

# SMMS Sternite Dataset Analysis

Gregory K. Ross

University of Florida

(mailto:#)gkross@ufl.edu (mailto:gkross@ufl.edu)

- 1 Introduction
- 2 Data
- 3 Analysis
  - 3.1 AEG/ALB Analysis (Both Species)
    - 3.1.1 AEG/ALB Data Summary
    - 3.1.2 AEG/ALB Data Description
    - 3.1.3 AEG/ALB Data Dispersion
    - 3.1.4 AEG/ALB Data Distribution
    - 3.1.5 AEG/ALB Hexbin of Scutum vs Stergum
    - 3.1.6 AEG/ALB Analysis Discussion
  - 3.2 AEG Analysis
    - 3.2.1 AEG Data Summary
    - 3.2.2 AEG Data Description
    - 3.2.3 AEG Data Frequency Counts
    - 3.2.4 AEG Data Dispersion
    - 3.2.5 AEG Data Distribution
    - 3.2.6 AEG Data Chi-Square Test for Independence
    - 3.2.7 AEG Data Mosaic Plot
    - 3.2.8 AEG Analysis Discussion
  - 3.3 ALB Analysis
    - 3.3.1 ALB Data Summary
    - 3.3.2 ALB Data Description
    - 3.3.3 ALB Data Frequency Counts
    - 3.3.4 ALB Data Dispersion
    - 3.3.5 ALB Data Distribution
    - 3.3.6 ALB Data Chi-Square Test for Independence
    - 3.3.7 ALB Data Mosaic Plot
    - 3.3.8 ALB Analysis Discussion
  - 3.4 Between Species Data (AEG and ALB)
    - 3.4.1 Between Species Data Frequency Counts
    - 3.4.2 Between Species Data Scutum Distribution
    - 3.4.3 Between Species Data Scutum Distribution - Stacked
    - 3.4.4 Between Species Data Scutum Spine Plot
    - 3.4.5 Between Species Data Scutum Chi-Square Test for Independence
    - 3.4.6 Between Species Data Scutum Mosaic Plot
    - 3.4.7 Between Species Data Stergum Distribution
    - 3.4.8 Between Species Data Stergum Distribution - Stacked
    - 3.4.9 Between Species Data Stergum Spine Plot
    - 3.4.10 Between Species Data Stergum Chi-Square Test for Independence
    - 3.4.11 Between Species Data Stergum Mosaic Plot
    - 3.4.12 Between Species Analysis Discussion

## 1 Introduction



This analysis was performed by the University of Florida - Florida Medical Entomology Laboratory, 200 9th Street SE, Vero Beach FL, 32962. This analysis is done in a reproducible research manner. All Tables, Figures and Statistical Analyses were created in the open source software package R.

**Date:** 10/26/2016

**Version:** 2.0

## 2 Data

A copy of this data analysis (PDF) may be downloaded here (<https://gatorbox.rc.ufl.edu/index.php/s/tYayMuuQ8iowUd8>).

A copy of the original data (XLSX) may be downloaded here (<https://gatorbox.rc.ufl.edu/index.php/s/aStOBibUNnTcPDJ>).

A copy of the cleaned data (CSV) used for the analysis may be downloaded here (<https://gatorbox.rc.ufl.edu/index.php/s/iMHCgooLtCEpN6Y>).

**\*\*Notes:\*\*** (Notes:\*\*) The original data had 1244 records. One (1) record was modified due to a data entry error. All values of 0.02 or 0.05 were changed to 0.10 for categorical data analysis.

# 3 Analysis

## 3.1 AEG/ALB Analysis (Both Species)

### 3.1.1 AEG/ALB Data Summary

Table 1. Summary of Sternite Dataset

SCUTUM	STERGUM
Min. :0.0000	Min. :0.0000
1st Qu.:0.1000	1st Qu.:0.8000
Median :0.2000	Median :1.0000
Mean :0.3372	Mean :0.8665
3rd Qu.:0.6000	3rd Qu.:1.0000
Max. :1.0000	Max. :1.0000

### 3.1.2 AEG/ALB Data Description

Table 2. Description of Sternite Dataset

	vars	n	mean	sd	median	trimmed	mad	min	max	range	skew	kurtosis	se
SCUTUM	1	1241	0.3372	0.3283	0.2	0.3015	0.2965	0	1	1	0.7975	-0.8095	0.00932
STERGUM	2	1241	0.8665	0.226	1	0.9224	0	0	1	1	-2.042	3.381	0.006415

### 3.1.3 AEG/ALB Data Dispersion

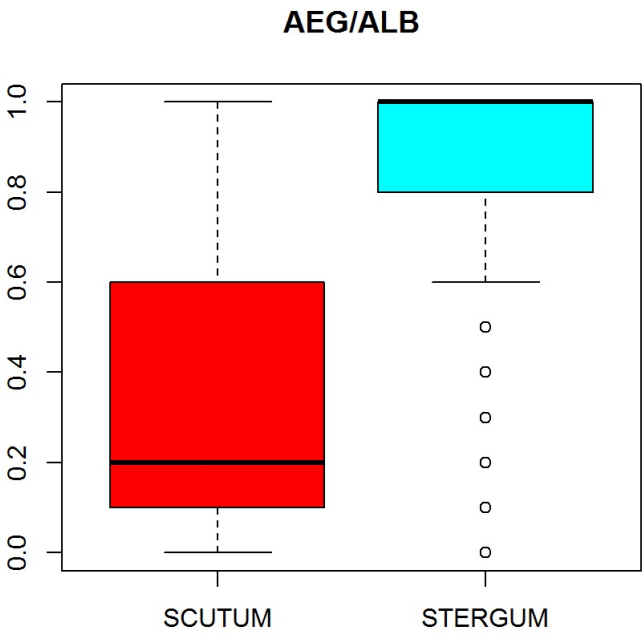


Figure 1.

### 3.1.4 AEG/ALB Data Distribution

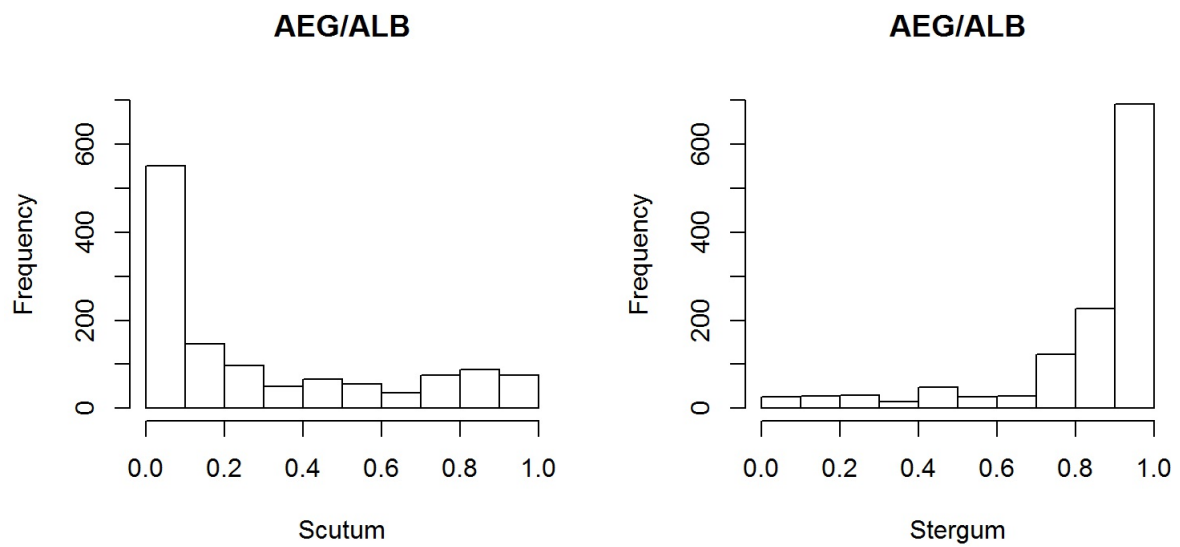


Figure 2.

### 3.1.5 AEG/ALB Hexbin of Scutum vs Stergum

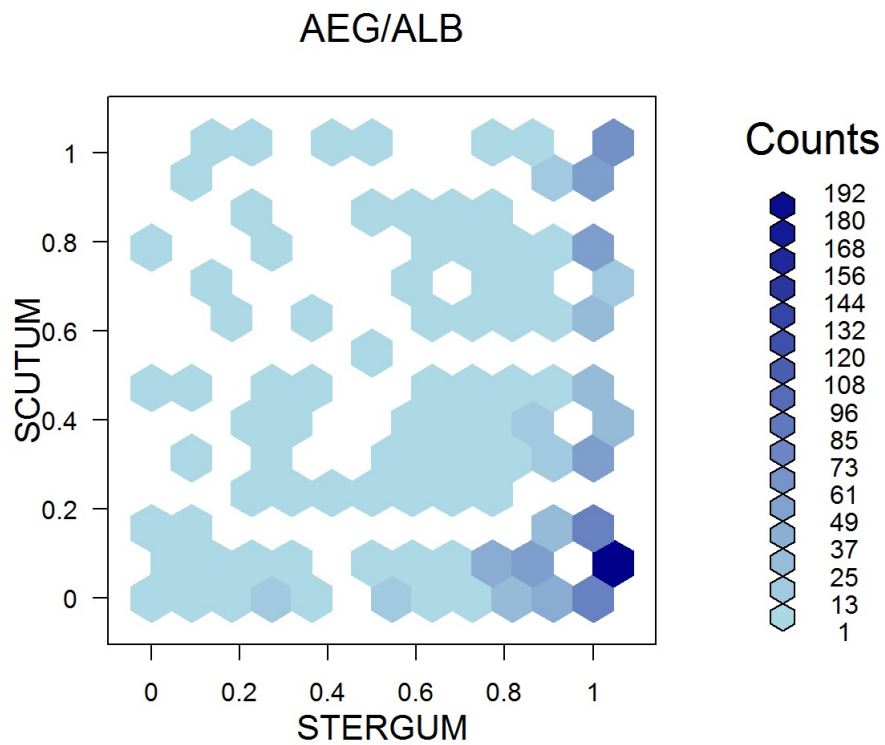


Figure 3.

### 3.1.6 AEG/ALB Analysis Discussion

When looking at both species as a collective dataset, the scutum and stergum data appear to be heavily skewed, though in opposite directions. Figure 3 indicates that when the stergum is skewed in the direction of 100% coverage and the scutum is skewed in the direction of 0% coverage.

3.2 AEG Analysis

3.2.1 AEG Data Summary

Table 3. Summary of AEG Dataset

SCUTUM	STERGUM
Min. :0.0000	Min. :0.0000
1st Qu.:0.4000	1st Qu.:0.9000
Median :0.6000	Median :1.0000
Mean :0.6026	Mean :0.9132
3rd Qu.:0.9000	3rd Qu.:1.0000
Max. :1.0000	Max. :1.0000

3.2.2 AEG Data Description

Table 4. Description of AEG Dataset

	vars	n	mean	sd	median	trimmed	mad	min	max	range	skew	kurtosis	se
SCUTUM	1	227	0.6026	0.2969	0.6	0.6224	0.2965	0	1	1	-0.4569	-0.8901	0.0197
STERGUM	2	227	0.9132	0.1735	1	0.9552	0	0	1	1	-3.094	10.62	0.01151

3.2.3 AEG Data Frequency Counts

Table 5. AEG Frequencies

AEG\$SCUTUM	0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1
AEG\$STERGUM											
0	1	0	0	0	0	1	0	0	0	0	0
0.1	0	0	0	1	0	0	0	0	0	0	0
0.2	1	0	0	0	0	0	0	0	0	0	0
0.3	2	0	0	0	1	0	0	0	1	0	0
0.5	0	1	0	0	0	1	0	0	0	1	0
0.6	0	0	0	0	1	1	1	0	1	0	0
0.7	1	0	0	1	1	0	1	0	0	1	0
0.8	3	3	1	1	1	3	1	1	4	1	1
0.9	1	0	4	6	3	6	4	5	4	11	2
1	4	6	4	11	11	10	16	14	22	22	21

3.2.4 AEG Data Dispersion

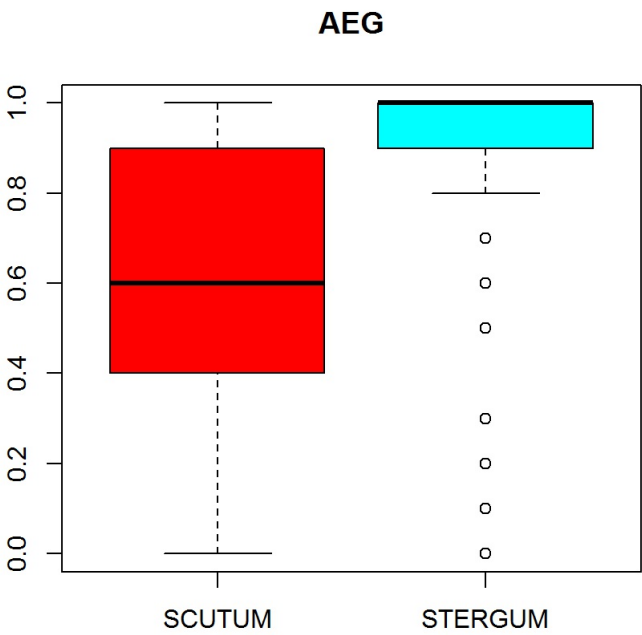


Figure 6.

3.2.5 AEG Data Distribution

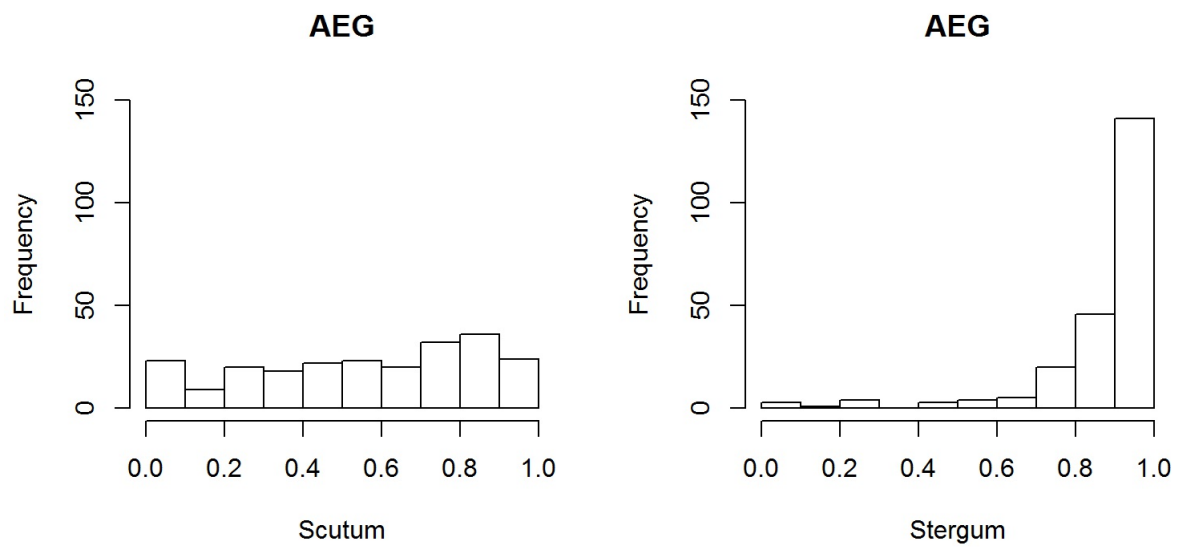


Figure 7.

### 3.2.6 AEG Data Chi-Square Test for Independence

Table 6. AEG Chi-Square Test for Independence

Test statistic	df	P value
111.2	90	0.06401

### 3.2.7 AEG Data Mosaic Plot

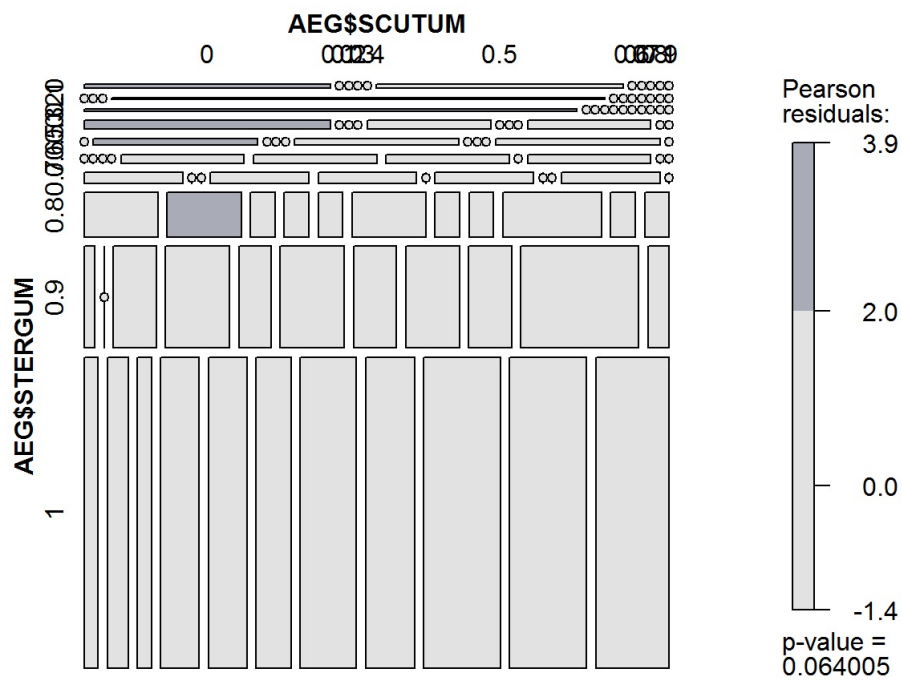


Figure 8.

### 3.2.8 AEG Analysis Discussion

When looking at the AEG species, the stergum dataset appears to be heavily skewed while the scutum appears to be closer to normal. Table 6 and Figure 8 indicate that there are no significant differences between AEG scutum and stergum coverages at the 0.05 level.

3.3 ALB Analysis

3.3.1 ALB Data Summary

Table 7. Summary of ALB Dataset

SCUTUM	STERGUM
Min. :0.0000	Min. :0.000
1st Qu.:0.1000	1st Qu.:0.800
Median :0.1000	Median :1.000
Mean :0.2778	Mean :0.856
3rd Qu.:0.4000	3rd Qu.:1.000
Max. :1.0000	Max. :1.000

3.3.2 ALB Data Description

Table 8. Description of ALB Dataset (continued below)

	vars	n	mean	sd	median	trimmed	mad	min
SCUTUM	1	1014	0.2778	0.3049	0.1	0.2287	0.1483	0
STERGUM	2	1014	0.856	0.2349	1	0.9121	0	0

	max	range	skew	kurtosis	se
SCUTUM	1	1	1.205	0.1296	0.009575
STERGUM	1	1	-1.886	2.623	0.007378

3.3.3 ALB Data Frequency Counts

Table 9. ALB Frequencies

ALB\$SCUTUM	0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1
ALB\$STERGUM											
0	1	1	1	0	0	0	0	0	1	0	0
0.1	5	4	4	0	0	1	0	1	0	2	2
0.2	10	3	8	0	1	0	1	0	0	2	1
0.3	12	7	3	2	0	1	0	0	1	0	0
0.4	9	0	1	0	0	1	3	0	0	0	1
0.5	15	9	7	4	0	5	1	0	0	2	2
0.6	4	10	3	1	0	0	1	1	1	1	0
0.7	6	10	3	0	1	2	0	0	1	0	0
0.8	28	39	9	8	4	4	2	2	5	2	0
0.9	39	60	21	15	11	5	7	3	6	13	1
1	70	186	78	48	14	25	17	9	29	30	44

3.3.4 ALB Data Dispersion

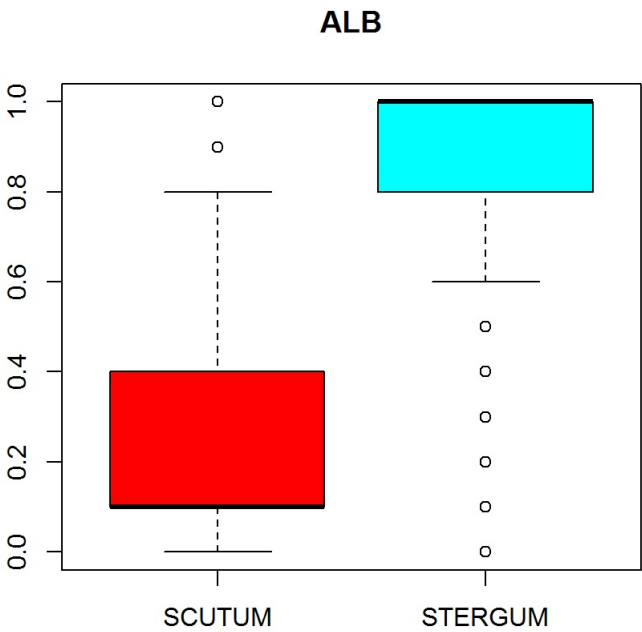


Figure 9.

3.3.5 ALB Data Distribution

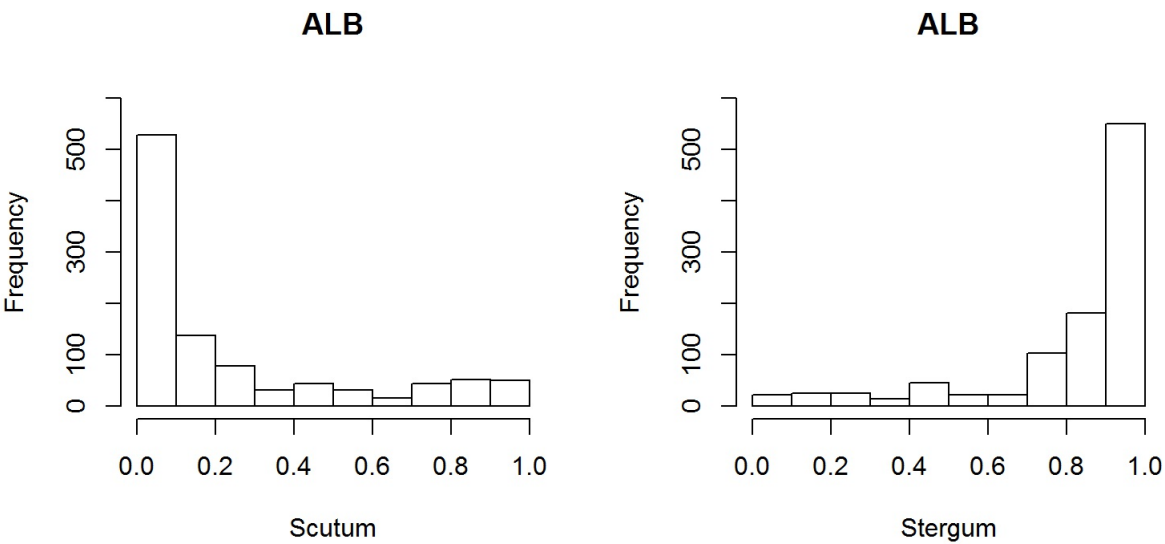


Figure 10.

3.3.6 ALB Data Chi-Square Test for Independence

Table 10. ALB Chi-Square Test for Independence

Test statistic	df	P value
168.8	100	2.073e-05 ***

3.3.7 ALB Data Mosaic Plot

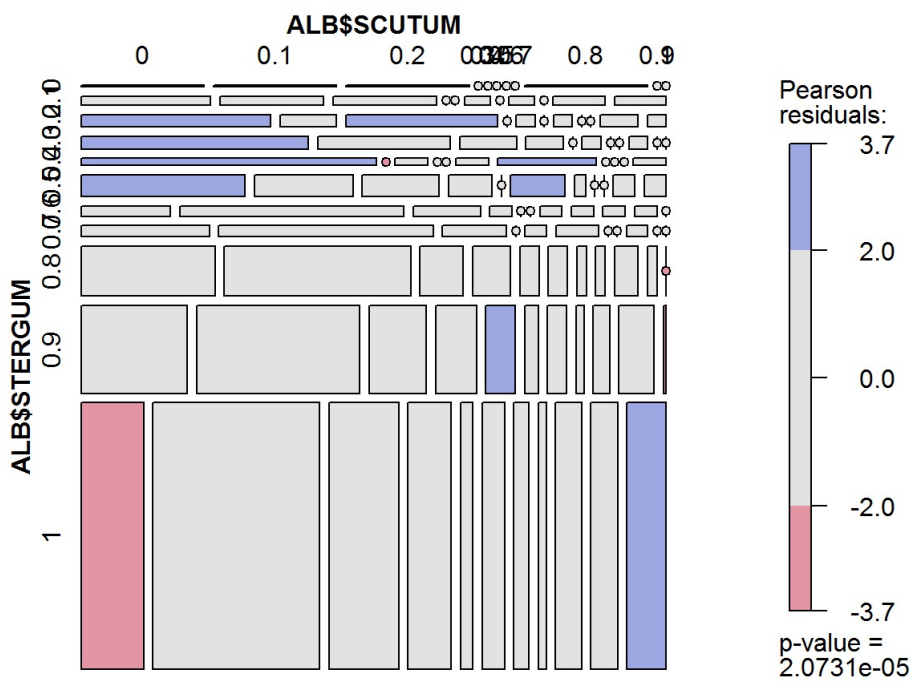


Figure 11.

3.3.8 ALB Analysis Discussion

When looking at the ALB species, the scutum and stergum datasets appear to be heavily skewed, though in opposite directions. Table 10 and Figure 11 indicate that there are significant differences between AEG scutum and stergum coverages at the 0.05 and 0.01 level.

3.4 Between Species Data (AEG and ALB)

3.4.1 Between Species Data Frequency Counts

Table 11. Stergum Frequencies

	STERGUM	0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1
SPECIES												
AEG		2	1	1	4	0	3	4	5	20	46	141
ALB		4	19	26	26	15	45	22	23	103	181	550

Table 12. Scutum Frequencies

	SCUTUM	0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1
SPECIES												
AEG		13	10	9	20	18	22	23	20	32	36	24
ALB		199	329	138	78	31	44	32	16	44	52	51

3.4.2 Between Species Data Scutum Distribution

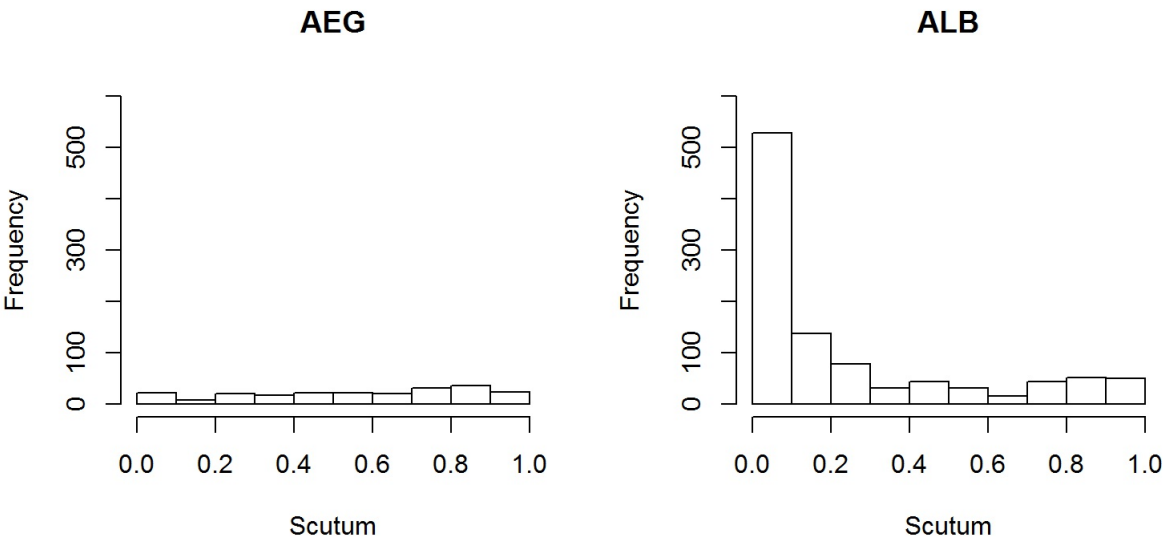


Figure 12.

3.4.3 Between Species Data Scutum Distribution - Stacked

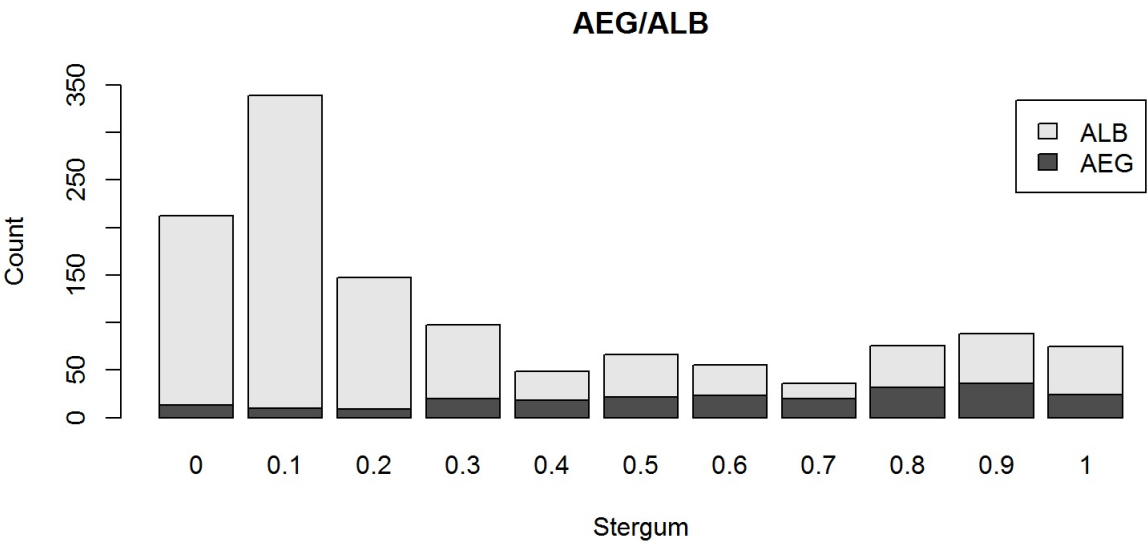


Figure 13.

3.4.4 Between Species Data Scutum Spine Plot



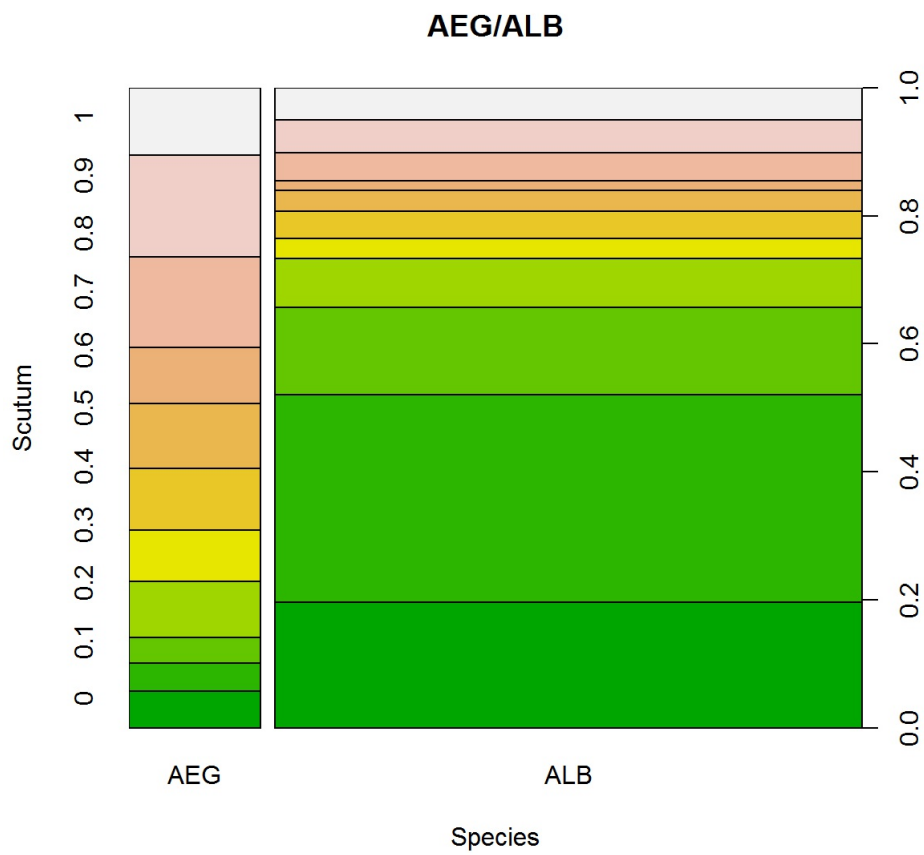


Figure 14.

### 3.4.5 Between Species Data Scutum Chi-Square Test for Independence

Table 13. Between Species Scutum Chi-Square Test for Independence

Test statistic	df	P value
232.6	10	2.492e-44 * * *

### 3.4.6 Between Species Data Scutum Mosaic Plot

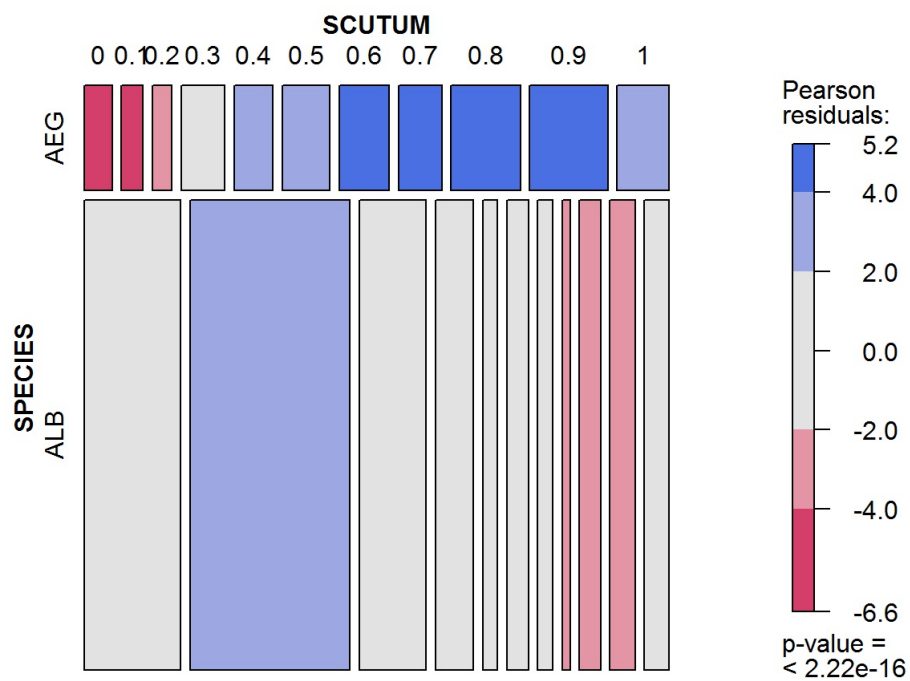


Figure 15.

3.4.7 Between Species Data Stergum Distribution

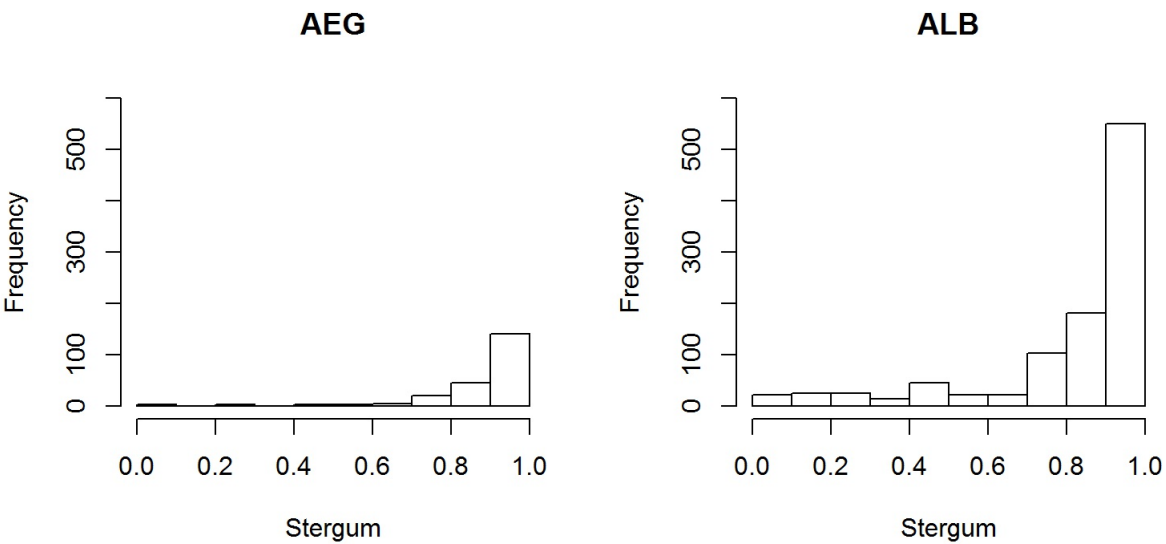


Figure 16.

3.4.8 Between Species Data Stergum Distribution - Stacked

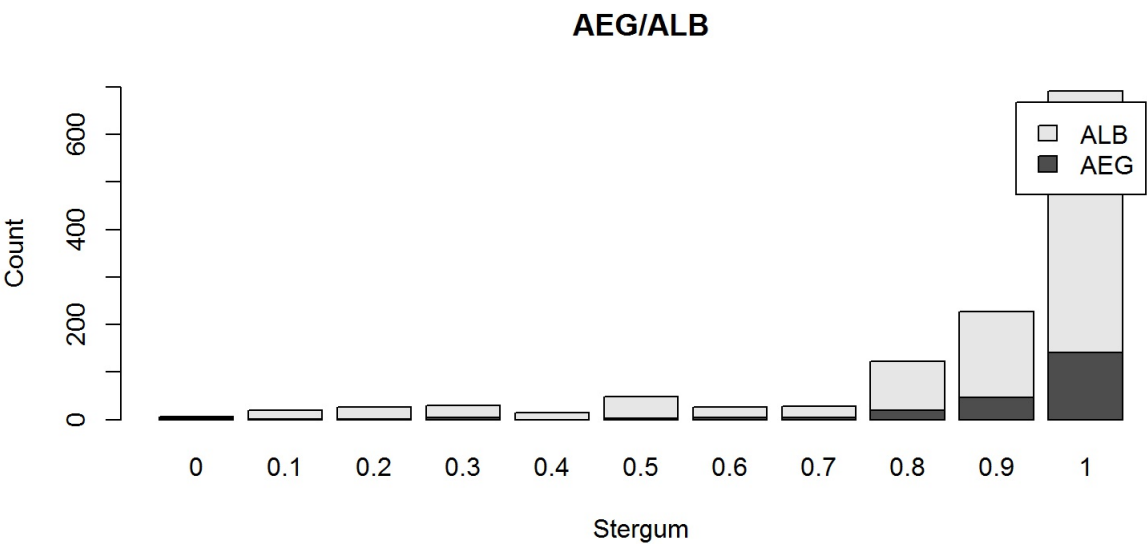


Figure 17.

3.4.9 Between Species Data Stergum Spine Plot

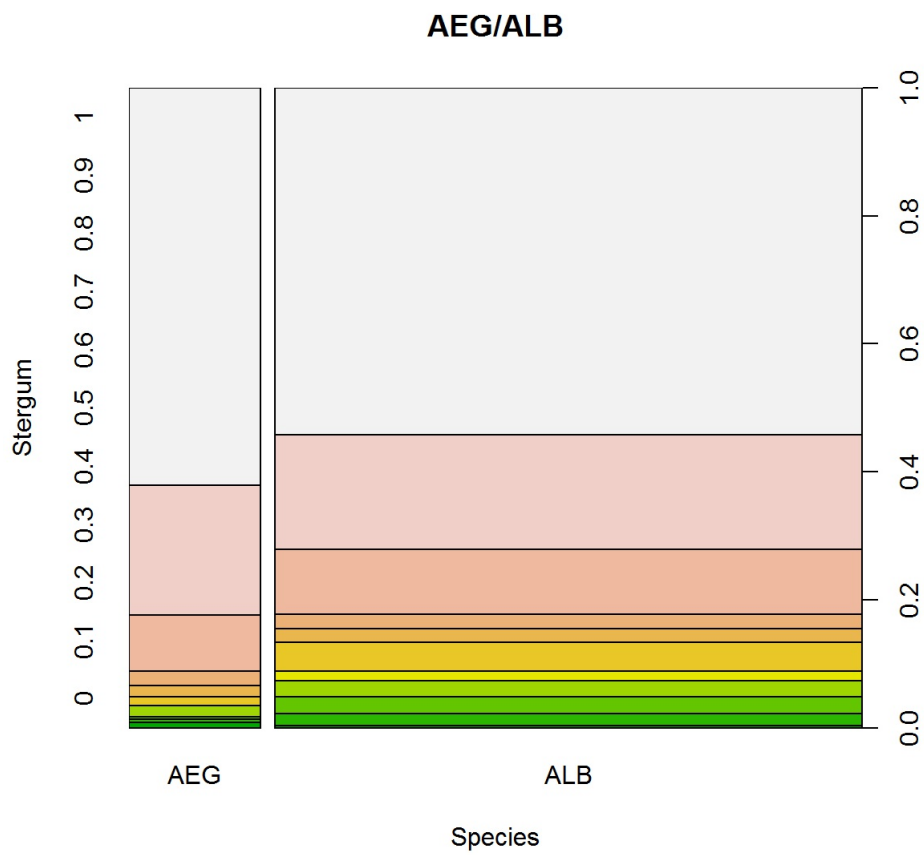


Figure 18.

### 3.4.10 Between Species Data Stergum Chi-Square Test for Independence

Table 14. Between Species Stergum Chi-Square Test for Independence

Test statistic	df	P value
18.77	10	0.04326 *

### 3.4.11 Between Species Data Stergum Mosaic Plot

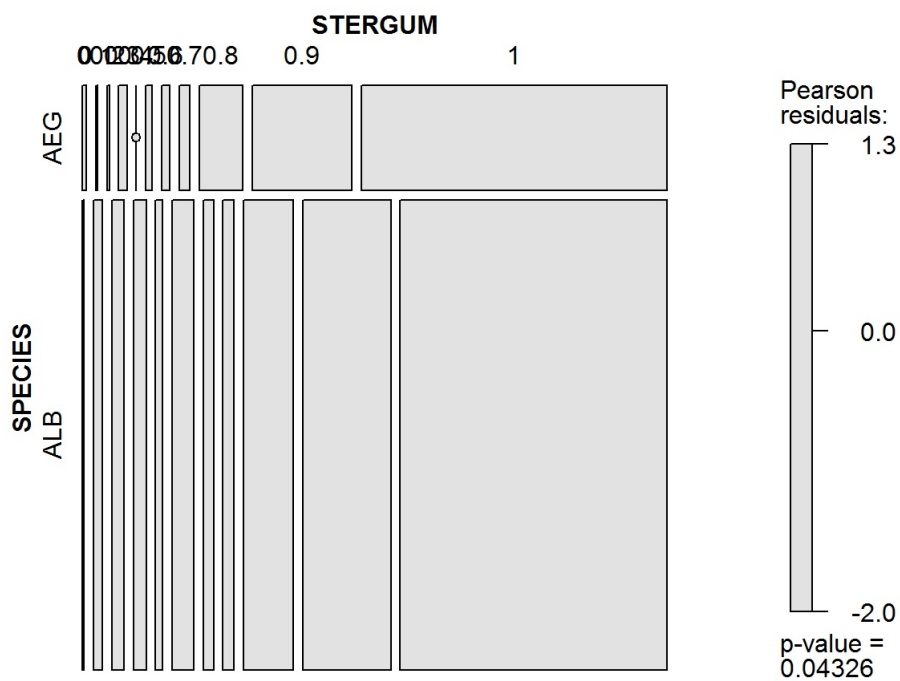


Figure 19.

### 3.4.12 Between Species Analysis Discussion

When looking between the AEG and ALB species, the stergum datasets appear to be heavily skewed in the same direction while the scutum datstets appear to be different between species. **Table 13** indicates that there are *significant differences between AEG and ALB species scutum coverages* at the 0.05 and 0.01 level. **Figure 15** shows that range of 0.6 - 0.9 scutum coverages for AEG had *significantly more* values than expected and the range of 0.0 - 0.1 scutum coverages for AEG had *significantly less* values than expected. **Figure 15** also indicates that for ALB the range of 0.7 - 0.9 scutum coverages had *less* values than expected **Table 14** indicates that there are *significant differences between AEG and ALB species stergum coverages* at the 0.05 level. **Figure 19** shows that no coverages for either species had *significantly* more or less values than expected.