ADA HW4

Liangquan Zhou (lz2377)

10/03/2014

Problem 2

From example 2 we know that $X_1, X_2, ...X_n \sim N(\mu_1, \sigma^2)$, $Y_1, Y_2, ...Y_n \sim N(\mu_2, \sigma^2)$, $\sigma = 10$, $\Delta = 4$, $\alpha = 0.05$ and $power = 0.8$

And the test is : $H_0 : |\mu_1 - \mu_2| = \Delta = 0$ vs. $H_1 : \Delta > 0$

a). Want to plot power as a function of sample size. Use normal approach to calculation the relation between power and sample size.

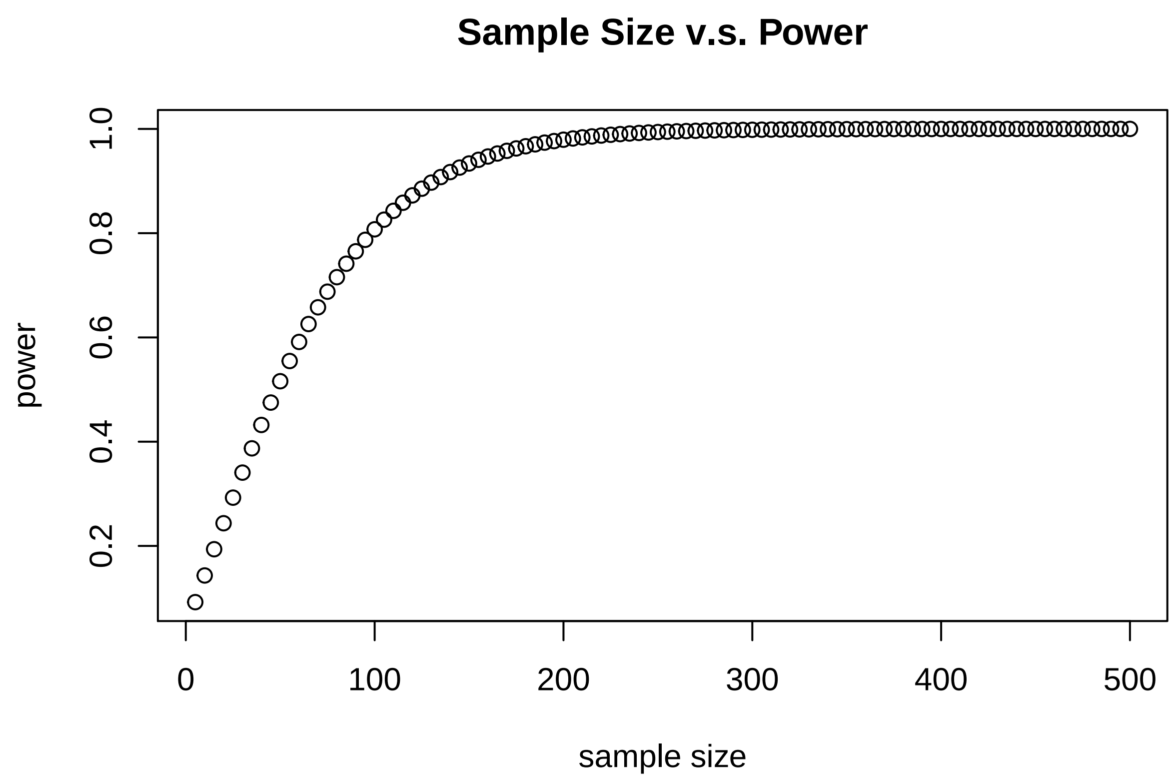
From the slides we know that

$$n = \frac{2(Z_{\alpha/2} + Z_\beta)^2\sigma^2}{\Delta^2}$$

then

$$Z_\beta = \frac{\sqrt{n}\Delta}{\sqrt{2}\sigma} - Z_{\alpha/2}$$

```
power_n = function(sample_size, alpha, delta, sigma){

  z_alpha_2 = qnorm(1 - alpha/2)

  z_beta = (sqrt(sample_size) * delta) / (sqrt(2) * sigma) - z_alpha_2

  power = pnorm(z_beta)

  return(power)
}

# alpha = 0.05, sigma = 10, delta = 4
sample_size = seq(5, 500, by = 5)

power = power_n(sample_size, 0.05, 4, 10)

plot(sample_size, power, type="b", xlab="sample size", ylab="power",
  main = "Sample Size v.s. Power")
```

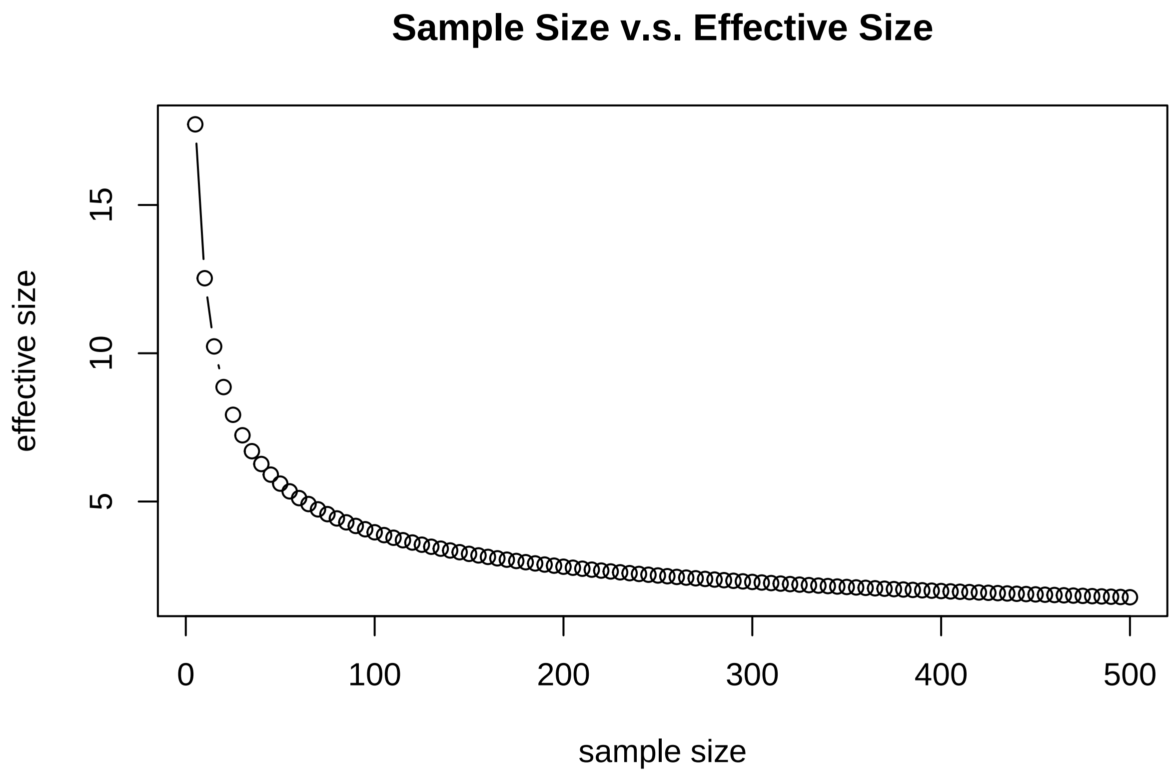


b). For effect size $\Delta = \mu_1 - \mu_2$, also use normal approach, we can get

$$\Delta = \sqrt{\frac{2(Z_{\alpha/2} + Z_\beta)^2\sigma^2}{n}}$$

and we can get the plot as follows:

```
power_e = function(sample_size, alpha, beta, sigma){
  ## calculate delta
  z_alpha_2 = qnorm(1 - alpha/2)

  z_beta = qnorm(1 - beta)

  delta = sqrt(2) * (z_alpha_2 + z_beta) * sigma / sqrt(sample_size)

  return(delta)
}

# power = 0.8, beta = 1 - power = 0.2, alpha = 0.05, sigma = 10
sample_size = seq(5, 500, by = 5)

delta = power_e(sample_size, 0.05, 0.2, 10)

plot(sample_size, delta, type = "b", xlab = "sample size", ylab = "effective size",
  main = "Sample Size v.s. Effective Size")
```



Problem 3

In this problem, $X \sim bin(n, p_1)$, $Y \sim bin(n, p_2)$, and X and Y are independent.

The test is $H_0 : p_1 = p_2$ vs. $H_1 : p_1 \neq p_2$. We want to do the same thing for proportion test.

a). Want to plot the power as a function of sample size.

1. Use normal approximation.

From the slides we know that

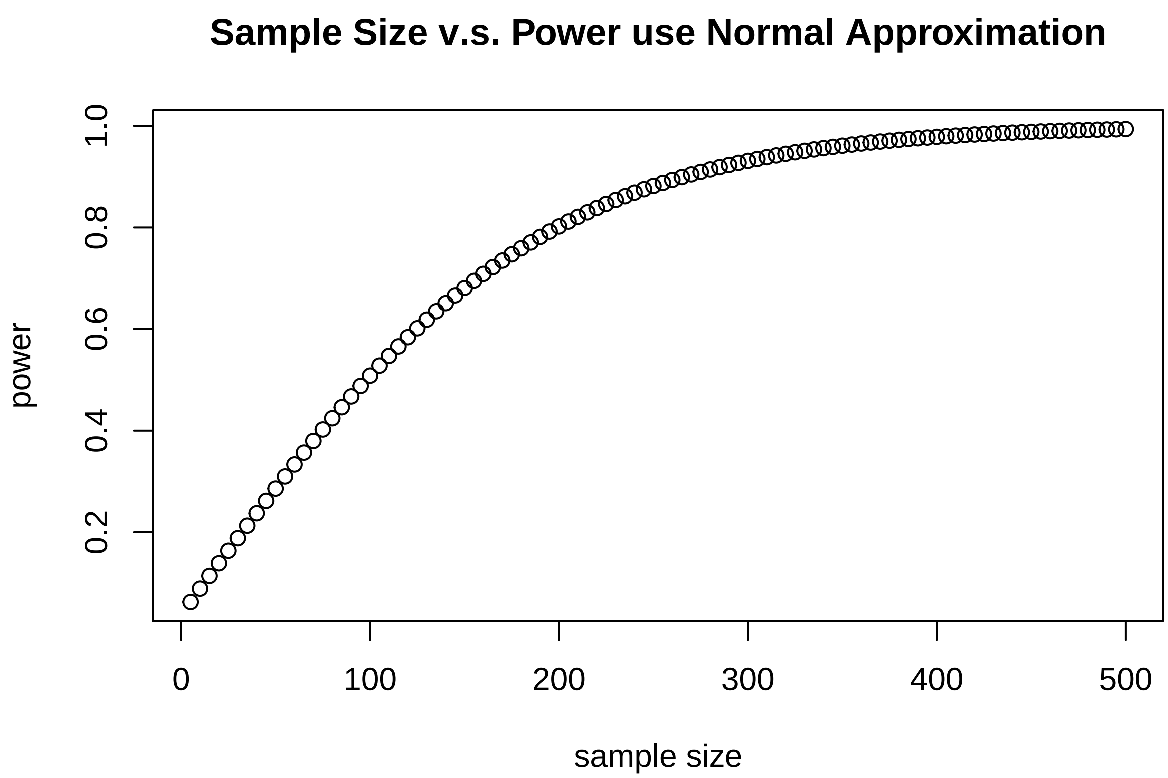
$$n = \frac{[Z_{\alpha/2}\sqrt{2\bar{p}\bar{q}} + Z_\beta\sqrt{p_1q_1 + p_2q_2}]^2}{(p_2 - p_1)^2}$$

and we can solve that

$$Z_\beta = \frac{\sqrt{n(p_2 - p_1)^2} - Z_{\alpha/2}\sqrt{2\bar{p}\bar{q}}}{\sqrt{p_1q_1 + p_2q_2}}$$

```
power_normal1 = function(sample_size, alpha, p1, p2){
  q1 = 1- p1

  q2 = 1- p2

  p_mean = (p1 + p2) / 2

  q_mean = 1 - p_mean
```

```
  z_alpha_2 = qnorm(1 - alpha/2)

  z_beta = (sqrt(sample_size * (p2 - p1)^2) - z_alpha_2 * sqrt(2 * p_mean *
      q_mean)) / sqrt(p1 * q1 + p2 * q2)

  power = pnorm(z_beta)

  return(power)
}

# alpha = 0.05, p1 = 0.8, p2 = 0.9
sample_size = seq(5, 500, by = 5)

power = power_normal1(sample_size, 0.05, 0.8, 0.9)

plot(sample_size, power, type="b", xlab="sample size", ylab="power", main =
    "Sample Size v.s. Power use Normal Approximation")
```



2. Use `arcsin transformation`

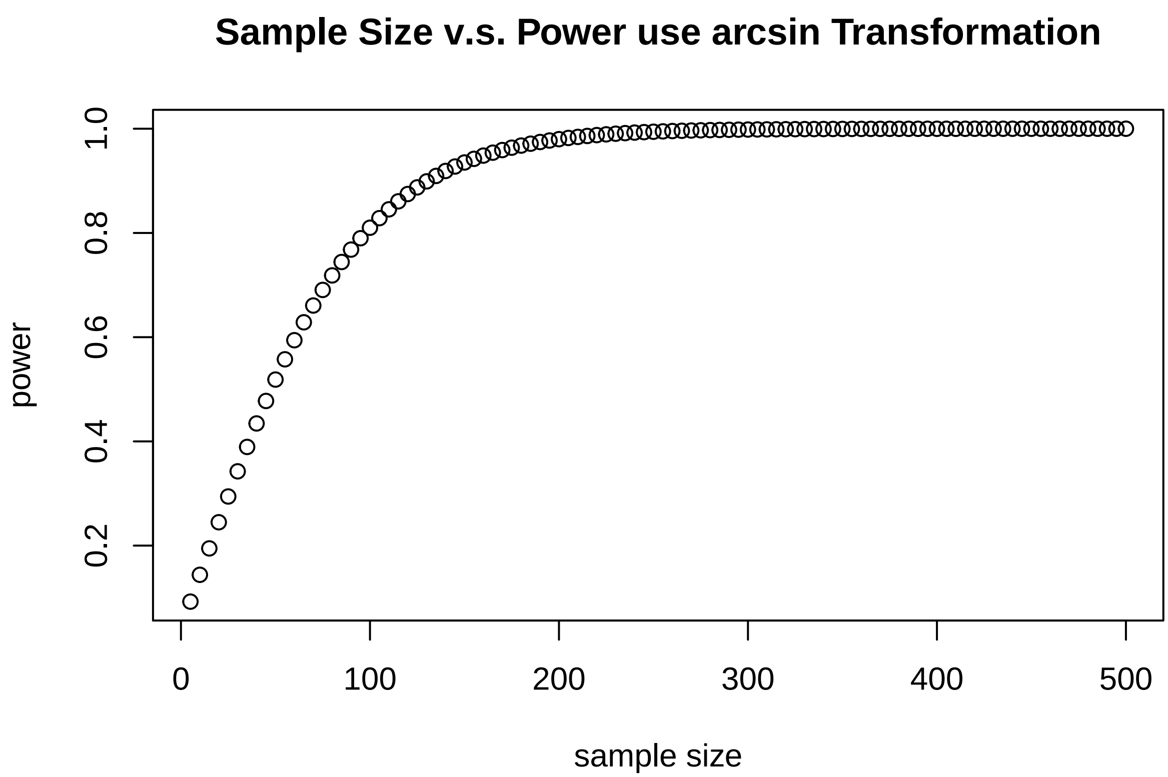
From the slides we know that

$$n = \frac{(Z_{\alpha/2} + Z_\beta)^2}{\Delta^2}$$

$\Delta = f(p1) - f(p2)$, then

$$Z_\beta = \sqrt{n}\Delta - Z_{\alpha/2}$$

```r
f = function(x) 2 * asin(sqrt(x))

power_arcsin1 = function(sample_size, alpha, p1, p2){

  delta = abs(f(p1) - f(p2))

  z_alpha_2 = qnorm(1 - alpha/2)

  z_beta = sqrt(sample_size) * delta - z_alpha_2

  power = pnorm(z_beta)

  return(power)
}

# alpha = 0.05, p1 = 0.8, p2 = 0.9
sample_size = seq(5, 500, by = 5)

power = power_arcsin1(sample_size, 0.05, 0.8, 0.9)
plot(sample_size, power, type = "b", xlab = "sample size", ylab = "power",
  main = "Sample Size v.s. Power use arcsin Transformation")
```



b). Want to plot sample size as a function of effective size.

1. Use normal approximation.

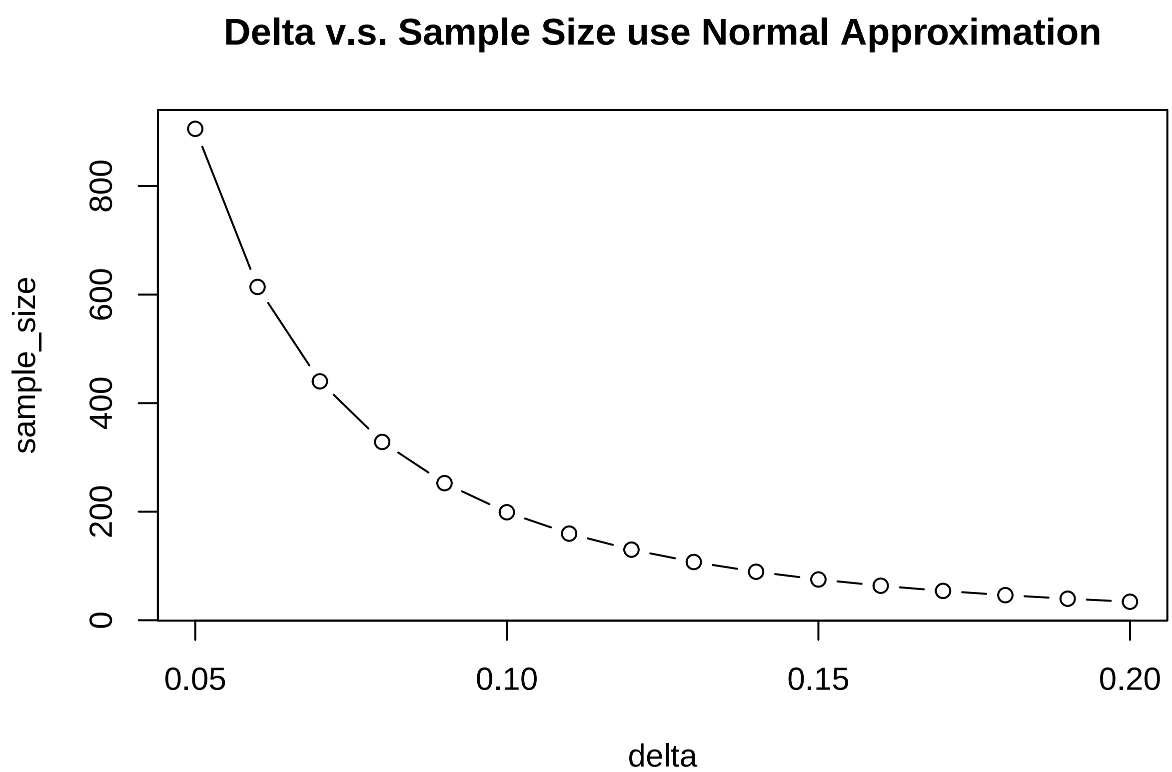
Fix p_1 and let the Δ change. (Because $\Delta = p_2 - p_1$, we can let p_2 change.)

Use

$$n = \frac{[Z_{\alpha/2}\sqrt{2\bar{p}\bar{q}} + Z_{\beta}\sqrt{p_1 q_1 + p_2 q_2}]^2}{(p_2 - p_1)^2}$$

to calculate the sample size.

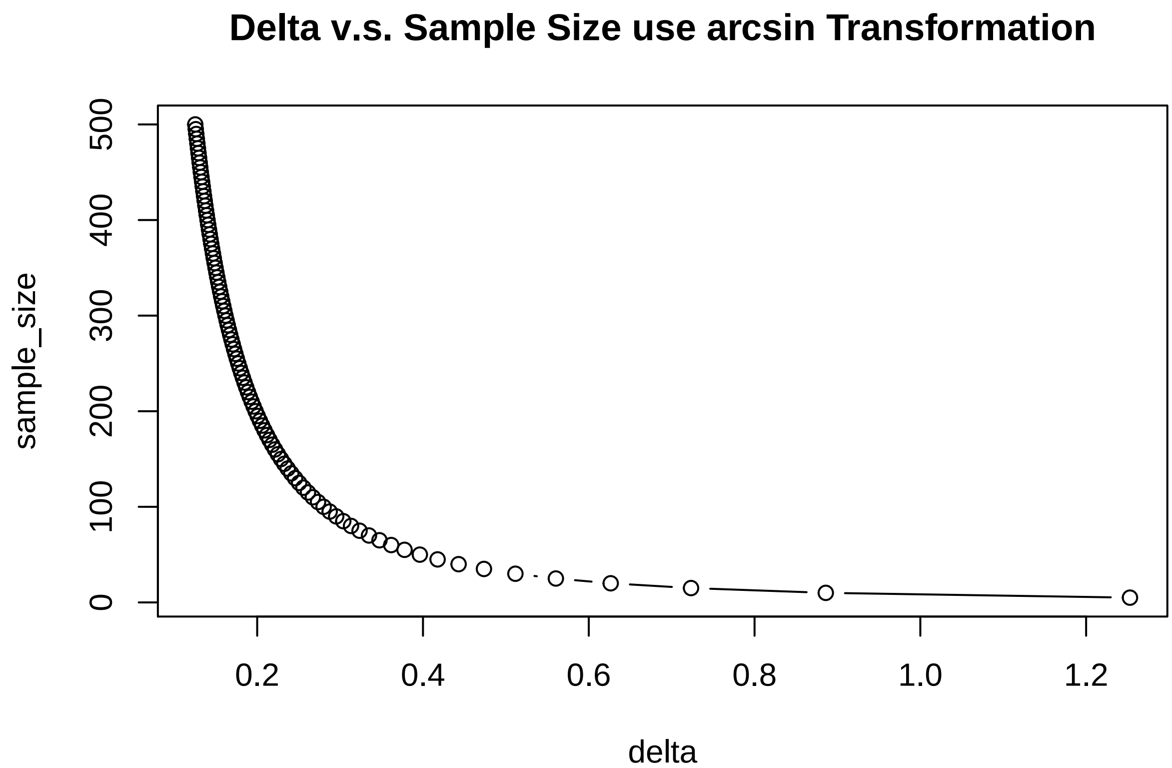
```
# fix p1 and let delta change, calculate sample size
power_normal2 = function(p1, p2, alpha, beta){

  q1 = 1- p1

  q2 = 1- p2

  delta = abs(p1 - p2)

  p_mean = (p1 + p2) / 2

  q_mean = 1 - p_mean

  z_alpha_2 = qnorm(1 - alpha/2)

  z_beta = qnorm(1- beta)

  sample_size = (z_alpha_2 * sqrt(2*p_mean*q_mean)+z_beta*sqrt(p1*q1+p2*q2))^2/(p2-p1)^2

  return(sample_size)
}

# set p2 > p1
p2 = seq(0.85, 1, by = 0.01)
p1 = 0.8
delta = p2 - p1
sample_size = power_normal2(p1, p2, 0.05, 1 - 0.8)

plot(delta, sample_size, type="b", main = "Delta v.s. Sample Size use Normal Approximation")
```



2. Use `arcsin transformation` We can get

$$\Delta = \frac{(Z_{\alpha/2} + Z_{\beta})^2}{\sqrt{n}}$$

```
f = function(x) 2 * asin(sqrt(x))

delta_arcsin = function(sample_size, alpha, beta){

  z_alpha_2 = qnorm(1 - alpha/2)

  z_beta = qnorm(1- beta)

  delta = (z_alpha_2 + z_beta)/sqrt(sample_size)

  return(delta)
}

# alpha = 0.05, beta = 1- power = 0.2
sample_size = seq(5, 500, by = 5)

delta = delta_arcsin(sample_size, 0.05, 0.2)
plot(delta,sample_size, type = "b", main = "Delta v.s. Sample Size use arcsin Transformation")
```


Delta v.s. Sample Size use arcsin Transformation
sample_size
0
100
200
300
400
500
0.2
0.4
0.6
0.8
1.0
1.2
delta