Statistics 581, Problem Set 3

Wellner; 10/14/2009

Reading: Lehmann & Casella, TPE, pages 54-61 and pages 75-78.
Ferguson, ACILST, pages 1 - 60.

Due: Wednesday, October 21, 2009.

1. 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?

2. Suppose that $X_1, \ldots, X_n$ are i.i.d. $N(\theta, \theta^2/r^2)$ where $\theta \in (0, \infty)$ and $r > 0$ is known. Thus the "signal-to-noise ratio" $\mu/\sigma = \theta/\sqrt{\theta^2/r^2} = r$ is a constant.
(a) Find the score function for θ (for $n = 1$).
(b) Find the information for θ (for $n = 1$).
(c) Express the likelihood equation for θ in the form of a polynomial in θ; what is the degree of this polynomial?
(d) Use standard results to show that the MLE $\widehat{\theta}_n$ of θ is asymptotically normal and find the asymptotic variance.
(e) Show that $g(x) = \log x$ is a variance stabilizing transformation for the limiting distribution you found in (d).

3. 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
$$\sqrt{n}\begin{pmatrix} \overline{X}_n - \mu \\ S_n^2 - \sigma^2 \end{pmatrix} \rightarrow_d \underline{Z} \sim N_2(0, \Sigma)$$
where
$$\begin{pmatrix} \sigma^2 & \mu_3 \\ \mu_3 & \mu_4 - \sigma^4 \end{pmatrix}.$$
(b) Suppose $\mu > 0$. Use (a) to find the limiting distribution of the sample *signal to noise ratio* $R_n \equiv \overline{X}_n/S_n$; i.e. show that $\sqrt{n}(R_n - r) \rightarrow_d N(0, V^2)$ with $r \equiv \mu/\sigma$ and find V^2.

4. 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 $\lambda > 0$: $p_k(\lambda) = P(X_1 = k) = \exp(-\lambda)\lambda^k/k!$, $k = 0, 1, \ldots$. Let J be a fixed positive integer (e.g. $J = 5$), and consider the following two estimators of $\underline{p} = (p_0, p_1, \ldots, p_J) \in \mathbb{R}^{J+1}$:
$$\underline{\widehat{p}} = (p_0(\hat{\lambda}), p_1(\hat{\lambda}), \ldots, p_J(\hat{\lambda}))$$
where $\hat{\lambda} = \overline{X}_n = n^{-1}\sum_1^n X_i$, and
$$\underline{\widetilde{p}} = n^{-1}\sum_{i=1}^n (1_{[X_i=0]}, 1_{[X_i=1]}, \ldots, 1_{[X_i=J]}).$$

(a) What is the joint asymptotic distribution of

$$\sqrt{n}\begin{pmatrix} \widehat{\underline{p}} - \underline{p} \\ \widetilde{\underline{p}}_n - \underline{p} \end{pmatrix}$$

as a vector in $\mathbb{R}^{2(J+1)}$? Is the resulting joint distribution nondegenerate? (Why or why not?)
(b) Consider $p_j(\lambda) \equiv P_\lambda(X_1 = j)$. What is the joint asymptotic distribution of $\hat{p}_j \equiv p_j(\hat{\lambda}_n)$ and $\tilde{p}_j$ where $\hat{\lambda}_n = \overline{X}_n$ and $\tilde{p}_j = n^{-1}\sum_{i=1}^n 1_{[X_i=j]}$?
(c) Compute the ratio of the asymptotic variances of the two estimators $\hat{p}_j$ and $\tilde{p}_j$ of p_j. How does this ratio behave as a function of j?
(d) Which estimator would you prefer if the Poisson model (assumption) holds? Which estimator would you prefer if the Poisson model (assumption) fails?

5. Let $X_{n1}, \ldots, X_{nn}$ be independent, $X_{nk} \sim$ Bernoulli(p_{nk}), and let $Y_n \sim$ Poisson$(\sum_{k=1}^n p_{nk})$. Let P_n be the distribution of $\sum_{k=1}^n X_{nk}$ and let Q_n be the distribution of Y_n. Show that

$$d_{TV}(P_n, Q_n) \equiv \sup_{A\in\mathcal{B}} |P(S_n \in A) - P(Y_n \in A)| \le \sum_{k=1}^n p_{nk}^2.$$

Note that when $p_{nk} = p_n \to 0$ for all k and $np_n \to \lambda$, then $\sum_{k=1}^n p_{nk}^2 = np_n^2 = (np_n)^2/n = O(n^{-1})$.
[Hint: construct S_n and Y_n on a common probability space as follows: let $T_{nk} \sim$ Poisson(p_{nk}), $k = 1, \ldots, n$ be independent, and let $Z_{nk} \sim$ Bernoulli$(1-(1-p_{nk})e^{p_{nk}})$, $k = 1, \ldots, n$ be independent and independent of the T_{nk}'s. Define

$$X_{nk} = 1_{[T_{nk}\ge 1]} + 1_{[T_{nk}=0]}1_{[Z_{nk}=1]}.$$

Set $S_n = \sum_{k=1}^n X_{nk}$, $Y_n = \sum_{k=1}^n T_{nk}$. Check that $X_{nk} \sim$ Bernoulli(p_{nk}), $Y_n \sim$ Poisson$(\sum_1^n p_{nk})$, and

$$\begin{aligned} P(T_{nk} = 0, X_{nk} = 1) &= e^{-p_{nk}} - (1 - p_{nk}) \\ P(T_{nk} \ge 1, X_{nk} = 0) &= 0 \\ P(T_{nk} \ge 2) &= 1 - e^{-p_{nk}} - p_{nk}e^{-p_{nk}}. \end{aligned}$$

Show that

$$d_{TV}(P_n, Q_n) \le P(S_n \ne Y_n) \le \sum_{k=1}^n P(X_{nk} \ne T_{nk}) \le \sum_{k=1}^n p_{nk}^2.]$$

6. **Optional bonus problem** 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.