Calcolo delle Probabilità e Statistica
Variabili Aleatorie Continue

Ionel Eduard Stan[1]

[1]Dip. di Matematica e Informatica, Università di Ferrara ioneleduard.stan@unife.it

Le variabili aleatorie che hanno un dominio continuo sono molto comuni. La velocità di un veicolo che viaggia su un'autostrada è un esempio; se la velocità è misurata da un velocimetro, possiamo vedere la lettura del velocimetro come una variabile aleatoria discreta, ma se vogliamo misurare la velocità esatta abbiamo bisogno di una variabile aleatoria continua.

Definizione 1 (Variabile aleatoria continua)**.** *Una variabile aleatoria X si dice* continua *se il suo codominio è continuo.* ♦

Definizione 2 (Funzione di densità di probabilità)**.** *Sia X una variabile aleatoria continua. La* funzione di densità di probabilità (probability density function*, in inglese), o semplicemente PDF, denotata con f_X, è definita come:*

$$P(\{X \in B\}) = \int_B f_X(x)dx,$$

per ciascun $B \subseteq \mathbb{R}$. ♦

In particolare, la probabilità che il valore di X cada dentro un intervallo è:

$$P(a \leq X \leq b) = \int_a^b f_X(x)dx,$$

e può essere interpretata come l'area sotto la curva della PDF. Si noti che, per un singolo valore a, abbiamo che $P(\{X = a\}) = \int_a^a f_X(x)dx = 0$. Per questo motivo, includere o escludere gli estremi dell'intervallo non ha nessun effetto sulla sua probabilità:

$$P(\{a \leq X \leq b\}) = P(\{a < X < b\}) = P(\{a \leq X < b\}) = P(\{a < X \leq b\}).$$

Sia X una variabile aleatoria continua con PDF f_X, possiamo elencare alcune proprietà su f_X:

- $f_X(x) \geq 0$, per ciascun x.
- $\int_{-\infty}^{\infty} f_X(x)dx = 1$.
- Se δ è molto piccolo, allora $P([x, x+\delta]) = P(\{x \leq X \leq x+\delta\}) \approx f_X(x) \cdot \delta$.
- Per ogni $B \subseteq \mathbb{R}$,

$$P(\{X \in B\}) = \int_B f_X(x)dx.$$

Definizione 3 (Valore atteso)**.** *Il* valore atteso *(*expected value*, in inglese) di una variabile aleatoria continua X è definita come:*

$$E[X] = \int_{-\infty}^{\infty} x \cdot f_X(x)dx.$$

♦

Sia X una variabile aleatoria continua con PDF f_X, elenchiamo alcune proprietà sul valore atteso:

- Il valore atteso di X è definito come:

$$E[X] = \int_{-\infty}^{\infty} x \cdot f_X(x)dx.$$

- La regola del valore atteso per una funzione $g(X)$ ha la forma:

$$E[g(X)] = \int_{-\infty}^{\infty} g(x)f_X(x)dx.$$

- La varianza di X è definita come:

$$var(X) = E[(X - E[X])^2] = \int_{-\infty}^{\infty} (x - E[X])^2 \cdot f_X(x)dx.$$

- Abbiamo che $0 \leq var(X)$.
- Se $Y = a \cdot X + b$, per a e b costanti, allora:

$$E[Y] = a \cdot E[X] + b, \quad var(Y) = a^2 \cdot var(X).$$

Definizione 4 (Variabile aleatoria uniforme)**.** *Sia X una variabile aleatoria continua. Se la PDF di X è definita come:*

$$f_X(x) = \begin{cases} \frac{1}{b-a} & \text{if } a \leq x \leq b, \\ 0 & \text{altrimenti,} \end{cases}$$

allora X è una variabile aleatoria uniforme *(*uniform random variable*, in inglese). Diciamo che X ha la* distribuzione uniforme *(*uniforma distribution*, in inglese) nell'intervallo $[a, b]$, in simboli $X \sim \mathcal{U}([a, b])$, $E[X] = \frac{a+b}{2}$ e $var(X) = \frac{(b-a)^2}{12}$.* ♦

Definizione 5 (Variabile aleatoria esponenziale)**.** *Una variabile aleatoria continua X si dice* variabile aleatoria esponenziale *(*exponential random variable*, in inglese) se la sua PDF è:*

$$f_X(x) = \begin{cases} \lambda e^{-\lambda x} & \text{if } x \geq 0, \\ 0 & \text{altrimenti,} \end{cases}$$

dove λ è un parametro positivo. Diciamo che X ha la distribuzione esponenziale *(*exponential distribution*, in inglese) con parametro λ, in simboli $X \sim \mathcal{EXP}(\lambda)$, $E[X] = \frac{1}{\lambda}$ e $var(X) = \frac{1}{\lambda^2}$.* ♦

Una variabile aleatoria esponenziale può, ad esempio, modellare il tempo necessario affinché un certo incidente di interesse succede, come un messaggio che arriva ad un computer, una lampadina che si brucia, ecc.

Esempio 1. Il tempo necessario finché un meteorite atterra per la prima volta nel deserto del Sahara è modellato come una variabile aleatoria esponenziale con una media di 10 giorni. Il tempo adesso è mezzanotte. Qual'è la probabilità che il meteorite atterra per la prima volta tra le 06 : 00 e le 18 : 00 del primo giorno? ♢

Definizione 6 (Funzione di distribuzione cumulativa)**.** *Sia X una variabile aleatoria. Se x è uno dei valori di X, la* funzione di distribuzione cumulativa *(*cumulative distribution function*, in inglese), o semplicemente CDF, denotata con $F_X(x)$, è la probabilità dell'evento $\{X \leq x\}$, cioè:*

$$F_X(x) = P(\{X \leq x\}) = \begin{cases} \sum_{k \leq x} p_X(x) & \text{se } X \text{ è discreta,} \\ \int_{-\infty}^{x} f_X(t)dt & \text{se } X \text{ è continua.} \end{cases}$$

♦

Esempio 2. Vi è stato detto che potete fare un certo test tre volte, e il vostro voto finale è il masso dei tre risultati. Dunque, $X = \max(X_1, X_2, X_3)$, dove X_1, X_2, X_3 sono i risultati dei tre test e X è il risultato finale. Assumiamo che il voto in ciascun test è un valore tra 1 e 10 con probabilità 1/10 indipendentemente dagli altri test. Qual'è la PMF p_X del voto finale? ◊

Esempio 3. Jane va in banca per fare un ritiro ed è equamente probabile di trovare 0 o 1 cliente davanti a lei. Il tempo di servizio del cliente davanti, se è presente, è distributio esponenzialmente con parametro λ. Qual'è la CDF del tempo di attesa di Jane? $\Diamond$

Definizione 7 (Variabile aleatoria normale)**.** *Una variabile aleatoria continua X è detta* normale oppure Gaussiana (normal or Gaussian random variable, *in inglese) se la sua PDF è:*

$$f_X(x) = \frac{1}{\sqrt{2\pi} \cdot \sigma} \cdot e^{-\frac{(x-\mu)^2}{2\sigma^2}}.$$

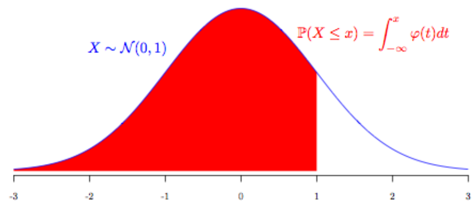
Diciamo che X ha la distribuzione normale o Gaussiana (normal or Gaussian distribution, *in inglese) con parametri μ e $\sigma^2 > 0$, in simboli $X \sim \mathcal{N}(\mu, \sigma^2)$, $E[X] = \mu$ e $var(X) = \sigma^2$.* ◆

Definizione 8 (Variabile aleatoria normale standard)**.** *Una variabile normale Y con media 0 e varianza 1, cioè $Y \sim \mathcal{N}(0,1)$, è detta* variabile aleatoria normale standard (standard normal random variable, *in inglese). La sua CDF denotata con Φ è:*

$$\Phi(y) = P(\{Y \leq y\}) = P(\{Y < y\}) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{y} e^{-\frac{t^2}{2}} dt.$$

La CDF è salvata in una tabella che ci aiuta nel calcolo; si guardi la Figura 1. ◆

Esempio 4. Sia $Y \sim \mathcal{N}(0,1)$, calcolare $\Phi(-0.5)$. ◇



	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.5000	0.5040	0.5080	0.5120	0.5160	0.5199	0.5239	0.5279	0.5319	0.5359
0.1	0.5398	0.5438	0.5478	0.5517	0.5557	0.5596	0.5636	0.5675	0.5714	0.5753
0.2	0.5793	0.5832	0.5871	0.5910	0.5948	0.5987	0.6026	0.6064	0.6103	0.6141
0.3	0.6179	0.6217	0.6255	0.6293	0.6331	0.6368	0.6406	0.6443	0.6480	0.6517
0.4	0.6554	0.6591	0.6628	0.6664	0.6700	0.6736	0.6772	0.6808	0.6844	0.6879
0.5	0.6915	0.6950	0.6985	0.7019	0.7054	0.7088	0.7123	0.7157	0.7190	0.7224
0.6	0.7257	0.7291	0.7324	0.7357	0.7389	0.7422	0.7454	0.7486	0.7517	0.7549
0.7	0.7580	0.7611	0.7642	0.7673	0.7704	0.7734	0.7764	0.7794	0.7823	0.7852
0.8	0.7881	0.7910	0.7939	0.7967	0.7995	0.8023	0.8051	0.8078	0.8106	0.8133
0.9	0.8159	0.8186	0.8212	0.8238	0.8264	0.8289	0.8315	0.8340	0.8365	0.8389
1.0	0.8413	0.8438	0.8461	0.8485	0.8508	0.8531	0.8554	0.8577	0.8599	0.8621
1.1	0.8643	0.8665	0.8686	0.8708	0.8729	0.8749	0.8770	0.8790	0.8810	0.8830
1.2	0.8849	0.8869	0.8888	0.8907	0.8925	0.8944	0.8962	0.8980	0.8997	0.9015
1.3	0.9032	0.9049	0.9066	0.9082	0.9099	0.9115	0.9131	0.9147	0.9162	0.9177
1.4	0.9192	0.9207	0.9222	0.9236	0.9251	0.9265	0.9279	0.9292	0.9306	0.9319
1.5	0.9332	0.9345	0.9357	0.9370	0.9382	0.9394	0.9406	0.9418	0.9429	0.9441
1.6	0.9452	0.9463	0.9474	0.9484	0.9495	0.9505	0.9515	0.9525	0.9535	0.9545
1.7	0.9554	0.9564	0.9573	0.9582	0.9591	0.9599	0.9608	0.9616	0.9625	0.9633
1.8	0.9641	0.9649	0.9656	0.9664	0.9671	0.9678	0.9686	0.9693	0.9699	0.9706
1.9	0.9713	0.9719	0.9726	0.9732	0.9738	0.9744	0.9750	0.9756	0.9761	0.9767
2.0	0.9772	0.9778	0.9783	0.9788	0.9793	0.9798	0.9803	0.9808	0.9812	0.9817
2.1	0.9821	0.9826	0.9830	0.9834	0.9838	0.9842	0.9846	0.9850	0.9854	0.9857
2.2	0.9861	0.9864	0.9868	0.9871	0.9875	0.9878	0.9881	0.9884	0.9887	0.9890
2.3	0.9893	0.9896	0.9898	0.9901	0.9904	0.9906	0.9909	0.9911	0.9913	0.9916
2.4	0.9918	0.9920	0.9922	0.9925	0.9927	0.9929	0.9931	0.9932	0.9934	0.9936
2.5	0.9938	0.9940	0.9941	0.9943	0.9945	0.9946	0.9948	0.9949	0.9951	0.9952
2.6	0.9953	0.9955	0.9956	0.9957	0.9959	0.9960	0.9961	0.9962	0.9963	0.9964
2.7	0.9965	0.9966	0.9967	0.9968	0.9969	0.9970	0.9971	0.9972	0.9973	0.9974
2.8	0.9974	0.9975	0.9976	0.9977	0.9977	0.9978	0.9979	0.9979	0.9980	0.9981
2.9	0.9981	0.9982	0.9982	0.9983	0.9984	0.9984	0.9985	0.9985	0.9986	0.9986
3.0	0.9987	0.9987	0.9987	0.9988	0.9988	0.9989	0.9989	0.9989	0.9990	0.9990

Figure 1: Tabella per la variabile normale standard.

Sia $X \sim \mathcal{N}(\mu, \sigma^2)$. Possiamo *standardizzare* X definiamo una nuova variabile Y come:

$$Y = \frac{X - \mu}{\sigma}.$$

Siccome Y è una funzione lineare di X, Y è normale. In più, abbiamo che:

$$E[Y] = \frac{E[X] - \mu}{\sigma} = 0, \quad var(X) = \frac{var(X)}{\sigma^2} = 1.$$

Esempio 5. La caduta annuale di neve in una certa area geografica è modellata come una variabile aleatoria normale con media $\mu = 60$ centimetri e deviazione standard $\sigma = 20$. Qual'è la probabilità che la caduta di quest'anno abbia almeno 80 centimetri? $\Diamond$

Da questo esempio otteniamo una procedura per calcolare la CDF di una variabile aleatoria normale. Per una variabile aleatoria normale X con media μ e varianza σ^2, dobbiamo:

1. Standardizzare X per ottenere Y, cioè:
$$Y = \frac{X - \mu}{\sigma}.$$
2. Leggere il valore della CDF dalla tabella per la variabile normale standard:
$$P(\{X \leq x\}) = P(\{\frac{X-\mu}{\sigma} \leq \frac{x-\mu}{\sigma}\}) = P(\{Y \leq \frac{x-\mu}{\sigma}\} = \Phi(\frac{x-\mu}{\sigma}).$$

Definizione 9 (Funzione di densità di probabilità congiunta)**.** *La* funzione di densità di probabilità congiunta (joint PDF*, in inglese), o semplicmente* PDF *congiunta, denotata con $f_{X,Y}$, per due variabile aleatorie continue X e Y è definita come:*
$$P(\{(X,Y) \in B\}) = \int\int_{(x,y)\in B} f_{X,Y}(x,y)dxdy,$$
per ciascun $B \subseteq \mathbb{R} \times \mathbb{R}$. Dunque, la PDF marginale *di Y è $P(\{Y = y\}) = \int_{-\infty}^{\infty} f_{X,Y}(x,y)dx$.* ⧫

In particolare, se B è un rettangolo del tipo $B = \{(x,y) \mid a \leq x \leq b, c \leq y \leq d\}$, abbiamo che:
$$P(\{a \leq X \leq b, c \leq Y \leq d\}) = \int_a^b \int_c^d f_{X,Y}(x,y)dxdy.$$

Siano X e Y variabili aleatorie continue congiunte con PDF congiunta $f_{X,Y}$, allora abbiamo che:

- La PDF congiunta è usata per calcolare probabilità:
$$P(\{(X,Y) \in B\}) = \int\int_{(x,y)\in B} f_{X,Y}(x,y)dxdy.$$
- La PDF marginale di X e Y si possono ottenere come:
$$f_X(x) = \int_{-\infty}^{\infty} f_{X,Y}(x,y)dy, \quad f_Y(y) = \int_{-\infty}^{\infty} f_{X,Y}(x,y)dx.$$

- La CDF congiunta è definita come $F_{X,Y}(x,y) = P(\{X \leq x, Y \leq y\})$, e determina la PDF congiunta tramite la formula:

$$f_{X,Y}(x,y) = \frac{\partial^2 F_{X,Y}}{\partial x \partial y}(x,y),$$

  per ciascun (x,y) per cui la PDF congiunta è continua.

- Una funzione $g(X,Y)$ di X e Y definisce una nuova variabile aleatoria continua, e:

$$E[g(X,Y)] = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} g(x,y) \cdot f_{X,Y}(x,y) dx dy.$$

  Se g è lineare della forma $aX + bY + c$, abbiamo che:

$$E[aX + bY + c] = aE[X] + bE[y] + c.$$

Esempio 6. Sia X una variabile aleatoria con PDF:

$$f_X(x) = \begin{cases} \frac{x}{4} & \text{se } 1 \leq x \leq 3, \\ 0 & \text{altrimenti,} \end{cases}$$

e sia A l'evento $\{X \geq 2\}$.

i) Trovare $E[X], P(A), f_{X|A}(x)$ e $E[X \mid A]$.

ii) Sia $Y = X^2$. Trovare $E[Y]$ e $var(Y)$.

$\Diamond$