World Journal of Advanced Research and Reviews
eISSN: 2581-9615 CODEN (USA): WJARAI
Cross Ref DOI: 10.30574/wjarr
Journal homepage: https://wjarr.com/

(Research Article)

A study of the effect of outliers on regression inference and the performance of the t-test statistic using the standard deviation method

I.A. Baba [1,*], A. A. Mohammad [2], M. B. Mohammed [3], R. M. Madaki [1] and H. Ahmad [4]

[1] *Department of Statistics, Faculty of Science, Abubakar Tafawa Balewa University Bauchi, Bauchi Sate, Nigeria.*
[2] *Department of Mathematics, Faculty of Science, Bauchi State University Gadau, Nigeria.*
[3] *Department of Mathematics and Statistics, Faculty of Science, Federal University Kashre, Gombe, Gombe State.*
[4] *Department of Statistics and Operations Research, Modibbo Adama University of Technology, Yola, Nigeria.*

World Journal of Advanced Research and Reviews, 2026, 30(02), 2389-2397

Publication history: Received on 17 April 2026; revised on 24 May 2026; accepted on 26 May 2026

Article DOI: https://doi.org/10.30574/wjarr.2026.30.2.1494

Abstract

One of the essential requirements for smooth regression analysis based on ordinary least squares (OLS) method is normality of the data. When the dataset passes the normality test, it is virtually free from unusual observations, OLS method can then be applied effectively to obtain the required estimate of the regression parameters and make inference. It is quite obvious that presence of outlier distorts regression inference when non-robust methods are used. This article highlighted on the consequences of including outlier in the analysis of regression models based on the t test statistics. Standard deviation (SD) method was employed in measuring the effect of outlier on t test statistics taking into account the sample sizes and the number of independent variables at different intensity of outlier both using real examples and simulation study. It was discovered that outlier affect the estimate of regression coefficient negatively which in return render the classical regression estimators inefficient. In addition, it can alter the odds of making both type I and type II error as well as influence the estimate of regression that are of essential interest.

Keywords: Outlier; Regression; Standard Deviation; T Statistic

1. Introduction

Consider the general linear regression model is of the form

$$y = X\beta + e \qquad (1)$$

Where X is $n \times (p+1)$ matrix of predictor variables, y is a vector of response observations, β is a column vector of the unknown regression parameters, and e is a column vector of unobservable random errors defined as:

$$X = \begin{bmatrix} 1 & x_{11} & \cdots & x_{1p} \\ 1 & x_{21} & \cdots & x_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & x_{n1} & \cdots & x_{np} \end{bmatrix}, y = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}, \beta = \begin{bmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_{p+1} \end{bmatrix}, e = \begin{bmatrix} e_1 \\ e_2 \\ \vdots \\ e_n \end{bmatrix}$$

---

[*] Corresponding author: I.A. Baba

The solution of Equation (1) can be achieved by minimizing the difference between the estimated and observed response, which can be obtained using the least squares loss function defined by:

$$Q(\beta) = \sum_{i=1}^{n} \left(y - \hat{y}\right)^2 \tag{2}$$

When a data set contains outliers and other uncleanliness, the OLS function in Equation (2) produces corrupt estimates and may lead to incorrect inference (Al Rezami, 2020, Kowalskie, 2025).

Outliers arise for several reasons, such as data from heavy-tailed distribution functions or erroneous measurements (Verardi and Vermandele, 2016, Klebanov and Volchenkova, 2019). Outliers can cause significant damage in statistical data analysis. According to Dash et al. (2023), outliers cause an increase in variance and reduce the power of statistical tests. In addition, they can alter the odds of making both Type I and Type II errors, as well as influence the estimates of regression coefficients that are of essential interest. Therefore, detecting and deleting or resolving them before the analysis is essential to prevent skewing. Several techniques have emerged as a way of tackling this problem. Nkechinyere et al. (2015) compared some outlier labeling techniques, which include the Standard Deviation (SD) method, the MAD (Median Absolute Deviation), and the Median formula. Hodge and Austin (2004) conducted a survey on outlier detection methodologies and identified their advantages and disadvantages in data analysis. For details on outlier detection methods and their effects in data analysis, see Dodge (1997), Hawkins et al. (2002), and Leys et al. (2013).

There are commonly two types of hypothesis tests used in multiple linear regression analysis: the t-test and the F-test statistics. The t-test checks the significance of individual regression coefficients, while the F-test is used to simultaneously check the significance of multiple regression coefficients. The F-test can also be used to test individual coefficients. It is worth mentioning that this paper is proposed to enlighten researchers and practitioners on the consequences of outlying points and to show the practical benefit of conducting outlier diagnostic checks before conducting any statistical analysis. In this paper, the Standard Deviation (SD) method is used to measure the effect of outliers on the t-statistic, taking into account the sample sizes and the number of regressors at different outlier intensities.

As established by (Isobe et al., 1990, Burton, 2021), although several methods can be used to find the estimates of regression coefficients in a linear model, the Ordinary Least Squares (OLS) method is the most commonly used in practice due to its computational simplicity and elegant mathematical properties. Its status as the Best Linear Unbiased Estimator (BLUE) under the Gauss–Markov theorem makes it theoretically superior and a natural choice. This holds even though real-life data often show that perfectly satisfying all BLUE assumptions is unrealistic. Thus, the OLS estimator in matrix notation is defined as

$$\hat{\beta} = \left(X^T X\right)^{-1} Xy \tag{3}$$

where $\hat{\beta}$ is the $(p+1)\times 1$ vector of estimated regression coefficients. For this unique algebraic solution to exist, the cross-product matrix $X^T X$ must be non-singular. This fundamentally requires the design matrix X to have full column rank $(\text{rank}\ (\text{X}) = \text{p}+1)$ which mathematically precludes the presence of perfect multicollinearity among the explanatory variables. This presents another problem that could be solved in future work. A key gap in the literature is to investigate the simultaneous effect of outliers and multicollinearity in regression analysis. Some of the recent papers on this topic include (Lukman et al., 2024, Oyeleke et al., 2024, Wasim et al., 2025). Further study in this direction would provide valuable guidance for social science and public health researchers and practitioners who frequently encounter both issues in real-world data.

2. Statistics use for testing significance of regression coefficients

2.1 Procedure for Test of Significance Using F Test Statistic

The F -statistic is widely utilized to evaluate the overall statistical significance of a multiple linear regression model, testing whether all explanatory variables jointly exert a significant effect on the dependent variable. The following steps are used for implementation of the F test statistic.

Step 1

Set the hypothesis

$$H_0 : \beta_1 = \beta_2 = \cdots \beta_p = 0$$

Versus

$$H_1 : \text{at least one } \beta_j \neq 0$$

Step 2

Compute the F statistic

$$F = \frac{MSR}{MSE}$$

where MSR is the regression mean square and MSE is the error mean square.

Step 3

$$\text{Reject } H_0 \text{ if } F > F_{\alpha, p, (n-p-1)}$$

2.2 Procedure for Test of Significance Using t Test Statistic

Unlike the F statistics, t is widely utilized to evaluate the statistical significance of individual regression coefficients in multiple linear regression models (Seber and Lee, 2003). In model specification, incorporating a statistically significant variable generally improves model efficiency and explanatory power. Conversely, adding an irrelevant or insignificant variable does not yield additional explanatory benefits; instead, it reduces the model's degrees of freedom and inflates the variances of the remaining estimated coefficients (Kim et al., 2024)

To test the statistical significance of a specific regression coefficient β_j, a two-tailed hypothesis test is conducted using the following systematic steps:

Step 1: Formulate the Hypotheses Set the null hypothesis H_0 against the alternative hypothesis H_1:

$$H_0 : \beta_j = 0$$

$$H_1 : \beta_j \neq 0$$

Step 2: Compute the t-Statistic using the ratio of the estimated coefficient to its standard error:

$$t = \frac{\hat{\beta}_j}{se(\beta_j)} \tag{3}$$

where $\hat{\beta}_j$ is the estimated coefficient and $se(\beta_j)$ is its corresponding standard error.

Step 3: Define the decision rule. The null hypothesis H_0 is rejected at the chosen significance level α if the calculated t-statistic lies in the rejection region, that is, if $|t| > t_{\frac{\alpha}{2}}, (n-p)$ where $t_{\frac{\alpha}{2}}, (n-p)$ denotes the critical value from the Student's t-distribution with $n-p$ degrees of freedom

3. Numerical Examples and Simulation Study

In this section, two real-life examples and a simulation study are discussed. The first dataset was used in (Becker et al., 1988), while the second dataset was used in (Harrison Jr and Rubinfeld, 1978) to study issues related to outliers. The simulation procedure was based on six different regression models with varying sizes, consisting of one, two, three, four, and ten regressors. Observations for the explanatory variables were generated from the uniform distribution. The error terms were generated from the standard normal distribution with mean zero and variance one $N(0,1)$

The parameters of the model in each case were set such that each parameter was significant. A t-test was conducted on the last parameter in each case. Two levels of outliers (5% and 10%) were considered under different sample sizes and varying outlier intensities. Outlier intensity, in this context, refers to the degree or extent of deviation of the outlying observations from the normal mean. Six successive outlier intensities were considered: 5SD (5 standard deviations from the mean), 6SD, 7SD, 8SD, 9SD, and 10SD.

The results presented in the Tables 1-4 correspond to the power of the t-test (expressed in percentages), with each result based on 10,000 replications implemented in R. The power of a statistical test refers to the ability of a test statistic to reject the null hypothesis H_0 when it is false. In this study, H_0 is always false; therefore, the effect of outliers becomes apparent through the non-rejection of H0 or, equivalently, through a reduction in statistical power. This implies that the reduction in power is proportional to the severity of the outlier effect.

From the results obtained, it can be deduced that the negative effect of outliers on statistical inference depends on the following factors:

3.1 Number of outliers per sample

The greater the number of outliers relative to the sample size, the lower the resistance of the test statistic to the effect of outliers on statistical inference. For example, in the single regressor case, at $n = 40$ with 5% outliers at 8SD, 9SD, and 10SD, the powers are 86.19, 78.12, and 70.58, respectively. Under the same condition but with 10% outliers, the powers further reduced to 59.10, 50.69, and 43.52, respectively.

3.2 Intensity of the outlier

The farther the outlying observations deviate from the normal observations, the greater the effect of the outliers in distorting statistical inference.

3.3 Sample size

The test statistic exhibits greater resistance to the negative effect of outliers in large samples than in small samples. For example, in the two-regressor case at $n = 60$ with 10% outliers at 5SD, 6SD, and 7SD, the powers are 84.21, 72.31, and 59.90, respectively. Under the same conditions but at n = 80, the powers improved to 93.18, 84.07, and 72.97, respectively.

3.4 Number of regressors

The test statistic is more resistant to the negative effect of outliers when the regression model contains fewer regressors than when it contains more. This implies that smaller- sized regression models are more resistant to outliers in statistical inference than larger sized models. For example, in the two regressors case at $n = 60$ with 5% outliers at 7SD, 8SD, and 9SD, the powers are 86.22, 77.82, and 69.64, respectively. Under the same conditions but with three regressors, the powers reduced to 71.37, 62.00, and 52.68, respectively. Furthermore, two real datasets were also considered. In each dataset, two levels of outliers, approximately 5% and 10%, were introduced separately. The outliers were inserted in place of original observations to contaminate the data. At each contamination level, outlier intensities ranging from 5SD to 10SD were considered. In the first dataset (Freeny Data), three of the parameters were significant at the 95% confidence level (at least) in the uncontaminated data. However, under both the 5% and 10% contamination levels, none of the parameters remained significant across all outlier intensities considered. This clearly demonstrates the adverse effect of outliers on statistical inference.

In the second real dataset (Boston Housing Data), six parameters were significant at the 95% confidence level (at least) in the uncontaminated data. With 5% outliers introduced into the data, the numbers of significant parameters at the 90% confidence level (at least) were 5, 5, 3, 2, 2, and 2 for 5SD, 6SD, 7SD, 8SD, 9SD, and 10SD outlier intensities, respectively. Similarly, with 10% outliers in the data, the numbers of significant parameters at the 90% confidence level (at least) were 3, 1, 1, 1, 1, and 1 for 5SD, 6SD, 7SD, 8SD, 9SD, and 10SD outlier intensities, respectively.

---

4 Results and discussion

This section presents the simulation results for the power of the t-test statistic under varying sample sizes, percentages of outliers, outlier intensities, and numbers of regressors. The powers reported are expressed in percentages and are based on 10,000 replications implemented in R. The simulation study considered five regression structures consisting of one, two, three, four, and ten regressors. Two levels of contamination (5% and 10%) were introduced into the datasets, while six outlier intensities ranging from 5SD to 10SD were examined. The results in Tables 1-6 generally indicate that the presence of outliers substantially reduces the power of the t-test statistic. The reduction in power becomes more severe as the percentage and intensity of outliers increase. However, larger sample sizes tend to improve the resistance of the test statistic against the adverse effect of outliers.

4.1 Single Regressor

For the single regressor model, the power of the test remained relatively high in larger sample sizes despite contamination. However, the power consistently declined as the outlier intensity increased from 5SD to 10SD. The reduction was more severe under 10% contamination than under 5% contamination. For example, at $n = 40$ and 5% contamination, the powers decreased from 99.78 at 5SD to 70.58 at 10SD. Under the same sample size but with 10% contamination, the powers reduced further from 90.74 to 43.52.

4.2 Two Regressors

The two regressors model exhibited a greater reduction in power compared to the single regressor model. The effect of outliers became increasingly pronounced as both the contamination level and outlier intensity increased. At $n = 60$ with 5% contamination, the powers declined from 98.33 at 5SD to 62.85 at 10SD. Under 10% contamination, the corresponding powers further decreased from 84.21 to 35.48.

4.3 Three Regressors

Table 2 show that the three regressors model exhibited a marked decline in power as both outlier intensity and contamination increased. For example, at $n = 60$ with 5% contamination, the power dropped from 47.56% at 5SD to 19.24% at 10SD. Under 10% contamination at the same sample size, the power declined from 31.70% at 5SD to 12.72% at 10SD, indicating that moderate-sized models already suffer substantially from outliers when contamination is present.

4.4 Four Regressors

Table 3 show that the influence of outliers became even more pronounced, especially at higher sample sizes. At $n = 60$ with 5% contamination, the power decreased from 53.80% at 5SD to 23.14% at 10SD. Under 10% contamination at n=80, the corresponding powers fell from 37.94% at 5SD to 15.06% at 10SD. These results show that, as the number of regressors increases from three to four, the t-test becomes less resistant to both the proportion and intensity of outliers.

4.5 Ten Regressors

The results in Table 4 show that the ten regressors model exhibited the poorest resistance to outliers among all models considered. At $n = 100$ with 5% contamination, the power dropped from 47.70% at 5SD to 18.60% at 10SD. Under 10% contamination at the same sample size, power declined from 29.86% at 5SD to 11.52% at 10SD. Overall, the simulation results demonstrate that the power of the t-test decreases systematically as the percentage and intensity of outliers increase. Moreover, although larger sample sizes partially improve robustness, regression models with many regressors remain substantially more sensitive to contamination than smaller, lower-dimensional models.

4.6 Freeny Dataset

Table 5 show the results for the Freeny dataset. The dataset was analysed under uncontaminated conditions and under two contamination levels of approximately 5% and 10%. Outlier intensities ranging from 5SD to 10SD were considered. The uncontaminated data showed that three parameters were statistically significant at the 95% confidence level. However, after introducing outliers into the dataset, the statistical significance of the parameters deteriorated substantially. Under the 5% contamination level, the p-values generally increased as the outlier intensity increased from

5SD to 10SD. Similarly, under the 10% contamination level, the p-values became even larger, indicating a greater loss of statistical significance. The results clearly demonstrate that the presence of outliers severely affects the reliability of hypothesis testing and statistical inference in regression analysis.

4.7 Boston Housing Dataset

Table 6 display the results for Boston Housing dataset. This dataset was also analysed under uncontaminated conditions and under contamination levels of approximately 5% and 10%. In the uncontaminated dataset, six parameters were statistically significant at the 95% confidence level. However, the introduction of outliers considerably altered the regression results. Under the 5% contamination level, the number of significant parameters progressively reduced as the outlier intensity increased from 5SD to 10SD. A similar but more severe effect was observed under the 10% contamination level.

Table 1 Simulation results for power test of t statistic

Outlier Intensity	**N20(5)**	**N20(10)**	**N40(5)**	**N40(10)**	**N60(5)**	**N60(10)**	**N80(5)**	**N80(10)**	**N100(5)**	**N100(10)**
1Regressor										
5SD	82.87	61.32	99.78	90.74	100.00	98.50	100.00	99.83	100.00	99.98
6SD	72.29	49.29	97.42	80.34	99.92	94.17	100.00	98.21	100.00	99.60
7SD	63.59	40.16	92.79	68.65	99.03	86.47	99.91	94.55	100.00	98.14
8SD	56.04	33.74	86.19	59.10	96.88	77.11	99.41	88.07	99.95	94.75
9SD	49.30	28.69	78.12	50.69	93.06	67.54	97.90	81.16	99.44	89.07
10SD	45.50	25.03	70.58	43.52	87.25	58.63	95.26	73.07	98.35	81.92
2 Regressors										
5SD	62.74	39.79	90.36	66.75	98.33	84.21	99.80	93.18	99.97	97.36
6SD	52.64	31.26	80.42	53.27	93.88	72.31	98.39	84.07	99.70	91.03
7SD	45.28	24.70	70.02	43.31	86.22	59.90	94.76	72.97	98.20	81.36
8SD	37.48	20.89	60.78	35.24	77.82	50.02	89.09	60.72	94.71	71.72
9SD	34.11	17.44	51.91	29.55	69.64	41.64	82.28	52.90	89.14	62.60
10SD	28.28	15.11	44.98	25.35	62.85	35.48	73.21	45.67	82.91	53.12

Table 2 Three regressors

Outlier	**N20(5)**	**N20(10)**	**N40(5)**	**N40(10)**	**N60(5)**	**N60(10)**	**N80(5)**	**N80(10)**	**N100(5)**	**N100(10)**
5SD	16.10	10.98	32.24	22.48	47.56	31.70	58.36	40.62	70.02	49.42
6SD	12.90	9.04	25.46	16.88	39.38	24.76	46.90	31.78	59.62	39.78
7SD	11.82	8.00	22.48	14.18	31.14	20.14	40.64	26.08	49.78	31.86
8SD	10.66	8.16	17.92	11.76	26.72	16.40	33.44	22.46	41.48	26.30
9SD	8.54	7.44	16.70	10.86	22.24	14.72	28.66	18.08	35.06	21.48
10SD	8.08	6.44	13.92	10.10	19.24	12.72	25.14	15.86	29.16	18.94

Table 3 Four regressors

Outlier	N20(5)	N20(10)	N40(5)	N40(10)	N60(5)	N60(10)	N80(5)	N80(10)	N100(5)	N100(10)
5SD	14.62	9.46	29.20	19.46	44.64	29.08	53.80	37.94	65.04	46.62
6SD	12.52	9.00	23.64	15.76	34.66	23.02	44.68	29.26	55.38	36.72
7SD	10.72	8.60	19.58	13.48	29.18	19.28	36.96	23.06	44.90	28.86
8SD	8.92	7.50	15.80	11.10	24.18	15.26	31.18	19.42	37.54	24.48
9SD	8.02	7.02	15.10	9.94	20.86	13.50	26.04	16.18	32.52	20.04
10SD	6.34	7.42	13.92	9.12	17.82	11.70	23.14	15.06	27.94	17.40

Table 4 Ten regressors

Outlier	N20(5)	N20(10)	N40(5)	N40(10)	N60(5)	N60(10)	N80(5)	N80(10)	N100(5)	N100(10)
5SD	8.28	6.80	19.04	12.88	28.68	20.08	39.84	25.22	47.70	29.86
6SD	7.64	6.36	15.58	10.10	22.24	15.40	30.62	19.62	38.12	23.38
7SD	7.22	6.06	12.86	9.48	19.76	12.60	26.40	15.80	30.32	19.22
8SD	6.26	5.84	11.48	7.88	15.62	10.48	21.10	13.48	25.34	16.92
9SD	5.78	5.18	10.32	6.80	13.82	9.78	17.30	11.46	22.94	12.80
10SD	5.38	5.50	9.72	7.04	11.88	8.52	15.00	10.48	18.60	11.52

Table 5 Freeny dataset results

Outliers	5SD	6SD	7SD	8SD	9SD	10SD	5SD	6SD	7SD	8SD	9SD	10SD
0.091	0.735	0.756	0.773	0.786	0.796	0.805	0.641	0.661	0.675	0.686	0.695	0.702
0.390	0.796	0.801	0.805	0.808	0.810	0.812	0.901	0.912	0.920	0.927	0.932	0.936
0.000	0.788	0.800	0.810	0.817	0.823	0.828	0.910	0.922	0.930	0.936	0.941	0.946
0.000	0.862	0.846	0.834	0.825	0.817	0.811	0.622	0.600	0.585	0.573	0.565	0.558
0.013	0.673	0.673	0.691	0.705	0.717	0.727	0.577	0.598	0.614	0.627	0.637	0.645

Table 6 Boston dataset results

Outliers	5SD	6SD	7SD	8SD	9SD	10SD	5SD	6SD	7SD	8SD	9SD	10SD
0.020	0.333	0.428	0.505	0.577	0.640	0.694	0.129	0.164	0.196	0.2245	0.250	0.273
0.017	0.945	0.890	0.764	0.668	0.594	0.535	0.717	0.835	0.929	0.996	0.936	0.888
0.444	0.659	0.598	0.553	0.520	0.494	0.474	0.346	0.307	0.281	0.262	0.249	0.238
0.526	0.518	0.458	0.416	0.384	0.361	0.343	0.297	0.261	0.238	0.220	0.208	0.199
0.051	0.442	0.532	0.610	0.678	0.734	0.784	0.303	0.362	0.412	0.455	0.492	0.523
0.483	0.436	0.451	0.465	0.478	0.489	0.499	0.944	0.981	0.991	0.969	0.951	0.936
0.935	0.058	0.047	0.040	0.036	0.033	0.031	0.544	0.529	0.520	0.513	0.507	0.503
0.000	0.046	0.077	0.112	0.149	0.187	0.224	0.083	0.124	0.167	0.209	0.247	0.283
0.000	0.011	0.039	0.093	0.172	0.266	0.368	0.089	0.200	0.338	0.482	0.618	0.743

0.464	0.418	0.435	0.452	0.465	0.478	0.489	0.135	0.138	0.141	0.145	0.148	0.151
0.147	0.759	0.846	0.917	0.975	0.978	0.939	0.754	0.824	0.878	0.921	0.956	0.984
0.041	0.068	0.084	0.101	0.117	0.131	0.145	0.120	0.147	0.170	0.191	0.209	0.226
0.096	0.009	0.010	0.011	0.012	0.013	0.014	0.047	0.053	0.059	0.065	0.070	0.075
0.001	0.974	0.797	0.630	0.509	0.420	0.355	0.862	0.961	0.828	0.726	0.649	0.588

The results indicate that outliers substantially distort regression estimates, and statistical significance is highly sensitive to contamination. Furthermore, larger outlier intensities produce more severe effects, meaning that contamination can ultimately lead to misleading statistical conclusions.

In both the simulation study and the real-life examples, different outlier intensities generated using the standard deviation (SD) method were considered under varying sample sizes, numbers of regressors, and percentages of outliers introduced into each dataset. The study was conducted to investigate the consequences of outliers on regression analysis and the performance of the t-test statistic. The results demonstrate clearly that outliers affect regression estimates, reduce the power of hypothesis tests, and distort statistical inference. It was also observed that the power of the significance test depends on the percentage of outliers introduced into the data, the number of regressors in the model, the sample size, and the intensity of the outliers.

Furthermore, the findings indicate that larger sample sizes improve the resistance of statistical inference to outliers, while smaller regression models are generally more robust to con- termination than larger models. Additionally, increasing outlier intensity leads to a progressive deterioration in statistical power. It is therefore important to examine datasets for the presence of outliers before conducting statistical analysis in order to avoid misleading conclusions and the possible occurrence of Type I and or Type II errors.

5 Conclusion

This study systematically investigated the consequences of outlier contamination on linear regression analysis and the empirical performance of the t-test statistic. By evaluating various outlier intensities generated via the standard deviation (SD) method across both simulated environments and real-life datasets (the Freeny and Boston Housing datasets), this research mapped the precise boundaries where classical statistical inference fails.

The empirical and simulation results comprehensively demonstrate that the presence of outliers severely distorts parameter estimates, artificially inflates standard errors, and diminishes the statistical power of hypothesis tests. Crucially, the degradation of test power is highly dependent on four interlocking structural factors: the percentage of outliers introduced, the number of regressors within the model, the overall sample size, and the dimensional intensity of the anomalies. Furthermore, the findings indicate that while larger sample sizes offer a buffer that improves asymptotic resistance to contamination, high-dimensional models with numerous regressors exhibit a catastrophic sensitivity to outliers compared to more parsimonious, lower-dimensional alternatives. Unchecked contamination systematically drives calculated test statistics toward zero, escalating the probability of committing Type II errors (false negatives) while occasionally inducing spurious significance (Type I errors). Consequently, it is imperative that applied researchers rigorously audit datasets for anomalous points prior to implementing classical inference pipelines.

Building upon the findings of this study, several concrete avenues for future research are proposed to extend the robustness of linear statistical inference. First, while this study maps the failures of standard Ordinary Least Squares (OLS), future work should evaluate how high breakdown robust estimators such as Huber M estimation, S estimation, or MM-estimation preserve the power of the t-test under identical contamination profiles, specifically investigating these robust frameworks when outliers and severe multicollinearity are present. Second, given that power degrades exponentially as the number of regressors increases, future studies should focus on deriving mathematically adjusted critical values for the t and F statistics that dynamically adapt based on sample to regressors ratios (n/p) and robust diagnostic leverage scores. Third, a valuable extension would be to test whether these observed power-degradation patterns hold true in non-linear configurations, such as Logistic, Poisson, or Beta regression models, where outliers in the link function can induce even more severe structural bias.

Finally, future research could explore pairing classical diagnostic checks with unsupervised anomaly detection algorithms, such as Isolation Forests or Support Vector Machines, to prefilter and weight observations before executing hypothesis tests.

Compliance with ethical standards

Disclosure of conflict of interest

No conflict of interest to be disclosed.

References

[1] Rezami, A. (2020). Effect of outliers on the coefficient of determination in multiple regression analysis with the application on the gpa for student. International Journal of Advanced and Applied Sciences, 7(10):30–37.

[2] Becker, R., Chambers, J., and Wilks, A. (1988). The news language. wadsworth & brooks/cole. Computer Science Series, Pacific Grove, CA.

[3] Burton, A. L. (2021). Ols (linear) regression. The encyclopedia of research methods in criminology and Criminal Justice, 2:509–514.

[4] Dash, C. S. K., Behera, A. K., Dehuri, S., and Ghosh, A. (2023). An outlier's detection and elimination framework in classification task of data mining. Decision Analytics Journal, 6:100164.

[5] Dodge, Y. (1997). Lad regression for detecting outliers in response and explana- tory variables. Journal of multivariate analysis, 61(1):144–158.

[6] Harrison Jr, D. and Rubinfeld, D. L. (1978). Hedonic housing prices and the demand for clean air. Journal of environmental economics and management, 5(1):81–102.

[7] Hawkins, S., He, H., Williams, G., and Baxter, R. (2002). Outlier detec- tion using replicator neural networks. In International conference on data warehousing and knowledge discovery, pages 170–180. Springer.

[8] Hodge, V. and Austin, J. (2004). A survey of outlier detection methodologies. Artificial intelligence review, 22(2):85–126.

[9] Isobe, T., Feigelson, E. D., Akritas, M. G., and Babu, G. J. (1990). Linear regression in astronomy. Astrophysical Journal, Part 1 (ISSN 0004-637X), vol. 364, Nov. 20, 1990, p. 104-113. Research supported by NASA., 364:104–113.

[10] Kim, J., Kim, D. H., and Kwak, S. G. (2024). Comprehensive guidelines for appropriate statistical analysis methods in research. Korean Journal of Anesthesiology, 77(5):503–517.

[11] Klebanov, L. and Volchenkova, I. (2019). Outliers and the ostensibly heavy tails. Mathematical Methods of Statistics, 28(1):74–81.

[12] Kowalskie, A. (2025). The impact of outliers on linear regression models: detection and correction strategies. OTS Canadian Journal, 4(6):108–118.

[13] Leys, C., Ley, C., Klein, O., Bernard, P., and Licata, L. (2013). Detecting outliers: Do not use standard deviation around the mean, use absolute deviation around the median. Journal of experimental social psychology, 49(4):764–766.

[14] Lukman, A. F., Mohammed, S., Olaluwoye, O., and Farghali, R. A. (2024). Handling multicollinearity and outliers in logistic regression using the robust kibria– lukman estimator. Axioms, 14(1):19.

[15] Nkechinyere, E. M., Andrew, I., and Idochi, O. (2015). Comparison of different methods of outlier detection in univariate time series data. International Journal for Research in Mathematics and Statistics, 1(1):55–83.

[16] Oyeleke, K., Olatayo, T., and Efuwape, B. (2024). Handling multi- collinearity and outliers: acomparative study of some one and two–parameter estimators using real-life data. Int J Dev Math (IJDM), 1(4):177–190.

[17] Seber, G. A. and Lee, A. J. (2003). Linear regression analysis. John Wiley & Sons.

[18] Verardi, V. and Vermandele, C. (2016). Outlier identification for skewed and/or heavy-tailed unimodal multivariate distributions. Journal de la Soci´et´e Fran¸caise de Statistique, 157(2):90–114.

[19] Wasim, D., Zaman, Q., Ahmad, M., and Kibria, B. G. (2025). Mitigating multicollinearity and outliers in regression: comparison of some new and old robust ridge m-estimators. Journal of Statistical Computation and Simulation, 95(16):3526–3547.