STOCKHOLMS UNIVERSITET
MATEMATISKA INSTITUTIONEN
Avd. Matematisk statistik

MT 5003
TENTAMEN
1 november 2019

Tentamen i Statistisk inferensteori
1 november 2019, kl. 9-14

Examinator: Chun-Biu Li, cbli@math.su.se.
Tillåtna hjälpmedel: Miniräknare tillhandahålles av institutionen, personlig ej tillåten. Formelsamling på tentamens sista sidor.
Återlämning: Meddelas i kurshemsidans forum.

Resonemang skall vara tydliga och lätta att följa. Varje korrekt och fullständigt löst uppgift ger 10 poäng. För betyg A-E krävs 20 poäng på Del 1, samt att följande gränser uppnås på Del 2:

A	B	C	D	E
25	19	13	7	0

Del 1

Låt $X = (X_1, \ldots, X_n)$ beteckna en vektor av n oberoende $Geometrisk(\theta)$-fördelade stokastiska variabler och $x = (x_1, \ldots, x_n)$ en realisering av densamma.

Uppgift 1

a) Bestäm ML-skattaren av oddset $\theta/(1-\theta)$ och visa att den är väntevärdesriktig. (4p)

b) Definiera begreppet tillräcklig (sufficient) stickprovsvariabel. (3p)

c) Bestäm en endimensionell tillräcklig (sufficient) stickprovsvariabel. (3p)

Uppgift 2

a) Beskriv hur standardfelet hos ML-skattaren $\hat{\theta}$ kan uppskattas med parametrisk Bootstrap. (5p)

b) Bestäm aposteriorifördelningen för θ givet Jeffreys prior, ange även en konjungerande familj av apriorifördelningar. (5p)

Uppgift 3

a) För ML-skattaren gäller att $\sqrt{n}(\hat{\theta}-\theta)$ konvergerar i fördelning mot $N(0, \sigma^2)$. Vad är σ^2 och hur kan resultatet användas för att bestämma ett standardfel för $\hat{\theta}$? (5p)

b) Härled att uttryck för Score statistikan $T_S(x)$ för hypotesen $H_0 : \theta = 1/2$. Kan nollhypotesen förkastas på nivån 0.05 om vi observerar $T_S(x) = 4.3$? Stickprovsstorleken kan antas stor nog för att asymptotiska resulat skall gälla med god noggrannhet. (5p)

Del 2

Uppgift 4

a) Formulera Cramér-Raos olikhet. Använd den för att visa att odds-skattaren i uppgift 1a) har minst varians av alla väntevärdesriktiga skattare, utan att hänvisa till asymptotiska resultat. (5p)

b) Härled den asymptotiska fördelningen hos ML-skattaren under regularitetsvillkor genom att göra en första ordningens Taylorapproximation av scorefunktionen. Högre ordningens termer kan antas försumbara utan närmare motivering. (5p)

Uppgift 5

Låt y utgöra ett datamaterial från en fördelning med parameter θ. För att skatta precisionen hos en Bayesiansk punktskattare $\hat{\theta}(y)$ av θ kan man använda medelkvadratfelet $E_{\theta|y}[(\theta-\hat{\theta}(y))^2] = \int f(\theta|y)(\theta-\hat{\theta}(y))^2 d\theta$. Här betecknar $f(\theta|y)$ tätheten för aposteriorifördelningen.

a) Låt $\mu = E_{\theta|y}(\theta)$ beteckna väntevärdet för aposteriorifördelningen. Visa att medelkvadratfelet $E_{\theta|y}[(\theta - \hat{\theta}(y))^2]$ kan skrivas på formen $E_{\theta|y}[(\theta - \hat{\theta}(y))^2] = Var_{\theta|y}(\theta) + (\mu - \hat{\theta}(y))^2$ genom att, som ett första steg, använda att $E_{\theta|y}[(\theta-\hat{\theta}(y))^2] = E_{\theta|y}[(\theta-\mu+\mu-\hat{\theta}(y))^2]$. Motivera varje steg tydligt. (6p)

b) Använd del a) för att hitta den Bayesianska punktskattare $\hat{\theta}(y)$ för vilken medelkvadratfelet $E_{\theta|y}[(\theta - \hat{\theta}(y))^2]$ är minimerat. Vad är det minimala värdet på $E_{\theta|y}[(\theta - \hat{\theta}(y))^2]$? (4p)

Kommentar: Både deluppgift a) och b) kan lösas i ungefär fem steg eller färre. Om din lösning blir väldigt lång så behöver du troligtvis testa ett annat angripssätt.

Uppgift 6

Rayleighfördelningen fås som specialfallet av Weibullfördelningen då formparametern $k = 2$ $\quad (Rayleigh(\lambda^2/2) = Weibull(\lambda, 2))$. I denna uppgift skall du anta att du har tillgång till n oberoende observationer från en $Weibull(\lambda, k)$-fördelning.

a) Bestäm score-vektorn. (2p)

b) Bestäm den observerade Fisherinformationsmatrisen. (3p)

c) Bestäm ett uttryck för profil-likelihooden för parametern k. (3p)

d) Beskriv hur man kan utföra ett Generalized-likelihood-ratio-test av hypotesen $H_0 : k = 2$ på nivån 5%. Du kan hänvisa till ML-skattarna $\hat{\lambda}_{\mathrm{ML}}$ och $\hat{k}_{\mathrm{ML}}$ utan att ange deras explicita uttryck. (2p)

Lycka till!

Formelsamling med användbara fördelningar

Betafördelningen $X \sim Beta(\alpha, \beta), \quad \alpha > 0, \beta > 0.$
Täthetsfunktion:

$$p(x|\alpha, \beta) = \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} x^{\alpha-1}(1-x)^{\beta-1}, \quad 0 \leq x \leq 1.$$

$E(X) = \alpha/(\alpha+\beta)$, $V(X) = \alpha\beta/((\alpha+\beta)^2(\alpha+\beta+1))$

Binomialfördelningen $X \sim Binomial(n, p), \quad 0 \leq p \leq 1, n = 1, 2, \ldots.$
Sannolikhetsfunktion:

$$p(x|n, p) = \binom{n}{x} p^x (1-p)^{n-x}, \quad x = 0, 1, \ldots, n.$$

$E(X) = np$, $V(X) = np(1-p)$.

χ^2**-fördelningen** $X \sim \chi^2(k), \quad k = 1, 2, 3, \ldots.$
Täthetsfunktion:

$$p(x|k) = \frac{1}{2^{k/2}\Gamma(k/2)} x^{k/2-1} \exp\left(-\frac{x}{2}\right), \quad x \geq 0.$$

$E(X) = k$, $V(X) = 2k$.
Några approximativa kvantiler:

$k = 1; \quad P(X > 3.84) = 0.05$
$k = 2; \quad P(X > 5.99) = 0.05$
$k = 3; \quad P(X > 7.81) = 0.05$

Gammafördelningen $X \sim Gamma(\alpha, \beta), \quad \alpha > 0, \beta > 0.$
Täthetsfunktion:

$$p(x|\alpha, \beta) = \frac{\beta^\alpha}{\Gamma(\alpha)} x^{\alpha-1} \exp(-\beta x), \quad x \geq 0.$$

$E(X) = \alpha/\beta$, $V(X) = \alpha/\beta^2$.

Geometriska fördelningen $X \sim Geometrisk(\theta), \quad 0 \leq \theta \leq 1.$
Sannolikhetsfunktion:

$$p(x; \theta) = (1-\theta)^x \theta, \quad x = 0, 1, 2, \ldots.$$

$E(X) = (1-\theta)/\theta$, $V(X) = (1-\theta)/\theta^2$.

Normalfördelningen $X \sim N(\mu, \sigma^2)$, $\quad -\infty < \mu < \infty, 0 < \sigma < \infty$.
Täthetsfunktion:

$$p(x|\mu,\sigma) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\Big(-\frac{(x-\mu)^2}{2\sigma^2}\Big), \quad -\infty < x < \infty.$$

$E(X) = \mu$, $V(X) = \sigma^2$.
Några approximativa kvantiler för $N(0,1)$:
$P(X > 2.58) = 0.005, P(X > 2.33) = 0.01, P(X > 1.96) = 0.025, P(X > 1.64) = 0.05$.

Poissonfördelningen $X \sim Poisson(\lambda)$, $\quad \lambda > 0$.
Sannolikhetsfunktion:

$$p(x|\lambda) = \frac{1}{x!}\lambda^x \exp(-\lambda), \quad x = 0, 1, 2, \ldots.$$

$E(X) = \lambda$, $V(X) = \lambda$.

Rayleighfördelningen $X \sim Rayleigh(\theta)$, $\quad \theta > 0$.
Täthetsfunktion:

$$p(x|\theta) = \frac{x}{\theta} \exp(-\frac{x^2}{2\theta}), \quad x \geq 0.$$

$E(X) = \sqrt{\pi\theta/2}$, $V(X) = (4-\pi)\theta/2$.

Weibullfördelningen $X \sim Weibull(\lambda, k)$, $\quad \lambda > 0, k > 0$.
Täthetsfunktion:

$$p(x|\lambda,k) = \frac{k}{\lambda}\left(\frac{x}{\lambda}\right)^{k-1} \exp\left(-\left(\frac{x}{\lambda}\right)^k\right), \quad x \geq 0.$$

$E(X) = \lambda\Gamma(1 + 1/k)$, $V(X) = \lambda^2(\Gamma(1 + 2/k) - \Gamma(1 + 1/k)^2)$.