

# COMPARISON OF APPROACHES TO TESTING EQUALITY OF EXPECTATIONS AMONG SAMPLES FROM POISSON AND NEGATIVE BINOMIAL DISTRIBUTION

Martin Tejkal<sup>1</sup>, Zuzana Hübnerová<sup>2</sup>

<sup>1</sup>Department of Econometrics, Faculty of Military Leadership, University of Defence in Brno, Kounicova 65, Brno, Czech Republic

<sup>2</sup>Department of Mathematics, Faculty of Mechanical Engineering, Brno University of Technology, Technická 2896/2, Brno, Czech Republic

## Abstract

TEJKAL MARTIN, HÜBNEROVÁ ZUZANA. 2018. Comparison of Approaches to Testing Equality of Expectations Among Samples from Poisson and Negative Binomial Distribution. *Acta Universitatis Agriculturae et Silviculturae Mendelianae Brunensis*, 66(4): 1025–1034.

The paper deals with testing of the hypothesis of equality of expectations among  $p$  samples from Poisson or negative binomial distribution. a comparison of two main approaches is carried out. The first approach is based on transforming the samples from either Poisson or negative binomial distribution in order to achieve normality or variance stability, and then testing the hypothesis of equality of expectations via the  $F$ -test. In the second approach, test statistics coming from the theory of maximum likelihood appearing in generalised linear models framework, specially designed for testing the hypothesis among samples from the respective distributions (Poisson or negative binomial), are used. The comparison is done graphically, by plotting the simulated power functions of the test of the hypothesis of equality of expectations, when first or second approach was used. Additionally, the relationship between the power functions obtained via the respective approaches and sample sizes is studied by evaluating the respective power functions as functions of a sample size numerically.

Keywords: Poisson distribution, negative binomial distribution, ANOVA,  $F$ -test statistic, generalized linear model, likelihood ratio, score statistic, variance stabilizing transformation, Yeo-Johnson transformation, power function

## INTRODUCTION

In applications, we often meet data involving counts. For modelling of such data Poisson distribution, or in case of heteroscedastic data, negative binomial distribution is often used. If we want to test the hypothesis of equality of expectations among a given number of independent samples of such data, we cannot use the classical analysis of variance framework, because the assumptions of the normality and variance stability of the data are not met. Hence a different strategy has to be developed.

One of the possible approaches is to transform the samples in order to meet the assumptions

of classical analysis of variance (see Scheffé, 1999). In the case of Poisson or negative binomial distribution, a logarithmic transformation is often suggested (see Moss and McPhee, 2006). The problem of zero observations is solved by adding 1 or another positive constant into the argument of the logarithm. Another possible solution is to use variance stabilising transformation (see Anděl, 2011; Tejkal, 2017), or some other transformation that assures, that the transformed random variable is approximately normally distributed. Such a transformation might be for

example of the Yeo-Johnson family (see Yeo and Johnson, 2000).

Another possible approach is to use wholly different test statistic, specially designed for testing the hypothesis among samples from a given distribution. Such test statistics come from the theory of maximum likelihood, and they appear in the framework of the generalised linear models (GLM).

The goal of this paper is to provide the reader with a comparison of the above mentioned approaches. This is done by comparing the power functions of the test statistics, when testing the hypothesis of equality of expectations among  $p$  samples of the same size from the Poisson and the negative binomial distribution. Additionally, the relationship between power functions and sample size is studied, and the comparison of power functions for different sample sizes is provided. The results of this comparison may be straightforwardly applied to experimental design to determine sample size.

## MATERIALS AND METHODS

### Assumed models

Let  $\mathbf{Y}_i = (Y_{i1}, \dots, Y_{in})^T$  for  $i = 1, \dots, p$  be the random samples of a size  $n$  from either Poisson distribution  $Po(\theta_i)$  with expectation parameter  $\theta_i$ , or negative binomial distribution  $NBi(\theta_i, \kappa)$  with expectation parameter  $\theta_i$  and shape parameter  $\kappa$ . For the sake of convenience, the parameter of the Poisson distribution and the first parameter of the negative binomial distribution are denoted by the same symbol. Furthermore, assume that for both Poisson and negative binomial case we have  $EY_{ij} = \theta_i$  for all  $i = 1, \dots, p$  and  $j = 1, \dots, n$ . For completeness, we present the probability density function of the negative binomial distribution under this parametrisation:

$$p(x; \theta, \kappa) = \begin{cases} \frac{\Gamma(x + \kappa) \left( \frac{\theta}{\kappa + \theta} \right)^x \left( \frac{\kappa}{\kappa + \theta} \right)^\kappa}{x! \Gamma(\kappa)} & \text{for } \forall x \in \mathbb{N}_0, \\ 0 & \text{otherwise.} \end{cases} \quad (1)$$

For details about this parametrisation of negative binomial distribution see Tejkal (2017).

Assume that the  $p$  random samples are mutually independent. We want to test the hypothesis

$$H_0 : \theta_1 = \dots = \theta_p \quad (2)$$

of equality of expectations among the  $p$  samples against the alternative

$$H_1 : \exists i, k \in \{1, \dots, p\}, i \neq k, \text{ such that } \theta_i \neq \theta_k \quad (3)$$

### Approach via Transformations

For the samples from Poisson distribution we consider the following transformations, the logarithmic transformation

$$Z = \ln(Y + 1), \quad (4)$$

the square root transformation

$$Z = \sqrt{Y + k}, \quad (5)$$

where the value of the constant  $k$  is chosen to be optimal according to Anscombe (1948), i.e.  $k = \frac{3}{8}$ . Finally, we consider the so called Yeo-Johnson transformation given by

$$Z = \begin{cases} \frac{(1+X)^\lambda - 1}{\lambda} & \text{for } X \geq 0, \quad \lambda \neq 0 \\ \ln(1+X) & \text{for } X \geq 0, \quad \lambda = 0 \\ -\frac{((1+X)^{2-\lambda} - 1)}{2-\lambda} & \text{for } X < 0, \quad \lambda \neq 2 \\ -\ln(1-X) & \text{for } X < 0, \quad \lambda = 2, \end{cases} \quad (6)$$

where the value of the parameter  $\lambda$  can be estimated by maximum likelihood estimation (see Yeo and Johnson, 2000).

For the samples from negative binomial distribution we consider again the logarithmic transformation (4), the Yeo-Johnson transformation (6), and additionally also the argument of hyperbolic sine transformation

$$Z = 2 \sinh^{-1} \left( \sqrt{\frac{Y+c}{\kappa+d}} \right), \quad (7)$$

where  $c = -\frac{1}{6} \left[ \sqrt{6\kappa(6\kappa^3 - 27\kappa^2 + 41\kappa - 21)} - 6\kappa^2 + 9\kappa \right] \left[ 2\kappa - 3 \right]$  and  $d = -2c$  are the optimal values of the constants  $c$  and  $\kappa$  according to Anscombe (1948).

The formulas (5) and (7) are generalised versions of variance stabilising transformations for the respective distributions (see Anděl, 2011). Full derivations of the optimal values of the constants can be found in Anscombe (1948) or Tejkal (2017).

It is assumed, that after applying one of the introduced transformations to the initial model, the classical one-way analysis of variance setting is obtained. I. e. for each  $i = 1, \dots, p$  is  $Z_i = (Z_{i1}, \dots, Z_{in})^T$  the random sample of a size  $n$  from  $N(\mu_i, \sigma^2)$  and, furthermore, the  $p$  random samples are mutually independent. The extent of how much this was satisfied in the respective cases of different transformations was checked by computing the estimation of the variance and the estimation of skewness of the transformed sample. The following estimators were used

$$s^2 = \frac{1}{n-1} \sum_{j=1}^n (Z_{ij} - \bar{Z}_i)^2, \quad (8)$$

for estimating variance, and

$$b_1 = \frac{\frac{1}{n} \sum_{j=1}^n (Z_{ij} - \bar{Z}_i)^3}{\left[ \frac{1}{n-1} \sum_{j=1}^n (Z_{ij} - \bar{Z}_i)^2 \right]^{\frac{3}{2}}}, \quad (9)$$

for estimating skewness, where  $\bar{Z}_i = \frac{1}{n} \sum_{j=1}^n Z_{ij}$  is the arithmetic mean of the  $i$ -th sample.

The  $F$ -test statistic used to test the hypothesis of the equality of expectations of the transformed samples is given by

$$F = \frac{p(n-1)}{p-1} \frac{\sum_{i=1}^p n(\bar{Z}_i - \bar{Z})^2}{\sum_{i=1}^p \sum_{j=1}^n (Z_{ij} - \bar{Z}_i)^2}, \quad (10)$$

where  $\bar{Z} = \frac{1}{np} \sum_{i=1}^p \sum_{j=1}^n Z_{ij}$ . In what follows, we will denote  $Q_{\mathcal{L},d}(p)$  a quantile of a distribution  $\mathcal{L}$  with  $d$  degrees of freedom. The  $F$  test is carried out by comparing the value of the test statistics with the respective quantile  $Q_{F_{p-1,p(n-1)}}(1-\alpha)$  for the selected level of significance  $\alpha$ .

### Approach via Tests from GLM Framework

In the case of Poisson distributed samples, the following two test statistics will be used to test the hypothesis of the equality of expectations. The likelihood ratio test statistic (see Hrdličková, 2002), which is given by

$$LR_{P_0} = 2 \sum_{i=1}^p \sum_{j=1}^n Y_{ij} \ln \left( \frac{\bar{Y}_i}{\bar{Y}} \right), \quad (11)$$

and the score test statistic (see Hrdličková, 2006), which is given by

$$Sc_{P_0} = n \sum_{i=1}^p \frac{(\bar{Y}_i - \bar{Y})^2}{\bar{Y}}, \quad (12)$$

where  $\bar{Y}_i = \frac{1}{n} \sum_{j=1}^n Y_{ij}$  and  $\bar{Y} = \frac{1}{np} \sum_{i=1}^p \sum_{j=1}^n Y_{ij}$ . In the case of negative binomially distributed samples, the following two test statistics will be used. The likelihood ratio test statistic (see Hübnerová and Doudová, 2011), that is given by

$$LR_{NBi} = 2 \sum_{i=1}^p n \left[ \bar{Y}_i \ln \left( \frac{\bar{Y}_i}{\bar{Y}} \right) - (\bar{Y}_i + \kappa) \ln \left( \frac{\bar{Y}_i + \kappa}{\bar{Y} + \kappa} \right) \right], \quad (13)$$

and the score test statistic, that for the negative binomial case (see Hübnerová and Doudová, 2011) is given by

$$Sc_{NBi} = \frac{\kappa}{\bar{Y}(\bar{Y} + \kappa)} \sum_{i=1}^p n(\bar{Y}_i - \bar{Y})^2. \quad (14)$$

For all the test statistics (11), (12), (13), and (14) the test is carried out by comparing the value of the test statistic with the respective quantile  $Q_{\chi^2_{p-1}}(1-\alpha)$  for the selected level of significance  $\alpha$ .

### Comparing Power Functions

Let  $\Omega$  be the parametric space, let  $\theta$  be an element of  $\Omega$ . Let us denote  $\beta(\theta)$  the conditional probability of rejecting the null hypothesis, given that the alternative, characterised by the value of parameter  $\theta$ , holds. The function  $\beta(\theta)$  with values  $\theta \in \Omega$  is called the power function of a test.

In this paper, the main focus will be on providing a comparison of power functions obtained via simulations. a more theoretical approach is developed in Tejkal (2017). Its use is, however, limited only for obtaining power functions of the  $F$ -test after either transformation (4) or (5) is applied in the Poisson case or transformation (4) or (7) is applied in the negative binomial case. Additionally, the  $p$  samples have to be of the same size. The method used for obtaining the results of the theoretical approach will be explained briefly in the following subsection. For details see Tejkal (2017).

The process of computing power functions for Poisson case by any of the approaches does not differ from the negative binomial case, therefore, when providing the description, it will not be differentiated between the two distributions. Recall that the expectation parameters of both distributions are denoted  $\theta$ .

### Theoretical Power Functions

The theoretical approach to obtaining the approximations of power functions of the  $F$ -test in cases described above is based on the following. The power function of the  $F$ -test can be approximated as described in the Proposition 1.

**Proposition 1.** The power of the  $F$ -test  $\beta_\alpha(\theta)$  on the level of significance  $\alpha$  may be written as follows

$$\beta_\alpha(\theta) = 1 - \mathcal{F}_{F_{p-1,n(p-1),\delta}}(Q_{F_{p-1,n(p-1)}}(1-\alpha)), \quad (15)$$

where  $\mathcal{F}_{F_{p-1,n(p-1),\delta}}$  is the distribution function of noncentral  $F$  distribution with degrees of freedom  $d_1 = p-1$ ,  $d_2 = p(n-1)$  and a noncentrality parameter  $\delta = \frac{n}{\sigma^2} \sum_{i=1}^p (\theta_i - \bar{\theta})^2$ , where  $\bar{\theta} = \frac{1}{p} \sum_{i=1}^p \theta_i$ .

*Proof.* Straightforward use of properties of conditional probability.  $\square$

Furthermore, it is assumed, that the expectation parameters of the distribution of the transformed random samples are known. In practical computations approximations of the numerical characteristics are used.

With these assumptions, the value of the power function of the  $F$ -test given by formula (15) can be computed at any point  $\theta$ . Approximations of the power functions of the tests in GLM can be found in Hrdličková (2008).

### Simulated Power Functions

The simulated power functions are computed in the following way. An initial value of expectation parameter  $\theta_1$  are chosen. Additionally, the value of the parameter  $\kappa$  is selected for the negative binomial case. The significance level  $\alpha = 0.05$  is considered hereinafter.

The  $p$  random samples  $\mathbf{Y}_i$  for  $i = 1, \dots, p$  of a size  $n$  from either Poisson or negative binomial distribution are generated. The values  $\theta_2, \dots, \theta_p$  are obtained from  $\theta_1$  by adding and subtracting multiples of a number  $h_i$  according to the formula

$$\theta_i = \begin{cases} \theta_1 + \frac{i}{2}h_j, & \text{for } i = 2, 4, \dots, p, \\ \theta_1 - \frac{i-1}{2}h_j, & \text{for } i = 3, 5, \dots, p, \end{cases} \quad (16)$$

where  $j$  is the number of the run of the algorithm. Denote  $s$  a fixed step, and set  $h_1 = s$  for  $j = 1$ . For  $j > 1$  the value of  $h_j$  is given by the following formula

$$h_{j+1} = h_j + s \quad (17)$$

that characterises  $\theta$  in  $\beta(\theta)$ . When using the approach via transformations, the random samples are transformed using (4), (5), and (6) for the Poisson case and (4), (7), and (6) for the negative binomial case, obtaining transformed random samples  $\mathbf{Z}_i$  for  $i = 1, \dots, p$ . The optimal value of the parameter  $\lambda$  of transformation (6) is determined via maximum likelihood method (see Yeo and Johnson, 2000). The  $F$  statistic is computed using formula (10). The transformed samples  $\mathbf{Z}_i = (Z_{i1}, \dots, Z_{in})^T$  for  $i = 1, \dots, p$  stacked one above each other are used as the input vector  $\mathbf{Z}_{np} = (Z_{11}, \dots, Z_{1n}, \dots, Z_{p1}, \dots, Z_{pn})^T$ . The value of the  $F$  statistic is then compared with quantile  $Q_{F-1, p(n-1)}(1-\alpha)$  to decide about the result of the test.

In case of the approach via test statistics of GLM framework, the original samples  $\mathbf{Y}_i$  for  $i = 1, \dots, p$  are used as input. The values of the test statistics (11) and (12) in the Poisson case and (13) and (14) in negative binomial case are computed and compared with quantile  $Q_{\chi^2_{k-1}}(1-\alpha)$ .

This process is repeated  $k$  times for the same setting of parameters in order to compute the relative frequency of rejecting hypothesis  $H_0$  (see (2)). The value of  $h_j$  increases with every run of the algorithm (see (17)). Hence, by re-running the algorithm the values of the simulated power function are obtained.

The practical computation was done for the values of the parameters collected in Tab. I, from which the most interesting cases will be presented in the Results section of this paper. The number of repetitions of the algorithm used in the practical computation was  $k = 1000$ .

Notice that for  $p = 3$  the value  $h_j$  in each run of the program is the difference between the fixed expectation parameter  $\theta_1$  and the expectation parameters  $\theta_2, \theta_3$  of the other two distributions, which increase and decrease respectively. I. e.  $h_1 = |\theta_1 - \theta_2| = |\theta_1 - \theta_3|$ .

## Comparison of Power Functions for Different Sample Sizes

Lastly, the relation between the power functions and the size of the samples was studied via simulations. The method is based on the computations described in the section Simulated Power Functions and was carried out for values of parameters given by the Tab. I. By doing these computations for varying sample size  $n$ , a set of points of the simulated power function for chosen approach for different sample sizes was obtained. I. e. the set of values describing the relation  $\beta = \beta(h_j, n)$ . By fixing the value  $h_j$ , meaning that the difference between the expectations  $\theta_1, \theta_2$ , and  $\theta_3$  is fixed, a set of points describing the relation  $\beta = \beta(n)$  for given  $h_j$  is obtained. The described procedure is done for each method of the two approaches, and the graphical comparison of the results is presented for each method for selected choices of parameters.

The fixed values of  $h_j$  were chosen to be in 1/4, 1/2 and 3/4 of the interval, on which the simulated power functions were computed. The value of  $h_j$  in the quarter of the computational interval was chosen to observe what happens for varying sample sizes, when the differences between the expectations  $\theta_1, \theta_2$ , and  $\theta_3$  are relatively small. The value of  $h_j$  in the half of the computational interval was picked to observe the changes, that happen for increasing differences between the expectations. Finally, the value in the three quarters of the computational interval was selected to observe what happens, when the differences between the expectations are relatively big. Nonparametric regression methods were applied to fit the obtained data.

## RESULTS

### Poisson Case

We will start with a comparison of the estimations of variance and skewness of the transformed random sample, when transformations (4), (5), and (6) were applied.

In the left graph of Fig. 1 one can see a comparison of sample variance estimates for  $\theta \in [0, 100]$ . Notice, that only the transformation (5) has truly the variance stabilising property.

The comparison of sample skewness is represented graphically by the right graph of Fig. 1. Here one can observe that both transformation (5) and (6) perform very well. The skewness of the samples transformed via (5) and (6) is very close to zero, and therefore it can be assumed, that

I: Values of the parameters

Distribution	Number of samples	Sample size	Expectation parameter	Shape parameter
$Po(\theta)$	$p = 3$	$n = 100$	$\theta = 5, 10, 20, 50$	–
$NBi(\theta, \kappa)$	$p = 3$	$n = 100$	$\theta = 5, 30, 50, 100$	$\kappa = 3, 5, 10$

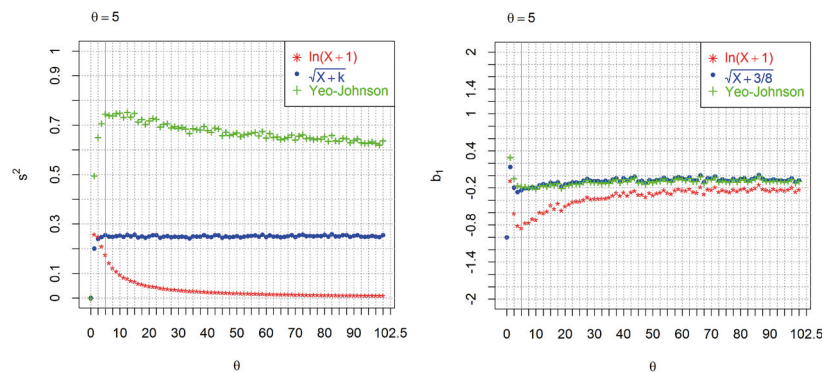
the transformed sample is approximately normally distributed. Observe that when the transformation (4) is applied, however, for small values of expectation parameter, the departure from normality may be significant.

We will continue by presenting the graphical comparison of the power functions for the Poisson case. The smooth lines in Fig. 2 represent the power functions of transformations (4) and (5) obtained via the theoretical approach. The points represent power functions obtained via simulations. Observe that the transformation (4) performs the worst, (5) and (6) both perform similarly, better than (4). The approach via likelihood ratio test statistic (11) gives slightly better results than the approach via transformations. The score statistic simulated power function (12) attains the highest values out of all the power functions in the points, where the numerical computation was carried out. However, it does not attain the value  $\beta = 0.05$  when the expectation parameters among the samples are equal. We conclude that the score test tends to have higher simulated test size than the chosen level of significance  $\alpha = 0.05$ . The simulation results for transformations (4) and (5) are in accordance with

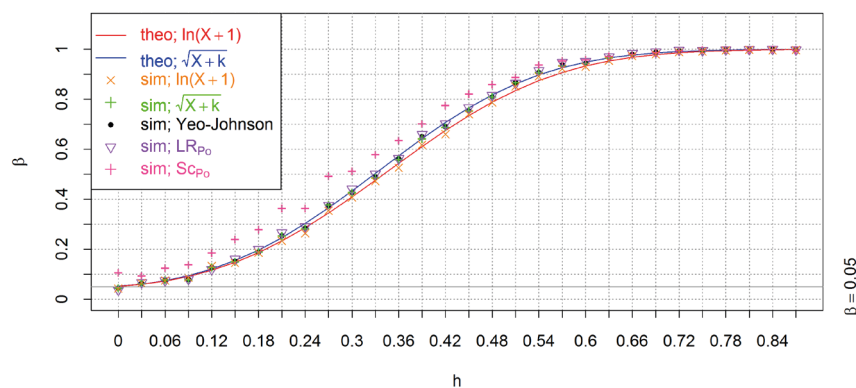
the theoretical power functions. For the increasing value of the parameter  $\theta$ , the differences between the simulated power functions become negligible.

We assume that the weak performance of the transformation (4) is caused by the fact that the skewness of the transformed sample is different from zero (see Fig. 1) and hence, the departure from the normality of the transformed sample might be significant. Furthermore, the transformation (4) lacks the variance stabilising property (see Fig. 1).

We will conclude the Poisson case by providing figures of the comparison of the power functions for different sample sizes. From this graphical comparison one can observe, what the necessary sample size is to distinguish the difference  $h_i$  with a required probability, for a given approach. Recall that the value of  $h_i$  is the difference between the fixed  $\theta_1$  and the computed  $\theta_2$  and  $\theta_3$ , respectively. The sample size needed to distinguish the difference  $h_i = 0.513$  with the probability  $\beta = 0.8$  by score test statistic (12) is  $n = 82$ . By likelihood ratio test statistic (11) it is  $n = 89$ . By the worst performing method using logarithmic transformation the required sample size is  $n = 99$  (see Fig. 3).

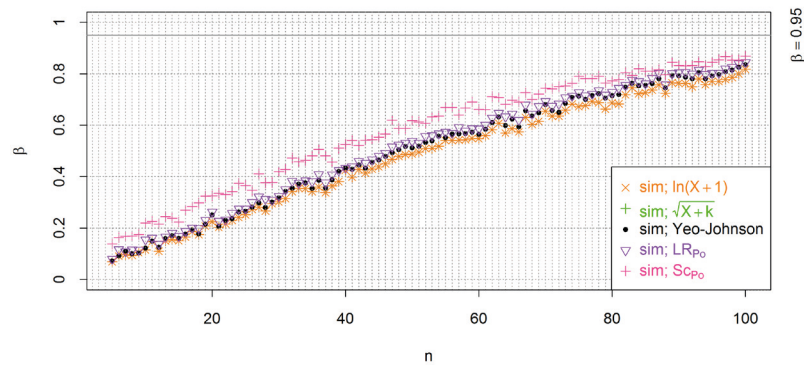


1: Poisson case - variance (left) and skewness (right) estimates for transformations (3) (red), (4) (blue) and (6) (black)



2: Poisson case – power functions comparison for  $\theta_1 = 5$ ,  $\theta_2 = \theta_1 + h$ ,  $\theta_3 = \theta_1 - h$ ,





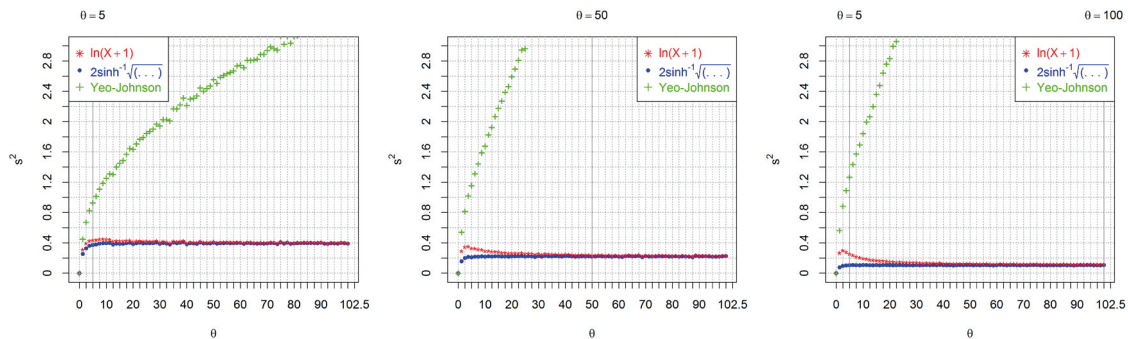
3: Poisson case – power functions as functions of a sample size for  $h_1 = 0.513$  (approximately a quarter of the computation interval),  $\theta_1 = 5, \theta_2 = \theta_1 + h = 5.513, \theta_3 = \theta_1 - h = 4.487$

### Negative Binomial Case

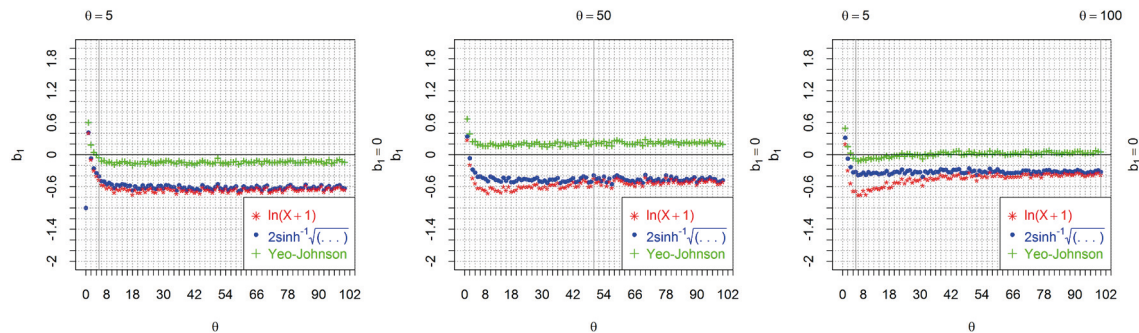
As before, we will start by comparing the variance and skewness estimates of the transformed samples when transformations (4), (7), and (6) were applied.

In Fig. 4 we can see the comparison of sample variances for increasing values of  $\kappa$ . Notice that both transformations (4) and (7) have the variance stabilising property, while (6) clearly does not.

The graphical comparison of sample skewness for increasing values of  $\kappa$  is provided in Fig. 5. Notice that the sample skewness of the samples transformed via (6) is always close to 0. For the samples obtained by applying the other two transformations this is not true. We may, however, observe, that with increasing value of  $\kappa$  the value of skewness of these two samples gets closer to 0. Additionally, for smaller



4: Negative binomial case – variance estimation via sample variance for transformations (4) (red), (7) (blue) and (6) (black) for values of parameter  $\kappa = 3, 5, 10$  from left to right



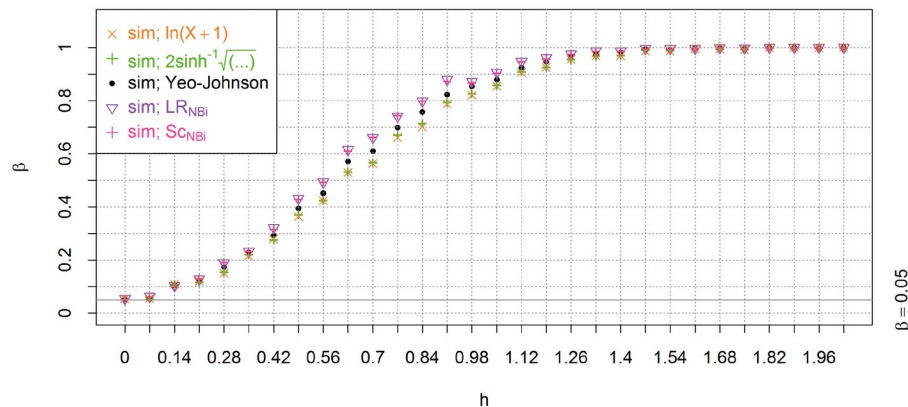
5: Fig. 5: Negative binomial case – skewness estimation via sample skewness for transformations (4) (red), (7) (blue) and (6) (black) for values of parameter  $\kappa = 3, 5, 10$  from left to right

values of  $\theta$  the transformation (7) outperforms (4) in the terms of skewness. In general we may say that for the samples transformed via (4) or (7) the departure from normality may be significant especially for small values of  $\kappa$ . For greater values of  $\kappa$  the performance of (7) slightly improves.

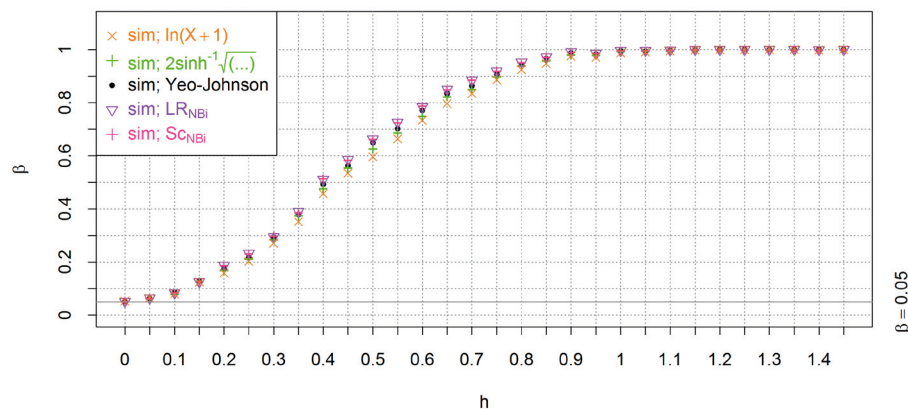
We will continue by presenting the graphical comparison of simulated power functions. Out of the multitude of computational results for various choices of parameters, we will present the cases for the following pairs of parameters  $\theta_1 = 5, \kappa = 3$ ;  $\theta_1 = 5, \kappa = 10$ ;  $\theta_1 = 100, \kappa = 10$ ; and  $\theta_1 = 50, \kappa = 5$ . The reason for the choice of the first pair is to observe, how the methods of the different approaches tackle with the worst case scenario – a relatively small expectation parameter and a small shape parameter. The second pair represents the case of a relatively small value of the expectation parameter but relatively big value of the shape parameter. The third pair represents the choice of big values of both parameters and

finally, the last pair represents the choice of moderately big expectation and shape parameters. Additionally, for the case and the comparison is enriched by the theoretical power functions for samples transformed via (4) and (7) (see Fig. 9, red and blue lines respectively). For the other two cases that are covered in the graphical comparison, the theoretical power functions are not presented because the approximations of numerical characteristics used in the computations do not behave well (for more detail see Tejkal, 2017).

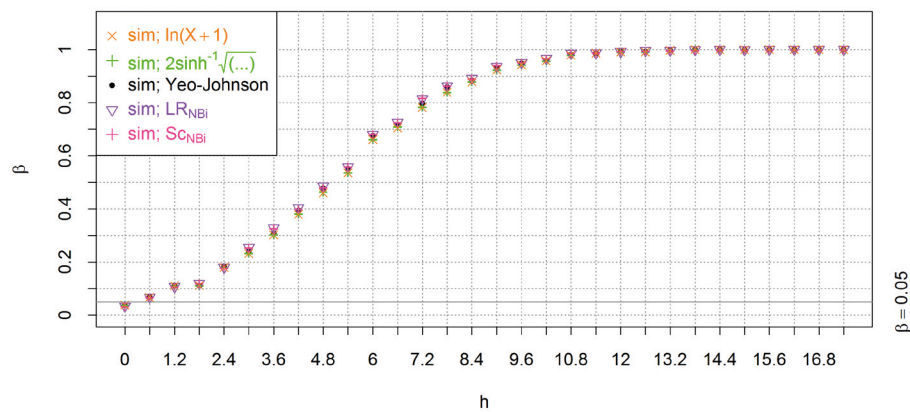
From the Figures 6, 7, and 9 we can see that the worst performance regardless of the setting of the parameters has the transformation (4). In case of small shape parameter  $\kappa$  (see Fig. 6), the difference between (4) and (7) is not very significant. However, in case of big value of the shape parameter  $\kappa$  (see Fig. 7) (7) performs better than (4). This ceases to be true for big values of the expectation parameter  $\theta$  (see Fig. 8). The best performance out of the approach via transformations has transformation (6).



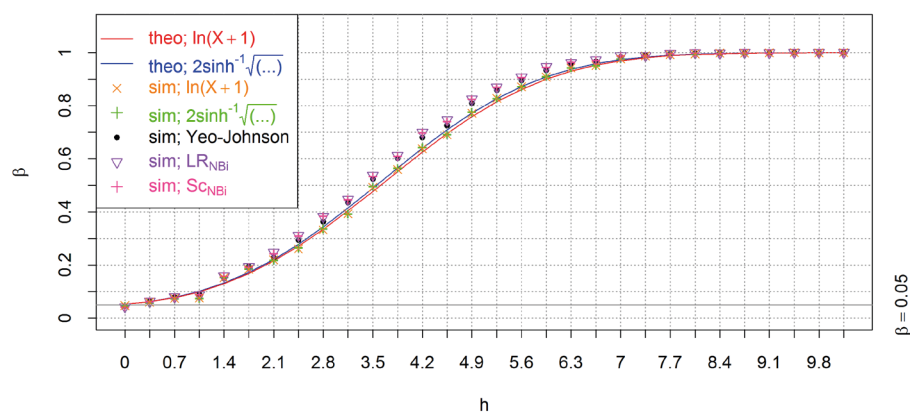
6: Negative binomial case – Power functions comparison for  $\kappa = 3, \theta_1 = 5, \theta_2 = \theta_1 + h, \theta_3 = \theta_1 - h$



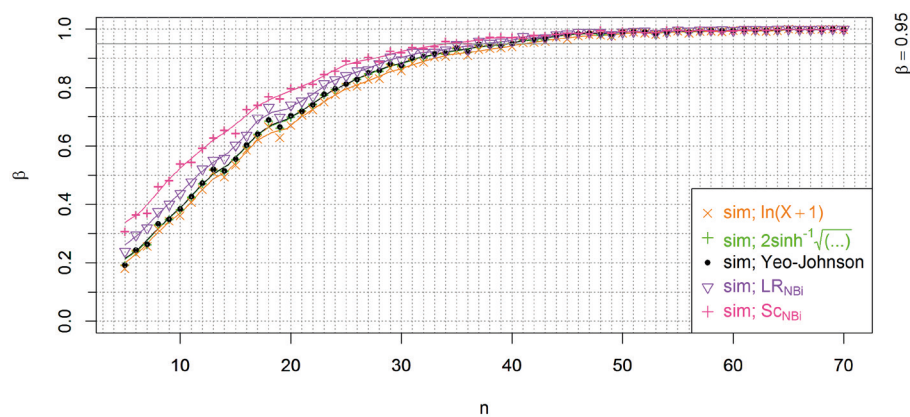
7: Negative binomial case – power functions comparison for  $\kappa = 10, \theta_1 = 5, \theta_2 = \theta_1 + h, \theta_3 = \theta_1 - h$



8: Negative binomial case – power functions comparison for  $\kappa = 10, \theta_1 = 100, \theta_2 = \theta_1 + h, \theta_3 = \theta_1 - h$



9: Negative binomial case – power functions comparison for  $\kappa = 5, \theta_1 = 50, \theta_2 = \theta_1 + h, \theta_3 = \theta_1 - h$



10: Negative binomial case – power functions as functions of a sample size at  $h_1 = 7.98$  (approximately a half of the computation interval) for the setting of parameters  $\kappa = 3$  and  $\theta_1 = 30, \theta_2 = \theta_1 + h = 37.98, \theta_3 = \theta_1 - h = 22.02$



Notice also, that for a good performance of a transformation it seems to be more important, that the skewness of the transformed sample is close to 0, than that the variances among the samples are equal. Observe that (6), that performs the best out of the considered transformations, does not stabilise variance in the particular studied case (see Fig. 4), however, the skewness of samples transformed via (6) is close to 0 (see Fig. 5).

Using test statistic coming from the theory of maximum likelihood seems to be the best approach since the power function of the test when either (13) or (14) is used attains higher values than any power function of the approach via transformations. The simulated power function of the test statistic (14) tends to attain slightly higher values than the one of (13). We may further observe, that for increasing values of  $\theta$  and  $\kappa$  the differences between the power functions become smaller (cf. Fig. 6 with 9, and 6 with 7).

Lastly, we will present the graphical comparison of the power functions for different sample sizes for all the studied approaches. Recall that the value of  $h_i$  is the difference between the fixed  $\theta_1$  and the computed  $\theta_2$  and  $\theta_3$ , respectively.

From Fig. 10 we can observe, that for the best performing method (likelihood ratio test statistic (13)), to distinguish the difference  $h_i = 7.98$  with probability  $\beta = 0.8$ , samples of a size  $n = 23$  are needed. For the worst performing pair of methods (logarithmic transformation (4) and argument of hyperbolic sine transformation (7)), to distinguish the difference  $h_i = 7.98$  with probability  $\beta = 0.8$ , samples of a size  $n = 28$  are needed.

## DISCUSSION

The early trend in tackling the non-normal heteroscedastic data involving counts in linear regression and the respective tests was to apply various transformations in order to achieve normality and homoscedasticity. To improve the desired effect

of the transformations of the random variables with Poisson and negative binomial distribution, Anscombe (1948) introduced generalisations of the variance stabilising transformations. For data involving counts it is often convenient to use the logarithmic scale, therefore the logarithmic transformation comes, perhaps, as a natural choice. The problem of zero observations is usually solved by adding a constant (usually one) into the argument of the logarithm. Some authors even suggest different constants. Anscombe (1948) derives the logarithmic transformation for negative binomially distributed random variable as an approximation of the more complicated argument of hyperbolic sine transformation and finds an optimal value of the constant added within the argument to be a function of the shape parameter  $\kappa$  of the distribution. Yamamura (1999) suggests using transformation  $Y = \ln(X + 0.5)$  instead of  $Y = \ln(X + 1)$ . Numerical analysis carried out when collecting data for this paper, however, did not suggest, that there is a significant improvement in the power of the test when Anscombe's or Yamamura's choice of the optimal constant was used instead of adding one within the argument of the logarithm. On the other hand, the family of transformations developed by Yeo and Johnson (2000) performed in our comparison significantly better than both the logarithmic transformation with one added within the argument or the transformations proposed by Anscombe.

More recent development (see O'Hara and Kotze, 2010) has shown that for modelling data involving discrete counts, models based on Poisson or negative binomial distribution attain better results than the approach via transformations. The results of this paper in general support this statement. The GLM based tests outperformed the approach via transformations in all studied cases. However, for some cases, the difference between the GLMs and the approach via Yeo-Johnson transformation, the best performing one out of the studied transformations, was rather small.

## CONCLUSION

The presented analysis provides a comparison of different methods of testing the hypothesis of equality of expectations for data involving counts. Poisson and negative binomial distributions were used to model the data. Two main approaches were used, the approach via normalising and variance stabilising transformations, followed by applying the classical ANOVA towards the transformed data, and the approach via GLM test statistics.

In the Poisson case, the GLM test statistics performed better than the transformational approach. By far the highest values were attained by the simulated power function of the score test statistic (12) followed by the power function of the likelihood ratio test statistic (11). The likelihood ratio test statistic, in turn, performed slightly better than any of the transformations. It should be noted, however, that the score test tends to have greater simulated test size than the chosen level of significance  $\alpha = 0.05$ , since the respective power function did not attain the value 0.05 for  $h = 0$ . The square root (5) and the Yeo-Johnson transformation (6) both performed similarly and both better than the logarithmic (4) transformation, which performed the worst. With increasing values of the expectation parameter  $\theta$  of the Poisson distribution, the differences between the power functions of the different approaches became negligible.

In the negative binomial case, the results were dependent both on the choice of the initial value of the expectation parameter  $\theta$  and the shape parameter  $\kappa$ . In all settings, better result was obtained when

the GLM test statistics were used. The simulated power function of the score statistic (14) attained higher values than the one of the likelihood ratio statistic (13). All of the transformations performed worse than the GLM test statistics. The Yeo-Johnson transformation (6) outperformed the argument of hyperbolic sine transformation (7) and the logarithmic transformation (4). The argument of hyperbolic sine achieved better results than the logarithm for small values of  $\theta$  and big values of  $\kappa$ . The difference between these two transformations for small values of  $\theta$  and  $\kappa$  and for big values of  $\theta$  independently of  $\kappa$  were not significant. For increasing values of the parameters  $\theta$  and  $\kappa$  the differences between the power functions of the different approaches became negligible.

Finally, by comparing the power functions for different sample sizes we saw, that by choosing the right approach, we can obtain a given power of the test with significantly smaller sample, than in the case of picking a less suitable approach. Namely, the difference between the worst and best performing method in the comparison provided in this paper was 10 samples in the Poisson case and 5 samples in the negative binomial case.

## REFERENCES

- ANDĚL, J. 2011. *Basics of mathematical statistics* [in Czech: *Základy matematické statistiky*]. Praha: Matfyzpress.
- ANSCOMBE, F. J. 1948. The transformation of Poisson, binomial and negative binomial data. *Biometrika*, 35(3–4): 246–254.
- HRDLIČKOVÁ, Z. 2006. Comparison of the power of the tests in one-way ANOVA type model with Poisson distributed variables. *Environmetrics*, 17(3): 227–237.
- HRDLIČKOVÁ, Z. 2008. Approximation of powers of some tests in one-way MANOVA type multivariate generalized linear model. *Computational Statistics & Data Analysis*, 52(8): 4059–4075.
- HRDLIČKOVÁ, Z. 2002. *Log-linear models with Poisson random variables* [in Czech: *Log-lineární modely s Poissonovskými proměnnými*]. Diploma Thesis. Masaryk University in Brno, Faculty of Science. Supervisor: doc. RNDr. Jaroslav Michálek, CSc.
- HÜBNEROVÁ, Z. and DOUDOVÁ, L. 2011. Influence of estimate of shape parameter on tests in models with negative binomial distribution. In: *Biometric Methods and Models in Current Science and Research*. Brno: UKZUZ, pp. 119–126.
- O'HARA, R. B. and KOTZE, D. J. 2010. Do not log-transform count data. *Methods in Ecology and Evolution*, 1(2): 118–122.
- MOSS, D. and MCPHEE, D. P. 2006. The impacts of recreational four-wheel driving on the abundance of the ghost crab (*Ocypode Cordimanus*) on sub-tropical sandy beaches in SE Queensland. *Coastal Management*, 34: 133–140.
- SCHEFFÉ, H. 1999. *The analysis of variance*. Wiley classics library ed. New York: Wiley-Interscience Publication.
- TEJKAL, M. 2017. *Selected random variables transformations used in classical linear regression*. Diploma thesis. Brno University of Technology, Faculty of Mechanical Engineering. Supervisor: Doc. Mgr. Zuzana Hübnerová, Ph. D.
- YAMAMURA, K. 1999. Transformation using  $(x + 0.5)$  to stabilize the variance of populations. *Researches on Population Ecology*, 41(3): 229–234.
- YEO, I., and JOHNSON R. A. 2000. A new family of power transformations to improve normality or symmetry. *Biometrika*, 87(4): 954–959.

## Contact information

Martin Tejkal: martin.tejkal@unob.cz

Zuzana Hübnerová: hubnerova@fme.vutbr.cz