SDS 387, Fall 2024
Homework 2

Due September 3, by midnight on Canvas.

1. Show that it X_n and Y_n are independent for all n and $X_n \xrightarrow{d} X$ and $Y_n \xrightarrow{d} Y$, then

$$\begin{bmatrix} X_n \\ Y_n \end{bmatrix} \xrightarrow{d} \begin{bmatrix} X \\ Y \end{bmatrix}.$$

2. In class we showed that, if $X_n \xrightarrow{d} X$ and $Y_n - x_n \xrightarrow{d} 0$ then $Y_n \xrightarrow{d} X$. Use this result to prove that, if $X_n \xrightarrow{d} X$ and $Y_n \xrightarrow{d} c$ for some constant, then

$$\begin{bmatrix} X_n \\ Y_n \end{bmatrix} \xrightarrow{d} \begin{bmatrix} X \\ c \end{bmatrix}.$$

Note that X_n and Y_n are not necessarily independent.

3. Consider the settings of the above problem. Prove the following results, known together as Slutsky's theorem:

$$X_n Y_n \xrightarrow{d} Xc \quad \text{and} \quad X_n + Y_n \xrightarrow{d} X + c.$$

4. **Polya's Theorem**. Let $\{X_n\}$ be a sequence of random variables in $\mathbb{R}$ converging to X, a random variable with a continuos c.d.f. F_X. Show that

$$\lim_n \sup_{x \in \mathbb{R}} |F_{X_n}(x) - F_X(x)| = 0,$$

where F_{X_n} is the c.d.f of X_n. The above result says that if X is continuous, then the convergence of the c.d.f.'s is uniform over $\mathbb{R}$, not just point-wise. You may (though you do not need to) proceed as follows.

   (a) Let $\epsilon \in (0,1)$ be arbitrary (small). Next, let $-\infty = x_0 < x_1 < \ldots, x_k < x_{k+1} = \infty$ be such that $F(x_i) - F(x_{i-1}) \leq \epsilon$ for all $i = 1, \ldots, k$. This is possible. Why?

   (b) For any $x \in \mathbb{R}$ there exists one $i \in \{1, \ldots, k\}$ such that $x \in [x_{i-1}, x_i]$. Show that $F_{X_n}(x) - F_X(x) \leq F_{X_n}(x_i) - F_X(x_i) + \epsilon$ and that $F_{X_n}(x) - F_X(x) \geq F_{X_n}(x_{i-1}) - F_X(x_{i-1}) - \epsilon$. Conclude that

$$\sup_{x \in \mathbb{R}} |F_{X_n}(x) - F_X(x)| \leq \max_{i=0,\ldots,k} |F_{X_n}(x_i) - F_X(x_i)| + \epsilon.$$

   (c) Deduce the result from the inequality above.

5. **Some O_P and o_P calculus**.

(a) Show that $O_p(1) + O_p(1) = O_P(1)$.

(b) Show that $o_p(1) + o_p(1) = o_P(1)$.

(c) Show that $O_P(o_p(1)) = o_P(O_P(1)) = o_p(1)$.

(d) If $X_n = o_p(1)$, can we conclude that $X_n = O_P(1)$? Explain.

(e) What can you say about the asymptotic behavior of the stochastic quantity $\frac{1}{O_P(1)}$?

6. Give an example of a sequence of independent, centered random variables $X_1, X_2, \ldots$, all with unit variances, such that $\sqrt{n}\overline{X}_n$ does not converge in distribution to $N(0,1)$. *Hint: Construct a sequence of independent centered random variables such that the probability that $X_n = 0$ converges to 1 exponentially.*

7. Let $Y_1, Y_2, \ldots$ be i.i.d. with mean zero and unit variance and let $X_k = \sigma^k Y_k$. Show that the LF condition in this case reduces to

$$\lim_n \frac{\max_{k=1,\ldots,n} \sigma_k^2}{\sum_{i=1}^n \sigma^2} = 0$$

8. Read the proofs of Theorem 1 and 2 in the paper *Variable selection via nonconcave penalized likelihood and its oracle properties,* by J. Fan and R. Li, Journal of American Statistical Association, 2001, 96, 1348-1360. This will show you how O_P and o_P notation is useful. Available here.

9. **Optional reading assignment.** In class, we saw an example of why the triangular array setup is desirable for proving CLTs when the data-generating distribution is not fixed and may change with n. Here is an example from the literature: Lemma 6 of the paper Hypothesis Testing For Densities and High-Dimensional Multinomials: Sharp Local Minimax Rates by S. Balakrishnan and L. Wasserman.