Lec 7

Friday, 8 August 2025 10:07 AM

$\underset{\sim}{x} \to$ single r.v.

x_1
$\vdots$
x_n

$\bar{x} = \frac{1}{n} \sum_{i=1}^{n} x_i$

$E(\bar{x}) = \mu$ ← pop mean.

↑ sample mean

$s^2 = Var(\underset{\sim}{x})$ $\quad s^2 = \frac{1}{n-1} \sum_{i=1}^{n} (x_i - \bar{x})^2$.

$E(s^2) = \sigma^2$

$\underset{\sim}{x} = \begin{bmatrix} x_1 \\ \vdots \\ x_p \end{bmatrix}$ $\quad \bar{x} = \begin{bmatrix} \bar{x}_1 \\ \vdots \\ \bar{x}_p \end{bmatrix}$. $\quad S = [s_{ij}]$

$s_{ij} = \frac{1}{(n-1)} \sum_{k=1}^{n} (x_{ki} - \bar{x}_i)(x_{kj} - \bar{x}_j)$.

$E(\bar{x}) = \mu_{p\times 1}$

① $E(\bar{x}) = \mu$.

pf- $\underset{\sim}{\bar{x}} = \begin{bmatrix} \bar{x}_1 \\ \vdots \\ \bar{x}_p \end{bmatrix}$

$\bar{x}_i = \frac{1}{n} \sum_{i=1}^{n} x_{ki}$ Sample is i.i.d.

$E(\bar{x}_i) = \frac{1}{n} \sum_{i=1}^{n} E(x_{ki})$.

$= \frac{1}{n} \left(E(x_{1i}) + E(x_{2i}) + \dots + E(x_{ni}) \right)$.

i.i.d. ⇒ they come from same popula^n; pdf of the popula^n " " " sample same; moments will be same.

$= \frac{1}{n} (\mu_i + \mu_i + \dots + \mu_i) = \mu$.

↑ ith feature.

$\left(E(\bar{x}) = \frac{1}{n} \sum_{k=1}^{n} E(\underset{\sim}{x}_k) \right.$

$\frac{1}{n} \left(E(x_1) + E(x_2) \right) \dots$

each sample has same mean

$= \frac{1}{n} (\mu + \dots \mu)$.

} mean of each sample → μ.

For any p, $\bar{x}$ is an unbiased estimator of $\mu \in \mathbb{R}^p$.

ⓑ $cov(\bar{x}) = \frac{\Sigma}{n}$.

↑ R.V; linear fns of r.v. (before taking a sample)

$\underset{p\times1}{cov(\underset{\sim}{\bar{x}})} = E\left[\underbrace{(\underset{\sim}{\bar{x}} - E(\underset{\sim}{\bar{x}}))}_{p\times1} \underbrace{(\underset{\sim}{\bar{x}} - E(\underset{\sim}{\bar{x}}))^T}_{1\times p} \right]$. ($p\times p$)

$(\underset{\sim}{\bar{x}}) = \begin{bmatrix} \bar{x}_1 \\ \bar{x}_2 \\ \vdots \\ \bar{x}_p \end{bmatrix}_{p\times1}$. These values change as samples change.

$[\quad]_{n\times p} [\quad]_{p\times n}$??

$E(\underset{\sim}{\bar{x}} - \mu)(\bar{x} - \mu)^T$

$= E\left(\frac{1}{n} \sum_{i=1}^{n} x_i - \mu \right) \left(\sum_{i=1}^{n} \frac{1}{n} x_i - \mu \right)$.

$= \frac{1}{n^2} E\left(\underset{p\times1}{\sum_{i=1}^{n} (x_i - \mu)} \underset{1\times p}{\sum_{i=1}^{n} (x_i - \mu)^T} \right]$

$$\begin{array}{c|ccccc} & x_1 & x_2 & \dots & x_p \\ \hline x_1 & & p\times1 & & \\ x_2 & & & & \\ \vdots & & & & \\ x_n & & & & \end{array}$$

$= \frac{1}{n^2} E\left(\sum_{i=1}^{n} \sum_{j=1}^{n} (x_i - \mu)(x_j - \mu)^T \right]$.

$= \frac{1}{n^2} E\left[\sum_{i=1}^{n} (x_i - \mu)(x_i - \mu)^T + \sum_{i=1}^{n} \sum_{\substack{j=1 \\ j \neq i}}^{n} \underbrace{(x_i - \mu)}_{p\times1}\underbrace{(x_j - \mu)^T}_{p\times1} \right]$. → sample pts kev apas me.

$= \frac{1}{n^2} \sum_{i=1}^{n} E\underset{p\times1}{(x_i - \mu)}\underset{p\times p}{(x_i - \mu)^T}$ $+ \frac{1}{n^2} \sum_{i=1}^{n} \sum_{\substack{j=1 \\ j \neq i}}^{n} E(x_i - \mu)(x_j - \mu)^T$. → $cov(x_i, x_j)$ $i \neq j$.

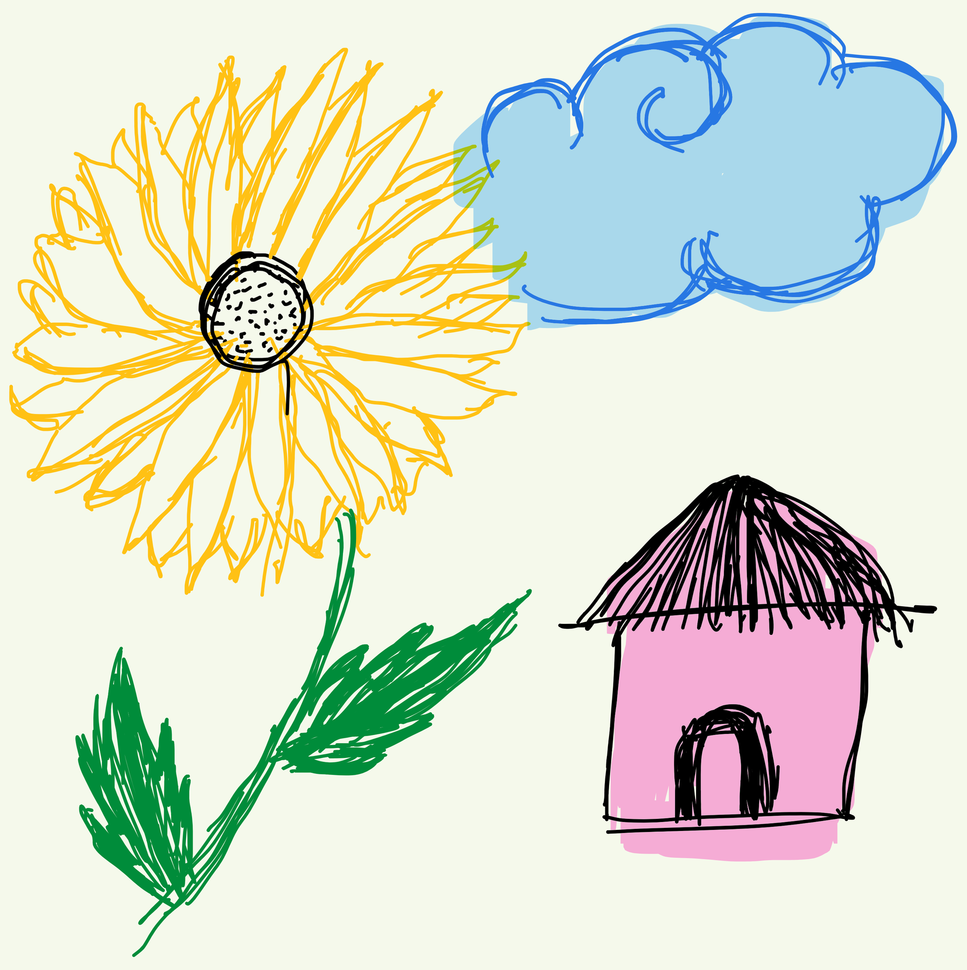
identically distributed used; now if two samples are independent ⇒ correlation = 0.

∵ Samples are independent so $\underset{\sim}{x_i}$ & $\underset{\sim}{x_j}$ & two independent R.V.s. in $\mathbb{R}^p$.

pop. mean.

$cov(\underset{\sim}{\bar{x}}) = \frac{1}{n^2} \sum_{i=1}^{n} E\left[\underbrace{\underset{p\times1}{(x_i - \mu)}\underset{p\times1}{(x_i - \mu)^T}}_{p\times p} \right]$. $p\times1$ deterministic no randomness.

$= \frac{1}{n^2} \sum_{i=1}^{n} \Sigma = \frac{\Sigma}{n}$.



eg $p = 2$. Var - cov matrix 2×2.

$$\begin{array}{c} \\ x_1 \\ x_2 \\ x_3 \end{array} \begin{array}{c} \begin{array}{cc} x_1 & x_2 \end{array} \\ \begin{bmatrix} a & \alpha \\ b & \beta \\ c & \gamma \end{bmatrix}_{3\times2} \end{array} \quad \begin{array}{c} \\ x_1 \\ x_2 \end{array} \begin{array}{c} \begin{array}{ccc} x_1 & x_2 & x_3 \end{array} \\ \begin{bmatrix} a & b & c \\ \alpha & \beta & \gamma \end{bmatrix}_{2\times3} \end{array} \begin{bmatrix} a & \alpha \\ b & \beta \\ c & \gamma \end{bmatrix}_{3\times2}$$

$$\begin{bmatrix} s_{11} & s_{12} \\ s_{21} & s_{22} \end{bmatrix}$$

Variation - covariance of a feature matrix