Variabile casuale uniforme

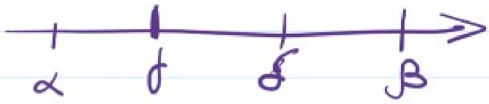
nedì 8 novembre 2021 17:55

$$X \sim U(\alpha, \beta) \qquad f(x) = \begin{cases} \frac{1}{\beta-\alpha} & \text{se } x \in [\alpha, \beta] \\ 0 & \text{altrove} \end{cases}$$

$$E[X] = \frac{\alpha+\beta}{2} \qquad F(x) = \begin{cases} 0 & \text{se } x < \alpha \\ \frac{x-\alpha}{\beta-\alpha} & \text{se } \alpha \leq x \leq \beta \\ 1 & \text{se } x > \beta \end{cases}$$

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

$$\alpha < \gamma < \delta < \beta$$



$$P(X \in [\gamma, \delta]) = \frac{\delta-\gamma}{\beta-\alpha}$$

Coppie di variabili casuali uniformi e indipendenti

Coppie di v.c. uniformi e indipendenti

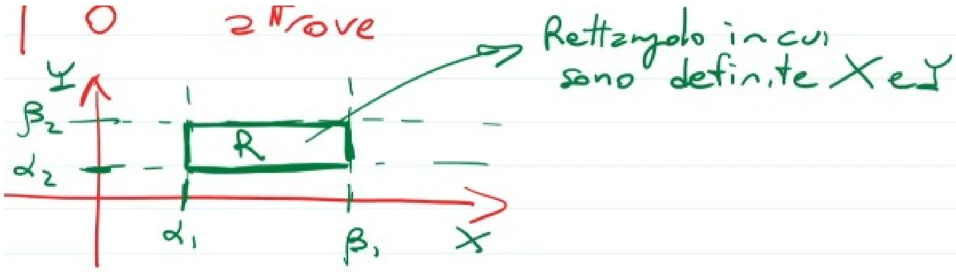
$$\left.\begin{array}{l} X \sim U(\alpha_1, \beta_1) \\ Y \sim U(\alpha_2, \beta_2) \end{array}\right\} \text{INDIP.} \qquad \text{se } \alpha_1, \alpha_2, \beta_1, \beta_2 \in \mathbb{R} \quad \begin{array}{l} \alpha_1 < \beta_1 \\ \alpha_2 < \beta_2 \end{array}$$

Sotto queste condizioni è facile scrivere la funzione di densità congiunta.

$$f(x,y) \underset{\text{INDIP.}}{=} f_X(x)\, f_Y(y) = \begin{cases} \frac{1}{\beta_1-\alpha_1} & \text{se } x \in [\alpha_1, \beta_1] \\ 0 & \text{altrove} \end{cases} \cdot \begin{cases} \frac{1}{\beta_2-\alpha_2} & \text{se } y \in [\alpha_2, \beta_2] \\ 0 & \text{altrove} \end{cases}$$

$$= \begin{cases} \frac{1}{(\beta_1-\alpha_1)(\beta_2-\alpha_2)} & \text{se } x \in [\alpha_1, \beta_1] \text{ e } y \in [\alpha_2, \beta_2] \\ 0 & \text{altrove} \end{cases}$$

È una funzione di densità congiunta diversa da 0 in un rettangolo di definizione che è il rettangolo che corrisponde ai valori di X e Y consentiti dalla definizione di variabile uniforme.

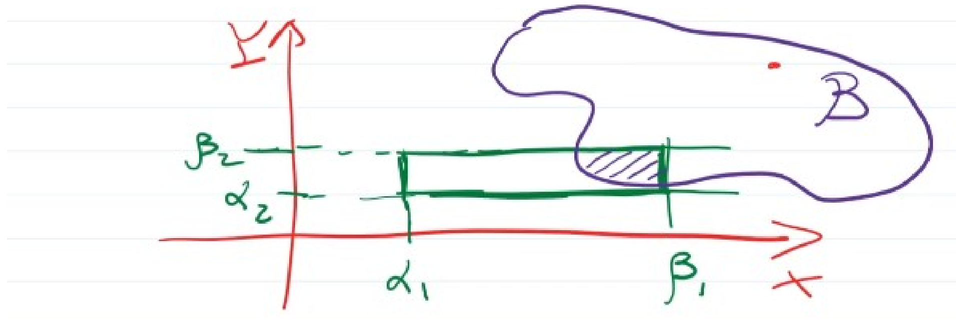


$$\frac{1}{(\beta_1-\alpha_1)(\beta_2-\alpha_2)}=\frac{1}{A_R}$$

Perché introdurre una coppia di variabili uniformi? Perché in molti problemi, quando bisogna rappresentare delle coordinate scelte a caso in un certo dominio oppure delle variabili casuali indipendenti che sono scelte a caso in certi intervalli, è naturale utilizzare una coppia di variabili casuali definite in questa maniera. Per la sua semplicità, questo caso gode di proprietà speciali.

$$A_R=(\beta_1-\alpha_1)(\beta_2-\alpha_2)=$$
$$= \text{AREA RETTANGOLO}$$

$$P((X,Y)\in B)? \quad \text{con } B\subset \mathbb{R}^2$$



È chiaro che per definizione le variabili casuali possono stare ciascuna nell’intervallo di definizione e come coppia nel rettangolo. Intanto, la probabilità che X e Y appartengano a B, è la probabilità che X e Y appartengano a B intersecato il rettangolo in cui sono definite. È chiaro che non possono essere definite nel puntino rosso perché è fuori dal loro dominio di definizione.

$$P((X,Y)\in B)=P((X,Y)\in B\cap R)$$

R rettangolo di definizione.

$$= \iint_{B \cap R} f(x,y)\, dx\, dy = \iint_{B \cap R} \frac{1}{A_R}\, dx\, dy =$$

$$= \frac{1}{A_R} \iint_{B \cap R} 1\, dx\, dy = \frac{\text{Area}(B \cap R)}{\text{Area } R}$$

In questo caso specifico, ogni volta che abbiamo necessità di calcolare la probabilità che la coppia sia in una regione, basta fare il rapporto tra l'area di questa regione intersecata il rettangolo di definizione diviso l'area del rettangolo.

es. 1

Autobus passa ad una certa fermata in un orario casuale tra le 8:10 e le 8:20.

Anna esce da casa in un orario casuale tra le 8:00 e le 8:15.

Qual è la probabilità che Anna prenda l'autobus se la fermata è sotto casa?

Ci sono tanti modi per studiare questo problema. Il più semplice è quello di modellare due variabili casuali (arrivo dell'autobus, arrivo di Anna).

Rimaneggiamo il tutto lavorando coi minuti (tutti gli orari sono compresi tra le 8:00 e le 8:20).

X = 'orario di arrivo dell'autobus'

$X \sim U(10, 20)$ in minuti

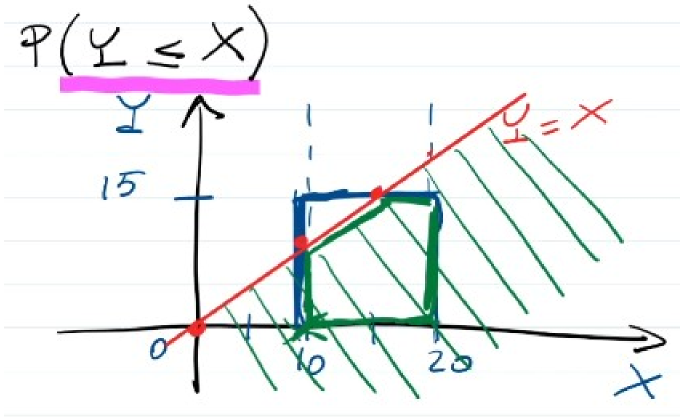
Y = 'arrivo di Anna'

$Y \sim U(0, 15)$ in minuti

$$P(Y \leq X)$$

Se l'arrivo di Anna avviene prima o in contemporanea dell'autobus immaginiamo che l'autobus possa essere preso.

Abbiamo due strade: o impostiamo l'integrale doppio o ragioniamo con i domini e il rettangolo di definizione.



Y <= X è un semipiano con bordo Y = X. La X e la Y vogliamo che siano nel pentagono evidenziato.

$$P(Y \leq X) = \frac{\text{Area pentagono}}{\text{Area } R}$$

$$= \frac{\text{Area } R - \text{Area } T}{(20-10)(15-0)} =$$

$$= \frac{150 - \frac{5^2}{2}}{150} =$$

$$= \frac{300 - 25}{300} =$$

$$= \frac{275}{300} = \frac{11}{12}$$

Oppure (poco consigliato):

$$P(Y \le X) = \int_{-\infty}^{\infty}\left(\int_{-\infty}^{x} f(x,y)\,dy\right)dx = \int_{10}^{20}\left(\int_{0}^{\min(x,15)} \frac{1}{150}\,dy\right)dx =$$

$$= \int_{10}^{15}\left(\int_{0}^{x} \frac{1}{150}\,dy\right)dx + \int_{15}^{20}\left(\int_{0}^{15} \frac{1}{150}\,dy\right)dx =$$

$$= \int_{10}^{15} \frac{x}{150}\,dx + \int_{15}^{20} \frac{15}{150}\,dx = \frac{1}{150}\left(\left[\frac{x^2}{2}\right]_{10}^{15} + 15\left[x\right]_{15}^{20}\right) =$$

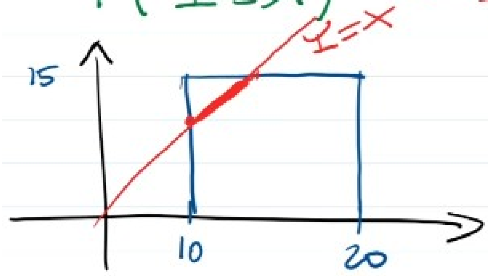
$$= \frac{1}{150}\left(\frac{15^2-100}{2} + 15(20-15)\right) = \frac{1}{150}\left[\frac{225-100}{2} + 75\right] =$$

$$= \frac{1}{300}(125+150) = \frac{275}{300} = \frac{11}{12}$$

Bisogna fare più attenzione agli estremi di integrazione.

Quale è la prob. che Anna e l'autobus arrivino simultaneamente?

$$P(Y = X) = \frac{\text{Area segmento}}{\text{Area } R} = 0$$



Questa osservazione grafica è un'osservazione generale: quando abbiamo due variabili casuali continue, la probabilità che siano identiche è nulla.

N.B. In generale se
(X, Y) coppia di v. c.
congiuntamente
continue

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

infatti:

$$P(X=Y)=\int_{-\infty}^{+\infty}\left(\int_{y}^{y} f(x,y)\,dx\right)dy$$

$$=\underbrace{\int_{-\infty}^{+\infty}\left(\int_{y}^{y} f(x,y)\,dx\right)}_{=0}dy=0$$

Che significato ha? X e Y assumono valori con continuità ed è praticamente impossibile osservare che assumono lo stesso valore reale, visto che possono assumere un numero di valori non numerabile (intervallo o asse reale).

Graficamente si vede che è possibile che X sia uguale Y. Certamente, però se faccio il rapporto tra un oggetto unidimensionale e un'area verrà sempre 0. Il segmento è unidimensionale, ha area nulla.

Queste coppie di variabili casuali vengono impiegate nei metodi Monte Carlo.

Metodi Monte Carlo

Sono stati sviluppati in periodi storici diversi e che ancora oggi vengono introdotti e utilizzati per stimare delle quantità deterministiche utilizzando la probabilità.

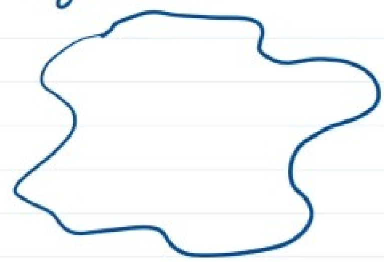
METODI MONTE CARLO

Metodi che stimano quantità deterministiche utilizzando la teoria della probabilità e gli eventi casuali.

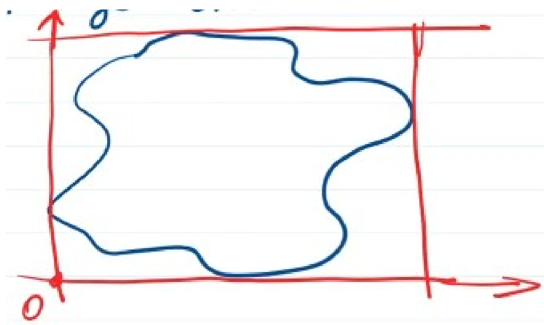
Il nome è stato dato da Enrico Fermi. Questi metodi erano stati introdotti per i primi elaboratori per simulazioni di fisica e ricorda il fatto che a Monte Carlo ci siano dei casinò, per ricordare che questi metodi abbiano alla base un fattore aleatorio.

Determinazione di un'area

es. lago dalla forma irregolare



Immaginiamo di racchiudere il lago in un rettangolo che lo contenga tutto. Ovviamente ce ne sono infiniti, ma cerchiamo quello che circoscrive il lago (o poco più grande).



Potremmo usare un sistema di riferimento.

Immaginiamo di voler misurare l'area della superficie del lago, lanciando delle palle di cannone all'interno del rettangolo.

Come misurare la superficie del lago lanciando palle di cannone all'interno del rettangolo che circoscrive il lago?

Immaginiamo di lanciare delle palle puntiformi. Ovviamente ci saranno dei puntini che finiranno all'interno del lago, mentre altre no.

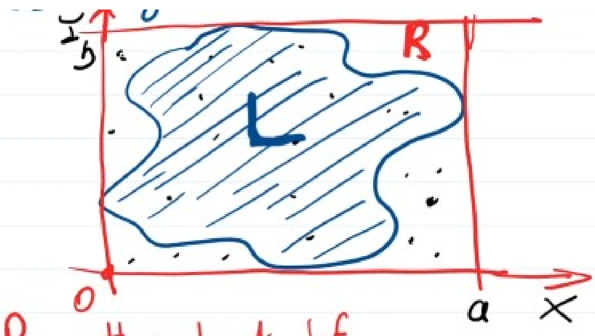
(X, Y) coordinate del punto di arrivo di una palla di cannone

$X \sim U(0, a)$, $Y \sim U(0, b)$ X e Y INDIPENDENTI

n_L = n° di palle di cannone che finiscono nel lago L

n_{TOT} = n° di lanci complessivi

Conoscendo nL e n totale si può stimare l'area del lago.



R = rettangolo di def.

$$P((X, Y) \in L) \underset{\substack{\downarrow \\ X \text{ e } Y \\ \text{uniformi e indip.}}}{=} \frac{\text{Area } L}{\text{Area } R}$$

Per il corollario di Bernoulli: la frequenza relativa di un evento converge in prob. alla sua prob. teorica quando il n° di prove $\to +\infty$

$$P((X, Y) \in L) = P(A)$$

A = 'una palla lanciata a caso nel rettangolo cade nel lago'

$$f_A = \frac{n_L}{n_{TOT}} \qquad \text{per } n_{TOT} \gg 1$$

$$f_A \approx P(A) \Rightarrow \frac{n_L}{n_{TOT}} \approx \frac{\text{Area } L}{\text{Area } R}$$

$$\Downarrow$$

$$\text{Area } L \simeq \frac{n_L \, \text{Area } R}{n_{TOT}}$$

Questa misura sarà tanto migliore quanto più esperimenti vengono fatti.

In un mondo settecentesco o ottocenteschi si poteva pensare di lanciare delle palle di cannone, oggi si fa una simulazione al computer, generando numeri casuali distribuiti uniformemente.

ANTICIPAZIONE: Se $U \sim U(0,1)$

$$\gamma U \sim U(0,\gamma)$$

Queste tecniche ricordano un metodo per stimare il pi greco.

Vediamo il trisavolo dei metodi Monte Carlo. Stima il pi greco in maniera naive, il problema degli spilli o degli aghi.

PROBLEMA DEGLI SPILLI (AGHI) DI BUFFON (1777)

Immaginiamo di avere un pavimento e di dividerlo in strisce tutte identiche, parallele ed equidistanti.

Superficie piana suddivisa da rette parallele equidistanti

La distanza tra 2 rette vicine è 2b.

Si lanciano a caso spilli su questa superficie piana: lunghezza di ogni spillo $2a$ con ($2a < 2b$)

Uno spillo perpendicolare alla direzione delle rette parallele potrebbe essere completamente all'interno di una striscia, senza toccare le rette.

Si può stimare il pi greco, come?

Prob. che uno spillo intersechi o tocchi una retta?

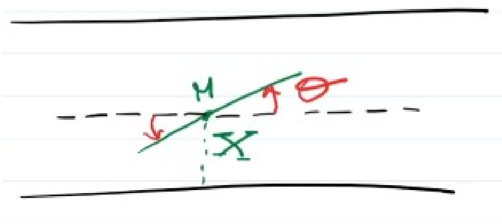
1) Considero il punto medio dello spillo e la sua distanza dalla retta più vicina

X = 'lunghezza del segmento di perpendicolare che va da M alla retta più vicina ovvero la distanza tra M e la retta più vicina'

È chiaro che X sia distribuito in maniera uniforme, non ci sono distanze favorite. In generale possiamo supporre che X sia distribuito tra 0 (M sta sulla retta) e b (metà della distanza tra una retta e l'altra).

$X \sim U(0, b)$ (se la distanza di M da una certa retta fosse $> b$ si considererebbe un'altra retta come retta più vicina)

Non ci da completamente la posizione dello spillo. Abbiamo bisogno di una seconda variabile che dica l'inclinazione dello spillo rispetto al fascio di rette. Disegniamo una retta con direzione comune passante per M e misuriamo theta.



2) Introduco θ angolo formato tra lo spillo e la direzione comune alle rette parallele

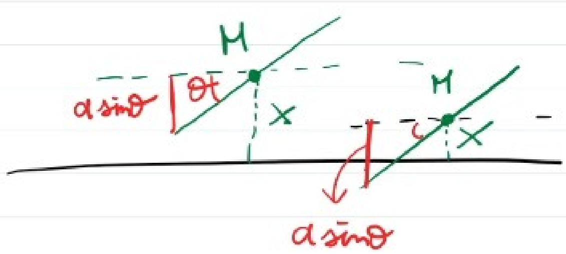
$$\theta \sim U(0,\pi)$$

3) È ragionevole pensare che θ ed X siano INDIPENDENTI

Inclinazione e distanza non ha senso che si influenzino.

$$(X,\theta) \quad \text{INDIP. CON} \quad f(x,\theta) = \begin{cases} \frac{1}{b-0}\,\frac{1}{\pi-0} & \text{se } x\in[0,b],\ \theta\in[0,\pi] \\ 0 & \text{altrimenti} \end{cases}$$

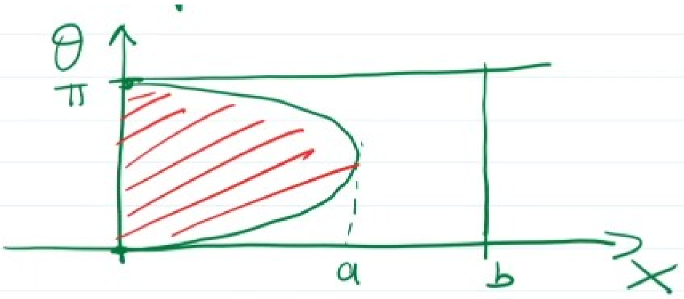
4) Sotto quali condizioni si osserva l'intersezione tra spillo e retta più vicina?



$$X \leq a\sin\theta$$

CONDIZIONE DI INTERSEZIONE

$$P\left(\begin{array}{c}\text{uno spillo intersechi}\\ \text{retta più vicina}\end{array}\right) = P(X \leq a \sin\theta)$$



È una regione individuata dal seno di theta, quindi si dovrà fare l'integrale doppio.

$$= \int_0^{\pi}\left(\int_0^{a\sin\theta} \frac{1}{b\pi}\,dx\right)d\theta = \int_0^{\pi} \frac{1}{b\pi}\, a\sin\theta\, d\theta =$$

$$= \frac{a}{b\pi}\int_0^{\pi} \sin\theta\, d\theta = \frac{a}{b\pi}\left[-\cos\theta\right]_0^{\pi} = \frac{2a}{b\pi}$$

Per il corollario di Bernoulli:

$A =$ 'lo spillo interseca la retta più vicina'

$$P(A) = \frac{2a}{b\pi} \approx f_A = \frac{n_A}{n_{TOT}}$$

$n_A \longrightarrow$ n° spilli che intersecano una retta

$n_{TOT} \longrightarrow$ n° spilli lanciati

vale se $n_{TOT} \gg 1$

$$\pi \approx \frac{2a}{b}\frac{n_{TOT}}{n_A}$$

Suscitò scalpore individuare pi greco in maniera probabilistica. C'è il problema di fare tanti esperimenti, qualcuno ha millantato di averlo fatto, ma ottenendo risultati forse forzati. In teoria si può simulare al computer, ma c'è il problema di dover generare un numero tra 0 e pi greco, quindi il pi greco bisogna conoscerlo già.