

# CRIPTOANÁLISE COMO PROPOSTA DIDÁTICA PARA O ENSINO DE ESTATÍSTICA

## CRYPTANALYSIS AS DIDACTIC PROSPOTE FOR TEACHING STATISTICS

**Deivison Porto de Sousa**

Instituto Federal do Rio Grande do Sul,  
deivison.sousa@vacaria.ifrs.edu.br

**Jandresson Dias Pires**

Instituto Federal do Norte de Minas Gerais  
jandresson.pires@ifnmg.edu.br

### Resumo

O objetivo deste artigo é apresentar uma proposta didática para o ensino da estatística na Educação Básica de forma lúdica por meio de resolução de problemas e atividades investigativas. A proposta consiste na elaboração de uma tabela de frequência das letras na língua portuguesa para então decifrar um texto cujas letras foram trocadas com intuito de esconder seu conteúdo. Os textos utilizados para a construção da tabela parâmetro serão notícias de fontes diversas que podem, dessa forma, facilitar o processo de interdisciplinaridade. Espera-se que os alunos se sintam mais envolvidos e dispostos a participar de uma atividade em grupo e percebam a importância da adoção de medidas adequadas para coleta de dados, organização em tabelas e representações gráficas de forma pertinente além da necessidade de estabelecer estratégias variadas para solução de problemas por não poder interpretar os dados estatísticos com garantia de exatidão.

**Palavras-chave:** Criptoanálise, estatística e frequência.

### Abstract

The purpose of this article is to present a didactic proposal for the teaching of statistics in Basic Education in a playful way through problem solving and research activities. The proposal consists in the elaboration of a table of frequency of the letters in the Portuguese language to decipher a text whose letters have been exchanged in order to hide its content. The texts used for the construction of the parameter table will be news from diverse sources that can, in this way, facilitate the process of interdisciplinarity. Students are expected to feel more involved and willing to participate in a group activity and realize the importance of adopting appropriate measures for data collection, organization of tables and graphical representations in a relevant way, and the need to establish varied strategies for problems of not being able to interpret the statistical data with guarantee of accuracy.

**Keywords:** Cryptanalysis, statistics and frequency.

## Introdução

Apesar da inegável importância e utilidade da estatística em diversas áreas do conhecimento, ela costuma estar restrita, na Educação Básica, a um mero conteúdo da disciplina de matemática em um dos anos do Ensino Médio e, normalmente, não vai muito além da construção de tabelas e estudos de gráficos. No ensino superior, ainda que comumente seja uma disciplina específica na grade, continua sendo associada pura e simplesmente a uma parte da matemática e pode não ser vista como nada além de um pesadelo para muitos estudantes.

Visto que todas as áreas do conhecimento em algum momento fazem levantamento de informações e coleta de dados, não é difícil justificar, em qualquer que seja a área, a importância da estatística na formação profissional e acadêmica dos alunos, pois, segundo Andrade (2013), a estatística é uma ciência que se preocupa com o planejamento de uma pesquisa, envolvendo desde a forma como os dados são coletados à organização e avaliação dos mesmos e, assim, o pesquisador deve possuir razoáveis conhecimentos em estatística ou consultar um estatístico para auxiliá-lo no desenvolvimento de suas pesquisas.

Muitos são os fatores que justificam o ensino da estatística desde os anos iniciais da Educação Básica, e autores e pesquisadores defendem que esse ensino seja feito por meio da resolução de problemas. Lopes e Souza (2016, p. 1469), assegura que: “Um dos aspectos que entendemos ser relevantes ao ensino da Estatística é ensinar através da resolução de problemas e processos investigativos”. Para Souza et al. (2013), é preciso estar preparado para identificar as ideias fundamentais em probabilidade e estatística abordando esses temas nas aulas de matemática a partir do interesse dos alunos por meio de atividades investigativas, para que estes aprendam a investigar e tirar conclusões próprias, não se deixando influenciar futuramente por informações tendenciosas.

Partindo desse ponto, pensamos na elaboração de uma atividade que permita aliar o interesse comum despertado por mensagens secretas com a coleta de dados, organização e análise dos mesmos que são etapas características dos levantamentos estatísticos. Couto (2008), afirma que desde pequenos somos acostumados a montar cifras bastante básicas e que, no tocante à privacidade, embora utilizemos muitas delas nos dias de hoje, essa é uma preocupação que vem desde a antiguidade com a criação e desaparecimento de muitas cifras e códigos ao longo da história. Ainda segundo Couto (2008), a análise estatística é uma fonte importante para a criptoanálise e cita como exemplo o estudo da frequência da ocorrência de letras para identificar o idioma usado na mensagem cifrada além de outros detalhes, como, por exemplo, o fato de, na língua portuguesa, a letra que Q vir sempre acompanhada pela letra U (Qua, Que, Qui e Quo).

A atividade consiste basicamente em analisar um conjunto de textos individualmente e depois juntar os dados obtidos para formar uma tabela que indique quais letras têm mais chance de aparecer em um texto na língua portuguesa, similarmente à Tabela 1 levantada, de acordo com Couto (2008), pela analista de sistemas Viktoria Tkotz em seu site Criptografia Numaboa, após a análise de textos de domínio público disponíveis na internet que totalizaram 157764 palavras compostas por 725511 letras. Após a organização desses dados, de forma conveniente, sua análise permitirá a “tradução” de um texto em que cada letra do texto original tenha sido substituída por outra, tornando-o incompreensível até que se descubra a correspondência entre as letras.

Tabela 1 - Percentual das frequências de letras

<b>Letra</b>	<b>A</b>	<b>E</b>	<b>O</b>	<b>I</b>	<b>R</b>	<b>S</b>	<b>N</b>	<b>D</b>	<b>M</b>	<b>T</b>	<b>C</b>	<b>U</b>	<b>L</b>
<b>Frequência (%)</b>	14,63	1,04	3,88	4,99	12,57	1,02	1,30	1,28	6,18	0,40	0,02	2,78	4,74
<b>Letra</b>	<b>P</b>	<b>F</b>	<b>G</b>	<b>V</b>	<b>B</b>	<b>H</b>	<b>Q</b>	<b>J</b>	<b>Z</b>	<b>Y</b>	<b>K</b>	<b>W</b>	<b>X</b>
<b>Frequência (%)</b>	5,05	10,73	2,52	1,20	6,53	7,81	4,34	4,63	1,67	0,01	0,21	0,01	0,47

Fonte: Couto 2008, p. 73 – adaptado pelos autores.

No entanto, a criptoanálise não é uma tarefa simplória e o simples estudo da frequência das letras não será o suficiente para quebrar códigos nos dias atuais. A orientação do professor se faz necessária em cada etapa, impedindo que sejam tomados caminhos errôneos que acabem por levar ao desinteresse e justificando a importância da estatística sem, contudo, colocá-la como uma solução exata para todos os problemas.

### **Procedimento metodológico para a sequência didática desenvolvida com a criptoanálise**

Com o intuito inicial de despertar a curiosidade dos alunos será entregue um texto para a turma como o apresentado no Quadro 1:

Quadro 1 – Texto codificado

*TEKDKREKRS D T D SRTOSRD FNT NKRQRXD ET UDE KTWJRDE  
 MJWPDPRQREKRSDE MDJD TLMQRSDJ D HJT FNTOSRD UD WSWJJTOSRD UT  
 TYTOKWE KDOKW TV TEKNUWE WPETJYDSRWODRE FNDOKW  
 TV TLMTJRVTOKWE MDJD VWUTQDJ D DQTDKWJRTUDUT T D ROSTJKTXD UT  
 HWJVD D TEKRVDJ WN MWEERPRQRKDJ D MJTYREDW UT HTOWVTOWE  
 HNKNJWE SWOHWJVT W SDEW.*

Possivelmente, ficarão surpresos. Não saberão do que se trata e o que deverá ser feito. Cabe ao professor informar que receberam a missão de decifrar a mensagem. Presumivelmente, alguns alunos já tenham feito ou visto brincadeiras semelhantes e com um pouco de busca na internet, podem tentar usar a Cifra de César<sup>1</sup> para quebrar o código. É importante salientar que acentos e pontuações foram ignorados.

Após alguns minutos para sugestões e tentativas o professor poderá intervir falando sobre a importância dos levantamentos estatísticos e da relação do estudo da frequência das letras com a criptoanálise.

A sequência da atividade consiste em dividir a turma em grupos e encontrar notícias, reportagens para cada um desses grupos. Em um primeiro momento, cada grupo deve fazer uma análise da frequência das letras no texto.

As notícias escolhidas para o exemplo abaixo têm um tamanho diminuto, mas tanto sua extensão quanto conteúdo podem variar para atender uma necessidade de

<sup>1</sup> **Cifra de César** - Em criptografia, a Cifra de César, também conhecida como cifra de troca, código de César ou troca de César, é uma das mais simples e conhecidas técnicas de criptografia. É um tipo de cifra de substituição na qual cada letra do texto é substituída por outra que está a um número fixo à frente.

interdisciplinaridade, maior precisão através de uma quantidade maior de dados ou, até mesmo, conveniência e facilidade de manuseio em virtude das ferramentas disponíveis.

Para o nosso exemplo, utilizaremos três textos curtos de fontes distintas, porém, salientamos que o trabalho se torna mais interessante com um número maior de grupos e, por consequência, textos.

#### Texto 1

*Duas meninas, uma com menos de um ano de idade e a outra com dois anos, foram resgatadas pelo Conselho Tutelar de Santos, no litoral de São Paulo, após vizinhos denunciarem a mãe das crianças. Ao chegar à residência na manhã desta terça-feira (18) os profissionais constataram a situação de abandono e levaram as duas crianças para um abrigo.*

*De acordo com informações da Delegacia da Mulher de Santos, onde o caso foi inicialmente tratado, essa não seria a primeira vez que a mãe abandona as filhas sozinhas em casa, no bairro Chico de Paula, enquanto sai durante a noite. As crianças foram encaminhadas até a delegacia e posteriormente foram levadas a um abrigo. Até o momento a mãe das crianças não foi localizada.*

Fonte: Página do G1, 2014.

Antes de partirmos para uma análise definitiva da frequência das letras, precisamos adotar algumas convenções para evitar interpretações dúbias. Por exemplo, letras acentuadas não serão consideradas diferentes umas das outras em virtude do acento e o c com cedilha será interpretado como C. Tendo isso em vista, a análise de frequência do texto acima está representada na Tabela 2.

Tabela 2 – Percentual da frequência de letras no primeiro texto

Letra	A	B	C	D	E	F	G	H	I	J	K	L	M
Frequência (%)	19,44	0,87	5,03	5,56	9,55	1,56	1,04	1,56	6,94	0,00	0,00	2,78	4,86
Letra	N	O	P	Q	R	S	T	U	V	W	X	Y	Z
Frequência (%)	7,12	9,90	1,39	0,35	5,90	7,99	3,99	2,78	0,69	0,00	0,00	0,00	0,69

Fonte: elaborado pelos autores.

#### Texto 2

*Após o Vaticano negar que o papa Francisco estaria trabalhando para mediar a crise entre os Estados Unidos e Coreia do Norte, o Pontífice disse nesta segunda-feira (30) que "a humanidade corre risco de suicídio" fazendo referência ao uso de armas nucleares.*

*Durante visita a sede do Dicastério para o Desenvolvimento Humano Integral, a subsecretária Flaminia Giovanelli, afirmou à rádio Vaticano, que ao falar sobre armas nucleares, a declaração de Francisco foi repleta de "palavras muito fortes".*

*"O papa ressaltou novamente, algo que faz habitualmente, o fenômeno do comércio das armas. Ele reiterou a denúncia desse comércio que estimula e sustenta esses conflitos, que não são surtos, mas, sim uma guerra real", acrescentou Giovanelli.*

*Para Jorge Mario Bergoglio, "a ameaça que, infelizmente, esteve presente há décadas, mas às vezes se torna mais aguda na atualidade", é um risco ao "suicídio da humanidade", finalizou a subsecretária do Dicastério.*

Fonte: Página do Jornal do Brasil, 2017

Tabela 3 – Percentual das frequências de letras no segundo texto

Letra	A	B	C	D	E	F	G	H	I	J	K	L	M
<b>Frequência (%)</b>	15,05	0,78	4,28	4,80	12,97	2,08	1,43	0,78	7,78	0,13	0,00	3,24	3,37
Letra	N	O	P	Q	R	S	T	U	V	W	X	Y	Z
<b>Frequência (%)</b>	5,58	8,82	1,56	0,91	7,26	7,78	4,67	4,67	1,43	0,00	0,00	0,00	0,65

Fonte: elaborado pelos autores.

### Texto 3

*O ministro do Planejamento, Dyogo de Oliveira, afirmou que o governo não tem "liberdade" para modificar o salário mínimo de forma diferente do que determina a legislação.*

*Na mensagem modificativa para o Orçamento de 2018, anunciada hoje, o Ministério diminuiu a projeção para o salário mínimo de R\$ 969 para R\$ 965 no ano que vem. Segundo ele, a diminuição ocorreu por conta da menor projeção para o INPC - que serve, junto com o PIB, de base de cálculo para o mínimo.*

*Segundo o ministro, trata-se apenas de uma projeção sobre o salário mínimo - e não da mudança em si - e a alteração de fato será feita em janeiro do ano que vem. "O governo não tem liberdade para definir salário mínimo. Ele será definido em janeiro, na forma prevista pela lei. A lei determina o salário mínimo", disse.*

Fonte: (PUPO, SIMÃO e GRANER, 2017, p.1)

Tabela 4 – Percentual das frequências de letras no terceiro texto

Letra	A	B	C	D	E	F	G	H	I	J	K	L	M
<b>Frequência (%)</b>	13,79	0,83	1,50	5,48	12,29	1,66	1,16	0,17	9,80	1,33	0,00	3,16	6,48
Letra	N	O	P	Q	R	S	T	U	V	W	X	Y	Z
<b>Frequência (%)</b>	6,98	12,62	2,82	0,83	7,81	3,82	3,16	2,82	1,33	0,00	0,00	0,17	0,00

Fonte: elaborado pelos autores.

Os grupos com a posse de sua respectiva tabela, podem iniciar uma das partes mais interessantes da atividade que é a ampliação da base de dados por meio da junção dos resultados de todos os grupos. Para fazer a tabela de seu texto, cada grupo teve de verificar a quantidade de vezes que cada letra aparece em sua respectiva notícia e, por consequência, obter o total de caracteres. Somando-se os totais de cada um dos grupos, obteremos uma tabela que, teoricamente, será melhor que as observações individuais e convergirá para os valores "ideais" à medida que a base de dados se amplia. O resultado, para os três textos analisados, se encontra na tabela 5.

Tabela 5 – Percentual das junções das frequências de letras dos textos

Letra	A	B	C	D	E	F	G	H	I	J	K	L	M
<b>Frequência (%)</b>	15,92	0,82	3,65	5,24	11,76	1,80	1,23	0,82	8,17	0,46	0,00	3,08	4,78
Letra	N	O	P	Q	R	S	T	U	V	W	X	Y	Z
<b>Frequência (%)</b>	6,47	10,32	1,90	0,72	7,04	6,63	3,95	3,54	1,18	0,00	0,00	0,05	0,46

Fonte: elaborado pelos autores.

Comparar a Tabela 5 com as tabelas que apresentam os valores encontrados em pesquisas na internet com a Tabela 1 nos possibilita realizar algumas observações

interessantes: enquanto a letra A esteve muito acima do valor de referência no primeiro texto, apresentou uma queda no terceiro e se aproximou mais do valor esperado no final. As letras C, G, J, M e O apresentam aproximações muito boas na tabela final, melhores do que quaisquer outras obtidas individualmente. Já para as letras R e S que são as consoantes mais frequentes nos textos temos um caso curioso: enquanto os dois primeiros textos tiveram “boas” aproximações, o terceiro apresentou uma discrepância muito grande para a ocorrência do S, para menos, e uma variação considerável do R para mais. Tamanho descompasso fez nossa tabela final inverter o posicionamento das duas letras, dando a impressão de que o R seria mais frequente do que o S.

Possivelmente, com um número maior de textos ou textos mais longos, teríamos um resultado mais próximo do esperado, porém, esse tipo de situação nos dá a oportunidade de fazer análises sobre a não exatidão das observações estatísticas, da necessidade de buscar estratégias diversificadas para soluções de problemas e da constante busca por melhorias nas observações.

Algumas questões podem ser levantadas neste momento: será que conseguimos decifrar o texto apenas com a tabela que produzimos? A organização das letras em ordem alfabética é adequada para o objetivo proposto ou podemos encontrar uma forma melhor? A confecção de gráficos seria interessante?

Para responder a primeira indagação, inicialmente, responderemos as outras duas apresentando a Tabela 6, que organiza as letras por ordem de frequência, e a Figura 1, que mostra um gráfico gerado com base nessa tabela.

Tabela 6 – Tabela 5 organizada por ordem de frequência

<b>Letra</b>	<b>A</b>	<b>E</b>	<b>O</b>	<b>I</b>	<b>R</b>	<b>S</b>	<b>N</b>	<b>D</b>	<b>M</b>	<b>T</b>	<b>C</b>	<b>U</b>	<b>L</b>
<b>Frequência (%)</b>	15,92	11,76	10,32	8,17	7,04	6,63	6,47	5,24	4,78	3,95	3,65	3,54	3,08
<b>Letra</b>	<b>P</b>	<b>F</b>	<b>G</b>	<b>V</b>	<b>B</b>	<b>H</b>	<b>Q</b>	<b>J</b>	<b>Z</b>	<b>Y</b>	<b>K</b>	<b>W</b>	<b>X</b>
<b>Frequência (%)</b>	1,90	1,80	1,23	1,18	0,82	0,82	0,72	0,46	0,46	0,05	0,00	0,00	0,00

Fonte: elaborado pelos autores.

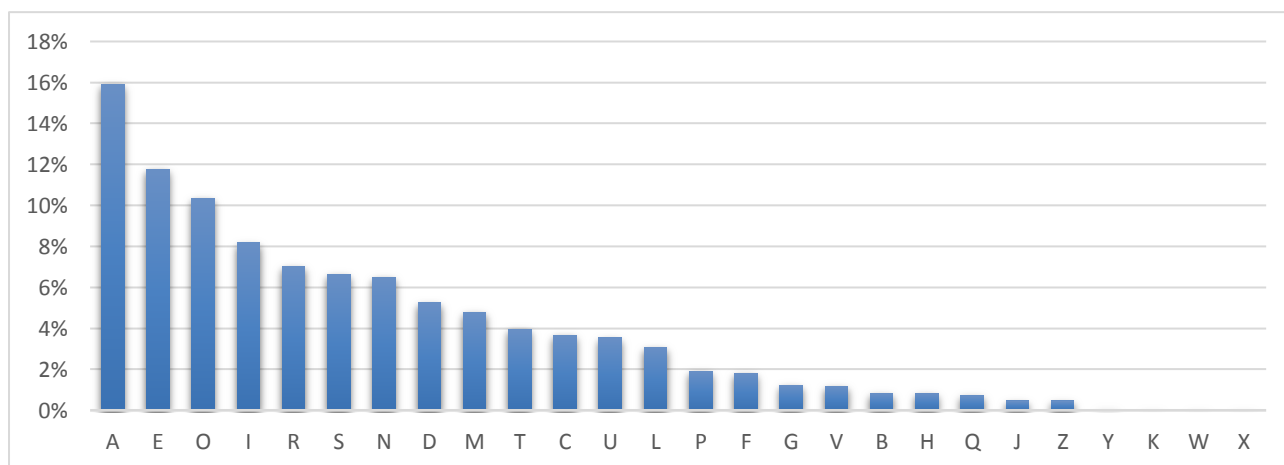


Figura 1 – Distribuição das letras por ordem de frequência  
Fonte: Elaborado pelos autores



Agora, analisando o texto criptografado, no Quadro 1, conseguimos, respectivamente, a Tabela 7 e a Figura 2.

Tabela 7 – Frequência do texto criptografado

<b>Letra</b>	<b>D</b>	<b>T</b>	<b>W</b>	<b>R</b>	<b>E</b>	<b>J</b>	<b>K</b>	<b>O</b>	<b>S</b>	<b>U</b>	<b>N</b>	<b>V</b>	<b>M</b>
<b>Frequência (%)</b>	14,17	13,39	9,45	9,06	7,87	7,48	6,69	4,72	4,72	3,54	3,15	3,15	2,76
<b>Letra</b>	<b>Q</b>	<b>H</b>	<b>P</b>	<b>F</b>	<b>Y</b>	<b>L</b>	<b>X</b>	<b>A</b>	<b>B</b>	<b>C</b>	<b>G</b>	<b>I</b>	<b>Z</b>
<b>Frequência (%)</b>	2,36	1,97	1,57	1,18	1,18	0,79	0,79	0,00	0,00	0,00	0,00	0,00	0,00

Fonte: elaborado pelos autores.

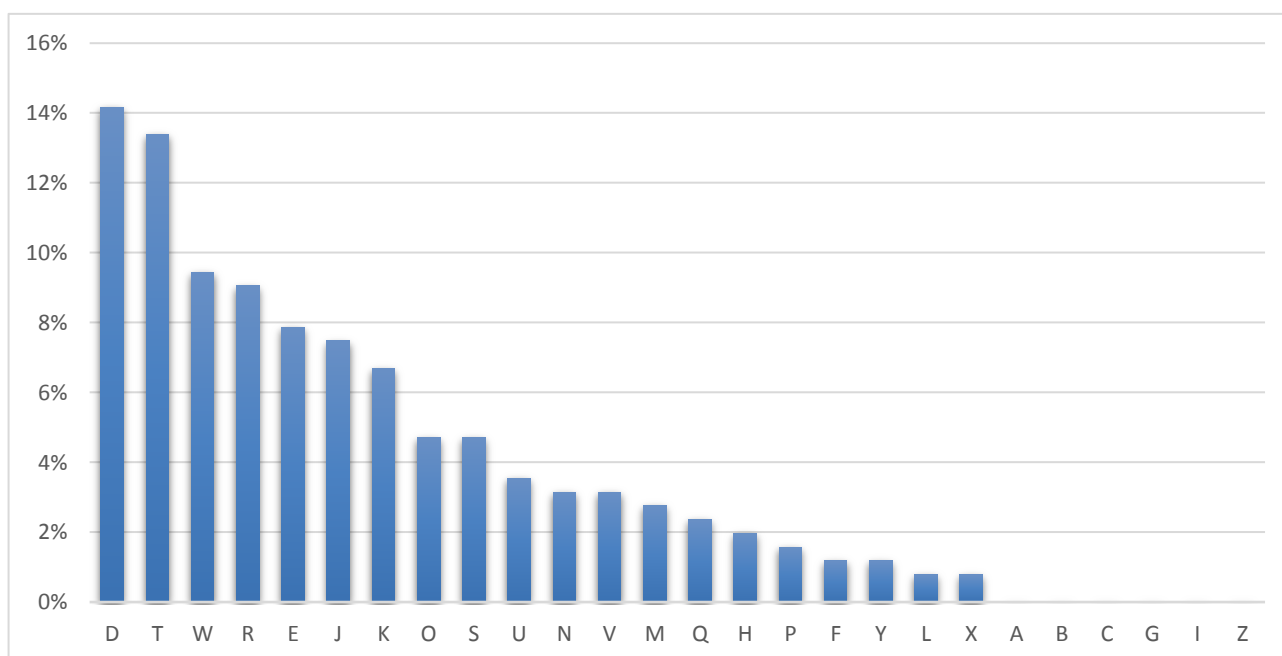


Figura 2 – Distribuição das letras por ordem de frequência

Fonte: Elaborado pelos autores

Em seguida, o objetivo é comparar as tabelas e ir fazendo substituições, ou seja, se a frequência estiver correta, o D deve corresponder ao A, enquanto que o T, W, R e E devem corresponder a E, O, I e R respectivamente. Após a primeira substituição, que deve ser o D pelo A, podemos procurar palavras comuns formadas por apenas duas ou três letras que podem nos ajudar a identificar outras letras. É importante considerar a possibilidade de as tabelas não estarem perfeitamente alinhadas para alguns casos e tentar opções mais próximas. Para isso, precisamos estudar cada substituição. Substituindo o T por E, já fica bastante provável que os dígrafos **Ua** e **Ue** (ao fazer a modificação, deixamos as letras em minúsculo) possivelmente indicam que a letra U está no lugar da letra D. Já a trigramma **FNe** deve indicar que F é igual a Q e N igual a U. Se essas modificações fizerem sentido, continuamos, caso contrário, deveremos verificar os valores mais próximos com base nas tabelas. Também estamos fortemente inclinados a substituir o V por M em virtude da ocorrência de **eV**. Supondo que as substituições estão corretas, as estruturas **quaOKW**, **Wu e W** indicam fortemente que o W deve ser um O, pois além dos dados informados no gráfico notamos a presença de um artigo no final do texto, pois o "a" e o "e" já foram

encontrados. Vejamos como nosso texto criptografado está após as primeiras modificações:

#### Quadro 2 – Texto criptografado

TEKDKREKRS D T D SRTOSRD FNT NKRQRXD ET UDE KTWJRDE  
MJWPDPRQREKRSDE MDJD TLMQRS DJ D HJTFNTOSRD UD WSWJJTOSRD UT  
TYTOKWE KDOKW TV TEKNUWE WPETJYDSRWODRE FNDOKW TV  
TLMTJRVTOKWE MDJD VWUTQDJ D DQTDKWJRTUDUT T D ROSTJKTXD UT  
HWJVD D TEKRVDJ WN MWEERPRQRKDJ D MJTYREDW UT HTOWVTOWE  
HNKNJWE SWOHWJVT W SDEW.

#### Quadro 3 – Texto com as primeiras modificações

eEKaKREKRSa e a SReOSRa que uKRQRXa Ee daE KeoJRaE  
MJoPaPRQREKRSaE MaJa eLMQRSaJ a HJequeOSRa da oSoJJJeOSRa de eYeOKoE  
KaOKo em eEKudoE oPEeJYaSRoOaRE FuaOKo em eLMeJRmeOKoE MaJa  
modeQaJ a aQeaKoJRTdade e a ROSeJKeXa de HoJma a eEKRVaJ ou  
MoEERPRQRKaJ a MJeYREao de HeOomeOoE HuKuJoE SoOHoJme o SaEo.

Logo após, **quaOKo**, **MaJa** e **modeQaJ** parecem substituíveis por **quando**, **para** e **modelar**. Apesar disso, pode ser que ainda haja outras palavras que passem despercebidas e se encaixem no contexto como, por exemplo, **quanto** no lugar de **quando**. A substituição do J por R como visto em **modeQaJ** torna-se mais evidente analisada em **HoJma** que poderia ser **torna** ou **forma** indicando, assim, mais duas possibilidades para a substituição do H, mas já sabemos que há o “m”.

Em virtude do pouco espaço para análises minuciosas, algo que o professor poderá fazer melhor ao considerar um texto estudado previamente, apenas substituiremos o J, O e M por R, N e P respectivamente. Tais alterações nos darão um final de texto (nas atuais circunstâncias) como se segue:

eEKaKREKRSa e a SReNSRa que uKRQRXa Ee daE KeorRaE  
proPaPRQREKRSaE para eLpQRSaR a HrequenSRa da oSorrenSRa de eYenKoE  
KanKo em eEKudoE oPEerYaSRonaRE FuanKo em eLperRmenKoE para modeQar a  
aQeaKorRTdade e a RnSerKeXa de Horma a eEKRVaR ou poEERPRQRKaR a  
preYREao de HenomenoE HuKuroE SonHorme o SaEo.

A substituição do Q por L como previsto anteriormente, reforça-se em **aQeaKorRedade**, que parece ser perfeitamente compatível com **aleatoriedade** que, por sua vez, nos indica que o R é substituível por I e o K por T, fazendo com que a tradução **estimar** seja perfeitamente plausível para **eEKRmar**. Se estivermos corretos, o E tem valor de S devido ao encontro consonantal em **poEERPRQRKaR** e tais substituições nos deixarão muito próximos do resultado final:

estatistiSa e a SienSia que utiliXa se das teorias proPaPilistiSas para eLpliSar a  
HrequenSia da oSorrenSia de eYentos tanto em estudos oPserYaSionais quanto em  
eLperimentos para modelar a aleatoriedade e a inSerteXa de Horma a estimar ou  
possiPilitar a preYisao de Henomenos Huturos SonHorme o Saso.



Finalmente, o texto parece perfeitamente compreensível e se revela conforme a seguinte informação extraída na internet.

*Estatística é a ciência que se utiliza das teorias probabilísticas para explicar a frequência da ocorrência de eventos tanto em estudos observacionais quanto em experimentos para modelar a aleatoriedade e a incerteza de forma a estimar ou possibilitar a previsão de fenômenos futuros conforme o caso.*

## **Considerações finais**

Buscamos com o presente trabalho ressaltar a importância da estatística como ferramenta norteadora para a tomada de decisões nas diversas áreas do conhecimento, aliar o interesse intrínseco pela escrita em códigos com a ludicidade de uma atividade que pode ser proposta como competição de grupos para obtenção do resultado sem perder, contudo, o foco nos aspectos característicos dos levantamentos estatísticos e do método científico como a identificação do problema, levantamento de hipóteses, coleta, observação e análise de dados e verificação de resultados.

Lopes (2008), afirma que é necessário trabalhar situações que envolvam a ideia de acaso e de aleatório para não reduzirmos o ensino da matemática à mera análise de valor verdadeiro ou falso de suas proposições. Pensando nisso, a situação proposta confronta o aluno com a necessidade de ponderar acerca de ocasiões em que a frequência das letras poderá não ocorrer exatamente como a tabela indica e, assim, será necessário desenvolver estratégias para contornar a situação e solucionar o problema. Logo, espera-se que o educando, por meio das orientações do professor, se aproprie dos conceitos estatísticos direcionando suas ações sem, contudo, garantia de exatidão dos resultados. Perceba-se dependente da adoção de estratégias conjuntas, bom senso, revisão e até redirecionamento dos passos adotados para obtenção dos objetivos.

Ressaltamos ainda a grande possibilidade para interdisciplinaridade da atividade proposta, dada à possibilidade das notícias se relacionarem com um tema específico de estudos ou de interesse da turma. Aqui também surge a possibilidade de observar e comentar mais um aspecto importante sobre os dados analisados, pois, embora a ocorrência de algumas letras seja ínfima na maioria dos textos, assuntos específicos de uma área ou matéria podem fazer com que a ocorrência de certas letras seja distorcida e apresente uma frequência muito maior do que a habitual, como a letra X em um texto sobre raios X, por exemplo. Essa é uma boa oportunidade para destacar, se necessário, questões de possível influência como época, regionalismos, jargões profissionais e outros que podem levar a uma necessidade de revisão do estudo de frequência obtido.

Com a construção das tabelas, espera-se que os alunos percebam que somente colocar a quantidade de vezes que a letra aparece no texto não é uma boa alternativa para

comparar com outros textos e então, a frequência relativa se apresenta como forma mais adequada e as tabelas começam a ficar mais parecidas com aquelas trabalhadas nas aulas de estatística tradicionais. Novamente, caberá ao professor estar atento e participativo para permitir que o aluno perceba com mais clareza este tipo de relação, ou seja levado a notá-las quando preciso.

A tarefa também abre oportunidade para explorar ferramentas computacionais. Turmas com conhecimentos em linguagens de programação poderão desenvolver programas para diversas tarefas. Sabendo que, na maioria dos casos, a contagem de letras por meio de ferramentas de busca de editores de texto, por exemplo, torna o tamanho irrelevante, sob certos aspectos. Uma planilha eletrônica já permitiria construir gráficos variados: gráficos da frequência das letras em cada texto; gráficos simultâneos e comparativos de textos; organização por ordem crescente ou decrescente de frequência da ocorrência e outros, permitindo rever certos aspectos sobre a organização dos dados de forma rápida e prática.

A princípio, durante as dificuldades que surgem nas substituições das letras, o estudante pode acreditar, ou ser levado a pensar a respeito por meio das intervenções do professor, que os erros são provocados porque sua tabela não é a mesma que aquela encontrada na internet, porém, mesmo que ele a usasse, as substituições, pura e simplesmente, não resultariam em um texto compreensível. Tal fato evidencia a necessidade de analisar, caso a caso, e verificar possibilidades próximas, pois, mais uma vez, fica claro que por melhor que sejam as aproximações ou medições, a estatística não pode garantir exatidão na maioria das ocasiões, mas sim, apontar os melhores caminhos a serem seguidos, os caminhos mais prováveis.

Por fim, considerando as dificuldades encontradas e as potencialidades observadas no desenvolvimento deste trabalho, ressaltamos a necessidade de se continuar os estudos analisando, cada vez mais, um número maior de textos, com o intuito de tentar perceber uma convergência das frequências das letras para valor ideal que seria aquele das tabelas encontradas por especialistas, e então surge a possibilidade de tratar sobre a interpretação frequentista do conceito de probabilidade e possivelmente ampliar a discussão para outras interpretações, buscando desenvolver novos trabalhos que venham a contribuir para uma aprendizagem mais significativa. Com isso, acreditamos ter fomentado no leitor o interesse pelo desenvolvimento de novos estudos sobre o tema, entendemos que a partir deste trabalho motivamos o desenvolvimento de novas pesquisas no ramo da estatística e probabilidade.

## Referências

ANDRADE, Dalton. F; OGLIARI, Paulo. J. **Estatística para ciências agrárias e biológicas: com noções de experimentação**. 3. ed. Florianópolis: Ed. da UFSC, 2013.

ANSA, Agência. **Humanidade corre risco de suicídio, diz Papa sobre arma nuclear**. Disponível em: <<http://www.jb.com.br/internacional/noticias/2017/10/30/humanidade-corre-risco-de-suicidiodiz-papa-sobre-arma-nuclear/>>. Acesso em: 04/11/2017.

COUTO, S. P. **Códigos e cifras - da antiguidade à era moderna**. Rio de Janeiro, RJ: Novaterra, 2008.

LOPES, Celi E. **O ensino da estatística e da probabilidade na educação básica e a formação dos professores**. Cadernos CEDES (Impresso), v. 28, p. 57-73, 2008.

LOPES, Celi E.; SOUZA, Leandro de O. **Aspectos filosóficos, psicológicos e políticos no estudo da Probabilidade e da Estatística na Educação Básica**. Educação Matemática Pesquisa (Online), v. 18, p. 1465-1489, 2016.

PUPO, Fábio; SIMÃO, Edna; GRANER, Fabio. **Governo prevê mínimo menor em 2018, mas ministro diz que pode mudar**. Disponível em: <<http://www.valor.com.br/brasil/5176942/governo-preve-minimo-menor-em-2018-mas-ministro-diz-que-pode-mudar>>. Acesso em: 04/11/2017.

RODRIGUES, L. **Mãe abandona duas crianças pequenas e vizinhos chamam polícia**. Disponível em: <<http://g1.globo.com/sp/santos-regiao/noticia/2014/03/mae-abandona-duas-criancas-pequenas-e-vizinhos-chamam-policia.html>>. Acesso em: 04/11/2017.

SOUZA, A. C. de. et al. **O ensino de Estatística e Probabilidade na Educação Básica: atividades e projetos gerados a partir de pesquisas de Mestrado Profissional**. Vidya (Santa Maria. Online), v. 33, p. 9-25, 2013.

Submissão: 10/11/2017

Aceite: 18/03/2018