Tratamento Estatístico de Dados em Física Experimental - Atividade 06

Teorema Central do Limite e intervalos de confiança

*Envie as respostas na forma de arquivos CSV ou arquivos de texto com colunas separadas por tabulação para o formulário https://forms.gle/aX1Kk8WeNek9Wgdx5. Essa atividade deve ser entregue até às **23h59** do dia 29/09 (quarta-feira). Até o final do prazo é possível mudar as respostas.*

Recomenda-se iniciar o gerador de números aleatórios com a semente igual ao seu número USP. No Python, isso é feito pelo comando np.random.seed(nU) com nU o seu número USP.

Questão 1) Considere a função densidade de probabilidade definida por:

$$f(x) = \begin{cases} \frac{3}{2}x^2 & se\ |x| \leq 1 \\ 0 & caso\ contrário \end{cases}$$

cujo valor verdadeiro é $x_0 = 0$ e o desvio-padrão (verdadeiro) é $\sigma_x = \sqrt{3/5}$ (verifique)

Faça uma rotina para gerar valores de x de acordo com essa função de densidade de probabilidade.

a) Gere um conjunto com $N = 10.000$ dados e preencha a coluna (1) do Quadro 1 com o número de valores observados em cada um dos intervalos indicados e com as estimativas amostrais dos indicadores de forma assimetria, $\boldsymbol{A}$ (scipy.stats.skew), e de curtose excedente, $\boldsymbol{K}$ (scipy.stats.kurtosis). Escreva os valores amostrais dos indicadores de forma usando 2 casas decimais.

Quadro 1 – Respostas numéricas da Questão 1.

$\boldsymbol{\lambda}$	(1) $n(\lvert x - x_0\rvert \leq \lambda\sigma_x)$	(2) $n(\lvert y - y_0\rvert \leq \lambda\sigma_y)$	(3) M=3 $n(\lvert S - S_0\rvert \leq \lambda\sigma_S)$	(4) M=5 $n(\lvert S - S_0\rvert \leq \lambda\sigma_S)$	(5) M=10 $n(\lvert S - S_0\rvert \leq \lambda\sigma_S)$	(6) M=100 $n(\lvert S - S_0\rvert \leq \lambda\sigma_S)$
1						
1,5						
2						
2,5						
3						
$\boldsymbol{A}$						
$\boldsymbol{K}$						

b) Considere agora que cada dado, y, seja a soma de dois valores que seguem a função $f(x)$ (isto é, que $y = x_1 + x_2$). Gere $N = 10.000$ valores independentes de y (o que implica em gerar, 20.000 valores de x) e preencha a coluna (2) do Quadro 1 com o número de valores de y observados em cada um dos intervalos indicados. Note que $y_0 = 2x_0$ e que $\sigma_y = \sqrt{2}\sigma_x$. Calcule também as estimativas amostrais dos indicadores de forma, $\boldsymbol{A}$ e $\boldsymbol{K}$, e os escreva com 2 casas decimais.

c) Faça uma rotina para calcular dados de uma variável S definida como $S = \sum_{i=1}^{M} x_i$, para os casos com M=3, 5, 10 e 100. Em seguida, preencha as colunas (3)-(6) do Quadro 1 com os resultados de conjuntos com $N = 10.000$ valores independentes de S (o item **a** corresponde a $M = 1$, quando $S = x = x_1$, e o item **b** ao caso $M = 2$, quando $S = y = x_1 + x_2$). Note que $S_0 = Mx_0$ e que $\sigma_S = \sqrt{M}\sigma_x$. Calcule as estimativas amostrais dos indicadores de forma, $\boldsymbol{A}$ e $\mathbf{K}$, e os escreva com 2 casas decimais.

Questão 2) Considere agora a função densidade de probabilidade definida por:

$$f(x) = \begin{cases} e^{-x} & se\ x \geq 0 \\ 0 & caso\ contrário \end{cases}$$

(Exemplo 3 do Tópico 2 com $L = 1$), cujo valor verdadeiro é $x_0 = 1$ e o desvio-padrão é $\sigma_x = 1$

Faça uma rotina para gerar valores de x de acordo com essa função de densidade de probabilidade. Note que é preciso usar o método da inversão neste caso, pois essa função tem domínio ilimitado.

a) Gere um conjunto com $N = 10.000$ dados e preencha a coluna (1) do Quadro 2 com o número de valores observados em cada um dos intervalos indicados e com as estimativas amostrais dos indicadores de forma assimetria, $\boldsymbol{A}$ (scipy.stats.skew), e de curtose excedente, $\boldsymbol{K}$ (scipy.stats.kurtosis) e os escreva com duas casas decimais.

b) Considere agora que cada dado, y, seja a soma de dois valores que seguem a função $f(x)$ (isto é, que $y = x_1 + x_2$). Gere $N = 10.000$ valores independentes de y (o que implica em gerar, 20.000 valores de x) e preencha a coluna (2) do Quadro 2 com o número de valores de y observados em cada um dos intervalos indicados. Note que $y_0 = 2x_0$ e que $\sigma_y = \sqrt{2}\sigma_x$. Calcule também as estimativas amostrais dos indicadores de forma, $\boldsymbol{A}$ e $\boldsymbol{K}$, e os escreva com duas casas decimais.

c) Faça uma rotina para calcular dados de uma variável S definida como $S = \sum_{i=1}^{M} x_i$, para os casos com M=3, 5, 10 e 100. Em seguida, preencha as colunas (3)-(6) do Quadro 2 com os resultados obtidos para conjuntos com $N = 10.000$ valores independentes de S. Lembre-se que $S_0 = Mx_0$ e que $\sigma_S = \sqrt{M}\sigma_x$. Calcule também as estimativas amostrais dos indicadores de forma, $\boldsymbol{A}$ e $\mathbf{K}$, e os escreva com duas casas decimais.

Quadro 2 – Respostas numéricas da Questão 2.

$\boldsymbol{\lambda}$	(1) $\boldsymbol{n(\lvert x-x_0\rvert \leq \lambda\sigma_x)}$	(2) $\boldsymbol{n(\lvert y-y_0\rvert \leq \lambda\sigma_y)}$	(3) M=3 $\boldsymbol{n(\lvert S-S_0\rvert \leq \lambda\sigma_S)}$	(4) M=5 $\boldsymbol{n(\lvert S-S_0\rvert \leq \lambda\sigma_S)}$	(5) M=10 $\boldsymbol{n(\lvert S-S_0\rvert \leq \lambda\sigma_S)}$	(6) M=100 $\boldsymbol{n(\lvert S-S_0\rvert \leq \lambda\sigma_S)}$
1						
1,5						
2						
2,5						
3						
$\boldsymbol{A}$						
$\boldsymbol{K}$						