Estimating the Repeatability of a Variable using ANOVA

Paul M. Magwene

Measurement error

- Nearly every measure of a continuous measurement or assay we apply in biology (and other sicences) has associated with it some "measurement error".
- This is not "biological variation" of interest, but rather "technical variation" associated with our ability to measure quantities of interest precisely

Examples:

- Have three people independently measure the length of a human femur tot he nearest millimeter. Unless the values are aggressively rounded, there is a high likelihood that you'll get three different values.

Repeatability

ANOVA can be used to estimate the **Repeatability** of a measure, which provides a way to quanify how much of the variance we observe between individuals is due to measurement error.

Experimental steps

- Repeatedly but independently measure the same variable in the same individual or other unit of observation
- Carry out same repeated measurements across individuals

Repeatability: Statistical steps

- Calculate MS_{groups} and MS_{error} where individuals are the grouping unit.
- Estimate the variance among groups, s_A^2 as:

$$s_A^2 = \frac{MS_{\text{groups}} - MS_{\text{error}}}{n}$$

  - where n is the number of replicates per individual

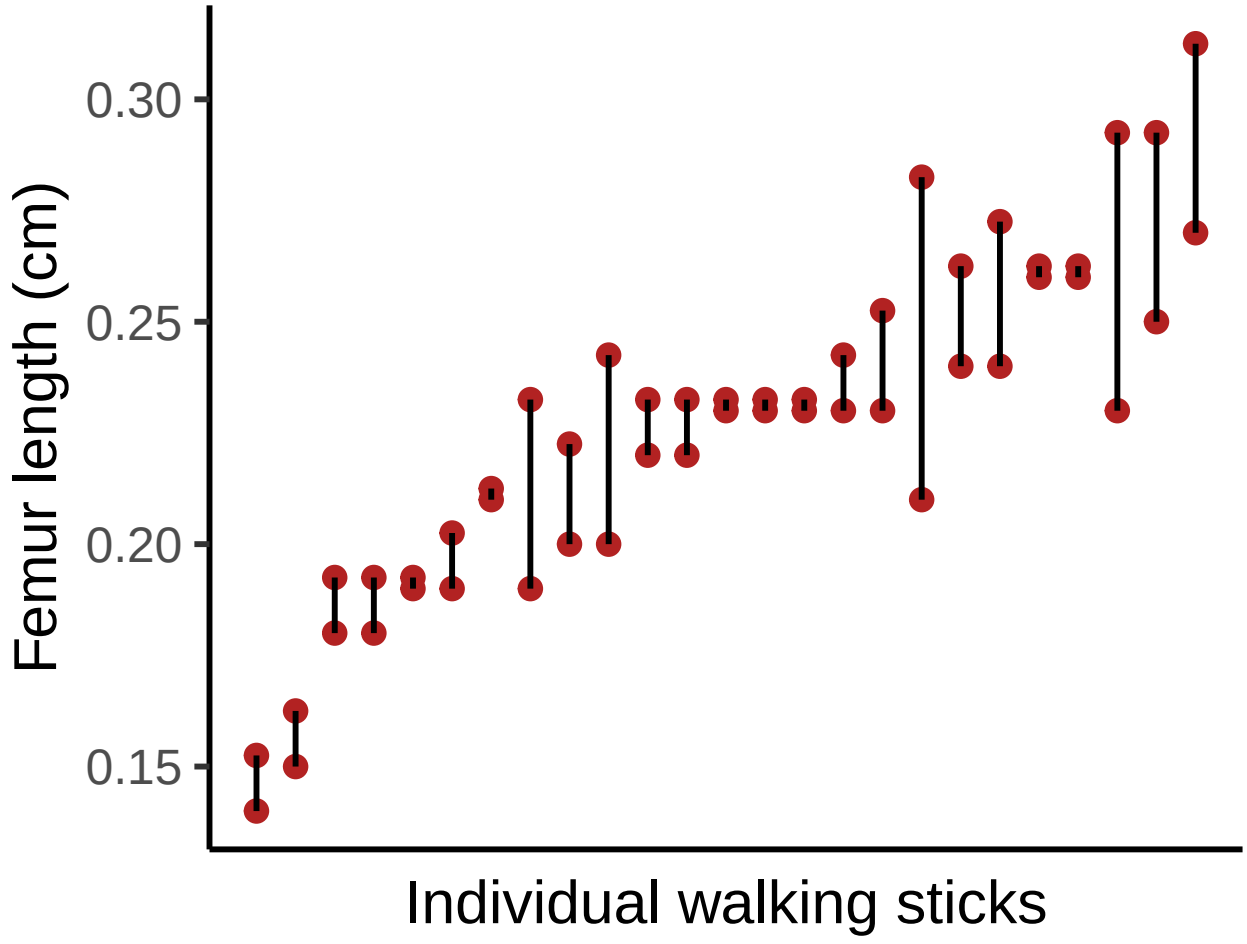
- Estimate the repeatability as:

$$\text{Repeatability} = \frac{s_A^2}{s_A^2 + MS_{\text{error}}}$$

- $0 \leq \text{Repeatability} \leq 1$; values near zero indicate nearly all variance is due to measurement error, values near one indicate small fraction of variance due to measurement error

Repeatability: Walking stick example

Nosil and Crespi measured various morphological features of walking stick insects from digital photographs. For each specimen they took two independent photographs and measured femur length (in cm) on each photograph.



Repeatability: Walking stick analysis

```
sticks.aov <- aov(femurLength ~ as.factor(specimen),
                  data = walking.sticks)
sticks.table <- tidy(sticks.aov)
sticks.table
```

```
# A tibble: 2 x 6
  term                 df  sumsq   meansq statistic     p.value
  <chr>             <dbl>  <dbl>    <dbl>     <dbl>       <dbl>
1 as.factor(specimen)    24 0.0591 0.00246       6.92  0.00000408
2 Residuals              25 0.0089 0.000356     NA    NA
```

```
MS.groups <- sticks.table$meansq[1]
MS.error <- sticks.table$meansq[2]
sA = (MS.groups - MS.error)/2
repeatability = sA/(sA + MS.error)
```

From the table we calculate:

- $s_A^2 = 0.0011$
- Repeatability $= 0.7475$