

# Inference for one proportion (using SE)

Stat 145

## Standard Error for a proportion

We have seen that we can compute p-values and confidence intervals using a randomization or bootstrap distribution. We have also noted that the bootstrap distributions often have a similar shape that appears to be approximately normal. This observation provides us another way to compute p-values and confidence intervals: If there is a way to know that the randomization distribution or bootstrap distribution is approximately normal and to determine the mean and standard deviation (called the standard error in this context) of that normal distribution, then we can use facts about the normal distribution to obtain p-values and confidence intervals.

We have already talked about where randomization distributions and bootstrap distributions will be centered, so it is the standard error that is the important missing piece. In many common situations, statisticians have worked out formulas for the standard errors. Here is our first example: The standard error for a sample proportion is

$$SE = \sqrt{\frac{p(1-p)}{n}}$$

where  $p$  is the actual proportion in the population. Unfortunately, we don't ever know  $p$ . So we need to approximate that too. Use what you know to fill in the blanks below.

- For a hypothesis test, we will use  $p \approx$  \_\_\_\_\_.

(Hint: What assumption is made when computing a p-value?)

- For a confidence interval, we will use  $p \approx$  \_\_\_\_\_.

(Hint: Why can't we do the same thing here that we do for p-values?)

## Is the approximation good enough?

This SE formula is an approximation (because we don't know the true value of  $p$ ). Furthermore, the normal distribution is also only approximately the correct shape. These approximations are better

- if  $n$  is larger, and
- if  $p$  is closer to 0.5.

Our rule of thumb will be that the SE methods for a proportion are good enough to use if

- $n\hat{p}$  and  $n(1-\hat{p})$  are both at least 10 [CI situation]
- $np_0$  and  $n(1-p_0)$  are both at least 10 [HT situation]

*IMS* calls this the **success-failure** condition because it can be interpreted as saying we should expect at least 10 "successes" and at least 10 "failures" (i.e., at least 10 of each of our two outcomes.)

## p-values and confidence intervals

- Traditional test statistic is the z-score for the sample proportion:  $z = \frac{\hat{p} - p_0}{SE}$
- Use `pnorm()` to convert  $\hat{p}$  or  $z$  into a p-value
  - for  $\hat{p}$ 
    - \* mean =  $p_0$
    - \* standard deviation =  $SE$
  - for  $z$ ,
    - \* mean = 0
    - \* standard deviation = 1
- Confidence interval: estimate  $\pm$  (critical value)  $\cdot$  SE
  - estimate:  $\hat{p}$
  - critical value
    - \* 1.96  $\approx$  2 for a 95% confidence interval
    - \* Find it using `qnorm()` for other confidence levels. Remember that `qnorm()` works with below, not between. So you will need to convert your question into “below language”.

### Example (From IMS 6.1.2)

A simple random sample of 826 payday loan borrowers was surveyed to better understand their interests around regulation and costs. 578 of the responses supported new regulations on payday lenders.

Before proceeding to compute a confidence interval for the proportion of payday loan borrowers who support new regulations, let make sure the normal approximation is good enough:

```
phat <- 578/826; phat
```

```
## [1] 0.6997579
```

```
phat * 826
```

```
## [1] 578
```

```
(1 - phat) * 826
```

```
## [1] 248
```

Both checks are well above 10.

Now let's estimate the standard error and use it to create a 95% confidence interval.

```
SE <- sqrt(phat * (1 - phat) / 826); SE
```

```
## [1] 0.01906539
```

```
ME <- 1.96 * SE; ME
```

```
## [1] 0.03736817
```

```
c(phat - ME, phat + ME) # c() combines the two ends so we can display them together
```

```
## [1] 0.6623897 0.7371260
```

So our confidence interval is 66 to 73%. Since 50% is not in this interval, we know we have enough evidence (at the  $\alpha = 0.50$ ) level to conclude that a majority of payday loan holders are in favor of new regulations. In fact, 50% is quite far outside the interval, so the p-value should be quite small.

Let's compute the p-value to see how small it is. Note the slightly different SE now.

```
# standard error uses 0.5 from null hypothesis
SE <- sqrt(0.5 * 0.5 / 826)
# standardized test statistic
z <- (phat - 0.5) / SE; z
```

```
## [1] 11.48217
```

```
# p-value -- two ways
2 * (1 - pnorm(z))
```

```
## [1] 0
```

```
2 * (1 - pnorm(phat, 0.5, SE))
```

```
## [1] 0
```

```
# picture
xpnorm(phat, 0.5, SE)
```

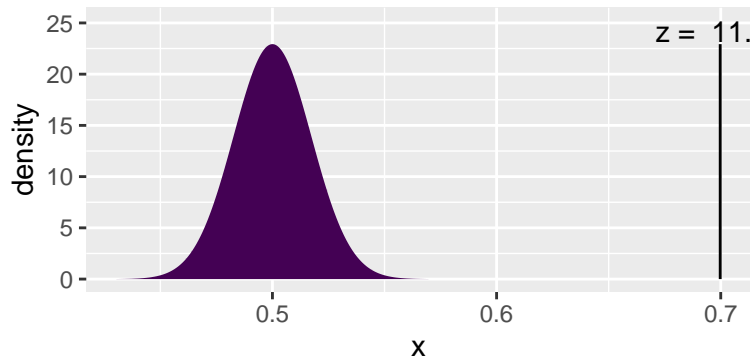
```
##
```

```
## If  $X \sim N(0.5, 0.0174)$ , then
```

```
##  $P(X \leq 0.6998) = P(Z \leq 11.48) = 1$ 
```

```
##  $P(X > 0.6998) = P(Z > 11.48) = 0$ 
```

```
##
```



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

The p-value isn't really 0, but it is incredibly small. Here's a trick to see how small it is that takes advantage of the symmetry of the normal distribution.

```
2 * pnorm(-z)
```

```
## [1] 1.621616e-30
```

```
# picture
xpnorm(-z)
```

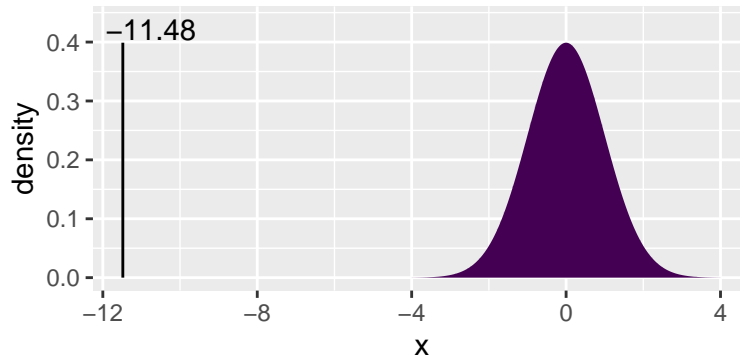
```
##
```

```
## If  $X \sim N(0, 1)$ , then
```

```
##  $P(X \leq -11.48) = P(Z \leq -11.48) = 8.108e-31$ 
```

```
##  $P(X > -11.48) = P(Z > -11.48) = 1$ 
```

```
##
```



```
## [1] 8.108079e-31
```

(R is able to work more accurately with numbers close to 0. This trick avoids starting with a number near 1 that computers don't store as precisely.)

### Some Practice

In each situation (a) check that the approximation is good enough to use, (b) compute the standard error and (c) compute the requested confidence interval or p-value

1. 95% confidence interval for germination rate if 257 out of 325 seeds germinate.
2. P-value for  $H_0 : p = 0.5$  vs  $H_a : p \neq 0.5$  in the situation above.
3. 90% confidence interval for proportion of people who choose "Rock" in the first round of Rock-Paper-Scissors based on data in which 66 of 119 people chose rock.
4. P-value for  $H_0 : p = 1/3$  vs  $H_a : p \neq 1/3$  in the situation above. (Why are we testing a proportion of  $1/3$ ?)

### Hugo

Here's an example based on the game *Mitternachtsparty* which uses a die with a ghost (Hugo) on it. Let's test

- $H_0 : p = 1/6$
- $H_a : p > 1/6$

Data: 16 Hugos in 50 tosses of the die.

5. Show that this situation does not meet our success-failure rule of thumb.
6. Do the test using randomization.
7. Do it again using SE (even though we shouldn't) to see how good/bad that approximation is.

### What makes SE large?

For a fixed sample size, what value of  $\hat{p}$  or  $p_0$  makes the standard error largest? (Feel free to try it with  $n = 100$  as an example.)

### Comparing with Randomization and Bootstrap methods

Go back and redo some of the previous problems using our simulation based methods to see how the results compare.

## Sample Size

One advantage to the formula approach is that we can use the formula to help plan a study. Suppose you want to estimate the portion of the earth that is covered by water. How large must your sample be?

To answer this question, we need to start with 3 numbers:

- Desired margin of error: \_\_\_\_\_
- Desired confidence level: \_\_\_\_\_
- Rough estimate for actual proportion: \_\_\_\_\_
  - When in doubt, estimate closer to 0.5 since the margin of error is larger when the proportion is nearer 0.5.

Given our answers above, we can proceed as follows:

- a. Is a sample of size 100 large enough? 1000? 10,000? [Determine the margin of error in each situation to see if it is small enough.]
- b. Get a more precise estimate either by trying more numbers (think about the higher/lower game) or by solving algebraically.

## Examples

1. If you want to estimate the proportion of people who will vote for candidate A with a margin of error of  $\pm 4\%$ , how large must your sample be?
2. If you want to estimate the proportion of people who will vote for candidate A with a margin of error of  $\pm 2\%$ , how large must your sample be?
3. Compare your two answers above. To make the margin of error half the size, how much larger did we have to make our sample size? Use this to see how large the sample would need to be to get a margin of error of  $\pm 1\%$ , then check your work using our original method.
4. Researchers want to estimate the positivity rate of COVID tests given to asymptomatic people. How large a sample must they obtain to estimate with a margin of error of  $\pm 1\%$ ? Use the fact that the researchers are quite confident that the positivity rate is less than 10%.
  - a. What would happen if the positivity rate was actually quite a bit higher than 10%?
  - b. What would happen if the positivity rate was actually quite a bit smaller than 10%?
5. The unemployment rate is usually under 10% and is reported with a margin of error of  $\pm 0.2\%$  each month. How many people must the government survey each month to get this level of accuracy for the unemployment rate? How many fewer people does it take if the unemployment rate is expected to be under 5%? (In practice, of course, they adjust the margin of error each month based on their sample size and the sample proportion, but they design the study to have a sample size that is likely to give them roughly the margin of error they are looking for.)