36 - 226 Introduction to Statistical Inference

Homework assignment 9

Due: Wednesday, March 27, 2013

- Write your full name, the course number, and the homework number at the top of each page.
- **STAPLE** your entire assignment together with a staple.
- Write clearly. Electronic submission of homework assignments is not accepted.

1. Sometimes we can have more than one method of moments estimator for a parameter (often happens when the same parameter is part of multiple population moments).

   Let Y_i be i.i.d. Gamma$(4, \beta)$ with $f_Y(y) = \frac{1}{6\beta^4} y^3 e^{-\frac{y}{\beta}}$ where $y > 0$.

   (a) Find the MVUE using the Factorization Theorem and the Rao-Blackwell Theorem, and find the MLE. Verify that these are the same.

   (b) Find the Cramer-Rao Lower Bound for $f_Y(y)$. Verify that the MVUE and MLE achieve the bound.

   (c) Use the first two moments to find two different estimates for β using the method of moments.

   (d) Verify that the method of moments estimator based on the first moment is unbiased and achieves the CRLB.

2. Wackerly 10.4.

3. Read pages 279 - 282. Wackerly 5.119.

4. Professor Greenhouse will be covering material related to this problem. According to genetic theory, blood types MM, NM, and NN should occur in a very large population with probability θ^2, $2\theta(1-\theta)$, and $(1-\theta)^2$, where θ is the unknown gene frequency. In other words the probability that a person randomly selected from the population is MM is $P(MM) = \theta^2$; is NM is $P(NM) = 2\theta(1-\theta)$; and is NN is $P(NN) = (1-\theta)^2$.

   Suppose you select a random sample of size n from this population and f_1 have blood type MM, f_2 have blood type NM, and f_3 have blood type NN, where $n = f_1 + f_2 + f_3$.

   (a) Write down the likelihood function and the log-likelihood function for θ.

   (b) Find $\hat{\theta}$, the maximum likelihood estimator of θ. Also find the second derivative of the log-likelihood with respect to θ. Is $\hat{\theta}$ a maximum?

   (c) Show that the observed information is $i(\hat{\theta}) = \frac{2n}{\hat{\theta}(1-\hat{\theta})}$. Find $V(\hat{\theta})$.

   (d) What is the approximate probability distribution of $\hat{\theta}$?

   Suppose a random sample of size $n = 500$ is taken from the population and each subject's blood type is determined. The result for the sample are given in the following table:

Blood type	MM	MN	NN	
Frequency	$f_1 = 125$	$f_2 = 225$	$f_3 = 150$	$n = 500$

(e) Based on this sample, what are $\hat{\theta}$ and $V(\hat{\theta})$? Find a 95% large sample confidence interval for θ.

(f) Based on your confidence interval in part (e), would you conclude that θ, the unknown gene frequency, is $\frac{1}{2}$? Explain.

5. Wackerly 9.98. Professor Greenhouse will be covering material related to this problem.

6. Read pages 483 - 484. Wackerly 9.100. Professor Greenhouse will be covering material related to this problem.