stichting
mathematisch
centrum

MC

AFDELING MATHEMATISCHE STATISTIEK   SD 115/83   SEPTEMBER

B.F. SCHRIEVER

ATTITUDE-ONDERZOEK OPENBAAR VERVOER IN ROTTERDAM EN OMGEVING: STATISTISCHE ANALYSE ENQUÊTE NAAR DENKBEELDIGE VERPLAATSINGEN

Attitude research of public transportation in Rotterdam and Suburbs'.
Statistical analysis inquiry of hypothetical movements

kruislaan 413 1098 SJ amsterdam

Printed at the Mathematical Centre, Kruislaan 413, Amsterdam, The Netherlands.

The Mathematical Centre, founded 11 February 1946, is a non-profit institution for the promotion of pure and applied mathematics and computer science. It is sponsored by the Netherlands Government through the Netherlands Organization for the Advancement of Pure Research (Z.W.O.).

Attitude-onderzoek openbaar vervoer in Rotterdam en omgeving: statistische analyse enquête naar denkbeeldige verplaatsingen

door

B.F. Schriever

SAMENVATTING

In dit rapport worden door reizigers gegeven beoordelingen van denkbeeldige verplaatsingen met het openbaar vervoer geanalyseerd met behulp van een variantie-analyse model. In het model is een afhankelijkheidsstruktuur tussen de beoordelingen verondersteld. De parameters van het model worden op eèn asymptotisch efficiente manier geschat.

TREFWOORDEN: *Variantie-analyse; Algemeen lineair model; Asymptotisch efficiënte schatters*

1. INLEIDING

Kennis van de attitudes van openbaar-vervoerreizigers is van nut bij de bevordering van openbaar-vervoergebruik en bij de planning van openbaar-vervoervoorzieningen. Een belangrijk aspect hiervan is de reactie van reizigers op mogelijke, veelal nog niet gerealiseerde, wijzigingen in de reisomstandigheden, bijvoorbeeld tariefswijzigingen en wijzigingen in de dienstregeling. Dit rapport beschrijft een statistische analyse van een enquête naar denkbeeldige situaties van openbaar-vervoer in Rotterdam en omgeving. De enquête is aan twee groepen respondenten voorgelegd, te weten een groep ontstaan door een steekproef uit de openbaar-vervoerreizigers te nemen en een (minder grote) contrastgroep ontstaan door een steekproef uit de niet-openbaar-vervoergebruikers. Een beschrijving van de steekproeftrekkingen en de opzet van de enquête wordt gegeven in het DHV-rapport [1].

In paragraaf 2 van het onderhavige rapport wordt de proefopzet en een bruikbaar wiskundig model beschreven. Een theoretische behandeling van de statistische methode, waarmee de bestanden met waarnemingen geanalyseerd zijn, wordt in paragraaf 3 gegeven. Om de resultaten van de analyses te kunnen interpreteren is het niet nodig deze paragraaf te lezen. De resultaten van de analyses van het openbaar-vervoergebruikersbestand en het bestand van de contrastgroep worden in de paragraven 4 en 5 besproken. Paragraaf 6 besluit met enkele opmerkingen van algemene aard.

2. PROEFOPZET EN MODEL

In de enquête zijn respondenten gevraagd om een aantal denkbeeldige verplaatsingen (situaties) met het openbaar-vervoer in Rotterdam te beoordelen. Het accent is hierbij gelegd op regelmatige dagelijkse verplaatsingen, waarbij de reizigers geacht worden een maand-abonnement te kopen als ze het openbaar-vervoer (beperkt tot bus) kiezen. Het enige alternative vervoermiddel bij deze denkbeeldige verplaatsingen is de fiets. Verder zijn nog enkele a pecten van de verplaatsingen voor alle denkbeeldige situaties hetzelfde: dagelijks vertrek van huis naar het bestemmingsadres, er is een bushalte op 100 meter van huis en ook van het bestemmingsadres, de verplaatsing met de bus is over twee zones en bovendien is er 's ochtends altijd een zitplaats in de bus. De denkbeeldige verplaatsingen zijn verder geconcretiseerd door wisselende aspecten, welke ontstaan door een zestal factoren op verschillende niveaus in te stellen. Deze factoren met bijbehorende niveaus zijn:

1. aantal malen overstappen	: 0 of 1 keer.
2. reistijd fiets	: 15, 30 of 45 minuten.
3. zitplaats terugreis	: al dan geen zitplaats.
4. reistijd bus	: laag of hoog.
5. volgtijd bus	: 5, 10 of 15 minuten.
6. kosten maandabonnement (bij vol tarief)	: ƒ34,-; ƒ51,- of ƒ59,50.

Het niveau van de factor "reistijd bus" is voor de respondenten uitgedrukt in minuten (afgeleid uit de niveaus van de faktoren "aantal malen overstappen", "reistijd fiets" en een omrijfactor (0% of 50%); de factor "reistijd bus" moet dus geïnterpreteerd worden als een omrijfactor). Op deze manier ontstaan 216 denkbeeldige situaties die gekarakteriseerd worden door de combinaties van de niveaus van de zes factoren. Schematies is dit weergegeven in tabel 1. Hierin correspondeert iedere denkbeeldige situatie met één cel. De waargenomen variabele is het door de respondent opgegeven kanspercentage (tussen 0% en 100%) dat hij in de betreffende denkbeeldige situatie de bus zal kiezen in plaats van de fiets, en dus een maandabonnement zal kopen.

De 216 denkbeeldige situaties zijn verdeeld over 12 blokken van ieder 18 situaties. In ieder blok zijn de niveaus van de factoren "aantal malen overstappen", "reistijd fiets" en "zitplaats terugreis" vastgehouden terwijl de niveaus van de overige factoren varieëren. In tabel 1 komt een blok dus overeen met een horizontale rij cellen. Een respondent beoordeelt niet alle 216 situaties, maar de 18 situaties die bij een bepaald blok horen. Het is de bedoeling dat bij het nemen van de steekproef de respondenten aselect over de blokken verdeeld worden.

In paragraaf 6 worden enkele opmerkingen bij de proefopzet gemaakt.

Het model ziet er als volgt uit:

$$\left.\begin{aligned} \underline{z}_{i_{abc}def} = {} & \mu + \alpha_a + \beta_b + \gamma_c + \delta_d + \varepsilon_e + \zeta_f + \\ & + \alpha\beta_{ab} + \alpha\gamma_{ac} + \alpha\delta_{ad} + \alpha\varepsilon_{ae} + \alpha\zeta_{af} + \\ & + \beta\gamma_{bc} + \beta\delta_{bd} + \beta\varepsilon_{be} + \beta\zeta_{bf} + \\ & + \gamma\delta_{cd} + \gamma\varepsilon_{ce} + \gamma\zeta_{cf} + \\ & + \delta\varepsilon_{de} + \delta\zeta_{df} + \\ & + \varepsilon\zeta_{cf} + \\ & + \underline{v}_{i_{abc}def} \end{aligned}\right\} \left.\begin{array}{l} d = 1,2; \\ e = 1,2,3; \\ f = 1,2,3; \\ i_{abc} = 1,\ldots,n_{abc}; \end{array}\right\} \begin{array}{l} a=1,2; \\ b=1,2,3; \\ c=1,2. \end{array} \tag{1}$$

met de randvoorwaarden:

$$\begin{gathered} \sum_a \alpha_a = \sum_b \beta_b = \sum_c \gamma_c = \sum_d \delta_d = \sum_e \varepsilon_e = \sum_f \zeta_f = 0, \\ \sum_a \alpha\beta_{ab} = \sum_b \alpha\beta_{ab} = \sum_a \alpha\gamma_{ac} = \sum_c \alpha\gamma_{ac} = \ldots = \sum_e \varepsilon\zeta_{ef} = \sum_f \varepsilon\zeta_{ef} = 0. \end{gathered} \tag{2}$$

Hier geeft de index a aan welk niveau van de factor "aantal malen overstappen" bij de denkbeeldige situatie hoort en analoog geven de indices b,c,d,e en f aan op welke niveaus de overige factoren ingesteld zijn. Deze relatie tussen waarden van de indices en niveaus van de factoren is als volgt:

a = 1,2	: respectivelijk niet en wel overstappen;
b = 1,2,3	: reistijd fiets resp. 15, 30 en 45 min.;
c = 1,2	: resp. geen of wel zitplaats op terugreis;
d = 1,2	: reistijd bus resp. laag en hoog;
e = 1,2,3	: volgtijd bus resp. 5, 10 en 15 min.;
f = 1,2,3	: kosten abonnement resp. ƒ34,-; ƒ51,- en ƒ59,50.

Zo correspondeert een blok met precies één combinatie van de indices a,b en c. Verder stelt n_{abc} het aantal respondenten voor in het blok wat correspondeert met de indexcombinatie (a,b,c) en geeft i_{abc} de i^e respondent uit dit blok aan. Het totale aantal respondenten is $n = \Sigma_a \Sigma_b \Sigma_c\, n_{abc}$.

De symbolen in (1) en (2) hebben de volgende betekenis:

$\underline{z}_{i_{abc}def}$	: de beoordeling van de denkbeeldige situatie, welke correspondeert met de indexcombinatie (a,b,c,d,e,f), door respondent i uit het blok corresponderende met (a,b,c). Om een onderscheid te maken tussen stochastische grootheden en realisaties ervan worden stochastische grootheden onderstreept.
μ	: de algemeen gemiddelde beoordeling van alle situaties.
α_a	: een afwijking van de algemeen gemiddelde beoordeling voor het niveau a van de factor "aantal malen overstappen" (hoofdeffect), dus:
$\mu+\alpha_a$	: de algemeen gemiddelde beoordeling voor het niveau a van de factor "aantal malen overstappen".
β_b	: een afwijking van de algemeen gemiddelde beoordeling voor het niveau b van de factor "reistijd fiets".
$\gamma_c,\delta_d,\varepsilon_e,\zeta_f$	: analoog.
$\alpha\beta_{ab}$	: een afwijking van de algemeen gemiddelde beoordeling voor de combinatie van niveau a van de factor "aantal malen overstappen" en niveau b van de factor "reistijd fiets" (interactie), dus:
$\mu+\alpha_a+\beta_b+\alpha\beta_{ab}$	: de algemeen gemiddelde beoordeling van alle situaties corresponderende met de indexcombinatie (a,b).
$\alpha\gamma_{ac},\ldots,\varepsilon\zeta_{ef}$	: analoog.
$\underline{v}_{i_{abc}def}$	: een stochastische storingsterm voor de beoordeling van de situatie, corresponderende met de indexcombinatie (a,b,c,d,e,f), door de respondent i_{abc} (residu).

De interpretatie van formule (1) is dat iedere respondent de beoordeling van een denkbeeldige situatie opbouwt uit een algemene beoordeling, effecten ten gevolge van niveaus (hoofdeffecten) en combinaties van niveaus (interacties) van de factoren plus een storingsterm. De restrictie (2) op de hoofdeffecten en interacties is nodig om deze parameters uniek te kunnen bepalen.

De stochastische grootheden $\underline{v}_{i_{abc}def}$ worden verondersteld een normale verdeling te bezitten met verwachting nul en een covariantiestructuur, zodanig dat de 18-dimensionale vectoren $\underline{v}_{i_{abc}} = (\underline{v}_{i_{abc}111}, \underline{v}_{i_{abc}112}, \ldots, \underline{v}_{i_{abc}232})'$ verdeeld zijn volgens:

(3) $\quad L(\underline{v}_{i_{abc}}) = N_{18}(0, \ddagger_{abc})$; onderling onafhankelijk voor

$i_{abc} = 1, \ldots, n_{abc}$ en $a = 1,2$; $b = 1,2,3$; $c = 1,2$.

De veronderstelling van normaliteit is in principe nodig voor de uit te voeren schattingen en toetsen. Aangezien er van een behoorlijke hoeveelheid respondenten waarnemingen zijn, is het niet zo'n probleem als de verdeling van $\underline{v}_{i_{abc}}$ in werkelijkheid niet normaal is. Immers de gebruikte schatters en toetsen (zie paragraaf 3) zijn gebaseerd op gemiddelden, welke volgens de centrale limiet stelling asymptotisch normaal verdeeld zijn, en dus zullen de schatters en toetsen asymptotisch dezelfde eigenschappen hebben als onder normaliteit. De in (3) veronderstelde covariantiestructuur heeft als betekenis dat alle respondenten de situaties onafhankelijk van elkaar beoordelen, maar dat de 18 beoordelingen van iedere respondent wel afhankelijk (gecorreleerd) zijn. Deze afhankelijkheid is zo dat beoordelingen van respondenten uit eenzelfde blok dezelfde covariantiematrix $\ddagger_{abc}$ hebben. De afhankelijkheidsstructuur tussen de beoordelingen is dus zo algemeen mogelijk gehouden. Aan de hand van de waarnemingen zal onderzocht worden of een speciale eenvoudige structuur verondersteld mag worden.

Het is mogelijk dat de respondenten tijdens het beoordelen van de 18 situaties een zekere vermoeidheid of vaardigheid krijgen (volgorde-effect). Indien alle respondenten de situaties altijd in dezelfde volgorde beoordelen is een eventueel volgorde-effect niet meer te onderscheiden van effecten ten gevolge van de factoren "reistijd bus", "volgtijd bus" en "kosten maandabonnement". Om deze reden zou het verstandig zijn de volgorde van de 18 situaties voor iedere respondent middels loting te bepalen (randomisatie). Aangezien dit technisch niet uitvoerbaar is, is in de proefopzet met randomisatie van de situatievolgorde voor ieder blok volstaan (zie ook DHV-rapport [1]).

De verdeling van een stochastische grootheid $\underline{x}$ wordt met $L(\underline{x})$ genoteerd.
Het symbool N staat voor de (multivariatie) normale verdeling.

3. SCHATTERS EN TOETSEN BIJ HET MODEL

In deze paragraaf wordt een theoretische beschrijving gegeven van de statistische methode waarmee schattingen voor de parameters (d.w.z. hoofd-effecten en interacties) verkregen worden en toetsen van hypotheses over de parameters uitgevoerd worden. Enige algemene statistische kennis wordt bekend verondersteld; deze kan eventueel gevonden worden in RAO (1973).

Eerst wordt het model in een andere, equivalente, vorm geschreven. In deze paragraaf wordt een blok niet meer met de drie indices (a,b,c), maar slechts met één index k aangeduid (k=1,...,12). Noteer met $\underline{z}_{i_k}$ de vector ter lengte 18 met de beoordelingen $\underline{z}_{i_k def}$ ($=\underline{z}_{i_{abc} def}$) van respondent i_k in de volgorde waarbij de index f het snelst loopt, gevolgd door de indices e en d. Analoog wordt met $\underline{v}_{i_k}$ (voorheen $\underline{v}_{i_{abc}}$) de vector met bijbehorende residuen genoteerd. Laat β de vector met alle parameters zijn. Wegens de randvoorwaarden (2) kan worden volstaan met de volgende 43 parameters:

$$
\begin{aligned}
\beta = (&\mu,\alpha_1,\beta_1,\beta_2,\gamma_1,\delta_1,\varepsilon_1,\varepsilon_2,\zeta_1,\zeta_2,\\
&\alpha\beta_{11},\alpha\beta_{12},\alpha\gamma_{11},\alpha\delta_{11},\alpha\varepsilon_{11},\alpha\varepsilon_{12},\alpha\zeta_{11},\alpha\zeta_{12},\\
&\beta\gamma_{11},\beta\gamma_{21},\beta\delta_{11},\beta\delta_{21},\beta\varepsilon_{11},\beta\varepsilon_{12},\beta\varepsilon_{21},\beta\varepsilon_{22},\\
&\beta\zeta_{11},\beta\zeta_{12},\beta\zeta_{21},\beta\zeta_{22},\gamma\delta_{11},\gamma\varepsilon_{12},\gamma\varepsilon_{12},\gamma\zeta_{11},\\
&\gamma\zeta_{12},\delta\varepsilon_{11},\delta\varepsilon_{12},\delta\zeta_{11},\delta\zeta_{12},\varepsilon\zeta_{11},\varepsilon\zeta_{12},\varepsilon\zeta_{21},\varepsilon\zeta_{22})'.
\end{aligned}
\tag{4}
$$

De formules (1) en (2) kunnen nu geschreven worden als:

$$
\underline{z}_{i_k} = D_k\beta + \underline{v}_{i_k}\,; \quad i_k = 1,\dots,n_k; \quad k = 1,\dots,12,
\tag{5}
$$

waarbij D_k een designmatrix is van de afmetingen 18×43. Deze designmatrix heeft als elementen 0, 1 en −1, en is zo gekozen dat het gemiddelde wat volgens (1) bij de beoordeling $\underline{z}_{i_k def}$ hoort samengesteld wordt uit de componenten van β. Als voorbeeld is achter in het rapport de matrix D_1 gegeven. De residuen $\underline{v}_{i_k}$ worden nog steeds verondersteld de normale verdeling (3) te hebben.

Schatten van de parameters β.

Een methode die in het algemeen nauwkeurige schatters oplevert is de methode van de grootste aannemelijkheid (maximum likelihood). Door eerst deze methode te bestuderen kan tenslotte een eenvoudigere, maar even goede, schatter voor de vector β gevonden worden.

Uit (5) en (3) volgt dat voor vectoren met realisaties z_{i_k}, $i_k = 1,\ldots,n_k$, $k = 1,\ldots,12$ (soms kort samengenomen in één grote vector z) de aannemelijkheidsfunktie, en tevens dichtheid, gegeven wordt door:

(6) $$L(z;\beta,\Sigma) = \prod_{k=1}^{12} \prod_{i_k=1}^{n_k} (2\pi)^{-9}(\det\Sigma_k)^{-\frac{1}{2}}\exp\{-\tfrac{1}{2}(z_{i_k}-D_k\beta)'\Sigma_k^{-1}(z_{i_k}-D_k\beta)\}.$$

De meest aannemelijke schattingen $\hat{\beta}$ en $\hat{\Sigma}_k$ voor de parameters β en Σ_k (k=1,...,12) zijn gedefinieerd als die waarden welke voor de gegeven realisaties z de functie $L(z;\beta,\Sigma)$ maximaliseren. Analoog aan MUIRHEAD (1982), p. 83-84 of RAO (1973), p. 529-531 volgt dat:

(7) $$\hat{\beta} = \left(\sum_{k=1}^{12} D_k' \hat{\Sigma}_k^{-1} D_k\right)^{-1} \sum_{k=1}^{12} D_k' \hat{\Sigma}_k^{-1} \bar{z}_k$$

en

(8) $$\hat{\Sigma}_k = \frac{1}{n_k}\sum_{i_k=1}^{n_k} (z_{i_k}-D_k\hat{\beta})(z_{i_k}-D_k\hat{\beta})', \quad k = 1,\ldots,12,$$

waarbij

(9) $$\bar{z}_k = \frac{1}{n_k}\sum_{i_k=1}^{n_k} z_{i_k}, \quad k = 1,\ldots,12.$$

Voor gegeven waarnemingen z_{i_k} is het niet mogelijk om $\hat{\beta}$ direkt uit (7) en (8) te halen. Het stelsel vergelijkingen (7) en (8) is wel iteratief op te lossen. Gezien de afmetingen van het probleem zal het erg veel rekentijd vergen voordat een stabiele oplossing van het iteratieproces gevonden is.

Het probleem om de meest aannemelijke schattingen van β te berekenen is moeilijk omdat de covariantiematrices Σ_k (k=1,...,12) onbekend zijn. Immers, indien Σ_k (k=1,...,12) bekend zouden zijn dan zou de meest aannemelijke schatting van β gegeven worden door:

$$\beta^* = \Big(\sum_{k=1}^{12} D_k' \Sigma_k^{-1} D_k\Big)^{-1} \sum_{k=1}^{12} D_k' \Sigma_k^{-1} \bar{z}_k. \tag{10}$$

Voor het probleem met bekende covariantiematrices is β^* de zuivere schatter met minimale variantie (zie RAO (1973), p. 223).

In het geval van onbekende covariantiematrices kan een eenvoudig te berekenen schatter $\underline{\tilde{\beta}}$ van β aangegeven worden, welke asymptotisch even goed is als $\underline{\beta}^*$ (en ook evengoed als $\underline{\hat{\beta}}$). Hiervoor wordt eerst een consistente schatting $\tilde{\Sigma}_k$ van Σ_k gegeven en vervolgens wordt, analoog aan (10), β geschat door

$$\tilde{\beta} = \Big(\sum_{k=1}^{12} D_k' \tilde{\Sigma}_k^{-1} D_k\Big)^{-1} \sum_{k=1}^{12} D_k' \tilde{\Sigma}_k^{-1} \bar{z}_k. \tag{11}$$

Uit onderstaande stelling 1 volgt nu dat $\underline{\tilde{\beta}}$ consistent is en asymptotisch even goed is als de zuiver schatter met minimale variantie; dat wil zeggen dat $\underline{\tilde{\beta}}$ asymptotisch efficient is. (Dit geldt ook voor $\underline{\hat{\beta}}$.) Als consistente schatting voor Σ_k is:

$$\tilde{\Sigma}_k = \frac{1}{n_k} \sum_{i_k=1}^{n_k} (z_{i_k} - \bar{z}_k)(z_{i_k} - \bar{z}_k)', \quad k = 1,\ldots,12, \tag{12}$$

gebruikt. Merk op dat de stochastische grootheden $\underline{\hat{\beta}}$, $\underline{\tilde{\beta}}$, $\underline{\beta}^*$, $\underline{\hat{\Sigma}}_k$, $\underline{\tilde{\Sigma}}_k$ en $\underline{\bar{z}}_k$ van de steekproefgroottes n_k $(k=1,\ldots,12)$ afhangen.

<u>STELLING 1</u>. *Indien* $n_k \to \infty$, *zodanig dat* $n/n_k \to$ *constante* $(k=1,\ldots,12)$, *dan convergeren de stochastische grootheden* $\sqrt{n}\,(\underline{\tilde{\beta}}-\beta)$, $\sqrt{n}\,(\underline{\beta}^*-\beta)$ *en* $\sqrt{n}\,(\underline{\hat{\beta}}-\beta)$ *in verdeling naar dezelfde normale verdeling met verwachting nul en covariantiematrix* $(\Sigma_{k=1}^{12} D_k' \Sigma_k^{-1} D_k)^{-1}$.

<u>BEWIJS</u>. Consistentie van de schatter $\underline{\tilde{\Sigma}}_k$ wil zeggen dat deze schatter in kans naar Σ_k convergeert als $n_k \to \infty$. Hieruit volgt dat:

$$\Big(\sum_{k=1}^{12} D_k' \underline{\tilde{\Sigma}}_k^{-1} D_k\Big)^{-1} D_k' \underline{\tilde{\Sigma}}_k^{-1} \to_p \Big(\sum_{k=1}^{12} D_k' \Sigma_k^{-1} D_k\Big)^{-1} D_k' \Sigma_k^{-1} \tag{13}$$

als $n_k \to \infty$ voor $k = 1,\ldots,12$. Beschouw vervolgens:

De notatie $\to_p$ staat voor convergentie in kans, en $\xrightarrow{L}$ voor convergentie in verdeling.

$$\sqrt{n}\ (\tilde{\underline{\beta}}-\underline{\beta}^*) = \sqrt{n}\ \{(\tilde{\underline{\beta}}-\beta) - (\underline{\beta}^*-\beta)\} =$$

(14)
$$= \sum_{k=1}^{12} \sqrt{\frac{n}{n_k}}\{(\sum_{k=1}^{12} D_k'\tilde{\Sigma}_k^{-1}D_k)^{-1}D_k'\tilde{\Sigma}_k - (\sum_{k=1}^{12} D_k'\Sigma_k^{-1}D_k)^{-1}D_k'\Sigma_k\}.$$
$$.\ \sqrt{n_k}\ (\bar{\underline{z}}_k - D_k\beta).$$

Dit convergeert in kans naar nul omdat $\sqrt{n_k}\ (\bar{\underline{z}}_k - D_k\beta)$ een limietverdeling heeft en omdat, volgens (13), de uitdrukking tussen accolades in kans naar nul gaat. Dus $\sqrt{n}\ (\tilde{\underline{\beta}}-\beta)$ en $\sqrt{n}\ (\underline{\beta}^*-\beta)$ hebben dezelfde limietverdeling. Analoog volgt, omdat $\hat{\underline{\Sigma}}_k$ een consistente schatter is van Σ_k, dat de limietverdeling van $\sqrt{n}\ (\hat{\underline{\beta}}-\beta)$ en $\sqrt{n}\ (\underline{\beta}^*-\beta)$ ook dezelfde is.
Uit (3) volgt dat:

(15) $$L(D_k'\Sigma_k^{-1}\underline{z}_{i_k}) = N_{43}(D_k'\Sigma_k^{-1}D_k\beta, D_k'\Sigma_k^{-1}D_k),\ \text{o.o},\ i_k=1,\dots,n_k;\ k=1,\dots,12,$$

zodat volgens de centrale limietstelling:

(16) $$\sqrt{n}\ D_k'\ \Sigma_k^{-1}\ (\bar{\underline{z}}_k-D_k\beta) \xrightarrow{L} N_{43}(0,D_k'\Sigma_k^{-1}D_k),\quad k=1,\dots,12.$$

Dit impliceert, wegens onafhankelijkheid van $\bar{\underline{z}}_k$ voor $k = 1,\dots,12$, dat:

(17)
$$\sqrt{n}\ (\underline{\beta}^*-\beta) = (\sum_{k=1}^{12} D_k'\Sigma_k^{-1}D_k)^{-1}\sum_{k=1}^{12}\sqrt{n}\,D_k'\Sigma_k^{-1}(\bar{\underline{z}}_k-D_k\beta) \xrightarrow{L}$$
$$N_{43}(0,(\sum_{k=1}^{12} D_k'\Sigma_k^{-1}D_k)^{-1}). \quad \square$$

Resultaten van dit type zijn niet verrassend; met name wordt in de econometrische literatuur voor veel verschillende regressie problemen bewezen dat met behulp van consistente schattingen van (co)varianties en de methode van de grootste aannemelijkheid asymptotisch efficiente schatters voor de regressie parameters verkregen worden. Dit verschijnsel is te verklaren uit een algemeen resultaat van DZHAPARIDZE (1983) en het feit dat de meest aannemelijke schatters voor de regressie parameters en de (co)varianties onafhankelijk zijn.

Toetsen van hypotheses over de parameter β.

Beschouw hypotheses over componenten van de vector β van de vorm:

(18) $$H_0:\ A\beta = 0,$$

waarbij A een $\ell \times 43$ matrix is met rang gelijk aan ℓ. Met (18) kunnen alle hypotheses over afwezigheid van hoofdeffecten of interacties geformuleerd worden. Bijvoorbeeld, door voor de matrix A te kiezen:

(19) $$A = (0,1,0,\ldots,0),$$

wordt de hypothese (18) gelijk aan:

(20) $$H_0:\ \alpha_1 = 0.$$

Wegens de randvoorwaarden (2) is dit equivalent aan:

(21) $$H_0:\ \alpha_1 = \alpha_2 = 0;$$

dat wil zeggen de factor "aantal malen overstappen" heeft geen invloed op de beoordeling van de situaties.

Het uitvoeren van de aannemelijkheidsquotient (likelihood ratio) toets is, evenals het berekenen van de meest aannemelijke schattingen $\hat{\beta}$ en $\hat{\Sigma}$, niet mogelijk. Uit de onderstaande stelling blijkt dat er een op $\underline{\tilde{\beta}}$ en $\tilde{\Sigma}_k$ gebaseerde toets bestaat, welke asymptotisch equivalent is aan de aannemelijkheidsquotient toets en bovendien in zekere zin asymptotisch optimaal is.

STELLING 2. *De toets voor de nulhypothese* (18) *welke verwerpt als*

(22) $$n\tilde{\beta}'A'\Big(A\Big(\sum_{k=1}^{12} D_k'\tilde{\Sigma}_k^{-1}D_k\Big)^{-1}A'\Big)^{-1}A\tilde{\beta}$$

groter is dan het rechter α_0-punt van een χ^2-verdeling met ℓ vrijheidsgraden, is asymptotisch equivalent aan de aannemelijkheidsquotient toets. Bovendien zijn beide toetsen consistent tegen alle vaste alternatieven en is het asymptotisch onderscheidingsvermogen tegen naburige (contiguous) *alternatieven gelijk aan dat van de uniform meest onderscheidende invariante toets.*

BEWIJS. In RAO (1973), p. 418-419, wordt bewezen dat de aannemelijkheidsquotient toets asymptotisch equivalent is aan Wald's toets. Asymptotisch equivalent is hier bedoeld in de zin dat beide toetsingsgrootheden (of transformaties ervan) onder de nulhypothese, en dus ook onder naburige alternatieven, dezelfde limiet verdeling hebben. Wald's toets verwerpt in het onderhavige geval voor "grote waarden" van

$$(23) \qquad n\hat{\beta}'A'\Gamma^{-1}A\hat{\beta}.$$

waarbij Γ de covariantiematrix is van de asymptotische nulhypothese verdeling van $\sqrt{n}\, A\hat{\underline{\beta}}$. Het volgt uit stelling 1 en het feit dat:

$$(24) \qquad A\Big(\sum_{k=1}^{12} D_k'\Sigma_k^{-1}D_k\Big)^{-1}A,$$

een consistente schatter is van Γ, dat de toets gebaseerd op (22) asymptotisch equivalent is aan de toets gebaseerd op (23). Eveneens volgt uit stelling 1 dat de limiet verdeling van

$$(25) \qquad n\tilde{\underline{\beta}}'A'\Big(A\Big(\sum_{k=1}^{12} D_k'\Sigma_k^{-1}D_k\Big)^{-1}A'\Big)^{-1}A\tilde{\underline{\beta}}$$

onder de nulhypothese (18) een chi-kwadraat verdeling met ℓ vrijheidsgraden is. Onder vaste alternatieven verdwijnt de massa van de verdeling van (25) naar oneindig als $n \to \infty$; dat wil zeggen dat de toets gebaseerd op (22) consistent is tegen alle vaste alternatieven.

Het toetsingsprobleem is ook asymptotisch equivalent aan eenzelfde probleem maar met bekende covariantiematrices Σ_k. Voor dit probleem wordt in LEHMANN (1959), p. 311, aangetoond dat het asymptotisch onderscheidingsvermogen tegen naburige alternatieven gelijk is aan dat van de uniform meest onderscheidende invariante toets. □

Berekenen van schattingen en uitvoeren van toetsen.

De berekening van de schatting $\tilde{\beta}$ en het uitvoeren van een (22) asymptotisch equivalente toets kan als volgt met een standaard regressie computerprogramma gebeuren. Na berekening van $\tilde{\Psi}_k$, $k = 1,\ldots,12$, worden de waarnemingen en designmatrices getransformeerd volgens:

$$
(26) \qquad
\begin{aligned}
y_{i_k} &= \tilde{\Psi}_k^{-\frac{1}{2}} z_{i_k}, \quad i_k = 1,\ldots,n_k;\ k = 1,\ldots,12 \\
X_k &= \tilde{\Psi}_k^{-\frac{1}{2}} D_k, \quad k = 1,\ldots,12.
\end{aligned}
$$

Formule (5) wordt dan gelijk aan:

$$
(27) \qquad \underline{y}_{i_k} = X_k\beta + \underline{\varepsilon}_{i_k}, \quad i_k = 1,\ldots,n_k;\ k = 1,\ldots,12,
$$

waarbij:

$$
(28) \qquad \mathcal{L}(\underline{\varepsilon}_{i_k}) = \mathcal{N}_{18}(0,I), \text{ o.o. voor } i_k = 1,\ldots,n_k;\ k = 1,\ldots,12.
$$

Door de formule (25) voor $i_k = 1,\ldots,n_k$; $k = 1,\ldots,12$ onder elkaar te schrijven ontstaat het standaard regressie probleem:

$$
(29) \qquad \underline{y} = X\beta + \underline{\varepsilon}.
$$

De kleinste kwadraten schatter van β in (29) is precies de schatter $\underline{\tilde{\beta}}$, immers:

$$
(30) \qquad \tilde{\beta} = (X'X)^{-1}X'y.
$$

De gebruikelijke F-toets voor de hypothese (18) verwerpt als:

$$
(31) \qquad \frac{18n-43}{\ell}\, \tilde{\beta}'A'(A(X'X)^{-1}A')^{-1}A\tilde{\beta}/(y-X\tilde{\beta})'(y-X\tilde{\beta})
$$

groot is (zie RAO (1973), p. 238). De toets (31) is weer asymptotisch equivalent aan de toets gebaseerd op (22).

4. ANALYSE OPENBAAR-VERVOERGEBRUIKERSBESTAND

Het openbaar-vervoergebruikersbestand bestaat uit 1316 respondenten.

Na verwijdering van de respondenten die niet alle 18 situaties beoordeeld hebben, blijft de volgende verdeling over de blokken over:

blok 1	$n_{111} = 60$;	blok 7	$n_{211} = 65$
blok 2	$n_{112} = 65$;	blok 8	$n_{212} = 126$
blok 3	$n_{121} = 91$;	blok 9	$n_{221} = 64$
blok 4	$n_{122} = 57$;	blok 10	$n_{222} = 75$
blok 5	$n_{131} = 70$;	blok 11	$n_{231} = 96$
blok 6	$n_{132} = 87$;	blok 12	$n_{232} = 80$.

Uit deze verdeling blijkt dat de respondenten niet uniform over de blokken verdeeld zijn. (Een χ^2-toets op gelijke blokkansen verwerpt.) Nader onderzoek naar de oorzaak van dit verschijnsel is wenselijk. Er is een ernstige vertekening in de steekproef ontstaan indien als oorzaak aangewezen wordt dat juist die respondenten die een bepaald antwoord zouden geven niet geantwoord hebben. Er zijn ook minder ernstige oorzaken denkbaar; bijvoorbeeld dat van ieder blok niet evenveel enquête formulieren in omloop zijn geweest.

De schattingen $\hat{\Sigma}_{111},\ldots,\hat{\Sigma}_{232}$ van de twaalf covariantiematrices zijn niet in dit rapport opgenomen. Met behulp van deze schattingen is onderzocht of de matrices $\Sigma_{111},\ldots,\Sigma_{232}$ mogelijkerwijs aan elkaar gelijk zijn. Hiervoor is de aannemelijkheidsquotient toets, zie MUIRHEAD (1982), p. 292, uitgevoerd, welke duidelijk verwerpt. Het blijkt zelfs dat deze matrices qua grootte van de elementen grofweg in twee groepen ingedeeld kunnen worden: enerzijds $\hat{\Sigma}_{111}$, $\hat{\Sigma}_{112}$, $\hat{\Sigma}_{211}$ en $\hat{\Sigma}_{212}$ en anderzijds de resterende matrices. Globaal gezien hebben de matrices uit de eerst genoemde groep iets grotere elementen. Merk op dat deze matrices corresponderen met de denkbeeldige situatie waarbij de reistijd met de fiets 15 minuten is. Dit verschil in (co-) varianties zou verklaard kunnen worden uit het feit dat voor situaties met een reistijd met de fiets van 15 minuten de gemiddelde beoordeling in het midden van het antwoordbereik (d.w.z. rond de 50%) ligt, terwijl voor situaties met langere fietstijden de gemiddelde beoordeling meer naar de rand van het antwoordbereik verschoven is en er dus minder variatie mogelijk is. Door voor de afhankelijke variabele niet de beoordeling zelf maar een transformatie ervan te nemen (bijvoorbeeld met de arcsinus functie) zullen de (co)varianties van waarnemingen waarvoor de verwachting in het middengebied ligt minder gaan verschillen van die van waarnemingen met een verwachting dicht bij de rand van het antwoordbereik. Zo'n transformatie heeft als nadeel dat de resultaten (schattingen van de hoofdeffecten en interacties) niet meer de oorspronkelijke eenheid hebben en dus moeilijk te interpreteren zijn. Om deze reden is geen arcsinus transformatie uitgevoerd.

De modelveronderstellingen in (3) betreffende de covariantiestructuur tussen de beoordelingen kan niet eenvoudiger gekozen worden. Er is geen bepaald patroon in de covariantiestructuur te vinden, zoals bijvoorbeeld een diagonaalstructuur (onafhankelijkheid) of een structuur met constante correlaties (deze structuur wordt in de modellen van WINER (1971) veelvuldig gebruikt). De covarianties tussen beoordelingen van een respondent zijn relatief hoog.

Aangezien de situatievolgorde voor alle respondenten uit een blok hetzelfde is, zal een volgorde-effect, indien aanwezig, binnen een blok optreden. Het is dan te verwachten dat covarianties tussen eerst gestelde vragen lager zijn dan tussen laatst gestelde vragen. Een dergelijk effect blijkt echter niet duidelijk aanwezig te zijn. Desalniettemin is het verstandig bij soort gelijke onderzoeken de situatievolgorde weer te randomiseren.

De in paragraaf 3 beschreven schatter $\underline{\tilde{\beta}}$ en toetsen (31) leveren de volgende schattingen voor de hoofdeffecten en interacties en overschrijdingskansen voor de hypotheses of de betreffende effecten en interacties gelijk aan nul zijn.

OPENBAAR-VERVOERGEBRUIKERS.

algemeen gemiddelde beoordeling

μ	= 55.90

"aantal malen overstappen" overschrijdingskans .866

	α
0 keer	0.14
1 keer	-0.14

"reistijd fiets" overschrijdingskans .000

	β
15 min.	-13.32
30 min.	3.91
45 min.	9.41

"zitplaats terugreis" overschrijdingskans .013

	γ
geen	-2.07
wel	2.07

"reistijd bus" overschrijdingskans .000

	δ
laag	1.45
hoog	-1.45

"volgtijd bus" overschrijdingskans .000

	ε
5 min.	5.76
10 min.	-0.28
15 min.	-5.48

"kosten abonnement" overschrijdingskans .000

	ζ
f34,--	6.85
f51,--	-1.65
f59,50	-5.19

"aantal malen overstappen" overschrijdingskans
*
"reistijd fiets" .265

$\alpha\beta$	15 min.	30 min.	45 min.
0 keer	0.11	1.46	-1.57
1 keer	-0.11	-1.46	1.57

"aantal malen overstappen" overschrijdingskans
*
"zitplaats terugreis" .625

$\alpha\gamma$	geen	wel
0 keer	0.40	-0.40
1 keer	-0.40	0.40

"aantal malen overstappen" overschrijdingskans
*
"reistijd bus" .000

$\alpha\delta$	laag	hoog
0 keer	0.81	-0.81
1 keer	-0.81	0.81

"aantal malen overstappen" overschrijdingskans
*
"volgtijd bus" .601

$\alpha\varepsilon$	5 min.	10 min.	15 min.
0 keer	0.05	0.07	-0.12
1 keer	-0.05	-0.07	0.12

"aantal malen overstappen" overschrijdingskans
*
"kosten abonnement" .000

$\alpha\zeta$	f34,--	f51,--	f59,50
0 keer	-0.89	0.49	0.41
1 keer	0.89	-0.49	-0.41

"reistijd fiets" overschrijdingskans
*
"zitplaats terugreis" .066

$\beta\gamma$	geen	wel
15 min.	2.91	-2.91
30 min.	-1.25	1.25
45 min.	-1.66	1.66

"reistijd fiets" overschrijdingskans
*
"reistijd bus" .103

$\beta\delta$	laag	hoog
15 min.	-0.10	0.10
30 min.	0.24	-0.24
45 min.	-0.14	0.14

"reistijd fiets" * "volgtijd bus" — overschrijdingskans .ØØ6

βε	5 min.	1Ø min.	15 min.
15 min.	Ø.73	-Ø.4Ø	-Ø.33
3Ø min.	Ø.ØØ	Ø.Ø6	-Ø.Ø6
45 min.	-Ø.73	Ø.34	Ø.39

"reistijd fiets" * "kosten abonnement" — overschrijdingskans .ØØØ

βζ	f34,--	f51,--	f59,5Ø
15 min.	-1.15	-Ø.21	1.37
3Ø min.	1.2Ø	-Ø.16	-1.Ø4
45 min.	-Ø.Ø5	Ø.37	-Ø.32

"zitplaats terugreis" * "reistijd bus" — overschrijdingskans .123

γδ	laag	hoog
geen	Ø.11	-Ø.11
wel	-Ø.11	Ø.11

"zitplaats terugreis" * "volgtijd bus" — overschrijdingskans .Ø14

γε	5 min.	1Ø min.	15 min.
geen	Ø.41	Ø.15	-Ø.56
wel	-Ø.41	-Ø.15	Ø.56

"zitplaats tergreis" * "kosten abonnement" — overschrijdingskans .Ø58

γζ	f34,--	f51,--	f59,5Ø
geen	Ø.13	-Ø.26	Ø.13
wel	-Ø.13	Ø.26	-Ø.13

"reistijd bus" * "volgtijd bus" — overschrijdingskans .ØØ1

δε	5 min.	1Ø min.	15 min.
laag	Ø.18	Ø.Ø3	-Ø.21
hoog	-Ø.18	-Ø.Ø3	Ø.21

"reistijd bus" * "kosten abonnement" — overschrijdingskans .Ø63

δζ	f34,--	f51,--	f59,5Ø
laag	Ø.13	Ø.ØØ	-Ø.13
hoog	-Ø.13	Ø.ØØ	Ø.13

"volgtijd bus" * "kosten abonnement" — overschrijdingskans .ØØØ

εζ	f34,--	f51,--	f59,5Ø
5 min.	Ø.79	Ø.Ø3	-Ø.82
1Ø min.	-Ø.Ø4	-Ø.15	Ø.19
15 min.	-Ø.75	Ø.12	Ø.64

De schattingen van de hoofdeffecten zijn goed te interpreteren. Bijvoorbeeld, naarmate de reistijd met de fiets groter wordt neemt de kans op het kopen van een abonnement toe. Ook van de andere hoofdeffecten komt de volgorde van de schattingen goed overeen met de intuïtieve volgorde. Het effect ten gevolge van de factor "reistijd fiets" is het belangrijkste; het algemeen gemiddelde varieert hier van 42.6% tot 65.3% kans op het kopen van een maandabonnement. Na "reistijd fiets" zijn "kosten abonnement" en "volgtijd bus" de belangrijkste factoren, gevolgd door respectivelijk "zitplaats terugreis" en "reistijd bus". Hoewel sommige interacties wel significant van nul verschillen, zijn ze niet erg interessant omdat ze in het algemeen een afwijking van minder dan 1% op het gemiddelde veroorzaken.

5. ANALYSE NIET-OPENBAAR-VERVOERGEBRUIKERSBESTAND

De contrastgroep van niet-openbaar-vervoergebruikers bestaat uit 292 respondenten. Na verwijdering van de niet volledig ingevulde formulieren zijn deze respondenten als volgt over de blokken verdeeld:

blok 1	: $n_{111} = 6$;	blok 7	: $n_{211} = 23$
blok 2	: $n_{112} = 25$;	blok 8	: $n_{212} = 9$
blok 3	: $n_{121} = 12$;	blok 9	: $n_{221} = 15$
blok 4	: $n_{122} = 25$;	blok 10	: $n_{222} = 24$
blok 5	: $n_{131} = 11$;	blok 11	: $n_{231} = 11$
blok 6	: $n_{132} = 16$;	blok 12	: $n_{232} = 12$.

Ook nu zijn de respondenten niet uniform over de blokken verdeeld, en is nader onderzoek wenselijk.

De aantallen in de blokken zijn te klein om twaalf verschillende covariantiematrices goed te kunnen schatten. Om deze reden wordt voor de contrastgroep in het model de extra veronderstelling gemaakt dat de twaalf covariantiematrices voor de verschillende blokken gelijk zijn; dat wil zeggen:

(32) $$\Sigma_{111} = \Sigma_{112} = \ldots = \Sigma_{232}.$$

De wijzigingen in de theorie van paragraaf 3 zijn voor de hand liggend. De gemeenschappelijke covariantiematrix word nu geschat met:

$$\tilde{\Sigma} = \frac{1}{n} \sum_{k=1}^{12} \sum_{i_k=1}^{n_k} (z_{i_k} - \bar{z}_k)(z_{i_k} - \bar{z}_k)'. \tag{33}$$

Analyse van de contrastgroep levert de volgende schattingen van effecten en overschrijdingskansen van de toetsen op.

NIET-OPENBAAR-VERVOERGEBRUIKERS.

algemeen gemiddelde beoordeling

μ	= 38.3Ø

"aantal malen overstappen" — overschrijdingskans .583

	α
Ø keer	1.22
1 keer	-1.22

"reistijd fiets" — overschrijdingskans .ØØØ

	β
15 min.	-12.9Ø
3Ø min.	1.22
45 min.	11.68

"zitplaats terugreis" — overschrijdingskans .ØØ8

	γ
geen	-5.92
wel	5.92

"reistijd bus" — overschrijdingskans .ØØØ

	δ
laag	1.88
hoog	-1.88

"volgtijd bus" — overschrijdingskans .ØØØ

	ε
5 min.	5.75
1Ø min.	-Ø.Ø8
15 min.	-5.67

"kosten abonnement" — overschrijdingskans .ØØØ

	ζ
f34,--	5.43
f51,--	-Ø.88
f59,5Ø	-4.55

"aantal malen overstappen" * "reistijd fiets" — overschrijdingskans .715

αβ	15 min.	3Ø min.	45 min.
Ø keer	-Ø.31	2.Ø5	-1.74
1 keer	Ø.31	-2.Ø5	1.74

"aantal malen overstappen" * "zitplaats terugreis" — overschrijdingskans .787

αγ	geen	wel
Ø keer	Ø.55	-Ø.55
1 keer	-Ø.55	Ø.55

"aantal malen overstappen" * "reistijd bus" — overschrijdingskans .ØØØ

αδ	laag	hoog
Ø keer	1.24	-1.24
1 keer	-1.24	1.24

"aantal malen overstappen" * "volgtijd bus" — overschrijdingskans .481

αε	5 min.	1Ø min.	15 min.
Ø keer	-Ø.23	Ø.27	-Ø.Ø4
1 keer	Ø.23	-Ø.27	Ø.Ø4

"aantal malen overstappen" * "kosten abonnement" — overschrijdingskans .195

αζ	f34,--	f51,--	f59,5Ø
Ø keer	-Ø.12	Ø.45	Ø.33
1 keer	Ø.12	-Ø.45	-Ø.33

"reistijd fiets" * "zitplaats terugreis" — overschrijdingskans .644

βγ	geen	wel
15 min.	-Ø.42	Ø.42
3Ø min.	2.41	-2.41
45 min.	-1.99	1.99

"reistijd fiets" * "reistijd bus" — overschrijdingskans .Ø99

βδ	laag	hoog
15 min.	-Ø.94	Ø.94
3Ø min.	Ø.73	-Ø.73
45 min.	Ø.21	-Ø.21

"reistijd fiets" * "volgtijd bus" — overschrijdingskans .522

βε	5 min.	1Ø min.	15 min.
15 min.	-Ø.Ø1	-Ø.17	Ø.18
3Ø min.	Ø.33	Ø.49	-Ø.82
45 min.	-Ø.32	-Ø.32	Ø.64

"reistijd fiets" * "kosten abonnement" — overschrijdingskans .Ø79

βζ	f34,--	f51,--	f59,5Ø
15 min.	-1.29	-Ø.2Ø	1.49
3Ø min.	1.42	-Ø.22	-1.2Ø
45 min.	-Ø.13	Ø.42	-Ø.29

"zitplaats terugreis" * "reistijd bus" — overschrijdingskans .785

βδ	laag	hoog
geen	-Ø.1Ø	Ø.1Ø
wel	Ø.1Ø	-Ø.1Ø

"zitplaats terugreis" * "volgtijd bus" — overschrijdingskans .762

γε	5 min.	1Ø min.	15 min.
geen	-Ø.13	Ø.17	-Ø.Ø4
wel	Ø.13	-Ø.17	Ø.Ø4

"zitplaats tergreis" * "kosten abonnement" — overschrijdingskans .612

γζ	f34,--	f51,--	f59,5Ø
geen	Ø.47	-Ø.25	-Ø.22
wel	-Ø.47	Ø.25	Ø.22

"reistijd bus" * "volgtijd bus" — overschrijdingskans .384

δε	5 min.	1Ø min.	15 min.
laag	-Ø.3Ø	Ø.17	Ø.13
hoog	Ø.3Ø	-Ø.17	-Ø.13

"reistijd bus" * "kosten abonnement" — overschrijdingskans .921

δζ	f34,--	f51,--	f59,5Ø
laag	-Ø.Ø4	Ø.Ø7	-Ø.Ø3
hoog	Ø.Ø4	-Ø.Ø7	Ø.Ø3

"volgtijd bus" * "kosten abonnement" — overschrijdingskans .ØØ9

εζ	f34,--	f51,--	f59,5Ø
5 min.	1.35	-Ø.36	-Ø.99
1Ø min.	-Ø.4Ø	Ø.15	Ø.25
15 min.	-Ø.95	Ø.21	Ø.74

Deze resultaten zijn goed in overeenstemming met de resultaten van de analyse van paragraaf 4. Het algemene gemiddelde van alle situaties ligt, zoals te verwachten, lager (38.3% in plaats van 55.9%). De hoofdeffecten hebben dezelfde orde van grootte als bij de openbaarvervoergebruikers; de factor "zitplaats terugreis" is misschien iets belangrijker geworden. De meeste interacties verschillen niet significant van nul.

6. SLOTOPMERKINGEN

Buiten inhoudelijke conclusies en opmerkingen over de resultaten van beide analyses, zijn de volgende opmerkingen van algemene aard te maken.

1. In de gevolgde proefopzet zijn de factoren volledig gekruist (volledig proefschema; engels: complete layout). Door het aantal factoren en niveaus beperkt te houden zijn er in iedere cel een behoorlijk aantal waarnemingen mogelijk. Hierdoor kan een beroep gedaan worden op asymptotische normaliteit van de schatters en op de "optimale" asymptotische eigenschappen van de analyse methode. Een groot voordeel van een volledig proefschema is dat er geen verwarring (eng.: confounding) plaats vindt tussen de effecten. In een onvolledig proefschema (incomplete layout) zijn er niet bij alle mogelijke combinaties van factoren waarnemingen gedaan. De cellen waarvoor waarnemingen gedaan worden zijn volgens een bepaald patroon gekozen; voorbeelden zijn latijnse vierkanten en geneste proefschema's. Het voordeel boven een volledig proefschema is dat met evenveel waarnemingen per (waar te nemen) cel meer denkbeeldige situaties (d.w.z. meer factoren met meer niveaus) bestudeerd kunnen worden. Een groot nadeel is echter dat voor de analyse methode de veronderstelling van geen interacties essentieel is. In de aanwezigheid van interacties treed verwarring op tussen de hoofdeffecten en interacties. Hierdoor kan de interpretatie bij onvolledige proefschema's misleidend zijn. Bij het onderhavige experiment kunnen de interacties niet nul verondersteld worden.

   Door iedere respondent meerdere beoordelingen te laten geven, wordt het experiment minder gevoelig voor verschillen tussen respondenten. Het is echter ondoenlijk om aan elke respondent alle 216 denkbeeldige situaties voor te leggen. Om deze reden beoordeelt een respondent de 18 situaties die bij een bepaald blok horen.

De in het experiment gekozen afhankelijke variabele (het kanspercentage op kopen van abonnement) heeft als nadeel dat veel respondenten een beoordeling aan de rand van het antwoordbereik zullen geven.

2. Onderzoek naar de oorzaak van de niet uniforme verdeling van respondenten onder de 12 blokken is wenselijk. Mogelijkerwijs geven de steekproeven vertekende beelden van de te onderzoeken populaties.
3. Het gebruikte wiskundige model lijkt realistisch en is goed hanteerbaar. Bovendien is een goede statistische analyse mogelijk, welke met standaard computerprogrammatuur uitgevoerd kan worden. De modelveronderstellingen van het model zijn redelijkerwijs vervuld en de resultaten zijn goed interpreteerbaar. Er is een goede overeenstemming tussen de resultaten van beide analyses.
4. Dezelfde vorm van analyse kan natuurlijk op deelgroepen van respondenten worden toegepast. Bijvoorbeeld de deelgroep bestaande uit respondenten waarvoor de denkbeeldige situaties realistisch zijn.

De auteur is Dr. R.D. GILL en Ir. J.P. ROOS erkentelijk voor hun waardevolle opmerkingen tijdens het uitvoeren van de analyse en de voorbereiding van dit rapport, en is de heer R. v.d. HORST erkentelijk voor het schrijven van de computerprogramma's en het uitvoeren van de berekeningen.

LITERATUUR

[1] *Attitude-onderzoek openbaar vervoer in Rotterdam en omgeving*, DHV Raadgevend Ingenieursbureau-rapport, dd. 15-2-1983.

[2] DZHAPARIDZE, K.O. (1983), *On iterative procedures of asymptotic inference*, Mathematisch Centrum rapport SW 94-83, Amsterdam. (Te verschijnnen in Statistica Neerlandica.)

[3] LEHMANN, E.L. (1959), *Testing Statistical Hypotheses*, Wiley, New York.

[4] MUIRHEAD, R.J. (1982), *Aspects of Multivariate Statistical Theory*, Wiley, New-York.

[5] RAO, C.R. (1973), *Linear Statistical Inference and its Applications*, Wiley, New-York.

[6] WINER, B.J. (1971), *Statistical Principals in Experimental Design*, McGraw-Hill, New-York.

aantal malen overstappen	reistijd fiets	reistijd bus / volgtijd bus / kosten / zitplaats	laag									hoog										aantal respond in blok
		volgtijd bus	5 min.			10 min.			15 min.			5 min.			10 min.			15 min.				
		kosten	ƒ34,--	ƒ51,--	ƒ59,50	ƒ34,--	ƒ51,--	ƒ59,50	ƒ34,--	ƒ51,--	ƒ59,50	ƒ34,--	ƒ51,--	ƒ59,50	ƒ34,--	ƒ51,--	ƒ59,50	ƒ34,--	ƒ51,--	ƒ59,50		
niet overstappen	15 min.	geen																			blok 1	n_{111}
		wel																			blok 2	n_{112}
	30 min.	geen	"	"	"	"	"	"	"	"	"	"	"	"	"	"	"	"	"	"	blok 3	n_{121}
		wel																				n_{122}
	45 min.	geen																				n_{131}
		wel																				n_{132}
1× overstappen	15 min.	geen																				n_{211}
		wel																				n_{212}
	30 min.	geen																				n_{221}
		wel																				n_{222}
	45 min.	geen																				n_{231}
		wel																			blok 12	n_{232}

Tabel 1. Schematisch overzicht denkbeeldige situaties

1	1	1	0	1	1	1	0	1	0	1	0	1	1	1	0	1	0	1	0	1	0	1	0	0	0	1	0	0	0	1	1	0	1	0	1	0	1	0	1	0	0	0
1	1	1	0	1	1	1	0	0	1	1	0	1	1	1	0	0	1	1	0	1	0	1	0	0	0	0	1	0	0	1	1	0	0	1	1	0	0	1	0	1	0	0
1	1	1	0	1	1	1	0	-1	-1	1	0	1	1	1	0	-1	-1	1	0	1	0	1	0	0	0	-1	-1	0	0	1	1	0	-1	-1	1	0	-1	-1	-1	-1	0	0
1	1	1	0	1	1	0	1	1	0	1	0	1	1	0	1	1	0	1	0	1	0	0	1	0	0	1	0	0	0	1	0	1	1	0	0	1	1	0	0	0	1	0
1	1	1	0	1	1	0	1	0	1	1	0	1	1	0	1	0	1	1	0	1	0	0	1	0	0	0	1	0	0	1	0	1	0	1	0	1	0	1	0	0	0	1
1	1	1	0	1	1	0	1	-1	-1	1	0	1	1	0	1	-1	-1	1	0	1	0	0	1	0	0	-1	-1	0	0	1	0	1	-1	-1	0	1	-1	-1	0	0	-1	-1
1	1	1	0	1	1	-1	-1	1	0	1	0	1	1	-1	-1	1	0	1	0	1	0	-1	-1	0	0	1	0	0	0	1	-1	-1	1	0	-1	-1	1	0	-1	0	-1	0
1	1	1	0	1	1	-1	-1	0	1	1	0	1	1	-1	-1	0	1	1	0	1	0	-1	-1	0	0	0	1	0	0	1	-1	-1	0	1	-1	-1	0	1	0	-1	0	-1
1	1	1	0	1	1	-1	-1	-1	-1	1	0	1	1	-1	-1	-1	-1	1	0	1	0	-1	-1	0	0	-1	-1	0	0	1	-1	-1	-1	-1	-1	-1	-1	-1	1	1	1	1
1	1	1	0	1	-1	1	0	1	0	1	0	1	-1	1	0	1	0	1	0	-1	0	1	0	0	0	1	0	0	0	-1	1	0	1	0	-1	0	1	0	1	0	0	0
1	1	1	0	1	-1	1	0	0	1	1	0	1	-1	1	0	0	1	1	0	-1	0	1	0	0	0	0	1	0	0	-1	1	0	0	1	-1	0	0	1	0	1	0	0
1	1	1	0	1	-1	1	0	-1	-1	1	0	1	-1	1	0	-1	-1	1	0	-1	0	1	0	0	0	-1	-1	0	0	-1	1	0	-1	-1	-1	0	-1	-1	-1	-1	0	0
1	1	1	0	1	-1	0	1	1	0	1	0	1	-1	0	1	1	0	1	0	-1	0	0	1	0	0	1	0	0	0	-1	0	1	1	0	0	-1	1	0	0	0	1	0
1	1	1	0	1	-1	0	1	0	1	1	0	1	-1	0	1	0	1	1	0	-1	0	0	1	0	0	0	1	0	0	-1	0	1	0	1	0	-1	0	1	0	0	0	1
1	1	1	0	1	-1	0	1	-1	-1	1	0	1	-1	0	1	-1	-1	1	0	-1	0	0	1	0	0	-1	-1	0	0	-1	0	1	-1	-1	0	-1	-1	-1	0	0	-1	-1
1	1	1	0	1	-1	-1	-1	1	0	1	0	1	-1	-1	-1	1	0	1	0	-1	0	-1	-1	0	0	1	0	0	0	-1	-1	-1	1	0	1	1	1	0	-1	0	-1	0
1	1	1	0	1	-1	-1	-1	0	1	1	0	1	-1	-1	-1	0	1	1	0	-1	0	-1	-1	0	0	0	1	0	0	-1	-1	-1	0	1	1	1	0	1	0	-1	0	-1
1	1	1	0	1	-1	-1	-1	-1	-1	1	0	1	-1	-1	-1	-1	-1	1	0	-1	0	-1	-1	0	0	-1	-1	0	0	-1	-1	-1	-1	-1	1	1	-1	-1	1	1	1	1

Matrix D_{111}