Math 3081 (Probability and Statistics)

Lecture #17 of 27 $\sim$ August 3rd, 2021

Interval Estimation (Part 2)

- Binomial Confidence Intervals
- Applications of Confidence Intervals

This material represents §3.2.3 from the course notes, and problems 13-20 from WeBWorK 5.

Antecedently, I

Last lecture, we introduced the notion of a confidence interval:

Definition

If X is a random variable and $0 < \alpha < 1$, a $100(1-\alpha)\%$ confidence interval for X is an interval (a, b) with $a < X < b$ such that $P(a < X < b) = 1 - \alpha$.

When θ is an unknown parameter, we interpret a confidence interval for θ as giving us a reasonable error range on the estimation $\hat{\theta}$ for θ that we have computed.

Antecedently, II

We also constructed confidence intervals using the normal distribution:

Proposition (Normal Confidence Intervals)

A $100(1-\alpha)\%$ confidence interval for the unknown mean μ of a normal distribution with known standard deviation σ is given by $\hat{\mu} \pm c\dfrac{\sigma}{\sqrt{n}} = (\hat{\mu} - c\dfrac{\sigma}{\sqrt{n}}, \hat{\mu} + c\dfrac{\sigma}{\sqrt{n}})$ where n sample points $x_1, \dots, x_n$ are taken from the distribution, $\hat{\mu} = \frac{1}{n}(x_1 + \cdots + x_n)$ is the sample mean, and c is the constant satisfying $P(-c < N_{0,1} < c) = 1-\alpha$.

Here are various pairs (α, c) where $P(-c < N_{0,1} < c) = 1-\alpha$:

$1-\alpha$	50%	80%	90%	95%	98%	99%	99.5%	99.9%
c	0.6745	1.2816	1.6449	1.9600	2.3263	2.5758	2.8070	3.2905

Binomial Confidence Intervals, I

Our goal today is to extend our discussion of confidence intervals to situations that are well approximated by the normal distribution.

Specifically, we are interested in using the normal approximation to the binomial distribution, which arises from repeated sampling of a Bernoulli random variable.

Binomial Confidence Intervals, II

So, suppose that we have a Bernoulli random variable with unknown success probability p that we sample n times, yielding sample values $x_1, x_2, \ldots, x_n$ with a total number of successes equal to $k = x_1 + x_2 + \cdots + x_n$.

- As we have shown, the sample success estimator $\hat{p} = k/n$ is unbiased and is the most efficient possible unbiased estimator of the true success probability p.
- Furthermore, the sample estimator $n\hat{p}$ (which counts the total number of successes in the n samples) will be binomially distributed with mean np and standard deviation $\sqrt{np(1-p)}$.

Binomial Confidence Intervals, III

To compute an exact confidence interval, we would need to determine the precise nature of the relationship between $\hat{p} = k/n$ and the parameter p itself, which is quite difficult to do directly.

- However, when np and $n(1-p)$ are both reasonably large, the binomial distribution will be well approximated by the corresponding normal distribution.
- That means $n\hat{p}$ will have an approximately normal distribution with mean np and standard deviation $\sqrt{np(1-p)}$.
- Equivalently, this says $\hat{p}$ will have an approximately normal distribution with mean p and standard deviation $\sqrt{p(1-p)/n}$.

Binomial Confidence Intervals, IV

We can now invert our focus and switch from using p to study the variation in $\hat{p}$ to using $\hat{p}$ to study the variation in p.

- This is the same thing we did last time for the normal distribution: we noted that $\hat{\mu}$ was normally distributed with mean μ and standard deviation σ, and used this to deduce that μ was normally distributed with mean $\hat{\mu}$ and standard deviation σ.
- We did this by observing that $\hat{\mu} - \mu$ was normally distributed with mean 0 and standard deviation σ.

Binomial Confidence Intervals, V

We can try using the same idea in our case here:

- Since $\hat{p}$ has an approximately normal distribution with mean p and standard deviation $\sqrt{p(1-p)/n}$, that means $\hat{p} - p$ has an approximately normal distribution with mean 0 and standard deviation $\sqrt{p(1-p)/n}$.
- So that means p is approximately normally distributed with mean $\hat{p}$ and standard deviation $\sqrt{p(1-p)/n}$.
- However, there is one crucial problem: with the normal distribution we were given the standard deviation of the distribution explicitly. But here, the standard deviation still depends on the (now) unknown parameter p.

Binomial Confidence Intervals, VI

It would seem that we are stuck going in circles: we cannot find the standard deviation $\sqrt{p(1-p)/n}$ without knowing the value of p, but p is exactly the quantity we are trying to set up a confidence interval for!

- We would like to avoid having to study the complicated way in which the exact distribution of p would depend on $\hat{p}$.
- Here is a way out of this conundrum: if we assume (reasonably) that $\hat{p}$ is fairly close to p, then the true standard deviation $\sqrt{p(1-p)/n}$ should be fairly close to the estimated standard deviation $\sqrt{\hat{p}(1-\hat{p})/n}$ that uses the sample proportion $\hat{p}$.
- So what we will do is use the estimate $\sqrt{\hat{p}(1-\hat{p})/n}$ for the standard deviation of our normal distribution.
- Once we use this estimate, we are back in the situation of writing down a normal confidence interval.

Binomial Confidence Intervals, VII

For completeness, here is a brief justification of why this replacement is acceptable:

- As we noted, $\hat{p}$ is roughly normally distributed with mean p and standard deviation $\sqrt{p(1-p)/n}$, which is typically much smaller than p.
- Indeed, 99.8% of the time, $|\hat{p}-p| < 3\sqrt{p(1-p)/n}$.
- By a basic calc-1 linearization, we can compute the estimate $\left|\sqrt{\hat{p}(1-\hat{p})/n} - \sqrt{p(1-p)/n}\right| \approx |3-6\hat{p}|/(2n) + O(n^{-2})$.
- This is quite small relative to the actual standard deviation, which is on the order of $\sqrt{\hat{p}/(2n)}$ at worst.
- Thus, when n is large, the relative error introduced by this replacement will be very small. And of course, since we are using the normal approximation to the binomial distribution, we will require n to be large anyway.

Binomial Confidence Intervals, VII

To summarize that lengthy discussion, we have the following:

Proposition (Binomial Confidence Intervals)

Suppose a Bernoulli random variable is sampled n times yielding k successes, for an overall sample success rate of $\hat{p} = k/n$. In situations where the normal approximation to the binomial distribution is accurate (heuristically, when k and $n-k$ are both larger than 5 or so), then a $100(1-\alpha)\%$ confidence interval for the true success probability p is given by
$\hat{p} \pm c\sqrt{\hat{p}(1-\hat{p})/n} = (\hat{p} - c\sqrt{\hat{p}(1-\hat{p})/n}, \hat{\mu} + c\sqrt{\hat{p}(1-\hat{p})/n})$,
where c is the constant satisfying $P(-c < N_{0,1} < c) = 1-\alpha$.

We use the same table of (α, c) as we did before:

$1-\alpha$	50%	80%	90%	95%	98%	99%	99.5%	99.9%
c	0.6745	1.2816	1.6449	1.9600	2.3263	2.5758	2.8070	3.2905

Binomial Confidence Intervals, VIII

More compactly, our best estimate if we observe k successes in n trials for the overall success rate is $\hat{p} = k/n$, and the margin of error at the $100(1-\alpha)\%$ confidence level is equal to $c\sigma$ where $\sigma = \sqrt{\hat{p}(1-\hat{p})/n}$ is the sample's standard deviation.

Examples, I

Example: A coin with unknown probability p of landing heads is flipped 100 times, yielding 64 heads. Find the following:

1. A 50% confidence interval for p.
2. An 80% confidence interval for p.
3. A 90% confidence interval for p.
4. A 99.5% confidence interval for p.
5. The reasonableness that the coin is actually fair.
6. The reasonableness that the coin actually has heads probability 2/3.

$1-\alpha$	50%	80%	90%	95%	98%	99%	99.5%	99.9%
c	0.6745	1.2816	1.6449	1.9600	2.3263	2.5758	2.8070	3.2905

Examples, II

<u>Example</u>: A coin with unknown probability p of landing heads is flipped 100 times, yielding 64 heads. Find the following:

$1-\alpha$	50%	80%	90%	95%	98%	99%	99.5%	99.9%
c	0.6745	1.2816	1.6449	1.9600	2.3263	2.5758	2.8070	3.2905

1. A 50% confidence interval for p.

Examples, II

Example: A coin with unknown probability p of landing heads is flipped 100 times, yielding 64 heads. Find the following:

$1-\alpha$	50%	80%	90%	95%	98%	99%	99.5%	99.9%
c	0.6745	1.2816	1.6449	1.9600	2.3263	2.5758	2.8070	3.2905

1. A 50% confidence interval for p.
   - Here, we have $n = 100$ and $\hat{p} = 64/100 = 0.64$, so that $\sigma = \sqrt{\hat{p}(1-\hat{p})/n} = 0.048$.
   - Thus, the 50% CI is $\hat{p} \pm 0.6745\sigma = (0.6076, 0.6724)$.
2. An 80% confidence interval for p.

Examples, II

Example: A coin with unknown probability p of landing heads is flipped 100 times, yielding 64 heads. Find the following:

$1-\alpha$	50%	80%	90%	95%	98%	99%	99.5%	99.9%
c	0.6745	1.2816	1.6449	1.9600	2.3263	2.5758	2.8070	3.2905

1. A 50% confidence interval for p.
   - Here, we have $n = 100$ and $\hat{p} = 64/100 = 0.64$, so that $\sigma = \sqrt{\hat{p}(1-\hat{p})/n} = 0.048$.
   - Thus, the 50% CI is $\hat{p} \pm 0.6745\sigma = (0.6076, 0.6724)$.
2. An 80% confidence interval for p.
   - This is $\hat{p} \pm 1.2816\sigma = (0.5785, 0.7015)$.
3. A 90% confidence interval for p.

Examples, II

Example: A coin with unknown probability p of landing heads is flipped 100 times, yielding 64 heads. Find the following:

$1-\alpha$	50%	80%	90%	95%	98%	99%	99.5%	99.9%
c	0.6745	1.2816	1.6449	1.9600	2.3263	2.5758	2.8070	3.2905

1. A 50% confidence interval for p.
   - Here, we have $n = 100$ and $\hat{p} = 64/100 = 0.64$, so that $\sigma = \sqrt{\hat{p}(1-\hat{p})/n} = 0.048$.
   - Thus, the 50% CI is $\hat{p} \pm 0.6745\sigma = (0.6076, 0.6724)$.
2. An 80% confidence interval for p.
   - This is $\hat{p} \pm 1.2816\sigma = (0.5785, 0.7015)$.
3. A 90% confidence interval for p.
   - This is $\hat{p} \pm 1.6449\sigma = (0.5610, 0.7190)$.
4. A 99.5% confidence interval for p.

Examples, II

Example: A coin with unknown probability p of landing heads is flipped 100 times, yielding 64 heads. Find the following:

$1-\alpha$	50%	80%	90%	95%	98%	99%	99.5%	99.9%
c	0.6745	1.2816	1.6449	1.9600	2.3263	2.5758	2.8070	3.2905

1. A 50% confidence interval for p.
   - Here, we have $n = 100$ and $\hat{p} = 64/100 = 0.64$, so that $\sigma = \sqrt{\hat{p}(1-\hat{p})/n} = 0.048$.
   - Thus, the 50% CI is $\hat{p} \pm 0.6745\sigma = (0.6076, 0.6724)$.
2. An 80% confidence interval for p.
   - This is $\hat{p} \pm 1.2816\sigma = (0.5785, 0.7015)$.
3. A 90% confidence interval for p.
   - This is $\hat{p} \pm 1.6449\sigma = (0.5610, 0.7190)$.
4. A 99.5% confidence interval for p.
   - This is $\hat{p} \pm 2.0870\sigma = (0.5053, 0.7747)$.

Examples, III

Example: A coin with unknown probability p of landing heads is flipped 100 times, yielding 64 heads. Find the following:

5. The reasonableness that the coin is actually fair.

Examples, III

Example: A coin with unknown probability p of landing heads is flipped 100 times, yielding 64 heads. Find the following:

5. The reasonableness that the coin is actually fair.

- The 99.5% confidence interval was $(0.5053, 0.7747)$, and even this does not contain $p = 0.5$.
- This suggests it is very unlikely that the coin is actually fair.

Examples, III

Example: A coin with unknown probability p of landing heads is flipped 100 times, yielding 64 heads. Find the following:

5. The reasonableness that the coin is actually fair.
   - The 99.5% confidence interval was $(0.5053, 0.7747)$, and even this does not contain $p = 0.5$.
   - This suggests it is very unlikely that the coin is actually fair.
6. The reasonableness that the coin actually has heads probability $2/3$.

Examples, III

Example: A coin with unknown probability p of landing heads is flipped 100 times, yielding 64 heads. Find the following:

5. The reasonableness that the coin is actually fair.
   - The 99.5% confidence interval was $(0.5053, 0.7747)$, and even this does not contain $p = 0.5$.
   - This suggests it is very unlikely that the coin is actually fair.
6. The reasonableness that the coin actually has heads probability $2/3$.
   - The 50% CI was $\hat{p} \pm 0.6745\sigma = (0.6076, 0.6724)$, which does contain $p = 2/3$. This suggests it is not unreasonable to think that the coin actually has heads probability $2/3$.
   - Although $p = 2/3$ is not the most likely estimate based on the observed data (that would be $\hat{p} = 0.64$), it is not so far away from the most likely estimate that we should find it strange if the coin really did have heads probability $2/3$.

Examples, IV

In tomorrow's lecture, we will develop the notion of hypothesis testing, and in Thursday's lecture, we will return to this example to discuss how we can quantify the "believability" of these two hypotheses ($p = 1/2$ and $p = 2/3$) that we just discussed.

Examples, V

Example: Last season, a professional basketball player attempted 1800 two-point shots in a season and made 753 of them. Find

1. An 80% confidence interval for his shooting average last season.
2. A 99% confidence interval for his shooting average last season.
3. An 80% confidence interval for the total number of shots he should expect to make this season if he attempts 1500 shots and his true shooting average stays the same as last season.
4. A 99% confidence interval for the number of shots made this season.

Examples, VI

Example: Last season, a professional basketball player attempted 1800 two-point shots in a season and made 753 of them. Find

1. An 80% confidence interval for his shooting average last season.

Examples, VI

Example: Last season, a professional basketball player attempted 1800 two-point shots in a season and made 753 of them. Find

1. An 80% confidence interval for his shooting average last season.
   - Here, we have $n = 1800$ and $\hat{p} = 753/1800 \approx 41.83\%$, so that $\sigma = \sqrt{\hat{p}(1-\hat{p})/n} \approx 1.163\%$.
   - We obtain the 80% confidence interval $\hat{p} \pm 1.2816\sigma = (40.34\%, 43.32\%)$.
2. A 99% confidence interval for his shooting average last season.

Examples, VI

Example: Last season, a professional basketball player attempted 1800 two-point shots in a season and made 753 of them. Find

1. An 80% confidence interval for his shooting average last season.
   - Here, we have $n = 1800$ and $\hat{p} = 753/1800 \approx 41.83\%$, so that $\sigma = \sqrt{\hat{p}(1-\hat{p})/n} \approx 1.163\%$.
   - We obtain the 80% confidence interval $\hat{p} \pm 1.2816\sigma = (40.34\%, 43.32\%)$.
2. A 99% confidence interval for his shooting average last season.
   - We obtain the 99% confidence interval $\hat{p} \pm 2.5758\sigma = (38.83\%, 44.83\%)$.

Examples, VI

Example: Last season, a professional basketball player attempted 1800 two-point shots in a season and made 753 of them. Find

3. An 80% confidence interval for the total number of shots he should expect to make this season if he attempts 1500 shots and his true shooting average stays the same as last season.

Examples, VI

Example: Last season, a professional basketball player attempted 1800 two-point shots in a season and made 753 of them. Find

3. An 80% confidence interval for the total number of shots he should expect to make this season if he attempts 1500 shots and his true shooting average stays the same as last season.

- The distribution of the total number of made shots out of 1500 attempts will be binomial with mean $1500p$ and standard deviation $\sqrt{1500p(1-p)}$.
- We apply the approximation $\sqrt{p(1-p)} \approx \sqrt{\hat{p}(1-\hat{p})}$ (which we also used, and justified as reasonable, in our analysis of the binomial distribution above).
- Then, using the normal approximation to the binomial, the number of made shots this season is distributed approximately normally with mean $1500p$ and standard deviation $\sigma' = \sqrt{1500\hat{p}(1-\hat{p})} = 19.105$.

Examples, VII

Example: Last season, a professional basketball player attempted 1800 two-point shots in a season and made 753 of them. Find

3. An 80% confidence interval for the total number of shots he should expect to make this season if he attempts 1500 shots and his true shooting average stays the same as last season.

- We are thus constructing confidence intervals with mean $1500p$ and standard deviation $\sigma' = \sqrt{1500\hat{p}(1-\hat{p})} = 19.105$.
- This yields the 80% CI $1500\hat{p} \pm 1.2816\sigma' = (603, 652)$.
- Alternatively, we could simply scale our previous 80% confidence interval by 1500. (Think about why.)

4. A 99% confidence interval for the number of shots made this season.

Examples, VII

Example: Last season, a professional basketball player attempted 1800 two-point shots in a season and made 753 of them. Find

3. An 80% confidence interval for the total number of shots he should expect to make this season if he attempts 1500 shots and his true shooting average stays the same as last season.

- We are thus constructing confidence intervals with mean $1500p$ and standard deviation $\sigma' = \sqrt{1500\hat{p}(1-\hat{p})} = 19.105$.
- This yields the 80% CI $1500\hat{p} \pm 1.2816\sigma' = (603, 652)$.
- Alternatively, we could simply scale our previous 80% confidence interval by 1500. (Think about why.)

4. A 99% confidence interval for the number of shots made this season.

- We obtain the 99% CI $1500\hat{p} \pm 2.5758\sigma' = (578, 677)$.

Examples, VIII

Example: To test for fairness, a six-sided die is rolled 2000 times, yielding the results below. Find:

Number	1	2	3	4	5	6
Occurrences	354	347	318	312	333	336

1. An 80% and 95% confidence interval for the true probability of rolling a 1.
2. An 80% and 95% confidence interval for the true probability of rolling a 4.
3. An approximate 80% and 95% confidence interval for the average value of one roll of the die.
4. Are all of these confidence intervals consistent with the die actually being fair?

Examples, IX

Example: To test for fairness, a six-sided die is rolled 2000 times, yielding the results below. Find:

Number	1	2	3	4	5	6
Occurrences	354	347	318	312	333	336

1. An 80% and 95% confidence interval for the true probability of rolling a 1.

Examples, IX

Example: To test for fairness, a six-sided die is rolled 2000 times, yielding the results below. Find:

Number	1	2	3	4	5	6
Occurrences	354	347	318	312	333	336

1. An 80% and 95% confidence interval for the true probability of rolling a 1.

- Here, we have $n = 2000$ and $\hat{p} = 354/2000 = 0.177$, so that $\sigma = \sqrt{\hat{p}(1-\hat{p})/n} \approx 0.00853$.
- We obtain the 80% CI $\hat{p} \pm 1.2816\sigma = (0.1661, 0.1879)$.
- We obtain the 95% CI $\hat{p} \pm 1.9600\sigma = (0.1603, 0.1937)$.

Examples, X

Example: To test for fairness, a six-sided die is rolled 2000 times, yielding the results below. Find:

Number	1	2	3	4	5	6
Occurrences	354	347	318	312	333	336

2. An 80% and 95% confidence interval for the true probability of rolling a 4.

Examples, X

Example: To test for fairness, a six-sided die is rolled 2000 times, yielding the results below. Find:

Number	1	2	3	4	5	6
Occurrences	354	347	318	312	333	336

2. An 80% and 95% confidence interval for the true probability of rolling a 4.

- Here, we have $n = 2000$ and $\hat{p} = 312/2000 = 0.156$, so that $\sigma = \sqrt{\hat{p}(1-\hat{p})/n} \approx 0.00811$.
- We obtain the 80% CI $\hat{p} \pm 1.2816\sigma = (0.1456, 0.1664)$.
- We obtain the 95% CI $\hat{p} \pm 1.9600\sigma = (0.1401, 0.1719)$.

Examples, XI

Example: To test for fairness, a six-sided die is rolled 2000 times, yielding the results below. Find:

Number	1	2	3	4	5	6
Occurrences	354	347	318	312	333	336

3. An approximate 80% and 95% confidence interval for the average value of one roll of the die.

Examples, XI

Example: To test for fairness, a six-sided die is rolled 2000 times, yielding the results below. Find:

Number	1	2	3	4	5	6
Occurrences	354	347	318	312	333	336

3. An approximate 80% and 95% confidence interval for the average value of one roll of the die.

- If the true distribution of one roll has mean μ and standard deviation σ, then the average of 2000 rolls will have mean μ and standard deviation $\sigma' = \sigma/\sqrt{2000}$, and be approximately normally distributed by the central limit theorem.
- Although we do not know the actual value of σ, it is reasonable to feel that it should be very close to the sample standard deviation $S = 1.7321$, which (as we showed) is an unbiased estimator of the true variance.

Examples, XII

Example: To test for fairness, a six-sided die is rolled 2000 times, yielding the results below. Find:

Number	1	2	3	4	5	6
Occurrences	354	347	318	312	333	336

3. An approximate 80% and 95% confidence interval for the average value of one roll of the die.

Examples, XII

Example: To test for fairness, a six-sided die is rolled 2000 times, yielding the results below. Find:

Number	1	2	3	4	5	6
Occurrences	354	347	318	312	333	336

3. An approximate 80% and 95% confidence interval for the average value of one roll of the die.

- The actual average of the die rolls is $\hat{\mu} = 3.4655$, and its standard deviation is then estimated to be $\sigma' = S/\sqrt{2000} = 1.7321/\sqrt{2000} \approx 0.03873$.
- This gives us the 80% CI $\hat{\mu} \pm 1.2816\sigma' = (3.4158, 3.5151)$ and also the 95% CI $\hat{\mu} \pm 1.9600\sigma' = (3.3896, 3.5414)$.

Examples, XIII

Example: To test for fairness, a six-sided die is rolled 2000 times, yielding the results below. Find:

Number	1	2	3	4	5	6
Occurrences	354	347	318	312	333	336

3. An approximate 80% and 95% confidence interval for the average value of one roll of the die.

- As a sanity check, no matter what the actual distribution of values is for the roll of one die, it is not so hard to see that the standard deviation is always at most 2.5, since that is half the maximum difference between two rolls of the die.
- The actual average of the die rolls is $\hat{\mu} = 3.4655$ and the standard deviation is at most $\sigma'' = 2.5/\sqrt{2000} \approx 0.05590$.
- This yields the "worst-case scenario"
  80% CI $\hat{\mu} \pm 1.2816\sigma'' = (3.3938, 3.5371)$ and
  95% CI $\hat{\mu} \pm 1.9600\sigma'' = (3.3559, 3.5751)$.

Examples, XIV

Example: To test for fairness, a six-sided die is rolled 2000 times, yielding the results below. Find:

4. Are all of these confidence intervals consistent with the die actually being fair?

- Rolling 1 had 80% CI $(0.1661, 0.1879)$.
- Rolling 4 had 80% CI $(0.1456, 0.1664)$.
- The average value had 80% CI $(3.4158, 3.5151)$.

Examples, XIV

Example: To test for fairness, a six-sided die is rolled 2000 times, yielding the results below. Find:

4. Are all of these confidence intervals consistent with the die actually being fair?

- Rolling 1 had 80% CI $(0.1661, 0.1879)$.
- Rolling 4 had 80% CI $(0.1456, 0.1664)$.
- The average value had 80% CI $(3.4158, 3.5151)$.
- We can see that the expected true value ($1/6$ for the probability of rolling 1 or 4, and 3.5 for the average value of one roll) lies inside, or just barely outside, the 80% confidence interval in each case.
- This doesn't provide strong evidence against the die being fair.

Examples, XIV

Example: To test for fairness, a six-sided die is rolled 2000 times, yielding the results below. Find:

4. Are all of these confidence intervals consistent with the die actually being fair?

- Rolling 1 had 80% CI $(0.1661, 0.1879)$.
- Rolling 4 had 80% CI $(0.1456, 0.1664)$.
- The average value had 80% CI $(3.4158, 3.5151)$.
- We can see that the expected true value ($1/6$ for the probability of rolling 1 or 4, and 3.5 for the average value of one roll) lies inside, or just barely outside, the 80% confidence interval in each case.
- This doesn't provide strong evidence against the die being fair.

In the next chapter, we will discuss methods of testing the goodness of fit of an expected distribution to an observed data sample using the χ^2 test.

Polling, I

An extremely common use of confidence intervals is in polling statistics, where a random sample of a population is used to estimate the proportion that support a particular measure.

- Typically, most polls report the margin of error associated with a 95% confidence interval.
- In popular parlance, it is usually referred to as simply "the margin of error", with no qualifier, but most reputable polls also include the confidence level with their statistics.

Polling, II

Example: A newspaper poll reports "45% of voters support X, with a margin of error of 6%".

- This typically means that the 95% confidence interval for the percent support of X is $(39\%, 51\%)$.
- It is important not to misinterpret what the confidence interval actually represents: it says that we are 95% confident that the true level of support for X is somewhere between 39% and 51%.
- It does *not* mean that we believe all the values in this range are equally likely!
- It also does *not* mean that the true value must be inside this range!

Polling, III

Example: A newspaper poll reports "45% of voters support X, with a margin of error of 6%".

- Although a portion of the 95% confidence interval (39%, 51%) does include outcomes where the support of X is above 50%, it is far more likely that the support for X is below 50% than above 50%.
- That is because we expect the true distribution to be approximately normal, with mean 45% and standard deviation $6\%/1.9600 = 3.0612\%$ (or so).
- Indeed, we can compute the actual expected probability that the support is above 50%: it is $P(N_{45,3.0612} \geq 50) = P(N_{0,1} \geq 1.3708) = 0.0852$. Moderately low, but still positive!

Polling, IV

<u>Example</u>: A pollster wishes to measure the statewide support for Proposition Q. He randomly samples 1000 likely voters and finds 540 of them support Proposition Q. Find

1. Find a 95% confidence interval, and its associated margin of error, for the true percentage of the population that supports Proposition Q.
2. Find a 99.9% confidence interval, and its associated margin of error, for the true percentage of the population that supports Proposition Q.
3. Estimate the probability that Proposition Q actually has at least 50% support in the general population.

Polling, V

Example: A pollster wishes to measure the statewide support for Proposition Q. He randomly samples 1000 likely voters and finds 540 of them support Proposition Q. Find

1. Find a 95% confidence interval, and its associated margin of error, for the true percentage of the population that supports Proposition Q.

Polling, V

Example: A pollster wishes to measure the statewide support for Proposition Q. He randomly samples 1000 likely voters and finds 540 of them support Proposition Q. Find

1. Find a 95% confidence interval, and its associated margin of error, for the true percentage of the population that supports Proposition Q.

- Here, we have $n = 1000$ and $\hat{p} = 540/1000 = 54\%$, so that $\sigma = \sqrt{\hat{p}(1-\hat{p})/n} \approx 1.576\%$.
- The margin of error for the 95% confidence interval is $1.9600\sigma \approx 3.09\%$.
- The confidence interval itself is $54\% \pm 3.09\% = (50.91\%, 57.09\%)$.

Polling, VI

Example: A pollster wishes to measure the statewide support for Proposition Q. He randomly samples 1000 likely voters and finds 540 of them support Proposition Q. Find

2. Find a 99.9% confidence interval, and its associated margin of error, for the true percentage of the population that supports Proposition Q.

Polling, VI

Example: A pollster wishes to measure the statewide support for Proposition Q. He randomly samples 1000 likely voters and finds 540 of them support Proposition Q. Find

2. Find a 99.9% confidence interval, and its associated margin of error, for the true percentage of the population that supports Proposition Q.

- From before, $\hat{p} = 540/1000 = 54\%$ and $\sigma \approx 1.576\%$.
- The margin of error for the 99.9% confidence interval is $3.2905\sigma \approx 5.19\%$.
- The confidence interval itself is $54\% \pm 5.19\% = (48.81\%, 59.19\%)$.

Polling, VII

Example: A pollster wishes to measure the statewide support for Proposition Q. He randomly samples 1000 likely voters and finds 540 of them support Proposition Q. Find

3. Estimate the probability that Proposition Q actually has at least 50% support in the general population.

Polling, VII

<u>Example</u>: A pollster wishes to measure the statewide support for Proposition Q. He randomly samples 1000 likely voters and finds 540 of them support Proposition Q. Find

3. Estimate the probability that Proposition Q actually has at least 50% support in the general population.

- From before, $\hat{p} = 540/1000 = 54\%$ and $\sigma \approx 1.576\%$.
- Since the difference $\hat{p} - p$ is approximately normally distributed with standard deviation σ, we can interpret this as saying that p is approximately normally distributed with mean $\hat{p}$ and standard deviation σ.
- Then $P(p \geq 50\%) = P(N_{54\%,1.576\%} \geq 50\%) = P(N_{0,1} \geq -2.5381) = 0.9944$. Quite likely!

Polling, VIII

Example: A pollster wishes to measure the support for the statewide support of Proposition R.

1. If she expects the support level for the proposition to be approximately 65%, what is the smallest number of people needed for the 95% confidence interval's margin of error to be at most $\pm 2\%$?
2. How would the answer change if the support level for the proposition is unknown?

Polling, IX

Example: A pollster wishes to measure the support for the statewide support of Proposition R.

1. If she expects the support level for the proposition to be approximately 65%, what is the smallest number of people needed for the 95% confidence interval's margin of error to be at most $\pm 2\%$?

Polling, IX

<u>Example</u>: A pollster wishes to measure the support for the statewide support of Proposition R.

1. If she expects the support level for the proposition to be approximately 65%, what is the smallest number of people needed for the 95% confidence interval's margin of error to be at most $\pm 2\%$?

- Here, the expected proportion is $\hat{p} = 0.65$, so $\sigma = \sqrt{\hat{p}(1-\hat{p})/n}$.
- At the 95% confidence level, the margin of error is 1.9600σ.
- Thus, we want $1.9600\sqrt{\hat{p}(1-\hat{p})/n} = 2\%$, which yields $n = \dfrac{\hat{p}(1-\hat{p})}{(0.02/1.9600)^2} \approx 2184.9$.
- Thus, the minimum number of people needed for the poll will be 2185 to achieve a 2% margin of error at the 95% confidence level.

Polling, X

Example: A pollster wishes to measure the support for the statewide support of Proposition R.

2. How would the answer change if the support level for the proposition is unknown?

Polling, X

<u>Example</u>: A pollster wishes to measure the support for the statewide support of Proposition R.

2. How would the answer change if the support level for the proposition is unknown?

- If the support level $\hat{p}$ is unknown, the largest possible value of $n = \dfrac{\hat{p}(1-\hat{p})}{(0.02/1.9600)^2}$ will occur when the numerator $\hat{p}(1-\hat{p})$ is maximized.
- Either by calculus or completing the square, we can see that this maximum occurs when $\hat{p} = 1/2$.
- The corresponding value of n is then $\dfrac{(1/2)\cdot(1/2)}{(0.02/1.9600)^2} \approx 2401.0$.

Polling, XI

Example: A political article states "Based on a recent poll, candidate Y has an approval rating of $43.1\% \pm 3.0\%$ (95% CI, $n = 750$), which means that it is impossible for their favorability rating to be 50% or above". Critique this statement.

Polling, XI

<u>Example</u>: A political article states "Based on a recent poll, candidate Y has an approval rating of $43.1\% \pm 3.0\%$ (95% CI, $n = 750$), which means that it is impossible for their favorability rating to be 50% or above". Critique this statement.

- Because the poll was conducted by sampling, there is always a possibility (however remote) that the actual favorability rating lies outside any given confidence interval.
- Thus, it is *always possible* that the poll is giving a result that is very far off from reality.

Polling, XI

<u>Example</u>: A political article states "Based on a recent poll, candidate Y has an approval rating of $43.1\% \pm 3.0\%$ (95% CI, $n = 750$), which means that it is impossible for their favorability rating to be 50% or above". Critique this statement.

- In this case, given that the sample size was $n = 750$ and that it is a 95% confidence interval with a success probability $\hat{p} = 0.431$, the actual distribution of the true favorability rating will be normal with mean 42.1% and standard deviation $\sqrt{\hat{p}(1-\hat{p})/n} \approx 1.80\%$. (Note that this is consistent with the quoted information since the margin of error would then be $1.96\sigma \approx 3.0\%$.)
- Using properties of the normal distribution, we can then compute $P(N_{43.1,1.80} > 50) = P(N_{0,1} > 3.833) \approx 0.06\%$. So, although it is fairly unlikely that candidate Y's favorability rating is actually 50% or above, it is certainly still possible.

Polling, XII

Example: In a two-candidate runoff, a poll of 1000 respondents indicates A has 48.2% support while B has 51.8% support.

1. Calculate the 95%-confidence margin of error for the poll.
2. Find the probability that A actually has at least 50% support (i.e., will win the race).
3. A pundit makes the following statement: "Because the two candidates are within the margin of error for the poll (a statistical tie), the race is a toss-up." Critique this statement.

Polling, XIII

Example: In a two-candidate runoff, a poll of 1000 respondents indicates A has 48.2% support while B has 51.8% support.

1. Calculate the 95%-confidence margin of error for the poll.

- The margin of error is $1.9600 \cdot \sqrt{\hat{p}(1-\hat{p})/n}$ where $\hat{p}$ is the support probability.
- For A, we obtain a margin of $1.9600 \cdot \sqrt{0.482 \cdot 0.518/1000} \approx 3.10\%$.
- For B, we obtain a margin of $1.9600 \cdot \sqrt{0.518 \cdot 0.482/1000} \approx 3.10\%$.
- It should make sense that the two candidates have the same margin of error, since knowing A's support percentage tells us B's, and vice versa.

Polling, XIV

Example: In a two-candidate runoff, a poll of 1000 respondents indicates A has 48.2% support while B has 51.8% support.

2. Find the probability that A actually has at least 50% support (i.e., will win the runoff).

- Here, we have $k = 482$ and $n - k = 518$, so the normal approximation to the binomial distribution will be good.
- Then the support proportion for A will be approximately normally distributed with mean $\hat{p} = 0.482$ and standard deviation $\sqrt{0.482 \cdot 0.518/1000} \approx 0.0158$.
- Then $P(A > 0.50) = P(N_{0.482,0.0158} > 0.50) = P(N_{0,1} > 1.139) \approx 0.1273$.
- Thus, A has about a 12.7% chance of winning the runoff, based on the results of this poll.

Polling, XV

Example: In a two-candidate runoff, a poll of 1000 respondents indicates A has 48.2% support while B has 51.8% support.

3. A pundit makes the following statement: "Because the two candidates are within the margin of error for the poll (a statistical tie), the race is a toss-up." Critique this statement.

- Despite the fact that the candidates are "within the margin of error", as we just calculated, A only has about a 1-in-8 chance of winning the election. So it is quite inaccurate to say that both candidates are equally likely to win!

Polling, XVI

Example: A poll on a national political issue is taken, and the results are broken down by various demographics:

Group	% Support	# Support	# Polled
All	52.1%	13,840	26,565
Men	47.2%	6,335	13,421
Women	56.8%	7,407	13,041
Age 18-29	67.9%	3,430	5,051
Age 30-44	51.8%	3,718	7,178
Age 45-64	48.3%	5,349	11,075
Age 65+	41.2%	1,343	3,261

Find the margin of error for the support in each group at the 95% confidence level.

Polling, XVII

We can calculate the 95%-CI margins using our earlier formula $1.9600 \cdot \sqrt{\hat{p}(1-\hat{p})/n}$:

Group	% Support	# Support	# Polled	% Margin
All	52.1%	13,840	26,565	0.60%
Men	47.2%	6,335	13,421	0.84%
Women	56.8%	7,407	13,041	0.85%
Age 18-29	67.9%	3,430	5,051	1.29%
Age 30-44	51.8%	3,718	7,178	1.16%
Age 45-64	48.3%	5,348	11,077	0.93%
Age 65+	41.2%	1,344	3,259	1.69%

Note that the smaller groups have a larger margin of error (as one would expect, of course, since less information is collected from smaller groups).

Polling, XVIII

Notice also in this last example that the support percentages vary quite substantially among different groups.

- This is a fairly common phenomenon: most demographic groups do not have identical levels of support for most issues.
- But now imagine you are trying to conduct a poll that will accurately represent the overall support among everyone: if, for example, your poll has 60% men and 40% women, the results will show a bias toward a lower support percentage than the true result.
- Avoiding such bias is a difficult issue, since it requires pollsters to conduct polls that are representative of the general population in very many ways (gender, race, ethnicity, age, income, education, national origin, etc.), which is a difficult problem in general.

Confidence Intervals Everywhere, I

We have reached the end of our official discussion of confidence intervals, and will start discussing hypothesis testing in earnest next lecture.

- For reasons of efficiency (in the colloquial sense, not the statistical sense), we have only dealt with confidence interval construction using the normal distribution.
- After we discuss the t distribution in the next chapter, we will also describe how to construct confidence intervals for normally-distributed variables whose standard deviation is not known, which is the most typical practical scenario.

Confidence Intervals Elsewhere, II

We will mention now, though, that it is entirely possible to construct confidence intervals for the other estimators we have analyzed so far.

- As a few particular examples, we could study the rescaled maximum estimator for the German tank problem, the average value estimator for the parameter of the Poisson distribution, and the reciprocal average value estimator for the parameter of the exponential distribution.
- However, each of these requires a similar level of effort in order to "invert" the analysis (as was needed with the normal and binomial distributions) to produce statements about the variation in the value of the true parameter in terms of the sample statistic.
- As such, we will leave a deeper discussion of general confidence intervals to another course.

Summary

We introduced how to construct confidence intervals for binomially-distributed variables.

We discussed applications of binomial confidence intervals to polling.

Next lecture: Hypothesis testing, z-tests