HW 3

Problem 1

```r
x <- c(-.86, -.3, -.05, .73)
n <- rep(5, 4)
y <- c(0, 1, 3, 5)
```

The posterior is described in (3.15) and we set the following prior

$$(\alpha, \beta) \sim \mathrm{N}\left(\begin{pmatrix} 0 \\ 10 \end{pmatrix}, \begin{pmatrix} 4 & 10 \\ 10 & 100 \end{pmatrix}\right).$$

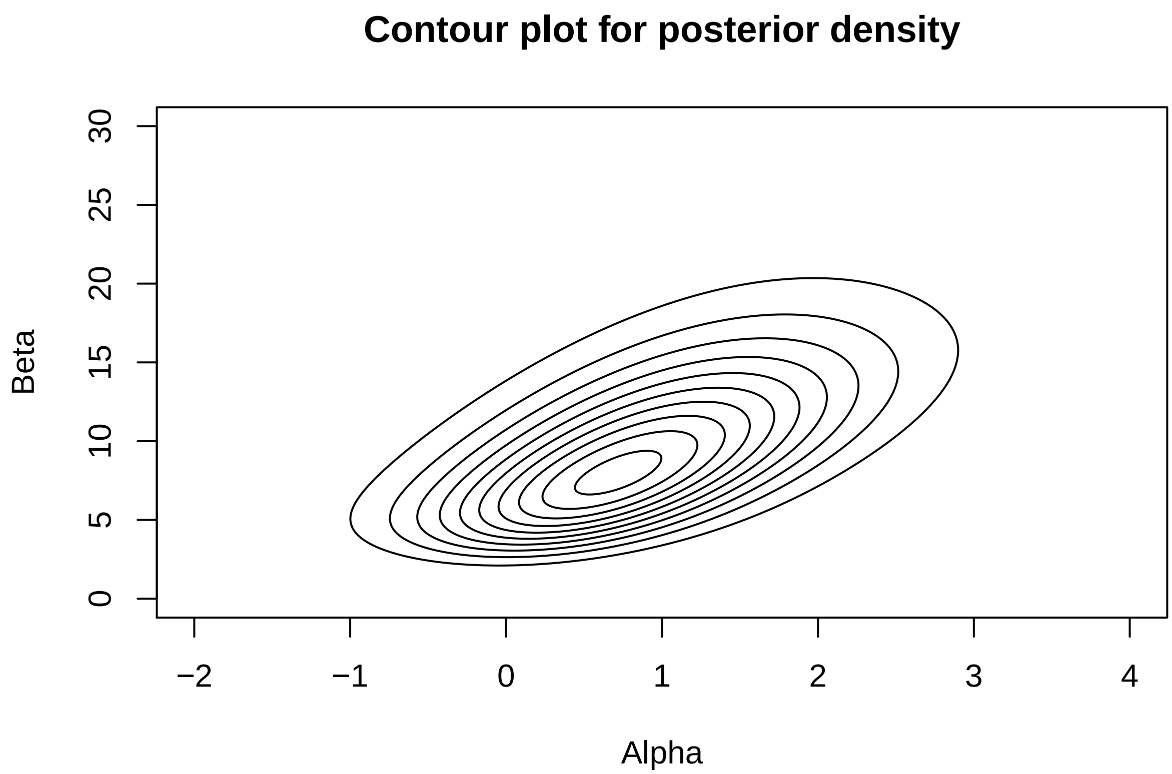
As we can not simplify the posterior, we will not write down the formula here, please use (3.15) for reference. Following is the sampling procedure. We will try to replicate all the figures in section 3.7.

```r
library(mvtnorm)
mu <- c(0, 10)
Sig <- matrix(c(4, 10, 10, 100), 2, 2)

link <- function(x) {
  1/(1+exp(-x))
}

log_posterior <- function(alpha, beta) {
  # x,n,y are given and fixed
  par <- c(alpha, beta)
  dmvnorm(par, mean=mu, sigma=Sig, log=T) +
    sum(log(link(par[1]+par[2]*x)^y*(1-link(par[1]+par[2]*x))^(n-y)))
}

iter <- 1e5
res <- matrix(nrow=iter+1, ncol=2)
res[1, ] <- c(0, 10) # init
for (i in 1:iter) {
  temp_alpha <- rnorm(1, res[i, 1], 2) # proposal
  log_ratio <- log_posterior(temp_alpha, res[i, 2]) -
    log_posterior(res[i, 1], res[i, 2]) +
    dnorm(temp_alpha, res[i, 1], 2, log=T) -
    dnorm(res[i, 1], temp_alpha, 2, log=T)
  temp_log_U <- log(runif(1))
  if (temp_log_U <= log_ratio) {
    res[i+1, 1] <- temp_alpha
  } else {
    res[i+1, 1] <- res[i, 1]
  }

  temp_beta <- rnorm(1, res[i, 2], 10) # proposal
  log_ratio <- log_posterior(res[i+1, 1], temp_beta) -
```

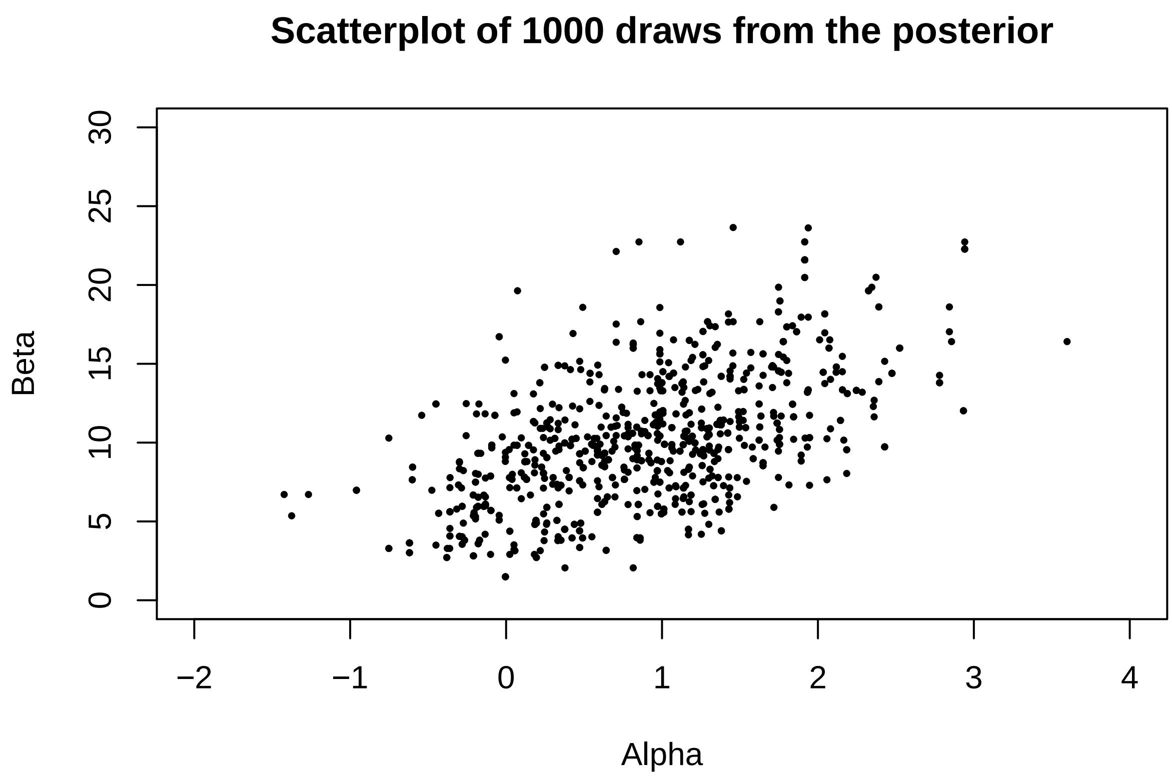
```
    log_posterior(res[i+1, 1], res[i, 2]) +
    dnorm(temp_beta, res[i, 2], 10, log=T) -
    dnorm(res[i, 2], temp_beta, 10, log=T)
  temp_log_U <- log(runif(1))
  if (temp_log_U <= log_ratio) {
    res[i+1, 2] <- temp_beta
  } else {
    res[i+1, 2] <- res[i, 2]
  }
}
```

```
library(ggplot2)
xgrid <- seq(-2, 4, .01)
ygrid <- seq(0, 30, .05)
z <- matrix(nrow=length(xgrid), ncol=length(ygrid))
for (i in 1:length(xgrid)) {
  for (j in 1:length(ygrid)) {
    z[i, j] <- exp(log_posterior(xgrid[i], ygrid[j]))
  }
}
```

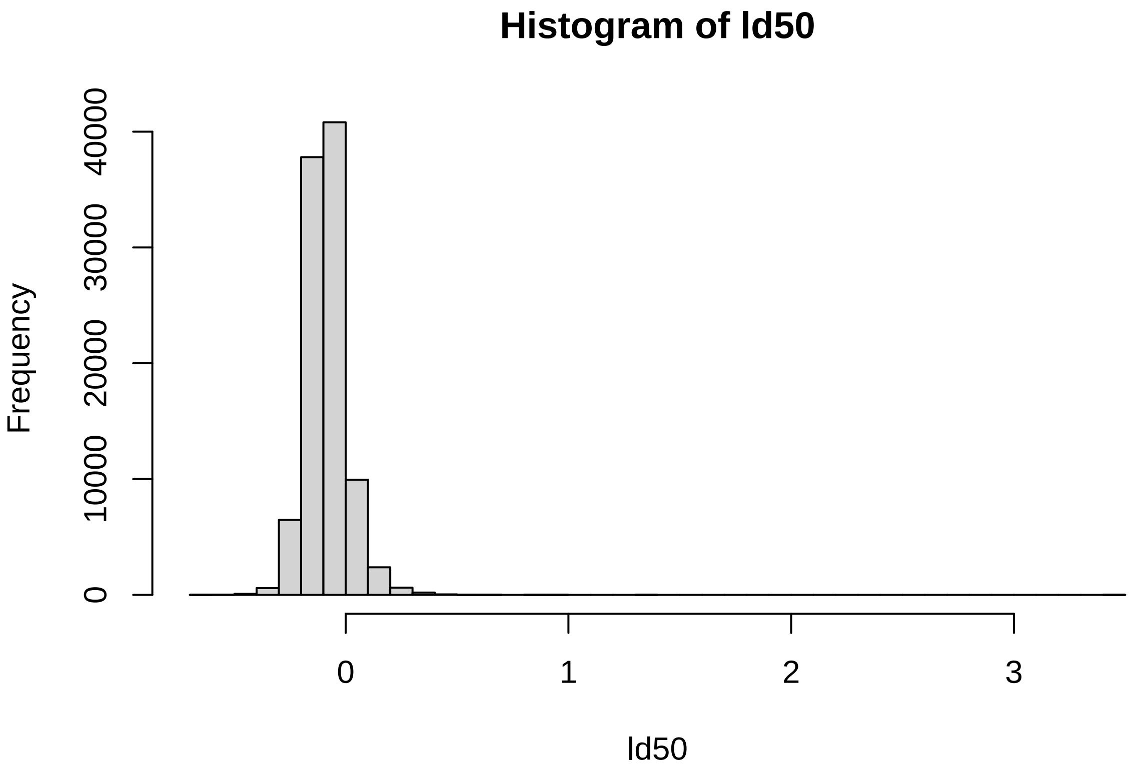
```
contour(x=xgrid, y=ygrid, z=z, drawlabels=F,
        main='Contour plot for posterior density',
         xlab='Alpha', ylab='Beta')
```




```
plot(tail(res[, 1], 1000), tail(res[, 2], 1000),
     xlim=c(-2, 4), ylim=c(0, 30), pch=16, cex=.5,
     main='Scatterplot of 1000 draws from the posterior',
     xlab='Alpha', ylab='Beta')
```




```
#LD50
res <- res[-c(1:1000), ] # burn-in
ld50 <- -res[, 1]/res[, 2]
ld50 <- ld50[res[, 2] > 0] #conditional on beta > 0

hist(ld50, breaks=50)
```



Comparing to the uniform prior, this normal prior allow you to specify the location of the parameters with your existing knowledge.

Problem 2

Suppose sample (x_i, n_i, y_i), $i = 1, ..., k$ satisfies three assumptions

$$\begin{aligned} &(1) \quad x_i \neq x_j \quad \forall i, j, \\ &(2) \quad k > 1, \\ &(3) \quad 0 < y_i < n_i \quad \forall i. \end{aligned}$$

The first assumption can always be achieved because if $\exists i, j$ such that $x_i = x_j$, then we can define $y^* = y_i + y_j$, $n^* = n_i + n_j$, $x^* = x_i = x_j$ and drop (x_i, n_i, y_i) and (x_j, n_j, y_j).

We will show that the posterior is proper under these assumptions. With uniform prior on α and β, the posterior is

$$p(\alpha, \beta | x, n, y) \propto p(\alpha, \beta) p(y | \alpha, \beta, x, n) \tag{1}$$

$$\propto \prod_{i=1}^{k} \left(\frac{1}{1 + \exp(-\alpha - \beta x_i)} \right)^{y_i} \left(1 - \frac{1}{1 + \exp(-\alpha - \beta x_i)} \right)^{n_i - y_i}. \tag{2}$$

To check whether it is proper, we show that

$$
\begin{aligned}
\int p(\alpha,\beta|x,n,y)d\alpha d\beta &\propto \int \prod_{i=1}^{k}\left(\frac{1}{1+\exp(-\alpha-\beta x_i)}\right)^{y_i}\left(1-\frac{1}{1+\exp(-\alpha-\beta x_i)}\right)^{n_i-y_i} d\alpha d\beta && (3)\\
&\le \int \prod_{i=1}^{2}\left(\frac{1}{1+\exp(-\alpha-\beta x_i)}\right)^{y_i}\left(1-\frac{1}{1+\exp(-\alpha-\beta x_i)}\right)^{n_i-y_i} d\alpha d\beta && (4)\\
&= \int \prod_{i=1}^{2} {p_i}^{y_i}(1-p_i)^{n_i-y_i}|\det(d\phi)(p_1,p_2)|dp_1dp_2, && (5)
\end{aligned}
$$

where $\phi : (p_1,p_2) \mapsto \left(\frac{x_2}{x_2-x_1}\text{logit}(p_1) - \frac{x_1}{x_2-x_1}\text{logit}(p_2), \frac{\text{logit}(p_1)-\text{logit}(p_2)}{x_2-x_1}\right)$ is the inverse function of $\varphi : (\alpha,\beta) \mapsto \left(\frac{1}{1+\exp(-\alpha-\beta x_1)}, \frac{1}{1+\exp(-\alpha-\beta x_2)}\right)$. The logit function is defined as $\log(x/(1-x))$. The second line is because $p_i \in (0,1)$ for all i and all α,β. Please note that this is only possible when we don't have duplicate x and $k \ge 2$.

Direct calculation gives

$$
|\det(d\phi)(p_1,p_2)| = \left|\det\begin{pmatrix} \frac{x_2}{x_2-x_1}\frac{1}{p_1(1-p_1)} & -\frac{x_1}{x_2-x_1}\frac{1}{p_2(1-p_2)} \\ -\frac{1}{x_2-x_1}\frac{1}{p_1(1-p_2)} & \frac{1}{x_2-x_1}\frac{1}{p_2(1-p_2)}\end{pmatrix}\right| = \left|\frac{1}{x_2-x_1}\frac{1}{p_1(1-p_1)}\frac{1}{p_2(1-p_2)}\right|,
$$

and above integral is just

$$
\begin{aligned}
&\left|\frac{1}{x_2-x_1}\right| \int \prod_{i=1}^{2} {p_i}^{y_i}(1-p_i)^{n_i-y_i}\frac{1}{p_1(1-p_1)}\frac{1}{p_2(1-p_2)}dp_1dp_2 && (6)\\
\propto &\left|\frac{1}{x_2-x_1}\right| < \infty, && (7)
\end{aligned}
$$

because we form the p.d.f. of two independent Beta random variables $\text{Beta}(y_1, n_1-y_1)$ and $\text{Beta}(y_2, n_2-y_2)$ in the integral (assumption (3) ensures that those are proper Beta random variables).

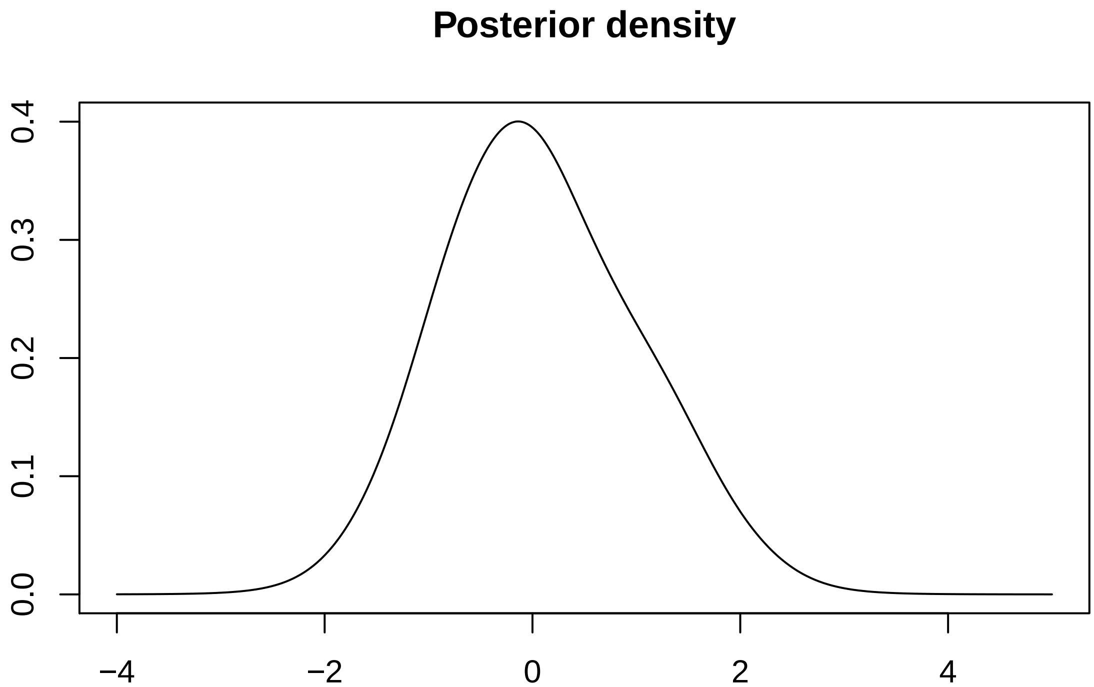
This answer is inspired by https://stats.stackexchange.com/questions/390993/verifying-a-posterior-is-proper. The assumption $0 < y_i < n_i$ for all i could be replaced by a weaker form $0 < \sum_{i=1}^{k} y_i < \sum_{i=1}^{k} n_i$ as pointed out by Christian Robert but it seems that in above proof we can not replace it.

Problem 3

```r
y <- c(-2, -1, 0, 1.5, 2.5)

temp <- function(theta) {
  1/pi^5/(1+(y[1]-theta)^2)/(1+(y[2]-theta)^2)/(1+(y[3]-theta)^2)/
    (1+(y[4]-theta)^2)/(1+(y[5]-theta)^2)
}
const <- integrate(temp, -100, 100)$value # Normalizing constant for the posterior

posterior <- function(theta) {
  prod(1/pi/(1+(y-theta)^2)) / const
}
```

Plot the posterior density

```r
grid <- seq(-4, 5, .01)
z <- c()
for (x in grid) z <- c(z, posterior(x))
plot(grid, z, type='l', main='Posterior density', xlab='', ylab='')
```



Derivatives

The log posterior is

$$l(\theta|y) = -\sum_{i=1}^{n} \log\left(1 + (y_i - \theta)^2\right),$$

and the derivatives are

$$\frac{d\ l(\theta|y)}{d\theta} = \sum_{i=1}^{n} \frac{2(y_i - \theta)}{1 + (y_i + \theta)^2}, \tag{8}$$

$$\frac{d^2\ l(\theta|y)}{d\theta^2} = \sum_{i=1}^{n} \frac{2(y_i - \theta)^2 - 2}{\left(1 + (y_i - \theta)^2\right)^2}. \tag{9}$$

Find the posterior mode

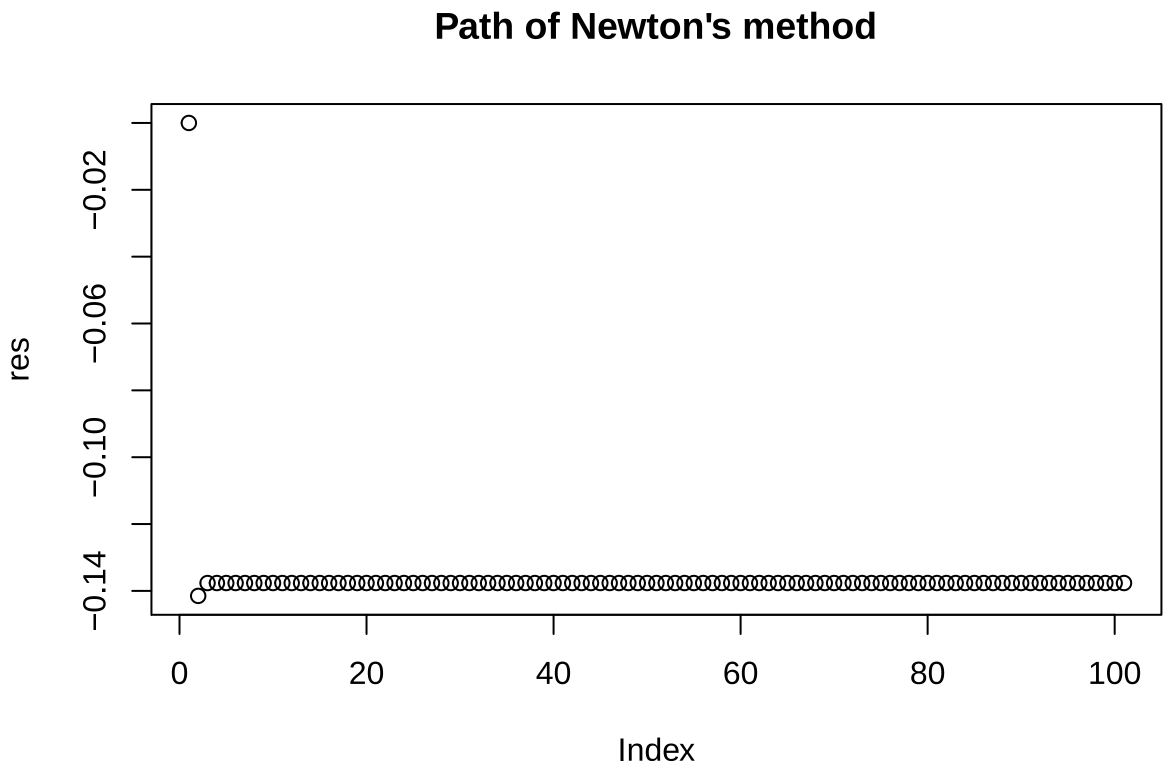
I don't think you can solve it by hand. Since we already have the second order derivative I think it's better to use Newton's method to get the solution.

```r
d1 <- function(x) {
  sum((y-x)/(1+(y-x)^2))
}

d2 <- function(x) {
  sum(((y-x)^2-1)/(1+(y-x)^2)^2)
}
```

```r
iter <- 1e2
res <- rep(0, iter)
for (i in 1:iter) {
  res[i+1] <- res[i] - d1(res[i])/d2(res[i])
}
tail(res)
```

```
## [1] -0.1376493 -0.1376493 -0.1376493 -0.1376493 -0.1376493 -0.1376493
```

```r
plot(res, main="Path of Newton's method")
```



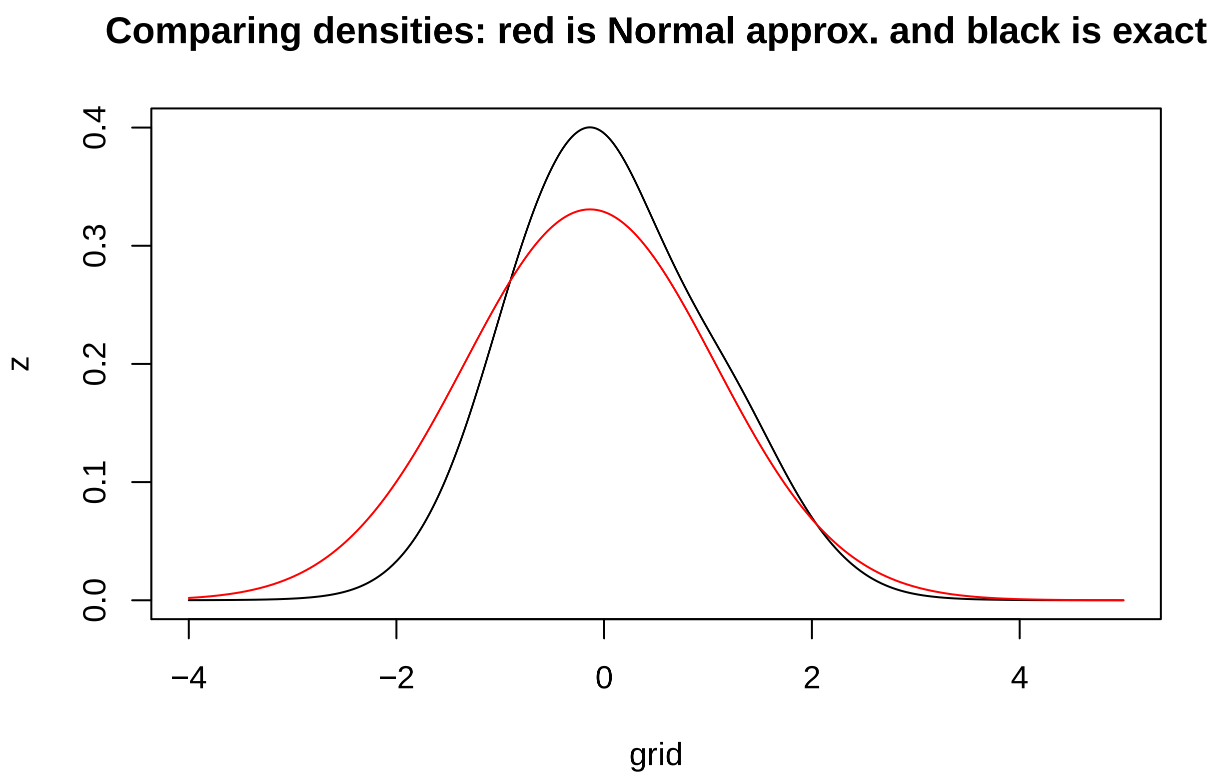
Normal approximation

Let $\hat{\theta}$ be the mode, then

$$p(\hat{\theta}|y) \approx \mathrm{N}\left(\hat{\theta}, I^{-1}(\hat{\theta})\right),$$

where $I(\hat{\theta}) = -\frac{d^2}{d\theta^2} \log p(\theta|y)|_{\hat{\theta}}$.

```r
mu <- tail(res, 1)
sig <- sqrt(-1/d2(mu))

grid <- seq(-4, 5, .01)
z2 <- c()
for (x in grid) z2 <- c(z2, dnorm(x, mean=mu, sd=sig))
plot(grid, z, type='l', main='Comparing densities: red is Normal approx. and black is exact')
lines(grid, z2, col='red')
```


Comparing densities: red is Normal approx. and black is exact
z
0.0
0.1
0.2
0.3
0.4
−4
−2
0
2
4
grid