2022-06-01
2022-05-31
2022-05-30
2022-05-27
2022-05-26
2022-05-25
2022-05-24
2022-05-23
2022-05-20
2022-05-19
2022-05-18
2022-05-17
2022-05-16
2022-05-13
2022-05-12
2022-05-11
2022-05-10
2022-05-09
2022-05-06
2022-05-05
2022-05-04
2022-05-03
2022-05-02
2022-04-29
2022-04-28
2022-04-27
2022-04-26
2022-04-22
2022-04-21
2022-04-20
2022-04-19
2022-04-15
2022-04-14
2022-04-13
2022-04-12
2022-04-11
2022-04-08
2022-04-07
2022-04-06
2022-04-05
2022-04-04
2022-04-01
2022-03-31
2022-03-30
2022-03-29
2022-03-28
2022-03-24

```
##
## Attaching package: 'xtable'

## The following object is masked from 'package:emayili':
##
##     display
```

Soluzioni all'esercizio del 2022-05-19 creato per luigi.miazzo

Sono date due variabili casuali discrete: X prende valori in $\{-1, 8.5, 9.5\}$, mentre $Y(\omega) \in \{-6.5, 8, 9.5\}$.

Si conoscono: la funzione di probabilità marginale $p_X(x) = P(X = x)$

-1	8.5	9.5
0.39	0.33	0.28

e le probabilità condizionate $p_{Y|X}(y|x)$ (righe della seguente tabella)

	-6.5	8	9.5
-1	0.36	0.30	0.34
8.5	0.04	0.24	0.72
9.5	0.23	0.10	0.67

Suggerimento: Potete estrarre dal sorgente le tabelle HTML `<table>...</table>` e, usando qualche sito, trasformarle in csv, per poi leggere i dati con `read.csv(file, row.names = 1)`.

Quesiti e soluzioni

Quesito 1

Qual è la probabilità della coppia $(9.5, 9.5)$, secondo la legge bivariata di (X, Y)?

Dalla definizione di probabilità condizionata:

$$p_{X|Y}(x|y) = \mathrm{P}(\{X = x\}|\{Y = y\}) = \frac{p_{X,Y}(x, y)}{p_Y(y)}$$

per ogni $y \in R_Y$ tale che $p_Y(y) > 0$; da cui

$$p_{X,Y}(x, y) = p_{X|Y}(X = x|Y = y) \cdot p_Y(y).$$

e, sostituendo ad x e y i valori dati, concludiamo.

- Risposta corretta: 0.1876
- Risposta inserita: 0.1876
- che corrisponde a 0.1876

Quesito 2

Determinare $\mathrm{Cov}(X, Y)$.

Sappiamo che la covarianza è

$$\begin{aligned}\mathrm{Cov}(X, Y) &= \mathbb{E}[(X - \mathbb{E}(X))(Y - \mathbb{E}(Y))] \\ &= \mathbb{E}(XY) - \mathbb{E}(X)\mathbb{E}(Y)\end{aligned}$$

quindi dobbiamo, per prima cosa, calcolare i valori attesi di X e Y.

$\mathbb{E}(X)$ si ottiene immediatamente, poiché la marginale di X è nota, $\mathbb{E}(X) = 5.075$.

Per calcolare $\mathbb{E}(Y)$ possiamo usare la regola dell'aspettazione totale $\mathbb{E}(Y) = \mathbb{E}(\mathbb{E}(Y|X))$ o ricavare la marginale di Y dalla funzione di probabilità congiunta ricavata nel punto precedente. I valori di $\mathbb{E}(Y|X = x)$ per ogni valore di x sono:

```
print(xtable(EY_X), include.rownames = TRUE, include.colnames = TRUE, type="html")
```

```
## <!-- html table generated in R 4.1.2 by xtable 1.8-4 package -->
## <!-- Thu Jun  9 22:32:16 2022 -->
## <table border=1>
## <tr> <th>  </th> <th> E(Y|X=x) </th>  </tr>
##   <tr> <td align="right"> -1 </td> <td align="right"> 3.29 </td> </tr>
##   <tr> <td align="right"> 8.5 </td> <td align="right"> 8.50 </td> </tr>
##   <tr> <td align="right"> 9.5 </td> <td align="right"> 5.67 </td> </tr>
##    </table>
```

Infine $\mathbb{E}(Y) = \sum_{i=1}^{3} \mathbb{E}(Y|X = x_i)p_X(x_i) = 5.6757$.

Dopodiché possiamo procedere in due modi:

- 1. Centriamo le variabili X e Y attorno alle rispettive medie, chiamiamo $\tilde{X} = X - \mathbb{E}(X)$ e $\tilde{Y} = Y - \mathbb{E}(Y)$ le rispettive v.a. centrate, e calcoliamo il valore atteso del prodotto (usando la funzione di densità discreta congiunta)

$$\mathbb{E}(\tilde{X}\tilde{Y}) = \sum_{i=1}^{3}\sum_{j=1}^{3} \tilde{x}_i \tilde{x}_j p_{X,Y}(x_i, x_j).$$

- 2. Calcoliamo il valore atteso del prodotto delle variabili non centrate (usando la funzione di densità discreta congiunta), a cui poi sottrarremo il prodotto dei valori attesi

$$\mathbb{E}(XY) = \sum_{i=1}^{3}\sum_{j=1}^{3} x_i y_j p_{X,Y}(x_i, y_j).$$

In entrambi i casi abbiamo bisogno della densità discreta congiunta di (X, Y), che riportiamo qui per comodità.

	-6.5	8	9.5
-1	0.14	0.12	0.13
8.5	0.01	0.08	0.24
9.5	0.06	0.03	0.19

- Risposta corretta: 8.8374225
- Risposta inserita: 8.8374225
- che corrisponde a 8.8374225

Quesito 3

Determinare l'indice di correlazione di Pearson, $\rho(X, Y)$.

La correlazione di X e Y è data da

$$\rho(X, Y) = \frac{\mathrm{cov}(X, Y)}{\sqrt{\mathrm{var}(X) \cdot \mathrm{var}(Y)}}$$

Dobbiamo quindi calcolare la varianze delle singole variabili.

La varianza di X si ricava direttamente usando la definizione e la marginale di X, $\mathbb{V}ar(X) = \sum_{i=1}^{3} \tilde{x}_i^2 p_X(x)$ (dove abbiamo usato la notazione precedente $\tilde{x}_i = x_i - \mathbb{E}(X)$).

Per la varianza di Y dobbiamo, prima, ricavarci la marginale e poi procedere come per la X. La marginale di Y, $p_Y(y)$ si ottiene prendendo la somma di ogni colonna nella tabella della distribuzione congiunta. In R, se abbiamo salvato le probabilità congiunta in una matrice chiamata `p_XY`, basta fare `colSums(p_XY)` e otteniamo 0.218, 0.2242, 0.5578.

- Risposta corretta: 0.2808802
- Risposta inserita: 0.2808802
- che corrisponde a 0.2808802