

Leipzig, 13. September 2006

# **Konfidenz- und Ablehnungsbereiche für multivariate Tests in klinischen Studien**

**Jürgen Läuter**

Otto-von-Guericke-Universität Magdeburg

und

**Joachim Röhmel**

Institut für Biometrie und Klinische Epidemiologie, Charité Universitätsmedizin Berlin

## 1 Einführung

Vortrag schließt an die Problemstellung von Joachim Röhmel, Christoph Gerlinger und Norbert Benda an.

Schmerzstudie an zwei Behandlungsgruppen:

Aktive Behandlung mit zwei Medikamenten und entsprechende Placebo-Behandlung.

Für jedes Medikament soll ein univariater Nichtunterlegenheitstest erfüllt werden, und für wenigstens ein Medikament soll die Überlegenheit nachgewiesen werden.

Dazu wird insbesondere ein zweidimensionaler Test gegen Unterschied 0 angewandt, der aber auch für alle negativen Fälle (Medikament schlechter als Placebo) das einseitige Signifikanzniveau  $\alpha = 0.025$  nicht überschreiten soll.

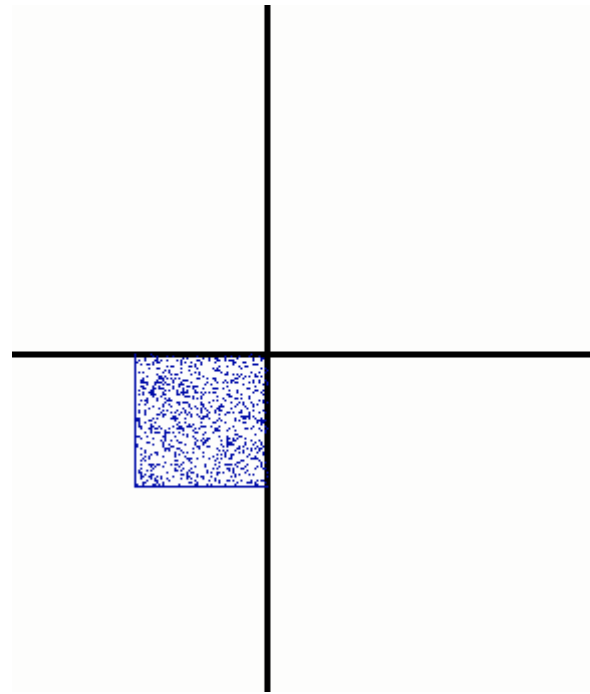
Da vorausgesetzt wird,

dass jedes Medikament zumindest die Nichtunterlegenheitsschranke  $\varepsilon_1 < 0$  bzw.  $\varepsilon_2 < 0$  erreicht,

braucht nur noch die „Nullhypothese“, dass  $\begin{pmatrix} \delta_1 \\ \delta_2 \end{pmatrix} = \begin{pmatrix} \mu_1^{(1)} - \mu_1^{(2)} \\ \mu_2^{(1)} - \mu_2^{(2)} \end{pmatrix}$  im Rechteck  $\begin{cases} \varepsilon_1 \leq \delta_1 \leq 0 \\ \varepsilon_2 \leq \delta_2 \leq 0 \end{cases}$  liegt,

verworfen zu werden. Alle sonst noch möglichen Fälle entsprechen der Zielsetzung der Studie,

mindestens ein Medikament hat dann einen positiven Unterschied zu Placebo.



Im Folgenden werden drei multivariate Verfahren untersucht:

- Der positiv gerichtete Konfidenzbereich nach dem Prinzip des Hotelling-Tests.

Die Methode ist an Follmann (1996), JASA **91**, 854-861, angelehnt, geht aber darüber hinaus.

- Der positiv gerichtete Konfidenzbereich nach dem konservativen PC-Test.

Siehe Läuter und Kropf (2002), Biocybernetics and Biomedical Engineering **22**, 5-16,

Läuter und Glimm (2005), Statistics **39**, 207-219.

- Verfahren auf der Grundlage des Standard-Summen-Tests (SS-Test).

Siehe gemeinsame Arbeit mit Röhmel, Gerlinger und Benda.

Alle folgenden Ausführungen beziehen sich auf den Vergleich der beiden Behandlungsgruppen, auch wenn das nicht immer besonders betont wird.

## 2 Die positiv gerichtete Hotelling-Methode

Der Hotelling-Test stützt sich auf quadratische Formen und ist daher zunächst nicht für gerichtete Untersuchungen geeignet.

Follmann fordert in seinem Test gegen 0 zwar,

dass zum Erreichen von positiver Signifikanz eine positive Summe der Mittelwertsdifferenzen,

$$\sum_{i=1}^p (\bar{x}_i^{(1)} - \bar{x}_i^{(2)}) \geq 0 ,$$

vorhanden sein muss.

Aber das bedeutet noch nicht, dass bei beliebigen negativen Parameterdifferenzen  $\delta_i = \mu_i^{(1)} - \mu_i^{(2)}$  eine Signifikanz höchstens mit Wahrscheinlichkeit  $\alpha$  möglich ist.

Es ist für uns wesentlich, dass die Anforderungen auf die **wahren Parameter** bezogen werden.

Der Test wird daher im Folgenden verschärft.

## Ungerichteter $p$ -dimensionaler $(1 - \alpha)$ -Konfidenzbereich im Fall des Zwei-Gruppen-Problems:

$$\frac{(n - p - 1) a}{p} (\boldsymbol{\delta} - \bar{\boldsymbol{x}})' \mathbf{G}^{-1} (\boldsymbol{\delta} - \bar{\boldsymbol{x}}) < F_{1-\alpha}(p, n - p - 1)$$

Hier bedeutet

$$\boldsymbol{\delta} = \begin{pmatrix} \delta_1 \\ \vdots \\ \delta_p \end{pmatrix} \quad \text{den Vektor der } p \text{ wahren Mittelwertsdifferenzen,}$$

$$\bar{\boldsymbol{x}} = \bar{\boldsymbol{x}}^{(1)} - \bar{\boldsymbol{x}}^{(2)} = \begin{pmatrix} \bar{x}_1^{(1)} - \bar{x}_1^{(2)} \\ \vdots \\ \bar{x}_p^{(1)} - \bar{x}_p^{(2)} \end{pmatrix} \quad \text{den Differenzvektor der Mittelwerte der Gruppen,}$$

$$\mathbf{G} = \sum_{j=1}^{n^{(1)}} (\mathbf{x}_{(j)}^{(1)} - \bar{\boldsymbol{x}}^{(1)})(\mathbf{x}_{(j)}^{(1)} - \bar{\boldsymbol{x}}^{(1)})' + \sum_{j=1}^{n^{(2)}} (\mathbf{x}_{(j)}^{(2)} - \bar{\boldsymbol{x}}^{(2)})(\mathbf{x}_{(j)}^{(2)} - \bar{\boldsymbol{x}}^{(2)})'$$

die interne gepoolte Produktsummenmatrix,

$$n = n^{(1)} + n^{(2)}, \quad a = \frac{n^{(1)} n^{(2)}}{n} .$$

Man sollte sich den **Sinn dieser Formel** klarmachen:

- Wenn der wahre Mittelpunkt  $\delta$  der Verteilung gegeben ist,  
liegen die Beobachtungen  $\bar{x}$  in elliptischer Gestalt um  $\delta$  herum.
- Dagegen: Nimmt man eine bestimmte Beobachtung  $\bar{x}$  heraus und fragt sich,  
welche wahren Mittelpunkte  $\delta$  damit verträglich sind,  
dann liegen diese  $\delta$  in elliptischer Gestalt um das feste  $\bar{x}$  herum.

**Also:**

**Man betrachtet hierbei zwei verschiedene Ellipsoide,**

**die aber beide der gleichen Formel folgen!**

Das ist der Zusammenhang von Streubereich und Konfidenzbereich.

### Positiv gerichteter $(1 - \alpha)$ -Streu- und Konfidenzbereich:

Der **gerichtete Streubereich** setzt sich aus einem  $p$ -dimensionalen Halbraum und einem halben Ellipsoid zusammen. Wenn der wahre Mittelpunkt  $\delta$  gegeben ist, dann liegen die Beobachtungen  $\bar{x}$  mit der Wahrscheinlichkeit  $1 - 2\alpha$  in dem **verkleinerten Ellipsoid**

$$\frac{(n - p - 1) a}{p} (\delta - \bar{x})' \mathbf{G}^{-1} (\delta - \bar{x}) < F_{1-2\alpha}(p, n - p - 1).$$

Schneidet man das Ellipsoid passend in der Mitte durch, gilt die Trefferwahrscheinlichkeit  $\frac{1}{2} - \alpha$ .

Als **Halbraum** wird die Menge der Beobachtungen  $\bar{x} = \begin{pmatrix} \bar{x}_1 \\ \vdots \\ \bar{x}_p \end{pmatrix}$  mit  $\sum_{i=1}^p \frac{\bar{x}_i - \delta_i}{\sqrt{g_{ii}}} < 0$  betrachtet.

Die  $g_{ii}$  sind die Diagonalelemente der Matrix  $\mathbf{G}$  (die gepoolten Quadratsummen der Variablen).

Der Rand des Halbraumes verläuft durch den Mittelpunkt  $\delta = \begin{pmatrix} \delta_1 \\ \vdots \\ \delta_p \end{pmatrix}$ , sodass die Wahrscheinlichkeit  $\frac{1}{2}$

realisiert wird. Beide Bereiche zusammen haben die Wahrscheinlichkeit  $1 - \alpha$ .



Der  $\bar{x}$ -Halbraum liegt „unterhalb“ von  $\delta$ . Das halbierte Ellipsoid ragt „nach oben“ heraus.

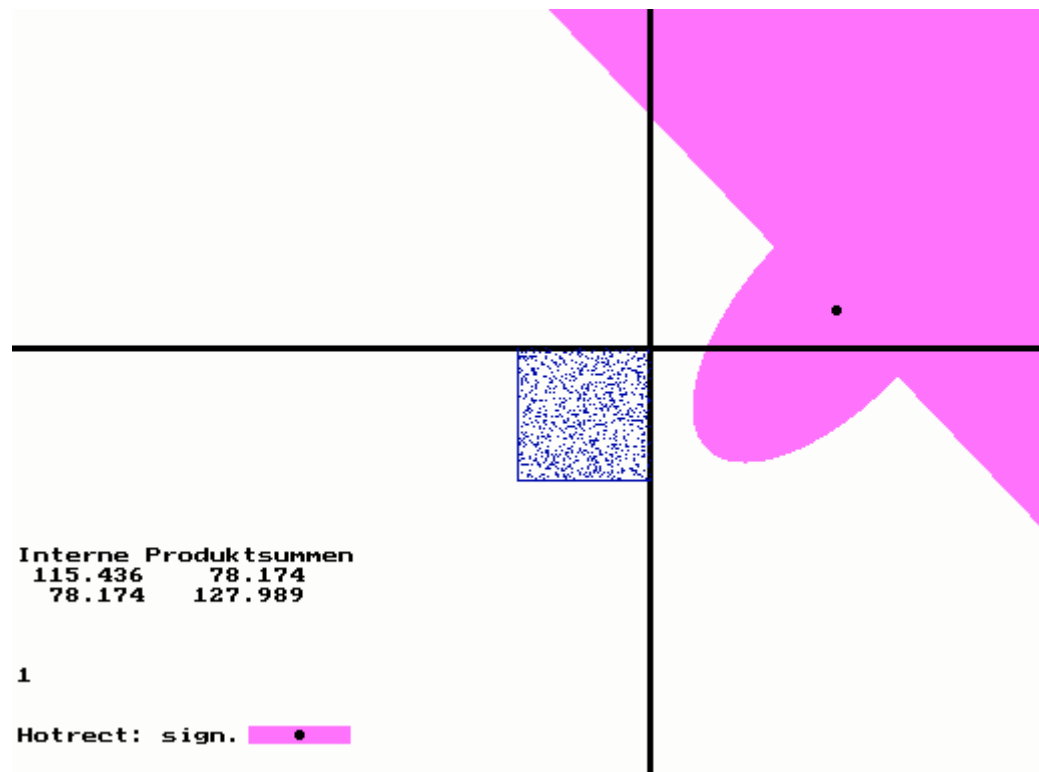
Geht man wieder den umgekehrten Weg, von  $\bar{x}$  zu den damit verträglichen Parametern  $\delta$ , dann liefern die obigen Formeln den entsprechenden **gerichteten Konfidenzbereich**.

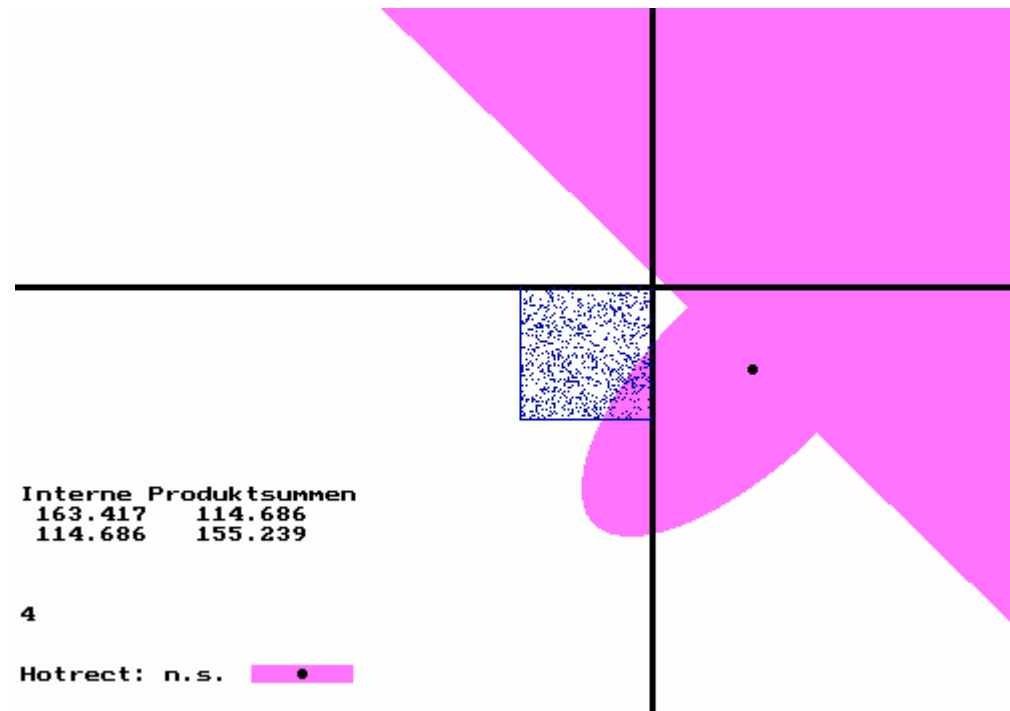
Der  $\delta$ -Halbraum ist durch  $\sum_{i=1}^p \frac{\delta_i - \bar{x}_i}{\sqrt{g_{ii}}} > 0$  gegeben, er enthält also die Mittelwerte  $\delta$  „oberhalb“ von  $\bar{x}$ . Hier ragt das halbierte Ellipsoid „nach unten“ heraus.

Man hat also mit **unterschiedlichen Gebilden** zu tun, die aber mathematisch das Gleiche bedeuten.

Bei unserem gerichteten Test wird „**Signifikanz**“ erreicht, wenn kein Parameterpunkt  $\delta$  des „Nullquaders“ innerhalb des Konfidenzbereichs liegt.

Es ist wesentlich, dass die Richtung der Trennfläche in Abhängigkeit von  $\mathbf{G}$  definiert wird (Unabhängigkeit von  $\bar{x}$  und  $\mathbf{G}$ ).





Der Nullpunkt liegt außerhalb des Konfidenzbereichs,  
 aber das Rechteck nicht!

Der einseitige Hotelling-Test erreicht in der vorliegenden Anwendung eine hohe Güte.  
Im Spezialfall  $p = 2$  hat man für den Test einige quadratische Gleichungen zu lösen.

Vergleich der verschiedenen Verfahren später.

Bei hoher Dimension  $p$  und kleinem  $n$  werden die bekannten Instabilitätsprobleme auftreten.  
Die Bedingung  $p < n - 1$  muss unbedingt eingehalten werden.

### 3 Die konservative Principal-Component-Test-Methode

Ungerichteter konservativer Test der Hypothese  $\delta = \delta_0$ :

$$F_{\text{conserv}} = (n - 2) \frac{d'(\bar{x} - \delta_0) a (\bar{x} - \delta_0)' d}{d' G d} \quad \text{ist stochastisch nicht größer als } F(1, n - 2),$$

wenn  $d$  die Maximalstelle der quadratischen Form  $\lambda = d' G d$  ist.

Dabei muss  $d$  einem bestimmten beschränkten  $p$ -dimensionalen Bereich entnommen sein.

Beispiele:

- Bedingung  $d'd = 1$ ; das führt auf das Eigenwertproblem  $Gd = d\lambda$ .
- Bedingung  $d'd = 1$  mit nichtnegativen Elementen.
- Nur Elemente 1 und 0 sind erlaubt.

Der Beweis dieser Teststrategie stützt sich auf unsere Theorie der sphärischen Tests.

Dabei wird dann die Maximierung der quadratischen Form mit der sog. totalen Produktsummenmatrix durchgeführt.

### Ungerichteter Konfidenzbereich:

$$\bar{x}'d - \sqrt{\frac{\lambda}{(n-2)a} F_{1-\alpha}(1, n-2)} < \delta'd < \bar{x}'d + \sqrt{\frac{\lambda}{(n-2)a} F_{1-\alpha}(1, n-2)},$$

Der Vektor  $\delta$  liegt dann im Zwischenraum zweier Hyperebenen im  $p$ -dimensionalen Raum.

### Entsprechender gerichteter Konfidenzbereich

für Einhaltung von  $\alpha$  bei Parametern  $\delta$  im Nullquader (mit Werten von 0 und darunter):

$$\bar{x}'d - \sqrt{\frac{\lambda}{(n-2)a} F_{1-2\alpha}(1, n-2)} < \delta'd$$

Hiernach liegt  $\delta$  in einem „oberen“ Halbraum, dessen untere Konfidenzgrenze in Abhängigkeit von  $\bar{x}$  bestimmt wird.

Es kommt jetzt auf die Orientierung von  $d$  an, wenn diese nicht ohnehin eindeutig festgelegt ist.

Man strebt nichtnegative Elemente von  $d$  an. Allerdings muss diese Orientierung, wenn sie nicht schon durch die Maximierung von  $\lambda = d'Gd$  feststeht, als eine Funktion von  $G$  definiert sein.

Will man die ganze „totalnegative Ecke“ von Parametern absichern,  
dann kommen nur vollkommen nichtnegative Vektoren  $\mathbf{d}$  in Betracht.

In der Anwendung von Röhmel et al. brauchen wir allerdings nur das eingeschränkte Rechteck unterhalb  
des Nullpunkts zu berücksichtigen.

Trotzdem sollen hier nur nichtnegative Gewichtsvektoren benutzt werden (nach Simulationen).

Das bedeutet für unsere praktische Situation:

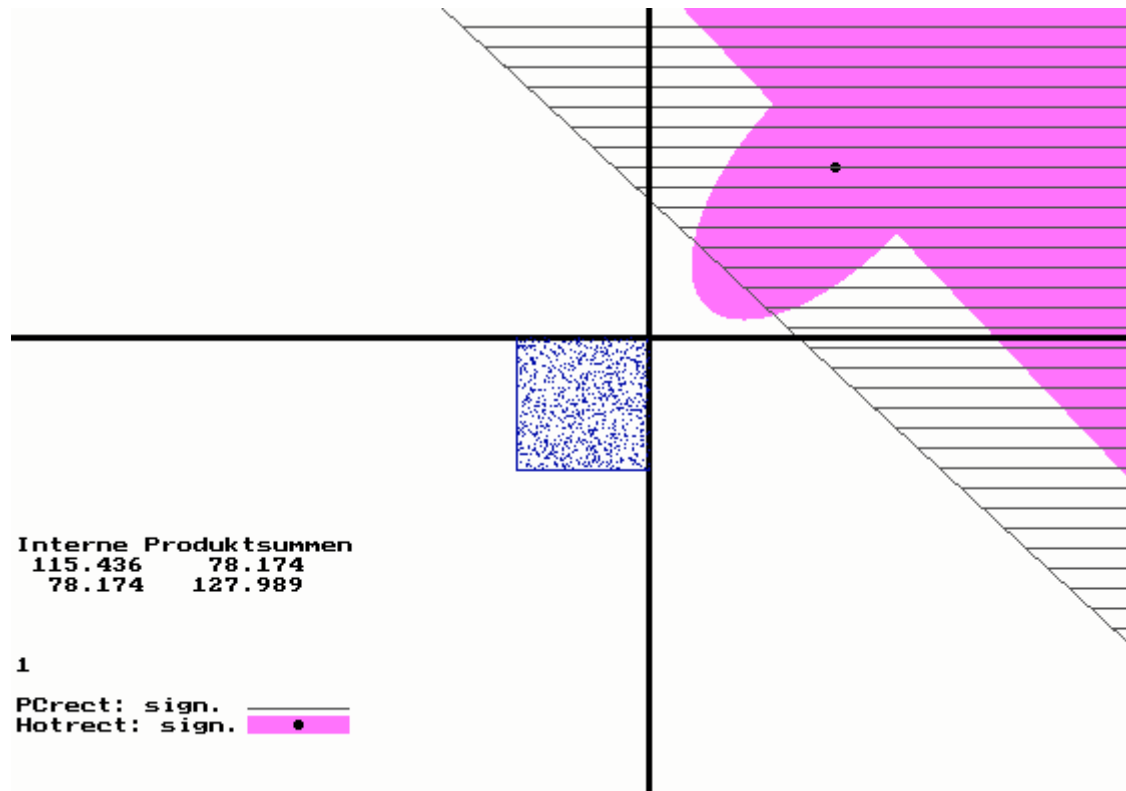
- Im Fall positiver interner Korrelation (d.h. bei  $g_{12} > 0$ ) wird das Eigenwertproblem

$$\begin{pmatrix} g_{11} & g_{12} \\ g_{21} & g_{22} \end{pmatrix} \begin{pmatrix} d_1 \\ d_2 \end{pmatrix} = \begin{pmatrix} d_1 \\ d_2 \end{pmatrix} \lambda$$

gelöst,

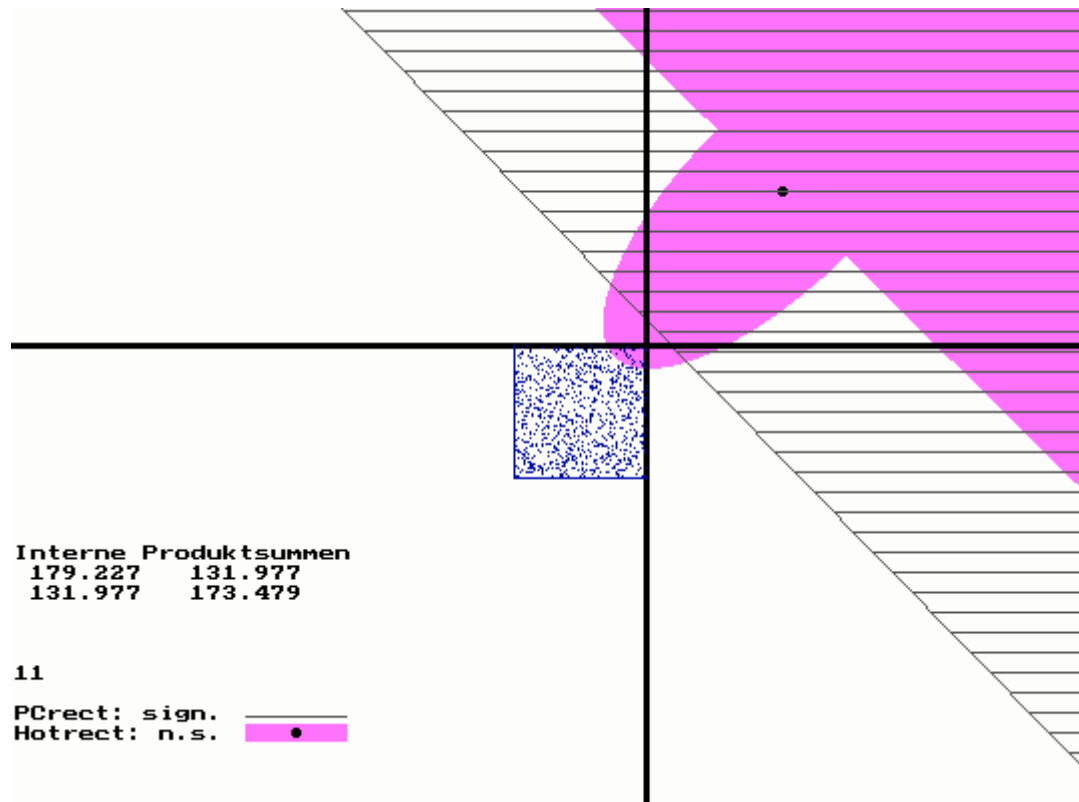
- bei  $g_{12} \leq 0$  wird  $\begin{pmatrix} d_1 \\ d_2 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$  für  $g_{11} \geq g_{22}$  bzw.  $\begin{pmatrix} d_1 \\ d_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$  für  $g_{11} < g_{22}$  gesetzt,

d.h., dann geht der Test in eine univariate Entscheidung über. Ein Nachteil dieser Methode!

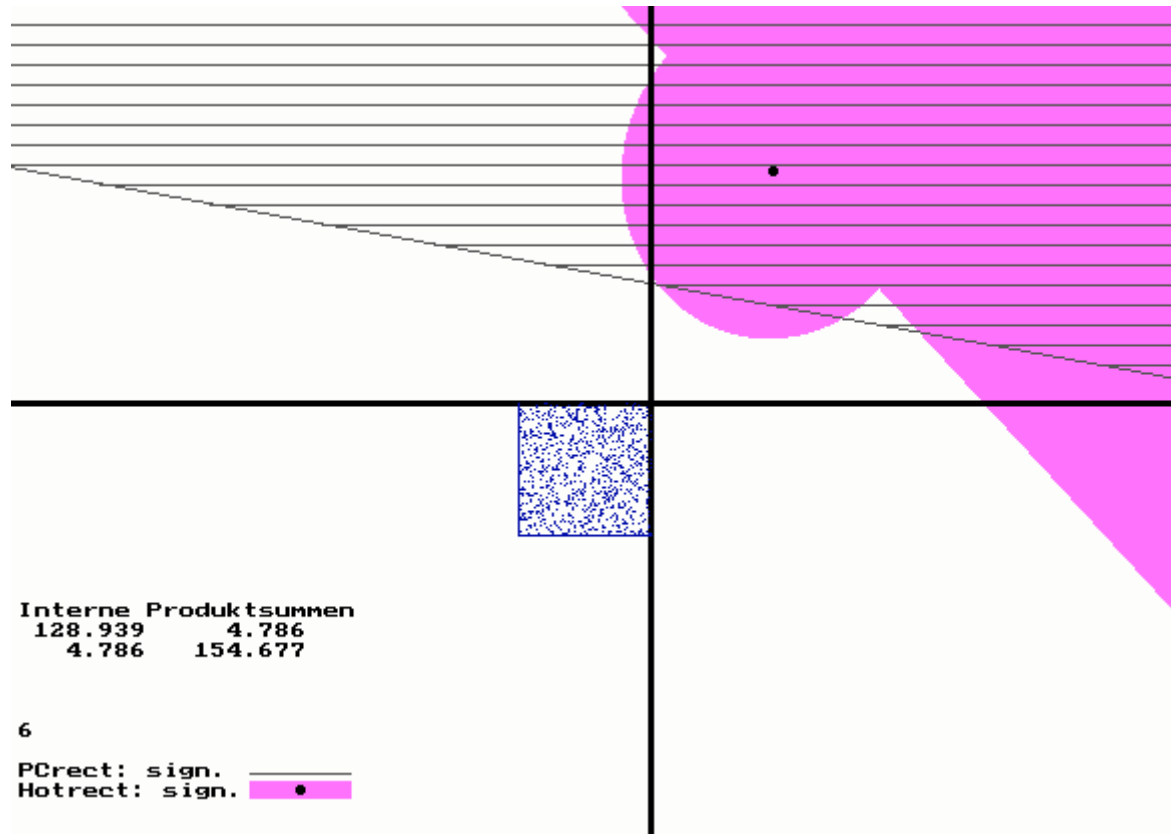


Zwei verschiedene Formen des  $(1 - \alpha)$ -Konfidenzbereichs

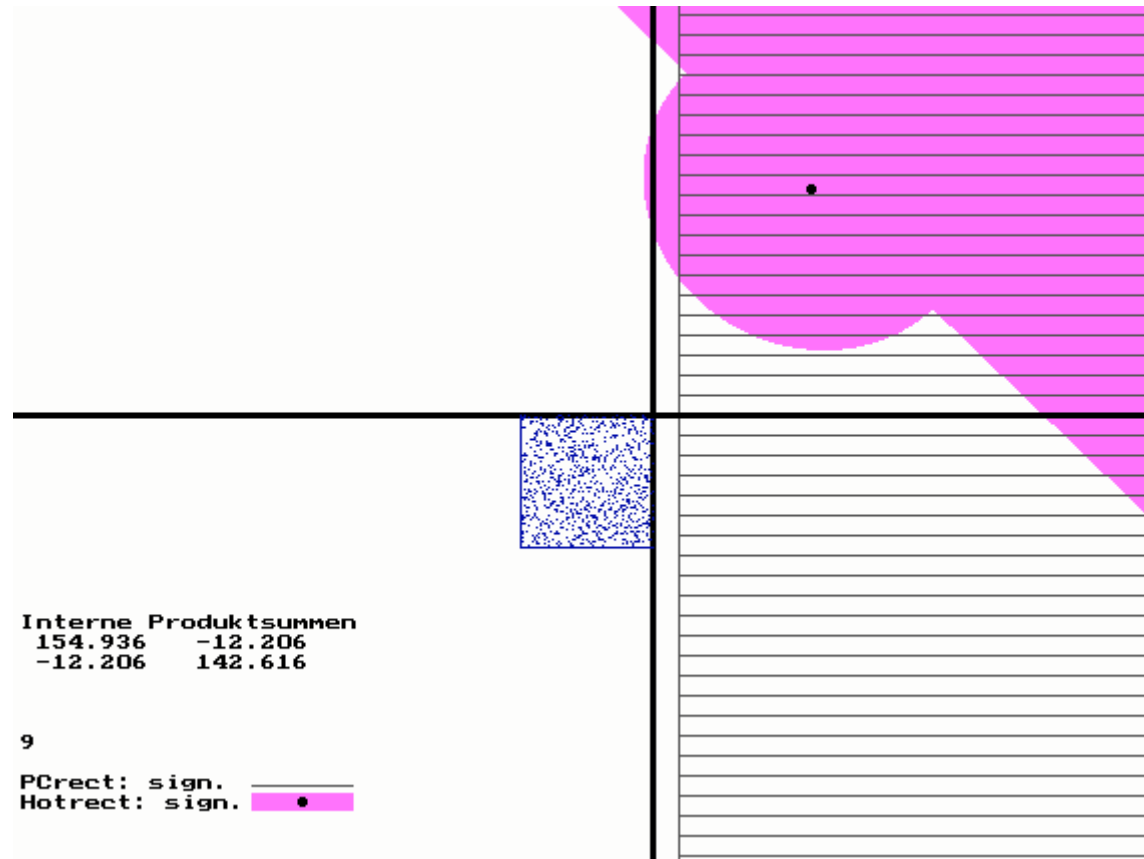




Hier der Vorteil der konservativen PC-Methode



Sehr unterschiedliche Steilheiten der Trenngeraden



PC-Test entartet zu einem eindimensionalen Test.  
 Trotzdem wird Signifikanz erreicht.

Der konservative PC-Test bewährt sich vor allem bei mehreren positiv korrelierten und in der Trennschärfe etwa gleich guten Variablen.

Er ist ein stabiler Test, d.h., die Vergrößerung der Dimension wird nicht zum Problem.

Ein Nachteil besteht darin, dass er maßstabsabhängig ist. Die Wahl der Skala ist wesentlich. Vielleicht sollte man anhand von vorliegenden Erfahrungen die Skala passend wählen.

#### 4 Anwendung des Standard-Summentests

Der SS-Test ist ein exakter sphärischer Test, bei dem ein linearer Score gebildet wird, dessen Koeffizienten positiv sind. Gerade deshalb ist er für gerichtete Tests prädestiniert.

Eine **Schwierigkeit** besteht allerdings darin, dass die Koeffizienten von der jeweiligen Nullhypothese  $\delta = \delta_0$  abhängen.

Wenn man also einen ganzen Quader von betrachteten Nullhypothesen hat, dann existieren entsprechend viele Gewichtsvektoren.

**SS-Test der Hypothese  $\delta = \delta_0$  :** 
$$t = \sqrt{(n-2)} a \frac{(\bar{x} - \delta_0)' d}{\sqrt{d' G d}} \sim t(n-2)$$

Die sog. totalen Abweichungsquadratsummen sind

$$g_{ii} + a(\bar{x}_i - \delta_{0i})^2 \quad \text{mit} \quad \bar{x}_i = \bar{x}_i^{(1)} - \bar{x}_i^{(2)} \quad (i = 1, \dots, p).$$

Daraus ergeben sich die Gewichte zur Bildung des Scores

$$d_i = \frac{1}{\sqrt{g_{ii} + a(\bar{x}_i - \delta_{0i})^2}} \quad (i = 1, \dots, p).$$

Hier ist die Abhängigkeit des Scores von der Nullhypothese zu erkennen.

Wir wollen den Test besser verstehen.

Definiert man für  $g_{ii} > 0$  die Winkel  $\xi_i$  im Intervall  $(0, \pi)$  durch

$$\sin \xi_i = \sqrt{g_{ii}} d_i \quad \text{und} \quad \cos \xi_i = \sqrt{a} (\bar{x}_i - \delta_{0i}) d_i \quad (i = 1, \dots, p),$$

so bekommt der  $t$ -Test die Gestalt

$$t = \sqrt{n-2} \frac{\sum_{i=1}^p \cos \xi_i}{\sqrt{(\sin \boldsymbol{\xi})' \mathbf{R} (\sin \boldsymbol{\xi})}},$$

wobei

die aus  $\mathbf{G}$  abgeleitete Korrelationsmatrix  $\mathbf{R} = [\text{Diag}(\mathbf{G})]^{-1/2} \mathbf{G} [\text{Diag}(\mathbf{G})]^{-1/2}$

$$\text{und der Sinusvektor } \sin \boldsymbol{\xi} = \begin{pmatrix} \sin \xi_1 \\ \vdots \\ \sin \xi_p \end{pmatrix}$$

auftreten. Zur Signifikanz gegen den Parameter  $\delta_0$  muss  $t \geq t_{1-\alpha}(n-2)$  gelten.

Die unbekannt Parameter  $\delta_{0i}$  verbergen sich in den  $\xi_i$ , sonst nirgends.

In der praktischen Anwendung wird man zunächst gegen  $\delta_0 = \mathbf{0}$  testen.

Wenn dort Signifikanz zustande kommt,

wird man fragen, ob auch Signifikanz in den anderen Parameterpunkten der negativen Ecke vorliegt.

Monotoniebetrachtungen ergeben, dass der Winkel  $\xi_i$  mit steigendem  $\bar{x}_i$  und mit fallendem  $\delta_{0i}$  kleiner wird. Das hilft bei der Aufklärung der Vorgänge.

In unserem Spezialfall  $p = 2$  kann man folgendes beweisen:

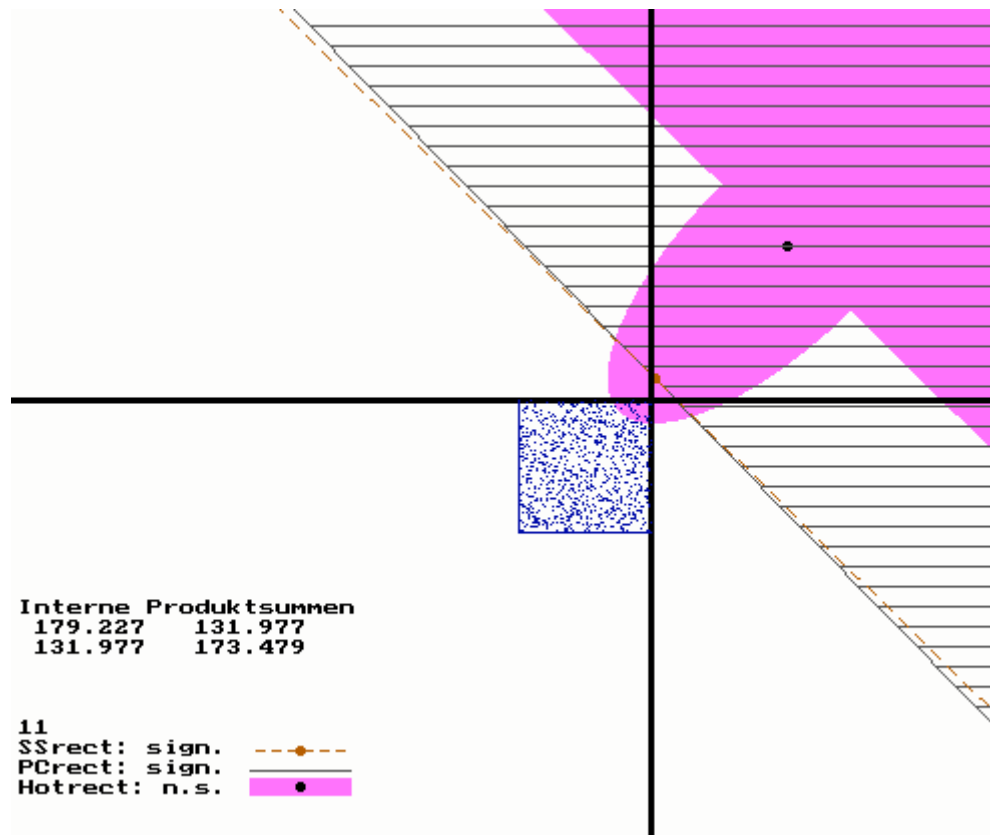
- Für  $g_{12} \geq 0$  bzw.  $r_{12} \geq 0$  hat die Statistik  $t$  bei festem  $\bar{\mathbf{x}}$  ihr Minimum an der Stelle  $\delta_0 = \mathbf{0}$ ,  
d.h., Signifikanz gegen null sichert auch die Signifikanz im ganzen Nullrechteck.
- Bei  $r_{12} < 0$  kann das Minimum von  $t$  auch noch an einer der beiden Stellen

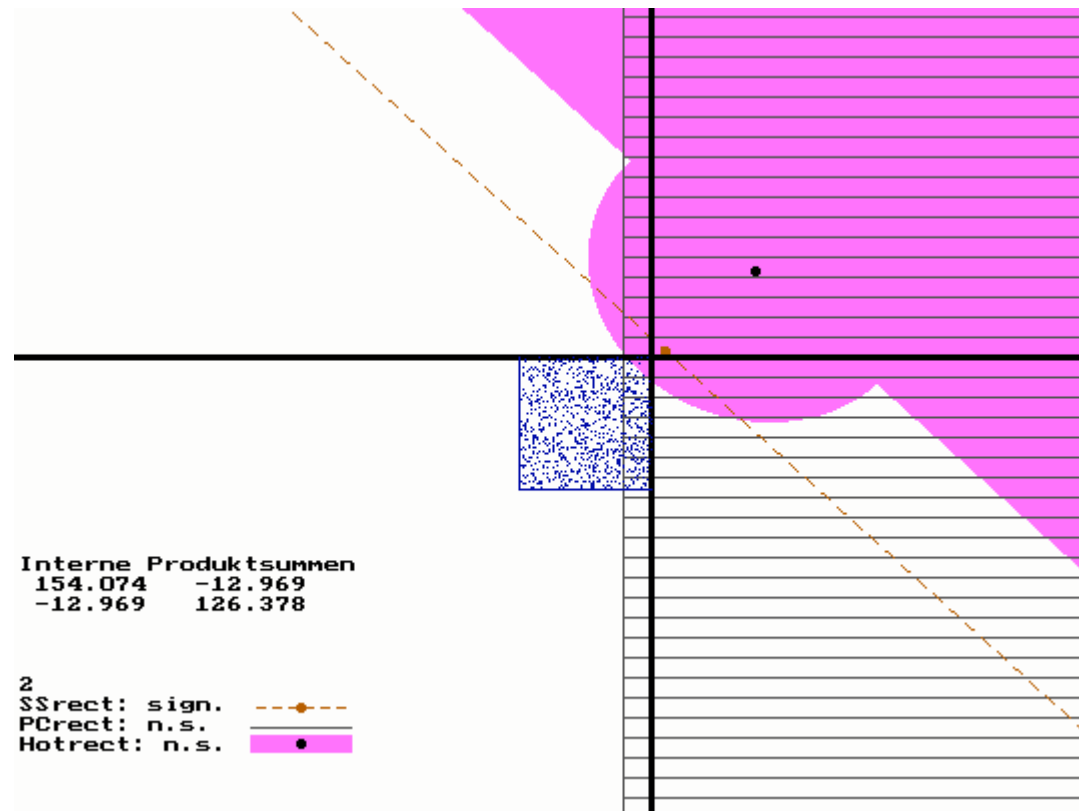
$\delta_0 = \begin{pmatrix} \varepsilon_1 \\ 0 \end{pmatrix}$  und  $\delta_0 = \begin{pmatrix} 0 \\ \varepsilon_2 \end{pmatrix}$  auftreten. Deshalb müssen mit den gegebenen Daten auch die  $t$ - Tests

gegen diese beiden zusätzlichen Nullhypothesen ausgeführt werden.

Das ist für unseren Spezialfall ausreichend.







Hier ist nur der SS-Test erfolgreich.

Ablehnungsquote einseitiger multivariater Tests,  $\alpha = 0.025$ ,  
zwei Variablen, Nullhypothesen  $-0.33 \leq \mu_{01}^{(1)} - \mu_{01}^{(2)} \leq 0$ ,  $-0.33 \leq \mu_{02}^{(1)} - \mu_{02}^{(2)} \leq 0$ ,  
 $\sigma_1^2 = \sigma_2^2 = 1$ ,  $\rho = 0.7$

$$n^{(1)} = n^{(2)} = 75$$

$$n^{(1)} = n^{(2)} = 100$$

Alternative $\delta_i = \mu_i^{(1)} - \mu_i^{(2)}$	SS-Test	PC-Test, konserv.	Hotelling, einseitig
Fall6: $\delta_1 = 0.33, \delta_2 = 0$	0.19 0.19	0.19 0.19	0.34 0.28
Fall2: $\delta_1 = 0.66, \delta_2 = 0$	0.56 0.48	0.58 0.48	0.96 0.51
Fall3: $\delta_1 = 0.66, \delta_2 = 0.166$	0.77 0.76	0.78 0.76	0.95 0.83
Fall5: $\delta_1 = 0.66, \delta_2 = 0.33$	0.90 0.90	0.90 0.90	0.94 0.93
Fall4: $\delta_1 = 0.33, \delta_2 = 0.33$	0.58 0.57	0.58 0.57	0.46 0.46

SS-Test	PC-Test, konserv.	Hotelling, einseitig
0.24 0.24	0.24 0.24	0.46 0.40
0.70 0.61	0.71 0.62	0.99 0.64
0.88 0.87	0.88 0.88	0.99 0.93
0.96 0.96	0.96 0.96	0.98 0.98
0.72 0.71	0.71 0.70	0.60 0.60

Wert 1: Multivariates Verfahren allein

Wert 2: Multivariates Verfahren in Kombination mit univariaten Tests

Ablehnungsquote einseitiger multivariater Tests,  $\alpha = 0.025$ ,  
zwei Variablen, Nullhypothesen  $-0.33 \leq \mu_{01}^{(1)} - \mu_{01}^{(2)} \leq 0$ ,  $-0.33 \leq \mu_{02}^{(1)} - \mu_{02}^{(2)} \leq 0$ ,  
 $\sigma_1^2 = \sigma_2^2 = 1$ ,  $\rho = 0$

$$n^{(1)} = n^{(2)} = 75$$

$$n^{(1)} = n^{(2)} = 100$$

Alternative $\delta_i = \mu_i^{(1)} - \mu_i^{(2)}$	SS-Test	PC-Test, konserv.	Hotelling, einseitig
Fall6: $\delta_1 = 0.33, \delta_2 = 0$	0.29 0.22	0.26 0.16	0.37 0.21
Fall2: $\delta_1 = 0.66, \delta_2 = 0$	0.79 0.49	0.57 0.34	0.95 0.49
Fall3: $\delta_1 = 0.66, \delta_2 = 0.166$	0.94 0.83	0.68 0.61	0.97 0.83
Fall5: $\delta_1 = 0.66, \delta_2 = 0.33$	0.99 0.97	0.83 0.82	0.99 0.97
Fall4: $\delta_1 = 0.33, \delta_2 = 0.33$	0.81 0.72	0.59 0.56	0.71 0.68

SS-Test	PC-Test, konserv.	Hotelling, einseitig
0.37 0.31	0.33 0.24	0.49 0.33
0.90 0.63	0.60 0.42	0.99 0.64
0.98 0.93	0.72 0.69	0.99 0.93
1.00 0.99	0.88 0.88	1.00 0.99
0.91 0.85	0.72 0.70	0.85 0.83

Wert 1: Multivariates Verfahren allein

Wert 2: Multivariates Verfahren in Kombination mit univariaten Tests

Ablehnungsquote einseitiger multivariater Tests,  $\alpha = 0.025$ ,  
zwei Variablen, Nullhypothesen  $-0.33 \leq \mu_{01}^{(1)} - \mu_{01}^{(2)} \leq 0$ ,  $-0.33 \leq \mu_{02}^{(1)} - \mu_{02}^{(2)} \leq 0$ ,  
 $\sigma_1^2 = \sigma_2^2 = 1$ ,  $\rho = -0.7$

$$n^{(1)} = n^{(2)} = 75$$

$$n^{(1)} = n^{(2)} = 100$$

Alternative $\delta_i = \mu_i^{(1)} - \mu_i^{(2)}$	SS-Test	PC-Test, konserv.	Hotelling, einseitig
Fall6: $\delta_1 = 0.33, \delta_2 = 0$	0.73 0.17	0.27 0.09	0.71 0.17
Fall2: $\delta_1 = 0.66, \delta_2 = 0$	1.00 0.51	0.50 0.27	1.00 0.51
Fall3: $\delta_1 = 0.66, \delta_2 = 0.166$	1.00 0.85	0.57 0.51	1.00 0.85
Fall5: $\delta_1 = 0.66, \delta_2 = 0.33$	1.00 0.98	0.74 0.73	1.00 0.98
Fall4: $\delta_1 = 0.33, \delta_2 = 0.33$	1.00 0.85	0.51 0.49	1.00 0.85

SS-Test	PC-Test, konserv.	Hotelling, einseitig
0.84 0.34	0.33 0.17	0.83 0.34
1.00 0.64	0.51 0.34	1.00 0.64
1.00 0.94	0.60 0.57	1.00 0.94
1.00 1.00	0.81 0.81	1.00 1.00
1.00 0.96	0.63 0.63	1.00 0.96

Wert 1: Multivariates Verfahren allein

Wert 2: Multivariates Verfahren in Kombination mit univariaten Tests

Ablehnungsquote einseitiger multivariater Tests,  $\alpha = 0.025$ ,  
zwei Variablen, Nullhypothesen  $-0.33 \leq \mu_{01}^{(1)} - \mu_{01}^{(2)} \leq 0$ ,  $-0.33 \leq \mu_{02}^{(1)} - \mu_{02}^{(2)} \leq 0$ ,  
 $\sigma_1^2 = \sigma_2^2 = 1$

Vergleich der drei Verfahren anhand der minimalen Gütwerte der Korrelationen 0.7, 0.0 und -0.7

$$n^{(1)} = n^{(2)} = 75$$

$$n^{(1)} = n^{(2)} = 100$$

Alternative	Bestes Verfahren	Bestes Verfahren
Fall6: $\delta_1 = 0.33, \delta_2 = 0$	Hot: 0.34 SS, Hot: 0.17	Hot: 0.46 Hot: 0.33
Fall2: $\delta_1 = 0.66, \delta_2 = 0$	Hot: 0.95 Hot: 0.49	Hot: 0.99 Hot: 0.64
Fall3: $\delta_1 = 0.66, \delta_2 = 0.166$	Hot: 0.95 Hot: 0.83	Hot: 0.99 Hot: 0.93
Fall5: $\delta_1 = 0.66, \delta_2 = 0.33$	Hot: 0.94 Hot: 0.93	Hot: 0.98 Hot: 0.98
Fall4: $\delta_1 = 0.33, \delta_2 = 0.33$	SS: 0.58 SS: 0.57	SS: 0.72 SS: 0.71

Wert 1: Multivariates Verfahren allein

Wert 2: Multivariates Verfahren in Kombination mit univariaten Tests

## Abschließende Bewertung

1. Wenn als wahrer Parameter irgendein Punkt des Nullquaders vorliegt, erfolgt höchstens mit der Wahrscheinlichkeit  $\alpha$  Ablehnung der Nullhypothese
2. Zur Frage einer „monotonen Entscheidung“:  
Nimmt man die ganze negative Ecke als Nullquader, dann wird die Ablehnung mit wachsenden Differenzen  $\bar{x}_i = \bar{x}_i^{(1)} - \bar{x}_i^{(2)}$  immer wahrscheinlicher.

Das gilt für alle vorgetragenen Verfahren bei jeder Dimension  $p$ , auch wenn die Verfahren im Detail nur für die Dimension  $p = 2$  beschrieben wurden.