Aufgabe 1

Betrachten Sie wieder den Datensatz `babyboom.dat` mit den Geburtszeiten $t_1, \ldots, t_n$ (Minuten ab Mitternacht) von $n = 44$ Kindern, die am 18. Dezember 1997 in Brisbane, Australien auf die Welt kamen. Die Zeiten x_i zwischen den Geburten, d.h.

$$x_i = t_i - t_{i-1}, \quad t_0 \overset{!}{=} 0$$

können als i.i.d. Realisierungen einer Zufallsvariable $X \sim \text{Exp}(\lambda)$ mit unbekanntem Parameter $\lambda > 0$ modelliert werden.

(a) Zeigen Sie, dass die Familie der Gamma-Verteilungen

$$\mathcal{F} = \left\{ f(x) = \frac{b^a}{\Gamma(a)} x^{a-1} \exp(-bx) : \ a, b > 0 \right\}$$

konjugiert ist zur $\text{Exp}(\lambda)$-Verteilung und bestimmen Sie die Posteriori $p(\lambda|x)$.

(b) Berechnen Sie den theoretischen Posteriori-Erwartungswert $\mathbb{E}(\lambda|x)$ und den Posteriori-Modus $\arg\max_\lambda p(\lambda|x)$. Vergleichen Sie die Bayesianischen Schätzer mit dem ML-Schätzer (vgl. Blatt 1, Aufgabe 2).

(c) Zeichnen Sie die Posteriori für den `babyboom` Datensatz, wenn als Priori-Parameter $a = b = 10^{-3}$ gewählt werden. Visualisieren Sie die drei Schätzer aus Aufgabe b).

(d) Ziehen Sie $N = 200$ Realisierungen aus der Posteriori und bestimmen Sie den Posteriori-Erwartungswert, -Median und -Modus empirisch.

Hinweis: Verwenden Sie für den Modus den Befehl `density`.

Aufgabe 2

Sei wieder $X \sim \text{N}(\mu, \kappa^{-1})$ mit bekanntem $\mu \in \mathbb{R}$ und unbekannter Präzision $\kappa \sim \text{Ga}(a, b)$ und seien $x_1, \ldots, x_n$ i.i.d. Realisierungen von X.

Ein HPD-Intervall für κ zum Niveau $1 - \alpha$ sei gegeben durch $H = [\kappa_u, \kappa_o]$ mit $0 < \kappa_u < \kappa_o$.

(a) Welche Optimalitätseigenschaft haben HPD-Intervalle unter allen $(1 - \alpha)$-Kredibilitätsintervallen? Welche Beziehung gilt für $p(\kappa_u|x)$ und $p(\kappa_o|x)$?

Hinweis: Die Dichte $p(\kappa|x)$ ist unimodal und der Modus liegt nicht auf dem Rand des Parameterraums.

(b) Berechnen Sie auf Basis der Posteriori-Verteilung aus Aufgabe 2 von Blatt 2 ein 95%-HPD-Intervall für die Präzision κ und den `r-jy-daily.dat` Datensatz.

Verwenden Sie dazu die in Moodle verfügbare Funktion `gammaHPD`, die für verschiedene Quantile $u \in [0, \alpha]$ die Länge des zugehörigen 95%-Kredibilitätsintervalls berechnet und mittels `optimize()` die Länge minimiert. Vergleichen Sie das Ergebnis mit dem gleichendigen 95%-Kredibilitätsintervall.

Aufgabe 3

(a) Zeigen Sie, dass für die quadratische Verlustfunktion $L(d, \theta) = (d - \theta)^2$ der Posteriori-Erwartungswert $\mathbb{E}(\theta|x)$ Bayes-optimal ist.

Sei nun x eine Realisierung von einer Zufallsgröße $X \sim \mathrm{N}(\mu, 1)$. Für den unbekannten Erwartungswert werde a priori $\mu \sim \mathrm{N}(0, s),\ s > 0$ angenommen.

(b) Bestimmen Sie den Posteriori-Erwartungswert $d^* = \mathbb{E}(\mu|x)$.

(c) Bestimmen Sie das Bayes-Risiko $r^*(p) = r(d^*, p|x)$.