**Come spiegato nel regolamento d'esame,
in questo tema d'esame possono comparire entrambi gli standard del punto decimale e della virgola decimale.
In ogni esercizio in cui nel quesito o nello svolgimento compaiono numeri che in italiano diciamo *con la virgola*, scrivere all'inizio dello svolgimento se è usato lo standard del punto o della virgola decimale.
Ovviamente se nel testo di un quesito c'è qualcuno di quei numeri, lo svolgimento va fatto continuando con lo stesso standard.**

Si consideri bene la nota sul punto decimale del regolamento d'esame

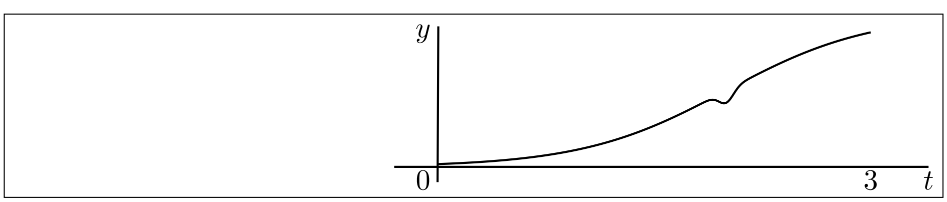
Chi si ritira, consegna **solo** questo foglio: col nome e una grande R.
Gli altri, tengono questo foglio, e consegnano la bella copia

RIQUADRARE ovvero incorniciare I RISULTATI

Legenda
$*$ è richiesto il valore esatto. Può anche essere $+\infty$, $-\infty$, o una frase.
$\approx$ è richiesta una ragionevole approssimazione.
$\%$ è richiesto il valore in percentuale, se serve ragionevolmente approssimato.
(R) è richiesto solo il risultato; negli altri esercizi riportare anche i calcoli.

Esercizio 0. Triplice – quesiti basici – chi non risolve almeno 2 non passa l'esame – per ricevere più di 18 risolvere tutti 3.

ES. 0a$_{\mu 2024}$ (R) $*$ La funzione definita in $[0, 3]$ a valori in $\mathbb{R}$ di cui è visibile il grafico (più tecnicamente: un disegno del grafico) in figura, è crescente o no?



no

(È del tutto evidente dal grafico che non è crescente, c'è una sorta di piccolo "avvallamento". Tecnicamente, non è crescente perché non verifica la definizione di crescenza: a numero minore dovrebbe associare valore minore, in simboli

se $x_1 < x_2 \Rightarrow f(x_1) < f(x_2)$ **allora** f **si dice crescente;**

ma questo non succede sempre: detta $f(x)$ la funzione, esiste una coppia di numeri $x_1 < x_2$ a cui corrispondono valori $f(x_1) > f(x_2)$).

ES. 0b$_{\mu 2024}$ (R) $*$ Qual è la parola qua sotto mancante, al posto dei puntini?

Per una variabile aleatoria continua – nelle solite ipotesi di sufficiente regolarità di cui non ci occupiamo in dettaglio – la probabilità corrisponde all'area del sottografico della funzione ... e precisamente, se essa è $f(t)$,

$$P(a < X \le b) = \int_a^b f(t)\,dt\,.$$

$$\boxed{\text{densità}}$$

(Naturalmente la $f(t)$, di cui si dice nel testo del quesito, è la *(funzione) densità* della variabile aleatoria considerata, e queste sono le basi del Calcolo delle Probabilità moderno: la variabile aleatoria – qua continua – e la sua funzione densità, spesso denotata proprio $f(t)$ come nel quesito).

ES. 0c$_{\mu 2024}$ (R) $*$ Qual è la parola qua sotto mancante, al posto dei puntini?

Un argomento fondamentale della Statistica Inferenziale è quello degli ... puntuali, in particolare non distorti.

$$\boxed{\text{stimatori}}$$

ES. 1$_{\mu 2024}$ $*$ Si scriva con gli usuali simboli matematici l'equazione qua scritta informaticamente (nella fattispecie, su WolframAlpha) e la si risolva:

$$E\text{^}4 + 2 * E\text{^}2 * x + x\text{^}2 = 1$$

SVOLGIMENTO

Si tratta dell'equazione di secondo grado in x

$$\mathrm{e}^4 + 2\,\mathrm{e}^2\,x + x^2 = 1$$

più ordinatamente scritta

$$x^2 + 2\,\mathrm{e}^2\,x + (\mathrm{e}^4 - 1) = 0$$

che si risolve subito con le 2 classicissime formule per Δ e per $x_{1,2}$:

$$\Delta = b^2 - 4\,a\,c =$$

(oppure ancora più semplicemente con la formula per $\frac{\Delta}{4}$, se la si conosce)

$$= (2\,\mathrm{e}^2)^2 - 4\,(\mathrm{e}^4 - 1) =$$

per le proprietà delle potenze

$$= 4\,\mathrm{e}^4 - 4\,\mathrm{e}^4 + 4 =$$

$$= 4$$

$$x_{1,2} = \frac{-b \pm \sqrt{\Delta}}{2\,a} =$$

$$= \frac{-2\,\mathrm{e}^2 \pm \sqrt{4}}{2} =$$

$$= \frac{-2\,\mathrm{e}^2 \pm 2}{2} =$$

$$\boxed{-\mathrm{e}^2 \pm 1}$$

soluzione altrimenti espressa con

$$\boxed{x = -\mathrm{e}^2 - 1 \vee x = -\mathrm{e}^2 + 1}$$

(e in vari altri modi equivalenti).
Nota. Soluzioni confermate da WolframAlpha: Link ->

ES. 2$_{\mu 2024}$ * Qual è la moda di questo dataset di 19 intervalli? (Tratto da doi: 10.1136/bmj.l6460) [0,1) [23,0) [10,11) [1,2) [12,13) [19,20) [9,10) [14,15) [9,10) [10,11) [19,20) [13,14) [14,15) [12,13) [10,11) [22,23) [8,9) [10,11) [14,15) (Senza che il risolutore debba occuparsene, sono 19 di intervalli temporali di un'ora, in cui più frequentemente vengono inviati articoli da 19 stati al British Medical Journal, un cui articolo indaga sugli orari di lavoro che appaiono risultare da questa e analoghe statistiche; in Italia l'ora di picco degli invii è fra le 19 e le 20 e in Spagna fra le 21 e le 22, e quest'ultimo dato è soggetto alla stessa distorsione di quella di dire – come spesso si fa – che gli spagnoli cenano tardissimo: potrà essere in parte vero ma hanno lo stesso fuso orario dell'Italia pur essendo molto più a ovest, col sole che – oltre a sorgere – tramonta molto dopo a quanto fa in Italia, a Madrid a Natale quasi un'ora e mezza più tardi che a Trieste: la Statistica ha le sue sottigliezze).

SVOLGIMENTO

La moda è il valore più frequente di un dataset non necessariamente numerico – e qua si tratta di intervalli – per esempio il cognome più diffuso in uno stato. Si vede che l'intervallo [10,11) ricorre 4 volte e nessun altro ricorre così tante volte, e allora la moda è

$$\boxed{[10, 11)}$$

ES. 3$_{\mu 2024}$ $\ast$ Calcolare c affinché

$$\int_1^{\mathrm{e}} \frac{2\,c}{t}\,\mathrm{d}t = 1$$

(Così, senza che il risolutore debba occuparsene, sarà una densità di probabilità la funzione che vale

$$f(t) := \frac{2\,c}{t} \quad se\ 1 \le t \le \mathrm{e}$$

e 0 altrimenti).

SVOLGIMENTO

La

$$(2\,c) \cdot \frac{1}{x} \qquad (\ast)$$

ha primitive ovvero ha integrale indefinito

$$2\,c\,\ln x + c \qquad (\diamond)$$

(se vogliamo dire, perchè derivando $(\diamond)$ si ottiene $(\ast)$ qualunque sia la costante c) e allora

$$1 =^{EQUAZIONE\ DATA} \int_1^{\mathrm{e}} \frac{2\,c}{t}\,\mathrm{d}t =$$

con l'integrale indefinito sopra detto

$$= [2\,c \ln t]_1^{\mathrm{e}} =$$

$$= 2\,c\,\ln \mathrm{e} - 2\,c\,\ln 1 =$$

$$= 2 \cdot c \cdot 1 - 0$$

e dall'ottenuta equazione $1 = 2\,c$ segue il valore di c:

$\boxed{\frac{1}{2}}$

ES. 4$_{\mu 2024}$ % Qual è la probabilità di ottenere almeno 1 volta il 5 in 2 lanci di dado? (Se si preferisce: la probabilità di prendere almeno 1 volta il topo n. 5 fra 6 topi, prendendone prima 1 a caso, poi rimettendolo in gabbia, e prendendone poi 1 a caso).

SVOLGIMENTO

Verrà usato lo standard della virgola decimale. (Ma si potrebbe usare lo standard del punto decimale, a scelta).

È

$$P(\text{almeno 1 successo}) =$$

con l'evento complementare

$$= 1 - P(\text{2 fallimenti}) =$$

$$= 1 - P(\text{fallimento al primo tentativo e fallimento al secondo tentativo}) =$$

per l'indipendenza

$$= 1 - P(\text{fallimento al primo tentativo}) \cdot P(\text{fallimento al secondo tentativo}) =$$

con gli eventi complementari

$$= 1-(1 - P(\text{successo al primo tentativo}))\cdot(1 - P(\text{successo al secondo tentativo})) =$$

$$= 1 - \left(1 - \frac{1}{6}\right) \cdot \left(1 - \frac{1}{6}\right) =$$

$$= 1 - \frac{5}{6} \cdot \frac{5}{6} =$$

$$= 1 - \frac{25}{36} =$$

$$= \frac{36 - 25}{36} =$$

$$= \frac{11}{36} \approx$$

$$\approx 0{,}3056 =$$

$$\boxed{\approx 30{,}56\%}$$

e con minore precisione

$$\approx 30{,}6\%$$

(Ovvero circa il 31%).

ES. 5$_{\mu 2024}$ $*$ Premettiamo che la soluzione richiesta è un intervallo espresso "con il trattino", come spesso si fa nelle Scienze Biomediche, come quando si scrive "70 000 - 80 000 contagiati". (E plausibilmente i valori estremi sono approssimazioni e non ha molto senso chiedersi se appartengano o no all'intervallo). Usando la consueta approssimazione con la radice quadrata, se supponiamo che lanceremo 40 000 volte una moneta, in quale intervallo dovrà stare il numero di teste affinchè non rifiutiamo al consueto livello della significatività statistica l'ipotesi della regolarità della moneta?

SVOLGIMENTO

Notoriamente con probabilità circa del 95% e più precisamente del 95.4% una moneta regolare lanciata n volte (con n sufficientemente grande) darà un numero di teste X_n compreso fra $n/2$ meno $\sqrt{n}$, e $n/2$ più $\sqrt{n}$, in simboli

$$P\left(\frac{n}{2} - \sqrt{n} \leq X_n \leq \frac{n}{2} + \sqrt{n}\right) \approx 95\% \qquad (*)$$

Abbiamo

$$n = 40\,000$$
$$\frac{n}{2} = 20\,000$$
$$\sqrt{n} = 200$$

e allora l'intervallo cercato, che ci farà respingere l'ipotesi della regolarità se il numero di teste non vi cadrà dentro, è approssimativamente

$$\boxed{19\,800 \text{ - } 20\,200}$$

Nota. Più tecnicamente della $(*)$, vale

$$P\left(\left|X_n - \frac{n}{2}\right| > \sqrt{n}\right) \approx 4.6$$

ovvero con l'evento complementare

$$P\left(\left|X_n - \frac{n}{2}\right| \leq \sqrt{n}\right) \approx 95.4$$

ovvero equivalentemente

$$P\left(\frac{n}{2} - \sqrt{n} \le X_n \le \frac{n}{2} + \sqrt{n}\right) \approx 95.4$$

e poi nella $(*)$ abbiamo approssimato 95.4% con 95%, e invece il 95% è valore esatto se si sostituisce $\sqrt{n}$ con una certa complicata funzione di n, di cui $\sqrt{n}$ è "buona" approssimazione; per questo il testo del quesito parla della "consueta approssimazione con la radice quadrata": usare la radice quadrata di n dà approssimazioni dei valori degli estremi, il cui calcolo esatto per il valore esatto 95% è molto complesso, e in pratica si farebbe col computer.

Due tipici esempi facili da tenere a mente:

– con un milione di lanci respingiamo l'ipotesi della regolarità della moneta se il numero di teste non sta nell'intervallo approssimativamente espresso da 499 000 - 501 000 (per non qua parlare del triste fenomeno dell'aborto selettivo e dell'infanticidio delle neonate, questione comunque molto delicata dal punto di vista statistico: questo è solo un accenno basico della statistica implicata, che detto ipersemplificatamente induce sospetti se, su grandissimi numeri, la percentuale delle neonate femmine non è *vicinissima* al 50%, e si noti che qua il test è bilatero, ancora nell'ottica dell'ipersemplificazione)

– prendendo a caso 20 Studenti del primo anno di CTF (o Farmacia) di Trieste, presenti un giorno in aula, come campione degli Studenti del primo anno di CTF (o rispettivamente Farmacia) delle università italiane (con tutte le problematiche e limitazioni del caso), respingiamo (circa al 95%) l'ipotesi che per tali studenti sia equiprobalile essere femmina o maschio se fra quei 20 il numero delle femmine è ≤ 5 o ≥ 15, come risulta da

$$\frac{20}{2} - \sqrt{20} < X_{20} < \frac{20}{2} + \sqrt{20}$$

ed equivalentemente, e non scrivendo infiniti decimali,

$$5.52... \le X_{20} \le 14.47...$$

ed equivalentemente, con numeri necessariamente interi,

$$6 \le X_n \le 14$$

ovvero, per l'evento complementare che ci serve per respingere l'ipotesi nulla dell'equiprobabilità,

$$X_n < 6 \vee X_n > 14$$

ovvero equivalentemente, di nuovo con numeri necessariamente interi,

$$X_n \le 5 \vee X_n \ge 15$$

proprio come sopra anticipato. (E di nuovo si noti che qua il test è bilatero, per massima semplicità, mentre si può dimostrare che per il test unilatero per l'ipotesi alternativa di più femmine la condizione è $X_n \ge 14$).