

Estatística Econômica II

Lupércio França Bessegato
Departamento de Estatística/UFJF

Amostragem



Roteiro Geral



1. Introdução
2. Modelos probabilísticos
3. **Amostragem**
4. Distribuições amostrais e estimação
5. Testes de significância
6. Comparações de médias
7. Tópicos especiais
8. Referências

Estatística Econômica II - 2018

2



Roteiro do Módulo



2. Amostragem:
 - a) Introdução
 - b) Tipos de estudos
 - c) Amostragem
 - d) Cuidados com as pesquisas por amostragem
 - e) Referências

Estatística Econômica II - 2018

4

Introdução

Amostragem

- Objetivo:
 - √ Obter informações sobre um grande grupo de indivíduos
- São coletadas informações sobre apenas parte desse grupo com a finalidade de obtenção de conclusões sobre o todo

Estatística Econômica II - 2018

Pesquisa Amostral



- Objetivo:
 - √ Conhecer a população com base em uma amostra

Estatística Econômica II - 2018

Conceitos

- População:
 - √ Grupo total de indivíduos sobre os quais desejamos informações
 - Nosso desejo de conhecimento
- Amostra:
 - √ Parte da população que é efetivamente examinada a fim de obter informações
 - Parte da qual extraímos informações sobre o todo

Estatística Econômica II - 2018





Dados

- Matéria-prima para a obtenção de conclusões confiáveis
- Objetivo:
 - √ Informar-se com dados e não com simples opiniões (impressões subjetivas)

9

Estatística Econômica II - 2018





Dados Secundários

- Vantagens:
 - √ Mais baratos
 - √ Exige menos tempo
- Desvantagens:
 - √ Podem estar desatualizados
 - √ Dados podem ser irrelevantes
- Tipos:
 - √ Interno
 - √ Externo

10

Estatística Econômica II - 2018





Dados Primários

- Vantagens:
 - √ Atualizados
 - √ Diretamente relacionados com a pesquisa
- Desvantagens:
 - √ Mais caros
 - √ Mais tempo para coleta
- Tipos:
 - √ Observação
 - √ Levantamento
 - √ Experimental

11

Estatística Econômica II - 2018





Fontes de Dados Secundários

- Órgãos que utilizam intensivamente técnicas estatísticas de amostragem:
 - √ Brasil:
 - INEP - www.inep.gov.br
 - IBGE - www.ibge.gov.br
 - Fundação João Pinheiro - www.fjp.gov.br
 - SEADE - www.seade.gov.br
 - √ Canadá:
 - Statistics Canada - www.statcan.ca

12



Estatística Econômica II - 2018



- √ EUA:
 - US Census Bureau - www.census.gov
 - Bureau of Labor Statistics - www.bls.gov
- √ Inglaterra:
 - Office for National Statistics - www.statistics.gov.uk
- √ Colômbia:
 - DANE - www.dane.gov.co
- √ Chile:
 - INE - www.ine.gov.cl
- √ Argentina:
 - INDEC - www.indec.gov.ar

13

Estatística Econômica II - 2018





Produção de Dados

- Produzir dados em escala menor
- Utilizar esses dados para concluir sobre um cenário maior
 - √ Inferência estatística
- Sucesso da inferência depende do **planejamento** da produção de dados

14

Estatística Econômica II - 2018

Tipos de Estudo





Coleta de Dados

- Tipos de coleta de dados
 - √ Estudo observacional
 - √ Levantamento amostral
 - √ Experimento planejado

16

Estatística Econômica II - 2018





Estudo # 1

Uso de celular causa câncer de cérebro?



- **Questão:**
 - √ Casos observados de câncer cerebral entre usuários de celulares são evidências de que os celulares causam câncer de cérebro?
- **Procedimento:**
 - √ Estudos cuidadosos devem verificar grandes grupos para tentar obter conclusão
 - Ao invés de estudar casos isolados

Estatística Econômica II - 2018 17





Como conduzir o estudo?

Estatística Econômica II - 2018 18




- **Estudo observacional:**
 - √ não é capaz de responder diretamente a questão de causa e efeito
- Caso os dados apontarem a existência de uma relação entre uso de celular e câncer cerebral
 - √ Os usuários de celular podem ter outras características em comum que contribuem para o desenvolvimento de câncer cerebral

Estatística Econômica II - 2018 19




Como verificar se a radiação do celular de fato provoca câncer no cérebro?

Estatística Econômica II - 2018 20




Confundimento




- Variáveis confundidas:
 - √ Variáveis explicativas ou ocultas
 - √ Seus efeitos sobre uma variável resposta não podem ser distinguidos um dos outros
 - √ Exemplo: uso de celulares e ocorrência de câncer
 - Qualquer efeito cancerígeno do celular fica confundido (ou misturado) com variáveis ocultas (idade, profissão, lugar de residência, etc.)

21

Estatística Econômica II - 2018



Estudo # 2




Quais são os efeitos de uma exposição contínua a uma mensagem publicitária?


- Resposta:
 - √ Pode depender tanto da duração do anúncio quanto da frequência com que ele é repetido
- Sujeitos:
 - √ Universitários

22

Estatística Econômica II - 2018




Procedimento:




- Procedimento:
 - √ Todos os sujeitos assistem a um programa de TV de 40 minutos, que inclui propaganda de câmera digital
 - √ Alguns veem um comercial de 30 seg.; outros, uma versão de 90 seg.
 - √ Comercial é repetido 1, 3 ou 5 vezes durante o programa

23

Estatística Econômica II - 2018



Fatores:





- Fatores:
 - √ Duração do comercial
 - 2 níveis
 - √ Repetições:
 - 3 níveis
- Tratamentos:
 - √ 6 combinações de níveis e fatores

		Fator B Repetições (vezes)		
		1 vez	3 vezes	5 vezes
Fator A Duração (seg.)	30	# 1	# 2	# 3
	90	# 4	# 5	# 6



24

Estatística Econômica II - 2018





- Variáveis resposta:
 - √ Após o programa, pergunta-se a todos os sujeitos:
 - O que se recordam do anúncio
 - Opinião em relação à câmera
 - Intenção de compra

Estatística Econômica II - 2018 25



- Interação:
 - √ Efeito simultâneo de diversos fatores
 - √ Ex.: **Talvez**
 - Comerciais mais longos aumentem o interesse
 - Repetições aumentem o interesse
 - Muitas repetições de comerciais mais diminuam interesse



Estatística Econômica II - 2018 26



Alguns Tipos de Pesquisas

Dados Secundários	}	• Pesquisa bibliográfica	}	Estudos Observacionais
		• Pesquisa histórica		
Dados Primários	}	• Estudo de caso		
		• Pesquisa (“survey”)		
		• Estudo caso-controle		
		• Experimento		



Estatística Econômica II - 2018 27



Estudo Observacional

- Observam os resultados para os sujeitos disponíveis nas variáveis
 - √ Pesquisador mede respostas, mas não controla a alocação dos sujeitos



Estatística Econômica II - 2018 28



Estudo Observacional - Características



- Não é possível estabelecer causa e efeito com estudo não experimental
 - √ Com amostra disponível ou com levantamento de dados
 - √ Alguma variável não mensurada pode ser a responsável por padrão nos dados
- Com esse tipo de estudo existe forte possibilidade de que a amostra não expresse muito bem a população

Estatística Econômica II - 2018 29



- Comparação de grupos:
 - √ Difícil com esse tipo de estudo
 - √ Grupos podem estar desequilibrados nas variáveis que afetam os resultados (mesmo com amostragem aleatória)
 - √ Variáveis chave podem não ter sido mensuradas no estudo

Estatística Econômica II - 2018 30





Exemplo

Estudo prospectivo sobre efeitos do fumo (1951)



- √ Hábitos de fumo dos sujeitos são estabelecidos no início do estudo
- √ Sujeitos da pesquisa
 - 34.440 médicos que responderam 60.000 questionários enviados
- √ Grupos:
 - Fumantes e não-fumantes

Estatística Econômica II - 2018 31



- √ Participantes foram acompanhados durante 11 anos
 - Taxas de morte por câncer e algumas outras causas nos dois grupos
- Resultados:
 - √ Em observações para cada 100.000 homens:
 - √ Câncer de pulmão:
 - Não-fumantes: 7 mortes por ano
 - Fumantes: 166 mortes por ano
 - (24 vezes maior)



Estatística Econômica II - 2018 32



- √ Ataque cardíaco:
 - Não-fumantes: 422 mortes por ano
 - Fumantes: 599 mortes por ano
 - (177 mortes a mais)
- √ Diferenças:
 - Câncer de pulmão: 159 mortes a mais
 - Ataque cardíaco: 177 mortes a mais

33



Estatística Econômica II - 2018



- Evidências não serviram como prova de que fumar era a causa de mortes a mais
- Grupos:
 - √ Tratamento:
 - Apenas pessoas que queriam fumar
 - √ Grupo de controle:
 - Pessoas que decidiram não fumar
 - √ Não havia mecanismo para assegurar que os grupos fossem essencialmente idênticos em todas as outras características, exceto por seu hábito de fumar

34

Estatística Econômica II - 2018





Estudos Observacionais e Causação

- Frequentemente útil na identificação das possíveis causas dos efeitos
- Não pode estabelecer causação de maneira adequada
- Somente experimentos desenhados e executados adequadamente podem demonstrar causação de maneira confiável

35

Estatística Econômica II - 2018





Observação versus Experimento

- Estudo observacional:
 - √ Investiga os indivíduos (sujeitos) e mede as variáveis de interesse, mas não tenta influenciar as respostas
- Experimento
 - √ Impõe deliberadamente algum tipo de tratamento sobre indivíduos, a fim de observar suas respostas

36



Estatística Econômica II - 2018



Questões Importantes



1. Qual é a maior diferença entre um estudo observacional e um experimento?
 - √ E sobre uma pesquisa (similar a uma pesquisa de opinião pública) e um experimento?

Estatística Econômica II - 2018 37





2. Pesquisa de opinião com amostra de 1.600 pessoas (selecionadas apropriadamente).
 - √ Interesse: saber se as pessoas assistem regularmente o Programa do Faustão
 - √ 800 pessoas da amostra (50%) respondem que SIM

Estatística Econômica II - 2018 38





- Qual a distância que essa percentagem fica da proporção populacional que assiste o programa?
 - √ 30%? 10%? 5%? 1%
 - √ Ou é exatamente igual?

Estatística Econômica II - 2018 39





3. Programa de TV faz uma pergunta e promove pesquisa de opinião por telefone.
 - √ É pedido para os espectadores ligarem para um determinado número para responder Sim e para outro número para responder Não
 - Os resultados desse tipo de pesquisa representam a opinião da comunidade?
 - √ Representam as opiniões de todos que estão assistindo aquela TV no momento? Outro grupo?

Estatística Econômica II - 2018 40



- Provavelmente irão ligar as pessoas com opiniões fortes sobre o assunto
 - √ Especialmente aquelas com opiniões ferrenhamente contrárias (ou favoráveis)



Estatística Econômica II - 2018 41



4. Você tem uma lista das empresas da cidade, classificada por ramo de negócio

- √ Você deseja telefonar para 100 dessas empresas para obter uma amostra representativa de todas as empresas da cidade.
- √ Como você escolheria as 100 empresas?
- √ Seria boa ideia usar simplesmente as 100 primeiras empresas listadas?



Estatística Econômica II - 2018 42



5. Institutos de pesquisa frequentemente publicam pesquisas com amostras de tamanho 1.200 (ou valor similar)

- √ Como os indivíduos são selecionados?



Estatística Econômica II - 2018 43



Survey



- Pesquisa por amostragem:
 - √ Coleta de dados de amostra de população
 - Entrevista pessoal ou telefone
 - Questionário autoaplicado
 - Etc.
- Resultados da amostra representam a população
 - √ Será verdade se amostra for escolhida corretamente

Estatística Econômica II - 2018 44



- Não há controle dos sujeitos da pesquisa
 - √ Sujeitos simplesmente respondem às perguntas (ou características são medidas)



Estatística Econômica II - 2018 45



Exemplos – Pesquisas Amostrais

- PNAD – Pesquisa Nacional por Amostra de Domicílios
- PME – Pesquisa Mensal de Emprego
- INPC – Índices de Preço ao Consumidor
- PAC – Pesquisa Anual de Comércio
- PMC – Pesquisa Mensal de Comércio
- PIA – Pesquisa Industrial Anual

Estatística Econômica II - 2018 46





Exemplo – SAEB

SAEB – Sistema de Avaliação da Educação Básica

- Objetivo:
 - √ Produzir estimativas da proficiência média dos alunos para grupos definidos
 - √ Estratos:
 - Séries: 4ª. e 9ª. (EF) e 3ª. (EM)
 - Estados
 - Dependência administrativa: Particular, estadual, municipal
 - Zona de localização: Capital, interior



Estatística Econômica II - 2018 47



Experimento Planejado



- Objetivo:
 - √ Comparar respostas dos sujeitos, sob condições diferentes (tratamentos)
 - √ Combinação dos níveis de variáveis que podem influenciar o resultado
- Pesquisador atribui os sujeitos às condições (aos tratamentos)
 - √ Controla o experimento

Estatística Econômica II - 2018 48





- Mede efeito no resultado de característica modificada (controlada) pelo pesquisador
 - √ Característica: variável explicativa
 - √ Resultado: variável resposta
 - √ Ex.:
 - receber um tipo de tratamento médico, fazer dieta especial, passar por programa de treinamento, etc.
- Maioria dos experimentos com pessoas envolve voluntários (candidaturas)
 - √ Não há como forçar a participação

Estatística Econômica II - 2018 49



- Delineamento experimental:
 - √ Aleatorização do tratamento que o sujeito recebe



Estatística Econômica II - 2018 50



Exemplo



- Analisar se a ingestão regular de uma aspirina reduz a mortalidade por doença do coração
 - √ Amostra: 22.000 médicos do sexo masculino
 - √ Grupos: aspirina/placebo
 - √ Tempo do estudo: 5 anos
 - √ Comparação das taxas de ataques cardíacos no dois grupos
 - √ Alocação aleatória

Estatística Econômica II - 2018 51



- √ Aleatorização permitiu equilíbrio aproximado dos fatores que poderiam afetar as taxas de ataque cardíaco (idade, hábitos alimentares, histórico familiar, etc.)

Estatística Econômica II - 2018 52





Experimento x Pesquisa

Experimento	Pesquisa
- Intervenção	- Observação controlada
- População infinita	- População finita
- Modelagem e teste de hipóteses	- Estimação de medidas descritivas

Estatística Econômica II - 2018

53





- Quer o estudo seja experimental ou não, é importante incorporar a aleatorização em todo estudo em que se planeja fazer inferências.

Estatística Econômica II - 2018

54

Amostragem





População

- População alvo:
 - √ População para a qual se gostaria de obter informação
- População de pesquisa:
 - √ População a ser realmente coberta pela pesquisa

Estatística Econômica II - 2018



56



Exemplo – PNAD

População alvo	População de pesquisa
– População residente no Brasil em data de referência	– População residente no Brasil em data de referência, excluídos os habitantes de setores rurais da região Norte (exceto Tocantins).



Estatística Econômica II - 2018 57



População de Pesquisa

- Especificação:
 - √ Unidades a serem pesquisadas
 - √ Características definidoras das unidades
 - √ Localização espacial das unidades
 - √ Período de referência considerado
 - √ Vinculação a cadastros

Estatística Econômica II - 2018 58





Exemplo

ECINF – Pesquisa Economia Informal Urbana

- População alvo:
 - √ “Pesquisa por amostragem de domicílios situados em áreas urbanas, onde se busca identificar os trabalhadores por conta própria e empregadores com até 5 empregados que desenvolvam atividades não agrícolas”

(Manual)



Estatística Econômica II - 2018 59



- √ “pertencem ao informal todas as unidades econômicas de propriedade de trabalhadores por conta própria e de empregadores com até 5 empregados, moradores de áreas urbanas, sejam elas a atividade principal de seus proprietários ou atividades secundárias.”



(IBGE, Economia Informal Urbana)

Estatística Econômica II - 2018 60

Unidades	Características Definidoras	Localização	Período
Pessoas	Habitando domicílios particulares permanentes	Bicas	Durante a semana da pesquisa
Empresas de comércio varejista	Classificadas como supermercados	Belo Horizonte	Em 2016
Pessoas	Maiores de 5 anos de idade	Que visitaram o Hopi Hari	Entre 01/09 a 31/12/2016
Alunos	Do curso de mestrado de administração	Juiz de Fora	2º semestre de 2016
Estabelecimentos agropecuários	Produtores de cachaça	Em Minas Gerais	Em 2016



Estatística Econômica II - 2018 61

Censo e Amostragem

- Censo:
 - √ coleta de informações sobre todas as unidades da população
 - Censo Escolar, Prova Brasil, Censo do IBGE.
- Pesquisa por amostragem:
 - √ coleta de informações sobre uma parte da população.
 - √ Exemplo: PNAD



Estatística Econômica II - 2018 62

Censo ou Amostragem?



- Em igualdade de condições:
 - √ censo produz resultados mais precisos que amostra
- Consideradas restrições orçamentárias:
 - √ Amostra pode produzir resultados mais informativos e precisos

Estatística Econômica II - 2018 63



- Censo:
 - √ Recomendado quando:
 - população é pequena
 - erros amostrais são grandes
 - Necessárias informações sobre todas as unidades
 - √ Vantagens:
 - Estudo de populações muito pequenas
 - Informações requeridas para pequenas áreas geográficas ou classificações muito detalhadas
 - Permite construção de cadastros

Estatística Econômica II - 2018 64



- Amostragem
 - √ Deve ser utilizada quando:
 - população é grande e/ou os custos é alto e/ou o tempo disponível é curto
 - obtenção da informação desejada implica na destruição da unidade pesquisada



Estatística Econômica II - 2018 65



Exemplo

- **Que anúncio de TV venderá mais pasta de dente?**
 - √ Não é possível perguntar a todos os consumidores sobre seus planos de consumo.



Estatística Econômica II - 2018 66



Conceitos



- **Unidade:**
 - √ Sujeito a ser medido ou observado
- **População:**
 - √ Conjunto de todas as unidades para o qual desejamos obter informações
- **Amostra:**
 - √ Unidades selecionadas para medir/observar

Estatística Econômica II - 2018 67





- **Cadastro:**
 - √ Lista de unidades de onde a amostra é selecionada
- **Variável de pesquisa:**
 - √ Característica de interesse, relativa a cada elemento da população pesquisada, mas que é observada apenas na amostra.
- **Plano amostral:**
 - √ Descrição de métodos e medidas para execução da amostragem.

Estatística Econômica II - 2018 68



- **Parâmetro:**
 - √ Alguma característica da população.
 - (Função do conjunto de valores da população)
- **Estimativa:**
 - √ Qualquer valor calculado tendo como base os dados provenientes de uma amostra



Estatística Econômica II - 2018 69



Definições Necessárias à Amostragem



- **Unidade de referência:**
 - √ Unidade de observação ou sobre a qual são obtidas informações de interesse
- **Unidade informante:**
 - √ Unidade que fornece a informação
- **Unidade de análise:**
 - √ Unidade à qual a inferência é dirigida

Estatística Econômica II - 2018 70



- **Unidade de amostragem:**
 - √ Unidade selecionada para a amostra
- **Domínio de análise (interesse):**
 - √ Grupo de unidades de análise agregadas com a finalidade de tabulação, inferência e análise



Estatística Econômica II - 2018 71



Exemplo – PNAD

- **Unidade(s) de referência:**
 - √ Pessoas, dentro de domicílios
- **Unidade informante:**
 - √ Pessoa adulta, moradora no domicílio
- **Unidade(s) de análise:**
 - √ Pessoas, famílias, domicílios



Estatística Econômica II - 2018 72



- Unidade(s) de amostragem:
 - √ Município, setor, domicílio
- Domicílios de análise:
 - √ 27 estados, regiões metropolitanas, país

Estatística Econômica II - 2018

73





Exemplo – POF

- Objetivo:
 - √ Informações sobre renda e despesa familiar
- Unidades de referência:
 - √ Pessoas dentro das famílias, famílias
- Unidade informante:
 - √ Pessoa de referência da família
- Unidade de análise:
 - √ Família
- Unidade amostral:
 - √ Setor, domicílio

Estatística Econômica II - 2018

74





Cadastro

- Qualquer material ou instrumento usado para identificar e obter acesso aos elementos da população-alvo.
- Unidades amostrais
 - √ Processo físico de aleatorização do plano amostral, aplicado às unidades do cadastro.
 - √ Pode ser um elemento ou um conjunto (conglomerado) de elementos da população

Estatística Econômica II - 2018

75



Tipos de Cadastros

- a) Cadastros compostos por listagem de elementos da população alvo
 - √ permite utilização de planos amostrais com seleção direta de elementos
- b) Cadastros compostos por listagem de conjuntos de elementos da população alvo
 - √ Para observar um elemento, é preciso fazer uso de planos amostrais em dois ou mais estágios
 - Comum em Pesquisas Educacionais

Estatística Econômica II - 2018

76

Escolha do Cadastro

- Critérios:
 - √ Inclusão de todos os elementos da população
 - √ Devem estar listadas apenas uma vez
 - √ Não deve incluir elementos que não pertencem à população alvo

77

Exemplo de Cadastro

- Base Operacional Geográfica do IBGE
 - √ Lista de unidades geo-estatísticas:
 - setores, subdistritos, distritos, municípios, microrregiões, unidades da federação;
 - √ Usado no Censo Demográfico e em pesquisas domiciliares e agropecuárias
 - PNAD, PME, etc.)

78

Amostragem

- “Amostragem consiste em selecionar parte de uma população para observar, de modo que seja possível estimar alguma coisa sobre toda a população”

(Steven K. Thompson)

79

Erros em Levantamento Amostrais

- Erro amostral:
 - √ Associado ao fato de a variável de interesse ser observada apenas na amostra.

Amostra 1: 1 bola preta

Amostra 2: 2 bolas pretas

Amostra 3: Nenhuma bola preta

Fonte: Vieira, M. T.; Silva, P. L. N., 2017.

80

• Erro não amostral:

- √ Associado ao processo operacional de um levantamento amostral

POPULAÇÃO

Cadastro

Amostra: 2 bolas pretas

erro de cobertura do cadastro

possível erro de processamento

Fonte: Vieira, M. T.; Silva, P. L. N., 2017.

Estatística Econômica II - 2018

81

• Tipos de erros não-amostrais:

- √ Erros de especificação na fase do planejamento amostral;
- √ Erros de cobertura do cadastro;
- √ Erros de processamento;
- √ Erros de mensuração
- √ Erros de não-observação.

Erro total = erro amostral + erro não-amostral

Estatística Econômica II - 2018

82

• **Como obter uma boa amostra para ser usada em uma pesquisa?**

Estatística Econômica II - 2018



83

Características Desejáveis de uma Amostra

- Capacidade de generalização:
 - √ Da estimativa amostral para toda a população
- Não tendenciosidade
 - √ “imparcialidade”
- Apresentar menor erro amostral possível
 - √ Considerados custo, tempo e restrições operacionais
- Capacidade de medir a precisão das estimativas

Estatística Econômica II - 2018



84



Importante

- Se esses critérios não forem satisfeitos, as amostras não permitirão generalizar inferências para a população
- Exemplos de amostras inadequadas
 - √ Amostra por conveniência (acidental)
 - √ Amostra por resposta voluntária
 - √ Amostras por julgamento (intencionais)
 - √ Amostras por quotas



Estatística Econômica II - 2018 85



Amostragem por Conveniência

- O pesquisador seleciona membros da população mais acessíveis
- Não é possível avaliar a qualidade da amostra em termos de sua representatividade da população



Estatística Econômica II - 2018 86



Amostra de Resposta Voluntária

- Pessoas escolhem a si próprias para responder algo que foi indagado
- São amostras enviesadas
 - √ Pessoas mais prováveis de responder
 - aquelas com opiniões fortes, especialmente as com opiniões bastante negativas



Estatística Econômica II - 2018 87



Amostragem por Julgamento

- O pesquisador usa o seu julgamento para selecionar os membros da população que são boas fontes de informação
 - √ Pressupõe que elas têm a propriedade de influenciar a opinião dos demais
 - √ Exemplo:
 - Pesquisa de opinião dos diretores de faculdades de comunicação sobre a extinção da obrigatoriedade do diploma universitário para o exercício da profissão de jornalista.



Estatística Econômica II - 2018 88



Amostragem por Quotas



- Amostragem por quotas pressupõe 3 etapas:
 - √ Classificação da população de acordo com características que se presume serem relevantes
 - √ Determinação da proporção da população que deve ser colocada em cada estrato
 - √ Entrevistador escolhe pessoalmente número predefinido sujeitos em cada categoria.

Estatística Econômica II - 2018 89



- Técnica não probabilística bastante utilizada em levantamentos de mercado, prévias eleitoral, sondagem de opinião pública



Estatística Econômica II - 2018 90



Amostragem Probabilística

- Amostra escolhida ao acaso
 - √ Possível definir o conjunto de todas as amostras possíveis de serem selecionadas
 - √ Possível calcular a probabilidade de seleção de cada uma dessas amostras

Estatística Econômica II - 2018 91



- Aleatorização da amostra
 - √ Amostra é selecionada usando mecanismo de sorteio
 - Amostra s é escolhida com probabilidade $p(s)$
 - √ Todas as unidades da população tem probabilidade não nula (> 0) de serem selecionadas.
 - √ Consideração dos pesos amostrais na estimação dos parâmetros de interesse
 - Probabilidade de inclusão das unidades selecionadas

Estatística Econômica II - 2018 92

Margem de Erro

- Expressa a quantidade de erro amostral aleatório nos resultados de uma pesquisa
- Característica importante da amostragem probabilística:
 - √ Estimação de parâmetros populacionais com margem de erro conhecida ou estimada

Estadística Econômica II - 2018 93

Amostragem Probabilística

- Amostragem aleatória simples
- Amostragem aleatória estratificada
- Amostragem por conglomerados
- Amostragem por captura e recaptura

Estadística Econômica II - 2018 94



Amostragem Aleatória Simples

- Seleção de n unidades de uma população (cadastro, lista) com N unidades
 - √ Todas as amostras possíveis (de tamanho n) têm mesma probabilidade de ser selecionada
 - √ Unidades são selecionadas sem reposição
 - √ Unidades têm mesma probabilidade de serem selecionadas (n/N)

Estadística Econômica II - 2018 95

Fonte: Vieira, M. T.; Silva, P. L. N., 2017.

Estadística Econômica II - 2018 96




- Amostra s tem tamanho fixo $n \leq N$.
 - √ Todas as amostras de tamanho n têm a mesma probabilidade de serem escolhidas.
 - √ Quantidade de amostras distintas

$$\binom{N}{n} = \frac{N!}{n!(N-n)!}$$
 - √ Probabilidade de seleção da amostra s :

$$p(s) = \frac{1!}{\binom{N}{n}}$$
 - √ Probabilidade de seleção de cada unidade populacional: $\frac{1}{N}$

97

Estatística Econômica II - 2018







- √ Tamanho da População: $N = 48$
- √ Tamanho da amostra: $n = 7$
- √ Quantidade de amostras possíveis: $\binom{48}{7} = 73.629.072$
- √ Probabilidade de seleção de cada unidade populacional: $7/48$

$$\frac{7}{48} = 0,146$$

98



Estatística Econômica II - 2018

- Comentários:
 - √ Método mais simples, porém o mais importante
 - √ Procedimento é a base para outros esquemas amostrais mais complexos.
 - √ Principais ideias de amostragem são desenvolvidas com ele

102

Estatística Econômica II - 2018






AAS – Vantagens

- Vantagens:
 - √ Simplicidade
 - √ Fácil estimar parâmetros da população e medir a precisão das estimativas
 - √ Propriedades conhecidas e testadas
- Implementação:
 - √ Cada unidade da população tem de estar cadastrada
 - √ Requer mecanismo de sorteio aleatório



103

Estatística Econômica II - 2018



- Questões importantes:
 - √ Como fazer AAS da população de um país?
 - √ Onde encontrar listas completas de pessoas da população
- Na prática, amostras aleatórias simples são difíceis de serem implementadas



Estatística Econômica II - 2018 104



Amostragem Probabilística – Outros Métodos

- Amostragem aleatória estratificada
- Amostragem com probabilidades desiguais
- Amostragem aleatória por conglomerados
- Amostragem aleatória sistemática



Estatística Econômica II - 2018 105



Amostragem Aleatória Estratificada

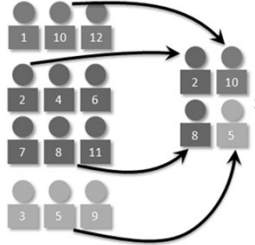
- Princípios Gerais:
 - √ População em estratos (mutuamente exclusivos e exaustivos) de acordo com características conhecidas.
 - √ Selecionadas unidades ao acaso em cada estrato
 - √ Em geral, os estratos são mais homogêneos que a população
 - Redução do erro amostral global.
 - Plano amostral será mais eficiente à medida em que os subgrupos forem mais homogêneos

Estatística Econômica II - 2018 106



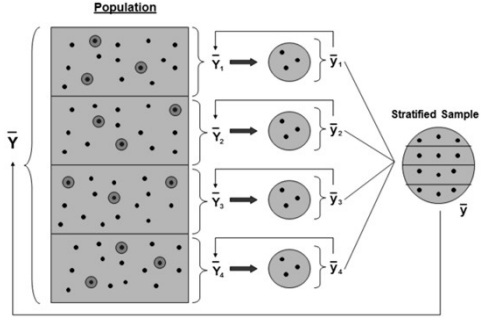
- Procedimento:
 1. Dividir a população em grupos homogêneos (estratos)
 2. Selecionar AAS de cada estrato
 3. Estimar parâmetros de cada estrato
 4. Combinar as estimativas para toda a população

Estatística Econômica II - 2018 107



- População dividida em 2 ou mais estratos
 - √ Grupos homogêneos com característica comum
- Amostra aleatória simples (ou outra) em cada estrato
- Sub-amostras combinadas em uma amostra total

108



Fonte: Vieira, M. T.; Silva, P. L. N., 2017.

109

Amostragem Aleatória Estratificada - Vantagens



- √ Possível obter estimativas de cada estrato
- √ Estimativas da população podem ser mais precisas, se a variável medida é mais homogênea dentro dos estratos
 - Permite também redução do tamanho da amostra para um nível de precisão fixo
- √ Coleta de menor custo se os estratos estão espalhados geograficamente
 - Pode ser desejável utilizar entrevistadores distintos nos diversos estratos

110

Amostragem Estratificada – Dificuldades

- √ Necessário conhecimento prévio das variáveis de estratificação para todas as unidades do cadastro
- √ Requer a reestruturação do cadastro antes da amostragem



111



Motivos para Estratificar

- Grupos naturais de interesse como estratos
 - √ Exemplo: regiões geográficas
- Grupos substantivos de interesse como estratos:
 - √ Exemplo: ramos de atividade econômica, etc.
- ‘Distribuição’ maior da amostra sobre toda a população
- Melhoria da eficiência amostral
 - √ Redução da variância amostral



Estatística Econômica II - 2018 112



Exemplo

- Pesquisas Econômicas Anuais do IBGE
 - √ Plano amostral:
 - Amostragem estratificada por corte de empresas, com 2 estratos de tamanho em cada estrato natural (UF x classe de atividade econômica).
- Estrato certo (Pessoal ocupado – $PO \geq 20$)
 - √ Todas as empresas são pesquisadas;
- Estrato amostrado
 - √ amostra selecionada.



Estatística Econômica II - 2018 113



Amostragem com Probabilidades Desiguais



- Planos amostrais anteriores consideram mesma probabilidade de seleção de todas as amostras possíveis
- É possível trabalhar com esquemas probabilísticos mais gerais
 - √ Unidades de amostragem têm variação de tamanho.
 - Não considerar essa variação de tamanho pode resultar em desenhos ineficientes.

Estatística Econômica II - 2018 114



- Quando utilizar:
 - √ Variação dos tamanhos for grande.
 - √ Disponível informação auxiliar precisa sobre tamanhos
 - √ Tamanho fortemente correlacionado com variáveis de interesse.
- Solução possível:
 - √ Amostragem com probabilidades proporcionais ao tamanho (PPT)



Estatística Econômica II - 2018 115

- Modelo probabilístico da PPT:
 - √ População: $U = \{1, 2, \dots, N\}$
 - √ x_i : variável auxiliar, $\forall i \in U$
 - Valores conhecidos para todos os elementos da população
 - Pode ser usada como medida de tamanho das unidades populacionais, se $x_i > 0, \forall i \in U$
 - √ π_i : probabilidade de seleção da unidade i
 - √ π_{ij} : \forall : probabilidade de seleção da unidade i , dado que a unidade j tenha sido escolhida

116



Estatística Econômica II - 2018

- √ Em geral, pode-se assumir que é possível selecionar amostras tal que:
 - $\pi_i \propto x_i, \forall i \in U$
 - $\pi_{ij} > 0, \forall i \neq j \in U$

117



Estatística Econômica II - 2018

- Procedimento:
 - √ População com N unidades
 - Podem ser conglomerados.
 - √ Medida de tamanho x_i associada a elemento i
 - √ Exemplos:
 - Unidades populacionais: hospitais
Medida de tamanho: número de leitos
 - Unidades populacionais: escolas
Medida de tamanho: número de alunos matriculados (ou turmas)
 - Unidades populacionais: indústrias
Medida de tamanho: número de empregados (ou faturamento.)

118

Estatística Econômica II - 2018



- Importante:
 - √ Medida de tamanho (X) correlacionada com variável de interesse (Y)
 - Amostragem PPT é mais eficiente que AAS.
 - √ Condição para o estimador de total não ser enviesado

$$w_i = \frac{1}{\pi_i}, i \in U$$

 - w_i : pesos das unidade i na amostra
 - π_i : probabilidade de inclusão da unidade i na amostra

119



Estatística Econômica II - 2018

- Casos particulares
 - √ Se $\pi_i = \frac{n}{N}$ e $\pi_{ij} = \frac{n(n-1)}{N(N-1)}$
 - Obtém-se as expressões válidas para caso AAS.
 - √ Se $\pi_i = \frac{n}{N}$
 - Obtém-se as expressões válidas para caso AES.

120

Estatística Econômica II - 2018






Pesos Amostrais

- Em geral,
 - √ Número de unidades populacionais que cada unidade amostral representa ($\frac{n}{N}$)
- Peso amostral inicial:
 - √ Igual ou proporcional ao inverso da probabilidade de seleção.
- Peso amostral final pode incorporar:
 - √ Ajustes para não-resposta
 - √ Uso de informações auxiliares

121

Estatística Econômica II - 2018






Amostragem Aleatória por Conglomerados

- Plano amostral em múltiplos estágios
 - √ Divisão da população em muitos conglomerados
 - Ex.: bairros de cidades
 - √ Seleção de AAS dos conglomerados
 - Unidade amostral: conglomerado
 - Cadastro é de conglomerados
 - Contém mais de um elemento populacional.
 - √ Amostra:
 - Sujeitos selecionados dos conglomerados

122


Estatística Econômica II - 2018


- Problemas:
 - √ Dentro de um mesmo conglomerado, unidades tendem a ter valores parecidos para as variáveis pesquisadas.
 - Planos amostrais deste tipo tendem a ser menos eficientes
 - √ Procedimento é menos eficiente quanto maior for a correlação dentro dos conglomerados

123

Estatística Econômica II - 2018




- Quanto maior a heterogeneidade dentro dos conglomerados, maior é a eficiência do procedimento amostral
 - √ Isto é o oposto do exigido na construção dos estratos!




124

Estatística Econômica II - 2018




Amostragem por Conglomerados em um Estágio



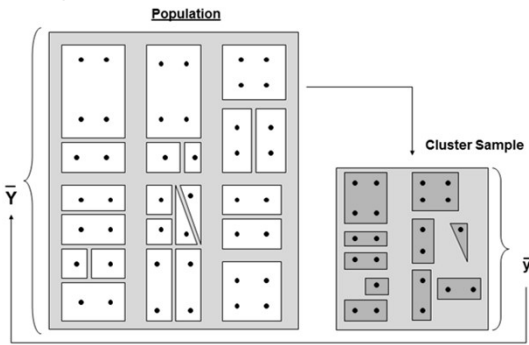
- Seleção de amostra de conglomerados de acordo com um plano amostral qualquer
- Amostra é composta por todos os elementos pertencentes aos conglomerados

125

Estatística Econômica II - 2018




- Amostragem por conglomerados em um estágio




Fonte: Vieira, M. T.; Silva, P. L. N., 2017.

126

Estatística Econômica II - 2018



Importante:



- Se os conglomerados são homogêneos
 - √ Não é aconselhável o uso de todos os elementos dos conglomerados selecionados
- Solução:
 - √ Sortear elementos dos conglomerados escolhidos

127

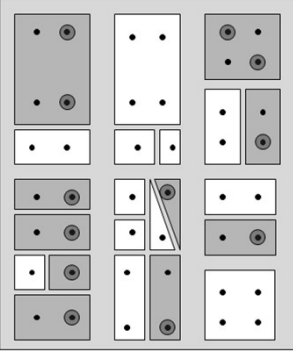
Estatística Econômica II - 2018

Amostragem por Conglomerados em Múltiplos Estágios

- Procedimento em 2 estágios:
 - √ 1º estágio: seleção de conglomerados
 - Unidades primárias de amostragem – UPA's
 - √ 2º estágio: sortear elementos de cada conglomerado selecionado:
 - Unidades secundárias de amostragem – USA's
- √ Podem ser usados planos amostrais distintos em cada um dos estágios

Estatística Econômica II - 2018 128

• Amostragem por conglomerados em 2 estágios.



Fonte: Vieira, M. T.; Silva, P. L. N., 2017. Estatística Econômica II - 2018 129



Exemplos

- 1) Conglomerados: áreas num mapa
Unidades elementares: fazendas
- 2) Conglomerados: hospitais
Unidades elementares: enfermeiras
- 3) Conglomerados: carros cruzando pedágio
Unidades elementares: pessoas nos carros

Estatística Econômica II - 2018 130

- 4) Unidades primárias – UPA's: indústrias
unidades elementares: unidades locais
- 5) Unidades primárias – UPA's: escolas
Unidades secundárias – USA's: turmas
Unidades elementares: crianças



Estatística Econômica II - 2018 131



Exemplo – CPS

- CPS – Current Population Survey
- Plano de amostragem em múltiplos estágios:
 - √ Estágio 1:
 - Dividir país em 2007 Unidades Primárias de Amostragem (UPA)
 - Selecionar amostra de 754 UPAs
 - (428 com as maiores populações e amostra estratificada de 326 das outras unidades)

Estatística Econômica II - 2018 134





√ Estágio 2:

- Dividir cada UPA selecionada em pequenas áreas (blocos censitários)
- Estratificar os blocos em relação às raças e outras informações relevantes
- Selecionar amostra estratificada de blocos em cada UPA

√ Estágio 3:

- Em cada bloco, agrupar as unidades habitacionais em conglomerados de 4 unidades próximas entre si
- Entrevistar os domicílios pertencentes a uma amostra aleatória desses conglomerados



Estatística Econômica II - 2018 135



Amostragem Aleatória Sistemática

- Procedimento:
 - √ Período da amostragem (k): $k = n/N$
 - √ Seleciona um sujeito ao acaso dos k primeiros nomes da população (cadastro)
 - √ Seleciona cada k -ésimo sujeito listado depois daquele.



Estatística Econômica II - 2018 136



Amostra Aleatória Sistemática



- Vantagens:
 - √ É mais fácil de ser selecionada pois usa somente um número ao acaso
 - √ Pode fornecer uma boa representação da população
 - Em listas em ordem alfabética os valores de cada variável flutuam aleatoriamente?

Estatística Econômica II - 2018 137



- As amostras de tamanho n não são igualmente prováveis
 - √ Dois sujeitos listados próximos um do outro na lista não podem ambos, aparecer na amostra



Estatística Econômica II - 2018 138



Amostra – Vantagens

- Custo menor
- Velocidade maior
- Precisão controlada
- Redução da carga de coleta sobre unidades da população
- Necessidades especializadas
 - √ Quando censo é impossível



Estatística Econômica II - 2018 141



Amostragem

- Uso de métodos apropriados de amostragem para selecionar 1.500 adultos de uma população com milhões de adultos
 - √ Estimação de proporção de característica de interesse da população, com erro inferior a 3%
 - √ Não depende do tamanho da população mas do tamanho da amostra



Estatística Econômica II - 2018 142



Margem de Erro

- Estimação de proporção da população que tem característica de interesse (ou opinião)
 - √ Pesquisas conduzidas apropriadamente oferecem estimativas com grande precisão
 - √ Medida de precisão:
 - Margem de erro
 - √ Grau de confiança da estimacão
 - Ex.: Diferença da proporção amostral para a proporção populacional é maior que a margem de erro 5% das vezes (confiança de 95%)
 - Menos de 1 vez a cada 20 pesquisas

Estatística Econômica II - 2018 143






• Regra geral:

- √ Proporção obtida com a amostra raramente vai diferir da proporção populacional mais que $\frac{1}{\sqrt{n}}$
- √ Ou seja, margem de erro = $\frac{1}{\sqrt{n}}$

144

Estatística Econômica II - 2018






Margem de Erro – Proporção

Tamanho amostral (n)	Margem de erro $1/\sqrt{n}$
100	10,0 %
400	5,00 %
1.600	2,50 %
3.200	1,77 %
6.400	1,25 %

145

Estatística Econômica II - 2018






Passos da Amostragem

1. Definição dos objetivos, conceitos e recursos
2. Obtenção e avaliação do cadastro
3. Planejamento, seleção e controle da amostra
4. Estimação das quantidades de interesse
5. Avaliação da qualidade das estimativas

146

Estatística Econômica II - 2018





Plano Amostral

- Descrição de métodos e medidas para efetuar a seleção dos elementos amostrais.
 - √ Necessária a definição prévia da população de pesquisa.
- Importante:
 - √ Planejamentos deficientes podem gerar conclusões enganadoras

147



Estatística Econômica II - 2018



Planejamento da Amostragem

- Detalhamento de:
 - √ Objetivos da pesquisa
 - √ Exigências de dados
 - √ Definições conceituais
 - √ Definições operacionais
 - √ Recurso disponíveis
 - Tempo é recurso quase sempre escasso

Estatística Econômica II - 2018 148





Exemplo – PNAD

PNAD – Pesquisa Nacional por Amostra de Domicílios

- Objetivo:
 - √ Produção de informações básicas para o estudo do desenvolvimento socioeconômico do país.

Estatística Econômica II - 2018 149




Exemplo – POF

POF – Pesquisa de Orçamentos Familiares

- Objetivo:
 - √ Estimar o rendimento mensal bruto total dos chefes de domicílio em cada área de pesquisa, com erro amostral máximo admissível inferior a 5%
 - √ Áreas cobertas: 11 regiões metropolitanas de primeiro nível e DF.

Estatística Econômica II - 2018 150



Plano Amostral

- Detalhamento de:
 - √ População alvo
 - √ População de pesquisa e cadastros
 - √ Unidade(s) de amostragem
 - √ Unidades de informação (de pesquisa)
 - √ Método(s) para seleção da amostra
 - √ Aspecto longitudinal (medidas repetidas)

Estatística Econômica II - 2018 151

Cuidados com as Pesquisas por Amostragem

Conceitos



- Parâmetro:
 - √ Resumo numérico da população
- Estatística:
 - √ Resumo numérico dos dados amostrais

Exemplo

- Pergunta dirigida por colunista a seus leitores:
 - √ “Se você tivesse de começar tudo de novo, teria filhos?”
 - √ 70% dos que responderam disseram que não
 - √ Conclusão: “70% dos pais respondem que não
 - √ Amostra foi autosselecionada
 - Pessoas que se dispuseram a responder
 - √ Pesquisa posterior planejada estatisticamente
 - 91% teriam filhos novamente

Erro Amostral



- Erro amostral de uma estatística
 - √ Diferença entre o seu valor e o parâmetro populacional
- Na prática, é desconhecido
 - √ Os valores dos parâmetros populacionais são desconhecidos
- AAS protege contra o viés
 - √ Erro amostral tende a flutuar em torno de zero



Vício Amostral

- Fontes potenciais de tendenciosidade em amostragem:
 - √ Tendenciosidade amostral
 - √ Tendenciosidade na resposta
 - √ Tendenciosidade da não resposta



Estatística Econômica II - 2018 156



Tendenciosidade Amostral

- Amostragem não probabilística:
 - √ Amostra pode não ser representativa da população
 - √ Não é possível determinar as probabilidades das possíveis amostras
 - √ Inferências usando tais amostras têm confiabilidade desconhecida
 - Conclusões podem estar equivocadas
 - √ Amostras grandes não atenuam o viés amostral



Estatística Econômica II - 2018 157



Subcobertura (sub-representação):

- Subcobertura (sub-representação):
 - √ Alguns grupos populacionais são excluídos do processo de seleção da amostra
 - √ Pode ocorrer mesmo com a amostragem aleatória
 - √ Exemplo: levantamento de dados por telefone
 - Não alcançará pessoas que não podem ter (ou não têm) telefone
 - Famílias que têm muitos telefones tenderão a ser representados em excesso



Estatística Econômica II - 2018 158



Tendenciosidade na Resposta

- O sujeito dá uma resposta incorreta (talvez minta)
- Respostas podem ser influenciadas por:
 - √ Formulação da pergunta
 - √ Questões mal colocadas ou confusas
 - √ Características do entrevistador
 - √ Forma com que o entrevistador faz as perguntas
 - √ Ordem da pergunta



Estatística Econômica II - 2018 159



Exemplo

- As fraldas descartáveis deveriam deixar de ser produzidas?
 - √ Estima-se que as fraldas descartáveis sejam responsáveis por menos de 2% do conteúdo dos depósitos de lixo. Por outro lado, calcula-se que os vasilhames, as cartas publicitárias e os trastes domiciliares respondam por 21% desse lixo. Baseando-se nesses fatos, em sua opinião seria justo proibir as fraldas descartáveis?



Estatística Econômica II - 2018 160



Tendenciosidade da Não Resposta

- Sujeitos amostrados:
 - √ Não podem ser encontrados
 - √ Se recusam a participar
 - √ Não respondem a algumas perguntas

Estatística Econômica II - 2018 161





Exemplo

Levantamento de dados com mulheres nos EUA



- Conclusão:
 - √ 70 % das mulheres que estão casadas há pelo menos 5 anos têm casos extraconjugais
- Amostragem
 - √ Questionário enviado pelo correio
 - Devolvidos: 4.500
 - Enviados: 100.000

Estatística Econômica II - 2018 162



- As mulheres que devolveram (4,5%) são representativas da população?



Estatística Econômica II - 2018 163



Dados Faltantes

- *Missing data:*
 - √ Problema em quase todos os grandes estudos
 - √ Alguns sujeitos não fornecem respostas para algumas das variáveis mensuradas
- Maioria dos softwares ignora casos para os quais há observações faltantes para pelo menos uma variável
 - √ Desperdício de informação
 - √ Possível tendenciosidade



Estatística Econômica II - 2018 164



Dificuldades em Amostragem

- Usar cadastro errado para amostragem
- Não contatar os indivíduos selecionados
- Não conseguir as respostas ou conseguir somente respostas voluntárias
- Obter amostra pequena demais



Estatística Econômica II - 2018 165



Desastres em Amostragem

- Obter amostra de voluntários
- Obter amostra de conveniência ou sem controle
- Ignorar o processo de amostragem ao estimar para a população



Estatística Econômica II - 2018 166



Cuidados com a Amostra

- Em qualquer estudo, avaliação cuidadosa do alcance das conclusões
- Formule as conclusões criticamente:
 - √ Observe a constituição da amostra
 - √ Como a amostra foi selecionada?
 - √ Qual era seu tamanho?
 - √ Como foram formuladas as perguntas?
 - √ Quem patrocinou e conduziu a pesquisa?



Estatística Econômica II - 2018 167



- Quanto menos informação disponível, menos você deve confiar
- Seja cauteloso com qualquer estudo que faça inferências para uma população maior do que é justificado pela amostra escolhida
 - √ Cuidado com generalizações não suportadas pela amostragem

Estatística Econômica II - 2018 168

Referências



Bibliografia Recomendada

- AGRESTI, A.; FINLAY, B. *Métodos estatísticos para as ciências sociais*. Penso, 2012.
- MALHOTRA, N. K. *Pesquisa de marketing: uma orientação aplicada*. Bookman, 2002.
- MOORE, D. S.; MCCABE, G. P.; DUCKWORTH, W. M.; SLOVE, S. L. *A prática da estatística empresarial: como usar dados para tomar decisões*. LTC, 2006.
- VIEIRA, M. T.; SILVA, P. L. N. *Notas de aula: amostragem*. UFJF, 2017.

Estatística Econômica II - 2018 171