

5 Aplicações

5.1 Dados Utilizados

Para exemplificar o modelo principal desta dissertação e compará-lo com outros modelos tradicionais, serão utilizados dados na forma incremental. Os dados da tabela 5.1 referem-se à “*Historical Loss Development Study*” – edição de 1991 – publicado pela Associação de Resseguros da America (RAA). Estes dados foram utilizados por Mack[12], por England & Verrall[5] e por De Jong[3]. Suas observações estão na unidade milhar de dólar. Por questões de facilidade, esses dados serão denominados simplesmente de ‘AFG’.

Ano de Origem w	Desenvolvimento d									
	0	1	2	3	4	5	6	7	8	9
1	5012	3257	2638	898	1734	2642	1828	599	54	172
2	106	4179	1111	5270	3116	1817	-103	673	535	
3	3410	5582	4881	2268	2594	3479	649	603		
4	5655	5900	4211	5500	2159	2658	984			
5	1092	8473	6271	6333	3786	225				
6	1513	4932	5257	1233	2917					
7	557	3463	6926	1368						
8	1351	5596	6165							
9	3133	2262								
10	2063									

Tabela 5.1: Dívidas Gerais Facultativas (excluindo ambiental) do Estudo da Evolução Histórica de Perdas (1991 - unidade milhar de dólar)–AFG.

Também serão utilizados os dados que se encontram em De Jong & Zehnwirth[2] como exemplo para o modelo por eles proposto, o qual serve de base para o nosso modelo. Segundo os autores, estes dados – que encontram-se na tabela 5.2 – referem-se às atividades de uma seguradora inglesa

durante o período de 1970 a 1974 inclusive. Esses dados serão referidos como ‘DJZ’.

Ano de Origem w	Desenvolvimento d				
	0	1	2	3	4
1	753,5	648,9	311,7	173,5	71,3
2	642,3	648,4	249,7	206,5	
3	715,8	661,1	309,4		
4	841,6	862,6			
5	968,8				

Tabela 5.2: Triângulo de *runoff* para uma seguradora inglesa (unidade milhar de libra)–DJZ.

5.2

Aplicação do método Chain Ladder

Primeiro será feita a estimação da reserva IBNR através do método *Chain Ladder* clássico, descrita na seção 2.3. Os resultados apresentados a seguir – para os dados AFG – são semelhantes aos encontrados em England & Verrall[5], nos quais não são levados em consideração qualquer *fator de cauda* e os *development factors* finais são os que resultaram da média demonstrada pela equação (2-2) (e não valores “selecionados” pelo atuário).

Desenvolvimento d								
0:1	1:2	2:3	3:4	4:5	5:6	6:7	7:8	8:9
2,999	1,624	1,271	1,172	1,113	1,042	1,033	1,017	1,009

Tabela 5.3: *Development factors* referentes aos dados AFG.

Os *development factors* dos dados AFG estão representados na tabela 5.3. Como explicado na seção 2.3, estes fatores são utilizados para completar o triângulo de *runoff*, e daí obter-se as previsões IBNR estimadas. As referidas estimativas encontram-se na tabela 5.4.

Para o conjunto de dados DJZ, os *development factors* encontram-se na tabela 5.5

5.3

Aplicação do Modelo de Espaço de Estado

Nesta seção, será ilustrada a técnica de espaço de estado e o filtro de Kalman descritos na seção 4.3. Como afirmado na referida seção, não serão levados em consideração *índices de volume* ou *índices de preço* nos exemplos a seguir. O motivo disto é poder comparar os resultados com os do método *chain ladder*, que não os leva em conta. Além disso, ao utilizar o *índice de preço*, é preciso prever seus valores futuros de forma a completar o triângulo de *runoff*. Neste caso, haveria a necessidade de um outro modelo para prever estes índices.

Portanto, será utilizada a equação (4-1), ao invés de (4-2) para modelar as observações $y_d(t)$.

5.3.1

Dados AFG

Tomem-se os dados da tabela 5.1 como primeiro exemplo. Em termos da notação introduzida em (4-5), os vetores das observações sucessivas $y(t)$

Ano de Origem	Reservas
1	0
2	154
3	617
4	1636
5	2747
6	3649
7	5435
8	10907
9	10650
10	16339
Total	52135

Tabela 5.4: Reservas de AFG calculadas através do método *Chain Ladder*.

Desenvolvimento d			
0:1	1:2	2:3	3:4
1,9552	1,2140	1,1168	1,0378

Tabela 5.5: *Development factors* referentes aos dados DJZ.

serão dados por

$$y(1) = (5012)$$

$$y(2) = (106, 3257)'$$

$$y(3) = (3410, 4179, 2638)'$$

$$y(4) = (5655, 5582, 1111, 898)'$$

$$y(5) = (1092, 5900, 4881, 5270, 1734)'$$

$$y(6) = (1513, 8473, 4211, 2268, 3116, 2642)'$$

$$y(7) = (557, 4932, 6271, 5500, 2594, 1817, 1828)'$$

$$y(8) = (1351, 3463, 5257, 6333, 2159, 3479, -103, 599)'$$

$$y(9) = (3133, 5596, 6926, 1233, 3786, 2658, 649, 673, 54)'$$

$$y(10) = (2063, 2262, 6165, 1368, 2917, 225, 984, 603, 535, 172)'$$

que nada mais são do que as diagonais sucessivas do triângulo de *runoff* da tabela 5.1.

A primeira observação $y(1)$ possui uma média deflacionada $m(1, 0)$, a qual é parametrizada em termos de $b(1)$. O segundo vetor $y(2)$ sofre influência de $m(2, 0)$ e $m(1, 1)$, que são parametrizados por $b(2)$ e $b(1)$ respectivamente. Para o t -ésimo vetor de observações $y(t)$ tem-se a influência de $m(t, 0)$, $m(t-1, 1)$, \dots , $m(1, t-1)$, os quais são parametrizados em termos de $b(t)$, $b(t-1)$, \dots , $b(1)$. Estes últimos parâmetros constituem o vetor $\beta(t)$ ou “estado” no instante t . Ou seja, o vetor de estado no instante t nada mais é do que todos os parâmetros desconhecidos $b(w)$ envolvidos nas observações realizadas no instante t .

Para este conjunto de dados, é necessário definir uma função básica que faça sentido. Como já argumentado no capítulo 4, essa função seria uma espécie de conhecimento *a priori* do comportamento decrescente dos

Ano de	
Origem	Reservas
1	0
2	65,985
3	268,03
4	693,45
5	1696,2
Total	2723,7

Tabela 5.6: Reservas de DJZ calculadas através do método *Chain Ladder*.

valores observados (em um triângulo de *runoff* incremental), à medida que o atraso (Desenvolvimento) aumenta. Para auxiliar na escolha desta função básica, foi analisado este padrão de decaimento das observações em função do período de desenvolvimento d , fixando-se o ano de origem w . O gráfico 5.1 (p. 47) mostra este comportamento. A linha pontilhada preta refere-se à função básica escolhida para os dados:

$$\phi(d) = (d + 1)e^{-\frac{d}{2}}, \quad d = 0, 1, \dots, 9 \quad (5-1)$$

Com isso, pôde-se iniciar a estimação dos parâmetros e hiperparâmetros do modelo através da função de verossimilhança $L(\mathbf{y}; \boldsymbol{\psi})$, calculada pelo filtro de Kalman.

Como explicado na seção 4.3, foram utilizadas duas abordagens para a especificação do termo aleatório da equação das observações para cada conjunto de dados: uma homocedástica e outra heterocedástica. Para os dados AFG, a função log-verossimilhança do caso homocedástico tende ao infinito em zero, como se pode ver na figura 5.2(a) (p. 47). Neste caso o modelo terá uma relação sinal/ruído $r = 0$, significando que $\beta(t) = \beta(t - 1)$. Portanto, não há desconto, e o modelo se reduz a um modelo de regressão linear.

Então, partindo da premissa que o termo aleatório das observações é heterocedástico, a função verossimilhança possuirá um claro ponto de máximo, como mostra a figura 5.2(b). A relação sinal/ruído r do ponto de máximo da função $\log L$ é 473,4700.

Após estimados os parâmetros do modelo, utiliza-se a equação (4-19) para estimar os valores faltantes da tabela de *runoff* de forma a “completar o quadrado”. Os resultados previstos pelo modelo encontram-se na tabela 5.7.

Testes & Diagnósticos

Em seguida realizaram-se testes de adequação do modelo aplicado aos dados AFG. De posse dos resíduos (inovações padronizadas), foram traçadas as FAC¹ (figura 5.3, p. 50) para os quatro primeiros desenvolvimentos, isto é, para $d = 0, 1, 2, 3$. Observando os gráficos, podemos inferir que os resíduos praticamente não possuem correlação serial. O grande problema é que existem poucos dados para análise da FAC. No caso dos dados AFG, temos apenas nove observações para $d = 0, 1$, oito para $d = 2$ e sete para

¹Função de autocorrelação.

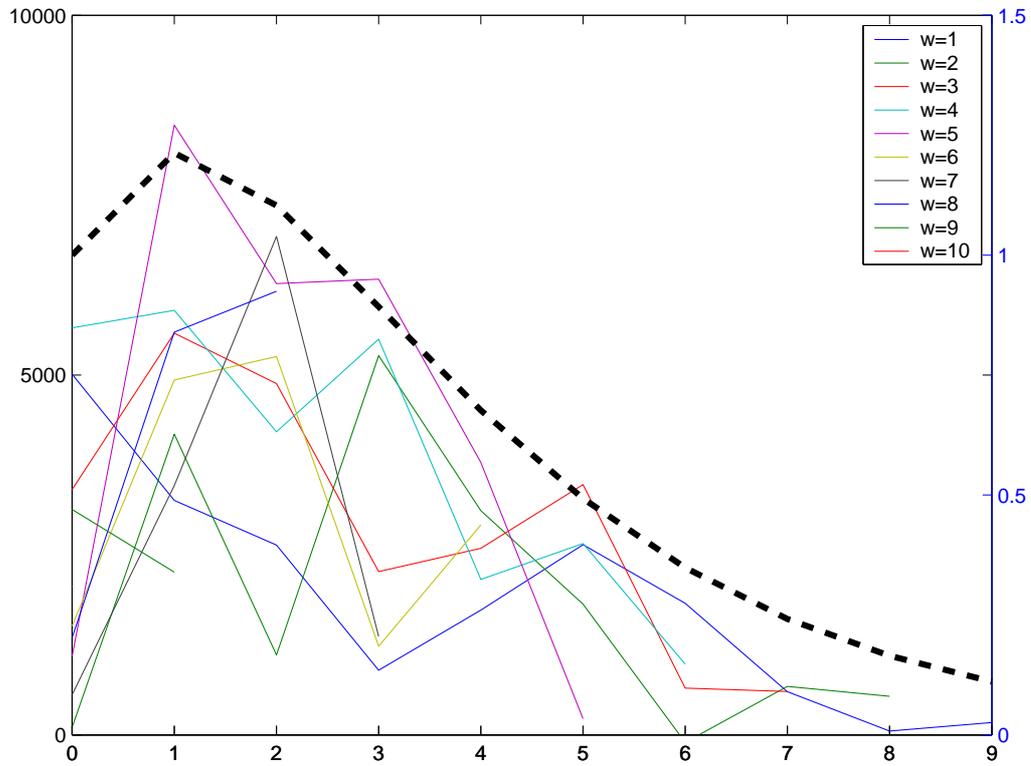
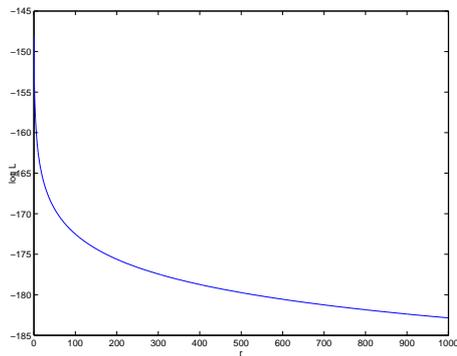
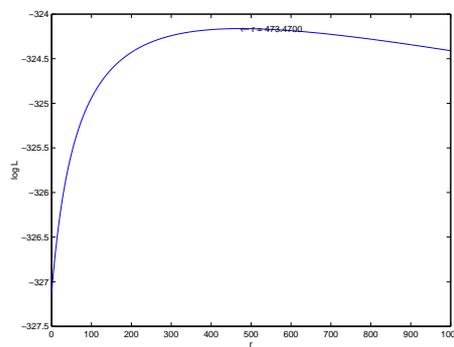


Figura 5.1: Distribuição dos valores fixando-se w (Dados AFG).



5.2(a): Homocedástico



5.2(b): Heterocedástico

Figura 5.2: Função $\log L$ para os dados AFG.

Ano Orig. w	Desenvolvimento d									
	0	1	2	3	4	5	6	7	8	9
1	5012	3257	2638	898	1734	2642	1828	599	54	172
2	106	4179	1111	5270	3116	1817	-103	673	535	460
3	3410	5582	4881	2268	2594	3479	649	603	686	462
4	5655	5900	4211	5500	2159	2658	984	1045	713	480
5	1092	8473	6271	6333	3786	225	1478	1025	699	471
6	1513	4932	5257	1233	2917	1825	1291	895	611	411
7	557	3463	6926	1368	2332	1697	1201	832	568	382
8	1351	5596	6165	3035	2301	1674	1185	821	560	377
9	3133	2262	3401	2750	2085	1517	1074	744	508	342
10	2063	3588	3265	2640	2001	1457	1031	714	487	328

Tabela 5.7: Triângulo de *runoff* completo para os dados AFG.

Estatística de Jarque-Bera		
Desenvolvimento	Valor	p-valor
$d = 0$	0,5453	0,7613
$d = 1$	0,6707	0,7151
$d = 2$	0,8218	0,6630
$d = 3$	1,1571	0,5607

Tabela 5.8: Resultado de teste de Jarque-Bera para os dados AFG.

$d = 3$. Vale lembrar também que para cada autocorrelação $r_v(\tau)$, teremos $T - \tau$ observações. Isto faz com que as estimativas da FAC possuam baixa precisão, podendo ser facilmente afetadas por valores extremos dos resíduos.

As estatísticas de Jarque-Bera (teste de normalidade) para os quatro desenvolvimentos em questão, que se encontram na tabela 5.8, reforçam a idéia de normalidade nos resíduos, apesar do número de observações ser pequeno.

Foi realizada também uma análise seccional dos resíduos padronizados, juntando todos em uma só amostra atemporal com um número maior de elementos. Observando o gráfico *quantil-quantil* (gráfico 5.4, p. 51), percebe-se que os resíduos têm distribuição normal.

5.3.2 Dados DJZ

Nesta seção será explorada a aplicação do modelo nos dados encontrados em De Jong & Zehnwirth[2]. Entretanto, há uma diferença de premissas adotadas entre o nosso enfoque e o dos autores:

- não utilização dos *índices de preço* e de *volume*.
- no trabalho dos autores eles não levaram em conta a influência da matriz covariância do erro da equação das observações $U(t)$ no desvio padrão do resultado final. No nosso enfoque a influência desta matriz está incluída.

A função básica utilizada é a mesma da definida em De Jong & Zehnwirth[2].

$$\phi(d) = (d + 1)e^{-d}, \quad d = 0, 1, \dots, 4$$

Assim como realizado para os dados AFG, foi traçado o gráfico das observações $y_d(t)$ pelos Desenvolvidos d , dado os anos de origem w fixos. O comportamento observado encontra-se na figura 5.5 (p. 51)

Os vetores das observações sucessivas $y(t)$ serão dados por

$$y(1) = (753, 5)$$

$$y(2) = (643, 3; 3257)'$$

$$y(3) = (715, 8; 648, 4; 311, 7)'$$

$$y(4) = (841, 6; 661, 1; 249, 6; 173, 5)'$$

$$y(5) = (968, 8; 862, 6; 309, 4; 206, 5; 71, 3)'$$

Assim como no exemplo anterior, os parâmetros do modelo foram estimados de duas maneiras: com ruído das observações homocedástico e heterocedástico. Também repetiu-se o mesmo comportamento do exemplo anterior nos gráficos de $\log L$ em função da relação sinal ruído r : no caso homocedástico o ponto de máximo converge para relação sinal/ruído zero, enquanto que no heterocedástico, a relação sinal/ruído fica em 55,9866 no ponto de máximo. Os gráficos encontram-se na página 53 (figura 5.6).

Utilizando então a relação sinal/ruído obtida no caso heterocedástico, obteve-se a estimação da previsão $\hat{y}(s + 1)$, completando, assim a seção inferior do triângulo de *runoff*, conforme mostra a tabela 5.9.

Testes & Diagnósticos

Ao se realizar os diagnósticos/testes neste conjunto de dados, é preciso ter em mente que temos poucas observações. Neste caso especificamente, temos apenas cinco instantes de tempo t . Por isso só serão aplicados os testes nas inovações padronizadas referentes a dois desenvolvimentos

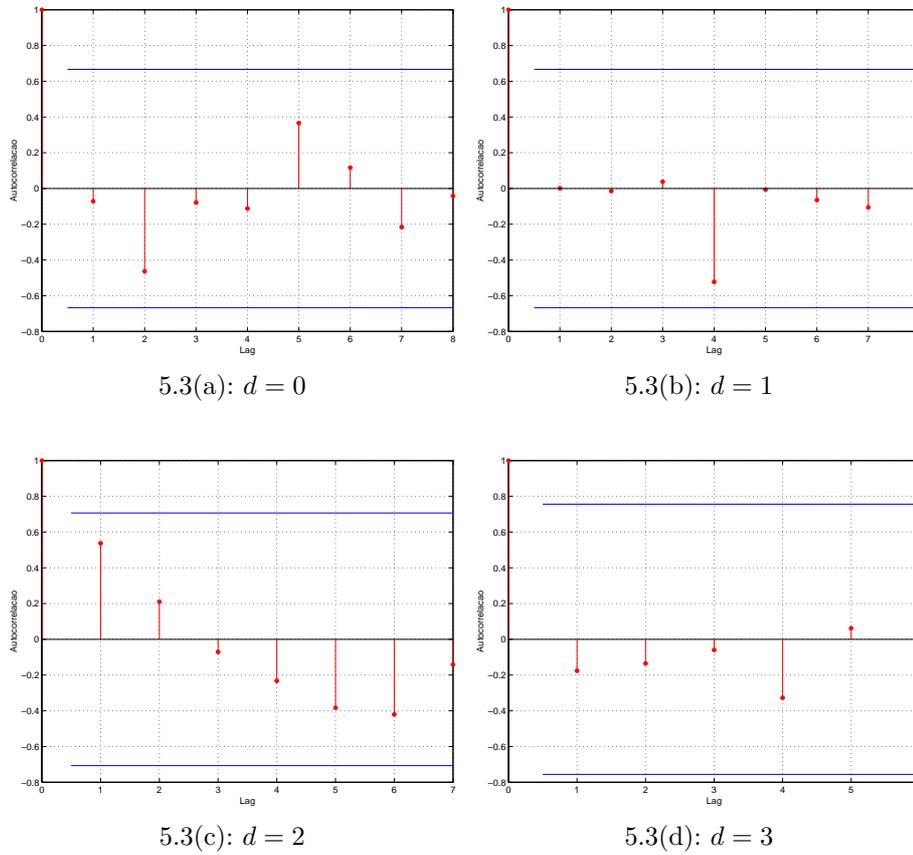


Figura 5.3: Autocorrelação das Inovações – dados AFG.

Ano de Origem w	Desenvolvimento d				
	0	1	2	3	4
1	753,5	648,9	311,7	173,5	71,3
2	642,3	648,4	249,7	206,5	70,9646
3	715,8	661,1	309,4	159,5727	73,3794
4	841,6	862,6	336,9318	165,2671	75,9979
5	968,8	617,5062	340,7517	167,1407	76,8596

Tabela 5.9: Triângulo de *runoff* completo para os dados DJZ.

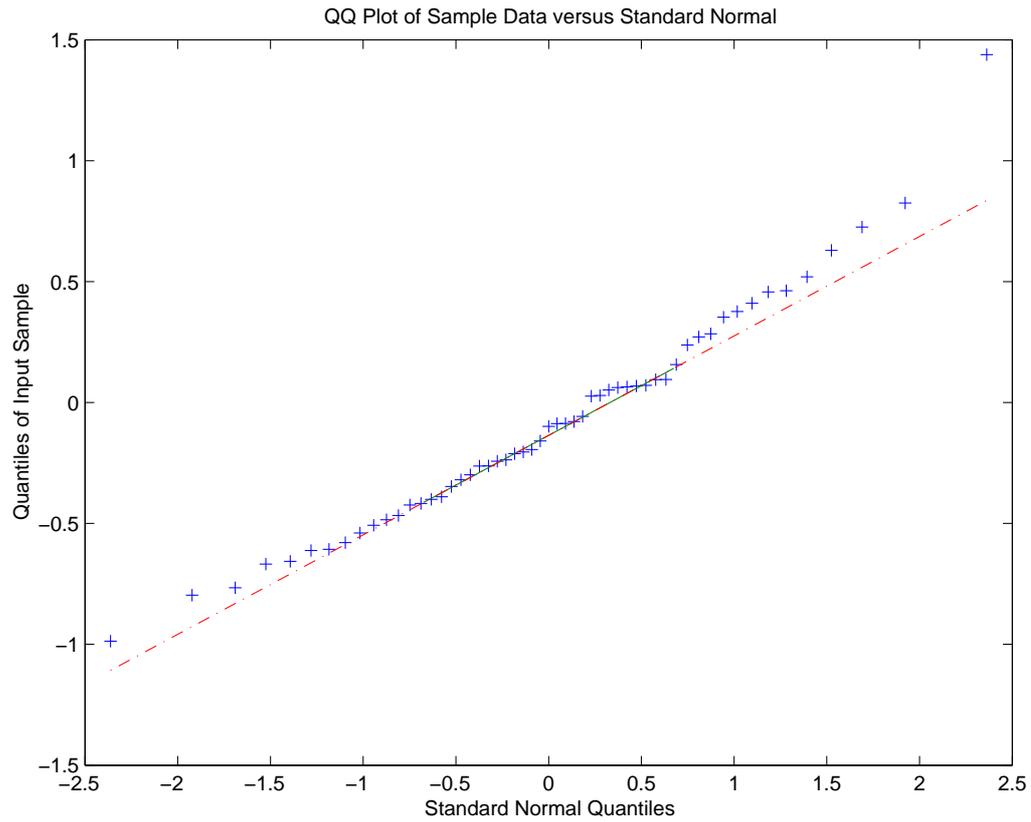


Figura 5.4: Gráfico quantil-quantil dos quantis dos resíduos pelos quantis teóricos (Dados AFG).

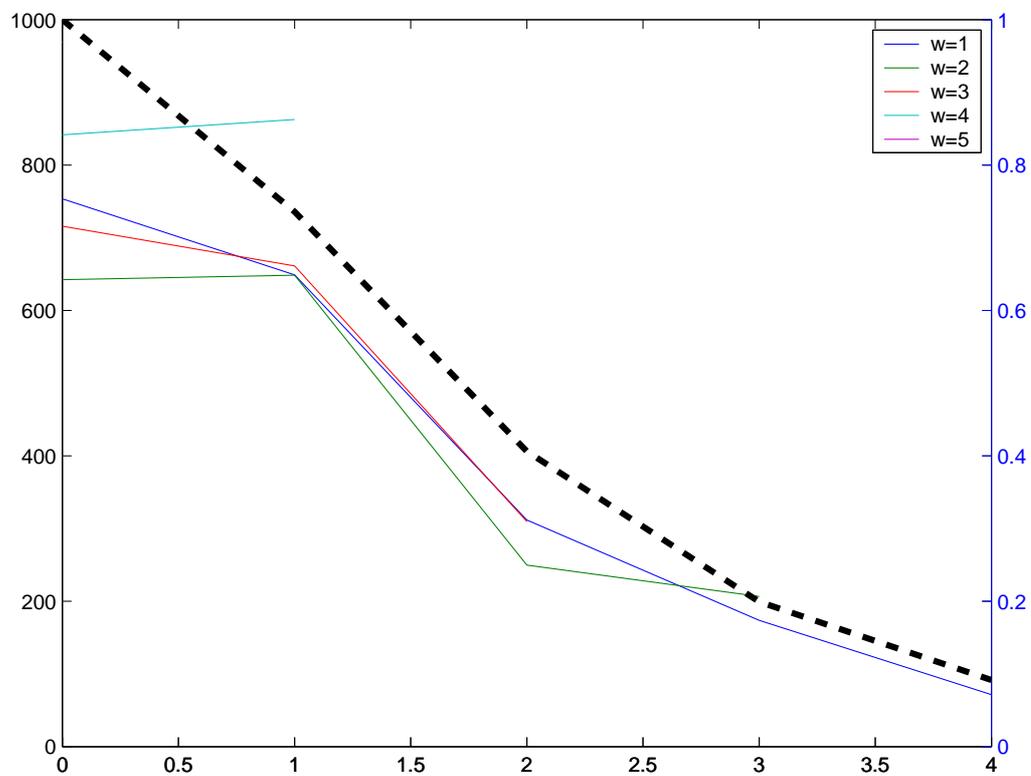


Figura 5.5: Distribuição dos valores fixando-se w (Dados DJZ)

Estatística de Jarque-Bera		
Desenvolvimento	Valor	p-valor
$d = 0$	0,7662	0,6817
$d = 1$	0,8452	0,6553

Tabela 5.10: Resultado de teste de Jarque-Bera para os dados DJZ.

apenas: $d = 0, 1$. A cada Desenvolvimento acima de $d = 1$, perdemos uma observação, tornando o resultado do teste deprecável ou pouco expressivo.

O primeiro teste gráfico a ser realizado é a observação da FAC. A figura 5.7 (p. 53) mostra a autocorrelação dos resíduos do modelo. Pelos gráficos pode-se inferir que os resíduos são serialmente descorrelatados.

Na tabela 5.10 encontra-se a estatística de Jarque-Bera para os resíduos dos dados DJZ. Apesar de poucas observações, o resultado indica um forte indício de normalidade (o p-valor revela um forte aceite da hipótese nula de normalidade).

Como realizado nos dados AFG, aqui também será realizada a análise seccional dos dados. O gráfico quantil-quantil (gráfico 5.8, p. 55), também reflete normalidade nos resíduos.

5.3.3 Considerações Finais

A tabela 5.12 consolida os resultados obtidos das estimações pelo modelo *chain ladder* e o modelo de espaço de estado proposto. O desvio padrão final da previsão IBNR está extremamente alto nos dados AFG, que possui um comportamento que não condiz com uma das premissas do modelo, que é a redução do valor das observações referentes à um ano de origem w , à medida que o Desenvolvimento d aumenta. O referido conjunto de dados possui um valor negativo (-103) e valores $y_d(t) > y_{d'}(t')$ onde $d > d'$. A consequência disto é um intervalo de confiança exageradamente alto, como mostra a mesma tabela. Em virtude disso resolvemos verificar a influência da *função básica* na previsão. A tabela 5.11 mostra o comportamento da previsão da média e do desvio padrão à medida que alteramos a função básica. Alterando o denominador do fator ‘exponenciado’, acelera-se o decaimento. A média vai diminuindo, assim como o desvio padrão, mas este não diminui significativamente como esperávamos.

Os dados DJZ já são um pouco mais “bem-comportados”, mas mesmo assim o intervalo de confiança é bem alto. Os valores médios estimados

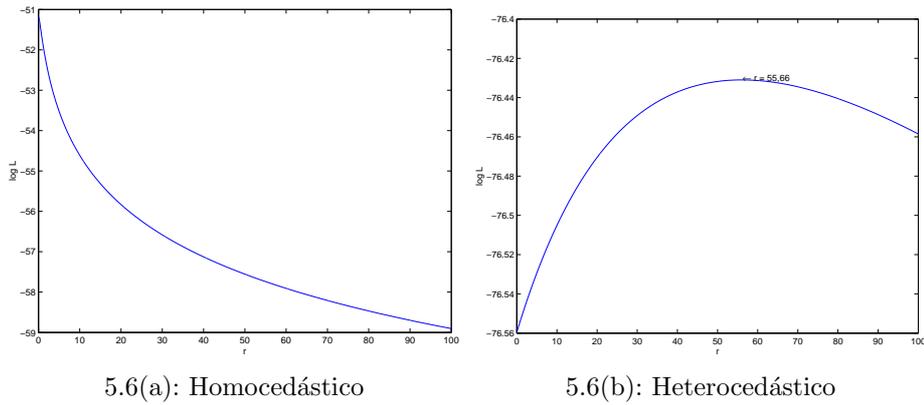


Figura 5.6: Função $\log L$ para os dados DJZ.

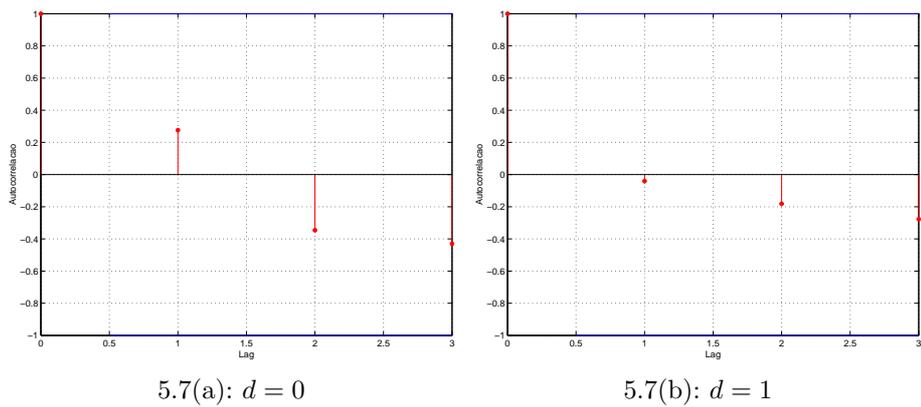


Figura 5.7: Autocorrelação das Inovações – dados DJZ.

Função Básica					
	$(d + 1)e^{-\frac{d}{1,9}}$	$(d + 1)e^{-\frac{d}{1,8}}$	$(d + 1)e^{-\frac{d}{1,7}}$	$(d + 1)e^{-\frac{d}{1,6}}$	$(d + 1)e^{-\frac{d}{1,5}}$
r	435,1623	393,4597	347,4241	289,4771	0
Reserva IBNR	52713	48115	4372,4	39667	–
Desvio Padrão	2289,2	2159,5	2067,2	2014,7	–

Tabela 5.11: Previsão IBNR variando a função básica dos dados AFG.

Reserva IBNR		
	AFG	DJZ
<i>Chain Ladder</i>	52135	2723,7
Modelo em Espaço de Estado	57472	2084,4
Desvio padrão do modelo	24524	383,4
Intervalo de Confiança (95%)		
Lim. Sup.	105540	2835,8
Lim. Inf.	9406,6	1332,9

Tabela 5.12: Previsão IBNR para os diferentes modelos.

diferem um pouco daqueles estimados pelo método *chain ladder*.

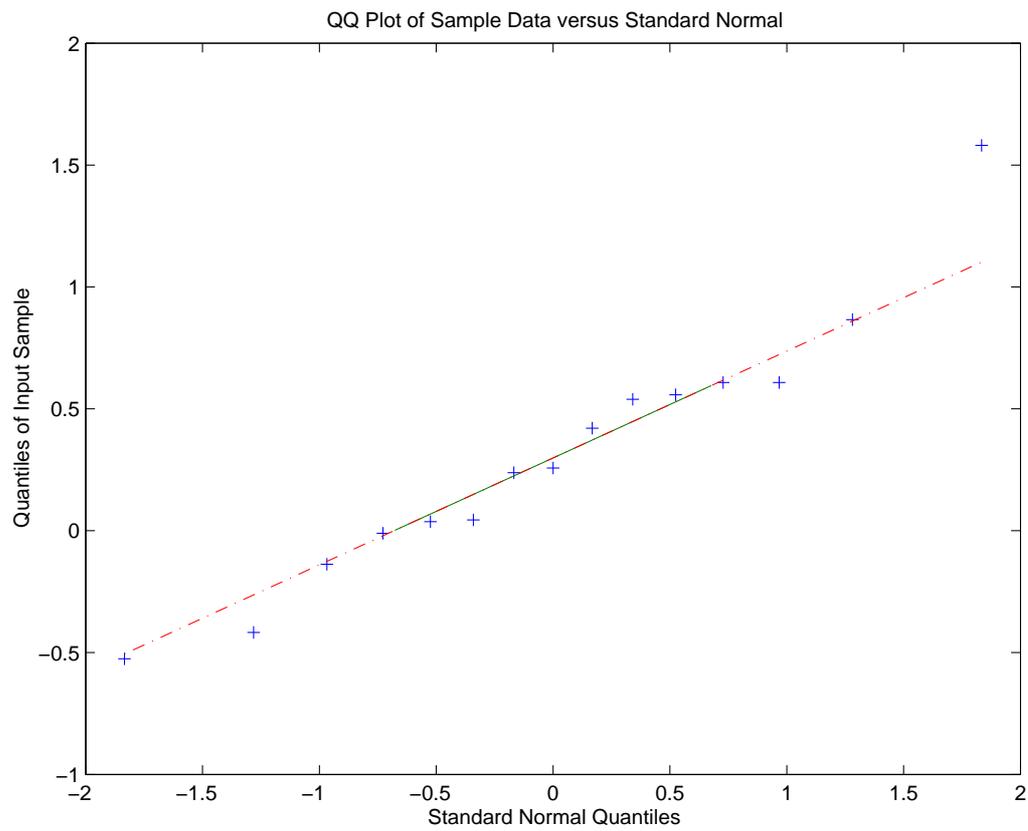


Figura 5.8: Gráfico quantil-quantil dos quantis dos resíduos pelos quantis teóricos (Dados DJZ).