

# 04\_homework\_nonparametric\_t\_test

Bill Perry

## Background

This assignment is based on Graham & Angilletta (2022) “Distinguishing between unreliability and dishonesty: A comparative study of aggressive communication in crayfish” published in Functional Ecology.

The study investigates whether claw size serves as an honest or dishonest signal of fighting ability in crayfish. The researchers compared claw strength between:

- **Stream-dwelling species** (tertiary burrowers: *Cambarus carinirostris*, *C. robustus*, *Faxonius obscurus*)
  - aggressive species that fight frequently and may use claws for signaling
- **Burrowing species** (primary burrowers: *Cambarus dubius*, *C. monongalensis*, *Lacunicambus thomai*)
  - docile species that construct burrows and do not fight or signal with claws

## Learning Objectives

By the end of this assignment, you will be able to:

1. State assumptions for non-parametric tests
  2. Formulate and test appropriate hypotheses
  3. Create exploratory visualizations using ggplot2
  4. Conduct and interpret Mann-Whitney U tests (Wilcoxon rank-sum tests)
  5. Create publication-quality figures
- 

## Assignment Tasks

### Task 1: Load Libraries and Data

```
# Load required libraries
library(tidyverse)
```

```
— Attaching core tidyverse packages — tidyverse 2.0.0 —
✓ dplyr      1.1.4      ✓ readr      2.1.5
✓ forcats    1.0.0      ✓ stringr     1.5.1
✓ ggplot2     3.5.2      ✓ tibble      3.3.0
✓ lubridate   1.9.4      ✓ tidyr       1.3.1
✓ purrr       1.1.0
— Conflicts — tidyverse_conflicts() —
✖ dplyr::filter() masks stats::filter()
✖ dplyr::lag()     masks stats::lag()
i Use the conflicted package (<http://conflicted.r-lib.org/>) to force all conflicts to become errors
```

```
library(ggplot2)
library(dplyr)
library(readr)
library(broom)
```

```
library(ggpubr)
library(scales)
```

Attaching package: 'scales'

The following object is masked from 'package:purrr':

discard

The following object is masked from 'package:readr':

col\_factor

```
library(viridis)
```

Loading required package: viridisLite

Attaching package: 'viridis'

The following object is masked from 'package:scales':

viridis\_pal

```
# Load the data
claw_df <- read_csv("data/graham_and_angilletta_claw_data.csv")
```

Rows: 619 Columns: 9

— Column specification —————

Delimiter: ","

chr (5): crayfish.id, species, lifestyle, sex, clawside

dbl (4): carapace.length.mm, claw.length, maxclawstr.N, PC1.claw.size

i Use `spec()` to retrieve the full column specification for this data.

i Specify the column types or set `show\_col\_types = FALSE` to quiet this message.

```
# Display data structure
```

```
head(claw_df)
```

```
# A tibble: 6 × 9
```

	crayfish.id	species	lifestyle	sex	clawside	carapace.length.mm	claw.length
	<chr>	<chr>	<chr>	<chr>	<chr>	<dbl>	<dbl>
1	CC05	c.carinir...	Stream-d...	Fema...	l	15.6	8.38
2	CC05	c.carinir...	Stream-d...	Fema...	r	15.6	9.07
3	CC06	c.carinir...	Stream-d...	Male	r	34.0	29.2
4	CC06	c.carinir...	Stream-d...	Male	l	34.0	28.3
5	CC07	c.carinir...	Stream-d...	Male	l	24	14.6
6	CC07	c.carinir...	Stream-d...	Male	r	24	14.6

```
# i 2 more variables: maxclawstr.N <dbl>, PC1.claw.size <dbl>
```

## Task 2: Data Exploration

Calculate the summary statistics

```
# Summary statistics
```

## Task 3: State Assumptions and Hypotheses

YOUR ANSWER HERE:

**Assumptions for Mann-Whitney U Test (Wilcoxon Rank-Sum Test):**

1. **Independence:** [Write your answer here]
2. **Ordinal or continuous data:** [Write your answer here]
3. **Similar distribution shape:** [Write your answer here]

**Hypotheses:**

**Null Hypothesis ( $H_0$ ):**

**Alternative Hypothesis ( $H_1$ ):**

**Biological rationale:** Stream-dwelling crayfish (tertiary burrowers) engage in frequent territorial fights using their claws as weapons and potentially as signals of competitive ability. Burrowing crayfish (primary burrowers) are docile and use their claws primarily for digging, with no territorial fighting or signaling behavior. If claw strength has been shaped by selection for fighting ability or signaling, we might expect differences between these lifestyle groups. However, both groups may show similar variation due to shared developmental constraints.

---

## Task 4: Exploratory Data Analysis

Create exploratory visualizations to understand the data distribution and patterns.

**Interpretation of exploratory plots:**

Write your interpretation focusing on:

1. Do the distributions appear normal or skewed?
  2. Are there obvious differences in central tendency between stream-dwelling and burrowing crayfish?
  3. Do you notice differences in variability between the two lifestyle groups?
  4. Are there any notable differences between males and females?
  5. Do the data appear suitable for non-parametric analysis?
- 

## Task 5: Non-parametric Statistical Tests

Conduct Mann-Whitney U tests (Wilcoxon rank-sum tests) to compare claw strength between lifestyles.

Note you can do this for

- all crayfish
- males only
- females only
- or all of the above

**Interpretation of statistical results:**

Describe what the test results mean, including:

1. Which group (stream-dwelling or burrowing) tends to have higher/lower claw strength?
2. Are the differences statistically significant?
3. What do the effect sizes tell us about practical significance?
4. How do these results relate to the paper's findings about dishonest signaling?

5. Do the results support or contradict your initial hypothesis?

---

## **Task 6: Publication-Quality Figure**

Create a publication-ready figure summarizing your results.

---

## **Submission Guidelines**

### **What to turn in -**

1. a quarto markdown file and dataframe. Note that your code should be able to run with what you turn in.
2. a self-contained word and html file showing the code and output
3. annotations in the quarto file that shows or tells what is being done in the r code chunks describing what you are trying to do - credit will be given even if it does not work as long as you detail what you are doing. As we start to move into more statistics you will be expected to interpret the results.

### **Points**

- summary stats - 1 points
- exploratory graphs - 1 point
- Hypotheses - 2 points
- Assumptions - 2 points
- Model and Test Results - 4
- interpretation - 3 points
- Assumption Tests - 4 points
- Final figure - 3 points