The simple crossover design, in which patients are randomly assigned to one of two groups with a different sequence of treatment-receipt, has often been employed to study the effect of a treatment on non-curable chronic diseases, including angina pectoris, epilepsy, hypertension or asthma.

Because each patient serves as his/her own control, use of the crossover design can be more efficient than the parallel groups design, especially when the response of interest varies substantially between patients. However, the crossover design would not be suitable for use if the studied treatment had a long-lasting effect. This is because the effect of the treatment administered at the latter period can be confounded with the residual effect of the treatment administered at the earlier period. Therefore, we should apply the crossover design to only those treatments with relatively short acting or no residual effects. In practice, we almost always employ a wash-out period between administering two treatments to nullify any residual effect of the treatment administered first. An excellent review of various aspects in use of the crossover design appears elsewhere. Most of these research focused attention on continuous data based on the normality assumptions or binary data based on a random effects logistic risk model.

The discussions on analysis of count data, such as the number of seizures in a study of epilepsy or the frequency of exacerbations in a trial of asthma, are few. Note that a distribution-free approach, which concentrates on interval estimation of the relative treatment effect, has been recently proposed for count data under a simple crossover design. However, these asymptotic distribution-free interval estimators derived on the basis of large sample theory are invalid for use when the number of patients in a trial is small (as one of the main cases considered here) under the assumed Poisson distribution. The exact test and exact interval estimator derived under the Poisson distribution assumption can be of use in these small-sample cases. Furthermore, when the frequency of occurrence follows the Poisson distribution and the number of patients per group is not large, asymptotic interval estimators (or tests) developed elsewhere without accounting for the Poisson assumption may lose accuracy or efficiency. Also, to the best of our knowledge we find that the discussions on hypothesis testing in count data have been so far truly limited.

The goal of this article is to provide a systematic discussion on testing non-equality, non-inferiority and equivalence, as well as interval estimation of the relative treatment effect when the frequency of patient response follows the commonly-assumed Poisson distribution under a crossover design. We present asymptotic and exact procedures for testing non-equality, non-inferiority and equivalence, as well as develop asymptotic and exact interval estimators for the ratio of mean frequencies between two treatments. We employ Monte Carlo simulation to evaluate and compare the performance of these test procedures with respect to Type I error and power, as well as the performance of these interval estimators with respect to the coverage probability and average length. Finally, we use the data comparing salmeterol with a placebo based on the number of exacerbations of asthma published elsewhere to illustrate the practical use of test procedures and interval estimators discussed here.

交叉設計實驗中，患者會被隨機分配到不同治療順序對應的組別，此設計經常被用來研究無法治癒的慢性疾病的治療效應，比如心絞痛、癲癇、高血壓或氣喘之類的疾病。由於每個患者都是自己的對照組，所以，在某些狀況下使用交叉設計會比平行設計來的更有效率，特別是當患者自身的反應變數差異很大的時候。

但若使用的治療方式有持久(long-lasting)的效果，則交叉設計將不適用於此，因為後期的治療效果可能會和早期治療產生的殘留效應產生混淆(confounding).

因此，交叉設計應用於治療效果作用時間較短或沒有殘留效應的治療方式。

在實務上，交叉設計會在兩次治療之間有一個清洗時間(washout period)，讓受試者於這期間將前一期的藥效代謝掉，以消除早期治療的殘留效應。

過去的文獻大多專注於連續型資料與二元資料的探討，連續型資料多用常態假設，二元型態的資料多使用隨機效應的logistic risk model去分析。

然而，當患者數量少的時候，這些基於大樣本理論推導出的無漸進分配的區間估計量在卜瓦松分配可能會失去準確性或效率，甚至是無效的。

本文的目的是對交叉設計之下患者的反應頻率服從卜瓦松分佈時的不相等性、非劣性與等校性檢驗以及相對治療效果的區間估計進行討論，作者提出了檢驗不相等性、非劣性與等效性的漸進檢驗程序與精確檢驗程序，並為平均頻率之比開發了漸進與精確的區間估計量。

最後使用蒙地卡羅模擬來評估和比較這些測試程度在型一錯誤和檢定力的表現，以及區間估計量在覆蓋率與平均長度的表現。最後，作者將本文提出的方法應用於氣喘患者使用沙美特羅與安慰劑下的發作頻率資料並進行說明。

An Example

Fieller’s Theorem

In large samples, asymptotic normality takes over and symmetric confidence intervals are appropriate. It’s not difficult to see that these variances are identical to the first-order approximate obtained from the corresponding linear models by the delta method.

But for small samples, the skewness of distribution may give rise to asymmetry of the exact confidence limits.

In large samples, the confidence interval given by the delta method is approximately the same as that obtained from Fieller's theorem. If g is small, then the large-sample confidence limits will be close to those given by Fieller's theorem. When the value of g is not small, however, the confidence limits given by Fieller's theorem are in fact usually wider than those obtained from the large-sample standard deviation, and are not centered on the estimate.

**euqivalence**

Estimated Type I error

Estimated Power

**Non-inferiority**

Estimated Type I error

Estimated Power

**Interval Estimator**

**Estimated Coverage probability(for measuring the accuracy)**

**Average length(for measuring the precision)**

Generate 10,000 repeated samples

Table1 type I error

Table2 power

Table3 type I error

Table4 power

**Table5**

**Coverage probability**

**Confidence interval**