

CENTRO DE ESTATÍSTICA APLICADA – CEA – USP
RELATÓRIO DE CONSULTA

TÍTULO DO PROJETO: Avaliação do aprendizado em cursos de estatística básica oferecidos pelo Instituto de Matemática e Estatística da Universidade de São Paulo – IME/USP para os cursos de serviço

PESQUISADOR: Marcos Nascimento Magalhães

COLABORADORES: Fábio Fujikawa e Thiago Vilhega

INSTITUIÇÃO: Departamento de Estatística do Instituto de Matemática e Estatística da Universidade de São Paulo – IME/USP

FINALIDADE DO PROJETO: Iniciação Científica

PARTICIPANTES DA ENTREVISTA: Marcos Nascimento Magalhães
Denise Aparecida Botter
Júlia Maria Pavan Soler
Luciana Xarim Mendes
Lúcio Franchi Cruz

DATA: 13/04/2004

FINALIDADE DA CONSULTA: Sugestões para o planejamento e dimensionamento amostral

RELATÓRIO ELABORADO POR: Luciana Xarim Mendes
Lúcio Franchi Cruz

1. Introdução

A avaliação da aprendizagem de qualquer disciplina é sempre uma difícil tarefa, uma vez que diversos fatores podem influenciar no quanto um determinado indivíduo foi capaz de aprender sobre um conteúdo programado. Além disso, há um grau de subjetividade na atribuição de uma variável para quantificar “aprendizagem”. Essa dificuldade torna-se ainda mais evidente quando consideramos a área de conhecimento de exatas, como Estatística.

O presente estudo tem por objetivo avaliar o aprendizado dos alunos que cursaram (e foram aprovados) no curso de Estatística Básica do Instituto de Matemática e Estatística da Universidade de São Paulo (IME-USP), verificando o quanto os alunos estão conseguindo aprender e apreender os conceitos básicos mais importantes e mais utilizados, que futuramente proporcionarão a interação entre a Estatística e a área de estudo de cada aluno de forma efetiva.

O pesquisador procurou o CEA a fim de obter sugestões para o planejamento e o dimensionamento amostral que possibilitem alcançar o objetivo proposto.

2. Descrição do Estudo

O universo da pesquisa é composto pelos alunos que foram aprovados nos cursos de serviço oferecidos pelo IME-USP, a saber:

- a) MAE 112 - Introdução à Probabilidade e à Estatística I
- b) MAE 121 - Introdução à Probabilidade e à Estatística I
- c) MAE 219 - Introdução à Probabilidade e à Estatística I
- d) MAE 122 - Introdução à Probabilidade e à Estatística II
- e) MAE 212 - Introdução à Probabilidade e à Estatística II
- f) MAE 229 - Introdução à Probabilidade e à Estatística II
- g) MAE 1511 – Estatística para Licenciatura I
- h) MAE 1512 – Estatística para Licenciatura II

Dos cursos listados acima, MAE 112 e MAE 122 são ministrados somente no período diurno enquanto que os demais são ministrados nos períodos diurno e noturno.

Para um aluno fazer parte do estudo, alguns critérios devem estar satisfeitos: o aluno deve ter sido aprovado e não ter cursado anteriormente nenhum outro curso introdutório de estatística oferecido pelo Instituto. Alunos que já tiverem cursado anteriormente qualquer um dos cursos acima mencionados e que não obtiveram aprovação não deverão fazer parte da amostra. Outros critérios de exclusão devem ser observados: alunos matriculados na disciplina MAE 116 – Noções de Estatística e alunos que tiverem cursado as disciplinas: MAE 112, MAE 121 ou MAE 219 há mais de um ano, não devem ser entrevistados. Além disso, alunos que tiverem cursado MAE 122, MAE 212, MAE 229 ou MAE 1512 há mais de um ano depois de terem sido aprovados nas disciplinas MAE 112, MAE 121, MAE 219 ou MAE 1511, respectivamente, também não devem ser entrevistados.

A coleta dos dados será realizada pelo pesquisador por meio de um questionário composto por 50 questões fechadas (com somente duas categorias de resposta, sim ou não). Esse questionário deverá ser aplicado no semestre seguinte à conclusão (e aprovação) do curso avaliado.

3. Descrição das variáveis

As variáveis serão definidas após a finalização do questionário e aplicação de um pré-teste (para validar o instrumento de avaliação) a ser realizado com os alunos do Bacharelado em Estatística do IME-USP.

Possivelmente serão 50 variáveis nominais classificadas nas categorias “Sim” e “Não”, associadas às questões do questionário, além das seguintes variáveis: Duração do Curso (semestral ou anual), Área de Conhecimento (humanas ou exatas), Unidade de Origem (FP, IME ou FEA) e Curso de Origem (Psicologia, Economia, etc.).

4. Situação do Projeto

O projeto está em fase de finalização da elaboração do questionário. A consulta ao CEA teve como objetivo a discussão de aspectos do dimensionamento da amostra e também do planejamento do experimento, a fim de determinar a forma de conduzir a coleta dos dados e a abordagem dos alunos que farão parte da amostra.

5. Sugestões do CEA

Primeiramente, recomenda-se que o questionário seja aplicado a alguns alunos (amostra piloto), por exemplo, dentre aqueles que tenham sido aprovados em algum dos cursos básicos de estatística oferecidos pelo Instituto. O objetivo desse pré-teste é o de fazer uma validação do questionário, ou seja, verificar se cada questão está sendo interpretada corretamente pelos entrevistados, além de verificar quais variáveis podem ser excluídas ou necessitam ser incluídas no estudo.

Para o experimento propriamente dito, recomenda-se uma amostragem, e não um censo, já que se trata de uma população de cerca de 800 estudantes. Esse valor é baseado no número de alunos matriculados nos cursos (itens a, b, c e g da Descrição do Estudo) no primeiro semestre de 2004, indicado pelo sistema Júpiter, do IME. Um censo, nesse caso, seria desaconselhável já que existem dificuldades logísticas, tais como localizar cada aluno em diferentes faculdades da universidade, gerando gastos desnecessários.

Alunos que cursaram somente as disciplinas introdutórias, conforme indicado na Descrição do Estudo (itens: a, b, c e g) serão considerados como alunos com cursos semestrais e aqueles que, além das disciplinas anteriormente mencionadas, cursaram também as disciplinas indicadas nos itens: d, e, f e h da mesma seção, serão considerados com cursos anuais, desde que as duas disciplinas tenham sido cursadas no mesmo ano.

O planejamento amostral sugerido, a fim de estimar a proporção de alunos que conseguiram aprender os conceitos básicos de estatística, é a

amostragem aleatória estratificada com alocação proporcional considerando-se quatorze estratos (Bolfarine e Bussab, 2000). Esses estratos são as diversas disciplinas (incluindo as divisões de cursos diurnos e noturnos e cursos anuais e semestrais) oferecidas para as Faculdades de Economia, Administração e Contabilidade (FEA), Psicologia (FP) e as oferecidas no IME, para os cursos de Licenciatura, Matemática Pura e Matemática Aplicada. Para um melhor entendimento do planejamento sugerido, consultar a Tabela A.1.

A amostra dos alunos deverá ser sorteada aleatoriamente, (utilizando-se as listas de alunos aprovados em cada curso), sem reposição, proporcionalmente ao número de alunos matriculados em cada disciplina, considerando a duração do curso (semestral ou anual) e o período (diurno ou noturno).

Para o cálculo do tamanho da amostra, considerou-se o número de alunos matriculados nas disciplinas oferecidas no primeiro semestre de 2004. (ver Apêndice C, que contém as tabelas com o número de alunos matriculados em cada disciplina. Os dados foram obtidos do Sistema Júpiter do IME-USP, no primeiro semestre de 2004). Estimou-se em seguida o número de matriculados para as disciplinas ministradas no segundo semestre. Para isto, assumiu-se que o índice de aprovação foi igual a 90% para os cursos do primeiro semestre. Assumiu-se também que o índice de aprovação foi de 80% dos alunos para os cursos do segundo semestre. O ideal, no entanto, seria considerar o número de alunos efetivamente aprovados em cada disciplina oferecida, por exemplo, do ano de 2003. Esse dado, no entanto, não está disponível no presente momento e conseguí-lo pode levar muito tempo. Além disso, usar os dados de 2004, ainda que parte deles estimados, não causa maiores prejuízos aos cálculos realizados para a conclusão deste estudo.

Na Tabela A.3 apresentou-se o número de alunos matriculados no primeiro semestre de 2004 em cada estrato e na Tabela A.4 o número estimado de alunos aprovados no primeiro e segundo semestres.

Utilizou-se, para o cálculo da amostra, diversas frações (valores de F) (ver Apêndice B) conforme apresentado na Tabela 1 a seguir. Cabe ao

pesquisador analisar qual o tamanho amostral mais adequado para alcançar o objetivo proposto, considerando as limitações logísticas e operacionais da coleta dos dados.

Para um melhor entendimento, deve-se considerar P_h (ver Apêndice B) como a proporção de alunos que foram classificados como os que aprenderam a matéria dada em cada disciplina e Q_h como $1 - P_h$. Considerou-se então, o valor de $P_h Q_h$ (produto das proporções) como sendo 0,25, já que não dispomos de informação preliminar sobre P_h e nem amostras piloto ou pesquisas anteriores (Bolfarine e Bussab, 2000). É importante observar que adotar essa consideração corresponde à solução mais conservadora, que leva a um maior tamanho amostral.

Usando Alocação Ótima de Neyman (Bolfarine e Bussab, 2000), (para maiores detalhes ver Apêndice B), chegou-se à seguinte tabela:

Tabela 1 – Tamanho Amostral

			Número de alunos que devem responder ao questionário			
Curso	Período	Amostra	Valor para F = 20%	Valor para F = 25%	Valor para F = 30%	Valor para F = 50%
MAE 219	Diurno	n_1	3	4	4	7
	Noturno	n_2	2	3	3	5
MAE 229	Diurno	n_3	2	3	3	6
	Noturno	n_4	2	2	3	4
MAE 1511	Diurno	n_5	14	17	21	35
	Noturno	n_6	20	25	29	49
MAE 1512	Diurno	n_7	11	14	17	28
	Noturno	n_8	16	20	24	39
MAE 121	Diurno	n_9	28	35	42	70
	Noturno	N_{10}	12	15	18	30
MAE 212	Diurno	N_{11}	22	28	34	56
	Noturno	N_{12}	10	12	14	24
MAE 112	Diurno	N_{13}	12	15	18	30
MAE 122	Diurno	N_{14}	10	12	14	24
TOTAL			163	204	245	408

Após a aplicação dos questionários, uma possível tabulação dos dados poderá ser realizada, por exemplo, utilizando uma planilha do Microsoft Excel para Windows conforme apresentado na Tabela A.2.

O pesquisador deve realizar seu estudo piloto e a primeira coleta de dados ainda neste primeiro semestre de 2004, a fim de obter dados anuais, pois este é um semestre imediatamente posterior a um semestre onde foram ministrados cursos como os dos itens e, d, f e h (denominados anuais). Caso isso não ocorra, o pesquisador poderá coletar somente dados dos cursos semestrais (cursos introdutórios, como os itens a, b, c e g da Descrição do Estudo), oferecidos no segundo semestre de 2004.

Recomenda-se um primeiro contato com os alunos selecionados para comporem a amostra, (levando-se em conta os critérios de inclusão e exclusão anteriormente apresentados), a fim de expor os objetivos do estudo e agendar uma data para aplicação do questionário. Para isso, sugere-se ir às faculdades onde os alunos estudam e, com prévio consentimento dos professores, usar alguns minutos de algumas aulas para explicar esse estudo.

Para a classificação de um aluno entre aqueles que “aprenderam” o conteúdo de um curso e aqueles que “não aprenderam”, serão utilizados os resultados dos questionários. Um critério a ser adotado é alocar no grupo dos “que aprenderam” os que obtiverem pelo menos 75% de acerto das questões propostas (38 questões), e os demais no outro grupo. É importante ressaltar que outras propostas poderão ser adotadas na descrição dos dados.

6. Referências Bibliográficas

BOLFARINE, H. e BUSSAB, W.O. (2000). **Elementos de Amostragem**. Versão Preliminar. São Paulo.

APÊNDICE A

Tabela A.1. Descrição dos estratos para o plano amostral

FEA – Faculdade de Economia, Administração e Contabilidade				Faculdade de Psicologia		IME – Instituto de Matemática e Estatística							
						Licenciatura				Computação, Matemática Pura e Matemática Aplicada			
Semestral MAE 219		Anual MAE 229		Semestral MAE 112	Anual MAE 122	Semestral MAE 1511		Anual MAE 1512		Semestral MAE 121		Anual MAE 212	
Diurno	Noturno	Diurno	Noturno	Diurno	Diurno	Diurno	Noturno	Diurno	Noturno	Diurno	Noturno	Diurno	Noturno

Tabela A.2. Exemplo de planilha dos dados

Id	Curso	Duração do curso	Área de Conhecimento	Unidade de origem	Curso de Origem	Período do curso	Var 1 (*)	Var 2 (*)	Var 50 (*)
1	MAE 112	Semestral	Humanas	FP	Psicologia	Diurno	1	0	1
2	MAE 121	Anual	Exatas	IME	Lic. Mat.	Noturno	0	0	0
3	MAE 121	Semestral	0	0	1
4	MAE 219	Anual	1	1	1
5	MAE 219	Semestral	0	1	0
6	MAE 112	Anual	0	0	1
.
.
.
N	MAE 219	Anual	Humanas	FEA	Economia	Noturno	0	1	0

(*) Codificação: 0 = Não e 1 = Sim

Tabela A.3. Número de alunos matriculados em cada curso (ver Apêndice C)*

FEA – Faculdade de Economia, Administração e Contabilidade				Faculdade de Psicologia		IME – Instituto de Matemática e Estatística							
						Licenciatura				Computação, Matemática Pura e Matemática Aplicada			
Semestral MAE 219		Anual MAE 229		Semestral MAE 112	Anual MAE 122	Semestral MAE 1511		Anual MAE 1512		Semestral MAE 121		Anual MAE 212	
Diurno	Noturno	Diurno	Noturno	Diurno	Diurno	Diurno	Noturno	Diurno	Noturno	Diurno	Noturno	Diurno	Noturno
16	12			67		77	109			156	67		

* Na presente data, só existem os números dos alunos matriculados nas disciplinas do primeiro semestre de 2004, fornecidos pelo Sistema Júpiter do IME-USP.

Tabela ^a 4. Tamanhos dos estratos (valores estimados)*

FEA – Faculdade de Economia, Administração e Contabilidade				Faculdade de Psicologia		IME – Instituto de Matemática e Estatística							
						Licenciatura				Computação, Matemática Pura e Matemática Aplicada			
Semestral MAE 219		Anual MAE 229		Semestral MAE 112	Anual MAE 122	Semestral MAE 1511		Anual MAE 1512		Semestral MAE 121		Anual MAE 212	
Diurno	Noturno	Diurno	Noturno	Diurno	Diurno	Diurno	Noturno	Diurno	Noturno	Diurno	Noturno	Diurno	Noturno
14	11	12	9	60	48	69	98	55	78	140	60	112	48

* **Diurno/Noturno** – 90% do número de alunos matriculados (considerando 10% de reprovação para os cursos do primeiro semestre)

Diurno/Noturno – 80% do número de alunos matriculados (considerando 20% de reprovação para os cursos do segundo semestre)

APÊNDICE B

Para a construção da Tabela 1 na Seção 5, usou-se a Alocação Ótima de Neyman (Bolfarine e Bussab, 2000).

Considerou-se, para tal, uma função de custo de forma linear, onde o custo C é dado por:

$$C = C_0 + \sum_{h=1}^H c_h n_h$$

onde C_0 denota o custo inicial, c_h o custo por unidade observada no estrato h e n_h o tamanho amostral.

Considerou-se, para este estudo, custo fixo, isto é:

$$C = C_0 + nc$$

onde c é o custo por unidade.

Assim, a Alocação Ótima se reduz a:

$$n_h = n \frac{N_h \sqrt{P_h Q_h}}{\sum_{h=1}^H N_h \sqrt{P_h Q_h}}$$

onde, nesse caso, P_h é a proporção de alunos que foram classificados como os que aprenderam a matéria dada em cada disciplina.

Lembrar que $Q_h = 1 - P_h$.

Considerou-se então, o valor de $P_h Q_h$ (produto das proporções) como sendo 0,25 (por ser um limite superior), já que não se dispõe de informação preliminar sobre P_h e nem amostras piloto ou pesquisas anteriores.

Assim, n_h se reduz a:

$$n_h = n \frac{N_h}{N}$$

Os valores de F , usados também na referida tabela, são diversos valores para a razão n/N . Dispôs-se valores para F de 20%, 25%, 30% e 50%.

APÊNDICE C

**Instituto de Matemática e Estatística
Estatística**

Disciplina: MAE1511 - Estatística para Licenciatura I

Disciplinas Oferecidas para o 1º Semestre de 2004

	Horário	Prof.
	ter 10:00 11:40	Mariane Streibel
	qui 08:00 09:40	Mariane Streibel

Código da Turma: 2004142
Início: 01/03/2004
Fim: 03/07/2004
Tipo da Turma: Teórica

	Vagas	Inscritos	Pendentes	Matriculados
Obrigatória	80	78	0	77

	Horário	Prof.
	ter 19:20 21:00	Chang Chiann
		Nikolai Valtchev Kolev
	qui 21:10 22:50	Chang Chiann
		Nikolai Valtchev Kolev

Código da Turma: 2004147
Início: 01/03/2004
Fim: 03/07/2004
Tipo da Turma: Teórica

	Vagas	Inscritos	Pendentes	Matriculados
Obrigatória	140	109	1	108
Optativa Livre	1	1	0	1
Qualquer Unidade da USP	0	0	-	-

Disciplina: MAE0112 - Introducao a Probabilidade e a Estatistica I

Disciplinas Oferecidas para o 1º Semestre de 2004

	Horário	Prof.
	ter 12:00 13:40	Vladimir Belitsky
	sex 12:00 13:40	Vladimir Belitsky

Código da Turma: 2004150
Início: 01/03/2004
Fim: 03/07/2004
Tipo da Turma: Teórica

	Vagas	Inscritos	Pendentes	Matriculados
Obrigatória	80	68	0	67
Optativa Livre	1	0	0	0
Qualquer Unidade da USP	0	0	-	-

Disciplina: MAE0121 - Introdução a Probabilidade e a Estatística I
Disciplinas Oferecidas para o 1º Semestre de 2004

Código da Turma: 2004141 **Horário** **Prof.**
Início: 01/03/2004 qua 10:00 11:40 Claudia Monteiro Peixoto
Fim: 03/07/2004 sex 08:00 09:40 Claudia Monteiro Peixoto
Tipo da Turma: Teórica

	Vagas	Inscritos	Pendentes	Matriculados
Obrigatória	92	90	1	89

Código da Turma: 2004151 **Horário** **Prof.**
Início: 01/03/2004 qua 10:00 11:40 Silvia Lopes de Paula Ferrari
Fim: 03/07/2004 sex 08:00 09:40 Silvia Lopes de Paula Ferrari
Tipo da Turma: Teórica

	Vagas	Inscritos	Pendentes	Matriculados
Obrigatória	80	66	0	66
Optativa Livre	1	1	0	1
Qualquer Unidade da USP	0	0	-	-

Código da Turma: 2004154 **Horário** **Prof.**
Início: 01/03/2004 seg 21:10 22:50 Antonio Carlos Pedroso de Lima
Fim: 03/07/2004 qua 19:20 21:00 Antonio Carlos Pedroso de Lima
Tipo da Turma: Teórica

	Vagas	Inscritos	Pendentes	Matriculados
Obrigatória	64	61	0	61
Optativa Livre	4	4	0	4
Qualquer Unidade da USP	0	0	-	-
Alunos Especiais	2	2	-	2

Disciplina: MAE0219 - Introdução à Probabilidade e a Estatística I

Disciplinas Oferecidas para o 1º Semestre de 2004

Código da Turma: 2004101 **Horário** **Prof.**
Início: 01/03/2004 ter 07:30 09:10 Clelia Maria de Castro Toloí
Fim: 03/07/2004 qui 09:20 11:00 Clelia Maria de Castro Toloí
Tipo da Turma: Teórica

	Vagas	Inscritos	Pendentes	Matriculados
Obrigatória	70	14	0	14
Optativa Livre	2	3	0	2
Qualquer Unidade da USP	0	0	-	-

Código da Turma: 2004121 **Horário** **Prof.**
Início: 01/03/2004 ter 19:30 21:10 Mariane Streibel
Fim: 03/07/2004 qui 21:20 23:00 Mariane Streibel
Tipo da Turma: Teórica

	Vagas	Inscritos	Pendentes	Matriculados
Obrigatória	70	10	0	10
Optativa Livre	2	2	0	2
Qualquer Unidade da USP	0	0	-	-