Butterfly Preference Analysis

Callin Switzer, 23 January 2017

* Model Output From R

* Model Equation

* Coefficient interpretation

1   Model Output

Here's the output from R

```
Generalized linear mixed model fit by maximum likelihood
(Laplace Approximation) ['glmerMod']
 Family: binomial  ( logit )
Formula: cbind(visits.LB, visits.other) ~
context * pol * array + (1 |      polID) + (1 | date)
    Data: polDS
Control: glmerControl(optimizer = "bobyqa")


     AIC      BIC   logLik deviance df.resid
  1589.3   1677.9   -774.7   1549.3      600


Scaled residuals:
    Min      1Q  Median      3Q     Max
-2.9246 -0.6236  0.1130  0.6624  3.4302


Random effects:
 Groups Name        Variance Std.Dev.
 polID  (Intercept) 0.22914  0.4787
 date   (Intercept) 0.04843  0.2201
Number of obs: 620, groups:  polID, 620; date, 30


Fixed effects:
                                  Estimate Std. Error z value Pr(>|z|)
0 (Intercept)                      0.87083    0.25560   3.407 0.000657 ***
1   contextCvD                    -1.11658    0.37180  -3.003 0.002672 **
2 contextDvDwC                     0.42750    0.36892   1.159 0.246534
3 polBAT                          -0.46801    0.25918  -1.806 0.070958 .
4 arrayDark Red                    1.54985    0.60928   2.544 0.010967 *
5 arrayLight Red                  -0.58868    0.37328  -1.577 0.114779
6 contextCvD:polBAT               -0.56694    0.42581  -1.331 0.183040
7 contextDvDwC:polBAT             -1.46165    0.42327  -3.453 0.000554 ***
8 contextCvD:arrayDark Red        -0.07595    0.71940  -0.106 0.915915
9 contextDvDwC:arrayDark Red      -0.78330    0.78903  -0.993 0.320833
10 contextCvD:arrayLight Red       1.12912    0.51716   2.183 0.029011 *
11 contextDvDwC:arrayLight Red    -0.87196    0.54906  -1.588 0.112261
12 polBAT:arrayDark Red           -1.34469    0.62071  -2.166 0.030284 *
13 polBAT:arrayLight Red           0.78708    0.40239   1.956 0.050467 .
14 contextCvD:polBAT:arrayDark Red   -0.42513    0.80410  -0.529 0.597009
15 contextDvDwC:polBAT:arrayDark Red  1.29409    0.81995   1.578 0.114505
16 contextCvD:polBAT:arrayLight Red  -2.13592    0.62114  -3.439 0.000585 ***
17 contextDvDwC:polBAT:arrayLight Red  1.07866    0.60635   1.779 0.075249 .
```

2 Model Equation (fixed effects only)

P_i = probability that $pollinator_i$ visits light blue flower

$logit(P_i) =$
$\beta_0 +$
β_1*`contextCvD` $+$
β_2*`contextDvDwC` $+$
β_3*`polBAT` $+$
β_4*`arrayDark Red` $+$
β_5*`arrayLight Red` $+$
β_6*`contextCvD:polBAT` $+$
β_7*`contextDvDwC:polBAT` $+$
β_8*`contextCvD:arrayDark Red` $+$
β_9*`contextDvDwC:arrayDark Red` $+$
β_{10}*`contextCvD:arrayLight Red` $+$
β_{11}*`contextDvDwC:arrayLight Red` $+$
β_{12}*`polBAT:arrayDark Red` $+$
β_{13}*`polBAT:arrayLight Red` $+$
β_{14}*`contextCvD:polBAT:arrayDark Red` $+$
β_{15}*`contextDvDwC:polBAT:arrayDark Red` $+$
β_{16}*`contextCvD:polBAT:arrayLight Red` $+$
β_{17}*`contextDvDwC:polBAT:arrayLight Red`

Indicator variables
All the variables in this model are categorical, and are represented as indicator variables. For instance

`contextCvD` denotes an indicator variable

$$\texttt{contextCvD} = \begin{cases} 1, & \text{context} = \text{CvD} \\ 0, & \text{otherwise} \end{cases}$$

`contextDvDwC:polBAT:arrayLight Red` denotes another indicator variable

$$\texttt{contextDvDwC:polBAT:arrayLight Red} = \begin{cases} 1, & \text{context} = \text{DvDwC} \\ & \text{and pol} = \text{BAT} \\ & \text{and array} = \text{Light Red} \\ 0, & \text{otherwise} \end{cases}$$

3 Coefficient Interpretations

The reference levels are SKIP, DvD, and Dark Blue, so everything should be interpreted relative to those things.

The coef for `contextCvD` (β_1) is -1.11658. This means that for a SKIP in a dark blue array, with context CvD, there is a significant difference in probability of choosing light blue, relative to a SKIP in a dark blue array when context = DvD. Specifically, a SKIP is less likely to choose blue flowers in a dark blue array when context = CvD than when context = DvD.

We can calculate the probability of choosing light blue for each of the groups, by plugging values into the model.

For instance, we can calculate the probability of choosing light blue for all the reference levels (SKIP, DvD, and Dark Blue) by inverting the model (using the inverse logit):

$$logit(x) = ln\left(\frac{x}{(1-x)}\right)$$

$$logit^{-1}(x) = \frac{exp(x)}{(exp(x)+1)}$$

To solve for the reference level, we would get this:

$logit(P_i) = \beta_0 + \beta_1*0 + \beta_2*0 + \beta_3*0 + \beta_4*0 + \beta_5*0 + \beta_6*0 + \beta_7*0 + \beta_8*0 + \beta_9*0 + \beta_{10}*0 + \beta_{11}*0 + \beta_{12}*0 + \beta_{13}*0 + \beta_{14}*0 + \beta_{15}*0 + \beta_{16}*0 + \beta_{17}*0$

Which simplifies to $logit(P_i) = \beta_0$. We invert to solve for the probabilty:

$$p_i = \frac{exp(\beta_0)}{(exp(\beta_0)+1)}$$

$$exp(0.87083)/(exp(0.87083)+1) = 0.7049184$$

The coef for `contextDvDwC` (β_2) is 0.42750. There is not enough evidence to say that this coefficient is significant – meaning for a SKIP in a Dark Blue array, there is no difference in the probability of choosing light blue in context DvDwC vs. context DvD.

The coef for `polBAT` (β_3) is -0.46801. There is not enough evidence to say that this coefficient is significant – meaning for DvD in a Dark Blue array, there is no difference in the probability of choosing light blue in SKIP vs. BAT.

The coef for `arrayDark Red` (β_4) is 1.54985. For SKIP in a DvD Context, there is a significant difference in the probability of choosing light blue in a Dark Red vs. Dark Blue. That is, Skippers in a DvD context are more likely to choose light blue in a Dark Red context than in a Dark Blue context.

The coef for `contextDvDwC:polBAT` (β_7) is -1.46165. For Color Dark Blue, The difference between context DvD and context DvDwC for polBAT is 1.46 less (on the logit scale) than the difference between context DvD and context DvDwC for polSKIP.

We can solve for the probability of visiting light flowers for BAT, context DvDwC, array Dark Blue like this:

$logit(P_i) = \beta_0 + \beta_1*0 + \beta_2*1 + \beta_3*1 + \beta_4*0 + \beta_5*0 + \beta_6*0 + \beta_7*1 + \beta_8*0 + \beta_9*0 + \beta_{10}*0 + \beta_{11}*0 + \beta_{12}*0 + \beta_{13}*0 + \beta_{14}*0 + \beta_{15}*0 + \beta_{16}*0 + \beta_{17}*0$

Which simplifies to $logit(P_i) = \beta_0 + \beta_2*1 + \beta_3*1 + \beta_7*1$

We can solve for the probability like this:

$$p_i = \frac{exp(\beta_0 + \beta_2*1 + \beta_3*1 + \beta_7*1)}{(exp(\beta_0 + \beta_2*1 + \beta_3*1 + \beta_7*1)+1)}$$

$$\frac{exp(0.87083 + 0.42750 + -0.46801 + -1.46165)}{exp(0.87083 + 0.42750 + -0.46801 + -1.46165)+1} = 0.347209$$

The coef for `polBAT:arrayDark Red` (β_{12}) is -1.34469. For array DvD, the difference between polBAT and polSKIP is 1.34 less (on the logit scale) than the difference between polBAT and polSKIP when array is Dark Red.

I'm not going to try to interpret the 3-way interactions, but I will show how different the probabilities are. Here's an example of calculating the probability for contextDvDwC, polBAT, arrayLight Red.

$logit(P_i) = \beta_0 + \beta_1 {}^*0 + \beta_2 {}^*1 + \beta_3 {}^*1 + \beta_4 {}^*0 + \beta_5 {}^*1 + \beta_6 {}^*0 + \beta_7 {}^*1 + \beta_8 {}^*0 + \beta_9 {}^*0 + \beta_{10} {}^*0 + \beta_{11} {}^*1 + \beta_{12} {}^*0 + \beta_{13} {}^*1 + \beta_{14} {}^*0 + \beta_{15} {}^*0 + \beta_{16} {}^*0 + \beta_{17} {}^*1$

Which simplifies to $logit(P_i) = \beta_0 + \beta_2 {}^*1 + \beta_3 {}^*1 + \beta_5 {}^*1 + \beta_7 {}^*1 + \beta_{11} {}^*1 + \beta_{13} {}^*1 + \beta_{17} {}^*1$

We can solve for the probability like this:

$$p_i = \frac{exp(\beta_0 + \beta_2 * 1 + \beta_3 * 1 + \beta_5 * 1 + \beta_7 * 1 + \beta_{11} * 1 + \beta_{13} * 1 + \beta_{17} * 1)}{(exp(exp(\beta_0 + \beta_2 * 1 + \beta_3 * 1 + \beta_5 * 1 + \beta_7 * 1 + \beta_{11} * 1 + \beta_{13} * 1 + \beta_{17} * 1) + 1)}$$

$$\frac{exp(0.87083 + 0.42750 + -0.46801 + -0.58868 + -1.46165 + -0.87196 + 0.78708 + 1.07866)}{exp(0.87083 + 0.42750 + -0.46801 + -0.58868 + -1.46165 + -0.87196 + 0.78708 + 1.07866) + 1}$$

$$= 0.4436825$$