



CURSO DE GRADUAÇÃO EM ENGENHARIA DE SOFTWARE
TRABALHO DA DISCIPLINA ENGENHARIA DE SOFTWARE EXPERIMENTAL

ROBSON PEREIRA DE MORAES

ANÁLISE DE DADOS EXPERIMENTAIS EM ENGENHARIA DE SOFTWARE USANDO O TESTE QUI-QUADRADO

Rio de Janeiro
2024

Sumário

INTRODUÇÃO.....	4
QUI-QUADRADO.....	4
Figura 1: Fórmula qui-quadrado.....	4
<i>Figura 2: Tabela da distribuição Qui-Quadrado - (“Tabela da distribuição Qui-Quadrado”, [s.d.]).....</i>	<i>5</i>
ETAPAS.....	6
Figura 3: Arquivo de dados “csv”.....	6
Figura 4: Resposta_def.py.....	7
Figura 5: Resposta.txt.....	7
RESULTADOS.....	8
REFERÊNCIAS.....	9

ANÁLISE DE DADOS EXPERIMENTAIS EM ENGENHARIA DE SOFTWARE USANDO O TESTE QUI-QUADRADO

Robson Pereira de Moraes – 202022467

Prof. Sidney Loyola de Sá

Universidade de Vassouras

Curso de Engenharia de Software

Trabalho da disciplina Engenharia de Software Experimental

10/05/2024

INTRODUÇÃO

Para a verificação técnica de dados, definimos sua hipótese a ser considerada, em vista que a Engenharia de Software Experimental possui o objetivo estabelecer cenários a serem testados e entender quais resultados serão obtidos a partir de sua coleta de amostras, analisando se uma hipótese do experimento é possível ou nula através dos dados retornados do experimento.

Para que tal estudo seja apresentado, a apresentação do teste Qui-Quadrado será a técnica utilizada neste documento para executar os testes e tornar a apresentação de resultados das hipóteses.

QUI-QUADRADO

Técnica de estatística que possui como objetivo a avaliação de diferenças significativas dentre as frequências esperadas e observadas de uma ou mais categorias, utilizando de uma premissa inicial para analisar as hipóteses a serem respondidas, como o teste de adequação de modelo teórico ou tabela de incontingência.

Sua fórmula de calculo é apresentada da seguinte maneira:

$$\chi^2 = \sum \frac{(O_{ij} - E_{ij})^2}{E_{ij}}$$

Figura 1: Fórmula qui-quadrado

- O_{ij} representa a frequência observadas.
- E_{ij} representa a frequência esperada.

Os graus de liberdade serão os parâmetros a serem considerados para ser representado pelo quadro de Distribuição Qui-Quadrado – Unicaudal à Direita, considerando sua representação será o guia que vai ditar o calculo em conjunto ao Nível de Significância. Caso não seja especificado, o Nível de Significância será considerado 0,05 ou 5% para o quadro.

- $gl = (linhas - 1) * (colunas - 1)$
- $\alpha = P(EI)$

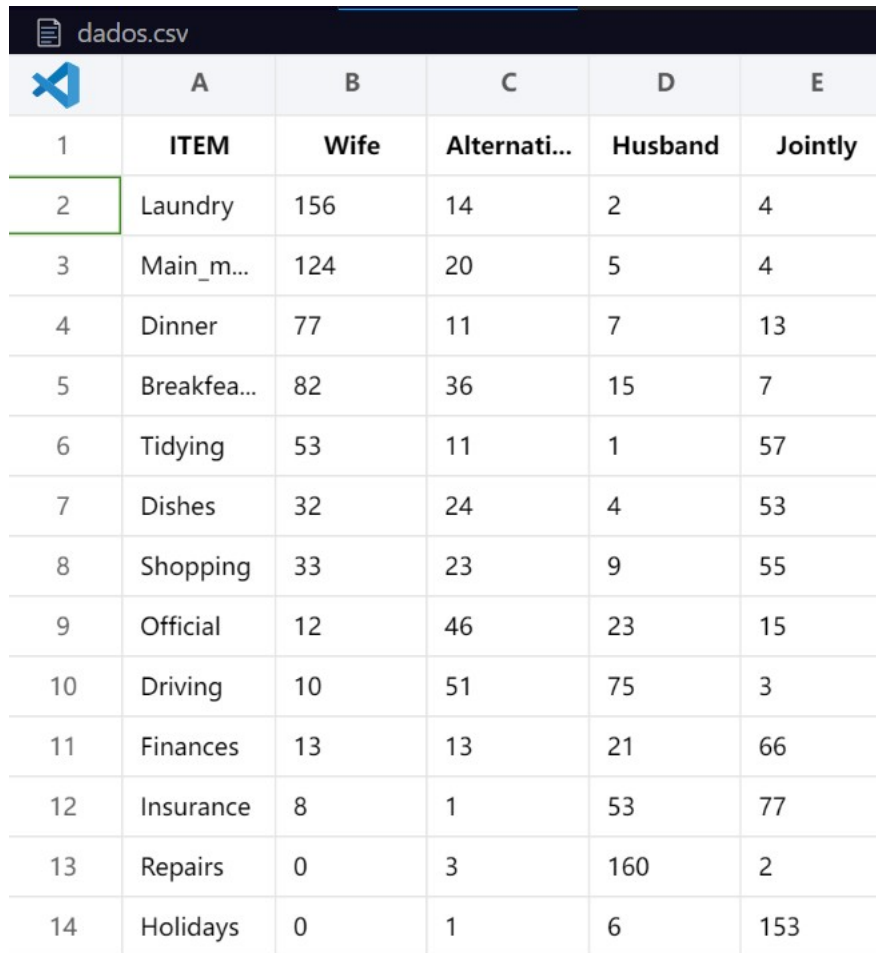
Abaixo segue a tabela o Grau de Liberdade e o Nível de Significância:

Distribuição Qui-Quadrado - Unicaudal à Direita											
Graus de Liberdade	p ►	99,5%	99%	97,5%	95%	90%	10%	5%	2,5%	1%	0,5%
	1	0,000	0,000	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
	6	0,676	0,872	1,237	1,635	2,204	10,645	12,592	14,449	16,812	18,548
	7	0,989	1,239	1,690	2,167	2,833	12,017	14,067	16,013	18,475	20,278
	8	1,344	1,646	2,180	2,733	3,490	13,362	15,507	17,535	20,090	21,955
	9	1,735	2,088	2,700	3,325	4,168	14,684	16,919	19,023	21,666	23,589
	10	2,156	2,558	3,247	3,940	4,865	15,987	18,307	20,483	23,209	25,188
	11	2,603	3,053	3,816	4,575	5,578	17,275	19,675	21,920	24,725	26,757
	12	3,074	3,571	4,404	5,226	6,304	18,549	21,026	23,337	26,217	28,300
	13	3,565	4,107	5,009	5,892	7,042	19,812	22,362	24,736	27,688	29,819
	14	4,075	4,660	5,629	6,571	7,790	21,064	23,685	26,119	29,141	31,319
	15	4,601	5,229	6,262	7,261	8,547	22,307	24,996	27,488	30,578	32,801
	16	5,142	5,812	6,908	7,962	9,312	23,542	26,296	28,845	32,000	34,267
	17	5,697	6,408	7,564	8,672	10,085	24,769	27,587	30,191	33,409	35,718
	18	6,265	7,015	8,231	9,390	10,865	25,989	28,869	31,526	34,805	37,156
	19	6,844	7,633	8,907	10,117	11,651	27,204	30,144	32,852	36,191	38,582
	20	7,434	8,260	9,591	10,851	12,443	28,412	31,410	34,170	37,566	39,997
	21	8,034	8,897	10,283	11,591	13,240	29,615	32,671	35,479	38,932	41,401
	22	8,643	9,542	10,982	12,338	14,041	30,813	33,924	36,781	40,289	42,796
	23	9,260	10,196	11,689	13,091	14,848	32,007	35,172	38,076	41,638	44,181
	24	9,886	10,856	12,401	13,848	15,659	33,196	36,415	39,364	42,980	45,559
	25	10,520	11,524	13,120	14,611	16,473	34,382	37,652	40,646	44,314	46,928
	26	11,160	12,198	13,844	15,379	17,292	35,563	38,885	41,923	45,642	48,290
	27	11,808	12,879	14,573	16,151	18,114	36,741	40,113	43,195	46,963	49,645
	28	12,461	13,565	15,308	16,928	18,939	37,916	41,337	44,461	48,278	50,993
	29	13,121	14,256	16,047	17,708	19,768	39,087	42,557	45,722	49,588	52,336
	30	13,787	14,953	16,791	18,493	20,599	40,256	43,773	46,979	50,892	53,672
	31	14,458	15,655	17,539	19,281	21,434	41,422	44,985	48,232	52,191	55,003
	32	15,134	16,362	18,291	20,072	22,271	42,585	46,194	49,480	53,486	56,328
	33	15,815	17,074	19,047	20,867	23,110	43,745	47,400	50,725	54,776	57,648
	34	16,501	17,789	19,806	21,664	23,952	44,903	48,602	51,966	56,061	58,964
	35	17,192	18,509	20,569	22,465	24,797	46,059	49,802	53,203	57,342	60,275
	36	17,887	19,233	21,336	23,269	25,643	47,212	50,998	54,437	58,619	61,581
	37	18,586	19,960	22,106	24,075	26,492	48,363	52,192	55,668	59,893	62,883
	38	19,289	20,691	22,878	24,884	27,343	49,513	53,384	56,896	61,162	64,181
	39	19,996	21,426	23,654	25,695	28,196	50,660	54,572	58,120	62,428	65,476
	40	20,707	22,164	24,433	26,509	29,051	51,805	55,758	59,342	63,691	66,766
	45	24,311	25,901	28,366	30,612	33,350	57,505	61,656	65,410	69,957	73,166
	50	27,991	29,707	32,357	34,764	37,689	63,167	67,505	71,420	76,154	79,490
	55	31,735	33,570	36,398	38,958	42,060	68,796	73,311	77,380	82,292	85,749
	60	35,534	37,485	40,482	43,188	46,459	74,397	79,082	83,298	88,379	91,952
	70	43,275	45,442	48,758	51,739	55,329	85,527	90,531	95,023	100,425	104,215
	80	51,172	53,540	57,153	60,391	64,278	96,578	101,879	106,629	112,329	116,321
	90	59,196	61,754	65,647	69,126	73,291	107,565	113,145	118,136	124,116	128,299
	100	67,328	70,065	74,222	77,929	82,358	118,498	124,342	129,561	135,807	140,169
	110	75,550	78,458	82,867	86,792	91,471	129,385	135,480	140,917	147,414	151,948
	120	83,852	86,923	91,573	95,705	100,624	140,233	146,567	152,211	158,950	163,648
	130	92,222	95,451	100,331	104,662	109,811	151,045	157,610	163,453	170,423	175,278

Figura 2: Tabela da distribuição Qui-Quadrado - (“Tabela da distribuição Qui-Quadrado”, [s.d.])

ETAPAS

Foi utilizado um arquivo de dados “csv” contendo os seguintes dados:



	A	B	C	D	E
1	ITEM	Wife	Alternati...	Husband	Jointly
2	Laundry	156	14	2	4
3	Main_m...	124	20	5	4
4	Dinner	77	11	7	13
5	Breakfea...	82	36	15	7
6	Tidying	53	11	1	57
7	Dishes	32	24	4	53
8	Shopping	33	23	9	55
9	Official	12	46	23	15
10	Driving	10	51	75	3
11	Finances	13	13	21	66
12	Insurance	8	1	53	77
13	Repairs	0	3	160	2
14	Holidays	0	1	6	153

Figura 3: Arquivo de dados “csv”

Os dados recebidos foram importados para o “python”, utilizando a biblioteca “scipy.stats” com o modulo “chi2_contingency”, organizados em sequência de listas como variável “data”, tratados pelo modulo em “chi2, p, dof, expected = chi2_contingency(data)” e demonstra o resultado no terminal ao executar, simultaneamente escrevendo os resultados para um arquivo txt “Resposta.txt” de acordo com a ilustração abaixo do código:

```

1 from scipy.stats import chi2_contingency
2 import numpy as np
3
4
5 data = np.array([[156, 124, 77, 82, 53, 32, 33, 12, 10, 13, 8, 0, 0],
6                 [14, 20, 11, 36, 11, 24, 23, 46, 51, 13, 1, 3, 1],
7                 [2, 5, 7, 15, 1, 4, 9, 23, 75, 21, 53, 160, 6],
8                 [4, 4, 13, 7, 57, 53, 55, 15, 3, 66, 77, 2, 153]])
9
10 chi2, p, dof, expected = chi2_contingency(data)
11
12 with open("Resposta.txt", "w") as f:
13     f.write(f"Chi-squared statistic: {chi2:.13f}\n")
14     f.write(f"P-value: {p:.2f}\n")
15     f.write(f"Degrees of freedom: {dof:.2f}\n")
16     f.write(f"Expected frequencies: \n")
17     np.savetxt(f, expected, fmt='%1.8f')
18
19 print(f"Chi-squared statistic: {chi2}")
20 print(f"P-value: {p}")
21 print(f"Degrees of freedom: {dof}")
22 print(f"Expected frequencies: \n, {expected}")

```

Figura 4: Resposta_def.py

```

1 Resposta_def.py
2
3 Chi-squared statistic: 1944.4561959955277
4 P-value: 0.0
5 Degrees of freedom: 36
6 Expected frequencies:
7 [[60.55045872 52.63761468 37.1559633 48.16513761 41.97247706 38.87614679
8 41.28440367 33.02752294 47.82110092 38.87614679 47.82110092 56.76605505
9 55.04587156]
10 [25.63302752 22.28325688 15.7293578 20.38990826 17.76834862 16.45756881
11 17.47706422 13.98165138 20.24426606 16.45756881 20.24426606 24.0309633
12 23.30275229]
13 [38.44954128 33.42488532 23.5940367 30.58486239 26.65252294 24.68635321
14 26.21559633 20.97247706 30.36639908 24.68635321 30.36639908 36.04644495
15 34.95412844]
16 [51.36697248 44.65424312 31.5206422 40.86009174 35.60665138 32.97993119
17 35.02293578 28.01834862 40.56823394 32.97993119 40.56823394 48.1565367
18 46.69724771]]

```

Figura 5: Resposta.txt

RESULTADOS

A questão definida é se há uma associação o tipo de pessoa e as tarefas executadas e descobrir a resposta da Formulação da Hipótese Nula e Alternativa:

- Hipótese Nula (H_0): Não há associação entre o tipo de pessoa (esposa, marido, alternadamente e em conjunto) e o tipo de tarefa doméstica.
- Hipótese Alternativa (H_1): Há uma associação entre o tipo de pessoa e o tipo de tarefa doméstica.

Podemos rejeitar a hipótese nula e concluir que há uma associação significativa entre o tipo de pessoa e o tipo de tarefa doméstica.

CONCLUSÃO

As pesquisas efetuadas para este documento expressam a vantagem do estudo de hipóteses e projetar ideias que necessitem ou não de ambientes controlados para esboçar um cenário onde todas as possibilidades possíveis devem ser consideradas, mesmo que pareçam não ser algo fácil de se observar. Toda linha de raciocínio deve ser analisada com todo o cuidado e hipótese segue uma premissa para que possam existir alternativas, ou seja, não existe apenas um resultado concreto, em vista que o contexto pode ou não invalidar uma ideia.

REFERÊNCIAS

Tabela da distribuição Qui-Quadrado. Disponível em:

<<https://www.professorguru.com.br/estatistica/tabela-qui-quadrado.html>>. Acesso em: 8 maio. 2024.

ROBSON. **HunterLowBit/Eng_Sof_EXP.** , 10 maio 2024. Disponível em:

<https://github.com/HunterLowBit/Eng_Sof_EXP>. Acesso em: 10 maio. 2024