STAT 620: Asymptotic Statistics **Spring 2022**

Lecture: Jan 20

Lecturer: Xianyang Zhang

1 Tightness

A random vector X is *tight* if for every $\epsilon > 0$ there exists $M > 0$ such that $P(\|X\| > M) < \epsilon$. A set of random vectors $\{X_a : a \in A\}$ is called *uniformly tight* if for every $\epsilon > 0$ there exists $M > 0$ such that

$$\sup_{a \in A} P(\|X_a\| > M) < \epsilon \,.$$

We often use the notation $X_n = O_p(1)$ to denote that $\{X_n : n \in \mathbb{N}\}$ is uniformly tight.

2 Helly's Lemma

Each given sequence F_n of cumulative distribution functions on $\mathbb{R}^k$ has a subsequence F_{n_j} with the property that $F_{n_j}(x) \to F(x)$ at each continuity point x of a possibly defective distribution function F (i.e., F has all the properties of a cumulative distribution function with the exception that it may not satisfy $\lim_{x \to \infty} F(x) = 1$ and $\lim_{x \to -\infty} F(x) = 0$.

3 Prohorov's Theorem

Let $\{X_n : n \in \mathbb{N}\}$ be a sequence of random vectors in $\mathbb{R}^k$.

1. If $X_n \xrightarrow{d} X$ for some X, then $\{X_n : n \in \mathbb{N}\}$ is uniformly tight.
2. If $\{X_n : n \in \mathbb{N}\}$ is uniformly tight, then there exists a subsequence with $X_{n_j} \xrightarrow{d} X$ as $j \to \infty$ for some random vector X.

3.1 Sketch of the proof

- As $X_n \xrightarrow{d} X$, by the continuous mapping theorem, we have $\|X_n\| \xrightarrow{d} \|X\|$. Fix $\epsilon > 0$. Then there exists $M > 0$ such that $P(\|X\| > M) < \epsilon$. Choose M to be a continuity point of the distribution function of $\|X\|$. Then there exists $N \in \mathbb{N}$ such that $P(\|X_n\| > M) < 2\epsilon$ for all $n \geq N$. As each of the finitely many random variables $\|X_n\|$ with $n < N$ are tight, the value of M can be suitably increased such that $P(\|X_n\| > M) < 2\epsilon$ for all $n \in \mathbb{N}$.
- By Helly's Lemma, there exists a subsequence F_{n_j} of the sequence of cdf's F_n that converges pointwise to a possibly defective distribution function F at all its continuity points. It suffices to show that F is a valid probability distribution function. Towards that end, note that for any fixed $\epsilon > 0$, there exists $M > 0$ such that $1 \geq F_n(M) > 1 - \epsilon$ for all $n \in \mathbb{N}$ (M can be chosen to be a continuity point of F). Then clearly $1 \geq F(M) = \lim_{j \to \infty} F_{n_j}(M) \geq 1 - \epsilon$ and the rest of the proof follows.

4 Stochastic o and O symbols

1. Write $X_n = o_p(1)$ if $X_n \xrightarrow{P} 0$. Write $X_n = O_p(1)$ if $\{X_n : n \in \mathbb{N}\}$ is uniformly tight.
2. $X_n = o_p(R_n)$ means $X_n = R_n\, o_p(1)$. Likewise $X_n = O_p(R_n)$ means $X_n = R_n\, O_p(1)$.
3. Some facts :

$$
\begin{aligned}
&o_p(1) + o_p(1) = o_p(1)\\
&o_p(1) + O_p(1) = O_p(1)\\
&o_p(1)\, O_p(1) = o_p(1)\\
&(1 + o_p(1))^{-1} = O_p(1)\\
&o_p(O_p(R_n)) = R_n\, o_p(1)\,.
\end{aligned}
$$

5 Lemma

Let R be a function defined on $\mathcal{U} \subset \mathbb{R}^k$ such that $R(0) = 0$. Let X_n be a sequence of random vectors taking values in $\mathcal{U}$ and $X_n = o_p(1)$. Then for every $p > 0$,

1. if $R(h) = o(\|h\|^p)$ as $h \to 0$, then $R(X_n) = o_p(\|X_n\|^p)$;
2. if $R(h) = O(\|h\|^p)$ as $h \to 0$, then $R(X_n) = O_p(\|X_n\|^p)$.

5.1 Sketch of the proof of Statment 1

Define $g(h) = R(h)/\|h\|^p$ for $h \neq 0$ and $g(0) = 0$. Clearly g is continuous at 0. By the continuous mapping theorem, $g(X_n) \xrightarrow{P} g(0) = 0$. The other proof can be derived similarly.

6 Distance between probability measures

1. For two probability measures defined on some measurable space $(\Omega, \mathcal{B})$, the *Total Variation Distance* between P and Q is defined as

$$\|P - Q\|_{TV} \;=\; \sup_{A \in \mathcal{B}} |P(A) - Q(A)|\,.$$

   If P and Q admit densities p and q respectively, then

$$\|P - Q\|_{TV} \;=\; \frac{1}{2}\int |p(x) - q(x)|\, dx\,.$$

2. *Hellinger Distance* between P and Q is defined as

$$H^2(P,Q) \;=\; \frac{1}{2}\int (\sqrt{p(x)} - \sqrt{q(x)}\,)^2\, dx \;=\; 1 - \int \sqrt{p(x)q(x)}\, dx.$$

3. *Kullback Leibler Divergence* between P and Q is defined as

$$D_{KL}(P\|Q) \;=\; \int p(x) \log \frac{p(x)}{q(x)}\, dx\,.$$