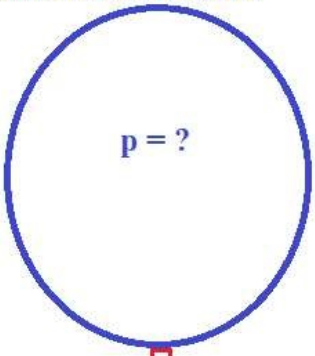
Topic 17

Hypothesis test for a population proportion based on a sample proportion.

Start with a population with an unknown proportion of some characteristic.

By the way, the population has around 50,000 items in it.



We have a null hypothesis that the proportion of items in the population with the desired characteristic is 0.23. The alternative hypothesis is that the proportion is less than 0.23. Thus, H_0: $p = 0.23$ and H_1: $p < 0.23$.

We want to test that null hypothesis at the 0.02 level of significance.

We take a random sample of size 147 and we compute the proportion of the sample that has the characteristic.

n=147 x=22 $\hat{p}$=22/147=0.14966

The first thing to do is to be sure that we can use the normal approximation to the distribution of sample proportions. Are we sampling less than 5% of the population? 0.05*4000 = 200 and 147<200. Second, is $n*p \geq 10$ and $n*(1-p) \geq 10$? 147*0.23=33.81 and 147*(1-0.23)=113.19. OK

Then the standard deviation of the normal distribution of sample proportions will be $\sqrt{p*(1-p)/n} = \sqrt{0.23*(1-0.23)/147} \approx 0.0347$. So, under the null hypothesis the sample proportions will be N(0.23, 0.0347).

Critical value approach: This is a one-tailed test. Find the value that has 0.02 as the area to its left for a N(0.23,0.0347) distribution.

```
qnorm( 0.02, mean=0.23, sd=0.0347)
```

gives the critical low value of 0.1587.

We look at our sample proportion, $\hat{p}$=22/147≈0.14966, and that is lower than our critical low value, 0.1587, so we reject H_0 at the 0.02 level of significance in favor of the alternative H_1: $p < 0.23$.

Attained significance approach: We know that the sample proportion is $\hat{p}$=22/147. If H0 is true, i.e., if p=0.23, then how "strange" would it be to get a sample proportion, in a sample of size 147, that is as strange or more strange than our sample proportion? We can use our normal approximation to answer this via

```
pnorm( 22/147, mean=0.23, sd=sqrt( 0.23*(1-0.23)/147 ) )
```

and that give us the value 0.01031654 as our attained significance. **Because our attained significance is less than the specified level of significance we reject H_0 at the 0.02 level in favor of the alternative H_1: $p < 0.23$.**

Or we could use the supplied function hypoth_test_prop to do both kinds of analysis in one step.

```
55 ##############################
56 #  Or we could just use our hypoth_test_prop() function
57 #    to compute both approaches
58 source("../hypo_prop.R")
59 #      we will use -3 to indicate that H1 is "<"
60 hypoth_test_prop( 0.23, 22, 147, -3, 0.02)
```

```
> ##############################
> #  Or we could just use our hypoth_test_prop() function
> #    to compute both approaches
> source("../hypo_prop.R")
> #      we will use -3 to indicate that H1 is "<"
> hypoth_test_prop( 0.23, 22, 147, -3, 0.02)
                 H0_p                   H1                    x
              "0.23"       "prop <  0.23"                 "22"
                    n            sig level         s.d. of prop
                "147"              "0.02"  "0.034709680274556"
              z-score           crit low            crit_high
 "2.05374891063182"  "0.158715031947752"               "n.a."
           samp prop                    z             attained
"0.149659863945578"  "-2.31463198217113" "0.0103165405232536"
             decision
             "Reject"
```