International journal of advanced scientific and technical research Issue 3 volume 1,January 2013
Available online on http://www.rspublication.com/ijst/index.html ISSN 2249-9954

On the determination of critical values and other statistical characteristics in case of restricted models-A Kernel Density Approach

A. Rahman*, and A. K. Majumder

Lecturer, Department of Statistics, Jagannath University, Dhaka-1100, Bangladesh
Professor, Department of Statistics, Jahangirnagar University, Savar, Dhaka-1000

Abstract

Econometric models are based on a set of assumptions. These assumptions are based either on underlying theory that is being modeled or on any statistical evidence, or more commonly on both. One of the important assumptions regarding distributional pattern is that the data should follow the normal distribution. Wolak, (1989) provides formula to determine the weights of weighted mixture distribution, which is very complicated. This paper contributes in determining critical values for making a decision in case of testing such restricted models whatever the number of restricted parameters.

Key words: Critical values, kernel density, weighted mixture distribution,

Introduction

In making decision it is essential to setup model which is based on some assumptions of the related field. Applications of model are numerous and occur in almost every field, including engineering, physical science, economics, management, life science. In such disciplines researcher or decision maker often used tools of statistical inference techniques.

Econometric models are based on a set of assumptions. These assumptions are based either on underlying theory that is being modeled or on any statistical evidence, or more commonly on both. One of the important assumptions regarding distributional pattern is that the data should follow the normal distribution. Suppose we are not certain that the data are in fact distributed normally. Then what is the solution? We solve this problem by the non-parametric density estimation. The nonparametric approach is quite different. The idea is to rid oneself of the need to specify in advance a particular functional form. (see for example, John F. (2004)).

During the past 25 years the theoretical development of nonparametric probability density estimation has been rapid. Several different approaches have received extensive treatment: kernel methods (Rosenblatt 1956, Przen 1962, Watson and Leadbetter 1963, Cacoullos 1966, and Epanechnikow 1969), orthogonal series methods (Kronmal and Tarter 1968, Watson 1969, Wahba 1971, and Brunk 1978), penalized-likelihood methods (Good and Graskins 1971) etc. See Tapia and Thompson (1978) for recent and more complete bibliography.

Econometric testing problems are based on prior information about the sign of some or all of the unknown parameters. Such parameter information is important to improve the quality of inference. The density estimation performs magic, in the sense that it permits statistical inference more flexible in very general circumstance. Another issue related to estimation and testing is the determination of weights of the sampling distribution of test statistics when parameters are restricted. The distribution of these test statistics under restricted alternatives follows weighted mixture of chi-square, t, F etc. In some cases, estimation can be very tedious or complicated.

Wolak, (1989) provides formula to determine these weights up to seven restricted parameters. In such a situation, among the other estimation technique Kernel density (KD) estimation technique provides the solution to determine the critical value of the distribution whatever the number of restricted parameters. In this study, a newly proposed algorithm of kernel density estimation is used in determination of critical value of the test statistics.

Methods and materials

The Basics of Kernel Density Estimation

It is remarkable that the histogram stood as the only non parametric density estimator until the 1950s, when substantial and simultaneous progress was made in density estimation and in spectral density estimation. In a little-known paper, Fix and Hodges (1951) introduced the basic algorithm of nonparametric density estimation. They addressed the problem of statistical discrimination when the parametric form of the sampling density was not known. During the following decade, several general algorithms and alternative theoretical modes of analysis were introduced by Cencov (1962). There followed a second wave of important and primarily theoretical papers by Watson and Leadbetter (1963), Loftsgaarden and Quesenberry (1965), Schwartz (1967), Epanechnikov (1969), Tarter and Kronmal (1970), and Wahba (1971). The natural multivariate generalization was introduced by Cacoullos (1966). Finally, in the 1970s came the first papers focusing on the practical application of these methods by Scott et al. (1978) and Silverman (1978b). Then and later multivariate applications awaited the computing revolution.

Kernel density estimates (KDEs) at their simplest can be thought of as an alternative to the histogram. They typically provide a smoother representation of the data and, unlike the histogram, their appearance does not depend on a choice of starting point. In this sense KDEs alleviate problems with the histogram that have been perceived by some archaeologists (Whallon, 1987). In addition to the visual advantage of being a smooth curve, the kernel estimate has an advantage over the histogram in terms of bias. The bias of a histogram estimator with bin width h is of order h, whereas centering the kernel each data point and using a symmetric kernel zeroes this term and as such produces a leading bias term for the kernel estimate of order h^2. (see for example, Sheather (2004)).

The smoothness of the KDE means that it is aesthetically more pleasing than the histogram. It also facilities the presentation of several data sets in a single figure, and makes it easier to compare data sets. This has been argued and illustrated in Baxter and Beardah (1995b)

It does, however, suffer from a slight drawback when applied to data from long tailed distributions. Because the window width is fixed across the entire sample, there is a tendency for spurious noise to appear in the tails of the estimates; if the estimates are smoothed sufficiently to deal with this, then essential detail in the main part of the distribution is masked. An example of this behavior is given by disregarding the fact that the suicide data are naturally non-negative and estimating their density treating them as observations on $(-\infty, \infty)$.

Some elementary properties of kernel estimates follow at once from the definition. Provided the kernel K is everywhere non negative and satisfies the condition-in other words is a probability density function-it will follow at once from the definition that $\hat{f}$ will itself be a probability density. Furthermore, $\hat{f}$ will inherit all the continuity and differentiability properties of the kernel K, so that if for example K is the normal density function, then $\hat{f}$ will be a smooth curve having derivatives of all orders. There are arguments for sometimes using kernels which take negative as well as positive values. If such a kernel is used, then the estimate may itself be negative in places. However, for most practical purposes nonnegative kernels are used.

Data used here obtained from different distribution by simulation process to measure the performance of usual method and proposed method in determination of critical values. A total of 1000 observations were generated by 20000 times simulation process.

Density estimation is completely computer intensive. Applications of the techniques presented in this thesis require some programming in statistical computing packages that make density estimation easier than others. The entire programs used to generate the results in this paper are written in R version 2.9.0 along with the “np” package. **(**Hayfield T. and Racine J. S. (2008)).

Likelihood based test with restricted and unrestricted parameter space

There are three most commonly used test procedure that are based on the likelihood function usually named Likelihood Ratio (LR), Lagrange’s Multiplier (LM) and Wald test. In this study such tests are considered in case of restricted and unrestricted parameter space. For example, if $X_1, \cdots, X_n$ is a random sample from a population with pdf or pmf $f(x/\theta)$ (θ may be a vector), the likelihood function is defined as

$$L(\theta/x_1, x_2, \cdots, x_n) = L(\theta/x) = f(x/\theta) = \prod_{i=1}^{n} f(x_i/\theta).$$

Now in case of unrestricted parameter space, we have the form of hypothesis is $H_o : \theta \in \theta_o$ and for restricted parameter space, $H_o : \theta \in \theta_c$ where $c \in R$. For a particular case the unrestricted means $H_0 : \theta = \theta_0$ and for restricted case it means $H_0 : \theta > \theta_0$ or θ could be a p-dimensional vector.

Thus the usual likelihood based test under restricted and unrestricted parameter space has the following functional form

$LR = 2\left[\log(\hat{\theta}) - \log(\tilde{\theta})\right]$, where $\log(\tilde{\theta})$ is the restricted log-likelihood and $\log(\hat{\theta})$ is unrestricted log-likelihood. LM= $\{s(\theta_0)\}'\left[I(\hat{\theta})\right]^{-1}\{s(\theta_0)\}$, where $s(\theta_0)$ and $I(\theta_0)$ is score and information matrix respectively. The Wald test statistics has the following form $W = (\hat{\theta} - \theta_0)' \left[I(\hat{\theta})\right] (\hat{\theta} - \theta_0)$. In the restricted case, the distributions of these tests follow asymptotically chi-square distribution for large sample case. But in small sample case, we do not know the distribution of these test statistics. Our density estimation technique provides the pattern of the density of the distribution. On the other hand, for restricted parameter case, the distribution of these test statistics follows weighted mixture of chi-square and the determination of weights is very complicated. Wolak (1989) provides formula for the determination of these weights at most seven restricted parameters.

There are many excellent references that complement and expand on this subject. In density estimation, the classic tests of Tapia and Thompson (1978), Wertz (1978), and Thompson and Tapia (1990) first indicated the power of the nonparametric approach for univariate and bivariate data. Silverman (1986) has provided a further look at applications in these settings.

Critical value determination of restricted alternative by kernel density

Let us consider the form of regression model as follows

$$y = f(x) + \varepsilon$$

In matrix notation $y = X'\beta + \varepsilon$, where $\beta = (\beta_1, \beta_2, \beta_3)'$

In this setting, we might be interested in testing the hypothesis

$H_0 : \beta = 0$ against $H_1 : \beta > 0$. From the earlier discussion, it is noticed that the distribution of the test statistic for testing this restricted parameter follows weighted mixture of χ^2 or F distribution. For example, the test statistics $-2\ln LR$ has asymptotic distribution under H_0, is given by $P_r(-2\ln LR < c) = \sum_{i=0}^{p} W(p,i) P_r(\chi_i^2 < c)$, $c \in R$ which is a probability mixture of independent chi-squared distributions, with different degrees of freedom. The weights $W(p,i), i = 0,\cdots,p$, represent the asymptotic probability of the event that under H_0 any i elements are strictly positive. It is suggested to determine the weights at α level of significance such that the positive value c satisfying $P_r(-2\ln LR > c) = \alpha$ asymptotically under H_0 can be found by solving (Wu and King (1994)) $\alpha = \sum_{i=0}^{p} W(p,i) P_r(\chi_i^2 > c)$. In some cases it is very difficult to determine the weights when more than seven parameters are to be estimated or tested. Unless we determine the weights it is impossible to determine the critical values of the test statistics. Which is a serious problem happened in most of the regression model in hypothesis testing. The Kernel Density (KD) estimation techniques provide flexible solution to this problem in determination of critical values.

Proposed algorithm of kernel density approach in determining critical value for restricted alternatives:

Proposed algorithm consists of the following steps:

Step 1 Generate data from any known distribution or load data series from any specified field

Step 2 Determine the standard deviation of data series and inter quartile range.

Step 3 Calculate appropriate bandwidth or smoothing parameter (h) using standard deviation and inter quartile range in case of rule of thumb introduced by Silverman(1996).

Step 4 Choose the appropriate kernel function usually (Gaussian Kernel) to determine density of the data series.

Step 5 Kernel function depends on choice of starting location value x_0 and calculate distance from each point as a center point in the data set.

Step 6 Centered at each data point is each point's contribution in the overall density estimate namely $\left(\frac{1}{nh}\right)K\left(\left(x_0 - x_i\right)/h\right)$

Step 7 Finally the density estimate is the sum of these scaled normal densities**.**

Step 8 Now generate two sets of uniform random number $U_1(0,b)$ and $U_2(0,b_1)$ of size N.

Step 9 Calculate the number of points that fall within the area of the estimated density function and estimate the approximate proportion $S_1 = N'/N$.

Step 10 Determine the critical value (point) at which it gives the .05 probability value.

Among the above steps, steps 1-7 are used as density estimation procedure and rest of the 3 steps are used to determine critical values.

Results and Discussions

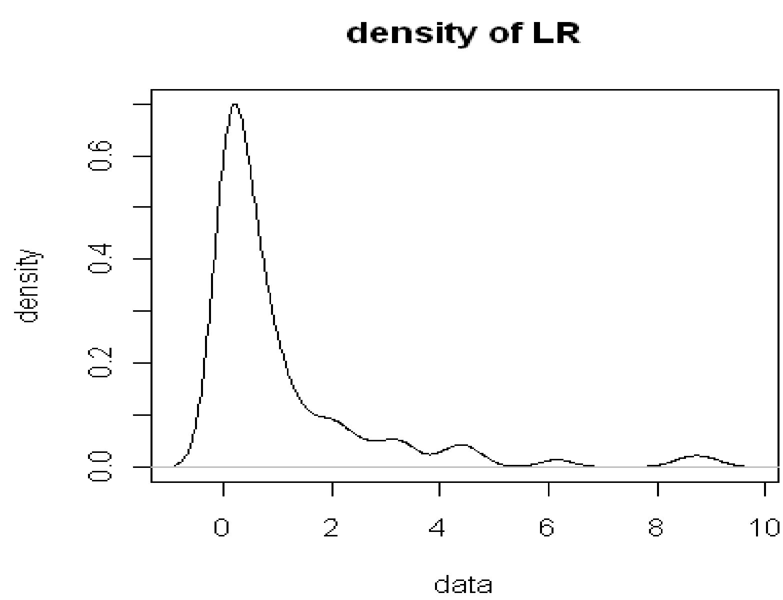
The literature about three asymptotically equivalent test statistics is discussed in Likelihood based test with restricted and unrestricted parameter space situation. Here we are interested to determine the density estimation of three asymptotic test statistics such as the LR, LM and the Wald test. It is known that they are the large sample test statistic. In some cases the distributional form of these test statistics in small sample are unknown. In that case we can obtain the density estimates of these test statistics. The results are in the following figures. It is observed from the figure 1 that the density of LR test statistics takes the skewed (positive) shape and long tail to the right indicates large values has less frequency than that of the small values occurs large at the begging of the zero values.

Also it is observed from the above figure 2 that the density of LR test statistics takes the skewed (positive) shape and long tail to the right indicates large values has less frequency than that of the small values occurs large at the begging of the zero values.

It is observed from the above figure 3 that the density of LR test statistics takes the skewed (positive) shape and long tail to the right indicates large values has less frequency than that of the small values occurs large at the begging of the zero values.

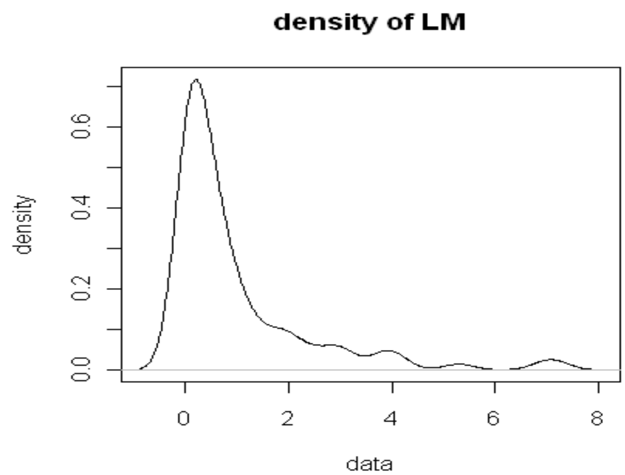
It is observed from the above figure 4 that the density of LR test statistics takes the skewed (positive) shape and long tail to the right indicates large values has less frequency than that of the small values occurs large at the begging of the zero values.

Figure 1: Density plot of data on simulated LR test statistic (weighted mixture of chi-square distribution).



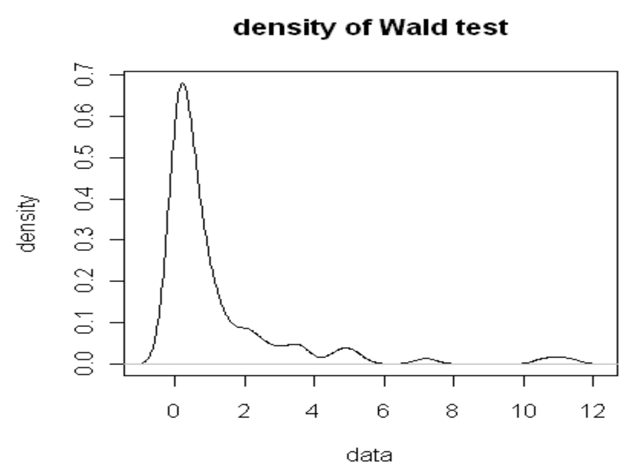
Comments: It is observed from the above figure that the density of LR test statistics takes the skewed (positive) shape and long tail to the right indicates large values has less frequency than that of the small values occurs large at the begging of the zero values.

Figure 2: Density plot of data on simulated LM test statistic (weighted mixture of Chi-square distribution) .



Comments: It is observed from the above figure that the density of LR test statistics takes the skewed (positive) shape and long tail to the right indicates large values has less frequency than that of the small values occurs large at the begging of the zero values.

Figure 3: Density plot of data on simulated Wald test statistic (which follows weighted mixture of Chi-square distribution)



Comments: It is observed from the above figure that the density of LR test statistics takes the skewed (positive) shape and long tail to the right indicates large values has less frequency than that of the small values occurs large at the begging of the zero values.

Here, the following results for simulated Z values, demonstrate that how our proposed algorithm of kernel density estimation provides the approximated critical value to the true value. So this procedure could be applicable to any of the complicated problem such as in restricted alternative case, determination of weights is difficult through usual procedure. In such a situation our proposed algorithm provides flexible solution in determination of critical value for restricted alternatives accept Monte Carlo simulation techniques.

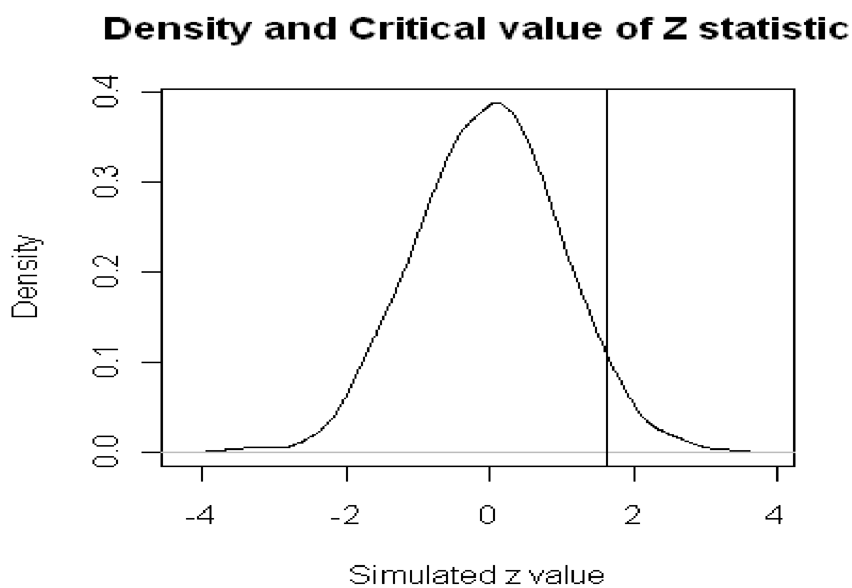
The following table will demonstrate the true critical value and critical value obtained by the Monte-Carlo simulation using density estimation technique.

Table 1 Critical value for standard normal distribution at 5% level of significance by density method.

Distribution of the test statistic(z)	Critical value at 5% level of significance	
	Usual Method	Proposed Method
Standard Normal	1.6453	1.641

And this is illustrated in figure 7 how the density estimation provides the 95% critical value.

Figure 7: Density plot for standard normal and the critical value at 5% level of significance by simulation based on n=2000.

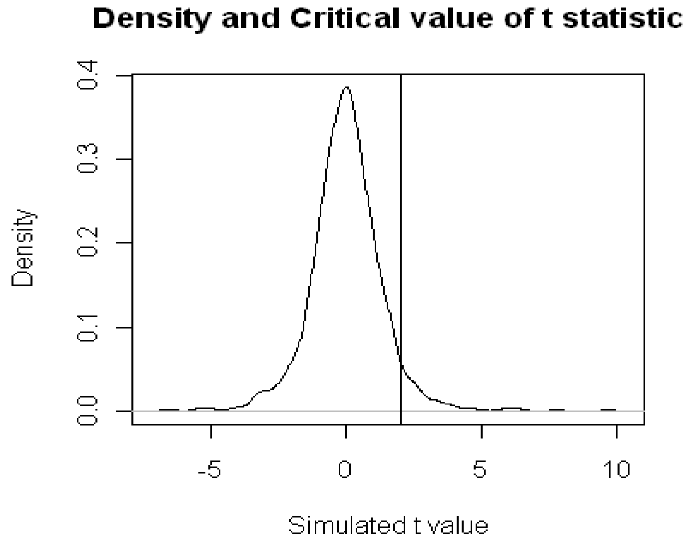


The following table will demonstrate the true critical value obtained by usual method and critical value obtained by the Monte-Carlo simulation using density estimation technique.

Table 2 Critical value for *t* at 5 % level of significance by density method

Distribution of the test statistic(*t*)	Critical value at 5% level of significance	
	Usual Method	Density Method
t distribution	1.943	1.921

Figure 8: Density of *t* statistic and its critical value obtained by simulation with n=2000 and d. f. = 6

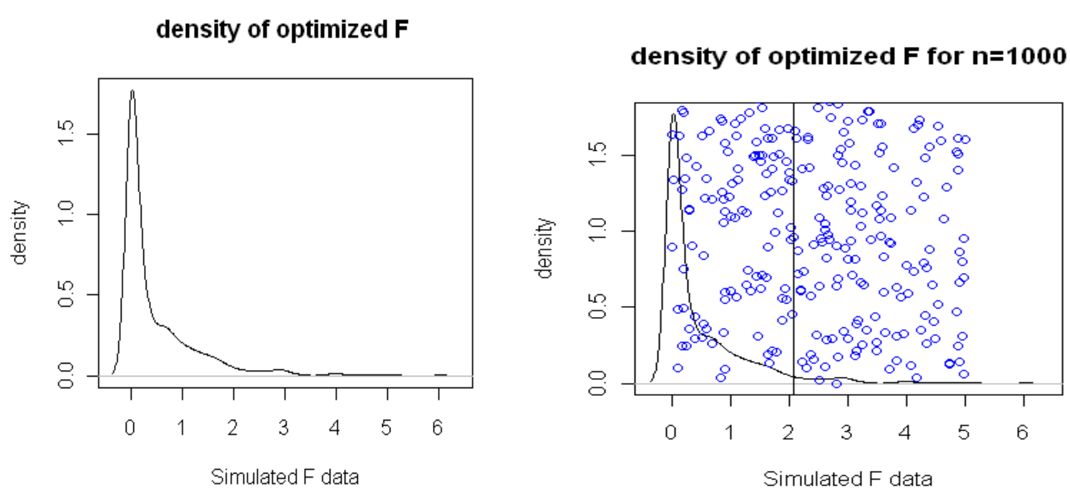


The following table will demonstrate the true critical value obtained by usual method and critical value obtained by the Monte-Carlo simulation using density estimation technique.

Table 3 Critical value for *F optimum* at 5 % level of significance by density method

Distribution of the test statistic*(F)*	Critical value at 5% level of significance	
	Usual Method	Density Method
F distribution	2.063089	2.063062

Figure 9 Density of *F* statistics and its critical value obtained by simulation with n=1000.



Conclusion

In conclusion we can say that once the determination of critical value of more than seven parameters is complicated in that case the use of our proposed algorithm of kernel density estimation provides the solution.

We perform the application of density in several test statistics whose distributional form is unknown in case of small sample. Finally, we wish to close by mentioning some related areas for further research. Determination of critical value of the restricted parameter testing problem in non-linear model should be an issue for further research. The restriction could be the linear restrictions of a non-linear model or non-linear restrictions of a non-linear model.

References

- Baxter, M. J. and Beardah, C. C. (1995b) Beyond the Histogram-Archaeological Applications of Kernel Density Estimation. Research Report 6/95, Nottingham Trent University Department of Mathematics, Statistics and Operational Research, Nottingham Trent University, Nottingham,U. K.

- Cacoullos, T. (1966), Estimation of a Multivariate Density, Annals of Institute for Statistics and Mathematics. 18 178-189.

- Epanechnikov, V. K. (1969), Nonparametric estimation of a multivariate probability Density. Theory of Probability and Application. 14 153-158.

- Hayfield T. and Racine J. S. (2008), The np package.

- John Fox (2004), Nonparametric regression. Department. of Sociology, McMaster University.

- Parzen, E. (1962) On estimation of probability density function and mode, Annals of Mathematical Statistics 33 1065-1076.

- Rosenblatt, M. (1956), Remarks on Some Nonparametric Estimates of a Density Function
- Annals of Mathematical Statistics. 27 832-837.

- Scott, D. W., Tapia, R. A. and Thompson, J. R. (1977), Kernel density estimation revisted, Nonlinear Analysis, 1, 339-72.
- Silverman B. W. (1986), Density Estimation for Statistics and Data Analysis, Chapman
- and Hall.

- Tarter, M. E. and Kronmal, R. A. (1970), On multivariate density estimates based on

- orthogonal expansions, Annals of Mathematical Statistics. 41 718-722.

- Watson, G. S. and Leadbetter, M. R. (1963), On the estimation of the probability density
- 1, Annals of Mathematical Statistics 34 480-491.

- Watson, G. S. and Leadbetter, M. R. (1963), On the estimation of the probability density
- 1, Annals of Mathematical Statistics 34 480-491.

- Wahba, G. (1971), A polynomial algorithm for density estimation, Annals of Mathematical Statistics.42 1870-1886.

- Wilks, S. S. (1932), Certain generalizations in the analysis of variance. Biometrika 24,
- 471-494.

- Wilks, S. S. (1960), Multidimensional statistical scatter. In contributions to probability
- and statistics in Honor of Harold Hotelling, ed. I. Olkin et al. Stanford University press,pp 486-503.