Exam 2 : Stat 426

12/13/2006

(1) The likelihood function $L(\theta)$ can be written as

$$L(\theta / x_1, \dots x_n) = \prod_{i=1}^{n} \frac{\theta^2}{2} 1\left(|x_i| \le \frac{1}{\theta^2}\right)$$

$$= \frac{\theta^{2n}}{2^n} 1\left(\max |x_i| \le \frac{1}{\theta^2}\right)$$

Thus, the value $\hat{\theta}_{MLE}$ that maximizes this function is

$$\hat{\theta}_{MLE} = \frac{1}{\sqrt{\max |x_i|}}$$

also, $|x_i| \sim \text{Unif}\left(0, \frac{1}{\theta^2}\right)$

Thus, $E[|x_i|] = \frac{1}{2\theta^2}$

i.e. $\theta = \frac{1}{\sqrt{2 E|x_i|}}$

Thus, a MOM estimator is

$$\hat{\theta}_{MOM} = \frac{1}{\sqrt{2 \cdot \frac{1}{n} \sum_{i=1}^{n} |x_i|}}$$

2.

$X_i \sim N(\mu_1, \sigma^2)$

$Y_i \sim N(\mu_2, \sigma^2)$

$cor(X_i, Y_i) = \rho$

Thus, $X_i - Y_i \sim N(\mu_1 - \mu_2,\ 2\sigma^2 + 2\sigma^2\rho)$

define, $\omega^2 = 2\sigma^2 + 2\sigma^2\rho$

then, $X_i - Y_i \sim N(\mu_1 - \mu_2, \omega^2)$ [ω^2 unknown]

(i) $$Pr\left[-t_{\alpha/2}^{(n)} \le \frac{(\bar{X} - \bar{Y}) - (\mu_1 - \mu_2)}{s/\sqrt{n}} \le t_{\alpha/2}^{(n)}\right] = 1 - \alpha$$

where $t_{\alpha/2}^{(n)}$ and $t_{(1-\alpha/2)}^{(n)}$ are quantiles of t' distribution with d.f. $= n$ and

$$s^2 = \frac{1}{n-1} \sum_{i=1}^{n} \{(X_i - Y_i) - (\bar{X} - \bar{Y})\}^2$$

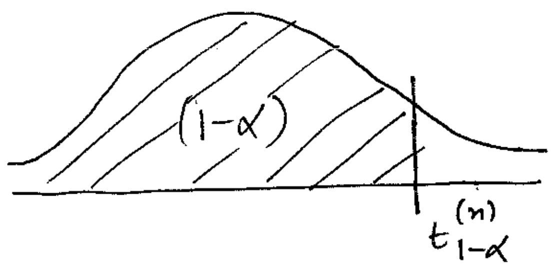
Thus, the required CI is

$$\left((\bar{X} - \bar{Y}) \pm t_{\alpha/2}^{(n)} \frac{s}{\sqrt{n}}\right)$$

(ii) We have ~~\~~ seen that,

$$T \doteq \frac{(\bar{X}-\bar{Y}) - (\mu_1-\mu_2)}{s/\sqrt{n}} \sim t^{(n)}$$

Now,



$$\Pr\left(T < t^{(n)}_{1-\alpha}\right) = 1-\alpha$$

where $t^{(n)}_{1-\alpha}$ is the $(1-\alpha)$ th quantile of a $t^{(n)}$ ~~r.v.~~ distribution.

Thus, $$\Pr\left(\frac{(\bar{X}-\bar{Y}) - (\mu_1-\mu_2)}{s/\sqrt{n}} < t^{(n)}_{1-\alpha}\right) = 1-\alpha$$

or, $$\Pr\left((\bar{X}-\bar{Y}) - t^{(n)}_{(1-\alpha)}\frac{s}{\sqrt{n}} < (\mu_1-\mu_2)\right) = 1-\alpha$$

Thus, $$\boxed{a = (\bar{X}-\bar{Y}) - t^{(n)}_{(1-\alpha)}\frac{s}{\sqrt{n}}}$$

where, as before,

$$s^2 = \frac{1}{n-1}\sum_{i=1}^{n}\left\{(X_i - Y_i) - (\bar{X}-\bar{Y})\right\}^2$$

(3) The likelihood function,

$$L(\theta / x_1, \dots, x_n) = \prod_{i=1}^{n} (\theta+1)\, x_i^{\theta}\, 1\{0 < x_i < 1\}\, 1\{\theta > 0\}$$

$$= (\theta+1)^n \left(\prod_{i=1}^{n} x_i\right)^{\theta} 1\{\min x_i > 0\}\, 1\{\max x_i < 1\}\, 1\{\theta > 0\}$$

$\therefore$ Log-likelihood $\ell(\theta) \equiv \ln [L(\theta / x_1, \dots, x_n)]$

$$\ell(\theta) = n \ln(\theta+1) - \theta \sum_{i=1}^{n} \ln x_i$$

$$\ell'(\theta) = \frac{n}{\theta+1} - \sum_{i=1}^{n} \ln x_i$$

$$\therefore \quad \hat{\theta}_{MLE} = \frac{n}{\sum_{i=1}^{n} \ln x_i} - 1 \quad \text{-------- (1)}$$

Fisher information,

$$I(\theta) = \cancel{-E_\theta\left[\frac{\partial^2 \ln}{\partial \theta^2} f(x,\theta)\right]}$$

$$= -E_\theta\left[\frac{\partial^2}{\partial \theta^2} \ln(f(X,\theta))\right]$$

$$= -E_\theta\left[\frac{\partial^2}{\partial \theta^2}[\ln(\theta+1) - \theta \ln X]\right]$$

$$= \frac{1}{(\theta+1)^2}$$

Now, we know

$$\sqrt{n}\,(g(\hat{\theta}_{MLE}) - g(\theta)) \Rightarrow N\left(0, \frac{[g'(\theta)]^2}{I(\theta)}\right)$$

Thus, if we choose

$$g'(\theta) = \sqrt{I(\theta)}$$

Then

$$\sqrt{n}\left(g(\hat{\theta}_{MLE}) - g(\theta)\right) \Rightarrow N(0,1)$$

and we have a pivot.

Thus, in this problem,

$$g'(\theta) = \frac{1}{\theta+1}$$

$$\therefore \; g(\theta) = \ln(\theta+1)$$

So,

$$\sqrt{n}\left(\ln(\hat{\theta}+1) - \ln(\theta+1)\right) \Rightarrow N(0,1)$$

Then, $\Pr\left[-z_{\alpha/2} \leq \sqrt{n}\,\ln\left(\frac{\hat{\theta}+1}{\theta+1}\right) \leq z_{\alpha/2}\right] = 1-\alpha$

i.e. $\Pr\left[\exp\left(-\frac{z_{\alpha/2}}{\sqrt{n}}\right) \leq \frac{\hat{\theta}+1}{\theta+1} \leq \exp\left(\frac{z_{\alpha/2}}{\sqrt{n}}\right)\right] = 1-\alpha$

i.e. $\Pr\left[\frac{\hat{\theta}+1}{\exp\left(-\frac{z_{\alpha/2}}{\sqrt{n}}\right)} \geq \theta+1 \geq \frac{\hat{\theta}+1}{\exp\left(\frac{z_{\alpha/2}}{\sqrt{n}}\right)}\right] = 1-\alpha$

Thus, the required C.I is

$$\left[\frac{\hat{\theta}+1}{\exp\left(\frac{z_{\alpha/2}}{\sqrt{n}}\right)} - 1, \; \frac{\hat{\theta}+1}{\exp\left(-\frac{z_{\alpha/2}}{\sqrt{n}}\right)} - 1\right]$$

where $\hat{\theta}$ is given by 1

(4) $X_i = \mathbb{1}(Y_i = 0)$

Thus, $X_i \sim \text{Bernoulli}(e^{-\theta})$

$$\left[\text{since } P(Y_i = 0) = \frac{e^{-\theta}\theta^0}{0!} = e^{-\theta}\right]$$

Now, if $X_i \sim \text{Bernoulli}(p)$

Likelihood fn.

$$L(\theta / x_1, \ldots, x_n) = \prod_{i=1}^{n} p^{x_i}(1-p)^{1-x_i}$$

Log-likelihood,

$$\ell(p) = \ln p \sum_{i=1}^{n} x_i + \ln(1-p)\sum_{i=1}^{n}(1-x_i)$$

$$\ell'(p) = \frac{\sum_{i=1}^{n} x_i}{p} - \frac{\sum_{i=1}^{n}(1-x_i)}{1-p}$$

For MLE, $\ell'(\hat{p}) = 0$

$$\therefore \quad \frac{\sum x_i}{\hat{p}} = \frac{\sum(1-x_i)}{1-\hat{p}}$$

or, $$\frac{\bar{X}}{\hat{p}} = \frac{1-\bar{X}}{1-\hat{p}}$$

or, $$\frac{1-\hat{p}}{\hat{p}} = \frac{1-\bar{X}}{\bar{X}}$$

or, $$\frac{1}{\hat{p}} - 1 = \frac{1}{\bar{X}} - 1 \Rightarrow \boxed{\hat{p} = \bar{X}}$$

Thus, since $p = e^{-\theta}$ in this question,

$$e^{-\hat{\theta}_X} = \overline{X} \qquad \text{(continuous mapping)}$$

$$\Rightarrow \boxed{\hat{\theta}_X = -\ln \overline{X}}$$

where $\hat{\theta}_X$ is the MLE based on X_i's.

Now, if Y_i's are available,

$$L(\theta / Y_1, \ldots, Y_n) = \prod_{i=1}^{n} \frac{e^{-\theta}\, \theta^{Y_i}}{\underline{Y_i}}$$

$$\therefore \quad l(\theta) = -n\theta + \ln \theta . \sum_{i=1}^{n} Y_i - \ln \left(\prod_{i=1}^{n} Y_i ! \right)$$

For MLE $l'(\hat{\theta}_Y) = 0$

i.e.

$$-n + \frac{1}{\hat{\theta}_Y} \sum_{i=1}^{n} Y_i = 0$$

$$\Rightarrow \boxed{\hat{\theta}_Y = \overline{Y}}$$, $\hat{\theta}_Y$ is the MLE based on Y_i's

<u>Qualitative comparison</u>:

Y_is are the full data, where X_is only tell if Y_i's are 0 or not. Thus, Y_i's have more information and an estimator based on Y_i's will be <u>more efficient</u>, in the sense that we will require fewer samples to get comparable standard error for our estimate than if we base our estimator on X_i's. This ratio of sample sizes can be computed explicitly, but is not necessary here.

(5) Let T_i be the time between $(i-1)$th and i'th emissions $(i \geqslant 1)$

then, $T_i \sim \exp(\text{mean} = \beta)$

Let $S_n = \sum_{i=1}^{n} T_i$

Thus, $S_n \sim \text{Gamma}\left(n, \frac{1}{\beta}\right)$

Thus, data available to the physicist $= S_n$.

Now, $\frac{S_n}{\beta} \sim \text{Gamma}(n, 1)$

Suppose c_1 and c_2 denote the $\alpha/2$ th and $(1-\alpha/2)$ th quantile's of Gamma $(n, 1)$

Thus, $\Pr\left[c_1 \leq \frac{S_n}{\beta} \leq c_2\right] = 1-\alpha$

Thus, the required CI is

$$\left[\frac{S_n}{c_2}, \frac{S_n}{c_1}\right]$$