Lecture 20: Single Proportion Test

Chapter 6.1

Discussion of Quiz

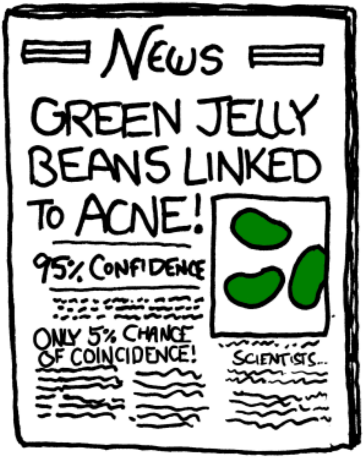
Question 1: Why did $\frac{1}{20}$ studies yield a positive/significant result i.e. that there is a link between jelly beans and acne?

Not that the p-value is 0.05, rather that $\alpha = 0.05$:

- significance level AKA
- type I error rate AKA
- false positive rate

i.e. we expect 1 out of 20 results to be significant even if there is no effect.

Publication Bias



Publication Bias

Publication bias: people only highlight significant/positive results. From Wikipedia: "Publication bias occurs when the publication of research results depends on their nature and direction."

To counter this, some prominent medical journals including

- New England Journal of Medicine
- The Lancet
- Journal of the American Medical Association

require registration of a trial before it starts so that unfavorable results are not withheld from publication.

Journal of Negative Results: `http://www.jnrbm.com/`

Publication Bias

From: Sterne JA, Davey Smith G (2001) Sifting the evidence - What's wrong with significance tests. BMJ 322: 226231.

Multiple Testing

A related issue is the statistical concept of multiple testing.

Say we are conducting many experiments, and H_0 is true for all of them. If you repeat experiments many times, you're bound to get a significant result eventually just by chance alone.

Multiple Testing

What do people do? Make the α stricter! i.e.

- make the α smaller
- i.e. less chance the p-value is smaller than α
- i.e. less chance of incorrectly rejecting H_0 when it is true

Use the Bonferroni correction to α: If you are conducting n tests, use $\alpha^* = \frac{\alpha}{n}$. More later...

Discussion of Quiz

Question 2: Say a very successful entrepreneur named Jamie puts out an autobiography called "How to be a success in life." In it, Jamie details a plan to become a success along the various dimensions of life. Jamie states "I followed these steps, and look at me now! You should do the same!" Critique this statement keeping the comic in mind.

There might have been 9999 people who did the same things but perhaps aren't as successful. Those people generally don't get book deals so we don't know about them.

Relatedly http://bit.ly/1EvIMOG

Question for Today

According to a poll done by the New York Times/CBS News in June 2012, only about 44% of the American public approved of the Supreme Court's performance.

The sample proportion $\widehat{p} = 0.44$ is point estimate of p: the true proportion of the American public who approves.

What are some next things to ask?

- What was n?
- What is the SE of $\widehat{p} = 44\% = 0.44$?
- What is the sampling distribution of $\widehat{p}$?

Question for Today

Just like with $\overline{x}$, if we want to use the normal model to

- build confidence intervals via z^*
- conduct hypothesis tests via the normal tables

we need the sampling distribution of $\widehat{p}$ to be nearly normal.

This happens when the population distribution of 0's and 1's is not too strongly skewed. As the sample size $n \longrightarrow \infty$, this is less of an issue by the CLT.

Note:

$$\widehat{p} = \frac{x_1 + \ldots + x_n}{n} = \frac{1}{n}\sum_{i=1}^{n} x_i$$

where each of the x_i's are 0/1 success/failure Bernoulli random variables.

Conditions for Sampling Dist'n of $\widehat{p}$ Being Nearly Normal

The sampling distribution of the sample proportion $\widehat{p}$ based on sample size n is nearly normal when

- The observations are independent: the 10% rule
- We expect to see at least 10 successes and 10 failures in our sample. This is called the success-failure condition:
  - $np \geq 10$
  - $n(1-p) \geq 10$

Conditions for Sampling Dist'n of $\widehat{p}$ Being Nearly Normal

If conditions are met, then the sampling distribution of $\widehat{p}$ is nearly normal with

- mean p (the true population proportion)
- standard error

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

Note the similarity of the previous formula for the sample mean $\overline{x}$:

$$SE_{\overline{x}} = \frac{\sigma}{\sqrt{n}} = \sqrt{\frac{\sigma^2}{n}}$$

What p to use?

But we don't know what p is. So what p do we use

- to check the success/failure condition?
- for the $SE_{\widehat{p}} = \sqrt{\frac{p(1-p)}{n}}$?

For

- Confidence intervals: plug in the point estimate $\widehat{p}$ of p
- Hypothesis tests: plug in the null value p_0 from $H_0 : p = p_0$

Confidence Intervals

Going back to the poll: $\widehat{p} = 0.44$ based on $n = 976$. What is a 95% confidence interval?

Check the conditions and find SE using $p = \widehat{p}$

- $976 < 10\%$ of 313 million $\Rightarrow$ independence
- Defining a success as a person approving of the job done by the Supreme Court:
  - $976 \times \widehat{p} = 976 \times .44 = 429$ successes ≥ 10
  - $976 \times (1 - \widehat{p}) = 976 \times .56 = 547$ failures ≥ 10

$$SE_{\widehat{p}} = \sqrt{\frac{p(1-p)}{n}} \approx \sqrt{\frac{\widehat{p}(1-\widehat{p})}{n}} = \sqrt{\frac{0.44(1-0.44)}{976}} = 0.016$$

Confidence Intervals

A 95% confidence interval using the normal model has $z^* = 1.96$, thus:

$$\text{point estimate} \pm 1.96 \times SE$$

In our case

$$\widehat{p} \pm 1.96 \times SE_{\widehat{p}} = 0.44 \pm 1.96 \times 0.016 = (0.409, 0.471)$$

Hypothesis Tests

Thomas Carcetti is running for mayor of Baltimore. His campaign manager claims he has more than 50% support of the electorate.

The Baltimore Sun collects a random sample of $n = 500$ likely voters and finds that 52% support him. Does this provide convincing evidence for the claim of Carcetti's manager at the 5% significance level?

Hypothesis Tests

The hypothesis test is, with the null value $p_0 = 0.5$

$$\begin{aligned} & H_0 : p = p_0 \\ \text{vs} \qquad & H_A : p > p_0 \end{aligned}$$

Check the conditions and find SE using $p = p_0$

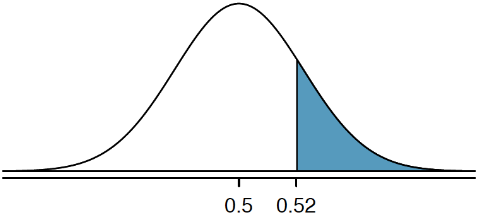
- $500 < 10\%$ of the population of Baltimore $\Rightarrow$ independence
- Success-failure condition
  - $np_0 = 500 \times 0.5 = 250 \geq 10$
  - $n(1 - p_0) = 500 \times (1 - 0.5) = 250 \geq 10$

$$SE_{\widehat{p}} = \sqrt{\frac{p(1-p)}{n}} \approx \sqrt{\frac{p_0(1-p_0)}{n}} = \sqrt{\frac{0.5(1-0.5)}{500}} = 0.022$$

Hypothesis Tests

$$z = \frac{\text{point estimate } \widehat{p} - \text{ null value } p_0}{SE_{\widehat{p}}} = \frac{0.52 - 0.50}{0.022} = 0.89$$

p-value is 0.1867. In the original %'age scale:



Hence we do not reject the null hypothesis, and we do not find convincing evidence to support the campaign manager's claim.

Next Time

Same as with the jump from

$$\mu \quad \text{to} \quad \mu_1 - \mu_2$$

i.e. from one to two-sample tests for means, we make the jump from

$$p \quad \text{to} \quad p_1 - p_2$$

i.e. from one to two-sample tests for proportions.