Homework 5

Sharleen Price spp2122

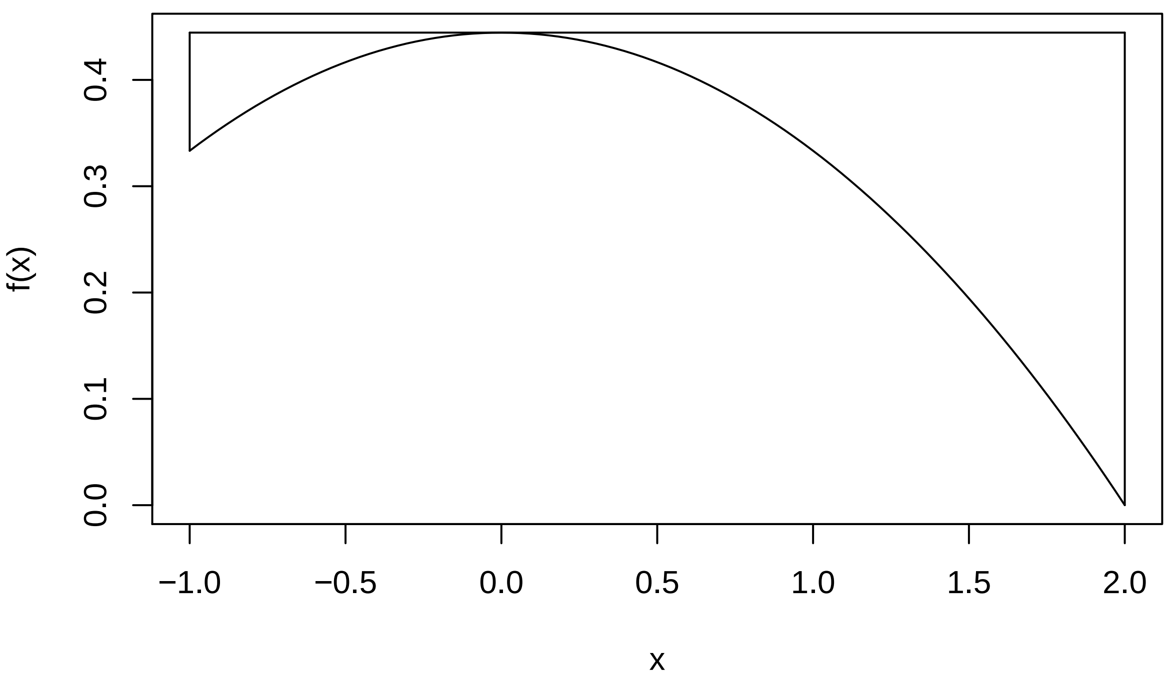
4/25/2018

1) reject-accept

```
#1) function
f<-  function(x){
  return(ifelse((-1<=x | 2>=x),(1/9)*(4-x^2),0))
}

x <- seq(-1, 2, length = 100)
f.max=max(f(x))
plot(x, f(x), type="l", ylab="f(x)")

#envelope
e <- function(x) {
  return(ifelse((-1 > x | x > 2), Inf, f.max))
}
lines(c(-1, -1), c(f(-1), e(-1)), lty = 1)
lines(c(2, 2), c(0, e(2)), lty = 1)
lines(x, e(x), lty = 1)
```



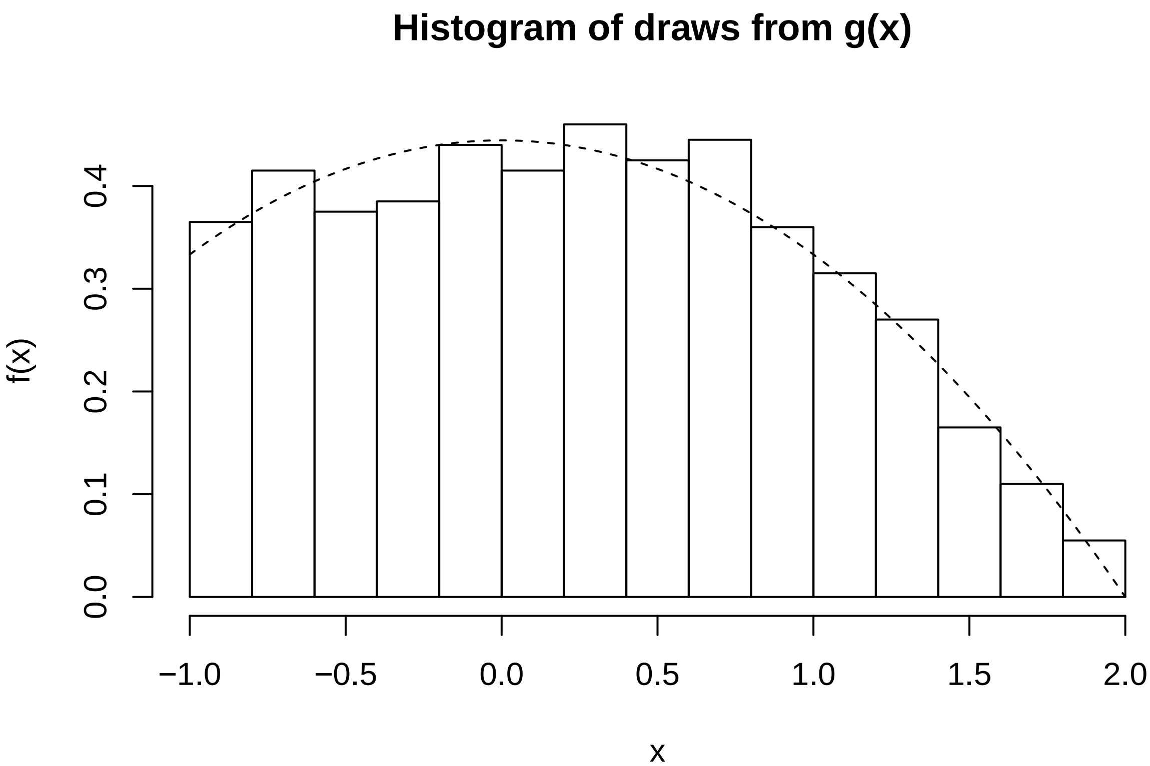

```
#accept-reject algorithm

n.samps <- 1000   # number of samples desired>
n       <- 0                      # counter for number samples accepted>
samps   <- numeric(n.samps) # initialize the vector of output>
while (n < n.samps)
  {
   y <- runif(1,-1,2)     #random draw from g
```

```
    u <- runif(1)
    if (u < f(y)/e(y))
      {n <- n + 1
      samps[n] <- y  } #simulated data
}

#histogram
hist(samps, prob = T, ylab = "f(x)", xlab = "x", main = "Histogram of draws from g(x) ")

lines(x, f(x), lty = 2)
```



2)

Regression and Empirical Size

Regression 1,2: Read in the grocery retailer dataset.Name the dataset grocery. Use the least squares equation^ = (XTX)−1XTYto estimate regression model (2).To estimate the model, use the linear model function inR, i.e., uselm.

```
grocery <- read.table("grocery.txt", header = TRUE, as.is = TRUE)
head(grocery)
```

```
##      Y      X1   X2 X3
## 1 4264 305.657 7.17  0
## 2 4496 328.476 6.20  0
## 3 4317 317.164 4.61  0
## 4 4292 366.745 7.02  0
## 5 4945 265.518 8.61  1
## 6 4325 301.995 6.88  0
```

```
X <- cbind(rep(1,52), grocery$X1, grocery$X2, grocery$X3)
beta_hat <- solve((t(X) %*% X)) %*% t(X) %*% grocery$Y
round(t(beta_hat), 2)
```

```
##         [,1] [,2]   [,3]   [,4]
## [1,] 4149.89 0.79 -13.17 623.55
```

```
lm0 <- lm(Y ~ X1 + X2 + X3, data = grocery)
lm0
```

```
##
## Call:
## lm(formula = Y ~ X1 + X2 + X3, data = grocery)
##
## Coefficients:
## (Intercept)           X1           X2           X3
##   4149.8872       0.7871     -13.1660     623.5545
```

Regression 3: Use R to estimate 2, i.e., computeMSE=1n−4 ni=1(Yi−ˆYi)2. To perform this task,use theresidualsfunction.

```
mse<- mean(lm0$residuals^2)
mse
```

```
## [1] 18952.5
```

```
#summary(lm0)
```

2.2 Test Slope

```
summary(lm0)
```

```
##
## Call:
## lm(formula = Y ~ X1 + X2 + X3, data = grocery)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -264.05 -110.73  -22.52   79.29  295.75
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept) 4149.8872   195.5654  21.220  < 2e-16 ***
## X1             0.7871     0.3646   2.159   0.0359 *
## X2           -13.1660    23.0917  -0.570   0.5712
## X3           623.5545    62.6409   9.954 2.94e-13 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 143.3 on 48 degrees of freedom
## Multiple R-squared:  0.6883, Adjusted R-squared:  0.6689
## F-statistic: 35.34 on 3 and 48 DF,  p-value: 3.316e-12
```

2.3 Sampling Distribution

1)

```
R<- 10000
stat.list<-rep(NA,R)
beta.list<-rep(NA,R)
for (i in 1:R){
  y.sim<- 4200-0*grocery$X1+15*grocery$X2+620*grocery$X3 + rnorm(52, mean=0, sd = 5)
  model.sim<-lm(y.sim~grocery$X1+grocery$X2+grocery$X3)
```

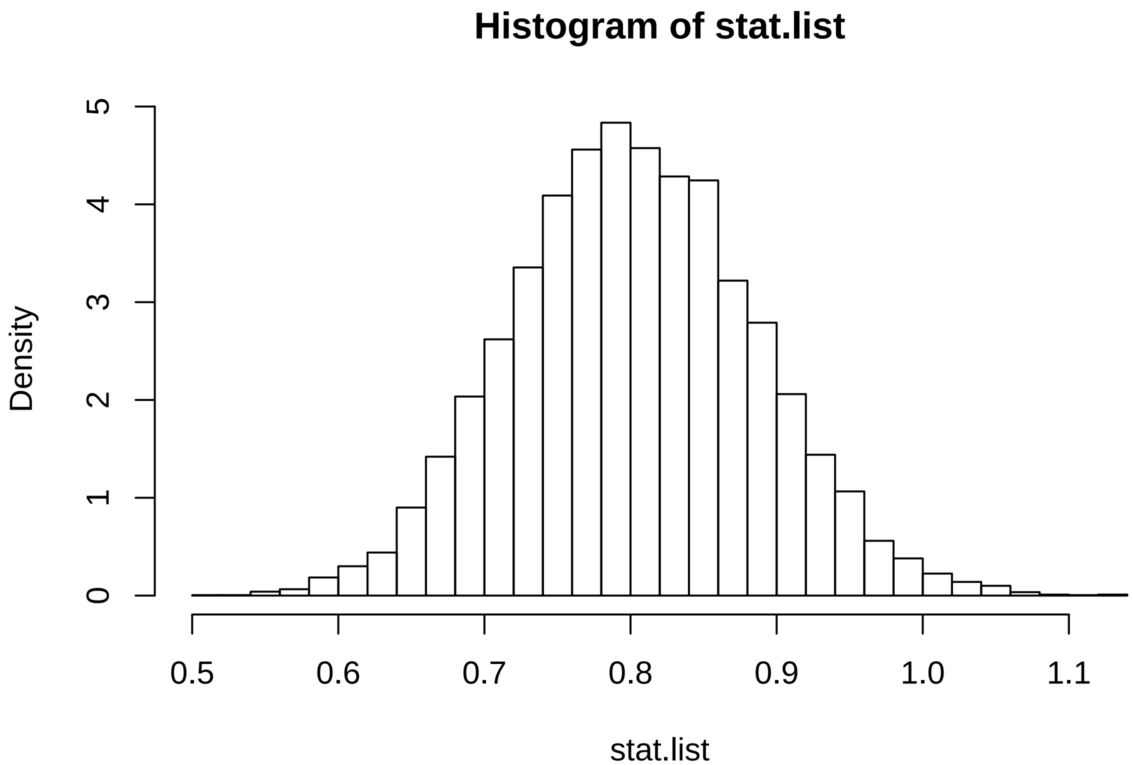
```
  stat.list[i]<-summary(model.sim)[[4]][3,2]

}
#stat.list
head(stat.list)
```

```
## [1] 0.7664888 0.8162480 0.6624979 0.8034457 0.7960991 0.8856065
```

2)

```
hist(stat.list, probability = TRUE, breaks = 40)
```




```
#t.list2<-seq(0.5,1.1,by=.01)
#lines(t.list2, dt(t.list2, 49), col = "green")


quantile(stat.list, .9) #simulated quantile
```

```
##       90%
## 0.9092806
```

```
qt(.9, 47) #true quantile
```

```
## [1] 1.299825
```

```
quantile(stat.list, .95) #simulated quantile
```

```
##       95%
## 0.9405444
```

```
qt(.95, 47) #true quantile
```

```
## [1] 1.677927
```

```r
quantile(stat.list, .99) #simulated quantile
```

```
##      99%
## 1.002643
```

```r
qt(.99, 47) #true quantile
```

```
## [1] 2.408345
```