



Redes. Revista do Desenvolvimento Regional

ISSN: 1414-7106

ISSN: 1982-6745

revistaredes@unisc.br

Universidade de Santa Cruz do Sul

Brasil

do Carmo, Roberto Luiz; Shimoda, Eduardo  
Análise da capacidade de discriminação de critérios da avaliação da Pós-  
Graduação na Área de Planejamento Urbano e Regional e Demografia da Capes  
Redes. Revista do Desenvolvimento Regional, vol. 23, núm. 2, 2018, -, pp. 358-385  
Universidade de Santa Cruz do Sul  
Brasil

DOL: <https://doi.org/10.17058/redes.v23i2.11962>

Disponível em: <http://www.redalyc.org/articulo.oa?id=552059721018>

- Como citar este artigo
- Número completo
- Mais informações do artigo
- Site da revista em [redalyc.org](http://redalyc.org)

UNESC  
[redalyc.org](http://redalyc.org)

Sistema de Informação Científica Redalyc  
Rede de Revistas Científicas da América Latina e do Caribe, Espanha e Portugal  
Sem fins lucrativos acadêmica projeto, desenvolvido no âmbito da iniciativa  
acesso aberto



# **Análise da capacidade de discriminação de critérios da avaliação da Pós-Graduação na Área de Planejamento Urbano e Regional e Demografia da Capes**

**Roberto Luiz do Carmo**

*Universidade Estadual de Campinas*

**Eduardo Shimoda**

*Universidade Estadual de Campinas*

*Universidade Candido Mendes*

## **Resumo**

Na avaliação dos programas de pós-graduação pela Capes, em função dos critérios adotados para enquadramento dos programas nos cinco estratos de nota, os itens demonstram diferentes capacidades de discriminar os programas. O presente artigo tem como objetivos propor uma metodologia para avaliar a capacidade de discriminação dos itens e identificar aqueles cujos critérios poderiam ser rediscutidos para tentar melhorar a capacidade de discriminar os programas. Para tal, foram analisadas as frequências de programas nos estratos em cada um dos itens, tomando por base as informações presentes no documento de área de Planejamento Urbano e Regional e Demografia do triênio 2010-2012. Para avaliação da capacidade de discriminação de cada item, foi proposta a utilização dos seguintes parâmetros estatísticos: média, amplitude, desvio-padrão, grau de assimetria e desvio médio em relação à distribuição uniforme. Os itens, com base nesses parâmetros estatísticos, foram classificados como tendo nenhuma, baixa, média e alta capacidade de discriminação. Ressalta-se que os resultados poderiam servir como subsídio preliminar para discussão, no âmbito da área, para possível implementação de critérios objetivos nos itens.

**Palavras-chave:** Avaliação da Capes. Área de Planejamento Urbano e Regional e Demografia. Avaliação da Produtividade em Educação.

## **Analysis of discrimination capacity of Postgraduate evaluation criteria in the area of Urban and Regional Planning and Demography of Capes**

## **Abstract**

In the evaluation of the postgraduate programs by Capes, according to the criteria adopted for framing the programs in the five note strata, the items demonstrate different capacities to discriminate the programs. The purpose of this article is to propose a methodology to evaluate the discrimination capacity of the items and to identify the items whose criteria could be rediscussed to try to improve the capacity to discriminate the programs. For that, we analyzed the frequency of programs in the strata in each one of the items, based on the

information presented in the area of Urban and Regional Planning and Demography for the triennium 2010-2012. To evaluate the discrimination capacity of each item, it was proposed to use the following statistical parameters: mean, amplitude, standard deviation, degree of asymmetry and mean deviation in relation to the uniform distribution. The items, based on these statistical parameters, were classified as having no, low, medium and high discrimination capacity. It should be emphasized that the results could serve as a preliminary subsidy for discussion within the area for possible implementation of objective criteria in the items.

**Keywords:** Capes evaluation. Urban and Regional Planning area and Demography. Evaluation of Productivity in Education.

### **Análisis de la capacidad de discriminación de criterios de la evaluación de la Post-Graduación en el área de Planeamiento Urbano y Regional y Demografía de Capes**

#### **Resumen**

En la evaluación de los programas de postgrado por la Capes, en función de los criterios adoptados para encuadrar los programas en los cinco estratos de nota, los ítems demuestran diferentes capacidades de discriminar los programas. El presente artículo tiene como objetivos proponer una metodología para evaluar la capacidad de discriminación de los ítems e identificar los ítems cuyos criterios podrían ser rediscutidos para intentar mejorar la capacidad de discriminar los programas. Para ello, se analizaron las frecuencias de programas en los estratos en cada uno de los ítems, tomando como base las informaciones presentes en el documento de área de Planificación Urbana y Regional y Demografía del trienio 2010-2012. Para evaluar la capacidad de discriminación de cada ítem, se propuso la utilización de los siguientes parámetros estadísticos: media, amplitud, desviación estándar, grado de asimetría y desviación media en relación a la distribución uniforme. Los ítems, sobre la base de estos parámetros estadísticos, fueron clasificados como teniendo, baja, media y alta capacidad de discriminación. Se resalta que los resultados podrían servir como subsidio preliminar para discusión, en el ámbito del área, para la posible implementación de criterios objetivos en los ítems.

**Palabras clave:** Evaluación de la Capes. Área de Planificación Urbana y Regional y Demografía. Evaluación de la Productividad en Educación.

#### **1 Introdução**

O número de Programas de Pós-Graduação *Stricto Sensu* aumentou de maneira expressiva ao longo das últimas décadas. Em 1998, existia um total de 1.259 programas de Pós-Graduação, sendo que, no ano de 2016, esse número chegou a 4.175 programas, ou seja, um crescimento de mais de 7% ao ano no período. Essa expansão da pós-graduação no Brasil foi acompanhada da preocupação em relação à qualidade dos cursos, o que impulsionou a construção e a consolidação de um sistema de avaliação do sistema de pós-graduação. A Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Capes é o órgão federal responsável pela avaliação, que foi realizada com periodicidade trienal por vários anos, até o período 2010-2012, tendo passado a quadrienal na avaliação mais recente (2013-2016). Os programas de pós-graduação estão agrupados em 50 áreas de avaliação, uma delas é a de Planejamento Urbano e Regional e Demografia - PLURed, que é o objeto principal deste texto.

O sistema de avaliação se baseia na análise de uma série de critérios distribuídos entre cinco quesitos. Alguns destes critérios são objetivos e

quantitativos, principalmente aqueles relacionados à produção científica e ao tempo de titulação dos mestrandos/doutorandos. Outros, como a inserção social, ainda são um tanto quanto subjetivos, embora exista um esforço dos comitês das áreas no sentido de tornar todos os critérios mais objetivos.

Em cada quesito, pode ser atribuída uma nota de 1 a 5, de acordo com as diretrizes definidas pela área de avaliação. Como a análise é, em termos, comparativa, seria interessante que cada quesito tivesse claras as diretrizes de forma a se poder distinguir os programas que atendem mais eficazmente ao critério.

Por exemplo, no caso de um quesito ter programas classificados em todos os cinco estratos de nota, distribuídos de forma relativamente uniforme, pode-se afirmar que as diretrizes do critério foram eficientes na discriminação dos programas. Por outro lado, se, em um quesito, todos os programas recebem a mesma nota, em princípio, as diretrizes não foram capazes de discriminar os programas quanto à eficácia no quesito.

Ao final de cada ciclo avaliativo, é gerado um relatório contendo uma série de informações a respeito da avaliação, incluindo a percentagem de programas em cada estrato de nota em cada quesito.

Com base nessa distribuição, é possível realizar uma análise para identificar quais quesitos conseguiram discriminar melhor os programas.

Dessa forma, os objetivos do presente trabalho são:

- (i) propor uma metodologia que permita avaliara capacidade de discriminação dos quesitos e itens de avaliação da pós-graduação da Capes
- (ii) no âmbito da avaliação da área de Planejamento Urbano e Regional e Demografia da Capes, classificar os itens de acordo com a capacidade de discriminação com base na avaliação dos programas acadêmicos no triênio 2010-2012.

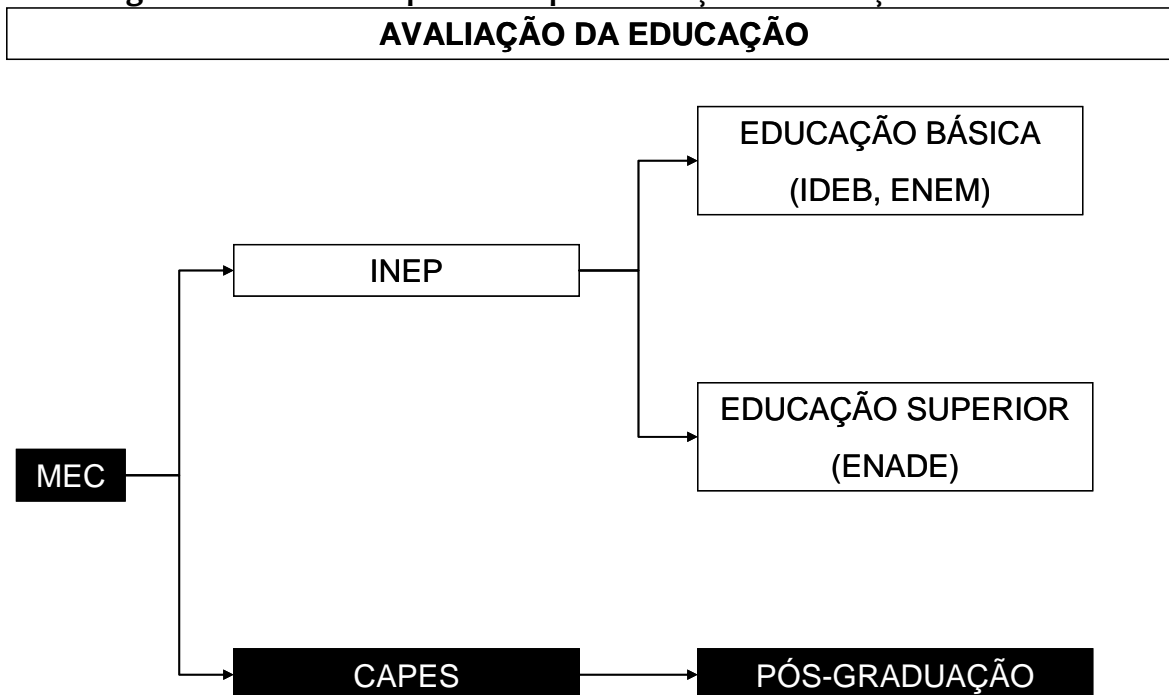
Destaca-se que os resultados iniciais deste trabalho permitiram construir um arcabouço metodológico que serviu como suporte para a avaliação realizada no período 2013-2016.

## **2 Referencial Teórico**

### **2.1 Avaliação da educação**

A avaliação da educação é realizada pelo Ministério da Educação - MEC por intermédio de duas de suas entidades: O Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira - Inep, que é responsável pela avaliação da educação básica e do ensino superior, e a Capes, que avalia a Pós-Graduação *Stricto Sensu* (Figura 1).

**Figura 1- Entidades responsáveis pela avaliação da educação no Brasil**



Quanto à Educação Básica, o Índice de Desenvolvimento da Educação Básica - Ideb visa medir a qualidade do aprendizado e estabelecer metas para a melhoria do ensino nos dois segmentos (anos iniciais e finais) do Ensino Fundamental (BRASIL, 2007). Ainda no que se refere à avaliação do último segmento da Educação Básica, existe o Exame Nacional do Ensino Médio - Enem (BRASIL, 1998). O Inep também é o responsável pela avaliação da Educação Superior, por meio do Exame Nacional de Desempenho dos Estudantes - Enade.

A pós-graduação *Stricto Sensu*, nos seus níveis de mestrado e doutorado, é avaliada pela Capes, que, mediante análise de um conjunto de critérios, atribui notas de 1 a 7 aos programas de pós-graduação.

## 2.2 Capes

A Fundação Capes é uma entidade cuja finalidade é formular e implementar a Política Nacional de Pós-Graduação. Foi criada em 1951 e, desde então, tem atuado na concessão de bolsas e no auxílio aos programas de pós-graduação. Em 1970, passou a ser um órgão autônomo e focou sua atuação nos programas de formação de pós-graduandos.

Ainda, a Capes:

é uma entidade pública vinculada ao Ministério da Educação - MEC, que tem como objetivo principal subsidiar o MEC na formulação das políticas de pós-graduação, coordenando e estimulando - mediante a concessão de bolsas de estudo, auxílios e outros mecanismos - a formação de recursos humanos altamente qualificados para a docência em grau superior, a pesquisa e o atendimento da demanda profissional dos setores públicos e privados. À CAPES compete ainda avaliar os programas de pós-graduação *stricto sensu* ministrados no país. (UFMG, 2003).

### 2.3 O Sistema de Avaliação da Capes

A Capes foi criada em 1951 com o objetivo de proporcionar condições para a formação de pessoal especializado que pudesse auxiliar ou promover o desenvolvimento nacional. A partir de 1976, foi instituído o Sistema de Avaliação dos cursos de pós-graduação. Em princípio, o objetivo desse sistema foi aumentar a qualidade da pós-graduação e classificar os programas de acordo com padrões que seriam estabelecidos (NEVES; COSTA, 2006).

A avaliação, em si, tinha um objetivo inicial relacionado à alocação de bolsas de pós-graduandos. A comunidade científica tinha o anseio, e a Capes acabou incumbida da missão de definir critérios que permitissem realizar a distribuição de bolsas para atender de forma mais justa às demandas existentes. Dessa forma, a avaliação da Capes veio, fundamentalmente, responder à necessidade de realizar a alocação de mais bolsas aos programas mais bem qualificados (CASTRO; SOARES, 1983, p. 64).

Ao final da década de 1990, a Capes passou a implementar substancial melhoria na informatização das informações e também iniciou o processo para tornar os critérios mais rigorosos e objetivos a fim de enquadrar os programas em conceitos. A avaliação foi gradativamente sendo incrementada, embora tenha certa estabilidade no que tange aos princípios fundamentais. Dentre as mudanças, muitas se referem à inclusão de novos quesitos e itens, bem como alteração da forma de mensuração destes itens e de atribuição de pesos nos quesitos e itens (MACCARI et al., 2014).

Há de se ressaltar a relativa dicotomia dos objetivos da pós-graduação no Brasil. Horta (2006), em um trabalho em que foram entrevistados coordenadores de programas de pós-graduação, identificou percepções de que existiria um dilema entre o foco na publicação dos resultados de pesquisa ou na formação dos mestrandos e doutorandos. Um dos coordenadores entrevistados por Horta (2006), da Área de Engenharias III, afirmou não ter dúvidas de que a produção científica qualificada perpassa pela produção de artigos em periódicos, mas questiona até que ponto a formação de recursos humanos (mestres e doutores) foi preterida em detrimento da publicação de artigos em periódicos científicos. Essa é uma discussão importante, que remonta à própria concepção das universidades criadas no Brasil, que são marcadas pela indissociabilidade entre ensino e pesquisa, que caracteriza o surgimento da “Universidade Moderna”, na perspectiva humboldiana, conforme explicita Castilho (2008). Essa busca por um equilíbrio na valorização da formação discente e divulgação dos resultados de pesquisa permanece nas avaliações recentes, refletindo-se no peso relativo atribuído aos quesitos e itens nos instrumentos de avaliação.

Outra vertente dos estudos realizados sobre a avaliação da pós-graduação no Brasil refere-se à natureza da avaliação. Gatti et al. (2003) fizeram um levantamento junto a colegiados de cursos de pós-graduação na Área de Educação e identificaram a perspectiva consensual nesta área de que não deveria ser imputada à avaliação um objetivo punitivo, mas, sim, diagnóstica e que pudesse conduzir a atitudes e sugestões orientadoras por parte dos programas mais bem qualificados junto aos cursos em situação menos favoráveis. Nesse sentido, a avaliação implicaria na detecção de itens críticos e no encaminhamento de



propostas efetivas a serem trabalhadas pelos programas cuja avaliação identificasse necessidade de evolução e melhorias.

Medeiros (2016) realizou uma análise da avaliação da Capes e a produção intelectual usando como recorte temporal o período de 1975 a 2002. Na dissertação, foram identificados alguns “ciclos” em que, intrinsecamente, embora com algumas alterações gradativas e evolutivas, as estratégias e a forma de avaliação vertiam aos mesmos objetivos. Entre 1974 e 1979, fundamentalmente, a avaliação tinha como objetivo precípuo a definição de cotas de bolsas para os programas, sendo o momento pioneiro no que tange à utilização da produção intelectual como critério de avaliação. No período posterior, compreendido entre os anos de 1979 e 1982, segundo a autora, foi realizado um ajuste no sistema de avaliação envolvendo a sistematização da mesma, a utilização de modelos econométricos e uma valorização, pelo menos em tese, da qualidade. De 1982 a 1989, ocorreu uma busca pela construção de métricas que balizariam a avaliação. Posteriormente, de 1990 a 1994, segundo a autora, o período foi marcado pela informatização da avaliação e por uma “saturação” do sistema de avaliação, em que havia a percepção de que a avaliação não conseguia “discriminar” adequadamente a “qualidade”. Por fim, considerando o recorte temporal utilizado na pesquisa, no período compreendido entre 1995 e 2002, foi instituído o novo modelo de avaliação da Capes, cujos moldes principais perduram até os dias atuais, em que a produção intelectual passou a ter destaque por meio da priorização de aspectos como o estabelecimento do Qualis Periódicos.

Com a evolução no sistema de avaliação, os critérios adotados passaram a servir como diretrizes para que os programas buscassem lograr um padrão de qualidade adequado, retroalimentando o sistema. As coordenações de programas passaram a utilizar essas diretrizes, bem como os relatórios de ciclos avaliativos progressos, como norteadores para adoção de planejamentos estratégicos para seus cursos (NEVES; COSTA, 2006).

O aspecto indutor da avaliação da pós-graduação é citado por Moritz, Moritz e Melo (2011), que comentam o papel decisivo desta avaliação na elevação da qualidade dos programas, o que, em última instância, promoveria o aprimoramento induzido dos cursos, além de sistematizar o desempenho e estabelecer um padrão mínimo de qualidade acadêmica destes. Corroborando esta opinião do efeito indutor da avaliação, Conto e Nunes (2017), de forma mais recente, ressaltam a relevância norteadora do Documento de Área. Segundo esses autores, a gestão dos programas, via de regra, leva em consideração esse documento para direcionar ações, contratar docentes, selecionar discentes e definir direcionamento das produções do curso. Por outro lado, haveria um questionamento quanto à definição e à utilização de indicadores e faixas (parâmetros) por parte da Comissão Avaliadora.

De acordo com Maccari (2008), a política adotada pela Capes tem sido eficaz sob diversos pontos de vista, incluindo a contribuição na quantidade de cursos de mestrado e doutorado, na melhoria da qualidade destes, o que pode ser notado, por exemplo, na evolução nos indicadores de qualidade representados pelos itens do nos diversos ciclos avaliativos. Ainda, segundo Martins et al. (2012), outro aspecto que permitiria uma indução à melhoria dos indicadores, além de conferir certa transparência à avaliação, refere-se à divulgação pública de fichas de avaliação

e da quantidade de produções de todos os cursos avaliados. De posse desses documentos, os programas podem avaliar, de forma comparativa, como estão relativamente aos demais programas da Área de Avaliação da Capes. Isso permitiria uma análise crítica do desempenho do seu próprio programa e poderia induzi-lo a melhorar em itens específicos em que o curso estaria aquém dos demais programas.

Maccari et al. (2009) também corroboram a opinião de que a avaliação e seus critérios podem ser usados como diretrizes para gestão estratégica dos programas de pós-graduação, bem como o sistema de avaliação tem papel importante na certificação da qualidade de uma instituição de ensino.

De maneira geral, a Avaliação da Capes pode ser considerada como um sistema relativamente complexo e sistematizado, com estruturação que permite avaliar diversos fatores (itens) ligados à pós-graduação. A nota atribuída a cada programa deriva da análise baseada em indicadores quantitativos, como número de docentes, discentes titulados, produções, dentre outros, bem como de relatórios e possíveis visitas realizadas por comitês (LEITE; VIANA; PEREIRA, 2006).

Esse Sistema de Avaliação da Capes, de acordo com Gazzola (2008), não possui semelhantes na América Latina e pode ser considerado como referência mundial. Segundo Schwartzman (2010), em termos de magnitude, a pós-graduação *Stricto Sensu* no Brasil possui proporção comparável ao encontrado nos países economicamente mais importantes do mundo, tanto em número de cursos, docentes e discentes, quanto à produção qualificada em periódicos, sendo que esta última tem crescido de forma sistemática e acelerada.

Esta avaliação, conforme comentado por Maccari (2008), apresentava a mesma sistemática entre todas as Áreas de Avaliação. À época, havia pesos referenciais para cada um dos quesitos avaliados (proposta do programa, corpo docente, corpo discente, teses e dissertações, produção intelectual e inserção social), sendo que cada área poderia alterar estas ponderações em até 5%. Esses quesitos eram divididos em itens, sendo que a Área de Avaliação apresentava relativa liberdade para incluir, excluir e definir pesos para estes itens.

A despeito dos pesos atribuídos aos itens e quesitos, via de regra, pondera-se fortemente aqueles relacionados à produção intelectual qualificada, como os artigos em periódicos. De acordo com Horta (2006), o estabelecimento de critérios mais objetivos de avaliação criou uma procura, por parte dos pesquisadores, pela produção de artigos em quantidade. No entanto, conforme discutido por esse autor, não deveria haver dissociação entre quantidade e qualidade, sendo importante estabelecerem-se pesos entre estas duas características dos trabalhos produzidos.

Um enfoque produtivista é citado como o ponto mais polêmico da Avaliação da Capes (HORTA, 2006; NASCIMENTO, 2010) e, segundo Oliveira e Almeida (2011), ocorreu, desde o princípio da pós-graduação, uma associação entre as políticas de financiamento e o sistema de avaliação, de forma que as práticas produtivistas se tornaram determinantes.

Horta (2006), ao entrevistar coordenadores de programas de pós-graduação, identificou um anseio oriundo da Área de Astronomia/Física no sentido de que deveria haver uma busca, por parte da Capes e do seu Comitê Técnico Científico - CTC, por aperfeiçoamento da avaliação, priorizando a qualidade dos



mestres e doutores titulados, atribuindo-se menos notoriedade a aspectos puramente quantitativos, como número de publicações, quantidade de docentes bolsistas de produtividade científica do Conselho Nacional de Desenvolvimento Científico e Tecnológico - CNPq, dentre outros. Segundo esses coordenadores entrevistado, analisando o nível dos candidatos a cargos de docentes em concursos públicos, podia ser identificada baixa qualidade dos mestres e doutores formados. Em tese, os programas estariam priorizando a formação de pós-graduandos que publicassem grande quantidade de artigos em pouco tempo, em detrimento de uma formação acadêmica e científica mais sólida, que permitiria melhor amadurecimento dos alunos para publicação de artigos de melhor qualidade.

Também demonstrando preocupação com a qualidade dos egressos dos programas de pós-graduação, Hortale (2003) afirmou que a Avaliação tem como foco principal a pesquisa em si, e não a qualidade do ensino.

De maneira semelhante, Horta (2006) afirma que coordenadores consideram relevante a utilização de indicadores quantitativos na avaliação, no sentido de torná-la mais objetiva, mas demonstram preocupação com o produtivismo. Na Área de Educação, um dos coordenadores cita a necessidade de se avaliar de forma mais qualitativa as produções. Na Área de Economia, um deles também afirma ter preocupação no superdimensionamento da produção intelectual, sugerindo que a formação de discentes qualificados deveria ser o foco central da avaliação. Corroborando outros coordenadores e indo além, coordenadores das áreas de Farmácia, Medicina I e Engenharias I criticam a mensuração da qualidade de um programa pelo impacto das suas produções, propondo que esta qualidade deveria ser medida pelo impacto na sociedade. Essa valorização das publicações científicas também é citada criticamente por Spagnolo e Souza (2004), Nascimento (2010) e Marengo (2015).

No caso da área de Educação, Gatti et al. (2003) sugerem que a contribuição social de um programa deveria ser considerada como fator importante na avaliação dos programas. Segundo esses autores, aspectos como: a formação de docentes para o ensino superior e de profissionais para gestão de instituições públicas, as atividades ligadas à rede pública de ensino, as atividades de extensão, bem como as participações em conselhos de educação e em comitês/associações científicas, deveriam contabilizar positivamente aos programas. Além disso, também foram citados aspectos como a nucleação/solidariedade (auxílio na implantação de novos programas) e monitoramento do impacto do programa por meio de informações sobre destino dos egressos.

Outro alvo de críticas por parte de coordenadores de programas se refere ao caráter de avaliação *a posteriori*, que ocorre a cada ciclo avaliativo. Conto e Nunes (2017) indagam de que maneira um gestor de um programa poderia nortear ações para algo que deveria ter sido realizado anteriormente. Spagnolo e Souza (2004) apontam que, embora a avaliação apresente muitos indicadores quantitativos, ocorre ausência de clareza nas informações utilizadas na avaliação.

Ainda na linha de sugestões de melhorias que poderiam ser implementadas à Avaliação da Capes, Guimarães (2007) aponta a necessidade de agregar componentes de prioridade definidos por especialistas tanto de dentro quanto de fora da comunidade científica na avaliação, a fim de que, em última instância, isso estimule a associação entre empresas e grupos de pesquisa, estimulando o

desenvolvimento e a inovação nas empresas. Essa perspectiva continua em evidência, tendo como desdobramento a criação da Empresa Brasileira de Pesquisa e Inovação Industrial - Embrapii em 2013, sediado no Ministério da Ciência, Tecnologia, Inovações e Comunicações - MCTIC e tendo o Ministério da Educação - MEC como “instituição interveniente”. A criação dos doutorados profissionais, que deve ocorrer em 2018, pode ser entendida como mais um movimento nesse sentido, com impactos que ainda não se pode vislumbrar em termos de organização do sistema de pós-graduação.

Conto e Nunes (2017) comentaram que os Documentos de Área servem como diretrizes que são estabelecidas por pesquisadores da área, embora haja restrição imposta pelos itens e quesitos que devem ser obedecidos por todos os colégios, grandes áreas e áreas de avaliação. Nesse contexto, em trabalho de Horta (2006), um coordenador de curso da área de Letras/Linguística, entrevistado em sua pesquisa, afirma que as áreas contidas no Colégio de Humanidades da Capes deveriam ter suas especificidades consideradas. Segundo esse coordenador, o processo de Avaliação é eminentemente quantitativo, o que não seria o adequado no caso das ciências humanas. No mesmo trabalho, um coordenador da área de Serviço Social sugere que os parâmetros considerados de excelência sejam definidos no âmbito da própria área após participação de representantes dos programas quanto a eventuais discordâncias. Gatti et al. (2003) também comentam que seria conveniente a existência de critérios gerais, mas uma flexibilização na sua operacionalização de forma a garantir as especificidades das áreas de avaliação.

Em síntese, a bibliografia evidencia a complexidade do sistema de avaliação da pós-graduação. Ao final de cada período de avaliação são realizadas “avaliações da avaliação”, indicando necessidades de mudança e redirecionamento.

Um dos aspectos destacados é a necessidade de “discriminação” dos quesitos da avaliação, no qual se pretende avançar com o presente artigo.

## **2.4 Sistema de avaliação da Capes nos dois ciclos recentes (2010-2012 e 2013-2016)**

Uma das incumbências da Capes é avaliar os programas de Pós-Graduação *Stricto Sensu* no Brasil, atribuindo notas que vão de 1 a 7, tanto de programas já em funcionamento quanto as propostas de novos cursos.

A avaliação de cursos já existentes é realizada em ciclos que variaram ao longo do tempo, tendo permanecido como trienais até o período 2010-2012, e passando a quadrienal no período 2013-2016. Cada área de avaliação possui um coordenador geral e dois coordenadores-adjuntos, um para programas acadêmicos e outro para mestrados profissionais. O coordenador de área nomeia uma comissão que elabora um Documento de Área com a forma de avaliação, seguindo algumas diretrizes básicas definidas pela Diretoria de Avaliação - DAV da Capes.

Os programas são avaliados quanto a uma série de critérios distribuídos em cinco quesitos: (1) Proposta do curso; (2) Corpo docente; (3) Corpo discente; (4) Produção intelectual; (5) Inserção social. Os quesitos são compostos a partir de itens, que, por sua vez, podem conter subitens para a sua definição. Para cada item ou subitem, cada programa recebe uma nota de 1 a 5. Cada item ou sub-item tem um peso dentro do quesito correspondente, sendo que a nota de um programa em um quesito é a média ponderada das notas recebidas nos itens. Cada quesito

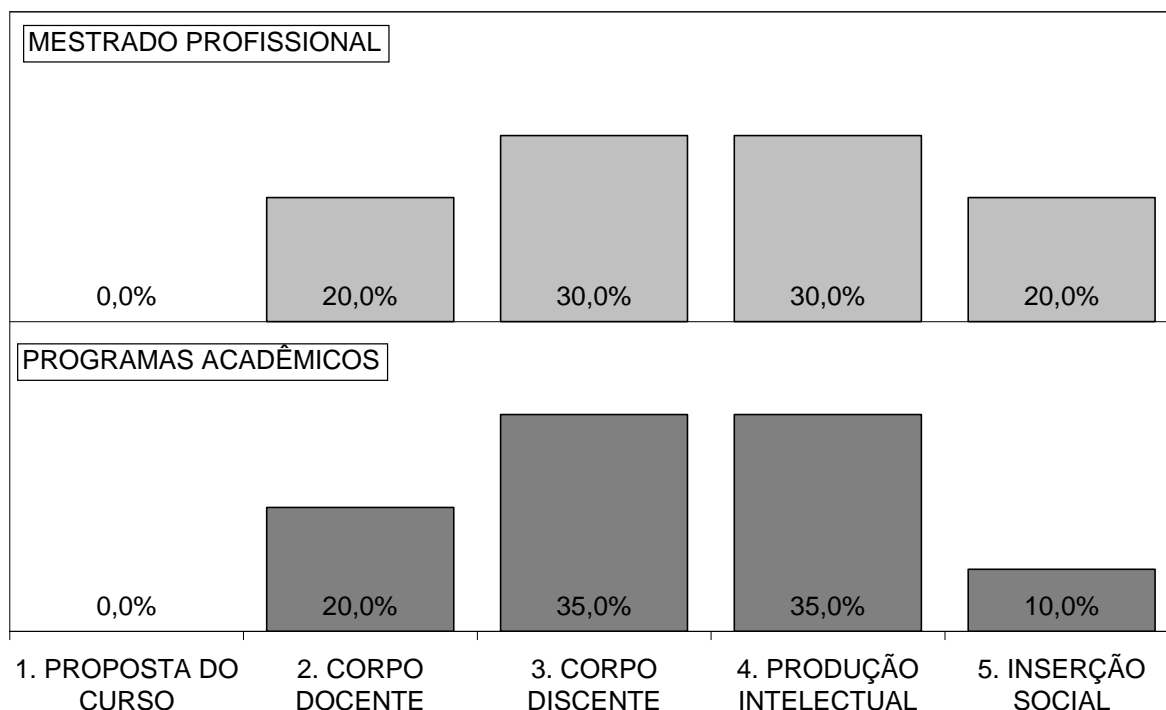
também tem um peso, sendo que a nota geral de um programa é a média ponderada destes quesitos. Em princípio, um programa recebe uma nota entre 1 e 5, sendo que programas com notas 3 ou superior continuam autorizadas a funcionar, enquanto os programas que obtêm nota 1 ou 2 recebem indicação para descredenciamento. Os programas com nota 5 são submetidos a uma outra rodada de avaliação, considerando critérios como internacionalização, nucleação, solidariedade e qualidade da produção científica, dentre outros. Nesse momento, são definidos os programas indicados a receber notas 6 e 7. A homologação dos resultados é realizada no âmbito da DAV.

Compete ressaltar que, no documento de área e no transcorrer da avaliação, existem regras específicas para programas acadêmicos e mestrados profissionais, inclusive com comissões diferentes para avaliação nestas duas modalidades. Neste artigo, centra-se a análise nos programas acadêmicos.

## 2.5 Documento da área de Planejamento Urbano e Regional e Demografia para avaliação da pós-graduação (período 2010-2012)

Os pesos dos quesitos e dos itens na avaliação da área de PLURed no período 2010-2012 são apresentados, respectivamente, nas Figuras 2 e 3.

**Figura 2. Pesos dos quesitos na avaliação da área de Planejamento Urbano e Regional e Demografia.**

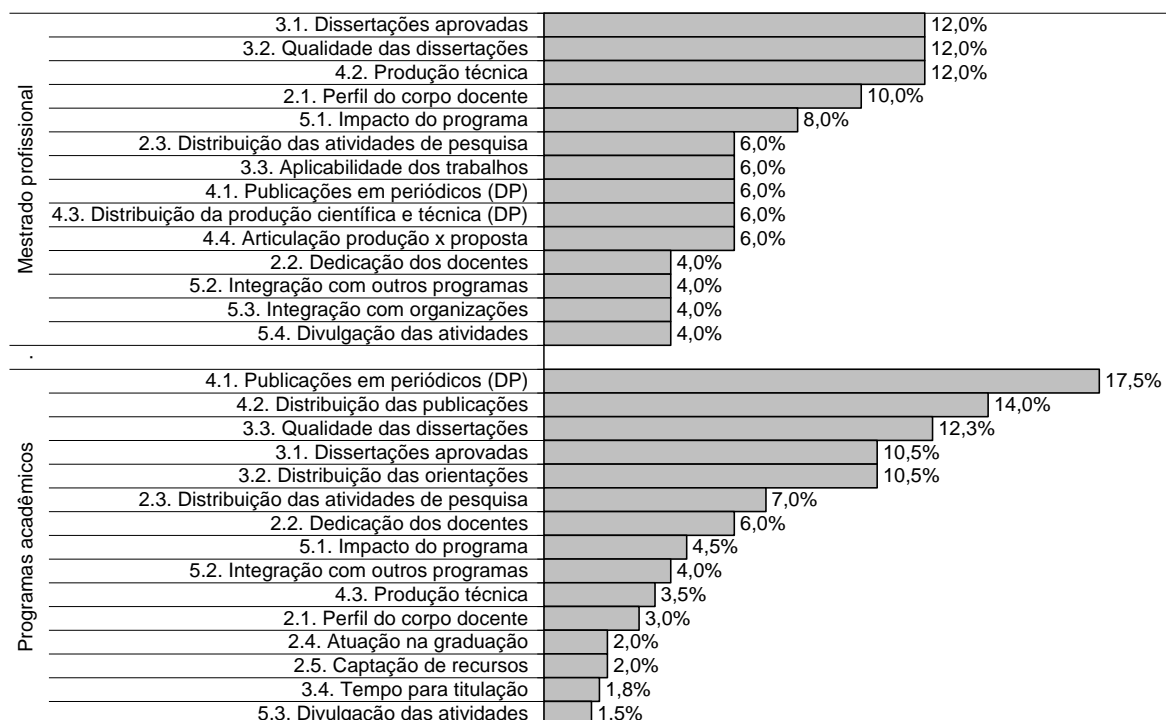


Fonte: Documento de área Capes (...).

Conforme pode ser observado na Figura 2, os quesitos que possuem maiores pesos na avaliação são (3) Corpo Discente e (4) Produção Intelectual, tanto na avaliação dos programas profissionais quanto dos mestrados profissionais. Comparando-se as duas modalidades, nesses quesitos, o peso é menor na avaliação

dos mestrados profissionais (30%) quando comparada aos programas acadêmicos (35%). Essa diferença de 5% em cada uma dessas duas dimensões encontra-se no quesito (5) inserção social, que tem peso de 20% nos mestrados profissionais. É importante frisar que o quesito (1) Proposta do Curso, embora tenha peso 0, na prática, funciona como uma “trava”, uma vez que é impreterível que o corpo docente, o corpo discente, a produção intelectual e a inserção social estejam alinhados com uma proposta de curso coerente.

**Figura 3. Pesos dos itens na avaliação da área de Planejamento Urbano e Regional e Demografia**



Fonte: Documento de área CAPES (<http://www.capes.gov.br/component/content/article/44-avaliacao/4680-planejamento-urbano-e-regionaldemografia>).

Observando a Figura 3, percebe-se que a área de PLURed atribui maiores pesos na avaliação dos programas acadêmicos aos itens relacionados ao quesito “(4) produção intelectual” (4.1. Produção em periódicos e 4.2. Distribuição das publicações), seguida de itens da dimensão “(3) Corpo discente” (3.3. Qualidade das dissertações; 3.1. Dissertações aprovadas e 3.2. Dissertações aprovadas). No caso da avaliação dos mestrados profissionais, dos três itens com maior peso, dois são da dimensão “(3) Corpo discente” (3.1. Dissertações aprovadas e 3.2. Qualidade das dissertações) e um deles da dimensão “(4) Produção intelectual” (4.2. Produção técnica). Sobressaem ainda, na avaliação do mestrado profissional, os itens 2.1. Perfil do corpo docente e 5.1. Impacto do programa.

### **3 PROPOSTA DE METODOLOGIA PARA AVALIAR A CAPACIDADE DE DISCRIMINAÇÃO**

A seguir, apresenta-se um conjunto de possíveis abordagens que poderiam ser utilizadas para melhorar a discriminação dos quesitos e itens, de maneira a indicar de forma mais precisa a diferenciação entre os programas de pós-graduação no momento da avaliação.

Para realizar essas abordagens, foram utilizados os dados relativos às notas obtidas em cada item e em cada quesito, de acordo com os resultados finais disponibilizados no Relatório da Avaliação do período 2010-2012, disponibilizados no site da Área PLURed, na página da Capes.

#### **3.1 Parâmetros estatísticos**

Com base nas frequências de programas nos estratos de nota (1 a 5) em cada item, presentes no Relatório da Avaliação do período 2010-2012, foram testados vários parâmetros estatísticos no sentido de verificar a capacidade de discriminação de um critério.

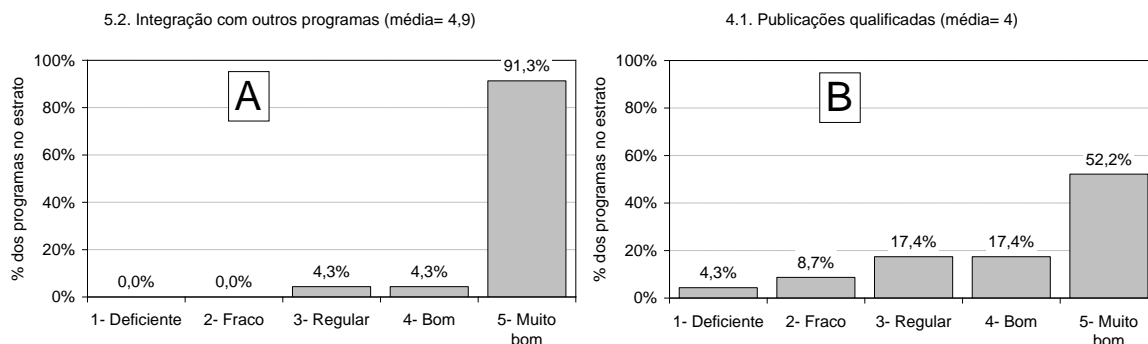
Em um item que teria capacidade de discriminação baixa (neste caso, nula), por exemplo, apresentou-se 100% dos programas em um mesmo estrato. Em princípio, o máximo de discriminação seria alcançado no caso de existirem 20% dos programas em cada um dos 5 estratos (1- deficiente; 2- fraco; 3- regular; 4- bom; 5- muito bom). Importante notar que não existe o objetivo de buscar uma distribuição uniforme nos estratos, mas esta equidade é um referencial interessante para medir a capacidade de discriminação de uma dimensão ou item.

Alguns indicadores da capacidade de discriminação são apresentados a seguir, sendo comentado o princípio básico e as limitações de sua utilização.

##### *Média*

O parâmetro média poderia ser um indicador de discriminação, uma vez que valores extremos (próximos de 1 ou de 5) poderiam indicar que praticamente todos os programas encontram-se classificados em um dos dois estratos. Na Figura 4A, verifica-se que a média do item foi de 4,9, sendo que 91,3% dos programas estão enquadrados no estrato “5- muito bom”, o que evidencia baixa capacidade de discriminação do item. Por outro lado, o item da Figura 4B apresenta média igual a 4,0, com uma melhor distribuição dos programas nos estratos.

**Figura 4. Utilização do parâmetro média para avaliação da capacidade de discriminação (A- baixa capacidade; B- alta capacidade)**

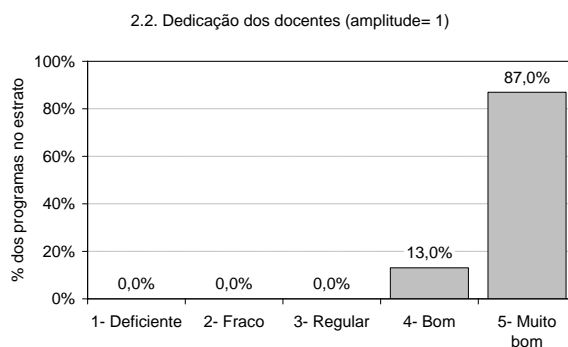


Fonte: Documento de área Capes (...).

### Amplitude

Entende-se como amplitude, neste caso, a diferença entre a nota mais alta que teve frequência diferente de zero e a nota mais baixa que teve frequência diferente de zero.

**Figura 5. Exemplo de distribuição de frequência para cálculo da amplitude**

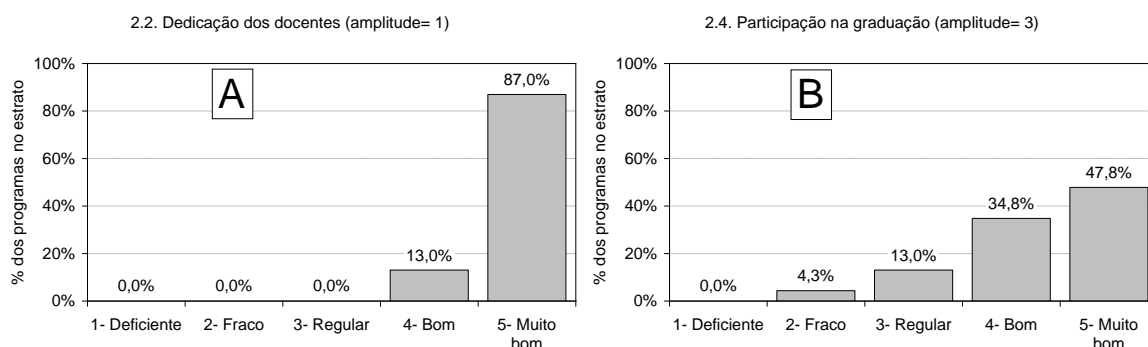


Na Figura 5, observa-se que a nota mais alta com frequência diferente de zero é o “5-muito bom”. A nota mais baixa com frequência diferente de zero é o “4-bom”. A amplitude é a diferença dos dois: Amplitude = 5 – 4 = 1.

Em princípio, quanto MAIOR a amplitude, MAIOR a capacidade de discriminação.



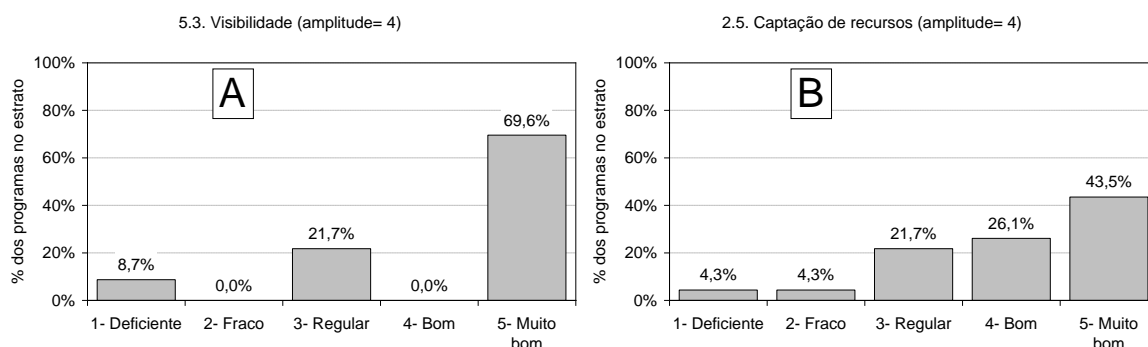
**Figura 6. Utilização do parâmetro amplitude para avaliação da capacidade de discriminação (A- baixa capacidade; B- alta capacidade)**



Na Figura 6A, observa-se que os programas receberam apenas notas 4 ou 5, o que representa uma baixa capacidade de discriminação e uma amplitude igual a 1. Na Figura 6B, por outro lado, percebe-se que existem programas com notas 2 até 5, o que representa boa capacidade de discriminação e uma amplitude igual a 3.

Uma restrição associada a esse indicador ocorre, pois ele considera apenas os extremos, sem levar em conta a existência de frequências nos valores intermediários, conforme pode ser constatado na Figura 7.

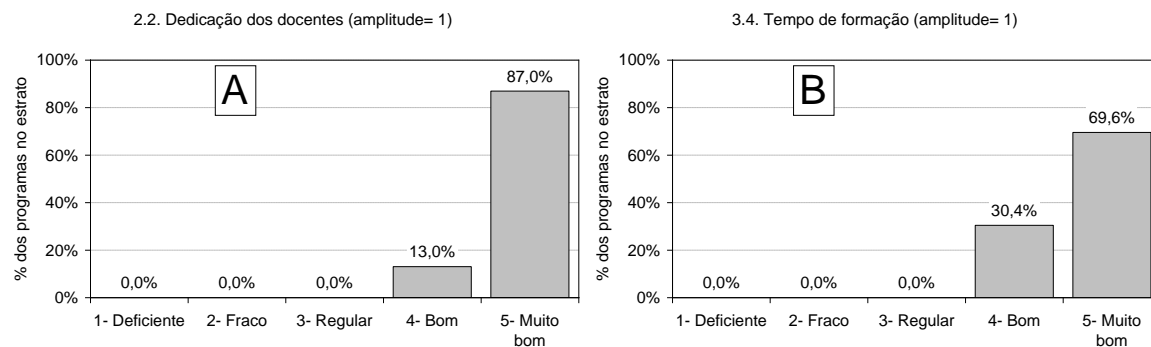
**Figura 7. Exemplo de distribuição de frequência em: (A) 3 classes de nota e com amplitude igual a 4 e (B) 4 classes de nota e com amplitude igual a 4**



As distribuições de notas mostradas nas Figuras 7A e 7B possuem a mesma amplitude. Porém, na Figura 7B, as frequências estão mais bem distribuídas, com mais classes de notas contempladas (notas 1, 2, 3, 4 e 5) e com frequências que vão de 4,3 a 43,5%. Na Figura 9, existem programas enquadrados nas notas 1, 3 e 5, com as frequências indo de 0 a 69,6%.

Outra restrição relaciona-se ao fato de que a amplitude não leva em consideração as frequências, mas somente os valores extremos.

**Figura 8. Exemplo de distribuição de frequência com amplitude igual a 1 e (A) alta concentração em uma classe de nota e (B) média concentração em uma classe de nota**

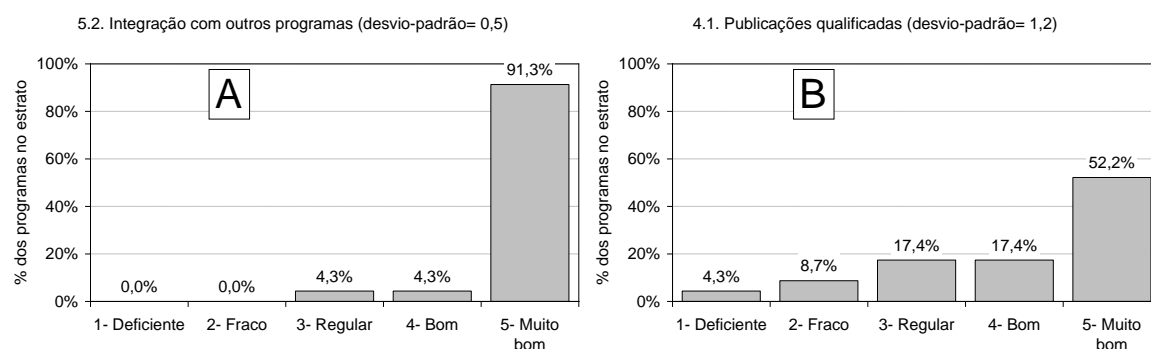


Nas Figuras 8A e 8B, as distribuições de frequência têm amplitude igual a 1, porém, percebe-se que a distribuição na Figura 11 é mais heterogênea, com aproximadamente 90% dos programas concentrados num único estrato de nota.

#### Desvio-padrão

Outro parâmetro estatístico que pode ajudar na identificação do grau de discriminação de um critério é o desvio-padrão. Em princípio, quanto maior o desvio-padrão, maior a capacidade de discriminação de um item.

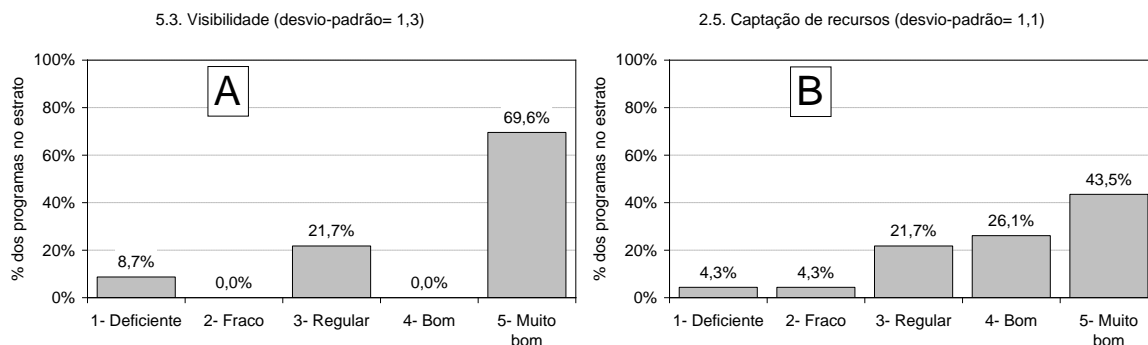
**Figura 9. Utilização do parâmetro desvio-padrão para avaliação da capacidade de discriminação (A- baixa capacidade; B- alta capacidade)**



Na Figura 9, observam-se exemplos de distribuição de frequência de critérios com baixa e alta capacidade de discriminação, respectivamente. Como se percebe, nos casos em que os dados são concentrados em torno de um dos estratos de nota (Figura 9A), o desvio padrão é menor do que os casos em que as frequências são mais bem distribuídas e, conseqüentemente, com maior capacidade de discriminação (Figura 9B).

Uma restrição ao indicador em situações em que se verificam altas frequências nos estratos extremos de nota (1 e 5). Nesse caso, o desvio-padrão acaba sendo alto, conforme pode ser verificado na Figura 10.

**Figura 10. Exemplo de distribuição de frequência em critério com alto desvio-padrão devido a: (A) frequência nos estratos 1 e 5 e (B) boa distribuição dos programas**



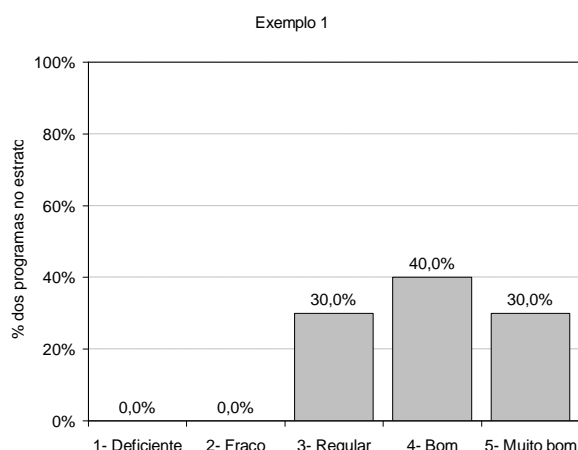
As frequências nos estratos 1 (4,3%) e 5 (43,5%) na Figura 10A são menores do que no exemplo da Figura 10B, levando o desvio-padrão a ser menor. No entanto, intuitivamente, é possível verificar que as frequências da Figura 10B são mais bem distribuídas.

#### *Grau de assimetria*

Este indicador demonstra se o conjunto de dados está mais concentrado à esquerda (assimetria à direita ou positiva) ou à direita (assimetria à esquerda ou negativa). Em princípio, uma assimetria muito baixa ou muito alta indica alta concentração dos dados próximos aos valores 5 ou 1, respectivamente, o que não seria interessante sob o ponto de vista de capacidade de discriminação.

Um problema nesse indicador é que, no cálculo do parâmetro, apenas são considerados os estratos que tiveram frequência diferente de 0 nas extremidades, como pode ser observado na Figura 11.

**Figura 11. Exemplo fictício de distribuição para cálculo do grau de assimetria**



Dessa forma, no exemplo da Figura 11, ocorre uma distribuição simétrica (assimetria=0), com ponto médio no estrato 4 (bom) e distribuição simétrica nos outros dois estratos (3-regular e 5-muito bom). No entanto, parece claro que a

distribuição está concentrada à direita, se forem considerados os estratos 1 (deficiente) e 2 (fraco).

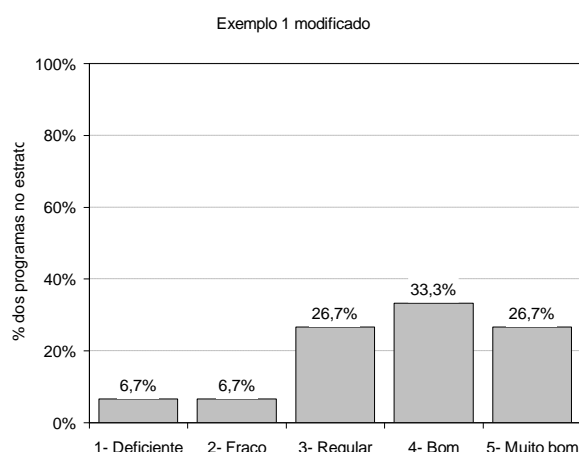
Para resolver esse problema, foi acrescentada 1 unidade em cada estrato, de forma que todas as porcentagens ficam diferente de 0, conforme apresentado na Tabela 1.

**Tabela 1- Redistribuição da frequência para cálculo do grau de assimetria**

Estrato	Situação real		Situação modificada	
	Nº de cursos	%	Nº de cursos (adicionando 1)	%
1- Deficiente	0	0%	1	6,7%
2- Fraco	0	0%	1	6,7%
3- Regular	3	30%	4	26,7%
4- Bom	4	40%	5	33,3%
5- Muito bom	3	30%	4	26,7%
TOTAL	10	100%	15	100,0%

Na Figura 12, é possível verificar a nova distribuição de frequências após acréscimo de 1 unidade em cada estrato de nota.

**Figura 12. Exemplo fictício de distribuição para cálculo do grau de assimetria.**



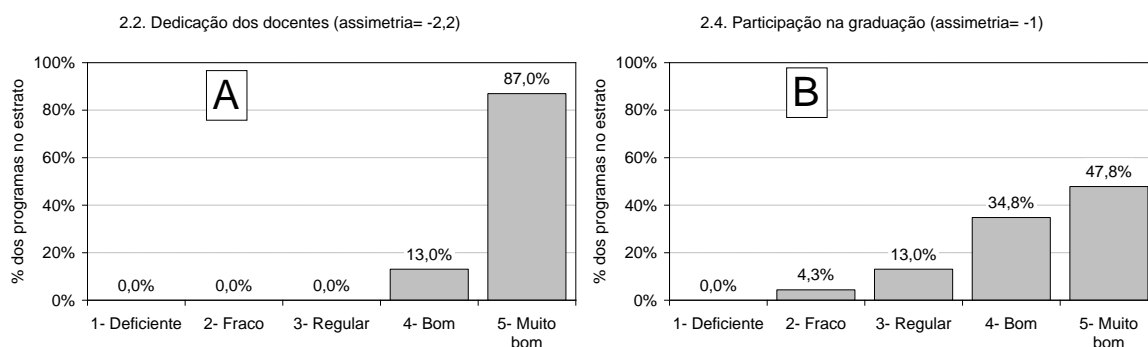
Adicionando 1 unidade a cada estrato, a distribuição ficou assimétrica, com dados concentrados à direita (assimetria à esquerda ou negativa). O valor do parâmetro, que antes era igual a 0, passou a ser igual a -0,57.

Compete ressaltar que os cálculos do grau de assimetria apresentados na seção 4 são todos realizados acrescentando-se 1 unidade a cada estrato de nota.

No entanto, exceto no exemplo fictício apresentado na Figura 12, todos os demais gráficos são apresentados sem acréscimo da unidade.

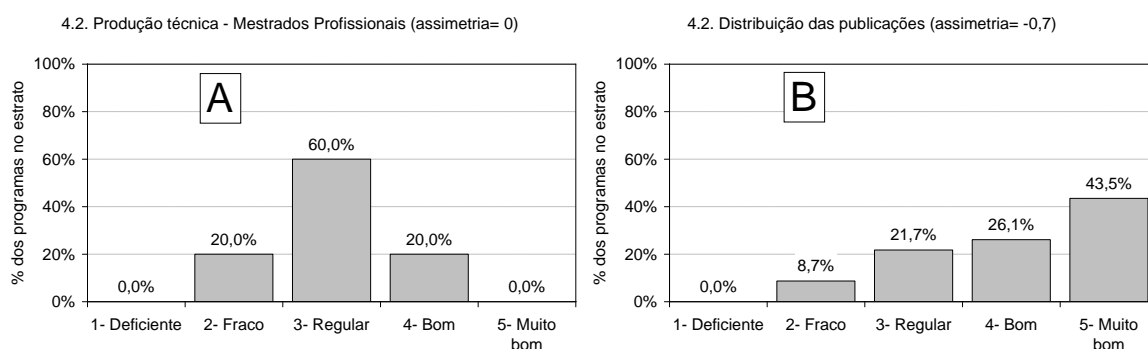
Feitas as considerações quanto à forma de cálculo do grau de assimetria, a distribuição da Figura 13A, comparativamente da Figura 13B, é evidente que os dados estão mais concentrados à direita (assimetria à esquerda), com grau de assimetria mais negativa, indicando que as frequências não estão bem distribuídas.

**Figura 13- Utilização do parâmetro assimetria para avaliação da capacidade de discriminação (A- baixa capacidade; B- alta capacidade)**



Uma restrição da utilização do grau de assimetria é que esse indicador mede se a distribuição de frequência está simétrica, mas não leva em conta, por exemplo, o número de estratos. No exemplo apresentado na Figura 14A (dados referentes aos mestrados profissionais), a distribuição é perfeitamente simétrica, mas não é tão bem distribuída por vários estratos como o presente na Figura 14B.

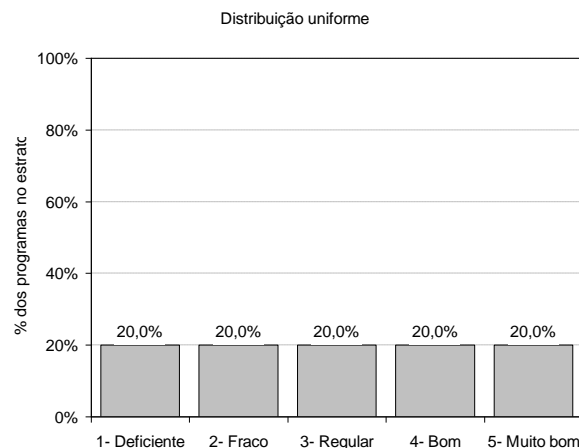
**Figura 14. Exemplo de distribuição de frequência com: (A) má distribuição nos estratos, mas simétrica e (B) boa distribuição nos estratos, mas assimétrica**



#### Desvio médio da distribuição uniforme

A distribuição uniforme seria aquela em que, para todos os estratos, as proporções são as mesmas. O máximo de discriminação possível seria obtido caso a distribuição fosse uniforme, conforme demonstrado na Figura 15.

**Figura 15- Distribuição uniforme de frequência de programas em dado critério fictício**



Para cada estrato, foi obtido o desvio da frequência do estrato em relação a 20%. Os módulos dos desvios foram somados, e este somatório foi dividido para obtenção do desvio médio da distribuição uniforme.

$$DM_{\text{Distr. Uniforme}} = \frac{\sum_{i=1}^5 |FR_i - 0,2|}{5}$$

DM = Desvio médio (Distr. Uniforme)

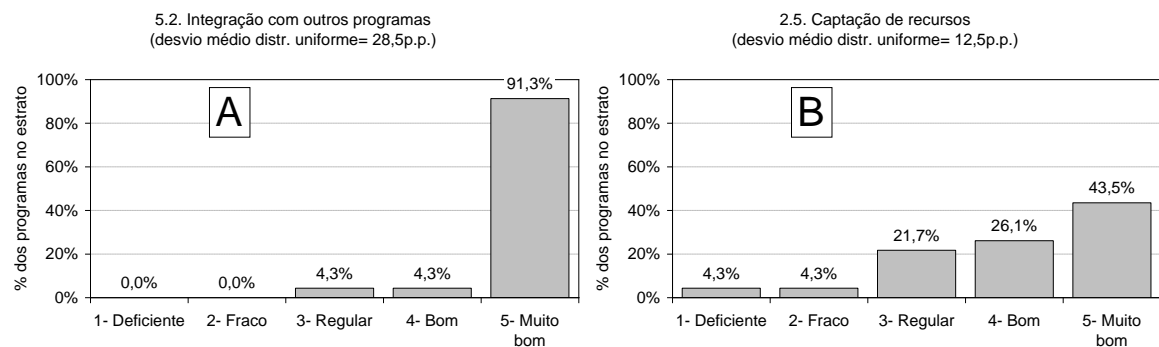
i = iésimo estrato de 1 a 5

FR = frequência relativa do estrato

É importante ressaltar que não se afirma que uma distribuição uniforme é a ideal, muito menos que deveria ser uma meta dos comitês das várias áreas de avaliação. No entanto, parece claro que esse tipo de discriminação seria o máximo teórico de discriminação. Não obstante, entende-se que alguns critérios podem nunca ter uma distribuição uniforme. Por exemplo, no caso da percentagem de docentes que participam de projetos de pesquisa, o ideal é que 100% dos programas tivessem a totalidade dos docentes nesta situação. Nesse caso, claramente, a distribuição uniforme não é interessante. No entanto, de forma geral, um critério cuja distribuição dos programas nos estratos de nota seja uniforme constitui forte indicador de que as diretrizes do critério foram eficazes na discriminação entre os programas, conforme pode ser observado nas Figuras 24 e 25.



**Figura 16- Utilização do parâmetro desvio médio em relação à distribuição uniforme para avaliação da capacidade de discriminação (A- baixa capacidade; B- alta capacidade)**



As frequências mostradas na Figura 16B são muito mais bem distribuídas do que as observadas na Figura 16A. Isso fica claro no desvio médio da distribuição uniforme (28,5 pontos percentuais contra 12,5 pontos percentuais). Em princípio, esse parâmetro estatístico será utilizado para classificar os critérios.

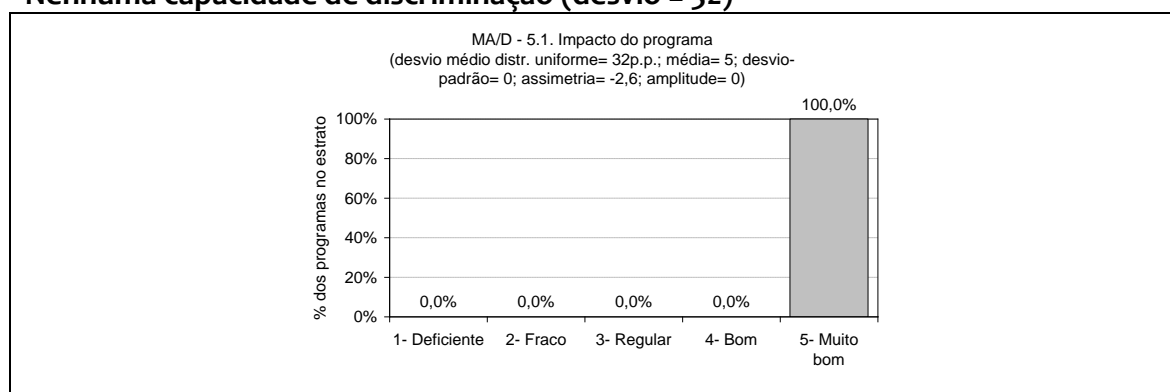
### 3.2 Classificação dos critérios quanto à capacidade de discriminação

Uma vez definidos diversos parâmetros estatísticos que são possíveis indicadores da capacidade de discriminação, propõe-se a classificação de um critério com base numa escala arbitrária de desvio em relação à distribuição uniforme:

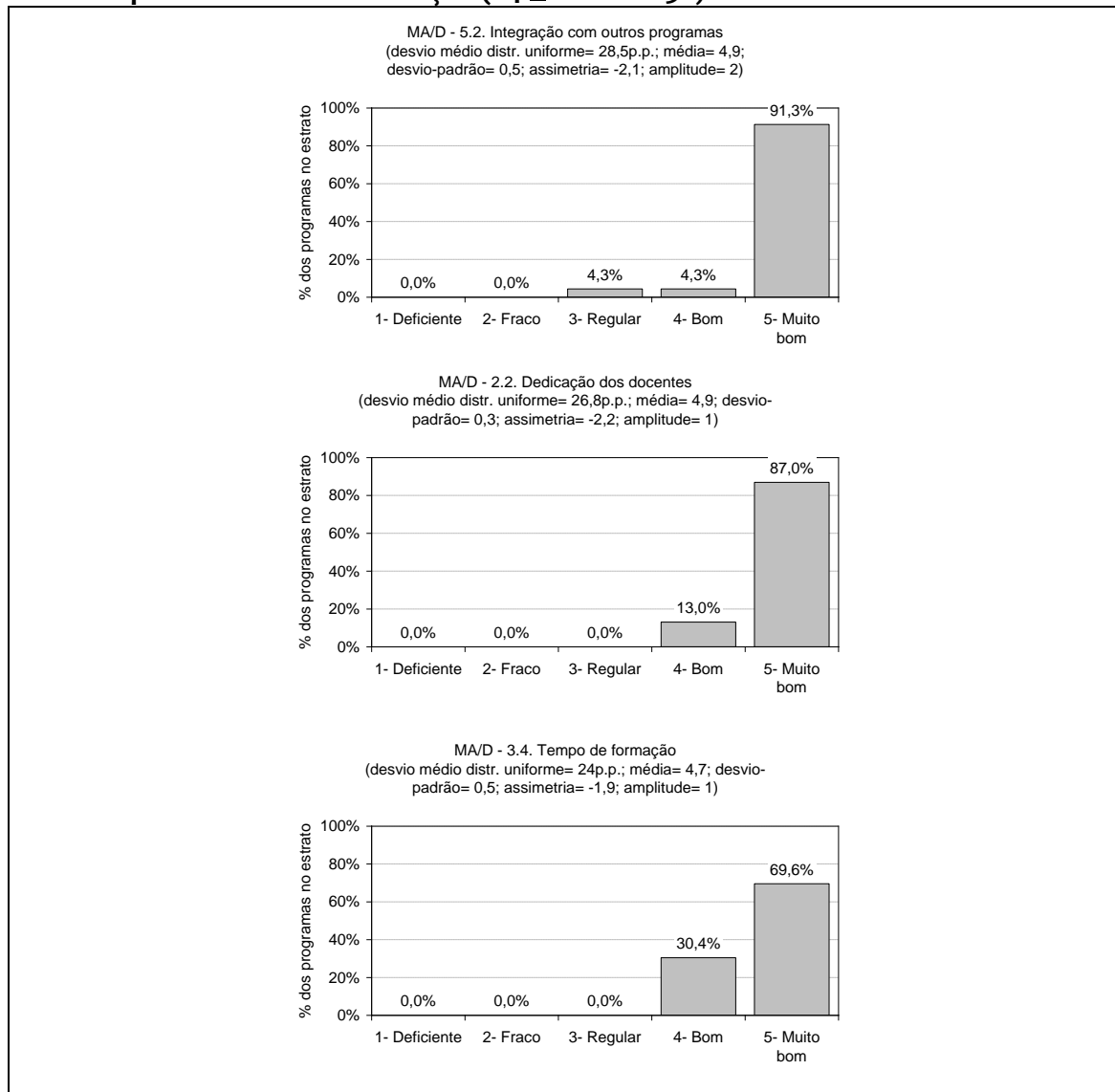
- nenhuma capacidade de discriminação: desvio = 32
- baixa capacidade de discriminação:  $24 \leq \text{desvio} < 32$
- média capacidade de discriminação:  $16 \leq \text{desvio} < 24$
- alta capacidade de discriminação: desvio < 16

## 4 Avaliação da capacidade de discriminação dos itens na área de Planejamento Urbano e Regional e Demografia (programas acadêmicos) – triênio 2010-2012

**Figura 17- Critérios com nenhuma capacidade de discriminação – área PLURed (programas acadêmicos)**  
**Nenhuma capacidade de discriminação (desvio = 32)**



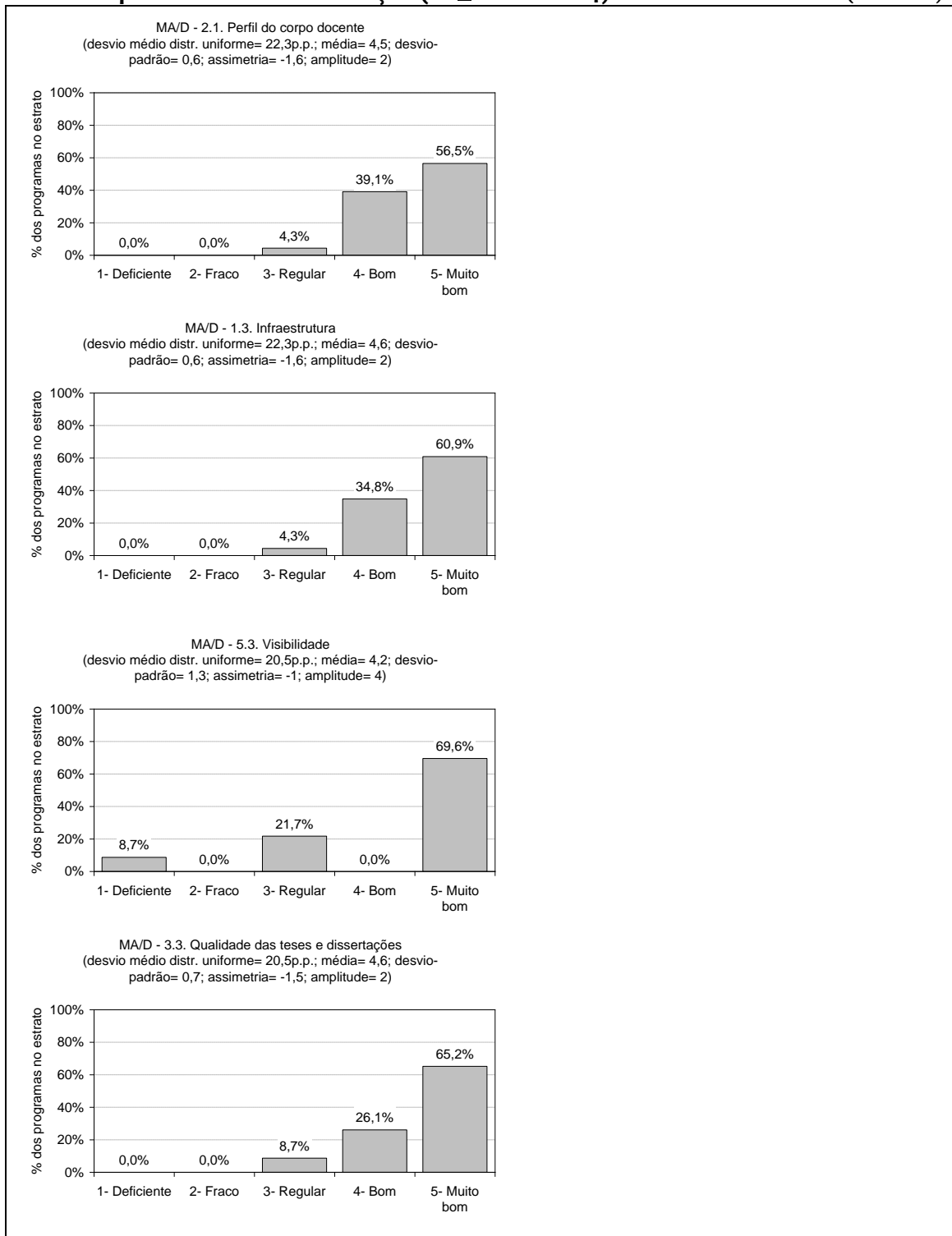
**Figura 18- Critérios com baixa capacidade de discriminação – área PLUREd  
(programas acadêmicos)**  
**Baixa capacidade de discriminação ( $24 \leq \text{desvio} < 32$ )**



**Figura 19- Critérios com média capacidade de discriminação – área PLURed (programas acadêmicos)**

**Média capacidade de discriminação ( $16 \leq \text{desvio} < 24$ )**

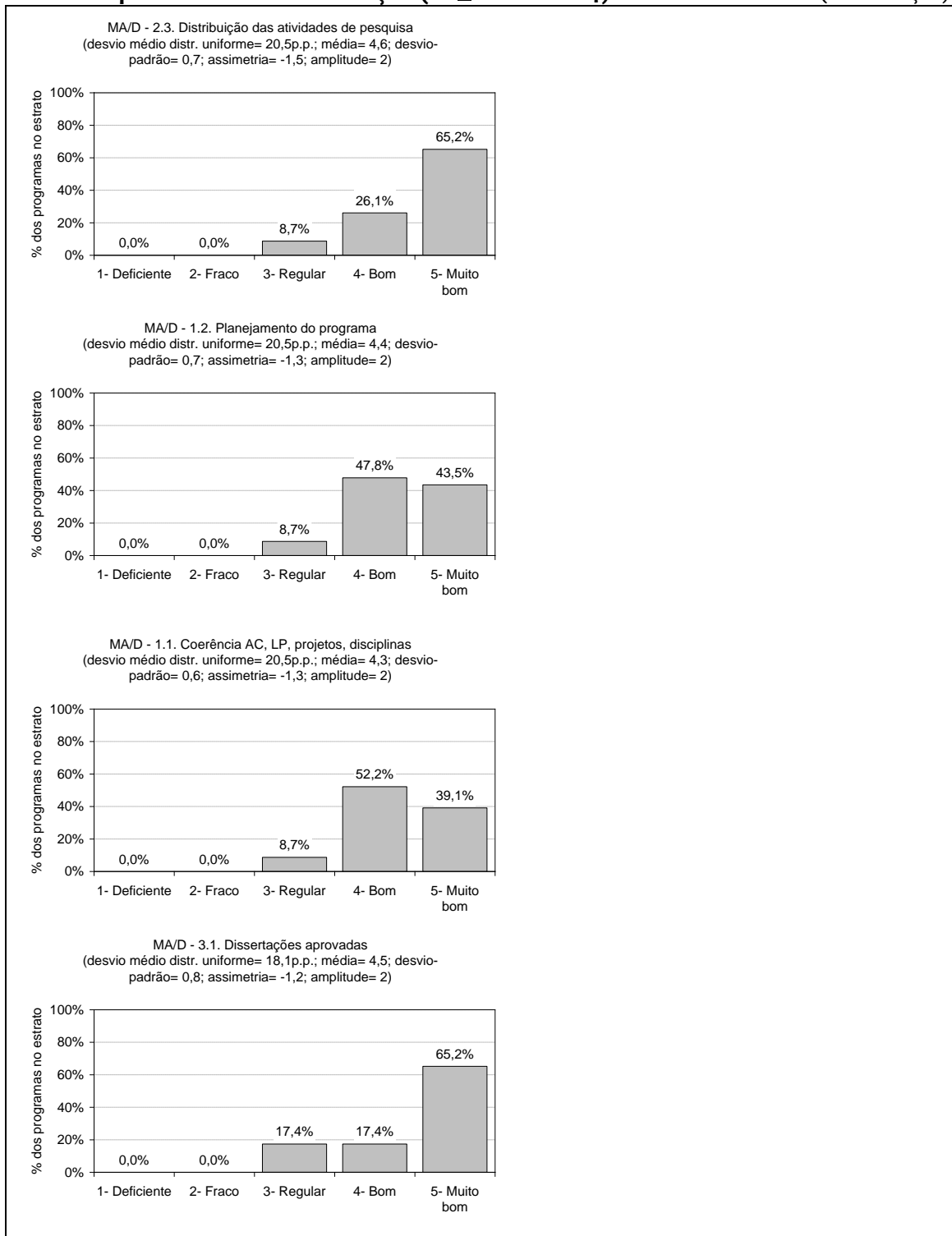
(continua)



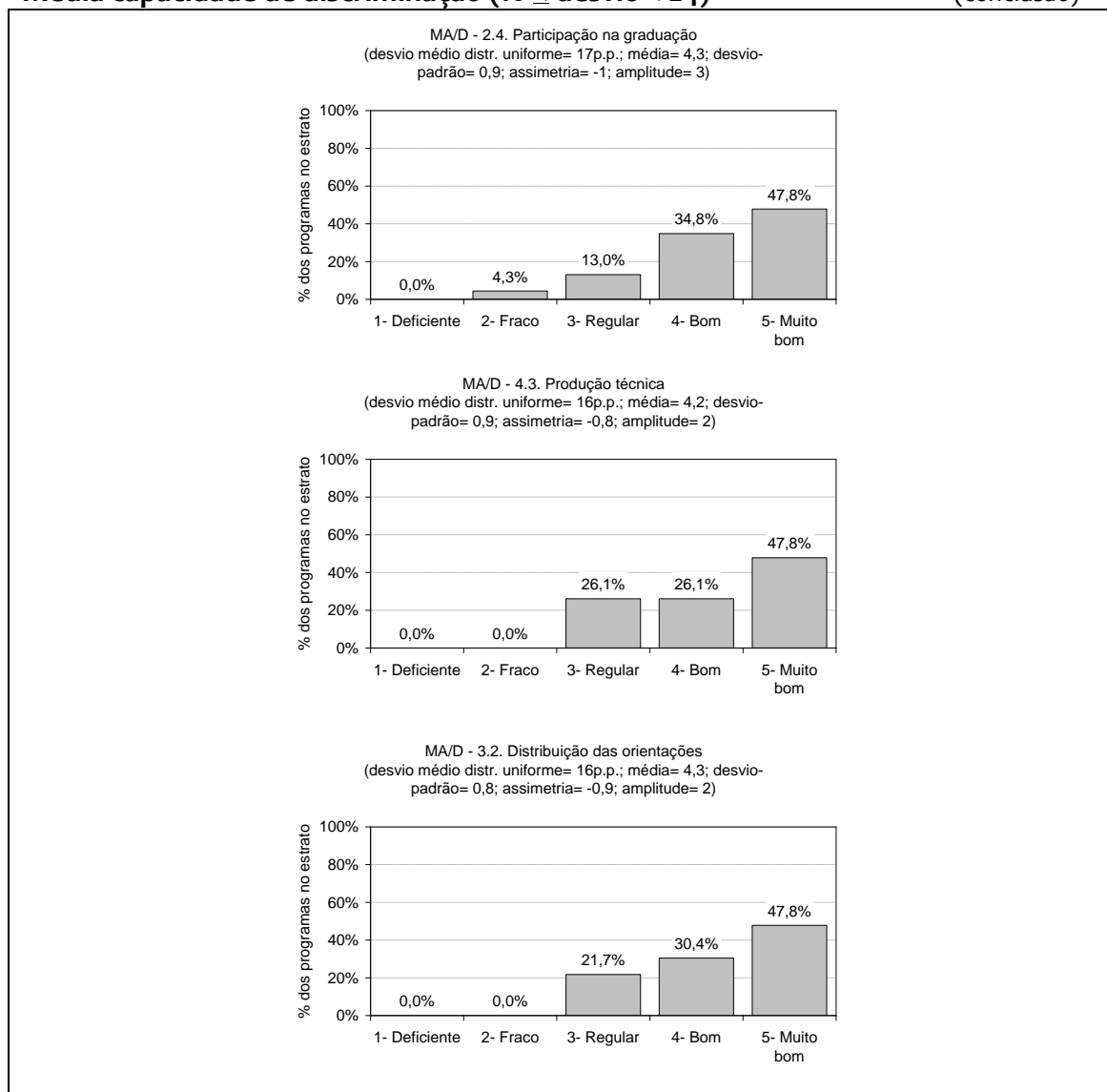
**Figura 19- Critérios com média capacidade de discriminação – área PLURed  
(programas acadêmicos)**

**Média capacidade de discriminação ( $16 \leq \text{desvio} < 24$ )**

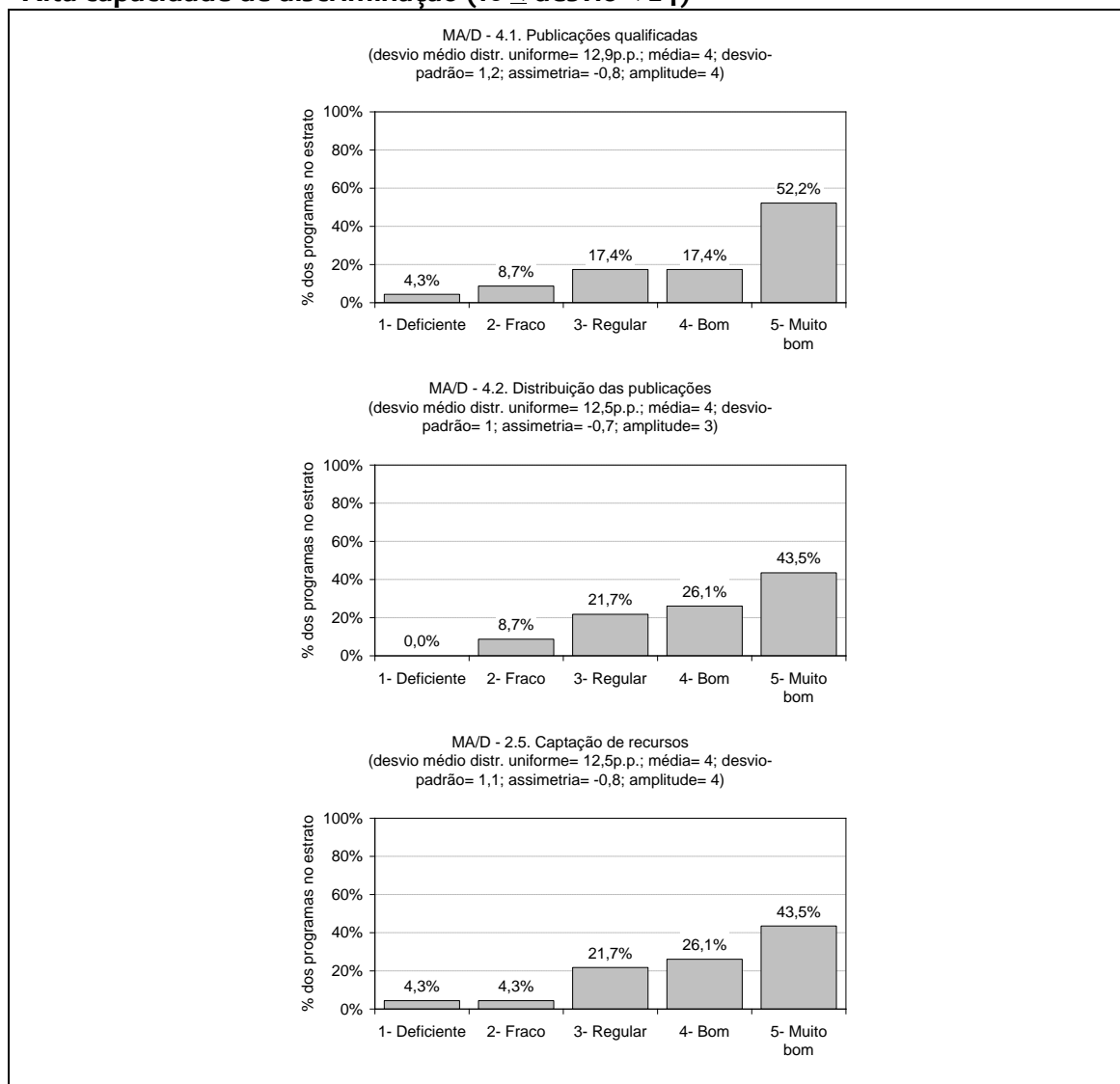
(continuação)



**Figura 19- Critérios com média capacidade de discriminação – área PLURed**  
**(programas acadêmicos)**  
**Média capacidade de discriminação ( $16 \leq \text{desvio} < 24$ )** (conclusão)



**Figura 21- Critérios com alta capacidade de discriminação – área PLURed  
(programas acadêmicos)**  
**Alta capacidade de discriminação ( $16 \leq \text{desvio} < 24$ )**



## 5 CONSIDERAÇÕES FINAIS

Um dos aspectos importantes no processo avaliativo é a construção de instrumentos que sejam capazes de classificar e diferenciar os entes que estão sendo avaliados. No caso da avaliação dos programas de pós-graduação que compõem o Sistema Nacional de Pós-Graduação, é importante que cada quesito e cada item ou subitem previsto no Documento de Área seja suficiente para permitir uma diferenciação entre os programas de pós-graduação. Nesse sentido, um quesito no qual mais de 90% dos programas sejam avaliados com nota 5 (muito bom) tem pouca capacidade de discriminação, o que acaba transferindo o peso relativo para os quesitos com maior capacidade de diferenciação. Esses quesitos com maior capacidade de diferenciação são, de maneira geral, aqueles com



indicadores quantitativos, o que aumenta o peso relativo desses quesitos no conjunto da avaliação.

Este artigo analisou a distribuição das notas atribuídas aos quesitos e itens da avaliação do período 2010-2012. Com base nessa distribuição, foi possível realizar uma análise para identificar quais quesitos conseguiram discriminar melhor os programas. Foram realizadas diversas abordagens no sentido de propor uma metodologia que permita avaliar a capacidade de discriminação dos quesitos e itens de avaliação da pós-graduação da Capes. Ao final, cada abordagem evidenciou suas potencialidades e limitações.

No âmbito da avaliação da área de Planejamento Urbano e Regional e Demografia da Capes, os itens foram classificados de acordo com a capacidade de discriminação com base na avaliação dos programas acadêmicos no triênio 2010-2012, evidenciando o peso preponderante que assumiram os quesitos compostos por elementos quantificáveis em relação aos demais quesitos. Salienta-se, nesse sentido de discriminação, a importância que tiveram os itens 4.1 (“Publicações Qualificadas”), 4.2 (“Distribuição das Publicações” entre o corpo docente permanente dos programas) e 2.5 (“Captação de Recursos”).

Observou-se também que o item 5.1 (“Impacto do Programa”) não resultou em discriminação, tendo todos os programas sido avaliados com nota 5 (“Muito Bom”). Certamente, a falta de indicadores mais precisos pesou nessa dificuldade de diferenciar os programas e exige uma atenção especial.

Outros três quesitos apresentaram baixa capacidade de discriminação: 5.2 (“Interação com outros programas”), no qual 91,3% dos programas foram avaliados com nota 5 (“muito bom”); 2.2 (“Dedicação dos docentes” aos programas de pós-graduação”), no qual 87% dos programas obtiveram nota 5; e o item 3.4 (“Tempo de formação”) com 69,6% obtendo nota 5 e 30,4% obtendo nota 4.

O item 3.4 já foi um item de grande importância no processo de avaliação, tendo em vista que o tempo de titulação foi muito elevado durante as primeiras décadas da pós-graduação no País. Entretanto, houve um grande esforço no sentido de reduzir e padronizar o tempo de titulação, fazendo com que, atualmente, esse item tenha pouca capacidade de diferenciação dos programas entre si.

Destaca-se que os resultados iniciais desse trabalho permitiram construir um arcabouço metodológico que serviu como suporte para a avaliação realizada no período 2013-2016.

## REFERÊNCIAS

CASTILHO, Fausto. **O conceito de universidade no projeto da UNICAMP**. Organização de Alexandre Guimarães Tadeu de Soares. Campinas, SP: Editora da Unicamp, 2008. 207 p.

CASTRO, C. M.; SOARES, G. A. D. Avaliando as avaliações da Capes. **Revista de Administração de Empresas**, Rio de Janeiro: v. 23, n. 3, p. 63-73, 1983.

CONTO, Sabrina Fonseca de; NUNES, Rogério da Silva. O documento de área da Capes e o impacto na gestão de cursos de pós-graduação Stricto Sensu: um estudo na Área de Comunicação e Informação. In: SIMPÓSIO AVALIAÇÃO DA EDUCAÇÃO SUPERIOR, 3., 2017, Florianópolis. Anais.... Florianópolis: Ufsc, p. 1 – 16, 2017.

GATTI, B. et al. O modelo de avaliação da CAPES. **Revista Brasileira de Educação**, n. 22, p. 137-144, 2003.

GUIMARÃES, R. O futuro da pós-graduação: avaliando a avaliação. **Revista Brasileira de Pós Graduação**, Brasília, v. 4, n. 8, p. 282-292, 2007

HORTA, J. S. B. Avaliação da Pós-graduação: com a palavra os Coordenadores de Programas. **Revista Perspectiva**, v. 24, nº 1, p. 19-47, jan/jun. 2006.

HORTA, J.S.B. Avaliação da Pós-graduação: com a palavra os Coordenadores de Programas. **Perspectiva**, Florianópolis, v. 24, n. 1, p. 19-47, 2006.

HORTALE, V. A. Modelo de avaliação CAPES: desejável e necessário, porém, incompleto. **Caderno de Saúde Pública**, v. 19, nº 6, p. 1837-1840, nov. 2003.

LEITE, M.F.B.; VIANA, A.B.N.; PEREIRA, G.G. Métodos quantitativos na avaliação da Capes: uma pesquisa bibliográfica. **FACEF Pesquisa**, v.9, n.2, p. 166-174, 2006.

MACCARI, E. A. et al. A Gestão Dos Programas De Pós-Graduação Em Administração Com Base no Sistema de Avaliação da CAPES. *Revista de Gestão USP*, v. 16, n. 4, p. 1–16, 2009.

MARENCO, A. When Institutions Matter: CAPES and Political Science in Brazil. **Revista de Ciência Política**, v. 35, n. 1, p. 33–46, 2015.

MARTINS. C.B. A influência do sistema de avaliação nos programas de pós-graduação *Stricto Sensu* brasileiro. **Revista GUAL**, Florianópolis, v. 5, n. 3, p. 155-178, 2012.

MEDEIROS, Michele Dias. **Avaliação da pós-graduação no Brasil e a produção intelectual: o "modelo da Capes" de 1975 a 2002**. 2016. 162 f. Dissertação (Mestrado) - Curso de Gestão Educacional, Universidade do Vale do Rio dos Sinos, São Leopoldo, 2016.

MORITZ, G. D. O.; MORITZ, M. O.; MELO, P. A. DE. A Pós-Graduação Brasileira: evolução e principais desafios no ambiente de cenários prospectivos. ANAIS do XI Colóquio Internacional Sobre Gestão Universitária na América do Sul e II Congresso Internacional IGLU. **Anais...** Florianópolis SC: 2011.

NASCIMENTO, L. F. Modelo CAPES de Avaliação: quais as consequências para o triênio 2010-2012?. **Administração: Ensino e Pesquisa**, v. 11, n. 4, p. 579–600, 2010.

NEVES, R.B.; COSTA, H.G. Avaliação de programas de pós-graduação: proposta baseada na integração ELECTRE TRI, SWOT e sistema CAPES. **Revista Eletrônica Sistemas & Gestão**, v.1, n. 3, p. 276-298, 2006.

OLIVEIRA, M. R.; ALMEIDA, J. Programas de pós-graduação interdisciplinares – contexto contradições e limites do processo de avaliação CAPES. **Revista Brasileira de Pós-Graduação**, v. 8, n. 15, p. 37–57, 2011.

SPAGNOLO, Fernando; SOUZA, Valdinei Costa. O Que mudar na avaliação da CAPES? **Revista Brasileira de Pós-graduação**, Brasília, v. 1, n. 2, p. 8-34, nov. 2004.

**Roberto Luiz do Carmo.** Professor Livre Docente do Departamento de Demografia, Instituto de Filosofia e Ciências Humanas da Universidade Estadual de Campinas (IFCH/UNICAMP) e pesquisador do Núcleo de Estudos de População (NEPO/UNICAMP). Diretor-Associado do Instituto de Filosofia e Ciências Humanas da Universidade Estadual de Campinas (IFCH/UNICAMP), 2017-2021. Doutorado em Demografia pela Universidade Estadual de Campinas. Coordenador da Área de Planejamento Urbano e Regional/Demografia da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES), para o período de 2014 a 2018. bobcarmo@uol.com.br

**Eduardo Shimoda.** Doutorado em Produção Animal pela UENF. Coordenou o curso de mestrado em Pesquisa Operacional e Inteligência Computacional (POIC) na UCAM-Campos, onde leciona e orienta na graduação e nos mestrados de POIC e Engenharia de Produção e no mestrado e doutorado em Planejamento Regional / Gestão da Cidade. Atua em pesquisas relacionadas a Indicadores de Qualidade na Educação e nas área de Qualidade em Serviços, Estatística aplicada à educação e à agropecuária. Também é consultor da Capes, tendo atuado nos Grupos de Trabalho "Mestrados Profissionais" e "Produção Técnica", além de ter participado dos comitês de elaboração do documento de área da Produção Técnica e da Avaliação Quadrienal na área de Planejamento Urbano e Regional e Demografia. shimoda@ucam-campos.br

Submetido em: 10/04/2018

Aprovado em: 27/04/2018