HW3

Yigao Li

September 23, 2017

Problem 17

```
n = 100000
data17 = rbeta(n, 0.5, 0.8)
cat("E(X) =", mean(data17))
```

```
## E(X) = 0.3848093
```

```
cat("\ns(X) =", sd(data17))
```

```
##
## s(X) = 0.3206881
```

```
cat("\nE(X^(-1/3)) =", mean(data17^(-1/3)))
```

```
##
## E(X^(-1/3)) = 2.659642
```

Problem 18

```
for (n in c(340000:350000)){
  data18 = rbeta(n, 0.5, 0.8)
  sample_mean = mean(sqrt(data18))
  sample_var = var(sqrt(data18))/n
  sample_sd = sqrt(sample_var)
  error = qnorm(0.975, mean = sample_mean, sd = sample_sd) - sample_mean
  if (error < 1e-3){
    sim = n
    break
  }
}
cat("Simulation of", sim, "times is necessary to approximate E(sqrt(x)) with an error of less than 10^-
```

```
## Simulation of 340879 times is necessary to approximate E(sqrt(x)) with an error of less than 10^-3.
```

By Central Limit Theorem, when n is as large as 340806 times, we are 95% confident that this simulation to approximate $E[\sqrt{X}]$ with an error of less than 10^{-3}.

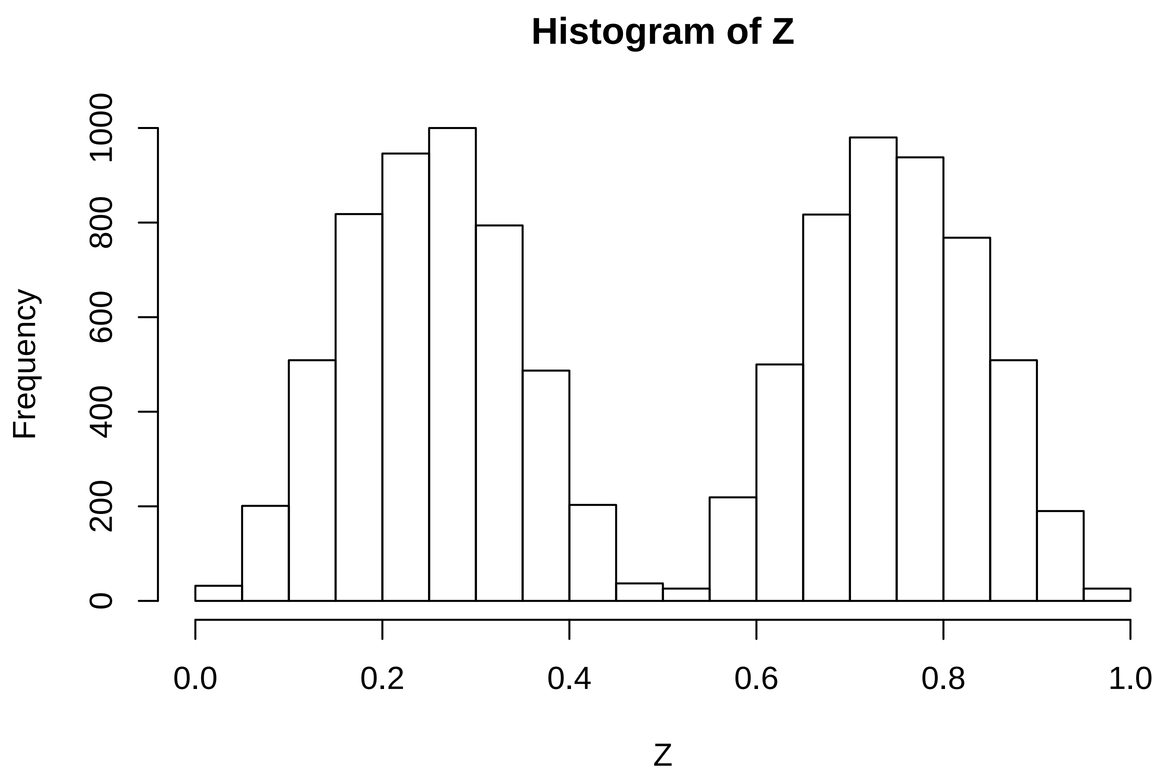
Problem 19

```
while (TRUE){
  x = runif(1)
  y = runif(1)
  if (y <= (sin(2*pi*x))^2){
    Z = x
    break
```

```
  }
}

while (TRUE){
  x = runif(1)
  y = runif(1)
  if (y <= (sin(2*pi*x))^2){
    Z = append(Z, x)
    if (length(Z) == 10000){
      break
    }
  }
}

hist(Z)
```



Histogram shows reflective 2 peaks around 0.25 and 0.75. Function $y = sin^2(2\pi x)$ has a closely similar plot with this histogram.

Problem 20

a) Sensitivity is the probability that a drunk driver is correctly identified as drunk. Specificity is the probability that a sober driver is correctly identified as being sober.

b) $Pr(Fail|Drunk) = 0.99,\ Pr(Fail|Sober) = 0.02,\ Pr(Drunk) = \frac{1}{500}$

$$Pr(Drunk|Fail) = \frac{Pr(Fail|Drunk) \times Pr(Drunk)}{Pr(Fail)}$$

where

$$\begin{aligned} Pr(Fail) &= Pr(Fail|Drunk) \times Pr(Drunk) + Pr(Fail|Sober) \times Pr(Sober) \\ &= Pr(Fail|Drunk) \times Pr(Drunk) + Pr(Fail|Sober) \times (1 - Pr(Drunk)) \\ &= 0.99 \times \frac{1}{500} + 0.02 \times \frac{499}{500} \\ &= 0.02194 \end{aligned}$$

$$Pr(Drunk|Fail) = \frac{0.99 \times \frac{1}{500}}{0.02194} \approx 0.09024$$

Problem 21

$X \sim Binom(n = 50, p = 0.2)$ (1)

```
s20 = pbinom(19, 50, 0.2)
g10s20 = pbinom(19, 50, 0.2) - pbinom(10, 50, 0.2)
g10cs20 = g10s20/s20
cat("Pr(X<20) =", s20)
```

```
## Pr(X<20) = 0.9990676
```

```
cat("\nPr(X>10|X<20) =", g10cs20)
```

```
##
## Pr(X>10|X<20) = 0.4158959
```

(2) $X \sim Binom(n = 500, p = 0.1)$

(3)

```
n = 1000
x = rbinom(n, 500, 0.1)
xs60 = sum(x<60)/n
xg30 = sum(x>30)/n
xg30s60 = sum((x>30)*(x<60))/n
xs60cg30 = xg30s60/xg30
xg30cs60 = xg30s60/xs60
cat("Pr(X<60) =", xs60)
```

```
## Pr(X<60) = 0.908
```

```
cat("\nPr(X<60|X>30) =", xs60cg30)
```

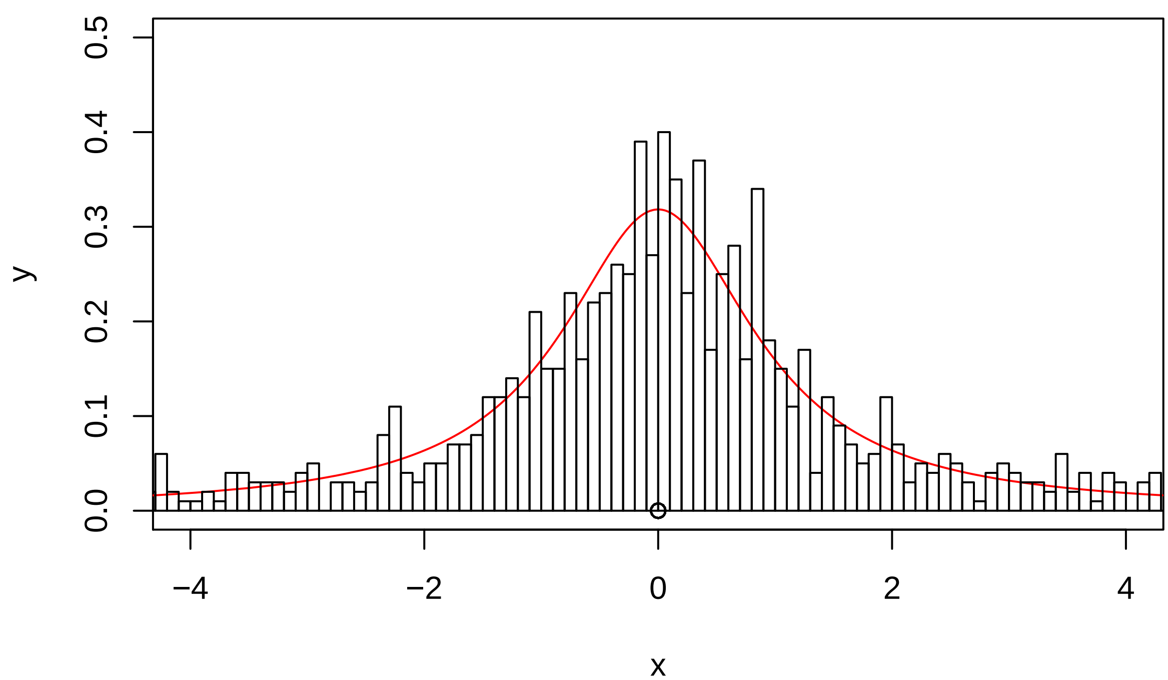
```
##
## Pr(X<60|X>30) = 0.908
```

```
cat("\nPr(X>30|X<60) =", xg30cs60)
```

```
##
## Pr(X>30|X<60) = 1
```

Problem 22

```
size = 1000
n = 10
x = integer(size)
y = integer(size)
plot(x, y, xlim = c(-4,4), ylim = c(0,0.5))
for (i in c(1:n)){
  xsub = rcauchy(size)
  x = x + xsub
  xsub = sort(xsub)
  ysub = dcauchy(xsub)
  if (i == n){
    hist(x/n, freq = FALSE, breaks = 4000, add = TRUE)
  }
  lines(xsub, ysub, col = rainbow(x)[i])
}
```



Problem 23

$$
\begin{aligned}
Pr(Y \le y) &= Pr[(X|X > A) - A \le y] \\
&= Pr[(X|X > A) \le A + y] \\
&= Pr(X \le A + y|X > A) \\
&= 1 - Pr(X > A + y|X > A)
\end{aligned}
$$

By Memoryless Property of Exponential Distribution,

$$\begin{aligned} Pr(Y \le y) &= 1 - Pr(X > y) \\ &= 1 - e^{-\lambda y} \end{aligned}$$