Confidence sets for optimal factor levels of a response surface

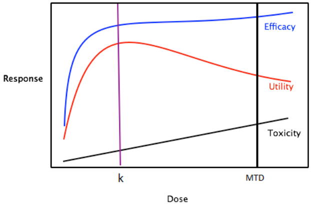
Fang Wan, Wei Liu, Frank Bretz, and Yang Han

Medical and Pharmaceutical Statistics Research Unit

Mathematics & Statistics | Lancaster University

Outline

Section



Dose response for efficacy and safety.

- Low doses, small efficacy; High doses, safety problem.
- CUI combines efficacy and safety into one response function.
- CUI is almost always unimodal.

Model specification

$Y = f(\boldsymbol{x}, \boldsymbol{\theta}) + \epsilon$

- $\epsilon \sim N(0, \sigma^2)$
- $f(\boldsymbol{x}, \boldsymbol{\theta}) = \mathbf{z}(\boldsymbol{x})^T \boldsymbol{\theta}$
- $\boldsymbol{x} = (x_1, \cdots, x_q)^T$, $\boldsymbol{\theta} = (\theta_1, \cdots, \theta_p)^T$
- a constraint covariate region χ
- $\boldsymbol{k} = \boldsymbol{k}(\boldsymbol{\theta})$ is a maximum point of $f(\boldsymbol{x}, \boldsymbol{\theta})$ in χ

Problem

Suppose $(Y_1, \boldsymbol{x}_1), (Y_2, \boldsymbol{x}_2), \cdots, (Y_n, \boldsymbol{x}_n)$ are n observations, we want to construct a $(1 - \alpha)$ level confidence set for $\boldsymbol{k}$.

Model specification

$Y = f(\boldsymbol{x}, \boldsymbol{\theta}) + \epsilon$

- $\epsilon \sim N(0, \sigma^2)$
- $f(\boldsymbol{x}, \boldsymbol{\theta}) = \mathbf{z}(\boldsymbol{x})^T \boldsymbol{\theta}$
- $\boldsymbol{x} = (x_1, \cdots, x_q)^T$, $\boldsymbol{\theta} = (\theta_1, \cdots, \theta_p)^T$
- a constraint covariate region χ
- $\boldsymbol{k} = \boldsymbol{k}(\boldsymbol{\theta})$ is a maximum point of $f(\boldsymbol{x}, \boldsymbol{\theta})$ in χ

Problem

Suppose $(Y_1, \boldsymbol{x}_1), (Y_2, \boldsymbol{x}_2), \cdots, (Y_n, \boldsymbol{x}_n)$ are n observations, we want to construct a $(1 - \alpha)$ level confidence set for $\boldsymbol{k}$.

Method given by Rao (1973)

The $(1-\alpha)$ conservative confidence set for $\boldsymbol{k}$ is given by

$$\mathbf{C}_R(\boldsymbol{Y}) \quad = \{\boldsymbol{k}(\boldsymbol{\theta}) \in \chi : \boldsymbol{\theta} \in C_{\boldsymbol{\theta}}\}$$

where $C_{\boldsymbol{\theta}}$ is a $(1-\alpha)$ level confidence set for $\boldsymbol{\theta}$.

A choice of $C_{\boldsymbol{\theta}}$

$$C_{\boldsymbol{\theta}} = \{\beta : (\hat{\boldsymbol{\theta}} - \beta)^T(\boldsymbol{X}^T\boldsymbol{X})(\hat{\boldsymbol{\theta}} - \beta) \leq p\hat{\sigma}^2 f^{\alpha}_{p,\nu}\}.$$

Method given by Rao (1973)

The $(1-\alpha)$ conservative confidence set for $\boldsymbol{k}$ is given by

$$\mathbf{C}_R(\boldsymbol{Y}) \quad = \{\boldsymbol{k}(\boldsymbol{\theta}) \in \chi : \boldsymbol{\theta} \in C_{\boldsymbol{\theta}}\}$$

where $C_{\boldsymbol{\theta}}$ is a $(1-\alpha)$ level confidence set for $\boldsymbol{\theta}$.

A choice of $C_{\boldsymbol{\theta}}$

$$C_{\boldsymbol{\theta}} = \{\beta : (\hat{\boldsymbol{\theta}} - \beta)^T(\boldsymbol{X}^T\boldsymbol{X})(\hat{\boldsymbol{\theta}} - \beta) \leq p\hat{\sigma}^2 f^{\alpha}_{p,\nu}\}.$$

Section

Theorem (Neyman, 1937)

Suppose random observation $\boldsymbol{Y}$ has distribution $h(\boldsymbol{y};\boldsymbol{\gamma})$, where $\boldsymbol{\gamma}$ is the unknown parameter. Let $\mathbb{B}$ and Ω be the parameter space and sample space, respectively. For each $\boldsymbol{\gamma}_0 \in \mathbb{B}$, let $A(\boldsymbol{\gamma}_0) \subset \Omega$ be the acceptance set of a size α test of $H_0 : \boldsymbol{\gamma} = \boldsymbol{\gamma}_0$. For each $\boldsymbol{Y} \in \Omega$, define a set $\boldsymbol{C}(\boldsymbol{Y}) \subset \mathbb{B}$ by $\boldsymbol{C}(\boldsymbol{Y}) = \{ \boldsymbol{\gamma}_0 : \boldsymbol{Y} \in A(\boldsymbol{\gamma}_0) \}$. Then the random set $\boldsymbol{C}(\boldsymbol{Y})$ is a $(1-\alpha)$ level confidence set for $\boldsymbol{\gamma}$.

A $(1-\alpha)$ level acceptance set for testing $H_0 : \boldsymbol{k} = \boldsymbol{k}_0$

$$A(\boldsymbol{k}_0) = \{\mathbf{Y} : f(\boldsymbol{k}_0, \hat{\boldsymbol{\theta}}) - f(\boldsymbol{x}, \hat{\boldsymbol{\theta}}) \geq -c(\boldsymbol{k}_0)\sqrt{v(\boldsymbol{k}_0, \boldsymbol{x}, \hat{\boldsymbol{\theta}})}, \ \forall \boldsymbol{x} \in \chi\},$$

where $v(\boldsymbol{k}_0, \boldsymbol{x}, \hat{\boldsymbol{\theta}})$ is the estimated variance of $f(\boldsymbol{k}_0, \hat{\boldsymbol{\theta}}) - f(\boldsymbol{x}, \hat{\boldsymbol{\theta}})$ and $c(\boldsymbol{k}_0)$ is chosen such that

$$P\{\boldsymbol{Y} \in A(\boldsymbol{k}_0)\} = 1 - \alpha.$$

Computation of $c(\boldsymbol{k}_0)$

- Analytical method (for univariate simple and quadratic linear functions)
- Simulation method

A $(1-\alpha)$ level acceptance set for testing $H_0 : \boldsymbol{k} = \boldsymbol{k}_0$

$$A(\boldsymbol{k}_0) = \{\mathbf{Y} : f(\boldsymbol{k}_0, \hat{\boldsymbol{\theta}}) - f(\boldsymbol{x}, \hat{\boldsymbol{\theta}}) \geq -c(\boldsymbol{k}_0)\sqrt{v(\boldsymbol{k}_0, \boldsymbol{x}, \hat{\boldsymbol{\theta}})}, \ \forall \boldsymbol{x} \in \chi\},$$

where $v(\boldsymbol{k}_0, \boldsymbol{x}, \hat{\boldsymbol{\theta}})$ is the estimated variance of $f(\boldsymbol{k}_0, \hat{\boldsymbol{\theta}}) - f(\boldsymbol{x}, \hat{\boldsymbol{\theta}})$ and $c(\boldsymbol{k}_0)$ is chosen such that

$$P\{\boldsymbol{Y} \in A(\boldsymbol{k}_0)\} = 1 - \alpha.$$

Computation of $c(\boldsymbol{k}_0)$

- Analytical method (for univariate simple and quadratic linear functions)
- Simulation method

A $(1-\alpha)$ level confidence set for $\boldsymbol{k}$

$$
\begin{aligned}
\mathbf{C}_E(\boldsymbol{Y}) &= \{\boldsymbol{k}_0 \in \chi : \boldsymbol{Y} \in A(\boldsymbol{k}_0)\} \\
&= \{\boldsymbol{k}_0 \in \chi : f(\boldsymbol{k}_0, \hat{\boldsymbol{\theta}}) - f(\boldsymbol{x}, \hat{\boldsymbol{\theta}}) \geq -c(\boldsymbol{k}_0)\sqrt{v(\boldsymbol{k}_0, \boldsymbol{x}, \hat{\boldsymbol{\theta}})}, \forall \boldsymbol{x} \in \chi\}.
\end{aligned}
$$

A by-product

Substitute $c_0 = \sqrt{pF^{\alpha}_{p,\nu}}$ for $c(\boldsymbol{k}_0)$, then we get a conservative solution

$$\mathbf{C}_0(\boldsymbol{Y}) = \{\boldsymbol{k}_0 \in \chi : f(\boldsymbol{k}_0, \hat{\boldsymbol{\theta}}) - f(\boldsymbol{x}, \hat{\boldsymbol{\theta}}) \geq -c_0\sqrt{v(\boldsymbol{k}_0, \boldsymbol{x}, \hat{\boldsymbol{\theta}})}, \forall \boldsymbol{x} \in \chi\}.$$

The function of set $\mathbf{C}_0$

- Conservative confidence set for $\boldsymbol{k}$, i.e., confidence level is larger than $(1-\alpha)$.
- Reduces the computation burden of constructing $\mathbf{C}_E$.

A by-product

Substitute $c_0 = \sqrt{pF^{\alpha}_{p,\nu}}$ for $c(\boldsymbol{k}_0)$, then we get a conservative solution

$$\mathbf{C}_0(\boldsymbol{Y}) = \{\boldsymbol{k}_0 \in \chi : f(\boldsymbol{k}_0, \hat{\boldsymbol{\theta}}) - f(\boldsymbol{x}, \hat{\boldsymbol{\theta}}) \geq -c_0\sqrt{v(\boldsymbol{k}_0, \boldsymbol{x}, \hat{\boldsymbol{\theta}})}, \forall \boldsymbol{x} \in \chi\}.$$

The function of set $\mathbf{C}_0$

- Conservative confidence set for $\boldsymbol{k}$, i.e., confidence level is larger than $(1-\alpha)$.
- Reduces the computation burden of constructing $\mathbf{C}_E$.

Section

GLMs

Confidence set for a maximum point $\boldsymbol{k}$ of $E(Y)$

Suppose we have the generalized linear model

$$g[E(Y)] = f(\boldsymbol{x}, \boldsymbol{\theta}),$$

where the link function g is strictly (increasing) monotone and differentiable.

Convert the GLMs problem to LMs problem

$$argmax_{\boldsymbol{x} \in \chi} E(Y) \Longleftrightarrow argmax_{\boldsymbol{x} \in \chi} f(\boldsymbol{x}, \boldsymbol{\theta})$$

GLMs

Confidence set for a maximum point $\boldsymbol{k}$ of $E(Y)$

Suppose we have the generalized linear model

$$g[E(Y)] = f(\boldsymbol{x}, \boldsymbol{\theta}),$$

where the link function g is strictly (increasing) monotone and differentiable.

Convert the GLMs problem to LMs problem

$$argmax_{\boldsymbol{x}\in\chi} E(Y) \Longleftrightarrow argmax_{\boldsymbol{x}\in\chi} f(\boldsymbol{x}, \boldsymbol{\theta})$$

Cox-Proportional hazard model

Confidence set for a minimum point of $h(t, \boldsymbol{x})$

Suppose we have the proportional hazard model

$$h(t, \boldsymbol{x}) = \lambda(t) exp(f(\boldsymbol{x}, \boldsymbol{\theta})).$$

The interest is in constructing a confidence set for a minimum point of $h(t, \boldsymbol{x})$.

Convert to LM problem

$$argmin_{\boldsymbol{x} \in \chi} h(t, \boldsymbol{x}) \iff argmin_{\boldsymbol{x} \in \chi} f(\boldsymbol{x}, \boldsymbol{\theta}) \iff argmax_{\boldsymbol{x} \in \chi} f(\boldsymbol{x}, -\boldsymbol{\theta})$$

Cox-Proportional hazard model

Confidence set for a minimum point of $h(t, \boldsymbol{x})$

Suppose we have the proportional hazard model

$$h(t, \boldsymbol{x}) = \lambda(t) exp(f(\boldsymbol{x}, \boldsymbol{\theta})).$$

The interest is in constructing a confidence set for a minimum point of $h(t, \boldsymbol{x})$.

Convert to LM problem

$$argmin_{\boldsymbol{x} \in \chi} h(t, \boldsymbol{x}) \Longleftrightarrow argmin_{\boldsymbol{x} \in \chi} f(\boldsymbol{x}, \boldsymbol{\theta}) \Longleftrightarrow argmax_{\boldsymbol{x} \in \chi} f(\boldsymbol{x}, -\boldsymbol{\theta})$$

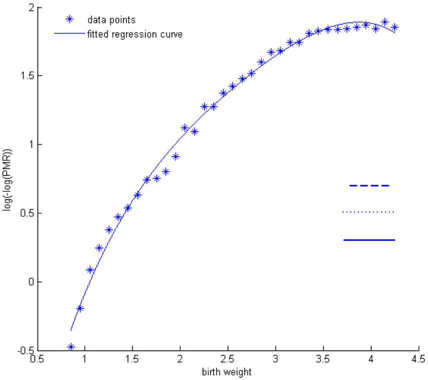
Section

Example 1

The transformed perinatal mortality rate (PMR) $Y = log(-log(PMR))$ and birth weight (BW) $x = BW$ of white infants at 35 different levels of x studied by Selvin (1998).

Table: Perinatal mortality rate data for white infants

x	Y	x	Y	x	Y	x	Y	x	Y
0.85	-0.4761	1.55	0.6340	2.25	1.2771	2.95	1.6751	3.65	1.8351
0.95	-0.1950	1.65	0.7391	2.35	1.2771	3.05	1.6830	3.75	1.8437
1.05	0.0849	1.75	0.7551	2.45	1.3731	3.15	1.7429	3.85	1.8527
1.15	0.2464	1.85	0.8042	2.55	1.4241	3.25	1.7429	3.95	1.8722
1.25	0.3791	1.95	0.9128	2.65	1.4775	3.35	1.8114	4.05	1.8437
1.35	0.4715	2.05	1.1204	2.75	1.5165	3.45	1.8269	4.15	1.8939
1.45	0.5364	2.15	1.0919	2.85	1.6018	3.55	1.8351	4.25	1.8527



The three 95% confidence sets:
$\mathbf{C}_E(\mathbf{Y}) = [3.75, 4.21]$,
$\mathbf{C}_0(\mathbf{Y}) = [3.72, 4.25]$,
$\mathbf{C}_c(\mathbf{Y}) = [3.72, 4.25]$.

Example 2

5FU+VM26 combination experiment (Stablein et al., 1983)

Treatment levels		Days of survival
5FU(mg/kg)	VM26(mg/kg)	
0.0	0.00	8,9(2),10(5)
0.0	9.71	10,13(5),14(2)
0.0	19.40	8,10,13,14(4),15
0.0	25.90	9,14(4),15(3)
35.6	9.71	13,14(3),15(3),17
48.5	4.85	9,13(2),14(3),15(2)
48.5	19.40	14(2),15(2),16(4)
97.1	0.00	8(2),10,11,12(2),14,16
97.1	3.56	8,9(2),11(2),13(2),16
97.1	9.71	8,10,11,16(2),17(2),18
97.1	25.9	16(3),17,18(3),19
194.0	0.00	10, 13(6),14
194.0	4.85	11(2),14(3),16,17
194.0	19.40	8,14,16,20(4),21
259.0	0.00	9,11,12(3),13(3)
259.0	9.71	16(2),17,18(2),19(2),20

The number in the parentheses indicates the number of animals dead on that day.

Model fitting

- $x_1 = (5FU - 130)/130$, $x_2 = (VM26 - 13)/13$
- $h(t, \boldsymbol{x}) = \lambda(t) exp(f(\boldsymbol{x}, \boldsymbol{\theta}))$
- $f(\boldsymbol{x}, \boldsymbol{\theta}) = \theta_1 x_1 + \theta_2 x_2 + \theta_3 x_1^2 + \theta_4 x_2^2 + \theta_5 x_1 x_2$
- $\chi = \{x_1, x_2 : x_1^2 + x_2^2 = 1\}$

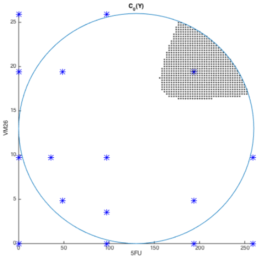


Figure: Confidence set $\mathbf{C}_0(\boldsymbol{Y})$

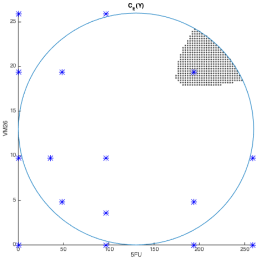


Figure: Confidence set $\mathbf{C}_E(\boldsymbol{Y})$

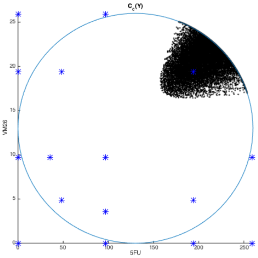


Figure: Confidence set $\mathbf{C}_R(\boldsymbol{Y})$

Section

Conclusion

The confidence set $\mathbf{C}_E(\mathbf{Y})$ is always smaller than $\mathbf{C}_0(\mathbf{Y})$ and $\mathbf{C}_R(\mathbf{Y})$. Hence it is recommended to always use the confidence set $\mathbf{C}_E(\mathbf{Y})$.

Future works

- Construction of a confidence set for the optimal factor levels of a function of a more general form.
- Construction of a confidence set by inverting acceptance sets in different forms, such as

$$A(\boldsymbol{k}_0) = \left\{ \boldsymbol{Y} : f(\boldsymbol{k}_0, \hat{\boldsymbol{\theta}}) - f(\boldsymbol{x}, \hat{\boldsymbol{\theta}}) \geq -c(\boldsymbol{k}_0)\hat{\sigma}, \forall \boldsymbol{x} \in \chi \right\}.$$

Conclusion

The confidence set $\mathbf{C}_E(\mathbf{Y})$ is always smaller than $\mathbf{C}_0(\mathbf{Y})$ and $\mathbf{C}_R(\mathbf{Y})$. Hence it is recommended to always use the confidence set $\mathbf{C}_E(\mathbf{Y})$.

Future works

- Construction of a confidence set for the optimal factor levels of a function of a more general form.
- Construction of a confidence set by inverting acceptance sets in different forms, such as

$$A(\boldsymbol{k}_0) = \left\{ \boldsymbol{Y} : f(\boldsymbol{k}_0, \hat{\boldsymbol{\theta}}) - f(\boldsymbol{x}, \hat{\boldsymbol{\theta}}) \geq -c(\boldsymbol{k}_0)\hat{\sigma}, \forall \boldsymbol{x} \in \chi \right\}.$$

Thanks!

DOI:10.1111/biom.12500

References

Neyman, J. (1937) Outline of a theory of statistical estimation based on the classical theory of probability. *Philosophical Transaction of the Royal Society of London, Series A, Mathematical and Physical Sciences*, **236**, 333–380.

Rao, C. (1973) *Linear Statistical Inference and Its Application, Second Edition*. John Wiley.

Selvin, S. (1998) *Modern Applied Biostatistical Methods Using S-Plus*. Oxford University Press.

Stablein, D. M., Carter, W. H. and Wampler, G. L. (1983) Confidence regions for constrained optima in response-surface experiments. *Biometrics*, **39**, 759–763.