

MAE116 – Noções de Estatística

Grupo BD – 2º semestre de 2020

Lista de exercícios 11 Qui-quadrado – C L A S S E - GABARITO

Exercício 1

Um grande banco sabe que, historicamente, 20% de seus clientes aplicam em caderneta de poupança, 35% aplicam em CDB, 18% aplicam em fundos de investimentos, 12% em fundos mútuos de ações e o restante em outras aplicações, tais como, títulos da dívida pública ou privada, entre outras. Com a instabilidade econômica e política durante os últimos anos, a diretoria do banco deseja saber se houve alteração no tipo de aplicação de seus clientes, a fim de estabelecer suas metas de vendas para o próximo período. Uma amostra aleatória de 1000 clientes forneceu os seguintes dados:

Poupança	CDB	FI	Ações	Outras
190	305	240	130	135

- (a) Dê uma estimativa pontual para a proporção atual de clientes desse banco que investem em poupança.

$$\frac{190}{1000} = 0,19$$

- (b) Se a participação dos clientes nos diversos produtos oferecidos pelo banco não se alterou, quantos clientes dessa amostra seriam esperados investindo em poupança? E em ações? Quantos foram observados em cada caso?

Número esperado de clientes, sob a hipótese de não ter havido mudança.

Poupança	CDB	FI	Ações	Outras
200	350	180	120	150
$0,2 \times 1000$	$0,35 \times 1000$	$0,18 \times 1000$	$0,12 \times 1000$	$0,15 \times 1000$

- (c) Utilizando um procedimento estatístico adequado, pode-se afirmar que a participação dos clientes nos diferentes produtos oferecidos por esse banco não se alterou nesses últimos anos, em relação ao que ocorria historicamente? Utilize o nível descritivo e adote nível de significância de 5%.

$$\chi_{obs}^2 = \frac{(190 - 200)^2}{200} + \frac{(305 - 350)^2}{350} + \frac{(240 - 180)^2}{180} + \frac{(130 - 120)^2}{120} + \frac{(135 - 150)^2}{150} \\ = 0,5 + 5,76 + 20 + 0,83 + 1,5 = 28,59$$

Nível descritivo: **valor-p** = $P(\chi_4^2 \geq 28,59) \cong 0,000 < \alpha = 0,05 \Rightarrow H_0$ é rejeitada.

Conclusão: Ao nível de significância de 5%, há evidências amostrais para concluir que a distribuição das aplicações dos clientes desse banco, não segue o padrão histórico, nos tempos atuais.

- (d) Construa um intervalo de confiança de 95% para a proporção de clientes desse banco que nos últimos anos aplicam em fundos de investimentos (FI).

$$IC(95\%) = 0,24 \pm 1,96 \sqrt{\frac{0,24 \times 0,76}{1000}} = 0,240 \pm 0,026 = [0,214; 0,266]$$

Obs.: Note que este IC não contém o valor histórico (18%), o que indica haver evidência amostral significativa de mudança nesta proporção. Esse resultado era esperado, pois foi a categoria que produziu a maior contribuição individual (20) no valor do χ_{obs}^2 . O mesmo deve ocorrer com a aplicação em CDB (contribuição CDB = 5,76).

MAE116 – Noções de Estatística

Grupo BD – 2º semestre de 2020

Lista de exercícios 11 Qui-quadrado – C L A S S E - GABARITO

Exercício 2

Uma pesquisa realizada entre estudantes da USP (Campus da Capital, Faculdade de Direito e Complexo de Saúde) foi conduzida no ano de 2011, sendo que um dos objetivos era verificar a existência de associação entre a área do curso (Exatas, Biológicas e Humanas) e o consumo de bebida alcoólica em mais de 2 dias por semana (Sim; Não). Para isso, uma amostra aleatória de 3715 alunos foi selecionada e os alunos responderam um questionário. Dos 958 alunos consultados de cursos da área biológica (exceto medicina), 630 responderam que consomem bebida alcoólica em mais de 2 dias por semana. Dentre os alunos de ciências exatas, 450 disseram não ingerirem bebida alcoólica mais de 2 dias por semana e 970 dos 1373 alunos de cursos de ciências humanas consultados relataram que fazem uso de bebida alcoólica em mais de 2 dias por semana.

(a) Escreva as informações da pesquisa em uma tabela de contingência.

Consumo \ Área	Biológicas	Exatas	Humanas	Total
Sim	630	934	970	2534
Não	328	450	403	1181
Total	958	1384	1373	3715

(b) Qual é uma estimativa para a proporção de alunos da USP que consomem bebida alcoólica em mais de 2 dias por semana? Dentre os alunos de ciências exatas, qual é essa proporção?

$$\hat{p}_A = \frac{2534}{3715} = 0.682; \quad \hat{p}_{A|E} = \frac{934}{1384} = 0.675$$

(c) Se não há associação entre área do curso e consumo de bebida alcoólica, quantos alunos do curso de ciências humanas esperaríamos que relatassem ingerir bebida alcoólica em mais de 2 dias por semana?

$$E_{HA} = \frac{1373 \times 2534}{3715} = 936.52$$

(d) Formule as hipóteses H_0 e H_1 de um teste de hipóteses estatístico adequado a esta situação.

H_0 : Consumo de bebida alcoólica e área do curso são independentes

H_1 : Existe associação (dependência) entre consumo de bebida alcoólica e área do curso

(e) Utilizando o valor do nível descritivo resultante, conclua o teste, adotando um nível de significância de 1%.

Colocando os dados no Rcmdr:

Estatísticas → Tabelas de Contingência → Digite e analise tabela de dupla entrada

Nível descritivo: **valor-p = 0,03427**

Conclusão: Como $\text{valor-p} > \alpha = 0,01$, então H_0 (variáveis são independentes) não é rejeitada, ou seja, não há evidência amostral suficiente, ao nível de significância de 1%, para afirmar que o padrão de consumo de bebida alcoólica difere por área do curso do estudante.

Exercício 3

Um questionário foi aplicado a 45 voluntários de uma ONG fornecendo informação sobre tempo de trabalho na ONG (até 3 anos ou mais de 3 anos) e grau de escolaridade completado (Fundamental - F, Médio - M ou Superior - S), sendo obtidos os seguintes dados:

MAE116 – Noções de Estatística

Grupo BD – 2º semestre de 2020

Lista de exercícios 11 Qui-quadrado – C L A S S E - GABARITO

Até 3 anos: *F F F S M M F F M M M F M S S F F F M S S S M M S*

Mais de 3 anos: *M M M S M S S F M F M M S S S S F M M M*

(a) Identifique as variáveis em estudo e classifique-as.

Tempo de ONG: qualitativa ordinal; Escolaridade: qualitativa ordinal

(b) Construa uma tabela de contingência para as variáveis do levantamento.

Tempo \ Escolaridade	F	M	S	Total
Até 3 anos	9	9	7	25
Mais de 3 anos	3	10	7	20
Total	12	19	14	45

(c) Se a formação acadêmica do voluntário não está associada ao tempo de *ONG*, quantos voluntários com até 3 anos de empresa seriam esperados com curso fundamental? E quantos com mais de 3 anos de *ONG* e curso superior? Quantos foram observados em cada caso?

$$E_{Até\ 3,F} = \frac{12 \times 25}{45} = 6,67; \quad E_{+3,S} = \frac{14 \times 20}{45} = 6,22$$

$$O_{Até\ 3,F} = 9; \quad O_{+3,S} = 7$$

(d) Conclua, utilizando algum procedimento estatístico, se existe associação entre grau de escolaridade e tempo de *ONG* do funcionário. Utilize $\alpha = 5\%$.

⇒ Teste de independência

Hipóteses:

H_0 : Tempo de *ONG* e grau de escolaridade dos voluntários são independentes

H_1 : Tempo de *ONG* e grau de escolaridade dos voluntários estão associados (não são independentes)

⇒ **Rcmdr**: *Estatísticas* → *Tabelas de Contingência* → *Digite e analise tabela de dupla entrada*

Saída do R: Pearson's Chi-squared test

data: .Table

X-squared = 2.5283, df = 2, **p-value = 0.2825**

Como valor- $p > \alpha=0,05$, então H_0 (variáveis são independentes) não é rejeitada, ou seja, não há evidência amostral suficiente, ao nível de significância de 5%, para afirmar que o tempo de *ONG* esteja associado ao grau de escolaridade do voluntário.