行政院國家科學委員會專題研究計畫 成果報告

兩個卜瓦松分佈的統計檢定方法
研究成果報告(精簡版)

計 畫 類 別 ：個別型
計 畫 編 號 ：NSC 99-2118-M-004-005-
執 行 期 間 ：99 年 08 月 01 日至 100 年 12 月 31 日
執 行 單 位 ：國立政治大學統計學系

計畫主持人：薛慧敏

計畫參與人員：碩士班研究生-兼任助理人員：卓達瑋
博士班研究生-兼任助理人員：許嫚荏

報 告 附 件 ：出席國際會議研究心得報告及發表論文

公 開 資 訊 ：本計畫涉及專利或其他智慧財產權，2 年後可公開查詢

中 華 民 國 101 年 03 月 29 日

中文摘要： 卜瓦松分佈被普遍運用在許多流行病學、工業品質管制問題上。Ng 與 Tang(2005)曾
介紹過一個實際範例，研究人員欲比較兩個群體中某個特定疾病的發生率，則研究人
員以卜瓦松分布來描述群體中疾病發生個數的分布狀態，而研究目的在比較此兩個群
體的平均發病率。已知當樣本數或平均發生個數夠大時，則統計人員得運用常態分布
近似在相關的漸近統計檢定方法的發展上。但上述條件不滿足時，則應運用確實統計
檢定方法(exact testing procedure)以獲得適當的統計結論。本研究的目的為在比較兩個
卜瓦松分佈的問題中，發展確實統計檢定方法。此時虛無假設參數空間將牽涉干擾參
數，而干擾參數的出現將大幅增加確實統計檢定方法的計算複雜度。我們將研究相關
的理論以減少檢定方法的計算量。我們將考慮受限最大概似估計方法在檢定統計量
上，另一方面，由 Berger 與 Boo (1994)提出的信賴域 p-value 也將被研究，其中將牽涉
在受限參數空間中的信賴域。我們將研究所提出的統計檢定方法的有效性性質，並將
透過電腦模擬來驗證這些方法。

中文關鍵詞： 卜瓦松分佈，確實統計檢定，干擾參數，漸進統計檢定，顯著 p 值。

英文摘要： The Poisson distribution is suitable in variety fields such as
biology, quality control, and so on. For example, to study the
incidence rate of some disease (Ng and Tang, 2005), the number of
disease occurrence during some trial period can be modeled by a
Poisson distribution. This study aims to identify a superiority of
one treatment group over the control group under Poisson
distributions. We consider two commonly-seen Wald ‘s test statistics.
We introduce the correspondent asymptotic p-values

for large
sample sizes or long duration. On the other hand, for small data
sets we propose two types of exact p-values, which own
computational efficiency and are build under conservatism. The
validity of the proposed p-values are theoretically justified. A
numerical study is conducted for their finite-sample performances.
The application of one real example is provided for illustration.

英文關鍵詞： Poisson distribution, exact test, nuisance parameter, asymptotic test, p-value.

Testing the Superiority under Poisson Distribution

Mingte Liu   Huey-Miin Hsueh

March 29, 2012

Abstract

The Poisson distribution is suitable in variety fields such as biology, quality control, and so on. For example, to study the incidence rate of some disease (Ng and Tang, 2005), the number of disease occurrence during some trial period can be modeled by a Poisson distribution. This study aims to identify a superiority of one treatment group over the control group under Poisson distributions. We consider two commonly-seen Wald's test statistics. We introduce the correspondent asymptotic p-values for large sample sizes or long duration. On the other hand, for small data sets we propose two types of exact p-values, which own computational efficiency and are build under conservatism. The validity of the proposed p-values are theoretically justified. A numerical study is conducted for their finite-sample performances. The application of one real example is provided for illustration.

1 Introduction

It is well known that the Poisson distribution is a suitable model for rare events in a variety of fields, such as biology, commerce, quality control, and so on. These applications are usually used to compare the means of two Poisson random variables. As the sample sizes or the mean parameters are sufficiently large, asymptotic tests are often recommended in the literature under Poisson distribution. (See Shiue and Bain(1982), Thode(1997), Ng and Tang(2005) and Gu *el al.*(2008)). On the other hand, the exact testing procedures are more appropriate when both assumptions fail. In comparing two Poisson means, there involves nuisance parameters, the unknown mean values, under the null hypothesis. So far, the existing literatures consider the problem with the null hypothesis of equality of the two means to ease the complexity. Here in this study, our testing procedures are constructed under the null hypothesis of non-superiority. The test statistics under consideration are two commonly-seen Wald's test statistics. Their correspondent asymptotic p-values and two exact p-values are proposed. Theoretical justification on the validity of these p-values are provided. In addition, a numerical study is performed for their finite-sample properties. Further, a real example data set is presented for illustrations.

2 Test Procedure

2.1 Notations and Test Statistics

Assume two independent Poisson random samples within a fixed duration,

$$Y_{1i} \overset{iid}{\sim} Poi(\lambda_1), \;\; Y_{2j} \overset{iid}{\sim} Poi(\lambda_2), \text{ for } i = 1 \cdots n_1, \; j = 1 \cdots n_2.$$

It's 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. The research interest is to test the superiority of the alternative procedure with the following hypotheses,

$$H_0 : \lambda_1 \leq \lambda_2, \;\; H_1 : \lambda_1 > \lambda_2.$$

We consider two types of Wald's statistic,

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

where $\tilde{\lambda}_0 = \frac{Y_1+Y_2}{n_1+n_2}$ is the restricted MLE of the common mean value under the boundary of H_0, $\lambda_1 = \lambda_2$. Then the null hypothesis is rejected if a sufficiently large value of the test statistic is observed.

2.2 Asymptotic P-values

When n_1, n_2 are sufficiently large, the asymptotic p-values are

$$p_{A,R} = 1 - \Phi(z_R), \;\; p_{A,U} = 1 - \Phi(z_U),$$

where $\Phi(\cdot)$ is the distribution function of $N(0,1)$. Denote $\delta = \lambda_1 - \lambda_2$ as the mean difference. Assume as $n_1, n_2 \to \infty$, $n_1/n_2 \to \rho > 0$. Then at $\delta = \delta_0 \in R$, it's easy to derive that approximately,

$$Z_R\sigma - \mu \xrightarrow{d} N(0,1), \;\; \text{and} \;\; Z_U - \mu \xrightarrow{d} N(0,1),$$

with

$$\mu = \frac{\delta_0}{\sqrt{\frac{(1+\rho)\lambda_2+\delta_0}{n_2\rho}}}, \quad \sigma = \sqrt{\frac{(1+\rho)\lambda_2 + \rho\delta_0}{(1+\rho)\lambda_2 + \delta_0}}.$$

Furthermore, at the significance level α, the asymptotic power functions of the two asymptotic tests are respectively,

$$\bar{\beta}_{A,R}(\delta_0, \lambda_2, \rho, n_2) = 1 - \Phi(z_\alpha \sigma - \mu), \text{ and } \bar{\beta}_{A,U}(\delta_0, \lambda_2, \rho, n_2) = 1 - \Phi(z_\alpha - \mu).$$

Theorem 1. *At any* $\delta_0 \leq 0$,

1. $\bar{\beta}_{A,U} \leq \alpha$, *for all* n_1, n_2.
2. $\bar{\beta}_{A,R} \leq \alpha$, *for* n_2, ρ *satisfy*

$$\sqrt{n_2 \rho} \geq \left(\frac{z_\alpha}{(-\delta_0)} \right) \left(\sqrt{(1+\rho)\lambda_2 + \delta_0} - \sqrt{(1+\rho)\lambda_2 + \rho\delta_0} \right). \quad (1)$$

2.3 Exact P-values

When the sample sizes are insufficient or the mean values are relatively small, exact tests are more appropriate for establishing the superiority. Given the observed z_R, z_U of the two test statistics, the exact p-values are

$$p_{E,R} = P(Z_R \geq z_R), \quad p_{E,U} = P(Z_U \geq z_U),$$

where the probabilities are evaluated under two independent Poisson distributions with means in the composite null parameter space $\lambda_1 \leq \lambda_2$. To control the size, one can consider the standard p-value,

$$p_{S,R}^{(\gamma)} = \sup_{\lambda_1 \leq \lambda_2} P(Z_R \geq z_R), \quad p_{S,U}^{(\gamma)} = \sup_{\lambda_1 \leq \lambda_2} P(Z_U \geq z_U).$$

However, because the null space is infinite, the computation of a standard exact p-value is a complicated and time-consuming task. We aim to develop an efficient testing procedure in this study.

Consider the following confidence-set p-values by Berger and Boos (1994),

$$p_{CI,R}^{(\gamma)} = \sup_{(\lambda_1,\lambda_2) \in C_\gamma} P(Z_R \geq z_R) + \gamma, \quad p_{CI,U}^{(\gamma)} = \sup_{(\lambda_1,\lambda_2) \in C_\gamma} P(Z_U \geq z_U) + \gamma,$$

where C_γ is a joint confidence set of (λ_1, λ_2) that guarantees $100(1-\gamma)\%$ confidence within the null parameter space.

$$C_\gamma = \{L_1 \leq \lambda_1 \leq \min(U_1, \lambda_2), \quad L_2 \leq \lambda_2 \leq U_2\},$$

where

$$(L_1, U_1) = \frac{1}{2n_1} \left(\chi^2_{[1-(1-\sqrt{1-\gamma})/2,\, 2y_1]},\ \chi^2_{[(1-\sqrt{1-\gamma})/2,\, 2(y_1+1)]} \right),$$

and

$$(L_2, U_2) = \frac{1}{2n_2}\left(\chi^2_{[1-(1-\sqrt{1-\gamma})/2,\ 2y_2]},\ \chi^2_{[(1-\sqrt{1-\gamma})/2,\ 2(y_2+1)]}\right),$$

and $\chi^2_{(\delta,k)}$ is the $100(1-\delta)$-th percentile of a chi-square distribution with degree of freedom k.

Theorem 2. *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 convexity condition. Then given s_0, the supremum of $P(S \geq s_1)$ occurs at a boundary point of the parameter space.*

Theorem 3. *Z_R, Z_U satisfy the convexity condition.*

Subsequently, the confidence-set p-values of Z_R and Z_U can be evaluated in the boundary of the confidence set C_γ. That is,

$$p^{(\gamma)}_{CI,R} = \sup_{(\lambda_1,\lambda_2)\in\partial C_\gamma} P(Z_R \geq z_R) + \gamma, \quad p^{(\gamma)}_{CI,U} = \sup_{(\lambda_1,\lambda_2)\in\partial C_\gamma} P(Z_U \geq z_U) + \gamma,$$

where ∂C_γ is the boundary of C_γ.

Alternatively, we propose the estimated exact p-values as

$$\tilde{p}_{E,R} = P(Z_R \geq z_R|\tilde{\lambda}_{01}, \tilde{\lambda}_{02}), \quad \tilde{p}_{E,U} = P(Z_U \geq z_U|\tilde{\lambda}_{01}, \tilde{\lambda}_{02}), \tag{2}$$

where the probabilities are computed on

$$(\tilde{\lambda}_{01}, \tilde{\lambda}_{02}) = \begin{cases} (\hat{\lambda}_1, \hat{\lambda}_2), & \text{if } \hat{\lambda}_1 \leq \hat{\lambda}_2; \\ (\tilde{\lambda}_0, \tilde{\lambda}_0), & \text{if } \hat{\lambda}_1 > \hat{\lambda}_2. \end{cases}$$

3 Numerical Study

In this numerical study, we investigate the performance of the asymptotic tests, denoted as p_A, by using Z_R, Z_U, respectively. The correspondent exact tests by the confidence-set p-values, denoted as p^{γ}_{CI}, and the estimated p-value in (2), denoted as $\tilde{p}_E$, are studied as well. The confidence-set p-value is constructed with two different confidence limits, $\gamma = 0.001$. That is, the smaller γ is mount to expand the rejection range of the procedure. Because the two Wald's test statistic are a function of the sufficient statistics, the exact type I error rate and power of their associated tests can be easily computed. Here, the exact type I error rate and the exact power of each test are calculated. We consider $\lambda_2 = 1$ and δ_0 ranged from -0.25 to 2.0 and the second sample size $n_2 = 10$ and three $\rho = 3/5,\ 1$. The nominal significance level α is set as 0.05. The type I error rate and power are presented in Table 1-2.

Table 1: Type I error rate of the asymptotic p-value and the exact p-value of Z_R, Z_U at $\lambda_2 = 1, n_2 = 10$ and $\alpha = 5\%$.

Test			δ_0				
ρ	Statistic	p-value	-0.25	-0.15	-0.1	-0.05	0.0
3/5	Z_R	p_A	0.0157	0.0266	0.0337	0.0421	0.0519
		$p_{CI}^{(0.001)}$	0.0096	0.0176	0.0231	0.0297	0.0375
		$\tilde{p}_E$	0.0137	0.0233	0.0297	0.0372	0.0460
	Z_U	p_A	0.0082	0.0153	0.0202	0.0262	0.0334
		$p_{CI}^{(0.001)}$	0.0129	0.0228	0.0293	0.0370	0.0461
		$\tilde{p}_E$	0.0145	0.0250	0.0318	0.0399	0.0493
1	Z_R	p_A	0.0126	0.0230	0.0301	0.0387	0.0489
		$p_{CI}^{(0.001)}$	0.0123	0.0227	0.0298	0.0384	0.0487
		$\tilde{p}_E$	0.0123	0.0227	0.0298	0.0384	0.0487
	Z_U	p_A	0.0126	0.0230	0.0301	0.0387	0.0489
		$p_{CI}^{(0.001)}$	0.0123	0.0227	0.0298	0.0384	0.0487
		$\tilde{p}_E$	0.0123	0.0227	0.0298	0.0384	0.0487

Table 2: Power of the asymptotic p-value and the exact p-value of Z_R, Z_U at $\lambda_2 = 1, n_2 = 10$ and $\alpha = 5\%$.

Test			δ_0				
ρ	Statistic	p-value	0.1	0.5	1.0	1.5	2.0
3/5	Z_R	p_A	0.0757	0.2298	0.5024	0.7432	0.8907
		$p_{CI}^{(0.001)}$	0.0574	0.1942	0.4524	0.6999	0.8655
		$\tilde{p}_E$	0.0675	0.2099	0.4728	0.7194	0.8781
	Z_U	p_A	0.0517	0.1833	0.4425	0.6942	0.8623
		$p_{CI}^{(0.001)}$	0.0682	0.2120	0.4743	0.7199	0.8782
		$\tilde{p}_E$	0.0721	0.2199	0.4871	0.7310	0.8841
1	Z_R	p_A	0.0748	0.2554	0.5773	0.8279	0.9477
		$p_{CI}^{(0.001)}$	0.0746	0.2544	0.5724	0.8223	0.9451
		$p_{CI}^{(0.005)}$	0.0646	0.2350	0.5554	0.8145	0.9422
		$\tilde{p}_E$	0.0747	0.2554	0.5773	0.8279	0.9477
	Z_U	p_A	0.0748	0.2554	0.5773	0.8279	0.9477
		$p_{CI}^{(0.001)}$	0.0746	0.2544	0.5724	0.8223	0.9451
		$\tilde{p}_E$	0.0747	0.2554	0.5773	0.8279	0.9477

4 Real Example

In this section, the methods are applied to the breast cancer study described in Ng and Tang (2005). 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 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,

$$H_0 : \lambda_1 \leq \lambda_2 \quad H_1 : \lambda_1 > \lambda_2.$$

See the following table for the p-values. All these p-values 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 statistical significance.

Table 3: The asymptotic, estimated and confidence-set p-values of Z_R and Z_U.

Test statistic	$z_R = 2.0818$	$z_U = 2.2047$
Asymptotic p-value	0.0187	0.0137
Estimated p-value	0.0177	0.0186
Confidence-set p-value ($\gamma = 0.001$)	0.0182	0.0188

5 Discussion

In this study, we focus on the problem of testing the superiority of one Poisson distribution over another from the difference in means. Differing with the problem with the equality null hypothesis, the current problem involves a more complicated null parameter space, and it introduces more difficulties on both theoretical justification and practical computation of a p-value. A test, that ignores the null space of inferiority, is likely to produce a liberal conclusion and may not adequately control its type I error rate. We here propose several asymptotic and exact testing procedures for large and small to moderate data sets respectively. Theoretical justifications on validity of these testing procedures are provided. In additional, a numerical study is conducted for their finite-sample performances.

References

[1] Gu, K., Ng, H. K., Tang, M. L. and Schucany, W. R. (2008) Testing the Ratio of Two Pisson Rates. *Biometrical Journal*, **50**, 283-298.

[2] Ng, H. K. and Tang, M. L.(2005) Testing The Equality of Two Poisson Means Using The Rate Ratio. *Satistics in Medicine*, **24**, 955-965.

[3] Shiue, W. and Bain, L. J.(1982) Experiment Size and Power Comparisons for Two-Sample Poisson Tests. *Applied Statistics*, **31**, 130-134.

[4] Thode, H. C.(1997) Power and Sample Size Requirements for Tests of Differences Between Two Poisson Rates. *The Statistican*, **46**, 227-230.

國科會補助專題研究計畫項下出席國際學術會議心得報告

日期：100 年 8 月 23 日

計畫編號	NSC 99－2118－M－004－005－		
計畫名稱	兩個卜瓦松分佈的統計檢定方法		
出國人員姓名	薛慧敏	服務機構及職稱	政治大學統計系
會議時間	100 年 7 月 30 日 至 100 年 8 月 4 日	會議地點	美國邁阿密市
會議名稱	(中文)北美共同統計會議 (英文)2011 Joint Statistical Meetings		
發表論文題目	(中文)基因集的顯著性檢定 (英文)Assessing the Significance of a Gene Set		

一、參加會議經過

第一天為報到程序。之後五天則為參加研討會議。每一天有四個場次，各有多個演講廳同時進行會議。另外在每天的上午與下午，在安排的場地上皆有海報發表。本人在 8/2 上午發表論文，在其他時間則至各演講廳，聆聽與會學者的文章發表，學習目前最新學術發展，並適時參與討論。

二、與會心得

本屆大會有約六千人與會參加。有來自統計學門與資訊工程學門專家、學者踴躍參與。此會議提供很好的機會讓不同領域的學者能夠在統計計算與計算統計上的理論、方法與實際運用上交流與分享。本人在此次會議上有豐富收穫。

三、考察參觀活動(無是項活動者略)

無。

四、建議

近年來，由於經濟因素，至歐美參加會議之旅費與生活費逐年高漲，國科會補助通常不敷使用，建議能因應客觀環境，適當提高補助經費，或增加計畫經費流用的彈性，才能提高參加會議意願。

五、攜回資料名稱及內容

包括書面資料一冊與光碟片一份。其中書面資料為會議議程，光碟片則為發表文章摘要。

六、其他

無。

國科會補助計畫衍生研發成果推廣資料表

日期:2012/03/29

國科會補助計畫	計畫名稱: 兩個卜瓦松分佈的統計檢定方法
	計畫主持人: 薛慧敏
	計畫編號: 99-2118-M-004-005- 學門領域: 生物統計

無研發成果推廣資料

99 年度專題研究計畫研究成果彙整表

計畫主持人：薛慧敏			計畫編號：99-2118-M-004-005-				
計畫名稱：兩個卜瓦松分佈的統計檢定方法							
成果項目			量化			單位	備註（質化說明：如數個計畫共同成果、成果列為該期刊之封面故事...等）
			實際已達成數（被接受或已發表）	預期總達成數(含實際已達成數)	本計畫實際貢獻百分比		
國內	論文著作	期刊論文	0	0	100%	篇	
		研究報告/技術報告	0	0	100%		
		研討會論文	0	0	100%		
		專書	0	0	100%		
	專利	申請中件數	0	0	100%	件	
		已獲得件數	0	0	100%		
	技術移轉	件數	0	0	100%	件	
		權利金	0	0	100%	千元	
	參與計畫人力（本國籍）	碩士生	0	0	100%	人次	
		博士生	0	0	100%		
		博士後研究員	0	0	100%		
		專任助理	0	0	100%		
國外	論文著作	期刊論文	0	1	100%	篇	
		研究報告/技術報告	0	0	100%		
		研討會論文	1	1	100%		
		專書	0	0	100%	章/本	
	專利	申請中件數	0	0	100%	件	
		已獲得件數	0	0	100%		
	技術移轉	件數	0	0	100%	件	
		權利金	0	0	100%	千元	
	參與計畫人力（外國籍）	碩士生	1	2	100%	人次	
		博士生	1	0	100%		
		博士後研究員	0	0	100%		
		專任助理	0	0	100%		

其他成果 (無法以量化表達之成果如辦理學術活動、獲得獎項、重要國際合作、研究成果國際影響力及其他協助產業技術發展之具體效益事項等，請以文字敘述填列。)	無。

	成果項目	量化	名稱或內容性質簡述
科教處計畫加填項目	測驗工具(含質性與量性)	0	
	課程/模組	0	
	電腦及網路系統或工具	0	
	教材	0	
	舉辦之活動/競賽	0	
	研討會/工作坊	0	
	電子報、網站	0	
	計畫成果推廣之參與（閱聽）人數	0	

國科會補助專題研究計畫成果報告自評表

請就研究內容與原計畫相符程度、達成預期目標情況、研究成果之學術或應用價值（簡要敘述成果所代表之意義、價值、影響或進一步發展之可能性）、是否適合在學術期刊發表或申請專利、主要發現或其他有關價值等，作一綜合評估。

1. 請就研究內容與原計畫相符程度、達成預期目標情況作一綜合評估

■達成目標

□未達成目標（請說明，以 100 字為限）

□實驗失敗

□因故實驗中斷

□其他原因

說明：

2. 研究成果在學術期刊發表或申請專利等情形：

論文：□已發表 □未發表之文稿 ■撰寫中 □無

專利：□已獲得 □申請中 ■無

技轉：□已技轉 □洽談中 ■無

其他：（以 100 字為限）

3. 請依學術成就、技術創新、社會影響等方面，評估研究成果之學術或應用價值（簡要敘述成果所代表之意義、價值、影響或進一步發展之可能性）（以 500 字為限）

學術研究：對於福從卜瓦松分佈的母體提供適當統計檢定方法。

應用方面：可應用於公衛資料的統計分析上。