Esercizio 3.1.4 pag. 12 dell'Eserciziario [Tema d'esame 1.2 del 29/06/04] Sia $X_1, \dots, X_n$ un campione casuale estratto dalla funzione di densità discreta

$$p(x, \theta) = \begin{cases} \frac{\mathrm{e}^{-\theta^2}\theta^{2x}}{x!} & x = 0, 1, \dots \\ 0 & \textit{altrove} \end{cases} \qquad \theta \in \mathbb{R} \setminus \{0\}$$

1. Determinate uno stimatore della caratteristica $\kappa(\theta) = \theta^2$ usando il metodo dei momenti.

Sia $\widehat{\kappa}$ lo stimatore individuato al punto 1.

2. Verificate se $\widehat{\kappa}$ è stimatore non distorto per $\kappa(\theta)$. Giustificate rigorosamente la risposta.
3. Calcolate la funzione di verosimiglianza del campione: $L_\theta(x_1, \dots, x_n)$.
4. Verificate se la varianza di $\widehat{\kappa}$ raggiunge il confine di Cramer-Rao.

Esercizio 4.2.7 pag. 22 dell'Eserciziario [Tema d'esame 3.2 del 12/09/06 AA 05/06]
È noto che il peso di una donna italiana di 25 anni è una variabile aleatoria gaussiana X di media 62 Kg e varianza 16 Kg^2. Il programma dimagrante `WW` consente di ridurre il peso di una quantità $Y \sim \mathcal{N}(\mu, 20)$, con μ parametro incognito. In altre parole, alla fine della dieta il peso di una donna che ha seguito il programma è $Z = X - Y$ e assumiamo che X, Y siano indipendenti.

Nel depliant pubblicitario del programma dimagrante `WW` si sostiene che chi segue scrupolosamente il programma perde mediamente 5 Kg, e noi, per stabilire se il messaggio è ingannevole o no, abbiamo deciso di verificare il programma `WW` su 64 donne; $Z_1, \dots, Z_{64}$ è il campione casuale ottenuto.

1. Qual è la distribuzione di Z?

2. Costruite un test su μ tale che sia al più pari al 4% la probabilità di commettere l'errore di prima specie di ritenere ingannevole il messaggio quando in realtà seguendo il programma `WW` si perdono effettivamente (almeno) 5 Kg.

3. Fornite analiticamente e rappresentate graficamente la funzione di potenza del test di livello 4% costruito al punto 2.

4. Se il campione dei dati ha fornito $\sum_{j=1}^{64} Z_j = 3712.0$, qual è il p-value del test?

Esercizio 5.1.11 pag. 34 [Tema d'esame 1.3 del 01/07/08, AA 07/08] Una delle densità tradizionalmente usate per modellare la distribuzione dei redditi è la densità di Pareto di parametri α, β data da

$$f(x; \alpha, \beta) = \begin{cases} \frac{\alpha\beta^\alpha}{x^{\alpha+1}} & \text{se } x > \beta \\ 0 & \text{altrove} \end{cases}, \qquad \alpha > 2, \beta > 0.$$

Nel modello di Pareto di parametri α, β con $\alpha > 2$ e $\beta > 0$, la media e il momento secondo sono

$$\mathrm{E}(X) = \frac{\alpha\beta}{(\alpha - 1)}, \qquad \mathrm{E}(X^2) = \frac{\alpha\beta^2}{\alpha - 2}.$$

Per verificare se un modello di Pareto ben si adatta ai redditi dei lavoratori dipendenti del quartiere `OQ` di Milano, abbiamo campionato i redditi di 500 lavoratori dipendenti che abitano nel quartiere e i dati ottenuti (espressi in migliaia di euro) sono sintetizzati nella seguente tabella:

A_k	$(0, 11.6]$	$(11.6, 13.4]$	$(13.4, 15.3]$	$(15.3, 18.3]$	$(18.3, 22.2]$	$(22.2, \infty)$
N_k	175	105	95	65	60	0

,

dove A_k indica la classe di reddito annuo netto, espresso in migliaia di euro, e N_k il numero di dipendenti con reddito appartenente alla classe A_k. Inoltre, il reddito medio campionario (espresso in migliaia di euro) e il momento secondo campionario dei dati raggruppati della precedente tabella valgono rispettivamente 12.00 e 169.64.

1. Determinate gli stimatori dei momenti di α e β sulla base dei valori di reddito medio campionario e momento secondo campionario dei dati raggruppati che vi sono stati forniti.

2. Determinate $P(x_1 < X \leq x_2)$ quando X ha densità di Pareto $f(x; \alpha, \beta)$.

3. Valutate con un opportuno test la bontà di adattamento del modello di Pareto ai dati sui redditi dei lavoratori dipendenti del quartiere `OQ` di Milano. (*Se non siete riusciti a risolvere il punto 1., scegliete voi dei valori per i parametri α e β ed eseguite un opportuno test.*)

Esercizio 5.2.3 pag. 36 [Tema d'esame 4.3 del 29/09/06, AA 05/06] Un'azienda possiede due stabilimenti A e B per la produzione di transistor. Per verificare che i prodotti dei due stabilimenti siano equivalenti, si esaminano due campioni indipendenti di 8 e 10 transistor. Denotiamo con $X_1, \ldots, X_8$ i tempi di vita degli 8 transistor campionati dallo stabilimento A e con $Y_1, \ldots, Y_{10}$ i tempi di vita dei 10 provenienti dallo stabilimento B.

La classifica dei 18 transitor esaminati, ordinati dal meno al più durevole è la seguente

Rank	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
	b	a	b	b	b	b	a	b	b	b	a	b	a	a	b	a	a	a

(a=stabilimento A, b=stabilimento B).

1. Impostate un test di verifica dell'ipotesi nulla H_0 :"*i prodotti dei due stabilimenti sono equivalenti*" contro l'alternativa H_1 :"*Il transistor dello stabilimento A dura più di quello dello stabilimento B*", di livello $\alpha = 5\%$. Sulla base dei dati cosa decidete?

In realtà, noi possiamo supporre che i transistor A e B abbiano tempi di vita esponenziali di medie θ_A, θ_B, entrambe incognite. Inoltre, conosciamo anche i valori delle medie campionarie: $\overline{X}_8 = 25.1$ e $\overline{Y}_{10} = 22.0$.

2. Determinate le distribuzioni di $\dfrac{16\overline{X}_8}{\theta_A}, \dfrac{20\overline{Y}_{10}}{\theta_B}$ e del rapporto $\overline{X}_8 / \overline{Y}_{10}$ nell'ipotesi $\theta_A = \theta_B$.

3. Traducete le ipotesi H_0 :"*i prodotti dei due stabilimenti sono equivalenti*" contro H_1 :"*Il transistor dello stabilimento A dura più di quello dello stabilimento B*" in ipotesi sui parametri θ_A, θ_B. Quindi, costruite un nuovo test (questa volta parametrico) di livello $\alpha = 5\%$ basato sulla statistica test $\overline{X}_8 / \overline{Y}_{10}$. Quale decisione prendete se $\overline{X}_8 = 25.1$ e $\overline{Y}_{10} = 22.0$?

Esercizio 5.3.2 pag. 40 dell'Eserciziario [Tema d'esame 2.4 del 15/07/04, AA 03/04] I dati disponibili sul sito del Ministero dell'Istruzione, dell'Università e della Ricerca rivelano che nell'anno solare 2002, i laureati e diplomati presso le facoltà di ingegneria del Politecnico di Milano sono stati 3502 di cui 2967 uomini. Fra le donne, 176 erano fuori corso da un anno, 72 da due anni e 61 da tre anni. Invece, fra gli uomini, i fuori corso da un anno erano 733, quelli da due anni 531 e quelli da tre anni 279. Tutti gli altri erano fuori corso da quattro anni o più; nel 2002 non ci sono stati laureati in corso. Verificate sulla base di questi dati se uomini e donne impiegano (più o meno) lo stesso tempo per laurearsi in ingegneria.

Tema d'esame 5.3 del 28/02/2011, AA 09/10 Un gruppo costituito da 5 programmatori deve consegnare un software. Per ciascun programmatore viene registrato il numero di righe di programma scritte giornalmente ottenendo il seguente campione di osservazioni:

$$x_1 = 50.97,\ x_2 = 71.58,\ x_3 = 340.29,\ x_4 = 112.06,\ x_5 = 76.44\ ;$$

l'i-esimo dato x_i riportato è la media aritmetica fatta in una prova di due giorni.

1. Stabilite con un opportuno test di livello 5% se la variabile aleatoria X del numero giornaliero di righe scritte abbia distri buzione gaussiana (normale).

Viene poi contattato un secondo gruppo di 6 programmatori e anche di ciascuno di questi viene registrato il numero giornaliero di righe di programma scritte ottenendo

$$y_1 = 65.30,\ y_2 = 187.48,\ y_3 = 111.84,\ y_4 = 2.60,\ y_5 = 23.67,\ y_6 = 42.60\ .$$

2. Sulla base dei dati forniti, il numero giornaliero di righe scritte da un programmatore del secondo gruppo segue lo ste sso modello del primo gruppo? Per rispondere costruite un opportuno test di significatività 5%. Nella scelta del test tenete conto di quanto risposto al punto 1.