and $\underline{g}(\underline{z}_n) \xrightarrow{D} \underline{g}(\underline{z})$ for continuous $\underline{g}$ by the continuous mapping Th.

25} Th} Let $\underline{x} = \begin{pmatrix} \underline{x}_1 \\ \underline{x}_2 \end{pmatrix} \in \mathbb{R}^k$ with $\underline{x}_1 \in \mathbb{R}^{k_1}$, $\underline{x}_2 \in \mathbb{R}^{k_2}$ and $k_1 + k_2 = k$.

Let $C_{\underline{x}}$, $C_{\underline{x}_1}$ and $C_{\underline{x}_2}$ be the char fns of $\underline{x}$, $\underline{x}_1$, and $\underline{x}_2$, Then $\underbrace{\underline{x}_1 \perp\!\!\!\perp \underline{x}_2}_{\text{independent}}$

iff $C_{\underline{x}}(\underline{t}) = C_{\underline{x}_1}(\underline{t}_1)\, C_{\underline{x}_2}(\underline{t}_2) \quad \forall\, \underline{t} = \begin{pmatrix} \underline{t}_1 \\ \underline{t}_2 \end{pmatrix} \in \mathbb{R}^k$

Severini p 82

Prove $\underline{x}_1 \perp\!\!\!\perp \underline{x}_2 \Rightarrow C_{\underline{x}}(\underline{t}) = C_{\underline{x}_1}(\underline{t}_1)\, C_{\underline{x}_2}(\underline{t}_2)$:

$$C_{\underline{x}}(\underline{t}) = E\left[e^{i\underline{t}^T\underline{x}}\right] = E\left[e^{i(\underline{t}_1^T\ \underline{t}_2^T)\begin{pmatrix} \underline{x}_1 \\ \underline{x}_2 \end{pmatrix}}\right]$$

$$= E\left[e^{i(\underline{t}_1^T\underline{x}_1 + \underline{t}_2\underline{x}_2)}\right] = E\left[e^{i\underline{t}_1^T\underline{x}_1}\, e^{i\underline{t}_2^T\underline{x}_2}\right]$$

$\overset{\text{ind}}{=} E\left[e^{i\underline{t}_1^T \underline{x}_1}\right] E\left[e^{i\underline{t}_2^T \underline{x}_2}\right] = C_{\underline{x}_1}(\underline{t}_1)\, C_{\underline{x}_2}(\underline{t}_2)$

(similar to proof for mgfs).

Note! Independence is an exception to 23} since ind $\underset{\sim}{RV}$s have a joint dist that does not affect the marginal dist and vice verca.

26} If $\underline{x}_n \xrightarrow{D} \underline{x}$ and $\underline{Y}_n \xrightarrow{P} \underline{c}$, a constant vector, then

$$\begin{pmatrix} \underline{x}_n \\ \underline{Y}_n \end{pmatrix} \xrightarrow{D} \begin{pmatrix} \underline{x} \\ \underline{c} \end{pmatrix}.$$

Severini p 346

Note! a constant vector $\underline{c} \perp\!\!\!\perp \underline{x}$ for any $\underset{\sim}{RV}$ $\underline{x}$. (Knowing 10=10 has no effect on $\underline{x}$.)

ex] Let $X \sim N(0,1)$,

Let $X_n = X \;\; \forall n$. Let $Y_n = \begin{cases} X & n \text{ odd} \\ -X & n \text{ even} \end{cases}$

So $Y_n \sim N(0,1)$. $X_n \xrightarrow{D} X$, $Y_n \xrightarrow{D} X$

Then $(1 \;\; 1)\begin{pmatrix} X_n \\ Y_n \end{pmatrix} = X_n + Y_n = \begin{cases} 2X & n \text{ odd} \\ 0 & n \text{ even} \end{cases}$

does not converge in dist as $n \to \infty$

$\therefore \begin{pmatrix} X_n \\ Y_n \end{pmatrix}$ does not converge in dist.

ex] Let $X \sim N(0,1)$ and $W \sim N(0,1)$.

Let $X_n = X \;\; \forall n$ and

$Y_n = -X \;\; \forall n$.

Then $\begin{pmatrix} X_n \\ Y_n \end{pmatrix} = \begin{pmatrix} X \\ -X \end{pmatrix} \;\; \forall n$ so $\begin{pmatrix} X_n \\ Y_n \end{pmatrix} \xrightarrow{D} \begin{pmatrix} X \\ -X \end{pmatrix}$,

but $X_n \xrightarrow{D} W$ and $Y_n \xrightarrow{D} W$.

Thus $\begin{pmatrix} X_n \\ Y_n \end{pmatrix} \not\xrightarrow{D} \begin{pmatrix} W \\ W \end{pmatrix}$ since

$(1 \;\; 1)\begin{pmatrix} X_n \\ Y_n \end{pmatrix} = X_n + Y_n = 0 \;\; \forall n$

$\Rightarrow (1 \quad 1)\begin{pmatrix} x_n \\ y_n \end{pmatrix} \xrightarrow{D} 0 \sim N(0,0)$

while $(1 \quad 1)\begin{pmatrix} W \\ W \end{pmatrix} = 2W \sim N(0,4)$

$\therefore (1 \quad 1)\begin{pmatrix} x_n \\ y_n \end{pmatrix} \cancel{\xrightarrow{D}} (1 \quad 1)\begin{pmatrix} W \\ W \end{pmatrix}$.

27} Th}

i) $\underline{x}_n \xrightarrow{wp1} \underline{x} \Rightarrow \underline{x}_n \xrightarrow{P} \underline{x}, \ \underline{x}_n \xrightarrow{D} \underline{x}$

ii) $\underline{x}_n \xrightarrow{r} \underline{x} \Rightarrow \underline{x}_n \xrightarrow{P} \underline{x}, \quad \underline{x}_n \xrightarrow{D} \underline{x}$.

iii) $\underline{x}_n \xrightarrow{P} \underline{x} \Rightarrow \underline{x}_n \xrightarrow{D} \underline{x}$

iv) $\underline{x}_n \xrightarrow{P} \underline{c}$ iff $\underline{x}_n \xrightarrow{D} \underline{c}$

constant vector

Proofs of ii) Let $\varepsilon > 0$.

a) $E[\|\underline{x}_n - \underline{x}\|^r] \geq E[\|\underline{x}_n - \underline{x}\|^r I(\|\underline{x}_n - \underline{x}\| \geq \varepsilon)]$

$\geq \varepsilon^r P[\|\underline{x}_n - \underline{x}\| \geq \varepsilon]$ so

$$P\left[\|\underline{x}_n - \underline{x}\| \geq \varepsilon\right] \leq \frac{E\left[\|\underline{x}_n - \underline{x}\|^r\right]}{\varepsilon^r} \to 0$$

as $n \to \infty$.

b) $P\left[\|x_n - x\| \geq \varepsilon\right] = P\left[\|\underline{x}_n - \underline{x}\|^r \geq \varepsilon^r\right]$

$$\leq \frac{E\left[\|\underline{x}_n - \underline{x}\|^r\right]}{\varepsilon^r} \to 0 \text{ by Gen Cheb.}$$

$\square$

"interchangeable"

28) For a 0.9 quantile = 90th percentile

$y_{0.9} = y_\alpha$, would like $P(Y \leq y_\alpha) = \alpha = 0.9$

$0 < \alpha < 1$. Let y_α be the α quantile = 100αth percentile. Then

$P(Y \leq y_\alpha) \geq \alpha$ and $P(Y > y_\alpha)$

$= 1 - P(Y \leq y_\alpha) \leq 1 - \alpha$.

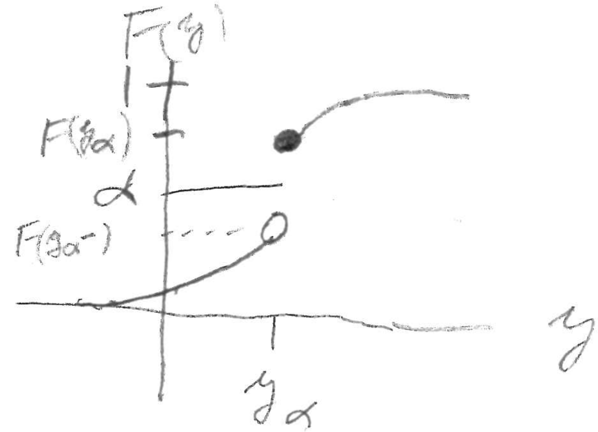
Then $P(Y \geq Y_\alpha) = P(Y = y_\alpha) + P(Y > y_\alpha)$.

Case 1

$P(Y > y_\alpha) \underset{\leq}{<} 1-\alpha$

$P(Y < y_\alpha) < \alpha$

$P(Y \leq y_\alpha) \overset{=}{\geq} \alpha$



Let $F(y-) = P(Y < y)$.

$F(y) = P(Y \leq y)$

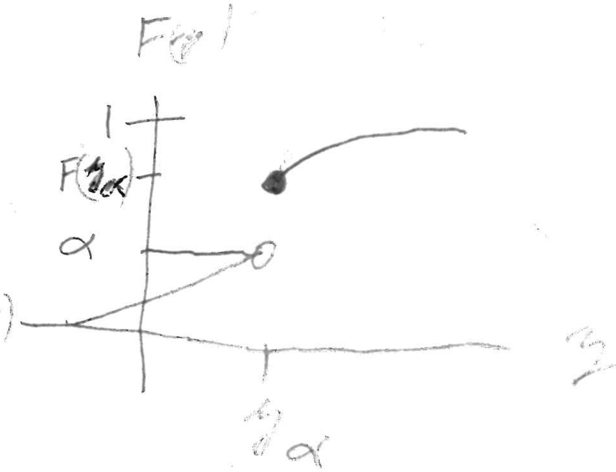
$P(Y > y)$
$= 1 - P(Y \leq y)$
$= 1 - F(y)$

case 2

$P(Y > y_\alpha) \underset{\leq}{<} 1-\alpha$

$P(Y < y_\alpha) = \alpha = F(y_\alpha-)$

$P(Y \leq y_\alpha) \overset{=}{>} \alpha$

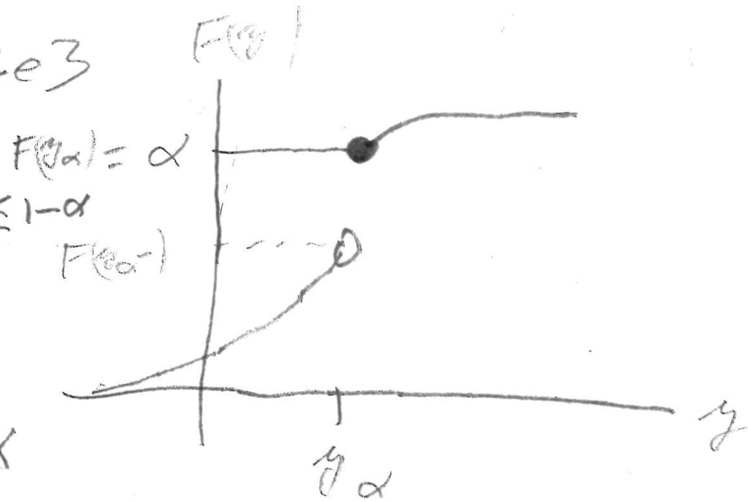


Case 3

$P(Y > y_\alpha) = 1-\alpha \leq 1-\alpha$

$P(Y < y_\alpha) < \alpha$

$P(Y \leq y_\alpha) = \alpha \geq \alpha$

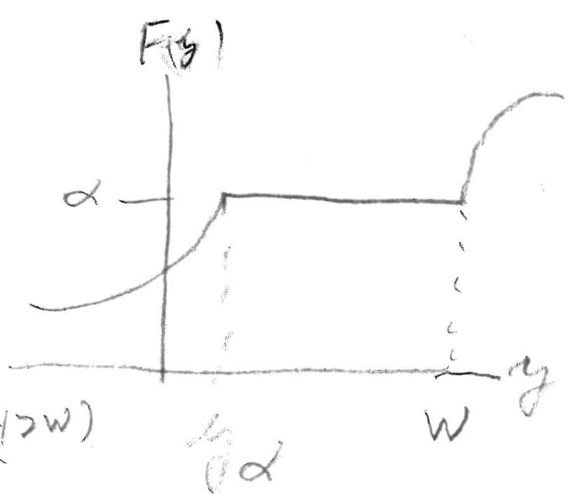


case 4)

want y_α to be unique.

$F(w) =$
$F(y_\alpha) = P(Y \leq y_\alpha) = \alpha$

$P(Y > y_\alpha) = 1-\alpha = P(Y > w)$



29] The Quantile Function

$$Q(u) = \inf\{x : u \le F(x)\}, \quad 0 < u < 1.$$

If F^{-1} exists, then $Q(u) = F^{-1}(u)$.

$Q(u) \le x$ iff $u \le F(x)$.

Define $y_\alpha = Q(\alpha)$ to get unique pop quantile when $F(x) = \alpha$ for $y_\alpha \le x < w$ or $\le w$

30] If Y has pdf $f_Y(y)$ such that $f(y) > 0$ in a neighborhood of y_α, then $P(Y \le y_\alpha) = \alpha$ and $P(Y > y_\alpha) = 1 - \alpha$.

Then $F(y)$ is continuous and strictly increasing in a neighborhood of y_α.

If $F(y_\alpha)$ is at a flat spot, then estimators $\hat{y}_\alpha$ could oscillate between y_α and w and not be consistent.

31} There are many estimators of y_α, we will use the sample percentile $\hat{y}_\alpha = Y_{(\lceil n\alpha \rceil)}$.

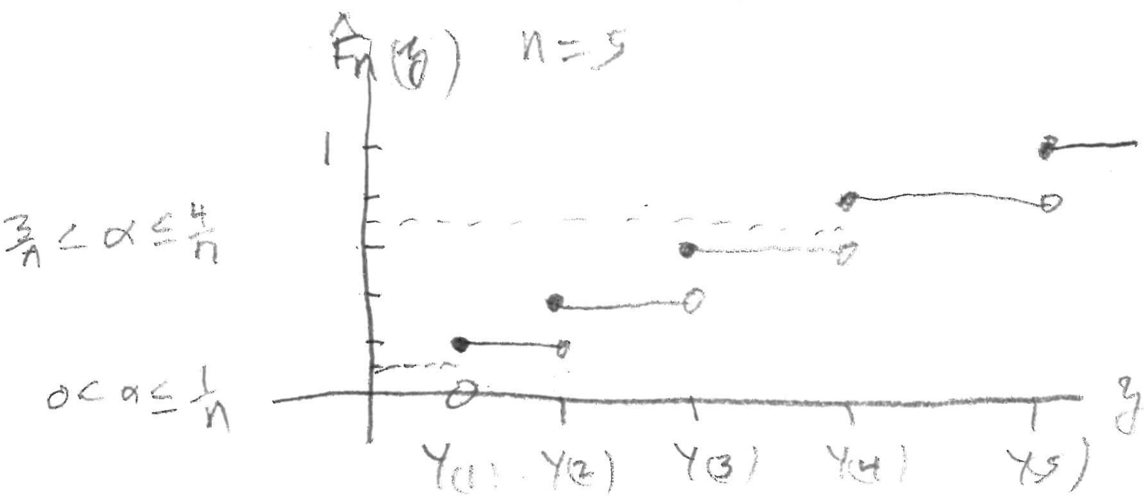
For $0 < \alpha < 1$, $1 \le \lceil n\alpha \rceil \le n$.

Here the <u>ceiling function</u> $\lceil x \rceil$ is the smallest integer $\ge x$. So $\lceil 0.7 \rceil = 1$, $\lceil 7.7 \rceil = 8$, $\lceil 8 \rceil = 8$ etc.

32} The <u>empirical cdf</u>

$$\hat{F}_n(y) = \frac{1}{n} \sum_{i=1}^{n} I(Y_i \le y)$$ and has jumps of $\frac{1}{n}$ at the order statistics.

Then $\hat{y}_\alpha = \hat{Q}(\alpha) = \inf\{y : \hat{F}(y) \ge \alpha\}$



If $0 < \alpha \leq \frac{1}{n}$ then $0 < n\alpha \leq 1$

and $\hat{Q}(\alpha) = Y_{(1)} = Y_{(\lceil n\alpha \rceil)}$

If $\frac{1}{n} < \alpha \leq \frac{2}{n}$ then $1 < n\alpha \leq 2$

and $\hat{Q}(\alpha) = Y_{(2)} = Y_{(\lceil n\alpha \rceil)}$

etc.

Note that $\hat{y}_{1-\alpha} = Y_{(\lceil n(1-\alpha) \rceil)}$.

Note that $\hat{y}_{0.5}$ is not quite the sample median.

33] P314 Quantile Theorem]

Let $0 < \alpha_1 < \alpha_2 < \dots < \alpha_k < 1$.

Suppose F has pdf f that is positive and continuous in neighborhoods of $y_{\alpha_1}, \dots, y_{\alpha_k}$. Then

$$\sqrt{n}\left[(\hat{y}_{n\alpha_1}, \dots, \hat{y}_{n\alpha_k})^T - (y_{\alpha_1}, \dots, y_{\alpha_k})^T\right] \xrightarrow{D}$$

$N_k(\underline{0}, \Sigma)$ where $\Sigma = (\sigma_{ij})$

and $\sigma_{ij} = \dfrac{\alpha_i(1-\alpha_j)}{f(y_{\alpha_i})\, f(y_{\alpha_j})}$ for

$i \le j$ and $\sigma_{ij} = \sigma_{ji}$ for $i > j$.

Note: often ξ_α is used for y_α.
↑ Greek letter xi

p315
34) **The Multivariate Delta Method:**

If $\sqrt{n}\,(T_n - \underline{\theta}) \xrightarrow{D} N_K(\underline{0}, \Sigma)$,

then $\sqrt{n}\left(\underline{g}(T_n) - \underline{g}(\underline{\theta})\right) \xrightarrow{D} N_d\left[\underline{0},\, D_{\underline{g}(\underline{\theta})}\, \Sigma\, D^T_{\underline{g}(\underline{\theta})}\right]$

where the $d \times K$ Jacobian matrix of partial derivatives

$$D_{\underline{g}(\underline{\theta})} = \begin{bmatrix} \frac{\partial}{\partial \theta_1} g_1(\underline{\theta}) & \cdots & \frac{\partial}{\partial \theta_K} g_1(\underline{\theta}) \\ \vdots & & \\ \frac{\partial}{\partial \theta_1} g_d(\underline{\theta}) & \cdots & \frac{\partial}{\partial \theta_K} g_d(\underline{\theta}) \end{bmatrix}$$

Also $D\underset{\sim}{g}(\underset{\sim}{\theta}) \Sigma D\underset{\sim}{g}(\underset{\sim}{\theta}) > 0$ ← positive definite

and the mapping $\underset{\sim}{g} : \mathbb{R}^k \to \mathbb{R}^d$ needs to be differentiable in a neighborhood of $\underset{\sim}{\theta}$.

ex} $\underset{\sim}{g}(\underset{\sim}{\theta}) = \underset{d \times k}{A} \; \underset{k \times 1}{\theta} = \begin{bmatrix} \underset{\sim}{a}_1^T \\ \vdots \\ \underset{\sim}{a}_d^T \end{bmatrix} \underset{\sim}{\theta}$ ← 1st row of A

$$= \begin{pmatrix} \underset{\sim}{a}_1^T \underset{\sim}{\theta} \\ \vdots \\ \underset{\sim}{a}_d^T \underset{\sim}{\theta} \end{pmatrix} = \begin{pmatrix} g_1(\underset{\sim}{\theta}) \\ \vdots \\ g_d(\underset{\sim}{\theta}) \end{pmatrix} = \begin{pmatrix} a_{11} & \cdots & a_{1k} \\ \vdots & & \\ a_{d1} & \cdots & a_{dk} \end{pmatrix} \begin{pmatrix} \theta_1 \\ \vdots \\ \theta_k \end{pmatrix}$$

$$= \begin{bmatrix} \sum_{i=1}^{k} a_{1i} \theta_i \\ \vdots \\ \sum_{i=1}^{k} a_{di} \theta_i \end{bmatrix}. \text{ So } D\underset{\sim}{g}(\underset{\sim}{\theta}) = \begin{bmatrix} a_{11} & \cdots & a_{1k} \\ \vdots & & \\ a_{d1} & \cdots & a_{dk} \end{bmatrix}$$

$= A$ and $\sqrt{n}(AT_n - A\underset{\sim}{\theta}) \xrightarrow{D} N_d[0, A \Sigma A^T]$ *

see 10 i).

35) The delta method is a special case since if T_n and θ are real valued, then $D_{g(\theta)} = g'(\theta)$.

see HW6

ex) suppose $\sqrt{n}\left[\begin{pmatrix}\hat{\lambda}\\ \hat{\eta}\end{pmatrix} - \begin{pmatrix}\lambda\\ \eta\end{pmatrix}\right] \xrightarrow{D} N\left[\begin{pmatrix}0\\0\end{pmatrix}, \underbrace{\begin{pmatrix}\sigma_{11} & \sigma_{12}\\ \sigma_{21} & \sigma_{22}\end{pmatrix}}_{\Sigma}\right]$

a) Find the limiting dist of

$$\sqrt{n}\,(\hat{\lambda}\hat{\eta} - \lambda\eta) = N(0, \tau^2).$$

soln $\underline{\theta} = \begin{pmatrix}\lambda\\ \eta\end{pmatrix}$ and $g(\underline{\theta}) = \lambda\eta$

$$D_{g(\underline{\theta})} = \left[\frac{\partial g}{\partial \lambda} \quad \frac{\partial g}{\partial \eta}\right] = [\eta \quad \lambda]$$

$$\tau^2 = D_{g(\underline{\theta})}\, \Sigma\, D_{g(\theta)}^T = [\eta \quad \lambda]\begin{pmatrix}\sigma_{11} & \sigma_{12}\\ \sigma_{21} & \sigma_{22}\end{pmatrix}\begin{pmatrix}\eta\\ \lambda\end{pmatrix}$$

$$= (\eta\sigma_{11} + \lambda\sigma_{21} \quad \eta\sigma_{12} + \lambda\sigma_{22})\begin{pmatrix}\eta\\ \lambda\end{pmatrix}$$

$= \eta^2 \sigma_{11} + \eta \lambda \sigma_{21} + \lambda \eta \sigma_{12} + \lambda^2 \sigma_{22}$

$$\boxed{\tau^2 = \eta^2 \sigma_{11}^2 + 2 \eta \lambda \sigma_{12} + \lambda^2 \sigma_{22}}$$

$\left(\sigma_{\hat{\lambda}\hat{\eta}} = \sigma_{12} = \sigma_{21}, \quad \sigma_{11} = \sigma^2_{\hat{\lambda}}, \quad \sigma_{22} = \sigma^2_{\hat{\eta}} \right)$

b) $\underbrace{(1 \quad 1)}_{A} \sqrt{n} \left[\begin{pmatrix} \hat{\lambda} \\ \hat{\eta} \end{pmatrix} - \begin{pmatrix} \lambda \\ \eta \end{pmatrix} \right] =$

$\sqrt{n} \left[(\hat{\lambda} + \hat{\eta}) - (\lambda + \eta) \right] \xrightarrow{D} N(0, A \Sigma A^T)$

where $A \Sigma A^T = \begin{bmatrix} 1 & 1 \end{bmatrix} \begin{bmatrix} \sigma_{11} & \sigma_{12} \\ \sigma_{21} & \sigma_{22} \end{bmatrix} \begin{bmatrix} 1 \\ 1 \end{bmatrix} =$

$\begin{bmatrix} \sigma_{11} + \sigma_{21} & \sigma_{12} + \sigma_{22} \end{bmatrix} \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \sigma_{11} + 2\sigma_{12} + \sigma_{22}$

b) $(1 \quad -1) \sqrt{n} \left[\begin{pmatrix} \hat{\lambda} \\ \hat{\eta} \end{pmatrix} - \begin{pmatrix} \lambda \\ \eta \end{pmatrix} \right] =$

$\sqrt{n} \left[(\hat{\lambda} - \hat{\eta}) - (\lambda - \eta) \right] \xrightarrow{D} N(0, B \Sigma B^T)$