

MINISTÉRIO DA ECONOMIA, FAZENDA E PLANEJAMENTO



FUNDAÇÃO INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA – IBGE

# TEXTOS DI DÁTII COS

ENCE



ESCOLA  
NACIONAL  
DE CIÊNCIAS  
ESTATÍSTICAS

AJUSTE SAZONAL DE SÉRIES DE TEMPO

O MÉTODO X-11

Victor Gouvêa

IBGE/ENCE

IBGE - NÚMERO DE PUBLICAÇÃO  
TÍTULOS DE PUBLICAÇÃO

11 de 11  
- exemplar  
M set  
(9)

AJUSTE SAZONAL DE SÉRIES DE TEMPO

O MÉTODO X-11

Victor Gouvêa

IBGE/ENCE

## ANÁLISE DE SÉRIES DE TEMPO

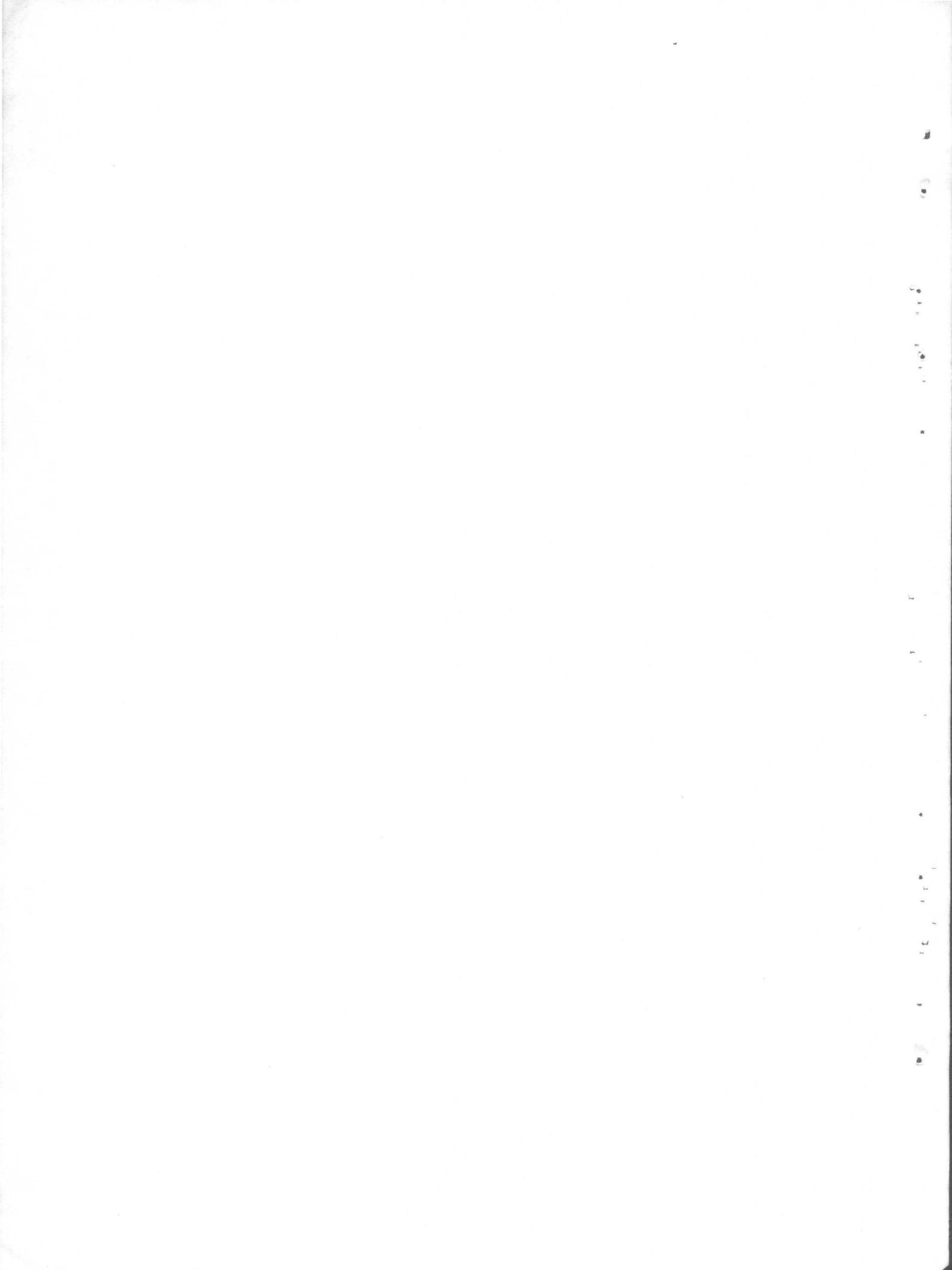
Em ciência e tecnologia, vários problemas ocorrem em que podemos idealizar modelos de análise onde uma variável  $X$  é descrita por suas observações ao longo do tempo. Vários métodos matemáticos vem sendo desenvolvidos para o estudo das séries de tempo. As flutuações existentes nestas séries podem ser devidas à varios fatores e os esforços no sentido de compreensão das séries caminham na direção da decomposição da série em algumas componentes para maior entendimento da importância de cada uma delas na série original.

1. VARIACIONES DE CALENDARIO: Lidando com séries de tempo é importante lembrarmo-nos que os dados fornecidos em base mensal raramente são comparáveis de mês a mês. Vejam-se como exemplo, os dados da tabela 1 que registra a quantidade produzida mensalmente de alcatrão de hulha no Brasil no ano de 1982.

TABELA 1  
PRODUÇÃO DE ALCATRÃO DE HULHA  
BRASIL - 1982

MESES	PRODUÇÃO (ton)
Janeiro.....	15406
Fevereiro.....	12422
Março.....	14043
Abril.....	14074
Maió.....	14342
Junho.....	16495
Julho.....	16400
Agosto.....	15680
Setembro.....	13764
Outubro.....	14599
Novembro.....	14203
Dezembro.....	13708

Deve-se observar que a quantidade produzida em fevereiro está bem abaixo dos valores de janeiro e de março; fevereiro por isso corresponde a cerca de 90% da extensão do mês de janeiro.



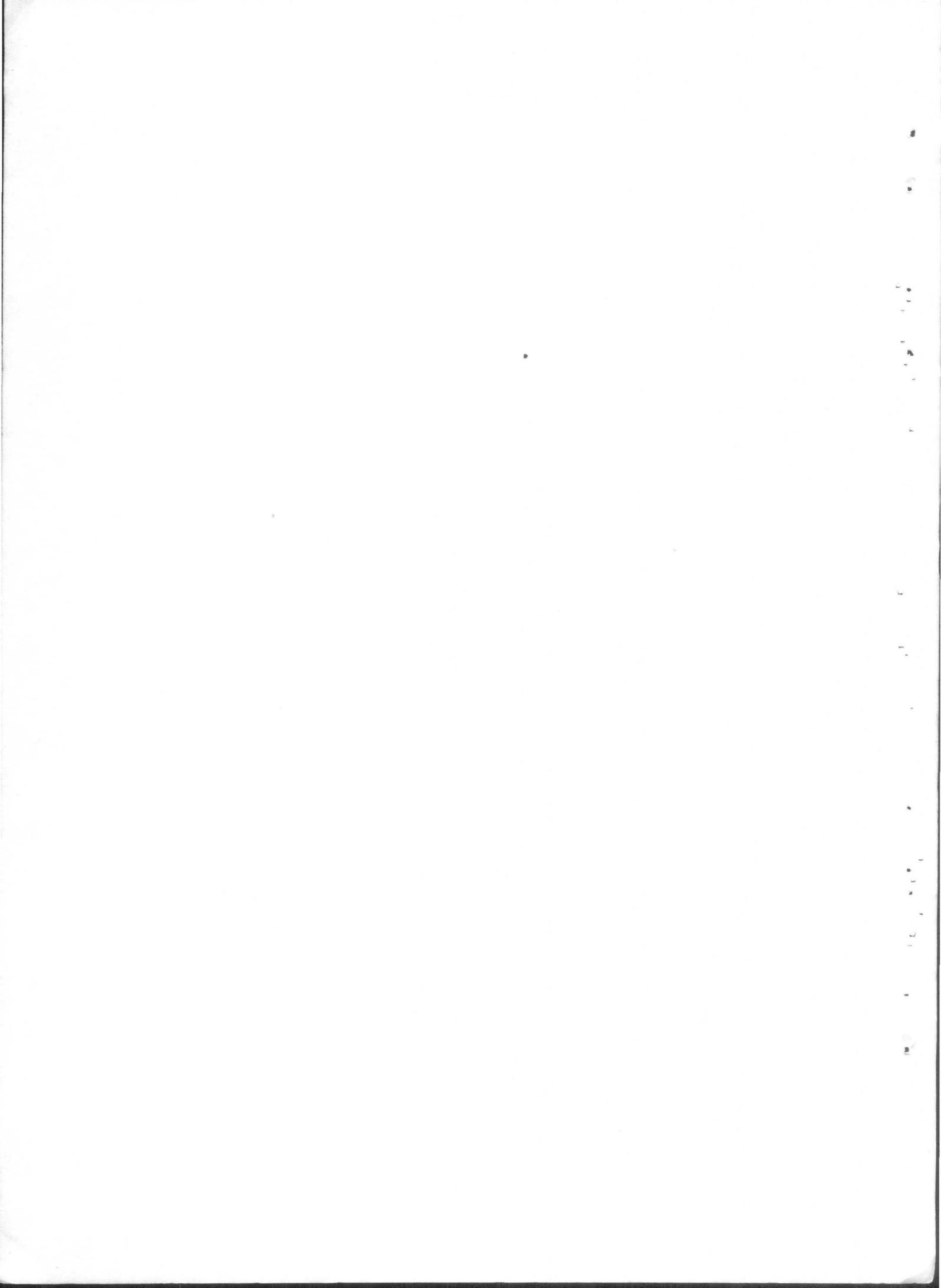
Para tornar os dados da tabela comparáveis seria útil apresentar uma outra tabela com a produção por dia, dividindo a produção de cada mês pelo número de dias do mes. O resultado está apresentado na tabela 2.

TABELA 2  
PRODUÇÃO DE ALCATRÃO DE HULHA DO  
BRASIL SEGUNDO OS MESES - 1982

MESES	PRODUÇÃO POR DIA
Janeiro.....	497
Fevereiro.....	443
Março.....	453
Abril.....	490
Maió.....	462
Junho.....	549
Julho.....	529
Agosto.....	506
Setembro.....	459
Outubro.....	471
Novembro.....	473
Dezembro.....	442

Em problemas econômicos a situação é muitas vezes mais complicada do que a que acabamos de expor. Torna-se necessário ajustar não somente as diferenças em números de dias de cada mes, mas também no número de dias de trabalho nos diversos meses e no mesmo mês em diversos anos. Num ano janeiro pode ter cinco domingos e no ano seguinte quatro. O carnaval pode cair em fevereiro ou março, etc. O modo correto de se fazerem comparações mensais seria o de se verificar o número de dias trabalhados em cada um dos meses e ponderar a produção de maneira a tornar os números comparáveis.

O número de dias de trabalho na produção varia, no entanto, de produto para produto. Algumas fabricas fecham no fim de semana, outras, trabalham 24 horas por dia durante todo o ano, como é o caso das textéis modernas, onde o simples piscar do fornecimento de energia elétrica cortando o fio dos teares pode trazer um grande prejuizo. Assim, os métodos utilizados para correção de um grande número de produtos fazem a correção de dias



de trabalho por meio de regressão múltipla:

$$Y = a_1X_1 + a_2X_2 + a_3X_3 + a_4X_4 + a_5X_5 + a_6X_6 + a_7X_7$$

onde Y é a variável a ser ajustada, X<sub>1</sub> é o número de domingos de cada mes, X<sub>2</sub> o número de segundas-feiras e assim por diante até X<sub>7</sub>, número de sabados de cada mes. Se houver relação entre o número de dias trabalhados e a produção do produto, a série será recorrigida e seus números tornados comparáveis.

2. COMPONENTE SAZONAL: Variações de diversos tipos podem ocorrer em dados históricos e suas diversidades podem ser melhor descritas por meio de ilustrações. Vejamos por exemplo, os dados de produção de açúcar cristal. A tabela 3 nos fornece a série de produção mensal brasileira de 1975 a 1982. O grafico 1 é a representação dos dados desta tabela. A série se caracteriza por se repetir em sua forma a intervalos regulares. A repetição é a essência do movimento cíclico. Quando um ciclo tem a duração de 12 meses é chamado de ciclo de estação ou sazonal. O movimento sazonal é portanto um determinado tipo de movimento cíclico - o que tem um período de 12 meses.

Tabela 3  
PRODUÇÃO FÍSICA DE AÇUCAR CRISTAL  
BRASIL - 1975 A 1982

	1975	1976	1977	1978	1979	1980	1981	1982
Jan...	2430	3319	2838	2448	2427	1696	2798	2451
Fev...	2077	2365	2032	1976	1520	976	2204	1571
Mar...	1518	1453	1384	1378	799	857	1034	970
Abr...	725	716	762	898	362	194	494	104
Mai...	207	550	370	183	30	40	96	0
Jun...	4835	5781	8064	7099	7776	8114	12315	10619
Jul...	7440	6453	8934	7721	8185	8048	10715	9841
Ago...	8219	7101	9800	9524	8621	9236	11799	8986
Set...	8509	7319	9158	9824	7093	9950	10346	10356
Out...	5942	8954	9513	9459	7610	10381	6241	7475
Nov...	3468	6566	6352	5139	5270	5238	3356	4713
Dez...	3181	4028	2977	4411	3803	3345	2447	3407

A característica mais notável desta série é o fato de se



verificarem variações periódicas francamente regulares. O movimento sazonal domina o gráfico. Os movimentos não são uniformes de ano para ano, de forma exata. Mas as diferenças são menos impressionantes que as semelhanças.

Se a série não tivesse outros movimentos diferentes do sazonal, seria possível descrever este movimento apenas obtendo a média para cada mes. O resultado para o nosso exemplo de açúcar cristal esta dado na tabela 4.

Tabela 4  
PRODUÇÃO MEDIA DE AÇUCAR CRISTAL  
BRASIL - 1975 A 1982

MESES	QUANTIDADE MEDIA
Janeiro.....	2551
Fevereiro.....	1840
Março.....	1174
Abril.....	532
Maió.....	182
Junho.....	8075
Julho.....	8417
Agosto.....	9161
Setembro.....	9069
Outubro.....	8196
Novembro.....	5019
Dezembro.....	3450

Na tabela, ficam bem marcadas as épocas de quantidades produzidas altas e baixas. Se nós desejássemos, poderíamos converter estas médias em índices percentuais, isto é, em indicadores da variação sazonal. De que modo? Calculando a média das 12 médias e convertendo cada uma das 12 médias em percentuais da média total. Ora, a média total é 4805 e a transformação, digamos, do mês de março seria feita do seguinte modo:

$$I_{\text{março}} = \frac{1174}{4805} \times 100 = 24.4\%$$

o que significa que o valor de março ficou 75,6% (100 - 24,4) abaixo da média anual. Este método de calculo não daria bons resultados se estivessem incluídos nos dados uma tendência de longo prazo ou um ciclo diferente do de 12 meses. Por isso, os estatísticos utilizam um procedimento diferente e um pouco mais complexo que consiste em fazer primeiro uma média móvel de 12



meses e eliminar esta média móvel. Desde que a média móvel seja de 12 meses ela não conterá nada do movimento sazonal. As variações mensais terão sido eliminadas, uma vez que em cada elemento da média móvel estarão somados os 12 meses. A tendência secular e os outros ciclos diferentes do sazonal continuarão atuando sobre a média móvel. Mas, junto com a componente sazonal, também a componente aleatória será eliminada. Os elementos da média móvel de 12 meses se encontram na tabela 6, para o exemplo do açúcar.

TABELA 5  
INDICES DE VARIAÇÃO SAZONAL DA  
PRODUÇÃO DE AÇUCAR CRISTAL

MESES	INDICE DO MOVIMENTO SAZONAL
Janeiro.....	53.1
Fevereiro.....	38.3
Março.....	24.4
Abril.....	11.1
Maió.....	3.8
Junho.....	168.1
Julho.....	175.2
Agosto.....	190.6
Setembro.....	188.7
Outubro.....	170.6
Novembro.....	104.3
Dezembro.....	71.8

TABELA 6  
MEDIAS MOVEIS

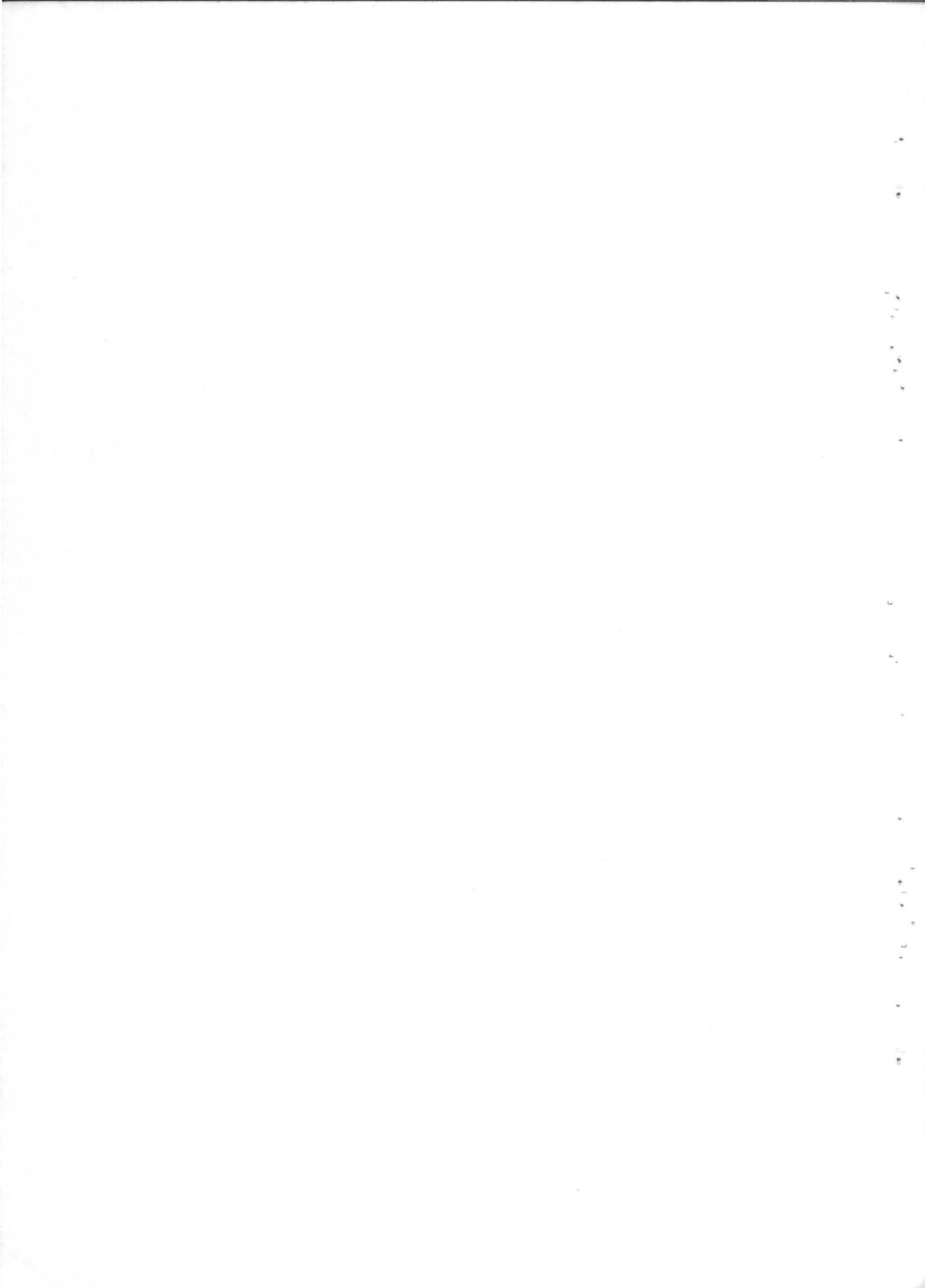
	1975	1976	1977	1978	1979	1980	1981	1982
Jan...	—	4204	4759	5007	4935	4366	5539	5015
Fev...	—	4116	4975	4945	4917	4386	5757	4862
Mar...	—	4020	5164	4962	4766	4530	5881	4745
Abr...	—	4096	5264	4987	4575	4765	5725	4797
Mai...	—	4351	5279	4934	4503	4879	5474	4905
Jun...	—	4515	5226	4944	4483	4859	5358	5001
Jul...	4083	4530	5166	5002	4428	4886	5306	—
Ago...	4132	4496	5147	4983	4374	4983	5265	—
Set...	4141	4480	5145	4939	4354	5041	5236	—
Out...	4138	4479	5150	4893	4350	5061	5217	—
Nov...	4152	4473	5147	4865	4343	5076	5197	—
Dez...	4206	4561	5098	4888	4358	5253	5122	—



Resumindo: tabulados os dados a estudar, calcula-se a seguir a média móvel de 12 meses. Isto fara com que primeira média obtida fique no centro do período, entre junho e julho. Poderíamos simplesmente situar esta média no mês de julho. Ou, alternativamente, podemos ajustar a média de modo que fique centrada exatamente em julho. De que modo? Fazendo a média com a metade de janeiro do primeiro ano e a metade de janeiro do segundo ano. A média móvel da série seria calculada para cada mês como uma média de 12 meses com o primeiro e o último mês contados pela metade. Em geral este ajuste não traz uma diferença significativa, a não ser que haja uma forte reverção de um mês para outro, o que é exatamente o caso da nossa série de açúcar, onde existe uma reverção abrupta de maio para junho. Fizemos o nosso cálculo por este processo mais preciso.

Na média móvel estão a componente de tendência de longo prazo (tendência secular) e ciclos diferentes do sazonal e o que desejamos saber é o que sobra, se estivermos interessados na componente sazonal. Existem dois procedimentos diferentes para obtenção desta "sobra". Uma é a de se fazer a diferença entre a série original e a média móvel. (a soma das componentes é igual a série original). Trata-se do método aditivo. Outro procedimento é o de se dividir a série original pela média móvel (o produto das componentes é igual à série original). É o chamado método multiplicativo. A experiência vem demonstrando que para séries econômicas o método multiplicativo se ajusta melhor, exceto nos casos de séries agrícolas ou ligadas à agricultura (caso do açúcar) onde a componente irregular ou aleatória é muito alta. O resultado do cálculo pelos dois métodos estão nas tabelas 7 e 8 sendo que nesta última os resultados se apresentam em termos percentuais.

Note-se que uma vez eliminada a componente secular cíclica contida na média móvel, nos resta ainda uma série onde se somam (método aditivo) ou se multiplicam (multiplicativo) as



componentes sazonal e irregular ou aleatória. Com efeito, a média móvel ao somar os 12 meses, elimina o efeito sazonal, mas, ao mesmo tempo, aplaina as irregularidades da série. Precisamos então separar nos dados das tabelas 7 e 8 a componente sazonal da componente irregular. O processo que seguimos até aqui é o de considerarmos a média do mês nos oito anos como o padrão de sazonalidade do mes. Calculemos estes padrões de sazonalidade. O resultado se encontra nas tabelas 9 e 10.

TABELA 7  
(SERIE ORIGINAL - MEDIAS MOVEIS)

	1975	1976	1977	1978	1979	1980	1981	1982
Jan...	-	-885	-1921	-2559	-2508	-2670	-2741	-2564
Fev...	-	-1751	-2943	-2969	-3397	-3410	-3553	-3291
Mar...	-	-2567	-3780	-3584	-3967	-3673	-4847	-3775
Abr...	-	-3380	-4502	-4089	-4213	-4571	-5231	-4693
Mai...	-	-3801	-4909	-4771	-4473	-4839	-5378	-4905
Jun...	-	1266	2838	2155	3293	3255	6957	5618
Jul...	3357	1923	3768	2719	3757	3163	5409	-
Ago...	4087	2605	4653	4541	4247	4253	6534	-
Set...	4368	2839	4013	4885	2739	4909	5110	-
Out...	1804	4475	4363	4566	3260	5320	1024	-
Nov...	-684	2093	1205	274	927	162	-1841	-
Dez...	-1025	-533	-2121	-477	-555	-1908	-2675	-

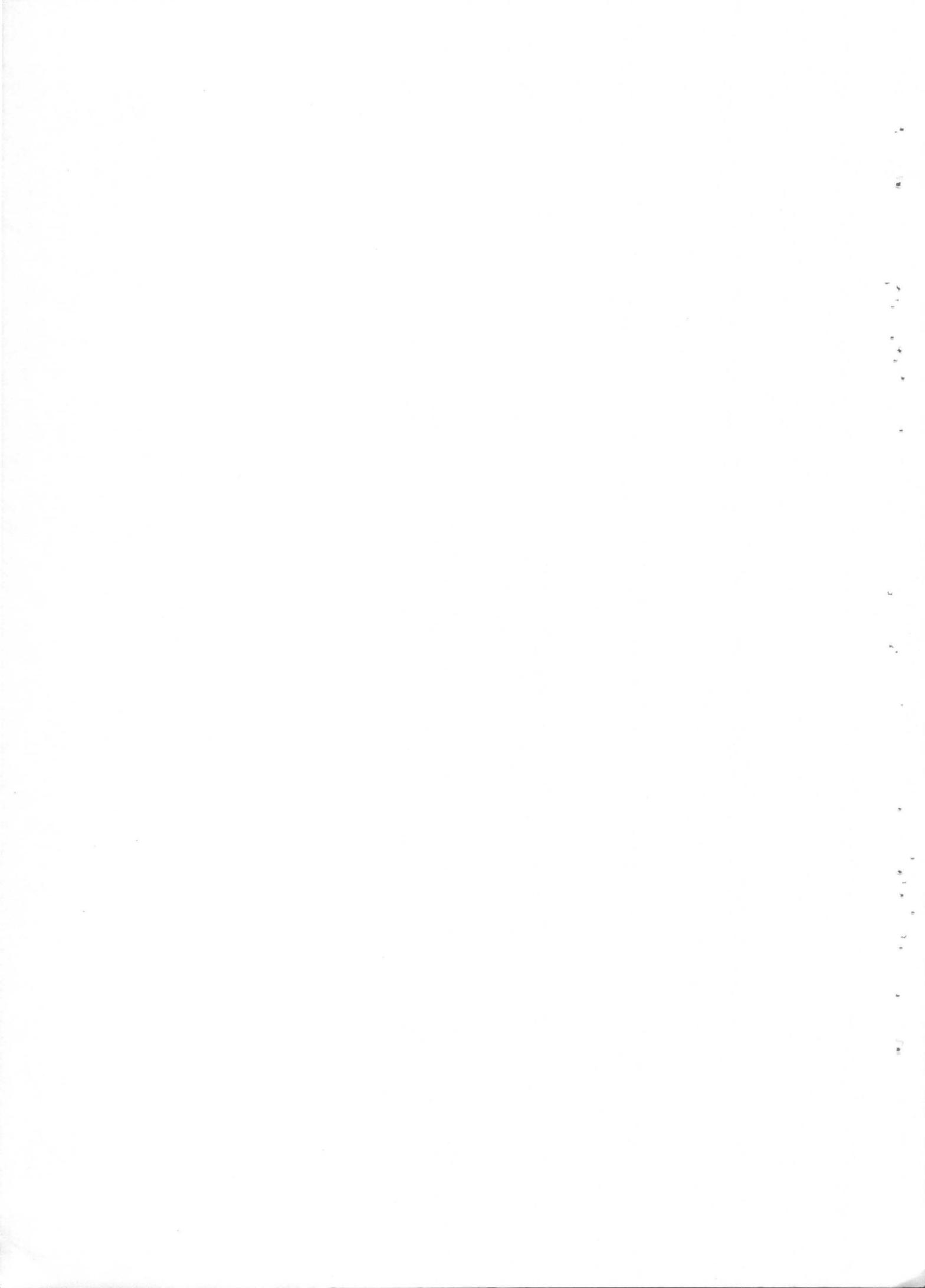


TABELA 8  
(SERIE ORIGINAL / MEDIAS MOVEIS) x100

	1975	1976	1977	1978	1979	1980	1981	1982
Jan...	-	79	60	49	49	39	51	49
Fev...	-	57	41	40	31	22	38	32
Mar...	-	36	27	28	17	19	18	20
Abr...	-	17	14	18	8	4	9	2
Mai...	-	13	7	3	1	1	2	0
Jun...	-	128	154	144	173	167	230	212
Jul...	182	142	173	154	185	165	202	-
Ago...	199	158	190	191	197	185	224	-
Set...	205	163	178	199	163	197	198	-
Out...	144	200	185	193	175	205	120	-
Nov...	84	147	123	106	121	103	65	-
Dez...	76	88	58	90	87	64	48	-

TABELA 9  
COMPONENTE SAZONAL PELO METODO  
ADITIVO. DADA EM VALOR ABSOLUTO

MESES	QUANTIDADE
Janeiro.....	-2264
Fevereiro.....	-3045
Março.....	-3742
Abril.....	-4383
Mai.....	-4725
Junho.....	3626
Julho.....	3442
Agosto.....	4417
Setembro.....	4123
Outubro.....	3545
Novembro.....	305
Dezembro.....	-1328

TABELA 10  
COMPONENTE SAZONAL PELO METODO  
MULTIPLICATIVO. EM PERCENTUAIS

MESES	PERCENTUAL
Janeiro.....	53
Fevereiro.....	37
Março.....	23
Abril.....	10
Mai.....	4
Junho.....	173
Julho.....	172
Agosto.....	192
Setembro.....	186
Outubro.....	174
Novembro.....	107
Dezembro.....	73

Salta aos olhos na comparação das duas tabelas que para os meses onde a componente sazonal é negativa no método aditivo, o percentual é inferior à 100 no método multiplicativo. Este resultado era de se esperar uma vez que ambos foram criados para medirem a mesma coisa. Vejamos para cada método como se faz para a obtenção da série sem a componente sazonal a que chamaremos de "série sazonalmente ajustada".



Pelo método aditivo basta subtrair de cada valor da série original a componente sazonal do mês correspondente, dada na tabela 9.

Pelo método multiplicativo, dividiremos o valor de cada mês na série original pelos percentuais correspondentes ao mês, dados pela tabela 10 (sem esquecermos de multiplicar por 100).

Deste modo, obteremos as séries sazonalmente ajustadas por cada um dos métodos. O resultado é apresentado nas tabelas 11 e 12.

Analisando o resultado do método multiplicativo verificaremos que em alguns pontos o resultado do ajuste não é de muito boa qualidade. Por exemplo a quantidade correspondente ao mês de maio de 1976 é de 14687, superior a qualquer elemento da série original. Isto ocorre porque na série original o mês de maio tem dados que vão desde 550 (maio de 1976) até zero (maio de 1982), que fomos obrigados a substituir por um para podermos aplicar o procedimento multiplicativo. Enquanto para o método aditivo a diferença em termos absolutos entre 550 e 1 é de 549, relativamente pequena dados os outros valores da série, o método multiplicativo "vê" o valor de maio de 1976 como 550 vezes maior do que o de maio de 82. Daí o método aditivo apresentar melhores resultados quando a irregularidade da série é muito alta.

Façamos agora algumas considerações sobre o método aditivo, para explicar melhor o seu funcionamento: tomemos na série original um elemento qualquer, por exemplo, o correspondente a setembro de 1980. Temos 9950. Pela tabela 6 a componente secular-cíclica deste mês e ano é igual a 5041. Onde, restam para as componentes sazonal e irregular  $9950 - 5041 = 4908$ . (dado na tabela 7). Como a componente sazonal correspondente é (tabela 10) igual a 4122 restam para a componente irregular  $4908 - 4122 = 786$ . Refazendo as contas:



5041 (componente de tendência de longo prazo)

4122 (componente sazonal)

787 (componente irregular)

9950 (valor da série original)

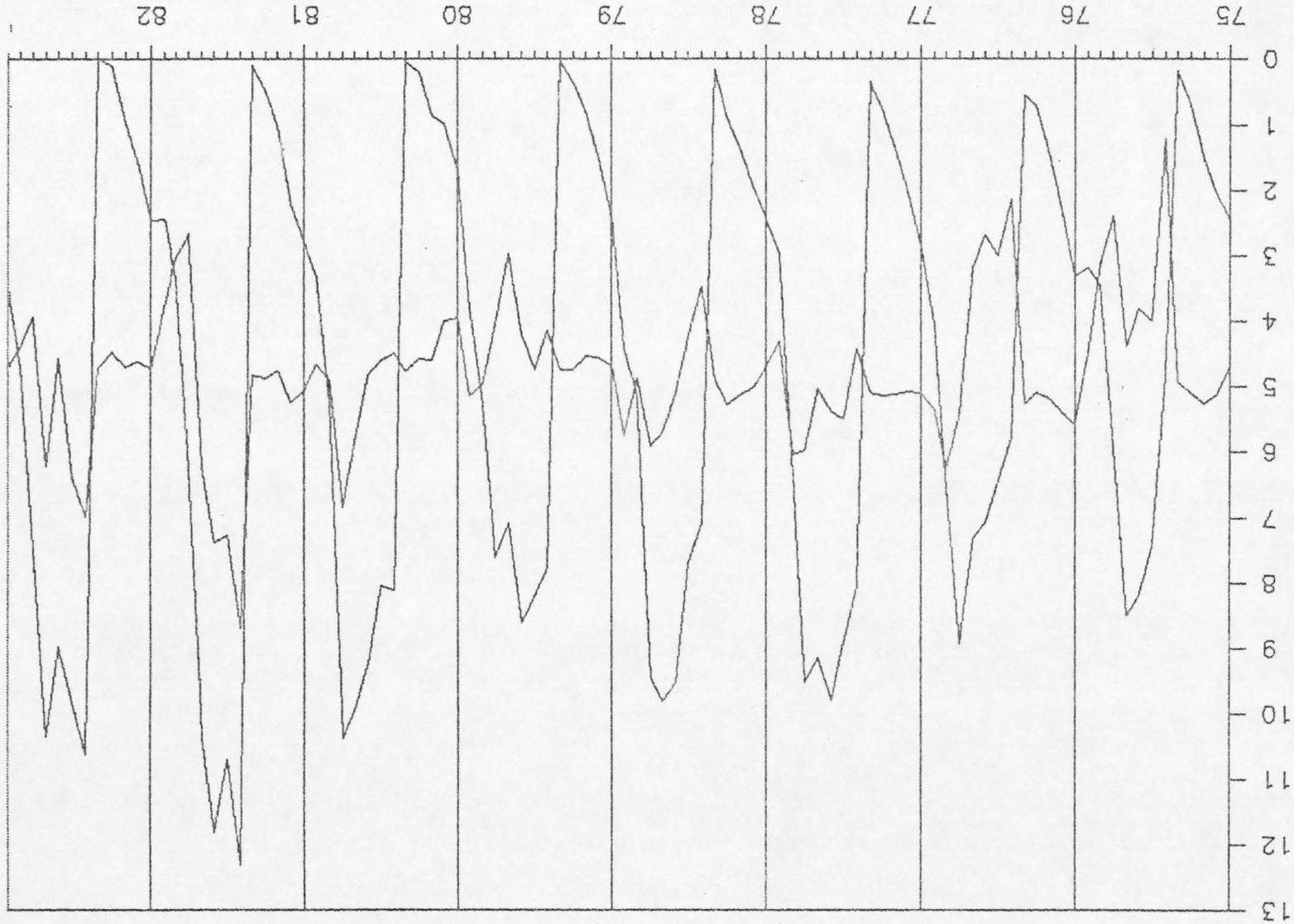
TABELA 11  
SERIE SAZONALMENTE AJUSTADA PELO METODO ADITIVO

	1975	1976	1977	1978	1979	1980	1981	1982
Jan...	4694	5583	5102	4712	4691	3960	5062	4715
Fev...	5122	5410	5077	5021	4565	4012	5249	4616
Mar...	5260	5195	5126	5120	4541	4599	4776	4712
Abr...	5108	5099	5145	5281	4745	4577	4877	4487
Mai...	4932	5275	5095	4888	4755	4765	4821	4726
Jun...	1209	2155	4438	3473	4150	4488	8699	6993
Jul...	3998	3011	5492	4279	4743	4606	7273	6399
Ago...	3802	2684	5385	5107	4204	4819	7382	4569
Set...	4386	3196	5035	5701	2970	5827	6223	6233
Out...	2398	5410	5969	5915	4066	6836	2697	3931
Nov...	3155	6253	6039	4880	4957	4925	3043	4400
Dez...	4509	5356	4305	5739	5131	4673	3775	4735



# PRODUCAO DE ACUCAR CRISTAL

Brasil - 1975/82



Mil toneladas

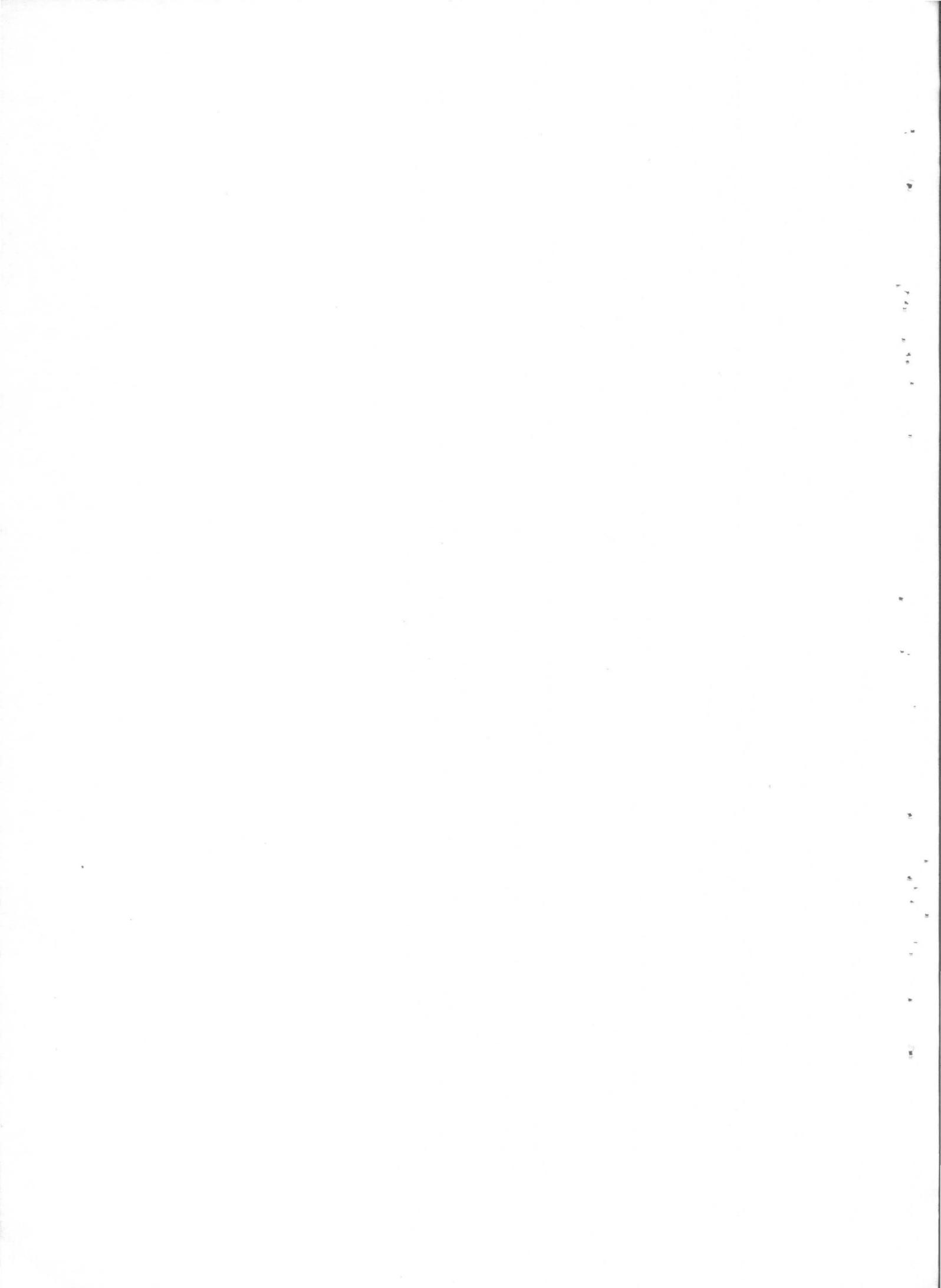


TABELA 12  
SERIE SAZONALMENTE AJUSTADA PELO METODO MULTIPLICATIVO

	1975	1976	1977	1978	1979	1980	1981	1982
Jan...	4538	6198	5300	4571	4532	3167	5225	4577
Fev...	5549	6319	5429	5279	4061	2607	5888	4197
Mar...	6463	6186	5892	5867	3401	3648	4402	4129
Abr...	6977	6890	7333	8642	3483	1867	4754	1000
Mai...	5527	14687	9880	4352	801	1068	2563	26
Jun...	2800	3348	4671	4112	4504	4700	7133	6151
Jul...	4327	3753	5194	4491	4761	4681	6233	5724
Ago...	4278	3696	5101	4957	4487	4807	6141	4677
Set...	4569	3930	4918	5275	3809	5343	5556	5561
Out...	3406	5133	5453	5422	4362	4951	3577	4285
Nov...	3239	6132	5932	4850	4922	4892	3142	4401
Dez...	4355	5515	4076	6039	5207	4580	3350	4665

O gráfico da série original junto com o da série sazonalmente ajustada pelo método aditivo se encontra a seguir e pode nos dar uma visão global do comportamento das duas séries.

A utilização do mesmo padrão de sazonalidade durante oito anos seguidos nem sempre é conveniente. Vejamos o caso da nossa série exemplo. O pico da produção de açúcar cristal se dava, até o ano de 1980 no mês de agosto. Nos últimos anos da série o ponto mais alto correspondente ao mês de junho. O que acontece quando se usa um padrão de sazonalidade fixo? A mudança de padrão de sazonalidade é tratada como se fosse a componente irregular.

Outro problema que pode existir com o uso do mesmo padrão de sazonalidade é o que ocorre quando há uma tendência de crescimento ou de queda ao longo do tempo. Neste caso a componente sazonal é, muitas vezes, proporcional ao valor da série original. Isto é, nos anos de baixa, a componente sazonal é menor; nos anos de alta, a componente é maior. Deste ponto de vista, o método multiplicativo é superior ao aditivo uma vez que pelo método multiplicativo a componente sazonal é calculada como uma percentagem da série original. O uso de um padrão fixo de



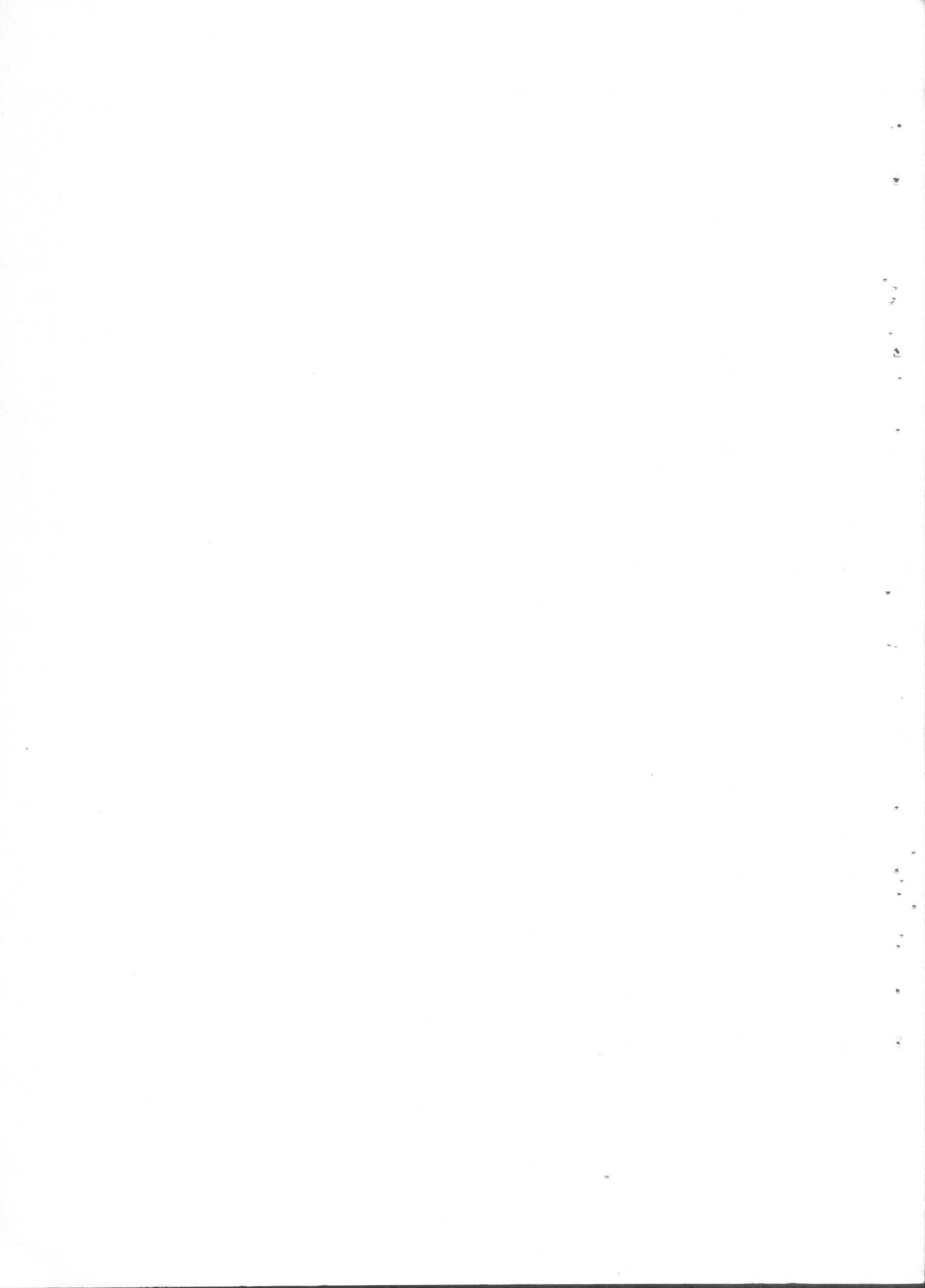
sazonalidade no método aditivo pode levar à distorções como o do mês de junho de 1975 na nossa série sazonalmente ajustada. Observe o gráfico: neste mês aparece uma queda brusca. Porque? Porque o padrão de sazonalidade é em parte determinado pelo fato de que os grandes picos nos anos finais da série são em junho. Como foi retirada a mesma quantidade para todos os meses de junho, a série sazonalmente ajustada vem com os meses de junho dos primeiros anos com esta distorção.

O que fazer? Usar um padrão móvel de sazonalidade. Vale dizer, após termos feito a média móvel de 12 meses e retirado da série a componente secular-cíclica, ficando com um composto das componentes sazonal e irregular, ao invés de fazermos uma média geral dos oito anos, fazermos uma média móvel, digamos, de três anos. Repetindo: vamos fazer, para cada mes, com os dados da tabela 7 uma média móvel de três em três anos. Assim, ao invés de termos o padrão fixo de sazonalidade dado na tabela 9 teremos uma tabela de padrões de sazonalidade variando com o tempo. Por exemplo: para janeiro de 1977 vou usar (veja a tabela 8)

$$\frac{-885 - 1921 - 2559}{3} = -1788$$

como padrão. Os padrões de sazonalidade são obtidos fazendo a média do ano igual a 100. (multiplicativo) ou zero (aditivo). Este será o valor a ser diminuído da série original para obtenção da série ajustada no mês de janeiro de 1977.

Todos os padrões de sazonalidade podem ser obtidos por este processo, do ano de 1977 até 1980. Resta-nos o problema do que fazer para obtenção dos padrões de sazonalidade para os anos de 1975, 1976, 1981 e 1982. Vamos por partes. Tratemus inicialmente o caso do segundo (1976) e do penúltimo (1981) ano. O que ocorre para estes anos é que como não temos a média móvel dos 6 primeiros meses de 1975 e dos últimos 6 meses de 1982 não posso obter o padrão de sazonalidade destes meses nos anos de 1976 e 1981 porque não tenho como obter a média dos três anos (fica faltando o 1º elemento da média móvel para 1976 e o último para 1981). O procedimento comumente adotado é o de se repetir os



padrões de sazonalidade de 1977 para os meses de 1976 e que não existem em 1976 e do mesmo modo, repetir os padrões de 1980 nos últimos meses de 1981. Com isso, teremos os padrões de sazonalidade de 1976 e 1981, bastando fazer a média dos padrões destes anos iguais a 100 (multiplicativo) ou zero (aditivo).

Resta o problema da obtenção de padrões para 1975 e 1982. O problema é tanto mais grave pelo fato de normalmente estarmos interessados em separar a componente sazonal justamente no último ano para podermos fazer análises de conjuntura. As soluções apontadas são:

Uso de pesos diferentes: Numa média móvel de três anos, cada elemento é contado com peso  $1/3$ . Assim, para o ano de 1981 o padrão de sazonalidade é obtido somando-se os valores dos anos de 1980, 1981 e 1982 (e dividindo por 3). No caso de 1982, onde não temos o ano de 1983, poderíamos fazer a média móvel, colocando-se o próprio ano de 1982 no lugar de 1983. Isto equivale a obter a média somando-se o ano de 1981 com duas vezes o ano de 1982, o que equivale a tomar o ano de 1981 com peso  $1/3$  e o ano de 1982 com peso  $2/3$ . No caso de uma média móvel de período mais longo, daremos pesos mais fortes para pontos próximos ao ano que estamos tentando determinar e mais fracos para os anos mais afastados. O padrões de sazonalidade incluindo 1975 e obtidos com peso  $1/3$  e  $2/3$  como explicado acima, são dados na tabela 13. A série sazonalmente ajustada por estes padrões é dada na tabela 14. O gráfico 2 nos fornece a série original, com a série ajustada sazonalmente.



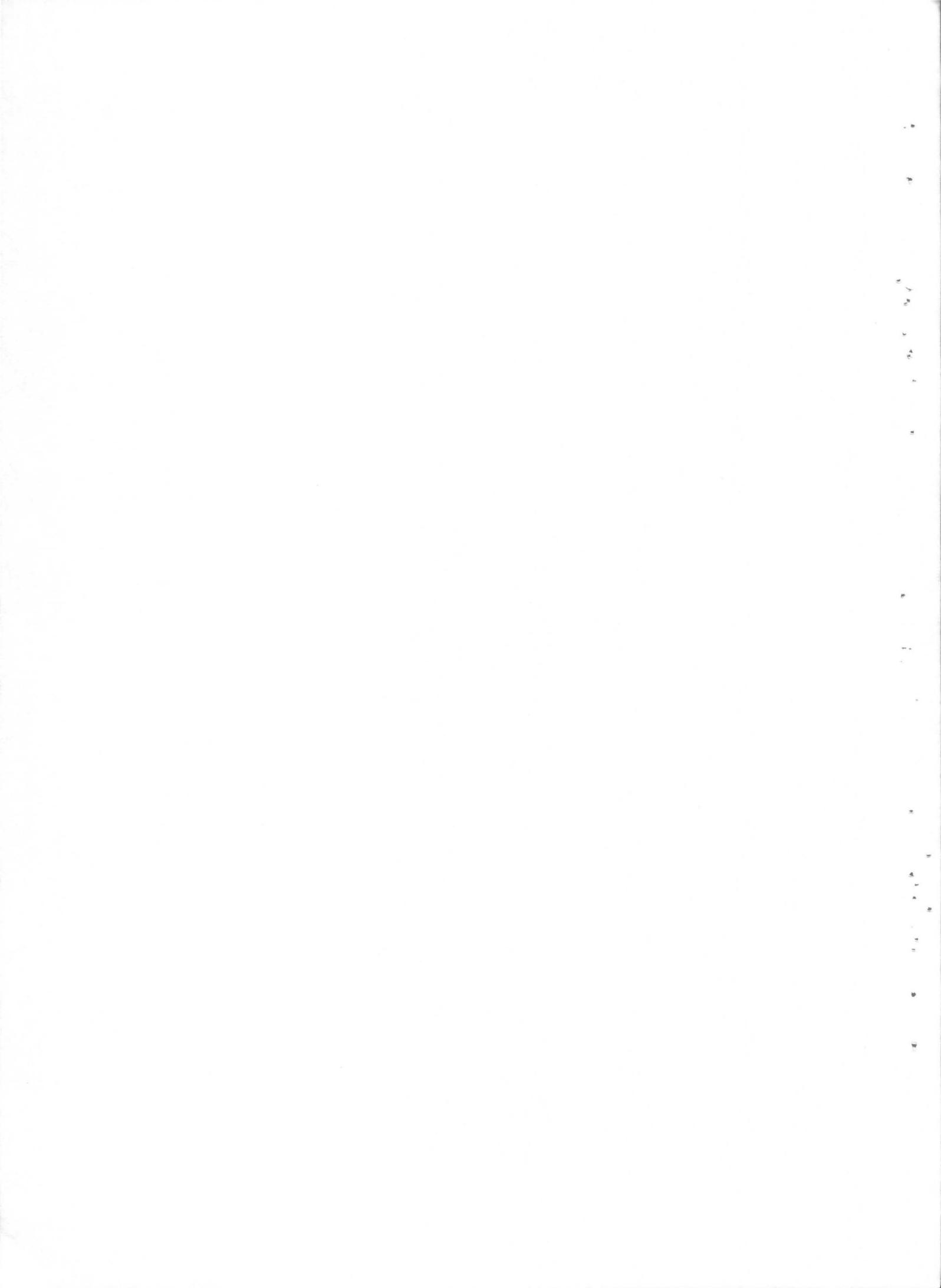
TABELA 13  
PADRÕES MOVEIS DE SAZONALIDADE - METODO ADITIVO

	1975	1976	1977	1978	1979	1980	1981	1982
Jan...	-973	-1840	-1890	-2327	-2586	-2589	-2643	-2579
Fev...	-1839	-2606	-2656	-3091	-3265	-3402	-3402	-3334
Mar...	-2655	-3462	-3412	-3765	-3748	-4112	-4082	-4807
Abr...	-3668	-4040	-4090	-4254	-4296	-4621	-4815	-4827
Mai...	-3883	-4545	-4595	-4705	-4701	-4845	-5024	-5017
Jun...	1177	2032	1982	2773	2893	4573	5299	6108
Jul...	2918	3155	2699	3426	3205	4181	4204	4871
Ago...	3632	4071	3823	4491	4336	5082	5105	4708
Set...	3898	3789	3808	3870	4169	4323	4346	5160
Out...	1939	3597	4364	4074	4374	3273	3296	1901
Nov...	282	920	1104	831	464	-199	-829	-891
Dez...	-820	-1175	-1146	-1039	-981	-1661	-1646	-1991

TABELA 14  
SERIE SAZONALMENTE AJUSTADA PELO METODO ADITIVO

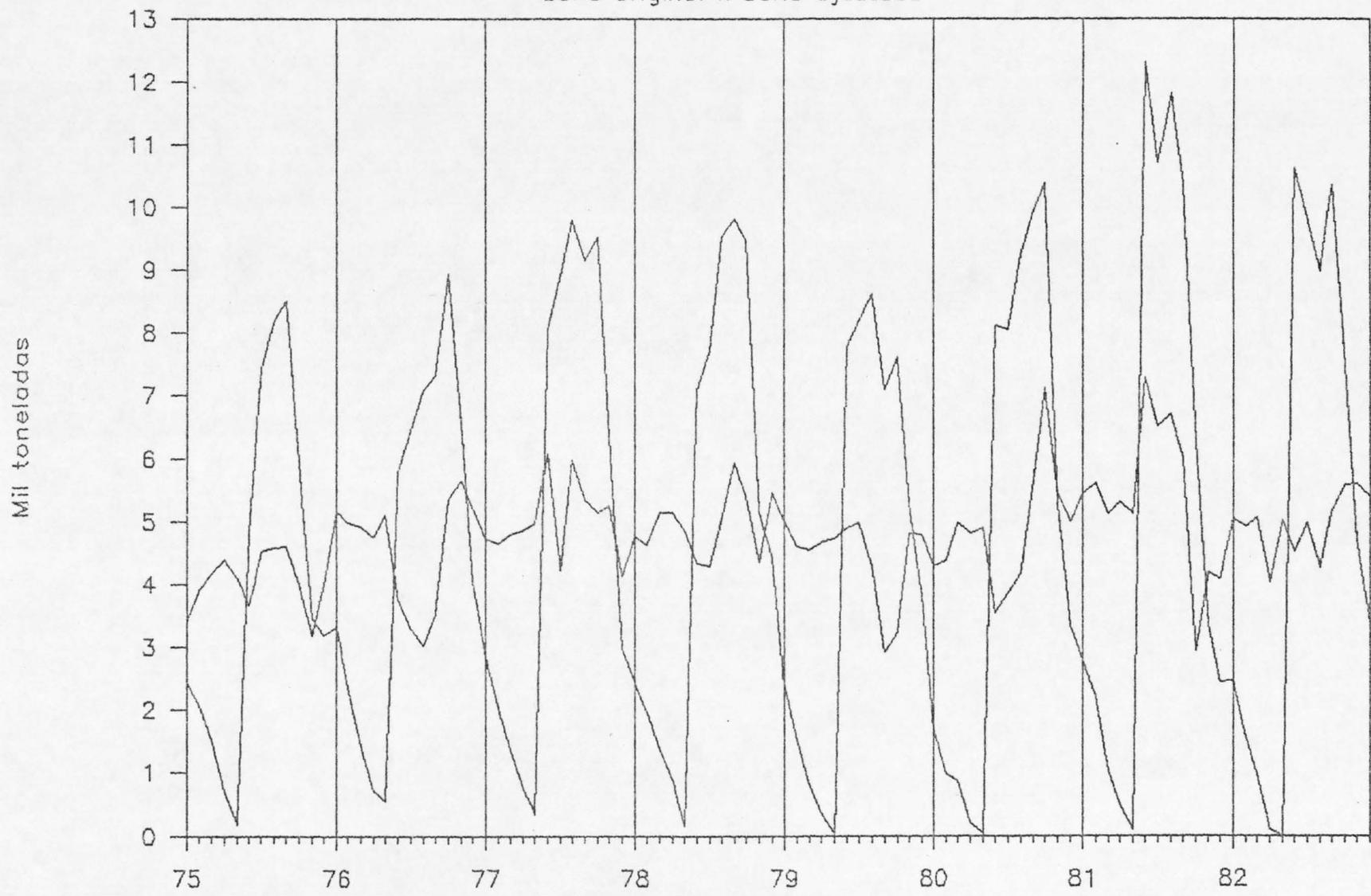
	1975	1976	1977	1978	1979	1980	1981	1982
Jan...	3403	5159	4728	4775	5013	4285	5441	5030
Fev...	3916	4971	4638	4627	4611	4378	560	4905
Mar...	4173	4915	4795	5143	4547	4969	511	5057
Abr...	4393	4756	4852	5152	4658	4815	530	4031
Mai...	4095	5095	4965	4868	4731	4885	512	5017
Jun...	3656	3749	6082	4326	4883	3541	729	4511
Jul...	4522	3298	4235	4295	4980	3867	651	4970
Ago...	4587	3030	5977	5033	4285	4154	669	4278
Set...	4611	3530	5350	5934	2924	5627	600	5196
Out...	4003	5357	5149	5385	3236	7108	294	5574
Nov...	3186	5646	5248	4362	4806	5437	418	5604
Dez...	4001	5199	4123	5450	4784	5006	409	5398

Análise auto-regressiva. O problema da obtenção da média móvel para 1982 estaria resolvido se dispuséssemos dos dados de 1983. Mas, uma vez que não os temos, porque não estimarmos o ano de 1983 por algum método de previsão estatística? Uma das



# PRODUCAO DE ACUCAR CRISTAL

serie original x serie ajustada





maneiras de se fazer projeções em séries de tempo é o de se raciocinar do seguinte modo: se todo ano existe uma componente que se repete aproximadamente da mesma maneira (sazonal) e se eu fizer uma regressão entre os dados de um ano e os dados do ano imediatamente anterior esta regressão terá necessariamente um coeficiente de correlação alto. Com esta equação de regressão eu posso então prever o ano seguinte. Se fizermos isto poderemos obter o ano de 1983 com base nos dados de 1982 e de uma equação de regressão. Como é achada exatamente esta equação? Toma-se como variável dependente a dos dados da série e como variável explicativa a dos dados da mesma série apenas com defasagem de 12 meses. Esta técnica é chamada de autoregressiva porque é feita da série histórica com ela própria. A equação de regressão é:

$$V_t = V_{t-12} + U_t$$

Traduzindo: o valor da série no tempo  $t$  é função do valor de doze meses antes mais uma componente aleatória.

Uma vez obtidos os padrões de sazonalidade para os anos centrais da série, obtemos o padrão dos anos da extremidade (1975 e 1982 no exemplo). Porque aplicar a equação na série de padrões de sazonalidade e não na série original?

Porque, na série original é de se esperar uma relação muito menos forte entre  $V_t$  e  $V_{t-12}$  do que na série de padrões. O problema aqui é que se confia demasiadamente no padrão do penúltimo ano.

Por este método ainda fica por definir que período usar para a regressão. A série toda? E' confiar demais em dados já distantes no passado. O último ano contra o anterior? Talvez hajam especificidades de curto prazo que influam decisivamente e que não seriam benéficas na obtenção de padrões de sazonalidade. O resultado da obtenção dos padrões de sazonalidade por este método, utilizando-se uma regressão que leva em conta os três últimos anos para estimar 1982 e os três primeiros para estimar 1975 é dado na tabela 15. A série ajustada é dada na tabela 16. Como os outros



anos tem os mesmos valores nos dois métodos, apresentaremos apenas os anos de 1975 e 1982.

Tabela 15  
Padrões de Sazonalidade por  
autoregressão

MÊS	1975	1982
Jan...	-1751	-2771
Fev...	-2488	-3571
Mar...	-3341	-4287
Abr...	-3896	-5060
Mai...	-4382	-5280
Jun...	1927	5613
Jul...	3015	4449
Ago...	3895	5397
Set...	3624	4598
Out...	3439	3492
Nov...	868	- 861
Dez...	-1144	-1775

Tabela 16  
Série ajustada pela  
tabela 15

MÊS	1975	1982
Jan...	4181	5222
Fev...	4565	5142
Mar...	4859	5257
Abr...	4621	5164
Mai...	4589	5280
Jun...	2908	5006
Jul...	4425	5392
Ago...	4324	3689
Set...	4885	5758
Out...	2503	3982
Nov...	2600	5574
Dez...	4325	5182

Uma comparação entre os dois métodos apresenta vantagem para o método de pesos (pelo processo autoregressivo as disparidades de certos meses é flagrante (por ex: outubro de 1982). O que não inviabiliza a tentativa nesta direção desde que nós utilizemos métodos autoregressivos mais sofisticados e de maior acurácia.

De todo modo, a análise do gráfico 2 nos mostra que um padrão móvel de três em três é muito instável. No gráfico 1 tínhamos utilizado um padrão fixo para os oito anos, o que é o outro extremo. A instabilidade do padrão de três anos se mostra por exemplo (tabela 13) no mês de julho dos anos de 1975, 1976 e 1977. Um padrão não pode ficar subindo e descendo. Tem que ser relativamente estável. A utilização de um padrão móvel de sazonalidade implica portanto em certas considerações: por um lado, se lançarmos mão de poucos anos para sua obtenção, o padrão fica muito sensível às variações conjunturais ou irregulares. Por outro lado, um número muito grande de anos pode levar ao problema já apontado de que mudanças no padrão não são levados em conta. Isto nos remete à necessidade de decidir o número ótimo de anos a ser utilizado. Para suavizar o problema de sensibilidade do padrão de sazonalidade às mudanças conjunturais (componente irregular) sempre se dá algum tratamento estatístico aos dados que



representam casos extremos de irregularidade. Vejamos como é feito este tratamento.

O procedimento classico em estatística é o de considerar como extremo todo dado acima ou abaixo da média por um certo número de desvios padrões. Por exemplo: podemos considerar como de irregularidade extrema todo dado acima da média mais dois desvios padrões ou abaixo da média menos dois desvios. Isto é, se  $x > M + 2\sigma$  ou  $x < M - 2\sigma$  então  $x$  sera considerado fora dos padrões de normalidade. Por trás destas considerações está implícito que a componente irregular esta sendo considerada como tendo distribuição normal. Assim, se cortaria todo dado que excedesse dois desvios da média. Para efeito da obtenção do padrão de sazonalidade utilizando-se apenas três anos, o ano considerado de irregularidade extrema seria substituído por uma média do ano anterior e do ano posterior. Assim, se fosse constatado que em setembro do ano de 1978 o dado esta fora da normalidade ele seria substituído pela média dos dados de setembro de 1977 e 1979. Um procedimento mais sofisticado para tratamento é o que se segue:

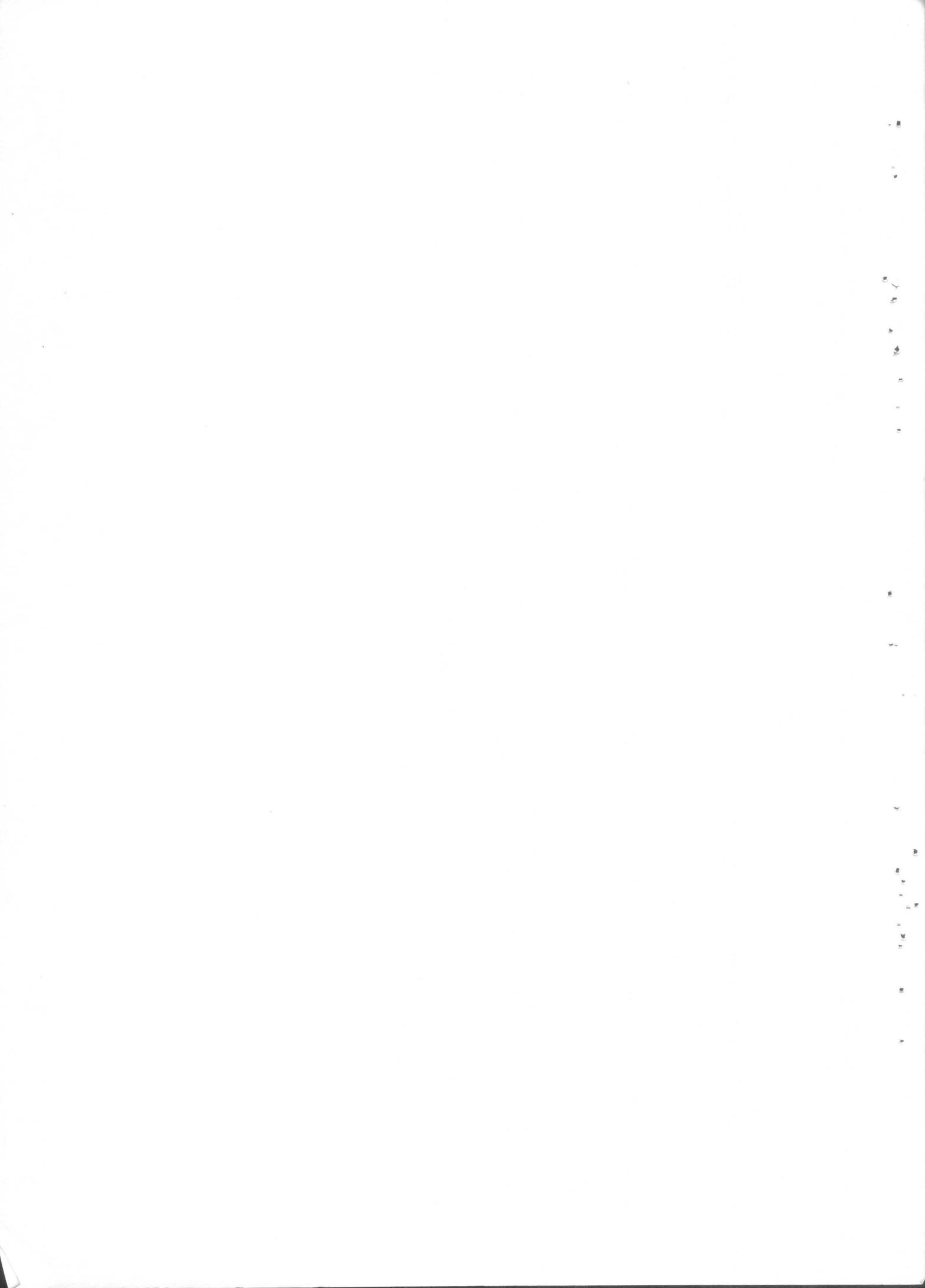
a) Considerar como extremo todo dado acima ou abaixo de 2.5 desvios padrões.

Neste caso substituir pela média dos anos posteriores e anteriores.

b) Considerar como normal todo dado abaixo de 1.5 desvios. Neste caso o dado é mantido.

c) Se o dado estiver na faixa de 1.5 a 2.5 desvios padrões dar a ele um peso proporcional à sua distância do ponto correspondente a 2.5 desvios. Vale dizer, até 1.5 desvios o dado tem peso um, entre 1.5 e 2.5 tem peso entre um e zero e acima de 2.5 o dado tem peso nulo.

Façamos um exemplo do caso intermediário (entre 1.5 e 2.5 desvios). Suponhamos que setembro de 1978 é um dado com irregularidade acima do normal e que se encontra à distância de 1.9 da média. Neste caso, 1.9 desvios corresponde a um peso de 0.6 (2.5 - 1.9)



O tratamento deste dado para o caso de uma média móvel de três anos será considerar em seu lugar a média ponderada:

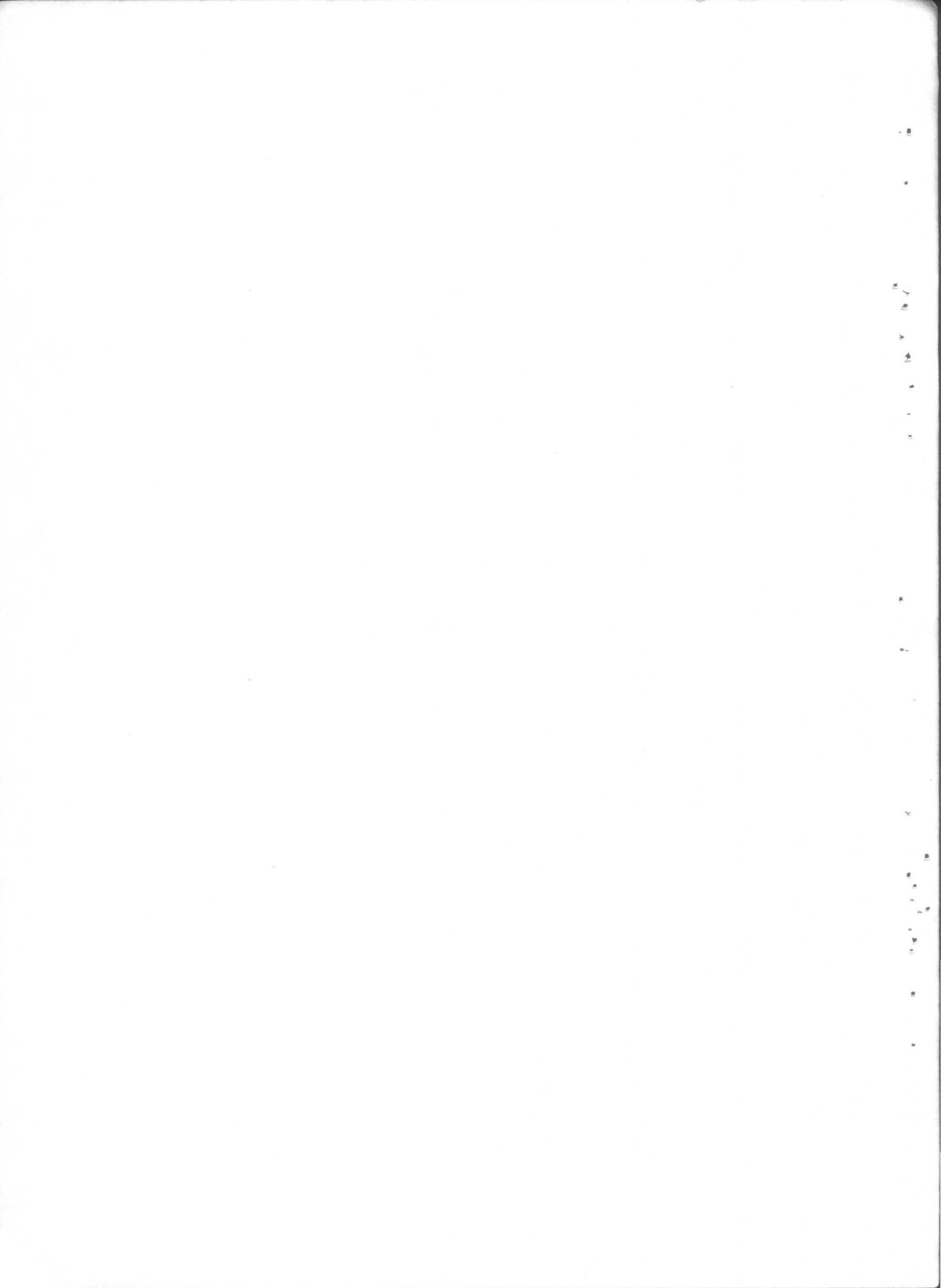
$$\frac{\text{setembro 77} + 0.6 \times \text{setembro 78} + \text{setembro 79}}{2.6}$$

Uma tabela dos pesos para este processo é:

Desvio em relação à média	Peso
Menos de 1.5 $\sigma$ .....	1.0
1.6 $\sigma$ .....	0.9
1.7 $\sigma$ .....	0.8
1.8 $\sigma$ .....	0.7
1.9 $\sigma$ .....	0.6
2.0 $\sigma$ .....	0.5
2.1 $\sigma$ .....	0.4
2.2 $\sigma$ .....	0.3
2.3 $\sigma$ .....	0.2
2.4 $\sigma$ .....	0.1
Mais de 2.5 $\sigma$ .....	0

Vale dizer, os pesos vão diminuindo se a componente irregular for muito alta e o dado estiver muito afastado da média. A suposição de normalidade da componente irregular, se correta, indicaria que apenas muitos poucos dados (cerca de 14 em cada 100) estariam na faixa acima de 1.5 desvios padrões e menos ainda (menos de 2 em cada 100) na faixa superior a 2.5 desvios padrões.

Uma coisa talvez não tenha ficado clara: em que ponto do processo de análise se faz esta correção? Resposta: no ponto imediatamente posterior a termos achado as médias móveis e descontado estas médias da série original ficando com uma "sobra" que é o conjunto da componente irregular com a componente sazonal. Neste momento, a correção de dados extremos aplicada à este conjunto, retira parte da irregularidade. Isto faz com que o método não fique tão sensível às variações de irregularidade, o que nos permite utilizar, para obtenção do padrão de sazonalidade, poucos anos.



Vamos repetir as diversas fases na obtenção de um padrão móvel de sazonalidade.

I) Média móvel de 12 meses para calcular a componente secular cíclica.

II) Divisão da série original pela média móvel (caso do método multiplicativo) ou subtração da média móvel na série original (método aditivo). Obtendo-se o conjunto composto pelas componentes sazonal e irregular.

III) Para os anos da extremidade (primeiro e último) obtenção da componente sazonalidade-irregularidade dos seis primeiros meses e dos últimos seis meses por repetição dos meses correspondentes ao segundo e penúltimo ano.

IV) O conjunto sazonalidade-irregularidade sofre tratamento para dados extremos.

V) Um padrão móvel de sazonalidade é obtido fazendo-se média móvel de alguns anos (normalmente, 3 anos ou 5 anos).

VI) Os anos que não estão na extremidade podem ser obtidos por média móvel simples. Ou por médias ponderadas atribuindo-se menos peso aos anos mais afastados.

VII) Os anos da extremidade (primeiro e últimos) podem ser obtidos por meio de pesos ou de um processo auto-regressivo.

Não incluímos aí o trabalho inicial de depuração da componente devida às variações de calendário. O item VI) tem de ser melhor explicado. Até o momento utilizamos unicamente médias móveis de três anos com pesos simples ( $1/3, 1/3$  e  $1/3$ ) para os anos do meio e  $1/3, 2/3$  para os da extremidade. No entanto, a escolha de um número maior de anos para a média, digamos, 5 anos, pode nos levar a utilizar pesos diferentes, privilegiando os anos próximos ao meio da média móvel e dando pesos menores aos da extremidade.

10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65  
66  
67  
68  
69  
70  
71  
72  
73  
74  
75  
76  
77  
78  
79  
80  
81  
82  
83  
84  
85  
86  
87  
88  
89  
90  
91  
92  
93  
94  
95  
96  
97  
98  
99  
100

Isto corresponde a fazer o que se convencionou chamar de "dupla média móvel".

Suponhamos que temos 8 elementos A, B, C, D, E, F, G e H para fazermos uma média móvel de três anos. As médias móveis serão:

$$\frac{A+B+C}{3} \quad \frac{B+C+D}{3} \quad \frac{C+D+E}{3} \quad \frac{D+E+F}{3} \quad \frac{E+F+G}{3} \quad \frac{F+G+H}{3}$$

destas médias móveis queremos fazer médias móveis. Teremos uma "dupla média móvel de 3 x 3":

$$\frac{\frac{A+B+C}{3} + \frac{B+C+D}{3} + \frac{C+D+E}{3}}{3} \quad \frac{\frac{B+C+D}{3} + \frac{C+D+E}{3} + \frac{D+E+F}{3}}{3} \quad \frac{\frac{C+D+E}{3} + \frac{D+E+F}{3} + \frac{E+F+G}{3}}{3}$$

e

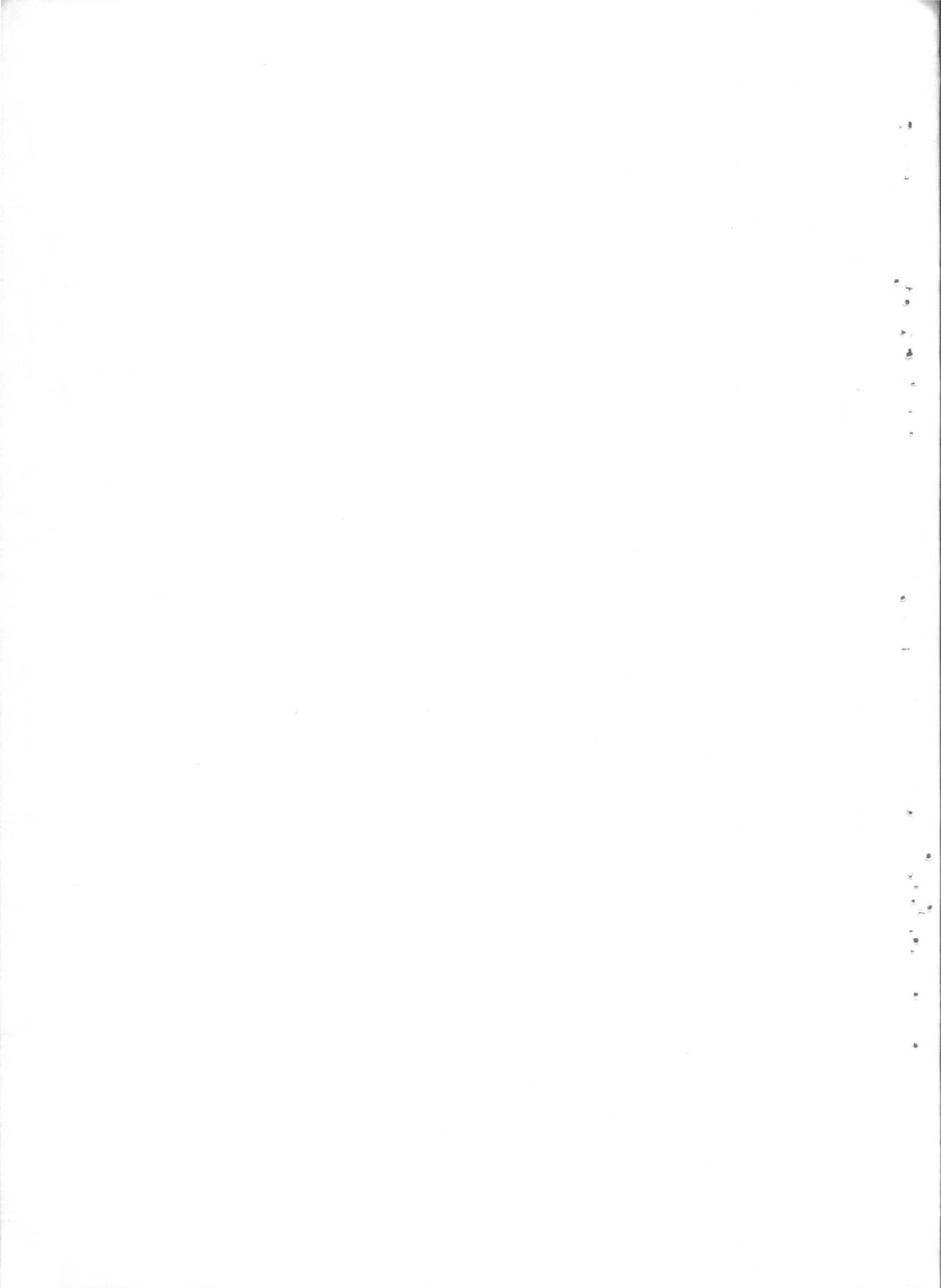
$$\frac{\frac{D+E+F}{3} + \frac{E+F+G}{3} + \frac{F+G+H}{3}}{3}$$

Se examinarmos uma destas médias móveis, a primeira por exemplo, veremos que ela pode se transformar em

$$\frac{(A+B+C) + (B+C+D) + (C+D+E)}{9} = \frac{A+2B+3C+2D+E}{9}$$

o que equivale a fazermos uma média móvel ponderada de 5 anos (pesos: 1/9, 2/9, 3/9, 2/9 e 1/9. O ano do meio (C) tem peso maior que os outros. Os afastados (A e E) tem peso menor.

Quais pesos serão utilizados numa média móvel de 5 anos para obtenção dos anos da extremidade? Quando estavamos utilizando uma média móvel de três anos os pesos eram 1/3 e 2/3. E era apenas um ano em cada extremidade a obter. No caso de uma média móvel de 5 anos temos que estimar os dois primeiros e os dois últimos anos. Coloquemos um sinal de interrogação na média móvel correspondente ao penúltimo ano:



$$\frac{(E+F+G) + (F+G+H) + (G+H+?)}{9}$$

O que colocar no lugar de nosso ponto de interrogação?

O natural seria, ou aumentar ainda mais o peso do penúltimo ano, já que a média móvel esta centrada nele, ou aumentar o peso dos três últimos anos colocando a média destes três últimos anos. Isto equivale a por no lugar do ponto de interrogação a média  $\frac{F+G+H}{3}$ . Donde a média seria:

$$\frac{(E+F+G) + (F+G+H) + (G+H + \frac{F+G+H}{3})}{9} = 0.111E + 0.259F + 0.307G + 0.259H$$

Os pesos são portanto (0.111, 0.259, 0.370 e 0.259). Resultados empíricos tem demonstrado que esta solução é a que dá melhores resultados (a alternativa sendo usar um método de auto-regressão).

E o último ano? Vejamos como seria a média móvel:

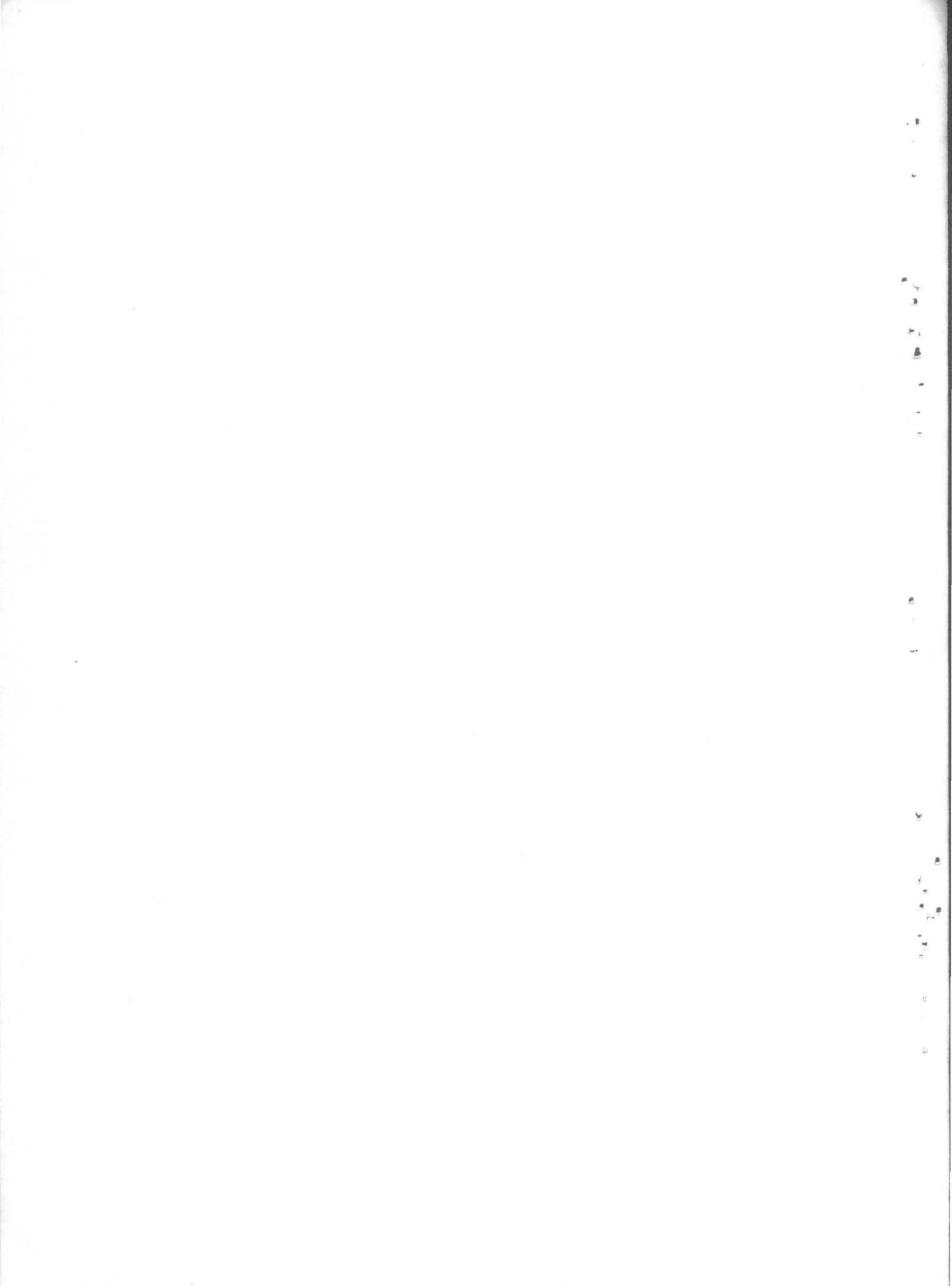
$$\frac{(F+G+H) + (G+H+?) + (H+?+??)}{9}$$

A substituição de dois elementos na parte (H + ? + ??) torna frágil demais este elemento. A solução heurística encontrada por alguns autores que advogam a utilização de pesos foi a de repetir o elemento  $\frac{F+G+H}{3}$  duas vezes na média móvel. O que nos leva a:

$$\frac{(F+G+H) + (G+H + \frac{F+G+H}{3}) + (G+H + \frac{F+G+H}{3})}{9} = 0.185F + 0.407G + 0.407H$$

atribuindo-se então os pesos 0.185, 0.407 e 0.407 aos três últimos anos para determinação do padrão do último ano.

Resumindo, o método mais comum de média ponderada usa os pesos (0.111, 0.222, 0.333, 0.222 e 0.111) para a média móvel dos pontos centrais, os pesos (0.111, 0.259, 0.370 e 0.259) para o penúltimo (e para o segundo) ano e os pesos (0.185, 0.407 e 0.407) para o último (e o primeiro) ano.



### O método X-11

O método X-11 é a última variante de uma sucessão de métodos de ajuste sazonal construídos no Bureau of the Census americano, conhecidos pela letra X numa sequência que começou em 1954 com o método X-1. O método X-11 foi introduzido em 1965, sendo agora amplamente utilizado internacionalmente para ajustar séries de índices.

Este método é um refinamento do método de médias móveis, que tem a vantagem de medir mais precisamente as componentes da série e permitir maior flexibilidade do que outros métodos usados antes dele. Além disso, ele permite a análise de cada um dos estágios sucessivos do processo de ajuste sazonal. Por esta razão, ele foi adotado (ou ele próprio ou sua variante X-11 ARIMA) por quase todos os grupos engajados no trabalho de ajuste sazonal em grande escala, a despeito do fato de ser relativamente trabalhoso.

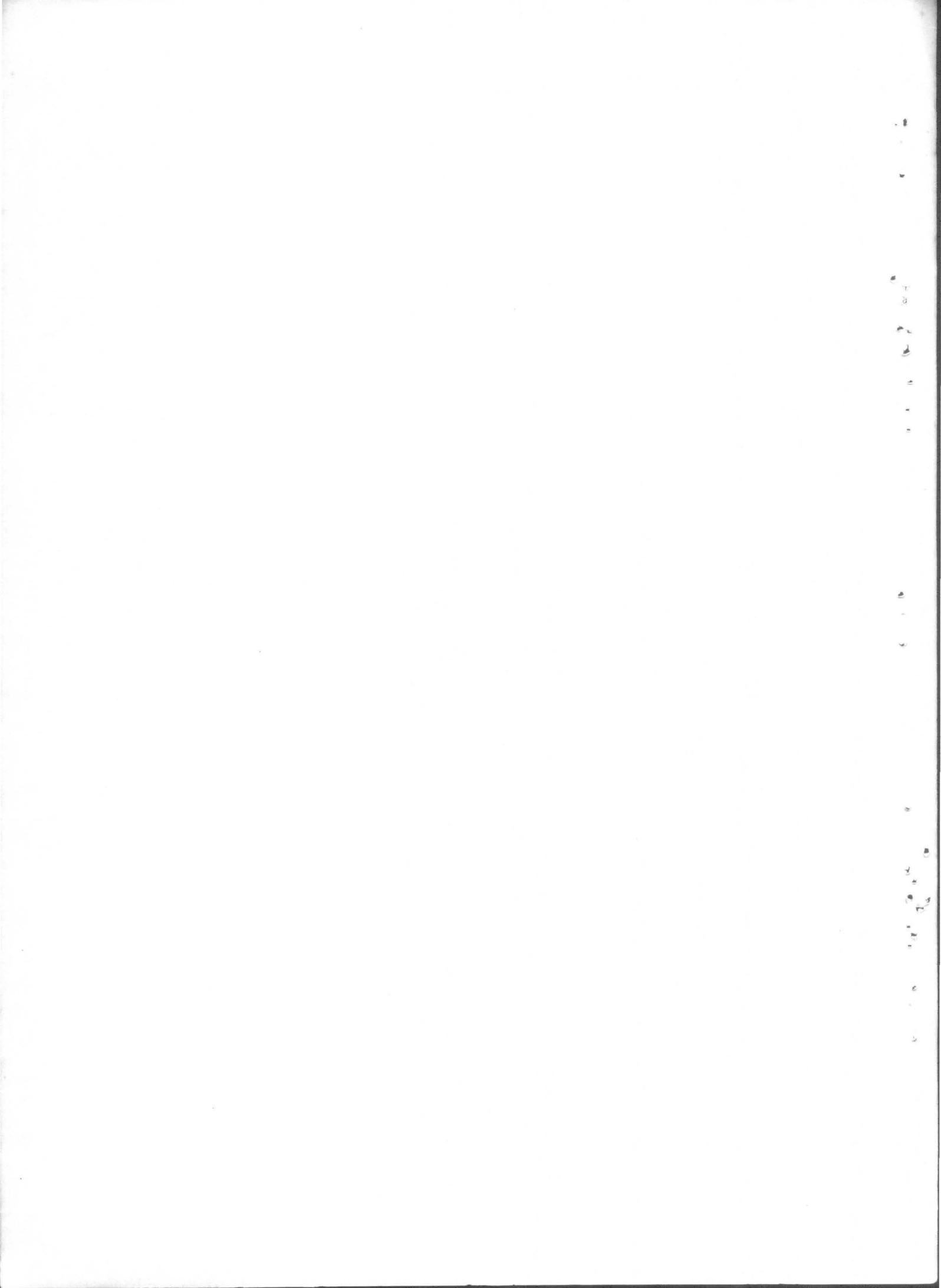
Em linhas gerais, o procedimento padrão do método X-11 consiste nos seguintes passos:

1 - Obtenção de estimativas preliminares da componente de tendência aplicando uma média móvel centrada de 12 meses (média móvel  $2 \times 12$ ).

2 - Obtenção da estimativa preliminar da componente sazonal através de média móvel ponderada de 5 anos (isto é, dupla média móvel  $3 \times 3$ ) em cada mês, de forma a obter os fatores e correção sazonal a serem aplicados ao mês.

3 - Substituição dos valores da componente sazonalidade irregularidade com valores de irregularidade extrema. Os critérios de substituição são os seguintes:

- a - obtenção de desvios móveis de 5 anos (60 meses);
- b - atribuição de pesos zero aos valores cuja irregularidade esteja acima de  $2.5\sigma$ ;
- c - atribuição de pesos proporcionais (entre 1 e 0) aos valores em que a componente irregular esteja entre  $1.5\sigma$  e  $2.5\sigma$ ;
- d - atribuição de peso 1 aos valores com irregularidade abaixo de  $1.5\sigma$ ;



e - substituição dos valores de peso inferior a 1, levando em consideração o peso e os valores de mesmo mês próximos do dado a ser substituído.

4 - Após a substituição dos valores com irregularidade extrema, recalculer os passos 1 e 2.

5 - Ajustamento da variação correspondente aos dias de trabalho (trading-days) por regressão entre a série obtida no item 4 e as variáveis  $X_1, X_2, \dots, X_7$ , em que  $X_1$  = número de domingos, ...,  $X_7$  = número de sábados no mês. Em seguida recalculer o passo 2.

6 - Ajustar a série pelo número de dias do mês (a menos que um comando opcional atribuindo diferentes pesos para os meses sejam dados à priori pelo operador).

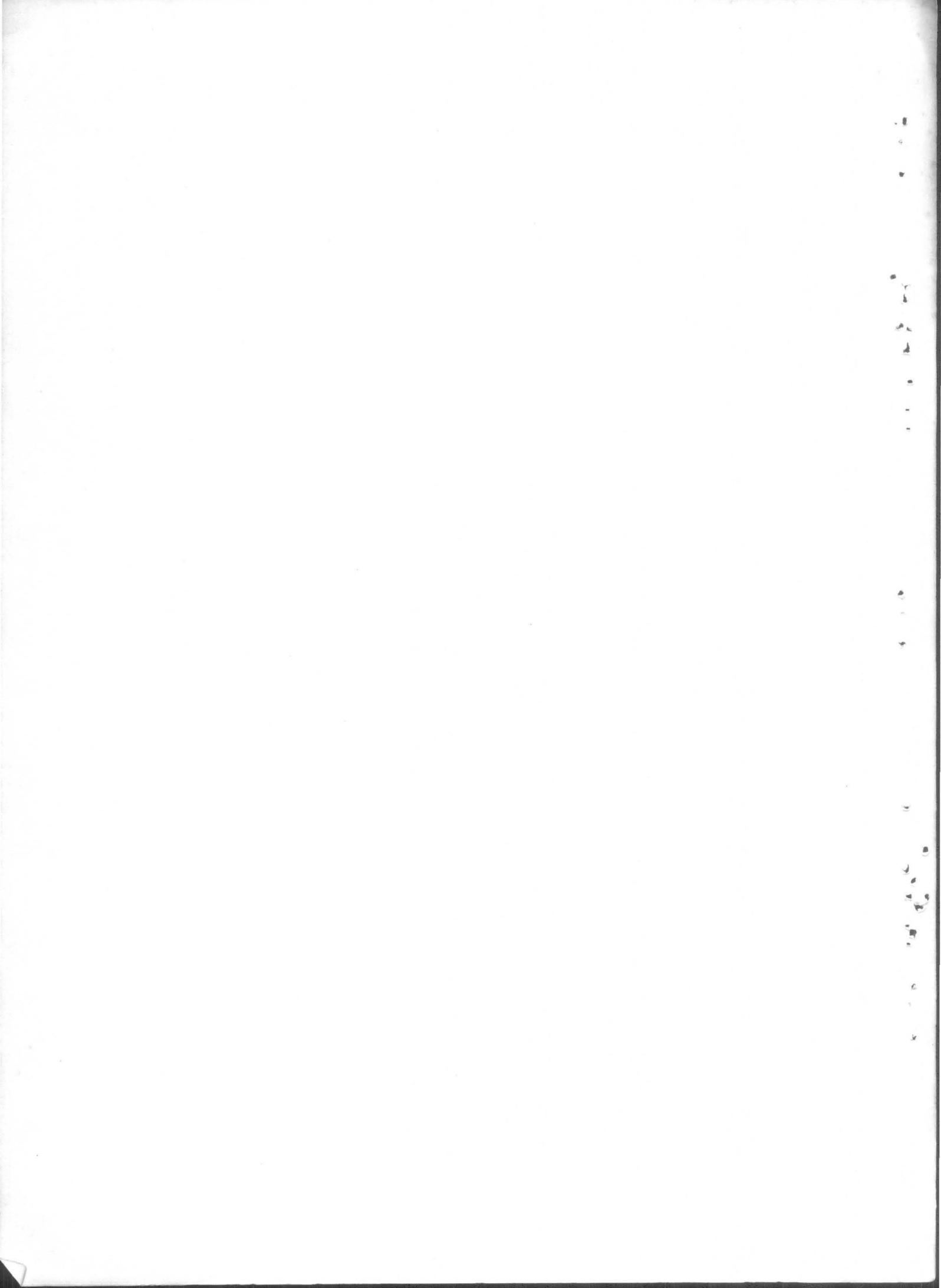
7 - Recalculer os passos de 1 a 4 para obtenção de das séries sazonalmente ajustadas finais.

8 - Aplicação de uma média móvel centrada de 13 meses à série sazonalmente ajustada para a obtenção da componente de tendência.

9 - Testar a existência de sazonalidade estável e a existência da componente "trading-days".

10 - Calcular estimativas de previsão dos fatores sazonais para o ano seguinte aos dados.

11 - Traçar gráficos, dar um sumário das medidas obtidas, etc.



## Bibliografia

SHISKIN, J.; YOUNG, A. & MUSGRAVE, S. - The X-II variant of the Census Method II Seasonal Adjustment Program. - United States Department of Commerce, Bureau of the Census - 1967.

DAGUM, Estela B. - The Magic Box and the four Golden Roules. Technical paper - Ottawa - Statistics Canada, 1980.

DAGUM, Estela B. - The effects of Time Varying Filters on Seasonal Factor Revisions. - Technical paper - Ottawa - Statistics Canada, 1981.

GOUVEA V.H.C. & GARCIA, R. M. - Ajuste Sazonal para os Indicadores de Produção Física - Revista Brasileira de Estatística n<sup>o</sup> 175 - pp 320 - 375 - Rio de Janeiro - 1983.

KENNY, P. B. & DURBIN, S. - Local Trend Estimation and Seasonal Adjustment of Economic and Social Time Series. - Journal of the Royal Statistical Society, series A, 145 - pp 1 -41 - 1982

