Bertil Wegmann
Avdelning statistik och maskininlärning (STIMA)
IDA
Linköpings universitet
bertil.wegmann@liu.se, 070-1128321
2018-03-05

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) Förklara på vilket sätt Bayes sats ligger till grund för Bayesiansk statistik genom att skriva om Bayes sats för händelserna A och B nedan till Bayes sats för en parameter θ givet data $x_1, \ldots, x_n$ och förklara vad uttrycken i denna omskrivning av Bayes sats innebär i ord.

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

   (b) Redogör för den huvudsakliga skillnaden i inferens mellan frekventistisk och Bayesiansk inferens.

   (c) Förklara varför likelihoodfunktionen inte är en täthetsfunktion för en parameter θ och vad likelihoodfunktionen behöver kombineras med i Bayesiansk statistik för att erhålla en giltig täthetsfunktion för θ.

   (d) Motivera vad som menas med en informativ priorfördelning samt ge **ett** exempel på en informativ priorfördelning för ett medelvärde θ.

   (e) Motivera vad som menas med en konjugerad priorfördelning samt ange **en** fördel med en konjugerad priorfördelning för en parameter θ i en valfri modell för data.

   (f) Beskriv vad priorelicitering innebär för en parameter θ samt ge **ett** exempel på hur priorelicitering kan användas utifrån experthjälp om en andel θ.

   (g) Motivera när kvadratisk approximation av aposteriorifördelningen för $\theta_1, \theta_2, \ldots, \theta_{20}$ kan vara ett bättre samt ett sämre alternativ än MCMC skattning av aposteriorifördelningen.

   (h) Beskriv vad ett 95 % kredibilitetsintervall innebär för en parameter θ och redogör för hur detta intervall skiljer sig från ett 95 % konfidensintervall för θ.

   (i) Redogör för hur man kan skapa ett 95 % kredibilitetsintervall för medelvärdet $\frac{\alpha}{\beta}$ i en gammafördelning om man känner till aposteriorifördelningen för α och β.

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

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

   $$\mu = \boldsymbol{\beta x'},$$

   där variansen σ^2 är okänd och med vektorn av förklaringsvariabler $\boldsymbol{x} = (1\, x_1 \ \ldots \ x_k)$ och vektorn med parametrar $\boldsymbol{\beta} = (\beta_0\, \beta_1 \ \ldots \ \beta_k)$.

(a) Antag en uniform prior för parametrarna $(\boldsymbol{\beta}, \ln\sigma)$:

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

Vad är det för fördel med att använda sig av denna uniforma prior för $(\boldsymbol{\beta}, \ln\sigma)$? Motivera.

(b) Oavsett prior, beskriv i ord hur man kan beräkna ett 95.2 % kredibilitetsintervall för μ utifrån samplade dragningar från posteriorfördelningen av $\boldsymbol{\beta}$.

(c) Oavsett prior, beskriv i ord hur man kan skapa 90.9 % kredibilitetsintervall för y som funktion av en dummyvariabel x_j utifrån samplade dragningar från posteriorfördelningen av $(\boldsymbol{\beta}, \sigma)$.

(d) Beskriv vad som skiljer en Bayesiansk förklaringsgrad från en frekventistisk förklaringsgrad för multipel linjär regression samt ange **en** fördel med den Bayesianska förklaringsgraden i Gelman et al (2017) jämfört med andra alternativ för Bayesiansk förklaringsgrad.

3. (7 poäng)

(a) Förklara vad det innebär att en modell överanpassar respektive underanpassar data.

(b) Motivera hur regulariserande priors kan användas för att minska överanpassning av data. Motivera även hur regulariserande priors kan ge en underanpassning av data.

(c) Förklara vad Akaikevikter kan användas till samt ange två informationskriterier som kan användas för att beräkna Akaikevikter och förklara vad som skiljer dessa informationskriterier åt.

(d) Vad kan traceplottar över parameterdragningarna från olika MCMC kedjor användas till? Hur kan en plott för ackumulerade posterior medelvärden av parametrarna över MCMC dragningar användas i MCMC diagnostik? Motivera.

(e) Förklara vad det är för skillnad mellan Bayesiansk logistisk regression och Bayesiansk binomialregression och hur dessa regressionsmodeller relaterar till varandra.

4. (4 poäng) Antag följande Bayesianska multilevelmodell med 2 nivåer för individ i som tillhör en skolklass $j = 1, \ldots, J$:

$$\begin{aligned}
y_i &\overset{iid}{\sim} N(\mu, \sigma_y) \\
\mu &= \alpha_j + \boldsymbol{\beta}\boldsymbol{x}', \\
\alpha_j &\sim N\left(\alpha, \sigma_\alpha\right), \\
\alpha &\sim N\left(0, 10\right), \\
\sigma_\alpha &\sim halfcauchy\left(0, 2\right), \\
\beta_k &\sim N\left(0, 10\right),\ k = 1, \ldots, p \\
\sigma_y &\sim halfcauchy\left(0, 2\right),
\end{aligned}$$

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

(a) Beskriv i ord hur man kan beräkna ett 95 % kredibilitetsintervall för μ för en individ i som tillhör en ny skolklass $J+1$ (som inte fanns med då man anpassade modellen) utifrån samplade dragningar från posteriorfördelningen av parametrarna i modellen.

(b) Antag att vi förändrar modellen ovan genom att sätta $\alpha = 0$ och $\sigma_\alpha = 10$. Skriv upp den förändrade modellen i sin helhet och redogör för skillnaden mellan denna modell och den ursprungliga modellen ovan med avseende på poolning av information mellan grupper.

(c) Förklara vad som menas med fullständig poolning av intercept i en regressionsmodell och hur modellen ovan hade förändrats om fullständig poolning av intercept hade varit gällande.