Processo stocastico di Poisson

PROCESSO STOCASTICO DI POISSON DI COSTANTE $\lambda \in \mathbb{R}^+$

$N(t)$ = 'n° eventi in $]0,t]$' $\quad t \in \mathbb{R}^+ \cup \{0\}$)

1) $N(0)=0$

2) INDIPENDENZA DEGLI INCREMENTI

3) STAZIONARIETÀ DEGLI INCREMENTI

4) $\lim_{h \to 0} \frac{P(N(h)=1)}{h} = \lambda \Rightarrow$ Per $h \ll 1$ $P(N(h)=1) \simeq \lambda h$

5) $\lim_{h \to 0} \frac{P(N(h) \geq 2)}{h} = 0 \Rightarrow$ per $h \ll 1$ $P(N(h)=2) \simeq 0$

N(t)

l'intervallo $]0,t]$ viene suddiviso in sottointervalli di uguale ampiezza $\frac{t}{m}$ con $m \gg 1$

$$P\left(N\left(\tfrac{t}{m}\right)=1\right) \underset{4)}{\simeq} \lambda \frac{t}{m} \quad ; \quad P\left(N\left(\tfrac{t}{m}\right) \geq 2\right) \underset{5)}{\simeq} 0$$

Il n° di eventi in ciascun intervallo di lunghezza $\frac{t}{m}$ si comporterà con la stessa distrib. di prob. (3))

inoltre il n° di eventi in uno dei sottointervalli non dipende dal n° di eventi in un altro sottointervallo (2))

$$P\left(N(t)=k\right) = P\left(\begin{array}{l}\text{in } k \text{ sottointervalli si osservi} \\ \text{un evento e nei restanti } m-k \\ \text{non si osservi nessun evento}\end{array}\right)$$

Cambiando t, cambia N(t), cambia il parametro della distribuzione. Per questo più di variabile dipendente dal tempo parliamo di una famiglia di variabili che hanno la stessa distribuzione con parametro variabile nel tempo.

$$\underset{\text{BINOMIALE}}{=} \binom{m}{k} P\left(N\left(\tfrac{t}{m}\right)=1\right)^k \left(1-P\left(N\left(\tfrac{t}{m}\right)=1\right)\right)^{m-k}$$

$$= \binom{m}{k}\left(\lambda\frac{t}{m}\right)^k\left(1-\frac{\lambda t}{m}\right)^{m-k}$$

$$\Rightarrow N(t) \sim B\left(m, \frac{\lambda t}{m}\right)$$

Questo m è introdotto artificialmente. Quello che ci interessa è il limite per h tendente a 0, cioè m tendente a infinito.

m è stato scelto
in maniera
arbitraria
con la condizione
$m \gg 1$

Siamo nelle condizione della legge degli eventi rari, il caso limite della poissoniana.

$$\Rightarrow N(t) \sim B\left(m, \frac{\lambda t}{m}\right)$$

$$\Bigg\downarrow \; m \to +\infty \; \left(\frac{\lambda t}{m} \to 0\right) \quad \text{legge degli eventi rari}$$

$$N(t) \sim Po\left(\cancel{m}\frac{\lambda t}{\cancel{m}}\right) = Po(\lambda t)$$

Parliamo di famiglia di variabili, perché ci sono tante poissoniane con parametro diverso, definito da lambda (costante della famiglia) e t (intervallo di tempo scelto per studiare la variabile casuale).

Poniamoci una seconda domanda: avrei anche bisogno di descrivere in una simulazione il tempo che intercorre tra un evento e l'altro. Occupiamoci degli intertempi.

Studio degli intertempi

STUDIO DEGLI INTERTEMPI

$X_1, X_2 \dots X_n \dots$

X_1 = tempo che intercorre tra l'istante iniziale e il verificarsi del primo evento

X_2 = tempo che intercorre tra il primo e il secondo evento

$\vdots$

X_k = tempo che intercorre tra il $(k-1)$esimo e il k-esimo evento

Dobbiamo pensarle come variabili casuali continue, perché possono verificarsi in qualunque tempo.

X_k V.C. CONTINUE

$s \in \mathbb{R}^+$

$$P(X_1 > s) = P(N(s) = 0) = \frac{(\lambda s)^0}{0!} e^{-\lambda s}$$

$$= e^{-\lambda s}$$

Questo significa che:

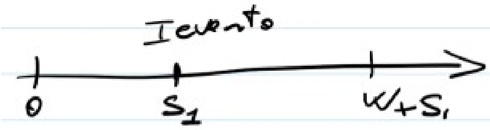
Se $s>0$

$$P(X_1\le s)=1-P(X_1>s)=\underbrace{1-e^{-\lambda s}}_{\text{FUNZIONE DI RIPARTIZIONE DI UNA V.C. ESPONENZIALE DI PARAMETRO } \lambda}$$

$$X_1 \sim \mathcal{E}(\lambda)$$

Se $w\in\mathbb{R}^+$ $\quad P(X_2>w)$

Vuol dire che è trascorso un certo lasso di tempo prima che si verificasse il primo evento. Ricomincio a misurare il tempo a partire dall'istante in cui si è verificato il primo evento e vedo cosa succede.



X2 non misura il tempo dall'istante iniziale, ma dall'istante in cui si è verificato il primo evento.

$$P(X_2>w)=P(X_2>w\,|\,X_1=s_1)$$

Questo perché in realtà io so dalle ipotesi del processo stocastico che intervalli di tempo disgiunti non si condizionano.

$$=P(\text{in }]s_1, s_1+w] \text{ non ci sono eventi})=$$

$$\underset{3)}{=}P(\text{in }]0,w] \text{ non ci sono eventi})$$

$$=P(N(w)=0)\underset{Po(\lambda w)}{=}\frac{(\lambda w)^0}{0!}e^{-\lambda w}=e^{-\lambda w}$$

$$\Downarrow$$

$$P(X_2\le w)=1-P(X_2>w)=\underbrace{1-e^{-\lambda w}}_{\text{FUNZIONE DI RIPARTIZIONE DI ESPONENZIALE DI PARAMETRO } \lambda}$$

$$\Rightarrow X_2 \sim \mathcal{E}(\lambda)$$

Ragionando nella stessa maniera otteniamo lo stesso risultato per qualunqu Xk.

$\Rightarrow$ GENERALIZZANDO IL RAGIONAMENTO

$$X_k \sim \mathcal{E}(\lambda)$$

C'è un collegamento tra poissoniana ed esponenziale: ogni volta che osservo un fenomeno che si comporta come una poissoniana, gli intertempi si comportano come un'esponenziale, perché siamo in presenza di un processo stocastico di Poisson.

Nel mondo reale, l'ipotesi 3 si verifica in un tempo non troppo lungo, lo stesso vale per l'indipendenza. Dobbiamo sempre tener conto che è un'approssimazione.

Teoria delle funzioni di variabili casuali

Cosa succede se considero in generale una funzione di una singola variabile casuale o una coppia o un vettore di variabili casuali (dal punto di vista matematico per quanto riguarda la funzione di ripartizione di probabilità, la funzione di densità di probabilità…).

CASO SINGOLA VARIABILE CASUALE

X v.c. $\quad Y = g(X)$

Sappiamo calcolare valor medio e varianza di Y conoscendo X.

X v.c. $\quad Y = g(X)$

$$E[Y] = E[g(X)] = \begin{cases} \sum_{k=1}^{m} g(x_k)\, p(x_k) & \text{se } X \in \{x_1 \dots x_m\} \text{ v.c. discreta} \\ \int_{-\infty}^{+\infty} g(x)\, f(x)\, dx & \text{se } X \text{ v.c. CONTINUA} \end{cases}$$

$$E[Y^2] = E[g^2(x)] = \begin{cases} \sum_{k=1}^{m} (g(x_k))^2\, p(x_k) & \text{se } X \text{ v.c. DISCRETA} \\ \int_{-\infty}^{+\infty} g(x)^2 f(x)\, dx & \text{se } X \text{ v.c. CONTINUA} \end{cases}$$

$\Downarrow$

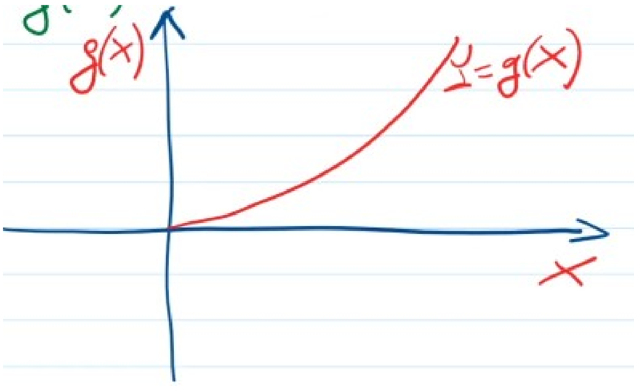
poi si determina

$$Var(Y) = E[Y^2] - E^2[Y]$$

Facciamoci domande sulla funzione di ripartizione, concentrandoci sulle variabili continue, le discrete sono molto specifiche, quindi è difficile ricavare proprietà generiche.

Concentrando l'attenzione sul caso di X v.c. continua, quanto vale $f_Y(a)$? $F_Y(a)$?

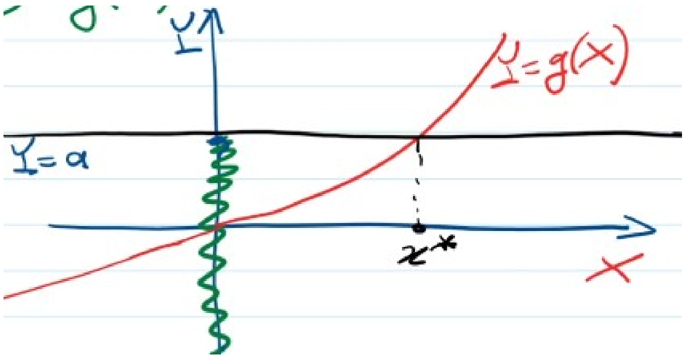
Dividiamo il caso continuo in 3 sottocasi (g monotona crescente, decrescente, non monotona).



Abbiamo una relazione invertibile crescente che lega X a Y.

X è nota insieme ad $f_X(x)$

$$F_Y(a) = P(Y \le a)$$



Questa condizione che X non possa essere più grande di x*.

$$a \in \mathbb{R}$$

$$F_Y(a) = P(Y \le a) = P(X \le x^*)$$

$$\boxed{\begin{array}{l} x^* : \ g(x^*) = a \\ x^* = g^{-1}(a) \end{array}}$$

È possibile, perché in quanto monotona, g è invertibile.

$$F_Y(a) = P(Y \leq a) = P(X \leq x^*) =$$

$$= P(X \leq g^{-1}(a)) = \int_{-\infty}^{g^{-1}(a)} f_X(x)\, dx$$

Posso anche calcolare la funzione di densità.

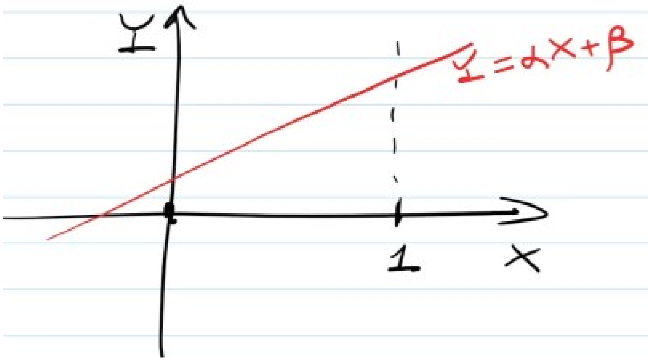
$$f_Y(a) = \frac{dF_Y(a)}{da} = \frac{d}{da} \int_{-\infty}^{g^{-1}(a)} f_X(x)\, dx$$

Torna utile il teorema fondamentale del calcolo integrale.

$$= f_X(g^{-1}(a)) \frac{d\,g^{-1}(a)}{da}$$

es. 1

$X \sim U(0,1)$ $\qquad$ $Y = \alpha X + \beta$ $\qquad$ con $\alpha > 0$, $\beta \in \mathbb{R}$



$$f_X(x) = \begin{cases} 1 & \text{se } 0 \leq x \leq 1 \end{cases}$$

$$Y = g(X) = \alpha X + \beta \Rightarrow X = \frac{Y - \beta}{\alpha} = g^{-1}(Y)$$

$$g^{-1}(a) = \frac{a-\beta}{\alpha} \Rightarrow \frac{d\,g^{-1}(a)}{d\,a} = \frac{1}{\alpha}$$

$$f_X(g^{-1}(a)) = \begin{cases} 1 & \text{se } 0 \leq g^{-1}(a) \leq 1 \\ 0 & \text{altrove} \end{cases}$$

$$f_X(g^{-1}(a)) = \begin{cases} 1 & \text{se } 0 \leq \frac{a-\beta}{\alpha} \leq 1 \Rightarrow \beta \leq a \leq \alpha+\beta \\ 0 & \text{altrove} \end{cases}$$

$$f_Y(a) = f_X(g^{-1}(a)) \cdot \frac{d g^{-1}(a)}{da} = \frac{1}{\alpha} \begin{cases} 1 & \text{se } \beta \leq a \leq \alpha+\beta \\ 0 & \text{altrove} \end{cases} =$$

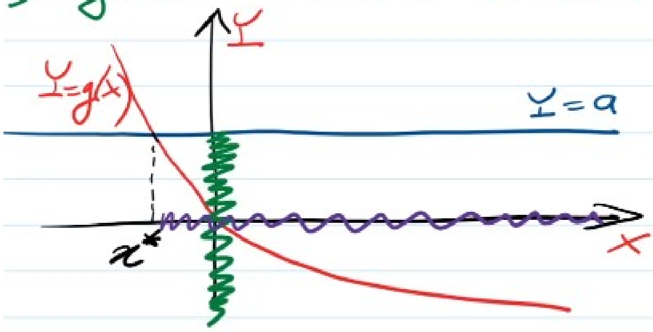
$$= \begin{cases} \frac{1}{\alpha} & \text{se } \beta \leq a \leq \alpha+\beta \\ 0 & \text{altrove} \end{cases}$$

Questi passaggi permettono di arrivare al risultato formale.

$$\Rightarrow Y \sim U(\beta, \alpha+\beta)$$

Y si comporta come un’uniforme definita tra beta e alfa + beta.

B) $g(x)$ FUNZIONE MONOTONA DECRESCENTE



$$\boxed{\begin{array}{l} x^*: \ g(x^*) = a \\ x^* = g^{-1}(a) \end{array}}$$

$$F_Y(a) = P(Y \leq a) = P(X \geq x^*)$$

La X deve assumere un valore tra x* e +infinito.

$$F_Y(a) = P(Y \leq a) = P(X \geq x^*) = \int_{x^*}^{+\infty} f_X(x)\,dx =$$

$$= \int_{g^{-1}(a)}^{+\infty} f_X(x)\,dx$$

Cerchiamo la funzione di densità della Y andando a derivare.

$$f_Y(a) = \frac{d F_Y(a)}{da} = \frac{d}{da}\left(- \int_{+\infty}^{g^{-1}(a)} f_X(x)\,dx \right) = -f_X(g^{-1}(a)) \frac{d g^{-1}(a)}{da}$$

Rispetto a prima c'è un meno, perché la derivata è negativa, una funzione di densità negativa è impossibile.

RIASSUMENDO

$$f_Y(a) = f_X(g^{-1}(a)) \frac{dg^{-1}(a)}{da} \quad \text{se } g \text{ MONOTONA CRESC.}$$

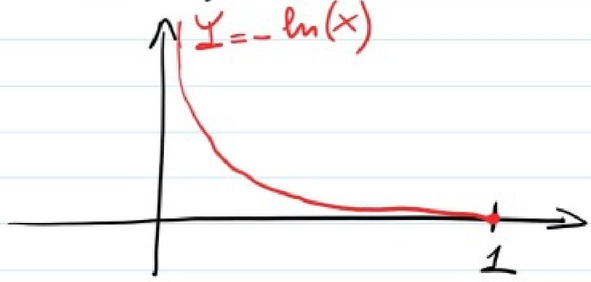
$$f_Y(a) = -f_X(g^{-1}(a)) \frac{dg^{-1}(a)}{da} \quad \text{se } g \text{ MONOTONA DECR.}$$

$\Downarrow$ SE g È UNA FUNZIONE MONOTONA

$$f_Y(a) = f_X(g^{-1}(a)) \left| \frac{dg^{-1}(a)}{da} \right|$$

es. 2

$X \sim U(0,1) \qquad Y = -\ln(X)$



$f_Y(a)$?

$$Y = g(X) = -\ln X \Rightarrow -Y = \ln X \Rightarrow X = e^{-Y} = g^{-1}(Y)$$

$$g^{-1}(a) = e^{-a} \Rightarrow \left| \frac{dg^{-1}(a)}{da} \right| = \left| -e^{-a} \right| = e^{-a}$$

$$f_X(g^{-1}(a)) = \begin{cases} 1 & \text{se } 0 \le g^{-1}(a) \le 1 \\ 0 & \text{altrove} \end{cases} = \begin{cases} 1 & \text{se } 0 \le e^{-a} \le 1 \\ 0 & \text{altrove} \end{cases}$$

$$= \begin{cases} 1 & \text{se } 0 \le e^{-a} \le 1 \text{ ovvero } a \ge 0 \\ 0 & \text{altrove} \end{cases}$$

$$f_Y(a) = f_X(g^{-1}(a)) \left| \frac{d g^{-1}(a)}{d a} \right| = e^{-a} \begin{cases} 1 & \text{se } a \ge 0 \\ 0 & \text{altrove} \end{cases} =$$

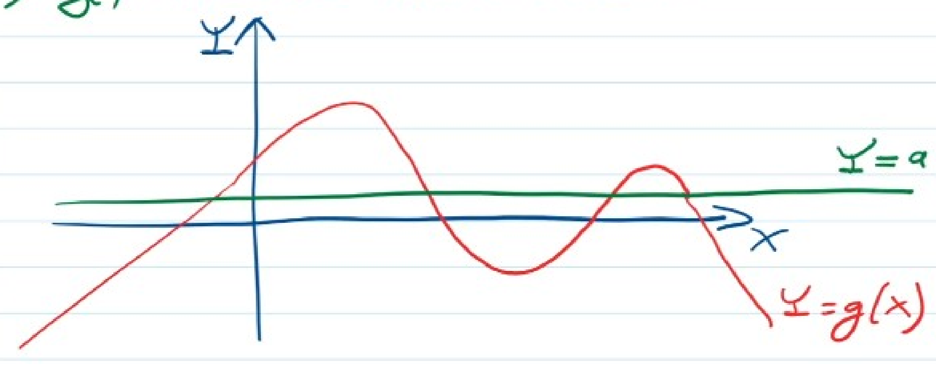
$$= \begin{cases} e^{-a} & \text{se } a \ge 0 \\ 0 & \text{altrove} \end{cases} \Rightarrow Y \sim E(1)$$

Partendo dalla variabile random, Y è il -ln, si determina che Y si comporta come un'esponenziale di parametro 1.

XCASA: COME POSSO TROVARE $Y \sim E(\lambda)$ partendo da $X \sim U(0,1)$?

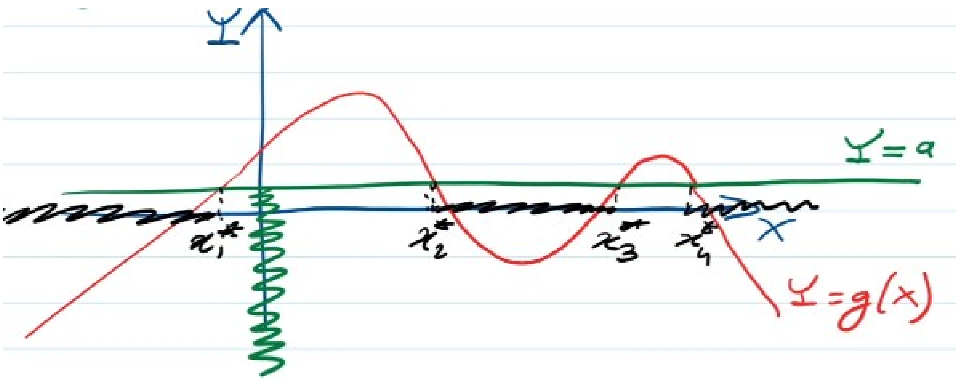
Rimane il caso di funzioni non monotone, diverse da trattare. Proprio in presenza di massimi e minimi relativi, c'è un modo diverso di trattare il problema.

c) $g(x)$ FUNZIONE NON MONOTONA



Localmente posso trovare parti monotone, ma nel complesso non lo è, quindi non si può invertire.

Ci possono essere più o 0 intersezioni con Y = a.



$$F_Y(a) = P(Y \le a) = P(X \le x_1^* \cup x_2^* \le X \le x_3^* \cup X \ge x_4^*)$$

$$= P(X \le x_1^*) + P(x_2^* \le X \le x_3^*) + P(X \ge x_4^*) =$$

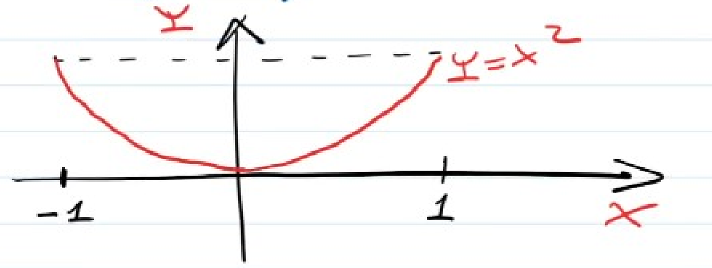
$$= \int_{-\infty}^{x_1^*} f_X(x)\,dx + \int_{x_2^*}^{x_3^*} f_X(x)\,dx + \int_{x_4^*}^{+\infty} f_X(x)\,dx$$

$$x_k^* = x_k^*(a) \qquad k = 1, 2, 3, 4$$

Dipende dalla funzione ovviamente.

es. 1

$X \sim U(-1, 1) \qquad Y = X^2$



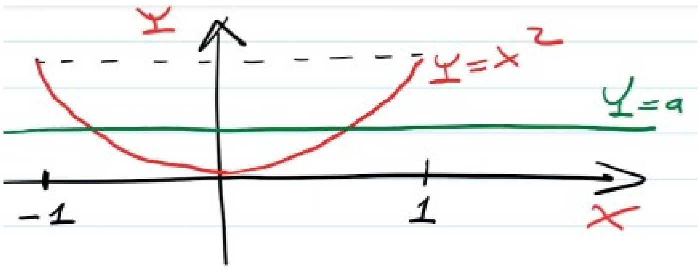
$F_Y(a)$?

$f_Y(a)$?

$$\text{Se } X \in [-1, 1] \Rightarrow Y \in [0, 1]$$

$$F_Y(a) = P(Y \leq a) = 0 \quad \text{se } a < 0$$

$$F_Y(a) = P(Y \leq a) = 1 \quad \text{se } a > 1$$



Vediamo il caso in cui a sia invece compreso nell'intervallo.

$$\boxed{\begin{array}{l} (x^*)^2 = a \\ \Rightarrow x^* = \sqrt{a} \end{array}}$$
$$\text{con } 0 \leq a \leq 1$$

$$f_X(a) = \begin{cases} \frac{1}{2} & \text{se } x \in [-1, 1] \\ 0 & \text{altrove} \end{cases}$$

½ perché l'intervallo va da -1 a 1, quindi è di lunghezza 2.

$$\text{se } 0 \leq a \leq 1$$
$$F_Y(a) = P(Y \leq a) = \int_{-x^*}^{+x^*} f_X(x)\,dx = \int_{-x^*}^{+x^*} \frac{1}{2}\,dx =$$
$$= \left[\frac{x}{2}\right]_{-x^*}^{+x^*} = x^* = \sqrt{a}$$

Riscriviamo nel complesso la funzione di ripartizione.

$$F_Y(a) = \begin{cases} 0 & \text{se } a < 0 \\ \sqrt{a} & \text{se } 0 \leq a \leq 1 \\ 1 & \text{se } a > 1 \end{cases}$$

$$f_Y(a) = \frac{dF_Y(a)}{da} = \begin{cases} 0 & \text{se } a < 0 \\ \frac{1}{2\sqrt{a}} & \text{se } 0 \leq a \leq 1 \\ 0 & \text{se } a > 1 \end{cases}$$

C’è un problema di convergenza, ma integrando è convergente.

Esercizi

$\boxed{5\ 2014,\ 4\ 2013,\ 3\ 2014}$

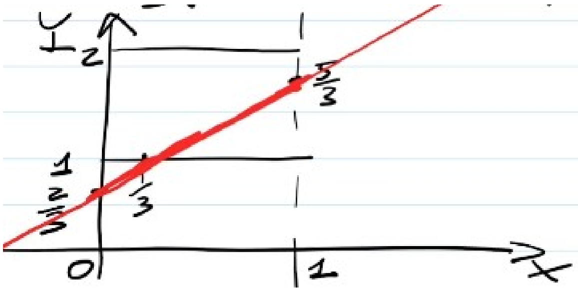
5 del 2014

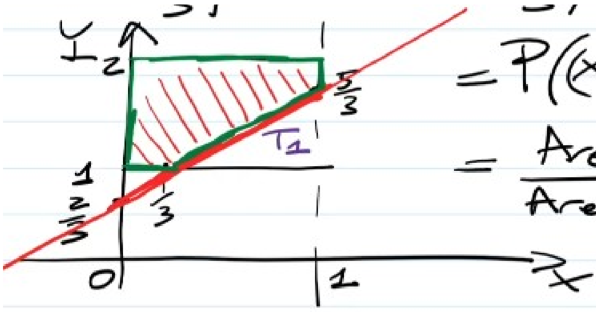
$X \sim U(0,1)$

$Y \sim U(1,2)$

$D = |Y - X| = Y - X$ ($Y \le X$ sempre per def.)

$$P\left(D > \frac{2}{3}\right) = P\left(Y - X > \frac{2}{3}\right) = P\left(Y > X + \frac{2}{3}\right)$$





$$= P((X,Y) \in \text{PENTAGONO}) =$$

$$= \frac{\text{Area PENTAGONO}}{\text{Area RETTANGOLO}} =$$

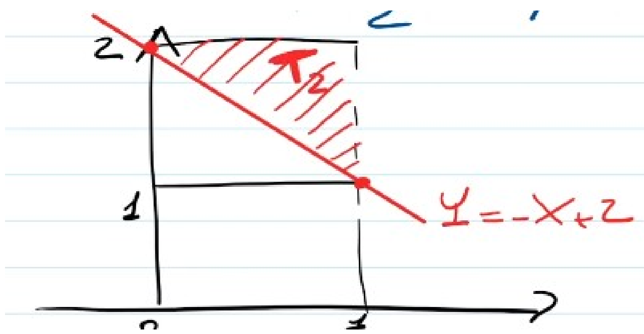
$$= \frac{\text{Area RETTANGOLO} - \text{Area } T_1}{\text{Area RETTANGOLO}}$$

$$= \frac{1^2 - \left(\frac{2}{3} \cdot \frac{2}{3} \cdot \frac{1}{2}\right)}{1^2} =$$

$$= 1 - \frac{2}{9} = \frac{7}{9}$$

$$Z = \frac{X+Y}{2} \qquad P(Z=1)=0$$

$$P(Z>1) = P\left(\frac{X+Y}{2}>1\right) = P(Y>-X+2)$$



$$= P((X,Y)\in T_2) =$$

$-X+2$

$$= \frac{\text{Area } T_2}{\text{Area Rettangolo}} =$$

$$= \frac{1\cdot 1\cdot \frac{1}{2}}{1^2} = \frac{1}{2}$$

In generale, la somma di due variabili casuali uniformi non è una variabile casuale uniforme, non sono riproducibili.

Non posso pensare che X + Y sia una variabile uniforme definita tra 0 e 2 oppure tra 1 e 3. Sicuramente X + Y avrà un range che va dalla somma dei valori minimi di X e Y alla somma dei valori massimi di X e Y.

La sua funzione di densità in generale non sarà una funzione di densità costante. Di solito, sommare 2 variabili casuali uniformi da luogo a una funzione di densità che si comporta come una funzione di densità a tenda, cioè a triangolo.

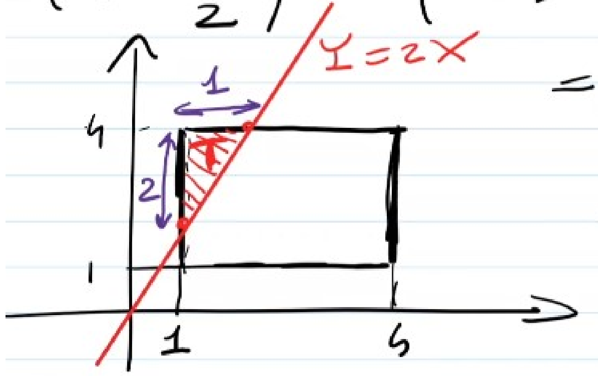
Q4 del 2013

$X, Y \sim U(1,4)$ $\quad$ X e Y INDIPENDENTI

$$P(X=Y)=0$$

$$P\left(X<\frac{Y}{2}\right)=P(Y>2X)$$

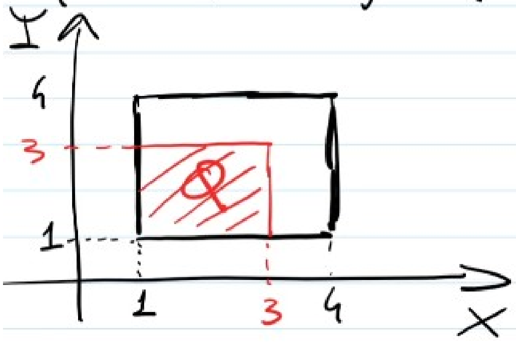
$$P\left(X<\frac{Y}{2}\right)=P(Y>2X)=P((X,Y)\in T)=$$



$$=\frac{\text{Area } T}{\text{Area Rettangolo}}=\frac{1\cdot 2\cdot \frac{1}{2}}{3^2}=$$

$$=\frac{1}{9}$$

$$P(X<3, Y<3)=P((X,Y)\in Q)=\frac{\text{Area } Q}{\text{Area Rettangolo}}=\frac{2^2}{3^2}=\frac{4}{9}$$



Ci sono delle domande che riguardano le funzioni della coppia X e Y.

$$T=2X+2Y$$

T non può essere uniforme, però posso sfruttare le proprietà del valor medio.

Se $X \sim U(\alpha,\beta)$

$$E[X]=\frac{\alpha+\beta}{2}=\frac{5}{2} \quad (\alpha=1, \beta=4)$$

$$Var(X)=\frac{(\beta-\alpha)^2}{12}=\frac{9}{12}=\frac{3}{4}$$

$$E[T] = E[2X + 2Y] = 2E[X] + 2E[Y] = 2\cdot\frac{5}{2} + 2\frac{5}{2} = 10$$

$$Var(T) \underset{\text{INDIP.}}{=} Var(2X) + Var(2Y) = 4\,Var(X) + 4\,Var(Y) =$$

$$= 4\cdot\frac{3}{4} + 4\cdot\frac{3}{4} = 6$$

INDIP. $+\downarrow$

$$= 4\cdot\frac{3}{4} + 4$$

$$Var(\alpha X) = \alpha^2\, Var(X)$$

$$E[X^2] = Var(X) + E^2[X] = \frac{3}{4} + \frac{25}{4} = \frac{28}{4} = 7$$

$$A = XY$$

$$E[A] = E[XY] \underset{\text{INDIPENDENZA}}{=} E[X]\,E[Y] = \left(\frac{5}{2}\right)^2 = \frac{25}{4}$$

$$Var(A) = E[A^2] - E^2[A] = E[(XY)^2] - \left(\frac{25}{4}\right)^2 =$$

$$= E[X^2Y^2] - \left(\frac{25}{4}\right)^2 = E[X^2]\,E[Y^2] - \left(\frac{25}{4}\right)^2$$

$$Var(A) = E[A^2] - E^2[A] = E[(XY)^2] - \left(\frac{25}{4}\right)^2 =$$

$$= E[X^2Y^2] - \left(\frac{25}{4}\right)^2 = E[X^2]\,E[Y^2] - \left(\frac{25}{4}\right)^2 =$$

$$= 7\cdot 7 - \frac{625}{16} = 49 - \frac{625}{16}$$

In generale non è banale rispondere alla domanda qual è la funzione di ripartizione di una somma o di un prodotto di una coppia di variabili casuali, a maggior ragione se dipendenti.

Q3 del 2014

$$\boxed{5B + 8R}$$

4 palline estratte senza reimm.

$X=$ 'n° palline rosse estratte'

$$Y=\begin{cases} 1 & \text{se } B_1 \\ 0 & \text{se } R_1 \end{cases}$$

Non posso usare le combinazioni perché dovrò pensare di estrarre una pallina alla volta.

$$X \in \{0, 1, 2, 3\}$$

Y \ X	0	1	2	3
0	0			
1				0

$$p(0,1) = P(B_1 B_2 B_3) = \frac{5 \cdot 4 \cdot 3}{13 \cdot 12 \cdot 11}$$

$$p(3,0) = P(R_1 R_2 R_3) = \frac{8 \cdot 7 \cdot 6}{13 \cdot 12 \cdot 11}$$

$$p(1,0) = P(R_1 B_2 B_3) = \frac{8 \cdot 5 \cdot 4}{13 \cdot 12 \cdot 11}$$

$$p(1,1) = P(B_1 R_2 B_3 \cup B_1 B_2 R_3) =$$

$$= \frac{(5 \cdot 8 \cdot 4 + 5 \cdot 4 \cdot 8)}{13 \cdot 12 \cdot 11}$$

$$p(2,0) = P(R_1 R_2 B_3 \cup R_1 B_2 R_3) =$$

$$= \frac{8 \cdot 7 \cdot 5 + 8 \cdot 5 \cdot 7}{13 \cdot 12 \cdot 11}$$

$$p(2,1) = P(B_1 R_2 R_3) = \frac{5 \cdot 8 \cdot 7}{13 \cdot 12 \cdot 11}$$

Da questo, i valori medi, le marginali, le varianze…

$$E[XY] = 1 \cdot 1 \cdot p(1,1) + 2 \cdot 1 \cdot p(2,1)$$