

Esercizio

Una azienda produce barre di materiale plastico. Dopo avere esaminato 5000 barre prodotte si conclude che la lunghezza è distribuita normalmente con media 11.5 cm e deviazione standard 1.15 cm.

- 1) Quale è la probabilità che la lunghezza di una barra assuma valore compreso tra 10.0 e 13.0 cm
- 2) Delle 5000 barre quante hanno lunghezza tra 10 e 13 cm; più di 10 cm; meno di 13 cm.
- 3) Vengono analizzate altre due barre prodotte. Quale è la probabilità che entrambe siano più lunghe di 13.5 cm?

Definiamo la VA X relativa alla lunghezza delle barre: $X:=N(11.5,1.15)$. Attenzione il dato è la deviazione standard non la varianza.

- 1) Si deve valutare la probabilità $P\{10 \leq X \leq 13\}$. Utilizzando Matlab la probabilità di tale evento si determina con il comando:

$$\text{normcdf}(13,11.5,1.15) - \text{normcdf}(10,11.5,1.15).$$

Si ottiene il valore 0.8079, quindi 80.79% del campione ha lunghezza compresa tra 10 e 13cm.

- 2) Punto 2

- a. Barre del campione con dimensione compresa tra 10 e 13 cm: $5000 \times 0.8079 = 4039$. Abbiamo ovviamente arrotondato all'intero più vicino.
- b. Barre di dimensione maggiore di 10cm: $5000 \times P\{X > 10\} = 5000 \times (1 - \text{normcdf}(10,11.5,1.15)) = 5000 \times 0.904 = 4519$
- c. Barre di dimensione minore di 13cm: $5000 \times P\{X < 13\} = 5000 \times \text{normcdf}(13,11.5,1.15) = 5000 \times 0.904 = 4519$

- 3) Evento A: la barra estratta ha dimensione maggiore di 13.5: $P(A) = P\{X > 13.5\} = 1 - \text{normcdf}(13.5,11.5,1.15) = 0.041$. Evento B: l'ulteriore barra estratta ha dimensione maggiore di 13.5: $P(B) = 0.041$. Vogliamo conoscere la probabilità che entrambi gli eventi si verifichino: $P(A \cap B)$. Gli eventi sono indipendenti quindi $P(A \cap B) = P(A)P(B) = 0.041^2$.

Esercizio

Siano $P(A) = 0.9$ e $P(B) = 0.8$. Dimostrate che $P(A \cap B) \geq 0.7$

Non abbiamo alcuna informazione sul legame tra A e B. Possiamo procedere con la proprietà:

$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$

$$P(A \cap B) = P(A \cup B) - P(A) - P(B)$$

Ma $P(A \cup B)$ può al massimo essere pari ad 1 quindi: $P(A \cap B) \leq 1 - P(A) - P(B) = 0.7$ CVD.

Esercizio

Si consideri una variabile aleatoria Gaussiana $Y = N(\mu_Y, \sigma_Y^2) = N(0.6, 0.14)$

Attraverso l'uso delle tabelle si chiede di calcolare:

- 1) la probabilità: $P\{0.35 \leq y \leq +\infty\}$
- 2) la probabilità: $P\{-0.98 \leq y \leq \mu_Y\}$
- 3) la costante c_1 reale tale che: $P\{-c_1 \leq y \leq c_1\} = 0.75$

Si consideri la seguente trasformazione: $Z = \frac{1}{Y^2 - 1}$

- 4) Si determini la PDF della variabile aleatoria Z
- 5) Si stimino media e varianza della Z

1) La Gaussiana va standardizzata per poter usare le tabelle:

$$P\{0.35 \leq Y \leq +\infty\} = P\left\{\frac{0.35-0.6}{\sqrt{0.14}} \leq Z \leq +\infty\right\} = P\{0.668 \leq Z \leq +\infty\}$$

Quindi: $P\{0.668 \leq Z \leq +\infty\} = 1 - P\{Z \leq 0.668\} = 1 - \Phi(0.668) = 0.747$

$$2) P\{-0.98 \leq Y \leq \mu_Y\} = P\left\{\frac{-0.98-0.6}{\sqrt{0.14}} \leq Z \leq 0\right\}$$

$$P\{-4.22 \leq Z \leq 0\} = F(0) - F(-4.22) = 0.5$$

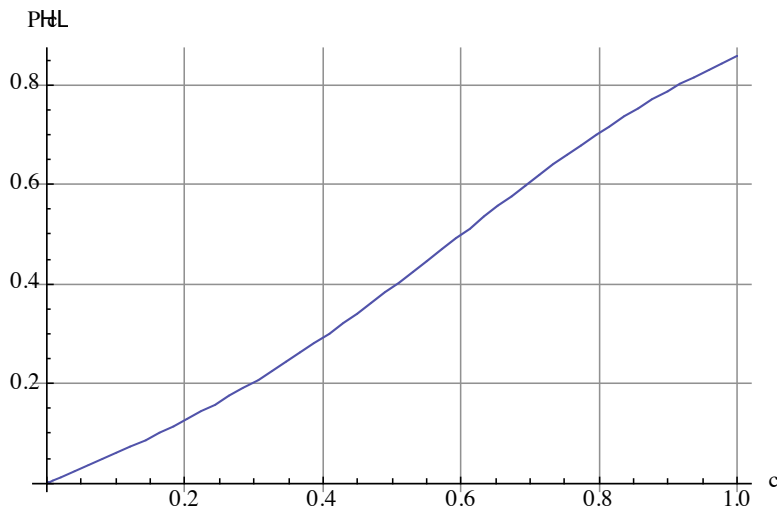
NB il valore -4.22 non compare nelle tabelle che si fermano a -3. Al di sotto di tale valore la F è praticamente nulla.

3) Anche in questo caso standardizziamo la gaussiana:

$$P\left\{\frac{-c-0.6}{\sqrt{0.14}} \leq Z \leq \frac{c-0.6}{\sqrt{0.14}}\right\} = 0.75$$

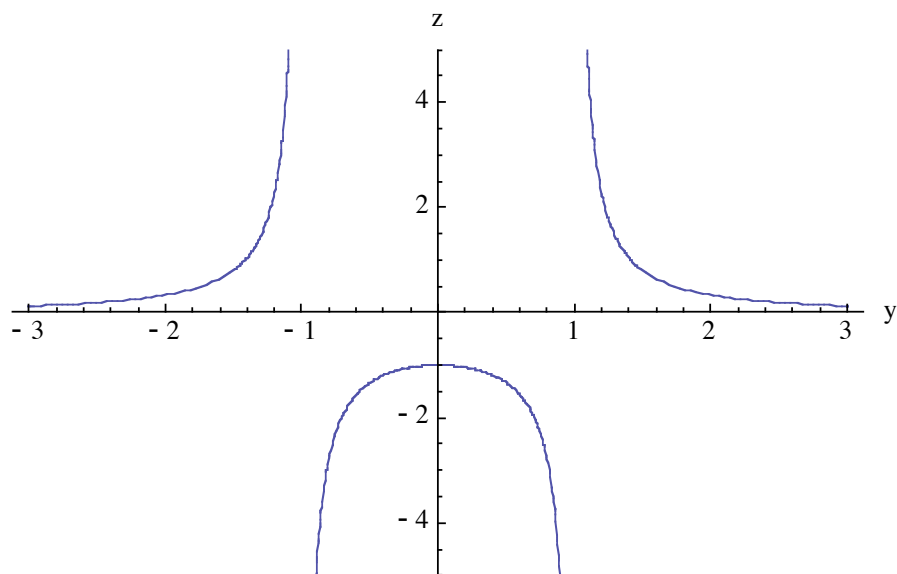
Si noti che i due estremi non sono simmetrici rispetto allo 0.

Si può costruire una $P(c)$ con Matlab e diagrammarla al variare di c :



La probabilità è circa 0.75 per c circa 0.85. In modo più accurato si può procedere alla determinazione del valore di c risolvendo per tentativi (numericamente) l'equazione $P(c)=0.85$. Il risultato accurato è $c=0.852$.

4) La trasformazione è non biunivoca e non lineare:



Le radici della trasformazione sono: $y = \pm \sqrt{\frac{1}{z} + 1}$. Si può facilmente verificare, anche dal grafico, che non ci sono soluzioni per $-1 < z \leq 0$, per cui in tale intervallo $f_z=0$. Al di fuori di tale intervallo ci sono due radici per cui:

$$f_z(z) = \sum_{i=1}^2 \frac{f_Y(y_i)}{\left| \frac{dz}{dy} \right|_{y_i}}$$

La derivata della trasformazione è:

$$\frac{dz}{dy} = -\frac{2y}{(y^2 - 1)^2}$$

di conseguenza si ottiene per $z > 0$ e $z \leq -1$:

$$f_z(z) = \frac{1}{\left| 2z^2 \sqrt{\frac{1}{z} + 1} \right|} \frac{1}{\sqrt{2\pi} 0.14} \left\{ \exp \left[-\frac{1}{2} \frac{\left(\sqrt{\frac{1}{z} + z} - 0.6 \right)^2}{0.14} \right] + \exp \left[-\frac{1}{2} \frac{\left(-\sqrt{\frac{1}{z} + z} - 0.6 \right)^2}{0.14} \right] \right\}$$

5) Per stimare la media e varianza della z possiamo usare le formule troncate:

$$\mu_z \approx g(\mu_y) = -1.56$$

$$\sigma_z^2 \approx \left(\frac{dz}{dy} \Big|_{\mu_y} \right)^2 \sigma_y^2 = \left(-\frac{2\mu_y}{(\mu_y^2 - 1)^2} \right)^2 \sigma_y^2 = 1.2$$

Esercizio

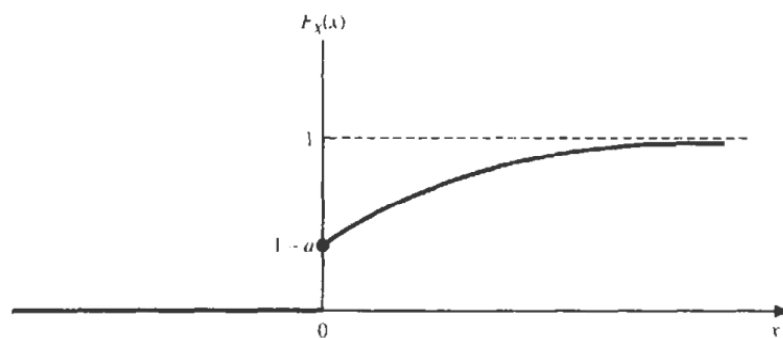
- 1) Determinate i valori delle costanti a e b tali che la funzione $F(x)$ possa essere una CDF

$$F(x) = \begin{cases} 1 - a \exp\left(-\frac{x}{b}\right) & x \geq 0 \\ 0 & x < 0 \end{cases}$$

- 2) Per una scelta plausibile di a e b diagrammate qualitativamente la funzione e calcolate: i) media, ii) moda e iii) mediana

1) Perché la F sia una CDF deve soddisfare la proprietà $0 \leq F(x) \leq 1$: Questa è verificata per $0 \leq a \leq 1$ e $b > 0$. Per queste condizioni sui parametri a e b la F soddisfa la proprietà $F(\infty) = 1$. Inoltre la proprietà $F(-\infty) = 0$ è automaticamente rispettata.

Per $0 \leq a \leq 1$ e $b > 0$ la funzione è



- 2) Scegliamo $a=1$ e $b=1$

Media = $\int_0^{\infty} xf(x)dx = 1$; Mediana = 0.693147 (si ottiene risolvendo $F(x)=0.5$); Moda=0 (si ottiene determinando il massimo di $f(x)$).

Esercizio

Si consideri la VA gaussiana X. Si sa che $P(3.5 < X < 4.5) = 0.19$ e $P(2 < X < 3) = 0.45$. Si determinino media e varianza della VA.

Per la stessa VA X si conoscono le probabilità di due diversi eventi, queste due informazioni permettono la definizione di due equazioni nelle incognite media e varianza della VA. Standardizzando la VA, le due equazioni sono :

$$F\left(\frac{4.5 - \mu_X}{\sigma_X}\right) - F\left(\frac{3.5 - \mu_X}{\sigma_X}\right) = 0.19$$

$$F\left(\frac{3 - \mu_X}{\sigma_X}\right) - F\left(\frac{2 - \mu_X}{\sigma_X}\right) = 0.45$$

Per risolvere le due equazioni in due incognite si può costruire un programma in Matlab bisogna prima creare la function con le equazioni:

```
function F = sistema(x)
F = [ normcdf((4.5-x(1))/x(2)) - normcdf((3.5-x(1))/x(2)) - 0.19
      normcdf((3.0-x(1))/x(2)) - normcdf((2.0-x(1))/x(2)) - 0.45 ];
```

dove x(1) e x(2) sono media e deviazione standard.

Poi si esegue il comando:

```
fsolve(@sistema, [1,1])
```

che risolve il sistema non lineare. Le soluzioni sono:

x(1)=2.9059 e x(2)=0.7144.

Esercizio

Il peso di confezioni di cemento è stato misurato su un campione casuale di dimensione 14. I dati sono riportati in tabella.

[7.95, 7.02, 8.15, 5.03, 8.08, 8.07, 8.24, 8.00, 8.07, 8.04, 8.08, 8.14, 7.86, 8.00]

- Costruite un istogramma per la frequenza relativa
- Calcolate media, mediana e moda del campione
- Calcolate la deviazione standard del campione.

Sostituire l'ultimo dato con il valore 15. Ricalcolate media e mediana e commentate i risultati

a) Procedendo con Matlab:

Si immettono nella variabile peso i dati. Si calcola la dimensione del campione e la deviazione standard

```
N=length(peso)  
sigma=std(peso)
```

Si calcola il numero ottimale di intervalli, per esempio

```
N=(max(peso)-min(peso))/(sigma*3.49*(N^(-1/3)))
```

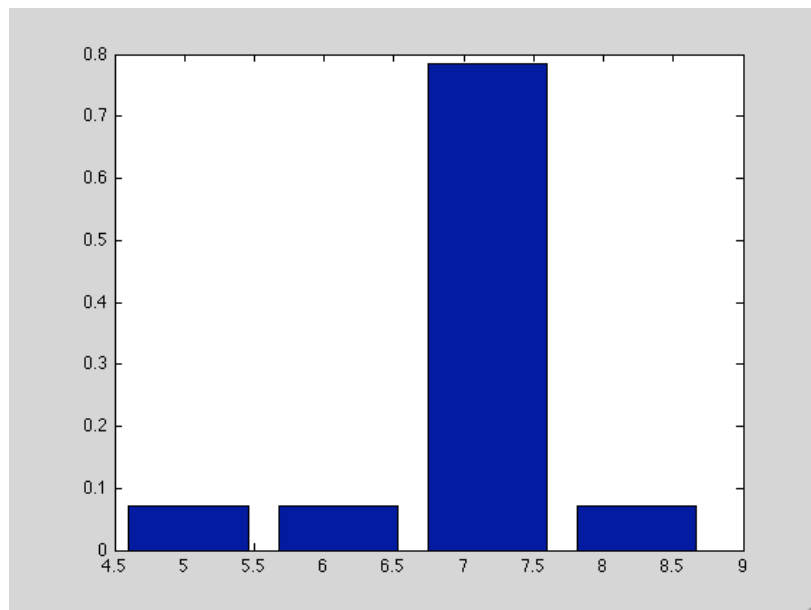
```
passo=(max(peso)-min(peso))/round(n)
```

Si determina il vettore dei sottointervalli

```
x=min(peso):passo:max(peso)
```

Si calcola la frequenza relativa e si costruisce l'istogramma

```
bar(x,histc(peso,x)/N)
```



b) media=mean(peso)=7.7664; mediana=median(peso)=8.0550; moda: il punto medio del 3° sottointervallo 7.17.

c) la deviazione standard è già stata calcolata al punto a)

d) sostituendo l'ultimo dato del campione con il valore 15 e ricalcolando la media e la mediana ci si rende conto che la media cambia significativamente (diventa 8.2664) mentre la mediana resta quasi invariata (8.0700). Ciò conferma la maggiore robustezza della mediana come indicatore della tendenza centrale di un campione.

Esercizio

Si consideri una variabile aleatoria X uniforme tra 1 e 3.

- 1) Si determini pdf della VA $Y=X^2$
- 2) Si diagrammi qualitativamente la pdf della Y
- 3) Si calcolino media e varianza della Y

1) La variabile aleatoria X ha pdf:

$$f_X = \begin{cases} \frac{1}{3-1} & \text{se } 1 \leq x \leq 3 \\ 0 & \text{altrove} \end{cases}$$

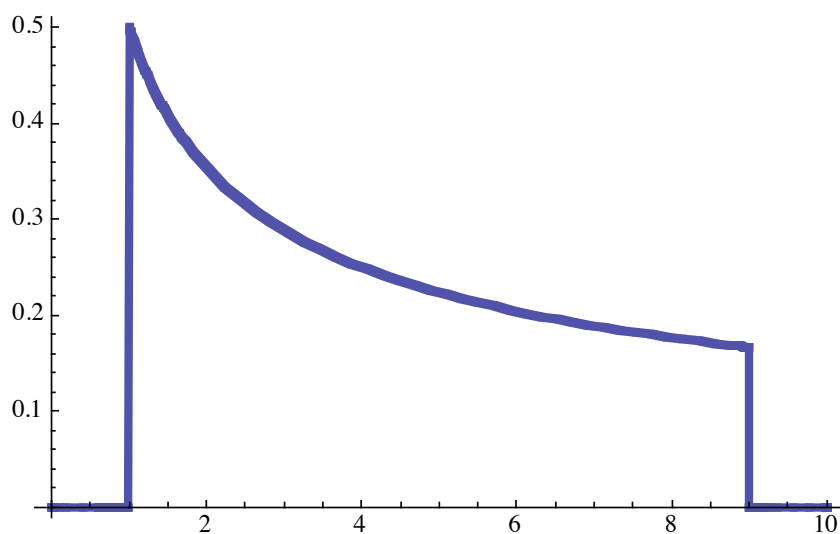
L'evento $A = \{Y \leq y\}$ è equivalente all'evento $B = \{-\sqrt{y} \leq X \leq \sqrt{y}\}$. Quindi se $y \leq 0$ $f_Y = 0$.

$$\text{In definitiva: } f_Y = \begin{cases} \frac{1}{2\sqrt{y}} [f_X(\sqrt{y}) + f_X(-\sqrt{y})] & \text{per } y > 0 \\ 0 & \text{per } y \leq 0 \end{cases}$$

ovvero

$$f_Y = \begin{cases} 0 & \text{per } y < 1 \\ \frac{1}{2\sqrt{y}} \left[\frac{1}{2} + \frac{1}{2} \right] = \frac{1}{2\sqrt{y}} & \text{per } 1 \leq y \leq 9 \\ 0 & \text{per } y > 9 \end{cases}$$

2) Il diagramma della funzione è



4) La media e la varianza della Y sono:

$$\mu_Y = E(Y) = \int_{-\infty}^{+\infty} y f_Y dy = \int_1^9 y \frac{1}{2\sqrt{y}} dy = \left[\frac{1}{3} y^{\frac{3}{2}} \right]_1^9 = 8.666$$

$$\sigma_Y^2 = E[(Y - \mu_Y)^2] = \int_{-\infty}^{+\infty} (y - \mu_Y)^2 f_Y dy = \int_1^9 (y - \mu_Y)^2 \frac{1}{2\sqrt{y}} dy = \left[\frac{1}{15} y^{\frac{1}{2}} (15\mu_Y^2 - 10\mu_Y y + 3y^2) \right]_1^9 = 48.4$$

Esercizio

Sia X una VA continua con pdf (k è una costante)

$$f_X(x) = \begin{cases} kx & 0 < x < 1 \\ 0 & \text{altrove} \end{cases}$$

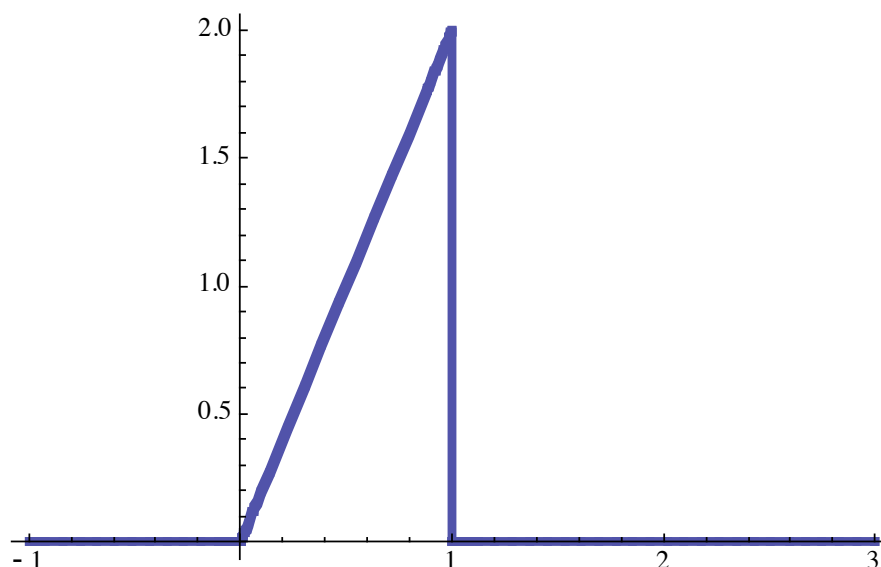
- 1) Determinate il valore di k
- 2) Diagrammate qualitativamente la pdf
- 3) Determinate $P(1.0 < x \leq 5.0)$

1) Perché la funzione sia una PDF è necessario che sia positiva quindi $k > 0$ ed il suo integrale tra meno infinito e più infinito unitario. Questa seconda condizione implica che:

$$\int_{-\infty}^{+\infty} f_X dx = \int_0^1 kx dx = k \left[\frac{x^2}{2} \right]_0^1 = \frac{k}{2},$$

di conseguenza $k=2$.

2) Il diagramma della PDF è



2) La $P(1 < x < 5)$ è pari a zero perché in questo intervallo la PDF è nulla

Esercizio

Si consideri una VA X di tipo chi-quadro a 60 gradi di libertà. Si determini la probabilità $P(50 < X < 70)$ applicando il teorema del limite centrale.

La definizione di una VA di tipo chi-quadro è: $Y = X_1^2 + X_2^2 + \dots + X_n^2$. Le X che compaiono nella somma sono tutte VA dello stesso tipo ovvero Gaussiane standard. Ciascun elemento della somma può essere considerato come una chi-quadro ad 1 grado di libertà, quindi con media 1 (NB media=numero dei gradi di libertà) e varianza 2 (varianza=2*numero di gradi di libertà). Abbiamo quindi la somma di 60 variabili aleatorie tutte dello stesso tipo con egual media ed eguale varianza.

Possiamo quindi applicare il TLC. Per calcolare la probabilità ci riferiamo quindi alla Gaussiana Standard. Si deve determinare

$$P\left\{\frac{50 - n \times \mu}{\sqrt{n} \times \sqrt{\sigma^2}} \leq Z \leq \frac{70 - n \times \mu}{\sqrt{n} \times \sqrt{\sigma^2}}\right\} = P\left\{\frac{50 - 60 \times 1}{\sqrt{60} \times \sqrt{2}} \leq Z \leq \frac{70 - 60 \times 1}{\sqrt{60} \times \sqrt{2}}\right\} = P\left\{-\sqrt{\frac{5}{6}} \leq Z \leq \sqrt{\frac{5}{6}}\right\} = 0.638$$

Esercizio

Si considerino due VA normali $X_1 = N(\mu_Y, \sigma_Y^2) = N(0, 16)$ ed $X_2 = N(0, 20)$ indipendenti tra loro.

- 1) Si determini la pdf della variabile $Y = X_1 + 3X_2$
- 2) Si calcoli la probabilità $P(Y < 5.6)$

1) La VA Y deriva dalla combinazione lineare di due VA Gaussiane, quindi essa è una Gaussiana con media pari alla stessa combinazione lineare delle medie, varianza data dalla somma delle varianze moltiplicate rispettivamente per il coefficiente al quadrato della trasformazione. La covarianza è nulla per l'indipendenza delle X . Quindi:

$$\begin{aligned}\mu_Y &= \mu_{X_1} + 3\mu_{X_2} = 0 \\ \sigma_Y^2 &= \sigma_{X_1}^2 + 9\sigma_{X_2}^2 = 196\end{aligned}$$

Quindi la PDF della Y è:

$$f_Y(y) = \frac{1}{2\pi\sqrt{196}} \exp\left[-\frac{1}{2} \frac{y^2}{196}\right]$$

2) Usando Matlab: $P\{Y < 5.6\} = \text{normcdf}(5.6, 0, \sqrt{196}) = 0.655$

Esercizio

Si consideri la VA bivariata (X, Y) con pdf congiunta data da:

$$f_{XY} = \begin{cases} k(x+y) & 0 < x < 2, 0 < y < 2 \\ 0 & \text{altrove} \end{cases}$$

- 1) Si determini il valore della costante k che rende la pdf ammissibile
- 2) Si determinino le due marginali
- 3) Le VA X ed Y sono indipendenti?

1) Perché la funzione sia una PDF essa deve essere positiva quindi $k > 0$, ed il suo integrale deve essere unitario. Quindi:

$$\int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} f_{XY} dx dy = \int_0^2 \int_0^2 k(x+y) dx dy = \int_0^2 k \left[xy + \frac{y^2}{2} \right]_0^2 dy = \int_0^2 k(2x+2) dx = k \left[2x + x^2 \right]_0^2 = 8k$$

da cui si desume che $k=1/8$.

2) Per definizione le due marginali sono

$$F_X(x) = P\{X \leq x, -\infty \leq Y \leq \infty\}, F_Y(y) = P\{-\infty \leq X \leq \infty, Y \leq y\}$$

Per la congiunta in questione le due marginali diventano:

$$F_X(x) = P\{X \leq x, -\infty \leq Y \leq \infty\} = \int_0^x \int_0^2 \frac{1}{8}(u+y) dy du = \frac{1}{8}(2x + x^2)$$

$$F_Y(y) = P\{-\infty \leq X \leq \infty, Y \leq y\} = \int_0^2 \int_0^y \frac{1}{8}(x+u) du dx = \frac{1}{8}(2y + y^2)$$

3) Le due VA X ed Y sono indipendenti se la congiunta è il prodotto delle marginali. Nel caso allo studio ciò non si verifica quindi le due VA non sono indipendenti.