STAT 801

Problems: Assignment 5

1. Suppose $X_1,\ldots,X_m$ are iid $N(\mu,\sigma^2)$ and $Y_1,\ldots,Y_n$ are iid $N(\chi,\tau^2)$. Assume the Xs are independent of the Ys.

   (a) Find complete and sufficient statistics.
   The log likelihood is

   $$-(n+m)\log(2\pi)/2-m\log(\sigma)-n\log(\tau)-\frac{\sum X_i^2}{2\sigma^2}-\frac{\sum Y_i^2}{2\tau^2}+\frac{m\bar{X}\mu}{\sigma^2}+\frac{n\bar{Y}\chi}{\tau^2}-\frac{m\mu^2}{2\sigma^2}-\frac{n\chi^2}{2\tau^2}$$

   It follows that

   $$(\bar{X},\sum X_i^2,\bar{Y},\sum Y_i^2)$$

   is complete and sufficient. There are many acceptable alternative one to one functions of this vector.

   (b) Find UMVUE's of $\mu-\chi$ and σ^2/τ^2.
   Since $E(\bar{X}-\bar{Y})=\mu-\chi$ the Lehmann-Shceffé theorem proves that the UMVUE of $\mu-\chi$ is $\bar{X}-\bar{Y}$.
   The random variable $[s_x^2/\sigma^2]/[s_y^2/tau^2]$ has an F distribution so that

   $$E(s_x^2/s_y^2)=c\sigma^2/\tau^2$$

   If we find c then the estimate $c^{-1}s_x^2/s_y^2$ will be the UMVUE by Lehmann-Scheffé. But

   $$\begin{aligned}
   E([s_x^2/\sigma^2]/[s_y^2/tau^2]) &= E(\chi^2_{m-1}/(m-1))E((n-1)/\chi^2_{n-1}) \\
   &= \int_0^\infty \frac{n-1}{u}\frac{1}{\Gamma((n-1)/2)}\left(\frac{u}{2}\right)^{(n-1)/2-1} e^{-u/2}d(u/2) \\
   &= \int_0^\infty \frac{n-1}{2y}\frac{1}{\Gamma((n-1)/2)}y^{(n-1)/2-1}e^{-y}dy \\
   &= \frac{n-1}{2\Gamma((n-1)/2)}\int_0^\infty y^{(n-3)/2-1}e^{-y}dy \\
   &= \frac{(n-1)\Gamma((n-3)/2)}{2\Gamma((n-1)/2)} \\
   &= \frac{n-1}{n-3}
   \end{aligned}$$

   So the UMVUE is

   $$\frac{n-3}{n-1}\frac{s_x^2}{s_y^2}$$

(c) Now suppose you know that $\sigma = \tau$. Find UMVUE's of $\chi - \mu$ and of $(\chi - \mu)/\sigma$. (You have already found the UMVUE for σ^2.)

You should check that for the smaller model the vector of complete sufficient statistics is the 3 vector

$$(\bar{X}, \bar{Y}, \sum (X_i - \bar{X})^2 + \sum (Y_i - \bar{Y})^2)$$

According to Lehmann-Scheffé $\bar{Y} - \bar{X}$ is still the UMVUE of $\chi - \mu$. Since $(\bar{X}, \bar{Y})$ is independent of

$$s^2 = [\sum (X_i - \bar{X})^2 + \sum (Y_i - \bar{Y})^2]/(n + m - 2)$$

we find

$$E[(\bar{Y} - \bar{X})/s] = \frac{\chi - \mu}{\sigma} E(\sigma/s)$$

Now $(n+m-2)s^2/\sigma^2$ has a χ^2_{n+m-2} distribution so

$$\begin{aligned} E(\sigma/s) &= \sqrt{n+m-2} \int_0^\infty u^{-1/2} \frac{1}{\Gamma((n+m-2)/2)} \left(\frac{u}{2}\right)^{(n+m-2)/2-1} e^{-u/2} d(u/2) \\ &= \sqrt{n+m-2} \int_0^\infty \frac{1}{2\Gamma((n+m-2)/2)} y^{(n+m-2)/2-1-1/2} e^{-y} dy \\ &= \frac{(n+m-2)\Gamma((n+m-3)/2)}{2\Gamma((n+m-2)/2)} \end{aligned}$$

It follows that the UMVUE of $(\chi - \mu)/\sigma$ is

$$\frac{2\Gamma((n+m-2)/2)}{(n+m-2)\Gamma((n+m-3)/2)} \frac{\bar{Y} - \bar{X}}{s}$$

WARNING: I haven't checked the algebra very carefully.

(d) Now suppose σ and τ are unknown but that you know that $\mu = \chi$. Prove there is no UMVUE for μ. (Hint: Find the UMVUE if you knew $\sigma/\tau = a$ with a known. Use the fact that the solution depends on a to finish the proof.)

We are analyzing the model $X_1, \ldots, X_m$ iid $N(\mu, \sigma^2)$ and independently $Y_1, \ldots, Y_n$ iid $N(\mu, \tau^2)$. I suggested studying the submodel $X_1, \ldots, X_m$ iid $N(\mu, \sigma^2)$ and independently $Y_1, \ldots, Y_n$ iid $N(\mu, a^2\sigma^2)$. For this model the complete sufficient statistics are

$$\sum X_i^2 + \sum Y_i^2/a^2$$

and

$$\sum X_i + \sum Y_i/a^2$$

The UMVUE of μ is then

$$\hat{\mu}_a \equiv (\sum X_i + \sum Y_i/a^2)/(m + n/a^2)$$

Now if there were a UMVUE $\hat{\mu}$ for the big model we would have to have

$$\text{Var}(\hat{\mu}) \leq \text{Var}(\hat{\mu}_a)$$

for all σ^2, τ^2 and μ. In particular, this would have to be true for each τ and σ for which $\tau = a\sigma$. But then $\hat{\mu}$ would be UMVUE in the small model. Since UMVUEs are unique in the small model (Lehmann-Scheffé and Rao-Blackwell and see the theorem in Casella and Berger) we find that

$$\hat{\mu} = \hat{\mu}_a$$

Since this can't be true for two different values of a there must not exist a UMVUE in the big model.

(e) Why doesn't the Lehmann-Scheffé theorem apply?

For any choice of a $\hat{\mu}_a$ is unbiased. Thus $\hat{\mu}_1 - \hat{\mu}_2$ has expected value identically equal to 0. Since $\hat{\mu}_1 - \hat{\mu}_2$ is a function of the minimal sufficient statistics, we see that these statistics are not complete, making Lehmann-Shceffé irrelevant.

2. Suppose $X_1, \ldots, X_n$ iid Poisson(λ). Find the UMVUE for λ and for $1 - \exp(-\lambda) = P(X_1 \neq 0)$.

$\bar{X}$ is complete and sufficient and unbiased for λ. Hence the UMVUE of λ is $\bar{X}$. IF

$$T = 1(X_1 \neq 0)$$

we see that T is unbiased for $\phi = 1 - \exp(-\lambda)$ then the UMVUE of ϕ is

$$E(T|\bar{X})$$

To compute this we note

$$\begin{aligned} E(T|\sum X_i = x) &= P(X_1 \neq 0|\sum X_i = x) \\ &= 1 - P(X_1 = 0|\sum X_i = x) \\ &= 1 - \frac{P(X_1 = 0, X_1 + \cdots + X - n = x)}{P(\sum X_i = x)} \end{aligned}$$

$$= 1 - \frac{P(X_1 = 0)P(X_2 + \cdots + X_n = x)}{P(\sum X_i = x)}$$

Remember that $X_1 + \cdots + X_n$ is Poisson$(n\lambda)$ and that $X_2 + \cdots + X_n$ is Poisson$((n-1)\lambda)$ to get

$$\begin{aligned} E(T|\sum X_i = x) &= 1 - \frac{e^{-\lambda}\exp(-(n-1)\lambda)((n-1)\lambda)^x/x!}{exp(-n\lambda)(n\lambda)^x/x!} \\ &= 1 - \left(\frac{n-1}{n}\right)^x \end{aligned}$$

Conditioning on $\sum X_i = x$ is the same as conditioning on $\bar{X} = x/n$ we see that the UMVUE of ϕ is

$$1 - (1 - 1/n)^{n\bar{X}}$$

3. Suppose $X_1, \ldots, X_n$ iid with

$$P(X_1 = k) = \text{Prob}(\text{Poisson}(\lambda) = k | \text{Poisson}(\lambda) > 0)$$

for $k = 1, 2, 3, \ldots$. For $n = 1$ and 2 find the UMVUE of $1 - \exp(-\lambda)$. (Hint: The expected value of any function of X is a power series in λ divided by $e^{\lambda} - 1$. Set this equal to $1 - \exp(-\lambda)$ and deduce that two power series are equal. Since this implies their coefficients are the same you can see what the estimate must be.)

For $n = 1$ if the UMVUE is $g(X)$ then

$$1 - e^{-\lambda} = \sum_{k=1}^{\infty} g(k) \frac{e^{-\lambda}}{1 - e^{-\lambda}} \lambda^k / k!$$

Rearrange to get

$$\sum g(k) \lambda^k / k! = e^{\lambda} + e^{-\lambda} - 2$$

The power series expansion of the right hand side is

$$\sum_{k \text{ even}} 2\lambda^k / k!$$

so that $g(k) = 0$ for odd k and $g(x) = 2$ for k even.

For $n = 2$ the complete sufficient statistic is $X = X - 1 + X_2$. If $g(X)$ is the UMVUE of $\phi = 1 - e^{-\lambda}$ then

$$1 - e^{-\lambda} = \sum_{k=2}^{\infty} g(k) P(X = k)$$

Now

$$\begin{aligned} P(X = k) &= \sum_{j=1}^{k-1} P(X_1 = j) P(X_2 = k - j) \\ &= \frac{e^{-2\lambda}}{(1 - e^{-\lambda})^2} \lambda^k \sum_{j=1}^{k-1} \frac{1}{j!(k-j)!} \end{aligned}$$

Temporarily let $c_k = \sum_{j=1}^{k-1} 1/(j!(k-j)!)$ and find

$$e^{2\lambda} - 3e^{\lambda} + 3 - e^{-\lambda} = \sum c_k g(k) \lambda^k / k!$$

From the power series expansion of the LHS we conclude

$$g(k) = (2^k - 3 - (-1)^k)/c_k$$

Some simplification of c_k is possible using the binomial expansion of $(1 + 1)^k$.

4. Suppose $\{X_{ij}; j = 1, \ldots, n_i; i = 1, \ldots, p\}$ are independent $N(\mu_i, \sigma^2)$ random variables. (This is the usual set-up for the one-way layout.)

(a) Find the MLE's for μ_i and σ .

$$\hat{\mu}_i = \sum_j X_{ij}/n_i$$

and

$$\hat{\sigma} = \sqrt{\frac{\sum(X_{ij} - \hat{\mu}_i)^2}{\sum_i n_i}}$$

(b) Find the expectations and variances of these estimators.

$$E(\hat{\mu}_1) = \mu_i$$

and

$$E(\hat{\sigma}^2) = \frac{n-p}{n}\sigma^2$$

Moreover

$$\text{Var}(\hat{\mu}_i) = \sigma^2/n_i$$

To find the mean and variance of $\hat{\sigma}$ (as opposed to those of $\hat{\sigma}^2$ which is what I meant to ask about) use the tactics in question 1 of this assignment remebering that $n\hat{\sigma}^2/\sigma^2$ has a χ^2_{n-p} distribution. Note that

$$\text{Var}(\hat{\sigma}^2) = \frac{2(n-p)}{n^2}\sigma^2$$

5. Let T_i be the error sum of squares in the ith cell in the previous question.

(a) Find the joint density of the T_i.
Use the fact that $Y_i \equiv T_i/\sigma^2$ has a $\chi^2_{n_i-1}$ distribution and the change of variables formula along with the independence of the T_i. Get

$$\sigma^{-2(n-p)} \prod \left(\frac{t_i}{2}\right)^{(n_i-1)/2-1} \frac{1}{2\Gamma((n_i-1)/2)} \exp(-\sum t_i/(2\sigma^2))$$

(b) Find the best estimate of σ^2 of the form $\sum_1^p a_i T_i$ in the sense of mean squared error.
You have to minimize

$$\sum a_i^2 \text{Var}(T_i) + \left(\sum a_i E(T_i) - \sigma^2\right)^2$$

This quantity is just σ^4 times

$$2\sum a_i^2(n_i-1) + (\sum a_i(n_i-1) - 1)^2$$

Take the derivative with respect to a_k to get

$$4a_k(n_k-1) + 2(n_k-1)(\sum a_i(n_i-1) - 1) = 0\,.$$

It follows that

$$a_k = c$$

for some constant c not depending on k. Replacing each a_i by c gives

$$4c + 2(c\sum(n_i - 1) - 1) = 0$$

or $c = 1/(\sum(n_i - 1) + 2) = 1/(n - p + 2)$. Thus the desired estimate is

$$\sum T_i/(n - p + 2) = ESS/(n - p + 2)$$

(c) Do the same under the condition that the estimator must be unbiased.
We now have to minimize

$$2\sum a_i^2(n_i - 1)$$

subject to $\sum a_i(n_i - 1) = 1$. Use Lagrange multipliers and take the derivative of

$$2\sum a_i^2(n_i - 1) + \lambda(\sum a_i(n_i - 1) - 1)$$

with respect ot a_k to get

$$4a_k(n_k - 1) + \lambda(n_k - 1) = 0$$

to learn again that $a_k = c$ for some c. The constraint then requires that $c\sum(n_i - 1) = c(n - p) = 1$. The desired estimate is the usual mean squared error

$$\sum T_i/(n - p)$$

(d) If only $T_1, \ldots T_p$ are observed what is the MLE of σ?
You have to minimize the log of the likelihood in the first part of this question, that is, the log of the joint density of $T_1, \ldots, T_p$. You get

$$\hat{\sigma}^2 = \sum T_i/(n - p)$$

(e) Find the UMVUE of σ^2 for the usual one-way layout model, that is, the model of the last two questions.
The solution is $\hat{\sigma}^2 = \sum T_i/(n - p) = ESS/(n - p) = MSE$. You can use the Lehmann-Scheffé theorem to prove it is UMVUE.

6. Exponential families: Suppose $X_1, \ldots, X_n$ are iid with density

$$f(x; \theta_1, \ldots, \theta_p) = c(\theta)\exp(\sum_1^p T_i(x)\theta_i)h(x).$$

(a) Find minimal sufficient statistics.

$(S_1, \ldots, S_p$ are minimal sufficient where

$$S_i = \sum_j T_i(X_j)$$

(b) If $S_1, \ldots, S_p$ are the minimal sufficient statistics show that setting $S_i = \mathrm{E}_\theta(S_i)$ and solving gives the likelihood equations. (Note the connection to the method of moments.)

The trick is to use the fact that the score function in this problem has kth component

$$S_k - \frac{\partial}{\partial \theta_k} \log(c(\theta))$$

and that the score has mean 0 so that

$$E(S_k) = \frac{\partial}{\partial \theta_k} \log(c(\theta))$$

7. In question 4 take $n_i = 2$ for all i and let $p \to \infty$. What happens to the MLE of σ?

We have

$$E(\hat\sigma^2) = \sigma^2/2$$

and

$$\mathrm{Var}(\hat\sigma^2) = \sigma^4/(2(n-p)) \to 0$$

Thus

$$\hat\sigma^2 \to \sigma^2/2$$

and

$$\hat\sigma \to \sigma/\sqrt{2}$$

8. Suppose that $Y_1, \ldots, Y_n$ are independent random variables and that $x_1, \ldots, x_n$ are the corresponding values of some covariate. Suppose that the density of Y_i is

$$f(y_i) = \exp\left(-y_i \exp(-\alpha - \beta x_i) - \alpha - \beta x_i\right) 1(y_i > 0)$$

where α, and β are unknown parameters.

(a) Find the log-likelihood, the score function and the Fisher information.

$$\ell(\alpha, \beta) = -\sum [Y_i \exp(-\alpha - \beta x_i) - \alpha - \beta x_i]$$

The score function has components

$$U_\alpha(\alpha, \beta) = \sum [Y_i \exp(-\alpha - \beta x_i) - 1]$$

and

$$U_\beta(\alpha, \beta) = \sum [x_i Y_i \exp(-\alpha - \beta x_i) - x_i]$$

The Fisher information matrix is

$$\mathcal{I} = \left[\begin{array}{cc} n & \sum x_i \\ \sum x_i & \sum x_i^2 \end{array}\right]$$

(b) For the data set in ~*/teaching/courses/801/data1* fit the model and produce a contour plot of the log-likelihood surface, the profile likelihood for β and an approximate 95% confidence interval for β.

9. Consider the random effects one way layout. You have data $X_{ij}; i = 1, \ldots, p; j = 1, \ldots, n$ and a model $X_{ij} = \mu + \alpha_i + \epsilon_{ij}$ where the α's are iid $N(0, \tau^2)$ and the ϵ's are iid $N(0, \sigma^2)$. The αs are independent of the ϵs.

(a) Compute the mean and variance covariance matrix of the vector you get by writing out all the X_{ij} as a vector.

This vector is MVN and the mean of X_{ij} is μ. The variance of any X_{ij} is $\sigma^2 + \tau^2$. We have

$$\text{Cov}(X_{ij}, X_{ij'}) = \tau^2$$

for $j \neq j'$ and

$$\text{Cov}(X_{ij}, X_{i'j'}) = 0$$

for $i \neq i'$. This makes the variance covariance matrix of the vector mentioned here be a block diagonal matrix with blocks of the form

$$\sigma^2 I + \tau^2 11^t$$

as in the next part of the question.

(b) Suppose that M is a matrix of the form $aI + b11^t$ where I is a $p \times p$ identity and 1 denotes a column vector of p ones. Show that M^{-1} is of the form $cI + d11^t$ and find c and d. In what follows you may use the fact that the determinant of M is $a^{p-1}(a + pb)$.

Simply multiply

$$(cI + d11^t)(aI + b11^t) = acI + (ad + bc + pbd)11^t = I$$

provided $c = 1/a$ and then $(a + pb)d + b/a = 0$ or $d = -b/(a(a + pb))$. This makes

$$(\sigma^2 I + \tau^2 11^t)^{-1} = \sigma^{-2} I - \frac{\tau^2}{\sigma^2(\sigma^2 + p\tau^2)} 11^t$$

To prove the determinant fact (which I didn't ask you to do) you can use the fact that $b11^t$ has $p - 1$ eigenvalues equal to 0 and 1 equal to pb. Then if N is any symmetric matrix with eigenvalues $\lambda_1, \ldots, \lambda_p$ and corresponding eigenvectors $x_1, \ldots, x_p$ you see that

$$(N + aI)x_i = (\lambda_i + a)x_i$$

so that the eigenvalues of $N + aI$ are $\lambda_1 + a, \ldots, \lambda_p + a$. Now use the fact that for a symmetric matrix N we have $det(N) = \prod \lambda_i$ to finish the proof.

(c) Write down the likelihood.

If X is the vector of all the X_{ij} strung out in the order $X_{11}, X_{12}, \cdots$ then X has a $MVN(\mu 1, \Sigma)$ distribution where $\Sigma = \text{diag}(M, M, \ldots, M)$ and $M = \sigma^2 I + \tau^2 11^t$. The log likelihood is

$$\ell(\mu, \sigma, \tau) = -\frac{1}{2}(X - \mu 1)^t \Sigma^{-1}(X - \mu 1) - \log(\det(\Sigma))/2$$

Define X_i to be the vector of data in cell i. Then the log likelihood simplifies to

$$-\frac{1}{2}\sum_i (X_i - \mu 1)^t \left(\frac{1}{\sigma^2} I - \frac{\tau^2}{\sigma^2(\sigma^2 + p\tau^2)} 11^t\right)(X_i - \mu) - \log(\sigma^2 + p\tau^2)/2 - p(n-1)\log(\sigma)$$

(d) Find minimal sufficient statistics.

(e) Are they complete?

(f) Data sets like this are usually analyzed based on the fixed effects ANOVA table. Use the formulas for expected mean squares in this table to develop "method of moments" estimates of the three parameters. (Because the data are not iid this is not going to be exactly the same technique as the examples in class.)

Here I wanted you to say that the ANOVA table has entries Between SS, Within SS (or Error SS), and Total SS. The expected mean square for error is

$$E(MSE) = \frac{1}{p(n-1)} E\left(\sum_{ij}(X_{ij} - \bar{X}_{i\cdot})^2\right)$$

which is just

$$\frac{1}{p(n-1)} E\left(\sum_{ij}(\epsilon_{ij} - \bar{\epsilon}_{i\cdot})^2\right) = \sigma^2$$

The expected mean square between is

$$E(MSB) = \frac{1}{p-1} E\left(\sum_{ij}(\bar{X}_{i\cdot} - \bar{X}_{\cdot\cdot})^2\right) = (n-1)\tau^2 + \sigma^2$$

The correpsonding "method of moments estimates" are then

$$\hat{\sigma}^2 = MSE$$

and

$$\hat{\tau}^2 = (MSB - MSE)/(n-1)$$

Notice that this estimate could be negative.

(g) Can you find the MLE's?

If you differentiate the log likelihood with resepct to μ you get

$$1^t \Sigma^{-1}(X - \mu 1) = 0$$

which gives

$$\hat{\mu} = 1^t\Sigma^{-1}X/1^t\Sigma^{-1}1$$

You can check that

$$\Sigma 1 = (\sigma^2 + p\tau^2)1$$

Multiply by Σ^{-1} to find that

$$\Sigma^{-1}1 = \frac{1}{\sigma^2 + p\tau^2}1$$

and then check that

$$\hat{\mu} = \bar{X}_{\cdot\cdot}$$

Since

$$\ell(\hat{\mu}, \sigma, \tau) = -\frac{1}{2}\sum_i (X_i - \hat{\mu}1)^t A$$

10. For each of the doses $d_1, \ldots, d_p$ a number of animals $n_1, \ldots, n_p$ are treated with the corresponding dose of some drug. The number dying at dose d is Binomial with parameter $h(d)$. A common model for $h(d)$ is $\log(h/(1-h)) = \alpha + \beta d$.

    (a) Find the likelihood equations for estimating α and β.

    $$\ell(\alpha, \beta) = \sum X_i(\alpha + \beta d_i) + \sum n_i \log(1 - h(d_i))$$

    so that the likelihood equations become

    $$\sum X_i = \sum \frac{n_i e^{\alpha+\beta d_i}}{1 + e^{\alpha+\beta d_i}}$$

    and

    $$\sum d_i X_i = \sum \frac{n_i d_i e^{\alpha+\beta d_i}}{1 + e^{\alpha+\beta d_i}}$$

    (b) Find the Fisher information matrix.

    $$\mathcal{I} = \begin{bmatrix} a & b \\ b & c \end{bmatrix}$$

    (c) Define the parameter LD50 as the value of d for which $h(d) = 1/2$; express LD50 as a function of α and β.
    Since $h(LD50)/(1 - h(LD50)) = 1$ we find $\alpha + \beta LD50 = 0$. Hence

    $$LD50 = -\alpha/\beta$$

    (d) Use a Taylor expansion to find large sample confidence limits for LD50.
    This is a delta method question. If $f(\alpha, \beta) = -\alpha/\beta$ then the gradient of f is $(-1/\beta, \alpha/\beta^2)^t$ and

(e) At each of the doses -3.204, -2.903, 2.602, -2.301 and -2.000 a sample of 40 mice were exposed to antipneumonococcus serum. The numbers surviving were 7, 18, 32, 35, and 38 respectively. Get numerical values for the theory above. You can use *glm* or get preliminary estimates based on linear regression of the MLE of $h(d_i)$ against dose.

11. Suppose $X_1, \ldots, X_n$ are a sample of size n from the density

$$f_{\alpha,\beta}(x) = \frac{1}{\beta\Gamma(\alpha)}\left(\frac{x}{\beta}\right)^{\alpha-1} \exp(-x/\beta)\,1(x>0).$$

In the following question the digamma function ψ is defined by $\psi(\alpha) = \frac{d}{d\alpha}\log(\Gamma(\alpha))$ and the trigamma function ψ' is the derivative of the digamma function. From the identity $\Gamma(\alpha+1) = \alpha\Gamma(\alpha)$ you can deduce recurrence relations for the digamma and trigamma functions.

(a) For $\alpha = \alpha_o$ known find the mle for β.

The log likelihood for the full model is

$$\ell(\alpha,\beta) = -n\log(\Gamma(\alpha)) - n\alpha\log(\beta) - \sum X_i/\beta$$

Replacing α by α_0 and differentiating wrt β gives the likelihood equation

$$n\alpha_0/\beta = \sum X_i/\beta^2$$

so that

$$\hat\beta = \bar X/\alpha_0$$

(b) When both α and β are unknown what equation must be solved to find $\hat\alpha$, the mle of α?

The α derivative of the log likelihood is

$$-n\psi(\alpha) - n\log(\beta)$$

Replace β with, $\bar X/\alpha$ as in a) to get

$$n\psi(\alpha) = n[\log(\alpha) - \log(\bar X)]$$

(c) Evaluate the Fisher information matrix.

$$\mathcal{I} = \begin{bmatrix} n\psi'(\alpha) & n/\beta \\ n/\beta & n\alpha/\beta^2 \end{bmatrix}$$

(d) A sample of size 20 is in the file

~/teaching/801/gamma.

Use this data in the following questions. First take $\alpha = 1$ and find the mle of β subject to this restriction.

(e) Now use $\mathrm{E}(X) = \alpha\beta$ and $\mathrm{Var}(X) = \alpha\beta^2$ to get method of moments estimates $\tilde{\alpha}$ and $\tilde{\beta}$ for the parameters. (This was done in class so I just mean get numbers.)

(f) Do two steps of Newton Raphson to get MLEs.

(g) Use Fisher's scoring idea, which is to replace the second derivative in Newton Raphson with the Fisher information (and then not change it as you run the iteration), to redo the previous question.

(h) Compute standard errors for the MLEs and compare the difference between the estimates in the previous 2 questions to the SEs.

(i) Do a likelihood ratio test of $H_o : \alpha = 1$.