METRIC TESTS (MORE-THAN-TWO-SAMPLE SITUATIONS)

UNIVERSITÄT
LEIPZIG

Erik Kusch

erik.kusch@uni-leipzig.de

Behavioural Ecology Research Group
University of Leipzig

27/05/2019

Introduction

Metric tests are used to compare parameters of metric/ordinal variable values among groups/individuals.

Prominent metric tests for more-than-two-sample situations include:

- **Kruskal-Wallis Test**
- **Friedman Test**
- t Test (dealt with in Seminar 12)
- ...

Terminology

Remember:

- **Got two samples?**
  - *Independent.* $\rightarrow$ Mann-Whitney U Test
  - *Dependent.* $\rightarrow$ Wilcoxon Signed Rank Test
- **Got more than two samples?**
  - *Independent.* $\rightarrow$ Kruskal-Wallis Test
  - *Dependent.* $\rightarrow$ Friedman Test

Purpose And Assumptions

Kruskal-Wallis Test

`kruskal.test()` in base `R`

Purpose: To identify whether groups of variable values are different from one another.

H_0 *There is no difference in characteristics of the response variable values in dependence of the classes of the predictor variable.*

Assumptions:

- Predictor variable is categorical (not binary!)
- Response variable is ordinal or metric
- Variable values are **independent** (not paired)

Minimal Working Example

Let's use the `kruskal.test()` function to test whether the medians of an unnamed variable of three unconnected populations (`a`, `b` and `c`) with 10 individuals each are truly different:

```
set.seed(42)
a <- rnorm(n = 10, mean = 15, sd = 3)
b <- rnorm(n = 10, mean = 10, sd = 3)
c <- rnorm(n = 10, mean = 5, sd = 3)
groups <- as.factor(rep(c("a", "b", "c"), each = 10))

kruskal.test(x = c(a, b, c), g = groups)
```

```
##
##  Kruskal-Wallis rank sum test
##
## data:  c(a, b, c) and groups
## Kruskal-Wallis chi-squared = 20, df = 2, p-value =
## 7e-05
```

Purpose And Assumptions

Friedman Test

`friedman.test()` in base `R`

Purpose:	To identify whether groups of variable values are different from one another.
H_0	*There is no difference in characteristics of the response variable values in dependence of the classes of the predictor variable.*
Assumptions:	■ Predictor variable is categorical (not binary!) ■ Response variable is ordinal or metric ■ Variable values are **dependent** (paired)

Minimal Working Example

Let's use the `friedman.test()` function to test whether the medians of an unnamed variable of three connected samples (`a`, `b` and `c`) with 10 individuals each (i.e. one re-sampled population) are truly different:

```
set.seed(42)
a <- rnorm(n = 10, mean = 15, sd = 3)
b <- rnorm(n = 10, mean = 10, sd = 3)
c <- rnorm(n = 10, mean = 5, sd = 3)
TestData <- matrix(c(a, b, c), nrow = 10, byrow = FALSE,
    dimnames = list(1:10, c("a", "b", "c")))

friedman.test(y = TestData)
```

```
##
##  Friedman rank sum test
##
## data:  TestData
## Friedman chi-squared = 10, df = 2, p-value = 7e-04
```

Variables We Can Use

Response variables (metric/ordinal)

- Weight
- Height
- Wing Chord
- Nesting Height
- Number of Eggs
- Egg Weight
- Home Range

Predictor variables (categorical but not binary)

- Home Range (3 levels - Small, Medium, Large)
- Site Index (11 levels)
- Predator Presence/Type (3 levels - Avian vs. Non-Avian vs. None)
- Climate (3 levels - Continental, Semi-Coastal, Coastal)

Research Questions And Hypotheses

So which of our major research questions (seminar 6) can we answer?

Kruskal-Wallis Test

- *Site-wise variation:* Do characteristics of *Passer domesticus* depend on the site they are recorded at?
- *Climate Warming/Extremes:* Does sparrow morphology depend on climate?
- *Predation:* Does nesting height depend on predator characteristics?
- *Competition:* Does home range depend on climate?

Friedman Test (suppose a resettling program)

- *Climate Warming/Extremes:* Does sparrow morphology change depend on climate?
- *Predation:* Does nesting height depend on predator characteristics?
- *Competition:* Does home range depend on climate?

$\rightarrow$ Find the data in
2a - Sparrow_ResettledSIMA_READY.rds
and *2b - Sparrow_ResettledSIUK_READY.rds*

Writng A Function In R

Establishing *user-defined* functions is at the **heart of 'R'**!

A function requires:

+ A **name**
+ **Arguments**
+ to be **called**

```
Fun <- function(argument) {
    print(argument)
}
Fun(argument = "Test")
```

```
## [1] "Test"
```