

AGA0505 - Análise de Dados em Astronomia I - Lista #6

Data de entrega: 16/05/2017 - 14:30

- Seja claro nas respostas, explicando seu raciocínio; não serão aceitas respostas sem justificativa;

★

1. Faça a análise PCA da tabela de dados QSO.dat (no STOA) seguindo os seguintes passos:
 - (a) “Normalize” cada coluna de dados (i.e. subtraia sua média e divida pelo seu desvio padrão). Veja se algum dado é um *outlier*.
 - (b) Obtenha e apresente a matriz de covariância (13×13) desses dados: $C_{i,j} = E[(X_i - \mu_i)(X_j - \mu_j)]$
 - (c) Obtenha os 13 autovetores e autovalores. Quantas dessas componentes principais (autovetores) são necessárias para descrever 95% da variância total desses dados?
 - (d) Correlacione a primeira componente principal com os dados observacionais originais e veja quais grandezas são as que mais contribuem para essa componente.
2. Suponha que a plausibilidade de um determinado efeito seja de 5%. Utilizando-se do valor $p = 0.05$ como critério de rejeição da hipótese nula, e tendo o experimento atingido este nível, qual é a chance de que um efeito real tenha sido observado? E se o valor for $p = 0.01$? E se a plausibilidade for 50% (para cada um dos valores p)? (*Dica importante: leia "Statistical Errors"; Nuzzo, R. Nature (2014), e "Calibration of p Values for Testing Precise Null Hypotheses"; Sellke et al. The Am. Stat. (2001)*)
3. Acesse <https://xkcd.com/882/>
 - (a) Qual é o erro conceitual de estatística cometido pelos pesquisadores apresentados no quadrinho?
 - (b) Qual é a probabilidade que algum efeito espúrio ocorra na quantidade de testes apresentados no quadrinho, com o valor $p = 0.05$ adotado?
 - (c) Qual é o maior valor p que deveria ser adotado para que, em tantos testes, a chance de um evento espúrio se parecer com uma rejeição da hipótese nula fosse menor do que 5%.
4. A imagem abaixo foi feita pela sonda Viking 1, em Marte e atraiu a curiosidade das pessoas pela semelhança com um rosto, desenhado pelo padrão das montanhas e os contrastes. Gerry Soffen, cientista chefe da missão, descreveu a aparição como “um truque de luz e sombra”, o que seria melhor percebido em imagens de sondas futuras como Mars Reconnaissance Orbiter e Mars Global Surveyor. Partindo do ponto de vista instrumental do ser humano como um reconhecedor de padrões, explique fenômenos desse tipo - de reconhecer faces em imagens de objetos - como um tipo de erro de estatística. Este tipo de erro estatístico é associado a que característica de um instrumento de medição? (*Dica: considere uma fotografia do céu noturno, quando a sensibilidade é muito grande, mesmo as regiões escuras da foto apresentam*

artefatos de imagem, que não correspondem diretamente a objetos emitindo.) Por que - ainda adotando o mesmo ponto de vista - fenômenos do mesmo tipo, a identificação "acidental" de rostos em imagens, são tão comuns? (Questões de natureza psicológica, sociológica, etc. de pessoas não precisam ser consideradas nessa resposta. O exercício é a discussão do ponto de vista estatístico e de medição.)

