ETH zürich

ETH Library

# Robust testing in linear models
## the infinitesimal approach

**Doctoral Thesis**

**Author(s):**
Ronchetti, Elvezio

**Publication date:**
1982

**Permanent link:**
https://doi.org/https://doi.org/10.3929/ethz-a-000278506

**Rights / license:**
In Copyright - Non-Commercial Use Permitted

This page was generated automatically upon download from the ETH Zurich Research Collection.
For more information, please consult the Terms of use.

Diss. ETH No. 7084

# ROBUST TESTING IN LINEAR MODELS: THE INFINITESIMAL APPROACH

DISSERTATION
*submitted to the*
SWISS FEDERAL INSTITUTE OF TECHNOLOGY ZURICH
for the degree of
DOCTOR OF MATHEMATICS

presented by

ELVEZIO RONCHETTI

Dipl.Math. ETH
born on November 4, 1955
citizen of Monte TI, Switzerland

Accepted on the recommendation of
Prof.Dr. F. Hampel, referee
Prof.Dr. I. Olkin, Prof.Dr. C. Field, co-referees

1982

Printed by: CPZ, Berne/Switzerland

I

E. RONCHETTI

# ROBUST TESTING IN LINEAR MODELS : THE INFINITESIMAL APPROACH


## Abstract

The infinitesimal approach to the theory of robust statistics is extended to robust testing. By means of the influence function and the change-of-power function we investigate the properties of infinitesimal robustness of test procedures and we establish an optimality criterion for infinitesimal robustness of tests. This formalism allows us to construct optimally robust tests that maximize the asymptotic power, within a certain class, under the side condition of a bounded influence function.

Further a new infinitesimal concept, the change-of-variance function, is discussed both for the location and the regression model, and the following optimality problem is solved: we find the estimator that minimizes the trace of the asymptotic covariance matrix, within a certain class, under the side condition of a bounded change-of-variance function.

These infinitesimal methods are used for constructing a robust version of the classical F-test for linear models. To accomplish this, we introduce a new general class of test procedures for linear models and compute their influence function and asymptotic power. Then, solving the optimality problem for infinitesimal robustness of tests, we find the

optimally robust test in this class.

One can achieve a robust version of the F-test in two steps. First, one looks for the optimally robust test with respect to the influence of the residual ($\rho_c$-test). Secondly, taking into account the influence of position in factor space, one imposes the condition of a bounded total influence, obtaining optimally bounded influence tests which are the natural counterpart to optimally bounded influence estimators.

In this framework we also propose a robust procedure for the selection of the variables.

An alternative class of procedures (called $C(\alpha)$-type tests for linear models) is introduced. It is shown that they are equivalent to the former ones, having the same influence function and the same asymptotic efficiency. Moreover, an interesting connection between the optimally robust $C(\alpha)$-type test and an asymptotically minimax test is shown.

Finally, we compare the asymptotic behaviour of several test procedures under different distributions and we compute two examples based on real data: both show the excellent performance of our optimally robust tests.

III

E. RONCHETTI

# ROBUSTE TESTS FUER LINEARE MODELLE : DER INFINITESIMALE ANSATZ

## Zusammenfassung

Der infinitesimale Ansatz in der robusten Statistik wird für das Testproblem verallgemeinert. Wir untersuchen die Robustheitseigenschaften von Tests mit Hilfe der Einflussfunktion und der "Machtänderungsfunktion" (change-of-power function) und formulieren ein Optimalitätskriterium für die infinitesimale Robustheit von Tests. Dieser Formalismus erlaubt die Konstruktion optimaler robuster Tests, welche unter der Nebenbedingung, dass die Einflussfunktion beschränkt sei, die asymptotische Macht in einer gewissen Klasse von Tests maximieren.

Ein weiterer neuer infinitesimaler Begriff, die "Varianzänderungsfunktion" (change-of-variance function), wird für das Lokations- und das Regressionsmodell diskutiert und je das folgende Optimalitätsproblem gelöst: wir finden den Schätzer, welcher unter der Nebenbedingung, dass die "Varianzänderungsfunktion" beschränkt sei, die Spur der asymptotischen Kovarianzmatrix in einer gewissen Klasse von Schätzern minimiert.

Diese infinitesimalen Methoden werden dazu benützt, eine robuste Version des klassischen F-Tests für lineare Modelle zu konstruieren. Das wird folgendermassen erreicht.
Wir führen eine neue allgemeine Klasse von Testverfahren für

lineare Modelle ein und berechnen ihre Einflussfunktion und ihre asymptotische Macht. Dann lösen wir das Optimalitätsproblem für die infinitesimale Robustheit von Tests und finden so den optimalen robusten Test in dieser Klasse.
Diese Robustifizierung des F-Tests wird in zwei Schritten durchgeführt. Im ersten Schritt sucht man ein optimales robustes Testverfahren bezüglich des Einflusses auf das Residuum ($\rho_c$-Test). In einer zweiten Phase betrachtet man auch den Einfluss auf die Lage im Faktorraum und verlangt, dass der totale Einfluss beschränkt sei. Auf diese Weise erhält man optimale robuste Tests, welche den optimalen robusten Schätzern entsprechen. Ferner schlagen wir in diesem Zusammenhang ein robustes Verfahren für die Variablenauswahl vor.

Eine alternative Klasse von Verfahren (Tests vom Typus $C(\alpha)$ für lineare Modelle) wird eingeführt. Es wird gezeigt, dass diese Tests zu den obigen äquivalent sind, da sie die gleiche Einflussfunktion und die gleiche asymptotische Effizienz besitzen. Ferner wird eine interessante Beziehung zwischen dem optimalen robusten Test vom Typus $C(\alpha)$ und einem asymptotischen minimax Test gezeigt.

Am Ende vergleichen wir das asymptotische Verhalten von verschiedenen Testverfahren unter verschiedenen Verteilungen und berechnen zwei auf realen Daten basierende Beispiele: beide zeigen die vorzügliche Leistung unserer optimalen robusten Tests.

E. RONCHETTI

# TESTS ROBUSTES POUR LES MODELES LINEAIRES ET TECHNIQUES INFINITESIMALES

## Résumé

L'application de techniques infinitésimales dans la statistique robuste est généralisée à la théorie des tests robustes. Nous étudions les propriétés de robustesse des tests en utilisant la fonction d'influence et la "fonction d'influence sur la puissance" (change-of-power function) et formulons un critère d'optimalité pour la robustesse infinitésimale des tests. Ce formalisme nous permet de construire des tests robustes optimaux, qui maximisent la puissance asymptotique parmi les tests ayant une fonction d'influence bornée.

Une autre notion infinitésimale, la "fonction d'influence sur la variance" (change-of-variance function), est définie pour le modèle d'un paramètre de position et pour le modèle de la régression linéaire. On résout également le problème d'optimalité suivant: on trouve l'estimateur qui minimise la trace de la matrice de covariance asymptotique parmi les estimateurs ayant une "fonction d'influence sur la variance" bornée.

Ces méthodes infinitésimales sont employées pour construire une version robuste du F-test classique pour les modèles linéaires. Cela s'effectue de la manière suivante: on introduit une classe nouvelle et générale de tests pour les modèles linéaires et on calcule la fonction d'influence et la

puissance asymptotique. Puis, on résout le problème d'optimalité pour la robustesse infinitésimale des tests en trouvant le test robuste optimal dans cette classe.
On peut obtenir cette version robuste du F-test en deux étapes. Premièrement on cherche le test robuste optimal par rapport à l'influence du résidu ($\rho_c$-test). Deuxièmement on considère l'influence de la position dans l'espace des facteurs et on impose la condition d'avoir une influence totale bornée. On obtient alors les tests robustes optimaux qui correspondent aux estimateurs robustes optimaux.
Dans ce contexte on propose un critère robuste pour la sélection des variables.

On introduit de plus une classe alternative de tests (tests du type $C(\alpha)$ pour les modèles linéaires). On montre qu'ils sont équivalent aux tests qu'on a introduits auparavant car ils ont la même fonction d'influence et la même efficacité asymptotique. On montre aussi une relation intéressante parmi le test robuste optimal du type $C(\alpha)$ et un test asymptotiquement minimax.

Finalement on compare le comportement asymptotique des tests sous divers distributions et on calcule deux exemples réels: l'un et l'autre montrent l'excellente performance de nos tests robustes optimaux.

VII

E. RONCHETTI

# TEST ROBUSTI PER I MODELLI LINEARI E TECNICHE INFINITESIMALI

## Riassunto

L'impiego di tecniche infinitesimali nel campo della statistica robusta è generalizzato alla teoria dei test robusti. Studiamo le proprietà di robustezza dei test per mezzo della funzione d'influenza e della "funzione d'influenza sulla potenza" (change-of-power function) e formuliamo un criterio di ottimalità per la robustezza infinitesimale dei test. Questo formalismo ci permette di costruire i test robusti ottimali che massimizzano la potenza asintotica in una classe di test con la funzione d'influenza limitata.

Un altro concetto infinitesimale, la "funzione d'influenza sulla varianza" (change-of-variance function), viene discusso per il modello di un parametro di localizzazione e per quello della regressione lineare. Viene inoltre risolto il problema seguente: si costruisce la stima che minimizza la traccia della matrice di covarianza asintotica in una classe di stime con la "funzione d'influenza sulla varianza" limitata.

Questi metodi infinitesimali vengono usati per costruire una versione robusta del test F classico per i modelli lineari. Per ottenere ciò viene introdotta una nuova classe di test per i modelli lineari e viene calcolata la loro funzione d'influenza e la loro potenza asintotica. Poi si risolve il problema di

ottimalità per la robustezza infinitesimale dei test, trovando così il test robusto ottimale in questa classe.
È possibile ottenere questa versione robusta del test F in due stadi. Dapprima si cerca il test robusto ottimale rispetto alla influenza del residuo (test $\rho_c$). In seguito, tenendo conto dell'influenza della posizione nello spazio dei fattori, si impone la condizione di avere un'influenza totale limitata. Si ottengono così test robusti ottimali che corrispondono alle stime robuste ottimali.
In questo contesto si propone pure un criterio robusto per la selezione delle variabili.

Si introduce inoltre una classe alternativa di test (test del tipo $C(\alpha)$ per i modelli lineari). Si dimostra che essi sono equivalenti a quelli introdotti precedentemente in quanto possiedono la stessa funzione d'influenza e la stessa efficienza asintotica. Si stabilisce pure una relazione interessante fra il test robusto ottimale del tipo $C(\alpha)$ e un test asintoticamente minimax.

Infine si confronta il comportamento asintotico di alcuni test per diverse distribuzioni e si calcolano due esempi basati su dati reali: ambedue dimostrano il rendimento eccellente dei nostri test robusti ottimali.