Problem Set #6 (Due December 5, 2018)

1. Let $(Y_1, X_1), \ldots, (Y_n, X_n)$ be an i.i.d. sequence of random variables such that $Y_i \in \mathbf{R}$, $X_i \in \mathbf{R}^{k+1}$, the first component of X_i equals one and

$$Y_i = X_i'\theta + \epsilon_i \ ,$$

where $\epsilon_1, \ldots, \epsilon_n$ are, independently of $X_1, \ldots, X_n$, i.i.d. with distribution $N(0, \sigma^2)$. Assume that

$$\sum_{i=1}^{n} X_i X_i'$$

is invertible.

   (a) What is the (conditional) likelihood function? What is the (conditional) log-likelihood function?

   (b) Find the ML estimators of θ and σ, $\hat{\theta}_n$ and $\hat{\sigma}_n$.

   (c) Consider testing $H_0 : \theta = \theta_0$ versus $H_1 : \theta \neq \theta_0$. Describe how you would carry this out using each of the three hypothesis testing methods described in class. How do they compare with the Wald tests and LM tests described earlier in class using ordinary least squares?

2. Let $X_1, \ldots, X_n$ be an i.i.d. sequence of random variables with distribution $U(\theta, 2\theta)$ with $\theta > 0$, i.e., uniformly distributed on the interval $[\theta, 2\theta]$.

   (a) Show that the maximum likelihood estimator of θ is given by

$$\hat{\theta}_n = \frac{1}{2} \max_{1 \leq i \leq n} X_i \ .$$

   (b) Is $\hat{\theta}_n$ an unbiased estimator of θ? Justify your answer.

   (c) Prove that $\hat{\theta}_n$ is a consistent estimator of θ.

   (d) Show that $n(\theta - \hat{\theta}_n)$ converges in distribution to an exponential distribution with parameter λ as $n \to \infty$. Express λ in terms of θ. (Hint: The exponential distribution with parameter λ has c.d.f.

$$F(x) = \begin{cases} 0 & \text{if } x < 0 \\ 1 - \exp(-x/\lambda) & \text{if } x \geq 0 \end{cases} .$$

   You may also use the fact that

$$\lim_{n \to \infty} \left(1 - \frac{c}{n}\right)^n = \exp(-c) \ .)$$

(e) Construct an (approximate) 95% confidence interval for θ. Justify your answer.

3. Recall the setup of the previous question.

   (a) Suggest an unbiased estimator, $\tilde{\theta}_n$, of θ. Justify your answer.

   (b) Use simulation to compare this estimator with the ML estimator in terms of the values of $E[(\hat{\theta}_n - \theta)^2]$ and $E[(\tilde{\theta}_n - \theta)^2]$ as follows: For each $\theta \in \{0.5, 1, 10\}$ and $n \in \{2, 5, 20, 100\}$, repeat the following 10^4 times:

      i. Draw n i.i.d. observations from $U(\theta, 2\theta)$.
      ii. Compute $\hat{\theta}_n$ and $\tilde{\theta}_n$.
      iii. Compute $(\hat{\theta}_n - \theta)^2$ and $(\tilde{\theta}_n - \theta)^2$.
      iv. Compute $|\hat{\theta}_n - \theta|$ and $|\tilde{\theta}_n - \theta|$.
      v. Compute $I\{|\hat{\theta}_n - \theta| < |\tilde{\theta}_n - \theta|\}$, an indicator telling us whether $\hat{\theta}_n$ is closer to θ than $\tilde{\theta}_n$.

      Average the values obtained in parts (iii)-(v) across the 10^4 simulations to obtain estimates of $E[(\hat{\theta}_n - \theta)^2]$, $E[(\tilde{\theta}_n - \theta)^2]$, $E[|\hat{\theta}_n - \theta|]$, $E[|\tilde{\theta}_n - \theta|]$, and $P\{|\hat{\theta}_n - \theta| < |\tilde{\theta}_n - \theta|\}$. Based on these simulations, which estimator do you prefer? Is unbiasedness always a desirable property?

   (c) Can you justify your preference on theoretical grounds? (Hint: Examine the limiting distribution of $\tilde{\theta}_n$. What do you need to scale by to get a non-degenerate limiting distribution? Compare this to part (d) of the previous question.)

4. Let $(Y_1, X_1), \ldots, (Y_n, X_n)$ be an i.i.d. sequence of random variables such that $Y_i \in \{0, 1\}$, $X_i \in \mathbf{R}^{k+1}$, the first component of X_i equals one, and

$$p_\theta(y|x) = P\{Y_i = y|X_i = x\} = \begin{cases} G(x'\theta) & \text{if } y = 1 \\ 1 - G(x'\theta) & \text{if } y = 0 \end{cases},$$

   where G is a c.d.f.

   (a) Find the likelihood function $\ell_n(\theta)$ and the log-likelihood function $L_n(\theta)$.

   (b) Under what conditions is $L(\theta) = E[\log p_\theta(Y_i|X_i)]$ uniquely maximized at θ_0?

(c) The Logit model specifies that

$$G(z) = \frac{\exp(z)}{1 + \exp(z)} \ .$$

Suppose the data consists of $\{(1, 1), (0, 0.8), (1, 2), (0, 0.5)\}$. Does the ML estimator, $\hat{\theta}_n$ exist for this data? Explain briefly.

(d) Assuming appropriate regularity conditions, what is the limiting distribution of the ML estimator, $\hat{\theta}_n$?

(e) Calculate $D_x P\{Y_i = 1|X_i = x\}$. Suggest an estimator for this quantity using the ML estimator. How does this estimator relate to the ML estimator for $D_x P\{Y_i = 1|X_i = x\}$?

(f) Again assuming appropriate regularity conditions, derive the limiting distribution for your estimator in part (d).

5. Let $X_i, i = 1, \ldots, n$ be an i.i.d. sequence of random variables with p.d.f. on $\mathbf{R}$ given by

$$f_\theta(x) = \begin{cases} (1+\theta)x^\theta & \text{for } 0 < x < 1 \\ 0 & \text{otherwise} \end{cases}$$

for some $\theta > -1$.

(a) Show that

$$\mu = E[X_i] = \frac{\theta + 1}{\theta + 2} \ .$$

Express θ as a function of μ.

(b) Write the log-likelihood function as a function of μ.

(c) Derive the ML estimator $\hat{\mu}_n$ of μ. Justify your answer.

(d) Show that $\hat{\mu}_n$ is consistent for μ as $n \to \infty$. (Hint: You may need to show that

$$E[\log X_i] = -\frac{1}{1+\theta} \ .$$

This can be done using integration by parts and the fact that

$$\lim_{x \downarrow 0} x^\lambda \log x = 0$$

for $\lambda > 0$.)

(e) Derive the limiting distribution of $\sqrt{n}(\hat{\mu}_n - \mu)$ as $n \to \infty$. Do this using only results from the first part of the class. In particular, do not appeal to "high level" results about maximum likelihood presented in second part of the class. (Hint: You may need to show that

$$E[(\log X_i)^2] = \frac{2}{(1+\theta)^2} \; .$$

This can be done using integration by parts (twice!) and the fact that

$$\lim_{x \downarrow 0} x^{\lambda} (\log x)^2 = 0$$

for $\lambda > 0$.)

(f) Compute the information matrix. How does this relate to your answer in part (e)?

(g) Describe the score test for testing $H_0 : \mu = \frac{2}{3}$ versus $H_1 : \mu \neq \frac{2}{3}$ at level α.