Solutions Homework 6

November 12, 2018

Solution to Exercise 3.3.1 (a) FALSE: Here is a counterexample to the statement: Let U and V be independent $N(0,1)$ random variables. Define a sequence of bivariate random vectors by

$$(X_n, Y_n) = (U, V), \quad \forall n.$$

Then $X_n \xrightarrow{D} U$, $Y_n \xrightarrow{D} U$ (since $X_n \stackrel{D}{=} Y_n \stackrel{D}{=} U$). However, (X_n, Y_n) does not converge in distribution to the bivariate random vector (U, U) (which has a singular normal distribution in $I\!R^2$ whereas the common distribution for all (X_n, Y_n) has a nonsingular normal distribution).

(b) FALSE: Counterexample: let $U \sim Unif(0,1)$, $X_n = nI_{(0,1/n)}(U)$. The $P[X_n \neq 0] = 1/n$ which $\rightarrow 0$, so $X_n \xrightarrow{P} 0$ and hence also $X_n \xrightarrow{D} 0$. However $E[X_n] = 1$ for all n while $E[0] = 0$, of course.

(c) TRUE: Use Continuous Mapping Principle with $h(x) = x^2$.

(d) TRUE: Apply the Cramer-Wold device. Suppose that $X_n \xrightarrow{D} X$ and $Y_n \xrightarrow{D} c$, where X_n, Y_n are random k-vectors. Let v be a fixed k-vector, they by the Continuous Mapping Principle, $v'X_n \xrightarrow{D} v'X$ and $v'Y_n \xrightarrow{D} v'c$. By the univariate Slutksky result, $v'X_n + v'Y_n = v'(X_n + Y_n) \xrightarrow{D} v'(X + c)$. Since this holds for all v', the result that $X_n + Y_n \xrightarrow{D} X + c$ follows.

Solution to Exercise 3.3.2: This should be a straightforward applicaton of the Cramer-Wold device (Theorem 3.3.7). Assume the random vectors are k dimensional and fix any $u \in I\!R^k$. Then by the continuous mapping principle, $u^T X_n \xrightarrow{D} u^T X$. We are assuming $X_n - Y_n \xrightarrow{D} 0$, so by continuous mapping principle again, $u^T(X_n - Y_n) = u^T X_n - u^T Y_n \xrightarrow{D} 0$, and hence the univariate random variabls $u^T X_n$ and $u^T Y_n$ are convergence equivalent. Thus, we conclude from the univariate version of the theorem that $u^T Y_n \xrightarrow{D} u^T X$. Since u was arbitrary, we conclude from Cramer-Wold that $Y_n \xrightarrow{D} X$, as desired.

Solution to Exercise 3.3.6 (a) This is an easy application of Cramer-

Wold. Let $\underline{v}$ be a fixed d-vector. Note that

$$\begin{aligned} E[\underline{v}^t \underline{X}_n] &= \underline{v}^t \underline{\mu} \\ \text{Var}[\underline{v}^t \underline{X}_n] &= \underline{v}^t V \underline{v}. \end{aligned}$$

An application of the univariate CLT gives

$$\sqrt{n}\left(\underline{v}^t \bar{\underline{X}}_n - \underline{v}^t \underline{\mu}\right) = \underline{v}^t \left[\sqrt{n}(\bar{\underline{X}}_n - \underline{\mu})\right] \xrightarrow{D} N(0, \underline{v}^t V \underline{v}).$$

Note that if $\underline{Z} \sim N(\underline{0}, V)$ then $\underline{v}^t \underline{Z} \sim N(0, \underline{v}^t V \underline{v})$. Thus, we have shown that for all $\underline{v}$ we have $\underline{v}^t \left[\sqrt{n}(\bar{\underline{X}}_n - \underline{\mu})\right] \xrightarrow{D} \underline{v}^t \underline{Z}$ where $\underline{Z} \sim N(\underline{0}, V)$ and so it follows that $\sqrt{n}(\bar{\underline{X}}_n - \underline{\mu}) \xrightarrow{D} N(\underline{0}, V)$.

(b) Let A be a symmetric matrix such that $A^2 = V$. Basically, if $V = U \Lambda U^t$ where U is orthogonal and $\Lambda = diag(\lambda_1, \ldots, \lambda_d)$ is diagonal, then $A = U \Lambda^{1/2} U^t$ where $\Lambda^{1/2}$ is the diagonal matrix $diag(\lambda_1^{1/2}, \ldots, \lambda_d^{1/2})$. Put $\underline{Y}_n = A^{-1}\sqrt{n}(\bar{\underline{X}}_n - \underline{\mu})$. If $\underline{Z} \sim N(\underline{0}, V)$ then since $\sqrt{n}(\bar{\underline{X}}_n - \underline{\mu}) \xrightarrow{D} \underline{Z}$ we have by the continuous mapping principle that $\underline{Y}_n \xrightarrow{D} A^{-1}\underline{Z} \sim N(\underline{0}, A^{-1} V A^{-1}) = N(\underline{0}, I)$. Let $\underline{Y} \sim N(\underline{0}, I)$. Applying the continuous mapping principle again, we get $n\,(\bar{\underline{X}}_n - \underline{\mu})' V^{-1} (\bar{\underline{X}}_n - \underline{\mu}) = \underline{Y}_n^t \underline{Y}_n \xrightarrow{D} \underline{Y}^t \underline{Y} = \sum_{i=1}^d Y_i^2 \sim \chi_d^2$ since Y_1, ..., Y_d are i.i.d. $N(0, 1)$.

Solution to Exercise 3.3.7 It is somewhat easier to do part (b) first, then part (a). So, to that end, assume X_n are random k-vectors, $X_n \xrightarrow{D} X$, and A_n are random $m \times k$ matrices, $A_n \xrightarrow{D} Z$, where A is a fixed $m \times k$ matrix. As pointed out in class, for any matrix B and vector x,

$$\|Bx\| \leq \|B\| \|x\|,$$

provided the multiplication makes sense (where $\|B\|^2 = \sum_i \sum_j B_{ij}^2$). Now, $A_n \xrightarrow{D} A$ means $A_n = A + o_P(1)$, by Proposition 3.3.4(b).Hence,

$$\begin{aligned} \|A_n X_n - A X_n\| &= \|(A_n - A) X_n\| \\ &\leq \|(A_n - A)\| \|X_n\| \\ &= o_P(1) \|X_n\| \\ &= o_P(1) O_P(1) \qquad \text{(tightness)} \\ &= o_P(1). \end{aligned}$$

Thus, $A_n X_n$ and AX_n are convergence equivalent. Now $AX_n \xrightarrow{D} AX$ by the continuous mapping principle (applied to $h(x) = Ax$). The desired result follows.

To do part (a), we need only show that $A_n \xrightarrow{D} A$ where A is nonsingular implies $A_n^{-1} \xrightarrow{D} A^{-1}$. To this end, it suffices to argue that $h(B) = B^{-1}$ is a continuous mapping on the space of nonsingular matrices, then apply continuous mapping principle. One can "handwave" this argument as follows. Using the formula for B^{-1} in terms of determinants of cofactors, one sees that B^{-1} can be expressed in terms of sums of products of entries of B (which are "clearly" continuous maps) divided by $\det B$, which is also a continuous map (same argument), so $h(B) = B^{-1}$ is continuous provided $\det B$ is nonzero.

Solution to Exercise 3.4.1: By the Continuous Mapping Principle (CMP) for convergence in distribution, we have that $h(\sqrt{n}[X_n - \mu]) \xrightarrow{D} h(Z) = Z^2$, where $Z \sim N(0, \sigma^2)$. Of course $Z^2 \sim \sigma^2 \chi_1^2$.

The limiting distribution of $h(X_n) = X_n^2$ is degenerate at μ^2. Since $\sqrt{n}[X_n - \mu]$ converges in distribution, we have by the Tightness Lemma that $\sqrt{n}[X_n - \mu] = O_P(1)$, and hence $X_n = \mu + O_P(n^{-1/2})$ which implies $X_n \xrightarrow{D} \mu$ and so by CMP, $X_n^2 \xrightarrow{D} \mu^2$, a degenerate random variable.

Finally, for the last part, we apply the δ-method. $Dh(\mu) = 2\mu$, so $\sqrt{n}[h(X_n) - h(\mu)] \xrightarrow{D} 2\mu Z$ where $Z \sim N(0, \sigma^2)$, and thus $2\mu Z \sim N(0, 4\mu^2\sigma^2)$.

Solution to Exercise 3.4.2: **(a)** Since $\mathrm{Var}[X_i] = \mu^2$, we have by the CLT

$$\sqrt{n}\left(\bar{X}_n - \mu\right) \xrightarrow{D} N(0, \mu^2).$$

(b) We apply the δ-method with

$$h(u) = \exp[-x_0/u], \quad u > 0.$$

Now

$$h'(u) = (x_0/u^2)\exp[-x_0/u],$$

so

$$\begin{aligned} &\sqrt{n}\left(\exp[-x_0/\bar{X}_n] - \exp[-x_0/\mu]\right) \\ &= \sqrt{n}\left(h(\bar{X}_n - \mu\right) \\ &\xrightarrow{D} N\left(0, [h'(\mu)]^2\mu^2\right) \end{aligned}$$

$$
\begin{aligned}
&= N\left(0, \frac{x_0^2}{\mu^4}\exp[-2x_0/u]\mu^2\right) \\
&= N\left(0, \frac{x_0^2}{\mu^2}\exp[-2x_0/u]\right)
\end{aligned}
$$

Note that the variance expression is dimensionless (e.g., assume X_i are in meters), as it should be since the first expression is basically a (dimensionless) probability.

Solution to Exercise 3.4.3 Following the hint in part (a),

$$nS_n^2/\sigma^2 \sim \chi^2_{n-1} \stackrel{D}{=} \sum_{i=1}^{n-1} Y_i, \quad \text{where the } Y_i \text{ are i.i.d. } \chi^2_1.$$

Defining

$$\bar{Y}_{n-1} = \frac{1}{n-1}\sum_{i=1}^{n-1} Y_i,$$

then by the CLT,

$$\sqrt{n-1}\left(\bar{Y}_{n-1} - 1\right) \stackrel{D}{\to} N(0,2). \tag{1}$$

(Note, we used the facts that $E[Y_i] = 1$ and $\text{Var}[Y_i] = 2$.) So let's relate $\bar{Y}_n$ to S_n^2.

$$nS_n^2/\sigma^2 \stackrel{D}{=} \sum_{i=1}^{n-1} Y_i = (n-1)\bar{Y}_{n-1},$$

so

$$S_n^2 \stackrel{D}{=} \frac{n-1}{n}\sigma^2\bar{Y}_{n-1},$$

and

$$
\begin{aligned}
\sqrt{n}\left(S_n^2 - \sigma^2\right) &\stackrel{D}{=} \sqrt{n}\left(\frac{n-1}{n}\sigma^2\bar{Y}_{n-1} - \sigma^2\right) \\
&= \sqrt{\frac{n}{n-1}}\sigma^2\sqrt{n-1}\left(\bar{Y}_{n-1} - 1\right) - \frac{1}{\sqrt{n}}\sigma^2\bar{Y}_{n-1}
\end{aligned}
$$

Now

$$\sqrt{\frac{n}{n-1}} \to 1,$$

so by Slutsky (equation (3.28) of the text) and (1) above,

$$\sqrt{\frac{n}{n-1}}\sigma^2\sqrt{n-1}\left(\bar{Y}_{n-1}-1\right) \stackrel{D}{\to} \sigma^2 N(0,2) \stackrel{D}{=} N(0,2\sigma^4). \qquad (2)$$

Also,

$$\frac{1}{\sqrt{n}}\sigma^2\bar{Y}_{n-1} = O_P\left(n^{-1/2}\right) = o_P(1),$$

so by Slutsky (equation (3.27) in the text) and (2) above,

$$\sqrt{\frac{n}{n-1}}\sigma^2\sqrt{n-1}\left(\bar{Y}_{n-1}-1\right) - \frac{1}{\sqrt{n}}\sigma^2\bar{Y}_{n-1} \stackrel{D}{\to} N(0,2\sigma^4),$$

which of course gives the asymptotic distribution for $\sqrt{n}[S_n^2-\sigma^2]$.

For S_n we apply the δ-method with $h(x)=\sqrt{x}$, $h'(x) = (2\sqrt{x})^{-1}$, and

$$\sqrt{n}[S_n-\sigma] \stackrel{D}{\to} N\left(0,(2\sqrt{\sigma^2})^{-2}2\sigma^4\right) \stackrel{D}{=} N(0,\sigma^2/2).$$

(Exercise: check that units are correct.)

Solution to Exercise 3.4.6 We first consider the univariate case. We assume without loss of generality that $\mu=0$, $\sigma^2=1$. Let $C_n=\max_{1\le i\le n} c_{ni}^2$. Verifying the Lindeberg condition:

$$\begin{aligned}
&\frac{1}{\sum_{i=1}^n c_{ni}^2}\sum_{i=1}^n c_{ni}^2 E\left[Y_i^2 I[c_{ni}^2 Y_i^2 > \epsilon\sum_{i=1}^n c_{ni}^2]\right] \\
&\le \frac{1}{\sum_{i=1}^n c_{ni}^2}\sum_{i=1}^n c_{ni}^2 E\left[Y_1^2 I[C_n Y_1^2 > \epsilon\sum_{i=1}^n c_{ni}^2]\right] \\
&= E\left[Y_1^2 I[Y_1^2 > \epsilon\sum_{i=1}^n c_{ni}^2/C_n]\right] \\
&\to 0.
\end{aligned}$$

as $n\to\infty$ for all $\epsilon>0$ since the quantity in the indicator in the last expression tends to ∞, so the indicator tends to 0. We are using the dominated convergence theorem with dominating function Y_1^2. Note how easy the verification of the Lindeberg condition is. Also, taking all $c_{ni}=1$ gives the i.i.d. CLT.

Turning to the vector valued case. Note that this is not a simple application of the Cramer-Wold device. We assume the mean is 0 and that the covariance matrix V is nonsingular. Let Λ and λ be the largest and smallest eigenvalues, respectively. Then for any vector c we have

$$\lambda\|c\|^2 \;\le\; c^TVc \;\le\; \Lambda\|c\|^2,$$

and all quantities are positive if c is nonzero. We have the events

$$[(c_{ni}^TY_i)^2 > \epsilon\sum_{i=1}^n c_{ni}^TVc_{ni}] \subset [\|c_{ni}\|^2\|Y_i\|^2 > \epsilon\lambda\sum_{i=1}^n \|c_{ni}\|^2].$$

Hence,

$$\begin{aligned}
&\frac{1}{\sum_{i=1}^n c_{ni}^TVc_{ni}}\sum_{i=1}^n E\left[(c_{ni}^TY_i)^2 I[(c_{ni}^TY_i)^2 > \epsilon\sum_{i=1}^n c_{ni}^TVc_{ni}]\right]\\
&\le \quad \frac{1}{\Lambda\sum_{i=1}^n\|c_{ni}\|^2}\sum_{i=1}^n\|c_{ni}\|^2 E\left[\|Y_1\|^2 I[\|Y_1\|^2 > \epsilon\lambda\sum_{i=1}^n\|c_{ni}\|^2/\max_{1\le i\le n}\|c_{ni}\|^2]\right]\\
&= \quad \frac{1}{\Lambda}E\left[\|Y_1\|^2 I[\|Y_1\|^2 > \epsilon\lambda\sum_{i=1}^n\|c_{ni}\|^2/\max_{1\le i\le n}\|c_{ni}\|^2]\right]\\
&\to \quad 0,
\end{aligned}$$

as $n\to\infty$ for all $\epsilon > 0$ by the same argument as before. This verifies the Lindeberg condition.