Statistics 523, Problem Set 4 Solutions

Wellner; 4/29/2020

1. Exercise 10.2.1, PfS (2017), page 235: the following are equivalent:
   (i) The random variables are *uan*; that is,
   $\max_{k\le n} P(|X_{nk}-\mu_{nk}|>\epsilon)\to 0$ for all $\epsilon>0$.
   (ii) $\max_{k\le n}|\phi_{nk}(t)-1|\to 0$ uniformly on every finite interval of t's.
   (iii) $\max_{k\le n} E(X_{nk}^2\wedge 1)\to 0$.

   Solution: First, (i) implies (ii): 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 (ii) holds.

   Now (ii) implies (i): 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:

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

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

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

and hence (ii) implies (i). 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

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

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

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

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 (iii) implies (i) 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, (i) implies (iii): 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.

2. Suppose that $\{b_i\}_{i=1}^N$ and $\{c_i\}_{i=1}^N$ are two sequences of real numbers, and write $c(i) \equiv c_i$. Suppose that $\underline{R} = (R_1, \ldots, R_N)$ is distributed uniformly over Π_N, the collection of all permutations of $\{1, \ldots, N\}$; i.e. $P(\underline{R} = \underline{r}) = 1/N!$ for all $\underline{r} \in \Pi_N$. Let $S \equiv S_N \equiv \sum_{j=1}^N b_j c(R_j)$. (i) Show that $Var(S) = (N-1)^{-1} B_N^2 \cdot C_N^2$ where $B_N^2 = \sum_{j=1}^N (b_j - \overline{b}_N)^2$ and $C_N^2 = \sum_{j=1}^N (c_j - \overline{c}_N)^2$.

   (ii) What is said in Chen, Goldstein, and Shao (2011) about the asymptotic normality of $(S_N - E(S_N))/\sqrt{Var(S_N)}$?
   (iii) Do they say anything about the rate of convergence to normality of $S_N - E(S_N))/\sqrt{Var(S_N)}$?

   Solution: (i) Since $P(R_i = j) = 1/N$ for $j = 1, \ldots, N$ and for $i = 1, \ldots, N$,

$$Ec(R_i) = N^{-1} \sum_{j=1}^N c(j) = \overline{c}_N \quad \text{for each} \quad i = 1, \ldots, N,$$

   and

$$Ec^2(R_i) = N^{-1} \sum_{j=1}^N c^2(j) = \overline{c^2}_N \quad \text{for each} \quad i = 1, \ldots, N,$$

   it follows that

$$Var(c(R_i)) = N^{-1} \sum_{j=1}^N (c_j - \overline{c}_N)^2.$$

   Since $\sum_{j=1}^N c(R_j) = \sum_{j=1}^N c(j) =$ a constant, we find that

$$\begin{aligned} 0 &= Var\left(\sum_{j=1}^N c(j)\right) = Var\left(\sum_{j=1}^N c(R_j)\right) \\ &= NVar(c(R_1)) + N(N-1)Cov(c(R_1), c(R_2)), \end{aligned}$$

and hence the covariances are all equal to

$$Cov(c(R_i), c(R_j)) = -\frac{1}{N(N-1)} \sum_{j=1}^{N} (c_j - \overline{c}_N)^2.$$

Thus

$$\begin{aligned} Var(S_N) &= \sum_{i=1}^{N} b_i^2 Var(c(R_i)) + \sum_{i \neq j} b_i b_j Cov(c(R_i), c(R_j)) \\ &= \frac{1}{N(N-1)} \sum_{j=1}^{N} (c_j - \overline{c})^2 \left\{ (N-1) \sum_{i=1}^{N} b_i^2 - \sum_{i \neq j} b_i b_j \right\} \\ &= \frac{1}{N(N-1)} \sum_{j=1}^{N} (c_j - \overline{c})^2 N \sum_{i=1}^{N} (b_i - \overline{b})^2 \\ &= \frac{1}{N-1} \sum_{j=1}^{N} (c_j - \overline{c})^2 \sum_{i=1}^{N} (b_i - \overline{b})^2 \\ &= \frac{1}{N-1} B_N^2 C_N^2, \end{aligned}$$

where we used the identity

$$(N-1) \sum_{i=1}^{N} b_i^2 - \sum_{i \neq j} b_i b_j = N \sum_{i=1}^{N} b_i^2 - N^2 \overline{b}_N^2 = N \sum_{i=1}^{N} (b_i - \overline{b})^2.$$

(ii) C-G-S give at least three theorems providing a rate of convergence to normality in the general case of an $N \times N$ array of numbers $\{a_{i,j}\}_{i,j=1}^N$:

- **Theorem 4.8, page 101:** Let $W \equiv (S_N - E(S_N))/\sigma$ where $E(S_N) \equiv \mu = N a_{\cdot,\cdot}$,

  $$Var(Y) \equiv \sigma_A^2 = \frac{1}{N-1} \sum_{i,j} (a_{i,j} - a_{i\cdot} - a_{\cdot,j} + a_{\cdot,\cdot})^2,$$

  and

  $$\gamma \equiv \gamma_A \equiv \sum_{i,j=1}^{N} |a_{i,j} - a_{i\cdot} - a_{\cdot,j} + a_{\cdot,\cdot}|^3.$$

  Then, for $N \geq 3$,

  $$\|F_W - \Phi\|_1 \leq \frac{\gamma}{(N-1)\sigma^3} \left(16 + \frac{56}{N-1} + \frac{8}{(N-1)^2} \right).$$

- **Theorem 6.1, page 168:** Let $W \equiv (S_N - E(S_N))/\sigma$ where $E(S_N) \equiv \mu = Na_{\cdot,\cdot}$,

$$Var(Y) \equiv \sigma_A^2 \equiv \sigma^2 = \frac{1}{N-1}\sum_{i,j}(a_{i,j} - a_{i\cdot} - a_{\cdot,j} + a_{\cdot,\cdot})^2,$$

and

$$C \quad \equiv \quad C_A = \max_{1\le i,j\le N} |a_{i,j} - a_{i\cdot} - a_{\cdot,j} + a_{\cdot,\cdot}|. \tag{2}$$

Then, for $N \geq 3$,

$$\|F_W - \Phi\|_\infty \leq 16.3\,\frac{C_A}{\sigma_A}.$$

- **Theorem 6.2, page 169:** Let $W \equiv (S_N - E(S_N))/\sigma$ where $E(S_N) \equiv \mu = Na_{\cdot,\cdot}$,

$$Var(Y) \equiv \sigma_A^2 \equiv \sigma^2 = \frac{1}{N-1}\sum_{i,j}(a_{i,j} - a_{i\cdot} - a_{\cdot,j} + a_{\cdot,\cdot})^2,$$

and

$$\gamma \quad \equiv \quad \gamma_A = \sum_{1\le i,j\le N} |a_{i,j} - a_{i\cdot} - a_{\cdot,j} + a_{\cdot,\cdot}|^3. \tag{3}$$

Then, for $N \geq 2$,

$$\|F_W - \Phi\|_\infty \leq \widetilde{C}\,\frac{\gamma_A}{n\sigma_A^3}.$$

These results all have corollaries for the special case $a_{i,j} = b_i c_j$, $i, j \in \{1, \ldots, N\}$. In this case, $\mu = Na_{\cdot\cdot} = N\overline{b}_N\overline{c}_N$, $a_{i,\cdot} = b_i\overline{c}_N$, $a_{\cdot,j} = \overline{b}_N c_j$, and we find that since

$$\begin{aligned}
&a_{i,j} - a_{i\cdot} - a_{\cdot,j} + a_{\cdot,\cdot} \\
&\quad = \quad b_i c_j - b_i\overline{c}_N - \overline{b}_N c_j + \overline{b}_N\overline{c}_N \\
&\quad = \quad b_i(c_j - \overline{c}_N) - \overline{b}_N(c_j - \overline{c}_N) = (b_i - \overline{b}_N)(c_j - \overline{C}_N),
\end{aligned}$$

$$\sigma^2 = (N-1)^{-1}\sigma_B^2\sigma_C^2 = (N-1)^{-1}B_N^2 C_N^2,$$

$$
\begin{aligned}
\gamma_A &= \sum_{i,j=1}^{N} |a_{i,j} - a_{i\cdot} - a_{\cdot,j} + a_{\cdot,\cdot}|^3 \\
&= \sum_{i,j=1}^{N} |(b_i - \overline{b}_N)(c_j - \overline{C}_N)|^3 \\
&= \sum_{i=1}^{N} |b_i - \overline{b}_N|^3 \cdot \sum_{j=1}^{N} |c_j - \overline{c}_N|^3 \equiv \gamma_B \cdot \gamma_C,
\end{aligned}
$$

and

$$
\begin{aligned}
C_A &\equiv \max_{1\le i,j\le N} |a_{i,j} - a_{i\cdot} - a_{\cdot,j} + a_{\cdot,\cdot}| \\
&= \max_{1\le i\le N} |b_i - \overline{b}_N| \cdot \max_{1\le j\le N} |c_j - \overline{c}_N|.
\end{aligned}
$$

Let

$$
\begin{aligned}
m^2_{N,B} &\equiv \frac{\max_{1\le i\le N} |b_i - \overline{b}_N|^2}{B^2_N}, \\
m^2_{N,C} &\equiv \frac{\max_{1\le i\le N} |c_j - \overline{b}_N|^2}{C^2_N}, \\
\gamma_B &\equiv \sum_{i=1}^{N} |b_i - \overline{b}_N|^3, \\
\gamma_C &\equiv \sum_{j=1}^{N} |c_j - \overline{c}_N|^3.
\end{aligned}
$$

Then the bounds in the first of the two general theorems of C-G-S become:

$$
\|F_W - \Phi|_1 \quad \le \quad \frac{\gamma}{(N-1)\sigma^3}\left(16 + \frac{56}{N-1} + \frac{8}{(N-1)^2}\right)
$$

where the rate of convergence is determined by $\gamma/((N-1)\sigma^3) \equiv \gamma_A/((N-1)\sigma^3_A)$. But now

$$
\begin{aligned}
\frac{\gamma}{(N-1)\sigma^3} &= \frac{\gamma_B \cdot \gamma_C}{(N-1)((N-1)^{-1}B^2_N C^2_N)^{3/2}} \\
&= (N-1)^{1/2} \cdot \frac{\gamma_B}{B^3_N} \cdot \frac{\gamma_C}{C^3_N} \\
&= O(N^{-1/2})
\end{aligned}
$$

since $\gamma_B/B^{3/2}_N = O(N^{-1/2})$ if we assume that $N^{-1}\gamma_B \to g_b$ and $N^{-1}B^2_N \to s^2_b > 0$, and similarly for γ_C and C^2_N. [Note that these hypotheses would hold (in probability or even a.s.) if the b_i's were i.i.d. with finite absolute third moment.]

In the case of the second of the three general theorems, the bound becomes

$$\|F_W - \Phi\|_\infty \le 16.3\, \frac{C_A}{\sigma_A}.$$

where $C = C_A$ is given by (2) and $\sigma^2 = \sigma_A^2$ is as before. Now we have

$$\begin{aligned}
\frac{C}{\sigma} &= \frac{C_b \cdot C_c}{((N-1)^{-1} B_N^2 C_N^2)^{1/2}} \\
&= \frac{\max_{1\le i\le N} |b_i - \overline{b}_N|}{B_N} \cdot \frac{\max_{1\le i\le N} |c_i - \overline{c}_N|}{C_N} \cdot (N-1)^{1/2} \\
&= O_p(N^{-1/2})
\end{aligned}$$

if we assume that the b_i's and c_j's are bounded by some constant M (uniformly in N) since then we could also expect that $B_N^2/N \to s_b^2 > 0$ and $C_N^2/N \to s_c^2 > 0$.

3. Suppose that $X_1, \ldots, X_n$ are the numbers resulting from sampling without replacement from an urn consisting of balls with the numbers $a_1, \ldots, a_N$ on the N balls. Let $\overline{a}_N \equiv \overline{a} \equiv N^{-1} \sum_1^N a_i$ and $\sigma_a^2 \equiv N^{-1} \sum_{i=1}^N (a_i - \overline{a})^2$. Let $T_n \equiv X_1 + \cdots + X_n$.
(i) Verify that for $j \ne k$, $j, k \in \{1, \ldots, N\}$,

$$Cov[X_j, X_k] = Cov[X_1, X_2] = -\frac{\sigma_a^2}{N-1}$$

and that

$$Var(T_n/n) = \frac{\sigma_a^2}{n}\left(1 - \frac{n-1}{N-1}\right).$$

The factor $(1-(n-1)/(N-1))$ is sometimes called the *finite-sampling correction factor*; note that the variance of the mean is *smaller* than the variance of the mean under sampling with replacement (namely $n^{-1}\sigma_a^2$).
(ii) Is there any connection with the previous problem, # 2?

Solution: (i) Here $T_n \stackrel{d}{=} S$ as in Problem 2 with $c_i = a_i$ for $1 \le i \le N$ and $b_i = 1_{\{1,\ldots,n\}}(i)$ for $1 \le i \le N$. Now $\overline{c}_N = \overline{a}_N$, $\overline{b}_N = n/N$. Furthermore $C_N^2 = \sum_{i=1}^N (a_i - \overline{a})^2$ while

$$\begin{aligned}
B_N^2 &= \sum_{i=1}^N (b_i - \overline{b})^2 = \sum_{i=1}^N b_i^2 - N\overline{b}^2 \\
&= n - N(n/N)^2 = \frac{n}{N}(N-n).
\end{aligned}$$

Thus it follows from the calculation in Problem 2 that

$$\begin{aligned} Cov(X_i, X_j) &= Cov(a(R_i), a(R_j)) = -\frac{1}{N(N-1)} \sum_{i=1}^{N} (a_i - \overline{a})^2 \\ &= -\frac{\sigma_a^2}{N-1} \end{aligned}$$

for $i \neq j$ since $\sigma_a^2 = N^{-1} \sum_1^N (a_i - \overline{a})^2$, and

$$\begin{aligned} Var(T_n/n) &= \frac{1}{n^2} \cdot \frac{1}{N-1} \frac{n}{N} (N-n) \sum_{i=1}^{N} (a_i - \overline{a})^2 \\ &= \frac{\sigma_a^2}{n} \cdot \frac{N-n}{N-1} = \frac{\sigma_a^2}{n} \cdot \left(1 - \frac{n-1}{N-1}\right). \end{aligned}$$

(ii) As noted in the solution of (i) above, this is the special case of $a_{i,j} = b_i c_j$ with $c_j = a_j$ for $j = 1, \ldots, N$ and $b_i = 1_{\{1,\ldots,n\}}(i)$ (so that $\underline{b} = (1, 1, \ldots, 1, 0, \ldots, 0)$ with n 1's and $(N-n)$ zeros. C-G-S (2011) give several results concerning rates of convergence of F_W to Φ. In particular, see Theorem 4.10, pages 112 - 113 for a somewhat complicated bound for $\|F_W - \Phi\|_1$.

4. Suppose that $Y_1, Y_2, \ldots$ are i.i.d. with distribution function G and characteristic function $\varphi(t) = E \exp(itY_1)$. Let N_λ be a random variable with Poisson(λ) distribution and assume that N_λ is independent of the $\{Y_i\}$'s. Let $S \equiv S_\lambda \equiv \sum_{j=1}^{N_\lambda} Y_j$. Find the characteristic function ϕ_S of S.

   Solution: By conditioning on N_λ we find that

$$\begin{aligned} \phi_S(t) &= E \exp(itS) = E[E(e^{it \sum_{j=1}^{N_\lambda} Y_j} | N_\lambda)] \\ &= E(\varphi(t)^{N_\lambda}) = \sum_{k=0}^{\infty} \varphi(t)^k e^{-\lambda} \frac{\lambda^k}{k!} \\ &= e^{-\lambda} \sum_{k=0}^{\infty} \frac{(\lambda \varphi(t))^k}{k!} = e^{\lambda(\varphi(t)-1)}. \end{aligned}$$