



Data Science Academy

www.datascienceacademy.com.br

Microsoft Power BI Para Data Science

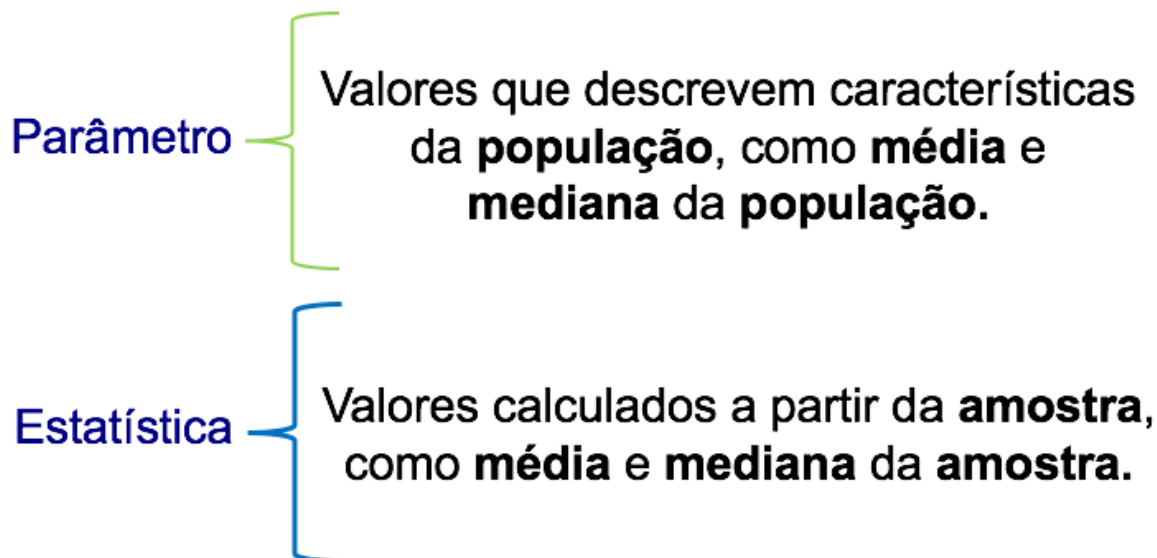
Erro de Amostragem

Não há dúvida de que uma amostra **não representa perfeitamente** uma população. Ou seja, a utilização de uma amostra implica na aceitação de uma margem de erro que denominaremos ERRO AMOSTRAL, que ocorrem durante o processo de amostragem.

O Erro Amostral é a diferença entre um resultado amostral e o verdadeiro resultado populacional. Tais erros são decorrentes da variabilidade natural das unidades amostrais. Ou seja esses erros são aleatórios.

Sem dúvida amostragem é uma técnica fabulosa, que nos permite obter informações sobre uma população inteira, analisando apenas uma porção dos dados e isso nos garante eficiência, além de tempo e custo. Ela é a base de vários algoritmos de Machine Learning.

A chave para compreender os erros de amostragem é compreender que Parâmetro e Estatística são coisas diferentes:



Como a Estatística é calculada a partir da amostra, que é uma parte da população, não seria razoável esperar que a média dos dados da amostra fosse igual a média dos dados da população? A diferença entre estes 2 valores é chamada “erro de amostragem da média da amostra”.

A diferença entre estes 2 valores é chamada “erro de amostragem da média da amostra”.

Erro de amostragem = $x - \mu$

Onde:

x = média da amostra

μ = média da população

Erros de amostragem podem diferir de uma amostra para outra e podem ser positivos ou negativos. Como regra geral, quanto maior o tamanho da amostra, menor será o erro de amostragem. Mas não há garantia de que isso sempre ocorrerá.

E agora talvez você entenda porque o Big Data é tão revolucionário. Com mais dados a nossa disposição e recursos tecnológicos, podemos coletar e utilizar cada vez mais dados e conduzir testes estatísticos cada vez mais precisos. Esse é o grande diferencial do Big Data Analytics. Os conceitos da Estatística existem há muito tempo, mas nunca houve um volume de dados tão grande, como o que vemos atualmente.

