

Please provide complete and well-written solutions to the following exercises.

Due November 18, 12PM noon PST, to be uploaded as a single PDF document to blackboard (under the Assignments tab).

## Homework 6

**Exercise 1.** Using any method you wish to use, find the number of ways that the integer 30 can be written as a sum of 5 distinct positive integers among the elements of  $\{1, 2, \dots, 10\}$ . (Hint: you should probably use a computer.)

**Exercise 2.** Suppose  $n = 3, m = 4, n + m = 7$  and we denote  $Z$  as the Mann-Whitney statistic. Suppose  $X_1, \dots, X_n$  are i.i.d. (treatment outcomes of the  $n$  people in the control group) and  $Y_1, \dots, Y_m$  are i.i.d. (treatment outcomes of the  $m$  people in the treatment group). By definition of  $Z$ ,  $1 + 2 + 3 \leq Z \leq 7 + 6 + 5$ . The null hypothesis is that the treatment has no effect on people (no “good” effect and no “bad” effect). We should reject the null hypothesis when  $Z$  is close to 6 or 18. Consider the (family of) hypothesis tests that rejects the null hypothesis when  $|Z - 12| \geq c$  for some constant  $c > 0$ .

- Compute the  $p$ -value for this hypothesis test when  $Z = 17$ .
- Compute the  $p$ -value for this hypothesis test when  $Z = 16$ .

Now, suppose  $n = m = 1000$ . Recall that  $\mathbf{E}Z = \frac{m(m+n+1)}{2}$ . Consider the (family of) hypothesis tests that rejects the null hypothesis when  $\left|Z - \frac{m(m+n+1)}{2}\right| \geq c$  for some constant  $c > 0$ . Using the limiting distribution of  $Z$  as an approximation to the distribution of  $Z$  itself:

- Approximately compute the  $p$ -value for this hypothesis test when you observe that  $Z = \frac{m(m+n+1)}{2} + 2\sqrt{nm(m+n+1)/12}$ . Are you confident in rejecting the null hypothesis in this case?
- Approximately compute the  $p$ -value for this hypothesis test when you observe that  $Z = \frac{m(m+n+1)}{2} + 3\sqrt{nm(m+n+1)/12}$ . Are you confident in rejecting the null hypothesis in this case?

**Exercise 3.** Let  $Y_1, \dots, Y_n$  be i.i.d random variables uniformly distributed in  $\{-1, 1\}$ . Let

$$W_n := \sum_{i=1}^n \max(iY_i, 0).$$

Explicitly write down the distribution of  $W_6$ . (Hint: you should probably use a computer.)

**Exercise 4.** Suppose you are running an experiment to test a new blood pressure drug. The first phase of your trial only involves four people. The following table summarizes the results (for e.g. systolic blood pressure).

Person	Blood pressure before treatment	Blood pressure after treatment
1	120	124
2	140	130
3	130	132
4	110	111

Compute the Wilcoxon ranked sign statistic  $W_4$  for this data.

Now, suppose you run a larger trial, and the data is the following.

Person	Blood pressure before treatment	Blood pressure after treatment
1	120	124
2	140	130
3	130	132
4	110	111
5	122	121
6	143	132
7	131	132
8	140	161
9	100	110
10	100	107
12	100	94
12	140	145
13	130	137
14	110	118
15	100	88
16	135	130
17	136	132
18	113	111
19	129	124
20	145	130

Let  $n = 20$ . Approximately compute the  $p$ -value for this data for the (family of) hypothesis test that reject when  $\frac{|W_n - \frac{n(n+1)}{4}|}{n^3/12} > c$  for some  $c > 0$ .

Are you confident in rejecting the null hypothesis?