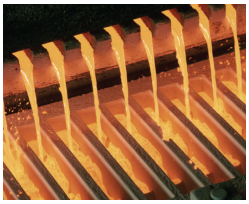
Lesson 35: What is the formal inference process for testing a hypothesis?

Is it 'innocent until proven guilty' or 'guilty until proven innocent'?

What difference does this makes? Explain.

Scenario: Ingot Manufacturing



- metals used in manufacturing cars, planes etc.
- preferred crack free
- cracked ingots are recycled but expensive to do so

One Plant: 80% ingots free of cracks

- wants to reduce cracking proportion (<20%)
- engineers try a new casting process

After 400 ingots made -- 17% cracked

these 400 could be considered a random sample

Is 17% just a product of sampling variability?

How can we state and test a hypothesis about ingot cracking?

Hypothesis - model we adopt temporarily

Null Hypothesis (starting hypothesis) $\left\{\right.$ To test - first assume new process has made no difference and any improvement was just sampling error.

Null Hypothesis H_0: parameter = hypothesized value
that we wish to learn about

For Ingots: H_0: $p = .20$

Alternate Hypothesis (H_A) has value of parameter that we consider possible if we **reject** the null hypothesis.

For Ingots: H_A: $p < .20$ (alternative: the new ingot casting process *reduced* cracking → reason for < symbol

What would it take for you to believe that the cracking rate has actually gone down? (Since 20% and 17% are so close, we should be skeptical)

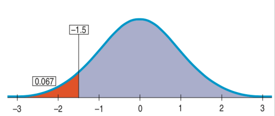
Let's use standard deviations to measure statistical significance!

400 ingots → random sample
$nq=(400)(.8)$
$np=(400)(.2)$ ≥ 10
→ Sampling Distribution Model

$p=.20$, $SD(\hat{p}) = \sqrt{\frac{pq}{n}} = \sqrt{\frac{(.2)(.8)}{400}} = .02$

$$z = \frac{.17-.2}{.02} = -1.5$$

How likely is it to observe a value at least 1.5 SD below the mean of a normal model.



$normalCdf(-\infty, -1.5) = .0668 \approx 6.7\%$

*Decision: Is something that happens 6.7% of the time by chance strong evidence to conclude that true cracking proportion has decreased?

Is 17% surprising??

Innocent until proven guilty

Is there evidence to raise reasonable doubt?

> If there isn't sufficient evidence, the we say not guilty.

not the same as saying innocent

Null Hypothesis H_0: innocent defendent
if evidence puts him/her on spot etc. --- reject null H_0
if not enough evidence --->fail to reject the null, and declare not guilty

P-value

probability value: measure of how surprising a result is

P-value High: events with high probability of happening (usual stuff -- fail to reject null)

P-value Low: says unlikely we'd observe data like these if null hypothesis were true -- make the choice to reject the null hypothesis.

Imagine a clinical trial testing the effectiveness of a new headache remedy.

- compare the treatment with placebo

Null H_0: the new treatment is no more effective than a placebo

- If we use only six people, the results will likely be unclear and we will be unable to reject the hypothesis.

*Does this mean that the drug does't work?

we just don't have enough evidence to reject the hypothesis.

reason why we don't start by assuming the drug is effective. (In unclear situations, we can't just say drug is effective)

ONE-PROPORTION z-TEST

The conditions for the one-proportion z-test are the same as for the one-proportion z-interval. We test the hypothesis H_0: $p = p_0$ using the statistic $z = \frac{(\hat{p} - p_0)}{SD(\hat{p})}$. We use the hypothesized proportion to find the standard deviation, $SD(\hat{p}) = \sqrt{\frac{p_0 q_0}{n}}$.

When the conditions are met and the null hypothesis is true, this statistic follows the standard Normal model, so we can use that model to obtain a P-value.

A 1996 report from the U.S. Consumer Product Safety Commission claimed that at least 90% of all American homes have at least one smoke detector. A city's fire department has been running a public safety campaign about smoke detectors consisting of posters, billboards, and ads on radio and TV and in the newspaper. The city wonders if this concerted effort has raised the local level above the 90% national rate. Building inspectors visit 400 randomly selected homes and find that 376 have smoke detectors. Is this strong evidence that the local rate is higher than the national rate?