

ANÁLISE QUANTITATIVA

AUTORES

Jyotikumarie Juggernath: African Vision Research Institute (AVRI), Durban, South Africa

Kovin S Naidoo: Brien Holden Vision Institute (BHVI), African Vision Research Institute (AVRI), Durban, South Africa

Urmilla Bob: University of KwaZulu Natal (UKZN), Durban, South Africa

Vadivelu Moodley: African Vision Research Institute (AVRI), Durban, South Africa

REVISOR

James Loughman: Dublin Institute of Technology (DIT), Dublin, Ireland

ANÁLISE QUANTITATIVA / ANÁLISE ESTATÍSTICA

Os dados quantitativos referem-se a dados que incluem valores numéricos, variando entre a frequência numérica de ocorrências à apresentação complexa de dados em termos de gráficos e figuras. A análise estatística é levada a cabo para interpretar informação numérica de forma científica. Os achados são quantificados e reduzidos para valores numéricos. A análise estatística fornece procedimentos para recolher, sumarizar, analisar, apresentar e interpretar dados empíricos.

TÉCNICAS ESTATÍSTICAS PRINCIPAIS	<ul style="list-style-type: none">• Estatística descritiva: sumariza (reduz uma grande quantidade de informação bruta [observações] numa quantidade de informação sumarizada), estima os parâmetros da população, apresenta e interpreta conjuntos de dados.• Estatística inferencial: faz afirmações sobre uma população baseada nos dados da amostra.• Estes tipos de estatística são levados a cabo com limites muito estreitos e incluem testes paramétricos e não-paramétricos, teste da hipótese, distribuição de amostra e probabilidades.<ul style="list-style-type: none">○ Testes paramétricos: normalmente efetuada em dados contínuos.○ Não paramétricos: são frequentemente utilizados para dados categóricos mas podem ser também usados em dados contínuos.○ Hipótese nula: É a hipótese que se quer rejeitar.○ Hipótese alternativa: É a hipótese que se quer provar como verdadeira.• Significância: determina se uma tendência de observação é significativa ou resulta apenas do acaso.• Previsão: o resultado dos processos pode ser previsto dentro de certos limites.
LIMITAÇÕES DA ANÁLISE ESTATÍSTICA	<p>A análise estatística não faz qualquer julgamento ou fornece conclusões concretas. Ela fornece meramente as probabilidades matemáticas. Os investigadores é que fazem juízos sobre os dados, e a sua interpretação é normalmente influenciada por algumas suposições, as técnicas de amostragem (incluindo o tamanho da amostra) e os diferentes testes usados para analisar os dados. O tipo de técnica estatística utilizada depende do conjunto de dados.</p> <p>A maioria dos programas de análise estatística são apenas ferramentas para serem usadas dentro de uma conjectura teórica particular e deve ser interpretada criticamente. Um desenho estatístico bem formulado, incluindo dados que são precisos é um pré-requisito para uma boa análise estatística. Estatística (e métodos quantitativos em geral) é uma ferramenta complementar dentro de outras abordagens disponíveis. É importante que o investigador esteja alertado pela falácia ecológica, i.e. os resultados de uma análise não devem ser extrapolados para além da escala e do ambiente representado pelos dados originais.</p>

Exemplo: Os achados da investigação entre os alunos de optometria em Singapura indicaram que a prevalência de miopia é de 10%. Este resultado não pode ser extrapolado para a população como um todo ou para comunidades rurais com menos estudos. O achado não é representativo de toda a população como um todo mas pode relacionar alunos de medicina nomes no espaço universitário.

DADOS

Os dados são a forma de informação mais básica, na sua forma bruta e não refinada/ não processada.

Medições/ observações/ indivíduos/ casos (fenómenos numéricos ou verbais sob estudo) são designados por dados em bruto (ou não processados)

O que está a ser medido é designado uma variável (exemplo: acuidade visual, temperatura, idade, tamanho do agregado familiar, etc.) – uma propriedade ou característica de cada observação que pode ser medida, classificada ou contabilizada.

Cada observação ou medida é designada uma variável individual. Cada conjunto de dados consiste em várias variáveis individuais para cada variável que é medida. Para uma variável “x” as variáveis são designadas por x_1 , x_2 , x_3 ...

Onde x_1 é a primeira variável individual e x_n é a última variável individual registada. O número total de variáveis individuais ou observações é n.

Se o número total de observações for $n=50$, então o intervalo de observações está entre 1 e 50.

Exemplo: As medições de acuidade visual são efetuadas entre os pacientes numa clínica e o seguinte são algumas das observações.

6/6; 6/9; 6/12; 6/24; 6/36; 6/60..... n_{50}

MEDIÇÕES DE DADOS

Medições Discretas

Unidades completas ou inteiros (exemplo: tamanho de propriedade, género, número de árvores numa plantação, etc.)

Medições Contínuas

Pode ser um valor, incluindo frações (exemplo: idade, medição da acuidade visual, equivalente esférico do erro refrativo). O nível de detalhe depende qual a precisão necessária e quão precisos são os instrumentos de medida.

ESCALAS DE MEDIÇÃO DE DADOS

ESCALA NOMINAL	<ul style="list-style-type: none"> As categorias não são ordenadas As observações são ordenadas em categorias mutuamente exclusivas as quais não têm relações inerentes. <p>Exemplo: raça, grupos linguísticos, género, áreas residenciais, perguntas de sim e não</p>
ESCALA ORDINAL	<ul style="list-style-type: none"> As categorias são ordenadas. As observações são colocadas numa ordem específica e ordenadas de acordo com um determinado critério. As categorias de medições têm uma ordem hierárquica. Análise estatística não paramétrica. <p>Exemplo: grupos de remuneração (alto, médio, baixo), ordem do rio, cidades organizadas por habitabilidade</p>
ESCALA DE INTERVALO	<ul style="list-style-type: none"> Dados contínuos sem um zero fixo A cada observação é dado um valor numérico exato A diferença entre duas quaisquer observações é conhecida Pode ser subjetiva a testes paramétricos <p>Exemplo: temperatura (não existe nenhuma entidade física tal como sem temperatura, 0°C representa uma posição relativa na escala Celsius)</p>
ESCALA DE RÁCIO	<ul style="list-style-type: none"> Dados contínuos com um zero fixo A cada observação é dado um valor numérico exato e existe um valor de zero verdadeiro num ponto na escala. O rácio entre quaisquer dois valores permanece igual com respeito à unidade usada. Pode ser sujeito a testes paramétricos exemplo: precipitação ou remuneração (0 significa ausência de precipitação ou de remuneração sendo um zero verdadeiro e não uma posição relativa).

MEDIÇÃO DA TENDÊNCIA CENTRAL E MEDIÇÃO DA DISPERSÃO (VARIABILIDADE)

	<p>As medições da tendência central fornecem uma ideia da média de um conjunto de observações enquanto as medidas da dispersão indicam a dispersão ou grau de variabilidade de um conjunto de medições.</p>
MÉDIA (MÉDIA ARITMÉTICA)	<ul style="list-style-type: none"> Normalmente utilizada. Cálculo através da adição de todos os valores e divisão pelo número total de valores. Exemplo: $93 + 97 + 65 + 66 + 41 + 56$ dividido por $6 = 418/6 = 69.7$ <p>Vantagem da medição é que é matematicamente sonante pois é baseada no cálculo. A desvantagem é que, se existir um valor muito elevado nos dados o qual leva a que estes estejam enviesados, a média não irá ser um bom indicador da média ou do valor central. Tendo em conta o exemplo acima, supor que a observação 56 é alterada para 256, o resultado do cálculo seria agora de $618/6=103$. Neste tipo de casos podem ser utilizadas outras medidas da tendência central tais como o modo ou a mediana.</p>
MEDIANA	<ul style="list-style-type: none"> Ponto central de um conjunto de dados ordenados i.e. o ponto no qual metade do número total de observações é alcançado. Os valores são organizados do menor para o maior Exemplo: 10, 70, 20, 50, 20, 30, 40, 40, 10, 60, 70, 80, 90, 90, 90 <p>Primeiro é ordenado para: 10, 10, 20, 20, 30, 40, 40, 50, 60, 70, 70, 80, 90, 90, 90. A mediana é a observação 8 que é a observação central à qual corresponde o 50.</p> <ul style="list-style-type: none"> Se o número de valores for par, a mediana é a média dos dois valores centrais. No exemplo anterior se o número 100 for adicionado ao conjunto de números, a mediana passa a ser 55. A desvantagem da mediana é que esta é insensível a valores extremos (discrepantes). Adicionalmente, uma condição necessária para utilizar a mediana é o valor necessita ser ordinal uma vez que as observações têm que ser ordenadas com significado.
MODA	<p>É a forma menos complicada de descrever o centro da distribuição.</p> <ul style="list-style-type: none"> É o valor que ocorre com maior frequência no conjunto de dados. Se existirem dois modos os dados são ditos bimodais, se existirem três ou mais modos os dados são ditos multimodais. A moda não quer qualquer ordem de ordenação das categorias ou distâncias iguais entre categorias. Pode ser aplicado a qualquer variável mas é tipicamente aplicado a variáveis nominais. A desvantagem da moda é que pode vezes é um valor difícil de determinar e por vezes é ambíguo. Exemplo: na distribuição de categorias por idade, aparecem dois grupos 50 vezes 20 e 24 anos, mas outras categorias de idade aparecem quase com a mesma frequência tal como 28 e 30 anos com 49 ocorrências as quais não irão estar representadas na moda.

MEDIÇÃO DA VARIABILIDADE /DISPERSÃO

O desvio padrão é a medida mais utilizada para medir a variabilidade e está relacionada com a distância entre as observações e a média. O desvio padrão é a distância média em relação à média e refere-se a uma distribuição usada frequentemente i.e. a distribuição normal a qual é importante para a maioria dos testes estatísticos. O desvio padrão é expresso nas mesmas unidades da própria medida como variável. Por exemplo, se a variável peso corporal é medida em quilogramas, o desvio padrão indica a média da distância à média em quilogramas. O desvio padrão é apenas aplicável para variáveis em intervalo ou rácio e também é sensível a pontos extremos (distribuições enviesadas à direita e à esquerda). É importante notar que o desvio padrão relacionado com a distribuição da amostragem não é designado por desvio padrão mas como erro padrão.

ORGANIZAÇÃO DOS DADOS

- **Ordenação das Variáveis**
 - Os dados são ordenados em ordem crescente ou decrescente.
 - A menor ou maior variável é definida como x_1 , a segunda como x_2 etc. A última variável como x_n .
- **Tabelas de Frequência**
 - O número de ocorrências de cada valor ou classe é contabilizado.
- **Diagramas de dispersão**
 - Útil para representar uma variável contra outra de forma a identificar tendências dentro do conjunto de dados.
 - Cada ponto tem duas variáveis ligadas a ele.
 - Pode ser identificado o tipo de correlação.
 - Importante decidir o tipo de análise estatística que pode ser efetuado.

MEDIÇÃO E AMOSTRAGEM

A medição é necessária em investigação quantitativa e é o processo de atribuição de palavras, símbolos e números a pessoas, objetos ou locais. Envolve a análise da diferença entre o valor atual e o valor medido. Considerações importantes na medição são:

- **Fiabilidade:** introdução de viés que podem emanar de diferentes fontes e necessitam ser minimizadas.
- **Validade:** trata de assuntos relativos à natureza, significado ou definição de um conceito ou variável (necessário para definições operacionais).
- **Precisão:** refere-se ao nível de exatidão associado com uma medida (muitas vezes determinado pela calibração do equipamento de medida, exemplo: tonómetro).
- **Exatidão:** refere-se à extensão do viés no sistema no processo de medição de forma a não serem obtidas medidas distorcidas da realidade.

FONTES DE ERROS DE MEDIÇÃO

Erros de cálculo: erros efetuados na execução simples de operações aritméticas.

Erros de medida: erros de incorreção no conjunto de dados (exemplo: codificação incorreta, codificação em falta, etc.).

Erros de Especificação: utilização de pressuposto incorretos e /ou aplicação incorreta de formulas ou equações matemáticas.

Erros de amostragem: se a amostragem da população não é correta, o erro é estendido para a generalização da população.

Ruído aleatório no meio: refere-se a ruído nos dados ou falta de compreensão.