30 september 2023 - lec. 1

BIVARIATE RANDOM VARIABLES

we extend the concept of random variable to 2 dimensions
bivariate random variable $(X,Y): \Omega \longrightarrow \mathbb{R}^2$

- The CDF is now a function $F_{(X,Y)}: \mathbb{R}^2 \longrightarrow [0,1]$

  $F_{(X,Y)}(x,y) = \mathbb{P}((X,Y) \in (-\infty,x) \times (-\infty,y)) = \mathbb{P}(X \leq x, Y \leq y)$

- discrete r.v.'s:

  joint probability function $\quad p_{(X,Y)}(x,y) = \mathbb{P}(X=x, Y=y)$

  marginal probability function $\quad p_X(s) = \mathbb{P}(X=s) = \sum_{y \in S_Y} \mathbb{P}(X=s, Y=y)$

- continuous r.v.'s:

  joint density function $\quad f_{(X,Y)}(x,y)$

  marginal density function $\quad f_X(x) = \int_{-\infty}^{+\infty} f_{(X,Y)}(x,v)\, dv$

- INDEPENDENCE: Two r.v.'s X and Y are independent $(X \perp\!\!\!\perp Y)$

  $\iff f_{(X,Y)}(x,y) = f_X(x) \cdot f_Y(y)$ (continuous case)

  $\iff p_{(X,Y)}(x,y) = p_X(x) \cdot p_Y(y)$ i.e. $\mathbb{P}(X=x, Y=y) = \mathbb{P}(X=x) \cdot \mathbb{P}(Y=y)$ (discrete case)

- COVARIANCE between X and Y: $\quad cov(X,Y) = \sigma_{XY} = \mathbb{E}[(X-\mathbb{E}[X])(Y-\mathbb{E}[Y])]$

  it expresses how the two variables change together

- CORRELATION: $corr(X,Y) = \rho_{XY} = \dfrac{cov(X,Y)}{\sqrt{var(X)\, var(Y)}} \in [-1,1]$

we can extend these concepts to a generic dimension $d \geq 1$.

MULTIVARIATE RANDOM VARIABLES (RANDOM VECTORS)

A multivariate r.v. is a column vector $\underline{X} = [X_1\ X_2\ \dots\ X_d]^T$ whose components are r.v.'s

$[X_1 \dots X_d]^T: \Omega \longrightarrow \mathbb{R}^d$

- CDF $\quad F_{\underline{X}}: \mathbb{R}^d \longrightarrow [0,1]$

  $F_{\underline{X}}(\underline{x}) = \mathbb{P}(X_1 \leq x_1, X_2 \leq x_2, \dots, X_d \leq x_d)$

- EXPECTED VALUE: $\mathbb{E}[\underline{X}] = \begin{bmatrix} \mathbb{E}[X_1] \\ \vdots \\ \mathbb{E}[X_d] \end{bmatrix}$ d-dimensional vector

- COVARIANCE MATRIX

$$\begin{aligned} var(\underline{X}) &= \mathbb{E}[(\underline{X}-\mathbb{E}[\underline{X}])(\underline{X}-\mathbb{E}[\underline{X}])^T] = \\ &= \mathbb{E}[\underline{X}\,\underline{X}^T - \underline{X}\,\mathbb{E}[\underline{X}]^T - \mathbb{E}[\underline{X}]\,\underline{X}^T + \mathbb{E}[\underline{X}]\,\mathbb{E}[\underline{X}]^T] = \\ &= \mathbb{E}[\underline{X}\,\underline{X}^T] - \mathbb{E}[\underline{X}]\,\mathbb{E}[\underline{X}]^T - \cancel{\mathbb{E}[\underline{X}]\,\mathbb{E}[\underline{X}^T]} + \cancel{\mathbb{E}[\underline{X}]\,\mathbb{E}[\underline{X}]^T} \\ &= \mathbb{E}[\underset{d\times 1}{\underline{X}}\,\underset{1\times d}{\underline{X}^T}] - \underset{d\times 1}{\mathbb{E}[\underline{X}]}\,\underset{1\times d}{\mathbb{E}[\underline{X}]^T} \Rightarrow d \times d \text{ matrix} \end{aligned}$$

what are the elements of this matrix?

$$\mathbb{E}\left[\begin{bmatrix} X_1-\mathbb{E}[X_1] \\ X_2-\mathbb{E}[X_2] \\ \vdots \\ X_d-\mathbb{E}[X_d] \end{bmatrix} \begin{bmatrix} X_1-\mathbb{E}[X_1] & X_2-\mathbb{E}[X_2] & \dots & X_d-\mathbb{E}[X_d] \end{bmatrix} \right]$$

$$= \mathbb{E}\left[\begin{bmatrix} (X_1-\mathbb{E}[X_1])^2 & (X_1-\mathbb{E}[X_1])(X_2-\mathbb{E}[X_2]) & \dots & (X_1-\mathbb{E}[X_1])(X_d-\mathbb{E}[X_d]) \\ (X_2-\mathbb{E}[X_2])(X_1-\mathbb{E}[X_1]) & (X_2-\mathbb{E}[X_2])^2 & \dots & (X_2-\mathbb{E}[X_2])(X_d-\mathbb{E}[X_d]) \\ \vdots & & \ddots & \vdots \\ (X_d-\mathbb{E}[X_d])(X_1-\mathbb{E}[X_1]) & (X_d-\mathbb{E}[X_d])(X_2-\mathbb{E}[X_2]) & & (X_d-\mathbb{E}[X_d])^2 \end{bmatrix} \right] =$$

$$= \begin{bmatrix} var(X_1) & cov(X_1,X_2) & \dots & cov(X_1,X_d) \\ cov(X_2,X_1) & var(X_2) & \dots & cov(X_2,X_d) \\ \vdots & \vdots & \ddots & \vdots \\ cov(X_d,X_1) & cov(X_d,X_2) & \dots & var(X_d) \end{bmatrix} = \begin{bmatrix} \sigma_1^2 & \sigma_{12} & \dots & \sigma_{1d} \\ \sigma_{21} & \sigma_2^2 & \dots & \sigma_{2d} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{d1} & \sigma_{d2} & \dots & \sigma_d^2 \end{bmatrix}$$

symmetric
positive semi-definite

MULTIVARIATE NORMAL DISTRIBUTION

generalization of the normal distribution to d dimensions

$\underline{X} = [X_1 \dots X_d]^T \sim N_d(\underline{\mu}, \Sigma)$

- support $\quad S_X = \mathbb{R}^d$
- parameters:
  - expected value $\quad \underline{\mu} = \mathbb{E}[\underline{X}] = [\mathbb{E}[X_1] \dots \mathbb{E}[X_d]]^T$ $\quad d$-dim vector
  - covariance matrix $\quad \Sigma = var(\underline{X})$ $\quad d \times d$ matrix
- density function

  $\phi_{\underline{X}}(x_1, \dots, x_d) = (2\pi)^{-d/2} \det(\Sigma)^{-1/2} \exp\left\{-\frac{1}{2}(\underline{x}-\underline{\mu})^T \Sigma^{-1} (\underline{x}-\underline{\mu})\right\}$

- marginal distributions: example $\quad [X_1, X_2, X_3]^T \sim N_3\left(\underline{\mu} = \begin{bmatrix} \mu_1 \\ \mu_2 \\ \mu_3 \end{bmatrix}, \Sigma = \begin{bmatrix} \sigma_1^2 & \sigma_{12} & \sigma_{13} \\ \sigma_{21} & \sigma_2^2 & \sigma_{23} \\ \sigma_{31} & \sigma_{32} & \sigma_3^2 \end{bmatrix}\right)$

  the marginal distributions are simply obtained by looking only at the components we are considering

  e.g.

  $X_1 \sim N_1(\mu_1, \sigma_1^2)$

  $[X_1, X_3]^T \sim N_2\left(\begin{bmatrix} \mu_1 \\ \mu_3 \end{bmatrix}, \begin{bmatrix} \sigma_1^2 & \sigma_{13} \\ \sigma_{31} & \sigma_3^2 \end{bmatrix}\right)$

MULTIVARIATE STANDARD NORMAL

$\underline{\mu} = \underline{0}$ and $\Sigma = I_d$ $\qquad \underline{Z} \sim N_d(\underline{0}, I_d)$

in this case, the Z_i $(i=1,\dots,d)$ are independent normal r.v.'s $Z_i \sim N(0,1)$

- general normal $\underline{X} \sim N_d(\underline{\mu}, \Sigma)$ from the standard normal $\underline{Z} \sim N_K(\underline{0}, I_K)$

  $\underline{\mu} \in \mathbb{R}^d$ d-dim vector, $\quad A$ $d \times k$ matrix such that $\Sigma = AA^T$

  $\underline{X} = A\underline{Z} + \underline{\mu} \quad \Rightarrow \quad \underline{X} \sim N_d(\underline{\mu}, \Sigma)$

  1) linear transformation of a normal r.v. is normal
  2) $\mathbb{E}[A\underline{Z} + \underline{\mu}] = A\,\cancel{\mathbb{E}[\underline{Z}]} + \underline{\mu} = \underline{\mu}$
  3) $var(A\underline{Z} + \underline{\mu}) = var(A\underline{Z}) = \mathbb{E}[(A\underline{Z} - A\mathbb{E}[\underline{Z}])(A\underline{Z} - A\mathbb{E}[\underline{Z}])^T] =$

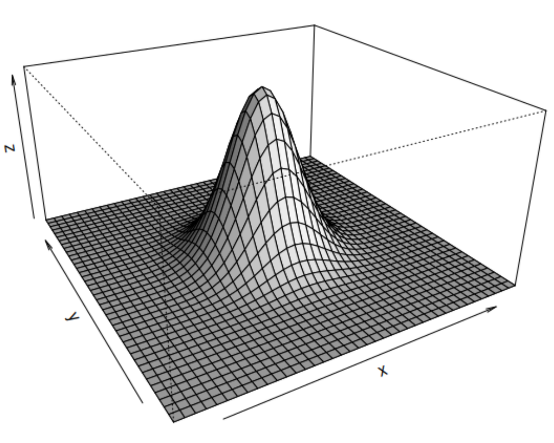
$$\begin{aligned} &= \mathbb{E}[A\underline{Z}\,\underline{Z}^T A^T - A\underline{Z}\,\mathbb{E}[\underline{Z}]^T A^T - A\mathbb{E}[\underline{Z}]\,\underline{Z}^T A^T + A\mathbb{E}[\underline{Z}]\,\mathbb{E}[\underline{Z}]^T A^T] = \\ &= A\,\mathbb{E}[\underline{Z}\,\underline{Z}^T]A^T - A\,\mathbb{E}[\underline{Z}]\,\mathbb{E}[\underline{Z}]^T A^T - A\,\cancel{\mathbb{E}[\underline{Z}]\,\mathbb{E}[\underline{Z}]^T} A^T + A\,\cancel{\mathbb{E}[\underline{Z}]\,\mathbb{E}[\underline{Z}]^T} A^T \\ &= A\,(\mathbb{E}[\underline{Z}\,\underline{Z}^T] - \mathbb{E}[\underline{Z}]\,\mathbb{E}[\underline{Z}]^T)\,A^T \\ &= A\,\underset{\underset{I_K}{\parallel}}{var(\underline{Z})}\,A^T = AA^T = \Sigma \end{aligned}$$

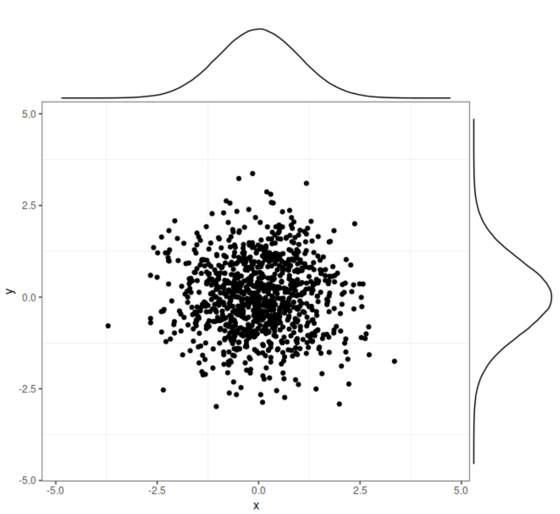
BIVARIATE NORMAL (d=2)

$$\begin{bmatrix} X_1 \\ X_2 \end{bmatrix} \sim N_2\left(\begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix}, \Sigma = \begin{bmatrix} \sigma_1^2 & \sigma_{12} \\ \sigma_{21} & \sigma_2^2 \end{bmatrix}\right)$$

EXAMPLES

$(X_1, X_2)^T \sim N_2(\underline{0}, I_2)$

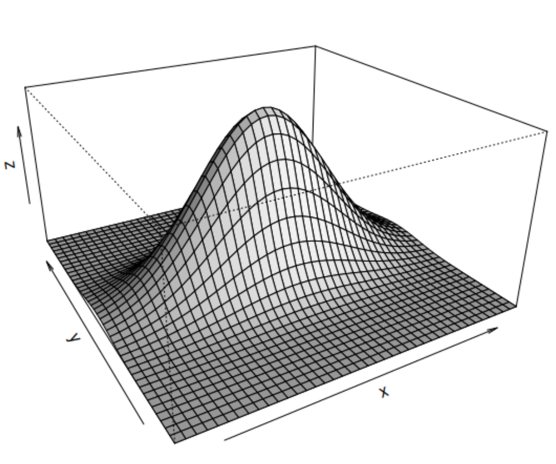


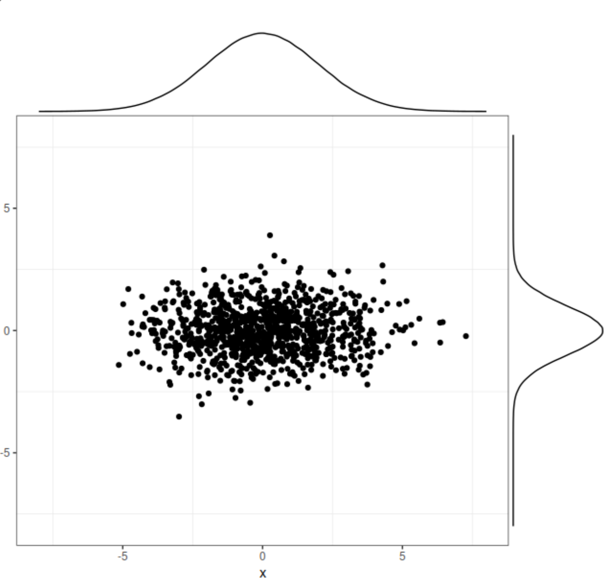


$X_1 \sim N(0,1) \qquad X_2 \sim N(0,1)$

$[X_1, X_2]^T \sim N$ and $cov(X_1, X_2) = 0 \Rightarrow X_1 \perp\!\!\!\perp X_2$

$(X_1, X_2)^T \sim N_2\left(\underline{0}, \Sigma = \begin{bmatrix} 4 & 0 \\ 0 & 1 \end{bmatrix}\right)$



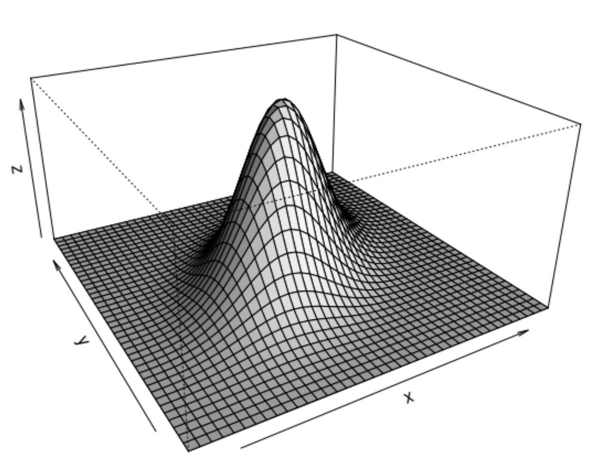


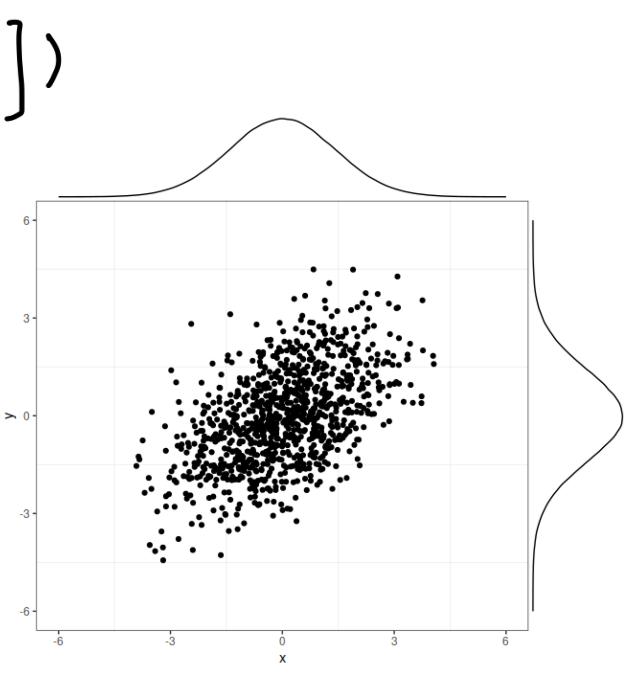
"wider" in the first component

$X_1 \sim N(0,4) \qquad X_2 \sim N(0,1)$

$X_1 \perp\!\!\!\perp X_2$

$(X_1, X_2)^T \sim N_2\left(\underline{0}, \Sigma = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}\right)$





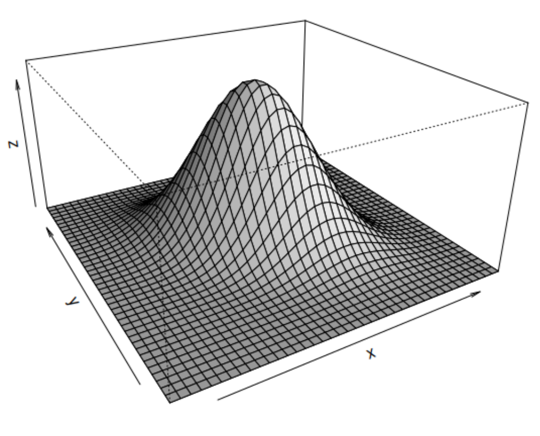
"oblique"

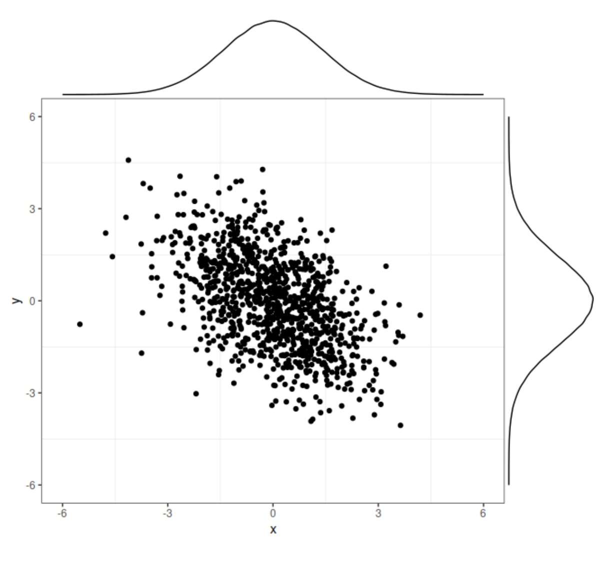
$X_1 \sim N(0,2) \qquad X_2 \sim N(0,2)$ same variance

they are not independent!

positive correlation: at large values of X_1 we expect large values of X_2

$(X_1, X_2)^T \sim N_2\left(\underline{0}, \Sigma = \begin{bmatrix} 2 & -1 \\ -1 & 2 \end{bmatrix}\right)$





$X_1 \sim N(0,2) \qquad X_2 \sim N(0,2)$ same variance

they are not independent!

negative correlation: at large values of X_1 we expect small values of X_2