CHAPTER 4

APPLICATIONS OF THE TESTS THAT EVALUATE THE PREDICTABILITY OF A LINEAR MODEL AND THAT COMPARE THE PREDICTABILITY OF TWO LINEAR MODELS BASED ON THE χ^2 AND THE CORRELATED GAMMA-RATIO DISTRIBUTIONS

4.1.1 APPLICATION OF χ^2 AND CORRELATED GAMMA RATIO TEST FOR THE INDIANA CROP-YIELD DATA

In this section, we make use of the data of corn crop at the state of Indiana in the USA, in the time interval 1931-1980.(Xekalaki and Katti 1984)

The crop yields for ten different districts (CRD) are given. Two different sets of variables have been used to predict the crop yield for the next years, for each of the district.

The first set of variables, is common for all the districts. The following sixteen predictors for 48 years (1932-1980, 1970 is not included), as well as the predictions and their standard errors for the 29 last years ([1951-1980]), are given.

MODEL A (M_A)

1. CPRE : the rainfall between September and June
2. CPRE-SQ : the square of CPRE
3. R7 : the rainfall during July
4. R7-SQ : the square of R7
5. R8 : the rainfall during August
6. R8-SQ : the square of R8
7. T6 : the temperature during June
8. T6-SQ : the square of T6
9. T7 : the temperature during July
10. T7-SQ : the square of T7
11. T8 : the temperature during August
12. T8-SQ : the square of T8
13. TREND 1 $= \begin{cases} \text{year} - 1929 & \text{if year} < 1961 \\ 31 & \text{otherwise} \end{cases}$
14. TREND 2 $= \begin{cases} \text{year} - 1960 & \text{if year} \geq 1961 \\ 0 & \text{otherwise} \end{cases}$
15. TREND 2-SQ = the square of TREND 2
16. INTERCEPT.

The second set of variables is different for every district. In this set predictors for 49 years (1931-1980, 1970 is not included) as well as the predictions and their standard errors for the 24 last years ([1956-1980]) are given:

MODEL B (M_B)

CRD 10.

1.TRIPE
2.TREND
3.TREND2
4.E7
5.E8
6.INTERCEPT

CRD 30.

1.TREND
2.TREND2
3.T8
4.DEF7
5.INTERCEPT

CRD 50.

1.TREND
2.TREND1
3.MR7
4.T8
5.INTERCEPT

CRD 70.

1.TRIPE
2.TREND2
3.INTERCEPT

CRD 90.

1.TRIPE
2.TREND
3.TREND 2
4. E6
5. E7
6. INTERCEPT

CRD 20.

1.TRIPE
2.TREND
3.TREND1
4.E7
5.CMR4
6.INTERCEPT

CRD 40.

1.TRIPE
2.TREND
3.TREND1
4.DEF7
5.DF2CRAUG
6.INTERCEPT

CRD 60.

1.TREND
2.TREND1
3.DEF7
4.INTERCEPT

CRD 80.

1.TRIPE
2.TREND2
3.R7
4.INTERCEPT

CRD 100.

1.TRIPE
2.TREND
3.TREND1
4.TREND 2
5. E6_E7
6. INTERCEPT

The aim of the application is to compare the predictability of the two models used by the USDA for predicting crop yield for the ten districts, by considering the test based on the Correlated Gamma Ratio distribution. Some practical problems in applying the hypothesis test, were the following:

- The predictions were estimated only for the last 29 and 24 years respectively, which means that we may not have a sufficient number of time points as the theorem of Brown-Kendall demands.
- The Correlated Gamma Ratio distribution requires the same number of time-points ($\kappa=n/2$).
- The estimation of the correlation coefficient also requires the same number of observations.
- The correlation coefficient has to be the correlation coefficient of the population and not the empirical one.

For all the above reasons, we had to consider the same number of years (24 years). We calculated the standardized residuals of predictions using (3.2.2) for each of the model,for the interval 1956-1980(tables 1,3).

We also estimated the average and the variance of the standardized residuals of predictions. These have to be approximately standard normal distributed for a large number of years. On the other hand the number of years is 24, not too large. Nevertheless, the average and the variance are quite often close to 0 and 1 respectively.

We also calculated the empirical correlation coefficients between the standardized residuals of the predictions for the two models for each of the district (table 2) as well as the ratios Z (table 4):

$$Z = X/Y = \frac{\sum_{t=1956}^{1980} r_A^2(t)}{\sum_{t=1956}^{1980} r_B^2(t)} \quad \text{or} \quad Z = Y/X = \frac{\sum_{t=1956}^{1980} r_B^2(t)}{\sum_{t=1956}^{1980} r_A^2(t)}$$

TABLE 1: *Standardized Residuals of Predictions of Model A for the ten districts*

	MODEL A									
	CRD10	**CRD20**	**CRD30**	**CRD40**	**CRD50**	**CRD60**	**CRD70**	**CRD80**	**CRD90**	**CRD100**
YEAR										
1956	-1,413	-0,183	0,732	-0,233	0,295	1,315	-1,682	-2,310	-0,226	-0,211
1957	-0,172	-0,141	1,354	1,692	1,203	0,966	1,162	-0,066	0,621	1,089
1958	-1,224	0,130	-0,114	-2,831	-0,763	-0,092	-0,494	-0,325	-1,466	-1,355
1959	-0,854	0,109	-1,295	-0,688	-0,983	-1,429	-0,524	-0,375	-0,232	-0,540
1960	-1,385	-0,403	0,247	-0,699	-0,496	-0,963	-0,847	-0,225	0,402	-1,196
1961	-2,671	-1,595	-0,840	-1,343	-1,053	-0,778	-1,996	-2,846	-2,641	-1,577
1962	-1,442	-1,828	-0,800	-3,794	-2,700	-3,700	-2,832	-2,095	-3,575	-2,213
1963	-1,179	-0,328	-0,490	0,867	0,782	3,140	-0,259	-1,197	0,317	-0,230
1964	1,152	1,575	1,570	1,658	1,778	2,103	0,782	1,420	0,285	1,967
1965	-1,080	-2,160	-2,434	-2,636	-3,824	-2,192	-2,428	-3,666	-3,908	-2,153
1966	0,905	0,414	0,314	0,734	0,490	0,014	1,989	1,324	-0,510	1,727
1967	-1,788	-1,503	-0,238	-0,720	-0,255	-0,703	-1,174	-0,462	-0,777	-0,835
1968	0,058	-0,908	-1,046	-1,229	-0,238	-0,050	0,504	-0,225	0,867	-0,358
1969	-1,338	0,239	0,196	-2,039	-1,941	0,702	-1,612	-1,670	-2,046	-1,509
1971	-2,866	-0,960	0,264	0,891	1,213	0,545	-0,189	0,932	0,863	1,151
1972	-0,632	-0,295	-0,420	-0,904	-0,908	-1,535	-1,893	-1,380	-0,641	-0,502
1973	-0,169	-0,846	0,368	-1,326	0,727	0,691	-0,426	-0,532	0,124	-0,179
1974	3,309	5,589	2,724	1,202	2,672	1,962	0,203	0,616	1,292	1,663
1975	-0,830	0,189	-0,245	-2,897	0,153	0,648	-1,123	2,522	0,073	-3,236
1976	-2,882	-1,774	-0,780	0,642	0,448	0,749	-0,715	-2,031	-3,225	-0,273
1977	-1,132	0,232	-0,832	-1,993	-1,203	-1,674	-0,892	-1,533	-1,981	-1,118
1978	0,015	0,736	-0,133	1,092	-1,744	-0,138	0,070	0,174	-0,654	1,471
1979	-0,465	-0,255	-0,268	-0,818	0,097	-2,271	-0,574	-0,390	-0,141	-0,274
1980	1,920	2,277	0,000	2,227	-0,271	0,021	1,002	1,056	0,196	1,135
aver	-0,673	-0,070	-0,090	-0,548	-0,272	-0,111	-0,581	-0,553	-0,708	-0,315
var.	2,085	2,546	1,063	2,716	2,054	2,420	1,384	2,176	2,150	1,919

TABLE 2: *Correlation Coefficients between the Standardized Residuals of Predictions for Model A and Model B for the ten districts.*

	ρ
CRD 10	0,803238
CRD 20	0,908718
CRD 30	0,885794
CRD 40	0,449055
CRD 50	0,620981
CRD 60	0,155735
CRD 70	0,56102
CRD 80	0,796186
CRD 90	0,669246
CRD 100	0,593629

TABLE 3: *Standardized Residuals of Predictions of Model B for the ten districts*

MODEL B										
	CRD10	CRD20	CRD30	CRD40	CRD50	CRD60	CRD70	CRD80	CRD90	CRD100
YEAR										
1956	-1,330	-0,266	0,628	-1,163	-0,312	0,122	-0,469	-0,315	-0,262	-0,273
1957	0,411	0,634	0,964	0,679	0,775	0,286	1,946	1,386	0,664	1,314
1958	-0,135	1,498	0,726	0,438	2,220	2,011	1,102	2,258	0,753	0,456
1959	0,273	-0,271	-0,558	-1,029	-2,019	-0,512	0,292	-0,285	-1,162	-0,492
1960	-1,110	0,060	0,048	-1,305	-1,182	-1,472	0,744	0,989	0,629	-0,252
1961	-0,797	-0,346	-0,652	-1,100	-1,068	-0,948	-0,226	-1,454	-1,869	-1,165
1962	-1,052	-1,356	-0,675	-2,071	-1,643	-0,888	-1,198	-0,512	-1,828	-2,516
1963	-1,070	-0,478	-0,653	-3,348	-1,646	-2,066	-1,869	-1,145	-0,147	-2,096
1964	1,218	1,231	1,641	0,567	1,060	1,301	-0,341	1,989	0,617	1,309
1965	-2,310	-2,550	-1,890	-2,790	-2,961	-1,773	-3,059	-2,432	-0,754	-2,989
1966	1,684	0,455	0,078	0,898	0,372	-0,494	1,032	0,028	-2,042	0,776
1967	1,476	0,146	0,410	-0,085	0,505	-0,067	-0,508	-0,443	-0,058	0,768
1968	0,654	-0,496	-1,653	0,375	-1,031	-0,859	0,340	-0,115	-0,484	-0,118
1969	-1,239	-1,081	-0,525	-1,031	-1,989	-1,161	-1,508	-1,883	-2,648	-1,728
1971	-0,697	-0,938	-0,186	-0,087	-0,684	-0,580	1,038	1,544	1,141	-0,121
1972	-0,059	-0,618	-1,002	-0,994	-1,470	-1,774	-0,746	-1,612	-0,467	-0,836
1973	0,760	-0,284	-0,215	-0,316	-0,811	-0,577	0,344	0,054	1,579	0,121
1974	7,035	5,734	4,092	4,816	3,299	1,907	4,507	1,945	2,582	6,462
1975	0,816	-0,463	-0,818	-0,572	-0,821	0,067	0,112	3,302	1,418	0,154
1976	0,101	-1,377	-0,464	-0,565	-0,924	-1,501	-0,253	-1,047	-1,281	-0,475
1977	0,714	-0,281	-1,488	1,029	0,670	-0,023	-0,137	-0,867	-2,176	-0,181
1978	0,983	0,406	-0,427	0,247	-0,592	-0,832	0,000	0,627	-0,731	-0,023
1979	0,601	-0,354	-1,128	1,074	1,254	1,465	1,611	2,161	0,371	0,047
1980	4,447	2,873	-0,228	2,808	-0,058	-0,738	0,479	0,406	-0,145	1,392
aver	0,474	0,078	-0,166	-0,147	-0,377	-0,379	0,135	0,191	-0,262	-0,020
var	3,800	2,585	1,509	2,877	2,070	1,275	2,112	2,233	1,746	3,215

TABLE 4: *Ratios (Z) of the Squared Standardized Residuals of Predictions of the two models, for the ten districts.*

	$\sum_{t=1}^{24} r_A^2(t)$	$\sum_{t=1}^{24} r_B^2(t)$	$Z = X/Y$	$Z = Y/X$
CRD 10	58,84408	92,79867		1,577027
CRD 20	58,68183	59,59543		1,015569
CRD 30	24,63873	35,35429		1,434907
CRD 40	69,67754	66,69173	1,04477	
CRD 50	49,00581	51,02836		1,041272
CRD 60	55,94915	32,78959	1,706308	
CRD 70	39,93337	49,01209		1,227347
CRD 80	57,39694	52,23236	1,098877	
CRD 90	61,46159	41,81048	1,470005	
CRD 100	46,51586	73,94394		1,58965

4.1.2 INFERENCE BASED ON THE χ^2-TEST FOR THE PREDICTABILITY OF ONE LINEAR MODEL

For each one of the models A and B we will test the hypothesis:

$$\begin{cases} H_0: \text{the model is appropriate for predictions} \\ H_A: \text{the model has lack of predictability} \end{cases}$$

According to 3.2.5 $\Sigma r^2_t \sim \chi^2_n$. The results of the hypothesis testing for each model are the following: (d.f : 24)

***TABLE 5** : P-Values of the hypothesis test for the appropriateness of the linear models A and B for predictions, for the ten districts.*

	MODEL A		MODEL B	
	Σr^2_t	P-VALUE	Σr^2_t	P-VALUE
CRD 10	58,8440	$9.3\ 10^{-5}$	92,7986	0,000
CRD 20	58,6818	$9.8\ 10^{-5}$	59,5954	$7.3\ 10^{-5}$
CRD 30	24,6387	***0,425597***	35,3542	***0,0633***
CRD 40	69,6775	$2\ 10^{-6}$	66,6917	$7\ 10^{-6}$
CRD 50	49,0058	$1.891\ 10^{-3}$	51,0283	$1.046\ 10^{-3}$
CRD 60	55,9491	$2.33\ 10^{-4}$	32,7895	***0,1086***
CRD 70	39,9333	0,021742	49,0120	$1.88\ 10^{-3}$
CRD 80	57,3969	$1.47\ 10^{-4}$	52,2323	$7.3\ 10^{-4}$
CRD 90	61,4615	$4\ 10^{-5}$	41,8104	0,01355
CRD 100	46,5158	$3.836\ 10^{-3}$	73,9439	$1\ 10^{-6}$

Rejecting the null hypothesis of a satisfactory predictability if p-value < 0.05, **we can conclude that only model A in CRD30 and models B in CRD30 and CRD60 are adequate for predictions.**

4.1.3 INFERENCE BASED ON THE CORRELATED GAMMA RATIO-TEST ABOUT THE PREDICTABILITY OF TWO COMPETING LINEAR MODELS

We will test the hypothesis:

$$\begin{cases} H_0 : M_B \text{ equivalent with } M_A \\ H_A : M_B \text{ better than } M_A \end{cases}$$

or the hypothesis :

$$\begin{cases} H_0 & : \ M_A \text{ equivalent with } M_B \\ H_A & : \ M_A \text{ better than } M_B \end{cases}$$

when M_B seems to be better than M_A or M_A seems to be better than M_B respectively.

CRD 10.

We test the hypothesis :

$$\begin{cases} H_0 & : \ M_A \text{ equivalent with } M_B \\ H_A & : \ M_A \text{ better than } M_B \end{cases}$$

where Z=1,577027, κ = 24/2 =12, ρ=0.803238 (tables 3-4). The p-value was estimated to be equal to 0.0355 < 0.1. So, we have enough evidence to reject the null hypothesis of the equivalence of the two models. This means that model A is better than model B in CRD 10.

CRD 20.

We test the hypothesis :

$$\begin{cases} H_0 & : \ M_A \text{ equivalent with } M_B \\ H_A & : \ M_A \text{ better than } M_B \end{cases}$$

where Z=1.015569, κ = 24/2 =12, ρ=0.908718. The p-value was estimated to be equal to 0.4656 > 0.1. So, we haven't enough evidence to reject the null hypothesis of the equivalence of the two models. This means that model A and model B are equivalent in CRD 20.

CRD 30.

We test the hypothesis :

$$\begin{cases} H_0 & : \; M_A \text{ equivalent with } M_B \\ H_A & : \; M_A \text{ better than } M_B \end{cases}$$

where Z=1,434907, κ = 24/2 =12, ρ=0.885794. The p-value was estimated to be equal to 0.0337 < 0.1. So, we have enough evidence to reject the null hypothesis of the equivalence of the two models. This means that model A is better than model B in CRD 30.

CRD 40.

We test the hypothesis :

$$\begin{cases} H_0 & : \; M_B \text{ equivalent with } M_A \\ H_A & : \; M_B \text{ better than } M_A \end{cases}$$

where Z=1,04477, κ = 24/2 =12, ρ=0.449055. The p-value was estimated to be equal to 0.453 > 0.1. So, we haven't enough evidence to reject the null hypothesis of the equivalence of the two models. This means that model A and model B are equivalent in CRD 30.

CRD 60.

We test the hypothesis :

$$\begin{cases} H_0 & : \; M_B \text{ equivalent with } M_A \\ H_A & : \; M_B \text{ better than } M_A \end{cases}$$

where Z=1,706308, κ = 24/2 =12, ρ=0.155735. The p-value was estimated to be equal to 0,0963 < 0.1. So, we have enough evidence to reject the null hypothesis of the equivalence of the two models. This means that model B is better than model A in CRD 60.

CRD 100.

We test the hypothesis :

$$\begin{cases} H_0 : M_A \text{ equivalent with } M_B \\ H_A : M_A \text{ better than } M_B \end{cases}$$

where Z=1,58965, $\kappa = 24/2 = 12$, $\rho = 0.593629$. The p-value was estimated to be equal to 0,0868 < 0.1. So, we can conclude that model A is better than model B in CRD 100.

The estimated p-values for all the CRD's can be summarized in the following table:

TABLE 6: *P-Values of the hypothesis test that compares the predictability of models A and B, for the ten districts.*

	P-VALUES	BEST MODEL
CRD 10	**0,0355**	**MODEL A**
CRD 20	0,4656	'EQUIVALENT'
CRD 30	**0,0337**	**MODEL A**
CRD 40	0,453	'EQUIVALENT'
CRD 50	0,45	'EQUIVALENT'
CRD 60	**0,0963**	**MODEL B**
CRD 70	0,275	'EQUIVALENT'
CRD 80	0,353	'EQUIVALENT'
CRD 90	0,1068	'EQUIVALENT'
CRD 100	**0,0868**	**MODEL A**

According to the results of table 5 and table 6, we can conclude that, based on the χ^2 and the correlated gamma-ratio test, we should consider model A for CRD30 and model B for CRD60.

4.1.4 COMPARISON OF THE CONCLUSIONS OF THE TEST BASED ON THE CORRELATED GAMMA-RATIO DISTRIBUTION FOR THE PREDICTABILITY WITH THE TEST BASED ON THE CROSS VALIDATION METHOD AND THE R^2 AND R^2_{ADJ} COEFFICIENTS.

We compared the conclusions of the analysis of the crop-yield data of the previous section with those of an analysis based on the Cross Validation method. Besides, the type of discrepancy used in a Cross Validation study resembles the one considered, used to evaluate the predictive ability of a linear model.

We also considered all the data (for 48 and 49 years for the two models respectively) but we took into consideration only the last 24 years to find the PRESS statistic for each model. The results are as follows:

***TABLE* 7** *: Press Criterion for the linear models A and B for the ten districts.*

	PRESS	
	MODEL A	**MODEL B**
CRD 10	**1134,037**	1206,723
CRD 20	**636,2727**	686,2955
CRD 30	**385,6956**	494,1782
CRD 40	**849,3338**	1076,734
CRD 50	**640,8159**	838,3118
CRD 60	891,8188	**542,6568**
CRD 70	**399,9135**	575,6925
CRD 80	**419,5981**	628,0908
CRD 90	**364,723**	425,8439
CRD 100	**508,3535**	635,9434

Comparing the results of table 6 and table 7 we notice that the two methods arrive at the same conclusions, for CRD's : 10, 30, 60 ,100, that is, that models A, A, B, A respectively are the most appropriate models for these districts.

As it concerns the other CRD's : 20, 40, 50, 70, 80, 90 the PRESS statistic considers that model A is better than model B while the Correlated Gamma-Ratio test indicates equivalence of the models.

For the above models, we estimated the coefficients R^2 and R^2_{adj} to evaluate the descriptive ability of the models:

TABLE 8: *Coefficients R^2 and R^2_{adj} for model A and model B for the ten districts.*

	MODEL A		**MODEL B**	
	R^2	**R^2_{adj}**	**R^2**	**R^2_{adj}**
CRD 10	0,9467	0,92171	0,919	0,91
CRD 20	0,96	0,94181	0,934	0,927
CRD 30	0,966	**0,95066**	0,92565	**0,91889**
CRD 40	0,967	0,95221	0,9299	0,92185
CRD 50	0,965	0,94875	0,92843	0,92192
CRD 60	0,966	0,95123	0,90	**0,89**
CRD 70	0,978	0,96783	0,948	0,946
CRD 80	0,97	0,95685	0,942	0,938
CRD 90	0,972	0,95980	0,95177	0,94616
CRD 100	0,973	0,96172	0,9419	0,9366

Table 8 shows that all the models describe data in a very satisfactory way. Comparing these results with these of table 5 we can notice that most of the models that describe well the data don't forecast well. Besides, models that present small R^2_{adj}, comparing with the others, seem to predict well.

4.2 ANOTHER APPLICATION OF THE TESTS BASED ON THE χ^2 AND THE CORRELATED GAMMA RATIO DISTRIBUTION FOR THE IOWA CROP-YIELD DATA

The method of testing the predictability of one and two linear models was applied to another set of real data used by Drapper and Smith (1981, p. 407). The data refer to corn crop yields at the state of Iowa in the USA, in the time interval 1930-1962.

The true crop-yield for every year is given and the following nine variables are used as possible predictors of the crop-yield (see Appendix, table 27):

PREDICTOR	NOTATION
1.Serial number of the years.	1
2.Preseason Precipitation.	2
3.May Temperature.	3
4.June Rainfall.	4
5.June Temperature.	5
6.July Rainfall.	6
7.July Temperature.	7
8.August Rainfall.	8
9.August Temperature.	9

In order to evaluate the predictive ability of the models two different methods are used:

- Firstly, we use the backward elimination procedure, and we get the results given in table 9:

***TABLE 9:** Coefficient of determination and Adjusted Coefficient of Determination for the models that arise from the Backward Elimination Procedure.(in the first column, on each line we have the variable retained)*

Model	R^2	R^2_{adj}
{1,2,3,4,5,6,7,8,9}	**0,74759**	0,6488
{1,2,3,4,5,6,8,9}	0,74751	0,663
{1,2,3,4,5,6,9}	0,74574	0,67454
{1,2,3,5,6,9}	0,73833	0,67794
{1,2,5,6,9}	0,72991	0,67990
{1,2,6,9}	0,72065	**0,68074**

According to the R^2_{adj} criterion, the subset of variables that give the best description of the data is the set {1,2,6,9} with the first variable as the most important one (the p-value of the t-test that the first variable is statistically different from zero is equal to 0.0). Nevertheless, the description of the data is not a satisfactory one since R^2_{adj} for the set of variables {1,2,6,9} is equal to 0.68074. We also notice that the values of R^2_{adj} for all the candidate models do not differ much.

Alternatively, for those models arising from the backward elimination procedure, we estimate the statistic Σr_t^2 (Squared Sum of the Standardized Residuals of Prediction) given in 3.2.2, in order to compare, each time, the predictability of the linear model to the one with one variable less.

We should stress that for estimating the statistic we consider only the residuals for the last 20 years (1943-1962). This because for the full model of the nine predictors we can only make predictions after a period of ten years since we use ten explanatory variables (9 predictors plus the intercept). We also do not take into consideration the results of the next two years because of the unreliability of the estimations. So, we consider the predictions for the last 20 years for all the models in order that the models are comparable. The results are given in the following table:

TABLE 10: Sum of the Squared Standardized Residuals of Predictions for the models that arise using the Backward Elimination procedure.

MODEL	$\sum_{t=1943}^{1962} r_t^2$
{1,2,3,4,5,6,7,8,9}	127,5744
{1,2,3,4,5,6,8,9}	**17,27544**
{1,2,3,4,5,6,9}	18,88685
{1,2,3,5,6,9}	21,39122
{1,2,5,6,9}	22,59598
{1,2,6,9}	23,03739

According to table 10 the model with the variables {1,2,3,4,5,6,8,9} seems to be the one with the best predictability. We have to compare this with the other models to find out if it is also statistically better. The ratio Z and the correlation coefficient of the above model and the one with the variables {1,2,3,4,5,6,8,9} are presented in the next table:

TABLE 11: Ratios Z and Correlation Coefficients of the residuals of the models that arise from the Backward Elimination Procedure with the one that seems to have the best predictive ability, that is, the model with variables {1,2,3,4,5,6,8,9}.

	{1,2,3,4,5,6,8,9}	
	Z	*ρ*
1. {1,2,3,4,5,6,7,8,9}	7,385	0,569
2. {1,2,3,4,5,6,9}	1,093	0,993
3. {1,2,3,5,6,9}	1,238	0,980
4. {1,2,5,6,9}	1,308	0,973
5. {1,2,6,9}	1,334	0,973

In table 12 we provide the percentiles of the Correlated Gamma-Ratio distribution for large values of the correlation coefficient: (for $\kappa = 20/2 = 10$)

TABLE 12: Percentiles of the Correlated Gamma-Ratio Distribution for κ=10 and ρ=0.90 to 0.99.

ρ	α = 0.1	α = 0.05	α = 0.01
0.90	1,3	1,4	1,63
0.91	1,28	1,38	1,60
0.92	1,27	1,36	1,56
0.93	1,25	1,33	1,52
0.94	1,23	1,31	1,47
0.95	1,21	1,28	1,43
0.96	1,18	1,25	1,37
0.97	1,16	1,21	1,32
0.98	1,13	1,17	1,26
0.99	1,09	1,12	1,18

For α = 0.05 we cannot reject the hypothesis of statistical equivalence : of the model with variables {1,2,3,4,5,6,8,9} to the one with variables {1,2,3,4,5,6,9}. This is so, since the critical value of the test statistic is $CGR_{10,\ 0.99,\ 0.05} = 1.12 > 1.093$ (see table 12).

Since, the model with variables {1,2,3,4,5,6,9} is equivalent to the model with variables {1,2,3,4,5,6,8,9} that is simpler in the sense that it contains fewer predictors, one should prefer the latter.

On the other hand, if one is searching for a more accurate model, one would prefer the model with the variables {1,2,3,4,5,6,8,9}.

We should also note that the models with variables {1,2,3,4,5,6,8,9} and {1,2,3,4,5,6,9} can be regarded as appropriate for predictions according to the χ^2-test since $\Sigma r^2_t = 17.27544 < \chi^2_{20;0.95} = 31.4$ and $\Sigma r^2_t = 18.88685 < \chi^2_{20;0.95} = 31.4$.

Remark: One may notice that the conclusions of the two procedures seem to differ. Nevertheless, both of them «judge» the model with variables {1,2,3,4,5,6,8,9} as a satisfactory one (look at tables 9 and 10). Besides, the procedures for the computation of R^2_{adj} and Σr^2_t use different types of residuals and different number of residuals (33 and 20 respectively).

• An alternative method is to apply the backward elimination procedure using the statistic Σr_t^2 as a criterion for eliminating variables. That is, in the first stage we have 9 different linear models with 8 predictors. We estimate the statistic Σr_t^2 for all the 9 models. The model for which the smallest Σr_t^2 is observed, indicates which variable can be omitted. In the next stages we proceed in the same way. The results of this approach for the Draper and Smith data are as follows:

1^{st} Stage

TABLE 13: Sum of the Squared Standardized Residuals of Predictions for the full model .

MODEL	$\sum_{t=1943}^{1962} r_t^2$
1. {1,2,3,4,5,6,7,8,9}	127,57*

(* For the computation of the statistic $\sum_{t=1943}^{1962} r_t^2$ of the model with variables {1,2,3,4,5,6,7,8,9} , see Appendix - table 28)

2nd Stage

TABLE 14: Sum of the Squared Standardized Residuals of Predictions for all the models resulting from the elimination of one predictor.

MODEL	$\sum_{t=1943}^{1962} r_t^2$
1. {2,3,4,5,6,7,8,9}	40,742161
2. {1,3,4,5,6,7,8,9}	132,90766
3. {1,2,4,5,6,7,8,9}	114,26441
4. {1,2,3,5,6,7,8,9}	91,006347
5. {1,2,3,4,6,7,8,9}	44,004969
6. {1,2,3,4,5,7,8,9}	101,97787
7. {1,2,3,4,5,6,8,9}	**17,275441****
8. {1,2,3,4,5,6,7,9}	105,25004
9. {1,2,3,4,5,6,7,8}	34,335191

(** For the computation of the statistic $\sum_{t=1943}^{1962} r_t^2$ of the model with variables {1,2,3,4,5,6,8,9}, see Appendix - table 29)

It is obvious that the smallest value of the statistic corresponds to model 7 (that results in from the elimination of predictor 7) at the 2nd stage of the procedure. To test whether this model has the best predictive ability of all the candidate models, we need to compare its predictability to that of the model of stage 1 and subsequently to the predictability of all the other models of stage 2. The ratios Z and the correlation coefficients are given in the following table:

TABLE 15 : Ratio Z and Correlation Coefficient of the residuals of the model of the first stage and the one of the second stage with the smallest value of the statistic Σr^2_t.

MODELS	Z	ρ
{12345689}-{123456789}	7,385	0,569

TABLE 16 : Ratios Z and Correlation Coefficients of the residuals of all the models of stage 2 with the model with variables {1,2,3,4,5,6,8,9}

	{1,2,3,4,5,6,8,9}	
	Z	***ρ***
1. {2,3,4,5,6,7,8,9}	2,358	0,86
2. {1,3,4,5,6,7,8,9}	7,693	0,60
3. {1,2,4,5,6,7,8,9}	6,6142	0,60
4. {1,2,3,5,6,7,8,9}	5,2679	0,67
5. {1,2,3,4,6,7,8,9}	2,547	0,84
6. {1,2,3,4,5,7,8,9}	5,903	0,62
7. {1,2,3,4,5,6,7,9}	6,092	0,61
8. {1,2,3,4,5,6,7,8}	1,987	0,88

From the above tables one may claim that the model with variables {1,2,3,4,5,6,8,9} has the best predictive ability of all the other models. **So, predictor 7 can be omitted.**

3rd Stage

We study models arrived at in stage 2 omitting also variable 7. The corresponding values of the test statistic are:

TABLE 17: Sum of the Squared Standardized Residuals of Predictions for all the models having eliminated two predictors.

MODEL	$\sum_{t=1943}^{1962} r_t^2$
1. {1,2,3,4,5,6,8}	**17,335565**
2. {1,2,3,4,5,6,9}	18,886848
3. {1,2,3,4,5,8,9}	20,729604
4. {1,2,3,4,6,8,9}	19,160255
5. {1,2,3,5,6,8,9}	19,478084
6. {1,2,4,5,6,8,9}	18,891581
7. {1,3,4,5,6,8,9}	19,753381
8. {2,3,4,5,6,8,9}	26,258426

The model corresponding to the smallest value of the statistic is the one resulting from the omission of predictor 9. Note that model 7 of the second stage corresponds to the smallest value of the statistic from all the models of stage 3. We now need to compare the predictability of model 7 of the second stage with model

1 of the stage 3 (the one that results in from the omission of predictor 9).

TABLE 18 : Ratio Z and Correlation Coefficient of the residuals of the models of the second and the third stage corresponding to the smallest value of the statistic Σr^2_t.

MODELS	Z	ρ
{1,2,3,4,5,6,8,9}-{1,2,3,4,5,6,8}	1,0035	0,87

Comparing the predictability of the models at α=0.05 we do not have enough evidence to reject the null hypothesis of the equivalence of the models. So, we do not omit variable 9 and the procedure terminates here. This means that the model that possesses the best predictability is the one that results in from the omission of variable 7.

Note: If we want to search for a model that is simpler and equivalent to the one with variables {1,2,3,4,5,6,8,9}, we may continue the procedure without using a statistical test but by eliminating that predictor, whose omission from the model gives the smallest Σr_t^2 statistic.

So, in stage 3 we eliminate predictor 9.

4th Stage

MODEL	$\sum_{t=1943}^{1962} r_t^2$
1. {1,2,3,4,5,6}	17,434134
2. {1,2,3,4,5,8}	23,947501
3. {1,2,3,4,6,8}	17,062579
4. {1,2,3,5,6,8}	**15,297831*** **
5. {1,2,4,5,6,8}	16,908574
6. {1,3,4,5,6,8}	20,21673
7. {2,3,4,5,6,8}	24,442708

(Variable 4 is eliminated)

(*** For the computation of the statistic $\sum_{t=1943}^{1962} r_t^2$ of the model with variables {1,2,3,5,6,8}, see Appendix - table 30)

5th Stage

MODEL	$\sum_{t=1943}^{1962} r_t^2$
1. {1,2,3,5,6}	16,433526
2. {1,2,3,5,8}	21,914002
3. {1,2,3,6,8}	17,644509
4. {1,2,5,6,8}	**16,213089**
5. {1,3,5,6,8}	17,011164
6. {2,3,5,6,8}	25,735334

(Variable 3 is eliminated)

6th Stage

MODEL	$\sum_{t=1943}^{1962} r_t^2$
1. {1,2,5,6}	**17,438941**
2. {1,2,5,8}	19,003137
3. {1,2,6,8}	17,975222
4. {1,5,6,8}	17,739044
5. {2,5,6,8}	27,273949

(Variable 8 is eliminated)

7th Stage

MODEL	$\sum_{t=1943}^{1962} r_t^2$
1. {1,2,5}	19,816709
2. {1,2,6}	19,923134
3. {1,5,6}	**19,26566**
4. {2,5,6}	29,437064

(Variable 2 is eliminated)

8th Stage

MODEL	$\sum_{t=1943}^{1962} r_t^2$
1. {1,5}	**18,800213**
2. {1,6}	18,894565
3. {5,6}	31,509477

(Variable 6 is eliminated)

9th Stage

MODEL	$\sum_{t=1943}^{1962} r_t^2$
{1}	**19,918106**
{5}	29,084934

From the above procedure we conclude that the model corresponding to the smallest value of the test statistic is the model with variables {1,2,3,5,6,8} of the 4th stage. This model contains fewer predictors than the one with variables {1,2,3,4,5,6,8,9} and is equivalent to it since:

MODELS	Z	ρ
{1,2,3,4,5,6,8,9}-{1,2,3,5,6,8}	1,129	0,93

and according to table 4 $CGR_{10,\ 0.93,\ 0.05} = 1.33 > 1.129$. So, the model with variables {1,2,3,5,6,8} can be used instead of the one with variables {1,2,3,4,5,6,8,9}.

Remark: Comparing the results of the two backward elimination procedures we notice a lot of similarities. Both of them eliminate variable 7 at the beginning and finally consider variable 1 as the most important variable.

4.3 SIMULATION STUDY

4.3.1 TESTING THE PREDICTABILITY OF A LINEAR MODEL BASED ON THE χ^2 DISTRIBUTION

According to the theorem of Brown and Kendall (see 3.2), the standardized residuals of the predictions (r_t) are mutually independent standard normal variables for large sample sizes. So:

$$r_t \sim N(0,1) \Rightarrow \sum_{t=1}^{n} r_t^2 \sim \chi_n^2$$

According to (3.2.5) we may test the predictability of a linear model using the statistic χ^2. We simulated ten samples from a Standard Normal distribution, using Microsoft Excel, for different sample sizes (30, 40, 50, 80 ,100). The value of the statistic Σr^2_t, for the different sample sizes is given in the following table:

Table 19: Value of the statistic Σr^2_t, simulated from a Standard Normal Distribution using Microsoft Excel, for different sample sizes.

	Sample Size				
Samples	**30**	**40**	**50**	**80**	**100**
1	30,47199	29,11291	51,83264	***98,77972***	95,91258
2	28,24655	50,41146	41,84502	78,36374	87,69239
3	29,45237	42,64283	40,52792	64,17798	113,8865
4	37,00891	39,52287	49,09213	***109,4855***	104,6006
5	16,69117	44,10941	40,72393	84,82636	115,5696
6	19,41776	47,98779	35,22784	88,17285	78,33449
7	24,57304	40,84238	42,09678	82,33548	106,964
8	24,58699	36,70731	56,23162	83,43191	73,81989
9	27,35341	43,74965	37,44991	***99,15593***	92,20017
10	22,45686	47,18792	28,86553	88,37012	***123,4442***

From the above 50 samples, we see that the null hypothesis of «appropriateness of the model for predictions» is rejected only in 4 cases (the ones in bold), at level of significance α=0.1.

4.3.2 COMPARING THE PREDICTABILITY OF TWO LINEAR MODELS BASED ON THE CORRELATED GAMMA-RATIO DISTRIBUTION

To simulate observations from a Correlated Gamma-Ratio distribution we will make use of the following facts:

- Let $r_{A_{(t+1)}}$ and $r_{B_{(t+1)}}$ be the standardized residuals of predictions for the t+1 time point, for two different linear models. Then for large sample, $r_{A_{(t+1)}}$ and $r_{B_{(t+1)}}$ are approximately standard normal variables which are mutually independent within each model (Brown’s (1975) and Kendall’s theorem(1985)). That is:

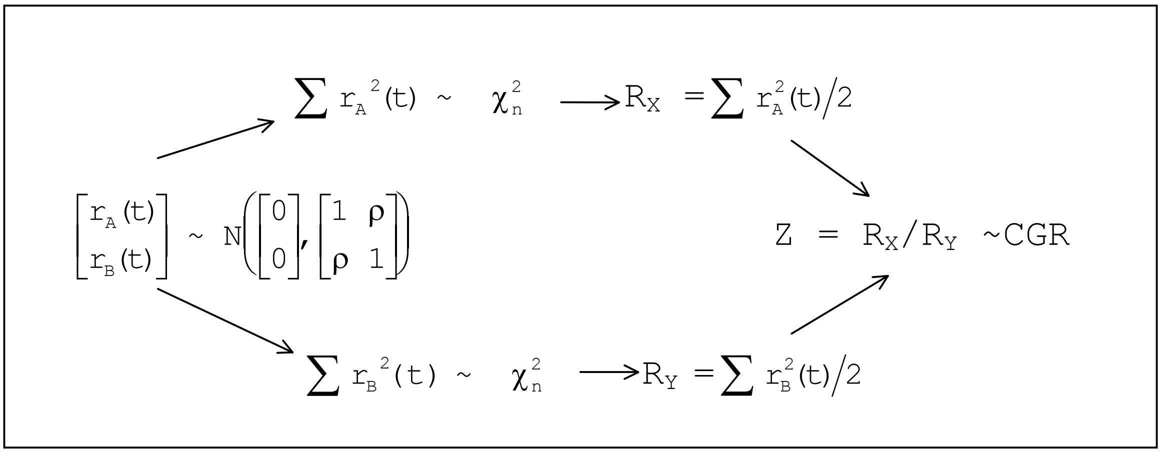
$$r_{A(t+1)} = \frac{\hat{Y}^{\circ}_{t+1} - Y^{\circ}_{t+1}}{S_t\sqrt{\left(\mathbf{X}^{\circ}_{t+1}\left(\mathbf{X}'_t\mathbf{X}_t\right)^{-1}\mathbf{X}^{\circ\,\prime}_{t+1} + 1\right)}} \sim N(0,1)$$

$$r_{B(t+1)} = \frac{\hat{Y}^{\circ}_{t+1} - Y^{\circ}_{t+1}}{S_t\sqrt{\left(\mathbf{x}^{\circ}_{t+1}\left(\mathbf{X}'_t\mathbf{X}_t\right)^{-1}\mathbf{x}^{\circ\,\prime}_{t+1} + 1\right)}} \sim N(0,1)$$

- The statistics $r_{A_{(t+1)}}$ and $r_{B_{(t+1)}}$ are not independent since they are based on predictions of the same response variable. We consider the correlation coefficient of the two variables as a measure of their association. The joint distribution of $r_{A_{(t+1)}}$, $r_{B_{(t+1)}}$ is a bivariate standard normal distribution.
- According to what we have already proved (paragraph 3.3), the ratio of the variables $X = \frac{\sum_{i=1}^{t} r_A^2(i)}{2}$ and $Y = \frac{\sum_{i=1}^{t} r_B^2(i)}{2}$ is Correlated Gamma Ratio distributed.

Based on the above we can construct a diagram (figure 7) which can lead us to the simulation steps:

Figure 7: Simulation steps of Correlated Gamma-Ratio variables.



From figure 7 it become obvious that the algorithm for generating a gamma-ratio variable amounts to the following steps.

- *Step 1* : Generate a sample of n observations (X_i, Y_i) on the vector (X,Y) that is distributed according to the $N\left(\begin{bmatrix}0\\0\end{bmatrix}, \begin{bmatrix}1 & \rho\\ \rho & 1\end{bmatrix}\right)$ distribution.

- *Step 2* : Compute $R_X = \sum_{i=1}^{n} X_i^2 / 2$ and $R_Y = \sum_{i=1}^{n} Y_i^2 / 2$.

- *Step 3* : Obtain $Z = R_X/R_Y$.

We simulated from a Bivariate Standard Normal using Statgraphics. We took five samples of the same size n and the same correlation coefficient. The values of n and p considered were n=24,60,100 and ρ=0.0, 0.1, … , 0.9. The results are summarized in tables 19-21. The entries of these tables are the values of the ratio $Z=R_X/R_Y$ for different sample sizes and different correlation coefficients. The values of Z that lead to the rejection of the null hypothesis of equivalence of the two models at the 0.05 level of significance appear in bold. Here we made use of the Appendix- «Percentiles of the Correlated Gamma Ratio distribution».

Table 20: Simulation results for the ratio Z, for different values of the correlation coefficient and κ=12 (n=24).

Correl. Coeff.	Sample 1 Z	Sample 2 Z	Sample 3 Z	Sample 4 Z	Sample 5 Z
0.0	1.1498	***2.39***	1.6688	1.1030	1.4024
0.1	***2.329***	1.338	1.4	1.118	1.45
0.2	1.257	1.1730	***2.458***	***2.226***	1.372
0.3	1.122	1.1955	1.122	1.8469	1.3469
0.4	***3.058***	1.209	1.4589	1.4968	1.8116
0.5	1.3101	1.3864	1.4236	1.7167	1.6567
0.6	1.5396	1.0738	***1.7436***	1.055	1.047
0.7	***1.8949***	1.03	1.327	1.187	1.1705
0.8	1.2411	1.0169	1.2595	1.098	1.263
0.9	1.277	1.003	***1.8429***	1.0732	1.0985

Table 21: Simulation results for the ratio Z, for different values of the correlation coefficient and κ=30 (n=60).

Correl. Coeff.	Sample 1 Z	Sample 2 Z	Sample 3 Z	Sample 4 Z	Sample 5 Z
0.0	1,1167	1,2933	***1,6422***	1,0627	***1,6779***
0.1	1,5032	1,3719	1,05	1,1815	1,0041
0.2	1,187	1,1929	1,14	***1,6497***	1,3921
0.3	1,0926	1,2667	1,012	1,4344	1,141
0.4	1,2148	1,156	1,065	1,2169	1,05
0.5	1,4103	1,181	1,183	1,3499	1,0158
0.6	1,182	1,0007	1,0807	1,4067	1,609
0.7	1,025	1,0104	1,0726	1,2426	1,0683
0.8	1,1415	1,163	***1,3019***	1,036	1,16
0.9	1,0711	1,1426	1,025	1,0346	1,1419

Table 22: Simulation results for the ratio Z, for different values of the correlation coefficient and κ=50 (n=100).

Correl. Coeff.	Sample 1 Z	Sample 2 Z	Sample 3 Z	Sample 4 Z	Sample 5 Z
0.0	1.05365	1.08108	1.13852	1.0164	1.3273
0.1	1.1598	1.1167	1.08145	1.2067	1.0457
0.2	**1.488**	**1.499**	**1.6535**	**1.6743**	1.11948
0.3	1.0257	1.13774	1.15945	1.0817	1.18346
0.4	1.048	1.269	1.151	1.3363	1.199
0.5	1.1328	1.1287	1.324	**1.459**	1.0522
0.6	1.0427	1.1062	1.0261	1.0888	1.019
0.7	1.0401	1.1789	1.092	1.205	1.188
0.8	1.126	1.079	1.046	1.172	1.023
0.9	1.058	1.093	1.0457	1.133	1.152

The following table summarizes the relative frequency of rejecting the null hypothesis of the equivalence of two linear models for the different values of the sample size and the correlation coefficient used in the simulation study.

Table 23: Summary table of the simulation results.

Corr.Coeff \κ	κ=12	κ=30	κ=50	Total
0.0	1/5	2/5	0/5	*3/15*
0.1	1/5	0/5	0/5	*1/15*
0.2	2/5	1/5	4/5	*7/15*
0.3	0/5	0/5	0/5	*0/15*
0.4	1/5	0/5	0/5	*1/15*
0.5	0/5	0/5	1/5	*1/15*
0.6	1/5	0/5	0/5	*1/15*
0.7	1/5	0/5	0/5	*1/15*
0.8	0/5	1/5	0/5	*1/15*
0.9	1/5	0/5	0/5	*1/15*
Total	*8/50*	*4/50*	*5/50*	*17/150*