

## Inlämningsuppgift - Bayesiansk statistik, 7.5 hp

### Moment 1 - introduktion till Bayesiansk statistik

**Inlämningsuppgiften skall lösas individuellt.** Valfritt program får användas för att lösa uppgifterna nedan, men mina instruktioner och ledtrådar gäller för programmeringsspråket R. **Bifoga all kod tillsammans med lösningen för varje deluppgift. Labbrapport** med lösningar på nedanstående uppgifter samt relevanta datorutskriftar **lämnas in på Lisam senast måndagen den 11:e november kl. 23:59.**

1. (9 poäng) I denna uppgift ska du använda data från följande sida med samtliga svenska väljarbarometrar: <http://novus.se/valjaropinionen/samtliga-svenska-valjarbarometrar/>. Klicka på länken SVENSKA VÄLJARBAROMETRAR och välj ett av partierna från september 2019 i Exceldokumentet.
  - (a) Redovisa vilket parti du valt. Tag fram **din konjugerade prior** för andelen väljare i september 2019 för det parti du valt ut. Känn dig fri att skapa **din konjugerade prior** utifrån ditt eget tycke, men du ska använda information från tidigare undersökningar på något valfritt sätt. Plotta fördelningen för **din konjugerade prior**.
  - (b) Använd data för *Demoskops* väljarundersökning i september 2019 och uppdatera din prior till posterior genom att dra 10000 värden direkt från den kända (exakta) posteriorfördelningen. Plotta både din prior och posterior i samma figur och kommentera skillnader mellan prior och posterior.
  - (c) Använd data för *Demoskops* väljarundersökning i september 2019 och simulera fram 10000 värden från posteriorfördelningen med hjälp av 1) gridapproximation och 2) kvadratisk approximation. Plotta posteriorfördelningarna från respektive approximation för denna deluppgift och posteriorfördelningen i uppgift b). Verkar fördelningarna lika? Vad kan det bero på?
  - (d) Gör känslighetsanalys i **din konjugerade prior** genom att presentera följande figurer och kommentera vad skillnaderna/likheterna i posteriorn kan bero på (använd koden Manipulate Beta-Bernoulli Posterior.R för att skapa figurerna):
    - i. figur med **din konjugerade prior**, normaliserade likelihooden och posteriorn.
    - ii. figur med en **uniform prior**, normaliserade likelihooden och posteriorn.
    - iii. figur med en **betydligt mer informativ prior** än **din konjugerade prior**, normaliserade likelihooden och posteriorn.
  - (e) Beräkna posteriorfördelningens medelvärde och standardavvikelse med hjälp av gridapproximationen i uppgift (c).
  - (f) Redovisa ett 90.9 % kredibilitetsintervall för andelen väljare av det parti du valt genom att använda 10000 dragna värden direkt från den kända (exakta) posteriorfördelningen i uppgift (b). Tolk kredibilitetsintervallet i ord och ange hur många gånger mer sannolikt det är med värden på intervallet än värden utanför intervallet.
  - (g) Beräkna fram och redovisa posteriorfördelningen för  $oddset = \frac{p}{1-p}$ , där  $p$  är andelen väljare av det parti du valt. Den kvadratiske approximationen av posteriorn i uppgift (c) ska användas. Kommentera vad posteriorfördelningen för *oddset* visar.

2. (4 poäng) Använd datamaterialet om vikter för kycklingar efter 6 veckors föda, se Kod\_Moment1.R
- (a) Simulera fram 10000 värden från följande posterior för en genomsnittsvikt  $\mu$  med hjälp av kvadratisk approximation:
- använd viktdata för kycklingar som fått födan "soybean" och antag att vikterna följer en normalfördelning med känd standardavvikelse som standardavvikelsen av vikterna.
  - använd en normalfördelad prior för medelvikten, där medelvärdet och standardavvikelsen i priorn är lika med medelvärdet och standardavvikelsen för vikterna på kycklingarna som inte fått varken födan "soybean" eller "meatmeal".
- (b) Simulera på samma sätt som i uppgift (a) fram 10000 värden från posteriorn för en genomsnittsvikt  $\mu$ , **men** använd här i stället viktdata för kycklingar som fått födan "meatmeal". Plotta posteriorfördelningen för  $\mu$  och jämför med posteriorfördelningen för  $\mu$  i (a) i samma figur och kommentera skillnaderna. Verkar det som att genomsnittsvikterna för de olika födorna skiljer sig åt?
- (c) Med hjälp av dom simulerade posteriorvärdena i uppgift (a) och (b), ta fram posteriorfördelningen för skillnaden mellan genomsnittsvikterna för födorna "meatmeal" och "soybean",  $\mu_{\text{meatmeal}} - \mu_{\text{soybean}}$ . Beräkna posteriorsannolikheten att  $\mu_{\text{meatmeal}} - \mu_{\text{soybean}} > 0 \iff \mu_{\text{meatmeal}} > \mu_{\text{soybean}}$  och tolka resultatet i ord.