Exam Asymptotic Statistics, Mathematische Statistiek

December 19, 2006

Department of Mathematics, Faculty of Sciences, *vrije* Universiteit

1. a) Consider a sequence of random variables X_n for which $EX_n \to \mu$ for a fixed value $\mu \epsilon \mathbb{R}$ and $\mathrm{Var}(X_n) \to 0$ as $n \to \infty$. Show that this implies $X_n \to^P \mu$ for $n \to \infty$.

   b) Give an example of a sequence of random variables X_n such that $X_n \to^P 1$ whereas $EX_n = 0$ for all n. Of course you are also expected to show your sequence to have these properties.

   c) Let X_n be a sequence of random variables and X such that $X_n \rightsquigarrow X$. Moreover, let Y_n be a sequence of random variables such that X_n are Y_n stochastically independent and $Y_n \rightsquigarrow Y$ for a random variable Y. The distribution functions of X and Y are continuous. Show that $(X_n, Y_n) \rightsquigarrow (U, V)$, where U has the same distribution as X, V has the same distribution as Y and U and V are stochastically independent. Hint: use directly the definition of convergence in distribution in $\mathbb{R}^2$.

   d) Let $X, X_1, X_2, \ldots$ be a sequence of random variables such that $X_n \to^P X$. Let $g : \mathbb{R} \to \mathbb{R}$ be a continuous function. Prove (directly from the definition of convergence in probability) that $g(X_n) \to^P g(X)$.

2. Consider the random vector $Y_n = (Y_{n,1}, \ldots, Y_{n,k})$, multinomially distributed with parameters n and $(p_1, \ldots, p_k)$

   a) Formulate a theorem for the asymptotic distribution (for $n \to \infty$) of the sequence of random variables X_n, where
   $$X_n = \sum_{i=1}^{k} \frac{(Y_{n,i} - np_i)^2}{np_i}$$

   b) Prove this theorem.

3. The random variables $X_1, X_2, \ldots$ are independent and distributed according to the Poisson distribution with (unknown) parameter $\theta > 0$:
   $$P_\theta(X_1 = k) = \frac{\theta^k}{k!} e^{-\theta}, \quad k = 0, 1, 2, \ldots$$
   This means that the expectation as well as the variance of X_1 is θ. Denote by $\bar{X}_n$ the mean of the first n X_i's in the sequence.

   a) Derive the asymptotic distribution of $T_n = \sqrt{n}(\sqrt{\bar{X}_n} - \sqrt{\theta})$.

   b) Based on a), construct an asymptotic confidence interval of level $1-\alpha$ for θ. (if you cannot answer a), just denote the asymptotic distribution function by G and express the confidence interval in terms of G).

4. Let $X_1, X_2, \ldots$ be independent and identically distributed random variables with probability density function
   $$p_\mu(x) = \frac{1}{2} e^{-|x-\mu|} \quad x \epsilon \mathbb{R}$$

The distribution is a shifted Laplace distribution with unknown, but fixed, shift μ. Define, for $\theta \epsilon I\!R$, the function

$$\psi_\theta(x) = (x-\theta)^3 \quad (x \epsilon I\!R)$$

and consider the Z-estimator $\hat{\theta}_n$ defined as a zero of the function

$$\Psi_n(\theta) = I\!P_n \psi_\theta = \frac{1}{n}\sum_{i=1}^{n} \psi_\theta(X_i)$$

a) Using properties of the function Ψ_n, show that $\hat{\theta}_n$ is well defined. In other words: show that Ψ_n has exactly one point where it becomes zero.

It is easily shown that $\int_{-\infty}^{\infty} x^2 e^{-|x|}\,dx = 4$ (you don't have to show this).

b) Prove that $\hat{\theta}_n \to^P \mu$.

Using the Gamma function it is also easy to derive that $\int_{-\infty}^{\infty} x^6 e^{-|x|}\,dx = 1440$.

c) The random variables $\sqrt{n}(\hat{\theta}_n - \mu)$ are asymptotically normally distributed. What parameters do you expect for this asymptotic normal distribution?

5. a) Suppose F is a continuous distribution function on $I\!R$ and (F_n) a sequence of distribution functions such that

$$\lim_{n\to\infty} F_n(x) = F(x) \quad \text{for all } x \epsilon I\!R$$

Prove that this implies the following (stronger) type of convergence:

$$\lim_{n\to\infty} \|F_n - F\|_\infty = \lim_{n\to\infty} \sup_{x \epsilon I\!R} |F_n(x) - F(x)| = 0.$$

b) It can be shown that the Mean Integrated Squared Error (MISE) of the kernel estimator $\hat{f}_{n,h}$ for the density f has the following form:

$$MISE_f(\hat{f}_{n,h}) = I_{f,K}(n,h) + J_{f,K}(h)$$

The two terms (variance and squared bias) have the following asymptotic behavior

$$nhI_{f,K}(n,h) \to c_1(f,K) > 0 \text{ en } h^{-4}J_{f,K}(h) \to c_2(f,K) > 0$$

for $n \to \infty$, $h \downarrow 0$ and $nh \to \infty$. We decide to choose $h = h_n = cn^{-\alpha}$ for some $0 < \alpha < 1$. Determine (and argue) which choice of α is 'asymptotically MISE optimal'.

c) Taking the optimal α, one can show that with $h = h_n = cn^{-\alpha}$,

$$\lim_{n\to\infty} n^{4\alpha} MISE_f(\hat{f}_{n,h_n}) = \frac{c_1(f,K)}{c} + c^4 \cdot c_2(f,K)$$

Derive the optimal choice of c.

Grading:

1a: 2	1c: 2	2a: 1	3a: 4	4a: 2	4c: 2	5b: 2
1b: 2	1d: 3	2b: 4	3b: 2	4b: 3	5a: 4	5c: 3

The final grade is computed as follows: $\frac{\text{number of points} + 4}{4}$. Good luck with the exam!