Statistics 523, Problem Set 2 Solutions

Wellner; 4/15/99

1. PfS, Exercise 14.1.2, page 309. Use chf's to show that if $X_1, X_2, \ldots$ are i.i.d. with mean μ, then $\overline{X}_n \rightarrow_p \mu$ as $n \rightarrow \infty$. Equivalently,

$$\overline{X}_n \rightarrow_d \delta_\mu \equiv \text{the degenerate distribution with mass 1 at } \mu .$$

Solution: Let $\mu \equiv E(X_1)$ and $\phi(t) \equiv E\exp(itX_1)$. Since $E|X_1| < \infty$, it follows from inequality 4.2 that

$$|\phi(t) - 1 - it\mu|/|t| \rightarrow 0 \qquad \text{as } \; t \rightarrow 0 .$$

Therefore we have

$$\begin{aligned} \phi_{\overline{X}_n}(t) &= \phi(t/n)^n = \left\{ 1 + \frac{it\mu}{n} + \frac{\phi(t/n) - 1 - i(t/n)\mu}{t/n} \frac{t}{n} \right\}^n \\ &= \left\{ 1 + \frac{it}{n} \left(\mu + \frac{\phi(t/n) - 1 - i(t/n)\mu}{it/n} \right) \right\}^n \\ &\rightarrow \exp(it\mu) , \end{aligned}$$

the characteristic function of μ. Hence by the Cramér - Lévy continuity theorem, $\overline{X}_n \rightarrow_d \mu$. But since convergence in distribution to a constant implies convergence in probability (to the same constant), it follows that $\overline{X}_n \rightarrow_p \mu$.

2. PfS, Exercise 14.1.4, page 309. Show that if $T_\lambda \equiv Poisson(\lambda)$, then $(T_\lambda - \lambda)/\sqrt{\lambda} \rightarrow_d N(0,1)$ as $\lambda \rightarrow \infty$.

Solution: Since $\phi_{T_\lambda}(t) = \exp(\lambda(e^{it} - 1))$, we find that

$$\begin{aligned} \phi_{(T_\lambda - \lambda)/\sqrt{\lambda}}(t) &= E\exp(itT_\lambda/\sqrt{\lambda})\exp(-i\sqrt{\lambda}t) \\ &= \exp(\lambda(e^{it/\sqrt{\lambda}} - 1 - it/\sqrt{\lambda})) \\ &= \exp(\frac{e^{it\eta} - 1 - it\eta}{\eta^2}) \end{aligned}$$

where $\lambda^{-1} \equiv \eta^2$. Now by L'Hopital's rule,

$$\begin{aligned}\lim_{\eta\to 0} \frac{e^{it\eta} - 1 - it\eta}{\eta^2}) &= \lim_{\eta \to 0} \frac{it(e^{it\eta} - 1)}{2\eta}) \\ &= \frac{1}{2}(it)^2 = -\frac{1}{2}t^2 .\end{aligned}$$

Thus

$$\lim_{\lambda\to\infty} \phi_{(T_\lambda - \lambda)/\sqrt{\lambda}}(t) = \exp(-t^2/2) .$$

Thus we conclude that $(T_\lambda - \lambda)/\sqrt{\lambda} \to_d N(0,1)$.

3. PfS, Exercise 14.2.9, page 318. The following are equivalent:
(a) $|X_{nk}|$'s are uan: $\max_{k\le n} P(|X_{nk}| > \epsilon) \to 0$ for all $\epsilon > 0$.
(b) $\max_{k\le n} |\phi_{nk}(t) - 1| \to 0$ uniformly on every finite interval of t's.
(c) $\max_{k\le n} \int \alpha(x) dF_{nk}(x) \to 0$ for $\alpha(x) \equiv x^2 \wedge 1$.

Solution: First, (a) implies (b): Fix $\delta > 0$ and $T > 0$. Choose $\epsilon \le \delta/T$. Note that

$$\phi_{nk}(t) = Ee^{itX_{nk}} = E(e^{itX_{nk}} 1_{[|X_{nk}|\le\epsilon]}) + E(e^{itX_{nk}} 1_{[|X_{nk}|>\epsilon]}) .$$

Moreover, from the proof of Lemma 4.2,

$$\sup_{|t|\le T} |(e^{itx} - 1) 1_{[|x|\le\epsilon]}| \le \sup_{|t|\le T} |tx| 1_{[|x|\le\epsilon]} \le T\epsilon .$$

It follows that

$$\begin{aligned}\max_{1\le k\le n} \sup_{|t|\le T} |\phi_{nk}(t) - 1| &\le \max_{1\le k\le n} \sup_{|t|\le T} |E(e^{itX_{nk}} 1_{[|X_{nk}|\le\epsilon]} - 1)| \\ &\qquad + 2 \max_{1\le k\le n} P(|X_{nk}| > \epsilon) \\ &\le \max_{1\le k\le n} \sup_{|t|\le T} |E[(e^{itX_{nk}} - 1) 1_{[|X_{nk}|\le\epsilon]}]| \\ &\qquad + 3 \max_{1\le k\le n} P(|X_{nk}| > \epsilon) \\ &\le \epsilon T + \ 3 \max_{1\le k\le n} P(|X_{nk}| > \epsilon) \\ &\to \epsilon T = \delta\end{aligned}$$

as $n \to \infty$. But $\delta > 0$ was arbitrary, so (b) holds.

Now (b) implies (a): an inequality in the same spirit as the one we used to prove the continuity theorem, Inequality 13.3.1, page 293, is as follows:

$$P(|X| \geq \epsilon) \leq \frac{\epsilon}{2} \int_{[|t| \leq 2/\epsilon]} |1 - \phi(t)|\, dt\,. \tag{1}$$

We will prove this below. Suppose that (b) holds. Note that (1) implies that

$$\begin{aligned}
\max_{k \leq n} P(|X_{nk}| \geq \epsilon) &\leq \frac{\epsilon}{2} \max_{k \leq n} \int_{[|t| \leq 2/\epsilon]} |1 - \phi_{nk}(t)|\, dt \\
&\leq 2 \max_{k \leq n} \sup_{|t| \leq 2/\epsilon} |1 - \phi_{nk}(t)| \\
&\to 0
\end{aligned}$$

and hence (b) implies (a). To see that (1) holds, note that for $T \in (0, \infty)$ we have, by Fubini's theorem,

$$\begin{aligned}
\frac{1}{2T} \int_{-T}^{T} \phi(t) dt &= \frac{1}{2T} \int_{-T}^{T} E(\cos(tX) + i \sin(tX))\, dt \\
&= \frac{1}{2T} E\left\{ \int_{-T}^{T} (\cos(tX) + i \sin(tX))\, dt \right\} \\
&= E\left(\frac{\sin(TX)}{TX} \right).
\end{aligned}$$

It follows that

$$\begin{aligned}
\left|\frac{1}{2T} \int_{-T}^{T} \phi(t) dt\right| &\leq E\left|\frac{\sin(TX)}{TX}\right| \\
&\leq E\left|\frac{\sin(TX)}{TX}\right| 1_{[|X| \geq \epsilon]} + E\left|\frac{\sin(TX)}{TX}\right| 1_{[|X| < \epsilon]} \\
&\leq \frac{1}{T\epsilon} P(|X| \geq \epsilon) + 1 - P(|X| \geq \epsilon)
\end{aligned}$$

since $|\sin(y)| \leq 1$ and $|\sin(y)/y| \leq 1$. Choosing $T = 2/\epsilon$ yields

$$\left|\frac{\epsilon}{4} \int_{-2/\epsilon}^{2/\epsilon} \phi(t) dt\right| \leq 1 - \frac{1}{2} P(|X| \geq \epsilon)$$

or, equivalently,

$$\begin{aligned} P(|X| \ge \epsilon) &\le 2 - \left|\frac{\epsilon}{2}\int_{-2/\epsilon}^{2/\epsilon} \phi(t)dt\right| \\ &= \frac{\epsilon}{2}\int_{|t|\le 2/\epsilon} dt - \left|\frac{\epsilon}{2}\int_{-2/\epsilon}^{2/\epsilon} \phi(t)dt\right| \\ &\le \frac{\epsilon}{2}\int_{-2/\epsilon}^{2/\epsilon} |1-\phi(t)|dt\,; \end{aligned}$$

i.e. (1) holds.

Note that (c) implies (a) easily since, for $\epsilon \in (0,1]$,

$$1_{[|x|\ge\epsilon]} \le \frac{x^2}{\epsilon^2} \wedge 1 \le \frac{x^2 \wedge 1}{\epsilon^2} = \frac{\alpha(x)}{\epsilon^2}\,,$$

and hence

$$P(|X_{nk}| \ge \epsilon) \le \epsilon^{-2} E\alpha(X_{nk})\,.$$

Finally, (a) implies (c): for any $\epsilon < 1$,

$$\begin{aligned} E\alpha(X_{nk}) &= E\alpha(X_{nk})1_{[|X_{nk}|\le\epsilon]} + E\alpha(X_{nk})1_{[|X_{nk}|>\epsilon]} \\ &\le \epsilon^2 + P(|X_{nk}| \ge \epsilon)\,, \end{aligned}$$

and hence

$$\max_{k\le n} E\alpha(X_{nk}) \le \epsilon^2 + \max_{k\le n} P(|X_{nk}| \ge \epsilon) \to \epsilon^2\,.$$

Since this holds for arbitrary $\epsilon > 0$, (c) holds.

4. Suppose that $X_{n1}, \ldots, X_{nn}$ are independent random variables with $X_{nk} \sim$ Bernoulli(p_{nk}), and let $Y_{nk} \sim$ Poisson(p_{nk}) for $k = 1, \ldots, n$. Let P_n be the distribution of $X_n \equiv \sum_{k=1}^n X_{nk}$ and let Q_n be the distribution of $Y_n \equiv \sum_{k=1}^n Y_{nk}$. Show that

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

Note that if $p_{nk} = \lambda_k/n$ for $k = 1, \ldots, n$, then the bound becomes $\bar{\lambda}/n$.
[Hint: Construct X_n and Y_n on a common probability space as follows: Let $T_{nk} \sim$ Poisson(p_{nk}), $k = 1, \ldots, n$ and $Z_{nk} \sim$ Bernoulli$(1-(1-p_{nk})e^{p_{nk}})$, $k = 1, \ldots, n$ all be independent, and define

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

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

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

Show that

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

Solution: First, note that

$$\begin{aligned} P_n(A) - Q_n(A) &= P(X_n \in A) - P(Y_n \in A) \\ &= P(X_n \in A, X_n = Y_n) + P(X_n \in A, X_n \ne Y_n) \\ &\qquad - (P(Y_n \in A, X_n = Y_n) + P(Y_n \in A, X_n \ne Y_n)) \\ &= P(X_n \in A, X_n = Y_n) + P(X_n \in A, X_n \ne Y_n) \\ &\qquad -(P(X_n \in A, X_n = Y_n) + P(Y_n \in A, X_n \ne Y_n)) \\ &\le P(X_n \ne Y_n) \end{aligned}$$

since the first and third terms cancel, the fourth term is negative (non-positive), and the second term is clearly bounded above by the term in the last line. This argument also works for $Q_n(A) - P_n(A)$, and we conclude that

$$|P_n(A) - Q_n(A)| \le P(X_n \ne Y_n)$$

for all Borel sets A; hence $d_{TV}(P_n, Q_n) \le P(X_n \ne Y_n)$. Now $\cap_{k=1}^n [X_{nk} = T_{nk}] \subset [X_n = Y_n]$, so taking complements yields $\cup_{k=1}^n [X_{nk} \ne T_{nk}] \supset$

$[X_n \neq Y_n]$, and this implies that

$$P(X_n \neq Y_n) \leq \sum_{k=1}^{n} P(X_{nk} \neq T_{nk})\,;$$

Thus the first two inequalities in (2) hold. Now for X_{nk} as defined in terms of T_{nk} and Z_{nk} we have

$$\begin{aligned} P(X_{nk}=1) &= P(T_{nk} \geq 1) + P(T_{nk}=0)P(Z_{nk}=1) \\ &= 1 - e^{-p_{nk}} + e^{-p_{nk}}(1-(1-p_{nk})e^{p_{nk}}) \\ &= p_{nk} \end{aligned}$$

and since X_{nk} takes on only the values 0 and 1, X_{nk} ~Bernoulli(p_{nk}). Now the joint distribution of (X_{nk}, T_{nk}) in this construction is given as follows:

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

$$P(T_{nk}=j, X_{nk}=0) = 0, \qquad \text{for} \;\; j = 1, 2, \ldots\,,$$

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

$$P(T_{nk}=1, X_{nk}=1) = P(T_{nk}=1) = p_{nk}e^{-p_{nk}}\,,$$

and

$$P(T_{nk} \geq 2) = 1 - e^{-p_{nk}} - p_{nk}e^{-p_{nk}}\,.$$

Thus we have

$$\begin{aligned} P(X_{nk} \neq T_{nk}) &= 1 - e^{-p_{nk}} - p_{nk}e^{-p_{nk}} + e^{-p_{nk}} - (1-p_{nk}) \\ &= p_{nk}(1-e^{-p_{nk}}) \\ &\leq p_{nk}^2 \qquad \text{since} \;\; 1 - e^{-x} \leq x\,. \end{aligned}$$