Solutions

STAT 613 Midterm (1 hour)

April 11th, 2011

NAME:

Total number of Marks: /25

Answer all the questions (questions are on both sides of the paper).

Marks will be given for clarity of the solution.

Write your solutions in the question paper.

Good Luck!

(1) Suppose that Z is a Weibull random variable with density $f(x;\phi,\alpha) = (\frac{\alpha}{\phi})(\frac{x}{\phi})^{\alpha-1}\exp(-(x/\phi)^{\alpha})$. Show that

$$\mathbb{E}(Z^r) = \phi^r\Gamma\left(1+\frac{r}{\alpha}\right).$$

[3]

Hint: Use

$$\int x^a \exp(-x^b)dx = \frac{1}{b}\Gamma(\frac{a}{b}+\frac{1}{b}) \qquad a,b>0.$$

This result may be useful in the questions below.

Change variables, let $y = x^b \Rightarrow x = y^{1/b}$ and

$$\frac{dx}{dy} = \frac{1}{b} y^{\frac{1}{b}-1}$$

Thus we have

$$\int x^a e^{-x^b} dx = \int y^{a/b+1/b-1} e^{-y} dy$$

(By given identity) $= \frac{1}{b}\Gamma\left(\frac{a}{b}+\frac{1}{b}\right)$.

Using the above we have that

$$\mathbb{E}(\cancel{Z}) = \phi\Gamma\left(1+\frac{1}{\alpha}\right)$$

$$\mathbb{E}(Z^2) = \phi^2\Gamma\left(1+\frac{2}{\alpha}\right)$$

Thus

$$\text{Var}(Z) = \phi^2\Gamma\left(1+\frac{2}{\alpha}\right) - \phi^2\Gamma\left(1+\frac{1}{\alpha}\right)^2$$

(2) Let us suppose that the random variable X is a mixture of Weibull distributions

$$f(x;\theta) = p(\frac{\alpha_1}{\phi_1})(\frac{x}{\phi_1})^{\alpha_1-1}\exp(-(x/\phi_1)^{\alpha_1}) + (1-p)(\frac{\alpha_2}{\phi_2})(\frac{x}{\phi_2})^{\alpha_2-1}\exp(-(x/\phi_2)^{\alpha_2}).$$

(i) Derive the mean and variance of X. [3]

define the r.v where $\delta = \begin{cases} 1 \\ 0 \end{cases}$ and $P(\delta=1)=p$, $P(\delta=0)=1-p$

$$\mathbb{E}(X) = \mathbb{E}\{\mathbb{E}(X|\delta)\} = p\phi_1\Gamma(1+\frac{1}{\alpha_1}) + (1-p)\phi_2\Gamma(1+\frac{1}{\alpha_2}).$$

To obtain variance use either. $\text{var}(X) = \mathbb{E}\{\mathbb{E}(X^2|\delta)\} - [\mathbb{E}(\mathbb{E}(X|\delta))]^2$ or

that $\text{var}(X) = \mathbb{E}\{\text{var}(X|\delta)\} + \text{var}[\mathbb{E}(X|\delta)]$, to obtain

$$\text{var}(X) = p\phi_1^2\Gamma(1+\frac{2}{\alpha_1}) + (1-p)\phi_2^2\Gamma(1+\frac{2}{\alpha_2}) - \{p\phi_1\Gamma(1+\frac{1}{\alpha_1}) + (1-p)\phi_2\Gamma(1+\frac{1}{\alpha_2})\}^2.$$

(ii) Obtain the exponential distribution which best fits the above mixture Weibull according to the Kullbach-Lieber criterion (recall that the exponential is $g(x;\lambda) = \frac{1}{\lambda}\exp(-x/\lambda)$). [3]

The Kullbach-Leibler criterion is defined as $\mathbb{E}_f\left\{\log\frac{g(X;\lambda)}{f(X;\theta)}\right\}$.

The best fitting exponential distribution is the λ which maxim

$$\hat{\lambda} = \arg\max_{\lambda} \mathbb{E}_f\left\{\log\frac{g(X;\lambda)}{f(X;\theta)}\right\} = \arg\max_{\lambda}\mathbb{E}_f\{\log g(X;\lambda)\}$$

= expectation of the misspecified log-likelihood.

$$\mathbb{E}_f\{\log g(X;\lambda)\} = \mathbb{E}_f\left\{-\frac{X}{\lambda} - \log\lambda\right\} = -\frac{\mathbb{E}_f(X)}{\lambda} - \log\lambda.$$

By diff the above wrt λ we see that it is maximised

when $\lambda = \mathbb{E}_f(X) = p\phi_1\Gamma(1+\frac{1}{\alpha_1}) + (1-p)\phi_2\Gamma(1+\frac{1}{\alpha_2})$.

Thus the best fitting exponential has mean.

(3) Let us suppose that $\{T_i\}_i$ are the survival times of lightbulbs. We will assume that $\{T_i\}$ are iid random variables with the density $f(\cdot;\theta_0)$ and survival function $\mathcal{F}(\cdot;\theta_0)$, where θ_0 is unknown. The survival times are censored, and $Y_i = \min(T_i, c)$ and δ_i are observed ($c > 0$), where $\delta_i = 1$ if $Y_i = T_i$ and is zero otherwise.

(a) (i) State the log-likelihood of $\{(Y_i, \delta_i)\}_i$. [1]

$$\mathcal{L}_T(\theta) = \sum_{i=1}^{n} \delta_i \log f(\underset{T_i}{Y_i};\theta) + \sum_{i=1}^{n}(1-\delta_i)\log \mathcal{F}(c;\theta_i).$$

(ii) We denote the above log-likelihood as $\mathcal{L}_T(\theta)$. Show that

$$-\mathbb{E}\left(\frac{\partial^2 \mathcal{L}_T(\theta)}{\partial\theta^2}\rfloor_{\theta=\theta_0}\right) = \mathbb{E}\left(\frac{\partial \mathcal{L}_T(\theta)}{\partial\theta}\rfloor_{\theta=\theta_0}\right)^2,$$

stating any important assumptions that you may use. [3]

Evaluate the first and second derivative of $\mathcal{L}_T(\theta)$ wrt θ.

$$\frac{\partial \mathcal{L}_T}{\partial\theta} = \sum_{i=1}^{n} \delta_i \frac{1}{f}\frac{\partial f}{\partial\theta} + \sum_{i=1}^{n}(1-\delta_i)\frac{1}{\mathcal{F}(c;\theta)}\frac{\partial \mathcal{F}(c;\theta)}{\partial\theta},$$

$$\frac{\partial^2 \mathcal{L}}{\partial\theta^2} = -\sum_{i=1}^{n}\delta_i \frac{1}{f^2}\left(\frac{\partial f}{\partial\theta}\right)^2 - \sum_{i}(1-\delta_i)\frac{1}{\mathcal{F}(c;\theta)^2}\left(\frac{\partial \mathcal{F}(c;\theta)}{\partial\theta}\right)^2 + \sum_{i=1}^{n}\delta_i\frac{1}{f}\frac{\partial^2 f}{\partial\theta^2} + \sum_{i=1}^{n}(1-\delta_i)\frac{1}{\mathcal{F}(c;\theta)}\frac{\partial^2 \mathcal{F}(c;\theta)}{\partial\theta^2}.$$

Regularity condition we can exchange derivative and integral ie $\frac{\partial}{\partial\theta}\int f(x;\theta)dx = \int \frac{\partial f}{\partial\theta}dx$

$$\mathbb{E}\left\{\frac{\partial^2 f}{\partial\theta^2}\right\} = \boxed{n\left\{\int_0^c -\frac{1}{f^2}\left(\frac{\partial f}{\partial\theta}\right)^2 f\,dx - \frac{1}{\mathcal{F}(c;\theta)}\left(\frac{\partial \mathcal{F}(c;\theta)}{\partial\theta}\right)^2\right\}} \quad (1)$$

$$+ \underbrace{n\left\{\int_0^c \frac{\partial^2 f}{\partial\theta^2}dx + \frac{\partial^2 \mathcal{F}(c;\theta)}{\partial\theta^2}\right\}}$$

$$= \int_0^c \frac{\partial^2 f}{\partial\theta^2}dx + \frac{\partial^2}{\partial\theta^2}\int_c^\infty f(x)dx = \int_0^\infty \frac{\partial^2 f}{\partial\theta^2}dx = 0$$

since they are iid r.vs are the $\mathbb{E}\{\delta_i \frac{1}{f}(\frac{\partial f}{\partial\theta}) + (1-\delta)\frac{1}{\dots}\} = 0$

We now consider $\mathbb{E}\left\{\left(\frac{\partial \mathcal{L}}{\partial\theta}\right)^2\right\}$. Since δ_i and $(1-\delta_i)$ are disjoint r.vs we have

$$\mathbb{E}\left\{\left(\frac{\partial \mathcal{L}}{\partial\theta}\right)^2\right\} = \sum_{i=1}^{n}\left[\mathbb{E}\left\{\delta_i^2 \frac{1}{f^2}\left(\frac{\partial f}{\partial\theta}\right)^2\right\} + \mathbb{E}\left\{(1-\delta_i)^2 \frac{1}{\mathcal{F}(c;\theta)^2}\left(\frac{\partial \mathcal{F}(c;\theta)}{\partial\theta}\right)^2\right\}\right]$$

Using the same argument in (1)

$$= n\left\{\int_0^c \frac{1}{f^2}\left(\frac{\partial f}{\partial\theta}\right)^2 f\,dx - \frac{1}{\mathcal{F}(c;\theta)}\left(\frac{\partial \mathcal{F}(c;\theta)}{\partial\theta}\right)^2\right\}$$

Thus we have the required result.

(b) Let us suppose that the above survival times satisfy a Weibull distribution $f(x;\phi,\alpha) = (\frac{\alpha}{\phi})(\frac{x}{\phi})^{\alpha-1}\exp(-(x/\phi)^{\alpha})$ and as in part (a) we observe and $Y_i = \min(T_i, c)$ and δ_i, where $c > 0$.

(i) Using your answer in part 2a(i), give the log-likelihood of $\{(Y_i,\delta_i)\}_i$ for this particular distribution (we denote this as $\mathcal{L}_T(\alpha,\phi)$) and derive the profile likelihood of α (profile out the nusiance parameter ϕ).
Suppose you wish to test $H_0 : \alpha = 1$ against $H_A : \alpha \neq 1$ using the log-likelihood ratio test, what is the limiting distribution of the test statistic under the null? [3]

$$\mathcal{L}_T(\alpha,\phi) = \sum_{i=1}^{n} \delta_i \left\{ \log\alpha - \log\phi + (\alpha-1)\log Y_i - (\alpha-1)\log\phi - \left(\frac{Y_i}{\phi}\right)^{\alpha} \right\} - \sum_i (1-\delta_i)\left(\frac{c}{\phi}\right)^{\alpha}.$$

We now 'profile' out the 'nuisance' parameter ϕ

$$\frac{\partial \mathcal{L}}{\partial \phi} = \sum_{i=1}^{n}\left[\delta_i\left\{-\frac{1}{\phi} - \frac{(\alpha-1)}{\phi} + \alpha\frac{Y_i^{\alpha}}{\phi^{\alpha+1}}\right\} + (1-\delta_i)\alpha\frac{c^{\alpha}}{\phi^{\alpha+1}}\right\}.$$

Thus keeping α fixed the mle of ϕ (with fixed α) is

$$\hat{\phi}_{\alpha} = \left[\frac{\sum_i \delta_i X_i^{\alpha} + (1-\delta_i)c^{\alpha}}{\sum_i \delta_i}\right]^{1/\alpha}.$$ Thus the Profile likelihood for α is

$$\mathcal{L}_T(\alpha,\hat{\phi}_{\alpha}) = \sum_{i=1}^{n}\delta_i\left\{\log\alpha - \log\hat{\phi}_{\alpha} + (\alpha-1)\log Y_i - (\alpha-1)\log\hat{\phi}_{\alpha} - \left(\frac{Y_i}{\hat{\phi}_{\alpha}}\right)^{\alpha}\right\} + \sum_i(1-\delta_i)\frac{c^{\alpha}}{\hat{\phi}_{\alpha}^{\alpha}}\Big\}.$$

To test $H_0: \alpha = 1$ against $H_A: \alpha = 1$ (test whether distribution is exponential).

Using the log-likelihood ratio test, under the null we have:

$$2\left\{\max_{\alpha}\mathcal{L}_T(\alpha,\hat{\phi}_{\alpha}) - \max_{\phi}\mathcal{L}_T(1,\phi)\right\} \xrightarrow{D} \chi^2_1.$$

(ii) Let $\hat{\phi}_T, \hat{\alpha}_T = \arg\max \mathcal{L}_T(\alpha, \phi)$ (maximum likelihood estimators involving the censored likelihood). Do the estimators $\hat{\phi}_T$ and $\hat{\alpha}_T$ converge to the true parameters ϕ and α (you can assume that $\hat{\phi}_T$ and $\hat{\alpha}_T$ converge to some parameters, and your objective is to find whether these parameters are ϕ and α). [3]

It is very hard to prove this result for exactly this example. However if we show the result under a general set of assumptions, and show that the Weibull satisfy these, then we have proven the result.

We know that (it is stated in the question) that the ~~general~~ mle estimators converge to some constants/parameters (we do not need to prove this). The object is to find what these parameters are. ~~The limiting parameters are~~

The general survival function censored likelihood expectation is

$$\mathbb{E}\left\{ \delta_i \log f(T_i;\theta) + (1-\delta_i)\log \mathcal{F}(c_i;\theta)\right\}.$$

The parameter which maximises this is

$$\alpha = \arg\max \mathbb{E}\left\{ \delta \log f(T_i;\theta) + (1-\delta_i)\log \mathcal{F}(c_i;\theta)\right\}.$$

Now in the notes we have shown that this parameter is the true parameter of the distribution.

Since the Weibull distribution satisfies all the assumptions (we can exchange integral and derivative), then the above result implies $\hat{\alpha}_T, \hat{\phi}_T \longrightarrow \alpha, \phi$ (the true parameters).

Please refer to notes for the details.

(iii) Obtain the (expected) Fisher information matrix of maximum likelihood estimators. [3]

Differentiate $\mathcal{L}_T(\alpha, \phi)$ twice with respect to α, ϕ. and take expectations (not possible to obtain an explicit expression).

$$I(\alpha, \phi) = \begin{bmatrix} -\mathbb{E}\left(\frac{\partial^2 \mathcal{L}_T(\alpha,\phi)}{\partial \alpha^2}\right) & -\mathbb{E}\left(\frac{\partial^2 \mathcal{L}_T(\alpha,\phi)}{\partial \alpha \partial \phi}\right) \\ -\mathbb{E}\left(\frac{\partial^2 \mathcal{L}_T(\alpha,\phi)}{\partial \alpha \partial \phi}\right) & -\mathbb{E}\left(\frac{\partial^2 \mathcal{L}_T(\alpha,\phi)}{\partial \phi^2}\right) \end{bmatrix}$$

$$= \begin{bmatrix} I_{\alpha\alpha} & I_{\alpha\phi} \\ I_{\alpha\phi} & I_{\phi\phi} \end{bmatrix}.$$

(iv) Using your answer in part 2b(iii) derive the limiting variance of the maximum likelihood estimator of $\hat{\alpha}_T$. [3]

The limiting variance (at least variance of the limiting normal distribution) is

$$\frac{I_{\phi\phi}}{I_{\alpha\alpha} I_{\phi\phi} - I_{\alpha\phi}^2}$$

ie inverse of $\begin{bmatrix} a & b \\ c & d \end{bmatrix}^{-1}$

$$= \frac{1}{ad - cb} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$$