Palmer's Agave & Lehmann Lovegrass

Jeff Oliver

March 25, 2020

Summary

A trio of questions on agave (*Agave palmeri*) survival and growth as effected by the invasive, Lehmann lovegrass (*Eragrostis lehmanniana*). To address these questions, I used a mixed-model approach, so the plot variable could be included as a random effect.

0. Data summary

Following data cleaning, the total number of replicates and plants in each treatment is:

Treatment	# Replicates	# Plants
C	4	20
H	5	25
J	5	25
J+H	4	20
J+S	5	25
J+W	6	30
S	5	25
S+H	5	25
S+W	7	35
W	4	20

1. How does treatment affect agave survival and size?

The two response variables are:

1. Survival of individual agave plants
2. The number of leaves on agave plants; the number of agaves measured varies among plot/treatment combinations

The inclusion of treatment in the models can be done in two different ways:

1. Treatment is treated as a single predictor variable, with 10 levels, including the control.
2. Treatment is separated into multiple binary predictor variables; indicating whether or not a specific treatment (i.e. Javalina protection) was applied.

1.1 How does treatment affect agave survival *s.s.*?

In this model, survival is treated as a binary response variable and a logistic regression model is applied (N = 250).

1.1.a Treatment as a single predictor

For the method of predicting based on treatment, there is a single fixed-effect in the model:

$$Log{-}odds\ \ Survival = \beta_0 + \beta_1 \times Treatment + b_0$$

Where b_0 is the random intercept for plot.

Predictor	Coefficient Estimate	Error	P-value	Survival Probability
(Intercept)	-0.81657	0.81719	0.31767	0.30649
H	0.16441	0.70243	0.81494	0.34250
J	0.16447	0.70242	0.81487	0.34252
J+H	-0.85779	0.76587	0.2627	0.15784
J+S	3.07192	0.88807	0.00054	0.90511
J+W	1.75895	0.73992	0.01744	0.71958
S	1.50570	0.74035	0.04197	0.66577
S+H	-0.86220	0.79210	0.27638	0.15726
S+W	-0.76281	0.67871	0.26105	0.17088
W	-1.69443	0.83332	0.04202	0.07509

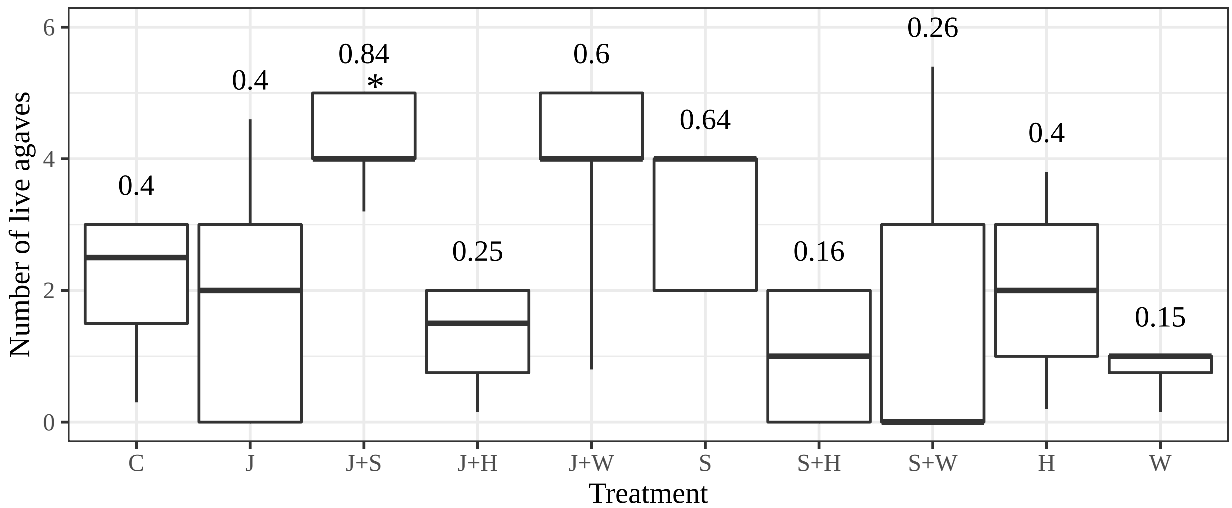
In this table, the row listed as (*Intercept*) represents the Control treatment.

Post-hoc comparisons among treatments, showing results of Tukey test for significant differences from the Control treatment:

Predictor	t value	P-value
H	-0.23406	1.00000
J	-0.23415	1.00000
J+H	1.12003	0.98281
J+S	-3.45909	0.01935
J+W	-2.37721	0.33952
S	-2.03377	0.57489
S+H	1.08849	0.98592
S+W	1.12391	0.98240
W	2.03334	0.57520

Note: The p-values from Tukey *post-hoc* tests will *always* be higher than p-value from the initial model. This is because the *post-hoc* test is correcting for the fact that we are making many pairwise comparisons (we have to compare each treatment to every other treatment), and the chance that we encounter a false positive (inferring significance where the low p-value is due to chance alone) is increasing. In this case, there are 45 total pairwise comparisons.

Figure 1.1. Agave survival by treatment



Boxplots show median (center line), 25th and 75th percentiles (lower and upper box boundaries, respectively), and 5th and 95th percentiles (lower and upper whisker, respectively) Asterisks indicate treatments that had a significantly different effect from the Control treatment in Tukey *post-hoc* tests.

Aside: interpreting logistic regression

Interpreting the results of logistic regression in terms of probabilities requires examination of how the log-odds model works. We are primarily interested in the probability of survival for given values of $Treatment$, i.e. $p(Survival|Treatment)$. We can model this with the logistic function:

$$p(Survival|Treatment) = \frac{e^{\beta_0+\beta_1 Treatment}}{1+e^{\beta_0+\beta_1 Treatment}}$$

However, the relationship between our variable of interest ($p(Survival)$) and the predictor ($Treatment$) is not linear. Using some algebraic rearranging, we arrive at the logit model:

$$p(Survival|Treatment) = \frac{e^{\beta_0+\beta_1 Treatment}}{1+e^{\beta_0+\beta_1 Treatment}}$$

$$p(Survival|Treatment) \times (1+e^{\beta_0+\beta_1 Treatment}) = e^{\beta_0+\beta_1 Treatment}$$

$$p(Survival|Treatment) + p(Survival|Treatment) \times e^{\beta_0+\beta_1 Treatment} = e^{\beta_0+\beta_1 Treatment}$$

$$p(Survival|Treatment) = e^{\beta_0+\beta_1 Treatment} - p(Survival|Treatment) \times e^{\beta_0+\beta_1 Treatment}$$

$$p(Survival|Treatment) = e^{\beta_0+\beta_1 Treatment} \times (1 - p(Survival|Treatment))$$

$$\frac{p(Survival|Treatment)}{1-p(Survival|Treatment)} = e^{\beta_0+\beta_1 Treatment}$$

Finally, we take the natural logarithm of both sides to get:

$$ln\left(\frac{p(Survival|Treatment)}{1-p(Survival|Treatment)}\right) = \beta_0 + \beta_1 Treatment$$

Which is the familiar log-odds model from above:

$$Log{-}odds\ \ Survival = \beta_0 + \beta_1 \times Treatment$$

So how do we interpret the β coefficients from the model? For this model, let us compare the probability of survival in the Control treatment and the probability of survival in the hand-pulling treatment. We can use the original probability model, but because it is the control, we drop the $\beta_1 Treatment$ terms because $Treatment$ in this case is zero.

$$p(Survival|Treatment = Control) = \frac{e^{\beta_0}}{1+e^{\beta_0}}$$

substituting in the value for $\beta_0 = -0.817$,

$$p(Survival|Treatment = Control) = \frac{e^{-0.817}}{1+e^{-0.817}}$$

And the estimated probability for survival in the Control group is then 0.306.

For the survival probability of the hand-pulling treatment, we start with the same probability model:

$$p(Survival|Treatment = Hand-pulling) = \frac{e^{\beta_0+\beta_1 Treatment}}{1+e^{\beta_0+\beta_1 Treatment}}$$

Set $Treatment$ equal to 1 and use the coefficient estimate from the table above for $\beta_1 = 0.164$:

$$p(Survival|Treatment = Hand-pulling) = \frac{e^{-0.817+0.164}}{1+e^{-0.817+0.164}}$$

And the estimated probability for survival in the Hand-pulling treatment is then 0.343.

Finally, we can compare the two by calculating the difference in probabilities, 0.343 - 0.306 = 0.037.

We can consider a similar comparison, between the Control treatment and the Javalina exclusion and shade treatment ("J + S"):

$$p(Survival|Treatment = Javalina + Shade) = \frac{e^{-0.817+3.072}}{1+e^{-0.817+3.072}}$$

So the estimated probability of survival in the J + S treatment is 0.905 and the difference in survival probability from the Control treatment is 0.905 - 0.306 = 0.599.

Note the explanations above do not discuss random effects of plot. So when reporting these values, it is best to say that this is the *average* effect of treatment on survival.

1.1.b Treatment as multiple predictors

The alternative approach is to treat each type of treatment as a separate variable:

$$Log{-}odds\ \ Survival = \beta_0 + \beta_1 \times Javalina + \beta_2 \times Shade + \beta_3 \times Weed\ \ eating + \beta_4 \times Hand\ \ pulling + b_0$$

Where b_0 is the random intercept for plot. This simple model does *not* incorporate interaction effects, although a richer model could include those. For example, the interaction between Javalina protection and Shade is significant (results not shown) and may explain the difference between Javalina protection alone and Javalina protection plus shade in 1.1.a, above.

Predictor	Coefficient Estimate	Error	P-value
(Intercept)	-0.6342	0.6085	0.2973
Javalina	1.1926	0.3432	5e-04
Shade	0.6755	0.3369	0.0449
Weed eating	-0.8751	0.3541	0.0135
Hand pulling	-1.1323	0.3907	0.0038

In this table, the row listed as (*Intercept*) represents the Control treatment.

1.1.c Treatment as multiple predictors with interaction effects

If this model is expanded to include interaction effects:

$$\begin{aligned} Log{-}odds\ \ Survival = {} & \beta_0 + \beta_1 \times Javalina + \beta_2 \times Shade + \beta_3 \times Weed\ \ eating + \beta_4 \times Hand\ \ pulling \\ & + \beta_5 \times Javalina \times Shade + \beta_6 \times Javalina \times Weed\ \ eating \\ & + \beta_7 \times Javalina \times Hand\ \ pulling + \beta_8 \times Shade \times Weed\ \ eating \\ & + \beta_9 \times Shade \times Hand\ \ pulling + b_0 \end{aligned}$$

Predictor	Coefficient Estimate	Error	P-value
(Intercept)	-0.8152	0.8172	0.3185
Javalina	0.1629	0.7025	0.8166
Shade	1.5041	0.7403	0.0422
Weed eating	-1.6962	0.8335	0.0418
Hand pulling	0.1631	0.7024	0.8164
Javalina x Shade	1.4038	1.0911	0.1982
Javalina x Weed eating	3.2909	1.0849	0.0024
Javalina x Hand pulling	-1.1846	1.0250	0.2478
Shade x Weed eating	-0.5717	1.0639	0.591
Shade x Hand pulling	-2.5305	1.0395	0.0149

In this table, the row listed as (*Intercept*) represents the Control treatment.

1.2. How does treatment affect agave size?

In this section, individual agave sizes (measured by number of leaves) is a continuous response variable in linear regression mixed-effects models. The data are restricted to cases where there was a live agave (N = 74). Furthermore, only measurements for a maximum of three agaves per row were used in analyses.

1.2.a Treatment as a single predictor

The first model considers Treatment as a single predictor variable.

$$\#Leaves = \beta_0 + \beta_1 \times Treatment + b_0$$

Where b_0 is the random intercept for plot.

Predictor	Coefficient Estimate	Error	P-value
(Intercept)	4.75985	1.01455	3e-05
H	2.18513	1.29757	0.09726
J	1.04089	1.47315	0.4824
J+H	2.25015	1.63663	0.17416
J+S	2.50682	1.18344	0.03807
J+W	2.14523	1.22489	0.0848
S	3.28090	1.20697	0.00848
S+H	0.70625	1.66647	0.67314
S+W	2.75130	2.13626	0.20249
W	2.74840	2.84783	0.33824

In this table, the row listed as $(Intercept)$ represents the Control treatment.

Note: There are several Plot by Treatment combinations with zero measurements for leaf count and at least one Treatment level (W) where there was a *single* leaf count measurement. This could potentially influence this particular analysis.

Aside: Interpreting linear regression

In contrast to logistic regression, linear regression is a straighforward additive model:

$$\#Leaves = \beta_0 + \beta_1 \times Treatment + b_0$$

For this explanation, we will ignore the random intercept effect, b_0, so this formula becomes

$$\#Leaves = \beta_0 + \beta_1 \times Treatment$$

In our model, β_0 is the intercept of the model, and β_1 is the coefficient for the Treatment of interest. For the Control treatment, $\beta_1 = 0$, so the formula for the number of leaves of agaves in the Control treatment becomes

$$\#Leaves(Treatment = Control) = \beta_0$$

or

$$\#Leaves(Treatment = Control) = 4.76$$

To calculate the agave size for any other treatment, we use the formula

$$\#Leaves = \beta_0 + \beta_1$$

substituting the β_1 with the coefficient of the Treatment of interest. For example, if we wanted to know the size of an agave in the Hand-pulling treatment, we substitute $\beta_0 = 4.76$ and $\beta_1 = 2.185$, for an average size of 6.945 leaves.

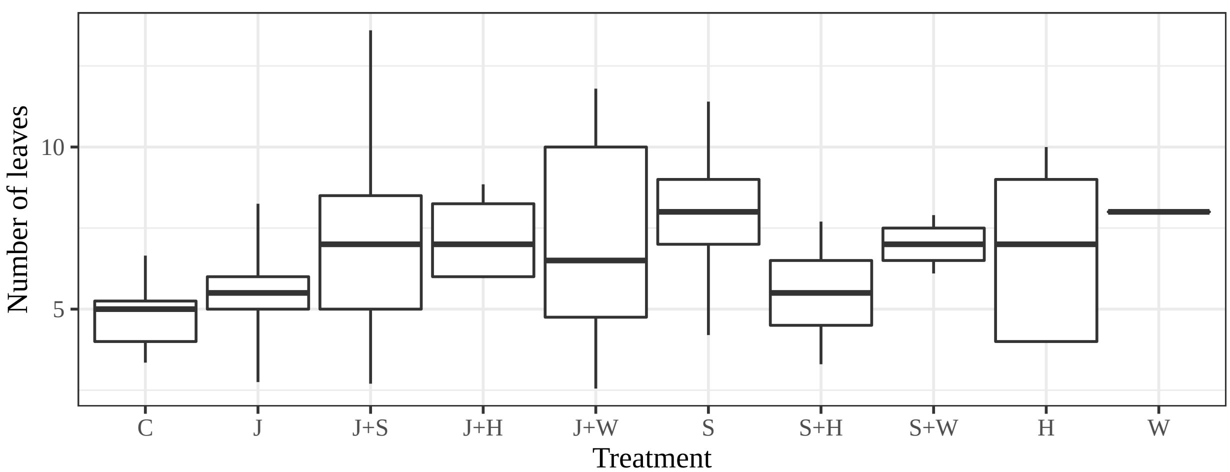
Another way of interpreting these coefficients is that they show the *difference* in size from the control treatment. That is, the coefficient for the Hand-pulling treatment, 2.185, is the difference in size between plants in the Control treatment and plants in the Hand-pulling treatment.

Post-hoc comparisons among treatments, showing results of Tukey test for significant differences from the Control treatment:

Predictor	t value	P-value
H	-1.67763	0.80335
J	-0.68726	0.99951
J+H	-1.36833	0.93212
J+S	-2.08411	0.54515
J+W	-1.73469	0.77155
S	-2.69210	0.19957
S+H	-0.41327	0.99999
S+W	-1.26585	0.95754
W	-0.95536	0.99369

Figure 1.2. Agave size by treatment

Numbers in plot show mean survivorship for treatment



Boxplots show median (center line), 25th and 75th percentiles (lower and upper box boundaries, respectively), and 5th and 95th percentiles (lower and upper whisker, respectively)

1.2.b Treatment as multiple predictors

The second model separates out the four treatment types into separate variables:

$$\#Leaves = \beta_0 + \beta_1 \times Javalina + \beta_2 \times Shade + \beta_3 \times Weed\ eating + \beta_4 \times Hand\ pulling + b_0$$

Where b_0 is the random intercept for plot.

Predictor	Coefficient Estimate	Error	P-value
(Intercept)	5.82051	0.83025	1.5673e-07

Predictor	Coefficient Estimate	Error	P-value
Javalina	0.17606	0.68238	0.79718
Shade	1.37375	0.72472	0.06226
Weed eating	0.90326	0.90924	0.32411
Hand pulling	0.43212	0.85794	0.61616

1.2.c Treatment as multiple predictors with interaction effects

A model with interaction effects is:

$$
\begin{aligned}
\#Leaves = \beta_0 + \beta_1 \times Javalina + \beta_2 \times Shade + \beta_3 \times Weed\ eating + \beta_4 \times Hand\ pulling \\
+ \beta_5 \times Javalina \times Shade + \beta_6 \times Javalina \times Weed\ eating \\
+ \beta_7 \times Javalina \times Hand\ pulling + \beta_8 \times Shade \times Weed\ eating \\
+ \beta_9 \times Shade \times Hand\ pulling + b_0
\end{aligned}
$$

Predictor	Coefficient Estimate	Error	P-value
(Intercept)	4.75985	1.01455	3e-05
Javalina	1.04089	1.47315	0.4824
Shade	3.28090	1.20697	0.00848
Weed eating	2.74840	2.84783	0.33824
Hand pulling	2.18513	1.29757	0.09726
Javalina x Shade	-1.81497	1.78028	0.31187
Javalina x Weed eating	-1.64406	3.13672	0.60206
Javalina x Hand pulling	-0.97587	2.18658	0.6569
Shade x Weed eating	-3.27800	3.54199	0.35822
Shade x Hand pulling	-4.75978	2.02176	0.02169

2. How do the W and H treatments and agaves affect the percent of Lehmann cover?

In these analyses, the effect of agaves is modeled in two ways:

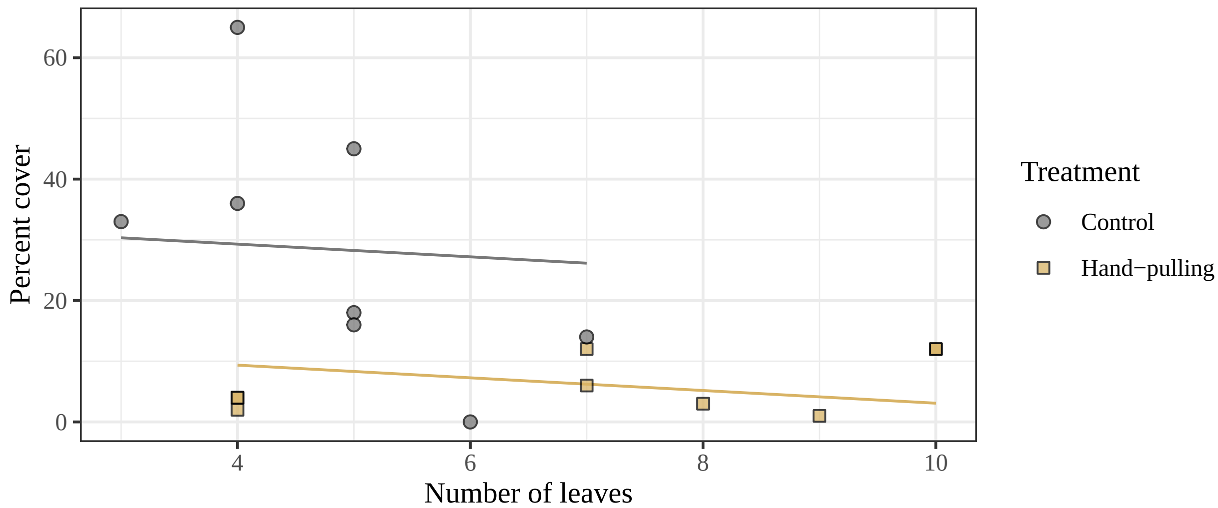
1. Agave size (i.e. the number of leaves)
2. Presence or absence of a live agave

Note in these analyses (Section 2), only treatments W and H *sensu stricto* were considered. That is, data from rows with treatments J+H, J+W, S+H, and S+W are not included in plots & analyses.

2.1 How do selected treatments and agave size affect percent cover?

Here we are interested to know how certain treatments and the *size* of the agave plants affect the percent cover of Lehman lovegrass (N = 17).

Figure 2.1. Percent Lehmann cover by agave size



Note in the plot above, there are some points with identical values, e.g. there are two observations where the number of leaves was 4 and percent cover was 12%.

Ideally we would use a mixed-effect model with linear regression:

$$\%Lehmann\ \ cover = \beta_0 + \beta_1 \times Treatment + \beta_2 \times \#Leaves + b_0$$

Where b_0 is the random intercept for plot. However, when considering *only* the samples from H and Control) treatments, there are too few samples (N = 17) to run a mixed effect model. A good old fashioned linear regression model is then:

$$\%Lehmann\ \ cover = \beta_0 + \beta_1 \times Treatment + \beta_2 \times \#Leaves$$

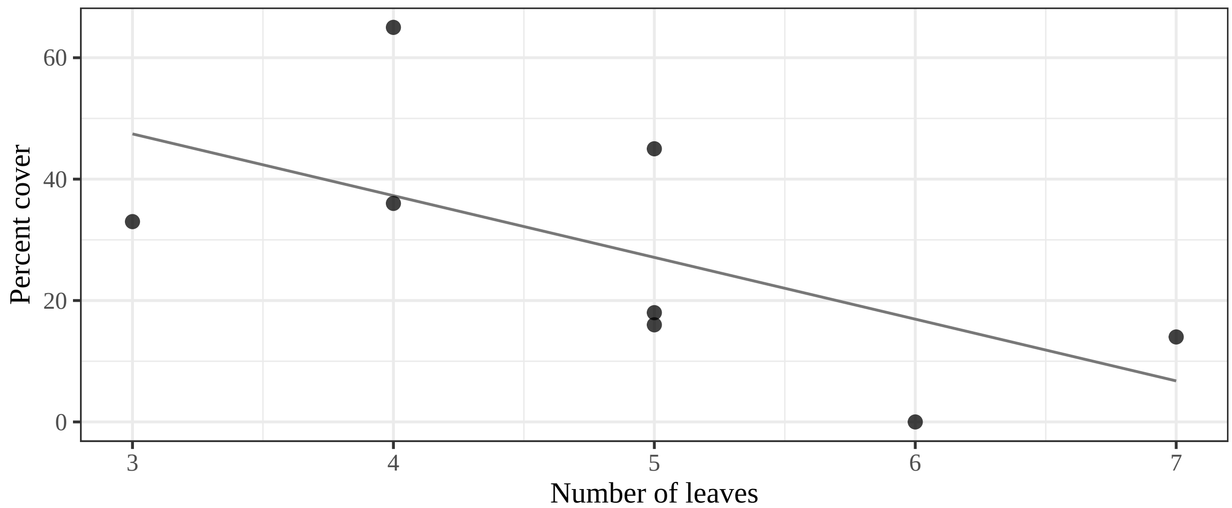
Predictor	Coefficient Estimate	Error	P-value
(Intercept)	33.47023	10.62923	0.00711
Hand-pulling	-19.93178	8.24733	0.02989
Agave size	-1.04517	1.89762	0.59047

Note: Similar to question 1.2, there are several H treatments with zero leaf counts. It may be necessary instead to evaluate a model that does not include the number of agave leaves.

2.1.b How does agave size in control treatments effect percent cover?

This analysis is identical to the one in 2.1, above, but includes only plants from the Control treatment.

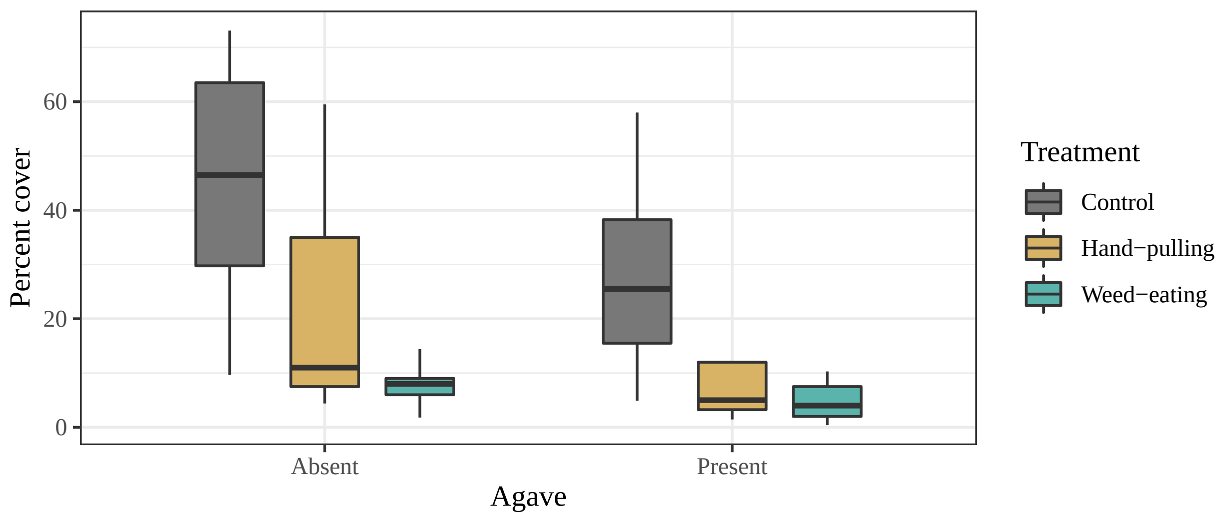
Figure 2.1. Percent Lehmann cover by agave size in Control Treatment



2.2 How do selected treatments and the presence of a live agave affect percent cover?

In contrast to 2.1, we instead use presence or absence of a live agave in the model (N = 65).

Figure 2.2. Percent Lehmann cover by agave presence



Boxplots show median (center line), 25th and 75th percentiles (lower and upper box boundaries, respectively), and 5th and 95th percentiles (lower and upper whisker, respectively)

We use a mixed-effect model:

$$\%Lehmann\ cover = \beta_0 + \beta_1 \times Treatment + \beta_2 \times Agave + b_0$$

where *Agave* is a binary predictor, indicating whether or not a live agave was present and b_0 is the random intercept for plot.

Predictor	Coefficient Estimate	Error	P-value
(Intercept)	37.86291	7.26623	0.00309

Predictor	Coefficient Estimate	Error	P-value
Hand-pulling	-20.26001	3.78197	1.5637e-06
Weed-eating	-23.23516	4.28922	1.2102e-06
Agave presence	-4.20726	3.74574	0.26593

3. How does percent cover affect survival and size of agaves?

3.1 How does percent cover affect agave survival?

Where survival is the total number of agaves alive. This is probably best addressed with an expanded version of the model presented in 1.1.a, above (N = 250).

$$Log\!-\!odds\ \ Survival = \beta_0 + \beta_1 \times Treatment + \beta_2 \times \%Cover + b_0$$

Where b_0 is the random intercept for plot.

Predictor	Coefficient Estimate	Error	P-value
(Intercept)	-0.0520	0.8835	0.953
H	-0.3416	0.7632	0.6545
J	-0.1651	0.7324	0.8217
J+H	-1.4110	0.8239	0.0868
J+S	2.7080	0.8760	0.002
J+W	1.5985	0.7861	0.042
S	1.3969	0.7475	0.0616
S+H	-1.3471	0.8432	0.1101
S+W	-0.9937	0.7022	0.157
W	-2.2401	0.8900	0.0118
% Cover Lehman lovegrass	-0.0232	0.0116	0.0462

3.2 How does percent cover affect agave size?

$$\#Leaves = \beta_0 + \beta_1 \times Treatment + \beta_2 \times \%Cover + b_0$$

Where b_0 is the random intercept for plot. This is a modification of the model presented in 1.2.a, above (N = 74).

Predictor	Coefficient Estimate	Error	P-value
(Intercept)	4.85929	1.26063	3e-04
H	2.10529	1.43626	0.14787
J	1.01393	1.49651	0.50056
J+H	2.18983	1.71069	0.20539
J+S	2.49701	1.19468	0.04069
J+W	2.08840	1.30416	0.11451
S	3.27985	1.21644	0.00903
S+H	0.64363	1.74228	0.71308
S+W	2.73633	2.15505	0.20891
W	2.66538	2.93845	0.36794
% Cover Lehman lovegrass	-0.00364	0.02719	0.89381