Bivariate Analysis using python

Notes of class

Iván Andrés Trujillo Abella

Facultad de ingeniería

---

1 Analysis

We need test if two variables have some degree of relation, we need specify the kind of variables involved, categorical or numerical.

1.1 Exact fisher Test

1.2 Contingency table

	H	**M**	**Total**
Event A	a	b	a+b
Event B	c	d	c+d
Total	a+c	b+d	

2 Insights

There are a relationship between marginal and joint probabilities

2.1 Null hypothesis

conditional probablities,

Note if H_0 it is true then $P(x, y) = P(x)P(y)$

the event A underwent $a + b$, note also that this ise the sum of marginal frecuencies. the total of individuals with the feature H is $a + c$ notice the are also marginal frecuencies. Note that the probability of a individual underwent the event A is $\frac{a+b}{(a+b)+(c+d)}$ now, the expected number that the feature H appear is $\frac{(a+b)}{(a+b)+(c+d)}(a+c)$. now we can fill the table with expected values and constrast with the observed values.

If we think in the contigency table as a grid, then we can said that f_{ij} is the frecuency in the row i and j column.

Now we can said f_{ij} to observed frecuencies and $\hat{f}_{ij}$ to the expected frecuencies.

To constract hypothesis

$$\chi^2 = \sum_i \sum_j \frac{f_{ij} - \hat{f}_{ij}}{\hat{f}_{ij}} \tag{1}$$

Degree of fredoms are $(rows - 1)(columns - 1)$ think in the number of expected observations to eestimate.

2.1.1 Chi square

observed - expected values.

Distribution, and degree of freedoms

H_0 is independence if you get a smaller p value then reject this hipothesis

restriction: expected cell count in each cell is at least five

What is Chocrans Rule?

All expected values are greater than 1 and at least 80% of expected values are greater than five.

2.2 Exact Fisher Test

One tail p-value.

2.3 Join, marginal and conditional

$F(x, y) = P(X < x, Y < y)$ therefore we can get $f(x, y) = \frac{\partial^2 F}{\partial x \partial y}$

2.4 Conditional

2.5 Uniform distribution

A variable that $X \sim U(a, b)$ means that $l(b - a) = 1$ then getting $l = \frac{1}{b-a}$ Therefore the pdf of uniform distribution is

$$f(x) = \begin{cases} \frac{1}{b-a}, & \text{if } a \leq x \leq b \\ 0, & \text{otherwise} \end{cases} \tag{2}$$

think in that $180° = \frac{\pi}{2} radians$ in the unit circle

2.6 Unconscious Statistician

Search the proof of this theorem.

2.7 Convolution of two formulas

2.8 SUM of continuous random variables

This is important due, Chi squared come from normal distribution:

Here is important to know convolution formula:

3 Laplace Transform

3.1 Potential series

A potential serie is defined as $\sum_{i=0}^{\infty} a_n (x - c)^n$.

for instance remember that in geometric series $\sum_{i=0}^{\infty} ax^i = \frac{a}{1-x}$ if $-1 < x < 1$ the last interval as named the *convergence interval.*