

MATH 557 - ASSIGNMENT 1 SOLUTIONS

1 Let $\mathcal{N}_\theta = \{0, 1, \dots, \theta\}$. For two vectors $\underline{x}, \underline{y} \in \mathbb{R}^n$, and any fixed $\theta \in \mathbb{N}$,

$$\frac{f_{\underline{X}|\theta}(\underline{x}|\theta)}{f_{\underline{Y}|\theta}(\underline{y}|\theta)} = \frac{\prod_{i=1}^n \mathbb{1}_{\mathcal{N}_\theta}(x_i)}{\prod_{i=1}^n \mathbb{1}_{\mathcal{N}_\theta}(y_i)} K(\underline{x}, \underline{y}) \theta^{(s_x - s_y)} \quad (1)$$

where

$$s_x = \sum_{i=1}^n x_i \quad s_y = \sum_{i=1}^n y_i \quad K(\underline{x}, \underline{y}) = \frac{\prod_{i=1}^n y_i!}{\prod_{i=1}^n x_i!}$$

and $\mathbb{1}_A(x)$ is the indicator function for set A . The ratio in (1) does not depend on θ (for **any** value of θ) if and only if $s_x = s_y$ and the ratio of indicator functions equals 1, which occurs if

$$X_{(n)} = \max\{x_1, \dots, x_n\} \leq \theta \quad \text{and} \quad Y_{(n)} = \max\{y_1, \dots, y_n\} \leq \theta$$

Thus, by the Theorem from lectures, $\underline{T}(\underline{X}) = \left(\sum_{i=1}^n X_i, X_{(n)} \right)$ is a minimal sufficient statistic.

5 MARKS

2 Let $Z = 0$ correspond to Tails, $Z = 1$ correspond to Heads. At the end of the study, we have **observed** the values of random variables, M, N, Z and $X_1, \dots, X_N$, that is, we have observed $M = m, N = t, Z = z$, and $X_i = x_i$ for $i = 1, \dots, t$.

(a) The full joint distribution can be factorized

$$f_Z(z) f_{N|Z, \theta}(t|z, \theta) f_{M|Z, N, \theta}(m|z, t, \theta) f_{X_1, \dots, X_t|N, M, Z, \theta}(x_1, \dots, x_t|t, m, z, \theta)$$

From the result in lectures

$$f_{X_1, \dots, X_t|N, M, Z, \theta}(x_1, \dots, x_t|t, m, z, \theta) \equiv f_{X_1, \dots, X_t|N, M}(x_1, \dots, x_t|t, m) = \binom{t}{m}^{-1}$$

as, given $N = t$ and $M = m$, the conditional distribution of $X_1, \dots, X_t$ is uniform on the set of binary sequences of length t that contains m 1s. Thus the joint distribution can be written

$$g(t, m, z|\theta) h(\underline{x})$$

where

$$g(t, m|\theta) = f_Z(z) f_{N|Z, \theta}(t|z, \theta) f_{M|Z, N, \theta}(m|z, t, \theta)$$

$$h(\underline{x}) = \binom{t}{m}^{-1}$$

Hence by the Neyman factorization theorem, (N, M, Z) is sufficient.

However, note further that, by Bayes theorem

$$f_{Z|N,M,\theta}(z|t, m, \theta) = \frac{f_{N,M|Z,\theta}(t, m|z, \theta) f_Z(z)}{f_{N,M|\theta}(t, m|\theta)}$$

Evaluating at $z = 1$, and computing the denominator using the Theorem of Total Probability conditioning in turn on $Z = 0$ and $Z = 1$ gives the following three cases:

- if $N = n$ and $M = k$

$$f_{Z|N,M,\theta}(1|n, m, \theta) = \frac{\frac{1}{2} \binom{n}{k} \theta^k (1 - \theta)^{n-k}}{\frac{1}{2} \binom{n}{k} \theta^k (1 - \theta)^{n-k} + \frac{1}{2} \binom{n-1}{k-1} \theta^n (1 - \theta)^{n-k}} = \frac{\binom{n}{k}}{\binom{n}{k} + \binom{n-1}{k-1}}$$

- if $N \neq n$ and $M = k$, then $f_{Z|N,M,\theta}(1|n, m, \theta) = 0$.
- if $N = n$ and $M \neq k$, then $f_{Z|N,M,\theta}(1|n, m, \theta) = 1$.

In all three cases, this distribution does not depend on θ (this answers part (b)). Therefore

$$f_{Z|N,M,\theta}(z|t, m, \theta) \equiv f_{Z|N,M}(z|t, m)$$

and amending the previous argument the joint distribution can be rewritten

$$f_{N,M|\theta}(t, m|\theta) f_{Z|N,M,\theta}(z|t, m, \theta) f_{X_1, \dots, X_t|N,M,Z,\theta}(x_1, \dots, x_t|t, m, z, \theta) = g(t, m|\theta) h(z, \underline{x})$$

where

$$g(t, m|\theta) = f_{N,M|\theta}(t, m|\theta) \quad h(z, \underline{x}) = f_{Z|N,M}(z|t, m) \binom{t}{m}^{-1}$$

and thus by the Neyman factorization theorem, (N, M) is a sufficient statistic. In fact, (N, M) is minimal sufficient.

From first principles:

By independence, we have the full joint distribution as

$$f_Z(z) f_{N|Z,\theta}(t|z, \theta) f_{M|Z,N,\theta}(m|z, t, \theta) f_{X_1, \dots, X_t|N,M,Z,\theta}(x_1, \dots, x_t|t, m, z, \theta)$$

in which $f_Z(z) = 1/2$, $z = 0, 1$, zero otherwise. Furthermore

$$f_{N|Z,\theta}(t|z, \theta) = \{\mathbb{1}_{\{n\}}(t)\}^z \left\{ \binom{k-1}{t-1} \theta^k (1 - \theta)^{n-k} \right\}^{1-z}$$

(as, given $z = 0$, the distribution of N given Z is negative binomial, whereas given $z = 1$, the distribution of N is degenerate at n). Secondly

$$f_{M|Z,N,\theta}(m|z, t, \theta) = \left\{ \binom{t}{m} \theta^m (1 - \theta)^{t-m} \right\}^z \{\mathbb{1}_{\{k\}}(m)\}^{1-z}$$

Finally,

$$f_{X_1, \dots, X_t|N,M,Z,\theta}(x_1, \dots, x_t|t, m, z, \theta) = \binom{t}{m}^{-1}$$

as conditional on $N = t$ and $M = m$, by the sufficiency result from lectures, the conditional distribution of $X_1, \dots, X_t$ is uniform on the set of binary sequences of length t that contains m 1s.

Thus the full joint distribution is

$$\frac{1}{2} \{\mathbb{1}_{\{n\}}(t)\}^z \left\{ \binom{k-1}{t-1} \theta^k (1-\theta)^{n-k} \right\}^{1-z} \left\{ \binom{t}{m} \theta^m (1-\theta)^{t-m} \right\}^z \{\mathbb{1}_{\{k\}}(m)\}^{1-z} \binom{t}{m}^{-1}$$

which can be re-written

$$C_1(k, n, t, m, z) \times \{\mathbb{1}_{\{n\}}(t)\}^z \{\mathbb{1}_{\{k\}}(m)\}^{1-z} \theta^{k(1-z)+mz} (1-\theta)^{(1-z)(n-k)+z(t-m)} \quad (2)$$

where

$$C_1(k, n, t, m, z) = \frac{1}{2} \binom{k-1}{t-1}^{1-z} \binom{t}{m}^z \binom{t}{m}^{-1}.$$

Now, if $z = 0$, the second term in (2) equals

$$\{\mathbb{1}_{\{k\}}(m)\} \theta^k (1-\theta)^{n-k}$$

which equals $\theta^m (1-\theta)^{n-m}$ if $m = k$, and zero otherwise. If $z = 1$, the second term in (2) equals

$$\{\mathbb{1}_{\{n\}}(t)\} \theta^k (1-\theta)^{t-m}$$

which equals $\theta^m (1-\theta)^{n-m}$ if $t = n$, and zero otherwise. Therefore, the second term equals

$$\theta^m (1-\theta)^{n-m}$$

if $(z = 0, m = k)$ or $(z = 1, t = n)$, and zero otherwise, and we can rewrite the full joint distribution as

$$C_1(k, n, t, m, z) \theta^m (1-\theta)^{n-m}$$

if $(z = 0, m = k)$ or $(z = 1, t = n)$, and zero otherwise. Hence, by the Neyman factorization theorem, the random variables (N, M) jointly form a sufficient statistic for θ .

4 MARKS

- (b) The conditional distribution of Z given N and M can be obtained by dividing the joint distribution of (Z, N, M) by the marginal for (N, M) . The marginal of N, M is obtained from above as

$$\frac{1}{2} \mathbb{1}_{\{n\}}(t) \binom{t}{m} \theta^m (1-\theta)^{t-m} + \frac{1}{2} \mathbb{1}_{\{k\}}(m) \binom{k-1}{t-1} \theta^k (1-\theta)^{n-k} \quad (3)$$

which is zero if both $t \neq n$ and $m \neq k$. If this is **not** the case, then the term in (3) is proportional to $\theta^m (1-\theta)^{n-m}$, with constant of proportionality

$$C_2(k, n, t, m) = \begin{cases} \frac{1}{2} \binom{n}{m} & t = n, m \neq k \\ \frac{1}{2} \binom{k-1}{t-1} & t \neq n, m = k \\ \frac{1}{2} \binom{n}{k} + \frac{1}{2} \binom{k-1}{n-1} & t = n, m = k \end{cases}$$

Therefore, the conditional distribution of Z given $N = t$ and $M = m$ is

$$\frac{C_1(k, n, t, m, z)}{C_2(k, n, t, m)}$$

with the denominator non-zero, which does not depend on θ .

4 MARKS

3 (a) The joint pmf for data $y_1, \dots, y_n$ is

$$\begin{aligned}
 f_{Y|\underline{\beta}}(\underline{y}|\underline{\beta}) &= \prod_{i=1}^n \binom{m_i}{y_i} \psi_i(\underline{\beta})^{y_i} (1 - \psi_i(\underline{\beta}))^{m_i - y_i} \\
 &= \prod_{i=1}^n \binom{m_i}{y_i} \frac{\exp\{y_i(\beta_0 + \beta_1 x_i)\}}{(1 + \exp\{\beta_0 + \beta_1 x_i\})^{m_i}} \\
 &= \frac{\prod_{i=1}^n \binom{m_i}{y_i}}{\prod_{i=1}^n (1 + \exp\{\beta_0 + \beta_1 x_i\})^{m_i}} \exp\left\{\beta_0 \sum_{i=1}^n y_i + \beta_1 \sum_{i=1}^n x_i y_i\right\}
 \end{aligned}$$

Therefore, for two data vectors $\underline{y}, \underline{y}^*$

$$\frac{f_{Y|\underline{\beta}}(\underline{y}|\underline{\beta})}{f_{Y|\underline{\beta}}(\underline{y}^*|\underline{\beta})} = \frac{\prod_{i=1}^n \binom{m_i}{y_i}}{\prod_{i=1}^n \binom{m_i}{y_i^*}} \exp\left\{\beta_0 \left(\sum_{i=1}^n y_i - \sum_{i=1}^n y_i^*\right) + \beta_1 \left(\sum_{i=1}^n x_i y_i - \sum_{i=1}^n x_i y_i^*\right)\right\}$$

This ratio does not depend on $\underline{\beta}$ if and only if

$$\sum_{i=1}^n y_i = \sum_{i=1}^n y_i^* \quad \text{and} \quad \sum_{i=1}^n x_i y_i = \sum_{i=1}^n x_i y_i^*$$

Therefore

$$\underline{T}(\underline{Y}) = \left(\sum_{i=1}^n Y_i, \sum_{i=1}^n x_i Y_i\right)$$

is a minimal sufficient statistic.

5 MARKS

(b) From the formula, denoting by $\psi(x)$ the probability ψ as an explicit function of x

$$\psi(0) = \frac{\exp\{\beta_0\}}{1 + \exp\{\beta_0\}} \quad \psi(1) = \frac{\exp\{\beta_0 + \beta_1\}}{1 + \exp\{\beta_0 + \beta_1\}}$$

so therefore

$$\frac{\psi(0)}{1 - \psi(0)} = \exp\{\beta_0\} \quad \frac{\psi(1)}{1 - \psi(1)} = \exp\{\beta_0 + \beta_1\}$$

and

$$\frac{\psi(1)/(1 - \psi(1))}{\psi(0)/(1 - \psi(0))} = \exp\{\beta_1\} \tag{4}$$

Recalling that $\psi_i(x)$ is the success probability for the i experiment, for $x = 0, 1$, the quantity

$$\frac{\psi(x)}{1 - \psi(x)}$$

is the **odds on success** for that experiment. Thus, from (4), we see that β_1 measures the **odds ratio** for success on the log scale when $x = 1$ and $x = 0$ for every experiment.

2 MARKS