$$Y = T_n(X_1, X_2, X_3, \ldots)$$

$\rightarrow$ sample mean: $\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$

sample variance: $S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X})^2$

What is the mean and variance $\bar{X}$?

$$\mathbb{E}[\bar{X}] = \mu = \mathbb{E}[X]$$

$$\mathbb{V}[\bar{X}] = \frac{\sigma^2}{n} \qquad \sigma^2 = \mathbb{E}[(X-\mu)^2]$$

Probability Inequalities

$$|X - \mu| > \epsilon, \quad \epsilon > 0$$

$$P(|X-\mu| < \epsilon) = 0 \text{ or } 1?$$

Theorem (Markov's Inequality)

Let Y be a non-negative RV. Then, for any $r>0$,

$$\boxed{\mathbb{P}(Y \geq r) \leq \frac{\mathbb{E}[Y]}{r}}$$

Proof: Assume Y is a continuous RV with pdf $p(y)$.

$$\mathbb{E}[Y] = \int_0^\infty y\, p(y)\, dy$$

$$= \int_0^h y\, p(y)\, dy + \int_h^\infty y\, p(y)\, dy$$

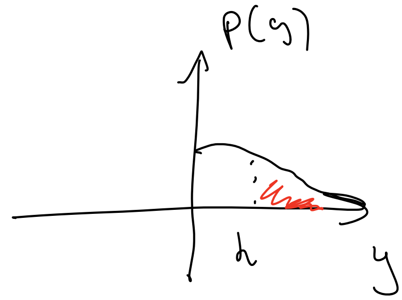
$$\geq \int_h^\infty y\, p(y)\, dy$$

$$\geq \int_h^\infty h\, p(y)\, dy = h \underbrace{\int_h^\infty p(y)\, dy}_{P(Y \geq h)}$$

$$= h\, P(Y \geq h)$$

$$\mathbb{E}[Y] \geq h\, P(Y \geq h)$$

$$P(Y \geq h) \leq \frac{\mathbb{E}[Y]}{h}$$



Theorem (Chebychev's Inequality)

Let Y be a RV with finite non-zero variance σ^2. Then, for any $k > 0$

$$\mathbb{P}(|Y-\mu| \geq k\sigma) \leq \frac{1}{k^2}$$

Proof

$$\mathbb{P}(|Y-\mu| \geq k\sigma) = \mathbb{P}\left((Y-\mu)^2 \geq k^2\sigma^2\right)$$

$$\leq \frac{\mathbb{E}[(Y-\mu)^2]}{k^2\sigma^2}$$

$$= \frac{1}{k^2} \quad \blacksquare$$

Moment generating function (MGF)

For a random variable X:

$$M_X(t) = \mathbb{E}\left[e^{tX}\right]$$

$$\left.\frac{d M_X^{(k)}(t)}{dt}\right|_{t=0} = \mathbb{E}\left[X^k\right]$$

Theorem (Chernoff Bounds)

Let Y be a RV with MGF $M_Y(t)$ where $|t| < h$. Then, for any a

$$\boxed{\mathbb{P}(Y \geq a) \leq e^{-at} M_Y(t), \quad 0 < t < h}$$

and

$$\mathbb{P}(Y \leq a) \leq e^{-at} M_Y(t)$$
$$-h < t < 0$$

Proof

$$P(Y \geq a) = P(Yt \geq ta) \qquad t>0$$

$$= P(e^{Yt} \geq e^{ta})$$

$$\leq \frac{\mathbb{E}[e^{Yt}]}{e^{ta}} = e^{-ta} M_Y(t)$$

$$P(Y \geq a) \leq \min_{t \in [0,h]} e^{-ta} M_Y(t)$$

---

1. WLLN $\longrightarrow$ 2. SLLN $\longrightarrow$ 3. CLT

Sample mean

$$\bar{X}_1 = X_1, \quad \bar{X}_2 = \frac{1}{2}\sum_{i=1}^{2} X_i, \quad \dots$$

$$\bar{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i, \quad \dots \longrightarrow \bar{X}_\infty$$

<u>Definition</u> (Convergence in Probability) A sequence of random variables $X_1, X_2, \ldots$ <u>converges in probability</u> to a random variable X if for every $\epsilon > 0$

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

OR

$$\lim_{n \to \infty} \mathbb{P}(|X_n - X| \leq \epsilon) = 1$$

$X_n \xrightarrow{P} X$.

<u>Theorem</u> (Weak Law of Large Numbers)
Let $X_1, X_2, \ldots$ be iid random variables such that $\mathbb{E}[X_i] = \mu$ and $\mathbb{V}[X_i] = \sigma^2 < \infty$. Define $\overline{X}_n = \frac{1}{n} \sum_{i=1}^{n} X_i$. Then, $\overline{X}_n \xrightarrow{P} \mu$.

$$\overline{X^2} = \frac{1}{n} \sum_{i=1}^{n} X_i^2 \xrightarrow{p} \mathbb{E}[X^2]$$

Yes, as long as $\mathbb{E}[X_i^2] = \mathbb{E}[X^2]$

and

$\mathbb{V}[X_i^2] < \infty$

<u>Proof.</u> $\mathbb{P}(|\bar{X}_n - \mu| \geq \epsilon) = \mathbb{P}((\bar{X}_n - \mu)^2 \geq \epsilon^2)$

$$\leq \frac{\mathbb{E}\left[(\bar{X}_n - \mu)^2\right]}{\epsilon^2} \quad \Big\} \, \mathbb{V}[\bar{X}] = \sigma^2/n$$

$$= \frac{\sigma^2}{n\epsilon^2}$$

$$\lim_{n\to\infty} \mathbb{P}(|\bar{X}_n - \mu| \geq \epsilon) \leq \underbrace{\lim_{n\to\infty} \frac{\sigma^2}{n\epsilon^2}}_{0}$$

$$\Rightarrow \lim_{n\to\infty} \mathbb{P}(|\bar{X}_n - \mu| \geq \epsilon) = 0$$

Definition (Consistent Estimator)

Let $\hat{\theta}$ be an estimator of θ s.t. $\hat{\theta} \xrightarrow{P} \theta$. Then, $\hat{\theta}$ is said to be a consistent estimator of θ.

Example: Consistency of Sample Variance

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

$$\mathbb{E}[S_n^2] = \sigma^2$$

$$\mathbb{P}(|S_n^2 - \sigma^2| \geq \epsilon) = \mathbb{P}((S_n^2 - \sigma^2)^2 \geq \epsilon^2)$$

$$\leq \frac{\mathbb{E}[(S_n^2 - \sigma^2)^2]}{\epsilon^2} \quad \mathbb{V}[S_n^2]$$

Sufficient condition

$\boxed{\text{If } \mathbb{V}[S_n^2] \to 0 \text{ as } n \to \infty}$, then

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

$$\sqrt{S_n^2} \xrightarrow{P} \sigma \quad ???$$

Theorem (Continuous Mapping Theorem)

Suppose $X_1, X_2, \ldots$ converge in probability to X. Let h be a continuous function. Then,

$$h(X_1), h(X_2), \ldots \xrightarrow{P} h(X)$$

Remark: h can have discontinuities but

$$\mathbb{P}(X \in \underbrace{D_h}_{\text{set of discontinuities}}) = 0$$

SLLN

Definition (Almost Sure Convergence)
A sequence of RVs $X_1, X_2, \ldots$ converges almost surely to a random variable X if for every $\epsilon > 0$:

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

$X_n \xrightarrow{a.s.} X$

Remark: Convergence in Probability

$$\lim_{n\to\infty} \mathbb{P}(|X_n - X| < \epsilon) = 1$$

$X : \Omega \to \mathbb{R}$: X_∞

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

X_∞

$\lim_{n\to\infty} |X_n - X| < \epsilon$

$P(\quad) = 1$

$\lim_{n\to\infty} |X_n - X| \geq \epsilon$

$P(\quad) = 0$

Theorem

Let $X_1, X_2, \ldots$ be a sequence of RVs such that $X_n \xrightarrow{a.s.} X$. Then, $X_n \xrightarrow{P} X$.

Theorem (Strong Law of Large Numbers)

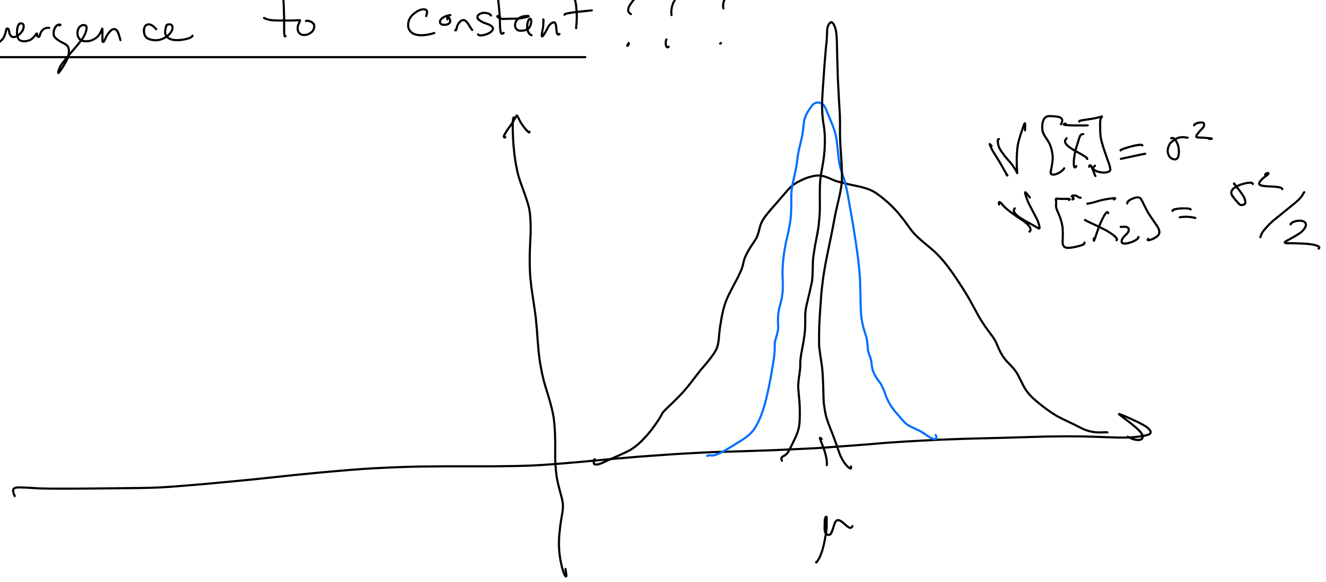
Let $X_1, X_2, \dots$ be i.i.d. random variables such that $\mathbb{E}[X_i] = \mu$ and $\mathbb{V}[X_i] = \sigma^2 < \infty$. Define $\bar{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i$. Then, for every $\epsilon > 0$:

$$\mathbb{P}\left(\lim_{n\to\infty} |\bar{X}_n - \mu| < \epsilon\right) = 1$$

or

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

Convergence to constant ???



Point mass OR Delta function OR impulse

OR Dirac measure

Random variable s.t.

$$\mathbb{P}(X = \mu) = 1$$

otherwise

$$\mathbb{P}(X \neq \mu) = 0$$

$$\delta_\mu(x) = \begin{cases} \infty, & x = \mu \\ 0, & \text{o.w.} \end{cases}$$

"sifting property"

$$\int f(x)\, \delta_\mu(x)\, dx = f(\mu)$$

CLT

<u>Definition</u> (Convergence in distribution)

A sequence of random variables $X_1, X_2, \ldots$ is said to converge in distribution to X if

$$\lim_{n\to\infty} \underbrace{F_{X_n}(x)}_{\mathbb{P}(X_n \le x)} = \underbrace{F_X(x)}_{\mathbb{P}(X \le x)}$$

We say $X_n \xrightarrow{d} X$.

- Weakest form of convergence → implied by convergence in probability

$$X_n \xrightarrow{a.s.} X \Rightarrow X_n \xrightarrow{p} X \Rightarrow X_n \xrightarrow{d} X$$

<u>Theorem</u>

The sequence of random variables $X_1, X_2, \ldots$ converges in probability to a constant μ <u>if and only if</u> the sequence converges in distribution to μ.

Theorem (Central limit theorem)

Let $X_1, X_2, \dots$ be a sequence of i.i.d. RVs whose MGFs exists. Let $\mathbb{E}[X_i] = \mu$ and let $\mathbb{V}[X_i] = \sigma^2$. Define $\bar{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i$. Let $G_n(x)$ denote the CDF of the RV:

$$\frac{\sqrt{n}(\bar{X}_n - \mu)}{\sigma}$$

Then for any x s.t. $-\infty < x < \infty$:

$$\lim_{n\to\infty} G_n(x) = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} e^{-y^2/2} \, dy$$

That is $\frac{\sqrt{n}(\bar{X}_n - \mu)}{\sigma} \xrightarrow{d} Z \sim N(0,1)$

$X_1, \dots, X_n$

$$\bar{X}_1 = X_1, \quad \bar{X}_2 = \frac{1}{2}\sum_{i=1}^{2} X_i, \quad \dots, \quad \bar{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i$$

Theorem (Convergence of MGFs)

Let $X_1, X_2, \dots$ be a sequence of RVs whose MGF exists and let X be another RV whose MGF also exists. Then, if

$$\lim_{n \to \infty} M_{X_n}(t) = M_X(t)$$

then

$$\lim_{n \to \infty} F_{X_n}(t) = F_X(t)$$

That is, $X_n \xrightarrow{d} X$.

The MGF of a standard Gaussian $Z \sim N(0,1)$

$$M_Z(t) = e^{t^2/2}$$

Proof (CLT)

First, lets define $Y_n = \left(\frac{X_n - \mu}{\sigma}\right) \Rightarrow \begin{array}{l} \mathbb{E}[Y_n] = 0 \\ \mathbb{V}[Y_n] = 1 \end{array}$

We now define $Z_n = \frac{\sqrt{n}\,(\bar{X}_n - \mu)}{\sigma}$

$$Z_n = \sqrt{n}\,\frac{(\bar{X}_n - \mu)}{\sigma} = \frac{\sqrt{n}\left(\left(\frac{1}{n}\sum_{i=1}^{n} X_i\right) - \mu\right)}{\sigma}$$

$$= \sqrt{n}\left(\frac{1}{n}\sum_{i=1}^{n}\frac{(X_i - \mu)}{\sigma}\right)$$

$$= \frac{1}{\sqrt{n}}\sum_{i=1}^{n} Y_i$$

$$M_{Z_n}(t) = M_{\frac{1}{\sqrt{n}}\sum_{i=1}^{n} Y_i}(t) = \mathbb{E}\left[e^{t\left(\frac{1}{\sqrt{n}}\sum_{i=1}^{n} Y_i\right)}\right]$$

$$= \mathbb{E}\left[\prod_{i=1}^{n} e^{\left(\frac{t}{\sqrt{n}}\right) Y_i}\right]$$

$$= \prod_{i=1}^{n} \mathbb{E}\left[e^{\left(\frac{t}{\sqrt{n}}\right) Y_i}\right] \quad \text{(by independence)}$$

$$Y = \frac{X-\mu}{\sigma}$$

$$= \prod_{i=1}^{n} \underbrace{\mathbb{E}\left[e^{\left(\frac{t}{\sqrt{n}}\right)Y}\right]}_{M_Y\left(\frac{t}{\sqrt{n}}\right)}$$

$$= \Bigg(\underbrace{M_Y\left(\frac{t}{\sqrt{n}}\right)}_{\text{Taylor series expansion!!}} \Bigg)^n$$

$$M_Y\left(\frac{t}{\sqrt{n}}\right) = \sum_{k=0}^{\infty} \underbrace{\left(\left.\frac{d^k M_Y\left(\frac{t}{\sqrt{n}}\right)}{dt^k}\right|_{t=0} \right) \frac{(t/\sqrt{n})^k}{k!}}$$

$$= \underbrace{1}_{k=0} + \underbrace{0}_{k=1} + \underbrace{\frac{(t/\sqrt{n})^2}{2}}_{k=2}(1) + R_Y\left(\frac{t}{\sqrt{n}}\right)$$

$$M_{Z_n}(t) = \left(1 + \frac{(t/\sqrt{n})^2}{2} + R_Y\left(\frac{t}{\sqrt{n}}\right)\right)^n$$

$$= \left(1 + \frac{1}{n}\left(\frac{t^2}{2} + nR_Y\left(\frac{t}{\sqrt{n}}\right)\right)\right)^n$$

$$\approx \left(1 + \frac{t^2}{2n}\right)^n$$

$$\lim_{n\to\infty}\left(1+\frac{a}{n}\right)^n = e^a$$

$$\lim_{n\to\infty}\left(1 + \frac{t^2}{2n}\right)^n = e^{t^2/2}$$

$$\lim_{n\to\infty} M_{Z_n}(t) = e^{t^2/2} = M_Z(t)$$

$Z \sim N(0, 1)$. This implies $Z_n \xrightarrow{d} Z$. $\square$

<u>Theorem (Slutsky's Theorem)</u>

If $X_n \xrightarrow{d} X$ and $Y_n \xrightarrow{P} a$, then:

- $X_n Y_n \xrightarrow{d} aX$
- $X_n + Y_n \xrightarrow{d} X + a$

Discussion: If σ is unknown:

→ we can replace σ in the CLT equation with a consistent estimator

$$\hat{\sigma} = \sqrt{S_n^2}$$

Theorem (Delta Method)

Let Y_n be a sequence of RVs that satisfies.

$$\sqrt{n}\,(Y_n - \theta) \xrightarrow{d} N(0, \sigma^2)$$

For a given function h whose derivative exists, and we denote it by h'. Then,

$$\sqrt{n}\,\big(h(Y_n) - h(\theta)\big) \xrightarrow{d} N\big(0, \sigma^2 (h'(\theta))^2\big)$$