Exercise 2

TKO_7093 Statistical Data Analysis
BIMA3015 Statistics in Biomedical Sciences

1. Let us consider the following three datasets

A: 34, 23, 52, 47, 34
B: 48, 28, 33, 46, 41, 35
C: 34, 53, 51, 35, 52, 32

Find P-values for suitable statistical test comparing A&B, A&C and B&C. Justify your selection of the test.

2. Browse through the following two articles

Article 2: W. He et al., High-salt diet inhibits tumour growth in mice via regulating myeloid-derived suppressor cell differentiation, PMID: 32265505

Article 3: M.S. Venäläinen et al., Easy-to-use tool for evaluating the elevated acute kidney injury risk against reduced cardiovascular disease risk during intensive blood pressure control, PMID: 31977572

For Article_2 determine
a) What kind of statistical tests were used in Figure 1?
b) What kind of statistical tests were used in Figure 2?
c) What kind of statistical tests were used in Figure 4 a,b?
d) What kind of statistical tests were used in Figure 4 e,f?

In Article_3, caption of table 1 it reads “There were no significant differences ($P < 0.05$) between training and test cohorts for any variable”. Which tests were used? (Or, which test do you think were used?)

3. Groups D and E have the same variable measured before (test) and after medication (retest). Select right statistical test and find P-values.
a) Compare test and retest results for group D.
b) Compare test and retest results for group E.
c) Compare test results for group D and E.
d) Compare retest results for group D and E.

Group_D
Test: 5.6, 3.1, 8.7, 4.5, 6.7, 4.5
Retest: 6.1, 5.8, 8.5, 5.3, 7.2, 5.1

Group E
Test: 4.5, 3.9, 7.1, 4.3, 6.9, 8.2, 7.6
Retest: 4.9, 4.7, 7.8, 4.8, 7.5, 7.8, 8.1

4. Consider correlation between group E test and group E retest in Exercise 3.
a) Person correlation and related P-value.
b) Spearman correlation and related P-value.
c) Rank the two groups. (For same values use half. E.g. if there are two values with the smallest value, then rank them both 1.5.)
d) Pearson correlation and related P-value of the ranking obtained in 3).

5. In this exercise we consider file simulated_data_2_5.csv. It has four groups of 100 observation: F, G, H and I. Find correlation and related P-value
a) Data file, compare columns F and G.
b) Data file simulated_data.csv, compare columns F and H.
c) Data file simulated_data.csv, compare columns F and I.