Anolis lizards of Bimini

---

51 marks

See the Anolis Lizards background file for the context of this question.

Of interest is whether the two species of lizard occupy different perch habitats.

Assuming you have downloaded the data files into a directory saved on `dataDirectory` you create `lizards` either as

or as

```
file <- file.path(dataDirectory, "AnolisLizards.csv")
lizards <- read.csv(file, header=TRUE)
```

Just set `eval = TRUE` for your preferred choice.

a. The **Problem** stage.

   i. *(3 marks)* What is the target population here? What are the units making up the population?

   - The target population here is the population of the 4 species of Anolis lizards that co-exists on Bimini
   - Anolissagrei, distichus, angusticeps and carolinensis species make up the population

b. The **Plan** stage.

   i. *(3 marks)* What is the study population? What are the units making up the population.

      - The study population are the two species of lizards that are observed during the period November 18 to 30, 1966, between 9:00 A.M. and 4:30 P.M.
      - Sagrei and Angusticeps species make up the population.

   ii. *(3 marks)* What are the variates? Which are response? Which are explanatory?

      - Variates: perch_height_ft, perch_diameter_inches, and count
      - response: count
      - explanatory: perch_height_ft, perch_diameter_inches

   iii. *(3 marks)* Describe the sampling plan.

- The sampling plan used in this study is simple random sampling. Because the samples are collected only if they are seen by the observers and the observers does not have any idea what species will show up next. However, I believe this sampling plan is different from anything we have learned in course.

   iv. *(2 marks)* Describe the measuring system mentioned.

      – The measuring system used here is feet for perch height and inches for perch diameters. This is because the perch height could be really tall, so it is better to use feet as a measurement, and the diameters of a perch is not expected to be too big, therefore it is measured in inches in this case.

   v. *(2 marks)* Is this study experimental or observational? Explain your reasoning.'

  - The study is observational because all of the data is observed and counted in a certain period of time.

c. The **Data** stage.

   i. *(1 mark)* The `lizards` data is a data frame having 8 rows with the counts in one column. Another way to represent this data is as a $2 \times 2 \times 2$ table of counts. Use the `xtabs()` function with an appropriate formula to construct such a table from `lizards`. Assign the table to the variable `lizardsTable`. That is complete the following:

```
lizardsTable <- xtabs(... your arguments here ...)
```

Note that `lizardsTable` will not be a `data.frame` but a `table` and an `array`

```
lizardsTable <- xtabs(count ~ perch_height_ft + perch_diameter_inches + species, data = lizard
lizardsTable
```

```
## , , species = angusticeps
##
##                perch_diameter_inches
## perch_height_ft  "<= 4.25"  "> 4.25"
##          "< 3"           3         2
##          ">= 3"         21         1
##
## , , species = sagrei
##
##                perch_diameter_inches
## perch_height_ft  "<= 4.25"  "> 4.25"
##          "< 3"          48        84
##          ">= 3"         15        18
```

   ii. *(2 marks)* Using `lizardTables` and the `sum` function, programmatically determine how many lizards there were of each species. Show how the same results can be obtained with `margin.table()`.

```
# angusticeps
sum(lizardsTable[1:4])
```

```
## [1] 27
```

```
# sagrei
sum(lizardsTable[5:8])
```

```
## [1] 165
```

```
margin.table(lizardsTable[1:4])
```

```
## [1] 27
```

```
margin.table(lizardsTable[5:8])
```

```
## [1] 165
```

d. The **Analysis** stage.

   i. *(10 marks)* Use `xtabs()` with an appropriate formula on the original `lizards` data frame to construct a smaller table that ignores species and just contains the counts at each perch habitat. Call this table `lizardsHabitat`. You will now do a small analysis based only on this table. That is,

      - Complete the following:

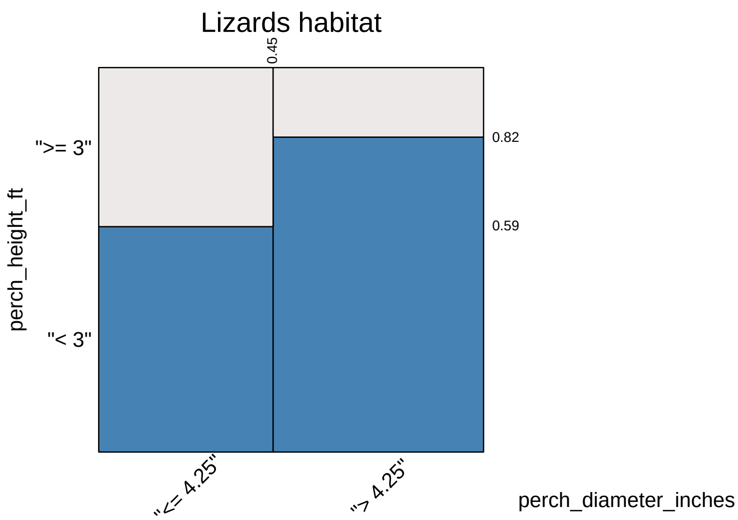
```r
lizardsHabitat <- xtabs(... your arguments here ...)
lizardsHabitat
```

- Construct an eikosogram of perch height versus perch diameter using the `lizardsHabitat` table.
- Call `summary(lizardsHabitat)`.
- Show your code.

```r
lizardsHabitat <- xtabs(count ~ perch_height_ft + perch_diameter_inches, data = lizards)
lizardsHabitat
```

```
##                perch_diameter_inches
## perch_height_ft  "<= 4.25"  "> 4.25"
##          "< 3"          51        86
##          ">= 3"         36        19
```

```r
# Eikosogram perch height versus perch diameter
library(eikosograms)
eikos(perch_height_ft ~ perch_diameter_inches, data = lizardsHabitat, xlab_rot = 45, ispace = list(
```




```r
summary(lizardsHabitat)
```

```
## Call: xtabs(formula = count ~ perch_height_ft + perch_diameter_inches,
##     data = lizards)
## Number of cases in table: 192
## Number of factors: 2
## Test for independence of all factors:
##    Chisq = 12.62, df = 1, p-value = 0.0003817
```

```
Based on all of the above results, what can you conclude about the distribution of the perch of th
```

- From the above results, we can easily see that the two heights of the two vertical bars are quite different, this suggests that the perch_height and the perch_diameter may not be independent variates for the anolis lizards. In the test for independence of all factors, we see small p-value, which is an evidence against the independent hypothesis.

```
What does this say about the preferred habitat of these lizards? Explain.
```

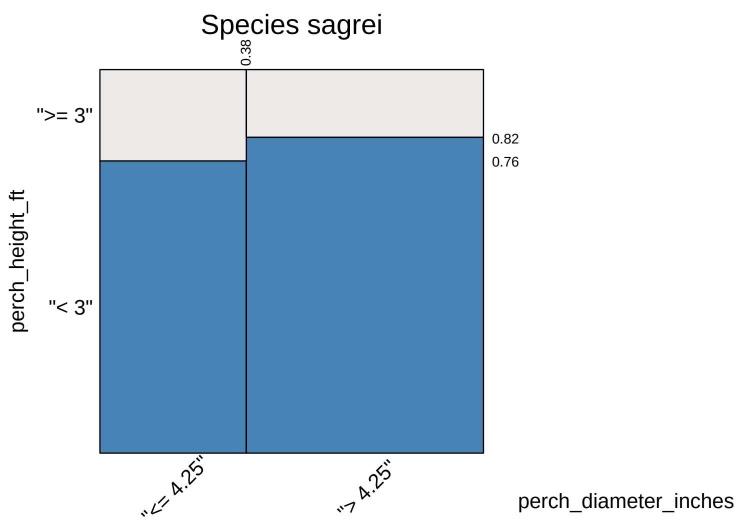
- The eikosogram above shows that the most preferred habitat of these lizards is when perch_diameter_inches > 4.25 and perch_height_ft <3.

ii. *(6 marks)* Repeat the analysis of part (i) above but this time on the sub-table of `lizardsTable` (i.e. no `xtabs()` call) of only the counts for `sagrei`. In addition to calling `summary()`, also call `chisq.test()` on the relevant table **but** use a significance level determined by simulation.

```
# only sagrei, count = 165
#sagreiDataframe <- lizards[lizards$species == "sagrei", -1]
#sagreiTable <- xtabs(count ~ perch_height_ft + perch_diameter_inches, data = sagreiDataframe)
sagreiTable <- lizardsTable[,1:2,2]
sagreiTable
```

```
##                  perch_diameter_inches
## perch_height_ft  "<= 4.25"  "> 4.25"
##           "< 3"        48        84
##           ">= 3"       15        18
```

```
eikos(perch_height_ft ~ perch_diameter_inches, data = sagreiTable, xlab_rot = 45, ispace = list(bottom =
```




```
summary(sagreiTable)
```

```
## Number of cases in table: 165
## Number of factors: 2
## Test for independence of all factors:
##  Chisq = 0.9244, df = 1, p-value = 0.3363
```

```
chisq.test(sagreiTable, simulate.p.value = TRUE)
```

```
##
##  Pearson's Chi-squared test with simulated p-value (based on 2000
##  replicates)
##
## data:  sagreiTable
## X-squared = 0.92437, df = NA, p-value = 0.4128
```

- From the above results, we can easily see that the two heights of the two vertical bars are quite similar, this suggests that the perch_height and the perch_diameter may be independent variates for the sagrei species. In the chi-square test for independence of sagrei lizard, we do not see evidence against the hypothesis.

iii. *(6 marks)* Construct the sub-table of `lizardsTable` (i.e. no `xtabs()` call) containing only the counts for `angusticeps`. Using this table, construct the eikosogram of perch height versus perch diameter. Also call `chisq.test()` on this table with a significance level determined by simulation.

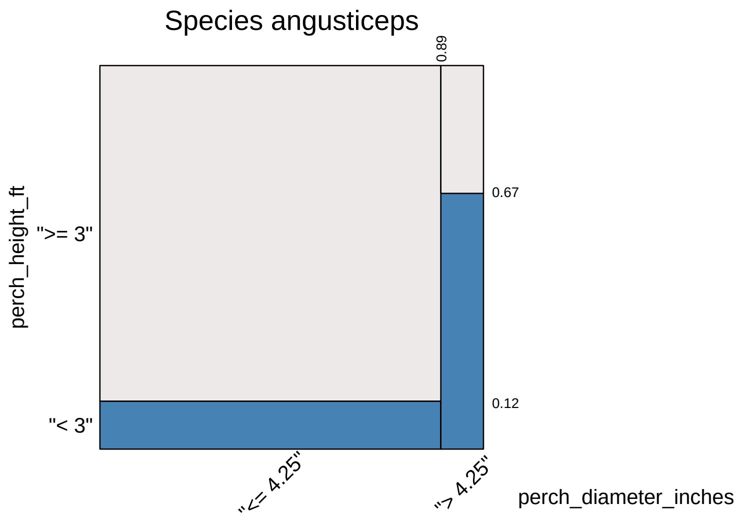
– Do the interpretations from the eikosogram and the chi-squared tests agree? How do they agree/disagree?

– Explain why this has occurred here.

```
#angusticepsDataframe <- lizards[lizards$species == "angusticeps", -1]
#angusticepsTable <- xtabs(count ~ perch_height_ft + perch_diameter_inches, data = angusticepsDataframe
angusticepsTable <- lizardsTable[,1:2,1]
angusticepsTable
```

```
##                perch_diameter_inches
## perch_height_ft  "<= 4.25"  "> 4.25"
##          "< 3"          3         2
##          ">= 3"        21         1
```

```
eikos(perch_height_ft ~ perch_diameter_inches, data = angusticepsTable, xlab_rot = 45, ispace = list(bo
```




```
summary(angusticepsTable)
```

```
## Number of cases in table: 27
## Number of factors: 2
## Test for independence of all factors:
##  Chisq = 5.185, df = 1, p-value = 0.02278
##  Chi-squared approximation may be incorrect
```

```
chisq.test(angusticepsTable,simulate.p.value = TRUE)
```

```
##
##  Pearson's Chi-squared test with simulated p-value (based on 2000
##  replicates)
##
## data:  angusticepsTable
## X-squared = 5.1852, df = NA, p-value = 0.08246
```

- From the above results, we can easily see that the two heights of the two vertical bars are not even close, this suggests that the perch_height and the perch_diameter are not independent variates for the angusticeps species. For the eikosogram and chi-square, they both suggest that the the perch_height and the perch_diameter are not independent variates. However, there appears a warning that "Chi-squared approximation may be incorrect". This is due to there are so little observations for the angusticeps species, with a total of 27. With this amount of observation, any tests will show no evidence against the test for independence.

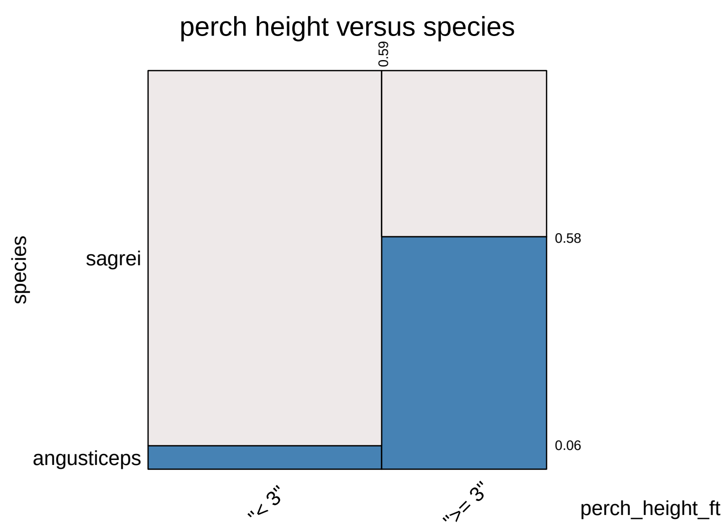
iv. *(4 marks)* Construct the sub-table of `lizardsTable` (i.e. no `xtabs()` call) containing only the counts for lizards on perches that are less than 4.25 inches in diameter. Using this table, construct the eikosogram of perch height versus species. Also call `chisq.test()` on this table with a significance level determined by simulation.

What do you conclude, if anything, about the preference in habitat for the two species? Justify your conclusions.

```
#lessDiaDataframe <- lizards[lizards$perch_diameter_inches == " \"<= 4.25\"", c(1,2,3,4)]
#lessDiaTable <- xtabs(count ~ perch_height_ft + species, data = lessDiaDataframe)
lessDiaTable <- lizardsTable[1:2,1,1:2]
lessDiaTable
```

```
##                  species
## perch_height_ft angusticeps sagrei
##          "< 3"            3     48
##          ">= 3"          21     15
```

```
eikos(species ~ perch_height_ft, data = lessDiaTable, xlab_rot = 45, ispace = list(bottom = 20), main =
```




```
chisq.test(lessDiaTable, simulate.p.value = TRUE)
```

```
##
##  Pearson's Chi-squared test with simulated p-value (based on 2000
##  replicates)
##
## data:  lessDiaTable
## X-squared = 29.063, df = NA, p-value = 0.0004998
```

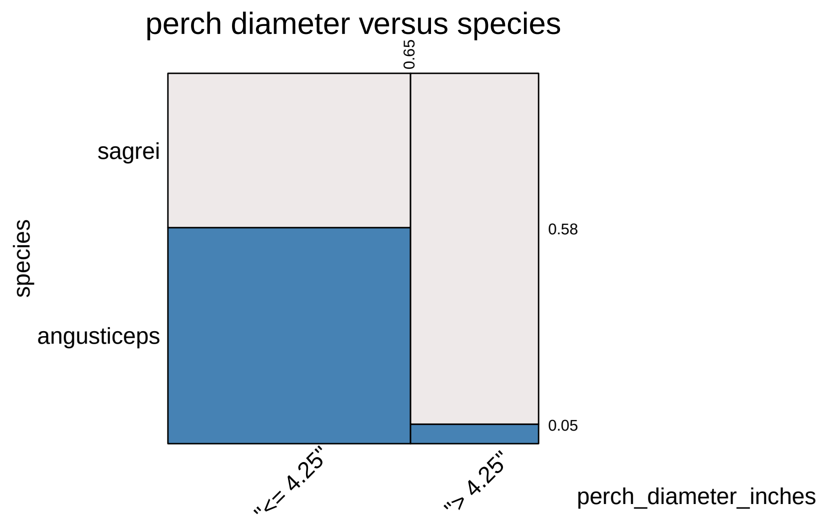
- We can conclude that the sagrei species prefer perch_height_ft that is less than 3 ft. On the other hand, the angusticeps species prefer perch_height_ft that is greater or equal to 3 ft. By looking at the eikosogram, it is obvious that the two bars are not at the same level, this indicates that the species and perch_height_ft are not independent. When performing chi-square independent test, the p-value is really small, which is an evidence against the independent hypothesis.

iv. *(4 marks)* Construct the sub-table of `lizardsTable` (i.e. no `xtabs()` call) containing only the counts for lizards on perches that are three feet or higher in height. Using this table, construct the eikosogram of perch diameter versus species. Also call `chisq.test()` on this table with a significance level determined by simulation.

```r
#greaterHeightDataframe <- lizards[lizards$perch_height_ft == " \">= 3\"", c(1,2,3,4)]
#greaterHeightTable <- xtabs(count ~ perch_diameter_inches + species, data = greaterHeightDataframe)
greaterHeightTable <- lizardsTable[2,1:2,1:2]
greaterHeightTable
```

```
##                       species
## perch_diameter_inches angusticeps sagrei
##             "<= 4.25"          21     15
##             "> 4.25"            1     18
```

```r
eikos(species ~ perch_diameter_inches, data = greaterHeightTable, xlab_rot = 45, ispace = list(bottom =
```




```r
chisq.test(greaterHeightTable, simulate.p.value = TRUE)
```

```
##
##  Pearson's Chi-squared test with simulated p-value (based on 2000
##  replicates)
##
## data:  greaterHeightTable
## X-squared = 14.594, df = NA, p-value = 0.0004998
```

- We can conclude that the sagrei species prefers a little more towards perch_diameter that is greater than 4.25 inches. On the other hand, the angusticeps species prefer perch_diameter that is less than or equal to 4.25 inches. By looking at the eikosogram, it is obvious that the two bars are not at the same level, this indicates that the species and perch_height_ft are not independent. When performing chi-square independent test, the p-value is really small, which is an evidence against the independent hypothesis.

e. The **Conclusions** stage:

   i. *(2 marks)* A common name for *anolis angusticeps* is the *Cuban Twig Anole*. In light of your analysis explain why, or why not, this name might have become common?

- This name might not become common because its prefer habitat is in a higher perch with a smaller diameter. Therefore, it is hard to observe the species, so it may not be well-knowned.