3A) $\hat{\lambda} = \bar{X} = \frac{1}{n}\sum_{i} X_i = 24.9, \ n = 1$

$\delta = 0.5, 1, 1.5, 2, 2.5$

$\mathbb{E}[X_i] = \mathrm{Var}[X_i] = \lambda_0$

CLT:

$$\sqrt{n}\,\frac{(\bar{X} - \lambda_0)}{\sqrt{\lambda_0}} \xrightarrow{n \to \infty} N(0,1)$$

$$\sqrt{\frac{\lambda_0}{n}} \approx \sqrt{\frac{\hat{\lambda}}{n}} = \sqrt{\frac{24.9}{23}} \boxed{\approx 1.04}$$

$$P(|\lambda_0 - \hat{\lambda}| > \delta) = P(|\lambda_0 - \bar{X}| > \delta)$$

$$= P\left(\sqrt{n}\,\frac{|\lambda_0 - \bar{X}|}{\sqrt{\lambda_0}} > \sqrt{n}\,\frac{\delta}{\sqrt{\lambda_0}}\right)$$

$$\approx P\left(|N(0,1)| > \sqrt{n}\,\frac{\delta}{\sqrt{\hat{\lambda}}}\right)$$

$$= P\left(|N(0,1)| > \frac{\sqrt{23}}{\sqrt{24.9}}\delta\right)$$

$$= 2\left(1 - \Phi\left(\sqrt{\frac{23}{24.9}}\,\delta\right)\right)$$

$\approx 1.315, 1.168, 1.075, 1.027, 1.008$

3B)

$$P(X=x) = \frac{\lambda^x e^{-\lambda}}{x!} = \frac{\hat{\lambda}^{x_i} e^{-\hat{\lambda}}}{x_i!}(300) = \hat{n}_i$$

MLE of λ, $\hat{\lambda} = \bar{X}$

$$\hat{\lambda} = \frac{292}{75} \approx 3.89\overline{3}$$

estimate $= 3.89\dot{3}$

The difference in the observed $(3.89\overline{3})$ and the expected (3.893) is quite close to one another, which shows this is a good fit

3 c)

$$\{x_1, x_2, \ldots x_n), \quad \hat{\mu}_k = \frac{1}{n}\sum_{i=1}^{n} X_i^k$$

$$E(\hat{\mu}_k) = \mu_k$$

Kth moment $\mu_k = E(X^k)$

$$\mathbb{E}(\hat{\mu}_k) = E\left(\frac{1}{n}\sum_{i=1}^{n} X_i^k\right)$$

$$= \frac{1}{n}\sum_{i=1}^{n} \mathbb{E}(X_i^k)$$

$$= \frac{1}{n}\sum_{i=1}^{n} \mu_k$$

$$= \frac{n\mu_k}{n}$$

$$\boxed{= \mu_k}$$

3D) $\hat{p}_{MM}$ $\{X_1, X_2, \ldots, X_n\} \sim Geo(p)$

$k=1$ $\mathbb{E}(X) = \frac{1}{p}$

$$\frac{1}{\hat{p}_{MM}} = \frac{1}{n} \sum_{i=1}^{n} x_i$$

$$= \bar{X}_n$$

$$\boxed{\hat{p}_{MM} = \frac{1}{\bar{X}_n}}$$

3E)

a) $\Gamma(r+1) = r\Gamma(r)$ $\qquad e^{-t} = \frac{d}{dt}(-e^{-t})$

$$\Gamma(r+1) = \int_0^\infty t^r e^{-t} dt$$

$$= \left[t^r(-e^{-t})\right]_0^\infty - \int_0^\infty r t^{r-1}(-e^{-t}) dt$$

$$= 0 + r \boxed{\int_0^\infty t^{r-1}(-e^{-t}) dt} = \Gamma(r)$$

$$= 0 + r\Gamma(r)$$

$$\boxed{\Gamma(r+1) = r\Gamma(r)}$$

b) $\Gamma\left(\frac{1}{2}\right) = \sqrt{\pi}$

$$\Gamma(1/2) = \int_{t=0}^{+\infty} t^{\frac{1}{2}-1} e^{-t} dt$$

$$= \int_{t=0}^{+\infty} \frac{e^{-t}}{t^{1/2}} dt$$

$y = \sqrt{t}, \; dy = \frac{dt}{2\sqrt{t}}$

$$= 2\int_{y=0}^{+\infty} e^{-y^2} dy$$

$$= \int_{y=-\infty}^{+\infty} e^{-y^2} dy \boxed{= \sqrt{\pi}}$$

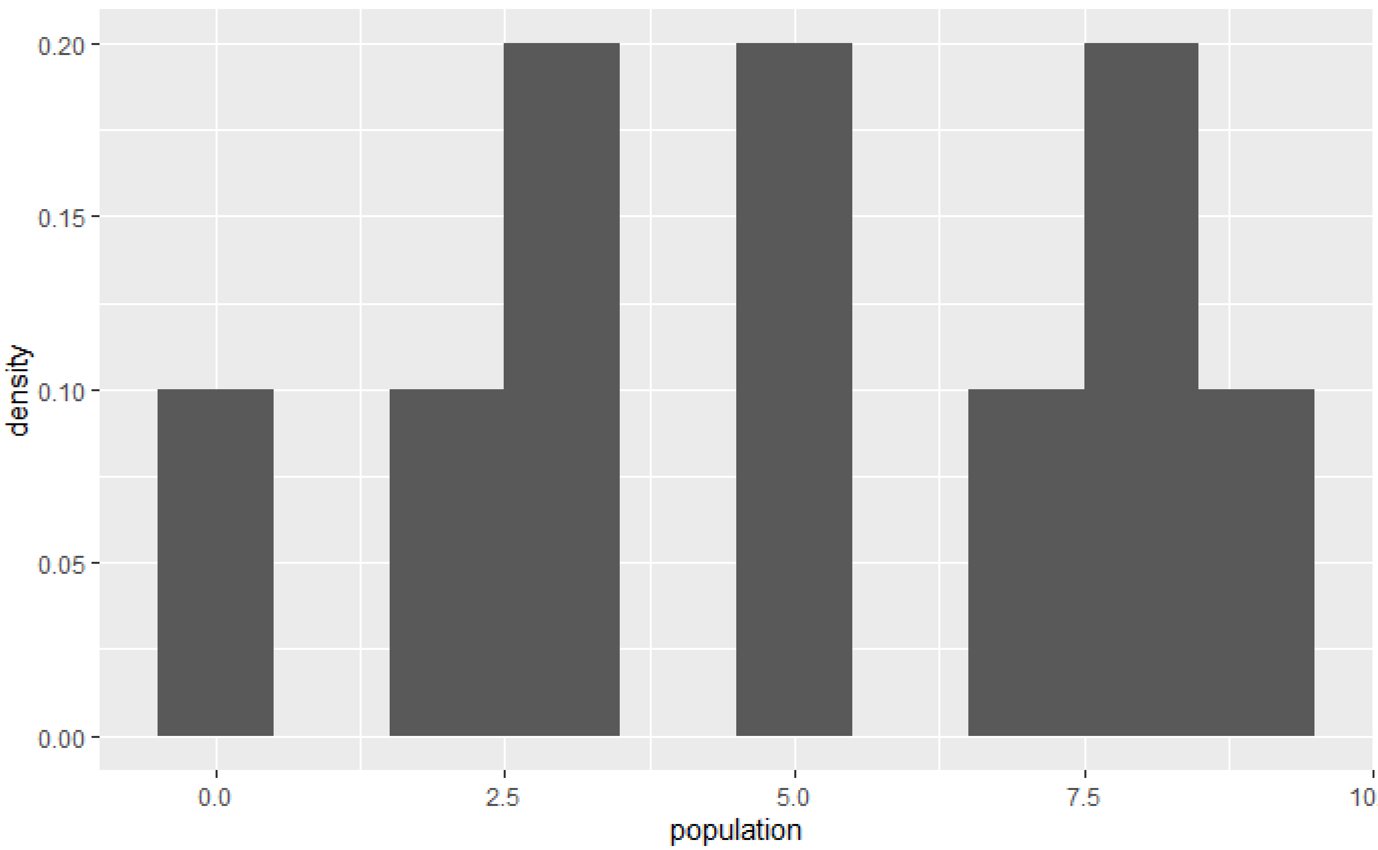
HW3 3F-J

Code

Problem 3F

Hide

```
library(ggplot2)
population <- c(3,0,3,5,8,5,7,9,2,8)
df_population <- data.frame(population)
ggplot(df_population, aes(x=population)) + geom_histogram(aes(y=..density..),binwidth = 1)
```



Hide

```
population_mean <- mean(population)
population_mean
```

```
[1] 5
```

Hide

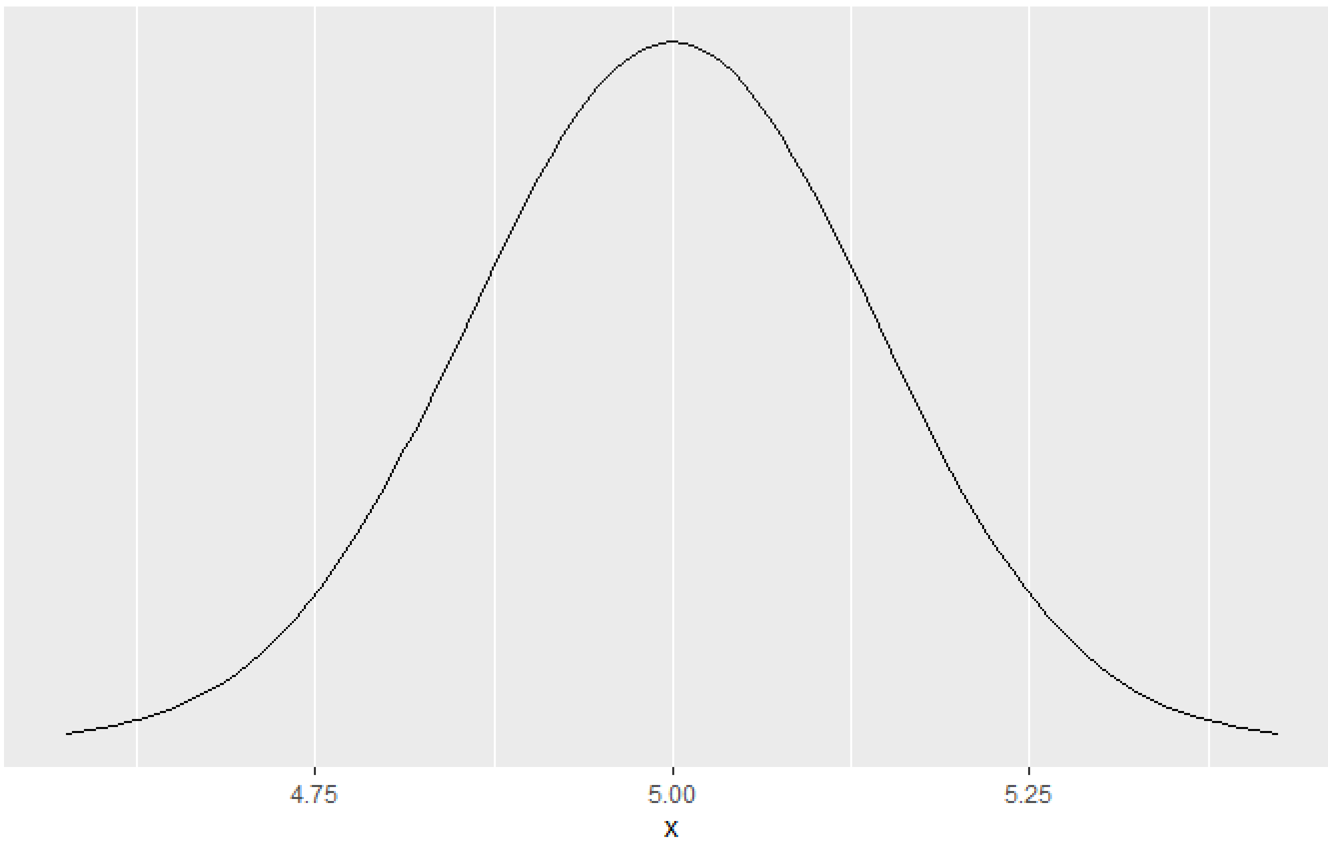
```
population_sd <- sqrt(mean((population-population_mean)^2))
population_sd
```

```
[1] 2.828427
```

Problem 3G

Hide

```
expected <- mean(population)
se <- population_sd/sqrt(400)
ggplot(data=data.frame(x=c(expected-3*se, expected+3*se)),aes(x))+ stat_function(fun= dnorm,n=10
1,args=list(mean= expected,sd= se))+ ylab("")+ scale_y_continuous(breaks=NULL)
```



Problem 3H

Hide

```
num_samples <- 100
sample_size <- 400

find_mean <- function(n){
  temp <- sample(population,size=n,replace=TRUE)
  mean(temp)}

means <- replicate(num_samples, find_mean(sample_size))
head(means)
```

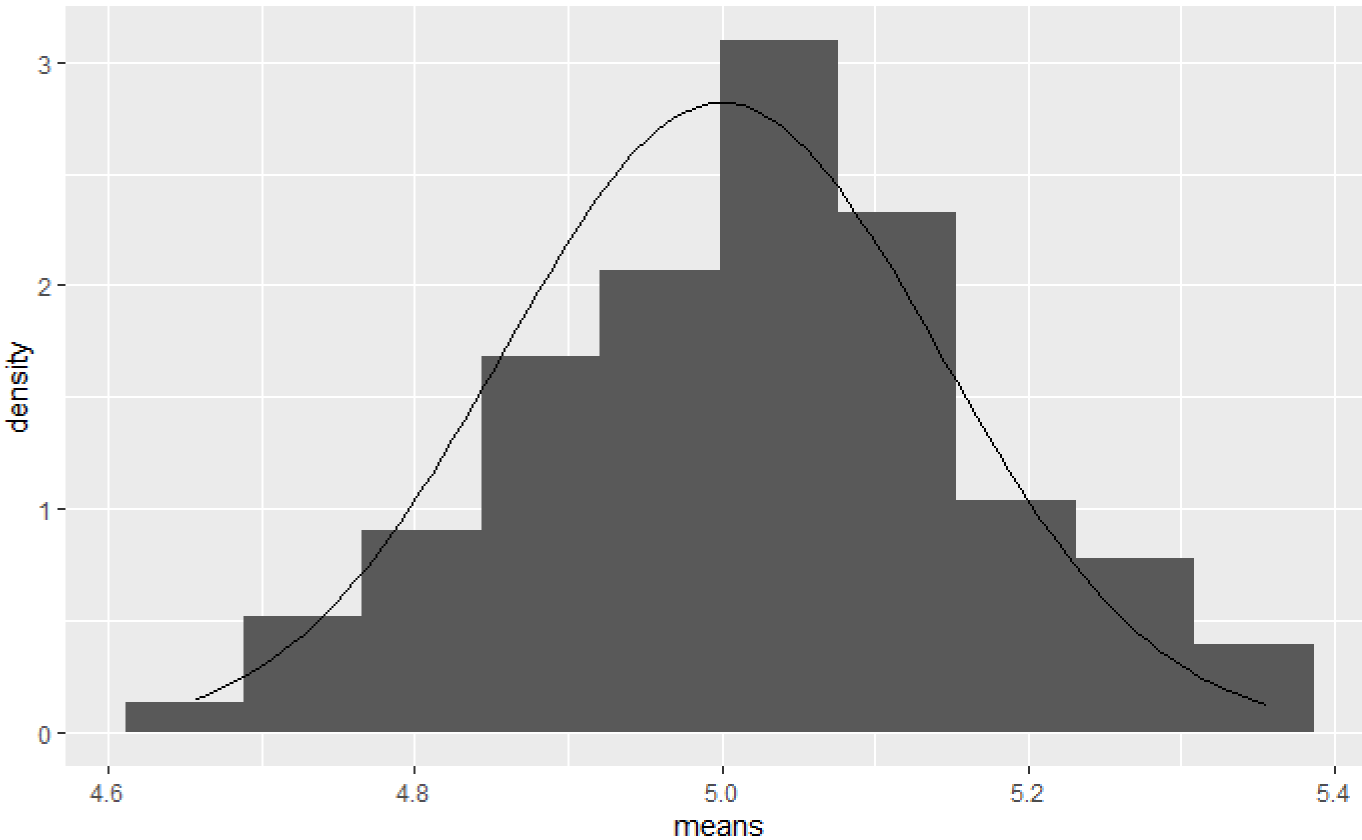
```
[1] 4.9300 5.0125 5.0500 5.0950 5.2925 4.9475
```

Hide

```
## [1] 4.9800 5.2100 4.9225 5.2950 5.0625 4.9675
```

```
new_data <- data.frame(means)

ggplot(data = new_data, aes(x = means)) +
  geom_histogram(aes(y = ..density..), bins=10) +
  stat_function(fun = dnorm, n = 101, args=list(mean = expected, sd = se))
```



Problem 3I

Hide

```
num_samples <- 100
sample_size <- 400

CI_range <- function(n){
  temp <- sample(population,size=n,replace=TRUE)
  CI_lower_bound <- mean(temp)-1.96*se
  CI_upper_bound <- mean(temp)+1.96*se
  ifelse(mean(population) < CI_upper_bound&mean(population) > CI_lower_bound, 1, 0)}

vec_good <- replicate(num_samples, CI_range(sample_size))
vec_good
```

```
  [1] 1 1 1 1 1 0 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 0 1 0 1 1 1 1 1
 [70] 1 1 1 1 1 1 1 1 1 0 1 0 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
```

Hide

```
num_good <- sum(vec_good)
num_good
```

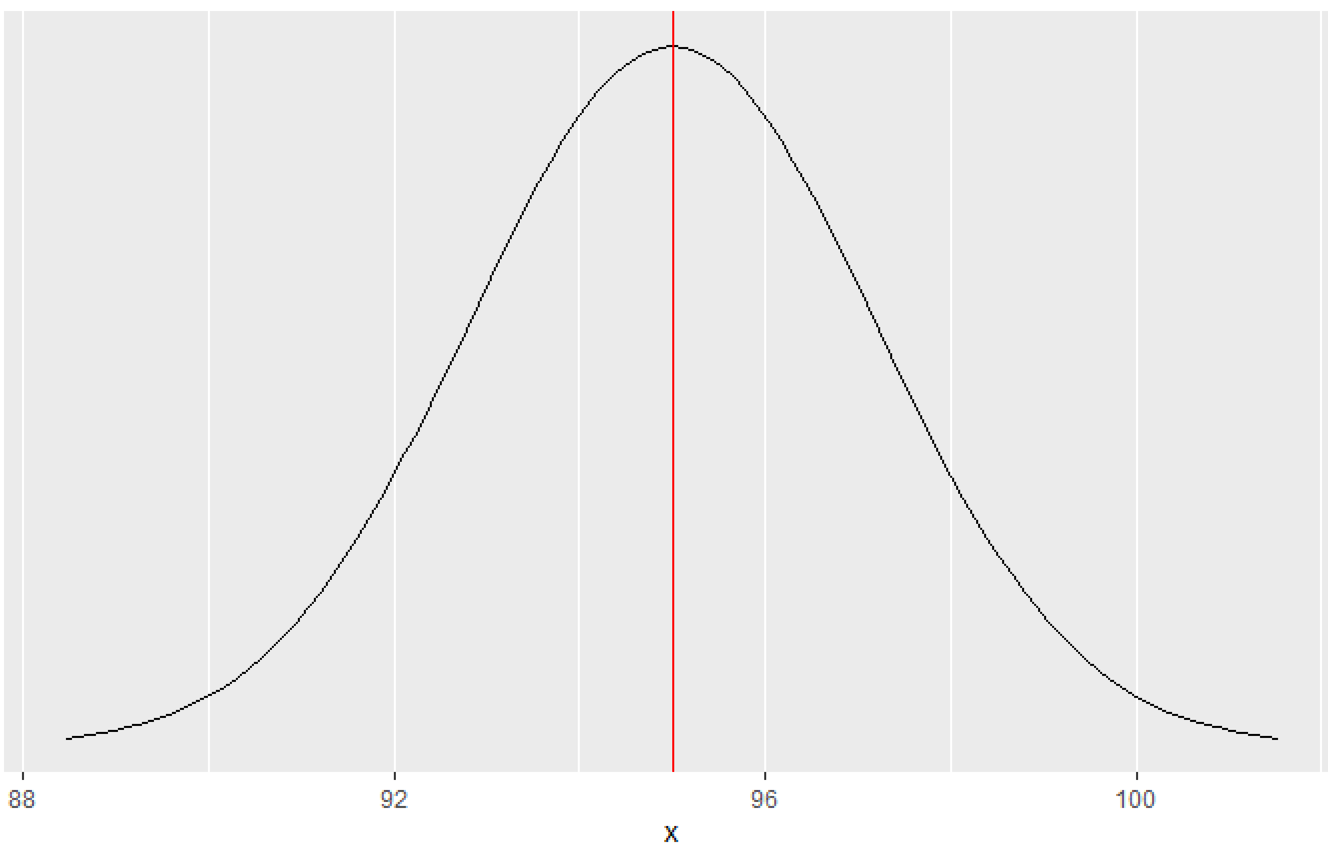
```
[1] 95
```

Problem 3J

Hide

```
expected = 100*(0.95)
se = sqrt(expected*(0.05))

ggplot(data = data.frame(x=c(expected-3*se, expected+3*se)), aes(x)) +
  stat_function(fun = dnorm, n = 101, args=list(mean=expected, sd=se)) +
  ylab("") +
  scale_y_continuous(breaks=NULL) +
  geom_vline(xintercept = sum(vec_good), col="red")
```



Based on the values of E(G) and SE(G), the mark looks exactly where it should be. It aligns directly in the middle of the normal distribution which, if all data is correct, is the most likely place for it to be. Overall, it is expected to be anywhere near the middle, which is what we see in this run of sampling.