MATH 557 - ASSIGNMENT 1
SOLUTIONS

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

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

where

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

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

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

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

4 MARKS

2. The joint pdf of the rvs is, for $\lambda > 0$,

$$f_{\mathbf{X}}(\mathbf{x};\lambda) = \frac{\prod_{i=1}^{n} \mathbb{1}_{(-\lambda,\lambda)}(x_i)}{2^n \lambda^n}$$

which indicates that a (minimal) sufficient statistic can be obtained by inspecting the numerator

$$\prod_{i=1}^{n} \mathbb{1}_{(-\lambda,\lambda)}(x_i) = \begin{cases} 1 & x_{(1)} > -\lambda \text{ and } x_{(n)} < \lambda \\ 0 & \text{otherwise} \end{cases}$$

where $x_{(1)}$ and $x_{(n)}$ are the observed minimum and maximum order statistic values. The numerator is therefore 1 if and only if

$$T = \max\{|X_{(1)}|, |X_{(n)}|\} \equiv \max_i |X_i| < \lambda$$

and so T is a sufficient statistic; furthermore, it is minimal sufficient by the minimal sufficiency theorem, as

$$\frac{f_{\mathbf{X}}(\mathbf{x};\lambda)}{f_{\mathbf{X}}(\mathbf{y};\lambda)} = \frac{\prod_{i=1}^{n} \mathbb{1}_{(-\lambda,\lambda)}(x_i)}{\prod_{i=1}^{n} \mathbb{1}_{(-\lambda,\lambda)}(y_i)} = \frac{\mathbb{1}_{(0,\lambda)}(T(\mathbf{x}))}{\mathbb{1}_{(0,\lambda)}(T(\mathbf{y}))}$$

is independent of λ if and only if $T(\mathbf{x}) = T(\mathbf{y})$.

4 MARKS

3. 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 $X_1, \ldots, X_N$, that is, we have observed $M = m$, $N = t$, $Z = z$, and $X_i = x_i$ for $i = 1, \ldots, t$.

   (a) The full joint distribution can be factorized

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

   Now

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

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

$$g(t,m,z;\theta)h(\mathbf{x})$$

   where

$$\begin{aligned} g(t,m|\theta) &= f_Z(z) f_{N|Z,\theta}(t|z,\theta) f_{M|Z,N,\theta}(m|z,t,\theta) \\ h(\mathbf{x}) &= \binom{t}{m}^{-1} \end{aligned}$$

   Hence by the factorization theorem, (N, M, Z) is sufficient. However, note further that, by Bayes theorem

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

   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 $M = k$

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

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

   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

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

   where

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

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

Alternatively, from first principles: By independence, we have the full joint distribution as

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

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

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

(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

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

Finally,

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

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

Thus the full joint distribution is

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

which can be re-written

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

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

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

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 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

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

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

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

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

4. The joint pmf for data $y_1, \ldots, y_n$ is

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

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

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

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

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

Therefore

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

is a minimal sufficient statistic.

4 MARKS