Project 1: Horse Kick – Hypothesis Test

Due on September 14th, Monday

The data set below was reported by von Bortkiewicz (1898, *Das Gesetz der Kleinen Zahlen*, Leipzig: Teubner) as a classic Poisson example, for the chance of a Prussian cavalryman being killed by the kick of a horse. Fourteen army corps (indicated by columns) were observed during 20 years (1875-1894, indicated by rows), giving a total of 280 death counts of one corps for a one-year period.

Year\Corps	1	2	3	4	5	6	7	8	9	10	11	12	13	14
1875	0	0	0	0	0	0	0	1	1	0	0	0	1	0
1876	2	0	0	0	1	0	0	0	0	0	0	0	1	1
1877	2	0	0	0	0	0	1	1	0	0	1	0	2	0
1878	1	2	2	1	1	0	0	0	0	0	1	0	1	0
1879	0	0	0	1	1	2	2	0	1	0	0	2	1	0
1880	0	3	2	1	1	1	0	0	0	2	1	4	3	0
1881	1	0	0	2	1	0	0	1	0	1	0	0	0	0
1882	1	2	0	0	0	0	1	0	1	1	2	1	4	1
1883	0	0	1	2	0	1	2	1	0	1	0	3	0	0
1884	3	0	1	0	0	0	0	1	0	0	2	0	1	1
1885	0	0	0	0	0	0	1	0	0	2	0	1	0	1
1886	2	1	0	0	1	1	1	0	0	1	0	1	3	0
1887	1	1	2	1	0	0	3	2	1	1	0	1	2	0
1888	0	1	1	0	0	1	1	0	0	0	0	1	1	0
1889	0	0	1	1	0	1	1	0	0	1	2	2	0	2
1890	1	2	0	2	0	1	1	2	0	2	1	1	2	2
1891	0	0	0	1	1	1	0	1	1	0	3	3	1	0
1892	1	3	2	0	1	1	3	0	1	1	0	1	1	0
1893	0	1	0	0	0	1	0	2	0	0	1	3	0	0
1894	1	0	0	0	0	0	0	0	1	0	1	1	0	0

1. For each corps, test if the death counts follow a Poisson distribution.

2. Assume the death counts follow a Poisson distribution for each corps, test if the parameters are the same across different corps.

3. Assume the death counts in the 14 corps follow the same distribution, test if the death counts of all the 14 corps follow a Poisson distribution.

4. Given the null hypothesis is Poisson distribution with some parameter, find a reasonable alternative hypothesis for your test and check the power of your test.

5. List at least two other tests for the first and third questions, discuss the differences of the tests, and recommend a test for the horse kick data.

Notes:

[1] Students are required to work in groups on course projects and submit their reports in pdf or doc format. The group size can be 1, 2 or 3.

[2] Each group is required to submit one hard copy of the report in the class of the due day. A list of names of the group members should be put on the cover page of the report.

[3] Each report should include a nontechnical part and a technical part. The nontechnical part should be fitted into one page. The technical part should include the statement of the problem, model assumptions, formulas used, results, conclusion, suggestion and discussion if necessary.

[4] Any statistical software and other tools can be used. If R program is used, the attachment of R code is recommended.