

3.1 Introduction → on slides

Ex 1: $\vec{Y}$ n samples, 1 measurement

Ex 2: $\vec{X}_1$ 1 sample, p measurements

3.2 Mean, Variance, Covariance (repetition)

$Y \sim f(y)$ continuous random variable

$$\text{Mean: } E(Y) = \mu = \int_{-\infty}^{\infty} y f(y) dy$$

Average of many samples from $f(y)$.

$$\text{Variance: } \text{Var}(Y) = \sigma^2 = E[(Y - \mu)^2]$$

average squared distance from μ .

$$\sigma = \sqrt{\sigma^2}, \text{ standard deviation}$$

$$\text{Recall: } Z = \frac{Y - \mu}{\sigma} \text{ 'standardized random variable'}$$

$$E(Z) = 0 \quad \text{Var}(Z) = 1.$$

'Standardized distance' (see examples on slides)

$$N(0,1) \quad N(0,2) \quad N(0,4)$$

$y = 2$ ← how far from mean $\mu = 0$?

$$Z = \frac{y - \mu}{\sigma}$$

$$Z_1 = \frac{2-0}{1} = 2 \quad Z_2 = \frac{2-0}{\sqrt{2}} = 1.4 \quad Z_3 = \frac{2-0}{\sqrt{4}} = 0.5$$

far

medium

close

Rules: $E(aY) = aE(Y) = a\mu$

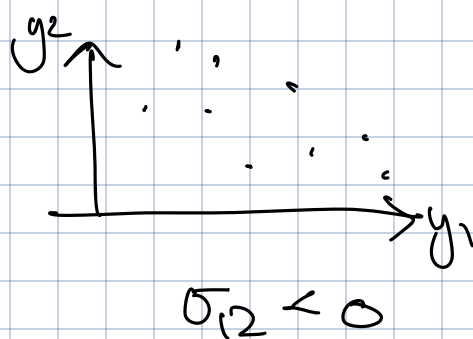
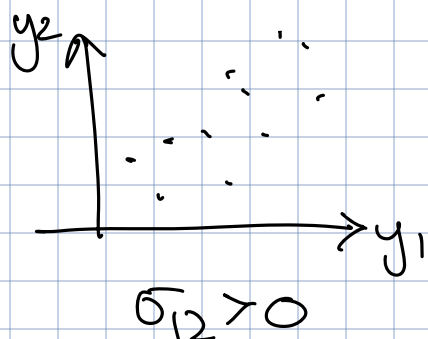
$$E(Y_1 + Y_2) = \mu_1 + \mu_2$$

$$\text{Var}(aY) = a^2 \sigma^2$$

$$\text{Var}(Y_1 + Y_2) = \sigma_1^2 + \sigma_2^2 + 2\text{Cov}(Y_1, Y_2)$$

Covariance: $\text{Cov}(Y_1, Y_2) = E((Y_1 - \mu_1)(Y_2 - \mu_2)) = \sigma_{12}$

↑ measures linear co-variation



Correlation: $\rho_{12} = \frac{\sigma_{12}}{\sigma_1 \sigma_2} \quad -1 \leq \rho_{12} \leq 1$

Examples on slide.

Note: Y_1, Y_2 independent $\rightarrow \text{Cov}(Y_1, Y_2) = 0$

(But we can see $\text{Cov}(Y_1, Y_2) = 0$ in other examples, slide).

Rule: $\sigma_{12} = E(Y_1 Y_2) - \mu_1 \mu_2$

3.3 Mean and covariance random vectors

$$\vec{Y} = \begin{pmatrix} Y_1 \\ \vdots \\ Y_n \end{pmatrix}$$

Def: mean vector

$$\vec{\mu} = \begin{pmatrix} E(Y_1) \\ \vdots \\ E(Y_n) \end{pmatrix} = \begin{pmatrix} \mu_1 \\ \vdots \\ \mu_n \end{pmatrix}$$

Also: $E(\vec{X} + \vec{Y}) = E(\vec{X}) + E(\vec{Y})$

Def: covariance matrix

$$\Sigma = \text{cov}(\vec{Y}) = \begin{bmatrix} \sigma_1^2 & \sigma_{12} & \dots & \sigma_{1n} \\ \sigma_{21} & \sigma_2^2 & & \\ \vdots & & \ddots & \\ \sigma_{n1} & \sigma_{n2} & \dots & \sigma_n^2 \end{bmatrix}$$

(n x n)

Diagonal: Variances $\sigma_i^2 = \sigma_{ii} = \text{cov}(Y_i, Y_i)$
off-diagonal: covariances $\sigma_{ij} = \sigma_{ji}$
 $= \text{cov}(Y_i, Y_j)$

Note: Σ Symmetric and

We require Σ positive definite (or semidef)

Why? For vector of scalars $\vec{a}$,

$$U = \vec{a}^T \vec{Y} \text{ univariate random variable}$$

$$U = a_1 Y_1 + \dots + a_n Y_n.$$

3.6 (Thursday) : $\text{Var}(v) = \underbrace{\vec{a}^T \Sigma \vec{a}}_{> 0 \text{ if } \Sigma \text{ pos. def.}}$

↗
should not
be negative.

Rule: $\text{Cov}(\vec{Y}) = E[(\vec{Y} - \vec{\mu})(\vec{Y} - \vec{\mu})^T] \quad (*)$

$$= E(\vec{Y} \vec{Y}^T) - \vec{\mu} \vec{\mu}^T$$

$(*)$ Expectation of random matrix:

diagonal: $(Y_i - \mu_i)^2$

off-diagonal: $(Y_i - \mu_i)(Y_j - \mu_j)$

3.3.3 Generalized variance

$$\text{Cov}(\vec{Y}) = \Sigma$$

$|\Sigma| = \det(\Sigma)$ measures 'overall variability'

both variability in each y_i ,
but also captures correlation,
high correlation means less overall
variability.

$$\text{Ex: a) } \vec{Y} = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} \quad \Sigma = \begin{bmatrix} 1 & 1.5 \\ 1.5 & 4 \end{bmatrix}$$

What is ρ_{12} ?

What is $|\Sigma|$?

$$\rho_{12} = \frac{\sigma_{12}}{\sigma_1 \sigma_2} = \frac{1.5}{1 \cdot \sqrt{4}} = \frac{1.5}{2} = 0.75$$

$$|\Sigma| = 4 - 1.5^2 = 4 - 2.25 = 1.75$$

$$\text{b) } \Sigma = \begin{bmatrix} 1 & 0.5 \\ 0.5 & 4 \end{bmatrix}$$

[SLIDE]

$$\rho_{12} = \frac{0.5}{2} = 0.25, \quad |\Sigma| = 4 - 0.5^2 = 3.75$$

3.3.4 Standardized distance (++)

Recall univariate: $z = \frac{y - \mu}{\sigma}$ ← distance of y from μ , comparable across different σ .

also $z^2 = \left(\frac{y - \mu}{\sigma}\right)^2$ (directionless)

Multivariate (vector $\vec{y}$)

$$d^2 = (\vec{y} - \vec{\mu})^T \Sigma^{-1} (\vec{y} - \vec{\mu})$$

'Mahalanobis distance'

Example (if time):

$$\vec{y} = \begin{pmatrix} 1 \\ -4 \end{pmatrix} \quad a) \quad \vec{\mu} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \quad \Sigma = \begin{bmatrix} 1 & 1.5 \\ 1.5 & 4 \end{bmatrix}$$

$$b) \quad \vec{\mu} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \quad \Sigma = \begin{bmatrix} 1 & 0.5 \\ 0.5 & 4 \end{bmatrix}$$

$$d_1^2 \approx 18$$

$$d_2^2 \approx 6.4$$

see slide

Also, standardized random var.

Univariate: $Z = \frac{Y - \mu}{\sigma}$ $E(Z) = 0$, $\text{Var}(Z) = 1$.

Multivariate: $\vec{Z} = \Sigma^{-1/2} (\vec{Y} - \vec{\mu})$ more on
thursday.

$$E(\vec{Z}) = \vec{0} \quad \text{Cov}(\vec{Z}) = \mathbf{I}$$

$$\Sigma^{-1/2} ?$$

Spectral decomposition of Σ :

$$\Sigma = C D C^T$$

$$D = \text{diag}(\lambda_1, \dots, \lambda_n)$$

λ_i positive since Σ pos. def.

$$\text{Def: } \Sigma^{1/2} = C D^{1/2} C^T$$

$$\text{and } \Sigma^{-1/2} = C D^{-1/2} C^T$$

$$D^{-1/2} = \text{diag}\left(\frac{1}{\sqrt{\lambda_1}}, \dots, \frac{1}{\sqrt{\lambda_n}}\right)$$

drop 2
if you want
repetition.