Statistics 581, Problem Set 3 Solutions

Wellner; 10/21/2010

1. A sequence of random variables Y_n is *bounded in probability* and we write $Y_n = O_p(1)$ if
$$\lim_{\lambda\to\infty} \limsup_{n\to\infty} P(|Y_n| > \lambda) = 0;$$
i.e. for each $\epsilon > 0$ there exist λ_ϵ and N_ϵ such that $P(|Y_n| > \lambda_\epsilon) < \epsilon$ for all $n > N_\epsilon$.
(a) Show that if $Y_n \to_d Y$ for some random variable Y, then Y_n is bounded in probability. (This is Lehmann and Casella, problem 8.24, page 77.)
(b) Give an example of a sequence of random variables Y_n that is bounded in probability, but does not converge in distribution.
(c) Lehmann and Casella, problem 8.25, page 77.

Solution: (a) Let F denote the limiting distribution, and let $Y \sim F$. Fix $\epsilon > 0$. Choose $\pm M \in C_F$ so large that $P(|Y| > M) \le 1 - F(M) + F(-M) < \epsilon/2$. Since $\pm M \in C_F$, and $Y_n \sim F_n$ converge in distribution to $Y \sim F$, we can find an $N = N_{M,\epsilon}$ so that $|F_n(\pm M) - F(\pm M)| < \epsilon/4$. Then for $n \ge N$ it follows that

$$\begin{aligned} P(|Y_n| > M) &= 1 - F_n(M) + F_n(-M-) \\ &\le 1 - F(M) + F(-M) + F(M) - F_n(M) + F_n(-M) - F(-M) \\ &< \frac{1}{2} + \frac{\epsilon}{4} + \frac{\epsilon}{4} = \epsilon; \end{aligned}$$

i.e. Y_n is bounded in probability.
(b) Suppose that $Y_n = (-1)^n Y$ where $Y \sim$ Exponential(1). Then $|Y_n| \le |Y| = O_p(1)$, but $Y_{2n} \to Y$ while $Y_{2n+1} \to -Y$, and therefore $Y_n \not\to_d$.
(c) Problem 8.21 holds with o and O replaced by o_p and O_p:
(a') If $R_n = o_p(1/k_n)$, then $R_n/(1/k_n) = k_n R_n = o_p(1)$, so $k_n R_n = O_p(1)$, and hence $R_n = O_p(1/k_n)$.
(b') If $R_n = O_p(1)$ then R_n is bounded in probability. This is just the definition of R_n being $O_p(1)$.
(c') If $R_n = o_p(1)$, then $R_n \to_p 0$; this is just the definition of $R_n = o_p(1)$.
(d') If $R_n = O_p(1/k_n)$, and $k_n'/k_n \to \rho \in (-\infty, \infty)$, then $R_n = O_p(1/k_n')$: for every $\epsilon > 0$ there exists $M = M_\epsilon$ and $N = N_\epsilon$ such the for $n \ge N_\epsilon$ we have

$$P(|k_n R_n| > M) < \epsilon$$

and, perhaps by choosing N_ϵ somewhat larger, $|k_n'/k_n - \rho| < \epsilon$. Thus with $M' = M(|\rho| + \epsilon)$,

$$P(|k_n' R_n| > M') = P(|(k_n'/k_n) k_n R_n| > M') \le P((|\rho| + \epsilon)|k_n R_n| > M') \le P\left(|k_n R_n| > M\right) < \epsilon.$$

Problem 8.22 holds with o and O replaced by o_p and O_p:
(a') If $R_n = O_p(1/k_n)$ and $R_n' = O_p(1/k_n)$, then $R_n + R_n' = O_p(1/k_n)$: by the hypothesis, for each $\epsilon > 0$ there exist $M = M_\epsilon, M_\epsilon'$ such that for $n \ge N_\epsilon$

$$P(|k_n R_n| > M) < \epsilon \qquad \text{and} \qquad P(|k_n R_n'| > M') < \epsilon.$$

Let $\tilde{M} \equiv M_{\epsilon/2} + M'_{\epsilon/2}$. Then

$$P(|k_n(R_n + R'_n)| > \tilde{M}) \le P(|k_n R_n| > M_{\epsilon/2}) + P(|k_n R'_n| > M'_{\epsilon/2}) < \epsilon/2 + \epsilon/2 = \epsilon.$$

Thus $R_n + R'_n = O_p(1/k_n)$.
(b') If $R_n = o_p(1/k_n)$ and $R'_n = o_p(1/k_n)$, then $k_n R_n \to_p 0$ and $k_n R'_n \to_p 0$, so $k_n(R_n + R'_n) \to_p 0$; i.e. $R_n + R'_n = o_p(1/k_n)$.
Problem 8.23 holds with o and O replaced by o_p and O_p: Suppose that $k'_n/k_n \to \infty$.
(a') If $R_n = O_p(1/k_n)$ and $R'_n = O_p(1/k'_n)$, then $R_n + R'_n = O_p(1/k_n)$: The hypotheses imply that $k_n R_n = O_p(1)$ and $R'_n = O_p(1/k'_n)$. Thus

$$k_n(R_n + R'_n) = k_n R_n + \frac{k_n}{k'_n} k'_n R'_n = O_p(1) + o_p(1) O_p(1) = O_p(1);$$

i.e. $R_n + R'_n = O_p(1/k_n)$.
(b') If $R_n = o_p(1/k_n)$ and $R'_n = o_p(1/k'_n)$, then $R_n + R'_n = o_p(1/k_n)$: This follows easily since

$$k_n(R_n + R'_n) = k_n R_n + \frac{k_n}{k'_n} k'_n R'_n \to_p 0 + 0 \cdot 0 = 0.$$

2. Suppose that X is a random variable with finite fourth moment; $E|X|^4 < \infty$. Then $\mu_4 = E(X-\mu)^4$ is the fourth central moment of X. The ratio $\mu_4/\sigma^4 \equiv \kappa$ is the *kurtosis* of X (or of the distribution function F of X), and $\gamma_2 \equiv \mu_4/\sigma^4 - 3$ is called the *excess of kurtosis*; note that for any $N(\mu, \sigma^2)$ random variable, $\gamma_2 = 0$. Investigate the value of γ_2 for various classical distributions (t_r, uniform, bernoulli, Poission(λ), ...). How big can γ_2 be? How small can γ_2 be?

   Solution: Note that $\mu_4^{1/4} = \{E(X-\mu)^4\}^{1/4} \ge \{E(X-\mu)^2\}^{1/2} = \sigma$ by Liapunov's inequality. Thus $\mu_4/\sigma^4 \ge 1$ always, or $\gamma_2 \equiv \mu_4/\sigma^4 \ge -2$ with equality if $X = \pm 1$ with probability 1/2 each: then $\mu = 0$, $\sigma^2 = 1$, $\mu_4 = 1$, and $\gamma_2 = -2$.
   For $X \sim N(0,1)$, $\gamma_2 = 0$ since $EX^4 = 3$.
   For $X \sim t_r$, $r > 4$, $\gamma_2 = 6/(r-4) \nearrow \infty$ as $r \searrow 4$; $\gamma_2 \searrow 0$ as $r \nearrow \infty$.
   For $X \sim$ Gamma(α, β), $\gamma_2 = 6/\alpha \nearrow \infty$ as $\alpha \searrow 0$.
   For $X \sim$ Poisson(λ), $\gamma_2 = 1/\lambda \nearrow \infty$ as $\lambda \searrow 0$.
   For $X \sim$ Bernoulli(p), $\gamma_2 = (1-p)^2/p + p^2/(1-p) - 3$ which $= -2$ when $p = 1/2$, and $\nearrow \infty$ when $p \to 0, 1$.

3. Suppose that $X, X_1, \ldots, X_n$ are i.i.d. with mean μ, variance σ^2, and $E|X|^4 < \infty$.
   (a) Show that the sample variance $S_n^2 = \sum_{i=1}^n (X_i - \overline{X}_n)^2/(n-1)$ satisfies

$$\sqrt{n}(S_n^2 - \sigma^2)/\sqrt{2}\sigma^2 \to_d N(0, 1 + \gamma_2/2).$$

   (b) Suppose that you want to test $H : \sigma \le \sigma_0^2$ versus $K : \sigma^2 > \sigma_0^2$ for σ_0 a fixed number, and you base your test on normal theory, but in fact the X's are *not normal* with $\gamma_2 \ne 0$. What effect does this have on the level (or size or actual type one

error) of the normal theory test?
Solution: (a) Since

$$\overline{S}_n^2 \equiv \frac{1}{n}\sum_{i=1}^n (X_i - \overline{X}_n)^2 = \frac{1}{n}\sum_{i=1}^n (X_i - \mu)^2 - (\overline{X}_n - \mu)^2$$

where $Y_i \equiv (X_i - \mu)^2$ are i.i.d. $(E(Y_1), Var(Y_1)) = (\sigma^2, \mu_4 - \sigma^4)$, it follows that

$$\begin{aligned}
\sqrt{n}(\overline{S}_n^2 - \sigma^2) &= \sqrt{n}(\overline{Y}_n - \sigma^2) - \sqrt{n}(\overline{X}_n - \mu)(\overline{X}_n - \mu) \\
&\to_d N(0, \mu_4 - \sigma^4) - N(0, \sigma^2) \cdot 0 \\
&= N(0, \mu_4 - \sigma^4).
\end{aligned}$$

Hence

$$\sqrt{n}(\overline{S}_n^2 - \sigma^2)/(\sqrt{2}\sigma^2) \to_d N(0, (\mu_4/\sigma^4 - 1)/2) = N(0, (\mu_4/\sigma^4 - 3 + 2)/2) = N(0, 1 + \gamma_2/2).$$

If instead of $\overline{S}_n^2$ we consider the more usual $S_n^2 = (n/(n-1))\overline{S}_n^2$, it is easily seen that

$$\sqrt{n}(S_n^2 - \overline{S}_n^2) = \sqrt{n}(\frac{n}{n-1} - 1)\overline{S}_n^2 = o(1)O_p(1) = o_p(1).$$

Thus we also have

$$\sqrt{n}(S_n^2 - \sigma^2)/(\sqrt{2}\sigma^2) \to_d N(0, 1 + \gamma_2/2).$$

(b) When the X_i's are normal, $\gamma_2 = 0$ and $(n-1)S_n^2/\sigma_0^2 \sim \chi_{n-1}^2$ when $\sigma = \sigma_0$ is true. Hence the size of the normal theory test when normal theory is true is

$$\begin{aligned}
\alpha &= P_{\sigma_0}((n-1)S_n^2/\sigma_0^2 \ge \chi_{n-1,\alpha}^2) \\
&= P_{\sigma_0}(\sqrt{n}(S_n^2/\sigma_0^2 - 1)/\sqrt{2} \ge \sqrt{\frac{n}{2}}(\frac{\chi_{n-1,\alpha}^2}{n-1} - 1)).
\end{aligned}$$

Since $\sqrt{n}(S_n^2/\sigma_0^2 - 1)/\sqrt{2} \to_d N(0,1)$ under normality, this forces

$$\sqrt{\frac{n}{2}}(\frac{\chi_{n-1,\alpha}^2}{n-1} - 1)) \to z_\alpha.$$

Thus when the X_i's *are not normal* we have

$$\begin{aligned}
P_{\sigma_0}((n-1)S_n^2/\sigma_0^2 \ge \chi_{n-1,\alpha}^2) &= P_{\sigma_0}(\sqrt{n}(S_n^2/\sigma_0^2 - 1)/\sqrt{2} \ge \sqrt{\frac{n}{2}}(\frac{\chi_{n-1,\alpha}^2}{n-1} - 1)) \\
&\to P(N(0, 1 + \gamma_2/2) \ge z_\alpha) \\
&= P(Z \ge \frac{z_\alpha}{\sqrt{1 + \gamma_2/2}}) = 1 - \Phi(\frac{z_\alpha}{\sqrt{1 + \gamma_2/2}}).
\end{aligned}$$

When $\gamma_2 = -2$, the asymptotic size is 0; when $\gamma_2 = 0$, the asymptotic size is α; when $\gamma_2 = \infty$, the asymptotic size is $1/2$. For $\gamma_2 \in [-2, 0)$ the asymptotic size is $< \alpha$ while for $\gamma_2 \in (0, \infty)$ the asymptotic size is $> \alpha$.

4. Suppose that $X_1, \ldots, X_n$ are independent Poisson(λ) random variables (so $P(X_1 = k) = e^{-\lambda}\lambda^k/k!$, $k = 0, 1, \ldots$).
(a) Show that $\sqrt{n}(\overline{X}_n - \lambda) \to_d N(0, \text{“}something\text{''})$.
(b) Show that the sequence $\{\sqrt{n}|\overline{X}_n - \lambda|\}$ is uniformly integrable and find $\lim_{n\to\infty} E(\sqrt{n}|\overline{X}_n - \lambda|)$.
(c) Let $g(x) = x^\gamma$ for $x \ge 0$ and $0 < \gamma < \infty$. Show that $\sqrt{n}(g(\overline{X}_n) - g(\lambda)) \to_d N(0, V^2)$ and compute V^2 explicitly in terms of λ and γ. For what γ is V^2 constant in λ? Is this the value of γ that makes $g(\overline{X}_n)$ “most nearly normal”?
Solution: (a) By the univariate CLT, since $E(X_1) = \lambda$ and $\mathrm{Var}(X_1) = \lambda$,

$$\sqrt{n}(\overline{X}_n - \lambda) \to_d Y \sim N(0, \lambda);$$

note that $Y =_d \sqrt{\lambda}Z$ where $Z \sim N(0,1)$.
(b) Let $Y_n = \sqrt{n}(\overline{X}_n - \lambda)$. Then $E(Y_n) = 0$ and $E(Y_n^2) = Var(Y_n) = n(\lambda/n) = \lambda$. Hence $\overline{lim}_n E(Y_n^2) = \lambda < \infty$. But this implies that $\{|Y_n|\}_{n\ge 1}$ is uniformly integrable. Hence by Vitali's theorem (applied to almost surely convergent versions Y_n^*, Y^* as described in section 2.3), or by the Helly-Bray theorem and truncation ($E|Y_n| = E\{|Y_n| \wedge M\} + E\{(|Y_n| - M)1_{[|Y_n|>M]}\}$),

$$E|Y_n| \to E|Y| = \sqrt{\lambda}E|Z| = \sqrt{2\lambda/\pi}$$

since

$$E|Z| = 2\int_0^\infty z\phi(z)dz = 2\int_0^\infty \{-\phi'(z)\}dz = 2\phi(0) = \frac{2}{\sqrt{2\pi}} = \sqrt{\frac{2}{\pi}}.$$

(c) Since $g'(x) = \gamma x^{\gamma-1}$ the g' theorem yields

$$\sqrt{n}(g(\overline{X}_n) - g(\lambda)) \to_d g'(\lambda)Y = \gamma\lambda^{\gamma-1}Y \sim N(0, \gamma^2\lambda^{2\gamma-1}).$$

When $\gamma = 1/2$ this yields

$$\sqrt{n}(\sqrt{\overline{X}_n} - \sqrt{\lambda}) \to_d N(0, 1/4).$$

Thus $g(x) = \sqrt{x}$ is variance stabilizing for Poisson. It is not the transformation which makes $g(\overline{X}_n)$ most nearly normal in terms of skewness; it turns out that with $g(x) = x^{2/3}$ the transformed variable $g(\overline{X}_n)$ has approximate skewness 0; see e.g. Anscombe (1948), *Biometrika* **35**, 246 - 254, or Efron (1982), *Ann. Statist.* **10**, 323 - 339. Of course this transformation is *not* variance stabilizing; $\sqrt{n}(\overline{X}_n^{2/3} - \lambda^{2/3}) \to_d N(0, (4/9)\lambda^{1/3})$ which depends on λ.

5. Suppose that $X_1, X_2, \ldots$ are i.i.d. positive random variables, and define $\overline{X}_n \equiv n^{-1}\sum_{i=1}^n X_i$, $H_n \equiv 1/(n^{-1}\sum_{i=1}^n (1/X_i))$, and $G_n \equiv \{\prod_{i=1}^n X_i\}^{1/n}$ to be the *arithmetic, harmonic, and geometric* means respectively. We know that $\overline{X}_n \to_{a.s.} E(X_1) = \mu$ if and only if $E|X_1| < \infty$.
(a) Use the SLLN together with appropriate additional hypotheses to show that $H_n \to_{a.s.} 1/\{E(1/X_1)\} \equiv h$, and $G_n \to_{a.s.} \exp(E\{\log X_1\}) \equiv g$.
(c) Use the multivariate CLT and the delta method to find the joint limiting distribution of $\sqrt{n}(\overline{X}_n - \mu, H_n - h, G_n - g)$. You will need to impose or assume

additional moment conditions to be able to prove this. Specify these additional assumptions carefully.

Solution: (a) If $0 < E(1/X_1) < \infty$, then

$$\frac{1}{n}\sum_{i=1}^{n}(1/X_i) \to_{a.s.} E(1/X_1) > 0.$$

If $E|\log(X_1)| < \infty$, then

$$\log G_n = \frac{1}{n}\sum_{i=1}^{n}\log(X_i) \to_{a.s.} E\log X_1.$$

Thus by the continuous mapping theorem if both $E(1/X_1) < \infty$ and $E|\log X_1| < \infty$, it follows that

$$(H_n, G_n) \to_{a.s.} (1/E(1/X_1), \exp(E\log X_1)) \equiv (h, g).$$

(c) By the multivariate CLT, if $EX_1^2 < \infty$, $E(1/X_1)^2 < \infty$, and $E(\log X_1)^2 < \infty$, then

$$\sqrt{n}\begin{pmatrix} \overline{X}_n - \mu \\ \overline{X^{-1}}_n - E(1/X_1) \\ \overline{\log X}_n - E\log X_1 \end{pmatrix} \to_d \underline{Z} \sim N_3(0, \Sigma)$$

where

$$\Sigma = \begin{pmatrix} Var(X_1) & Cov(X_1, 1/X_1) & Cov(X_1, \log(X_1)) \\ Cov(X_1, 1/X_1) & Var(1/X_1) & Cov(1/X_1, \log X_1) \\ Cov(X_1, \log(X_1)) & Cov(1/X_1, \log X_1) & Var(\log(X_1)) \end{pmatrix}.$$

Hence by the delta method with $g(x, y, z) = (x, 1/y, \exp(z))$ so that $\nabla g(x, y, z) = \text{diag}(1, -y^{-2}, \exp(z))$ and $\nabla g(\mu, E(1/X_1), E(\log X_1)) = \text{diag}(1, -h^2, g)$, it follows that

$$\sqrt{n}\begin{pmatrix} \overline{X}_n - \mu \\ H_n - h \\ G_n - g \end{pmatrix} \to_d \nabla g \cdot \underline{Z} \sim N_3(0, \nabla g \Sigma \nabla g^T).$$