ch8 Large-Sample Theory

- Convergence in probability

Def. Y_n converges in probability to Y as $n \to +\infty$

$Y_n \xrightarrow{P} Y$ if for every $\varepsilon > 0$

$P(|Y_n - Y| \geq \varepsilon) \to 0$ as $n \to +\infty$.

- Chebyshev's Inequality: $P(|X| \geq a) \leq \dfrac{EX^2}{a^2}$

For any $a > 0$.

$\left(I\{|X| \geq a\} \leq \dfrac{X^2}{a^2} \right)$

- if $E(Y_n - Y)^2 \to 0$ as $n \to \infty$, then $Y_n \xrightarrow{P} Y$.

$\left(P(|Y_n - Y| \geq \varepsilon) \leq \dfrac{E(Y_n - Y)^2}{\varepsilon^2} \to 0 \right)$

ex. $X_1 \cdots X_n \overset{iid}{\sim}$ mean μ, var σ^2. $\bar{X}_n = \frac{1}{n}(X_1 + \cdots + X_n)$.

$E(\bar{X}_n - \mu)^2 = Var(\bar{X}_n) = \dfrac{\sigma^2}{n} \to 0$.

$\bar{X}_n \xrightarrow{P} \mu$. Weak law of Large numbers

- If f is continuous at c and if $Y_n \xrightarrow{P} c$ then $f(Y_n) \xrightarrow{P} f(c)$

(f continuous at c $\Rightarrow$ $\forall \varepsilon > 0$. $\exists\ \delta_\varepsilon > 0$ s.t. $|f(y) - f(c)| < \varepsilon$ if $|y - c| < \delta_\varepsilon$. $\Rightarrow P(|Y_n - c| < \delta_\varepsilon) \le P(|f(Y_n) - f(c)| < \varepsilon)$

$\Rightarrow P(|f(Y_n) - f(c)| \ge \varepsilon) \le P(|Y_n - c| \ge \delta_\varepsilon) \to 0$)

- Def. A sequence of estimators δ_n, $n \ge 1$, is <u>consistent</u> for $g(\theta)$ if for any $\theta \in \Omega$. $\delta_n \xrightarrow{P_\theta} g(\theta)$. as $n \to +\infty$

- Mean squared error: $R(\theta, \delta_n) = E_\theta(\delta_n - g(\theta))^2$

  bias: $b_n(\theta) = E_\theta \delta_n - g(\theta)$

  $R(\theta, \delta_n) = b_n^2(\theta) + Var_\theta(\delta_n)$.

  if $b_n(\theta) \to 0$. $Var_\theta(\delta_n) \to 0$ $\Rightarrow$ consistency

- convergence in distribution

$Y_n, n \geq 1$ ($\underline{\text{cdf } H_n}$) converges in distribution to Y (cdf H) if

$H_n(y) \longrightarrow H(y)$ as $n \to \infty$ whenever H is continuous at y. $Y_n \Rightarrow Y$ or $Y_n \xrightarrow{D} Y$.

ex $Y_n = \frac{1}{n}$. $H_n(y) = P(Y_n \leq y) = I\{\frac{1}{n} \leq y\}$

$H(y) = P(Y \leq y) = I(0 \leq y)$.

$H_n(0) = 0 \to 0 \neq 1 = H(0)$. but $Y_n \Rightarrow Y$.

Thm: $Y_n \Rightarrow Y$ iff $Ef(Y_n) \to Ef(Y)$ for all bounded continuous function f.

(def)

Coro: if g is continuous and $Y_n \Rightarrow Y$, then $g(Y_n) \Rightarrow g(Y)$

(f is bounded and continuous. then $f \circ g$ is all bounded and continuous. $Y_n \to Y$, $Ef(g(Y_n)) \to Ef(g(Y))$

$\Rightarrow g(Y_n) \Rightarrow g(Y)$)

- Central Limit Theorem. $X_1 \cdots X_n$ iid mean μ var. σ^2

$\overline{X}_n = \frac{1}{n}(X_1 + \cdots + X_n)$.

$$\Longrightarrow \quad \sqrt{n}\,(\overline{X}_n - \mu) \Rightarrow N(0, \sigma^2)$$

$$P\left(\mu - \frac{a}{\sqrt{n}} < \overline{X}_n \le \mu + \frac{a}{\sqrt{n}}\right) = P\left(-a < \sqrt{n}\left(\overline{X}_n - \mu\right) \le a\right)$$

$$= H_n(a) - H_n(-a) \rightarrow \Phi\left(\frac{a}{\sigma}\right) - \Phi\left(-\frac{a}{\sigma}\right)$$

- Thm 8.13. $Y_n \Rightarrow Y \quad A_n \xrightarrow{P} a \quad B_n \xrightarrow{P} b$.

then $\quad A_n + B_n Y_n \Rightarrow a + bY$

- Deta Method: $\quad f(\overline{X}_n) \approx f(\mu) + f'(\mu)(\overline{X}_n - \mu)$

f is differentiable at μ. then

$$\sqrt{n}\left(f(\overline{X}_n) - f(\mu)\right) \Rightarrow N\left(0, (f'(\mu))^2 \sigma^2\right)$$

($f(\overline{X}_n) = f(\mu) + f'(\mu_n)(\overline{X}_n - \mu)$. $\quad \mu_n$ is between $\overline{X}_n$ and μ.

$|\mu_n - \mu| \le |\overline{X}_n - \mu|$. $\quad \overline{X}_n \xrightarrow{P} \mu$. $\quad \mu_n \xrightarrow{P} \mu \quad f'(\mu_n) \xrightarrow{P} f'(\mu)$

$\sqrt{n}\left(f(\overline{X}_n) - f(\mu)\right) = f'(\mu_n)\left(\sqrt{n}(\overline{X}_n - \mu)\right) \Rightarrow f'(\mu) Z$)

- $X_n \Rightarrow X$ and f is bounded and continuous

then $Ef(X_n) \to Ef(X)$.

If f is continuous but unbounded, $Ef(X_n)$ may fail.

- Def. X_n, $n \geq 1$ are uniformly integrable if

$$\sup_{n \geq 1} E\left[|X_n| I\{|X_n| \geq t\} \right] \to 0 \quad \text{as } t \to \infty.$$

$$E|X_n| \leq t + \underbrace{E\{|X_n| I\{|X_n| \geq t\}\}}_{\text{if finite for some } t.}$$

$\sup_n E|X_n| < \infty$. Thus uniform integrability implies $\sup_n E|X_n| < \infty$

The converge can fail. $Y_n \sim \text{Bernoulli}(\frac{1}{n})$

$X_n = nY_n$. $E|X_n| = 1$. X_n are not uniformly integrable

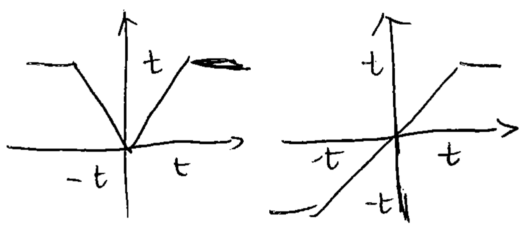
- Thm. If $X_n \Rightarrow X$. then $E|X| \le \liminf E|X_n|$

$$\left(\liminf_{n\to\infty} X_n := \lim_{n\to\infty}\left(\inf_{m\ge n} X_m \right)\right)$$

If X_n $n\ge 1$ are uniformly integrable and $X_n \Rightarrow X$, then $EX_n \longrightarrow EX$.

If X and X_n $n\ge 1$ are nonnegative and integrable with $X_n \Rightarrow X$ and $EX_n \to EX$. then X_n $n \ge 1$ are uniformly integrable

($t>0$. $g_t(x) = |x| \wedge t$. $h_t(x) = -t \vee (x \wedge t)$



continuous and bounded.

$X_n \Rightarrow X$. $E g_t(X_n) \to E g_t(X)$.

$E h_t(X_n) \to E h_t(X)$.

$\liminf E|X_n| \ge \liminf E|X_n| \wedge t = E|X| \wedge t \to E|X|$ as $t \to \infty$

$|X_n - h_t(X_n)| \le |X_n| \, I\{|X_n| \ge t\}$.

Maximum Likelihood Estimation

- likelihood function: $L(\theta) = P_\theta(x)$.

  maximum likelihood estimator (MLE): $\hat{\theta} = \hat{\theta}(x)$ maximizing $L(\theta)$.

  log-likelihood: $l(\theta) = \log L(\theta)$

ex. $p_\eta(x) = e^{\eta T(x) - A(\eta)} h(x)$

$l(\eta) = \eta T - A(\eta) + \log h(x)$. $l''(\eta) = -A''(\eta) = -Var_\eta(T) < 0$

$0 = l'(\eta) = T - A'(\eta)$. $\hat{\eta} = \psi(T)$. unique..

$X_1 \dots X_n \overset{iid}{\sim} P_\eta$.

$l(\eta) = \eta \sum_{i=1}^n T(X_i) - n A(\eta) + \log \prod_{i=1}^n h(X_i)$

$0 = l'(\eta) = \sum_{i=1}^n T(X_i) - n A'(\eta)$

$\hat{\eta} = (A')^{-1}(\bar{T})$

$\hat{\eta} = \psi(\bar{T})$ $E_\eta T(X_i) = A'(\eta)$.

$A'(\hat{\eta}) = A'(\psi(\bar{T})) = \bar{T}$

- Medians and Percentiles.

$X_1 \cdots X_n$

order statistics : $X_{(1)} \le X_{(2)} \le \cdots \le X_{(n)}$.

$X_{(1)} = \min(X_1, \cdots X_n)$. $\quad\quad X_{(n)} = \max(X_1 \cdots X_n)$.

Sample median : $\tilde{X} = \begin{cases} X_{(m)}. & n = 2m-1 \\ \frac{1}{2}(X_{(m)} + X_{(m+1)}). & n = 2m. \end{cases}$

robust to outliers.

population median : $F(\theta) = \frac{1}{2}$.

$$P\left(\sqrt{n}(\tilde{X}_n - \theta) \le a\right) = P\left(\tilde{X}_n \le \theta + \frac{a}{\sqrt{n}}\right)$$

$S_n = \#\{ i \le n : X_i \le \theta + \frac{a}{\sqrt{n}} \}$

$\tilde{X}_n \le \theta + \frac{a}{\sqrt{n}}$ iff $S_n \ge m$

$S_n \sim Bin\left(n, F(\theta + \frac{a}{\sqrt{n}})\right)$

$$\left(Y_n \sim Bin(n.p). \quad \sqrt{n}\left(\frac{Y_n}{n} - p\right) = \frac{Y_n - np}{\sqrt{n}} \Rightarrow N(0, p(1-p)) \right)$$

$$P\left(\frac{Y_n - np}{\sqrt{n}} > y\right) = 1 - P\left(\frac{Y_n - np}{\sqrt{n}} \le y\right)$$

$$\longrightarrow 1 - \Phi\left(\frac{y}{\sqrt{p(1-p)}}\right) = \Phi\left(-\frac{y}{\sqrt{p(1-p)}}\right)$$

$$P\left(\sqrt{n}(\tilde{X} - \theta) \le a\right) = P\left(S_n > m-1\right)$$

$$= P\left(\frac{S_n - nF(\theta + \frac{a}{\sqrt{n}})}{\sqrt{n}} > \frac{m-1-nF(\theta+\frac{a}{\sqrt{n}})}{\sqrt{n}}\right)$$

$$= \Phi\left(\frac{[nF(\theta+\frac{a}{\sqrt{n}}) - m + 1]/\sqrt{n}}{\sqrt{F(\theta+\frac{a}{\sqrt{n}})(1-F(\theta+\frac{a}{\sqrt{n}}))}}\right) + o(1)$$

If F is continuous at θ. $\sqrt{F(\cdot)(1-F(\cdot))} \to \frac{1}{2}$.

$$\frac{nF(\theta+\frac{a}{\sqrt{n}}) - m + 1}{\sqrt{n}} = a\,\frac{F(\theta+\frac{a}{\sqrt{n}}) - F(\theta)}{a/\sqrt{n}} + \frac{nF(\theta) - m + 1}{\sqrt{n}}$$

$$= a\,\frac{F(\theta+\frac{a}{\sqrt{n}}) - F(\theta)}{a/\sqrt{n}} + \frac{1}{2\sqrt{n}} \longrightarrow aF'(\theta)$$

$$P\left(\sqrt{n}(\tilde{X}_n - \theta) \le a\right) \to \Phi\left(2aF'(\theta)\right)$$

$$\sqrt{n}(\tilde{X}_n - \theta) \Rightarrow N\left(0, \frac{1}{4[F'(\theta)]^2}\right)$$

Thm 8.18. $r \in (0,1)$. $\tilde{\theta}_n$ — $\lfloor rn \rfloor^{th}$ order statistic

$F(\theta) = r$. $F'(\theta)$ exists, finite and positive

$$\Rightarrow \sqrt{n}(\tilde{\theta} - \theta) \Rightarrow N\left(0, \frac{r(1-r)}{[F'(\theta)]^2}\right).$$

- Asymptotic Relative Efficiency

mean vs. median.

$X_1 \cdots X_n \overset{iid}{\sim} f(x-\theta)$ location family.

f symmetric about zero

$P_\theta(X_i < \theta) = P_\theta(X_i > \theta) = \frac{1}{2}$.

$E_\theta X_i = \theta$.

By the CLT, $\sqrt{n}(\bar{X}_n - \theta) \Rightarrow N(0, \sigma^2)$

$$\sigma^2 = \int x^2 f(x)\, dx.$$

$$\sqrt{n}(\tilde{X}_n - \theta) \Rightarrow N\left(0, \frac{1}{4f^2(0)}\right)$$

f — standard normal. $f(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$.

$\sigma^2 = 1$. $\frac{1}{4f^2(0)} = \frac{\pi}{2}$

- Asymptotic relative efficiency (ARE) of $\bar{X}_n$ w.r.t. $\tilde{X}_n$ is $\frac{\frac{\pi}{2}}{1} = \frac{\pi}{2}$.

if $\sqrt{n}(\hat{\theta}_n - \theta) \Rightarrow N(0, \sigma_{\hat{\theta}}^2)$

$\sqrt{n}(\tilde{\theta}_n - \theta) \Rightarrow N(0, \sigma_{\tilde{\theta}}^2)$

the ARE of $\hat{\theta}_n$ w.r.t. $\tilde{\theta}_n$ is $\frac{\sigma_{\tilde{\theta}}^2}{\sigma_{\hat{\theta}}^2}$.

- f — double exponential $f(x) = \frac{1}{2} e^{-|x|}$

$\sigma^2 = \int x^2 \frac{1}{2} e^{-|x|} dx = \int_0^\infty x^2 e^{-x} dx = \Gamma(3) = 2! = 2$

$\sqrt{n}(\bar{X}_n - \theta) \Rightarrow N(0.2)$

$\sqrt{n}(\tilde{X}_n - \theta) \Rightarrow N(0.1)$

the ARE of $\bar{X}_n$ w.r.t. $\tilde{X}_n$ is $\frac{1}{2}$.

ex $X_1 \cdots X_n \overset{iid}{\sim} N(\theta, 1)$

estimate. $p = P_\theta(X_i \le a) = \Phi(a-\theta)$

$\hat{p} = \Phi(a - \bar{X})$

$\tilde{p} = \frac{1}{n} \#\{ i \le n : X_i \le a \} = \frac{1}{n} \sum_{i=1}^{n} I(X_i \le a)$

$\sqrt{n}\,(\tilde{p} - p) \Rightarrow N(0, \tilde{\sigma}^2)$

$\tilde{\sigma}^2 = Var(\, I(X_i \le a)) = \Phi(a-\theta)\left(1 - \Phi(a-\theta)\right)$

Δ-method:

$\sqrt{n}\,(\hat{p} - p) \Rightarrow N(0, \hat{\sigma}^2)$

$\hat{\sigma}^2 = \left[\frac{d}{dx} \Phi(a-x) \Big|_{x=\theta} \right]^2 = \phi^2(a-\theta)$

The ARE of $\hat{p}$ w.r.t. $\tilde{p}$ is

$$ARE = \frac{\Phi(a-\theta)\left(1-\Phi(a-\theta)\right)}{\phi^2(a-\theta)}$$

When $\theta = a$. $ARE = \frac{\pi}{2}$. the ARE increase without bound as $|\theta - a|$ increases.

$\tilde{p}$ — robustness.