Lecture 06

* $X_i \overset{iid}{\sim} N(0,1)$, $\left.\begin{array}{l}\bar{X} = \frac{1}{n}(X_1+\cdots+X_n) \sim N(0,\frac{1}{n}) \\ \sum_{i=1}^{n}(X_i-\bar{X})^2 \sim \chi^2_{n-1}\end{array}\right]$ indep.

$$X = \begin{bmatrix} X_1 \\ \vdots \\ X_n \end{bmatrix} \sim N\left(\begin{bmatrix} 0 \\ \vdots \\ 0 \end{bmatrix}, \begin{bmatrix} 1 & & 0 \\ & \ddots & \\ 0 & & 1 \end{bmatrix}\right) \overset{d}{=} N(0, I)$$

$$\begin{bmatrix} Y_1 \\ \vdots \\ Y_n \end{bmatrix} = Y = PX \qquad \text{where } P = \begin{bmatrix} \frac{1}{\sqrt{n}}, & \frac{1}{\sqrt{n}}, & \frac{1}{\sqrt{n}}, & \cdots, & \frac{1}{\sqrt{n}} \\ \frac{1}{\sqrt{n(n-1)}}, & \frac{1}{\sqrt{n(n-1)}}, & \frac{1}{\sqrt{n(n-1)}}, & \cdots, & -\frac{n-1}{\sqrt{n(n-1)}} \\ \vdots & \vdots & \vdots & & \vdots \\ \frac{1}{\sqrt{3\times 4}}, & \frac{1}{\sqrt{3\times 4}}, & \frac{1}{\sqrt{3\times 4}}, & \cdots, & 0 \\ \frac{1}{\sqrt{2\times 3}}, & \frac{1}{\sqrt{2\times 3}}, & -\frac{1}{\sqrt{2\times 3}}, & \cdots, & 0 \\ \frac{1}{\sqrt{1\times 2}}, & -\frac{1}{\sqrt{1\times 2}}, & 0, & \cdots, & 0 \end{bmatrix} \text{ norm } 1$$

$(P^TP = PP^T = I)$
orthogonal

$$Y \sim N(P\mathbb{E}(X), P\,Var(X)P^T)$$
$$\overset{d}{=} N(0, I)$$

Multiply by P : Rotation

Specifically, $Y_1 = \frac{1}{\sqrt{n}}X_1 + \cdots + \frac{1}{\sqrt{n}}X_n = \sqrt{n}\bar{X}$

$$Y_1^2 + Y_2^2 + \cdots + Y_n^2 = \|Y\|_2^2 = \|X\|_2^2 = X_1^2 + X_2^2 + \cdots + X_n^2$$

So
$$\begin{aligned} Y_2^2 + \cdots + Y_n^2 &= X_1^2 + \cdots + X_n^2 - Y_1^2 \\ &= X_1^2 + \cdots + X_n^2 - n\bar{X}^2 \\ &= \sum_{i=1}^{n}(X_i - \bar{X})^2 \end{aligned}$$

We have shown that $Y_i \overset{iid}{\sim} N(0,1)$,

so $\sqrt{n}\bar{X} \sim N(0,1)$, $\bar{X} \sim N(0,\frac{1}{n})$

Plus, $Y_2^2 + \cdots + Y_n^2 \sim \chi^2_{n-1}$, so $\sum_{i=1}^{n}(X_i-\bar{X})^2 \sim \chi^2_{n-1}$

Since Y_1 is independent of $(Y_2, \ldots, Y_n)$,

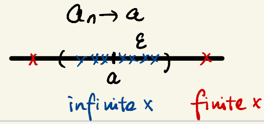
we know $\bar{X}$ is independent of $\sum_{i=1}^{n}(X_i - \bar{X})^2$.

$X \perp\!\!\!\perp Y \Rightarrow g(X) \perp\!\!\!\perp g(Y)$

$\downarrow$

$P(X\in A, Y\in B) = P(X\in A)\cdot P(Y\in B)$, $\forall A, B$

$$\text{Then } \begin{aligned} P(g(X)\in A', g(Y)\in B') &= P(X\in g^{-1}(A'), Y\in g^{-1}(B')) \\ &= P(X\in g^{-1}(A'))\cdot P(Y\in g^{-1}(B')) \\ &= P(g(X)\in A')\cdot P(g(Y)\in B') \end{aligned}$$



* Modes of convergence

$Y_1, Y_2, \dots, Y_n, \dots$

① Convergence in probability

$\lim_{n\to\infty} P(|Y_n - Y| > \varepsilon) = 0. \quad \forall \varepsilon > 0.$

$[\forall \varepsilon > 0,\ \exists N < \infty$, such that $P(|Y_n - Y| > \delta) < \varepsilon$ for any $n \geq N]$

Ex. $Y_n = \begin{cases} 1 & \text{w/ prob. } \frac{1}{n} \\ 0 & \text{w/ prob. } 1-\frac{1}{n} \end{cases}, \quad Y \equiv 0$ w/ prob. 1

$P(|Y_n - Y| > \frac{1}{2}) = P(Y_n = 1) = \frac{1}{n}$

* Weak law of large numbers

$X_1, X_2, \dots, \overset{iid}{\sim} f(x), \quad \mu = \mathbb{E}(X_i),\ \sigma^2 = Var(X_i)$

let $Y_n = \bar{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i,\ n \in \{1, 2, 3, \dots\}$

WLLN: $\bar{X}_n \xrightarrow{P} \mu$

conv in prob.

Pf. $\mathbb{E}(\bar{X}_n) = \mathbb{E}(X_i) = \mu,\ Var(\bar{X}_n) = \frac{1}{n} Var(X_i)$

$P(|\bar{X}_n - \mu| > \varepsilon) \leq \frac{Var(\bar{X}_n)}{\varepsilon^2} = \frac{\sigma^2}{n\varepsilon^2} \downarrow 0,\ \forall \varepsilon$

Chebyshev's

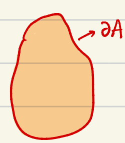
$P(|W - \mathbb{E}W| > \varepsilon) \leq \frac{Var(W)}{\varepsilon^2}$

② Convergence in distribution

$Y_n \Longrightarrow Y$ if $\lim_{n\to\infty} P(Y_n \in A) = P(Y \in A)$

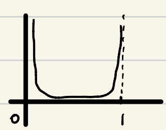
for every $A \in \mathcal{B}$ such that $P(Y \in \partial A) = 0$

Y cannot on the boundary.



eg. $Y_n \sim Be(\frac{1}{n}, \frac{1}{n})$,

$Y_n \Rightarrow Y,\ Y \sim Ber(\frac{1}{2})$



which is $Y_n \Rightarrow Ber(\frac{1}{2})$

which is $Be(\frac{1}{n}, \frac{1}{n}) \Rightarrow Bern(\frac{1}{2})$

We can show: $P(Y_n \in (0.5, 0.8)) \to P(Y \in (0.5, 0.8))$

$P(Y_n \in (0.5, 1.5)) \to P(Y \in (0.5, 1.5))$

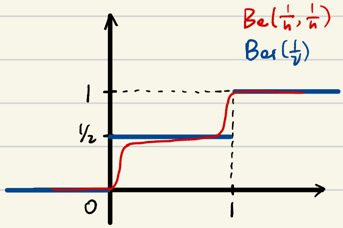
Check all intervals.

Shortcut:

$Y_n \Rightarrow Y$ if and only if $F_{Y_n}(y) \xrightarrow{n\to\infty} F_Y(y)$ for every $y \in \mathbb{R}$, s.t. $P(Y=y)=0$.

WTS:

$F_{Y_n}(y) \longrightarrow F_Y(y)$ for all $y \notin \{0,1\}$.



eg. $(X_1, \dots, X_d) \sim \text{Multi}(n, p)$

$\downarrow$

$(p_1, \dots, p_d) \in \Delta^d$

$$Y_n = \sum_{j=1}^{d} \frac{(X_j - np_j)^2}{np_j} \Rightarrow \chi^2_{d-1}$$

Should show: $F_{Y_n}(y) \rightarrow F_{\chi^2_d}(y)$, $\forall y \in \mathbb{R}$.

* CLT

$X_1, X_2, \dots \overset{iid}{\sim} f(x)$, $\mu = \mathbb{E}(X_i)$, $\sigma^2 = Var(X_i)$

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

which is $\sqrt{n}\,\dfrac{\bar{X} - \mu}{\sigma} \Longrightarrow N(0,1)$

not fixed.

$\downarrow$

which is $\bar{X} \sim AN(\mu, \frac{\sigma^2}{n})$ [NOT allowed: $\bar{X}_n \Rightarrow N(\mu, \frac{\sigma^2}{n})$]