

**REVISTA DE
EMPREENDEDORISMO,
NEGÓCIOS E INOVAÇÃO**

ISSN 2448-3664

**Marcos Ricardo Rosa
Georges**

Centro de Economia e
Administração. Pontifícia
Universidade Católica de
Campinas.

UNIVERSIDADE FEDERAL DO ABC
AVENIDA DOS ESTADOS, 5001
BAIRRO BANGU, SANTO ANDRÉ - SP.
CEP 09210-580

E-MAIL: RENI@UFABC.EDU.BR

COORDENAÇÃO
AGÊNCIA DE INOVAÇÃO INOVAUFABC



AMOSTRAGEM ESTRATIFICADA E DISTRIBUIÇÃO MULTINOMIAL NA PESQUISA DE OPINIÃO: APLICAÇÃO COM ALUNOS DE ADMINISTRAÇÃO

STRATIFIED SAMPLING AND MULTINOMIAL DISTRIBUTION IN OPINION RESEARCH: APPLICATION WITH ADMINISTRATIVE STUDENTS

RESUMO

Este artigo apresenta uma pesquisa de opinião feita com alunos de administração elaborada com o objetivo de conhecer a opinião dos alunos sobre infra-estrutura, professores, expectativas futuras e sobre seu envolvimento nas atividades de ensino. O principal interesse do artigo reside na definição do tamanho da amostra usando a técnica de amostragem estratificada para se proceder a pesquisa de opinião, bem como a presença de uma distribuição multinomial. Os estratos foram concebidos a partir da linha de formação do curso de administração e do período oferecido, constituindo oito diferentes estratos na população objeto. Já a distribuição multinomial surge em função de perguntas com três alternativas de respostas. A presença de uma distribuição multinomial e o uso de estratificação na amostragem exigem a aplicação de um procedimento específico para determinação do tamanho da amostra. Detalhes do cálculo do tamanho da amostra são mostrados no artigo, bem como a apresentação dos resultados que revelam a opinião dos alunos.

Palavras-Chave: Ensino em Administração, Pesquisa de Opinião, Metodologia de Pesquisa, Amostragem Estratificada, Distribuição Multinomial.

ABSTRACT

This article presents a survey carried out with business students developed in order to know the opinion of students on infrastructure, teachers, future expectations about their involvement in teaching. The main interest of this article is the sample size definition and the use of stratified sampling technique to make the survey, as well as the presence of a multinomial distribution. The strata were designed by course and the period, representing eight different strata in the population object. Already a multinomial distribution arises due to questions with three possible answers. The presence of a multinomial distribution and use of stratification sample technique requiring the application of a specific procedure for determining the sample size. Details of the calculation of sample size are shown in the article, as well as the presentation of results that reveal the students' opinion.

Keywords: Business Teaching, Opinion Poll, Research Methodology, Stratified Sampling, Multinomial Distribution.

JEL Classification: I21

1. INTRODUÇÃO

A pesquisa de opinião é um instrumento amplamente usado por empresas e indivíduos. Em empresas, é comum o uso de pesquisas de opinião para se saber o quão satisfeito estão seus clientes, para se levantar opiniões dos funcionários sobre o clima organizacional e em diversas outras situações de natureza mercadológica onde se deseja conhecer a opinião dos clientes em relação a características de um produto a ser lançado.

Já estudantes e pesquisadores em administração usam com frequência a pesquisa de opinião como instrumento de coleta de dados nos estudos e pesquisas em administração quando se deseja levantar a opinião de pessoas, gestores ou clientes sobre algum fenômeno de interesse. Porém, nenhuma outra situação é tão emblemática para representar o uso da pesquisa de opinião como nas pesquisas de intenção de voto que antecedem as eleições.

Pesquisas de opinião são feitas para representar as opiniões de uma população fazendo-se uma série de perguntas a um pequeno número de pessoas e então extrapolando as respostas para um grupo maior dentro do intervalo de confiança. É um levantamento estatístico realizado sobre uma amostra da população.

Numa perspectiva histórica, o primeiro exemplo de pesquisa de opinião que se tem conhecimento foi uma enquete conduzida pelo *The Harrisburg Pennsylvanian* em 1824, que indicava o candidato Andrew Jackson à frente de John Quincy Adams por 335 a 169 votos, na disputa pela presidência dos Estados Unidos (MOON, 1999).

Mesmo sem rigor científico e sem permitir qualquer generalização mais precisa que a mera adivinhação, este tipo de enquete tornava-se cada vez mais popular a cada nova eleição presidencial norte-americana.

Em 1916, o *Literary Digest* iniciou uma pesquisa nacional postando milhões de cartões postais aos seus leitores que, após indicarem seu candidato, os remetiam de volta. Contando os cartões que retornaram,

o *Literary Digest* previu corretamente o vencedores na eleição daquele ano e também nas quatro eleições presidenciais seguintes. Woodrow Wilson em 1916; Warren Harding em 1920; Calvin Coolidge em 1924; Herbert Hoover em 1929, e Franklin Roosevelt em 1932 foram os presidentes americanos corretamente previstos pela *Literary Digest* usando este sistema simples de enquete com base na contagem de cartões remetidos de volta a redação (MOON, 1999).

No entanto, este tipo de pesquisa de opinião pública sem rigor científico estava prestes a revelar-se traiçoeiro e, em 1936, a *Literary Digest* protagonizou um dos maiores fiascos da história ao prever, erroneamente, o candidato vitorioso daquele ano (MADONNA e YOUNG, 2002),

Segundo Bryson (1976) e Squire (1988), em 1936 o *Literary Digest* foi surpreendido! Seus 2,3 milhões de “eleitores” constituíam uma vasta amostra, porém, eram pessoas de maior poder aquisitivo e que tendiam a ser simpatizantes do Partido Republicano. O *Literary Digest* nada fez para corrigir esse viés e divulgaram que Alf Landon seria o candidato vitorioso. Mas quem ganhou a eleição foi Franklin D. Roosevelt.

Para tornar completo o fiasco da previsão da *Literary Digest*, George Gallup efetuou uma pesquisa para a mesma eleição, usando uma amostra muito menor, porém com base científica. A pesquisa de Gallup apontou corretamente a vitória de Roosevelt. Foi o início da bancarrota da *Literary Digest* e o início da era das pesquisas de opinião com base estatística!

Segundo Bryson (1976), o sucesso na pesquisa de Gallup para a eleição de 1936 foi a seleção de uma amostra que representasse a distribuição demográfica dos eleitores segundo sua orientação política.

Atualmente, no âmbito da política, a pesquisa de opinião é amplamente utilizada por todos os países democráticos no mundo. Para além da política, a pesquisa de opinião também se tornou um instrumento valioso para as organizações, seja na pesquisa de satisfação dos clientes ou em pesquisas de mercado.

O embasamento científico na realização de uma pesquisa de opinião é respaldado na estatística que, no século XX, evoluiu sobremaneira a ponto de estabelecer diversas técnicas de amostragem e tratamento dos dados para assegurar uma previsão confiável e representativa da opinião de toda a população que se queira saber.

2. PESQUISA, AMOSTRAGEM E DISTRIBUIÇÕES

O objetivo de uma pesquisa de opinião, ou simplesmente pesquisa, é conhecer a opinião da população a respeito de algum fenômeno. No entanto, o custo e o tempo necessário para perguntar a todos os indivíduos da população a sua opinião tornam os censos inviáveis, sendo necessário o questionamento a somente uma parcela da população. Esta parcela é chamada de amostra.

Existem diferentes meios de se estabelecer o tamanho da amostra e o campo da estatística que reúne estes métodos é chamado de amostragem. Basicamente há duas formas distintas de se estabelecer a

amostra de uma pesquisa: probabilística e não-probabilista.

Uma amostra não-probabilística não faz uso de procedimentos estatísticos que garantirão a escolha de indivíduos de forma aleatória. A escolha dos indivíduos da amostra pode ser por conveniência ou por interesse do pesquisador. Já a amostra probabilística, os critérios de inclusão dos indivíduos na amostra são aleatórios e não estão sujeitos ao julgamento e nem a vontade do pesquisador.

A figura 1 a seguir apresenta um esquema que resume as diferentes técnicas de amostragem. Cada método de amostragem traz vantagens e desvantagens em relação aos demais. No entanto, em termos de precisão e confiabilidade, os métodos de amostragem probabilística são mais vantajosos que os métodos de amostragem não-probabilísticas.

Na categoria dos métodos probabilísticos, a amostragem estratificada traz vantagens em relação a amostragem aleatória simples e amostragem sistemática por permitir uma pesquisa com mesma margem de erro usando-se uma amostra menor (COCHRAN, 1977). Na verdade,

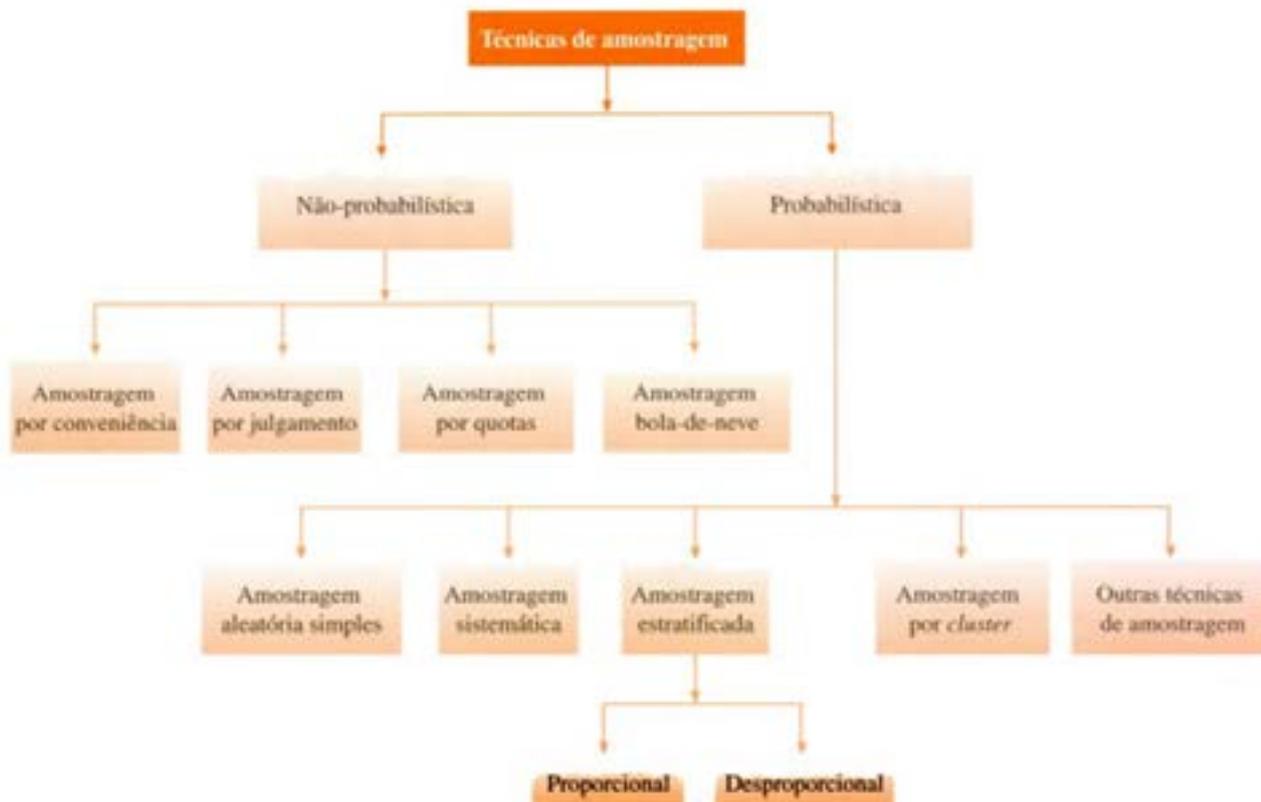


Figura 1 - Classificação das Técnicas de Amostragem (fonte: Malhotra, 2006)

quanto maior a discrepância do comportamento de cada estrato, menor será o tamanho da amostra. Esta característica pode ser útil se bem manipulada.

A amostragem estratificada pode ser aplicada para se realizar uma pesquisa quando, na população objeto, pode ser observada a existência de estratos, ou seja, de grupos bem determinados que particionam a população objeto em grupos (estratos) distintos (DOWNING & CLARK, 2002).

Portanto, a técnica de amostragem adotada influenciará o tamanho da amostra necessária para se realizar a pesquisa e, dentre as técnicas de amostragem probabilística, a amostragem estratificada possibilita um tamanho de amostra menor para uma mesma margem de erro se comparada a uma amostra determinada por outras técnicas.

Mas, além da técnica de amostragem adotada, outra característica fundamental e determinante para se definir o tamanho da amostra é a distribuição de probabilidade do fenômeno a ser pesquisado.

O tipo mais simples de distribuição de probabilidade é a distribuição de Bernoulli, aquela que define somente duas possíveis respostas para uma pergunta, como sim e não. Uma amostra constituída de diversos questionários com distribuição de Bernoulli (duas alternativas de resposta) terá distribuição binomial. Outra distribuição de probabilidade bem conhecida é a distribuição normal, uma distribuição contínua e que representa uma quantidade enorme de fenômenos físicos na natureza.

Na maioria absoluta da literatura científica sobre pesquisa de opinião, pesquisa de marketing e até mesmo em livros específicos sobre estatística e amostragem, as formulas para cálculo do tamanho da amostra são definidas para uma distribuição normal ou para uma distribuição de binomial.

Logo, a determinação do tamanho da amostra só será preciso se a pesquisa for elaborada para se ter apenas respostas dicotômicas (sim e não) ou para se medir variáveis contínuas normalmente distribuídas. No entanto, o que se observa com frequência são pesquisas com questionários

com escalas de mensuração discretas e com mais de duas alternativas de resposta, constituindo assim, uma distribuição multinomial.

Portanto, determinar uma amostra a partir de uma distribuição multinomial e fazendo-se uso da técnica de amostragem estratificada é o desafio que motivou este artigo e espera-se que a pesquisa exemplo apresentada a seguir possa servir de referência para os administradores que queiram realizar suas pesquisas, mas que esbarram na dificuldade de somente encontrar formulas de determinação de tamanho de amostra que consideram apenas distribuições normais ou binomiais.

3. A PESQUISA DE OPINIÃO COM ALUNOS DE ADMINISTRAÇÃO: USO DE AMOSTRA ESTRATIFICADA COM DISTRIBUIÇÃO MULTINOMIAL

A pesquisa de opinião realizada neste artigo é bem simples e tinha como objetivo conhecer o quão satisfeito os alunos estão com a instituição, com seus professores e sua expectativa em relação ao mercado de trabalho. Também há questões que buscam conhecer o perfil do aluno. São perguntas simples e as respostas são múltipla escolha, com três alternativas (sim, não, em parte) e com duas alternativas (sim e não). O uso da metodologia de amostragem estratificada com a presença de uma distribuição multinomial é um dos objetivos da pesquisa, cuja metodologia está detalhada a seguir.

A pesquisa tem como **população objeto** todos os alunos do curso de Administração. Essa população pode ser dividida em diversos agrupamentos, como o período que o aluno cursa (matutino ou noturno), o ano que ele está (do primeiro ao quarto ano), a modalidade de curso que o aluno está matriculado (administração, comercio exterior ou logística e serviços) e ainda se o aluno pertence ao novo currículo ou se pertence ao currículo em extinção.

Estas diversas categorias que a **população objeto** pode ser dividida

constituem em estratos. Um estrato é uma partição da população objeto, os estratos são mutuamente excludentes e exaustivos, ou seja, os elementos a serem amostrados (unidades de observação) pertencem a um, e somente um dos estratos e não há unidade de observação que não pertença a um estrato (os estratos cobrem toda a população objeto).

A pesquisa será aplicada em uma amostra da população objeto. Esta amostra da população objeto é chamada de população estudo que será determinada através de um procedimento de pesquisa por amostragem estratificada. Embora a pesquisa por amostragem estratificada seja mais complexa e exige maior conhecimento da população objeto, a sua aplicação trás maiores vantagens em oposição à técnica de amostragem aleatória simples.

A tabela 1 a seguir apresenta os estratos da população objeto, são eles: ADM – alunos que cursam Administração de Empresas (currículo novo); COMEX – alunos do curso de Comércio Exterior; LOG – alunos do curso de Logística e Serviços e ADM EXT – alunos do curso de Administração de Empresas (currículo em extinção).

Tabela 1 – População Objeto e seus Estratos

ESTRATO	Matutino	Noturno	TOTAL
ADM	283	981	1264
COMEX	148	255	403
LOG	57	233	290
ADM EXT	86	473	559
TOTAL	574	1942	2516

Muitos outros estratos naturais poderiam ser definidos, como série, sexo, religião, entre outros, mas a partição da tabela 1 já fornece oito estratos bem definidos e o suficiente para aplicar a metodologia pretendida.

A teoria de amostragem nos diz que a técnica de amostragem estratificada trás maior precisão quanto maior for à heterogeneidade entre os estratos. Ou seja, quanto maior for a variabilidade e o valor médio entre os estratos maior será a precisão do resultado final em oposição ao uso da técnica de amostragem aleatória simples

(COCHRAN, 1977).

Porém, como não se tem nenhuma informação sobre a distribuição nos estratos, foi realizada uma amostra piloto para se conhecer a distribuição dentro dos estratos para depois realizar a amostragem definitiva.

3.1. REALIZAÇÃO DA AMOSTRA PILOTO

Como não se tem nenhuma informação de como é a distribuição das respostas para as perguntas da pesquisa foi necessário realizar uma amostra piloto. De posse das respostas da amostra piloto será possível determinar o número da amostra definitiva com base na distribuição das respostas obtidas na amostra piloto.

Porém, a pesquisa é constituída de perguntas que têm 2 e 3 alternativas (veja o questionário na figura 2 a seguir) e sabe-se que, quanto maior a heterogeneidade entre os estratos, menor será o tamanho da amostra.

Dessa forma, a determinação do tamanho da amostra piloto será feita levando-se em consideração o pior caso, ou seja, quando há a mínima heterogeneidade entre as repostas, e também se levando em consideração o fato de se ter respostas com distribuição binomial e multinomial.

3.1.1. TAMANHO DA AMOSTRA USANDO DISTRIBUIÇÃO BINOMIAL

A determinação do tamanho da amostra é determinada por diversos fatores, entre eles: a distribuição do parâmetro de interesse, o nível de confiança do intervalo de confiança e a margem de erro, e também do tipo de informação que se deseja determinar.

No caso da avaliação institucional o parâmetro de interesse é a proporção de alunos que responderam sim, não ou em parte (ou outra resposta equivalente).

Observa-se que o questionário a ser respondido é constituído de nove perguntas, sendo as cinco primeiras com respostas tricotômicas e as quatro últimas dicotômicas, e são as cinco primeiras questões onde reside o maior interesse desta pesquisa.

É interessante observar que na amostra piloto não se tem nenhuma informação sobre a variabilidade dos estratos, portanto para a amostra piloto o procedimento adotado será o pior caso para uma amostra aleatória simples. Para estimar a proporção na amostra aleatória simples a técnica de amostragem fornece a seguinte expressão matemática para o cálculo do tamanho da amostra:

Equação 1 – determinação do tamanho da amostra aleatória simples para proporção

$$n = \frac{\frac{z^2 PQ}{d^2}}{1 + \frac{1}{N} \left(\frac{z^2 PQ}{d^2} - 1 \right)}$$

Figura 2 – Questionário da pesquisa realizada

1. Como você avalia a infra-estrutura da faculdade?

Adequada Parcialmente adequada Não adequada

2. Como você avalia a competência dos professores?

Adequada Parcialmente adequada Não adequada

3. Como você avalia o conteúdo das disciplinas ministradas na grade curricular do curso?

Adequada Parcialmente adequada Não adequada

4. Como você avalia a expectativa do mercado de trabalho em relação a um profissional formado pela poi esta faculdade?

Alta Média Baixa

5. Como você avalia o seu envolvimento nas atividades acadêmico propostas pelos professores?

Alta Médio Baixo

6. Você tem computador?

Sim Não

7. Você vem para faculdade com automóvel próprio?

Sim Não

8. Você está empregado atualmente?

Sim Não

9. Você é fumante?

Sim Não

Esta equação 1 pode ser simplificada se chamarmos o numerador de n_0 e reescrever a equação, o resultado final será:

$$n = \frac{n_0}{1 + \frac{1}{N}(n_0 - 1)}, \text{ onde } n_0 = \frac{z^2 PQ}{d^2}$$

Equação 2 – simplificação da equação 1

Nesta equação z representa o nível de confiança do intervalo de confiança a ser construído para a estimativa de interesse, d representa a margem de erro e PQ representam a proporção de alunos que responderam sim e não respectivamente.

Como não se tem qualquer informação sobre a distribuição do parâmetro de interesse, portanto será assumido o pior caso, quando há um equilíbrio na preferência entre as respostas (lembre-se que são respostas dicotômicas: sim e não), ou seja quando há uma intenção de 50% para a resposta sim e 50% para a resposta não. Esta distribuição equilibrada fornece o pior caso em termos de número de unidades amostrais a ser pesquisada, isto pode ser visto claramente no gráfico 1 abaixo.

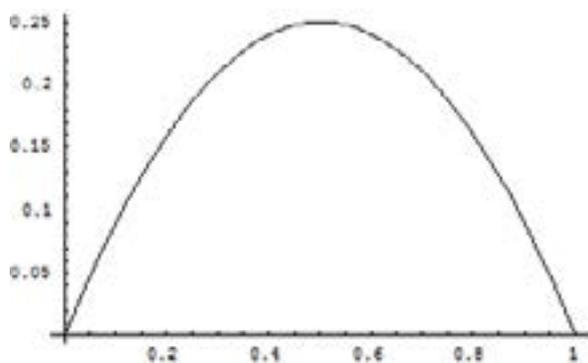


Gráfico 1 – valor de $p.(1-p)$

O nível de confiança para construir o intervalo de confiança para a estimativa de interesse será fixada arbitrariamente em 95%. Observe que quanto maior será o nível de confiança maior será o número de unidades amostrais a ser pesquisado.

A margem de erro será definida pela coordenação do curso de Administração, que teve como alternativas de escolha as margens

de erro mostradas na tabela 2 a seguir.

Tabela 2 – Margem de Erro e conseqüente tamanho da amostra

Margem de erro	Tamanho da amostra
1	1993.87
1.5	1583.18
2	1228.83
2.5	954.23
3	749.52
3.5	597.92
4	484.79
4.5	399.18
5	333.39
6	241.29
7	181.90
8	141.67
9	113.27
10	92.54

O cálculo para o tamanho da amostra apresentado nesta tabela 2 foi obtido da equação 1, adotando $P = Q = 0,5$; $z = 1,96$; $N = 2516$. A margem de erro escolhida foi de 7%, o que resulta em um tamanho de amostra igual a 182 alunos para o pior caso (quando $P = Q = 0,5$). Espera-se que a proporção entre P e Q seja mais desigual, reduzindo o número de unidades amostrais a serem pesquisadas.

3.1.2. TAMANHO DA AMOSTRA USANDO DISTRIBUIÇÃO MULTINOMIAL

Um método bastante simples para se estimar o número de amostras necessárias para estimar a proporção em uma distribuição multinomial com um determinado nível de confiança e margem de erro é proposto por Thompson (1987).

Este método determina o tamanho da amostra para o pior caso, análogo ao feito no item anterior, mas aqui é para uma distribuição multinomial e não para uma distribuição binomial.

O método consiste em determinar o nível de confiança, que neste caso é de 95%, e depois basta olhar em uma tabela e encontrar o número correspondente e dividir este número pelo quadrado da margem de erro. O interessante deste método é que o tamanho

da amostra para o pior caso não depende do número de categorias da distribuição multinomial, seja para três (como o caso desta pesquisa) ou um n qualquer, o pior caso será o mesmo.

Nota-se também que para o pior caso não se necessita nenhuma informação a priori do parâmetro de interesse, assim como no item anterior, o pior caso ocorre quando há um equilíbrio entre as respostas.

Tabela 3 - Pior Caso para Tamanho da Amostra

a	d2n
0.50	.44129
0.40	.50729
0.30	.60123
0.20	.74739
0.10	1.00635
0.05	1.27359
0.25	1.55963
0.01	1.96986

Para o caso desta pesquisa o $a = 0.05$ e portando $d2n = 1.27359$. Dividindo-se $d2n$ por $(0.07)^2$, que foi a margem de erro escolhida obtém: $n = d2n / (0.07)^2 = 1.27359/0.0049 = 259.91$ ou 260. Logo o tamanho da amostra para o pior caso será de 260 alunos.

3.2. TAMANHO DA AMOSTRA PILOTO EM CADA ESTRATO

A Amostra Piloto tem a finalidade de obter uma estimativa aproximada para a proporção de alunos que responderam cada alternativa, para então determinar o tamanho definitivo para a amostra.

Para a amostra piloto será adotado arbitrariamente o valor de 40% do número total de unidades amostrais do pior caso para a distribuição multinomial, ou seja, $0.4 * 260 = 104$ alunos. Assim a amostra piloto será constituída de 104 alunos sorteados aleatoriamente e em número proporcional de seus estratos, ou seja, para a amostra piloto será adotado o procedimento de **amostragem estratificada proporcional**.

A tabela 4 a seguir indica a quantidade de alunos a serem amostrados de cada um dos

seis estratos baseados na regra de proporção.

Tabela 4 – unidades amostrais para a amostra piloto

	Matutino	Noturno	
ADM	11.7	40.6	52.2
COMEX	6.1	10.5	16.7
LOG	2.4	9.6	12.0
ADM EXT	3.6	19.6	23.1
TOTAL	23.7	80.3	104.0

Arredondando-se os valores obtêm-se novos valores para a quantidade de unidades amostrais a serem pesquisadas na amostra piloto, estes valores estão apresentados na tabela 5 a seguir.

Tabela 5 – unidades amostrais definitivas para a amostra piloto

	Matutino	Noturno	TOTAL
ADM	12	41	53
COMEX	7	11	18
LOG	3	10	13
ADM EXT	4	20	24
TOTAL	26	82	108

3.3. O PROCEDIMENTO DE SORTEIO

O procedimento de sorteio das unidades amostrais dentre todas as unidades de observação foi feita de posse do cadastro dos alunos fornecida pela secretaria de graduação da faculdade de Administração. Este cadastro contém as seguintes informações: RA do aluno; nome completo; curso (ADM, ADM EXT, LOG, COMEX); série (1, 2,..., 8); classe (1, 2, 3,...); turno (1 ou 3) e situação acadêmica (matriculado, trancado, afastado, etc.).

De posse da relação completa dos alunos o primeiro passo foi introduzir esta listagem no software Minitab, e estratificar esta listagem por curso e por turno, resultando uma listagem completa dos alunos para cada um dos seis estratos definidos inicialmente.

Para cada um dos seis estratos foi adotado o procedimento de amostragem aleatória simples para sortear as unidades amostrais, este procedimento consistem em gerar números aleatórios com distribuição uniforme entre 0 e 1 e depois ordenar a relação dos alunos em ordem crescente e

entrevistar o número desejado de alunos a partir desta lista.

De posse da lista dos alunos sorteados em cada um dos estratos para amostra piloto o passo seguinte foi identificar qual é a sala de cada aluno. Cada estrato é constituído de diversas salas de aulas, pois são quatro anos de curso e algumas salas por ano, assim, um estrato é constituído por mais de uma sala, e um aluno sorteado para responder ao questionário deve ser encontrado em sua sala. Prevendo-se algumas ausências dos alunos no ato da entrevista, a lista dos alunos sorteados contém mais alunos do que a amostra piloto necessita. Caso um aluno falte, o aluno seguinte da lista será chamado.

3.4. O RESULTADO DA AMOSTRA PILOTO

O resultado da amostra é de extrema valia por fornecer alguns parâmetros iniciais, especialmente a distribuição da característica de interesse. Também se apresenta como resultado imediato da amostra piloto o reconhecimento de dificuldades encontrada durante o procedimento de coleta de dados.

Por ora, somente a distribuição da característica de interesse será mencionada,

pois o principal objetivo é o cálculo do tamanho final da amostra.

A tabela 6 a seguir revela os resultados obtidos da amostra piloto sem distingui-los em estratos. Para cada uma das oito questões é calculado a proporção de respostas sim, +/- e não. A resposta sim equivale a resposta Adequada ou Alta do questionário, a resposta +/- equivale às respostas parcialmente adequadas ou médias, e a resposta não equivale às respostas não adequadas ou baixas do questionário.

Tabela 6 - Proporções das respostas na Amostra Piloto

	P (SIM)	P (+/-)	P (NÃO)
Q1	0.240741	0.712963	0.046296
Q2	0.259259	0.685185	0.055556
Q3	0.259259	0.62963	0.111111
Q4	0.351852	0.601852	0.046296
Q5	0.231481	0.657407	0.111111
Q6	0.925926	-	0.074074
Q7	0.490741	-	0.509259
Q8	0.685185	-	0.314815
Q9	0.185185	-	0.814815

A tabela 7 a seguir mostra os resultados da amostra piloto estratificada para cada um dos 8 estratos.

Tabela 7 - Respostas da Amostra Piloto

	Estrato 1 - Adm. Mat.			Estrato 2 - Comex Mat.			Estrato 3 - Log. Mat.			Estrato 4 - Adm. Ext. Mat.		
	Sim	+/-	não	sim	+/-	não	sim	+/-	não	sim	+/-	não
Q1	4	8	0	1	5	1	2	1	0	0	3	1
Q2	3	8	1	2	5	0	2	1	0	0	4	0
Q3	2	7	3	1	4	2	1	2	0	0	3	1
Q4	5	7	0	1	6	0	1	2	0	0	3	1
Q5	3	6	3	1	4	2	1	2	0	1	3	0
Q6	12	0	0	5	2	2	3	0	0	4	0	0
Q7	8	4	2	2	5	3	3	0	1	1	1	3
Q8	6	6	0	0	7	2	2	1	1	1	1	3
Q9	5	7	0	0	7	0	0	3	3	3	0	1
	Estrato 5 - Adm. Not.			Estrato 6 - Comex Not.			Estrato 7 - Log. Not			Estrato 8 - Adm. Ext. Not.		
	sim	+/-	não	sim	+/-	não	sim	+/-	não	sim	+/-	não
Q1	12	28	1	0	10	1	4	6	0	3	16	1
Q2	14	26	1	1	10	0	4	6	0	2	14	4
Q3	13	26	2	2	9	0	6	4	0	3	13	4
Q4	14	25	2	6	5	0	8	2	0	3	15	2
Q5	9	30	2	3	7	1	4	6	0	3	13	4
Q6	38	3	0	11	0	0	7	3	3	20	0	0
Q7	21	20	5	5	6	3	3	7	7	10	10	10
Q8	32	9	9	9	2	8	8	2	2	16	4	4
Q9	7	34	1	10	0	0	0	10	4	4	16	16

Observa-se na tabela 6 que a resposta que obteve o maior grau de equilíbrio na opinião dos alunos foi a questão nº 7, que pergunta se o aluno vem para a faculdade com automóvel próprio. Mas são as cinco primeiras questões onde reside o maior interesse, então o número definitivo da amostra final deverá ser calculado em função das cinco primeiras questões.

3.5. DETERMINAÇÃO DA AMOSTRA DEFINITIVA

A determinação do tamanho da amostra definitiva será feita em função de apenas uma das cinco primeiras questões, exatamente a questão que apresentar o maior equilíbrio entre as respostas dadas pelos alunos. Isto se justifica em função de quanto mais equilibradas forem as proporções, maior será o tamanho da amostra. Isto é claramente observado para a distribuição binomial e Thompson (1987) afirma que o mesmo ocorre com distribuições multinomiais.

4. O MÉTODO DE CÁLCULO DE N

Este método é extraído de Thompson

(1987) e determina o tamanho da amostra para distribuições multinomiais com k categorias e com informações prévias sobre a distribuição de cada categoria. É um método iterativo composto de duas etapas onde se procura satisfazer as equações $\Theta_i = 2(1 - \Phi(z_i))$, e $z_i = d_i(n)^{1/2} / (p_i(1 - p_i))^{1/2}$,

onde d_i é margem de erro para a i -ésima proporção e $\Phi(z_i)$ é a distribuição normal padrão acumulada.

Para o caso deste trabalho a margem de erro é de 7% para todas as proporções e o nível de confiança é de 5%. A princípio espera-se que o maior n será encontrado para a questão número 3, pois apresenta o maior equilíbrio. Mas para a surpresa o cálculo foi efetuado para todas as questões e foi para a quarta questão que resultou no maior número da amostra.

Na tabela 8 abaixo estão todos os n 's obtidos para cada uma das cinco questões tricotômicas, todos calculados pelo método descrito acima; também estão nesta tabela 8 os tamanhos da amostra para as quatro questões dicotômicas também presentes no questionário, mas estas calculadas segundo o método da distribuição binomial.

Tabela 8 - Cálculo de n para cada uma das questões

	P (SIM)	P (+/-)	P (NÃO)	n
Q1	0.240741	0.712963	0.046296	199
Q2	0.259259	0.685185	0.055556	210
Q3	0.259259	0.62963	0.111111	221
Q4	0.351852	0.601852	0.046296	240
Q5	0.231481	0.657407	0.111111	210
Q6	0.925926	-	0.074074	53
Q7	0.490741	-	0.509259	182
Q8	0.685185	-	0.314815	159
Q9	0.185185	-	0.814815	114

Logo o tamanho da amostra a ser usado neste trabalho será de 240 alunos.

4.1. O CÁLCULO DA AMOSTRA PARA CADA ESTRATO

Para calcular o tamanho da amostra para cada um dos oito estratos será utilizado o procedimento de partilha ótima, ou partilha de Neyman (COCHRAN, 1977). Este procedimento requer a informação de quantos alunos pertencem a cada um dos oito estratos e também a variabilidade de cada um dos oito estratos.

A variabilidade de cada um dos oito estratos será calculada a partir da questão nº4, justamente a questão que forneceu o pior caso para o tamanho da amostra. A variabilidade é calculada através do software Minitab, através do comando Display Descriptive Statistics, e os resultados estão expostos a seguir.

resultado piloto (estrato = 1)

Variable	Mean	StDev
Q1	1.333	0.492
Q2	1.167	0.577
Q3	0.917	0.669
Q4	1.417	0.515
Q5	1.000	0.739

resultado piloto (estrato = 2)

Variable	Mean	StDev
Q1	1.000	0.577
Q2	1.286	0.488
Q3	0.857	0.690
Q4	1.143	0.378
Q5	0.857	0.690

resultado piloto (estrato = 3)

Variable	Mean	StDev
Q1	1.667	0.577
Q2	1.667	0.577
Q3	1.333	0.577
Q4	1.333	0.577
Q5	1.333	0.577

resultado piloto (estrato = 4)

Variable	Mean	StDev
Q1	0.750	0.500
Q2	1.0000	0.000
Q3	0.750	0.500
Q4	0.750	0.500
Q5	1.250	0.500

resultado piloto (estrato = 5)

Variable	Mean	StDev
Q1	1.2683	0.501
Q2	1.3171	0.521
Q3	1.2683	0.548
Q4	1.2927	0.558
Q5	1.1707	0.495

resultado piloto (estrato = 6)

Variable	Mean	StDev
Q1	0.9091	0.301
Q2	1.0909	0.301
Q3	1.182	0.405
Q4	1.545	0.522
Q5	1.182	0.603

resultado piloto (estrato = 7)

Variable	Mean	StDev
Q1	1.400	0.516
Q2	1.400	0.516
Q3	1.600	0.516
Q4	1.800	0.422
Q5	1.400	0.516

resultado piloto (estrato = 8)

Variable	Mean	StDev
Q1	1.100	0.447
Q2	0.900	0.553
Q3	0.950	0.605
Q4	1.050	0.510
Q5	0.950	0.605

Observando somente a questão nº 04 obtêm-se as seguintes variâncias:

Tabela 9 - Variâncias Amostrais dos Estratos

S_1	S_2	S_3	S_4	S_5	S_6	S_7	S_8
0,2652	0,1439	0,3330	0,2500	0,3114	0,2725	0,1781	0,2601

O peso de cada um dos estratos é dado por:

Tabela 10 - Peso de Cada Estrato

W_1	W_2	W_3	W_4	W_5	W_6	W_7	W_8	SW
0.1125	0.0588	0.0227	0.0342	0.3899	0.1013	0.0926	0.1880	1.000

O produto da variância e do peso fornece:

Tabela 11 - Produto do Peso e Variância Amostral de Cada Estrato

S_1W_1	S_2W_2	S_3W_3	S_4W_4	S_5W_5	S_6W_6	S_7W_7	S_8W_8	SSW
0.0298	0.0085	0.0076	0.0086	0.1214	0.0276	0.0165	0.0489	0.2688

Assim, para cada estrato o tamanho da amostra será dado por:

Tabela 12 - Tamanho da Amostra para cada Estrato

n_1	n_2	n_3	n_4	n_5	n_6	n_7	n_8	S
26.64	7.55	6.75	7.63	108.40	24.65	14.72	43.66	240
27	8	7	8	109	25	15	44	243

Mas como os questionários da amostra piloto serão aproveitados, então a amostra definitiva para cada estrato será dada por:

Tabela 13 - Tamanho da Amostra Final para cada Estrato

n_1	n_2	n_3	n_4	n_5	n_6	n_7	n_8	S
15	1	4	4	68	14	5	24	135

4.2. O PROCEDIMENTO DE SORTEIO

O procedimento de sorteio será o mesmo para definir os alunos sorteados para a amostra piloto. Ou seja, de posse do cadastro é feito um procedimento de geração de números aleatórios com distribuição uniforme (0,1) e depois estes números são ordenados. Os alunos associados aos primeiros números serão escolhidos para responder o questionário.

5. O RESULTADO DA PESQUISA

O resultado da pesquisa está apresentado na tabela 14 a seguir e revela que, com relação a infra-estrutura da faculdade, 25,5% dos alunos consideram-na adequada, 66,7% consideram parcialmente adequada e 7,8% consideram a infra-estrutura

inadequada.

Com relação a competência de seus professores, 25,9% dos alunos a considera adequada, 63,8% consideram parcialmente adequada e 10,3% consideram inadequada.

Quanto ao conteúdo ministrado nas disciplinas, 21,4% consideram adequado; 65,8% consideram parcialmente adequado e 12,8% consideram inadequado. Com relação a expectativa do mercado, 40,3% tem alta expectativa, 56,4% tem uma expectativa parcial e 3,3% não tem expectativas quanto ao mercado de trabalho após a conclusão do curso.

Quanto ao envolvimento nas atividades propostas pelos professores, 18,5% dos alunos consideram que se envolvem adequadamente nas atividades, 67,5% tem envolvimento parcial e 14% não se envolvem

adequadamente.

Mais além, 81,4% dos alunos possuem computadores, 51,9% vão a faculdade com automóvel próprio, 67,5% dos alunos estão empregados atualmente e 18,1% são fumantes.

Tabela 14 - resultados finais da pesquisa

	P_{SIM}	$P_{+/-}$	$P_{NÃO}$
Q1	0.255	0.667	0.078
Q2	0.259	0.638	0.103
Q3	0.214	0.658	0.128
Q4	0.403	0.564	0.033
Q5	0.185	0.675	0.140
Q6	0.814	-	0.186
Q7	0.519	-	0.481
Q8	0.675	-	0.325
Q9	0.181	-	0.819

6. CONCLUSÃO E COMENTÁRIOS FINAIS

Este artigo fez uma pesquisa de opinião com alunos de administração usando um questionário com três respostas, necessitando usar um procedimento mais elaborado para se determinar o tamanho da amostra que quando se faz uma pergunta com somente duas respostas. O procedimento adotado, extraído de Thompson (1987), se mostrou difícil de ser aplicado num primeiro momento, porém acredita-se que os benefícios de se elaborar questões com mais de duas respostas sejam compensadores para se optar um procedimento mais difícil.

O uso da amostragem estratificada também se mostrou um recurso interessante, por exemplo, foi observada que o oitavo estrato (alunos do período noturno cursando currículo de administração em extinção) se mostrou muito mais insatisfeitos com a competência dos professores que os demais estratos.

Sabendo do histórico deste estrato e dos professores que ministraram aula, é possível analisar este fato de modo mais detalhado que quando se tem somente este índice de satisfação global, sem a estratificação por período e por curso.

Muitas outras análises seriam possíveis

de serem feitas se o questionário fosse mais bem elaborado e versasse sobre assuntos mais específicos, no entanto, neste primeiro momento o objetivo era conhecer um pouco mais os alunos, sua opinião e, sobretudo, aplicar a técnica de amostragem estratificada e questões com distribuição multinomial.

7. BIBLIOGRAFIA

- Bryson, Maurice C. The Literary Digest poll: Making of a statistical myth. *The American Statistician*, v. 30, n.4, p.184-185, 1976.
- Chocran, W. G. - *Sampling Techniques*. John Wiley & Sons, New York, 1977
- Downing D. & Clark, J. - *Estatística Aplicada*. 2ed. Ed. Saraviva, São Paulo, 2002.
- MADONNA, G.T. e YOUNG, M. - The First Political Poll, In *Politically Uncorrected Column*, 2002. Disponível em <<http://www.fandm.edu/politics/politically-uncorrected-column/2002-politically-uncorrected/the-first-political-poll>> acessado em 05 de junho de 2013.
- Malhotra, Naresh K. *Pesquisa de Marketing: Uma Orientação Aplicada*. 4ed., Ed. Bookman, Porto Alegre, 2006.
- MOON, Nick - *Opinion Polls: history, theory and practice*. Oxford, 1999.
- SQUIRE, Peverill. Why the 1936 Literary Digest Poll Failed. *Public Opinion Quarterly*, p. 125-133, 1988.
- THOMPSON, Steven K. Sample size for estimating multinomial proportions. *The American Statistician*, v. 41, n. 1, p. 42-46, 1987.