Regression Analysis as Analytical Procedures: A Simulation Study of The Effect of Error Term Nonnormality on Auditor Decisions

Dr. Arlette C. Wilson, Accounting, Auburn University

Abstract

Regression analysis has been shown through research to be a useful audit tool. One underlying assumption of regression theory is error-term normality. Decision rules were applied to models developed using simulated data with normal and nonnormal error terms. Material errors were seeded into the audit period data, and incorrect rejections and acceptances were tabulated for both types of models. The results of this study suggest that nonnormality of error terms may not significantly affect auditor decisions.

Introduction

Statement on Auditing Standards No. 56 -- Analytical Procedures [AICPA, 1988] requires the use of analytical procedures in audits of financial statements. The basis for analytical procedures can range from subjective assessment of the reasonableness of account balances based on past information to the application of various statistical methods. Some auditors use statistical analysis as a more objective basis for deciding whether account balances are reasonable as Stringer (1975) notes:

I am convinced that regression analysis provides a more objective basis for performing analytical review, and thereby a means for reasonable quantification of the reliability that may be assigned to this class of substantive tests (p.4).

Regression analysis is currently being used to a limited extent in practice. Research has shown that regression analysis as an analytical procedure may be more efficient than traditional, nonstatistical procedures (Akresh and Wallace, 1980), that auditor judgment models compare favorably with automatic stepwise models (Neter, 1980), that regression-based decision rules are superior to nonstatistical approaches (Knechel, 1986), and in general, that regression is a useful audit tool.

One underlying assumption of regression theory that allows for valid interpretation of this confidence interval is that the error terms are normally distributed, which may be expressed as follows:

$$e \sim N(\mu, V)$$

All accounting data series which auditors analyze may not result in regression models with normally distributed error terms; however, the upper precision limit of the decision rule is still calculated using the normal variate. The purpose of this study is to evaluate the effect on auditor decisions when using regression models with nonnormal error terms to construct regression-based decision rules. Two 100 data series are simulated--one with nonnormal error terms and a corresponding one with normal error terms. Material and fraction of material errors are randomly generated into the audit period months. The number of incorrect rejections and incorrect acceptances using the STAR decision rule are tabulated and compared for the corresponding models.

The Decision Rule

The Statistical Technique for Analytical Review (STAR) regression-based decision model is used to evaluate the relative performance of regression models with and without error term normality. The simple linear regression model is

$$Y = \alpha + \beta + e$$

where Y is the audited balance of the account under audit consideration, X is the amount of a related financial or nonfinancial variable for the same month, α and ß are regression coefficients relating X to Y, and e is the residual or unexplained part of Y.

The parameters α and ß are estimated by applying ordinary least squares to past observations of X and Y, generally 36-monthly observations referred to as base period months. Given that this same relationship is expected to continue for the next 12 months, the auditor can calculate the expected amount for this account in the audit period as:

$$Y = A + B X$$

where A and B are the estimates for the parameters α and ß. Error terms for each of the audit period months are calculated as:

$$m = Y(b) - Y$$

where Y(b) is the book amount. An upper precision limit (UPL) which is used to signal accounts for investigation is calculated as:

$$UPL = m + t \cdot \sigma(m)$$

The t-value is based upon degrees of freedom (n-2) and the risk level desired by the auditor. Assuming the use of a 36-month base period and a desired incorrect acceptance risk level of A, a set of (1-A) upper precision limits on each m is calculated as:

$$[1] \quad UPL_{(1-A)}(e, n) = m + t_{(1-n\sqrt{A}),34} \cdot \sigma_m$$

for n=1,2,...,12

The standard error of prediction of m for each audit period month is:

$$[2] \quad \sigma_m = \sigma \sqrt{1 + 1/36 + (X - \bar{X})^2 / \sum_{i=1}^{36} (X_i - \bar{\bar{X}})^2}$$

The n associated with the t-value is an index denoting possible distributions of accounting error. The error could be concentrated all in one month or could be spread evenly or unevenly throughout two or more months. Even though there is no basis to assume an even spread in error, this assumption results in the most conservative decision for signaling excesses using the STAR approach. (For further discussion, see Stringer and Stewart, 1985, pp. 81-88)

The upper precision limit for n=1 is compared with the amount considered material (M) by the auditor. If the book amount includes a material error, the rule would indicate investigation with a (1-A) probability and leave an incorrect acceptance error risk of A. For n=2, a new t-value results in a new UPL which is compared with M/2. If the upper precision limit equals or exceeds M/2, an investigation of the month is indicated.

Methodology

One hundred data series with 48 observations each for X and Y were simulated as follows:

$Y = .333 + 2X + e$
$X \sim NID\ (4, .5776)$
$e \sim NID\ (0,V)$

The variance (V) for the normally distributed error terms ranged from .05 to .45 allowing for regression models with various levels of model significance. A corresponding 100 data series with 48 observations each for X and Y were simulated as follows:

$Y = .333 + 2X + e$
$X \sim NID\ (4,.5776)$
$e \sim UNIFORM\ (0,V)$

The variances for the uniformly distributed error terms corresponded to the variances for the normally distributed error terms. The simulation model yielded an expectation of Y of 8.333 and an audited value for the year of 12. (8.333) or \$100 (2). The first 36 observations represented 36 monthly amounts which were the base period used for model construction. The last 12 observations represented the 12-month audit period.

Six error conditions were seeded into the audit period months:

1) No error
2) Error equal to materiality in one month
3) Error equal to 1/2 of materiality in each of two months
4) Error equal to 1/4 of materiality in each of four months
5) Error equal to 1/6 of materiality in each of six months
6) Error equal to 1/12 of materiality in each of twelve months

A material error (M) was defined as \$2 or 2 percent of the expected audit balance of the account (3). The months to which errors were added were randomly selected.

The STAR investigation rule is applied to the simulated data with seeded error, and incorrect decisions were defined as follows:

Annual incorrect rejection. Any year which has at least one month with no error signaled for investigation is classified as an annual incorrect rejection.

Monthly incorrect rejection. Any month with no error which is signaled for investigation is classified as a monthly incorrect rejection.

Annual incorrect acceptance. Any audit year which does not signal at least one month with an error will leave the equivalent of a material error in the year, and the year is classified as an annual incorrect acceptance.

Monthly incorrect acceptance. Any month with an error (whether equal to materiality or a fraction of materiality) that is not signaled for investigation is classified as a monthly incorrect acceptance.

The number of monthly and annual incorrect rejections and acceptances are tabulated for each of the six error conditions for the models with error term normality and the corresponding models with nonnormal error terms.

Results and Analysis

The regression models constructed from the simulated data were tested for error term normality by calculating the Shapiro-Wilk test statistic, W. The W statistic is the ratio of the best estimator of the variance (based on the square of a linear combination of the order statistics) to the usual corrected sum-of-squares estimator of the variance. Small values of W lead to the rejection of the null hypothesis that the error terms are a random sample from a normal distribution. Table 1 presents a summary of the results of the error term normality test for those models constructed with simulated data without error term normality and those corresponding models constructed with simulated data with error term normality. The majority of the models constructed with the non-normal error term data resulted in small values of W, with 75% of the models with a W-statistic no more than .10. The models constructed with the normal error term data exhibited higher values of W although 15% of the models produced a W-statistic of .10 or less. The test does indicate that, in general, one group of models tended to exhibit a much greater degree of error term normality than the other group allowing for a comparison of the two types of models.

Errors were generated into the audit period months, the STAR decision rule was used to attempt detection of these errors, and the number of incorrect rejections and

FIGURE 1

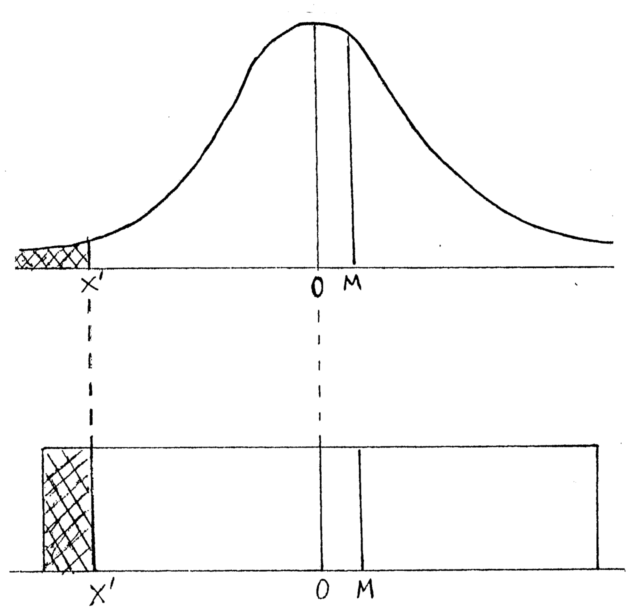


TABLE 1

SUMMARY SHAPIRO-WILK TEST FOR ERROR TERM NORMALITY

W STATISTIC	MODELS CONSTRUCTED WITH DATA SIMULATED WITH NORMAL ERROR TERMS	MODELS CONSTRUCTED WITH DATA SIMULATED WITH NONNORMAL ERROR TERMS
.00-.10	15	75
.10-.20	4	11
.20-.30	10	2
.30-.40	8	6
.40-.50	15	3
.50-.60	12	1
.60-.70	14	2
.70-.80	9	0
.80-.90	5	0
.90-1.00	8	0

TABLE 2

SUMMARY OF NUMBER OF INCORRECT DECISIONS
USING THE MONTHLY DECISION RULE

	Nonnormal Error Terms			Normal Error Terms					
	MONTHS INVESTIGATED	INCORRECT REJECTION	INCORRECT ACCEPTANCE	MONTHS INVESTIGATED	INCORRECT REJECTION	INCORRECT ACCEPTANCE	INCORRECT ACCEPTANCE DIFFERENCE	NUMBER OF INCORRECT ACCEPTANCES POSSIBLE	% OF POSSIBLE
NO ERROR	24	24	--	54	54	--	--	--	--
ERROR=M+	121	24	3	147	50	3	0	100	0.00%
ERROR=1/2*M	152	21	69	172	43	71	2	200	1.00%
ERROR=1/4*M	163	17	254	164	35	271	17	400	4.25%
ERROR=1/6*M	151	13	452	119	28	509	47	600	7.83%
ERROR=1/12*M	125	--	1075	102	--	1098	23	1200	1.92%

+M=Materiality amount

incorrect acceptances were tabulated. Table 2 presents a summary of these monthly incorrect decisions by model type and by the amount of the generated error.

Incorrect Rejections

The models with error-term normality resulted in a greater number of months investigated when no error was generated into the audit period. The number of incorrect rejections declined as the number of months affected by generated errors increased since some of the error-free months originally signaled now have a generated error and are no longer classified as an incorrect rejection. However, the number of incorrect decisions remained relatively higher for the normal error-term models. Figure 1 may help to explain why models with error-term normality may produce a greater number of incorrect rejections than models with non-normal error terms when using the STAR decision rule. When normality exists, the error terms are more concentrated around the mean than the error terms which are uniformly distributed. Using the normal variate (Z-value) to calculate the confidence interval will result in a greater number of error terms beyond the critical value (X') for the uniform distribution. Therefore e is more likely to be located beyond X' for the uniform distribution than for the normally distributed residuals expressed as:

$$e < X'$$

Since $X' = M - Z(\alpha)\cdot\sigma$, then $e < M - Z(\alpha)\cdot\sigma$, and $e + Z(\alpha)\cdot\sigma < M$ and fewer months are likely to be signaled for investigation. The normal error-term model also resulted in a greater number of incorrect rejections using the annual decision rule (summarized in Table 3).

Incorrect Acceptances

Referring to Table 2 there were only three incorrect acceptances for each of the two types of models when the seeded error was equal to materiality. An error of that magnitude concentrated into one month was easy to detect regardless of the degree of error-term normality. As the generated error became smaller, the difference between the number of incorrect acceptances for the two models became greater. However, compared to the number of incorrect acceptances possible also included in Table 2, this difference is relatively small. The smallest percentage difference occurred when the seeded error was equal to materiality or 1/2 of materiality which was fairly easy to detect, or when the seeded error was equal to 1/12 of a material error which is difficult to detect for both types of models.

Table 4 presents a more in-depth analysis of these differences by summarizing the total number of normal error-term models for which there were greater and fewer incorrect acceptances than the nonnormal error-

TABLE 3

SUMMARY OF NUMBER OF INCORRECT DECISIONS
USING THE ANNUAL DECISION RULE

	NONNORMAL ERROR TERMS		NORMAL ERROR TERMS	
	INCORRECT REJECTIONS	INCORRECT ACCEPTANCES	INCORRECT REJECTIONS	INCORRECT ACCEPTANCES
NO ERROR	17	--	39	--
ERROR=M	17	3	37	3
ERROR=1/2*M	15	15	33	19
ERROR=1/4*M	13	20	28	26
ERROR=1/6*M	12	26	25	42
ERROR=1/12*M	--	28	--	39

TABLE 4

SUMMARY OF INCORRECT ACCEPTANCE DIFFERENCE
BY SEEDED ERROR AMOUNT

	NUMBER OF NONNORMAL ERROR-TERM MODELS RESULTING IN A GREATER NUMBER OF INCORRECT ACCEPTANCES THAN NORMAL ERROR-TERM MODELS	NUMBER OF NORMAL ERROR TERM MODELS RESULTING IN A GREATER NUMBER OF INCORRECT ACCEPTANCES THAN NONNORMAL ERROR-TERM MODELS	NO. OF MODELS PRODUCING DIFFERENT RESULTS
ERROR = M			
GREATER BY 1	2	2	4
ERROR=1/2*M			
GREATER BY 1	21	25	
GREATER BY 2	5	4	55
ERROR=1/4*M			
GREATER BY 1	20	23	
GREATER BY 2	7	11	
GREATER BY 3	2	4	67
ERROR=1/6*M			
GREATER BY 1	18	27	
GREATER BY 2	7	12	
GREATER BY 3	1	6	
GREATER BY 4	0	2	
GREATER BY 5	0	1	74
ERROR=1/12*M			
GREATER BY 1	20	20	
GREATER BY 2	8	12	
GREATER BY 3	4	9	
GREATER BY 4	1	1	75

TABLE 5

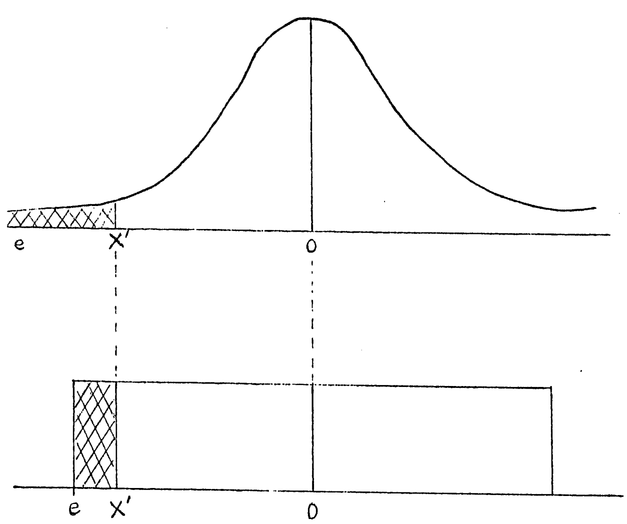
MONTHS INVESTIGATED IN EXCESS OF INCORRECT REJECTION MONTHS

	NONNORMAL ERROR-TERM MODELS	NORMAL ERROR-TERM MODELS
ERROR=M	97	93
ERROR=1/2*M	128	118
ERROR=1/4*M	139	110
ERROR=1/6*M	127	65
ERROR=1/12*M	101	48

term models. Although the difference in the number of incorrect decisions was relatively small, many models were involved which netted to this small difference. When the seeded error was equal to 1/2 of materiality, the total difference of incorrect acceptance errors was two; however, a total of 55 models produced different results. Twenty-one of the nonnormal error-term models resulted in one more incorrect acceptance than the normal error-term models, and five models resulted in a greater number of incorrect decisions by two. On the other hand, there were 25 and four normal error-term models which produced one and two more incorrect acceptances, respectively. Even though the total difference may not be relatively large, the corresponding models did not consistently produce the same results.

As presented in Table 2 the number of months investigated is greater for the normal error-term models when the seeded error is equal to materiality, 1/2 of materiality and 1/4 of materiality, but starts shifting at that point and has fewer months investigated for seeded errors of 1/6 and 1/12 of materiality. However, Table 5 reveals that if the months that were signaled for investigation when no error was present were removed, then the normal error term model consistently has fewer months signaled for investigation. The incorrect rejection months were removed since those months would continue to be signaled no matter what amount of error was added to the monthly data. Figure 2 attempts to explain why this situation may occur. The normally distributed error terms could theoretically be very large since the normal curve expands from -B to +B. The uniformly distributed error terms are truncated based on the mean and variance. The lowest possible value for e on the uniform curve is closer than some of the possible e values on the normal curve. When a finite error amount is added to the e, more uniform e's may exceed the critical value (X') than the normal e's since the uniform e may be closer to X'. Fewer additional months signaled results in a greater number of incorrect acceptances.

FIGURE 2



These results provide preliminary evidence that using normal curve theory to construct confidence intervals for decision rules when error terms are nonnormal may not significantly affect the auditor's decisions. This study did use a special case of nonnormality by simulating error terms that were uniformly distributed. Had the error terms been simulated in other nonnormal distributions, results may have differed from the results in this study. The results of this study do suggest normality of error terms is not essential.

Conclusion

Auditors may use regression-based decision rules as analytical procedures in an audit of financial statements. One underlying assumption of regression theory that may affect interpretation of confidence intervals is that the error terms are normally distributed. The results of this study suggest that auditor's decision when using the STAR regression-based investigation rule developed by Deloitte Haskins & Sells.

Data were simulated to result in regression models with and without error term normality. Material errors were seeded into the audit period data, and the regression-based decision rule was used in an attempt to detect the seeded errors. Although the normal error-term and corresponding nonnormal error-term models did not consistently produce the same results, the total difference of incorrect rejections and acceptances was relatively small. The results of this study provide a preliminary indication that normality of error terms may not be essential when auditors use regression-based decision rules.

Footnotes

1 There are generally 34 degrees of freedom which allows for a good approximation of the t-value by using the Z-value.

2 The values chosen for the above parameters, except for the variance of the error term, were chosen based upon a simulation study done by Kinney and Salamon (1982).

3 The material error was defined based upon the amount set as materiality in a simulation study done by Kinney and Salamon (1982).

References

1 Akresh, A. O. and W. A. Wallace, "The Application of Regression Analysis for Limited Review and Audit Planning," *Symposium on Auditing Research IV*, University of Illinois, pp. 67-167, 1980.

2 Kinney, W. R. and G. L. Salamon, "Regression Analysis in Auditing: A Comparison of Alternative Investigation Rules," *Journal of Accounting Research Autumn*, pp. 350-366, 1982.

3 Knechel, W. Robert, "A Simulation Study of the Relative Effectiveness of Alternative Analytical Review Procedures," *Decision Sciences*, volume 17, pp. 376-394, 1986.

4 Neter, J., "Two Case Studies on the Use of Regression for Analytical Review," *Symposium on Auditing Research IV*, University of Illinois, 1980.

5 Stringer, K. W., "A Statistical Technique for Analytical Review," *Journal of Accounting Research--Supplement*, pp. 1-9, 1975.

6 Stringer, K. W. and T. R. Stewart, *Statistical Techniques for Analytical Review in Auditing*, John Wiley & Sons, New York, N.Y., 1985.