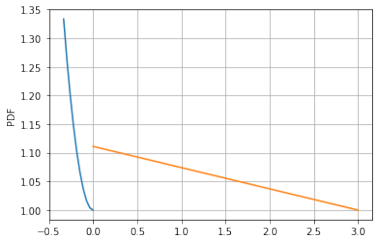
1 I parte 31/01/2022

Es. 1.1 . Sia $f(x)$ la densità di probabilità di una variabile casuale continua x definita come segue:

$$f(x) = \begin{cases} 1 + A\,x^2 & \text{per } x \in \left[-\frac{1}{A}, 0\right] \\ k\left(1 - \dfrac{1}{A}x\right) & \text{per } x \in]0, A] \\ 0 & \text{altrove} \end{cases}$$

dove A è parametro e k è parametro reale da determinarsi. Si definiscano e si stimino a) la speranza matematica, b) la deviazione standard, c) la funzione di distribuzione (CDF), d) la probabilità che $x \in [0, A/2]$. Motivare le risposte.

R: 1.1. Anzitutto procediamo ad un tentativo di graficare la funzione, che ci puo' aiutare:



La condizione affinché $f(x)$ sia PDF e' che sia positiva sul dominio e l'integrale sul dominio sia pari a 1. La prima condizione é verificata se $k \geq 0$. Per la seconda, l'integrale per valori minori di 0 e'

$$\int_{-1/A}^{0} f(x)dx = \left[x + a\frac{x^3}{3}\right]_{-1/A}^{0} = \frac{3A+1}{3A^2}$$

mentre per la parte a destra basta pensare che si tratta di un triangolo quindi l'integrale (l'area) é pari a $Ak/2$ e quindi la condizione si puó scrivere:

$$\frac{3A+1}{3A^2} + \frac{Ak}{2} = 1$$

Il valore di k é positivo solo per $A > 1$. I valori di k sono riportati sotto. a) Per calcolare il valore di aspettazione bisogna svolgere l'integrale:

$$\mu = \mathrm{E}[X] = \int_{-1/A}^{A} xf(x)dx = \int_{-1/A}^{0}(x + Ax^3)dx + \int_0^A k\Big(x - \frac{x^2}{A}\Big)dx = -\frac{1}{2A^2} - \frac{1}{4A^3} + \frac{kA^2}{6}$$

I risultati sono riportati sotto. b) Per calcolare la deviazione standard bisogna calcolare la varianza

$$\mathrm{VAR}[X] = \mathrm{E}[X^2] - \mu^2$$

$$\mathrm{E}[X^2] = \int_{-1/A}^{A} x^2 f(x)dx = \int_{-1/A}^{0}(x^2 + Ax^4)dx + \int_0^A k\Big(x^2 - \frac{x^3}{A}\Big)dx$$

e poi la sua radice. Vedi sotto per i risultati. c) La funzione di distribuzione si calcola con

$$F(x) = \int_{-1/A}^{x} f(x')dx' = \begin{cases} \frac{Ax^3}{3} + x + \frac{3A+1}{3A^2} & \text{per } x \in \left[-\frac{1}{A}, 0\right] \\ F(0) - \frac{kx^2}{2A} + kx & \text{per } x \in [0, A] \end{cases}$$

d). Per la probabilità si poteva fare il conto esplicito

$$P(x \in [0, A/2]) = \int_0^{A/2} f(x)dx$$

oppure usare la $F(x)$ appena calcolata:

$$P(x \in [0, A/2]) = F[A/2] - F[0] = \frac{3}{8}AK$$

oppure pensare ancora alla forma triangolare della PDF per valori positivi e geometricamente trovare:

$$P(x \in [0, A/2]) = \frac{3}{4} \cdot \frac{AK}{2}$$

Raccogliamo i risultati per i diversi A:

A	2	3	4	5
k	0.417	0.420	0.365	0.315
μ	0.121	0.565	0.937	1.289
σ	0.563	0.800	1.036	1.272
P	0.312	0.472	0.547	0.590

Es. 1.2 . Sia

$$W(x, y, z) = Aye^x + yz$$

una grandezza dipendente da x, y e z grandezze fisiche definite in $\mathbb{R}$, A sia parametro dato. Si effettuano 20 misure ripetute di x e z e 15 misure ripetute di y ottenendo come valori medi

$$\overline{x} = 1.00 \pm 0.01; \qquad \overline{y} = 2.00 \pm 0.02; \qquad \overline{z} = 3.00 \pm 0.03$$

con x, y, z covarianti tali che $\mathrm{COV}[X, Y] = 0.01$ e $\mathrm{COV}[X, Z] = 0.001$. Si riporti: *a*) la miglior stima di W a partire dalle misure effettuate; *b*) l'incertezza casuale attesa per W in un generico punto x_0, y_0, z_0 entro la relativa deviazione standard dalla miglior stima di x, y e z sia con che senza i termini di covarianza, discutendone il risultato; *c*) l'incertezza da associare a $W(1.01, 2.00, 3.00)$ nel caso in cui x e z siano adimensionali, y abbia le unità di misura di m/s

R: 1.2. *a*) La miglior stima di W é fatta a partire dalle medie delle misure, e quindi sul punto $\overline{P} = (\overline{x}, \overline{y}, \overline{z})$. Si veda sulla tabella per i rispettivi valori in funzione di A. La formula di propagazione delle incertezze in forma implicita é in generale (usando la notazione $W'_i = \partial W / \partial i(\overline{P})$):

$$s_W^2(P) \simeq (W'_x)^2\, s_x^2 + \left(W'_y\right)^2 s_y^2 + (W'_z)^2\, s_z^2 + 2\left(W'_x W'_y\right) r_{XY} s_x s_y + 2\left(W'_x W'_z\right) r_{XZ} s_x s_z$$

e si tratta di definire correttamente le incertezze di P e le correlazioni. Partiamo comunque con il calcolo delle derivate parziali prime, che sono sempre valutate nella miglior stima dei valori veri di X, Y, Z ovvero $\overline{P}$.

$$W'_x = Aye^x \qquad W'_y = Ae^x + z \quad W'_z = y$$

I valori per i diversi parametri di A sono riportati sotto. *b*) Ora, in questo punto si chiede di propagare l'incertezza attorno ad un punto P_0 non distante dalla media, in questo caso allora la formula diventa:

$$s_W^2(P_0) \simeq (W'_x)^2\, s_{x_0}^2 + \left(W'_y\right)^2 s_{y_0}^2 + (W'_z)^2\, s_{z_0}^2 + 2\left(W'_x W'_y\right) r_{XY} s_{x_0} s_{y_0} + 2\left(W'_x W'_z\right) r_{XZ} s_{x_0} s_{z_0}$$

dove

$$s_{x_0} = s_{\overline{x}}\sqrt{20} = 0.04\ m/s \qquad s_{y_0} = s_{\overline{y}}\sqrt{15} = 0.08\ m/s \qquad s_{z_0} = s_{\overline{z}}\sqrt{20} = 0.13\ m/s$$

$$r_{XY} s_{x_0} s_{y_0} = \mathrm{COV}[X, Y] = 0.01 \qquad r_{XZ} s_{x_0} s_{z_0} = \mathrm{COV}[X, Z] = 0.001$$

Sostituendo con e senza i termini di covarianza si ottengono i valori riportati in tabella. Si vede quindi che se avessimo trascurato la covarianza avremmo commesso una sottostima importante della incertezza della misura di anche il 100%.

Qualche studentessa ha considerato che la domanda a) chiedesse di stimare anche la incertezza. In questo caso, propagando le incertezze attorno alle medie, il valore dei termini con le derivate parziali non sarebbe cambiato, ma come incertezza si prende l'incertezza delle medie:

$$s_W^2(\overline{P}) \simeq (W_x')^2 s_{\overline{x}}^2 + \left(W_y'\right)^2 s_{\overline{y}}^2 + (W_z')^2 s_{\overline{z}}^2 + 2\left(W_x' W_y'\right) r_{XY} s_{\overline{x}} s_{\overline{y}} + 2\left(W_x' W_z'\right) r_{XZ} s_{\overline{x}} s_{\overline{z}}$$

dove

$$s_{\overline{x}} = 0.01\ m/s \qquad s_{\overline{y}} = 0.02\ m/s \qquad s_{\overline{z}} = 0.03\ m/s$$

$$r_{XY} s_{\overline{x}} s_{\overline{y}} = \frac{\mathrm{COV}[X,Y]}{\sqrt{20}\sqrt{15}} = 5.8\,10^{-4} \qquad r_{XZ} s_{\overline{x}} s_{\overline{z}} = \frac{\mathrm{COV}[X,Z]}{\sqrt{20}\sqrt{20}} = 5\,10^{-5}$$

Portando ai risultati in tabella.

c) Il punto chiede la propagazione attorno ad un terzo punto $P = (1.01, 2.00, 3.00)$. Se si intendono le tre misure come singole, si torna al caso precedente (il primo del punto b). Altrimenti, si poteva anche considerare $x = 1.01$ come misura singola e y, z come valori medi. A questo punto

$$s_W^2(P) \simeq (W_x')^2 s_x^2 + \left(W_y'\right)^2 s_{\overline{y}}^2 + (W_z')^2 s_{\overline{z}}^2 + 2\left(W_x' W_y'\right) r_{XY} s_x s_{\overline{y}} + 2\left(W_x' W_z'\right) r_{XZ} s_x s_{\overline{z}}$$

dove

$$s_x = s_{\overline{x}}\sqrt{20} = 0.04\ m/s \qquad s_{\overline{y}} = 0.02\ m/s \qquad s_{\overline{z}} = 0.03\ m/s$$

$$r_{XY} s_x s_{\overline{y}} = \frac{\mathrm{COV}[X,Y]}{\sqrt{15}} = 2.6\,10^{-3} \qquad r_{XZ} s_x s_{\overline{z}} = \frac{\mathrm{COV}[X,Z]}{\sqrt{20}} = 2.2\,10^{-4}$$

Ecco i dati riassunti in tabella per i diversi parametri di A.

A	2	3	4	5
$W_x'(\overline{P})$	10.87	16.31	21.75	27.18
$W_y'(\overline{P})$	8.44	11.15	13.87	16.59
$W_z'(\overline{P})$	2	2	2	2
$s_{W(P_0)}$ [m/s]	1.62	2.25	2.88	3.51
$s_{W(P_0)}^{no-corr}$ [m/s]	0.86	1.16	1.47	1.79
$W(\overline{P})$ [m/s]	16.87	22.31	27.75	33.18
$s_{W(\overline{P})}$ [m/s]	0.39	0.54	0.69	0.84
$s_{W(\overline{P})}^{no-corr}$ [m/s]	0.21	0.28	0.36	0.43
$W(P)$ [m/s]	16.98	22.47	27.96	33.46
$s_{W(P)}$ [m/s]	0.87	1.24	1.61	1.99

Es. 1.3 . Nel gioco Malefiz ciascun giocatore possiede una pedina che può entrare nel circuito del gioco solo se, lanciando un dado a sei facce numerate da uno a sei, osserva proprio il numero 6. Dopo esser entrato nel circuito di gioco, quando arriva il suo turno, il giocatore lancia il dado muove la pedina di un numero di passi pari alla faccia del dado uscita. Per raggiungere la meta deve muovere la pedina di almeno 24 passi in totale. a) calcolare qual è il numero minimo di lanci che gli garantiscono di andare in meta; b) supponendo che il giocatore abbia potuto fare 3 lanci del dado complessivamente, calcolare qual è la probabilità che abbia fatto 7 passi; c) calcolare qual è la probabilità che con il quarto lancio arrivi a compiere 12 passi se al terzo lancio si trova ad aver fatto 7 passi. d) stimare se gli eventi descritti ai punti in b) e c) sono statisticamente indipendenti. Motivare le risposte nell'elaborato.

R: 1.3. a) Il numero minimo di passi per raggiungere metà e' 5: 1 per entrare e 4 per fare 24 passi; b) Per poter avanzare di 7 passi il giocatore deve essere entrato nel circuito. Si tratta quindi di un prodotto logico tra i due eventi aleatori 'E1=il giocatore e' entrato nel gioco' e 'E2=il giocatore ha fatto 7 passi'. Per fare 7 passi servono almeno due lanci, e quindi il giocatore deve entrare necessariamente al primo lancio. La probabilità di E1 e' 1/6 mentre l'evento E2 si puo' realizzare in 6 maniere (1+6,6+1,2+5,5+2,3+4,4+3) su 36, e quindi la probabilità composta é

$$P = \frac{1}{6} \cdot \frac{6}{36} = \frac{1}{36}$$

c) se al terzo lancio ha fatto 7 passi significa che al quarto lancio deve essere uscito la faccia '5', e quindi $P = 1/6$. Se la domanda veniva interpretta come la probabilità composta di entrambi gli eventi allora $P = 1/6 \cdot 1/36$ *d*) Gli eventi sono statisticamente indipendenti.

Es. 1.4 .

Si consideri il seguente campione di misure ripetute della lunghezza L di un filo metallico sottoposto a trazione a forza costante fissata in cui f è la frequenza assoluta delle misure effettuate. Definire e stimare rispettivamente: a) la deviazione standard della singola misura b) la moda e la mediana; c) la miglior stima il valor medio e relativa incertezza; d) il valore atteso di f assumendo che la densità di probabilità gaussiana ben rappresenti nella classe 45.20mm. e) indicare se è necessario effettuare la reiezione di dati qualora si assuma il criterio del 3σ; f) l'intervallo di confidenza centrato sul valor medio e di ampiezza 2σ; g) assumendo la lunghezza a riposo $L_0 = (A.10 + 45) \pm 0.01$ mm, calcolare l'allungamento medio ΔL subito dal filo e la relativa incertezza (A è parametro fornito dal primo esercizio); h) la compatibilità di tale valore con il valore di riferimento (0.24 ± 0.01) mm. Motivare le risposte nell'elaborato.

L (mm)	f
45.00	7
45.10	12
45.20	14
45.30	5
45.40	8

R: 1.4 La numerosità del campione e' $N = 46$; La media aritmetica $\overline{L} = 45.1891$ mm; La deviazione della singola misura $s_L = 0.1303$ mm[1] da cui $s_{\overline{L}} = s_x\sqrt{N} = 0.0192$ mm. La moda, ovvero il valore più ricorrente è $= 45.2$ mm e la mediana é la mnedia tra la $N/2 = 23$esima e la 24esima misura, comunque pari a $= 45.2$ mm. La miglior stima della grandezza e' quindi $L = (45.189 \pm 0.019)$ mm ovvero $L = (45.19 \pm 0.02)$ mm. Per la classe 45.20 il valore aspettato di conteggi é:

$$f^*(45.20) = \frac{Nw}{\sqrt{2\pi}s_L} e^{-\frac{1}{2}\left(\frac{45.20-\overline{L}}{s_L}\right)^2} = 9.81$$

da confrontare con i 5 misurati. $w = 0.1$ e' la larghezza della classe di frequenza. L'intervallo dei 3σ é $[\overline{x} - 2s_L, \overline{x} + 2s_L] = [44.7981, 45.5802]$ mm. Nessuna misura e' da scartare come ci aspettavamo dal fatto che al di fuori di 3σ ci aspettiamo $(1-99.7) \cdot 46 = 0.14$ misure. L'intervallo dei 2σ e' invece $[\overline{x} - 2s_L, \overline{x} + 2s_L]$. Dato ora $L_0 = (A.10 = 45) \pm 0.01$ mm calcoliamo

$$\Delta L = \overline{L} - L_0 \qquad s_{\Delta L} = \sqrt{s_{\overline{L}}^2 + s_0^2}$$

. A quel punto era da calcolare

$$r = \frac{|\Delta L - 0.24|}{\sqrt{s_{\Delta L}^2 + 0.01^2}} \gg 3$$

e quindi con forte incompatibilità.

[1] Alcune hanno considerato una deviazione standard basata su ipotesi uniforme, che non poteva essere usata in questo caso

Es. 1.5 . Sia X una variabile casuale gaussiana centrata in $\mu = 20$ e varianza pari a 4 Si calcoli la probabilità di trovare un valore compreso in $[18; 22]$ e in $[18, 24]$. Assumendo che $y = A(x^2 + x - 1)$ con A il risultato del primo esercizio, calcolare la covarianza $\mathrm{COV}[X, Y]$ per $x = 18$ e $x = 24$.

R: 1.5 Per il calcolo delle probabilità si osserva come nel primo caso si chiede $P[-1\sigma, +1\sigma] = 0.683$ mentre nel secondo si chiede $P[-1\sigma, +2\sigma] \sim 0.819$. Non serviva quindi usare la tabella della funzione degli errori in quanto i primi margini di confidenza della Gaussiana sono noti. Ad ogni modo, attraverso la tabella il conto sarebbe stato, ricordando che $F_{\mathcal{G}}(x) = \frac{1}{2}\Big[1 - \mathtt{erf}\Big(\frac{x-\mu}{\sqrt{2}\sigma}\Big)\Big]$

$$F(22) - F(18) = \frac{1}{2}\Big[\mathtt{erf}\Big(\frac{22-20}{\sqrt{2}2}\Big) - \mathtt{erf}\Big(\frac{18-20}{\sqrt{2}2}\Big)\Big]$$

$$= \frac{1}{2}\Big[\mathtt{erf}\Big(\frac{1}{\sqrt{2}}\Big) - \mathtt{erf}\Big(\frac{-1}{\sqrt{2}}\Big)\Big] = \frac{1}{2}\Big[2\mathtt{erf}\Big(\frac{1}{\sqrt{2}}\Big)\Big] = \mathtt{erf}(0.707) \simeq 0.68$$

e analogamente nell'altro caso:

$$F(24) - F(18) = \frac{1}{2}\Big[\mathtt{erf}\Big(\frac{2}{\sqrt{2}}\Big) - \mathtt{erf}\Big(\frac{-1}{\sqrt{2}}\Big)\Big] = \frac{1}{2}[0.95 + 0.7] \simeq 0.83$$

Si noti anche che si poteva 'traslare la Gaussiana' mettendo a zero la media e non cambiava il conto.

Per il calcolo della covarianza, che non dipende dal punto in cui e' calcolata, si procedeva con:

$$\mathrm{COV}[Y, X] = \mathrm{COV}[AX^2 + AX - A, X] = \mathrm{COV}[AX^2, X] + \mathrm{COV}[AX, X] - \mathrm{COV}[A, X]$$

$$= A\,\mathrm{COV}[X^2, X] + A\,\underbrace{\mathrm{VAR}[X]}_{\sigma^2}$$

Il secondo addendo e' noto, mentre per il primo:

$$\mathrm{COV}[X^2, X] = \mathrm{E}[X^3] - \underbrace{\mathrm{E}[X^2]}_{\mathrm{VAR}[X]+\mathrm{E}[X]^2}\ \underbrace{\mathrm{E}[X]}_{\mu}$$

e quindi:

$$\mathrm{COV}[Y, X] = A\,\mathrm{E}[X^3] - A\mu(\sigma^2 + \mu^2) + A\sigma^2$$

Dove il terzo momento algebrico in realtà (non era richiesto) e' noto algebricamente per la Gaussiana e vale $\mathrm{E}[X^3] = \mu^3 + 3\mu\sigma^2$ che si puo' verificare pensando che se X normale $Z = \frac{X-\mu}{\sigma}$ e' normale standard: $\mu = \mathrm{E}[Z] = 0$, $\sigma^2 = \mathrm{VAR}[Z] = 1$. Inoltre essendo simmetrica vale $\mathrm{E}[Z^3] = 0$. Allora

$$\mathrm{E}[X^3] = \mathrm{E}[(\mu + \sigma Z)^3] = \mathrm{E}[\mu^3 + 3\mu^2\sigma Z + 3\mu\sigma^2 Z^2 + Z^3] = \mu^3 + 3\mu\sigma^2$$