Exercícios de programação

Estes exercícios serão propostos durante as aulas sobre o Mathematica. Caso você use outra linguagem para os exercícios e problemas do curso de estatística, resolva estes problemas, a fim de verificar sua habilidade em lidar com as funções estatísticas do programa que adotou.

1) Um exercício para resolver graficamente

Os polinômios : p1[x] = 4 x^3 - 8 x^2 - 15 x e p2[x] = 4 x^6 - 16 x^5 + 4 x^4 + 44 x^3 - 24 x^2 - 47 x são praticamente coincidentes em duas regiões contidas no intervalo - 2 < x < 3. Encontre essas regiões a partir dos gráficos dos dois polinômios.

2) Simulação do lançamento de um projétil – parte 1.

Gerar um conjunto de dados que, uma vez graficado, represente a posição vertical medida (com erros, portanto) de um projétil em movimento livre.

i. Monte uma tabela com a posição vertical afetada da incerteza de medida
ii. Grafique o resultado do item i
iii. Escreva uma rotina que faça tanto o sorteio quanto o gráfico
iv. Verifique o que acontece cada vez que você recalcular a linha de comando do item iii

3) Distribuição da média de uma medida

Verificar, por simulação, como se distribui a média de uma medida e examinar a variância da média depende do número de dados e do número de simulações.

i. Sorteie 10 números aleatórios gaussianos de media 10 e desvio padrão 3
ii. Calcule a média desses valores
iii. Escreva uma rotina que sorteie os valores e faça a média de 10 aleatorios gaussianos
iv. Faça uma lista com 100 medias de 10 aleatórios gaussianos
v. Histograme essa lista
vi. Refaça a lista com as medias de 1000 medidas de 10 aleatórios gaussianos e histograme
vii. Faça uma lista com as medidas de 1000 medidas de 90 aleatórios gaussianos e histograme
viii. Compare os histogramas dos itens v, vi e vii e discuta o resultado

4) A função de probabilidade binomial

Comparar a simulação da quantidade de faces 2 em um lançamento de 5 dados de seis faces com as respectivas probabilidades calculadas.

i. Gere um número aleatório binomial com n=5 e p=1/6.
ii. Prepare uma lista com 100 desses números aleatórios.
iii. Faça um histograma dessa tabela
iv. Realize todas essas operações todas em uma única rotina.
v. Faça uma tabela com as probabilidades de obter 0, 1, 2 ... 5 faces 2 no lançamento de 5 dados.
vi. Faça um gráfico com o número esperado de faces 2 em 100 lançamentos de 5 dados.
vii. Superponha os gráficos dos itens iii e vi

5) A função densidade de probabilidade uniforme

Investigar a distribuição do erro de arredondamento na soma de um número grande de parcelas

i. Sorteie um conjunto de números aleatórios reais *u* com 12 elementos; sorteie valores maiores que 1.
ii. Crie outro conjunto com os valores de *u* arredondados de forma a apresentarem uma única casa decimal.
iii. Determine a diferença entre a soma de todos os elementos de *u* e a soma de todos os elementos do conjunto arredondado
iv. Reúna todos os itens anteriores em uma única rotina "erroDeArredondamento”, que tem como parâmetro o número de elementos de *u*, a fim de investigar como o erro de arredondamento depende desse número
v. Histograme um número grande dos erros de arredondamento e verifique se o resultado combina com a previsão teórica. Explore a dependência do erro com o número de parcelas da soma.

6) Propagar incerteza por simulação

Desenvolver um método de propagar incerteza de uma variável aleatória para uma função dessa variável, baseado em simulação. Investigar o comportamento do algoritmo com funções suaves e com a função exponencial.

i. Defina uma função da variável x, por exemplo f(x) = sen(5x)
ii. Simule uma medida de x com 10 dados, {x}, com média e desvio-padrão x0 e s0 a sua escolha.
iii. Defina duas variáveis xm e sm que correspondam à média e ao desvio-padrão da média dessa medida.
iv. Construa um grande conjunto de valores aleatórios, {z}, de média xm e desvio-padrão sm.
v. Determine o conjunto dos valores calculados de $f(z_i)$.
vi. Determine a média e o desvio-padrão dos valores do conjunto do item anterior.

vii. Compare o resultado obtido por simulação com aquele obtido com a fórmula analítica aproximada de propagação de incerteza.
viii. Prepare uma função que reúna os itens ii a vi e tenha como parâmetros o nome da função f e os valores x0 e s0.
ix. Teste a função com as escolhas dos itens i e ii. Explique porque os valores obtidos variam a cada repetição do cálculo e se essa variação está de acordo com o que lidamos na prática.
x. Aplique sua função nos casos:
   a) $s(x) = \pi x^2$, com x0 = 10.0 e s0 = 1.0
   b) g(x) = exp(- 5 t) com t0 = 2,110 e s0 = 0,012
   c) h(x) = exp(-5 t) com t0 = 2,11 e s0 = 0,12

   Nesses casos, compare os resultados de simulação com aqueles da fórmula analítica.

7) Ler um arquivo de dados e fazer gráficos com os valores obtidos

i. Leia o arquivo muge.dat, que está copiado na mesma pasta em que está este documento.
ii. Selecione os valores numéricos
iii. Separe os valores das 4 primeiras colunas. A primeira corresponde à variável x e as outras 3, a y1, y2 e y3.
iv. Selecione as unidades das grandezas das 4 primeiras colunas, que estão entre parênteses na 6ª linha do arquivo. Se precisar, releia o arquivo como texto.
v. Faça, no mesmo sistema de eixos, os três gráficos {x,y1}, {x,y2} e {x,y3}. Vale a pena usar escalas logarítmicas em ambos os eixos.
vi. Decore o gráfico, com os nomes dos eixos (x ou y e a unidade) e dos gráficos (y1,y2,y3)
vii. Selecione do conjunto de dados {x,y2} apenas os pontos em que y2>0.1 Faça um gráfico em escala linear desse novo conjunto, para verificar sua seleção.

8) Ajuste de parâmetros de um modelo

Considere a função $y = a \,\mathrm{Log}[\lambda t + b] + c/(t+0{,}5)$ e o conjunto de dados do arquivo "dadosExp1.dat", na forma $\{ (t, y, \sigma) \}$ em que σ é o desvio-padrão de y e t está em segundos, com y adimensional.

a) Modelo linear.

Sabendo que $\lambda = 0{,}2$/s e b = 0,5, ajuste os parâmetros a e c, determine seus desvios-padrões, calcule χ^2 e faça os gráficos necessários para avaliar a qualidade do ajuste. Analise criticamente os valores dos parâmetros obtidos.

b) Modelo não-linear

Considere todos os parâmetros (a, λ, b e c) desconhecidos. Nessa situação, ajuste todos esses parâmetros, determine seus desvios-padrões, calcule χ^2 e faça os gráficos necessários para avaliar a qualidade do ajuste. Analise criticamente os valores dos parâmetros obtidos.

9) Observação da distribuição de qui-quadrado na determinação da variância de uma medida

Considere uma medida com N dados distribuídos normalmente com mesma média e variância. Vamos verificar por simulação qual a distribuição da estimativa da variância.

i. Escolha um valor de N típico, N=6.
ii. Sorteie N números aleatórios gaussianos de média 20 e desvio-padrão 3.
iii. Calcule a variância desses N números.
iv. Escreva as operações ii e iii numa única instrução
v. Construa uma tabela com M repetições (M=1000, digamos) das estimativas da variância.
vi. Histograme a tabela anterior.
vii. Grafique a fdp correspondente à estimativa da variância para os valores de M e N escolhidos.
viii. Superponha o gráfico ao histograma.
ix. Repita as operações dos itens v a viii acima para outro valor de N
x. Escreva uma função que realize todas as operações v a viii em função dos parâmetros N e M, e retorne o gráfico superposto ao histograma.

10) Construção de uma tabela de t de Student

Em uma medição com N dados, todos distribuídos normalmente com média x_0 e desvio-padrão σ_0, os níveis de confiança associados aos intervalos

$$\bar{x} - \sigma_{m,0} \leq x_0 \leq \bar{x} + \sigma_{m,0}$$

$$\bar{x} - 2\,\sigma_{m,0} \leq x_0 \leq \bar{x} + 2\,\sigma_{m,0}$$

$$\bar{x} - 3\,\sigma_{m,0} \leq x_0 \leq \bar{x} + 3\,\sigma_{m,0}$$

são 68,3%; 95,4% e 99,7%, respectivamente.

Se o desvio-padrão verdadeiro da média é desconhecido e usamos sua estimativa convencional no lugar de $\sigma_{m,0}$ nos intervalos acima, a fim de manter esses níveis de confiança usuais precisamos usar intervalos

$$\bar{x} - t_I\,\sigma_{m,0} \leq x_0 \leq \bar{x} + t_I\,\sigma_{m,0}$$

$$\bar{x} - t_{II}\,\sigma_{m,0} \leq x_0 \leq \bar{x} + t_{II}\,\sigma_{m,0}$$

$$\bar{x} - t_{III}\,\sigma_{m,0} \leq\ x_0 \leq \bar{x} + t_{III}\,\sigma_{m,0}$$

com valores adequados de t_I, t_{II} e t_{III}, que devem ser obtidos a partir da distribuição de *t* de Student.

Construa uma tabela de t_I, t_{II} e t_{III}, para um número de dados $2 < N < 10$ e N = 20, 30, 40, ...100 (de 10 em 10). Você pode seguir qualquer caminho que deseje para montar essa tabela, mas se quiser uma orientação passo a passo, siga os itens abaixo.

i. Determine os níveis de probabilidade convencionais para m = 1, 2 e 3 desvios-padrões, usando a CDF[NormalDistribution[ ], m]
ii. Escreva uma função que retorna o valor de t' tal que $P(-t' \leq t \leq t') = p$ dado a probabilidade p e o número de dados N. Use FindRoot para resolver a equação.
iii. Monte uma tabela de valores críticos para N entre 2 e 10, outra para N entre 20 e 100 de 10 em 10, e junte (Join[ ]) as tabelas.
iv. Use TableForm para exibir a tabela, usando como legendas das linhas os valores de N, obtidos com a função Range[ ].

11) Correlação entre valores interpolados por uma função calculada com parâmetros ajustados pelo Método dos Mínimos Quadrados

Os parâmetros a e b da função $\omega(\theta; a, b) = a + b\cos^2\theta$ foram ajustados a um certo conjunto de dados; os valores estimados e a matriz de covariâncias estão no arquivo "DataExercise11.txt". Deseja-se obter os valores interpolados para θ_1 = 125º e θ_2 = 35º, bem como sua matriz de covariâncias. Você pode proceder como bem entender, mas abaixo está detalhado um procedimento possível.

i. Leia as estimativas e a matriz de variâncias do arquivo "DataExercise11.txt"
ii. Determine os valores y_1 e y_2 interpolados para θ_1 = 125º e θ_2 = 35º; aproveite para treinar a escrever funções e crie uma função capaz de calcular o valor interpolado para qualquer ângulo.
iii. Construa os vetores necessários para calcular as variâncias de y_1 e y_2.
iv. Determine as variâncias dos valores interpolados para θ_1 = 125º e θ_2 = 35º.
v. Determine a matriz de covariâncias de y_1 e y_2: use $V_y = F\,V_{ab}\,F^t$, com a matriz F apropriada.
vi. Determine o coeficiente de correlação entre y_1 e y_2.