STAT 410 **Examples for 10/23/2020 (1)** Fall 2020
A. Stepanov

Chebyshev's Inequality:

Let X be any random variable with mean μ and variance σ^2. For any $\varepsilon > 0$,

$$P(|X - \mu| < \varepsilon) \ \geq \ 1 - \frac{\sigma^2}{\varepsilon^2}$$

or, equivalently,

$$P(|X - \mu| \geq \varepsilon) \ \leq \ \frac{\sigma^2}{\varepsilon^2}$$

Def Let $U_1, U_2, \ldots$ be an infinite sequence of random variables, and let W be another random variable. Then the sequence $\{U_n\}$ **converges in probability** to W, if for all $\varepsilon > 0$,

$$\lim_{n \to \infty} P\left(\left| U_n - W \right| \geq \varepsilon \right) = 0,$$

and write $U_n \xrightarrow{P} W$.

Def An estimator $\hat{\theta}$ for θ is said to be **consistent** if $\hat{\theta} \xrightarrow{P} \theta$, i.e.,

for all $\varepsilon > 0$, $P\left(\left| \hat{\theta} - \theta \right| \geq \varepsilon \right) \to 0$ as $n \to \infty$.

The (Weak) Law of Large Numbers:

Let $X_1, X_2, \ldots$ be a sequence of independent random variables, each having the same mean μ and each having variance less than or equal to $v < \infty$. Let

$$M_n = \frac{X_1 + \ldots + X_n}{n}, \quad n = 1, 2, \ldots .$$

Then $M_n \xrightarrow{P} \mu$. That is, for all $\varepsilon > 0$, $\lim\limits_{n \to \infty} P\left(\left| M_n - \mu \right| \geq \varepsilon \right) = 0$.

① Let $X_1, X_2, \ldots$ be i.i.d. with mean μ and standard deviation σ. Let

Let $\overline{X}_n = \dfrac{X_1 + \ldots + X_n}{n}$, $n = 1, 2, \ldots$. $\quad E(\overline{X}_n) = \mu$, $\quad \mathrm{Var}(\overline{X}_n) = \dfrac{\sigma^2}{n}$.

Let $\varepsilon > 0$. $\quad 0 \le P(|\overline{X}_n - \mu| \ge \varepsilon) \le \dfrac{\mathrm{Var}(\overline{X}_n)}{\varepsilon^2} = \dfrac{\sigma^2}{n\varepsilon^2} \to 0$ as $n \to \infty$.

Then $\overline{X}_n \xrightarrow{P} \mu$, since $\lim\limits_{n\to\infty} P(|\overline{X}_n - \mu| \ge \varepsilon) = 0$ for all $\varepsilon > 0$.

As the sample size, n, increases, the sample mean, $\overline{X}$, "tends to gets closer and closer" to the population mean μ.

② Let Y_n be the number of "successes" in n independent Bernoulli trials with probability p of "success" on each trial. $\quad E(Y_n) = np$, $\quad \mathrm{Var}(Y_n) = np(1-p)$.

Let $\varepsilon > 0$. $\quad 0 \le P\left(\left| \dfrac{Y_n}{n} - p \right| \ge \varepsilon \right) \le \dfrac{p(1-p)}{n\varepsilon^2} \to 0$ as $n \to \infty$.

Then $\dfrac{Y_n}{n} \xrightarrow{P} p$.

As the number of trials, n, increases, the sample proportion of "successes", Y/n, "tends to gets closer and closer" to the probability of "success" p.

- $X_n \xrightarrow{P} X$, $Y_n \xrightarrow{P} Y \Rightarrow X_n + Y_n \xrightarrow{P} X + Y$
- $X_n \xrightarrow{P} X$, $a = \text{const} \Rightarrow aX_n \xrightarrow{P} aX$
- $X_n \xrightarrow{P} a$, g is continuous at $a \Rightarrow g(X_n) \xrightarrow{P} g(a)$
- $X_n \xrightarrow{P} X$, g is continuous $\Rightarrow g(X_n) \xrightarrow{P} g(X)$
- $X_n \xrightarrow{P} X$, $Y_n \xrightarrow{P} Y \Rightarrow X_n \cdot Y_n \xrightarrow{P} X \cdot Y$

Example 1: Let X_n have p.d.f. $f_n(x) = n\,e^{-n\,x}$, for $x > 0$, zero otherwise.

Then $X_n \xrightarrow{P} 0$, since

if $\varepsilon > 0$, $P(|X_n - 0| \ge \varepsilon) = P(X_n \ge \varepsilon) = e^{-n\,\varepsilon} \to 0$ as $n \to \infty$.

Example 2: Let X_n have p.d.f. $f_n(x) = n\,x^{n-1}$, for $0 < x < 1$, zero otherwise.

Then $X_n \xrightarrow{P} 1$, since

if $0 < \varepsilon \le 1$, $P(|X_n - 1| \ge \varepsilon) = P(X_n \le 1 - \varepsilon) = (1 - \varepsilon)^n \to 0$ as $n \to \infty$,

and if $\varepsilon > 1$, $P(|X_n - 1| \ge \varepsilon) = 0$.

Example 3: Let X_n have p.m.f. $P(X_n = 3) = 1 - \dfrac{1}{n}$, $P(X_n = 7) = \dfrac{1}{n}$.

Then $X_n \xrightarrow{P} 3$, since

if $0 < \varepsilon \le 4$, $P(|X_n - 3| \ge \varepsilon) = \dfrac{1}{n} \to 0$ as $n \to \infty$, and

if $\varepsilon > 4$, $P(|X_n - 3| \ge \varepsilon) = 0$.

Example 4: Suppose $U \sim \text{Uniform}(0, 1)$.

Let

$$X_n = \begin{cases} 1 & \text{if} \quad U \in \left(0, \dfrac{1}{3} + \dfrac{1}{n}\right) \\[2ex] 2 & \text{if} \quad U \in \left[\dfrac{1}{3} + \dfrac{1}{n}, \dfrac{2}{3} - \dfrac{1}{n}\right) \\[2ex] 3 & \text{if} \quad U \in \left[\dfrac{2}{3} - \dfrac{1}{n}, 1\right) \end{cases} \qquad X = \begin{cases} 1 & \text{if} \quad U \in \left(0, \dfrac{1}{3}\right) \\[2ex] 2 & \text{if} \quad U \in \left[\dfrac{1}{3}, \dfrac{2}{3}\right) \\[2ex] 3 & \text{if} \quad U \in \left[\dfrac{2}{3}, 1\right) \end{cases}$$

Then $P(|X_n - X| \ge \varepsilon) = \dfrac{2}{n}$, $0 < \varepsilon \le 1$, $P(|X_n - X| \ge \varepsilon) = 0$, $\varepsilon > 1$.

Therefore, $X_n \xrightarrow{P} X$.

1. Let $X_1, X_2, \ldots, X_n$ be a random sample of size n from the distribution with probability density function

$$f(x;\theta) = \begin{cases} \dfrac{1}{\theta} \cdot x^{\frac{1}{\theta} - 1} & 0 \le x \le 1 \\ 0 & \text{otherwise} \end{cases} \qquad 0 < \theta < \infty.$$

Recall: Method of moments estimator of θ is $\tilde{\theta} = \dfrac{1-\overline{X}}{\overline{X}} = \dfrac{1}{\overline{X}} - 1.$

Maximum likelihood estimator of θ is $\hat{\theta} = -\dfrac{1}{n} \cdot \sum\limits_{i=1}^{n} \ln X_i$.

$$E(X) = \frac{1}{1+\theta}, \qquad E(-\ln X) = \theta.$$

a) Show that $\hat{\theta}$ is a consistent estimator of θ.

Let $W_i = -\ln X_i$, $i = 1, 2, \ldots, n.$ Then $\hat{\theta} = \overline{W}$.

By WLLN, $\hat{\theta} = \overline{W} \xrightarrow{P} E(W) = \theta$.

b) Show that $\tilde{\theta}$ is a consistent estimator of θ.

By WLLN, $\overline{X} \xrightarrow{P} \mu = E(X) = \dfrac{1}{1+\theta}$.

$g(x) = \dfrac{1-x}{x}$ is continuous at $\dfrac{1}{1+\theta}$.

$g(\overline{X}) = \tilde{\theta}$, $\qquad g\left(\dfrac{1}{1+\theta}\right) = \theta$.

$\Rightarrow \quad \tilde{\theta} \xrightarrow{P} \theta$.

Similarly to Chebyshev's Inequality,

$$\mathrm{P}\left(\left| \hat{\theta} - \theta \right| \geq \varepsilon \right) \leq \frac{\mathrm{E}\left[\left(\hat{\theta} - \theta \right)^2 \right]}{\varepsilon^2} = \frac{\mathrm{MSE}\left(\hat{\theta} \right)}{\varepsilon^2}.$$

2. Let $X_1, X_2, \ldots, X_n$ be a random sample of size n from a uniform distribution on the interval $(0, \theta)$.

a) Show that $\tilde{\theta} = 2\,\overline{X}$ is a consistent estimator of θ.

By WLLN, $\overline{X}_n \xrightarrow{P} \mu = \mathrm{E}(X) = \frac{\theta}{2}$. $\quad\Rightarrow\quad \tilde{\theta} = 2\,\overline{X} \xrightarrow{P} \theta$.

OR

$\Rightarrow$ $\mathrm{MSE}(\tilde{\theta}) = \dfrac{\theta^2}{3n} \to 0$ as $n \to \infty$.

$\Rightarrow$ $\mathrm{P}\left(\left| \tilde{\theta} - \theta \right| \geq \varepsilon \right) \to 0$ as $n \to \infty$ for any $\varepsilon > 0$.

$\Rightarrow$ $\tilde{\theta} = 2\,\overline{X}$ is a consistent estimator for θ.

b) Show that $\hat{\theta} = \max X_i$ is a consistent estimator of θ.

$$f_{\max X_i}(x) = \mathrm{F}'_{\max X_i}(x) = \frac{n \cdot x^{n-1}}{\theta^n}, \qquad 0 < x < \theta.$$

$$\mathrm{P}\left(\left| \hat{\theta} - \theta \right| \geq \varepsilon \right) = \int_0^{\theta-\varepsilon} \frac{n \cdot x^{n-1}}{\theta^n}\, dx = \left. \frac{x^n}{\theta^n} \right|_0^{\theta-\varepsilon} = \left(1 - \frac{\varepsilon}{\theta} \right)^n \to 0 \quad \text{as } n \to \infty.$$

$\Rightarrow$ $\hat{\theta} = \max X_i$ is a consistent estimator for θ.

OR

$$\Rightarrow \qquad \text{MSE}(\hat{\theta}) = \frac{2\theta^2}{(n+1)(n+2)} \to 0 \quad \text{as } n \to \infty.$$

$$\Rightarrow \qquad \text{P}\left(\left| \hat{\theta} - \theta \right| \ge \varepsilon \right) \to 0 \quad \text{as } n \to \infty \quad \text{for any } \varepsilon > 0.$$

$\Rightarrow \qquad \hat{\theta} = \max X_i$ is a consistent estimator for θ.

2½. Let $X_1, X_2, \ldots, X_n$ be a random sample of size n from a Geometric(p) distribution (the number of independent trials until the first "success"). That is,

$$\text{P}(X_1 = k) = (1-p)^{k-1} p, \quad k = 1, 2, 3, \ldots .$$

Show that $\hat{p} = \tilde{p} = 1/\overline{X}$ is a consistent estimator of p.

By WLLN, $\overline{X} \xrightarrow{P} \mu = \text{E}(X) = 1/p$.

Since $g(x) = 1/x$ is continuous at $1/p$,

$$\hat{p} = g(\overline{X}) \xrightarrow{P} g\left(1/p\right) = p.$$

$\Rightarrow \qquad \hat{p}$ is a consistent estimator for p.

3. Let $\lambda > 0$ and let $X_1, X_2, \dots, X_n$ be a random sample from the distribution with the probability density function

$$f(x;\lambda) = 2\lambda^2 x^3 e^{-\lambda x^2}, \qquad x > 0.$$

a) Recall that the maximum likelihood estimator of λ is $\hat{\lambda} = \dfrac{2n}{\sum_{i=1}^{n} X_i^2}$.

Is $\hat{\lambda}$ a consistent estimator of λ?

$$E(X^2) = \lambda^{-2/2}\Gamma\left(\frac{2}{2}+2\right) = \lambda^{-1}\cdot\Gamma(3) = \lambda^{-1}\cdot 2! = \frac{2}{\lambda}.$$

OR

$W = X^2$ has Gamma$(\alpha = 2, \theta = \frac{1}{\lambda})$ distribution.

$$E(X^2) = E(W) = \alpha\theta = \frac{2}{\lambda}.$$

By WLLN, $\overline{X^2} = \frac{1}{n}\cdot\sum_{i=1}^{n} X_i^2 \xrightarrow{P} E(X^2) = \frac{2}{\lambda}.$

$$X_n \xrightarrow{P} a,\ g \text{ is continuous at } a \ \Rightarrow\ g(X_n) \xrightarrow{P} g(a)$$

Since $g(x) = \frac{2}{x}$ is continuous at $\frac{2}{\lambda}$,

$$\hat{\lambda} = g(\overline{X^2}) \xrightarrow{P} g\left(\frac{2}{\lambda}\right) = \lambda.$$

b) Construct a consistent estimator for λ based on $\sum\limits_{i=1}^{n} X_i^4$.

Hint: Recall that $E(X^k) = \lambda^{-k/2}\Gamma\left(\frac{k}{2}+2\right), \quad k > -4.$

$$E(X^4) = \lambda^{-4/2}\Gamma\left(\frac{4}{2}+2\right) = \lambda^{-2}\cdot\Gamma(4) = \lambda^{-2}\cdot 3! = \frac{6}{\lambda^2}.$$

By WLLN, $\overline{X^4} = \frac{1}{n}\cdot\sum\limits_{i=1}^{n} X_i^4 \xrightarrow{P} E(X^4) = \frac{6}{\lambda^2}.$

$$X_n \xrightarrow{P} a, \ g \text{ is continuous at } a \ \Rightarrow \ g(X_n) \xrightarrow{P} g(a)$$

Consider $g(x) = \sqrt{\frac{6}{x}}$. Since $g(x) = \sqrt{\frac{6}{x}}$ is continuous at $\frac{6}{\lambda^2}$,

$$\hat{\hat{\lambda}} = \sqrt{\frac{6n}{\sum\limits_{i=1}^{n} X_i^4}} = \sqrt{\frac{6}{\overline{X^4}}} = g(\overline{X^4}) \xrightarrow{P} g\left(\frac{6}{\lambda^2}\right) = \lambda.$$

Example 5:

Suppose $P(X_n = 0) = 1 - \frac{1}{n}$ and $P(X_n = n) = \frac{1}{n}$, $n = 1, 2, 3, \ldots$.

Then $E(X_n) = 1$, $n = 1, 2, 3, \ldots$.

Let $\varepsilon > 0$. Then for large n, $P(|X_n - 0| \ge \varepsilon) = \frac{1}{n} \to 0$ as $n \to \infty$.

$\Rightarrow \quad X_n \xrightarrow{P} 0.$

However, $E(X_n)$ does not approach 0 as $n \to \infty$.