Exercise 6A.1

For a random sample of 800 males, 35 were observed to have a particular type of cyst. Calculate a 95% confidence interval for the proportion of males in the population who have the cyst.

$n = 800, \hat{p} = \dfrac{35}{800} = 0.04375$

CI: $\hat{p} \pm z^* \sqrt{\dfrac{\hat{p}(1-\hat{p})}{n}}$

Critical value from standard normal with 0.05/2=0.025 in upper tail $\rightarrow z^* = 1.96$

$0.04375 \pm 1.96 \sqrt{\dfrac{0.04375(1-0.04375)}{800}}$

$0.04375 \pm 1.96 \times 0.00723$

$(0.030, 0.058)$

Exercise 6A.2

Suppose that out of 10,000 50- to 54-year-old women sampled whose mother had breast cancer, 400 had breast cancer at some time in their lives. What is the 99% confidence interval for the incidence rate of breast cancer in such a population?

$n = 10000, \hat{p} = \frac{400}{10000} = 0.04$

CI: $\hat{p} \pm z^* \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$

Critical value from standard normal with 0.01/2=0.005 in upper tail $\rightarrow z^* = 2.58$

$0.04 \pm 2.58 \sqrt{\frac{0.04(1-0.04)}{10000}}$

$0.04 \pm 2.58 \times 0.001960$

$(0.035, 0.045)$

Exercise 6A.3

In the University of New Hampshire student survey, students reported their GPA. We are interested in the proportion of students who have a GPA lower than 2.5. Summary of the data are in the Stata output on the next page.

(a) What is the estimate of the desired proportion?

$$\hat{p} = \frac{47}{243} = 0.1934$$

(b) Use this information to test whether the proportion of all students (the population) is equal to 25%.

$H_0 : p = 0.25$ $\quad$ $p =$ *proportion of students at UNH who have a GPA below 2.5*

$H_a : p \neq 0.25$

$$z = \frac{\hat{p} - p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}} = \frac{0.1934 - 0.25}{\sqrt{\frac{0.25(1-0.25)}{243}}} = \frac{-0.0566}{0.02778} = -2.04$$

Under H_0, $Z \sim N(0, 1)$

p-value $= 2 \times P(Z > |-2.04|) = 2 \times 0.0207 = 0.0414$

$0.0414 < 0.05$, *Reject* H_0

We have evidence that the proportion of students with GPA below 2.5 is significantly different from 25%.

(For problem on previous page)

```
. gen lowgpa = 0 if gpa > 0

. replace lowgpa = 1 if gpa < 2.5
(47 real changes made)

. tab lowgpa

     lowgpa |      Freq.     Percent        Cum.
------------+-----------------------------------
          0 |        196       80.66       80.66
          1 |         47       19.34      100.00
------------+-----------------------------------
      Total |        243      100.00
```

Exercise 6A.4

In the University of New Hampshire student survey we have information on each student's sex. Use the Stata output below to test whether the proportion of females is significantly different from 50%.

```
. tab gender

     Gender |
     (male) |      Freq.     Percent        Cum.
------------+-----------------------------------
     Female |        134       55.14       55.14
       Male |        109       44.86      100.00
------------+-----------------------------------
      Total |        243      100.00
```

$$\hat{p} = \frac{134}{243} = 0.5514$$

$H_0 : p = 0.5$ $\quad$ $p =$ *proportion of students at UNH who are female*
$H_a : p \neq 0.5$

$$z = \frac{\hat{p} - p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}} = \frac{0.5514 - 0.5}{\sqrt{\frac{0.5(1-0.5)}{243}}} = \frac{0.0514}{0.0321} = 1.60$$

Under H_0, $Z \sim N(0, 1)$
p-value $= 2 \times P(Z > |1.60|) = 2 \times 0.0548 = 0.1096$
$0.1096 > 0.05$, *Fail to reject* H_0
There is no evidence that the proportion of female students in the population is significantly different from 50%.