THE AUSTRALIAN NATIONAL UNIVERSITY
SCHOOL OF FINANCE AND APPLIED STATISTICS

STATISTICAL INFERENCE - STAT3013/STAT8027

Mid-Semester Examination 2001

Total Marks: 50

Reading Period: 15 Minutes
Time Allowed: Two Hours
Permitted Materials: Course Brick, Lecture Notes, Non-Programmable Calculator

---

Question 1

Indicate whether each statement is either TRUE or FALSE (please write the entire word, as hasty F's can sometimes look like T's). Two marks will be awarded for each correct response, two marks will be deducted for each incorrect response, and zero marks will be given for no response.

(a) If $T = t(X_1, \ldots, X_n)$ is a UMVU estimator of $\tau(\theta)$, then $Var_\theta(T) = \frac{\{\tau'(\theta)\}^2}{I(\theta)}$, where $I(\theta) = ni(\theta)$ is the Fisher information for θ based on the data $X_1, \ldots, X_n$.

(b) If $T = t(X_1, \ldots, X_n)$ is an unbiased estimator of $\tau(\theta)$ with $Var_\theta(T) = \frac{\{\tau'(\theta)\}^2}{I(\theta)}$, where $I(\theta) = ni(\theta)$ is the Fisher information for θ based on the data $X_1, \ldots, X_n$, then, provided the standard regularity conditions are satisfied, T is a UMVU estimator of $\tau(\theta)$.

(c) If $\hat{\theta} = \hat{\theta}(X_1, \ldots, X_n)$ is an unbiased estimator of θ with variance equal to the Cramér-Rao lower bound, then the δ-method variance of $\hat{\tau} = \tau(\hat{\theta})$, an estimator of $\tau(\theta)$, is again equal to the Cramér-Rao lower bound and thus cannot be an overestimate of the true variance.

(d) If we observe an *iid* sample, $X_1, \ldots, X_n$, from a distribution having *CDF* F and we estimate some functional of the *CDF*, $\theta(F)$, using the estimator $\theta(\hat{F})$, where $\hat{F}$ is the empirical distribution function, then the bootstrap bias estimate of the estimator $\theta(\hat{F})$, based on B re-samples, will not be exactly zero even if $\theta(\hat{F})$ is truly unbiased [i.e., if $E_F\{\theta(\hat{F})\} = \theta(F)$ for any distribution F]; however, if $\theta(\hat{F})$ is truly unbiased then the bootstrap bias estimate will tend towards zero as B increases.

(e) If θ is a parameter taking values in a finite interval $[a, b]$ and we assign a uniform prior to θ, $\pi(\theta) = (b - a)^{-1}$ for $\theta \in [a, b]$, this prior can be used to represent the idea of having "no prior information" about the parameter θ.

[10 marks]

Question 2

Let $X_1, \ldots, X_n$ be an *iid* sample from a Poisson distribution with rate parameter θ and let the parameter of interest be $\tau(\theta) = e^{-2\theta}$.

(a) Show that the estimator $T_1 = \frac{1}{n}\sum_{i=1}^{n}(-1)^{X_i}$ is unbiased for $\tau(\theta)$. **[4 marks]**

(b) The conditional distribution of X_i given $S = \sum_{j=1}^{n} X_j$ can be shown to be binomial with parameters S and n^{-1}. Use this fact to find a UMVU estimator of $\tau(\theta)$. [HINT: Recall that $\sum_{x=0}^{k} \frac{k!}{x!(k-x)!} a^x b^{k-x} = (a+b)^k$.] **[6 marks]**

(c) The sequences $Z_n = \sqrt{n}(T_1 - e^{-2\theta})$ and $W_n = \sqrt{n}(T_U - e^{-2\theta})$, where T_U is the UMVU estimator, are known to be asymptotically normal. Furthermore, $Var_\theta(T_U) = e^{-4\theta}(e^{4\theta/n} - 1)$. Show that the asymptotic relative efficiency of T_1 with respect to T_U is $e_{T_1,T_U} = \frac{4\theta}{e^{4\theta}-1}$ and interpret this quantity. [HINT: Recall that $\lim_{n\to\infty} n(e^{a/n} - 1) = a$.] **[5 marks]**

Question 3

The DooDad Manufacturing Company sells its product in lots of size N. Furthermore, they claim that their DooDad-making machines produce a known proportion of defectives, p. You have just been hired as Chief DooDad Purchaser for the West Australian Whatzit-Works Corporation, and you take your new duties very seriously. In particular, you need to estimate the actual number of defective DooDads, D, in the most recent lot you purchased. You decide upon a Bayesian analysis procedure, and you pick a binomial prior for D, with parameters N and p.

(a) Briefly discuss the sensibility of this choice of prior distribution. **[4 marks]**

(b) Suppose that you sample n DooDads from your lot of size N and test how many are defective. The distribution of d, the number of defective DooDads, is hypergeometric so that:

$$f(d|D) = \frac{D!(N-D)!n!(N-n)!}{d!(D-d)!(n-d)!(N-D-n+d)!N!}, \qquad 0 \leq d \leq D.$$

If you observe no defective DooDads, find the posterior Bayes estimate of D. **[7 marks]**

(c) Interpret your result in part (*b*) and suggest what the posterior Bayes estimate would have been if you had observed 2 defective DooDads instead of none. **[4 marks]**

Question 4

Let $X_1, \ldots, X_n$ be an *iid* sample from a distribution F. We want to estimate a functional of this distribution, $\theta(F)$, and we choose the estimator $\theta(\hat{F})$, were $\hat{F}$ is the empirical distribution function.

(a) Briefly discuss how you would use the bootstrap to estimate the *MSE* of $\theta(\hat{F})$. **[5 marks]**

(b) Now, suppose we know that, for any distribution F:

$$E_F\{\theta(\hat{F})\} = \theta(F) + \frac{a(F)}{n},$$

where $a(F)$ is a functional of the distribution F which does not depend on the value of n, and which itself satisfies:

$$E_F\{a(\hat{F})\} = a(F) + \frac{b(F)}{n},$$

where $b(F)$ is again a functional of the distribution F which does not depend on the value of n. Calculate the bias of $\tilde{\theta}_B = 2\theta(\hat{F}) - \frac{1}{B}\sum_{b=1}^{B}\theta(\hat{F}_b^\star)$, the usual bootstrap bias-corrected estimator, and discuss your result. [HINT: Recall that the law of the iterated expectation shows that $E_F\{g(\hat{F}^\star)\} = E_F\big(E_{\hat{F}}\{g(\hat{F}^\star)\}\big)$.] **[5 marks]**

END OF EXAMINATION