



ANÁLISE DO PERFIL E DESEMPENHO DOS CANDIDATOS AO VESTIBULAR DA UNESPAR/FECILCAM COMO ESTRATÉGIA PARA O ENSINO DA ESTATÍSTICA NOS CURSOS DE GRADUAÇÃO

Gislaine Aparecida Peričaro, UNESPAR/FECILCAM, gapericaro@fecilcam.br
Solange Regina dos Santos, UNESPAR/FECILCAM, solaregina@fecilcam.br
Tatiane Cazarin da Silva, UTFPR – Campo Mourão, tatianecazarin@utfpr.edu.br

Resumo: O presente trabalho tem o objetivo de apresentar uma forma de aplicação de técnicas estatísticas para análise das informações relacionadas ao perfil sócio-educacional dos candidatos ao vestibular da UNESPAR/FECILCAM, bem como o desempenho destes nas provas que fazem parte do processo seletivo para ingresso nessa Instituição. Os dados são tabulados e analisados por meio da construção de gráficos, aplicação do teste de independência entre variáveis e análise de variância.

Palavras-chave: Estatística. Vestibular. UNESPAR/FECILCAM.

1 INTRODUÇÃO

A disciplina de Estatística está presente em vários cursos de graduação. Apenas para exemplificar, entre os dez cursos de graduação ofertados pela UNESPAR/FECILCAM, sete deles possuem pelo menos uma disciplina dedicada ao estudo de técnicas estatísticas para análise de dados. O conhecimento de tais técnicas torna-se importante para auxiliar na compreensão da grande quantidade de informações numéricas a qual somos expostos diariamente e na tomada de decisões. Nesse sentido, acreditamos que deve haver uma preocupação por parte dos docentes em fornecer subsídios para que os acadêmicos tornem-se capazes de aplicar os conhecimentos proporcionados pela disciplina a problemas que possivelmente serão encontrados no exercício de suas profissões.

Nos livros didáticos podemos encontrar exemplos de aplicações de ferramentas estatísticas a problemas de diversas áreas, tais como administração, economia, engenharia, medicina, educação, entre outras. No entanto, é interessante que o professor explore também outras aplicações, preferencialmente aquelas que estejam mais próximas da realidade dos acadêmicos. Neste sentido, o presente trabalho tem o objetivo de mostrar uma possibilidade de aplicação de algumas técnicas estatísticas básicas à análise do perfil sócio-educacional e do desempenho dos candidatos ao vestibular de uma Instituição de

ensino superior, a partir dos dados oriundos do questionário sócio-educacional preenchidos pelos candidatos ao vestibular.

O preenchimento de um questionário sócio-educacional é uma prática comum no processo de inscrição ao vestibular de Instituições de Ensino Superior – IES. Com as informações contidas nesses questionários, pode-se traçar o perfil dos candidatos, bem como estabelecer relações entre as variáveis sócio-educacionais e o desempenho dos mesmos nas provas que fazem parte do processo de seleção. Tais informações podem auxiliar os administradores das IES na tomada de decisões visando a melhoria da qualidade do ensino e na elaboração de projetos que atendam as necessidades dos acadêmicos, principalmente no que diz respeito à permanência destes na Instituição.

O banco de dados utilizado nesse trabalho contém as informações referentes aos 1186 candidatos ao Vestibular de Verão 2007 (para ingresso em 2008) da UNESPAR/FECILCAM - Faculdade Estadual de Ciências e Letras de Campo Mourão - PR, que na época oferecia nove cursos: Administração (Adm.), Ciências Contábeis (C.C.), Ciências Econômicas (C.E.), Engenharia de Produção Agroindustrial (EPA), Geografia (Geo.), Letras, Matemática (Mat.), Pedagogia (Ped.) e Turismo e Meio Ambiente (TURMA). Hoje a Instituição conta com mais um curso, o de Licenciatura em História. No entanto, o processo seletivo ainda mantém o mesmo formato, sendo realizados dois vestibulares do tipo vocacionado por ano, o Vestibular de Inverno e o Vestibular de Verão, em junho e dezembro, respectivamente, do ano que antecede o ingresso dos aprovados na Instituição. No total, foram ofertadas 265 vagas por vestibular.

O presente trabalho tem apenas o objetivo de apresentar a possibilidade de aplicação de testes estatísticos estudados na graduação à análise do perfil dos candidatos ao vestibular de uma Instituição de ensino superior, bem como à análise do desempenho destes. Não se pretende definir as técnicas estatísticas utilizadas, mas simplesmente aplicá-las. Além disso, o banco de dados foi escolhido com o propósito de exemplificar a aplicação de tais técnicas e suas interpretações e por isso não houve a preocupação em analisar os dados dos alunos ingressantes em 2013. Tal análise pode ser desenvolvida a partir da discussão proposta aqui. Além disso, o bom entendimento desse exemplo pode contribuir para o desenvolvimento de aplicações em outras áreas.

2 O VESTIBULAR DA UNESPAR/FECILCAM

As provas do vestibular da UNESPAR/FECILCAM compreendem duas etapas que são desenvolvidas em dois dias consecutivos. Na primeira etapa os candidatos realizam as provas de formação específica determinadas pelo curso para o qual se inscreveram, com o objetivo de valorizar aspectos da área do curso, de acordo com os conhecimentos

adquiridos até o Ensino Médio. As provas são compostas por 15 questões objetivas valendo 3 pontos cada uma. Na segunda etapa, os candidatos realizam as provas de formação geral, comum a todos os cursos, também formadas por 15 questões objetivas, exceto a Redação. Nessa etapa, as questões da prova de Língua Portuguesa valem 4 pontos cada, enquanto as demais continuam valendo 3 pontos. A nota da redação pode variar de 0 a 80 pontos. A Tabela 1 apresenta a composição das provas de acordo com o curso escolhido pelo candidato.

Cursos	Formação Específica	Formação Geral
Grupo 1: Administração Ciências Contábeis Ciências Econômicas	Geografia História Matemática	Redação Língua Portuguesa Literatura da Língua Portuguesa Língua Estrangeira (Inglês/Espanhol)
Grupo 2: Geografia Letras Pedagogia Turismo e Meio Ambiente	Geografia História Biologia	Redação Língua Portuguesa Literatura da Língua Portuguesa Língua Estrangeira (Inglês/Espanhol)
Grupo 3: Matemática Eng. de Prod. Agroindustrial	Química Física Matemática	Redação Língua Portuguesa Literatura da Língua Portuguesa Língua Estrangeira (Inglês/Espanhol)

Tabela 1: Composição das provas

Todos os candidatos, exceto ao curso de Letras, têm a opção de escolha na prova de Língua Estrangeira entre Inglês e Espanhol. Os candidatos ao curso de Letras devem responder, obrigatoriamente, à prova de Língua Inglesa. Não existe uma nota de corte para a aprovação dos candidatos, sendo que a classificação é dada pela soma dos pontos obtidos nas provas de formação geral e específica. Porém, se o candidato zerar em alguma das provas, este é desclassificado.

O questionário sócio-educacional preenchido pelos candidatos no ato da inscrição está composto por 27 questões que fornecem informações que se destinam, após o devido tratamento estatístico, à pesquisa de aspectos sócio educacionais dos vestibulandos. Este questionário aborda questões relacionadas à renda mensal da família, nível de instrução dos

pais, moradia, tipo de formação no Ensino Médio, entre outras. A partir dessas informações podemos aplicar técnicas estatísticas para responder a questões como:

- Existe diferença significativa entre o desempenho dos candidatos que estudaram em escolas públicas e particulares?
- Existe diferença significativa entre o desempenho dos candidatos que fizeram ou não curso pré-vestibular?
- Existe uma relação de dependência entre a raça do candidato e o resultado?
- Existe diferença significativa entre as notas dos candidatos dos diferentes cursos em cada prova?

Silva e Perizaro (2009) usaram esse mesmo banco de dados e empregaram técnicas estatísticas multivariadas de reconhecimento de padrões para classificar os candidatos como aprovados ou reprovados de acordo com as variáveis sócio-educacionais. Isso exemplifica a possibilidade de explorar tais informações empregando técnicas estatísticas mais sofisticadas. No entanto, como mencionado anteriormente, neste trabalho nos limitaremos às técnicas básicas que são abordadas nos cursos de graduação, principalmente naqueles ofertados pela UNESPAR/FECILCAM.

3 ANÁLISES ESTATÍSTICAS BASEADAS NAS INFORMAÇÕES CONTIDAS NO QUESTIONÁRIO

Os dados utilizados nesse trabalho foram disponibilizados em uma planilha do Excel, que contém as respostas aos itens propostos no questionário sócio-educacional, o registro das notas nas provas realizadas pelos candidatos, além de outras informações extraídas da ficha de inscrição. A partir desse banco de dados pode-se trabalhar inicialmente com a construção de gráficos estatísticos. Por exemplo, ao analisar a variável qualitativa sexo, pode-se construir um gráfico de setores, como o apresentado na Imagem 1, construído no Excel. Vale lembrar que vários outros *softwares* estatísticos poderiam ser utilizados, por exemplo, o Minitab. Este gráfico indica que 58% dos candidatos ao Vestibular de Verão 2008 são do sexo feminino, o que evidencia o que muitos estudos têm mostrado em relação ao avanço da escolarização feminina. Pode-se também analisar o sexo de acordo com o curso escolhido.

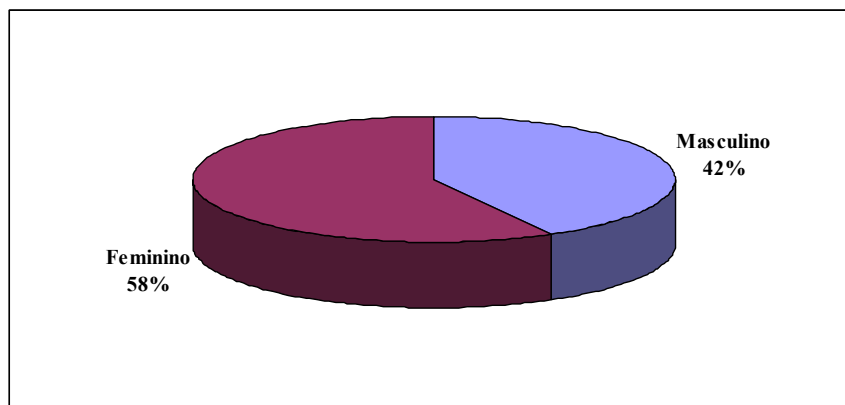


Imagem 1: Sexo dos candidatos ao vestibular da UNESPAR/FECILCAM

Uma das informações disponíveis nesse banco de dados diz respeito ao tipo de escola de origem dos candidatos, pública ou particular. Este conhecimento tem sua importância justificada pela crescente discussão existente nas Instituições de Ensino Superior acerca da implantação do sistema de cotas para alunos oriundos de escolas públicas. Consta-se que 1,94% dos candidatos inscritos no vestibular cursaram a maior parte do Ensino Médio em escolas particulares; 2,78% cursaram a maior parte em escolas públicas; 4,30% cursaram todo o Ensino Médio em escolas particulares e 90,98% cursaram somente em escolas públicas. Observa-se assim que a maioria, 93,76%, dos candidatos inscritos, cursou o Ensino Médio, ou pelo menos a maior parte dele, em escolas públicas.

Mesmo diante de um percentual baixo de alunos oriundos de escolas particulares, é interessante verificar a existência de dependência entre as variáveis relacionadas ao tipo de escola e ao resultado final no vestibular, o que pode auxiliar na tomada de decisão referente à implantação do sistema de cotas na referida Instituição. Esses dados podem ser organizados em uma tabela de contingência (Tabela 2) e a partir daí pode-se verificar se as variáveis são independentes por meio do teste χ^2 (Qui-Quadrado) (Ara, Mussetti e Schneiderman, 2003), possibilitando assim a aplicação de um Teste Não-Paramétrico.

Resultado/Escola	Pública	Particular
Aprovado	269	34
Reprovado	819	35

Tabela 2: Dados para o teste de independência entre tipo de escola e aprovação

O teste χ^2 testa a hipótese nula de que o resultado final no vestibular independe do tipo de escola em que o candidato cursou o Ensino Médio, contra a hipótese alternativa de que há dependência entre o tipo de escola e o resultado. Obteve-se o seguinte valor para a

estatística do teste $\chi^2 = 20,23$. O valor- p obtido é menor que o nível de significância $\alpha = 0,05$, portanto, deve se rejeitar a hipótese nula. Dessa forma, conclui-se que o tipo de escola em que o candidato cursou o Ensino Médio influencia em sua aprovação no vestibular da UNESPAR/FECILCAM. Isso pode levantar uma discussão sobre a qualidade do ensino nas escolas públicas.

Inicialmente foi citado que no referido vestibular houve um número de 1186 candidatos e que a UNESPAR/FECILCAM oferecia, na época, 265 vagas por vestibular, dividida entre seus nove cursos. Porém, no ano de 2007 somente o vestibular de Verão ofereceu vagas para o curso de Geografia diurno, dobrando assim o número de vagas oferecidas normalmente para esse curso. Dessa forma, totalizaram-se 305 vagas para este vestibular. Porém, na Tabela 2 pode-se observar que 303 candidatos foram aprovados e também que 1157 candidatos efetivamente fizeram a prova, os demais são considerados desistentes.

Explorando ainda o teste não-paramétrico χ^2 , pode-se verificar se o fato do candidato ter frequentado um curso pré-vestibular influencia em sua aprovação. A Tabela 3 apresenta os resultados referentes ao número de candidatos que frequentaram um curso pré-vestibular ou não e o resultado final no vestibular. O teste de independência revela que o curso pré-vestibular não influencia na aprovação. Pode-se notar 82,71% dos candidatos que realmente fizeram a prova, não buscaram um curso preparatório.

Resultado/Curso Pré-Vestibular	Não	Sim
Aprovado	248	55
Reprovado	709	145

Tabela 3: Dados para o teste de independência entre curso pré-vestibular e aprovação

O questionário proporciona também o conhecimento da raça dos candidatos ao vestibular. A Imagem 2 apresenta um gráfico de barras construído no *software* Minitab, indicando a porcentagem de candidatos brancos, negros, pardos, amarelos e indígenas. Pode-se perceber a predominância de candidatos da raça/cor branca. Essa informação se torna útil também para discussões referentes ao sistema de cotas baseado na raça do candidato.

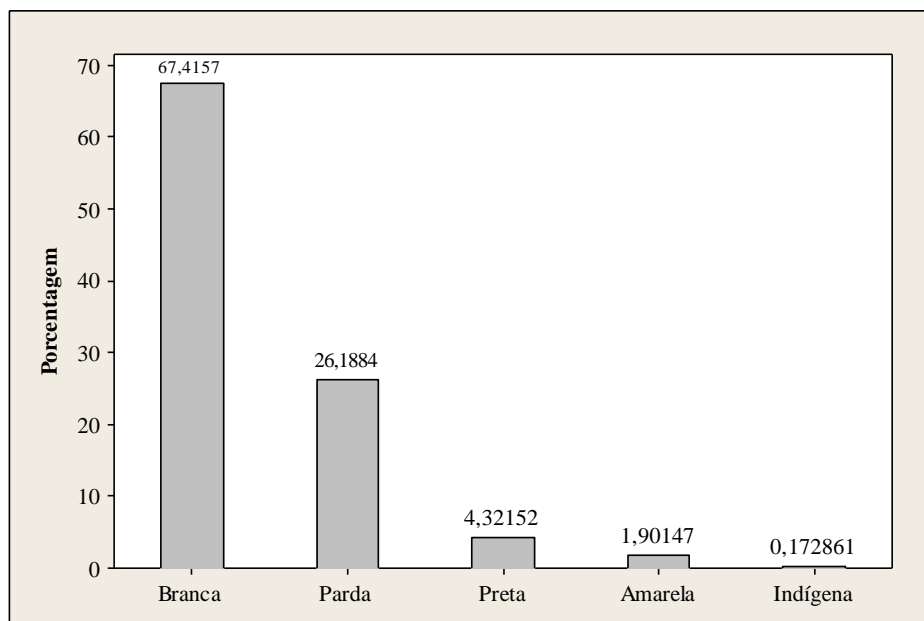


Imagem 2: Raça/cor dos candidatos

Uma outra questão contida no questionário sócio-educacional se refere à necessidade de um candidato aprovado trabalhar durante a graduação. Esta informação torna-se importante para que o corpo docente conheça a realidade dos ingressantes, uma vez que quando se trata de “alunos trabalhadores” o tempo para estudo é limitado, sendo assim pode-se trabalhar de forma que o acadêmico aproveite o máximo possível o tempo de aula. A questão apresentada no questionário é a seguinte: “Durante o curso superior, você terá que trabalhar?”. Entre os aprovados, 28,05% responderam “não sei”; 5,62% responderam “não”; 25,41% responderam “sim, desde o primeiro ano em tempo parcial”; 40,26% responderam “sim, desde o primeiro ano em tempo integral”; e 0,66% responderam “sim, apenas nos últimos anos”.

Pode-se notar que a maioria dos alunos que ingressaram terá que trabalhar no decorrer do curso, o que muitas vezes compromete a formação, pois para muitos desses alunos não será possível participar de algumas atividades extracurriculares oferecidas pela Instituição. Por outro lado, sabe-se que existem alunos que trabalham o dia todo e muitas vezes têm um aproveitamento superior ao de outros que, teoricamente, teriam tempo integral para se dedicar aos estudos.

Várias outras informações podem ser obtidas por meio do questionário sócio-educacional, sendo possível explorar os recursos existentes em *softwares* estatísticos que podem auxiliar na organização dos dados, visando a aplicação de ferramentas da estatística descritiva e inferencial aos mesmos.

4 ANÁLISE DO DESEMPENHO DOS CANDIDATOS

Os dados referentes às notas obtidas em cada uma das provas possibilitam a realização de comparações entre os desempenhos de candidatos dos diferentes cursos ou grupos. A primeira análise a ser feita diz respeito ao desempenho dos candidatos a cada um dos nove cursos nas provas de formação geral, pois estas são comuns a todos os vestibulandos. Assim, pode-se, por exemplo, levantar a seguinte questão: Existe diferença significativa entre o desempenho na Redação para os candidatos de cada curso?

Para responder a essa questão pode-se construir um gráfico (Imagem 3) para facilitar a visualização e aplicar a ANOVA (Análise de Variância), para verificar a existência de diferença significativa entre as médias (Bussab e Morettin, 2002). Pode-se perceber ao analisar a Imagem 3, que os candidatos ao curso de Ciências Econômicas obtiveram o melhor desempenho na Redação. Por outro lado, os candidatos ao curso de Turismo e Meio Ambiente obtiveram a menor média. Pode-se observar também que o curso de Turismo e Meio Ambiente é o que apresenta maior variação entre as notas dos candidatos. Os resultados da ANOVA revelaram que existe diferença significativa, ao nível de 5%, entre as notas médias dos candidatos na Redação.

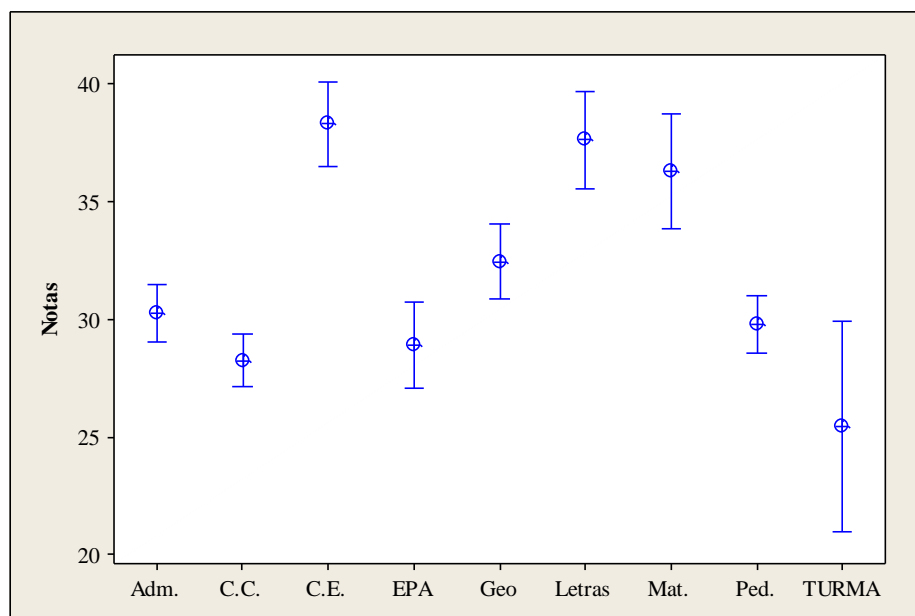


Imagem 3: Comparação entre o desempenho dos os candidatos na prova de Redação

Aplicando o teste de Tukey para comparação entre as médias pode-se verificar entre quais cursos existe diferença significativa entre o desempenho dos candidatos. A Tabela 3 apresenta entre quais cursos existe essa diferença, sendo indicada por um "X".

Cursos	Adm.	C.C.	C.E.	EPA	Geo	Letras	Mat.	Ped.	TURMA
Adm.			X			X	X		
C.C.			X		X	X	X		
C.E.	X	X		X	X			X	X
EPA			X			X	X		
Geo.		X	X			X			X
Letras	X	X		X	X			X	X
Mat.	X	X		X				X	X
Ped.			X			X	X		
TURMA			X		X	X	X		

Tabela 3: Diferença entre o desempenho médio dos candidatos a cada curso na prova de Redação

Um outro gráfico (Imagem 4) foi construído com o objetivo de comparar o desempenho na Redação entre os candidatos aprovados. Pode-se notar que entre os aprovados os que obtiveram o melhor desempenho foram aqueles que escolheram o curso de Matemática, o que contradiz a ideia de que estudantes que escolhem esse curso possuem dificuldades com a escrita. Já a menor média continua sendo do curso de Turismo e Meio Ambiente. Ao aplicar a ANOVA aos dados do desempenho dos aprovados, concluiu-se que também existe diferença significativa entre as médias, ao nível de 5% de significância.

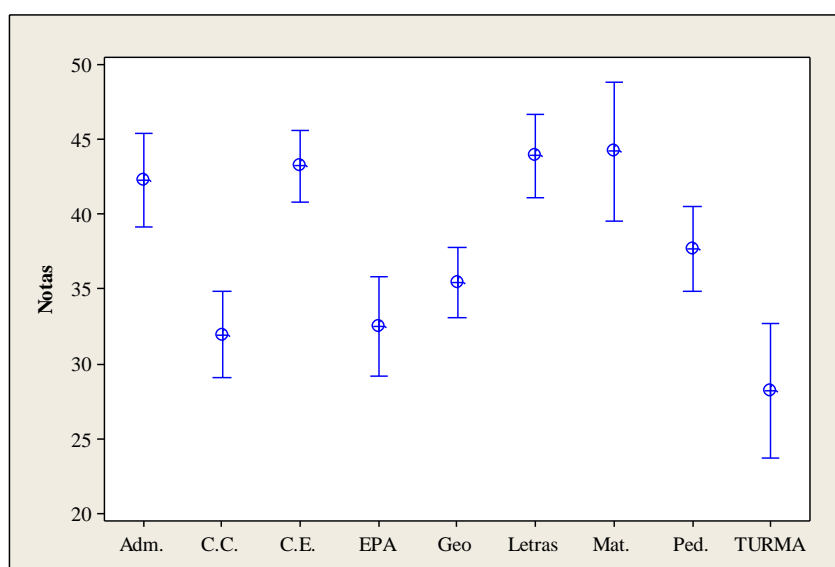


Imagem 4: Comparação entre o desempenho dos os candidatos aprovados na prova de Redação

Análises como as apresentadas anteriormente podem ser realizadas considerando as demais disciplinas de formação geral, ou ainda pode-se realizar comparações entre o desempenho dos candidatos de um mesmo grupo nas provas de formação específica.

Outra questão que pode ser levantada diz respeito ao desempenho dos candidatos aprovados a um determinado curso nas provas realizadas por estes. Por exemplo, os candidatos ao curso de Matemática realizaram as provas de Língua Portuguesa (Pot.), Literatura da Língua Portuguesa (Lit.), Língua Estrangeira (L. Estr.), Redação (Red.), Matemática (Mat.), Física (Fís.) e Química (Quí.), sendo assim pode-se verificar se realmente acontece o que todos esperam, ou seja, o desempenho desses candidatos na prova de Matemática é melhor que o desempenho nas provas das demais disciplinas, já que esta disciplina está inteiramente ligada ao curso escolhido.

Um *boxplot* (Imagem 5) foi construído a partir dos dados referentes às notas dos candidatos aprovados para o curso de Matemática em todas as provas por eles realizadas. Esta mesma análise poderia ser feita levando em consideração todos os candidatos a esse curso.

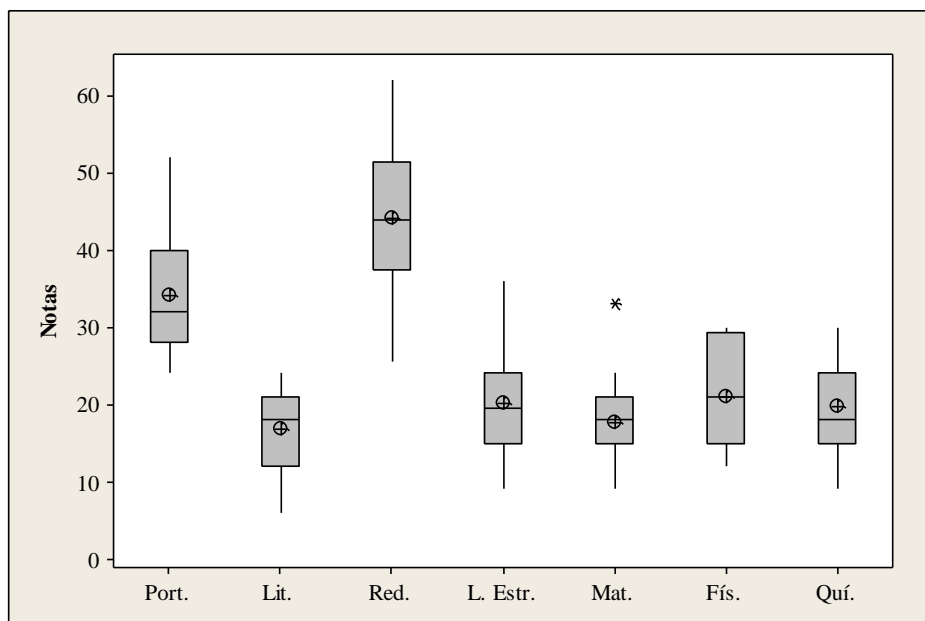


Imagem 5: Desempenho dos candidatos ao curso de Matemática nas diferentes provas

Por meio da Imagem 5 pode-se constatar que os candidatos ao curso de Matemática obtiveram melhor desempenho na Redação, porém entre as provas com questões objetivas os candidatos se saíram melhor em Língua Portuguesa, com média de 34,2, apresentando pior desempenho em Literatura da Língua Portuguesa, com média 16,8. Como cada questão da prova de Língua Portuguesa tem peso quatro, pode-se concluir que em média os candidatos acertaram aproximadamente 8 questões entre as 15 questões

propostas. Já as questões da prova de Literatura têm peso três, assim os candidatos acertaram em média aproximadamente 6 questões.

É interessante observar que a média das notas dos candidatos na prova de Matemática foi a segunda menor comparada às demais provas, com média 17,7, ou seja, os candidatos acertaram em média 5 questões. Além disso, o *boxplot* permite verificar a presença de um *outlier* (valor discrepante) entre as notas nesta prova. Esse *outlier* corresponde à maior nota obtida por um dos candidatos, que foi de 33. Isto significa que o maior número de acertos nessa prova foi de 11 questões. Não se pode esquecer que o grau de dificuldade das provas influencia na diferença entre o desempenho dos candidatos.

Como apresentado anteriormente, a maioria dos vestibulandos cursaram o Ensino Médio em escolas públicas, sendo assim pode-se utilizar os dados referentes ao desempenho dos candidatos em diferentes disciplinas como um instrumento de avaliação do ensino destas. Como exemplo, podemos considerar as provas de formação específica realizadas pelos candidatos aos cursos do Grupo 1: Administração, Ciências Contábeis e Ciências Econômicas. A Tabela 4 apresenta a média obtida em cada uma das provas realizadas por estes candidatos.

	Matemática	História	Geografia
Média	10,287	14,528	12,492

Tabela 4: Média dos candidatos aos cursos do Grupo 1 nas provas de formação específica

Em média os candidatos aos cursos do Grupo 1 acertaram aproximadamente 3 questões em Matemática, 5 questões em História e 4 em Geografia. A ANOVA indica que existe diferença significativa entre o desempenho dos candidatos nas provas de formação específica. A maior média obtida foi na prova de História, mas mesmo assim é preocupante o desempenho dos candidatos nessas provas, pois o número de acertos é muito baixo em cada uma delas. Algumas questões poderiam ser levantadas: “o problema está no ensino dessas disciplinas?” ou “as questões propostas tinham alto grau de dificuldade?” ou ainda “os candidatos não estavam preparados?”.

5 CONSIDERAÇÕES FINAIS

Diante do exposto neste trabalho, pode-se perceber que muitos dos conteúdos estudados na disciplina de Estatística, particularmente daqueles abordados nos cursos de graduação da UNESPAR/FECILCAM, podem ser explorados a partir dos dados referentes ao questionário sócio-educacional aplicado aos vestibulandos e aos dados referentes às notas obtidas nas provas que fazem parte do processo de seleção da Instituição. O objetivo

aqui não era o de apresentar o perfil dos candidatos ao vestibular da UNESPAR/FECILCAM, mas sim mostrar as possibilidades de análises que podem ser feitas a partir desses dados, proporcionando aos acadêmicos um momento de discussões acerca da realidade dos estudantes da Instituição.

Como podemos notar a partir dos exemplos apresentados, esse tema também proporciona a discussão sobre diversos temas, tais como, sistema de cotas, a predominância do sexo feminino buscando ensino superior, qualidade das escolas públicas, desempenho dos candidatos em cada uma das provas realizadas, etc. Além disso, para realizar análises como as propostas neste trabalho, os acadêmicos terão que trabalhar com *softwares* estatísticos que facilitam a tabulação dos dados, já que existe uma grande quantidade destes. Dessa forma, pode-se atingir o objetivo do ensino da Estatística, que não é somente o de fornecer subsídios para que os alunos saibam quais técnicas estatísticas podem ser aplicadas a determinadas situações, mas principalmente saber compreender os resultados apresentados para utilizá-los da melhor forma possível.

REFERÊNCIAS

ARA, A. B.; MUSETTI, A. V. e SCHNEIDERMAN, B. **Introdução à Estatística**. São Paulo: Edgard Blücher LTDA, 2003.

BUSSAB, W. O. e MORETTIN, P. A. **Estatística Básica**. São Paulo: Saraiva, 2002.

SILVA, T. C. e PERICARO, G. A. Classificação dos candidatos ao vestibular da UNESPAR/FECILCAM via técnicas estatísticas multivariadas. *Anais do CNMAC* v. 2, 2009, p. 566-571.