

Exercise: Chi-Square Test

Human Computer Interaction

Fulvio Corno, Luigi De Russis

Academic Year 2019/2020

Portale della Didattica: A/B Testing

Will the "Community" link lead to significantly more students using the functionality versus the original "Forum" link?



Experimental Design: online A/B test, **randomly** show each student one version of the page. We are measuring the "engagement rate", i.e., how many students open that functionality

Portale della Didattica: Collected Data

	Button on the page	
	Community	Forum
Students use the functionality	30	20
Students didn't use the functionality	70	100
Total visitors:	100	120

Portale della Didattica: Collected Data

	Button on the page	
	Community	Forum
Students use the functionality	30 (30%)	20 (17%)
Students didn't use the functionality	70	100

Total visitors:

100

120

30% is higher than 17%, so we are done!

Right? "Community" is clearly better!

Portale della Didattica: Collected Data

- Not so fast!
- What if this happened just by chance since we had quite few visitors?
- Let's use the Chi-square test to understand whether this difference is significant:
 - *null hypothesis*: _____
 - *alternative hypothesis*: _____

Portale della Didattica: Collected Data

- Not so fast!
- What if this happened just by chance since we had quite few visitors?
- Let's use the Chi-square test to understand whether this difference is significant:
 - *null hypothesis*: the "Community" link will lead to no significant change in the number of students using the functionality versus the original "Forum" link
 - *alternative hypothesis*: the "Community" link will lead to significant more students using the functionality versus the original "Forum" link

Portale della Didattica: Collected Data

	Button on the page		Totals
	Community	Forum	
Students use the functionality	30	20	50
Students didn't use the functionality	70	100	170
Total visitors:	100	120	220

What do we expect to have if the null hypothesis is true?

Portale della Didattica: Collected Data

	Button on the page		Totals
	Community	Forum	
Students use the functionality	30	20	50
Students didn't use the functionality	70	100	170
Total visitors:	100	120	220

Expected

	Button on the page		Totals
	Community	Forum	
Students use the functionality	$(50/220)*100 = 22.7$	$(50/220)*120 = 27.3$	50
Students didn't use the functionality	$(170/220)*100 = 77.3$	$(170/220)*120 = 92.7$	170

Portale della Didattica: Collected Data

	Button on the page		Totals
	Community	Forum	
Students use the functionality	30	20	50
Students didn't use the functionality	70	100	170
Total visitors:	100	120	220

Expected

	Button on the page		Totals
	Community	Forum	
Students use the functionality	22.7	27.3	50
Students didn't use the functionality	77.3	92.7	170

Chi-Square Test: Process

1. Calculate

- $\chi^2 = \sum_{i=1}^n \frac{(O_i - E_i)^2}{E_i}$

- where O_i is the i -th observation and E_i is the expected (theoretical) count of type i

- $$\chi^2 = \frac{(30-22.7)^2}{22.7} + \frac{(20-27.3)^2}{27.3} + \frac{(70-77.3)^2}{77.3} + \frac{(100-92.7)^2}{92.7} =$$
$$2.34 + 1.95 + 0.69 + 0.57 = 5.55$$

Chi-Square Test: Process

2. Determine the degrees of freedom, df , of that statistic:
 - With a single variable, $df = (Cols - 1)$
 - We have one variable with two "columns", so...
 - $df = (2 - 1) = 1$

Chi-Square Test: Process

- Look for the level of confidence (p-value) related to the χ^2 result (5.55) and df (1) in a Probability Table:

df	0.995	0.99	0.975	0.95	0.90	0.10	0.05	0.025	0.01	0.005
1	---	---	0.001	0.004	0.016	2.706	3.841	5.024	6.635	7.879
2	0.010	0.020	0.051	0.103	0.211	4.605	5.991	7.378	9.210	10.597
3	0.072	0.115	0.216	0.352	0.584	6.251	7.815	9.348	11.345	12.838
4	0.207	0.297	0.484	0.711	1.064	7.779	9.488	11.143	13.277	14.860
5	0.412	0.554	0.831	1.145	1.610	9.236	11.070	12.833	15.086	16.750

from <https://people.richland.edu/james/lecture/m170/tbl-chi.html>

- Coin example:
 - first row, $p \approx 0.025$

Chi-Square Test: Process

4. **Sustain** or **reject** the null hypothesis

- we usually reject the null hypothesis at $p < 0.05$ or $p < 0.01$
 - i.e., we are confident that 95% or 99% of the time the test result correctly applies to the entire population
- In our case:
- we can reject the null hypothesis (if we choose $p < 0.05$)!
 - so, we can say that the (imaginary) "Community" link leads to significant more students using the functionality



License

- These slides are distributed under a Creative Commons license “**Attribution-NonCommercial-ShareAlike 4.0 International (CC BY-NC-SA 4.0)**”
- **You are free to:**
 - **Share** — copy and redistribute the material in any medium or format
 - **Adapt** — remix, transform, and build upon the material
 - The licensor cannot revoke these freedoms as long as you follow the license terms.
- **Under the following terms:**
 - **Attribution** — You must give [appropriate credit](#), provide a link to the license, and [indicate if changes were made](#). You may do so in any reasonable manner, but not in any way that suggests the licensor endorses you or your use.
 - **NonCommercial** — You may not use the material for [commercial purposes](#).
 - **ShareAlike** — If you remix, transform, or build upon the material, you must distribute your contributions under the [same license](#) as the original.
 - **No additional restrictions** — You may not apply legal terms or [technological measures](#) that legally restrict others from doing anything the license permits.
- <https://creativecommons.org/licenses/by-nc-sa/4.0/>

