Superiority or Non-inferiority Testing Procedures for Two Independent Poisson Samples

Liu, Mingte [1]   Hsueh, Huey-miin.[2]

July 27, 2010

[1] General Education Center, Tatung Institute of Commerce and Technology.
[2] Department of Statistics, National Cheng-Chi University.

List of Tables

2.1 The type I error rate ($\delta_0 = 0$) of asymptotic p-value test (p_A) and exact p-value test (p_{CI}, p_E) based on T, Z_R, Z_U respectively for $n_2 = 10$. 21

2.2 The type I error rate ($\delta_0 = 1$) of asymptotic p-value test (p_A) and exact p-value test (p_{CI}, p_E) based on T, Z_R, Z_U respectively for $n_2 = 10$. 22

2.3 The type I error rate ($\delta_0 = 0$) of asymptotic p-value test (p_A) and exact p-value test (p_{CI}, p_E) based on T, Z_R, Z_U respectively for $n_2 = 30$. 23

2.4 The type I error rate ($\delta_0 = 1$) of asymptotic p-value test (p_A) and exact p-value test (p_{CI}, p_E) based on T, Z_R, Z_U respectively for $n_2 = 30$. 24

2.5 To achieve 80% power at $\delta_0 = 0.6$, the required sample size of the second group n_2^* of Z_R, Z_U, T for $\rho = 3/5$. Based on the required samples n_2^*, the power and the type I error rate (in parentheses) are given. 25

2.6 To achieve 80% power at $\delta_0 = 0.6$, the required sample size of the second group n_2^* of Z_R, Z_U, T for $\rho = 1$. Based on the required samples n_2^*, the power and the type I error rate (in parentheses) are given. 26

2.7 To achieve 80% power at $\delta_0 = 0.6$, the required sample size of the second group n_2^* of Z_R, Z_U, T for $\rho = 5/3$. Based on the required samples n_2^*, the power and the type I error rate (in parentheses) are given. 27
2.8 To achieve 80% power at $\delta_0 = 1$, the required sample size of the second group n_2^* of Z_R, Z_U, T for $\rho = 3/5$. Based on the required samples n_2^*, the power and the type I error rate (in parentheses) are given. 28
2.9 To achieve 80% power at $\delta_0 = 1$, the required sample size of the second group n_2^* of Z_R, Z_U, T for $\rho = 1$. Based on the required samples n_2^*, the power and the type I error rate (in parentheses) are given. 29
2.10 To achieve 80% power at $\delta_0 = 1$, the required sample size of the second group n_2^* of Z_R, Z_U, T for $\rho = 5/3$. Based on the required samples n_2^*, the power and the type I error rate (in parentheses) are given. 30

3.1 Type I error rate and power of asymptotic p-value and exact p-value at $\lambda_2 = 1, n_2 = 10$, these p-values are based on test statistics T, Z_R, Z_U respectively. 44
3.2 Type I error rate and power of asymptotic p-value and exact p-value at $\lambda_2 = 2, n_2 = 10$, these p-values are based on test statistics T, Z_R, Z_U respectively. 45

4.1 Type I error rate and power of asymptotic p-value and exact p-value at $\lambda_2 = 1, n_2 = 10$, these p-values are based on test statistics Z_{R^*}, Z_{U^*} respectively. 62
4.2 Type I error rate and power of asymptotic p-value and exact p-value at $\lambda_2 = 2, n_2 = 10$, these p-values are based on test statistics Z_{R^*}, Z_{U^*} respectively. 63

4.3 To achieve 80% power at $\delta_0^* = 0.6, \rho = 3/5$, the required sample size of the second group n_2 of the asymptotic p-values and exact p-value which are conducted at Z_{R^*}, Z_{U^*}. Based on the required samples n_2, the power and the type I error rate (in parentheses) are given at various δ_0^* in Ω_{03} 64

4.4 To achieve 80% power at $\delta_0^* = 0.6, \rho = 1$, the required sample size of the second group n_2 of the asymptotic p-values and exact p-value which are conducted at Z_{R^*}, Z_{U^*}. Based on the required samples n_2, the power and the type I error rate (in parentheses) are given at various δ_0^* in Ω_{03}. 65

4.5 To achieve 80% power at $\delta_0^* = 0.6, \rho = 5/3$, the required sample size of the second group n_2 of the asymptotic p-values and exact p-value which are conducted at Z_{R^*}, Z_{U^*}. Based on the required samples n_2, the power and the type I error rate (in parentheses) are given at various δ_0^* in Ω_{03}. 66

4.6 To achieve 80% power at $\delta_0^* = 1.0, \rho = 3/5$, the required sample size of the second group n_2 of the asymptotic p-values and exact p-value which are conducted at Z_{R^*}, Z_{U^*}. Based on the required samples n_2, the power and the type I error rate (in parentheses) are given at various δ_0^* in Ω_{03}. 67

4.7 To achieve 80% power at $\delta_0^* = 1.0, \rho = 1$, the required sample size of the second group n_2 of the asymptotic p-values and exact p-value which are conducted at Z_{R^*}, Z_{U^*}. Based on the required samples n_2, the power and the type I error rate (in parentheses) are given at various δ_0^* in Ω_{03}. 68

4.8 To achieve 80% power at $\delta_0^* = 1.0, \rho = 5/3$, the required sample size of the second group n_2 of the asymptotic p-values and exact p-value which are conducted at Z_{R^*}, Z_{U^*}. Based on the required samples n_2, the power and the type I error rate (in parentheses) are given at various δ_0^* in Ω_{03}. 69

5.1 The asymptotic, estimated and confidence-set p-value of the Wald Z-test Z_R, Z_U. 76

List of Figures

2.1 The joint parameter space Ω is all, the null parameter space Ω_{01} for testing the equality, and the null parameter space Ω_{02} for testing superiority, the null parameter space Ω_{03} for testing non-inferiority. 10

2.2 As $n_2 = 10, \lambda_2 = 0.03, \rho = 8, 20, 50$, the asymptotic power of Z_R over $\delta_0 \in (0, 0.1)$. 31

2.3 As $n_2 = 10, \lambda_2 = 0.3, \rho = 3/5, 1, 5/3$, the asymptotic powers of the Z_R(the dotted and dashed line) and Z_U(the solid line) over $\delta_0 \in (0, 1)$. 32

3.1 The asymptotic power function of the Z_R (dotted and dashed line) and the Z_U (dashed line) when $n_2 = 5, \lambda_2 = 0.3, \delta_0 = -0.3 : 0.05 : 0, \rho = 3/5, 1, 5/3$ in the left panel, $\rho = 18, 25, 30$ in the right panel. 46

4.1 As $n_2 = 2, \lambda_2 = 0.2, \Delta_0 = 0.2\lambda_2, \rho = 0.2, 0.5, 0.8, 1.2, 1.4, 1.6, \delta_0 = -0.16 : 0.001 : 0$, the asymptotic type I error rate of the Z_{R^*}(solid line). 70

4.2 As $n_2 = 2, 7, \lambda_2 = 0.2, \Delta_0 = 0.2\lambda_2, \rho = 1.7, 3, 5, \delta_0^* = -0.16 : 0.001 : 0$, the asymptotic type I error rate of the Z_{R^*}(solid line). 70

4.3 As $n_2 = 2, \lambda_2 = 0.02, \Delta_0 = 0.2\lambda_2, \rho = 0.2, 0.4, 0.6, 0.8, 1, 1.2, \delta_0^* = 0 : 0.001 : 0.05$, the asymptotic power of the Z_{R^*}(solid line). . . 71

4.4 As $n_2 = 2, 7, \lambda_2 = 0.02, \Delta_0 = 0.2\lambda_2, \rho = 1.3, 1.6, 2, \delta_0^* = 0 : 0.001 : 0.05$, the asymptotic power of the Z_{R^*}(solid line). . . . 71

4.5 As $n_2 = 2, \lambda_2 = 100, 200, \Delta_0 = 0.2\lambda_2, \rho = 0.5, 5, 50, \delta_0^* = 0 : 1 : 10$, the asymptotic power of the Z_{R^*}(solid line). 72
4.6 As $n_2 = 10, \Delta_0 = 2, \rho = 0.6$, a contour map of $Z_{U^*} = 2, 3, 4, 5, 6, 7, 8, 9, 10$. 72
4.7 As $n_2 = 10; \Delta_0 = 0.2; \rho = 0.6$, a contour map of $Z_{U^*} = k$ for $k = 2, 3, 4$. 73
4.8 As $n_2 = 10; \Delta_0 = 2; \rho = 1$, a contour map of $Z_{U^*} = k$ for $k = 2, 3, 4, 5, 6, 7, 8, 9, 10$. 73
4.9 As $n_2 = 10; \Delta_0 = 2; \rho = 5/3$, a some contour map of $Z_{U^*} = k$ for $k = 2, 3, 4, 5, 6, 7, 8, 9, 10$. 74

Notation

$(Y_{11}, \cdots, Y_{1n_1})$	Independent Poisson random samples.
$(Y_{21}, \cdots, Y_{2n_2})$	Independent Poisson random samples.
$Y_i, \bar{Y}_i$	Sample sum and sample mean of group i.
ρ	$\frac{n_1}{n_2}$
λ_i	True mean rate of group i.
Ω	Full parameter space in Poisson.
Ω_{01}	Null parameter space of the null hypothesis of equality.
Ω_{02}	Null parameter space of the null hypothesis of non-superiority.
Ω_{03}	Null parameter space of the null hypothesis of inferiority.
δ	The difference between the true mean rate of group 1 and group 2.
$\hat{\delta}$	Maximum likelihood estimator of δ under Ω.
$se(\delta)$	Asymptotic standard error of δ.
Z	Wald statistic.
Z_R	Wald statistic with constrained MLE of asymptotic standard error.
Z_U	Wald statistic with unconstrained MLE of asymptotic standard error.
$\tilde{\lambda}_0$	Restricted maximum likelihood estimator under $\lambda_1 = \lambda_2$.
T	Two-independent-sample random variable.
$p_{A,(\cdot)}$	Asymptotic p-value based on $(\cdot)$.
$\mu,\ \sigma$	Mean and standard error of asymptotic distribution of Z_R.

Notation

$\bar{\beta}_{(\cdot)}$	Asymptotic power function of $(\cdot)$.
$Z_{\cdot,c}$	Z_R, Z_U Continuity corrected.
C_γ	$100(1-\gamma)\%$ confidence interval of γ under Ω_{01}.
$poi(\cdot,\nu)$	Probability of Poisson distribution with mean ν.
$p^{(\gamma)}_{CI,\cdot}$	Confidence interval$p$-value based on $Z_\cdot$.
$p_{E,R}$	Estimated p-value based on Z_R.
$p_{E,U}$	Estimated p-value based on Z_U.
C^*_γ	$100(1-\gamma)\%$ confidence interval of γ under Ω_{02}.
$C_{\gamma,0}$	$100(1-\gamma)\%$ cross product of λ_1, λ_2 under Ω.
(L_i, U_i)	Independent $100\sqrt{(1-\gamma)}\%$ confidence interval of λ_1, λ_2 under Ω respectively.
$\tilde{\lambda}_{0i}$	Restricted maximum likelihood estimator of λ_i on Ω_{02}.
Δ_0	Non-inferiority limit.
Z_{i^*}	Wald test statistic with the unconstricted estimator of the standard error under Ω_{03}.
$\tilde{\lambda}_i$	Restricted maximum likelihood estimator of λ_i with respect to $\lambda_1 - \lambda_2 + \Delta_0 = 0$.
μ^*, σ^*	Asymptotic mean and standard error of Z_{R^*}.
$\bar{\beta}_{Z_{i^*}}$	Asymptotic power function of Z_{i^*}.
n_{2,Z_i}	The minimum sample size of the second group required for Z_i at significance level α.
$n_{2,Z_{i^*}}$	The minimum sample size of the second group required for Z_{i^*} at significance level α.
$p^{(\gamma)}_{CI,Z_{i^*}}$	Confidence interval p-value based on Z_{i^*}.
$p_{E,Z_{i^*}}$	Estimated p-value based on Z_{i^*}.
C^{**}_γ	$100(1-\gamma)\%$ confidence interval of γ under Ω_{03}.
$\tilde{\lambda}_{i3}$	Some estimator of λ_i under the restricted null parameter space Ω_{03}.

Abstract

The Poisson distribution is a well-known suitable model for modeling a rare events in variety fields such as biology, commerce, quality control, and so on. Many applications involve comparisons of two treatment groups and focus on showing the superiority of the new treatment to the conventional one, or the non-inferiority of the experimental implement to the standard implement upon the cost consideration. We aim to develop statistical tests for testing the superiority and non-inferiority by two independent random samples from Poisson distributions. In developing these tests, both computational and theoretical difficulties arise from presence of nuisance parameters. In this study, we first consider the problems with the null hypothesis of equality for simplicity. The problems are extended to have a regular null hypothesis of non-superiority next. Subsequently, the proposed methods are further investigated in establishing the non-inferiority.

Two types of Wald test statistics are of our main research interest. The correspondent asymptotic testing procedures are developed by using the normal limiting distribution. In our study, the asymptotic distribution of the test statistics are derived. The asymptotic power functions and the sample size formula are further obtained. Given the power functions, we justify the validity and unbiasedness of the tests. The adequate continuity correction term for these tests is also found to reduce inflation of the type I error rate. On the other hand, the exact testing procedures based on two exact p-values, the confidence-interval p-value (Berger and Boos (1994)), and the estimated p-value (Krishnamoorthy and Thomson (2004)), are also applied in our study. It is known that an exact testing procedure tends to involve complex computations. In this thesis, several strategies are proposed to lessen the computational burden. For the confidence-interval p-value, a truncated confidence set is used to narrow the area for finding the p-value. Further, the test statistic is verify whether they fulfill the property of convexity. It is shown that under the convexity the exact p-value occurs somewhere of the boundary of the null parameter space. On the other hand, for the estimated p-value, a simpler point estimate is applied instead of the use of the restricted maximum likelihood estimators, which are less straightforward in this problem. The estimated p-value is shown to provide a conservative conclusion. The calculations of the sample sizes required by using the two exact tests are discussed.

Intensive numerical studies show that the performances of the asymptotic tests depend on the fraction of the two sample sizes and the continuity correction can be useful in some cases to reduce the inflation of the type I error rate. However, with small samples, the two exact tests are more adequate in the sense of having a well-controlled type I error rate. A data set of breast cancer patients is analyzed by the proposed methods for illustration.

keywords: *A*symptotic test, Barnard convexity condition, exact test, non-inferiority, Poisson, p-value, restricted maximum likelihood estimator(RMLE), superiority, unbiasedness, validity.

Chapter 1

Introduction

1.1 Motivation

It is well known that the Poisson distribution is a suitable model for rare events in variety fields such as biology, commerce, quality control, and so on. Those applications are usually used to compare two population means, and some practical examples have been illustrated in literature. For example, to compare the rate of breast cancer of the group with/without X-ray fluoroscopy examination during treatment for tuberculosis, the equality of the mean numbers of cases in a given person-years at risk of the two groups are tested (Ng and Tang (2005)). Another example investigates whether the failure rate of the new component is less than the current one in planes (Shiue and Bain, 1982). Sometimes, a severe conclusion may be unnecessary as adopting some consideration. For instance, in air filter system one wants to know whether the experimental air filter is not inferior than the standard one, when the former one is relatively cheaper (Lui, 2005). Actually, these comparison can be described by statistical hypothesis in terms of either the difference of the two Poisson means or their ratio. Here, the comparison is considered in terms of difference of the two Poisson means.

Gail (1974) introduced two different experiments. In the first experiment, the total number of the two Poisson variables is predetermined. In the other experiment, the length of experiment duration is fixed instead. The exact test based on the conditional distribution given the fixed total number, which was proposed by Przyborowski and Wilenski in 1940, is an adequate testing method in the former experiment. This test is uniformly most powerful among unbiased tests. In the later experiment, which is more common in practice, an unconditional test is more suitable. When the sample sizes or the mean parameters are large, a normal approximation is considered for the unconditional test to lessen the computation.

Sometimes the experiment durations of the two Poisson variables are unequal. For example, one is interested in the comparison of failure rate of an airplane component between war time and peace time. The simulating condition of war time is more expensive than that of peace time, see Shiue and Bain (1982). The authors generalized the conditional exact test and a normal approximated test to the unequal interval cases. An approximation formula of the experiment length required to achieve a specified power is also proposed and is shown to be useful through an empirical study. Thode (1997) provided an alternative normal approximated test and showed that the new test is more powerful than the test proposed by Shiue and Bain (1982) when the mean rate is large for a lengthy experiment. Basically, these proposed methods were developed in terms of the difference of the two Poisson means in literatures. Alternative, some authors expressed the comparison in terms of the ratio of the two positive means, see Ng and Tang (2005), Gu *et al.* (2008). Ng and Tang (2005) tested the unity of the mean ratio. They compared two normal approximated tests, which apply the logarithmic-transformed rate ratio in the numerator of the test statistic, and adopt two different estimations for the standard error. They found that two specific test statistics perform well, especially when the means values are large. Gu *et al*. (2008) extended the numerical comparisons to more tests. However, all the existing the procedures were studied and compared through numerical studies in most literatures. In this paper, we consider a comparison between two independent Poisson random samples with a fixed experiment duration. When the sample sizes are unbalanced, the scenario is equivalent to the unequal duration case.

In application of Poisson model, testing the non-inferiority is an important problem as well when the endpoint is count data. For instance, in a medical study one aims to justify that the efficiency of an experimental drug is non-inferior to some control drug with a given non-inferiority margin(Song, 2009). Lui (2005) studied the calculation of the sample sizes required and power by exact tests for testing non-inferiority. The author further derived the formulae of calculation of sample sizes and power by large sample theory, in which a test statistic involves a logarithmic-transformation was proposed. Corinna and Jochen (2005) studied the calculation of sample sizes and power by the likelihood ratio test, the score test, and the exact conditional test, in which the power calculations were illustrated graphically. These authors express the hypothesis of non-inferiority in terms of the ratio of two group means. Here, we will develop statistical tests in testing the non-inferiority in terms of the difference of two group means.

This study investigates two types of testing method: asymptotic, and exact tests. The first aim is to investigate the performance of the two types Wald test. The validity and unbiasedness of the two tests will be studied in Poisson problem. The asymptotic power and sample size formula of the two tests will be derived, too. Further, the test will be compared with the two-independent-sample T-test. Which is originally proposed for testing two normal population means with an unknown, equal variance. To improve a mild inflation of type I error rate, we modify the three tests by adding some continuity correction term. Pirie and Hamdan (1972) derived a continuity correction term when the two Poisson random samples are of equal size. In this paper, adequate continuity correction term for general cases will also be derived. There are two important theoretical properties for a testing procedure: Validity and unbiasedness. Given a test statistic, the correspondent p-value can be found and it shows the strength of evidence to reject the null hypothesis. The statistical conclusion can be drawn based on the p-value. Berger and Boos (1994) called a p-value valid if it satisfies $P_\theta(p \leq \alpha) \leq \alpha$, for each $\alpha \in [0, 1]$, for all θ in null parameter space. On the other hand, a p-value is called unbiased if $P_\theta(p \leq \alpha) \geq \alpha$, for every θ over the alternative parameter space (Lehmann, 1986). So far, the proposed tests of this problem in literatures are rarely justified for these theoretical properties. In this study, the asymptotic testing procedures will be explored whether they satisfy the validity and unbiasedness.

When the sample sizes are small or the mean parameter are insufficiently large, the uses of an asymptotic test is inadequate. The exact methods based on the exact sampling distribution of the test statistic will be proposed. In the problem of comparing two Poisson means, nuisance parameters present in the sampling distribution. Casella and Berger (1990) define the standard p-value that considers the least favorable case under the principle of conservativeness. However, the standard p-value is less powerful and tends to be unnecessarily over-conservative by not taking the data information into consideration. Moreover, the computation becomes complex and inefficient when the null parameter space is an infinite set. Berger and Boos (1994) showed that the p-value constructed as the maximum over a confidence region of the nuisance parameters is valid. The associated confidence-set p-value has been shown to be valid and will be considered here. Although the extent of searching the maximum has been reduced, intensive calculations are necessary to find out the maximum. Röhmel and Mansmann (1999) showed that in a binomial problem, once the test statistic satisfies the Barnard convexity condition, the supremum of the p-value occurs at the boundaries and the calculations of confidence-set p-value can be hence greatly reduced. In the study, we will generalize previous result to Poisson problems. Two types of Wald test will be examined whether they satisfy the convexity condition or not. Hence, more efficient confidence-set p-values will be obtained.

On the other hand, Krishnamoorthy and Thomson (2004) inspired by Storer and Kim (1990) developed a nearly exact testing methods. The associated p-value is exact because it is evaluated under Poisson distribution. The authors use an point estimate of the nuisance parameter in calculation of the exact p-value. The same test was studied in Gu *et al*. (2008). Although the estimated p-value was shown to perform well and can control its exact type I error rate below the nominal level in selected settings in these papers. However, this testing procedure could not guarantee a well-controlled type I error rate theoretically. Here, the estimated p-value proposed by Krishnamoorthy and Thomson (2004) will be adapted. However, the restricted estimation will be modified for handy applications.

Basically, the content of the null parameter space determines the complexity of computation of a p-value. In this study, we are interested in testing superiority and non-inferiority. These associated null parameter space are infinite regions in concluding diagonal line or others in the first quadrant. Then, the calculation of searching p-value is quite complicated. In next chapter, we first consider the null hypothesis of equality for simplicity. The investigations will be extended to a conventional superiority in Chapter 3. The validity and the power of the proposed testing methods will be derived theoretically. Intensive numerical studies will be provided as well. Subsequently, these proposed testing procedure will further be applied to testing non-inferiority in Chapter 4. Similarly, the validity and unbiasedness of these testing procedure will be explored, and the performances between them will be compared.

1.2 Outline

This articles is organized as follows. In Chapter 2, we will focus on testing the null hypothesis of equality. We will give the asymptotic properties and the sample size formula of two types Wald test and T-test in Section 2.2. Adequate continuity correction terms will be derived. In Section 2.3, several exact testing procedures will be introduced. Subsequently, numerical studies will be presented in Section 2.4. The power and the type I error rate of the proposed tests will be compared. In Chapter 3, the problem will be extended to testing superiority. Further we will study the validity of the asymptotic tests and exact tests proposed in Chapter 2. More issues on the exact tests will be discussed. Similarly, some numerical study will be given. In Chapter 4, two types Wald test statistic will be redefined at the null hypothesis of testing non-inferiority. There are two asymptotic tests and exact tests based on this two test statistics are explored. Similarly, the validity and unbiasedness of two testing procedure will be examined and the correspondent sample size formulae will be derived, respectively. In Chapter 5, our proposed methods will be applied on a real example of breast cancer. Last, a brief conclusion will be presented. In this study, all numerical studies are conducted by MATLAB software and C^{++} language.

Chapter 2

Testing the null hypothesis of equality

Assume two independent Poisson random samples within a fixed duration, $(Y_{11}, \cdots, Y_{1n_1})$, $(Y_{21}, \cdots, Y_{2n_2})$,

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

where $Poi(\cdot)$ is a Poisson distribution with the mean rate $(\cdot)$. Then, the full parameter space is the first quadrant on $\mathcal{R}^2$,

$$\Omega = \{(\lambda_1, \lambda_2) | \lambda_1 > 0, \lambda_2 > 0\}.$$

This study mainly focuses on three types of one-sided hypothesis testing problems on comparing the two Poisson distributions. The first two problems are the so-called superiority tests, while the third one is the non-inferiority test. An essential difference between these problems is the extent of the associated null parameter space, which determines the complexity of the problem as explained in Chapter 1. See Figure 2.1 for the plots of the three null parameter spaces. In this chapter, for simplicity, we consider the null hypothesis of equality. The associated null parameter space includes only the diagonal ($\Omega_{01} = \{0 < \lambda_1 = \lambda_2\}$ in Figure 2.1). Next chapter, the test of superiority will be explored. The correspondent null space be extended to Ω_{02} which is the region above and including the diagonal line. Subsequently, the problem of testing non-inferiority correspondent to the null space Ω_{03} will be studied.

2.1 Statistical Hypothesis and Test Statistics

If prior knowledge indicates the equality of the two population, the statistical hypothesis can be expressed as follows,

$$H_{01} : \lambda_1 = \lambda_2, \quad \text{vs.} \quad H_1 : \lambda_1 > \lambda_2.$$

It's seen that $Y_1 = \sum_{i=1}^{n_1} Y_{1i},\ Y_2 = \sum_{j=1}^{n_2} Y_{2j}$ are sufficient statistics, and the maximum likelihood estimator(MLE) of $\delta = \lambda_1 - \lambda_2$ can be derived as $\hat{\delta} = \bar{Y}_1 - \bar{Y}_2$, where $\bar{Y}_1, \bar{Y}_2$ are the MLE of λ_1 and λ_2 under Ω, respectively.

Dividing the MLE $\hat{\delta}$ by its estimated asymptotic standard error $se(\hat{\delta})$, one obtains the Wald's test statistic,

$$Z = \frac{\hat{\delta}}{se(\hat{\delta})},$$

where $se(\hat{\delta})$ is obtained by plugging some consistent estimators of λ_1, λ_2 in the standard error of $\hat{\delta}$. In general, two common estimators are employed, one with constrained MLE is

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

where $\tilde{\lambda}_0 = \frac{Y_1+Y_2}{n_1+n_2}$ is RMLE(restricted maximum likelihood estimator)under

$H_{01} : \lambda_1 = \lambda_2 = \lambda$. The other one with unconstrained MLE is

$$Z_U = \frac{\bar{Y}_1 - \bar{Y}_2}{\sqrt{\frac{\bar{Y}_1}{n_1} + \frac{\bar{Y}_2}{n_2}}}.$$

On the other hand, when testing the equality of two normal means, the two-independent sample T-test is commonly used. We will study the applicability of this test in the comparison of Poisson means. Let S_1^2, S_2^2 be the sample variances of the two random samples, respectively. The two-independent-sample T statistic is

$$T = \frac{\bar{Y}_1 - \bar{Y}_2}{S_p\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}, \text{ where } S_p^2 = \frac{(n_1 - 1)S_1^2 + (n_2 - 1)S_2^2}{n_1 + n_2 - 2}$$

is the pooled sample variance. The null hypothesis H_{01} is rejected if a sufficiently large value of Z or T is observed.

The asymptotic p-values of the two Wald's tests can be computed straightforward under normality, while the asymptotic p-value of the T-test is found under a t-distribution with degrees of freedom $(n_1 + n_2 - 2)$. The theoretical performance of the asymptotic power function of the three p-values will be studied in next section.

2.2 Asymptotic p-values

In the following, the asymptotic p-values of the observed z_R, z_U, t_0 are

$$p_{A,R} = 1 - \Phi(z_U), \quad p_{A,U} = 1 - \Phi(z_R), \quad p_T = 1 - t_{(n_1+n_2-2)}(t_0)$$

where $\Phi(\cdot)$ is the distribution function of $N(0, 1)$, and $t_\nu(\cdot)$ is the t-distribution with degrees of freedom ν . The null hypothesis is rejected if the p-value is not greater than the significance level α. In the following, we will explore the

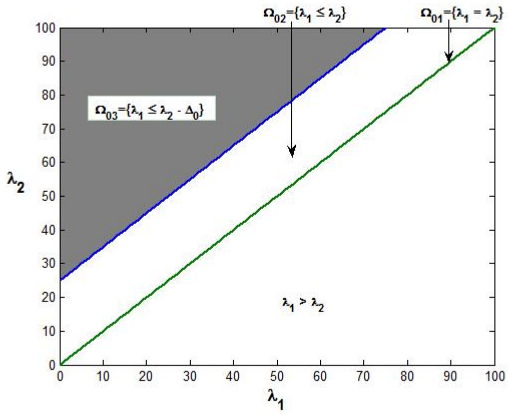

Figure 2.1: The joint parameter space Ω is all, the null parameter space Ω_{01} for testing the equality, and the null parameter space Ω_{02} for testing superiority, the null parameter space Ω_{03} for testing non-inferiority.

validity and asymptotic power function of the three asymptotic tests, and deriving formula of required sample sizes for these tests.

Theorem 1. Let δ_0 be the true value of δ, and $\rho = n_1/n_2 \in (0,1)$ be the sample size fraction of the first group to the second group. As $n_1,\ n_2 \to \infty$,

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

In which,

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

At significance level α, H_{01} is rejected if the test statistic exceeds z_α, where z_α is the $100(1-\alpha)\%$-th percentile of $N(0,1)$. Then the asymptotic power functions of Z_R, Z_U can be found as follows,

$$\bar{\beta}_{Z_R}(\delta_0,\lambda_2,n_2,\rho) = 1-\Phi(z_\alpha\sigma-\mu), \quad \bar{\beta}_{Z_U}(\delta_0,\lambda_2,n_2,\rho) = 1-\Phi\left(z_\alpha-\mu\right).$$

Under H_{01}, $\delta_0 = 0$, then $\mu = 0, \sigma = 1$, and further $\bar{\beta}_{Z_U} = \bar{\beta}_{Z_R} = \alpha$. That is, both the two asymptotic tests successfully control their type I error rate at the significance level. The correspondent p-values are called asymptotic valid.

When $\delta_0 > 0$, $\mu > 0$, the asymptotic power $\bar{\beta}_{Z_U}$ can be shown always greater than α. It indicates that the testing procedure Z_U is an unbiased test approximately. Nevertheless, the unbiasedness of Z_R is not always true. When the first group has a smaller size than the second group, i.e. $\rho \leq 1$, $\sigma \leq 1$, the asymptotic power $\bar{\beta}_{Z_R}$ is always above the nominal level α and increases as δ_0. On the contrary, if $\rho > 1$, the power may not exceed the nominal level. In the following we explore the behavior of the asymptotic power $\bar{\beta}_{Z_R}$ at some extreme λ_2 as $\delta_0 > 0, \rho > 1$. As λ_2 approaches to infinity,

$$\mu = \frac{\delta_0}{\sqrt{\frac{\lambda_2(1+\rho)+\delta_0}{n_2\rho}}} \to 0, \quad \sigma = \sqrt{\frac{\lambda_2(1+\rho)+\rho\delta_0}{\lambda_2(1+\rho)+\delta_0}} \to 1.$$

Then the asymptotic power of Z_R converges to the level α. As $\lambda_2 \to 0$,

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

Hence,

$$\lim_{\lambda_2 \to 0} \bar{\beta}_{Z_R} = 1 - \Phi\left(z_\alpha\sqrt{\rho} - \sqrt{n_2\rho\delta_0}\right). \tag{2.1}$$

In this case, one can see that $\bar{\beta}_{Z_R}$ increases as δ_0 increases. However it's easy to derive that the power is less than α when

$$\delta_0 < \left\{\frac{z_\alpha(\sqrt{\rho}-1)}{\sqrt{n_2\rho}}\right\}^2.$$

Hence, Z_R tends to be biased when the sample sizes are extremely unbalanced and the means of group are relatively small, i.e. $\rho >> 1, \lambda_1 \approx 0, \lambda_2 \approx 0$. See Figure 2.2 for the plots of the asymptotic power function of Z_R for

$\rho = 8, 20, 50, \lambda_2 = 0.03$ and $n_2 = 10$. In summary, Z_R is not always an unbiased test for $\rho > 1$.

In the next theorem, the asymptotic distribution of T is shown the same as that of Z_R in this Poisson problem. It's known that the mean and the variance coincide in a Poisson population. Hence the two test statistics use a sample estimate of standard error of $\hat{\delta}$ in the denominator under a common constraint.

Theorem 2. Let δ_0 be the true value of δ, and $\rho = n_1/n_2$ be the sample size fraction of the first group to the second group. As $n_1,\ n_2 \to \infty$,

$$T\sigma - \mu \xrightarrow{d} N(0,1).$$

When $n_1,\ n_2$ are sufficiently large, the critical value of the T test approximates to that of the Wald test, $t_{(n_1+n_2-2,\alpha)} \approx z_\alpha$. In addition, from Theorem 2, T and Z_R have the same asymptotic distribution. Consequently, the asymptotic power of T can be derived to be equal to the power of Z_R,

$$\bar{\beta}_T(\delta_0, \lambda_2, \rho, n_2) = \bar{\beta}_{Z_R} = 1 - \Phi\left(z_\alpha\sigma - \mu\right).$$

Hence T has the same performance as Z_R approximately. According to the discussion in previous paragraphs, T is a valid test, and is unbiased as $\rho \leq 1$. As $\rho > 1$, T is not necessarily unbiased.

Based on the power function of a testing procedure, the necessary sample size for achievement of a prespecified power at some alternative setting at significance level can be further determined. Given ρ, to achieve a prespecified power level $1 - \beta_0$ at $\lambda_2, \delta_0 > 0$, the minimal sample size of the second group required for the Z_U and Z_R at significant level α is given as

$$n^*_{2,Z_R} \geq \left\{ \frac{z_\alpha \sigma + z_{\beta_0}}{\delta_0} \right\}^2 \left\{ \frac{\lambda_2(1+\rho)+\delta_0}{\rho} \right\}, \tag{2.2}$$

and

$$n^*_{2,Z_U} \geq \left\{ \frac{z_\alpha + z_{\beta_0}}{\delta_0} \right\}^2 \left\{ \frac{\lambda_2(1+\rho)+\delta_0}{\rho} \right\}, \tag{2.3}$$

respectively. The size of the first group is found as $n^*_1 = [n^*_2 \cdot \rho] + 1$, in which $[a] = q$, the q is the maximum integer less than or equal to a. The formulae of sample sizes for T is equivalent to the equation (2.2).

It can be seen that the powers and sample size formulae of the three tests mainly differ in the multiple of z_α, σ. When $\rho = 1$, the sample sizes are balanced, $\sigma = 1$ and the three tests are equivalent in terms of the power function and the sample size formula. When $\delta_0 = 0$, all $\bar{\beta}_{Z_R} = \bar{\beta}_T = \bar{\beta}_{Z_U} = \alpha$. When $\delta_0 > 0$, we discover that $\bar{\beta}_{Z_U} < \bar{\beta}_{Z_R} = \bar{\beta}_T$ if $\rho < 1$, $\bar{\beta}_{Z_U} > \bar{\beta}_{Z_R} = \bar{\beta}_T$, if $\rho > 1$. See Figure 2.3. It indicates that the $Z_R - /T$-tests are more powerful and required less observations for a specified power than the Z_U-test when there are less observations in the first group. The result is opposite when the samples size of the first group is more than that of the second group. Hence, when the sampling cost for a subject from the first group is more expensive than from the second group, one may consider a study of $\rho < 1$, and the use of Z_R or T is suggested.

In this study, the sampling fraction $\rho \in (0, \infty)$ is considered a fixed constant exactly or approximately. It requires that the two group sizes n_1, n_2 have the same converging rates. Otherwise, as both sizes converge to infinity, the statistic correspondent to the larger sample converges to a constant faster than others. The subsequent asymptotic distribution of the testing statistic becomes trivial and is less worthy to derive. On the other hand, in the design stage, the sampling fraction ρ should be specified a priori for sample size determination. In practice, the information, as well as λ_2, δ_0, are obtained after a consultation with experts of the related field and after taking consideration of a realistic situation on applications.

When testing a parameter of a discrete distribution, a continuity correction is often added in the test statistic when one applies an approximation by some continuous distribution. The continuity correction revised by Pirie and Hamdan (1972) is employed in the Poisson problem. It's known that given an unbiased and sufficient estimator $\hat{\delta}$ for δ, the continuity corrected test statistic is

$$\frac{\hat{\delta} - \frac{1}{2}b}{se(\hat{\delta})},$$

provided that the support of $\hat{\delta}$ has equal spacings with space b.

Pirie and Hamdan (1972) indicated that for two independent Poisson random samples, the MLE $\hat{\delta}$ has equal spacings if one of $n_1,\ n_2$ is an integer multiple of the other. Specifically, when $n_1 = n_2, b = 1$. In the following theorem, we extend the results of Pirie and Hamdan (1972) to any $n_1,\ n_2$.

Theorem 3. For any $n_1,\ n_2$, the sampling distribution of $\hat{\delta}$ has equal spacings with space

$$b = \frac{1}{2m},$$

where m is the least common multiple of $n_1,\ n_2$.

Consequently, the continuity-corrected two Wald's test and T-test are defined as

$$Z_c = \frac{\hat{\delta} - \frac{1}{2m}}{se(\hat{\delta})},\ \ T_c = \frac{\hat{\delta} - \frac{1}{2m}}{s_p\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}},$$

respectively. In which Z_c can be either $Z_{R,c}$ or $Z_{U,c}$.

2.3 Exact p-values

When the sample sizes are insufficient or the mean values are relatively small, exact testing procedures are more adequate than asymptotic ones. Given a realization of a test statistic, an exact p-value is defined and calculated under the exact null distribution. In many applications, the null distribution often involves an unknown nuisance parameter(s). Here, both the Wald statistics Z_U, Z_R are functions of the sufficient statistics (Y_1, Y_2). Under the null hypothesis, $H_{01} : \lambda_1 = \lambda_2 = \lambda > 0$, Y_1, Y_2 independently follow a Poisson distribution with mean $n_1\lambda, n_2\lambda$, respectively. Given an observed z_0 of the Wald statistic Z, where Z can be either Z_U or Z_R, an exact p-value is defined under the true null distribution, which involves the unknown common mean value λ,

$$p_\lambda(z_0) = P(Z \geq z_0 | \lambda_1 = \lambda_2 = \lambda) = \sum_{y_1 \geq 0} \sum_{y_2 \geq 0} poi(y_1, n_1\lambda) poi(y_2, n_2\lambda) I_{\{Z \geq z_0\}}, \tag{2.4}$$

where $poi(y, \lambda')$ is the probability function of Poisson distribution with mean λ' and I is the indicator function. The common λ is a nuisance parameter. In the following, several testing procedures to deal with unknown nuisance parameters in literature are reviewed.

Casella and Berger (1990) defined the following standard p-value that considers the most conservative scenario and guarantees the validity,

$$p_s = \sup_{\lambda_1, \lambda_2 \in \Omega_{01}} P(Z \geq z_0 | \lambda_1 = \lambda_2 = \lambda),$$

where $\Omega_{01} = \{(\lambda_1,\ \lambda_2) :\ \lambda_1 = \lambda_2 > 0\}$ is the null parameter space of H_{01}. Ω_{01} is unbounded in a Poisson problem, hence the computation of the standard p-value is difficult in real-world applications. In addition, not taking the data information into consideration, one may obtain an unnecessarily conservative conclusion.

To ease the computational burden brought by searching the supremum over an infinite interval, Berger and Boos (1994) proposed a confidence-set p-value and showed that it is valid. The confidence-set p-value is the supremum over a confidence-set of the nuisance parameter. Here, given an observation z_R of Z_R, the confidence-set p-value is defined as

$$p^{(\gamma)}_{CI,R} = \sup_{\lambda \in C_\gamma} P(Z_R \geq z_R \mid \lambda_1 = \lambda_2 = \lambda) + \gamma, \tag{2.5}$$

where C_γ is a $100(1-\gamma)\%$ confidence interval for the nuisance parameter λ. On the other hand, given z_U, the confidence-set p-value based on Z_U is

$$p^{(\gamma)}_{CI,U} = \sup_{\lambda \in C_\gamma} P(Z_U \geq z_U \mid \lambda_1 = \lambda_2 = \lambda) + \gamma. \tag{2.6}$$

In which, γ is a positive real number and is far less than α for a non-trivial conclusion. In this study, we consider the following $100(1-\gamma)\%$ exact confidence interval C_γ of λ,

$$\frac{1}{2(n_1+n_2)}(\chi^2_{(1-\gamma/2,\, 2(Y_1+Y_2))},\ \chi^2_{(\gamma/2,\, 2(Y_1+Y_2+1))}),$$

where $\chi^2_{\alpha,v}$ is the $100(1-\alpha)$-th percentile of a chi-square distribution with degrees of freedom v (Casella and Berger, 1990). The confidence interval is based on the following equivalent relationship between Poisson and Chi-square random variables,

$$\begin{aligned} \frac{\gamma}{2} &= P(Y \leq y_0) = P(\chi^2_{2(y_0+1)} > 2(n_1+n_2)\lambda), \\ \frac{\gamma}{2} &= P(Y \geq y_0) = P(\chi^2_{2y_0} < 2(n_1+n_2)\lambda), \end{aligned}$$

where Y follows $Poi((n_1+n_2)\lambda)$, $\chi^2_{2(\cdot)}$ is a random variable with degrees of freedom $2(\cdot)$.

Krishnamoorthy and Thomson (2004) proposed an alternative exact p-value by using the RMLE $\tilde{\lambda}_0$ of the nuisance parameter λ. That is, given z_R, z_U, the estimated p-value are defined as

$$p_{E,R} = P(Z_R \geq z_R \mid \tilde{\lambda}_0), \;\; p_{E,U} = P(Z_U \geq z_U \mid \tilde{\lambda}_0),$$

respectively. The estimated p-value has great reduction in computation and performs well empirically. Although the estimator owns many pleasant properties in the inference of point estimation under H_{01}, but the resultant p-value does not guarantee a valid test theoretically.

As the Wald statistic depends on the data only through the two sufficient statistics (Y_1, Y_2), the exact power of the test correspondent to the p-value, p, is given by

$$\sum_{y_1 \geq 0} \sum_{y_2 \geq 0} poi(y_1, n_1\lambda_1) poi(y_2, n_2\lambda_2) I_{\{p \leq \alpha\}}.$$

Given a predetermined power level $1 - \beta_0$ at some specific λ_2 and $\delta_0 > 0$, the required sample size of the second group is the smallest integers such that the exact power achieves the level, and it is found as follows

$$n_2^* = \min\{n_2 : \sum_{y_1 \geq 0} \sum_{y_2 \geq 0} poi(y_1, ([n_2\rho] + 1)(\lambda_2 + \delta_0)) poi(y_2, n_2\lambda_2) I_{\{p \leq \alpha\}} \geq 1 - \beta_0\}, \tag{2.7}$$

for some $\rho > 0$. Further $n_1^* = [n_2^*\rho] + 1$.

2.4 Numerical study

In this section, we investigate the performance of the two test statistics Z_R, Z_U, as well as T. The asymptotic testing procedures by using the asymptotic $p-$values are considered. The effect of a continuity correction are explored in these asymptotic tests. Denote the p-value as p_A if it is without a continuity correction; as p_{A_c} if it is with a continuity correction term. The exact tests by using the confidence-set p-value, denoted as p_{CI}, and the estimated p-value, denoted as p_E, of Z_R, Z_U are further studied. As described in previous section, the calculation of the exact power is straightforward when the test statistic depends on the data only through the two sufficient statistics Y_1, Y_2. Here, except the T-test, the exact type I error rate and the exact power of each test are calculated. The power of the T-test is found through $100,000$ replicates. In this numerical analysis, we consider $\lambda_2 = 0.3, 0.4, 0.6, 1, 2, 3, n_2 = 10, 30, \delta_0 = 0,\ 1, \rho = 3/5,\ 1,\ 5/3$ and $\alpha = 0.05$. The calculated type I error rate and power are presented in Table 2.1-2.4. The required samples sizes of the second group to achieve $1 - \beta_0 = 80\%$ power at $\delta_0 = 0.6, 1$ are provided in Table 2.5-2.10.

We first compare the three asymptotic tests in Table 2.1 to 2.4. Although Z_R and T are found to have different results in the finite sample cases from the tables, we find that the two tests have quite consistent patterns. It justifies the theoretical results given in Section 2.2 that the two test statistics have the same asymptotic distributions. Theoretically, at $\delta_0 = 0$ the asymptotic sizes of the three tests are independent of ρ and equal to the nominal significance level α. However, the finite-sample results in Table 2.1 and Table 2.3 appear to be more consistent with the asymptotic power functions under the alternative hypothesis. When $\rho = 3/5 < 1$, Z_R and T have more chance to reject the null hypothesis than Z_U. The trend becomes the opposite when $\rho > 1$. Basically, the type I error rate of the three tests sometimes exceed the nominal level $\alpha = 5\%$. Although as the sample sizes increase, there are some improvement in the type I error rate, the differences are not obvious. When $\rho = 3/5$, the sizes of Z_R and T are not well-controlled at $\alpha = 5\%$, and T is more liberal than Z_R at small λ_2 and $n_2 = 10$. For $\rho = 5/3$, the inflation of the type I error rate of Z_U is even worse. For the three tests, at $\rho > 1, \rho < 1$, adding a continuity correction or increasing the sample size entail limited improvement. Overall speaking, Z_R and T is more robust to the choice of ρ than Z_U. Z_U is too liberal for $\rho > 1$ and is too conservative for $\rho < 1$.

Next the two exact p-values, p_{CI}, p_E, are studied. Note that in finding the confidence-set p-value, the supremum is searched over 16 grids of the confidence interval of the common mean value λ. Moreover, we consider $\gamma = 0.001$. Table 2.1 and 2.3 show that the two exact approaches have their type I error rate well-controlled. The confidence-interval p-value is more conservative than the estimated p-value. The computations involved are greatly reduced for the estimated p-value. One should keep in mind that the estimated p-value is not a valid test theoretically. Although in these selected scenarios of our simulation, its type I error rate does not exceed the nominal level. It is possible that the estimated p-value has an inflated type I error rate in other cases.

Table 2.5-2.10 present the required sample size of the second group for 80% power at $\delta_0 = 0.6, 1.0$. The results for the three asymptotic tests are based on the asymptotic sample size formulae (2.2) and (2.3). For the two exact tests, the figures are the minimal integers such that the exact power achieves the level by (2.7). All the tests need less required sample size of the second group for 80% power when the δ_0 increases. Between the three asymptotic tests, Z_U needs a slightly smaller sample than Z_R and T for $\rho > 1$. On the contrary, however, with the smaller sample size, the exact type I error rate of three asymptotic tests often exceeds the nominal level α. The inflation is more severe in the application of Z_U and showed limited improvement with the continuity correction.

Moreover, the sample sizes obtained for the two exact tests are near that of the asymptotic tests and the differences are within 3 units in all cases. With the calculated sample size, every exact test achieves the prespecified power level and has a well-controlled type I error rate. In summary, although the exact tests are more time-consuming, they guarantee more adequate statistical conclusions. The asymptotic sample sizes (2.2) and (2.3) can be regarded as an efficient alternative of (2.7) for the exact tests. A much quicker solution can be obtained and the result is found to be close to the exact sample size.

Table 2.1: The type I error rate ($\delta_0 = 0$) of asymptotic p-value test (p_A) and exact p-value test (p_{CI}, p_E) based on T, Z_R, Z_U respectively for $n_2 = 10$.

ρ	Test Statistic	p-value	λ_2					
			0.3	0.4	0.6	1	2	3
3/5	T	$p_{A,T}$	0.0660	0.0600	0.0544	0.0514	0.0525	0.0513
		$p_{A_c,T}$	0.0652	0.0582	0.0505	0.0486	0.0509	0.0497
	Z_R	$p_{A,R}$	0.0540	0.0528	0.0537	0.0519	0.0529	0.0524
		$p_{A_c,R}$	0.0540	0.0528	0.0537	0.0505	0.0498	0.0496
		$p_{CI,R}^{(\gamma=0.001)}$	0.0385	0.0412	0.0359	0.0375	0.0446	0.0483
		$p_{E,R}$	0.0406	0.0476	0.0502	0.0466	0.0493	0.0483
	Z_U	$p_{A,U}$	0.0219	0.0232	0.0266	0.0334	0.0410	0.0426
		$p_{A_c,U}$	0.0219	0.0232	0.0265	0.0308	0.0378	0.0403
		$p_{CI,U}^{(\gamma=0.001)}$	0.0385	0.0417	0.0402	0.0461	0.0482	0.0483
		$p_{E,U}$	0.0389	0.0433	0.0459	0.0487	0.0482	0.0483
1	T	$p_{A,T}$	0.0473	0.0455	0.0482	0.0509	0.0498	0.0475
		$p_{A_c,T}$	0.0302	0.0344	0.0383	0.0408	0.0427	0.0417
	Z_R	$p_{A,R}$	0.0497	0.0508	0.0515	0.0489	0.0496	0.0497
		$p_{A_c,R}$	0.0331	0.0358	0.0371	0.0396	0.0414	0.0437
		$p_{CI,R}^{(\gamma=0.001)}$	0.0448	0.0421	0.0454	0.0487	0.0475	0.0471
		$p_{E,R}$	0.0448	0.0421	0.0454	0.0487	0.0496	0.0497
	Z_U	$p_{A,U}$	0.0497	0.0508	0.0515	0.0489	0.0496	0.0497
		$p_{A_c,U}$	0.0331	0.0358	0.0371	0.0397	0.0414	0.0437
		$p_{CI,U}^{(\gamma=0.001)}$	0.0448	0.0421	0.0454	0.0487	0.0475	0.0471
		$p_{E,U}$	0.0448	0.0421	0.0454	0.0487	0.0496	0.0497
5/3	T	$p_{A,T}$	0.0417	0.0420	0.0444	0.0457	0.0472	0.0480
		$p_{A_c,T}$	0.0354	0.0403	0.0430	0.0439	0.0454	0.0465
	Z_R	$p_{A,R}$	0.0455	0.0474	0.0461	0.0484	0.0462	0.0467
		$p_{A_c,R}$	0.0369	0.0388	0.0434	0.0447	0.0457	0.0460
		$p_{CI,R}^{(\gamma=0.001)}$	0.0455	0.0474	0.0461	0.0482	0.0459	0.0467
		$p_{E,R}$	0.0455	0.0474	0.0461	0.0484	0.0470	0.0491
	Z_U	$p_{A,U}$	0.0799	0.0711	0.0644	0.0632	0.0563	0.0548
		$p_{A_c,U}$	0.0724	0.0697	0.0629	0.0589	0.0562	0.0543
		$p_{CI,U}^{(\gamma=0.001)}$	0.0353	0.0335	0.0324	0.0420	0.0457	0.0467
		$p_{E,U}$	0.0455	0.0474	0.0461	0.0484	0.0470	0.0491

Table 2.2: The type I error rate ($\delta_0 = 1$) of asymptotic p-value test (p_A) and exact p-value test (p_{CI}, p_E) based on T, Z_R, Z_U respectively for $n_2 = 10$.

ρ	Test Statistic	p-value	λ_2 0.3	0.4	0.6	1	2	3
3/5	T	$p_{A,T}$	0.7263	0.6709	0.5833	0.4693	0.3332	0.2638
		$p_{A_c,T}$	0.7148	0.6579	0.5733	0.4642	0.3273	0.2594
	Z_R	$p_{A,R}$	0.7576	0.7037	0.6129	0.5024	0.3516	0.2834
		$p_{A_c,R}$	0.7575	0.7034	0.6093	0.4873	0.3438	0.2723
		$p_{CI,R}^{(\gamma=0.001)}$	0.6841	0.6209	0.5469	0.4524	0.3341	0.2705
		$p_{E,R}$	0.7429	0.6816	0.5899	0.4848	0.3380	0.2705
	Z_U	$p_{A,U}$	0.6505	0.5996	0.5275	0.4425	0.3139	0.2497
		$p_{A_c,U}$	0.6500	0.5971	0.5162	0.4268	0.3005	0.2453
		$p_{CI,U}^{(\gamma=0.001)}$	0.7042	0.6560	0.5878	0.4743	0.3374	0.2705
		$p_{E,U}$	0.7282	0.6817	0.6013	0.4751	0.3374	0.2705
1	T	$p_{A,T}$	0.8068	0.7565	0.6714	0.5494	0.3887	0.3102
		$p_{A_c,T}$	0.7729	0.7206	0.6327	0.5136	0.3625	0.2901
	Z_R	$p_{A,R}$	0.8387	0.7872	0.6996	0.5773	0.4073	0.3274
		$p_{A_c,R}$	0.8044	0.7532	0.6641	0.5364	0.3818	0.3050
		$p_{CI,R}^{(\gamma=0.001)}$	0.8323	0.7847	0.6992	0.5724	0.3988	0.3193
		$p_{E,R}$	0.8323	0.7847	0.6994	0.5773	0.4073	0.3263
	Z_U	$p_{A,U}$	0.8387	0.7872	0.6996	0.5773	0.4073	0.3274
		$p_{A_c,U}$	0.8044	0.7532	0.6641	0.5364	0.3818	0.3050
		$p_{CI,U}^{(\gamma=0.001)}$	0.8323	0.7847	0.6992	0.5724	0.3988	0.3193
		$p_{E,U}$	0.8323	0.7847	0.6994	0.5773	0.4073	0.3263
5/3	T	$p_{A,T}$	0.8687	0.8212	0.7395	0.6142	0.4396	0.3458
		$p_{A_c,T}$	0.8633	0.8147	0.7328	0.6065	0.4332	0.3415
	Z_R	$p_{A,R}$	0.8948	0.8522	0.7733	0.6419	0.4524	0.3657
		$p_{A_c,R}$	0.8869	0.8422	0.7615	0.6275	0.4522	0.3589
		$p_{CI,R}^{(\gamma=0.001)}$	0.8948	0.8521	0.7709	0.6339	0.4524	0.3657
		$p_{E,R}$	0.8948	0.8523	0.7743	0.6499	0.4549	0.3729
	Z_U	$p_{A,U}$	0.9208	0.8832	0.8086	0.6749	0.4876	0.3917
		$p_{A_c,U}$	0.9140	0.8750	0.8002	0.6695	0.4875	0.3846
		$p_{CI,U}^{(\gamma=0.001)}$	0.8693	0.8313	0.7595	0.6275	0.4524	0.3656
		$p_{E,U}$	0.8948	0.8523	0.7743	0.6499	0.4549	0.3729

Table 2.3: The type I error rate ($\delta_0 = 0$) of asymptotic p-value test (p_A) and exact p-value test (p_{CI}, p_E) based on T, Z_R, Z_U respectively for $n_2 = 30$.

ρ	Test Statistic	p-value	λ_2 0.3	0.4	0.6	1	2	3
3/5	T	$p_{A,T}$	0.0529	0.0515	0.0526	0.0517	0.0517	0.0506
		$p_{A_c,T}$	0.0516	0.0504	0.0500	0.0497	0.0505	0.0496
	Z_R	$p_{A,R}$	0.0523	0.0525	0.0536	0.0524	0.0516	0.0510
		$p_{A_c,R}$	0.0516	0.0493	0.0501	0.0496	0.0499	0.0502
		$p_{CI,R}^{(\gamma=0.001)}$	0.0356	0.0405	0.0426	0.0483	0.0484	0.0489
		$p_{E,R}$	0.0467	0.0476	0.0499	0.0483	0.0499	0.0496
	Z_U	$p_{A,U}$	0.0314	0.0370	0.0404	0.0426	0.0447	0.0455
		$p_{A_c,U}$	0.0297	0.0336	0.0374	0.0403	0.0431	0.0441
		$p_{CI,U}^{(\gamma=0.001)}$	0.0450	0.0463	0.0477	0.0483	0.0482	0.0486
		$p_{E,U}$	0.0490	0.0471	0.0477	0.0483	0.0499	0.0496
1	T	$p_{A,T}$	0.0514	0.0490	0.0509	0.0515	0.0487	0.0491
		$p_{A_c,T}$	0.0389	0.0386	0.0422	0.0453	0.0442	0.0458
	Z_R	$p_{A,R}$	0.0492	0.0489	0.0498	0.0497	0.0497	0.0500
		$p_{A_c,R}$	0.0394	0.0396	0.0408	0.0437	0.0453	0.0465
		$p_{CI,R}^{(\gamma=0.001)}$	0.0486	0.0487	0.0475	0.0471	0.0486	0.0488
		$p_{E,R}$	0.0486	0.0489	0.0498	0.0497	0.0497	0.0498
	Z_U	$p_{A,U}$	0.0492	0.0489	0.0498	0.0497	0.0497	0.0500
		$p_{A_c,U}$	0.0394	0.0396	0.0408	0.0437	0.0453	0.0465
		$p_{CI,U}^{(\gamma=0.001)}$	0.0486	0.0487	0.0475	0.0471	0.0486	0.0488
		$p_{E,U}$	0.0486	0.0489	0.0498	0.0497	0.0497	0.0498
5/3	T	$p_{A,T}$	0.0477	0.0468	0.0477	0.0486	0.0481	0.0486
		$p_{A_c,T}$	0.0429	0.0441	0.0460	0.0474	0.0472	0.0479
	Z_R	$p_{A,R}$	0.0456	0.0467	0.0482	0.0470	0.0486	0.0491
		$p_{A_c,R}$	0.0452	0.0451	0.0464	0.0469	0.0477	0.0480
		$p_{CI,R}^{(\gamma=0.001)}$	0.0456	0.0467	0.0482	0.0470	0.0486	0.0490
		$p_{E,R}$	0.0484	0.0497	0.0496	0.0475	0.0497	0.0499
	Z_U	$p_{A,U}$	0.0639	0.0621	0.0593	0.0554	0.0547	0.0536
		$p_{A_c,U}$	0.0598	0.0576	0.0570	0.0553	0.0537	0.0529
		$p_{CI,U}^{(\gamma=0.001)}$	0.0393	0.0414	0.0459	0.0469	0.0479	0.0484
		$p_{E,U}$	0.0453	0.0475	0.0496	0.0475	0.0497	0.0499

Table 2.4: The type I error rate ($\delta_0 = 1$) of asymptotic p-value test (p_A) and exact p-value test (p_{CI}, p_E) based on T, Z_R, Z_U respectively for $n_2 = 30$.

ρ	Test Statistic	p-value	λ_2 0.3	0.4	0.6	1	2	3
3/5	T	$p_{A,T}$	0.9894	0.9771	0.9477	0.8685	0.6846	0.5612
		$p_{A_c,T}$	0.9887	0.9759	0.9457	0.8648	0.6809	0.5574
	Z_R	$p_{A,R}$	0.9905	0.9805	0.9518	0.8765	0.6935	0.5676
		$p_{A_c,R}$	0.9896	0.9791	0.9497	0.8716	0.6893	0.5640
		$p_{CI,R}^{(\gamma=0.001)}$	0.9871	0.9768	0.9477	0.8697	0.6852	0.5595
		$p_{E,R}$	0.9896	0.9788	0.9478	0.8697	0.6892	0.5618
	Z_U	$p_{A,U}$	0.9865	0.9745	0.9415	0.8577	0.6709	0.5460
		$p_{A_c,U}$	0.9853	0.9722	0.9376	0.8537	0.6658	0.5424
		$p_{CI,U}^{(\gamma=0.001)}$	0.9889	0.9784	0.9478	0.8690	0.6852	0.5577
		$p_{E,U}$	0.9889	0.9784	0.9478	0.8697	0.6892	0.5635
1	T	$p_{A,R}$	0.9977	0.9949	0.9825	0.9361	0.7872	0.6591
		$p_{A_c,R}$	0.9971	0.9937	0.9796	0.9293	0.7746	0.6459
	Z_R	$p_{A,R}$	0.9984	0.9956	0.9842	0.9415	0.7918	0.6668
		$p_{A_c,R}$	0.9979	0.9946	0.9816	0.9343	0.7814	0.6545
		$p_{CI,R}^{(\gamma=0.001)}$	0.9983	0.9953	0.9832	0.9393	0.7885	0.6612
		$p_{E,R}$	0.9984	0.9956	0.9842	0.9405	0.7918	0.6654
	Z_U	$p_{A,U}$	0.9984	0.9956	0.9842	0.9415	0.7918	0.6668
		$p_{A_c,U}$	0.9979	0.9946	0.9816	0.9343	0.7814	0.6545
		$p_{CI,U}^{(\gamma=0.001)}$	0.9983	0.9953	0.9832	0.9393	0.7885	0.6612
		$p_{E,U}$	0.9984	0.9956	0.9842	0.9405	0.7918	0.6654
5/3	T	$p_{A,T}$	0.9998	0.9989	0.9945	0.9710	0.8572	0.7370
		$p_{A_c,T}$	0.9997	0.9988	0.9942	0.9702	0.8551	0.7345
	Z_R	$p_{A,R}$	0.9998	0.9991	0.9953	0.9730	0.8625	0.7457
		$p_{A_c,R}$	0.9998	0.9991	0.9953	0.9720	0.8604	0.7425
		$p_{CI,R}^{(\gamma=0.001)}$	0.9998	0.9991	0.9953	0.9726	0.8619	0.7447
		$p_{E,R}$	0.9998	0.9992	0.9953	0.9746	0.8646	0.7479
	Z_U	$p_{A,U}$	0.9999	0.9994	0.9963	0.9769	0.8730	0.7586
		$p_{A_c,U}$	0.9998	0.9994	0.9963	0.9760	0.8716	0.7565
		$p_{CI,U}^{(\gamma=0.001)}$	0.9998	0.9991	0.9953	0.9726	0.8612	0.7438
		$p_{E,U}$	0.9998	0.9992	0.9953	0.9746	0.8646	0.7479

Table 2.5: To achieve 80% power at $\delta_0 = 0.6$, the required sample size of the second group n_2^* of Z_R, Z_U, T for $\rho = 3/5$. Based on the required samples n_2^*, the power and the type I error rate (in parentheses) are given.

Test Statistic	p-value		λ_2 0.3	0.4	0.6	1	2
T	$p_{A,T}$	n_2^*	27	31	41	59	105
		Power	0.8241	0.8111	0.8124	0.8055	0.7995
		(Size)	(0.0526)	(0.0533)	(0.0536)	(0.0527)	(0.0512)
	$p_{A_c,T}$	Power	0.8210	0.8040	0.8088	0.8025	0.7976
		(Size)	(0.0495)	(0.0495)	(0.0507)	0.0514)	(0.0506)
Z_R	$p_{A,R}$	n_2^*	27	31	41	59	105
		Power	0.8285	0.8088	0.8105	0.8045	0.8037
		(Size)	(0.0514)	(0.0543)	(0.0528)	(0.0516)	(0.0508)
	$p_{A_c,R}$	Power	0.8179	0.8048	0.8067	0.8021	0.8017
		(Size)	(0.0506)	(0.0499)	(0.0509)	(0.0500)	(0.0499)
	$p_{CI,R}^{(\gamma=0.001)}$	n_2^*	28	33	42	61	108
		Power	0.8147	0.8107	0.8118	0.8074	0.8055
		(Size)	(0.0449)	(0.0480)	(0.0477)	(0.0484)	(0.0488)
	$p_{E,R}$	n_2^*	27	32	41	59	107
		Power	0.8078	0.8141	0.8018	0.8001	0.8068
		(Size)	(0.0451)	(0.0466)	(0.0485)	(0.0496)	(0.0498)
Z_U	$p_{A,U}$	n_2^*	31	36	45	63	109
		Power	0.8301	0.8295	0.8216	0.8068	0.8053
		(Size)	(0.0345)	(0.0399)	(0.0426)	(0.0443)	(0.0469)
	$p_{A_c,U}^{(\gamma=0.001)}$	Power	0.8212	0.8188	(0.8184)	0.8039	0.8036
		(Size)	(0.0304)	(0.0368)	(0.0395)	(0.0428)	(0.0464)
	$p_{CI,U}$	n_2^*	28	33	42	61	108
		Power	0.8023	0.8097	0.8072	0.8056	0.8050
		(Size)	(0.0440)	(0.0442)	(0.0458)	(0.0483)	(0.0488)
	$p_{E,U}$	n_2^*	27	32	41	60	107
		Power	0.8083	0.8141	0.8018	0.8086	0.8068
		(Size)	(0.0466)	(0.0469)	(0.0485)	(0.0499)	(0.0498)

Table 2.6: To achieve 80% power at $\delta_0 = 0.6$, the required sample size of the second group n_2^* of Z_R, Z_U, T for $\rho = 1$. Based on the required samples n_2^*, the power and the type I error rate (in parentheses) are given.

Test Statistic	p-value		λ_2 0.3	0.4	0.6	1	2
T	$p_{A,T}$	n_2^*	21	25	31	45	79
		Power	0.8159	0.8190	0.8014	0.8003	0.7978
		(Size)	(0.0491)	(0.0494)	(0.0496)	(0.0498)	(0.0490)
	$p_{A_c,T}$	Power	0.7865	0.7933	0.7821	0.7871	0.7904
		(Size)	(0.0368)	(0.0393)	(0.0417)	(0.0448)	(0.0463)
Z_R	$p_{A,R}$	n_2^*	21	25	31	45	79
		Power	0.8244	0.8279	0.8084	0.8059	0.8017
		(Size)	(0.0512)	(0.0489)	(0.0498)	(0.0505)	(0.0499)
	$p_{A_c,R}$	Power	0.7971	0.8012	0.7909	0.7931	0.7940
		(Size)	(0.0373)	(0.0396)	(0.0410)	(0.0448)	(0.0471)
	$p_{CI,R}^{(\gamma=0.001)}$	n_2^*	20	24	31	45	80
		Power	0.8070	0.8088	0.8014	0.8032	0.8025
		(Size)	(0.0454)	(0.0487)	(0.0475)	(0.0488)	(0.0489)
	$p_{E,R}$	n_2^*	20	24	31	45	79
		Power	0.8073	0.8140	0.8084	0.8058	0.8011
		(Size)	(0.0454)	(0.0487)	(0.0498)	(0.0497)	(0.0499)
Z_U	$p_{A,U}$	n_2^*	21	25	31	45	79
		Power	0.8244	0.8279	0.8084	0.8059	0.8017
		(Size)	(0.0512)	(0.0489)	(0.0498)	(0.0505)	(0.0499)
	$p_{A_c,U}$	Power	0.7971	0.8012	0.7909	0.7931	0.7940
		(Size)	(0.0373)	(0.0396)	(0.0410)	(0.0448)	(0.0471)
	$p_{CI,U}^{(\gamma=0.001)}$	n_2^*	20	24	31	45	80
		Power	0.8070	0.8088	0.8014	0.8032	0.8025
		(Size)	(0.0454)	(0.0487)	(0.0475)	(0.0488)	(0.0489)
	$p_{E,U}$	n_2^*	20	24	31	45	79
		Power	0.8073	0.8140	0.8084	0.8058	0.8011
		(Size)	(0.0454)	(0.0487)	(0.0498)	(0.0497)	(0.0499)

Table 2.7: To achieve 80% power at $\delta_0 = 0.6$, the required sample size of the second group n_2^* of Z_R, Z_U, T for $\rho = 5/3$. Based on the required samples n_2^*, the power and the type I error rate (in parentheses) are given.

Test Statistic	p-value		λ_2 0.3	0.4	0.6	1	2
T	$p_{A,T}$	n_2^*	18	20	26	37	64
		Power	0.8265	0.8094	0.8117	0.8057	0.7984
		(Size)	(0.0491)	(0.0494)	(0.0496)	(0.0498)	(0.0490)
	$p_{A_c,T}$	Power	0.8194	0.8038	0.8074	0.8026	0.7966
		(Size)	(0.0368)	(0.0393)	(0.0417)	(0.0448)	(0.0463)
Z_R	$p_{A,R}$	n_2^*	18	20	26	37	64
		Power	0.8376	0.8151	0.8240	0.8090	0.8022
		(Size)	(0.0479)	(0.0457)	(0.0474)	(0.0496)	(0.0492)
	$p_{A_c,R}$	Power	0.8339	0.8080	0.8131	0.8054	0.8003
		(Size)	(0.0415)	(0.0431)	(0.0452)	(0.0496)	(0.0484)
	$p_{CI,R}^{(\gamma=0.001)}$	n_2^*	18	20	26	37	65
		Power	0.8376	0.8151	0.8165	0.8040	0.8066
		(Size)	(0.0479)	(0.0456)	(0.0474)	(0.0487)	(0.0489)
	$p_{E,R}$	n_2^*	17	20	25	36	64
		Power	0.8188	0.8201	0.8004	0.8069	0.8030
		(Size)	(0.0467)	(0.0456)	(0.0495)	(0.0476)	(0.0499)
Z_U	$p_{A,U}$	n_2^*	15	18	23	34	62
		Power	0.8182	0.8148	0.8030	0.7967	0.8008
		(Size)	(0.0753)	(0.0645)	(0.0598)	(0.0579)	(0.0532)
	$p_{A_c,U}$	Power	0.8072	0.8058	0.8006	0.7927	0.7988
		(Size)	(0.0749)	(0.0630)	(0.0574)	(0.0579)	(0.0524)
	$p_{CI,U}^{(\gamma=0.001)}$	n_2^*	18	20	26	37	65
		Power	0.8213	0.8078	0.8131	0.8038	0.8062
		(Size)	(0.0374)	(0.0374)	(0.0451)	(0.0449)	(0.0488)
	$p_{E,U}$	n_2^*	17	20	25	36	64
		Power	0.8188	0.8201	0.8004	0.8069	0.8030
		(Size)	(0.0401)	(0.0446)	(0.0495)	(0.0476)	(0.0499)

Table 2.8: To achieve 80% power at $\delta_0 = 1$, the required sample size of the second group n_2^* of Z_R, Z_U, T for $\rho = 3/5$. Based on the required samples n_2^*, the power and the type I error rate (in parentheses) are given.

Test Statistic	p-value		λ_2 0.3	0.4	0.6	1	2
T	$p_{A,T}$	n_2^*	13	15	18	25	41
		Power	0.8239	0.8243	0.8129	0.8072	0.8027
		(Size)	(0.0570)	(0.0582)	(0.0542)	(0.0518)	(0.0508)
	$p_{A_c,T}$	Power	0.8206	0.8184	0.8093	0.8025	0.7999
		(Size)	(0.0496)	(0.0515)	(0.0508)	(0.0493)	(0.0495)
Z_R	$p_{A,R}$	n_2^*	13	15	18	25	41
		Power	0.8342	0.8393	0.8134	0.8219	0.8014
		(Size)	(0.0589)	(0.0537)	(0.0554)	(0.0522)	(0.0508)
	$p_{A_c,R}$	Power	0.8155	0.8346	0.8050	0.8133	0.8014
		(Size)	(0.0412)	(0.0537)	(0.0494)	(0.0501)	(0.0504)
	$p_{CI,R}^{(\gamma=0.001)}$	n_2^*	14	15	19	25	42
		Power	0.8383	0.8040	0.8217	0.8118	0.8064
		(Size)	(0.0426)	(0.0359)	(0.0435)	(0.0477)	(0.0486)
	$p_{E,R}$	n_2^*	12	15	19	25	42
		Power	0.8010	0.8191	0.8217	0.8118	0.8068
		(Size)	(0.0421)	(0.0445)	(0.0446)	(0.0482)	(0.0486)
Z_U	$p_{A,U}$	n_2^*	16	17	21	27	44
		Power	0.8581	0.8385	0.8335	0.8200	0.8108
		(Size)	(0.0292)	(0.0265)	(0.0386)	(0.0410)	(0.0455)
	$p_{A_c,U}$	Power	0.8488	0.8275	0.8297	0.8141	0.8098
		(Size)	(0.0208)	(0.0257)	(0.0356)	(0.0395)	(0.0440)
	$p_{CI,U}^{(\gamma=0.001)}$	n_2^*	14	15	19	25	42
		Power	0.8363	0.8260	0.8217	0.8118	0.8059
		(Size)	(0.0426)	(0.0402)	(0.0446)	(0.0482)	(0.0481)
	$p_{E,U}$	n_2^*	12	15	19	25	42
		Power	0.8010	0.8305	0.8217	0.8118	0.8068
		(Size)	(0.0421)	(0.0459)	(0.0446)	(0.0482)	(0.0488)

Table 2.9: To achieve 80% power at $\delta_0 = 1$, the required sample size of the second group n_2^* of Z_R, Z_U, T for $\rho = 1$. Based on the required samples n_2^*, the power and the type I error rate (in parentheses) are given.

Test Statistic	p-value		λ_2 0.3	0.4	0.6	1	2
T	$p_{A,T}$	n_2^*	10	12	14	19	31
		Power	0.8050	0.8261	0.8067	0.8039	0.7986
		(Size)	(0.0462)	(0.0497)	(0.0488)	(0.0502)	(0.0495)
	$p_{A_c,T}$	Power	0.7714	0.7954	0.7814	0.7845	0.7866
		(Size)	(0.0296)	(0.0322)	(0.0395)	(0.0427)	(0.0453)
Z_R	$p_{A,R}$	n_2^*	10	12	14	19	31
		Power	0.8387	0.8486	0.8258	0.8159	0.8032
		(Size)	(0.0497)	(0.0518)	(0.0495)	(0.0497)	(0.0498)
	$p_{A_c,R}^{(\gamma=0.001)}$	Power	0.8044	0.8237	0.7978	0.7988	0.7934
		(Size)	(0.0331)	(0.0364)	(0.0391)	(0.0411)	(0.0454)
	$p_{CI,R}$	n_2^*	10	11	14	19	31
		Power	0.8323	0.8192	0.8227	0.8089	0.8001
		(Size)	(0.0448)	(0.0422)	(0.0484)	(0.0474)	(0.0487)
	$p_{E,R}$	n_2^*	10	11	14	19	31
		Power	0.8323	0.8192	0.8227	0.8117	0.8001
		(Size)	(0.0448)	(0.0422)	(0.0484)	(0.0475)	(0.0487)
Z_U	$p_{A,U}$	n_2^*	10	12	14	19	31
		Power	0.8387	0.8486	0.8258	0.8159	0.8032
		(Size)	(0.0497)	(0.0518)	(0.0495)	(0.0497)	(0.0498)
	$p_{A_c,U}$	Power	0.8044	0.8237	0.7978	0.7988	0.7934
		(Size)	(0.0331)	(0.0364)	(0.0391)	(0.0411)	(0.0454)
	$p_{CI,U}^{(\gamma=0.001)}$	n_2^*	10	11	14	19	31
		Power	0.8323	0.8192	0.8227	0.8089	0.8001
		(Size)	(0.0448)	(0.0422)	(0.0484)	(0.0475)	(0.0487)
	$p_{E,U}$	n_2^*	10	11	14	19	31
		Power	0.8323	0.8192	0.8227	0.8117	0.8001
		(Size)	(0.0448)	(0.0422)	(0.0484)	(0.0475)	(0.0487)

Table 2.10: To achieve 80% power at $\delta_0 = 1$, the required sample size of the second group n_2^* of Z_R, Z_U, T for $\rho = 5/3$. Based on the required samples n_2^*, the power and the type I error rate (in parentheses) are given.

Test Statistic	p-value		λ_2 0.3	0.4	0.6	1	2
T	$p_{A,T}$	n_2^*	9	10	12	16	26
		Power	0.8362	0.8329	0.8208	0.8151	0.8105
		(Size)	(0.0462)	(0.0497)	(0.0488)	(0.0502)	(0.0495)
	$p_{A_c,T}$	Power	0.8275	0.8229	0.8134	0.8111	0.8080
		(Size)	(0.0296)	(0.0322)	(0.0395)	(0.0427)	(0.0453)
Z_R	$p_{A,R}$	n_2^*	9	10	12	16	26
		Power	0.8638	0.8522	0.8398	0.8180	0.8146
		(Size)	(0.0404)	(0.0474)	(0.0468)	(0.0479)	(0.0486)
	$p_{A_c,R}$	Power	0.8605	0.8422	0.8335	0.8128	0.8117
		(Size)	(0.0394)	(0.0388)	(0.0446)	(0.0466)	(0.0475)
	$p_{CI,R}^{(\gamma=0.001)}$	n_2^*	8	9	11	15	26
		Power	0.8101	0.8134	0.8027	0.8023	0.8145
		(Size)	(0.0366)	(0.0429)	(0.0445)	(0.0474)	(0.0475)
	$p_{E,R}$	n_2^*	8	9	11	15	26
		Power	0.8222	0.8206	0.8099	0.8023	0.8146
		(Size)	(0.0366)	(0.0429)	(0.0481)	(0.0477)	(0.0486)
Z_U	$p_{A,U}$	n_2^*	7	8	10	14	24
		Power	0.8244	0.8159	0.8086	0.8020	0.8030
		(Size)	(0.0922)	(0.0802)	(0.0644)	(0.0596)	(0.0551)
	$p_{A_c,U}$	Power	0.8059	0.8143	0.8002	0.8016	0.8000
		(Size)	(0.0651)	(0.0749)	(0.0629)	(0.0578)	(0.0532)
	$p_{CI,U}^{(\gamma=0.001)}$	n_2^*	9	10	12	16	26
		Power	0.8427	0.8313	0.8288	0.8128	0.8132
		(Size)	(0.0394)	(0.0335)	(0.0383)	(0.0469)	(0.0468)
	$p_{E,U}$	n_2^*	8	9	11	15	26
		Power	0.8222	0.8122	0.8099	0.8023	0.8146
		(Size)	(0.0366)	(0.0389)	(0.0446)	(0.0474)	(0.0486)

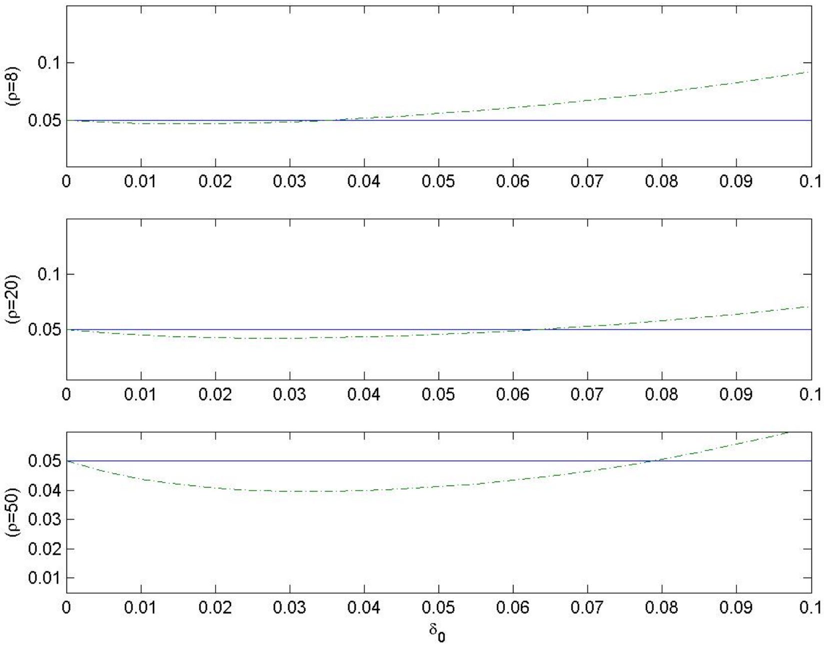

Figure 2.2: As $n_2 = 10, \lambda_2 = 0.03, \rho = 8, 20, 50$, the asymptotic power of Z_R over $\delta_0 \in (0, 0.1)$.

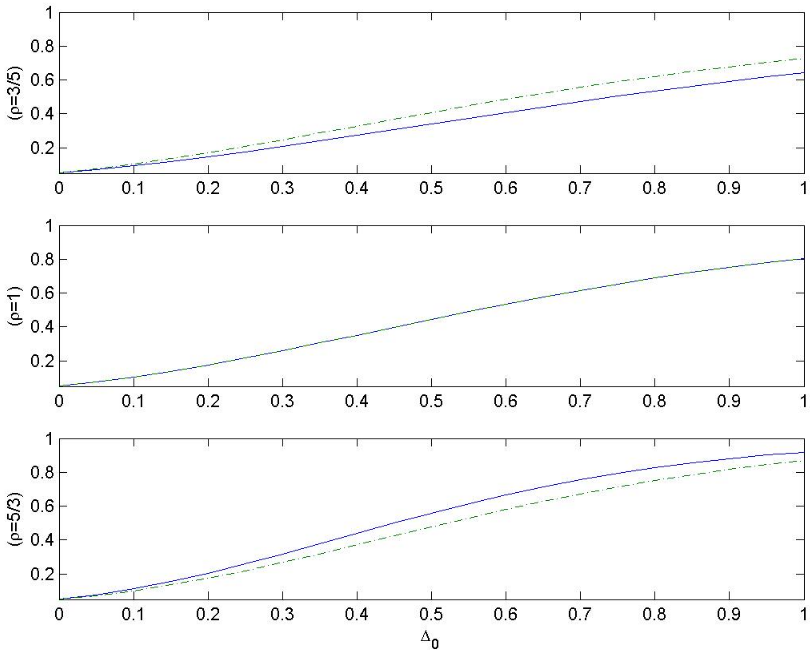

Figure 2.3: As $n_2 = 10, \lambda_2 = 0.3, \rho = 3/5, 1, 5/3$, the asymptotic powers of the Z_R(the dotted and dashed line) and Z_U(the solid line) over $\delta_0 \in (0, 1)$.

Chapter 3

Testing the superiority

3.1 Statistical hypothesis and Test Statistics

In this chapter, we consider testing the superiority with the conventional complementary null hypothesis,

$$H_{02} : \lambda_1 \leq \lambda_2 \quad \text{vs.} \quad H_1 : \lambda_1 > \lambda_2.$$

Recall that the null parameter space is denoted as $\Omega_{02} = \{(\lambda_1, \lambda_2) : \lambda_1 \leq \lambda_2, \lambda_1 > 0\}$, which is region above and includes the diagonal line, see Figure 2.1 in Chapter 2. The two types of Wald statistic, Z_R, Z_U are employed as test statistics. First, their correspondent asymptotic testing procedures will be investigated. Since this chapter and Chapter 2 only differ in the null hypothesis, which affects the validity property of a test. Hence, in next section, we will focus on justifying the validity of the two asymptotic tests. The two exact tests based on the confidence-set p-value and the estimated p-value will be introduced in this chapter. Because the null parameter space becomes wider here, computation of an exact test increases and becomes more complicated. Hence one important goal of our study is to develop efficient exact tests with successful reduction in computations. The details will be given in Section 3.3. Later the results of a numerical study will be presented and discussed in Section 3.4.

3.2 Asymptotic p-values

If the means of two groups are relative large or sample sizes are sufficiently enough, an asymptotic test under normality can be considered in this problem. Since the alternative hypothesis remains the same as Chapter 2, we have the same results in the property of unbiasedness for the testing procedures and it suffices to investigate their validity here. In this section, we study the two asymptotic testing procedures based on the p-values $p_{A,R}$ and $p_{A,U}$ defined in Chapter 2. From Theorem 1 in Chapter 2, recall that the asymptotic distributions of Z_R and Z_U are expressed as follows,

$$Z_R \cdot \sigma - \mu \ \overset{d}{\to} \ N(0,1) \quad \text{and} \quad Z_U - \mu \ \overset{d}{\to} \ N(0,1) \quad \text{as} \quad n_1, n_2 \to \infty.$$

In which,

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

Consequently, the asymptotic power functions of the two asymptotic tests are respectively represented as follows:

$$\bar{\beta}_{Z_R}(\delta_0, \lambda_2, \rho, n_2) = 1 - \Phi\left(z_\alpha \sigma - \mu\right),$$

and

$$\bar{\beta}_{Z_U}(\delta_0, \lambda_2, \rho, n_2) = 1 - \Phi\left(z_\alpha - \mu_0\right).$$

Under H_{02}, we have $\delta_0 = \lambda_1 - \lambda_2 \leq 0$. If the sampling fraction $\rho \leq 1$, then the component σ in $\bar{\beta}_{Z_R}$ is easily found greater than 1, and $z_\alpha \sigma - \mu \geq z_\alpha$

is always true. Further, taking the partial derivative of $\bar{\beta}_{Z_R}$ with respect to δ_0, we find that $\bar{\beta}_{Z_R}$ increases as δ_0. Hence, the maximum of $\bar{\beta}_{Z_R}$ occurs at $\delta_0 = 0$ and is equal to α. It means that the asymptotic test by using the p-value $p_{A,R}$ is asymptotic valid when $\rho \leq 1$. However, when $\rho > 1$, the power $\bar{\beta}_{Z_R}$ may exceed the level α whenever

$$z_\alpha \sigma - \mu < z_\alpha. \tag{3.1}$$

It is likely to happen with an extremely large ρ when the true λ_1 is close to zero and δ_0 is nearly $-\lambda_2$. Figure 3.4 give the plots of the asymptotic power function $\bar{\beta}_{Z_R}$ for $\delta_0 \in (-0.3, 0)$ with $n_2 = 5, \lambda_2 = 0.3$ at various scenarios of ρ. In the left panel, one can see that $\bar{\beta}_{Z_R}$ has maximal value α at $\delta_0 = 0$ as $\rho < 1$ or ρ is not far greater than 1. On the other hand, in the right panel, where the power functions are evaluated at $\rho = 18, 25, 30$, we discover that $\bar{\beta}_{Z_R}$ can exceed α in the area where δ_0 is close to the boundary $-\lambda_2 = -0.3$. The magnitude and area of the inflation of the type I error rate become severe as ρ increases. However, this fault can be improved when sample sizes increase sightly. Note that we have shown that the asymptotic test based on Z_R is asymptotic valid under the null hypothesis of equality H_{01} in Chapter 2. Here in this section, the test is found not able to control its type I error rate at significance level when the first group has an extremely larger sample size than the second group. Note that in Chapter 2, we have shown that the two-independent-sample T-test has the same asymptotic distribution as Z_R. They own the same asymptotic properties for sufficiently large n_1, n_2. On the other hand, we find that $\bar{\beta}_{Z_U}$ increases as δ_0. Hence its maximum occurs at $\delta_0 = 0$ and is equal to α. Thus the asymptotic test by using Z_U is asymptotic valid under H_{02}.

In summary, we find from Chapter 2 and Chapter 3 that the asymptotic test correspondent to Z_U is asymptotic valid under both H_{01}, H_{02} and is always unbiased. On the other hand, the tests by using the p-values $p_{A,R}$

and p_T of the test statistics Z_R, T are asymptotic valid under H_{01}, but no longer valid under H_{02} when ρ is extremely large. Further, recall from last chapter, the two tests are biased under such circumstances.

3.3 Exact p-values

When the sample sizes are insufficient or the mean values are relatively small, exact testing procedures are more appropriate for establishing the superiority. Consider the Wald statistic Z, where Z can be either Z_R or Z_U. Under the null hypothesis of non-superiority, H_{02}, the exact p-value given an observed z_0 is defined as follows,

$$\begin{aligned} p_{(\lambda_1,\lambda_2)}(z_0) &= P(Z \geq z_0 | H_{02} : 0 < \lambda_1 \leq \lambda_2) \\ &= \sum_{y_1 \geq 0} \sum_{y_2 \geq 0} poi(y_1, n_1\lambda_1) poi(y_2, n_2\lambda_2) I_{\{Z \geq z_0\}}, \end{aligned}$$

where $poi(y, \lambda')$ is the probability function of Poisson distribution with mean λ', and y_1, y_2 are possible outcomes of Y_1, Y_2, respectively. The exact p-value $p(\lambda_1, \lambda_2)$ depends on two nuisance parameters. Again to control the size of a testing procedure, one can consider the standard p-value, which is defined as the supremum of the exact p-value over the null parameter space. Recall that in last chapter, an exact p-value involves only single nuisance parameter, the common mean value under H_{01}. Hence, the supremum is searched only along the main diagonal $\lambda_1 = \lambda_2$. Now because the null parameter space becomes wider, computing a standard exact p-value is a more complicated task here. In this chapter we aim to find the testing procedures that are efficient in reducing calculations of p-values. The first strategy is to reduce the range for the supremum search. Again, the confidence-set p-value and a revised estimated p-value are proposed in this section.

Under H_{02}, the confidence-set p-value by Berger and Boos (1994) is defined as,

$$p_{CI} = \sup_{(\lambda_1,\lambda_2) \in C^*_\gamma} P(Z \geq z_0 | 0 < \lambda_1 \leq \lambda_2) + \gamma.$$

In which C^*_γ is a joint confidence set of (λ_1, λ_2) that guarantees $100(1-\gamma)\%$

confidence within the null parameter space Ω_{02}. That is,

$$P((\lambda_1, \lambda_2) \in C^*_\gamma | \lambda_1, \lambda_2) \geq 1 - \gamma, \ \text{ for any } (\lambda_1, \lambda_2) \in \Omega_{02},$$

where $\Omega_{02} = \{(\lambda_1, \ \lambda_2) : 0 < \lambda_1 \leq \lambda_2\}$. The construction of a confidence set in a restricted parameter space is less straight forward. Subsequently, we propose to truncate a confidence set that is build under the unrestricted parameter space. Consider the following cross-product set,

$$C_{\gamma,0} = \{(\lambda_1, \lambda_2) : L_1 \leq \lambda_1 \leq U_1, \ \ L_2 \leq \lambda_2 \leq U_2\},$$

where (L_1, U_1) and (L_2, U_2) are two independent $100\sqrt{(1-\gamma)}\%$ confidence interval of λ_1 and λ_2 respectively. Then it is easily shown that $C_{\gamma,0}$ is a $100(1-\gamma)\%$ confidence set of (λ_1, λ_2) in the unrestricted parameter space Ω. Next theorem shows that the coverage probability of the truncated confidence set, which is defined as the intercept of $C_{\gamma,0}$ and Ω_{02}, is at least $1-\gamma$ under Ω_{02}.

Theorem 4. Let $C^*_\gamma = C_{\gamma,0} \cap \Omega_{02}$ be the truncated confidence set. Then

$$P((\lambda_1, \lambda_2) \in C^*_\gamma \mid \lambda_1, \lambda_2) \geq 1 - \gamma, \quad \text{for all } (\lambda_1, \lambda_2) \in \Omega_{02}.$$

Again (L_1, U_1) and (L_2, U_2) are derived through the relation between Poisson distribution and chi-squared distribution, and are respectively represented as follows,

$$(L_1, U_1) = \frac{1}{2n_1}\left(\chi^2_{(1-(1-\sqrt{1-\gamma})/2,\ 2Y_1)},\ \chi^2_{((1-\sqrt{1-\gamma})/2,\ 2(Y_1+1))}\right),$$

and

$$(L_2, U_2) = \frac{1}{2n_2}\left(\chi^2_{(1-(1-\sqrt{1-\gamma})/2,\ 2Y_2)},\ \chi^2_{((1-\sqrt{1-\gamma})/2,\ 2(Y_2+1))}\right).$$

Furthermore, C^*_γ is of the following form,

$$C^*_\gamma = \{L_1 \leq \lambda_1 \leq \min(U_1, \lambda_2), \;\; L_2 \leq \lambda_2 \leq U_2\}.$$

Consequently, the correspondent confidence-set p-value of Z_R is given by

$$p^{(\gamma)}_{CI,R} = \sup_{(\lambda_1,\lambda_2)\in C^*_\gamma} P(Z_R \geq z_R | 0 < \lambda_1 \leq \lambda_2) + \gamma,$$

and the correspondent confidence-set p-value of Z_U is given by

$$p^{(\gamma)}_{CI,U} = \sup_{(\lambda_1,\lambda_2)\in C^*_\gamma} P(Z_U \geq z_U | 0 < \lambda_1 \leq \lambda_2) + \gamma,$$

as long as the realization of C^*_γ is not empty. When the observed $C_{\gamma,0}$ is completely outside of Ω_{02}, then C^*_γ is empty. In this case, we define $p_{CI} = \gamma < \alpha$, and reject the null hypothesis H_{02}. This confidence-set p-value is always valid. However, C^*_γ is still a large region and the calculations required for the p-value p_{CI} are not as easy as before. In the following, a sufficient condition on the test statistic for its correspondent p-value to own some kind of monotonicity is introduced. The monotonicity ensures that the supremum occurs at the boundary. As a consequence, computational burden is greatly reduced.

Barnard (1947) proposed the so-called convexity condition for a test statistic in a bivariate discrete distribution. The condition is described as follows:

$$S(s_1, s_2) \leq S(s_1 + 1, s_2) \quad \text{and} \quad S(s_1, s_2) \leq S(s_1, s_2 - 1),$$

where (s_1, s_2) is a realization of the two discrete random variables. The condition means that: If an outcome leads to reject the null hypothesis, then the outcome with greater value of the random variable in the first population or smaller value of the random variable in the second population, leads to reject the null hypotheses as well. Röhmel and Mansmann (1999) derived the property that whenever the test statistic S satisfies the convexity condition, the supremum of the exact p-value is a maximum and is attained at a boundary point under the Binomial distribution. We show that the property holds in comparing two Poisson means in next theorem.

Theorem 5. Let S be a test statistic that depends on the data only through the two sufficient statistics (Y_1, Y_2) in comparing two Poisson means. Suppose S satisfies the convexity condition. Then given s_0, the supremum of $P(S \geq s_0|\lambda_1, \lambda_2)$ occurs at a boundary point of the parameter space.

Theorem 6. Z_R, Z_U satisfy the convexity condition.

The convexity of Z_U, Z_R in Theorem 6 is shown from the monotonicity of Z_U and Z_R with respect to Y_1 and Y_2, see Appendix A.6. Hence, by Theorem 5 and 6, we obtain that the confidence-set p-values of Z_R and Z_U are evaluated in the boundary of the confidence set C^*_γ. That is,

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

where ∂C^*_γ is the boundary of C^*_γ. Therefore, we discover that the two associated confidence-set p-values based on Z_R and Z_U can have their computations dramatically reduced.

Moreover, the probabilities, $P(Z_R \geq z_R|\lambda_1, \lambda_2), P(Z_U \geq z_U|\lambda_1, \lambda_2)$ can be shown to be increasing as λ_1 increases and λ_2 decreases, see the proof of Theorem 5 in Appendix. Hence, when C^*_γ is non-empty, the supremums for the confidence-set p-values based on Z_R, Z_U either occur at the point

(U_1, L_2) or somewhere on the diagonal. Again the supremums are found by grid-search method in the latter case.

In testing the null hypothesis of equality, the estimated exact p-value proposed by Krishnamoorthy and Thomson (2004) although does not guarantee theoretical validity, but has great computational efficiency and gives satisfactory performance in numerical studies. In the following, we adapt the idea and propose an estimated exact p-value for testing the null hypothesis of non-superiority. Define the estimated exact p-values as

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

where $(\tilde{\lambda}_{01}, \tilde{\lambda}_{02})$ are some estimators of (λ_1, λ_2) under the restricted null parameter space Ω_{02}. Potential candidates for $(\tilde{\lambda}_{01}, \tilde{\lambda}_{02})$ are the RMLEs on Ω_{02}. However, because directly solving for RMLEs is quite difficult, we consider a revised procedure. First solve for the unrestricted MLEs $(\hat{\lambda}_1, \hat{\lambda}_2)$. If it happens that $(\hat{\lambda}_1, \hat{\lambda}_2) \in \Omega_{02}$, i.e. $\hat{\lambda}_1 \leq \hat{\lambda}_2$, $(\hat{\lambda}_1, \hat{\lambda}_2)$ are exactly the RMLEs under Ω_{02} and let $(\tilde{\lambda}_{01}, \tilde{\lambda}_{02}) = (\hat{\lambda}_1, \hat{\lambda}_2)$. However, if $\hat{\lambda}_1 > \hat{\lambda}_2$, we take the RMLE under the diagonal $\lambda_1 = \lambda_2$, i.e. $(\tilde{\lambda}_{01}, \tilde{\lambda}_{02}) = (\tilde{\lambda}_0, \tilde{\lambda}_0)$. In summary,

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

The reason for selecting the RMLE under the diagonal is for a conservative conclusion. It's known that the exact p-value is an increasing function as the parameter point (λ_1, λ_2) moves toward the down-right direction. To avoid a liberal conclusion, the p-value is evaluated at the most down-right location of Ω_{02}, which is on the main diagonal. In next section, we will conduct extensive numerical studies to compare the performance of these proposed testing procedures.

3.4 Numerical Studies

In the numerical studies, the test statistics used are Z_R, Z_U and T. For Z_R, Z_U, the asymptotic test by using the asymptotic p-value, denoted as p_A, and the two exact tests by using the confidence-set p-value and the estimated p-value, denoted as p_E, are investigated. Two confidence-set p-values are constructed at $\gamma = 0.001, 0.005$, and denoted as $p_{CI,\cdot}^{(\gamma=0.001)}, p_{CI,\cdot}^{(\gamma=0.005)}$, respectively. For the two-independent-sample T statistic, only the test by using p_A calculated from a t-distribution is studied. Because the Wald statistics are functions of the two sufficient statistics, the exact powers of the associated tests can be easily computed. On the other hand, the power of the T-test is found through 100,000 replicates. We consider $\lambda_2 = 1, 2, n_2 = 10, \rho = 3/5, 1, 5/3$, and δ_0 is ranged within -0.25 to 2. The scenarios of $\delta_0 \leq 0$ correspond to null cases, while that of $\delta > 0$ are the alternative ones. Table 3.1-3.2 present the power at 5% significant level.

First, we compare the three asymptotic tests in Table 3.1 and 3.2. We find that although Z_R and T have different numerical results in the finite sample case, they have quite consistent patterns as presented in Chapter 2. When $\rho = 1$, Z_R and Z_U are of the same form and have completely the same results.

Theoretically, as $\rho \leq 1$, the type I error rate of Z_R (T) increases as δ_0, and has its maximum $\alpha = 5\%$ occurred at $\delta_0 = 0$ approximately. One finds the consistent trend in the finite-sample cases from Table 3.1 and 3.2. However, the maximal type I error rate, occurred at $\delta_0 = 0$, exceeds the nominal level for $\rho = 3/5$. On the other hand, Z_U is found being not able to control its type I error rate when $\rho = 5/3$. Recall that Z_U is always valid asymptotically. In Table 3.1 and 3.2, all the power of the three tests increase with $\delta_0 > 0$. When

$\rho = 3/5 < 1$, Z_R and T have more chance to reject the null hypothesis than Z_U. The trend is contrary when $\rho = 5/3 > 1$. In summary, the findings on the comparison between the three asymptotic tests are the same as Chapter 2.

On the other hand, we can find that the two exact p-values almost have their sizes well controlled at $\alpha = 5\%$. The only exception is at $\lambda_2 = 2, \delta_0 = 0, \rho = 3/5$, at where the estimated p-value of Z_U has a type I error rate 5.3%. Using the same test statistic, the size of the estimated p-value p_E is always more close to the nominal level and is more efficient in computations than the confidence-set p-value p_{CI}. However, the estimated p-value is not theoretically valid and sometimes exceeds the nominal level as found in the exception. For the estimated p-value, the use of Z_U brings about more powerful results than Z_R when $\rho \neq 1$.

For a confidence-set p-value, a larger γ leads to less computations involved for the supremum search. However, with a trade-off term, which adjusts for the selection of γ, of the confidence-set p-value, the performance of the test is not significantly affected by γ. As other testing procedures, the test statistic used in the confidence-set p-value causes some effect on the performance. Interestingly, the trend is totally opposite to that of the asymptotic tests. Here compared with the use of Z_R, the employment of Z_U is more powerful at $\rho < 1$, and less powerful at $\rho > 1$.

Table 3.1: Type I error rate and power of asymptotic p-value and exact p-value at $\lambda_2 = 1, n_2 = 10$, these p-values are based on test statistics T, Z_R, Z_U respectively.

ρ	Test Statistic	p-value	δ -0.25	-0.15	-0.1	-0.05	0.0	0.1	0.5	1.0	1.5	2.0
3/5	T	$p_{A,T}$	0.0170	0.0279	0.0348	0.0430	0.0526	0.0772	0.2166	0.4709	0.7053	0.8627
	Z_R	$p_{A,R}$	0.0157	0.0266	0.0337	0.0421	0.0519	0.0757	0.2298	0.5024	0.7432	0.8907
		$p_{CI,R}^{(\gamma=0.001)}$	0.0096	0.0176	0.0231	0.0297	0.0375	0.0574	0.1942	0.4524	0.6999	0.8655
		$p_{CI,R}^{(\gamma=0.005)}$	0.0097	0.0176	0.0232	0.0297	0.0375	0.0574	0.1941	0.4520	0.6990	0.8643
		$p_{E,R}$	0.0137	0.0233	0.0297	0.0372	0.0460	0.0675	0.2099	0.4728	0.7194	0.8781
	Z_U	$p_{A,U}$	0.0082	0.0153	0.0202	0.0262	0.0334	0.0517	0.1833	0.4425	0.6942	0.8623
		$p_{CI,U}^{(\gamma=0.001)}$	0.0129	0.0228	0.0293	0.0370	0.0461	0.0682	0.2120	0.4743	0.7199	0.8782
		$p_{CI,U}^{(\gamma=0.005)}$	0.0107	0.0197	0.0258	0.0330	0.0416	0.0628	0.2037	0.4629	0.7085	0.8706
		$p_{E,U}$	0.0145	0.0250	0.0318	0.0399	0.0493	0.0721	0.2199	0.4871	0.7310	0.8841
1	T	$p_{A,T}$	0.0140	0.0248	0.0320	0.0408	0.0507	0.0762	0.2443	0.5479	0.7995	0.9315
	Z_R	$p_{A,R}$	0.0126	0.0230	0.0301	0.0387	0.0489	0.0748	0.2554	0.5773	0.8279	0.9477
		$p_{CI,R}^{(\gamma=0.001)}$	0.0123	0.0227	0.0298	0.0384	0.0487	0.0746	0.2544	0.5724	0.8223	0.9451
		$p_{CI,R}^{(\gamma=0.005)}$	0.0104	0.0191	0.0251	0.0326	0.0415	0.0646	0.2350	0.5554	0.8145	0.9422
		$p_{E,R}$	0.0123	0.0227	0.0298	0.0384	0.0487	0.0747	0.2554	0.5773	0.8279	0.9477
	Z_U	$p_{A,U}$	0.0126	0.0230	0.0301	0.0387	0.0489	0.0748	0.2554	0.5773	0.8279	0.9477
		$p_{CI,U}^{(\gamma=0.001)}$	0.0123	0.0227	0.0298	0.0384	0.0487	0.0746	0.2544	0.5724	0.8223	0.9451
		$p_{CI,U}^{(\gamma=0.005)}$	0.0104	0.0191	0.0251	0.0326	0.0415	0.0646	0.2350	0.5554	0.8145	0.9422
		$p_{E,U}$	0.0123	0.0227	0.0298	0.0384	0.0487	0.0747	0.2554	0.5773	0.8279	0.9477
5/3	T	$p_{A,T}$	0.0110	0.0203	0.0285	0.0370	0.0474	0.0735	0.2712	0.6269	0.8719	0.9699
	Z_R	$p_{A,R}$	0.0101	0.0196	0.0264	0.0351	0.0457	0.0736	0.2831	0.6497	0.8918	0.9782
		$p_{CI,R}^{(\gamma=0.001)}$	0.0101	0.0196	0.0265	0.0351	0.0457	0.0736	0.2831	0.6495	0.8912	0.9776
		$p_{CI,R}^{(\gamma=0.005)}$	0.0092	0.0183	0.0248	0.0329	0.0429	0.0694	0.2731	0.6391	0.8852	0.9757
		$p_{E,R}$	0.0112	0.0216	0.0289	0.0379	0.0488	0.0771	0.2858	0.6560	0.8954	0.9792
	Z_U	$p_{A,U}$	0.0159	0.0292	0.0384	0.0496	0.0629	0.0964	0.3250	0.6888	0.9100	0.9831
		$p_{CI,U}^{(\gamma=0.001)}$	0.0082	0.0168	0.0231	0.0312	0.0411	0.0674	0.2678	0.6323	0.8864	0.9771
		$p_{CI,U}^{(\gamma=0.005)}$	0.0082	0.0168	0.0231	0.0312	0.0411	0.0674	0.2676	0.6293	0.8806	0.9747
		$p_{E,U}$	0.0111	0.0216	0.0291	0.0385	0.0499	0.0795	0.2945	0.6592	0.8957	0.9793

Table 3.2: Type I error rate and power of asymptotic p-value and exact p-value at $\lambda_2 = 2, n_2 = 10$, these p-values are based on test statistics T, Z_R, Z_U respectively.

ρ	Test Statistic	p-value	δ = -0.25	-0.15	-0.1	-0.05	0.0	0.1	0.5	1	1.5	2
3/5	T	$p_{A,T}$	0.0246	0.0345	0.0403	0.0458	0.0527	0.0666	0.1582	0.3345	0.5299	0.7120
	Z_R	$p_{A,R}$	0.0242	0.0338	0.0395	0.0459	0.0529	0.0694	0.1669	0.3516	0.5618	0.7433
		$p_{CI,R}^{(\gamma=0.001)}$	0.0191	0.0274	0.0325	0.0382	0.0446	0.0597	0.1524	0.3341	0.5451	0.7302
		$p_{CI,R}^{(\gamma=0.005)}$	0.0183	0.0261	0.0309	0.0363	0.0424	0.0567	0.1453	0.3219	0.5310	0.7179
		$p_{E,R}$	0.0215	0.0305	0.0359	0.0419	0.0486	0.0644	0.1595	0.3432	0.5540	0.7368
	Z_U	$p_{A,U}$	0.0177	0.0253	0.0299	0.0351	0.0409	0.0546	0.1396	0.3110	0.5177	0.7060
		$p_{CI,U}^{(\gamma=0.001)}$	0.0214	0.0303	0.0356	0.0415	0.0482	0.0637	0.1570	0.3374	0.5467	0.7308
		$p_{CI,U}^{(\gamma=0.005)}$	0.0195	0.0277	0.0326	0.0382	0.0444	0.0590	0.1481	0.3240	0.5320	0.7183
		$p_{E,U}$	0.0232	0.0329	0.0388	0.0455	0.0530	0.0706	0.1811	0.4002	0.6415	0.8260
1	T	$p_{A,T}$	0.0204	0.0304	0.0366	0.0439	0.0500	0.0680	0.1744	0.3880	0.6236	0.8075
	Z_R	$p_{A,R}$,	0.0201	0.0295	0.0354	0.0420	0.0496	0.0675	0.1818	0.4073	0.6508	0.8333
		$p_{CI,R}^{(\gamma=0.001)}$	0.0190	0.0281	0.0337	0.0401	0.0475	0.0649	0.1769	0.3988	0.6413	0.8267
		$p_{CI,R}^{(\gamma=0.005)}$	0.0178	0.0264	0.0317	0.0378	0.0448	0.0613	0.1683	0.3853	0.6285	0.8177
		$p_{E,R}$	0.0201	0.0295	0.0354	0.0420	0.0496	0.0675	0.1818	0.4073	0.6508	0.8333
	Z_U	$p_{A,U}$	0.0201	0.0295	0.0354	0.0420	0.0496	0.0675	0.1818	0.4073	0.6508	0.8333
		$p_{CI,U}^{(\gamma=0.001)}$	0.0190	0.0281	0.0337	0.0401	0.0475	0.0649	0.1769	0.3988	0.6413	0.8267
		$p_{CI,U}^{(\gamma=0.005)}$	0.0178	0.0264	0.0317	0.0378	0.0448	0.0613	0.1683	0.3853	0.6285	0.8177
		$p_{E,U}$	0.0201	0.0295	0.0354	0.0420	0.0496	0.0675	0.1818	0.4073	0.6508	0.8333
5/3	T	$p_{A,T}$	0.0178	0.0273	0.0328	0.0410	0.0480	0.0657	0.1925	0.4464	0.7058	0.8804
	Z_R	$p_{A,R}$	0.0173	0.0268	0.0328	0.0398	0.0479	0.0677	0.1993	0.4608	0.7274	0.8981
		$p_{CI,R}^{(\gamma=0.001)}$	0.0172	0.0265	0.0325	0.0393	0.0473	0.0665	0.1954	0.4573	0.7266	0.8980
		$p_{CI,R}^{(\gamma=0.005)}$	0.0161	0.0248	0.0303	0.0366	0.0442	0.0624	0.1879	0.4495	0.7148	0.8875
		$p_{E,R}$	0.0182	0.0279	0.0341	0.0413	0.0497	0.0700	0.2060	0.4702	0.7311	0.8987
	Z_U	$p_{A,U}$	0.0220	0.0335	0.0407	0.0491	0.0587	0.0817	0.2287	0.4987	0.7554	0.9118
		$p_{CI,U}^{(\gamma=0.001)}$	0.0161	0.0252	0.0311	0.0379	0.0458	0.0651	0.1952	0.4584	0.7270	0.8980
		$p_{CI,U}^{(\gamma=0.005)}$	0.0152	0.0235	0.0289	0.0351	0.0424	0.0602	0.1851	0.4459	0.7109	0.8859
		$p_{E,U}$	0.0183	0.0280	0.0342	0.0415	0.0498	0.0703	0.2089	0.4793	0.7377	0.9003

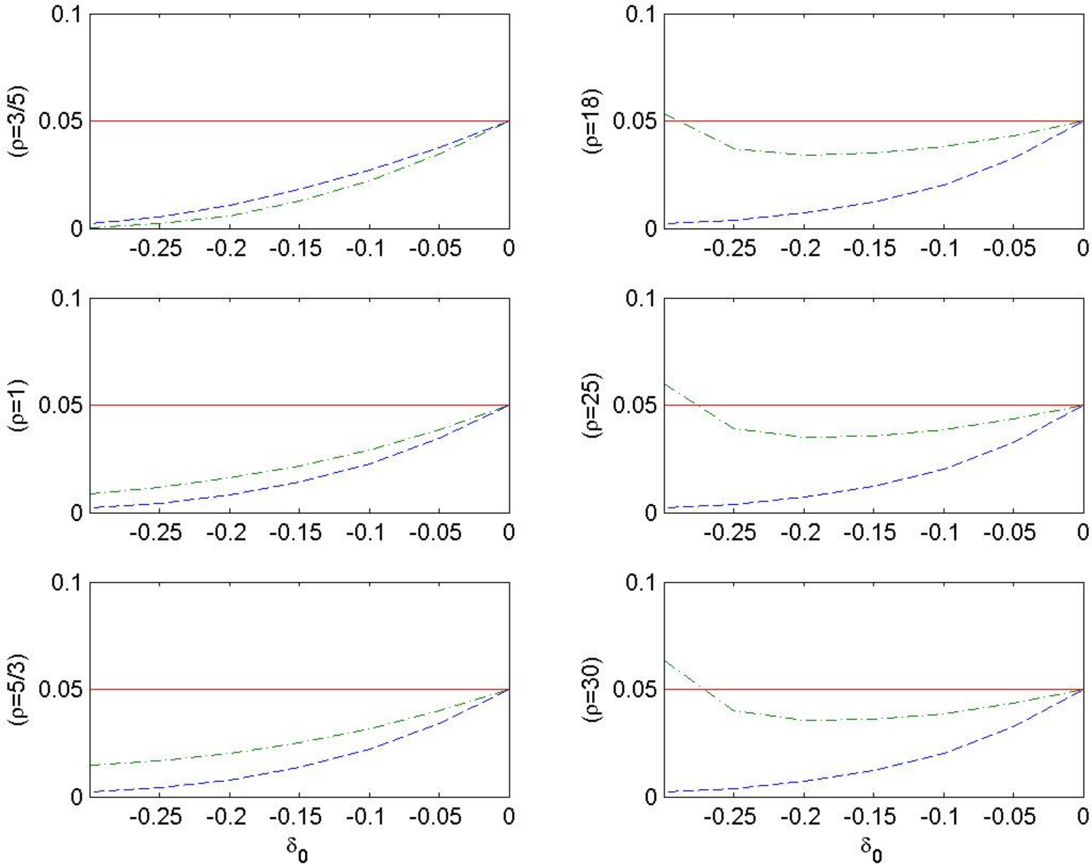

Figure 3.1: The asymptotic power function of the Z_R (dotted and dashed line) and the Z_U (dashed line) when $n_2 = 5, \lambda_2 = 0.3, \delta_0 = -0.3 : 0.05 : 0$, $\rho = 3/5, 1, 5/3$ in the left panel, $\rho = 18, 25, 30$ in the right panel.

Chapter 4

Non-inferiority Test

So far, our study focuses on identification of the superiority. It is sometimes unnecessary to draw such a strong conclusion. Instead, the non-inferiority test is of interest. For example (Lui, 2005), there are two air filter systems in an air pollution research, it is examined that the cheaper system is not inferior than the other one. That is, one aims to achieve the following alternative hypothesis,

$$H_a : \lambda_1 > \lambda_2 - \Delta_0.$$

In which, the non-inferiority limit Δ_0 is a positive real number and is predetermined by the investigators or experts of the related professional fields. In a clinical non-inferiority trial, it is commonly chosen as $0.2\lambda_2$ (Lui, 2005).

In next section, the Wald test statistics will be redefined first due to the presence of the non-zero non-inferiority limit. Their correspondent asymptotic testing procedures will be explored as well as two types of exact testing procedures later in this chapter. Regarding the asymptotic tests, we will derive their asymptotic distribution and power function for further verification on validity and unbiasedness. For the exact tests, the confidence-set p-value will be considered. It has been shown in Chapter 3 that once a test statistic satisfies the convexity condition, there is a great reduction in computation of a confidence-set p-value. The convexity of the two new-defined Wald test statistics will be justified in later section. On the other hand the estimated p-value will be applied for this problem. This chapter will end up with numerical studies on the type I error rate and power, as well as the sample size formulae of these proposed testing procedures.

4.1 Statistical Hypothesis and Test Statistics

Given some $\Delta_0 > 0$, consider the following hypothesis

$$\begin{cases} H_{03} : \lambda_1 \leq \lambda_2 - \Delta_0, \\ H_{a3} : \lambda_1 > \lambda_2 - \Delta_0. \end{cases}$$

The null space corresponding H_{03} is $\Omega_{03} = \{\lambda_1 \leq \lambda_2 - \Delta_0\}$, see Figure 2.1.

The Wald test statistic with respect to the non-inferiority test can be easily derived and has the following form:

$$Z_* = \frac{\hat{\delta} + \Delta_0}{se(\hat{\delta})},$$

where $\hat{\delta} = \bar{Y}_1 - \bar{Y}_2$ is the MLE of $\delta = \lambda_1 - \lambda_2$, and $se(\hat{\delta})$ is obtained by plugging some consistent estimators of λ_1, λ_2 in the standard error of $\hat{\delta}$. In this study, two estimators of λ_1, λ_2 are considered: The unconstrained and constrained MLE. The test statistic with the unconstrained estimator of the standard error can be easily seen and given as

$$Z_{U^*} = \frac{\bar{Y}_1 - \bar{Y}_2 + \Delta_0}{\sqrt{\frac{\hat{\lambda}_1}{n_1} + \frac{\hat{\lambda}_2}{n_2}}}.$$

On the other hand, the constrained MLE is solved by maximizing the likelihood

$$L(\lambda_1, \lambda_2) = Y_1 \ln \lambda_1 - n_1\lambda_1 + Y_2 \ln \lambda_2 - n_2\lambda_2,$$

subject to $\lambda_1 = \lambda_2 + \Delta_0$. The restricted MLE(RMLE) of λ_2 and λ_1 can be found as follows (see Appendix A.7 for details),

$$\tilde{\lambda}_2 = \frac{1}{2}\left\{\tilde{\lambda}_0 + \Delta_0 + \sqrt{(\tilde{\lambda}_0 + \Delta_0)^2 - \frac{4\hat{\lambda}_2\Delta_0}{1+\rho}}\right\}$$

and

$$\tilde{\lambda}_1 = \frac{1}{2}\left\{\tilde{\lambda}_0 - \Delta_0 + \sqrt{(\tilde{\lambda}_0 + \Delta_0)^2 - \frac{4\hat{\lambda}_2\Delta_0}{1+\rho}}\right\}.$$

Consequently, the Wald test statistic with the constrained estimator of the standard error is given as follows,

$$Z_{R^*} = \frac{\bar{Y}_1 - \bar{Y}_2 + \Delta_0}{\sqrt{\frac{\tilde{\lambda}_1}{n_1} + \frac{\tilde{\lambda}_2}{n_2}}}.$$

Note that in previous chapters, the two Wald test statistics correspondent to the superiority test are shown exactly of the same form when $\rho = 1$. However, the property is no longer true with respect to Z_{R^*} and Z_{U^*}.

4.2 Asymptotic p-values

First of all, consider the asymptotic testing procedures based on the two asymptotic p-values at some observed z_{R^*}, z_{U^*},

$$p_{A,R^*} = 1 - \Phi(z_{R^*}), \;\; p_{A,U^*} = 1 - \Phi(z_{U^*}).$$

The following theorem gives the asymptotic distributions of Z_{R^*} and Z_{U^*} in this Poisson problem. Subsequently, the correspondent asymptotic power function and the behavior of type I error rate of two asymptotic tests can be further investigated. Define $\delta^* = \lambda_1 - \lambda_2 + \Delta_0$ and δ_0^* be the correspondent true value.

Theorem 7. As $n_1,\ n_2 \to \infty$,

$$Z_{R^*}\sigma^* - \mu^* \xrightarrow{d} N(0,1), \quad Z_{U^*} - \mu^* \xrightarrow{d} N(0,1),$$

where

$$\sigma^{*2} = \frac{(1+\rho)\lambda_2 - \Delta_0 + \rho\delta_0^* + \sqrt{((1+\rho)\lambda_2 + \Delta_0 + \rho\delta_0^*)^2 - 4\lambda_2\Delta_0(1+\rho)}}{2((1+\rho)\lambda_2 - \Delta_0 + \delta_0^*)},$$

and

$$\mu^* = \frac{\delta_0^*}{\sqrt{\frac{\lambda_2(1+\rho)+\delta_0^*}{n_2\rho}}}.$$

By Theorem 7, we can show that the asymptotic tests of Z_{R^*} and Z_{U^*} have their power functions as follows,

$$\bar{\beta}_{Z_{R^*}}(\delta_0^*, \lambda_2, n_2, \rho, \Delta_0) = 1 - \Phi(z_\alpha\sigma^* - \mu^*), \tag{4.1}$$

and

$$\bar{\beta}_{Z_{U^*}}(\delta_0^*, \lambda_2, n_2, \rho, \Delta_0) = 1 - \Phi(z_\alpha - \mu^*), \tag{4.2}$$

approximately.

By (4.1) and (4.2), under $\delta_0^* = 0$, $\sigma^* = 1, \mu^* = 0$, and

$$\bar{\beta}_{Z_{R^*}} = \bar{\beta}_{Z_{U^*}} = 1 - \Phi(z_\alpha) = \alpha.$$

That is, the type I error rates of both tests achieve the significance level α at the boundary of the null parameter space. Further by (4.2), we can find that the maximal type I error rate of Z_{U^*} occurred at $\delta_0^* = 0$ and is equal to α. Hence the asymptotic test based on Z_{U^*} is a valid test asymptotically.

On the other hand, with a complicated component σ^* involved, the justification of validity of Z_{R^*} is less straight forward. By simple algebra the following inequality about σ^* can be shown,

$$\sqrt{\frac{(1+\rho)\lambda_2 - \Delta_0 + \rho\delta_0^*}{(1+\rho)\lambda_2 - \Delta_0 + \delta_0^*}} \leq \sigma^*. \tag{4.3}$$

When $\rho \leq 1$, from (4.3), then $\sigma^* \geq 1$, and hence $z_\alpha\sigma^* - \mu^* \geq z_\alpha$, for any $\delta_0^* < 0$. We can find that the maximum of $\bar{\beta}_{Z_{R^*}}$ occurred at $\delta_0^* = 0$ and is equal to α. Therefore, the type I error rate of Z_{R^*} is controlled at the significance level α, and the correspondent asymptotic p-value is asymptotically valid whenever $\rho \leq 1$. For example, Figure 4.1 gives the plots of the asymptotic type I error rate of Z_{R^*} versus δ_0^* at $\lambda_2 = 0.2, n_2 = 2, \Delta_0 = 0.2\lambda_2$ and $\alpha = 5\%$. The three plots on the left panel are correspondent to $\rho = 0.2, 0.5, 0.8$. One can see that the type I error rate increases with δ_0^* and the maximum, equal to α, occurs at the boundary $\delta_0^* = 0$. We further find that as long as ρ is not too unbalanced, the type I error rate can be still controlled. See the right panel of Figure 4.1 for $\rho = 1.2, 1.4, 1.6$. In contrast, when $\rho > 1$, the type I error rate can exceed the nominal level α especially when ρ is extremely large, and n_2 is relatively small. See the left panel of Figure 4.2 for the type I error rate of Z_{R^*} with $\rho = 1.7, 3, 5$ and $n_2 = 2$. However, as the sample sizes are slightly increased, the inflation of the type I error rate can be successfully improved. In the previous example, if n_2 is increased from 2 to 7, the type I error rates are then controlled within the level α, see the right panel of Figure 4.2. In summary, the asymptotic test based on Z_{R^*} is not always valid when the first group is extremely larger than the second group, $\rho >> 1$, and the group sizes are small.

Next we focus on the investigation on the power function of the two asymptotic testing procedures over the alternative parameter space. It can be easily shown that the power function of Z_{U^*} is always greater than or equal to α. That is, it is asymptotically unbiased. However, similar to previous results, the performance of Z_{R^*} is more complicated.

First, we examine the case that $\lambda_2 \to 0$. When one considers that Δ_0 is proportional to λ_2, the non-inferiority limit approaches to 0 as λ_2. Given δ_0^*, n_2, we can find that $\mu^* \to \sqrt{n_2\rho\delta_0^*}, \sigma^* \to \sqrt{\rho}$ as λ_2 approaches to 0, then we have

$$\lim_{\lambda_2 \to 0} \bar{\beta}_{Z_{R^*}} = 1 - \Phi(z_\alpha\sqrt{\rho} - \sqrt{n_2\rho\delta_0^*}).$$

As $\rho \leq 1$, the limit always exceeds α. But, it is not necessarily true when $\rho > 1$. In Figure 4.3, all the power functions $\bar{\beta}_{Z_{R^*}}$ are above the level $\alpha = 5\%$ when $\lambda_2 = 0.02, n_2 = 2$ and $\rho = 0.2, 0.4, 0.6, 0.8, 1, 1.2$. In the left panel of Figure 4.4, we see that the unbiasedness breaks down when ρ exceeds 1.3. Again, the problem can be improved with a slight increment in the sample size. In this example, the power function becomes no less than α when n_2 is increased from 2 to 7, see the right panel of Figure 4.4.

Next, we study the case that $\lambda_2 \to \infty$. It follows that $\mu^* \to 0, \sigma^* \to 1$ as λ_2 approaches to infinite given some δ_0^*, n_2. Then, the power converges to

$$\lim_{\lambda_2 \to \infty} \bar{\beta}_{Z_{R^*}} = 1 - \Phi(z_\alpha) = \alpha.$$

The limit is then independent with ρ as λ_2 approaches to infinite. For $\lambda_2 = 100, 200, n_2 = 2$ and $\rho = 0.5, 5, 50$, the power is found decreasing as λ_2 increases. And, all the powers are above the nominal level and increase as ρ increases, see Figure 4.5.

In summary, while the asymptotic test of Z_{U^*} is always unbiased, the power of the asymptotic test of Z_{R^*} may be below the nominal level when λ_2 is relatively small and ρ is larger than one.

Based on power function of a testing procedure, the necessary sample size for achievement of a prespecified power at some alternative setting at significance level can be further determined. Given ρ, to achieve a prespecified power level $1-\beta_0$ at λ_2, $\delta_0^* > 0$, the minimal sample size of the second group required for the Z_{R^*} and Z_{U^*} at significance level α is given as

$$n_{2,Z_{R^*}} \geq \left(\frac{z_\alpha + z_\beta}{\delta_0^*}\right)^2 \left\{\frac{\lambda_2(1+\rho) - \Delta_0 + \delta_0^*}{\rho}\right\}. \tag{4.4}$$

and,

$$n_{2,Z_{U^*}} \geq \left(\frac{z_\alpha + z_\beta}{\delta_0^*}\right)^2 \left\{\frac{\lambda_2(1+\rho) - \Delta_0 + \delta_0^*}{\rho}\right\}. \tag{4.5}$$

respectively. The size of the first group is fund as $n_1 = [n_2 \cdot \rho] + 1$.

4.3 Exact p-values

In testing the superiority, we have found that the confidence-set p-value has advantage of validity, and the revised estimated p-value has benefit of convenient use. Further both have satisfactory performances in numerical studies. Therefore, we adopt the two exact p-value in testing the non-inferiority. The exact testing procedures of Z_{U^*} and Z_{R^*} based on the correspondent confidence-set p-value and estimated p-value are proposed and studied. It is known that the null parameter space of a non-inferiority test is different from that of a superiority test. An exact p-value is defined as follows, given an observed z_0,

$$\begin{aligned} p^*_{(\lambda_1,\lambda_2)}(z_0) &= P(Z \geq z_0 | H_{03} : 0 < \lambda_1 \leq \lambda_2 - \Delta_0) \\ &= \sum_{y_1 \geq 0} \sum_{y_2 \geq 0} poi(y_1, n_1\lambda_1) poi(y_2, n_2\lambda_2) I_{\{Z \geq z_0\}}, \end{aligned}$$

where $poi(y, \lambda')$ is the probability function of Poisson distribution with mean λ', and y_1, y_2 are possible outcomes of Y_1, Y_2, respectively.

To solve for the computational difficulty brought by an infinite null parameter space, a confidence-set p-value is considered. The confidence-set p-value of Z_{R^*} is presented as follows,

$$p^{(\gamma)}_{CI,Z_{R^*}} = \sup_{(\lambda_1,\lambda_2)\in C^{**}_\gamma} P(Z_{R^*} \geq z_{R^*} | \lambda_1, \lambda_2) + \gamma,$$

and the confidence-set p-value of Z_{U^*} is presented as follows,

$$p^{(\gamma)}_{CI,Z_{U^*}} = \sup_{(\lambda_1,\lambda_2)\in C^{**}_\gamma} P(Z_{U^*} \geq z_{U^*} | \lambda_1, \lambda_2) + \gamma,$$

where C^{**}_γ is a $100(1-\gamma)\%$ confidence interval of λ_1 and λ_2 over the null parameter space Ω_{03}. Following from Chapter 3, we first consider the cross product set $C_{\gamma,0} = (L_1, U_1) \times (L_2, U_2)$, where (L_1, U_1) is the $100\sqrt{(1-\gamma)}\%$

confidence interval of λ_1 and (L_2, U_2) is the $100\sqrt{(1-\gamma)}\%$ confidence interval of λ_2. Here, the two exact interval estimates are applied,

$$(L_1, U_1) = \frac{1}{2n_1}\left(\chi^2_{(1-(1-\sqrt{1-\gamma})/2,\ 2Y_1)},\ \chi^2_{((1-\sqrt{1-\gamma})/2,\ 2(Y_1+1))}\right),$$

and

$$(L_2, U_2) = \frac{1}{2n_2}\left(\chi^2_{(1-(1-\sqrt{1-\gamma})/2,\ 2Y_2)},\ \chi^2_{((1-\sqrt{1-\gamma})/2,\ 2(Y_2+1))}\right).$$

Then $C_{\gamma,0}$ is a $100(1-\gamma)\%$ confidence interval of λ_1 and λ_2 in the unrestricted parameter space Ω. Subsequently, the confidence set C^{**}_{γ} is constructed as the intersection of the cross product set $C_{\gamma,0}$ and Ω_{03}. That is,

$$C^{**}_{\gamma} = C_{\gamma,0} \cap \Omega_{03} = \{L_1 \le \lambda_1 \le \min(U_1, \lambda_2 - \Delta_0),\ \ L_2 \le \lambda_2 \le U_2\}.$$

Note that when the observed interval estimate $C_{\gamma,0}$ is completely outside of Ω_{03}, C^{**}_{γ} is empty. In this case, we define $p_{CI} = \gamma < \alpha$, and reject the null hypothesis H_{03}.

It is known that once a test statistic satisfies the Barnard convexity condition, the computation of the correspondent confidence-set p-value can be further reduced due to the monotonic property of Poisson distribution. In the following, the two Wald test statistics are investigated to confirm whether they satisfy the Barnard convexity condition.

Theorem 8. Z_{R^*} satisfy the convexity condition.

The convexity of Z_{R^*} in Theorem 8 is shown from the monotonicity of Z_{R^*} with respect to Y_1 and Y_2, see Appendix A.9. As a consequence, from Theorem 8 and 5 of Chapter 3, the confidence-set p-value of Z_{R^*} is evaluated in the boundary of the confidence set C^{**}_{γ}. That is,

$$p^{(\gamma)}_{CI,Z_{R^*}} = \sup_{(\lambda_1,\lambda_2)\in\partial C^{**}_{\gamma}} P(Z_{R^*} \ge z_{R^*} | \lambda_1, \lambda_2) + \gamma,$$

where ∂C_γ^{**} is the boundary of C_γ^{**}. The associated confidence-set p-value based on Z_{R^*} can has its computation dramatically reduced. Furthermore, the probabilities $P(Z_{R^*} > z_{R^*} \mid \lambda_1, \lambda_2)$ can be shown to be increasing as λ_1 increases and λ_2 decreases, see proof of Theorem 5 of Chapter 3. Therefore, when C_γ^{**} is not empty, the supremum in $p_{CI,Z_{R^*}}$ either occurs at the point (U_1, L_2) or somewhere on the intersect of $C_{\gamma,0}$ and the line $\lambda_1 = \lambda_2 - \Delta_0$.

Next, to check the convexity condition on Z_{U^*}, we consider the partial derivative of Z_{U^*} w.r.t. Y_1 and Y_2 respectively,

$$\frac{\partial Z_{U^*}}{\partial Y_1} = \frac{\frac{1}{n_1^2}(\frac{Y_1}{n_1} + \frac{Y_2}{n_2} - \Delta_0) + \frac{2Y_2}{n_1 n_2^2}}{2(\frac{Y_1}{n_1^2} + \frac{Y_2}{n_2^2})\sqrt{\frac{Y_1}{n_1^2} + \frac{Y_2}{n_2^2}}}, \tag{4.6}$$

$$\frac{\partial Z_{U^*}}{\partial Y_2} = -\frac{\frac{Y_1}{n_1 n_2}(\frac{2}{n_1} + \frac{1}{n_2}) + \frac{Y_2}{n_2^3} + \frac{\Delta_0}{n_2^2}}{2(\frac{Y_1}{n_1^2} + \frac{y_2}{n_2^2})\sqrt{\frac{Y_1}{n_1^2} + \frac{y_2}{n_2^2}}}. \tag{4.7}$$

Since the numerator and denominator are both positive in (4.7), we can find that the partial derivative of Z_{U^*} w.r.t. Y_2 is negative. Then Z_{U^*} is decreasing in Y_2. But, (4.6) can not be showed always positive because $\frac{1}{n_1}(\frac{Y_1}{n_1} + \frac{Y_2}{n_2} - \Delta_0) + \frac{2Y_2}{n_2^2} < 0$ may occurs at some Y_1, Y_2 in the numerator of (4.6). Hence, one can not conclude the monotonicity of Z_{U^*} in Y_1. Several contour plots of $Z_{U^*} = k$ for k ranged from 2 to 10, are given in Figure 4.6-4.9 for $n_2 = 10, \rho = 3/5, \Delta_0 = 0.2, 2$. In Figure 4.6, the point marked symbol of star indicates a break down of monotonicity. One can see that the failure of the convexity condition Z_{U^*} is likely to occur for small Y_1, Y_2. Further, the content depends on the non-inferiority limit, Δ_0 and ρ. When $\Delta_0 = 0.2$, Z_{U^*} satisfies the convexity condition in the full sample space, see Figure 4.7. On the other hand, Figure 4.8 and 4.9 are the contour plots for $\rho = 1, 5/3$ and $\Delta_0 = 2$.

Since Z_{U^*} fails to satisfy the Barnard convexity, the supremum in $p_{CI,Z_{U^*}}$ may occur somewhere outside the boundary of C_γ^{**}. However, since it is observed that the break-down of convexity is not severe from the numerical study. For simplicity, we suggest to find the confidence-set p-value of Z_{U^*} at the boundary ∂C_γ^{**},

$$p_{CI,Z_{U^*}}^{(\gamma)} = \sup_{(\lambda_1,\lambda_2)\in\partial C_\gamma^{**}} P(Z_{R^*} \geq z_{R^*}|\lambda_1,\lambda_2) + \gamma.$$

Following Chapter 3.3, we propose a revised estimated p-value in testing the non-inferiority. The estimated exact p-values based on Z_{R^*} and Z_{U^*} are redefined as

$$p_{E,Z_{R^*}} = P(Z_{R^*} \geq z_{R^*}|\tilde{\lambda}_{13},\tilde{\lambda}_{23}),$$

and,

$$p_{E,Z_{U^*}} = P(Z_{U^*} \geq z_{U^*}|\tilde{\lambda}_{13},\tilde{\lambda}_{23}),$$

respectively. In which, $\tilde{\lambda}_{13}$ and $\tilde{\lambda}_{23}$ are some estimators of λ_1, λ_2 under the restricted null parameter space Ω_{03}. Again, similar to Chapter 3.3, we consider a revised RMLE. First, find the unrestricted MLE $\hat{\lambda}_1$ and $\hat{\lambda}_2$ of λ_1 and λ_2. If $\hat{\lambda}_1 \leq \hat{\lambda}_2 - \Delta_0$, then $\hat{\lambda}_1, \hat{\lambda}_2$ are exact the RMLEs under Ω_{03} and let $(\tilde{\lambda}_{13},\tilde{\lambda}_{23}) = (\hat{\lambda}_1,\hat{\lambda}_2)$. If $\hat{\lambda}_1 > \hat{\lambda}_2 - \Delta_0$, we consider the RMLE on the boundary $\lambda_1 = \lambda_2 - \Delta_0$, that is, $(\tilde{\lambda}_{13},\tilde{\lambda}_{23}) = (\tilde{\lambda}_1,\tilde{\lambda}_2)$. In summary,

$$(\tilde{\lambda}_{13},\tilde{\lambda}_{23}) = \begin{cases} (\hat{\lambda}_1,\hat{\lambda}_2), & \text{if } \hat{\lambda}_1 \leq \hat{\lambda}_2 - \Delta_0; \\ (\tilde{\lambda}_1,\tilde{\lambda}_2), & \text{if } \hat{\lambda}_1 > \hat{\lambda}_2 - \Delta_0. \end{cases}$$

In this chapter, the exact p-value bases on Z_{R^*} is increasing as λ_1 and decreasing as λ_2. respectively. The exact p-value is an increasing function as (λ_1, λ_2) moves toward at the down-right direction. Hence, adopting the MLE on the boundary leads to a more conservative conclusion. In next section, the performance of these proposed testing procedures will be compared through numerical studies.

As the Wald statistic depends on the data only through the two sufficient statistics (Y_1, Y_2), the exact power of the test correspondent to an exact p-value, p, is given by

$$\sum_{y_1 \geq 0} \sum_{y_2 \geq 0} poi(y_1, n_1\lambda_1) poi(y_2, n_2\lambda_2) I_{\{p \leq \alpha\}},$$

where p is either p_{CI}, p_E of Z_{R^*} or Z_{U^*}. Given a predetermined power level $1 - \beta_0$, at some specific $\lambda_2,\ \Delta_0,\ \delta_0$, the required sample size of the second group is the smallest integers such that the exact power achieves level,

$$n_2 = \min\{n_2 : \sum_{y_1 \geq 0} \sum_{y_2 \geq 0} poi(y_1, n_1\lambda_1) poi(y_2, n_2\lambda_2) I_{\{p \leq \alpha\}} \geq 1 - \beta_0\}, \quad (4.8)$$

for some $\rho > 0$. Further $n_1 = [n_2 \cdot \rho] + 1$.

4.4 Numerical Studies

Based on the two test statistics, Z_{U^*}, Z_{R^*}, the asymptotic test using the asymptotic p-value(denoted as p_A), and the two exact tests using the confidence-set p-value and the estimated p-value(denoted as p_E) are explored in this section. There are two confidence-set p-values constructed with $\gamma = 0.001, 0.005$, and denoted as $p_{CI,\cdot}^{(\gamma=0.001)}, p_{CI,\cdot}^{(\gamma=0.001)}$, respectively. Because the Wald statistics are function of the two sufficient statistics Y_1, Y_2, the powers of the associated tests can be directly calculated through their sampling distribution. In testing the non-inferiority, the maximal acceptable non-inferiority limit Δ_0 is chosen as $0.2\lambda_2$. We consider $n_2 = 10, \rho = 3/5, 1, 5/3, \alpha = 0.05$. The type I error rate and power of four test procedures are computed at true difference $\delta_0^* = \lambda_1 - \lambda_2 + \Delta_0$ ranged within -0.25 to 2 for $\lambda_2 = 1, 2$. Table 4.1 - 4.2 show the type I error rate and power calculated. On the other hand, the required sample sizes of the second group to achieve 80% power at $\delta_0^* = 0.6, 1$ are presented in Table 4.3 - 4.8. In which, only the results of the confidence-set p-value with $\gamma = 0.001$ are reported.

We first compare the two asymptotic tests in Table 4.1 and 4.2. The two tests have monotone increasing power with δ_0^*. As $\delta_0^* \leq 0$, the maximal type I error rates of Z_{R^*}, Z_{U^*} occur at $\delta_0^* = 0$, the boundary of the null parameter space for testing the non-inferiority. However, the finite sample results in Table 4.1 and 4.2 show that Z_{R^*} has more chance in rejecting the null hypothesis than Z_{U^*} when $\rho = 3/5 < 1$. It is contrary when $\rho \geq 1$. As $\delta_0^* = 0, \lambda_2 = 1$, the type I error rate of $Z_{R^*}(Z_{U^*})$ exceeds the significance level $\alpha = 0.05$ for $\rho = 3/5(5/3)$. As the mean value increases, the inflation of the type I error rate is reduced, but the improvement is not significant. In summary, Z_{U^*} is liberal as $\rho = 5/3, 1$, and Z_{R^*} is liberal as $\rho = 3/5, 1$.

Next the two exact p-values, p_{CI}, p_E are investigated. In last section, although we have justified numerically that due to the breakdown to the convexity condition, the supremum in p_{CI} of Z_{U^*} does not guarantee to occur at the boundary of the confidence-set. However, in this thesis, for simplicity we propose to search for the supremum over the boundary of the confidence-set. Here the supremums of the two confidence-set exact p-values are searched over 16 grids on the boundary of the truncated confidence-set in the null parameter space. From Table 4.1 to 4.2, we discover that the two exact p-values are always well-controlled at $\alpha = 0.05$. By Table 4.1 and 4.2, we find that the power of p_{CI} by using Z_{R^*} is greater than that of p_{CI} by using Z_{U^*}. The trend is not in accordance with that of the asymptotic tests. On the other hand, in applying the estimated p-value, the two test statistics Z_{R^*} and Z_{U^*} generate indifferent performances.

Table 4.3 - 4.8 present the required sample size of the second group for 80% power at $\delta_0^* = 0.6, 1.0$. And, based on the required sample sizes, the type I error rate at δ_0^*=-0.2,-0.1,-0.05,0, and the power at δ_0^*=0.6 or 1 of these tests are also reported. The results for the two asymptotic tests are based on the asymptotic sample size formulae (4.4) and (4.5). For the two exact tests, the figures are the minimal integers such that the exact power achieves the level by (4.8). All the tests need less sample size when δ_0^* increases, as expected. Between the two asymptotic tests, Z_{U^*} needs a slightly smaller sample than Z_{R^*} for $\rho > 1$. It is the contrary as $\rho < 1$. On the other hand, we find that the exact type I error rate of two asymptotic tests often exceeds the nominal level $\alpha = 5\%$ as $\rho \geq 1$. The inflation is more severe in the application of Z_{U^*}.

With the calculated sample size, every exact test achieves the prespecified power level and has a well-controlled type I error rate. Similarly, for the application of testing inferiority,the asymptotic sample sizes (4.4) and (4.5) can be regarded as an efficient alternative of (4.8) for the exact tests. A much quicker solution can be obtained and the result is found to be close to the exact sample size.

Table 4.1: Type I error rate and power of asymptotic p-value and exact p-value at $\lambda_2 = 1, n_2 = 10$, these p-values are based on test statistics Z_{R^*}, Z_{U^*} respectively.

ρ	Test Statistic	p-value	δ_0^*								
			-0.25	-0.15	-0.1	-0.05	0.0	0.5	1.0	1.5	2.0
3/5	Z_{R^*}	p_{A,R^*}	0.0140	0.0262	0.0344	0.0442	0.0557	0.2542	0.5301	0.7630	0.9024
		$p_{CI,R^*}^{(\gamma=0.001)}$	0.0098	0.0185	0.0245	0.0319	0.0406	0.2182	0.5010	0.7474	0.8951
		$p_{CI,R^*}^{(\gamma=0.005)}$	0.0096	0.0179	0.0237	0.0306	0.0388	0.2062	0.4841	0.7352	0.8886
		p_{E,R^*}	0.0103	0.0198	0.0264	0.0345	0.0441	0.2318	0.5136	0.7531	0.8968
	Z_{U^*}	p_{A,U^*}	0.0096	0.0177	0.0233	0.0300	0.0380	0.1978	0.4684	0.7223	0.8817
		$p_{CI,U^*}^{(\gamma=0.001)}$	0.0098	0.0185	0.0245	0.0319	0.0406	0.2182	0.5010	0.7474	0.8951
		$p_{CI,U^*}^{(\gamma=0.005)}$	0.0096	0.0177	0.0233	0.0300	0.0380	0.1979	0.4692	0.7242	0.8838
		p_{E,U^*}	0.0103	0.0198	0.0264	0.0345	0.0441	0.2318	0.5136	0.7531	0.8968
1	Z_{R^*}	p_{A,R^*}	0.0121	0.0233	0.0311	0.0405	0.0517	0.2726	0.6027	0.8465	0.9559
		$p_{CI,R^*}^{(\gamma=0.001)}$	0.0112	0.0212	0.0282	0.0368	0.0471	0.2614	0.5876	0.8361	0.9523
		$p_{CI,R^*}^{(\gamma=0.005)}$	0.0095	0.0185	0.0249	0.0329	0.0425	0.2482	0.5750	0.8279	0.9486
		p_{E,R^*}	0.0113	0.0212	0.0282	0.0368	0.0471	0.2617	0.5905	0.8404	0.9545
	Z_{U^*}	p_{A,U^*}	0.0134	0.0245	0.0322	0.0415	0.0526	0.2727	0.6027	0.8465	0.9559
		$p_{CI,U^*}^{(\gamma=0.001)}$	0.0095	0.0185	0.0249	0.0329	0.0425	0.2490	0.5802	0.8346	0.9521
		$p_{CI,U^*}^{(\gamma=0.005)}$	0.0080	0.0161	0.0219	0.0292	0.0383	0.2381	0.5627	0.8228	0.9477
		p_{E,U^*}	0.0113	0.0213	0.0282	0.0368	0.0472	0.2659	0.5983	0.8420	0.9537
5/3	Z_{R^*}	p_{A,R^*}	0.0090	0.0185	0.0254	0.0342	0.0449	0.2833	0.6603	0.9006	0.9807
		$p_{CI,R^*}^{(\gamma=0.001)}$	0.0092	0.0186	0.0256	0.0343	0.0449	0.2832	0.6588	0.8958	0.9776
		$p_{CI,R^*}^{(\gamma=0.005)}$	0.0087	0.0174	0.0238	0.0317	0.0416	0.2774	0.6451	0.8854	0.9756
		$p_{E,R}$	0.0106	0.0210	0.0285	0.0379	0.0493	0.2911	0.6614	0.8997	0.9797
	Z_{U^*}	p_{A,U^*}	0.0155	0.0294	0.0390	0.0508	0.0648	0.3346	0.7035	0.9187	0.9851
		$p_{CI,U^*}^{(\gamma=0.001)}$	0.0064	0.0140	0.0198	0.0273	0.0369	0.2774	0.6583	0.8958	0.9776
		$p_{CI,U^*}^{(\gamma=0.005)}$	0.0064	0.0140	0.0198	0.0273	0.0368	0.2676	0.6281	0.8810	0.9753
		p_{E,U^*}	0.0106	0.0210	0.0285	0.0379	0.0493	0.2911	0.6615	0.9006	0.9806

Table 4.2: Type I error rate and power of asymptotic p-value and exact p-value at $\lambda_2 = 2, n_2 = 10$, these p-values are based on test statistics Z_{R^*}, Z_{U^*} respectively.

ρ	Test Statistic	p-value	δ -0.25	-0.15	-0.1	-0.05	0.0	0.5	1	1.5	2
3/5	Z_{R^*}	p_{A,R^*}	0.0228	0.0327	0.0387	0.0454	0.0529	0.1759	0.3738	0.5917	0.7712
		$p_{CI,R^*}^{(\gamma=0.001)}$	0.0170	0.0251	0.0302	0.0360	0.0425	0.1575	0.3535	0.5740	0.7584
		$p_{CI,R^*}^{(\gamma=0.005)}$	0.0165	0.0244	0.0292	0.0348	0.0410	0.1510	0.3413	0.5600	0.7466
		p_{E,R^*}	0.0209	0.0302	0.0358	0.0422	0.0493	0.1667	0.3598	0.5768	0.7593
	Z_{U^*}	p_{A,U^*}	0.0176	0.0257	0.0307	0.0363	0.0427	0.1510	0.3364	0.5530	0.7408
		$p_{CI,U^*}^{(\gamma=0.001)}$	0.0170	0.0251	0.0302	0.0360	0.0425	0.1575	0.3535	0.5740	0.7584
		$p_{CI,U^*}^{(\gamma=0.005)}$	0.0153	0.0225	0.0270	0.0321	0.0379	0.1417	0.3282	0.5491	0.7403
		p_{E,U^*}	0.0209	0.0302	0.0358	0.0422	0.0493	0.1667	0.3598	0.5768	0.7593
1	Z_{R^*}	p_{A,R^*}	0.0183	0.0277	0.0336	0.0404	0.0482	0.1895	0.4297	0.6782	0.8539
		$p_{CI,R^*}^{(\gamma=0.001)}$	0.0178	0.0271	0.0329	0.0395	0.0471	0.1837	0.4184	0.6673	0.8473
		$p_{CI,R^*}^{(\gamma=0.005)}$	0.0159	0.0243	0.0296	0.0357	0.0427	0.1730	0.4046	0.6554	0.8396
		p_{E,R^*}	0.0183	0.0277	0.0336	0.0404	0.0481	0.1884	0.4256	0.6727	0.8506
	Z_{U^*}	p_{A,U^*}	0.0202	0.0303	0.0366	0.0438	0.0520	0.1951	0.4327	0.6791	0.8543
		$p_{CI,U^*}^{(\gamma=0.001)}$	0.0168	0.0259	0.0317	0.0383	0.0459	0.1831	0.4183	0.6673	0.8473
		$p_{CI,U^*}^{(\gamma=0.005)}$	0.0157	0.0241	0.0294	0.0356	0.0426	0.1730	0.4046	0.6554	0.8396
		p_{E,U^*}	0.0183	0.0277	0.0336	0.0404	0.0481	0.1884	0.4256	0.6727	0.8506
5/3	Z_{R^*}	p_{A,R^*}	0.0159	0.0251	0.0311	0.0381	0.0462	0.2029	0.4733	0.7366	0.9008
		$p_{CI,R^*}^{(\gamma=0.001)}$	0.0164	0.0256	0.0315	0.0385	0.0466	0.2030	0.4733	0.7368	0.9014
		$p_{CI,R^*}^{(\gamma=0.005)}$	0.0153	0.0243	0.0300	0.0366	0.0442	0.1854	0.4437	0.7195	0.8960
		p_{E,R^*}	0.0171	0.0264	0.0324	0.0394	0.0474	0.2032	0.4735	0.7382	0.9037
	Z_{U^*}	p_{A,U^*}	0.0220	0.0333	0.0404	0.0487	0.0583	0.2314	0.5096	0.7644	0.9153
		$p_{CI,U^*}^{(\gamma=0.001)}$	0.0155	0.0248	0.0308	0.0379	0.0461	0.2029	0.4733	0.7368	0.9014
		$p_{CI,U^*}^{(\gamma=0.005)}$	0.0143	0.0227	0.0280	0.0343	0.0415	0.1809	0.4384	0.7103	0.8899
		p_{E,U^*}	0.0164	0.0256	0.0315	0.0385	0.0466	0.2030	0.4735	0.7382	0.9037

Table 4.3: To achieve 80% power at $\delta_0^* = 0.6, \rho = 3/5$, the required sample size of the second group n_2 of the asymptotic p-values and exact p-value which are conducted at Z_{R^*}, Z_{U^*}. Based on the required samples n_2, the power and the type I error rate (in parentheses) are given at various δ_0^* in Ω_{03} .

Test Statistic	p-value		δ_0^*	λ_2 = 0.3	0.4	0.6	1	2
Z_{R^*}	p_{A,R^*}	n_2		25	29	37	53	93
			0.6	0.8132	0.8125	0.8076	0.7969	0.7960
			0	(0.0485)	(0.0527)	(0.0519)	(0.0507)	(0.0506)
			-0.05	(0.0237)	(0.0284)	(0.0291)	(0.0300)	(0.0312)
			-0.1	(0.0093)	(0.0134)	(0.0148)	(0.0166)	(0.0182)
			-0.2	(0.0004)	(0.0016)	(0.0026)	(0.0039)	(0.0052)
	$P_{CI,R^*}^{(\gamma=0.005)}$			27	32	39	58	100
			0.6	0.8166	0.8084	0.8003	0.8103	0.8067
			0	(0.0439)	(0.0437)	(0.0444)	(0.0445)	(0.0446)
			-0.05	(0.0207)	(0.0221)	(0.0239)	(0.0253)	(0.0266)
			-0.1	(0.0077)	(0.0094)	(0.0116)	(0.0133)	(0.0149)
			-0.2	(0.0003)	(0.0007)	(0.0018)	(0.0028)	(0.0039)
	P_{E,R^*}			25	29	39	55	98
			0.6	0.8039	0.8009	0.8113	0.8099	0.8125
			0	(0.0485)	(0.0474)	(0.0487)	(0.0500)	(0.0501)
			-0.05	(0.0237)	(0.0245)	(0.0266)	(0.0291)	(0.0304)
			-0.1	(0.0093)	(0.0108)	(0.0130)	(0.0157)	(0.0174)
			-0.2	(0.0007)	(0.0011)	(0.0020)	(0.0035)	(0.0047)
Z_{U^*}	p_{A,U^*}	n_2		29	33	41	57	97
			0.6	0.8298	0.8191	0.8125	0.8096	0.8035
			0	(0.0398)	(0.0408)	(0.0418)	(0.0455)	(0.0477)
			-0.05	(0.0182)	(0.0205)	(0.0222)	(0.0260)	(0.0289)
			-0.1	(0.0064)	(0.0089)	(0.0105)	(0.0137)	(0.0165)
			-0.2	(0.0002)	(0.0008)	(0.0015)	(0.0029)	(0.0045)
	$P_{CI,U^*}^{(\gamma=0.005)}$			27	32	40	58	100
			0.6	0.8087	0.8157	0.8131	0.8079	0.8067
			0	(0.0439)	(0.0416)	(0.0432)	(0.0443)	(0.0446)
			-0.05	(0.0207)	(0.0203)	(0.0230)	(0.0252)	(0.0266)
			-0.1	(0.0077)	(0.0082)	(0.0109)	(0.0133)	(0.0149)
			-0.2	(0.0003)	(0.0006)	(0.0016)	(0.0028)	(0.0039)
	P_{E,U^*}			25	29	39	55	97
			0.6	0.8039	0.8009	0.8113	0.8099	0.8105
			0	(0.0484)	(0.0474)	(0.0487)	(0.0500)	(0.0501)
			-0.05	(0.0236)	(0.0245)	(0.0266)	(0.0291)	(0.0304)
			-0.1	(0.0091)	(0.0108)	(0.0129)	(0.0157)	(0.0175)
			-0.2	(0.0002)	(0.0011)	(0.0020)	(0.0035)	(0.0048)

Table 4.4: To achieve 80% power at $\delta_0^* = 0.6, \rho = 1$, the required sample size of the second group n_2 of the asymptotic p-values and exact p-value which are conducted at Z_{R^*}, Z_{U^*}. Based on the required samples n_2, the power and the type I error rate (in parentheses) are given at various δ_0^* in Ω_{03}.

Test Statistic	p-value		δ_0^*	λ_2 = 0.3	0.4	0.6	1	2
Z_{R^*}	p_{A,R^*}	n_2		19	23	29	41	72
			0.6	0.8103	0.8152	0.8038	0.7976	0.7975
			0	(0.0501)	(0.0472)	(0.0509)	(0.0494)	(0.0500)
			-0.05	(0.0260)	(0.0247)	(0.0287)	(0.0293)	(0.0308)
			-0.1	(0.0116)	(0.0115)	(0.0147)	(0.0162)	(0.0180)
			-0.2	(0.0007)	(0.0016)	(0.0026)	(0.0039)	(0.0052)
	$p_{CI,R^*}^{(\gamma=0.005)}$			20	24	30	44	79
			0.6	0.8020	0.8102	0.8009	0.8081	0.8161
			0	(0.0412)	(0.0439)	(0.0440)	(0.0444)	(0.0448)
			-0.05	(0.0201)	(0.0228)	(0.0242)	(0.0256)	(0.0266)
			-0.1	(0.0079)	(0.0103)	(0.0121)	(0.0137)	(0.0149)
			-0.2	(0.0005)	(0.0012)	(0.0021)	(0.0030)	(0.0039)
	p_{E,R^*}			20	23	32	44	78
			0.6	0.8154	0.8118	0.8389	0.8225	0.8253
			0	(0.0480)	(0.0458)	(0.0498)	(0.0492)	(0.0498)
			-0.05	(0.0244)	(0.0243)	(0.0273)	(0.0286)	(0.0300)
			-0.1	(0.0106)	(0.0113)	(0.0135)	(0.0154)	(0.0171)
			-0.2	(0.0017)	(0.0014)	(0.0022)	(0.0035)	(0.0046)
Z_{U^*}	p_{A,U^*}	n_2		19	22	28	41	72
			0.6	0.8103	0.8045	0.7967	0.8023	0.8002
			0	(0.0503)	(0.0537)	(0.0539)	(0.0517)	(0.0507)
			-0.05	(0.0263)	(0.0301)	(0.0313)	(0.0308)	(0.0313)
			-0.1	(0.0122)	(0.0149)	(0.0167)	(0.0171)	(0.0183)
			-0.2	(0.0019)	(0.0021)	(0.0034)	(0.0042)	(0.0053)
	$p_{CI,U^*}^{(\gamma=0.005)}$			21	24	30	44	78
			0.6	0.8122	0.8089	0.8008	0.8081	0.8115
			0	(0.0221)	(0.0356)	(0.0416)	(0.0435)	(0.0448)
			-0.05	(0.0087)	(0.0175)	(0.0226)	(0.0249)	(0.0267)
			-0.1	(0.0031)	(0.0074)	(0.0111)	(0.0133)	(0.0150)
			-0.2	(0.0004)	(0.0007)	(0.0018)	(0.0029)	(0.0040)
	p_{E,U^*}			20	23	29	42	75
			0.6	0.8154	0.8118	0.8035	0.8051	0.8123
			0	(0.0451)	(0.0454)	(0.0485)	(0.0494)	(0.0498)
			-0.05	(0.0210)	(0.0236)	(0.0275)	(0.0290)	(0.0303)
			-0.1	(0.0076)	(0.0106)	(0.0142)	(0.0159)	(0.0175)
			-0.2	(0.0001)	(0.0010)	(0.0026)	(0.0038)	(0.0049)

Table 4.5: To achieve 80% power at $\delta_0^* = 0.6, \rho = 5/3$, the required sample size of the second group n_2 of the asymptotic p-values and exact p-value which are conducted at Z_{R^*}, Z_{U^*}. Based on the required samples n_2, the power and the type I error rate (in parentheses) are given at various δ_0^* in Ω_{03}.

Test Statistic	p-value		δ_0^*	λ_2 0.3	0.4	0.6	1	2
Z_{R^*}	p_{A,R^*}	n_2		16	19	24	34	60
			0.6	0.8026	0.8076	0.8034	0.7949	0.8009
			0	(0.0389)	(0.0444)	(0.0472)	(0.0504)	(0.0489)
			-0.05	(0.0198)	(0.0234)	(0.0267)	(0.0302)	(0.0301)
			-0.1	(0.0093)	(0.0108)	(0.0138)	(0.0169)	(0.0176)
			-0.2	(0.0007)	(0.0014)	(0.0028)	(0.0042)	(0.0051)
	$P_{CI,R^*}^{(\gamma=0.005)}$			17	20	25	36	64
			0.6	0.8187	0.8171	0.8062	0.8052	0.8102
			0	(0.0394)	(0.0434)	(0.0421)	(0.0422)	(0.0448)
			-0.05	(0.0192)	(0.0227)	(0.0238)	(0.0243)	(0.0269)
			-0.1	(0.0084)	(0.0103)	(0.0124)	(0.0131)	(0.0152)
			-0.2	(0.0006)	(0.0012)	(0.0023)	(0.0031)	(0.0041)
	P_{E,R^*}			16	19	24	35	63
			0.6	0.8109	0.8146	0.8119	0.8089	0.8197
			0	(0.0462)	(0.0460)	(0.0499)	(0.0499)	(0.0499)
			-0.05	(0.0232)	(0.0254)	(0.0286)	(0.0296)	(0.0304)
			-0.1	(0.0103)	(0.0126)	(0.0150)	(0.0164)	(0.0175)
			-0.2	(0.0036)	(0.0017)	(0.0030)	(0.0040)	(0.0049)
Z_{U^*}	p_{A,U^*}	n_2		13	16	21	31	57
			0.6	0.7797	0.7871	0.7866	0.7846	0.7960
			0	(0.0891)	(0.0733)	(0.0630)	(0.0595)	(0.0543)
			-0.05	(0.0574)	(0.0469)	(0.0388)	(0.0372)	(0.0342)
			-0.1	(0.0347)	(0.0279)	(0.0222)	(0.0219)	(0.0205)
			-0.2	(0.0116)	(0.0071)	(0.0056)	(0.0062)	(0.0063)
	$P_{CI,U^*}^{(\gamma=0.005)}$			17	20	25	36	63
			0.6	0.8064	0.8153	0.8062	0.8052	0.8054
			0	(0.0203)	(0.0347)	(0.0388)	(0.0421)	(0.0444)
			-0.05	(0.0095)	(0.0165)	(0.0213)	(0.0242)	(0.0267)
			-0.1	(0.0059)	(0.0062)	(0.0105)	(0.0130)	(0.0153)
			-0.2	(0.0019)	(0.0003)	(0.0017)	(0.0029)	(0.0042)
	P_{E,U^*}			16	19	24	35	62
			0.6	0.8027	0.8146	0.8119	0.8089	0.8140
			0	(0.0382)	(0.0460)	(0.0486)	(0.0499)	(0.0497)
			-0.05	(0.0186)	(0.0253)	(0.0273)	(0.0296)	(0.0304)
			-0.1	(0.0076)	(0.0126)	(0.0140)	(0.0164)	(0.0176)
			-0.2	(0.0002)	(0.0015)	(0.0028)	(0.0040)	(0.0050)

Table 4.6: To achieve 80% power at $\delta_0^* = 1.0, \rho = 3/5$, the required sample size of the second group n_2 of the asymptotic p-values and exact p-value which are conducted at Z_{R^*}, Z_{U^*}. Based on the required samples n_2, the power and the type I error rate (in parentheses) are given at various δ_0^* in Ω_{03}.

Test Statistic	p-value		δ_0^*	λ_2 = 0.3	0.4	0.6	1	2
Z_{R^*}	p_{A,R^*}	n_2		12	13	16	22	36
			1.0	0.8492	0.7810	0.7918	0.7993	0.7883
			0	(0.0608)	(0.0466)	(0.0572)	(0.0509)	(0.0512)
			-0.05	(0.0364)	(0.0309)	(0.0408)	(0.0364)	(0.0381)
			-0.1	(0.0185)	(0.0191)	(0.0278)	(0.0252)	(0.0277)
			-0.2	(0.0012)	(0.0057)	(0.0108)	(0.0108)	(0.0136)
	$P_{CI,R^*}^{(\gamma=0.005)}$			14	15	18	24	40
			1.0	0.8347	0.8286	0.8001	0.8050	0.8143
			0	(0.0371)	(0.0373)	(0.0437)	(0.0444)	(0.0444)
			-0.05	(0.0228)	(0.0243)	(0.0293)	(0.0310)	(0.0321)
			-0.1	(0.0122)	(0.0146)	(0.0184)	(0.0208)	(0.0226)
			-0.2	(0.0000)	(0.0033)	(0.0056)	(0.0082)	(0.0104)
	P_{E,R^*}			13	14	17	24	38
			1.0	0.8120	0.8050	0.8098	0.8189	0.8007
			0	(0.0407)	(0.0451)	(0.0510)	(0.0493)	(0.0502)
			-0.05	(0.0230)	(0.0289)	(0.0340)	(0.0349)	(0.0370)
			-0.1	(0.0106)	(0.0167)	(0.0212)	(0.0238)	(0.0266)
			-0.2	(0.0003)	(0.0031)	(0.0062)	(0.0098)	(0.0128)
Z_{U^*}	p_{A,U^*}	n_2		14	16	18	24	39
			1.0	0.8124	0.8317	0.7960	0.8037	0.8042
			0	(0.0262)	(0.0451)	(0.0359)	(0.0424)	(0.0453)
			-0.05	(0.0147)	(0.0281)	(0.0230)	(0.0295)	(0.0330)
			-0.1	(0.0073)	(0.0159)	(0.0138)	(0.0198)	(0.0234)
			-0.2	(0.0006)	(0.0034)	(0.0038)	(0.0079)	(0.0109)
	$P_{CI,U^*}^{(\gamma=0.005)}$			13	15	18	24	40
			1.0	0.8046	0.8285	0.8083	0.8042	0.8129
			0	(0.0365)	(0.0268)	(0.0412)	(0.0424)	(0.0444)
			-0.05	(0.0225)	(0.0154)	(0.0269)	(0.0295)	(0.0321)
			-0.1	(0.0120)	(0.0080)	(0.0165)	(0.0198)	(0.0226)
			-0.2	(0.0009)	(0.0014)	(0.0048)	(0.0079)	(0.0104)
	P_{E,U^*}			12	14	17	24	38
			1.0	0.8059	0.8095	0.8098	0.8189	0.8007
			0	(0.0226)	(0.0420)	(0.0510)	(0.0493)	(0.0502)
			-0.05	(0.0114)	(0.0264)	(0.0340)	(0.0349)	(0.0370)
			-0.1	(0.0046)	(0.0152)	(0.0212)	(0.0238)	(0.0266)
			-0.2	(0.0001)	(0.0029)	(0.0062)	(0.0098)	(0.0128)

Table 4.7: To achieve 80% power at $\delta_0^* = 1.0, \rho = 1$, the required sample size of the second group n_2 of the asymptotic p-values and exact p-value which are conducted at Z_{R^*}, Z_{U^*}. Based on the required samples n_2, the power and the type I error rate (in parentheses) are given at various δ_0^* in Ω_{03}.

Test Statistic	p-value		δ_0^*	λ_2 = 0.3	0.4	0.6	1	2
Z_{R^*}	p_{A,R^*}	n_2		9	10	13	17	28
			1.0	0.8152	0.8030	0.8149	0.7985	0.7936
			0	(0.0391)	(0.0504)	(0.0513)	(0.0478)	(0.0498)
			-0.05	(0.0225)	(0.0335)	(0.0351)	(0.0342)	(0.0370)
			-0.1	(0.0108)	(0.0212)	(0.0227)	(0.0238)	(0.0269)
			-0.2	(0.0004)	(0.0070)	(0.0077)	(0.0103)	(0.0133)
	$P_{CI,R^*}^{(\gamma=0.005)}$			10	11	14	18	30
			1.0	0.8318	0.8083	0.8246	0.8009	0.8038
			0	(0.0386)	(0.0369)	(0.0423)	(0.0440)	(0.0446)
			-0.05	(0.0221)	(0.0227)	(0.0287)	(0.0311)	(0.0326)
			-0.1	(0.0106)	(0.0128)	(0.0186)	(0.0212)	(0.0232)
			-0.2	(0.0004)	(0.0025)	(0.0063)	(0.0089)	(0.0110)
	P_{E,R^*}			10	10	13	18	29
			1.0	0.8446	0.8010	0.8141	0.8174	0.8064
			0	(0.0287)	(0.0377)	(0.0465)	(0.0492)	(0.0496)
			-0.05	(0.0138)	(0.0237)	(0.0320)	(0.0352)	(0.0366)
			-0.1	(0.0050)	(0.0137)	(0.0210)	(0.0243)	(0.0264)
			-0.2	(0.0000)	(0.0028)	(0.0074)	(0.0105)	(0.0128)
Z_{U^*}	p_{A,U^*}	n_2		9	10	12	17	28
			1.0	0.8163	0.8032	0.7950	0.8038	0.7969
			0	(0.0572)	(0.0658)	(0.0577)	(0.0532)	(0.0515)
			-0.05	(0.0403)	(0.0478)	(0.0417)	(0.0387)	(0.0384)
			-0.1	(0.0259)	(0.0336)	(0.0291)	(0.0273)	(0.0280)
			-0.2	(0.0034)	(0.0155)	(0.0126)	(0.0124)	(0.0139)
	$P_{CI,U^*}^{(\gamma=0.005)}$			10	11	14	18	30
			1.0	0.8154	0.8025	0.8246	0.8008	0.8038
			0	(0.0385)	(0.0314)	(0.0373)	(0.0416)	(0.0446)
			-0.05	(0.0221)	(0.0217)	(0.0238)	(0.0291)	(0.0326)
			-0.1	(0.0106)	(0.0145)	(0.0141)	(0.0198)	(0.0232)
			-0.2	(0.0004)	(0.0053)	(0.0036)	(0.0081)	(0.0110)
	P_{E,U^*}			10	10	13	18	29
			1.0	0.8446	0.8010	0.8091	0.8174	0.8064
			0	(0.0287)	(0.0377)	(0.0455)	(0.0488)	(0.0496)
			-0.05	(0.0138)	(0.0237)	(0.0316)	(0.0347)	(0.0366)
			-0.1	(0.0050)	(0.0137)	(0.0208)	(0.0239)	(0.0264)
			-0.2	(0.0000)	(0.0028)	(0.0073)	(0.0100)	(0.0128)

Table 4.8: To achieve 80% power at $\delta_0^* = 1.0, \rho = 5/3$, the required sample size of the second group n_2 of the asymptotic p-values and exact p-value which are conducted at Z_{R^*}, Z_{U^*}. Based on the required samples n_2, the power and the type I error rate (in parentheses) are given at various δ_0^* in Ω_{03}.

Test Statistic	p-value		δ_0^*	λ_2 0.3	0.4	0.6	1	2
Z_{R^*}	p_{A,R^*}	n_2		8	9	10	14	23
			1.0	0.8221	0.8264	0.7806	0.7906	0.7896
			0	(0.0380)	(0.0446)	(0.0427)	(0.0465)	(0.0479)
			-0.05	(0.0223)	(0.0288)	(0.0301)	(0.0334)	(0.0359)
			-0.1	(0.0103)	(0.0169)	(0.0206)	(0.0233)	(0.0263)
			-0.2	(0.0002)	(0.0032)	(0.0084)	(0.0104)	(0.0133)
	$P_{CI,R^*}^{(\gamma=0.005)}$			8	9	11	15	25
			1.0	0.8136	0.8086	0.8082	0.8011	0.8072
			0	(0.0378)	(0.0352)	(0.0420)	(0.0426)	(0.0443)
			-0.05	(0.0223)	(0.0227)	(0.0289)	(0.0304)	(0.0323)
			-0.1	(0.0103)	(0.0138)	(0.0191)	(0.0210)	(0.0230)
			-0.2	(0.0002)	(0.0030)	(0.0073)	(0.0091)	(0.0109)
	P_{E,R^*}			8	9	11	15	24
			1.0	0.8348	0.8296	0.8242	0.8242	0.8080
			0	(0.0380)	(0.0446)	(0.0489)	(0.0485)	(0.0497)
			-0.05	(0.0223)	(0.0288)	(0.0336)	(0.0345)	(0.0366)
			-0.1	(0.0103)	(0.0169)	(0.0220)	(0.0239)	(0.0263)
			-0.2	(0.0002)	(0.0032)	(0.0078)	(0.0104)	(0.0127)
Z_{U^*}	p_{A,U^*}	n_2		6	7	9	12	22
			1.0	0.7863	0.7829	0.7941	0.7721	0.7916
			0	(0.1250)	(0.0864)	(0.0728)	(0.0631)	(0.0567)
			-0.05	(0.0975)	(0.0724)	(0.0546)	(0.0482)	(0.0432)
			-0.1	(0.0684)	(0.0619)	(0.0399)	(0.0361)	(0.0323)
			-0.2	(0.0102)	(0.0450)	(0.0192)	(0.0188)	(0.0171)
	$P_{CI,U^*}^{(\gamma=0.005)}$			9	10	12	16	25
			1.0	0.8324	0.8326	0.8323	0.8151	0.8040
			0	(0.0467)	(0.0284)	(0.0329)	(0.0403)	(0.0443)
			-0.05	(0.0377)	(0.0204)	(0.0205)	(0.0278)	(0.0323)
			-0.1	(0.0291)	(0.0146)	(0.0116)	(0.0186)	(0.0230)
			-0.2	(0.0069)	(0.0056)	(0.0025)	(0.0075)	(0.0109)
	P_{E,U^*}			8	9	11	15	24
			1.0	0.8262	0.8296	0.8242	0.8243	0.8080
			0	(0.0378)	(0.0446)	(0.0489)	(0.0485)	(0.0497)
			-0.05	(0.0223)	(0.0288)	(0.0336)	(0.0345)	(0.0366)
			-0.1	(0.0103)	(0.0169)	(0.0219)	(0.0239)	(0.0263)
			-0.2	(0.0002)	(0.0032)	(0.0076)	(0.0104)	(0.0127)

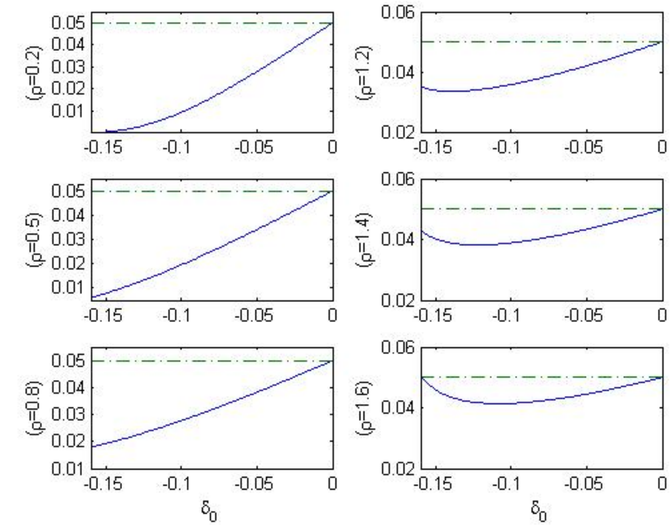

Figure 4.1: As $n_2 = 2, \lambda_2 = 0.2, \Delta_0 = 0.2\lambda_2, \rho = 0.2, 0.5, 0.8, 1.2, 1.4, 1.6, \delta_0 = -0.16 : 0.001 : 0$, the asymptotic type I error rate of the Z_{R^*}(solid line).

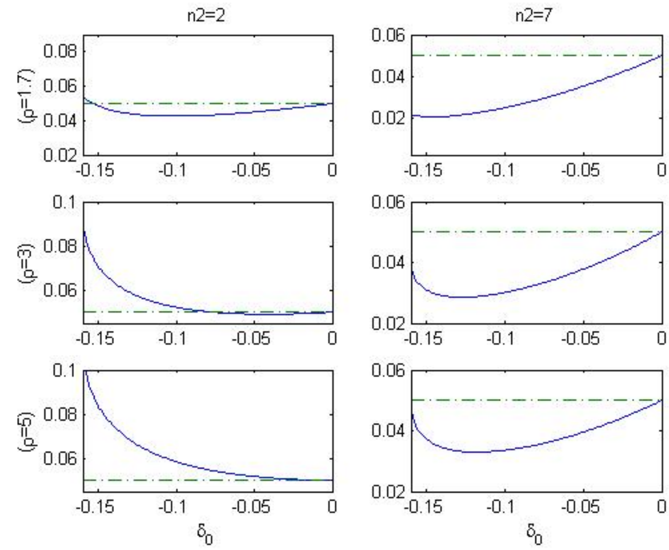

Figure 4.2: As $n_2 = 2, 7, \lambda_2 = 0.2, \Delta_0 = 0.2\lambda_2, \rho = 1.7, 3, 5, \delta_0^* = -0.16 : 0.001 : 0$, the asymptotic type I error rate of the Z_{R^*}(solid line).

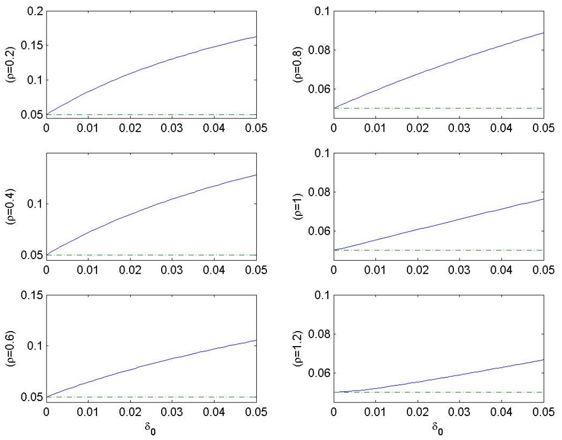

Figure 4.3: As $n_2 = 2, \lambda_2 = 0.02, \Delta_0 = 0.2\lambda_2, \rho = 0.2, 0.4, 0.6, 0.8, 1, 1.2, \delta_0^* = 0 : 0.001 : 0.05$, the asymptotic power of the Z_{R^*}(solid line).

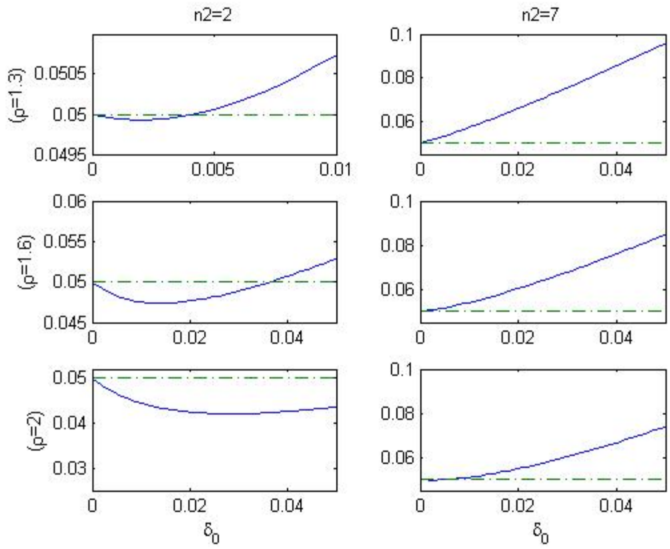

Figure 4.4: As $n_2 = 2, 7, \lambda_2 = 0.02, \Delta_0 = 0.2\lambda_2, \rho = 1.3, 1.6, 2, \delta_0^* = 0 : 0.001 : 0.05$, the asymptotic power of the Z_{R^*}(solid line).

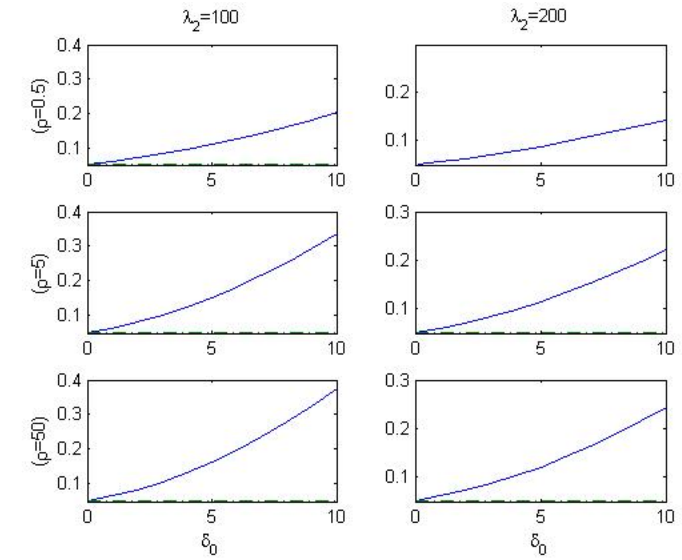

Figure 4.5: As $n_2 = 2, \lambda_2 = 100, 200, \Delta_0 = 0.2\lambda_2, \rho = 0.5, 5, 50, \delta_0^* = 0 : 1 : 10$, the asymptotic power of the Z_{R^*}(solid line).

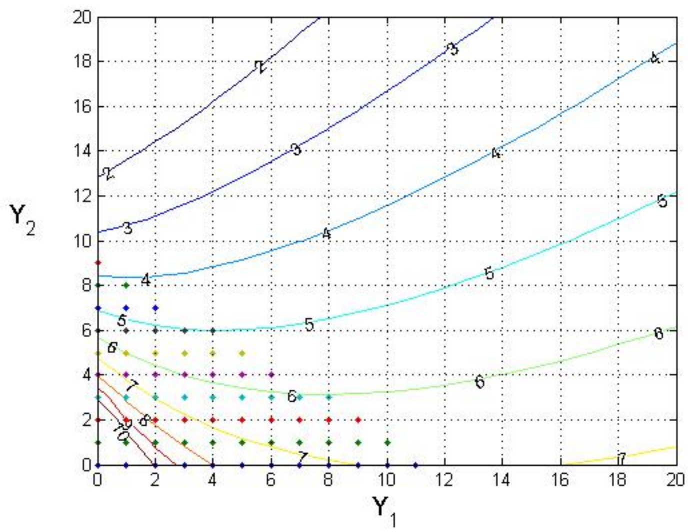

Figure 4.6: As $n_2 = 10, \Delta_0 = 2, \rho = 0.6$, a contour map of $Z_{U^*} = 2, 3, 4, 5, 6, 7, 8, 9, 10$.

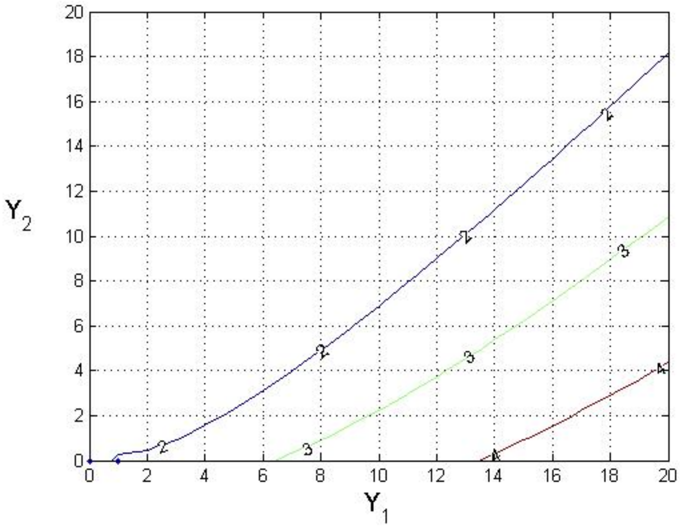

Figure 4.7: As $n_2 = 10; \Delta_0 = 0.2; \rho = 0.6$, a contour map of $Z_{U^*} = k$ for $k = 2, 3, 4$.

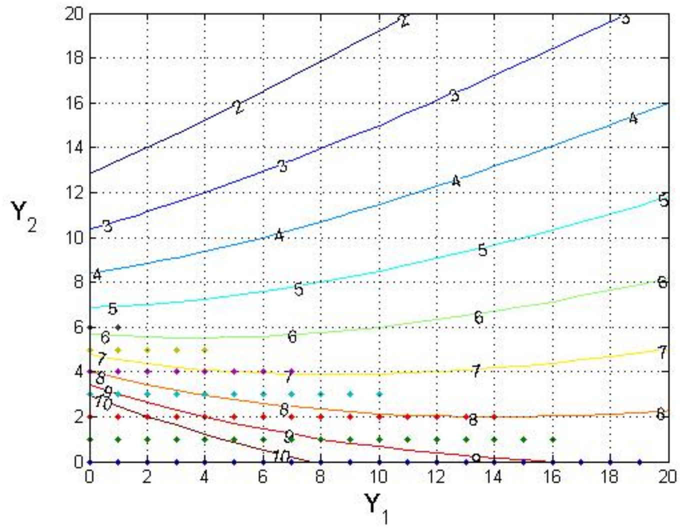

Figure 4.8: As $n_2 = 10; \Delta_0 = 2; \rho = 1$, a contour map of $Z_{U^*} = k$ for $k = 2, 3, 4, 5, 6, 7, 8, 9, 10$.

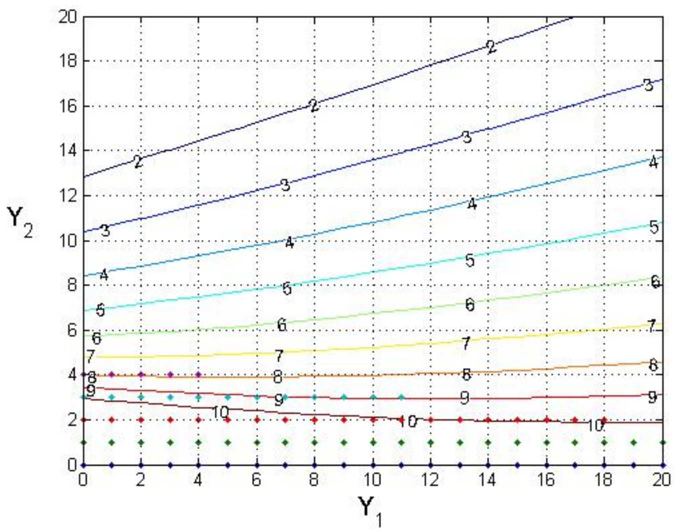

Figure 4.9: As $n_2 = 10; \Delta_0 = 2; \rho = 5/3$, a some contour map of $Z_{U^*} = k$ for $k = 2, 3, 4, 5, 6, 7, 8, 9, 10$.

Chapter 5

Real Example

5.1 Real Example

In this section, the methods introduced are applied to the breast cancer study described in Ng and Tang (2005). Female subjects were classified according to whether 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 patients had received X-ray; and λ_2 be the mean incidence number per person-year of the control group, in which patients were not examined by X-ray. Then we test the following hypothesis for establishing the superiority,

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

On the other hand, the procedures proposed can be easily extend to the case where every observation has a different experimental duration. Assume Y_{ij} be the Poisson random variable in the i-th group with m_{ij} units of duration, $i =$

$1, 2, j = 1, 2, \cdots, n_i$. Define $n_i^* = \sum_{j=1}^{n_i} m_{ij}, i = 1, 2$. Then n_i can be replaced by n_i^* in the test statistic, one can employ the approach straightforward.

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. It was found that $\hat{\lambda}_1 = 1.464, \hat{\lambda}_2 = 0.789$ and $\tilde{\lambda}_0 = 1.191$ per 1000 person-year. Consequently, $z_U = 2.2047, z_R = 2.0818$ with asymptotic p-value $0.0137, 0.0187$, respectively. The finding that the p-value of z_U is smaller than the p-value of z_R is consistent with our numerical results. When $\rho > 1$(here, 28010/19017=1.47), Z_U tends to have a more liberal conclusion than Z_R in an asymptotic test. The estimated p-value is evaluated at $\lambda_1 = \lambda_2 = \tilde{\lambda}_0 = 0.0011$. For the confidence-set p-value, the joint 99.9%(with $\gamma = 0.001$) confidence set of (λ_1, λ_2) is $\{0.0008 \leq \lambda_1 = \lambda_2 \leq 0.00177\}$. And the supremum of the p-value of Z_R occurs at $\lambda_1 = \lambda_2 = 0.0014$, and the supremum of the p-value of Z_U occurs at $\lambda_1 = \lambda_2 = 0.0010$. The calculated p-value are reported in Table 5.1. All these p-values are less than $\alpha = 0.05$ 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 5.1: The asymptotic, estimated and confidence-set p-value of the Wald Z-test Z_R, Z_U.

p-value	$Z_U = 2.2047$	$Z_R = 2.0818$
Asymptotic	0.0137	0.0187
Estimated	0.0186	0.0177
Confidence-set	0.0188	0.0182

Chapter 6

Concluding Remarks

In this study, we investigate several asymptotic and exact statistical procedures for comparing two Poisson means in identifying the superiority and non-inferiority. Two types of Wald test are considered, and they give different forms with respect to the superiority and the non-inferiority test respectively. The asymptotic power functions of the asymptotic procedures are derived and the correspondent asymptotic sample size formula are provided in the two testing problems. Consequently, the two asymptotic tests are compared in terms of the asymptotic power function and the required sample size. One concludes that the performances of the tests depend on the fraction of the group sizes. Moreover, the trends of asymptotic power function of testing superiority are consistent with that of testing non-inferiority. In this study, an exact test does not mean the use of the conventional p-value, which is denoted as the standard p-value in Chapter 2. The test is exact in the sense that the calculation of the p-value or is based on the exact sampling distribution of the test statistic. In fact, in the Poisson problem, the calculation of the exact standard $p-$value is rather difficult because the null parameter space is unbounded. Two alternative procedures, in which the computation of a p-value is taken either over a bounded space or a single point, are considered and proposed. The exact procedures under investigation are the confidence-set p-value and the estimated p-value. The definition and the computation of the exact p-values are introduced in details. The correspondent exact sample sizes for power requirement are shown to be found numerically. In this study, intensive numerical studies are provided and it is concluded that the asymptotic tests tend to have inflated type I error rates. On the contrary, the exact procedures have adequate performance overall, and dominate the asymptotic tests. Moreover, the quick solutions based on the asymptotic sample size formulae are found to provide good approximations to the exact sample sizes for testing superiority or non-inferiority.

The confidence-interval p-value is the sum of the supremum over a $100(1-\gamma)\%$ confidence region of the nuisance parameter(s) and γ. In which, the figure γ must be far less than the nominal level α. If not, the resultant p-value is easy to exceed α, and the test tends to give an insignificant conclusion. The testing procedure becomes powerless and is meaningless. Further, from Table 4.1 and 4.2, we find that the confidence interval p-values with $\gamma = 0.001$ are more powerful than that with $\gamma = 0.005$. It seems that the choice of γ affects the performance of the testing procedure. It is worthy to have more intensive investigations on the effect of γ in future study.

Bibliography

[1] Barnard, G. A.(1947) Significance Test for 2×2 Tables. *Biometrika*, **34**, 123-138.

[2] Berger, R. L. and Boos, D. D.(1994) P Values Maximized Over a Confidence Set for the Nuisance Parameter, *Journal of the American Statistical Association*, **89**, 1012-1016.

[3] Casella, G. and Berger, R. L.(1990) Statistical Inference. Pacific Grove, CA: Wadsworth.

[4] Corinna, M. and Jochen, M. C.(2005) Power Calcualtion for Non-inferiority Trials Comparing Two Poisson Distributions. *SAS Phuse Papers* , http://www.lexjansen.com/Phuse/2005/pk/pk01.pdf.

[5] Gail, M.(1974) Power Computations for Designing Comparative Poisson Trails. *Biometrics*, **30**, 231-237.

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

[7] Krishnamoorthy, K. and Thomson, J.(2004) More Power Test for Two Poisson Means. *Journal of Statistical Planning and Inference*, **119**, 23-35.

[8] Lehmann, E. L.(1986) *Testing Statistical Hypotheses*, **2nd edition**, Wiley, New York.

[9] Lui, K. J.(2005) Sample Size Calculation for Testing Non-inferiority and Equivalence Under Poisson Distribution. *Statistical Methodology*, **2**, 37-48.

[10] 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.

[11] Przyborowski, H. and Wilenski, H.(1940) Homogeneity of Results in Testing Samples from Poisson Series. *Biometrika*, **31**, 313-323.

[12] Pirie, W. R. and Hamdan, M. A.(1972) Some Revised Continuity Corrections for Discrete Distributions. *Biometrics*, **28**, 3, 693-701.

[13] Röhmel, J and Mansmann, U.(1990) 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.

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

[15] Storer, B. E. and Kim, C.(1990) Exact Properties of Some Exact Test Statistics for Comparing Two Binomial Proportions. *Journal of the American Statistical Association*, **85**, 409, 146-155.

[16] Song, J. X.(2009) Sample Size for Simultaneous Testing of Rate Differences in Non-inferiority Trials With Multiple Endpoints. *Computational Statistics and Data Analysis*, **53**, 1201-1207.

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

Appendix

A.1

Theorem 1. Let δ_0 be the true value of δ, and $\rho = n_1/n_2 \in (0,1)$ be the sample size fraction of the first group to the second group. As $n_1,\ n_2 \to \infty$,

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

Under $\lambda_1 = \lambda_2 = \lambda$, define the testing statistic

$$Z_R = \frac{\bar{Y}_1 - \bar{Y}_2}{\sqrt{\tilde{\lambda}_0(\frac{1}{n_1} + \frac{1}{n_2})}},$$

where $\tilde{\lambda}_0 = \frac{Y_1+Y_2}{n_1+n_2}$. By C. L. T, we have that

$$Z_R = \frac{\bar{Y}_1 - \bar{Y}_2}{\sqrt{\frac{\lambda_1}{n_1} + \frac{\lambda_2}{n_2}}} = \frac{\bar{Y}_1 - \bar{Y}_2}{\sqrt{\lambda(\frac{1}{n_1} + \frac{1}{n_2})}} \xrightarrow{d} N(0,1).$$

And,

$$\frac{Y_1}{n_1} = \hat{\lambda}_1 \xrightarrow{p} \lambda, \ \ \frac{Y_2}{n_2} = \hat{\lambda}_2 \xrightarrow{p} \lambda,$$

$$\frac{Y_1 + Y_2}{n_1 + n_2} \xrightarrow{p} \lambda,$$

$$\frac{\sqrt{\frac{Y_1+Y_2}{n_1+n_2}}}{\sqrt{\lambda}} \xrightarrow{p} 1.$$

Then,

$$
\begin{aligned}
Z_R &= \frac{\bar{Y}_1 - \bar{Y}_2}{\sqrt{\tilde{\lambda}_0(\frac{1}{n_1} + \frac{1}{n_2})}} \\
&= \frac{\frac{\bar{Y}_1 - \bar{Y}_2}{\sqrt{\lambda(\frac{1}{n_1}+\frac{1}{n_2})}}}{\frac{\sqrt{\frac{Y_1+Y_2}{n_1+n_2}}}{\sqrt{\lambda}}} \quad \text{(By Slutsky's theorem)} \\
&\xrightarrow{d} N(0,1).
\end{aligned}
$$

Hence, we have that

$$P(Z_R \geq z_\alpha | \lambda_1 = \lambda_2 = \lambda) \leq \alpha.$$

Therefore, the Z_R is valid.

Alternative, if $\lambda_1 > \lambda_2$ is true, the $\delta_0 = \lambda_1 - \lambda_2$ is defined. The following asymptotic distribution be hold,

$$Z^* = \frac{\bar{Y}_1 - \bar{Y}_2 - \delta_0}{\sqrt{\frac{\lambda_1}{n_1} + \frac{\lambda_2}{n_2}}} \xrightarrow{d} N(0,1).$$

And,

$$
\frac{Y_1}{n_1} = \hat{\lambda}_1 \xrightarrow{p} \lambda_1, \quad \frac{Y_2}{n_2} = \hat{\lambda}_2 \xrightarrow{p} \lambda_2,
$$

$$
\frac{Y_1 + Y_2}{n_1 + n_2} \xrightarrow{p} \frac{\rho\lambda_1 + \lambda_2}{1 + \rho},
$$

$$
\frac{\sqrt{\frac{Y_1+Y_2}{n_1+n_2}}}{\sqrt{\frac{\rho\lambda_1+\lambda_2}{1+\rho}}} \xrightarrow{p} 1.
$$

Hence, we have that

$$
\frac{\frac{\bar{Y}_1-\bar{Y}_2-\delta_0}{\sqrt{\frac{\lambda_1}{n_1}+\frac{\lambda_2}{n_2}}}}{\frac{\sqrt{\frac{Y_1+Y_2}{n_1+n_2}}}{\sqrt{\frac{\rho\lambda_1+\lambda_2}{1+\rho}}}} \xrightarrow{d} N(0,1).
$$

Subsequently, we can find that

$$\frac{\frac{\bar{Y}_1-\bar{Y}_2-\delta_0}{\sqrt{\frac{\lambda_1}{n_1}+\frac{\lambda_2}{n_2}}}}{\frac{\sqrt{\frac{Y_1+Y_2}{n_1+n_2}}}{\sqrt{\frac{\rho\lambda_1+\lambda_2}{1+\rho}}}} = \frac{\bar{Y}_1-\bar{Y}_2}{\sqrt{\frac{Y_1+Y_2}{n_1+n_2}}}\frac{\sqrt{\frac{\rho\lambda_1+\lambda_2}{1+\rho}}}{\sqrt{\frac{\lambda_1}{n_1}+\frac{\lambda_2}{n_2}}} - \frac{\delta_0}{\sqrt{\frac{\lambda_1}{n_1}+\frac{\lambda_2}{n_2}}}\frac{\sqrt{\frac{\rho\lambda_1+\lambda_2}{1+\rho}}}{\sqrt{\frac{Y_1+Y_2}{n_1+n_2}}}$$

$$= \frac{\bar{Y}_1-\bar{Y}_2}{\sqrt{\frac{Y_1+Y_2}{n_1+n_2}(\frac{1}{n_1}+\frac{1}{n_2})}}\frac{\sqrt{\frac{\rho\lambda_1+\lambda_2}{1+\rho}}\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}}{\sqrt{\frac{\lambda_1}{n_1}+\frac{\lambda_2}{n_2}}} - \frac{\delta_0}{\sqrt{\frac{\lambda_1}{n_1}+\frac{\lambda_2}{n_2}}}(1+o_p(1))$$

$$= Z_R\sqrt{\frac{\rho\lambda_1+\lambda_2}{\lambda_1+\rho\lambda_2}} - \frac{\delta_0}{\sqrt{\frac{\lambda_1+\rho\lambda_2}{n_2\rho}}}(1+o_p(1)).$$

Then, we have

$$Z_R\sigma-\mu_0 \xrightarrow{d} N(0,1), \quad \text{as} \quad n_1,n_2 \longrightarrow \infty,$$

where $\sigma=\sqrt{\frac{\rho\lambda_1+\lambda_2}{\lambda_1+\rho\lambda_2}}$, $\mu_0=\frac{\delta_0}{\sqrt{\frac{\lambda_1+\rho\lambda_2}{n_2\rho}}}$. We can find that the behavior of Z_R is the same as the T. And, the power function of the Z_R can be derived as follows,

$$\begin{aligned}\bar{\beta}_{Z_R}(\delta_0,\lambda_2,n_2,\rho) &= P(Z_R \geq \alpha|\delta_0)\\ &= P(Z_R\sigma-\mu_0 \geq z_\alpha\sigma-\mu_0|\delta_0)\\ &= 1-\Phi(z_\alpha\sigma-\mu_0).\end{aligned}$$

Given ρ,δ_0 and λ_2, the required sample sizes of the second group satisfied the power is greater than $1-\beta_0$ is

$$n^*_{2,Z_R} \geq \left(\frac{z_\alpha\sqrt{\frac{\lambda_2(1+\rho)+\rho\delta_0}{\lambda_2(1+\rho)+\delta_0}}+z_\beta}{\delta_0}\right)^2\frac{\lambda_2(1+\rho)+\delta_0}{\rho}.$$

$$Z_U=\frac{\hat{\delta}-\delta}{se(\hat{\delta})}=\frac{\hat{\delta}}{se(\hat{\delta})}\xrightarrow{d} N(0,1), \quad \text{as } n_1,n_2\to\infty,$$

where $se(\hat{\delta}) = \frac{\bar{Y}_1}{n_1} + \frac{\bar{Y}_2}{n_2}$. Hence,

$$P(Z_U \geq z_\alpha \mid H_0) \leq \alpha,$$

that is, the type I error is controlled at α asymptotically.

Under $\delta = \delta_0$, the variance of the MLE of δ, is given by

$$\sigma^2_{\hat{\delta}} = \frac{\lambda_1}{n_1} + \frac{\lambda_2}{n_2},$$

since $\bar{Y}_1$ and $\bar{Y}_2$ are independent. Then the estimated standard error satisfies

$$\frac{se(\hat{\delta})}{\sigma_{\hat{\delta}}} \xrightarrow{p} 1, \text{ provided } \lambda_1, \lambda_2 > 0.$$

Then according to the asymptotic normality of the MLE, Z's easily derived that for

$$Z_U - \frac{\delta_0}{\sigma^2_{\hat{\delta}}} \xrightarrow{d} N(0,1).$$

Hence the asymptotic power of Z can be derived too as

$$\begin{aligned} \bar{\beta}_{Z_U(\delta_0,\lambda_1,\rho,n_2)} &= P(Z_U \geq z_\alpha \mid \delta = \delta_0) \\ &= P(Z_U - \frac{\delta_0}{\sigma_{\hat{\delta}}} \geq z_\alpha - \frac{\delta_0}{\sigma_{\hat{\delta}}} \mid \delta_0) \\ &\approx 1 - \Phi(z_\alpha - \frac{\delta_0}{\sigma_{\hat{\delta}}}), \end{aligned}$$

where $\frac{\delta_0}{\sigma_{\hat{\delta}}} = \mu_0 = \frac{\delta_0}{\sqrt{\frac{\lambda_1+\rho\lambda_2}{n_2\rho}}}$.

Given ρ, δ_0 and λ_2, for the power greater than $1-\beta_0$, i.e.,

$$1 - \Phi(z_\alpha - \frac{\delta_0}{\sqrt{\frac{\lambda_2(1+\rho)+\delta_0}{n_2\rho}}}) \geq 1 - \beta_0,$$

the required size of the second group should satisfies

$$z_\alpha - \frac{\delta_0}{\sqrt{\frac{\lambda_2(1+\rho)+\delta_0}{n_2\rho}}} \leq -z_{\beta_0}.$$

The necessary asymptotic sample size is thus

$$n^*_{2,Z_U} \geq \left(\frac{z_\alpha + z_{\beta_0}}{\delta_0} \right)^2 \frac{\lambda_2(1+\rho) + \delta_0}{\rho}.$$

A.2

Theorem 2. Let δ_0 be the true value of δ, and $\rho = n_1/n_2$ be the sample size fraction of the first group to the second group. As $n_1,\ n_2 \to \infty$,

$$T\sigma - \mu \xrightarrow{d} N(0, 1).$$

Under $H_0 : \lambda_1 = \lambda_2 = \lambda$, for some $\lambda > 0$, We have

$$Y_{11}, \cdots, Y_{1n_1},\ \ Y_{21}, \cdots, Y_{2n_2} \sim Poi(\lambda).$$

Then

$$Y_1 = \sum_{i=1}^{n_1} Y_{1i} \sim Poi(n_1\lambda) \ \text{ and } \ Y_2 = \sum_{i=1}^{n_2} Y_{2i} \sim Poi(n_2\lambda).$$

By C.L.T., as $n_1, n_2 \to \infty$, and $n_1 = n_2\rho$,

$$\sqrt{n_1}(\bar{Y}_1 - \lambda) = \sqrt{n_2\rho}(\bar{Y}_1 - \lambda) \xrightarrow{d} N(0, \lambda),$$

and

$$\sqrt{n_2}(\bar{Y}_2 - \lambda) \xrightarrow{d} N(0, \lambda),$$

$$\sqrt{n_2}(\sqrt{\rho}(\bar{Y}_2 - \lambda)) \xrightarrow{d} N(0, \rho\lambda),$$

The sampling distribution of $\sqrt{n_2}(\bar{Y}_1 - \bar{Y}_2)$ can be derived straight forward

$$\sqrt{n_2}(\bar{Y}_1 - \bar{Y}_2) = \frac{1}{\sqrt{\rho}} \left\{ \sqrt{n_2\rho}(\bar{Y}_1 - \lambda) - \sqrt{n_2}(\sqrt{\rho}(\bar{Y}_2 - \lambda)) \right\} \xrightarrow{d} N\left(0, \lambda(1 + \frac{1}{\rho})\right),$$

and hence,

$$\sqrt{n_2}\frac{(\bar{Y}_1-\bar{Y}_2)}{\sqrt{\lambda(1+\frac{1}{\rho})}}=\frac{\bar{Y}_1-\bar{Y}_2}{\sqrt{\lambda(\frac{1}{n_1}+\frac{1}{n_2})}}\xrightarrow{d}N(0,1).$$

Let

$$s_1^2=\frac{\sum_{i=1}^{n_1}(Y_{1i}-\bar{Y}_1)^2}{n_1-1},\quad s_2^2=\frac{\sum_{i=1}^{n_2}(Y_{2i}-\bar{Y}_2)^2}{n_2-1},$$

and the polled variance estimate is,

$$s_p^2=\frac{(n_1-1)s_1^2+(n_2-1)s_2^2}{n_1+n_2-2}.$$

By WLLN, it can be shown that, as $n_1, n_2\to\infty$,

$$\frac{s_p^2}{\lambda}\xrightarrow{p}1,\quad s_p\xrightarrow{p}\sqrt{\lambda}.$$

By Slustky Theorem, the T-test statistic has an asymptotical standard normal distribution,

$$\frac{\bar{Y}_1-\bar{Y}_2}{s_p\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}}=\frac{\frac{\bar{Y}_1-\bar{Y}_2}{\sqrt{\lambda(\frac{1}{n_1}+\frac{1}{n_2})}}}{\frac{s_p}{\sqrt{\lambda}}}\xrightarrow{d}N(0,1).$$

On the other hand, it's known that as $n_1,\ n_2\to\infty$, we can have $t_{(n_1+n_2-2,\alpha)}\approx z_\alpha$. Consequently, the T-test has asymptotical level α, i.e.

$$P(T\geq t_{(n_1+n_2-2,\alpha)}\mid H_0)\approx P(Z\geq z_\alpha)\leq\alpha.$$

Under $H_1:\lambda_1-\lambda_2=\delta=\delta_0>0$, then we have

$$\frac{\bar{Y}_1-\bar{Y}_2-\delta_0}{\sqrt{\frac{\lambda_1}{n_1}+\frac{\lambda_2}{n_2}}}\xrightarrow{d}N(0,1),$$

as $n_1,\ n_2\to\infty$. Further since $s_1^2\xrightarrow{p}\lambda_1,\ s_2^2\xrightarrow{p}\lambda_2$, and $\rho=\frac{n_1}{n_2}$,

$$s_p^2=\frac{(n_1-1)s_1^2+(n_2-1)s_2^2}{n_1+n_2-2}\xrightarrow{p}\frac{\rho\lambda_1+\lambda_2}{1+\rho},\quad\frac{s_p}{\sqrt{\frac{\rho\lambda_1+\lambda_2}{1+\rho}}}\xrightarrow{p}1.$$

Hence

$$\frac{\frac{\bar{Y}_1-\bar{Y}_2-\delta_0}{\sqrt{\frac{\lambda_1}{n_1}+\frac{\lambda_2}{n_2}}}}{\frac{s_p}{\sqrt{\frac{\rho\lambda_1+\lambda_2}{1+\rho}}}} \xrightarrow{d} N(0,1).$$

Moreover, since $\lambda_1 = \lambda_2 + \delta_0$,

$$\begin{aligned}
\frac{\frac{\bar{Y}_1-\bar{Y}_2-\delta_0}{\sqrt{\frac{\lambda_1}{n_1}+\frac{\lambda_2}{n_2}}}}{\frac{s_p}{\sqrt{\frac{\rho\lambda_1+\lambda_2}{1+\rho}}}} &= \frac{\bar{Y}_1-\bar{Y}_2}{s_p}\frac{\sqrt{\frac{\rho\lambda_1+\lambda_2}{1+\rho}}}{\sqrt{\frac{n_2\lambda_1+n_1\lambda_2}{n_1 n_2}}} - \frac{\delta_0}{s_p}\frac{\sqrt{\frac{\rho\lambda_1+\lambda_2}{1+\rho}}}{\sqrt{\frac{\lambda_1}{n_1}+\frac{\lambda_2}{n_2}}} \\
&= \frac{\bar{Y}_1-\bar{Y}_2}{s_p\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}}\sqrt{\frac{(\rho\lambda_1+\lambda_2)n_1 n_2}{(1+\rho)(n_2\lambda_1+n_1\lambda_2)}}\sqrt{\frac{1}{n_1}+\frac{1}{n_2}} - \frac{\delta_0}{\sqrt{\frac{\lambda_1}{n_1}+\frac{\lambda_2}{n_2}}}(1+o_p(1)) \\
&= \frac{\bar{Y}_1-\bar{Y}_2}{s_p\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}}\sqrt{\frac{\rho\lambda_1+\lambda_2}{\lambda_1+\rho\lambda_2}} - \frac{\delta_0}{\sqrt{\frac{\lambda_1+\rho\lambda_2}{n_2\rho}}}(1+o_p(1)) \\
&= T\sigma_T - \mu_0(1+o_p(1)),
\end{aligned}$$

where $\sigma_T = \sqrt{\frac{\rho\lambda_1+\lambda_2}{\lambda_1+\rho\lambda_2}}$, $\mu_0 = \frac{\delta_0}{\sqrt{(\lambda_1+\rho\lambda_2)/n_2\rho}}$. Hence we have

$$T\sigma_T - \mu_0 \xrightarrow{d} N(0,1), \quad \text{as } n_1, n_2 \to \infty.$$

Then the asymptotic power of the T-test can be derived as,

$$\begin{aligned}
\bar{\beta}_T(\delta_0, \lambda_2, \rho, n_2) &= P(T \geq t_{(n_1+n_2-2,\alpha)} \mid \delta_0) \\
&= P(T\sigma_T - \mu_0 \geq t_{(n_1+n_2-2,\alpha)}\sigma_T - \mu_0) \\
&\approx 1 - \Phi(Z_\alpha\sigma_T - \mu_0).
\end{aligned}$$

Consequently, given ρ, δ_0 and λ_2, for the power greater than $1-\beta_0$,

$$1 - \Phi\left(z_\alpha\sqrt{\frac{\lambda_2(1+\rho)+\rho\delta_0}{\lambda_2(1+\rho)+\delta_0}} - \frac{\delta_0}{\sqrt{\frac{\lambda_2(1+\rho)+\delta_0}{n_2\rho}}}\right) \geq 1-\beta_0,$$

the required size of the second group should satisfies

$$z_\alpha\sqrt{\frac{\lambda_2(1+\rho)+\rho\delta_0}{\lambda_2(1+\rho)+\delta_0}} - \frac{\delta_0}{\sqrt{\frac{\lambda_2(1+\rho)+\delta_0}{n_2\rho}}} \leq z_{1-\beta_0} = -z_{\beta_0}.$$

The necessary asymptotic sample size is thus

$$n_2^* \geq \left(\frac{z_\alpha \sqrt{\frac{\lambda_2(1+\rho)+\rho\delta_0}{\lambda_2(1+\rho)+\delta_0}} + z_{\beta_0}}{\delta_0} \right)^2 \frac{\lambda_2(1+\rho)+\delta_0}{\rho}.$$

A.3

Theorem 3. For any n_1, n_2, the sampling distribution of $\hat{\delta}$ has equal spacings with space

$$b = \frac{1}{2m},$$

where m is the least common multiple of n_1, n_2.

Assume $n_1 = kn_2$, where $k = \frac{p}{q}$ is a fraction of two relatively prime integers, p, q, $(p, q) = 1$. Then $m = qn_1 = pn_2$ is the least common multiple(LCM) of the n_1, n_2. It's known that

$$\hat{\delta} = \bar{Y_1} - \bar{Y_2} = \frac{1}{n_1}\sum Y_{1i} - \frac{1}{n_2}\sum Y_{2i} = \frac{1}{m}(q\sum Y_{1i} - p\sum Y_{2i}).$$

Let $A = \{(t_1, t_2) : t_i \in N\bigcup\{0\}\}$ be the support of $(\sum Y_{1i}, \sum Y_{2i})$. The support of the estimator $\hat{\delta} = \frac{1}{m}(q\sum Y_{1i} - p\sum Y_{2i})$ can be obtained by considering all possible (t_1, t_2) in A. In the following, we first show that the support of $(q\sum Y_{1i} - p\sum Y_{2i})$ is exactly the set of integer values and thus has unity space, $b = 1$. Consequently, $\hat{\delta}$ has equal spacings with $b = \frac{1}{m}$.

For $(q, p) = 1$, there exist integers s_1 and s_2 to satisfy

$$qs_1 + ps_2 = 1, \tag{1}$$

where one of s_1, s_2 is negative (Yang and Yang, 1983). If $s_1 > 0$ and $s_2 < 0$, by letting $t_1 = s_1$ and $t_2 = -s_2$, one can rewrite (1) by

$$qt_1 - pt_2 = 1, \tag{2}$$

and $(t_1, t_2) \in A$. Multiplying (2) by -1, we have $-qt_1 + pt_2 = -1$. Adding the left hand side of the equation by $qp(t_1 + t_2) - pq(t_1 + t_2)$, then

$$q(t_1(p-1) + t_2 p) - p(t_2(q-1) + t_1 q) = -1, \text{ or } qt_1' - pt_2' = -1, \qquad (3)$$

where $t_1' = t_1(p-1) + t_2 p > 0$ and $t_2' = t_2(q-1) + t_1 q > 0$, and $(t_1', t_2') \in A$. That is, for any p, q, such that $(p, q) = 1$, there exist $(t_1, t_2), (t_1', t_2') \in A$ such that (2), (3) are true.

Similarly, when $s_1 < 0$ and $s_2 > 0$, we can find (t_1^*, t_2^*) and (t_1^{**}, t_2^{**}) in A such that $qt_1^* - pt_2^* = 1$ and $qt_1^{**} - pt_2^{**} = -1$. Hence, $(q \sum Y_{1i} - p \sum Y_{2i})$ has positive mass at $1, -1$ and their multiples. The support is the set of integers and the space $b = 1$. So, $\hat{\delta}$ has equal spacings with $b = \frac{1}{m}$. The continuity corrected Z-test and T-test are

$$Z_C = \frac{\hat{\delta} - \frac{1}{2m}}{se(\hat{\delta})}, \quad T_C = \frac{\hat{\delta} - \frac{1}{2m}}{s_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}.$$

A.4

Theorem 4. Let $C_\gamma^* = C_{\gamma,0} \cap \Omega_{02}$ be the truncated confidence set. Then

$$P((\lambda_1, \lambda_2) \in C_\gamma^* \mid \lambda_1, \lambda_2) \geq 1 - \gamma, \quad \text{for all } (\lambda_1, \lambda_2) \in \Omega_{02}.$$

First, we express C_γ^* as the following form,

$$L_1 \leq \lambda_1 \leq \min(U_1, \lambda_2), \quad L_2 \leq \lambda_2 \leq U_2.$$

Note that the two intervals are build on two independent statistics. Then, at any (λ_1, λ_2) such that $\lambda_1 \le \lambda_2$,

$$
\begin{aligned}
& P(L_1 \le \lambda_1 \le \min(U_1, \lambda_2), \;\; L_2 \le \lambda_2 \le U_2 | \lambda_1, \lambda_2) \\
= \;& P(L_1 \le \lambda_1 \le \min(U_1, \lambda_2) | \lambda_1, \lambda_2) P(L_2 \le \lambda_2 \le U_2 | \lambda_1, \lambda_2).
\end{aligned}
$$

And

$$
\begin{aligned}
& P(L_1 \le \lambda_1 \le \min(U_1, \lambda_2) | \lambda_1, \lambda_2) \\
= \;& P(L_1 \le \lambda_1 \le U_1, U_1 < \lambda_2 | \lambda_1, \lambda_2) + P(L_1 \le \lambda_1 \le \lambda_2, U_1 \ge \lambda_2 | \lambda_1, \lambda_2) \\
= \;& P(L_1 \le \lambda_1 \le U_1, U_1 < \lambda_2 | \lambda_1, \lambda_2) + P(L_1 \le \lambda_1, U_1 \ge \lambda_2 | \lambda_1, \lambda_2).
\end{aligned}
$$

Since under Ω_{02}, $\lambda_1 \le \lambda_2$, $\{U_1 \ge \lambda_2\}$ is a subset of $\{U_1 \ge \lambda_1\}$, and

$$
\begin{aligned}
& P(L_1 \le \lambda_1, U_1 \ge \lambda_2 | \lambda_1, \lambda_2) \\
= \;& P(L_1 \le \lambda_1, U_1 \ge \lambda_1, U_1 \ge \lambda_2 | \lambda_1, \lambda_2) \\
= \;& P(L_1 \le \lambda_1 \le U_1, U_1 \ge \lambda_2 | \lambda_1, \lambda_2).
\end{aligned}
$$

Consequently, for $\lambda_1 \le \lambda_2$,

$$
P(L_1 \le \lambda_1 \le \min(U_1, \lambda_2) | \lambda_1, \lambda_2) = P(L_1 \le \lambda_1 \le U_1 | \lambda_1, \lambda_2),
$$

and

$$
\begin{aligned}
& P(L_1 \le \lambda_1 \le \min(U_1, \lambda_2), \;\; L_2 \le \lambda_2 \le U_2 | \lambda_1, \lambda_2) \\
= \;& P(L_1 \le \lambda_1 \le U_1 | \lambda_1, \lambda_2) P(L_2 \le \lambda_2 \le U_2 | \lambda_1, \lambda_2) \\
\ge \;& 1 - \gamma.
\end{aligned}
$$

A.5

Theorem 5. Let S be a test statistic that depends on the data only through the two sufficient statistics (Y_1, Y_2) in comparing two Poisson means. Suppose

S satisfies the convexity condition. Then given s_0, the supremum of $P(S \geq s_0|\lambda_1, \lambda_2)$ occurs at a boundary point of the parameter space.

Consider the probability function of the Poisson distribution, $poi(x|\lambda)$, then

$$\frac{\partial}{\partial \lambda} P(X|\lambda) = P(X-1|\lambda) - P(X|\lambda), \quad \text{for } x = 1, 2, \cdots.$$

Given one the test statistic S and one observation (y_{10}, y_{20}), then the p-value is

$$P_S(\lambda_1, \lambda_2) = \sum_{S(y_1,y_2) \geq S_0(y_{10},y_{20})} P(y_1 \mid \lambda_1) P(y_2 \mid \lambda_2),$$

where, the $\{S(y_1, y_2) \geq S_0(y_{10}, y_{20})\}$ is rejection region, and, it can be rewritten as $\{(y_1, y_2) : S(y_1, y) \geq S_0\}$. We can derive a function $h : \{y_2 : 0, 1, 2, 3, \cdots\} \rightharpoonup \{y_1 : a, a+1, a+2, a+3, \cdots\}$ such that $\{(y_1, y_2) : S(y_1, y_2) \geq S_0\} = \{(y_1, y_2) : y_1 \geq h(y_2)\}$, and also can find the other function $h^* : \{y_1 : a, a+1, a+2, a+3, \cdots\} \rightharpoonup \{y_2 : 0, 1, 2, 3, \cdots\}$ such that $\{(y_1, y_2) : S(y_1, y_2) \geq S_0\} = \{(y_1, y_2) : y_2 \leq h^*(y_1)\}$. Hence, The P_S can be shown having the following two expressions.

$$\begin{aligned} P_S(\lambda_1, \lambda_2) &= \sum_{S(y_1,y_2) \geq S(y_{10},y_{20})} poi(y_1|\lambda_1) poi(y_2|\lambda_2) \\ &= \sum_{y_2=0}^{\infty} \sum_{y_1 \geq h(y_2)}^{\infty} poi(y_1|\lambda_1) poi(y_2|\lambda_2), \qquad (1) \end{aligned}$$

and

$$\begin{aligned} P_S(\lambda_1, \lambda_2) &= \sum_{S(y_1,y_2) \geq S(y_{10},y_{20})} poi(y_1|\lambda_1) poi(y_2|\lambda_2) \\ &= \sum_{y_1=a}^{\infty} \sum_{y_2 \leq h^*(y_1)} poi(y_1|\lambda_1) poi(y_2|\lambda_2), \qquad (2) \end{aligned}$$

by (1),

$$\begin{aligned}
P_S(\lambda_1,\lambda_2) &= poi(y_2=0|\lambda_2)\sum_{y_1\geq h(0)} poi(y_1|\lambda_1) \\
&+ poi(y_2=1|\lambda_2)\sum_{y_1\geq h(1)} poi(y_1|\lambda_1) \\
&+ \cdots \\
&+ poi(y_2=N_2|\lambda_2)\sum_{y_1\geq h(N_2)} poi(y_1|\lambda_1) \\
&+ \cdots,
\end{aligned}$$

and

$$\begin{aligned}
\frac{\partial P_S(\lambda_1,\lambda_2)}{\partial\lambda_2} &= -poi(y_2=0|\lambda_2)\sum_{y_1\geq h(0)} poi(y_1|\lambda_1) \\
&+ (poi(y_2=0|\lambda_2)-poi(y_2=1|\lambda_2))\sum_{y_1\geq h(1)} P(y_1|\lambda_1) \\
&+ (poi(y_2=1|\lambda_2)-poi(y_2|\lambda_2))\sum_{y_1\geq h(2)} P(y_1|\lambda_1) \\
&+ \cdots \\
&+ (poi(y_2=N_2-1|\lambda_2)-poi(y_2=N_2|\lambda_2))\sum_{y_1\geq h(N_2)} poi(y_1|\lambda_1) \\
&+ \vdots \\
&= -\sum_{y_2=0}^{\infty} poi(y_2|\lambda_2)poi(y_1=h(y_2)|\lambda_1) \\
&< 0.
\end{aligned}$$

By (2)

$$\begin{aligned}
P_S(\lambda_1,\lambda_2) &= poi(y_1=a|\lambda_1)\sum_{y_2\leq h^*(a)} poi(y_2|\lambda_2) \\
&+ poi(y_1=a+1|\lambda_1)\sum_{y_2\leq h^*(a+1)} poi(y_2|\lambda_2) \\
&+ \cdots \\
&+ poi(y_1=N_1|\lambda_1)\sum_{y_2\leq h^*(N_1)} poi(y_2|\lambda_2),
\end{aligned}$$

and

$$
\begin{aligned}
& \frac{\partial P_S(\lambda_1, \lambda_2)}{\partial \lambda_1} \\
= \; & (poi(y_1 = a-1|\lambda_1) - poi(y_1 = a|\lambda_1)) \sum_{y_2 \le h^*(a)} poi(y_2|\lambda_2) \\
& + (poi(y_1 = a|\lambda_1) - poi(y_1 = a+1|\lambda_1)) \sum_{y_2 \le h^*(a+1)} poi(y_2|\lambda_2) \\
& + (poi(y_1 = a|\lambda_1) - poi(y_1 = a+1|\lambda_1)) \sum_{y_2 \le h^*(a+1)} poi(y_2|\lambda_2) \\
& + \cdots \\
& + (poi(y_1 = N_1|\lambda_1) - poi(y_1 = N_1|\lambda_1)) \sum_{y_2 \le h^*(N_1)} poi(y_2|\lambda_2) \\
= \; & (poi(y_1 = a-1|\lambda_1) - poi(y_1 = a|\lambda_1)) poi(y_2 = 0|\lambda_2) \\
& + (poi(y_1 = a|\lambda_1) - poi(y_1 = a+1|\lambda_1))(poi(y_2 = 0|\lambda_2) + poi(y_2 = 1|\lambda_2)) \\
& + (poi(y_1 = a+1|\lambda_1) - poi(y_1 = a+2|\lambda_1)) \\
& (poi(y_2 = 0|\lambda_2) + poi(y_2 = 1|\lambda_2) + poi(y_{=}2|\lambda_2)) \\
& + (poi(y_1 = a+2|\lambda_1) - poi(y_1 = a+3|\lambda_1)) \\
& (poi(y_2 = 0|\lambda_2) + poi(y_2 = 1|\lambda_2) + poi(y_2 = 2|\lambda_2) + poi(y_2 = 3|\lambda_2)) \\
& + \cdots \\
& + (poi(y_1 = N_1 - 1|\lambda_1) - poi(y_1 = N_1|\lambda_1)) \\
& (poi(y_2 = 0|\lambda_2) + poi(y_2 = 1|\lambda_2) + poi(y_2 = 2|\lambda_2) + \cdots + poi(y_2 = h^*(N_1))) \\
= \; & \sum_{y_1 = a}^{N_1} (poi(y_1|\lambda_1) - poi(N_1|\lambda_1)) poi(y_2 = h^*(y_1)|\lambda_2) \\
> \; & 0,
\end{aligned}
$$

where $N_1 \to \infty$, such that $poi(N_1|\lambda_1) \doteq 0$. Moreover, the space of the null hypothesis H_{02} is a compact set, then the supremum of the P_θ is maximum can be shown in the Poisson problem.

A.6

Theorem 6. Z_R, Z_U satisfy the convexity condition.

For the test statistic

$$Z_U(Y_1, Y_2) = \frac{\frac{Y_1}{n_1} - \frac{Y_2}{n_2}}{\sqrt{\frac{Y_1}{n_1^2} + \frac{Y_2}{n_2^2}}},$$

we have

$$\begin{aligned}
\frac{\partial Z_U(Y_1, Y_2)}{\partial Y_1} &= \frac{\frac{1}{n_1}\sqrt{\frac{Y_1}{n_1^2} + \frac{Y_2}{n_2^2}} - (\frac{Y_1}{n_1} - \frac{Y_2}{n_2})\frac{\frac{1}{n_1^2}}{2\sqrt{\frac{Y_1}{n_1^2}+\frac{Y_2}{n_2^2}}}}{\frac{Y_1}{n_1^2} + \frac{Y_2}{n_2^2}} \\
&= \frac{\frac{2}{n_1}(\frac{Y_1}{n_1^2} + \frac{Y_2}{n_2^2}) - \frac{1}{n_1^2}(\frac{Y_1}{n_1} - \frac{Y_2}{n_2})}{2(\frac{Y_1}{n_1^2} + \frac{Y_2}{n_2^2})\sqrt{\frac{Y_1}{n_1^2} + \frac{Y_2}{n_2^2}}} \\
&= \frac{\frac{Y_1}{n_1^3} + \frac{Y_2}{n_1 n_2}(\frac{1}{n_1} + \frac{2}{n_2})}{2(\frac{Y_1}{n_1^2} + \frac{Y_2}{n_2^2})\sqrt{\frac{Y_1}{n_1^2} + \frac{Y_2}{n_2^2}}} \\
&> 0.
\end{aligned}$$

and

$$\begin{aligned}
\frac{\partial Z_U(Y_1, Y_2)}{\partial Y_2} &= \frac{-\frac{1}{n_2}\sqrt{\frac{Y_1}{n_1^2} + \frac{Y_2}{n_2^2}} - (\frac{Y_1}{n_1} - \frac{Y_2}{n_2})\frac{\frac{1}{n_2^2}}{2\sqrt{\frac{Y_1}{n_1^2}+\frac{Y_2}{n_2^2}}}}{\frac{Y_1}{n_1^2} + \frac{Y_2}{n_2^2}} \\
&= \frac{\frac{-2}{n_2}(\frac{Y_1}{n_1^2} + \frac{Y_2}{n_2^2}) - \frac{1}{n_2^2}(\frac{Y_1}{n_1} - \frac{Y_2}{n_2})}{2(\frac{Y_1}{n_1^2} + \frac{Y_2}{n_2^2})\sqrt{\frac{Y_1}{n_1^2} + \frac{Y_2}{n_2^2}}} \\
&= -\frac{\frac{Y_1}{n_1 n_2}(\frac{2}{n_1} + \frac{1}{n_2}) + \frac{Y_2}{n_2^3}}{2(\frac{Y_1}{n_1^2} + \frac{Y_2}{n_2^2})\sqrt{\frac{Y_1}{n_1^2} + \frac{Y_2}{n_2^2}}} \\
&< 0,
\end{aligned}$$

So $Z_U(Y_1, Y_2)$ is increasing in Y_1, and decreasing in Y_2, hence we have $Z_U(Y_1, Y_2) \leq Z_U(Y_1+1, Y_2)$, and $Z_U(Y_1, Y_2) \leq Z_U(Y_1, Y_2-1)$. The $Z_U(Y_1, Y_2)$ satisfies convexity condition.

For the test statistic

$$Z_R(Y_1, Y_2) = \frac{\frac{Y_1}{n_1} - \frac{Y_2}{n_2}}{\sqrt{(\frac{1}{n_1} + \frac{1}{n_2})(\frac{Y_1+Y_2}{n_1+n_2})}}$$

, we have

$$\begin{aligned}
\frac{\partial Z_R(Y_1, Y_2)}{\partial Y_1} &= \frac{\frac{1}{n_1}\sqrt{(\frac{1}{n_1} + \frac{1}{n_2})(\frac{Y_1+Y_2}{n_1+n_2})} - (\frac{Y_1}{n_1} - \frac{Y_2}{n_2})\frac{(\frac{1}{n_1}+\frac{1}{n_2})(\frac{1}{n_1+n_2})}{2\sqrt{(\frac{1}{n_1}+\frac{1}{n_2})(\frac{Y_1}{n_1}+\frac{Y_2}{n_2})}}}{(\frac{1}{n_1} + \frac{1}{n_2})(\frac{Y_1+Y_2}{n_1+n_2})} \\
&= \frac{\frac{1}{n_1}(\frac{1}{n_1} + \frac{1}{n_2})(\frac{Y_1+Y_2}{n_1+n_2}) - (\frac{Y_1}{n_1} - \frac{Y_2}{n_2})(\frac{1}{n_1} + \frac{1}{n_2})(\frac{1}{n_1+n_2})}{2(\frac{1}{n_1} + \frac{1}{n_2})(\frac{Y_1+Y_2}{n_1+n_2})\sqrt{(\frac{1}{n_1} + \frac{1}{n_2})(\frac{Y_1+Y_2}{n_1+n_2})}} \\
&= \frac{\frac{Y_2}{n_1+n_2}(\frac{1}{n_1} + \frac{1}{n_2})}{2(\frac{Y_1+Y_2}{n_1+n_2})\sqrt{(\frac{1}{n_1} + \frac{1}{n_2})(\frac{Y_1+Y_2}{n_1+n_2})}} \\
&= \frac{\frac{Y_2}{n_1 n_2}}{2(\frac{Y_1+Y_2}{n_1+n_2})\sqrt{(\frac{1}{n_1} + \frac{1}{n_2})(\frac{Y_1+Y_2}{n_1+n_2})}} \\
&> 0,
\end{aligned}$$

and

$$\begin{aligned}
\frac{\partial Z_R(Y_1, Y_2)}{\partial Y_2} &= \frac{\frac{-1}{n_2}\sqrt{(\frac{1}{n_1} + \frac{1}{n_2})(\frac{Y_1+Y_2}{n_1+n_2})} - (\frac{Y_1}{n_1} - \frac{Y_2}{n_2})\frac{(\frac{1}{n_1}+\frac{1}{n_2})(\frac{1}{n_1+n_2})}{2\sqrt{(\frac{1}{n_1}+\frac{1}{n_2})(\frac{Y_1}{n_1}+\frac{Y_2}{n_2})}}}{(\frac{1}{n_1} + \frac{1}{n_2})(\frac{Y_1+Y_2}{n_1+n_2})} \\
&= -\frac{\frac{1}{n_1}(\frac{1}{n_1} + \frac{1}{n_2})(\frac{Y_1+Y_2}{n_1+n_2}) + (\frac{Y_1}{n_1} - \frac{Y_2}{n_2})(\frac{1}{n_1} + \frac{1}{n_2})(\frac{1}{n_1+n_2})}{2(\frac{1}{n_1} + \frac{1}{n_2})(\frac{Y_1+Y_2}{n_1+n_2})\sqrt{(\frac{1}{n_1} + \frac{1}{n_2})(\frac{Y_1+Y_2}{n_1+n_2})}} \\
&= -\frac{\frac{Y_1}{n_1+n_2}(\frac{1}{n_1} + \frac{1}{n_2})}{2(\frac{Y_1+Y_2}{n_1+n_2})\sqrt{(\frac{1}{n_1} + \frac{1}{n_2})(\frac{Y_1+Y_2}{n_1+n_2})}} \\
&= -\frac{\frac{Y_1}{n_1 n_2}}{2(\frac{Y_1+Y_2}{n_1+n_2})\sqrt{(\frac{1}{n_1} + \frac{1}{n_2})(\frac{Y_1+Y_2}{n_1+n_2})}} \\
&< 0,
\end{aligned}$$

hence Z_R is increasing in Y_1 and decreasing in Y_2, then it can be provided $Z_R(Y_1, Y_2) \leq Z_R(Y_1+1, Y_2)$ and $Z_R(Y_1, Y_2) \leq Z_R(Y_1, Y_2-1)$, hence the Z_R satisfies the convexity condition.

A.7

The derivation of restricted MLE of λ_1 and λ_2 on H_{03}.

The constrained MLE maximizes the following likelihood function,

$$L(\lambda_1, \lambda_2) = Y_1 \ln \lambda_1 - n_1\lambda_1 + Y_2 \ln \lambda_2 - n_2\lambda_2, \quad \text{subject to} \quad \lambda_1 = \lambda_2 - \Delta_0.$$

The likelihood function can be rewritten to as the following function of λ_2,

$$L(\lambda_2) = Y_1 \ln(\lambda_2 - \Delta_0) - n_1(\lambda_2 - \Delta_0) + Y_2 \ln \lambda_2 - n_2\lambda_2,$$

taking the derivative of the likelihood function $L(\lambda_2)$ with respect to λ_2, we have

$$\frac{\partial L(\lambda_2)}{\partial \lambda_2} = \frac{Y_1}{\lambda_2 - \Delta_0} - n_1 + \frac{Y_2}{\lambda_2} - n_2 = 0.$$

The RMLE of λ_2 satisfies

$$(n_1+n_2)\lambda_2^2 - [(n_1+n_2)\Delta_0 + Y_1 + Y_2]\lambda_2 + Y_2\Delta_0 = 0,$$

and hence we have possible multiple solutions of λ_2,

$$\frac{[(n_1+n_2)\Delta_0 + Y_1 + Y_2] \pm \sqrt{[(n_1+n_2)\Delta_0 + Y_1 + Y_2]^2 - 4(n_1+n_2)Y_2\Delta_0}}{2(n_1+n_2)},$$

The solution with negative squared term leads to a negative RMLE of λ_1 and thus is not a valid RMLE. In the following, we have the RMLEs of λ_2 and λ_1 on $\lambda_1 = \lambda_2 - \Delta_0$. Define

$$\hat{\lambda}_1 = \frac{Y_1}{n_1}, \quad \hat{\lambda}_2 = \frac{Y_2}{n_2}, \quad \tilde{\lambda}_0 = \frac{Y_1+Y_2}{n_1+n_2} = \frac{\rho}{1+\rho}\hat{\lambda}_1 + \frac{1}{1+\rho}\hat{\lambda}_2.$$

Then the RMLEs are

$$\tilde{\lambda}_2 = \frac{1}{2}\left\{\tilde{\lambda}_0 + \Delta_0 + \sqrt{(\tilde{\lambda}_0 + \Delta_0)^2 - 4\frac{1}{1+\rho}\hat{\lambda}_2\Delta_0}\right\}$$

and

$$\tilde{\lambda}_1 = \frac{1}{2}\left\{\tilde{\lambda}_0 - \Delta_0 + \sqrt{(\tilde{\lambda}_0 + \Delta_0)^2 - 4\frac{1}{1+\rho}\hat{\lambda}_2\Delta_0}\right\}.$$

A.8

Theorem 7. As $n_1,\ n_2 \to \infty$,

$$Z_{R^*}\sigma^* - \mu^* \xrightarrow{d} N(0,1), \quad Z_{U^*} - \mu^* \xrightarrow{d} N(0,1),$$

where

$$\sigma^{*2} = \frac{(1+\rho)\lambda_2 - \Delta_0 + \rho\delta_0^* + \sqrt{((1+\rho)\lambda_2 + \Delta_0 + \rho\delta_0^*)^2 - 4\lambda_2\Delta_0(1+\rho)}}{2((1+\rho)\lambda_2 - \Delta_0 + \delta_0^*)},$$

and

$$\mu^* = \frac{\delta_0^*}{\sqrt{\frac{\lambda_2(1+\rho)+\delta_0^*}{n_2\rho}}}.$$

Under $\lambda_1 = \lambda_2 - \Delta_0$, the restricted MLE of λ_1 and λ_2 can be derived as follows,

$$\tilde{\lambda}_2 = \frac{1}{2}\left\{\tilde{\lambda}_0 + \Delta_0 + \sqrt{(\tilde{\lambda}_0 + \Delta_0)^2 - 4\frac{1}{1+\rho}\hat{\lambda}_2\Delta_0}\right\}$$

and

$$\tilde{\lambda}_1 = \frac{1}{2}\left\{\tilde{\lambda}_0 - \Delta_0 + \sqrt{(\tilde{\lambda}_0 + \Delta_0)^2 - 4\frac{1}{1+\rho}\hat{\lambda}_2\Delta_0}\right\}.$$

And the testing statistic is defined as follows,

$$Z_{R^*} = \frac{\bar{Y}_1 - \bar{Y}_2 + \Delta_0}{\sqrt{\frac{\tilde{\lambda}_1}{n_1} + \frac{\tilde{\lambda}_2}{n_2}}}$$

Similarly, under $\lambda_1 = \lambda_2 - \Delta_0$ we have that

$$Z^{**} = \frac{\bar{Y}_1 - \bar{Y}_2 + \Delta_0}{\sqrt{\frac{\lambda_1}{n_1} + \frac{\lambda_2}{n_2}}} = \frac{\bar{Y}_1 - \bar{Y}_2 + \Delta_0}{\sqrt{\frac{\lambda_2 - \Delta_0}{n_1} + \frac{\lambda_2}{n_2}}} = \frac{\bar{Y}_1 - \bar{Y}_2 + \Delta_0}{\sqrt{\lambda_2(\frac{1}{n_1} + \frac{1}{n_2}) - \frac{\Delta_0}{n_1}}} \xrightarrow{d} N(0,1).$$

We know that

$$\tilde{\lambda}_0 \xrightarrow{p} \lambda_2 - \frac{\rho\Delta_0}{1+\rho},$$

and

$$\sqrt{(\tilde{\lambda}_0 + \Delta_0)^2 - 4\frac{1}{1+\rho}\hat{\lambda}_2\Delta_0} \xrightarrow{p} \lambda_2 - \frac{\Delta_0}{1+\rho},$$

$$\tilde{\lambda}_2 \xrightarrow{p} \lambda_2,$$

and

$$\frac{\tilde{\lambda}_2}{\lambda_2} \xrightarrow{p} 1,$$

then

$$\frac{\sqrt{\tilde{\lambda}_2(\frac{1}{n_1} + \frac{1}{n_2}) - \frac{\Delta_0}{n_2\rho}}}{\sqrt{\lambda_2(\frac{1}{n_1} + \frac{1}{n_2}) - \frac{\Delta_0}{n_2\rho}}} \xrightarrow{p} 1.$$

Hence, we have as follows

$$\begin{aligned}
Z_{R^*} &= \frac{\bar{Y}_1 - \bar{Y}_2 + \Delta_0}{\sqrt{\frac{\tilde{\lambda}_1}{n_1} + \frac{\tilde{\lambda}_2}{n_2}}} \\
&= \frac{\bar{Y}_1 - \bar{Y}_2 + \Delta_0}{\sqrt{\frac{\tilde{\lambda}_2 - \Delta_0}{n_1} + \frac{\tilde{\lambda}_2}{n_2}}} \\
&= \frac{\bar{Y}_1 - \bar{Y}_2 + \Delta_0}{\sqrt{\tilde{\lambda}_2(\frac{1}{n_1} + \frac{1}{n_2}) - \frac{\Delta_0}{n_2\rho}}} \\
&= \frac{\frac{\bar{Y}_1 - \bar{Y}_2 + \Delta_0}{\sqrt{\lambda_2(\frac{1}{n_1} + \frac{1}{n_2}) - \frac{\Delta_0}{n_2\rho}}}}{\sqrt{\frac{\tilde{\lambda}_2(\frac{1}{n_1} + \frac{1}{n_2}) - \frac{\Delta_0}{n_2\rho}}{\lambda_2(\frac{1}{n_1} + \frac{1}{n_2}) - \frac{\Delta_0}{n_2\rho}}}} \\
&\xrightarrow{d} N(0,1).
\end{aligned}$$

Therefor, we can derive that

$$P(Z_{R^*} \geq z_\alpha | \lambda_1 = \lambda_2 - \Delta_0) \leq \alpha,$$

then the validity of Z_{R^*} is hold.

As the alternative of H_{03} is true, that is $\lambda_1 > \lambda_2 - \Delta_0$, we have as follows,

$$Z^{***} = \frac{\bar{Y}_1 - \bar{Y}_2 + \Delta_0 - \delta_0^*}{\sqrt{\frac{\lambda_1}{n_1} + \frac{\lambda_2}{n_2}}} \xrightarrow{d} N(0,1),$$

where $\delta_0^* = \lambda_1 - \lambda_2 + \Delta_0$. Because $\hat{\lambda}_1 \xrightarrow{p} \lambda_1$ and $\hat{\lambda}_2 \xrightarrow{p} \lambda_2$, we can derive the following the limit converge form (Jun Shao, 1998)

$$\tilde{\lambda}_1 \xrightarrow{p} q_1(\lambda_1, \lambda_2),\ \tilde{\lambda}_2 \xrightarrow{p} q_2(\lambda_1, \lambda_2),$$

where

$$q_1(\lambda_1, \lambda_2) = \frac{1}{2}\left\{\frac{\rho\lambda_1 + \lambda_2}{1+\rho} - \Delta_0 + \sqrt{(\frac{\rho\lambda_1 + \lambda_2}{1+\rho} + \Delta_0)^2 - 4\frac{\lambda_2\Delta_0}{1+\rho}}\right\},$$

and

$$q_2(\lambda_1, \lambda_2) = \frac{1}{2}\left\{\frac{\rho\lambda_1 + \lambda_2}{1+\rho} + \Delta_0 + \sqrt{(\frac{\rho\lambda_1 + \lambda_2}{1+\rho} + \Delta_0)^2 - 4\frac{\lambda_2\Delta_0}{1+\rho}}\right\}.$$

Further, we have

$$\frac{\tilde{\lambda}_1}{n_1} + \frac{\tilde{\lambda}_2}{n_2} \xrightarrow{p} \frac{q_1(\lambda_1, \lambda_2)}{n_1} + \frac{q_2(\lambda_1, \lambda_2)}{n_2},$$

and

$$\frac{\sqrt{\frac{\tilde{\lambda}_1}{n_1} + \frac{\tilde{\lambda}_2}{n_2}}}{\sqrt{\frac{q_1(\lambda_1,\lambda_2)}{n_1} + \frac{q_2(\lambda_1,\lambda_2)}{n_2}}} \xrightarrow{p} 1$$

Then, the limit distribution can be derived as follows

$$\frac{\frac{\bar{Y}_1 - \bar{Y}_2 + \Delta_0 - \delta_0^*}{\sqrt{\frac{\lambda_1}{n_1} + \frac{\lambda_2}{n_2}}}}{\frac{\sqrt{\frac{\tilde{\lambda}_1}{n_1} + \frac{\tilde{\lambda}_2}{n_2}}}{\sqrt{\frac{q_1(\lambda_1,\lambda_2)}{n_1} + \frac{q_2(\lambda_1,\lambda_2)}{n_2}}}} \xrightarrow{d} N(0,1).$$

Next, the above equation can be rewritten as follows

$$
\begin{aligned}
&\frac{\frac{\bar{Y}_1-\bar{Y}_2+\Delta_0-\delta_0^*}{\sqrt{\frac{\lambda_1}{n_1}+\frac{\lambda_2}{n_2}}}}{\frac{\sqrt{\frac{\tilde{\lambda}_1}{n_1}+\frac{\tilde{\lambda}_2}{n_2}}}{\sqrt{\frac{q_1(\lambda_1,\lambda_2)}{n_1}+\frac{q_2(\lambda_1,\lambda_2)}{n_2}}}} \\
&= \frac{\bar{Y}_1-\bar{Y}_2+\Delta_0-\delta_0^*}{\sqrt{\frac{\tilde{\lambda}_1}{n_1}+\frac{\tilde{\lambda}_2}{n_2}}} \frac{\sqrt{\frac{q_1(\lambda_1,\lambda_2)}{n_1}+\frac{q_2(\lambda_1,\lambda_2)}{n_2}}}{\sqrt{\frac{\lambda_1}{n_1}+\frac{\lambda_2}{n_2}}} \\
&= Z_{R^*}\frac{\sqrt{\frac{q_1(\lambda_1,\lambda_2)}{n_1}+\frac{q_2(\lambda_1,\lambda_2)}{n_2}}}{\sqrt{\frac{\lambda_1}{n_1}+\frac{\lambda_2}{n_2}}} - \frac{\delta_0^*}{\sqrt{\frac{\lambda_1}{n_1}+\frac{\lambda_2}{n_2}}} \frac{\sqrt{\frac{\tilde{\lambda}_1}{n_1}+\frac{\tilde{\lambda}_2}{n_2}}}{\sqrt{\frac{q_1(\lambda_1,\lambda_2)}{n_1}+\frac{q_2(\lambda_1,\lambda_2)}{n_2}}} \\
&= Z_{R^*}\frac{\sqrt{\frac{q_1(\lambda_1,\lambda_2)}{n_1}+\frac{q_2(\lambda_1,\lambda_2)}{n_2}}}{\sqrt{\frac{\lambda_1}{n_1}+\frac{\lambda_2}{n_2}}} - \frac{\delta_0^*}{\sqrt{\frac{\lambda_1}{n_1}+\frac{\lambda_2}{n_2}}}(1+o_p(1)).
\end{aligned}
$$

Then, we can obtain

$$
Z_{R^*}\frac{\sqrt{\frac{q_1(\lambda_1,\lambda_2)}{n_1}+\frac{q_2(\lambda_1,\lambda_2)}{n_2}}}{\sqrt{\frac{\lambda_1}{n_1}+\frac{\lambda_2}{n_2}}} - \frac{\delta_0^*}{\sqrt{\frac{\lambda_1}{n_1}+\frac{\lambda_2}{n_2}}} \xrightarrow{d} N(0,1).
$$

Let

$$
Z_{U^*} = \frac{\bar{Y}_1-\bar{Y}_2+\Delta_0}{\sqrt{\frac{\hat{\lambda}_1}{n_1}+\frac{\hat{\lambda}_2}{n_2}}}.
$$

Similarly, the asymptotic distribution of Z_{U^*} can be derived as follows:

$$
Z_{U^*} - \frac{\delta_0^*}{\sqrt{\frac{\lambda_1}{n_1}+\frac{\lambda_2}{n_2}}} \xrightarrow{d} N(0,1).
$$

where

$$
\sigma^* = \frac{\sqrt{\frac{q_1(\lambda_1,\lambda_2)}{n_1}+\frac{q_2(\lambda_1,\lambda_2)}{n_2}}}{\sqrt{\frac{\lambda_1}{n_1}+\frac{\lambda_2}{n_2}}},
$$

the other,

$$\frac{q_1(\lambda_1,\lambda_2)}{n_1}+\frac{q_2(\lambda_1,\lambda_2)}{n_2}$$
$$=\frac{1}{n_2\rho}\left\{q_1(\lambda_1,\lambda_2)+\rho q_2(\lambda_1,\lambda_2)\right\}$$
$$=\frac{1}{2n_2\rho}\left\{\frac{\rho\lambda_1+\lambda_2}{1+\rho}-\Delta_0+\sqrt{\left(\frac{\rho\lambda_1+\lambda_2}{1+\rho}+\Delta_0\right)^2-4\frac{\lambda_2\Delta_0}{1+\rho}}\right\}$$
$$+\frac{\rho}{2n_2\rho}\left\{\frac{\rho\lambda_1+\lambda_2}{1+\rho}+\Delta_0+\sqrt{\left(\frac{\rho\lambda_1+\lambda_2}{1+\rho}+\Delta_0\right)^2-4\frac{\lambda_2\Delta_0}{1+\rho}}\right\}$$
$$=\frac{1}{2n_2\rho}\left\{\rho\lambda_1+\lambda_2-(1-\rho)\Delta_0+\sqrt{(\rho\lambda_1+\lambda_2+(1+\rho)\Delta_0)^2-4\lambda_2\Delta_0(1+\rho)}\right\},$$

then,

$$\frac{\frac{q_1(\lambda_1,\lambda_2)}{n_1}+\frac{q_2(\lambda_1,\lambda_2)}{n_2}}{\frac{\lambda_1}{n_1}+\frac{\lambda_2}{n_2}}$$
$$=\frac{\frac{1}{2n_2\rho}\left\{\rho\lambda_1+\lambda_2-(1-\rho)\Delta_0+\sqrt{(\rho\lambda_1+\lambda_2+(1+\rho)\Delta_0)^2-4\lambda_2\Delta_0(1+\rho)}\right\}}{\frac{\lambda_1+\rho\lambda_2}{n_2\rho}}$$
$$=\frac{\rho\lambda_1+\lambda_2-(1-\rho)\Delta_0+\sqrt{(\rho\lambda_1+\lambda_2+(1+\rho)\Delta_0)^2-4\lambda_2\Delta_0(1+\rho)}}{2(\lambda_1+\rho\lambda_2)}.$$

Given $\lambda_1=\lambda_2-\Delta_0+\delta_0^*,\delta_0^*>0$, the σ^* can be rewritten as

$$\sigma^*=\sqrt{\frac{\rho\lambda_1+\lambda_2-(1-\rho)\Delta_0+\sqrt{(\rho\lambda_1+\lambda_2+(1+\rho)\Delta_0)^2-4\lambda_2\Delta_0(1+\rho)}}{2(\lambda_1+\rho\lambda_2)}}$$
$$=\sqrt{\frac{(1+\rho)\lambda_2-\Delta_0+\rho\delta_0^*+\sqrt{((1+\rho)\lambda_2+\Delta_0+\rho\delta_0^*)^2-4\lambda_2\Delta_0(1+\rho)}}{2((1+\rho)\lambda_2-\Delta_0+\delta_0^*)}}.$$

So, the power function of Z_{R^*} be can fund as follows

$$\begin{aligned}\bar{\beta}_{Z_{R^*}}(\delta_0^*,\lambda_2,n_2,\rho,\Delta_0)&=P(Z_{R^*}\geq z_\alpha|\lambda_1=\lambda_2-\Delta_0+\delta_0^*)\\&=P(Z_{R^*}\sigma^*-\mu^*\geq z_\alpha\sigma^*-\mu^*|\lambda_1=\lambda_2-\Delta_0+\delta_0^*)\\&=1-\Phi(z_\alpha\sigma^*-\mu^*).\end{aligned}$$

Similarly, the power function of Z_{U^*} be can fund as follows:

$$\bar{\beta}_{Z_{U^*}}(\delta_0^*, \lambda_2, n_2, \rho, \Delta_0) = 1 - \Phi(z_\alpha - \mu^*),$$

where

$$\mu^* = \frac{\delta_0^*}{\sqrt{\frac{\lambda_1}{n_1} + \frac{\lambda_2}{n_2}}} = \frac{\delta_0^*}{\sqrt{\frac{\lambda_2(1+\rho) - \Delta_0 + \delta_0^*}{n_2 \rho}}}.$$

Given $\rho, \delta_0^*, \Delta_0$ and λ_2, the required sample sizes of the second group satisfied the power is greater than $1 - \beta_0$ is

$$n_{2,Z_{R^*}} \geq \left(\frac{z_\alpha \sigma^* + z_{\beta_0}}{\delta_0^*} \right)^2 \frac{\lambda_2(1+\rho) - \Delta_0 + \delta_0^*}{\rho}$$

,and

$$n_{2,Z_{U^*}} \geq \left(\frac{z_\alpha + z_{\beta_0}}{\delta_0^*} \right)^2 \frac{\lambda_2(1+\rho) - \Delta_0 + \delta_0^*}{\rho}.$$

A.9

Theorem 8. Z_{R^*} satisfy the convexity condition.

Firstly, we execute the partial derivative of Z_{R^*} wrt $\hat{\lambda}_1$:

$$\begin{aligned}
\frac{\partial Z_{R^*}}{\partial \hat{\lambda}_1} &= \frac{1}{\frac{\tilde{\lambda}_1}{n_1} + \frac{\tilde{\lambda}_2}{n_2}} \left\{ \sqrt{\frac{\tilde{\lambda}_1}{n_1} + \frac{\tilde{\lambda}_2}{n_2}} - (\hat{\lambda}_1 - \hat{\lambda}_2 + \Delta_0) \frac{1}{2\sqrt{\frac{\tilde{\lambda}_1}{n_1} + \frac{\tilde{\lambda}_2}{n_2}}} \left(\frac{\partial \frac{\tilde{\lambda}_1}{n_1}}{\partial \hat{\lambda}_1} + \frac{\partial \frac{\tilde{\lambda}_2}{n_2}}{\partial \hat{\lambda}_1} \right) \right\} \\
&= \frac{1}{2(\frac{\tilde{\lambda}_1}{n_1} + \frac{\tilde{\lambda}_2}{n_2})^{\frac{3}{2}}} \left\{ 2\left(\frac{\tilde{\lambda}_1}{n_1} + \frac{\tilde{\lambda}_2}{n_2}\right) - (\hat{\lambda}_1 - \hat{\lambda}_2 + \Delta) \frac{1}{n_2} \frac{\tilde{\lambda}_2}{\tilde{\lambda}_1 + \tilde{\lambda}_2 - \tilde{\lambda}_0} \right\} \\
&= \frac{1}{2(\frac{\tilde{\lambda}_1 + \rho\tilde{\lambda}_2}{n_2 \rho})^{\frac{3}{2}}} \left\{ \frac{2}{n_2 \rho}(\tilde{\lambda}_1 + \rho\tilde{\lambda}_2) - \frac{1}{n_2}(\hat{\lambda}_1 - \hat{\lambda}_2 + \Delta) \frac{\tilde{\lambda}_2}{\tilde{\lambda}_1 + \tilde{\lambda}_2 - \tilde{\lambda}_0} \right\} \\
&= \frac{1}{2\sqrt{\frac{(\tilde{\lambda}_1 + \rho\tilde{\lambda}_2)^3}{n_2 \rho}}} \left\{ 2\left((1+\rho)\tilde{\lambda}_2 - \Delta_0\right) - \rho(\hat{\lambda}_1 - \hat{\lambda}_2 + \Delta) \frac{\tilde{\lambda}_2}{\tilde{\lambda}_1 + \tilde{\lambda}_2 - \tilde{\lambda}_0} \right\}.
\end{aligned}$$

since

$$\frac{\partial \frac{\tilde{\lambda}_1}{n_1}}{\partial \hat{\lambda}_1} = \frac{1}{2n_2}\left(\frac{1}{1+\rho} + \frac{1}{1+\rho}\frac{\tilde{\lambda}_0 + \Delta_0}{\sqrt{(\tilde{\lambda}_0+\Delta_0)^2 - 4\frac{1}{1+\rho}\hat{\lambda}_2\Delta_0}}\right),$$

and

$$\frac{\partial \frac{\tilde{\lambda}_2}{n_2}}{\partial \hat{\lambda}_1} = \frac{1}{2n_2}\left(\frac{\rho}{1+\rho} + \frac{\rho}{1+\rho}\frac{\tilde{\lambda}_0 + \Delta_0}{\sqrt{(\tilde{\lambda}_0+\Delta_0)^2 - 4\frac{1}{1+\rho}\hat{\lambda}_2\Delta_0}}\right),$$

and

$$\begin{aligned}
&\frac{\partial \frac{\tilde{\lambda}_1}{n_1}}{\partial \hat{\lambda}_1} + \frac{\partial \frac{\tilde{\lambda}_2}{n_2}}{\partial \hat{\lambda}_1} \\
&= \frac{1}{2n_2}\left(1 + \frac{\tilde{\lambda}_0 + \Delta_0}{\sqrt{(\tilde{\lambda}_0+\Delta_0)^2 - 4\frac{1}{1+\rho}\hat{\lambda}_2\Delta_0}}\right) \\
&= \frac{1}{n_2}\frac{\tilde{\lambda}_2}{\tilde{\lambda}_1+\tilde{\lambda}_2-\tilde{\lambda}_0}.
\end{aligned}$$

On the other hand, consider the partial derivative of Z_{R^*} wrt $\hat{\lambda}_2$,

$$\begin{aligned}
\frac{\partial Z_{R^*}}{\partial \hat{\lambda}_2} &= \frac{1}{\frac{\tilde{\lambda}_1}{n_1}+\frac{\tilde{\lambda}_2}{n_2}}\left\{-\sqrt{\frac{\tilde{\lambda}_1}{n_1}+\frac{\tilde{\lambda}_2}{n_2}} - (\hat{\lambda}_1-\hat{\lambda}_2+\Delta_0)\frac{1}{2\sqrt{\frac{\tilde{\lambda}_1}{n_1}+\frac{\tilde{\lambda}_2}{n_2}}}\left(\frac{\partial \frac{\tilde{\lambda}_1}{n_1}}{\partial \hat{\lambda}_2} + \frac{\partial \frac{\tilde{\lambda}_2}{n_2}}{\partial \hat{\lambda}_2}\right)\right\} \\
&= \frac{-1}{2(\frac{\tilde{\lambda}_1}{n_1}+\frac{\tilde{\lambda}_2}{n_2})^{\frac{3}{2}}}\left\{2\left(\frac{\tilde{\lambda}_1}{n_1}+\frac{\tilde{\lambda}_2}{n_2}\right) + (\hat{\lambda}_1-\hat{\lambda}_2+\Delta_0)\frac{1}{n_2\rho}\frac{\tilde{\lambda}_1}{\tilde{\lambda}_1+\tilde{\lambda}_2-\tilde{\lambda}_0}\right\} \\
&= \frac{-1}{2(\frac{\tilde{\lambda}_1+\rho\tilde{\lambda}_2}{n_2\rho})^{\frac{3}{2}}}\left\{\frac{2}{n_2\rho}(\tilde{\lambda}_1+\rho\tilde{\lambda}_2) + \frac{1}{n_2\rho}(\hat{\lambda}_1-\hat{\lambda}_2+\Delta_0)\frac{\tilde{\lambda}_1}{\tilde{\lambda}_1+\tilde{\lambda}_2-\tilde{\lambda}_0}\right\} \\
&= \frac{-1}{2\sqrt{\frac{(\tilde{\lambda}_1+\rho\tilde{\lambda}_2)^3}{n_2\rho}}}\left\{2((1+\rho)\tilde{\lambda}_2-\Delta_0) + (\hat{\lambda}_1-\hat{\lambda}_2+\Delta_0)\frac{\tilde{\lambda}_1}{\tilde{\lambda}_1+\tilde{\lambda}_2-\tilde{\lambda}_0}\right\}.
\end{aligned} \tag{1}$$

Where

$$\frac{\partial \frac{\tilde{\lambda}_1}{n_1}}{\partial \hat{\lambda}_2} = \frac{1}{2n_2\rho}\left(\frac{1}{1+\rho} + \frac{1}{1+\rho}\frac{\tilde{\lambda}_0 - \Delta_0}{\sqrt{(\tilde{\lambda}_0+\Delta_0)^2 - 4\frac{1}{1+\rho}\hat{\lambda}_2\Delta_0}}\right),$$

and

$$\frac{\partial \frac{\tilde{\lambda}_2}{n_2}}{\partial \hat{\lambda}_2} = \frac{1}{2n_2}\left(\frac{1}{1+\rho} + \frac{1}{1+\rho}\frac{\tilde{\lambda}_0 - \Delta_0}{\sqrt{(\tilde{\lambda}_0 + \Delta_0)^2 - 4\frac{1}{1+\rho}\hat{\lambda}_2\Delta_0}}\right),$$

and

$$\begin{aligned}
&\frac{\partial \frac{\tilde{\lambda}_1}{n_1}}{\partial \hat{\lambda}_2} + \frac{\partial \frac{\tilde{\lambda}_2}{n_2}}{\partial \hat{\lambda}_2} \\
&= \frac{1}{2n_2\rho}(1 + \frac{\tilde{\lambda}_0 - \Delta_0}{\sqrt{(\tilde{\lambda}_0 + \Delta_0)^2 - 4\frac{1}{1+\rho}\hat{\lambda}_2\Delta_0}}) \\
&= \frac{1}{n_2\rho}\frac{\tilde{\lambda}_1}{\tilde{\lambda}_1 + \tilde{\lambda}_2 - \tilde{\lambda}_0}.
\end{aligned}$$

In (1), as $\hat{\lambda}_1 - \hat{\lambda}_2 + \Delta_0 \geq 0$, then $\frac{\partial Z_{R^*}}{\partial \hat{\lambda}_2} < 0$ must be obtained. Similarly, we only check the sign of $\frac{\partial Z_{R^*}}{\partial \hat{\lambda}_2}$ if $\hat{\lambda}_1 - \hat{\lambda}_2 + \Delta_0 < 0$.

The following can be derived,

$$\begin{aligned}
&\frac{\tilde{\lambda}_1}{\tilde{\lambda}_1 + \tilde{\lambda}_2 - \tilde{\lambda}_0} \\
=&\frac{1}{2}\frac{\left(\tilde{\lambda}_0 - \Delta_0 + \sqrt{(\tilde{\lambda}_0 + \Delta_0)^2 - 4\frac{1}{1+\rho}\hat{\lambda}_2\Delta_0}\right)}{\sqrt{(\tilde{\lambda}_0 + \Delta_0)^2 - 4\frac{1}{1+\rho}\hat{\lambda}_2\Delta_0}} \\
=&\frac{1}{2}\left(1 + \frac{\tilde{\lambda}_0 - \Delta_0}{\sqrt{(\tilde{\lambda}_0 + \Delta_0)^2 - 4\frac{1}{1+\rho}\hat{\lambda}_2\Delta_0}}\right).
\end{aligned}$$

And, we have follows as,

$$
\begin{aligned}
&\sqrt{(\tilde{\lambda}_0+\Delta_0)^2-4\frac{1}{1+\rho}\hat{\lambda}_2\Delta_0} \\
=&\sqrt{(\frac{\rho}{1+\rho}\hat{\lambda}_1+\frac{1}{1+\rho}\hat{\lambda}_2+\Delta_0)^2-4\frac{1}{1+\rho}\hat{\lambda}_2\Delta_0} \\
=&\sqrt{(\frac{\rho}{1+\rho}\hat{\lambda}_1+\frac{1}{1+\rho}\hat{\lambda}_2-\Delta_0)^2+4\frac{\rho}{1+\rho}\hat{\lambda}_1\Delta_0} \\
\geq&\sqrt{(\frac{\rho}{1+\rho}\hat{\lambda}_1+\frac{1}{1+\rho}\hat{\lambda}_2-\Delta_0)^2} \\
=&|\tilde{\lambda}_0-\Delta_0|,
\end{aligned}
$$

then,

$$\frac{|\tilde{\lambda}_0-\Delta_0|}{\sqrt{(\tilde{\lambda}_0+\Delta_0)^2-4\frac{1}{1+\rho}\hat{\lambda}_2\Delta_0}}\leq 1,$$

and

$$-1\leq\frac{\tilde{\lambda}_0-\Delta_0}{\sqrt{(\tilde{\lambda}_0+\Delta_0)^2-4\frac{1}{1+\rho}\hat{\lambda}_2\Delta_0}}\leq 1.$$

Hence, we have

$$
\begin{aligned}
&\Rightarrow\ 0\leq 1+\frac{\tilde{\lambda}_0-\Delta_0}{\sqrt{(\tilde{\lambda}_0+\Delta_0)^2-4\frac{1}{1+\rho}\hat{\lambda}_2\Delta_0}}\leq 2 \\
&\Rightarrow\ 0\leq\frac{1}{2}\left\{1+\frac{\tilde{\lambda}_0-\Delta_0}{\sqrt{(\tilde{\lambda}_0+\Delta_0)^2-4\frac{1}{1+\rho}\hat{\lambda}_2\Delta_0}}\right\}\leq 1 \\
&\Rightarrow\ 0\leq\frac{\tilde{\lambda}_1}{\tilde{\lambda}_1+\tilde{\lambda}_2-\tilde{\lambda}_0}\leq 1.
\end{aligned}
$$

If $\hat{\lambda}_1-\hat{\lambda}_2+\Delta_0<0$, then

$$(\hat{\lambda}_1-\hat{\lambda}_2+\Delta_0)\leq(\hat{\lambda}_1-\hat{\lambda}_2+\Delta_0)\frac{\tilde{\lambda}_1}{\tilde{\lambda}_1+\tilde{\lambda}_2-\tilde{\lambda}_0}\leq 0,$$

and

$$
\begin{aligned}
&2((1+\rho)\tilde{\lambda}_2 - \Delta_0) + (\hat{\lambda}_1 - \hat{\lambda}_2 + \Delta_0) \\
\leq &2((1+\rho)\tilde{\lambda}_2 - \Delta_0) + (\hat{\lambda}_1 - \hat{\lambda}_2 + \Delta_0)\frac{\tilde{\lambda}_1}{\tilde{\lambda}_1 + \tilde{\lambda}_2 - \tilde{\lambda}_0} \\
\leq &2((1+\rho)\tilde{\lambda}_2 - \Delta_0),
\end{aligned}
$$

and,

$$
\begin{aligned}
&2\left((1+\rho)\tilde{\lambda}_2 - \Delta_0\right) + \hat{\lambda}_1 - \hat{\lambda}_2 + \Delta_0 \\
=&2\left(\frac{1+\rho}{2}(\tilde{\lambda}_0 + \Delta_0 + \sqrt{(\tilde{\lambda}_0 + \Delta_0)^2 - 4\frac{1}{1+\rho}\hat{\lambda}_2\Delta_0}) - \Delta_0\right) + \hat{\lambda}_1 - \hat{\lambda}_2 + \Delta_0 \\
=&(1+\rho)\left(\frac{\rho}{1+\rho}\hat{\lambda}_1 + \frac{1}{1+\rho}\hat{\lambda}_2 + \Delta_0 + \sqrt{(\tilde{\lambda}_0 + \Delta_0)^2 - 4\frac{1}{1+\rho}\hat{\lambda}_2\Delta_0}\right) - 2\Delta_0 + \hat{\lambda}_1 - \hat{\lambda}_2 + \Delta_0 \\
=&(1+\rho)\hat{\lambda}_1 + (1+\rho)\Delta_0 - \Delta_0 + \sqrt{(\tilde{\lambda}_0 + \Delta_0)^2 - 4\frac{1}{1+\rho}\hat{\lambda}_2\Delta_0} \\
=&(1+\rho)\hat{\lambda}_1 + \rho\Delta_0 + \sqrt{(\tilde{\lambda}_0 + \Delta_0)^2 - 4\frac{1}{1+\rho}\hat{\lambda}_2\Delta_0} \\
>&0.
\end{aligned}
$$

Therefore, as $\hat{\lambda}_1 - \hat{\lambda}_2 + \Delta_0 < 0$ we still can provide that

$$
2\left((1+\rho)\tilde{\lambda}_2 - \Delta_0\right) + (\hat{\lambda}_1 - \hat{\lambda}_2 + \Delta_0)\frac{\tilde{\lambda}_1}{\tilde{\lambda}_1 + \tilde{\lambda}_2 - \tilde{\lambda}_0} > 0.
$$