

# Probabilità e Processi Stocastici (455AA)

## Lezione 8

Dario Trevisan – <https://web.dm.unipi.it/trevisan>

15/10/2025

## **Varianza e deviazione standard**

# Varianza

La moda, la mediana ed il valore medio sono tutti indicatori *puntuali*, riassumono tutta la legge con un singolo valore.

Per descrivere in modo più efficace una variabile  $X$  si affianca un indicatore della dispersione, ossia della “concentrazione” della sua legge intorno ad un indicatore puntuale.

- Un indicatore di dispersione molto usato è la *deviazione standard* (o scarto quadratico medio), definita come

$$\sigma_X = \sqrt{\text{Var}(X)},$$

dove la *varianza* è

$$\text{Var}(X) = \mathbb{E} \left[ (X - \mathbb{E}[X])^2 \right].$$

L'unità di misura di  $\sigma_X$  è la stessa di  $X$ , mentre  $\text{Var}(X)$  ha come unità di misura il quadrato dell'unità di  $X$ .

Operativamente:

1. si calcola il valor medio di  $X$ ,

L'unità di misura di  $\sigma_X$  è la stessa di  $X$ , mentre  $\text{Var}(X)$  ha come unità di misura il quadrato dell'unità di  $X$ .

Operativamente:

1. si calcola il valor medio di  $X$ ,
2. si considera lo scarto (variabile aleatoria centrata)

$$X - \mathbb{E}[X],$$

L'unità di misura di  $\sigma_X$  è la stessa di  $X$ , mentre  $\text{Var}(X)$  ha come unità di misura il quadrato dell'unità di  $X$ .

Operativamente:

1. si calcola il valor medio di  $X$ ,
2. si considera lo scarto (variabile aleatoria centrata)

$$X - \mathbb{E}[X],$$

3. se ne prende il quadrato e il valor medio

$$\text{Var}(X) = \mathbb{E} \left[ (X - \mathbb{E}[X])^2 \right].$$

L'unità di misura di  $\sigma_X$  è la stessa di  $X$ , mentre  $\text{Var}(X)$  ha come unità di misura il quadrato dell'unità di  $X$ .

Operativamente:

1. si calcola il valor medio di  $X$ ,
2. si considera lo scarto (variabile aleatoria centrata)

$$X - \mathbb{E}[X],$$

3. se ne prende il quadrato e il valor medio

$$\text{Var}(X) = \mathbb{E}[(X - \mathbb{E}[X])^2].$$

4. Per la deviazione standard si prende la radice quadrata

$$\sigma_X = \sqrt{\mathbb{E}[(X - \mathbb{E}[X])^2]}.$$

$$\text{Var}(X) = \mathbb{E} \left[ (X - \mathbb{E}[X])^2 \right].$$

Partendo dalla densità di  $X$  e usando  $g(x) = (x - \mathbb{E}[X])^2$ , si trova la formula

$$\text{Var}(X) = \begin{cases} \sum_{x \in \mathbb{R}} (x - \mathbb{E}[X])^2 P(X = x) & \text{se } X \text{ ha densità discreta} \\ \int_{-\infty}^{\infty} (x - \mathbb{E}[X])^2 p(X = x) dx & \text{se } X \text{ ha densità continua.} \end{cases}$$



## Espressione alternativa per la varianza

Vale l'identità

$$\text{Var}(X) = \mathbb{E}[X^2] - (\mathbb{E}[X])^2.$$

# Dimostrazione

## Un altro esempio

# Diseguaglianza di Chebyshev

Sia  $X \in \mathbb{R}$  una variabile aleatoria. Allora per ogni costante  $k > 0$ , si ha

$$P(|X - \mathbb{E}[X]| > k) \leq \frac{\text{Var}(X)}{k^2},$$

► o, equivalentemente, per ogni  $k \geq 1$ ,

$$P(\mathbb{E}[X] - k\sigma_X \leq X \leq \mathbb{E}[X] + k\sigma_X) \geq 1 - \frac{1}{k^2}.$$

# Diseguaglianza di Chebyshev

Sia  $X \in \mathbb{R}$  una variabile aleatoria. Allora per ogni costante  $k > 0$ , si ha

$$P(|X - \mathbb{E}[X]| > k) \leq \frac{\text{Var}(X)}{k^2},$$

► o, equivalentemente, per ogni  $k \geq 1$ ,

$$P(\mathbb{E}[X] - k\sigma_X \leq X \leq \mathbb{E}[X] + k\sigma_X) \geq 1 - \frac{1}{k^2}.$$

► Una conseguenza è che, se  $\sigma_X = 0$  la variabile  $X$  è costante con probabilità 1

# Disuguaglianza di Chebyshev

Sia  $X \in \mathbb{R}$  una variabile aleatoria. Allora per ogni costante  $k > 0$ , si ha

$$P(|X - \mathbb{E}[X]| > k) \leq \frac{\text{Var}(X)}{k^2},$$

- ▶ o, equivalentemente, per ogni  $k \geq 1$ ,

$$P(\mathbb{E}[X] - k\sigma_X \leq X \leq \mathbb{E}[X] + k\sigma_X) \geq 1 - \frac{1}{k^2}.$$

- ▶ Una conseguenza è che, se  $\sigma_X = 0$  la variabile  $X$  è costante con probabilità 1
- ▶ Informalmente scriviamo

$$X \approx \mathbb{E}[X] \pm \sigma_X$$



Data  $X$ , la variabile  $Y = X - \mathbb{E}[X]$  è **centrata**, ossia  $\mathbb{E}[Y] = 0$ :

- La **standardizzazione**  $\hat{X}$  è la variabile

$$\hat{X} = \frac{X - \mathbb{E}[X]}{\sigma_X},$$

che è centrata e ha deviazione standard  $\sigma_{\hat{X}} = 1$ .









## Covarianza

# La dispersione nel caso vettoriale

- ▶ L'estensione del valor medio al caso vettoriale  $X \in \mathbb{R}^d$  è immediata:

$$\mathbb{E}[X] = (\mathbb{E}[X_1], \mathbb{E}[X_2], \dots, \mathbb{E}[X_d]).$$

# La dispersione nel caso vettoriale

- ▶ L'estensione del valor medio al caso vettoriale  $X \in \mathbb{R}^d$  è immediata:

$$\mathbb{E}[X] = (\mathbb{E}[X_1], \mathbb{E}[X_2], \dots, \mathbb{E}[X_d]).$$

- ▶ Estensione della varianza? Le varianze delle componenti non sono “sufficienti” a descrivere la dispersione della legge.

## Esempio

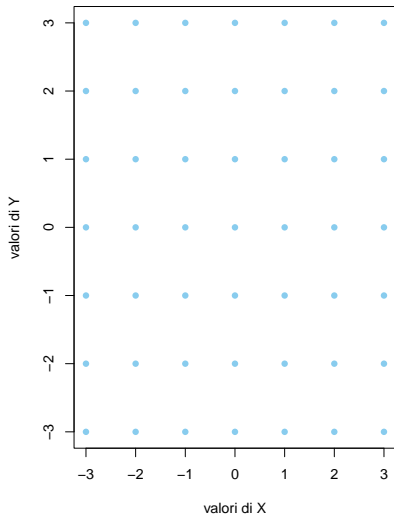
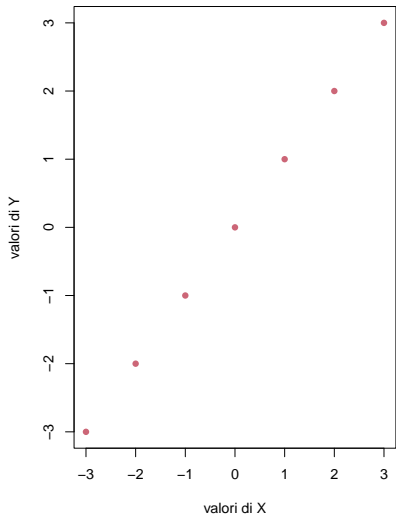
Si considerino due variabili  $X, Y$  uniformi discrete sui valori  $\{-3, -2, -1, 0, 1, 2, 3\}$  (in modo che siano già centrate). La deviazione standard risulta  $\sigma_X = \sigma_Y = 2$ .

Ma non abbiamo indicazioni sulla la “dispersione” della variabile congiunta  $(X, Y)$ :

- ▶ potrebbe essere noto che  $X = Y$ , e quindi la densità discreta della variabile congiunta è “concentrata” sulla diagonale principale;

Ma non abbiamo indicazioni sulla la “dispersione” della variabile congiunta  $(X, Y)$ :

- ▶ potrebbe essere noto che  $X = Y$ , e quindi la densità discreta della variabile congiunta è “concentrata” sulla diagonale principale;
- ▶ oppure le due variabili potrebbero essere indipendenti, e quindi la densità è “diffusa” su tutte le possibili coppie di valori.



# La covarianza: definizione e proprietà

Date due variabili aleatorie reali  $X$ ,  $Y$ , si definisce la covarianza come la quantità (reale)

$$\text{Cov}(X, Y) = \mathbb{E}[(X - \mathbb{E}[X])(Y - \mathbb{E}[Y])].$$

- La covarianza è una estensione della varianza

# La covarianza: definizione e proprietà

Date due variabili aleatorie reali  $X$ ,  $Y$ , si definisce la covarianza come la quantità (reale)

$$\text{Cov}(X, Y) = \mathbb{E}[(X - \mathbb{E}[X])(Y - \mathbb{E}[Y])].$$

- ▶ La covarianza è una estensione della varianza
- ▶ è una funzione *bilineare* (ossia separatamente lineare) di  $X$  e  $Y$

# La covarianza: definizione e proprietà

Date due variabili aleatorie reali  $X$ ,  $Y$ , si definisce la covarianza come la quantità (reale)

$$\text{Cov}(X, Y) = \mathbb{E}[(X - \mathbb{E}[X])(Y - \mathbb{E}[Y])].$$

- ▶ La covarianza è una estensione della varianza
- ▶ è una funzione *bilineare* (ossia separatamente lineare) di  $X$  e  $Y$
- ▶ A volte si indica anche con  $K_{XY}$

Date variabili aleatorie reali  $X$ ,  $Y$ ,  $Z$  e una costante  $a > 0$ , valgono le seguenti proprietà:

1.  $\text{Cov}(X, X) = \text{Var}(X)$

Date variabili aleatorie reali  $X$ ,  $Y$ ,  $Z$  e una costante  $a > 0$ , valgono le seguenti proprietà:

1.  $\text{Cov}(X, X) = \text{Var}(X)$
2. (simmetria)  $\text{Cov}(X, Y) = \text{Cov}(Y, X)$

Date variabili aleatorie reali  $X$ ,  $Y$ ,  $Z$  e una costante  $a > 0$ , valgono le seguenti proprietà:

1.  $\text{Cov}(X, X) = \text{Var}(X)$
2. (simmetria)  $\text{Cov}(X, Y) = \text{Cov}(Y, X)$
3. (bilinearità)  $\text{Cov}(X + Z, Y) = \text{Cov}(X, Y) + \text{Cov}(Z, Y)$  e similmente  $\text{Cov}(X, Y + Z) = \text{Cov}(X, Y) + \text{Cov}(X, Z)$ .  
Inoltre  $\text{Cov}(aX, Y) = a \text{Cov}(X, Y) = \text{Cov}(X, aY)$ .

Date variabili aleatorie reali  $X$ ,  $Y$ ,  $Z$  e una costante  $a > 0$ , valgono le seguenti proprietà:

1.  $\text{Cov}(X, X) = \text{Var}(X)$
2. (simmetria)  $\text{Cov}(X, Y) = \text{Cov}(Y, X)$
3. (bilinearità)  $\text{Cov}(X + Z, Y) = \text{Cov}(X, Y) + \text{Cov}(Z, Y)$  e similmente  $\text{Cov}(X, Y + Z) = \text{Cov}(X, Y) + \text{Cov}(X, Z)$ .  
Inoltre  $\text{Cov}(aX, Y) = a \text{Cov}(X, Y) = \text{Cov}(X, aY)$ .
4. (varianza della somma)  
 $\text{Var}(X + Y) = \text{Var}(X) + \text{Var}(Y) + 2 \text{Cov}(X, Y)$ .

Date variabili aleatorie reali  $X$ ,  $Y$ ,  $Z$  e una costante  $a > 0$ , valgono le seguenti proprietà:

1.  $\text{Cov}(X, X) = \text{Var}(X)$
2. (simmetria)  $\text{Cov}(X, Y) = \text{Cov}(Y, X)$
3. (bilinearità)  $\text{Cov}(X + Z, Y) = \text{Cov}(X, Y) + \text{Cov}(Z, Y)$  e similmente  $\text{Cov}(X, Y + Z) = \text{Cov}(X, Y) + \text{Cov}(X, Z)$ .  
Inoltre  $\text{Cov}(aX, Y) = a \text{Cov}(X, Y) = \text{Cov}(X, aY)$ .
4. (varianza della somma)  
 $\text{Var}(X + Y) = \text{Var}(X) + \text{Var}(Y) + 2 \text{Cov}(X, Y)$ .
5. (formula alternativa)  $\text{Cov}(X, Y) = \mathbb{E}[XY] - \mathbb{E}[X] \mathbb{E}[Y]$ .



# Indipendenza e covarianza

Se due variabili reali  $X$ ,  $Y$  sono indipendenti (rispetto ad una informazione  $I$ ), allora sono *non correlate*, ossia

$$\text{Cov}(X, Y) = 0.$$

► In particolare,

$$\text{Var}(X + Y) = \text{Var}(X) + \text{Var}(Y).$$

Questo fatto segue dalla formula alternativa per la covarianza

$$\text{Cov}(X, Y) = \mathbb{E}[XY] - \mathbb{E}[X]\mathbb{E}[Y]$$

- e la proprietà vista del valor medio del prodotto di variabili indipendenti,

$$\mathbb{E}[XY] = \mathbb{E}[X]\mathbb{E}[Y].$$

## Segno della covarianza

- ▶ Se  $X$ ,  $Y$  sono indipendenti, allora sono *non correlate*, ossia  $\text{Cov}(X, Y) = 0$

# Segno della covarianza

- ▶ Se  $X$ ,  $Y$  sono indipendenti, allora sono *non correlate*, ossia  $\text{Cov}(X, Y) = 0$
- ▶ Il viceversa non è vero in generale: se  $\text{Cov}(X, Y) = 0$ , non necessariamente  $X$ ,  $Y$  sono indipendenti.

## Segno della covarianza

- ▶ Se  $X, Y$  sono indipendenti, allora sono *non correlate*, ossia  $\text{Cov}(X, Y) = 0$
- ▶ Il viceversa non è vero in generale: se  $\text{Cov}(X, Y) = 0$ , non necessariamente  $X, Y$  sono indipendenti.
- ▶ Se  $\text{Cov}(X, Y) > 0$ ,  $X, Y$  sono *positivamente* correlate

# Segno della covarianza

- ▶ Se  $X, Y$  sono indipendenti, allora sono *non correlate*, ossia  $\text{Cov}(X, Y) = 0$
- ▶ Il viceversa non è vero in generale: se  $\text{Cov}(X, Y) = 0$ , non necessariamente  $X, Y$  sono indipendenti.
- ▶ Se  $\text{Cov}(X, Y) > 0$ ,  $X, Y$  sono *positivamente* correlate
- ▶ Se  $\text{Cov}(X, Y) < 0$ , sono *negativamente* correlate.

## Covarianza di variabili Bernoulli

Siano  $X \in \{0, 1\}$ , indicatrice dell'evento  $A$ ,  $Y \in \{0, 1\}$  indicatrice dell'evento  $B$ .

- ▶ Allora  $XY \in \{0, 1\}$  è indicatrice di “ $A$  e  $B$ ” e

$$\text{Cov}(X, Y) = \mathbb{E}[XY] - \mathbb{E}[X]\mathbb{E}[Y] = P(A \text{ e } B) - P(A)P(B).$$

## Covarianza di variabili Bernoulli

Siano  $X \in \{0, 1\}$ , indicatrice dell'evento  $A$ ,  $Y \in \{0, 1\}$  indicatrice dell'evento  $B$ .

- ▶ Allora  $XY \in \{0, 1\}$  è indicatrice di “ $A$  e  $B$ ” e

$$\text{Cov}(X, Y) = \mathbb{E}[XY] - \mathbb{E}[X]\mathbb{E}[Y] = P(A \text{ e } B) - P(A)P(B).$$

- ▶  $X$  e  $Y$  sono positivamente correlate se e solo se

$$P(A \text{ e } B) - P(A)P(B) > 0, \quad \text{ossia} \quad \frac{P(A \text{ e } B)}{P(A)P(B)} > 1,$$

## Covarianza di variabili Bernoulli

Siano  $X \in \{0, 1\}$ , indicatrice dell'evento  $A$ ,  $Y \in \{0, 1\}$  indicatrice dell'evento  $B$ .

- ▶ Allora  $XY \in \{0, 1\}$  è indicatrice di “ $A$  e  $B$ ” e

$$\text{Cov}(X, Y) = \mathbb{E}[XY] - \mathbb{E}[X]\mathbb{E}[Y] = P(A \text{ e } B) - P(A)P(B).$$

- ▶  $X$  e  $Y$  sono positivamente correlate se e solo se

$$P(A \text{ e } B) - P(A)P(B) > 0, \quad \text{ossia} \quad \frac{P(A \text{ e } B)}{P(A)P(B)} > 1,$$

- ▶ ossia  $P(A|B) > P(A)$  ( $B$  è una informazione a favore di  $A$ )

# Covarianza di variabili Bernoulli

Siano  $X \in \{0, 1\}$ , indicatrice dell'evento  $A$ ,  $Y \in \{0, 1\}$  indicatrice dell'evento  $B$ .

- ▶ Allora  $XY \in \{0, 1\}$  è indicatrice di “ $A$  e  $B$ ” e

$$\text{Cov}(X, Y) = \mathbb{E}[XY] - \mathbb{E}[X]\mathbb{E}[Y] = P(A \text{ e } B) - P(A)P(B).$$

- ▶  $X$  e  $Y$  sono positivamente correlate se e solo se

$$P(A \text{ e } B) - P(A)P(B) > 0, \quad \text{ossia} \quad \frac{P(A \text{ e } B)}{P(A)P(B)} > 1,$$

- ▶ ossia  $P(A|B) > P(A)$  ( $B$  è una informazione a favore di  $A$ )
- ▶ sono negativamente correlate se e solo se il rapporto è  $< 1$ ,

# Covarianza di variabili Bernoulli

Siano  $X \in \{0, 1\}$ , indicatrice dell'evento  $A$ ,  $Y \in \{0, 1\}$  indicatrice dell'evento  $B$ .

- ▶ Allora  $XY \in \{0, 1\}$  è indicatrice di “ $A$  e  $B$ ” e

$$\text{Cov}(X, Y) = \mathbb{E}[XY] - \mathbb{E}[X]\mathbb{E}[Y] = P(A \text{ e } B) - P(A)P(B).$$

- ▶  $X$  e  $Y$  sono positivamente correlate se e solo se

$$P(A \text{ e } B) - P(A)P(B) > 0, \quad \text{ossia} \quad \frac{P(A \text{ e } B)}{P(A)P(B)} > 1,$$

- ▶ ossia  $P(A|B) > P(A)$  ( $B$  è una informazione a favore di  $A$ )
- ▶ sono negativamente correlate se e solo se il rapporto è  $< 1$ ,
- ▶ sono non correlate se e solo se il rapporto vale 1 **ossia**  $A$ ,  $B$  sono indipendenti.

# Valor medio e varianza di variabili Binomiali

# Intepretazione del segno della covarianza

- ▶ Come intepretare il segno della covarianza nel caso di variabili generali?

# Interpretazione del segno della covarianza

- ▶ Come interpretare il segno della covarianza nel caso di variabili generali?
- ▶ Vedremo una spiegazione precisa trattando la regressione lineare.

# Interpretazione del segno della covarianza

- ▶ Come interpretare il segno della covarianza nel caso di variabili generali?
- ▶ Vedremo una spiegazione precisa trattando la regressione lineare.
- ▶  $X$  e  $Y$  sono positivamente correlate se, sapendo che  $X > \mathbb{E}[X]$  allora è più probabile che sia anche  $Y > \mathbb{E}[Y]$  (e similmente, sapendo  $X \leq \mathbb{E}[X]$ , è più probabile che sia  $Y \leq \mathbb{E}[Y]$ ).

# Interpretazione del segno della covarianza

- ▶ Come interpretare il segno della covarianza nel caso di variabili generali?
- ▶ Vedremo una spiegazione precisa trattando la regressione lineare.
- ▶  $X$  e  $Y$  sono positivamente correlate se, sapendo che  $X > \mathbb{E}[X]$  allora è più probabile che sia anche  $Y > \mathbb{E}[Y]$  (e similmente, sapendo  $X \leq \mathbb{E}[X]$ , è più probabile che sia  $Y \leq \mathbb{E}[Y]$ ).
- ▶ Graficamente, la densità congiunta tra  $(X, Y)$  è più concentrata nel primo e terzo quadrante cartesiano, avendo posto l'origine nel vettore dei valor medi

# Interpretazione del segno della covarianza

- ▶ Come interpretare il segno della covarianza nel caso di variabili generali?
- ▶ Vedremo una spiegazione precisa trattando la regressione lineare.
- ▶  $X$  e  $Y$  sono positivamente correlate se, sapendo che  $X > \mathbb{E}[X]$  allora è più probabile che sia anche  $Y > \mathbb{E}[Y]$  (e similmente, sapendo  $X \leq \mathbb{E}[X]$ , è più probabile che sia  $Y \leq \mathbb{E}[Y]$ ).
- ▶ Graficamente, la densità congiunta tra  $(X, Y)$  è più concentrata nel primo e terzo quadrante cartesiano, avendo posto l'origine nel vettore dei valor medi
- ▶ La correlazione negativa indica che è più concentrata nel secondo e quarto quadrante.

# Matrice delle covarianze

Dato un vettore aleatorio  $X = (X_1, \dots, X_d) \in \mathbb{R}^d$ , si definisce la matrice delle covarianze di  $X$  la matrice di numeri reali  $\Sigma_X \in \mathbb{R}^{d \times d}$

$$(\Sigma_X)_{i,j} = \text{Cov}(X_i, X_j) = \mathbb{E}[X_i X_j] - \mathbb{E}[X_i] \mathbb{E}[X_j] \quad \text{per } i, j \in \{1, \dots, d\}.$$

► Notazioni alternative:  $\text{Var}(X)$ ,  $K_{XX}$  o  $Q_X$ .

# Proprietà

- ▶ La matrice delle covarianze è simmetrica  $\Sigma_X = \Sigma_X^T$ , dove  $T$  indica l'operazione di trasposizione.

# Proprietà

- ▶ La matrice delle covarianze è simmetrica  $\Sigma_X = \Sigma_X^T$ , dove  $T$  indica l'operazione di trasposizione.
- ▶ (trasformazioni affini) Sia  $X \in \mathbb{R}^d$  e

$$Y = AX + b \quad \text{ossia} \quad Y_i = \sum_{j=1}^d A_{ij} X_j + b_i,$$

dove  $A \in \mathbb{R}^{k \times d}$  è una matrice e  $b \in \mathbb{R}^k$  è un vettore (costanti).  
Vale

$$\Sigma_{AX+b} = A \Sigma_X A^T.$$

# Proprietà

- ▶ La matrice delle covarianze è simmetrica  $\Sigma_X = \Sigma_X^T$ , dove  $T$  indica l'operazione di trasposizione.
- ▶ (trasformazioni affini) Sia  $X \in \mathbb{R}^d$  e

$$Y = AX + b \quad \text{ossia} \quad Y_i = \sum_{j=1}^d A_{ij}X_j + b_i,$$

dove  $A \in \mathbb{R}^{k \times d}$  è una matrice e  $b \in \mathbb{R}^k$  è un vettore (costanti).  
Vale

$$\Sigma_{AX+b} = A\Sigma_X A^T.$$

- ▶ In particolare, se  $k = 1$  e  $A = v^T$ , con  $v \in \mathbb{R}^d$ , si ottiene che

$$\text{Var}(v \cdot X) = \Sigma_{v \cdot X} = v^T \Sigma_X v,$$

ossia  $\Sigma_X$  è (semi-)definita positiva.



## Conseguenze: il coefficiente di correlazione

- ▶ Nel caso  $d = 2$ , scrivendo  $(X, Y)$  per la variabile congiunta di due variabili reali  $X, Y$ , la matrice delle covarianze è esplicitamente

$$\Sigma_{(X,Y)} = \begin{pmatrix} \text{Var}(X) & \text{Cov}(X, Y) \\ \text{Cov}(X, Y) & \text{Var}(Y) \end{pmatrix}.$$

## Conseguenze: il coefficiente di correlazione

- ▶ Nel caso  $d = 2$ , scrivendo  $(X, Y)$  per la variabile congiunta di due variabili reali  $X, Y$ , la matrice delle covarianze è esplicitamente

$$\Sigma_{(X,Y)} = \begin{pmatrix} \text{Var}(X) & \text{Cov}(X, Y) \\ \text{Cov}(X, Y) & \text{Var}(Y) \end{pmatrix}.$$

- ▶ Essendo semidefinita positiva, il suo determinante è positivo (o nullo):

$$\det(\Sigma_{(X,Y)}) = \text{Var}(X) \text{Var}(Y) - (\text{Cov}(X, Y))^2 \geq 0,$$

## Conseguenze: il coefficiente di correlazione

- ▶ Nel caso  $d = 2$ , scrivendo  $(X, Y)$  per la variabile congiunta di due variabili reali  $X, Y$ , la matrice delle covarianze è esplicitamente

$$\Sigma_{(X,Y)} = \begin{pmatrix} \text{Var}(X) & \text{Cov}(X, Y) \\ \text{Cov}(X, Y) & \text{Var}(Y) \end{pmatrix}.$$

- ▶ Essendo semidefinita positiva, il suo determinante è positivo (o nullo):

$$\det(\Sigma_{(X,Y)}) = \text{Var}(X) \text{Var}(Y) - (\text{Cov}(X, Y))^2 \geq 0,$$

- ▶ ossia, dopo alcune operazioni elementari

$$\rho_{XY} := \frac{\text{Cov}(X, Y)}{\sigma_X \sigma_Y} \in [-1, 1].$$

## Conseguenze: il coefficiente di correlazione

- ▶ Nel caso  $d = 2$ , scrivendo  $(X, Y)$  per la variabile congiunta di due variabili reali  $X, Y$ , la matrice delle covarianze è esplicitamente

$$\Sigma_{(X,Y)} = \begin{pmatrix} \text{Var}(X) & \text{Cov}(X, Y) \\ \text{Cov}(X, Y) & \text{Var}(Y) \end{pmatrix}.$$

- ▶ Essendo semidefinita positiva, il suo determinante è positivo (o nullo):

$$\det(\Sigma_{(X,Y)}) = \text{Var}(X) \text{Var}(Y) - (\text{Cov}(X, Y))^2 \geq 0,$$

- ▶ ossia, dopo alcune operazioni elementari

$$\rho_{XY} := \frac{\text{Cov}(X, Y)}{\sigma_X \sigma_Y} \in [-1, 1].$$

- ▶  $\rho_{XY}$  è il *coefficiente di correlazione* (o indice di correlazione di Pearson).

# Standardizzazione nel caso vettoriale

Il teorema spettrale permette di decomporre

$$\Sigma_X = U^T D U,$$

dove  $U \in \mathbb{R}^{d \times d}$ , è ortogonale  $U^T U = Id$  e  $D$  è diagonale (gli autovalori di  $\Sigma_X$ ).

- La trasformazione  $UX$ , corrisponde ad un cambio di coordinate e trasforma la covarianza

$$\Sigma_{UX} = U \Sigma_X U^T = D$$

ossia le componenti di  $UX$  sono a due a due non correlate.

- Se  $D$  è invertibile si può definire una *standardizzazione* di  $X$

$$\hat{X} = \sqrt{D}^{-1} U(X - \mathbb{E}[X])$$

dove  $\sqrt{D}$  è la matrice diagonale con entrate date dalla radice quadrata di quelle di  $D$ .

- Se  $D$  è invertibile si può definire una *standardizzazione* di  $X$

$$\hat{X} = \sqrt{D}^{-1} U(X - \mathbb{E}[X])$$

dove  $\sqrt{D}$  è la matrice diagonale con entrate date dalla radice quadrata di quelle di  $D$ .

- Usando le proprietà del vettore delle medie e della varianza, si ha

$$\mathbb{E}[\hat{X}] = 0 \in \mathbb{R}^d \quad \text{e} \quad \Sigma_{\hat{X}} = Id.$$

- ▶ Se  $D$  è invertibile si può definire una *standardizzazione* di  $X$

$$\hat{X} = \sqrt{D}^{-1} U(X - \mathbb{E}[X])$$

dove  $\sqrt{D}$  è la matrice diagonale con entrate date dalla radice quadrata di quelle di  $D$ .

- ▶ Usando le proprietà del vettore delle medie e della varianza, si ha

$$\mathbb{E}[\hat{X}] = 0 \in \mathbb{R}^d \quad \text{e} \quad \Sigma_{\hat{X}} = Id.$$

- ▶ Attenzione! quando si passa alle osservazioni di un campione, la *standardizzazione* si riferisce all'operazione effettuata sulle marginali (comando `scale()` in R).



