REC3

jx

2019/9/21

problem1

```r
P <- 0.05
# x ~ Bin(500,p), P(x>=20) = 1- P(x<20)
1 - pbinom(19,500,0.05)
```

```
## [1] 0.8727655
```

```r
#  so the probability of identifying at least 20 glaucoma cases is 0.8727655
```

problem2

```r
# 5.21
# F-smoking(x=2.5) = 0.04779035
pnorm(2.5,3.5,0.6)
```

```
## [1] 0.04779035
```

```r
# 5.22
# F-nonsmoking(x=2.5) = 0.001349898
pnorm(2.5,4.0,0.5)
```

```
## [1] 0.001349898
```

```r
# 5.23
# 4.0 - 30x < 2.5 ---->  x > 0.05
#  P(x > 0.05) = 1 - F(x=0.05) =  0.1586553
1 - pnorm(0.05,0.03,0.02)
```

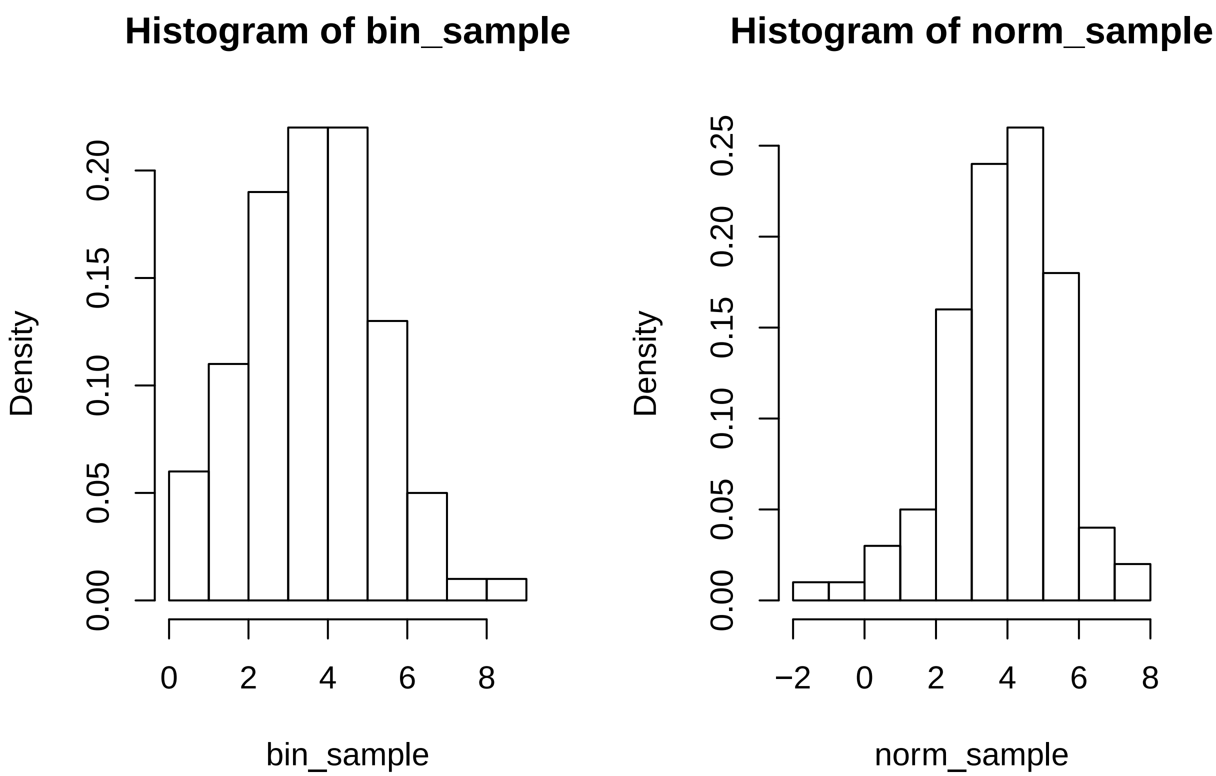
```
## [1] 0.1586553
```

```r
# 5.24
# 4.0 - 50x < 2.5 ---->  x > 0.03
#  P(x > 0.03) = 1 - F(x=0.03) = 0.5
1 - pnorm(0.03,0.03,0.02)
```

```
## [1] 0.5
```

```r
library(ggplot2)
library(ggpubr)
# 5.75
bin_sample <- rbinom(100,10,.4)
norm_sample <- rnorm(100,4,sqrt(2.4))

par(mfrow=c(1,2))
hist(bin_sample,freq = FALSE)
hist(norm_sample,freq = FALSE)
```




```r
dev.off()
```

```
## null device
##           1
```

```r
# np=4 , n(1-p)=6 both < 10, so may not be adequate

# 5.76  = 20 and p = .4 20
bin_sample <- rbinom(100,20,.4)
norm_sample <- rnorm(100,8,sqrt(4.8))

par(mfrow = c(1,2))
hist(bin_sample,freq = FALSE)
hist(norm_sample,freq = FALSE)

# np=8 , n(1-p)=12  so may not be adequate also

# # 5.77  = 50 and p = .4
```

```r
bin_sample <- rbinom(100,50,.4)
norm_sample <- rnorm(100,20,sqrt(12))

par(mfrow = c(1,2))
hist(bin_sample,freq = FALSE)
hist(norm_sample,freq = FALSE)
```

```r
# 6.5
# SE =SD/sqrt(n)
0.5/sqrt(40)
```

```
## [1] 0.07905694
```

```r
0.4/sqrt(32)
```

```
## [1] 0.07071068
```

```r
# 6.6
# it means the mean of the samples' mean follow a normal distribution

# 6.7
pt(0.99, df=16)
```

```
## [1] 0.8315447
```

```r
# upper 1 = 0.8315447

# 6.8
pt(0.1, df=28)
```

```
## [1] 0.5394716
```

```r
# lower 10 =  0.5394716

# 6.9
pt(0.975, df=7)
```

```
## [1] 0.8189791
```

```r
# upper 2.5 =  0.8189791


# 6.10
pchisq(0.975, df=2) # upper 0.3858401
```

```
## [1] 0.3858401
```

```
pchisq(0.025, df=2) # upper 0.0124222
```

```
## [1] 0.0124222
```

```
# 6.11
age <- c(30,73,40,47,25,82,60,56,43,50,59,4,22,33,30,32,36,69,47,22,11,19,67,43,41)
mean(age) #41.64
```

```
## [1] 41.64
```

```
sd(age)
```

```
## [1] 19.75871
```

```
s <- (19.75871 * sqrt(25))/sqrt(24)   ##  20.16615
sqrt(var(age)*25/24)   #double-check
```

```
## [1] 20.16615
```

```
mean(age) - qt(0.975,24) * (s/sqrt(25)) #  lower = 33.31582
```

```
## [1] 33.31582
```

```
mean(age) - qt(0.025,24) * (s/sqrt(25)) #  upper = 49.96418
```

```
## [1] 49.96418
```

```
# 6.12
wbc <- c(8,5,12,4,11,6,8,7,7,12,7,3,11,14,11,9,6,6,5,6,10,14,4,5,5)
mean(wbc) #7.84
```

```
## [1] 7.84
```

```
sd(wbc)
```

```
## [1] 3.2104
```

```
s <- (3.2104 * sqrt(25))/sqrt(24)   ##  3.276601
sqrt(var(wbc)*25/24)   #double-check
```

```
## [1] 3.276601
```

```
mean(wbc) - qt(0.975,24) * (s/sqrt(25)) #  lower = 6.487486
```

```
## [1] 6.487486
```

```r
mean(wbc) - qt(0.025,24) * (s/sqrt(25)) #  upper = 9.192514
```

```
## [1] 9.192514
```

```r
# 6.13
mean(wbc) - qt(0.95,24) * (s/sqrt(25)) #  lower =  6.718824
```

```
## [1] 6.718824
```

```r
mean(wbc) - qt(0.05,24) * (s/sqrt(25)) #  upper =  8.961176
```

```
## [1] 8.961176
```

```r
# 6.14
# the
```