

ESTATÍSTICA: definição e conceitos básicos

A **Estatística** é um **conjunto de técnicas** úteis para a tomada de decisão sobre um processo ou população, baseada na análise da informação contida em uma amostra desta população. A Estatística é uma ciência que oferece uma coleção de métodos para **planejar** experimentos e levantamentos para obter dados, **organizar, resumir, analisar, interpretar dados** e deles extrair conhecimento.

Deste modo, a Estatística **contribui para que dados gerem conhecimento** e, como tal, deve ter como objetivo não só a produção de dados, como também a interpretação de dados já existentes, utilizando a combinação de gráficos, tabelas e medidas numéricas que permitam interpretar o que esses dados significam.

Pensamento estatístico: Raciocínio analítico que enfoca a onipresença da variação; a investigação das fontes de variação; o planejamento de coleta de dados com a variação em mente; a quantificação da variação; e a explicação da variação.



Os dados em geral são números, mas não são "apenas números". Os dados são números com um contexto.

- Procure entender o que os dados dizem em cada contexto específico. Todos os métodos estatísticos nada mais são do que instrumentos que nos ajudam a entender os dados.
- Deixe para uma calculadora ou um computador o máximo possível dos cálculos e gráficos e procurar concentrar-se no que fazer? E por que fazer?
- Enfoque as grandes ideias da estatística, e não apenas regras e receitas.

Por que precisamos coletar dados?

- Para oferecer o insumo necessário a uma pesquisa;
- Para avaliar o desempenho de um processo de produção ou de um serviço em andamento;
- Para assessorar na formulação de cursos de ação alternativos num processo de tomada de decisão; e
- Para satisfazer nossa curiosidade.

Conceitos básicos

Uma **população** é uma coleção completa de todos os elementos (valores, pessoas, medidas, etc) a serem estudados.

Um **censo** é uma coleção de dados relativos a todos os elementos de uma população.

Uma **amostra** é uma sub-coleção de elementos extraídos de uma população.

Um **parâmetro** é uma medida numérica que descreve uma característica de uma população.

Uma **estatística** é uma medida numérica que descreve uma característica de uma amostra.

Os dados **quantitativos** consistem em números que representam contagens ou medidas.

Os dados **qualitativos** podem ser separados em diferentes categorias que se distinguem por alguma característica não-numérica.

Os **dados** são a matéria prima da Estatística. Definido o assunto de interesse, os dados são obtidos da medição de determinada característica ou propriedade desse objeto, pessoa ou coisa.

Os **dados discretos** resultam de um conjunto finito de valores possíveis, ou de um conjunto enumerável desses valores.

Os **dados contínuos** resultam de um número infinito de valores possíveis que podem ser associados a pontos em uma escala contínua de tal maneira que não haja lacunas ou interrupções.

Variável - Qualquer conjunto de dados contém informações sobre algum grupo de indivíduos. As informações são organizadas em variáveis. Uma variável é uma característica, propriedade ou atributo de uma unidade da população, cujo valor pode variar entre as unidades da população.

Varição – O padrão de variação de uma variável constitui a sua distribuição. A distribuição de uma variável quantitativa registra seus valores numéricos e a frequência de ocorrência de cada valor.

Como devemos coletar dados?

Os dados são parte crucial no estudo da variabilidade, pois, assim como são geradores de resultados podem também ser geradores de incerteza, estimulando e motivando mais aprofundamento e estudo sobre seu comportamento e sobre sua distribuição. Aos dados também estão associadas às fontes de erros que interferem diretamente na aplicação dos métodos estatísticos.

O **Planejamento** é crucial numa coleta de dados que vise um estudo estatístico.

Os processos ou padrões definidos para coletar dados são chamados de planejamentos. Os planejamentos devem abordar principalmente: Como vamos selecionar os indivíduos a serem estudados (tipo de amostragem ou de delineamento experimental); Quantos indivíduos devemos estudar (tamanho da amostra); Se há necessidade de composição de grupos e como eles devem ser formados para que possam ser comparados (alinhamento da amostra com os objetivos e restrições); como serão feitas as medições (procedimentos e instrumentos de medição); etc. O planejamento sistemático para gerar dados é um dos primeiros passos para a realização de um estudo com base científica. A falta de planejamento pode levar a tendenciosidades, à falta de dados ou a resultados confusos e imprecisos.

Produção de dados: parte crucial para a inferência estatística, para responder questões específicas formuladas antes dos dados serem produzidos.

- **Estudos observacionais ou levantamentos:** visam retratar a população o menos distorcida possível pelo ato da coleta de informações. Não tentamos manipular, influenciar ou modificar as respostas dos elementos a serem estudados.
- **Experimentos:** decorre da aplicação de determinado tratamento para posterior observação de seus efeitos. Preocupação com a relação de causa e efeito sobre os elementos pesquisados. Impõe deliberadamente algum tratamento aos indivíduos, a fim de observar sua reação e medir respostas. Requer planejamento apropriado para serem válidos cientificamente.
- **Simulação:** Uso de modelo matemático ou físico para reproduzir as condições de uma situação ou processo, usando métodos computacionais. A simulação é uma alternativa quando é impraticável ou mesmo perigoso estudar os fenômenos de interesse em condições reais.

A **Inferência Estatística** objetiva *estudar a população* por meio de *evidências* decorrentes de uma avaliação minuciosa *da amostra*. É parte da inferência estatística, produzir estimativas e testar hipóteses.

Contexto geral da análise estatística

