Robust inference for generalized linear models by flipscores approach

Riccardo De Santis

University of Padova

13 October 2023

Co-Authors:
Jelle J. Goeman - Jesse Hemerik - Livio Finos

Introduction

A linear regression model is defined as

$$Y = X\beta + Z\gamma + \varepsilon$$

where

- Y is the outcome
- X, Z are observed covariates
- β, γ are regression coefficients, β is of direct interest
- $\varepsilon \sim N(0, \sigma^2 I)$ is an error term
- the variance σ^2 is assumed common among all the units
- Consider $\dim(\beta) = 1$ and $\dim(\gamma) \geq 1$

Introduction: testing problem

We want to test $H_0 : \beta = 0$ against a one or two-sided alternative

- β is the parameter of interest
- γ, σ are nuisance parameters, not of direct interest but we have to estimate them

What happens if we ignore some existing heteroscedasticity?

Introduction: example

Some simulated data with

- $y_i \sim N(\mu_i, \sigma_i^2)$
- $\mu_i = x_i\beta + z_i\gamma$
- $\text{cor}(x_i, z_i) = 0.5$
- $\beta = 0$
- $\gamma = 1$
- $\sigma_i^2 = 4x_i^2$.

We fit a linear model **assuming common variance**, testing $H_0 : \beta = 0$ with significance level $\alpha = 0.05$

Introduction: simulation

Sample size	Proportion of rejection
25	0.20
50	0.21
100	0.21
200	0.21
500	0.22
1000	0.21

Much higher than 0.05, we reject too often (no type I error control!)

Generalized linear models

Generalized linear models are a flexible tool introduced to extend linear regression models. Some examples are

- normal regression with logarithmic link
- poisson regression
- logistic regression

Usually we have a strong assumption on the variance structure (e.g. homoscedastic Normal model, Poisson model, ...).

Further problem: when the model variance is not constant it is difficult to check the validity of the assumptions

Hypothesis testing

The regression model is

$$g(\mu_i) = \eta_i = x_i\beta + z_i\gamma.$$

Aim: we consider univariate test of the form

$$H_0 : \beta = 0$$

against a one or two-sided alternative.

We want to build a test robust against variance misspecification

Sign-flip tests

Sign-flip tests offer an alternative way to do hypothesis testing. Usually they

- require less assumptions (semi-parametric tests)
- converge to the parametric counterpart (when it exists)
- have exact control of type I error

Sign-flips

What are sign-flips?

Suppose we have a sample of n observations. Sign-flips are n-dimensional vectors of 1 and -1. Example, $n = 6$:

$$
\begin{aligned}
I = F_1 &= (1, 1, 1, 1, 1, 1) \\
F_2 &= (1, 1, 1, 1, 1, -1) \\
&\vdots \\
F_{64} &= (-1, -1, -1, -1, -1, -1)
\end{aligned}
$$

In general the total amount is 2^n different flips.

Sign-flip tests

The idea is to use a conditional (flipping) distribution of the data. What does it mean?

- Let $T(I)$ be any observed test statistic.
- Call $T(F)$ a flipped test statistic. It is obtained by multiplying the data (or other appropriate quantities) by a sign-flip F.
- We have a flipping distribution with a total of 2^n test statistics.
- We can perform valid hypothesis testing if

$$T(I) \overset{d}{=} T(F) \text{ (equality in distribution)}$$

for all sign-flips, when H_0 is true

Sign-flip tests: example

- We observe a sample of independent observations $y_1, \ldots, y_n$, with $y_i \sim N(\mu, \sigma_i)$
- We test $H_0 : \mu = 0$ vs $H_1 : \mu > 0$, significance level of α
- When H_0 is true, $y_i \overset{d}{=} -y_i$ (equality in distribution).
- Use the test statistic $T(I) = \sum_{i=1}^{n} y_i$
- The flipped test statistic is $T(F) = \sum_{i=1}^{n} f_i y_i$
- We have $T(I) \overset{d}{=} T(F)$

Sign-flip tests: example

- Call $G = 2^n$ the total amount of transformations.
- We order them
$$T^{(1)}(F) \leq \cdots \leq T^{(G)}(F)$$
- We reject H_0 if $T(I) > T^{(\lceil (1-\alpha) \cdot G \rceil)}(F)$ (which is the $1 - \alpha$ quantile of the flipping distribution)
- Exact control of type I error

Sign-flip tests: example

We observe a sample of 6 observations ($2^6 = 64$)

$$Y = (-1.63, 1.61, 0.13, 0.66, 0.01, -0.65)$$

$$\begin{aligned}
T(I) &= -1.63 + 1.61 + 0.13 + 0.66 + 0.01 - 0.65 = 0.13 \\
T(F_2) &= 1.63 + 1.61 + 0.13 + 0.66 + 0.01 - 0.65 = 3.39 \\
T(F_3) &= 1.63 - 1.61 + 0.13 + 0.66 + 0.01 - 0.65 = 0.17 \\
&\dots \\
T(F_{64}) &= 1.63 - 1.61 - 0.13 - 0.66 - 0.01 + 0.65 = -0.13
\end{aligned}$$

We reject H_0 if $T(I) > T^{(61)}(F)$

Sign-flip test for GLMs

How to apply the idea of sign-flip tests for GLMs?

- The outcome Y cannot be used. Under $H_0 : \beta = 0$

$$\mu_i = g^{-1}\left(0 + z_i\gamma\right)$$

- In general

$$\mu_i \neq -\mu_i$$

$$\implies y_i \overset{d}{\neq} -y_i$$

$$\implies T(I) \overset{d}{\neq} T(F)$$

Effective score

New proposal: use the effective score. It is defined as

$$T(F) = n^{-1/2} F(S_\beta - \mathcal{I}_{\beta\gamma}\mathcal{I}_{\gamma\gamma}^{-1} S_\gamma) = n^{-1/2} \sum_{i=1}^{n} f_i \nu_i^* \Big|_{\beta=0,\gamma=\hat{\gamma}}$$

We have

$$\mathbb{E}[T(I)] = \mathbb{E}[T(F)] = 0, \quad \mathbb{V}[T(F)] \xrightarrow{n \to \infty} \mathbb{V}[T(I)]$$

Standardized score

Easy solution to improve the convergence of the test statistic.
Standardized test statistic

$$T_s(F) = T(F)/\mathbb{V}(T(F))^{1/2}.$$

We have

$$\mathbb{E}[T_s(I)] = \mathbb{E}[T_s(F)] = 0, \quad \mathbb{V}[T_s(I)] = \mathbb{V}[T_s(F)] = 1$$

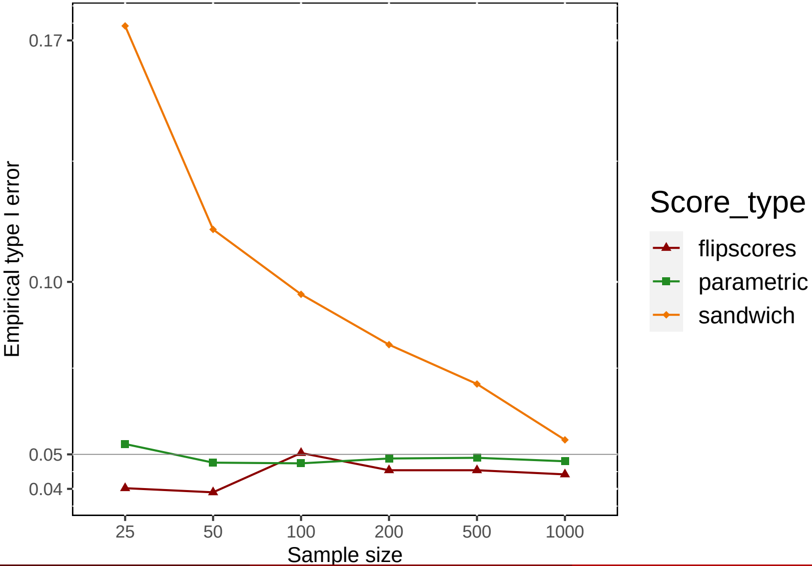
Simulation: Poisson model

Example we simulate n observations from the model $y_i \sim Poisson(\mu_i)$

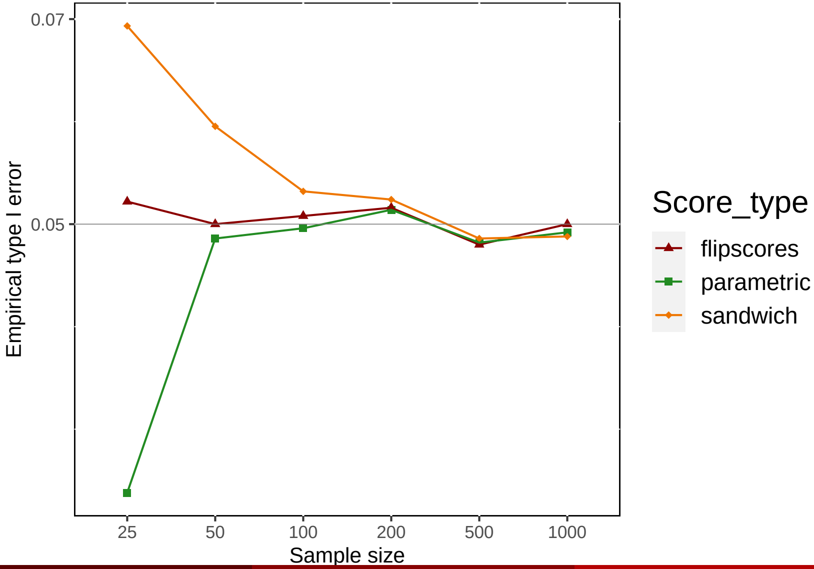
- $\log(\mu_i) = x_i\beta + z_i\gamma$
- $(\beta, \gamma) = (0, 1, 1, 1)$
- $cor(x_i, z_i) = (0.5, 0.1, 0.1)$

We have two competitors: the **standard parametric test** and the **sandwich estimator**. See some simulations!

Simulation: Poisson model



Simulation: Logit model

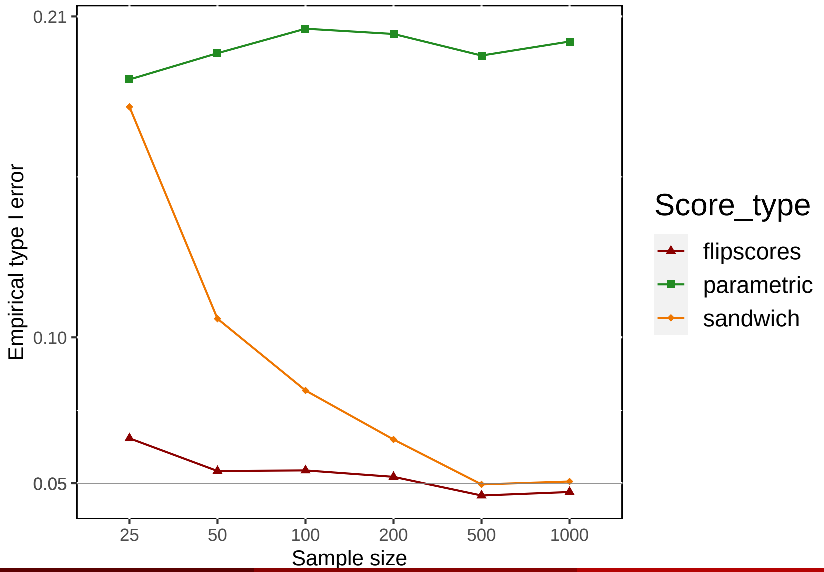


Simulation: Ignored Heteroscedastic Normal

Linear model with ignored heteroscedasticity. Some simulated data with

- $\beta = 0$
- $\gamma = (1, 1, 1)$
- $\sigma_i^2 = 4x_i^2$
- $cor(x_i, z_i) = (0.5, 0.1, 0.1)$

Ignored Heteroscedastic Normal



Simulation: True Negative binomial, fitted Poisson

Poisson distribution

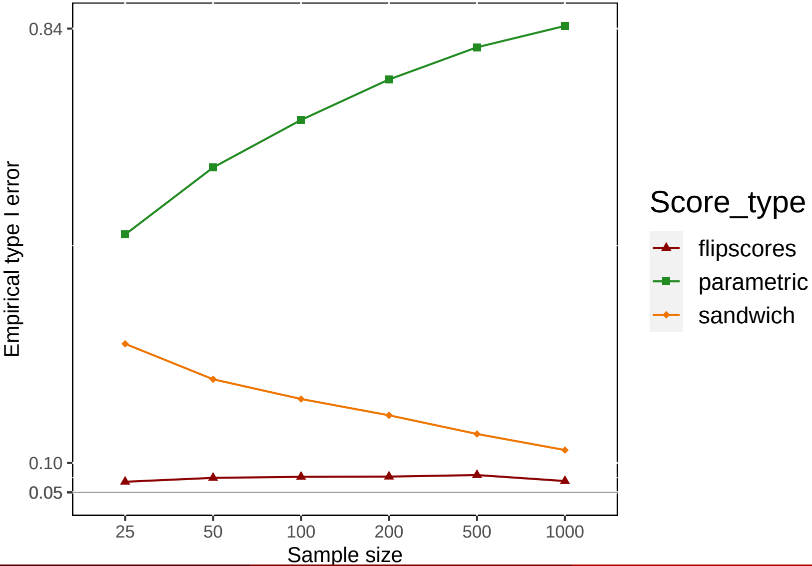
- $\mathcal{Y} = \mathbb{N}$
- $\mathbb{E}[y_i] = \mu_i$
- $\mathbb{V}[y_i] = \mu_i$

Negative binomial distribution

- $\mathcal{Y} = \mathbb{N}$
- $\mathbb{E}[y_i] = \mu_i$
- $\mathbb{V}[y_i] = \mu_i (1 + \phi \mu_i)$

We fit a Poisson regression model while the true distribution is Negative binomial. We set $\phi = 1$

True Negative binomial, fitted Poisson

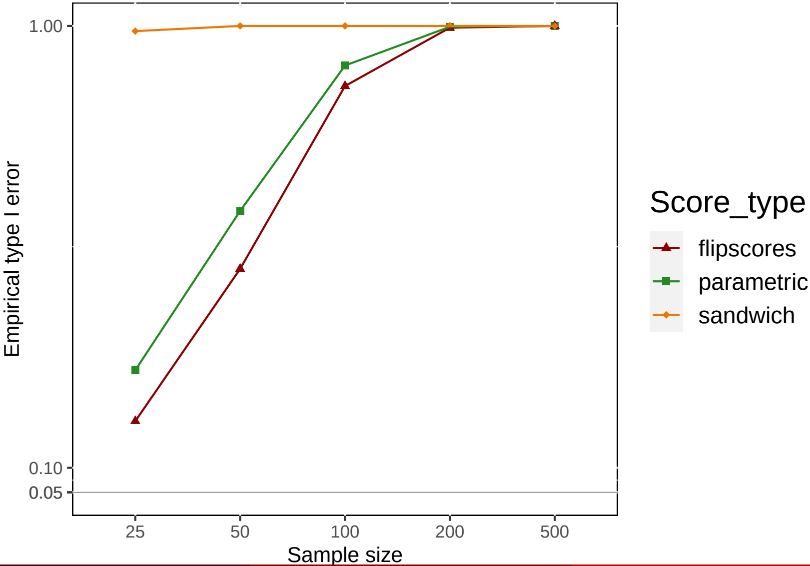


Simulation: power comparison

If the test is able to control the type I error, it should have good power

For a meaningful comparison we fit a correctly specified Poisson model setting $\beta = 0.3$, while we test $H_0 : \beta = 0$

Power comparison



Code example

```
##install.packages("flipscores")
library(flipscores)
set.seed(1)
X<-rnorm(25)
Z<-rnorm(25)
Y<-rpois(25,abs(Z))
mod<-flipscores(Y~X+Z,family="poisson"(link="log"),
                   score_type = "standardized")
summary(mod)
```

With $n = 1000$ and 4 covariates the computational time is ≈ 4.2 seconds

Code example

```
> summary(mod)

Call:
flipscores(formula = Y ~ X + Z, family = poisson(link = "log"),
    score_type = "standardized")

Deviance Residuals:
    Min       1Q   Median       3Q      Max
-1.0842  -0.9145  -0.7003   0.2292   2.3617

Coefficients:
            Estimate   Score Std. Error z value eff_size Pr(>|z|)
(Intercept)  -1.0072 -2.5093     0.6527 -3.8442   -0.630   0.0012 **
X             0.7292  1.3630     0.9683  1.4076    0.389   0.2058
Z            -1.0395 -1.1693     0.5107 -2.2896   -0.465   0.0372 *
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

References I

R. De Santis, J.J. Goeman, J. Hemerik, and L. Finos.
Inference in generalized linear models with robustness to misspecified variances.
https://arxiv.org/abs/2209.13918, 2022.

J. Hemerik and J.J. Goeman.
Exact testing with random permutations.
TEST, 27:811–825, 2018.

J. Hemerik, J.J. Goeman, and L. Finos.
Robust testing in generalized linear models by sign flipping score contributions.
Journal of The Royal Statistical Society Series B-statistical Methodology, 82(3):841–864, 2020.