Econ 6190 Problem Set 3

Fall 2024

1. [Hong 6.8] Establish the following recursion relations for sample means and sample variances. Let $\bar{X}_n$ and s_n^2 be the sample mean and sample variances based on random sample $\{X_1, X_2 \dots X_n\}$. Then suppose another observation, X_{n+1}, becomes available. Show:

   (a) $\bar{X}_{n+1} = \frac{X_{n+1} + n\bar{X}_n}{n+1}$.

   (b) $ns_{n+1}^2 = (n-1)s_n^2 + \frac{n}{n+1}(X_{n+1} - \bar{X}_n)^2$.

2. [Hong 6.6] Suppose $\mathbf{X}^n = (X_1, \dots, X_n)$ is an iid $\mathrm{N}(\mu, \sigma^2)$ random sample, $\mathbf{Y}^n = (Y_1, \dots, Y_n)$ is an iid $\mathrm{N}(\mu, \sigma^2)$ random sample, and the two random samples are mutually independent. Let $\bar{X}_n$ and $\bar{Y}_n$ be the sample means of the first and second random samples, respectively, and let s_X^2 and s_Y^2 be the sample variances of the first and second random samples respectively. Find:

   (a) the distribution of $(\bar{X}_n - \bar{Y}_n)/\sqrt{2\sigma^2/n}$;

   (b) the distribution of $(\bar{X}_n - \bar{Y}_n)/\sqrt{2s_X^2/n}$;

   (c) the distribution of $(\bar{X}_n - \bar{Y}_n)/\sqrt{2s_Y^2/n}$;

   (d) the distribution of $(\bar{X}_n - \bar{Y}_n)/\sqrt{(s_X^2 + s_Y^2)/n}$;

   (e) the distribution of $(\bar{X}_n - \bar{Y}_n)/\sqrt{s_n^2/n}$, where s_n^2 is the sample variance of the difference sample $\mathbf{Z}^n = (Z_1, Z_2 \dots Z_n)$, where $Z_i = X_i - Y_i$, $i = 1, 2 \dots n$.

3. [Hong 6.9] Let $X_i, i = 1, 2, 3$ be independent with $\mathrm{N}(i, i^2)$ distributions. For each of the following situations, use X_1, X_2, X_3 to construct a statistic with the indicated distribution:

   (a) Chi-square distribution of 3 degrees of freedom;

   (b) t distribution with 2 degrees of freedom;

4. [Final exam, 2022] Let $\{X_1, \dots X_n\}$ be i.i.d with pdf $f(x \mid \theta) = e^{-(x-\theta)}\mathbf{1}\{x \geq \theta\}$. Show $Y = \min\{X_1, \dots X_n\}$ is a sufficient statistic for θ **without** using the Factorization Theorem.

5. Let $\{X_1, \ldots, X_n\}$ be a random sample with the pdf for each X_i

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

   Show $\min_i \left(\frac{X_i}{i}\right)$ is a sufficient statistic for θ.

6. Show that the following claim is true: any one-to-one function of a sufficient statistic is a also sufficient statistic.

7. Let X be one observation from $\mathrm{N}(0, \sigma^2)$. Is $|X|$ a sufficient statistic for σ^2? Give your reasoning clearly.

1\. (a). $(n+1)\,\bar{X}_{n+1} = n\,\bar{X}_n + X_{n+1}$

then $\bar{X}_{n+1} = \dfrac{X_{n+1} + n\,\bar{X}_n}{n+1}$

(b).
$$n\,S_{n+1}^2 = \sum_{i=1}^{n+1} (X_i - \bar{X}_{n+1})^2$$
$$= \sum_{i=1}^{n+1} (X_i - \bar{X}_n + \bar{X}_n - \bar{X}_{n+1})^2$$
$$= \sum_{i=1}^{n+1} (X_i - \bar{X}_n)^2 + \sum_{i=1}^{n+1} (\bar{X}_n - \bar{X}_{n+1})^2 + 2\sum_{i=1}^{n+1} (X_i - \bar{X}_n)\cdot(\bar{X}_n - \bar{X}_{n+1})$$
$$= \sum_{i=1}^{n+1} (X_i - \bar{X}_n)^2 - (n+1)(\bar{X}_n - \bar{X}_{n+1})^2$$
$$= \sum_{i=1}^{n} (X_i - \bar{X}_n)^2 + (X_{n+1} - \bar{X}_n)^2 - (n+1)\underline{(\bar{X}_n - \bar{X}_{n+1})^2}$$

$$(\bar{X}_n - \bar{X}_{n+1})^2 = \left(\bar{X}_n - \frac{X_{n+1} + n\bar{X}_n}{n+1}\right)^2$$
$$= \left(\frac{1}{n+1}\right)^2 (\bar{X}_n - X_{n+1})^2$$

$$= \sum_{i=1}^{n} (X_i - \bar{X}_n)^2 + (X_{n+1} - \bar{X}_n)^2 - \frac{1}{n+1}(\bar{X}_n - X_{n+1})^2$$
$$= (n-1)\,S_n^2 + \frac{n}{n+1}(X_{n+1} - \bar{X}_n)^2$$

2\. (a). $\bar{X}_n \sim N(\mu, \frac{\sigma^2}{n})$, $\bar{Y}_n \sim N(\mu, \frac{\sigma^2}{n})$

$(\bar{X}_n - \bar{Y}_n) \sim N(0, \frac{2\sigma^2}{n})$

Hence. $\dfrac{\bar{X}_n - \bar{Y}_n}{\sqrt{2\sigma^2/n}} \sim N(0, 1)$

(b). $$\frac{\bar{X}_n - \bar{Y}_n}{\sqrt{2\sigma^2/n}} \sim N(0, 1)$$

$$S_X^2 = \frac{1}{n-1} \sum_{i=1}^{n} (X_i - \bar{X}_n)^2$$

and $$\frac{(n-1) S_X^2}{\sigma^2} \sim \chi^2_{n-1}$$

$$\frac{\bar{X}_n - \bar{Y}_n}{\sqrt{2 S_X^2/n}} = \frac{(\bar{X}_n - \bar{Y}_n)/\sqrt{2\sigma^2/n}}{\sqrt{2S_X^2/n} \Big/ \sqrt{2\sigma^2/n}} = \frac{(\bar{X}_n - \bar{Y}_n)/\sqrt{2\sigma^2/n}}{\sqrt{(n-1) S_X^2/\sigma^2} \Big/ \sqrt{n-1}}$$

then $$\frac{\bar{X}_n - \bar{Y}_n}{\sqrt{2 S_X^2/n}} \sim t_{n-1}$$ since $\bar{X}_n - \bar{Y}_n$ and S_X^2 are mutually independent.

(c). similar as above.

$$\frac{\bar{X}_n - \bar{Y}_n}{\sqrt{2 S_Y^2/n}} \sim t_{n-1}$$

(d). $$\frac{(\bar{X}_n - \bar{Y}_n)}{\sqrt{(S_X^2 + S_Y^2)/n}} = \frac{(\bar{X}_n - \bar{Y}_n)/\sqrt{2\sigma^2/n}}{\sqrt{(S_X^2 + S_Y^2)/2\sigma^2}}$$

$$= \frac{(\bar{X}_n - \bar{Y}_n)/\sqrt{2\sigma^2/n}}{\sqrt{\dfrac{(n-1)(S_X^2 + S_Y^2)}{(2n-2)\sigma^2}}} \sim t_{2n-2}$$

[since $\bar{X}_n - \bar{Y}_n$ and $S_X^2 + S_Y^2$ are mutually independent.]

(e). $$Z_i \sim N(0, 2\sigma^2)$$

$$S_n^2 = \frac{1}{n-1} \sum_{i=1}^{n} (Z_i - \bar{Z}_n)^2$$

$$\frac{(n-1) S_n^2}{2\sigma^2} \sim \chi^2_{n-1}$$

then $$\frac{(\bar{X}_n - \bar{Y}_n)}{\sqrt{S_n^2/n}} = \frac{(\bar{X}_n - \bar{Y}_n)/\sqrt{2\sigma^2/n}}{\sqrt{(n-1)S_n^2/(2\sigma^2)(n-1)}} \sim t_{n-1}$$

[since $\bar{X}_n - \bar{Y}_n$ and S_n^2 are mutually independent.]

3.

(a). $X_i \sim N(i, i^2)$. then $\frac{X_i - i}{i} \sim N(0,1)$

then we know $\sum_{i=1}^{3} \left(\frac{X_i - i}{i}\right)^2 \sim \chi^2_3$

(b). $\sum_{i=1}^{2} \left(\frac{X_i - i}{i}\right)^2 \sim \chi^2_2$ and $\frac{X_3 - 3}{3} \sim N(0,1)$

then $T = \frac{(X_3 - 3)/3}{\sqrt{\sum_{i=1}^{2} \left(\frac{X_i - i}{i}\right)^2 / 2}} \sim t_2$

4.

Answer: The joint density of $\{X_1, \dots X_n\}$ *is*

$$
\begin{aligned}
f(x_1, \dots x_n) &= \Pi_{i=1}^n e^{-(x_i - \theta)} \mathbf{1}\{x_i \geq \theta\} \\
&= e^{-\sum_{i=1}^n (x_i - \theta)} \left(\Pi_{i=1}^n \left(\mathbf{1}\{x_i \geq \theta\}\right)\right) \\
&= \begin{cases} e^{-\sum_{i=1}^n x_i} e^{n\theta} & \min\{x_1, \dots x_n\} \geq \theta \\ 0 & otherwise \end{cases},
\end{aligned}
$$

which can be written as

$$
f(x_1, \dots x_n) = e^{-\sum_{i=1}^n x_i} \mathbf{1}\left\{\min\{x_1, \dots x_n\} \geq \theta\right\} e^{n\theta}.
$$

Next, we derive the pdf of $Y = \min\{X_1, \dots X_n\}$. *To do so, note the cdf of* Y *is*

$$
\begin{aligned}
P\{Y \leq y\} &= P\{\min\{X_1, \dots X_n\} \leq y\} \\
&= 1 - P\{\min\{X_1, \dots X_n\} > y\} \\
&= 1 - P\{X_1 > y, \dots X_n > y\} \\
&= 1 - \prod_{i=1}^n P\{X_i > y\} \\
&= 1 - \prod_{i=1}^n \int_y^\infty e^{-(t-\theta)} \mathbf{1}\{t \geq \theta\} dt.
\end{aligned}
$$

Note

$$\int_y^\infty e^{-(t-\theta)}\mathbf{1}\{t \geq \theta\}dt = \begin{cases} \int_\theta^\infty e^{-(t-\theta)}dt & y < \theta \\ \int_y^\infty e^{-(t-\theta)}dt & y \geq \theta \end{cases}$$
$$= \begin{cases} 1 & y < \theta \\ e^{-(y-\theta)} & y \geq \theta \end{cases}.$$

Thus,

$$P\{Y \leq y\} = \left(1 - e^{-n(y-\theta)}\right)\mathbf{1}\{y \geq \theta\}.$$

And the pdf of Y *is* $f_Y(y) = \left(ne^{-n(y-\theta)}\right)\mathbf{1}\{y \geq \theta\}$ *by taking derivatives. Therefore,*

$$\frac{f(x_1,\ldots x_n)}{f_Y(y)} = \frac{e^{-\sum_{i=1}^n x_i}\mathbf{1}\{\min\{x_1,\ldots x_n\} \geq \theta\}e^{n\theta}}{\left(ne^{-n(\min\{x_1,\ldots x_n\}-\theta)}\right)\mathbf{1}\{\min\{x_1,\ldots x_n\} \geq \theta\}}$$
$$= \frac{e^{-\sum_{i=1}^n x_i}}{ne^{-n(\min\{x_1,\ldots x_n\}}}, \text{ for } \min\{x_1,\ldots,x_n\} \geq \theta$$

which does not depend on θ*. Thus,* $\min\{X_1,\ldots X_n\}$ *is a sufficient statistic.*

5.

$$f(\vec{x}, \theta) = \begin{cases} \prod_i e^{i\theta - x_i} & \forall i,\ \frac{x_i}{i} \geq \theta \iff \min\left(\frac{x_i}{i}\right) \geq \theta \\ 0 & \text{otherwise} \end{cases}$$

$$= \begin{cases} e^{\theta\Sigma i - \Sigma x_i} & \min\left(\frac{x_i}{i}\right) \geq \theta \\ 0 & 0 \end{cases}$$

then let $g(\vec{x} \mid \theta) = \begin{cases} e^{\theta\Sigma i} & \min\left(\frac{x_i}{i}\right) \geq \theta \\ 0 & 0 \end{cases}$

$$h(\vec{x}) = e^{-\Sigma x_i}$$

By factorization theorem, we know $\min\left(\frac{x_i}{i}\right)$ is s.s.

6.

suppose $T(\vec{x})$ is a sufficient statistic

then $f(\vec{x} | \theta) = g(T(\vec{x}) | \theta) \cdot h(\vec{x})$

let S be the one-to-one function of $T(\vec{x})$

and $S[T(\vec{x})] = T_1(\vec{x})$

then $T(\vec{x}) = S^{-1}(T_1(\vec{x}))$ due to one-to-one mapping

then $f(\vec{x} | \theta) = g(S^{-1}[T_1(\vec{x})] | \theta) \cdot h(\vec{x})$

$= g_1(T_1(\vec{x}) | \theta) \cdot h(\vec{x})$ #.

7.

$$f(x, \sigma^2) = \frac{1}{\sqrt{2\pi}} \cdot e^{-\frac{x^2}{2\sigma^2}}$$

$$= \frac{1}{\sqrt{2\pi}} \cdot e^{-\frac{|x|^2}{2\sigma^2}}$$

$$= g(|x| \mid \sigma^2) \cdot h(x)^{=1}$$, then Yes !