# R Short Course | MBBC Students 18/19
# Non-Parametric Inference

**EXERCISES**

1. The following data reports readings from 12 home radon detectors exposed to 105 picocuries per liter of radon:

   91.9 97.8 111.4 122.3 105.4 95.0 103.8 99.6 96.6 119.3 104.8 101.7

   We would like to test hypotheses about the median reading from home radon detectors:

   $$H_0: \text{median}=105 \qquad vs \qquad H_1: \text{median}\neq105$$

   To do this, apply the Wilcoxon signed-rank statistic to the differences between the observations and 105. What do you conclude?

2. The weight of 5 grains was measured from two experimental varieties designated *premier* and *super*. Each grain was measured in mg to one decimal place. It is intended to determine whether there are significant differences between the two weights varieties.

   | *premier* | 24.5 | 23.4 | 25.3 | 23.4 | 22.1 |
   |---|---|---|---|---|---|
   | *super* | 26.4 | 27.0 | 25.2 | 25.8 | 27.1 |

   Given that the number of observations is very small, use a non-parametric test to answer the question.

1

3. Nurses in an inner-city hospital were unknowingly observed on their use of latex gloves during procedures for which glove use is recommended. The nurses then attended a presentation on the importance of glove use. One month after the presentation, the same nurses were observed again. Here are the proportions of procedures for which each nurse wore gloves:

| Nurse | Before | After | Nurse | Before | After |
|-------|--------|-------|-------|--------|-------|
| 1 | 0.857 | 0.500 | 8 | 0.000 | 1.000 |
| 2 | 0.500 | 0.833 | 9 | 0.000 | 0.667 |
| 3 | 1.000 | 1.000 | 10 | 0.167 | 1.000 |
| 4 | 0.000 | 1.000 | 11 | 0.000 | 0.750 |
| 5 | 0.000 | 1.000 | 12 | 0.000 | 1.000 |
| 6 | 0.000 | 1.000 | 13 | 0.000 | 1.000 |
| 7 | 1.000 | 1.000 | 14 | 1.000 | 1.000 |

   (a) Why is a one-sided alternative proper here? Why must matched pairs methods be used?

   (b) Does the test indicate that the presentation was helpful? Use an adequate hypothesis testing verifying the assumptions.

4. A chicken pathologist wanted to determine if testosterone in three strains of mature roosters differs.

| Strain 1: | 339 | 468 | 134 | 497 | 229 | 329 |
|-----------|-----|-----|-----|-----|-----|-----|
| Strain 2: | 103 | 115 | 198 | 126 | 115 | 120 |
| Strain 3: | 107 | 199 | 102 | 105 | 89 | 110 |

   (a) Construct side-by-side boxplots of these data. Based on these plots, do means differ by strain? Do variances differ by strains? Would you perform a standard ANOVA for these data? Explain.

   (b) It appears that variances differ by strain. In addition, the sample sizes are small and data do not appear Normal. Use a Kruskal-Wallis test to determine if there are differences among strains.

2

5. The number of birds observed before finding one with yellow feathers is believed to follow a Geometric distribution with unknown probability of success $\theta$.

| Number of birds | 0 | 1 | 2 | 3 | 4 | 5 | 6 | $\geq 7$ |
|---|---|---|---|---|---|---|---|---|
| Number of observations | 12 | 27 | 49 | 66 | 67 | 32 | 11 | 5 |

 (a) What is the estimate for parameter $\theta$?

 (b) Does the following data provide sufficient evidence for this belief?

6. Hair color and eye color are categorized for a sample of 1000 people.

| **Eye color** | **Hair** Brunette | **color** Red | Blonde | Total |
|---|---|---|---|---|
| Blue | 100 | 20 | 180 | 300 |
| Brown | 450 | 30 | 120 | 600 |
| Green | 50 | 20 | 30 | 100 |
| Total | 600 | 70 | 330 | 1000 |

 (a) Was this study performed in order to test independence or homogeneity? Justify your answer.

 (b) Test the hypothesis established above.