**Análise descritiva das variáveis utilizadas no modelo**

Nesse sentido, o próximo passo consiste em analisar os resultados das principais medidas de análise em estatística descritiva a partir da análise de tendência central, dispersão e amplitude que são dados relevantes para compreender o comportamento das variáveis do modelo. Na tabela **001** retrata as estatísticas descritivas para as variáveis dependente, independente e de controle. As análises foram calculadas considerando as 40 empresas, selecionadas no período 2010 a 2018. Ressalta-se que nessa primeira análise dos dados não foram submetidos a tratamento para os *outliers* e tampouco utilizou-se de transformação nos dados.

**Tabela 001:** Estatística descritiva das variáveis utilizadas na pesquisa

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Variável** | **Média** | **Desvio Padrão** | **Mínimo** | **Máximo** |
| LegNE | 28,754 | 3,892 | 15,750 | 38,53 |
| LegCPC | 45,901 | 1,188 | 43,308 | 49,26 |
| TAM | 5.29e+07 | 1.22e+08 | 804536 | 9.00e+08 |
| COMPLEX | 3,444 | 1,579 | 1,000 | 7 |
| CAPIT | 18,625 | 9,180 | 0 | 31 |
| EXT | 86,328 | 35,544 | 20 | 231 |
| GC\* | - | - | 0 | 1 |
| AUDIT\* | - | - | 0 | 1 |
| RevCPC\* | - | - | 0 | 1 |
| ADR\* | - | - | 0 | 1 |

Notas: As variáveis RevCPC\*, GC\*, AUDIT\* e ADR\* são binárias.

Fonte: Elaborado pelo autor

Ao analisar os dados da tabela **001,** as variáveis LegNE, LegCPC e COMPLEX apresentaram variabilidade discreta o que é indicando que os dados estão bem distribuídos e pode ser indício de que eles seguem uma distribuição normal. No que tange as demais variáveis categóricas e discretas, notou-se que estas apresentavam elevada variabilidade para as variáveis tamanho do ativo (TAM), total de páginas da nota explicativa (EXT) e tempo de negociação na bolsa de valores (CAPIT). Em relação ao tamanho do ativo, verifica-se que o tamanho médio das empresas oscila 80.4536 a 9.00e+08, assim evidenciando que se trata de uma amostra heterogênea quando se trata de tamanho de ativos. Para melhor compreensão dessa variável nas próximas seções ela será analisada após passar por transformação, sendo a mais indicado por alguns estudos da área utilizar a função logarítmica na base 10. Com relação ao total de páginas da nota explicativa, evidenciou-se uma variabilidade considerável no qual temos empresa com cerca de 20 até 231 páginas. Devido à alta variabilidade entre as variáveis estudadas faz-se necessário tratamento por meio de transformações e a análise dos valores extremos(*outliers*).

**Procedimento de Tratativa das variáveis**

Como procedimento para a transformação e tratamento de *outlier* nos dados que serão apresentados a seguir, utilizou-se no *software* @*Stata* as técnicas *ladder* e *glander* para todas as variáveis dependentes, independentes e de controle. O comando *ladder* foi utilizado com a finalidade realizar várias transformações e testar a normalidade das distribuições após as referidas transformações. Já o comando *glander* foi utilizado para gerar os histogramas com as variáveis já transformadas (os histogramas foram anexados na pesquisa ver anexo I).

O quadro 1 apresenta as variáveis pelas quais tiveram que ser tratadas afim de corrigir problemas de dispersão e problemas de normalidade e na seguida utilizou-se a técnica *winsor* para correção de *outliers*, desse modo, deslocando esses dados para mais próximo do percentil definido que normalmente é 2,5% e 97,5% afim de suprir e controlar os dados que se encontram nas extremidades no qual cada uma das variáveis tratadas iniciaram com um teste de p-value de 0,05, e aumentou-se de 0,05 gradativamente até não existir mais dados nos extremos.

**Quadro 1:** Transformações e tratamento de outlier (*winsorizadas*)

|  |  |  |  |
| --- | --- | --- | --- |
| **Nome variável** | **Nome após transformação** | **Nome da transformação** | **Nome variável após tratamento *outlier*** |
| TAM | lTAM | logaritmo | wlTAM -> **TAM’** |
| COMPLEX | sCOMPLEX | quadrado | wsCOMPLEX -> **COMPLEX’** |
| CAPIT | sqCAPIT | raiz quadrada | WsqCAPIT -> **CAPIT’** |
| EXT | sqEXT | raiz quadrada | WsqEXT -> **EXT’** |

Nota: As variáveis winsorizadas foram nomeadas com aspas simples para simplificação.

Fonte: Elaborado pelo autor.

Logo após as tratativas nas variáveis de interesse de estudo foram realizados novamente a verificação da estatística descritiva para analisar se os resultados de tendência central, dispersão e amplitude obtiveram significativas melhoras. O que por ventura pode ser constatado pelos dados apresentados na tabela **002**.

**Tabela 002:** Estatística descritiva das variáveis utilizadas na pesquisa **após transformações**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Variáveis** | **Média** | **Desvio Padrão** | **Mínimo** | **Máximo** |
| wlTAM | 16,83 | 1,08 | 15,03 | 18,97 |
| wsCOMPLEX | 14,18 | 10,60 | 1,00 | 36,00 |
| wsqCAPIT | 4,18 | 1,13 | 2,24 | 5,57 |
| WsqEXT | 9,06 | 1,63 | 5,83 | 11,79 |

Nota: As variáveis *winsorizadas* foram nomeadas com aspas simples para simplificação.

Fonte: Elaborado pelo autor.

No anexo 1 está o *Box Plots* gerados e por meio a analises destes podemos evidencia melhoras em termos de normalização e simetria dos dados após a aplicação das técnicas mencionadas.

Na próxima seção serão apresentados resultados referentes aos pressupostos para análise de correlação em dados em Painel, quando se utilizar as variáveis com transformações elas serão identificas com as aspas.

**4.1 Análise de correlação múltipla**

Nessa seção serão apresentados os testes de especificação das variáveis de estudo, sendo estes: normalidade e assimetria, multicolinearidade, homocedasticidade e autocorrelação serial de dados em painel.

**Teste de Normalidade**

Primeiramente foi verificado se os dados seguiam distribuição normal. Portanto, através do teste Shapiro-Wilk as variáveis dependente e independente foram submetidas à estatística W.

**Tabela 003:** Teste de Normalidade (Shapiro-Wilk) – Antes do tratamento

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Variável** | **W** | **V** | **Z** | **Prob>z** |
| LegNE | 0,990 | 2,240 | 1,899 | 0,029 |
| LegCPC | 0,990 | 2,148 | 1,800 | 0,036 |
| TAM | 0,378 | 140,331 | 11,641 | 0,000 |
| COMPLEX | 0,982 | 4,101 | 3,323 | 0,000 |
| CAPIT | 0,921 | 17,894 | 6,792 | 0,000 |
| EXT | 0,949 | 11,421 | 5,735 | 0,000 |

Nota: As variáveis não foram tratadas nesse teste.

Fonte: Elaborado pelo autor

A tabela **003** apresenta os resultados para o teste de normalidade antes da transformação e tratamento das variáveis de estudo. Podemos notar fortes indícios de que os dados não possuem distribuição normal na variável TAM, CAPIT e EXT, algo já esperado visto a variabilidade apresentada na tabela **001**. Com a finalidade de corrigir tais dispersões foi realizado as tratativas como mencionados em seções anteriores. Desse modo, as variáveis foram restadas já tratadas como apresentada na tabela **004**. Assim sendo, evidenciou que houve uma melhora significativa nos índices.

**Tabela 004:** Teste de Normalidade (Shapiro-Wilk) – Após do tratamento

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Variável** | **W** | **V** | **Z** | **Prob>z** |
| LegNE | 0,990 | 2,240 | 1,899 | 0,029 |
| LegCPC | 0,990 | 2,148 | 1,800 | 0,036 |
| TAM’ | 0,987 | 2,918 | 2,521 | 0,005 |
| COMPLEX’ | 0,971 | 6,450 | 4,389 | 0,000 |
| CAPIT’ | 0,911 | 20,020 | 7,056 | 0,000 |
| EXT’ | 0,989 | 2,377 | 2,039 | 0,020 |

Fonte: Elaborado pelo autor

Segundo os resultados apresentados na tabela **004**, rejeitamos a hipótese nula de dados normalmente distribuídos ao nível de significância de 5%. Os valores de grande de V´ indicam a não normalidade dos dados. Nesse caso, notou-se que não houve valores significativamente grandes nas variáveis testadas, mesmo assim, ainda com a finalidade de comprovar que os dados possuem a distribuição normal, realizou-se os testes de assimetria que curtose que serão apresentados na tabela **005**.

Com a finalidade de avaliar se a amostra coletada é oriunda de uma população da qual apresenta normalidade, Hair Jr. *et al.* (2009) e Kline (2011) recomendam que sejam analisadas as medidas de assimetria e curtose da amostra. Ainda segundo os mesmos autores, a amostra segue uma distribuição normal quando os valores de assimetria são inferiores a 3 e os de curtose são menores do que 8.

**Tabela 005:** Teste de Assimetria e Curtose

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  |  |  |  |  |
| **Variáveis** | **Assimetria** | **Curtose** | **chi2(2)** | **Prob>chi2** |
| LegNE | 0,148 | 0,120 | 4,520 | 0,105 |
| LegCPC | 0,035 | 0,790 | 4,500 | 0,106 |
| TAM’ | 0,339 | 0,000 | 13,930 | 0,001 |
| COMPLEX’ | 0,000 | 0,002 | 23,880 | 0,000 |
| CAPIT’ | 0.0090 | 0,000 | 1012.38 | 0,000 |
| EXT’ | 0,103 | 0,000 | 22,630 | 0,000 |

Fonte: Elaborado pelo autor

Os resultados das análises dos cálculos dos construtos usados nesta pesquisa indicam que seus valores de assimetria e curtose estão dentro dos limites apontados pelos autores, conforme se pode verificar Tabela **005**. No anexo **I**, é possível notar esses achados de forma visual através dos histogramas.

**4.1 Análise de correlação múltipla**

**4.1.2 Teste de verificação de Multicolinearidade**

Empregou-se a estatística VIF (*Variance Inflation Factor*) para avaliar possíveis problemas relacionados à multicolinearidade, para que nenhuma das variáveis independentes ou exploratórias expliquem a mesma função, ou seja, possuírem correlação alta entre si (GUJARATI; PORTER, 2011).

**Tabela 10:** Teste VIF (*Variance Inflation Factor*)

|  |  |  |
| --- | --- | --- |
| **Variável** | **VIF** | **1/VIF** |
| **TAM’** | 3,290 | 0,3039 |
| **CAPIT’** | 2,580 | 0,387 |
| ADR | 2,380 | 0,420 |
| **EXT’** | 1,690 | 0,591 |
| GC | 1,480 | 0,675 |
| **COMPLEX’** | 1,490 | 0,671 |
| AUDIT | 1,370 | 0,729 |
| LegCPC | 1,340 | 0,746 |
| RevCPC | 1,040 | 0,961 |
| VIF médio | 1,851 |  |

Nota: As variáveis winsorizadas apresentam aspas simples.

Fonte: Elaborado pelo autor.

O teste VIF considera para ausência de multicolinearidade valores da estatística do teste inferiores a 10 (dez). Nesse sentido, observando os resultados da Tabela 10, nenhuma variável apresentou multicolinearidade. Como observado não há problemas relacionados a multicolinearidade entre as variáveis, sendo assim, não houve indicativo para remoção de nenhuma das variáveis do modelo.

**4.1.3 Teste de autocorrelação**

Nessa etapa será utilizado o teste de Wooldridge, com a finalidade de verificar se existe autocorreção ou correlação serial quando os erros ou perturbações da regressão são correlacionados ao longo do tempo, assim violando a hipótese de que os erros são aleatórios ou não correlacionados. Sendo assim, as hipóteses deste teste são apresentas a seguir. Como hipótese nula (H0): ausência de autocorrelação no modelo e, para hipótese alternativa (H1): presença de autocorrelação no modelo.

**Tabela XV:** Teste Wooldridge para autocorrelação

|  |  |  |
| --- | --- | --- |
| **Equação** | **F** | **Prob > F** |
| **01** | F( 1, 39) = 11.112 | 0.0019 |

Fonte: Elaborado pelo autor

Com base nos resultados apresentados na tabela XV, podemos notar que o teste de Wooldrige apresentou Prob > F 0.0000, desse modo, rejeitamos a hipótese nula de ausência de autocorrelação a um nível de significância de 5%, e assim sendo, será necessário utilizar o parâmetro *robust* para rodar os modelos de regressão. Esse parâmetro refere-se ao método denominado como *robust* que tem por fundamento rodar várias vezes a mesma regressão para diferentes amostras aleatoriamente extraídas da base original.

**4.1.4 Teste de heterocedasticidade**

**Tabela XVI:** Teste Wald para heterocedasticidade

|  |  |  |
| --- | --- | --- |
| **Equação** | **chi2** | **Prob > chi2** |
| **01** | chi2(40) = 2747.02 | 0.0000 |

Fonte: Elaborado pelo autor

Nessa seção, testou-se a hipótese nula de homocedasticidade dos resíduos. Para isso serão apresentas as hipóteses a seguir: hipótese nula (H0): ausência de heterocedasticidade no modelo e, para hipótese alternativa (H1): presença de heterocedasticidade no modelo.

Com relação ao teste de Wald retornou um Prob> chic2 =0.0000, tal resultado permite rejeitar a hipótese nula, a um nível de significância de 5%, portanto foi observada a presença de heterocedasticidade para a equação da pesquisa o que colabora com o teste anterior que recomenda a utilização do método *robust* para rodar o modelo com a finalidade de amenizar os resíduos homocedásticos, isto é, distribuição de resíduos como uma variância constante.

**4.1.4 Análise da Matriz de Correlação**

A análise de correlação múltipla visa verificar as associações das variáveis estudadas por meio da análise do grau de associação que irá auxiliar na compreensão da técnica de análise. Sendo assim, quanto mais próximo de zero, menor será a associação entre as variáveis e caso seja positiva, as variáveis variam na mesma direção. Quanto mais negativa, indicam variação na direção oposta. Em um estudo no qual podemos considerar a hipótese nula a qual considera que não há associação entre as variáveis, tenderá a ser rejeitada quanto maior for o índice de correlação. As variáveis winsorizadas apresentam aspas simples para diferenciá-la da variável que não passou pelo referido tratamento com demostrando na estas anteriores.

Nota-se na tabela **XVII** uma correlação positiva e significativa (5%) entre o índice de legibilidade na norma (LegCPC) e o índice de legibilidade da nota explicativa (LegNE). Esse resultado indica que aumentos nos níveis de legibilidade da norma da CPC provocariam aumentos significativos no índice de legibilidade das notas explicativas. Observa-se uma correlação negativa e significante (5%) da variável total de páginas da nota explicativa (**EXT’**) com a variável LegNE. Esse resultado pressupõe que acréscimos no número no total de páginas das notas explicativas resultariam em diminuição do índice de legibilidade na nota explicativa.

O índice de legibilidade da norma da CPC (LegCPC) está associado negativamente ao nível de significância de 5% com a variável complexidade operacional (**COMPLEX’**) e com a variável total de páginas da Nota explicativa (**EXT’**). Além disso, essa variável apresenta associação significativa e positiva com a variável **CAPIT’**. Com relação a variável tamanho da empresa (**TAM’**) verificou-se que esta apresentou moderada a associação com as variáveis **COMPLEX’**, **CAPIT’** e **EXT**’. Já a variável **COMPLEX’** está associada com **CAPIT’** e **EXT**’. E por fim, notou-se associação significativa e positiva também para a variável **CAPIT’** sobre a variável **EXT**’.

Podemos perceber que a variável total de páginas da nota explicativa (**EXT’**) possui correlação significante ao nível de 5% com todas as variáveis estudadas o que pode ser indício para um problema de colinearidade no qual uma variável do modelo está associada fortemente as demais variáveis independentes.

**Tabela XVII:** Matriz de Correlação das variáveis de pesquisa

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | **LegNE** | **LegCPC** | **TAM’** | **CAPIT’** | **COMPLEX’** | **EXT’** |
| **LegNE** | 1.0000 |  |  |  |  |  |
| **LegCPC** | 0.1898\* | 1.0000 |  |  |  |  |
| **TAM’** | -0.0445 | -0.0349 | 1.0000 |  |  |  |
| **CAPIT’** | 0.0637 | 0.1552\* | 0.6439\* | 1.0000 |  |  |
| **COMPLEX’** | -0.0738 | -0.1790\* | 0.3895\* | 0.1863\* | 1.0000 |  |
| **EXT’** | -0.1962\* | -0.3127\* | 0.4142\* | 0.1120\* | 0.3874\* | 1.0000 |

Correlação significativa a um nível de 5%

Fonte: dados da pesquisa.

De modo geral os resultados adquiridos até esta etapa foram de suma importância para evidenciar os primeiros sinais de que as variáveis LegCPC e de controle influenciam de algum modo no índice de legibilidade da nota explicativa. Sendo assim, essas análises permitiram uma compreensão inicial para poster estudos mais robustos com a aplicação de técnicas mais sofisticas como dados em painel.

**4.2 Modelo Econométrico**

**4.2.1 Escolha entre modelos de regressão pool, efeito fixo ou efeito aleatório**

Para dados em painel existem modelos diferentes que podem ser utilizados, sendo assim, para a escolha do modelo em painel, nesse estudo, decidiremos entre modelo por efeito fixo, aleatório ou *Pooled*. Com a finalidade de selecionar o modelo adequado foi aplicado os teses de Breush-pagan, Chow e Hausman. Para o primeiro teste rejeitou-se a menos de 1% a hipótese nula (H0). Portanto, o modelo estimado por efeitos aleatórios mostrou-se mais adequado que que o modelo *pooled*.

Logo em seguida foi aplicado o teste de Chow no qual foi possível rejeitar a menos de 1% a hipótese nula para a equação analisada. Sendo assim, o modelo estimado por modelos fixos é considerado mais aquedado do que o modelo *polled*. Portanto, após a realização dos testes mencionados anteriormente descarta-se o modelo *polled*.

Desse modo, foi feito o teste de Hausman para a escolher entre o modelo de efeito fixo ou aleatório. Sendo assim, não se rejeita a hipótese nula e o modelo de efeito fixo é considerado o mais adequado para a criação do modelo.

Adotadas as verificações e correções descritas anteriormente, no que tange as variáveis escolhidas no modelo podemos destacar que devido a problemas de colinearidade as variáveis (**GC** e **AUDIT** ) foram retiradas do modelo que será apresentado a seguir. Para estimar o modelo regressão proposto nessa tese foi utilizando o modelo aleatório (*robust*), para análise da hipótese principal de que a legibilidade da norma contábil afeta a legibilidade das notas explicativas emitidas pelas companhias abertas brasileiras.

As conduções dos testes foram realizadas por meio do software Stata 12 e as estimações apresentadas foram realizadas utilizando comando *robust* com a finalidade de corrigir qualquer tipo de heterocedasticidade.

**4.2.2 Análise da Regressão por efeitos aleatórios**

A tabela **XIV** apresenta os resultados da estimação da regressão para o modelo de dados em painel estimado em que parte da variabilidade, estatisticamente significativa, do índice de legibilidade das notas explicativas pode ser justificada pelas variáveis independentes utilizadas na regressão. Três variáveis se mostraram importantes para o modelo, já que os valores críticos da estatística T foram estatisticamente significativos ao nível de 5% e 10%, sendo elas: Legibilidade da norma contábil (**LegCPC**), Tempo de Negociação na Bolsa de Valores (**CAPIT**) e a *dummy* para Negociação na bolsa de valores de Nova York (**ADR**).

**Tabela XIV:** Modelo de Regressão Aleatório - *Robust*

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Variáveis** | **Beta** | **Erro Padrão Robusto** | **Estat. t** | **p-valor** | **Intervalo de Confiança 95%** | |
| LegCPCMedio | 0,524 | 0,247 | 2,12 | **0,034\*** | 0,039 | 1,010 |
| wlTAM | 0,008 | 0,677 | 0,01 | 0,989 | -1,319 | 1,336 |
| wsqCAPIT | 0,962 | 0,474 | 2,03 | **0,043\*** | -0,031 | 1,893 |
| wsCOMPLEX | 0,226 | 0,404 | 0,56 | 0,575 | -0,056 | 0,101 |
| RevCPC | -0,320 | 0,269 | -1,19 | 0,253 | -0,8497 | 0,208 |
| Reg\_Nreg | -0,776 | 1,221 | -0,64 | 0,525 | -3,170 | 1,617 |
| ADR | -2,181 | 1,294 | -1,69 | **0,092\*\*** | -4,718 | 0,355 |
| WsqEXT | -0,914 | 0,165 | -1,17 | 0,240 | -0,518 | 0,129 |
| \_cons | 3,138 | 11,981 | 0,26 | 0,793 | -20,345 | 26,622 |

Notas: \*\*\*Significante ao nível de \*p<0.05 , \*\* p<0.10, ausência de asterisco representa coeficientes não significativos. Todas as estimativas foram calcular por efeitos aleatórios tendo em vista o teste Hausman.

R² within= 0,0982, R² between= 0,0984, R² overall= 0,0954, F= 17,93.

Fonte: Elaborado pelo autor.

A principal variável explicativa *LegCPC* que mensurou a facilidade de leitura da norma contábil registrou relação positiva significativo de pelo menos 5% para as empresas estudadas, com o índice de legibilidade das notas explicativas, confirmando a hipótese central da tese e indicando que para cada unidade aumentada no índice Flesch de um pronunciamento contábil, em média, consegue-se um ganho de 0,524 na legibilidade da nota explicativa respectiva. Esse resultado é de suma importância para o Comitê de Pronunciamentos Contábeis, sendo esse o órgão responsável pela convergência das normas internacionais, ao passo que tornar os pronunciamentos técnicos mais legíveis resulta em maior compreensão da norma pelo elaborador das demonstrações financeiras, e por fim, implica em notas explicativas menos complexas. Dessa forma, apoiando-se nas dimensões técnica, semântica e pragmática da Teoria da Comunicação de Shannon (1948), o objetivo principal da contabilidade que é de gerar informações úteis aos usuários, tende a ser atingido com mais eficiência. Ademais, sistemas de evidenciação que possuem alto padrão de qualidade fornecem aos investidores confiabilidade e credibilidade nas demonstrações financeiras (LEVITT, 1998).

Com relação a variável que representa em anos o tempo de negociação na bolsa CAPIT (*p-value* 0.043) tem-se efeito positivo e significativo sobre o índice de legibilidade da nota explicativa. Tal resultado retrata que aumentos no tempo de negociação em bolsa de valores provocariam acrescimentos no índice de legibilidade das notas explicativas. Para a variável *dummy* que representa se a empresa possui ações na bolsa de valores de nova York (ADR) apresentou-se efeito negativo e significativo sobre o índice de legibilidade da nota explicativa.

Para a variável tamanho do ativo TAM tem-se efeitos negativos, porém não significativos, desse modo, não se pode presumir que acréscimos nessas variáveis ocasionariam redução no índice de legibilidade das notas explicativas (LegNE). Esperava-se uma relação positiva entre o tamanho da empresa e o índice de legibilidade da nota explicativa (LegNE). Um resultado que contrapõem outros trabalhos (citar)...

Para a variável *dummy* que representa se a norma foi revisada naquela ano pela CPC ( RevCPC) tem-se efeito negativos e não significativo, assim, não é possível inferir que o fato de ser revisa ou não revida possa ocasionar redução no índice de legibilidade das notas explicativas (LegNE) para as empresas estudadas. Por fim, para a variável complexidade operacional (COMPLEX) notou-se um efeito positivo e não significativo sobre o índice de legibilidade das notas explicativas (LegNE), sendo assim, não se pode afirmar que aumentos na complexidade operacional provocariam aumentos no índice de legibilidade das notas explicativas.

A quadro VV representa os sinais para as variáveis de controle encontrados na literatura (esperado) e os sinais informados pelo modelo econométrico aqui testado (observado).

....

**Quadro:** Relação Esperada e Observada para as variáveis

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Variável** | |  | | **Estudo base** | | **Relação**  **Esperada** | | **Relação Observada** | |
| *LegCPC* |  | | Contribuição da tese (H2) | | Positiva | | Positiva | |
| *Comple*x |  | | Miller (2010); Lundholm, Rogo, Zhang, (2014) | | Negativa | | Positiva n.s. | |
| *GC* |  | | - | | Positiva | | Negativa n.s. | |
| *Big4* |  | | Lundholm, Rogo, Zhang, (2014) | | Positiva n.s. | | Negativa | |
| *Temp* |  | | Li (2008) | | Positiva | | Negativa n.s. | |
| *Tam* |  | | Jones (1988); Li (2008) | | Negativa | | Negativa | |
| *Ext* |  | | Miller (2010); Sheidman (2013) | | Negativa | | Negativa n.s. | |

Nota: n.s. = não significativa

Fonte: Elaborado pelo autor (2018)

Diferença de Média

No presente trabalho, aplicou-se o teste t Student testar se duas amostras independentes foram retiradas de populações com médias iguais. Esse teste consiste em um tipo de método paramétrico o qual verifica suposições acerca do parâmetro populacional chamado média, quando há normalidade dos dados analisados (FÁVERO et al., 2009). Sendo assim, este teste tem como pressuposições que o modelo matemático normalidade, homogeneidade e independência dos resíduos foram são satisfeitas. Desse modo, como já foi demonstrado em seções anteriores, os dados da amostra desta pesquisa que serão testados seguem uma distribuição normal, portanto, o teste t de *Student* pôde ser aplicado. Foi um nível de significância de 10%, se o resultado do p-value analisado fosse maior que 0,10 não se rejeita a hipótese nula (Ho) e caso fosse menor que 0,05 rejeita-se a hipótese nula (Ho).

Foram testadas as médias com relação a dois grupos definidos por variáveis binárias em relação ao índice de legibilidade das notas explicativas (LegNE). O primeiro foi a relação de governança corporativa (GC), ou seja, se a empresa possuía ou não para aquele ano especifico. A segunda variáveis o Setor Regulamentado e Não Regulamentado (RegNreg) e, por fim, para a variável que relaciona se a empresa faz negociação na bolsa de valores de Nova York (ADR).

Considerando o Teste t Student, as hipóteses da pesquisa são expressas da seguinte forma:

H1.0: A médias dos escores do índice de legibilidade das notas explicativas é igual para empresas que possuem Governança Corporativa;

H2.0: A médias dos escores do índice de legibilidade das notas explicativas é igual para empresas que são do setor Regulamentado e não Regulamentado;

H3.0: A médias dos escores do índice de legibilidade das notas explicativas é igual para empresas que negociam na bolsa de valores de nova;

Conforme as hipóteses traçadas, aplicaram-se os testes estatísticos paramétricos ANOVA para analisar as diferenças de médias nos grupos selecionados. É importante salientar que devido a estar trabalhando com dados das mesmas empresas para anos diferentes, faz-se necessário calcular a diferença de média para cada período de ano especifico.

A tabela 10 registra a comparação entre o Governança Corporativa e o índice de legibilidade das notas explicativas. Verificou-se que o índice de legibilidade das notas explicativas não é estatisticamente diferente entre os grupos para nenhum dos anos analisados (2010 a 2017), quando observada se a empresa adota ou não Governança Corporativa (ex. p-valor < 0,242 ano 2010), desse modo, não rejeitando-se a hipótese de igualdade H1.0.

**Tabela 10:** Índice de legibilidade da nota explicativa X Governança Corporativa

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Ano** | **x e σ GC = Não** | **x e σ GC =Sim** | **x da diferença** | **t(df)** | **p-value** |
| 2010 | 27,08 (0,83) | 28,36 (0,69) | -1,270 | -1,1892 (38) | 0,242 |
| 2011 | 28,08 (0,85) | 28,97 (0,74) | -0,891 | -0,7935 (38) | 0,432 |
| 2012 | 28,33 (0,99) | 28,82(0,80) | -0,493 | -0,3893 (38) | 0,699 |
| 2013 | 28,67 (1.06) | 28,65 (0.87) | 0,017 | 0,0128 (38) | 0,990 |
| 2014 | 29,16 (1,07) | 28,35 (0,84) | 0,810 | 0,6005(38) | 0,552 |
| 2015 | 28,89 (0,98) | 28,37 (0,87) | 0,515 | 0,3909(38) | 0,698 |
| 2016 | 28,36 (0,68) | 30,20 (1,02) | -1,832 | -1,4587(38) | 0,153 |
| 2017 | 29,81 (0,92) | 29,80(0,72) | 0,010 | 0,0088(38) | 0,993 |

Notas: Significante ao nível 5%. x representa a média e σ representa o desvio padrão.

Fonte: Elaborado pelo autor.

A tabela 11 registra a comparação entre o Setor Regulamentado e Não Regulamentado e o índice de legibilidade das notas explicativas. Notou-se que o índice de legibilidade das notas explicativas não é estatisticamente diferente entre os grupos testados (exceto para os anos de 2014 e 2015) quando observada se a empresa que estão em setor regulamentado e não Regulamentado, desse modo, não rejeitando-se a hipótese de igualdade H2.0.

**Tabela 11:** Índice de legibilidade da nota explicativa X Setor Regulamentado e Não Regulamentado

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Ano** | **x e σ RegNreg = Não** | **x e σ RegNreg =Sim** | **x da diferença** | **t(df)** | **p-value** |
| 2010 | 28,10 (0,63) | 26,93 (1,01) | 1,170 | 0,9940 (38) | 0,327 |
| 2011 | 28,90 (0,59) | 27,73 (1,26) | 1,165 | 0,9551 (38) | 0,346 |
| 2012 | 29,25 (0,61) | 27,04 (1,46) | 2,211 | 1,6545 (38) | 0,106 |
| 2013 | 29,29 (0,74) | 27,21 (1,35) | 2,080 | 1,4390 (38) | 0,158 |
| 2014 | 29,55 (0,70) | 26,76 (1,37) | 2,795 | 1,9946 (38) | 0,053 |
| 2015 | 29,43 (0,74) | 26,68 (1,15) | 2,746 | 2,0151 (38) | 0,051 |
| 2016 | 29,95 (0,70) | 27,88 (1,28) | 2,067 | 1,5137 (38) | 0,138 |
| 2017 | 30,35 (0,62) | 28,57 (1,14) | 1,761 | 1,4526 (38) | 0,155 |

Notas: Significante ao nível 10%. x representa a média e σ representa o desvio padrão.

Foi considerado um intervalo de confiança de 90%

Fonte: Elaborado pelo autor.

Podemos destacar que para os anos de 2014 e 2015 o p-valorfoi alcançado (p-valor < 0,053 e p-valor <0,051 para 2014 e 2015 respectivamente), desse modo, rejeita-se a hipótese de igualdade H2.0, o que porventura é indício para concluir que houve diferença entre as médias quando observada as empresas que estão em setor regulamentado e não Regulamentado, ou seja, estar no setor regulamentado nesse anos sugere que a empresa têm um maior índice de legibilidade da nota explicativa.

**Tabela 12:** Índice de legibilidade da nota explicativa X Negociação na Bolsa de Valores

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Ano** | **x e σ ADR = Não** | **x e σ ADR =Sim** | **x da diferença** | **t(df)** | **p-value** |
| 2010 | 28,11 (0,68) | 27,15 (0,87) | 0,954 | 0,8536 (38) | 0,399 |
| 2011 | 28,67 (0,69) | 28,36 (0,96) | 0,308 | 0,2635 (38) | 0,794 |
| 2012 | 29,18 (0,77) | 27,59 (1,03) | 1,592 | 1,2395 (38) | 0,223 |
| 2013 | 29,07 (0,89) | 27,98 (0,99) | 1,093 | 0,7841 (38) | 0,438 |
| 2014 | 29,55 (0,70) | 26,76 (1,37) | 2,795 | 1,9946 (38) | 0,053 |
| 2015 | 29,38 (0,88) | 27,32 (0,84) | 2,059 | 1,5657 (38) | 0,126 |
| 2016 | 30,00 (0,87) | 28,20 (0,82) | 1,801 | 1,3866 (38) | 0,174 |
| 2017 | 30,09 (0,78) | 29,32 (0,75) | 0,769 | 0,6555 (38) | 0,516 |

Notas: Significante ao nível 5%. x representa a média e σ representa o desvio padrão.

Fonte: Elaborado pelo autor.

Por fim, na tabela 12 registra a comparação entre a negociação na bolsa de valores e o índice de legibilidade das notas explicativas. Para esse teste, como os anteriores notou-se que o índice de legibilidade das notas explicativas não é estatisticamente diferente entre os grupos testados para nenhum dos anos analisados (2010 a 2017), quando observada se a empresa que possuíam negociação na bolsa de valores, desse modo, não rejeitando-se a hipótese de igualdade H3.0. Salienta-se que para o ano de 2014 p-valor, assim sendo, rejeita-se a hipótese de igualdade H3.0 e concluir-se que houve diferença entre as médias quando observada as empresas das quais negociariam na bolsa de Nova York no q que refere-se ao índice de legibilidade das notas explicativas.

**REFERÊNCIAS**

FÁVERO, Luiz Paulo et al. **Análise de Dados:**Modelagem multivariada para tomada de decisões. Rio de Janeiro: Elsevier, 2009.

HAIR JR.; Joseph F.; BLACK, William C.; BABIN, Barry J.; ANDERSON, Rolph E.; TATHAM, Ronald L. **Análise Multivariada de Dados**. 6. ed. Porto Alegre: Bookman, 2009.

KLINE, Rex B. **Principles and Practice of Structural Equation Modeling**. 3. ed. New York: The Guilford Press, 2011.

**ANEXOS**

**ANEXOS I**: Procedimento de tratativas nas variáveis do modelo.







