Taras Bodnar
Department of Mathematics
Stockholm University
E-Mail: taras.bodnar@math.su.se

Theory of Statistical Inference

Exam, 2021/10/27

The only allowed aid is a pocket calculator provided by the department. The answers to the tasks should be clearly formulated and structured. All non-trivial steps need to be commented. The solutions should be given in English or Swedish.

The written exam is divided into two parts. The first part considers the most central of the course concepts and it is related to standard problems. The second part consists of problems that requires a higher level of understanding, the ability to generalize and to combine methods. Each part consists of three problems and will worth a maximum of 50 points. In order to receive grades A-E, a minimum of 35 points is required in the first part. The second part is only graded for students passing the first part. Given a minimum of 35 points in the first part, the final grade is determined by the sum of regular points in both parts of the exam and bonus points according to the following table:

Grade	A	B	C	D	E	F
Points	$\geq$90	(90-80]	(79-70]	(69-60]	< 60 and ≥ 35 in Part I	< 35 in Part I

Up to 10 bonus points (i.e., in addition to the ordinary 100 points) are given for the active participation in the problem sessions. A half of the bonus points will be used for the first part of the exam, while the second half of the bonus points will be used in the second part of the exam.

Part I:

Problem 1 [23P]

Suppose that we have an iid sample $X_{1:n} = (X_1, X_2, ..., X_n)$ from a Borel distribution with probability mass function of X_i given by

$$\mathbb{P}(X_i = x; \beta) = \frac{1}{x!}(\beta x)^{x-1}e^{-\beta x} \qquad \text{for} \qquad x = 1, 2, ...$$

where $\beta \in (0, 1)$ is an unknown parameter. It also holds that

$$\mathbb{E}(X_i) = (1-\beta)^{-1} \quad \text{and} \quad \mathbb{V}ar(X_i) = \beta(1-\beta)^{-3}.$$

(a) Derive the maximum likelihood estimate $\hat{\beta}_{ML}$ for β. **[4P]**

(b) Derive the ordinary Fisher information $I_{1:n}(\beta)$, the observed Fisher information $I_{1:n}(\hat{\beta}_{ML})$, and the expected Fisher information $J_{1:n}(\beta)$. **[5P]**

(c) Find a minimum sufficient statistic for β and explain your answer.**[3P]**

(d) Determine the asymptotic distribution of $\hat{\beta}_{ML}$ as $n \to \infty$ and provide the analytical expressions of its parameters expressed as functions of β. **[3P]**

(e) Construct a 95% two-sided Wald confidence interval for β. **[3P]**

(f) Based on the observed data

$$x_1 = 2, x_2 = 7, x_3 = 5, x_4 = 3, x_5 = 2, x_6 = 9, x_7 = 2, x_8 = 3, x_9 = 5, x_{10} = 6, x_{11} = 4, x_{12} = 4,$$

perform the Wald test of the hypothesis

$$H_0: \; \beta = 0.5 \quad \text{against} \quad H_1: \; \beta \neq 0.5$$

at significance level of 5%. **[5P]**

Hint: *Important quantiles of the standard normal distribution are:*

$z_{0.9}$	$z_{0.95}$	$z_{0.975}$	$z_{0.95}$
1.28	1.64	1.96	2.33

Problem 2 [17P]

Let $X_{1:n} = (X_1, ..., X_n)^T$ denote a random sample from a Poisson distribution with probability mass function of X_i given by

$$\mathbb{P}(X_i = x; \lambda) = \frac{\lambda^x}{x!}\exp(-\lambda) \qquad \text{for} \qquad x = 0, 1, 2, ... \quad \text{and} \quad \lambda > 0.$$

It also holds that $\mathbb{E}(X_i) = \lambda$ and $\mathbb{V}ar(X_i) = \lambda$.

(a) Show that the conjugate prior for λ is given by the gamma distribution with shape parameter $\alpha > 0$ and rate parameter $\beta > 0$, that is

$$f(\lambda) = \frac{\beta^\alpha}{\Gamma(\alpha)}\lambda^{\alpha-1}\exp(-\beta\lambda) \quad \text{for} \quad \lambda > 0 \quad \text{and} \quad \alpha, \beta > 0$$

with prior mean $\mathbb{E}(\lambda) = \alpha/\beta$. Determine the parameters in the corresponding posterior distribution. **[5P]**

(b) Compute the posterior mean of λ when the conjugate prior is used. **[2P]**

(c) Find the expression of the Jeffreys prior for λ and computed the corresponding posterior distribution. **[6P]**

(d) Calculate the posterior mean when the Jeffreys prior is used and compare it to the expression of the MLE estimator for λ obtained in the frequentist statistics. **[4P]**

Problem 3 [10P]

Let $J_\theta(\theta)$ denote the expected Fisher information of a scalar parameter θ and suppose that $\phi = h(\theta)$ is a one-to-one continuously differentiable transformation of θ. Derive the relation between the expected Fisher information $J_\phi(\phi)$ of ϕ and $J_\theta(\theta)$ of θ. Explain each step in detail.

Part II:

Problem 4 [17P]

Let X_1 and X_2 be independent random variables which are negative binomially distributed with probability mass functions expressed as

$$\mathbb{P}(X_1 = x; \pi_1) = \frac{(x+k-1)!}{x!(k-1)!}\pi_1^k(1-\pi_1)^x \qquad \text{for} \qquad x = 0, 1, 2, ...,$$

and

$$\mathbb{P}(X_2 = x; \pi_2) = \frac{(x+k-1)!}{x!(k-1)!}\pi_2^k(1-\pi_2)^x \qquad \text{for} \qquad x = 0, 1, 2, ...,$$

where X_1 and X_2 can be interpreted as the numbers of failures in Bernoulli trials until k successes with success probabilities $\pi_1 \in (0,1)$ and $\pi_2 \in (0,1)$, respectively.

The aim is to test the null hypothesis:

$$H_0: \ \pi_1 = \pi_2. \tag{1}$$

(a) Derive the generalized likelihood ratio statistic for testing H_0 in (2). **[12P]**

(b) Determine the distribution of the test statistics derived in part (a). **[1P]**

(c) Perform the generalized likelihood ratio test at significance level of 10% when $k = 30$ and the realizations of X_1 and X_2 are $x_1 = 12$ and $x_2 = 21$, respectively. **[4P]**

Hint: *Important quantiles of the χ^2-distribution at various degrees of freedom are:*

d	1	2	3	4	5
$\chi^2_{0.9}(\text{df} = d)$	2.71	4.61	6.25	7.78	9.24
$\chi^2_{0.95}(\text{df} = d)$	3.84	5.99	7.81	9.49	11.07
$\chi^2_{0.975}(\text{df} = d)$	5.02	7.38	9.35	11.14	12.83

Problem 5 [18P]

Let $\boldsymbol{X}_{1:n} = (\boldsymbol{X}_1, \boldsymbol{X}_2, ..., \boldsymbol{X}_n)$ be an independent sample from a multivariate 2-dimensional normal distribution with density function of $\boldsymbol{X}_i = (X_{1,i}, X_{2,i})^\top$ given by

$$f(\boldsymbol{x}; \boldsymbol{\mu}) = (2\pi)^{-1}|\boldsymbol{\Sigma}|^{-1/2} \exp\left(-\frac{1}{2}(\boldsymbol{x}-\boldsymbol{\mu})^\top \boldsymbol{\Sigma}^{-1}(\boldsymbol{x}-\boldsymbol{\mu})\right), \quad \boldsymbol{\mu} = (\mu_1, \mu_2)^\top \in \mathbb{R}^2,$$

with known covariance matrix $\boldsymbol{\Sigma} = (\sigma_{ij})_{i,j=1,...,2}$. The notation $|\boldsymbol{\Sigma}|$ stands for the determinant of matrix $\boldsymbol{\Sigma}$ and $(\boldsymbol{x}-\boldsymbol{\mu})^\top$ denotes the transpose of $(\boldsymbol{x}-\boldsymbol{\mu})$.

(a) Derive the maximum likelihood estimate $\hat{\boldsymbol{\mu}}_{ML}$ for $\boldsymbol{\mu}$. **[3P]**

(b) Compute the expected Fisher information $\mathbf{J}_{1:n}(\boldsymbol{\mu})$.**[4P]**

(c) Derive a 95% score confidence region for $\boldsymbol{\mu}$. **[6P]**

(d) Determine the asymptotic distribution of $\hat{\boldsymbol{\mu}}_{ML}$ as $n \to \infty$ and provide the analytical expressions of its parameters expressed as functions of $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$. **[3P]**

(e) Construct a 95% Wald confidence interval for each component of the parameter vector $\boldsymbol{\mu}$.**[2P]**

Problem 6 [15P]

Formulate the statement about the Cramér-Rao lower bound and prove it.

Taras Bodnar
Department of Mathematics
Stockholm University
E-Mail: taras.bodnar@math.su.se

Statistisk inferensteori

Tentamen, 2021/10/27

Det enda tillåtna hjälpmedlet är en miniräknare som tillhandahålls av institutionen. Lösningarna på uppgifterna skall vara välstrukturerade och tydligt redovisade. Alla icke-triviala steg skall motiveras. Lösningarna får endast vara skrivna på engelska eller svenska.

Den här tentamen är skriftlig och är indelad i två delar. Den första delen fokuserar på de mest centrala delarna av kursen och består av standardproblem. Den andra delen består av problem som kräver en djupare förståelse, samt förmågan att generalisera och kombinera metoder. De två delarna består var och en av tre frågor och kan som mest ge 50 poäng. Totalt kan de två delarna alltså som mest ge 100 poäng. För att uppnå betygen A-E krävs minst 35 poäng på den första delen. Den andra delen rättas enbart om 35 poäng uppnåtts på den första delen, och betyget ges då av summan av antalet poäng på de två delarna enligt:

Betyg	A	B	C	D	E	F
Poäng	$\geq$90	(90-80]	(79-70]	(69-60]	$<$ 60 och $\geq$ 35 i Del I	$<$ 35 i Del I

Bonusuppgifterna kan totalt bidra med upp till 10 poäng. Hälften av de uppnådda bonuspoängen adderas till poängen för del I, och den andra hälften adderas till poängen för del II.

Del I:

Problem 1 [23P]

Låt $X_{1:n} = (X_1, X_2, ..., X_n)$ vara ett slumpmässigt stickprov (dvs. $X_1, X_2, ..., X_n$ är oberoende och likafördelade) från en Borelfördelning med sannolikhetsfunktion

$$\mathbb{P}(X_i = x; \beta) = \frac{1}{x!}(\beta x)^{x-1} e^{-\beta x} \qquad \text{för} \qquad x = 1, 2, ...$$

där parametern $\beta \in (0, 1)$ är okänd. Då gäller

$$\mathbb{E}(X_i) = (1-\beta)^{-1} \quad \text{och} \quad \mathbb{V}ar(X_i) = \beta(1-\beta)^{-3}.$$

(a) Hitta maximum-likelihood-skattaren $\hat{\beta}_{ML}$ av β. **[4P]**

(b) Bestäm Fisherinformationen $I_{1:n}(\beta)$, den observerade Fisherinformationen $I_{1:n}(\hat{\beta}_{ML})$, och den förväntade Fisherinformationen $J_{1:n}(\beta)$. **[5P]**

(c) Ange en minimalt tillräcklig statistika för β och motivera ditt svar.**[3P]**

(d) Hitta den asymptotiska fördelningen för $\hat{\beta}_{ML}$ då $n \to \infty$. Ange de analytiska uttrycken for gränsfördelningens parametrar som funktioner av β. **[3P]**

(e) Konstruera ett 95%-igt tvåsidigt Wald-konfidensintervall för β. **[3P]**

(f) Utför ett Waldtest av hypoteserna

$$H_0: \ \beta = 0.5 \quad \text{mot} \quad H_1: \ \beta \neq 0.5$$

på signifikansnivån 5% baserat på följande observationer:

$x_1 = 2, x_2 = 7, x_3 = 5, x_4 = 3, x_5 = 2, x_6 = 9, x_7 = 2, x_8 = 3, x_9 = 5, x_{10} = 6, x_{11} = 4, x_{12} = 4.$

[5P]

Hjälp: *Några kvantiler för standard-normalfördelningen ges av:*

$z_{0.9}$	$z_{0.95}$	$z_{0.975}$	$z_{0.95}$
1.28	1.64	1.96	2.33

Problem 2 [17P]

Låt $X_{1:n} = (X_1, ..., X_n)^T$ utgöra ett slumpmässigt stickprov från en Poissonfördelning med sannolikhetsfunktion

$$\mathbb{P}(X_i = x; \lambda) = \frac{\lambda^x}{x!} \exp(-\lambda) \qquad \text{för} \qquad x = 0, 1, 2, ... \quad \text{och} \quad \lambda > 0.$$

Vi har att $\mathbb{E}(X_i) = \lambda$ och $\mathbb{V}ar(X_i) = \lambda$.

(a) Visa att den konjugerande apriorifördelningen för λ ges av Gammafördelningen. Bestäm även den motsvarande aposteriorifördelningens parametrar. Du kan använda att om λ är Gammafördelad med shape-parameter $\alpha > 0$ och rate-parameter $\beta > 0$ så har λ täthet

$$f(\lambda) = \frac{\beta^\alpha}{\Gamma(\alpha)} \lambda^{\alpha-1} \exp(-\beta\lambda) \quad \text{för} \quad \lambda > 0 \quad \text{och} \quad \alpha, \beta > 0$$

och väntevärde $\mathbb{E}(\lambda) = \alpha/\beta$. **[5P]**

(b) Hitta aposteriori-väntevärdet av λ om den konjugerande apriorifördelningen används. **[2P]**

(c) Hitta Jeffrey's apriorifördelning för λ och den motsvarande aposteriorifördelningen. **[6P]**

(d) Hitta aposteriori-väntevärdet av λ om Jeffrey's apriorifördelningen används. Jämför ditt svar med Maximum-likelihood-skattaren av λ från det frekventistiska ramverket. **[4P]**

Problem 3 [10P]

Låt $J_\theta(\theta)$ beteckna den förväntade Fisherinformationen för en en-dimensionell parameter θ, och låt $\phi = h(\theta)$ där funktionen h är en-till-en och kontinuerligt differentierbar. Härled sambandet mellan den förväntade Fisherinformationen $J_\phi(\phi)$ för ϕ den förväntade Fisherinformationen $J_\theta(\theta)$ för θ. Förklara och motivera noggrant stegen i härledningen.

Del II:

Problem 4 [17P]

Låt X_1 och X_2 vara oberoende slumpvariabler från varsin negativ-binomialfördelning vars sannolikhetsfunktioner ges av

$$\mathbb{P}(X_1 = x; \pi_1) = \frac{(x+k-1)!}{x!(k-1)!}\pi_1^k(1-\pi_1)^x \qquad \text{för} \qquad x = 0, 1, 2, ...,$$

respektive

$$\mathbb{P}(X_2 = x; \pi_2) = \frac{(x+k-1)!}{x!(k-1)!}\pi_2^k(1-\pi_2)^x \qquad \text{för} \qquad x = 0, 1, 2, ..., .$$

Slumpvariabeln X_i kan ses som antalet misslyckande försök som krävs innan k lyckade försök uppnåtts i en sekvens av oberoende Bernoulli-försök med parameter $\pi_i \in (0, 1)$, för $i = 1, 2$.

Målet med uppgiften är att testa nollhypotesen:

$$H_0 : \ \pi_1 = \pi_2. \tag{2}$$

(a) Bestäm den generaliserade likelihood-ratio-statistikan för att testa H_0 i (2). **[12P]**

(b) Ange fördelningen för teststatistikan i del (a). **[1P]**

(c) Genomför ett generaliserat-likelihood-ratio-test på signifikansnivån 10% givet realisationerna $x_1 = 12$ och $x_2 = 21$ av X_1 respektive X_2, och att $k = 30$. **[4P]**

Ledtråd: *Några kvantiler för χ^2-fördelningen ges av:*

d	1	2	3	4	5
$\chi^2_{0.9}(\mathrm{df} = d)$	2.71	4.61	6.25	7.78	9.24
$\chi^2_{0.95}(\mathrm{df} = d)$	3.84	5.99	7.81	9.49	11.07
$\chi^2_{0.975}(\mathrm{df} = d)$	5.02	7.38	9.35	11.14	12.83

Problem 5 [18P]

Antag att $\boldsymbol{X}_{1:n} = (\boldsymbol{X}_1, \boldsymbol{X}_2, ..., \boldsymbol{X}_n)$ utgör ett oberoende stickprov från en multivariat 2-dimensionell normalfördelning, och att $\boldsymbol{X}_i = (X_{1,i}, X_{2,i})^\top$ har täthet

$$f(\boldsymbol{x}; \boldsymbol{\mu}) = (2\pi)^{-1} |\boldsymbol{\Sigma}|^{-1/2} \exp\left(-\frac{1}{2} (\boldsymbol{x} - \boldsymbol{\mu})^\top \boldsymbol{\Sigma}^{-1} (\boldsymbol{x} - \boldsymbol{\mu}) \right) , \quad \boldsymbol{\mu} = (\mu_1, \mu_2)^\top \in \mathbb{R}^2,$$

med känd kovariansmatris $\boldsymbol{\Sigma} = (\sigma_{ij})_{i,j=1,...,2}$. Här betecknar $|\boldsymbol{\Sigma}|$ determinanten av matrisen $\boldsymbol{\Sigma}$ och $(\boldsymbol{x} - \boldsymbol{\mu})^\top$ betecknar transponatet av $(\boldsymbol{x} - \boldsymbol{\mu})$.

(a) Hitta maximul-likelihood-skattningen $\hat{\boldsymbol{\mu}}_{ML}$ for $\boldsymbol{\mu}$. **[3P]**

(b) Bestäm den förväntade Fisherinformationen $\mathbf{J}_{1:n}(\boldsymbol{\mu})$.**[4P]**

(c) Bestäm en 95% score-konfidensregion för $\boldsymbol{\mu}$. **[6P]**

(d) Ange den asymptotiska fördelningen för $\hat{\boldsymbol{\mu}}_{ML}$ då $n \to \infty$ och ange de analytiska uttrycken for gränsfördelningens parametrar som funktioner av $\boldsymbol{\mu}$ och $\boldsymbol{\Sigma}$. **[3P]**

(e) Konstruera varsitt 95%-igt Wald-konfidensintervall för elementen i parametervektorn $\boldsymbol{\mu}$.**[2P]**

Problem 6 [15P]

Formulera satsen om Cramér-Raos undre gräns och bevisa den.