Zhuoqun Huang (Calvin) 908525
Tutorial information: Anubhav Kaphle, Thur 9-10

October 13, 2019

MAS Assignment 3 Submission

revision 3

1 Question1

(a)
```
library(MASS)
data(quine)
## (a)
mle_p <- function(data, r=1.5) {
  return(
    sum(data) / (length(data) * r + sum(data))
  )
}
(mle_p(quine$Days))

## [1] 0.916476
```

According to Theorem 1,

$$\frac{\sum_{i=1}^{n} x_i}{nr + \sum_{i=1}^{n} x_i} = 0.916476$$

(b)
```
## (b)
bayesian_p.basic <- function(data, r=1.5, a=0.5, b=0.5) {
  return(
    c(sum(data)+a, length(data) * r+b)
  )
}
(alp = (bayesian_p.basic(quine$Days))[1])

## [1] 2403.5

(bet = (bayesian_p.basic(quine$Days))[2])

## [1] 219.5

(e_bayesian = alp / (alp + bet))

## [1] 0.9163172
```

According to Theorem 2,

$$p|X \sim beta(\sum_{i=1}^{n} x_i + a, nr + b) = beta(2403.5, 219.5)$$

with expected value of

$$E(p|X) = \frac{\sum_{i=1}^{n} x_i + a}{\sum_{i=1}^{n} x_i + a + nr + b} = 0.9163172 < 0.916476$$

(c)

$$
\begin{aligned}
p|X_{mle} &= \frac{\sum_{i=1}^{n} x_i}{\sum_{i=1}^{n} x_i + nr} \\
E(p|X) &= \frac{\sum_{i=1}^{n} x_i + a}{\sum_{i=1}^{n} x_i + nr + a + b} \\
p|X_{mle} - E(p|X) &= \frac{(\sum_{i=1}^{n} x_i)(\sum_{i=1}^{n} x_i + nr + a + b) - (\sum_{i=1}^{n} x_i + a)(\sum_{i=1}^{n} x_i + nr)}{(\sum_{i=1}^{n} x_i + nr)(\sum_{i=1}^{n} x_i + nr + a + b)} \\
&= \frac{b\sum_{i=1}^{n} x_i - anr}{(\sum_{i=1}^{n} x_i + nr)(\sum_{i=1}^{n} x_i + nr + a + b)} = \Delta
\end{aligned}
$$

$$
\begin{aligned}
&\because (\sum_{i=1}^{n} x_i + nr)(\sum_{i=1}^{n} x_i + nr + a + b) > 0 \\
&\therefore sign(\Delta) = sign(b\sum_{i=1}^{n} x_i - anr) \\
&\because \text{let } a \to \infty, sign(b\sum_{i=1}^{n} x_i - anr) < 0
\end{aligned}
$$

Thus from the above derivation, it will NOT always be true. When a is very large compared to b or r gets very large, the estimate can be larger instead. Intuitively, the prior of MLE has a mean $\frac{\alpha}{\alpha+\beta} = 0.5 < p|X_{mle}$, and the combined distribution $p|X$ will be smaller than original $p|X_{mle}$.

(d) $\because \ E(r) = 1.5$
$\therefore \frac{1}{\lambda} = 1.5 \Rightarrow \lambda = \frac{2}{3}$
According to Theorem 3,

$$
\begin{aligned}
f(r|y) &\propto [\prod_{i=1}^{n} \frac{\Gamma(y_i + r)}{\Gamma(r)}]\beta(\sum_{i=1}^{n} y_i + a, nr + b)e^{-\lambda r} \qquad (\because a = b = \tfrac{1}{2}) \\
&= [\prod_{i=1}^{n} \frac{\Gamma(y_i + r)}{\Gamma(r)}]\beta(\sum_{i=1}^{n} y_i + \frac{1}{2}, nr + \frac{1}{2})e^{-\frac{2}{3}r}
\end{aligned}
$$

2 Question2

(a) This is true, according to Theorem 4.

(b) This is true, according to Theorem 5.

(c) The Algorithm can be formulated as follows:

Algorithm 1 Random Gamma Distribution Generation

$randomNormal(\mu, \sigma)$ ▷ Generate data from normal distribution with mean of μ, variance of σ^2

$randomUnif(lo, hi)$ ▷ Generate data from uniform distribution with lower bound lo and upper bound hi

Require: α ▷ Shape parameter of gamma distribution

Require: λ ▷ Rate parameter of gamma distribution

procedure RGAMMA(α, λ)
  $d \leftarrow \alpha - 1/3$
  $c \leftarrow \frac{1}{\sqrt{9d}}$
  $u(x) \leftarrow \lambda(x) \Rightarrow exp(-\frac{x^2}{2})$
  $h(x) \leftarrow \lambda(x) \Rightarrow d(1 + cx)^3$
  $h'(x) \leftarrow \lambda(x) \Rightarrow 3cd(1 + cx)^2$
  $g(x) \leftarrow \lambda(x) \Rightarrow d \cdot log((1 + cx)^3) - d(1 + cx)^3 + d$
  $h^*(x) \leftarrow \lambda(x) \Rightarrow (h(x)^3)^{\alpha-1}e^{-h(x)}h'(x)$
  $ratio \leftarrow \frac{h^*(1)}{exp(g(1))}$ ▷ Need to compute a ratio to multiply $u(x)$ by to guarantee $ratio \cdot u(x) > h^*(x)$
  $x = randomNormal(0, 1)$
  $y = randomUnif(0, 1)$
  while $x < -\frac{1}{c}$ **OR** $y > \frac{h^*(x)}{ratio \cdot u(x)}$ **do**
    $x = randomNormal(0, 1)$
    $y = randomUnif(0, 1)$
  end while
  return $\frac{h(x)}{\lambda}$
end procedure

All operations $O(1)$ unless explicitly stated, we denote the complexity of search for problem also here.

The algorithm overall will have a complexity close to O(1)

Correctness of this algorithm is shown in Figure 1

The corresponding R code is as follows

```
## (c)
h.me <- function(x, c, d) {
  return(d * (1 + c * x)^3)
}

g.me <- function(x, c, d) {
  return(
    d * log((1 + c * x) ^ 3) - d * (1 + c * x)^3 + d
  );
}


f.star.me <- function(x, alpha) {
  d = alpha - 1/3;
  c = 1 / sqrt(9 * d);
  h.prime.me <- function(x) {
    return(3 * d * c * (1 + c * x) ^ 2)
  }
  return(
    h.me(x, c, d) ^ (alpha - 1) * exp(-h.me(x, c, d))* h.prime.me(x)
  );
}

yang.ge.star.me <- function(x, alpha) {
  d = alpha - 1/3;
  c = 1 / sqrt(9 * d);
  return(
    exp(
      g.me(x, c, d)
    )
  )
}

ratio.comp <- function(alpha) {
  return(f.star.me(1, alpha) / yang.ge.star.me(1, alpha))
}

h.star.me <- function(x) {
  return(exp(-x^2/2));
}

rgamma.me_ <- function(alpha=1, beta=1) {
  ratio = ratio.comp(alpha)
  d = alpha - 1/3;
  c = 1 / sqrt(9 * d);
  while (TRUE) {
    y = runif(1);
    x = rnorm(1);
```

```
    if (x > -1/c && y < f.star.me(x, alpha) / (h.star.me(x) * ratio)) {
      break;
    }
  }
  return(h.me(x, c, d)/beta);
}

rgamma.me <- function(n, alpha=1, beta=1) {
  return(replicate(n, rgamma.me_(alpha, beta)))
}


set.seed(6); # ;) Not the truth to the universe
plot(qgamma(1:1000/1001, 1.2, 3), sort(rgamma.me(1000, 1.2, 3)),
     xlab="Gamma Quantile", ylab="gamma.me Quantile", cex.lab=0.7)
abline(0, 1, col="red")
```

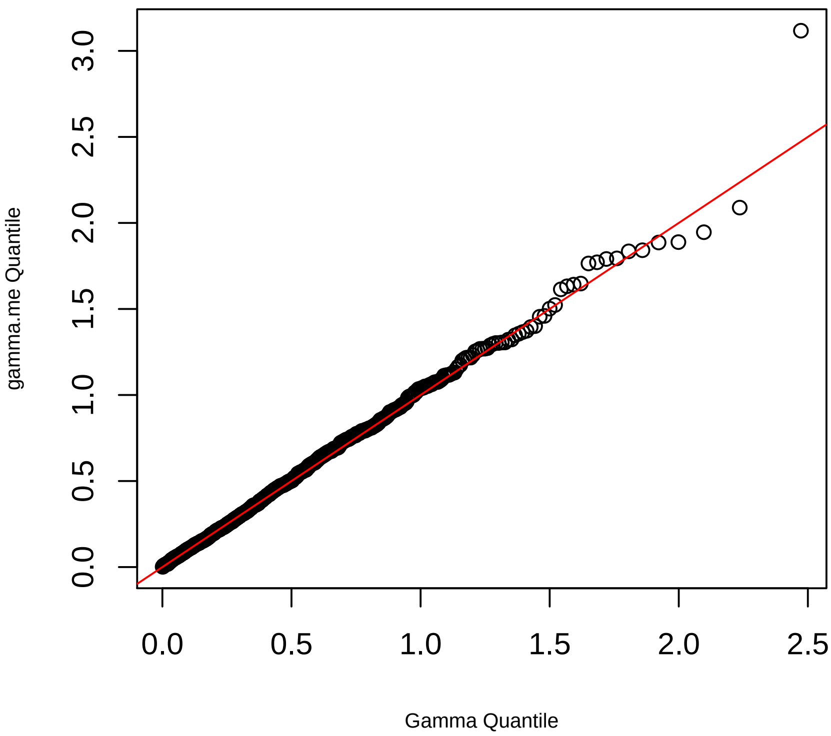

Figure 1: QQ plot

Appendix

Theorem 1. *The Maximum Likelihood Estimator for parameter p of Negative Binomial Distribution $NB(p,r)$ given r and a set of data $X = \{x_i; i \leq n\}$ is*

$$\hat{p} = \frac{\sum_{i=1}^{n} x_i}{nr + \sum_{i=1}^{n} x_i} \tag{1}$$

Proof of Theorem 1.

$$L = \prod_{i=1}^{n} f_X(x_i) = \prod_{i=1}^{n} \frac{\Gamma(x_i + r)}{\Gamma(x_i + 1)\Gamma(r)}(1-p)^r p^{x_i} \tag{*}$$

$$\begin{aligned}
l = log(L) &= \sum_{i=1}^{n} log(\frac{\Gamma(x_i + r)}{\Gamma(x_i + 1)\Gamma(r)}(1-p)^r p^{x_i}) \\
&= \sum_{i=1}^{n} log(\Gamma(x_i + r)) - log(\Gamma(x_i + 1)) - log(\Gamma(r)) + rlog(1-p) + x_i log(p) \\
\frac{\partial l}{\partial p} &= \frac{\partial}{\partial p}(\sum_{i=1}^{n} log(\Gamma(x_i + r)) - log(\Gamma(x_i + 1)) - log(\Gamma(r)) + rlog(1-p) + x_i log(p)) \\
&= \frac{\partial}{\partial p}(\sum_{i=1}^{n}(c + rlog(1-p) + x_i log(p)) \\
&= \sum_{i=1}^{n}(\frac{-r}{1-p} + \frac{x_i}{p}) \\
\frac{\partial l}{\partial p} &= 0 \Rightarrow p = \frac{\sum_{i=1}^{n} x_i}{nr + \sum_{i=1}^{n} x_i}
\end{aligned}$$

□

Theorem 2. *The Bayesian Posterior Distribution for parameter $(p|X)$ of Negative Binomial Distribution $NB(p,r)$ given r and prior $p \sim beta(a,b)$ and a set of data $X = \{x_i; i \leq n\}$ is*

$$p|X \sim beta(\sum_{i=1}^{n} x_i + a, nr + b) \tag{2}$$

with a mean value of

$$E(p|X) = \frac{\sum_{i=1}^{n} x_i + a}{\sum_{i=1}^{n} x_i + a + nr + b} \tag{3}$$

Proof of Theorem 2.

$$\begin{aligned}
p &\sim beta(a,b) \\
f(p) &\propto p^{a-1}(1-p)^{b-1} \\
f(p|X) &\propto f(X|p)f(p) \\
&= p^{a-1}(1-p)^{b-1} \prod_{i=1}^{n} \frac{\Gamma(x_i + r)}{\Gamma(x_i + 1)\Gamma(r)}(1-p)^r p^{x_i} \\
&\propto p^{a+\sum_{i=1}^{n} x_i - 1}(1-p)^{b+nr-1}
\end{aligned}$$

This is the same form as a beta distribution

$$
\begin{aligned}
p|X &\sim beta(\sum_{i=1}^{n} x_i + a, nr + b) \\
\Rightarrow E(p|X) &= \frac{\sum_{i=1}^{n} x_i + a}{\sum_{i=1}^{n} x_i + a + nr + b}
\end{aligned}
$$

□

Theorem 3. *The Bayesian Posterior Distribution for parameter $(r|X)$ of Negative Binomial Distribution $NB(p, r)$ given prior for $r \sim Exp(\lambda)$ and prior for $p \sim beta(a, b)$ and a set of data $X = \{x_i; i \leq n\}$ has the form*

$$
f(r|X) \propto [\prod_{i=1}^{n} \frac{\Gamma(x_i + r)}{\Gamma(r)}]\beta(\sum_{i=1}^{n} x_i + a, nr + b)e^{-\lambda r} \tag{4}
$$

Proof of Theorem 3.

$$
\begin{aligned}
p &\sim beta(a, b) \\
f(p) &\propto p^{a-1}(1-p)^{b-1} \\
r &\sim Exp(\lambda) \\
f(r) &\propto e^{-\lambda r} \\
f(r|X) &\propto f(X|r)f(r) \\
&\propto \int_0^1 f(X, p|r)f(r)dp \qquad \text{(Law of total Probability)} \\
&= \int_0^1 f(X|r, p)f(p)f(r)dp \\
&\propto [\prod_{i=1}^{n} \frac{\Gamma(x_i + r)}{\Gamma(x_i + 1)\Gamma(r)}(1-p)^r p^{x_i}]p^{a-1}(1-p)^{b-1}e^{-\lambda r} \\
&\propto \int_0^1 [\prod_{i=1}^{n} \frac{\Gamma(x_i + r)}{\Gamma(r)}]p^{a+\sum_{i=1}^{n} x_i - 1}(1-p)^{b+nr-1}e^{-\lambda r}dp
\end{aligned}
$$

Consider the following random variable Z from beta distribution

$$
\begin{aligned}
Z &\sim beta(\sum_{i=1}^{n} x_i + a, nr + b) \\
\int_0^1 f_Z(z) &= \int_0^1 \frac{p^{a+\sum_{i=1}^{n} x_i - 1}(1-p)^{b+nr-1}}{\beta(\sum_{i=1}^{n} x_i + a, nr + b)} = 1 \\
&\Rightarrow \int_0^1 p^{a+\sum_{i=1}^{n} x_i - 1}(1-p)^{b+nr-1} = \beta(\sum_{i=1}^{n} x_i + a, nr + b) \\
f(r|X) &\propto \int_0^1 [\prod_{i=1}^{n} \frac{\Gamma(x_i + r)}{\Gamma(r)}]p^{a+\sum_{i=1}^{n} x_i - 1}(1-p)^{b+nr-1}e^{-\lambda r}dp \\
&= [\prod_{i=1}^{n} \frac{\Gamma(x_i + r)}{\Gamma(r)}]\beta(\sum_{i=1}^{n} x_i + a, nr + b)e^{-\lambda r}
\end{aligned}
$$

□

Theorem 4. *Given $X \sim gamma(\alpha, 1)$ then*

$$\frac{X}{\lambda} \sim gamma(\alpha, \lambda) \tag{5}$$

Proof of Theorem 4.

$$\begin{aligned}
&\text{For } Z \sim gamma(a, b) \\
&\quad F_Z(z) = \frac{\gamma(a, bx)}{\Gamma(a)} && \text{(Source: WikiPedia)} \\
&\quad F_X(x) = \frac{\gamma(\alpha, x)}{\Gamma(\alpha)} \\
&\quad F_Y(y) = F(Y < y) = F(\frac{X}{\lambda} < y) \\
&\quad\quad = F(X < \lambda y) = \frac{\gamma(\alpha, \lambda y)}{\Gamma(\alpha)} \\
&\quad \Rightarrow Y \sim gamma(\alpha, \lambda) && \text{(Based on the previous CDF)}
\end{aligned}$$

□

Theorem 5. *Given $f_X(x) = \frac{h(x)^{\alpha-1}e^{-h(x)}h'(x)}{\Gamma(\alpha)}$ then*

$$Y = h(X) \sim gamma(\alpha, 1) \tag{6}$$

Proof of Theorem 5.

$$\begin{aligned}
f_X(x) &= \frac{h(x)^{\alpha-1}e^{-h(x)}h'(x)}{\Gamma(\alpha)} \\
F_X(x) &= \int_0^x f_x(t)dt \\
F_Y(y) &= F(Y < y) = F(h(X) \leq y) \\
&= F(X < h^{-1}(y)) \\
&(\because h(x) \text{ is monotonically increasing} \Rightarrow \exists h^{-1}(x) \text{ such that } h^{-1}(h(x)) = x) \\
&= \int_0^{h^{-1}(y)} f_X(x)dx \\
f_Y(y) &= \frac{d}{dy}F_y(y) = \frac{d}{dy}\int_0^{h^{-1}(y)} f_X(x)dx \\
&= f_x(h^{-1}(y))\frac{d}{dy}h^{-1}(y) - f_X(0) \cdot \frac{d}{dy}0 + 0 && \text{(Leibniz Rule)} \\
&= \frac{y^{\alpha-1}e^{-y}h'(h^{-1}(y))}{\Gamma(\alpha)}\frac{d}{dy}h^{-1}(y) \\
&(\because \tfrac{d}{dy}h(h^{-1}(y) = \tfrac{d}{dy}y = 1) \\
&(\therefore h'(h^{-1}(y)) \cdot \tfrac{d}{dy}h^{-1}(y) = 1)
\end{aligned}$$

$$
\begin{aligned}
f_Y(y) &= \frac{y^{\alpha-1}e^{-y}}{\Gamma(\alpha)}(h'(h^{-1}(y)) \cdot \frac{d}{dy}h^{-1}(y)) \\
&= \frac{y^{\alpha-1}e^{-y}}{\Gamma(\alpha)} \\
Y &\sim gamma(\alpha, 1)
\end{aligned}
$$

($\because$ same form of pdf as gamma distribution and pdf is always unique.)

$\square$