

**Universidade Federal do Paraná
Setor de Ciências Sociais Aplicadas
Departamento de Administração**

**Curso de Administração de Empresas
Disciplina: Estatística
Professor: Ricardo Valgas**

**Alunos: Augusto Escudero
Bernardo Adão
Bruno Patias Volpi
Filipe Neves
Flavia Oliveira**

**Curitiba
2007**

**Universidade Federal do Paraná
Setor de Ciências Sociais Aplicadas
Departamento de Administração**

Amostragem Estratificada

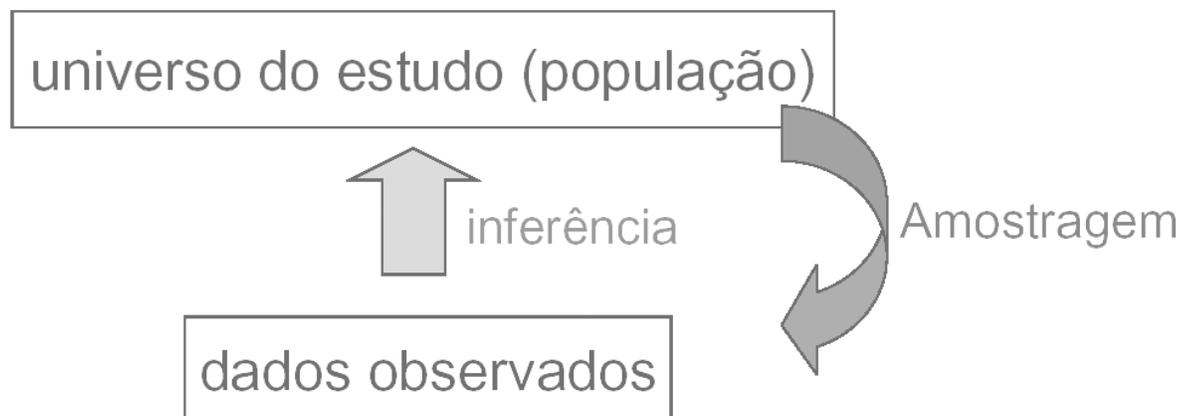
**Trabalho entregue à disciplina de Estatística,
como requisito para obtenção de nota
parcial no Curso de Administração da
Universidade Federal do Paraná.**

Orientador: Profº Ricardo Valgas.

**Curitiba
2007**

1.0 Introdução

O elemento fundamental em Estatística é a estimativa. Isto é, em uma pesquisa muitas vezes há a necessidade de se fazer previsões: Luis Inácio Lula da Silva, vai ganhar a eleição com uma margem de 2,5% de erro; a inflação do mês que vem deve chegar aos 17%. Em quase todos os casos, a estimativa está associada a uma pesquisa ou a uma verificação de características que não é realizada sobre todos os elementos da população, mas sim sobre uma parte dela, chamada de amostra. Assim, um dos objetivos da estatística é tirar conclusões sobre o “todo” (população) a partir das informações fornecidas pela “parte representativa” do todo (amostra). Assim, realizadas as fases de descrição dos dados (estatística descritiva), é feita uma análise dos resultados, obtidos através dos métodos da Estatística Inferencial ou Indutiva, que tem por base a indução (dedução de dados com base em um raciocínio que se acredita em uma precisão genérica a partir de fatos particulares) obtida por meio da teoria da probabilidade.



Quanto mais se conhecer sobre a população, melhores serão as informações colhidas pela observação de uma amostra. Por exemplo, uma cozinheira para verificar se o ensopado que ela está preparando tem ou não uma quantidade de sal desejada, experimenta apenas uma colher de ensopado, pois se sabe que a distribuição do sal em todo o ensopado é homogênea, e de qualquer lugar que se tivesse retirado a amostra do ensopado, ela seria “representativa”. Ou como em um exame de sangue, quando informações sobre a saúde de todo nosso corpo é baseado nas informações colhidas de algumas gotas de sangue, supondo que o sangue que circula em nosso organismo seja homogeneamente misturado. Mas nem sempre a escolha de uma amostra representativa é imediata. Quando a população da qual se deseja informações não é homogênea, são necessárias técnicas que assegurem amostras confiáveis para o estudo. A maneira de se escolher os elementos para compor uma amostra é denominada Amostragem.

Como as conclusões relativas à população vão estar baseadas nos resultados obtidos nas amostras dessa população, vemos o quão importante é a fase de coleta dos dados (amostragem), pois se erros forem cometidos no momento de selecionarmos os elementos da amostra o trabalho fica comprometido e os resultados finais serão provavelmente incorretos. Devemos, portanto, tomar cuidados especiais quanto aos critérios que serão utilizados na seleção da amostra.

No processo de seleção de amostras, as seguintes condições básicas deverão ser observadas para que a amostra seja representativa da população:

- Definição da **população da pesquisa** (grupo completo de objetos ou elementos relevantes para o projeto de pesquisa, por possuírem as informações que o projeto de pesquisa se propõe a coletar);
- Obtenção da lista de todas as **unidades amostrais da população** (os elementos ou objetos disponíveis para seleção durante o processo de amostragem);
- Decisão sobre o tamanho da amostra;
- Seleção de um procedimento específico através do qual a amostra será determinada ou selecionada e,
- Seleção física da amostra, tendo com base os procedimentos anteriores.

Como já dito, a técnica da amostragem tem como vantagens: o custo reduzido graças à necessidade de menos pessoas qualificadas, um menor aparato tecnológico, uma logística facilitada, entre outros fatores; e o tempo reduzido, já que o tempo de coleta e de apuração dos dados se faz de maneira mais ágil; e a flexibilidade oferecida em casos nos quais o censo (pesquisa onde os dados são coletados integralmente) é de certa forma impraticável.

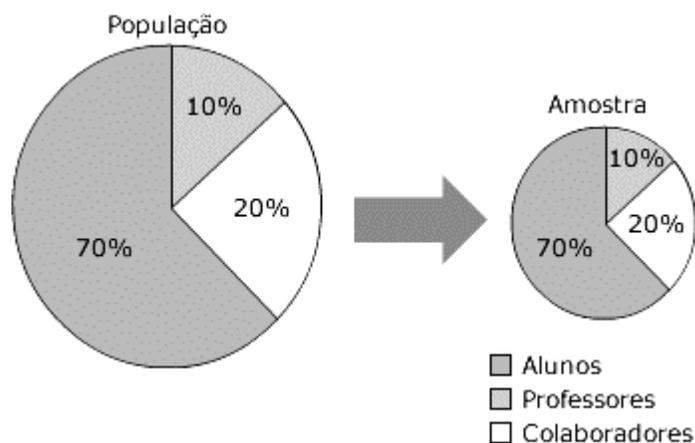
Porém existem casos em que a realização de uma pesquisa amostral não convém, por exemplo, em uma eleição política quando, apesar de todas as previsões serem baseadas em pesquisas amostrais, o real resultado deve ser altamente preciso e

uma pesquisa com toda a população (eleitores) é realizada; ou quando o tamanho da amostra é muito grande em relação ao tamanho da população; ou ainda quando já se dispõe dos dados da população. Nesses casos a realização de um censo é aconselhável, pois as facilidades trazidas inicialmente pela amostragem não surtem mais efeito e sua viabilidade deixa de ser válida.

2.0 Amostragem Estratificada

2.1 Definição

Quando a distribuição da variável que desejamos estimar é muito assimétrica, ou seja, tem uma alta variedade, impossibilita o emprego da amostra aleatória simples porque o tamanho da amostra seria praticamente igual ao tamanho da população. Quando isso ocorre há outro método que faz uso de uma variável auxiliar de modo a melhorar a precisão da estimativa. Por esse método, a população é dividida em grupos denominados de estratos, os quais são formados para ter o máximo de homogeneidade (dada por diversos fatores como topografia, religião, idade, etnia, etc.), esse método é chamado de amostragem estratificada.



As funções da amostra estratificada são inúmeras, um exemplo é a representação do universo das ações judiciais transitadas em julgado e não definidas

como de guarda permanente, com base em técnica de amostragem estratificada para a seleção de amostras representativas, com base na heterogeneidade desses processos.

Porém, embora pareça uma amostragem totalmente nova, essa técnica está fortemente relacionada à amostragem simples, pois após dividir a população em subgrupos/estratos, utiliza-se deste processo para retirar uma amostra de cada subgrupo. Pode-se dizer então que a técnica da estratificação é basicamente uma maneira de adequar uma população ao principal pré-requisito necessário para que se possa aplicar a amostragem simples: a homogeneidade.

Existem três tipos de amostragem estratificada que podem ser utilizados:

1. Uniforme: sorteia-se igualmente o número de elementos em cada estrato, é usada quando os estratos são de, relativamente, mesmo tamanho;
2. Proporcional o número de elementos sorteados em cada estrato é proporcional ao número de elementos existentes no estrato, quando os subgrupos não têm o mesmo tamanho, esse tipo gera uma maior representatividade da população em relação à uniforme.
3. Ótima: é a mais confiável, toma em cada estrato, um número de elementos proporcional ao número de elementos do estrato e também a variação da variável de interesse no estrato, medida pelo seu desvio-padrão. Pretende-se assim otimizar a obtenção de informações sobre a população, com base no princípio de que, onde a variação é menor, menos elementos são necessários para bem caracterizar o comportamento da variável. Dessa forma, com um

menor número total de elementos na amostra, se conseguiria uma quantidade de informação equivalente à obtida nos demais casos. As principais dificuldades para a utilização desse tipo de amostragem residem nas complicações teóricas relacionadas com a análise dos dados e em que, muitas vezes não podemos avaliar de antemão o desvio-padrão da variável nos diversos estratos.

2.2 Processo

Como visto na amostragem estratificada a amostra de cada subgrupo deve ser tirado proporcionalmente em relação ao tamanho desse estrato. Porém deve ser visto também a variabilidade em cada estrato, pois quanto maior a variabilidade maior o número de amostras para se ter um resultado confiável.

Passos:

1. Determinar a fração da amostra: dada por $f = n/N$, sendo f (fração da amostra), n (tamanho da amostra geral) e N (tamanho da população).
2. Calcular o número de elementos a serem retirados em cada estrato, proporcionalmente: dado por $n_1 = N_1 \cdot f$, sendo n_1 (tamanho da mostra no estrato 1), N_1 (tamanho do estrato) e f (fração da amostra). Lembrando-se que é necessário repetir este passo para cada estrato identificado.
3. Utilização da amostra simples para a conclusão da retirada da amostra.

2.3 Vantagens

A técnica da amostragem estratificada é largamente utilizada. Isso se deve a alguns fatores que a viabilizam:

1. Quando se deseja dados de determinada precisão sobre certas subdivisões da população esse método é aconselhável visto que cada uma dessas subdivisões passa a ser tratada como um grupo diferenciado.
2. As conveniências administrativas ocasionadas são um ponto positivo nesse tipo de amostragem, pois as divisões de tarefas no momento das pesquisas de campo ou até mesmo das apurações facilitam um melhor controle e uma melhor organização a pesquisa.
3. Nos casos em que dentro da população a ser analisada existe um grupo que não pode ser considerado da mesma maneira que os outros, através dessa técnica, pode-se separá-lo dos demais tornando-o um estrato.
4. Graças à divisão com o propósito de aumentar a homogeneidade da população, pode-se dizer ainda que a amostragem estratificada garante, de certa forma, uma maior precisão nos resultados.

3.0 Paralelo entre amostragem estratificada e amostragem por conglomerados

Ao se comparar a amostragem estratificada com a amostragem por conglomerados, pode-se encontrar alguma dificuldade em identificar as diferenças que às separam. Porém após um melhor entendimento a respeito, é possível observar que além do fato de que a primeira divide a população em estratos homogêneos e a segunda em subpopulações heterogêneas, existe uma diferença básica na maneira de se proceder: na técnica por conglomerados são utilizados todos os elementos de cada conglomerado (por isso é dito que para sua realização o ideal é que se tenha vários conglomerados e que em cada conglomerado existam poucos elementos) e não existe nenhuma proporcionalidade determine a quantidade ideal de conglomerados a serem analisados; já a amostragem estratificada seleciona amostras de cada estrato de proporcionalmente ao tamanho de cada um deles.

4.0 Exercícios

Exercício 1

Pretende-se estudar o nível salarial dos empregados de certa indústria. Suponhamos que esta indústria tem 2000 empregados, assim distribuídos: 1400 operários, 400 mestres e 200 supervisores.

N: 2000

1400 operários : 70%

400 mestres : 20%

200 supervisores : 10%

(seleção de funcionários - amostra – proporcional ao tamanho de cada estrato)

Operários	Mestres	Supervisores
salários	salários	salários
400	1200	3000
450	1300	4000
500	1500	
300	1000	
500		
400		
450		
700		

600		
400		
500		
500		
400		
600		
média 1	média 2	média 3
478,57	1250	3500

- média de cada estrato: $\bar{X}_i = \frac{\sum_{j=1}^{n_i} X_{ij}}{n_i}$, $i=1,2,3$

$$\bar{X}_1 = \frac{\sum_{j=1}^{14} X_{1j}}{14} = 478,57$$

$$\bar{X}_2 = \frac{\sum_{j=1}^4 X_{2j}}{4} = 1250$$

$$\bar{X}_3 = \frac{\sum_{j=1}^2 X_{3j}}{2} = 3500$$

- média geral(estratificada): $\bar{X} = \frac{\sum_{i=1}^k n_i \bar{X}_i}{N}$ onde $N=n_1+n_2+n_3$

$$\bar{X} = \frac{1400 \cdot 478,57 + 400 \cdot 1250 + 200 \cdot 3500}{2000} = 975$$

Em média, o salário dos funcionários da empresa é de 975 reais.

- variância de cada estrato i: $s_i^2 = \frac{\sum_{j=1}^{n_i} (X_{ij} - \bar{X}_i)^2}{n_i - 1}$

- desvio padrão do estrato i: $S_i = \sqrt{s_i^2}$

Diferença entre salários e média		
-8,5714	-50	-500
-28,5714	50	500
21,4286	250	
-78,5714	-250	
21,4286		
-78,5714		
-28,5714		
221,4286		
121,4286		
-78,5714		
21,4286		
21,4286		
-78,5714		
121,4286		

Diferença entre salários e média ao quadrado			
6173,464898	2500	250000	
816,324898	2500	250000	
459,184898	62500		
31887,7449	62500		
459,184898			
6173,464898			
816,324898			
49030,6249			
14744,9049			
6173,464898			
459,184898			
459,184898			
6173,464898			
14744,9049			
Variância			
10659,34066	43333,3		500000
Desvio padrão			
103,2440829	208,167	707,1067812	

- variância geral (estratificada): $S_{ST}^2 = \sum_{i=1}^k \text{proporção estrato } i \cdot S_i^2$

$$S_{ST}^2 = 0,7 \cdot 10659 + 0,2 \cdot 43333 + 0,1 \cdot 500000 = 66127,9$$

- desvio padrão geral (estratificada): $S_{ST} = \sqrt{S_{ST}^2}$

$$S_{ST} = \sqrt{66127,9} = 257,1535$$

O salário dos funcionários em média é de 975, porém podem variar, na maioria das vezes, em torno de 257 reais.

Exercício 2

Uma amostra aleatória simples de 120 famílias, tomadas o início de um ano, fornece as seguintes informações sobre gastos em entretenimentos, nos três meses anteriores:

Região	Número de Famílias	Gasto Médio \$	Desvio Padrão \$
Metropolitana	52	61,23	9,99
Urbana	33	38,79	10,64
Rural	35	23,51	5,73
Total*	120	44,06	18,52

* Os valores desta linha estão relacionados aos de cada região. Os valores são da amostra.

A população consistia de 10.000 famílias e sabia-se que a sua distribuição entre os três estratos estava na razão de 5:3:2, respectivamente. Estimar os limites de confiança de 95% para o gasto médio da população e para o total dos gastos do trimestre dado.

Tratando a amostra como aleatória simples de 120 famílias, os limites de confiança de 95% para a media serão:

$$\bar{X} \pm 1,96 s/\sqrt{N}$$

ou

$$\$44,06 \pm 3,31$$

Ou

$$\$ 40,75 \text{ a } 47,37$$

E os limites de confiança de 95% para o gasto total será o gasto médio vezes 10.000, ou seja :

$$\$407.500 \text{ a } \$ 473.700$$

Entretanto, se estratificarmos a amostra após a seleção, com base na distribuição dada da população entre os estratos, obteremos como uma estimativa dos gastos médios para o trimestre em questão:

$$\begin{aligned}\bar{X}^* &= \sum \alpha_s X_s \\ &= \$ 46,95\end{aligned}$$

E os limites de confiança de 95% serão

$$\bar{X} \pm 1,96 \sqrt{\sum (\alpha_s^2 s_s^2 / N_s)}$$

Ou seja

\$ 46,95 ± 1,78

Ou

\$ 45,17 a \$ 48,73

Os limites de confiança de 95% para o gasto total serão 10.000 vezes os limites de gasto médio, isto é;

\$ 451.700 a \$ 487.300

Aqui, a estratificação após a seleção estreitou o intervalo dos limites de confiança. Também aumentou a estimativa do gasto médio, pois o estrato metropolitano esta sub-representado na amostra aleatória simples.

5.0 Referências Bibliográficas

COCHRAN, G. William; *Técnicas de Amostragem*; Editora Fundo de Cultura S.A.; Rio de Janeiro (1968).

<http://www.unipar.br>; Acessado em 22 de outubro de 07.

<http://www.fag.edu.br>; Acessado em 23 de outubro de 07

Hoffmann, Rodolfo; *Estatística para Economistas*; Editora Thomson; São Paulo (2006)

7. M. POLASEK, P. H. KARMEEL ; *Estatística Geral e Aplicada*. São Paulo: Atlas, 1972€