Increasing Reliability of the F-Test in the Gauss-Markov Model when Outliers are Small

Şerif Hekimoğlu

Abstract
There are two populations of observations that have the same variance and are independent. Assume that one population includes outliers, whereas the other one does not. For testing the equality of the variances of two populations, the F-test is used. To measure the reliability of a test, the minimum mean success rate (minimum MSR) was introduced. The minimum MSRs of the F-test in the Gauss-Markov models are small when outliers are small. To increase the MSRs of the F-test, we propose a new F-test where the weights of all the observations in one sample with outliers are multiplied by a certain positive number k, such as 1.75. This new F-test was tested on a linear regression by a simulation. A thousand samples were generated by means of normally distributed random errors. Random and influential outliers are considered in the tail regions and in the whole region of a sample. These outliers are randomly generated 500 times for each sample. Using the new F-test, the minimum MSRs of the F-test are increased on the average by 24 % for a simple regression and by 37 % for a multiple regression using a significance level of $\alpha = 0.05$ when the outliers lie between 3σ and 6σ.

Zusammenfassung
Es sollen zwei Gruppen von Beobachtungen vorliegen, in denen sämtliche Beobachtungen gleiche Varianzen besitzen und voneinander unabhänging sind. Es wird vorausgesetzt, dass eine Gruppe keine Ausreißer und die andere Gruppe Ausreißer besitzt. Der F-Test wird angewendet, um zu testen, ob die beiden Gruppen die gleichen Varianzen besitzen. Um die Zuverlässigkeit eines Tests zu messen, wurde eine minimale mittlere Erfogsrate (minimale MER) eingeführt. Die minimale MER des F-Tests für kleine Ausreißer ist klein. Um sie zu erhöhen, wird ein neuer F-Test empfohlen, bei dem die Gewichte aller Beobachtungen in einer Gruppe, die Ausreißer enthält, mit einer positiven Anzahl k, zum Beispiel 1.75, multipliziert werden. Dieser neue Test wurde in einer linearen Regression durch eine Simulation erprobt. Tausend Stichproben mit normal verteilten Zufallsfehlern wurden erzeugt. Die zufälligen und einflussreichen Ausreißer wurden am Rande und in der Mitte einer Stichprobe angenommen. Diese Ausreißer wurden zufällig fünfhundertmal für jede Stichprobe erzeugt. Die minimale MER der beiden Tests erhöhte sich ungefähr um 24 % für die einfache und um 37 % für die mehrfache Regression bei einem Signifikanzniveau von $\alpha = 0.05$, wenn die Ausreißer zwischen 3σ and 6σ liegen.

1 Introduction

In order to measure the global reliability of a test procedure in robust statistics, the concept of the breakdown point, especially the power and level breakdown points are used (Ylvisaker 1977, Hampel et al. 1986, He et al. 1990, He 1991, Markatou and He 1994). The power breakdown point gives the amount of contamination that can drive the test statistic to its null value regardless of the true alternative value. The level breakdown point shows the amount of contamination that carries the test statistic to any value in the alternative space.

The F-test is extremely sensitive to distributions that are not normally distributed (Triola 2001). The F-test of a hypothesis, comparing two sample variances from the same distribution, is known to be nonrobust (Shorack 1969; Markatou and He 1994; He et al. 1990). This F-test is used mostly in geodesy as the global congruency test in the detection of deformation (Pelzer 1971).

In order to measure the reliability of a test for outliers, the minimum mean success rate (minimum MSR) was introduced in a given outlier interval for a certain number of outliers by Hekimoğlu and Koch (2000). Applied for this study, it means that the number of successful rejections of the null hypothesis in case of outliers is divided by the number of experiments. The minimum MSR may be interpreted as a finite sample version of the power breakdown point of a test. The minimum MSR gives more information about the reliability of a test for outliers than the breakdown point does, especially if the observations have small outliers.

Let two populations of observations be given that have the same variance and are independent. Assume that one population includes outliers, whereas the other one does not. Due to the spreading effect of the least squares estimation (LSE), the magnitude of an outlier is not completely reflected in the corresponding residual and in the estimated variance of unit weight. This is much more valid if the corresponding sample includes multiple outliers. Therefore, the F-test fails for some cases to distinguish the null hypothesis H_0 from the alternative hypothesis H_1 when only one of the two samples includes outliers. Hence, we expect that the minimum MSRs of the F-test will also be relatively small when outliers are small. In this case, we will investigate how the minimum MSRs of the F-test can be increased assuming that only one of the two independent samples with the same variances is contaminated by outliers. We argue that the estimated variance for the sample that includes outliers will be increased when the weights of all observations of this contaminated sample are multiplied by any positive number k. Thus, the null hypothesis will be rejected more successfully than before. This F-test is called here a new F-test where the weights of all observations from the contaminated sample are multiplied by any positive number k ($k > 1$).

2 Outlier concept

The outlier concept is the same as given in Hekimoğlu (1997). The one-dimensional bad observations which lie between $-\infty$ and $\mu - 3\sigma$ or between $\mu + 3\sigma$ and $+\infty$ are called outliers with σ being the standard deviations of the observations. They can take on any value from this space which is the outlier region. In this study, each group of outliers is divided into two broad categories, the *random* and *jointly influential* outliers. The outliers in each category can be further divided into the categories small and large. The small outliers lie between 3σ and 6σ and the large outliers between 6σ and 10σ.

3 Test of equality of the variances of two populations

There are two samples, consisting of the multidimensional observations $\mathbf{l}_1^T = [l_{11}, l_{12}, \ldots, l_{1n1}]$ and $\mathbf{l}_2^T = [l_{21}, l_{22}, l_{23}, \ldots, l_{2n2}]$. The assumption is that the sample $\mathbf{l}_1$ comes from the normally distributed population $N(\mathbf{A}_1\mathbf{x}, \sigma_1^2\mathbf{I}_1)$ and the sample $\mathbf{l}_2$ from the other normally distributed population $N(\mathbf{A}_2\mathbf{x}, \sigma_2^2\mathbf{I}_2)$. We assume that both populations are independent and have the same variance, i. e. $\sigma_1^2 = \sigma_2^2$. In addition, we assume that the sample $\mathbf{l}_2$ is contaminated by outliers, whereas the sample $\mathbf{l}_1$ is not. Furthermore, the variances of both populations are unknown. Considering the Gauss-Markov model for both samples, we can write

$$\mathbf{l}_1 = \mathbf{A}_1\mathbf{x} + \mathbf{e}_1, \text{ with } \mathbf{C}_1 = \sigma_1^2\mathbf{I}_1, \tag{1}$$

$$\mathbf{l}_2 = \mathbf{A}_2\mathbf{x} + \mathbf{e}_2, \text{ with } \mathbf{C}_2 = \sigma_2^2\mathbf{I}_2, \tag{2}$$

$$E(\mathbf{e}_1) = 0 \text{ and } (\mathbf{e}_2 E) = 0, \tag{3}$$

where the index 1 or 2 refers to the sample 1 or 2, respectively. $\mathbf{A}_1$ is the $n_1 \times 1$ design matrix, $\mathbf{x}$ is the $u \times 1$ unknown parameter vector, $\mathbf{l}_1$ is the $n_1 \times 1$ observation vector, $\mathbf{e}_1$ is the $n_1 \times 1$ random error vector assumed to be normally distributed, $\mathbf{C}_1$ is the $n_1 \times n_1$ covariance matrix of the observations $\mathbf{l}_1$, σ_1^2 is the variance of unit weight of the sample $\mathbf{l}_1$, u is the number of unknowns, n_1 is the number of observations, $\mathbf{I}_1$ is the $n_1 \times n_1$ unit matrix and E( ) is the expected value. Let $\mathbf{A}_1$ have full column rank, i. e. rank $\mathbf{A}_1 = u$.

The estimated value $\hat{\sigma}_1^2$ of the variance σ_1^2 of unit weight from the sample $\mathbf{l}_1$ and the estimated value $\hat{\sigma}_2^2$ of the variance σ_2^2 of unit weight from the sample $\mathbf{l}_2$ is given respectively by

$$\hat{\sigma}_1^2 = \frac{\mathbf{v}_1^T \mathbf{v}_1}{n_1 - u}, \tag{4}$$

$$\hat{\sigma}_2^2 = \frac{\mathbf{v}_2^T \mathbf{v}_2}{n_2 - u}, \tag{5}$$

where $\mathbf{v}_1$ is the $n_1 \times 1$ residual vector and $\mathbf{v}_2$ is the $n_2 \times 1$ residual vector.

To test the equality of the variances of two populations, the F-test is used as follows:

$$H_0\colon \sigma_2^2 = \sigma_1^2, \tag{6}$$

$$H_1\colon \sigma_2^2 > \sigma_1^2. \tag{7}$$

The test statistic is

$$T = \frac{(n_2 - u)\hat{\sigma}_2^2}{(n_1 - u)\hat{\sigma}_1^2} \sim F_{(1-\alpha, n_2-u, n_1-u)}, \tag{8}$$

where n_2 is the number of observations in the sample $\mathbf{l}_2$. If $T > F_{(1-\alpha, n_2-u, n1-u)}$, the null hypothesis is rejected at the level of significance α.

4 Multiplying the weights of all the observations in one sample

Theorem: If the weights p_i, $i \in (1, 2, \ldots, n)$, of all observations are multiplied by any positive number k, the estimated variance $(\hat{\sigma}')^2$ of unit weight is multiplied by k as well. If $k > 1$, the new variance $(\hat{\sigma}')^2$ becomes always greater than $\hat{\sigma}^2$, i. e. $(\hat{\sigma}')^2 > \hat{\sigma}^2$. If $0 < k < 1$, the new variance $(\hat{\sigma}')^2$ becomes always smaller than $\hat{\sigma}^2$, i. e. $(\hat{\sigma}')^2 < \hat{\sigma}^2$.

Proof of the theorem: If the weights p_i, $\{i \in (1, 2, \ldots, n)\}$ of all observations are multiplied by a positive number k ($k > 0$), i. e. $p_i' = k\,p_i$, all the results of LSE are given easily as follows:

$$\mathbf{P}' = k\,\mathbf{P}, \tag{9}$$

$$\mathbf{x}' = \mathbf{x}, \tag{10}$$

$$\mathbf{v}' = \mathbf{v}, \tag{11}$$

$$(\mathbf{v}')^T\mathbf{P}'\mathbf{v}' = k\,\mathbf{v}^T\mathbf{P}\mathbf{v}, \tag{12}$$

$$(\hat{\sigma}')^2 = k\hat{\sigma}^2, \tag{13}$$

where $\mathbf{P}'$, $\mathbf{x}'$, $\mathbf{v}'$ and $(\hat{\sigma}')^2$ denote the new weight matrix, the new unknown vector, the new residual vector and the new variance respectively. Thus, we can see easily that if $k > 1$, the new variance $(\hat{\sigma}')^2$ becomes always greater than $\hat{\sigma}^2$, i. e. $(\hat{\sigma}')^2 > \hat{\sigma}^2$. If $0 < k < 1$, the new variance $(\hat{\sigma}')^2$ becomes always smaller than $\hat{\sigma}^2$, i. e. $(\hat{\sigma}')^2 < \hat{\sigma}^2$.

Now this theorem may be applied to the two samples given in (1) and (2) respectively. Since we assume that all the observations of two samples $\mathbf{l}_1$ and $\mathbf{l}_2$ have the same variance, the weight matrix $\mathbf{P}$ in (9) and (12) is replaced by the unit matrix $\mathbf{I}$. If the weights of all the observations in the sample $\mathbf{l}_2$ are multiplied by a positive number k $(k > 1)$, the new variance $(\hat{\sigma}_2')^2$ from LSE becomes always greater than $\hat{\sigma}_2^2$, i. e. $(\hat{\sigma}_2')^2 > \hat{\sigma}_2^2$. Hence, the new test statistic T′ is always bigger than T in (8) as follows:

$$T' = \frac{(n_2 - u)(\hat{\sigma}_2')^2}{(n_1 - u)\hat{\sigma}_1^2} \sim F_{(1-\alpha, n_2-u, n_1-u)} \,. \tag{14}$$

The null hypothesis will be rejected more successfully than before due to the increase of the variance $(\hat{\sigma}_2')^2$. Thus, we argue that the minimum MSRs of the F-test can be increased.

5 Monte-Carlo Method

5.1 Observations without outliers

To investigate increasing of the reliability of the F-test, two linear regressions are chosen. For the F-test, two samples $\mathbf{l}_1$ and $\mathbf{l}_2$ as defined by (1) and (2) are simulated with choosing $n_1 = n_2$. We assume that only the sample $\mathbf{l}_2$ is contaminated by outliers and the sample $\mathbf{l}_1$ is not. For this purpose, a simple straight line

$y_i = a_0 + a_1 x_i$, $i = 1, 2, 3, \ldots, n_y$
with $a_0 = 1$, $a_1 = 1$, $n_y = 10$,

and a multiple regression are chosen:

$z_i = a_0 + a_1 x_{1i} + a_2 x_{2i} + a_3 x_{3i} + a_4 x_{4i}$, $i = 1, 2, 3, \ldots, n_z$
with
$a_0 = 2$, $a_1 = -1$, $a_2 = 0.5$, $a_3 = 1.2$, $a_4 = 1.5$ and $n_z = 13$.

The random errors e^y_{1i}, $i = 1, 2, \ldots, n_{1y}$ and e^y_{2i}, $i = 1, 2, \ldots, n_{2y}$ for the simple regression were generated from the normal distribution $\mathbf{e} \sim N(\mu = 0,\ \sigma^2 = 4\ \text{cm}^2)$ by a random number generator of the IMLS subroutine, and also e^z_{1i}, $i = 1, 2, \ldots, n_{1z}$ and e^z_{2i}, $i = 1, 2, \ldots, n_{2z}$ for the multiple regression, where $n_{1y} = n_{2y} = n_y$ and $n_{1z} = n_{2z} = n_z$. These random errors are regarded as observation errors.

In order to obtain two samples, the corresponding random errors e^y_{1i}, e^y_{2i} and e^z_{1i}, e^z_{2i} are added to the y_i, y_j and z_i, z_j values as follows:

for the simple regression

$$l_{1i} = y_i + e^y_{1i}, \quad i = 1, 2, \ldots, n_{1y}\,, \tag{15a}$$

$$l_{2j} = y_j + e^y_{2j}, \quad j = 1, 2, \ldots, n_{2y}\,, \tag{15b}$$

and also for the multiple regression

$$l_{1i} = z_i + e^z_{1i}, \quad i = 1, 2, \ldots, n_{1z}\,, \tag{16a}$$

$$l_{2j} = z_j + e^z_{2j}, \quad j = 1, 2, \ldots, n_{2z}\,. \tag{16b}$$

One thousand samples for the simple and multiple regressions were generated separately.

5.2 Bad Observations

A contaminated sample contains a few bad observations in the second sample $\mathbf{l}_2$. To simulate a »bad« observation $\bar{l}_{2i}$, the random error of an observation is replaced by an outlier δy. This means that the magnitude δy of an outlier is added to the y_i- and z_j-values, e. g., $\bar{l}_{2i} = y_i + \delta y_i$ and $\bar{l}_{2j} = z_j + \delta y_j$.

Outliers are divided into two broad categories, random and influential. The number of outliers is denoted by m.

a) Random Outliers:
The magnitude δy of one random outlier (i. e. $m = 1$) is generated by a uniform distribution for a given interval $\text{int}(\sigma)$ in the outlier region as follows:

$\text{int}(\sigma) = 3\sigma < \delta y < 6\sigma$ for each sample of simple and multiple regressions,

$$\delta y_k = \text{sign}(t_1)\ \delta\bar{y}_k\,, \tag{17}$$

$$sign(t_1) = \begin{Bmatrix} + & t_1 > 0.5 \\ - & t_1 \le 0.5 \end{Bmatrix}, \quad 0 < t_1 \le 1\,, \tag{18}$$

$$\delta\bar{y}_k = 3\sigma + t_1\Delta,$$
$$k = n_3\, t_2, \quad \Delta = 6\sigma - 3\sigma = 3\sigma, \quad 0 < t_2 \le 1, \tag{19}$$

where t_1 and t_2 are distributed uniformly, and Δ is the length of the outlier interval $\text{int}(\sigma)$. In addition, n_3 becomes n_{2y} for (15b), and n_{2z} for (16b) respectively.

The interval $\text{int}(\sigma)$ for small outliers lies between 3σ and 6σ and for large outliers between 6σ and 10σ. In the last case (19) is changed to

$$\delta\overline{y}_k = 6\sigma + t_1\Delta,$$
$$k = n_3\, t_2, \quad \Delta = 10\sigma - 6\sigma = 4\sigma, \quad 0 < t_2 \le 1. \tag{20}$$

The magnitudes δy_k and δy_j of two random outliers (i. e. m = 2), the magnitudes δy_k, δy_j and δy_t of three random outliers, i. e. m = 3, and the magnitudes δy_k, δy_j, δy_t and δy_q of four random outliers, i. e. m = 4, are generated correspondingly by the uniform distribution for a given interval int(σ) in the outlier region as shown in Hekimoğlu and Koch (2000).

This algorithm has been computed 500 times for each sample of the simple and multiple regressions respectively. If the test statistic T can distinguish H_0 from H_1 hypothesis, the new F-test is considered as successful. The success rate of a contaminated working sample is obtained by dividing the total number of these successful cases by 500.

One thousand different contaminated samples have been simulated for the simple and multiple regressions respectively. Thus, the MSRs of the new F-test can be estimated as a mean value from these 1000 different success rates for random outliers or for their subkinds. It means that a MSR value is a mean value of 1000×500 different experiments.

b) Influential Outliers:
The magnitude δy of the influential outlier is also generated by the uniform distribution for a given interval in outlier region as done for random outliers. However, in this case they all have the same sign, i. e., all the plus or all the minus.

Outliers of each category are divided again into two subcategories as follows:

- Outliers are randomly distributed in the whole region of observations,
- outliers are randomly distributed only in the tail regions of observations.

5.3 The estimation of the MSRs of the F-test

First, the F-test was applied to 1000 samples without outliers in order to verify whether the F-test accepts the null hypothesis. The results are given in the second row under heading »0« in Tab. 1. According to these results, the F-test accepts the null hypothesis for $\alpha = 0.05$ with the MSR of 92 % for the simple regression and with the MSR of 92 % for the multiple regression. Thus, the risk of rejecting H_0 is 8 % for the simple regression and 8 % for the multiple regression although the sample does not include any outlier.

Secondly, the F-test was applied to 1000 contaminated samples where only the sample $\mathbf{l}_2$ is contaminated. The MSRs of the F-test are computed for different numbers of outliers, for the two subkinds of random and influential outliers, for the levels of significance $\alpha = 0.05$ and $\alpha = 0.01$, and for the intervals of $3\sigma - 6\sigma$, and $6\sigma - 10\sigma$ for the simple and multiple regressions. The MSRs increase rapidly as the number of outliers increases. The MSRs for the influential outliers are smaller than the ones for random outliers. The MSRs for the tail region of the observations are smaller than the MSRs for the whole region of the observations. The greater the number of unknowns, the smaller the MSRs. In addition, the MSRs for $\alpha = 0.05$ are greater than ones for $\alpha = 0.01$. Only the minimum MSRs of the F-test are given in Tab. 1.

Tab. 1: Minimum mean success rates of the F-Test

Number of outliers	Magnitude of outliers	For simple regression		For multiple regression	
		$\alpha = 0.05$	$\alpha = 0.01$	$\alpha = 0.05$	$\alpha = 0.01$
0		92 %	98 %	92 %	98 %
1	$3\sigma - 6\sigma$	44	17	35	12
2		72	34	60	25
3		86	53	75	40
4		90	60	82	49
Total		292	164	252	126
1	$6\sigma - 10\sigma$	93	67	83	52
2		100	94	97	83
3		100	98	99	93
4		100	99	99	96
Total		393	358	378	324

5.4 The estimation of the MSRs of the new F-test

First, the weights of all the observations in the sample $\mathbf{l}_2$ are multiplied by $k = 1.75$. Then, the F-test is applied to the contaminated sample. This is called the new F-test.

Secondly, the new F-test was applied to the 1000 samples without outliers in order to verify whether the new F-test accepts the null hypothesis. The results are given in the second row under heading »0« in Tab. 2. The new F-test accepts the null hypothesis for $\alpha = 0.05$ with the MSR of 83 % for the simple regression and with the MSR of 83 % for the multiple regression. Thus, the risk of rejecting H_0 is 17 % for the simple regression and 17 % for the multiple regression although the sample does not include any outlier. These risks are greater than the ones of the F-test given in Tab. 1.

Thirdly, the new F-test was applied to the 1000 contaminated samples where only the sample $\mathbf{l}_2$ is contaminated. The MSRs of the new F-test are computed for different numbers of outliers, for the two subkinds of random and influential outliers, for $\alpha = 0.05$ and $\alpha = 0.01$, and for the intervals of $3\sigma - 6\sigma$ and $6\sigma - 10\sigma$ for the simple and multiple regressions. In addition, the MSRs are obtained for different numbers of outliers for the given intervals of $3\sigma - 6\sigma$ and $6\sigma - 10\sigma$. The minimum MSRs are given in Tab. 2.

Tab. 2: Minimum mean success rates of the new F-test with k = 1.75

Number of outliers	Magnitude of outliers	For simple regression		For multiple regression	
		$\alpha = 0.05$	$\alpha = 0.01$	$\alpha = 0.05$	$\alpha = 0.01$
0		83 %	96 %	83 %	96 %
1	$3\sigma - 6\sigma$	78	44	67	35
2		95	72	88	60
3		98	86	94	74
4		99	90	96	81
Total		370	292	345	250
1	$6\sigma - 10\sigma$	100	93	97	83
2		100	100	100	97
3		100	100	100	99
4		100	100	100	99
Total		400	393	397	378

For the simple regression: If comparing the results in Tab. 2 with the ones in Tab. 1, the minimum MSRs of the new F-test procedure increase on the average by 24 % for the small outliers that lie between $3\sigma - 6\sigma$ with $\alpha = 0.05$ and by 78 % with $\alpha = 0.01$. They increase slightly for large outliers.

For the multiple regression: If comparing the results in Tab. 2 with the ones of in Tab. 1, the minimum MSRs of the new F-test procedure increase on the average by 37 % for the small outliers with $\alpha = 0.05$ and by 100 % with $\alpha = 0.01$. They increase slightly for large outliers.

5.5 Discussion

How can the number k be chosen to increase the MSRs of the new F-test? To answer this question we have investigated the change of MSRs for the simple and multiple regression when the number k changes. For this purpose, firstly, the new F-test was applied to 1000 samples including only observations without outliers, secondly, it was applied to 1000 contaminated samples where only the sample $\mathbf{l}_2$ is contaminated, as done in the subsection 5.4. The MSRs for different k-values for $\alpha = 0.05$ are given in Tab. 3 for the simple regression and in Tab. 4 for the multiple regressions when random outliers lie between 3σ and 6σ. The greater the number k, the greater the MSRs. If assuming that only the sample $\mathbf{l}_2$ is contaminated, there is no problem to increase the MSRs up to 100 %. However, the greater the number k, the smaller the MSRs in case that the sample does not include any outlier. It means that the risk of rejecting H_0 increases as the number k increases when the sample does not include any outlier. Therefore, we should find a compromise between both situations. In this study, we choose $k = 1.75$. In this case, the risk of rejecting H_0 is increasing by 9 % (i. e., $0.92 - 0.83 = 0.09$) for the simple regression with $\alpha = 0.05$ while the MSR of the new F-test is increasing from 44 % to 78 % for one outlier. It is increasing by 9 % (i. e., $0.92 - 0.83 = 0.09$) for the multiple regression with $\alpha = 0.05$ while the MSR of the new F-test for one outlier increases from 35 % to 67 %.

Tab. 3: MSRs of the new F-test for different number k for the simple regression

m^* \ k	1.00	1.25	1.50	1.75	2.00	2.25	2.50	2.75	3.00	5.00
0	92 %	90 %	87 %	83 %	78 %	73 %	68 %	64 %	60 %	29 %
1	44	58	69	78	84	88	91	94	95	100
2	83	92	96	98	99	99	100	100	100	100

* m is the number of outliers

Tab. 4: MSRs of the new F-test for different number k for the multiple regression

m^* \ k	1.00	1.25	1.50	1.75	2.00	2.25	2.50	2.75	3.00	5.00
0	92 %	92 %	88 %	83 %	78 %	74 %	68 %	65 %	60 %	29 %
1	35	48	58	67	74	80	84	87	90	98
2	68	80	87	92	95	96	98	98	99	100

* m is the number of outliers

May the new F-test be applied to the samples $\mathbf{l}_1$ and $\mathbf{l}_2$ when they have the same weight matrix $\mathbf{P}$, i. e. $\mathbf{P}_1 = \mathbf{P}_2$ for $n_1 = n_2$? In this case, $\mathbf{C}_1 = \sigma_1^2\mathbf{P}^{-1}$, $\mathbf{C}_2 = \sigma_2^2\mathbf{P}^{-1}$. The F-test may be used to test that both samples have the same variance factor:

$$H_0:\ \sigma_2^2 = \sigma_1^2, \tag{21}$$

$$H_1:\ \sigma_2^2 > \sigma_1^2. \tag{22}$$

For this purpose, 1000 samples for the simple and multiple regressions are used as done in subsection 5.4. The $\mathbf{P}$ matrix is given as $\mathbf{P} = \mathrm{diag}(1.0, 1.3, 1.9, 2.3, 1.5, 1.9, 1.2, 2.1, 1.7, 1.4)$ for the simple regression and $\mathbf{P} = \mathrm{diag}(0.9, 1.2, 1.9, 1.5, 2.5, 2.1, 1.9, 0.8, 2.0, 1.5, 1.0, 1.8, 1.6)$ for the multiple regression. The algorithm given in subsection 5.2 for outliers is computed 500 times for each sample as done in subsection 5.4. Only the MSRs of the simple regression are given in Tab. 5. It shows that the new F-test may also be used for this general case.

Tab. 5: Minimum MSRs of the F-test and the new F-test with k = 1.75 for the simple regression

Number of outliers	Magnitude of outliers	The F-Test		The new F-Test	
		$\alpha = 0.05$	$\alpha = 0.01$	$\alpha = 0.05$	$\alpha = 0.01$
0		99 %	100 %	98 %	99 %
1	$3\sigma - 6\sigma$	48	22	77	48
2		60	34	88	62
3		85	58	97	85
4		86	59	97	86
Total		279	173	359	281
1	$6\sigma - 10\sigma$	90	67	98	90
2		98	76	100	98
3		100	97	100	100
4		100	98	100	100
Total		388	338	398	388

6 Conclusion

The MSRs of the F-test have been obtained for simple and multiple regressions. They are relatively small especially for the outliers whose magnitudes lie between 3σ and 6σ. Therefore, the F-test is not very sensitive against small outliers.

In this paper, we have proved that if the weights of all the observations are multiplied by a positive number k, the corresponding estimated variance $(\hat{\sigma}')^2$ in the Gauss-Markov model increases always, i. e. $(\hat{\sigma}')^2 > \hat{\sigma}^2$, if $k > 1$. If $0 < k < 1$, it becomes always smaller, i. e. $(\hat{\sigma}')^2 < \hat{\sigma}^2$.

Using this information we have developed a new F-test for the Gauss-Markov model in order to increase the minimum MSRs (the reliability) of the F-test when the outliers are small. We assume that only one of two different samples is contaminated. This new approach is based on multiplying the weights of all the observations in the contaminated sample by a certain number k. An optimal number k was found to be 1.75. We have numerically shown how the minimum MSRs of the F-test can be increased significantly for different kinds of outliers for a given interval and for a certain number of outliers when the outliers are small.

Acknowledgements

The author is grateful to Professor K. R. Koch for the English corrections and helpful comments and to Res. Asst. Cuneyt Aydin and Res. Asst. R. Cuneyt Erenoglu for the proof-reading.

References

Hampel, F., Ronchetti, E., Rousseeuw, P., and Stahel, W.: Robust statistics: the approach based on influence functions. John Wiley & Sons, Inc. New York, N.Y., 1986.

He, X.: A local breakdown property of robust tests in linear regression. Journal of Multivariate Analysis, 38, 294–305, 1991.

He, X., Simpson, D. G., and Portnoy, S. L.: Breakdown robustness of tests. J. Am. Statistical. Assn., 85, 446–452, 1990.

Hekimoğlu, Ş.: The finite sample breakdown points of the conventional iterative outlier detection procedures, J. of Surv. Engrg., ASCE, 123(1), 15–31, 1997,

Hekimoğlu, Ş., Koch, K. R.: How can reliability of tests for outliers be measured? AVN, 107(7): 247–254, 2000.

Koch, K. R.: Parameter estimation and hypothesis testing in linear models. Springer-Verlag, New York, N.Y. 2nd Edition, 1999.

Markatou, M., He, X.: Bounded influence and high breakdown point testing procedures in linear models. J. Am. Statistical. Assn., 89: 543–49, 1994.

Pelzer, H.: Zur Analyse geodätischer Deformationsmessungen, D. Geod. Komm., Reihe C, Nr. 164, München, 1971.

Shorack, G. R.: Testing and estimating ratios of scale parameters, J. Am. Statistical. Assn., 64, 999–1013, 1969.

Ylvisaker, D.: Test resistance. J. Am. Statistical Assn., 72, 551–557, 1977.

Triola, M. F.: Elementary Statistics, eighth edition, Addison Wesley Longman, Inc. Boston, 2001

Authors' address:
Prof. Dr. Şerif Hekimoğlu
Yildiz Teknik Üniversitesi Insaat Fakültesi
Jeodezi ve Fotogrametri Müh. Böl.
Yildiz – Istanbul, 34349 Türkiye
Fax No: +90 212 2610767
hekim@yildiz.edu.tr