Statistics and Probability Letters 83 (2013) 1339–1345

Contents lists available at SciVerse ScienceDirect

Statistics and Probability Letters

journal homepage: www.elsevier.com/locate/stapro

Exact tests of the superiority under the Poisson distribution

MingTe Liu [a], Huey-Miin Hsueh [b,*]

[a] General Education Center, Tatung Institute of Commerce and Technology, Chiayi, Taiwan

[b] Department of Statistics, National Chengchi University, Taipei 11605, Taiwan

ARTICLE INFO

Article history:
Received 4 September 2012
Received in revised form 9 January 2013
Accepted 12 January 2013
Available online 5 February 2013

Keywords:
Exact test
p-value
Superiority
Type I error rate
Wald statistic

ABSTRACT

We propose two exact *p*-values of two commonly-used test statistics for testing the superiority under Poisson populations. We show that, the computationally-intensive confidence-set *p*-value involves at most a supremum search over a closed interval of a single argument. On the other hand, the estimated *p*-value has adequate performances empirically.

1. Introduction

The Poisson distribution is suitable to model a rare event in a variety of fields, such as biology, commerce, quality control, and so on. Real examples include a breast cancer study in Ng and Tang (2005) and an evaluation of a new vaccine in Chan and Wang (2009). Recently, the Poisson distribution has been used to model the number of mapping reads for each gene in an RNA-seq experiment. See Wang et al. (2010).

In comparing two Poisson distributions, the asymptotic test can be too liberal in finite sample cases. When the sample scales are not large, the exact testing procedure is more appropriate for prevention of an inflated type I error rate. One major challenge in the development of an exact test is the presence of nuisance parameters. When the null hypothesis states the equality of the two populations, the classical conditional test uses conditioning to get rid of the nuisance parameter(s). However, with a null hypothesis of non-superiority, the conditioning fails to eliminate the nuisance parameter(s) completely.

One easy way to deal with the nuisance parameters is to estimate the *p*-value by plugging in some consistent estimates of the nuisance parameters. Krishnamoorthy and Thomson (2004) first introduced the use of the restricted maximum likelihood estimate (RMLE) of the common Poisson mean under the null hypothesis of equal means. Ng et al. (2007) extended the idea to the problems with a nonzero difference null hypothesis and proposed the numerical approximation *p*-value. Chan and Wang (2009) use the RMLEs at the boundary of the null space for stratified data. In this study, we will propose an exact method under the estimation principle as well. The RMLEs of the nuisance parameters by taking the null space of non-superiority into account is employed in estimating an exact *p*-value. Although the estimated *p*-value is easy and quick to implement, it does not guarantee a well-controlled type I error rate.

For a strict control of the type I error rate, the maximization principle is commonly suggested for a *p*-value instead. Much more effort is required for calculation. Since the Poisson mean has an infinite parameter space, it even increases the

* Corresponding author. Tel.: +886 2 29393091x81138; fax: +886 2 29398024.
E-mail addresses: mingte@ms2.ttc.edu.tw (M. Liu), hsueh@nccu.edu.tw (H.-M. Hsueh).

0167-7152/$ – see front matter
doi:10.1016/j.spl.2013.01.030

computational difficulty. To limit the search space, we consider the confidence-set p-value proposed by Berger and Boos (1994). Chan and Wang (2009) had applied this test for stratified analysis as well. In this study, we will show that, from the beneficial properties of the Poisson distribution and of the two popular test statistics, the calculation of the corresponding confidence-set p-values can be simplified.

In this study, we consider the two commonly-used test statistics as ordering criterion among all possible samples. The test statistics and the proposed exact p-values are introduced in Section 2. Section 3 presents a numerical study to justify the proposed tests. Section 4 applies the proposed tests on a real example for illustration. A discussion is given in Section 5.

2. Test procedures

2.1. Notations and test statistics

Assume two independent Poisson random samples observed within a fixed duration, $(Y_{11}, \ldots, Y_{1n_1})$, $(Y_{21}, \ldots, Y_{2n_2})$, where for $i = 1, \ldots, n_1, j = 1, \ldots, n_2$,

$$Y_{1i} \overset{iid}{\sim} Poi(\lambda_1), \qquad Y_{2j} \overset{iid}{\sim} Poi(\lambda_2),$$

and λ_1, λ_2 are the mean incidence numbers within the duration, respectively. It is known that the sums, $Y_1 = \sum_{i=1}^{n_1} Y_{1i}$, $Y_2 = \sum_{j=1}^{n_2} Y_{2j}$, are the sufficient statistics, and the sample means, $\bar{Y}_1 = Y_1/n_1$, $\bar{Y}_2 = Y_2/n_2$, are the MLE of λ_1, λ_2, respectively.

In application, the rare event under study may be some adverse effect, incidence or recurrence of some disease. A treatment is superior in terms of lowering the mean incidence rate. Hence Y_{2j}s stand for the realizations of incidences from a treatment group, while Y_{1i}s are incidences from some reference or control group. The research interest is to test the superiority of the treatment with the following hypotheses, i.e.

$$H_0 : \lambda_1 \le \lambda_2, \quad \text{versus } H_1 : \lambda_1 > \lambda_2.$$

Denote the null space as $\Omega_0 = \{\lambda_1 \le \lambda_2\}$. Two common test statistics based on the difference of the two MLEs by using different estimated standard errors in the denominator are considered,

$$Z_R = \frac{\bar{Y}_1 - \bar{Y}_2}{\sqrt{\frac{\tilde{\lambda}_0}{n_1} + \frac{\tilde{\lambda}_0}{n_2}}}, \quad \text{and} \quad Z_U = \frac{\bar{Y}_1 - \bar{Y}_2}{\sqrt{\frac{\bar{Y}_1}{n_1} + \frac{\bar{Y}_2}{n_2}}},$$

where $\tilde{\lambda}_0 = \frac{Y_1+Y_2}{n_1+n_2}$ is the RMLE of the common mean at the boundary of Ω_0, $\lambda_1 = \lambda_2$. Note that, the latter one is the Wald statistic. The null hypothesis is rejected if a sufficiently large value of the test statistic is observed; or given an observed value of a test statistic, the p-value is not greater than the significance level α. When the sample sizes are sufficiently large, an approximated p-value is calculated under normal limiting distribution. On the other hand, if the sample scales are of small to moderate size or the experimental duration is not lengthy, an exact p-value based on the exact null sampling distribution is more adequate in prevention of an inflated type I error rate.

2.2. Exact P-values

Given some observed values, z_R, z_U, of the two test statistics, the exact p-values are

$$p_{E,R} = P(Z_R \ge z_R|\lambda_1, \lambda_2), \qquad p_{E,U} = P(Z_U \ge z_U|\lambda_1, \lambda_2),$$

where the probabilities are evaluated under Poisson distributions with $(\lambda_1, \lambda_2) \in \Omega_0$. Note that, it can be shown that Z_R, Z_U are functions of Y_1, Y_2, which independently follow a Poisson distribution with mean $n_1\lambda_1, n_2\lambda_2$, respectively. Hence the exact probability is calculated by

$$P(A|\lambda_1, \lambda_2) = \sum_{y_1=0}^{\infty} \sum_{y_2=0}^{\infty} p(y_1|n_1\lambda_1)p(y_2|n_2\lambda_2)I_A,$$

where $p(\cdot|l)$ is the probability function of a Poisson distribution with mean l; I_A is the indicator of the event A.

To strictly control the size, one can consider the standard p-value, defined as the supremum of an exact p-value over the null space, see Casella and Berger (1990). However, because the null space is infinite, the computation of a standard exact p-value is a complicated and time-consuming task. Alternatively, we consider the confidence-set p-values proposed by Berger and Boos (1994),

$$p_{CI,R}^{(\gamma)} = \sup_{(\lambda_1,\lambda_2)\in C_\gamma} P(Z_R \ge z_R|\lambda_1, \lambda_2) + \gamma, \tag{1}$$

$$p_{CI,U}^{(\gamma)} = \sup_{(\lambda_1,\lambda_2)\in C_\gamma} P(Z_U \ge z_U|\lambda_1, \lambda_2) + \gamma, \tag{2}$$

where C_γ is a joint confidence set of (λ_1, λ_2) that guarantees $100(1-\gamma)\%$ confidence within Ω_0. Note that, the tradeoff γ should be far less than α for the possibility of obtaining a significant result. The construction of a confidence set in a restricted null parameter space is less straightforward. We propose using the intersection of the null parameter space Ω_0 and a $100(1-\gamma)\%$ unconstrained confidence set $C_{U,\gamma}$ built under the whole parameter space. That is, $C_\gamma = \Omega_0 \cap C_{U,\gamma}$. Similar approaches can be found in Wang (2008). For the unconstrained confidence set, we consider the following cross-product set, $C_{U,\gamma} = \{L_1 \le \lambda_1 \le U_1, L_2 \le \lambda \le U_2\}$, where (L_1, U_1) and (L_2, U_2) are independent $100\sqrt{1-\gamma}\%$ confidence intervals of λ_1 and λ_2, respectively. The conventional confidence intervals derived through the relation between a Poisson distribution and a chi-square distribution are employed,

$$(L_i, U_i) = \frac{1}{2n_i}\left(\chi^2_{\{1-(1-\sqrt{1-\gamma})/2, 2y_i\}}, \chi^2_{\{(1-\sqrt{1-\gamma})/2, 2(y_i+1)\}}\right), \quad i = 1, 2,$$

where $\chi^2_{\{\delta,k\}}$ is the $100(1-\delta)$-th percentile of a chi-square distribution with degree of freedom k. As a consequence, if $C_{U,\gamma}$ and Ω_0 are not mutually disjoint, $C_\gamma = \{L_1 \le \lambda_1 \le \min(U_1, \lambda_2). L_2 \le \lambda_2 \le U_2\}$. Otherwise, $C_\gamma = \emptyset$. In latter case, we then define the p-values as equal to γ and reject the null hypothesis. Note that, this case occurs when the sample evidence shows a great deviation from the null scenario. From Berger and Boos (1994), the confidence-set p-values are always valid. In the following, we show that the calculations of $p^{(\gamma)}_{CI,R}$, $p^{(\gamma)}_{CI,U}$ can be further simplified.

Barnard (1947) proposed the following convexity condition for a test statistic under a bivariate discrete distribution:

$$S(s_1, s_2+1) \le S(s_1, s_2) \le S(s_1+1, s_2),$$

for any sample point (s_1, s_2) of the two discrete random variables. With no occurrence of ties, the strict convexity is defined by

$$S(s_1, s_2+1) < S(s_1, s_2) < S(s_1+1, s_2).$$

The condition indicates that the statistic has a monotonicity in ordering the sample points of the support. Röhmel and Mansmann (1999) proved that under the Binomial distribution, a test statistic satisfying the convexity condition, the supremum of the correspondent exact p-value is attained at a boundary point over a compact parameter space. The calculation of the maximal exact p-value can be simplified to solely take the boundary into consideration. The following theorem and corollary are generalizations of the problem of comparing two Poisson means.

Theorem 1. *In comparing two Poisson means, let S be a test statistic depending on the data only through the sufficient statistics (Y_1, Y_2). Suppose S satisfies the strict convexity condition. Then given s_0, the supremum of $P(S \ge s_0|\lambda_1, \lambda_2)$ over a compact set in the parameter space occurs at a boundary point.*

In fact, it can be shown that Z_R and Z_U both satisfy the strict convexity condition, and subsequently the maximums for their confidence-set p-values occur in the boundary of C_γ. Further, in proving Theorem 1, it is seen that the probability $P(S \ge s_0|\lambda_1, \lambda_2)$ increases as λ_1 increases and λ_2 decreases. Hence, the maximum occurs at the bottom right corner. It either occurs at one specific point or on a closed interval over the diagonal. The supremum can be found by a grid-search in the latter case. Please see the following corollary.

Corollary 2. *If $C_\gamma \ne \emptyset$, let $C_d = [\max(L_1, L_2), \min(U_1, U_2)]$. Then the confidence-set p-values of Z_R and Z_U, in (1) and (2), are equal to*

$$p^{(\gamma)}_{CI,R} = \begin{cases} P(Z_R \ge z_R|U_1, L_2) + \gamma, & \text{if } C_\gamma = C_{U,\gamma} \\ \sup\limits_{\lambda \in C_d} P(Z_R \ge z_R|\lambda, \lambda) + \gamma, & \text{if } C_\gamma \subset C_{U,\gamma} \end{cases} \tag{3}$$

and

$$p^{(\gamma)}_{CI,U} = \begin{cases} P(Z_U \ge z_U|U_1, L_2) + \gamma, & \text{if } C_\gamma = C_{U,\gamma} \\ \sup\limits_{\lambda \in C_d} P(Z_U \ge z_U|\lambda, \lambda) + \gamma, & \text{if } C_\gamma \subset C_{U,\gamma}. \end{cases} \tag{4}$$

On the other hand, the easiest way to solve for the presence of nuisance parameters in finding an exact p-value is to replace the parameters with some adequate estimates under the null space Ω_0, which produces an estimated exact p-value. We consider the RMLEs for a most likely outcome given the observed data. The null space Ω_0 is a simple star-shaped scenario, Ng et al. (2007) indicated that the RMLEs of λ_1, λ_2 over Ω_0 are equal to

$$(\tilde\lambda_{01}, \tilde\lambda_{02}) = \begin{cases} (\hat\lambda_1, \hat\lambda_2), & \text{if } \hat\lambda_1 \le \hat\lambda_2; \\ (\tilde\lambda_0, \tilde\lambda_0), & \text{if } \hat\lambda_1 > \hat\lambda_2. \end{cases}$$

Then the estimated exact p-values are given by

$$\tilde p_{E,R} = P(Z_R \ge z_R|\tilde\lambda_{01}, \tilde\lambda_{02}), \qquad \tilde p_{E,U} = P(Z_U \ge z_U|\tilde\lambda_{01}, \tilde\lambda_{02}). \tag{5}$$

Table 1
The type I error rate (for $\delta_0 \le 0$) and the power (for $\delta_0 > 0$) of the proposed p-values of Z_R, Z_U at $\delta_0 \in [-.25, 2]\lambda_2 = 1$, $n_2 = 10$ and $\alpha = 5\%$.

n_1	Statistic	p-value	δ_0									
			−0.25	−0.15	−0.1	−0.05	0.0	0.1	0.5	1.0	1.5	2.0
6	Z_R	$p_{CI,R}^{(0.001)}$	0.0096	0.0176	0.0231	0.0297	0.0375	0.0574	0.1942	0.4524	0.6999	0.8655
		$\tilde{p}_{E,R}$	0.0137	0.0233	0.0297	0.0372	0.0460	0.0675	0.2099	0.4728	0.7194	0.8781
		$p_{A,R}$	0.0157	0.0266	0.0337	0.0421	0.0519	0.0757	0.2298	0.5024	0.7432	0.8907
	Z_U	$p_{CI,U}^{(0.001)}$	0.0129	0.0228	0.0293	0.0370	0.0461	0.0682	0.2120	0.4743	0.7199	0.8782
		$\tilde{p}_{E,U}$	0.0145	0.0250	0.0318	0.0399	0.0493	0.0721	0.2199	0.4871	0.7310	0.8841
		$p_{A,U}$	0.0082	0.0153	0.0202	0.0262	0.0334	0.0517	0.1833	0.4425	0.6942	0.8623
10	Z_R/Z_U	$p_{CI}^{(0.001)}$	0.0123	0.0227	0.0298	0.0384	0.0487	0.0746	0.2544	0.5724	0.8223	0.9451
		$\tilde{p}_E$	0.0123	0.0227	0.0298	0.0384	0.0487	0.0747	0.2554	0.5773	0.8279	0.9477
		p_A	0.0126	0.0230	0.0301	0.0387	0.0489	0.0748	0.2554	0.5773	0.8279	0.9477
17	Z_R	$p_{CI,R}^{(0.001)}$	0.0101	0.0196	0.0265	0.0351	0.0457	0.0736	0.2831	0.6495	0.8912	0.9776
		$\tilde{p}_{E,R}$	0.0112	0.0216	0.0289	0.0379	0.0488	0.0771	0.2858	0.6560	0.8954	0.9792
		$p_{A,R}$	0.0101	0.0196	0.0264	0.0351	0.0457	0.0736	0.2831	0.6497	0.8918	0.9782
	Z_U	$p_{CI,U}^{(0.001)}$	0.0082	0.0168	0.0231	0.0312	0.0411	0.0674	0.2678	0.6323	0.8864	0.9771
		$\tilde{p}_{E,U}$	0.0111	0.0216	0.0291	0.0385	0.0499	0.0795	0.2945	0.6592	0.8957	0.9793
		$p_{A,U}$	0.0159	0.0292	0.0384	0.0496	0.0629	0.0964	0.3250	0.6888	0.9100	0.9831

One can see that the RMLEs do not necessarily occur at the diagonal boundary. Since an exact p-value of the test statistic satisfying the convexity condition is an increasing function as the parameter point (λ_1, λ_2) moves toward bottom right, our estimated exact p-value is more powerful than the other estimated p-value evaluated at the diagonal, as proposed by Krishnamoorthy and Thomson (2004) and by Chan and Wang (2009).

3. Numerical study

In this numerical study, we investigate the performance of the proposed exact tests. Denote $\delta_0 = \lambda_1 - \lambda_2$. We consider $\lambda_2 = 1$, δ_0 ranged from -0.25 to 2, $n_2 = 10$ and three $n_1 = 6, 10, 17$. The nominal significance level α is set at 0.05. Table 1 reports the exact probabilities of rejecting the null hypothesis, which are the type I error rate, when $\delta_0 \le 0$, and are the power, when $\delta_0 > 0$. The confidence-set p-value is constructed with $(1 - \gamma) = 99.9\%$. For comparison, the results of the asymptotic tests by using the asymptotic normality, denoted by p_A, are also listed in the tables. The tables of $\lambda_2 = 2$ are provided in the supplementary materials. Note that, $n_1 = n_2$, Z_R and Z_U are actually of the same form and have the same results.

We find that, the two exact p-values have their sizes well controlled at $\alpha = 5\%$. On the contrary, the type I error rate of the asymptotic method can exceed the significance level. Using the same test statistic, the size of the estimated p-value $\tilde{p}_E$ is always closer to the nominal level and is more efficient in computations than the confidence-set p-value p_{CI}. However, the validity of the estimated p-value is not theoretically justified and hence is not guaranteed.

The test statistic used in the procedure affects the performance of the exact test. For the confidence-set p-value, the employment of Z_U is more powerful than that of Z_R at $n_1 < n_2$, and less powerful than Z_R at $n_1 > n_2$. For the estimated p-value, the use of Z_U always brings about more powerful results than Z_R when $n_1 \neq n_2$.

For a confidence-set p-value, a larger γ leads to less computations involved for the supremum search and subsequently, a smaller supremum obtained. However, by adding this trade-off term, the resultant confidence-set p-value hardly varies. Our numerical study indicates that the test is not significantly affected by γ. Please refer to the supplementary materials.

4. Real example

Consider the breast cancer study in Ng and Tang (2005). Female subjects were classified according to whether or not they had been examined by using X-ray fluoroscopy during treatment for tuberculosis. The investigators suspect that, the use of X-ray fluoroscopy will lead to a higher occurrence rate of breast cancer. Define λ_1 as the mean incidence number of breast cancer per person-year of the treatment group, in which the patients had received X-ray; and let λ_2 be the mean incidence number per person-year of the control group, in which the patients were not examined by X-ray. The research problem is to test the following hypothesis, $H_0 : \lambda_1 \le \lambda_2$ versus $H_1 : \lambda_1 > \lambda_2$. The procedures proposed are extended to the case where observations have various experimental durations. Assume Y_{ij} is the Poisson random variable in the i-th group with t_{ij} units of duration, for $i = 1, 2, j = 1, 2, \ldots, m_i$. Define $n_i^* = \sum_{j=1}^{m_i} t_{ij}$, $i = 1, 2$. Then, replacing n_i's by n_i^*'s in the test statistics, we can apply the proposed approaches directly.

From Ng and Tang (2005), it was reported that the treatment group had $y_1 = 41$ cases of breast cancer in $n_1^* = 28010$ persons-year at risk and the control group had $y_2 = 15$ cases of breast cancer in $n_2^* = 19017$ person-years at risk. The MLEs of λ_1, λ_2 are $\hat{\lambda}_1 = 1.464$, $\hat{\lambda}_2 = 0.789$, and the RMLE of the common mean value is $\tilde{\lambda}_0 = 1.191$ per 1000 person-year. In the following, all the estimates are expressed in the unit of 1000 person-year. Consequently, the observed Z_R, Z_U are $z_R = 2.0818$, $z_U = 2.2047$ with correspondent asymptotic p-values 0.0187, 0.0137, respectively.

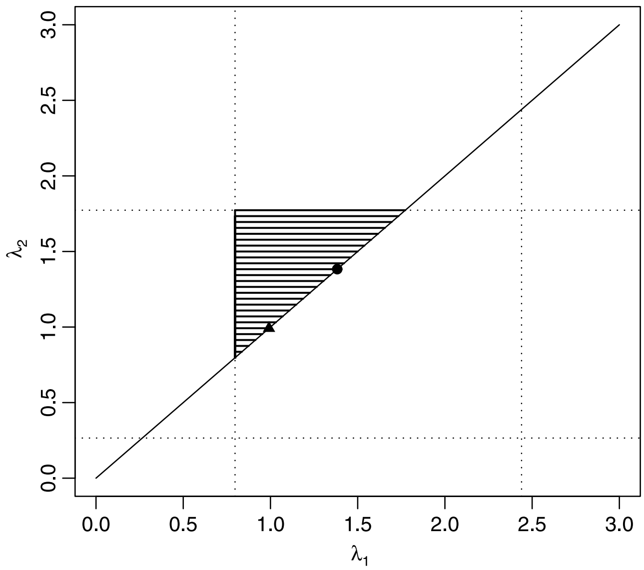

Fig. 1. The shaded area is the constrained joint 99.9% confidence-set C_γ of (λ_1, λ_2) in the example. The two maximal exact p-values occur at the diagonal. The p-value is evaluated at $\lambda_1 = \lambda_2 = 1.3831$ (the circle) for Z_R, at $\lambda_1 = \lambda_2 = 0.9916$ (the triangle) for Z_U.

Table 2
The asymptotic, estimated, and confidence-set p-values of Z_R and Z_U.

Test statistic	Z_R	Z_U
Confidence-set p-value ($\gamma = 0.001$)	0.0183	0.0187
Estimated p-value	0.0179	0.0184
Asymptotic p-value	0.0187	0.0137

For the confidence-set p-value, the unconstrained joint 99.9% (with $\gamma = 0.001$) confidence set of (λ_1, λ_2) is $C_{U,0.001} = \{0.7959 \le \lambda_1 \le 2.4420, 0.2657 \le \lambda_2 \le 1.7746\}$. Therefore, $C_{0.001} \ne \emptyset$ and $C_{0.001} \subset C_{U,0.001}$. See Fig. 1. As a result, the confidence-set p-value is evaluated somewhere on the diagonal $\lambda_1 = \lambda_2$ between $C_d = [\max(L_1, L_2) = 0.7959, \min(U_1, U_2) = 1.7746]$. Via a grid-search, it is found that the confidence-set p-value of Z_R occurs at $\lambda_1 = \lambda_2 = 1.3831$, and the confidence-set p-value of Z_U occurs at $\lambda_1 = \lambda_2 = 0.9916$. The calculated p-values are reported in Table 2. On the other hand, because the MLEs $\hat{\lambda}_1 > \hat{\lambda}_2$ are outside the null hypothetical space, the estimated exact p-value is evaluated at the RMLE on the diagonal, $\tilde{\lambda}_0 = 1.191$.

All p-values in Table 2 are less than $\alpha = 5\%$ and lead to the conclusion of rejecting the null hypothesis. The increase in the incidence rate of breast cancer by using the X-ray fluoroscopy achieves a statistical significance.

5. Discussion

This study focuses on testing the superiority of one Poisson distribution against another with a smaller mean value for a small to moderate data set. Differing with the majority of the existing literature, we consider a non-superiority null hypothesis and hence deal with a broader null parameter space, which introduces more difficulties for both theoretical justification and practical computation of a p-value. Using a test that considers only the boundary of equality and ignores the null space of inferiority risks an inflation of its type I error rate. For a strict control of the error rate, we consider the confidence-set p-value. As the two proposed test statistics satisfy the convexity condition strictly, the calculations for the confidence-set p-values are greatly simplified to involve at most a supremum search over a closed interval of a single argument. On the other hand, we propose the estimated exact p-value by using the RMLEs under the null space. In the numerical study, both the proposed p-values perform adequately.

Acknowledgments

The authors sincerely thank the editor-in-chief, Prof. Hira Koul, and the referee for their helpful suggestions and assistance in improving their manuscript. The authors would also like to thank Drew McNeil for his careful editing of their manuscript. This work was supported by the National Science Council of Taiwan, R.O.C. under the grant (NSC 99-2118-M-004-005).

Appendix A. Proofs

A.1. Proof of Theorem 1

It is known that for the Poisson probability function $p(x|\lambda)$, by letting $p(-1|\lambda) = 0$, for $x \geq 0$,

$$\frac{\partial}{\partial \lambda} p(x|\lambda) = p(x-1|\lambda) - p(x|\lambda).$$

Suppose S is a function of the sufficient statistics (Y_1, Y_2). If given the observed value (y_{10}, y_{20}), $S = s_0$, then the exact p-value based on S is

$$P_S(n_1\lambda_1, n_2\lambda_2) \equiv P(S \geq s_0|\lambda_1, \lambda_2) = \sum_{S(y_1, y_2) \geq s_0} p(y_1 \mid n_1\lambda_1) p(y_2 \mid n_2\lambda_2).$$

If S satisfies the convexity condition, then there exists a function h such that, at each fixed $y_2 \geq 0$, $h(y_2)$ is the lower limit of y_1 that satisfies $\{y_1 : S(y_1, y_2) \geq s_0\} = \{y_1 \geq h(y_2)\} \equiv R_{y_2}$. Similarly, for some non-negative integer a, there exists a function h^* such that at each fixed $y_1 \geq a$, $h^*(y_1)$ is the upper limit of y_2 that satisfies $\{y_2 : S(y_1, y_2) \geq s_0\} = \{y_2 \leq h^*(y_1)\} \equiv R^*_{y_1}$. Furthermore, if S satisfies the strict convexity condition, for any y_1, y_2 in the support,

$$R_{y_2+1} \subset R_{y_2}, \quad \text{and} \quad R^*_{y_1} \subset R^*_{y_1+1}. \tag{6}$$

Consequently, the p-value has the following expression:

$$P_S(n_1\lambda_1, n_2\lambda_2) = \sum_{y_2=0}^{\infty} p(y_2 \mid n_2\lambda_2) P(R_{y_2}) = \sum_{y_1=a}^{\infty} p(y_1 \mid n_1\lambda_1) P(R^*_{y_1}).$$

Then we can derive that

$$\frac{\partial}{\partial \lambda_2} P_S(n_1\lambda_1, n_2\lambda_2) = n_2 \sum_{y_2=0}^{\infty} p(y_2 \mid n_2\lambda_2) \left\{ P(R_{y_2+1}) - P(R_{y_2}) \right\}$$

which is always negative from (6). On the other hand,

$$\frac{\partial}{\partial \lambda_1} P_S(n_1\lambda_1, n_2\lambda_2) = \begin{cases} n_1 \displaystyle\sum_{y_1=0}^{\infty} p(y_1 \mid n_1\lambda_1) \left\{ P(R^*_{y_1+1}) - P(R^*_{y_1}) \right\}, & \text{if } a = 0 \\ n_1 \left(p(a-1|n_1\lambda_1) P(R^*_a) + \displaystyle\sum_{y_1=a}^{\infty} p(y_1 \mid n_1\lambda_1) \left\{ P(R^*_{y_1+1}) - P(R^*_{y_1}) \right\} \right), & \text{if } a \geq 1 \end{cases}$$

which is always positive from (6). Hence the maximum of the p-value cannot occur at any inner point of a compact subset in the parameter space.

A.2. Proof of Corollary 2

It suffices to show that, the two test statistics satisfy the strict convexity condition. Note that, at $y_1 = 0, y_2 = 0$, the two statistics are defined as 0. In the following, we exclude this case from discussion. First, $Z_R(y_1, y_2)$ is investigated. At $y_2 = 0$, $Z_R(y_1, 0) = \sqrt{(n_2/n_1)y_1}$, which obviously is strictly increasing in y_1. For $y_2 > 0$, it is derived that

$$\frac{\partial}{\partial y_1} Z_R(y_1, y_2) = \frac{\frac{y_2}{n_1 n_2}}{2\left(\frac{y_1+y_2}{n_1+n_2}\right)\sqrt{\left(\frac{1}{n_1}+\frac{1}{n_2}\right)\left(\frac{y_1+y_2}{n_1+n_2}\right)}} > 0.$$

Hence, this statistic satisfies the convexity condition strictly along the first coordinate. Next, for $y_1 = 0$, $Z_R(0, y_2) = -\sqrt{(n_1/n_2)y_2}$, which is strictly decreasing in y_2. For $y_1 > 0$,

$$\frac{\partial}{\partial y_2} Z_R(y_1, y_2) = -\frac{\frac{y_1}{n_1 n_2}}{2\left(\frac{y_1+y_2}{n_1+n_2}\right)\sqrt{\left(\frac{1}{n_1}+\frac{1}{n_2}\right)\left(\frac{y_1+y_2}{n_1+n_2}\right)}} < 0.$$

Hence, this statistic satisfies the convexity condition strictly along the second coordinate.

On the other hand, similarly, we found that for $Z_U(y_1, y_2)$,

$$\frac{\partial}{\partial y_1} Z_U(y_1, y_2) = \frac{\frac{y_1}{n_1^3} + \frac{y_2}{n_1 n_2}\left(\frac{1}{n_1}+\frac{2}{n_2}\right)}{2\left(\frac{y_1}{n_1^2}+\frac{y_2}{n_2^2}\right)\sqrt{\frac{y_1}{n_1^2}+\frac{y_2}{n_2^2}}} > 0$$

and

$$\frac{\partial}{\partial y_2} Z_U(y_1, y_2) = -\frac{\frac{y_1}{n_1 n_2}\left(\frac{2}{n_1}+\frac{1}{n_2}\right)+\frac{y_2}{n_2^3}}{2\left(\frac{y_1}{n_1^2}+\frac{y_2}{n_2^2}\right)\sqrt{\frac{y_1}{n_1^2}+\frac{y_2}{n_2^2}}} < 0.$$

$Z_U(y_1, y_2)$ is strictly increasing in y_1 and strictly decreasing in y_2, hence it satisfies the convexity condition strictly.

Appendix B. Supplementary data

Supplementary material related to this article can be found online at http://dx.doi.org/10.1016/j.spl.2013.01.030.

References

Barnard, G.A., 1947. Significance tests for 2 × 2 tables. Biometrika 34, 123–138.
Berger, R.L., Boos, D.D., 1994. *P* values maximized over a confidence set for the nuisance parameter. Journal of the American Statistical Association 89, 1012–1016.
Casella, G., Berger, R.L., 1990. Statistical Inference. Wadsworth, Pacific Grove, CA.
Chan, I.S.F., Wang, W.W.B., 2009. On analysis of the difference of two exposure-adjusted poisson rates with stratification: from asymptotic to exact approaches. Statistics in Bioscience 1, 65–79.
Krishnamoorthy, K., Thomson, J., 2004. A more powerful test for two poisson means. Journal of Statistical Planning and Inference 119, 23–35.
Ng, H.K.T., Gu, K., Tang, M.L., 2007. A comparative study of tests for the difference of two poisson means. Computational Statistics & Data Analysis 51, 3085–3099.
Ng, H.K., Tang, M.L., 2005. Testing the equality of two poisson means using the rate ratio. Satistics in Medicine 24, 955–965.
Röhmel, J., Mansmann, U., 1999. Unconditional non-asymptotic one-sided tests for independent binomial proportions when the interest lies in showing non-inferiority and/or superiority. Biometrical Journal 41, 149–170.
Wang, H., 2008. Confidence intervals for the mean of a normal distribution with restricted parameter space. Journal of Statistical Computation and Simulation 78, 829–841.
Wang, L., Feng, Z., Wang, X., Wang, X., Zhang, X., 2010. DEGseq: an R package for identifying differentially expressed genes from RNA-seq data. Bioinformatics 26, 136–138.