BAYESEKSAMEN

December 10, 2015

1 Insekter

a)

I denne oppgaven antas det følgende funksjonaler av eksponensialfordelinga med parameter θ er kjent: Forventinga (θ^{-1}), variansen (θ^{-2}) og medianen ($\theta^{-1}\log 2$). La $X \sim \exp(\theta)$ og sett $Y = X^2$. Da er $P(Y \le y) = P(X \le \sqrt{y}) = 1 - \exp(-\theta\sqrt{y})$. 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 $P(X^2 \le \alpha) = 0.5$ hvis og bare hvis $P(X \le \sqrt{\alpha}) = 0.5$, så $\text{median}(Y) = (\theta^{-1}\log 2)^2$.

b)

Log likelihooden er

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

Når vi deriverer med hensyn på θ og setter resultatet lik 0 ender vi opp med

$$\widehat{\theta} = \frac{n}{\sum_{i=1}^{n}\sqrt{y_i}}. \tag{1.1}$$

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

c)

Etter Bayes' regel er

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

hvor proporsjonaliteten er i θ. 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 $p(\theta \mid y)$ har proporsjonalitetskonstant $\frac{(b+\sqrt{y})^{a+1}}{\Gamma(a+1)}$, og proporsjonalitetskonstanten (m.h.p θ) i $p(y \mid \theta)p(\theta)$ 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 likhetheten 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 parameterestimater.

```
> 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 ≈ 2 og .9-kvantil i ≈ 4.5.

f)

ML-estimatet er $\widehat{\theta} = 0.856$, det følger av (1.1). Siden forventinga til en Gamma(a,b)-fordelt variabel er $\frac{a}{b}$, følger det av resultatet $p(\theta \mid y) = \Gamma(a+n, b+\sum_{i=1}^n \sqrt{y_i})$ 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 $\widehat{\theta} = 0.856$. Posteriortopppunktet er $\frac{a-1}{b} = 0.978$, som er litt nærmere ML-estimatet.

g)

De late Bayesianerne bruker approksimasjonen $p(\theta \mid y) = N(\theta^\pi, (I^\pi)^{-1})$, hvor θ^π er posteriortoppunktet og I^π er den generaliserte Fisherinformasjonen. Nå kan vi enten være ekstra late, og anta at π 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 $\frac{d^2}{d\theta^2} \log f(\theta) = \theta^{-2}$, $\frac{d^2}{d\theta^2} \log[p(\theta \mid \mathbf{y})] = -\theta^{-2}(a+n-2)$ og $\widehat{\theta} = 0.978$ 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.

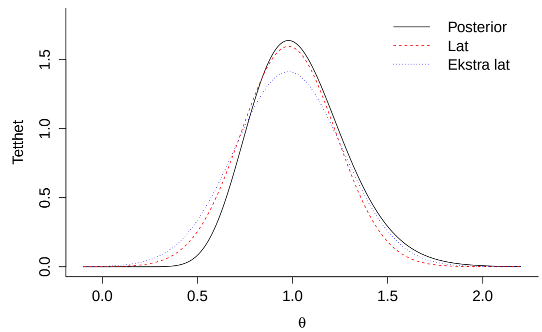

Figure 1.1: Plot av posterioren til θ sammen med late og esktra late varianter.

h)

La oss finne den posteriore prediktive tettheten i y.

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

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

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

hvor $\gamma'(\theta)$ er tettheten til 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

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

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 θ er $\widehat{\theta}_{ML} = n(\sum_{i=1}^n \sqrt{y})^{-1}$ og at $e^{-x} = \lim_{n\to\infty}(1 + \frac{x}{n})^{-n}$. Vi skriver om $\frac{\beta^\alpha}{(\beta+\sqrt{y})^\alpha}$ til $\left(1+\frac{\sqrt{y}}{\beta}\right)^{-\alpha}$ og bytter tilbake α og β for å få

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

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

	$(5.31, 2.66)$	$(4,4)$	$\exp(\widehat{\theta}_{ML})$
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 $(1+\frac{x}{n})^{-n} \to \exp(-x)$ 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

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

som er det k-te momentet til $E_{\theta|\mathbf{y}}(p(y \mid \theta))$. 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

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

hvor $h(y) = E_{\theta|\mathbf{y}}(p(y \mid \theta))$, 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, fordelingsfunksjonen 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 $\exp(\widehat{\theta}_{ML})$. Som allerede indikert har tettheten $\exp(\widehat{\theta}_{ML})$ seriøst tynnere haler enn $E_{\theta|\mathbf{y}}(p(y \mid \theta))$, noe som kan forklares ved at den ikke tar hensyn til usikkerheten i estimatet $\widehat{\theta}_{ML}$.

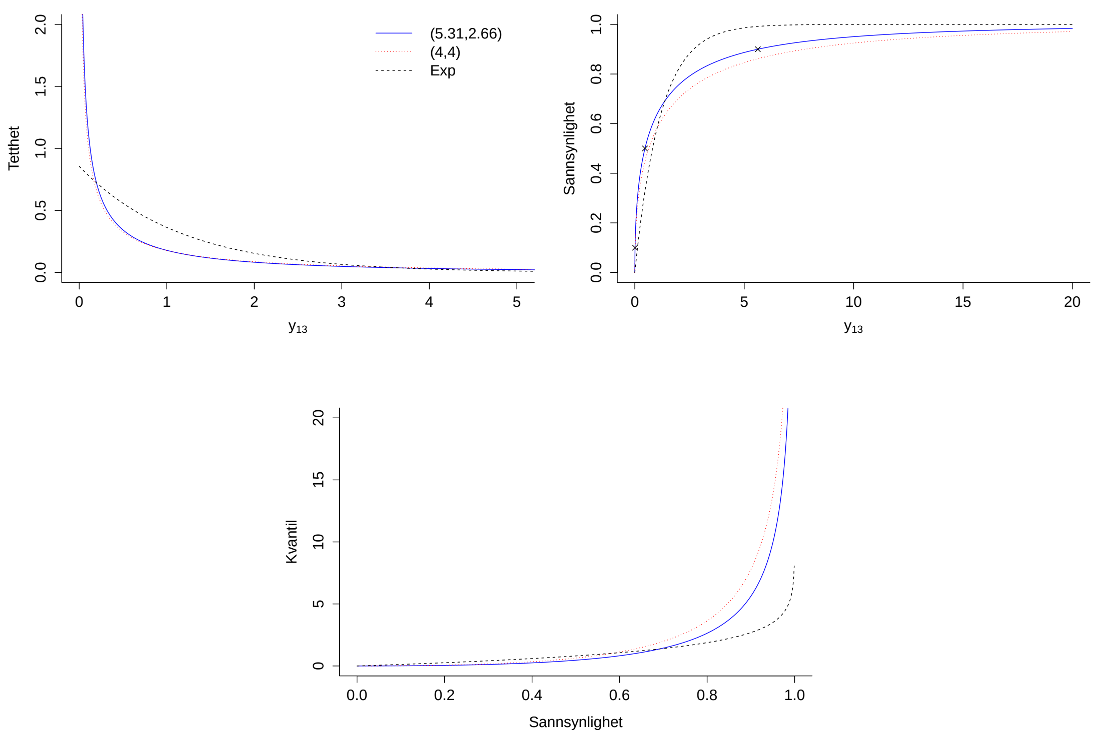

Figure 1.2: (venstre) Posterior prediktiv tetthet til y_{13}, den røde linja kan knapt ses. (høyre) Kumulativ tetthet. Kvantilene 0.1, 0.5 og 0.9 som hører til $(a, b) = (5.31, 2.66)$ er markert. (bunn) Kvantilfunksjonen.

2 Hierarkisk modell

a)

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

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

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

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

Siden forventinga til $\text{Beta}(\alpha, \beta)$ er $\frac{\alpha}{\alpha+\beta}$, er posterior forventing 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

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

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

b)

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

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

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

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

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.

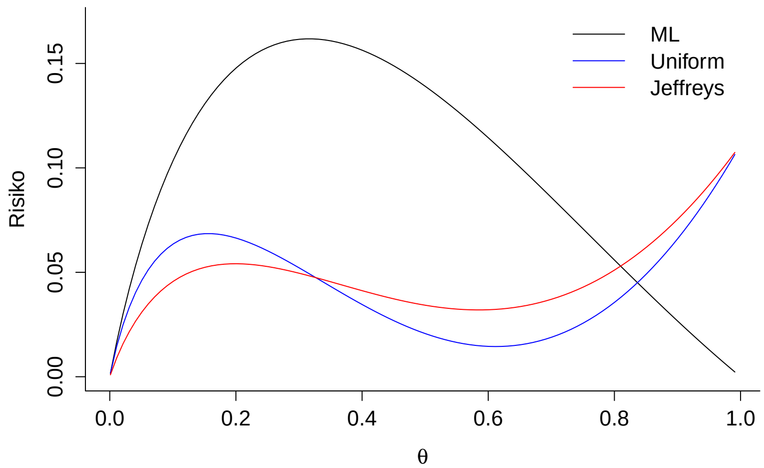

Figure 2.1: Sammenligning av kvadratrisikoen for de tre forskjellige estimatorne. ML er bedre enn uniform Bayes når $\theta > 0.84$.

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 Beta$(0,\frac{1}{2})$. Dette er en *improper* prior, ettersom θ^{-1} vokser for fort når $\theta \to 0$. Prioren legger altså ekstremt stor vekt på små θ-verdier og rimelig stor vekt på store θ-verdier.

Posterioren er proporsjonal med $(1-\theta)^{y-\frac{1}{2}-1}$, som svarer til Beta$(1, y-\frac{1}{2})$-fordelinga. Denne har forventing 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 $p(\theta \mid y) \sim$ Beta$(2, y)$ når $(a,b) = (1,1)$. Siden forventinga til Betafordelinga er gitt ved $\frac{a}{a+b}$, er posterior forventing 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 θ er nærme 1: For ML-estimatet er større enn posterior forventing for både uniform of Jeffreys.

Når $a+b = k \in \mathbb{N}$ kan en uttrykke risikoen ved hjelp av polylogaritmen 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 $\sum_{i=1}^{\infty}(\frac{a}{a+b+i}-x)^2x^i$ 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.

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	

e)

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

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

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

```r
ys = c(1, 1, 2, 2, 2,  2,  2,  3,  3,  3,
       3, 3, 5, 6, 6, 14, 15, 15, 20, 75)

llik = function(p,ys){
  a = p[1]
  b = p[2]
  sum(lbeta(a+n,b+ys-n) - lbeta(a,b))
}

ab = nlm(function(p) -llik(p,ys),p=c(1,1))$estimate
```

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

g)

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

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

Det er en smal sak å sample fra denne fordelinga. Vi trekker N observasjoner fra $\mathcal{U}([0, 100] \times [0, 100])$, beregner $\prod_{i=1}^{n} \frac{\mathrm{B}(a+1,b+y_i-1)}{\mathrm{B}(a,b)}$ 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 3: Tabell over verdier.

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

h)

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

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

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 θ_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

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

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

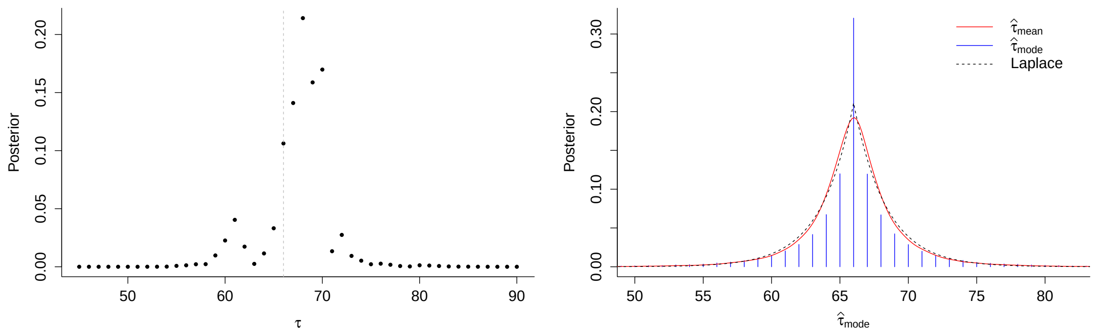

Figure 3.1: (venstre) Posteriorfordeling for $\tau \mid \mathbf{y}$. Den treffer ganske bra. (høyre) Samplingfordelingene til posteriortopppunket og posterior forventing, sammen med Laplacefordelinga med parametere estimert gjennom ML. Tettheten er plottet ved hjelp av `density`-funksjonen i `R`. Dn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

c)

Vi kjører en simulering med $n = 100$ og får posterioren i figur 3.1. Topppunktet 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 forventing og stolpediagrammet til posteriortopppunktet. Se figur 3.1 igjen. Interessant nok ser posterior forventing ut som den er Laplacefordelt, mens topppunktet 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 τ, 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 illusteres det hvordan log likelihooden ser ut, i tillegg til hvordan ML-estimatene $\widehat{\theta_L}$ og $\widehat{\theta_R}$ endres ettersom bruddpunktet τ forandres.

e)

Via kjerneregelen for kondisjonering får vi

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

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

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

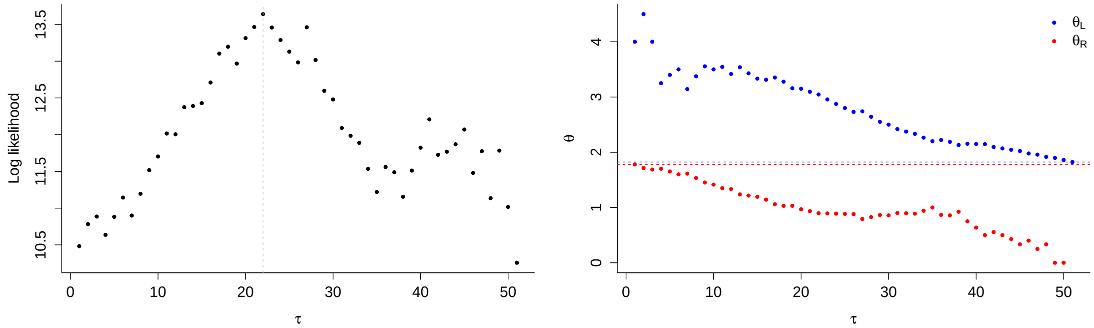

Figure 3.2: (venstre) Plot over max-log likelihooden for ulike valg av τ. (høyre) Her kan vi se hvordan θ_L og θ_R forandrer seg ettersom τ forandres. Legg merke til at $\theta_L > \theta_R$ for alle valg av τ, noe som støtter hypotesen om at det finnes ett bruddpunkt.

Vi kan identifisere de ulike betingede fordelingene ved å finne proporsjonalitetskonstantene. For å finne $p(\tau \mid \mathbf{y})$ integrerer vi ut θ_L, θ_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:

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

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 `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. Topppunktet $\tau \approx 22$ svarer heller til en mindre dramatisk endring fra θ_L til θ_R.

f)

Om $(\theta_L^1, \theta_R^1), (\theta_L^2, \theta_R^2), ..., (\theta_L^N, \theta_R^N)$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

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

Kall $p(\tau = i \mid \mathbf{y})$ for t_i. Da er $p(\theta_L, \theta_R \mid \mathbf{y}) = \sum_{i=1}^{n} t_i p(\theta_L, \theta_R \mid \mathbf{y}, \tau)$. Vi kan sample fra denne ved først å "sample" fra $\tau \mid \mathbf{y}$ 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 $\tau \mid \mathbf{y}$ har domene $1, 2, ..., n$, er det ingen vits i å sample fra den, vi trenger bare å ta antall $\tau = i$ omtrent lik Nt_i, hvor N er antall samples.

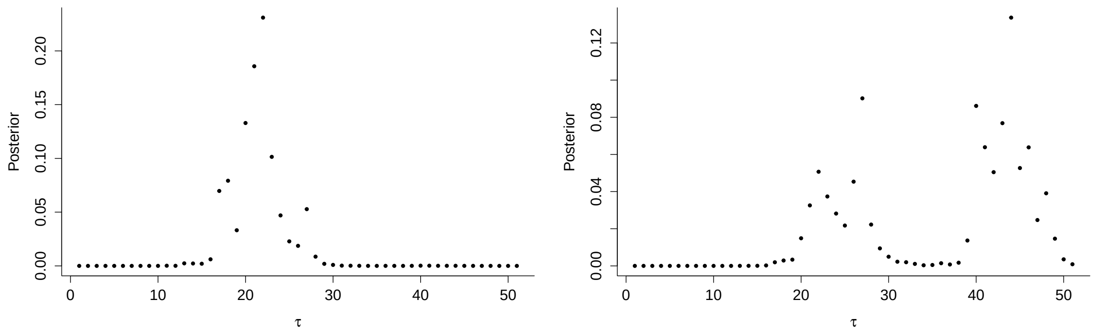

Figure 3.3: (venstre) Posterioren $p(\tau \mid \mathbf{y})$, hvor θ_L, θ_R er uavhengig exp(1)-fordelte og τ er diskret uniform på $1, ..., 51$. (høyre) Samme posterior, hvor θ_L, θ_R er uavhengig Gamma$(2, 10)$-fordelte. Eksponensialprioren legger all vekta rundt bumpen i likelihooden omkring $\tau = 20$, mens Gamma$(2, 10)$ finner for godt å priorietere det lokale maksimumet omkring $\tau = 45$.

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

```
# tis er vektoren med posterior sannsynligheter for tau.

a = 1
b = 1
N = 1000000
counts = floor(tis*N)

sampleGammas = function(tau,count){
        sum1 = sum(data[1:tau])
        sum2 = sum(data[(tau+1):n+1],na.rm=TRUE)
        # Sikrer at tau = 51 også funker.
        thetaR = rgamma(count,sum2+a,n-tau+b)
        thetaL = rgamma(count,sum1+a,tau+b)
        thetaR/thetaL
}

result = unlist(sapply(1:51,function(i) sampleGammas(i,counts[i])))
```

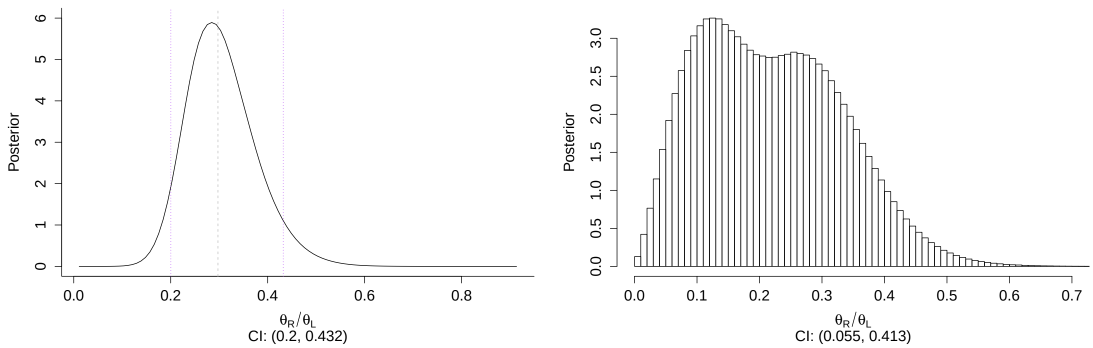

Figure 3.4: (venstre) Posteriorentettheten til $\frac{\theta_R}{\theta_L} \mid \mathbf{y}$, glattet ut med `locfit`. Den stiplede linjen er medianen, mens de prikkete er 0.15- og 0.95-kvantilene. Her er $(a, b) = (1, 1)$. (høyre) Samme $\frac{\theta_R}{\theta_L} \mid \mathbf{y}$ med hyperparameterne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

Som punktestimat 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

```
# Oppgave f) ---------------------------------------------------------------
# ML, posterior forventing, posterior topppunkt
insects  = c(0.03, 0.15, 0.18, 0.44, 0.58, 0.63,
             1.37, 1.91, 2.60, 3.58, 4.10, 7.46)
n        = length(insects)
MLest    = sum(sqrt(insects))^-1*length(insects)
postMean = (A+n)/(B+sum(sqrt(insects)))
postMode = (A+n-1)/(B+sum(sqrt(insects)))

# Oppgave g) ---------------------------------------------------------------
# Late Bayesianere

ys = insects
tt = seq(-0.1,2.2,by=0.001)
ys = insects
n  = length(ys)
thetaML = n/sum(sqrt(ys))
top = (A+n-1)/(B+sum(sqrt(ys)))
plot(tt,dgamma(tt,A+n,B+n/thetaML),type="l",ylim=c(0,1.8),bty="l",
     xlab=expression(theta),ylab="Tetthet")
lines(tt,dnorm(tt,top,sqrt(top^2/(A+n-2))),type="l",lty=2,col="red" )
lines(tt,dnorm(tt,top,top/(sqrt(n)))   ,type="l",lty=3,col="blue")
legend("topright",c("Posterior","Lat","Ekstra lat"),
       col=c("black","red","blue"),lty=c(1,2,3),bty="n")

# Oppgave h) ---------------------------------------------------------------
# Posterior predictive density

ys = insects
a = A
b = B
n = length(ys)

alpha = a + n
beta = b + sum(sqrt(ys))

dppd = function(x,alpha,beta) {
  alpha/(2*sqrt(x))*(beta/(beta+sqrt(x)))^alpha*1/(beta+sqrt(x))
}
```

```r

pppd = function(x,alpha,beta) {
  1 - (beta/(beta+sqrt(x)))^alpha
}

qppd = function(x,alpha,beta) {
  (beta*(1-x)^(-1/alpha)-beta)^2
}

plot(yy,dppd(yy,alpha,beta),type="l",col="blue",bty="l",
     xlab=expression(y[13]),ylab="Tetthet",ylim=c(0,2),xlim=c(0,5))
lines(yy,dexp(yy,1/mean(sqrt(insects))),lty=2)
lines(yy,dppd(yy,4+n,4+sum(sqrt(ys))),col="red",lty=3)
legend("topright",c("(5.31,2.66)","(4,4)","Exp"),
       col=c("blue","red","black"),lty=c(1,3,2),bty="n")

yy = seq(0,20,by=0.01)
q01 = qppd(0.1,alpha,beta)
q05 = qppd(0.5,alpha,beta)
q09 = qppd(0.9,alpha,beta)
plot(yy,pppd(yy,alpha,beta),type="l",col="blue",bty="l",
     xlab=expression(y[13]),ylab="Sannsynlighet",ylim=c(0,1))
lines(yy,pppd(yy,4+n,4+sum(sqrt(ys))),col="red",lty=3)
lines(yy,pexp(yy,1/mean(sqrt(ys))),lty=2)
points(q01,pppd(q01,alpha,beta),pch=4)
points(q05,pppd(q05,alpha,beta),pch=4)
points(q09,pppd(q09,alpha,beta),pch=4)

plot(xs,qppd(xs,alpha,beta),col="blue",bty="l",
     xlab="Sannsynlighet",ylab="Kvantil",type="l",ylim=c(0,20))
lines(xs,qppd(xs,4+n,4+sum(sqrt(ys))),col="red",lty=3)
lines(xs,qexp(xs,1/mean(sqrt(ys))),lty=2)

round(cbind(qppd(c(0.1,0.5,0.9),alpha,beta),
qppd(c(0.1,0.5,0.9),n+4,4+sum(sqrt(ys))),
qexp(c(0.1,0.5,0.9),1/mean(sqrt(ys)))),4)
```

B Kode for oppgave 2

```r
# Oppgave d) -------------------------------------------------------------------
# Beregning av risikofunksjoner
thetas = seq(0.001,0.999,by=0.01)

risky = function(theta,f) {
  ys = 1:10000
  sum((theta-f(ys))^2*(1-theta)^(ys-1)*theta)
}

A = 10
B = 10

f_ML = function(x) 1/x
f_UP = function(x) 1/(1+x/2)
f_JP = function(x) 1/(1/2+x)

risk_ML = sapply(thetas,function(theta) risky(theta,f_ML))
risk_UP = sapply(thetas,function(theta) risky(theta,f_UP))
risk_JP = sapply(thetas,function(theta) risky(theta,f_JP))
plot(thetas,risk_ML,type="l",ylim=c(0,0.17),bty="l",
     ylab="Risiko",xlab=expression(theta))
lines(thetas,risk_UP,col="blue")
lines(thetas,risk_JP,col="red")
legend("topright",c("ML","Uniform","Jeffreys"),bty="n",
       col=c("black","blue","red"),lty=c(1,1,1))

# Oppgave f) -------------------------------------------------------------------
# Beregning av a og b fra et empirisk bayes perspektiv.
ys = c(1, 1, 2, 2, 2, 2, 2, 3, 3, 3,
       3, 3, 5, 6, 6, 14, 15, 15, 20, 75)

llik = function(p,ys){
  a = p[1]
  b = p[2]
  sum(lbeta(a+1,b+ys-1) - lbeta(a,b))
```

```
}

ab = nlm(function(p) -llik(p,ys),p=c(1,1))$estimate
a  = ab[1]
b  = ab[2]

# Oppgave g) -------------------------------------------------------------------------------
# Simulering fra posterioren p(a,b|y).

N = 1000000
n = length(ys)

sampleit = function() {
  a = runif(1,0,100)
  b = runif(1,0,100)
  c(a,b,prod(beta(a+1,b+ys-1)/beta(a,b)))
}

samples = replicate(N,sampleit())
samples[3,] = samples[3,]/mean(samples[3,])

(expected_a = mean(samples[1,]*samples[3,]))
(expected_b = mean(samples[2,]*samples[3,]))

sd_a =  sqrt(mean(samples[1,]^2*samples[3,]) - expected_a^2)
sd_b =  sqrt(mean(samples[2,]^2*samples[3,]) - expected_b^2)

corr = (mean(samples[1,]*samples[2,]*samples[3,]) -
          expected_a*expected_b)/(sd_a*sd_b)

# Oppgave h) -------------------------------------------------------------------------------
# Tabell over verdier.

parti = function(i,x) {
  mean(sapply(1:N,function(j) dbeta(x,samples[1,j]+1,
                     samples[2,j]+ys[i]-1)*samples[3,j]))
}

thetas = seq(0,1,by=0.001)

cols = function(i) {
  means  = sapply(thetas,function(x) parti(i,x))
  quants = cumsum(means)/length(means)
  q005   = thetas[which.min(abs(quants-0.05))]
  q050   = thetas[which.min(abs(quants-0.50))]
  q095   = thetas[which.min(abs(quants-0.95))]
  c(q005,q050,q095)
}

results = t(sapply(1:20,cols))
col1    = ys
col2    = 1/ys
col3    = (a+1)/(a+b+ys)
res     = cbind(col1,col2,col3,results)
round(res,3)
```

C Kode for oppgave 3

```
# Oppgave c) (Posterior) --------------------------------------------------------------

mu1 = 1.1
mu2 = 2.2
tau0 = 66
n = 100
yy = c(rnorm(tau0,mu1,1),rnorm(n-tau0,mu2,1))

dens = function(yy,tau) {
  prod(dnorm(yy[1:tau],mu1,1))*prod(dnorm(yy[(tau+1):n],mu2,1))
}

C = sum(sapply(1:(n-1),function(tau) dens(yy,tau)))

plot(45:90,sapply(45:90,function(tau) dens(yy,tau)/C),bty="l",
     xlab=expression(tau),ylab="Posterior",pch=20)
abline(v=66,lty=2,col="grey")
```

```
# Oppgave c) (Frekventistiske egenskaper) ------------------------------------
N = 100000

maxes = c()
means = c()

for (i in 1:N) {
  yy    = c(rnorm(tau0,mu1,1),rnorm(n-tau0,mu2,1))
  vals  = sapply(1:(n-1),function(tau) dens(yy,tau))
  C     = sum(vals)
  means = c(means,sum((1:(n-1))*vals/C))
  maxes = c(maxes,which.max(vals))
}

is = min(maxes[maxes>0]):max(maxes[maxes>0])
ts = seq(1,100,by=0.01)

plot(is,sapply(is,function(i) mean(maxes==i)),type="h",bty="l",
     col="blue",xlab=expression(widehat(tau)[mode]),ylab="Posterior",
     xlim=c(50,82),ylim=c(0,mean(maxes==66)))
lines(density(means,adjust=2),type="l",bty="l",col="red",main=NA,
      xlab=expression(widehat(tau)),ylab="Tetthet / Sannsynlighet",
      xlim=c(50,82),ylim=c(0,mean(maxes==66)))

b = mean(abs(means-66))
lines(ts,1/(2*b)*exp(-abs(ts-66)/b),lty=2,col="black")
legend("topright",c(expression(widehat(tau)[mean]),
                    expression(widehat(tau)[mode]),"Laplace"),bty="n",
       col=c("red","blue","black"),lty=c(1,1,2))

# Oppgave d) (Maximum likelihood) ---------------------------------------------

data = c(4, 5, 3, 1, 4, 4, 1, 5, 5, 3, 4, 2,
         5, 2, 2, 3, 4, 2, 1, 3, 2, 2, 1, 1,
         1, 1, 3, 0, 0, 1, 0, 1, 1, 0, 0, 3,
         1, 0, 3, 2, 2, 0, 1, 1, 1, 0, 1, 0,
         1, 0, 0)

n = length(data)

taus = 1:n

ml = function(tau) {
  thetaL = mean(data[1:tau])
  thetaR = mean(data[(tau+1):n])
  loglik = sum(dpois(data[1:tau],thetaL)) +
           sum(dpois(data[(tau+1):(n+1)],thetaR),na.rm=TRUE)
  loglik
}

thetas = function(tau) {
  thetaL = mean(data[1:tau])
  thetaR = mean(data[(tau+1):(n+1)],na.rm=TRUE)
  c(thetaL,thetaR)
}

mls = sapply(taus,ml)
thetaLs = sapply(taus,thetas)[1,]
thetaRs = sapply(taus,thetas)[2,]

plot(taus,mls,bty="l",ylab="Log likelihood",xlab=expression(tau),pch=20)
abline(v=22,lty=2,col="grey")
plot(taus,thetaLs,col="blue",ylim=c(0,4.5),bty="l",
     ylab=expression(theta),xlab=expression(tau),pch=20)
points(taus,thetaRs,col="red",pch=20)
abline(h=min(thetaLs),lty=2,col="blue")
abline(h=max(thetaRs,na.rm=TRUE),lty=2,col="red")
legend("topright",c(expression(theta[L]),expression(theta[R])),bty="n",
       col=c("blue","red"),pch=c(20,20))

# Opppgave e -------------------------------------------------------------------

a = 1
b = 1

posterior = function(tau) {
```

```r
  sum1 = sum(data[1:tau])
  sum2 = sum(data[(tau+1):n+1],na.rm=TRUE)
  b^(2*a)/(n*prod(factorial(data))*gamma(a)^2)*
    gamma(sum1+a)/(tau+b)^(sum1+a)*
    gamma(sum2+a)/(n-tau+b)^(sum2+a)
}


prop = sum(sapply(taus,posterior))

posts = sapply(taus,posterior)/prop

plot(taus,sapply(taus,posterior)/prop,xlab=expression(tau),pch=20,
     ylab="Posterior",bty="l")

# Oppgave f ------------------------------------------------------------------------

library("locfit")
N = 10000000
a = 1
b = 1

sampleGammas = function(tau,count) {
  sum1 = sum(data[1:tau])
  sum2 = sum(data[(tau+1):n+1],na.rm=TRUE)
  thetaR = rgamma(count,sum2+a,n-tau+b)
  thetaL = rgamma(count,sum1+a,tau+b)
  thetaR/thetaL
}

counts = floor(posts*N)
res = unlist(sapply(2:50,function(i)  sampleGammas(i,counts[i])))
quants = round(quantile(res,c(0.05,0.95)),3)

plot(1,type="n",bty="l",xlab=expression(theta[R]/theta[L]),ylab="Posterior",
     xlim=c(min(res),max(res)),ylim=c(0,6),
     sub=paste0("CI: (",quants[1],", ",quants[2],")"))
lines(locfit(~res),main=NA,sub=NA,bty="l")
abline(v=median(res),lty=2,col="grey")
abline(v=quantile(res,0.05),col="purple",lty=3)
abline(v=quantile(res,0.95),col="purple",lty=3)
```