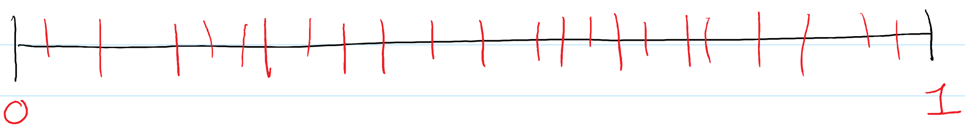
Ex. $X_1, X_2, X_3, \ldots \overset{iid}{\sim} U(0,1)$

$$Y_n = \max_{i=1,\ldots,n} X_i$$

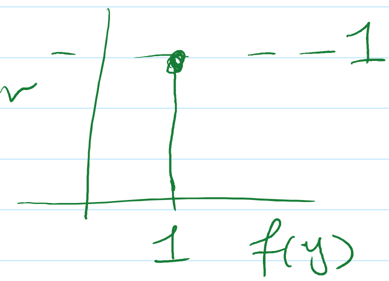


Intuition: $Y_n \approx 1$ for large n.

① $\boxed{Y_n \xrightarrow{P} 1}$ $\quad (X_n \xrightarrow{P} X)$

degenerate r.v. Y when

$\mathbb{P}(Y=1) = 1$



Recall: $Y_n \xrightarrow{P} Y$ means $\forall \varepsilon > 0$

$$\underbrace{\mathbb{P}(|Y_n - Y| \geq \varepsilon)}_{} \xrightarrow{n} 0$$

$$\mathbb{P}(|Y_n - 1| \geq \varepsilon) = \mathbb{P}(|1 - Y_n| \geq \varepsilon) \qquad 1 \geq X_n \geq 0$$

$$= \mathbb{P}(1 - Y_n \geq \varepsilon) \qquad 1 \geq Y_n \geq 0$$

$$= \mathbb{P}(Y_n \leq 1 - \varepsilon) \qquad Y_n = \max_{i=1,\ldots,n} X_n$$

$$= \mathbb{P}(X_1 \leq 1-\varepsilon, X_2 \leq 1-\varepsilon, \ldots, X_n \leq 1-\varepsilon)$$

$$= \prod^{n} \mathbb{P}(X_i \leq 1 - \varepsilon)$$

$$= \prod_{i=1}^{n} P(X_i \le 1-\varepsilon)$$

Since $X_i \sim U(0,1)$ then $\underbrace{P(X_i \le x) = x}_{\text{CDF of } U(0,1)}$

$$= \prod_{i=1}^{n} (1-\varepsilon)$$

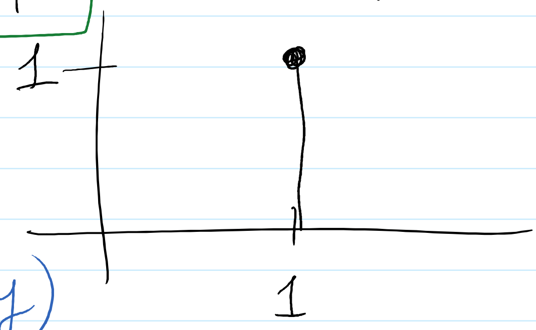
$$\boxed{P(|Y_n - 1| \ge \varepsilon) = (1-\varepsilon)^n} \quad (\text{for } \varepsilon \le 1)$$

So $P(|Y_n - 1| \ge \varepsilon) \xrightarrow{n} 0$

By defn $\boxed{Y_n \xrightarrow{p} 1.}$

$\boxed{Y_n \xrightarrow{d} 1}$ $\quad$ $\boxed{F_{Y_n} \xrightarrow{\text{ptwise}} F_Y}$

pmf



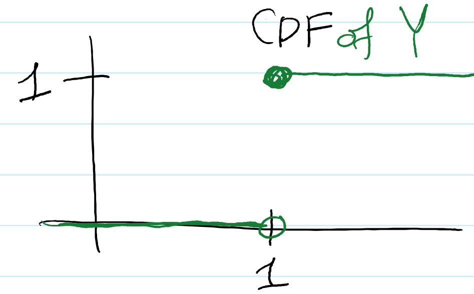
$$F_{Y_n}(y) = P(Y_n \le y)$$

$$= P(X_1 \le y, X_2 \le y, \dots, X_n \le y)$$

$$= \prod_{i=1}^{n} P(X_i \le y)$$

$$\boxed{F_{Y_n}(y) = \begin{cases} 0 & y < 0 \\ y^n & 0 \le y \le 1 \\ 1 & y > 1 \end{cases}}$$

CDF of Y



Step fn at 1

as $n \to \infty$

$$\underline{F_{Y_n}(y)} \xrightarrow{n} \begin{cases} 0 & y < 0 \\ 0 & 0 \le y \le 1 \\ 1 & y > 1 \end{cases} \qquad \text{b/c } y^n \to 0 \text{ as } n \to \infty \text{ for } y \in [0,1)$$

$$= \begin{cases} 0 & y < 1 \\ 1 & y \ge 1 \end{cases}$$

$$= \underline{\text{CDF of } Y}$$

So by defn $Y_n \xrightarrow{d} 1$.

Theorem: Convergence in prob. implies convergence in dist.

If $X_n \xrightarrow{P} X$ then $X_n \xrightarrow{d} X$.

Chain: The converse implications generally are not true.

$$\text{a.s.} \Rightarrow \text{i.p.} \Rightarrow d$$

almost sure   in prob.   in dist.

Revisit our example.

$$\boxed{Z_n = n(1 - Y_n)} \leftarrow \text{distributional limit?}$$

$$Y_n = \max_{i=1,\dots,n} X_n \quad \text{and} \quad X_i \overset{iid}{\sim} U(0,1)$$

$$F_{Z_n}(z) = \mathbb{P}(Z_n \le z) = \mathbb{P}(n(1-Y_n) \le z)$$
$$= \mathbb{P}(Y_n \ge 1 - z/n)$$
$$= 1 - \mathbb{P}(Y_n \le 1 - z/n)$$
$$\rightarrow \quad = 1 - \prod_{i=1}^{n} \underbrace{\mathbb{P}(X_i \le 1 - z/n)}$$

recall $X_i \overset{iid}{\sim} U(0,1)$

$$= 1 - \begin{cases} (1 - z/n)^n & 0 < z/n < 1 \\ 0 & z/n \ge 1 \\ 1 & z/n \le 0 \end{cases}$$

$$F_{Z_n}(z) = \begin{cases} 1 - (1 - z/n)^n & 0 < z/n < 1 \\ \cancel{1 \quad z/n \ge 1} \\ 0 & z/n \le 0 \end{cases}$$

What is the limit of our CDF?

$$\boxed{F_{Z_n}(z) \longrightarrow \begin{cases} 1 - e^{-z} & z > 0 \\ 0 & z \le 0 \end{cases}}$$

Recall: $\displaystyle e^c = \lim_{n \to \infty} \left(1 + \frac{c}{n}\right)^n$

So $\displaystyle \lim_{n \to \infty} 1 - (1 - z/n)^n = e^{-z}$

$\rightarrow$ Recognize as the CDF of a

→ Recognize as the CDF of a $\text{Exp}(1)$

So $Z_n = n(1 - Y_n) \xrightarrow{d} \text{Exp}(1)$

Theorem:

Know $\underline{a.s.} \Rightarrow \underline{i.p} \Rightarrow \underline{d}$

partial converse:

If $X_n \xrightarrow{d} c$ then $X_n \xrightarrow{P} c$.

where c is a constant (i.e. a degenerate distribution)

Algebraic Properties

If $X_n \to X$ and $Y_n \to Y$

(i) $aX_n + bY_n \to aX + bY$

(ii) $X_nY_n \to XY$

this is true for
1. almost sure convergence
2. convergence in prob.

not generally true for convergence in dist.

Slutsky's Theorem (algebraic props for

Slutsky's Theorem (algebraic props for conv. in dist)

If $X_n \xrightarrow{d} X$ and $Y_n \xrightarrow{p} c$ then

1. $X_n + Y_n \xrightarrow{d} X + c$
2. $X_n + Y_n \xrightarrow{d} Xc$
3. $X_n / Y_n \xrightarrow{d} X/c$

Theorem: Continuous Mapping Theorem

If g is a continuous fn $\mathbb{R} \to \mathbb{R}$

and $X_n \to X$ (by any mode)

then $g(X_n) \to g(X)$. For any type of convergence

Aside: defn of cts fn g is that

$\forall\ x_n \to x$ then $g(x_n) \to g(x)$

Intuition: $X_i \overset{iid}{\sim} f$ and $\mathbb{E}[X_i] = \mu$

then $\overline{X}_n = \frac{1}{n} \sum_{i=1}^{n} X_i$

Should be a good estimator of μ.

$\bar{X}_n \approx \mu$ for large n.

Theorem: Weak Law of Large Numbers (WLLN)

If X_i that are uncorrelated and

$$\left.\begin{array}{l} \text{① } E[X_i] = \mu \\ \text{② } Var(X_i) < \infty \end{array}\right\}$$

then $\boxed{\bar{X}_n \xrightarrow{P} \mu.}$

(called weak b/c conv. in prob)

---

pf. $X_i \overset{iid}{\sim} f$ and $E[X_i] = \mu$; $Var(X_i) = \sigma^2 < \infty$

① $E[\bar{X}_n] = E\left[\frac{1}{n}\sum_{i=1}^{n} X_i\right]$

$$= \frac{1}{n}\sum_{i=1}^{n} E(X_i] = \frac{1}{n}\sum_{i=1}^{n}\mu = \frac{1}{n} n \cdot \mu = \mu$$

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

$$= \frac{1}{n^2}\sum_{i=1}^{n} Var(X_i) = \frac{1}{n} n\sigma^2 = \frac{\sigma^2}{n}$$

Chebyshev:

$$\boxed{\mathbb{P}(|X - \mathbb{E}[X]| \geq \varepsilon) \leq \frac{Var(X)}{\varepsilon^2}}$$

Apply to $\bar{X}_n$

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

$Var(\bar{X}_n) = \sigma^2/n$

$$\boxed{\mathbb{P}(|\bar{X}_n - \mu| \geq \varepsilon) \leq \frac{\sigma^2}{n\varepsilon^2}}$$

notice as $n \to \infty$ $\frac{\sigma^2}{n\varepsilon^2} \longrightarrow 0$

$\boxed{\text{Can show } \bar{X}_n \xrightarrow{P} \mu.}$

$$\forall \varepsilon > 0 \quad \mathbb{P}(|\bar{X}_n - \mu| \geq \varepsilon) \xrightarrow{n} 0$$

b/c $\mathbb{P}(|\bar{X}_n - \mu| \geq \varepsilon) \leq \frac{\sigma^2}{n\varepsilon^2} \xrightarrow{n} 0$

Theorem: Strong Law of Large Numbers (SLLN)

If X_i are iid w/ $\mathbb{E}[X_i] = \mu$ and $Var(X_i) = \sigma^2 < \infty$ then

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

Theorem: Central Limit Theorem

If X_i are iid $E[X_i]=\mu$ and $Var(X_i)=\sigma^2<\infty$

then

$$\boxed{\sqrt{n}\left(\frac{\bar{X}_n-\mu}{\sigma}\right) \xrightarrow{d} N(0,1).}$$

Intro stats:

$$\boxed{\bar{X}_n \approx N(\mu, \sigma^2/n) \quad \text{for large } n}$$

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

$$\sqrt{n}\left(\frac{\bar{X}_n-\mu}{\sigma}\right) \sim N(0,1)$$

Ex CLT

$X_i \overset{iid}{\sim} \text{Bernoulli}(p)$

$E[X_i]=p$ and $Var(X_i)=p(1-p)$

CLT says:

$$\sqrt{n}\left(\frac{\bar{X}_n-p}{\sqrt{p(1-p)}}\right) \xrightarrow{d} N(0,1)$$

Cheat: $\sqrt{n}\left(\frac{\bar{X}_n-p}{\sqrt{p(1-p)}}\right) \approx N(0,1)$

then $\bar{X} \approx N\left(p, \frac{p(1-p)}{n}\right)$

[illegible]

Estimate a proportion

$$\hat{p} = \bar{X}_n \approx N\left(p, \frac{p(1-p)}{n}\right)$$

$$\text{CI}: \quad \hat{p} \pm 2\sqrt{\frac{p(1-p)}{n}}$$

Theorem: Taylor's Theorem

$g: \mathbb{R} \to \mathbb{R}$ that is k-times differentiable

then the k^{th}-order Taylor polynomial about a is

$$\boxed{T_k(x) = \sum_{r=0}^{k} \frac{g^{(r)}(a)}{r!}(x-a)^r}$$

Call $R = g(x) - T_k(x) \to 0$ as $x \to a$
the remainder quickly

So for $\boxed{x \approx a}$ $g(x) \approx T_k(x)$.

CLT: $X_i \overset{iid}{\sim}$ w/ $E[X_i] = \mu$ and

$Var(X_i) = \sigma^2 < \infty$

then $\bar{X} = \frac{1}{n}\sum^{n} X_i$

then $\bar{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i$

$$\underbrace{\sqrt{n}\left(\frac{\bar{X}_n - \mu}{\sigma}\right)}_{Y} \xrightarrow{d} N(0,1)$$

Pf.

$$\boxed{Y = \sqrt{n}\left(\frac{\bar{X}_n - \mu}{\sigma}\right) = \frac{1}{\sqrt{n}}\sum_{i=1}^{n} Y_i}$$

; $Y_i = \frac{X_i - \mu}{\sigma}$ → $E[Y_i] = 0$, $Var(Y_i) = 1$

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

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

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

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

$Y = \frac{1}{\sqrt{n}}\sum_{i=1}^{n} Y_i$ the Y_i are independent, MGF of Y_i

$$\boxed{M_Y(t) = M_{\frac{1}{\sqrt{n}}\sum_{i=1}^{n} Y_i}(t) = M(t/\sqrt{n})^n}$$

$$M_{\sum X_i}(t) = M(t)^n$$
$$M_{aX}(t) = M(at)$$

Taylor second-order expansion of M around zero $(a = 0)$

$$\rightarrow M_{Y_i}(t) = \underbrace{\sum_{r=0}^{2} \frac{M^{(r)}(0)}{r!}(t-0)^r}\ + R(t)$$

$R(t) \to 0$ as $t \to 0$

$$= \frac{(1)(t)^0}{0!} + \frac{M^{(1)}(0)\, t^1}{1!} + \frac{M^{(2)}(0)\, t^2}{2!} + R(t)$$

$$= 1 + \cancel{E[Y_i]}^{\,0}\, t + \frac{\cancel{E[Y_i^2]}^{\,1}\, t^2}{2} + R(t)$$

$$= 1 + t^2/2 + R(t)$$

① $M(t) = 1 + t^2/2 + R(t)$

② $M_Y(t) = M(t/\sqrt{n})^n = \left(1 + \frac{t^2}{2n} + R(t/\sqrt{n})\right)^n$

$$\boxed{Y = \sqrt{n}\left(\frac{\bar{X}_n - \mu}{\sigma}\right)}$$

0 as $n \to \infty$ b/c $t/\sqrt{n} \to 0$

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

$$\boxed{= e^{t^2/2}}$$

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

MGF of a $N(0,1)$

Since MGF of $\sqrt{n}\left(\frac{\bar{X}_n - \mu}{\sigma}\right)$ converges to the MGF of a $N(0,1)$

the MGF of a $N(0,1)$

then the CDFs also converge (*)

So $\sqrt{n}\left(\frac{\bar{X}_n-\mu}{\sigma}\right) \xrightarrow{d} N(0,1)$.