Using **pretend made up data** to pretend to evaluate impacts of mean annual precipitation and collection method on insect biomass in Hawaiʻi

Andy Rominger

## Background

The insects of Hawaiʻi represent an especially unique assemblage owing to their high levels of endemicity (Gillespie and Roderick 2002) and evolutionary divergence to occupy novel ecosystems through adaptive radiation. Up to 98% of insect in Hawaiʻi are endemic, meaning they are found nowhere else on Earth.

This high level of endemicity and multiple anthropocentric stressers—including habitat destruction, climate change, and introduction of invasive species—make Hawaiian insects a high priority for conservation and restoration (Medeiros et al. 2013). Invasive insects pose multiple threats both to endemic and native biodiversity on the islands, and also agricultural systems.

Despite the priority of Hawaiian insects for conservation, and impact of invasive insects, there remains insufficient study of how insect populations and species respond to environmental gradients (Medeiros et al. 2013), and even which methods are most effective for monitoring invasive species (Borges et al. 2018).

Here we analyze a *totally fake simulated dataset* to help answer two questions:

1. how does total biomass of insects change across gradients of rainfall across the Hawaiian islands?
2. are malaise traps or pitfall traps more effective in detecting the concerning invasive *Species 1*?

## Analysis and results

We will make use of several packages to analyze and visualize these data. We load those packages like this:

library(dplyr)  
library(ggplot2)  
library(cowplot)

Next we read-in the dataset:

dat <- read.csv("data/insect\_map.csv")  
head(dat)

We can see our data have columns indicating the name of the site where data where collected, the mean annual precipitation (MAP) at that site, and then information about the species identities of the insects collected, and their sampled masses from each of the two collecting methods: malaise and pitfall.

Now we can analyze the data to address our two questions.

### How does total biomass of insects change across gradients of rainfall across the Hawaiian islands?

To evaluate how biomass changes across MAP, we first need to prepare a suitably-formatted data.frame.

Now we can visualize the relationship between MAP and biomass

ggplot(dat\_no\_spp, aes(x = MAP, y = total\_mass)) +  
 geom\_point() +  
 theme\_cowplot() +  
 xlab("Mean annual precipitation (mm/yr)") +  
 ylab("Total insect biomass (g)")

|  |
| --- |
|  |

These data seem well suited to a correlation test, and visually do not appear to need transformation to meet the normality assumptions of correlation. To test if there is a significant correlation we use the cor.test function and find:

cor.test(dat\_no\_spp$MAP, dat\_no\_spp$total\_mass)

Pearson's product-moment correlation  
  
data: dat\_no\_spp$MAP and dat\_no\_spp$total\_mass  
t = 8.3598, df = 8, p-value = 3.177e-05  
alternative hypothesis: true correlation is not equal to 0  
95 percent confidence interval:  
 0.7870093 0.9877622  
sample estimates:  
 cor   
0.9472517

The null hypothesis is that **there is zero correlation between MAP and total mass** (i.e. ). The alternative hypothesis is that **MAP and total mass have a non-zero correlation** (i.e. ).

The test statistic is the sample correlation coefficient 0.9472517. And the null distribution is the distribution with 8. The -value is which is less than so we reject the null of no correlation. Furthermore, the 95% confidence interval for is [0.787, 0.988], which is positive and far from 0, so we conclude there is strong evidence for a positive correlation between MAP and biomass.

### Are malaise traps or pitfall traps more effective in detecting the concerning invasive *Species 1*?

*Species 1* is a known invasive that does such and such bad things. We would like to know which sampling method works best to monitor this species.

Here again we will need to do some data manipulations to properly format the data for our visualization and analysis needs

Now we can visualize the biomass for *Species 1* found in each trap type

# notice that we're subsetting the data to just look at species 1  
ggplot(subset(dat\_by\_trap, dat\_by\_trap$species == "species\_1"),   
 aes(x = trap\_type, y = total\_mass)) +  
 geom\_boxplot() +   
 theme\_cowplot() +  
 xlab("Trap type") +  
 ylab("Biomass of Species 1 (g)")

|  |
| --- |
|  |

Here we have a categorical explanatory variable and a numerical response variable. This is well-suited for a -test. Because we have two categories, we might be tempted to do a two-sample -test.

Our hypotheses are:

* :
* :

Because each replicate of the pitfall traps and malaise traps are located **at the same site**, we will use a *paired* -test. The results of which are

t.test(total\_mass ~ trap\_type,   
 data = subset(dat\_by\_trap, dat\_by\_trap$species == "species\_1"),  
 paired = TRUE)

Paired t-test  
  
data: total\_mass by trap\_type  
t = -5.9705, df = 9, p-value = 0.00021  
alternative hypothesis: true mean difference is not equal to 0  
95 percent confidence interval:  
 -4.923264 -2.217636  
sample estimates:  
mean difference   
 -3.57045

The -value is far less than so we reject the null of no difference between the traps. From the visualization and the 95% confidence interval of the mean difference between groups, we see that malaise traps catch much less biomass of *Species 1* so we conclude pitfall traps are the more effective way to monitor for this species.

## References

Borges, Paulo AV, Pedro Cardoso, Holger Kreft, Robert J Whittaker, Simone Fattorini, Brent C Emerson, Artur Gil, et al. 2018. “Global Island Monitoring Scheme (GIMS): A Proposal for the Long-Term Coordinated Survey and Monitoring of Native Island Forest Biota.” *Biodiversity and Conservation* 27: 2567–86.

Gillespie, Rosemary G, and George K Roderick. 2002. “Arthropods on Islands: Colonization, Speciation, and Conservation.” *Annual Review of Entomology* 47 (1): 595–632.

Medeiros, Matthew J, Jesse A Eiben, William P Haines, Raina Kaholoaa, Cynthia King, Paul D Krushelnycky, Karl N Magnacca, Daniel Rubinoff, Forest Starr, and Kim Starr. 2013. “The Importance of Insect Monitoring to Conservation Actions in Hawaii.”