Correlação entre Avaliações por Testes de Múltipla Escolha e por Provas Analítico-Expositivas: Crítica e Proposta Metodológica

(Correlation between Evaluations through Multiple-choice Tests and through Analytic-Expositive Exams: Critique and a Methodological Proposal)

Fernando Lang da Silveira
Instituto de Física, Universidade Federal do Rio Grande do Sul
Av. Bento Gonçalves, 9500, Caixa Postal 15051, CEP 91501-970
Porto Alegre. RS. Brasil
Correio eletrônico: Lang@if.ufrgs.br

Trabalho recebido em 17 de janeiro de 1996

O objetivo deste trabalho é fazer uma crítica à metodologia utilizada por Pinho (1995), no estudo da correlação entre as avaliaçõeses por testes de escolha múltipla e por provas analítico-expositivas no concurso vestibular da FUVEST em 1994. É proposta uma metodologia adequada, exemplificando-se com escores obtidos por simulação através do método de Monte Carlo e por vestibulandos da UFRGS em 1995.

Abstract

The purpose of this paper is to criticize the methodoloy used by Pinho (1995) in the study of the correlation between the evaluations through multiple-choice tests and through analytic-expositive exams of the FUVEST Vestibular in 1994. A suitable methodology is proposed and exampled with scores obtained by simulation through the Monte Carlo method and by candidates for admission at UFRGS in 1995.

I. Introdução

A comparação entre os escores obtidos em avaliaçõeses com testes de escolha múltipla e em provas analítico-expositivas é um assunto extremamente importante, tendo em vista que diversas universidades adotam os dois tipos em seus concursos vestibulares. Aqui no Rio Grande do Sul é possível que o concurso vestibular da UFRGS venha também a ser realizado nos moldes do da FUVEST a partir de 1997. Pinho (1995) tinha como objetivo *"comparar o resultado obtido pelos candidatos que lograram matrícula nas vagas disponíveis nos dois tipos de avaliação"* no vestibular da FUVEST de 1994.

O objetivo deste artigo é fazer uma crítica, apontando alguns aspectos que merecem ser revistos no referido trabalho. O autor dessa crítica acredita que os dois tipos de provas possam convergir para os mesmos resultados; apesar disso, discorda, em diversos pontos, da metodologia utilizada por Pinho (1995). Também este trabalho delineia um procedimento para a realização da comparação pretendida. Serão apresentados resultados de simulação numérica pelo método de Monte Carlo com o objetivo de ilustrar as críticas e a metodologia por nós proposta. Finalmente, a metodologia proposta será aplicada aos escores de candidatos do concurso vestibular da UFRGS em 1995.

II. Qualidade do ajuste e correção para acerto casual

No referido trabalho foram utilizados, como escores a serem correlacionados, as médias, por carreira, do número de acertos em provas de escolha múltipla e da nota nas provas analítico-expositivas. O autor houve por bem aplicar uma correção para acerto casual à média nas provas de escolha múltipla. A correção efetivamente utilizada é a dada pela equação 2 em Pinho (1995), que é uma **equação linear**. Como bem apontou o autor, o coeficiente de correlação linear entre duas variáveis mede a qualidade do ajuste da reta dos mínimos quadrados que relaciona uma variável com a outra (reta de regressão). Ele afirmou que a qualidade do ajuste, quando foram utilizadas as médias sem a correção para acerto casual, era muito inferior à qualidade do ajuste com as médias corrigidas. Ora, sabe-se que o coeficiente de correlação linear é **invariante frente a transformações lineares** em qualquer das duas variáveis (Silveira, 1994); transformaçõeses lineares em qualquer das variáveis apenas afetam a declividade da reta de regressão, bem como a constante independente, não tendo qualquer efeito sobre o coeficiente de correlação linear (coeficiente de correlação de Pearson). Não necessitamos calcular os dois coeficientes de correlação para saber que eles são **exatamente** os mesmos (com e sem a correção para acerto casual), ou seja, ***não é verdade*** que a qualidade do ajuste com a correção, dada pela equação 2, seja superior à qualidade do ajuste sem ela.

III. Correlação entre médias

Como já notamos, foram utilizadas médias por carreira nos dois tipos de avaliação como variáveis a serem submetidas à análise de correlação e regressão. Quase todos os coeficientes de correlação encontrados foram elevados (superiores a +0,90). Como se deve interpretar tal resultado ? Significa esse resultado que um particular aluno que possua um alto índice de acertos na prova com itens de múltipla escolha tenha, provavelmente, uma nota elevada na prova analítico-expositiva?

O resultado significa **apenas** que quando cresce a média do índice de acertos de uma carreira para a outra, cresce concomitantemente a nota média na prova analítico-expositiva. Nada se pode inferir sobre a relação entre as duas variáveis dentro de cada carreira, ou seja, não necessariamente o que acontece entre as diferentes carreiras, acontece entre os candidatos que buscam uma carreira. As relações entre coeficientes de correlação obtidos com resultados individuais e com médias podem ser encontradas, por exemplo, em Guilford e Fruchter (1973). Contentamo-nos aqui em apresentar, na Figura 1, os resultados de uma simulação numérica pelo método de Monte Carlo. Os pontos que aparecem na figura simulam escores de indivíduos dentro de cada carreira.

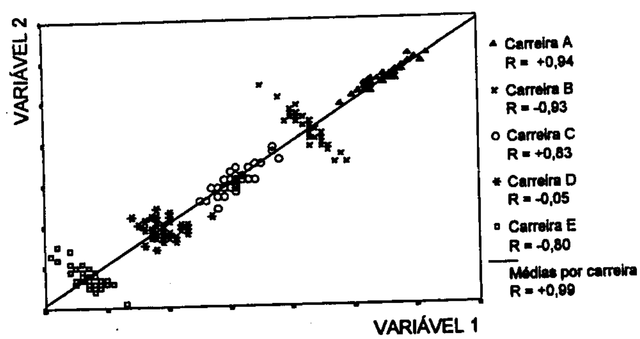

Figura 1. Diagrama de dispersão para as duas variáveis por carreira.

A figura 1 exemplifica diversas possibilidades para a correlação entre as duas variáveis dentro de cada carreira, sendo a correlação entre as médias das carreiras praticamente perfeita (R = +0,99); a reta apresentada na figura é a reta de regressão (reta dos mínimos quadrados) da média por carreira da variável 2 contra a média por carreira da variável 1. Nas carreiras A e C as correlações entre as duas variáveis são elevadas e positivas; dentro dessas duas carreiras a reta de regressão da variável 2 contra a variável 1 é praticamente igual à reta de regressão das médias apresentada na figura. Já nas carreiras B e E as correlações são negativas, significando que o crescimento de uma variável é acompanhado pelo decréscimo da outra. Na carreira D praticamente não há correlação entre as duas variáveis. Fica assim evidenciado que **nada se pode inferir sobre a relação entre os dois tipos de prova dentro de cada carreira, a partir do conhecimento da**

relação entre as médias nas diferentes carreiras. Não é possível repassar a correlação observada entre as médias nos dois tipos de avaliação para os escores individuais dos candidatos.

IV. Probabilidade de aprovação de um candidato

Na parte final de seu artigo, Pinho apresenta *"uma outra maneira de verificar a correlação entre os resultados obtidos nos dois tipos de avaliação"*. Faz um estudo de como varia a probabilidade de aprovação dos candidatos convocados para a segunda fase, com o índice de acertos na prova de escolha múltipla; são apresentados os resultados em apenas quatro das sessenta e nove carreiras. Não explica ele porque apenas as quatro carreiras, referindo que nas demais os resultados *"repetem-se, com pequenas variações"*. Temos razões teóricas para acreditar que tal não acontece em todas as carreiras, principalmente naquelas muito procuradas (mais adiante exporemos tais razões). Porque Pinho não continuou utilizando a mesma abordagem que na etapa anterior ? O cálculo dos coeficientes de correlação entre os escores obtidos, nos dois tipos de prova, pelos candidatos dentro de cada carreira seria o procedimento "natural" para quem utilizou tal abordagem com as médias.

V. Como investigar a correlação entre os dois tipos de avaliações ?

A maneira de investigar a correlação entre os dois tipos de avaliações é calculando o coeficiente de correlação linear de Pearson entre os escores (índice de acertos e nota) dos candidatos classificados para a segunda etapa em cada carreira. A alternativa de investigar a probabilidade de aprovação desses mesmos candidatos, a partir dos índices de acerto na prova com itens de escolha múltipla, é menos recomendável; a desvantagem de tal abordagem está em que a variável "nota do candidato na prova analítico-expositiva" é convertida em uma variável binomial (aprovado ou não-aprovado). Esta conversão determina uma perda de informação: candidatos aprovados não são mais diferenciados entre si; idem para os não- aprovados. Além disso, o objetivo precípuo é investigar a correlação; nada é mais coerente com este objetivo do que a obtenção dos coeficientes de correlação.

A literatura sobre medidas psicológicas e educacionais alerta para um fato importante, relativo à obtenção de coeficientes de correlação em subgrupos de indivíduos com variabilidade (variância) reduzida na variável independente: esses coeficientes serão sistematicamente menores do que aqueles obtidos para o grupo total. Isto acontece quando nos referimos ao grupo dos candidatos selecionados para a segunda etapa do vestibular em relação ao grupo de todos os candidatos. Os classificados para a segunda etapa, dentro de cada carreira, são exatamente aqueles que obtiveram os maiores índices de acerto na prova com itens de escolha múltipla; portanto a variância (quadrado do desvio padrão) do índice de acertos entre os classificados será menor do que para a totalidade dos candidatos.

Idealmente gostaríamos de saber qual seria o coeficiente de correlação entre os escores nos dois tipos de avaliação para a totalidade dos candidatos; esse coeficiente mostraria se escores elevados (baixos) na primeira avaliação estão ou não associados com escores elevados (baixos) na segunda avaliação. Tal é impossível de ser realizado diretamente; somente para os classificados para a segunda etapa é que se conhece a nota na prova analítico- expositiva. Mesmo assim é possível se estimar o coeficiente de correlação para a totalidade dos candidatos, através de uma fórmula conhecida como **"correção para a restrição em variabilidade"** (Ghiselii, 1964; Guilford e Fruchter, 1973) que apresentamos abaixo.

$$(R_{XY})_{\text{est}} = \frac{r_{XY}\Sigma_x}{S_X\sqrt{1 - r^2_{XY}\left(1 - \left[\frac{\Sigma_X}{S_X}\right]^2\right)}} \tag{1}$$

onde:

$(R_{XY})_{\text{est}}$ - coeficiente de correlação estimado para o grupo total.

r_{XY} - coeficiente de correlação calculado no grupo restrito.

Σ_X - desvio padrão da variável independente no grupo total.

S_X - desvio padrão da variável independente no grupo restrito.

É interessante notar que, em decorrência da equação (1), o coeficiente de correlação entre duas variáveis cresce quando aumenta a variabilidade (medida pelo desvio padrão) dos escores. Desta forma, o coeficiente de correlação entre duas variáveis não é uma propriedade exclusiva delas, dependendo também do grupo (amostra) de escores.

Sabe-se que erros de medida em qualquer das duas variáveis afetam o coeficiente de correlação entre elas. A literatura especializada em teoria da medida psicológica e educacional apresenta uma equação, conhecida como **"correção para atenuação"** (Ghiselli,1964; Guilford e Fruchter, 1973), que permite estimar qual seria o coeficiente de correlação entre as duas variáveis caso elas fossem medidas sem erro. Essa estimativa é feita a partir do coeficiente de correlação obtido com as variáveis medidas com erro e dos coeficientes de fidedignidade de cada variável. O coeficiente de fidedignidade *(reliability)* de uma variável é a razão entre a variância verdadeira (variância da variável sem erro) e a variância efetivamente obtida, isto é, variância da variável medida com erro. O coeficiente de fidedignidade tem valor compreendido no intervalo fechado de zero a um, podendo também ser interpretado como uma autocorrelação. Silveira (1981b) mostra que estatísticas utilizadas na comparação entre grupos (razão t de Student e razão F de Snedecor) também são atenuadas quando há erros de medida, apresentando a forma de corrigi-las.

A equação (1), não levando em consideração a existência de erro de medida na variável independente, produz um coeficiente de correlação estimado menor para todo o grupo do que na verdade ele é; havendo erro de medida na variável independente, a equação (1) produzirá subestimativas cada vez mais afastadas do valor real ao decrescer o desvio padrão no grupo restrito. Apresentamos a seguir uma equação inédita na literatura, que leva em conta o erro de medida na variável independente. Omitiremos a dedução dessa equação (assunto de um artigo ainda não publicado), contentando-nos em verificar a sua pertinência, seja em situações simuladas pelo método de Monte Carlo, seja em situações reais com escores de vestibulandos da UFRGS em 1995.

$$(R_{XY})_{\text{est}} = \frac{r_{XY} F_X \Sigma_X}{S_X \sqrt{f_X^2 - r_{XY}^2 \left(f_X - F_X \left[\frac{\Sigma_X}{S_X} \right]^2 \right)}} \qquad (2)$$

onde:

F_X - coeficiente de fidedignidade da variável independente no grupo total.

f_X - coeficiente de fidedignidade da variável independente no grupo restrito.

É fácil verificar que a equação (2) reduz-se à equação (1) quando os coeficientes de fidedignidade são iguais à unidade (quando não há erro na variável independente). Um pouco mais complicado é constatar que a equação (2) produzirá um coeficiente de correlação estimado maior para o grupo total do que o obtido pela equação (1) (isto acontecerá nas simulações apresentadas adiante e também nos estudos com os escores de vestibulandos da UFRGS em 1995).

Cabe ainda notar um importante pressuposto nas deduções das equações (1) e (2): a verdadeira regressão de uma variável sobre a outra no grupo total deve ser linear. Caso a verdadeira regressão não seja linear, as duas equações não permitirão estimar o coeficiente de correlação linear no grupo total a partir do grupo restrito. Ou seja, se houverem razões para acreditar em uma regressão curvilínea no grupo total, as equações (1) e (2) não podem ser utilizadas.

A preocupação com a fidedignidade dos escores obtidos em instrumentos de medida psicológicos e educacionais não pode ser omitida; é comum nesses domínios estarem presentes erros de medida importantes. A forma de se estimar coeficientes de fidedignidade pode ser encontrada na bibliografia especializada em teoria da medida psicológica e educacional; esse assunto é abordado na literatura anteriormente citada e também em Moreira e Silveira (1993) e em Silveira (1981a, 1983, 1985, 1994).

Passamos agora aos resultados de simulações numéricas pelo método de Monte Carlo, com o objetivo de ilustrar os pontos discutidos anteriormente e testar a equação (2).

VI. Simulações pelo método de Monte Carlo

Inicialmente simulamos duas variáveis aleatórias, pelo método de Monte Carlo (Sobol, 1983), com coeficientes de fidedignidade $F_X = F_Y = 0,90$ (esta alta fidedignidade é às vezes alcançada em algumas provas do concurso vestibular da UFRGS) e desvio padrão $\Sigma_X = \Sigma_Y = 10$; foram simulados vinte mil pontos com auxílio do software "SPSS for Windows - versão 6.0" (Norusis, 1993). O **"verdadeiro coeficiente de correlação"** (coeficiente de correlação entre as variáveis sem erro de medida) escolhido para elas foi +0,95; esta alta correlação **verdadeira** é coerente com a hipótese de que essas duas variáveis medem quase o mesmo construto, tal como se admite em provas com itens de escolha múltipla e analítico-expositivas versando sobre o mesmo conteúdo. O coeficiente de correlação entre as duas variáveis com erro estará atenuado em relação ao verdadeiro; um teorema, demonstrável dentro da teoria da medida psicológica e educacional, afirma que **o coeficiente de correlação entre as duas variáveis com erro será igual ao verdadeiro multiplicado pela raiz quadrada do produto dos coeficientes de fidedignidade** (Ghiselli,1964; Guilford e Fruchter, 1973). Esta propriedade que os erros de medida têm de atenuar o coeficiente de correlação, enfatiza a necessidade de haver preocupação com a fidedignidade das variáveis envolvidas em estudos correlacionais. Nessa primeira simulação, apesar da verdadeira correlação entre as variáveis ser +0,95, espera-se, pelo teorema anteriormente referido, um coeficiente de correlação de +0,855 entre elas. Quando efetivamente calculado nos vinte mil pontos simulados, resultou em $R_{XY} = +0,854$.

O próximo passo foi o de selecionar 5% dos pontos gerados, portanto mil pontos, para o grupo restrito; desta forma estamos simulando uma carreira bastante concorrida, em que apenas 1 em cada 20 candidatos é selecionado. Estes pontos foram selecionados construindo-se uma terceira variável, com coeficiente de correlação verdadeiro igual a +0,90 com a variável independente. Os mil escores mais altos desta terceira variável foram utilizados na determinação do grupo restrito. Neste grupo foram calculados o coeficiente de fidedignidade e o desvio padrão da variável independente, bem como o coeficiente de correlação desta com a variável dependente (estatísticas necessárias para se utilizar a equação (2)). Os valores encontrados foram os seguintes:

$$f_X = 0,73$$

$$S_X = 6,11$$

$$r_{XY} = 0,62$$

São dignas de nota as reduções, no grupo restrito em relação ao grupo total, do coeficiente de fidedignidade (0,73 para 0,90) e do coeficiente de correlação (0,62 para 0,85).

Utilizando-se então a equação (2), estimamos o coeficiente de correlação para o grupo total em $(R_{XY})_{\text{est2}} = +0,84$. A estimativa realizada pela equação (1) levou a $(R_{XY})_{\text{est1}} = +0,79$. O coeficiente de correlação para o grupo total, já apresentado, é $R_{XY} = +0,85$.

Uma segunda simulação foi realizada. A única diferença em relação à anterior é que os coeficientes de fidedignidade das duas variáveis, no grupo total, foram menores: $F_X = F_Y = 0,80$. O coeficiente de correlação entre elas no grupo total resultou em $R_{XY} = +0,760$, que é igual ao previsto pelo teorema anteriormente referido. Os valores encontrados no grupo restrito de mil pontos (5%) foram os seguintes:

$$f_X = 0,53$$

$$S_X = 6,53$$

$$r_{XY} = 0,45$$

Novamente, frisamos as reduções, no grupo restrito em relação ao grupo total, do coeficiente de fidedignidade (0,53 para 0,80) e do coeficiente de correlação (0,45 para 0,76). Note-se que agora as reduções são mais pronunciadas do que na simulação anterior; tal se deve a que nessa segunda simulação os erros de medida são maiores do que na primeira.

Utilizando-se então a equação (2), estimamos o coeficiente de correlação para o grupo total em $(R_{XY})_{est2} = +0,74$. A estimativa realizada pela equação (1) levou a $(R_{XY})_{est1} = +0,61$. O coeficiente de correlação para o grupo total, já apresentado, é $R_{XY} = +0,76$.

Uma última simulação foi realizada. A diferença em relação à anterior é que o grupo restrito foi diminuído para quatrocentos pontos (2% dos vinte mil pontos), simulando uma carreira com alta concorrência (cinqüenta candidatos por vaga). Os valores encontrados neste grupo restrito foram:

$$f_X = 0,43$$

$$S_X = 5,90$$

$$r_{XY} = 0,35$$

Também aqui, observamos reduções, maiores ainda, no grupo restrito em relação ao grupo total, do coeficiente de fidedignidade (0,43 para 0,80) e do coeficiente de correlação (0,35 para 0,76).

Utilizando-se então a equação (2) estimamos o coeficiente de correlação para o grupo total em $(R_{XY})_{est2} = +0,74$. A estimativa realizada pela equação (1) levou a $(R_{XY})_{est1} = +0,54$. O coeficiente de correlação para o grupo total, já apresentado, é $(R_{XY}) = +0,76$.

Na tabela 1 são apresentados os resultados das três simulações.

$F_X = F_Y$	R_{XY}	Σ_X	P	f_X	s_X	r_{XY}	$(R_{XY})_{est1}$	$(R_{XY})_{est2}$
0,90	0,85	10	5%	0,73	6,11	0,62	0,79	0,85
0,80	0,76	10	5%	0,53	6,53	0,45	0,61	0,74
0,80	0,76	10	2%	0,43	5,90	0,35	0,54	0,76

TABELA 1 - Resultados das simulações.

F_X, F_Y - coeficientes de fidedignidade de X e Y no grupo total.
R_{XY} - coeficiente de correlação entre X e Y no grupo total.
Σ_X - desvio padrão de X no grupo total.
P - proporção de pontos selecionados para o grupo restrito.
f_X - coeficiente de fidedignidade de X no grupo restrito.
S_X - desvio padrão de X no grupo restrito.
r_{XY} - coeficiente de correlação entre X e Y no grupo restrito.
$(R_{XY})_{est1}$ - coeficiente de correlação estimado pela equação (1).
$(R_{XY})_{est2}$ - coeficiente de correlação estimado pela equação (2).

Observa-se, nas três simulações, que os coeficientes de fidedignidade e os coeficientes de correlação nos grupos restritos são menores do que para o grupo total e menores do que o verdadeiro coeficiente para o grupo total (0,95). Desta forma, mesmo que exista um forte coeficiente de correlação verdadeiro entre duas variáveis (correlação entre as variáveis medidas sem erro), os coeficientes obtidos estarão atenuados (serão menores). Tais atenuações ocorrem quando no grupo restrito houver uma redução no desvio padrão da variável independente em relação ao grupo total, e, serão tanto mais intensas quanto menor forem os coeficientes de fidedignidade (estas afirmações são demostráveis teoricamente e são verificadas empiricamente através das simulações). A equação (2) permite estimar, com grande precisão, o coeficiente de correlação entre as duas variáveis no grupo total; a equação (1), por não levar em conta os erros de medida na variável independente, produz valores bastante subestimados.

Passamos finalmente a apresentar alguns resultados reais com os escores de vestibulandos da UFRGS em 1995.

VII. Estudo com vestibulandos da UFRGS em 1995

Na UFRGS, em 1995, o concurso vestibular era constituído por uma única etapa, contendo uma prova de Redação e oito provas (Biologia, Física, Química, Matemática, História, Geografia, Língua Estrangeira, Literatura) com 35 questões de múltipla escolha. Portanto, não havia condições de estudar a relação entre os dois tipos de avaliações. Com o mero intuito de exemplificar o que foi discutido na secção III, realizamos um estudo correlacional entre os escores de Língua Estrangeira e Redação e entre os escores de Física e Química.

Para as cinqüenta e uma carreiras calculamos as médias dos escores obtidos pelos candidatos classificados (aprovados) nas quatro provas acima referidas; portanto, escores médios em grupos restritos. Em seguida, calculamos os coeficientes de correlação entre estas médias para Língua Estrangeira com Redação e Física com Química; eles resultaram respectivamente em 0,953 e 0,948. Desta forma, foram constatadas altas correlações entre os escores médios dos classificados nas diversas carreiras, tais como as encontradas por Pinho (1995). Calculamos a seguir os coeficientes de correlação entre os dois pares de escores para os classificados dentro de cada carreira, obtendo 51 coeficientes para cada par. Estes resultados estão representados nos histogramas das figuras 2 e 3.

Nota-se, nos dois histogramas, que os coeficientes de correlação dentro das carreiras são muito menores do que entre as médias nas carreiras (0,953 e 0,948). O coeficiente de correlação médio correspondente ao histograma de Redação e Literatura é 0,282 e o correspondente ao de Física e Química é 0,430; observam-se muitos coeficientes de correlação baixos (próximos de zero) dentro das carreiras. Ou seja, *as altas correlações entre as médias das carreiras não podem ser repassadas para os escores dentro de cada carreira;* porque estamos trabalhando com grupos restritos (apenas os classificados), ocorrem coeficientes pequenos. Esses resultados corroboram o que foi afirmado nas secções III e V.

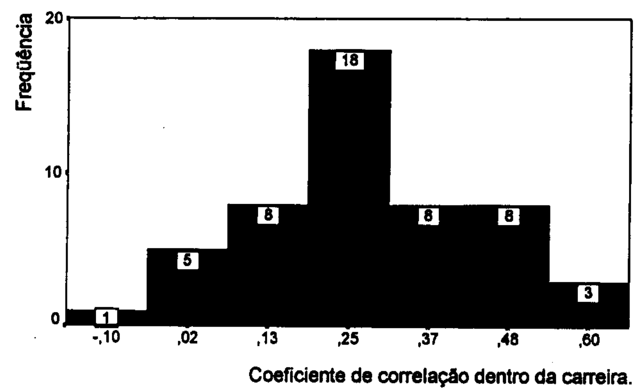

Figura 2. Histograma para os coeficientes de correlação entre os escores de Redação e Língua Estrangeira para os classificados nas 51 carreiras.

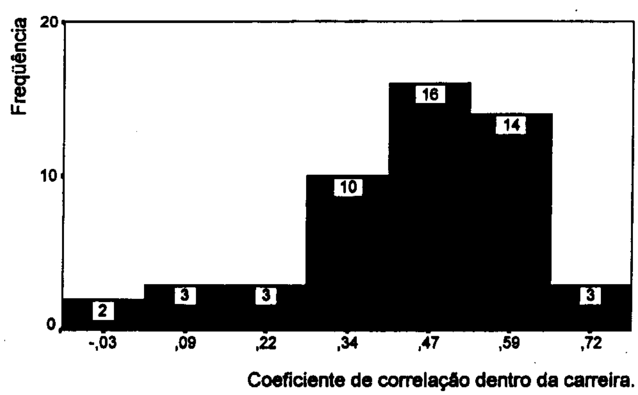

Figura 3. Histograma para os coeficientes de correlação entre os escores de Física e Química para os classificados nas 51 carreiras.

Adicionalmente calculamos os mesmos coeficientes de correlação dentro das diversas carreiras *para todos os candidatos.* Estes estão representados nos histogramas das figuras 4 e 5.

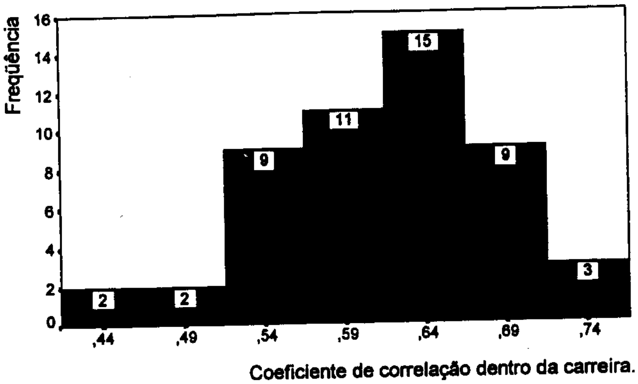

Figura 4. Histograma para os coeficientes de correlação entre os escores de Redação e Língua Estrangeira para todos os candidatos nas 51 carreiras.

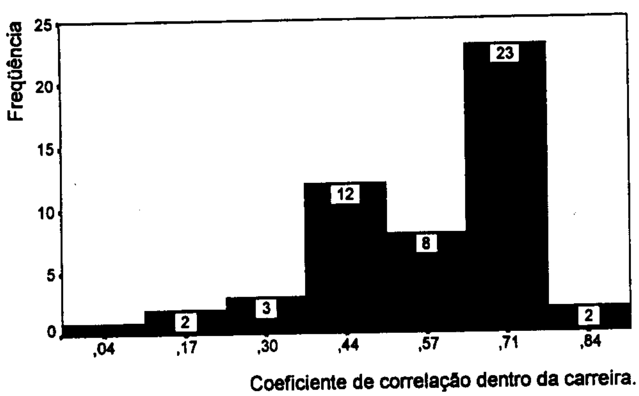

Figura 5. Histograma para os coeficientes de correlação entre os escores de Física e Química para todos os candidatos nas 51 carreiras.

Os histogramas das figuras 4 e 5, quando comparados com os das figuras 2 e 3 respectivamente, mostram coeficientes de correlação mais elevados. Isto é perfeitamente coerente com as previsões teóricas originadas nas considerações de **"restrição para variabilidade"** da secção V. O coeficiente de correlação médio correspondente ao histograma de Redação e Literatura (figura 4) é 0,607 (no histograma da figura 2 era 0,282) e o correspondente ao de Física e Química (na figura 5) é 0,572 (no histograma da figura 3 era 0,430). Apesar das correlações terem aumentado, ainda continuam menores do que 0,953 e 0,948 respectivamente (correlações entre as médias dos classificados(aprovados) por carreira). Todos esses resultados enfatizam que: **1) nada se pode inferir sobre as correlações dentro de cada carreira a partir das correlações entre os escores médios nas careiras; 2) os coeficientes de correlação em grupos restritos são menores (às vezes muito menores) do que no grupo total.**

A fim de testar a equação (2) e exemplificar a metodologia proposta, fizemos mais dois estudos com os escores dos vestibulandos.

O primeiro deles envolveu as provas de Língua Estrangeira (X) e Redação (Y). Escolhemos duas carreiras: Ciências Jurídicas, que apresentava uma taxa alta de 23,4 candidatos por vaga e Geologia, com apenas 2,6 candidatos por vaga.

O segundo estudo envolveu as provas de Física (X) e Química (Y). Escolhemos duas carreiras: Medicina, que apresentava uma taxa alta de 19,3 candidatos por vaga e Física,com apenas 1,9 candidatos por vaga.

Os coeficientes de fidedignidade da variável independente foram estimados através do **"coeficiente beta"** (Silveira, 1985), que é um coeficiente de consistência interna (para maiores detalhes sobre coeficientes de consistência interna pode-se consultar a bibliografia já referida e também o importante artigo de Cronbach (1951)). A tabela 2 apresenta os resultados obtidos.

Verifica-se, nesses dois estudos, que a equação (2) sempre produz estimativas mais próximas do valor real do que a equação (1). Como já notamos teoricamente e nas simulações, repete-se aqui que a equação (1) pode produzir valores muito subestimados para o coeficiente de correlação no grupo total.

Nos dois estudos, os desvios padrão, os coeficientes de fidedignidade e os coeficientes de correlação entre as duas variáveis no grupo de aprovados em carreiras com alta concorrência (Ciências Jurídicas e Medicina) são muito menores do que para todos os candidatos (tal também foi observado na secção VI). Isto sempre tenderá a ocorrer se uma parcela pequena dos candidatos for selecionada, pois haverá nesse grupo uma grande restrição em variabilidade. É por isso que nos surpreende Pinho sempre encontrar *"notável correlação entre os resultados nas provas analítico-expositivas, que determinam a conquista da vaga pelo candidato, e o resultado nos testes de escolha múltipla realizados cerca de um mês antes"* (Pinho, 1995).

Carreira	Σ_X	F_X	s_X	f_X	R_{XY}	r_{XY}	$(R_{XY})_{est1}$	$(R_{XY})_{est2}$
C. Jurídicas	7,02	0,87	3,01	0,22	0,70	0,22	0,25	0,70
Geologia	4,98	0,68	4,37	0,57	0,51	0,42	0,47	0,53
[redacted]								
Medicina	7,57	0,90	2,79	0,11	0,88	0,22	0,52	0,91
Física	5,17	0,72	4,78	0,65	0,63	0,56	0,59	0,63

Tabela 2 - Resultados para vestibulandos da UFRGS em 1995.

Σ_X - desvio padrão de X no grupo de todos os candidatos.
F_X - coeficiente de fidedignidade (beta) no grupo de todos os candidatos.
S_X - desvio padrão de X no grupo dos candidatos classificados.
f_X - coeficiente de fidedignidade (beta) no grupo dos candidatos classificados.
R_{XY} - coeficiente de correlação no grupo de todos os candidatos.
r_{XY} - coeficiente de correlação no grupo dos candidatos classificados.
$(R_{XY})_{\text{est1}}$ - coeficiente de correlação estimado pela equação (1).
$(R_{XY})_{\text{est2}}$ - coeficiente de correlação estimado pela equação (2).

VIII. Conclusão

Nesse trabalho, além de criticarmos uma metodologia utilizada para estudar a existência ou não de relação entre as provas analítico-expositivas e por testes de múltipla escolha no concurso vestibular à universidade, apresentamos uma proposta metodológica adequada. Mostramos como inferir sobre o coeficiente de correlação entre duas variáveis em um grupo todo a partir do coeficiente de correlação em um grupo restrito, apresentando uma equação inédita que leva em consideração a fidedignidade da variável independente. Certamente esta metodologia pode ser aplicada em outras pesquisas que envolvam restrição de variabilidade em uma das variáveis.

Ainda que se verifique a suposta convergência entre os dois tipos de avaliações no concurso vestibular, justificar-se-iam as provas analítico-expositivas, do estrito ponto de vista da teoria da medida educacional (podem haver outras justificativas que se apóiam em considerações que não dizem respeito à teoria da medida), se fosse constatado empiricamente que elas são mais fidedignas do que as de escolha múltipla: já que os candidatos devem ser selecionados, deseja-se que a medida utilizada como critério de seleção seja a mais fidedigna possível.

Agradecimentos

À professora Maria Cristina Varriale pela leitura crítica deste trabalho; ao árbitro da RBEF por suges tões que permitiram aprimorá-lo.

Referências

1. L. J. Cronbach, Coefficient Alpha and the internal structures of tests. Psychometrika, **16**, 297 (1951).
2. J. P. Guilford e B. Fruchter, *Fundamental statistics in psychology and education.* New York: McGraw-Hill, 1973.
3. E. E. Ghiselli, *Theory of psychological measurement.* New Delhi: Tata McGraw-Hill, 1964.
4. M. A. Moreira e F. L. Silveira, *Instrumentos de pesquisa em ensino e aprendizagem.* Porto Alegre: EDIPUCRS, 1993.
5. M. J. Norusis, SPSS for Windows, Release 6.0. Chicago: SPSS Inc., 1993.
6. A. G. Pinho, Correlação entre avaliações por testes de múltipla escolha e por provas analítico-

expositivas. Revista Brasileira de Ensino de Física, São Paulo, **17**(2), 169 (1995).

7. F. L. Silveira, Relação entre os índices de discriminação de itens em testes psicométricos e duas outras estatísticas: variância do escore total e coeficiente de fidedignidade. Ciência e Cultura, São Paulo, **33**(2), 256 (1981a.)
8. F. L. Silveira, Fidedignidade das medidas e diferenças entre grupos em psicologia e educação. Ciência e Cultura, São Paulo, **33**(5), 704 (1981b).
9. F. L. Silveira, Considerações sobre os índices de discriminação de itens em testes educacionais. Educação e Seleção, São Paulo, **8**, 54 (1983).
10. F. L. Silveira, Coeficiente beta: estimativa do coeficiente de fidedignidade de uma variável compósita. Educação e Seleção, São Paulo, **11**, 105 (1985).
11. F. L. Silveira, Validação de testes de papel e lápis. In: MOREIRA, M. A. (org.) *Atas da II Escola Latinoamericana sobre Pesquisa em Ensino de Física.* Porto Alegre: Instituto de Física da UFRGS, 1994.
12. I. Sobol, *O método de Monte Carlo.* Moscou: MIR, 1983.