

# BAYESEKKSAMEN

December 11, 2015

## 1 Insekter

a)

I denne oppgaven antas det følgende funksjonaler av eksponensialfordelinga med parameter  $\theta$  er kjent: Forventinga ( $\theta^{-1}$ ), variansen ( $\theta^{-2}$ ) og medianen ( $\theta^{-1} \log 2$ ). La  $X \sim \exp(\theta)$  og sett  $Y = X^2$ . Da er  $P(Y \leq y) = P(X \leq \sqrt{y}) = 1 - \exp(-\theta\sqrt{y})$ . Deriver og få tettheten vi er ute etter:

$$f(y; \theta) = \frac{\theta}{2\sqrt{y}} e^{-\theta\sqrt{y}}.$$

Forventingen er  $EX^2 = \text{Var}X + (EX)^2 = 2\theta^{-2}$ , mens medianen finner vi ved å observere at  $P(X^2 \leq \alpha) = 0.5$  hvis og bare hvis  $P(X \leq \sqrt{\alpha}) = 0.5$ , så  $\text{median}(Y) = (\theta^{-1} \log 2)^2$ .

b)

Log likelihooden er

$$\sum_{i=1}^n l(y_i; \theta) = n \log \theta - \sum_{i=1}^n \theta \sqrt{y_i} - \sum_{i=1}^n \log(2\sqrt{y_i}).$$

Når vi deriverer med hensyn på  $\theta$  og setter resultatet lik 0 ender vi opp med

$$\hat{\theta} = \frac{n}{\sum_{i=1}^n \sqrt{y_i}}. \quad (1.1)$$

Det er velkjent at  $\sqrt{n}(\hat{\theta} - \theta_0) \xrightarrow{d} N(0, I^{-1})$ , hvor  $I = \theta^{-2}$  er Fisherinformasjonen. Dessuten er  $\sqrt{Y_i} \stackrel{i.i.d.}{\sim} \exp(\theta)$ , så  $\sum_{i=1}^n \sqrt{Y_i} \sim \text{Gamma}(n, \theta)$ . Dermed er  $n^{-1} \sum_{i=1}^n \sqrt{Y_i} \sim \text{Gamma}(n, \frac{\theta_0}{n})$ , og følgelig  $\hat{\theta} \sim \text{Inv-Gamma}(n, \frac{n}{\theta_0})$ .

c)

Etter Bayes' regel er

$$\begin{aligned} p(\theta | y) &= \left[ \prod_{i=1}^n y_i^{-\frac{1}{2}} \right] 2^{-n} \theta^n e^{-\theta \sum_{i=1}^n \sqrt{y_i}} \frac{b^a}{\Gamma(a)} \theta^{a-1} e^{-b\theta}, \\ &\propto \theta^{(a+n)-1} e^{-\theta(b + \sum_{i=1}^n \sqrt{y_i})}, \end{aligned}$$

hvor proporsjonaliteten er i  $\theta$ . Det følger at posterioren er fordelt som  $\text{Gamma}(a + n, b + \sum_{i=1}^n \sqrt{y_i})$ .

d)

Nå skal vi finne  $p(y) = \frac{p(y|\theta)p(\theta)}{p(\theta|y)}$ . Siden  $p(\theta | y)$  har proporsjonalitetskonstant  $\frac{(b+\sqrt{y})^{a+1}}{\Gamma(a+1)}$ , og proporsjonalitetskonstanten (m.h.p  $\theta$ ) i  $p(y | \theta)p(\theta)$  er gitt ved  $\frac{b^a}{\sqrt{y}2\Gamma(a)}$ , er

$$\begin{aligned} p(y) &= \frac{b^a}{\sqrt{y}2\Gamma(a)} \frac{\Gamma(a+1)}{(b+\sqrt{y})^{a+1}} \\ &= \frac{ab^a}{2\sqrt{y}(b+\sqrt{y})^{a+1}}. \end{aligned}$$

Her har vi brukt at  $\Gamma(a+1) = a\Gamma(a)$ . Når  $(a, b) = (1, 1)$  får vi  $p(y) = \frac{1}{2\sqrt{y}(1+\sqrt{y})^2}$ . For å finne fordelingsfunksjonen bruker vi substitusjonen  $u = \sqrt{y}$ , så  $dy = 2udu$ , og  $F(y) = \int_0^{\sqrt{y}} \frac{1}{(1+u)^2} du = 1 - (1 + \sqrt{y})^{-1}$ .

e)

Med enkle manipulasjoner kan det vises at  $(a, b)$  tilfredsstiller de oppgitte kvantilulikhetene hvis og bare hvis

$$P(\theta \leq 1) = 0.10 \text{ og } P(\theta \leq \sqrt{10}) = 0.90.$$

Den andre liketheten følger av  $\frac{2}{\theta^2} \leq 0.2 \iff \theta \geq \sqrt{10}$ , ettersom Gamma lever på  $[0, \infty)$ . Den andre likheten vises tilsvarende. Nå kan vi bruke R for å finne gode parameterestimer.

---

```
> f = function(p) (qgamma(0.1, p[1], p[2]) - 1)^2 +
> (qgamma(0.9, p[1], p[2]) - sqrt(10))^2
> nlm(f, c(4, 4))$estimate
[1] 5.315377 2.656497
```

---

I resten av oppgaven bruker vi verdiene  $a = 5.31$  og  $b = 2.66$ . Denne prioren har topppunkt i 1.62, snitt  $\approx 2$  og .9-kvantil i  $\approx 4.5$ .

f)

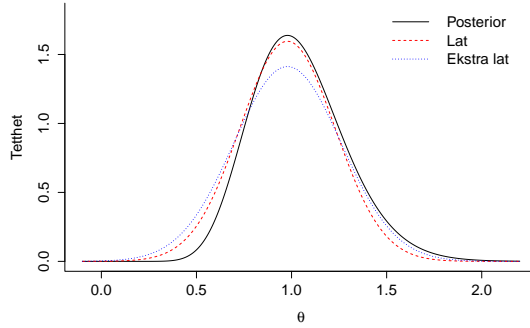
ML-estimatet er  $\hat{\theta} = 0.856$ , det følger av (1.1). Siden forventinga til en  $\text{Gamma}(a, b)$ -fordelt variabel er  $\frac{a}{b}$ , følger det av resultatet  $p(\theta | y) = \Gamma(a + n, b + \sum_{i=1}^n \sqrt{y_i})$  at posteriorsnittet er

$$\frac{5.31 + 12}{2.66 + 14.02} = 1.04.$$

Posteriorsnittet er høyere enn ML-estimatet, da prioren legger mye vekt til høyre for  $\hat{\theta} = 0.856$ . Posteriotoppunktet er  $\frac{a-1}{b} = 0.978$ , som er litt nærmere ML-estimatet.

g)

De late Bayesianerne bruker approksimasjonen  $p(\theta | y) = N(\theta^\pi, (I^\pi)^{-1})$ , hvor  $\theta^\pi$  er posteriotoppunktet og  $I^\pi$  er den generaliserte Fisherinformasjonen. Nå kan vi enten være ekstra late, og anta at  $\pi$  er flat, som gir oss  $I = \frac{d^2}{d\theta^2} = \theta^{-2}$ , eller være passe late og få  $I^\pi = a\theta^{-2}$ . Siden  $\frac{d^2}{d\theta^2} \log f(\theta) = \theta^{-2}$ ,  $\frac{d^2}{d\theta^2} \log [p(\theta | \mathbf{y})] = -\theta^{-2}(a + n - 2)$  og  $\hat{\theta} = 0.978$  får vi henholdsvis  $N(0.978, \frac{1}{n}0.978^2)$  og  $N(0.978, \frac{1}{a+n-2}0.978^2)$ . Begge approksimasjonene, som kan ses i figur 1.1, ser ut til å funke bra. Ikke overraskende funker den late approksimasjonen bedre enn den ekstra late approksimasjonen.



**Figure 1.1:** Plot av posterioren til  $\theta$  sammen med late og ekstra late varianter.

h)

La oss finne den posteriore prediktive tettheten i  $y$ .

$$\begin{aligned} E_{\theta|\mathbf{y}}(p(y|\theta)) &= \int_0^\infty p(y|\theta)p(\theta|\mathbf{y})d\theta, \\ &= \int_0^\infty \frac{\theta}{2\sqrt{y}} e^{-\theta\sqrt{y}} \gamma(\theta) d\theta. \end{aligned}$$

Her er  $\gamma(\theta)$  tettheten til  $\text{Gamma}(\alpha, \beta)$ , hvor  $\alpha = a + n$  og  $\beta = b + \sum_{i=1}^n \sqrt{y_i}$ . Integranden kan identifiseres som

$$\frac{1}{2\sqrt{y}} \frac{\Gamma(\alpha+1)\beta^\alpha}{\Gamma(\alpha)(\beta+\sqrt{y})^{\alpha+1}} \gamma(\theta),$$

hvor  $\gamma'(\theta)$  er tettheten til  $\text{Gamma}(\alpha+1, \beta+\sqrt{y})$ . Bruker vi  $\Gamma(x+1) = x\Gamma(x)$  og  $\int_0^\infty \gamma'(\theta) d\theta = 1$  får vi

$$E_{\theta|\mathbf{y}}(p(y|\theta)) = \frac{\alpha}{2\sqrt{y}} \frac{\beta^\alpha}{(\beta+\sqrt{y})^{\alpha+1}}.$$

For å finne fordelingsfunksjonen bruker vi substitusjonen  $u = \sqrt{y}$ , som gir oss integralet

$$\begin{aligned} \int_0^{\sqrt{y}} \frac{\alpha\beta^\alpha}{(\beta+u)^{\alpha+1}} du &= \left[ -\frac{\beta^\alpha}{(\beta+u)^\alpha} \right]_0^{\sqrt{y}}, \\ &= 1 - \frac{\beta^\alpha}{(\beta+\sqrt{y})^\alpha}. \end{aligned}$$

Husk at  $F(\theta) = 1 - e^{-\theta\sqrt{y}}$ , at ML-estimatet til  $\theta$  er  $\hat{\theta}_{ML} = n(\sum_{i=1}^n \sqrt{y_i})^{-1}$  og at  $e^{-x} = \lim_{n \rightarrow \infty} (1 + \frac{x}{n})^{-n}$ . Vi skriver om  $\frac{\beta^\alpha}{(\beta+\sqrt{y})^\alpha}$  til  $\left(1 + \frac{\sqrt{y}}{\beta}\right)^{-\alpha}$  og bytter tilbake  $\alpha$  og  $\beta$  for å få

$$\begin{aligned} \left(1 + \frac{\sqrt{y}}{n} \frac{n}{b + \sum_{i=1}^n \sqrt{y_i}}\right)^{-(a+n)} &= \left(1 + \frac{\sqrt{y}}{n} \left(\frac{b}{n} + 1/\hat{\theta}_{ML}\right)^{-1}\right)^{-(a+n)}, \\ &\approx \exp(-\sqrt{y}\hat{\theta}_{ML}). \end{aligned}$$

**Table 1:** Kvantiler for  $E_{\theta|\mathbf{y}}(p(y | \theta))$  med  $(a, b) = (5.31, 2.66)$ ,  $(a, b) = (4, 4)$  og fordelinga  $\exp(\hat{\theta}_{ML})$ .

	$(5.31, 2.66)$	$(4, 4)$	$\exp(\hat{\theta}_{ML})$
0.1	0.0104	0.0142	0.1231
0.5	0.4637	0.6363	0.8096
0.9	5.6233	7.7765	2.6895

Hvor “ $\approx$ ” krever at  $n$  er veldig stor. Det er mer på grunn av tregheten i konvergensen  $(1 + \frac{x}{n})^{-n} \rightarrow \exp(-x)$  enn på grunn av prior-parameterne  $(a, b)$ . Dette kan en se i haleatferden til fordelinga. La  $a > 0$  og  $b > 0$  være gitt, og betrakt

$$\int_0^\infty \frac{\alpha y^k}{2\sqrt{y}} \frac{\beta^\alpha}{(\beta + \sqrt{y})^{\alpha+1}} dy,$$

som er det  $k$ -te momentet til  $E_{\theta|\mathbf{y}}(p(y | \theta))$ . Bruk samme substitusjon som over for å få

$$\int_0^\infty u^{2k} \frac{\alpha \beta^\alpha}{(\beta + u)^{\alpha+1}} dy.$$

Anta  $u \gg \beta$ , slik at  $u^{\alpha+1}$  dominerer de andre termene i  $(\beta + u)^{\alpha+1}$  fullstendig. For slike  $u$  er integralet omtrentlig proporsjonalt med

$$\frac{u^{2k}}{u^{\alpha+1}} = u^{2k-a-n-1}.$$

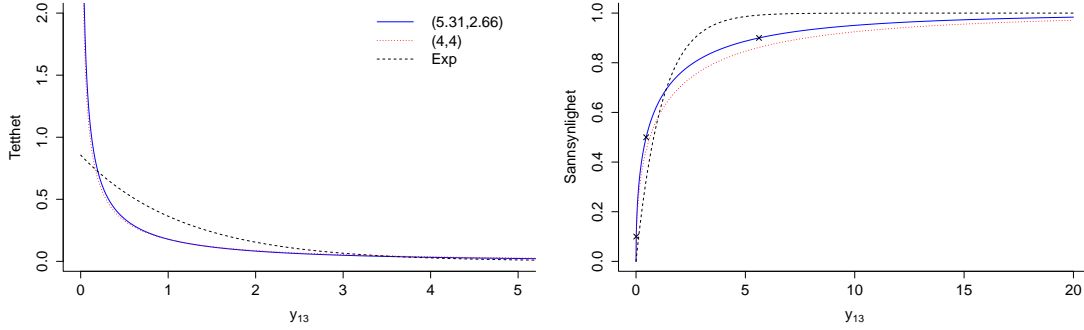
Integralet av denne termen diverger hvis og bare hvis om  $2k > a + n$ , noe som også gjelder for vårt opprinnelige integral. Dette er i skarp kontrast til eksponensialfordelinga, som selvfølgelig har endelige momenter for enhver  $k$ . Momentatferden kun avhengig av  $a$  og  $n$ , ikke de faktiske observasjonene i  $\mathbf{y}$ .

Siden fordelingsfunksjonen er såpass fin, er det enkelt å finne kvantilfunksjonen: Sett  $1 - \frac{\beta^\alpha}{(\beta + \sqrt{y})^\alpha} = z$ , løs for  $y$ , og end opp med  $y = \beta^2(1 - (1 - z)^{-\frac{1}{\alpha}})^2$ . Oppsummert er

$$\begin{aligned} h(y) &= \frac{\alpha}{2\sqrt{y}} \frac{\beta^\alpha}{(\beta + \sqrt{y})^{\alpha+1}}, \\ H(y) &= 1 - \left(1 + \frac{\sqrt{y}}{\beta}\right)^{-\alpha}, \\ Q(y) &= \beta^2(1 - (1 - z)^{-\frac{1}{\alpha}})^2, \end{aligned}$$

hvor  $h(y) = E_{\theta|\mathbf{y}}(p(y | \theta))$ ,  $H$  den tilhørende fordelingsfunksjonen og  $Q$  den tilhørende kvantilfunksjonen. Med den siste formelen er det enkelt å finne kvantilene det ble spurt om i oppgaven, som vi putter i Tabell 1.

Vi viser tettheten og kvantilfunksjonen for det aktuelle datasettet i figur 1.2. I tillegg til  $(a, b) = (5.31, 2.66)$  illustrer vi valget  $(a, b) = (4, 4)$ . Dessuten viser vi tettheten  $\exp(\hat{\theta}_{ML})$ . Som allerede indikert har tettheten  $\exp(\hat{\theta}_{ML})$  seriøst tynnere haler enn  $E_{\theta|\mathbf{y}}(p(y | \theta))$ , noe som kan forklares ved at den ikke tar hensyn til usikkerheten i estimatet  $\hat{\theta}_{ML}$ .



**Figure 1.2:** (venstre) Posterior prediktiv tetthet til  $y_{13}$ , den røde linja kan knapt ses. (høyre) Kumulativ tetthet. Kvantilene 0.1, 0.5 og 0.9 som hører til  $(a, b) = (5.31, 2.66)$  er markert.

## 2 Hierarkisk modell

a)

La  $\mathbf{y} = y_1, \dots, y_n$  i.i.d. observasjoner fra  $f(\theta)$ . Posterioren  $p(\theta | \mathbf{y})$  er proporsjonal med

$$\prod_{i=1}^n f(y_i | \theta) \pi(\theta) \propto \theta^{a+n-1} (1-\theta)^{b+\langle \mathbf{y}, 1 \rangle - n - 1},$$

så  $p(\theta | \mathbf{y}) \sim \text{Beta}(a+n, b+\langle \mathbf{y}, 1 \rangle - n)$ . Det følger at

$$\begin{aligned} p(\theta | y) &= \text{Beta}(a+1, b+y-1) \\ &= \frac{\theta^a (1-\theta)^{b+y-2}}{B(a+1, b+y-1)}. \end{aligned} \quad (2.1)$$

Siden forventinga til  $\text{Beta}(\alpha, \beta)$  er  $\frac{\alpha}{\alpha+\beta}$ , er posterior forventning lik  $\frac{a+n}{a+b+\langle \mathbf{y}, 1 \rangle}$ , mens posterior topppunkt  $(\frac{\alpha-1}{\alpha+\beta-1})$  er  $\frac{a+n-1}{a+b+\langle \mathbf{y}, 1 \rangle - 2}$ . Da

$$\log \prod_{i=1}^n f(y_i; \theta) = \sum_{i=1}^n [(y_i - 1) \log(1 - \theta)] + n \log(\theta)$$

har derivert  $-\sum_{i=1}^n \frac{y_i - 1}{1 - \theta} + \frac{n}{\theta}$ , er ML-estimatet lik  $\hat{\theta}_{ML} = \frac{n}{\langle \mathbf{y}, 1 \rangle}$ .

b)

Først beregner vi marginalfordelinga  $p(\mathbf{y})$ . Vi bruker proporsjonaliteten  $p(\mathbf{y} | \theta) p(\theta) \propto p(\theta | \mathbf{y})$  for å finne

$$\begin{aligned} p(\mathbf{y}) &= \frac{p(\mathbf{y} | \theta) p(\theta)}{p(\theta | \mathbf{y})}, \\ &= \frac{B(a+n, b+\langle \mathbf{y}, 1 \rangle - n)}{B(a, b)}. \end{aligned} \quad (2.2)$$

Her dukker betafunksjonen  $B(x, y) = \frac{\Gamma(x)\Gamma(y)}{\Gamma(x+y)}$  opp ettersom den er proporsjonalitetskonstanten i betafordelingen. Det er kjent at  $\mathcal{U}(0, 1) = \text{Beta}(1, 1)$ , så etter (2.1) er  $p(\theta | y) \sim \text{Beta}(2, y)$  i dette tilfellet. Siden  $\Gamma(n+1) = n!$  og  $\text{Beta}(1, 1) = 1$  er den ubetingede fordelinga  $p(y)$  gitt ved

$$\begin{aligned}
\frac{p(y | \theta)\pi(\theta)}{p(\theta | y)} &= \frac{B(a+1, b+y-1)}{B(a, b)} \\
&= \frac{\Gamma(2)\Gamma(y)}{\Gamma(2+y)} \\
&= \frac{(y-1)!}{(y+1)!} = \frac{1}{y(y+1)}.
\end{aligned}$$

I det andre spørsmålet er vi ute etter  $\sum_{i=1}^{\infty} yp(y)$ . Denne summen divergerer, ettersom  $\sum_{y=1}^{\infty} \frac{1}{y+1} = \infty$ . Forventet antall knips er dermed uendelig.

c)

Vi finner Fisherinformasjonen til  $f(y; \theta)$  via

$$\frac{d^2}{d\theta^2}((y-1)\log(1-\theta) + \log(\theta)) = -\frac{y-1}{(1-\theta)^2} - \frac{1}{\theta^2}.$$

Dermed er  $E\left[-\frac{d^2}{d\theta^2}l\right] = \frac{1}{\theta(1-\theta)^2} - \frac{1}{(1-\theta)^2} + \frac{1}{\theta^2}$ , som forenklet blir  $\frac{\theta-\theta^2+(1-\theta^2)}{\theta^2(1-\theta)^2} = \theta^{-2}(1-\theta)^{-1}$ . Dermed er Jeffreys prior proporsjonal med  $\text{Beta}(0, \frac{1}{2})$ . Dette er en *improper* prior, ettersom  $\theta^{-1}$  vokser for fort når  $\theta \rightarrow 0$ . Prioren legger altså ekstremt stor vekt på små  $\theta$ -verdier og rimelig stor vekt på store  $\theta$ -verdier.

Posterioren er proporsjonal med  $(1-\theta)^{y-\frac{1}{2}-1}$ , som svarer til  $\text{Beta}(1, y - \frac{1}{2})$ -fordelinga. Denne har forventning i  $(y + \frac{1}{2})^{-1}$ , som er en mindre verdi enn ML-estimatet  $y^{-1}$ . Dette veier opp for en slags “bias” når  $y = 1$ , hvor ML er ganske ekstrem og sier at 1 er det beste estimatet, noe som er litt søkt. Posterioren er forskjøvet mot venstre, ettersom prioren var det.

d)

Vi har allerede vist at  $p(\theta | y) \sim \text{Beta}(2, y)$  når  $(a, b) = (1, 1)$ . Siden forventinga til Betafordelinga er gitt ved  $\frac{a}{a+b}$ , er posterior forventning lik  $\frac{1}{1+\frac{y}{2}}$ . Vi simulerer kvardattapsrisikoen i R og presenterer kurven i figur 2.1. Risikoen til både Bayesestimatorne med både uniform og Jeffreys prior ser mye bedre ut enn MLs. Ikke overraskende er ML best når  $\theta$  er nærme 1: For ML-estimatet er større enn posterior forventning for både uniform og Jeffreys.

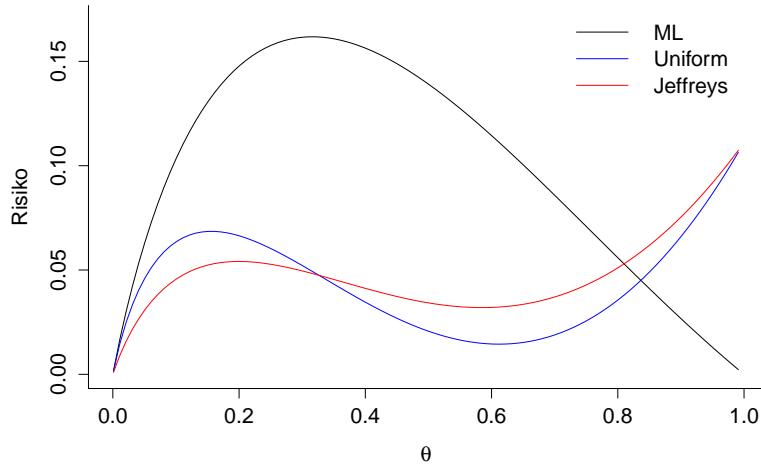
Når  $a + b = k \in \mathbb{N}$  kan en uttrykke risikoen ved hjelp av polylogaritmen  $\text{Li}_2$  og logaritmen. Om  $a + b \notin \mathbb{N}$  kan den uttrykkes ved hjelp av Lerchs transcendent. Dette blir klart om en putter uttrykket  $\sum_{i=1}^{\infty} (\frac{a}{a+b+i} - x)^2 x^i$  inn i Mathematica. Hvis en ønsker å analysere dypere hvordan riskoen endrer seg med  $a, b$  (og kanskje  $n$ ?) er det i det minste mulig å uttrykke som det et kombinasjon av elementære funksjoner og Lerchs transcendent.

e)

Nå skal vi finne  $\arg\min_{\delta} \int (\delta - \theta)^2 p(\theta | y) d\theta$ . Om vi ekspanderer  $(\delta - \theta)^2$ , deriverer med hensyn på  $\delta$  og setter uttrykket lik 0 får vi  $2E_{\theta|y}\theta = 2\delta$ . Dermed er posterior forventning Bayesestimatoren under kvadratfeiltapet. Siden  $p(\theta | y_i) \sim \text{Beta}(a+1, b+y_i-1)$  etter (2.1) og forventinga til  $\text{Beta}(a, b)$  er  $\frac{a}{a+b}$ , er posterior forventning er  $\frac{a+1}{a+b+y_i}$ .

f)

Vi har allerede vist at  $p(y) = \frac{\Gamma(a+b)\Gamma(a+1)\Gamma(b+y-1)}{\Gamma(a)\Gamma(b)\Gamma(a+b+y)}$ , se likning (2.2). Nå bruker vi  $\Gamma$ -egenskapen  $\Gamma(z+1) = z\Gamma(z)$  for alle  $z$  og får



**Figure 2.1:** Sammenligning av kvadratriskoen for de tre forskjellige estimatorne. ML er bedre enn uniform Bayes når  $\theta > 0.84$ .

$$p(y) = a \frac{\Gamma(a+b)\Gamma(b+y-1)}{\Gamma(b)\Gamma(a+b+y)}.$$

For å finne parameterestimatene bruker vi formuleringen  $\frac{B(a+1, b+y-1)}{B(a, b)}$ , sammen med R-koden:

```

1 ys = c(1, 1, 2, 2, 2, 2, 2, 3, 3, 3,
2       3, 3, 5, 6, 6, 14, 15, 15, 20, 75)
3
4 llik = function(p, ys){
5   a = p[1]
6   b = p[2]
7   sum(lbeta(a+n, b+ys-n) - lbeta(a, b))
8 }
9
10 ab = nlm(function(p) -llik(p, ys), p=c(1,1))$estimate
```

Dette gir oss parameterestimer på omtrent  $a = 1.71$  og  $b = 6.35$ . Dermed er posterior forventning lik  $\frac{1.71}{6.35+y_i}$ . De resulterende verdiene er i tabell 3.

**g)**

I likning (2.2) har vi tettheten til  $p(y | a, b)$ . Vi kan bruke denne til å finne

$$p(a, b | \mathbf{y}) \propto \prod_{i=1}^n \frac{B(a+1, b+y_i-1)}{B(a, b)} 1_{[0,100]}(a) 1_{[0,100]}(b).$$

Det er en smal sak å sample fra denne fordelinga. Vi trekker  $N$  observasjoner fra  $\mathcal{U}([0, 100] \times [0, 100])$ , beregner  $\prod_{i=1}^n \frac{B(a+1, b+y_i-1)}{B(a, b)}$  for hver av dem, og avslutter med å dele på snittet av alle observasjonene. Da har vi en sannsynlighet forbundet med hver observasjon  $(a, b)$ , som vi kan bruke til å beregne de ønskede verdiene (i tabell 2).

**Table 2:** Tabell over posterior snitt, standardavvik og korrelasjon.  $N = 1000000$ . De store verdiene på  $a, b$  kan oversettes til spissere priorer enn det vi fikk for  $a = 1.71$  og  $b = 6.35$ .

	$a$	$b$
Snitt	7.43	44.5
SD	4.41	28.1
Korrelasjon	0.894	

**Table 3:** Tabell over verdier.

	$y_i$	$\hat{\theta}_i$	$\theta_i^*$	$\theta_i^{0.05}$	$\theta_i^{0.50}$	$\theta_i^{0.95}$		$y_i$	$\hat{\theta}_i$	$\theta_i^*$	$\theta_i^{0.05}$	$\theta_i^{0.50}$	$\theta_i^{0.95}$
1	1	1	0.299	0.077	0.164	0.36	11	3	0.333	0.245	0.073	0.156	0.319
2	1	1	0.299	0.077	0.164	0.36	12	3	0.333	0.245	0.073	0.156	0.319
3	2	0.5	0.27	0.075	0.16	0.337	13	5	0.2	0.208	0.068	0.148	0.29
4	2	0.5	0.27	0.075	0.16	0.337	14	6	0.167	0.193	0.067	0.145	0.279
5	2	0.5	0.27	0.075	0.16	0.337	15	6	0.167	0.193	0.067	0.145	0.279
6	2	0.5	0.27	0.075	0.16	0.337	16	14	0.071	0.123	0.053	0.123	0.221
7	2	0.5	0.27	0.075	0.16	0.337	17	15	0.067	0.118	0.052	0.121	0.216
8	3	0.333	0.245	0.073	0.156	0.319	18	15	0.067	0.118	0.052	0.121	0.216
9	3	0.333	0.245	0.073	0.156	0.319	19	20	0.05	0.097	0.046	0.111	0.197
10	3	0.333	0.245	0.073	0.156	0.319	20	75	0.013	0.033	0.019	0.059	0.114

**h)**

La  $b'(x; a, b)$  være betatettheten med parametre  $a, b$ . Siden

$$\begin{aligned}
 p(\theta_1, \theta_2, \dots, \theta_n \mid \mathbf{y}, a, b) &= \prod_{i=1}^n b'(\theta_i; a+1, b+y_i-1) \\
 &= \prod_{i=1}^n p(\theta_i \mid y_i, a, b),
 \end{aligned}$$

er  $p(\theta_i \mid \mathbf{y}, a, b) = p(\theta_i \mid y_i, a, b)$ . Dette er intuitivt opplagt, ettersom all informasjonen  $y_j$ ,  $j \neq i$  kan gi om  $\theta_i$  er informasjon om hyperparameterne  $a, b$ . Nå bruker vi

$$\begin{aligned}
 p(\theta_i \mid \mathbf{y}) &= \int p(\theta_i \mid \mathbf{y}, a, b) p(a, b \mid \mathbf{y}) da db \\
 &= \int p(\theta_i \mid y_i, a, b) p(a, b \mid \mathbf{y}) da db,
 \end{aligned}$$

som gjør det mulig å finne de ønskede kvantilene ved å bruke verdiene fra forrige oppgave. Verdiene presenteres i tabell 3. Under den hierarkiske modelleringa får vi større “shrinkage” på parameterne enn i empirisk bayes.

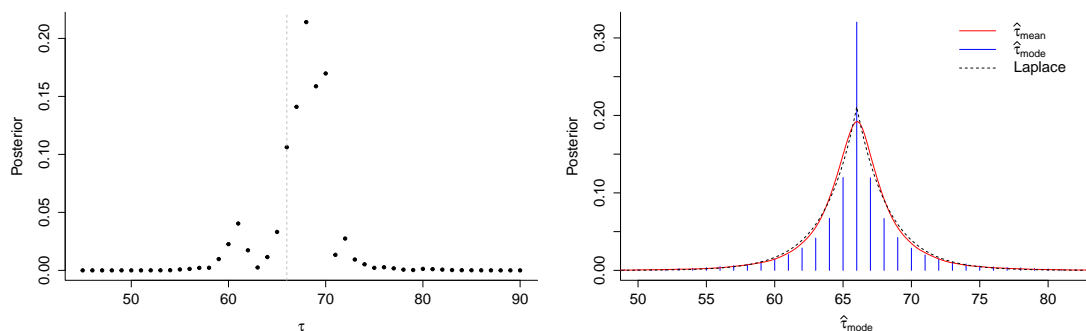
### 3 Bruddpunkt

**a)**

Tettheten er gitt ved

$$p(\mathbf{y} \mid \tau) = \prod_{i=1}^{\tau} f_L(y_i) \prod_{i=\tau+1}^n f_R(y_i),$$





**Figure 3.1:** (venstre) Posteriorfordeling for  $\tau \mid \mathbf{y}$ . Den treffer ganske bra. (høyre) Samplingfordelingene til posteriortoppunktet og posterior forventning, sammen med Laplacefordelinga med parametere estimert gjennom ML. Tettheten er plottet ved hjelp av `density`-funksjonen i R. Dn røde linja ser glattere ut enn den stiplede i toppunktet. Dette er en konsekvens av glattingen som kommer gjennom kjernetetthetsestimeringen, ikke en egenskap ved tettheten til posterior forventning.

så posterioren er

$$p(\tau \mid \mathbf{y}) = C_{\mathbf{y}} \prod_{i=1}^{\tau} f_L(y_i) \prod_{i=\tau+1}^n f_R(y_i),$$

hvor  $C_{\mathbf{y}} = \sum_{\tau=1}^n \prod_{i=1}^{\tau} f_L(y_i) \prod_{i=\tau+1}^n f_R(y_i)$ .

c)

Vi kjører en simulering med  $n = 100$  og får posterioren i figur 3.1. Toppunktet er ganske nærme den sanne  $\tau = 66$ , men posterioren er temmelig uelegant — den er asymmetrisk med flere “topppunkter”. Slik atferd ser ut til å være vanlig. For å forstå litt av de frekventistiske egenskapene til modellen kjører vi forsøket  $N = 10000$  ganger, og plotter tettheten til posterior forventning og stolpediagrammet til posteriortoppunktet. Se figur 3.1 igjen. Interessant nok ser posterior forventning ut som den er Laplacefordelt, mens toppunktet kanskje er “symmetrisk geometrisk” fordelt? Laplacefordelinga er tykkere haler enn normalfordelinga, men er spissere rundt snittet. Dette viser at ikke er veldig usannynlig å få estimater som gir veldig galt inntrykk, men likevel vil de fleste observasjonene gi et veldig godt og presist inntrykk.

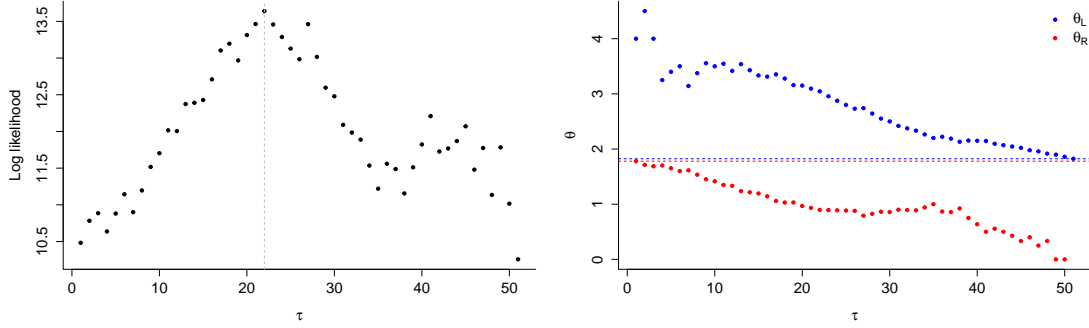
d)

Log likelihooden er  $\sum_{i=1}^{\tau} \log f_L(y_i; \theta_L) + \sum_{i=\tau+1}^n \log f_R(y_i; \theta_R)$ . En naiv men okay måte å maksimere denne på er iterere gjennom alle lovlige verdier av  $\tau$ , hvor vi på hvert steg beregner log likelihooden på hver side ved å plugge in  $\widehat{\theta}_L = \tau^{-1} \sum_{i=1}^{\tau} y_i$  og  $\widehat{\theta}_R = (n - \tau)^{-1} \sum_{i=\tau+1}^n y_i$ . Dette er enkelt å programmere i R. Løsningen blir  $(\tau, \theta_L, \theta_R) = (22, 3.05, 0.896)$ . I figur 3.2 illustreres det hvordan log likelihooden ser ut, i tillegg til hvordan ML-estimatene  $\widehat{\theta}_L$  og  $\widehat{\theta}_R$  endres ettersom bruddpunktet  $\tau$  forandres.

e)

Via kjerneregelen for kondisjonering får vi

$$p(\tau, \theta_L, \theta_R, \mathbf{y}) = p(\mathbf{y} \mid \tau, \theta_L, \theta_R) p(\tau, \theta_L, \theta_R).$$



**Figure 3.2:** (venstre) Plot over max-log likelihooden for ulike valg av  $\tau$ . (høyre) Her kan vi se hvordan  $\theta_L$  og  $\theta_R$  forandrer seg ettersom  $\tau$  forandres. Legg merke til at  $\theta_L > \theta_R$  for alle valg av  $\tau$ , noe som støtter hypotesen om at det finnes ett bruddpunkt.

Leddet  $p(\mathbf{y} \mid \tau, \theta_L, \theta_R)$  er kjent, mens  $p(\tau, \theta_L, \theta_R) = n^{-1} 1_{\{1, \dots, n\}}(\tau) \gamma(\theta_L; a, b) \gamma(\theta_R; a, b)$ , hvor  $\gamma(x; a, b)$  er Gamma( $a, b$ )-tettheten. Skrevet ut får vi

$$p(\tau, \theta_L, \theta_R, \mathbf{y}) \propto \frac{b^{2a}}{n \prod_{i=1}^n y_i! (\Gamma(a))^2} \theta_R^{\sum_{i=\tau+1}^n y_i + a - 1} \exp(-\theta_R(n - \tau + b)) \theta_L^{\sum_{i=1}^{\tau} y_i + a - 1} \exp(-\theta_L(\tau + b)). \quad (3.1)$$

Vi kan identifisere de ulike betingede fordelingene ved å finne proporsjonalitetskonstantene. For å finne  $p(\tau \mid \mathbf{y})$  integrerer vi ut  $\theta_L, \theta_R$ , hvor vi bruker at proporsjonalitetskonstanten til  $\theta_L^{\sum_{i=1}^{\tau} y_i + a - 1} \exp(-\theta_L(\tau + b))$  er og  $\theta_R^{\sum_{i=\tau+1}^n y_i + a - 1} \exp(-\theta_R(n - \tau + b))$  er kjente:

$$p(\tau \mid \mathbf{y}) \propto \frac{b^{2a}}{n \prod_{i=1}^n y_i! (\Gamma(a))^2} \frac{\Gamma(\sum_{i=1}^{\tau} y_i + a)}{(\tau + b)^{\sum_{i=1}^{\tau} y_i + a}} \frac{\Gamma(\sum_{i=\tau+1}^n y_i + a)}{(n - \tau + b)^{\sum_{i=\tau+1}^n y_i + a}}. \quad (3.2)$$

Fra dette følger det også at  $p(\mathbf{y})$  har følgende flotte fordeling:

$$p(\mathbf{y}) = \sum_{\tau=1}^n \left[ \frac{b^{2a}}{n \prod_{i=1}^n y_i! (\Gamma(a))^2} \frac{\Gamma(\sum_{i=1}^{\tau} y_i + a)}{(\tau + b)^{\sum_{i=1}^{\tau} y_i + a}} \frac{\Gamma(\sum_{i=\tau+1}^n y_i + a)}{(n - \tau + b)^{\sum_{i=\tau+1}^n y_i + a}} \right].$$

Når  $(a, b) = (1, 1)$  blir posterioren  $p(\tau \mid \mathbf{y})$  til

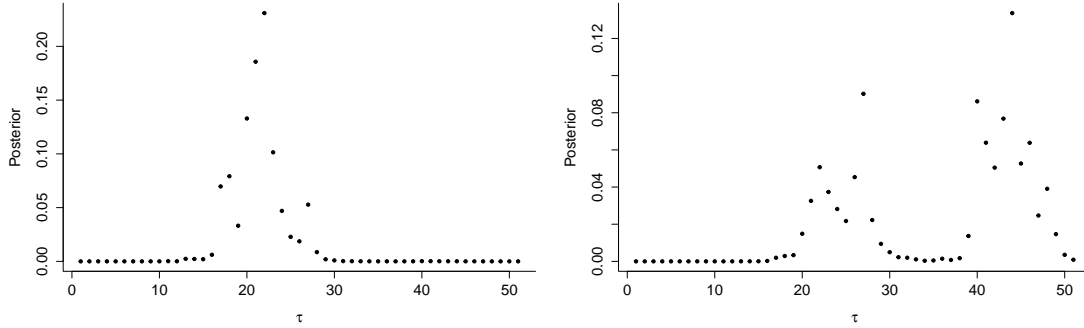
$$p(\tau \mid \mathbf{y}) \propto \frac{1}{n \prod_{i=1}^n y_i!} \frac{\Gamma(\sum_{i=1}^{\tau} y_i + 1)}{(\tau + 1)^{\sum_{i=1}^{\tau} y_i + 1}} \frac{\Gamma(\sum_{i=\tau+1}^n y_i + 1)}{(n - \tau + 1)^{\sum_{i=\tau+1}^n y_i + 1}}.$$

Det går raskt å beregne denne fordelinga  $\mathbf{R}$ , og presenterer resultatene i figur 3.3. Her tar vi også med posterioren hvor  $(a, b) = (2, 10)$  for å illustrere at andre ting kan skje. Her anslås bruddpunktet å være der hvor dataene slutter å ha verdier større enn 1. Toppunktet  $\tau \approx 22$  svarer heller til en mindre dramatisk endring fra  $\theta_L$  til  $\theta_R$ .

**f)**

Om  $(\theta_L^1, \theta_R^1), (\theta_L^2, \theta_R^2), \dots, (\theta_L^N, \theta_R^N)$ s er en trekning fra  $p(\theta_L, \theta_R \mid \mathbf{y})$  kan vi approksimere fordelingen til  $\frac{\theta_R}{\theta_L} \mid \mathbf{y}$  med den empiriske fordelinga til  $\left( \frac{\theta_R^i}{\theta_L^i} \right)_{i=1}^n$ . Etter (3.1) er

$$p(\theta_L, \theta_R \mid \mathbf{y}, \tau) = \gamma(\theta_L, \sum_{i=1}^{\tau} y_i + 1, \tau + 1) \gamma(\theta_R, \sum_{i=\tau+1}^n y_i + 1, n - \tau + 1).$$



**Figure 3.3:** (venstre) Posterioren  $p(\tau | \mathbf{y})$ , hvor  $\theta_L, \theta_R$  er uavhengig  $\exp(1)$ -fordelte og  $\tau$  er diskret uniform på  $1, \dots, 51$ . (høyre) Samme posterior, hvor  $\theta_L, \theta_R$  er uavhengig  $\text{Gamma}(2, 10)$ -fordelte. Eksponensialpriorer legger all vekten rundt bumpen i likelihooden omkring  $\tau = 20$ , mens  $\text{Gamma}(2, 10)$  finner for godt å prioritere det lokale maksimumet omkring  $\tau = 45$ .

Kall  $p(\tau = i | \mathbf{y})$  for  $t_i$ . Da er  $p(\theta_L, \theta_R | \mathbf{y}) = \sum_{i=1}^n t_i p(\theta_L, \theta_R | \mathbf{y}, \tau)$ . Vi kan sample fra denne ved først å “sample” fra  $\tau | \mathbf{y}$  og deretter fra  $\gamma(\theta_L, \sum_{i=1}^{\tau} y_i + 1, \tau + 1)$  og  $\gamma(\theta_R, \sum_{i=\tau+1}^n y_i + 1, n - \tau + 1)$ . Siden  $\tau | \mathbf{y}$  har domene  $1, 2, \dots, n$ , er det ingen vits i å sample fra den, vi trenger bare å ta antall  $\tau = i$  omtrent lik  $N t_i$ , hvor  $N$  er antall samples.

---

**Algorithm 1** Program i R for å beregne posterioren  $\frac{\theta_R}{\theta_L} | \mathbf{y}$ .

---

```

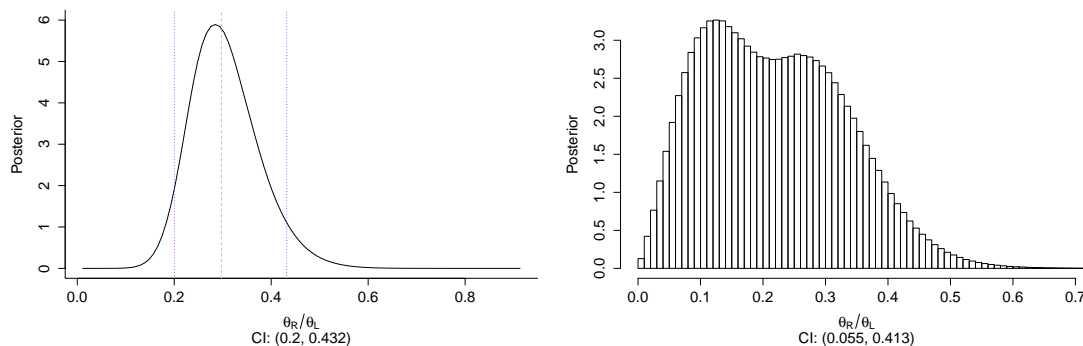
1 # tis er vektoren med posterior sannsynligheter for tau.
2
3 a = 1
4 b = 1
5 N = 1000000
6 counts = floor(tis*N)
7
8 sampleGammas = function(tau, count){
9     sum1 = sum(data[1:tau])
10    sum2 = sum(data[(tau+1):n+1], na.rm=TRUE)
11    # Sikrer at tau = 51 også fungerer.
12    thetaR = rgamma(count, sum2+a, n-tau+b)
13    thetaL = rgamma(count, sum1+a, tau+b)
14    thetaR/thetaL
15 }
16
17 result = unlist(sapply(1:51, function(i) sampleGammas(i, counts[i])))

```

---

Som punkttestimat på  $\frac{\theta_R}{\theta_L}$  bruker vi medianen i posterioren, som er 0.297. Dette er veldig nære ML-estimatet på 0.292. Kredibilitetsintervallet til 0.15- og 0.95-kvantilene er (0.200, 0.432). I figur (3.4) illustreres posterioren.

## A Kode for oppgave 1



**Figure 3.4:** (venstre) Posteriorentettheten til  $\frac{\theta_R}{\theta_L} | \mathbf{y}$ , glattet ut med `locfit`. Den stiplede linjen er medianen, mens de prikkete er 0.15- og 0.95-kvantilene. Her er  $(a, b) = (1, 1)$ . (høyre) Samme  $\frac{\theta_R}{\theta_L} | \mathbf{y}$  med hyperparametrene  $(a, b) = (2, 10)$ . Ikke overraskende har histogrammet to toppunkter, jamfør figur (3.3). Vi har brukt et histogram da `locfit` smoother vekker det andre toppunktet.

```

1 # Oppgave f) -----
2 # ML, posterior forventning, posterior toppunkt
3 insects = c(0.03, 0.15, 0.18, 0.44, 0.58, 0.63,
4             1.37, 1.91, 2.60, 3.58, 4.10, 7.46)
5 n = length(insects)
6 Mlest = sum(sqrt(insects))^-1*length(insects)
7 postMean = (A+n)/(B+sum(sqrt(insects)))
8 postMode = (A+n-1)/(B+sum(sqrt(insects)))
9
10 # Oppgave g) -----
11 # Late Bayesianere
12
13 ys = insects
14 tt = seq(-0.1, 2.2, by=0.001)
15 ys = insects
16 n = length(ys)
17 thetaML = n/sum(sqrt(ys))
18 top = (A+n-1)/(B+sum(sqrt(ys)))
19 plot(tt, dgamma(tt, A+n, B+n/thetaML), type="l", ylim=c(0, 1.8), bty="l",
20       xlab=expression(theta), ylab="Tetthet")
21 lines(tt, dnorm(tt, top, sqrt(top^2/(A+n-2))), type="l", lty=2, col="red")
22 lines(tt, dnorm(tt, top, top/(sqrt(n))), type="l", lty=3, col="blue")
23 legend("topright", c("Posterior", "Lat", "Ekstra_lat"),
24        col=c("black", "red", "blue"), lty=c(1, 2, 3), bty="n")
25
26 # Oppgave h) -----
27 # Posterior predictive density
28
29 ys = insects
30 a = A
31 b = B
32 n = length(ys)
33
34 alpha = a + n
35 beta = b + sum(sqrt(ys))
36
37 dppd = function(x, alpha, beta) {
38   alpha/(2*sqrt(x))*(beta/(beta+sqrt(x)))^alpha*1/(beta+sqrt(x))
39 }
40
41 pppd = function(x, alpha, beta) {
42   1 - (beta/(beta+sqrt(x)))^alpha
43 }
44
45 qppd = function(x, alpha, beta) {
46   (beta*(1-x)^(-1/alpha)-beta)^2
47 }
48
49 plot(yy, dppd(yy, alpha, beta), type="l", col="blue", bty="l",

```

```

50     xlab=expression(y[13]), ylab="Tetthet", ylim=c(0,2), xlim=c(0,5))
51 lines(yy, dexp(yy, 1/mean(sqrt(insects))), lty=2)
52 lines(yy, dppd(yy, 4+n, 4+sum(sqrt(ys))), col="red", lty=3)
53 legend("topright", c("(5.31, 2.66)", "(4,4)", "Exp"),
54       col=c("blue", "red", "black"), lty=c(1,3,2), bty="n")
55
56 yy = seq(0,20, by=0.01)
57 q01 = qppd(0.1, alpha, beta)
58 q05 = qppd(0.5, alpha, beta)
59 q09 = qppd(0.9, alpha, beta)
60 plot(yy, pppd(yy, alpha, beta), type="l", col="blue", bty="l",
61       xlab=expression(y[13]), ylab="Sannsynlighet", ylim=c(0,1))
62 lines(yy, pppd(yy, 4+n, 4+sum(sqrt(ys))), col="red", lty=3)
63 lines(yy, pexp(yy, 1/mean(sqrt(ys))), lty=2)
64 points(q01, pppd(q01, alpha, beta), pch=4)
65 points(q05, pppd(q05, alpha, beta), pch=4)
66 points(q09, pppd(q09, alpha, beta), pch=4)
67
68 plot(xs, qppd(xs, alpha, beta), col="blue", bty="l",
69       xlab="Sannsynlighet", ylab="Kvantil", type="l", ylim=c(0,20))
70 lines(xs, qppd(xs, 4+n, 4+sum(sqrt(ys))), col="red", lty=3)
71 lines(xs, qexp(xs, 1/mean(sqrt(ys))), lty=2)
72
73 round(cbind(qppd(c(0.1, 0.5, 0.9), alpha, beta),
74             qppd(c(0.1, 0.5, 0.9), n+4, 4+sum(sqrt(ys))),
75             qexp(c(0.1, 0.5, 0.9), 1/mean(sqrt(ys))), 4)

```

## B Kode for oppgave 2

```

1  # Oppgave d) -----
2  # Beregning av risikofunksjoner
3  thetas = seq(0.001, 0.999, by=0.01)
4
5  risky = function(theta, f) {
6    ys = 1:10000
7    sum((theta-f(ys))^2*(1-theta)^(ys-1)*theta)
8  }
9
10 A = 10
11 B = 10
12
13 f_ML = function(x) 1/x
14 f_UP = function(x) 1/(1+x/2)
15 f_JP = function(x) 1/(1/2+x)
16
17 risk_ML = sapply(thetas, function(theta) risky(theta, f_ML))
18 risk_UP = sapply(thetas, function(theta) risky(theta, f_UP))
19 risk_JP = sapply(thetas, function(theta) risky(theta, f_JP))
20 plot(thetas, risk_ML, type="l", ylim=c(0, 0.17), bty="l",
21       ylab="Risiko", xlab=expression(theta))
22 lines(thetas, risk_UP, col="blue")
23 lines(thetas, risk_JP, col="red")
24 legend("topright", c("ML", "Uniform", "Jeffreys"), bty="n",
25       col=c("black", "blue", "red"), lty=c(1,1,1))
26
27 # Oppgave f) -----
28 # Beregning av a og b fra et empirisk bayes perspektiv.
29 ys = c(1, 1, 2, 2, 2, 2, 2, 3, 3, 3,
30       3, 3, 5, 6, 6, 14, 15, 15, 20, 75)
31
32 llik = function(p, ys){
33   a = p[1]
34   b = p[2]
35   sum(lbeta(a+1, b+ys-1) - lbeta(a, b))
36 }
37
38 ab = nlm(function(p) -llik(p, ys), p=c(1,1))$estimate
39 a = ab[1]
40 b = ab[2]
41
42 # Oppgave g) -----
43 # Simulering fra posterioren p(a,b/y).
44
45 N = 1000000

```

```

46 n = length(ys)
47
48 sampleit = function() {
49   a = runif(1,0,100)
50   b = runif(1,0,100)
51   c(a,b,prod(beta(a+1,b+ys-1)/beta(a,b)))
52 }
53
54 samples = replicate(N,sampleit())
55 samples[3,] = samples[3,]/mean(samples[3,])
56
57 (expected_a = mean(samples[1,]*samples[3,]))
58 (expected_b = mean(samples[2,]*samples[3,]))
59
60 sd_a = sqrt(mean(samples[1,]^2*samples[3,]) - expected_a^2)
61 sd_b = sqrt(mean(samples[2,]^2*samples[3,]) - expected_b^2)
62
63 corr = (mean(samples[1,]*samples[2,]*samples[3,]) -
64         expected_a*expected_b)/(sd_a*sd_b)
65
66 # Oppgave h) -----
67 # Tabell over verdier.
68
69 parti = function(i,x) {
70   mean(sapply(1:N,function(j) dbeta(x,samples[1,j]+1,
71                                     samples[2,j]+ys[i]-1)*samples[3,j])))
72 }
73
74 thetas = seq(0,1,by=0.001)
75
76 cols = function(i) {
77   means = sapply(thetas,function(x) parti(i,x))
78   quants = cumsum(means)/length(means)
79   q005 = thetas[which.min(abs(quants-0.05))]
80   q050 = thetas[which.min(abs(quants-0.50))]
81   q095 = thetas[which.min(abs(quants-0.95))]
82   c(q005,q050,q095)
83 }
84
85 results = t(sapply(1:20,cols))
86 col1 = ys
87 col2 = 1/ys
88 col3 = (a+1)/(a+b+ys)
89 res = cbind(col1,col2,col3,results)
90 round(res,3)

```

## C Kode for oppgave 3

```

1 # Oppgave c) (Posterior) -----
2
3 mu1 = 1.1
4 mu2 = 2.2
5 tau0 = 66
6 n = 100
7 yy = c(rnorm(tau0,mu1,1),rnorm(n-tau0,mu2,1))
8
9 dens = function(yy,tau) {
10   prod(dnorm(yy[1:tau],mu1,1))*prod(dnorm(yy[(tau+1):n],mu2,1))
11 }
12
13 C = sum(sapply(1:(n-1),function(tau) dens(yy,tau)))
14
15 plot(45:90,sapply(45:90,function(tau) dens(yy,tau)/C),bty="l",
16      xlab=expression(tau),ylab="Posterior",pch=20)
17 abline(v=66,lty=2,col="grey")
18
19 # Oppgave c) (Frekventistiske egenskaper) -----
20 N = 100000
21
22 maxes = c()
23 means = c()
24
25 for (i in 1:N) {
26   yy = c(rnorm(tau0,mu1,1),rnorm(n-tau0,mu2,1))
27   vals = sapply(1:(n-1),function(tau) dens(yy,tau))

```

```

28   C      = sum( vals )
29   means = c( means, sum((1:(n-1))*vals/C) )
30   maxes = c( maxes, which.max( vals ) )
31 }
32
33 is = min( maxes[ maxes > 0 ] ) : max( maxes[ maxes > 0 ] )
34 ts = seq( 1, 100, by=0.01 )
35
36 plot( is, sapply( is, function( i ) mean( maxes==i ) ), type="h", bty="l",
37       col="blue", xlab=expression( widehat( tau ) [mode] ), ylab="Posterior",
38       xlim=c( 50, 82 ), ylim=c( 0, mean( maxes==66 ) ) )
39 lines( density( means, adjust=2 ), type="l", bty="l", col="red", main=NA,
40        xlab=expression( widehat( tau ) ), ylab="Tetthet_\u00d7 Sannsynlighet",
41        xlim=c( 50, 82 ), ylim=c( 0, mean( maxes==66 ) ) )
42
43 b = mean( abs( means-66 ) )
44 lines( ts, 1/(2*b)*exp(-abs( ts-66)/b ), lty=2, col="black" )
45 legend( "topright", c( expression( widehat( tau ) [mean] ),
46                       expression( widehat( tau ) [mode] ), "Laplace" ), bty="n",
47        col=c( "red", "blue", "black" ), lty=c( 1, 1, 2 ) )
48
49 # Oppgave d) (Maximum likelihood) -----
50
51 data = c( 4, 5, 3, 1, 4, 4, 1, 5, 5, 3, 4, 2,
52          5, 2, 2, 3, 4, 2, 1, 3, 2, 2, 1, 1,
53          1, 1, 3, 0, 0, 1, 0, 1, 1, 0, 0, 3,
54          1, 0, 3, 2, 2, 0, 1, 1, 1, 0, 1, 0,
55          1, 0, 0 )
56
57 n = length( data )
58
59 taus = 1:n
60
61 ml = function( tau ) {
62   thetaL = mean( data[ 1:tau ] )
63   thetaR = mean( data[ (tau+1):n ] )
64   loglik = sum( dpois( data[ 1:tau ], thetaL ) ) +
65           sum( dpois( data[ (tau+1):(n+1) ], thetaR ), na.rm=TRUE )
66   loglik
67 }
68
69 thetas = function( tau ) {
70   thetaL = mean( data[ 1:tau ] )
71   thetaR = mean( data[ (tau+1):(n+1) ], na.rm=TRUE )
72   c( thetaL, thetaR )
73 }
74
75 mls = sapply( taus, ml )
76 thetaLs = sapply( taus, thetas ) [ 1, ]
77 thetaRs = sapply( taus, thetas ) [ 2, ]
78
79 plot( taus, mls, bty="l", ylab="Log_likelihood", xlab=expression( tau ), pch=20 )
80 abline( v=22, lty=2, col="grey" )
81 plot( taus, thetaLs, col="blue", ylim=c( 0, 4.5 ), bty="l",
82       ylab=expression( theta ), xlab=expression( tau ), pch=20 )
83 points( taus, thetaRs, col="red", pch=20 )
84 abline( h=min( thetaLs ), lty=2, col="blue" )
85 abline( h=max( thetaRs, na.rm=TRUE ), lty=2, col="red" )
86 legend( "topright", c( expression( theta [L] ), expression( theta [R] ) ), bty="n",
87        col=c( "blue", "red" ), pch=c( 20, 20 ) )
88
89 # Opppgave e -----
90
91 a = 1
92 b = 1
93
94 posterior = function( tau ) {
95   sum1 = sum( data[ 1:tau ] )
96   sum2 = sum( data[ (tau+1):n+1 ], na.rm=TRUE )
97   b^(2*a) / ( n*prod( factorial( data ) ) * gamma(a)^2 ) *
98   gamma( sum1+a ) / ( tau+b )^( sum1+a ) *
99   gamma( sum2+a ) / ( n-tau+b )^( sum2+a )
100 }
101
102
103 prop = sum( sapply( taus, posterior ) )
104

```

```

105 posts = sapply(taus,posterior)/prop
106
107 plot(taus,sapply(taus,posterior)/prop,xlab=expression(tau),pch=20,
108      ylab="Posterior",bty="l")
109
110 # Oppgave f -----
111
112 library("locfit")
113 N = 10000000
114 a = 1
115 b = 1
116
117 sampleGammas = function(tau,count) {
118   sum1 = sum(data[1:tau])
119   sum2 = sum(data[(tau+1):n+1],na.rm=TRUE)
120   thetaR = rgamma(count,sum2+a,n-tau+b)
121   thetaL = rgamma(count,sum1+a,tau+b)
122   thetaR/thetaL
123 }
124
125 counts = floor(posts*N)
126 res = unlist(sapply(2:50,function(i) sampleGammas(i,counts[i])))
127 quants = round(quantile(res,c(0.05,0.95)),3)
128
129 plot(1,type="n",bty="l",xlab=expression(theta[R]/theta[L]),ylab="Posterior",
130      xlim=c(min(res),max(res)),ylim=c(0,6),
131      sub=paste0("CI: ",quants[1],"-",quants[2],"))
132 lines(locfit(~res),main=NA,sub=NA,bty="l")
133 abline(v=median(res),lty=2,col="grey")
134 abline(v=quantile(res,0.05),col="purple",lty=3)
135 abline(v=quantile(res,0.95),col="purple",lty=3)

```