THE AUSTRALIAN NATIONAL UNIVERSITY
SCHOOL OF FINANCE AND APPLIED STATISTICS

STATISTICAL INFERENCE - STAT3013/STAT8027

Mid-Semester Examination 2001 - SOLUTIONS

Total Marks: 50

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

---

Question 1

(a) FALSE - UMVU estimators need not have variances which achieve the Cramér-Rao lower bound.

(b) TRUE - If an estimator has a variance which achieves the Cramér-Rao lower bound it will be UMVU (provided the appropriate regularity conditions are satisfied).

(c) FALSE - The δ-method variance will indeed be equal to the Cramér-Rao lower bound for unbiased estimators of $\tau(\theta)$; however, $\hat{\tau} = \tau(\hat{\theta})$ is not necessarily an *unbiased* estimator of $\tau(\theta)$ and thus its variance may actually be smaller than the bound.

(d) TRUE - If $\theta(\hat{F})$ is truly unbiased, then $E_F\{\theta(\hat{F})\} = \theta(F)$ for any distribution F. In particular, if we set $F = \hat{F}$ in this identity, we see that $E_{\hat{F}}\{\theta(\hat{F}^\star)\} = \theta(\hat{F})$, where $\hat{F}^\star = \widehat{\hat{F}}$ is the empirical distribution function of a re-sample from the original sample data. Thus, the theoretical bootstrap bias estimate is $E_{\hat{F}}\{\theta(\hat{F}^\star)\} - \theta(\hat{F}) = 0$. Of course, the actual bootstrap bias estimate based on B re-samples is $\frac{1}{B}\sum_{b=0}^{B} \theta(\hat{F}_b^\star) - \theta(\hat{F})$ which will not be exactly zero, though it will approach zero as B approaches infinity.

(e) FALSE - Uniform or "flat" priors still represent some prior information about the parameter. In particular, if we examine any non-linear reparameterisation of θ, the transformation of the flat prior indicates that we do have some non-uniform prior for that new parameter.

[10 marks]

Question 2

(a) We have:

$$E_\theta(T_1) = E_\theta\left\{\frac{1}{n}\sum_{i=1}^{n}(-1)^{X_i}\right\} = \frac{1}{n}\sum_{i=1}^{n}E_\theta\{(-1)^{X_i}\} = E_\theta\{(-1)^{X_1}\} = \sum_{x=0}^{\infty}(-1)^x\frac{\theta^x e^{-\theta}}{x!}$$
$$= e^{-\theta}\sum_{x=0}^{\infty}\frac{(-\theta)^x}{x!} = e^{-\theta}(e^{-\theta}) = e^{-2\theta} = \tau(\theta).$$

[NOTE: It is tempting to write $E\{(-1)^{X_i}\} = E\{e^{\ln(-1)X_i}\}$ and then use the *mgf* of the Poisson distribution, $m_X(t) = \exp\{\lambda(e^t - 1)\}$. While this approach appears to "give the right answer", we must be careful since $\ln(-1)$ does not actually exist!]

(b) Since the Poisson distributions form a one-parameter exponential family with $d_1(x) = x$ (see Example 2.8 on page 26 of the course notes), we know that $S = \sum_{j=1}^{n} d_1(X_j) = \sum_{j=1}^{n} X_j$ is a complete and sufficient statistic. Therefore, the UMVU for $\tau(\theta)$ can be found as:

$$\begin{aligned}
T_U = E_\theta(T_1|S) &= E_\theta\left\{\frac{1}{n}\sum_{i=1}^{n}(-1)^{X_i}\middle|S\right\} = E_\theta\{(-1)^{X_i}|S\} \\
&= \sum_{x=0}^{S}(-1)^x\frac{S!}{x!(S-x)!}\left(\frac{1}{n}\right)^x\left(1-\frac{1}{n}\right)^{S-x} \\
&= \sum_{x=0}^{S}\frac{S!}{x!(S-x)!}\left(-\frac{1}{n}\right)^x\left(1-\frac{1}{n}\right)^{S-x} \\
&= \left\{\left(-\frac{1}{n}\right)+\left(1-\frac{1}{n}\right)\right\}^S \\
&= \left(\frac{n-2}{n}\right)^S,
\end{aligned}$$

where the fourth equality follows from the given conditional distribution of X_i given S and the penultimate equality follows from the fact given in the hint.

(c) To find the asymptotic relative efficiency, we need to calculate the limit of the variances of the sequences Z_n and W_n. Now, for W_n, we see (using the hint) that

$$\lim_{n\to\infty} Var_\theta(W_n) = \lim_{n\to\infty} nVar_\theta(T_U) = \lim_{n\to\infty} ne^{-4\theta}(e^{4\theta/n} - 1) = 4\theta e^{-4\theta}.$$

Next, we note that $Var_\theta\{(-1)^{X_i}\} = E_\theta\{(-1)^{2X_i}\} - [E_\theta\{(-1)^{X_i}\}]^2 = 1 - e^{-4\theta}$. Thus,

$$\lim_{n\to\infty} Var_\theta(Z_n) = \lim_{n\to\infty} nVar_\theta(T_1) = \lim_{n\to\infty} Var_\theta\{(-1)^{X_i}\} = 1 - e^{-4\theta}.$$

Therefore, the asymptotic relative efficiency of T_1 with respect to T_U is

$$e_{T_1,T_U} = \frac{4\theta e^{-4\theta}}{1-e^{-4\theta}} = \frac{4\theta}{e^{4\theta}-1},$$

which is close to one for small θ's but which can be much less than one for large θ's. Thus, for small values of θ both estimators are approximately asymptotically equivalent in terms of their efficiency (though the UMVU is still slightly more efficient, as it must be since it has uniformly minimum variance); however, for large values of θ, the UMVU estimator is much more efficient.

Question 3

(a) If we truly believe the manufacturer's claim that the proportion of defectives from their machine is p, and we assume that lots are made up of randomly chosen DooDads, then the binomial prior seems precisely the correct one. Of course, if we don't believe their claim, we might put a prior on the parameter p itself, and then use this to compose a different prior for D (which would lead to an "over-dispersed" binomial, for those who know this term).

(b) The posterior distribution is given by:

$$\begin{aligned}\pi(D|d) &= \frac{f(d|D)\pi(D)}{\sum_{D=0}^{N} f(d|D)\pi(D)} \\ &= C(d)\left\{\frac{D!(N-D)!n!(N-n)!}{d!(D-d)!(n-d)!(N-D-n+d)!N!}\right\}\left\{\frac{N!}{D!(N-D)!}p^D(1-p)^{N-D}\right\} \\ &= C(d)\frac{n!(N-n)!}{d!(D-d)!(n-d)!(N-D-n+d)!}p^D(1-p)^{N-D},\end{aligned}$$

where $\{C(d)\}^{-1} = \sum_{D=0}^{N} f(d|D)\pi(D)$. Now, when $d = 0$, this posterior reduces to:

$$\pi(D|d=0) = C(0)\frac{(N-n)!}{D!(N-D-n)!}p^D(1-p)^{N-D},$$

which clearly has the form of a binomial distribution with parameters $N-n$ and p, meaning that $C(0) = (1-p)^{-n}$. As such, the posterior Bayes estimator, which is just the mean of the posterior distribution, is $\hat{D}_\pi = E(D|d=0) = (N-n)p$.

(c) The result is seen to show that, given no defectives in the sampled n DooDads, we should just estimate the number of defectives as binomially distributed among the remaining $N-n$ unsampled DooDads with defective probability p. In other words, since we accepted the value of p as fixed, and we believe that the DooDads are independent, then our posterior belief should still be that p is the appropriate probability of defectives among the unsampled DooDads, and we should expect $(N-n)p$ of them to be defective. If we had seen some other number, d, of defectives, the same reasoning suggests (and indeed it turns out to be correct) that the posterior Bayes estimate of D in that case is just $(N-n)p+d$.

Question 4

(a) The *MSE* of $\theta(\hat{F})$ is $MSE_F\{\theta(\hat{F})\} = E_F\left[\{\theta(\hat{F}) - \theta(F)\}^2\right]$. Thus, the bootstrap paradigm indicates that we should estimate this value by:

$$MSE_F\{\theta(\hat{F})\} \approx MSE_{\hat{F}}\{\theta(\hat{F}^\star)\} = E_{\hat{F}}\left[\{\theta(\hat{F}^\star) - \theta(\hat{F})\}^2\right].$$

To actually calculate an estimate, of course, we must approximate this value using re-samples from the data, so that our actual bootstrap estimator of *MSE* is:

$$\widehat{MSE}_B\{\theta(\hat{F})\} = \frac{1}{B}\sum_{b=1}^{B}\{\theta(\hat{F}_b^\star) - \theta(\hat{F})\}^2.$$

Of course, we could also just recall that the *MSE* is equal to the variance plus the square of the bias and use the usual bootstrap variance and bias estimates to make up the pieces.

This approach will lead to a slightly different answer if we use the usual bootstrap variance estimator; however, it will yields exactly the same answer if we replace the value $B-1$ in the usual bootstrap variance formula by the value B instead (i.e., we opt not to make the usual degrees-of-freedom-type correction).

(b) Using the hint, we see that:

$$\begin{aligned}
E_F(\tilde{\theta}_B) &= E_F\left\{2\theta(\hat{F}) - \frac{1}{B}\sum_{b=1}^{B}\theta(\hat{F}_b^\star)\right\} \\
&= 2E_F\{\theta(\hat{F})\} - \frac{1}{B}\sum_{b=1}^{B}E_F\{\theta(\hat{F}_b^\star)\} \\
&= 2\left\{\theta(F) + \frac{a(F)}{n}\right\} - E_F\left[E_{\hat{F}}\{\theta(\hat{F}^\star)\}\right] \\
&= 2\theta(F) + \frac{2a(F)}{n} - E_F\left\{\theta(\hat{F}) + \frac{a(\hat{F})}{n}\right\} \\
&= 2\theta(F) + \frac{2a(F)}{n} - \left[\left\{\theta(F) + \frac{a(F)}{n}\right\} + \frac{1}{n}\left\{a(F) + \frac{b(F)}{n}\right\}\right] \\
&= \theta(F) - \frac{b(F)}{n^2},
\end{aligned}$$

where the fourth equality follows from the fact that we have been told that $E_F\{\theta(\hat{F}) = \theta(F) + n^{-1}a(F)$ for *any* distribution F, including $F = \hat{F}$. Thus, the bias is $-\frac{b(F)}{n^2}$, which shows how the bootstrap bias corrected estimator achieves a generally smaller bias [depending, of course, on the relative sizes of $a(F)$, $b(F)$ and n; though as n increases, the relative sizes of $a(F)$ and $b(F)$ become irrelevant and we see that the bias is indeed dramatically reduced for large sample cases].

END OF EXAMINATION