Robuuste variantie-analyse

THE-RC 66096

Eindhoven University of Technology
Computing Centre Note 30

Robuuste Variantie-analyse

Jan B. Dijkstra

Samengesteld voor de Statistische Dag (VVS)
3 april 1986
Katholieke Hogeschool Tilburg

1. Inleiding.

Een veel voorkomend probleem in de toegepaste statistiek is het simultaan vergelijken van een aantal steekproefgemiddelden.
Het gebruikelijke model is:

$$\underline{y}_{ij} = \mu_i + \underline{e}_{ij}$$

Hierbij wordt doorgaans verondersteld dat $\underline{e}_{ij} \sim N(0, \sigma^2)$ en dat bovendien de fouten onafhankelijk zijn. De index i geeft het groepsnummer aan ($i = 1, \ldots, k$) en j identificeert de elementen binnen iedere groep ($j = 1, \ldots, n_i$). De nulhypothese luidt:

$$H_0: \mu_1 = \ldots = \mu_k$$

Onder de hierboven beschreven voorwaarden kan deze hypothese getoetst worden met klassieke enkelvoudige variantie-analyse. Het schema staat hieronder vermeld. N geeft het totale aantal waarnemingen aan; er geldt dus $N = \sum_{i=1}^{k} n_i$.

bron	kwadraten-som KS	vrijheids-graden	gemiddeld kwadraat GK	F
tussen groepen	$\sum_{i=1}^{k} n_i(\bar{y}_i - \bar{y})^2$	k-1	$KS_t/(k-1)$	GK_t/GK_b
binnen groepen	$\sum_{i=1}^{k} \sum_{j=1}^{n_i} (y_{ij} - \bar{y}_i)^2$	N-k	$KS_b/(N-k)$	
totaal	$\sum_{i=1}^{k} \sum_{j=1}^{n_i} (y_{ij} - \bar{y})^2$	N-1		

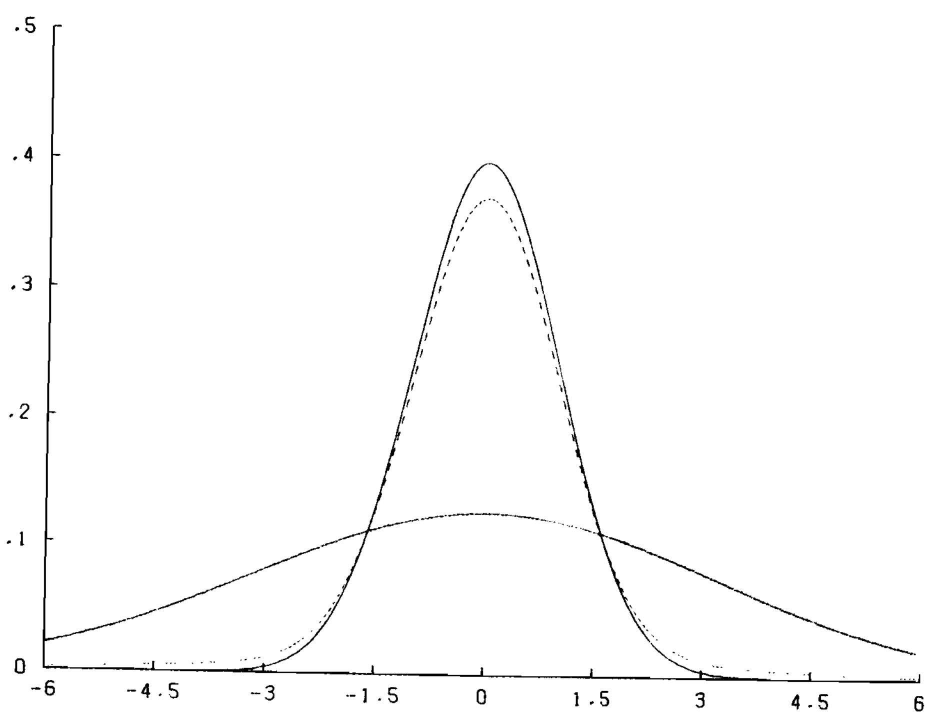
Hierin geeft $\bar{y}_i$ het gemiddelde in groep i weer en $\bar{y}$ het gemiddelde over alle waarnemingen. De beslissingsregel is nu: verwerp H_0 als $F > F_{N-k}^{k-1}(\alpha)$ bij gekozen onbetrouwbaarheid α.

Het onderwerp van deze notitie is het toetsen van H_0 als rekening gehouden moet worden met de mogelijke aanwezigheid van enkele uitschieters in de waarnemingen. Een model hiervoor is: met een kleine kans ϕ geldt $\underline{e}_{ij} \sim N(0, a\sigma^2)$ met $a >> 1$ en met kans $1-\phi$ blijft de situatie ongewijzigd. Ook de onafhankelijkheid van de fouten wordt gehandhaafd.

Als hier klassieke enkelvoudige variantie-analyse wordt toegepast, dan krijgt de verwerpingskans onder de nulhypothese mogelijk een andere waarde dan de gekozen onbetrouwbaarheid. Met andere woorden: de methode is niet robuust ten aanzien van deze afwijking van de modelveronderstellingen. In de volgende paragrafen worden vier manieren genoemd om met dit probleem om te gaan. Deze zijn: (i) Verdelingsvrije methoden, (ii) Trimming en winsorizing, (iii) Uitschieters verwijderen en (iv) Uitschieters dempen.

2. Verdelingsvrije methoden.

Hier wordt in feite een andere nulhypothese getoetst, namelijk "de steekproeven zijn afkomstig uit dezelfde continue verdeling".
H_A luidt nu dat de verwachtingen verschillen. De eis van normaliteit is dus vervallen en het is evident dat de hierboven beschreven mengvorm van twee normale verdelingen continu is.



In bovenstaande figuur hoort de steile top bij N(0, 1) en de vlakste grafiek bij N(0, 10). Er tussendoor loopt het hier onderzochte model met ϕ = 0.1. De bijbehorende kansdichtheidsfunctie is:

$$f(x) = \phi \frac{1}{\sigma\sqrt{a2\pi}} \exp[-\frac{x^2}{2a\sigma^2}] + (1-\phi) \frac{1}{\sigma\sqrt{2\pi}} \exp[-\frac{x^2}{2\sigma^2}]$$

met a = 10 en σ^2 = 1.

Voor de simultane vergelijking van een aantal lokatieparameters zijn verscheidene verdelingsvrije methoden bekend. Een geschikte keuze lijkt de toets van Van der Waerden (1952), die gebaseerd is op de volgende toetsingsgrootheid:

$$Q = \frac{N-1}{h} \sum_{i=1}^{k} \frac{1}{n_i}[\sum_{\ell \in S_i} \Phi^{-1}(\frac{R_\ell}{N+1})]^2 \text{ met } h = \sum_{\ell=1}^{N} [\Phi^{-1}(\frac{\ell}{N+1})]^2.$$

Hierbij stelt $y_1, \ldots, y_N$ de gecombineerde steekproef voor waarbinnen de groepen worden geïdentificeerd door indexverzamelingen S_i voor i = 1, ..., k. Verder is Φ de standaardnormale verdelingsfunctie en is R_ℓ de index van y_ℓ. Asymptotisch volgt Q een χ^2-verdeling met k-1 vrijheidsgraden en voor kleinere steekproeven bestaan tabellen.

De reden om uit de ruime collectie verdelingsvrije methoden voor het simultaan vergelijken van locatieparameters nu juist de toets van Van der Waerden te kiezen ligt in het feit dat dit de enige is die voor ϕ = 0 asymptotisch dezelfde efficiency heeft als de klassieke toets [Hajek (1969)]. Door deze verdelingsvrije methode te gebruiken kan men zich dus verzekeren tegen de eventuele aanwezigheid van uitschieters, waarbij de premie uitsluitend bestaat uit het verlies aan onderscheidend vermogen voor kleine steekproeven. En voor k = 2 valt dit verlies nogal mee [Van der Laan en Oosterhoff (1967)].

Als er sprake is van veel uitschieters zodat de staarten van de foutverdeling dikker worden, dan is de toets van Van der Waerden niet meer de optimale keuze. Een bekend voorbeeld van dikke staarten wordt gegeven door de dubbelexponentiële verdeling. Hiervoor is de toets van Mood en Brown (1950) optimaal. De mate van superioriteit laat zich goed uitdrukken in de Asymptotische Relatieve Efficiency (ARE). Er geldt: $ARE_{vdW,MB}$ (dubbelexponentieel) = $\frac{2}{\pi}$ en $ARE_{vdW,MB}$ (normaal) = $\frac{\pi}{2}$ [Hodges en Lehman (1961)]. Deze waarden wijken dermate veel van 1 af dat het de moeite waard lijkt om te zoeken naar een adaptieve verdelingsvrije methode die bij afwezigheid van uitschieters veel weg heeft van de toets van Van der Waerden, maar die zich ook aan andere omstandigheden goed aanpast.

Bij het gekozen model voor de foutverdeling is dit zeer wel mogelijk. Voor σ^2 bestaan verscheidene robuuste schatters die vrijwel geen hinder van de uitschieters ondervinden, en daarna kunnen ϕ en a simultaan uit de waarnemingen geschat worden door middel van de momentenmethode [Linders (1986)]. Voor de gevonden foutverdeling kan dan een verdelingsvrije toets geconstrueerd worden met optimale scorefuncties [Huber (1972)]. Een simulatie suggereert dat deze aanpak nogal grote steekproeven vereist. Het is nog zeer onduidelijk wat deze methode waard is als de feitelijke verdeling van de uitschieters anders is dan het hier besproken model.

3. Trimming en winsorizing.

Voor deze Engelse vaktermen zijn mij helaas geen Nederlandse alternatieven bekend. Toepassingen van deze technieken op de t-toets voor twee steekproeven zijn reeds gepubliceerd. Deze t-toets is de volgende:

$$t = \frac{\bar{y}_1 - \bar{y}_2}{\sqrt{\frac{KS_1+KS_2}{n_1+n_2-2}} \cdot \sqrt{\frac{1}{n_1}+\frac{1}{n_2}}} \quad \text{met } KS_i = \sum_{j=1}^{n_i} (\bar{y}_{ij}-\bar{y}_i)^2$$

De nulhypothese H_0: $\mu_1 = \mu_2$ wordt verworpen als $(|t| > t_{N-2}(\alpha)$, waarbij men een tabel voor tweezijdige toetsing dient te gebruiken. Deze methode is equivalent voor k=2 met de klassieke enkelvoudige variantie-analyse: t^2 is gelijk aan de F uit de inleiding en voor de kritieke waarden geldt hetzelfde verband $(t_\nu^2 = F_\nu^1)$.

Fung en Rahman (1980) hebben de t-toets door winsorizing ongevoelig gemaakt voor uitschieters. Dit gaat als volgt: Laat $a_1, \ldots, a_n$ een steekproef zijn die monotoon niet-dalend geordend is. Dan worden het gemiddelde en de kwadratensom, na tweezijdige winsorizing met parameter g, als volgt gedefinieerd:

$$\bar{a}_{wg} = \frac{1}{n}\{(g+1)a_{g+1} + a_{g+2} + \ldots + a_{n-g-1} + (g+1)a_{n-g}\}$$

$$KS_{wg} = (g+1)(a_{g+1} - \bar{a}_{wg})^2 + (a_{g+2} - \bar{a}_{wg})^2 + \ldots + (a_{n-g-1} - \bar{a}_{wg})^2 + (g+1)(a_{n-g} - \bar{a}_{wg})^2.$$

Het aantal relevante waarnemingen wordt hierdoor gereduceerd tot h = n-2g. Toepassing van deze techniek op de t-toets geeft de volgende formule:

$$t_{wg} = \frac{\bar{y}_{1wg} - \bar{y}_{2wg}}{\sqrt{\frac{KS_{1wg} + KS_{2wg}}{h_1+h_2-2}}\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}}$$

Deze waarde wordt vergeleken met een t-verdeling met h_1+h_2-2 vrijheidsgraden.

Bij winsorizing worden de staartelementen gelijk gemaakt aan de uiterste waarnemingen die niet tot de staart gerekend worden. Trimming is nog wat rigoreuzer: hier worden de staartelementen gewoon weggelaten. Yuen en Dixon (1973) hebben het gedrag van de t-toets na trimming reeds onderzocht. Beide technieken blijken in een simulatie met $n_i \geqslant 10$ dezelfde goede eigenschappen te vertonen: de verwerpingskans onder de nulhypothese is vrijwel gelijk aan de gekozen onbetrouwbaarheid, en het onderscheidend vermogen voor normale verdelingen ligt nauwelijks onder dat van de klassieke t-toets. Voor verdelingen met dikke staarten neemt het onderscheidend vermogen zelfs toe zolang g klein blijft ten opzichte van de steekproefgroottes.

Op grond hiervan lijkt het aantrekkelijk om deze technieken ook toe te passen op klassieke enkelvoudige variantie-analyse, wat in feite niets meer inhoudt dan generalisatie van het bovenstaande tot modellen met $k > 2$. Een simulatie gaf echter het volgende teleurstellende resultaat: de controle over de gekozen onbetrouwbaarheid, die zeer goed is voor k=2, neemt bij toenemende k snel af. En dit geldt vooral als de steekproefgroottes nogal verschillen.

De neiging tot conservativiteit (dat wil zeggen $P(\text{verwerpen}|H_0) < \alpha$) kan voor k=2 verkregen worden door niet van t_{wg} uit te gaan, maar van t^*_{wg} wat een generalisatie is van de toets voor één steekproef van Dixon en Tukey (1968). Er geldt:

$$t^*_{wg} = t_{wg}\sqrt{\frac{h_1 + h_2 - 2}{n_1 + n_2 - 2}}$$

Voor $k > 2$ blijkt dit echter geen oplossing. De controle over de gekozen onbetrouwbaarheid blijft onbevredigend. Bovendien lijkt het onaanvaardbaar om steeds uit elke groep 2 of meer waarnemingen te verwijderen, terwijl men zich slechts wil wapenen tegen een model met een kleine kans ϕ op uitschieters. Aantrekkelijker lijkt het om te zoeken naar een adaptieve variant op deze methoden. Daarover gaat de volgende paragraaf.

4. Uitschieters opsporen en verwijderen.

Voor deze en de volgende paragraaf moet het model voor variantie-analyse worden herschreven tot een regressiemodel:

$$\underline{y} = \beta_0 + \beta_1 x_1 + \ldots + \beta_{k-1} x_{k-1} + \underline{e}$$

De waarnemingen worden gerepresenteerd door y. Voor iedere waarneming wordt de bijbehorende groep geïdentificeerd door de dummy-variabelen $x_1, \ldots, x_{k-1}$. Er geldt $x_i = 1$ als y in groep i zit en anders geldt $x_i = 0$. Voor groep nummer k is geen dummy meer nodig omdat de identificatie al eenduidig is. Als nu zou gelden $\underline{e} \sim N(0, \sigma^2)$ met onafhankelijke fouten, dan komt toetsing van $H_0: \beta_1 = \ldots = \beta_{k-1} = 0$ neer op toetsing van $H_0: \mu_1 = \ldots = \mu_k$ bij klassieke enkelvoudige variantie-analyse. De F-waarden met bijbehorende vrijheidsgraden zijn dezelfde.

We gaan nu terug naar het bekende model met kans ϕ op uitschieters. Voor het opsporen van de uitschieters kan een methode gebruikt worden die door Leroy en Rousseeuw (1985) ontwikkeld is voor regressie-analyse.

De methode heet "Least Median of Squares" en kan hier op twee manieren worden toegepast:

(i) Per steekproef afzonderlijk. Dan worden de lokatieparameters μ_i geschat als oplossing van

$$\min_{\mu_i} \operatorname{med}_j (y_{ij} - \mu_i)^2.$$

Deze methode is zeer robuust: tot 50% uitschieters hebben geen invloed op de gevonden waarde.

(ii) Toepassing op alle waarnemingen gezamenlijk. De schatting voor $\beta = \beta_0, \ldots, \beta_{k-1}$ wordt dan bepaald door

$$\min_{\beta} \operatorname{med}_i (y_i - \hat{y}_i)^2$$

met $\hat{y}_i = \hat{\beta}_0 + \hat{\beta}_1 x_{i1} + \ldots + \hat{\beta}_{k-1} x_{ik-1}$.

In dit geval zijn covariabelen toegestaan. Bijzonder prettig hierbij is dat de methode ongevoelig is voor "leverage points" [Belsley, Kuh en Welsch (1980)]: ook in de covariabelen mogen uitschieters voorkomen.

Methode (i) en (ii) geven aangepaste waarden die bij $\hat{y}_i^*$ genoemd worden. Hiermee kan op robuuste wijze σ geschat worden door $\hat{\sigma}^*$. De volgende stap is het verwijderen van waarnemingen y_i als

$$\left| \frac{y_i - \hat{y}_i^*}{\hat{\sigma}} \right| > 2.5.$$

En daarna volgt een klassieke toetsing voor de overgebleven waarnemingen. Als de foutverdeling niet in uitschieters voorziet ($\phi = 0$), dan zal men minder dan 2% van de waarnemingen verliezen. En dat lijkt aanvaardbaar. Werkelijke extreme uitschieters worden hier altijd verwijderd en waar deze in de praktijk doorgaans ontstaan door fouten in het experiment of bij het noteren van de uitkomsten lijkt dit een goede keuze.

5. Uitschieters dempen.

Bij het klassieke regressiemodel wordt $\sum_{i=1}^{N} (y_i - x_i\beta)^2$ geminimaliseerd als functie van β, waarbij x_i de vector $(x_{i0}, \ldots, x_{ik-1})$ voorstelt en β de vector $(\beta_0, \ldots, \beta_{k-1})$. In tegenstelling tot de andere x-en is x_{i0} geen element van een dummy-variabele die de groepen identificeert, maar heeft x_{i0} altijd de waarde 1. Het moge duidelijk zijn dat uitschieters in y de schatting $\hat{\beta}$ in belangrijke mate bepalen; hun toch al grote bijdrage wordt ook nog gekwadrateerd. Men kan de regressie robuuster maken door over te gaan op een andere doelfunctie:

$$\min_{\beta} \sum_{i=1}^{N} \rho\left(\frac{y_i - x_i\beta}{\sigma}\right)$$

In het klassieke geval geldt $\rho(r) = r^2$ maar bij een robuuste methode wordt hiervoor een functie genomen die de bijdrage van grote residuen zeer beperkt. In Holland en Welsch (1977) worden acht verschillende functies ρ genoemd die deze eigenschap hebben. Het minimum van de doelfunctie wordt bereikt als

$$\sum_{i=1}^{N} x_{ij}\, \psi\left(\frac{y_i - x_i\beta}{\sigma}\right) = 0 \quad \text{voor } j = 0, \ldots, k-1$$

waarbij $\psi = \rho'$. Dit stelsel kan bijvoorbeeld worden opgelost door een iteratief herwogen kleinste kwadratenproces [Beaton en Tukey (1974)] met gewichtsfunctie $w(r) = \frac{\psi(r)}{r}$. Beginschattingen voor $\beta_0, \ldots, \beta_{k-1}$ kunnen verkregen worden door een klassieke regressie, waarna σ geschat kan worden als

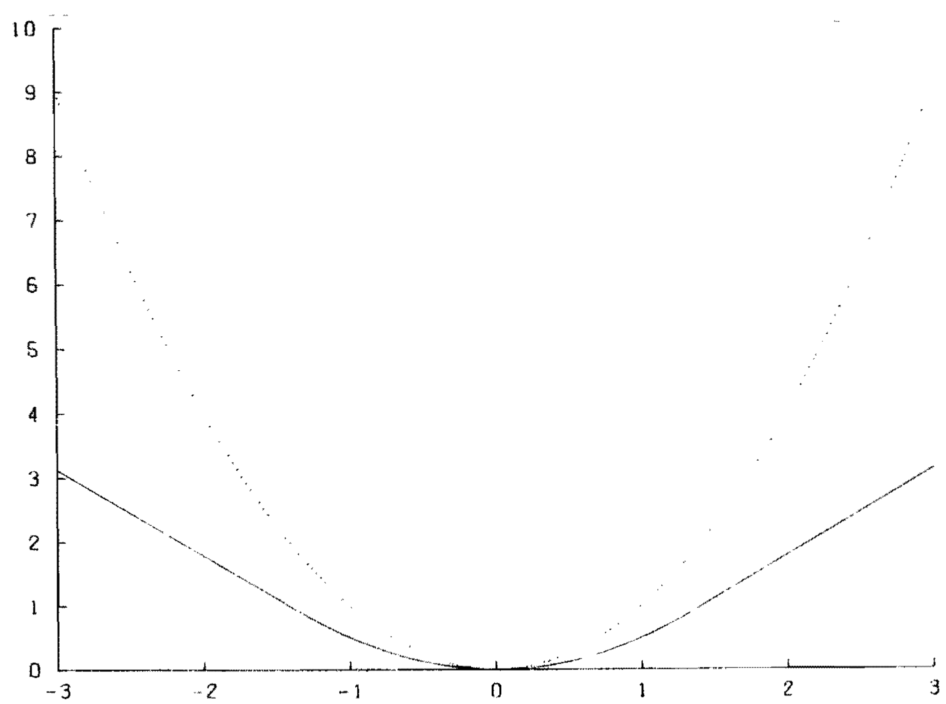
$$\hat{\sigma} = 1.4826 \left[\underset{j}{\text{med}} \left| (y_j - x_j\hat{\beta}) - \underset{i}{\text{med}}\, (y_i - x_i\hat{\beta}) \right| \right]$$

In het algemeen kan convergentie alleen gegarandeerd worden als σ niet mee itereert. Een gunstige uitzondering hierop vormt de ρ van Huber (1973):

$$\rho(r) = \frac{r^2}{2} \quad \text{voor } |r| < H$$

$$\rho(r) = H|r| - \frac{H^2}{2} \quad \text{voor } |r| > H$$

De gevoeligheid voor uitschieters kan worden ingesteld door de keuze van H. Als H = 1.345 is de efficiency 95% voor de normale verdeling.



Bovenstaande figuur geeft de ρ van Huber met H = 1.345. Ter vergelijking is ook de klassieke parabool aangegeven. Het is van geen belang dat Huber rond het centrum al vlakker is door de schalingsfactor ½; alleen de overgang op lineariteit voor $|r| > H$ maakt hem minder gevoelig voor uitschieters. Hoewel Huber zeker niet de grootst mogelijke robuustheid zal opleveren (sommige auteurs prefereren een ρ die constant is voor voldoend grote $|r|$), is deze methode ten opzichte van de niet-robuuste al een aanzienlijke verbetering.

Toepassing van iteratief herwogen kleinste kwadraten met bijvoorbeeld de ρ van Huber op het bekende model met kans ϕ op uitschieters geeft aangepaste waarden $\hat{y}^{\bullet}$ en een schatter $\hat{\sigma}^{\bullet}$ voor σ. Naar analogie met de vorige paragraaf kan men dan natuurlijk overwegen om waarnemingen y_i te verwijderen als:

$$\left| \frac{y_i - \hat{y}_i^{\bullet}}{\hat{\sigma}^{\bullet}} \right| > 2.5$$

Maar een andere aanpak is ook mogelijk. Daartoe gaan we even terug naar het klassieke geval waarbij het model niet in uitschieters voorziet. De toetsingsgrootheid is dan:

$$F = \frac{\sum_{i=1}^{k} n_i(\bar{y}_i - \bar{y})^2/(k-1)}{\sum_{i=1}^{k} \sum_{j=1}^{n_i} (y_{ij} - \bar{y}_i)^2/(N-k)}$$

Het lijkt aantrekkelijk om te streven naar een $F^{\bullet}$ die hieraan ontleend is, maar waar uitschieters weinig invloed op hebben. In de teller ligt de eerste stap voor de hand: vervang $\bar{y}_i$ door $\hat{y}_i^{\bullet}$ (de door robuuste regressie aangepaste waarde). Voor $\bar{y}$ beveelt Huber (1981) in een ruimer model een gewone kleinste kwadratenaanpassing aan met $\hat{y}^{\bullet}$ in plaats van y. In dit geval (zonder covariabelen) komt dat neer op een gewogen gemiddelde:

$$\bar{y}^{\bullet} = \frac{\sum_{i=1}^{k} n_i \hat{y}_i^{\bullet}}{N}$$

Na schaling heeft de zo verkregen teller onder milde voorwaarden asymptotisch een χ^2-verdeling met hetzelfde aantal vrijheidsgraden als bij de klassieke toets.

Eenvoudig valt in te zien dat we er bij de noemer niet zo gemakkelijk vanaf komen: één uitschieter kan voor een zeer grote waarde zorgen waardoor de nulhypothese mogelijk ten onrechte wordt geaccepteerd. Huber (1981) stelt voor de noemer te vervangen door de volgende uitdrukking, waarin de uitschieters gedempt worden:

$$\frac{1}{N-k}\,\frac{c^2 \sum_{i=1}^{N} \psi\left(\frac{r_i}{\hat{\sigma}^{\bullet}}\right)^2 \hat{\sigma}^{\bullet 2}}{\left[\frac{1}{N}\sum_{i=1}^{N} \psi'\left(\frac{r_i}{\hat{\sigma}^{\bullet}}\right)\right]^2} \qquad \text{met } r_i = y_i - \hat{y}_i^{\bullet}$$

Hierbij geldt het volgende:

$$c = 1 + \frac{k\ \hat{\mathrm{var}}(\psi')}{N[\hat{E}(\psi')]^2}$$

$$\hat{E}(\psi') = \frac{1}{N}\sum_{i=1}^{N} \psi'\left(\frac{r_i}{\sigma^{\bullet}}\right)$$

$$\hat{\mathrm{var}}(\psi') = \frac{1}{N}\sum_{i=1}^{N} \left[\psi'\left(\frac{r_i}{\hat{\sigma}^{\bullet}}\right) - \hat{E}(\psi')\right]^2$$

De feitelijke berekening kan wat handiger door over te gaan op een soort pseudo-waarnemingen, maar op de resulterende $F^{\bullet}$ heeft dat geen invloed. Net als in het klassieke geval wordt ook hier H_0: $\mu_1 = \ldots = \mu_k$ verworpen als $F^{\bullet} > F_{N-k}^{k-1}(\alpha)$ bij gekozen onbetrouwbaarheid α. Huber (1981) stelt dat de toetsingsgrootheid redelijk door een F-verdeling wordt benaderd als $n_i \geq 5$ voor $i = 1, \ldots, k$. Deze eis is niet nieuw: voor het gebruik van een χ^2-verdeling bij verdelingsvrije toetsen eist men doorgaans hetzelfde.

In tegenstelling tot de methode uit de vorige paragraaf is de hier beschreven aanpak zeer gevoelig voor "leverage points". Covariabelen zijn wel toegestaan, maar mogen beslist geen uitschieters bevatten. Het proces is generaliseerbaar tot complexere designs, evenveel met interacties. En dat laatste staat bij verdelingsvrije alternatieven nog maar in zijn kinderschoenen [De Kroon en Van der Laan (1981)].

In een stage van Hontelez (1984) op de Technische Hogeschool Eindhoven wordt de robuuste methode van Huber vergeleken met de klassieke toets voor enkelvoudige variantie-analyse. Onder de nulhypothese zijn 32 situaties gesimuleerd, en daarnaast zijn 10 alternatieven onderzocht. Iedere simulatie berust op 500 herhalingen.

Deze studie leidt tot de volgende conclusies:

(i) Als $\phi = 0$ (geen uitschieters) dan is bij Huber de controle over de gekozen onbetrouwbaarheid vrijwel perfect.

(ii) Voor $\phi = 0$ is het onderscheidend vermogen van Huber minder dan van de klassieke methode. De waargenomen verschillen zijn echter zeer klein.

(iii) Als er wel uitschieters zijn is de klassieke methode redelijk robuust, tenzij de uitschieters zich in één groep concentreren. Ook in die gevallen blijft bij Huber de controle over de onbetrouwbaarheid goed, behalve als de groepsgroottes erg verschillen en de uitschieters in een kleine groep vallen.

(iv) In alle gevallen met uitschieters heeft Huber een hoger onderscheidend vermogen dan de klassieke toets, en de verschillen zijn soms indrukwekkend.

6. Conclusies.

Voor definitieve conclusies zijn verdergaande simulaties nodig. Voorlopig lijken de volgende uitspraken verantwoord:

- Adaptieve verdelingsvrije toetsing vergt grotere steekproeven dan in de praktijk doorgaans voorkomen. Bovendien eist deze aanpak een specifiek uitschietermodel. Als dit model niet goed past zal het onderscheidend vermogen daaronder lijden.
- Trimming en winsorizing met vaste parameter zijn af te raden indien er meer dan twee groepen zijn.
- De methode van Huber lijkt zeer aantrekkelijk en is bovendien generaliseerbaar tot complexere gevallen. Met extreme uitschieters in relatief kleine steekproeven heeft deze methode nog wel problemen. Uitschieters in covariabelen zijn desastreus.
- Bij het opsporen en verwijderen van uitschieters volgens de methode van Leroy en Rousseeuw kan een grotere robuustheid verwacht worden. Het onderscheidend vermogen zal vermoedelijk wat lager zijn. Een systematisch vergelijkend simulatie-onderzoek lijkt hier zeer gewenst.

7. Literatuur.

[1] Beaton, A.E. and J.W. Tukey (1974)
The fitting of power series, meaning polynomials, illustrated on band-spectroscopic data
Technometrics (16) 147-185

[2] Belsley, D.A., E. Kuh and R.E. Welsh (1980)
Regression Diagnostics
John Wiley & Sons, New York

[3] Brown, G.W. and A.M. Mood (1950)
On median tests for linear hypotheses
Proc. 2nd. Berkeley Symposium, 159-166

[4] Dixon, W.J. and J.W. Tukey (1968)
Approximate behaviour of the distribution of winsorized t
Technometrics (10) 83-98

[5] Fung, K.Y. and S.M. Rahman (1980)
The two-sample winsorized t
Communications in Statistics (B9. no. 4), 337-347

[6] Hajek, J. (1969)
A course in nonparametric statistics
Holden-Day, San Francisco

[7] Hodges, J.L. and E.L. Lehman (1961)
Comparison of the normal scores and Wilcoxon tests
Proc. 4th. Berkely Symposium, 307-317

[8] Holland, P.W. and R.E. Welsh (1977)
Robust regression using iteratively reweighted least-squares
Communications in Statistics A6(9), 813-827

[9] Hontelez, J. (1984)
Een uitschieter-resistente procedure voor enkelvoudige klassieke variantie-analyse
Computing Centre Note 21, Eindhoven University of Technology
(stage o.l.v. prof.dr. R. Doornbos en drs. J.B. Dijkstra)

[10] Huber, P.J. (1972)
Robust statistics, a review
The Annals of Math. Stat. (43, no. 4), 1041-1067

[11] Huber, P.J. (1973)
Robust regression: asymptotics, conjectures, and Monte Carlo
Ann. Statist. (1), 799-821

[12] Huber, P.J. (1981)
Robust Statistics
John Wiley & Sons, New York

[13] De Kroon, J. and P. van der Laan (1981)
Distribution-free test procedures in two-way layouts; a concept of rank-interaction
Statistica Neerlandica (35, no. 4), 189-213

[14] Van der Laan, P. and J. Oosterhoff (1967)
Experimental determination of the power functions of the two-sample rank tests of Wilcoxon, Van der Waerden and Terry by Monte Carlo techniques
Statistica Neerlandica (21, no. 1), 55-68

[15] Leroy, A. and P. Rousseeuw (1985)
A multiple regression technique for detecting outliers
Kwantitatieve Methoden (6, no. 18), 41-58

[16] Linders, H. (1986)
Geen publikatie maar een volgeschreven schoolbord
Zie appendix

[17] Van der Waerden, B.L. (1952)
Order tests for the two-sample problem and their power
Indagationes Math. (14), 453-458

[18] Yuen, K.K. and W.J. Dixon (1973)
The approximate behaviour and performance of the two-sample trimmed t
Biometrika (60), 369-374

Appendix: schatting van σ^2, ϕ en a.

Het model is $\underline{y}_{ij} = \mu_i + \underline{e}_{ij}$ met voor $\underline{e}_{ij}$ de volgende verdeling: met kleine kans ϕ geldt $\underline{e}_{ij} \sim N(0, a\sigma^2)$ voor a >> 1 en met kans $1-\phi$ geldt $\underline{e}_{ij} \sim N(0, \sigma^2)$. De variantie σ^2 kan robuust geschat worden, bijvoorbeeld zoals in paragraaf 4 en 5 beschreven is. Deze aanpak levert tevens residuen op: $e_i = y_i - \hat{y}_i$ voor $i = 1, \ldots, N$. En hiermee zijn ϕ en a te schatten, zoals hieronder wordt beschreven.

Als een stochastische variabele $\underline{x}$ verdeeld is als $N(0, \sigma^2)$, dan geldt $E|\underline{x}| = \sigma\sqrt{\frac{2}{\pi}}$ en $E\underline{x}^2 = \sigma^2$. Deze momenten zijn te schatten als $\frac{1}{N}\sum_{i=1}^{N} |e_i|$ en $\frac{1}{N}\sum_{i=1}^{N} e_i^2$ respectievelijk. Dit geeft met de bekende dichtheid van de foutverdeling de volgende twee vergelijkingen

$$\frac{1}{N}\sum_{i=1}^{N} |e_i| = (1-\phi)\,\sigma\sqrt{\frac{2}{\pi}} + \phi\sigma\sqrt{a\frac{2}{\pi}}$$

$$\frac{1}{N}\sum_{i=1}^{N} e_i^2 = (1-\phi)\sigma^2 + \phi a\sigma^2$$

Na substitutie van de robuust geschatte $\hat{\sigma}^2$ blijven de parameters ϕ en a over en door dit stelsel op te lossen krijgt men de schatters $\hat{\phi}$ en $\hat{a}$.