



---

# TRABALHO - PARTE 1

## RELATÓRIO - QUESTÃO 3

---

ELIANE RAMOS DE SIQUEIRA RA:155233

GUILHERME PAZIAN RA:160323

HENRIQUE CAPATTO RA:146406

MURILO SALGADO RAZOLI RA:150987

Disciplina: **ME731 - Análise Multivariada**

Professor: **Caio Lucidius Naberezny Azevedo**

Campinas - SP

18 de Novembro de 2017

# Introdução

Ter uma boa saúde mental é um aspecto importante na construção da vida humana. Sem isto, pode-se pressupor que uma boa condição financeira contribui para existência de uma boa saúde mental. Com este trabalho, desejamos medir o grau de associação entre o estado da saúde mental e o status sócio-econômico do país. o objetivo específico deste trabalho é encontrar um padrão nos dados que indiquem quais estados de saúde mental estão mais relacionados ao status sócio- economico do país através de análise de correspondência simples.

O banco de dados que vamos analisar é constituído de 1760 observações, com duas variáveis categorizadas que são: *Saúde Mental*, dividida em quatro categorias (Boa, Presença fraca de sintomas, Presença moderada de sintomas, Debilitado) e *Status Sócio-Econômico dos Pais*, dividido em cinco ordens (A(Boa),B,C,D e E(Baixo)). Os softwares utilizados foram: *R* <sup>1</sup>, versão 3.4.0 e *Rstudio* <sup>2</sup>, versão 1.0.1.

Primeiramente, realizamos um Teste Qui-Quadrado para testarmos a hipótese de independência entre as duas variáveis citadas acima. Depois, Caso indique-se a dependência, auxiliadas por uma análise descritiva dos perfis médios, procuraremos indícios de relações entre as variáveis. Depois analisaremos a estrutura deste comportamento utilizando análise correspondência (ver Azevedo(2017) e Johnson e Wichern (2007)).

## Análise descritiva

Abaixo, Tabela 1, encontra-se os dados obtidos. O Modelo probabilístico gerador da tabela de contingência é a Multinomial, com tamanho 1760 e 20 categorias.

---

<sup>1</sup><https://cran.r-project.org/>

<sup>2</sup><https://www.rstudio.com/>

	A (Alto)	B	C	D	E (Baixo)	Total
Boa	121	57	72	36	21	307
Presença fraca de Sintomas	188	105	141	97	71	602
Presença Moderada de Sintomas	112	65	77	54	54	362
Debilitado	186	60	94	78	71	489
Total	607	287	384	265	217	1760

Primeiramente, vamos realizar o teste Qui-quadrado para testarmos independência e hipótese que desejamos testar é:

$H_0 : p_{ij} = p_{i.}p_{.j}$  vs  $H_1 : p_{ij} \neq p_{i.}p_{.j}$ , para pelo menos um par  $(i, j)$ , onde  $p_{ij}$  é a proporção de indivíduos na população total para a  $i$ -ésima categoria da variável *Sáude Mental* e da  $j$ -ésima categoria da variável *Status Sócio-Ecônomico do País*.  $p_{i.}$  é a proporção de indivíduos na categoria “i” da variável *Sáude Mental*.  $p_{.j}$  é a proporção de indivíduos na categoria “j” da variável *Status Sócio-Ecônomico do País*.

Os resultados para o teste qui-quadrado usual são: O quantil teórico de uma distribuição qui - quadrado é:  $\chi^2_{12} = 21,03$ . O valor da estatística qui-quadrado é:  $Q = 31,18$  e seu respectivo p-valor é 0,0529. Pelos resultados anteriores vemos que o teste rejeitou a hipótese de que há independência entre as variáveis.

Como discutidos anteriormente, pelo fato do teste qui-quadrado indicar dependência, vamos estudar essa estrutura calculando os perfis médios para cada variável, analisando-as de forma abrangente possível e, posteriormente, aplicando a Análise de Correspondência.

Na Tabela 2, observamos a proporção estimadas pelo total. Pode-se depreender que para 34,49% dos indivíduos vem de países com alto status sócio-econômico. 34,2% dos indivíduos têm presença fraca de sintomas. Podemos ver que 10,68% dos indivíduos tem presença fraca de sintomas e tem pais com alto status sócio-econômico e 10,57 dos indivíduos é debilitados e os pais tem alto status sócio-econômico.

	A (Alto)	B	C	D	E (Baixo)	Total
Boa	6,88	3,24	4,09	2,05	1,19	17,44
Presença fraca de Sintomas	10,68	5,97	8,01	5,51	4,03	34,20
Presença Moderada de Sintomas	6,36	3,69	4,38	3,07	3,07	20,57
Debilitado	10,57	3,41	5,34	4,43	4,03	27,78
Total	34,49	16,31	21,82	15,06	12,33	100,00

Na Tabela 3, há o rankeamento para cada variável. Como podemos observar que