Hypothesis Testing

Hypothesis tests give a way of using samples to test whether or not statistical claims are likely to be true or not.

Dr. Unsnora prescribes SnoreCull to 15 of her patients and records whether it cured them or not after 2 weeks. She found that 11 were cured and 4 were not.

If the drug maker claimed that 90% get cured, 13.5 or 14 patients should have been cured. Is the company making false claims or is the doctor's sampling biased?

Step 1: Decide on the hypothesis

SnoreCull cures 90% of the patients within 2 weeks.

This is called Null Hypothesis and is represented by H_0.

In this case, H_0: $p = 0.9$

If Null Hypothesis is rejected based on evidence, an Alternate Hypothesis, H_1, needs to be accepted. **We always start with the assumption that Null Hypothesis is true.**

In this case, H_1: $p < 0.9$

Examples of Hypotheses

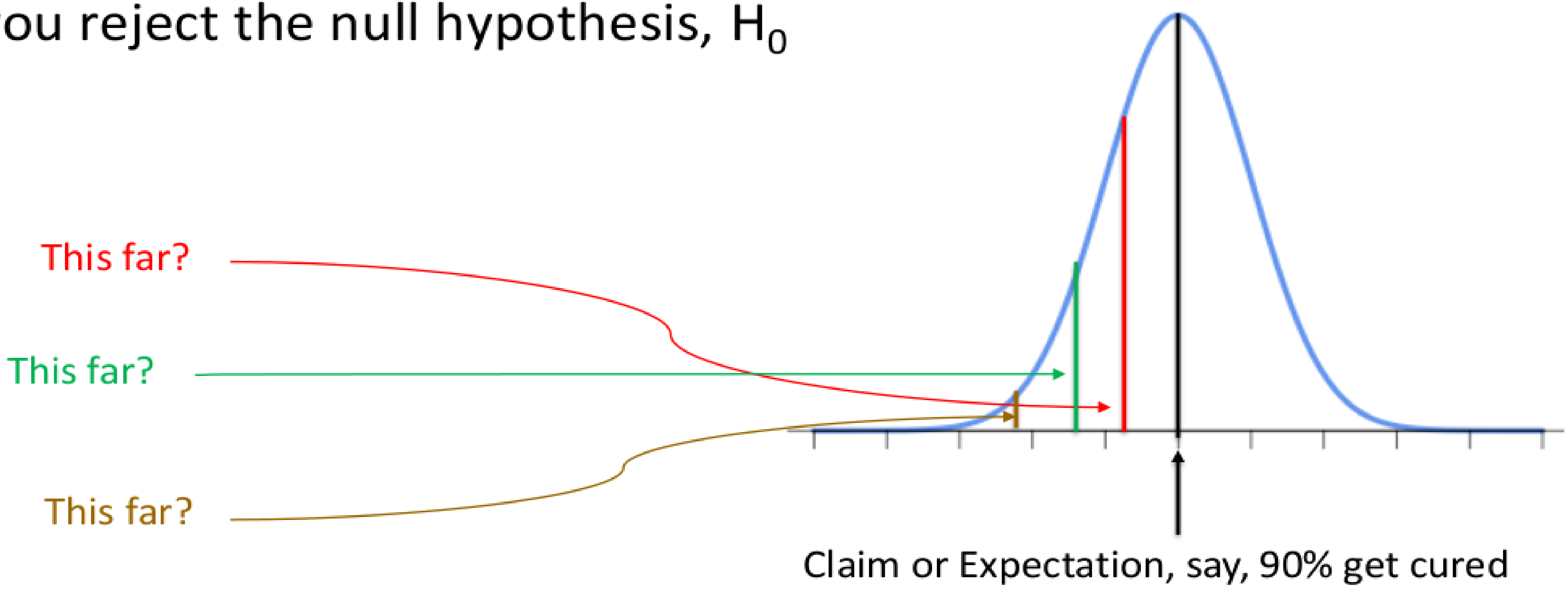
- Two hypotheses in competition:
  - H_0: The NULL hypothesis, usually the most conservative.
  - H_1 or H_A: The ALTERNATIVE hypothesis, the one we are actually interested in.
- Examples of NULL Hypothesis:
  - The coin is fair
  - The new drug is no better (or worse) than the placebo
- Examples of ALTERNATIVE hypothesis:
  - The coin is biased (either towards heads or tails)
  - The coin is biased towards heads
  - The coin has a probability 0.6 of landing on tails
  - The drug is better than the placebo

Step 2: Choose your statistic

$X \sim B(15, 0.9)$

Step 3: Specify the Significance Level

First, we must decide on the Significance Level, α. It is a measure of how unlikely you want the results of the sample to be before you reject the null hypothesis, H_0



Step 4: Determine the critical region

If X represents the number of snorers cured, the critical region is defined as $P(X < c) < \alpha$ where $\alpha = 5\%$.

Recall that in a 95% CI, there is a 5% chance that the sample will not contain the population mean. Hence if the sample falls in the critical region, the null hypothesis that 90% snorers are cured, is rejected.

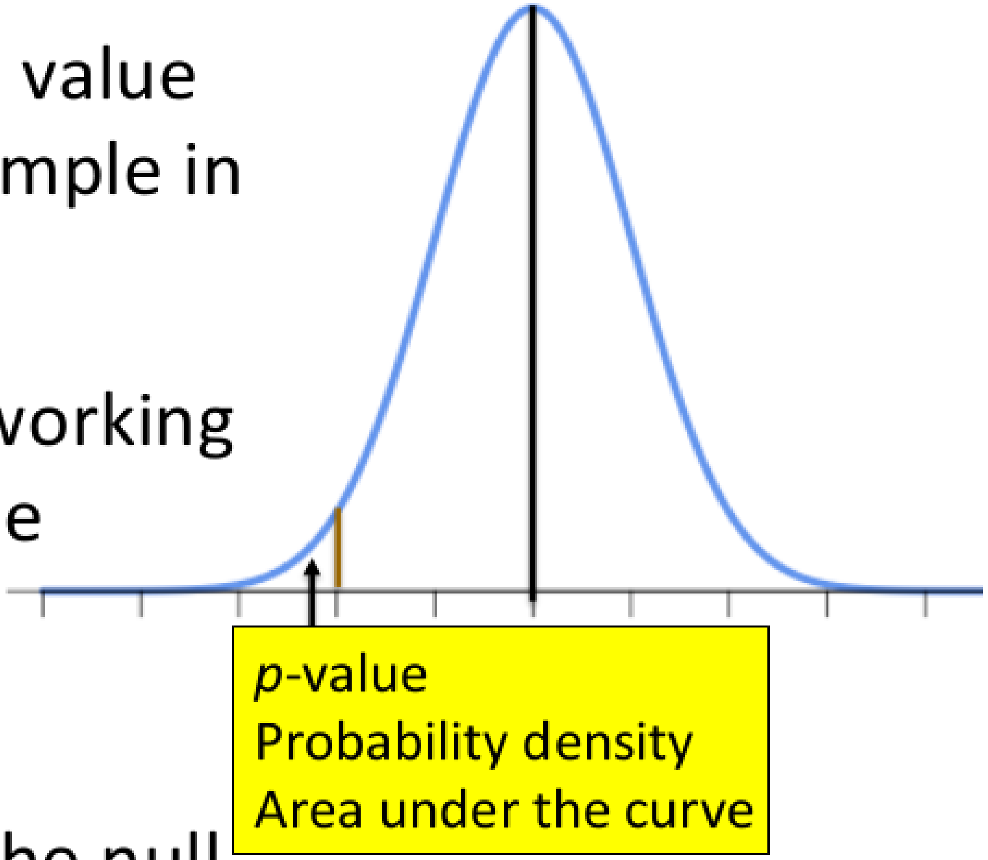
That is the reason 5% or 0.05 is called the Significance Level. In a 99% CI, 0.01 is the Significance Level.

Step 5: Find the *p*-value

p-value is the probability of getting a value up to and including the one in the sample in the direction of the critical region.

It is a way of taking the sample and working out whether the result falls within the critical region of the hypothesis test.

Essentially, this is the value used to determine whether or not to reject the null hypothesis.



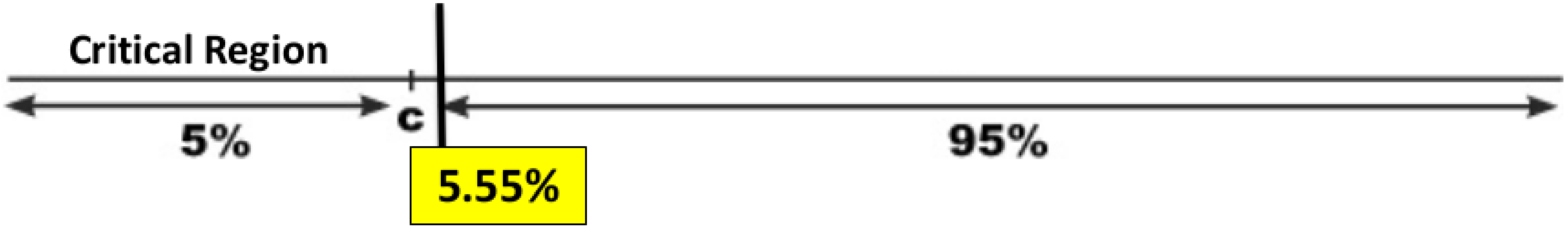
Step 5: Find the *p*-value

In the SnoreCull test done by Dr. Unsnora, 11 people were cured. This means our *p*-value is $P(X \leq 11)$, where X is the distribution of the number of people cured in the sample.

If $P(X \leq 11) < 0.05$ (Significance Level), it indicates that 11 is inside the critical region, and hence H_0 can be rejected.

Given that $X \sim B(15, 0.9)$, $P(X \leq 11) = 1 - P(X \geq 12) = 0.0555$

Step 6: Is the sample result in the critical region?



Step 7: Make your decision

There isn't sufficient evidence to reject the null hypothesis and so, the claims of the company are accepted.

Dr. Unsnora is not convinced and did another test with 100 people where 80 got cured and 20 didn't. What is your decision going to be now?

What are the null and alternate hypotheses?

H_0: $p = 0.9$

H_1: $p < 0.9$

What is the test statistic?

$X \sim B(100, 0.9)$ Oh! Dear

What probability distribution can be used to approximate the Binomial distribution?

Since np>5 and nq>5, Central Limit Theorem can be applied to sampling proportions.

What is the probability of 80% or fewer getting cured?

CONTINUITY CORRECTION FACTOR

$$z = \frac{\hat{p} + \frac{0.5}{n} - p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}} = \frac{0.805 - 0.9}{\sqrt{\frac{0.9 * 0.1}{100}}} = \frac{0.805 - 0.9}{\sqrt{0.0009}} = -3.17$$

p-value = P(Z < -3.17) = 0.0008

What is your decision?

Since the *p*-value (0.0008) is less than the Significance Level of 0.05, the null hypothesis can be rejected.

Attention Check

In hypothesis testing, do you assume the null hypothesis to be true or false?

True.

If there is sufficient evidence against the null hypothesis, do you accept it or reject it?

Reject it.

Attention Check

If the *p*-value is less than 0.05 for the above significance level, will you accept or reject the null hypothesis?

Reject it.

Do you need weaker evidence or stronger to reject the null hypothesis if you were testing at the 1% significance level instead of the 5% significance level?

Stronger.

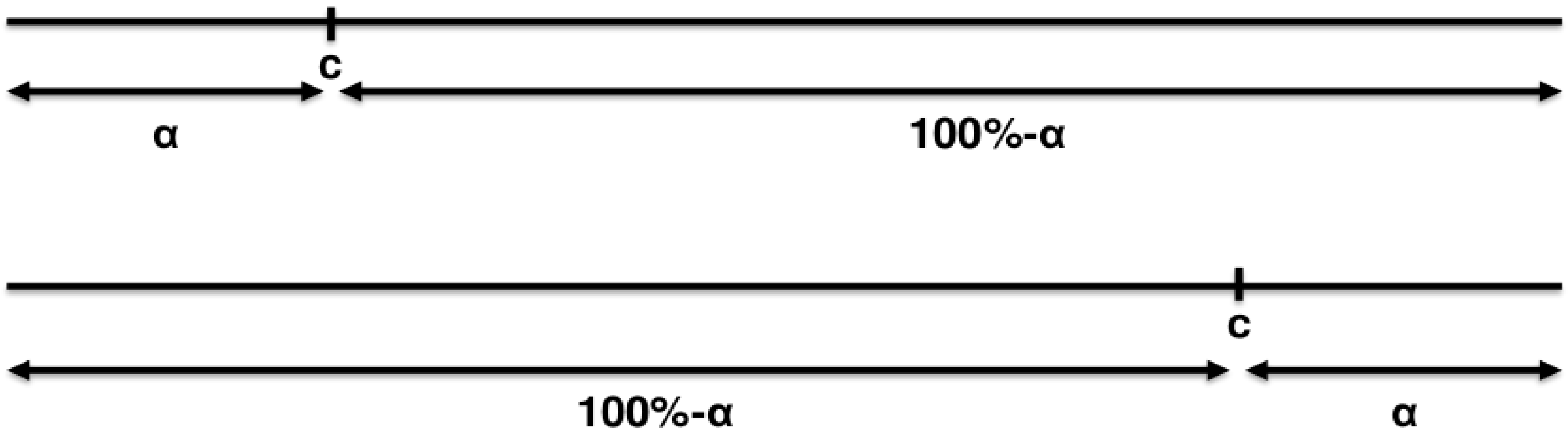
Critical Region Up Close

One-tailed tests

The position of the tail is dependent on H_1.

If H_1 includes a $<$ sign, then the **lower tail** is used.

If H_1 includes a $>$ sign, then the **upper tail** is used.



Critical Region Up Close

Two-tailed tests

Critical region is split over both ends. Both ends contain $\alpha/2$, making a total of α.

If H_1 includes a $\neq$ sign, then the two-tailed test is used as we then look for a change in parameter, rather than an increase or a decrease.

Critical Region Up Close

For each of the scenarios below, identify what type of test you would require.

- SnoreCull hypothesis test as discussed till now.
  One-tailed/Lower-tailed
- If we were checking whether significantly more or significantly fewer than 90% patients had been cured, i.e., H_1: $p \neq 0.9$.
  Two-tailed test
- The coin is biased.
  Two-tailed test
- The coin is biased towards heads with probability 0.8.
  One-tailed/Upper-tailed

The hypothesis test doesn’t answer the question whether the coin is biased or not; it only states whether the evidence is enough to reject the null hypothesis or not at the chosen significance level.