a1q1

a

i.

$$Pr(X=a)/Pr(X=b) = \frac{\binom{n}{a}p^a(1-p)^{n-a}}{\binom{n}{b}p^b(1-p)^{n-b}} = \frac{\binom{n}{a}}{\binom{n}{b}}p^{a-b}(1-p)^{a-b}$$

$$= \frac{\binom{n}{a}}{\binom{n}{b}}odd^{a-b}$$

ii.

```
prob_ratio1 <- function (n, a, b, odds = 1) {
    (choose(n,a)/choose(n,b))*odds^(a-b)
}
```

2nd function

```
prob_ratio2 <- function (n, a, b, odds = 1) {
   p = odds/(1+odds)
   dbinom(a,n,p)/dbinom(b,n,p)
}
```

iii.

```
# choose() function
prob_ratio1(50, a = 5, b = 45)
```

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

```
prob_ratio1(50, a = 5, b = 45, odds = 9)
```

```
## [1] 6.765496e-39
```

```
# dbinom() function
prob_ratio2(50, a = 5, b = 45)
```

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

```
prob_ratio2(50, a = 5, b = 45, odds = 9)
```

```
## [1] 6.765496e-39
```

b.

c.

$$\frac{Pr(\tilde{p}_x = 0)}{Pr(\tilde{p}_y = 0)} = \frac{Pr(x/n = 0)}{Pr(y/n = 0)} = \frac{Pr(x = 0)}{Pr(y = 0)} = \frac{\binom{n}{0} p^0 (1-p)^(n-0)}{\binom{m}{0} p^0 (1-p)^(m-0)} = \frac{(1-p)^n}{(1-p)^m}$$

$$\frac{Pr(\tilde{p}_x = 1)}{Pr(\tilde{p}_y = 1)} = \frac{Pr(x/n = 1)}{Pr(y/n = 1)} = \frac{Pr(x = n)}{Pr(y = n)} = \frac{\binom{n}{n} p^n (1-p)^(n-n)}{\binom{m}{m} p^m (1-p)^(m-m)} = \frac{p^n}{p^m}$$

ii.

$$\frac{Pr(\tilde{p}_x = 0)}{Pr(\tilde{p}_y = 0)}$$

increases as m increases.

$$\frac{Pr(\tilde{p}_x = 1)}{Pr(\tilde{p}_y = 1)}$$

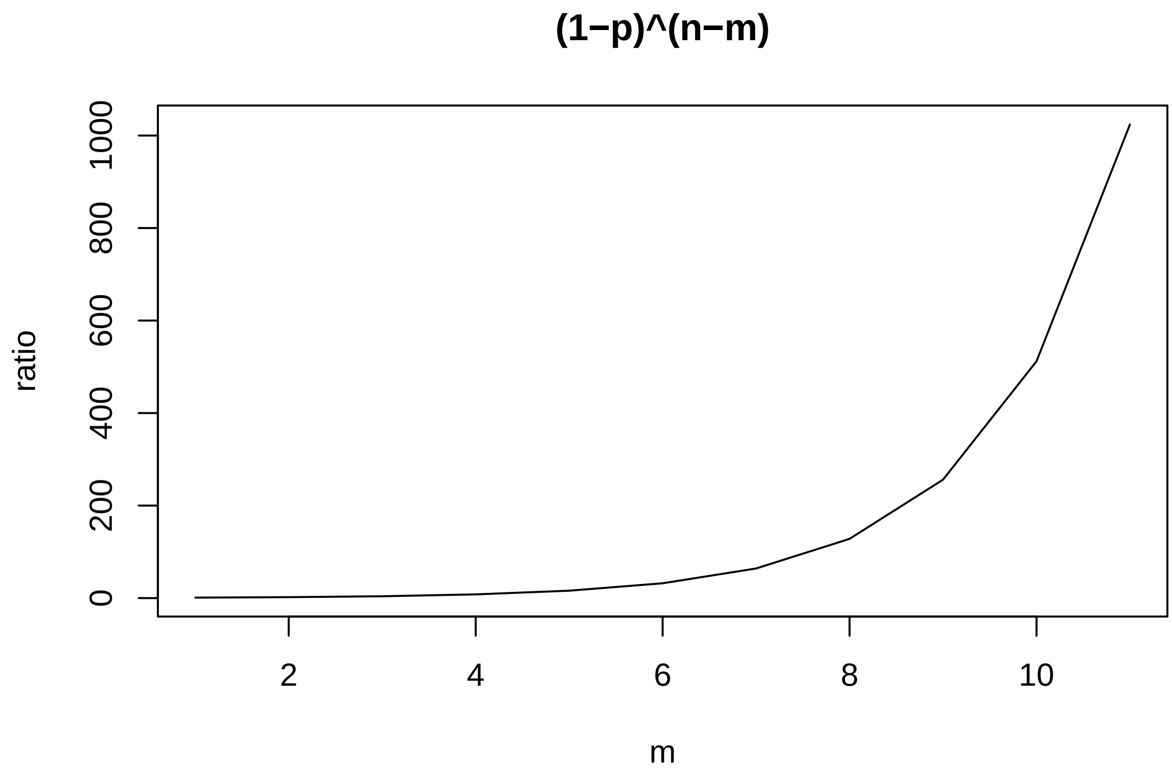
increases as m increases.

iii.

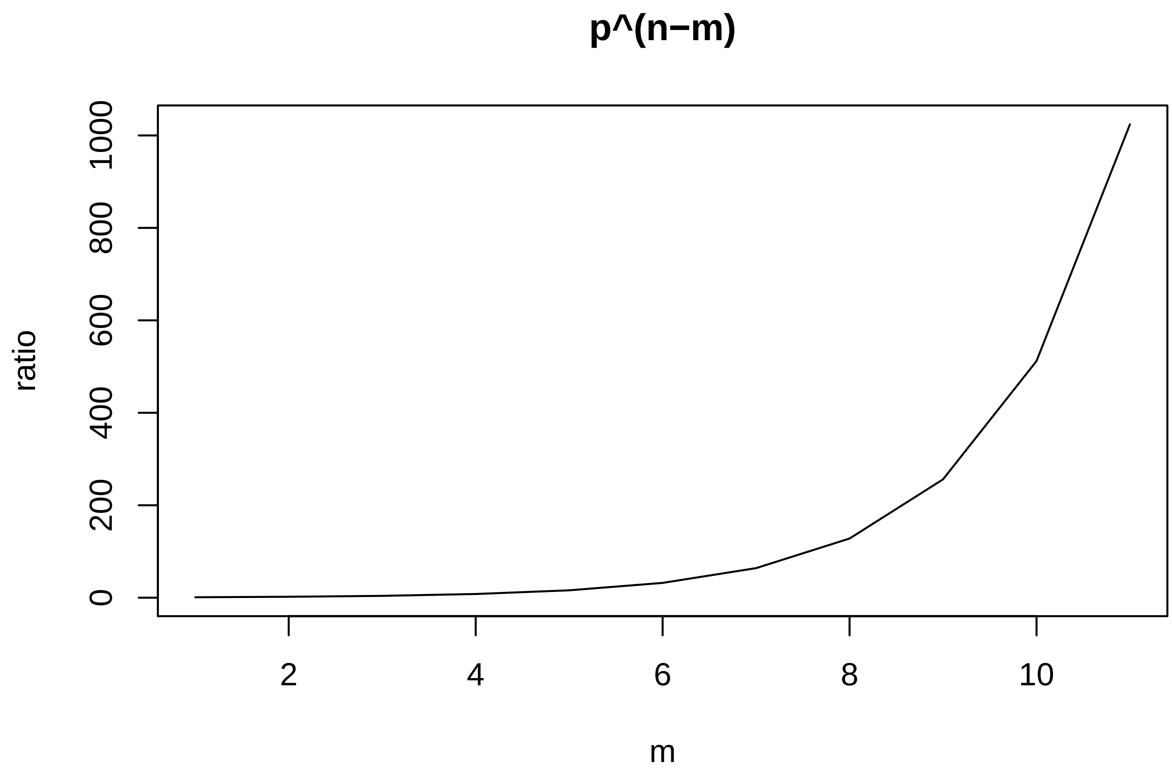
```
pxy_fun1<-function(n,m,p) {
  return ((1-p)^(n-m))
}
pxy_fun2<-function(n,m,p) {
  return (p^(n-m))
}
pxy_fun1(5,7,0.5)
```

```
## [1] 4
```

```
plot(pxy_fun1(5,5:15,0.5),main="(1-p)^(n-m)",xlab="m",ylab="ratio",type = 'l')
```




```r
plot(pxy_fun2(5,5:15,0.5),main="p^(n-m)",xlab="m",ylab="ratio",type = 'l')
```



iv. $\widehat{p}$ is likely to be 1 when the sample size is extremly small. We should be surprised when the sample size is big and $\widehat{p}$ is 1.

v.

vi.

$$E(\tilde{p}) = E(X/n) = (1/n)E(X) = (1/n)np = p$$

$$SD(\tilde{p}) = \sqrt{Var(X/n)} = \sqrt{(1/n^2)Var(X)} = \sqrt{(1/n^2)np(1-p)} = \sqrt{p(1-p)/n}$$

vii.

```
        sd_p_wig<-function(n, p) {
            sqrt(p*(1-p)/n)
    }
```

d.

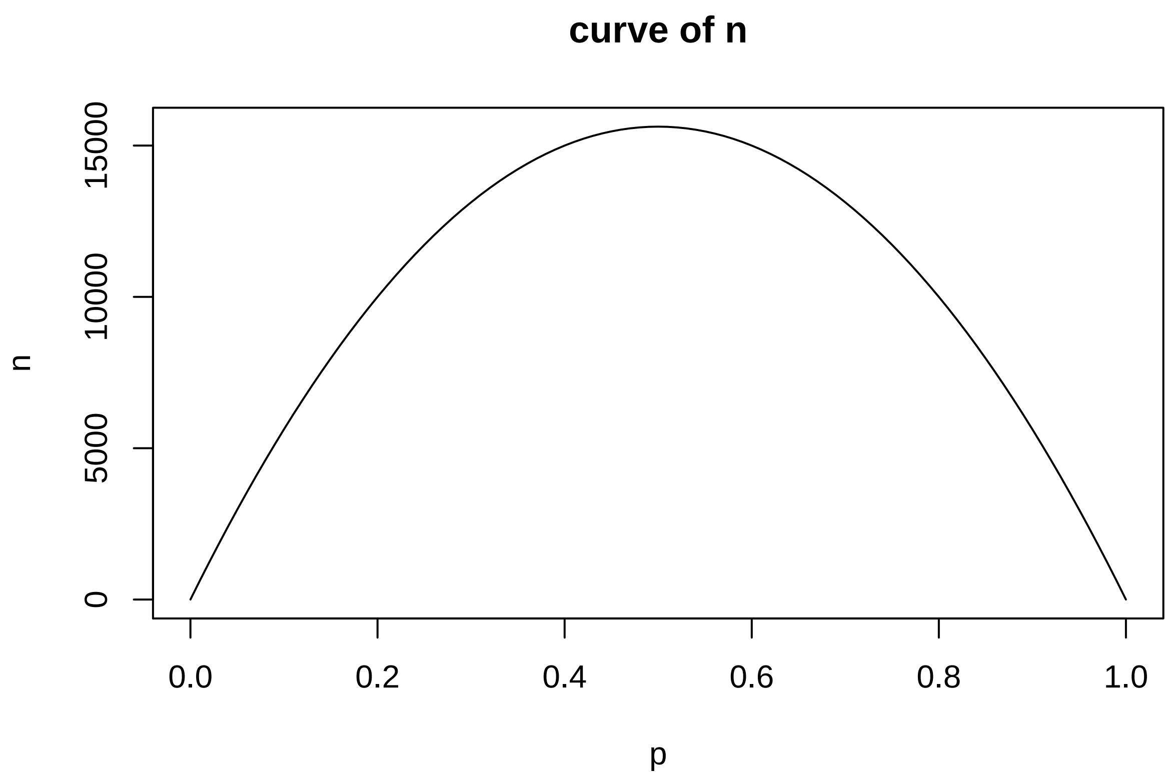
e.

$$Pr(|\tilde{p} - p| \geq k\sqrt{p(1-p)/n}) \leq 1/k^2$$

ii.

$$n = 250^2 p(1-p)$$

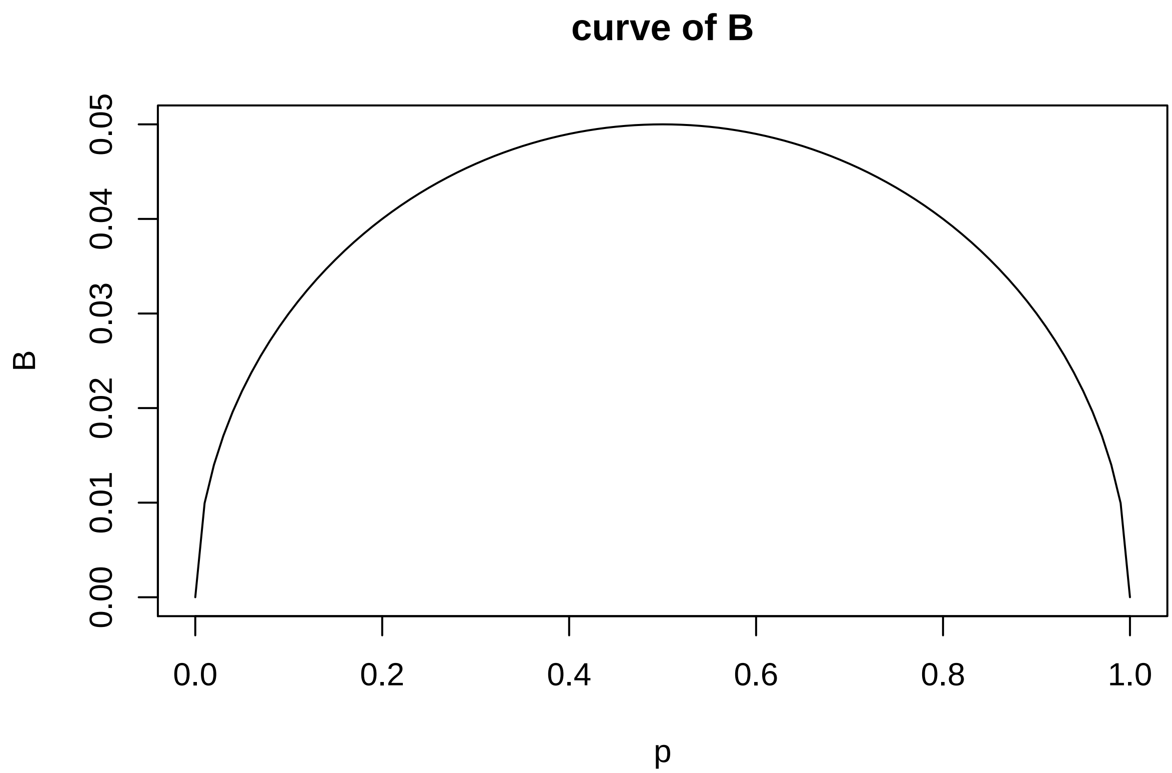
```r
n<-function( p) {
    250^2 * p * (1-p)
}
p = seq(0,1,0.01)
plot(p,n(p), main="curve of n", xlab="p", ylab="n", type = 'l')
```



iii.

$$B = k\sqrt{p(1-p)/n} = 1/10\sqrt{p(1-p)}$$

```r
B<-function( p) {
    1/10 * sqrt(p * (1-p))
}
p = seq(0,1,0.01)
plot(p,B(p), main="curve of B", xlab="p", ylab="B", type = 'l')
```

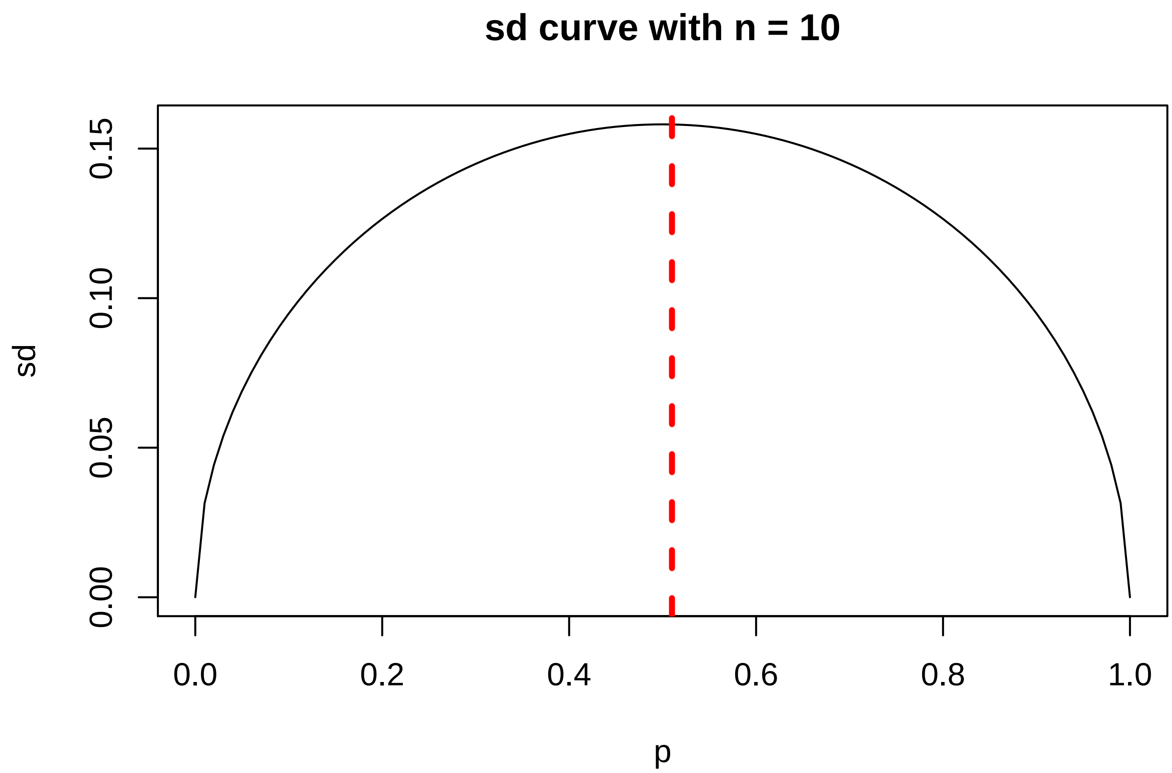


iv. According to Chebyshev's inequality, the largest B says p is likely to be 0.5 when $n = 2500$.

e.

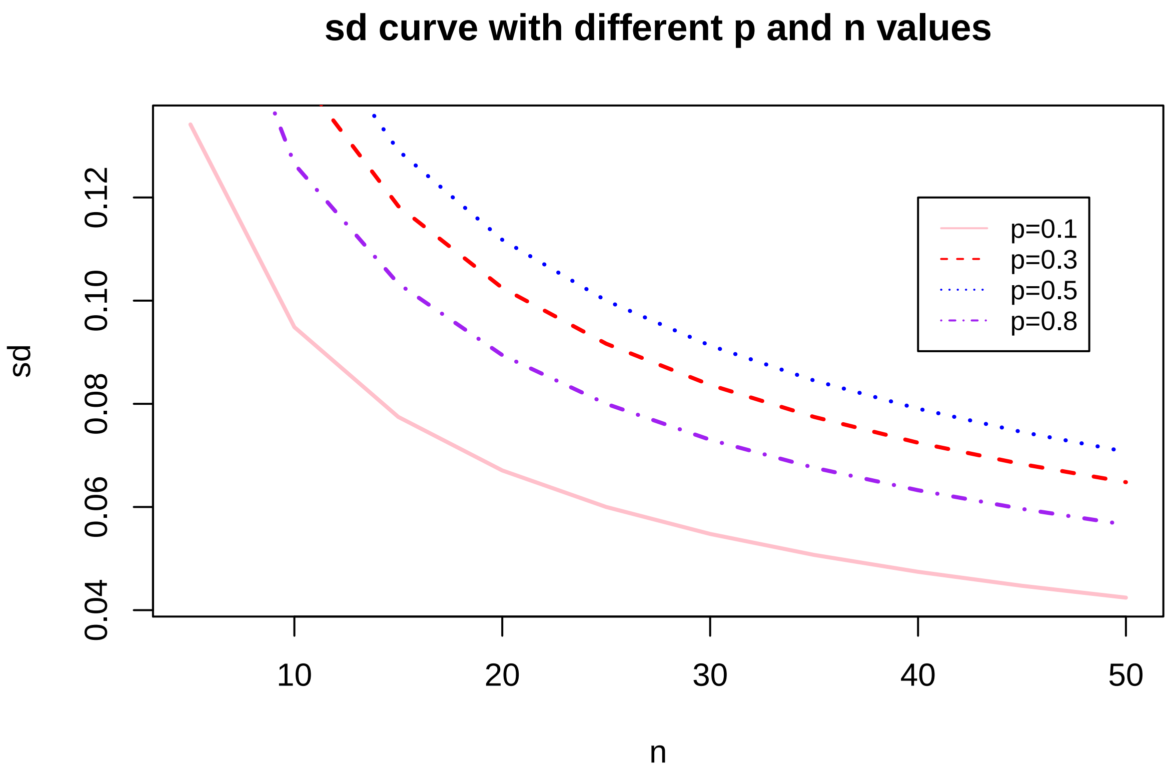
f.

```
p = seq(0,1,0.01)
r = sd_p_wig(10,p)
plot(p,r, main="sd curve with n = 10", xlab="p", ylab="sd", type = 'l')
abline(v = which.max(r)/100, col ="red",lwd=3,lty=2)
```



ii.

```
n = seq(5,50,5)
r1 = sd_p_wig(n,0.1)
r2 = sd_p_wig(n,0.3)
r3 = sd_p_wig(n,0.5)
r4 = sd_p_wig(n,0.8)
plot(n,r1, col="pink", main="sd curve with different p and n values", xlab="n",
     ylab="sd", type =          'l',lwd = 2,lty=1)
lines(n,r2, col="red",lwd = 2, lty = 2)
lines(n,r3, col="blue",lwd = 2, lty = 3)
lines(n,r4, col="purple",lwd = 2, lty = 4)
legend(40, 0.12, legend=c("p=0.1", "p=0.3","p=0.5","p=0.8"),
       col=c("pink", "red","blue","purple"),         lty=1:4, cex=0.8)
```



iii. When p is 0.5, the standard deviation is the highest. When p is smaller, the standard deviation decreases.

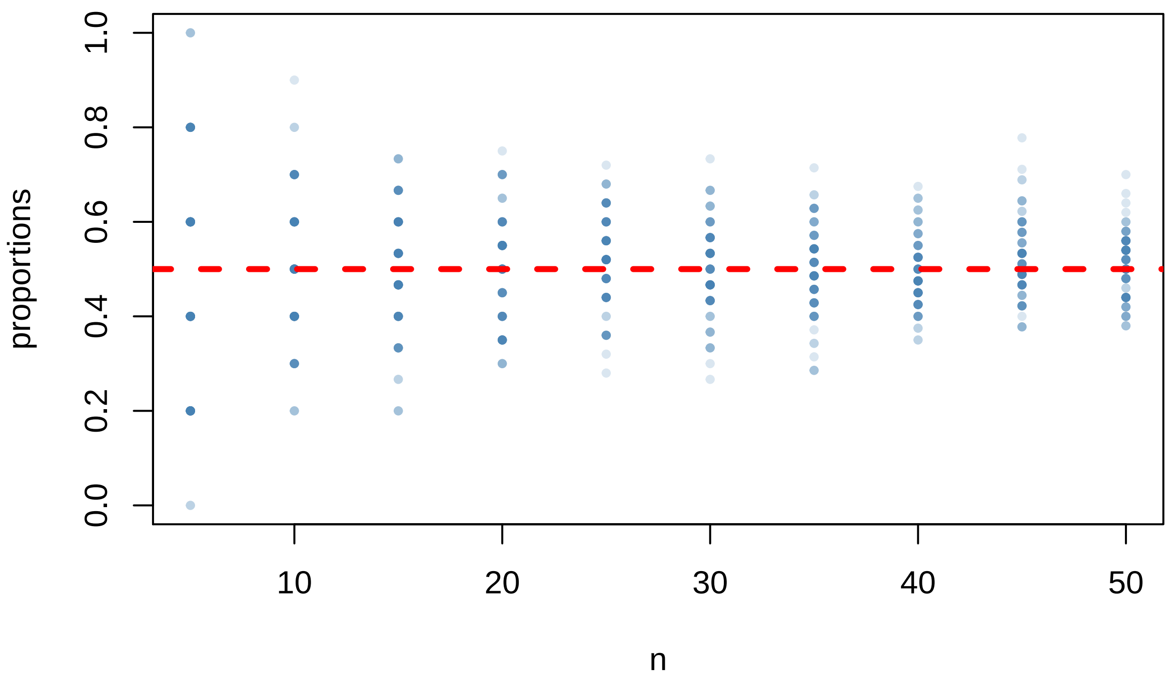
f.

g.

```
n = seq(5,50,5)
n1 = rep(n,each=100)
r = rbinom(length(n1), size=n1, 0.5)
```

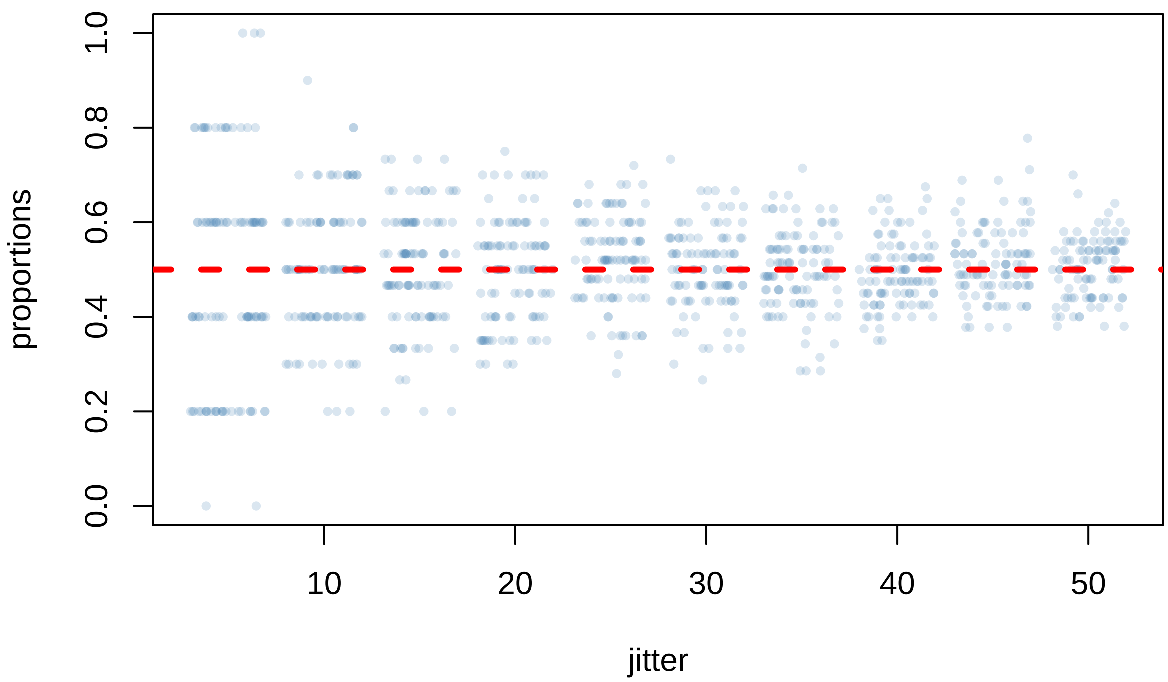
ii.

```
plot(n1,r/n1,xlab = "n",ylab = "proportions",ylim=c(0,1),pch=19,
     col= adjustcolor("steelblue",0.2), cex =0.5)
abline(h = 0.5, col ="red",lwd=3,lty=2)
```



iii.

```
plot(jitter(n1,2),r/n1,xlab = "jitter", ylab = "proportions", ylim=c(0,1),
     pch = 19, col = adjustcolor("steelblue",0.2), cex =0.5)
abline(h = 0.5, col ="red",lwd=3,lty=2)
```



jitter() adds an amount of noise in order to break ties and to see the data more clearly.

iv. As n increases, the standard deviation decreases and the proportions are closer to 0.5.

g.

h. It is hardest to estimate p when $p = 0.5$. Because according to the figures above, the standard deviation is the highest when p is 0.5. The easiest true p to estimate is when $p = 0.1$ because the standard deviation is the lowest. And we do not need to use large sample size.

ii. Law of large numbers (LLN) is a theorem that describes the result of performing the same experiment a large number of times. According to the law, the average of the results obtained from a large number of trials should be close to the expected value, and will tend to become closer to the expected value as more trials are performed. When the sample size n increases, the standard deviation decreases.

iii. There aren't enough small numbers to meet the many demands made of them. In this case, when data is small, the standard deviation is very high. The estimate $\widehat{p}$ is not accurate.