Probability-3 Lecture-5

06 August 2024 14:21

Fisher-Cochran Theorem:

$Q_1(x) = \underset{\sim}{x}^T A_1 \underset{\sim}{x}, \ldots, Q_n(x) = \underset{\sim}{x}^T A_n \underset{\sim}{x}$.

are quadratic forms satisfying:

$Q_1(x) + \cdots + Q_n(x) = \underset{\sim}{x}^T \underset{\sim}{x}$.

$[\text{ie, } A_1 + \cdots + A_n = I]$.

Let $r_j := r(A_j)$, $1 \le j \le n$.

Let $X_1, X_2, \ldots, X_n \overset{iid}{\sim} N(0,1)$

Then a necessary & sufficient condition for $Q_1(\underset{\sim}{X}), \ldots, Q_n(\underset{\sim}{X})$ to be independent with

$$Q_j(\underset{\sim}{X}) \sim \chi^2_{r_j}, \quad 1 \le j \le n.$$

is: $\boxed{r_1 + \cdots + r_n = k}$.

Proof: "Necessity" is obvious.

$$Q_j(\underset{\sim}{X}) \sim \chi^2_{r_j}$$

$$\therefore \sum Q_j(\underset{\sim}{X}) \sim \chi^2_{\sum r_j} = \chi^2_k$$

"Sufficiency": $\forall j = 1, \ldots, n$,

$\exists$ linearly independent vectors

$\underset{\sim}{l}_{j,1}, \underset{\sim}{l}_{j,2}, \ldots, \underset{\sim}{l}_{j,r_j}$ s.t.

$$Q_j(\underset{\sim}{x}) = \pm(\underset{\sim}{l}_{1j}^T \cdot \underset{\sim}{x})^2 \pm \cdots \pm (l_{j r_j}^T \cdot \underset{\sim}{x})^2$$

[Diagonalizing. Refer V.M.-2]

Let B be the matrix.

first r_1 rows: $\underset{\sim}{l}_{1,1}, \ldots, \underset{\sim}{l}_{1,r_1}$

next r_2 rows: $\underset{\sim}{l}_{2,1}, \ldots, l_{2,r_2}$,

$\vdots$

& so on.

So, B is a $k \times k$ matrix.

$$\underbrace{\sum_j Q_j(\underset{\sim}{x})}_{= x^T x} = \underset{\sim}{x}^T B^T \Delta B \underset{\sim}{x},$$

where $\Delta = \text{diag}(\pm 1, \pm 1, \ldots, \pm 1)$.

$\Rightarrow x^T x = x^T B^T \Delta B x.$

$\Rightarrow B^T \Delta B = I.$

$\Rightarrow \boxed{r(B) = k}$.

$\Rightarrow B$ is non-singular.

$\Delta = (B^T)^{-1} \cdot B^{-1}$

$\Rightarrow \Delta$ is positive definite
$\therefore$ &, $\Delta = \text{diag}(\pm 1, \ldots, \pm 1)$

this forces that $\Delta = \text{diag}(1, 1, \ldots, 1)$

$\Rightarrow \boxed{\Delta = I}$.

Also, B is orthogonal

$\Rightarrow \underset{\sim}{Y} = B\underset{\sim}{X} \Rightarrow Y_1, \ldots, Y_k \overset{iid}{\sim} N(0,1).$

$$\therefore \quad Q_j(\underset{\sim}{x}) = Y^2_{\sum_{i=1}^{j-1} r_i + 1} + \cdots + Y^2_{\sum_{i=1}^{j-1} r_i + r_j}$$

$$\sim \chi^2_{r_j} \qquad \blacksquare$$

<u>Corollary</u>: $X_1, \ldots, X_k \overset{iid}{\sim} N(0,1)$. A - real, symmetric matrix.

$Q(\underset{\sim}{X}) = \underset{\sim}{X}^T A \underset{\sim}{X}$ has a χ^2 distribution

$\Longleftrightarrow A$ is idempotent,

& in that case, degrees of freedom of

$\chi^2 = r(A) = tr(A)$

& in that case, 'degrees of freedom of $\chi^2 = r(A) = tr(A)$

<u>Proof</u>: ($\Leftarrow$)

Assume $A^2 = A$.

$\Rightarrow A(I-A) = 0$.

$\Rightarrow k \leq r(A) + r(I-A)$.

$\leq r(A + (I-A)) = k$.

Sylvester's inequality

$\Longrightarrow$ By Fisher-Cochran $\underset{\sim}{X}^T A \underset{\sim}{X} \sim \chi^2_{r(A)}$.

"$\Rightarrow$" Suppose, $\underset{\sim}{X}^T A \underset{\sim}{X} \sim \chi^2_d$

<u>To prove</u>: A is idempotent, & $d = r(A)$.

Let $r = r(A)$.

$\exists$ orthogonal B st.

$$B^T A B = diag(\lambda_1, \ldots, \lambda_r, 0 \cdots 0),$$

where $\lambda_1, \ldots, \lambda_r$ are the non-zero eigenvalues of A.

We make the orthogonal transformation:

$$\underset{\sim}{Y} = B\underset{\sim}{X} \Rightarrow Y_1, \ldots, Y_k \overset{iid}{\sim} N(0,1)$$

Further,

$$\underset{\sim}{X}^T A \underset{\sim}{X} = \lambda_1 y_1^2 + \cdots + \lambda_r y_r^2$$

Also, $\because$ the transformation is orthogonal,

$$\underset{\sim}{Y} = B\underset{\sim}{X}$$

$$\therefore \underset{\sim}{Y}^T \cdot \underset{\sim}{Y} = (B\underset{\sim}{X})^T (BX)$$

$$= X^T \underbrace{B^T B}_{I} X$$

$$= X^T B^T B X$$ with $\underbrace{B^T B}_{I}$

$$\underset{\sim}{Y}^T \underset{\sim}{Y} = \underset{\sim}{X}^T \underset{\sim}{X} \,.$$

Now, mgf: $E\left(e^{t \underset{\sim}{X}^T A \underset{\sim}{X}}\right) = E\left(e^{t(\lambda_1 y_1^2 + \cdots + \lambda_r y_r^2)}\right)$

$$= \prod_{j=1}^{r} E\left(e^{t\lambda_j Y_j}\right)$$

$$= \prod_{j=1}^{r} (1-2\lambda_j t)^{-1/2} \longleftarrow \text{Mgf of } \chi^2$$

$\therefore$ Now, suppose $\underset{\sim}{X}^T A X \sim \chi^2_d$,

$$\prod_{j=1}^{r} (1-2\lambda_j t)^{-1/2} = (1-2t)^{-d/2}$$

$$\Rightarrow \boxed{d = r} \;\; \& ,\;\; \lambda_j = 1 \;\; \forall j$$

$d = r$: equating highest powers of t on both sides.

$\lambda_j = 1$: equating the coefficients on both sides

Also, $B^T A B = \text{diag}(1, 1, \ldots, 1, 0, \ldots, 0)$.

$\therefore$ Clearly, $B^T A B$ is idempotent.

$$\Rightarrow B^T A \underbrace{B B^T}_{I} A B = B^T A B$$

$$\Rightarrow B^T A^2 B = B^T A B$$

$$\Rightarrow A^2 = A \,. \quad \therefore A \text{ is idempotent.}$$

(*) Suppose, $\{a_n\}$ - non-ve sequence.

Refer

Suppose, $\sum^{\infty} a_n t^n$ converges for $0 \le t < 1$

Refer pgf question from Sem-2 midsem.

Suppose, $\sum_{n=1}^{\infty} a_n t^n$ converges for $0 \le t < 1$

then, $\lim_{t \uparrow 1} \sum_{n=1}^{\infty} a_n t^n$ exists & equals to $\sum_{n=1}^{\infty} a_n$.

firstly, as $t \uparrow 1$, $\sum_{n=1}^{\infty} a_n t^n \uparrow$, hence, limit exists & is the supremum.

$$\therefore \lim_{t \uparrow 1} \sum_{n=1}^{\infty} a_n t^n = \sup \left\{ \sum_{n=1}^{\infty} a_n t^n, t < 1 \right\}.$$

$$\stackrel{?}{=} \sum_{n=1}^{\infty} a_n$$

One side is clear:

$$\sup \left\{ \sum_{n=1}^{\infty} a_n t^n, t < 1 \right\} \le \sum_{n=1}^{\infty} a_n$$

ie, this is an upper bound.

Take $\alpha < \sum_{n=1}^{\infty} a_n$. [if $\sum_n a_n$ diverges, take any $\alpha \in \mathbb{R}_{\ge 0}$]

we want to show, $\sum_{i=1}^{\infty} a_n$ is the lowest upper bound.

$\exists n_0$ s.t.

$$\sum_{n=1}^{N_0} a_n > \alpha \quad \text{[Archimedean Property]}$$

$$\Rightarrow \sum_{n=1}^{N_0} a_n t^n \longrightarrow \sum_{n=1}^{N_0} a_n. \text{ as } t \nearrow 1.$$

$\therefore$ for some $t \in [0,1)$,

$$\sum_{n=1}^{N_0} a_n t^n > \alpha.$$

$$\Rightarrow \sum_{n=1}^{\infty} a_n t^n > \alpha. \checkmark$$

$$\Rightarrow \sum_{n=1}^{\infty} a_n = \sup_{t \in [0,1)} \left\{ \sum_{n=1}^{\infty} a_n t^n \right\}.$$

PROBABILITY THEORY - 3
(officially begins)

(officially begins)

$(\Omega, \mathcal{A}, P)$ - Probability space.

$\{X_n\}$ - a sequence of real random variables, all on $(\Omega, \mathcal{A}, P)$.

Definition: Say that:

$\{X_n\}$ converges to X almost surely (a.s.)
or, with probability 1 (w.p. 1)

$$\text{if } P\left(\{w : X_n(w) \longrightarrow X(w)\}\right) = 1$$

$x_n \longrightarrow x$
iff $\forall \varepsilon > 0$,
$\exists\, n \in \mathbb{N}$ s.t.
$\forall m > N$,
$|x_m - x| < \varepsilon \equiv \frac{1}{j}$
(can be replaced)

Now, $\{w : X_n(w) \longrightarrow X(w)\} =$

$$\bigcap_{j \in \mathbb{N}} \bigcup_{n \in \mathbb{N}} \bigcap_{m \geq n} \left\{w : |X_m(w) - X(w)| < \frac{1}{j}\right\}$$

$= \{w :$ for every $j \in \mathbb{N}$, $\exists\, n \in \mathbb{N}$ s.t., for every $m \geq n$, $|X_m(w) - X(w)| < \frac{1}{j}\}$

Now, for every $\{w : |X_m(w) - X(w)| < \frac{1}{j}\}$,

$X_m \to$ r.v. $\quad X \to$ r.v

$|X_m - X|$: also an r.v. in $(\Omega, \mathcal{A}, P)$.

$\therefore (X_m - X)^{-1} = (-1/j, 1/j)$, which in a Borel set

$\therefore$ each such set $\in \mathcal{A}$.

$\therefore$ Countable unions & intersections of such sets $\in \mathcal{A}$.

Fact: $\left.\begin{array}{l} X_n \longrightarrow X \text{ a.s.} \\ X_n \longrightarrow Y \text{ a.s.} \end{array}\right\} X = Y$ a.s.

Proof. $X_n \longrightarrow X$ a.s. $\Rightarrow P(\underbrace{\{w : X_n(w) \longrightarrow X(w)\}}_{\text{"A"}}) = 1.$

$X \longrightarrow Y$ a.s $\Rightarrow P(\{w : X_n(w) \longrightarrow Y(w)\}) = 1.$

Proof. $X_n \longrightarrow X$ a.s. $\Rightarrow P(\{w : X_n(w) \longrightarrow X(w)\}) = 1.$

$X_n \longrightarrow Y$ a.s $\Rightarrow P(\underbrace{\{w : X_n(w) \longrightarrow Y(w)\}}_{\text{"B"}}) = 1.$

$\therefore P(A) = 1 \Rightarrow P(A^c) = 0$

$P(B) = 1 \Rightarrow P(B^c) = 0$

$\Rightarrow P(A^c \cup B^c) = 0$

$\Rightarrow P((A^c \cup B^c)^c) = 1$

$\Rightarrow P(A \cap B) = 1.$

$\Rightarrow P(\{w : X(w) \geq Y(w)\}) = 1$

&, $P(\{w : X(w) \leq Y(w)\}) = 1$

$\therefore P(\{w : X(w) = Y(w)\}) = 1$

Results:

① $X_n \xrightarrow{a.s.} X \Rightarrow c \cdot X_n \xrightarrow{a.s} cX$

Proof: $\{w : X_n(w) \longrightarrow X(w)\} \subseteq \{w : c \cdot X_n(w) \longrightarrow c \cdot X(w)\}$

$\therefore P(\{w : X_n(w) \longrightarrow X(w)\}) = 1 = P(\{w : cX_n(w) \longrightarrow cX(w)\})$

[ie, if $c = 0$, then this becomes entire Ω. $\because 0 \longrightarrow 0 \ \forall w \in \Omega$]

② $\left.\begin{array}{l} X_n \xrightarrow{a.s} X \\ Y_n \xrightarrow{a.s} Y \end{array}\right\} \Rightarrow X_n + Y_n \xrightarrow{a.s.} X + Y$

Proof: $P(\underbrace{\{w : X_n(w) \longrightarrow X(w)\}}_{A}) = 1 \quad \Rightarrow P(A) = 1.$

$P(\underbrace{\{w : Y_n(w) \longrightarrow Y(w)\}}_{B}) = 1 \quad \Rightarrow P(B) = 1.$

$\Rightarrow P(A \cap B) = 1.$

$\Rightarrow P\{w : X_n(w) + Y_n(w) \longrightarrow X(w) + Y(w)\} = 1$

$$\Rightarrow P\{w : X_n(w) + Y_n(w) \longrightarrow X(w) + Y(w)\} = 1$$

$$\Rightarrow X_n + Y_n \longrightarrow X + Y \text{ a.s.}$$

(3) $X_n \xrightarrow{a.s.} X \Rightarrow f(X_n) \xrightarrow{a.s} f(X)$

if f — continuous.

Proof: Exercise.

(4) $X_n \xrightarrow{a.s.} X$, $P(X=0) = 0$. ie, $P(X \neq 0) = 1$.

ie, X is non-zero a.s.

ie, $\frac{1}{X}$ is defined a.s.

then, $\frac{1}{X_n} \xrightarrow{a.s} \frac{1}{X}$.

Proof: Exercise.

[Here, note: $\frac{1}{X_n}$ are not "just" real r.v, but extended real r.v.]