HW1

Jiahao Tian

2023-01-15

Question 1

part a: method of moments estimators of (α, β).

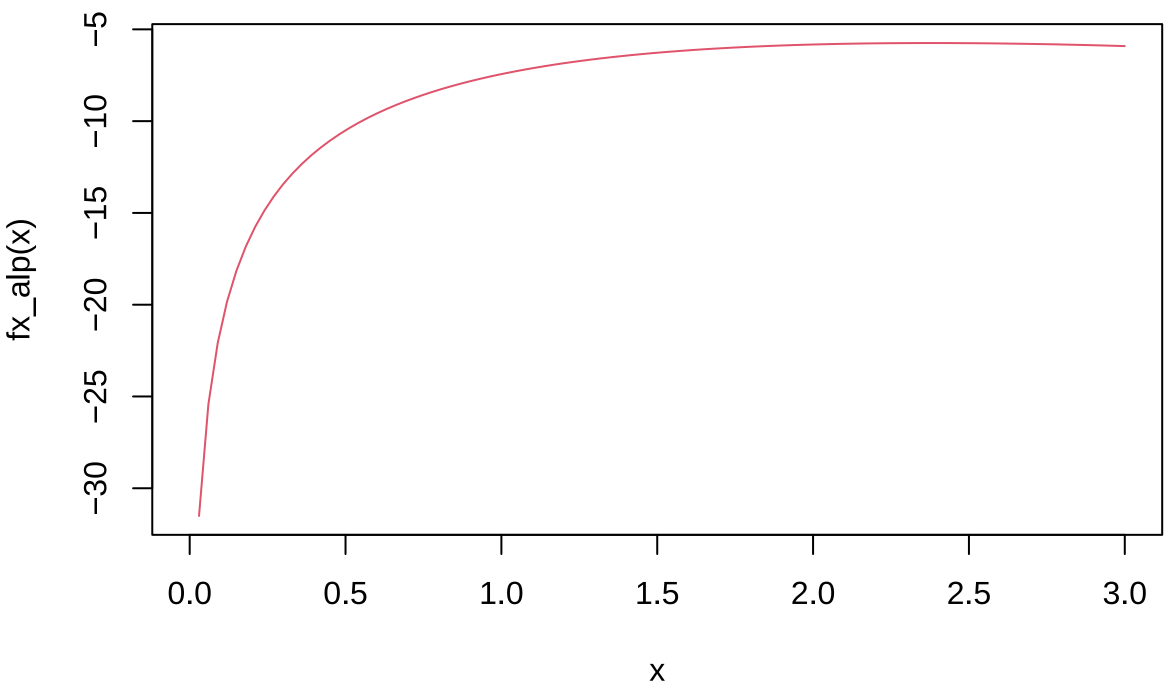
$$
\begin{aligned}
x \sim \Gamma(\alpha, \beta) \ \ with \ \ E(x) = \frac{\alpha}{\beta} \ \ and \ \ Var(x) = \frac{\alpha}{\beta^2} \\
E(x^2) &= \int_0^\infty \frac{\beta^\alpha}{\Gamma(a)} x^{\alpha+2-1} e^{-\beta x} dx \\
&= \frac{\Gamma(\alpha+2)}{\Gamma(\alpha)\beta^2} \int_0^\infty \frac{\beta^{\alpha+2}}{\Gamma(\alpha+2)} x^{\alpha+2-1} e^{-\beta x} dx \\
&= \frac{\alpha(\alpha+1)}{\beta^2} \\
&= \frac{\alpha^2}{\beta^2} + \frac{\alpha}{\beta^2} \\
Var(x) &= \frac{\alpha}{\beta^2} = E(x^2) - (E(x))^2 \\
\beta &= \frac{\alpha}{\beta} / \frac{\alpha}{\beta^2} = \frac{E(x)}{E(x^2) - (E(x))^2} \\
which \ \hat{\beta}_{mom} &= \frac{\bar{x}}{\frac{1}{n}\sum_{i=1}^n (x_i - \bar{x})^2} \\
\alpha &= \frac{\alpha}{\beta} \cdot \left(\frac{\alpha}{\beta} / \frac{\alpha}{\beta^2}\right) = \frac{(E(x))^2}{E(x^2) - (E(x))^2} \\
which \ \hat{\alpha}_{mom} &= \frac{(\bar{x})^2}{\frac{1}{n}\sum_{i=1}^n (x_i - \bar{x})^2}
\end{aligned}
$$

part b: maximum likelihood estimator of β in closed form.

$$
\begin{aligned}
Lf(x|\alpha, \beta) &= \left(\frac{\beta^\alpha}{\Gamma(\alpha)}\right)^n \Pi_{i=1}^n xi^{\alpha-1} e^{-\beta \sum xi} \\
\nabla log(Lf(x|\beta)) &= \frac{n\alpha}{\beta} - \sum xi, \ \ set \ \ to \ \ eqaul \ \ 0 \\
\hat{\beta}_{mle} &= \frac{\alpha}{\bar{x}}, \ \ by \ \ invariance \ \ of \ \ mle \\
\nabla^2 log(Lf(x|\beta)) &= -\frac{n\alpha}{\beta^2} < 0
\end{aligned}
$$

part c

```r
# Consider the following data set:
# x = (1.33, 1.60, 0.68, 0.28, 1.22, 0.72, 0.16, 0.32, 0.97, 0.46);
fx_alp = function(alp) {
  n = length(x)
  betahat = alp / mean(x)
  n * alp * log(betahat) - n * log(gamma(alp)) + (alp - 1) * sum(log(x)) - (betahat * sum(x))
  }

x = c(1.33, 1.60, 0.68, 0.28, 1.22, 0.72, 0.16, 0.32, 0.97, 0.46)
alp = seq(0.1, 10, by = 0.1)
post = fx_alp(alp)

curve(fx_alp, from = 0, to = 3, col = 2)
```



Newton-Raphson Method

The process repeated as $x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}$

- First, by calculate the first and second partial derivative of the log-likelihood function with respect to α
- Second, generate the random number data which Gamma distributed use to calculate the MLE
- Third, create the loop function to calculate the sum of the partial derivatives, the gradient vector, the Hessian matrix, and the MLE approximated value
- Last, until the difference between x^{n+1} and x^n is smaller than epsilon (very small value), such that use the MME for the initial value of β^0, and stop the approximation when $|\hat{\beta}^{(n+1)} - \hat{\beta}^{(n)}| < 0.0000001$. The MLE of α can be found.

Question 2

(a) derive the posterior distribution of $(\mu,\ \sigma^2)$

$$x \sim N(\mu, \sigma^2)$$

$$\begin{aligned}
P(\mu, \sigma^2|x_1, ..., x_n) &\propto P(x_1, ...x_n|\mu, \sigma^2)P(\mu)P(\sigma^2)\\
&\propto (\sigma^2)^{-\frac{n}{2}} exp\Bigg[-\frac{1}{2\sigma^2}\sum(x_i-\mu)^2\Bigg] exp\Bigg[-\frac{1}{2\frac{\sigma^2}{\lambda_\mu}}(\mu-\epsilon^2)\Bigg](\sigma^2)^{-(\lambda_\sigma+1)}\cdot exp(-\frac{\alpha}{\sigma^2})
\end{aligned}$$

$$\begin{aligned}
P(\sigma^2|x_1, ...x_n) &\propto (\sigma^2)^{-\frac{n}{2}} exp\Bigg[-\frac{1}{2\sigma^2}\sum(x_i-\mu)^2\Bigg](\sigma^2)^{-(\lambda_\sigma+1)}\cdot exp(-\frac{\alpha}{\sigma^2})\\
&\propto (\sigma^2)^{-\frac{n}{2}+\lambda_\sigma+1} exp\Bigg[-\frac{1}{\sigma^2}(\frac{1}{2}\sum(x_i-\mu)^2+\alpha\Bigg]\\
&\quad Subsititute\ parameters\ from\ distribution\ of\ \mu\\
\sigma^2 &\sim \Gamma^{-1}\Bigg(\lambda_\sigma+\frac{n}{2}, \alpha+\frac{1}{2}\sum(x_i-\bar{x})^2+\frac{n\lambda_\mu}{2(n+\lambda_\mu)}(\bar{x}-\epsilon)^2\Bigg)\\
P(\mu|\sigma^2, x_1, ..., x_n) &\propto exp\Bigg[-\frac{1}{2\sigma^2}\sum(x_i-\mu)^2\Bigg] exp\Bigg[-\frac{1}{2\frac{\sigma^2}{\lambda_\mu}}(\mu-\epsilon^2)\Bigg]\\
&\propto exp\Bigg[-\frac{1}{2\sigma^2}(\sum(x_i-\mu)^2+\lambda_\mu(\mu-\epsilon^2)\Bigg]\\
\mu &\sim N\Bigg(\frac{n\bar{x}+\lambda_\mu\epsilon}{n+\lambda_\mu}, \frac{\sigma^2}{n+\lambda_\mu}\Bigg)
\end{aligned}$$

(b) show marginal prior on μ is $\mathbf{T}(2\lambda_\alpha,\ \epsilon,\ \frac{\alpha}{\lambda_\mu\lambda_\sigma}\ )$

- From joint prior (Normal-inverse-gamma distribution) take integral from 0 to ∞, $\frac{df}{d\sigma^2}$.

$$\begin{aligned}
f(\mu, \sigma^2) &= \frac{1}{\sqrt{2\pi\frac{\sigma^2}{\lambda_\mu}}}\frac{\alpha^{\lambda_\sigma}}{\Gamma(\lambda_\sigma)}\Bigg(\frac{1}{\frac{\sigma^2}{\lambda_\sigma}}\Bigg) exp\Bigg(-\frac{2\alpha+(x-\epsilon)^2}{2\frac{\sigma^2}{\lambda_\mu}}\Bigg)\\
f(\mu) &= \frac{1}{\sqrt{2\pi}}\frac{\alpha^{\lambda_\sigma}}{\Gamma(\lambda_\sigma)}\int_0^\infty\Bigg(\frac{1}{\frac{\sigma^2}{\lambda_\sigma}}\Bigg)^{\lambda_\sigma+1/2+1} exp\Bigg(-\frac{2\alpha+(x-\epsilon)^2}{2\frac{\sigma^2}{\lambda_\mu}}\Bigg)d\sigma^2\\
since &\int_0^\infty \Gamma^{-1}(x)dx = 1\ \ and\ \ \int_0^\infty x^{-(a+1)}e^{-b/x}dx = \Gamma(a)b^{-a}\\
f(\mu) &\propto \Gamma(\lambda_\sigma+\frac{1}{2})\Bigg(\frac{2\alpha+(x-\epsilon)^2}{\frac{2}{\lambda_\mu}}\Bigg)^{-(\lambda_\sigma+\frac{1}{2})}\\
f(\mu) &\propto \Bigg(1+\frac{(x-\epsilon)^2}{\frac{2\alpha}{\lambda_\mu}}\Bigg)^{-(\lambda_\sigma+\frac{1}{2})}\\
\mu &\sim T(2\lambda_\sigma, \epsilon, \frac{\alpha}{\lambda_\mu\lambda_\sigma})
\end{aligned}$$

(c) give the corresponding marginal prior on σ^2.

$$f(\mu,\sigma^2) = \frac{1}{\sqrt{2\pi\frac{\sigma^2}{\lambda_\mu}}}\frac{\alpha^{\lambda_\sigma}}{\Gamma(\lambda_\sigma)}\bigg(\frac{1}{\frac{\sigma^2}{\lambda_\sigma}}\bigg)exp\bigg(-\frac{2\alpha+(x-\epsilon)^2}{2\frac{\sigma^2}{\lambda_\mu}}\bigg)$$

$$f(\sigma^2) \propto \frac{\alpha^{\lambda_\sigma}}{\Gamma(\lambda_\sigma)}\bigg(\frac{1}{\frac{\sigma^2}{\lambda_\sigma}}\bigg)^{\lambda_\sigma+1} exp\bigg(-\frac{\alpha}{\frac{\sigma^2}{\lambda_\mu}}\bigg)$$

$$Since\ \sigma^2\ is\ not\ dependent\ on\ \mu$$

$$\sigma^2 \sim \Gamma^{-1}(\lambda_\sigma,\alpha)$$

Question 3

part i:

Diaconis, Ylvisaker

$$y|\theta \sim Bin(n,\theta)$$

$$\begin{aligned}
P(y|\theta) &= \binom{n}{y}\theta^y(1-\theta)^{n-y}\\
&\propto exp\bigg[ylog(\theta)+(n-y)log(1-\theta)\bigg]\\
&\propto exp\bigg[ylog(\frac{\theta}{1-\theta})+nlog(1-\theta)\bigg]\\
&\propto exp\bigg[\phi log(\frac{\theta}{1-\theta})+n\lambda log(1-\theta)\bigg]\\
&\propto exp\bigg[\phi log(\theta)+\lambda log(1-\theta)-\phi log(1-\theta)\bigg]\\
&\propto \theta^\phi(1-\theta)^{\lambda-\phi} \sim Beta(\phi,\lambda-\phi)
\end{aligned}$$

Jeffreys

$$\begin{aligned}
f(y|\theta) &= \binom{n}{y}\theta^y(1-\theta)^{n-y}\\
logLf(y|\theta) &\propto ylog(\theta)+(n-y)log(1-\theta)\\
\triangledown^2 logLf(y|\theta) &\propto -\frac{y}{\theta^2}-\frac{n-y}{(1-\theta)^2}\\
Fisher\ Information: &\propto \frac{1}{\theta(1-\theta)}\\
by\ \ |I(\theta)|^{\frac{1}{2}}&\\
&\propto \theta^{-\frac{1}{2}}(1-\theta)^{-\frac{1}{2}}\\
Jeffreys\ prior\ is\ \ &\theta^{\frac{1}{2}-1}(1-\theta)^{\frac{1}{2}-1} \sim Beta(\frac{1}{2},\frac{1}{2})
\end{aligned}$$

part ii:

Diaconis, Ylvisaker

$$
\begin{aligned}
y|\theta &\sim Poisson(\theta) \\
P(y|\theta) &= \frac{e^{-\theta}\theta^{y}}{y!} \\
&\propto exp\Big[-\theta + ylog(\theta)\Big] \\
&\propto exp\Big[-\theta\lambda + \phi log(\theta)\Big] \\
&\propto e^{-\theta\lambda}\theta^{\phi+1-1} \sim \Gamma(\phi+1, \lambda)
\end{aligned}
$$

Jeffreys

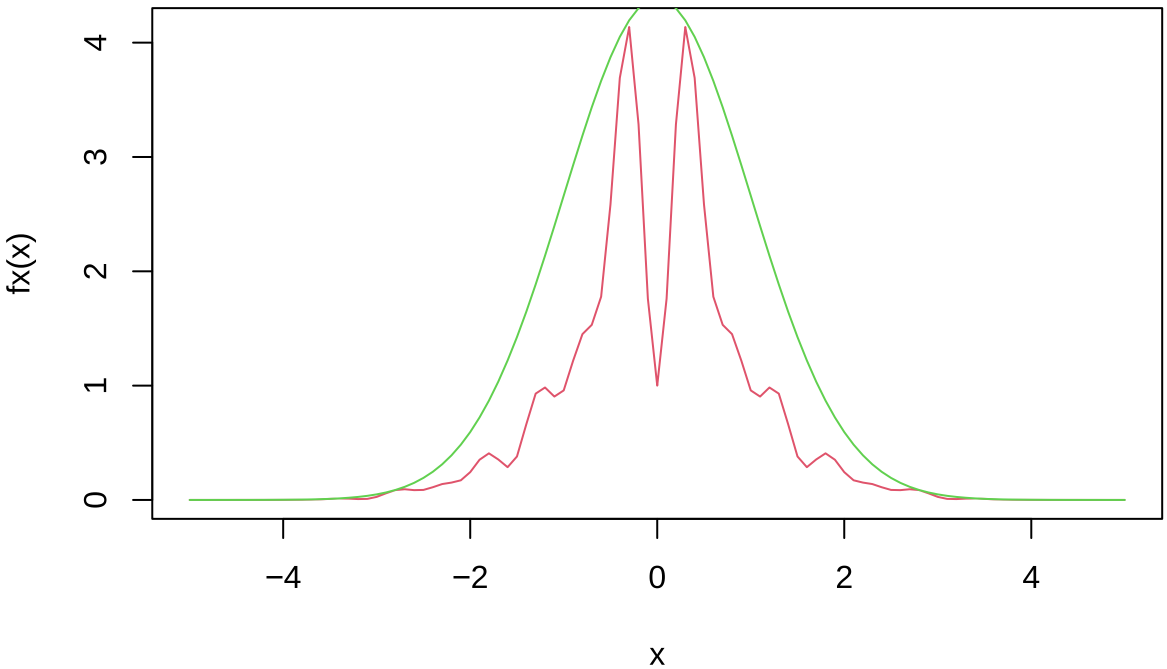
$$
\begin{aligned}
P(y|\theta) &= \frac{e^{-\theta}\theta^{x}}{y!} \\
logLf(y|\theta) &\propto -\theta + ylog(\theta) \\
\nabla^2 logLf(y|\theta) &\propto -\frac{y}{\theta^2} \\
Fisher\ Information: &\propto \frac{1}{\theta} \\
by\ \ &|I(\theta)|^{\frac{1}{2}} \\
Jeffreys\ prior\ is\ &\frac{1}{\sqrt{\theta}} \sim \Gamma(\frac{1}{2}, 0)
\end{aligned}
$$

Question 4

(a) Plot f(x) and show that it can be bounded by Mg(x), where $g = \frac{exp(-x^2/2)}{(2\pi)}$. Find the an acceptable, if not optimal value for M. (Hint: use optimize()).

```
fx = function(x) {
  exp(-x^2 / 2) * (sin(6 * x)^2 + 3 * cos(x)^2 * sin(4 * x)^2 + 1 )
}

gx = function(x) {
  exp(-x^2 / 2) / sqrt(2 * pi)
}

Mgx = function(x) {
  11 * (exp(-x^2 / 2) / sqrt(2 * pi)) #plug M back in
}
```

```r
set.seed(1166)
curve(fx, from = -5, to = 5, col = 2)
curve(Mgx, from = -5, to = 5, col = 3, add = TRUE)
```




```r
opfx = function(x) {
  fx(x) / gx(x)
}

max = optimise(opfx, c(0.1, 10), maximum = TRUE)
max
```

```
## $maximum
## [1] 3.464734
##
## $objective
## [1] 10.94031
```

- f(x) can be bounded by Mg(x) when M = 11.

(b)

```r
AR = function (f, g, m, n_iter) {
  #x1 = rep(NA, n_iter)
  accp = 0
  for (i in 1:n_iter) {
    while (accp) {
      x = rnorm(1)
      u = runif(1)
```

```r
      if (u <= (f(x) / (m * g(x)))) {
        accp = accp + 1
        #x1[i] = x
      }
    }
  }
  list(accp = accp)
  }
```

```r
set.seed(123)
ar = AR(f = fx, g = gx, m = 11, n_iter = 2500)
ar$accp
```

```
## [1] 0
```

(c)

- All rejected, accept rate is $\frac{2500}{2500+2500} = 0.5$.
- Constant is $\frac{1}{(0.5*11)} = 0.18$.

Question 5

(a) Find the Laplace approximation for the following integrals.

- (i)

$$\begin{aligned}
f(x) &= \frac{1}{2^{\frac{k}{2}}\Gamma(\frac{k}{2})}x^{\frac{k}{2}-1}exp(-\frac{x}{2}) \\
&= exp\left\{(\frac{k}{2}-1)log(x) - \frac{k}{2}log(2) - log(\Gamma(\frac{k}{2})) - \frac{x}{2}\right\} \\
h(x) &= (\frac{k}{2}-1)log(x) - \frac{k}{2}log(2) - log(\Gamma(\frac{k}{2})) - \frac{x}{2} \\
h^{''}(x) &= \frac{k-2}{2x^2}
\end{aligned}$$

$$\begin{aligned}
&\int_a^b \frac{1}{2^{\frac{k}{2}}\Gamma(\frac{k}{2})}x^{\frac{k}{2}-1}exp(-\frac{x}{2})dx \\
by\ Laplace\ approximation\ we\ get: \propto &\frac{\hat{x}^{\frac{k}{2}-1}exp(-\frac{\hat{x}}{2})}{2^{\frac{k}{2}}\Gamma(\frac{k}{2})}\int_a^b exp\left\{\frac{(x-\hat{x})^2}{2}(\frac{k-2}{2\hat{x}^2})\right\}dx \\
&\hat{x} = max(k-2, 0)
\end{aligned}$$

- (ii)

$$
\begin{aligned}
f(x) &= x^{\alpha-1}(1-x)^{\beta-1},\ for\ \alpha,\beta>1 \\
&= exp\Big\{(\alpha-1)log(x)+(\beta-1)log(1-x)\Big\} \\
h(x) &= (\alpha-1)log(x)+(\beta-1)log(1-x) \\
h^{''}(x) &= -\frac{\alpha-1}{x^2}-\frac{\beta-1}{(1-x)^2}
\end{aligned}
$$

$$
\begin{aligned}
&\int_a^b x^{\alpha-1}(1-x)^{\beta-1}dx \\
by\ Laplace\ approximation\ we\ get: &\propto \hat{x}^{\alpha-1}(1-\hat{x})^{\beta-1}\int_a^b exp\Big\{\frac{(x-\hat{x})^2}{2}(-\frac{\alpha-1}{\hat{x}^2}-\frac{\beta-1}{(1-\hat{x})^2})\Big\}dx \\
&\propto \frac{(\alpha-1)^{\alpha-1}(\beta-1)^{\beta-1}}{(\alpha+\beta-2)^{\alpha+\beta-2}}\int_a^b exp\Big\{\frac{(x-\hat{x})^2}{2}\Big(\frac{-(\alpha+\beta-2)^2}{(\alpha-1)}+\frac{-(\alpha+\beta-2)^2}{(\beta-1)}\Big)\Big\}dx \\
\hat{x} &= \frac{\alpha-1}{\alpha+\beta-2},\ for\ \alpha,\beta>1
\end{aligned}
$$

(b) Use the Laplace approximation to find the appropriate normalizing constant for $(\alpha = 2, \beta = 2)$ and $(\alpha = 1, \beta = 2)$, and compare with the exact evaluation of such constant.

- When $(\alpha = 2, \beta = 2)$.

$$
\begin{aligned}
\hat{x} &= \frac{1}{2} \\
plug\ \ \hat{x}\ \ and\ \ \alpha,\beta\ \ in &= \frac{(2-1)^{2-1}(2-1)^{2-1}}{(2+2-2)^{2+2-2}}\int_0^1 exp\Big\{\frac{(x-\frac{1}{2})^2}{2}\Big(\frac{-(2+2-2)^2}{(2-1)}+\frac{-(2+2-2)^2}{(2-1)}\Big)\Big\}dx \\
&= 0.25\int_0^1 exp(-4(x-0.5)^2)dx
\end{aligned}
$$

```
bx = function (x) {
  0.25 * exp(-4 * (x - 0.5)^2)
}
b1 = integrate(bx, 0, 1)
b1
```

```
## 0.186706 with absolute error < 2.1e-15
```

```
b2 = beta(2, 2)
b2
```

```
## [1] 0.1666667
```

```
##very close
```

- When $(\alpha = 1, \beta = 2)$

$$
\begin{aligned}
\hat{x} &= 0 \\
plug\ \ \hat{x}\ \ and\ \ \alpha,\beta\ \ in &= \frac{(1-1)^{1-1}(2-1)^{2-1}}{(1+2-2)^{1+2-2}} \int_0^1 exp\left\{\frac{(x-0)^2}{2}\left(\frac{-(1+2-2)^2}{(1-1)} + \frac{-(1+2-2)^2}{(2-1)}\right)\right\}dx \\
&= (1-x)\int_0^1 exp(\frac{x^2}{2(1-x)^2})dx
\end{aligned}
$$

- α here is 1 and not satisfy the condition that $\alpha > 1$. So when $(\alpha = 1, \beta = 2)$ this function is not proper.

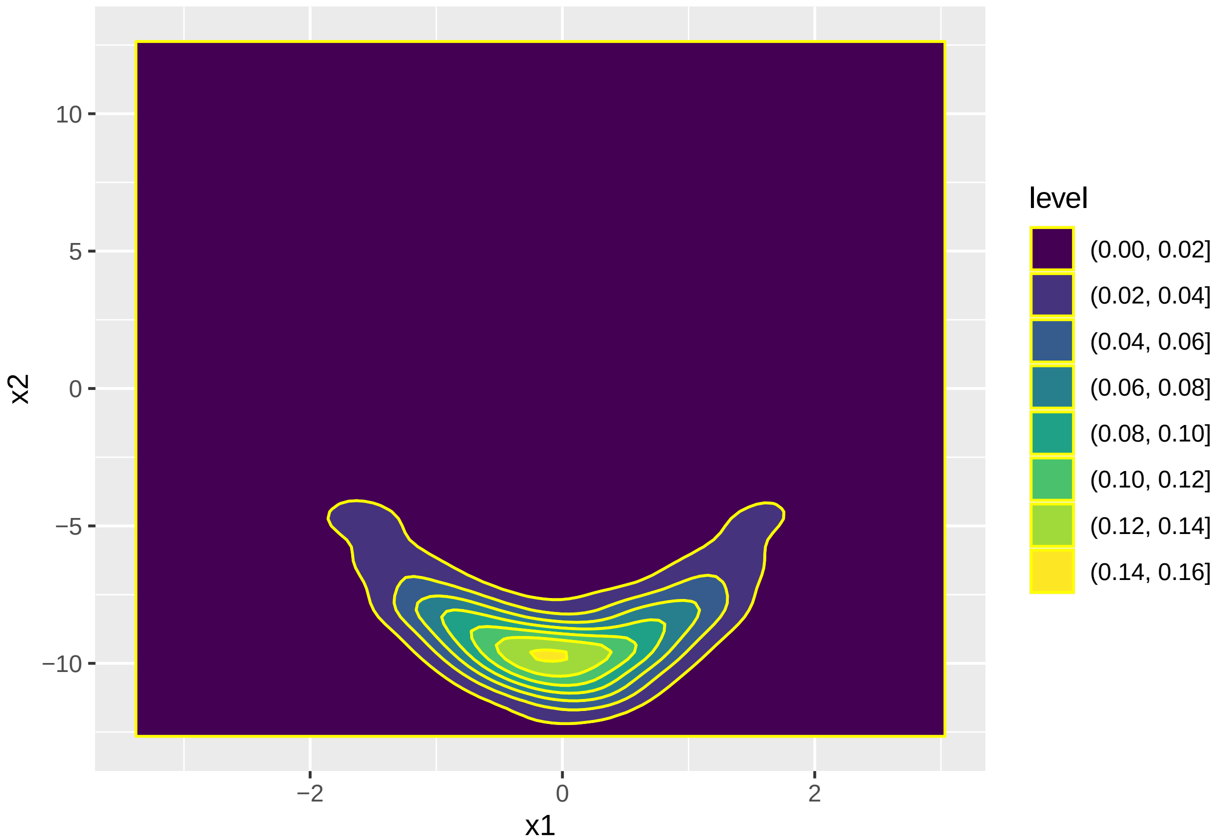
Question 6

(a) Describe how to simulate directly from $\pi(x_1, x_2)$.

First by derive the marginal distribution of x_1 and conditional distribution of x_2.

$$
\begin{aligned}
\pi(x_1, x_2) &\propto exp\left\{-\frac{x_1^2}{2}\right\}exp\left[-\frac{\{x_2 - 2(x_1^2 - 5)\}^2}{2}\right] \\
f(x_1) &= exp\left\{-\frac{x_1^2}{2}\right\}\int exp\left[-\frac{\{x_2 - 2(x_1^2 - 5)\}^2}{2}\right]dx_2 \\
1.\quad f(x_1) &\propto exp\left\{-\frac{x_1^2}{2}\right\},\ \ \pi(x_1) \sim N(0,1) \\
2.\quad f(x_2|x_1) &\propto exp\left[-\frac{\{x_2 - 2(x_1^2 - 5)\}^2}{2}\right],\ \ \pi(x_2|x_1) \sim N(2(x_1^2 - 5), 1)
\end{aligned}
$$

```
library(ggplot2)
set.seed(44) # set the random seed for reproducibility
x1 = rnorm(1000, 0, 1)
x2 = rnorm(1000, 2 * (x1^2 - 5), 1)
banana = as.data.frame(exp(-x1^2 / 2) * exp(-(x2 - 2 * (x1^2 - 5.0))^2 / 2))
bananadd <- as.data.frame(cbind(x1, x2))
ggplot(banana, aes(x = x1, y = x2)) + geom_density_2d_filled(color = "yellow")
```




```r
head(bananadd)
```

```
##             x1        x2
## 1  0.65391826 -9.829455
## 2  0.01905227 -9.986141
## 3 -1.84950398 -3.342322
## 4 -0.13276331 -9.876034
## 5 -1.19881818 -5.191604
## 6 -1.32974147 -6.380804
```

(b) Describe how to simulate from $\pi(x_1, x_2)$ using the accept-reject method. Implement your algorithm and discuss sampling efficiency in relation to the chosen instrument distribution.

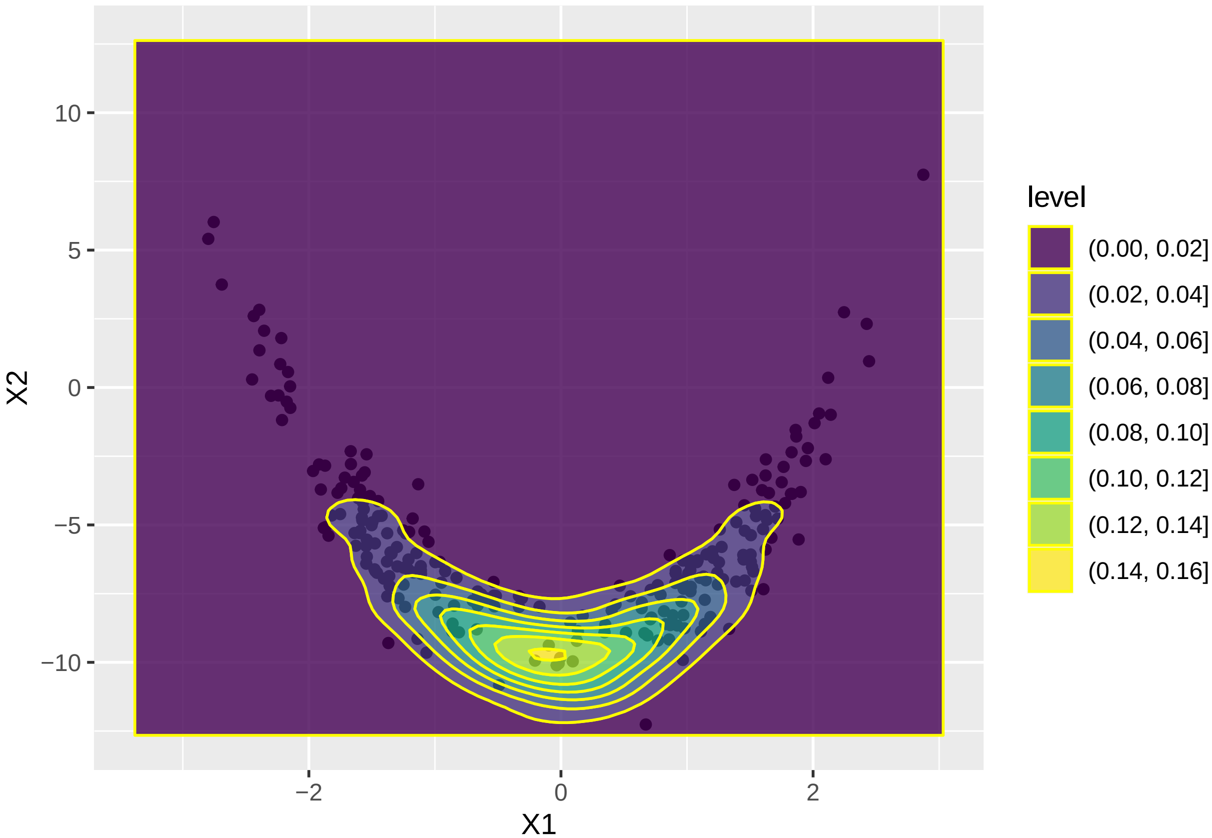
```r
ba = function(x1, x2) {
  exp(-x1^2 / 2) * exp(-(x2 - 2 * (x1^2 - 5.0))^2 / 2)
}

ba_ar = function(n_iter, sig, m) {

  ## step 1, initialize
  mvn = rmvnorm(n = n_iter, sigma = sig)
  n_iter = nrow(mvn)
```

```r
  ba_now = dmvnorm(mvn, sigma = sig)
  ba_cand = rep(NA, n_iter)

  ## step 2, iterate
  for (i in 1:n_iter) {
    ## step 2a
    x1_cand = mvn[i, 1] # draw a candidate
    x2_cand = mvn[i, 2] # draw a candidate

    ## step 2b
    ba_cand[i] = ba(x1 = x1_cand, x2 = x2_cand) # evaluate with the candidate
    c = (ba_cand / ba_now) * (1 / m) # ratio

    ## step 2c
    u = runif(n_iter)
    # draw a uniform variable which will be less than c with probability
    x_ar = mvn[c >= u, ]
  }

  ## return a list of output
  list(x_ar)
}
```

```r
set.seed(4355)  # set the random seed for reproducibility
sig = matrix(c(1.5, 0, 0, 10), 2)
x_ar = ba_ar(n_iter = 10000, sig = sig, m = 100)
post = as.data.frame(x_ar)
head(post)
```

```
##           X1        X2
## 1  1.1408968 -7.725557
## 2  0.7901188 -7.541967
## 3 -1.9049951 -3.710869
## 4  0.9614242 -7.019218
## 5 -1.7724500 -3.830764
## 6 -0.9186176 -6.684883
```

```r
ggplot() +
  geom_point(data = post, mapping = aes(x = X1, y = X2)) +
  geom_density_2d_filled(mapping = aes(x = x1, y = x2),
                         data = banana, alpha = 0.8, color = "yellow")
```



Sampling Efficiency

```
nrow(post) / 10000
```

```
## [1] 0.0239
```

- From the plot above, the points simulated from Aaccept-reject method. The Sampling Efficiency is 2.39%. If using distribution which close to Banana distribution might get a better result.

(c) Describe how to simulate from $\pi(x_1, x_2)$ using sampling importance resampling. Implement your algorithm and discuss sampling efficiency in relation to the chosen importance distribution.

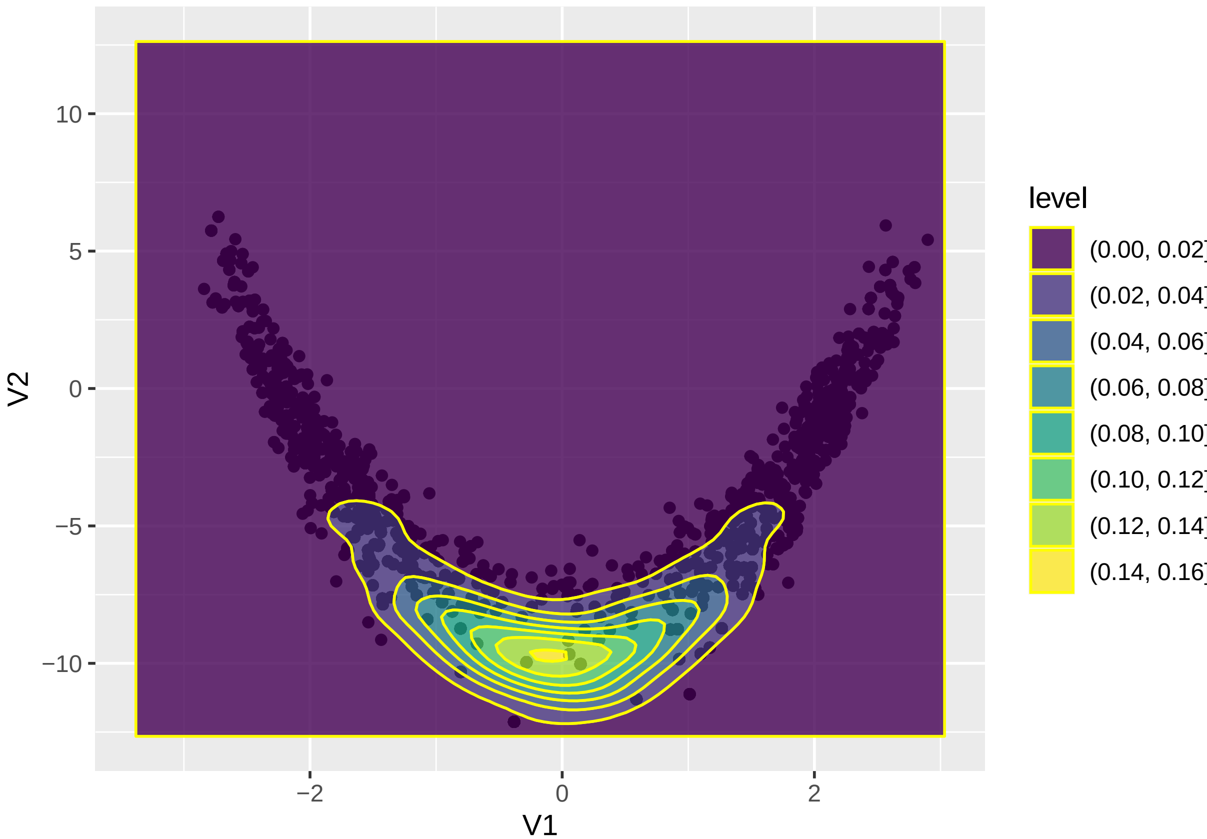
```
## step 1, initialize
mu1 = c(0, 1)
mu2 = c(0, -1)
sig1 = matrix(c(1.5, 0, 0, 10), 2)
sig2 = matrix(c(1.5, 0, 0, 10), 2)
n_iter = 10000
mvn1 = rbind(rmvnorm(n_iter / 2, mu1, sig1), rmvnorm(n_iter / 2, mu2, sig2))
mvn1 = rmvnorm(n = n_iter, sigma = sig)
ba_now1 = rep(NA, n_iter)
ba_cand1 = rep(NA, n_iter)
```

```r
## step 2, iterate
for(i in 1:n_iter) {
  ba_now1[i] = 0.5 * dmvnorm(mvn1[i,], mu1, sig1) +
    0.5 * dmvnorm(mvn1[i,], mu2, sig2)
  ba_cand1[i] = exp(-mvn1[i,1]^2 / 2) * exp(-(mvn1[i, 2] -
                                              2 * (mvn1[i, 1]^2 - 5.0))^2 / 2)
}
```

```r
set.seed(4355)
w = ba_cand1 / ba_now1
xx = seq(1, n_iter)
x_is  = mvn1[sample(xx, prob = w, replace = T), ]
post1 = as.data.frame(x_is)
head(post1)
```

```
##            V1         V2
## 1  0.1456008 -10.025856
## 2 -2.2838976   1.060188
## 3 -2.2168291  -1.536293
## 4  0.6853384  -7.522846
## 5  1.0346231  -7.316371
## 6 -0.4762098  -7.776759
```

```r
ggplot() +
  geom_point(data = post1, mapping = aes(x = V1, y = V2)) +
  geom_density_2d_filled(mapping = aes(x = x1, y = x2),
                         data = banana, alpha = 0.8, color = "yellow")
```



Sampling Efficiency

```
nrow(unique(post1)) / 10000
```

```
## [1] 0.0831
```

- From above results, for IS, Sampling Efficiency is 7.68%. Also, from both Sampling Efficiency and plot shows that IS is better than AR sampling. Resampling will yield a better results.

(d) Use the algorithms in (a, b, c) to estimate the following.

- $1.E(x_1^2)$

```
x1_2 = list(ba = bananadd$x1^2,
            ar = post$X1^2,
            ir = post1$V1^2)

means = lapply(x1_2, mean)
vars = lapply(x1_2, var)
x1_2 = cbind(means, vars)
x1_2
```

```
##    means     vars
## ba 0.9990152 2.005538
## ar 2.039095  2.51557
## ir 1.276256  2.07418
```

- For AR algorithm shows higher mean value, and other two's mean values are very close.
- $2.E(x_2^2)$.

```
x2 = list(ba = bananadd$x2,
          ar = post$X2,
          ir = post1$V2)

means = lapply(x2, mean)
vars = lapply(x2, var)
x2 = cbind(means, vars)
x2
```

```
##    means     vars
## ba -7.997502 9.180658
## ar -5.522164 9.813363
## ir -7.336305 9.09824
```

- All methods provided similar estimates for the mean of X_2^2.
- $3.P(x_1 + x_2 > 0)$.

```
p = list(ba = as.numeric(bananadd$x1 + bananadd$x2 > 0),
         ar = as.numeric(post$X1 + post$X2 > 0),
         ir = as.numeric(post1$V1 + post1$V2 > 0))
means = lapply(p, mean)
vars = lapply(p, var)
p = cbind(means, vars)
p
```

```
##    means      vars
## ba 0.031      0.03006907
## ar 0.06276151 0.05906965
## ir 0.041      0.03932293
```

- All methods provided similar estimates for the mean when $P(x_1 + x_2 > 0)$.