DIPLOMATURA D'ESTADÍSTICA

ESTADÍSTICA MATEMÀTICA 2

Guadalupe Gómez Melis

FACULTAT DE MATEMÀTIQUES I ESTADÍSTICA

DIPLOMATURA D'ESTADÍSTICA

ESTADÍSTICA MATEMÁTICA 2

Professora coordinadora:
Guadalupe Gómez i Melis

Professors:
Guadalupe Gómez i Jaume Coll

Consultes:
Guadalupe Gómez: Dilluns 5–6, Dijous 5–6 i hores convingudes al despatx 415.
Jaume Coll: Dimecres 2:00–3:00, Dijous 11–12 i hores convingudes al despatx 407.

PROGRAMA

1. ESTIMACIÓ PUNTUAL

1.1 La distribució mostral d'un estadístic.

1.2 El mètode dels moments. El concepte de consistència.

1.3 El mètode de la màxima versemblança. Invariància i consistència.

1.4 El concepte de suficiència. Teorema de factorització.

1.5 Famílies exponencials.

1.6 Estadístics basats en la llei normal. Distribució conjunta de la mitjana mostral i de la variança mostral.

1.7 Estimadors U.M.V.U. Teorema de Rao-Blackwell.

1.8 La informació de Fisher, el concepte d'eficiència i la cota de Cramer-Rao.

1.9 Propietats asimptòtiques dels estimadors de màxima versemblança.

2. PROVES D'HIPÒTESIS INTERVALS DE CONFIANÇA

2.1 Introducció. Conceptes bàsics en la teoría de les proves d'hipòtesis.

2.2 Proves d'hipòtesis òptimes. El lema de Neyman–Pearson.

2.3 Proves d'hipòtesis U.M.P.La prova de la raó de versemblança.

2.4 La distribució t. Aplicació a un problema de regressió. La distribució t descentrada.

2.5 Prova de la raó de versemblança generalitzada. La prova de la t de Student.

2.6 La distribució F. La prova de la F.

2.7 Comparació de dues mostres independents. La prova de la t per dues mostres.

2.8 Intervals de confiança.

3. PROVES PER A LA VALIDESA D'UN MODEL

3.1 La prova de la raó de versemblança per a distribucions multinomials.

3.2 La prova de χ^2.

3.3 Prova de Kolmogorov-Smirnov.

3.4 Papers i gràfics de probabilitat.

4. INTRODUCCIÓ A L'INFERÈNCIA BAYESIANA.

4.1 Lleis de Bayes i Minimax.

4.2 Anàlisi per a trobar una llei de Bayes.

4.3 Estimadors admissibles.

4.4 Inferència Bayesiana per a una distribució normal.

4.5 Inferència Bayesiana per a una distribució binomial.

5. MÈTODES NO PARAMÈTRICS

5.1 La prova dels signes

5.2 La prova dels rangs signats de Wilcoxon.

5.3 La prova dels rangs de Wilcoxon, Mann i Whitney.

AVALUACIÓ

⋆ Examen Parcial (**EP**) corresponent al tema 1.

⋆ Examen Final (**EF**) corresponent als temes 1, 2, 3, 4 i 5. Data prevista: 27 de maig de 1997.

⋆ La Qualificació Final (**QF**) es calculará de la següent manera:

Si Nota EP$\geq$ Nota EF: **QF=0.4 EP+0.6 EF**

Si Nota EP$<$ Nota EF: **QF= EF**

BIBLIOGRAFIA

- Cuadras,C. "Problemas de Probabilidades y Estadística. Volum 2". Promociones y Publicaciones Universitarias, 1990.
- DeGroot, M.H. "Probability and Statistics". Addison-Wesley, 1989.
- Hogg, R.V. i Craig, A.T. "Introduction to Mathematical Statistics". Macmillan Publishing Co., Inc, 1978.
- Kalbfleisch, J.G. "Probability and Statistical Inference. Volum 2". Springer-Verlag, 1985.
- Kiefer, J.C. "Introduction to Statistical Inference". Springer–Verlag, 1987.
- Lee, P.M. "Bayesian Statistics: An Introduction". Edward Arnold, 1989.
- Peña, D. "Estadística. Modelos y métodos 1. Fundamentos". Alianza Universidad Textos, 1989.
- Rice, J.A. "Mathematical Statistics and Data Analysis". Wadsworth i Brooks/Cole, 1988.

1. ESTIMACIÓ PUNTUAL

1.1 La distribució mostral d'un estadístic.

1.2 El mètode dels moments. El concepte de consistència.

1.3 El mètode de la màxima versemblança. Invariància i consistència.

1.4 El concepte de suficiència. Teorema de factorització.

1.5 Famílies exponencials.

1.6 Estadístics basats en la llei normal. Distribució conjunta de la mitjana mostral i de la variança mostral.

1.7 Estimadors U.M.V.U. Teorema de Rao-Blackwell.

1.8 La informació de Fisher, el concepte d'eficiència i la cota de Cramer-Rao.

1.9 Propietats asimptòtiques dels estimadors de màxima versemblança.

EXERCICIS:

1.1 Suposeu que X segueix una distribució geomètrica de paràmetre p. Prenem una mostra de grandària n d'aquesta distribució.

a) Trobeu l'estimador de p pel mètode dels moments.

b) Trobeu l'estimador de p pel mètode de la màxima versemblança.

c) Calculeu la variància asimptòtica de l'estimador de màxima versemblança.

1.2 Suposeu que X segueix una distribució de Rayleigh (cas particular de la Weibull) de paràmetre θ. Prenem una mostra de grandària n d'aquesta distribució.

a) Trobeu l'estimador de θ pel mètode dels moments.

b) Trobeu l'estimador de θ pel mètode de la màxima versemblança.

c) Calculeu la variància asimptòtica de l'estimador de màxima versemblança.

1.3 Suposeu que X segueix una distribució uniforme $[0,\theta]$. Prenem una mostra de grandària n d'aquesta distribució.

a) Trobeu l'estimador de θ pel mètode dels moments, així com la seva mitjana i variància.

b) Trobeu l'estimador de θ pel mètode de la màxima versemblança.

c) Trobeu la funció de densitat de l'estimador de màxima versemblança, i calculeu la seva mitjana i la seva variància. Compareu la variància, el biaix i l'error quadràtic mitjà amb l'obtingut pel mètode dels moments.

d) Trobeu una modificació de l'estimador de màxima versemblança que el faci no esbiaixat.

1.4 Suposeu que X segueix una distribució exponencial de mitjana τ. Prenem una mostra de grandària n d'aquesta distribució.

a) Trobeu l'estimador de τ pel mètode dels moments. Anomeneu-lo $\hat{\tau}_1$.

b) Trobeu l'estimador de τ pel mètode de la màxima versemblança. Anomeneu-lo $\hat{\tau}_2$.

c) Calculeu la seva variància. És $\hat{\tau}_2$ eficient?

d) Calculeu la distribució asimptòtica de $\hat{\tau}_2$.

e) Doneu la forma d'un interval de confiança exacta per a τ.

1.5 Suposeu que X segueix una distribució exponencial amb taxa de falla λ. Prenem una mostra de grandària n d'aquesta distribució.

a) Trobeu l'estimador de λ pel mètode dels moments. Anomeneu-lo $\hat{\lambda}_1$.

b) Trobeu l'estimador de λ pel mètode de la màxima versemblança. Anomeneu-lo $\hat{\lambda}_2$.

c) Demostreu que $\hat{\lambda}_2$ té biaix i calculeu la seva variància.

d) Trobeu la fita de Cràmer-Rao i compareu-la amb la variància trobada a c).

e) Doneu la forma d'un interval de confiança aproximat per a λ.

f) Doneu la forma d'un interval de confiança per a λ.

1.6 Suposeu que X segueix una distribució lognormal de paràmetres μ i σ^2. Prenem una mostra de grandària n d'aquesta distribució.

a) Trobeu els estimadors de μ i σ^2 pel mètode dels moments.

b) Trobeu els estimadors de μ i σ^2 pel mètode de la màxima versemblança.

c) Calculeu la matriu de variances asimptòtiques dels estimadors de màxima versemblança trobats a l'apartat b).

1.7 Suposeu que X segueix una distribució de Weibull paràmetres a i b. Prenem una mostra de grandària n d'aquesta distribució.

a) Trobeu els estimadors de a i b pel mètode de la màxima versemblança.

b) Calculeu la matriu de variances asimptòtiques dels estimadors de màxima versemblança trobats a l'apartat a).

1.8 Si les frequències genètiques estan en equilibri, els genotipus AA, Aa i aa ocorren amb probabilitats $(1-\theta)^2$, $2\theta(1-\theta)$ i θ^2. Plato i *al.*(1964) publicaren les següents dades sobre el tipus d'haptoglobina en una mostra de 190 persones.

tipus	Hp1-1	Hp1-2	Hp2-2
frequència	10	68	112

Trobeu l'estimador de màxima versemblança de θ i la seva variància asimptòtica. Doneu un interval de confiança aproximat al 99% per a θ.

1.9 L'estimador de màxima versemblança no és únic.

Suposeu que X segueix una distribució uniforme $[\theta, \theta+1]$. Prenem una mostra de grandària n d'aquesta distribució. Demostreu que l'estimador de θ obtingut pel mètode de la màxima versemblança no queda unívocament especificat.

1.10 L'estimador de màxima versemblança no sempre existeix

Considerem una v.a. X que pot provenir d'una distribució $N(0,1)$ amb probabilitat 0.5 i d'una distribució $N(\mu, \sigma^2)$, amb μ i σ^2 desconeguts, amb probabilitat 0.5 . Demostreu que no existeix l'estimador de màxima versemblança.

1.11 L'estimador de màxima versemblança no sempre existeix

Considerem una v.a. X que prové d'una distribució amb funció de densitat

$$f(x;\theta) = ae^{-ax}\mathbf{1}\{0 \leq x \leq \tau\} + be^{-a\tau - b(x-\tau)}\mathbf{1}\{x > \tau\},$$

on $x \in \mathcal{R}$ i $\theta = (a, b, \tau) \in \Theta = (0, \infty)^3$. Demostreu que no existeix l'estimador de màxima versemblança.

1.12 Sigui $X_1, \ldots, X_n$ una mostra d'una població amb distribució subjacent governada per un paràmetre θ. Estant interessats en l'estimació d'una funció $q(\theta)$ es proposa un estimador $T(X_1, \ldots, X_n)$. Demostreu que per n gran, i sota les hipòtesis que siguin necessàries per la correcta aplicació del Teorema del Límit Central,

$$E\{|T(X_1, \ldots, X_n) - q(\theta)|\} = \sqrt{2/\pi}\sqrt{R(\theta, T)}$$

a on

$$R(\theta, T) = E\{(T(X_1, \ldots, X_n) - q(\theta))^2\}.$$

1.13 Sigui X una v.a. normal de mitjana θ i de variància 1. Definim $T_{a,b}(X) = aX + b$ per a i b nombres reals qualsevols.

a) Calculeu l'error quadràtic mig, $R(\theta, T_{a,b})$, de $T_{a,b}$.

b) Compareu $R(\theta, T_{0.5,0})$ amb $R(\theta, T_{1,0})$. Quina conclusió treieu?

c) Existeix algun estimador de la forma $aX + b$ que sigui millor que X per θ?

d) Proveu que X és l'únic estimador no esbiaxat de θ de la forma $aX + b$.

1.14 Considereu el següent mètode per a estimar el paràmetre λ d'una distribució de Poisson. Observem que

$$p_0 = \text{Prob}(X = 0) = e^{-\lambda}$$

Sigui Y la v.a. que representa el número de zeros d'una mostra de grandària n. Podem estimar λ mitjançant

$$\hat{\lambda} = -\log(\frac{Y}{n})$$

a) Feu servir el mètode del'a per a obtenir expressions aproximades per a la variància i el biaix d'aquest estimador.

b) Compareu la variància d'aquest estimador amb la variància de l'estimador de màxima versemblança, i calculeu l'eficiència relativa per a diferents valors de λ. Indicació: $Y \sim Bin(n, p_0)$.

1.15 Sigui $X_1, \ldots, X_n$ una mostra d'una distribució normal amb mitjana desconeguda θ i variància σ^2. Suposant que $\theta \neq 0$, determineu la distribució asimptòtica de $\bar{X}_n^3$.

1.16 Sigui $X_1, \ldots, X_n$ una mostra d'una distribució de Poisson de mitjana λ, i sigui $T = \sum_{i=1}^n X_i$. Demostreu que la distribució de $X_1, \ldots, X_n$ donat T és independent de λ, i concluiu que T és suficient per a λ. Demostreu que X_1 no és suficient.

1.17 Sigui $X_1, \ldots, X_n$ una mostra d'una distribució geomètrica, i sigui $T = \sum_{i=1}^n X_i$. Demostreu que T és un estadístic suficient.

1.18 Useu el teorema de factorització per a trobar un estadístic suficient per a la distribució exponencial.

1.19 Useu el teorema de factorització per a trobar un estadístic suficient per a la distribució de Rayleigh.

1.20 Demostreu que la distribució gamma pertany a la família exponencial.

1.21 Siguin T_1 i T_2 dos estadístics suficients per a θ, i suposeu que $T_2 = g(T_1)$ per alguna funció g. Sigui U un estimador no esbiaixat de θ, i sigui

$$\begin{aligned} V_1 &= E(U \mid T_1) \\ V_2 &= E(V_1 \mid T_2) \end{aligned}$$

Demostreu que $\mathrm{Var}(V_2) \leq \mathrm{Var}(V_1)$.

1.22 Sigui $\bar{X}$ la mitjana d'una mostra de 16 v.a. independents i normals (0,1). Determineu c tal que

$$\mathrm{Prob}(\mid \bar{X} \mid < c) = 0.5$$

1.23 Useu la distribució chi quadrat per a calcular

$$\mathrm{Prob}(a < \frac{S^2}{\sigma^2} < b),$$

basant-vos en una mostra de grandària 18. Calculeu per a $a = 0.51$ i $b = 1.62$.

1.24 Suposeu que hem de prendre una mostra d'una distribució normal amb mitjana μ desconeguda i desviació estàndar 2. Calculeu la grandària de la mostra necessària per tal que

a)

$$E_\mu(\mid \bar{X}_n - \mu \mid^2) \leq 0.1$$

per a cada possible valor de μ.

b)

$$E_\mu(\mid \bar{X}_n - \mu \mid) \leq 0.1$$

per a cada possible valor de μ.

c)

$$\mathrm{Prob}_\mu(\mid \bar{X}_n - \mu \mid \leq 0.1) \geq 0.95$$

per a cada possible valor de μ.

1.25 Suposeu que les v.a. X_1, X_2, X_3 són i.i.d., cada una d'elles N(0,1). Definim les següents v.a.

$$\begin{aligned} Y_1 &= 0.8X_1 + 0.6X_2 \\ Y_2 &= \sqrt{2}(0.3X_1 - 0.4X_2 - 0.5X_3) \\ Y_3 &= \sqrt{2}(0.3X_1 - 0.4X_2 + 0.5X_3) \end{aligned}$$

Trobeu la llei conjunta de Y_1, Y_2, Y_3.

1.26 Si les v.a. X_1 i X_2 són independents i $N(\mu, \sigma^2)$, proveu que les v.a. $X_1 + X_2$ i $X_1 - X_2$ són independents.

1.27 Suposem que $X_1, \ldots, X_n$ formen una mostra aleatòria d'una distribució normal de mitjana μ i variància σ^2. Trobeu la distribució de

$$\frac{n(\bar{X}_n - \mu)^2}{\sigma^2}.$$

1.28 Suposem que $X_1, \ldots, X_6$ formen una mostra aleatòria d'una distribució normal estàndar. Sigui

$$Y = (X_1 + X_2 + X_3)^2 + (X_4 + X_5 + X_6)^2.$$

Determineu un valor c tal que la v.a cY segueixi una distribució χ^2.

1.29 Suposem que $X_1, \ldots, X_{16}$ formen una mostra aleatòria d'una distribució normal de mitjana μ i variància σ^2. Calculeu les següents probabilitats:

a)

$$\text{Prob}[\frac{\sigma^2}{2} \leq \frac{1}{n}\sum_{i=1}^{n}(X_i-\mu)^2 \leq 2\sigma^2]$$

b)

$$\text{Prob}[\frac{\sigma^2}{2} \leq \frac{1}{n}\sum_{i=1}^{n}(X_i-\bar{X}_n)^2 \leq 2\sigma^2]$$

1.30 Sigui X una v.a. normal de mitjana 0 i amb desviació estàndar σ ($\sigma > 0$) desconeguda.

a) Calculeu la informació de Fisher $I(\sigma)$ continguda en X.

b) Calculeu la informació de Fisher $I(\sigma^2)$ continguda en X.

1.31 Sigui $X_1, \ldots, X_n$ una mostra aleatòria d'una distribució de Poisson de paràmetre λ desconegut. Demostreu que $\bar{X}_n$ és el U.M.V.U. de λ.

1.32 Sigui $X_1, \ldots, X_n$ una mostra aleatòria d'una distribució que té per funció de densitat:

$$f(x;\theta) = \theta x^{\theta-1}, \quad 0 < x < 1, \quad \theta > 0.$$

a) Cerqueu un estadístic suficient $T(X_1 \ldots, X_n)$.

b) Calculeu $E(T)$ i $\text{Var}(T)$. Indicació: Feu servir el fet de pertànyer a la família exponencial.

c) Calculeu $I(\theta)$ i demostreu que $-T/n$ és el U.M.V.U. de $1/\theta$.

1.33 Suposem que fem una única observació X d'una distribució normal de mitjana 0 i amb desviació estàndar σ ($\sigma > 0$) desconeguda.

a) Trobeu un estimador no esbiaixat per a σ.

b) Calculeu la variància d'aquest estimador.

c) Demostreu que aquesta variància és superior a $1/I(\sigma)$ per cada valor $\sigma > 0$.

1.34 Sigui $X_1, \ldots, X_n$ una mostra aleatòria d'una distribució de Bernoulli de paràmetre p desconegut. Demostreu que $\bar{X}_n$ és un estimador eficient de p.

1.35 PROBLEMA D'EXAMEN

Suposeu que prenem una mostra aleatòria d'una distribució exponencial de mitjana θ desconeguda. Calculeu la grandària de mostra necessària per tal que

$$E\{\mid \bar{X}_n - \theta \mid^2\} \leq 0.01$$

1.36 PROBLEMA D'EXAMEN

Suposem que (X,Y) és un punt que s'ha d'escollir del pla real i que X i Y indiquen dues variables aleatòries independents amb distribució normal estàndar. Quin ha d'ésser el radi del cercle C (centrat a l'origen) per tal que el punt (X,Y) sigui dins de C amb probabilitat 0.99?

1.37 PROBLEMA D'EXAMEN

La distribució de Pareto s'utilitza en ciències econòmiques com a model per a una funció de densitat amb una cua que decreix a poc a poc:

$$f(x; x_0, \theta) = \theta x_0^\theta x^{-\theta-1}, \quad x \geq x_0, \quad \theta > 1.$$

Suposem que el valor x_0 és conegut.

a) Calculeu l'esperança d'una v.a. X que segueix una distribució de Pareto. Disposem d'una mostra $X_1, \ldots, X_n$ d'una llei de Pareto.

b) Trobeu l'estimador $\hat{\theta}_1$ de θ pel mètode dels moments.

c) Trobeu l'estimador $\hat{\theta}_2$ de θ pel mètode de la màxima versemblança.

d) Quina llei asimptòtica segueix $\hat{\theta}_2$.? Precisseu en particular la seva variança.

1.38 PROBLEMA D'EXAMEN

Disposem d'una mostra $X_1, \ldots, X_n$ d'una llei que té per funció de densitat

$$f(x; \theta) = \frac{\theta}{(1+x)^{\theta+1}}, \quad 0 < \theta < \infty \quad 0 \leq x < \infty.$$

Doneu un estadístic suficient per a θ. Raoneu la resposta.

1.39 PROBLEMA D'EXAMEN

Disposem d'una mostra $X_1, \ldots, X_n$ d'una llei que té per densitat

$$f(x; \theta) = (\theta + 1)x^\theta, \quad 0 \leq x \leq 1.$$

a) Calculeu l'esperança i la variança de X.

b) Trobeu l'estimador $\hat{\theta}_1$ de θ pel mètode dels moments.

c) Quina llei asimptòtica segueix $\hat{\theta}_1$.? Precisseu en particular la seva variança. Indicació: Feu servir el Teorema Central del Límit i el mètode Delta.

d) Trobeu l'estimador $\hat{\theta}_2$ de θ pel mètode de la màxima versemblança.

e) Quina llei asimptòtica segueix $\hat{\theta}_2$.? Precisseu en particular la seva variança.

f) Calculeu l'eficiència relativa asimptòtica de $\hat{\theta}_1$. respecte $\hat{\theta}_2$. i deduiu quin dels dos estimadors és més eficient.

1.40 PROBLEMA D'EXAMEN

Sigui $X_1, \ldots, X_n$ una mostra d'una llei que té per densitat

$$f(x; \theta) = \frac{1}{6\theta^4} x^3 e^{\frac{-x}{\theta}}, \quad 0 < x < \infty, 0 < \theta < \infty.$$

a) Trobeu un estadístic suficient per a θ.

b) Utilitzeu el teorema de Rao-Blackwell per a trobar un estadístic suficient que ens millori la variança.

1.41 PROBLEMA D'EXAMEN

Aquest problema fa referència a l'estimació de la variança d'una distribució normal de mitjana desconeguda a partir d'una mostra $X_1, \ldots, X_n$ de v.a. normals. Considerem tres possibles estimadors de la variança poblacional:

$$S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X})^2$$

$$\hat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^{n}(X_i - \bar{X})^2$$

$$\sigma_*^2 = \rho\sum_{i=1}^{n}(X_i - \bar{X})^2$$

(1) Calculeu el biaix de cadascun d'aquests tres estimadors. Quins d'aquests són no esbiaixats?

(2) Calculeu la variança de cadascun d'aquests tres estimadors.

(3) Calculeu l'error quadràtic mitjà dels dos primers estimadors.

(4) Per quin valor de ρ l'estimador σ_*^2 té error quadràtic mitjà més petit?

(5) Quin estimador escollirieu i per què?

1.42 PROBLEMA D'EXAMEN

Siguin $X_1, \ldots, X_n$ v.a. independents distribuïdes segons una llei uniforme a l'interval $[0, \theta]$. Denotem per $X_{(n)}$ el màxim de les n v.a.

(1) Calculeu la distribució de $\frac{X_{(n)}}{\theta}$.

(2) Doneu un interval de confiança per a θ de nivell 0.9.

1.43 PROBLEMA D'EXAMEN

Disposem d'una mostra $X_1, \ldots, X_n$ d'una llei que té per densitat

$$f(x;\theta) = e^{-(x-\theta)}, \quad \theta \in R, \quad \theta \leq x < \infty.$$

1) Trobeu l'estimador $\hat{\theta}_1$ de θ pel mètode dels moments.

2) Trobeu l'estimador $\hat{\theta}_2$ de θ pel mètode de la màxima versemblança.

3) Calculeu la funció de densitat de l'estimador de màxima versemblança, $\hat{\theta}_2$. És $\hat{\theta}_2$ suficient? Raoneu la resposta.

4) Utilitzeu el teorema de Rao-Blackwell per a trobar un estadístic suficient, no esbiaixat i amb variança més petita que la de $\hat{\theta}_2$.

1.44 PROBLEMA D'EXAMEN

Sigui $X_1, \ldots, X_n$ una mostra d'una llei exponencial de paràmetre λ. Demostreu que la v.a.

$$U = \frac{X_1}{\sum_{i=1}^{n} X_i}$$

segueix una llei beta de paràmetres $(1, n-1)$.

1.45 PROBLEMA D'EXAMEN

La vida (en hores) d'una component electrònica és una variable aleatòria T que es distribueix com una exponencial amb taxa de falla λ. Estem interessats en la probabilitat de que aquesta component falli (mori) abans de 3 hores. Per això prenem n components electròniques i anotem llurs vides; és a dir, disposem d'una mostra de grandària n, $T_1, \ldots, T_n$.

a) Calculeu la probabilitat $p = p(\lambda)$ que una component electrònica falli abans de 3 hores.

b) Estimeu p pel mètode dels moments. Doneu una interpretació a l'estimador obtingut.
Indicació: Definir
$$X_i = \begin{cases} 1 & \text{si } T_i \leq 3 \\ 0 & \text{si } T_i > 3 \end{cases}$$

c) Cerqueu un estadístic suficient i complet per aquesta família. Raoneu la resposta.

d) Trobeu un estimador U.M.V.U. per p.
Indicació 1: Sigui $S(X_1) = \mathbf{1}(X_1 \leq 3)$. Tenim que $E(S(X_1)|T = t) = \text{Prob}(X_1 \leq 3|T = t)$.
Indicació 2: Feu servir el resultat del problema 1.44

1.46 PROBLEMA D'EXAMEN

Sigui $X_1, \ldots, X_n$ v.a. independents i idènticament distribuïdes amb funció de densitat

$$f(x;\theta) = \frac{1}{6\theta^4} x^3 \exp\{-x/\theta\} \quad \text{per} \quad 0 < x < \infty \quad \text{i} \quad \theta > 0.$$

a) Obtingueu, mitjançant el mètode dels moments, un estimador per a θ.

b) Calculeu la variància de l'estimador obtingut a l'apartat a).

c) Obtingueu, mitjançant el mètode de la màxima versemblança, un estimador per a θ.

d) Doneu la llei asimptòtica de l'estimador de màxima versemblança.

e) Podeu concloure que l'estimador de màxima versemblança és eficient? Raoneu la resposta.

1.47 PROBLEMA D'EXAMEN Disposem d'una mostra $X_1, \ldots, X_n$ d'una llei que té per densitat ($\theta > 0$, $\tau > 0$)

$$f(x;\theta,\tau) = \begin{cases} \frac{\theta\tau^\theta}{x^{\theta+1}} & x \geq \tau \\ 0 & x < \tau \end{cases}.$$

Suposem primer que el valor de τ és conegut i el de θ desconegut.

(a) Calculeu l'estimador $\hat{\theta}_1$ de màxima versemblança de θ.

(b) És aquest estimadors $\hat{\theta}_1$ suficient pel paràmetre θ.

(c) Calculeu la distribució aproximada de $\hat{\theta}_1$.

Suposem ara que el valor de τ és desconegut i el de θ conegut.

(d) Calculeu l'estimador $\hat{\tau}_1$ de màxima versemblança de τ.

(e) És l'estimador $\hat{\tau}_1$ suficient pel paràmetre τ.

Suposem per últim que els dos valor de τ i de θ són desconeguts.

(f) Calculeu els estimadors de màxima versemblança de θ i de τ.

1.48 PROBLEMA D'EXAMEN

Disposem d'una mostra de grandària n d'una població Weibull amb paràmetre $\beta = 0.25$ i $\alpha > 0$ desconegut.

(a) Calculeu l'estimador de α, $\hat{\alpha}_1$, pel mètode dels moments.

(b) Calculeu l'estimador de α, $\hat{\alpha}_2$, pel mètode de la màxima versemblança.

(c) Calculeu la variància asimptòtica de $\hat{\alpha}_2$.

1.49 PROBLEMA D'EXAMEN

Considerem la variable aleatòria T que mesura el temps de supervivència en mesos d'un pacient desprès d'un tractament. Suposem que T es distribueix segons un model exponencial traslladat.

El model exponencial traslladat de paràmetres λ ($\lambda > 0$) i G ($G > 0$) té per funció de densitat:

$$f(t;\lambda,G) = \begin{cases} \lambda \exp\{-\lambda(t-G)\} & \text{si } t \geq G \geq 0 \\ 0 & \text{si } t < G \end{cases}$$

G s'interpreta com el temps de garantia o el mínim temps de vida abans del qual no hi han morts.

Hem fet un estudi amb 11 pacients i llurs temps de supervivència han estat:

$$11, 13, 13, 13, 13, 13, 14, 14, 15, 15, 17.$$

Suposem que aquests pacients són una mostra d'una llei exponencial traslladada de paràmetres λ i G.

a) Calculeu la funció de versemblança d'aquestes dades en funció dels paràmetres desconeguts.

b) Suposeu G conegut. Estimeu la taxa de fallada λ mitjançant el mètode de la màxima versemblança. Calculeu el valor de l'estimador de λ en funció de G pels pacients de l'estudi.

c) Suposeu λ conegut. Estimeu G mitjançant el mètode de la màxima versemblança. Calculeu el valor de l'estimador de G pels pacients de l'estudi.

d) Suposeu λ i G desconeguts.

d1) Trobeu els estimadors de màxima versemblança per λ i G. Calculeu-los pels pacients de l'estudi.

2) Feu servir les estimacions obtingudes per a estimar la probabilitat de sobreviure 18 mesos desprès del tractament.

e) Suposeu G conegut (podeu fer servir el valor trobat a l'apartat c)).

e1) Trobeu l'estimador de màxima versemblança de la mediana del temps de supervivència i anomeneu-lo $\hat{m}$. Calculeu la mediana del temps de supervivència dels 11 pacients.

e2) Determineu la distribució asimptòtica de l'estimador $\hat{m}$. Calculeu la variància asimptòtica de $\hat{m}$.

1.50 PROBLEMA D'EXAMEN

La vida (en minuts) d'un prototipus d'una component electrònica en un test de vida accelerat és una variable aleatòria X que es distribueix com una exponencial amb mitjana θ. Estem interessats en la probabilitat de que aquesta component falli abans de 5 minuts. Per això prenem n components i anotem llurs vides; és a dir, disposem d'una mostra de grandària n, $X_1, \ldots, X_n$.

a) Calculeu la probabilitat $q = q(\theta)$ que una component electrònica falli abans de 5 minuts.

b) Estimeu q pel mètode dels moments. Doneu una interpretació a l'estimador obtingut.
Indicació: Definir

$$Y_i = \begin{cases} 1 & \text{si} \quad X_i \leq 5 \\ 0 & \text{si} \quad X_i > 5 \end{cases}$$

c) Cerqueu un estadístic suficient per aquesta família. Raoneu la resposta.

d) Trobeu un estimador U.M.V.U. per q.
Indicació 1: Sigui $S(Y_1) = \mathbf{1}(Y_1 \leq 5)$. Tenim que $E(S(Y_1)|X = x) = \text{Prob}(Y_1 \leq 5|X = x)$.

1.51 PROBLEMA D'EXAMEN

Una variable aleatòria X segueix una distribució normal de mitjana μ i de variança 1. Es fan 20 observacions de X però en comptes d'escriure el valor de X només s'observa si X és negativa o no.

a) Estimar μ per màxima versemblança basant–se tan sols en la informació disponible.

b) Suposant que l'esdeveniment $\{X < 0\}$ ha ocurregut exactament 14 vegades, trobeu el valor de l'estimador de màxima versemblança de μ.

1.52 PROBLEMA D'EXAMEN

Considerem una mostra $X_1, \ldots, X_n$ d'una llei que té per densitat

$$f(x;\theta) = \frac{3\theta^3}{(x+\theta)^4} \qquad x > 0,$$

i l'espai de paràmetres és $\Omega = \{\theta : \theta > 0\}$

Indicació: Molts dels apartats d'aquest problema poden fer–se independentment dels altres.

a) Cerqueu un estimador de θ pel mètode dels moments. Anomeneu–lo $\hat{\theta}_1$.

b) Calculeu la variància de $\hat{\theta}_1$. Indicació: Pot resultar més senzill començar calculant $E(X+\theta)^2$.

c) Cerqueu un estimador de θ pel mètode de la máxima versemblança. Anomeneu–lo $\hat{\theta}_2$. Deixeu-lo indicat si no podeu resoldre les eqüacions de versemblança.

d) Calculeu la informació de Fisher continguda en la mostra $X_1, \ldots, X_n$.

e) És l'estimador $\hat{\theta}_1$ eficient?

f) Doneu la distribució asimptòtica de $\hat{\theta}_2$ i calculeu la seva variància asimptòtica.

g) Compareu $\hat{\theta}_1$ i $\hat{\theta}_2$. Quin dels dos estimadors és asimptòticament més eficient.?

h) Calculeu la mediana poblacional en funció de θ i anomeneu-la $M(\theta)$.

i) Doneu un estimador per $M(\theta)$ i determineu la seva distribució asimptòtica.

1.53 PROBLEMA D'EXAMEN

Un científic pren n mesures $X_1, X_2, \ldots, X_n$ d'una constant μ fent servir la tècnica T_1 amb una variància de l'error coneguda igual a σ^2. També pren m mesuraments addicionals $Y_1, Y_2, \ldots, Y_m$ de μ fent servir la tècnica T_2 amb variància de l'error coneguda igual a $k\sigma^2$, éssent k una constant positiva coneguda. Els mesuraments es suposen independents i amb distribució normal.

a) Escriure la funció de versemblança pel paràmetre μ basada en les $n+m$ observacions $X_1, X_2, \ldots, X_n$, $Y_1, Y_2, \ldots, Y_m$.

b) Trobeu l'estimador de màxima versemblança de μ. Anomeneu–lo $\hat{\mu}$.

c) Calculeu la distribució de $\hat{\mu}$.

d) Demostreu que si $n = m$ i $k > 1$, aleshores $\hat{\mu}$ està més a prop de $\bar{X}$ que de $\bar{Y}$ i expliqueu perquè és això desitjable.

1.54 PROBLEMA D'EXAMEN

La vida (en hores) d'una component electrònica és una variable aleatòria X que es distribueix com una exponencial amb mitjana μ. Es vol estudiar la probabilitat que aquesta component falli (mori) abans de 10 hores. Per a això prenem $n = 30$ components electròniques i anotem llurs vides: $X_1, \ldots, X_{30}$.

a) Calculeu la probabilitat $q = q(\mu)$ que una component electrònica falli abans de 10 hores.

b) Estimeu q pel mètode dels moments. Doneu una interpretació a l'estimador obtingut.

c) Cerqueu un estadístic suficient per aquesta família. Raoneu la resposta.

d) Trobeu un estimador U.M.V.U. per a q.

1.55 PROBLEMA D'EXAMEN

Considerem una mostra $X_1, \ldots, X_n$ d'una llei que té per funció de densitat

$$f(x; \alpha, \beta) = \begin{cases} \frac{1}{\alpha} \exp^{-\frac{x-\beta}{\alpha}} & x \geq \beta \\ 0 & \text{altrament} \end{cases}$$

i l'espai de paràmetres és $\Omega = \{(\alpha, \beta) : \alpha > 0, \beta > 0\}$.

Indicació: Molts dels apartats d'aquest problema poden fer–se amb independència dels altres.

a) Escriviu la funció de versemblança $L(\alpha, \beta)$ donat els valors mostrals $x_1, \ldots, x_n$.

b) Suposant $\alpha = \alpha_0$ conegut,

1) Cerqueu l'estimador de màxima versemblança de β i anomeneu–lo $\hat{\beta}$.

2) Caracteritzeu la llei de $\hat{\beta}$.

c) Suposant $\beta = \beta_0$ conegut,

1) Cerqueu un estimador de α pel mètode de la màxima versemblança i anomeneu–lo $\hat{\alpha}$.

2) Quina distribució segueix $\hat{\alpha}$ quan n és prou gran?

3) Calculeu la informació de Fisher pel paràmetre α continguda en la mostra $X_1, \ldots, X_n$.

4) Trobeu un estimador U.M.V.U. per a α.

5) És l'estimador $\hat{\alpha}$ eficient?

d) Suposeu que $(\hat{\alpha}, \hat{\beta})$ és el màxim de versemblança conjuntament de (α, β).

1) Calculeu la mediana poblacional en funció de α i β i anomeneu–la M.

2) Calculeu l'estimador de màxima versemblança de M i anomeneu–lo $\hat{M}$.

1.56 PROBLEMA D'EXAMEN

Sigui X una variable aleatòria d'una població que es distribueix segons una llei de Weibull amb paràmetre d'escala α desconegut i paràmetre de forma $\beta = 3$.

1) Demostreu que aquesta llei pertany a la família exponencial. Identifiqueu les funcions $c(\alpha)$ i $d(\alpha)$.

2) Doneu un estadístic suficient per a α.

3) Doneu la forma natural de la família i anomeneu $\eta = c(\alpha)$. Identifiqueu la funció $d_0(\eta)$.

4) Considerem una mostra $X_1, \ldots, X_n$ de X.

   a) Estimeu $\frac{1}{\eta}$ pel mètode dels moments. Anomeneu $\hat{r}$ el corresponent estimador.

   b) Calculeu la variància de $\hat{r}$.

   Indicació: Utilitzeu les propietats de la funció $d_0(\eta)$.

1.57 PROBLEMA D'EXAMEN

Disposem de n variables aleatòries, $Y_1, \ldots, Y_n$ independents i uniformement distribuïdes a l'interval $[0, \theta]$. Imaginem que enviem un observador cada dia, durant n dies successius, amb l'encàrrec de que observi, i anoti, els valors de les Y_j per a un subconjunt de m ($1 \leq m \leq n$) d'elles; és a dir, anotarà els m valors $Y_{j_1}, \ldots, Y_{j_m}$, on $j_1, \ldots, j_m \in \{1, \ldots, n\}$. Aquest observador, amb bons coneixements d'estadística, decideix anotar tan sols el màxim de $Y_{j_1}, \ldots, Y_{j_m}$.

a) Per quina raó decideix l'observador anotar tan sols el màxim de $Y_{j_1}, \ldots, Y_{j_m}$.?

b) Sigui X_i el màxim calculat el dia i. Anomenem m el número d'observacions d'entre les observacions $Y_j, 1 \leq j \leq n$ que ha escollit. Les n variables aleatòries, $X_1, \ldots, X_n$, descrites més amunt són independents i tenen per funció de densitat (no cal que ho demostreu)

$$f(x_1; m, \theta) = \begin{cases} \theta^{-m} m x_1^{m-1} & \text{si } 0 < x_1 \leq \theta \\ 0 & \text{altrament} \end{cases}$$

on $m > 0, \theta > 0$.

b1) Cerqueu l'estimador de màxima versemblanca per θ suposant que l'observador ens ha dit el valor de m. Anomeneu-lo $\hat{\theta}$.

b2) Per desgràcia aquest observador és molt despistat i s'oblida de dir-nos el valor de m, així que ara m també és desconegut. Cerqueu l'estimador de màxima versemblanca per m suposant que el valor de θ és conegut. Anomeneu-lo $\hat{m}$.

b3) Sabent que $(\hat{m}, \hat{\theta})$ és el màxim de versemblança de (m, θ), calculeu el màxim de versemblança de $\mathrm{Prob}\{X_1 > 1\}$. Raoneu la resposta.

NOTA: L'apartat B es pot fer independentment del A usant la funció de densitat $f(x_1; m, \theta)$ per a la mostra $X_1, \ldots, X_n$.

2. PROVES D'HIPÒTESIS I INTERVALS DE CONFIANÇA

2.1 Introducció. Conceptes bàsics en la teoría de les proves d'hipòtesis.

2.2 Proves d'hipòtesis òptimes. El lema de Neyman–Pearson.

2.3 Proves d'hipòtesis U.M.P.La prova de la raó de versemblança.

2.4 La distribució t. Aplicació a un problema de regressió. La distribució t descentrada.

2.5 Prova de la raó de versemblança generalitzada. La prova de la t de Student.

2.6 La distribució F. La prova de la F.

2.7 Comparació de dues mostres independents. La prova de la t per dues mostres.

2.8 Intervals de confiança.

EXERCICIS:

2.1 Suposeu que la proporció p de items defectuosos en una gran població és desconeguda i que volem realitzar la següent prova d'hipòtesi:
$H_0 : p = 0.2$
$H_A : p \neq 0.2$.
Suposeu que s'ha agafat una mostra de grandària 20 d'aquesta població. Denoteu per X el nombre d'items defectuosos en la mostra, i considereu un procediment δ que rebutja quan $X \geq 7$ o $X \leq 1$.

a) Calculeu la probabilitat d'un error de tipus I.

b) Calculeu la potència en els punts $p = 0, 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, 1$. Feu una gràfica de la funció de potència en funció de p.

c) És aquest procediment no esbiaixat?

2.2 Suposeu que la variable aleatòria X segueix una distribució binomial de paràmetres $n = 100$ i p desconegut. Considereu el procediment que rebutja $H_0 : p = 0.5$ en favor de $H_A : p \neq 0.5$ quan $| X - 50 |> 10$. Utilitzeu l'aproximació normal a la binomial amb la correcció per continuitat per a fer tots els càlculs.

a) Calculeu la probabilitat d'un error de tipus I.

b) Feu una gràfica de la funció de potència en funció de p.

2.3 Suposeu que es pren una observació X d'una distribució uniforme en l'interval $[\theta - 0.5, \theta + 0.5]$, i suposeu que voleu resoldre la següent prova d'hipòtesi:
$H_0 : \theta \leq 3$
$H_A : \theta \geq 4$.
Doneu un procediment que tingui una funció de potència que prengui els següents valors: $\pi(\theta) = 0$ per $\theta \leq 3$ i $\pi(\theta) = 1$ per $\theta \geq 4$.

2.4 La v.a. X segueix una de les següents distribucions:

X	H_1	H_2
x_1	0.2	0.1
x_2	0.3	0.4
x_3	0.3	0.1
x_4	0.2	0.4

a) Compareu la raó de versemblança per a cada valor possible de X i ordeneu els valors de X d'acord amb aquesta raó.

b) Utilitzeu el mètode de la raó de versemblança de H_1 versus H_2 a un nivell $\alpha = 0.2$. Determineu la regió de rebuig també per un $\alpha = 0.5$.

2.5 Sigui $X_1, \ldots, X_n$ una mostra de la distribució de Poisson.

a) Utilitzeu el mètode de la raó de versemblança per a contrastar $H_0 : \lambda = \lambda_0$ versus $H_1 : \lambda = \lambda_1$, a on $\lambda_1 > \lambda_0$. Indicació: La suma de v.a. independents de Poisson segueix una distribució de Poisson.

b) Demostreu que el procediment trobat a l'apartat a) és uniforment més potent per a contrastar $H_0 : \lambda = \lambda_0$, versus $H_A : \lambda > \lambda_0$.

c) Demostreu que el valor de $\alpha(\delta) + \beta(\delta)$ el minimitza un procediment δ que rebutja H_0 quan $\bar{X}_n > c$. Trobeu el valor de la constant c i feu els càlculs per $n = 20$, $\lambda_0 = 0.25$ i $\lambda_1 = 0.5$.

2.6 Sigui $X_1, \ldots, X_n$ una mostra de la distribució uniforme en l'interval $(0, \theta)$. Volem resoldre $H_0 : \theta \geq 2$ versus $H_1 : \theta < 2$. Sigui $Y_n = \max(X_1, \ldots, X_n)$ i considerem aquell procediment que té per regió crítica tots aquells resultats tals que $Y_n \leq 1.5$.

a) Determineu la funció de potència d'aquest procediment.

b) Determineu la talla del procediment.

2.7 Sigui $X_1, \ldots, X_{25}$ una mostra de la distribució normal amb mitjana μ desconeguda i variància igual a 100.

a) Trobeu la regió de rebuig per a un procediment amb $\alpha = 0.10$ de $H_0 : \mu = 0$ versus $H_A : \mu = 1.5$.

b) Quina és la potència d'aquest procediment?

c) Repetiu el mateix per $\alpha = 0.01$.

d) Feu la prova de la raó de versemblança generalitzada per a contrastar $H_0 : \mu = 0$ versus $H_A : \mu \neq 0$ per a $\alpha = 0.10$.

e) Feu una gràfica de la funció de potència en funció de μ.

f) Repetiu el mateix per $\alpha = 0.05$.

g) Repetiu el mateix per $n = 100$.

2.8 Considerem les següents funcions de densitat:

$$f_0(x) = \begin{cases} 1 & \text{per } 0 \leq x \leq 1 \\ 0 & \text{altrament} \end{cases}$$

$$f_1(x) = \begin{cases} 2x & \text{per } 0 \leq x \leq 1 \\ 0 & \text{altrament} \end{cases}$$

Prenem una observació d'una distribució que té per densitat $f(x)$ igual a $f_0(x)$ o $f_1(x)$ i volem resoldre:
$H_0 : f(x) = f_0(x)$
$H_1 : f(x) = f_1(x)$.
Nota: Seguim la notació utilitzada en el lema de Neyman–Pearson.

a) Descriure un procediment δ_1 pel qual el valor $\alpha(\delta_1) + 2\beta(\delta_1)$ és un mínim.

b) Determineu el valor de $\alpha(\delta_1) + 2\beta(\delta_1)$ per aquest procediment.

c) Descriure un procediment δ_2 pel qual el valor $3\alpha(\delta_2) + \beta(\delta_2)$ és un mínim.

d) Determineu el valor de $3\alpha(\delta_2) + \beta(\delta_2)$ per aquest procediment.

e) Descriure un procediment δ_3 pel qual el valor $\alpha(\delta_3) \leq 0.1$ i $\beta(\delta_3)$ és un mínim.

f) Determineu el valor de $\beta(\delta_3)$ per aquest procediment.

2.9 Suposeu que $X_1, \ldots, X_n$ és una mostra d'una distribució de Poisson de paràmetre λ desconegut ($\lambda > 0$).

a) Demostreu que la funció de probabilitat conjunta de $(X_1, \ldots, X_n)$ té una raó de versemblança monòtona en l'estadístic $\sum_{i=1}^{n} X_i$.

b) Demostreu que per $n = 10$ existeix un procediment U.M.P per la prova d'hipòtesi: $H_0 : \lambda \leq 1$ versus $H_A : \lambda > 1$ amb nivell de significació $\alpha_0 = 0.0143$.

c) Demostreu que per $n = 10$ existeix un procediment U.M.P per la prova d'hipòtesi: $H_0 : \lambda \geq 1$ versus $H_A : \lambda < 1$ amb nivell de significació α_0 per alguns α_0 tals que $0 < \alpha_0 < 0.03$.

2.10 Suposeu que $X_1, \ldots, X_n$ és una mostra d'una distribució gamma de paràmetres α ($\alpha > 0$) i β.

a) Demostreu que la funció de probabilitat conjunta de $(X_1, \ldots, X_n)$ quan α és desconegut i β és conegut té una raó de versemblança monòtona en l'estadístic $\prod_{i=1}^{n} X_i$.

b) Suposeu ara que α és conegut i β és desconegut. Cerqueu un estadístic $T(X_1, \ldots, X_n)$ tal que la funció de probabilitat conjunta de $(X_1, \ldots, X_n)$ tingui una raó de versemblança monòtona en l'estadístic $T(X_1, \ldots, X_n)$.

2.11 Suposeu que $X_1, \ldots, X_n$ és una mostra d'una distribució exponencial de paràmetre λ desconegut. Es desitja fer la prova d'hipòtesi: $H_0 : \lambda \geq 0.5$ versus $H_A : \lambda < 0.5$.

a) Demostreu que per qualsevol nivell de significació α_0 ($0 < \alpha_0 < 1$), existeix un procediment U.M.P. que rebutja H_0 quan $\bar{X}_n > c$, per alguna constant c.

b) Trobeu el valor de c en aquesta prova per a $n = 10$.

2.12 Suposeu que $X_1, \ldots, X_n$ és una mostra d'una distribució normal de mitjana μ desconeguda i de variància 1 i es desitja resoldre la prova d'hipòtesi: $H_0 : 0.1 \leq \mu \leq 0.2$ versus $H_1 : \mu < 0.1$ *o* $\mu > 0.2$. Considerem un procediment δ que rebutja l'hipòtesi nul.la si $\bar{X}_n \leq c_1$ o si $\bar{X}_n \geq c_2$, i sigui $\pi(\mu|\delta)$ la funció de potència de δ. Suposeu que $n = 25$.

a) Determineu els valors de c_1 i de c_2 per tal que $\pi(0.1|\delta) = \pi(0.2|\delta) = 0.07$.

b) Determineu els valors de c_1 i de c_2 per tal que $\pi(0.1|\delta) = 0.02$ i $\pi(0.2|\delta) = 0.05$.

2.13 Suposeu que $X_1, \ldots, X_n$ és una mostra d'una distribució uniforme en l'interval $(0, \theta)$, a on el valor de θ és desconegut i es desitja realitzar la prova d'hipòtesi: $H_0 : \theta \leq 3$ versus $H_A : \theta > 3$.

a) Demostreu que per qualsevol nivell de significació α_0 ($0 < \alpha_0 < 1$), existeix un procediment U.M.P. que rebutja H_0 quan $\max\{X_1, \ldots, X_n\} > c$, per alguna constant c.

b) Trobeu el valor de c per cada valor possible de α_0.

c) Per a una grandària n dibuixeu la funció de potència del procediment anterior.

2.14 Suposeu que $X_1, \ldots, X_8$ és una mostra d'una distribució normal de mitjana μ i de variància σ^2 ambdues desconegudes. Volem contrastar $H_0 : \mu = 0$ versus $H_A : \mu \neq 0$. Suposem que les dades mostrals són tals que $\sum_{i=1}^{8} X_i = -11.2$ i $\sum_{i=1}^{8} X_i^2 = 42.7$

a) Resoldre aquesta prova d'hipòtesi per $\alpha = 0.10$.

b) Fem una prova t tal que H_0 es rebutja si l'estadístic

$$U = \frac{\sqrt{n}(\bar{X}_n - \mu_0)}{S}$$

és tal que $U < c_1$ o $U > c_2$, a on $\text{Prob}(U < c_1) = 0.01$ i $\text{Prob}(U > c_2) = 0.09$. Amb les dades de l'apartat a), hauriem de rebutjar H_0?.

2.15 Sigui $X_1, \ldots, X_m$ una mostra d'una distribució exponencial amb tasa de falla θ. Feu un test de la raó de versemblança per a contrastar
$H_0 : \theta = \theta_0$
$H_1 : \theta \neq \theta_0$.
Indicació: La regió de rebuig és de la forma $\{\bar{X} \exp\{-\theta_0 \bar{X}\} \leq c\}$, a on c és una constant donada pel nivell de significació.

2.16 Demostreu que si la v.a. T es distribueix com una t amb n g.ll, aleshores T^2 segueix una $F_{1,n}$.

2.17 Demostreu que si la v.a. X es distribueix com una $F_{n,m}$ aleshores X^{-1} segueix una $F_{m,n}$.

2.18 Demostreu que si les v.a. X i Y són independents i exponencials amb $\lambda = 1$, aleshores $\frac{X}{Y}$ segueix una llei F. Identifiqueu els graus de llibertat.

2.19 Suposeu que les v.a. $X_1, \ldots, X_n$ són independents i que a cada una d'elles li correspón una distribució F_i. Definim

$$Y = -2 \sum_{i=1}^{n} \log F_i(X_i).$$

Demostreu que Y segueix una distribució χ^2 amb 2n g.ll.

2.20 Siguin X_1 i X_2 v.a. independents i $N(0, \sigma^2)$. Calculeu

$$\text{Prob}\left(\frac{(X_1 + X_2)^2}{(X_1 - X_2)^2} < 4\right)$$

2.21 Suposem que $X_1, \ldots, X_{17}$ formen una mostra aleatòria d'una distribució normal de mitjana μ i variància σ^2, ambdues desconegudes. Siguin $\bar{X}_n$ i S^2 la mitjana i la variància mostrals. Trobeu el valor de k tal que

$$\text{Prob}(\bar{X}_n > \mu + kS) = 0.95$$

2.22 Suposem que una v.a. X segueix una distribució F amb 1 i 8 graus de llibertat. Useu la taula de la distribució t per a determinar el valor de c tal que $\text{Prob}(X > c) = 0.3$.

2.23 Suposem que una v.a. X segueix una distribució F amb m i n graus de llibertat ($n > 2$). Demostreu que $E(X) = \frac{n}{n-2}$. Indicació: Trobeu el valor de $E(\frac{1}{Z})$, a on $Z \sim \chi^2_n$.

2.24 Suposem que una v.a. X segueix una distribució F amb m i m graus de llibertat. Calculeu la mediana de X.

2.25 Siguin U_1 i U_2 dues variables aleatòries independents distribuïdes segons lleis gamma de paràmetres (α_1, β) i (α_2, β), respectivament. Demostreu:

a) La variable aleatòria $U = \frac{U_1}{U_1 + U_2}$ segueix una llei beta de paràmetres (α_1, α_2).

b) U és independent de $U_1 + U_2$.

2.26 Suposem que una v.a. X segueix una distribució F amb m i n graus de llibertat. Demostreu que la v.a. $Y = \frac{mX}{mX+n}$ segueix una distribució Beta amb paràmetres $\alpha = m/2$ i $\beta = n/2$.

2.27 Considerem dues poblacions normals amb mitjanes μ_1 i μ_2 i variances σ_1^2 i σ_2^2 desconegudes i suposem que volem fer la següent prova d'hipòtesi: $H_0 : \sigma_1^2 \leq \sigma_2^2$ versus $H_A : \sigma_1^2 > \sigma_2^2$. Suposem que prenem una mostra de grandària 16 de la primera població donant com a resultats $\sum_{i=1}^{16} X_i = 84$ i $\sum_{i=1}^{16} X_i^2 = 563$, i una mostra de grandària 10 de la segona població tal que $\sum_{i=1}^{10} Y_i = 18$ i $\sum_{i=1}^{10} Y_i^2 = 72$

a) Determineu els estimadors de màxima versemblança de σ_1^2 i de σ_2^2.

b) Feu una prova de la F amb $\alpha = 0.05$ i concluiu si H_0 pot o no rebutjar-se.

c) Suposem ara que volem contrastar $H_0 : \sigma_1^2 \leq 3\sigma_2^2$ versus $H_A : \sigma_1^2 > 3\sigma_2^2$. Descriure com dur a terme un test F per aquestes hipòtesis.

2.28 Sigui $X_1, \ldots, X_m$ una mostra d'una distribució normal de mitjana desconeguda μ_1 i de variància desconeguda σ^2. Sigui $Y_1, \ldots, Y_n$ una mostra d'una distribució normal de mitjana desconeguda μ_2 i de variància desconeguda σ^2. Aquestes dues mostres són independents. Donat un valor constant λ $(-\infty < \lambda < \infty)$,

a) Construïu una prova de la t amb $m + n - 2$ graus de llibertat per a: $H_0 : \mu_1 - \mu_2 = \lambda$ versus $H_A : \mu_1 - \mu_2 \neq \lambda$.

Hem donat una certa droga A a 8 malalts seleccionats a l'atzar, i després d'una estona hem mesurat la concentració d'aquesta droga en la sang, obtenint: $1.23, 1.42, 1.41, 1.62, 1.55, 1.51, 1.60, 1,76$. Suposem que administrem una altra droga B a sis malalts diferents i els resultats obtinguts són: $1.76, 1.41, 1.87, 1.49, 1.67$, 1.81.

b) Suposant que totes les observacions provenen d'una distribució normal amb la mateixa variància, feu una prova d'hipòtesi per a saber si podem concloure que la mitjana de la concentració de la droga B és més gran que la de la droga A.

c) Feu servir la part a) per a donar un interval de confiança per a la diferència entre les dues mitjanes de les dues concentracions amb un nivell de confiança del 90%.

2.29 Suposem que hem de fer n mesuraments, $X_1, \ldots, X_n$, sota les condicions d'un tractament i uns altres n mesuraments, $Y_1, \ldots, Y_n$, independents dels primers sota condicions de control. Se suposa que la desviació estàndar d'una observació és aproximadament 10 sota qualsevol de les condicions.

a) Calculeu el valor de n per tal que un interval de confiança al 95% per $\mu_X - \mu_Y$ tingui amplada igual a 2. Indicació: Useu l'aproximació normal sense problemes perquè el valor de n resultarà bastant gran.

b) Calculeu el valor de n per tal que una prova de $H_0 : \mu_X = \mu_Y$ versus $H_1 : \mu_X > \mu_Y$ tingui potència igual a 0.5 si $\mu_X - \mu_Y = 2$ i $\alpha = 0.10$.

c) Considereu una prova bilateral per $H_0 : \mu_X = \mu_Y$. Compareu les següents corbes de potència: i) per $\alpha = 0.05$ i $n = 20$; ii) per $\alpha = 0.10$ i $n = 20$; iii) per $\alpha = 0.05$ i $n = 40$; iv) per $\alpha = 0.10$ i $n = 40$.

2.30 El nombre X de vegades que es pot encendre i apagar un endoll fins que falla segueix una distribució geomètrica amb paràmetre p. Donada una mostra aleatòria de 10 endolls i donat que la suma de vegades del total dels 10 endolls que es van poguer encendre i apagar fins que van fallar va ser de 15169 vegades, rebutjarieu l'hipòtesi $H_0 : p = 0.00005$ en favor de l'alternativa $H_1 : p > 0.00005$ amb un nivell de significació del 0.05? Indicació: Feu servir el teorema 3.5.

2.31 PROBLEMA D'EXAMEN

a) Enuncieu i demostreu, en el cas discret, el Lema de Neymann-Pearson. Utilitzeu una notació clara i definiu tot els conceptes que utilitzeu.

b) Suposeu que una observació X s'ha d'agafar d'una distribució amb funció de densitat:

$$f(x;\theta) = \begin{cases} 2(1-\theta)x + \theta & \text{per } 0 \leq x \leq 1 \\ 0 & \text{altrament,} \end{cases}$$

a on el valor del paràmetre θ és desconegut ($0 \leq \theta \leq 2$). Suposem que volem resoldre el següent contrast:

$$H_0 : \theta = 2,$$
$$H_1 : \theta = 0.$$

Nota: Els tres apartats que ara segueixen es poden resoldre independentment.

b1) Determineu el procediment δ_1 tal que $\alpha(\delta_1) + 2\beta(\delta_1)$ és un mínim, i calculeu aquest valor mínim.

b2) Suposeu que el valor de α ha sigut donat ($0 < \alpha < 1$). Determineu el procediment δ_2 tal que $\beta(\delta_2)$ és un mínim, i calculeu aquest valor mínim.

b3) Suposem ara que volem resoldre el següent contrast:

$$H_0 : \theta \geq 1,$$
$$H_1 : \theta < 1.$$

Determineu la funció de potència del procediment δ_3 que rebutja l'hipòtesi nul·la si $X > 0.9$. Quina és la talla d'aquest procediment.?

2.32 PROBLEMA D'EXAMEN

Sigui $X_1, \ldots, X_n$ una mostra d'una població normal de mitjana μ desconeguda i de variança σ^2 desconeguda. Desenvolupeu la prova de la t de Student per a contrastar

$$H_0 : \mu \geq \mu_0,$$
$$H_1 : \mu < \mu_0.$$

mitjançant el procediment de la raó de versemblança. Discutiu quines són les propietats de la prova t de Student.

2.33 PROBLEMA D'EXAMEN

Sigui $X_1, \ldots, X_n$ una mostra d'una població normal de mitjana μ_X desconeguda i de variança σ_X^2 desconeguda i sigui $Y_1, \ldots, Y_n$ una mostra d'una població normal de mitjana μ_Y desconeguda i de variança σ_Y^2 desconeguda. Sigui $\Delta > 0$ un valor conegut.

Desenvolupeu la prova de la t de Student per a contrastar

$$H_0 : \mu_X = \mu_Y + \Delta,$$
$$H_1 : \mu_X \neq \mu_Y + \Delta.$$

Especifiqueu l'estadístic que utilitzeu, la distribució nul·la d'aquest estadístic, la regió de rebuig,i doneu expressions per a la talla de la prova i per a la potència de la prova.

2.34 PROBLEMA D'EXAMEN

Suposem que $Y_1, Y_2, Y_3, Y_4, Y_5, Y_6$ són variables aleatòries independents amb distribució normal estàndar. Determineu una constant k tal que la variable aleatòria

$$\frac{k(Y_1 + Y_2)}{(Y_3^2 + Y_4^2 + Y_5^2 + Y_6^2)^{1/2}}$$

segueixi una distribució t de Student.

2.35 PROBLEMA D'EXAMEN

La v.a. X segueix una distribució exponencial de paràmetre $\lambda = 1$ o $\lambda = 2$. Volem contrastar amb una mostra de mida 1 les hipòtesis $H_0 : \lambda = 1$ contra $H_1 : \lambda = 2$.

(1) Calculeu les probabilitats dels errors de tipus I i de tipus II quan la regió d'acceptació és $A_0 = \{x \leq 1\}$.

(2) Calculeu les probabilitats dels errors de tipus I i de tipus II quan la regió d'acceptació és $A_1 = \{x \geq 0.07\}$.

(3) Quina de les dues proves definides per les regions d'acceptació dels apartats 1 i 2 triarieu?

(4) Trobeu el valor de k per tal que la suma dels errors sigui mínima quan la regió d'acceptació és $A_k = \{x \geq k\}$.

2.36 PROBLEMA D'EXAMEN

(a) Siguin U i V dues v.a. independents distribuïdes segons una llei normal estàndar. Demostreu que U/V segueix una distribució t de Student amb un grau de llibertat.

(b) Siguin U i V dues v.a. independents distribuïdes segons una llei normal amb mitjana 0 i variància σ^2. Demostreu que les v.a. $U + V$ i $U - V$ són independents.

(c) Siguin U i V dues v.a. independents distribuïdes segons una llei normal amb mitjana 0 i variància σ^2. Demostreu que $\frac{U+V}{U-V}$ segueix una distribució t de Student amb un grau de llibertat.

Indicació: La funció de densitat corresponent a una t de Student amb ν graus de llibertat és:

$$f(x;\nu) = \frac{\Gamma(\frac{\nu+1}{2})}{(\nu\pi)^{1/2}\Gamma(\frac{\nu}{2})}\left(1+\frac{x^2}{\nu}\right)^{-(\nu+1)/2} \qquad -\infty < x < \infty$$

Nota: Els tres apartats poden fer-se de forma independent i cada apartat pot utilitzar els resultats dels apartats anteriors encara que aquests no s'hagin demostrat.

2.37 PROBLEMA D'EXAMEN

Sigui $X_1, \ldots, X_n$ una mostra d'una distribució que té per funció de densitat::

$$f(x;\theta) = \theta x^{\theta-1} \qquad 0 < x < 1$$

a on el paràmetre $\theta > 0$ és desconegut. Es desitja saber si es pot concloure que $\theta > 1$ basant-nos en una mostra de grandària 8.

(a) Plantejeu formalment el problema.

(b) Demostreu que la funció de densitat conjunta té una raó de versemblança monòtona en l'estadístic T. Precisseu l'expressió d'aquest estadístic.

(c) Demostreu que existeix una prova uniformement més potent (UMP) i determineu genèricament la regió de rebuig d'aquesta prova.

(d) Es desitja concretar la regió de rebuig per un nivell de significació $\alpha_0 = 0.05$. Indicació: Sigui X una v.a. distribuïda segons una llei gamma de paràmetres (n,t). Cal utilitzar la següent relació:

$$\text{Prob}(X \leq a) = \text{Prob}(Y \geq n)$$

on Y és una v.a. distribuïda segons una llei de Poisson de paràmetre at, i utilitzar les taules de la Poisson.

2.38 PROBLEMA D'EXAMEN

La v.a. X segueix una distribució exponencial amb mitjana $\theta = 1$ o $\theta = 0.5$. Volem contrastar amb una mostra de mida 1 les hipòtesis $H_0 : \theta = 1$ contra $H_1 : \theta = 0.5$.

(a) Calculeu les probabilitats dels errors de tipus I i de tipus II quan la regió de rebuig és $R_0 = \{x \geq 1\}$.

(b) Calculeu les probabilitats dels errors de tipus I i la funció de potència quan la regió de rebuig és $R_1 = \{x \leq 0.07\}$.

c) Quina de les dues proves definides per les regions de rebuig dels apartats 1 i 2 triarieu?

(d) Trobeu el valor de x per tal que la suma dels errors sigui mínima quan la regió de rebuig és $R_x = \{x \leq x\}$.

2.39 PROBLEMA D'EXAMEN

Donada una v.a. X amb funció de densitat

$$f(x;\theta) = \begin{cases} \frac{1}{6\theta^4} x^3 e^{-x} & \text{si } x > 0 \\ 0 & \text{si } x \leq 0 \end{cases},$$

es desitja realitzar la següent prova d'hipòtesis:

$$H_0 : \theta = 2$$
$$H_1 : \theta = 3$$

basant-se en una mostra $X_1, \ldots, X_n$ d'aquesta població.

a1) Determineu un procediment δ que minimitzi $a\alpha(\delta) + b\beta(\delta)$.

a2) Calculeu el valor de $\alpha(\delta)$ per $a = 11$, $b = 2$, $n = 3$.

b) Es desitja ara realitzar la següent prova d'hipòtesis:

$$H_0 : \theta \geq 2$$
$$H_1 : \theta < 2$$

Determineu un procediment UMP amb nivell de significació $\alpha_0 = 0.05$ per $n = 3$.

2.40 PROBLEMA D'EXAMEN

Quan les morts a causa d'una malaltia rara, però no contagiosa, esdevenen aleatòriament i uniformement dins d'una població, el nombre de morts en una regió amb població igual a P (és a dir, amb un nombre d'habitants igual a P) es distribueix segons una llei de Poisson de mitjana μP.

a) Suposem que el nombre de morts observades en n regions amb poblacions $P_1, P_2, \ldots, P_n$ són $Y_1, Y_2, \ldots, Y_n$. Deriveu una expressió pel màxim de versemblança de μ. Anomeneu-lo $\hat{\mu}$.

b) La següent taula dóna el nombre d'homes morts de càncer de fetge (el càncer de fetge es pot considerar una malaltia rara) durant 4 anys a les regions d'Ontario.

Regions i	P_i	Morts	Regions i	P_i	Morts
Regió 1	423.447	37	Regió 6	242.810	14
Regió 2	175.685	11	Regió 7	213.591	16
Regió 3	1.245.379	72	Regió 8	166.045	9
Regió 4	413.465	40	Regió 9	265.880	15
Regió 5	216.476	12	Regió 10	116.371	12

Trobeu $\hat{\mu}$ per a aquestes dades i calculeu el nombre de morts esperades a cada regió. Creieu que aquestes dades són consistents amb les suposicions fetes a l'apartat a)?.

c) Suposem ara que $Y_1, Y_2, \ldots, Y_n$ són variables aleatòries de Poisson de mitjanes $\lambda_1, \lambda_2, \ldots, \lambda_n$ i siguin $P_1, P_2, \ldots, P_n$ constants positives conegudes. Considereu la hipòtesi

$$H_0 : \lambda_1 = \mu P_1, \lambda_2 = \mu P_2, \ldots, \lambda_n = \mu P_n$$

on μ és desconegut. Demostreu que la prova de la raó de versemblança generalitzada per contrastar aquesta hipòtesi és

$$-2 \log \Lambda = 2 \sum_{i=1}^{n} Y_i \log(Y_i / \hat{\lambda}_i)$$

on $\hat{\lambda}_i = \hat{\mu} P_i$ i $\hat{\mu}$ és l'estimador de màxima versemblança trobat a l'apartat a).

d) Apliqueu aquest últim resultat per concloure si les taxes de mort (proporció de morts a cada població) de les 10 regions són proporcionals a les poblacions de les regions.

2.41 PROBLEMA D'EXAMEN

a) Suposeu que una observació X s'ha d'agafar d'una distribució amb funció de densitat:

$$f(x;\theta) = \begin{cases} 2(1-\theta)x + \theta & \text{per } 0 \le x \le 1 \\ 0 & \text{altrament,} \end{cases}$$

a on el valor del paràmetre θ és desconegut ($0 \le \theta \le 2$). Suposem que volem resoldre el següent contrast:

$$\begin{aligned} H_0 &: \theta = 2, \\ H_1 &: \theta = 0. \end{aligned}$$

Nota: Els tres apartats que ara segueixen es poden resoldre independentment.

a1) Determineu el procediment δ_1 tal que $2\alpha(\delta_1) + \beta(\delta_1)$ és un mínim, i calculeu aquest valor mínim.

a2) Suposeu que el valor de α ha sigut donat ($0 < \alpha < 1$). Determineu el procediment δ_2 tal que $\beta(\delta_2)$ és un mínim, i calculeu aquest valor mínim.

a3) Suposem ara que volem resoldre el següent contrast:

$$\begin{aligned} H_0 &: \theta \ge 1, \\ H_1 &: \theta < 1. \end{aligned}$$

Determineu la funció de potència del procediment δ_3 que rebutja l'hipòtesi nul.la si $X > 0.9$. Quina és la talla d'aquest procediment?.

2.42 PROBLEMA D'EXAMEN

Les errades de mesurament d'un aparell segueixen una distribució normal de mitjana 0 i de variància σ^2. Fem n mesuraments i volem resoldre la prova

$$\begin{aligned} H_0 &: \sigma^2 = 1 \\ H_1 &: \sigma^2 = 2 \end{aligned}$$

a) Descriure un procediment δ_1 pel qual el valor de $2\alpha(\delta_1) + 4\beta(\delta_1)$ sigui un mínim.

b) Expliqueu el significat de $2\alpha(\delta_1) + 4\beta(\delta_1)$ i calculeu-lo deixant-lo indicat en funció de n i de la llei que resulti adequada.

c) S'han fet 7 mesuraments i s'han detectat els següents errors:

$$0.3 \quad 0.7 \quad -1.1 \quad 2.0 \quad 1.7 \quad -0.8 \quad -0.5.$$

Expliciteu el procediment δ_1 i el valor de $2\alpha(\delta_1) + 4\beta(\delta_1)$.

d) Cerqueu un procediment U.M.P., amb nivell de significació 0.025, anomeneu-lo δ_2, per a resoldre la prova

$$\begin{aligned} H_0 &: \sigma^2 \ge 1 \\ H_1 &: \sigma^2 < 1 \end{aligned}$$

i expliciteu-lo pels mesuraments de l'apartat c.

3. PROVES PER A LA VALIDESA D'UN MODEL

3.1 La prova de la raó de versemblança per a distribucions multinomials.

3.2 La prova de χ^2.

3.3 Prova de Kolmogorov-Smirnov.

3.4 Papers i gràfics de probabilitat.

EXERCICIS:

3.1 La distribució de Poisson es pot utilitzar com a model de tràfic quan aquest és lleuger, ja que es basaria en el fet que si la taxa d'arribada és aproximadamaent constant i el tràfic no és gaire dens (és a dir, els cotxes poden circular independentment un dels altres), la distribució del nombre de cotxes en un interval de temps donat, o en una area donada, és aproximadament Poisson. A continuació donem una taula que mostra el nombre de girs a la dreta durant 300 intervals de 3 minuts a un encreuament donat. Aquests 300 intervals estan distribuïts durant vàries hores del dia i varis dies de la setmana. Ajusteu una distribució de Poisson i feu una prova d'ajustament usant l'estadístic χ^2 de Pearson.

n	0	1	2	3	4	5	6	7	8	9	10	11	12	13+
frequència	14	30	36	68	43	43	30	14	10	6	4	1	1	0

3.2 En un estudi ecològic sobre el comportament alimentici dels ocells, es va comptar el nombre de salts (n) entre vols per a uns quants ocells. A partir de les dades següents, ajusteu la distribució geomètrica, trobeu un interval de confiança per a p i feu una prova d'ajustament.

n	1	2	3	4	5	6	7	8	9	10	11	12
frequència	48	31	20	9	6	5	4	2	1	1	2	1

3.3 La següent taula dóna el número de morts degut a caigudes accidentals per a cada mes de l'any 1970. Trobeu alguna evidència per a pensar que el número de morts per mes no es distribueix uniformement? Trobeu alguna tendència estacional?

mes	núm. morts
Gener	1668
Febrer	1407
Març	1370
Abril	1309
Maig	1341
Juny	1338
Juliol	1406
Agost	1446
Setembre	1332
Octubre	1363
Novembre	1410
Decembre	1526

3.4 Considereu una prova d'ajustament per a una distribució multinomial amb dues classes. Denotem el número d'observacions a cada classe per X_1 i X_2 i les probabilitats de cada classe per p_1 i p_2. L'estadístic χ^2 de Pearson ve donat per $Q = \sum_{i=1}^{2} \frac{(X_i - np_i)^2}{np_i}$. Demostreu que

$$Q = \frac{(X_1 - np_1)^2}{np_1(1-p_1)}$$

i que sota $H_0 : p_1 = p_1^0$, $Q \sim \chi_1^2$.

3.5 Al 1965, un diari va publicar una història sobre un estudiant que afirmava que havia llençat una moneda 17950 vegades i que havia obtingut 9207 cares i 8743 creus.

a) Representa aquest resultat una discrepància important sobre la hipòtesi $H_0 : p = 0.5$?

b) Un estadístic es va posar en contacte amb ell i li va preguntar que com havia fet l'experiment. L'estudiant per estalviar-se temps havia fet grups de 5 monedes i havia registrat els següents resultats:

número cares	freqüència
0	100
1	524
2	1080
3	1126
4	655
5	105

Són aquestes dades consistents amb l'hipòtesi que les monedes eren equilibrades?

c) Són aquestes dades consistents amb l'hipòtesi que les 5 monedes tenien la mateixa probabilitat de cares encara que aquesta no fos necessàriament 0.5?

3.6 En una mostra aleatòria de 1800 valors d'una distribució uniforme a l'interval $(0, 1)$, es va trobar que 391 valors estaven entre 0 i 0.2; 490 valors entre 0.2 i 0.5; 580 valors entre 0.5 i 0.8; i 339 valors entre 0.8 i 1. Contrasteu l'hipòtesi nulla H_0: La mostra prové d'una $U(0, 1)$, mitjançant una prova d'ajustament de χ^2 amb probabilitat d'un error de tipus I igual a 0.01.

3.7 La distribució de les alçades dels homes d'una gran ciutat se suposa que es distribueix normalment amb mitjana 173 cm. i desviació estàndar 2.54 cm. Vam agafar una mostra de 500 homes d'aquesta ciutat i vam obtenir la següent distribució: 18 homes mesuraven menys de 168 cm.; 177 homes mesuraven entre 168 cm i 171 cm.; 198 homes mesuraven entre 171 cm i 174 cm.; 102 homes mesuraven entre 174 cm i 178 cm.; 5 homes mesuraven més de 178 cm.

a) Penseu que aquesta mostra és representativa dels homes que viuen en aquesta ciutat en el que fa referència a l'alçada?.

b) Abans d'haver agrupat en intervals aquestes alçades es va trobar que la mitjana d'aquests 500 homes era $\bar{X}_n = 171.7$ cm i la variança mostral $\frac{n}{n-1}S^2 = 2.54cm^2$. Torneu a contestar la pregunta de l'apartat a) fent servir ara la nova informació.

3.8 La següent taula dóna 50 valors. Feu servir la prova de Kolmogorov-Smirnov per a contrastar que:

a) Les 50 dades són una mostra d'una distribució normal de mitjana 26 i de variança 4.

b) Les 50 dades són una mostra d'una distribució normal de mitjana 24 i de variança 4.

25.088	26.615	25.468	27.453	23.845	24.466
25.996	26.516	28.240	25.980	30.432	25.153
26.560	25.844	26.964	23.382	25.282	25.893
24.432	23.593	24.644	26.849	26.801	26.796
26.303	23.016	27.378	25.351	23.601	24.743
24.317	29.778	29.585	22.147	28.352	
29.263	27.924	21.579	25.320	28.129	
28.478	23.896	26.020	23.750	24.904	
24.078	27.228	27.433	23.341	28.923	

Mostra 1, Dades exercici 3.8

3.9 Se seleccionen 25 observacions (Mostra 1) a l'atzar de la població 1 amb distribució desconeguda F. També seleccionem 20 observacions (Mostra 2) de la població 2 amb funció de distribució desconeguda G. Feu servir la prova de Kolmogorov-Smirnov per a contrastar que:

a) Les dues poblacions tenen la mateixa distribució.

b) Si X denota la v.a. que representa la població 1 i Y la v.a. que representa la població 2, feu la prova d'hipòtesi que les funcions de distribució de $X + 2$ i de Y són idèntiques.

c) Amb la mateixa notació que a l'apartat b) feu la prova d'hipòtesi que les funcions de distribució de X i de $3Y$ són idèntiques.

061	0.29	0.06	0.59	-1.73
-0.74	0.51	-0.56	-0.39	1.64
0.05	-0.06	0.64	-0.82	0.31
1.77	1.09	-1.28	2.36	1.31
1.05	-0.32	-0.40	1.06	-2.47

Mostra 1, Exercici 3.9

2.20	1.66	1.38	0.20	0.36
0.00	0.96	1.56	0.44	1.50
-0.30	0.66	2.31	3.29	-0.27
-0.37	0.38	0.70	0.52	-071

Mostra 2, Exercici 3.9

3.10 Les duracions en hores de 37 components han sigut les següents: 10 15 20 22 32 40 42 46 48 51 60 71 76 87 93 105 112 116 127 131 172 195 207 219 238 260 300 342 382 435 460 490 520 600 630 670 770

a) Feu servir un paper de probabilitat semilogarítmic per a veure si aquestes duracions són una mostra d'una llei exponencial de paràmetre θ. Calculeu amb l'ajuda del gràfic el valor del paràmetre θ .

b) Calculeu ara el valor del paràmetre θ mitjançant el mètode de la màxima versemblança.

c) Feu una prova χ^2 prenent com a valor de θ el trobat a l'apartat b).

d) Feu la prova de la raó de versemblança generalitzada per a decidir si la mostra prové d'una llei exponencial.

e) Discutiu les diferències entre c) i d) i raoneu quina de les dues proves és més adequada.

3.11 Es consideren els cicles de funcionament de 15 relés elèctrics observats fins a fallar, on el criteri de falla consisteix en superar un determinat nivell de resistència al contacte. Els temps de falla observats són: 6.2, 9.0, 10.2, 12.1, 12.6, 14.4, 14.7, 16.1, 18.6, 20.5, 20.6, 23.0, 26.7, 27.6 34.3. Ajusteu un model de Weibull per aquestes dades mitjançant

a) un paper de probabilitat

b) el mètode de la màxima versemblança (i l'ajut del paquet estadístic que preferiu). Calculeu en ambdós casos els paràmetres α i β.

3.12 PROBLEMA D'EXAMEN

a) Sigui $X_1, \ldots, X_n$ una mostra d'una població que es distribueix segons una funció de distribució F contínua i desconeguda. Sigui F^* una funció de distribució donada. Discutiu els diferents procediments que ens permeten saber si les dades queden ben ajustades per F^*. En cada cas expliqueu sota quines condicions utilitzarieu un procediment o un altre.

b) Un grup de 23 rates que s'ha utilitzat amb finalitat experimental té un pes mitjà de $\mu_0 = 370.6$gr i una desviació estàndar $\sigma_0 = 29.1$gr. Les dades són les següents:

356.4 362.5 394.7 356.0 387.6 305.1 385.1 383.2 346.6 314.2 394.8 370.7 370.8 434.2 365.2 377.1 365.9 384.4 297.4 404.3 412.0 349.1 344.5

Utilitzeu la prova de Kolmogorov-Smirnov per a decidir si la distribució dels pes d'aquestes rates és $N(\mu_0, \sigma_0)$.

3.13 PROBLEMA D'EXAMEN

a) Els temps de falla en hores d'una component d'un avió són les següents:

0.88 1.33 1.32 0.22 1.54 1.00 2.50 0.50 3.00 1.76

a1) Creieu que el model de Weibull és apropiat? Raoneu la resposta.

a2) En cas afirmatiu, estimeu el valor dels paràmetres α i β.

a3) Quina proporció de components funcionen correctament més d'una hora?

a4) Quin és el percentil del 10%?

b) Les següents dades descriuen el nivell de colesterol a la sang de dos grups d'homes. El primer grup està format per 11 homes entre 20 i 30 anys i el segon grup està format per 11 homes entre 40 i 50 anys. Les dades són les següents:

x (20–30 anys)	y (40–50 anys)
135	294
222	311
251	286
260	264
269	277
235	336
386	208
252	346
352	239
173	172
156	254

Podeu concloure que el nivell de colesterol és significativament més alt en els homes més grans? Utilitzeu la prova que us sembli més apropiada i raoneu el perquè de la vostra elecció.

3.14 PROBLEMA D'EXAMEN

La següent taula dóna els comptatges observats en intervals d'un segon a l'experiment dut a terme per Berkson sobre l'emissió de partícules α. Es tracta de comprovar si la distribució de Poisson ajusta be aquestes dades. Les dades són les següents:

n	observat
0	5267
1	4436
2	1800

3	534
4	111
5+	21

3.15 PROBLEMA D'EXAMEN

S'experimentà amb 12 rellotges idèntics amb l'intenció de conèixer la distribució del seu temps de vida. Les duracions en mesos fins que van fallar són les següents:
30.5; 33; 33.1; 36; 42; 55; 55.5: 75.8; 76; 106; 106.5; 107.5.

(a) Estudieu gràficament aquestes dades mitjançant un plot de la Weibull.

(b) Creieu que la distribució de Weibull ajusta correctament aquestes dades?

(c) Estimeu gràficament els paràmetres de forma i escala.

(d) Utilitzeu el gràfic per a estimar la mediana de la distribució.

(e) Penseu que seria raonable donar una garantia de 5 anys a aquests rellotges?

3.16 PROBLEMA D'EXAMEN

Les dades següents de Lieblein & Zelen (1956) donen el nombre de revolucions (en milions) fins a la fallada de 23 coixinets de boles. Les dades ordenades de menor a major són:

17.88	28.92	33.00	41.52	42.12	45.60	48.40
51.84	51.96	54.12	55.56	67.80	68.64	68.64
68.88	84.12	93.12	98.64	105.12	105.84	127.92
128.04	173.40					

a) Raoneu si el model exponencial és adequat per a aquesta situació.
Indicació: Compareu la mitjana i la variància mostral.

b) Feu un ajust gràfic amb un paper de Weibull i calculeu el valor dels paràmetres α i β.

c) Empreu el mètode de la màxima versemblança per estimar els paràmetres α i β. Anomeneu $\hat{\alpha}$ i $\hat{\beta}$ els estimadors obtinguts.

d) Estimeu per màxima versemblança $t_{0.2}$, el percentil del 20% d'aquesta distribució. Anomeneu $\hat{t_{0.2}}$ l'estimador obtingut.

e) Suposeu que el valor de β és conegut i igual a 2.1018. Utilitzeu el mètode delta per calcular la variància asimptòtica de $\hat{t_{0.2}}$.

Indicacions: Els següents valors són necessaris per a poder acabar alguns dels càlculs: $\hat{\beta} = 2.1018$, $\sum_{i=1}^{23} t_i^{\hat{\beta}} = 241426$, $\mathrm{Var}(\hat{\alpha}) = 60.42$.

3.17 PROBLEMA D'EXAMEN

Disposem d'una xarxa de 165 cel.les. Es va comptar el nombre de grans de grafit en cada cel.la obtenint els resultats de la següent taula:

Número de grans per cel.la	observats
0	1
1	1
2	5
3	7
4	20

5	34
6	30
7	17
8	22
9	21
10	4
11	2
12	1

Provar la hipòtesi que el nombre de grans per cel.la és una variable aleatòria que segueix una distribució de Poisson.

3.18 PROBLEMA D'EXAMEN

A una certa zona del mediterrani es treballa amb l'hipòtesi que l'alçada de les crestes de les onades (mesurades en metres) segueixen, el mes de desembre, una distribució de Raleigh amb funció de distribució acumulada

$$F(x) = 1 - \exp\left\{\frac{-x^2}{2\gamma^2}\right\}; \quad \gamma = 0.7; 0 < x < +\infty.$$

Per tal de poder analitzar si aquest model havia deixat d'ésser vàlid –cosa que podria ser produïda per un canvi climàtic– es van prendre 10 mostres en dies diferents del mes de desembre i sempre a la mateixa hora, obtenint–se els següents valors:

$$0.91, 0.87, 1.01, 0.76, 0.84, 1.05, 1.12, 0.96, 0.89, 0.57.$$

a) Estudieu gràficament aquestes dades de la forma que us sembli més adequada. Creieu que la distribució de Raleigh ajusta correctament aquestes dades?.

b) Realitzeu una prova de la raó de versemblança generalitzada per decidir si es pot acceptar o no que $\gamma = 0.7$ amb un nivell de significació $\alpha = 0.05$.

3.19 PROBLEMA D'EXAMEN

Una companyia de telèfons fa un experiment pilot amb 12 clients i els hi instala un nou sistema de telecomunicació. Aquests sistemes poden fallar per diversos motius: a) per deixar de funcionar, b) per sentir–se soroll en la transmissió o c) per interferències.

A continuació donem el número de setmanes que triguen en fallar.

$$34 \quad 3 \quad 26 \quad 6 \quad 0.7 \quad 9 \quad 3 \quad 16 \quad 5 \quad 35 \quad 22 \quad 17.$$

Nota: El valor 0.7 indica una falla ocurreguda després de 5 dies.

L'equip d'estadístics de la companyia ha d'explicar a l'equip d'enginyers el comportament de les falles del nou sistema de telecomunicació i ha de respondre a les següents dues qüestions plantejades pels enginyers:

QÜESTIÓ 1: Quantes setmanes de garantia es podríen donar si volguessim que com a màxim el 20% fallés durant la garantia?

QÜESTIÓ 2: Quina es la mitjana i la variància del número de falles?

a) Feu un resum gràfic d'aquestes dades que sigui útil pels enginyers.

b) Feu servir el gràfic anterior per a contestar la primera de les preguntes dels enginyers.

c) Inferiu a partir del gràfic quina és la llei que millor s'ajusta als valors de falles observats i utilitzeu aquesta llei per a contestar la segona pregunta dels enginyers.

4. INTRODUCCIÓ A L'INFERÈNCIA BAYESIANA.

4.1 Lleis de Bayes i Minimax.

4.2 Anàlisi per a trobar una llei de Bayes.

4.3 Estimadors admissibles.

4.4 Inferència Bayesiana per a una distribució normal.

4.5 Inferència Bayesiana per a una distribució binomial.

EXERCICIS:

4.1 Suposem que la proporció θ de components defectuoses d'un gran lot és o bé 0.1 o bé 0.2, i la funció de probabilitat a priori de θ és $\psi(0.1) = 0.7$ i $\psi(0.2) = 0.3$. Es seleccionen 8 components del lot a l'atzar i es troben 2 defectuoses. Determineu la funció de probabilitat a posteriori de θ.

4.2 Suposem que la proporció θ de components defectuoses d'un gran lot és desconeguda, i la distribució a priori de θ és uniforme en l'interval $(0, 1)$. Es seleccionen 8 components del lot a l'atzar i es troben 3 defectuoses. Determineu la llei a posteriori de θ.

4.3 Suposem que tenim una observació X d'una distribució uniforme a l'interval $(\theta - 0.5, \theta + 0.5)$; suposem que el valor de θ és desconegut i que la llei a priori de θ és l'uniforme en l'interval $(10, 20)$.

a) Si el valor observat de X és $x = 12$, quina és la distribució a posteriori de θ?.

b) Si ara seleccionem sis observacions i observem $x_1 = 11.0, x_2 = 11.5, x_3 = 11.7, x_4 = 11.1, x_5 = 11.4, x_1 = 10.9$, quina és la distribució a posteriori de θ?.

4.4 Suposem que la proporció θ de components defectuoses d'un gran lot és desconeguda, i la distribució a priori de θ és beta$(\alpha = 2, \beta = 200)$. Es seleccionen 100 components del lot a l'atzar i es troben 3 defectuoses.

a) Determineu la llei a posteriori de θ.

b) Suposem que després d'observar 3 peces defectuoses la distribució a posteriori és una beta de mitjana .0392 i de variància 3.658×10^{-4}. Determineu la llei a priori de θ.

4.5 Sigui θ el nombre mig de defectes per cada 100 peus d'una cinta magnètica. Se suposa que el valor de θ és desconegut i que la distribució a priori de θ és una gamma de paràmetres $\alpha = 2$ i $\beta = 10$. Després d'inspeccionar una cinta de 1200 peus es troben 4 defectes. Determineu la llei a posteriori de θ.

4.6 Les alçades dels individus d'una certa població segueixen una distribució normal amb mitjana desconeguda θ i amb desviació estàndar 5cm. Se suposa que la distribució a priori de θ és una normal de mitjana 170 cm i de desviació estàndar 2.5cm.

a) La mitjana de 10 persones seleccionades a l'atzar és de 174 cm. Determineu la llei a posteriori de θ.

b) Calculeu l'interval d'amplada 2.5 cm que contingui el valor de θ amb la màxima probabilitat a priori i doneu aquesta probabilitat.

c) Calculeu l'interval d'amplada 2.5 cm que contingui el valor de θ amb la màxima probabilitat a posteriori i doneu aquesta probabilitat.

4.7 El temps (en minuts) que es triga en atendra un client en una botiga segueix una llei exponencial de paràmetre θ desconegut. la distribució a priori de θ és una gamma amb mitjana 0.2 i desviació estàndar igual a 1. Si el temps que s'ha trigat en atendre a 20 clients ha sigut de 3.8 minuts, quina és la distribució a posteriori de θ.?

4.8 Suposem que la proporció θ de components defectuoses d'un gran lot és desconeguda, i la distribució a priori de θ és beta($\alpha = 5, \beta = 10$). Es seleccionen 20 components del lot a l'atzar i es troben 1 sola component defectuosa. Determineu l'estimador de Bayes de θ si fem servir com a funció de pèrdua l'error quadràtic.

4.9 Les alçades dels individus d'una certa població segueixen una distribució normal amb mitjana desconeguda θ i amb desviació estàndar 5 cm. Se suposa que la distribució a priori de θ és una normal de mitjana 170 cm i de desviació estàndar 2.5 cm. La mitjana de 10 persones seleccionades a l'atzar és de 174 cm.

a) Determineu l'estimador de Bayes de θ si fem servir com a funció de pèrdua l'error quadràtic.

b) Determineu l'estimador de Bayes de θ si fem servir com a funció de pèrdua l'error absolut.

4.10 Denotem per θ la proporció de votants d'una gran ciutat que estan d'acord amb una certa llei. Suposem que aquest valor θ és desconegut i que dos estadístics A i B assignen a θ les següents densitats a priori $\psi_A(\theta)$ i $\psi_B(\theta)$, respectivament:

$$\psi_A(\theta) = 2\theta \quad per \quad 0 < \theta < 1$$

$$\psi_B(\theta) = 4\theta^3 \quad per \quad 0 < \theta < 1.$$

En una mostra de 1000 votants hi han 710 en favor de la llei.

a) Determineu la llei a posteriori de θ que cada estadístic assigna a θ.

b) Determineu l'estimador de Bayes de θ que cada estadístic calcularia si fessim servir com a funció de pèrdua l'error quadràtic.

c) Demostreu que després de conèixer les opinions de 1000 votants els estimadors de Bayes proposats pels dos estadístics no poden diferir en més de 0.002 amb independència del nombre de votants que haguessin votat en favor de la llei.

4.11 PROBLEMA D'EXAMEN

Suposem que la v.a. X segueix una llei geomètrica de paràmetre p; és a dir,

$$f(x;p) = (1-p)^{x-1}p \quad x = 1, 2, \ldots$$

Suposem que a p li assignem una llei a priori uniforme a l'interval (0,1).

(a) Calculeu la distribució a posteriori de p.

(b) Quin és l'estimador de Bayes de p si fem servir com a funció de pèrdua l'error quadràtic.

4.12 PROBLEMA D'EXAMEN

Denotem per τ la proporció de nens d'un casal d'estiu que estan d'acord amb anar d'excursió. Suposem que aquest valor τ és desconegut i que els dos monitors A i B assignen a τ les següents densitats a priori $\psi_A(\tau)$ i $\psi_B(\tau)$, respectivament:

$$\psi_A(\tau) = 3\tau^2 \quad per \quad 0 < \tau < 1$$

$$\psi_B(\tau) = 4\tau^3 \quad per \quad 0 < \tau < 1.$$

En aquest s'han apuntat 100 nens i d'aquests 60 volen anar d'excursió. En una mostra de 1000 votants hi han 710 en favor de la llei.

a) Determineu la llei a posteriori de τ que cada monitor assigna a τ.

b) Determineu l'estimador de Bayes de τ que cada monitor calcularia si fessim servir com a funció de pèrdua l'error quadràtic.

c) Demostreu que després de conèixer les preferències dels 100 nens els estimadors de Bayes proposats pels dos estadístics no poden diferir en més de 0.01 amb independència del nombre de nens que prefereixin anar d'excursió.

4.13 PROBLEMA D'EXAMEN

Suposem que el nombre de minuts que una persona ha d'esperar l'autobus cada matí té una distribució uniforme a l'interval $(0, \theta)$, a on el valor de θ és desconegut. Suposem que la funció de densitat a priori de θ ve donada per:

$$\xi(\theta) = \begin{cases} \frac{192}{\theta^4} & \text{per } \theta \geq 4, \\ 0 & \text{altrament} \end{cases} .$$

Si els temps esperats en tres matins successius són de $5, 3, 8$ minuts.

a) Calculeu la funció de densitat a posteriori de θ. Especifiqueu el domini de definició d'aquesta funció i les constants que en ella apareixen.

b) Si es vol estimar el valor de θ usant com a funció de pèrdua l'error quadràtic, quina és la forma de l'estimador de Bayes de θ. Calculeu el valor estimat de θ a partir delss tres temps esperats donats més amunt.

4.14 PROBLEMA D'EXAMEN

Suposem que $X_1, \ldots, X_n$ són variables aleatòries independents que segueixen una llei de Poisson de paràmetre θ i suposem que la llei a priori sobre θ pertany a la família gamma.

a) Demostreu que la llei a priori gamma és conjugada de la Poisson.

b) Determineu l'estimador de Bayes de θ sota una funció de pèrdua quadràtica.

4.15 PROBLEMA D'EXAMEN

La variable aleatòria que compta el nombre de morts per asma cada 100.000 habitants en un any es suposa que segueix una llei de Poisson de paràmetre θ, on θ representa la veritable taxa de mortalitat per asma cada 100.000 habitants.

A la ciutat C dels Estats Units, amb una població de 200.000 persones, es revisen les causes de mort durant l'any 1994, i es denota per $Y94$ la variable aleatòria que compta el nombre de morts per asma a C durant l'any 1994.

a) Quina llei segueix la variable aleatòria $Y94$?.

b) Quina llei *a priori* proposarieu per θ, la taxa de mortalitat per asma? Raoneu la resposta.

c) Sabem que la taxa de mortalitat per asma arreu del mon té mitjana 0.6 i variància 0.12. Basant-vos en la llei *a priori* proposada a b), calculeu el seus paràmetres.

d) Es troba que hi han hagut 3 persones que moriren d'asma a C l'any 1994, (*i.e.*, $Y94 = 3$). Deriveu la llei *a posteriori* per la taxa de mortalitat per asma? En particular, doneu la mitjana *a posteriori* i discutiu aquest resultat comparant–lo amb la mitjana *a priori*.

e) L'any 1995 es repeteix l'estudi i es tornen a donar 3 morts (*i.e.*, $Y95 = 3$). Feu servir els càlculs de l'apartat d) per a calcular la mitjana *a posteriori*? Que podeu concloure?

5. MÈTODES NO PARAMÈTRICS

5.1 La prova dels signes

5.2 La prova dels rangs signats de Wilcoxon

5.3 La prova de Wilcoxon-Mann-Whitney

EXERCICIS:

5.1 Un estudi es realitza amb l'objectiu de comparar coixinets fets de dos tipus (tipus I i tipus II). Es proven 10 coixinets de cada tipus. La següent taula dóna els temps fins que fallen (les unitats són milions de cicles):

tipus I	tipus II	tipus I	tipus II
3.03	3.19	16.84	12.75
5.53	4.26	16.04	9.37
5.60	4.47	15.21	6.79
9.30	4.53	12.95	12.78
9.92	4.67	12.51	4.69

Es tracta de plantejar la hipòtesi que no hi ha diferència entre els dos tipus de coixinets i mirar si la podem rebutjar fent servir:

a) La teoria basada en la suposició que la vida dels coixinets segueix una distribució normal,

b) Una prova noparamètrica. Quin dels dos mètodes és millor?

5.2 Trobeu la distribució nulla exacte de l'estadístic dels rangs signats de Wilcoxon quan n=4.

5.3 Es pretén comparar un mètode microbiològic (Mètode A) i un altre hidroxilamin (Mètode B) per analitzar les dosis d'ampicilina. En una sèrie d'experiments parells de tabletes s'analitzen pels dos mètodes. En la següent taula (dades de Lin, Sutton i Qurashi' 79) donem els percentatges de quantitat d'ampicilina trobada fent servir els dos mètodes. Analitzeu les dades per a determinar si hi ha una diferència sistemàtica entre els dos mètodes.

Mètode A	Mètode B	Mètode A	Mètode B
97.2	97.2	91.0	100.2
105.8	97.8	96.2	99.0
99.5	96.2	96.2	98.0
100.0	101.8	90.8	95.8
93.8	88.0	95.2	94.8
79.2	74.0	20.5	21.2
72.0	75.0	69.5	65.8
72.0	67.5		

5.4 Volem comparar l'efectivitat de dues drogues A i B que serveixen per a reduir la concentració de glucosa a la sang. La droga A s'administra a 25 pacients i la droga B a 15 pacients. Les dades es donen en les taules de més avall. Creieu que les dues drogues són igual d'efectives o pel contrari una és més efectiva que l'altre. Respongueu a aquesta pregunta fent servir tots els tipus de proves que coneixeu i contrasteu els resultats.

0.35	1.12	1.54	0.13	0.77
0.16	1.20	0.40	1.38	0.39
0.58	0.04	0.44	0.75	0.71
1.64	0.49	0.90	0.83	0.28
1.50	1.73	1.15	0.72	0.91

Dades droga A, Exercici 5.4

1.78	1.25	1.01	1.82	1.95
1.81	0.68	1.48	1.59	0.89
0.86	1.63	1.26	1.07	1.31

Dades droga B, Exercici 5.4

5.5 PROBLEMA D'EXAMEN

Volem estudiar l'efecte que té una certa metodologia d'estudi en la població estudiantil. Escollim 9 parelles de bessons idèntics, i decidim que un membre de la parella apliqui el nou mètode d'estudi, mentre que l'altre segueix amb el mètode tradicional. La suma de totes les notes obtingudes per un estudiant que aplica el nou mètode és una v.a. X, i la suma de totes les notes obtingudes per un estudiant que aplica el mètode tradicional és una altra v.a. Y. Els resultats obtinguts a final del curs van ser el següents:

Parella		1	2	3	4	5	6	7	8	9
Mètode Nou	X	128	132	114	212	107	196	175	148	131
Mètode Tradicional	Y	110	118	116	205	96	184	178	152	126

a) Apliqueu el test no paramètric que us sembli més adient per tal de decidir si el nou mètode d'estudi és més eficient.

b) Suposeu que les nostres dades provenen d'una llei Normal. Apliqueu el test de la t-student que creieu adient per tal de decidir si el nou mètode d'estudi és més eficient.

c) Utilitzeu un test no paramètric per tal de contrastar si la millora obtinguda amb el nou mètode es distribueix segons una llei Normal de mitjana 4 i variància 9.

d) A la vista de les conclusions tretes a l'apartat c), quin dels dos procediments, el donat a a) i el donat a b), aplicarieu per decidir quin mètode d'estudi és més eficient.

ESTADISTICA MATEMATICA II
EXAMEN 1
22 d'abril de 1992

PROBLEMA 1

Suposeu que prenem una mostra aleatòria d'una distribució exponencial de mitjana θ desconeguda. Calculeu la grandària de mostra necessària per tal que

$$E\{|\bar{X}_n - \theta|^2\} \leq 0.01$$

PROBLEMA 2

Suposem que (X,Y) és un punt que s'ha d'escollir del pla real i que X i Y indiquen dues variables aleatòries independents amb distribució normal estàndar. Quin ha d'ésser el radi del cercle C (centrat a l'origen) per tal que el punt (X,Y) sigui dins de C amb probabilitat 0.99?

PROBLEMA 3

Suposem que $Y_1, Y_2, Y_3, Y_4, Y_5, Y_6$ són variables aleatòries independents amb distribució normal estàndar. Determineu una constant k tal que la variable aleatòria

$$\frac{k(Y_1+Y_2)}{(Y_3^2+Y_4^2+Y_5^2+Y_6^2)^{1/2}}$$

segueixi una distribució t de Student.

PROBLEMA 4

La distribució de Pareto s'utilitza en ciències econòmiques com a model per a una funció de densitat amb una cua que decreix a poc a poc:

$$f(x;x_0,\theta) = \theta x_0^{\theta} x^{-\theta-1}, \quad x \geq x_0, \quad \theta > 1.$$

Suposem que el valor x_0 és conegut.

a) Calculeu l'esperança d'una v.a. X que segueix una distribució de Pareto. Disposem d'una mostra $X_1, \ldots, X_n$ d'una llei de Pareto.

b) Trobeu l'estimador $\hat{\theta}_1$ de θ pel mètode dels moments.

c) Trobeu l'estimador $\hat{\theta}_2$ de θ pel mètode de la màxima versemblança.

d) Quina llei asimptòtica segueix $\hat{\theta}_2$.? Precisseu en particular la seva variança.

PROBLEMA 5

Disposem d'una mostra $X_1, \ldots, X_n$ d'una llei que té per funció de densitat

$$f(x;\theta) = \frac{\theta}{(1+x)^{\theta+1}}, \quad 0 < \theta < \infty \quad 0 \leq x < \infty.$$

Doneu un estadístic suficient per a θ. Raoneu la resposta.

ESTADISTICA MATEMATICA II
EXAMEN 2
25 de maig de 1992

PART 1

a) Enunicieu i demostreu, en el cas discret, el Lema de Neymann-Pearson. Utilitzeu una notació clara i definiu tot els conceptes que utilitzeu.

b) Suposeu que una observació X s'ha d'agafar d'una distribució amb funció de densitat:

$$f(x;\theta)=\begin{cases}2(1-\theta)x+\theta & \text{per } 0\le x\le 1\\ 0 & \text{altrament,}\end{cases}$$

a on el valor del paràmetre θ és desconegut $(0\le\theta\le 2)$. Suposem que volem resoldre el següent contrast:

$$H_0:\theta=2,$$
$$H_1:\theta=0.$$

Nota: Els tres apartats que ara segueixen es poden resoldre independentment.

b1) Determineu el procediment δ_1 tal que $\alpha(\delta_1)+2\beta(\delta_1)$ és un mínim, i calculeu aquest valor mínim.

b2) Suposeu que el valor de α ha sigut donat $(0<\alpha<1)$. Determineu el procediment δ_2 tal que $\beta(\delta_2)$ és un mínim, i calculeu aquest valor mínim.

b3) Suposem ara que volem resoldre el següent contrast:

$$H_0:\theta\ge 1,$$
$$H_1:\theta<1.$$

Determineu la funció de potència del procediment δ_3 que rebutja l'hipòtesi nulla si $X>0.9$. Quina és la talla d'aquest procediment.?

PART 2

Sigui $X_1,\ldots,X_n$ una mostra d'una població normal de mitjana μ desconeguda i de variança σ^2 desconeguda. Desenvolupeu la prova de la t de Student per a contrastar

$$H_0:\mu\ge\mu_0,$$
$$H_1:\mu<\mu_0.$$

mitjançant el procediment de la raó de versemblança. Discutiu quines són les propietats de la prova t de Student.

PART 3

a) Sigui $X_1,\ldots,X_n$ una mostra d'una població que es distribueix segons una funció de distribució F contínua i desconeguda. Sigui F^* una funció de distribució donada. Discutiu els diferents procediments que ens permeten saber si les dades queden ben ajustades per F^*. En cada cas expliqueu sota quines condicions utilitzarieu un procediment o un altre.

b) Un grup de 23 rates que s'ha utilitzat amb finalitat experimental té un pes mitjà de $\mu_0=370.6$gr i una desviació estàndar $\sigma_0=29.1$gr. Les dades són les següents:

356.4 362.5 394.7 356.0 387.6 305.1 385.1 383.2 346.6 314.2 394.8 370.7 370.8 434.2 365.2 377.1 365.9 384.4 297.4 404.3 412.0 349.1 344.5

Utilitzeu la prova de Kolmogorov-Smirnov per a decidir si la distribució dels pes d'aquestes rates és $N(\mu_0, \sigma_0)$.

PART 4

a) Els temps de falla en hores d'una component d'un avió són les següents:

0.88 1.33 1.32 0.22 1.54 1.00 2.50 0.50 3.00 1.76

a1) Creieu que el model de Weibull és apropiat? Raoneu la resposta.

a2) En cas afirmatiu, estimeu el valor dels paràmetres α i β.

a3) Quina proporció de components funcionen correctament més d'una hora?

a4) Quin és el percentil del 10%?

b) Les següents dades descriuen el nivell de colesterol a la sang de dos grups d'homes. El primer grup està format per 11 homes entre 20 i 30 anys i el segon grup està format per 11 homes entre 40 i 50 anys. Les dades són les següents:

x (20–30 anys)	y (40–50 anys)
135	294
222	311
251	286
260	264
269	277
235	336
386	208
252	346
352	239
173	172
156	254

Podeu concloure que el nivell de colesterol és significativament més alt en els homes més grans? Utilitzeu la prova que us sembli més apropiada i raoneu el perquè de la vostra elecció.

ESTADISTICA MATEMATICA II
EXAMEN FINAL
22 de juny de 1992

PROBLEM 1

La següent taula dóna els comptatges observats en intervals d'un segon a l'experiment dut a terme per Berkson sobre l'emissió de partícules α. Es tracta de comprovar si la distribució de Poisson ajusta be aquestes dades.

Les dades són les següents:

n	observat
0	5267
1	4436
2	1800
3	534
4	111
5+	21

PROBLEM 2

Disposem d'una mostra $X_1, \ldots, X_n$ d'una llei que té per densitat

$$f(x;\theta) = (\theta+1)x^\theta, \quad 0 \leq x \leq 1.$$

a) Calculeu l'esperança i la variança de X.

b) Trobeu l'estimador $\hat{\theta}_1$ de θ pel mètode dels moments.

c) Quina llei asimptòtica segueix $\hat{\theta}_1$.? Precisseu en particular la seva variança. Indicació: Feu servir el Teorema Central del Límit i el mètode Delta.

d) Trobeu l'estimador $\hat{\theta}_2$ de θ pel mètode de la màxima versemblança.

e) Quina llei asimptòtica segueix $\hat{\theta}_2$.? Precisseu en particular la seva variança.

f) Calculeu l'eficiència relativa asimptòtica de $\hat{\theta}_1$. respecte $\hat{\theta}_2$. i deduiu quin dels dos estimadors és més eficient.

PROBLEM 3

a) Enunicieu i demostreu el teorema de Rao-Blackwell.

b) Sigui $X_1, \ldots, X_n$ una mostra d'una llei que té per densitat

$$f(x;\theta) = \frac{1}{6\theta^4}x^3 e^{\frac{-x}{\theta}}, \quad 0 < x < \infty, 0 < \theta < \infty.$$

a) Trobeu un estadístic suficient per a θ.

b) Utilitzeu el teorema de Rao-Blackwell per a trobar un estadístic suficient que ens millori la variança.

PROBLEM 4

Sigui $X_1, \ldots, X_n$ una mostra d'una població normal de mitjana μ_X desconeguda i de variança σ_X^2 desconeguda i sigui $Y_1, \ldots, Y_n$ una mostra d'una població normal de mitjana μ_Y desconeguda i de variança σ_Y^2 desconeguda. Sigui $\Delta > 0$ un valor conegut.

Desenvolupeu la prova de la t de Student per a contrastar

$$H_0 : \mu_X = \mu_Y + \Delta,$$
$$H_1 : \mu_X \neq \mu_Y + \Delta.$$

Especifiqueu l'estadístic que utilitzeu, la distribució nul·la d'aquest estadístic, la regió de rebuig,i doneu expressions per a la talla de la prova i per a la potència de la prova.

ESTADISTICA MATEMATICA II
EXAMEN FINAL
7 de setembre de 1992

La duració de l'examen és de 3 hores (sense interrupció).

Només podeu tenir amb vosaltres les taules estadístiques i la calculadora.

PROBLEMA 1

Aquest problema fa referència a l'estimació de la variança d'una distribució normal de mitjana desconeguda a partir d'una mostra $X_1, \ldots, X_n$ de v.a. normals. Considerem tres possibles estimadors de la variança poblacional:

$$S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X})^2$$
$$\hat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^{n}(X_i - \bar{X})^2$$
$$\sigma_*^2 = \rho\sum_{i=1}^{n}(X_i - \bar{X})^2$$

(1) Calculeu el biaix de cadascun d'aquests tres estimadors. Quins d'aquests són no esbiaixats?

(2) Calculeu la variança de cadascun d'aquests tres estimadors.

(3) Calculeu l'error quadràtic mitjà dels dos primers estimadors.

(4) Per quin valor de ρ l'estimador σ_*^2 té error quadràtic mitjà més petit?

(5) Quin estimador escollirieu i per què?

PROBLEMA 2

Siguin $X_1, \ldots, X_n$ v.a. independents distribuïdes segons una llei uniforme a l'interval $[0, \theta]$. Denotem per $X_{(n)}$ el màxim de les n v.a.

(1) Calculeu la distribució de $\frac{X_{(n)}}{\theta}$.

(2) Doneu un interval de confiança per a θ de nivell 0.9.

PROBLEMA 3

La v.a. X segueix una distribució exponencial de paràmetre $\lambda = 1$ o $\lambda = 2$. Volem contrastar amb una mostra de mida 1 les hipòtesis $H_0 : \lambda = 1$ contra $H_1 : \lambda = 2$.

(1) Calculeu les probabilitats dels errors de tipus I i de tipus II quan la regió d'acceptació és $A_0 = \{x \leq 1\}$.

(2) Calculeu les probabilitats dels errors de tipus I i de tipus II quan la regió d'acceptació és $A_1 = \{x \geq 0.07\}$.

(3) Quina de les dues proves definides per les regions d'acceptació dels apartats 1 i 2 triarieu?

(4) Trobeu el valor de k per tal que la suma dels errors sigui mínima quan la regió d'acceptació és $A_k = \{x \geq k\}$.

PROBLEMA 4

Disposem d'una mostra $X_1, \ldots, X_n$ d'una llei que té per densitat

$$f(x;\theta) = e^{-(x-\theta)}, \quad \theta \in R, \quad \theta \leq x < \infty.$$

1) Trobeu l'estimador $\hat{\theta}_1$ de θ pel mètode dels moments.

2) Trobeu l'estimador $\hat{\theta}_2$ de θ pel mètode de la màxima versemblança.

3) Calculeu la funció de densitat de l'estimador de màxima versemblança, $\hat{\theta}_2$. És $\hat{\theta}_2$ suficient? Raoneu la resposta.

4) Enunancieu i demostreu el teorema de Rao-Blackwell. Utilitzeu el teorema de Rao-Blackwell per a trobar un estadístic suficient, no esbiaixat i amb variança més petita que la de $\hat{\theta}_2$.

ESTADISTICA MATEMATICA II
EXAMEN 1
19 d'abril de 1993

PROBLEM 1

Sigui $X_1, \ldots, X_n$ una mostra d'una llei exponencial de paràmetre λ. Demostreu que la v.a.

$$U = \frac{X_1}{\sum_{i=1}^n X_i}$$

segueix una llei beta de paràmetres $(1, n-1)$.

PROBLEM 2

La vida (en hores) d'una component electrònica és una variable aleatòria T que es distribueix com una exponencial amb taxa de falla λ. Estem interessats en la probabilitat de que aquesta component falli (mori) abans de 3 hores. Per això prenem n components electròniques i anotem llurs vides; és a dir, disposem d'una mostra de grandària n, $T_1, \ldots, T_n$.

a) Calculeu la probabilitat $p = p(\lambda)$ que una component electrònica falli abans de 3 hores.

b) Estimeu p pel mètode dels moments. Doneu una interpretació a l'estimador obtingut.
Indicació: Definir
$$X_i = \begin{cases} 1 & \text{si} \quad T_i \leq 3 \\ 0 & \text{si} \quad T_i > 3 \end{cases}$$

c) Cerqueu un estadístic suficient i complet per aquesta família. Raoneu la resposta.

d) Trobeu un estimador U.M.V.U. per p.
Indicació 1: Sigui $S(X_1) = \mathbf{1}(X_1 \leq 3)$. Tenim que $E(S(X_1)|T = t) = \text{Prob}(X_1 \leq 3|T = t)$.
Indicació 2: Feu servir el resultat del problema 1.

PROBLEM 3

(Opcional)

Sigui $X_1, \ldots, X_n$ v.a. independents i idènticament distribuïdes amb funció de densitat

$$f(x;\theta) = \frac{1}{6\theta^4} x^3 \exp\{-x/\theta\} \quad \text{per} \quad 0 < x < \infty \quad \text{i} \quad \theta > 0.$$

a) Obtingueu, mitjançant el mètode dels moments, un estimador per a θ.

b) Calculeu la variància de l'estimador obtingut a l'apartat a).

c) Obtingueu, mitjançant el mètode de la màxima versemblança, un estimador per a θ.

d) Doneu la llei asimptòtica de l'estimador de màxima versemblança.

e) Podeu concloure que l'estimador de màxima versemblança és eficient? Raoneu la resposta.

EXAMEN FINAL. PRIMER PARCIAL
22 de juny de 1993

PROBLEM 1 (3 punts)

(a) Siguin U i V dues v.a. independents distribuïdes segons una llei normal estàndar. Demostreu que U/V segueix una distribució t de Student amb un grau de llibertat.

(b) Siguin U i V dues v.a. independents distribuïdes segons una llei normal amb mitjana 0 i variància σ^2. Demostreu que les v.a. $U+V$ i $U-V$ són independents.

(c) Siguin U i V dues v.a. independents distribuïdes segons una llei normal amb mitjana 0 i variància σ^2. Demostreu que $\frac{U+V}{U-V}$ segueix una distribució t de Student amb un grau de llibertat.

Indicació: La funció de densitat corresponent a una t de Student amb ν graus de llibertat és:

$$f(x;\nu) = \frac{\Gamma(\frac{\nu+1}{2})}{(\nu\pi)^{1/2}\Gamma(\frac{\nu}{2})}\left(1+\frac{x^2}{\nu}\right)^{-(\nu+1)/2} \qquad -\infty < x < \infty$$

Nota: Els tres apartats poden fer-se de forma independent i cada apartat pot utilitzar els resultats dels apartats anteriors encara que aquests no s'hagin demostrat.

PROBLEM 2 (6 punts)

Disposem d'una mostra $X_1,\ldots,X_n$ d'una llei que té per densitat ($\theta > 0$, $\tau > 0$)

$$f(x;\theta,\tau) = \begin{cases} \frac{\theta\tau^\theta}{x^{\theta+1}} & x \geq \tau \\ 0 & x < \tau \end{cases}.$$

PART 1: Suposem primer que el valor de τ és conegut i el de θ desconegut.

(a) Calculeu l'estimador $\hat{\theta}_1$ de màxima versemblança de θ.

(b) És aquest estimadors $\hat{\theta}_1$ suficient pel paràmetre θ.

(c) Calculeu la distribució aproximada de $\hat{\theta}_1$.

PART 2: Suposem ara que el valor de τ és desconegut i el de θ conegut.

(d) Calculeu l'estimador $\hat{\tau}_1$ de màxima versemblança de τ.

(e) És l'estimador $\hat{\tau}_1$ suficient pel paràmetre τ.

PART 3: Suposem per últim que els dos valor de τ i de θ són desconeguts.

(f) Calculeu els estimadors de màxima versemblança de θ i de τ.

QÜESTIÓ: (1 punt)

La informació de Fisher, el concepte d'eficiència i la cota de Cramer-Rao.

ESTADISTICA MATEMATICA II
EXAMEN FINAL. SEGON PARCIAL
22 de juny de 1993

PROBLEM 1 (3 punts)

S'experimentà amb 12 rellotges idèntics amb l'intenció de conèixer la distribució del seu temps de vida. Les duracions en mesos fins que van fallar són les següents:
30.5; 33; 33.1; 36; 42; 55; 55.5: 75.8; 76; 106; 106.5; 107.5.

(a) Estudieu gràficament aquestes dades mitjançant un plot de la Weibull.

(b) Creieu que la distribució de Weibull ajusta correctament aquestes dades?

(c) Estimeu gràficament els paràmetres de forma i escala.

(d) Utilitzeu el gràfic per a estimar la mediana de la distribució.

(e) Penseu que seria raonable donar una garantia de 5 anys a aquests rellotges?

PROBLEM 2 (4 punts)

Sigui $X_1, \ldots, X_n$ una mostra d'una distribució que té per funció de densitat::

$$f(x;\theta) = \theta x^{\theta-1} \quad 0 < x < 1$$

a on el paràmetre $\theta > 0$ és desconegut. Es desitja saber si es pot concloure que $\theta > 1$ basant-nos en una mostra de grandària 8.

(a) Plantejeu formalment el problema.

(b) Demostreu que la funció de densitat conjunta té una raó de versemblança monòtona en l'estadístic T. Precisseu l'expressió d'aquest estadístic.

(c) Demostreu que existeix una prova uniformement més potent (UMP) i determineu genèricament la regió de rebuig d'aquesta prova.

(d) Es desitja concretar la regió de rebuig per un nivell de significació $\alpha_0 = 0.05$. Indicació: Sigui X una v.a. distribuïda segons una llei gamma de paràmetres (n,t). Cal utilitzar la següent relació:

$$\text{Prob}(X \leq a) = \text{Prob}(Y \geq n)$$

on Y és una v.a. distribuïda segons una llei de Poisson de paràmetre at, i utilitzar les taules de la Poisson.

PROBLEM 3 (2 punts)

Suposem que la v.a. X segueix una llei geomètrica de paràmetre p; és a dir,

$$f(x;p) = (1-p)^{x-1} p \quad x = 1, 2, \ldots$$

Suposem que a p li assignem una llei a priori uniforme a l'interval (0,1).

(a) Calculeu la distribució a posteriori de p.

(b) Quin és l'estimador de Bayes de p si fem servir com a funció de pèrdua l'error quadràtic.

QÜESTIÓ: (1 punt)

La prova dels Rangs Signats de Wilcoxon.

ESTADISTICA MATEMATICA II
1 de setembre de 1993

PROBLEMA 1 (3 punts)

Disposem d'una mostra de grandària n d'una població Weibull amb paràmetre $\beta = 0.25$ i $\alpha > 0$ desconegut.

(a) Calculeu l'estimador de α, $\hat{\alpha}_1$, pel mètode dels moments.

(b) Calculeu l'estimador de α, $\hat{\alpha}_2$, pel mètode de la màxima versemblança.

(c) Calculeu la variància asimptòtica de $\hat{\alpha}_2$.

PROBLEMA 2 (3 punts)

La v.a. X segueix una distribució exponencial amb mitjana $\theta = 1$ o $\theta = 0.5$. Volem contrastar amb una mostra de mida 1 les hipòtesis $H_0 : \theta = 1$ contra $H_1 : \theta = 0.5$.

(a) Calculeu les probabilitats dels errors de tipus I i de tipus II quan la regió de rebuig és $R_0 = \{x \geq 1\}$.

(b) Calculeu les probabilitats dels errors de tipus I i la funció de potència quan la regió de rebuig és $R_1 = \{x \leq 0.07\}$.

c) Quina de les dues proves definides per les regions de rebuig dels apartats 1 i 2 triarieu?

(d) Trobeu el valor de x per tal que la suma dels errors sigui mínima quan la regió de rebuig és $R_x = \{x \leq x\}$.

PROBLEMA 3 (3 punts)

Denotem per τ la proporció de nens d'un casal d'estiu que estan d'acord amb anar d'excursió. Suposem que aquest valor τ és desconegut i que els dos monitors A i B assignen a τ les següents densitats a priori $\psi_A(\tau)$ i $\psi_B(\tau)$, respectivament:

$$\psi_A(\tau) = 3\tau^2 \quad per \quad 0 < \tau < 1$$

$$\psi_B(\tau) = 4\tau^3 \quad per \quad 0 < \tau < 1.$$

En aquest s'han apuntat 100 nens i d'aquests 60 volen anar d'excursió. En una mostra de 1000 votants hi han 710 en favor de la llei.

a) Determineu la llei a posteriori de τ que cada monitor assigna a τ.

b) Determineu l'estimador de Bayes de τ que cada monitor calcularia si fessim servir com a funció de pèrdua l'error quadràtic.

c) Demostreu que després de conèixer les preferències dels 100 nens els estimadors de Bayes proposats pels dos estadístics no poden diferir en més de 0.01 amb independència del nombre de nens que prefereixin anar d'excursió.

QÜESTIÓ (1 punt)

Mètode de la màxima versemblança. Propietats dels estimadors de màxima versemblança.

ESTADISTICA MATEMATICA II
EXAMEN 1
9 de maig de 1994

PROBLEMA 1

Considerem la variable aleatòria T que mesura el temps de supervivència en mesos d'un pacient després d'un tractament. Suposem que T es distribueix segons un model exponencial traslladat.

El model exponencial traslladat de paràmetres λ ($\lambda > 0$) i G ($G > 0$) té per funció de densitat:

$$f(t; \lambda, G) = \begin{cases} \lambda \exp\{-\lambda(t - G)\} & \text{si } t \geq G \geq 0 \\ 0 & \text{si } t < G \end{cases}$$

G s'interpreta com el temps de garantia o el mínim temps de vida abans del qual no hi han morts.

Hem fet un estudi amb 11 pacients i llurs temps de supervivència han estat:

$$11, 13, 13, 13, 13, 13, 14, 14, 15, 15, 17.$$

Suposem que aquests pacients són una mostra d'una llei exponencial traslladada de paràmetres λ i G.

a) Calculeu la funció de versemblança d'aquestes dades en funció dels paràmetres desconeguts.

b) Suposeu G conegut. Estimeu la taxa de fallada λ mitjançant el mètode de la màxima versemblança. Calculeu el valor de l'estimador de λ en funció de G pels pacients de l'estudi.

c) Suposeu λ conegut. Estimeu G mitjançant el mètode de la màxima versemblança. Calculeu el valor de l'estimador de G pels pacients de l'estudi.

d) Suposeu λ i G desconeguts.

   d1) Trobeu els estimadors de màxima versemblança per λ i G. Calculeu–los pels pacients de l'estudi.

   2) Feu servir les estimacions obtingudes per a estimar la probabilitat de sobreviure 18 mesos després del tractament.

e) Suposeu G conegut (podeu fer servir el valor trobat a l'apartat c)).

   e1) Trobeu l'estimador de màxima versemblança de la mediana del temps de supervivència i anomeneu–lo $\hat{m}$. Calculeu la mediana del temps de supervivència dels 11 pacients.

   e2) Determineu la distribució asimptòtica de l'estimador $\hat{m}$. Calculeu la variància asimptòtica de $\hat{m}$.

PROBLEMA 2

NOTA: En aquest problema heu d'explicar amb detall tots els resultats que feu servir i les hipòtesis en què aquests es basen.

Donada una v.a. X amb funció de densitat

$$f(x; \theta) = \begin{cases} \frac{1}{6\theta^4} x^3 e^{-x} & \text{si } x > 0 \\ 0 & \text{si } x \leq 0 \end{cases},$$

es desitja realitzar la següent prova d'hipòtesis:

$$H_0 : \theta = 2$$
$$H_1 : \theta = 3$$

basant–se en una mostra $X_1, \ldots, X_n$ d'aquesta població.

a)

a1) Determineu un procediment δ que minimitzi $a\alpha(\delta) + b\beta(\delta)$.

a2) Calculeu el valor de $\alpha(\delta)$ per $a = 11$, $b = 2$, $n = 3$.

b) Es desitja ara realitzar la següent prova d'hipòtesis:

$$H_0 : \theta \geq 2$$
$$H_1 : \theta < 2$$

Determineu un procediment UMP amb nivell de significació $\alpha_0 = 0.05$ per $n = 3$.

ESTADISTICA MATEMATICA II
EXAMEN FINAL
8 de juny de 1994

PROBLEMA 1

Les dades següents de Lieblein & Zelen (1956) donen el nombre de revolucions (en milions) fins a la fallada de 23 coixinets de boles. Les dades ordenades de menor a major són:

17.88	28.92	33.00	41.52	42.12	45.60	48.40
51.84	51.96	54.12	55.56	67.80	68.64	68.64
68.88	84.12	93.12	98.64	105.12	105.84	127.92
128.04	173.40					

a) Raoneu si el model exponencial és adequat per a aquesta situació.
Indicació: Compareu la mitjana i la variància mostral.

b) Feu un ajust gràfic amb un paper de Weibull i calculeu el valor dels paràmetres α i β.

c) Empreu el mètode de la màxima versemblança per estimar els paràmetres α i β. Anomeneu $\hat{\alpha}$ i $\hat{\beta}$ els estimadors obtinguts.

d) Estimeu per màxima versemblança $t_{0.2}$, el percentil del 20% d'aquesta distribució. Anomeneu $\hat{t_{0.2}}$ l'estimador obtingut.

e) Suposeu que el valor de β és conegut i igual a 2.1018. Utilitzeu el mètode delta per calcular la variància asimptòtica de $\hat{t_{0.2}}$.

Indicacions: Els següents valors són necessaris per a poder acabar alguns dels càlculs: $\hat{\beta} = 2.1018$, $\sum_{i=1}^{23} t_i^{\hat{\beta}} = 241426$, $\text{Var}(\hat{\alpha}) = 60.42$.

PROBLEMA 2

Suposem que el nombre de minuts que una persona ha d'esperar l'autobus cada matí té una distribució uniforme a l'interval $(0, \theta)$, a on el valor de θ és desconegut. Suposem que la funció de densitat a priori de θ ve donada per:

$$\xi(\theta) = \begin{cases} \frac{192}{\theta^4} & \text{per } \theta \geq 4, \\ 0 & \text{altrament} \end{cases}.$$

Si els temps esperats en tres matins successius són de $5, 3, 8$ minuts.

a) Calculeu la funció de densitat a posteriori de θ. Especifiqueu el domini de definició d'aquesta funció i les constants que en ella apareixen.

b) Si es vol estimar el valor de θ usant com a funció de pèrdua l'error quadràtic, quina és la forma de l'estimador de Bayes de θ. Calculeu el valor estimat de θ a partir delss tres temps esperats donats més amunt.

PROBLEMA 3

La vida (en minuts) d'un prototipus d'una component electrònica en un test de vida accelerat és una variable aleatòria X que es distribueix com una exponencial amb mitjana θ. Estem interessats en la probabilitat de que aquesta component falli abans de 5 minuts. Per això prenem n components i anotem llurs vides; és a dir, disposem d'una mostra de grandària n, $X_1, \ldots, X_n$.

a) Calculeu la probabilitat $q = q(\theta)$ que una component electrònica falli abans de 5 minuts.

b) Estimeu q pel mètode dels moments. Doneu una interpretació a l'estimador obtingut.
Indicació: Definir

$$Y_i = \begin{cases} 1 & \text{si} \quad X_i \leq 5 \\ 0 & \text{si} \quad X_i > 5 \end{cases}$$

c) Cerqueu un estadístic suficient per aquesta família. Raoneu la resposta.

d) Trobeu un estimador U.M.V.U. per q.
Indicació 1: Sigui $S(Y_1) = \mathbf{1}(Y_1 \leq 5)$. Tenim que $E(S(Y_1)|X = x) = \text{Prob}(Y_1 \leq 5|X = x)$.

QÜESTIÓ

La prova de la F.

ESTADÍSTICA MATEMÀTICA II

9 de setembre de 1994

La duració de l'examen és de 3 hores (sense interrupció).

Només podeu tenir amb vosaltres les taules estadístiques, els papers de probabilitat i la calculadora.

PROBLEMA 1

Una variable aleatòria X segueix una distribució normal de mitjana μ i de variança 1. Es fan 20 observacions de X però en comptes d'escriure el valor de X només s'observa si X és negativa o no.

a) Estimar μ per màxima versemblança basant–se tan sols en la informació disponible.

b) Suposant que l'esdeveniment $\{X < 0\}$ ha ocurregut exactament 14 vegades, trobeu el valor de l'estimador de màxima versemblança de μ.

PROBLEMA 2

Disposem d'una xarxa de 165 cel.les. Es va comptar el nombre de grans de grafit en cada cel.la obtenint els resultats de la següent taula:

Número de grans per cel.la	observats
0	1
1	1
2	5
3	7
4	20
5	34
6	30
7	17
8	22
9	21
10	4
11	2
12	1

Provar la hipòtesi que el nombre de grans per cel.la és una variable aleatòria que segueix una distribució de Poisson.

PROBLEMA 3

Les dades següents de Lieblein & Zelen (1956) donen el nombre de revolucions (en milions) fins a la fallada de 23 coixinets de boles. Les dades ordenades de menor a major són:

17.88	28.92	33.00	41.52	42.12	45.60	48.40
51.84	51.96	54.12	55.56	67.80	68.64	68.64
68.88	84.12	93.12	98.64	105.12	105.84	127.92
128.04	173.40					

a) Raoneu si el model exponencial és adequat per a aquesta situació.
Indicació: Compareu la mitjana i la variància mostral.

b) Feu un ajust gràfic amb un paper de Weibull i calculeu el valor dels paràmetres α i β.

c) Empreu el mètode de la màxima versemblança per estimar els paràmetres α i β. Anomeneu $\hat{\alpha}$ i $\hat{\beta}$ els estimadors obtinguts.

d) Estimeu per màxima versemblança $t_{0.2}$, el percentil del 20% d'aquesta distribució. Anomeneu $\hat{t_{0.2}}$ l'estimador obtingut.

e) Suposeu que el valor de β és conegut i igual a 2.1018. Utilitzeu el mètode delta per calcular la variància asimptòtica de $\hat{t_{0.2}}$.

Indicacions: Els següents valors són necessaris per a poder acabar alguns dels càlculs: $\hat{\beta} = 2.1018$, $\sum_{i=1}^{23} t_i^{\hat{\beta}} = 241426$, $\text{Var}(\hat{\alpha}) = 60.42$.

ESTADÍSTICA MATEMÀTICA II
EXAMEN PARCIAL
24 d'abril de 1995

La duració de l'examen és de 2 hores. Només podeu tenir amb vosaltres la calculadora.

QÜESTIÓ

Siguin $(X_1, Y_1), (X_2, Y_2), \ldots, (X_n, Y_n)$ parells de variables aleatòries relacionades segons el següent model lineal:

$$X_i = \alpha + \beta(Y_i - \bar{Y}_n) + \epsilon_i,$$

on $\bar{Y}_n = \sum_{j=1}^n Y_j/n$, i ϵ_i es distribueix segons una llei normal centrada i de variància σ^2.

Q.1) Demostreu que els estimadors de màxima versemblança de α, β i σ^2 són respectivament:

$$\begin{aligned}
\hat{\alpha} &= \bar{X}_n \\
\hat{\beta} &= \frac{\sum_{i=1}^n (Y_i - \bar{Y}_n)X_i}{\sum_{i=1}^n (Y_i - \bar{Y}_n)^2} \\
\hat{\sigma}^2 &= \frac{1}{n}\sum_{i=1}^n [X_i - \hat{\alpha} - \hat{\beta}(Y_i - \bar{Y}_n)]^2
\end{aligned}$$

Q.2) Quina distribució segueix la variable aleatòria $n\hat{\sigma}^2/\sigma^2$?

Q.3) Demostreu l'apartat b).

Indicació: La funció de versemblança $L(\alpha, \beta, \sigma^2)$ pot expressar-se com la funció de densitat conjunta de $\epsilon_1, \epsilon_2, \ldots, \epsilon_n$ i $\epsilon_i = X_i - \alpha + \beta(Y_i - \bar{Y}_n)$.

PROBLEMA

Considerem una mostra $X_1, \ldots, X_n$ d'una llei que té per densitat

$$f(x; \theta) = \frac{3\theta^3}{(x+\theta)^4} \quad x > 0,$$

i l'espai de paràmetres és $\Omega = \{\theta : \theta > 0\}$

Indicació: Molts dels apartats d'aquest problema poden fer-se independentment dels altres.

P.1) Cerqueu un estimador de θ pel mètode dels moments. Anomeneu-lo $\hat{\theta}_1$.

P.2) Calculeu la variància de $\hat{\theta}_1$. Indicació: Pot resultar més senzill començar calculant $E(X+\theta)^2$.

P.3) Cerqueu un estimador de θ pel mètode de la màxima versemblança. Anomeneu-lo $\hat{\theta}_2$. Deixeu-lo indicat si no podeu resoldre les eqüacions de versemblança.

P.4) Calculeu la informació de Fisher continguda en la mostra $X_1, \ldots, X_n$.

P.5) És l'estimador $\hat{\theta}_1$ eficient?

P.6) Doneu la distribució asimptòtica de $\hat{\theta}_2$ i calculeu la seva variància asimptòtica.

P.7) Compareu $\hat{\theta}_1$ i $\hat{\theta}_2$. Quin dels dos estimadors és asimptòticament més eficient.?

P.8) Calculeu la mediana poblacional en funció de θ i anomeneu–la $M(\theta)$.

P.9) Doneu un estimador per $M(\theta)$ i determineu la seva distribució asimptòtica.

ESTADÍSTICA MATEMÀTICA II
EXAMEN FINAL
22 de juny de 1995

La duració de l'examen és de 3.5 hores. Només podeu tenir amb vosaltres la calculadora, les taules estadístiques i els papers de Weibull.

PROBLEMA 1

Un científic pren n mesures $X_1, X_2, \ldots, X_n$ d'una constant μ fent servir la tècnica T_1 amb una variància de l'error coneguda igual a σ^2. També pren m mesuraments addicionals $Y_1, Y_2, \ldots, Y_m$ de μ fent servir la tècnica T_2 amb variància de l'error coneguda igual a $k\sigma^2$, éssent k una constant positiva coneguda. Els mesuraments es suposen independents i amb distribució normal.

a) Escriure la funció de versemblança pel paràmetre μ basada en les $n+m$ observacions $X_1, X_2, \ldots, X_n$, $Y_1, Y_2, \ldots, Y_m$.

b) Trobeu l'estimador de màxima versemblança de μ. Anomeneu-lo $\hat{\mu}$.

c) Calculeu la distribució de $\hat{\mu}$.

d) Demostreu que si $n = m$ i $k > 1$, aleshores $\hat{\mu}$ està més a prop de $\bar{X}$ que de $\bar{Y}$ i expliqueu perquè és això desitjable.

PROBLEMA 2

Quan les morts a causa d'una malaltia rara, però no contagiosa, esdevenen aleatòriament i uniformement dins d'una població, el nombre de morts en una regió amb població igual a P (és a dir, amb un nombre d'habitants igual a P) es distribueix con una llei de Poisson de mitjana μP.

a) Suposem que el nombre de morts observades en n regions amb poblacions $P_1, P_2, \ldots, P_n$ són $Y_1, Y_2, \ldots, Y_n$. Deriveu una expressió pel màxim de versemblança de μ. Anomeneu-lo $\hat{\mu}$.

b) La següent taula dóna el nombre d'homes morts de càncer de fetge (el càncer de fetge es pot considerar una malaltia rara) durant 4 anys a les regions d'Ontario.

Regions i	P_i	Morts
Regió 1	423.447	37
Regió 2	175.685	11
Regió 3	1.245.379	72
Regió 4	413.465	40
Regió 5	216.476	12
Regió 6	242.810	14
Regió 7	213.591	16
Regió 8	166.045	9
Regió 9	265.880	15
Regió 10	116.371	12

Trobeu $\hat{\mu}$ per a aquestes dades i calculeu el nombre de morts esperades a cada regió. Creieu que aquestes dades són consistents amb les suposicions fetes a l'apartat a)?.

c) Suposem ara que $Y_1, Y_2, \ldots, Y_n$ són variables aleatòries de Poisson de mitjanes $\lambda_1, \lambda_2, \ldots, \lambda_n$ i siguin $P_1, P_2, \ldots, P_n$ constants positives conegudes. Considereu la hipòtesi

$$H_0 : \lambda_1 = \mu P_1, \lambda_2 = \mu P_2, \ldots, \lambda_n = \mu P_n$$

on μ és desconegut. Demostreu que la prova de la raó de versemblança generalitzada per contrastar aquesta hipòtesi és

$$-2 \log \Lambda = 2 \sum_{i=1}^{n} Y_i \log(Y_i / \hat{\lambda}_i)$$

on $\hat{\lambda}_i = \hat{\mu} P_i$ i $\hat{\mu}$ és l'estimador de màxima versemblança trobat a l'apartat a).

d) Apliqueu aquest últim resultat per concloure si les taxes de mort (proporció de morts a cada població) de les 10 regions són proporcionals a les poblacions de les regions.

PROBLEMA 3

Volem estudiar l'efecte que té una certa metodologia d'estudi en la població estudiantil. Escollim 9 parelles de bessons idèntics, i decidim que un membre de la parella apliqui el nou método d'estudi, mentre que l'altre segueix amb el mètode tradicional. La suma de totes les notes obtingudes per un estudiant que aplica el nou mètode és una v.a. X, i la suma de totes les notes obtingudes per un estudiant que aplica el mètode tradicional és una altra v.a. Y. Els resultats obtinguts a final del curs van ser el següents:

Parella		1	2	3	4	5	6	7	8	9
Mètode Nou	X	128	132	114	212	107	196	175	148	131
Mètode Tradicional	Y	110	118	116	205	96	184	178	152	126

a) Apliqueu el test no paramètric que us sembli més adient per tal de decidir si el nou mètode d'estudi és més eficient.

b) Suposeu que les nostres dades provenen d'una llei Normal. Apliqueu el test de la t-student que creieu adient per tal de decidir si el nou mètode d'estudi és més eficient.

c) Utilitzeu un test no paramètric per tal de contrastar si la millora obtinguda amb el nou mètode es distribueix segons una llei Normal de mitjana 4 i variància 9.

d) A la vista de les conclusions tretes a l'apartat c), quin dels dos procediments, el donat a a) i el donat a b), aplicarieu per decidir quin mètode d'estudi és més eficient.

ESTADÍSTICA MATEMÀTICA II

8 de setembre de 1995

La duració de l'examen és de 3 hores (sense interrupció).

Només podeu tenir amb vosaltres les taules estadístiques, els papers de probabilitat i la calculadora.

PROBLEMA 1 (3 punts)

A una certa zona del mediterrani es treballa amb l'hipòtesi que l'alçada de les crestes de les onades (mesurades en metres) segueixen, el mes de desembre, una distribució de Raleigh amb funció de distribució acumulada

$$F(x) = 1 - \exp\left\{\frac{-x^2}{2\gamma^2}\right\}; \quad \gamma = 0.7; 0 < x < +\infty.$$

Per tal de poder analitzar si aquest model havia deixat d'ésser vàlid -cosa que podria ser produïda per un canvi climàtic- es van prendre 10 mostres en dies diferents del mes de desembre i sempre a la mateixa hora, obtenint-se els següents valors:

$$0.91, 0.87, 1.01, 0.76, 0.84, 1.05, 1.12, 0.96, 0.89, 0.57.$$

a) Estudieu gràficament aquestes dades de la forma que us sembli més adequada. Creieu que la distribució de Raleigh ajusta correctament aquestes dades?.

b) Realitzeu una prova de la raó de versemblança generalitzada per decidir si es pot acceptar o no que $\gamma = 0.7$ amb un nivell de significació $\alpha = 0.05$.

PROBLEMA 2 (3 punts)

La vide (en hores) d'una component electrònica és una variable aleatòria X que es distribueix com una exponencial amb mitjana μ. Es vol estudiar la probabilitat que aquesta component falli (mori) abans de 10 hores. Per a això prenem $n = 30$ components electròniques i anotem llurs vides: $X_1, \ldots, X_{30}$.

a) Calculeu la probabilitat $q = q(\mu)$ que una component electrònica falli abans de 10 hores.

b) Estimeu q pel mètode dels moments. Doneu una interpretació a l'estimador obtingut.

c) Cerqueu un estadístic suficient per aquesta família. Raoneu la resposta.

d) Trobeu un estimador U.M.V.U. per a q.

PROBLEMA 3 (2 punts)

Suposem que $X_1, \ldots, X_n$ són variables aleatòries independents que segueixen una llei de Poisson de paràmetre θ i suposem que la llei a priori sobre θ pertany a la família gamma.

a) Demostreu que la llei a priori gamma és conjugada de la Poisson.

b) Determineu l'estimador de Bayes de θ sota una funció de pèrdua quadràtica.

PROBLEMA 4 (2 punts)

a) Enunencieu i demostreu, en el cas discret, el Lema de Neymann-Pearson. Utilitzeu una notació clara i definiu tot els conceptes que utilitzeu.

b) Suposeu que una observació X s'ha d'agafar d'una distribució amb funció de densitat:

$$f(x;\theta) = \begin{cases} 2(1-\theta)x + \theta & \text{per } 0 \leq x \leq 1 \\ 0 & \text{altrament,} \end{cases}$$

a on el valor del paràmetre θ és desconegut ($0 \leq \theta \leq 2$). Suposem que volem resoldre el següent contrast:

$$H_0 : \theta = 2, \qquad H_1 : \theta = 0.$$

Nota: Els tres apartats que ara segueixen es poden resoldre independentment.

b1) Determineu el procediment δ_1 tal que $2\alpha(\delta_1) + \beta(\delta_1)$ és un mínim, i calculeu aquest valor mínim.

b2) Suposeu que el valor de α ha sigut donat ($0 < \alpha < 1$). Determineu el procediment δ_2 tal que $\beta(\delta_2)$ és un mínim, i calculeu aquest valor mínim.

b3) Suposem ara que volem resoldre el següent contrast:

$$H_0 : \theta \geq 1, \qquad H_1 : \theta < 1.$$

Determineu la funció de potència del procediment δ_3 que rebutja l'hipòtesi nul.la si $X > 0.9$. Quina és la talla d'aquest procediment?.

ESTADÍSTICA MATEMÀTICA II
EXAMEN PARCIAL
10 d'abril de 1996

La duració de l'examen és de 2 hores. Només podeu tenir amb vosaltres la calculadora.

PROBLEMA 1

Considerem una mostra $X_1, \ldots, X_n$ d'una llei que té per funció de densitat

$$f(x;\alpha,\beta) = \begin{cases} \frac{1}{\alpha} \exp^{-\frac{x-\beta}{\alpha}} & x \geq \beta \\ 0 & \text{altrament} \end{cases}$$

i l'espai de paràmetres és $\Omega = \{(\alpha,\beta) : \alpha > 0, \beta > 0\}$.

Indicació: Molts dels apartats d'aquest problema poden fer–se amb independència dels altres.

A) Escriviu la funció de versemblança $L(\alpha,\beta)$ donat els valors mostrals $x_1, \ldots, x_n$.

B) Suposant $\alpha = \alpha_0$ conegut,

1) Cerqueu l'estimador de màxima versemblança de β i anomeneu–lo $\hat{\beta}$.
2) Caracteritzeu la llei de $\hat{\beta}$.

C) Suposant $\beta = \beta_0$ conegut,

1) Cerqueu un estimador de α pel mètode de la màxima versemblança i anomeneu–lo $\hat{\alpha}$.
2) Quina distribució segueix $\hat{\alpha}$ quan n és prou gran?
3) Calculeu la informació de Fisher pel paràmetre α continguda en la mostra $X_1, \ldots, X_n$.
4) Trobeu un estimador U.M.V.U. per a α.
5) És l'estimador $\hat{\alpha}$ eficient?

D) Suposeu que $(\hat{\alpha}, \hat{\beta})$ és el màxim de versemblança conjuntament de (α, β).

1) Calculeu la mediana poblacional en funció de α i β i anomeneu–la M.
2) Calculeu l'estimador de màxima versemblança de M i anomeneu–lo $\hat{M}$.

PROBLEMA 2

Sigui X una variable aleatòria d'una població que es distribueix segons una llei de Weibull amb paràmetre d'escala α desconegut i paràmetre de forma $\beta = 3$.

1) Demostreu que aquesta llei pertany a la família exponencial. Identifiqueu les funcions $c(\alpha)$ i $d(\alpha)$.

2) Doneu un estadístic suficient per a α.

3) Doneu la forma natural de la família i anomeneu $\eta = c(\alpha)$. Identifiqueu la funció $d_0(\eta)$.

4) Considerem una mostra $X_1, \ldots, X_n$ de X.

   a) Estimeu $\frac{1}{\eta}$ pel mètode dels moments. Anomeneu $\hat{\tau}$ el corresponent estimador.

   b) Calculeu la variància de $\hat{\tau}$.

   Indicació: Utilitzeu les propietats de la funció $d_0(\eta)$.

ESTADÍSTICA MATEMÀTICA II
EXAMEN FINAL
18 de juny de 1996

La duració de l'examen és de 3 hores. Només podeu tenir amb vosaltres la calculadora, les taules estadístiques i els papers de probabilitat.

PROBLEMA 1 (2 punts)

La variable aleatòria que compta el nombre de morts per asma cada 100.000 habitants en un any es suposa que segueix una llei de Poisson de paràmetre θ, on θ representa la veritable taxa de mortalitat per asma cada 100.000 habitants.

A la ciutat **C** dels Estats Units, amb una població de 200.000 persones, es revisen les causes de mort durant l'any 1994, i es denota per $Y94$ la variable aleatòria que compta el nombre de morts per asma a **C** durant l'any 1994.

a) Quina llei segueix la variable aleatòria $Y94$?.

b) Quina llei *a priori* proposarieu per θ, la taxa de mortalitat per asma? Raoneu la resposta.

c) Sabem que la taxa de mortalitat per asma arreu del mon té mitjana 0.6 i variància 0.12. Basant-vos en la llei *a priori* proposada a b), calculeu el seus paràmetres.

d) Es troba que hi han hagut 3 persones que moriren d'asma a **C** l'any 1994, (*i.e.*, $Y94 = 3$). Deriveu la llei *a posteriori* per la taxa de mortalitat per asma? En particular, doneu la mitjana *a posteriori* i discutiu aquest resultat comparant–lo amb la mitjana *a priori*.

e) L'any 1995 es repeteix l'estudi i es tornen a donar 3 morts (*i.e.*, $Y95 = 3$). Feu servir els càlculs de l'apartat d) per a calcular la mitjana *a posteriori*? Que podeu concloure?

PROBLEMA 2 (3 punts)

Disposem de n variables aleatòries, $Y_1, \ldots, Y_n$ independents i uniformement distribuïdes a l'interval $[0, \theta]$. Imaginem que enviem un observador cada dia, durant n dies successius, amb l'encàrrec de que observi, i anoti, els valors de les Y_j per a un subconjunt de m $(1 \leq m \leq n)$ d'elles; és a dir, anotarà els m valors $Y_{j_1}, \ldots, Y_{j_m}$, on $j_1, \ldots, j_m \in \{1, \ldots, n\}$. Aquest observador, amb bons coneixements d'estadística, decideix anotar tan sols el màxim de $Y_{j_1}, \ldots, Y_{j_m}$.

A) Per quina raó decideix l'observador anotar tan sols el màxim de $Y_{j_1}, \ldots, Y_{j_m}$.?

B) Sigui X_i el màxim calculat el dia i. Anomenem m el número d'observacions d'entre les observacions $Y_j, 1 \leq j \leq n$ que ha escollit. Les n variables aleatòries, $X_1, \ldots, X_n$, descrites més amunt són independents i tenen per funció de densitat (no cal que ho demostreu)

$$f(x_1; m, \theta) = \begin{cases} \theta^{-m} m x_1^{m-1} & \text{si } 0 < x_1 \leq \theta \\ 0 & \text{altrament} \end{cases}$$

on $m > 0, \theta > 0$.

b1) Cerqueu l'estimador de màxima versemblanca per θ suposant que l'observador ens ha dit el valor de m. Anomeneu–lo $\hat{\theta}$.

b2) Per desgràcia aquest observador és molt despistat i s'oblida de dir–nos el valor de m, així que ara m també és desconegut. Cerqueu l'estimador de màxima versemblanca per m suposant que el valor de θ és conegut. Anomeneu-lo $\hat{m}$.

b3) Sabent que $(\hat{m}, \hat{\theta})$ és el màxim de versemblança de (m, θ), calculeu el màxim de versemblança de $\text{Prob}\{X_1 > 1\}$. Raoneu la resposta.

NOTA: L'apartat B es pot fer independentment del A usant la funció de densitat $f(x_1; m, \theta)$ per a la mostra $X_1, \ldots, X_n$.

PROBLEMA 3 (3 punts)

Les errades de mesurament d'un aparell segueixen una distribució normal de mitjana 0 i de variància σ^2. Fem n mesuraments i volem resoldre la prova

$$H_0 : \sigma^2 = 1$$
$$H_1 : \sigma^2 = 2$$

a) Descriure un procediment δ_1 pel qual el valor de $2\alpha(\delta_1) + 4\beta(\delta_1)$ sigui un mínim.

b) Expliqueu el significat de $2\alpha(\delta_1) + 4\beta(\delta_1)$ i calculeu–lo deixant–lo indicat en funció de n i de la llei que resulti adequada.

c) S'han fet 7 mesuraments i s'han detectat els següents errors:

$$0.3 \quad 0.7 \quad -1.1 \quad 2.0 \quad 1.7 \quad -0.8 \quad -0.5.$$

Expliciteu el procediment δ_1 i el valor de $2\alpha(\delta_1) + 4\beta(\delta_1)$.

d) Cerqueu un procediment U.M.P., amb nivell de significació 0.025, anomeneu–lo δ_2, per a resoldre la prova

$$H_0 : \sigma^2 \geq 1$$
$$H_1 : \sigma^2 < 1$$

i expliciteu–lo pels mesuraments de l'apartat c.

PROBLEMA 4 (2 punts)

Una companyia de telèfons fa un experiment pilot amb 12 clients i els hi instala un nou sistema de telecomunicació. Aquests sistemes poden fallar per diversos motius: a) per deixar de funcionar, b) per sentir–se soroll en la transmissió o c) per interferències.

A continuació donem el número de setmanes que triguen en fallar.

$$34 \quad 3 \quad 26 \quad 6 \quad 0.7 \quad 9 \quad 3 \quad 16 \quad 5 \quad 35 \quad 22 \quad 17.$$

Nota: El valor 0.7 indica una falla ocurreguda després de 5 dies.

L'equip d'estadístics de la companyia ha d'explicar a l'equip d'enginyers el comportament de les falles del nou sistema de telecomunicació i ha de respondre a les següents dues qüestions plantejades pels enginyers:

QÜESTIÓ 1: Quantes setmanes de garantia es podríen donar si volguessim que com a màxim el 20% fallés durant la garantia?

QÜESTIÓ 2: Quina es la mitjana i la variància del número de falles?

a) Feu un resum gràfic d'aquestes dades que sigui útil pels enginyers.

b) Feu servir el gràfic anterior per a contestar la primera de les preguntes dels enginyers.

c) Inferiu a partir del gràfic quina és la llei que millor s'ajusta als valors de falles observats i utilitzeu aquesta llei per a contestar la segona pregunta dels enginyers.

VALORS QUE PODEN SER UTILS

$\text{Prob}\{\chi^2_6 \leq 3.46\} = 0.25$ $\text{Prob}\{\chi^2_6 \leq 6.24\} = 0.60$ $\text{Prob}\{\chi^2_6 \leq 6.93\} = 0.67$ $\text{Prob}\{\chi^2_6 \leq 12.48\} = 0.95$
$\text{Prob}\{\chi^2_7 \leq 3.46\} = 0.16$ $\text{Prob}\{\chi^2_7 \leq 6.24\} = 0.49$ $\text{Prob}\{\chi^2_7 \leq 6.93\} = 0.56$ $\text{Prob}\{\chi^2_7 \leq 12.48\} = 0.91$

ESTADÍSTICA MATEMÀTICA II

3 de setembre de 1996

La duració de l'examen és de 3 hores. Només podeu tenir amb vosaltres la calculadora, les taules estadístiques i els papers de probabilitat.

PROBLEMA 1 (6 punts)

La probabilitat que una de les components d'un sistema electrònic falli després de x dies ve donada per:

$$\text{Prob}\{X = x\} = k \left(\frac{\theta}{1+\theta}\right)^x ; \quad x = 0, 1, 2, \ldots ; \quad \theta \geq 0. \qquad (*)$$

a) Calculeu el valor de k i caracteritzeu la llei de la variable aleatòria X.

b) n sistemes electrònics s'han posat en funcionament i s'han enregistrat els valors dels n temps de fallada. Amb aquesta informació cerqueu l'estimador de màxima versemblança pel paràmetre θ.

c) Considerem la llei de probabilitat que té per funció de probabilitat la funció donada a (*). Pertany aquesta llei a la família exponencial?. Raoneu la resposta.

d) Deriveu un estimador suficient per a θ.

e) Calculeu la informació de Fisher que conté la mostra dels n sistemes electrònics sobre el paràmetre θ.

f) Basant–vos en els resultats anteriors que siguin necessaris, i fent les comprovacions oportunes, decidiu si l'estimador de màxima versemblança calculat a l'apartat b) és eficient.

×

PROBLEMA 2 (4 punts)

Una agència de viatges vol determinar si la durada de les trucades telefòniques dels seus clients segueix una distribució exponencial. La darrera setmana aquesta agència va enregistrar les durades de totes les trucades i obtingué els següents resultats en segons:

$$4 \quad 6 \quad 5 \quad 8 \quad 9 \quad 10 \quad 12 \quad 8 \quad 16 \quad 20 \quad 24 \quad 27$$

$$33 \quad 37 \quad 43 \quad 50 \quad 58 \quad 68 \quad 70 \quad 78 \quad 88 \quad 100 \quad 120 \quad 130.$$

a) Creieu que aquestes dades indiquen que la durada de les trucades telefòniques queden governades per una distribució exponencial.? Utilitzeu la prova de Kolmogorov–Smirnov per a resoldre–ho.

b) Independentment de quin sigui el resultat a l'apartat a), suposeu que la durada X de les trucades telefòniques queden governades per una distribució exponencial amb mitjana θ. Basant-vos en una mostra aleatòria $X_1, \ldots, X_n$, determineu un procediment per a resoldre la prova d'hipòtesi

$$H_0 : \theta = \theta_0$$
$$H_1 : \theta = \theta_1$$

amb $\theta_0 < \theta_1$ tal que minimitzi l'error de tipus II i que mantingui un nivell de significació com a màxim del 5%.

c) Estem interessats en la següent prova d'hipòtesi:

$$H_0' : \theta \leq \theta_0$$
$$H_2 : \theta > \theta_0.$$

Existeix algun procediment uniformement més potent? En cas afirmatiu doneu aquest procediment amb nivell de significació del 5%.

d) Amb les dades de l'apartat a), resolgueu les següents proves d'hipòtesis i feu les conclusions pertinents.

$$H_0 : \theta = 30 \quad \text{versus} \quad H_1 : \ \theta = 45$$
$$H_0' : \theta \leq 30 \quad \text{versus} \quad H_2 : \ \theta > 30.$$

VALORS QUE PODEN SER UTILS

$\text{Prob}\{\Gamma(23; \frac{1}{30}) \leq 471\} = 0.05$ $\text{Prob}\{\Gamma(23; \frac{1}{30}) \leq 880\} = 0.90$ $\text{Prob}\{\Gamma(23; \frac{1}{30}) \leq 944\} = 0.95$

$\text{Prob}\{\Gamma(23; \frac{1}{35}) \leq 707\} = 0.05$ $\text{Prob}\{\Gamma(23; \frac{1}{35}) \leq 1320\} = 0.90$ $\text{Prob}\{\Gamma(23; \frac{1}{35}) \leq 1415\} = 0.95$

$\text{Prob}\{\Gamma(24; \frac{1}{30}) \leq 496\} = 0.05$ $\text{Prob}\{\Gamma(24; \frac{1}{30}) \leq 914\} = 0.90$ $\text{Prob}\{\Gamma(24; \frac{1}{30}) \leq 979\} = 0.95$

$\text{Prob}\{\Gamma(24; \frac{1}{35}) \leq 744\} = 0.05$ $\text{Prob}\{\Gamma(24; \frac{1}{35}) \leq 1371\} = 0.90$ $\text{Prob}\{\Gamma(24; \frac{1}{35}) \leq 1468\} = 0.95$