

## Exercício de programação:

### Simulação e intervalos de credibilidade aplicados à análise de pesquisas eleitorais:

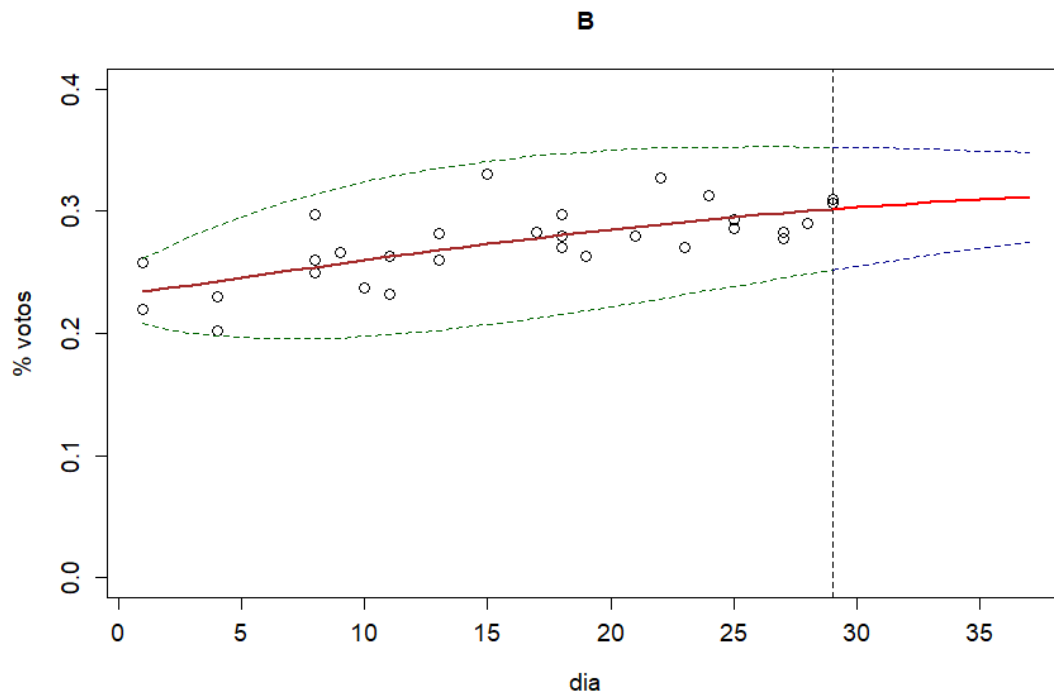
Problema: predição do resultado das eleições a partir de histórico de pesquisas eleitorais

Data sugerida para entrega: 26/11/2018

Usaremos os dados do arquivo pesquisa\_1T.txt contendo as pesquisas eleitorais do 1º turno, realizadas de 1 a 29/09/2018. Consideraremos os candidatos: B, H, C, O, BN.

Tarefas:

1. Ajustar as séries de pesquisas eleitorais por uma regressão Dirichlet. Considere como covariáveis o dia da pesquisa (número sequencial em que 01/09 corresponde ao dia 1) e a raiz quadrada do dia  
Use a função `DirichReg` do pacote `DirichletReg`
2. A partir do modelo ajustado, extrapolar as proporções de votos nos candidatos para o dia 7/10/2015 (o que equivale a prever as proporções no dia 37)  
Usar a função `predict` do pacote `DirichletReg`
3. Calcular os intervalos de credibilidade dos votos em cada candidato por data (do dia 1 até o dia 37), via máxima densidade a posteriori (HPD).  
Plotar o gráfico para cada candidato, conforme figura abaixo:



Para calcular os intervalos de máxima densidade a posteriori (HPD):

a) Recupere a matriz alpha resultante da função `predict`; essa matriz conterá os hiperparâmetros das distribuições Dirichlet para cada dia e cada candidato, ou seja:

$\alpha_{i,j}$  = hiperparâmetro correspondente à votação do candidato  $j$  no dia  $i$  (não confundir  $\alpha_{i,j}$  com  $\mu_{i,j}$ , que é a proporção de votos!)

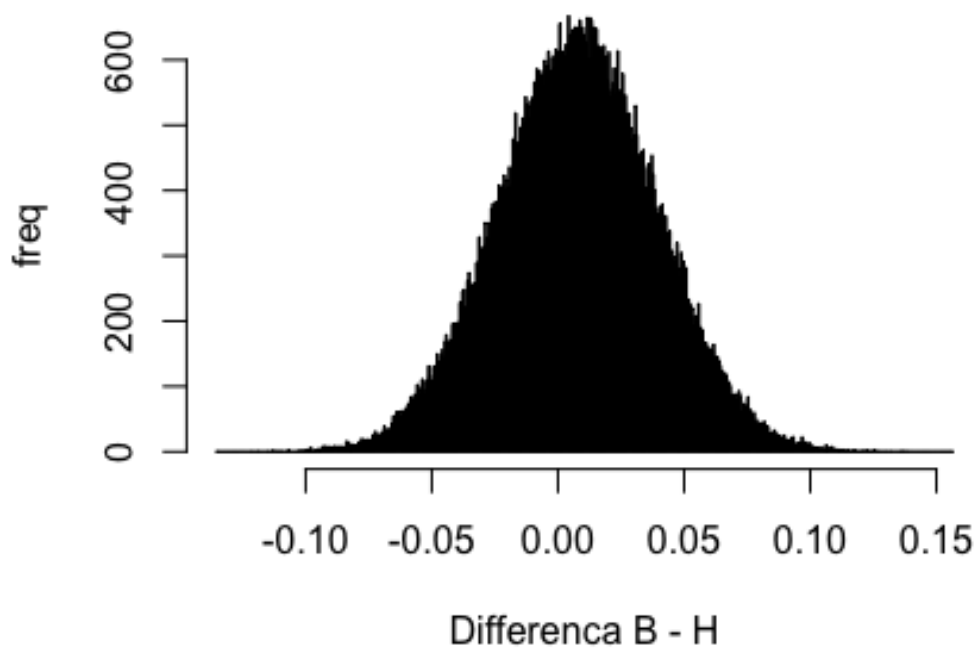
b) Para cada dia  $i$  e cada candidato  $j$ , converter a distribuição dos votos no candidato  $j$  “marginalizando” a distribuição Dirichlet, ou seja:

A distribuição de votos do candidato  $j$  no dia  $i$  será uma distribuição  $Beta(\alpha_{i,j}, \sum_{j^* \neq j} \alpha_{i,j^*})$  (notação:  $j^*$  corresponde a todos os demais candidatos diferentes do candidato  $j$ ).

c) Adapte o script “intcred\_hpd\_beta.r”, disponível no subdiretório EP2\_Pesquisa\_Eleitoral do Google Drive. Você precisará “encapsular” o trecho principal do cálculo do intervalo HPD, pois ele será executado para cada dia e cada candidato.

4. Sobre os valores de alpha projetados para o dia 7/10, simule vetores theta com a respectiva distribuição dirichet; sobre esses vetores theta simulados:

4.1. Plote um histograma das diferenças entre a proporção de votos no candidato B e a proporção de votos no candidato H (ver exemplo abaixo);



4.2. Calcule a probabilidade do candidato B ter votação maior do que o candidato H.