Exercise 3.1, Binomial and multinomial models:

$y = (y_1, \dots, y_J)$ följer en multinomial fördelning med parametrar $\theta = (\theta_1, \dots, \theta_J)$, vilket innebär att

$$p(y|\theta) = \binom{n}{y_1 \cdot y_2 \cdots y_J} \cdot \theta_1^{y_1} \cdot \theta_2^{y_2} \cdot \dots \cdot \theta_J^{y_J} \overset{\theta}{\propto} \prod_{i=1}^{J} \theta_i^{y_i}$$

Antag att $\theta = (\theta_1, \dots, \theta_J)$ följer en Dirichlet fördelning, så att priorfördelningen av θ med parametrar $a = (a_1, \dots, a_J)$ är

$$p(\theta) = \frac{\Gamma(a_1 + a_2 + \dots + a_J)}{\Gamma(a_1) \cdot \Gamma(a_2) \cdot \dots \cdot \Gamma(a_J)} \cdot \theta_1^{a_1 - 1} \cdot \theta_2^{a_2 - 1} \cdot \dots \cdot \theta_J^{a_J - 1}$$

$$\propto \prod_{i=1}^{J} \theta_i^{a_i - 1}.$$

Detta ger posteriorfunktionen för θ som en Dirichlet fördelning med parametrar $(a_1 + y_1, a_2 + y_2, \dots, a_J + y_J)$, ty

$$p(\theta|y) \propto p(\theta) \cdot p(y|\theta) \propto \prod_{i=1}^{J} \theta_i^{a_i - 1} \cdot \prod_{i=1}^{J} \theta_i^{y_i}$$

$$= \prod_{i=1}^{J} \theta_i^{(a_i + y_i) - 1}$$

a) Bestäm marginella posteriorfunktionen för α, där

$$\alpha = \frac{\theta_1}{\theta_1 + \theta_2}.$$

Exercise 3.1 forts.

②

α är en funktion av parametrarna θ_1 och θ_2.
Därför kan vi först bestämma marginella posteriorfunktionen för $(\theta_1, \theta_2) \mid y$ och sedan med variabelbyte bestämma marginella posteriorfunktionen för $(\alpha, f(\theta_1, \theta_2))$, där funktionen f av parametrarna θ_1 och θ_2 väljs på ett lämpligt sätt.
Generellt, om $\theta = (\theta_1, \theta_2, \dots, \theta_J)$ följer en Dirichlet förd. med par. $(\alpha_1, \alpha_2, \dots, \alpha_J)$, så har vi ett resultat från appendix A, s. 582, att

$$(\theta_i, \theta_j, 1-\theta_i-\theta_j) \sim \text{Dirichlet}\left(\alpha_i, \alpha_j, \sum_{k=1}^{J}\alpha_k - \alpha_i - \alpha_j\right),$$

d.v.s.

$$p[(\theta_1, \theta_2) \mid y] = \text{Dirichlet}\left(a_1+y_1, a_2+y_2, (a_3+y_3)+\dots+(a_J+y_J)\right)$$

$$\propto \theta_1^{a_1+y_1-1} \cdot \theta_2^{a_2+y_2-1} \cdot (1-\theta_1-\theta_2)^{(a_3+y_3)+\dots+(a_J+y_J)-1}$$

Från denna marginella posteriorfunktion genomför vi nu ett variabelbyte från (θ_1, θ_2) till $\left(\alpha = \frac{\theta_1}{\theta_1+\theta_2}, f(\theta_1, \theta_2)\right)$
och bestämmer den marginella täthetsfunktionen för $(\alpha, f(\theta_1, \theta_2))$.
Då har vi att

$$p_{\alpha, f}[(\alpha, f(\theta_1, \theta_2)) \mid y] = p_{\theta_1, \theta_2}\left[\alpha(\theta_1, \theta_2), f(\theta_1, \theta_2)\right] \cdot \left| J[\alpha, f(\theta_1, \theta_2)] \right|$$

Exercise 3.1 forts.

där $\left| J(\alpha, f(\theta_1,\theta_2)) \right|$ är areaförstoringen under variabelbytet och definieras som

$$J(\alpha, f) = \frac{d(\theta_1,\theta_2)}{d(\alpha, f)} = \frac{1}{\frac{d(\alpha,f)}{d(\theta_1,\theta_2)}} =$$

$$= \frac{1}{\begin{vmatrix} \frac{d\alpha}{d\theta_1} & \frac{d\alpha}{d\theta_2} \\ \frac{df}{d\theta_1} & \frac{df}{d\theta_2} \end{vmatrix}}$$

Vi har att $\alpha = \frac{\theta_1}{\theta_1+\theta_2}$, men hur ska vi välja $f(\theta_1,\theta_2)$ på ett lämpligt sätt? Om vi väljer f så att α och f är oberoende av varandra ges den marginella täthetsfunktionen för α direkt, ty då gäller det att den simultana täthetsfunktionen för α, f är

$$p[(\alpha,f) \mid y] = p(\alpha \mid y) \cdot p(f \mid y) .$$

Notera att den marginella posteriorfunktionen för $(\theta_1,\theta_2) \mid y$ innehåller termerna θ_1, θ_2, $\theta_1+\theta_2$, som ska ersättas med parametrarna α och $f(\theta_1,\theta_2)$. Vi har att

$$\alpha = \frac{\theta_1}{\theta_1+\theta_2} \implies \theta_1 = \alpha \cdot (\theta_1+\theta_2)$$, vilket ger direkt från $p[(\theta_1,\theta_2) \mid y]$ att f definierat som $f(\theta_1,\theta_2) = (\theta_1+\theta_2)$

medför en "uppdelning" av termerna α och f vid substitution av $\theta_1 = \alpha \cdot f$, $\theta_1 + \theta_2 = f$, eftersom

$$\theta_1^{a_1+y_1-1} = \alpha^{a_1+y_1-1} \cdot f(\theta_1,\theta_2)^{a_1+y_1-1} \text{, och}$$

$$(1-\theta_1-\theta_2)^{(a_3+y_3)+\dots+(a_J+y_J)-1} = \left(1-f(\theta_1,\theta_2)\right)^{(a_3+y_3)+\dots+(a_J+y_J)-1} .$$

Den kvarvarande termen, θ_2, från $p[(\theta_1,\theta_2)|y]$ ger också denna "uppdelning", ty $\theta_2 = f - \alpha \cdot f = f \cdot (1-\alpha)$, så att

$$\theta_2^{a_2+y_2-1} = \left(f \cdot (1-\alpha)\right)^{a_2+y_2-1} = f^{a_2+y_2-1} \cdot (1-\alpha)^{a_2+y_2-1} .$$

Med $f(\theta_1,\theta_2) = \theta_1 + \theta_2$ beräknar vi till sist $J(\alpha,f)$. Vi har att

$$\frac{d\alpha}{d\theta_1} = \frac{\theta_2}{(\theta_1+\theta_2)^2} , \quad \frac{d\alpha}{d\theta_2} = -\frac{\theta_1}{(\theta_1+\theta_2)^2}$$

$$\frac{df}{d\theta_1} = 1 \qquad \frac{df}{d\theta_2} = 1 \quad \text{, så att}$$

$$J(\alpha,f) = \frac{1}{\begin{vmatrix} \frac{\theta_2}{(\theta_1+\theta_2)^2} & -\frac{\theta_1}{(\theta_1+\theta_2)^2} \\ 1 & 1 \end{vmatrix}} = \frac{1}{\frac{\theta_2}{(\theta_1+\theta_2)^2} - \left(-\frac{\theta_1}{(\theta_1+\theta_2)^2}\right)} = \frac{1}{\frac{1}{\theta_1+\theta_2}}$$

$= \theta_1 + \theta_2 = f(\theta_1,\theta_2)$. Detta ger den simultana täthetsfunktionen för α och f som

$$p[\alpha, f(\theta_1,\theta_2)|y] = \alpha^{a_1+y_1-1} \cdot f(\theta_1,\theta_2)^{a_1+y_1-1} \cdot f(\theta_1,\theta_2)^{a_2+y_2-1} \cdot (1-\alpha)^{a_2+y_2-1} \cdot$$

...

$$\cdot \left(1-f(\theta_1,\theta_2)\right)^{(a_3+y_3)+\cdots+(a_J+y_J)-1} =$$

$$= \underbrace{\alpha^{a_1+y_1-1}\cdot(1-\alpha)^{a_2+y_2-1}}_{=p(\alpha|y)} \cdot p\left[f(\theta_1,\theta_2)|y\right].$$

Alltså, $\alpha|y \sim \text{Beta}(a_1+y_1, a_2+y_2)$.

Exercise 3.1 b

$$y_1 \sim \text{Bin}(n=y_1+y_2, \alpha) \Rightarrow p(y_1|\alpha) \propto \alpha^{y_1}\cdot(1-\alpha)^{y_1+y_2-y_1} = \alpha^{y_1}\cdot(1-\alpha)^{y_2}$$

Visa att

$$\alpha|y_1 \sim \text{Beta}(a_1+y_1, a_2+y_2)$$, d.v.s. samma förd. som i a).

Vi har att

$$p(\alpha|y_1) \propto p(\alpha)\cdot p(y_1|\alpha) \propto p(\alpha)\cdot\alpha^{y_1}\cdot(1-\alpha)^{y_2}$$

Välj priorn $p(\alpha)$ som $\alpha \sim \text{Beta}(a_1,a_2)$, vilket ger

$$p(\alpha|y_1) \propto \alpha^{a_1-1}\cdot(1-\alpha)^{a_2-1}\cdot\alpha^{y_1}\cdot(1-\alpha)^{y_2}$$

$$= \alpha^{a_1+y_1-1}\cdot(1-\alpha)^{a_2+y_2-1}$$, vilket skulle visas.

Binomialfördelningen för y_1 är ett specialfall av multinomialfördelningen. Den Beta-fördelade priorn är ett specialfall av Dirichlets fördelning. Specialfall med tanke på att endast två utfall. Alltså visar detta att om man utgår ifrån ett generellt antal utfall J stycken och ignorerar informationen från övriga utfall utom två stycken erhåller man samma resultat som ignorering av det övriga datamaterialet $y_3,\ldots,y_J$ från början.

Exercise 3.2, Comparison of two multinomial observations

Två undersökningar om röstberättigades preferenser utfördes innan och efter en presidentdebatt under en och samma kväll. 639 stycken valdes ut slumpmässigt och oberoende av varandra från populationen av alla röstberättigade.
Resultat av undersökningarna, se tabell 3.2.
Undersökningarna gav 3 utfall, preferens för Bush, Dukakis eller övriga preferenser.

Låt

$\theta_{1,f/e}$ = andelen preferenser för Bush före/efter debatten

$\theta_{2,f/e}$ = —"— Dukakis —"—

$\theta_{3,f/e}$ = —"— övriga —"—

$Y_{i,f/e}$ = antalet preferenser för Bush/Dukakis/övriga före/efter debatten

$$Y_{f/e} = (Y_{1,f/e}, Y_{2,f/e}, Y_{3,f/e}), \quad \theta_{f/e} = (\theta_{1,f/e}, \theta_{2,f/e}, \theta_{3,f/e})$$

Jobbar med "före" och på samma sätt fungerar det med "efter".

Vi har att $Y_{före} \mid \theta_{före}$ följer en multinomial fördelning

som

$$p(Y_{före} \mid \theta_{före}) \propto \prod_{i=1}^{3} \theta_{i,f}^{Y_{i,f}}$$

Exercise 3.2 forts.

Ansätt en Dirichlet fördelning som konjugerad priorfördelning med parametrar $(a_{1,f}, a_{2,f}, a_{3,f})$. Då ges posteriorfördelningen $\theta_f \mid y_f$ som $p(\theta_f \mid y_f) \propto \prod_{i=1}^{3} \theta_{i,f}^{a_{i,f} + y_{i,f} - 1}$.

Antag att vi inte känner till någonting om röstberättigades preferenser inför undersökningarna. Då är det lämpligt att ansätta en icke-informativ priorfördelning för $\theta_{i,f}$, t.ex. $a_{i,f} = 1$, $i = 1,2,3$

Det ger $p(\theta_f \mid y_f) \propto \prod_{i=1}^{3} \theta_{i,f}^{y_{i,f}}$ så att

$$\theta_f \mid y_f \sim \text{Dirichlet}\left(y_{1,f}+1, y_{2,f}+1, y_{3,f}+1\right) =$$

$$= \text{Dirichlet}\left(295, 308, 39\right).$$

Vi söker dock posteriorfördelningen för $\alpha_f = \dfrac{\theta_{1,f}}{\theta_{1,f} + \theta_{2,f}}$, d.v.s. andelen Bushpreferenser av Bush och Dukakispreferenserna sammanlagda andel.

Den får vi genom att utnyttja resultatet från exercise 3.1a, Vi har då att

$$\alpha_f \mid y \sim \text{Beta}\left(a_1 + y_{1,f}, a_2 + y_{2,f}\right) = \text{Beta}\left(y_{1,f}+1, y_{2,f}+1\right)$$

$$= \text{Beta}\left(295, 308\right).$$

Exercise 3.2 forts.

P. s.s. får vi

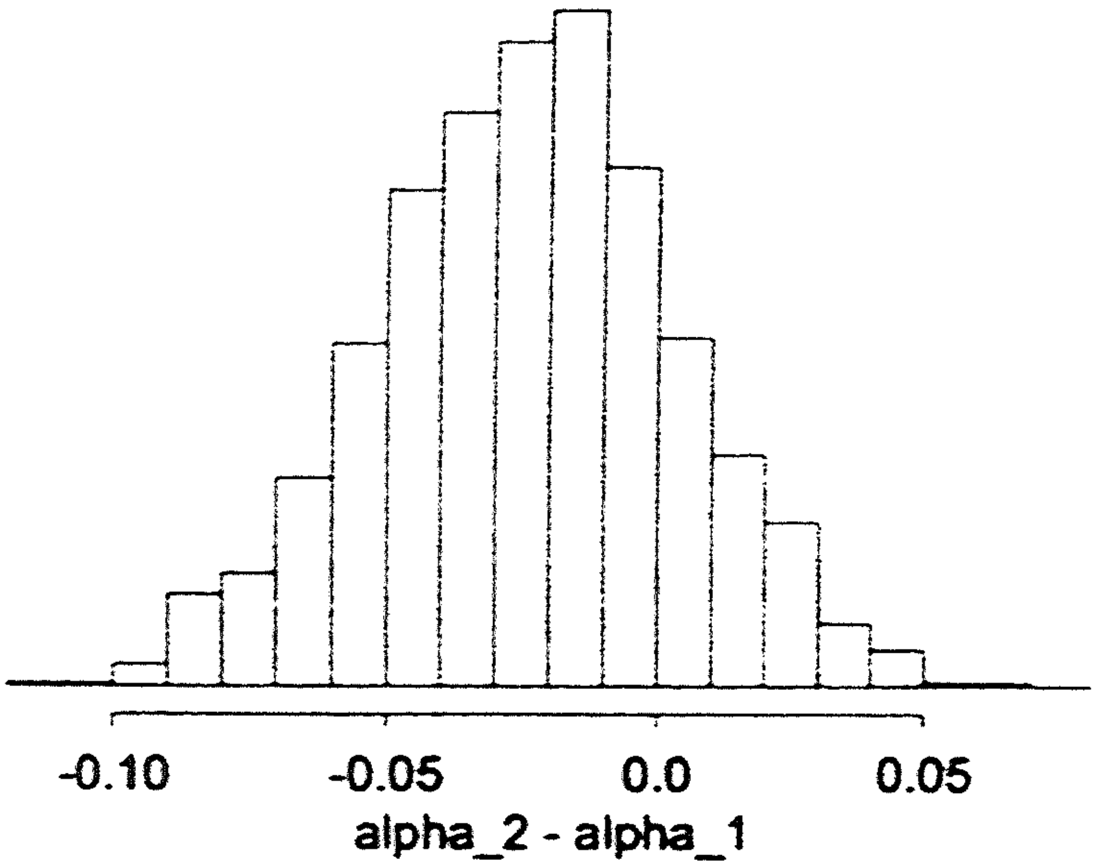
$\alpha_e | y \sim \text{Beta}(289, 333)$.

Nu söker vi posteriorfördelningen för det "skift",

$\alpha_e | y - \alpha_f | y$ som uppstod av debatteffekten.

Slumpa därför 1000 värden från respektive Betafördelning $\theta_e | y$ och $\theta_f | y$, bilda differensen $\theta_{e,(i)} - \theta_{f,(i)}$ för varje slumpmässigt värde där $i = 1, 2, \ldots, 1000$ och plotta differenserna i ett histogram, se $\boxed{\text{figur } 3.1}$

Från histogrammet kan vi utläsa posteriorsannolikheten, att debatten ledde till ett skift mot Bush, som arean under posteriortätheten till höger om 0.



figur 3.1

Exercise 3.5, Rounded data

Ett objekt vägs 5 gånger och de olika vikterna avrundas till närmaste pound. Sätt $y_{round} = y$ och $y_{exakt} = x$. Vi har från uppgiften att

$$y = (y_1, y_2, \dots, y_5) = (10, 10, 12, 11, 9)$$

och vi antar att

$$x_i \overset{i.i.d.}{\sim} N(\mu, \sigma^2)$$

med en icke-informativ prior fördelning för (μ, σ^2) (gavs i kap. 2) som $p(\mu, \sigma^2) \propto \sigma^{-2}$

a) Posteriorfördelningen för (μ, σ^2) ges uttryckligen i kap. 2 om man behandlar de avrundade observationerna som exakta, d.v.s. $(y_1, \dots, y_5) = (x_1, \dots, x_5)$, se sid. 74, 75.

Det gäller att den betingade posteriorfördelningen för μ och den marginella posteriorfördelningen för σ^2, ges som

$$\mu \mid \sigma^2, y \sim N(\bar{y}, \sigma^2/n) = N(10.4, \sigma^2/5)$$

$$\sigma^2 \mid y \sim \text{Inv-}\chi^2(n-1, s^2) = \text{Inv-}\chi^2(4, 1.3)$$

där vi får från $y = (y_1, \dots, y_5)$ att $n = 5$, $\bar{y} = 10.4$, $s^2 = 1.3$

Exercise 3.5 forts.

b) Om däremot observationerna $y=(y_1,\dots,y_5)$ behandlas korrekt som avrundade värden ges posteriorfunktionen för (μ,σ^2) som

$$p(\mu,\sigma^2|y) \propto p(\mu,\sigma^2)\cdot p(y|\mu,\sigma^2)$$

Observera att vi endast känner till fördelningen för de exakta värdena $(x_1,\dots,x_5)$, men täthetsfunktionen $p(y|\mu,\sigma^2)$ inkluderar "alla" täthetsfunktioner för x på intervallet $[y-\frac{1}{2}, y+\frac{1}{2}]$, eftersom

$y_i = 10 \Rightarrow 9{,}5 < x < 10{,}5$. Därför bestämmer vi den summerade täthetsfunktionen på varje intervall genom att integrera täthetsfunktionen för varje icke avrundad observation x_i över intervallet. Vi kan därför skriva

$$p(\mu,\sigma^2|y) \propto \sigma^{-2}\cdot\prod_{i=1}^{5} p(y_i|\mu,\sigma^2) = \sigma^{-2}\cdot\prod_{i=1}^{5}\left[\int_{y_i-\frac{1}{2}}^{y_i+\frac{1}{2}} p(x_i|\mu,\sigma^2)\,dx_i\right]$$

$$= \sigma^{-2}\cdot\prod_{i=1}^{5}\left[\mathbb{P}\left(y_i-\tfrac{1}{2}<x_i\le y+\tfrac{1}{2}\right)\right] = \sigma^{-2}\cdot$$

$$\prod_{i=1}^{5}\left[\mathbb{P}\left(\frac{y_i-\frac{1}{2}-\mu}{\sigma} < \frac{x_i-\mu}{\sigma} \le \frac{y_i+\frac{1}{2}-\mu}{\sigma}\right)\right] =$$

$$= \sigma^{-2}\cdot\prod_{i=1}^{5}\left[\Phi\left(\frac{y_i+\frac{1}{2}-\mu}{\sigma}\right)-\Phi\left(\frac{y_i-\frac{1}{2}-\mu}{\sigma}\right)\right]$$

Exercise 3.5, forts.

c) Simulerar 1000 värden från respektive posteriorfördelning i a) och b).

Posteriorfördelningen från a) är lätt att simulera ifrån. Vi har kända fördelningar att simulera ifrån för den betingade täthetsfunktionen över μ och marginella täthetsfunktionen över σ^2.

Drag först ett slumpmässigt värde $\sigma^2_{(i)}$ från den inversa $-\chi^2$ förd. $\chi^2(4, 1.3)$ och använd sedan detta värde för att dra ett värde $\mu_{(i)}$ från $N(10.4, \sigma^2_{(i)}/5)$. X dras först från $\chi^2(4)$ och sedan beräknas $\sigma^2_{(i)} = \frac{4 \cdot 1.3}{X}$.

Posteriorfördelningen från b) är inte känd och kan inte delas upp i några "komponenter" som i a). Bestäm därför först $p(\sigma^2|y)$ genom att approximera denna marginella täthetsfunktion som $p(\sigma^2|y) \approx \sum_{\beta} p(\sigma^2|y, \beta)$.

Beräkna den kumulativa fördelningen,

$$\sum_{\sigma^2} p(\sigma^2|y)$$

för respektive värde σ^2.

Normalisera den kumulativa fördelningen som

$$P_N(\sigma^2|y) = \frac{p(\sigma^2|y)}{p_{max}(\sigma^2|y)}$$

. Slumpa sedan fram ett värde u från $U[0,1]$ och finn det värde σ^2 där $u = p_N(\sigma^2|y)$.

Exercise 3.5 forts..

På liknande sätt kan man simulera fram värden $\mu_{(i)}$ genom att approximera den betingade täthetsfunktionen $p(\mu | \sigma^2, y)$ givet värden $\sigma^2_{(i)}$.

$\boxed{\text{Tabell 3.1}}$ ger skillnaden mellan posteriorfördelningarna och $\boxed{\text{figur 3.2}}$ och $\boxed{3.3}$ åskådliggör contour-plottarna för respektive posteriorfördelning.

d) Drag simuleringar från posteriorfördelningen av x.

Vi har att

$$x_i \overset{\text{i.i.d.}}{\sim} N(\mu, \sigma) \quad ; \quad -\infty < x_i < \infty$$

Givet observation y_i, vet vi att $y_i - \frac{1}{2} < x_i < y_i + \frac{1}{2}$.

Det ger att $x_i | \mu, \sigma^2, y_i$ är en trunkerad normalfördelning på intervallet ovan. Alltså,

$$x_i | \mu, \sigma^2, y_i \sim N(\mu, \sigma^2) \quad ; \quad y_i - \frac{1}{2} < x_i < \frac{1}{2}.$$

På liknande sätt som tidigare kan vi approximera och normalisera fördelningsfunktionen för varje $N(\mu_{(i)}, \sigma^2_{(i)})$ och sedan bestämma det korresponderande värdet $x_{(i)}$ där den trunkerande fördelningsfunkt. för $N(\mu_{(i)}, \sigma^2_{(i)})$ är lika med ett slumpmässigt värde u från $U[0,1]$.

Exercise 3.12, Poisson regression model

Betrakta exercise 2.13 a)

I tabell 2.2 ges ant. livshotande olyckor för schemalagda flygningar per år mellan åren 1976 – 1985

Låt y_i = ant. livshotande olyckor för år i ; $i = 1, 2, \dots, 10$

$y_i \sim \text{Poisson}(\theta = \alpha + \beta \cdot i)$, tid. $y_i \overset{\text{i.i.d.}}{\sim} \text{Poisson}(\theta)$

$$\Longrightarrow \quad p(y_i \mid \theta = \alpha + \beta \cdot i) = \frac{1}{y_i!} \cdot (\alpha + \beta \cdot i)^{y_i} \cdot \exp\big(-(\alpha + \beta \cdot i)\big)$$

Vi vill dock försäkra oss om att $\theta > 0$, så att låt istället $\log \theta = \alpha + \beta \cdot i \Rightarrow \theta = \exp(\alpha + \beta \cdot i)$

och

$$p\big(y_i \mid \theta = \exp(\alpha + \beta \cdot i)\big) = \frac{1}{y_i!} \cdot \big(\exp(\alpha + \beta \cdot i)\big)^{y_i} \cdot \exp\big(-\exp(\alpha + \beta \cdot i)\big)$$

a) Välj $p(\alpha, \beta) \propto 1$ som icke-informativ priorfördelning.

En annan icke-informativ priorfördelning för (α, β) skulle kunna väljas som en bivariat normalfördelning med höga varianser σ_α^2 och σ_β^2 för att skapa en så pass flack priorfördelning att man kan använda den som en icke-informativ priorfördelning.

Exercise 3.12 forts.

b) En realistisk informativ priorfördelning för (α, β) kan vara att utgå ifrån y_0, d.v.s. antalet livshotande flygolyckor år 1975 och uppskatta förändringen av antalet olyckor med tidigare data eller erfarenhet. En bivariat normalfördelning, centrerad kring $E[\alpha] \approx \log(y_0)$ och $E[\beta] \approx \hat{\Delta} y_i$, med kovariansmatris

där y_0 mäts i hundratal

$$V_{ij} = \begin{bmatrix} \sigma_\alpha^2 & p \cdot \sigma_\alpha \cdot \sigma_\beta \\ p \cdot \sigma_\alpha \cdot \sigma_\beta & \sigma_\beta^2 \end{bmatrix}.$$

Exempelvis $\mu_\alpha = -1.6$, $\mu_\beta = 0$

$$V_{ij} = \begin{bmatrix} 0.4 & 0.3 \\ 0.3 & 0.25 \end{bmatrix}$$

En contour plot av denna bivariata normalfördelning ges i figur 3.4.

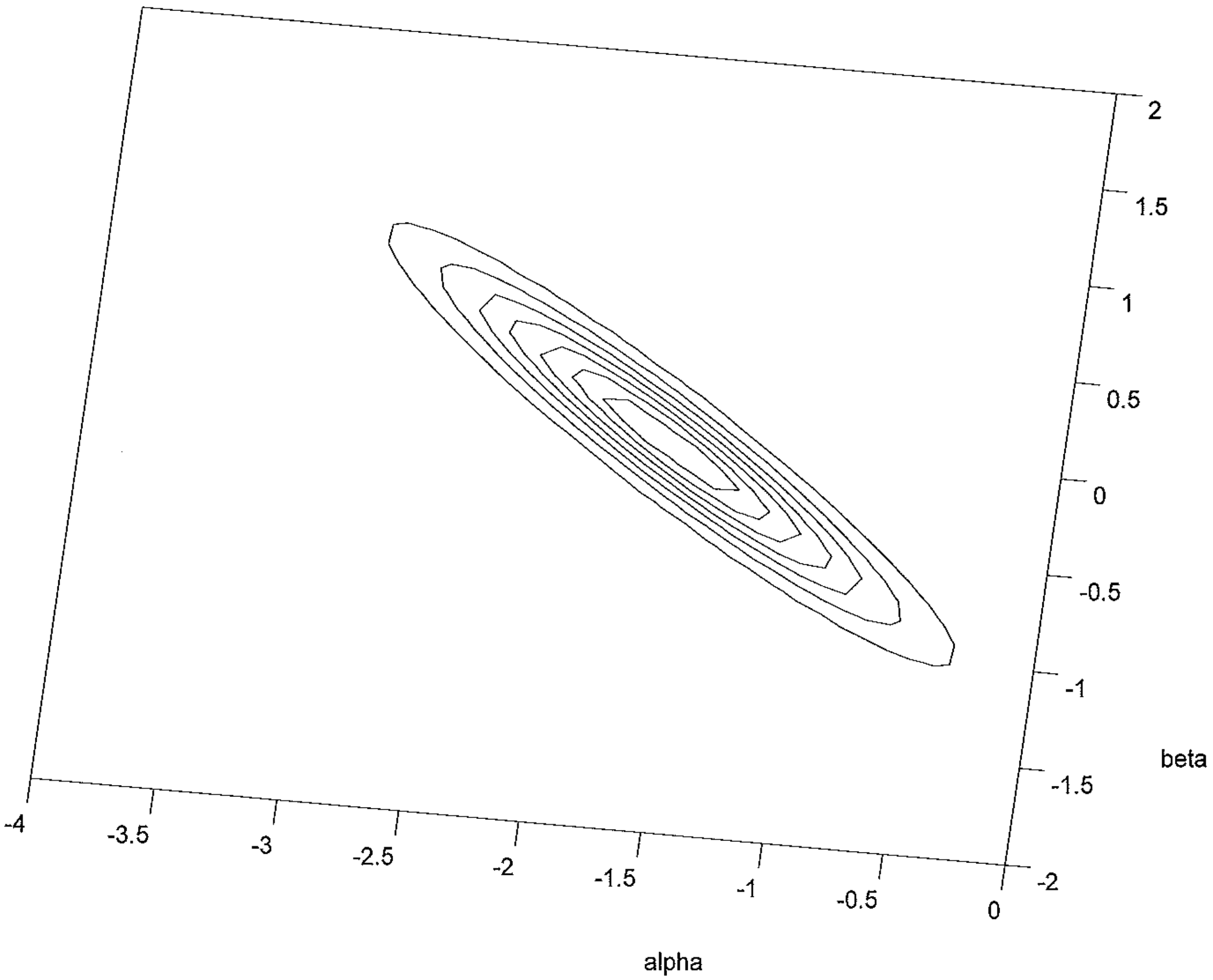
c) Vi har att $p(\alpha, \beta) \propto 1$. Alltså blir posteriorfunktionen för (α, β) proportionell mot Likelihoodfunktionen

$y = (y_1, y_2, \ldots, y_{10}) \mid \alpha, \beta$, så att

$$p(\alpha, \beta \mid y, t) \propto \prod_{i=1}^{10} p\left(y_i \mid \theta = e^{\alpha + \beta \cdot i}\right)$$

$$\propto \prod_{i=1}^{10} \left(\exp(\alpha + \beta \cdot i)\right)^{y_i} \cdot \exp\left(-\exp(\alpha + \beta \cdot i)\right) =$$

Figur 3.4: Contour plot för (alpha,beta) med en informativ bivariat normalfördelning som prior



Exercise 3.12 forts.

c) forts.

$$= e^{\alpha \cdot \sum_{i=1}^{10} y_i} \cdot e^{\beta \cdot \sum_{i=1}^{10} i \cdot y_i} \cdot \prod_{i=1}^{10} \exp\left(-\exp\left(\alpha + \beta \cdot i\right)\right) ,$$

vilket ger de tillräckliga statistikorna som

$$\sum_{i=1}^{10} y_i \quad \text{och} \quad \sum_{i=1}^{10} i \cdot y_i$$

d)

Posteriorfunktionen är proper, ty den består av en produkt av Poisson-sannolikheter, vilket innebär att

$$p(\alpha, \beta \mid y, t) < 1 < \infty \quad , \text{och}$$

då $(\alpha, \beta) \longrightarrow \pm\infty$ går $p(\alpha, \beta \mid y, t)$ mot noll.

Detta ger att $\displaystyle\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} p(\alpha, \beta \mid y, t) \, d\alpha \, d\beta < \infty$

Alltså ges arean under posteriortätheten av ett positivt ändligt tal och det innebär att posteriortätheten är proper.

Exercise 3.12 forts.

e) Vi har att

$$\log \theta = \alpha + \beta \cdot t \,,$$

så att vi kan erhålla en approximativ skattning $(\hat{\alpha}, \hat{\beta})$ genom linjär regression av $\log(y_i)$ på tidspunkterna t_i för de 10 olika åren, där $t_1 = 1, \ldots, t_{10} = 10$.

Datamaterialet ges från tabell 2.2 och regressionsresultatet blev

	skattning	st.avv. skattning
α	$-1{,}231$	$0{,}1153$
β	$-0{,}040$	$0{,}0186$

$\boxed{y_i \text{ mättes i hundratal}}$

f) tabellen i uppg. e) ger oss ledtrådar för vilka värden på (α, β) vi ska skapa en contour plot utifrån, $\boxed{\text{se figur 3.5}}$

$$\sum_{i=1}^{10} y_i = 2{,}38 \,, \quad \sum_{i=1}^{10} i \cdot y_i = 12{,}33 \,.$$ Alltså har vi

$$p(\alpha, \beta \mid y, t) \propto e^{2{,}38 \cdot \alpha + 12{,}33 \cdot \beta} \cdot \prod_{i=1}^{10} \exp\left(-\exp(\alpha + \beta \cdot i)\right)$$

1000 slumpmässiga dragningar från posteriorfunktionen bestäms enligt följande: Approximera den marginella täthetsfunktionen av α som $p(\alpha \mid y, t) \simeq \sum_{\beta} p(\alpha, \beta \mid y, t)$, d.v.s. summera ut β för resp. α

Bestäm den kumulativa fördelningsfunktionen för varje α, i.e.

$\bar{0}3$

$$P_k(\alpha_{(i)}|y,t) \approx \sum_{\alpha=\alpha_{min}}^{\alpha(i)} p(\alpha|y,t)$$ och tag fram den

normerande kumulativa fördelningsfunktionen som

$$P_{kN}(\alpha_{(i)}|y,t) = \frac{P_k(\alpha_{(i)}|y,t)}{P_k(\alpha_{max}|y,t)},$$

definierad på intervallet $[0,1]$.

Nu utnyttjar vi den inversa fördelningsfunktion metoden, så slumpa fram ett värde u från $U[0,1]$ och finn det värde $\alpha_{(i)}$ som uppfyller likheten

$$u = P_{kN}(\alpha_{(i)}|y,t),$$ vilket är det slumpade värdet för α.

Utnyttja det slumpade värdet $\alpha_{(i)}$ för att bestämma den betingade posteriorfunktionen

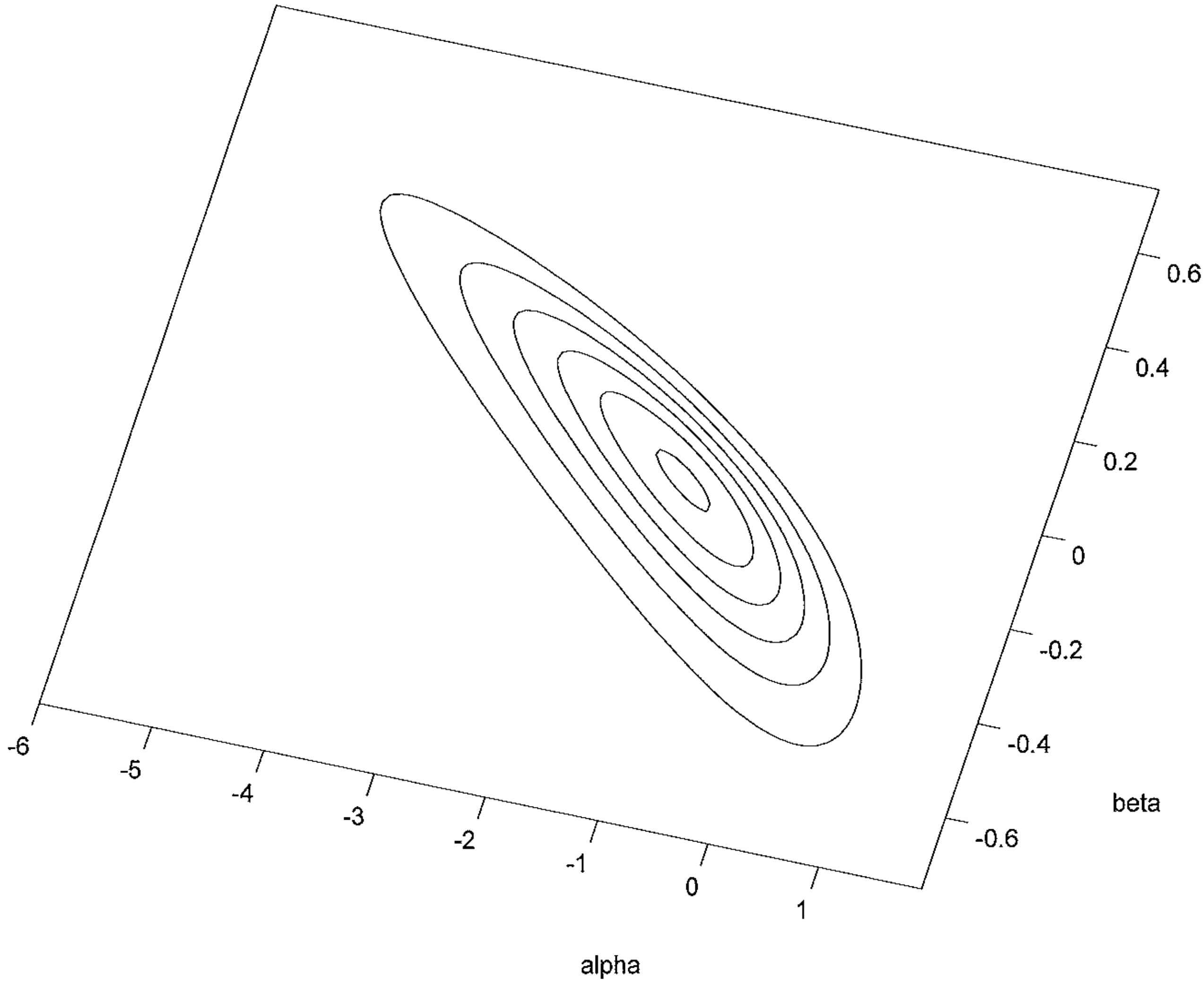
$$p(\beta_{(i)}|y,t,\alpha_{(i)})$$ och den kumulativa fördelningsfunktionen

för varje β. Fortsätt sedan p.s.s. som ovan, givet $\alpha_{(i)}$, för att slumpa fram värdet $\beta_{(i)}$.

Denna procedur upprepas 1000 gånger vilket gav

$$(\alpha_{(i)}, \beta_{(i)}) \quad \text{för} \quad i = 1,2,\dots,1000.$$

Figur 3.5: Contour plot av posteriorfunktion för (alpha,beta) med en icke-informativ prior



Exercise 3.12 forts.

g) Ett histogram för posteriortätheden av

$$\theta_{1986} = E\left[y_i \mid t = 1986\right] = e^{\alpha + \beta \cdot 11}$$

plottas utifrån de slumpade värdena $(\alpha_{(i)}, \beta_{(i)})$, $i = 1, 2, \ldots, 1000$ så att värdena

$$\theta_{(i),1986} = e^{\alpha_{(i)} + \beta_{(i)} \cdot 11} \quad \text{för} \quad i = 1, 2, \ldots, 1000$$

används för att plotta histogrammet, se $\boxed{\text{figur 3.6}}$.

h) Ett 95%-igt prediktionsintervall för y_{1986} ges från värdena $\theta_{(i),1986}$, $i = 1, 2, \ldots, 1000$, där varje $\tilde{y}_{(i),1986}$ slumpas fram från en Poissonfördelning med väntevärde $\theta_{(i),1986}$, ty vi har att

$\tilde{y}_{1986,i} \sim Po(\theta_i)$, där y_i är oberoende observationer för respektive år i. Ordna sedan de predikterade värdena $\tilde{y}_{(i),1986}$ i storleksordning och skatta den undre och övre intervallgränsen som $\hat{\pi}_{0{,}025} = \tilde{y}_{(i),(25)}$ och $\hat{\pi}_{0{,}975} = \tilde{y}_{(i),(975)}$ där $\hat{\pi}_p = \tilde{y}_{(i),(1000 \cdot p)}$ är den skattade p%-percentilen för den $(1000 \cdot p)$:e tagna observationen.

Exercise 3.12 forts.

h)

Om man jämför contour plots mellan

figur 3.4: Contour plot för (α, β) vid användning av den informativa priorfördelningen och

figur 3.5: Contour plot för posteriorfördelning av (α, β) vid användning av en icke-informativ priorfördelning

kan man tydligt urskilja att fördelningen för (α, β) i figur 3.4 är mer koncentrerad för α än i figur 3.5, vilket kan förklaras av att den icke-informativa priorn är tillräckligt dominerande för att posteriorfördelningen av (α, β) innehar en hög variation, vilket innebär att det lilla datamaterialet för y_i inte verkar ha fått den effekt för att justera den icke-informativa priorn.

Notera dock att för β är istället posteriorfunktionen för (α, β) mer koncentrerad än fördelningen för den informativa priorfördelningen för (α, β), men om den informativa priorn istället hade använts, skulle förmodligen posteriorfunktionen vara mer koncentrerad än den med icke-informativ prior.