GR5241 HW1

JIE LI

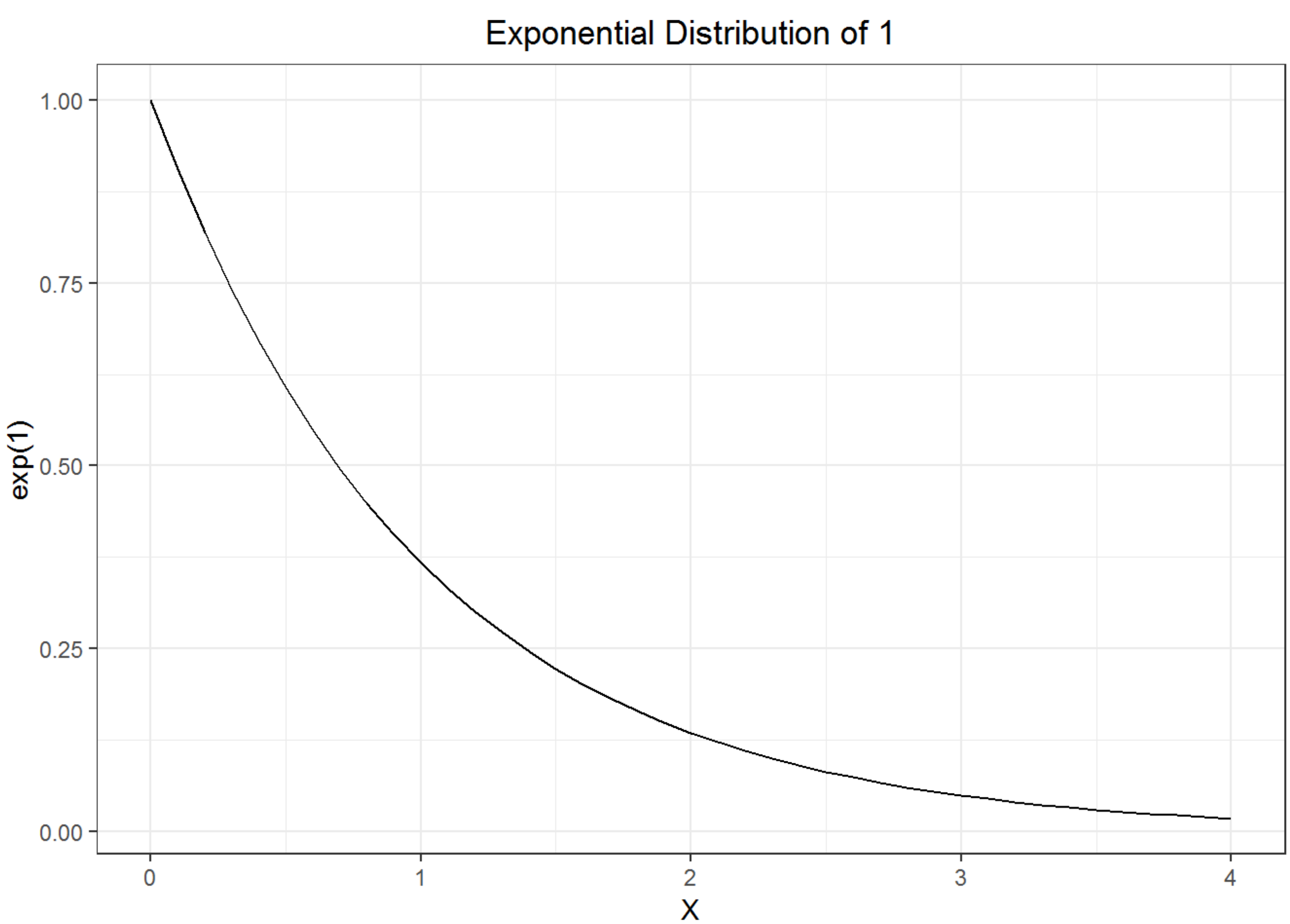
January 28, 2019

Problem 1 (Bayesian inference and online learning)

1

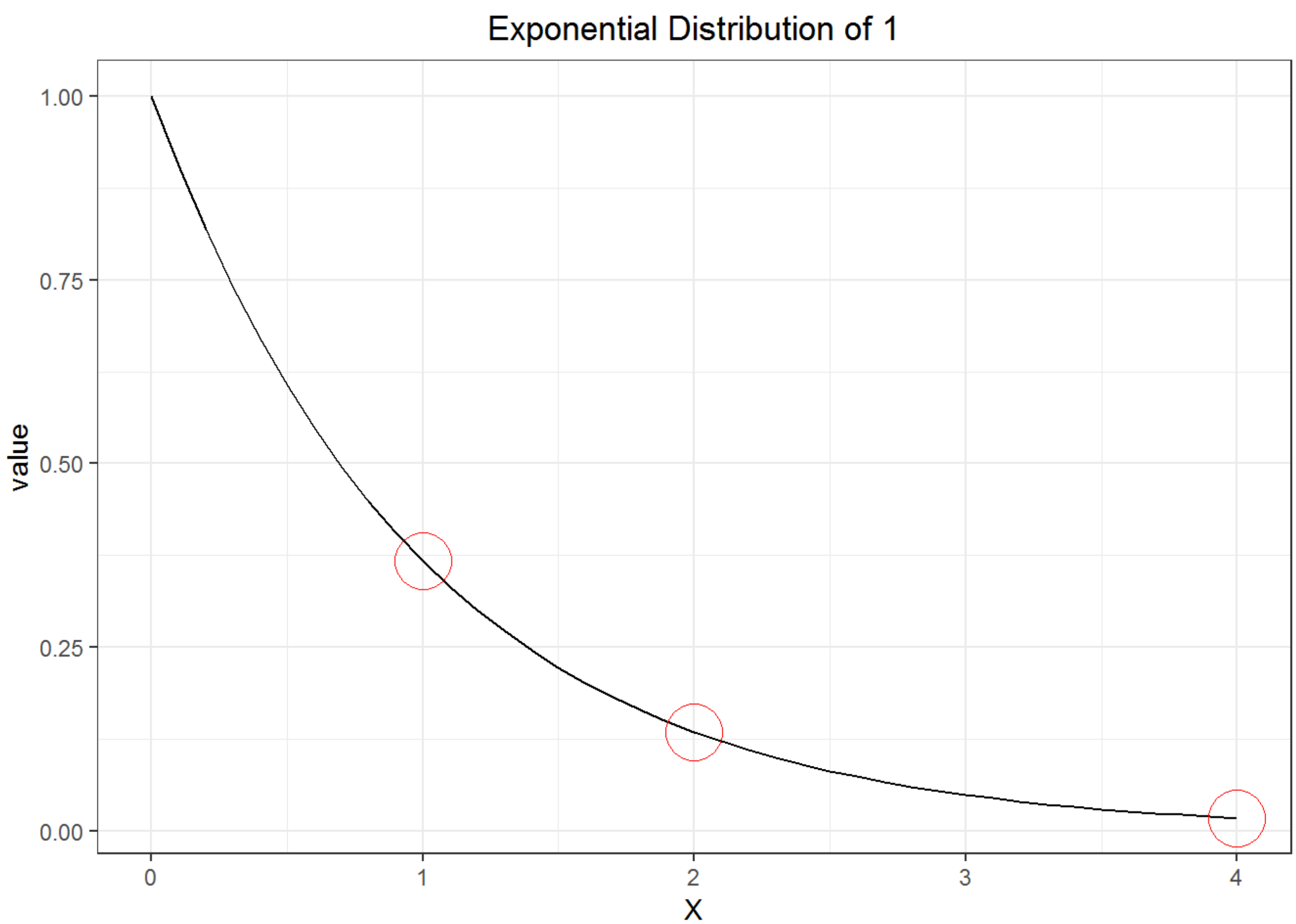
```
library(ggplot2)
x = seq(0,4,0.1)
y = dexp(x,1)

ggplot()+
  geom_line(aes(x,y))+
  theme_bw()+
  labs(title = "Exponential Distribution of 1", x = "X", y = "exp(1)")+
  theme(plot.title = element_text(hjust = 0.5))
```



2

```
ggplot()+
  geom_line(aes(x,y))+
  geom_point(aes(1, dexp(1,1)), color = "red", shape = 1, size = 10)+
  geom_point(aes(2, dexp(2,1)), color = "red", shape = 1, size = 10)+
  geom_point(aes(4, dexp(4,1)), color = "red", shape = 1, size = 10)+
  theme_bw()+
  labs(title = "Exponential Distribution of 1", x = "X", y = "value")+
  theme(plot.title = element_text(hjust = 0.5))
```

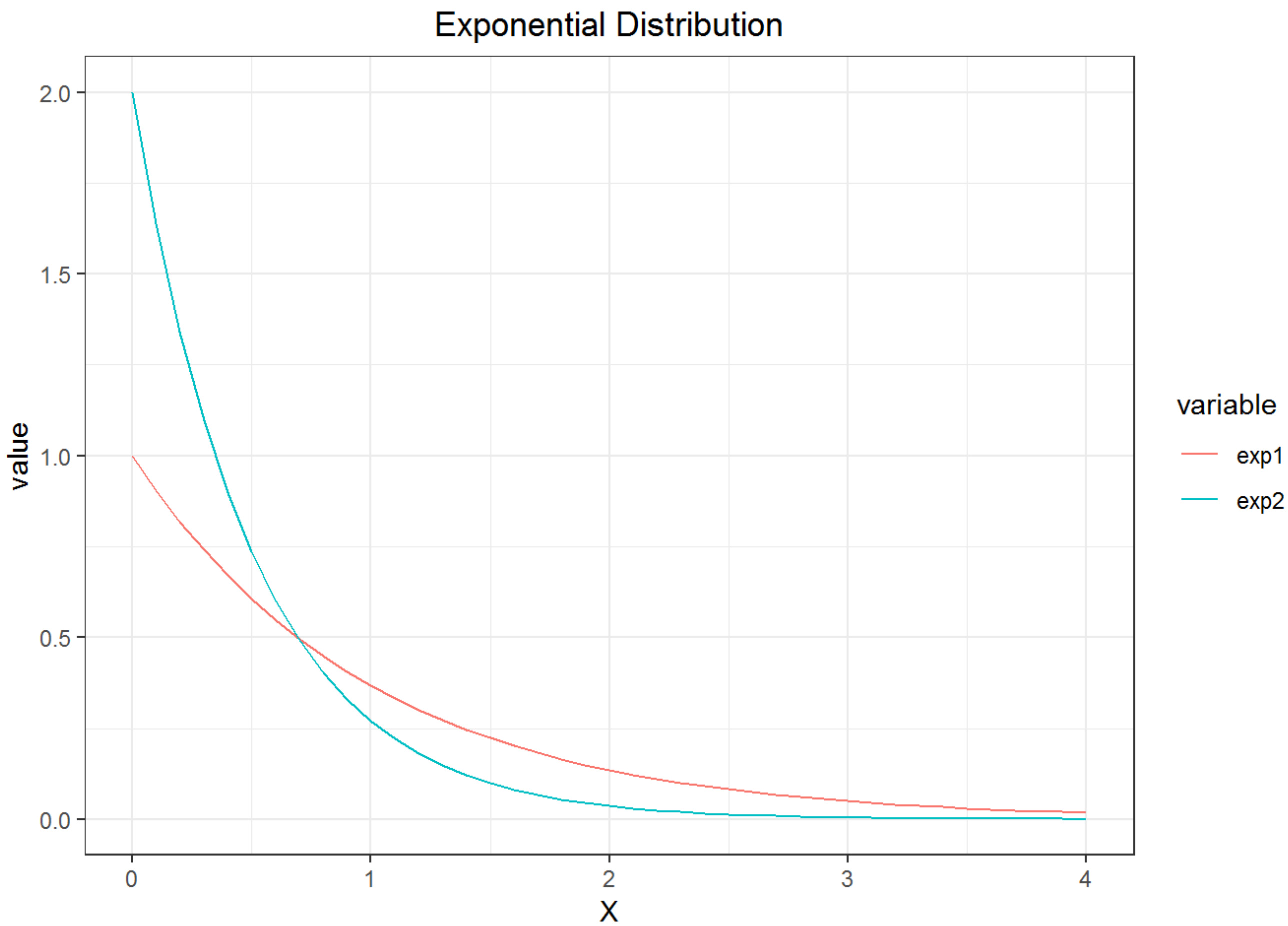


3

```
library(reshape2)
library(purrr)
df = data.frame(exp1 = y, exp2 = dexp(seq(0,4,0.1),2)) %>% melt() %>% cbind(x=rep(seq(0,4,0.1),2), .)
```

```
## No id variables; using all as measure variables
```

```
ggplot(df)+
  geom_line(aes(x, value, color = variable))+
  theme_bw()+
  labs(title = "Exponential Distribution", x = "X")+
  theme(plot.title = element_text(hjust = 0.5))
```



When x = 0,1, exp(2) is higher, but rest of data is lower.

Question 1

$$x_1, x_2, x_3, \ldots, x_n \sim exp(\theta)$$

$$q(\theta) \sim Gamma(\alpha_0, \beta_0) = \frac{\beta^\alpha}{\Gamma(\alpha)}\theta^{\alpha-1}e^{-\beta\theta}(\theta > 0)$$

$$l(x_{1\ldots n}|\theta) = \prod_i^n \theta e^{-\theta x_i} = \theta^n e^{-\theta\sum_i^n x_i}$$

$$P(\theta|x_{1\ldots n}) \propto q(\theta)l(x_{1\ldots n}|\theta) \propto \theta^{\alpha-1}e^{-\beta\theta}\theta^n e^{-\theta\sum_i^n x_i} \propto \theta^{\alpha+n-1}e^{-\theta(\beta+\sum_i^n x_i)} \sim Gamma(\alpha_0+n, \beta_0+\sum_i^n x_i)$$

Question 2

a

$$q(\theta|x_{1\ldots n-1}) = \theta^{\alpha-1+n-1}e^{-\theta(\beta+\sum_i^{n-1} x_i)} \sim Gamma(\alpha_{n-1}, \beta_{n-1})$$

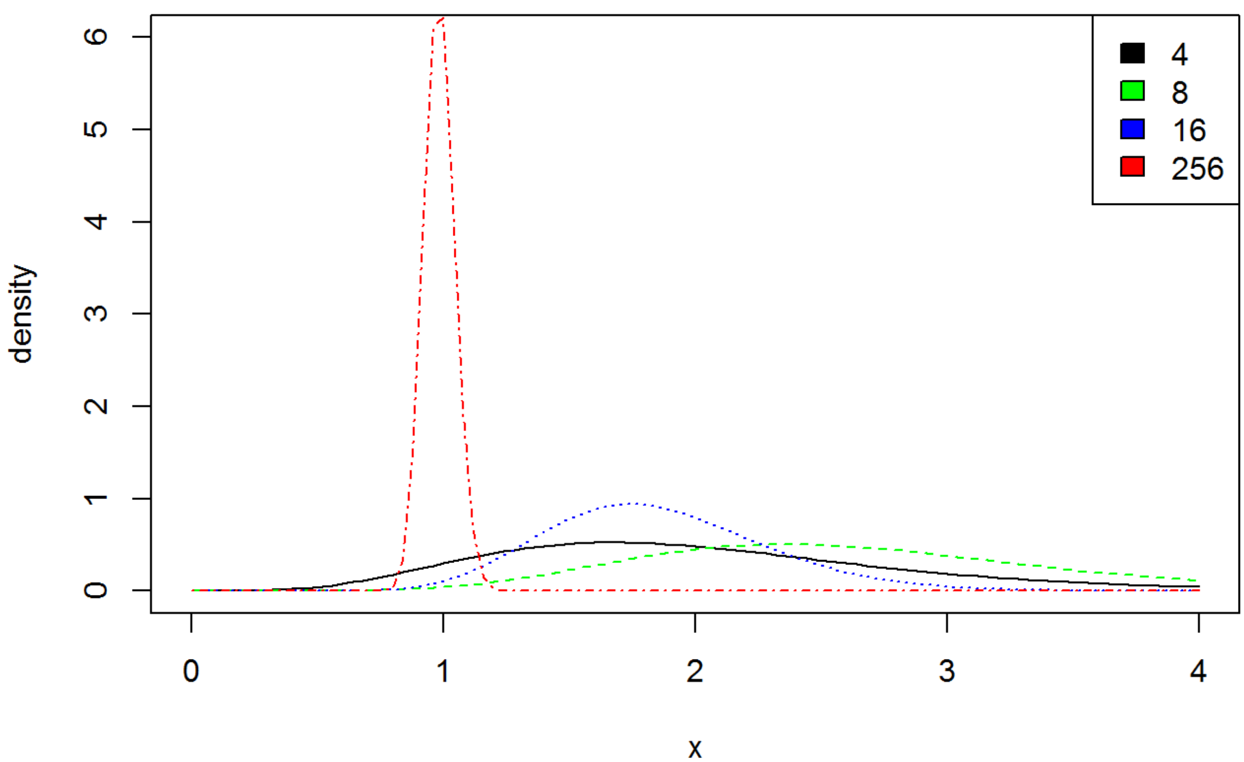
$$P(\theta|x_n) = q(\theta|x_{1\ldots n-1})\theta e^{-\theta x_n}$$

$$\alpha_n = \alpha_{n-1}+1, \beta_n = \beta_{n-1}+x_n$$

b

```
set.seed(123)
rv = rexp(256, 1)

curve(dgamma(x, shape=4+2, rate=0.2+sum(rv[1:4])), 0, 4, col = "black", ylim = c(0,6), ylab = "density")
curve(dgamma(x, shape=8+2, rate=0.2+sum(rv[1:8])), 0, 4, col = "green", ylim = c(0,6), ylab = "density", lt
y = 2, add = T)
curve(dgamma(x, shape=16+2, rate=0.2+sum(rv[1:16])), 0, 4, col = "blue", ylim = c(0,6), ylab = "density", l
ty = 3, add = T)
curve(dgamma(x, shape=256+2, rate=0.2+sum(rv)), 0, 4, col = "red", ylim = c(0,6), ylab = "density", lty = 4
, add = T)
legend("topright", legend=c(4,8,16,256), fill=c("black","green", "blue", "red"))
```



When we observed more data to update posterior distribution, its variance becomes smaller, so getting more precise on observed data.

Problem 2

$$P(\pi^1, \pi^2 | Y^{T_i}, T_i) = \frac{P(Y^{T_i}, T_i | \pi^1, \pi^2) q(\pi^1, \pi^2)}{\int P(Y^{T_i}, T_i | \pi^1, \pi^2) d(\pi^1, \pi^2)}$$

$$P(\pi^1, \pi^2|Y^{T_i}, T_i) \propto P(Y^{T_i}, T_i|\pi^1, \pi^2) \propto P(Y^{T_i}|T_i, \pi^1, \pi^2)P(T_i|\pi^1, \pi^2)$$
$$\propto \prod_i^n [\pi_1^{Y_i^1}(1-\pi_1)^{\sum(T_i=1)-Y_i^1}][\pi_2^{Y_i^2}(1-\pi_2)^{\sum(T_i=2)-Y_i^2}]$$

$$\pi^1|Y_{1\ldots n}^{T_i}, T_i \sim Beta(\sum_i^n Y_i^1 + 1, \sum_i^n I(T_i = 1) - \sum_i^n Y_i^1 + 1)$$

$$\pi^2|Y_{1\ldots n}^{T_i}, T_i \sim Beta(\sum_i^n Y_i^2 + 1, \sum_i^n I(T_i = 2) - \sum_i^n Y_i^2 + 1)$$

Problem 3

a

$$E(\bar{X}) = E(\sum_i^n X_i/n) = \frac{1}{n}\sum_i^n E(X_i) = \frac{n\lambda}{n} = \lambda$$

b

Let $\hat{\lambda}$ be another unbiased estimator

$$MSE: E(\hat{\lambda}-\lambda)^2 = E(\hat{\lambda}-\hat{X}+\hat{X}-\lambda)^2$$
$$= E(\hat{\lambda}-\bar{X})^2 + E(\bar{X}-\lambda)^2 + 2E(\hat{\lambda}-\bar{X})(\bar{X}-\lambda)$$
$$= E(\hat{\lambda}-\bar{X})^2 + E(\bar{X}-\lambda)^2$$

$E(\bar{X}-\lambda)^2$ is fixed biase^2 of estimator. Therefore, only $\hat{\lambda} = \bar{X} \rightarrow E(\hat{\lambda}-\bar{X})^2 = 0 \rightarrow MSE$is min