QUIZ # 3

Statistics 305

Term 2, 2006-2007

Tuesday, February 27, 2007

Time: 2:00pm – 2:30pm

Student Name (**Please print in caps**): SOLUTIONS

Student Number: ______________________

Notes:

- This quiz has 3 problems on the 4 following pages, plus 1 page of statistical tables. Check to ensure that you have a complete paper.
- The amount each part of each question is worth is shown in [ ] on the left-hand side of the page.
- Where appropriate, record your answers in the blanks provided on the right-hand side of the page.
- Your solutions **must be justified**; show **all the work** and state **all the reasons** leading to your answer for each question in the space provided immediately under the question.
- Clear and complete solutions are essential; little partial credit will be given.
- This is a closed book exam.
- A single one-sided 8.5 x 11 page of notes is allowed.
- Calculators are allowed (but not for symbolic differentiation or integration).
- No devices (including calculators) that can store text or send/receive messages are allowed.

Problem	Total Available	Score
1.	6	
2.	7	
3.	12	
Total	25	

1. Consider $T = T(X_1, X_2, ..., X_n)$, any estimator of the parameter θ. As discussed in class, the mean squared error and bias of T are defined as $MSE(T) = E[(T-\theta)^2]$ and $Bias(T) = E(T) - \theta$.

[4] a) Show that $MSE(T) = Var(T) + [Bias(T)]^2$. ________________

$$
\begin{aligned}
MSE(T) &= E\left[(T-\theta)^2\right] \\
&= E\left[\left(T-E(T)+E(T)-\theta\right)^2\right] \\
&= E\left[\left(T-E(T)\right)^2 + 2\left(T-E(T)\right)\left(E(T)-\theta\right) + \left(E(T)-\theta\right)^2\right] \\
&= E\left[\left(T-E(T)\right)^2\right] + 2\left(E(T)-\theta\right)E\left(T-E(T)\right) + E\left(\left(E(T)-\theta\right)^2\right) \\
&= Var(T) + 2\left(E(T)-\theta\right)\cdot\phi + \left(E(T)-\theta\right)^2 \\
&= Var(T) + Bias^2(T)
\end{aligned}
$$

(since there is nothing random in that term: E(constant) = constant)

[2] b) Explain the practical importance of both terms in the above expression for $MSE(T)$ when evaluating the performance of the estimator T. ________________

- Var(T) describes the precision of the estimator T: how variable is T in repeated samples of size n? If highly variable, the estimator could take a value far from the target in any particular sample of size n, even if the estimator is unbiased.
- Bias(T) describes the accuracy of the estimator T: how far "off the target" is the estimator, on average, in repeated samples of size n?

2. Suppose Y has a binomial distribution with parameters n and θ; that is,

$$P(Y = y) = \binom{n}{y}\theta^y(1-\theta)^{n-y} \qquad \text{for } y = 0, 1, \ldots, n.$$

Assume the value of n is known but θ is an unknown parameter.

[5] a) Find the expression for $\hat{\theta}_{ML}$, the maximum likelihood estimator (MLE) of θ based on the single binomial random variable Y.

$$\underline{\hat{\theta}_{ML} = Y/n}$$

$$L(\theta) = \binom{n}{y}\theta^y(1-\theta)^{n-y}$$

$$\Rightarrow \ell(\theta) = \log L(\theta) = \log\binom{n}{y} + y\log\theta + (n-y)\log(1-\theta)$$

$$\Rightarrow \ell'(\theta) = \frac{y}{\theta} - \frac{(n-y)}{1-\theta}$$

$\boxed{\text{Inequality remains the same because } 0<\theta<1}$

$$\ell'(\theta) > 0 \Leftrightarrow \frac{y}{\theta} > \frac{n-y}{1-\theta} \Leftrightarrow y(1-\theta) > (n-y)\theta$$

$$\Leftrightarrow y - \cancel{y\theta} > n\theta - \cancel{y\theta}$$

$$\Leftrightarrow \theta < y/n$$

$$\Rightarrow \hat{\theta}_{ML} = \frac{Y}{n}$$ and it is clear from this that this does correspond to a max of $\ell(\theta)$ $\Rightarrow$ a max of $L(\theta)$

Alternately, $\ell''(\theta) = -y/\theta^2 - \frac{(n-y)}{(1-\theta)^2} < 0$

In particular, $\ell''(y/n) = \frac{-n^3}{y(n-y)} < 0$

$\Rightarrow \theta = y/n$ corresponds to a value of θ that maximizes $\ell(\theta)$ $\Rightarrow$ maximizes $L(\theta)$

[2] b) Provide a rough sketch of the likelihood function for $n = 10$ and $y = 5$.

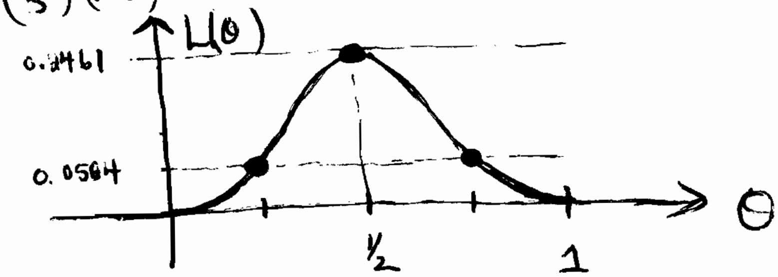
$$L(\theta) = \binom{10}{5}[\theta(1-\theta)]^5 \text{ for } 0<\theta<1$$

$\Rightarrow$ symmetric around $\theta = 1/2$

$\sim \binom{10}{5}\theta^5$ as $\theta \downarrow 0$

$\sim \binom{10}{5}(1-\theta)^5$ as $\theta \uparrow 1$

$\Rightarrow$



3. Suppose $X_1, X_2, \ldots, X_n$ is a simple random sample from an exponential distribution with a rate of θ; that is, from the density function:

$$f_\theta(x) = \theta \exp(-\theta x) \qquad \text{for } x > 0.$$

[3] a) Find the expression for $\hat{\theta}_{MM}$, the method of moments estimator (MME) of θ. $\hat{\theta}_{MM} = \frac{1}{\bar{X}}$

This density is a Gamma$(1, \theta)$ $\Rightarrow$ $\underset{\mu_1}{\overset{|||}{E(X)}} = \frac{1}{\theta}$, the 1st population moment

$$\Rightarrow \theta = \frac{1}{\mu_1}$$

$$\Rightarrow \hat{\theta}_{MM} = \frac{1}{\hat{\mu}_1} \leftarrow \text{(the 1st sample moment}$$

$$\Rightarrow \hat{\theta}_{MM} = \frac{1}{\bar{X}}$$

[5] b) Find an approximate expression for the variance of the MME $\hat{\theta}_{MM}$. $Var(\hat{\theta}_{MM}) \approx \frac{\theta^2}{n}$

For any random variable Y, the delta method yields

$$Var[g(Y)] \approx [g'(E(Y))]^2 \cdot Var(Y)$$

We have $\hat{\theta}_{MM} = \frac{1}{\bar{X}} \equiv g(\bar{X})$, where $g(x) = \frac{1}{x}$ $\rightarrow g'(x) = -\frac{1}{x^2}$

$$\Rightarrow Var(\hat{\theta}_{MM}) \approx \left[-\frac{1}{E(\bar{X})^2}\right]^2 \cdot Var(\bar{X})$$

$$= \left[\frac{1}{E(X)}\right]^4 \cdot \frac{Var(X)}{n}$$

But $Var(\bar{X}) \overset{\text{Always!}}{=} \frac{1}{n} Var(X)$

and $Var(X) = \frac{1}{\theta^2}$ for $G(1, \theta)$

$$= \theta^4 \cdot \frac{1}{\theta^2 n} = \frac{\theta^2}{n}$$

Also $E(\bar{X}) \underset{\text{Always!}}{=} E(X) = \frac{1}{\theta}$ for $G(1, \theta)$

[4] c) Find a second-order approximate expression for the bias of the MME $\hat{\theta}_{MM}$. $\text{Bias}(\hat{\theta}_{MM}) \approx \theta/n$

For any random variable Y, the delta method yields

As in b), $$E[g(Y)] \approx g(E(Y)) + \frac{1}{2} g''(E(Y)) \cdot \text{Var}(Y)$$

We have $\hat{\theta}_{MM} \equiv g(\bar{X})$, where $g(x) = \frac{1}{x} \Rightarrow g''(x) = \frac{2}{x^3}$

$$\Rightarrow E(\hat{\theta}_{MM}) \approx \frac{1}{E(\bar{X})} + \frac{1}{2} \frac{2}{[E(\bar{X})]^3} \text{Var}(\bar{X})$$

$$= \theta + \theta^3 \cdot \frac{1}{\theta^2 \cdot n}$$

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

$$\Rightarrow \text{Bias}(\hat{\theta}_{MM}) = E(\hat{\theta}_{MM}) - \theta$$

$$\approx \theta + \theta/n - \theta = \underline{\underline{\theta/n}}$$

In particular, Bias $\to 0$ as $n \to \infty$ $\to$ asymptotically unbiased!

Note that $\text{Bias}^2(\hat{\theta}_{MM}) \approx \theta^2/n^2$

$$\to \text{MSE}(\hat{\theta}_{MM}) = \text{Var}(\hat{\theta}_{MM}) + \text{Bias}^2(\hat{\theta}_{MM})$$

$$\approx \theta^2/n + \theta^2/n^2 = \theta^2/n \text{ to leading order (for large } n)$$

$\to$ For large n, MSE$(\hat{\theta}_{MM})$ is completely determined by the variance because $\hat{\theta}_{MM}$ has bias $\sim 1/n$

THE END

(This is typical for reasonable estimators: if they are biased, the bias contribution to MSE is negligible for large n.)