

Testes para periodicidades

Dois testes de periodicidades que utilizam as ordenadas do periodograma.

Suponha que um modelo adequado para a série temporal observada seja

$$X_t = \sum_{i=1}^K R_i \cos(\omega_i t + \phi_i) + \varepsilon_t,$$

$t = 1, \dots, N$, em que K , R_i e ω_i , $i = 1, \dots, K$, são constantes, $\phi_i \sim U(-\pi; \pi)$ independentes, $\{\varepsilon_i\}$ sequência de ruídos brancos independentes de X_t , com média zero e variância σ_ε^2 .

Suponha que o processo X_t seja Gaussiano, Fisher (1929) apresenta um procedimento para testar a hipótese:

$H_0: R_i = 0$, todo i (não existe periodicidade)

baseado na estatística

$$g = \frac{\max_j I_j^{(N)}}{\sum_{j=1}^{\lfloor \frac{N}{2} \rfloor} I_j^{(N)}} = \frac{I^{(1)}}{\sum_{j=1}^{\lfloor \frac{N}{2} \rfloor} I_j^{(N)}}$$

Fisher mostrou que, para N ímpar, a distribuição exata de g , sob H_0 , é dada por

$$P(g > a) = n(1-a)^{n-1} - \binom{n}{2}(1-2a)^{n-1} + \dots + (-1)^x \binom{n}{x}(1-xa)^{n-1}, \quad (16.52)$$

em que

$$n = \left\lfloor \frac{N}{2} \right\rfloor \text{ e } x \text{ é o maior inteiro menor que } \frac{1}{a} \left(x = \left\lfloor \frac{1}{a} \right\rfloor \right).$$

Assim, para um dado nível de significância α , podemos utilizar a equação (16.52) para encontrar o valor crítico $a(\alpha)$ tal que $P(g > a(\alpha)) = \alpha$. Se o valor observado da estatística (g_{obs}) for maior que $a(\alpha)$, rejeitaremos H_0 , o que significa afirmar que a série apresenta uma periodicidade igual a $1/\omega^*$, sendo ω^* a frequência em ciclos, correspondente a $I^{(1)}$.

Uma boa aproximação para a expressão (16.52) é obtida utilizando somente o primeiro termo da expansão, isto é,

$$P(g > a) \equiv n(1-a)^{n-1}, \quad (16.53)$$

que pode ser utilizada para encontrar o p -valor aproximado do teste, fazendo $a = g_{\text{obs}}$.

Fisher (1929) obteve a expressão (16.52) utilizando argumentos geométricos; Grenander e Rosenblatt (1957) e Hannan (1970) apresentaram demonstrações analíticas.

Whittle (1952) sugeriu uma extensão para o teste de Fisher, que fornece um teste para a segunda maior ordenada do Periodograma ($I^{(2)}$), omitindo o termo $I^{(1)}$ do denominador da estatística g :

$$g' = \frac{I^{(2)}}{\sum_{j=1}^{\lfloor \frac{N}{2} \rfloor} I_j^{(N)} - I^{(1)}} \quad (16.54)$$

E utilizando a distribuição g de Fisher com n substituído por $(n-1)$.

Se a segunda maior ordenada for significativa, pode-se aplicar o procedimento de Whittle para testar a terceira maior ordenada e assim por diante, até que se obtenha um resultado não significativo.

O teste de Whittle é bom se todas as componentes periódicas tiverem frequências iguais ou muito próximas a múltiplos de $2\pi/N$.

Exemplo 16.7. Vamos aplicar o teste de Fisher à série A_1 - Cananéia.

A Tabela 16.1 apresenta as ordenadas do periodograma. Podemos notar que $\max_j I_j^{(120)} = 67,05$, para $j = 10$. Utilizando (16.51) podemos calcular o valor da estatística de Fisher

$$g_{\text{obs}} = \frac{67,05}{\sum_{j=1}^{60} I_j^{(120)}} = \frac{67,05}{75,65} = 0,8863$$

e o p -valor, utilizando (16.53), é

$$\alpha^* = P(g > g_{\text{obs}}) \cong n(1 - g_{\text{obs}})^{n-1},$$

ou seja,

$$\alpha^* = P(g > 0,8863) \cong 60(1 - 0,8863)^{59} = 1,1695 \times 10^{-54},$$

que é menor do que qualquer valor usual de nível de significância. Esse p -valor nos leva a rejeitar a hipótese de não existência de periodicidade e a concluir que a série de temperatura em Cananéia possui uma periodicidade de $N/j = 120/10 = 12$ meses.

j	$I_j^{(120)}$	j	$I_j^{(120)}$	j	$I_j^{(120)}$	j	$I_j^{(120)}$
1	0,13667568	16	0,040034	31	0,00560722	46	0,01067816
2	0	17	0,06870146	32	0,00090925	47	0,02509832
3	0,37729525	18	0,18475946	33	0,01394776	48	0,12607798
4	0,4443844	19	0,18127692	34	0,01435663	49	0,08271761
5	0,01765445	20	0,45529573	35	0,22100755	50	0,26626876
6	0,24634056	31	0,07874526	36	0,24970452	51	0,04673396
7	0,94905469	32	0,29463036	37	0,07654535	52	0,04417737
8	0,38951211	23	0,13239664	38	0,06224509	53	0,2718604
9	0,23746182	24	0,07490943	39	0,0470564	54	0,04708572
10	67,04292	25	0,0797324	40	0,13276383	55	0,06200695
11	0,39026704	26	0,26798099	41	0,15847224	56	0,03227634
12	0,11612016	27	0,1465865	42	0,06690099	57	0,05022613
13	0,08815388	28	0,10279556	43	0,0159557	58	0,09308288
14	0,46013097	29	0,10890914	44	0,0064507	59	0,03516205
15	0,01897867	30	0,12572323	45	0,10737339	60	0,02238609

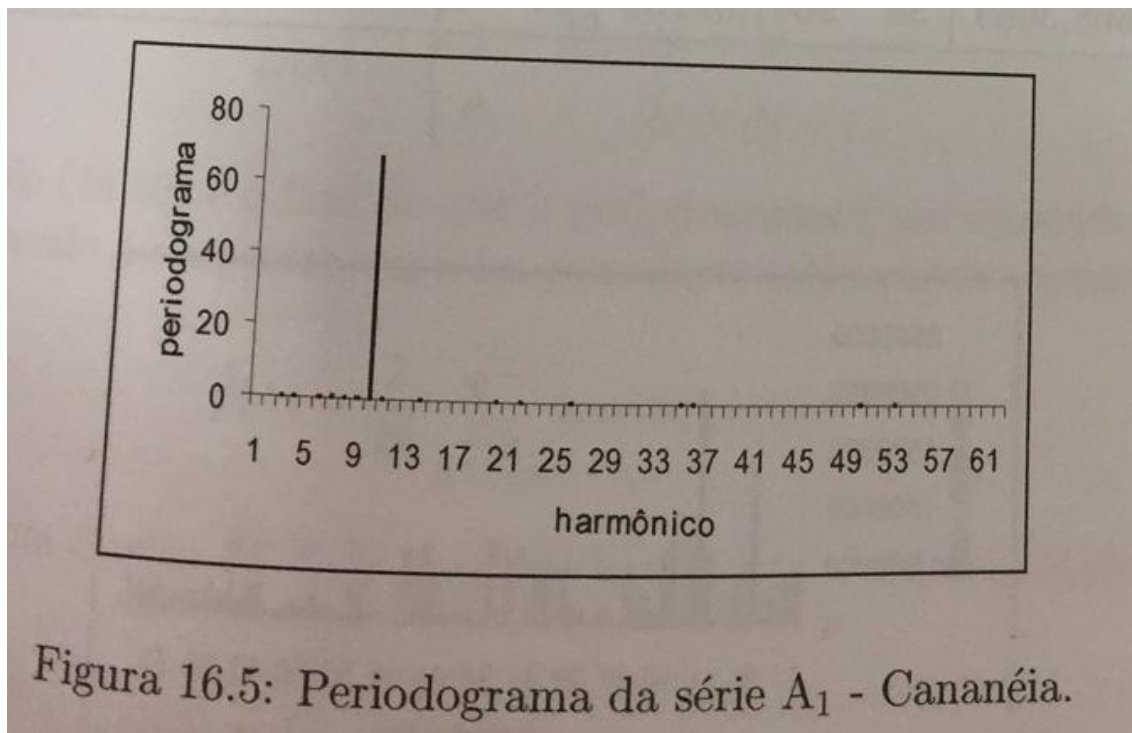


Figura 16.5: Periodograma da série A_1 - Cananéia.

Exemplo 16.8. Vamos utilizar, agora, a série A_3 - Fortaleza.
 As ordenadas do periodograma estão na Tabela 16.2. Dessa tabela podemos ver que

$$\max I_j^{(149)} = 312.911,17, \quad j = 12,$$

$$\sum_{j=1}^{75} I_j^{(149)} = 2.902.343,46.$$

Assim,

$$g_{\text{obs}} = \frac{312.911,17}{2.902.343,46} = 0,1078$$

e

$$\alpha^* = P(g > 0,1078) \approx 74 \times (1 - 0,1078)^{73} = 0,0179.$$

Fixando um nível de significância de 0,02, podemos concluir que existe uma periodicidade de $\frac{149}{12} = 12,41$ anos na série de chuvas em Fortaleza.

Para verificar a existência de uma segunda periodicidade, utilizamos a estatística g' , expressão (16.55), com

$$I^{(2)} = 211.756,89, \quad j = 6,$$

e

$$\sum_{j=1}^{\lfloor \frac{N}{2} \rfloor} I_j^{(149)} - I^{(1)} = 2.902.343,46 - 312.911,17 = 2.589.432,29.$$

Assim,

$$g'_{\text{obs}} = \frac{211.756,89}{2.589.432,29} = 0,08178$$

e

$$\alpha^* = P(g' > 0,0818) \approx 73(1 - 0,0818)^{72} = 0,1566,$$

indicando a não existência de uma segunda periodicidade na série.

j	$I_j^{(149)}$	j	$I_j^{(149)}$	j	$I_j^{(149)}$	j	$I_j^{(149)}$
1	71825,2778	20	75446,1267	39	5882,89354	58	2088,33069
2	54577,5753	21	5010,34324	40	26671,1439	59	3958,33957
3	25042,4995	22	7699,12301	41	52317,622	60	50554,7869
4	35083,4391	23	21954,2128	42	101904,875	61	35311,9109
5	6134,29402	24	2679,93863	43	64338,9275	62	20435,1454
6	211756,893	25	13169,2096	44	90852,7367	63	39293,0689
7	34983,4602	26	27716,5361	45	2320,59059	64	21472,3538
8	56387,493	27	72802,2969	46	1119,59624	65	34332,9211
9	2270,71428	28	5662,20033	47	12480,9849	66	15737,201
10	13740,1166	29	75872,9111	48	5922,79608	67	27449,8348
11	90433,7081	30	65851,7202	49	45336,603	68	23127,1768
12	312911,171	31	24695,2296	50	44273,7432	69	19528,1665
13	65138,2003	32	941,392541	51	2149,52817	70	9809,83456
14	9561,99929	33	82205,4893	52	5307,47647	71	39907,0876
15	113937,929	34	10620,548	53	13883,16	72	19403,1295
16	93864,3853	35	112482,346	54	27344,5755	73	31527,3867
17	13538,6463	36	2378,17506	55	21200,4976	74	40318,1897
18	40247,671	37	18050,3804	56	19050,0286	75	1705,75352
19	25608,3669	38	9041,59166	57	12801,4516		

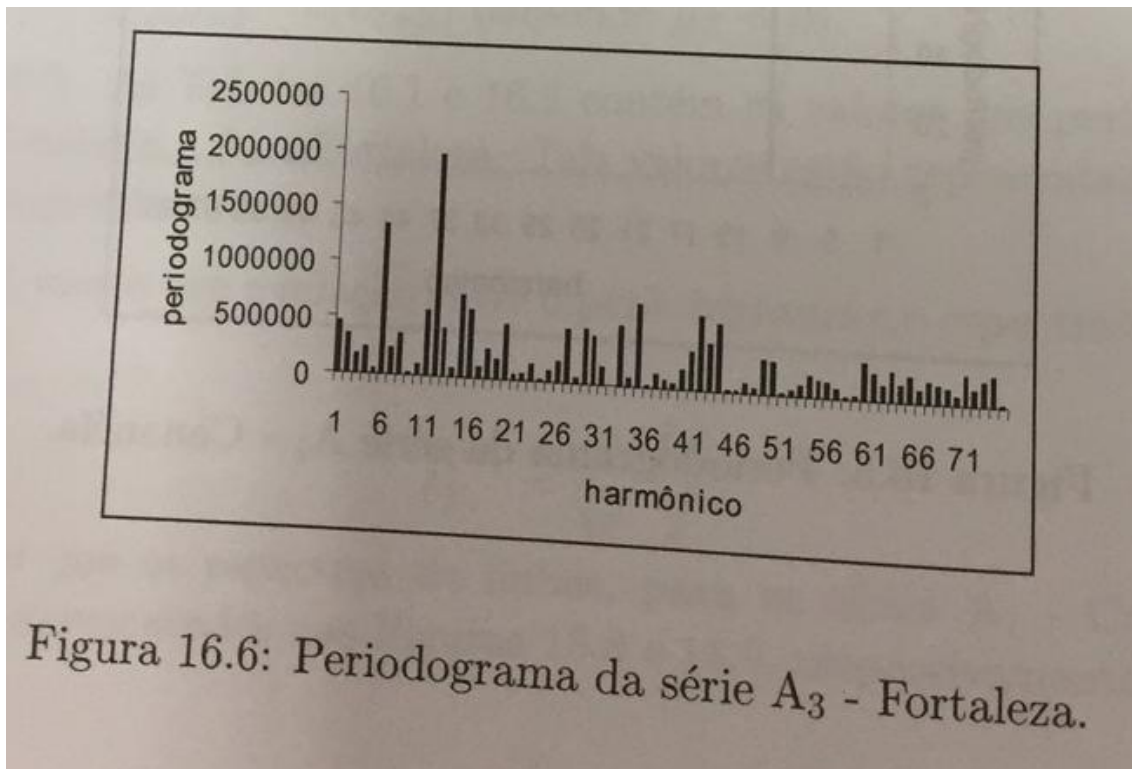


Figura 16.6: Periodograma da série A_3 - Fortaleza.