STAT 210B HWK #1 SOLUTIONS (DUE FEB 7)

JING LEI, DAVID ROSENBERG

(1) Given densities p_n and q_n with respect to some measure μ, define the likelihood ratio $L_n(x)$ as $L_n(x) = q_n(x)/p_n(x)$ for $p_n(x) > 0$, $L_n(x) = 1$ if $p_n(x) = q_n(x) = 0$ and $L_n(x) = \infty$ otherwise. Show that the likelihood ratio is a uniformly tight sequence.

Note that $E(L_n) = \int_{x:p_n(x)>0} \frac{q_n(x)}{p_n(x)} p_n(x)dx \leq \int q_n(x)dx = 1$. Then for any $\epsilon > 0$, choose M such that $M > 1/\epsilon$, by Markov's inequality, we have $P(L_n > M) \leq E(L_n)/M < \epsilon, \quad \forall n$, i.e., $\{L_n\}$ is UT. **(1) Show that if $EX_n \to \mu$ and Var$X_n \to 0$, then $X_n \xrightarrow{P} \mu$.**

By Chebyshev's inequality, for any $\epsilon > 0$,

$$P(|X_n - E(X_n)| > \epsilon) = P(|X_n - E(X_n)|^2 > \epsilon^2) \leq \frac{\mathrm{Var}(X_n)}{\epsilon^2} \to 0,$$

so $X_n - E(X_n) \xrightarrow{P} 0$, combine this with the assumption $EX_n - \mu \to 0$, we have $X_n - \mu = (X_n - EX_n) + (EX_n - \mu) \xrightarrow{P} 0$.

(2) Let $X, X_1, \ldots, X_n$ be i.i.d. from Beta$(\theta, 1)$, where $\theta > 0$. Let $\bar{X}_n$ denote the sample mean. The method of moments estimator of θ is $\hat{\theta}_n = \bar{X}_n/(1-\bar{X}_n)$. Find its asymptotic distribution.

First note that $EX = \theta/(1+\theta)$ and $\mathrm{Var}X = \theta/(1+\theta)^2(2+\theta)$.

$$\begin{aligned}
\sqrt{n}\left(\hat{\theta}_n - \theta\right) &= \sqrt{n}\left(\frac{\bar{X}_n - \theta(1-\bar{X}_n)}{1-\bar{X}_n}\right) = \sqrt{n}\left(\frac{(1+\theta)\bar{X}_n - \theta}{1-\bar{X}_n}\right) \\
&= (1+\theta)\frac{\sqrt{n}(\bar{X}_n - EX)}{1-\bar{X}_n}
\end{aligned}$$

By the Central Limit Theorem

$$\sqrt{n}(\bar{X}_n - EX) \xrightarrow{d} \mathcal{N}(0, \mathrm{Var}X)$$

and by SLLN $1 - \bar{X}_n \xrightarrow{P} 1 - EX = 1/(1+\theta)$. So by Slutsky's Theorem,

$$\frac{\sqrt{n}(\bar{X}_n - EX)}{1-\bar{X}_n} \xrightarrow{d} (1+\theta)\mathcal{N}(0, \mathrm{Var}X) \stackrel{d}{=} \mathcal{N}(0, \theta/(2+\theta))$$

So $\sqrt{n}(\hat{\theta}_n - \theta) \xrightarrow{d} \mathcal{N}(0, (1+\theta)^2\theta/(2+\theta)$.

(3) X_n uniform on $\{1/n, 2/n, \ldots, 1\}$. Show $X_n \xrightarrow{d} X$, where $X \sim$ Uniform$(0,1)$. Does $X_n \xrightarrow{P} X$?

For $x \in [0,1]$ we have

$$P(X_n \leq x) = \frac{\lfloor xn \rfloor}{n} \to x = P(X \leq x)$$

Date: Feb 9, 2008 .

For $x < 0$ and $x > 1$, $P(X_n \leq x)$ is independent of n and the convergence to $P(X \leq x)$ is obvious. Therefore $X_n \xrightarrow{d} X$.

We can't ask about convergence in probability unless X_n and X are defined on the same probability space (or convergence is to a degenerate random variable). Even if they are defined on the same probability space, one can have a counterexample: let $Y_n = \sum_1^n \frac{k}{n} \mathbb{1}\left(\frac{k-1}{n} < X \leq \frac{k}{n}\right)$, $X_n = \frac{n+1}{n} - Y_n$, then $X_n \stackrel{d}{=} Y_n$ and $Y_n \xrightarrow{P} X$ but X_n does not converge in probability to X.

(4) Suppose that $F_n(x) \to F(x)$ for all x, F is continuous, and strictly increasing so that $F^{-1}(\alpha)$ is unique for all $0 < \alpha < 1$. Show that

$$\sup\{|F^{-1}(\alpha) - F^{-1}(\alpha)| : \epsilon \leq \alpha \leq 1-\epsilon\} \to 0$$

for all $\epsilon > 0$. Here $F_n^{-1}(\alpha) = \inf\{x : F_n(x) \geq \alpha\}$

First we show that $F_n^{-1}(\alpha_0) \to F^{-1}(\alpha_0)$. In fact for any $\beta > 0$, since $F_n(F^{-1}(\alpha_0) + \beta) \to F(F^{-1}(\alpha_0) + \beta) > \alpha_0$, so $F_n^{-1}(\alpha_0) \leq F^{-1}(\alpha_0) + \beta$ for all large enough n. Similarly we can show that $F_n^{-1}(\alpha_0) \geq F^{-1}(\alpha_0) - \beta$ for all large enough n. So we have $F_n^{-1}(\alpha_0) \to F^{-1}(\alpha_0)$.

For any $\epsilon > 0$, argue by contradiction: if $\limsup_n \sup\{|F_n^{-1}(\alpha) - F^{-1}(\alpha)| : \alpha \in [\epsilon, 1-\epsilon]\} = \delta > 0$, then there exists $\alpha_{n_k} \in [\epsilon, 1-\epsilon]$, s.t., $|F_{n_k}^{-1}(\alpha_{n_k}) - F^{-1}(\alpha_{n_k})| \geq \delta$. By compactness of $[\epsilon, 1-\epsilon]$ we can with out loss of generality assume that $\alpha_{n_k} \to \alpha_0$, and furthermore $|F_{n_k}^{-1}(\alpha_{n_k}) - F^{-1}(\alpha_0)| \geq \delta/2$.

Since for each n_k, we have either $F_{n_k}^{-1}(\alpha_{n_k}) > F^{-1}(\alpha_0) + \delta/2$ or $F_{n_k}^{-1}(\alpha_{n_k}) < F^{-1}(\alpha_0) - \delta/2$.

case 1: there exits subsequence n_{k_j} such that $F_{n_{k_j}}^{-1}(\alpha_{n_{k_j}}) > F^{-1}(\alpha_0) + \delta/2$. But

$$\begin{aligned}
F_{n_{k_j}}(F^{-1}(\alpha_0) + \delta/2) &\to F(F^{-1}(\alpha_0) + \delta/2) \\
&\geq \alpha_{n_{k_j}} \quad \text{for large enough } j \\
&\Rightarrow F_{n_{k_j}}^{-1}(\alpha_{n_{k_j}}) \leq F^{-1}(\alpha_0) + \delta/2 \\
&\Rightarrow \text{Contradiction!}
\end{aligned}$$

case 2: there exits subsequence n_{k_j} such that $F_{n_{k_j}}^{-1}(\alpha_{n_{k_j}}) < F^{-1}(\alpha_0) - \delta/2$. But

$$\begin{aligned}
F_{n_{k_j}}(F^{-1}(\alpha_0) 2\delta/2) &\to F(F^{-1}(\alpha_0) 2\delta/2) \\
&\leq \alpha_{n_{k_j}} \quad \text{for large enough } j \\
&\Rightarrow F_{n_{k_j}}^{-1}(\alpha_{n_{k_j}}) \geq F^{-1}(\alpha_0) + \delta/2 \\
&\Rightarrow \text{Contradiction!}
\end{aligned}$$

Based on the assumption, at least one of the above 2 cases will happen, so the assumption is false.

(5) Let $(X_1, Y_1), ..., (X_n, Y_n)$ be i.i.d sample of 2-D random vector (X, Y) where $0 < EX^4 < \infty$, $0 < EY^4 < \infty$. Let $\rho^2 = \mathbf{cov}^2(X, Y)/\sigma_1^2\sigma_2^2$, where $\sigma_1^2 = \mathbf{var}(X)$, $\sigma_2^2 = \mathbf{var}(Y)$; and let $r^2 = \hat{C}^2/\hat{\sigma}_X^2\hat{\sigma}_Y^2$ where

$$\hat{C} = n^{-1}\sum(X_i - \bar{X})(Y_i - \bar{Y}), \quad \hat{\sigma}_X^2 = n^{-1}\sum(X_i - \bar{X})^2, \quad \hat{\sigma}_Y^2 = n^{-1}\sum(Y_i - \bar{Y})^2$$

show that

- **(a) If $(X, Y) \sim N(\mu_1, \mu_2, \sigma_1^2, \sigma_2^2, \rho)$, then $\sqrt{n}(r^2 - \rho^2) \xrightarrow{d} N(0, 4\rho^2(1-\rho^2)^2)$ and, if $\rho \neq 0$, then $\sqrt{n}(r - \rho) \xrightarrow{d} N(0, (1-\rho^2)^2)$.**
- **(b) If $\rho = 0$, $\sqrt{n}(r - \rho) \xrightarrow{d} N(0, 1)$.**

Sketch of proof:

(1) WOLG, can assume $E(X) = E(Y) = 0$, $\mathrm{Var}(X) = \mathrm{Var}(Y) = 1$

(2) Let $\tilde{C} = n^{-1}\sum X_i Y_i$, $\tilde{\sigma}_X^2 = n^{-1}\sum X_i^2$, $\tilde{\sigma}_Y^2 = n^{-1}\sum Y_i^2$, then r^2 and $\tilde{r}^2 = \tilde{C}^2/\tilde{\sigma}_X^2\tilde{\sigma}_Y^2$ have the same asymptotic distribution.

(3) show that

$$\sqrt{n}\begin{pmatrix} \tilde{C}-\rho \\ \tilde{\sigma}_X^2 - 1 \\ \tilde{\sigma}_Y^2 - 1 \end{pmatrix} \xrightarrow{d} N\left(0, \begin{pmatrix} 1+\rho^2 & 2\rho & 2\rho \\ 2\rho & 2 & 2\rho^2 \\ 2\rho & 2\rho^2 & 2 \end{pmatrix}\right)$$

(4) consider the function

$$f(a,b,c) = a^2 b^{-1} c^{-1},$$

use Delta-method, we have

$$\sqrt{n}(\tilde{r}^2 - \rho^2) \xrightarrow{d} N(0, 4\rho^2(1-\rho^2)^2)$$

(5) consider the function

$$g(a,b,c) = ab^{-1/2}c^{-1/2},$$

use Delta-method, we have

$$\sqrt{n}(\tilde{r} - \rho) \xrightarrow{d} N(0, (1-\rho^2)^2).$$

Part (b) follows easily.

Remark 1. It seems that part(b) needs the assumption of normality (and hence independence). Other wise the asymptotic variance should be $E(X^2Y^2)$ other than 1.