STATISTICA A – K (63 ore)

Marco Riani

mriani@unipr.it

http://www.riani.it

Generalizzazione del paradosso di Monty Hall a 10 porte

- Il concorrente ne sceglie 1 (ad es la porta 1)
- Il presentatore ne scarta 8 (ad es le porte 2-9)
- Pr che il premio sia nella porta 1?
- Pr che il premio sia nella porta 10?

- Generalizzare il paradosso di Monty Hall a k porte

Il paradosso di Monty Hall

- A_i = l'evento che l'auto (il premio di valore) si trovi dietro la porta *i* (*i*=1, 2, …, 3).
- Per simmetria, supponiamo che il giocatore scelga la porta 1 e che il conduttore abbia mostrato che dietro le porte 2-9 2 c'è il caprone.
- T= il conduttore ha aperto le porte 2-9.

- Obiettivo: calcolare $P(A_1| T)$ e $P(A_{10}|T)$

Il concorrente ha scelto A_1 Calcolo di $P(A_1|T)$

- A_i = l'evento che l'auto (il premio di valore) si trovi dietro la porta *i* (*i*=1, 2, …, 10).
- T= il conduttore apre le porte 2-9.

$$P(A_1|T)=\frac{P(T|A_1)P(A_1)}{\sum_{i=1}^{10}P(T|A_i)P(A_i)}=\frac{\frac{1}{C_{9,8}}\times\frac{1}{10}}{\frac{1}{C_{9,8}}\times\frac{1}{10}+0\times\frac{1}{10}+\cdots+0\times\frac{1}{10}+1\times\frac{1}{10}}=1/10$$

Il concorrente ha scelto A_1 Calcolo di $P(A_{10}|T)$

- A_i = l’evento che l’auto (il premio di valore) si trovi dietro la porta *i* (*i*=1, 2, …, 10).
- T= il conduttore apre la porte 2-9.

$$P(A_{10}|T)=\frac{P(T|A_{10})P(A_{10})}{\sum_{i=1}^{10}P(T|A_i)P(A_i)}=\frac{1\times\frac{1}{10}}{\frac{1}{C_{9,8}}\times\frac{1}{10}+0\times\frac{1}{10}+\cdots+0\times\frac{1}{10}+1\times\frac{1}{10}}=9/10$$

Esercizio

- Dimostrare che
- $f(x)=2(x-10)/50$ se $10<x<15$
- $f(x)=2(20-x)/50$ se $15<x<20$

  è una densità
- Rappresentare graficamente la funzione di densità e di ripartizione

Verificare che è una densità

f(x)=2(x-10)/50 se 10<x<15

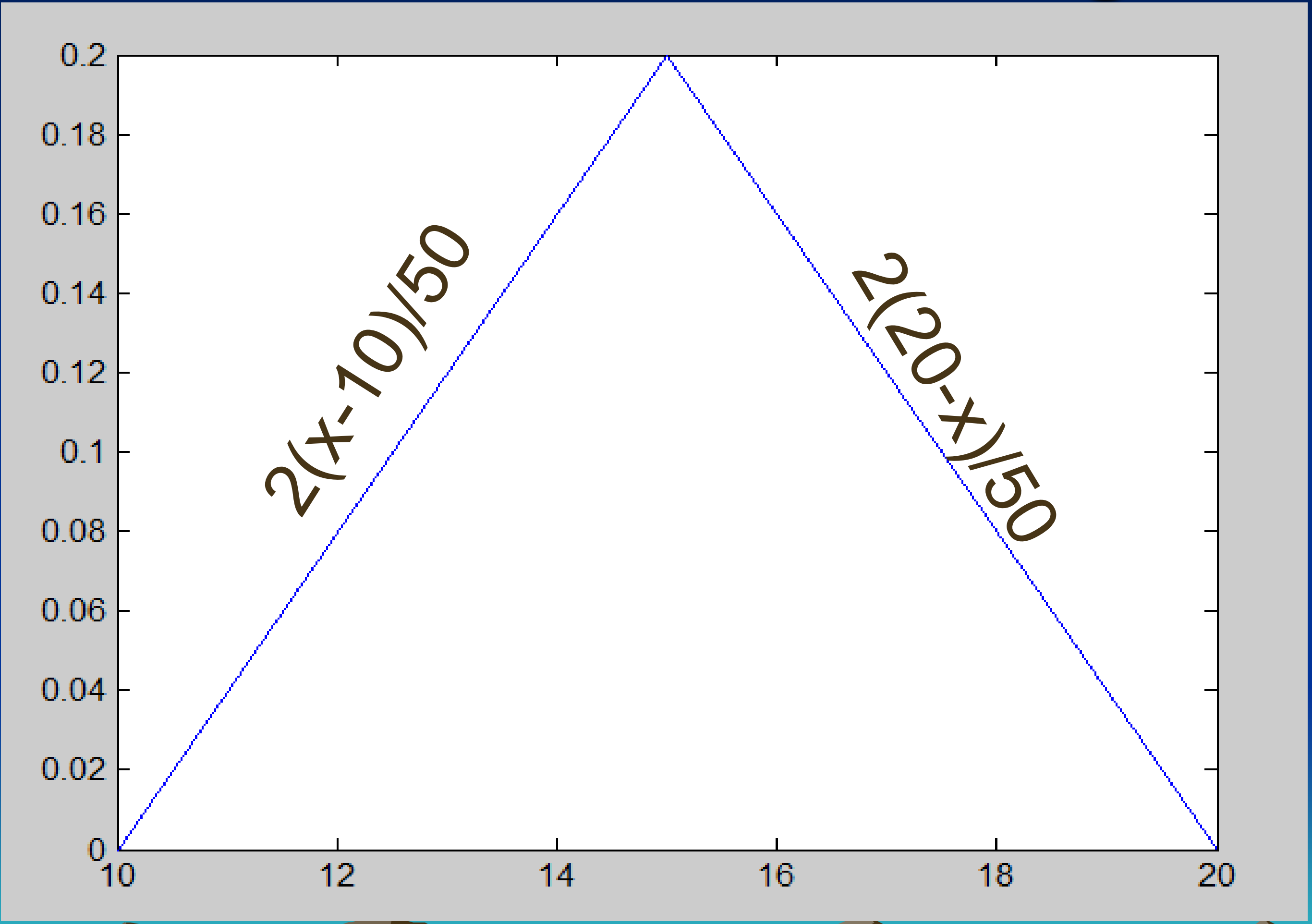
f(x)=2(20-x)/50 se 15<x<20

$$f(x) \geq 0$$

$$\int_{-\infty}^{+\infty} f(x)dx = 1$$

$$\frac{1}{25}\left[\int_{10}^{15}(x-10)dx + \int_{15}^{20}(20-x)dx\right] = 1$$

Rappresentazione grafica f(x)= densità triangolare



f(x)=2(x-10)/50 se 10<x<15

f(x)=2(20-x)/50 se 15<x<20

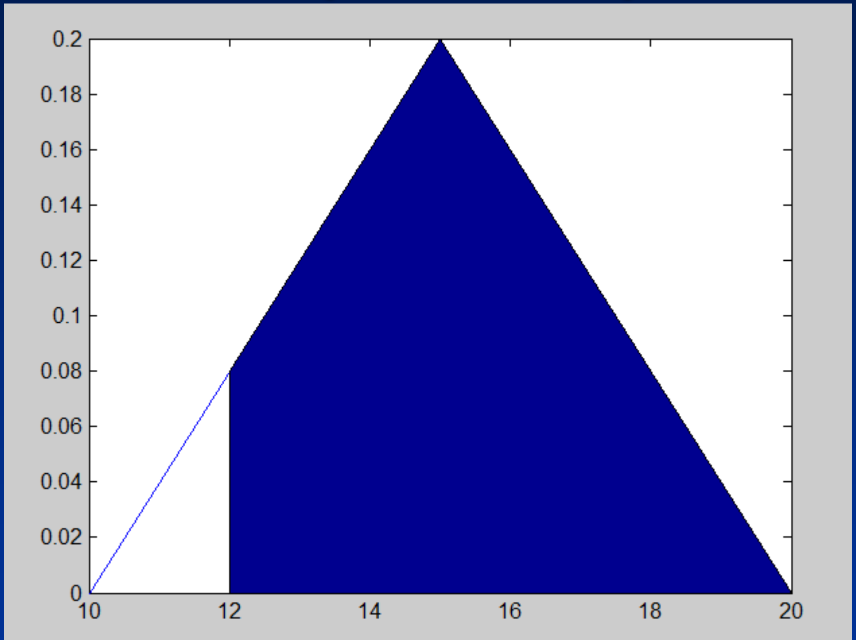
- Area triangolo =10×0,2/2=1

Calcolare

- $Pr(X>12)$
- $Pr(X<10)$
- $Pr(X<11)$
- $Pr(14 < X < 18)$
- $E(X)$?
- $VAR(X)$?
- Calcolare il quantile $x_{0,95}$ ossia la coordinata x che lascia alla sua destra una probabilità pari a 0,05 e a sinistra una probabilità pari a 0,95

Pr(X>12)

f(x)=2(x-10)/50
se 10<x<15



f(x)=2(20-x)/50
se 15<x<20

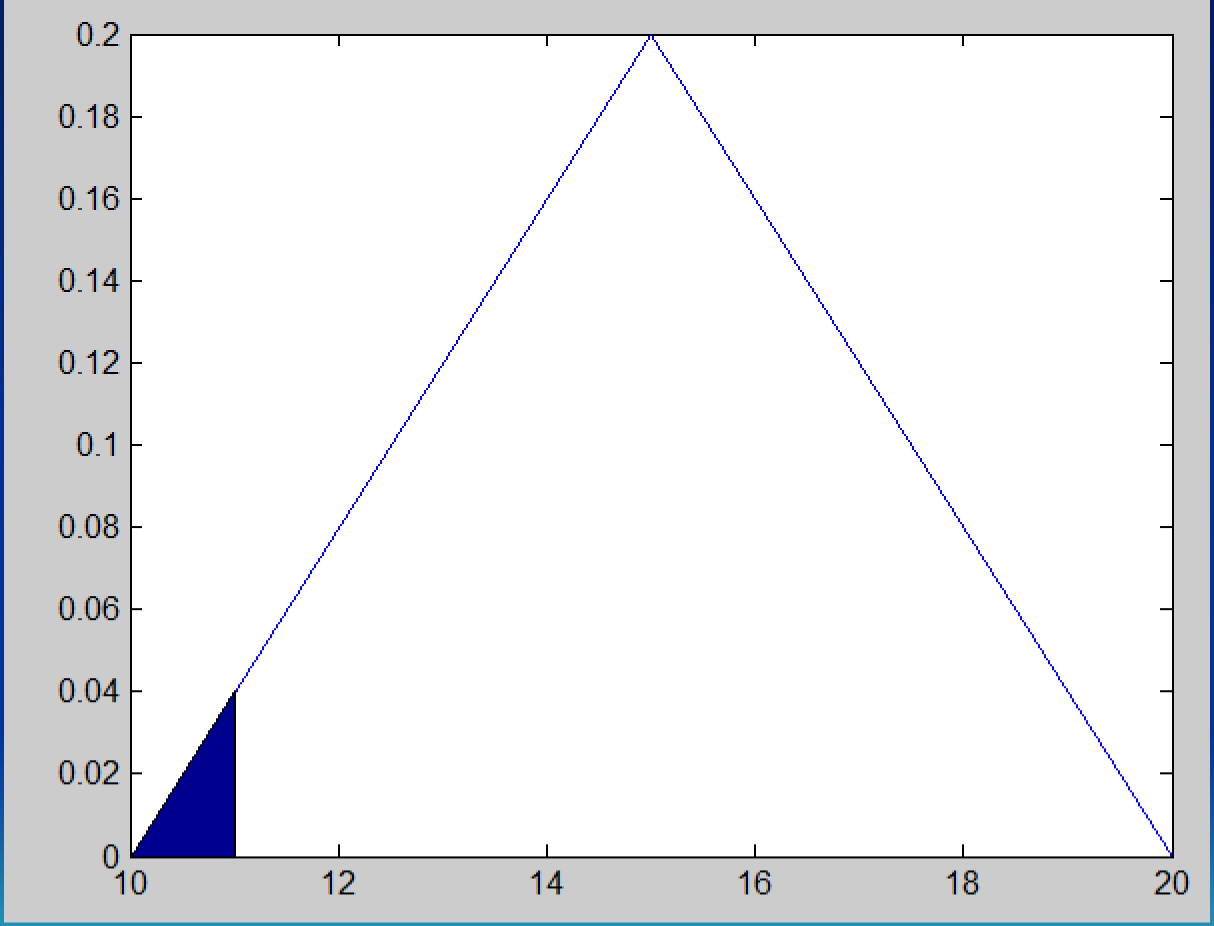
$$Pr(X > 12) = \frac{1}{25}\left[\int_{12}^{15}(x-10)dx + \int_{15}^{20}(20-x)dx\right] =$$

$$Pr(X > 12) = 1 - Pr(X < 12) = \frac{1}{25}\left[\int_{10}^{12}(x-10)dx\right] =$$

1- (Area triangolo con base 2 e altezza 0,08)=
1-2 × 0,08/2=0,92

Pr(X<11)

$f(x)=2(x-10)/50$
se $10<x<15$

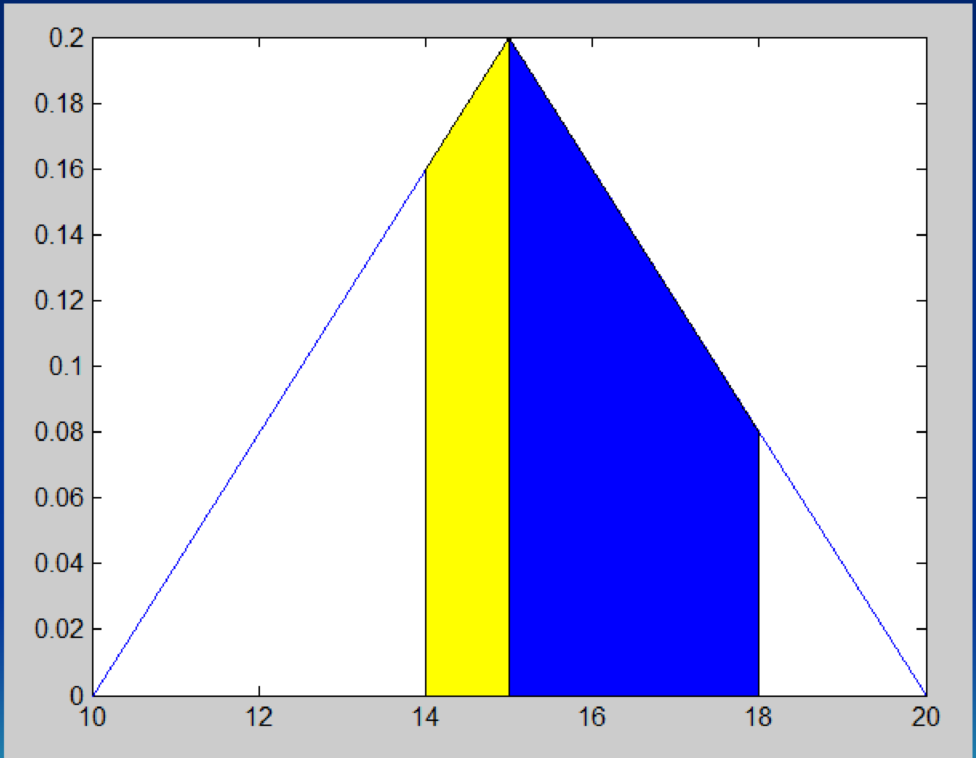


Area triangolo con base 1 e altezza 2/50=0.04
Area = 1 × 0,04/2=0,02

f(x)=(x-10)/25
se 10<x<15

f(x)=(20-x)/25
se 15<x<20

Pr(14 < X < 18)



Pr(14 < X < 18)=0,6 (utilizzando le aree dei due trapezi oppure il calcolo integrale)

E(*X*)

- Occorre calcolare il seguente integrale:

$$E(X) = \int_{-\infty}^{+\infty} x f(x) dx$$

f(x)=(x-10)/25
se 10<x<15

f(x)=(20-x)/25
se 15<x<20

$$E(X) = \frac{1}{25}\left[\int_{10}^{15} x(x-10)dx + \int_{15}^{20} x(20-x)dx\right] =$$

E(*X*)=15

VAR(*X*)

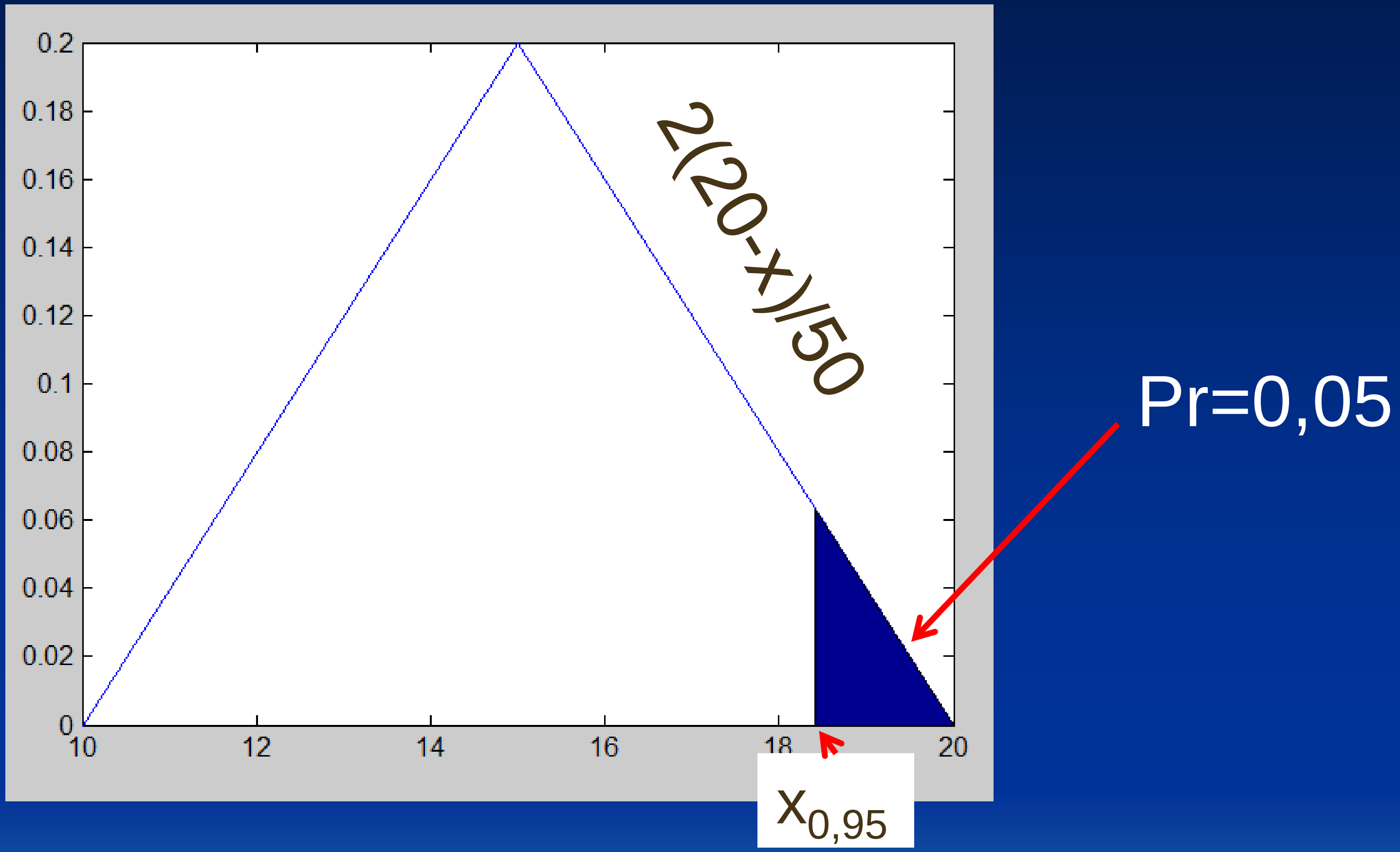
- Occorre calcolare il seguente integrale:

$$VAR(X) = \int_{-\infty}^{+\infty} [x - E(X)]^2 f(x)dx$$

$$VAR(X) = \frac{1}{25}\left[\int_{10}^{15} (x - E(X))^2 (x-10)dx + \int_{15}^{20} (x - E(X))^2 (20-x)dx\right] =$$

VAR(*X*)=4,17

Calcolo del quantile $x_{0,95}$



- $[(20-x_{0,95})\ 2(20-x_{0,95})/50]\ /2=0,05$
- $x_{0,95}=18,42$

Metodo alternativo basato sul calcolo integrale

$$\frac{1}{25}\int_{15}^{x_{0,95}}(20-x)dx = 0,45$$

- Si ottiene:

$$-x_{0,95}^2 + 40x_{0,95} - 397,5 = 0$$

- Le soluzioni dell'equazione di secondo grado sono 21,58 e 18,42. Naturalmente escludiamo 21,58 in quanto esterna all'intervallo di definizione della densità

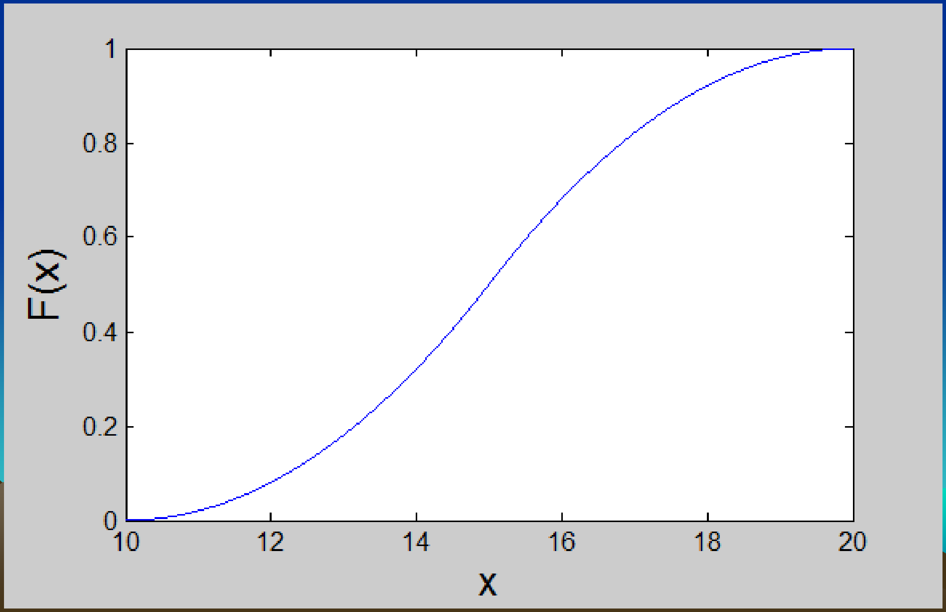
Funzione di ripartizione

$$F(x) = \int_{-\infty}^{x} f(w)dw$$

$$F(x) = \begin{cases} 0 & \text{per} \quad x \leq 10 \\ \frac{1}{25}\int_{10}^{x}(t-10)dt & \text{per} \quad 10 \leq x \leq 15 \\ \frac{1}{2} + \frac{1}{25}\int_{15}^{x}(20-t)dt & \text{per} \quad 15 \leq x \leq 20 \\ 1 & \text{per} \quad x \geq 20 \end{cases}$$

Funzione di ripartizione

$$F(x) = \begin{cases} 0 & \text{per} \quad x \leq 10 \\ \frac{1}{50}(x^2 - 20x + 100) & \text{per} \quad 10 \leq x \leq 15 \\ \frac{1}{2} + \frac{1}{50}(-x^2 + 40x - 375) & \text{per} \quad 15 \leq x \leq 20 \\ 1 & \text{per} \quad x \geq 20 \end{cases}$$



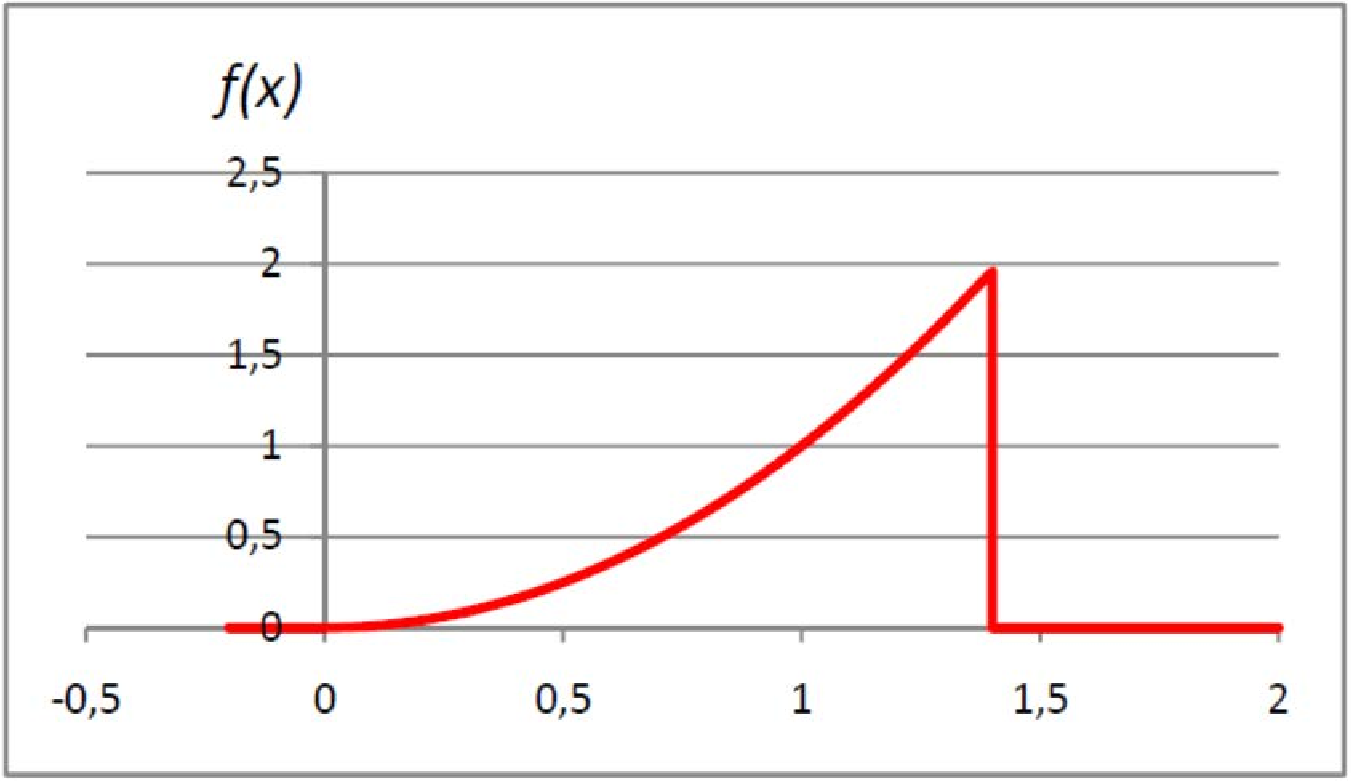
Dimostrare che la seguente funzione è una funzione di densità

$$f(x) = \begin{cases} 0 & x < 0 \\ x^2 & 0 \le x \le \sqrt[3]{3} \\ 0 & x > \sqrt[3]{3} \end{cases}$$

- Rappresentarla graficamente. Calcolare E(X), VAR(X) e Me(X)

Rappresentazione grafica

$$\int_0^{\sqrt[3]{3}} x^2 \mathrm{d}x = \left[\frac{x^3}{3}\right]_0^{\sqrt[3]{3}} = \frac{\left(\sqrt[3]{3}\right)^3}{3} - 0 = 1$$



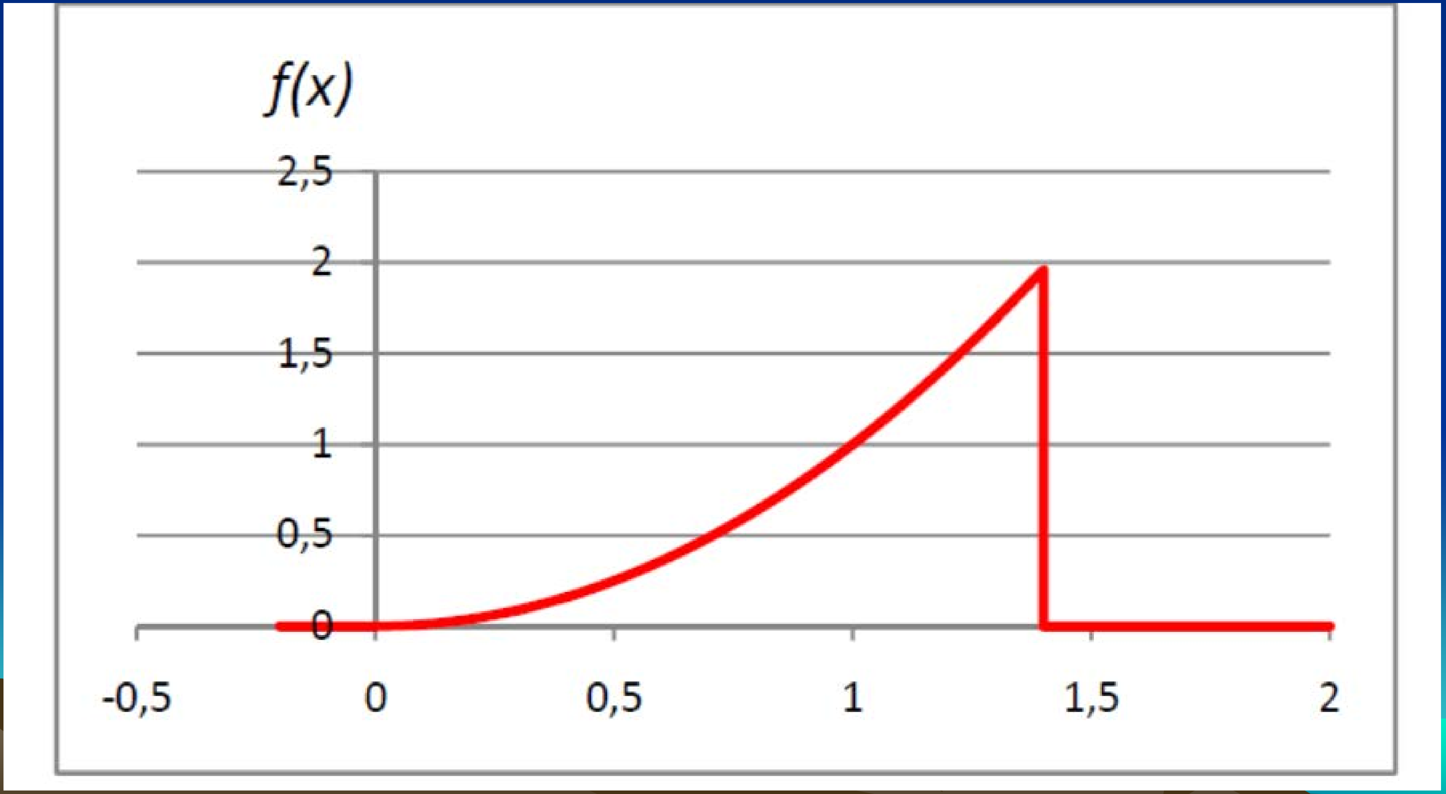
E(X)

$$f(x) = \begin{cases} 0 & x < 0 \\ x^2 & 0 \le x \le \sqrt[3]{3} \\ 0 & x > \sqrt[3]{3} \end{cases}$$

$$\mathrm{E}(X) = \int_{-\infty}^{+\infty} x \cdot f(x)\mathrm{d}x = \int_0^{\sqrt[3]{3}} x \cdot x^2 \mathrm{d}x = \int_0^{\sqrt[3]{3}} x^3 \mathrm{d}x = \left[\frac{x^4}{4}\right]_0^{\sqrt[3]{3}} = \frac{\left(\sqrt[3]{3}\right)^4}{4} - 0 = \frac{3^{\frac{4}{3}}}{4} \approx 1{,}0817$$

Mediana(X)

$$\int_{-\infty}^{\text{mediana}} f(x)\mathrm{d}x = 50\% \quad \text{e} \quad \int_{\text{mediana}}^{+\infty} f(x)\mathrm{d}x = 50\%$$



Mediana(X)

$$f(x) = \begin{cases} 0 & x < 0 \\ x^2 & 0 \le x \le \sqrt[3]{3} \\ 0 & x > \sqrt[3]{3} \end{cases}$$

$$\int_{-\infty}^{\text{mediana}} f(x)\mathrm{d}x = \int_{0}^{\text{mediana}} x^2\mathrm{d}x = \left[\frac{x^3}{3}\right]_0^{\text{mediana}} = \frac{(\text{mediana})^3}{3} - 0 = \frac{(\text{mediana})^3}{3} = 0{,}50$$

$$(\text{mediana})^3 = 1{,}50 \qquad \text{mediana} = \sqrt[3]{1{,}5} \approx 1{,}1447$$

VAR(X)

$$\mathrm{Var}(X) = \int_{-\infty}^{+\infty} (x - \mathrm{E}(X))^2 \cdot f(x)\mathrm{d}x$$

$$\int_{-\infty}^{+\infty} (x - 1{,}0817)^2 \cdot f(x)\mathrm{d}x = \int_0^{\sqrt[3]{3}} (x - 1{,}0817)^2 \cdot x^2 \mathrm{d}x$$

$$= \int_0^{\sqrt[3]{3}} (x - 1{,}0817)^2 \cdot x^2 \mathrm{d}x = \int_0^{\sqrt[3]{3}} x^4 - 2{,}1634x^3 + 1{,}170075x^2 \mathrm{d}x$$

$$= \left[\frac{x^5}{5} - 2{,}1634\frac{x^4}{4} + 1{,}170075\frac{x^3}{3}\right]_0^{\sqrt[3]{3}} = \frac{\left(\sqrt[3]{3}\right)^5}{5} - 2{,}1634\frac{\left(\sqrt[3]{3}\right)^4}{4} + 1{,}170075\frac{\left(\sqrt[3]{3}\right)^3}{3} - 0$$

$$\approx 1{,}248 - 2{,}3401 + 1{,}170075 = 0{,}077975$$

Esercizio

- In un paese scandinavo il 60% delle ragazze ha i capelli Biondi (Evento B), il 30% li ha Rossi (Evento R), il 10% Mori (Evento M). Risulta poi che ha gli occhi Scuri il 10% delle Bionde, il 25% delle Rosse e il 50% delle More. Qual è la probabilità che la ragazza con cui ho fatto amicizia tramite Internet abbia occhi scuri. (Evento oS)? Che probabilità c'è che sia Bionda, dato che ha occhi scuri?

In un paese scandinavo il 60% delle ragazze ha i capelli Biondi (Evento B), il 30% li ha Rossi (Evento R), il 10% Mori (Evento M). Risulta poi che ha gli occhi Scuri il 10% delle Bionde, il 25% delle Rosse e il 50% delle More.

P(B)=	0,6		P(oS|B)=	0,1
P(R)=	0,3		p(oS|R)=	0,25
P(M)=	0,1		p(oS|M)=	0,5

P(oS?)

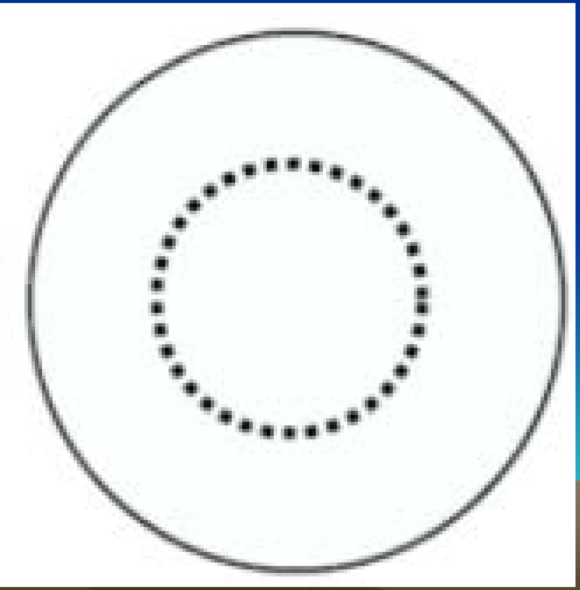
P(oS)=P(oS|B)*P(B)+P(oS|R)*P(R)+P(oS|M)*P(M)=0.185

P(B|oS?)

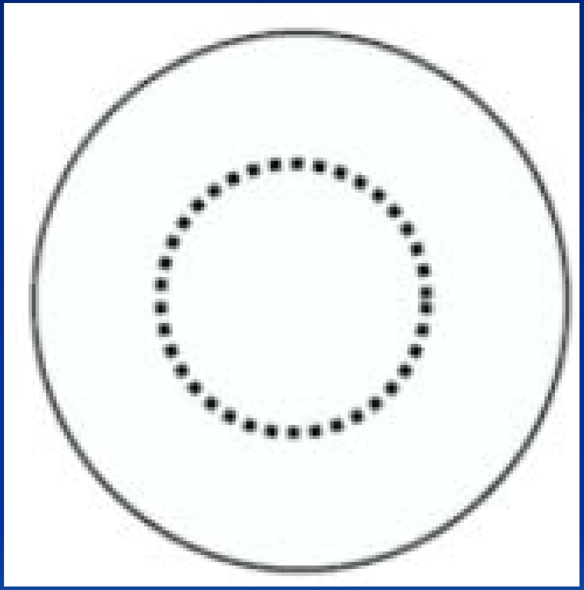
P(B|oS)=P(oS|B)*P(B)/P(oS)=0.324

Esercizio

- Consideriamo un bersaglio circolare con un diametro pari a 2r. Qual è la probabilità di colpire a caso un punto più vicino al centro che alla circonferenza?



Consideriamo un bersaglio circolare con un diametro pari a 2r. Qual è la probabilità di colpire a caso un punto più vicino al centro che alla circonferenza?



Area cerchio interno	**(r/2)*(r/2)*π**
Area cerchio esterno	**r*r*π**
Probabilità richiesta=	**1/4**

ESERCIZIO 2

- Un’azienda che assembla computer rileva difetti di assemblaggio nel 20% dei casi. Con riferimento ad un campione di 30 computer:
- si descrivano le caratteristiche delle variabili aleatorie “numero di difetti” e “frequenza relativa di difetti”;
- si scriva l’espressione che consente di determinare la probabilità che nel campione vi sia un numero di pezzi difettosi maggiore di 2 e un numero di pezzi difettosi compreso fra 2 e 5.

i) Universo Bernoulliano π=0.2

X = numero di difetti	P =X/n= frequenza relativa di difetti
X ~ B(30; 0,2)	P ~ B(30; 0,2)/30
E(X) = 30· 0,2 =6	E(P) = 0,2
VAR (X)= 30·0,2·0,8 = 4,8	VAR(P)=0,2·0,8/30= 0,053

ii)

$$\Pr(X > 2) = \sum_{x=3}^{30} \binom{30}{x} 0{,}2^{x} \cdot 0{,}8^{30-x} = 0{,}9645$$

$$\Pr(2 \leq X \leq 5) = \sum_{x=2}^{5} \binom{30}{x} 0{,}2^{x} \cdot 0{,}8^{30-x} = 0{,}4087$$