Statistics 581, Midterm Exam Solution

Wellner; 11/05/2001

1. (24 points) **Define** *any three* of the following terms. In each case, provide an appropriate context for your definition.
(a) A σ−field $\mathcal{A}$ of subsets of a set Ω.
(b) The event $[A_n \ i.o.]$.
(a) A measurable function X.
(d) Convergence in probability.
(e) Convergence in distribution (of a sequence of random variables).
(f) The inverse or quantile function F^{-1} of a distribution function F.

   Solution: See chapters 0 and 1.

2. (24 points) **State** *any three* of the following results:
(a) The monotone convergence theorem.
(b) The dominated convergence theorem.
(c) Fatou's lemma.
(d) The Mann-Wald theorem.
(e) Scheffé's theorem.
(f) The Liapunov central limit theorem.

   Solution: See chapters 0, 1, and 2

3. (30 points) A sequence of random variables X_n is "bounded in probability", which we express in symbols as $X_n = O_p(1)$, if for every $\epsilon > 0$ there exist M and n_0 such that $P(|X_n| > M) < \epsilon$ for all $n > n_0$; i.e. if

$$\lim_{M \to \infty} \limsup_{n \to \infty} P(|X_n| > M) = 0 \, .$$

   We write $X_n = O_p(b_n)$ for a sequence of positive real numbers b_n if $X_n/b_n = O_p(1)$.
(a) Show that if $X_n \to_d X$, then $X_n = O_p(1)$.
Now suppose that $X_1, X_2, X_3, \ldots$ are i.i.d. with mean μ and variance σ^2 (so $E(X^2) < \infty$. Let $S_n = X_1 + \ldots + X_n$ and $\overline{X}_n = S_n/n$.
(b) Is it true that:
(i) $S_n = O_p(1)$? (ii) $S_n = O_p(n^{1/2})$? (iii) $\overline{X}_n = O_p(n^{-1/2})$?
(iv) $n^{1/2}(\overline{X}_n - \mu) = O_p(1)$?
(v) $\sin(S_n) = O_p(1)$?

Solution: (a) Let F be the distribution function of X and let $\epsilon > 0$. Fix M large with $\pm M \in C_F = \{x \in R : F \text{ is continuous at } x\}$ and both $F(-M) < \epsilon/2$ and $1 - F(M) < \epsilon/2$. This is possible since the discontinuity points of F are at most countable and $F(-M) \to 0$, $1 - F(M) \to 0$ as $M \to \infty$. Then

$$\begin{aligned} P(|X_n| > M) &= P([X_n > M] \cup [X_n < -M]) \\ &\leq P(X_n > M) + P(X_n \leq -M) \\ &= 1 - F_n(M) + F_n(-M) \\ &\to 1 - F(M) + F(-M) \\ &< \epsilon/2 + \epsilon/2 = \epsilon\,. \end{aligned}$$

Thus for this choice of M we have

$$\limsup_{n\to\infty} P(|X_n| > M) \leq \epsilon\,.$$

Hence $X_n = O_p(1)$.
(b) (i) No; $S_n = O_p(n)$ in general (if $\mu \neq 0$), and $S_n = O_p(n^{1/2})$ if $\mu = 0$; $S_n = O_p(1)$ fails in either case.
(b) (ii) As noted in (i), this is true if $\mu = 0$, since then it follows that $n^{-1/2} S_n \to_d N(0, \sigma^2)$ so that $n^{-1/2} S_n = O_p(1)$ by the result of (a). If $\mu \neq 0$ then $S_n = O_p(n)$, but it is not $O_p(n^{1/2})$.
(b) (iii) This has the same answer as (b)(ii) since $\sqrt{n}\overline{X}_n = S_n/\sqrt{n}$.
(b) (iv) Since $n^{1/2}(\overline{X}_n - \mu) \to_d N(0, \sigma^2)$, it follows from (a) that $n^{1/2}(\overline{X}_n - \mu) = O_p(1)$.
(b) (v) Yes. Since $|\sin(x)| \leq 1$ taking $M = 1$ we have $P(|\sin(S_n)| > 1) = 0$ for every $n \geq 1$, and hence $\sin(S_n)$ is trivially $O_p(1)$.

4. (30 points) Let $(\Omega, \mathcal{A}, P)$ be a probability space. Let $\{A_n\}$ be a sequence of events, $A_n \subset \Omega$, $A_n \in \mathcal{A}$ for $n = 1, 2, \ldots$, and let $X_n = 1_{A_n}$ be the indicator functions of the events A_n. Suppose that $\sum_{n=1}^{\infty} P(A_n) < \infty$. Show that $X_n \to_{a.s.} 0$.

   Solution: Let $\epsilon \in (0, 1)$. Then $P(|X_n| > \epsilon) = P(1_{A_n} > \epsilon) = P(A_n)$ and moreover

$$P(\cup_{m=n}^{\infty}[|X_m| > \epsilon]) = P(\cup_{m=n}^{\infty} A_m) \leq \sum_{m=n}^{\infty} P(A_m) \to 0$$

   as $n \to \infty$ since $\sum_{m=}^{\infty} P(A_m) < \infty$. This implies that $X_n \to_{a.s.} 0$ by Corollary 1 of Proposition 2.6.

Do **either** problems 5 **or** problem 6.

5. (30 points) Suppose that $X, X_1, X_2, \ldots$ are i.i.d. exponential(λ) random variables: hence the distribution function of all the X_i's is $F(x) = 1 - \exp(-\lambda x)$ for $x \geq 0$. Let $M_n = \max\{X_1, \ldots X_n\}$.
(a) Find $P(\lambda^{-1} < X \leq (5/2)\lambda^{-1})$.
(b) Find the inverse distribution function (or quantile function) F^{-1} corresponding to F.
(c) Find the distribution function of M_n.
(d) Compute the distribution function of $Y_n = M_n - F^{-1}(1 - 1/n)$, show that $Y_n \rightarrow_d Y$ for some random variable Y, and find the limiting distribution.
(e) Compute the density function f_n of Y_n and show that $f_n(t) \rightarrow f(t)$ for each $t \in R$ where f is the density of the random variable Y.
(f) What can you conclude from (e) and one of the results from chapter 0?

Solution: (a)

$$\begin{aligned} P(\lambda^{-1} < X \leq (5/2)\lambda^{-1}) &= F((5/2)\lambda^{-1}) - F(\lambda^{-1}) \\ &= (1 - \exp(-(5/2)) - (1 - \exp(-1)) \\ &= \exp(-1) - \exp(-5/2)\,. \end{aligned}$$

(b) Setting $t = F(x) = 1 - \exp(-\lambda x)$ and solving for x yields $F^{-1}(t) = -(1/\lambda)\log(1-t)$.
(c) Now $[M_n \leq t] = [X_1 \leq t, \ldots, X_n \leq t]$, so it follows that

$$\begin{aligned} P(M_n \leq t) &= P(X_1 \leq t, \ldots, X_n \leq t) \\ &= P(X_1 \leq t) \cdots P(X_n \leq t) \\ &= P(X_1 \leq t)^n = (1 - \exp(-\lambda t))^n\,. \end{aligned}$$

(d) From (b) we have

$$F^{-1}(1 - 1/n) = -(1/\lambda)\log(1/n) = (1/\lambda)\log(n)$$

and hence from (c) it follows that

$$\begin{aligned} P(Y_n \leq t) &= P(M_n - (1/\lambda)\log n \leq t) \\ &= P(M_n \leq t + (1/\lambda)\log n) \\ &= (1 - \exp(-\lambda(t + (1/\lambda)\log n)))^n \\ &= (1 - \exp(-\lambda t)\exp(-\log n))^n \\ &= \left(1 - \frac{\exp(-\lambda t)}{n}\right)^n \\ &\rightarrow \exp(-\exp(-\lambda t)) \end{aligned}$$

as $n \to \infty$ since $(1 - x/n)^n \to e^{-x}$.
(e) From (d) we compute

$$\begin{aligned} f_n(t) &= \frac{d}{dt} P(Y_n \le t) = n \left(1 - \frac{\exp(-\lambda t)}{n}\right)^{n-1} \frac{-1}{n} \exp(-\lambda t)(-\lambda) \\ &= \lambda \exp(-\lambda t) \left(1 - \frac{\exp(-\lambda t)}{n}\right)^{n-1} \\ &\to \lambda \exp(-\lambda t) \exp(-\exp(-\lambda t)) = f(t)\,, \end{aligned}$$

where $f(t)$ is the density of the limiting distribution $F(t) = \exp(-\exp(-\lambda t))$.
(f) Since $f_n(t) \to f(t)$ for each fixed t, it follows from Scheffé's theorem that $d_{TV}(P_n, P) = (1/2)\int |f_n(t) - f(t)| dt \to 0$.

6. (30 points) Suppose that $X, X_1, X_2, \ldots, X_n$ are independent Poisson(λ) random variables:

$$P(X = k) = e^{-\lambda} \frac{\lambda^k}{k!}, \quad k = 0, 1, 2, \ldots.$$

(a) Use the weak law of large numbers to show that the random vector

$$\underline{\overline{Y}}_n \equiv \frac{1}{n} \sum_{i=1}^{n} (X_i, 1_{[X_i=0]}, 1_{[X_i=1]})'$$

converges in probability to some vector $(a, b, c)' \equiv \underline{y}$ where (a, b, c) depends on λ. Give (a, b, c) explicitly in terms of λ.
(b) Use the multivariate CLT to show that

$$\sqrt{n}(\underline{\overline{Y}}_n - \underline{y}) \to_d \underline{W} \sim N_3(0, \Sigma)$$

for some covariance matrix Σ; compute Σ explicitly in terms of λ.
(c) The usual estimator of λ is $\widehat{\lambda}_n = \overline{X}_n$. A friend suggests the following alternative estimator of λ:

$$\tilde{\lambda}_n = \frac{\sum_{i=1}^n 1_{[X_i=1]}}{\sum_{i=1}^n 1_{[X_i=0]}} = \frac{\overline{Y}_{3,n}}{\overline{Y}_{2,n}}\,.$$

Is $\tilde{\lambda}_n$ a consistent estimator of λ? If the answer is yes, find the asymptotic variance of this estimator of λ.

Solution: (a) Now $E(X) = \lambda$, $E(1_{[X=0]}) = P(X = 0) = e^{-\lambda}$, and $E(1_{[X=1]}) = P(X = 1) = \lambda e^{-\lambda}$, so the Weak Law of Large Numbers yields

$$\underline{\overline{Y}}_n \to_p (\lambda, e^{-\lambda}, \lambda e^{-\lambda}) \equiv (a, b, c) = \underline{y}\,.$$

(b) It follows immediately from the Multivariate CLT that

$$\sqrt{n}(\underline{\overline{Y}}_n - \underline{y}) \to_d \underline{W} \sim N_3(0, \Sigma)$$

where

$$\Sigma = \begin{pmatrix} \lambda & -\lambda e^{-\lambda} & \lambda e^{-\lambda} - \lambda^2 e^{-\lambda} \\ -\lambda e^{-\lambda} & e^{-\lambda}(1 - e^{-\lambda}) & -\lambda e^{-2\lambda} \\ \lambda e^{-\lambda} - \lambda^2 e^{-\lambda} & -\lambda e^{-2\lambda} & \lambda e^{-\lambda}(1 - \lambda e^{-\lambda}) \end{pmatrix}.$$

(c) Note that $\tilde{\lambda}_n = g(\underline{\overline{Y}}_n)$ where $g(x, y, z) = z/y$, and $\lambda = g(\underline{y}) = g(\lambda, e^{-\lambda}, \lambda e^{-\lambda})$. Thus $\tilde{\lambda}_n$ is a consistent estimator of λ by (a) and the continuous mapping theorem. To find the asymptotic variance we can use the delta-method:

$$\begin{aligned} \sqrt{n}(\tilde{\lambda}_n - \lambda) &= \sqrt{n}(g(\underline{\overline{Y}}_n) - g(\underline{y})) \\ &\to \nabla g \underline{W} \sim N(0, \nabla g \Sigma \nabla g') \end{aligned}$$

where

$$\nabla g = (0, -z/y^2, 1/y) = y^{-1}(0, -z/y, 1)$$

at $(x, y, z) = (\lambda, e^{-\lambda}, \lambda e^{-\lambda})$. That is $\nabla g = e^{\lambda}(0, -\lambda, 1)$. Hence we compute

$$\nabla g \Sigma (\nabla g)' = \lambda(1 + \lambda)e^{\lambda}.$$

Note that the asymptotic relative efficiency of this estimator with respect to the usual estimator is

$$\frac{\lambda}{\lambda(1 + \lambda)e^{\lambda}} = (1 + \lambda)^{-1} e^{-\lambda} < 1$$

for all $\lambda > 0$. Here is a plot of this relative efficiency as a function of λ. Thus you would probably try to dissuade your friend from using $\tilde{\lambda}_n$ if $\overline{X}_n$ is available.

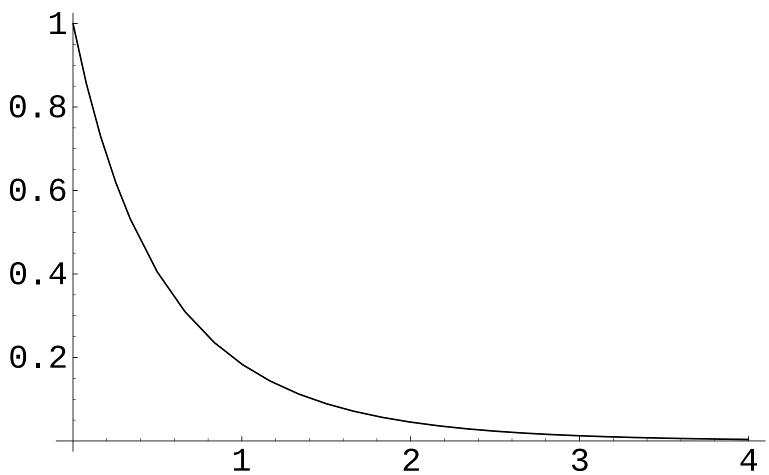

Figure 1: Asymptotic relative efficiency, $\tilde{\lambda}_n$ relative to $\hat{\lambda}_n$.