HW4

Joon

April 12, 2018

Q1

defining the equation for variance The expected value equation simply came out to Pa

```
variance <- function(n,p) {
ret <- (1/n) * (p*(1-p))

return(ret)

}
```

1B

Finding the variance e(x) = .5

```
variance(50,.5)
```

```
## [1] 0.005
```

1C

```
 pnorm(0.6,mean =0.5,sd = sqrt(.005), lower.tail = FALSE)
```

```
## [1] 0.0786496
```

1D

We accept the null hypothesis as the probability is greater than our threshold

1E

We reject the null hypothesis as the probability is less than our threshold

```
v2 <- variance(100,.5)

1 - pnorm(.6,.5,sqrt(v2))
```

```
## [1] 0.02275013
```

1F the smallest value

```
qb<-qnorm(.05,mean=.5,sd=.05,lower.tail = FALSE)
qb
```

```
## [1] 0.5822427
```

1G

Outputs the samllest detectable improvement

```r
qb - .5
```

```
## [1] 0.08224268
```

Q2

2a)

expected value is .75 according to the derived mean formula we calculate the variance using the defined formula with n = 50 and p = .75

```r
q2v1 <- variance(50,.75)
q2v1
```

```
## [1] 0.00375
```

2B)

```r
pnorm(0.6,mean =0.75,sd = sqrt(q2v1),lower.tail = FALSE)
```

```
## [1] 0.9928471
```

2C)

We fail to reject the null hypothesis as our result is way above the threshold. It makes sense as the probability the pa click rate is higher than our new click rate is almost guaranteed

2D)

we reject the null hypothesis in this case as increasing the sample size causes our result to be less than our threshold

```r
q2v2 <- variance(100,.75)

1 - pnorm(.6,.5,sqrt(q2v2))
```

```
## [1] 0.01046067
```

2E) smallest pb value

```r
qb2<-qnorm(.05,mean=.75,sd=q2v2,lower.tail = FALSE)

qb2
```

```
## [1] 0.7530841
```

2F)

```r
qb2-.75
```

```
## [1] 0.003084101
```

3

The samllest detectable improvement for question 1 is larger as 0.08224268 > 0.003084101

This makes sense as both the standard deviation and variance for q1 are larger than those of q2 thus improvements would also be larger in magnitude.

4

Q4B)

since the expected value of a bernoulli variable is p the p in our case would be the mean of a and b so 35+35/55+45 = 70/100 = .7

Q4C)

my derived formula ends up being 1/nb - 1/na so it would be 1/45 -1/55

```r
1/45 - 1/55
```

```
## [1] 0.004040404
```

Q4D)

estimate I derived from estimating pb and pa and then subtracting them. pa = mean of a pb = mean of b

```r
est <- 35/45 - 35/55

est
```

```
## [1] 0.1414141
```

Q4E

```r
pnorm(est,mean =0,sd = sqrt(.004),lower.tail = FALSE)
```

```
## [1] 0.0126774
```

Q4 F

Yes we can reject the null hypothesis as our click rate did increase. In that sense it is almost intuitive that we can reject the null hypothesis