Statistics

Feb 10, 2021

Standard setup: Collect data $x_1, x_2, \dots, x_n$, compute a statistic:

$$\frac{1}{n}\sum_{i=1}^{n} x_i = \hat{\mu}$$

numbers

If x_i were realizations of iid random variables $X_1, \dots, X_n$, then

$$\mathbb{E}\left(\frac{1}{n}\sum_{i=1}^{n} X_i\right) = \frac{1}{n}\sum_{i=1}^{n}\mathbb{E}(X_i)$$

$$= \frac{1}{n}\sum \mu$$

$$= \mu.$$

Call $\overline{X}_n = \frac{1}{n}\sum X_i$.

How is $\overline{X}_n$ distributed? What can we say about $\overline{X}_n$ as $n \to \infty$?

$\Rightarrow$ Convergence of random variables.

Ex: $X_1, X_2, \dots$ are all $N(0,1)$, and let $X \sim N(0,1)$.

Does X_n converge to X?

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

Types of Convergence

Let $X_1, X_2, \ldots$ be a sequence of random variables, and let X be another random variable.

Let F_n = CDF for X_n $\quad F_n(t) = \mathbb{P}(X_n \le t)$.

F = CDF for X $\quad F(t) = \mathbb{P}(X \le t)$.

① X_n converges <u>in probability</u> to X, $X_n \xrightarrow{P} X$, if for any $\epsilon > 0$,

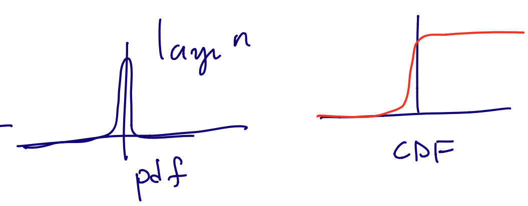
$$\mathbb{P}(|X_n - X| > \epsilon) \to 0 \quad \text{as } n \to \infty.$$

<u>Ex:</u> $X_n = X + \frac{1}{n} \underbrace{Z}_{N(0,1)}$.

② X_n converges <u>in distribution</u> to X, $X_n \rightsquigarrow X$, if $\lim_{n \to \infty} F_n(t) = F(t)$, for all t for which F is continuous.

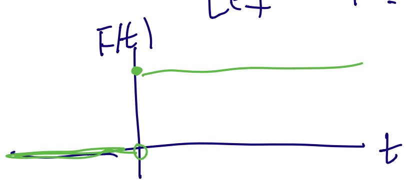
③ X_n converges to X in <u>quadratic mean</u> (convergence in L_2), $X_n \xrightarrow{qm} X$ if

$$\mathbb{E}\left((X_n - X)^2 \right) \to 0.$$



<u>Example</u> Let $X_n \sim N(0, \frac{1}{n})$

Let F = CDF for a point mass at 0.



$\mathbb{P}(X \ge 0) = 1 \qquad \mathbb{P}(X > 0) = 0$

$\mathbb{P}(X < 0) = 0 \qquad = 1 - \mathbb{P}(X \le 0)$

For $t < 0$:

$$F_n(t) = \mathbb{P}(X_n \le t)$$

$$= \mathbb{P}(\sqrt{n}X_n \le \sqrt{n}\,t)$$

$$= \mathbb{P}(\underset{\substack{\uparrow \\ N(0,1)}}{Z} \le \sqrt{n}\,t) \to 0 \quad \text{since } \sqrt{n}\,t \to -\infty.$$

For $t > 0$: $F_n(t) = \mathbb{P}(Z \le \sqrt{n}\,t) \to 1$ since $\sqrt{n}\,t \to \infty$.

$\Rightarrow \lim_{n\to\infty} F_n(t) = F(t)$ except at $t = 0$.

$\Rightarrow$ $X_n \rightsquigarrow X$ in distribution.

Example

Let $X_n \sim N(0, 1/n)$.

For any $\epsilon > 0$,

$$\mathbb{P}(|X_n - 0| > \epsilon) = \mathbb{P}(|X_n|^2 > \epsilon^2)$$

use Markov's Inequality

$$\le \frac{\mathbb{E}(X_n^2)}{\epsilon^2} = \frac{1/n}{\epsilon^2}$$

$$= \frac{1}{n\epsilon^2} \to 0 \quad \text{as } n \to \infty$$

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

Thm (a) If $X_n \xrightarrow{qm} X$, then $X_n \xrightarrow{P} X$

(b) If $X_n \xrightarrow{P} X$, then $X_n \rightsquigarrow X$.

(c) If $X_n \rightsquigarrow X$ and $\mathbb{P}(X = c) = 1$ for some real c, then $X_n \xrightarrow{P} X$.

In general, the converse of these statements is <u>not true</u>.

If $\mathbb{P}(X=c)=1$.

Convergence flow chart:

$$\text{quadratic mean} \longrightarrow \text{probability} \longrightarrow \text{distribution}$$

<u>Thm</u> (5.5 in AoS). Let X_n, X, Y_n, Y be r.v.'s, and g a continuous function.

(a) If $X_n \xrightarrow{P} X$, $Y_n \xrightarrow{P} Y$, then $X_n + Y_n \xrightarrow{P} X + Y$.

(b) If $X_n \xrightarrow{qm} X$, $Y_n \xrightarrow{qm} Y$, then $X_n + Y_n \xrightarrow{qm} X + Y$.

* (c) If $X_n \rightsquigarrow X$ and $Y_n \rightsquigarrow c$, then $X_n + Y_n \rightsquigarrow X + c$.

(d) If $X_n \xrightarrow{P} X$, $Y_n \xrightarrow{P} Y$, then $X_n Y_n \xrightarrow{P} XY$.

* (e) If $X_n \rightsquigarrow X$, $Y_n \rightsquigarrow c$, then $X_n Y_n \rightsquigarrow cX$.

(f) $X_n \xrightarrow{P} X$, then $g(X_n) \xrightarrow{P} g(X)$.

(g) $X_n \rightsquigarrow X$, then $g(X_n) \rightsquigarrow g(X)$.

These are known as Slutsky's Theorem.

<u>Laws of Large Numbers</u>

<u>Weak Law of Large Numbers</u> (WLLN): Let $X_1, \ldots, X_n$ be IID random variables with finite mean and variance. Then

$$\bar{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i \xrightarrow{P} \mu = \mathbb{E}(X_i).$$

$$\Rightarrow \quad \mathbb{P}(|X_n - \mu| > \epsilon) \to 0 \quad \text{as } n \to \infty.$$

__Strong Law of Large Numbers (SLLN)__: Let $X_1, \dots, X_n$ be IID r.v.s with finite mean and variance. Then

$$\bar{X}_n \xrightarrow{a.s.} \mu \; :$$

$$\mathbb{P}\left(\lim_{n\to\infty} |\bar{X}_n - \mu| < \epsilon\right) = 1 \quad \text{for any } \epsilon > 0.$$

almost sure convergence.

__Central Limit Theorem__ Let $X_1, \dots, X_n$ be IID r.v.'s with mean $\mu < \infty$ and variance $\sigma^2 < \infty$.

Then

$$Z_n = \frac{\bar{X}_n - \mu}{\sqrt{Var(\bar{X}_n)}} = \frac{\bar{X}_n - \mu}{\sigma/\sqrt{n}} \rightsquigarrow Z, \text{ with } Z \sim N(0,1).$$

Other forms:

$$Z_n \approx N(0,1)$$

$$\bar{X}_n \approx N(\mu, \sigma^2/n)$$

$$\bar{X}_n - \mu \approx N(0, \sigma^2/n)$$

__Proof__

Moment generating function:

$$M_X(t) = \mathbb{E}(e^{tX}).$$

__Lemma__ If $\lim_n M_{X_n}(t) = M_X(t)$ for all t, then $\lim_n F_{X_n}(t) = F_X(t)$ for all t at which F_X is continuous.

Assume that $\mu = 0$, $\sigma^2 = 1$.

$$\Rightarrow \quad \frac{\bar{X}_n - \mu}{\sigma/\sqrt{n}} = \frac{\frac{1}{n}\sum X_i}{1/\sqrt{n}} = \underline{\frac{1}{\sqrt{n}}\sum X_i}$$

Let $M(t) = E(e^{tX_i})$,

then $\mathbb{E}\left(e^{tX_i/\sqrt{n}}\right) = M(t/\sqrt{n})$

and $\mathbb{E}\left(e^{t\sum X_i/\sqrt{n}}\right) = \left(M(t/\sqrt{n})\right)^n$.

To simplify things, introduce $L(t) = \log M(t)$

<u>Note</u> $L(0) = \log M(0) = \log 1 = 0$

$$L'(t) = \frac{1}{M(t)} M'(t) \Rightarrow L'(0) = \frac{M'(0)}{M(0)} = E(X_i) = 0$$

$$L''(t) = \frac{M M'' - (M')^2}{M^2}$$

$$\Rightarrow L''(0) = \frac{M(0)M''(0) - M'(0)^2}{M(0)^2}$$

$$= \frac{1 \cdot \mathbb{E}(X_i^2) - 0}{1} = 1$$

To prove the CLT, show that

$$\lim_{n\to\infty} \left(M(t/\sqrt{n})\right)^n = e^{t^2/2}$$

MGF of $N(0,1)$ r.v.

or equivalently:

$$\lim_{n\to\infty} n\, L(t/\sqrt{n}) = t^2/2$$

Compute the limit directly:

$$\lim_{n\to\infty} n L(t/\sqrt{n}) = \lim \frac{L(t/\sqrt{n})}{1/n}$$

by L'Hopital

$$= \lim \frac{-L'(t/\sqrt{n})\,\frac{1}{2}\,t/n^{3/2}}{-1/n^2}$$

$$= \lim \frac{t\,L'(t/\sqrt{n})}{2/\sqrt{n}}$$

by L'Hopital

$$= \lim \frac{+t\,L''(t/\sqrt{n})\,\frac{1}{2}\,\frac{1}{\cancel{n^{3/2}}}\,t}{+\cancel{2}\cdot\cancel{\tfrac{1}{2}}\cdot 1/\cancel{n^{3/2}}}$$

$$= \lim \frac{t^2}{2} L''(t/\sqrt{n})$$

$$= \frac{t^2}{2} L''(0) = \frac{t^2}{2}$$

<u>The Delta Method</u> If Y_n has limiting normal distribution, then we can find the limiting distribution of $g(Y_n)$, when g is any smooth function.

<u>Theorem</u> Suppose that $\dfrac{Y_n - \mu}{\sigma/\sqrt{n}} \rightsquigarrow N(0,1)$, and g is differentiable such that $g'(\mu) \neq 0$. Then,

$$\frac{g(Y_n) - g(\mu)}{|g'(\mu)|\,\sigma/\sqrt{n}} \rightsquigarrow N(0,1).$$

I.e. $Y_n \approx N(\mu, \sigma^2/n)$

$$\Rightarrow \quad g(Y_n) \simeq N\left(g(\mu),\ g'(\mu)^2\, \sigma^2/n\right)$$

<u>Sketch of Proof</u>

Expand g in Taylor series about μ:

$$g(x) = g(\mu) + g'(\mu)(x-\mu) + \tfrac{1}{2} g''(\mu_0)(x-\mu)^2.$$

$$\Rightarrow \quad g(Y_n) = g(\mu) + g'(\mu)(Y_n-\mu) + \tfrac{1}{2} g''(\mu_0)(Y_n-\mu)^2.$$

Motivation

$$\Rightarrow \quad \frac{g(Y_n) - g(\mu)}{g'(\mu)} = (Y_n - \mu) + \frac{1}{2}\frac{g''(\mu_0)}{g'(\mu)}(Y_n-\mu)^2.$$

Since we assumed $\dfrac{Y_n - \mu}{\sigma/\sqrt{n}} \rightsquigarrow N(0,1) \Rightarrow Y_n - \mu \rightsquigarrow N(0, \sigma^2/n)$

$Y_n \rightsquigarrow N(\mu, \sigma^2/n)$.

It can be shown that this implies that

$Y_n \xrightarrow{P} \mu$.

<u>Ex:</u> If $Y_n = \bar{X}_n = \frac{1}{n}\sum X_i$, by CLT $Y_n \approx N(\mu, \sigma^2/n)$,

by the WLLN, $Y_n = \bar{X}_n \xrightarrow{P} \mu$.

And then, it can be shown that

$$R_n = \tfrac{1}{2} g''(\mu_0)(Y_n-\mu)^2 \xrightarrow{P} 0.$$

Rearranging

$$\sqrt{n}\left(g(Y_n)-g(\mu)\right) = g'(\mu)\overbrace{\sqrt{n}\,(Y_n-\mu)}^{\rightsquigarrow N(0,\sigma^2)}$$

$$+\ g''(\mu_0)\underbrace{\sqrt{n}\,(Y_n-\mu)}_{\rightsquigarrow N(0,\sigma^2)}\overbrace{(Y_n-\mu)}^{\xrightarrow{P} 0}.$$

Invoking Slutsky's Theorem:

$$\Rightarrow \sqrt{n}\left(g(Y_n)-g(\mu)\right) \rightsquigarrow N(0, g'(\mu)^2\sigma^2).$$

$$\Rightarrow \frac{g(Y_n)-g(\mu)}{|g'(\mu)|\sigma/\sqrt{n}} \rightsquigarrow N(0,1).$$

Simple example: $g(x) = a + bx.$

$$X \sim N(\mu,\sigma^2)$$

$$Y = g(X) = a + bX$$

$$E(Y) = E(a+bX) = a + b\mu.$$

$$Var(Y) = Var(a+bX) = b^2\sigma^2$$

$g'(x) = b.$

$$\Rightarrow Y \sim N(a+b\mu,\ b^2\sigma^2)$$

$$\sim N(g(\mu),\ (g'(\mu))^2\sigma^2).$$