Palmer's Agave & Lehmann Lovegrass

Jeff Oliver

March 6, 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:

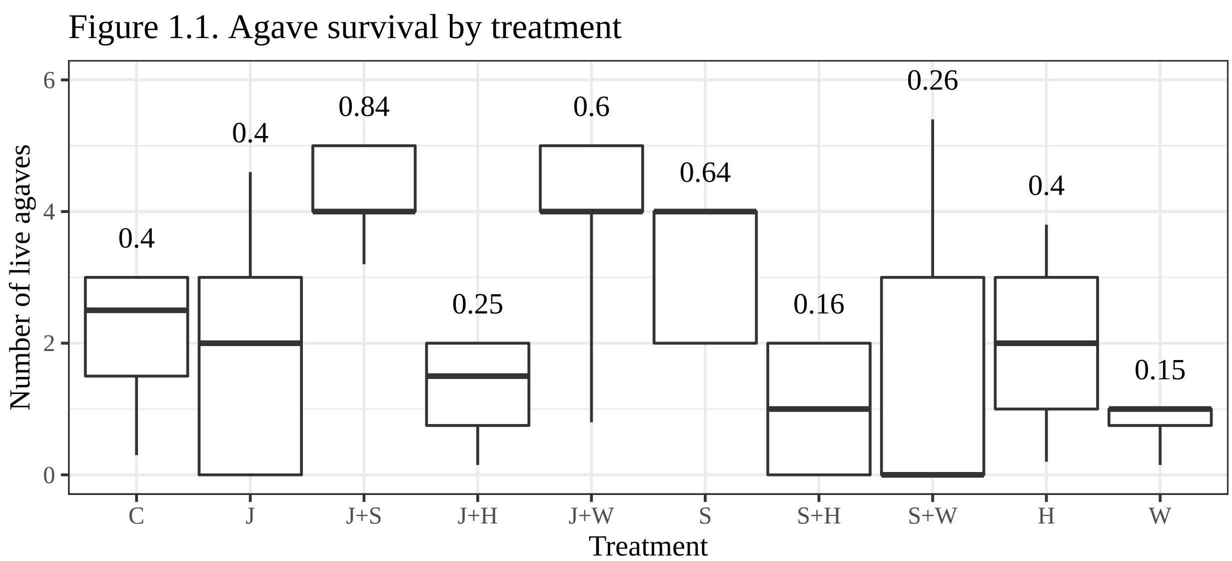
1. The total number of live plants in a row (all those not marked as dead or predated)
2. The number of leaves on agave plants; the number of agaves measured varies among plot/treatment combinations

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).



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)

1.1.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
(Intercept)	-0.81657	0.81719	0.31767
H	0.16441	0.70243	0.81494
J	0.16447	0.70242	0.81487
J+H	-0.85779	0.76587	0.26270
J+S	3.07192	0.88807	0.00054
J+W	1.75895	0.73992	0.01744
S	1.50570	0.74035	0.04197
S+H	-0.86220	0.79210	0.27638
S+W	-0.76281	0.67871	0.26105
W	-1.69443	0.83332	0.04202

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.

Note the explanation above does 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)	-0.63421	0.60853	0.29732
Javalina	1.19261	0.34325	0.00051
Shade	0.67555	0.33686	0.04492
Weed eating	-0.87507	0.35406	0.01345
Hand pulling	-1.13233	0.39069	0.00375

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)	-0.81520	0.81721	0.31850
Javalina	0.16294	0.70245	0.81657
Shade	1.50410	0.74034	0.04219
Weed eating	-1.69616	0.83346	0.04184
Hand pulling	0.16307	0.70244	0.81642
Javalina x Shade	1.40380	1.09110	0.19824
Javalina x Weed eating	3.29089	1.08486	0.00242
Javalina x Hand pulling	-1.18459	1.02498	0.24780
Shade x Weed eating	-0.57167	1.06388	0.59103
Shade x Hand pulling	-2.53050	1.03952	0.01492

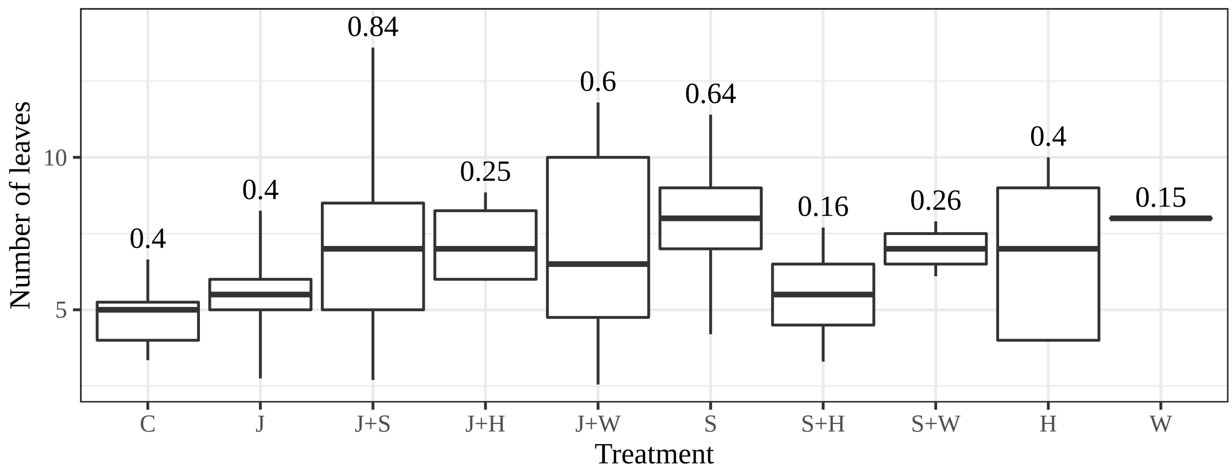
1.2. How does treatment affect agave size?

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.

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.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	0.00003
H	2.18513	1.29757	0.09726
J	1.04089	1.47315	0.48240
J+H	2.25015	1.63663	0.17416
J+S	2.50682	1.18344	0.03807
J+W	2.14523	1.22489	0.08480
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

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.

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	0.00000
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

Note: P-values reported as 0 are artifacts of rounding; they are not truly zero.

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	0.00003
Javalina	1.04089	1.47315	0.48240
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	0.65690
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?

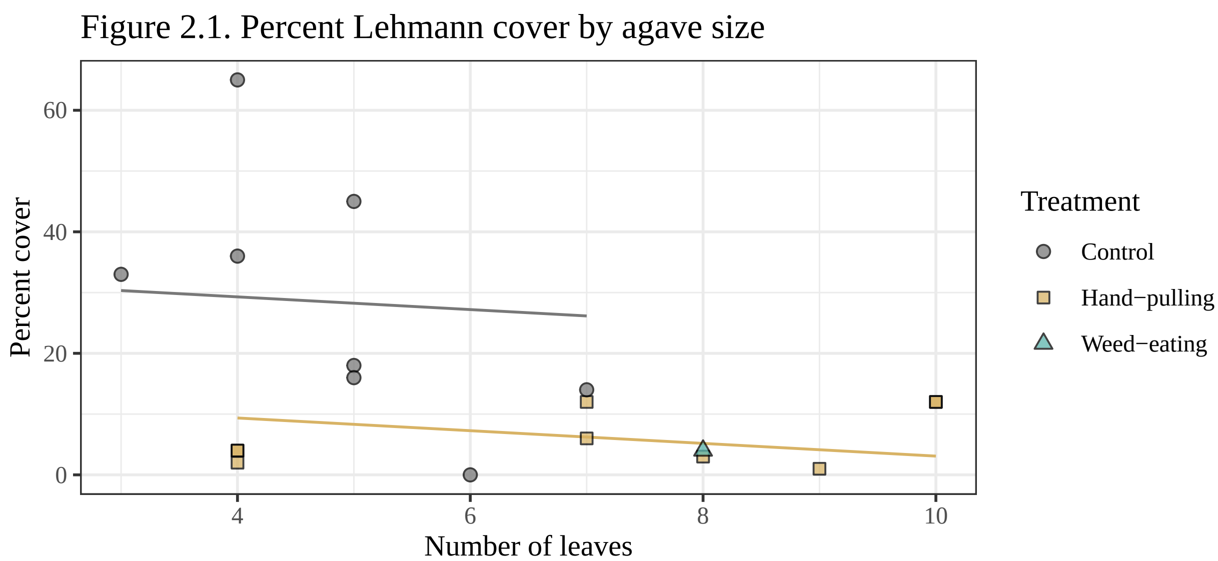
In these analyses, the effect of agaves is modeled in two ways:

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 = 18).



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 W (and control) treatments, there are too few samples (N = 18) to run a mixed effect model. A good old fashioned linear regression model is then:

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

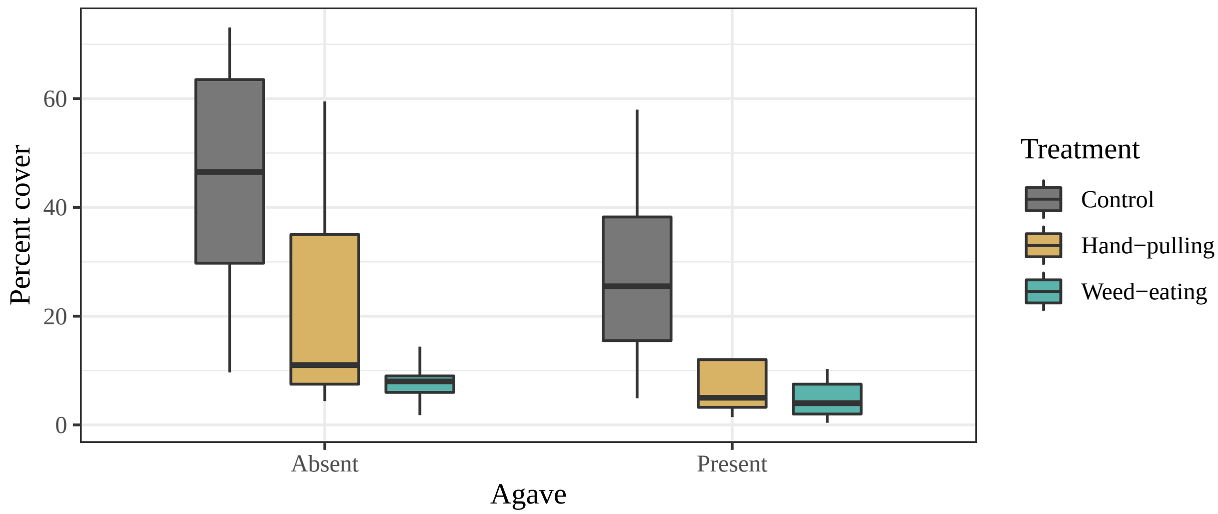
Predictor	Coefficient Estimate	Error	P-value
(Intercept)	33.47023	10.62923	0.00711
Hand-pulling	-19.93178	8.24733	0.02989
Weed-eating	-21.10883	16.78619	0.22914
Agave size	-1.04517	1.89762	0.59047

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

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
Hand-pulling	-20.26001	3.78197	0.00000
Weed-eating	-23.23516	4.28922	0.00000
Agave presence	-4.20726	3.74574	0.26593

Note: P-values reported as 0 are artifacts of rounding; they are not truly zero.

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)	-0.05204	0.88346	0.95303
H	-0.34161	0.76323	0.65445
J	-0.16508	0.73243	0.82168

Predictor	Coefficient Estimate	Error	P-value
J+H	-1.41096	0.82394	0.08681
J+S	2.70797	0.87600	0.00199
J+W	1.59853	0.78615	0.04201
S	1.39688	0.74747	0.06165
S+H	-1.34714	0.84319	0.11012
S+W	-0.99372	0.70215	0.15699
W	-2.24014	0.88998	0.01183
% Cover Lehman lovegrass	-0.02316	0.01162	0.04624

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	0.00030
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