

Tratamento Estatístico de Dados em Física Experimental - Atividade 12

Atividade – Estimativas amostrais do desvio-padrão e da variância

Faça as questões abaixo e depois transcreva suas respostas no formulário correspondente <https://forms.gle/ARbmKsQqKf7z57>. Esta atividade poder ser entregue até às **23h59** do dia 10/11.

Esta atividade visa discutir as estimativas amostrais do desvio-padrão, $s = \sqrt{\frac{1}{N-1} \sum_{i=1}^N (x_i - x_m)^2}$, e da variância, $V = s^2$, em termos do número de dados do conjunto.

Nesta atividade iremos considerar dados estatisticamente independentes com função densidade de probabilidade gaussiana, de valor verdadeiro $x_0 = 0$ e desvio-padrão verdadeiro $\sigma_0 = 1$. Gere $M = 10.000$ conjuntos de $N = 5$ dados cada e, para cada conjunto, calcule o desvio-padrão amostral, s , e a variância amostral, V .

a) Faça histogramas dos M valores de s e de V (**imagens desses histogramas para os casos de $N=100$, $N=10$ e $N=3$ deverão ser enviados**).

b) Usando os M valores de s e V , calcule as grandezas pedidas a seguir, preenchendo a coluna correspondente da tabela:

b.1) Calcule s_m , a média dos M valores do desvio-padrão amostral, s . Estime a incerteza de s_m .

b.2) Calcule V_m , o valor médio dos M valores da variância amostral, $V = s^2$. Estime a incerteza de V_m .

b.3) Quantos dos M valores de s são menores que o desvio-padrão verdadeiro, σ_0 ? E quantos valores de V são menores que a variância verdadeira, $V_0 = \sigma_0^2$? Estime a incerteza nessas contagens.

Repita para os casos em que $N = 100, 50, 10, 4, 3$ e 2 (sempre para $M = 10.000$ conjuntos) e preencha toda a tabela (**uma foto ou arquivo pdf da tabela preenchida deverá ser enviada**).

	(1) N=100	(2) N=50	(3) N=10	(4) N=5	(5) N=4	(6) N=3	(7) N=2
s_m (média de s)							
V_m (média de V)							
Número de casos em que $s \leq \sigma_0$							
Número de casos em que $V \leq V_0$							

1) Avalie a forma dos histogramas de s e de V . Para um mesmo valor de N , qual das duas grandezas tem o histograma mais assimétrico?

2) O desvio-padrão amostral é um estimador tendencioso? E a variância amostral? Um estimador é dito tendencioso quando seu valor esperado não é compatível com o valor verdadeiro da grandeza. Na prática, o valor esperado pode ser aproximado pela média de muitas simulações independentes (mas a incerteza da média precisa ser considerada na hora de tirar conclusões sobre a tendenciosidade do estimador).

3) Com base nos seus conhecimentos sobre o método da máxima verossimilhança e nos resultados numéricos obtidos neste experimento indique qual dos complementos propostos para a afirmação a seguir é o correto (há apenas uma opção correta): ***“Na expressão para estimar o desvio-padrão a partir***

de N medições independentes de uma mesma grandeza, $s = \sqrt{\frac{\sum_{i=1}^N (x_i - x_m)^2}{N-1}}$, o “-1” no denominador é necessário porque ...”

- A) “... desta forma a função densidade de probabilidade do desvio-padrão amostral se torna aproximadamente gaussiana.”
- B) “... assim o valor esperado para o desvio-padrão amostral coincide com o valor verdadeiro do desvio-padrão.”
- C) “... dessa forma a chance de subestimar o desvio-padrão amostral (isto é, obter um valor de s que seja menor que o valor verdadeiro do desvio-padrão) é a mesma de superestimá-lo.”
- D) “... assim o valor esperado para a variância amostral (o quadrado do desvio-padrão amostral) coincide com o valor verdadeiro da variância.”
- E) “... dessa forma a chance de subestimar a variância amostral (isto é, obter um valor de V que seja menor que o valor verdadeiro da variância) é a mesma de superestimá-la.”