

UNIVERSIDADE DE SÃO PAULO - INSTITUTO DE GEOCIÊNCIAS
GSA-0200 Introdução à Análise de Dados em Geologia Sedimentar
2021

19/04- Introdução à análise de dados em Geologia Sedimentar. Especificidades de diferentes tipos de dados geológicos. População, amostra e interpretações geológicas. A importância de conhecer e propagar incertezas. Erro aleatório e erro sistemático. Introdução às ferramentas numéricas e aos pacotes de análise de dados em Python (Scipy, Numpy, Pandas, Pyplot, Seaborn e Plotly).

26/04- Dados contínuos unidimensionais. Estatística descritiva e formas de representação (histogramas, curvas de densidade de probabilidades, boxplots e violin plots). Identificação e representação de outliers. Distribuição normal e log-normal.

03/05 -Outros modelos de distribuição utilizados em Geologia Sedimentar. Desvio padrão, erro amostral, erro da média e intervalo de confiança. Prática de determinação e propagação de erro.

10/05 - Formatação de dados para o pacote Pandas e prática de importação e visualização com Pyplot e Seaborn. Modelos estatísticos com funções em Python. Exemplos em Geologia Sedimentar - dados de granulometria, permeabilidade.

17/05 - Testes de hipóteses, conceitos e aplicações. Testes paramétricos e não paramétricos. Testes de hipóteses para mais de duas amostras. Montecarlo e “bootstrapping”. Prática com funções em Python em dados geológicos.

24/05 - Dados multidimensionais 1. Covariância. Crossplots, coeficiente de determinação e coeficiente de correlação. Regressão linear e Método dos Mínimos Quadrados. Propagação de erro em regressões lineares.

31/05 - Regressão não-linear: usos e equívocos. Prática - Ferramentas de regressão linear e não linear para aproximação de funções em Python.

07/06 - Dados multidimensionais 2. Análise de principais componentes e redução da dimensionalidade. Métodos de classificação em pacotes de ‘machine learning’ em Python.

14/06 - Dados composicionais. O problema do fechamento. Diagramas ternários e alternativas. Erro amostral e propagação em razões de elementos e de classes composicionais. Características dos diagramas discriminantes (geoquímica, contagens modais). Substituição de zeros, razão logarítmica e propagação de erros. Prática de substituição de zeros e razão logarítmica.

21/06 - Dados direcionais. Particularidades - média vetorial contornos e variância direcional (estatística de dados circulares). Distribuição de dados de rumo - a função von Mises. Efeito da rotação de camadas na variância direcional. Propagação de erro amostral e analítico (erro da bússola). Representações - estereogramas e histogramas circulares. Distribuições bimodais - falhas e fraturas. Paleocorrentes e variabilidade direcional dependente da escala. Prática - análise e representação de dados de paleocorrentes. Pacotes de visualização de histogramas e de diagramas em roseta.

28/06 - Dados composicionais com erro analítico independente - o exemplo de dados de datação. Funções de densidade de probabilidade multimodais. Comparação de amostras

considerando o erro. Prática análise de dados geocronológicos de componentes detríticos.
05/07 - Análise de dados em seções estratigráficas - tempo, modelo de idades, taxa de sedimentação e propagação de erro. Curvas de valores em seções estratigráficas: efeito da subamostragem no comprimento de onda.

12/07 - Curvas de valores em seções estratigráficas: efeito da subamostragem na determinação do erro amostral em sucessões estratigráficas. Eventos e hiatos: a distribuição Pareto o efeito Sadler em taxas de sedimentação.

Avaliação:

Média do conceito dos exercícios práticos.