Inference in generalized linear models with robustness to misspecified variances

Riccardo De Santis, Jelle J. Goeman, Jesse Hemerik, Livio Finos

Statistical methods and models for complex data — Poster Session: 21 September 2022

Intro: testing problem

Consider a Poisson regression model

$$g(\mu) = X\beta + Z\gamma$$

while the true distribution is negative binomial with dispersion parameter $\phi = 1$. That is:

- we correctly specify $\mathbb{E}[y_i] = \mu_i$
- we assume $\mathbb{V}[y_i] = \mu_i$ but the truth is $\mathbb{V}[y_i] = \mu_i + \mu_i^2$

Testing problem:

- Take the true $\beta = 0$
- do some simulations fitting a Poisson model (ignoring the overdispersion)
- do the test

$$H_0 : \beta = 0$$

for $\alpha = 0.05$ with the `glm` function in `R`.

Consequences?

Sample size	Proportion of rejections
25	0.489
50	0.603
100	0.684
200	0.753
500	0.808
1000	0.845

Table 1. Type I error control of the Wald test

Figure 1 shows the proportion of rejections of the test. **No type I error control!** (Remember: $\alpha = 0.05$)

Variance specification in GLMs

Generalized linear models (GLM) are a widely used tool in statistical inference. The model specification is divided in several steps, which include the assumption on the **variance structure** ($\Longleftrightarrow$ assumption on the **dispersion parameters**). Generally, it is quite restricting:

- **Normal linear regression** ($\mathbb{V}[y_i] = \sigma^2 \; \forall i$). Homoscedasticity is assumed. Some extensions are possible but we have always to make some assumptions on the structure
- **Poisson regression**. ($\mathbb{V}[y_i] = \mu_i$). Variance is assumed equal to the mean
- **Overdispersed Poisson model** ($\mathbb{V}[y_i] = \phi\mu_i$). Quasi-likelihood approach allows for a common additional dispersion parameter. One single parameter can still be too limiting

Further problem: when the model variance is not constant is difficult to check the assumptions made.

Parametric solution

As shown above, the standard parametric approach (Wald test, likelihood ratio test, score test) fails in the control of type I error. The parametric solution consists in the use of the **sandwich estimator** of the variance.

Sample size	Proportion of rejections
25	0.303
50	0.243
100	0.209
200	0.181
500	0.149
1000	0.122

Table 2. Type I error control of the Sandwich estimator

Figure 2 shows the **(slow)** convergence to the nominal level. It is quite unsatisfactory!

Group-invariant tests

We look for a different approach to perform statistical testing. How?

The idea is to define **null-invariant (appropriate) transformations** of the data:

- Define with $\mathcal{F}$ a **group of transformations**. Examples are permutations, rotations, sign flips (multiply part of the data by -1)
- **Null invariance**. Let $T(Y)$ be any statistical test. We can perform a valid test if
$$T(Y) \overset{d}{=} T(FY) \; \forall F \in \mathcal{F}$$
under H_0
- **Observed statistic**. It is identified by the identity element $F = I$

Why this approach? Usually these tests require less assumptions (non-parametric or semi-parametric tests)

How to adapt this idea to GLMs?

Sign-flip test for GLMs

Take the **effective score** as test statistic. It is the score of the profile likelihood (in our case for β), asymptotically independent from the estimation of the nuisance parameters. Formula:

$$T(F) = n^{-1/2} X^T W^{1/2} (I - H) W^{-1/2} F (y - \hat{\mu}) = \sum_i f_i \nu_i^*$$

where

- $H = W^{1/2} Z (Z^T W Z)^{-1} Z^T W^{1/2}$ is the projection matrix for GLMs
- W is the diagonal matrix with the weights
- $\hat{\mu}$ are the fitted value under the null model (i.e. with $\beta_0, \hat{\gamma}$)
- F is a sign-flip matrix, i.e. a diagonal matrix with elements -1 or 1.

Properties?

- $\mathbb{E}[T(F)] = 0, \; \forall F \in \mathcal{F}$
- $\mathbb{V}[T(F)] \xrightarrow{n \to \infty} \mathbb{V}[T(I)], \; \forall F \in \mathcal{F}$
- $\mathbb{V}[T(I)] > \mathbb{V}[T(F)]$ for finite sample size
- the test is asymptotically exact but anti-conservative for finite sample size

Need for a further improvement!

Sign-flip standardized score test

Next step: **standardization** of the test statistic:

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

Properties?

- $\mathbb{E}[T_s(F)] = 0, \; \forall F \in \mathcal{F}$
- $\mathbb{V}[T_s(F)] = 1, \; \forall F \in \mathcal{F}$
- the test is exact for the normal model, asymptotically otherwise, but the convergence involves only the third and further moments.

Robustness to variance misspecification

Main property: proven robustness to misspecified variance under minimal assumptions. That is, if you misspecify the variance (i.e. the dispersion parameters)

- $\mathbb{E}[T_s(F)] = 0, \; \forall F \in \mathcal{F}$
- $\mathbb{V}[T_s(F)] \xrightarrow{n \to \infty} \mathbb{V}[T_s(I)], \; \forall F \in \mathcal{F}$
- the test is asymptotically exact

See some simulations!

Simulation study

Simulation setting

We test a univariate hypothesis $H_0 : \beta = 0$ against a two-sided alternative. The model contains the true $\beta = 0$ and three nuisance parameters $\gamma = (1, 1, 1)$. The correlation between X and the nuisance Z is $\text{Cor}(X, Z) = (0.5, 0.1, 0.1)$. The nominal level of the test is $\alpha = 0.05$.

Misspecified model

Again, true model is Negative Binomial, we fit a Poisson. The results are in Figure 1

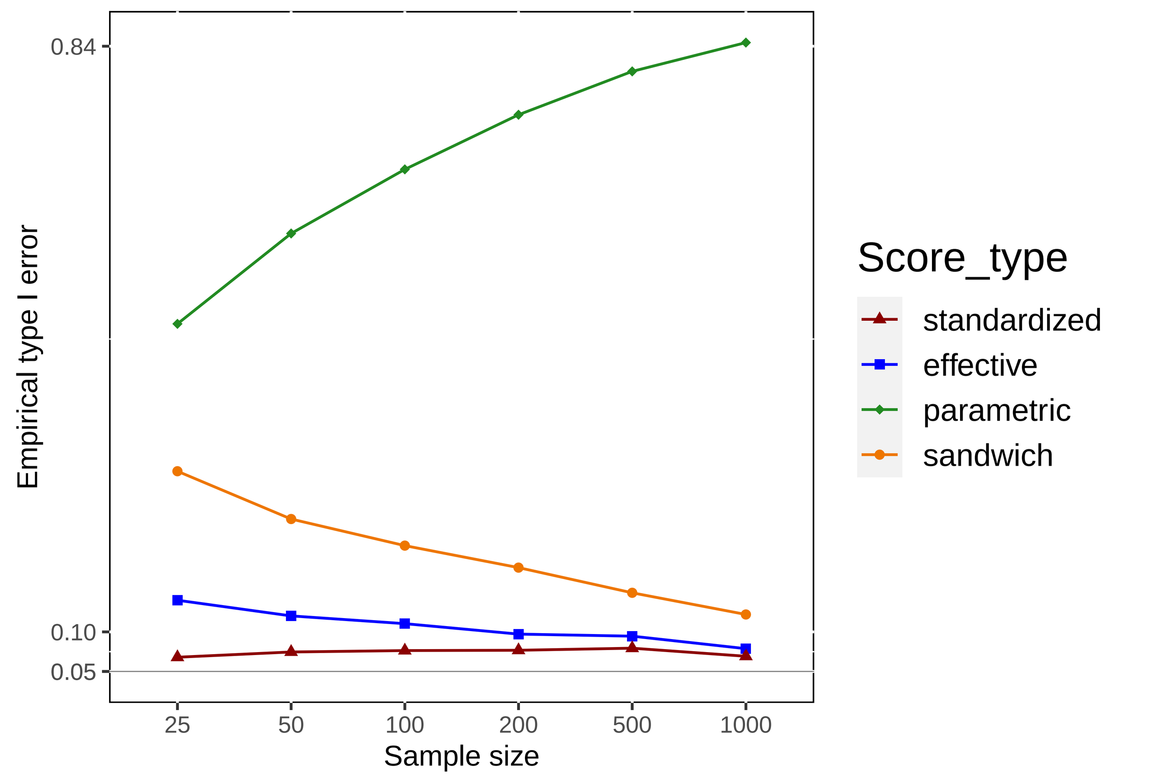

Figure 1. Comparison of type I error control for a misspecified Poisson model

Well-specified model

What happens in a correctly specified model? See in Figure 2 the simulation for a Poisson model

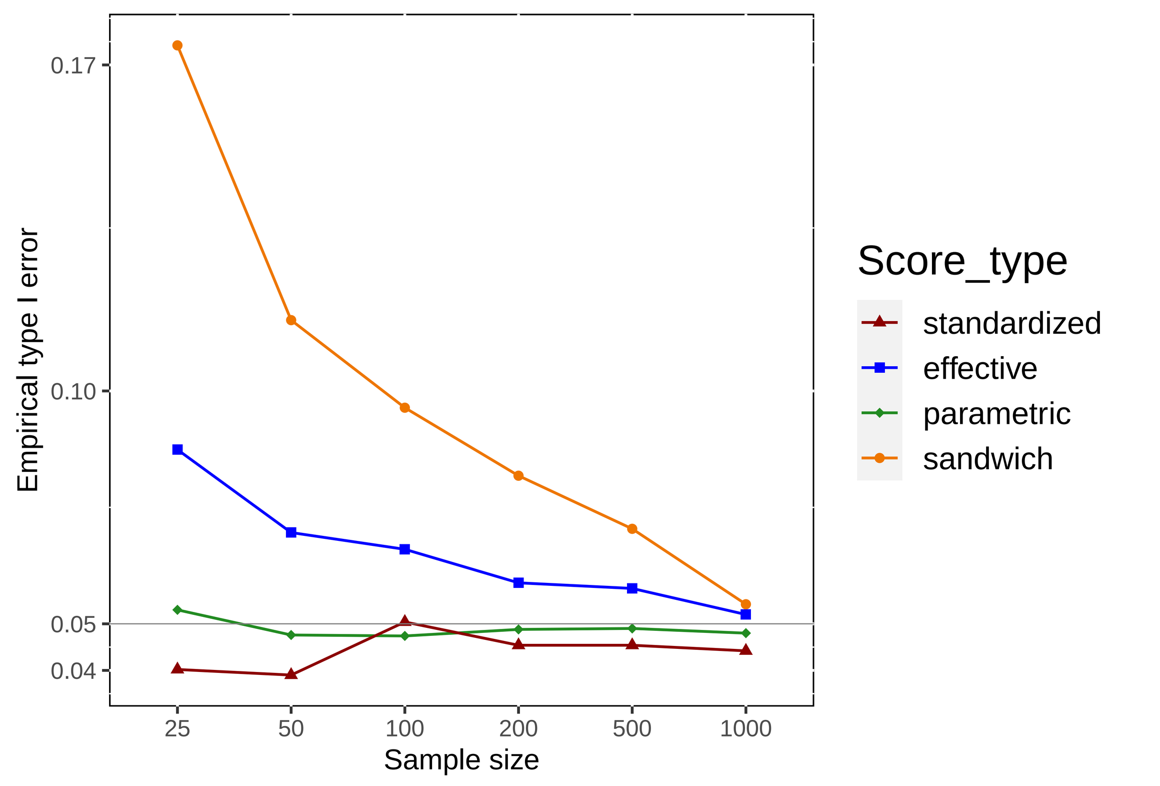

Figure 2. Comparison of type I error control for a well-specified Poisson model

Comments

When the model is **misspecified**, the parametric test is unreliable, the sandwich has a slow convergence while our standardized score test is the closest to the nominal level.

When the model is **well-specified**, our standardized score test is as reliable as the parametric test. The sandwich has still a slow convergence.

References

Hemerik, J., Goeman, J.J. (2018). Exact testing with random permutations. *TEST*, **27**, 811 – 825.

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

Contact information

Riccardo De Santis, Ph.D. student
Department of Statistical Sciences, University of Padova
riccardo.desantis.1@phd.unipd.it