Stat 543 Final Exam

May 3, 2005 **Prof. Vardeman**

WRITE ALL YOUR ANSWERS ON THIS EXAM!!!

1. Suppose that $X_1, X_2, \ldots, X_n$ are iid exponential with mean $\beta > 0$, i.e. have marginal pdf

$$f(x \mid \beta) = I[x > 0]\frac{1}{\beta}\exp\left(-\frac{x}{\beta}\right)$$

on $\Re$. Sadly, all one can observe are "right-censored" versions of the X_i,

$$Y_i = X_i \cdot I[X_i < T] + T \cdot I[X_i \geq T]$$

for some fixed (known) constant T, and inference about β must be based on $Y_1, Y_2, \ldots, Y_n$. (X_i is censored with β-probability $\exp\left(-\frac{T}{\beta}\right)$, and in such a case $Y_i = T$.) Define

$$n_T = \text{the number of } Y_i \text{ that are } T \quad \text{and} \quad n_{\text{u}} = n - n_T$$

The Y_i are neither discrete nor continuous, so the simple definitions of likelihood used in Stat 543 must be extended to cover this case. When this is done appropriately, a likelihood turns out to be

$$L_n(\beta) = \frac{1}{\beta^{n_{\text{u}}}}\exp\left(-\frac{\sum y_i}{\beta}\right)$$

a) A likelihood for $\gamma = 1/\beta$ is $L_n^*(\gamma) = \gamma^{n_{\text{u}}}\exp\left(-\gamma\sum y_i\right)$. A Bayesian uses an exponential prior distribution with mean 1 for γ. Set up completely but do not evaluate a ratio of definite integrals giving this person's (squared error loss) estimator for γ based on $Y_1, Y_2, \ldots, Y_n$.

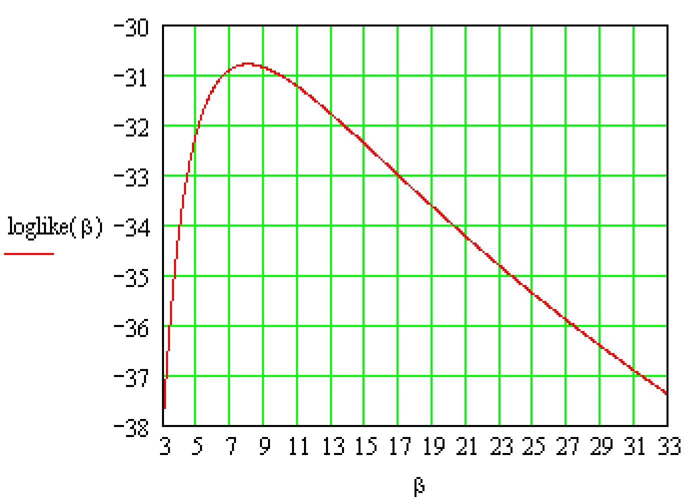
b) Find the maximum likelihood estimator of β.

c) Argue that your estimator from **b)** is consistent for $\beta \in (0,\infty)$. (Use Stat 542 facts about convergence in probability and/or distribution. You may use without proof the fact that $\mathrm{E}_\beta Y_1 = \beta\left(1-\exp\left(-T/\beta\right)\right)$.)

In a case where $n = 20$ and $T = 5$, it turns out that $n_{\mathrm{u}} = 10$ and $\sum y_i = 80$.

d) What are approximate 90% confidence limits for β in this case, based on your answer to part **b)** and the "observed Fisher information" in the sample?

The figure below is a plot of the log-likelihood $l_{20}(\beta) = \ln\left(L_{20}(\beta)\right)$ for the values of T, n_{u}, and $\sum y_i$ indicated above. Use it to answer **e)** and **f)** on the next page.



e) What are approximate 90% confidence limits for β (different from those in part **d)**) based on the figure? Explain.

f) If I wish to test $\mathrm{H}_0\!:\!\beta = 17$ versus $\mathrm{H}_0\!:\!\beta \neq 17$ using $\alpha \approx .05$, does the figure indicate acceptance or does it indicate rejection of the null hypothesis? Explain.

In parts **g)** and **h)** below, you may use the fact that (after making an appropriate definition for this case where Y_1 is neither discrete nor continuous) the Fisher information in any single censored observation is

$$I_1(\beta) = \frac{1}{\beta^2}\left(1 - \exp\left(-\frac{T}{\beta}\right)\right)$$

g) Estimation of β will be of interest in a future study. Investigators can choose n and T. Three (n,T) pairs are under consideration (as requiring about the same experimental effort/total resources). These are:

Plan A	**Plan B**	**Plan C**
$n = 80$ and $T = 4$	$n = 40$ and $T = 10$	$n = 20$ and $T = 22$

Investigators will choose between the plans on the basis of a "pre-study best guess" at the value of β. Which plan seems best if the "pre-study best guess"/"planning value" is $\beta \approx 18$? Justify your answer quantitatively. (Even a correct guess without justification gets no credit.)

h) Consider estimation of the variance of an uncensored observation, β^2 (still on the basis of n censored observations). What is the Cramér-Rao lower bound on the variance of an unbiased estimator of β^2?

2. For $\theta \in (0,\infty)$, suppose that $X_1, X_2, \ldots, X_n$ are iid $\mathrm{U}(0,\theta)$, that is, with marginal pdf on $\Re$ $f(x|\theta) = \theta^{-1} I[0 \le x \le \theta]$. Consider the MLE of θ,

$$\hat{\theta}_n = \max\{X_1, X_2, \ldots, X_n\}$$

a) Prove that in this "non-regular" problem $n(\theta - \hat{\theta}_n) \xrightarrow{\mathcal{L}_\theta} \mathrm{Exp}(\theta)$ (exponential with mean θ). (Hint: Evaluate $P_\theta\left[n(\theta - \hat{\theta}_n) > t\right]$ and remember that $\lim_{s\to\infty}\left(1+\frac{a}{s}\right)^s = \exp(a)$.)

b) Argue carefully that $n\hat{\theta}_n / (n - \ln(20))$ can be used as an approximately 95% upper confidence bound for θ.

3. Suppose that a discrete random variable X has pmf $f(x\,|\,\theta)$ specified below for $\theta \in \{1,2,3,4\}$.

θ \ x	1	2	3	4	5
4	5/15	4/15	3/15	2/15	1/15
3	4/30	5/30	6/30	7/30	8/30
2	1/15	2/15	3/15	4/15	5/15
1	.2	.2	.2	.2	.2

a) Carefully derive the size $\alpha = .4$ likelihood ratio test of $\mathrm{H}_0{:}\theta = 1$ or 2 versus $\mathrm{H}_\mathrm{a}{:}\theta = 3$ or 4. (Show your reasoning!)

b) A Bayesian (using 0-1 loss) wishes to test $\mathrm{H}_0{:}\theta = 1$ vs $\mathrm{H}_\mathrm{a}{:}\theta \neq 1$ and adopts a prior distribution with pmf specified by $g(1) = 1/2, g(2) = 1/6, g(3) = 1/6,$ and $g(4) = 1/6$. What test is optimal for this person? (Specify $\phi(x)$ for all 5 values of x.)

4. For $\eta \in \Re$, consider (X,Y) with (joint) pdf on $\Re^2$

$$f(x,y\,|\,\eta) = I[0 < x < 1, 0 < y < 1]\,C(\eta)\exp(\eta|x-y|)$$

(for some normalizing constant $C(\eta)$).

a) Suppose that Vardeman wants to evaluate the $\eta = 2$ mean of X using a Gibbs sampling algorithm. Carefully describe how to do this. (What conditionals does he sample from? You only need to specify them up to a multiplicative constant. Then what does he do with the generated sequence of $(X,Y)_i^*$'s?)

b) Identify a UMP size $\alpha = .25$ test of $\mathrm{H}_0\!:\eta \le 0$ vs $\mathrm{H}_a\!:\eta > 0$ based on a single observation (X,Y) and argue carefully that your test is indeed UMP of its size.

c) If $(X_1,Y_1),(X_2,Y_2),\ldots,(X_n,Y_n)$ are iid with marginal pdf $f(x,y\,|\,\eta)$, find a minimal sufficient statistic and argue carefully that it is indeed minimal sufficient.