EP5 - FBST

George Othon - NUSP 103xxxxx
Felipe Zaffalon - NUSP 103xxxxx

Junho de 2020

1 Introdução

Nesse EP deveriamos implementar um teste de Hipótese apresentado no artigo "Evidence and Credibility: Full Bayesian Signicance Test for Precise Hypotheses", para isso deveriamos reproduzir os resultados do e-valor da hypothese de Hardy-Weinberg para os resultados observacionais constantes na Tabela 2 do Artigo.

2 Função e hipótese nula

Para testar a Hipótese Nula precisariamos primeiro otimizar a função:

$$f(\theta_1, \theta_2, \theta_3, x1, x2, x3) = \theta_1^{x1}.\theta_2^{x2}.\theta_3^{x3}$$

Com as condições de $1 = \theta_1 + \theta_2 + \theta_3, \theta_1, \theta_2$, $\theta_3 > 0$, e a condição da hipótese de θ_3 = (1 - $\sqrt{\theta_1}$)^2, com x1 e x3 tabelos e x2 = n - x1 - x3 com n = 20. Para isso escrevemos a função f em torno de θ_1 e um índice i para definir os x1, x2 e x3 necessarios para os parametros, na otimização utilizamos a função optimise() do r, com maximum = TRUE para receber o valor do máximo da função, com isso temos dois vetores de resultados, um deles com os θ_{max} para cada índice de x1,x2,x3, e um com os f_{max} para cada um desses θ_{max}.

3 Função de integração por MCMC

Nessa parte precisavamos calcular o e-valor apresentado pela função em cada um dos trios usando o metódo de integração por MCMC. Para a segunda parte do ep, a integração por MCMC, escrevemos uma função que recebe os x1 e x3 tabelados e os valores s1 e s3, sendo esses a variância usada para o cálculo da normal. Num primeiro momento geramos θ_1, θ_3 aleatórios, e θ_2 em função de ambos, sendo assim o trio (θ_1, θ_2, θ_3) sempre respeitasse as condições ditas anteriormente, depois disso ocorria o cálculo da normal e se gerava mais um trio de θ's, e comparando ambos os trios se decidia qual deles entraria para a lista retornada usando o aplha de aceitação de metrópolis, após a finalização da lista comparavamos agora com os valores obtidos após a otimização no primeiro momento, depois calculávamos a média dessa iteração nos devolvendo o objetivo.

4 Calibragem do parâmetro da normal

Por muitos testes empíricos e gerações com multiplos parametros decidimos que um valor aceitável era o valor de 1.

5 Escolha de passos

Para uma boa precisão fizemos 5 iterações utilizando a média dos 5 como resultado para cálculo do e-valor, sem necessidade de burn-in, os resultados ao final de todo o processo se mostraram consistentemente próximos.

6 Método de Metropolis

Para a escolha de α, usamos o Método de Metropolis dado por:

$$\alpha = min(1, \frac{g(\theta_{1proposto}, \theta_{3proposto}, índice)}{g(\theta_{1atual}, \theta_{3atual}, índice)})$$

sendo indice referente aos valores da tabela para x1 e x3.