Stat 543 Exam II Spring 1998

April 1, 1998 Prof. Vardeman

1. Consider the simple discrete distribution specified by the probability mass function given below in tabular form for $\theta \in \Theta = [0, .302]$.

x	0	1	2	3
$f(x\|\theta)$	θ	2θ	θ^2	$1 - 3\theta - \theta^2$

Consider first estimation of θ based on a single observation from this distribution, X.

a) Show explicitly that the mean of the score function here is 0 (for all θ).

b) Write out (but do not take the time to simplify) an explicit expression for the Fisher information in X about θ. (As it turns out, I believe that $I(\theta) = \frac{3+4\theta-6\theta^2-3\theta^3}{\theta(1-3\theta-\theta^2)}$.)

c) One possible estimator of θ is

$$\widehat{\theta}(X) = \begin{cases} .302 & \text{if } X = 0 \\ .232 & \text{if } X = 1 \\ .016 & \text{if } X = 2 \\ 0 & \text{if } X = 3 \end{cases} .$$

Find a better estimator of θ (in terms of MSE) and argue carefully that it is better.

d) Consider now a Bayesian approach to estimation of θ with a prior distribution uniform on the interval $[0, .2]$. Suppose that $X = 2$.

- i) What is the posterior probability that $\theta > .25$ (based on the outcome $X = 2$)?
- ii) Find the posterior mean of θ (based on the outcome $X = 2$).
- iii) Find a 90% highest posterior density credible region for θ (based on the outcome $X = 2$).
- iv) Carefully describe how you could use the rejection algorithm and a stream of iid Uniform $(0, 1)$ random variables to create a sequence $\{\theta_i^*\}$ of iid observations from the posterior distribution (based on the outcome $X = 2$).

Now consider the possibility of estimating θ based on n iid observations $X_1, X_2, ..., X_n$. Henceforth suppose $\delta_n(X)$ is the MLE of θ. (**Don't** try to actually find the form of the MLE of θ.)

e) Using again the prior distribution for θ described in d), what, *in qualitative terms*, would you expect to happen (as $n \to \infty$) to the posterior distributions of θ

- i) if in fact $\theta = .15$?
- ii) if in fact $\theta = .25$?

The graph attached to this exam is for the loglikelihood in this problem, based on a sample of $n = 100$ observations with the frequency distribution below

x	0	1	2	3
frequency	19	34	7	40

f) Use the graph and give a large sample 90% confidence interval for θ based on the large sample distribution of $\delta_n(X)$ and the *observed* (not expected) Fisher information in this sample.

g) Use the graph of the loglikelihood $\ell_n(\theta)$ and give a large sample 90% confidence interval for θ based on the large sample distribution of $\ell_n(\delta_n(X)) - \ell_n(\theta)$.

2. Consider the problem of estimation in the Weibull family of distributions based on n iid observations $X_1, X_2, ..., X_n$. For scale and location parameters α and β, the loglikelihood here is :

$$\ell_n(\alpha, \beta) = n\ln\beta - n\beta\ln\alpha + (\beta - 1)\left(\sum_{i=1}^{n}\ln x_i\right) - \frac{1}{\alpha^\beta}\left(\sum_{i=1}^{n} x_i^\beta\right)$$

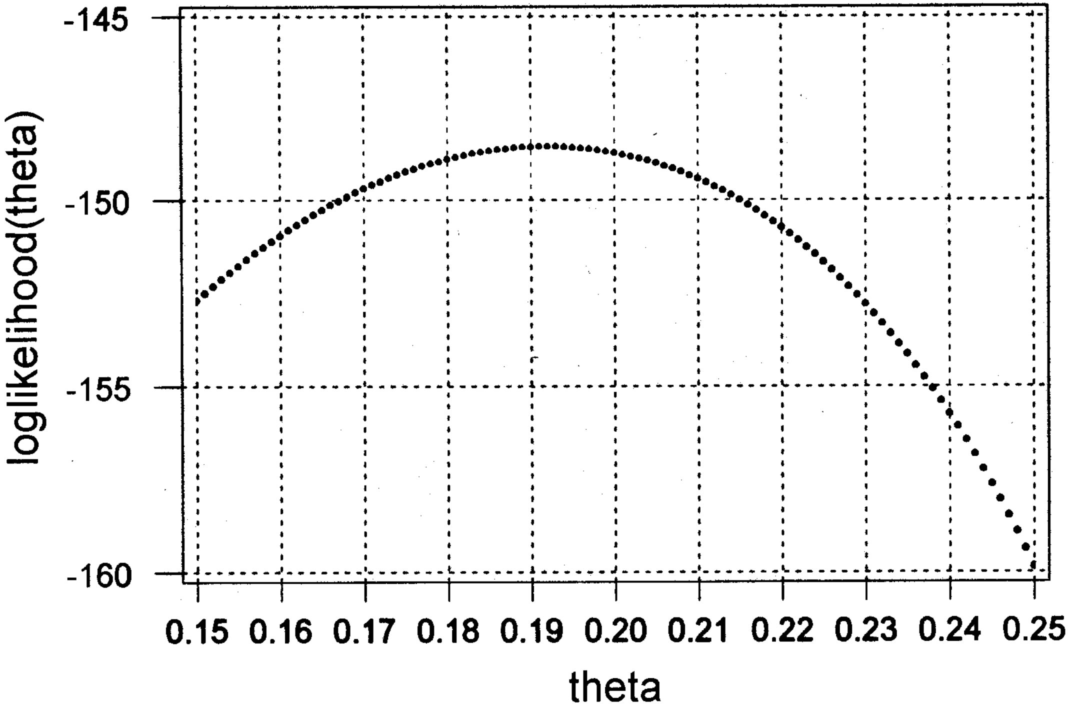
and the likelihood equations reduce to

$$\beta = \left(\frac{\sum_{i=1}^{n} x_i^\beta \ln x_i}{\sum_{i=1}^{n} x_i^\beta} - \frac{\sum_{i=1}^{n}\ln x_i}{n}\right)^{-1}$$

and

$$\alpha = \left(\frac{\sum_{i=1}^{n} x_i^\beta}{n}\right)^{\frac{1}{\beta}} .$$

There are no closed form solutions for these equations (i.e. there are no explicit formulas for MLEs here). However, suppose that I want to make a 90% approximate confidence interval for β (the shape parameter). Describe in as much detail as possible how you would go about producing such an interval.


-145
-150
-155
-160
loglikelihood(theta)
0.15
0.16
0.17
0.18
0.19
0.20
0.21
0.22
0.23
0.24
0.25
theta