$X \sim POIS(\lambda) \simeq N(\lambda, \lambda)$ if $\lambda \geq 9$.

So $X_n \sim bin(n, P_n) \simeq N(n\hat{P}_n, n\hat{P}_n(1-\hat{P}_n))$

if $n\hat{P}_n \geq 9$ and $n(1-\hat{P}_n) \geq 9$.

74} Can't, in general, take a limit of a sequence because the limit may not exist. The limit does exist if the sequence is nonincreasing or nondecreasing where $\pm\infty$ are allowed.

75} Let $\{a_n\}_{n=m}^{\infty} = a_m, a_{m+1}, \ldots$ be a sequence of numbers.

$\sup a_n$ = least upper bound of $\{a_n\}$

$\inf a_n$ = greatest lower bound of $\{a_n\}$

$\limsup_n a_n = \overline{\lim}\, a_n$ is the limit of the nonincreasing sequence $\sup_{k \geq m} a_k, \sup_{k \geq m+1} a_k, \ldots$

$\liminf_n a_n = \underline{\lim}\, a_n$ is the limit of the nondecreasing sequence

$$\inf_{k \geq m} a_k, \quad \inf_{k \geq m+1} a_k, \ldots .$$

$$\overline{\lim}\, a_n = \inf_n \sup_{k \geq n} a_k = \lim_{k \to \infty} \sup(a_n, n \geq k)$$

$$\underline{\lim}\, a_n = \sup_n \inf_{k \geq n} a_k = \lim_{k \to \infty} \inf(a_n, n \geq k)$$

If a limit point of a sequence $\{a_n\}$ is any number including $\pm\infty$ that is a limit of some subsequence, then $\liminf_n a_n$ and $\limsup_n a_n$ are the inf and sup of the set of limit points, often the smallest and largest limit points.

warning: a common error is to take the limit of both sides of an equation $a_n = b_n$ or an inequality $a_n \leq b_n$. This is an error if the existence of the limit has not been shown.

74} i) If $\pm\infty$ are allowed, $\underline{\lim}\, a_n$ and $\overline{\lim}\, a_n$
76} always exist,

LS 36

ii) $\lim_{n\to\infty} a_n = L$, possibly $\pm\infty$

iff $\limsup_n a_n = \liminf_n a_n$.

Then $\limsup_n a_n = \liminf_n a_n = L = \lim_n a_n$.

iii) $\liminf_n a_n \leq \limsup_n a_n$.

iv) Given 2 sequences $\{a_n\}$ and $\{b_n\}$,

let $\lim^* a_n$ be $\liminf_n a_n$ or $\limsup_n a_n$.

a) If $a_n \leq b_n$, $\lim^* a_n \leq \lim^* b_n$.

If $a_n \geq b_n$, $\lim^* a_n \geq \lim^* b_n$.

b) If $a_n < b_n$, $\lim^* a_n \leq \lim^* b_n$.

If $a_n > b_n$, $\lim^* a_n \geq \lim^* b_n$.

Hence the $\overline{\lim}$ or $\underline{\lim}$ of a strict inequality

$<$ or $>$ must be replaced by $\leq$ or $\geq$.

(Also true for limit if the limits exist.)

v) $\limsup_n (-a_n) = -\liminf_n a_n$

ex) a) $a_n = (-1)^n$, $\limsup_n a_n = 1$, $\liminf_n a_n = -1$

b) $a_n = \frac{(-1)^n}{n}$, $\limsup_n a_n = \liminf_n a_n = \lim_n a_n = 0$.

c) $\frac{1}{n+1} < \frac{1}{n}$ but

$\lim^* \frac{1}{n+1} \le \lim^* \frac{1}{n}$, in fact,

$\overline{\lim} \frac{1}{n+1} = 0 = \underline{\lim} \frac{1}{n+1} = \overline{\lim} \frac{1}{n} = \underline{\lim} \frac{1}{n}$.

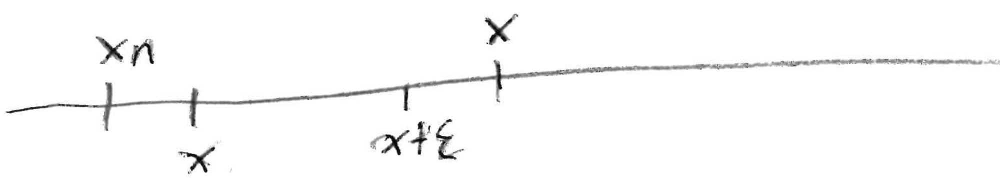
So $\overline{\lim} \frac{1}{n+1}$ is <u>not</u> less than $\overline{\lim} \frac{1}{n}$.

77) Prove $X_n \xrightarrow{P} X \Rightarrow X_n \xrightarrow{D} X$.

proof) $F_n(x) = P(X_n \le x) = \underbrace{P(X_n \le x, X > x+\varepsilon) + P(X_n \le x, X \le x+\varepsilon)}_{\text{partition}}$

see diagram

$\le P(|X_n - X| \ge \varepsilon) + P(X \le x+\varepsilon)$



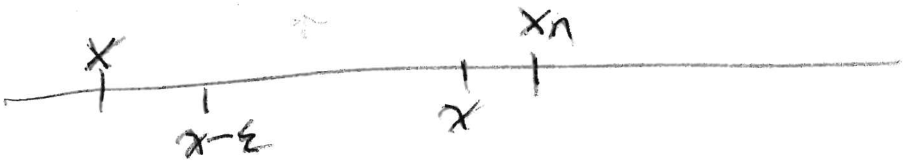
$= P(|X_n - X| \ge \varepsilon) + F_X(x+\varepsilon)$.

$F_X(x-\varepsilon) = P(X \le x-\varepsilon) =$

$\underbrace{P(X \le x-\varepsilon, X_n > x) + P(X \le x-\varepsilon, X_n \le x)}_{\text{partition}}$

see diagram

$\le P(|X_n - X| \ge \varepsilon) + P(X_n \le x)$



$= P(|X_n - X| \geq \varepsilon) + F_n(x).$

$\therefore F_X(x-\varepsilon) - P(|X_n - X| \geq \varepsilon) \leq F_n(x) \leq$

$$P(|X_n - X| \geq \varepsilon) + F_X(x+\varepsilon).$$

Since $X_n \xrightarrow{P} X$, $P(|X_n - X| \geq \varepsilon) \to 0$ as $n \to \infty$.

If $F_X(x)$ is continuous at x, then

$F_X(x-\varepsilon) \to F_X(x)$ and $F_X(x+\varepsilon) \to F_X(x)$ as $\varepsilon \to 0$.

Taking lim inf and lim sup gives

$F_X(x-\varepsilon) \leq \liminf_n F_n(x) \leq \limsup_n F_n(x) \leq F_X(x+\varepsilon).$

$\therefore F_n(x) \to F_X(x)$ if $F_X(x)$ is contin at x. $\square$

Ash stat

78] Prove $X_n \xrightarrow{P} c$ iff $X_n \xrightarrow{D} c$.

proof] showed $X_n \xrightarrow{P} c \Rightarrow X_n \xrightarrow{D} c$ in 77].

$P[|X_n - c| \geq \varepsilon] = P[X_n \geq c+\varepsilon] + P[X_n \leq c-\varepsilon]$

$= 1 - P(X_n < c+\varepsilon) + P(X_n \le c-\varepsilon) = RHS$ 3Z.5

Now $P(X_n < c+\varepsilon) \ge P(X_n \le c+\frac{\varepsilon}{2})$. $\therefore$

$$P[|X_n - c| \ge \varepsilon] = RHS \le 1 - P(X_n \le c+\tfrac{\varepsilon}{2}) + P(X_n \le c-\varepsilon)$$

$$= 1 - \underbrace{F_n(c+\tfrac{\varepsilon}{2})}_{\to 1} + \underbrace{F_n(c-\varepsilon)}_{\to 0}$$

since $F_n(t) \to F_X(t)$ for $t \ne c$ where $F_X(t) = \begin{cases} 1 & t \ge c \\ 0 & t < c. \end{cases}$

$\therefore P[|X_n - c| \ge \varepsilon] \to 1 - 1 + 0 = 0$ as $n \to \infty$

$\square$

79) Let X_n have pdf $f_{X_n}(x)$ and let X have pdf $f_X(x)$.

(a) If $f_{X_n}(x) \to f_X(x)$ $\forall x$, then $X_n \xrightarrow{D} X$.

b) If $f_{X_n}(x) \to f_X(x)$ $\underbrace{\text{ae wrt Lebesgue measure}}_{\text{for } x \text{ outside a set of Lebesgue measure } 0}$ then $X_n \xrightarrow{D} X$.

80] Suppose X_n and X are integer valued RVs.

Then $X_n \xrightarrow{D} X$ iff $P(X_n=k) \to P(X=k)$ for every integer k iff

$\underbrace{f_{X_n}(x) \to f_X(x)}_{\text{pmfs}}$ $\forall$ real x.

Extension to Random Vectors

1] Let $\underline{X} = \begin{pmatrix} X_1 \\ \vdots \\ X_k \end{pmatrix} \in \mathbb{R}^k$,

$$E(\underline{X}) = \underset{\sim}{\mu} = \begin{pmatrix} E(X_1) \\ \vdots \\ E(X_k) \end{pmatrix},$$

$$\text{cov}(\underline{X}) = \Sigma = E\left[(\underline{X}-E(\underline{X}))(\underline{X}-E(\underline{X}))^T\right] = (\sigma_{ij})$$

where $\sigma_{ij} = \text{cov}(X_i, X_j)$ and $\text{cov}(X_i, X_i) = \sigma_i^2 = \sigma_{ii} = V(X_i)$.

p278

2] The characteristic function of $\underset{\sim}{x}$ is $C_{\underset{\sim}{x}}(\underset{\sim}{t}) = E[e^{i\underset{\sim}{t}^T\underset{\sim}{x}}]$. The moment generating function of $\underset{\sim}{x}$ is $m_{\underset{\sim}{x}}(\underset{\sim}{t}) = E[e^{\underset{\sim}{t}^T\underset{\sim}{x}}]$ provided the expectation exists for all $\underset{\sim}{t}$ in a neighborhood of $\underset{\sim}{0}$.

The cdf $F_{\underset{\sim}{x}}(\underset{\sim}{t}) = P(X_1 \le t_1, \ldots, X_k \le x_k)$
row vector

if $\underset{\sim}{x}, \underset{\sim}{t} \in \mathbb{R}^k$.

3] Let $\|\underset{\sim}{x}\| = \sqrt{\underset{\sim}{x}^T\underset{\sim}{x}} = \sqrt{x_1^2 + \cdots + x_k^2} = \sqrt{\sum_{i=1}^{k} x_i^2}$.

4) know Def] Let $\underset{\sim}{X}_n \in \mathbb{R}^k$ be a sequence of random vectors $\underset{\sim}{RVs}$ with

joint cdfs $F_{\underline{X}_n}(\underline{x})$ and let $\underline{X} \in \mathbb{R}^k$ be a RV with cdf $F_{\underline{X}}(\underline{x})$.

p. 81 a) $\underline{X}_n$ converges in distribution to $\underline{X}$,

$\underline{X}_n \xrightarrow{D} \underline{X}$, if $F_{\underline{X}_n}(\underline{x}) \to F_{\underline{X}}(\underline{x})$ for all continuity points $\underline{x}$ of $F_{\underline{X}}(\underline{x})$.
$\underline{X}$ is the limiting or asymptotic dist of $\underline{X}_n$, and does not depend on n.

b) $\underline{X}_n$ converges in probability to $\underline{X}$

$\underline{X}_n \xrightarrow{P} \underline{X}$, if $P(\|\underline{X}_n - \underline{X}\| \geq \varepsilon) \to 0$ as $n \to \infty$ $\forall \varepsilon > 0$.

c) Let $r > 0$. $\underline{X}_n$ converges in rth mean to $\underline{X}$, $\underline{X}_n \xrightarrow{r} \underline{X}$, if $E[\|\underline{X}_n - \underline{X}\|^r] \to 0$ as $n \to \infty$.

d) $\underline{x}_n$ converges with probability one to $\underline{x}$

$\underline{x}_n \xrightarrow{wp1} \underline{x}$, if $P\left(\lim_{n\to\infty} \underline{x}_n = \underline{x}\right) = 1.$

p280

e) Let $B =$ wp1, p, r, or D.

For $\underset{\sim}{x}_n \xrightarrow{B} \underset{\sim}{c}$, replace $\underline{x}$ by $\underline{c}$.

f) $\xrightarrow{D} \; = \; \xrightarrow{L}$, $\quad \xrightarrow{wp1} \; = \; \xrightarrow{ae} \; = \; \xrightarrow{as}$

5) If $\underline{x}$ and $\underline{y}$ are $k\times 1$ RVs, $\underline{a}$ a conformable constant vector and A and B conformable constant matrices,

then $E[\underline{a} + \underline{x}] = \underline{a} + E(\underline{x})$,

$E[\underline{x} + \underline{y}] = E(\underline{x}) + E(\underline{y})$,

$E[A\underline{x}] = A\, E(\underline{x})$,

$E(A\underline{x}B) = \underset{g\times k}{A}\; \underset{k\times 1}{E(\underline{x})}\; \underset{1\times p}{B}$ and

$\mathrm{cov}(\underline{a} + A\underline{x}) = \mathrm{cov}(A\underline{x}) = A\, \mathrm{cov}(\underline{x})\, A^T.$

6) Generalized Chebyshev's Th:

Let $\underset{k\times 1}{\underline{X}} \in \mathbb{R}^k$ and $U: \mathbb{R}^k \to [0,\infty)$ be a nonnegative function. Then for any $c>0$,

$$P[U(\underline{X}) \geq c] \leq \frac{E[U(\underline{X})]}{c}.$$

proof when $\underline{X}$ has pdf $f_{\underline{X}}(\underline{x})$:

$$E[\underset{\substack{\uparrow \\ \text{col. vector}}}{U(\underline{X})}] = \int_{\mathbb{R}^k} \underbrace{U(\underline{x}) f(\underline{x})\, d\underline{x}}_{\text{row vectors}}$$

$$\geq \int_{\{\underline{x}: U(\underline{x}) \geq c\}} U(\underline{x}) f(\underline{x})\, d\underline{x} \qquad \left(\int_{\{\underline{x}: U(\underline{x}) < c\}} U(\underline{x}) f(\underline{x})\, d\underline{x} \geq 0 \text{ since } U(\underline{x}) f(\underline{x}) \geq 0 \right)$$

$$\geq c \int_{\{\underline{x}: U(\underline{x}) \geq c\}} f(\underline{x})\, d\underline{x} = c P[U(\underline{X}) \geq c].$$

$\square$

7) Let $E(\underline{X}) = \underline{\mu}$ and $U(\underline{X}) = \|\underline{X} - \underline{\mu}\|$.

Then $P[\|\underline{X} - \underline{\mu}\| \geq c] = P(\|\underline{X} - \underline{\mu}\|^r \geq c^r)$

TUE)

$$\le \frac{E[\|\underset{\sim}{X}-\underset{\sim}{\mu}\|^r]}{c^r} \quad \text{for } c>0.$$

can replace $\underset{\sim}{\mu}$ by $\underset{\sim}{a}$.

8) $\underset{k\times 1}{\underset{\sim}{X}}$ has a k dimensional multivariate normal (MVN) dist $N_k(\underset{\sim}{\mu}, \Sigma)$ if $\underset{\sim}{t}^T\underset{\sim}{X}$ has a univariate normal dist for any $k\times 1$ constant vector $\underset{\sim}{t}$.

Univariate normal $k=1$, $W \sim N(\mu,\sigma^2)$

$E(\underset{\sim}{X}) = \underset{\sim}{\mu}$, $\text{cov}(\underset{\sim}{X}) = \Sigma$

$\underset{\sim}{t}^T\underset{\sim}{X} \sim N(\underset{\sim}{t}^T\underset{\sim}{\mu}, \underset{\sim}{t}^T\Sigma\underset{\sim}{t})$ (write N_1 as N)

p310-1

9) If $\underset{\sim}{X} \sim N_k(\underset{\sim}{\mu}, \Sigma)$ and A is a $g\times k$ constant matrix, b a constant, $\underset{\sim}{a}$ a $k\times 1$ constant vector and $\underset{\sim}{d}$ a $g\times 1$ constant vector, then

i) $A\underset{\sim}{X} \sim N_q(A\underline{\mu}, A\Sigma A^T)$.

ii) $\underline{a} + b\underline{X} \sim N_K(\underline{a} + b\underline{\mu}, b^2\Sigma)$

Note $b\underline{X} = b I_K \underline{X}$ with $A = bI_K$.

iii) $A\underline{X} + \underline{d} \sim N_q(A\underline{\mu} + \underline{d}, A\Sigma A^T)$.

10} Suppose $\underline{X}_n \xrightarrow{D} \underset{\sim}{X} \sim N_K(\underset{\sim}{\mu}, \Sigma)$. Then

i) $A\underline{X}_n \xrightarrow{D} A\underline{X} \sim N_q(A\underline{\mu}, A\Sigma A^T)$

ii) $\underline{a} + b\underline{X}_n \xrightarrow{D} N_K(\underset{\sim}{a} + b\underset{\sim}{\mu}, b^2\Sigma)$

iii) $A\underline{X}_n + \underline{d} \xrightarrow{D} N_q(A\underline{\mu} + \underline{d}, A\Sigma A^T)$

Note} 10} can be proved using the Multivariate Delta Method.

p313

11} Multivariate Central Limit Th (MCLT):

If $\underline{X}_1, \ldots, \underline{X}_n$ are iid $k \times 1$ RVs with $E(\underline{X}) = \underline{\mu}$ and $\text{Cov}(\underline{X}) = \Sigma$, then

$$\sqrt{n}\left(\underline{\overline{X}}_n - \underline{\mu}\right) \xrightarrow{D} N_K\left(\underline{0}, \Sigma\right).$$

Note! $\underline{\overline{X}}_n = \frac{1}{n}\sum_{i=1}^{n} \underline{X}_i = \begin{pmatrix} \overline{X}_{1n} \\ \vdots \\ \overline{X}_{Kn} \end{pmatrix}$. $\left(\begin{array}{c}\text{coordinatewise} \\ \text{sample} \\ \text{means}\end{array}\right)$

usual CLT is a special case with $K=1$.

ex) know for final) Suppose $\underline{X}_1, \dots, \underline{X}_n$ are iid $p \times 1$ RVs with $E(\underline{X}_i) = e^{0.5}\underset{\sim}{1}$ and $\text{Cov}(\underline{X}_i) = (e^2 - e) I_p$. Find the limiting dist of $\sqrt{n}\left(\underline{\overline{X}}_n - \underline{c}\right)$ for appopriate $\underline{c}$.

Soln $$\sqrt{n}\left(\underline{\overline{X}}_n - \underbrace{e^{0.5}\underset{\sim}{1}}_{E(\underline{X}_i) = \mu}\right) \xrightarrow{D} N_{\underset{\uparrow}{p}}\left(\underline{0}, \underbrace{(e^2 - e) I_p}_{\text{Cov}(\underline{X}_i) = \Sigma}\right)$$

common error! get dimension p wrong

See Quiz 5
HW 5

12} Th If $0 < \delta \le 1$ and
$n^{\delta}(\underline{X}_n - \underset{\sim}{c}) \xrightarrow{D} \underline{X}$, then $\underline{X}_n \xrightarrow{P} \underline{c}$.

13} Th: If $\underline{X}_1, \dots, \underline{X}_n$ are iid with
$E(\underline{X}) = \underline{\mu} \in \mathbb{R}^k$, then

a) WLLN: $\overline{\underline{X}}_n \xrightarrow{P} \underline{\mu}$

b) SLLN: $\overline{\underline{X}}_n \xrightarrow{wp1} \underline{\mu}$.

14} know
__Continuity Th__: Let $\underline{X}_n, \underline{X}$ be $k \times 1$
with char fns $c_{\underline{X}_n}(\underline{t})$ and $c_{\underline{X}}(\underline{t})$, Then

$\underline{X}_n \xrightarrow{D} \underline{X}$ iff $c_{\underline{X}_n}(\underline{t}) \to c_{\underline{X}}(\underline{t}) \ \forall \underline{t} \in \mathbb{R}^k$.

15} p284 __know__ Th __Cramér Wold Device__:
Let $\underline{X}_n, \underline{X}$ by $k \times 1$. Then
$\underline{X}_n \xrightarrow{D} \underline{X}$ iff $\underline{t}^T \underline{X}_n \xrightarrow{D} \underline{t}^T \underline{X} \ \forall \underline{t} \in \mathbb{R}^k$.

4.2.15

<u>proof</u> $C_{\underbrace{\underline{t}^T \underline{x}_n}_{W_n}}(y) = E\left[e^{iy\,\underline{t}^T \underline{x}_n}\right] = C_{\underline{x}_n}(y\underline{t}),\ y \in \mathbb{R}$

and $C_{\underbrace{\underline{t}^T \underline{x}}_{W}}(y) = C_{\underline{x}}(y\underline{t}), \quad y \in \mathbb{R}.$

If $\underline{x}_n \xrightarrow{D} \underline{x}$, then $C_{\underline{x}_n}(\underline{t}) \to C_{\underline{x}}(\underline{t})\ \forall \underline{t} \in \mathbb{R}^k$.

Fix $\underline{t}$. Then $C_{\underline{x}_n}(y\underline{t}) \to C_{\underline{x}}(y\underline{t})\ \forall y \in \mathbb{R}$.

$\therefore\ \underline{t}^T \underline{x}_n \xrightarrow{D} \underline{t}^T \underline{x}$, ($\underline{t}$ arbitrary)

If $\underline{t}^T \underline{x}_n \xrightarrow{D} \underline{t}^T \underline{x} \quad \forall \underline{t} \in \mathbb{R}^k$,

then $C_{\underline{x}_n}(y\underline{t}) \to C_{\underline{x}}(y\underline{t}) \quad \forall y \in \mathbb{R}$
$\forall \underline{t} \in \mathbb{R}^k$.

Take $y=1$ to get $C_{\underline{x}_n}(\underline{t}) \to C_{\underline{x}}(\underline{t})$
$\forall \underline{t} \in \mathbb{R}^k$. $\therefore\ \underline{x}_n \xrightarrow{D} \underline{x}$

by the Continuity Th. $\square$

16) Proof of MCLT: For fixed $\underline{t}$, the $\underline{t}^T \underline{X}_i$ are iid with mean $\underline{t}^T \underline{\mu}$ and variance $\underline{t}^T \Sigma \underline{t}$. Hence by the CLT, $\underline{t}^T \sqrt{n}(\overline{\underline{X}}_n - \underline{\mu}) \xrightarrow{D} \underline{t}^T X \sim N(0, \underline{t}^T \Sigma \underline{t})$

scalars

where $\underset{\sim}{X} \sim N_k(\underline{0}, \Sigma)$. Hence by the Cramér Wold device, $\sqrt{n}(\overline{\underline{X}}_n - \underline{\mu}) \xrightarrow{D} N_k(\underline{0}, \Sigma)$.

$\square$

17) Continuous Mapping Th: Let $\underline{x}_n, \underline{x} \in \mathbb{R}^k$. If $\underline{x}_n \xrightarrow{D} \underline{x}$ and if $\underset{\sim}{g}: \mathbb{R}^k \to \mathbb{R}^j$ is continuous, then $\underset{\sim}{g}(\underline{x}_n) \xrightarrow{D} \underline{g}(\underline{x})$.

18) Let $W_1, \ldots, W_n$ be $m \times m$ random matrices where $W_n = (w_{ij,n})$. Let $W = (w_{ij})$ be a random matrix.

Then $W_n \xrightarrow{P} W$ iff $w_{ij,n} \xrightarrow{P} w_{ij}$.

$w_{ij,n} = {}_n w_{ij} = w_{ij}^{(n)} = w_{ij}(n)$ etc

Often $W = C$ where C is a constant matrix.

severini p346-8

see p302

14) Let W_n be a sequence of $m \times m$ random matrices and let C be an $m \times m$ constant matrix.

a) $W_n \xrightarrow{P} C$ iff

$$\underline{a}^T W_n \underline{b} \xrightarrow{P} \underline{a}^T C \underline{b} \quad \forall$$

constant vectors $\underline{a}, \underline{b} \in \mathbb{R}^m$.

b) If $W_n \xrightarrow{P} C$, then

$$\det(W_n) = |W_n| \xrightarrow{P} |C| = \det(C).$$

c) If W_n^{-1} exists for each n and if C^{-1} exists, then

$$W_n \xrightarrow{P} C \quad \text{iff} \quad W_n^{-1} \xrightarrow{P} C^{-1}.$$

20} If $\underset{k \times m}{A} = (\underset{\sim}{a}_1, \ldots, \underset{\sim}{a}_m)$ then

$$\mathrm{vec}(A) = \underset{km \times 1}{\begin{pmatrix} \underset{\sim}{a}_1 \\ \underset{\sim}{a}_2 \\ \vdots \\ \underset{\sim}{a}_m \end{pmatrix}}$$ is a vector.

So we can think of a matrix as a vector that has been unstacked.

If A is symmetric $A = \begin{pmatrix} a_{11} & a_{12} & a_{13} & \cdots & a_{1m} \\ a_{21} & a_{22} & a_{23} & \cdots & a_{2m} \\ \vdots & & & & \\ a_{k1} & a_{k2} & & & a_{km} \end{pmatrix}$

could stack the circled elements, that are on or below the diagonal, giving vech (A).

Searle 1978!

ex} $A = \begin{bmatrix} a & b \\ b & c \end{bmatrix}$, $\mathrm{vec}(A) = \begin{pmatrix} a \\ b \\ b \\ c \end{pmatrix}$

$\mathrm{vech}(A) = \begin{pmatrix} a \\ b \\ c \end{pmatrix}$.

21} Th: Let $\underline{x}_n = (x_{1n}, \ldots, x_{kn})^T$,

let $\underline{y}_n \in \mathbb{R}^k$, let $\underline{x} = (x_1, \ldots, x_k)^T$,

Let W_n be a sequence of $k \times k$ nonsingular matrices and let G be a $k \times k$ constant nonsingular matrix.

a) $\underline{x}_n \xrightarrow{P} \underline{x}$ iff $x_{in} \xrightarrow{P} x_i$ for $i = 1, \ldots, k$.

($\underline{x}_n$ gets close to $\underline{x}$ iff x_{in} gets close to x_i $\forall i$.)

b) Slutsky's Th! If $\underline{x}_n \xrightarrow{D} \underline{x}$, $\underline{y}_n \xrightarrow{P} \underline{a}$ for some constant vector $\underline{a}$ and if $W_n \xrightarrow{P} G$, then

i) $\underline{x}_n + \underline{y}_n \xrightarrow{D} \underline{x} + \underline{a}$

ii) $\underline{y}_n^T \underline{x}_n \xrightarrow{D} \underline{a}^T \underline{x}$

iii) $W_n \underline{x}_n \xrightarrow{D} G \underline{x}$

iv) $\underline{x}_n^T W_n \xrightarrow{D} \underline{x}^T G$

v) $\underline{W}_n^{-1} \underline{X}_n \xrightarrow{D} G^{-1} \underline{X}$ — if inverses exist

vi) $\underline{X}_n^T \underline{W}_n^{-1} \xrightarrow{D} \underline{X}^T G^{-1}$.

22} If $\underline{X}_n \xrightarrow{D} \underline{X}$, then

$X_{in} \xrightarrow{D} X_i$ for $i=1,\dots,k$.

proof} use the Cramér Wold device

with $\underline{t}_i = (0,\dots,0,\underset{\uparrow \atop i\text{th}}{1},0,\dots,0)^T$.

$\therefore \underline{t}_i^T \underline{X}_n = X_{in} \xrightarrow{D} X_i = \underline{t}_i^T \underline{X}$. $\square$

23} In general, $\underline{X}_{in} \xrightarrow{D} \underline{X}_i$ for $i=1,\dots,m$

does <u>not</u> imply that

$$\begin{pmatrix} \underline{X}_{1n} \\ \vdots \\ \underline{X}_{mn} \end{pmatrix} \xrightarrow{D} \begin{pmatrix} \underline{X}_1 \\ \vdots \\ \underline{X}_m \end{pmatrix}.$$ That is,

marginal convergence in distr does not imply joint convergence in dist.

24) Suppose $\underline{X_n} \perp\!\!\!\perp \underline{Y_n}$ for $n = 1, 2, \ldots$

Suppose $\underline{X_n} \xrightarrow{D} \underline{X}$ and $\underline{Y_n} \xrightarrow{D} \underline{Y}$.

Then $\begin{pmatrix} \underline{X_n} \\ \underline{Y_n} \end{pmatrix} \xrightarrow{D} \begin{pmatrix} \underline{X} \\ \underline{Y} \end{pmatrix}$ where

$\underline{X} \perp\!\!\!\perp \underline{Y}$.

Proof) $\{\underline{X_n}\} \perp\!\!\!\perp \{\underline{Y_n}\} \Rightarrow \underline{X} \perp\!\!\!\perp \underline{Y}$

Let $\underline{t} = \begin{pmatrix} \underline{t_1} \\ \underline{t_2} \end{pmatrix}$, $\underline{Z_n} = \begin{pmatrix} \underline{X_n} \\ \underline{Y_n} \end{pmatrix}$ and $\underline{Z} = \begin{pmatrix} \underline{X} \\ \underline{Y} \end{pmatrix}$.

Since $\underline{X_n} \perp\!\!\!\perp \underline{Y_n}$ and $\underline{X} \perp\!\!\!\perp \underline{Y}$,

$$C_{Z_n}(\underline{t}) = C_{\underline{X_n}}(\underline{t_1})\, C_{\underline{Y_n}}(\underline{t_2}) \to C_{\underline{X}}(\underline{t_1}) \cdot C_{\underline{Y}}(\underline{t_2}) = C_{\underline{Z}}\begin{pmatrix} \underline{t_1} \\ \underline{t_2} \end{pmatrix}.$$

$\therefore$ $\underline{Z_n} \xrightarrow{D} \underline{Z}$ by the continuity th