Solution to BIOS 760 Final 2008

1. (a) i. By using change of variables $U = \log(1/X)$, we obtain that U is exponential with mean $(\alpha+1)^{-1}$. Thus $-E\log(X) = (\alpha+1)^{-1}$ and $\text{var}(\log X_1) = (\alpha+1)^{-2}$. Clearly, $g_0(m_0(\alpha)) = \alpha$.

   ii. Note that for any $t \geq 0$, $\int_0^1 u^t(\alpha+1)u^\alpha du = (\alpha+1)(\alpha+t+1)^{-1}$. Hence $E(X) = (\alpha+1)(\alpha+2)^{-1}$. Also $E(X^2) = (\alpha+1)(\alpha+3)^{-1}$. Simple algebra yields that $g_1(m_1(\alpha)) = \alpha$ and $\text{var}(X) = (\alpha+1)(\alpha+2)^{-2}(\alpha+3)^{-1}$.

   (b) The law of large numbers yields that both $M_{0n} \rightarrow_p m_0(\alpha)$ and $M_{1n} \rightarrow_p m_1(\alpha)$. Since u^{-1} is continuous except at $u = 0$, $g_0(M_{0n}) \rightarrow_p \alpha$ by the continuous mapping theorem for convergence in probability. A similar argument establishes $g_1(M_{1n}) \rightarrow_p \alpha$.

   (c) i. First $\sqrt{n}(M_{0n} - m_0(\alpha)) \rightarrow_d N(0, (\alpha+1)^{-2})$ by the central limit theorem and results on moments from Part 1.(a)i. Now the derivative of g_0 is $\dot{g}_0(u) = -u^{-2}$ which, evaluated at $u = (1+\alpha)^{-1}$ is $-(\alpha+1)^2$. Hence the delta method yields the desired result.

   ii. The arguments are almost the same except that M_{1n} and $m_1(\alpha)$ replace M_{0n} and m_0, respectively. The derivative of g_1 is $(1-u)^{-2}$ which, when evaluated at $u = (\alpha+1)(\alpha+2)^{-1}$, is $(\alpha+2)^2$. Thus $\sigma_1^2(\alpha) = (\alpha+2)^4\text{var}(X) = (\alpha+1)(\alpha+2)^2(\alpha+3)^{-1}$

   (d) From above,
$$r(\alpha) = \frac{(\alpha+2)^2}{(\alpha+1)(\alpha+3)}.$$

   i. Let $\nu = \alpha+1$. Since $r(\alpha) = \tilde{r}(\nu) = 1 + \nu^{-1}(\nu+2)^{-1}$ and $\nu > 0$, $r(\alpha) > 1$. Cleary, as $\nu \downarrow 0$, $\tilde{r}(\nu) \rightarrow \infty$.

   ii. The above statement means that $g_0(M_{0n})$ is always asymptotically more precise that $g_1(M_{1n})$ for estimating α and that this relative precision can be arbitrarily bad for $g_1(M_{1n})$.

2. (a) By Taylor expansion, $\sqrt{n}(T_n - \alpha) = \dot{g}_0(\tilde{m})\sqrt{n}(M_{0n} - m_0(\alpha))$ for some $\tilde{m}$ on the line segment between M_{0n} and $m_0(\alpha)$. Thus $\tilde{m} \rightarrow_p m_0(\alpha)$. Since $\dot{g}_0$ is continuous, $\dot{g}_0(\tilde{m}) \rightarrow_p -(\alpha+1)^2$. Thus by Slutsky's theorem $[(\dot{g}_0(\tilde{m}) + (\alpha+1)^2]\sqrt{n}(M_{0n} - m_0(\alpha)) = o_P(1)$. This means that
$$\sqrt{n}(T_n - \alpha) = -(\alpha+1)^2\sqrt{n}(M_{0n} - m_0(\alpha)) + o_P(1),$$
and the desired result follows.

(b) The score $\dot{\ell}_\alpha(X)$ for α is $\log X + (\alpha+1)^{-1}$ which has mean zero. Since the variance of $\log X$ is $(\alpha+1)^{-2}$ and since the information I_α is equal to $E(\dot{\ell}^2_\alpha(X)) = \text{var}(\log X)$, we have that the efficient influence function $I_\alpha^{-1}\dot{\ell}_\alpha(X)$ is precisely $H_\alpha(X)$.

(c) This is Proposition 4.7 (which does not need to be remembered). What needs to be remembered is that this is one of the main dogmas of efficiency theory.

(d) If we use $\hat{\alpha}_n = \tilde{\alpha}_n + I_{\tilde{\alpha}_n}^{-1} n^{-1}\sum_{i=1}^n \dot{\ell}_{\tilde{\alpha}_n}(X_i)$, then, for some α^* on the line segment between $\tilde{\alpha}_n$ and α,

$$\begin{aligned}\sqrt{n}(\hat{\alpha}_n - \alpha) &= \sqrt{n}(\tilde{\alpha}_n - \alpha) + I_{\tilde{\alpha}_n}^{-1} n^{-1/2} \sum_{i=1}^n \left[\dot{\ell}_\alpha(X_i) + \ddot{\ell}_{\alpha^*}(X_i)(\tilde{\alpha}_n - \alpha)\right] \\ &= I_\alpha^{-1} n^{-1/2} \sum_{i=1}^n \dot{\ell}_\alpha(X_i) + o_P(1),\end{aligned}$$

since I_α is continuous in α, and where somewhat careful analysis is needed to verify that $n^{-1}\sum_{i=1}^n \ddot{\ell}_{\alpha^*}(X_i) = -I_\alpha + o_P(1)$.

3. (a) Note that the given joint density is $k_\alpha(\delta, y) = I\{0 \le y \le 1\} f_\alpha^{1-\delta}(y)/2$. If we now sum $k_\alpha(\delta, y)$ over $\delta = 0, 1$, we obtain the desired result.

(b) Note that

$$\begin{aligned}E(\Delta|Y = y, \alpha) &= \text{pr}(\Delta = 1|Y = y, \alpha) \\ &= \frac{k_\alpha(1, y)}{\sum_{\delta=0,1} k_\alpha(\delta, y)} \\ &= (1 + f_\alpha(y))^{-1}.\end{aligned}$$

(c) When we take logarithms of the likelihood, we can ignore the indicator part since it will be true for all observations in the sample. The result now follows since $\log(k_\alpha(\delta, Y)) = -\log(2) + (1-\delta)\left(\alpha \log(Y) + \log(\alpha+1)\right)$.

(d) Let $q(\delta, y)$ be a measurable function of (δ, y). By independence,

$$E\left[q(\Delta_i, Y_i)|Y_1, \ldots, Y_n, \alpha^{(k)}\right] = E\left[q(\Delta_i, Y_i)|Y_i, \alpha^{(k)}\right],$$

for $i = 1, \ldots, n$. This gives us that

$$E\left[\ell_n(\alpha)|Y_1, \ldots, Y_n, \alpha\right] = -n\log(2) + \sum_{i=1}^n \left(1 - E(\Delta_i|Y_i, \alpha)\right)\dot{\ell}_\alpha(Y_i),$$

and the desired result now follows from Part (b) above.

(e) This follows from Part (d) and a little algebra and calculus.

4. (a) A complete statistic T for the parameter θ satisfies the condition: If, for a measurable function g, $E_\theta g(T) = 0$ for all θ, then $g = 0$.

   (b) A martingale is a sequence of random variables $Y_1, Y_2, \ldots$ associated with a sequenct of increasing σ-fields $F_1, F_2, \ldots$ such that (i) Y_j is measurable with respect to F_j and (ii) $E[Y_{j+1}|F_j] = Y_j$, $j = 1, 2, \ldots$.