

MAP2212 - Laboratório de Computação e Simulação

Relatório - EP04

Ayrton Amaral Alves Vítor - 11288131
Vítor Garcia Comissoli - 11810411

1 Introdução

O objetivo desse EP é a estimação da área ($u(v)$) de uma função dirichlet dado um certo limite superior V . Para estimar tal área deve-se utilizar uma aproximação por "histograma", ou seja, a área deve ser dividida em um número k de retângulos no eixo y (retângulos empilhados uns acima dos outros na horizontal ao invés de lado a lado na vertical), denominados bandas (ou bins). Para isso deve-se obter um valor para k , e depois somar as áreas de todas as k bandas (bins) geradas.

Além disso, deve-se obter um erro de, no máximo, 0.05%, com 95% de confiança, ou seja, que fique dentro do Intervalo de Confiança $[u(v) - 0.0005, u(v) + 0.0005]$ com $\alpha = 0.95$.

2 Discussão Teórica

Seja \mathbf{X} uma v.a. de Dirichlet e tomemos uma amostra aleatória simples $\{X_i | 1 \leq i \leq n\}$ de \mathbf{X} . Queremos estimar o integral $\int_{T(v)} f(\theta|x, y) d\theta$ em que $T(v) = \{\theta \in \Theta | f(\theta|x, y) \leq v\}$.

Temos, pelo enunciado, que $\epsilon = 0.0005$. Pode-se então dizer que o erro se dá por: $|u(v) - \hat{u}(v)| \leq 0.0005$, onde $u(v)$ é a área da função real e $\hat{u}(v)$ é a estimativa para a área da função estimada pelo programa.

Primeiramente, devemos estabelecer um valor para k para que, fixando ele, possamos encontrar um valor para n que satisfaz o intervalo de confiança do erro proposto. Como k depende de n , e ainda não temos um valor para n , tomaremos um k múltiplo de n . Assim, fixaremos $k = \frac{1}{2} \cdot n$ (já que assim cada retângulo do "histograma" conterá 2 pontos com distribuição dirichlet).

2.1 Obtenção Amostral de n e k

Fixando os valores do input (\mathbf{x} , \mathbf{y} e \mathbf{v}) e tomando $k = \frac{n}{2}$, criamos uma função que utilizou-se do desvio padrão de vários valores de n , até que fosse encontrado um valor que garantisse um erro (desvio padrão) ≤ 0.0005 . Desse teste, obteve-se que $n = 36000000$ é o menor valor para n testado que garante o erro esperado. Assim, tomaremos n como 36000000

Como para $n = 36000000$, $k = \frac{n}{2}$ se mostrou coerente com o erro de 0.05%, temos que $k = 18000000$ é um valor suficientemente preciso para k . Tomaremos assim k como 18000000

Assim, encontramos valores tanto para n quando para k que satisfazem o intervalo de confiança do erro estipulado pelo exercício.

3 Discussão do Código

O código da classe Estimador consiste de duas funções de classe (`_init_` e `U`) e 3 funções adicionais (`dirichlet`, `f_dirichlet` e `gera_v`).

A função `dirichlet` recebe **alpha** e **n** e devolve um array com dimensão (n,3) que possui valores gerados por uma função geradora de probabilidade dirichlet

A função `f_dirichlet` recebe **alpha** e **theta** e devolve um array com os valores da função densidade de probabilidade de uma dirichlet.

A função `gera_v` recebe **k** e os valores simulados da Dirichlet, os ordena e divide em k bins, então retorna o valor máximo de cada bin.

Ao instanciar um objeto Estimador, nos utilizamos das funções anteriores para obter os valores da `f_dirichlet`. Então estimamos o integral pela função U, que computa a razão entre os valores que são menores que o valor **v** dado e o número total de valores gerados.

4 Resultado

Como já discutido na subsection 2.1, os valores de **n** e **k** encontrados são suficientemente precisos para que o erro esteja dentro do intervalo estipulado pelo enunciado do EP.

Já quanto a estimativa $u(\hat{v})$, pode-se dizer que ela se aproxima ao valor de real de $u(v)$, tanto que, com um valor de **n** suficientemente grande, o erro da estimativa tende a 0, o que demonstra a eficácia da estimativa obtida.

5 Conclusão

Pode-se concluir que o programa atingiu os objetivos que lhe foram propostos, já que os valores obtidos pelos testes acima se mostraram coerentes com aquilo que era esperado pela parte teórica (pela obtenção de valores para **n** e **k** que assegurassem o intervalo de confiança de erro proposto para **u(v)**), e com aquilo que foi solicitado no enunciado do EP (obtenção de $u(\hat{v})$).