STAT431_HW4

Taiga Hasegawa(taigah2)

2019/3/9

Question1

(a)

$$E[(\pi - E(\pi|y))^2|Y = y] = \int_0^1 (\pi - E(\pi|y))^2 p(\pi|y) d\pi = \frac{\int_0^1 (\pi - E(\pi|y))^2 p(\pi) L(\pi; y) d\pi}{\int_0^1 p(\pi) L(\pi; y)}$$

(b)

```
like <- function(pi) pi^12 * (1-pi)^(58)

prior <- function(pi) dbeta(pi, 0.5, 0.5)

posterior_mean=(12+0.5)/(70+0.5+0.5)
```

```
numerator=integrate(function(pi) (pi-posterior_mean)^2 * prior(pi) * like(pi), 0, 1)
denominator=integrate(function(pi) prior(pi) * like(pi), 0, 1)
numerator$value/denominator$value
```

```
## [1] 0.002014729
```

(c)

```
((12+0.5)*(58+0.5))/((12+58+0.5+0.5)^2*(12+58+0.5+0.5+1))
```

```
## [1] 0.002014729
```

I got the same answer with my approximation in (b).

Question2

(a)

```
x=c(136,119,114,113,113,112,112,108,102)
y=c(120,93,87,85,76)
ybar1=mean(x)
s1=sd(x)
ybar2=mean(y)
s2=sd(y)
```

Sample mean for small car is 114.3333333 and for SUVs is 92.2. Smaple standard deviations for small car is 9.3407708 and for SUVs is 16.6943104.

(b)

```
Nsim=100000
n1=9
n2=5
sigma1.2s = 1 / rgamma(Nsim, (n1-1)/2, (n1-1)*s1^2/2)
sigma2.2s = 1 / rgamma(Nsim, (n2-1)/2, (n2-1)*s2^2/2)
```

```r
mu1s = rnorm(Nsim, ybar1, sqrt(sigma1.2s/n1))
mu2s = rnorm(Nsim, ybar2, sqrt(sigma2.2s/n2))

quantile(mu1s - mu2s, c(0.025, 0.975))
```

```
##       2.5%      97.5%
##  0.2117792 43.7843854
```

Because this credible interval doesn't include 0, the means for small cars and SUVs are different.

(c)

```r
mean(mu1s < mu2s)
```

```
## [1] 0.02432
```

Posterior probability that the mean for small cars does not exceed the mean for SUVs is 2.379%.

(d)

```r
t.test(mu1s,mu2s,alternative = "greater",  var.equal=FALSE)
```

```
##
##  Welch Two Sample t-test
##
## data:  mu1s and mu2s
## t = 625.68, df = 122730, p-value < 2.2e-16
## alternative hypothesis: true difference in means is greater than 0
## 95 percent confidence interval:
##  22.03972      Inf
## sample estimates:
## mean of x mean of y
## 114.34730  92.24949
```

```r
t.test(mu1s,mu2s,alternative = "greater",  var.equal=TRUE)
```

```
##
##  Two Sample t-test
##
## data:  mu1s and mu2s
## t = 625.68, df = 2e+05, p-value < 2.2e-16
## alternative hypothesis: true difference in means is greater than 0
## 95 percent confidence interval:
##  22.03972      Inf
## sample estimates:
## mean of x mean of y
## 114.34730  92.24949
```

In both Welch t-test and usual t-test, the null hypothesis that mean for small cars does not exceed the mean for SUVs was rejected. This result is consistent with the answer in (b).

(e)

```r
quantile(sigma1.2s/sigma2.2s, c(0.025, 0.975))
```

```
##       2.5%      97.5%
## 0.03477796 1.57836479
```

Because this credible interval includes 1, we can say the variances of small cars and SUVs appear to be same.