Stat 513
Emerson, Winter 04
Midterm #2 Key
March 1, 2004

Instructions:

- **This exam is closed book, closed notes. No use of calculators is permitted.**
- **Write answers to the following questions on the front side of separate sheets of paper, starting each problem at the top of a new page. Be sure to write your name on the top of each page.**
- **In order to receive full credit, you must make clear how you derived the answers to the problems.**
- **If any of the problems seem unsolvable with the information provided, clearly state reasonable assumptions that would allow solution, and use those assumptions to solve the problem. If no such reasonable assumptions are obvious, simply state that and proceed to other problems.**
- **You are allowed 50 minutes for this exam. At the end of the in-class exam, you may complete or correct the solution to any problem as a take-home exam subject to the conditions specified on the last page of the exam.**
- **Write out and sign the following pledge for the in-class exam:**

  "On my honor, I have neither given nor received any unauthorized aid on this in-class examination."

  If for any reason you can not honestly sign the pledge, please discuss this with me on Monday.

1. (50 points) Let $X_1, X_2, \ldots, X_n$ be a sequence of i.i.d. random variables having Weibull distribution with cumulative distribution function

$$Pr(X_i \leq x) = (1 - \exp(-(\lambda x)^p))\mathbf{1}_{[x>0]}$$

and density

$$f(x; p, \lambda) = p\lambda(\lambda x)^{p-1} \exp(-(\lambda x)^p)\, \mathbf{1}_{[x>0]},$$

with $p > 0$ known and $\lambda > 0$ an unknown parameter.

a. For specified $\lambda_0 < \lambda_1$, find the most powerful level α test of $H_0 : \lambda = \lambda_0$ versus $H_1 : \lambda = \lambda_1$. (For this part, you need not find the critical value.)

Ans: By the Neyman-Pearson lemma, the MP-α test will be of the form

$$\psi(\vec{X}) = \begin{cases} 1 & f(\vec{X}; \lambda_1) > kf(\vec{X}; \lambda_0) \\ 0 & f(\vec{X}; \lambda_1) < kf(\vec{X}; \lambda_0) \end{cases}$$

Due to common support for $\vec{X}$ across all values of λ, we can consider the ratio of likelihoods

$$\begin{aligned}\frac{f(\vec{X};\lambda_1)}{f(\vec{X};\lambda_0)} &= \frac{p^n\lambda_1^{np}\left(\prod_{i=1}^n x_i\right)^p \exp\left(-\lambda_1^p\sum_{i=1}^n x_i^p\right)}{p^n\lambda_0^{np}\left(\prod_{i=1}^n x_i\right)^p \exp\left(-\lambda_0^p\sum_{i=1}^n x_i^p\right)} \\ &= \frac{\lambda_1^{np}}{\lambda_0^{np}}\exp\left(-(\lambda_1^p-\lambda_o^p)\sum_{i=1}^n x_i^p\right).\end{aligned}$$

Now

$$\frac{\lambda_1^{np}}{\lambda_0^{np}}\exp\left(-(\lambda_1^p-\lambda_0^p)\sum_{i=1}^n x_i^p\right) > k$$

$$\text{if and only if } (\lambda_1^p-\lambda_0^p)\sum_{i=1}^n x_i^p < k_1 = -\log\left(\frac{\lambda_0^{np}}{\lambda_1^{np}}k\right)$$

$$\text{if and only if } \sum_{i=1}^n x_i^p < k_2 = \frac{k_1}{(\lambda_1^p-\lambda_0^p)},$$

where the last step is valid because $\lambda_1 > \lambda_0 > 0$ and $p > 0$. Hence the MP-α test is

$$\psi(\vec{X}) = \begin{cases} 1 & \sum_{i=1}^n x_i^p \le k_2 \\ 0 & \sum_{i=1}^n x_i^p > k_2 \end{cases}$$

where k_2 is chosen to provide $E[\psi(\vec{X})\,|\,\lambda=\lambda_0]=\alpha$.

b. Find the uniformly most powerful level α test of $H_0:\lambda=\lambda_0$ versus $H_1:\lambda>\lambda_0$.

Ans: As the answer to part a depended only on λ_1 being greater than λ_0, that same test is MP-α for all $\lambda_1>\lambda_0$, and hence it is UMP-α for these hypotheses.

c. Derive an expression for the critical value for the test in part b.

Ans: We need to find k_2 such that $Pr[\sum_{i=1}^n X_i^p \le k_2\,|\,\lambda=\lambda_0]=\alpha$. I first find the distribution of $Y_i = X_i^p$ as

$$\begin{aligned}Pr[X_i^p \le y\,|\,\lambda] &= Pr[X_i \le y^{1/p}\,|\,\lambda] \\ &= \left(1-\exp\left(-(\lambda(y^{1/p})^p\right)\right)\mathbf{1}_{[y^{1/p}>0]} \\ &= \left(1-\exp\left(-\lambda^p y\right)\right)\mathbf{1}_{[y>0]},\end{aligned}$$

so X_i^p is exponentially distributed with hazard parameter λ^p. We thus immediately know that X_i^p has expectation $1/\lambda^p$ and variance $1/\lambda^{2p}$. We also know that the sum of these independent exponentials has the gamma distribution with parameters n and λ^p (in the parameterization of the gamma where the exponential is gamma with parameters 1 and λ^p). The critical value k_2 is thus the αth quantile of the gamma cumulative distribution function. That is, k_2 satisfies

$$\int_0^{k_2}\frac{u^{n-1}\lambda_0^{np}}{\Gamma(n)}\exp(-\lambda_0^p u)\,du = \alpha.$$

We could also use the normal approximation based on the asymptotic results. From the central limit theorem, we know that

$$\sum_{i=1}^{n} X_i^p \dot\sim \mathcal{N}\left(\frac{1}{\lambda^p}, \frac{1}{n\lambda^{2p}}\right)$$

so we can find an approximate value for k_2 by

$$\begin{aligned}
Pr\Big[\sum_{i=1}^{n} X_i^p \le k_2 \,|\, \lambda_0\Big] &= Pr\Big[\frac{\sum_{i=1}^{n} X_i^p - \frac{1}{\lambda_0}}{\frac{1}{\lambda_0\sqrt{n}}} \le \frac{k_2 - \frac{1}{\lambda_0}}{\frac{1}{\lambda_0\sqrt{n}}} \,|\, \lambda_0\Big] \\
&= Pr\Big[\sqrt{n}\left(\lambda_0 \sum_{i=1}^{n} X_i^p - 1\right) \le \sqrt{n}\,(\lambda_0 k_2 - 1) \mid \lambda_0\Big] \\
&\doteq Pr\Big[\mathcal{N}(0,1) \le \sqrt{n}\,(\lambda_0 k_2 - 1)\Big] \\
&= \Phi\left(\sqrt{n}\,(\lambda_0 k_2 - 1)\right) \\
&= \alpha.
\end{aligned}$$

Hence, $\sqrt{n}(\lambda_0 k_2 - 1) \doteq z_\alpha$ and $k_2 \doteq (z_\alpha/\sqrt{n} + 1)/\lambda_0$.

d. Derive an expression for the power function for the test in part b.

Ans: The power function is merely the cumulative distribution function for a gamma distributed random variable with parameters n and λ^p

$$Pwr(\lambda) = \int_0^{k_2} \frac{u^{n-1}\lambda^{np}}{\Gamma(n)} \exp(-\lambda^p u)\, du.$$

e. Show that the test in part b is also uniformly most powerful level α test of $H_0 : \lambda \le \lambda_0$ versus $H_1 : \lambda > \lambda_0$.

Ans: All we need to show here is that the power (probability of rejecting H_0) is less than α for $\lambda < \lambda_0$. That is, we need to show that for $\lambda_* < \lambda_0$

$$Pr\Big[\sum_{i=1}^{n} X_i^p \le k_2 \,|\, \lambda_*\Big] \le Pr\Big[\sum_{i=1}^{n} X_i^p \le k_2 \,|\, \lambda_0\Big].$$

There are several different ways we can do this. Here I consider rescaling the random variables. (This does not require that we know the form of the gamma cdf.) If $X_i^p \sim \mathcal{E}(\lambda_*^p)$, then $W_i = \lambda_*^p X_i^p / \lambda_0^p \sim \mathcal{E}(\lambda_0^p)$. To see this, note that

$$\begin{aligned}
Pr\Big[W_i \le w \,|\, \lambda_*\Big] &= Pr\Big[\frac{\lambda_*^p}{\lambda_0^p} X_i^p \le w \,|\, \lambda_*\Big] \\
&= Pr\Big[X_i^p \le \frac{\lambda_0^p}{\lambda_*^p} w \,|\, \lambda_*\Big] \\
&= 1 - \exp\left(-\lambda_*^p \frac{\lambda_0^p}{\lambda_*^p} w\right) = 1 - \exp\left(-\lambda_0^p w\right)
\end{aligned}$$

The sum of these rescaled exponential random variables will then be a rescaled gamma random variable, and

$$\begin{aligned} Pr\Big[\sum_{i=1}^{n} X_i^p \le k_2 \,|\, \lambda_*\Big] &= Pr\Big[\frac{\lambda_*^p}{\lambda_0^p}\sum_{i=1}^{n} X_i^p \le \frac{\lambda_*^p}{\lambda_0^p} k_2 \,|\, \lambda_*\Big] \\ &= Pr\Big[\sum_{i=1}^{n} X_i^p \le \frac{\lambda_*^p}{\lambda_0^p} k_2 \,|\, \lambda_0\Big] \\ &\le Pr\Big[\sum_{i=1}^{n} X_i^p \le k_2 \,|\, \lambda_0\Big] = \alpha \end{aligned}$$

because $\lambda_*^p \le \lambda_0^p$ and a cumulative distribution function is nondecreasing.

(An alternative approach uses the formula for the power function. We merely differentiate the power function with respect to λ and show that it is decreasing as λ increases.)

2. (45 points) Let $X_1, X_2, \ldots, X_n$ be a sequence of i.i.d. normal random variables having distribution $X_i \sim \mathcal{N}(\theta, \theta)$ with unknown $\theta \in \Theta = (0, \infty)$.

a. Find a complete sufficient statistic for θ. Justify your answer.

Ans: The joint density for $\vec{X}$ is

$$f_{\vec{X}}(\vec{x};\theta) = \left(\frac{1}{\sqrt{2\pi\theta}}\right)^n \exp\left(-\frac{\sum_{i=1}^n (x_i - \theta)^2}{2\theta}\right),$$

which can be written in the exponential family form as

$$\begin{aligned} f_{\vec{X}}(\vec{x};\theta) &= \exp\left(-\frac{n}{2}\log(2\pi) - \frac{n}{2}\log\theta - \frac{\sum_{i=1}^n x_i^2}{2\theta} + \sum_{i=1}^{n} x_i - \frac{n}{2}\theta\right) \\ &= \exp\Big(c_0(\theta) + T_0(\vec{X}) + c_1(\theta)T_1(\vec{X})\Big) \end{aligned}$$

with $c_0(\theta) = -(n/2)(\log(\theta) + \theta)$, $T_0(\vec{X}) = -(n/2)\log(2\pi) + n\overline{X}_n$, $c_1(\theta) = 1/(2\theta)$, and $T_1(\vec{X}) = \sum_{i=1}^n X_i^2$. Hence, we have a 1 parameter exponential family with complete sufficient statistic $T_1(\vec{X}) = \sum_{i=1}^n X_i^2$, because $\theta \in \mathcal{R}^1$, which contains a 1-dimensional open interval.

b. Find a uniform minimum variance unbiased estimator (UMVUE) for $g(\theta) = \theta(\theta + 1)$. Justify your answer.

Ans: I start by finding the expectation of the complete sufficient statistic. Now $E[X^2] = Var(X) + E^2[X]$, so for these data we have $E[X_i^2] = \theta + \theta^2$. We thus have $E[T_1(\vec{X})] = ng(\theta)$, so by the Lehmann-Scheffé theorem, we know that $\hat{\theta} = T_1(\vec{X})/n$ is an unbiased function of the complete sufficient statistic, and hence is the UMVUE.

c. Find the efficiency of the estimator in part b.

Ans: For regular problems, if the score function can be factored as

$$\mathcal{U}(\theta) = A(\theta)\left(T(\vec{X}) - g(\theta)\right),$$

then $T(\vec{X})$ is best regular unbiased estimator (BRUE) for $g(\theta)$. We thus consider the score function

$$L(\theta) = \exp\left(-\frac{n}{2}\log(2\pi) - \frac{n}{2}\log\theta - \frac{\sum_{i=1}^n x_i^2}{2\theta} + \sum_{i=1}^n x_i - \frac{n}{2}\theta\right)$$

$$\mathcal{L}(\theta) = \log(L(\theta)) = -\frac{n}{2}\log(2\pi) - \frac{n}{2}\log\theta - \frac{\sum_{i=1}^n x_i^2}{2\theta} + \sum_{i=1}^n x_i - \frac{n}{2}\theta$$

$$\mathcal{U}(\theta) = \frac{\partial}{\partial\theta}\mathcal{L}(\theta) = -\frac{n}{2\theta} - \frac{\sum_{i=1}^n x_i^2}{2\theta^2} - \frac{n}{2}$$

$$= \frac{n}{2\theta^2}\left(\frac{1}{n}\sum_{i=1}^n X_i^2 - \theta(\theta+1)\right).$$

By inspection, we see that our UMVUE $\hat{\theta}$ is also BRUE, so it is fully efficient (i.e., the efficiency is 1). (This was drastically easier than figuring out what the variance of $\hat{\theta}$ is.)

3. (45 points) Let $X_1, X_2, \ldots$ be a sequence of i.i.d. random variables having mean μ and variance μ^2, where $\mu \in (-\infty, \infty)$ is unknown. We are interested in inference about $\theta = \log(\mu)$.

a. Find a consistent estimator $\hat{\theta}$ of θ, and find its asymptotic distribution.

Ans: We know that sample moments are consistent for population moments, and we know that the Mann-Wald theorem provides for continuous functions of consistent statistics to be consistent for the corresponding continuous function of the parameters. Hence, by the weak law of large numbers, we know that $\hat{\mu} = \overline{X}_n$ is consistent for μ, so $\hat{\theta} = \log(\overline{X}_n)$ will be consistent for $\theta = \log(\mu)$ (providing, of course, that $\mu > 0$– it was a typo on the exam that I did not restrict $\mu > 0$.) Other estimators could have been based on the first and second sample moments, but as a general rule, it will be best to use the first moment when possible.)

b. Construct an approximate level α hypothesis test of $H_0 : \theta \leq \theta_0$ versus $H_1 : \theta > \theta_0$. Be sure to make clear the test statistic used, the critical value, and the power function for arbitrary θ.

Ans: From the central limit theorem, we have that

$$\sqrt{n}(\hat{\mu} - \mu) \rightarrow_d \mathcal{N}\left(0, \mu^2\right).$$

Using the delta method with $g(\mu) = \log(\mu)$ (so $g'(\mu) = 1/\mu$), we obtain

$$\sqrt{n}(\log(\hat{\mu}) - \log(\mu)) \rightarrow_d \frac{1}{\mu}\mathcal{N}(0, \mu^2) \sim \mathcal{N}(0, 1).$$

This then yields approximate asymptotic distribution

$$\hat{\theta} \dot{\sim} \mathcal{N}(\theta, 1/n).$$

An intuitive test would reject H_0 in favor of H_1 when $\hat{\theta}$ is large. So we define a rule

$$\psi(\vec{X}) = \mathbf{1}_{[\hat{\theta} \geq k]}$$

where critical value k satisfies

$$\sup\{E[\psi(\vec{X}) \,|\, \theta] \,:\, \theta \leq \theta_0\} \leq \alpha.$$

Now,

$$\begin{aligned} E[\psi(\vec{X}) \,|\, \theta] &= Pr[\hat{\theta} \geq k \,|\, \theta] \\ &= Pr[\sqrt{n}(\hat{\theta} - \theta) \geq \sqrt{n}(k - \theta) \,|\, \theta] \\ &\doteq 1 - \Phi\left(\sqrt{n}(k - \theta)\right), \end{aligned}$$

which is increasing in θ for fixed k and n (as θ gets large, $\sqrt{n}(k - \theta)$ gets small, and the nondecreasing nature of a cdf says that $\Phi\left(\sqrt{n}(k - \theta)\right)$ gets small, leading to higher power). This argues that we should choose k based on the largest value of θ in the null hypothesis. So we choose

$$\begin{aligned} \alpha &= E[\psi(\vec{X}) \,|\, \theta_0] \\ &\doteq 1 - \Phi\left(\sqrt{n}(k - \theta_0)\right), \end{aligned}$$

to obtain $\sqrt{n}(k - \theta_0) \doteq z_{1-\alpha}$, or $k \doteq \theta_0 + z_{1-\alpha}/\sqrt{n}$. The power function for arbitrary θ is

$$Pwr(\theta) = Pr[\hat{\theta} \geq k \,|\, \theta] \doteq 1 - \Phi\left(z_{1-\alpha} - \sqrt{n}(\theta - \theta_0)\right).$$

c. Presuming the approximate distribution of your test statistic is a good approximation, is the test you found in part b unbiased? Is it consistent?

Ans: Because the power function is increasing in θ and we have $\theta_1 > \theta_*$ for all $\theta_1 \in \Theta_1$ and all $\theta_* \in \Theta_0$, the test is (to a good approximation) unbiased.

For arbitrary $\theta_1 > \theta_0$, as $n \to \infty$, $z_{1-\alpha} - \sqrt{n}(\theta - \theta_0) \to -\infty$ and $Pwr(\theta_1) \to 1$. Hence the test is consistent.

Instructions for take-home option: At the end of the in-class exam, you may complete or correct the solution to any problem as a take-home exam subject to the following conditions:

- **By submitting a corrected problem, you can earn up to half the total number of points taken off from your solution to the problem as submitted during the in-class portion of the exam.**
- **There is no limit to the amount of time you can spend on the take-home portion of the exam, except that any corrections/additions must be handed in by the time that the clock in CMU B006 registers 10:30 am on Monday, March 1, 2004. There are no possible exceptions to this rule. (If you anticipate being hit by a meteor on the way to class, I would encourage your making other arrangements to get your exam to class on time.)**
- **You may work on the problems using the texts Dudewicz and Mishra, Casella and Berger, lecture notes that you took in class, or notes obtained from the class web pages. You may not use any other book or notes from any other class. You may not talk about the problems with anyone else prior to turning in your corrections/additions. This includes me and the TAs. We will not answer any questions either in person or via email about any matters regarding this examination. Furthermore, neither I nor the TAs will be holding office hours until after class on Wednesday.**
- **In order for the take-home portion to be accepted, it must be accompanied by the following (truthful) pledge written out and signed by you:**

  > **"On my honor, I have neither given nor received any unauthorized aid on this take-home examination."**

  In the event that you have inadvertently violated this pledge (e.g., if you overheard a discussion of the problems by someone who is not turning in corrections/additions), you should not sign the pledge and instead discuss the situation with me. Again, there are no exceptions to this policy.