STAT 721

Solutions to Midterm #1

marks

(10) (1) $\left.\begin{array}{l} y=\sqrt{x_1/x_2} \\ z=x_2 \end{array}\right\}$ is a 1-1 map from $(\mathbb{R}^+)\times\mathbb{R}^+$ to $\mathbb{R}^+\times\mathbb{R}^+$.

The inverse map is $\begin{array}{l} x_1 = zy^2 \\ x_2 = z \end{array}$ ← (2), $J = \begin{vmatrix} 2zy & y^2 \\ 0 & 1 \end{vmatrix} = 2zy$ ← (2)

$$f(x_1,x_2) = \frac{1}{2\pi}(x_1x_2)^{-1/2} e^{-(x_1+x_2)/2}, \quad x_1>0,\ x_2>0.$$

So the joint pdf of Y and Z is:

$$g(y,z) = \frac{1}{2\pi}(z^2y^2)^{-1/2} e^{-(zy^2+z)/2}\, 2zy$$

$$= \frac{1}{\pi} e^{-z(y^2+1)/2}, \quad 0<z<\infty,\ 0<y<\infty$$ ← (2)

So $g(y) = \int g(y,z)\,dz = \int_0^\infty e^{-z(y^2+1)/2}\,dz$ ← (2)

$$= \frac{1}{\pi}\int_0^\infty \frac{2}{y^2+1}\, d\left(-e^{-z(y^2+1)/2}\right)$$

$$= \frac{2}{\pi(y^2+1)}, \quad 0<y<\infty$$ ← (2)

[This is a "half-Cauchy" distribution].

(10) 2 4 (a) $\prod_{i=1}^{n} f(x_i|\theta) = e^{-\sum x_i + n\theta}$ if $x_1 \geq \theta, x_2 \geq \theta, \ldots, x_n \geq \theta$

$= 0$ elsewhere

Since $\{x_1 \geq \theta, \ldots, x_n \geq \theta\} \Leftrightarrow x_{(1)} \geq \theta$, we can write the joint pdf as:

$$\prod f(x_i|\theta) = e^{-\sum x_i} e^{n\theta} I_{[x_{(1)} - \theta]}$$

By the factorization theorem, $X_{(1)}$ is sufficient for θ.

3 → (b) For $x \in [\theta, \infty)$, the dist. function of $X_{(1)}$ at x is:

$$F_\theta(x) = 1 - P_\theta(X_{(1)} \geq x) = 1 - \prod_{i=1}^{n} P_\theta[X_i \geq x]$$

$$= 1 - [P_\theta(X_{(1)} - \theta \geq x - \theta)]^n = 1 - \left[\int_{x-\theta}^{\infty} e^{-v}\, dv\right]^n$$

$$= 1 - e^{-n(x-\theta)}.$$

So the pdf of $X_{(1)}$ is $f_\theta(x) = F'_\theta(x) = n e^{-n(x-\theta)}$, $x \geq \theta$

$= 0$ elsewhere.

3 → (c) By the definition of completeness, we need to show that if $E_\theta\, g(X_{(1)}) = 0$ for all θ, then it follows that $g(x) = 0$ for all x.

(continued)

(2) (c) (continued)

$$E_\theta g(X_{(1)}) = 0 \text{ for all } \theta \Rightarrow \int_\theta^\infty n e^{-nx} e^{n\theta} g(x)\,dx = 0 \text{ for all } \theta$$

$$\Rightarrow \int_\theta^\infty e^{-nx} g(x)\,dx = 0 \text{ for all } \theta \quad (\text{since } n e^{n\theta} > 0 \text{ for all } \theta).$$

$$\Rightarrow 0 = \frac{d}{d\theta}\int_0^\infty e^{-nx} g(x)\,dx = -e^{-n\theta} g(\theta) = 0 \text{ for all } \theta$$

$$\Rightarrow g(x) = 0 \text{ for all } x, \ -\infty < x < \infty.$$

Hence $X_{(1)}$ is complete

(10) → (3)

(5) → (a) $\prod f(x_i|\theta) = 2^n \prod x_i / \theta^{2n}$ $\quad 0 < X_{(n)} < \theta$

$= 0$ $\quad$ elsewhere

or $\prod f(x_i|\theta) = \left[2^n \left(\prod x_i\right) / \theta^{2n}\right] I_{[X_{(n)} - \theta]}$

By the Factorization Theorem, $X_{(n)}$ is sufficient. ← (2)

$P[X_i \le \theta] = \frac{x^2}{\theta^2}$ for $0 < x < \theta$, so

$P[X_{(n)} \le \theta] = \frac{x^{2n}}{\theta^{2n}}$, $0 < x < \theta$.

So the pdf of $X_{(n)}$ is

$$\boxed{f(x) = \frac{2n x^{2n-1}}{\theta^{2n}}, \ 0 < x < \theta} \leftarrow (1)$$

(continued)

(3) (a) (continued)

(2)

If $E_\theta[g(X_{(n)})] = 0$ for all θ,

then $\int_0^\theta \frac{2n x^{2n-1}}{\theta^{2n}} g(x)\,dx = 0$ for all θ,

then $\int_0^\theta x^{2n-1} g(x)\,dx = 0$ for all θ,

then $0 = \frac{d}{d\theta}\int_0^\theta x^{2n-1} g(x)\,dx = 0 = \theta^{2n-1} g(\theta)$ for all $\theta > 0$,

so $g(\theta) = 0$ for all $\theta > 0$, OR $g(x) = 0$ for all $x > 0$.
This shows that $X_{(n)}$ is complete.

(5) → (b) The pdf of a sample X_i is

$$f(x_i) = \theta x_i^{\theta-1} = \theta e^{(\theta-1)\log x_i}, \quad 0 < x_i < 1$$

This is therefore a one-dimensional regular exponential family.

It follows by a Theorem about random samples $X_1, \ldots, X_n$ from a regular exponential family that $T = \sum_{i=1}^n \log X_i$ (or equivalently $T' = \prod_{i=1}^n X_i$) is a complete sufficient statistic for θ.

(10) 4 $\prod f(x_i|\theta) = \left(\frac{1}{\sqrt{2\pi\theta^2}}\right)^n e^{-\frac{1}{2\theta^2}\sum(x_i-\theta)^2}$

$$= \left(\frac{1}{\sqrt{2\pi\theta^2}}\right)^n e^{-n/2}\, e^{-\frac{1}{2\theta^2}\sum x_i^2 + \frac{1}{\theta}\sum x_i}$$

~~By the factorization theorem:~~

(4) Hence the family of distributions is a two-dimensional exponential family with sufficient statistic $(\sum X_i, \sum X_i^2)$. Since $(\bar{X}, S^2)$ is a 1-1 function of $(\sum X_i, \sum X_i^2)$, $(\bar{X}, S^2)$ is also a sufficient statistic for θ.

[Another solution is to rewrite the exponential family representation so that $\bar{X}, S^2$ are seen directly to be sufficient statistics.]

(3) The family of distributions is not complete, because the set of parameter $\{(\theta, \theta^2) : \theta > 0\}$ does not contain a 2-dimensional rectangle.

[Another solution is to construct a function of $\bar{X}$ and S^2, $g(\bar{X}, S^2)$ which does not vanish but for which $E_\theta[g(\bar{X}, S^2)] = 0$, for all θ. There are many such functions g.]