Afdeling Wiskunde	Exam Mathematische Statistiek=Asymptotic Statistics
Vrije Universiteit	31 January 2006

Give clear but brief explications of your answers!
You may write your answers in Dutch, English or French.
This exam has 5 problems on 2 pages. The (nonuniform) credits are listed at the end of the exam.

1. The random vectors $(X_1, Y_1), \ldots, (X_n, Y_n)$ are independent and bivariate-normally distributed with $\mathrm{E}X_1 = \mathrm{E}Y_1 = \mu$, $\operatorname{var} X_1 = \operatorname{var} Y_1 = 1$, and $\operatorname{cov}(X_1, Y_1) = \rho$. We define estimators

$$\hat\mu_n = \tfrac{1}{2}\bar X_n + \tfrac{1}{2}\bar Y_n,$$
$$\hat\rho_n = \frac{1}{n}\sum_{i=1}^n (X_i - \hat\mu_n)(Y_i - \hat\mu_n).$$

   a. Show that $\hat\rho_n$ is asymptotically consistent for ρ.
   b. Show that $\operatorname{var}\big((X_1-\mu)(Y_1-\mu)\big) = 1+\rho^2$ and $\operatorname{cov}\big(X_1-\mu, (X_1-\mu)(Y_1-\mu)\big) = 0$.
   c. Show that the sequence $\sqrt{n}(\hat\rho_n - \rho)$ converges in distribution to the $N(0, 1+\rho^2)$-distribution.
   d. Give the form of a test for the null hypothesis $H_0\colon \rho = 0$ of asymptotic level 5 %.
   e. Show that this test rejects with probability tending to one as $n \to \infty$ if $\rho > 0$.
   f. The sample correlation coefficient r_n is known to satisfy that $\sqrt{n}(r_n - \rho)$ tends to a $N\big(0, (1-\rho^2)^2\big)$-distribution. Which of the estimators $\hat\rho_n$ or r_n is preferable?

2. Assume that $(Y_1, X_1), \ldots, (Y_n, X_n)$ are independent and identically distributed stochastic vectors distributed according to the non-linear regression model

$$Y_i = \sin(\theta_0 X_i) + e_i,$$

for $\theta_0 \in (0,1)$ and independent random variables X_i and e_i with X_i uniformly distributed on $(0,1)$ and e_i normally distributed with mean zero ($i = 1, 2, \ldots, n$). Let $\hat\theta_n$ be the point of minimum of

$$\theta \mapsto \sum_{i=1}^n \big(Y_i - \sin(\theta X_i)\big)^4, \qquad \theta \in [0,1].$$

   a. Give a heuristic argument showing that the sequence $\hat\theta_n$ converges in probability to θ_0.
   b. Which theorems and/or lemmas can be used to make this argument mathematically rigorous?
   c. Which limit distribution do you expect for the sequence $\sqrt{n}(\hat\theta_n - \theta_0)$?

3. Assume that X_n possesses a multinomial distribution with parameters n and $p = (p_1, \ldots, p_k)$.
   a. Formulate a theorem concerning the limit in distribution of the sequence of variables $\sum_{i=1}^k (X_{ni} - np_i)^2/(np_i)$ as $n \to \infty$.
   b. Prove this theorem.

4. Let $X_1, \ldots, X_n$ be independent and identically distributed random variables with density f.
   a. Give the formula for a kernel estimator $\hat f_n$ for f.
   b. Give the definition of the mean integrated square error (MISE) of $\hat f_n$ and derive its decomposition in square bias and variance.
   c. If it is known that f is once differentiable with derivative satisfying $\int f'(x)^2\,dx < \infty$, then it can be shown that the square bias is bounded above by a multiple of $C_f h^2$ for h the bandwidth of the kernel estimator and $C_f := \int f'(x)^2\,ds$. Which (asymptotic) choice of bandwidth can be recommended as $n \to \infty$ if no more is a-priori known about f?

5. Suppose that $(Y_1, X_1), \ldots, (Y_n, X_n)$ are i.i.d. stochastic vectors with $X_i \in \mathbb{R}$, $Y_i \in \{0, 1\}$ and a discrete distribution given by

$$P_\theta(X_i = x, Y_i = y) = h(x)\, \psi(\theta x)^y \big(1 - \psi(\theta x)\big)^{1-y}.$$

Here h and ψ are known functions and $\theta \in \mathbb{R}$ is an unknown parameter.

a. Determine the score function of the model.

b. Show that the Fisher information for θ in a single observation (X_i, Y_i) is given by

$$i_\theta = \mathrm{E} \frac{\psi'(\theta X_1)^2 X_1^2}{\psi(\theta X_1)\big(1 - \psi(\theta X_1)\big)}.$$

c. Give a general formula for an asymptotic confidence interval based on the maximum likelihood estimator.

d. Numerical maximization of the likelihood function given 100 observed values $(x_1, y_1), \ldots, (x_{100}, y_{100})$ yields the maximum likelihood estimator $\hat{\theta} = 1$. The observed values $x_1, \ldots, x_{100}$ are 50 times 0 and 50 times $\log 2$. We choose $\psi(x) = (1 + e^x)^{-1}$. Determine an approximate 95 % confidence interval for θ based on the maximum likelihood estimator.

Credits:

1a: 2	2a: 4	3: 8	4a: 2	5a: 2
1b: 4	2b: 2		4b: 3	5b: 2
1c: 7	2c: 4		4c: 3	5c: 2
1d: 2				5d: 2
1e: 2				
1f: 2				

Mark = total/53*9+1.

Graded exams can be inspected at the student administration of the Vrije Universiteit from three weeks after the exam onwards.