

ANÁLISE DO SISTEMA DE VESTIBULAR DA UNIVERSIDADE FEDERAL DO CEARÁ¹

Raimundo Alberto Normando*
Alberto Flávio Alves Aguiar**
Mauro Villar de Queiroz***

INTRODUÇÃO

A quase totalidade das instituições de ensino superior brasileiras adota, em seus concursos vestibulares, quesitos de quatro ou cinco alternativas, acarretando tal prática um acerto casual de vinte e cinco ou vinte por cento, respectivamente.

A partir do segundo concurso vestibular de 1977, a Universidade Federal do Ceará começou a se preocupar com essa inconveniência dos sistemas tradicionais e passou a estudar um novo tipo de provas em que se reduzisse, tanto quanto possível, o acerto por mera questão de sorte. Em julho de 1978, foi implantado o atual sistema de provas, que consta de quesitos com 100 alternativas, ao invés das 4 ou 5 opções do sistema habitualmente em vigor em quase todas as outras universidades, com exceção da Universidade de Brasília, Universidade Federal da Bahia e Universidade Federal de Santa Catarina. O primeiro resultado prático foi a drástica redução do acerto aleatório para apenas 1%.

O presente trabalho é o primeiro passo no sentido de construir indicadores adequados à análise do novo tipo de questões. A inexistência de bibliografia específica atrasou a avaliação do vestibular desenvolvido na Universidade Federal do Ceará. Por outro lado, somente com a implantação do atual computador tornou-se exequível o processamento da enorme massa de dados provenientes do atual sistema de provas em vigor na UFC.

(¹) Trabalho apresentado no VIII SAESUNN – Seminário de Acesso ao Ensino Superior das Instituições do Norte-Nordeste. 23 a 26 de agosto de 1983. Maceió. Sob patrocínio da Universidade Federal de Alagoas – CESMAC – Escola de Ciências Médicas de Alagoas.

* Pró-Reitor de Graduação da UFC

** Coordenador de Controle de Sistemas da PR-GR da UFC

*** Presidente da Comissão Coordenadora do Vestibular da UFC

TIPO DE QUESTÕES¹

O sistema compõe-se de dois tipos de questões: abertas e de proposições múltiplas. As questões abertas, mais adequadas para as provas de Matemática, Física, Química e alguns conteúdos de Biologia, são, na verdade, problemas que admitem por resposta um número inteiro pertencente ao intervalo [0,99]. O aluno resolve a questão e marca o resultado diretamente no cartão-resposta. As questões de proposições múltiplas constam de, no máximo, 7 proposições, das quais algumas são verdadeiras e outras falsas, admitindo-se, também, que todas possam ser falsas ou verdadeiras.

As proposições são numeradas segundo a progressão geométrica 1, 2, 4, 8, 16, 32 e 64, sendo a resposta correta, por definição, a soma dos números associados às proposições verdadeiras. É importante salientar que qualquer resposta é obtida de uma e apenas uma maneira, como combinação linear dos elementos da progressão geométrica.

Quer para as questões abertas, quer para as de proposições múltiplas, a delimitação das respostas ao conjunto {00, 01, 02, ..., 99} prende-se exclusivamente ao *lay-out* do cartão-resposta. Na verdade, esta restrição não causa qualquer embaraço, porque, quando o resultado de um problema for uma fração, poderá ser solicitado o produto daquele resultado por um número que seja um múltiplo conveniente do denominador, de forma a que se obtenha uma resposta inteira para o problema. Analogamente, quando o resultado de um quesito for, por exemplo, uma raiz quadrada, poderá ser pedido o quadrado do valor encontrado e não o próprio valor.

INDICADORES

O índice de facilidade de um item é definido tradicionalmente por $I_F = \frac{A}{N}$, onde A é o número de acertos no item e N o número de candidatos.

Em uma prova de 4 alternativas em que os candidatos ignorassem totalmente o conteúdo, o índice de facilidade de cada quesito seria aproximadamente 0,25.

Em qualquer prova objetiva, I_F está inflacionado por um contingente de candidatos que chegam à resposta correta de modo puramente casual.

No sistema em vigor na UFC o acerto casual se situa em torno de 0,01, valor 25 vezes menor do que o correspondente no sistema de 4 opções.

Aplicadas as provas, utilizando-se o novo sistema de quesitos, o índice de acerto caiu significativamente, apesar de a Universidade ter mantido o mesmo nível na elaboração dos quesitos.

Cabe fazermos a seguinte indagação: poderiam os candidatos ter tido uma queda instantânea no nível de sua competência? A resposta é não, inquestionavelmente.

A queda acentuada do número de acertos deve-se principalmente ao fato de os candidatos, no novo sistema, não terem os "benefícios" do acerto casual na mesma proporção em que o tinham anteriormente. Além disso, como nos sistemas tradicionais só há uma alternativa correta, o simples fato de o candidato identificar uma opção errada aumenta sua chance de acerto para 33%, enquanto a certeza de que duas alternativas são incorretas amplia a possibilidade de acerto para 50%.

Ainda dentro da mesma linha de raciocínio, os professores, mesmo com a assistência de técnicos em medidas educacionais, dificilmente conseguirão elaborar 3 alternativas incorretas tão solicitantes quanto a opção certa, o que leva o candidato a encontrar a resposta adequada, muito mais por uma questão de bom senso do que por competência.

Todos esses fatos aumentam o acerto casual para níveis superiores a 25%, no caso de quesitos com 4 alternativas.

¹ NORMANDO, R.A.; PORTO, V.C. & QUEIROZ, M.V. de. (1981) *O Vestibular na Universidade Federal do Ceará: novo modelo*. Educação e Seleção (3): 85-92, jan/jul.

É oportuno salientarmos que nada mudou em relação ao processo de solução das questões. Por exemplo, se em uma prova de Matemática há uma questão em que se solicita ao 8º termo de uma progressão geométrica, o candidato deverá usar a expressão $a_8 = a_1 q^7$, calcular o termo pedido e examinar qual das 4 alternativas coincide com o valor encontrado.

No sistema da UFC o aluno calcula o 8º termo da progressão e marca o valor encontrado no lugar adequado do cartão-resposta.

Se, anteriormente, no sistema tradicional, as respostas estavam sendo produzidas conscientemente, a porcentagem de acerto nas questões não deveria ter caído. No entanto, caiu e a queda se deve tanto ao fato de o novo sistema não permitir pistas para inferir a resposta, quanto, principalmente, ao fato de, antes, o acerto com certa margem de inconsciência estar alto demais.

Em suma, antes o aluno procurava "descobrir" a resposta e agora ele tem de produzi-la.

SIMULAÇÃO

Algumas vezes, foi questionada a possibilidade de um candidato obter resultado nulo em um quesito, do qual havia sido capaz de identificar o valor verdadeiro de quase todas as proposições, errando apenas uma delas.

No trabalho antes mencionado, mostramos que esse fato não é específico do sistema da UFC, porque ocorre nos sistemas tradicionais de questões objetivas, qualquer que seja o número de opções por quesito.

Para avaliar, nessas circunstâncias, em que grau poderiam os candidatos estar sendo prejudicados, globalmente, procedemos a uma correção simulada, levando-se em conta cada uma das proposições que o aluno acertasse e atribuindo-se ao quesito um valor proporcional ao número de acertos de proposições. Evidentemente, modificaram-se os escores brutos e, com eles, a média, o desvio padrão e o escore padronizado.

Corrigidas as provas, o computador emitiu a listagem de classificação simulada para cada curso. A Tabela 1 mostra os coeficientes de correlação ordinal (ρ) para os cursos nos vestibulares 82/1, 82/2 e 83/1. Os coeficientes médios de correlação para 82/1, 82/2 e 83/1 são, respectivamente, 0,81, 0,84 e 0,82, com desvios-padrão respectivamente iguais a 0,08, 0,06 e 0,08. É importante ressaltar que o coeficiente médio de correlação é praticamente constante, não havendo, nesse particular, diferença significativa, quer entre o primeiro e o segundo vestibular, quer entre o vestibular de um ano e o do ano seguinte, ocorrendo o mesmo em relação ao desvio padrão.

Os coeficientes de correlação são, em geral, bastante grandes e, para os cursos com alta relação candidato/vaga, ultrapassam a 0,90.

Esses resultados mostram a improcedência da argumentação de que os candidatos poderiam estar sendo prejudicados, globalmente, porque, de cada questão, sabiam quase tudo e, por não saberem tudo, obtinham zero em todos os quesitos.

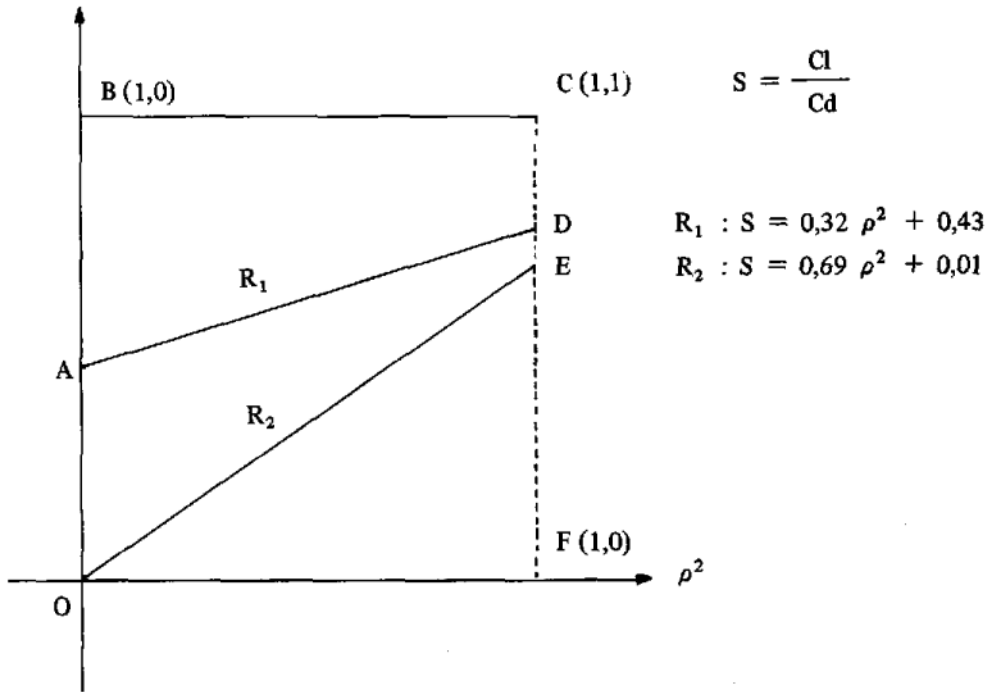
De posse dos dados contidos na Tabela 1 e da relação classificáveis/candidados $\left(\frac{Cl}{Cd}\right)$ para cada curso, encontramos a reta de regressão entre $\frac{Cl}{Cd}$ e ρ^2 para os vestibulares 82/1, 82/2 e 83/1. A reta relativa ao vestibular 82/1 é praticamente coincidente com aquela representativa do concurso de 83/1. Para o segundo vestibular de 1982, obtivemos uma reta bastante diferente (Fig. 1).

TABELA 1

Coeficientes de Correlação

Curso	82/1	82/2	83/1
Administração	0,79	—	0,79
Agronomia	0,88	0,88	0,85
Arquitetura	0,94	0,92	0,92
Biblioteconomia	0,69	0,80	0,69
Licenciatura em Ciências	0,70	—	—
Ciências Biológicas	0,76	0,81	0,78
Ciências Econômicas	0,84	0,85	0,85
Ciências Contábeis	0,78	0,81	0,77
Ciências Sociais	0,77	—	0,70
Comunicação Social	0,77	—	0,79
Direito	0,79	0,83	0,81
Economia Doméstica	0,80	0,80	0,72
Enfermagem	0,80	0,82	0,85
Engenharia Civil	0,95	0,92	0,92
Engenharia Elétrica	0,91	—	0,91
Engenharia Mecânica	0,89	—	0,91
Engenharia Química	0,91	—	0,90
Estatística	0,79	0,85	0,78
Farmácia	0,84	0,82	0,83
Física	0,88	0,87	0,88
Geologia	0,77	0,85	0,81
Geografia	0,66	0,82	0,70
História	—	0,79	—
Letras	0,73	0,74	0,79
Matemática	0,76	0,94	0,86
Medicina	0,92	0,93	0,92
Odontologia	0,89	0,90	0,91
Pedagogia	0,69	0,68	0,72
Psicologia	0,82	0,85	0,84
Química	0,80	—	0,74
Química Industrial	0,84	0,82	0,86
Engenharia de Pesca	0,88	0,82	0,82
Processamento de Dados	0,91	—	0,92
Tecnologia de Alimentos	0,73	0,75	0,78

Figura 1



Do gráfico, vemos que, para qualquer valor de ρ^2 , o valor de S para o primeiro vestibular (S_1) é maior do que para o segundo (S_2). Daí o valor médio de S_1 ser maior que o de S_2 .

Como $S_1 > S_2$, temos que $\frac{Cl_1}{Cd_1} > \frac{Cl_2}{Cd_2}$

O número de candidatos no primeiro vestibular (Cd_1) é sempre maior que no segundo (Cd_2), isto é $Cd_1 > Cd_2$.

Então, o número de classificáveis no primeiro vestibular é maior que no segundo, ou seja, $Cl_1 > Cl_2$.

Ora, S pode mudar, tanto por uma variação em Cl , quanto em Cd .

Para iguais variações em ambos os parâmetros, a de S é mais afetada pela do denominador, porque a variação em S é inversamente proporcional ao quadrado de Cd e apenas proporcional a Cl .

Como Cd_1 é da ordem de 50% maior que Cd_2 , temos que $Cl_1 > Cl_2$.

Concluímos, pois, que o desempenho dos candidatos de janeiro é melhor que o dos candidatos de julho.

A reta de regressão para o primeiro vestibular é menos inclinada do que a reta relativa ao segundo, porque, conforme já vimos, embora os coeficientes médios de correlação entre S e ρ^2 , para os dois concursos, sejam da mesma ordem de grandeza, S para o primeiro vestibular é muito maior do que para o vestibular de julho.

Na figura, o segmento BC representa a situação ideal, pois isto significa que nenhum candidato teria sido eliminado com zero.

Para efeito de análise, vamos associar a área unitária do quadrado $OBCF$ à situação ideal. Assim, as áreas $ABCD$ e $OBCE$ são medidas que estimam a diferença entre a situação ideal e a encontrada no vestibular de janeiro e de julho, respectivamente.

Vamos, agora, calcular a diferença entre a situação ideal e a real para ambos os vestibulares. Temos que $\delta_1 = 0,41$ e $\delta_2 = 0,65$. A razão $\frac{\delta_2}{\delta_1}$ mede o grau de decaimento de desempenho dos candidatos, o qual é da ordem de 60%.

Este resultado não recomenda às instituições de ensino superior fazerem um único vestibular no início do ano para classificar os primeiros colocados para o ingresso em março e "guardar" os subsequentes para agosto, porque, com a matrícula já assegurada para o segundo semestre, o futuro aluno, em geral, não manterá o mesmo interesse pelo estudo e até poderá desaprender alguns conteúdos.

O segmento \overline{OF} representa a situação em que todos os candidatos estão eliminados com zero, situação exatamente oposta àquela representada pela área $OBCF$.

À medida que cai o desempenho dos candidatos, a reta de regressão se afasta cada vez mais do segmento \overline{BC} e a área que representa a diferença entre o desempenho ideal e o real tende para 1.

DISTRIBUIÇÃO DE FREQUÊNCIAS

Seja N o número de quesitos da prova, n o número de alternativas de cada uma das questões e i o número de quesitos que o candidato acerta.

A probabilidade do acerto aleatório $P_n(i)$ é dada por $P_n(i) = \frac{(n-1)^{N-i}}{n^N} \cdot C_i^N$ onde C_i^N é o número de combinações de N elementos i a i .

Vamos considerar apenas os casos em que $n = 4$ e $n = 100$, calcular a probabilidade aleatória de acertar todos os quesitos e de errar todos eles, fixando $N = 30$ que é o número de quesitos de cada prova no sistema em vigor na UFC.

$$\text{Assim, } P_4(30) = \left(\frac{1}{4}\right)^{30}$$

$$P_4(0) = \left(\frac{3}{4}\right)^{30}$$

$$P_{100}(30) = \left(\frac{1}{100}\right)^{30}$$

$$P_{100}(0) = \left(\frac{99}{100}\right)^{30}$$

As Figuras 2 e 3, respectivamente para 4 e 100 alternativas, representam a distribuição de candidatos em função do número de quesitos respondidos corretamente, na hipótese de 20.000 candidatos. As curvas interrompidas traduzem o comportamento puramente aleatório, enquanto as cheias representam o desempenho real.

Figura 2

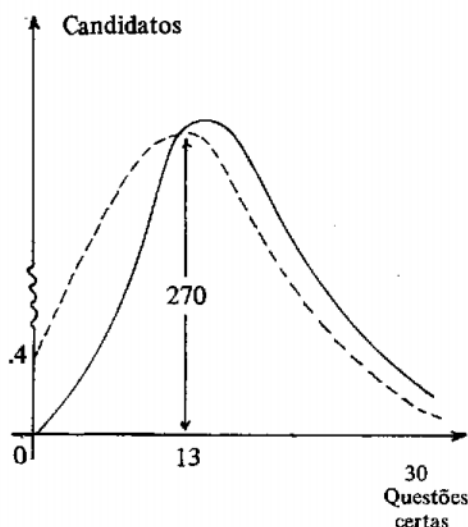
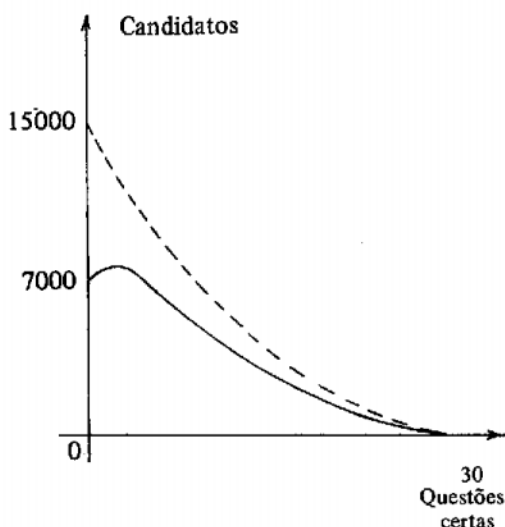


Figura 3



Ora, os candidatos se preparam para o vestibular. Por isso, o desempenho real é diferente daquele obtido em função apenas da aleatoriedade. Na situação real, há uma diminuição do número de zeros e elevação do número de acertos, embora esse aumento seja pouco significativo, em virtude da chance de acertar todas as questões ser muito baixa.

Em qualquer situação, a curva tende a zero para o acerto total, porém muito mais rapidamente para o sistema de 100 alternativas do que para o de 4, quer no modelo aleatório, quer na realidade do vestibular, produzindo uma assimetria de tal ordem que nos impede de considerar a massa inferior com o mesmo número de indivíduos da massa superior.

Em suma, a discrepância entre a curva obtida e a curva normal não nos autoriza a desprezar os 46% de candidatos que apresentam desempenho em torno da média no sistema tradicional.

ÍNDICE DE DISCRIMINAÇÃO

Partindo-se da hipótese de uma distribuição normal nos sistemas de 4 ou 5 alternativas, é comum dividirmos o grupo total de candidatos em três subgrupos, isto é, a massa superior (S), que é integrada por 27% dos candidatos que têm melhor desempenho, a massa inferior (I), que congrega 27% dos candidatos que apresentam pior desempenho e a massa central, que consta justamente dos restantes 46% de candidatos cujo desempenho se situa em torno do médio.

Abandonamos a massa média e calculamos o índice de facilidade para as massas superior e inferior, I_{FS} e I_{FI} , respectivamente. Definimos, então, o índice de discriminação $I_D = I_{FS} - I_{FI}$. O índice de dificuldade do item é aproximadamente a média aritmética entre I_{FS} e I_{FI} , isto é,

$$I_F \cong \frac{I_{FS} + I_{FI}}{2}$$

No sistema de 100 alternativas, é grande o número de candidatos que têm nota zero em um quesito. Por isso, a curva de distribuição de notas é fortemente assimétrica. Por essa razão, não podemos separar o grupo em três subconjuntos, dos quais os conjuntos extremos são compostos por igual número de candidatos.

Tal fato nos impede de calcular o índice de discriminação, usando a definição convencional. Vamos então definir um novo índice, capaz de medir a fuga dos candidatos à resposta

correta e a intensidade com que eles se dispersam pelo conjunto das 100 alternativas da questão. A tal indicador chamamos de índice de espalhamento.

ÍNDICE DE ESPALHAMENTO

O índice de espalhamento de um quesito é definido por

$$I_E = \frac{\sum_{i=0}^{2^N-1} \left[\frac{|X_i \cup X_k| - |X_i \cap X_k|}{N} P_i \right]^{P_k}}{2^N - 1}, \text{ onde } P_i \text{ é a porcentagem de}$$

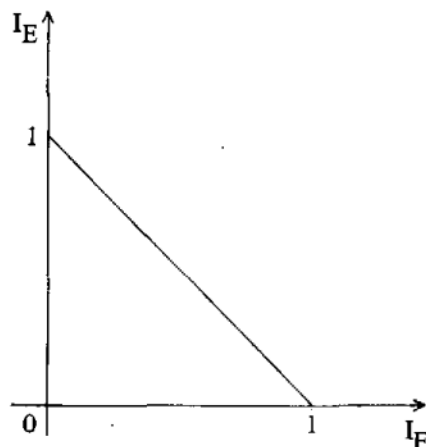
candidatos que marcou, incorretamente, i como resposta do quesito; P_k é a porcentagem de candidatos que marcou, corretamente, k como resposta do quesito; N é o número de proposições do quesito; X_i é o conjunto de elementos da progressão geométrica, com os quais é formado o número i , $i = 0, 1, 2, \dots, 2^N - 1$ com $i \neq k$; X_k é o conjunto de elementos da progressão geométrica com os quais é formado o número k (resposta do item); $|X_i \cup X_k|$ é o número de elementos da união $\{X_i\} \cup \{X_k\}$; $|X_i \cap X_k|$ é o número de elementos da intersecção $\{X_i\} \cap \{X_k\}$.

Vemos, pois, que o índice de espalhamento pode depender de um máximo de 201 variáveis, isto é, cada um dos 99 P_i ($i = 0, 1, 2, \dots, k - 1, k + 1, \dots, 2^N - 1$), P_k , cada um dos 100 elementos $\frac{|X_i \cup X_k| - |X_i \cap X_k|}{N}$, com k fixado, e mais a variável N .

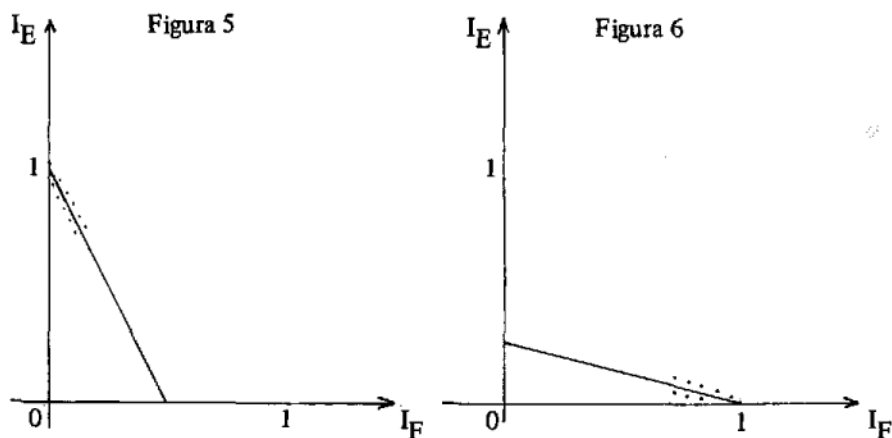
É evidente que quando $N = 7$, o valor de $2^N - 1$ é maior que 99. Por isso, na hipótese de $N \neq 7$, o denominador de I_E deve ser considerado igual a 99. De fato, na definição de I_E , incluímos o denominador com o único objetivo de normalizar o índice, situando-o entre 0 e 1, isto é, $0 \leq I_E \leq 1$.

Da própria definição do índice de espalhamento, podemos ver que I_E é tanto menor, quanto maior for a facilidade do quesito e vice-versa. Assim, para uma prova homogênea deveremos ter poucas questões com I_E próximo de zero, poucas questões com I_E próximo de um e a maioria dos quesitos com I_E compreendido entre 0,40 e 0,60. Nessas circunstâncias, a curva I_E versus I_F é, razoavelmente, a reta vista na Figura 4.

Figura 4



Se a maioria dos quesitos tem I_F próximos de 0 ou próximos de 1, as retas de regressão são aquelas vistas nas Figuras 5 e 6, respectivamente.



Estas duas situações são indesejáveis, porque caracterizam provas muito difíceis ou muito fáceis.

Na UFC, obtivemos as seguintes retas de regressão para os vestibulares 82/1, 82/2 e 83/1.

$$82/1: I_E = -2,30I_F + 0,84$$

$$82/2: I_E = -2,96I_F + 0,92$$

$$83/1: I_E = -2,79I_F + 0,87.$$

Calculando-se a média aritmética dos elementos da reta para os vestibulares de julho, temos:

$$1^{\circ} \text{ Vestibular: } I_E = -2,55I_F + 0,86$$

$$2^{\circ} \text{ Vestibular: } I_E = -2,96I_F + 0,92$$

Assim, a reta relativa ao segundo vestibular está mais afastada da reta padrão do que a reta representativa do primeiro. Como o nível das questões foi mantido, é imperioso concluir que o grupo de julho teve desempenho inferior ao grupo de janeiro. A razão desse resultado é o fato de a clientela do segundo vestibular ser o resíduo do grupo que tenta ingressar no início do ano.

ÍNDICE DE EQUÍLBRIO

Vamos agora definir um novo índice com o objetivo de medir o grau de adequação da prova ao grupo.

Vamos chamá-lo índice de equilíbrio e representá-lo por I_e .

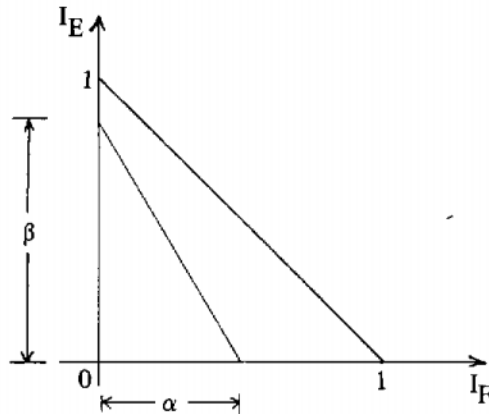
O desvio da reta em relação à reta padrão é medido pelo fato de tanto a abscissa à origem quanto a ordenada à origem serem diferentes de 1.

$\alpha - \beta$ é uma medida de assimetria, ou seja, do afastamento da reta real em relação à reta padrão cuja equação é $I_E = I_F + 1$. Vamos definir o índice de equilíbrio por $I_e = \alpha^2 - \beta^2$.

Podemos ver que I_e é proporcional à assimetria e proporcional, também, à média aritmética de α e β .

Para provas muito fáceis, α tende a 1 e β tende a zero. Assim, I_e tende a 1. Para provas muito difíceis, vemos que I_e tende a -1 . Para uma prova idealmente adequada ao grupo, $I_e = 0$.

Figura 7



CONCLUSÕES

O acerto casual no sistema da UFC é consideravelmente menor do que nos sistemas tradicionais de quatro ou cinco opções.

O sistema não permite a inferência da resposta, principalmente para as questões abertas.

Os altos índices de correlação encontrados indicam que é diminuto o número de candidatos que obtêm resultado nulo por errarem apenas uma das proposições.

O desempenho dos candidatos de janeiro é indiscutivelmente superior ao dos candidatos de julho.

A curva de distribuição de frequência é sensivelmente diferente da curva normal, em virtude da alta porcentagem de resultados nulos, oriundos da diminuição acentuada do acerto aleatório.

O sistema permite definir índices específicos, o índice de espalhamento (I_E) e o índice de equilíbrio (I_e), o primeiro para medir a fuga dos candidatos à resposta correta e o segundo para avaliar o grau de adequação da prova ao grupo.