An Application of the Closed Testing Principle to Enhance One-Sided Confidence Regions for a Multivariate Location Parameter

Michael Vock

Universität Bern
Institut für mathematische Statistik und Versicherungslehre

MCP 2007, Vienna, July 10, 2007

Research supported by the Swiss National Science Foundation

Outline

One-Sided Confidence Regions

- $\boldsymbol{X}_1, \ldots, \boldsymbol{X}_n$ i. i. d. random vectors in $\mathbb{R}^p$
- $\boldsymbol{X}_i \sim \mathrm{P}_{\boldsymbol{\vartheta}}$, $\boldsymbol{\vartheta} \in \Theta$ unknown
- $\mathrm{P}_{\boldsymbol{\vartheta}}$ (at least directionally) symmetric w. r. t. $\boldsymbol{\vartheta}$

Problem: Find a $1 - \alpha$ confidence region for $\boldsymbol{\vartheta}$ that is

- as strict as possible in specific directions
- possibly unbounded in "irrelevant" directions

(e. g. a cone or an orthant).

Connection with One-Sided Location Tests

Let φ_α be a non-randomized level α test for

$$H_0 : \boldsymbol{\vartheta} \in \Theta_0(\boldsymbol{\gamma}) \qquad \text{vs.}$$
$$H_1 : \boldsymbol{\vartheta} \in \Theta \setminus \Theta_0(\boldsymbol{\gamma}).$$

(E. g. $\Theta_0(\boldsymbol{\gamma}) = \boldsymbol{\gamma} + (-\infty, 0]^p$)

Inversion of φ_α

$\Rightarrow \mathcal{C}_{1-\alpha}(\boldsymbol{X}_1, \ldots, \boldsymbol{X}_n) = \{\boldsymbol{\gamma} : \varphi_\alpha((\boldsymbol{X}_1, \ldots, \boldsymbol{X}_n), \boldsymbol{\gamma}) = 0\}$, and

$$\mathrm{P}_{\boldsymbol{\vartheta}}(\mathcal{C}_{1-\alpha}(\mathrm{X}) \ni \boldsymbol{\gamma}) \geq 1 - \alpha \quad \forall\, \boldsymbol{\vartheta} \in \Theta_0(\boldsymbol{\gamma}) \quad \forall\, \boldsymbol{\gamma} \in \Theta.$$

$\mathcal{C}_{1-\alpha}$ is a $1-\alpha$ confidence region for the meta-parameter $\boldsymbol{\gamma}$.

Outline

Introduction

Direct Derivation of Confidence Regions for ϑ

Enhanced Confidence Regions for ϑ

Direct Derivation of Confidence Regions for ϑ

Assume that $\boldsymbol{\gamma} \in \Theta_0(\boldsymbol{\gamma}),\ \forall\, \boldsymbol{\gamma} \in \Theta$.
Then

$$\mathrm{P}_{\boldsymbol{\vartheta}}(\mathcal{C}_{1-\alpha}(\mathrm{X}) \ni \boldsymbol{\vartheta}) \geq 1 - \alpha \quad \forall\, \boldsymbol{\vartheta} \in \Theta.$$

$\mathcal{C}_{1-\alpha}$ is also a $1-\alpha$ confidence region for the location parameter $\boldsymbol{\vartheta}$.

Direct Derivation of Confidence Regions for ϑ

Assume that $\boldsymbol{\gamma} \in \Theta_0(\boldsymbol{\gamma}),\ \forall\, \boldsymbol{\gamma} \in \Theta$.
Then

$$\mathrm{P}_{\boldsymbol{\vartheta}}(\mathcal{C}_{1-\alpha}(\mathrm{X}) \ni \boldsymbol{\vartheta}) \geq 1-\alpha \quad \forall\, \boldsymbol{\vartheta} \in \Theta.$$

$\mathcal{C}_{1-\alpha}$ is also a $1-\alpha$ confidence region for the location parameter $\boldsymbol{\vartheta}$.

Problems:

- conservative
- unpleasant shape – to be illustrated...

Min and Max Tests

Min Test

Reject $H_0 : \exists\, j \in \{1, \ldots, p\} : \vartheta_j \leq \gamma_j$ in favor of $H_1 : \boldsymbol{\vartheta} > \boldsymbol{\gamma}$ at the level α if and only if

$$\varphi_{j,\alpha}((X_{1j}, \ldots, X_{nj}), \gamma_j) = 1, \forall\, j \in 1, \ldots, p.$$

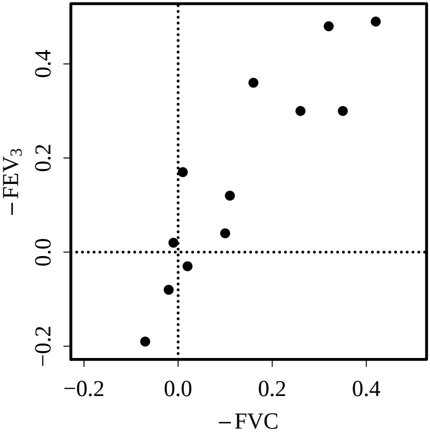
Bonferroni Max Test

Reject $H_0 : \boldsymbol{\vartheta} \leq \boldsymbol{\gamma}$ in favor of $H_1 : \exists\, j \in \{1, \ldots, p\} : \vartheta_j > \gamma_j$ at the level α if and only if

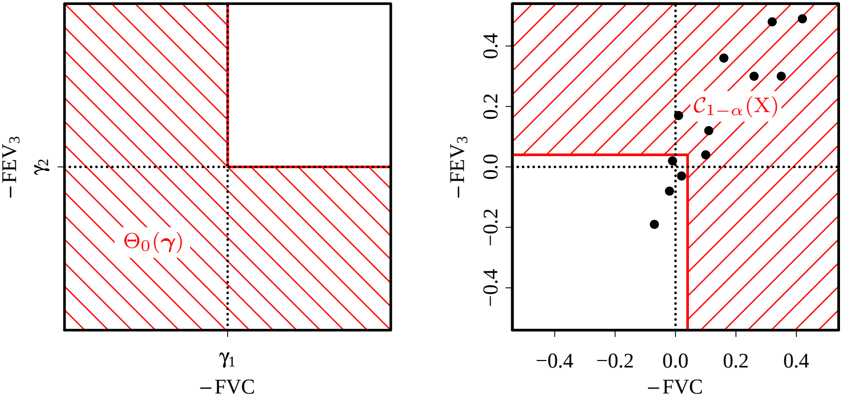
$$\exists\, j \in 1, \ldots, p : \varphi_{j,\alpha/p}((X_{1j}, \ldots, X_{nj}), \gamma_j) = 1.$$

Example

Two variables of the pulmonary function data by Randles (1989) (slightly modified from Merchant et al., 1975).

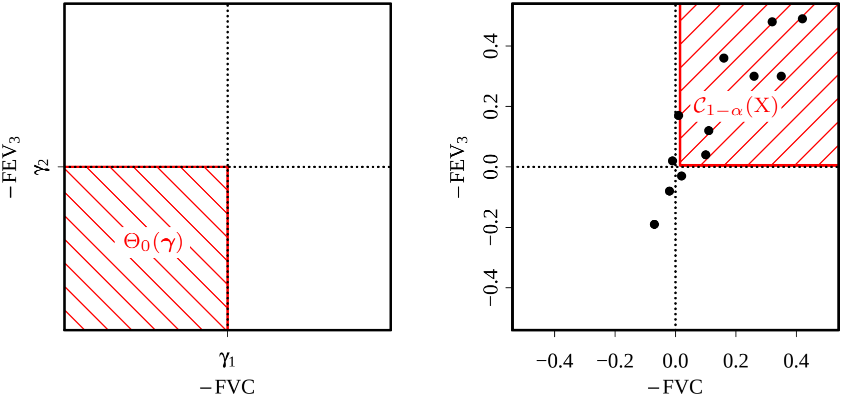


Pulmonary Function Data, Wilcoxon Min Test



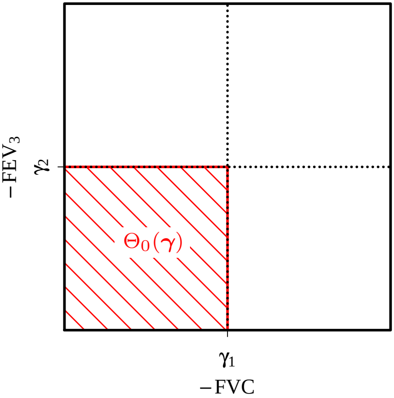
$$\mathcal{C}_{1-\alpha}(\mathrm{X}) = \boldsymbol{c}_{1-\alpha}(\mathrm{X}) - \Theta_0(\mathbf{0})$$

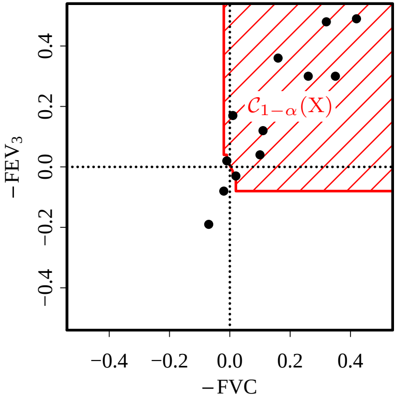
P. F. Data, Wilcoxon Bonferroni Max Test



$$\mathcal{C}_{1-\alpha}(\mathrm{X}) = \boldsymbol{c}_{1-\alpha}(\mathrm{X}) - \Theta_0(\mathbf{0})$$

P. F. Data, Sign Test by Larocque/Labarre (2004)





$\mathcal{C}_{1-\alpha}(\mathrm{X}) \approx \boldsymbol{c}_{1-\alpha}(\mathrm{X}) - \Theta_0(\mathbf{0})$ (outside a sufficiently large ball)

Back to the Drawbacks of the Direct Approach

If confidence regions for γ are directly used as confidence regions for ϑ, they are

- usually conservative and
- similar in shape to $-\Theta_0$, rather than to $\Theta_1 = \Theta \setminus \Theta_0$.

Outline

Enhanced Confidence Regions for ϑ (1)

Temptation: $\tilde{\mathcal{C}}_{1-\alpha}(\mathrm{X}) = \bigcap_{\boldsymbol{\gamma} \notin \mathcal{C}_{1-\alpha}(\mathrm{X})} \Theta_1(\boldsymbol{\gamma})$

Enhanced Confidence Regions for ϑ (1)

Temptation: $\tilde{\mathcal{C}}_{1-\alpha}(\mathrm{X}) = \bigcap_{\boldsymbol{\gamma} \notin \mathcal{C}_{1-\alpha}(\mathrm{X})} \Theta_1(\boldsymbol{\gamma})$

$\Rightarrow$ liberal – multiple testing problem!

Solution: Reduce the set of possible meta-parameters in advance.

Enhanced Confidence Regions for ϑ (2)

Let $\mathcal{C}_{1-\alpha} : \mathcal{X} \to \mathcal{P}(\mathbb{R}^p)$ be a $1-\alpha$ confidence region for γ based on $(\Theta_0(\boldsymbol{\gamma}))_{\boldsymbol{\gamma}\in\mathbb{R}^p}$.
Let $\Theta_0(\boldsymbol{\gamma}) = \boldsymbol{\gamma} + \Theta_0(\mathbf{0}), \forall\, \boldsymbol{\gamma} \in \mathbb{R}^p$, closed, $\Theta_1(\boldsymbol{\gamma}) = \mathbb{R}^p \setminus \Theta_0(\boldsymbol{\gamma})$.

Assume that $\Theta_0(\boldsymbol{\gamma}) \subset \Theta_0(\boldsymbol{\gamma} + (\delta, \ldots, \delta)^{\mathrm{T}}), \forall\, \boldsymbol{\gamma} \in \mathbb{R}^p, \delta > 0$.

With $\boldsymbol{\gamma}_i = (i, \ldots, i)^{\mathrm{T}} \in \mathbb{R}^p, \forall\, i \in I = [\ell, \infty)$, define

$$\tilde{\mathcal{C}}_{1-\alpha}(\mathrm{X}) := \bigcap_{i\in I:\boldsymbol{\gamma}_{i'}\notin\mathcal{C}_{1-\alpha}(\mathrm{X})\ \forall\, i'\leq i} \Theta_1(\boldsymbol{\gamma}_i).$$

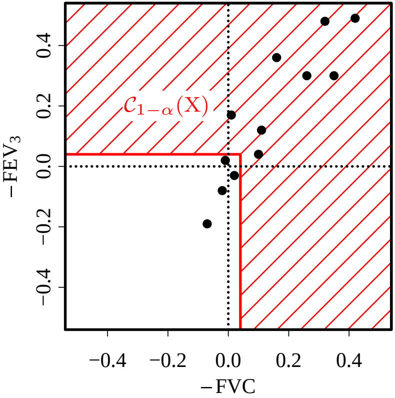
Then

$$\mathrm{P}_{\boldsymbol{\vartheta}}\left(\tilde{\mathcal{C}}_{1-\alpha}(\mathrm{X}) \ni \boldsymbol{\vartheta}\right) \geq 1-\alpha \quad \forall\, \boldsymbol{\vartheta} \in \mathbb{R}^p.$$

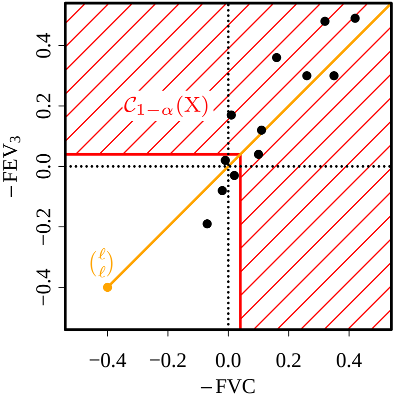
Idea of the Proof

- $(\Theta_0(\boldsymbol{\gamma}_i))_{i\in I}$ is closed under (finite and infinite) intersections.
- Apply the closed testing principle (Marcus, Peritz, and Gabriel, 1976).
- Translate to confidence regions.

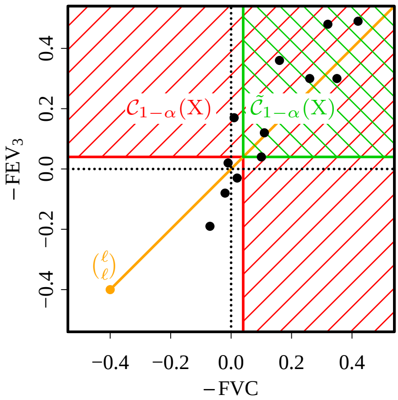
Pulmonary Function Data, Wilcoxon Min Test



Pulmonary Function Data, Wilcoxon Min Test



Pulmonary Function Data, Wilcoxon Min Test



Enhanced Confidence Regions: Properties

- $\tilde{\mathcal{C}}_{1-\alpha}(\mathrm{X}) = \boldsymbol{c}_{1-\alpha}(\mathrm{X}) + \Theta_1(\mathbf{0})$ (by definition)
- $\tilde{\mathcal{C}}_{1-\alpha}(\mathrm{X}) \subset \mathcal{C}_{1-\alpha}(\mathrm{X})$ under suitable conditions ($\Theta_1(\mathbf{0})$ convex cone, translation invariance, and a monotonicity property)

Enhanced Confidence Regions: Properties

- $\tilde{\mathcal{C}}_{1-\alpha}(\mathrm{X}) = \boldsymbol{c}_{1-\alpha}(\mathrm{X}) + \Theta_1(\mathbf{0})$ (by definition)
- $\tilde{\mathcal{C}}_{1-\alpha}(\mathrm{X}) \subset \mathcal{C}_{1-\alpha}(\mathrm{X})$ under suitable conditions ($\Theta_1(\mathbf{0})$ convex cone, translation invariance, and a monotonicity property)

Disadvantage: Restricted set of possible confidence regions (search on the diagonal)

Summary

- A confidence region $\mathcal{C}_{1-\alpha}(\mathrm{X})$ obtained by inversion of a test for a composite null hypothesis is for a meta-parameter.
- Even if $\mathcal{C}_{1-\alpha}(\mathrm{X})$ may also be a confidence region for the parameter $\boldsymbol{\vartheta}$ itself, it is not very useful.
- The proposed method based on the closed testing principle yields a confidence region $\tilde{\mathcal{C}}_{1-\alpha}(\mathrm{X})$ with a more useful shape.
- $\tilde{\mathcal{C}}_{1-\alpha}(\mathrm{X})$ is also less conservative than $\mathcal{C}_{1-\alpha}(\mathrm{X})$ under suitable conditions.

References

- Larocque, D., Labarre, M. (2004). A conditionally distribution-free multivariate sign test for one-sided alternatives. *Journal of the American Statistical Association*, **99**, 499–509.
- Marcus, R., Peritz, E., Gabriel, K. R. (1976). On closed testing procedures with special reference to ordered analysis of variance. *Biometrika*, **63**, 655–660.
- Merchant, J. A., Halprin, G. M., Hudson, A. R., Kilburn, K. H., McKenzie, W. N., Hurst, D. J., Bermazohn, P. (1975). Responses to cotton dust. *Archives of Environmental Health*, **30**, 222–229.
- Randles, R. H. (1989). A distribution-free multivariate sign test based on interdirections. *Journal of the American Statistical Association*, **84**, 1045–1050.
- Vock, M. (2007). *Enhanced one-sided confidence regions for a multivariate location parameter*, submitted.