} = \dfrac{m_1}{m_2 - m_1^2}$.

2. (35 points) Let X be a random variable with negative binomial distribution $NB(r, \theta)$. Assume that r is known but θ is unknown. Let the prior distribution of θ be a Beta distribution, $Be(\alpha, \beta)$, with p.d.f. $f(x|\alpha, \beta) = \frac{1}{B(\alpha,\beta)} x^{\alpha-1}(1-x)^{\beta-1}$, for $0 \leq x \leq 1$, where $B(\alpha, \beta)$ is the beta function with $\alpha > 0$ and $\beta > 0$. Note that the mean of $Be(\alpha, \beta)$ is $\frac{\alpha}{\alpha+\beta}$.

a) Suppose only a single observation X_1 was taken from $NB(r, \theta)$. Show that the posterior distribution of θ is also a beta distribution and find the parameters of this beta distribution.

Answer: The likelihood function is $\binom{r+x-1}{x}\theta^r(1-\theta)^x$.

The posterior p.d.f. of θ is:

$$\xi(\theta|x) \propto f(x|\theta)\xi(\theta) \propto \theta^r(1-\theta)^x\theta^{\alpha-1}(1-\theta)^{\beta-1} = \theta^{r+\alpha-1}(1-\theta)^{x+\beta-1}$$

Thus, the posterior distribution of θ is a beta distribution with parameters $r + \alpha$ and $x + \beta$.

b) Find the Bayes estimator for θ with respect to the squared error loss function.

Answer: The Bayes estimator for θ is the mean of the posterior distribution, which is $(r+\alpha)/(r+\alpha+x+\beta)$.

c) True or false for each of the following is correct?

_____ (i) The family of beta distribution is a conjugate family of prior distributions for samples from a negative binomial distribution with a known value of the parameter r and an unknown value of the parameter p.

Answer: True. The likelihood based on n observations is $f_n(\mathbf{x}|p) \propto p^{nr}(1-p)^{\sum_i x_i}$, the posterior p.d.f. of p is $\xi(p|\mathbf{x}) \propto f_n(\mathbf{x}|p)\xi(p) \propto p^{nr}(1-p)^{\sum_i x_i}p^{\alpha-1}(1-p)^{\beta-1} = p^{nr+\alpha-1}(1-p)^{\sum_i x_i+\beta-1} \sim Be(r+\alpha, \sum_i x_i + \beta)$.

_____ (ii) The family of beta distribution is a conjugate family of prior distributions for samples from a negative binomial distribution whether the value of r is known or not.

Answer: False. The posterior p.d.f. of p is not the p.d.f. of any beta distribution if r is unknown.

_____ (iii) The family of negative binomial distribution with a known value of the parameter r and an unknown value of the parameter p is a conjugate family of prior distributions for samples from a beta distribution.

Answer: False. The beta distribution has two parameters whereas the negative binomial distribution has one.