

UNIVERSIDADE DE LISBOA
INSTITUTO DE EDUCAÇÃO



Survey e análise estatística

Autore(s)

Catarina João Morgado

Fernando César dos Santos

Fernando Manuel Correia Maia

Flávio Alves da Silva

Gabriela Filipa Silva Ribeiro

Graciella Watanabe

Maria Margarida Urbano Geirinha Ramanhos

MESTRADO EM EDUCAÇÃO E

TECNOLOGIAS DIGITAIS

2014/2015

DECLARAÇÃO DE AUTENTICIDADE

Para o mestrado em causa, realizado para a Unidade Curricular Introdução à Análise de dados, do ano lectivo 2014/2015, o(s) autor(es) declara(m) que:

(i) Todo o conteúdo das páginas seguintes é de autoria própria, decorrendo do estudo, investigação e trabalho do(s) seu(s) autores.

(ii) Quaisquer materiais utilizados para produção deste trabalho não coloca em causa direitos de Propriedade Intelectual de terceiras entidades ou sujeitos.

(iii) Este trabalho, as partes dele, não foi previamente submetido como elemento de avaliação nesta ou em outra instituição de ensino/formação.

(iv) Caso o presente trabalho tenha sido desenvolvido em regime de trabalho de grupo, o que foi previamente definido ou acordado com os docentes da Unidade Curricular, não é submetido nenhuma versão que se revele totalmente igual ao trabalho de outro(s) grupo(s) de aluno(s).

(v) Foi tomado conhecimento das definições relativas ao regime de avaliação sobre o qual este trabalho será avaliado, pelo que se atesta que o mesmo cumpre as orientações que lhe foram impostas.

(vi) Foi tomado conhecimento que este trabalho deve ser submetido em versão digital, no espaço especificadamente criado para o efeito, e que essa versão poderá ser utilizada em actividades de detecção electrónica de plágio, por processos de análise comparativa com outros trabalhos, no presente e/ou no futuro.

(vii) O trabalho em causa apresenta-se, assim, de acordo com o regulamento de propriedade intelectual da Universidade de Lisboa (Despacho 45 2008, 28 de Outubro de 2008), encontrando-se sob a sua aplicação.

Data, 20 de Março de 2015

Assinatura

1. Pesquisa Survey

O método survey é uma pesquisa que busca investigar determinadas situações fazendo algumas questões como: O quê? Por quê? Como? Assim, procura-se através desse método determinar variáveis dependentes e independentes e a pesquisa realiza-se no momento presente e trata situações reais no local onde elas ocorrem (Silva, 2013).

Em geral, esse tipo de pesquisa possui interesses nas opiniões públicas, envolvem coleta de dados através, principalmente, de amostras selecionadas e utilizam como instrumento de coleta os questionários.

Os desafios a serem trilhados nesse tipo de pesquisa relacionam-se com custo, tempo e forma de garantir com que haja uma demanda pela resposta aceitável ao estudo. Para isso é necessário que o conceito a ser tratado na pesquisa esteja claro tanto para o investigador como para o investigado.

2. Construção do Survey

Plano de Investigação

A validação é um tema muito importante para a pesquisa survey, entendendo até que ponto se pode confiar em dados quantitativos. Para isso, no processo de entrevista devem ser evitados: modos de apresentar as questões tendenciosas, dúbios ou culturalmente desqualificados. Evita-se isso, também, escolhendo de modo cuidadoso a amostra a ser trabalhada, definindo critérios de seleção para a escolha dos entrevistados.

Um método possível de escolha de amostragem é o probabilístico, ou seja, a escolha é feita ao acaso e toda a unidade de amostra pode ter a mesma hipótese de escolha. Essa seleção é aleatória (ou randômica) e é feita por sorteio. Outro método, não probabilístico, procura fazer uma escolha estratificada e garante que todos os tipos de participantes sejam escolhidos.

A análise de dados para obtenção dos resultados e conclusões da pesquisa é combinado com o método estatístico para interpretação dos resultados. Esse tipo de método possibilita transformar escalas subjetivas em números e consequentemente, conduz a uma reflexão baseada em diferentes aspectos quantitativos e qualitativos, para entender algumas conclusões.

3. Amostra

A amostragem é uma etapa muito importante no processo de pesquisa e está associada a decisão de quem entrevistar e que grupos essas pessoas devem ser originárias (FLICK, 2009). Uwen Flick (2009) faz uma breve organização das decisões referentes a amostragem numa pesquisa:

<i>Tabela:</i> <i>Decisões relativas à amostragem no processo de pesquisa</i>	
Na coleta de dados	Amostragem de casos Grupos de amostragem de casos
Na interpretação dos dados	Amostragem do material Amostragem dentro do material
Na apresentação das descobertas	Amostragem da apresentação

Tabela retirada de Flick (2009)

Existem dois tipos de amostragem, segundo Flick (2009):

1. A determinação ***a priori*** da estrutura da amostra: os critérios são abstratos e partem da ideia da escolha da amostra através dos tipos e distribuição do objeto a ser

pesquisado. Conhecida mais comumente na lógica estatística, o material é reunido a partir de alguns critérios.

2. A **amostragem teórica**: são estratégias graduais que se baseiam em decisões relativas à seleção e reunião de material empírico e que são tomadas no processo de análise e interpretação dos dados.

4. Questionários

1. Deve ser elaborada uma versão inicial do questionário, incluindo um nível de itens superiores aqueles que realmente queremos.

2. Deve-se submeter esta versão inicial também denominada “estudo-piloto” à apreciação de todos.

3. Deve ser realizado um pré-teste com uma amostra relativamente grande de modo a perceber quais os itens que irão integrar a versão final.

4. A versão final deve ter a amostra adequada ao objetivo do estudo.

Num questionário, cada item mede uma variável diferente, sendo que qualquer fator estranho, poderá interferir com as respostas, colocando assim em causa a avaliação dessa variável.

Num questionário, o investigador deve definir com clareza qual a tarefa que o entrevistado terá de realizar.

O questionário demonstrará ter problemas quando o entrevistado que lê o item não o compreenda, o entrevistado poderá interpretar de maneira diferente o que o investigador espera.

Sugestões para escrever bons itens

Ainda que o questionário piloto possa ser um instrumento importante para melhorar o questionário final, há três regras que Sudman e Bradburn (GÜNTHER, 2003) apontam como relevantes para evitar alguns problemas:

a) controle o impulso de escrever itens específicos antes de haver refletido completamente sobre as perguntas da pesquisa;

b) anote as perguntas da pesquisa e mantenha-as perto enquanto estiver desenvolvendo o questionário;

c) cada vez que escrever um item, indague "Por que quero saber disto?"

Os desafios e sugestões apresentados no âmbito da construção dos questionários também são apresentados como (GÜNTHER, 2003):

- Linguagem e ambiguidade: evitar gírias ou palavras complexas;
- Tendenciosidade e ênfase: evite direcionar as respostas com palavras, exemplo "proibir" e "não permitir" possuem influências diferentes nas respostas.
- Perguntas abertas vêm no início para estabelecer clima receptivo e no fim para capturar opiniões que não foram cobertas pelos itens

Considerações técnicas

As opções de respostas são diversas e são classificadas pelos níveis de informação e parâmetros que apresentam:

- *Escala nominal*: utiliza-se número ou símbolo para identificar pessoas, objetos ou categorias;
- *Escala ordinal*: identifica pessoas, objetos ou categorias numa dimensão subjacente (hierarquização de preferência ou importância, status social ou ordem de chegada);

- *Escala intervalar*: são ordenadas conforme uma dimensão subjacente e os intervalos entre as alternativas têm tamanho conhecido e podem ser comparados;
- *Escala de razão*: são utilizados como escalas de salário ou tempo gasto com alguma tarefa;
- *Escala Likert*: mais utilizada, mede atitudes, opiniões e avaliações. Geralmente são 5 alternativas: aplica-se totalmente, aplica-se, nem sim nem não, não se aplica, definitivamente não se aplica (também pode ser inserido não sei) (GÜNTHER, 2003).

5. Análise Estatística

A análise estatística é um dos fundamentos mais importante nas pesquisas do tipo survey. Ela representa uma abordagem quantitativa que busca, através dos dados, também qualitativos, apresentar opiniões, atitudes e categorias de um conjunto numérico relevante de pessoas pesquisadas.

Os estudiosos desse tipo de análise procuram tratar com cautela os resultados obtidos e usam em geral a moda e mediana em lugar da média dos dados. No âmbito da estatística, também, se opta por estatísticas paramétricas para fins inferenciais (Günther, 2003). Nesse breve texto iremos abordar de modo superficial, mas espera-se, suficientemente abrangente, para mostrar um pouco das dinâmicas e debates decorrentes dessa perspectiva de análise estatística.

A análise estatística ganhou relevância nas ciências sociais quando se tornou necessário compreender diferentes relações entre modos de vida e propriedades na Sociologia. Pierre Bourdieu ganhou relevância nesses estudos quando publicou o livro *A Distinção* em 1976 onde tratou de temáticas estatísticas específicas para analisar seus

dados (ROUANET, ACKERMAN, & ROUX, 2000). Trabalhando com estatísticos, o grupo de sociólogos e matemáticos, construíram uma nova forma de analisar estatística conhecida como *abordagem geométrica*. Sua análise de dados é tão marcante no âmbito das Ciências Humanas, que hoje o autor é considerado um divisor de águas nos estudos sobre a sociedade, dando "dimensões de cientificidade" à sociologia (ROUANET, ACKERMAN, & ROUX, 2000).

Nesse momento, trataremos as questões básicas de estatísticas de modo a garantir alguns aspectos gerais dessa ciência.

Na estatística temos a média aritmética que representa a soma de todos os números e a divisão pelo número de vezes apresentados. Um exemplo: se temos uma sequência de números:

1 5 8 3 7

Para a média aritmética, somamos todos e dividimos por 5 vezes, ou seja:

$$(1+5+8+3+7)/5 = 4,8$$

A mediana é o ponto que divide 2 partes de mesma área, por exemplo, abaixo vemos um conjunto de dados expostos em gráfico, a área toda do gráfico (que não é simétrico) é dividido ao meio, possibilitando que se observe no gráfico a metade da área representada. Esse ponto é o que chamamos de mediana, a moda, por sua vez, representa o valor mais frequente de um conjunto de dados (VANIN, GOUFFON, & HELENE, 2005).

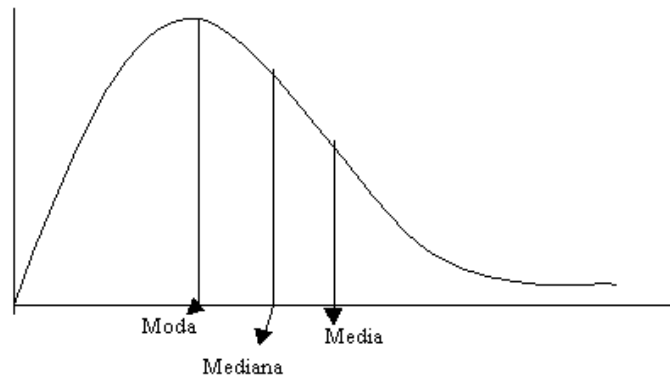


Figura: <http://web.cortland.edu/flteach/stats/glos-sp.html>

Assim, ao observarmos o gráfico que representa os pontos estatísticos apresentados, entende-se o motivo pelo qual Günther (2003) aponta a necessidade de utilizar nos dados estatísticos a moda ao invés de medianas ou médias dos dados. A moda representa a relevância do que está a ser tratado no tema enquanto as outras apontam uma representação fictícia dos dados no âmbito da análise de dados.

No que se refere a dados paramétricos e não paramétrico, a abordagem através da perspectiva da estimativa de valor. O modelo normal paramétrico está associado a estimativa do valor verdadeiro de um dado, ou seja, ela objetiva estimar a média.

Para que tal situação ocorra há possibilidades estatísticas de análise de dados, que não abordaremos aqui, mas que produzem uma distribuição normal como apresentado no gráfico abaixo. Inicialmente vemos diferentes colunas distribuídas de forma não homogênea, mas um gráfico que apresenta uma distribuição normal.

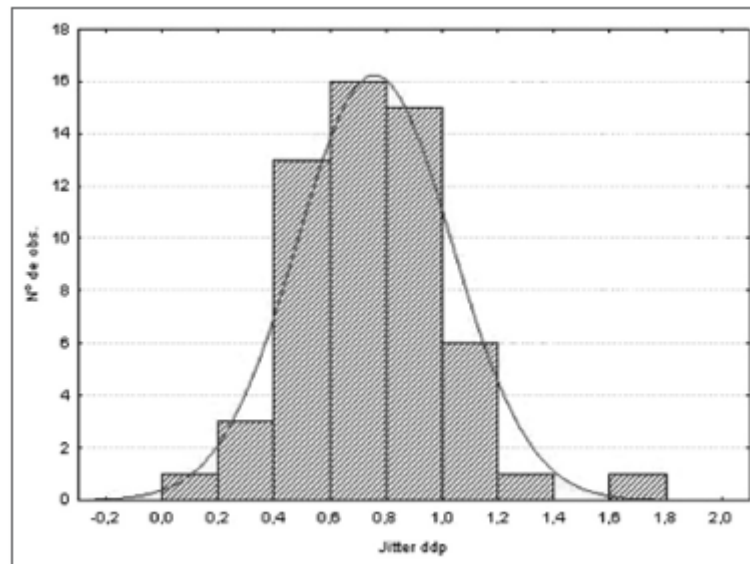


Figura 1: http://oldfiles.bjorl.org/conteudo/acervo/print_acervo.asp?id=3850

Se fosse retirado o histograma, teríamos algo como o da figura abaixo.

Obsevando esse gráfico podemos perceber que existe uma simetria onde há, no meio, uma média que representa também, uma mediana. Essa é uma característica significativa de estatística paramétrica. A curva é conhecida como gaussiana e possui a simetria como o principal instrumento de análise. Utilizar esse tipo de análise estatística refere-se a compreender o maior número possível de respostas ao que foi estudado e compreender a distribuição em função dessa tal preferência nos dados obtidos.

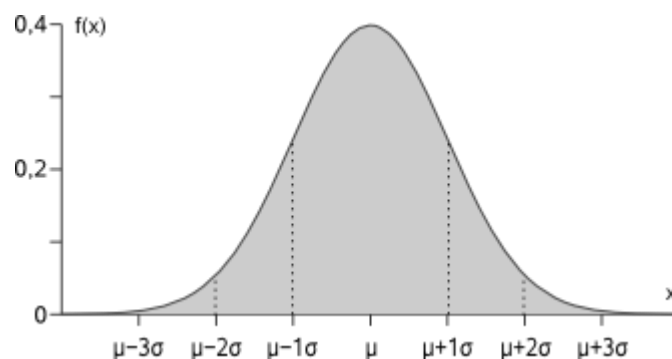


Figura 2: http://www.mspc.eng.br/matm/prob_est240.shtml

Uma análise estatística não paramétrica refere-se a uma curva, desenhada a partir de uma equação que impõem um tipo de desenho baseado nos dados apresentados, em geral, por histogramas, mas que obedece de modo sistemático a curva representativa dos dados.

No que se refere a inferência estatística, ela possibilita fazer informações sobre características de um determinado grupo populacional baseado na análise de uma amostra. Suas características são: amostragem, estimação por ponto e intervalo de confiança (desvio padrão, por exemplo) e teste de hipótese.

O teste de hipótese é definido formalmente como:

(...) uma regra que, aplicada aos dados experimentais, leva ou à decisão de aceitar a hipótese em consideração ou à decisão de a rejeitar. Exatamente pela decisão ser extrema - ou bem se aceita a hipótese ou é rejeitada, há possibilidade da flutuação estatística conduzir à conclusão trocada (VANIN, GOUFFON, & HELENE, 2005, p. 118)

O mais utilizado teste de hipótese é o teste $|t|$ que representa uma regra bastante simples, o teste estatístico consiste em aceitar a proposta como verdadeira se $|t|$ for pequeno e rejeitá-la se $|t|$ for grande. O quão grande ou pequeno $|t|$ deve ser para o aceitar ou rejeitar está habitualmente relacionado com o valor crítico, que em termos gráficos representa a região onde os dados estão muito distantes da média de uma curva gaussiana.

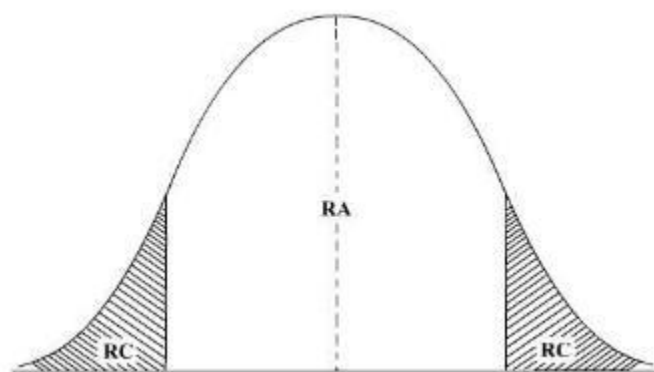


Figura 3: <http://www.ebah.com.br/content/ABAAABnXcAL/teste-hipoteses>

Em suma, tentou-se nesta apresentação referir alguns instrumentos para compreender de modo geral o sentido das análises estatísticas que representam as pesquisas do tipo survey.

Há diferentes modos de analisar com softwares que produzem as curvas de análise sem a necessidade de cálculos pelos pesquisadores. No entanto, é sempre relevante lembrar que o conhecimento aprofundado da estatística é primordial para reconhecer e analisar os gráficos apresentados pelos programas como SPSS, NVivo, Stata, entre outros.

Referências

- Almeida, L.; & Freire, T. (2003). *Metodologia de Investigação em psicologia e educação*. Braga: Psiquilíbrio.
- Flick, U. (2009). *Introdução à pesquisa qualitativa*. Porto Alegre: Artmed.
- Freitas, H. et al. O método de pesquisa survey. *Revista de administração*, São Paulo, 35, n. 3, julho/setembro 2000. 105-112.
- Günther, H. (2003). *Como elaborar um questionário*. (Universidade de Brasília) Acesso em 9 de Março de 2015, disponível em Planejamento de pesquisa nas ciências sociais: www.psi-ambiental.net/pdf/01Questionário.pdf

Rouanet, H., Ackerman, W., & Roux, B. L. (2000). A análise geométrica de questionários: a lição de La Distinction de Bourdieu. *Bulletin de Méthodologie Sociologique*, 65, pp. 5-15.

Silva, R. K. (09 de Dezembro de 2013). *Método de Pesquisa Survey*. Acesso em 09 de Março de 2015, disponível em P@rtes: <http://www.partes.com.br/2013/12/09/metodo-de-pesquisa-survey/#.VP3SjuHTiSo>

Tuckman, B. (2000). Elaboração de questionários. In: B. Tuckman, *Manual de investigação em educação* (p. 744). Lisboa: Fundação Calouste Gulbenkian.

Tuckmann, B.W. (2000). *Manual de Investigação em educação*. Lisboa: FCG.

Vanin, V. R., Gouffon, P., & Helene, O. (2005). *Análise estatística de medidas em ciências exatas*. São Paulo: Universidade de São Paulo.