Försättsblad till skriftlig tentamen vid Linköpings universitet

Datum för tentamen	2018-01-19
Sal (1)	TER1(7)
Tid	8-12
Kurskod	732G43
Provkod	TEN1
Kursnamn/benämning **Provnamn/benämning**	Bayesiansk statistik Tentamen
Institution	IDA
Antal uppgifter som ingår i tentamen	4
Jour/Kursansvarig Ange vem som besöker salen	Bertil Wegmann
Telefon under skrivtiden	070-112 83 21
Besöker salen ca klockan	ca 9.30
Kursadministratör/kontaktperson (namn + tfnr + mailaddress)	Annelie Almquist, tel 013-282934, annelie.almquist@liu.se
Tillåtna hjälpmedel	Inga
Övrigt	
Antal exemplar i påsen	

Bertil Wegmann
Avdelning statistik och maskininlärning (STIMA)
IDA
Linköpings universitet
bertil.wegmann@liu.se, 070-1128321
2018-01-19

Tentamen - Bayesiansk statistik, 732G43, 7.5 hp

Betygsgränser: Tentamen omfattar totalt 25 poäng. Slutbetyget på kursen baseras på en sammanvägning av tentamen (50%) och inlämningsuppgifter (50%).

För godkänt betyg på tentamen krävs minst 15 poäng.

Redovisa och motivera tydligt alla dina svar!

1. (9 poäng)

(a) Bayesiansk statistik bygger på Bayes sats. Bayes sats för händelserna A och B kan skrivas som

$$p(A|B) = \frac{p(B|A)\,p(A)}{p(B)}.$$

Skriv om Bayes sats för händelserna ovan till Bayes sats för variabler på ett sådant sätt att du får ett uttryck för sannolikhetsfördelningen aposteriori för en parameter θ givet data $x_1, \ldots, x_n$. Redovisa likelihoodfunktionen och priorfördelningen för θ från detta uttryck. Beskriv i ord vad uttrycken för likelihoodfunktionen och priorfördelningen för θ innebär i ord.

(b) Redogör för den huvudsakliga skillnaden i inferens från enbart likelihoodfunktionen för θ jämfört med inferens från aposteriorifördelningen för θ.

(c) Motivera vad som menas med en icke-informativ priorfördelning samt ge **ett** exempel på en icke-informativ priorfördelning för en andel θ.

(d) Motivera vad som menas med en konjugerad priorfördelning samt ge **ett** exempel på en konjugerad priorfördelning för en parameter θ i en valfri modell för data.

(e) Redogör för den huvudsakliga nackdelen med gridapproximation av en aposteriorifördelning för en uppsättning av 10 stycken parametrar $\theta_1, \theta_2, \ldots, \theta_{10}$. Motivera när kvadratisk approximation av aposteriorifördelningen för $\theta_1, \theta_2, \ldots, \theta_{10}$ kan vara ett bättre samt ett sämre alternativ än MCMC skattning av aposteriorifördelningen.

(f) Beskriv vad 95 % kredibilitetsintervall och Highest Posterior Density Interval (HPDI) innebär för en parameter θ och redogör för vad som skiljer dessa intervall från ett 95 % konfidensintervall för θ.

(g) Redogör för hur man kan skapa ett 95 % kredibilitetsintervall för $odds = \frac{p}{1-p}$, där p är en okänd andel, om man känner till aposteriorifördelningen för p.

2. (5 poäng) Antag följande modell för Bayesiansk linjär regression:

$$Y_1, \ldots, Y_n | \mu, \sigma^2, x \overset{iid}{\sim} N(\mu, \sigma^2)$$

$$\mu = \beta x',$$

där variansen σ^2 är okänd och med vektorn av förklaringsvariabler $x = (1\, x_1\, \ldots\, x_k)$ och vektorn med parametrar $\beta = (\beta_0\, \beta_1\, \ldots\, \beta_k)$. Om β och σ antas oberoende apriori gäller

$$p(\beta, \sigma) = \prod_{j=0}^{k} p(\beta_j) p(\sigma).$$

Uniform prior för parametrarna $(\beta, \ln \sigma)$:

$$p(\beta, \sigma^2 | x) \propto \frac{1}{\sigma^2}.$$

(a) Vad är det för fördelar respektive nackdelar med att använda sig av den uniforma priorn för parametrarna ovan? Motivera.

(b) Antag en annan prior för parametrarna än den uniforma priorn ovan. Beskriv i ord hur man kan beräkna ett 90.9 % kredibilitetsintervall för μ utifrån samplade dragningar från posteriorfördelningen av β.

(c) Antag en annan prior för parametrarna än den uniforma priorn ovan. Beskriv i ord hur man kan beräkna ett 90.9 % kredibilitetsintervall för y utifrån samplade dragningar från posteriorfördelningen av (β, σ).

(d) Antag en annan prior för parametrarna än den uniforma priorn ovan. Beskriv i ord hur man kan skapa ett 90.9 % kredibilitetsintervall för y som funktion av en kontinuerlig förklaringsvariabel x_j utifrån samplade dragningar från posteriorfördelningen av (β, σ).

3. (6 poäng)

(a) Motivera vad det innebär att en modell överanpassar respektive underanpassar data. Ange två approacher för att motverka underanpassning respektive överanpassning av data.

(b) Förklara vad en modells Deviance innebär och vad det kan användas till.

(c) Vad är syftet med att använda sig av regulariserande priors? Motivera.

(d) Vad används informationskriterierna AIC, DIC och WAIC för och vad skiljer dessa informationskriterier åt?

(e) Ange 2 mått för MCMC konvergens till posteriorn som kan användas för att avgöra hur många dragningar som är tillräckligt från posteriorn. Beskriv även innebörden av dessa mått.

4. (5 poäng) Antag följande Bayesianska multilevelmodell med 2 nivåer för individ i som tillhör grupp j:

$$\begin{aligned}
y_i &\overset{iid}{\sim} N(\mu, \sigma_y) \\
\mu &= \alpha_j + \beta x', \\
\alpha_j &\sim N(\alpha, \sigma_\alpha), \\
\alpha &\sim N(0, 10), \\
\sigma_\alpha &\sim halfcauchy(0, 1), \\
\beta_k &\sim N(0, 10), \; k = 1, \ldots, p \\
\sigma_y &\sim halfcauchy(0, 1),
\end{aligned}$$

där $x = (x_1 \ldots x_p)$ är en radvektor med förklaringsvariabler och $\beta = (\beta_1 \ldots \beta_p)$ är en radvektor med parametrar.

(a) Redogör för 2 stycken fördelar med den Bayesianska multilevelmodellen med 2 nivåer.

(b) Beskriv i ord hur man kan beräkna ett 95 % kredibilitetsintervall för μ för en individ i utifrån samplade dragningar från posteriorfördelningen av parametrarna i modellen.

(c) Antag att vi förändrar modellen ovan genom att ändra priorn för α_j till följande prior: $\alpha_j \sim N(0, 10)$. Skriv upp den förändrade modellen i sin helhet och redogör för skillnaden mellan denna modell och modellen ovan med avseende på pooling av information mellan grupper.

(d) Redogör för hur posteriorfördelningen för intraklasskorrelationen ρ_{ICC} kan beräknas utifrån samplade dragningar från posteriorfördelningen av parametrarna i modellen.