

Auto-Avaliação de IES - Base estatística para Índices de Desempenho

Maria Elisa Ehrhardt Carbonari
Doutora em Educação - UNICAMP
Diretora de Pesquisa, Extensão e Pós-Graduação da Anhanguera Educacional S.A.
e-mail: elisa@unianhanguera.edu.br

Braz Bello Junior*
Doutor em Química (Físico-Química) - USP
Professor da Faculdade Comunitária de Campinas - Unidade 3
e-mail: braz.bello@unianhanguera.edu.br

Luiz Renato Ribeiro Ferreira
Mestre em Ciência Política - UNICAMP
Assessor acadêmico - Anhanguera Educacional S/A no Departamento de Avaliação Institucional
e-mail: luiz.renato@unianhanguera.edu.br

Claudette Maria Medeiros Vendramini
Doutora em Educação - UNICAMP
Professora da Universidade São Francisco
e-mail: cvendramini@uol.com.br

*Bolsista FUNADESP

Resumo

Neste trabalho é apresentada a fundamentação teórica dos cálculos estatísticos que são utilizados no Programa de Avaliação Institucional - PAI - aplicado nas unidades mantidas pela Anhanguera Educacional. São apresentados a forma como são escolhidas as amostras de alunos que participam do PAI, e como são montados os índices que manifestam a satisfação dos alunos.

Palavras-chave: Auto-Avaliação Institucional, Instituições de Ensino Superior, Estatística e Índices de Desempenho.

Introdução

A Anhanguera Educacional S.A., empresa mantenedora de 25 (vinte e cinco) faculdades, 2 (dois) centros universitários e 1 (uma) universidade, distribuídos nos Estados de São Paulo, Rio Grande do Sul, Goiás e Mato Grosso do Sul, possui um programa de auto-avaliação institucional instituído desde 1997.

Esta ação, intitulada Programa de Avaliação Institucional - PAI, representa a principal atividade do

Abstract

In this work we show the Statistical concepts used on Institutional Evaluation Program - PAI - applied in colleges of the Anhanguera Educacional. We show the proceedings to obtain the samples of students for apply the PAI and how is built the performance index that record satisfaction of the students.

Key-words: Institutional Self-Evaluation, Colleges, Statistical and Performance Index.

Departamento de Avaliação Institucional - DAI, que realiza os procedimentos para os cursos de graduação, pós-graduação e extensão, avaliando, também, os formandos, egressos e o programa de capacitação docente, além de demandas pontuais provenientes de projetos específicos das áreas acadêmica e administrativa.

A criação do PAI, em 1997, contribuiu para que a AESA antecipasse boa parte dos preceitos contidos no Sistema Nacional de Avaliação do Ensino Superior,

o SINAES¹, possibilitando de um lado uma melhor adaptação à nova legislação e, por outro lado, um considerável avanço em relação ao trabalho que já vinha sendo desenvolvido.

Em função de questionamentos internos e externos sobre a exatidão dos procedimentos estatísticos utilizados no Programa de Avaliação Institucional, o Departamento de Avaliação Institucional constituiu um grupo de trabalho composto por educadores, especialistas e técnicos internos e externos com o objetivo de discutir e fundamentar cientificamente em relatório técnico-científico os procedimentos e técnicas estatísticas utilizadas para produzir os índices de satisfação dos discentes, nominados de ISD - Índice de Satisfação Discente e PNA - Percentual de Notas Altas.

Procedimentos para análise

Foram objeto do presente estudo os seguintes instrumentos utilizados para na auto-avaliação institucional:

- Avaliação de Desempenho Docente pelos alunos (ADDA);
- Avaliação dos Coordenadores de Cursos pelos Alunos (ACCA);
- Avaliação dos Coordenadores de Curso pelos professores (ACCP);

Normalmente, os processos de avaliação envolvem levantamentos de dados obtidos a partir de amostras, como é o caso PAI. Dessa maneira, os procedimentos para determinação das amostras são igualmente fundamentais para a obtenção dos índices de desempenho.

Igualmente importante é o desempenho dos instrumentos utilizados para a obtenção dos dados. Instituições como a *American Educational Research Association* (AERA), *American Psychological Association* (APA) e *National Council on Measurement in Education* (NCME) defendem que instrumentos de avaliação psicológica e educacional de qualidade são aqueles que possuem evidências empíricas e teórico-conceituais que fundamentem a validade das inferências feitas em situações específicas de determinados contextos.

Para que os instrumentos utilizados na coleta de dados sejam válidos para medir ao que se propõem, é fundamental que os mesmos sejam claros, bem definidos e possuam qualidades de um bom instrumento de medida. Ao se construir um instrumento de medida, para um levantamento de dados por amostragem, deve-se:

1. Explicitar claramente o objetivo da pesquisa;

2. Explicitar claramente o objetivo dos instrumentos de medida.

3. Conceituar “o quê” se pretende medir.

4. Elaborar itens em função do conceito definido.

5. Explicitar a população alvo.

6. Escolher os procedimentos adequados de amostragem aleatória.

No que se refere ao objetivo da pesquisa, foram realizadas análises qualitativas das questões que compunham os questionários de avaliação interna da instituição, adequando-as para uma avaliação que fosse, ao mesmo tempo, simples de ser realizada e refletisse todos os aspectos importantes da vida acadêmica do discente, conceituando “o quê” se pretende medir.

No que se refere ao objetivo dos instrumentos de medida, foram analisados os procedimentos estatísticos utilizados no processo de amostragem, coleta e interpretação dos dados obtidos, buscando dar condições para a consolidação das alterações propostas para o processo de avaliação institucional.

Com relação à avaliação do docente pelos estudantes - ADDa, os resultados são tratados estatisticamente e publicados na forma de índices de percentuais de desempenho do docente e que são chamados de *Índice de Satisfação Discente (ISD)* e *Percentual de Notas Altas (PNA)*.

Ao avaliar os docentes, os discentes aplicam notas que estão entre 1 (um) e 5 (cinco). A nota 1 corresponde a muito fraco, a nota 2 corresponde a fraco, a nota 3 corresponde a regular, a nota 4 corresponde a bom e a nota 5 corresponde a muito bom, qualificações semelhantes ao utilizado nos instrumentos de avaliação aplicados pelo MEC-INEP.

Dessas notas se extrai uma média, e essas médias são convertidas em percentuais estatísticos de forma a produzir o ISD e o PNA.

O ISD registra a porcentagem de notas iguais ou maiores do que 3 (três), ou seja, aquelas em que o discente considerou o docente regular, bom ou muito bom. ISD acima de 50% são aceitáveis e acima de 70% são desejáveis.

Já o PNA corresponde à porcentagem de notas iguais ou maiores do que 4 (quatro), ou seja, aquelas em que o discente considerou o docente bom ou muito bom. Não existem níveis de aceitação para o PNA porque o índice objetiva verificar o quão concentradas ou não estão as notas atribuídas com base no valor do desvio padrão obtidos dos dados.

Vale ressaltar que o desempenho do docente é inferido pela percepção que o aluno tem do seu desempenho, que pode sofrer influências de variáveis

subjetivas que são difíceis de serem controladas, tais como antipatia do aluno para com o professor, baixa habilidade necessária para desempenhar as atividades da disciplina com sucesso, pouco tempo para estudar, cansaço físico, atitudes negativas em relação à disciplina, entre outras.

Por esse motivo, é importante estimar um indicador de qualidade docente que mais se aproxime da realidade e que possa dar maior segurança para a tomada de decisões dos administradores, ou seja, outro objetivo do PAI é o de utilizar os resultados como ferramenta de gestão acadêmica.

Portanto, a construção de um índice confiável a ser utilizado como referência para todas as decisões pertinentes à gestão acadêmica e administrativa de qualquer instituição de ensino superior é fundamental para o sucesso da investigação científica.

Nessa linha, o ISD, principalmente, representa a tradução estatística das percepções dos alunos e dos professores sobre uma infinidade de variáveis, as quais, em seu conjunto, oferecem uma resposta fidedigna às indagações traduzidas nos instrumentos de avaliação. Em outros termos, o ISD é a quantificação da “fala” qualitativa dos entrevistados ou o desempenho percebido pelo aluno.

Dessa maneira, os três passos iniciais de construção de um instrumento de medida por levantamento amostral citados neste documento já foram atendidos:

1. o objetivo da pesquisa é fornecer dados fidedignos sobre o desempenho dos docentes e coordenadores de curso da instituição para subsidiar a tomada de decisão de seus gestores;
2. cada instrumento de medida está claramente explicitado como sendo para avaliar o desempenho docente pelos alunos e os coordenadores de cursos pelos alunos e pelos professores.
3. ao avaliar o docente ou o coordenador de curso, o que se pretende medir é a “qualidade” de cada um deles.

Sobre as percepções que a Anhanguera Educacional considera como características de um docente de qualidade aquele que:

1. domina o conteúdo que leciona e o expõe com dinamismo, clareza, organização e seqüência lógica os temas (didática);
2. relaciona os aspectos teóricos e práticos;
3. relaciona a sua disciplina com outras do curso;
4. tem postura ética;
5. é motivado, pontual, incentiva os alunos para a

participação de atividades extra-classe; tem interesse em esclarecer as dúvidas dos alunos; e

6. contribui para o projeto de vida dos alunos.

Além destes aspectos, um bom docente deve apresentar o plano de ensino ao estudante, cumpri-lo no prazo previsto, utilizar em suas aulas recursos adicionais à aula expositiva, e incentivar a leitura de livros complementares e de revistas publicadas pela instituição, além do livro texto. Para avaliar os resultados da aprendizagem o bom docente deve utilizar critérios de avaliação previstos no início do curso, com questões redigidas clara e adequadamente ao conteúdo ministrado.

As percepções positivas sobre o trabalho de um coordenador são aquelas que refletem se ele:

1. tem postura ética;
2. tem disponibilidade para atender alunos e professores;
3. busca resolver as questões encaminhadas pelos alunos e as questões emergenciais do seu curso;
4. está presente, com regularidade, na sala de aula e em reuniões com representantes de classe;
5. tem controle emocional nas situações de conflito ocorridas no cotidiano escolar; resolve as necessidades didático-pedagógicas do curso;
6. incentiva os alunos a participar de atividades de Extensão e de Pesquisa de Iniciação Científica; e
7. cumpre os horários disponibilizados para a coordenação do curso.

Ao mesmo tempo, um coordenador de qualidade deve: incorporar ao curso as inovações da área profissional; atuar na busca pelos padrões de qualidade exigidos pela Instituição e pelo MEC/INEP; discutir o projeto pedagógico do curso em reuniões com os professores; organizar atividades extra-classe de pesquisa, extensão, estudos em grupo, visitas técnica, entre outras; incentivar os professores à produção científica e à formação continuada; buscar o cumprimento das ações de melhoria identificadas na Avaliação Institucional; atuar nas situações de conflito do cotidiano escolar; incentivar o desenvolvimento de projetos e parcerias para a melhoria do curso.

Com o objetivo de avaliar os docentes e coordenadores da instituição foram elaborados itens em função do conceito de “qualidade” esperada para um docente ou coordenador de curso. O desempenho do docente e do coordenador são ambos inferidos através do ISD.

Como já salientado neste documento, é importante pensar no desempenho do processo de avaliação como um todo, desde a idéia inicial até a avaliação e discussão

de seus resultados. Nesse processo, é importante explicitar claramente quais são os participantes (alunos, professores, gestores, entre outros) da avaliação, se a população toda, que deve ser claramente definida, ou se uma parte da população alvo. Em caso de amostras, o processo de amostragem deve ser aleatório e representativo da população, para garantir inferências confiáveis para a população alvo, a partir de estimativas que apresentem erros de medida tão pequenos quanto possíveis. A vantagem de se utilizar procedimentos estatísticos adequados para medir as variáveis de interesse é que o erro associado ao processo de medida pode ser conhecido.

Neste sentido, são apresentados a seguir alguns dos procedimentos do Programa de Avaliação Institucional, disponíveis no site institucional (<http://www.unianhanguera.edu.br/avaliacao/avaliacao-procedimentos.php>) e alguns conceitos estatísticos que fundamentam as inferências que são feitas a partir dos índices calculados no processo de avaliação.

O Programa de Avaliação Institucional

Um dos instrumentos mais importantes de gestão educacional das unidades que formam o grupo Anhanguera Educacional é o Programa de Avaliação Institucional - PAI, com o objetivo de avaliar todas as dimensões que compreendem o ensino, pesquisa, extensão e gestão acadêmica/administrativa. Nesse programa são avaliadas as atividades acadêmicas, a gestão e a infra-estrutura da graduação e da pós-graduação. As avaliações são feitas no primeiro (maio) e segundo semestre (outubro) de cada ano letivo, e envolvem os egressos, programas institucionais e serviços terceirizados em todas as unidades de ensino da instituição.

O Departamento de Avaliação Institucional, em conjunto com as Comissões Próprias de Avaliação (CPA) de cada instituição coleta a opinião dos discentes por meio de um conjunto de instrumentos aplicados via sistema *online*, de maneira voluntária e sigilosa. Os resultados são disponibilizados à comissão de avaliação, professores, coordenadores, diretores, gestores e toda a comunidade acadêmica, no sentido de promover planos de melhoria de ensino, e mudanças nas atitudes, valores e comportamentos na instituição. O processo de avaliação docente (ISD) é por meio de amostragem aleatória de no mínimo 30 alunos por turma, ou por todos os alunos na ocorrência de turmas pequenas, com menos de 30 alunos. Para que uma amostra seja considerada aleatória todos os estudantes devem ter probabilidade

conhecida e diferente de zero de ser escolhido. Procedimentos como convite, participação voluntária, escolha por conveniência ou intencional, não são procedimentos de amostragem aleatória.

No primeiro semestre de 2006 foi avaliado um total de 98 cursos, distribuídos em 11 unidades, e envolvendo 81 coordenadores, 799 professores e 20.621 alunos.

Estatística - Fundamentação teórica

Pode-se supor que a pontuação que os alunos atribuem aos docentes segue uma distribuição desconhecida com média μ e desvio padrão σ , ou que ela possui uma distribuição normal.

Geralmente tanto a média populacional μ como o desvio padrão populacional σ , denominados de parâmetros populacionais porque descrevem as características da população, são desconhecidos e, portanto, devem ser estimados a partir de dados obtidos em amostras da população. Um estimador é uma característica da amostra, ou seja, é uma função dos valores da amostra tomado como verdadeiro para a população.

Se fossem realizadas várias amostras da mesma população poderiam ser encontrados diferentes resultados para o desempenho do docente. Por esse motivo, busca-se um estimador que tenha pouco desvio do valor verdadeiro populacional, já que a média amostral é uma variável aleatória em que a maioria dos valores se concentra em torno da média populacional, com os valores distantes da média populacional ocorrendo com menor probabilidade.

A diminuição da variabilidade da distribuição da média amostral (\bar{X}) não acontece por acaso, pois

$$\bar{X} = \sum_{i=1}^n \frac{X_i}{n} \Rightarrow E(X) = \mu \Rightarrow V(\bar{X}) = \frac{\sigma^2}{n}$$

Isso implica que à medida que o tamanho da amostra tende ao tamanho da população a variância da média amostral $V(\bar{X})$ tende para zero. No caso extremo quando amostramos a população inteira, não existe diferença da média amostral para a populacional. O desvio padrão da média amostral é igual ao desvio padrão da população dividido pela raiz quadrada do tamanho da população, e é denominado **erro padrão da média amostral**. Esse erro pode ser calculado a partir da seguinte fórmula quando se extraem amostras de populações finitas.

$$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}} \left(\frac{N-n}{N-1} \right)$$

Quando a população segue uma distribuição normal, a média amostral segue uma distribuição normal com a mesma média e com a variância dividida pelo tamanho da amostra, onde utiliza-se o teorema do limite central para garantir que a média amostral pode substituir a média populacional.

O **Teorema do Limite Central** garante que a distribuição da média amostral tende para uma distribuição normal à medida que o tamanho da amostra n tende para infinito, ou seja, quando o tamanho da amostra é suficientemente grande ($n \geq 30$). Com amostras maiores ou iguais a trinta a distribuição das médias amostrais tem uma boa aproximação da curva da distribuição normal. A convergência para a normalidade é mais rápida quando a distribuição dos dados é simétrica; quando a distribuição é muito assimétrica ou bimodal, a convergência é mais lenta.

Geralmente a variância populacional não é conhecida e é necessário estimá-la a partir da amostra. Neste caso utiliza-se outra distribuição de probabilidade denominada *t de Student*, que é simétrica em torno de zero e parecida com a distribuição normal, com a diferença que as caudas são mais pesadas, o que a torna mais rigorosa nos testes de hipóteses, quando a variância deve ser estimada a partir de dados amostrais. No entanto, para o tamanho da amostra maior ou igual a 30 pode-se usar a distribuição normal, que é bem próxima da *t de Student* nestes casos.

A figura 1 mostra o Teorema do Limite Central em ação para uma *distribuição exponencial*, com maioria de valores próximos a zero. A média dessa distribuição é 1 e seu desvio padrão é 1. As figuras (b), (c) e (d) são as curvas de densidade das médias amostrais com $n = 2$, $n = 10$ e $n = 25$. À medida que n aumenta, a distribuição torna-se mais próxima à curva normal. A média permanece 1 e o desvio padrão diminui, tomando o valor $1/\sqrt{n}$. A curva de densidade para $n=10$ ainda é um tanto assimétrica à direita, mas já se assemelha a uma curva normal com média $\mu = 1$ e desvio padrão finito $s = 1/\sqrt{10} = 0,32$. Para $n = 25$, a curva aproxima-se ainda mais da distribuição normal, com média $\mu = 1$ e desvio padrão finito $\sigma = 1/\sqrt{25} = 0,20$.

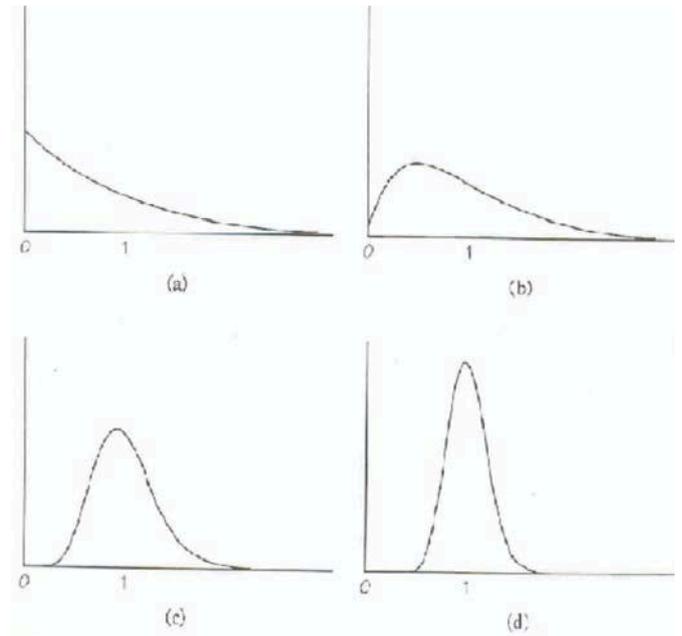


Figura 1 - Função densidade de probabilidade de uma exponencial com $\mu = 1$ e $\sigma = 1$ (a) e da distribuição das médias amostrais para $n = 2$ (b), $n = 3$ (c) e $n = 25$ (d)

Descrição do Plano de Amostragem

O procedimento de amostragem adotado foi o de amostragem aleatória estratificada com seleção de amostras aleatórias simples de cada um dos estratos.

Considerando-se a avaliação docente pelos estudantes, a população de N estudantes que pode ser dividida em subpopulações (denominadas estratos) de $N_1, N_2, N_3, N_4, \dots, N_L$ estudantes, sendo que essas subpopulações não se superpõem e, juntas, representam a totalidade da população. Esses estratos correspondem às turmas que o docente avaliado leciona, de tal modo que a soma das turmas (estratos) corresponda ao total de estudantes das turmas que o docente leciona:

$$N_1 + N_2 + N_3 + N_4 + \dots + N_L = N$$

Portanto, são conhecidos todos os valores de N_h .

Para a estimação do desempenho médio do docente, são extraídas amostras aleatórias simples de tamanho fixo ($n = 30$) das k populações com mais de 30 estudantes e a população total das $(L - k)$ turmas com menos de 30 estudantes, portanto:

$$n_1 = n_2 = n_3 = n_4 = \dots = n_k = 30 \quad \text{e} \quad n_{k+1} = N_{k+1}, \\ n_{k+2} = N_{k+2}, \dots, n_L = N_L$$

A técnica de amostragem aleatória estratificada foi escolhida por desejar dados de determinada precisão sobre cada turma, com informações específicas de cada

turma que pode apresentar características mais homogêneas entre os estudantes da turma e mais heterogêneas entre as turmas. Além disso, a amostragem estratificada aumenta a precisão das estimativas que se deseja fazer.

Será adotada a seguinte notação para a apresentação das fórmulas:

N_h número total de estudantes da turma h
 n_h número da amostra de estudantes da turma h
 $W_h = \frac{N_h}{N}$ Peso (proporção) do estrato (turma) h
 $f_h = \frac{n_h}{N_h}$ fração amostral na turma h
 $\bar{Y}_h = \frac{\sum_{i=1}^{N_h} y_{hi}}{N_h}$ pontuação média verdadeira atribuída ao docente pela turma h
 $\bar{y}_h = \frac{\sum_{i=1}^{n_h} y_{hi}}{n_h}$ pontuação média da amostra atribuída ao docente pela turma h
 $S_h^2 = \frac{\sum_{i=1}^{N_h} (y_{hi} - \bar{Y}_h)^2}{N_h - 1}$ variância verdadeira da pontuação atribuída ao docente pela turma h

Para descrever as propriedades das estimativas que serão feitas da pontuação média atribuída aos docentes pelos estudantes a partir da técnica de amostragem aleatória estratificada deve-se observar que a estimativa da pontuação média verdadeira, de um modo geral, não é igual à pontuação média amostral, considerando que:

$\hat{Y} = \frac{\sum_{h=1}^L N_h \cdot \bar{y}_h}{N}$ estimativa da pontuação média verdadeira atribuída ao docente
 $\bar{y} = \frac{\sum_{h=1}^L n_h \cdot \bar{y}_h}{n}$ pontuação média amostral atribuída ao docente

$V\left(\hat{Y}\right) = \frac{1}{N^2} \sum_{h=1}^L N_h (N_h - n_h) \cdot \frac{S_h^2}{n_h}$ variância da estimativa da pontuação média verdadeira atribuída ao docente

$s_h^2 = \frac{1}{n_h - 1} \sum_{i=1}^{n_h} (y_{hi} - \bar{y}_h)^2$ estimativa da variância verdadeira da pontuação atribuída ao docente pela turma h

A diferença é que as estimativas \hat{Y} da pontuação média das turmas individuais recebem pesos N_h/N corretos, mas no caso deste estudo estes pesos não coincidem com as frações amostrais n_h/n , dado que são retiradas amostras $n = 30$ para todas as turmas. Neste caso específico, a amostragem aleatória estratificada não é proporcional.

Algumas afirmações podem ser feitas a partir de teoremas demonstrados em livros de amostragem como Cochran (1965), Bolfarine e Bussab (2005), tais como para uma amostragem aleatória estratificada, a variância da estimativa da pontuação média verdadeira atribuída ao docente é dada pela expressão de $V\left(\hat{Y}\right)$. É possível tornar mínima essa variância, dentro de um limite de custo pré-estabelecido para a seleção das amostras de cada turma, ou para tornar mínimo o custo referente a um valor esperado de $V\left(\hat{Y}\right)$, mas para que isso ocorra deve-se promover estudos que estabeleçam a função custo que se deseja minimizar.

Como as amostras são extraídas aleatoriamente de cada um dos estratos tem-se a estimativa s_h^2 sem tendência para S_h^2 , sendo que a variância estimada de \hat{Y} é dada por:

$$s^2\left(\hat{Y}\right) = \sum_{h=1}^L \frac{W_h^2 \cdot s_h^2}{n_h} - \sum_{h=1}^L \frac{W_h \cdot s_h^2}{N}$$

Portanto têm-se as seguintes estimativas:

Estimativa da pontuação média de um item:	\hat{Y}	com erro amostral igual $s\left(\hat{Y}\right)$
---	-----------	---

Exemplo de aplicação

Considerando que o questionário de avaliação docente pelos estudantes contém 15 itens que indicam o desempenho dos docentes a partir do ISD. Cada docente

pode atingir uma pontuação total numa escala de 15 a 75 pontos, considerando que cada característica do professor é avaliada pelo estudante atribuindo-se uma pontuação numa escala de 1 a 5. Os critérios para avaliar o desempenho docente podem ser estabelecidos a partir da divisão da escala em quatro partes iguais, como nas figuras 2 e 3.

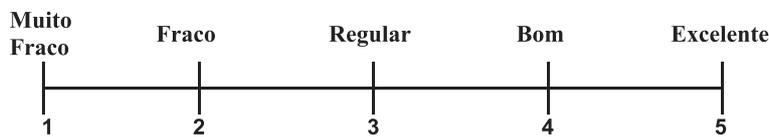


Figura 2 - Escala de avaliação docente para um item

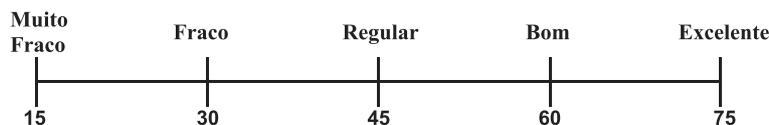


Figura 3 - Escala de avaliação docente total

Pode-se considerar um docente com qualidade *inaceitável* em um item de avaliação aquele que obtém pontuação média inferior a 3 pontos, um docente com qualidade *aceitável* no item aquele com pontuação média no item superior ou igual a 3 pontos.

Da mesma maneira pode-se considerar um docente com qualidade *inaceitável* aquele com pontuação total média inferior a 45 pontos, um docente com qualidade *aceitável* aquele com pontuação total média superior ou igual a 45 pontos.

Considerando, como exemplo, que um professor P leciona duas disciplinas (d_1 e d_2) em duas unidades de ensino (u_1 e u_2), em dois cursos (c_1 e c_2) de mesma série e uma turma cada, totalizando 8 turmas, com um total de alunos igual a:

$N_1 = 100$ estudantes $N_2 = 80$ estudantes $N_3 = 60$ estudantes
 $N_4 = 52$ estudantes $N_5 = 40$ estudantes $N_6 = 30$ estudantes
 $N_7 = 22$ estudantes $N_8 = 12$ estudantes

Totalizando $N_1 + N_2 + N_3 + N_4 + N_5 + N_6 + N_7 + N_8 = N = 396$ estudantes

Os pesos de cada estrato são dados por:

$$W_1 = \frac{100}{396} = 0,2525 \quad W_2 = \frac{80}{396} = 0,2020 \quad W_3 = \frac{60}{396} = 0,1515$$

$$W_4 = \frac{52}{396} = 0,1313 \quad W_5 = \frac{40}{396} = 0,1010 \quad W_6 = \frac{30}{396} = 0,0758$$

$$W_7 = \frac{22}{396} = 0,0556 \quad W_8 = \frac{12}{396} = 0,0303$$

Para a estimação do desempenho média do docente P são extraídas amostras aleatórias simples de tamanho fixo ($n = 30$) das 5 populações com mais de 30 estudantes e a quantidade total de estudantes das outras 3 turmas com 30 ou menos estudantes, portanto:

$$k = 5 \text{ turmas} \quad L - k = 8 - 5 = 3 \text{ turmas}$$

$$n_1 = n_2 = n_3 = n_4 = n_5 = 30 \quad n_6 = N_6 = 30$$

$$n_7 = N_7 = 22 \quad n_8 = N_8 = 12$$

totalizando uma amostra $n = 214$ estudantes, correspondente a 54% da população.

A pontuação média atribuída ao professor P pelos 30 estudantes selecionados aleatoriamente da primeira turma ($h = 1$) é dada por, considerando a variável pontuação média de cada estudante:

$$\bar{y}_1 = \frac{\sum_{i=1}^{n_1} y_{1i}}{n_1} = \frac{\sum_{i=1}^{30} y_{1i}}{30} = 3,92$$

Da mesma maneira, para as outras turmas a média será:

$$\bar{y}_2 = 3,84, \quad \bar{y}_3 = 3,51, \quad \bar{y}_4 = 4,52, \quad \bar{y}_5 = 4,24,$$

$$\bar{y}_6 = 3,28, \quad \bar{y}_7 = 3,60, \quad \bar{y}_8 = 3,54$$

A estimativa da variância verdadeira da pontuação atribuída ao professor P pela primeira turma é igual a:

$$s_1^2 = \frac{1}{n_1 - 1} \sum_{i=1}^{n_1} (y_{1i} - \bar{y}_1)^2 = \frac{1}{29} \sum_{i=1}^{30} (y_{1i} - 3,92)^2 = 0,9635^2$$

e a estimativa do desvio-padrão é igual a $s_1 = 0,9635$.

A partir do mesmo procedimento podem ser calculados os valores dos outros desvios padrões:

$$s_1 = 0,9635 \quad s_2 = 0,9635 \quad s_3 = 0,9635 \quad s_4 = 0,9635$$

$$s_5 = 0,9635 \quad s_6 = 0,9635 \quad s_7 = 0,9635 \quad s_8 = 0,9635$$

A estimativa da pontuação média verdadeira atribuída pelo estudante ao docente P é igual a:

$$\hat{Y} = \frac{\sum_{h=1}^L N_h \cdot \bar{y}_h}{N} = \frac{100 \cdot 3,92 + 80 \cdot 3,84 + 60 \cdot 3,51 + 52 \cdot 4,52 + 40 \cdot 4,24 + 30 \cdot 3,28 + 22 \cdot 3,60 + 12 \cdot 3,54}{396}$$

$$\hat{Y} = 3,45$$

A variância e o desvio-padrão da estimativa são dados por:

$$s^2\left(\frac{\Delta}{Y}\right) = \sum_{h=1}^8 \frac{W_h^2 \cdot s_h^2}{n_h} - \frac{\sum_{h=1}^8 W_h \cdot s_h^2}{N} = 0,005175 - 0,002344 = 0,00283 \Rightarrow s\left(\frac{\Delta}{Y}\right) = 0,05320$$

Portanto, pode-se afirmar com 95% de certeza que a pontuação média atribuída pelos estudantes ao professor P está entre $(3,45 \pm 1,96 \cdot 0,0532)$ ou seja, entre 3,35 e 3,55. Ou ainda que a pontuação média dada ao professor é igual a 3,45 com um erro igual a 0,10 para mais ou para menos.

Considerações finais

Todo processo de auto-avaliação institucional relaciona-se com inúmeras variáveis, muitas vezes de difícil controle, como, por exemplo, a motivação, as atitudes, organização e resistências que porventura se apresentam à frente do avaliador. Dentre as notórias dificuldades e desafios encontrados estão o processo de análise dos dados e a administração das reações da comunidade acadêmica em relação aos resultados obtidos após a aplicação dos instrumentos.

Para minimizar o impacto das resistências - um predicado dos processos de avaliação - é necessário tratar as amostras e cálculos de composição estatísticos com o maior rigor científico possível. Esta característica, aliada ao tratamento dos dados com transparência e participação da comunidade, representam o caráter científico e político que toda auto-avaliação deve supor.

Este artigo procurou apresentar as estratégias utilizadas pela Anhanguera Educacional no que se refere ao necessário rigor científico da auto-avaliação. Tal modelo pode ser aplicado em diferentes instituições de ensino superior e seu sucesso dependerá, em um outro momento, da capacidade de articulação e composição política, aliadas à precisão dos dados obtidos.

Uma das qualidades essenciais deste programa está justamente em realizar esta articulação com a comunidade acadêmica, fazendo com que os índices obtidos possam ser discutidos entre alunos, professores, coordenadores e funcionários, confrontando os percentuais estritamente matemáticos com o clima organizacional, as expectativas e sentimentos dos atores, aliando, portanto, variáveis exatas com variáveis subjetivas.

A ação de estabelecer tais tipos de comparações resulta, na absoluta maioria das vezes, em uma reiteração dos resultados obtidos, comprovando sua precisão e, simultaneamente, diminuindo a resistência por parte de

alguns setores da comunidade acadêmica.

Com a periodicidade dos processos e a reprodução permanente das estratégias científicas e políticas estabelece-se ao longo do tempo a cultura da auto-avaliação, elemento fundamental para o sucesso do Programa de Avaliação Institucional da Anhanguera Educacional e, evidentemente, de quaisquer outras IES que de fato pretendam utilizar seus resultados como ferramenta de gestão estratégica.

Referências Bibliográficas

- BOLFARINE, H., BUSSAB, W. O. (2005). *Elementos de amostragem*. São Paulo: Edgard Blücher
- BRASIL. Lei nº 10.861, de 14 de abril de 2004. Institui o Sistema Nacional de Avaliação da Educação Superior - SINAES e dá outras providências. Diário Oficial da União, Brasília, 15 abril 2004. p. 3, Col.1.
- COCHRAN, W. G. (1965). *Técnicas de amostragem*. Rio de Janeiro: Editora Fundo de Cultura.
- CARBONARI, Maria Elisa E. Avaliação institucional: espaço de controle, gestão ou planejamento emancipatório? Marcos históricos e fundamentos filosóficos. *Revista de Educação*, Pirassununga, v.5, n.5, out. 2002.
- CARBONARI, Maria Elisa E. Programa de Avaliação Institucional: possibilidades, desafios e perspectivas. *Revista da Rede de Avaliação Institucional da Educação Superior - RAIES*, Ano11, vol.11 n.1, março 2006.
- CENTRO UNIVERSITÁRIO ANHANGUERA. Programa de Avaliação Institucional. Departamento de Avaliação Institucional. Leme. 1996.
- DIAS SOBRINHO, José; RISTOFF, Dilvo. Universidade desconstruída: Avaliação Institucional e Resistência. Florianópolis: Ed. Insular, 2000.
- SINAES - Sistema Nacional de Avaliação da Educação Superior. Bases para uma nova proposta de avaliação da educação superior. Brasília: INEP, 2003.

Notas

- ¹ Criado pela Lei nº10.861, de 14 de abril de 2004, o SINAES, segundo o MEC, é “formado por três componentes principais: a avaliação das instituições, dos cursos e do desempenho dos estudantes. O SINAES avalia todos os aspectos que giram em torno desses três eixos: o ensino, a pesquisa, a extensão, a responsabilidade social, o desempenho dos alunos, a gestão da instituição, o corpo docente, as instalações e vários outros aspectos”. Do ponto de vista do interesse deste artigo, trataremos especificamente da “avaliação institucional”.

Recebido em 03 de julho de 2007 e aprovado em 24 de agosto de 2007.