Statistics 581, Problem Set 3

Wellner; 10/11/2017

Reading: Course Notes: Chapter 2, pages 1 - 25;
Ferguson ACILST pages 26-60;
Van der Vaart Asymp. Statistics, sections 2.2 - 2.9 & 3.1 - 3.2, pages 12 - 31.
Due: Wednesday, October 18, 2017.

1. Ferguson, ACILST, page 34, problem 1(a) (modified slightly)
Suppose that $X_1, X_2, \ldots$ are i.i.d. in R^2 with distribution giving probability θ_1 to $(1,0)'$, probability θ_2 to $(0,1)'$, θ_3 to $(0,0)'$ and θ_4 to $(-1,-1)'$ where $\theta_j \ge 0$ for $j = 1,2,3,4$ and $\theta_1 + \theta_2 + \theta_3 + \theta_4 = 1$.
(a) Find $\mu = E(X_1)$.
(b) Compute $E(X_1 X_1^T)$ and $\Sigma = E(X_1 - \mu)(X_1 - \mu)^T$.
(c) Find the limiting distribution of $\sqrt{n}(\overline{X}_n - \mu)$ and describe the resulting approximation to the distribution of $\overline{X}_n$.
(d) Find values of $(\theta_1, \ldots, \theta_4)$ such that Σ has rank 1 and $\det(\Sigma) = 0$.

2. Suppose that $X_1, X_2, \ldots$ are i.i.d. (μ, σ^2) with $\mu_4 < \infty$. Let $\overline{X}_n = n^{-1}\sum_{i=1}^n X_i$ and $S_n^2 = (n-1)^{-1}\sum_{i=1}^n (X_i - \overline{X}_n)^2$ be the sample mean and sample variance respectively.
(a) Show that
$$\sqrt{n}\begin{pmatrix} \overline{X}_n - \mu \\ S_n^2 - \sigma^2 \end{pmatrix} \to_d \underline{Z} \sim N_2(0, \Sigma)$$
where
$$\begin{pmatrix} \sigma^2 & \mu_3 \\ \mu_3 & \mu_4 - \sigma^4 \end{pmatrix}.$$
(b) Suppose $\mu \ne 0$. Use (a) to find the limiting distribution of the sample *coefficient of variation* $C_n \equiv S_n/\overline{X}_n$; i.e. show that $\sqrt{n}(C_n - c) \to_d N(0, V^2)$ with $c \equiv \sigma/\mu$ and find V^2.

3. Ferguson, ACILST, page 34, problem 1(b) (modified slightly)
Suppose that $X_1, \ldots, X_n$ is a sample from the Poisson distribution with parameter $\lambda > 0$: $P(X_1 = k) = \exp(-\lambda)\lambda^k/k!$, $k = 0, 1, \ldots$. Let $Z_n = (1/n)\sum_{i=1}^n 1_{[X_i = 0]}$.
(a) What is the joint asymptotic distribution of
$$\sqrt{n}((\overline{X}_n, Z_n)' - (\lambda, e^{-\lambda})')?$$
(b) Let $p_0(\lambda) \equiv P_\lambda(X_1 = 0)$. What is the asymptotic distribution of $\hat{p}_0 \equiv p_0(\hat{\lambda}_n)$ where $\hat{\lambda}_n = \overline{X}_n$?
(c) What is the joint asymptotic distribution of $(Z_n, \hat{p}_0)$ (after centering and rescaling)?
(d) Compute the ratio of the asymptotic variances of the two estimators Z_n and $\hat{p}_0$ of $p_0(\lambda)$. Which estimator would you prefer if the Poisson model (assumption) holds? Which estimator would you prefer if the Poisson model (assumption) fails?

4. 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) Let $X_n \sim t_n$, the $t-$distribution with n degrees of freedom; thus $X_n \stackrel{d}{=} Z/\sqrt{Y_n/n}$ where $Y_n \sim \chi^2_n$ is independent of Z. Fix $r > 0$ (large). Compute $E|X_n|^r$ exactly as a function of n and r in terms of the Gamma function $\Gamma(r) = \int_0^\infty t^{r-1}e^{-t}dt$. Use Stirling's formula, $\Gamma(r+1) \sim \sqrt{2\pi r}(r/e)^r$ to show that $\limsup_{n\to\infty} E|X_n|^r < \infty$. Use this to show that $\{|X_n|^p : \ n \ge 1\}$ is uniformly integrable for any integer $p \ge 1$, and hence that $E(X_n^p) \to E(Z^p)$.

5. 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?

6. **Optional bonus problem 1.** Suppose that $X_1, \ldots, X_n$ are independent $N(0,1)$ random variables, and let $Y_i = X_i^2$, for $i = 1, \ldots, n$. Thus $\sum_1^n Y_i \sim \chi^2_n$.
(a) Show that $\sqrt{n}(\overline{Y}_n - 1) \to_d N(0, \text{"}\textit{something}\text{"})$, and find "something".
(b) Show that for each $r > 0$, $\sqrt{n}(\overline{Y}_n^r - 1) \to_d N(0, V^2(r))$ and find $V^2(r)$ as a function of r.
(c) Show that $V^2(1/3) = 2/9$ and that

$$\frac{\sqrt{n}(\overline{Y}_n^{1/3} - (1 - 2/(9n)))}{\sqrt{2/9}} \to_d N(0,1)\,.$$

(d) Make normal probability plots to compare the approximations in (a) and (c). [The transformation in (c) is called the "Wilson-Hilferty" transformation of a χ^2 random variable.]

7. **Optional bonus problem 2.** (Van der Vaart, page 24)
(a) Suppose that X_n and Y_n are independent random vectors with $X_n \to_d X$ and $Y_n \to_d Y$. Show that $(X_n, Y_n) \to (X, Y)$ where X and Y are independent.
(b) Suppose that $P(X_n = i/n) = 1/n$ for $i = 1, 2, \ldots, n$. From HW #1 we know that $X_n \to_d X \sim$ Uniform$(0,1)$. Show that there exists a Borel set B such that $P(X_n \in B) = 1$ but $P(X \in B) = 0$. In particular, with $P_n = \mathcal{L}(X_n)$ and $P = \mathcal{L}(X)$, $d_{TV}(P_n, P) = 1$ for each n.