1 1

$X_i \overset{\text{iid}}{\sim} Pois(\lambda), \lambda > 0$. Use only the definition of sufficiency to prove that $T(X) = \sum_{i=1}^{n} X_i$ is sufficient for λ

<u>Class Definition</u>: A statistic T is sufficient if the distribution of $X|T$ does not depend on θ.

Because the X_is are independent, the sum is also Poisson: $T(X) \sim Pois(n\lambda)$

$P(\underset{\sim}{\boldsymbol{X}} = \underset{\sim}{\boldsymbol{x}}|T(X)) = \frac{P(\underset{\sim}{\boldsymbol{X}}=\underset{\sim}{\boldsymbol{x}})}{P(T(X)=t)} = \frac{\prod \frac{e^{-\lambda}\lambda^{x_i}}{x_i!}}{\frac{e^{-n\lambda}(n\lambda)^t}{t!}} = \frac{e^{-n\lambda}\lambda^{\sum x_i} t!}{\prod x_i! e^{-n\lambda} n^t \lambda^t}$

We defined above that in this case, $t = \sum x_i$, otherwise the above is zero/doesn't mean anything:

$\therefore = \frac{\lambda^{\sum x_i} t!}{\prod x_i! n^t \lambda^{\sum x_i}} = \underline{\frac{t!}{\prod x_i! n^{\sum x_i}}}$

That final expression does not depend on λ, thus $T(X) = \sum x_i$ is a sufficient statistic.

2 CB2 6.1

X is one observation from $N(0, \sigma^2)$. Is $|X|$ a sufficient statistic?

We can ignore the front coefficient that only depends on σ^2:

$X = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$ with $\mu = 0 \implies X = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{x^2}{2\sigma^2}}$

Because the mean is 0, that leaves only x^2 in the numerator of the exponent, which we can also rewrite as $|x|^2$, thus:

$X = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{|x|^2}{2\sigma^2}} = g(T(X), \sigma^2)h(x), h(x) = 1$

Because we can factor it, that means $T(X) = |X|$ is a sufficient statistic (though I think that only holds when $\mu = 0$).

3 CB2 6.2

$f_{X_i}(x|\theta) = \begin{cases} e^{i\theta - x} & x \geq i\theta \\ 0 x < i\theta \end{cases}$

Prove that $T = min_i(\frac{X_i}{i})$ is a suffucient statistic for θ

$f(\underset{\sim}{\boldsymbol{x}}|\theta) = \prod e^{i\theta - x} I_{(x_i \geq i\theta)}$

For the indicator to hold, we need all $x_i \geq i\theta$, which means we just need the $min(x_i) \geq \theta$, so we can plug in $min(x_i)$ in that portion:

$= e^{i\theta - \sum x_i} I_{x_{i_{(1)}} \geq i\theta} = \underline{[e^{i\theta} I_{x_{i_{(1)}} \geq i\theta}] e^{-\sum x_i}} = g(T(X), \theta)h(x)$

4 CB2 6.6

$X \sim Gamma(\alpha, \beta)$. Find a 2D sufficient statistic

Gamma is an exponential family, so we can just put it into exponential family form:

$L(\alpha, \beta|X) \propto \prod x_i^{\alpha-1} e^{-\beta x_i} = \prod e^{(\alpha-1)ln(x_i) - \beta x_i} = e^{(\alpha-1)\sum ln(x_i) - \beta \sum x_i}$

The sufficient statistics are all the $T_i(X)$ in the exponent, so the sufficient statistic is:

$T(X) = (\sum x_i, \sum ln(x_i))$

5 CB2 6.13

$X_1, X_2 \overset{\text{iid}}{\sim} \alpha x^{\alpha-1} e^{-x^\alpha}, x > 0, \alpha > 0$. Show that $\frac{ln(X_1)}{ln(X_2)}$ is ancillary.

Let $y = ln(x) \implies f(y|\alpha) = \alpha(e^y)^{\alpha-1} e^{-(e^y)^\alpha} e^y = \alpha e^{\alpha y - y + y - e^{-y\alpha}} = \alpha e^{\alpha y - e^{-y\alpha}}$

This can be rewritten as a scale-family: $\frac{1}{1/\alpha} e^{\frac{y}{1/\alpha} - e^{(-\frac{y}{1/\alpha})}}$

This is a scale transformation of some variable Z which is not dependent on α (ie. $Z \sim e^{z - e^{-z}}$)

Thus: $\frac{ln(X_1)}{ln(X_2)} = \frac{Y_1}{Y_2} = \frac{(1/\alpha)Z_1}{(1/\alpha)Z_2} = \frac{Z_1}{Z_2}$. This does not depend on α, thus $\frac{ln(X_1)}{ln(X_2)}$ is an ancillary statistic.

6 6

$X \sim Pois(\lambda)$. Cannot use 6.2.25, the theorem of completeness for exponential families.

i)

Show that $T(X) = \sum_{i=1}^{n} X_i$ is a complete statistic when $\lambda > 0$

Assume $g(t)$ is continuous, $0 = \mathbb{E}_\theta[g(T)] = \int_0^\lambda g(t) \sum X_i dt = \int_0^\lambda g(t) \frac{(n\lambda)^t e^{-n\lambda}}{t!} dt$

Take the derivative of both sides:

$0 = \frac{d}{d\lambda} e^{-n\lambda} \int_0^\lambda g(t) \frac{(n\lambda)^t}{t!} dt = -n\lambda e^{-n\lambda} \int_0^\lambda g(t) \frac{(n\lambda)^t}{t!} dt + e^{-n\lambda} g(\lambda) \frac{(n\lambda)^\lambda}{\lambda!} = \int_0^\lambda g(t) \frac{(n\lambda)^t e^{-n\lambda}}{t!} dt + e^{-n\lambda} g(\lambda) \frac{(n\lambda)^\lambda}{\lambda!}$

Because we defined that integral portion to be 0: $\implies 0 = e^{-n\lambda} g(\lambda) \frac{(n\lambda)^\lambda}{\lambda!}$

$e^{-n\lambda}$, $\frac{(n\lambda)^\lambda}{\lambda!}$ can never be 0 (because $\lambda > 0$), thus we must have $g(\lambda) = 0$, implying $T(X)$ is complete.

ii)

Show that for any integer $r > 0$: $\mathbb{E}[X_i(X_i - 1)(X_i - 2)...(X_i - r + 1)] = \lambda^r$

$\mathbb{E}[X_i(X_i - 1)(X_i - 2)...(X_i - r + 1)] = \sum_{x=0}^{\infty} x(x-1)(x-2)...(x-r+1) \frac{e^{-\lambda}\lambda^x}{x!} = \sum_{x=0}^{\infty} \frac{x!}{(x-r)!} \frac{e^{-\lambda}\lambda^x}{x!} = \sum_{x=0}^{\infty} \frac{e^{-\lambda}\lambda^x}{(x-r)!}$

Let $j = x - r$: $\implies \mathbb{E}[X_i(X_i - 1)(X_i - 2)...(X_i - r + 1)] = \sum_{j=0}^{\infty} \frac{e^{-\lambda}\lambda^{j+r}}{j!} = e^{-\lambda}\lambda^r \sum_{j=0}^{\infty} \frac{\lambda^j}{j!}$

I needed help to see that that summation term is the Taylor Series expansion of e^λ:

$\implies \underline{\mathbb{E}[X_i(X_i - 1)(X_i - 2)...(X_i - r + 1)] = e^{-\lambda}\lambda^r e^\lambda = \lambda^r}$

iii)

Restrict $\lambda \in 1, 3$. Show that the $T(X)$ from part (i) is not a complete statistic.

As shown above, $P_\lambda(T=t) = \sum \frac{(n\lambda)^t e^{-n\lambda}}{t!}$

$\mathbb{E}_\lambda[g(T)] = \sum_{t=0}^{\infty} g(t)\frac{(n\lambda)^t e^{-n\lambda}}{t!} \implies \sum_{t=0}^{\infty} g(t)\frac{(n\lambda)^t}{t!} = 0$

This must be equal to zero for both $\lambda = 1, 3$.

(I had to get help to see that I needed to just set anything beyond $t > 2$ to just $g(t) = 0$). Let $g(t) = 0$ when $t > 2$. We then have:

$\lambda = 1 : g(0)\frac{(n)^0}{0!} + g(1)\frac{(n)^1}{1!} + g(2)\frac{(n)^2}{2!} = g(0) + ng(1) + \frac{n^2}{2}g(2)$

$\lambda = 3 : g(0)\frac{(3n)^0}{0!} + g(1)\frac{(3n)^1}{1!} + g(2)\frac{(3n)^2}{2!} = g(0) + 3ng(1) + \frac{9n^2}{2}g(2)$

Subtracting the two we have:

$2ng(1) + \frac{8n^2}{2}g(2) = 0 \implies g(1) = -2ng(2)$

$\implies g(0) - 2n^2g(2) + \frac{n^2}{2}g(2) = 0 \implies g(0) = \frac{3n^2}{2}g(2)$

In order for all the baove to hold, pick $g(2) = 2$. This then means $g(1) = -4n, g(0) = 3n^2$.

Thus we have constructed a $g(t)$ that is not just the 0 function, but that fulfills all the requirements where the expected value is 0. Thus, T is not a complete statistic when $\lambda \in \{1, 3\}$

7 CB2 6.15

$X_i \overset{\text{iid}}{\sim} N(\theta, a\theta^2), a$ is a known constant $> 0, \theta > 0$

a)

Show the parameter space does not contain a two-dimensional open set.

I can't easily draw a graph here, but when imagining the parameter space, we basically have a quadratic equation of $\theta, a\theta^2$, meaning for a θ, the parameter space is just a line, which is not an open set (ie. we can't draw a tiny circle and stay in the set, to use the description Dr. Petersen provided).

b)

Show that $T = (\bar{X}, S^2)$ is sufficient, but not complete.

$L(\theta|\underset{\sim}{\boldsymbol{X}}) = (2\pi a\theta^2)^{n/2} e^{-\frac{1}{2a\theta^2}\sum(x-\theta)^2} = (2\pi a\theta^2)^{n/2} e^{-\frac{1}{2a\theta^2}\sum(x-\bar{x}+\bar{x}-\theta)^2} = (2\pi a\theta^2)^{n/2} e^{-\frac{1}{2a\theta^2}\left(\sum(x_i-\bar{x})^2 + n(\bar{x}-\theta)^2\right)}$

$t_1 = \bar{X}, t_2 = S^2 = \frac{1}{n-1}\sum(x_i - \bar{X})^2 \implies L(\theta|\underset{\sim}{\boldsymbol{X}}) = (2\pi a\theta^2)^{n/2} e^{-\frac{1}{2a\theta^2}\left((n-1)t_2 + n(t_1-\theta)^2\right)}$

Let $h(x) = 1$:

Thus, we can factor this into $L(\theta|\underset{\sim}{\boldsymbol{X}}) = h(x)g(t_1, t_2, \theta)$ implying that $T = (\bar{X}, S^2)$ is sufficient.

To show it's incomplete, we can just use the fact that this is part of the exponential family, and since the parameter space does not include an open set, we know that T is not complete.

8 CB2 6.22

$f(x|\theta) = \theta x^{\theta-1}, \theta > 0$

a) Is $\sum X_i$ a sufficient statistic?

$L(\theta|x) = \prod \theta x_i^{\theta-1} = \theta^n \prod e^{(\theta-1)ln(x_i)} = \theta^n e^{(\theta-1)\sum ln(x_i)}$

$\sum ln(x_i)$ is sufficient, but there is no direct linear transformation to $\sum x_i$, meaning $\sum x_i$ is not sufficient.

b) Find a complete sufficient statistic

We just showed about that $\sum ln(x_i)$ is sufficient. We also can show that $f(x|\theta) = \theta x^{\theta-1} = \theta e^{(\theta-1)\ln(x)}$ which means it is part of the exponential family. Because the parameter space is $\theta > 0$, that means that it contains an open set. Thus, we can use the exponential family rule of open sets to say that $\sum ln(x_i)$ is also complete.

Thus, $\sum ln(x_i)$ is a complete, sufficient statistic.