Convergence in Probability

Dr. Syring

February 27, 2020

Definition

A sequence of random variables $X_1, X_2, \ldots, X_n$ converges in probability to a random variables X (written $X_n \stackrel{i.p.}{\rightarrow} X$) if

$$\lim_{n \rightarrow \infty} P(|X_n - X| > \epsilon) = 0$$

for all choices of $\epsilon > 0$.

Examples: Exponential random variables

Suppose $X_n \sim Exp(n)$, meaning $f(X_n) = \frac{1}{n}\exp(-x/n)$. Then,

$$P(X_n > \epsilon) = e^{-\epsilon/n} \to 0$$

so X_n converges in probability to the constant 0. We can always think of a constant as a degenerate random variable. That means we can say $X = 0$ with probability 1, so X is effectively a constant.

Examples: Additive vanishing noise

Suppose $X_n = X + Y_n$ where $E(Y_n) = 1/n$ and $V(Y_n) = \sigma^2/n$. Recall Chebyshev's Inequality:

$$P(|X - E(X)| > \epsilon) \leq \frac{V(X)}{\epsilon^2}.$$

Using this we have

$$P(|X_n - X| > \epsilon) = P(|Y_n - 0| > \epsilon)$$

$$\leq P(|Y_n - \frac{1}{n}| > \epsilon - 1/n)$$

$$\leq \frac{\sigma^2/n}{(\epsilon - 1/n)^2} \to 0.$$

Therefore $X_n \stackrel{i.p.}{\to} X$.

Examples: Bounded X , small multiplicative noise

Suppose $X_n = XY_n$ where $|X| \leq M$ for a positive constant M , $E(Y_n) = a$ and $V(Y_n) = \sigma^2/n$. Again, using Chebyshev's Inequality:

$$P(|X_n - aX| > \epsilon) = P(|X(Y_n - a)| > \epsilon)$$

$$\leq P(|Y_n - a| > \epsilon/M)$$

$$\leq \frac{\sigma^2/n}{(\epsilon/M)^2} \to 0.$$

Therefore $X_n \stackrel{i.p.}{\to} aX$.

Sum of convergent sequences limits to sum of limits

If $X_n \stackrel{i.p.}{\to} X$ and $Y_n \stackrel{i.p.}{\to} Y$ then (using the Triangle Inequality)

$$P(|X_n + Y_n - X - Y| > \epsilon) \leq P(|X_n - X| + |Y_n - Y| > \epsilon)$$

$$\leq P(|X_n - X| > \epsilon) + P(|Y_n - Y| > \epsilon)$$

$$\to 0$$

So, $X_n + Y_n \stackrel{i.p.}{\to} X + Y$.

Continuous functions of convergent sequences converge

Suppose $X_n \stackrel{i.p.}{\to} a$ and g is a continuous function. Then $|g(x) - g(a)| \geq \epsilon \Rightarrow |x - a| \geq \delta$. So,

$$P(|g(X_n) - g(a)| \geq \epsilon) \leq P(|X_n - a| \geq \delta) \to 0$$

Applications: Sample variance

Write ${S^\star}^2 = \frac{1}{n}\sum_{i=1}^n (X_i - \bar{X})^2$. Using the shourtcut formula, rewrite as

$${S^\star}^2 = \frac{1}{n}\sum_{i=1}^n X_i^2 - \bar{X}^2.$$

Then, notice that $\frac{1}{n}\sum_{i=1}^n X_i^2 \overset{i.p.}{\to} E(X^2)$ and since $\bar{X} \overset{i.p.}{\to} \mu$ then $\bar{X}^2 \overset{i.p.}{\to} \mu^2$ by continuity. We can sum these so that ${S^\star}^2 \overset{i.p.}{\to} E(X^2) - \mu^2 = \sigma^2$.

Further, we can also consider $S^2 = \frac{n}{n-1}{S^\star}^2$. Then, $|S^2 - {S^\star}^2| = |{S^\star}^2 \frac{1}{n-1}|$ and by Markov's Inequality

$$P(|{S^\star}^2 \frac{1}{n-1}| > \epsilon) \leq \frac{E(|{S^\star}^2 \frac{1}{n-1}|)}{\epsilon} = \frac{\sigma^2}{n\epsilon} \to 0$$

hence $S^2 \overset{i.p.}{\to} = \sigma^2$ as well.

Applications: Fixed design regression

Consider the response variable Y with mean $E(Y) = a + bx$ for constants a and b and predictor/covariate variable x (non-random). Also, $V(Y) = \sigma^2$.

For data (y_i, x_i) we can estimate the mean of Y by the line $\hat{a} + \hat{b}x$ where

$$\hat{a} = \sum v_i y_i, \quad \hat{b} = \sum w_i y_i$$

where

$$v_i = \frac{1}{n} - \bar{x} w_i$$

$$w_i = \frac{x_i - \bar{x}}{\sum (x_i - \bar{x})^2}.$$

Not hard to show these are unbiased since $\sum x_i w_i = 1$ and $\sum w_i = 0$.

Then, Chebyshev implies convergence in probability to a and b if the variances vanish… The variances are

$$V(\hat{a}) = \sigma^2 \sum v_i^2$$

and

$$V(\hat{b}) = \sigma^2 \sum w_i^2 .$$

Consider a “fixed design” in which the x values are sampled on a grid (i/n, …,n/n). Then, show that the variances vanish in n .