Analysis on the Effects of Aloe and pH Treatments on Fungal Growth

Written Report

Miami University - Department of Statistics
311 Upham Hall
100 Bishop Circle
Oxford, OH 45056

James Dudley
Hanwei Zhao
Bohan Zhao
Mary Yu

STA 475 | Spring 2018
May 10th, 2018

Contents

1. Abstract

A statistical data analysis was performed on a dataset consisting of fungal interactions with varying levels of pH and seven different species of aloe. The purpose of the experiment was to find how these different pH levels and aloe species affected the germination of a specific fungus, *Candida albicans*. The research hypotheses asked more specifically if the germination decreased at low pH levels, increased at higher pH levels, and if one specific aloe species, *Aloe arborescens*, significantly changed the germination at specific pH levels. It was ultimately determined that both the presence of pH buffers and the presence of certain aloe species both have a statistically significant impact on the germination of *Candida albicans*.

2. Introduction

The dataset at hand contains the results of various experiments run to ultimately determine the effects of aloe and pH on fungal morphogenesis. The experiments were overseen by Dr. Marcia Lee of the Miami University Department of Microbiology and Dr. Richard Bretz of the Miami University Department of Chemistry.

Altogether, seven different species of aloe were analyzed in this experiment: *Aloe arborescens* (gold tooth aloe), *Aloe barbadensis*, *Aloe cameronii*, *Aloe distans*, *Aloe ferox*, *Aloe striata*, and *Gasteria verrucosa*. The aloe species were obtained from the Krohn Conservatory in Cincinnati, Ohio, and were then housed and propagated in the Belk Greenhouse on Miami University's campus under fungicide-free conditions. These aloe experiments were conducted to find their effects on the growth of *Candida albicans*, a yeast that is commonly found on the human body. Experiments with each of the aloe types were conducted one aloe at a time, and the level of three different response variables to fungal interactions were recorded. These variables are as follows:

1. the formation of germ tubes
2. the formation of pseudohyphae
3. no germination at all (no growth)

These three response variables were measured with other effects as well, including varied, controlled pH levels of 5.0, 5.5, 6.0, 6.5, 7.0 and an unbuffered "wild" pH level. Each pH level is observed with its own control. For each aloe, 16 repetitions in individual test tubes were performed at each of the 6 pH levels: eight repetitions with that include the aloe at the given pH, and eight "control" repetitions that do not include the aloe at the given pH. Thus, each aloe plant had eight tests at 6 different pH levels for a total of 96 tests. The total for all seven aloe vera plant experiments was 672 tests (96 x 7).

The specific research hypotheses addressed in this analysis are as follows:

1. The pH of the incubation medium influences the germination frequency of *C. albicans*. More specifically, germination frequency is decreased with the presence of pH buffers of 5.0 and 5.5, stagnant for a pH buffer of 6.0, and increased at a pH of 6.5 and 7.0.
2. The incubation of *C. albicans* with *A. arborescens* filtrate in the presence of buffers with a pH of 5.0 or 5.5 does not significantly change germination frequency. However,

incubation of *C. albicans* with *A. arborescens* filtrate in the presence of buffers that have pHs of 6.0, 6.5, and 7.0 significantly diminishes the impact of *A. arborescens* alone on germination frequency.

These two research hypotheses are additionally referred to as tasks 1 and 2, respectively.

3. Methods

For task 1, the effect of pH among the controls is performed separately for each aloe experiment using a beta regression model of the form

$$log\left(\frac{p}{1-p}\right) = \beta_0 + \beta_i(pH_i) + \varepsilon_{ij}$$

where

p = Germination proportion
pH_i = ith level of pH(5.0, 5.5, etc.)
β_0 = Mean value of log odds of germination proportion
β_i = Effect of ith level of pH on log odds of germination proportion
ε_{ij} = Random error component

The significance of the pH effect will be tested using a likelihood ratio Chi-Square test and any necessary post-hoc multiple comparisons were performed using a Dunnett adjustment of the significance level, letting a pH of 6.0 serve as the benchmark level for comparison.

For task 2, focusing on *Aloe arborescens*, the effect of pH and aloe on germination frequency will be investigated using a beta regression with two categorical factors: aloe condition (present/absent) and pH level (5.0, 5.5, 6.0, 6.5, 7.0, wild). The fitted model form is

$$log\left(\frac{p}{1-p}\right) = \beta_0 + \beta_i(pH_i) + \gamma_j(Aloe) + \delta_{ij}(pH \times Aloe) + \varepsilon_{ijk}$$

where

p = Germination proportion
pH_i= ith level of pH(5.0, 5.5, etc.)
β_0 = Mean value of log odds of germination proportion
β_i= Effect of ith level of pH on log odds of germination proportion
γ_j = jth Aloe level (presents/absents)
δ_{ij} = Effect of aloe at the ith level of pH on the log odds of germination proportion
ε_{ijk} = Random error component.

The significance of the aloe/ph interaction will be tested first using a likelihood ratio Chi-Square test to assess if the effect of *Aloe arborescens* is consistent, regardless of pH level. Post-hoc contrasts were conducted where necessary. All significance tests were performed at the 0.05 significance level. All analyses were performed using R version 3.3.1 and packages betareg, emmeans, ggplot2, and lmtest.

3.1 Beta Regression

The beta regression model is a tool for modeling continuous variables which assumes all values are inside the open standard unit interval (0,1). As a result, it is an appropriate method for modeling responses that are proportional in nature. The model is based on an alternative parameterization of the beta density in terms of the variate mean and a precision parameter. In the present context, the beta regressions are performed using the logit link function.

3.2 Likelihood Ratio Chi-square Test

The likelihood ratio test is a statistical test that aims to compare the fitting performance of two statistical models. These two models are based on the statistical hypothesis test that one is the null model, which is a special case of the other model, the alternative model. Model-based likelihoods are compared to assess the impact of imposing the statistical hypothesis on the alternative model. The statistic, which is calculated by the likelihood ratio and distributed like chi-square, expresses how many times more likely the data are under one model than the other following the chi-square distribution.

3.3 Estimated Marginal Means

Estimated marginal means (EMMs), also known as least-squares means, can be calculated for many linear, generalized linear, and mixed models. In contrast, a raw or an observed mean is a simple average of the values from original data, which is used without a model. Estimated marginal means are adjusted for other terms in the assumed model, like covariates, and are low probability for missing data. Theoretically, estimated marginal means are better than observed means to estimate the true population mean.

4. Results

4.1 Task 1

Aloe arborescens

After performing the likelihood ratio test, a significant difference was observed between the model with pH as a response variable and the null model containing only the intercept. This indicates that there is a significant pH effect within the *Aloe arborescens* controls (likelihood ratio chi-square (5) = 168.41, p-value < 0.05). The pseudo R-squared value from the model with pH as a response is 0.943, indicating a good fit with a 94.3% accuracy in predicting the variability of the germination proportion. *Figure 1* and *Table 1.1* show the estimated means, standard errors and confidence intervals for the germination proportion of *C. albican*s with respect to the pH levels for *Aloe arborescens*.

The comparison between a pH level of 6.0 and other pH levels can be observed from *Table 1.2*. A positive difference in the estimated mean germination proportion indicates that the mean germination proportion of that specific pH level is smaller than the estimated mean germination proportion at a pH level of 6.0. The p-value indicates the significance of this result. If the p-value is less than 0.05, it is safe to conclude that the difference between the two estimated mean germination proportions is significant. Looking back at the table, it can be observed that buffers with pH levels 5.0 and 5.5 have decreases in the germination proportion, and these decreases are significant since the p-values are less than 0.05. For buffers with pH

level of 6.5, a slight increase in the germination proportion (5.1%) is recorded, however the p-value is not small enough to conclude that this difference is significant. For buffers with pH level 7.0 and uncontrolled pH levels, there is a significant increase in the germination proportions.
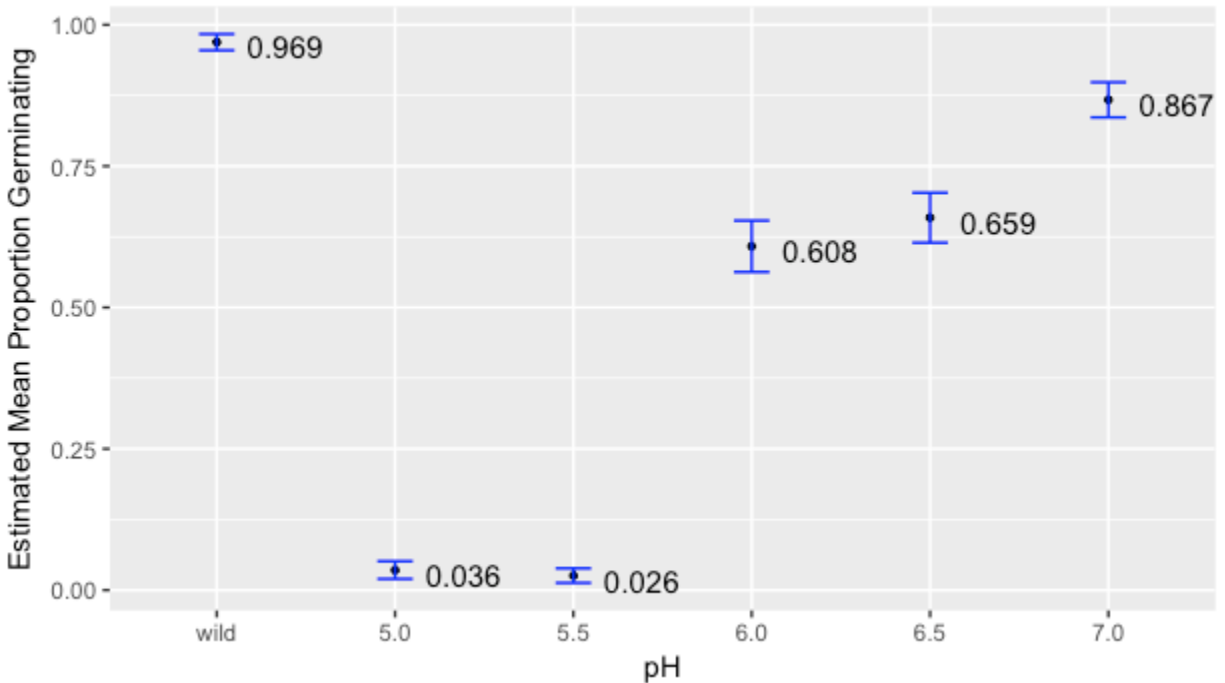

Figure 1: 95% Confident Interval for Mean Proportion Germinating vs. pH for *A.arborescens*

Table 1.1: Germination Proportion of *C. albicans* for *A.arborescen*s Absent

pH	Estimated Mean Proportion Germinating	Standard Error	95% Confidence Interval
wild	0.969	0.007	[0.955, 0.984]
5.0	0.036	0.008	[0.020, 0.051]
5.5	0.026	0.007	[0.013, 0.038]
6.0	0.608	0.023	[0.563, 0.654]
6.5	0.659	0.023	[0.615, 0.703]
7.0	0.867	0.016	[0.955, 0.984]

Table 1.2: Difference in Germination Proportion of *C. albicans* for *A.arborescens* Absent

Contrast	Difference in Estimated Mean Proportion Germinating	Standard Error	Adjusted P-value
6.0 - 5.0	0.573	0.025	< 0.0005***
6.0 - 5.5	0.583	0.024	< 0.0005***
6.0 - 6.5	-0.051	0.032	1
6.0 - 7.0	-0.259	0.028	< 0.0005***
6.0 - wild	-0.361	0.024	< 0.0005***

Other Aloes

 The likelihood ratio Chi-square test result for *Aloe cameronii* (pseudo r-square value: 0.966), *Aloe ciliaris* (pseudo r-square value: 0.969), *Aloe distans* (pseudo r-square value: 0.974), *Aloe ferox* (pseudo r-square value: 0.972), *Aloe squarrosa* (pseudo r-square: 0.971), and *Gasteria verrucosa* (pseudo r-square: 0.962) are all significant. Individual tests for the comparison between pH level of 6.0 and other pH levels has shown consistency across these aloe species.

 The pH results that were observed for *Aloe arborescens* were found to be consistent in all of the 6 aloes listed above, with the exception of the difference between pH levels 6.0 and 6.5. It was found that from *Aloe arborescens* the difference between the germination proportion at pH level 6.0 and 6.5 was insignificant, however for the other aloe species, the germination proportion for pH level 6.5 significantly increased. For specific figures and test results, please refer to *Tables 2.1-7.2, Figures 2-7* in the appendix.

4.2 Task 2

 After performing the likelihood ratio chi-squared test comparing a null model with main effects only (i.e., no interaction) to the fitted model that involves the interaction between pH and the presence of *Aloe arborescens*, the test concluded that the interaction is significant (LR Chi-Square = 187.37, df = 5, p-value < 0.0001), thus there is statistical evidence that the effect of aloe on germination frequency changes with pH level. *Table 8.1* and *Figure 8* show the estimated mean germination proportion of *C. albicans* for *Aloe arborescens*, present and absent. A response of "Y" indicates that *Aloe arborescens* is present in the buffer. The significance of differences due to the presence of *Aloe arborescens* can be observed from *Table 8.2*. For each row of *Table 8.2*, the difference in estimated mean proportion for germination was found by subtracting the mean of the samples with the *Aloe arborescens* filtrate from the mean of the samples without the *Aloe arborescens* filtrate for each pH level.

 For the group with wild, uncontrolled pH, there is a decrease of 80.5% germination between the proportions before and after the aloe was added to the buffer, and the standard error of the difference is 0.024. The p-value for this result is less than 0.05, therefore this difference is significant. For the pH level of 5.0, a small increase could be observed on the germination proportions, however the p-value is not small enough to testify the significance of this difference. For the pH level of 5.5, the decrease for germination proportions also increased after the aloe

was added into the buffer, and the p-value (0.0022 < 0.05) shows that this result is significant. The pH levels of 6.0 and 6.5 both showed increases in the estimated mean germination after the addition of *Aloe arborescens*, and both p-values are small, so it is safe to conclude that the differences are significant. Finally, for a pH level of 7.0, a small decrease could be observed in the germination proportions after adding the aloe to the buffer, however the p-value is not small enough to support the significance of this result.
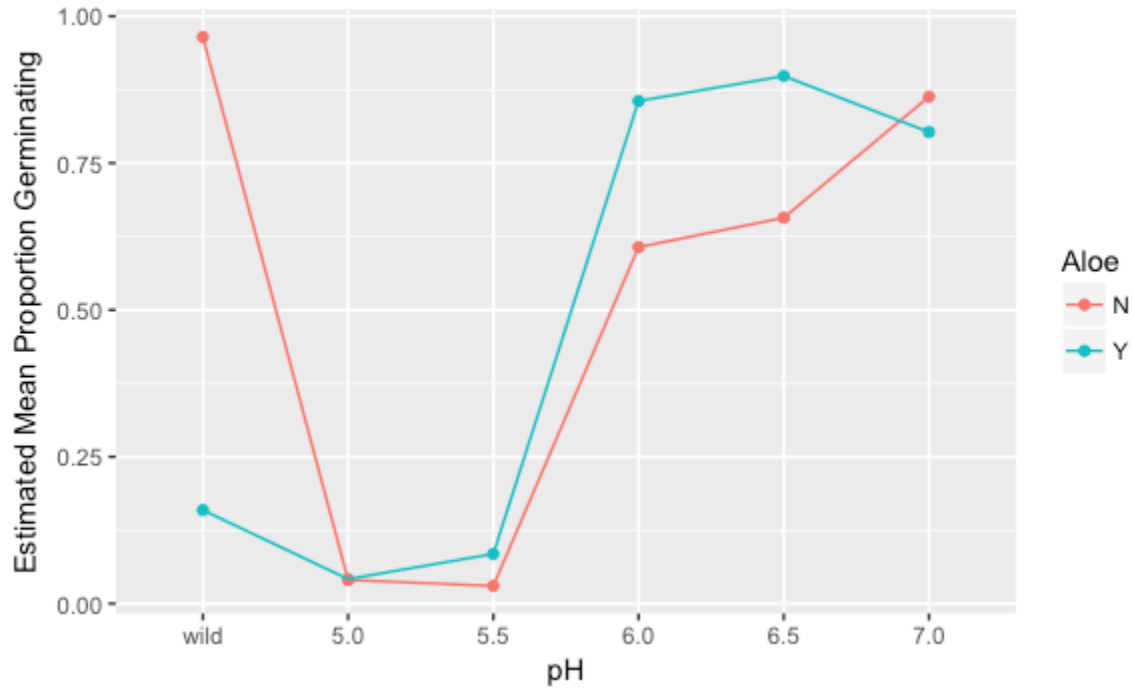

Figure 2: Interaction for *A.arborescens* Present and Absent in pH

Table 2.1: Germination Proportion of *C. albicans* for *A. arborescens,* Present and Absent

pH	Aloe arborescens (N / Y)	Estimated Mean Proportion Germinating	Standard Error	95% Confidence interval
wild	N	0.965	0.009	[0.946, 0.983]
	Y	0.159	0.022	[0.117, 0.202]
5.0	N	0.040	0.010	[0.020, 0.060]
	Y	0.042	0.010	[0.021, 0.062]
5.5	N	0.030	0.008	[0.014, 0.047]
	Y	0.085	0.016	[0.054, 0.116]
6.0	N	0.607	0.030	[0.549, 0.664]
	Y	0.856	0.021	[0.815, 0.896]
6.5	N	0.657	0.029	[0.601, 0.713]
	Y	0.898	0.017	[0.864, 0.932]
7.0	N	0.863	0.020	[0.824, 0.903]
	Y	0.803	0.024	[0.756, 0.849]

Table 2.2: Difference in Germination Proportion of *C. albicans* for *A. arborescens,* Present and Absent

pH	Estimated Aloe Induced Reduction in Germination Proportion	Standard Error	Adjusted P-value
wild	0.805	0.024	< 0.0001***
5.0	-0.001	0.014	0.9177
5.5	-0.054	0.018	0.0022***
6.0	-0.249	0.036	< 0.0001***
6.5	-0.241	0.033	< 0.0001***
7.0	0.060	0.031	0.0522

5. General Discussion

From the results of the analysis, one factor that gave interesting results was the pH level wild (uncontrolled). *Figure 2* is the graph that shows the differences of the germination proportion with *Aloe arborescens*, present and absent. Also, looking back at *Figures 4.1, 5.1, 6.1, 7.1, 8.1* and *9.1* in the appendix, it could be seen that the germination proportion for the uncontrolled pH level was very high. According to Dr. Lee, the uncontrolled pH group probably has pH levels that are very low, expectedly below 5.0. However, looking at the trends in *Figure 2*, the trend will make more sense if the uncontrolled pH group was placed after pH level 7.0, in that case, the uncontrolled pH group will have pH levels greater than 7.0. Therefore, recording the pH level for the uncontrolled pH group could be beneficial to the study.

6. References

Achim Zeileis, Torsten Hothorn (2002). Diagnostic Checking in Regression Relationships. R News 2(3), 7-10. URL https://CRAN.R-project.org/doc/Rnews/

Francisco Cribari-Neto, Achim Zeileis (2010). Beta Regression in R. Journal of Statistical Software 34(2), 1-24. URL http://www.jstatsoft.org/v34/i02/.

Gregory R. Warnes, Ben Bolker, Lodewijk Bonebakker, Robert Gentleman, Wolfgang Huber Andy Liaw, Thomas Lumley, Martin Maechler, Arni Magnusson, Steffen Moeller, Marc Schwartz and Bill Venables (2016). gplots: Various R Programming Tools for Plotting Data. R package version 3.0.1. https://CRAN.R-project.org/package=gplots

H. Wickham. ggplot2: Elegant Graphics for Data Analysis. Springer-Verlag New York, 2009.

Lenth, R. (2018, March 31). Interaction analysis in emmeans. Retrieved from https://cran.r-project.org/web/packages/emmeans/vignettes/interactions.html

Likelihood ratio chi-square test - Oxford Reference. (2016, December 02). Retrieved from http://www.oxfordreference.com/view/10.1093/acref/9780199534067.001.0001/acref-9780199534067-e-4654

Millard SP (2013). _EnvStats: An R Package for Environmental Statistics_. Springer, New York. ISBN 978-1-4614-8455-4, <URL: http://www.springer.com>. (n.d.). Retrieved from http://dawg.utk.edu/glossary/g_least_squares_means.htm

Russell Lenth (2018). emmeans: Estimated Marginal Means, aka Least-Squares Means. R package version 1.1.2. https://CRAN.R-project.org/package=emmeans

7. Appendix

***Aloe arborescens* absent**

Table 3: Beta Regression Coefficients

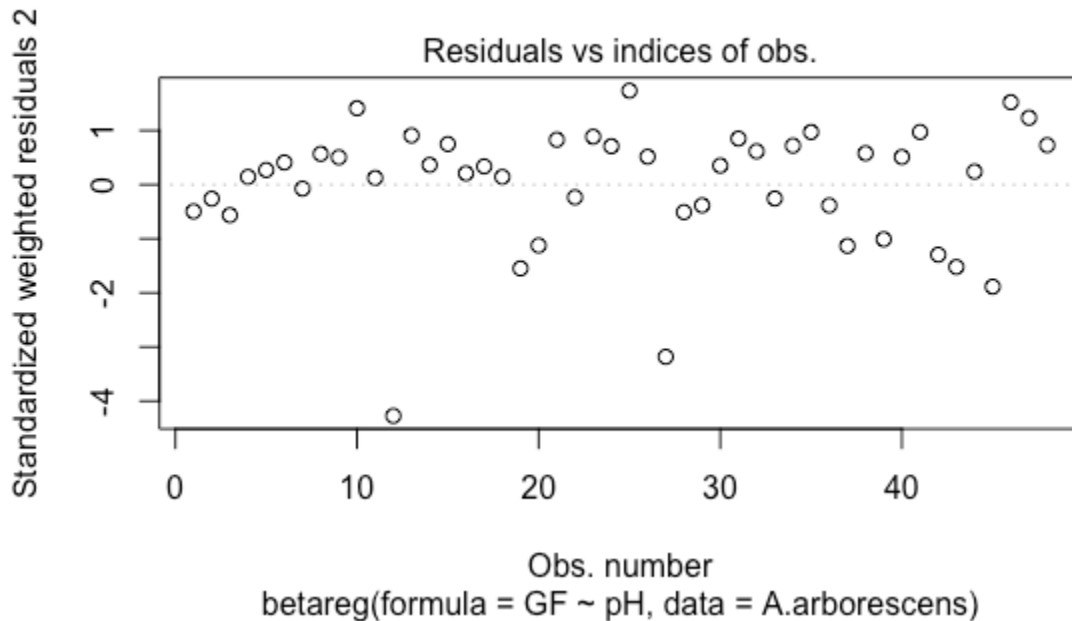|	**Estimate**	**Std. Error**	**z value**	**Pr (>|z|)**
(Intercept)	0.440	0.097	4.512	6.43e-06 ***
pH 5.0	-3.740	0.253	-14.782	< 2e-16 ***
pH 5.5	-4.077	0.280	-14.565	< 2e-16 ***
pH 6.5	0.218	0.140	1.560	0.119
pH 7.0	1.437	0.169	8.522	< 2e-16 ***
pH wild	3.001	0.264	11.399	< 2e-16 ***

Pseudo R-squared: 0.9426

Figure 3: Residuals vs. Observations of A.arborescens Absent
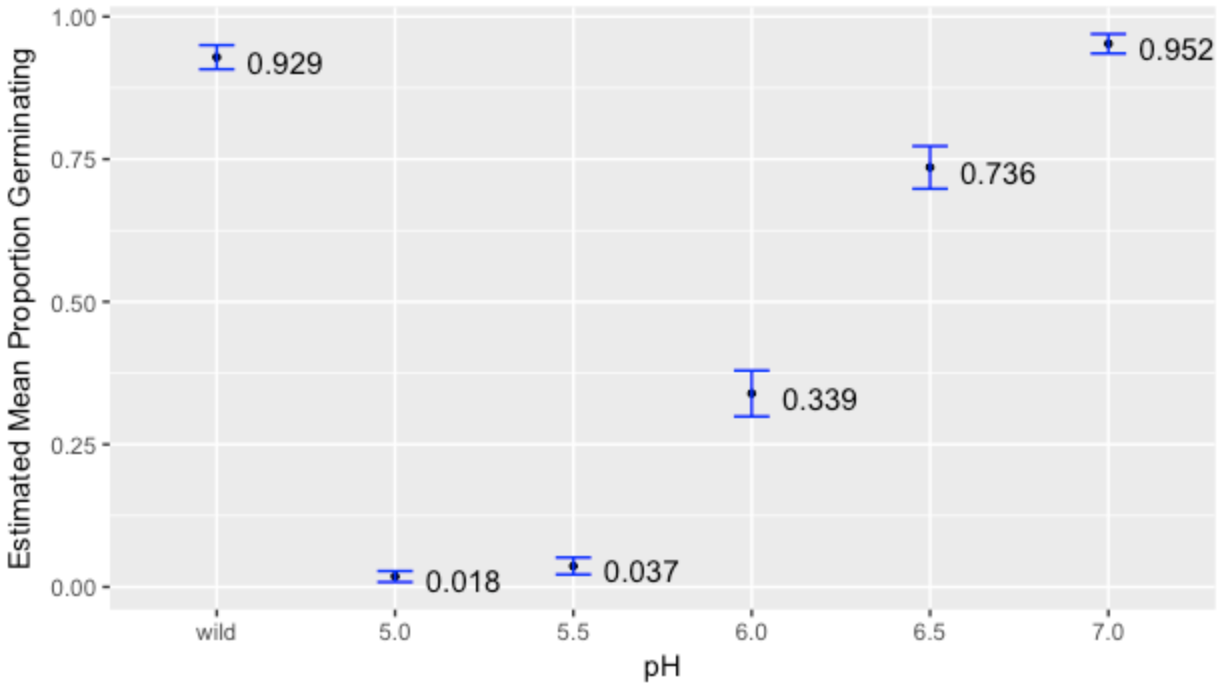
A.cameronii **absent**

Figure 4.1: Estimated Mean Proportion of Germination vs. pH for *A.cameronii*

Table 4.1: Germination Proportion of C. albicans for *A.cameronii* Absent

pH	Estimated Mean Proportion Germinating	Standard Error	95% Confidence Interval
wild	0.929	0.011	[0.908, 0.950]
5.0	0.018	0.005	[0.009, 0.028]
5.5	0.037	0.008	[0.022, 0.052]
6.0	0.339	0.021	[0.299, 0.379]
6.5	0.736	0.019	[0.698, 0.773]
7.0	0.952	0.011	[0.908, 0.950]

Table 4.2: Difference in Germination Proportion of C. albicans for *A.cameronii* Absent

Contrast	Difference in Estimated Mean Proportion Germinating	Standard Error	Adjusted P-value
6.0 - 5.0	0.321	0.021	< 0.0005***
6.0 - 5.5	0.302	0.022	< 0.0005***
6.0 - 6.5	-0.396	0.028	< 0.0005***
6.0 - 7.0	-0.613	0.022	< 0.0005***
6.0 - wild	-0.590	0.023	< 0.0005***

Table 4.3: Beta Regression Coefficients

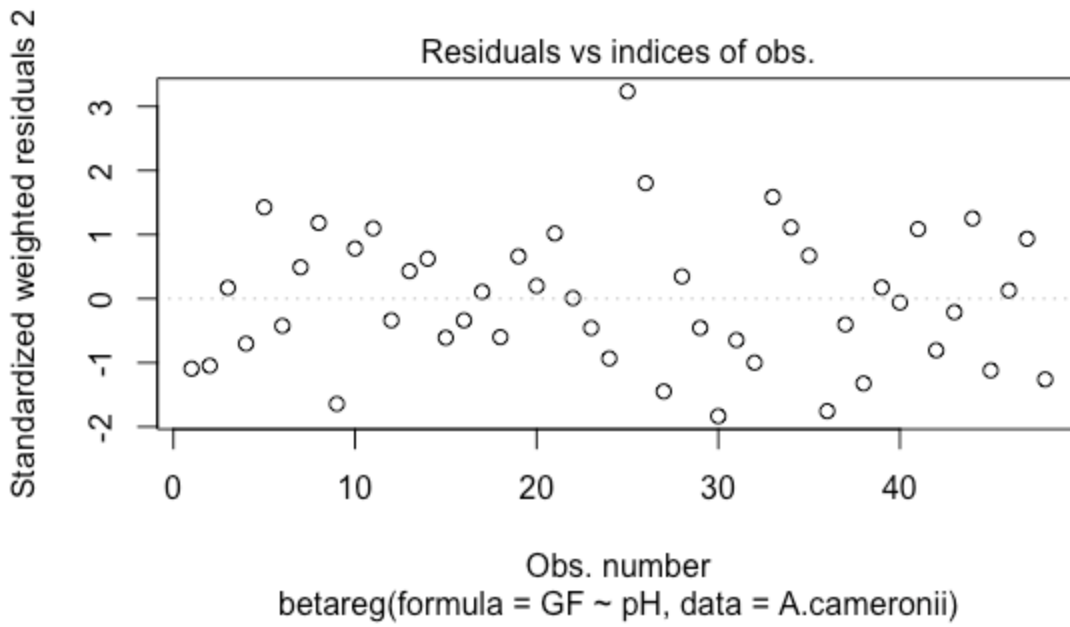|	Estimate	Std. Error	z value	Pr (>|z|)
(Intercept)	-0.667	0.092	-7.284	3.24e-13 ***
pH 5.0	-3.322	0.290	-11.475	< 2e-16 ***
pH 5.5	-2.600	0.232	-11.212	< 2e-16 ***
pH 6.5	1.690	0.134	12.587	< 2e-16 ***
pH 7.0	3.662	0.213	17.166	< 2e-16 ***
pH wild	3.235	0.187	17.289	< 2e-16 ***

Pseudo R-squared: 0.9662

Figure 4.2: Residuals vs. Observations of A.cameronii Absent
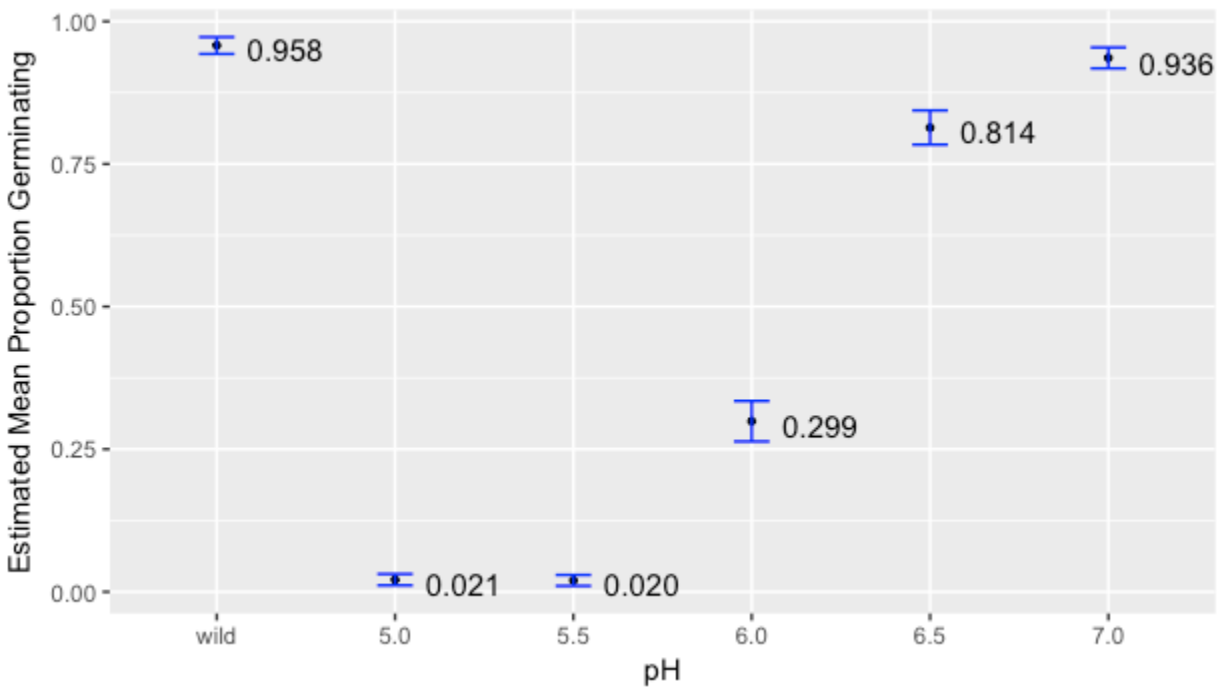
A.ciliaris absent

Figure 5.1 Estimated Mean Proportion of Germination vs. pH for *A.ciliaris*

Table 5.1: Germination Proportion of *C. albicans* for *A.ciliaris* Absent

pH	Estimated Mean Proportion Germinating	Standard Error	95% Confidence Interval
wild	0.958	0.008	[0.943, 0.972]
5.0	0.021	0.005	[0.011, 0.031]
5.5	0.020	0.005	[0.010, 0.030]
6.0	0.229	0.018	[0.264, 0.334]
6.5	0.814	0.015	[0.784, 0.844]
7.0	0.936	0.009	[0.917, 0.954]

Table 5.2: Difference in Germination Proportion of *C. albicans* for *A.ciliaris* Absent

Contrast	Difference in Estimated Mean Proportion Germinating	Standard Error	Adjusted P-value
6.0 - 5.0	0.278	0.019	< 0.0005***
6.0 - 5.5	0.279	0.019	< 0.0005***
6.0 - 6.5	-0.515	0.024	< 0.0005***
6.0 - 7.0	-0.637	0.020	< 0.0005***
6.0 - wild	-0.659	0.020	< 0.0005***

Table 5.3: Beta Regression Coefficients

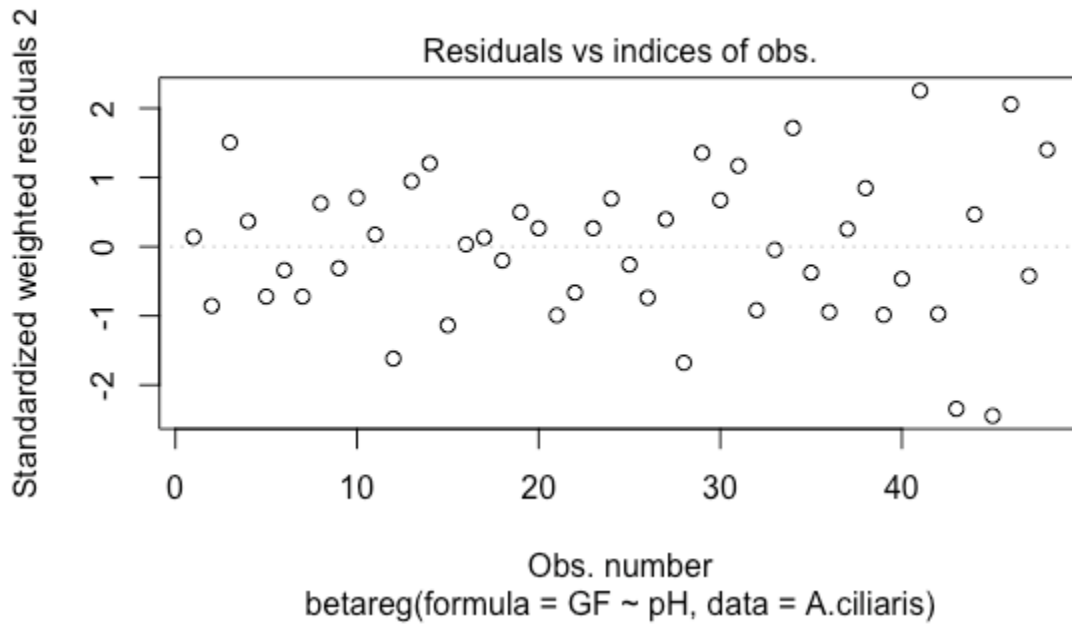|	**Estimate**	**Std. Error**	**z value**	**Pr (>|z|)**
(Intercept)	-0.852	0.086	-9.872	< 2e-16 ***
pH 5.0	-2.975	0.258	-11.527	< 2e-16 ***
pH 5.5	-3.039	0.263	-11.539	< 2e-16 ***
pH 6.5	2.327	0.133	17.501	< 2e-16 ***
pH 7.0	3.533	0.179	19.759	< 2e-16 ***
pH wild	3.971	0.205	19.325	< 2e-16 ***

Pseudo R-squared: 0.943

Figure 5.2: Residuals vs. Observations of A.ciliaris Absent
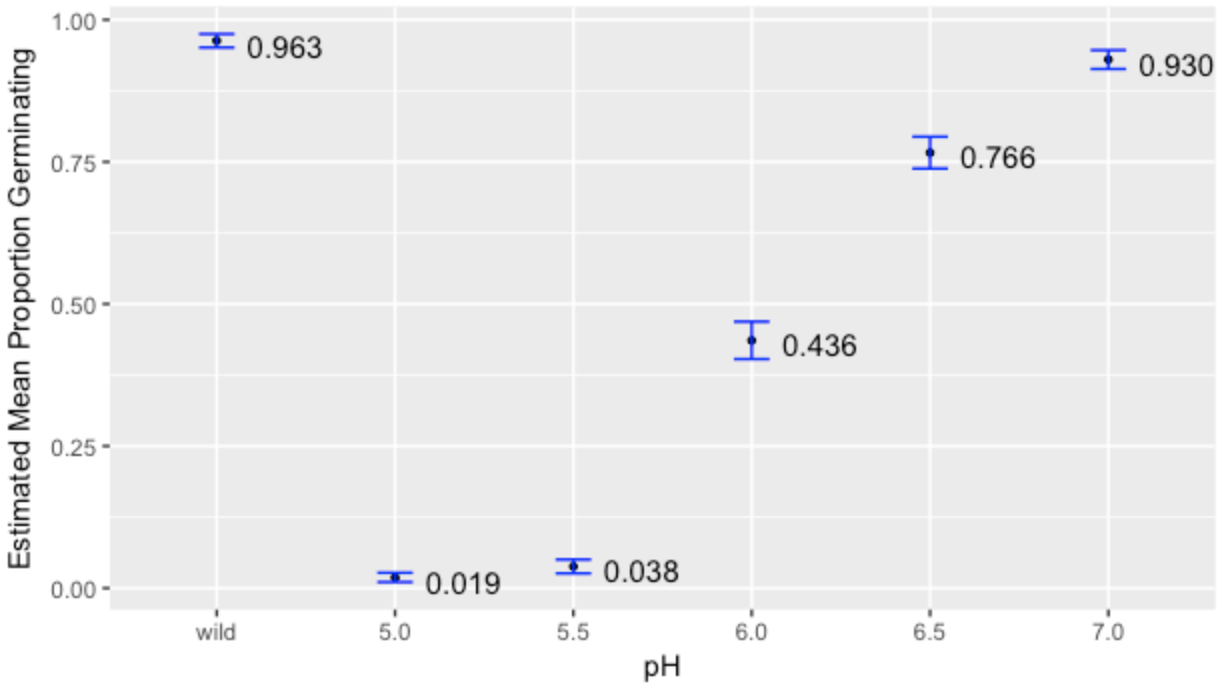
***A.distans* absent**

Figure 6.1 Estimated Mean Proportion of Germination vs. pH for *A. distans*

Table 6.1: Germination Proportion of *C. albicans* for *A. distans* Absent

pH	Estimated Mean Proportion Germinating	Standard Error	95% Confidence Interval
wild	0.963	0.006	[0.952, 0.975]
5.0	0.019	0.004	[0.011, 0.027]
5.5	0.038	0.006	[0.026, 0.050]
6.0	0.436	0.017	[0.403, 0.369]
6.5	0.766	0.014	[0.738, 0.794]
7.0	0.930	0.008	[0.913, 0.947]

Table 6.2: Difference in Germination Proportion of *C. albicans* for *Aloe distans* Absent

Contrast	Difference in Estimated Mean Proportion Germinating	Standard Error	Adjusted P-value
6.0 - 5.0	0.417	0.017	< 0.0005***
6.0 - 5.5	0.398	0.018	< 0.0005***
6.0 - 6.5	-0.330	0.022	< 0.0005***
6.0 - 7.0	-0.494	0.019	< 0.0005***
6.0 - wild	-0.527	0.018	< 0.0005***

Table 6.3: Beta Regression Coefficients

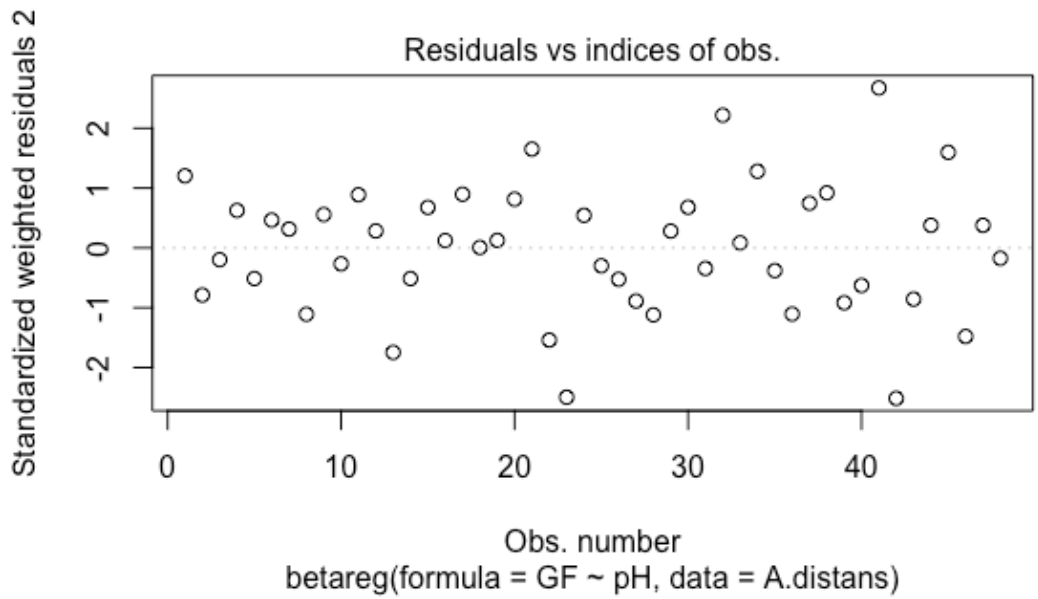|	Estimate	Std. Error	z value	Pr (>|z|)
(Intercept)	-0.258	0.068	-3.774	0.000161 ***
pH 5.0	-3.693	0.236	-15.677	< 2e-16 ***
pH 5.5	-2.971	0.182	-16.309	< 2e-16 ***
pH 6.5	1.445	0.105	13.767	< 2e-16 ***
pH 7.0	2.845	0.147	19.392	< 2e-16 ***
pH wild	3.520	0.185	19.079	< 2e-16 ***

Pseudo R-squared: 0.9743

Figure 6.2: Residuals vs. Observations of A.distans Absent
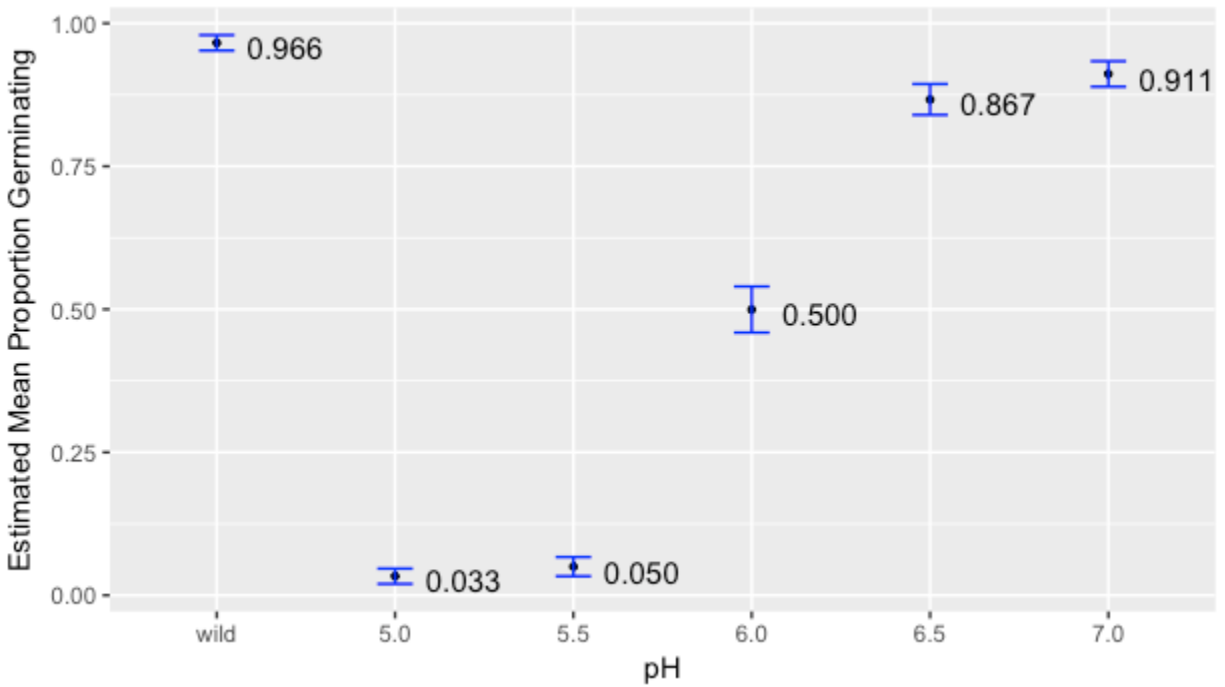
A.ferox absent

Figure 7.1: Estimated Mean Proportion of Germination vs. pH for A.ferox

Table 7.1: Germination Proportion of *C. albicans* for *A.ferox* Absent

pH	Estimated Mean Proportion Germinating	Standard Error	95% Confidence Interval
wild	0.966	0.007	[0.952, 0.979]
5.0	0.033	0.007	[0.020, 0.047]
5.5	0.050	0.009	[0.033, 0.067]
6.0	0.500	0.021	[0.459, 0.540]
6.5	0.867	0.014	[0.840, 0.894]
7.0	0.911	0.011	[0.889, 0.934]

Table 7.2: Difference in Germination Proportion of *C. albicans* for *A.ferox* Absent

Contrast	Difference in Estimated Mean Proportion Germinating	Standard Error	Adjusted P-value
6.0 - 5.0	0.466	0.022	< 0.0005***
6.0 - 5.5	0.449	0.022	< 0.0005***
6.0 - 6.5	-0.367	0.025	< 0.0005***
6.0 - 7.0	-0.412	0.024	< 0.0005***
6.0 - wild	-0.466	0.022	< 0.0005***

Table 7.3: Beta Regression Coefficients

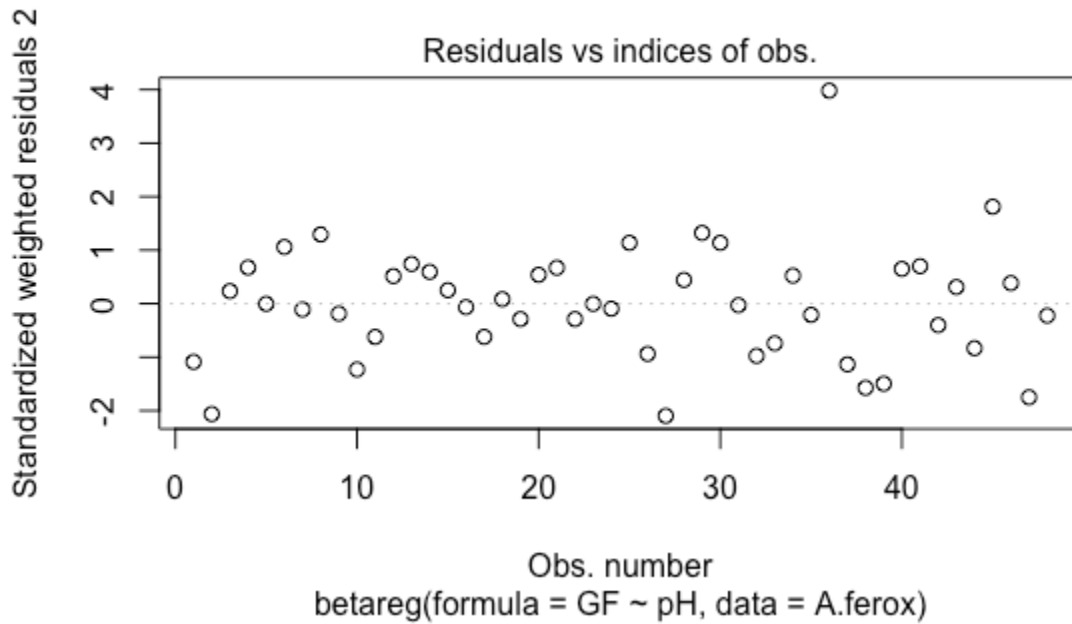|	Estimate	Std. Error	z value	Pr (>|z|)
(Intercept)	-0.001	0.082	-0.02	0.984
pH 5.0	-3.365	0.227	-14.81	<2e-16 ***
pH 5.5	-2.940	0.197	-14.90	<2e-16 ***
pH 6.5	1.875	0.145	12.93	<2e-16 ***
pH 7.0	2.333	0.163	14.27	<2e-16 ***
pH wild	3.344	0.225	14.84	<2e-16 ***

Pseudo R-squared: 0.9722

Figure 7.2: Residuals vs. Observations of A.ferox Absent
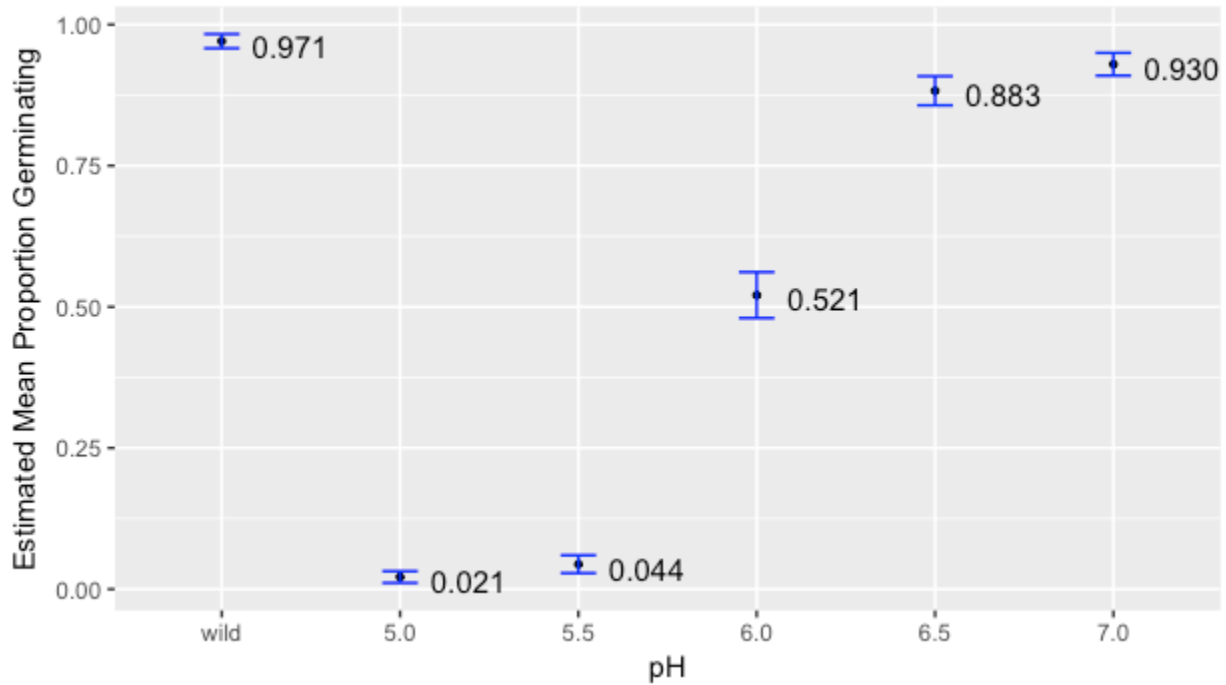
A.squarrosa absent

Figure 8.1: Estimated Mean Proportion of Germination vs. pH for A.squarrosa

Table 8.1: Germination Proportion of *C. albicans* for *A.squarrosa* Absent

pH	Estimated Mean Proportion Germinating	Standard Error	95% Confidence Interval
wild	0.971	0.006	[0.958, 0.983]
5.0	0.021	0.005	[0.011, 0.032]
5.5	0.044	0.008	[0.028, 0.060]
6.0	0.521	0.021	[0.480, 0.561]
6.5	0.883	0.013	[0.857, 0.909]
7.0	0.930	0.010	[0.909, 0.950]

Table 8.2: Difference in Germination Proportion of *C. albicans* for *A.squarrosa* Absent

Contrast	Difference in Estimated Mean Proportion Germinating	Standard Error	Adjusted P-value
6.0 - 5.0	0.499	0.021	< 0.0005***
6.0 - 5.5	0.477	0.022	< 0.0005***
6.0 - 6.5	-0.362	0.025	< 0.0005***
6.0 - 7.0	-0.409	0.023	< 0.0005***
6.0 - wild	-0.450	0.022	< 0.0005***

Table 8.3: Beta Regression Coefficients

	Estimate	Std. Error	z value	Pr (>|z|)
(Intercept)	0.083	0.083	0.997	0.319
pH 5.0	-3.904	0.266	-14.680	<2e-16 ***
pH 5.5	-3.156	0.208	-15.150	<2e-16 ***
pH 6.5	1.936	0.152	12.735	<2e-16 ***
pH 7.0	2.498	0.178	14.035	<2e-16 ***
pH wild	3.413	0.239	14.252	<2e-16 ***

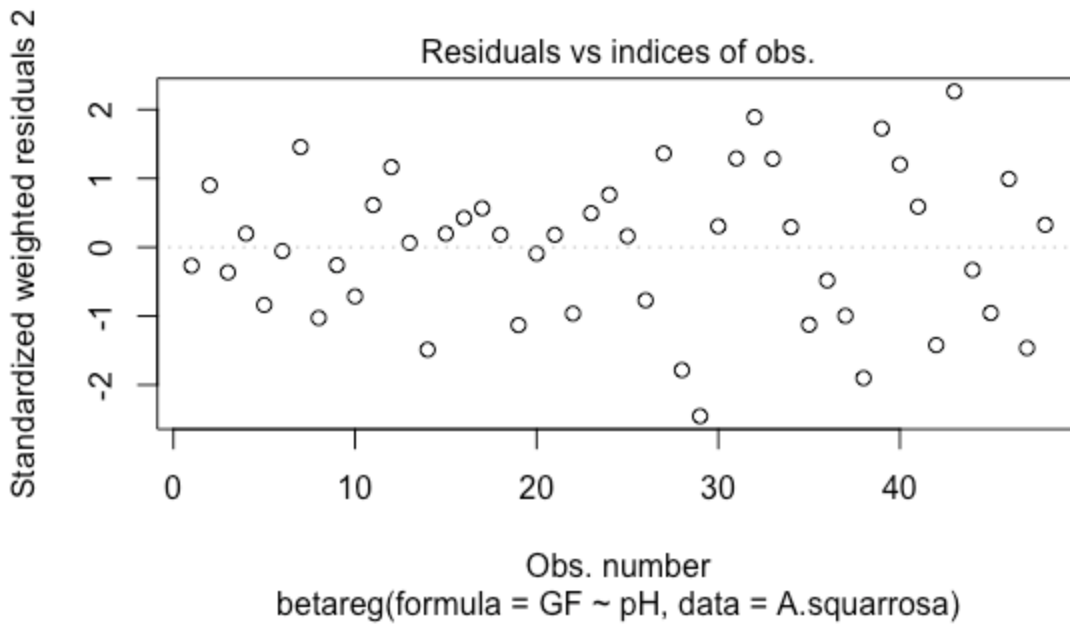**Pseudo R-squared:** 0.9711

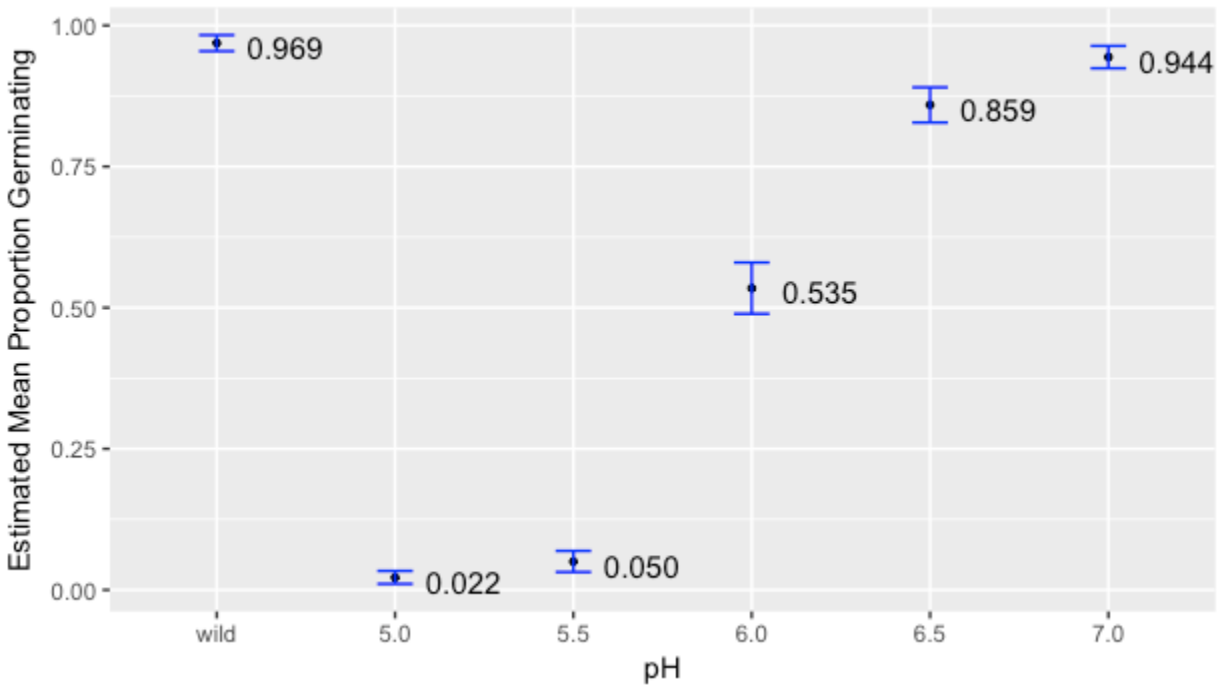**Figure 8.2:** Residuals vs. Observations of A.squarrosa Absent

G.verr absent

Figure 9.1: Estimated Mean Proportion of Germination vs. pH for *G.verrucosa*

Table 9.1: Germination Proportion of *C. albicans* for *G.verrucosa* Absent

pH	Estimated Mean Proportion Germinating	Standard Error	95% Confidence Interval
wild	0.969	0.007	[0.954, 0.983]
5.0	0.022	0.006	[0.011, 0.034]
5.5	0.050	0.010	[0.032, 0.069]
6.0	0.535	0.023	[0.489, 0.580]
6.5	0.859	0.016	[0.828, 0.890]
7.0	0.944	0.010	[0.924, 0.964]

Table 9.2: Difference in Germination Proportion of *C. albicans* for G.*verrucosa* Absent

Contrast	Difference in Estimated Mean Proportion Germinating	Standard Error	Adjusted P-value
6.0 - 5.0	0.512	0.024	< 0.0005***
6.0 - 5.5	0.484	0.025	< 0.0005***
6.0 - 6.5	-0.325	0.028	< 0.0005***
6.0 - 7.0	-0.409	0.025	< 0.0005***
6.0 - wild	-0.434	0.024	< 0.0005***

Table 9.3: Beta Regression Coefficients

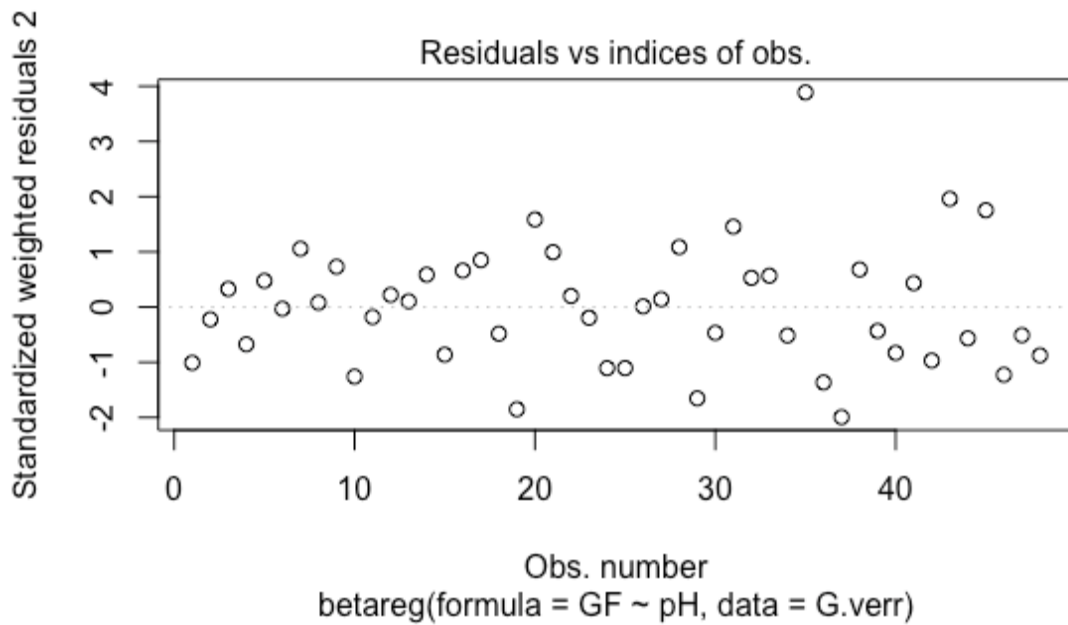|	Estimate	Std. Error	z value	Pr (>|z|)
(Intercept)	0.139	0.093	1.488	0.137
pH 5.0	-3.926	0.286	-13.719	<2e-16 ***
pH 5.5	-3.075	0.220	-13.957	<2e-16 ***
pH 6.5	1.670	0.161	10.363	<2e-16 ***
pH 7.0	2.685	0.213	12.629	<2e-16 ***
pH wild	3.294	0.257	12.810	<2e-16 ***

Pseudo R-squared: 0.962

Figure 9.2: Residuals vs. Observations of G.verr Absent

R code:

```
library(ggplot2)
library(EnvStats)
library(betareg)
library(emmeans)
library(lmtest)
library(gplots)

#task 1
#read data
Aloe<-read.csv("aloe.csv")
#create gemination frequency variable
Aloe$GF<-(Aloe$GermTubes)/500
Aloe$GF <- ifelse(Aloe$GF == 1, 0.999, Aloe$GF)
Aloe$GF <- ifelse(Aloe$GF == 0.000, 0.001, Aloe$GF)
Aloe$pH <- relevel(Aloe$pH, ref = "6")
#only control groups
control<-Aloe[which(Aloe$Aloe=="N"),]
#subset by Aloes
A.arborescens<-control[which(control$Species=="A.arborescens"),]
A.cameronii<-control[which(control$Species=="A.cameronii"),]
A.ciliaris<-control[which(control$Species=="A.ciliaris"),]
A.distans<-control[which(control$Species=="A.distans"),]
A.ferox<-control[which(control$Species=="A.ferox"),]
A.squarrosa<-control[which(control$Species=="A.squarrosa"),]
G.verr<-control[which(control$Species=="G.verr"),]

##A.arborescens
#fit models
beta.fit1<-betareg(GF ~ pH,data=A.arborescens)
beta.fitinter<-betareg(GF ~ 1,data=A.arborescens)
#test the effect of pH
lrtest(beta.fitinter,beta.fit1)
summary(beta.fit1)
#check residuals
plot(beta.fit1)
#emmeans
emmeans1<-emmeans(beta.fit1,pairwise ~pH,model="response")
cl1<-as.data.frame(emmeans1$emmean)
ggplot(cl1, aes(x=pH,y=emmean,label=sprintf("%0.3f", round(emmean, digits = 3))))+
  geom_point(size=0.9)+
  geom_errorbar(width=0.2,aes(ymin=asymp.LCL,ymax=asymp.UCL),color="blue")+
  scale_x_discrete(limits=c("wild","5","5.5","6","6.5","7"),
            labels=c("5"="5.0","6"="6.0","7"="7.0"))+
```

```
  labs(x="pH",y="Estimate Mean Proportion for Germination")+
  geom_text(size  = 4, hjust=-0.4,vjust = 0.7)


##A.cameronii
#fit models
beta.fit2<-betareg(GF ~ pH,data=A.cameronii)
beta.fitinter2<-betareg(GF ~ 1,data=A.cameronii)
#test the effect of pH
lrtest(beta.fitinter2,beta.fit2)
summary(beta.fit2)
#check residuals
plot(beta.fit2)
#emmeans
emmeans2<-emmeans(beta.fit2,pairwise ~pH,model="response")
cl2<-as.data.frame(emmeans2$emmean)
ggplot(cl2, aes(x=pH,y=emmean,label=sprintf("%0.3f", round(emmean, digits = 3))))+
  geom_point(size=0.9)+
  geom_errorbar(width=0.2,aes(ymin=asymp.LCL,ymax=asymp.UCL),color="blue")+
  scale_x_discrete(limits=c("wild","5","5.5","6","6.5","7"),
            labels=c("5"="5.0","6"="6.0","7"="7.0"))+
  labs(x="pH",y="Estimate Mean Proportion for Germination")+
  geom_text(size  = 4, hjust=-0.4,vjust = 0.7)


##A.ciliaris
#fit models
beta.fit3<-betareg(GF ~ pH,data=A.ciliaris)
beta.fitinter3<-betareg(GF ~ 1,data=A.ciliaris)
#test the effect of pH
lrtest(beta.fitinter3,beta.fit3)
summary(beta.fit3)
#check residuals
plot(beta.fit3)
#emmeans
emmeans3<-emmeans(beta.fit3,pairwise ~pH,model="response")
cl3<-as.data.frame(emmeans3$emmean)
ggplot(cl3, aes(x=pH,y=emmean,label=sprintf("%0.3f", round(emmean, digits = 3))))+
  geom_point(size=0.9)+
  geom_errorbar(width=0.2,aes(ymin=asymp.LCL,ymax=asymp.UCL),color="blue")+
  scale_x_discrete(limits=c("wild","5","5.5","6","6.5","7"),
            labels=c("5"="5.0","6"="6.0","7"="7.0"))+
  labs(x="pH",y="Estimate Mean Proportion for Germination")+
  geom_text(size  = 4, hjust=-0.4,vjust = 0.7)


##A.distans
```

```r
beta.fit4<-betareg(GF ~ pH,data=A.distans)
beta.fitinter4<-betareg(GF ~ 1,data=A.distans)
#test the effect of pH
lrtest(beta.fitinter4,beta.fit4)
summary(beta.fit4)
#check residuals
plot(beta.fit4)
#emmeans
emmeans4<-emmeans(beta.fit4,pairwise ~pH,model="response")
cl4<-as.data.frame(emmeans4$emmean)
ggplot(cl4, aes(x=pH,y=emmean,label=sprintf("%0.3f", round(emmean, digits = 3))))+
  geom_point(size=0.9)+
  geom_errorbar(width=0.2,aes(ymin=asymp.LCL,ymax=asymp.UCL),color="blue")+
  scale_x_discrete(limits=c("wild","5","5.5","6","6.5","7"),
              labels=c("5"="5.0","6"="6.0","7"="7.0"))+
  labs(x="pH",y="Estimate Mean Proportion for Germination")+
  geom_text(size = 4, hjust=-0.4,vjust = 0.7)

##A.ferox
beta.fit5<-betareg(GF ~ pH,data=A.ferox)
beta.fitinter5<-betareg(GF ~ 1,data=A.ferox)
#test the effect of pH
lrtest(beta.fitinter5,beta.fit5)
summary(beta.fit5)
#check residuals
plot(beta.fit5)
#emmeans
emmeans5<-emmeans(beta.fit5,pairwise ~pH,model="response")
cl5<-as.data.frame(emmeans5$emmean)
ggplot(cl5, aes(x=pH,y=emmean,label=sprintf("%0.3f", round(emmean, digits = 3))))+
  geom_point(size=0.9)+
  geom_errorbar(width=0.2,aes(ymin=asymp.LCL,ymax=asymp.UCL),color="blue")+
  scale_x_discrete(limits=c("wild","5","5.5","6","6.5","7"),
              labels=c("5"="5.0","6"="6.0","7"="7.0"))+
  labs(x="pH",y="Estimate Mean Proportion for Germination")+
  geom_text(size = 4, hjust=-0.4,vjust = 0.7)

##A.squarrosa
beta.fit6<-betareg(GF ~ pH,data=A.squarrosa)
beta.fitinter6<-betareg(GF ~ 1,data=A.squarrosa)
#test the effect of pH
lrtest(beta.fitinter6,beta.fit6)
summary(beta.fit6)
#check residuals
```

```
plot(beta.fit6)
#emmeans
emmeans6<-emmeans(beta.fit6,pairwise ~pH,model="response")
cl6<-as.data.frame(emmeans6$emmean)
ggplot(cl6, aes(x=pH,y=emmean,label=sprintf("%0.3f", round(emmean, digits = 3))))+
  geom_point(size=0.9)+
  geom_errorbar(width=0.2,aes(ymin=asymp.LCL,ymax=asymp.UCL),color="blue")+
  scale_x_discrete(limits=c("wild","5","5.5","6","6.5","7"),
            labels=c("5"="5.0","6"="6.0","7"="7.0"))+
  labs(x="pH",y="Estimate Mean Proportion for Germination")+
  geom_text(size = 4, hjust=-0.4,vjust = 0.7)


##G.verr
beta.fit7<-betareg(GF ~ pH,data=G.verr)
beta.fitinter7<-betareg(GF ~ 1,data=G.verr)
#test the effect of pH
lrtest(beta.fitinter7,beta.fit7)
summary(beta.fit7)
#check residuals
plot(beta.fit7)
#emmeans
emmeans7<-emmeans(beta.fit7,pairwise ~pH,model="response")
cl7<-as.data.frame(emmeans7$emmean)
ggplot(cl7, aes(x=pH,y=emmean,label=sprintf("%0.3f", round(emmean, digits = 3))))+
  geom_point(size=0.9)+
  geom_errorbar(width=0.2,aes(ymin=asymp.LCL,ymax=asymp.UCL),color="blue")+
  scale_x_discrete(limits=c("wild","5","5.5","6","6.5","7"),
            labels=c("5"="5.0","6"="6.0","7"="7.0"))+
  labs(x="pH",y="Estimate Mean Proportion for Germination")+
  geom_text(size = 4, hjust=-0.4,vjust = 0.7)


#task 2
#read data
Aloe.arborescens<-read.csv("Aloe arborescens_April15_2018.csv")
#create germinated proportion variable
Aloe.arborescens$GF<-Aloe.arborescens$X..cells.forming.germ.tubes/500
Aloe.arborescens$GF <- ifelse(Aloe.arborescens$GF == 1, 0.999, Aloe.arborescens$GF)
Aloe.arborescens$GF <- ifelse(Aloe.arborescens$GF == 0.000, 0.001, Aloe.arborescens$GF)
names(Aloe.arborescens)
head(Aloe.arborescens)
#reoder the pH for plot
Aloe.arborescens$pH<-revalue(Aloe.arborescens$pH, c("5"="5.0", "6"="6.0","7"="7.0"))
Aloe.arborescens$pH1<-factor(Aloe.arborescens$pH, levels=c("wild", "5.0",
"5.5","6.0","6.5","7.0"))
```

```r
#change the variable name
colnames(Aloe.arborescens)[colnames(Aloe.arborescens)=="Aloe..Y.or.N."] <- "AloePresence"
names(Aloe.arborescens)
#fit beta regression model
beta.fit<-betareg(GF ~ pH1*AloePresence,data=Aloe.arborescens)
beta.fitmain<-betareg(GF~pH+AloePresence,data=Aloe.arborescens)
#compare models
lrtest(beta.fitmain,beta.fit)
summary(beta.fit)
#interaction plot
emmip(beta.fit,AloePresence~pH1,xlab="pH",ylab="Estimate Mean Proportion for Germination")
#check residuals
plot(beta.fit)
#emmeans
emmeans(beta.fit,pairwise ~ AloePresents| pH1,model="response")

#citations
citation()
citation("ggplot2")
citation("EnvStats")
citation("betareg")
citation("emmeans")
citation("lmtest")
citation("gplots")
```