# Tutorial 3

1. Let $X_1, \ldots, X_n$ be a random sample from a Poiss($\theta$) population where $\theta > 0$.

   (a) Show directly that $\sum_{j=1}^{n} X_j$ is sufficient for $\theta$.

   (b) Establish the same result using the Factorisation Theorem.

2. Suppose that $X_1, \ldots, X_n$ is a random sample from a population with the following density:

$$p(x, \theta) = \begin{cases} \theta a x^{a-1} \exp\{-\theta x^a\} & x > 0, \quad \theta > 0, \quad a > 0 \\ 0 & \text{otherwise} \end{cases}$$

   where $a$ is fixed. This is known as the *Weibull* density. Find a real-valued sufficient statistic for $\theta$.

3. Let $X$ be a random variable with state space $\mathcal{X} = \{v_1, \ldots, v_k\}$ and probability distribution $\mathbb{P}_\theta(X = v_i) = \theta_i$ for $i \in \{1, \ldots, k\}$ (so that $\sum_{i=1}^{k} \theta_i = 1$) and suppose that $\theta_i \in (0, 1)$ for each $i = 1, \ldots, k$. Let $X_1, \ldots, X_n$ be a random sample from $X$. Let

$$N_j = \sum_{i=1}^{n} \mathbf{1}_{\{v_j\}}(X_i).$$

   (the number of trials such that $X_i = v_j$).

   (a) What is the distribution of $(N_1, \ldots, N_k)$?

   (b) Show that $N = (N_1, \ldots, N_k)$ is sufficient for $\theta = (\theta_1, \ldots, \theta_k)$.

4. Let $X_1, \ldots, X_n$ be a random sample from a population with density $p(x, \theta)$ where:

$$p(x, \theta) = \frac{1}{\sigma} \exp\left\{ - \left( \frac{x - \mu}{\sigma} \right) \right\} \mathbf{1}_{[\mu, +\infty)}(x).$$

   Here $\theta = (\mu, \sigma)$, $\Theta = (-\infty, +\infty) \times (0, +\infty)$.

   (a) Show that $\min(X_1, \ldots, X_n)$ is sufficient for $\mu$ when $\sigma$ is fixed.
   Note: you cannot use the factorisation theorem for this part since the support of the density depends on $\mu$.

   (b) Find a one-dimensional sufficient statistic for $\sigma$ when $\mu$ is fixed.

   (c) Find a two-dimensional sufficient statistic for $\theta = (\mu, \sigma)$.

5. Let $X_1, \ldots, X_n$ be a random sample from a distribution $F$. Treating $F$ as a parameter, show that the order statistic $(X_{(1)}, \ldots, X_{(n)})$ is sufficient for $F$.

6. Let $X_1, \ldots, X_n$ be a random sample from $f(t - \theta), \theta \in \mathbb{R}$. Show that the order statistic is *minimal sufficient* for $f$ when $f$ is the Cauchy density

$$f(t) = \frac{1}{\pi(1 + t^2)} \qquad t \in \mathbb{R}.$$

7. Let $X_1, \ldots, X_m; Y_1, \ldots, Y_n$ be independent and distributed according to $N(\mu, \sigma^2)$ and $N(\eta, \tau^2)$ respectively. Find minimal sufficient statistics in the following three cases, where $(\mu, \eta, \sigma, \tau) \in \mathbb{R} \times \mathbb{R} \times \mathbb{R}_+ \times \mathbb{R}_+$:

   (a) $\mu, \eta, \sigma, \tau$ arbitrary.

   (b) $\sigma = \tau$, $\mu, \eta, \sigma$ arbitrary.

   (c) $\mu = \eta$, $\mu, \sigma, \tau$ arbitrary.

8. Let $Y = (Y_1, \ldots, Y_n)^t$ be a multivariate Gaussian random vector with distribution

$$Y \sim N(X\beta, \sigma^2 I)$$

   where $X$ is an $n \times p$ design matrix (values are given) and $\beta = (\beta_1, \ldots, \beta_p)^t$ is a parameter vector. Compute a $p + 1$ dimensional sufficient statistic for $(\beta, \sigma^2)$.

9. Let $Y_1, \ldots, Y_n$ be independent Bernoulli trials, where

$$\mathbb{P}(Y_j = 1) = \frac{1}{1 + \exp\left\{-\sum_{k=1}^{p} X_{jk}\beta_k\right\}} \qquad j = 1, \ldots, n.$$

   Compute a $p$ dimensional sufficient statistic for $\beta$.

## Short Answers

1. (a) Let $T = \sum_{j=1}^n X_j$. Note: $T \sim \text{Poiss}(n\theta)$. For $x_1 + \ldots + x_n = t$

$$\begin{aligned} \mathbb{P}_\theta((X_1, \ldots, X_n) = (x_1, \ldots, x_n) | T = t) &= \frac{\mathbb{P}_\theta((X_1, \ldots, X_n) = (x_1, \ldots, x_n))}{\mathbb{P}(T = t)} \\ &= \frac{\theta^{\sum_{j=1}^n x_j} \prod_{j=1}^n \frac{1}{x_j!} e^{-n\theta}}{\theta^t \frac{1}{t!} e^{-n\theta}} = \frac{(\sum_j x_j)!}{\prod_{j=1}^n x_j!} \end{aligned}$$

which does not depend on $\theta$.

(b)

$$\mathbb{P}_\theta((X_1, \ldots, X_n) = (x_1, \ldots, x_n)) = \frac{\theta^{\sum_{j=1}^n x_j}}{\prod_{j=1}^n x_j!} \exp\{-n\theta\}$$

This factorises as $g(\sum_{j=1}^n x_j, \theta) h(\underline{x})$ where $g(t, \theta) = \theta^t e^{-n\theta}$ and $h(\underline{x}) = \frac{1}{\prod_{j=1}^n x_j!}$.

2.

$$f(x_1, \ldots, x_n; \theta) = \begin{cases} \theta^n a^n \left(\prod_{j=1}^n x_j\right)^{a-1} \exp\left\{-\theta \sum_{j=1}^n x_j^a\right\} & x_1 > 0, \ldots, x_n > 0 \\ 0 & \text{other} \end{cases}$$

Set $t(x_1, \ldots, x_n) = \sum_{j=1}^n x_j^a$ then

$$f(x_1, \ldots, x_n; \theta) = g(t(x_1, \ldots, x_n), \theta) h(x_1, \ldots, x_n)$$

where

$$g(t, \theta) = \theta^n e^{-\theta t}, \qquad h(x_1, \ldots, x_n) = a^n \mathbf{1}_{\{x_1 > 0, \ldots, x_n > 0\}} \left(\prod_{j=1}^n x_j\right)^{a-1}$$

Hence $t(X_1, \ldots, X_n)$ is sufficient for $\theta$.

3. (a)

$$(N_1, \ldots, N_k) \sim \text{mult}(n; \theta_1, \ldots, \theta_k).$$

(b)

$$\mathbb{P}_\theta\left((X_1, \ldots, X_n) = (v_{a_1}, \ldots, v_{a_n})\right) = \theta_1^{n_1} \ldots \theta_k^{n_k}$$

where $n_j = \sum_{i=1}^n \mathbf{1}(a_i = j)$. This is in the required form from the factorisation theorem.

4. (a) The joint density is:

$$p(x_1, \ldots, x_n, \theta) = \frac{1}{\sigma^n} \exp\left\{-\sum_{j=1}^n \frac{x_j - \mu}{\sigma}\right\} \mathbf{1}_{\{\min_j x_j \geq \mu\}}.$$

The factorisation theorem cannot be used, since the support of the density depends on $\mu$. We therefore show that the conditional density $p(x_1, \ldots, x_n | \min_j X_j = y)$ does not depend on $\mu$.

Since $X - \mu \sim \text{Exp}(\frac{1}{\sigma})$, therefore $\min_{j \in \{1,\ldots,n\}} X_j - \mu \sim \text{Exp}(\frac{n}{\sigma})$ so that the density of $Y := \min_j X_j$ is:

$$p_Y(y) = \frac{n}{\sigma} \exp\left\{-\frac{n(y-\mu)}{\sigma}\right\} \mathbf{1}_{\{y \geq \mu\}}$$

and therefore

$$p(x_1, \ldots, x_n | \min_j X_j = y) = \frac{p(x_1, \ldots, x_n)}{p_Y(y)} = \frac{1}{n\sigma^{n-1}} \exp\left\{-\frac{1}{\sigma}\sum_{j=1}^{n}(x_j - \min_i x_i)\right\}$$

which does not depend on $\mu$, hence $\min_i X_i$ is sufficient for $\mu$ when $\sigma$ is fixed.

(b) From the factorisation theorem, it follows that $\sum_{j=1}^{n} X_j$ is sufficient for $\sigma$ when $\mu$ is fixed.

(c) From the factorisation theorem, applied to the conditional density $p_\sigma(x_1, \ldots, x_n | \min_j X_j = y)$, $\sum_{j=1}^{n} X_j$ is sufficient for $\sigma$, conditioned on $\min_j X_j = y$ for any $y$.
Hence $p(x_1, \ldots, x_n | \sum_j x_j = z, \min_j = y)$ depends neither on $\mu$ nor on $\sigma$, hence from the definition of sufficiency $(\min_j X_j, \sum_j X_j)$ is sufficient for $\theta = (\mu, \sigma)$.

5. Once the order statistics $x_{(1)}, \ldots, x_{(n)}$ are given, the problem is then the random assignment (without replacement) of $x_1, \ldots, x_n$ to $x_{(1)}, \ldots, x_{(n)}$. There are $n!$ permutations, each with equal probability. Suppose that there are $m$ groups, group $j$ contains $n_j$ so that $n_1 + \ldots + n_m = n$, and the order statistics are equal within each group. Then

$$\mathbb{P}((X_1, \ldots, X_n) = (x_1, \ldots, x_n) | x_{(1)}, \ldots, x_{(n)}) = \frac{\prod_{j=1}^{m} n_j!}{n!}$$

which does not depend on $F$.

6. We use the Dynkin Lehman Scheffe lemma; a statistic $T$ is minimal sufficient if $\frac{L(\theta;x)}{L(\theta;y)}$ does not depend on $\theta$ for $T(x) = T(y)$ and does depend on $\theta$ for $T(x) \neq T(y)$.

$$L(\theta; x_1, \ldots, x_n) = \frac{1}{\pi^n} \prod_{j=1}^{n} \frac{1}{(1 + (x_j - \theta)^2)}$$

$$\frac{L(\theta; \underline{x})}{L(\theta; \underline{y})} = \prod_{j=1}^{n} \frac{(1 + (y_j - \theta)^2)}{(1 + (x_j - \theta)^2)}$$

Firstly, $L(x_1, \ldots, x_n; \theta) = L(x_{(1)}, \ldots, x_{(n)}; \theta)$ so that if $\underline{y}$ is a permutation of $\underline{x}$, then $\frac{L(\theta;\underline{x})}{L(\theta;\underline{y})} = 1$. To see that this is minimal, the function does not depend on $\theta$ only if the roots of the numerators and denominators are the same (considering as functions of $\theta$), These are: $\theta = y_j \pm i$ for $j = 1, \ldots, n$ (for the numerator) and $\theta = x_j \pm i$ for $j = 1, \ldots, n$ for the denominator (where $i = \sqrt{-1}$). These are the same if and only if $(y_{(1)}, \ldots, y_{(n)}) = (x_{(1)}, \ldots, x_{(n)})$.

7. We use the Dynkin Lehman Scheffe lemma; a statistic $T$ is minimal sufficient if $\frac{L(\theta;x)}{L(\theta;y)}$ does not depend on $\theta$ for $T(x) = T(y)$ and does depend on $\theta$ for $T(x) \neq T(y)$. This is equivalent to these properties holding for the log likelihood; $\log L(\theta; x) - \log L(\theta; y)$.

   The log likelihood function is:

$$\log L(\mu, \eta; \sigma, \tau; \underline{x}, \underline{y}) = -\frac{(n+m)}{2}\log(2\pi) - m\log\sigma - n\log\tau$$
$$-\frac{1}{2\sigma^2}\left(\sum_{j=1}^m x_j^2 - \mu\sum_{j=1}^m x_j + m\mu^2\right) - \frac{1}{2\tau^2}\left(\sum_{j=1}^n y_j^2 - \eta\sum_{j=1}^n y_j + n\eta^2\right).$$

   Write out

$$\begin{aligned}\log L(\theta; \underline{x}_1, \underline{y}_1) - \log L(\theta; \underline{x}_2, \underline{y}_2) &= -\frac{1}{2\sigma^2}\left(\sum_{j=1}^m x_{1j}^2 - \sum_{j=1}^m x_{2j}^2\right) + \frac{\mu}{2\sigma^2}\left(\sum_{j=1}^m x_{1j} - \sum_{j=1}^m x_{2j}\right)\\ &\quad -\frac{1}{2\tau^2}\left(\sum_{j=1}^n y_{1j}^2 - \sum_{j=1}^n y_{2j}^2\right) + \frac{\eta}{2\tau^2}\left(\sum_{j=1}^n y_{1j} - \sum_{j=1}^n y_{2j}\right)\end{aligned}$$

   and obtain:

   (a) $\sum_{j=1}^m X_j$, $\sum_{j=1}^m X_j^2$, $\sum_{j=1}^n Y_j$, $\sum_{j=1}^n Y_j^2$.
   (b) $\sum_{j=1}^m X_j$, $\sum_{j=1}^n Y_j$, $\sum_{j=1}^m X_j^2 + \sum_{j=1}^n Y_j^2$.
   (c) $\sum_{j=1}^m X_j$, $\sum_{j=1}^n Y_j$, $\sum_{j=1}^n X_j^2$, $\sum_{j=1}^n Y_j^2$. (same as part (a)).

8. Density is:

$$f(y_1, \ldots, y_n) = \frac{1}{(2\pi)^{n/2}\sigma^n}\exp\left\{-\frac{1}{2\sigma^2}(y^t y - 2y^t X\beta + \beta^t X^t X\beta)\right\}$$

   so, by the factorisation theorem, a $p+1$ dimensional sufficient statistic is $y^t X, y^t y$.

9. For an outcome $(Y_1, \ldots, Y_n) = (y_1, \ldots, y_n)$ where $(y_1, \ldots, y_n)$ is a vector of 0's and 1's, we have

$$\mathbb{P}_{Y_1,\ldots,Y_n}(y_1, \ldots, y_n) = \frac{\prod_{j=1}^n \exp\{(1-y_j)\sum_{k=1}^p X_{jk}\beta_k\}}{\prod_{j=1}^n(1+\exp\{-\sum_{k=1}^p X_{jk}\beta_k\})}$$

   The sufficient statistic is therefore $(\sum_{j=1}^n y_j X_{jk} : k = 1, \ldots, p)$.