Harjoitustehtävät 2

1. Oletetaan, että y on binäärinen muuttuja, joka saa arvot 1 ja 2, ja sen jakauma on $p(y|\theta)$. Tuntematon parametri θ voi saada arvot θ_1 ja θ_2. Havainnon y jakauman oletetaan olevan

$$p(1|\theta_1) = 0.2, p(2|\theta_1) = 0.8,$$

kun parametri on θ_1, ja parametrin arvolla θ_2 jakauma on

$$p(1|\theta_2) = 0.5, p(2|\theta_2) = 0.5.$$

Oletetaan, että θ:n priorijakauma on $p(\theta_1) = 0.4$ ja $p(\theta_2) = 0.6$.

a) Määritä θ:n posteriorijakauma, ts. laske $p(\theta_1|1)$, $p(\theta_2|1)$, $p(\theta_1|2)$ ja $p(\theta_2|2)$.

b) Laske priori(vedonlyönti)suhde $p(\theta_1)/p(\theta_2)$.

c) Oletetaan, että havainto on $y = 2$. Laske posteriori(vedonlyönti)suhde $p(\theta_1|2)/p(\theta_2|2)$.

2. Oletetaan, että tuntemattoman suureen θ posteriorijakaumaksi saadaan

$$p(\theta|y) = \frac{1}{2}\theta^2 e^{-\theta}, \quad \theta > 0,$$

missä y on havainto. Kuvaile tätä posterioria: Laske jakauman odotusarvo, keskihajonta sekä 90% symmetrinen Bayes-väli.

(Vihje: kyseessä on Gamma-jakauma, jonka odotusarvolla ja varianssilla on helpot kaavat. Välin laskemisessa voit käyttää R-funktiota `qgamma`).

3. Laadi R-funktio, joka palauttaa kutsulla `hdi_gamma(prob,shape,rate)` suurimman tiheyden välin Gamma-jakaumalle annetulle todennäköisyydelle `prob`, kun Gamma-jakauman parametrit ovat `shape` ja `rate`. Sovella tätä edelliskohdan tehtävään ja vertaa saatua väliä symmetriseen Bayes-väliin.

4. Olkoot havainnot $y = (y_1, ..., y_n), y_i|\theta \sim N(\theta, v)$, missä varianssi v on tunnettu ja y_i:t riippumattomia ehdolla θ. Parametrille θ oletetaan priori $\theta \sim N(m_0, w_0)$, missä m_0 ja w_0 ovat tunnettuja. Parametrin θ posteriorijakauman odotusarvo voidaan ymmärtää painotettuna keskiarvona priorijakauman odotusarvosta ja otoskeskiarvosta. Millä otoskoolla n datan vaikutus posteriorijakauman odotusarvoon on yhtä suuri kuin priorin vaikutus?

5. Oletetaan, että $y_i|\mu \sim N(\mu, 4.0)$. Asetetaan odotusarvon priorijakaumaksi $N(8.0, 9.0)$. Poimitaan $n = 9$ havainnon otos:

   ```
   18.36 16.17 10.21 11.07 18.32 6.03 18.42 8.04 13.53
   ```

   ($\bar{y} = 13.35$). Määritä posteriorijakauma μ:lle. Esitä myös 95 % Bayes-väli μ:lle.

6. Oletetaan, että $\theta \sim \text{Tas}(3, 5)$ ja $y|\theta \sim \text{Tas}(0, \theta)$. Määritä satunnaismuuttujan (havainnon) y reunajakauma.

7. (jatkoa edelliseen) Simuloi R:llä 1000 kpl havaintoa y:n reunajakaumasta ja tulosta histogrammi, jossa on suhteelliset frekvenssit. Tarkista edellisen tehtävän tulos lisäämällä histogrammiin tiheysfunktiota osoittava käyrä.

   (Vihje: Reunajakaumasta $p(y)$ simulointi on helpointa niin, että ensin generoi parametrin θ sen reunajakaumasta ja sitten havainnon y ehdollisesta jakaumasta $p(y|\theta)$).

8. Oletetaan, että satunnaismuuttujat X ja Y ovat diskreettejä. Todista seuraavat aputulokset:

$$\begin{aligned} \mathrm{E}(Y) &= \mathrm{E}[\mathrm{E}(Y|X)] \\ \mathrm{Var}(Y) &= \mathrm{Var}[\mathrm{E}(Y|X)] + \mathrm{E}[\mathrm{Var}(Y|X)] \end{aligned}$$