 Q1. Consider the **fev.csv** data set in our Example 1 (and Example1.SAS) for studying the effect of a test drug on improving the resting FEV1 after six weeks of treatment. With the same primary endpoint as in Example 1, the investigator is interested in conducting a **non-inferiority (NI) test** to check **if the effect of the placebo is not inferior to the test drug** by a tolerance (NI margin) of **M=0.5** liter. Write down the null and alternative hypotheses for this NI test, and use the current data set (by omitting the missing values) to conduct the test with **α=0.025.** Clearly define the notation that you use, and explain your conclusion.

H0: The mean effect of the placebo is inferior to the mean effect of the test drug more than M = 0.5 liters.

Ha: The mean effect of the placebo is NOT inferior to the mean effect of the test drug by more than M = 0.5.

FEV1 indicates lung capacity and the higher FEV1 means better lung capacity, healthier.

We expect greater FEV1 (positive value)

Let d = the difference of REV1 between the T0 (reading fev0) and T6 (reading fev6) = fev6 – fev0.

* d= fev6 – fev0

We are expecting fev6 – fev0 to be positive value, indicating improvement (greater lung capacity).

* μ\_P: Average change in FEV1 for the placebo group (week 6 - week 0)
* μ\_A: Average change in FEV1 for the new drug group (week 6 - week 0)
* M: Non-inferiority margin (specified as 0.5 liters)

If the mean of d.placebo (u.p) is greater than (or equal to) the mean of d.test (u.A), the improvement is greater with placebo. This means placebo effect is greater or equal to new drug effect. This means placebo is not inferior to new drug. (u.p ≥ u.A : P is not inferior)

Thus, to test whether the effect of placebo is not inferior to the test drug by M=0.5,

**H0**: μ.P ≤ μ.A - 0.5, thus u.A – u.P ≥ 0.5 (**the effect of the placebo is inferior to the test drug)**

**Ha:** μ.P > μ.A - 0.5, thus u.A – u.P < 0.5 (**the effect of the placebo is NOT inferior to the test drug)**

**Data setup with omitting the missing values:**

**A screenshot of a computer program

Description automatically generated**

A screenshot of a computer program

Description automatically generated

A screenshot of a table

Description automatically generated A screenshot of a computer screen

Description automatically generated

* **The mean difference between the new drug and placebo (u.A – u.P) is 0.2187 L.**
* The 95% confidence limits by both pooled and Satterthwaite methods for the mean difference are from -0.2515 to 0.6889 liters.
* Since this is non-inferiority test, we consider the lower bound -0.2515 (α=0.025) against the NI margin M=0.5. Since -0.2515 is greater than -0.5, the lowest confidence limit is not in the rejection region. Thus, there is not enough evidence to reject H0.
* **Thus, we do not reject H0 and conclude that the effect of the placebo is inferior to the test drug by the NI margin of 0.5.**

Q2. Use the parameter estimates that you obtained from Q1 to calculate the needed sample size for the another NI trial under the same setting. The NI margin for the next trial is again M=0.5, the target statistical power is 90%, and the type I error rate is α=0.025. You may assume an equal variance between the two groups and consider a balanced design having an equal sample size for both treatment groups.

**From the results above:**

A screenshot of a data

Description automatically generated

**SAS codes and result:**

A screenshot of a computer program

Description automatically generated A screenshot of a computer

Description automatically generated

**Results:**

Conclusion: The required total sample size is 144, with the power of >0.9.

Q3. An equivalence trial with two parallel groups is conducted to monitor the incidence of gastro-intestinal (GI) adverse drug reactions of a new antibiotic used in lower respiratory tract infections (LRTI). In this study, 66 LRTI patients are randomized to receive the new treatment, and 52 LRTI patients are randomized into the control group to receive erythromycin. There are 36 patients in the test group, and 28 patients in the control group who reported one or more GI complaints during 7 days of treatment. Let pt and pc be the proportions of LRTI patients who experience such adverse events in the test and control groups, respectively. If the absolute difference of pt and pc is within 5%, we say that the two drugs have similar GI adverse drug reactions. Use this equivalence margin (5%) and the data presented in this question to conduct an equivalence test with a type I error rate of **α=0.05**. Clearly state the null and alternative hypotheses, the method that you choose to use, and your conclusion.

Reference. This is Example 16.1 of [R5] with modifications.

*\* To test GI-reaction by new LRTI-drug*

**66** LRTI patients - case (new drug)

**52** LRTI patients - control group (Erythromycin)

New\_Treatment GI\_Complaints **36**

New\_Treatment No\_GI\_Complaints **30**

Erythromycin GI\_Complaints **28**

Erythromycin No\_GI\_Complaints **24** \*/

**data** AE\_NewDrug;

do GI = **0** to **1**; */\*GI complain, no:yes \*/*

do group = **1** to **2**; */\*Group, case:control\*/*

input count @@;

output;

end;

end;

*\*/GI=0, 66-36 : 52-28, case:control \*/*

*\*/GI=1, 36 : 28, case:control \*/;*

cards;

**30** **24**

**36** **28**

;

**run**;

**proc** **format**;

title 'GI adverse effect by a new drug';

value GIfmt **0** = 'No GI complaint'

**1** = 'GI complaints (>=1)';

value groupfmt **1** = 'New drug'

**2** = 'Control(Ery)';

**run**;

**proc** **print** data=AE\_NewDrug;

**format** GI GIfmt. group groupfmt.;

**run**;

A screenshot of a computer

Description automatically generated

*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\**

*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\* Risk difference \*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\**

*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\**

• pt=36/66

• pc=28/52

Ho: pt - pc =< **0**.**05** (similar AE between New drug and control)

Ha: pt - pc > **0**.**05** (More adverse effect is likely happen due to the new drug);

ods trace on;

**proc** **freq** data=AE\_NewDrug;

title 'Contingency Table: Farrington-Manning Score Test and Confidence Intervals';

**format** GI GIfmt. group groupfmt.;

weight count;

tables group\*GI / nopercent nocol

riskdiff(column=2 equal method=FM cl=FM cl=NEWCOMBE) chisq;

*\* column=1, GI=0, No GI compaining*

*\* column=2, GI=1, GI compaining;*

**run**;

ods trace off;

title;

*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\**

Screens screenshot of a computer screen

Description automatically generated

A screenshot of a computer screen

Description automatically generated

Based on the output, the risk difference is -0.0070 with a 95% confidence interval from the Farrington-Manning method ranging from -0.1741 to 0.1881. Since 1) the confidence interval includes zero, and 2) the absolute value of the risk difference is less than the equivalence margin of 5%, the results suggest that **there is no significant difference between the two treatments regarding the incidence of GI adverse reactions**.

* pt=36/66
* pc=28/52
* **H0:** |pt - pc| =< 0.05 (similar AE between New drug and control)
* **Ha:** |pt - pc| > 0.05 (More adverse effect is likely happen due to the new drug);

**Risk difference test**:

Both one-sided and the two-sided p-value is greater than 0.05 (=α). Thus, the test failed to reject H0, indicating the absolute difference of pt and pc is within 5%.

In conclusion, with 95% confidence, there is no evidence suggesting the risk difference is greater than 5% between the new drug and erythromycin (control) in terms of the incidence of GI adverse reactions. Thus, the two drugs have similar GI adverse drug reactions based on the data from this study.

Q4. In this question, we set the vaccine efficacy (VE) as 𝑉𝐸 = 1 − p𝑣/𝑝𝑐, where pj is the probability of experiencing the condition of interest for subjects in Group j=v, c.

Group v is the vaccinated group, and Group c is the control group. We also use 𝑝̂𝑗 = x𝑗 / 𝑛𝑗 as an estimator of pj, where xj is the number of infected subjects, and nj is the sample size for Group j. In a COVID-19 vaccine study (Polack et al., 2020, The New England Journal of Medicine, 383, p. 2603-2615), 21,720 subjects are vaccinated (Group v), and 21,728 subjects are with placebo (Group c). Among these subjects, the investigators observed 8 cases of COVID in Group v, and 162 infected cases in Group c.

Use this data set, and the methods that we introduced for studying **relative risks to find 95% two-sided confidence intervals (CIs) for VE**. Use both the Wald and score approaches.

From these confidence intervals, **would you reject H0: VE ≤ 0.3 (vs. Ha: VE > 0.3) to claim that the vaccine is effective?** Clearly specify the type I error rate α for this test.

Note 1. You may directly use SAS to obtain your answers, but the SAS results may be for 𝑅𝑅 = 𝑝𝑣/𝑝𝑐. In this case, you need a simple calculation to obtain the CI for VE=1-RR from the CI for RR.

Note 2. We consider here a similar setting as that in O'Neill (1988) Statistics in Medicine, 7, p.1279-1288. This is different from the slightly more advanced approach that we introduced in class for the Pfizer's SARS-CoV2 vaccine study, although we use the same COVID-19 data for this question.

A screenshot of a computer code

Description automatically generatedA screenshot of a screen

Description automatically generated

A screenshot of a computer screen

Description automatically generated

* A screenshot of a computer screen

  Description automatically generated**The RR (pv/pc) of getting COVID in vaccinated group over the placebo group is 0.0494 with p-value <0.0001**. This means that the infection rate in the vaccinated group is about 4.94% of the placebo group, indicating significant reduction in the COVID risk due to the vaccine.
* **The vaccine effect (VE=1-RR) is estimated to be 95.06% (1−0.0494= 0.9506),** which is significantly high.
* Since 97.5% confidence limits of lower or upper RR is 0.0246 or 0.0990 by score method, **the 95% CI of VE by score method is (0.901, 0.9754).**
* Since 97.5% confidence limits of lower or upper RR is 0.0243 or 0.1004 by Wald method, **the 95% CI of VE by Wald method is (0.8996, 0.9757).**

**H0: VE ≤ 0.3 (RR≥0.7)**

**Ha: VE > 0.3 (RR<0.7)**

Given that both CI by the Score and Wald methods, **we reject the null hypothesis H0: VE ≤ 0.3** and conclude that there is enough evidence suggesting the VE > 0.3 with the type I error rate α<0.05 (we used 95% 2-sided CI). The estimated **VE is 95.06%**, with the two-sided 95% CI for VE ranging from **90.10% to 97.54% by the Score method**, or from **89.96% to 97.57% by the Wald method**, with a p-value < 0.0001.

Q5. Consider the same setting as in Q4 with same hypotheses H0 and Ha. Suppose that the investigator plans to use the Farrington-Manning score test for relative risk. **Calculate the sample size** needed for a target **power of 90% at VE=0.6** with **one-sided α=0.025**. As in the Pfizer’s SARS-CoV2 vaccine study, assume that the disease attack rate (in 6 month) is pc=0.65%.

Note 1. You may considering rewriting the problem (and the hypotheses) in terms of 𝑅𝑅 = 𝑝𝑣/𝑝𝑐 = 1 − 𝑉𝐸, and modifying the last portion of Example3.SAS to calculate the sample size.

Note 2. In the Pfizer’s study, it is expected that 20% of the subjects are not evaluable. If we use this expectation to increase the sample size obtained from Q5, the final sample size will turn out to be similar to that in the Pfizer’s study (stated in Q4).

**H0: VE ≤ 0.3 (RR≥0.7)**

**Ha: VE > 0.3 (RR<0.7)**

A computer code with text

Description automatically generated with medium confidence

A screenshot of a computer

Description automatically generated

The total sample size needed to achieve the target power of 90% at the specified significance level is 33758.