ELSEVIER

Statistics & Probability Letters 33 (1997) 95–103

STATISTICS & PROBABILITY LETTERS

Testing for constant variance in a linear model

Angela Diblasi *, Adrian Bowman

Department of Statistics, University of Glasgow, G12 8QQ, UK

Received October 1995; revised June 1996

Abstract

A nonparametric test of constant variance for the errors in a linear model is constructed through nonparametric smoothing of the residuals on a suitably transformed scale. Standard results on quadratic forms allow accurate distributional calculations to be made.

Keywords: Homoscedasticity; Linear model; Nonparametric regression; Quadratic form; Reference band; Variance

1. Introduction

It is a common assumption of linear regression models that the error terms all have equal variances. Checks on this assumption are often carried out by visual examination of appropriate residual plots. Where the interpretation of these is in doubt, a few more formal procedures exist. One example is the proposal of Cook and Weisberg (1983) which tests for the presence of an exponential trend in the variance. Where lack of fit in the proposed regression function is of interest, recent work has used nonparametric smoothing techniques to identify trends in residuals plots. For example, Raz (1990), Eubank and Spiegelman (1990) and Azzalini and Bowman (1993) describe this approach. It is the object of the present paper to apply techniques of nonparametric smoothing to the assumption of constant variance, by examining trends in the squared residuals.

A variety of papers over the last two decades have dealt with the estimation of a variance function under the assumption of heteroscedasticity in regression models. Some of these papers used nonparametric smoothing techniques to tackle this problem. Müller and Stadtmüller (1987) used kernel smoothers to obtain estimators of the variance function in the general regression model. Müller and Zhao (1995) proposed a methodology to estimate both parametric and nonparametric components in models where the regression relationship may also be estimated in a nonparametric manner. A test for heteroscedasticity, using the asymptotic distributions of the estimated parameters, was also described.

* Corresponding author. Supported by the Universidad Nacional de Cuyo, Universidad Tecnologica Nacional and CONICET in Argentina and by an Overseas Research Student award from the UK.

PII S0167-7152(96)00115-0

In this paper, a smoothing technique is applied to the residuals r_i on the transformed scale, $t_i = |r_i|^{1/2}$. The transformation is used to induce approximate normality. In Section 2, the behaviour of these variables under the hypotheses of homoscedasticity and of heteroscedasticity of errors is examined informally. A test statistic suitable for quantifying heteroscedasticity is proposed and its approximate distribution under the null hypothesis of homoscedasticity is calculated. This calculation uses accurate chi-squared approximations, rather than asymptotic theory. A simpler distributional approximation, based on ignoring the correlation among residuals, is examined in Section 3, along with a bootstrap proposal. Simulations are used to assess the sizes and powers of the tests in Section 4. A reference band is proposed in Section 5 as a graphical aid to the identification of heteroscedasticity. Some final discussion is given in Section 6.

2. A description of the test

2.1. The test statistic

As a first approach we consider the simple linear model

$$y_i = \beta_0 + \beta_1 x_i + \varepsilon_i, \quad i = 1,\dots,n, \tag{1}$$

where ε_i has a normal distribution with mean 0 and variance σ_i^2, and $\text{cov}(\varepsilon_i, \varepsilon_j) = 0$ for $i \neq j$. Since the ideas described below employ residuals they can easily be extended to the general linear model. The hypotheses to be tested can be written as

$$\begin{aligned} &\mathrm{H}_0 : \sigma_i^2 = \sigma_0^2, \quad \text{a constant},\\ &\mathrm{H}_1 : \sigma_i^2 = \sigma^2(x_i), \quad \text{a smooth function of } x_i. \end{aligned}$$

Using least squares, the fitted regression model $\hat{y}_i = \hat{\beta}_0 + \hat{\beta}_1 x_i$ produces the residuals $r_i = y_i - \hat{y}_i$. Under an assumption of normally distributed errors, these residuals are also normally distributed with mean zero.

It would be natural to examine the behaviour of $|r_i|$ or r_i^2 in order to check scale changes in the errors as suggested by Carroll (1982) and Carroll and Ruppert (1988). However, these variables have skewed distributions and the techniques to be used in this paper require approximate normality. Some authors, such as Carroll and Ruppert (1988) and McCullagh and Nelder (1989) suggest the transformation $|r_i|^{2/3}$, based on an original proposal by Wilson and Hilferty (1931). In the present paper the transformation $t_i = |r_i|^{1/2}$ is used, as described by Cleveland (1993) among others. This transformation has the same basic shape as the function $\Phi^{-1}F$ which provides an exact means of creating a normal random variable, with distribution function Φ, from a chi-squared random variable, with distribution function F. Since the r_i all have slightly different variances, even when the variance of ε_i is constant, the transformed quantities all have slightly different means. It is, therefore, more appropriate to deal with the adjusted variables $s_i = |r_i|^{1/2} - E_0(|r_i|^{1/2})$, where the subscript 0 denotes that the calculation is carried out under the null hypothesis. The expectation $E_0(|r_i|^{1/2})$ can be calculated easily, and an explicit expression is given in Section 2.2 below.

Under H_0 the values of s_i will lie close to their overall average $\bar{s}$, whereas under H_1 local changes in the mean of s_i are to be expected. If it is reasonable to assume that these local variations change in a smooth manner, then it becomes natural to employ nonparametric smoothing to identify the trends in a more powerful way, without making any assumptions about the shape of these trends. The kernel method of nonparametric regression provides a simple means of doing this. Wand and Jones (1995) give an introduction to this technique. In its simplest form, a smooth curve is defined across the design space as

$$\tilde{s}(x) = \sum_{j=1}^{n} w_j(x) s_j, \tag{2}$$

where the weights $w_j(x)$ are defined by the kernel function and sum to one, to provide a weighted average. In this paper a normal kernel function is used, giving weights

$$w_j(x) = \frac{\exp\left(\frac{x - x_j}{h}\right)^2}{\sum_{k=1}^{n} \exp\left(\frac{x - x_k}{h}\right)^2}.$$

If the values $\tilde{s}(x_i)$ of the smooth curve at each design point x_i are denoted by $\tilde{s}_i$ then a suitable test statistic to assess the degree of scale variation in the data is

$$T = \frac{\sum_{i=1}^{n}(s_i - \bar{s})^2 - \sum_{i=1}^{n}(s_i - \tilde{s}_i)^2}{\sum_{i=1}^{n}(s_i - \tilde{s}_i)^2}.$$

This is an analogue of the "lack-of fit" test statistic in parametric regression. Under the null hypothesis, there will be little difference in the size of the terms $\sum_{i=1}^{n}(s_i - \bar{s})^2$ and $\sum_{i=1}^{n}(s_i - \tilde{s}_i)^2$. Under the alternative, the first of these terms should become systematically larger than the second and so the test statistic will tend towards large positive values.

2.2. The distribution of the test statistic

In considering the distribution of the statistic T it is important to note that expression (2) is linear in the variables s_i. If we denote the column vector of s_i's by s, the column vector of $\tilde{s}_i$'s by $\tilde{s}$, and the $n \times n$ matrix of weights w_{ij} by W, then we can write $\tilde{s} = Ws$ and the quadratic form $\sum_{i=1}^{n}(s_i - \tilde{s}_i)^2$ can be expressed as $(s - Ws)^{\mathrm{T}}(s - Ws) = s^{\mathrm{T}}Bs$, where B is the matrix $(I - W)^{\mathrm{T}}(I - W)$, and I is the $n \times n$ unit matrix. The proposed statistic can therefore be written as

$$T = \frac{s^{\mathrm{T}}As - s^{\mathrm{T}}Bs}{s^{\mathrm{T}}Bs},$$

where A is the matrix $I - (1/n)L$, and L is the $n \times n$ matrix with all of its entries equal to one. This shows that T is the ratio of two quadratic forms, with $T = s^{\mathrm{T}}Cs/s^{\mathrm{T}}Bs$ where $C = A - B$.

Properties of quadratic forms in normal variables are well documented. Johnson and Kotz (1970) give a variety of results. In this case, the random variables in the quadratic forms are approximately normal and so the distribution of T can be approximated easily, following the techniques used by King et al. (1991) and Azzalini and Bowman (1993). Where ratios are involved, a standard manipulation can reduce calculations to a single quadratic form. Here, if t_1 denotes an observed value of T, the corresponding p-value for the test can be written as

$$\begin{aligned} p &= P(T > t_1 \mid \mathrm{H}_0 \text{ true}) \\ &= P\left(\frac{s^{\mathrm{T}}Cs}{s^{\mathrm{T}}Bs} > t_1 \,\middle|\, \sigma_i^2 = \sigma_0^2\right) \\ &= P(s^{\mathrm{T}}(C - t_1B)s > 0 \mid \sigma_i^2 = \sigma_0^2). \end{aligned}$$

The distribution of the final quadratic form $s^{\mathrm{T}}(C - t_1B)s$ can be handled through an accurate approximation, by matching moments to a shifted χ^2 distribution of the form $a + b\chi_c^2$. The cumulants of the quadratic form, and hence the moments, are available through the expression $k_j = 2^{j-1}(j-1)!\mathrm{tr}\{((C - t_1B)\Sigma)^j\}$, for $j = 1,2,3$, where Σ is the variance–covariance matrix of $s = (s_1, s_2, \ldots, s_n)$.

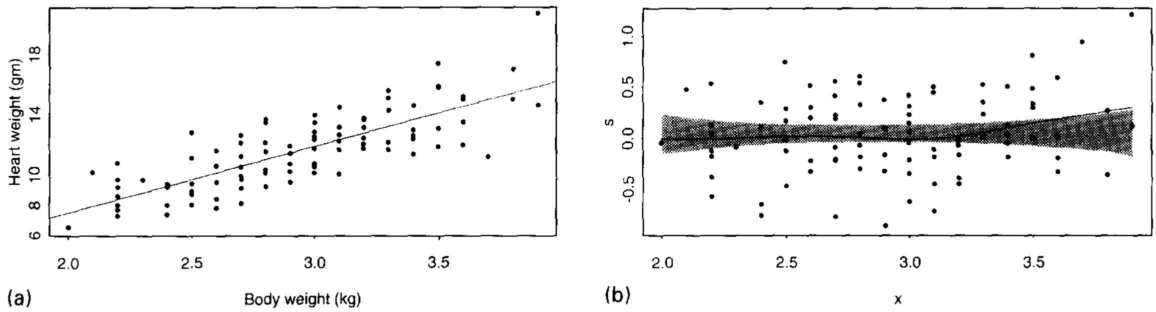

Fig. 1. Reference band for the cats data. (a) Observed data with fixed regression line. (b) Reference band and observed curve.

An explicit expression for the density function of $t_i = \sqrt{|r_i|}$ is

$$f_{t_i}(x) = \frac{4x}{\sqrt{2\pi\sigma_{r_i}^2}} \exp\left(-\frac{x^4}{2\sigma_{r_i}^2}\right), \quad x \geqslant 0,$$

where $\sigma_{r_i}^2$ is the variance of the residual r_i. The expected values and variances are

$$E(t_i) = \frac{2^{1/4}\Gamma(3/4)\sqrt{\sigma_{r_i}}}{\sqrt{\pi}},$$

$$\operatorname{var}(t_i) = \frac{\sqrt{2}(\sqrt{\pi} - \Gamma^2(3/4))\sigma_{r_i}}{\pi}.$$

The entries σ_{ij} of Σ are then $\sigma_{ii} = \operatorname{var}(s_i) = \operatorname{var}(t_i)$, and

$$\sigma_{ij} = \operatorname{cov}(s_i, s_j) = E(t_i t_j) - \frac{\sqrt{2}\Gamma^2(3/4)\sigma_{r_i}\sigma_{r_j}}{\pi}, \quad i \neq j.$$

The quantity $E(t_i t_j)$ can be calculated directly by numerical integration, using the joint distribution of (r_i, r_j).

2.3. An example

Fisher (1947) describes a dataset which includes the heart weights and body weights of a group of cats. Venables and Ripley (1994) also analyse these data. Fig. 1 shows a plot for the male cats. Aitchison (1986) suggested that a log transformation of both variables is appropriate. Such a transformation is often necessary in comparing weights of this kind.

For the nonparametric test, the smoothing parameter was chosen to be one-eighth of the range of the x-values. Since this value refers to the standard deviation of the normal kernel function, each kernel covers approximately half the observations from tail to tail. The effect of changing the value of the smoothing parameter will be discussed in Section 5. The test produces a p-value of 0.084. The mild change in variance for large body weights does not, of itself, provide convincing evidence of changing variance. However, the small p-value, in conjunction with other factors, does lend some support to the adoption of log scales for these data.

3. Alternative versions

3.1. Ignoring the covariances of the residuals

The covariances among the residuals r_i induce covariances among the derived variables s_i. The values of these covariances are difficult to calculate analytically, and so numerical integration was used, as described at the end of Section 2.2. Since the correlations of the residuals r_i are generally weak, this raises the question of whether ignoring these terms would substantially alter the accuracy of the test. The removal of the need to carry out numerical integration would provide a welcome simplification of the operation of the test. The results of approximating the covariance matrix of s by a diagonal matrix, using the formula for $\mathrm{var}(s_i)$ provided in Section 2.2, are shown in Table 1 of Section 4 below.

3.2. A bootstrap approach

Another possible tool, which has been widely used in the statistical literature during the last decade, is the bootstrap. Its advantage in the present setting is that the empirical distribution of the test statistic can be generated directly, without the use of moment approximations. The numerically intensive approach of the bootstrap is likely to lead to slower execution of the test. It does however provide a helpful means of checking on the accuracy of the proposed test. The present context involves a parametric model with normal errors, as defined in (1). Data can then be generated from the fitted model and used to study the distribution of the test statistic under the null hypothesis. Simulation procedures of this type have been referred to as "parametric bootstrap" methods.

An algorithm to implement the bootstrap in the present setting is as follows:

1. Simulate a set of observed values $y_1^*, y_2^*, \ldots, y_n^*$ from the fitted model (1).
2. Obtain the corresponding least-squares fitted values $\hat{y}_1^*, \hat{y}_2^*, \ldots, \hat{y}_n^*$ and the corresponding residuals $r_1^*, r_2^*, \ldots, r_n^*$.
3. Calculate the observed value t^* of the test statistic T.
4. Repeat steps 1 to 3 a large number of times.
5. Calculate the p-value of the observed data from the empirical distribution of T.

Note that the operation of the test does not depend on the true values of the parameters β_0, β_1 and σ. This allows the generation of the bootstrap residuals r_i^* to be simplified.

The performance of the bootstrap test is described in Section 4 below.

4. A power study

A small simulation study was carried out to analyse the size and power of the test. Different values of the sample size n, as well as different functions for the variance of the errors, were considered. A design based on equally spaced values of the explanatory variable in the interval $[a,b] = [0,1]$ was used. A wide range of bandwidths for estimating the variance function as a smooth curve has been used. The values for the regression parameters were $\beta_0 = 1$, $\beta_1 = 2$ and the simulated model was $y_i = \beta_0 + \beta_1 x_i + \varepsilon_i$, $i = 1, \ldots, n$ where ε_i were independent normal random variables with zero mean. In each case 500 samples were generated and the number of times the observed significance fell below 0.05 was counted. In Table 1, the results for the case where the correlations among residuals are ignored, and for the bootstrap, are also shown. The functions used for the variances were $\sigma_1(x) = 1$, $\sigma_2(x) = 0.25 + x$, $\sigma_3(x) = 0.25 + (x - 0.5)^2$ and $\sigma_4(x) = 0.25\ \exp(x \log(5))$. All these functions, apart from $\sigma_1(x)$, have the same minimum and maximum, to make comparison easier.

Table 1
Size and power for the three versions of the nonparametric test and the parametric test of Cook and Weisberg (1983), using simulated data with a variety of variance functions, sample sizes n and smoothing parameters h. The largest standard error for the entries in the table is 2%

		$\sigma_1(x) = 1$				$\sigma_2(x) = 0.25 + x$			
n	h	Non-ind. resid. %	Indep. resid. %	Bootstrap %	Cook and Weisb. %	Non-ind. resid. %	Indep. resid. %	Bootstrap %	Cook and Weisb. %
30	0.08	5.6	5.8	6.0	3.8	41.0	41.2	47.4	63.2
	0.16	5.2	6.0	5.2		52.8	53.8	59.4	
	0.32	3.6	4.2	5.2		61.4	62.4	70.2	
50	0.08	3.6	3.6	7.0	5.0	71.4	72.8	75.6	92.1
	0.16	4.8	3.6	4.6		85.6	86.0	89.6	
	0.32	4.0	4.6	6.4		89.6	90.0	93.8	
70	0.08	6.4	6.0	7.2	5.8	87.8	88.0	88.0	97.6
	0.16	5.6	5.4	6.2		94.2	94.4	96.4	
	0.32	6.4	6.0	7.0		96.2	96.0	98.2	
		$\sigma_3(x) = 0.25 + 4(x - 0.5)^2$				$\sigma_4(x) = 0.25\exp(\ln(5)x)$			
30	0.08	55.8	56.2	69.2	30.0	49.4	50.6	50.2	86.0
	0.16	63.2	64.6	75.0		63.0	64.8	55.8	
	0.32	32.0	35.4	45.2		73.6	75.4	82.0	
50	0.08	88.2	88.4	93.2	30.4	81.4	82.2	84.4	98.6
	0.16	92.0	92.2	96.4		89.0	89.6	92.6	
	0.32	69.2	70.4	89.2		94.0	94.2	96.6	
70	0.08	97.2	97.2	97.8	35.6	93.4	93.8	93.0	99.6
	0.16	98.4	98.2	98.4		96.8	96.6	98.8	
	0.32	90.4	90.8	97.6		98.2	98.2	94.2	

The fact that the test involves a ratio leads to invariance with respect to σ_0, since any multiplicative constant will cancel in the test statistic T.

In all cases, the size of the test, indicated by the results for $\sigma_1(x)$, are close to the target value of 5%. In addition, there is very little difference between the performance of the bootstrap and the other two tests. Since the bootstrap provides a direct means of generating the empirical distribution of the test statistic T, these results therefore provide confirmation that the approximation considered in this paper for the distribution of T is effective. The results in Table 1 document the power achieved by the nonparametric test and allow comparison with the parametric test of Cook and Weisberg (1983) whose assumption for the variance of errors is $\mathrm{var}(\varepsilon_i) = \sigma_0^2 \exp(\lambda x_i)$, $i = 1,\ldots,n$ where λ is an unknown parameter. Clearly, when the parametric assumption is appropriate the power of the parametric test is superior. When the parametric assumption is inappropriate the reverse is true.

For the nonparametric test, the range of smoothing parameters considered is very large, changing by a factor of four. Despite this, the change in power is relatively small, with a small increase as h increases for $\sigma_2(x)$ and a small decrease for $\sigma_3(x)$. These simulation studies, and others not reported here, allow us to conclude that the approximation of the variance–covariance matrix of the residuals by a diagonal matrix does not lead to significant inaccuracy in the size of the test. It does, however, allow very straightforward techniques of calculation to be used. Even where the correlation among the residuals is taken into account, the computer time necessary for the calculation of the p-value is not an obstacle.

5. A reference band

The idea of smooth kernel estimation can also be used as a graphical tool to explore the local behaviour of the variance as a follow-up to a global test. Bowman and Young (1996) explore the use of *reference bands* for nonparametric curves in a variety of settings. In the present case, the variables s_i can be used to build reference bands which indicate where the smooth curve $s(x)$ should lie if the null hypothesis of constant variance is true. Since $s(x)$ is approximately normal with mean 0 and known variance, then, for each value x in the design space, and under the null hypothesis of constant variance, we have

$$-q(x) = -2\sqrt{\mathrm{var}(\tilde{s}(x)-\bar{s})} \leqslant \tilde{s}(x)-\bar{s} \leqslant 2\sqrt{\mathrm{var}(\tilde{s}(x)-\bar{s})} = q(x)$$

with approximate probability 95%. The 95% reference band is therefore defined as

$$\{(x,y)\colon \bar{s}-q(x)\leqslant y\leqslant \bar{s}+q(x);\ a\leqslant x\leqslant b\}.$$

This can be used as a graphical follow-up to the global test. The band is constructed from pointwise intervals and so it does not provide simultaneous inference. It does, however, display the mean and variance structure under the null hypothesis. Where the test suggests that the variance is not constant, the reference band can therefore be used to explore which features of the data are contributing to this result. Where the test does not produce a significant result, the reference band might explain why apparent changes in scale do not in fact provide convincing evidence against the null hypothesis.

A simple calculation shows that the variance term can be written as

$$\begin{aligned}\mathrm{var}(\tilde{s}(x)-\bar{s}) = {} & \sum_{k=1}^{n} a_k(x)^2 0.12191\sqrt{\mathrm{var}(r_k)} \\ & + \sum_{l=1}^{n}\sum_{m=1}^{n} a_l(x)a_m(x)(E(t_l t_m) - 0.67598\sqrt{\mathrm{var}(r_l)\mathrm{var}(r_m)}),\end{aligned}$$

where $a_k(x) = w_k(x) - 1/n$, and $w_k(x)$ denotes the weights defined by the kernel function. Under the null hypothesis, $\mathrm{var}(r_k)$ is the kth diagonal entry of the matrix $(I - X(X^{\mathrm{T}}X)^{-1}X^{\mathrm{T}})$ multiplied by σ_0^2. The quantity $E(t_l t_m)$ in the second term of the expression can be computed by numerical integration, as discussed in Section 2.2. However, the calculation of the reference band is simplified if the correlations among the residuals are ignored. This further approximation was adopted in the example discussed below.

Fig. 1(b) displays a reference band for the cats data. The global test provides suggestive, but not convincing, evidence for changing variance. The reference band contains the observed curve for most values of x. It shows in particular that the variance in estimation of the smooth curve increases with the sparsity of x-values at the upper range of the data, and so caution must be exercised in interpreting the rise of the curve in that region.

6. Discussion

One of the most important advantages of using nonparametric techniques to check constant variance is that no particular shape of variance pattern is assumed. This widens the scope of the available tools. However this generalization will necessarily lead to some reduction in power over parametric methods when the parametric assumptions provide a good description of the true pattern. Data simulated from some of the variance functions used in Section 4 do not show marked changes in variance which can easily be identified visually. Despite

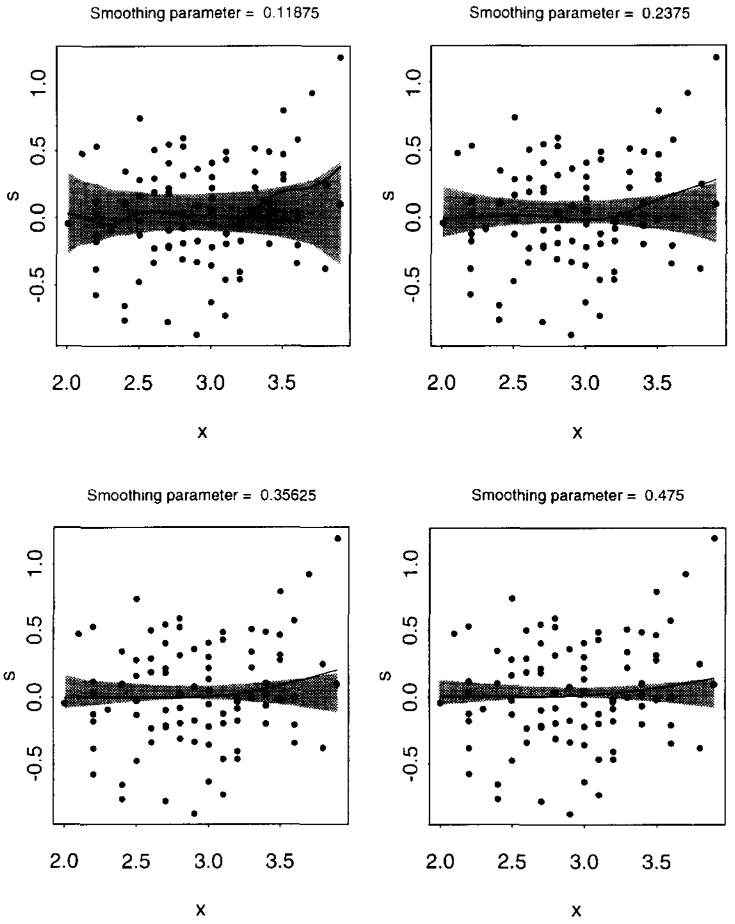

Fig. 2. Reference bands for a range of smoothing parameters with the cats data.

this, the results of Section 4 show that the power of the nonparametric test can reach very reasonable levels even in these cases.

The bandwidth h is an important parameter in smoothing techniques. Bowman and Young (1996) reviewed the role of the smoothing parameter in a number of nonparametric tests. In the present case the bandwidth seems to have little influence on the power of the test for reasonably large sample sizes ($n = 70$ in Table 1). For smaller values of n, the power of the test may be affected by the shape of the variance function under the alternative hypothesis as illustrated in Table 1. Large h gives the highest power for σ_2 and σ_4 whereas $h = 0.16$ seems to be best for σ_3. Fig. 2 shows reference bands for the cats data (in which $n = 95$), corresponding to a wide range of smoothing parameters. The values used are $i/16$ time the range of x, for $i = 1,\ldots,4$. In each case the information conveyed by the reference bands is the same, demonstrating that the particular choice of this parameter is not crucial. This is also true for the p-values from the test which, for the values of smoothing parameter used in Fig. 2, are 0.279, 0.084, 0.075 and 0.048, respectively. Apart from the case of the very small value of the smoothing parameter h, the evidence against constant variance is very similar across a wide range of smoothing parameters.

As an automatic technique, a "plug-in" bandwidth selection technique, such as the one described by Ruppert et al. (1996), could be used. These techniques assume independent observations, but it has already been demonstrated in Section 4 that the adoption of this approximation does not greatly affect the results. However, with the use of an automatic bandwidth selection procedure introduces an additional source of variation into the problem and so may have an effect on the distributional properties of the test.

Acknowledgements

The authors wish to acknowledge the valuable comments of an Associate Editor and a referee.

References

Aitchison, J. (1986), *The Statistical Analysis of Compositional Data* (Chapman & Hall, London).

Azzalini, A. and A.W. Bowman (1993), On the use of nonparametric regression for checking linear relationships, *J. Roy. Statist. Soc. Ser. B.* **55**, 549–557.

Bowman, A. and S. Young (1996), Graphical comparison of nonparametric curves, to appear in: *Appl. Statist.*

Carroll, R.J. (1982), Adapting for heteroscedasticity in linear models, *Ann. Statist.* **10**, 1224–1233.

Carroll, R.J. and D. Ruppert (1988), *Transformation and Weighting in Regression* (Chapman & Hall, New York).

Cleveland, W. (1993), *Visualising Data* (Hobart Press, New Jersey).

Cook, R.D. and S. Weisberg (1983), Diagnostics for heteroscedasticity in regression, *Biometrika* **70**, 1–10.

Eubank, R.L. and C.H. Spiegelman (1990), Testing the goodness-of-fit of a linear model via nonparametric techniques, *J. Amer. Statist. Assoc.* **85**, 387–392.

Fisher, R.A. (1947), The analysis of covariance method for the relation between a part and the whole, *Biometrics* **3**, 65–68.

Johnson and Kotz (1970), Continuous univariate distributions, Vol. 2 (Houghton Mifflin, Boston).

King, E., J.D. Hart and T.E. Wehrly (1992), Testing the equality of two regression curves using linear smoothers, *Statist. Probab. Lett.* **12**, 239–247.

McCullagh, P. and J.A. Nelder (1989), *Generalised Linear Models* (Chapman & Hall, London, 2nd ed.)

Müller, H.G. and U. Stadtmüller (1987), Estimation of heteroscedasticity in regression analysis, *Ann. Statist.* **15**, 610–625.

Müller, H.G. and P.H. Zhao (1995), On a semiparametric variance function model and a test for heteroscedasticity, to appear in: *Ann. Statist.*

Raz, J. (1990), Testing for no effect when estimating a smooth function by nonparametric regression: a randomisation approach, *J. Amer. Statist. Assoc.* **85**, 132–138.

Ruppert, D., S.J. Sheather and M.P. Wand (1996), An effective bandwidth selector for local least squares regression, to appear in: *J. Amer. Statist. Assoc.*

Venables, W.N. and B.D. Ripley (1994), *Modern Applied Statistics with S-Plus* (Springer, New York).

Wand, M. and M.C. Jones (1995), *Introduction to Kernel Smoothing* (Chapman & Hall, London).

Wilson, E.B. and M.M. Hilferty (1931), The distribution of chi-square, *Proc. Nat. Acad. Sci.* **17**, 684–688.