



## AVALIAÇÃO DOCENTE COMO MEDIÇÃO DA QUALIDADE DE ENSINO?

Carlos Alberto de Almeida Vilela<sup>1</sup>

### RESUMO

A avaliação docente pode ser uma ferramenta de alto valor desde que seja elaborada de forma confiável. Como qualquer pesquisa de opinião, deve-se observar alguns aspectos matemáticos formais importantes para validação da pesquisa, caso contrário corre-se o risco de ter em mãos um resultado totalmente fora da realidade e com isso a tomada de decisão pode ser desastrosa. Com base na análise de dados de avaliação docente o objetivo deste trabalho é levantar a discussão da credibilidade estatística deste tipo de avaliação e apontar alguns procedimentos para sua validação.

**Palavras-chave:** Educação Superior; Avaliação docente; Análise estatística; Credibilidade matemática.

### INTRODUÇÃO

Toda avaliação docente discutida até hoje e amplamente utilizada pelas IES estão fundamentadas no princípio básico de que o aluno está de forma honrosa e responsável respondendo aos questionamentos com a verdade dos fatos. Toda a pesquisa está baseada na boa vontade, honestidade e responsabilidade do aluno, algo que na opinião do autor é discutível por muitos fatores:

- a) o aluno não tem obrigatoriedade de responder à pesquisa e talvez por isso várias IES tenham baixíssimos índices de participação;
- b) as informações em geral são sigilosas e por isso o aluno se sente seguro para responder qualquer coisa;
- c) não há punição alguma para o aluno que fraudar ou não participa da pesquisa;
- d) nem todos os alunos estão agindo com honestidade.

O conceito da avaliação docente está inserido nas instituições de ensino já há bastante tempo, mas teve em 2004 um importante marco na sua concepção a partir da criação do Sistema Nacional de Avaliação da Educação Superior (SINAES) que regulamenta vários procedimentos para avaliação dos cursos no Brasil com o objetivo de quantificar e garantir a qualidade do ensino superior no Brasil. Os seus objetivos são “a melhoria da qualidade da

---

<sup>1</sup> Professor do curso de Engenharia Mecânica - UFG, carlosavgeo@yahoo.com;



educação superior, a orientação da expansão da sua oferta, o aumento permanente da sua eficácia institucional e efetividade acadêmica e social”.

No universo que compõe a avaliação docente encontra-se a sua avaliação docente pelo aluno, que deveria ser um canal de retorno de informações sobre o que acontece dentro da sala de aula sob o olhar do outro, entretanto na maioria das vezes a forma como é realizada a pesquisa hoje, majoritariamente com perguntas subjetivas, perde-se esta capacidade de feedback positivo e transforma-se em um momento de retaliações do professor versus aluno e vice-versa e com isso a desmotivação e mal estar docente é inevitável, (SANTO e SANTOS, 2010).

Deve ser lembrado que o programa de avaliação de desempenho está intrinsecamente relacionado com o Planejamento Estratégico e fornece subsídios imprescindíveis para os programas de gestão de pessoas, cargos e salários, treinamento e desenvolvimento, qualidade, financeiro e participação nos lucros ou resultados, (PONTES, 2002).

Para que o objetivo inicial da avaliação docente seja atingido, para que situações constrangedoras entre docentes e alunos sejam evitadas e que principalmente tomadas de decisão sejam feitas baseadas em dados verdadeiros, como garantir a veracidade dos resultados obtidos na avaliação docente? Uma forma que o autor propõe é basear-se na análise estatística dos resultados.

Lembremos do panorama geral que é uma sala de aula: tem-se um professor convivendo com vários alunos, ou seja são várias pessoas que passam pelas mesmas situações, nos mesmos locais, nos mesmos horários, nas mesmas circunstâncias, nos mesmos dias. Isto deveria resultar em uma homogeneidade nas respostas, sendo elas positivas ou negativas, pelo menos nos quesitos objetivos já que a relação entre professor-aluno é a mesma para todos. O que o professor faz para um aluno ele faz para o outro, segundo o princípio da isonomia.

## **BACKGROUND MATEMÁTICO**

A quase totalidade das IES hoje utiliza como parâmetro principal de medição da qualidade da atividade docente, o valor da média aritmética das notas atribuídas pelos alunos. Talvez por ser um parâmetro fácil de calcular ou pelo fato de que o conceito “média” estar formado no consciente coletivo desde a educação básica, mas este não é o único e nem é o melhor parâmetro para representar um conjunto de dados. A seguir apresenta-se alguns outros



parâmetros que poderiam ser aplicados para melhor entender o universo das opiniões na pesquisa.

**Média Aritmética** é a forma mais usual de cálculo de um valor representativo de um conjunto de dados. É usada para distribuições simétricas ou quase simétricas, ou para distribuições que têm um único pico dominante. É calculada segundo a equação abaixo somando-se as observações e dividindo-se pelo número de observações.

$$\bar{X} = \frac{\sum_{i=1}^n x_i}{n} \quad (1)$$

**Mediana** de um conjunto ordenado é o valor do elemento central. Este valor coincide com o elemento central quando o conjunto tem número ímpar de elementos, e quando o conjunto tem número par de elementos a mediana é calculada como sendo a média aritmética dos dois elementos centrais. A mediana é utilizada para reduzir os efeitos dos valores extremos encontrados no conjunto.

$$Med = x_{\frac{n+1}{2}} \quad \text{ou} \quad Med = \frac{x_{\frac{n}{2}} + x_{\frac{n}{2}+1}}{2} \quad (2)$$

**Moda** é o valor que ocorre com maior frequência dentro do conjunto considerado. É frequentemente utilizada para representar distribuições extremamente assimétricas, situações irregulares onde dois picos são encontrados, ou para eliminar os efeitos dos valores extremos. Em um conjunto onde dois ou mais valores se repetem com a mesma frequência este é chamado de multimodal.

**Amplitude total** é a diferença entre os valores máximo e mínimo do conjunto de dados. Como a amplitude é baseada em dois números, ela se torna representativa na análise quando o número de elementos no conjunto normalmente é pequeno.

$$A = (x_i|_{\max} - x_i|_{\min})_{i=1}^n \quad (3)$$

**Frequência** é a medida da ocorrência de um valor dentro do conjunto. Quando classificados ordenadamente os elementos em classes e colocadas estas informações em forma de tabela, é criada uma tabela de distribuição de frequência.

**Desvio médio** é definido como sendo a média dos valores absolutos dos desvios em relação a média aritmética.

$$D_{med} = \frac{1}{n} \sum_{i=1}^n |x_i - \bar{X}| \quad (4)$$

### Características do Desvio Médio

- Depende de todos os valores da distribuição
- Seu cálculo pode ser efetuado a partir da média ou da mediana
- Não leva em consideração a existência de desvios negativos

**Desvio padrão** é a mais importante medida de dispersão estatística. O cálculo é feito segundo a equação abaixo que é a raiz quadrada da média do quadrado dos desvios em relação a média aritmética, quando o número de elementos considerado for pequeno.

$$\sigma = \sqrt{\frac{1}{n} \sum_{i=1}^n (x_i - \bar{X})^2} \quad (5)$$

Quando o número de elementos do conjunto for grande, em geral maior que 30 elementos, a equação pode ser aproximada para a seguinte forma:

$$s = \sqrt{\frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{X})^2} \quad (6)$$

O desvio padrão como sendo a medida mais importante de variação, ela determina a dispersão dos valores em relação à média e mede a dispersão absoluta ou a variabilidade de uma distribuição. Um conjunto de dados com desvio padrão pequeno representa uma homogeneidade dos valores e um conjunto de dados com desvio padrão grande representa uma heterogeneidade dos valores. Em uma pesquisa de opinião isso representa o consenso das opiniões ou não.

**Coefficiente de Variação** é uma medida de dispersão relativa que indica a relação percentual entre o desvio padrão e a média aritmética do conjunto.

$$C_{\text{var}} = 100 \frac{\sigma}{\bar{X}} \quad (7)$$

**Erro Padrão** da média é um valor que corrige a inexatidão da média como valor representativo do conjunto. Ele é calculado segundo a equação abaixo:

$$E_{\bar{X}} = \frac{\sigma}{\sqrt{n}} \quad (8)$$

Algumas características deste valor são:

- Quanto maior o desvio padrão, pior é a média como medida característica da amostra;
- Quanto maior a amostra, menor será o erro padrão e, teoricamente, a média torna-se, para grandes amostras, uma medida característica da tendência central dos dados.



O erro padrão representa matematicamente que a média aritmética não pode ser considerada um valor representativo de um conjunto se o mesmo não for homogêneo. A irregularidade do conjunto é expressa pelo desvio padrão – quanto maior o desvio padrão, menos homogêneos são seus elementos.

## UMA PESQUISA DE QUALIDADE

É de senso comum que alguns aspectos de uma pesquisa para torná-la confiável devam ser considerados e aplicados com o maior rigor possível, caso contrário todo o trabalho realizado poderá ser contestado pela falta de um ou mais aspectos que garantam a sua credibilidade. O (SEBRAE-MG, 2003) em material publicado direcionado à pesquisa de mercado, apresenta alguns aspectos importantes que devem ser seguidos em uma pesquisa:

- A redação precisa ser clara, simples e objetiva.
- As perguntas devem ser curtas e de fácil entendimento.
- As perguntas devem evitar distorções.

Segundo (FERNANDES, 1999), podem surgir dificuldades no processo de amostragem e dois tipos de erros podem ocorrer:

1. *Erros de amostragem aleatórios* que aparecem no processo de amostragem. Dão origem a resultados diferentes dos que seriam obtidos se fosse usado um censo. São devidos à aleatoriedade da repetição de uma experiência e são fáceis de ultrapassar, bastando para isso repetir várias vezes a experiência para obter outras amostras.

2. Os *erros de amostragem não aleatórios* que não estão relacionados com a seleção da amostra. Muitas vezes são devidos à complexidade do comportamento humano. São exemplos típicos destes erros, que surgem em inquéritos:

- a falta de dados, que pode aparecer quando não é possível contatar um sujeito (que tenha sido selecionado para a amostra) ou o sujeito se recusa a responder ao questionário;
- erros nas respostas, quando o sujeito pode deliberadamente mentir ao responder às questões colocadas, ou pode não ter entendido a questão;
- erros no processamento de dados, que ocorrem no processo mecânico do cálculo de quantidades numéricas e no processo de introdução de dados para um ficheiro;
- efeitos do método usado para a obtenção dos dados, que têm a ver com os inquiridores, com a altura em que é realizado o inquérito, com a linguagem utilizada nas perguntas e com o meio utilizado (correios, telefones ou contatos diretos).



Consideremos os resultados de uma avaliação docente realizada na Universidade Veiga de Almeida, apresentados por (BOCLIM, 2004).

Na referida pesquisa, as repostas variavam em uma escala de um (1) a cinco (5), sendo 1 o valor representativo para baixa qualidade e 5 o valor representativo para alta qualidade. A Tabela (1) apresenta os valores percentuais dos alunos que responderam de acordo com a nota atribuída ao professor no quesito assiduidade as aulas. Importante lembrar que este é um ponto objetivo, ou seja, ou o professor compareceu as aulas ou não compareceu as aulas, além do que esta informação pode ser facilmente verificada em uma folha de ponto da instituição.

Tabela (1). Assiduidade às aulas, segundo (BOCLIN, 2004)

Curso	1	2	3	4	5
Administração	9,7%	7,1%	12,8%	27,0%	43,4%
Ciências Contábeis	3,6%	8,4%	6,0%	15,7%	66,3%
Comunicação Social	4,3%	4,1%	11,0%	18,1%	62,4%
Turismo	9,7%	4,5%	6,7%	20,0%	59,1%
Serviço Social	2,5%	2,1%	7,0%	17,7%	70,7%
Gestão Empresarial	1,0%	3,6%	6,1%	29,9%	59,4%
Marketing	0,0%	2,0%	3,9%	2,0%	92,2%
Gestão Imobiliária	0,6%	1,1%	5,9%	20,8%	71,6%
Técnica Propaganda e Marketing	10,3%	5,7%	11,1%	16,8%	56,1%

Analisando a turma de Técnica Propaganda e Marketing onde houve a maior amplitude nas respostas sendo que 10,3% disseram que o professor não tem assiduidade as aulas e 56,1% dos alunos disseram que o mesmo professor tem excelente assiduidade as aulas. Se todas as variáveis são as mesmas, o mesmo professor, a mesma turma, os mesmos horários de aula, os mesmos alunos, as mesmas salas de aula, etc, como pode ser possível uma parte da turma dizer que o professor não compareceu as aulas e uma parte da turma dizer exatamente o contrário? Este é um quesito onde se espera a maior uniformidade nas respostas sejam elas positivas ou negativas, e não uma amplitude como a apresentada na tabela.

O que os resultados nos diz? Os alunos não entenderam a pergunta e responderam assim mesmo ou os alunos de forma deliberada não responderam com a verdade dos fatos. De qualquer forma, os resultados não correspondem com a verdade e deveriam ser descartados da pesquisa, entretanto não foi isso que aconteceu e decisões estratégicas foram tomadas



baseadas em informações erradas. Se isto ocorreu com um quesito objetivo, o que pode ocorrer com uma pergunta subjetiva?

Esta é a importância da utilização de ferramentas para detectar possíveis anomalias nas respostas apresentadas pelos alunos em uma avaliação docente.

## METODOLOGIA

O objetivo deste trabalho não é discutir a metodologia de aplicação dos questionários utilizados no processo de avaliação institucional, isto cabe às CAD (Comissões de Avaliação Docente) que aprova o formato mais adequado, mas uma vez aplicado o processo, a proposta do trabalho é avaliar por meio de parâmetros estatísticos se aqueles resultados obtidos são dignos de credibilidade ou não.

Com base nos dados disponíveis das avaliações docentes, a metodologia de análise consiste na aplicação dos conceitos estatísticos apresentados anteriormente e posterior discussão dos resultados obtidos.

O formulário de avaliação utilizado e disponibilizado pela instituição está apresentado na Tabela (2) onde estão listadas na ordem sequencial as perguntas às quais o aluno é submetido.

Tabela 2: Formulário de avaliação UFG – Universidade Federal de Goiás

N.	Questão	Objetividade
1	Comparece às aulas / orientações / chats sob sua responsabilidade.	Objetivo. Ele foi ou não foi à aula.
2	Respeita os horários de início e término das aulas.	Objetivo. Ele foi ou não foi à aula no horário previsto.
3	Cumprir o plano de ensino da disciplina.	Objetivo. É o seu plano de ensino.
4	Estabelece relação entre conteúdos abordados na prática profissional.	Subjetivo. Nem sempre o aluno consegue verificar esta conexão pois ele nem sempre tem conhecimento da prática profissional.
5	Domina e apresenta com clareza os conteúdos abordados.	Subjetivo. Como o discente que está aprendendo vai identificar se o docente sabe ou não sabe do assunto?
6	Utiliza metodologias diversificadas e/ou instigadoras de aprendizagem (ex. estudo de caso, seminários, grupos de discussão, aulas práticas, pesquisa de campo, outras).	Subjetivo. Isto só mostra que o professor é multimídia. Se o docente for tradicional, o que não significa ser ruim, ele será mal avaliado.
7	É coerente ao avaliar a aprendizagem.	Subjetivo. Se o aluno não souber responder a avaliação ele vai considerar que não foi coerente.
8	Devolve as avaliações e esclarece as dúvidas.	Objetivo.
9	Age de forma cortês e ética com os estudantes.	Subjetivo. Se o docente for de outra cultura onde o tratamento interpessoal for diferente do tratamento que se dá no Brasil, ele será mal avaliado.
10	Você faria outra disciplina com este professor?	Que tipo de pergunta é essa???



Foram utilizados os resultados da avaliação docente de um professor da instituição considerando as quatro disciplinas sob sua responsabilidade no ano. A escolha do referido professor não foi aleatória, e sim intencional motivada pela suspeição da anormalidade dos resultados. Resultados tais que foram considerados válidos pela instituição para compor sua avaliação de desempenho anual.

## ANÁLISE DOS DADOS

Foram analisados dados obtidos a partir da pesquisa realizada referentes a quatro turmas e os resultados das notas atribuídas estão apresentados a seguir.

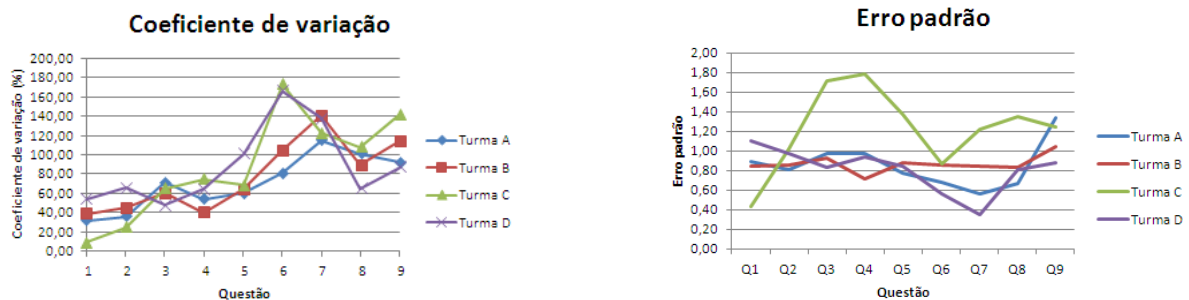


Figura 1. Coeficiente de variação e Erro padrão

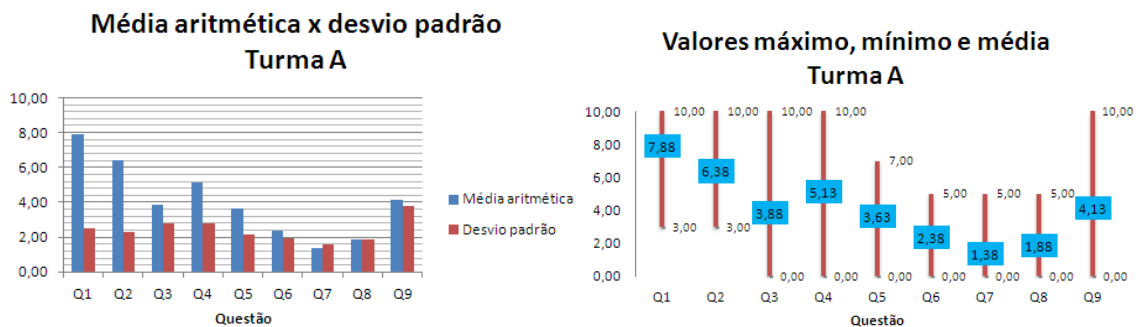


Figura 2. Média aritmética x Desvio Padrão e Valores Máximos, Mínimos e Média Aritmética. Turma A.

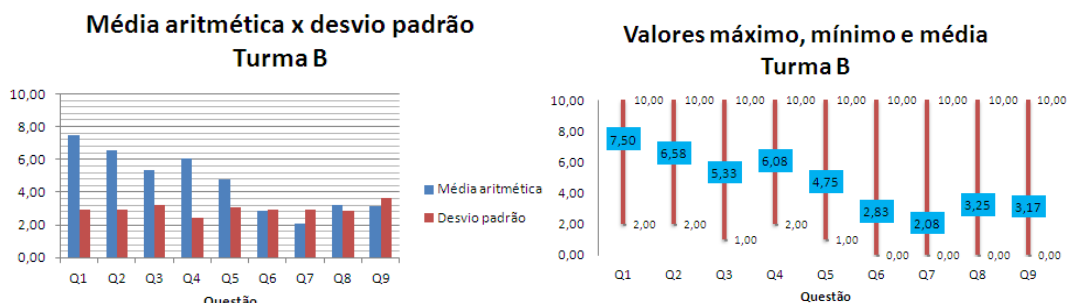
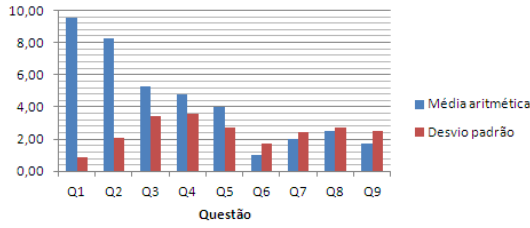


Figura 3. Média aritmética x Desvio Padrão e Valores Máximos, Mínimos e Média Aritmética. Turma B.



**Média aritmética x desvio padrão Turma C**



**Valores máximo, mínimo e média Turma C**

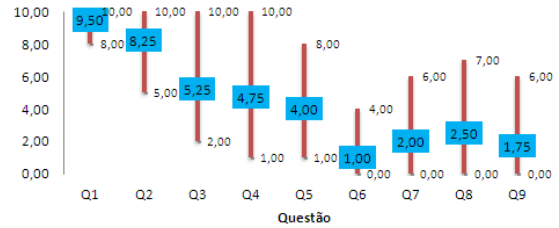
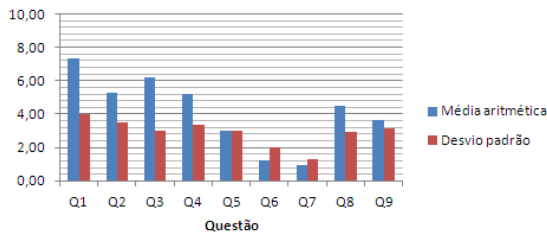


Figura 4. Media aritmética x Desvio Padrão e Valores Máximos, Mínimos e Média Aritmética. Turma C.

**Média aritmética x desvio padrão Turma D**



**Valores máximo, mínimo e média Turma D**

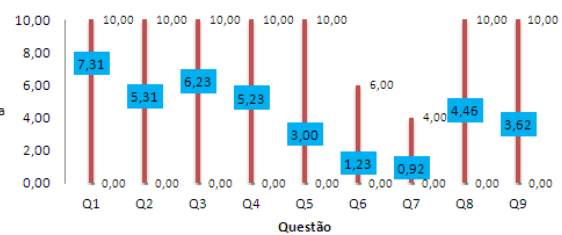


Figura 5. Media aritmética x Desvio Padrão e Valores Máximos, Mínimos e Média Aritmética. Turma D.

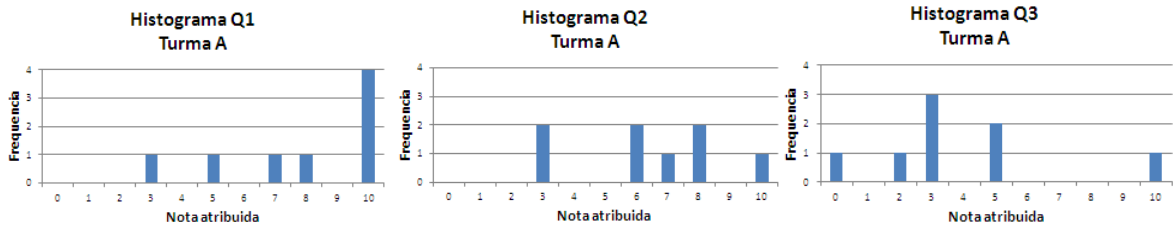


Figura 6. Frequência das notas. Turma A.

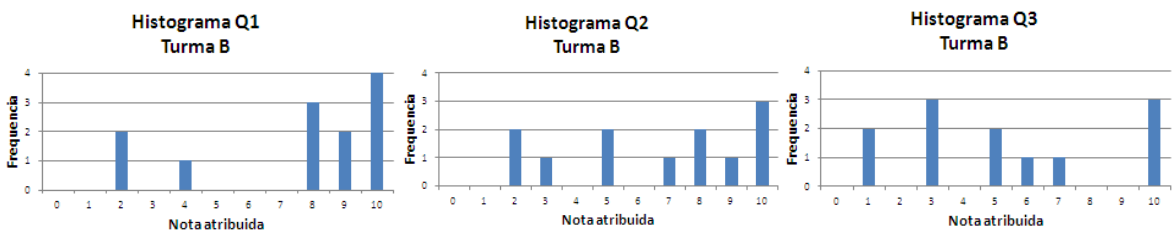


Figura 7. Frequência das notas. Turma B.

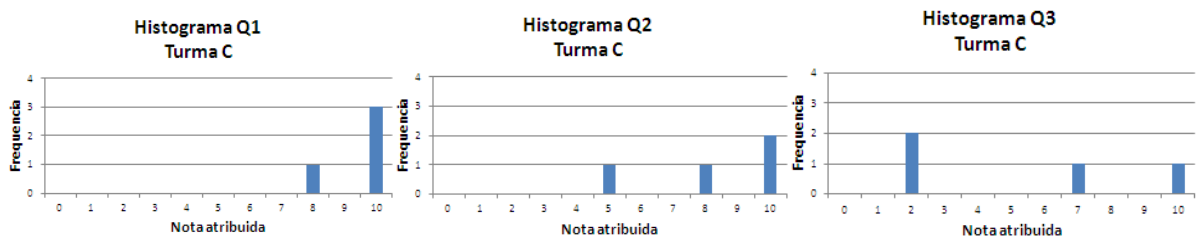


Figura 8. Frequência das notas. Turma C.

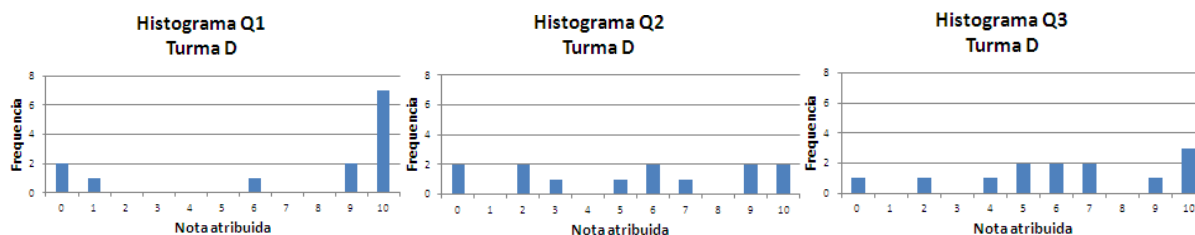


Figura 9. Frequência das notas. Turma D.

Uma análise cuidadosa dos histogramas apresentados nas Figuras 6 a 9 mostra que em todas as questões objetivas, Q1 a Q3, em praticamente todos os casos tem-se avaliações em lados extremos da escala, com exceção do histograma Q1 - Turma C. O que deve ser evidenciado aqui não é o valor da frequência das notas atribuídas, mas sim o fato de ter notas em lados extremos da escala em uma questão objetiva, que não é dependente de uma opinião do discente. Com base nesta observação conclui-se que: o discente não tem a menor ideia do que está respondendo no questionário de avaliação ou deliberadamente atribuiu notas que não refletem a realidade dos fatos.

Em qualquer um dos casos, a análise revela que o discente não tem condições de realizar uma avaliação deste tipo, sendo por não entender a pergunta ou por não responder de forma correta, e isto coloca em dúvida a credibilidade da pesquisa, lembrando que estas são questões objetivas. A pergunta que se faz agora é: o que pode acontecer nas questões subjetivas que envolvem uma análise pedagógica?

É comum encontrar em distribuições de dados quantitativos uma distribuição assimétrica e isto pode ser observado quando os valores da média e da mediana são muito diferentes ou quando o valor do desvio padrão é da mesma ordem de grandeza do valor da média. Casos onde a distribuição é assimétrica quando um grupo está concentrado em um extremo do conjunto, a mediana se torna um valor mais representativo da distribuição que a média.

Nos gráficos de comparação entre média e desvio padrão, em apenas um caso o valor do desvio padrão está uma ordem de grandeza menor que a média. Se considerado um critério de validação que é a comparação entre a média aritmética e o desvio padrão, apenas 2,77% dos casos poderiam ser considerados representativos. A pergunta que se faz agora é: o que acontece com os dados que estão nesta condição?

Analisando os gráficos de Valores máximo, mínimo e média, observa-se que em todas as questões consideradas foram obtidos conceitos máximos e mínimos, tanto para questões



objetivas quanto para as subjetivas. Em relação às questões objetivas, ou seja, não é uma expressão de opinião pessoal, não deveria haver tal amplitude, pois o fato constatado é o mesmo para todos os discentes. A amplitude apresentada mostra que discentes responderam verdadeiro para um fato ou responderam falso para o mesmo fato. Este comportamento coloca a credibilidade das respostas em dúvida e os resultados deveriam ser analisados com bastante critério até mesmo considerando-se as possibilidades de anulação dos dados ou nova realização de pesquisa.

## CONCLUSÕES

Não foi discutido aqui, e nem é o objetivo deste trabalho, o mérito da ferramenta administrativa Avaliação Docente, se é ou não importante, mas o que se deseja expor é a fragilidade que se tem na execução deste tipo de pesquisa se não forem observados os resultados com cuidado.

Como garantir que os resultados obtidos na pesquisa realmente refletem os acontecimentos reais dentro da sala de aula? Este é o ponto principal da discussão. Não se podem tomar decisões baseadas em informações duvidosas. Isto é evidente quando analisamos os resultados referentes às perguntas objetivas tais como "O professor devolve as avaliações?" e temos com resultado uma dispersão muito grande nas opiniões. Valores de vão desde o zero (não devolve) até o dez (devolve). Como é um ponto objetivo, ou o professor devolve ou não devolve para todos os alunos da turma, ele não faz isso de forma seletiva para alguns escolhidos, ou seja, o aluno não entendeu a pergunta ou o aluno mentiu na resposta. De qualquer forma este é um resultado que não deveria ser contabilizado na avaliação do docente, entretanto não é isso que se observa na prática.

A forma como são executadas as pesquisas deixa muita margem para erros e falsas impressões o que pode levar a decisões errôneas. Os parâmetros estatísticos aqui apresentados deveriam servir como base para validação da pesquisa e não simplesmente calculado, tabelado e sem uso prático posterior.

Algumas sugestões podem ser utilizadas para melhoramento do processo de avaliação:

- Uso de questões objetivas que possam ser confrontadas por meio de documentos.
- Número mínimo de discentes na avaliação para assegurar a representatividade.
- Valores limite máximo de desvio padrão e amplitude total para assegurar a uniformidade das opiniões.



- Cruzamento dos resultados obtidos com resultados de auto avaliação docente e discente.

Enquanto alterações desta magnitude não são aplicadas, ainda fica a pergunta: o que se faz com pesquisas cujos resultados não estão dentro do mínimo aceitável de qualidade?

O presente autor não tem relato de nenhuma instituição de ensino que frente a resultados desta natureza, tenha refeito o processo de avaliação, e então os resultados não significativos são utilizados assim mesmo. Enquanto este tipo de ocorrência continuar a acontecer, as instituições de ensino correm o sério risco de tomarem decisões baseadas em informações erradas. Talvez seja este um dos motivos de tantas universidades hoje estarem com médias avaliativas tão baixas do ENADE e a causa de um índice de reprovação tão alto no exame da ordem da OAB mesmo quando observadas as instituições com altos índices de aprovação docente.

Na opinião do autor, hoje a avaliação docente como é realizada e na forma que é utilizada, é totalmente desacreditada.

Após uma análise rápida das notas de desempenho dos cursos na avaliação ENADE, não é difícil encontrar cursos com conceitos baixos mas que apresentam uma aprovação quase que unânime do seu corpo discente.

## **REFERÊNCIAS**

- BOCLIN, R.; Avaliação de Docentes do Ensino Superior: Um Estudo de Caso, Ensaio: aval. pol. públ. Educ., Rio de Janeiro, v.12, n.45, p. 959-980, out./dez. 2004.
- FERNANDES, E. M. G. P.; Apostila de Estatística Aplicada, Universidade do Minho, Braga, 1999.
- PONTES, B. R.; Avaliação de desempenho: nova abordagem. 8ª ed. São Paulo: LTR, 2002.
- SANTO, E. E., SANTOS, F. M. G.; Avaliação de Desempenho Docente: um Estudo de Caso Numa Instituição de Ensino Superior Privado em Salvador – Bahia, Brasil, X Coloquio Internacional Sobre Gestion Universitária en America del Sur, Mar del Plata, 8-10 de Diciembre de 2010.