Comparison of Asymptotic Confidence Sets for Regression in Small Samples: Supplementary Materials

Dmitry Kolobkov[*,1,2], Oleg Demin[1], Evgeny Metelkin[1]

[1]Institute for Systems Biology, Nauchny proezd, 20, bldg 2, Moscow, Russia

[2]Department of Bioengineering and Bioinformatics, Lomonosov Moscow State University, Leninskie Gory, 1, Moscow, Russia

[*]Corresponding author e-mail: kolobkov@insysbio.ru

1 Appendix A

Theorem 1. Confidence region $\Omega = \{f(\boldsymbol{\theta}) > c\}$ at level p is the unique optimal confidence region at level p.

Proof.

Let us consider another unique optimal confidence region Ω' at level p.

If $\Omega' \subset \Omega$ then $p = \int\limits_{\Omega} f(\boldsymbol{\theta})\mathrm{d}\boldsymbol{\theta} = \int\limits_{\Omega'} f(\boldsymbol{\theta})\mathrm{d}\boldsymbol{\theta} + \int\limits_{\Omega\setminus\Omega'} f(\boldsymbol{\theta})\mathrm{d}\boldsymbol{\theta}$, thus, $\int\limits_{\Omega'} f(\boldsymbol{\theta})\mathrm{d}\boldsymbol{\theta} < p$. Contradiction.

Analogously, if $\Omega \subset \Omega'$ then $\int\limits_{\Omega'} f(\boldsymbol{\theta})\mathrm{d}\boldsymbol{\theta} > p$.

Hence, $\Omega \not\subset \Omega'$ and $\Omega' \not\subset \Omega$.

We have $\int\limits_{\Omega} f(\boldsymbol{\theta})\mathrm{d}\boldsymbol{\theta} = \int\limits_{\Omega'} f(\boldsymbol{\theta})\mathrm{d}\boldsymbol{\theta} \Leftrightarrow \int\limits_{\Omega\setminus\{\Omega\cap\Omega'\}} f(\boldsymbol{\theta})\mathrm{d}\boldsymbol{\theta} = \int\limits_{\Omega'\setminus\{\Omega\cap\Omega'\}} f(\boldsymbol{\theta})\mathrm{d}\boldsymbol{\theta}$.

Using the mean value theorem we get $\boldsymbol{\theta_1} \in \Omega\setminus\{\Omega\cap\Omega'\} : \int\limits_{\Omega\setminus\{\Omega\cap\Omega'\}} f(\boldsymbol{\theta})\mathrm{d}\boldsymbol{\theta} = f(\boldsymbol{\theta_1})V(\Omega\setminus\{\Omega\cap\Omega'\})$, where $V()$ is the region volume.

Analogously, $\exists\boldsymbol{\theta_2} \in \Omega'\setminus\{\Omega\cap\Omega'\} : \int\limits_{\Omega'\setminus\{\Omega\cap\Omega'\}} f(\boldsymbol{\theta})\mathrm{d}\boldsymbol{\theta} = f(\boldsymbol{\theta_2})V(\Omega'\setminus\{\Omega\cap\Omega'\})$.

Since $f(\boldsymbol{\theta_2}) \le c < f(\boldsymbol{\theta_1})$, $V(\Omega\setminus\{\Omega\cap\Omega'\}) < V(\Omega'\setminus\{\Omega\cap\Omega'\}) \Leftrightarrow V(\Omega) < V(\Omega')$.

2 Appendix B

B.1 Estimating actual coverage of confidence intervals for the linear model.

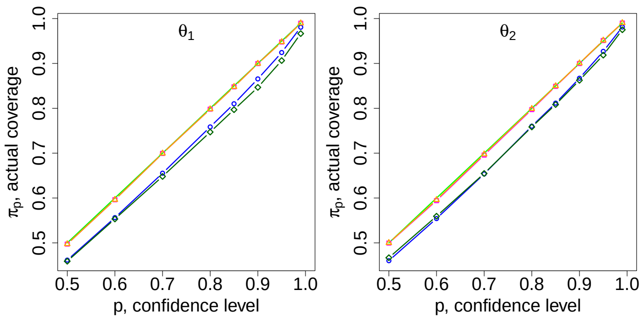

Figure 1: The linear model. Purple squares stand for the linearization method, orange circles – F-test, blue triangles – LR-test, dark green diamonds – bootstrap. Green color indicates the asymptote.

B.2 Estimating actual coverage of pointwise confidence bands for the linear model.

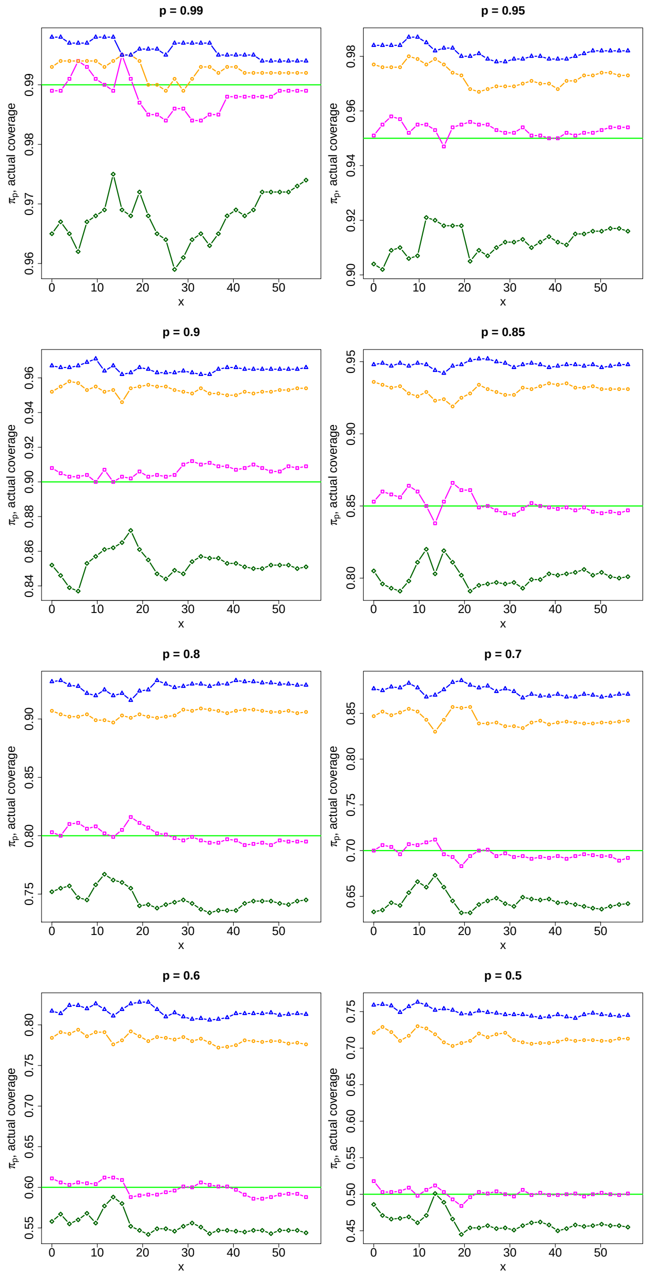

Figure 2: The linear model. Purple squares stand for the linearization method, orange circles – F-test, blue triangles – LR-test, dark green diamonds – bootstrap. Green color indicates the asymptote.

B.3 Altering experimental conditions: accuracy of 90% confidence intervals for the linear model.

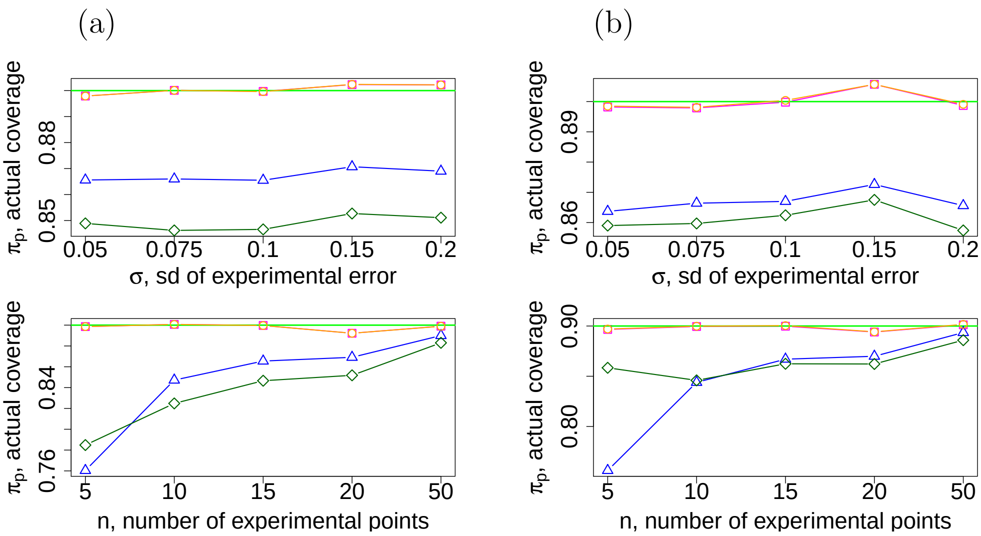

Figure 3: Altering experimental conditions: (a) accuracy of confidence intervals for θ_1, (b) accuracy of confidence intervals for θ_2. Purple color corresponds to the linearization method, orange – F-test, blue – LR-test, dark green – bootstrap. Green color indicates the asymptote.

3 Appendix C

C.1 Estimating actual coverage of confidence intervals for the Hill model.

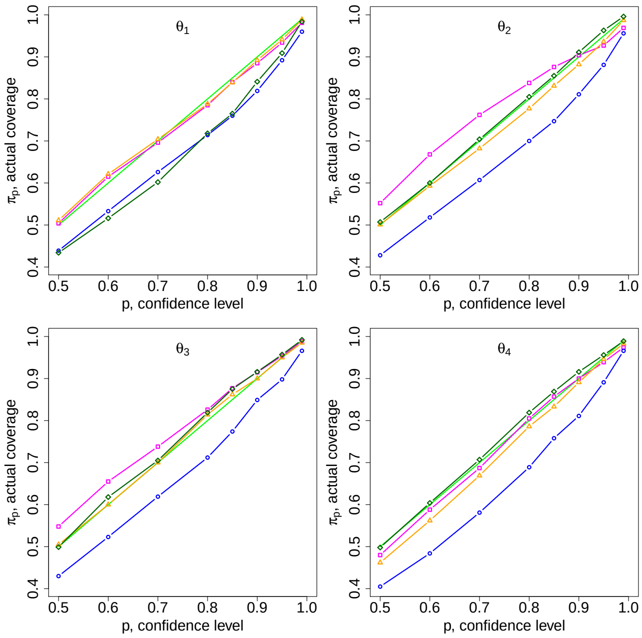

Figure 4: The Hill model. Purple squares stand for the linearization method, orange circles – F-test, blue triangles – LR-test, dark green diamonds – bootstrap. Green color indicates the asymptote.

C.2 Estimating actual coverage of pointwise confidence bands for the Hill model.

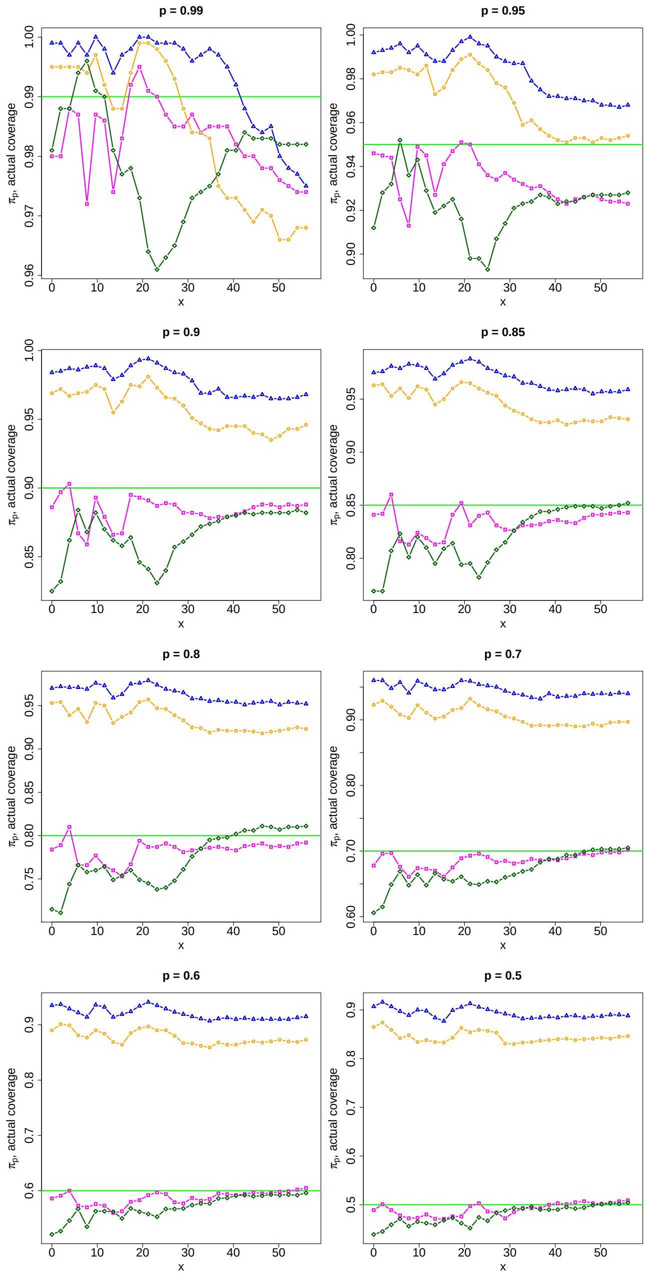

Figure 5: The Hill model. Purple squares stand for the linearization method, orange circles – F-test, blue triangles – LR-test, dark green diamonds – bootstrap. Green color indicates the asymptote.

C.3 Altering experimental conditions: accuracy of 90% confidence intervals for the Hill model.

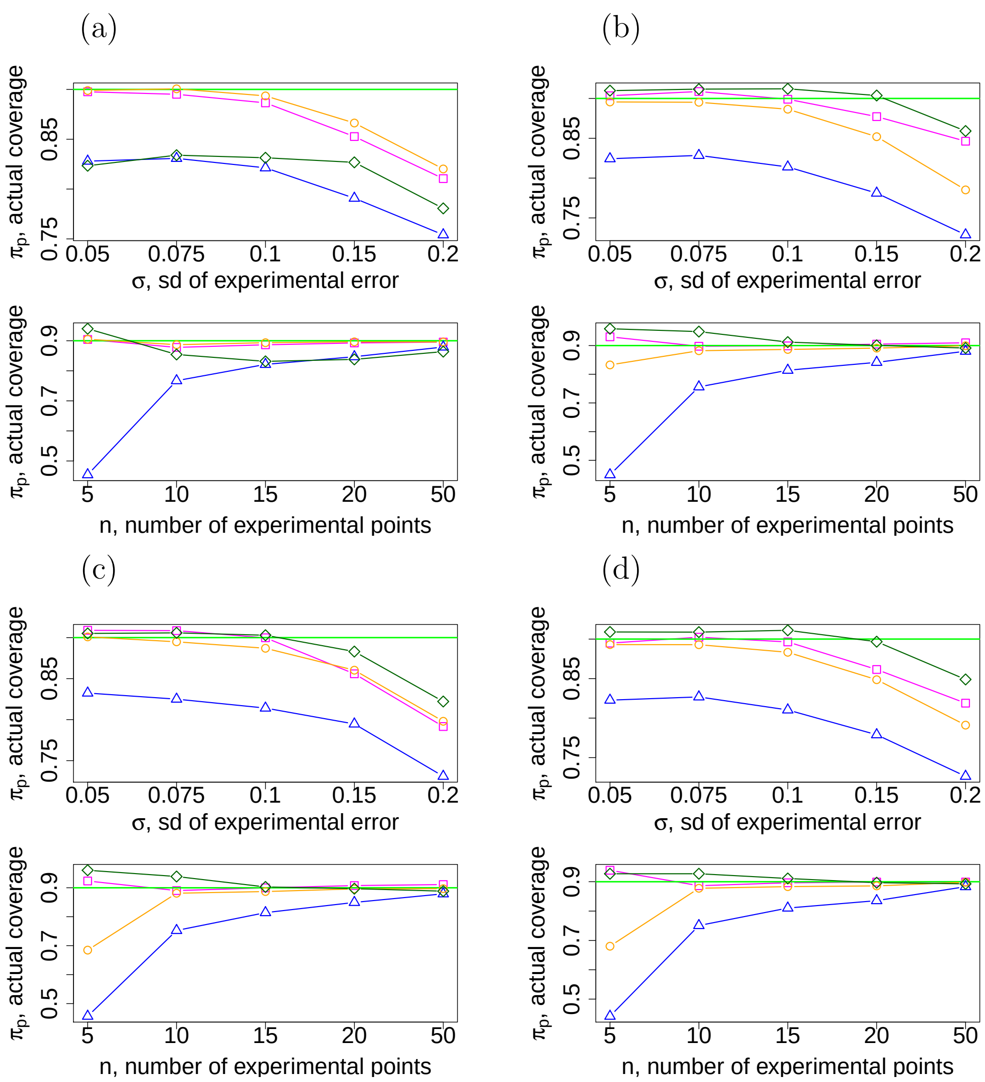

Figure 6: Altering experimental conditions: (a) accuracy of confidence intervals for θ_1, (b) accuracy of confidence intervals for θ_2, (c) accuracy of confidence intervals for θ_3, (d) accuracy of confidence intervals for θ_4. Purple color corresponds to the linearization method, orange – F-test, blue – LR-test, dark green – bootstrap. Green color indicates the asymptote.

4 Appendix D

$D.1$ Estimating actual coverage of confidence intervals for the Gomperz model.

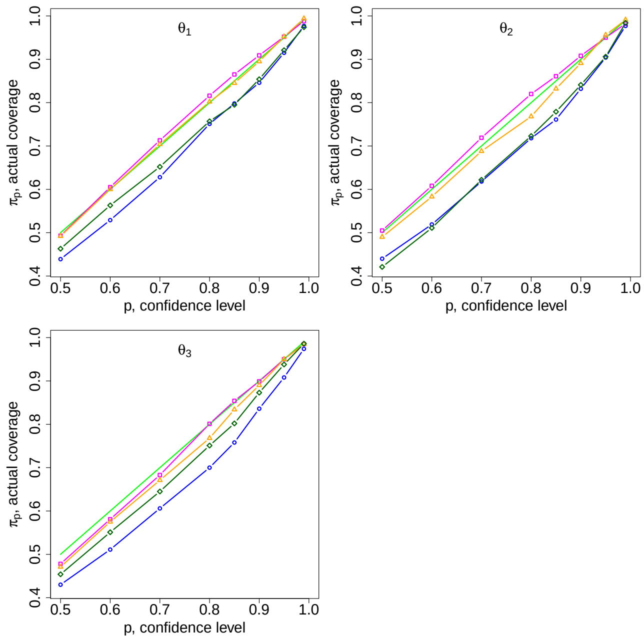

Figure 7: The Gomperz model. Purple squares stand for the linearization method, orange circles – F-test, blue triangles – LR-test, dark green diamonds – bootstrap. Green color indicates the asymptote.

D.2 Estimating actual coverage of pointwise confidence bands for the Gomperz model.

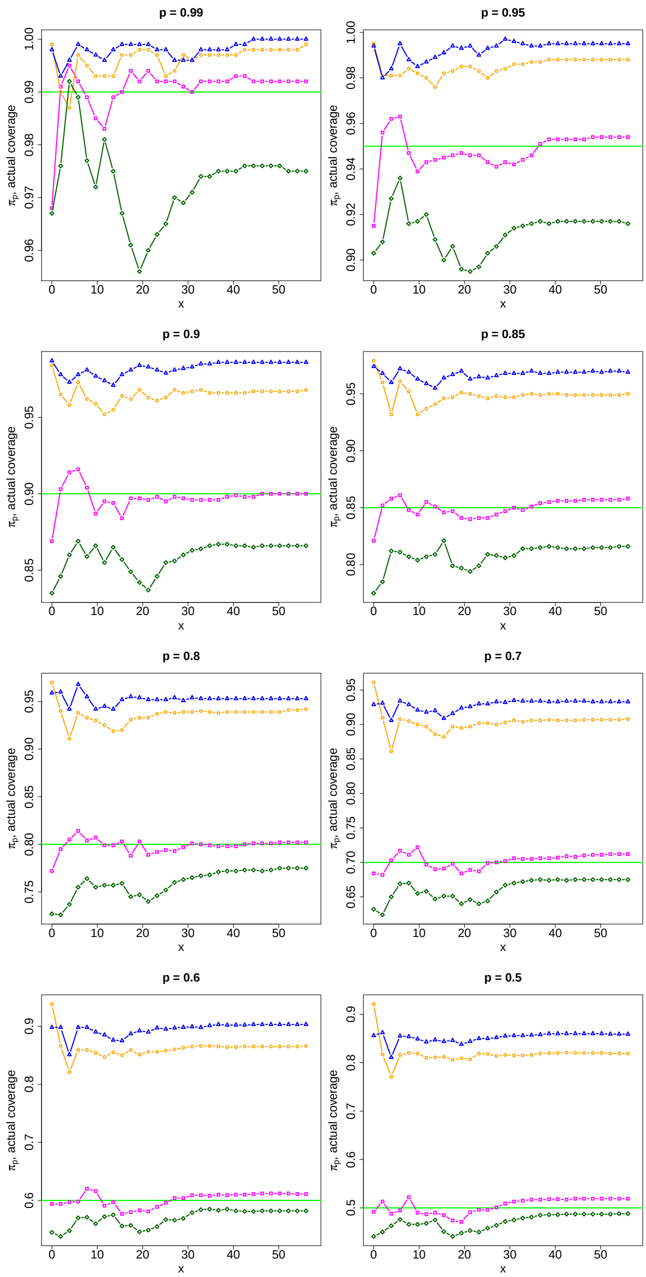

Figure 8: The Gomperz model. Purple squares stand for the linearization method, orange circles – F-test, blue triangles – LR-test, dark green diamonds – bootstrap. Green color indicates the asymptote.

D.3 Altering experimental conditions: accuracy of 90% confidence intervals for the Gompertz model.

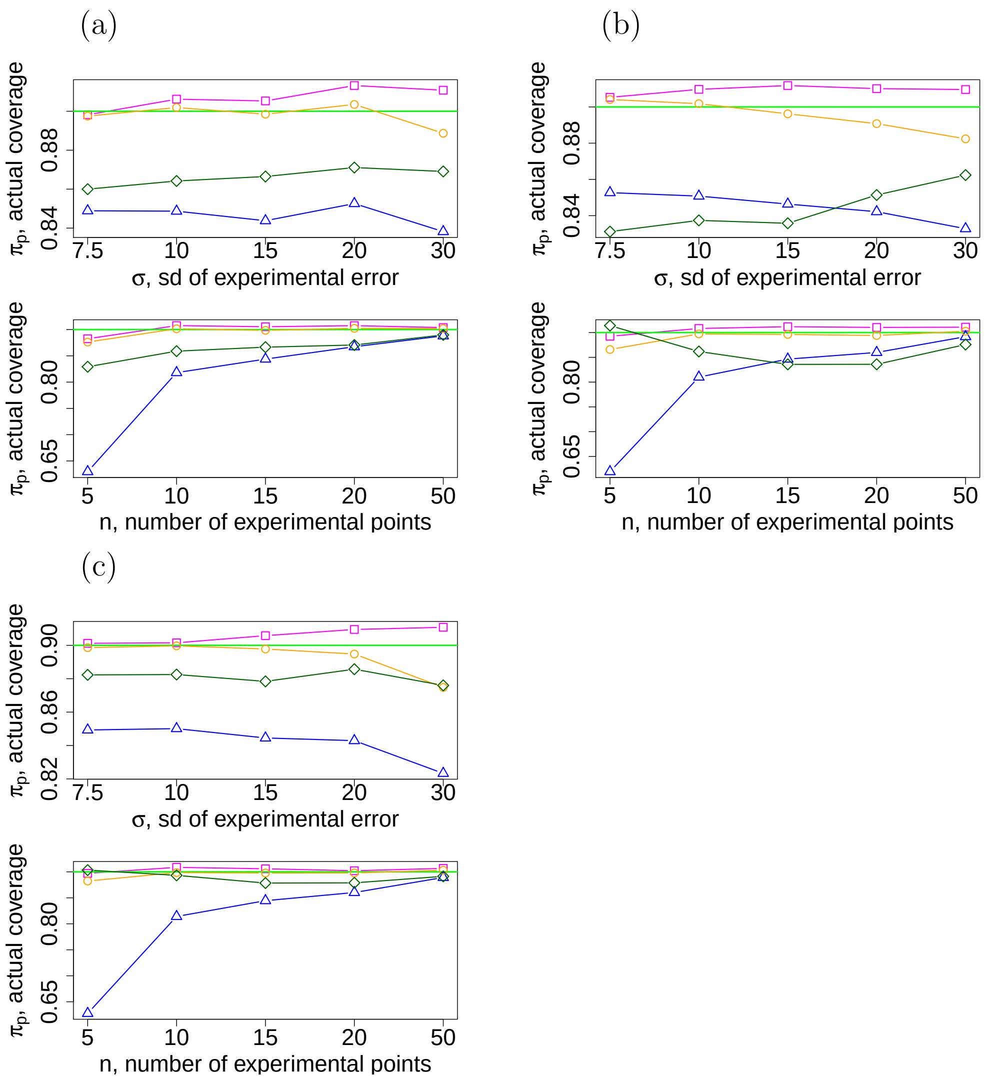

Figure 9: Altering experimental conditions: (a) accuracy of confidence intervals for θ_1, (b) accuracy of confidence intervals for θ_2, (c) accuracy of confidence intervals for θ_3. Purple color corresponds to the linearization method, orange – F-test, blue – LR-test, dark green – bootstrap. Green color indicates the asymptote.

5 Appendix E

Here we present the π_p estimation results for confidence regions of the the Gomperz problem with 100000 15-point pseudoexperimental datasets.

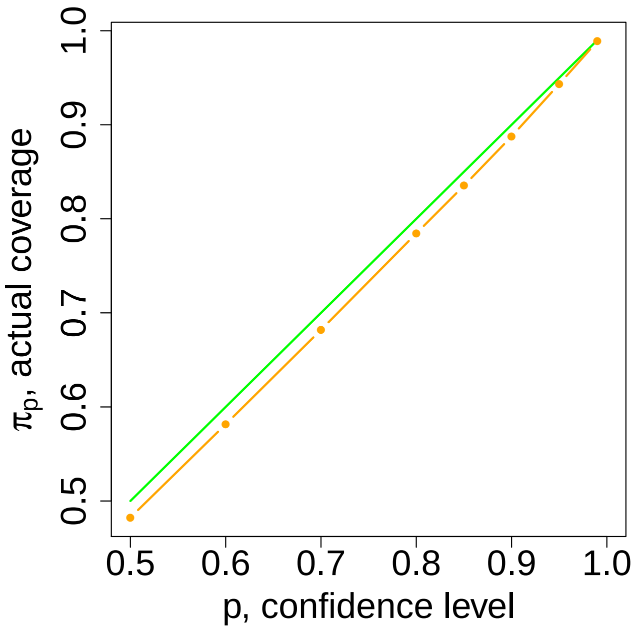

Figure 10: The Gomperz model, F-test confidence region. Green line indicates the asymptote.