Large Sample Theory for Finite Population Sampling

Study Notes April 26, 2015 Kenneth Chu

A Cumulative distribution functions

Definition A.1 *Let $X : (\,\Omega, \mathcal{A}, \mu\,) \longrightarrow \mathbb{R}$ be a $\mathbb{R}$-valued random variable. The* ***cumulative distribution function*** *of X is, by definition, the function $F_X : \mathbb{R} \longrightarrow [0,1]$ defined as follows:*

$$F_X(x) \;:=\; P(\,X \leq x\,) \;=\; \mu(\{\,\omega \in \Omega \mid X(\omega) \leq x\,\})\,, \quad \text{for each } x \in \mathbb{R}.$$

Definition A.2 *A function $f : D \subseteq \mathbb{R} \longrightarrow \mathbb{R}$ is said to be*

- ***non-decreasing*** *if $f(x) \leq f(y)$, for any $x, y \in D$ with $x \leq y$.*
- ***non-increasing*** *if $f(x) \geq f(y)$, for any $x, y \in D$ with $x \leq y$.*
- ***monotone*** *if f is either non-decreasing or non-increasing.*

Theorem A.3 *A function $F : \mathbb{R} \longrightarrow [0,1]$ is a cumulative distribution function of some $\mathbb{R}$-valued random variable if and only if each of following four conditions holds:*

- *F is non-decreasing.*
- *F is right-continuous.*
- *$\lim_{x\to-\infty} F(x) = 0$.*
- *$\lim_{x\to+\infty} F(x) = 1$.*

PROOF If $F : \mathbb{R} \longrightarrow [0,1]$ is a cumulative distribution function of some $\mathbb{R}$-valued random variable $X : (\,\Omega, \mathcal{A}, \mu\,) \longrightarrow \mathbb{R}$, then the four conditions follow immediately from the property of the probability measure μ. Conversely, suppose the four conditions hold. Let $\Omega := (0,1)$ and $\mathcal{B}(\Omega)$ the Borel subsets of Ω. Let μ be the Lebesgue measure on $(\Omega, \mathcal{B}(\Omega))$, i.e. μ is determined by:

$$\mu(\,(0,\omega]\,) \;:=\; \omega, \quad \text{for each } \omega \in \Omega = (0,1).$$

Define the random variable $X : (\Omega, \mathcal{B}(\Omega), \mu) \longrightarrow \mathbb{R}$ by:

$$X(\omega) \;:=\; \sup F^{-1}(\,(0,\omega]\,)\,, \quad \text{for each } \omega \in \Omega = (0,1).$$

Then,

□

Theorem A.4 *Let $f : D \subseteq \mathbb{R} \longrightarrow \mathbb{R}$ be a monotone function. Then,*

$$\lim_{x\to a^-} f(x) \quad \text{and} \quad \lim_{x\to a^+} f(x)$$

exist for every $a \in \text{interior}(D)$.

Definition A.5 *Let $f : D \subseteq \mathbb{R} \longrightarrow \mathbb{R}$. A point $a \in \text{interior}(D)$ is a* ***jump discontinuity*** *of f if both*

$$\lim_{x\to a^-} f(x) \quad \text{and} \quad \lim_{x\to a^+} f(x)$$

exist but they are unequal.

Corollary A.6 *A monotone $\mathbb{R}$-valued function defined on an interval of $\mathbb{R}$ can have only jump discontinuities.*

Theorem A.7 (Darboux-Froda)
The set of discontinuities of a monotone $\mathbb{R}$-valued function defined on an interval of $\mathbb{R}$ is at most countable.

B The O_P and o_P notations; convergence in distribution implies boundedness in probability

[1]

Definition B.1 (The Big-O_P notation)
Let $\{ X_n : (\Omega_n, \mathcal{A}_n, \mu_n) \longrightarrow \mathbb{R}^k \}_{n\in\mathbb{N}}$ be a sequence of $\mathbb{R}^k$-valued random variables. Let $\{ a_n \}_{n\in\mathbb{N}}$ be sequence of positive numbers. The notation $X_n = O_p(a_n)$ means:

For every $\varepsilon > 0$, there exist $C_\varepsilon > 0$ and $n_\varepsilon \in \mathbb{N}$ such that $P(\,|X_n| \leq C_\varepsilon \cdot a_n\,) > 1-\varepsilon$, for every $n \geq n_\varepsilon$.

Proposition B.2 *The following are equivalent:*

(a) $X_n = O_P(a_n)$.

(b) *For every $\varepsilon > 0$, there exists $C_\varepsilon > 0$ such that $P(\,|X_n| \leq C_\varepsilon \cdot a_n\,) > 1-\varepsilon$, for each $n \in \mathbb{N}$.*

(c) *For every $\varepsilon > 0$, there exists $C_\varepsilon > 0$ such that $\displaystyle\limsup_{n\to\infty} P(\,|X_n| > C_\varepsilon \cdot a_n\,) \leq \varepsilon$.*

(d) *For every $\varepsilon > 0$, there exists $C_\varepsilon > 0$ such that $\displaystyle\sup_{n\in\mathbb{N}} P(\,|X_n| > C_\varepsilon \cdot a_n\,) \leq \varepsilon$.*

(e) $\displaystyle\lim_{C\to\infty}\limsup_{n\to\infty} P(\,|X_n| > C \cdot a_n\,) = 0$.

(f) $\displaystyle\lim_{C\to\infty}\sup_{n\in\mathbb{N}} P(\,|X_n| > C \cdot a_n\,) = 0$.

PROOF

<u>(a) $\Longrightarrow$ (b)</u>

Let $\varepsilon > 0$ be given. By (a), there exist $B_\varepsilon > 0$ and $n_\varepsilon \in \mathbb{N}$ such that $P(\,|X_n| \leq B_\varepsilon \cdot a_n\,) > 1-\varepsilon$, for each $n \geq n_\varepsilon$.

> **Claim**: Let Y be an $\mathbb{R}^k$-valued random variable. Then, for each $\varepsilon > 0$, there exists $A_\varepsilon > 0$ such that $P(\,|Y| \leq A_\varepsilon\,) > 1-\varepsilon$.
>
> Proof of Claim: Suppose the Claim were false. Then, there exists some $\varepsilon > 0$ such that $P(\,|Y| \leq A\,) \leq 1-\varepsilon$, for every $A > 0$; equivalently, $P(\,|Y| > A\,) > \varepsilon$, for every $A > 0$. This implies $\lim_{A\to\infty} P(\,|Y| > A\,) = \limsup_{A\to\infty} P(\,|Y| > A\,) \geq \varepsilon > 0$. But this is a contradiction since $\lim_{A\to\infty} P(\,|Y| > A\,) = 0$, for every $\mathbb{R}^k$-valued random variable Y. This proves the Claim.

By the Claim, for each $i = 1, 2, \ldots, n_\varepsilon - 1$, there exists $B_\varepsilon^{(i)} > 0$ such that $P\left(|X_i| \leq B_\varepsilon^{(i)} \cdot a_i\right) > 1-\varepsilon$. Now, let $C_\varepsilon := \max\left\{ B_\varepsilon^{(1)}, B_\varepsilon^{(1)}, \ldots, B_\varepsilon^{(n_\varepsilon-1)}, B_\varepsilon \right\}$. Then, $P(\,|X_n| \leq C_\varepsilon \cdot a_n\,) > 1-\varepsilon$, for every $n \in \mathbb{N}$. This proves the implication (a) $\Longrightarrow$ (b).

<u>(b) $\Longrightarrow$ (a)</u> Trivial: Suppose (b) holds. Then (a) immediately follows with $n_\varepsilon = 1$.

<u>(a) $\Longleftrightarrow$ (c)</u> Let $\varepsilon > 0$ be given.

$$\begin{aligned}
\text{(a)} \quad &\Longleftrightarrow \quad \text{There exist } C_\varepsilon > 0 \text{ and } n_\varepsilon \in \mathbb{N} \text{ such that } P(\,|X_n| \leq C_\varepsilon \cdot a_n\,) > 1-\varepsilon, \text{ for every } n \geq n_\varepsilon. \\
&\Longleftrightarrow \quad \text{There exist } C_\varepsilon > 0 \text{ and } n_\varepsilon \in \mathbb{N} \text{ such that } P(\,|X_n| > C_\varepsilon \cdot a_n\,) \leq \varepsilon, \text{ for every } n \geq n_\varepsilon. \\
&\Longleftrightarrow \quad \text{There exist } C_\varepsilon > 0 \text{ such that } \limsup_{n\to\infty} P(\,|X_n| > C_\varepsilon \cdot a_n\,) \leq \varepsilon \quad \Longleftrightarrow \quad \text{(c)}
\end{aligned}$$

(b) $\Longleftrightarrow$ (d) Let $\varepsilon > 0$ be given.

$$\begin{aligned}
\text{(b)} \quad &\Longleftrightarrow \quad \text{There exists } C_\varepsilon > 0 \text{ such that } P(\,|X_n| \leq C_\varepsilon \cdot a_n\,) > 1 - \varepsilon, \text{ for every } n \in \mathbb{N}. \\
&\Longleftrightarrow \quad \text{There exists } C_\varepsilon > 0 \text{ such that } P(\,|X_n| > C_\varepsilon \cdot a_n\,) \leq \varepsilon, \text{ for every } n \in \mathbb{N}. \\
&\Longleftrightarrow \quad \text{There exist } C_\varepsilon > 0 \text{ such that } \sup_{n\in\mathbb{N}} P(\,|X_n| > C_\varepsilon \cdot a_n\,) \leq \varepsilon \quad \Longleftrightarrow \quad \text{(d)}
\end{aligned}$$

(d) $\Longleftrightarrow$ (f) Let $\varepsilon > 0$ be given. We first establish that (f) $\Longrightarrow$ (d).

$$\begin{aligned}
\text{(f)} \quad &\Longleftrightarrow \quad \text{There exists } C_\varepsilon > 0 \text{ such that } \sup_{n\in\mathbb{N}} P(\,|X_n| > C \cdot a_n\,) \leq \varepsilon, \text{ for each } C \geq C_\varepsilon. \\
&\Longrightarrow \quad \text{There exists } C_\varepsilon > 0 \text{ such that } \sup_{n\in\mathbb{N}} P(\,|X_n| > C_\varepsilon \cdot a_n\,) \leq \varepsilon \quad \Longleftrightarrow \quad \text{(d)}
\end{aligned}$$

Conversely, suppose (d) holds and $C \geq C_\varepsilon$. Then, $|X_n| > C \cdot a_n \Longrightarrow |X_n| > C_\varepsilon \cdot a_n$. Hence, $\{\,|X_n| > C \cdot a_n\,\} \subseteq \{\,|X_n| > C_\varepsilon \cdot a_n\,\}$, which in turn implies $P(\,|X_n| > C \cdot a_n\,) \leq P(\,|X_n| > C_\varepsilon \cdot a_n\,)$. Thus, we have

$$\sup_{n\in\mathbb{N}} P(\,|X_n| > C \cdot a_n\,) \quad \leq \quad \sup_{n\in\mathbb{N}} P(\,|X_n| > C_\varepsilon \cdot a_n\,) \quad \leq \quad \varepsilon,$$

i.e. (f) holds.

(c) $\Longleftrightarrow$ (e) Let $\varepsilon > 0$ be given. We first establish that (e) $\Longrightarrow$ (c).

$$\begin{aligned}
\text{(e)} \quad &\Longleftrightarrow \quad \text{There exists } C_\varepsilon > 0 \text{ such that } \limsup_{n\to\infty} P(\,|X_n| > C \cdot a_n\,) \leq \varepsilon, \text{ for each } C \geq C_\varepsilon. \\
&\Longrightarrow \quad \text{There exists } C_\varepsilon > 0 \text{ such that } \limsup_{n\to\infty} P(\,|X_n| > C_\varepsilon \cdot a_n\,) \leq \varepsilon \quad \Longleftrightarrow \quad \text{(c)}
\end{aligned}$$

Conversely, suppose (c) holds and $C \geq C_\varepsilon$. Then, $|X_n| > C \cdot a_n \Longrightarrow |X_n| > C_\varepsilon \cdot a_n$. Hence, $\{\,|X_n| > C \cdot a_n\,\} \subseteq \{\,|X_n| > C_\varepsilon \cdot a_n\,\}$, which in turn implies $P(\,|X_n| > C \cdot a_n\,) \leq P(\,|X_n| > C_\varepsilon \cdot a_n\,)$. Thus, we have

$$\limsup_{n\to\infty} P(\,|X_n| > C \cdot a_n\,) \quad \leq \quad \limsup_{n\to\infty} P(\,|X_n| > C_\varepsilon \cdot a_n\,) \quad \leq \quad \varepsilon,$$

i.e. (e) holds.

This completes the proof of the Proposition. $\square$

Definition B.3 (Bounded in probability)

A sequence $\{\, X_n : (\,\Omega_n, \mathcal{A}_n, \mu_n\,) \longrightarrow \mathbb{R}^k\,\}_{n\in\mathbb{N}}$ *of* $\mathbb{R}^k$*-valued random variables is said to be* ***bounded in probability*** *if* $X_n = O_P(1)$.

Theorem B.4

If a sequence $\{\, X_n : (\,\Omega_n, \mathcal{A}_n, \mu_n\,) \longrightarrow \mathbb{R}^k\,\}_{n\in\mathbb{N}}$ *of* $\mathbb{R}^k$*-valued random variables converges in distribution to some random variable* $X : (\Omega, \mathcal{A}, \mu) \longrightarrow \mathbb{R}^k$*, then the sequence* $\{\, X_n\,\}$ *is bounded in probability.*

Proof Let

$\square$

References

[1] Deville, J.-C., and Särndal, C.-E. Calibration estimators in survey sampling. *Journal of the American Statistical Association 87*, 418 (1992), 376–382.