
ENEE 621
SPRING 2016
DETECTION AND ESTIMATION THEORY

ANSWER KEY TO FINAL EXAM:

---

1. Consider a Borel mapping $\Psi : \mathbb{R} \to \mathbb{R}$ such that

$$\mathbb{E}_\theta \left[|\Psi(Y)| \right] < \infty, \quad \theta \in (0,1).$$

This equivalent to the absolute summability conditions

$$\theta |\Psi(-1)| + \sum_{y=0}^{\infty} (1-\theta)^2 \theta^y |\Psi(y)| < \infty, \quad \theta \in (0,1).$$

Assume now that

$$\mathbb{E}_\theta \left[\Psi(Y) \right] = 0, \quad \theta \in (0,1).$$

In other words, assume that

$$\theta \Psi(-1) + \sum_{y=0}^{\infty} (1-\theta)^2 \theta^y \Psi(y) = 0, \quad \theta \in (0,1).$$

Elementary calculations show that

$$\begin{aligned}
&\theta \Psi(-1) + \sum_{y=0}^{\infty} (1-\theta)^2 \theta^y \Psi(y) \\
&= \theta \Psi(-1) + \sum_{y=0}^{\infty} (1 - 2\theta + \theta^2) \theta^y \Psi(y) \qquad (1.1) \\
&= \theta \Psi(-1) + \sum_{y=0}^{\infty} \theta^y \Psi(y) - 2 \sum_{y=0}^{\infty} \theta^{y+1} \Psi(y) + \sum_{y=0}^{\infty} \theta^{y+2} \Psi(y) \\
&= \theta \Psi(-1) + \Psi(0) + \theta \Psi(1) - 2\theta \Psi(0) \\
&\quad + \sum_{y=2}^{\infty} \theta^y \Psi(y) - 2 \sum_{y=1}^{\infty} \theta^{y+1} \Psi(y) + \sum_{y=0}^{\infty} \theta^{y+2} \Psi(y) \\
&= \Psi(0) + \theta \left(\Psi(-1) - 2\Psi(0) + \Psi(1) \right) \\
&\quad + \sum_{y=2}^{\infty} \theta^y \left(\Psi(y) - 2\Psi(y-1) + \Psi(y-2) \right)
\end{aligned}$$

These manipulations are permitted because the absolute summability of the infinite series allows them to be handled as finite sums would.

As we impose the conditions

$$\theta\Psi(-1) + \sum_{y=0}^{\infty}(1-\theta)^2\theta^y\Psi(y) = 0, \quad \theta \in (0,1)$$

we conclude that $\Psi(0) = 0$, $\Psi(-1) - 2\Psi(0) + \Psi(1) = 0$ and

$$\Psi(y) - 2\Psi(y-1) + \Psi(y-2) = 0, \quad y = 2, 3, \ldots$$

by standard analyticity arguments for power series. It follows that $\Psi(-1) + \Psi(1) = 0$, and setting $\Psi(1) \equiv F$, we conclude $\Psi(-1) = -F$. It is easy to see by induction that $\Psi(y) = yF$ for each $y = 2, 3, \ldots$, hence for all $y = -1, 0, 1, \ldots$. In particular, for each θ in $(0,1)$, we find that

$$\mathbb{P}_\theta\left[\Psi(Y) = 0\right] = \mathbb{P}_\theta\left[YF = 0\right] = \begin{cases} 1 & \text{if } F = 0 \\ \\ \mathbb{P}_\theta\left[Y = 0\right] = (1-\theta)^2 & \text{if } F \neq 0 \end{cases}$$

and the family $\{F_\theta, \; 0 < \theta < 1\}$ is *not* a complete family.

2.

2.a. Here, $\theta = \sigma^2$, $\Theta = \mathbb{R}_+$ and F_θ is a probability distribution on $\mathbb{R}^k$ with probability density function given by

$$f_\theta(\boldsymbol{y}) = \left(\frac{1}{\sqrt{2\pi\theta}}\right)^k e^{-\frac{1}{2\theta}\sum_{i=1}^k (y_i - \mu a_i)^2}, \quad \boldsymbol{y} = (y_1, \ldots, y_k) \in \mathbb{R}^k.$$

2.b. Writing

$$f_\theta(\boldsymbol{y}) = \left(\frac{1}{\sqrt{2\pi\theta}}\right)^k e^{-\frac{\mu^2}{2\theta}\sum_{i=1}^k a_i^2} \cdot e^{\frac{\mu}{\theta}\sum_{i=1}^k a_i y_i - \frac{1}{2\theta}\sum_{i=1}^k y_i^2}, \quad \boldsymbol{y} \in \mathbb{R}^k \tag{1.2}$$

and the Factorization Theorem immediately implies that the family $\{F_\theta, \; \theta > 0\}$ is an exponential family.

2.c. We have

$$\mathbb{E}_\theta\left[g(\boldsymbol{Y})\right] = \frac{1}{k-1}\mathbb{E}_\theta\left[K_{\sigma^2}(\boldsymbol{Y}) - \frac{K_\mu(\boldsymbol{Y})^2}{\sum_{i=1}^k a_i^2}\right] \tag{1.3}$$

with

$$\mathbb{E}_\theta\left[K_{\sigma^2}(\boldsymbol{Y})\right] = k\sigma^2 + \mu^2\left(\sum_{i=1}^k a_i^2\right) \tag{1.4}$$

and

$$\mathbb{E}_\theta \left[K_\mu(\boldsymbol{Y})^2 \right] = \left(\sum_{i=1}^k a_i^2 \right) \left(\sigma^2 + \mu^2 \left(\sum_{i=1}^k a_i^2 \right) \right) \tag{1.5}$$

by the calculations carried out in the Lecture Notes. The conclusion $\mathbb{E}_\theta [g(\boldsymbol{Y})] = \theta$ is now immediate and the estimator $g : \mathbb{R}^k \to \mathbb{R}$ is indeed an unbiased estimator of θ on the basis of $\boldsymbol{Y}$.

It is not an MVU estimator because the statistic $K : \mathbb{R}k \to \mathbb{R}^2$ is a not a complete sufficient statistic. This can be *intuited* from the fact that here the subset (of $\mathbb{R}^2$)

$$Q(\Theta) = \left\{ \left[\begin{array}{c} \frac{\mu}{\theta} \\ -\frac{1}{2\theta} \end{array} , \right], \ \theta > 0 \right\}$$

is a half-line originating from the origin and therefore does not contain a two-dimensional rectangle!

However the result used to build this intuition is only a sufficient condition! An ironclad argument is as follows: Fix $\theta > 0$ and observe from the calculations above that

$$\mathbb{E}_\theta \left[\frac{1}{k} \left(K_{\sigma^2}(\boldsymbol{Y}) - \mu^2 \left(\sum_{i=1}^k a_i^2 \right) \right) \right] = \sigma^2 \tag{1.6}$$

and

$$\mathbb{E}_\theta \left[\frac{K_\mu(\boldsymbol{Y})^2}{\sum_{i=1}^k a_i^2} - \mu^2 \left(\sum_{i=1}^k a_i^2 \right) \right] = \sigma^2. \tag{1.7}$$

Therefore, the statistic $\psi : \mathbb{R}^k \to \mathbb{R}$ given by

$$\psi(\boldsymbol{y}) = \frac{1}{k} \left(K_{\sigma^2}(\boldsymbol{y}) - \mu^2 \left(\sum_{i=1}^k a_i^2 \right) \right) - \left(\frac{K_\mu(\boldsymbol{y})^2}{\sum_{i=1}^k a_i^2} - \mu^2 \left(\sum_{i=1}^k a_i^2 \right) \right), \quad \boldsymbol{y} \in \mathbb{R}^k$$

has the property

$$\mathbb{E}_\theta [\psi(\boldsymbol{Y})] = 0, \quad \theta > 0$$

and yet

$$\mathbb{P}_\theta [\psi(\boldsymbol{Y}) = 0] \neq 0, \quad \theta > 0.$$

Therefore, the statistic $K : \mathbb{R}k \to \mathbb{R}^2$ is a not a complete sufficient statistic.

2.d. Note that the statistic $K_{\text{other}} : \mathbb{R}^k \to \mathbb{R}$ given by

$$K_{\text{other}}(\boldsymbol{y}) \equiv \sum_{i=1}^k (y_i - \mu a_i)^2, \quad \boldsymbol{y} \in \mathbb{R}$$

is clearly sufficient for the family $\{F_\theta, \ \theta \in \Theta\}$. Indeed, we have

$$f_\theta(\boldsymbol{y}) = \left(\frac{1}{\sqrt{2\pi\theta}} \right)^k e^{-\frac{2}{\theta} K_{\text{Other}}(\boldsymbol{y})}$$

so the family $\{F_\theta,\ \theta \in \Theta\}$ is an exponential family. Noting that

$$Q(\Theta) = \left\{ -\frac{2}{\theta}, \quad \theta > 0 \right\} = (-\infty, 0)$$

we conclude that the statistic $K_{\text{other}} : \mathbb{R}^k \to \mathbb{R}$ is a complete statistic for he family $\{F_\theta,\ \theta \in \Theta\}$.

Furthermore,

$$\mathbb{E}_\theta \left[K_{\text{other}}(\boldsymbol{Y}) \right] = \sum_{i=1}^{k} \mathbb{E}_\theta \left[(Y_i - \mu a_i)^2 \right] = k\theta$$

Thus the statistic $g^\star : \mathbb{R} \to \mathbb{R}$ given by

$$g^\star(\boldsymbol{y}) \equiv \frac{1}{k} \sum_{i=1}^{k} (y_i - \mu a_i)^2, \quad \boldsymbol{y} \in \mathbb{R}$$

is an unbiased estimator of θ on the basis of the observation $\boldsymbol{Y}$. Being obtained as a function of a complete sufficient statistic, it is necessarily a MVU estimator by virtue of the Rao-Blackwell Theorem and the uniqueness lemma.

3.

3.a. Here, $\theta = (p(1), \ldots, p(A))$ so that

$$\Theta \equiv \left\{ \boldsymbol{p} = (p(1), \ldots, p(A)) \in (0,1)^A : \sum_{a=1}^{A} p(a) = 1 \right\}.$$

Furthermore, for each θ in Θ, under $\mathbb{P}_\theta$, the rv Y is a discrete rv taking values in S with

$$f_\theta(s_a) = \mathbb{P}_\theta \left[Y = s_a \right] = p(a), \quad a = 1, \ldots, A.$$

3.b. Fix θ in Θ. Obviously, the probability distribution F_θ being discrete with support S, we find

$$\mathbb{P}_\theta \left[Y = y \right] = \prod_{a=1}^{A} p(a)^{\mathbf{1}[y = s_a]} = e^{\sum_{a=1}^{A} \mathbf{1}[y = s_a] \log p(a)}, \quad y \in S$$

so that the family $\{F_\theta,\ \theta \in \Theta\}$ is indeed an exponential family with

$$q(y) = \sum_{a=1}^{A} \mathbf{1}\left[y = s_a \right] \quad \text{and} \quad K(y) = \left(\mathbf{1}\left[y = s_1 \right], \ldots, \mathbf{1}\left[y = s_A \right] \right)', \quad y \in \mathbb{R}$$

while

$$C(\theta) = 1 \quad \text{and} \quad Q(\theta) = (\log p(1), \ldots, \log p(A)), \quad \theta \in \Theta.$$

3.c. A non-trivial sufficient statistic is given by

$$K(y) = \left(\mathbf{1}\left[y = s_1 \right], \ldots, \mathbf{1}\left[y = s_A \right] \right)', \quad y \in \mathbb{R}.$$

3.d. It suffices to consider $y_1, \ldots, y_k$ in S. On that range, consider the problem

$$\text{Maximize } \prod_{i=1}^k f_\theta(y_i) \text{ subject to } \theta \text{ in } \Theta \ .$$

This is equivalent to

$$\text{Maximize } \sum_{i=1}^k \log f_\theta(y_i) \text{ subject to } \theta \text{ in } \Theta$$

with

$$\begin{aligned} \sum_{i=1}^k \log f_\theta(y_i) &= \sum_{i=1}^k \left(\sum_{a=1}^A \mathbf{1}\left[y_i = s_a\right] \log p(a) \right) \\ &= \sum_{a=1}^A \left(\sum_{i=1}^k \mathbf{1}\left[y_i = s_a\right] \right) \log p(a) \\ &= \sum_{a=1}^A \log p(a) \cdot N_k(a; y_1, \ldots, y_k) \end{aligned} \tag{1.8}$$

where we have set

$$N_k(a; y_1, \ldots, y_k) = \sum_{i=1}^k \mathbf{1}\left[y_i = s_a\right], \qquad \begin{array}{c} a = 1, \ldots, A \\ y_1, \ldots, y_k \in \mathbb{R}. \end{array}$$

This quantity counts the number of times the value s_a appears amongst the observations $y_1, \ldots, y_k$.

A standard Lagrangian argument leads to considering the problem

$$\text{Maximize} \quad \begin{array}{c} \sum_{a=1}^A \log p(a) \cdot N_k(a; y_1, \ldots, y_k) \\ -\lambda \left(\sum_{a=1}^A p(a) - 1 \right) \end{array} \quad \text{subject to } \theta \in \mathbb{R}_+^A \text{ and } \lambda > 0.$$

Its solution is easily seen to be given by

$$g_{k,\mathrm{ML}}(y_1, \ldots, y_k) = \left(\frac{N_1(a; y_1, \ldots, y_k)}{k}, \ldots, \frac{N_K(a; y_1, \ldots, y_k)}{k} \right)'$$

Note that this solution in an element of the closure $\bar{\Theta}$, an acceptable fact despite the constraint that $0 < p(a) < 1$ for each $a = 1, \ldots, A$ imposed on the problem.

3.e. The ML estimator $g_{k,\mathrm{ML}} : \mathbb{R}^k \to \mathbb{R}^A$ is unbiased. It is consistent by virtue of the Strong law of Large Numbers and displays asymptotic normality by virtue of the Central Limit Theorem.

4. ______

With $\theta > 0$, note that

$$f_\theta(y) = \begin{cases} 0 & \text{if } y \le 0 \\ \theta a^\theta \left(a + y\right)^{-\theta - 1} & \text{if } y > 0. \end{cases}$$

Also, recall that if $Y \sim F_\theta$, then

$$\log\left(\frac{a+Y}{a}\right) \sim \text{Exp}(\theta).$$

4.a. Fix $0 < p \leq 1$. Recall that

$$V(p) = J_p(d_{\eta(p)}) \quad \text{with} \quad \eta = \frac{\Gamma_0(1-p)}{\Gamma_1 p}$$

where for each $\eta > 0$, the test $d_\eta : \mathbb{R} \to \{0,1\}$ is given by

$$d_\eta(y) = 0 \quad \text{iff} \quad f_{\theta_1}(y) < \eta f_{\theta_0}(y), \quad y > 0.$$

Simple calculations show that

$$\begin{aligned} d_\eta(y) = 0 \qquad & \text{iff} \qquad \left(\frac{a+y}{a}\right)^{\theta_0-\theta_1} < \eta\frac{\theta_0}{\theta_1} \\ & \text{iff} \qquad \log\left(\frac{a+y}{a}\right) > T(\eta;\theta_0,\theta_1), \quad y > 0 \end{aligned} \tag{1.9}$$

where

$$T(\eta;\theta_0,\theta_1), \equiv -\frac{1}{\theta_1-\theta_0}\log\left(\eta\cdot\frac{\theta_0}{\theta_1}\right) = \frac{1}{\theta_1-\theta_0}\log\left(\eta^{-1}\cdot\frac{\theta_1}{\theta_0}\right).$$

Thus,

$$\begin{aligned} P_{\text{F}}(d_\eta) &= \mathbb{P}_{\theta_0}\left[d_\eta(Y) = 1\right] \\ &= \mathbb{P}_{\theta_0}\left[\log\left(\frac{a+Y}{a}\right) \leq T(\eta;\theta_0,\theta_1)\right] \\ &= 1 - e^{-\theta_0 T(\eta;\theta_0,\theta_1)^+} \end{aligned} \tag{1.10}$$

and

$$\begin{aligned} P_{\text{D}}(d_\eta) &= \mathbb{P}_{\theta_1}\left[d_\eta(Y) = 1\right] \\ &= \mathbb{P}_{\theta_1}\left[\log\left(\frac{a+Y}{a}\right) \leq T(\eta;\theta_0,\theta_1)\right] \\ &= 1 - e^{-\theta_1 T(\eta;\theta_0,\theta_1)^+}. \end{aligned} \tag{1.11}$$

Therefore, as discussed in the Lecture Notes, for each p in $(0,1]$ we have

$$J_p(d_\eta) = pC(1,1) + (1-p)C(0,0) + \widehat{J}_p(d_\eta)$$

with

$$\begin{aligned} \widehat{J}_p(d_\eta) &= \Gamma_0(1-p)\cdot P_{\text{F}}(d_\eta) + \Gamma_1 p\cdot(1 - P_{\text{D}}(d_\eta)) \\ &= \Gamma_0(1-p)\cdot\left(1 - e^{-\theta_0 T(\eta;\theta_0,\theta_1)^+}\right) + \Gamma_1 p\cdot e^{-\theta_1 T(\eta;\theta_0,\theta_1)^+}. \end{aligned} \tag{1.12}$$

4.b. By direct inspection it is easy to check that

$$\{T(\eta;\theta_0,\theta_1), \quad \eta > 0\} = [0,\infty)$$

whencce $\{P_{\rm F}(d_\eta),\ \eta > 0\} = [0,1)$ and $\{P_{\rm D}(d_\eta),\ \eta > 0\} = [0,1)$ with $\lim_{\eta\downarrow 0} P_{\rm F}(d_\eta) = P_{\rm F}(d_0) = 1$ and $\lim_{\eta\downarrow 0} P_{\rm D}(d_\eta) = P_{\rm D}(d_0) = 1$. Obviously the ROC curve is defined on the interval $p_{\rm F}$ interval $[0,1]$.

5.

5.a. Under the foregoing assumptions, we note that

$$d_\eta(y) = 0 \quad \text{iff} \quad \frac{f_b(y)}{f_a(y)} < \eta.$$

But, by continuity and strict monotonicity we readily conclude

$$\left\{ y \in \mathbb{R} : \frac{f_b(y)}{f_a(y)} < \eta \right\} = (-\infty, t(\eta))$$

where

$$t(\eta) = \sup\left\{ y \in \mathbb{R} : \frac{f_b(y)}{f_a(y)} < \eta \right\} = \left(\frac{f_b(\cdot)}{f_a(\cdot)}\right)^{-1}(\eta).$$

Obviously, the mapping $y \to \frac{f_b(y)}{f_a(y)}$ is a bijection from $\mathbb{R}$ to $\mathbb{R}_+$ under the assumptions made here – There is a one-to-one correspondence between η and $t(\eta)$ with

$$\frac{f_b(t(\eta))}{f_a(t(\eta))} = \eta.$$

5.b. We seek $\eta > 0$ such that $\mathbb{P}_a\left[d_\eta(Y) = 1\right] = \alpha$. In view of Part **a**, we get

$$\alpha = \mathbb{P}_a\left[Y \geq t(\eta)\right] = 1 - F_a(t(\eta)), \tag{1.13}$$

i.e., $F_a(t(\eta)) = 1-\alpha$, and the requisite $\eta = \eta_{a,b}(\alpha)$ is therefore given through

$$\eta_{a,b}(\alpha) = F_a^{-1}(1-\alpha).$$

It follows that

$$d_{\rm NP}(\alpha; a, b) = 0 \quad \text{iff} \quad y \in (-\infty, t(\eta_{a,b}(\alpha))) = (-\infty, F_a^{-1}(1-\alpha))$$

Note that these acceptance regions do not depend on b as soon as $a < b$.

5.c. From the discussion in Part **b**, it is immediate that there exists a UMP test $d_{\rm UMP}(\alpha; a, \Theta_b^+)$ of size α (in $(0,1)$) to test the null simple hypothesis $H_0 \equiv H_a$ against the non-null composite hypothesis $H_1 \equiv (H_c,\ c \in \Theta_b^+)$. It is given by

$$d_{\rm NP}(\alpha; a, \Theta_b^+)(y) = 0 \quad \text{iff} \quad y \in (-\infty, F_a^{-1}(1-\alpha)).$$

6.

We begin by noting that

$$\begin{aligned} f_{\vartheta|Y}(\theta|y) &= \frac{f_{Y|\vartheta}(y|\theta) f_\vartheta(\theta)}{f_Y(y)} \\ &= \mathbf{1}\left[y \ge \theta\right] \frac{e^{-(y-\theta)} f_\vartheta(\theta)}{f_Y(y)}, \quad \theta, y \in \mathbb{R} \end{aligned} \tag{1.14}$$

with

$$\begin{aligned} g_{\mathrm{MAP}}(y) &= \arg\max \left(\mathbf{1}\left[y \ge \theta\right] e^{-(y-\theta)} f_\vartheta(\theta) : \ \theta \in \mathbb{R}\right) \\ &= \arg\max \left(e^{-(y-\theta)} f_\vartheta(\theta) : \ \theta \le y\right) \\ &= \arg\max \left(e^{\theta} f_\vartheta(\theta) : \ \theta \le y\right), \quad y \in \mathbb{R}. \end{aligned} \tag{1.15}$$

6.a. Here we have

$$f_\vartheta(\theta) = \frac{1}{\pi(1+\theta^2)}, \quad \theta \in \mathbb{R}.$$

Thus, with y in $\mathbb{R}$ given, we need to solve the optimization problem

$$\text{Maximize } \tfrac{e^\theta}{1+\theta^2} \text{ subject to } \theta \le y.$$

Taking derivatives we get

$$\frac{d}{d\theta}\left(\frac{e^\theta}{1+\theta^2}\right) = e^\theta \left(\frac{1}{1+\theta^2} - \frac{2\theta}{(1+\theta^2)^2}\right) = e^\theta \frac{(1-\theta)^2}{(1+\theta^2)^2} \ge 0, \quad \theta \in \mathbb{R}.$$

In other words, the function $\theta \to \frac{e^\theta}{1+\theta^2}$ is non-decreasing on $\mathbb{R}$, whence its maximum on $(-\infty, y]$ is achieved at $\theta = y$, i.e.,

$$g_{\mathrm{MAP}}(y) = y, \quad y \in \mathbb{R}. \tag{1.16}$$

6.b. More generally, we consider an arbitrary probability density function $f_\vartheta : \mathbb{R} \to \mathbb{R}_+$ such that $f_\vartheta(\theta) > 0$ for all θ in $\mathbb{R}$. For any y in $\mathbb{R}$ given, we need to solve the optimization problem

$$\text{Maximize } e^\theta f_\vartheta(\theta) \text{ subject to } \theta \le y.$$

Taking derivatives we get

$$\frac{d}{d\theta}\left(e^\theta f_\vartheta(\theta)\right) = e^\theta \left(f_\vartheta(\theta) + \frac{d}{d\theta} f_\vartheta(\theta)\right) > 0, \quad \theta \in \mathbb{R}.$$

In other words, the function $\theta \to \frac{e^\theta}{1+\theta^2}$ is non-decreasing on $\mathbb{R}$, whence its maximum on $(-\infty, y]$ is achieved at $\theta = y$, i.e.,

$$g_{\mathrm{MAP}}(y) = y, \quad y \in \mathbb{R}. \tag{1.17}$$