

**FLORIANÓPOLIS, 12 DE JULHO**

PAMELA CRISTINA DOS SANTOS MONTEIRO

Análise estatística dos preços de marcas de bebidas  
Trabalho Final de Probabilidade

Trabalho da Disciplina INE 5405  
– Estatística e Probabilidade,  
sob Orientação do Prof. José  
Francisco D. de G. C. Fletes.

Florianópolis, 12 de Julho de 2024

## Introdução

Este relatório apresenta os resultados e discussões para a atividade 3. A situação a ser analisada foi a de comparar um segmento de mercado com duas marcas concorrentes. O segmento escolhido foi o 1 com o canal 3. As marcas escolhidas foram a COCA COLA e GUARANA ANTARCTICA. Os dados foram analisados com linguagem Python usando biblioteca Pandas e os resultados gerados em Excel.

### Objetivo geral:

Ao analisar esse modelo, observamos que os preços individuais das duas marcas tendem a se desviar da média com uma pequena variação, indicando que os erros ( $\epsilon$ ) são relativamente pequenos. Isso sugere que:

- **Estabilidade dos Preços:** Os preços das bebidas não variam muito em relação à média, indicando uma certa estabilidade no mercado.
- **Previsibilidade:** Para o consumidor, isso significa que os preços são relativamente previsíveis e não apresentam grandes flutuações inesperadas.
- **Análise Comparativa:** Comparando duas marcas, se ambas têm variações pequenas em torno da média, pode-se inferir que a competição de preços entre elas é equilibrada e que o consumidor não encontrará grandes diferenças de preços entre as duas.

Essa análise é útil para o consumidor ao considerar suas opções de compra, pois fornece uma visão de como os preços se comportam em relação a uma média central e quão consistentes são essas variações.

## **PARTE I: ESTATÍSTICA DESCRITIVA E EXPLORATÓRIA (Vale 1,0)**

### **I.1- Os modelos empíricos (distribuições/tabelas de frequências)**

#### **Para COCA COLA:**

Menor  $X = 6,79$

Maior  $X = 12,00$

Range = 5,21

$k = \text{raiz}(n) = 10,91$

$C = R/k = 0,477$

k	DADOS AGRUPADOS			HISTOGRAMA	<del>f</del> freq (n <sub>i</sub> )	<del>p</del> pi	<del>P</del> Pi	<del>X</del> Xi	<del>X</del> Xi*pi	<del>Q</del> QDP
1ª	6,79	/----	7,29	///	3	2,52%	2,52%	<b>7,04</b>	0,1775	0,041
2ª	7,29	/----	7,79	////	31	26,05%	28,57%	<b>7,54</b>	1,9642	0,154
3ª	7,79	/----	8,29	////	17	14,29%	42,86%	8,04	1,1486	0,010
4ª	8,29	/----	8,79	////	4	3,36%	46,22%	8,54	0,2871	0,002
5ª	8,79	/----	9,29	////	28	23,53%	69,75%	9,04	2,1271	0,125
6ª	9,29	/----	9,79	////	5	4,20%	73,95%	9,54	0,4008	0,064
7ª	9,79	/----	10,29	////	24	20,17%	94,12%	10,04	2,0249	0,604
8ª	10,29	/----	10,79		0	0,00%	94,12%	10,54	0,0000	0,000
9ª	10,79	/----	11,29	///	3	2,52%	96,64%	11,04	0,2783	0,188
10ª	11,29	/----	11,79		0	0,00%	96,64%	11,54	0,0000	0,000
11ª	11,79	/----	12,29	////	4	3,36%	100,00%	12,04	0,4047	0,052
TOTAL					119	100,00%		7,57	8,81	1,4
							MODA	Média	POD.	VAR
										<b>1,18</b>
									DESVPAD	

Os preços concentram-se na classe 2ª, ou seja, a maior parte está dentro da faixa de valores de 7,29 a 7,79. Existem poucos preços abaixo de 6,79 reais.

### Para Guaraná Antártico:

Menor  $X = 4,59$

Maior  $X = 10,00$

Range = 5,41

$k = \text{raiz}(n) = 9,85$

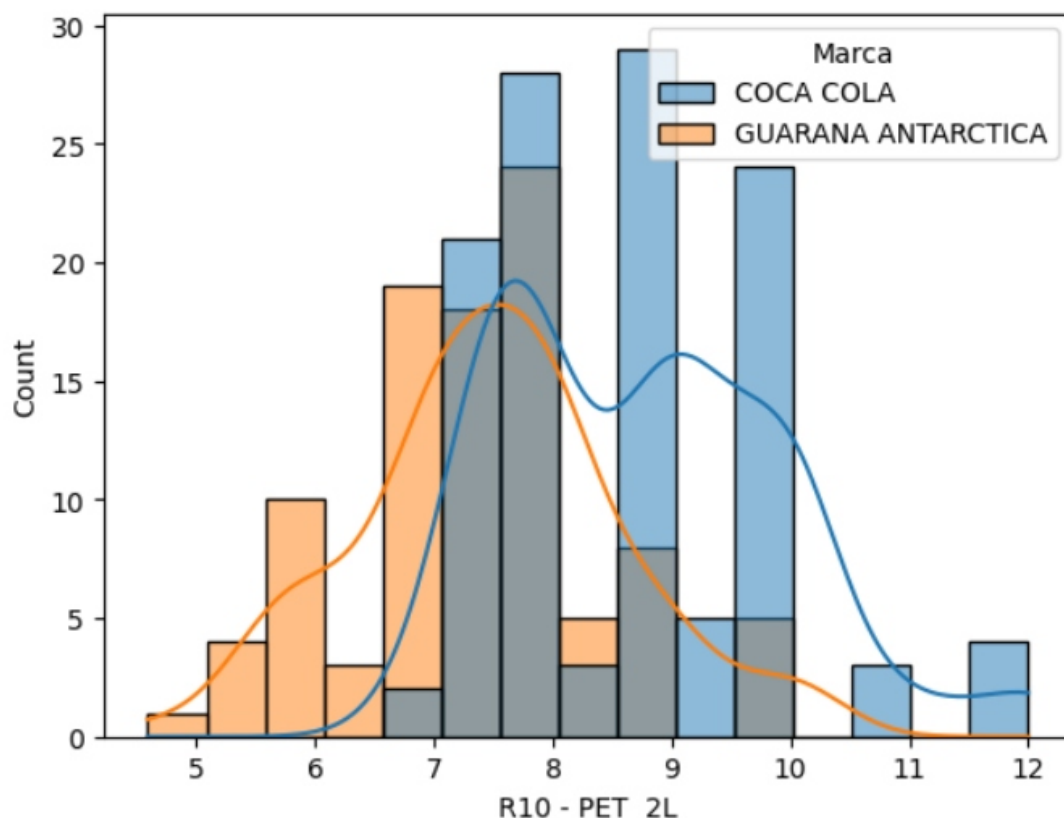
$C = R/k = 0,54$

k	DADOS AGRUPADOS				HISTOGRAMA	Freq (ni)	pi	Pi	Xi	Xi*pi	QDP
1ª	4,59	/----	5,14	/		1	1,03%	1,03%	4,865	0,0502	0,0754
2ª	5,14	/----	5,69	////		4	4,12%	5,15%	5,415	0,2233	0,1913
3ª	5,69	/----	6,24	///// /////		10	10,31%	15,46%	5,965	0,6149	0,2652
4ª	6,24	/----	6,79	///		3	3,09%	18,56%	6,515	0,2015	0,0344
5ª	6,79	/----	7,34	///// ///// ///// ///// ///		23	23,71%	42,27%	7,065	1,6752	0,0602
6ª	7,34	/----	7,89	///// ///// ///// ///		18	18,56%	60,82%	7,615	1,4131	0,0004
7ª	7,89	/----	8,44	///// ///// ///// /////		20	20,62%	81,44%	8,165	1,6835	0,0732
8ª	8,44	/----	8,99	///// //		7	7,22%	88,66%	8,715	0,6289	0,0948
9ª	8,99	/----	9,54	///// /		6	6,19%	94,85%	9,265	0,5731	0,1779
10ª	9,54	/----	10,09	/////		5	5,15%	100,00%	9,815	0,5059	0,2600
Total						97	100,00%		7,65	7,569	1,23
									MODA	Media Pod.	VAR
											1,109
											DESVPAD

Os preços concentram-se na classe 5ª, ou seja, a maior parte está dentro da faixa de valores de 6,79 a 7,74. Existem poucos preços abaixo de 5,00 reais.

## Histogramas das duas Marcas

Análise do Histogramas obtidos quanto ao **padrão da distribuição de preços**, para a situação.



Do ponto de vista do consumidor, o produto mais barato é da marca GUARANA ANTARCTICA. Como o desvio padrão e variância é similar em ambas as marcas, a variação de preços se comporta de forma similar. Portanto, o consumidor que está acostumado com um preço de ambas as marcas, não terá surpresas ao tentar comprar em outro estabelecimento. A Tabela 3 apresenta as cidades com os menores preços para cada marca. A COCA COLA apresenta 4 cidades com menor preço, enquanto o GUARANA ANTARCTICA apresenta somente 1 cidade. Isso na visão do consumidor mostra, que, dependendo da cidade, está em vantagem pois é mais fácil encontrar COCA COLA na faixa de frequência de menor preço.

**Tabela 3 - Lista de cidades que apresentam os menores valores para cada marca**

Código da cidade	Cidades	Marca
174	Asa norte	COCA COLA
940	Planaltina	COCA COLA
2656	Samambaia	GUARANA ANTARCTICA
2935	Santa maria	COCA COLA
28581	Sao Sebastião	COCA COLA

### **1.1 - Estatísticas descritivas**

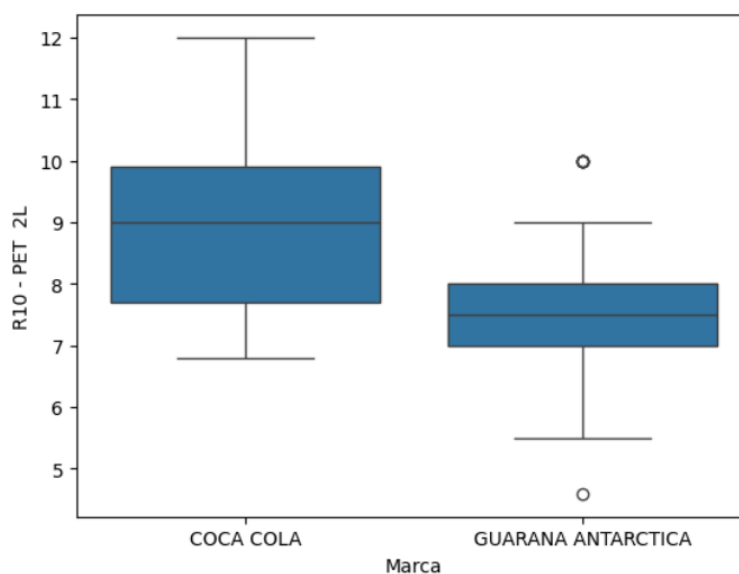
A Tabela 4 apresenta as estatísticas descritivas extraídas a parte do modelo empírico. Para os valores de média simples, mediana simples e média ponderada, os valores da marca COCA COLA ficam sempre acima dos valores da marca GUARANA ANTARCTICA por pelo menos 1R\$, indicando que os preços da COCA COLA são mais caros. No entanto, o valor da moda fica bem próxima em ambas as marcas, indicando que para a marca COCA COLA existem valores que puxam as médias e mediana para cima. Mas olhando o valor do erro relativo, mostra que a margem de erro da COCA COLA fica dentro do permitido (abaixo dos 0,50%). Enquanto isso, o erro relativo do GUARANA ANTARCTICA extrapola, mostrando que pode haver erros no preço da marca. O coeficiente se mostrou bem próximo da dispersão nula. Por fim, ambos mostraram assimetria muito forte pois os valores ficaram acima de 1.

**Tabela 4 - Estatísticas descritivas**

Estatísticas descritivas		
	MARCA - COCA COLA	MARCA - GUARANA ANTARCTICA
Média Simples	8,77	7,48
Mediana Simples	9,00	7,49
Média Ponderada	8,81	7,6
Moda (MOD0)	7,54	7,065
Erro relativo agrupado	0,49%	1,60%
VAR	1,40	1,23
DESVPAD	1,18	1,1
CV	0,13	0,14
Assimetria	1,497	2,46

#### 1.4. Diagrama de caixas

A Figura 2 apresenta o boxplot da coluna R10 - PET 2L de ambas as marcas. É possível notar que a mediana da COCA COLA está acima do superior do GUARANA ANTARCTICA. Isso mostra que, no geral, a COCA COLA apresenta um preço muito elevado se comparado a outra marca. No entanto, o limite superior da COCA COLA é acima do outlier do GUARANA ANTARCTICA, reforçando que a COCA COLA é um produto bem mais caro do que a outra marca escolhida.



## **PARTE II: ESTATÍSTICA NÃO-PARAMÉTRICA (Vale 3,0)**

- **SITUAÇÃO DE INTERESSE:** A situação de interesse é verificar se os dados de preços de bebidas seguem uma distribuição normal;
- **HIPÓTESE NULA:**  $H_0$ : Os dados mostram aderência à frequência esperada da distribuição normal;
- **HIPÓTESE ALTERNATIVA:**  $H_1$ : Os dados não mostram aderência à frequência esperada da distribuição normal;
- **ESTATÍSTICA:** A estatística do teste no caso do qui-quadrado é calculada usando a fórmula:  $\chi^2 = \sum ((O_i - E_i)^2 / E_i)$ ;
- **CRITÉRIO REJEIÇÃO  $H_0$ :** Se o  $\chi^2$  calculado for maior que o valor crítico, a hipótese nula é rejeitada. O valor crítico depende do número de graus de liberdade, que é geralmente  $k-1$ , onde  $k$  é o número de categorias ou intervalos.

### **QQ: (Marca COCA - COLA)**

k	QUI					Ei	QUI
	Oi	z1s	z2s	P(z1<Z<z2)			
1°	8	-1,6734	-1,0887	0,1381		16,4385	4,3318
2°	43	-1,0887	-0,5040	0,1690		20,1089	26,0581
3°	4	-0,5040	0,0807	0,2250		26,7780	19,3755
4°	33	0,0807	0,6654	0,2149		25,5783	2,1535
5°	24	0,6654	1,2501	0,1473		17,5251	2,3922
6°	0	1,2501	1,8347	0,0724		8,6118	8,6118
7°	3	1,8347	2,4194	0,0255		3,0344	0,0004
8°	4	2,4194	3,0041	0,0078		0,9249	10,2243
TOTAL	119			1,0		119	73,15

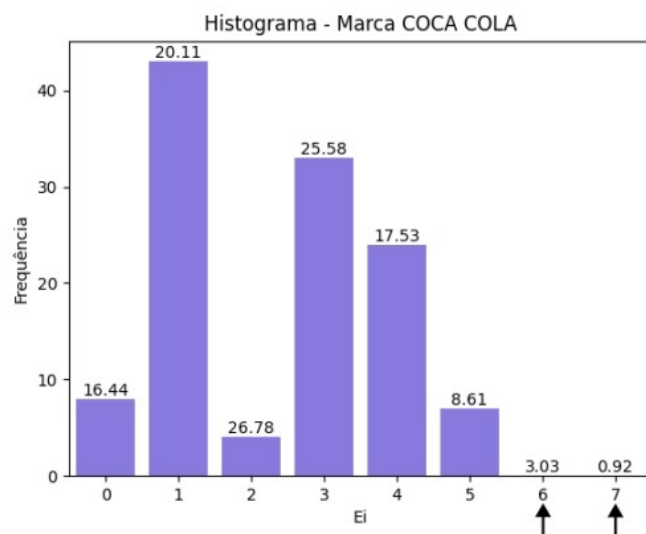
### **Hipótese nula: $H_0: O_i = E_i$**

Ou seja, a forma do modelo é gaussiana

- As frequências esperadas ( $E_i$ ) < 5 para todas as classes.
- Agrupe essas classes com as classes adjacentes até que a soma das frequências esperadas seja maior que 5.



## Histograma da tabela do Qui - Quadratico ( Marca Coca - Cola )



### Cálculos:

Para garantir que as classes no histograma tenham uma frequência esperada ( $E_i$ ) maior que 5, é necessário agrupar classes adjacentes que não atendam a essa condição. No seu caso, as classes 7ª e 8ª precisam ser ajustadas porque não obedecem à regra  $E_i > 5.00$ .

Assim,

Como a Classe 6ª (3.03) e a Classe 7ª (0.92) têm  $E_i < 5$ , podemos agrupar essas tres classes com a Classe 5ª para garantir que o novo grupo tenha uma  $E_i > 5$ .

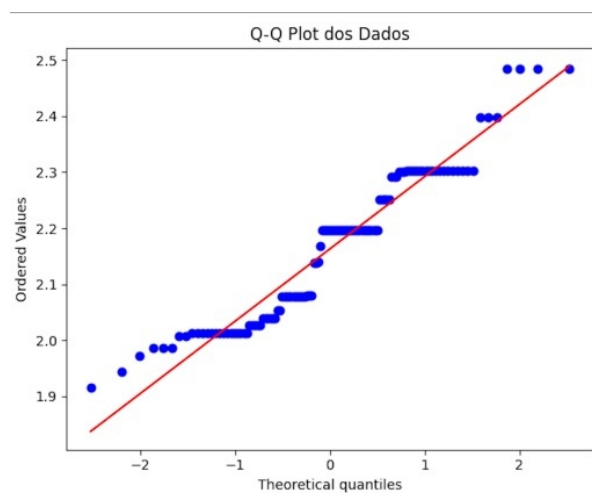
Ou seja  $8.71 + 3.03 + 0.92 = 12.5711$ , se tornando uma nova classe 6ª. Como mostrado na tabela abaixo.

QUI			
MARCA COCA – COLA			
k	O i	Ei Novo	QUI
1º	8	16,4385	4,3318
2º	43	20,1089	26,0581
3º	4	26,7780	19,3755
4º	33	25,5783	2,1535
5º	24	17,5251	2,3922
6º	7	12,5711	2,4690
TOTAL	119	119	56,7800

### Distribuição logNormal:

Distribuição logNormal: MARCA COCA – COLA						
k	O <sub>i</sub> (observado)	z1s	z2s	E <sub>i</sub>	P	Q <sub>U</sub> I
1°	6	-1,87840	-1,30384	11,44112	0,09614	2,5877
2°	30	-1,30384	-0,72928	16,27586	0,13677	11,5725
3°	18	-0,72928	-0,15472	24,46726	0,20561	1,7094
4°	29	-0,15472	0,41985	26,67727	0,22418	0,2022
5°	8	0,41985	0,99441	21,09713	0,17729	8,1307
6°	21	0,99441	1,56897	12,10044	0,10168	6,5454
7°	7	1,56897	2,14354	5,03275	0,04229	0,0005
8°	0	2,14354	2,71810	1,90816	0,01603	2,5877
TOTAL	119,00			119,00	1,00	33,3361
Qui Critico						10,90751285

### GRAFICO DO LOGNORMAL DA MARCA COCA – COLA



### 7. Conclusão/Decisão estatística:

Valor calculado com base na amostra → Qui-quadrado = 56,78

para erro  $\alpha$  de 5% = 10,90, sendo os graus de liberdade;  $k - p - 1 = 3$  já que temos  $k=6$  classes,  $p=2$  parâmetros estimados do modelo com a amostra. Assim, ao nível de 5% não se rejeita  $H_0$ , podendo empregar o modelo de Gauss para a situação.

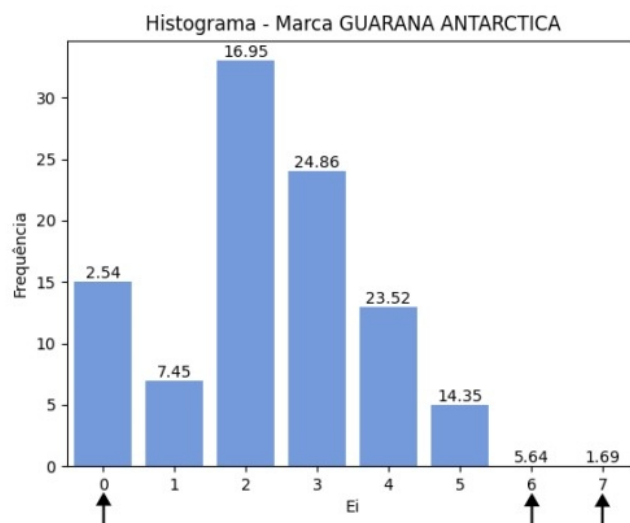
## MARCA (GUARANA ANTARCTICA)

- **SITUAÇÃO DE INTERESSE:** A situação de interesse é verificar se os dados de preços de bebidas seguem uma distribuição normal;
- **HIPÓTESE NULA:**  $H_0$ : Os dados mostram aderência à frequência esperada da distribuição normal;
- **HIPÓTESE ALTERNATIVA:**  $H_1$ : Os dados não mostram aderência à frequência esperada da distribuição normal;
- **ESTATÍSTICA:** A estatística do teste no caso do qui-quadrado é calculada usando a fórmula:  $\chi^2 = \sum ((O_i - E_i)^2 / E_i)$ ;
- **CRITÉRIO REJEIÇÃO  $H_0$ :** Se o  $\chi^2$  calculado for maior que o valor crítico, a hipótese nula é rejeitada. O valor crítico depende do número de graus de liberdade, que é geralmente  $k-1$ , onde  $k$  é o número de categorias ou intervalos.

QUI Marca Guaraná Antarctica						
$k$	$O_i$	$z1s$	$z2s$	$P(z1 < Z < z2)$	$E_i$	QUI
1º	1	-2,6151	-1,9400	0,02619	2,54022	0,9339
2º	14	-1,9400	-1,2649	0,07676	7,44574	5,7695
3º	7	-1,2649	-0,5898	0,17470	16,94637	5,8378
4º	33	-0,5898	0,0853	0,25633	24,86358	2,6626
5º	24	0,0853	0,7604	0,24251	23,52318	0,0097
6º	13	0,7604	1,4355	0,14794	14,35012	0,1270
7º	0	1,4355	2,1106	0,05817	5,64259	5,6426
8º	5	2,1106	2,7857	0,01740	1,68820	6,4969
<b>TOTAL</b>	<b>97</b>			<b>1,0</b>	<b>97</b>	<b>27,48</b>

Hipótese nula:  $H_0: O_i = E_i$  - Ou seja, a forma do modelo é gaussiana

### Histograma da tabela do Qui - Quadrático ( Marca Guaraná Antarctica)



### Cálculos:

Considerando a regra prática que recomenda que  $E_i > 5$  e como a 1ª - 2.54, e a 8ª -1.68 classe não obedecem à regra, agrupamos com as classes adjacentes de forma a obter mais de 5. Assim,

- Como a  $E_i$  novo a Classe 1ª (9.986) e a Classe 6ª (7.331)

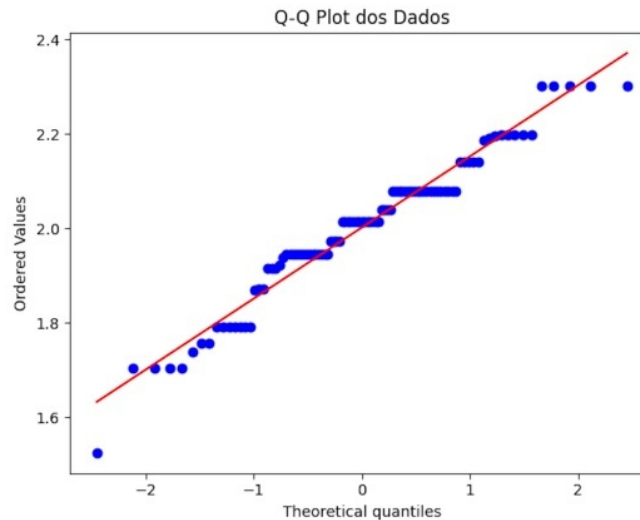
Ou seja  $2.54 + 7.45 = 9.986$ , se tornando uma nova classe 1ª e  $5.64 + 1.69 = 7.331$  se tornando a nova classe 6ª de  $E_i$ . Como mostrado na tabela abaixo.

MARCA (Guaraná Antarctica)			
k	O i	Ei Novo	Qui
1º	15	9,986	2,51760
2º	7	16,946	5,83785
3º	33	24,864	2,66258
4º	24	23,523	0,00967
5º	13	14,350	0,12702
6º	5	7,331	0,74106
TOTAL	97	97	11,8958

### Distribuição logNormal:

Distribuição logNormal: MARCA Guaraná Antarctica						
k	Oi (observado)	z1s	z2s	Ei	P	Qui
1º	15	-3,16366	-2,45127	0,69042	0,00712	0,003392
2º	22	-2,45127	-1,73888	3,28932	0,03391	0,002513
3º	22	-1,73888	-1,02649	10,79639	0,11130	0,916780
4º	25	-1,02649	-0,31410	21,76612	0,22439	0,963679
5º	13	-0,31410	0,39829	26,97273	0,27807	0,000325
6º	0	0,39829	1,11069	20,54989	0,21185	0,000000
7º	0	1,11069	1,82308	9,62298	0,09921	0,0000
8º	0	1,82308	2,53547	3,31216	0,03415	0,0000
TOTAL	97,00			97,00	1,00	1,8867
Qui Critico						10,4461

## GRAFICO DO LOGNORMAL DA MARCA **GUARANA ANTARCTICA**



**CONCLUSÃO:** Os dados observados não se encaixam bem na distribuição normal esperada. Há evidências estatísticas de que as frequências observadas diferem significativamente das frequências esperadas de acordo com a distribuição normal. Também não se encaixa bem na distribuição logNormal.

### **Parte III - Estatística paramétrica (Vale 3,0)**

#### *3.1) Razão de variâncias*

- *SITUAÇÃO DE INTERESSE: A situação de interesse é verificar a razão de variâncias entre as marcas.*
- *HIPÓTESE NULA:  $H_0$ : As variâncias de preços para as duas marcas são iguais.*
- *HIPÓTESE ALTERNATIVA:  $H_1$ : As variâncias de preço das duas marcas são diferentes.*
- *ESTATÍSTICA: A estatística do teste (Teste F)*

- **CRITÉRIO REJEIÇÃO  $H_0$ :** Rejeita a hipótese nula se o valor- $p < \alpha$  ou rejeita a hipótese nula se o valor absoluto da estatística de teste ( $f$ ) for maior que o valor crítico.

Realidade (desconhecida)	Decisão do teste	
	Aceita $H_0$	Rejeita $H_0$
$H_0$ verdadeira	Decisão correta (probab = $1 - \alpha$ )	Erro tipo I (probab = $\alpha$ )
$H_0$ falsa	Erro tipo II (probab = $\beta$ )	Decisão correta (probab = $1 - \beta$ )

## 2ª etapa: Hipótese Nula ou de Igualdade: $H_0$ :

### ● CÁLCULOS:

- Valor Crítico F (Inferior): 0,68439
- Estatística F: 1,14
- Tamanho da amostra da Coca-Cola ( $N_1$ ): 119.00
- Tamanho da amostra da Guaraná Antarctica ( $N_2$ ): 97.00
- F Crítico: 1,47

### Resultado do Teste:

- **Rejeição  $H_0$  (variâncias iguais): Falso**

### • CONCLUSÃO:

Como  **$f = 1.14$**  é menor que  **$f_c = 1.47$** , não rejeitamos a **hipótese nula ( $H_0$ )**. Isso indica que não há evidências suficientes para afirmar que as variâncias dos preços entre Coca-Cola e Guaraná Antarctica são diferentes. Outra abordagem confirma essa conclusão: o valor de  $p$  é maior que o nível de significância ( **$\alpha = 0.05$** ), levando-nos também a aceitar  $H_0$ . Com base nas análises realizadas, concluímos que não há diferenças significativas entre as variâncias dos preços das marcas Coca-Cola e Guaraná Antarctica. Ambas as abordagens do **teste F** indicam que as **variâncias são iguais**, reforçando a hipótese de que os preços das duas marcas têm variabilidade semelhante.

### **3ª etapa e 4ª etapas: Hipótese Alternativa ou Experimental: $H_1$ e Estatística do Teste:**

● *HIPÓTESE ALTERNATIVA:  $H_1$ : As diferenças de médias de preços para as duas marcas são diferentes.*

● **ESTATÍSTICA:** A estatística do teste (teste T):

● **CRITÉRIO REJEIÇÃO  $H_0$ :** Rejeita a hipótese nula se o valor-p  $< \alpha$  ou rejeita a hipótese nula se o valor absoluto da estatística de teste (t) for maior que o valor crítico.

### ***CÁLCULOS:***

#### **Estatística do Teste:**

- Média da Coca-Cola: 8,77
- Média do Guaraná Antarctica: 7,48
- Diferença das médias:  $8,77 - 7,48 = 1,29$
- Estatística t calculada: 8,1818
- Valor crítico  $t_c$  (para  $\alpha = 0,05$ , bicaudal): 1,9711
- Graus de liberdade:  $(119 + 97 - 2) = 214$

### ***5ª etapa: Critério de Rejeição de $H_0$ se: Pela abordagem Clássica***

Para um nível de significância  $\alpha = 0,05$ , a abordagem clássica de rejeição da  $H_0$  é:

- Se  $t > t_c$  (valor crítico), rejeitamos  $H_0$ .

#### ***Resultado do Teste:***

- Rejeição de  $H_0$  (hipótese de médias iguais): **Verdadeira**
- Como o valor de p é menor que o valor de significância (0,05), rejeitamos  $H_0$  e aceitamos  $H_1$ .
- Como  $t > t_c$ , o teste também rejeita.

### 6ª etapa: Cálculos do Excel:

Os cálculos foram realizados utilizando Python e a biblioteca Pandas, com os dados processados e plotados em uma planilha Excel para determinar a estatística t e os valores críticos. Todos os dados foram calculados de forma a garantir precisão nas análises.

#### Estatística paramétrica

Estatística F	Valor crítico F (Limite Inferior)	Valor crítico F (Limite Superior)	Rejeição H0 variâncias iguais
1,14583713319326	0,684387776996406	1,47235602585669	FALSO

Estatística t	Valor crítico t	Rejeição H0 médias iguais
8,18189844201217	1,97111125766066	VERDADEIRO

### 7ª etapa: Conclusões:

Com base nos cálculos, como o valor de p é menor que o nível de significância **(0,05)**, **rejeitamos a hipótese nula (H0)** e aceitamos a hipótese alternativa (H1). Além disso, como a estatística **t calculada (8,18)** é maior que o **valor crítico tc (1,9711)**, também rejeitamos H0. Portanto, concluímos que existem diferenças significativas entre os preços das marcas Coca-Cola e Guaraná Antarctica, indicando que os preços das duas marcas não são equivalentes. Com base na análise estatística paramétrica, podemos concluir que os preços médios das duas marcas são significativamente diferentes, rejeitando assim a hipótese nula de que as médias são iguais.



**III.2 Intervalo de Confiança de 95% para a razão entre *variâncias* e para a diferença de *médias***

**III.1 Para razão de variâncias:**

**Fórmula da razão de variância**

$$\text{IC Razão Variância (95\%)} = \text{Estatística F} / \text{Crítico F}$$

**1. Marca Coca - Cola**

- Estatísticas F = 1,14
- Crítico F = 1,47

$$\text{IC Razão Variância (95\%)} = (\text{Estatísticas F} / \text{Crítico F}) = 1,14 / 1,47 = 0,778$$

**2. Marca Guaraná Antártica**

- Estatística F = 1,14
- Crítico F = 1,47

$$\text{IC Razão Variância (95\%)} = (\text{Estatísticas F} * \text{Crítico F}) = 1,14 * 1,4 = 1,68$$

$$\text{Intervalo de Confiança - IC} = [0,778, 1,68]$$

### III.II - Para diferença de médias:

#### **IC Diferença das Medias (95%)**

Diferença nas médias de duas distribuições normais  $\mu_1 - \mu_2$ , com variâncias  $\sigma_1^2 \neq \sigma_2^2$  e desconhecidas

$$\bar{x}_1 - \bar{x}_2$$

$$\begin{aligned} \bar{x}_1 - \bar{x}_2 - t_{\alpha/2, v} \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}} &\leq \mu_1 - \mu_2 \\ &\leq \bar{x}_1 - \bar{x}_2 + t_{\alpha/2, v} \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}} \\ \text{em que } v &= \frac{(s_1^2/n_1 + s_2^2/n_2)^2}{\frac{(s_1^2/n_1)^2}{n_1 + 1} + \frac{(s_2^2/n_2)^2}{n_2 + 1}} - 2 \end{aligned}$$

- Média COCA - COLA - 8,77
- Média do Guaraná Antártica - 7,48
- Diferença das médias = (Media COCA - COLA - Media do Guaraná Antártica) = 8,77 - 7,48 = 1,29
- N1 (tamanho da amostra da COCA- COLA) - 119.00
- N1 (tamanho da amostra da GUARANA ANTARCTICA) - 97.00
- Valor crítico t - 1,9711

Intervalo de Confiança - **IC = [0,97, 1,59]**

#### **PARTE IV: CONCLUSÕES GERAIS (Vale 2,0)**

Com base nas análises realizadas, as marcas Coca-Cola e Guaraná Antarctica apresentam estatísticas semelhantes, como médias, variâncias e desvios padrão próximos. No entanto, o histograma revela que ambas não seguem uma distribuição normal. Do ponto de vista do consumidor, o Guaraná Antarctica se destaca como a opção mais econômica. A similaridade nas variâncias e desvios padrão indica que a variação de preços é comparável, permitindo que consumidores habituados a uma marca não tenham surpresas ao adquirir a outra.

Embora as médias apresentem diferenças significativas, as variâncias permanecem próximas, com intervalos de confiança que refletem essa relação. A maior parte dos preços do Guaraná Antarctica se concentra na classe 5° (6,79 a 7,74), enquanto os preços da Coca-Cola se agrupam na classe 2° (7,29 a 7,79), com raras ocorrências abaixo de 5,00 reais.

Além disso, nenhuma das marcas se adere à distribuição LogNormal. Em relação às estatísticas descritivas e exploratórias, ambas apresentam valores baixos para a variância e o desvio padrão, e os gráficos de caixas indicam poucos valores fora do padrão.

Os testes de hipótese mostraram que as marcas não aderem à distribuição normal, com valores do QQ bem acima do esperado. Em relação às estatísticas paramétricas, verifica-se que existem diferenças entre as médias, embora as variâncias permaneçam muito próximas. O intervalo de confiança indica os intervalos para a diferença de médias em função das variâncias.