Stat 401 - 11/22/17 

Hint on 3.7.1 (mean of X)
Do 3 cases for $\beta$: $0<\beta<1$
$\beta=1$
$1<\beta$

## Sec 5.1 - Convergence in Probability

Motivation Example.

- Suppose we flip a coin $H=1$, $T=0$

Calculate sample mean as

$$\bar{X}_n = \frac{X_1 + \cdots + X_n}{n}$$

sample size

We would expect / hope that

$$\bar{X}_n \to \frac{1}{2}.$$

In what sense is this convergence?

Does $\bar{X}_n \to \frac{1}{2}$ for sure?

Or is $\left|\bar{X}_n - \frac{1}{2}\right|$ very small with large probability?

def] Let $X_n$ be a sequence of RVs.
Let $X$ be another (target) RV.

We say $X_n$ converges in probability

to $X$ if for any $\varepsilon > 0$

$$\lim_{n \to \infty} \mathbb{P}\left\{ |X_n - X| > \varepsilon \right\} = 0$$

OR

$$\lim_{n \to \infty} \mathbb{P}\left\{ |X_n - X| < \varepsilon \right\} = 1$$

ex] Coin toss

1 0 0 1 1 1 1 0 0

$X_1 = 1 \qquad X_2 = \frac{1+0}{2} = \frac{1}{2} \qquad X_3 = \frac{1+0+0}{3} = \frac{1}{3}$

$X_4 = \frac{1+0+0+1}{4} = \frac{1}{2}$

Notation: $X_n \xrightarrow{P} X$

or $X_n \longrightarrow X$ i.p.

**Thm** Weak Law of Large Numbers (WLLN)

Let $X_n$ be a sequence of iid RVs with mean $\mu$ and finite variance $\sigma^2$.

Then $\bar{X}_n = \dfrac{X_1 + \cdots + X_n}{n} = \frac{1}{n} \sum X_i \longrightarrow \mu$ i.p.

**Proof** Recall $\mathbb{E}[\bar{X}_n] = \mu$

$$\mathrm{Var}[\bar{X}_n] = \frac{\sigma^2}{n}$$

$$\mathbb{P}\{|\bar{X}_n - \mu| > \varepsilon\} \le \frac{\mathrm{Var}[\bar{X}_n]}{\varepsilon^2} \quad \text{Chebyshev}$$

$0 \le \mathbb{P}(A) \ge 0$

$$= \frac{\sigma^2/n}{\varepsilon^2}$$

$\lim_{n\to\infty} \mathbb{P}(A) = 0$

$$= \frac{\sigma^2}{n\varepsilon^2} \longrightarrow 0 \text{ as } n \to \infty.$$

$\therefore \bar{X}_n \xrightarrow{\mathbb{P}} \mu$

# Properties of Convergence I.P.

① If $X_n \xrightarrow{P} X$, $Y_n \xrightarrow{P} Y$ then

$$(X_n + Y_n) \xrightarrow{P} (X+Y)$$

Proof by triangle ineq.

② If $X_n \xrightarrow{P} X$ and $a$ is a constant then

$$aX_n \xrightarrow{P} aX$$

Proof) If $a=0$, then $0 \to 0$
If $a \neq 0$, let $\varepsilon > 0$. then

$$\mathbb{P}\{|aX_n - aX| > \varepsilon\}$$

$$= \mathbb{P}\{|a||X_n - X| > \varepsilon\}$$

$$= \mathbb{P}\left\{|X_n - X| > \frac{\varepsilon}{|a|}\right\}.$$

$$\longrightarrow 0 \quad \text{since } X_n \xrightarrow{P} X \qquad \square$$

$$\frac{\varepsilon}{\varepsilon} = 1 \quad > \varepsilon$$

$$\frac{\varepsilon}{1/2} = 2\varepsilon > \varepsilon \qquad \frac{\varepsilon}{2}$$

3 If $X_n \xrightarrow{P} X$ and $g$ is a continuous function then

$$g(X_n) \xrightarrow{P} g(X)$$

4 If $X_n \xrightarrow{P} X$, $Y_n \xrightarrow{P} Y$ then

$$X_n Y_n \xrightarrow{P} XY$$

Proof:

$$X_n Y_n = \frac{X_n^2 - X_n^2 + Y_n^2 - Y_n^2 + 4X_nY_n}{4}$$

$$= \frac{(X_n^2 + 2X_nY_n + Y_n^2) - (X_n^2 - 2X_nY_n + Y_n^2)}{4}$$

$g(Z_n) = \frac{Z^2}{4}$

$$= \frac{(X_n + Y_n)^2 - (X_n - Y_n)^2}{4}$$

$$\xrightarrow[\text{by } 3]{P} \frac{(X+Y)^2 - (X-Y)^2}{4}$$

$$= XY$$

def) Let $X$ be a RV with cdf

$$F(x, \theta)$$

parameter.

Let $X_1, \dots, X_n$ be samples from $X$.

Let $T_n$ denote an estimator of $\theta$ obtained from the samples.
(i.e. $T_n$ is a statistic).

We say $T_n$ is a consistent estimator of $\theta$ if

$$T_n \xrightarrow{P} \theta$$

Interpretation: Sampling dist of $T_n$ gets more and more concentrated as $n \to \infty$.

$\mathbb{E}[T_n] = \theta$ unbiased.

ex) $\overline{X}_n$ is a consistent estimator of $\mu$ by WLLN.

ex) Let sample variance be denoted by

$$S_n^2 = \frac{1}{n-1} \sum_{i=1}^{n} (X_i - \overline{X}_n)^2$$

In Sec 2.8. we showed

$$\mathbb{E}[S_n^2] = \sigma^2 \quad \text{unbiased for } \sigma^2$$

We can show $S_n^2$ is a consistent estimator of $\sigma^2$

$$S_n^2 \xrightarrow{P} \sigma^2$$

Proof) $$S_n^2 = \frac{1}{n-1} \sum_{i=1}^{n} (X_i - \bar{X}_n)^2$$

$$= \frac{n}{n-1} \left( \frac{1}{n} \sum_{i=1}^{n} X_i^2 - \bar{X}_n^2 \right)$$

$$\xrightarrow{P} (1) \left( \mathbb{E}[X^2] - \mu^2 \right)$$

$$= \sigma^2$$

$\therefore S_n^2 \xrightarrow{P} \sigma^2$.

and $S_n \xrightarrow{P} \sigma$

Remark)

unbiased estimators ~~generall~~ generally are not invariant under transformations.

i.e. If $T_n$ is unbiased for $\theta$ then

$g(T_n)$ is NOT unbiased for $g(\theta)$.

However consistent estimators have this property.

Thm) Continuous Mapping Thm.

Let $g$ be a continuous function. If $T_n$ is consistent for $\theta$ then $g(T_n)$ is consistent for $g(\theta)$.