1 Esercizi per l'esame finale

1.1 Stima puntuale

1 Sia $(X_1, ..., X_n)$ un campione casuale estratto da una distribuzione $U[0, \theta], \quad \theta > 0$.

(a) Scrivere la funzione di verosimiglianza del campione;

(b) trovare uno stimatore per θ con il metodo dei momenti e verificare che è non distorto e consistente;

(c) trovare lo stimatore di massima verosimiglianza θ e verificare se è non distorto (in caso contrario correggerlo affinché lo diventi).

2 In una fabbrica ci sono n macchinari che lavorano in parallelo producendo ciascuno la stessa quantità di prodotto. Ciascun macchinario è soggetto a guasti indipendentemente dagli altri e, indicato con T_i il tempo di vita dell'i-simo macchinario, risulta $T_i \sim Exp(\theta)$. Indichiamo inoltre con Y il tempo in cui la fabbrica lavora con tutti i macchinari in funzione.

(a) Calcolare la distribuzione di Y;

(b) utilizzando la statistica Y, trovare uno stimatore non distorto della durata media di vita di un macchinario;

(c) trovare uno stimatore di max verosimiglianza della durata media di vita di un macchinario e confrontarlo con quello trovato in b). Qual è preferibile?

3 Un casinò per invogliare la gente a giocare, regala il primo milione per partecipare ad un gioco d'azzardo in cui si vince la somma puntata con probabilità $\theta, \quad 0 < \theta < 1$. Ipotizziamo che ad ogni puntata la probabilità di vittoria sia sempre la stessa e che sia indipendente dai risultati precedenti. Un giocatore decide di partecipare scegliendo la seguente strategia:

prende il milione offerto dal casinò e gioca, finché vince, puntando ogni volta un milione. Smette di giocare dopo la prima perdita.

Sia X la v.a. che indica la vincita totalizzata dal giocatore; calcolare:

(a) la distribuzione di X ed il suo valore atteso.

Supponiamo che un altro giocatore per decidersi a giocare, osserva n partecipanti al gioco che seguono tutti la strategia descritta e prende nota delle loro vincite. Calcolare:

(b) lo stimatore di max verosimoglianza $\widehat{\theta}$ di θ e lo stimatore θ^* ottenuto con il metodo dei momenti;

(c) lo stimatore di max verosimiglianza T della vincita attesa, e verificare che è corretto e consistente;

(d) controllare se T raggiunge il limite inferiore di Cramer-Rao.

4 Un contatore emette segnali seguendo una distribuzione di Poisson di parametro θ incognito. Si vuole stimare la probabilità che in un giorno il numero X dei segnali emessi sia minore di un certo k intero positivo. A questo proposito si osservano

n contatori identici. Trovare lo stimatore di max verosimiglianza del parametro di interesse.

5 Sia $(X_1, ..., X_n)$ un campione casuale estratto dalla densità

$$f(x, \theta) = \begin{cases} \dfrac{1}{2} & x = -1; \\ \dfrac{1-\theta}{2} & x = 0; \\ \dfrac{\theta}{2} & x = 1. \end{cases}$$

Si determini lo stimatore di max verosimiglianza per il parametro θ.

6 Sia $f(x, \theta) = \frac{1}{\sqrt{\theta}} e^{-\frac{x}{\sqrt{\theta}}} I_{[0,\infty)}(x), \;\; \theta > 0$. Dato un campione casuale $X_1, ..., X_n$ estratto da $f(x, \theta)$ calcolare:

(a) lo stimatore di max verosimiglianza del parametro;

(b) si verifichi se questo stimatore corretto.

7 Dovendo controllare se un dado è truccato in modo da favorire la realizzazione dell'evento "esce un numero pari ", si considera il modello statistico

$$f(x, \theta) = \begin{cases} \dfrac{1-\theta}{6} & x \in \{2, 4, 6\}; \\ \dfrac{1+\theta}{6} & x \in \{1, 3, 5\}. \end{cases}$$

Dato un campione casuale $X_1, ..., X_n$ estratto da $f(x, \theta)$, determinare:

(a) lo stimatore di max verosimiglianza del parametro;

(b) si verifichi se questo stimatore raggiunge il limite inferiore di Cramer-Rao.

8 Sospettando che un fiume sia inquinato, si estraggono dei campioni per verificare se la presenza di microrganismi sia irregolare. Si sa che il numero X dei microrganismi presenti nell'acqua è una v.a. di Poisson di parametro θ incognito. Calcolare:

(a) lo stimatore di max verosimiglianza di θ;

(b) si verifichi se questo stimatore è efficiente.

(c) Si effettuano 30 rilevamenti ottenendo i seguenti valori:

12	15	16	17	19	11	10	15	19	11
20	10	9	16	18	9	12	21	13	16
16	15	17	13	21	13	11	12	18	15

Stimare θ sulla base dei valori osservati;

(d) trovare un intervallo di confidenza per θ a livello 95%.

9 Sia $f(x, \theta) = 2\theta x e^{-\theta x^2} I_{[0,\infty)}(x), \quad \theta > 0.$ Dato un campione casuale $X_1, ..., X_n$ estratto da $f(x, \theta)$ calcolare:

(a) lo stimatore di max verosimiglianza del parametro;

(b) si verifichi se questo stimatore efficiente.

10 Sia $(X_1, ..., X_n)$ un campione casuale estratto dalla densità geometrica $f(x, \theta) = \theta(1-\theta)^x \quad x \in \mathbb{N}, \quad \theta \in (0,1)$. Calcolare:

(a) lo stimatore θ^* di θ ottenuto con il metodo dei momenti;

(b) lo stimatore di max verosimoglianza $\widehat{\theta}$ di θ;

(c) verificare se $\widehat{\theta}$ corretto ed efficiente.

1.2 Stima intervallare

1 In un laboratorio di ricerca, a 12 cavie viene somministrata una dieta specifica per una settimana ottenendo i seguenti aumenti di peso in grammi:

$$15, 11, 16, 13, 12, 20, 17, 18, 16, 17, 14, 15$$

Si calcoli un intervallo di confidenza al 90% per l'aumento di peso, ipotizzando che

(a) la v.a. X che indica l'aumento di peso sia una v.a. normale con media incognita e varianza 9;

(b) la v.a. X sia una v.a. normale con media e varianza incognite.

2 Su un campione casuale di 10.000 giovani di 11 anni, 8.000 risultano avere la licenza elementare. Costruire un intervallo al 95% per il valore della percentuale p di giovani di 11 anni con licenza elementare.

3 Due appezzamenti di uno stesso frutteto sono stati trattati con 2 diversi fertilizzanti. In ciascun appezzamento stato scelto a caso un campione di piante ed stato controllato il peso del raccolto in Kg ottenendo la seguente tabella

I campione 26,3 32,6 18,7 29,4

II campione 31,5 23,4 29,2 34,6 27,5

Supponendo che nella I popolazione il peso X_1 abbia distribuzione $N(\mu_1, \sigma^2)$ mentre nella II popolazione il peso X_2 abbia distribuzione $N(\mu_2, \sigma^2)$, con μ_1, μ_2, σ^2 incognite determinare:

(a) l'intervallo di confidenza al 95% di σ^2 ;

(b) l'intervallo di confidenza al 95% per la differenza tra le medie $\mu_1 - \mu_2$.

4 Sia $X_1, ..., X_n$ un campione estratto da una distribuzione di Poisson di parametro θ e sia $Y_n = \sum_{i=1}^n X_i$. Calcolare:

(a) la distribuzione di $Y_n, E(Y_n), Var(Y_n)$;

(b) il valore atteso di $S_n^2 = \displaystyle\sum_{i=1}^n \frac{(X_i - \overline{X}_n)^2}{n-1}$;

(c) stabilire quale stimatore di θ preferibile tra $\frac{Y_n}{n}$ e S_n^2;

(d) si supponga che il numero di studenti iscritti ogni anno al corso di laurea di Ingegneria Informatica segua una distribuzione di Poisson di parametro θ e che il numero degli studenti iscritti in anni diversi siano indipendenti tra loro. Si dia una stima puntuale di θ in base ai dati:

ANNO	**N° ISCRIZIONI**
2002/2003	82
2003/2004	155
2004/2005	197
2005/2006	257
2006/2007	272
2007/2008	248

(e) sulla base dei dati del punto precedente, si costruisca un intervallo di confidenza al livello $0,85$ per θ.

5 Sia $X \sim U[0, \theta]$. Si trovi un intervallo di confidenza al 95% per il parametro θ basato su un campione di ampiezza n=12 utilizzando una trasformata pivotale basata sullo stimatore di massima verosimiglianza.

6 In uno stabilimento si osservato che il tempo richiesto affinché un operaio compia un certo lavoro si distribuisce normalmente con media incognita e varianza $\sigma^2 = 144$;

(a) sulla base di un campione di 25 operai, calcolare l'ampiezza dell' intervallo di confidenza al 95% per il tempo medio impiegato a compiere il lavoro;

(b) di quanto si riduce l'intervallo considerando un campione di 100 operai?

7 Sia X una v.a. geometrica di parametro θ. Utilizzando le proprietà asintotiche degli stimatori di max verosimiglianza, determinare un intervallo di confidenza per al livello 0,9 per il parametro θ.

8 Un campione di 16 sigarette di una certa marca stato sottoposto ad analisi per misurarne il contenuto di nicotina. Si sono ottenuti i seguenti risultati: $\overline{x}_n = 20mg$, $s_n^2 = 4mg$. Si calcoli un intervallo di confidenza al 95% per il contenuto medio di nicotina X di ogni sigaretta, supponendo che $X \sim N(\mu, \sigma^2)$, μ, σ parametri incogniti.

9 Sia $X_1, ..., X_n, X_{n+1}$ un campione estratto da una distribuzione $N(\mu, \sigma^2)$, con media e varianza incognite. Sulla base dell' osservazione di $X_1, ..., X_n$, si vuole determinare un intervallo, detto *intervallo di predizione*, che conterrà il valore X_{n+1} con un livello di confidenza assegnato pari ad $1 - \alpha$. A tal fine, dette $\overline{X}_n$ ed S_n^2 la media e la varianza campionaria dell' osservazione $X_1, ..., X_n$ calcolare:

(a) la legge di $X_{n+1} - \overline{X}_n$;

(b) la legge di $\dfrac{X_{n+1} - \overline{X}_n}{S_n \sqrt{\frac{n+1}{n}}}$;

(c) l 'intervallo di predizione per X_{n+1}.

10 Un ingegnere civile vuole misurare la resistenza alla compressione di due diversi tipi di calcestruzzo. Viene provato un campione di 10 esemplari per ciascuno dei due tipi di materiale ottenendo i dati seguenti (in libbre per pollice quadrato):

Tipo 1	3250	3268	4302	3184	3266	3297	3332	3502	3064	3116
Tipo 2	3094	3106	3004	3066	2984	3124	3316	3212	3380	3018

Ipotizziamo che i campioni provengano da popolazioni normali con la stessa varianza. Determinare per la differenza delle medie delle due popolazioni, gli intervalli di confidenza al 95% seguenti:

(a) bilatero;

(b) unilaterale destro;

(c) unilaterale sinistro.

1.3 Test d'ipotesi statistiche

1 La degenza ospedaliera in giorni per il trattamento di una certa malattia per i pazienti della classe di età tra i 20 ed i 40 anni si distribuisce normalmente con media μ incognita e varianza 2,1. Sia n la dimensione del campione casuale che viene osservato per verificare l'ipotesi nulla $H_0 : \mu = 7$. Fissato un livello di significatività $\alpha = 5\%$, si costruisca:

(a) un test per la verifica dell'ipotesi nulla $H_0 : \mu = 7.$ in alternativa all'ipotesi $H_1 : \mu \neq 7$ quando n=16;

(b) si calcoli la funzione di potenza del test quando $\mu = 7.6$ e quando $\mu = 9$; motivare i cambiamenti osservati.

(c) riscrivere i risultati precedenti per n=49 e motivare il cambiamento della regione critica.

2 In uno studio fatto per accertare l'efficacia dell'esercizio fisico per dimagrire, sono stati osservati su 9 campioni i seguenti risultati in termini di riduzione di peso in Kg:

$$-5.9 \quad -3.2 \quad 0.5 \quad -2.3 \quad -1.4 \quad -0.9 \quad 0.5 \quad 0 \quad 3.6$$

Assumendo che la riduzione del peso segua una legge normale con media e varianza incognita, si verifichi l'efficacia del programma al livello dell'1% e del 5%.

3 Il contenuto effettivo in litri di 16 confezioni di latte prelevati a caso da una linea di produzione sono:

1.03	1.01	0.94	0.99	1.10	1.03	0.89	1.07
1.12	1.06	1.02	0.94	0.99	1.09	1.01	1.12

Supponendo che siano determinazioni di una v.a. normale, si costruisca un test:

(a) per la verifica dell'ipotesi nulla $H_0 : \sigma^2 = 0.0049$ in alternativa all'ipotesi $H_1 : \sigma^2 \neq 0.0049$ ad un livello di significatività dell'1%;

(b) la verifica dell'ipotesi che il contenuto effettivo di una confezione sia pari ad un litro ad un livello di significatività del 5%.

4 Si supponga che per determinare a quale epoca appartengono alcuni oggetti rinvenuti in uno scavo archeologico, si ricorre ad un metodo basato sulla quantità di radiazioni X emesse in un dato intervallo di tempo da ciascun oggetto. Si assuma che X sia una v.a. di Poisson di parametro incognito θ. Supponendo di poter fare 25 misurazioni, si determini un test ottimale con livello di significatività $\alpha = 0.05$ per le due ipotesi:

H_0: ciascun oggetto appartiene ad un'epoca A caratterizzata da $\theta = 9$

H_1: ciascun oggetto appartiene ad un'epoca B caratterizzata da $\theta = 4$

5 Una ditta produce cavi i cui carichi di rottura hanno una distribuzione che si può considerare normale con media pari a 300 kg e deviazione standard di 24 Kg. Viene proposta l'installazione di un nuovo processo produttivo mediante il quale il carico di rottura medio verrebbe aumentato, restando invariata la deviazione standard. Per decidere se il caso di sostituire il vecchio processo si adotta il seguente criterio, provando 64 cavi: Se il carico medio di rottura supera i 307 Kg allora si prende in considerazione la proposta di installare il nuovo processo altrimenti si lascia la situazione invariata. Si discutano le implicazioni di tale decisione, in particolare valutando:

(a) la probabilità di accettare una proposta non valida;

(b) la probabilità di rifiutare una proposta valida, se le prove effettuate indicano che il nuovo processo produce un aumento del 5% del carico di rottura medio.

6 Si consideri la seguente distribuzione secondo il sesso e l'atteggiamento nei confronti del fumo effettuata su un campione di 191 soggetti:

	FAVOREVOLE	CONTRARIO	INDIFFERENTE	Totale
MASCHI	9	55	19	83
FEMMINE	10	71	27	108
Totale	19	126	46	191

Si verifichi l'ipotesi di indipendenza tra i due caratteri ad un livello di significatività $\alpha = 0,01$.

7 Un algoritmo per la simulazione di una v.a. N(0,1) fornisce i seguenti valori:

0.464	0.137	2.455	−0.323	−0.068
0.906	−0.513	−0.525	0.595	0.881
−0.482	1.678	−0.057	−1.229	−0.486
−1.787	−0.261	1.237	1.046	−0.508

Verificare tramite un test di ampiezza $\alpha = 0,05$ se la procedura di simulazione è corretta.

8 In un periodo temporale di 72 ore vengono rilevati il numero di incidenti automobilistici per ora. I dati ottenuti sono i seguenti:

Num. incidenti per ora	Num. di ore
0 o 1	5
2	10
3	15
4	12
5	12
6	6
7	5
≥ 8	7

Verificare tramite un test opportuno se il numero di incidenti per ora è una variabile aleatoria di Poisson.

9 Un dado è lanciato 120 volte con i seguenti risultati:

	1	2	3	4	5	6
Frequenze:	20	30	20	25	15	10

Costruire un test per verificare, a livello $\alpha = 0,01$, se il dado sia o meno truccato.

10 Un esperimento consiste nel lanciare una moneta finché non esce testa. L' esperimento viene ripetuto 100 volte con i seguenti risultati

Num. lanci effettuati:	1	2	3	4	≥ 5
Frequenze:	40	32	15	7	6

Possiamo concludere, ad un livello di significatività $\alpha = 0,01$ che la moneta è equa?