

Universidade Estadual de Campinas
Instituto de Matemática, Estatística e Computação Científica
Departamento de Estatística

ME 607 SÉRIES TEMPORAIS
Trabalho Prático 1

Professor: Mauricio Zevallos

Primeiro Semestre 2020

Instruções

1. O trabalho deverá ser desenvolvido somente pelos integrantes do grupo e sem consultar a outros colegas ou ao professor da disciplina.
2. O trabalho será disponibilizado no dia 21/07/20 às 16:00 horas na plataforma Moodle.
3. O trabalho deverá ser entregue até as 15:59 horas do dia 23/07/20 somente via plataforma Moodle. Envios fora do período especificado não serão permitidos na plataforma nem serão considerados na avaliação. Não haverá exceções.
4. O relatório do trabalho deve ser enviado em formato pdf e deve ser redigido em processador de texto (LaTeX, Word, etc). O número máximo de páginas é 15 (incluindo a capa, se tiver).
5. Adicionalmente, devem ser enviados os códigos R utilizados (comentados, de forma que as respostas sejam facilmente reproduzíveis), em formato txt. Os resultados apresentados no relatório (estatísticas, ajustes, tabelas e figuras) devem coincidir exatamente com os obtidos através dos códigos R enviados.
6. O relatório deverá ser redigido com clareza, com tabelas e figuras referenciadas. Esses aspectos terão bastante peso na avaliação. As saídas de computador (do R) devem ser colocadas em forma de tabela. Os resultados apresentados tem que ser interpretados e discutidos.
7. No envio do trabalho, identifique o arquivo pelo número do Grupo da seguinte forma: por exemplo, se seu Grupo é o 50, o nome do arquivo do relatório é T1-Grupo50.pdf, e o nome do arquivo dos códigos R é T1R-Grupo50.txt.
8. Não é necessário que todos os integrantes do grupo enviem o relatório e os códigos R. É suficiente um relatório e um arquivo de programas. Contudo, se o grupo preferir enviar mais de um (por exemplo, para não ter problemas de envio pela internet), não há problema desde que os arquivos enviados sejam idênticos, e o professor escolherá qualquer um.

Questão 1

(8 pts.) Interessa fazer a modelagem de séries de mortalidade em cidades para duas faixas etárias: entre 65 e 74 anos (nomeada **65ate74**) e maiores de 75 (nomeada **Maior75**). São disponíveis dados de mortalidade diária¹ de três cidades dos Estados Unidos: Chicago, New York e Los Angeles.

Este trabalho consiste na modelagem da mortalidade semanal construída a partir da série asignada ao seu Grupo (ver arquivo **Grupos-T1.pdf**) utilizando os modelos que sejam convenientes dentre os estudados na disciplina (conteúdos da disciplina especificados para a Prova 1). Assim, cada Grupo trabalhará com uma série de mortalidade semanal para uma cidade e faixa etária específicas². Considere como mortalidade semanal a soma das mortalidades diárias de domingo a sábado.

- (a) Descreva as características empíricas da série.
- (b) Ajuste um modelo para a série descartando as últimas 13 observações (aproximadamente 3 meses). O ajuste é satisfatório?
- (c) Faça o diagnóstico do modelo.
- (d) Calcule as previsões para as semanas descartadas no item (b) e compare as previsões com os valores reais. Avalie a qualidade do modelo em termos de previsão.
- (e) Um pesquisador afirma que a mortalidade está diminuindo no tempo. Você concorda com o pesquisador? Justifique sua resposta.

¹Se a série tiver NA's impute os valores considerando a média das observações vizinhas (observações anterior e posterior ao NA).

²Cada arquivo de dados contém duas colunas; na primeira é consignada a data e na segunda a mortalidade diária.

Questão 2

(2 pts.) Seja y_1, \dots, y_n uma amostra gerada pelo processo

$$y_t = z_t \varepsilon_t \quad (1)$$

$$z_t = \alpha_0 + \alpha_1 y_{t-1}^2 \quad (2)$$

onde $\{\varepsilon_t\}$ é uma sequência de variáveis aleatórias independentes cada uma com distribuição Normal com esperança zero e variância um. É possível demonstrar que

$$y_t^2 = \alpha_0 + \alpha_1 y_{t-1}^2 + \eta_t, \quad (3)$$

onde $\{\eta_t\}$ é uma sequência de variáveis aleatórias não-correlacionadas com esperança zero.

- (a) Proponha estimadores para α_0 e α_1 descrevendo a forma (método) pelo qual foram obtidos.
- (b) Implemente um programa na linguagem R para calcular os estimadores de α_0 e α_1 propostos em (a). Use seu programa para calcular as estimativas de α_0 e α_1 com base na série de 1000 observações assignada ao seu Grupo. Seja k o número do seu Grupo, então sua série corresponde à k -ésima coluna no arquivo **Series.txt**. Apresente os gráficos que considere relevantes.