

Introdução à Análise de Dados – Duarte Trigueiros

Objectivos: Esta introdução fornece uma base a alunos de licenciatura, mestrado e doutoramento nas áreas da Gestão e que venham a precisar de utilizar metodologias estatísticas nos seus trabalhos. Serve também de introdução, junto com outros, a cursos de *Business Intelligence* e de *Data Mining* para gestores e outros profissionais.

Programa:

1. Tipos de dados: nominais, ordinais, escalares (intervalares e racionais). Noção de medição. Variáveis contínuas e discretas, proporções.
2. O acaso e a probabilidade, leis da soma e da multiplicação. Aleatoriedade, variáveis aleatórias. Processo de Bernulli e a distribuição Binomial. Noção intuitiva de grau de liberdade. Processo e distribuição de Poisson.
3. A distribuição Normal como limite. Parametrização da distribuição Normal. Média, desvio padrão, enviezamento, curtose. Distribuição Normal acumulada, probabilidades associadas a intervalos e caudas.
4. População e amostra, amostragem aleatória e atribuição aleatória. Limite Central, a distribuição das médias, Intervalos de confiança para a média. Estimção vs medição.
5. Diferença entre duas médias, a hipótese nula, o teste t . Dois tipos de erro. Custos associados aos erros.
6. Diferença entre duas proporções, a aproximação do Qui-Quadrado. Tabelas cruzadas.
7. Covariância e correlação. Causação. Matriz de variância-covariância. Gráficos de dispersão. Distribuições correladas.
8. Regressão linear simples, método dos mínimos quadrados, pressupostos. Modelação linear, Variabilidade explicada e não explicada. Teste F de Fisher. Interpretação de parâmetros de uma regressão e testes de hipóteses. Análise de resíduos.
9. Transformação logística para probabilidades. Regra de Bayes. Introdução breve a outras metodologias de modelação e ao *Data Mining*.
10. [Opcional] Recolha, preparação, validação, armazenamento de dados. Gestão de dados e evolução das bases de dados: sequenciais, indexadas, relacionais, OLAP. Organização de dados: *Data Warehouses*. Produtos existentes.

Duração e equipamento complementar: O curso dura 18 a 24 horas lectivas dependendo do grau de profundidade pretendido e da preparação prévia dos alunos. Destas, dois terços são em sala com computadores.

Bibliografia: fornecida pelo docente.