Shaikh Shawon Arefin Shimon 830201247
STAT511 HW 8

1) n=22, $\pi = 0.72$

A) Mean μ_Y = n π = 22*0.72 = 15.84
SD ,$\sigma_Y = \sqrt{n\,\pi(1-\pi)} = \sqrt{22 * 0.72(1-0.72)}$ = 2.105991

B) P(Y ≤ 16) = 0.6100871

C) P(Y < 16) = 0.422552

D) P(16 ≤ Y < 18) = 0.1701999

E) P(Y = 18) = 0.1215714

F) P(Y ≥ 18) = 0.2197129

G) without continuity correction,
P(Y≥18) = 1 – P(Y<18) = 1 - P(Y≤17) = 1-pnorm(17,mean=15.84,sd=2.105991,lower.tail = TRUE)
so, without continuity correction , P(Y≥18) = 0.2908821

H) with continuity correction,

P(Y≥18) = 1 – P(Y<18) = 1 - P(Y≤17.5) = 1-pnorm(17.5,mean=15.84,sd=2.105991,lower.tail = TRUE)
so, without continuity correction , P(Y≥18) = 0.2152818

2) n = #of students = 75 (trial)
y = #of STEM majors = 51 (success)

A) $Sample\ proportion, \hat{\pi} = \frac{y}{n} = \frac{51}{75} = 0.68$ [Estimate of STEM population proportion]

Estimate of standard deviation, SE($\hat{\pi}$) = $\sqrt{\frac{\hat{\pi}\ (1-\hat{\pi})}{n}}$ = = $\sqrt{\frac{(0.68)*(1-0.68)}{75}}$ = 0.05386403

B) 95% confidence interval for true population proportion π = $\hat{\pi} \pm Z_{\alpha/2} * \mathrm{SE}(\pi) = 0.68 \pm (1.96 * 0.05386403) = (0.5744265, 0.7855735)$

C) $H_0 : \pi \le 0.5$
$H_A : \pi > 0.5$

skshimon@cs.colostate.edu

$\pi_0 = 0.5$, $\hat{\pi} = 0.68$,

Test statistics z = $\frac{\hat{\pi}-\pi_0}{\sqrt{\frac{\pi_0(1-\pi_0)}{n}}} = \frac{0.68-0.5}{\sqrt{\frac{0.5*0.5}{75}}} = 3.117691$

Rejection region, z$\geq z_\alpha$, or z z≥ 1.645

In this case, $3.117691 \geq 1.645$, so we can reject the null hypothesis with 95% confidence, which means we can conclude (with 95% certainty) that the population proportion of STEM majors is greater than 0.5.

From running prop.test, p-value = 0.00091114. (Ans)

```
        1-sample proportions test without continuity correction

data:  51 out of 75, null probability 0.5
X-squared = 9.72, df = 1, p-value = 0.0009114
alternative hypothesis: true p is greater than 0.5
95 percent confidence interval:
 0.5864651 1.0000000
sample estimates:
   p 
0.68
```

3) n = #of items = 60 (trial)

y = #of defective items = 10 (success)

A) $Sample\ proportion, \hat{\pi} = \frac{y}{n} = \frac{10}{60} = 0.1666667$ [Estimate of population proportion]

B) α = 0.1, 90% confidence interval = (0.0958273,0.2691848)

```
        1-sample proportions test with continuity correction

data:  10 out of 60, null probability 0.5
X-squared = 25.35, df = 1, p-value = 4.782e-07
alternative hypothesis: true p is not equal to 0.5
90 percent confidence interval:
 0.0958273 0.2691848
sample estimates:
        p 
0.1666667
```

C) 90% confidence interval = (0.09330693 0.26629080)

```
Exact binomial test

data:  10 and 60
number of successes = 10, number of trials = 60, p-value =
1.616e-07
alternative hypothesis: true probability of success is not equal to 0.5
90 percent confidence interval:
 0.09330693 0.26629080
sample estimates:
probability of success 
             0.1666667
```

Shaikh Shawon Arefin Shimon 830201247

STAT511 HW 8

4) A)

i) n=9

```
power.anova.test(groups = 4, between.var = between_var_i, within.var = within
_Var, sig.level = 0.05,power=0.9)

     Balanced one-way analysis of variance power calculation 

         groups = 4
              n = 8.139055
    between.var = 66.66667
     within.var = 100
      sig.level = 0.05
          power = 0.9

NOTE: n is number in each group
```

ii) n= 16

```
power.anova.test(groups = 4, between.var = between_var_ii, within.var = withi
n_Var, sig.level = 0.05,power=0.9)

     Balanced one-way analysis of variance power calculation 

         groups = 4
              n = 15.18834
    between.var = 33.33333
     within.var = 100
      sig.level = 0.05
          power = 0.9

NOTE: n is number in each group
```

iii) n= 11

```
power.anova.test(groups = 4, between.var = between_var_iii, within.var = with
in_Var, sig.level = 0.05,power=0.9)

     Balanced one-way analysis of variance power calculation 

         groups = 4
              n = 10.48319
    between.var = 50
     within.var = 100
      sig.level = 0.05
          power = 0.9

NOTE: n is number in each group
```

ii) scenario ii requires the largest sample size.

We know that power increases as sample size and differences among the true means increase, and decrease as error standard deviation decreases.

skshimon@cs.colostate.edu

This means, that if the power stays the same, increasing differences among true means will reduce the sample size. And decreasing differences among true means will increase the sample size.
Of the given three scenario, scenario ii has the lowest differences among true means (between variance = 33.33), which results in the largest sample size.