

Modeling Flight Response Summary File

Data Cleaning And Exploration

Read Libraries

```
library(lme4) # fit regressions
library(rethinking) # Bayesian data analysis and plotting
library(popbio) # logistic regression plotting
library(binom) # binomial confidence intervals
```

Read Source Files

```
source_path = paste0(dir, "/Rsrc/")

script_names = c("center_flight_data.R", # 1 function: center_data()
                 "clean_flight_data.R", # 1 function: clean_flight_data()
                 "unique_flight_data.R", # 1 function: create_delta_data()
                 "compare_models.R", # 1 function: model_comparisonsAIC()
                 "AICprobabilities.R") # 1 function: AICprobs()

for (script in script_names) {
  path = paste0(source_path, script)
  source(path)
}
```

Read the Data

```
data_path = paste0(dir, "/Dispersal/Winter_2020/stats/data/all_flight_data-Winter2020.csv")

data <- read_flight_data(data_path)
data_all <- data[[1]]
data_tested <- data[[2]]
d <- create_delta_data(data_tested, tested_more_than_once=FALSE)
```

Repeating Plot Parameters & Functions

```
# scale/magnifications
c1 = 1.3*2 # size of points
c2 = 1.2*2 # size of text
c3 = 2*2 # size of title

# compute confidence interval
get_CI = function(x,y,m) {
  x.seq = seq(min(x) - sd(x), max(x) + sd(x), length.out=100)
  prd <- data.frame(x=x.seq) # newdata
```

```

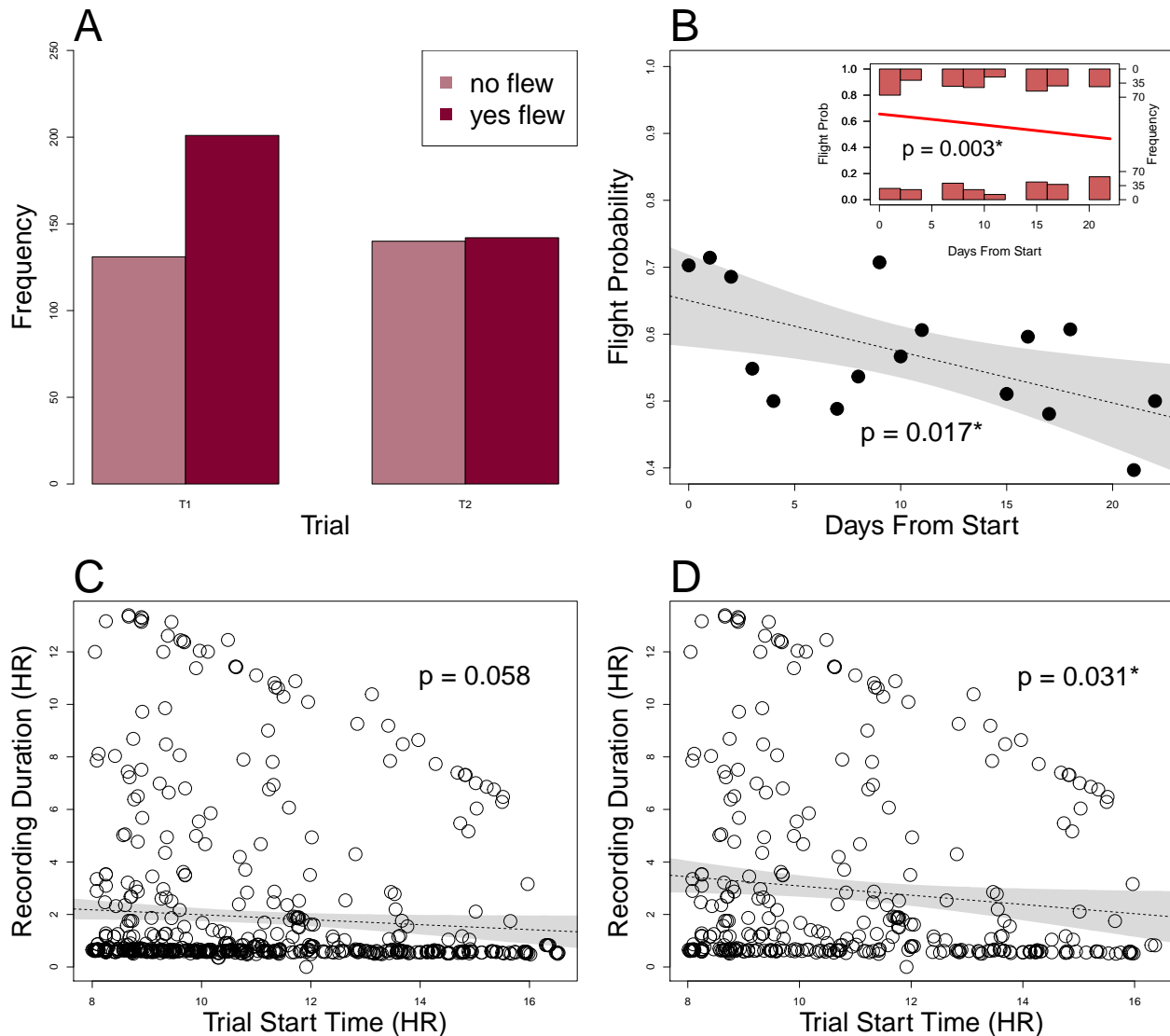
err <- predict(m, newdata = prd, se.fit = TRUE)
prd$lci <- err$fit - 1.96 * err$se.fit
prd$fit <- err$fit
prd$uci <- err$fit + 1.96 * err$se.fit
mu_ci <- t(matrix(c(prd$lci,prd$uci), ncol=2))
return(list(mu_ci, prd))
}

# tailoring variables for plotting
d$mass_block<-round(d$average_mass/0.005)*0.005
d$wing2body_block<-round(d$wing2body, digits=2)
d$days_block<-round(d$avg_days, digits=0)

```

Across-Trial Flight Response

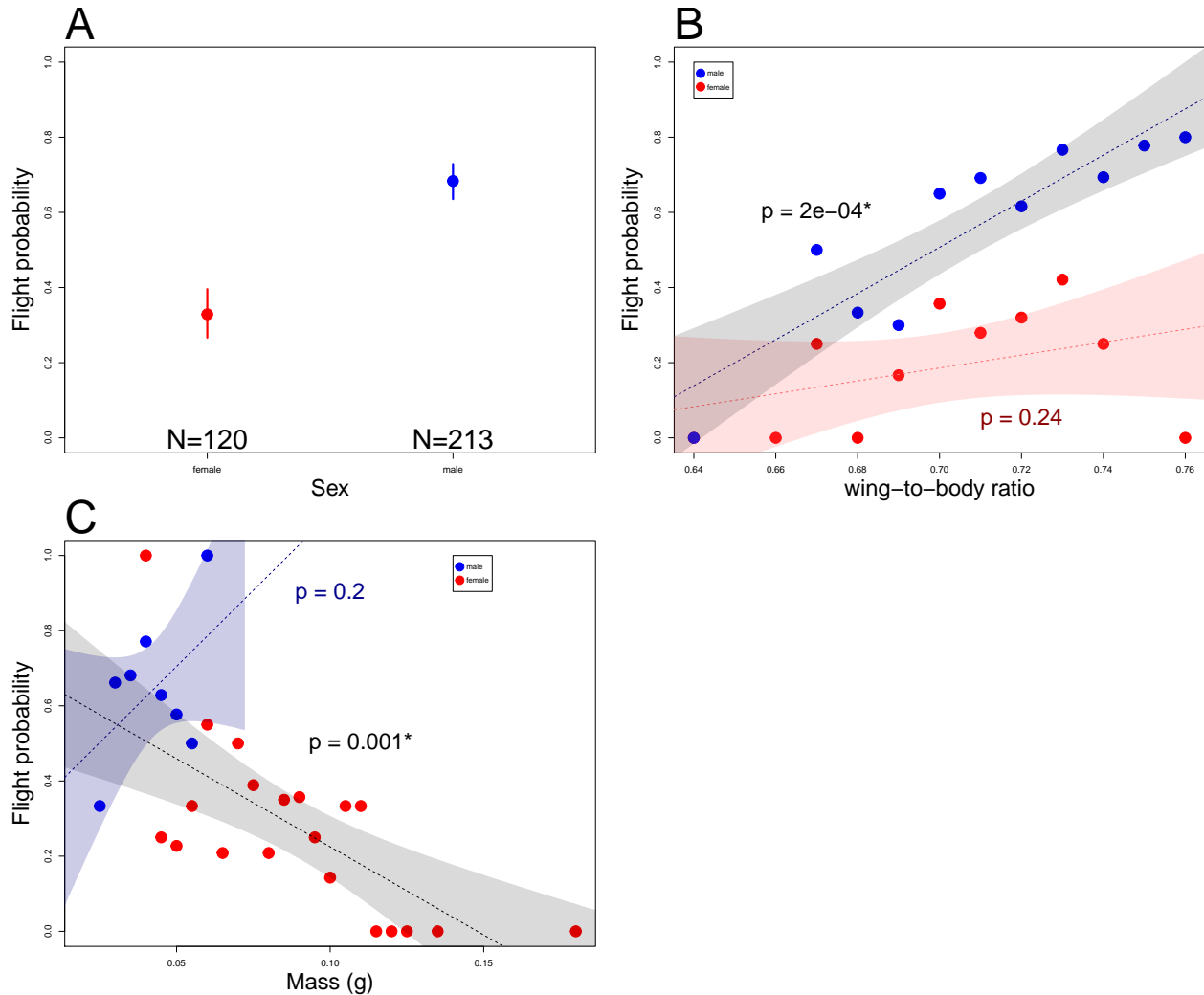
Experimental Effects



A & B. There was a negative effect of day a bug was tested but only when the full dataset is considered

(not the unique dataset). **C & D.** There was a negative effect of the trial start time but only after removing bugs that didn't fly.

Single-Variate Effects



Between-Trial Flight Response (T1 vs. T2)

Read Libraries

```
library(dplyr) # data manipulation
library(zoo)   # data manipulation
library(nnet)  # multinomial modeling
library(kableExtra) # table formatting
```

Read Source Files

```
script_names = c( "get_warnings.R", # 1 function: withWarnings()
                  "multinom_functions.R") # 9 functions: calculate_P(),
                                           # calculate_P2(), calculate_P3(),
```

```

# prediction_equations(),
# prediction_equations2(),
# get_significant_models(),
# get_significant_models2(),
# getting_odds()
# getting_odds2()

for (script in script_names) {
  path = paste0(source_path, script)
  source(path)
}

```

Read the Data

```
d <- create_delta_data(data_tested, tested_more_than_once=TRUE)
```

Delta Flight Response Key	
Event	Encoding
flew in both trials	2
flew in T2 only	1
flew in neither trials	0
flew in T1 only	-1

We wanted to model the probability of different delta flight response cases with sex, host plant, percent changes in mass, and percent changes in egg-laying response as predictors. Since the outcomes (or response variables) were no longer binomial, we used multicategorical logit models. See the Appendix for an explanation and examples of computing multicategorical logit models.

Delta Mass Key	
Event	Sign
gained mass from T1 to T2	pos
no mass change between trails	0
lost mass from T1 to T2	neg

Multinomial Modeling

Baseline

```

df <- d %>%
  filter(!is.na(mass_diff), !is.na(flight_case))

df <- df[with(df, order(mass_per)),]
n_trials = nrow(df)

df$flight_case <- relevel(as.factor(df$flight_case), ref = "0")

```

Null Model

```
null <- multinom(flight_case ~ 1, data = df)
```

```
## # weights:  8 (3 variable)
## initial value 385.389832
## iter  10 value 319.269929
## final value 319.269680
## converged
```

Compare Models

```
data <- data.frame(R = df$flight_case,
  A = df$mass_per,
  B = df$sex_c,
  C = df$host_c)
model_script = paste0(source_path, "generic multinomial models- multinom 1RF + 3 FF.R")
model_comparisonsAIC(model_script)
```

flight case ~ mass %, sex, host

```
##          [,1]      [,2]      [,3]      [,4]
## AICs    587.5607  591.9016  592.3168  592.4231
## models  4          7         13         12
## probs   0.7141852 0.0815063 0.06622882 0.06280119
##
## m4  multinom(formula = R ~ A + B, data = data, trace = FALSE)
## m7  multinom(formula = R ~ A + B + C, data = data, trace = FALSE)
## m13 multinom(formula = R ~ B * C + A, data = data, trace = FALSE)
## m12 multinom(formula = R ~ A * C + B, data = data, trace = FALSE)
anova(m4, m7, test="Chisq") # Adding C (host plant) does not improve fit
anova(m4, m8, test="Chisq") # Adding A*B does not improve fit
```

Likelihood ratio tests of Multinomial Models

##

Response: R

##	Model	Resid. df	Resid. Dev	Test	Df	LR stat.	Pr(Chi)
## 1	A + B	825	569.5607				
## 2	A + B + C	822	567.9016	1 vs 2	3	1.659076	0.6460701

Likelihood ratio tests of Multinomial Models

##

Response: R

##	Model	Resid. df	Resid. Dev	Test	Df	LR stat.	Pr(Chi)
## 1	A + B	825	569.5607				
## 2	A * B	822	569.4209	1 vs 2	3	0.1398496	0.9866598

```
delta_mass_model <- multinom(flight_case ~ mass_per + sex_c, data = df)
model_table = calculate_P2(delta_mass_model, "mass_per", "sex_c")
```

```
## # weights:  16 (9 variable)
## initial value 385.389832
## iter  10 value 286.869825
## iter  20 value 284.809036
## iter  30 value 284.797822
## final value 284.780360
```

```
## converged
##
## AIC: 587.5607
##      (Intercept) mass_per sex_c DF   SEi   SE1   SE2      zi      z1      z2
## -1      -1.015     0.043 -0.692  9 0.239 0.010 0.203  -4.248  4.390  -3.408
##  1      -6.820    -0.009 -5.626  9 0.183 0.026 0.183 -37.245 -0.348 -30.721
##  2       0.124     0.019 -0.902  9 0.167 0.008 0.159   0.742  2.334  -5.684
##      waldi  wald1   wald2 Pi > |z| P1 > |z| P2 > |z|
## -1   18.049 19.272  11.617  0.000  0.000  0.001
##  1  1387.197  0.121 943.764  0.000  0.728  0.000
##  2    0.551  5.447  32.310  0.458  0.020  0.000
```

```
prediction_equations2(model_table, " Mass Percent Change", " Sex ")
```

```
## [1] "Where F = 1"
```

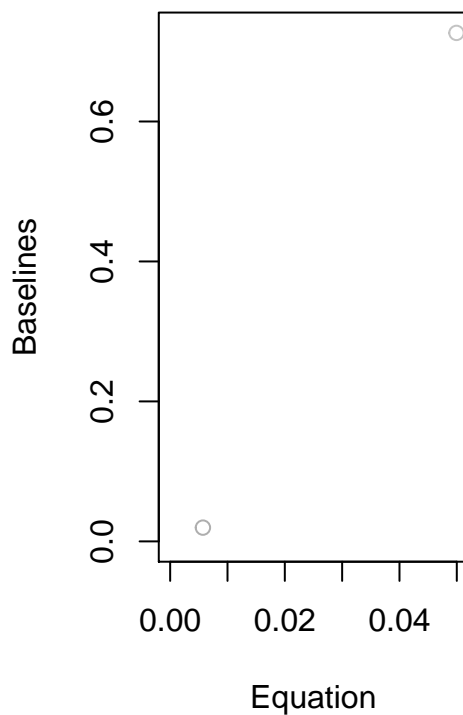
```
## [1] "log(pi_-1 / pi_1) = 5.81 + 0.05 Mass Percent Change + 4.93 Sex      Flew in T1, rather than T2"
```

```
## [2] "log(pi_2 / pi_-1) = 1.14 + -0.02 Mass Percent Change + -0.21 Sex    Flew in both, rather than T2"
```

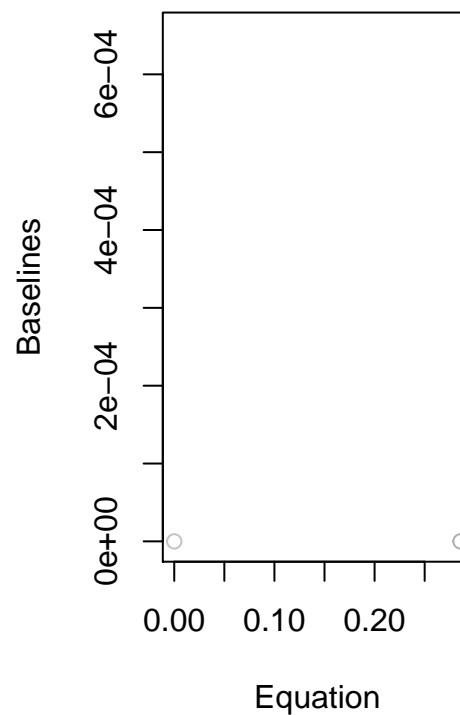
```
## [3] "log(pi_2 / pi_1) = 6.94 + 0.03 Mass Percent Change + 4.72 Sex      Flew in both, rather than T2"
```

```
run_multinom_model = function(d) {
  m <- multinom(flight_case ~ mass_per + sex_c, trace=FALSE, data = d)
  model_table = calculate_P2(m, "mass_per", "sex_c", print_table=FALSE)
  return(model_table)
}
par(mfrow=c(1,2))
MASS_ML = get_significant_models(15) # mass per
SEX_ML = get_significant_models(16) # sex
```

p-values



p-values



```
df <- df[with(df, order(mass_per)),]
df$wing2body_c = df$wing2body - mean(df$wing2body)
df$wing2body_scaled = df$wing2body_c/sd(df$wing2body)*100 # normalized and then multiplied by 100

data <- data.frame(R = df$flight_case,
  A = df$mass_per,
  B = df$sex_c,
  C = df$wing2body_c)
model_script = paste0(source_path,"generic multinomial models- multinom 1RF + 3 FF.R")
model_comparisonsAIC(model_script)
```

flight case ~ mass %, sex, wing2body

```
##          [,1]      [,2]      [,3]
## AICs    582.2678  585.1197  587.133
## models  7         12         13
## probs   0.6671688 0.1603139 0.05858546
##
## m7    multinom(formula = R ~ A + B + C, data = data, trace = FALSE)
## m12   multinom(formula = R ~ A * C + B, data = data, trace = FALSE)
## m13   multinom(formula = R ~ B * C + A, data = data, trace = FALSE)
```

```
anova(m7, m12, test="Chisq") # adding A*C does not improve fit
anova(m7, m13, test="Chisq") # Adding B*C does not improve fit
```

Likelihood ratio tests of Multinomial Models

##

Response: R

##	Model	Resid. df	Resid. Dev	Test	Df	LR stat.	Pr(Chi)
## 1	A + B + C	822	558.2678				
## 2	A * C + B	819	555.1197	1 vs 2	3	3.148182	0.3693379

Likelihood ratio tests of Multinomial Models

##

Response: R

##	Model	Resid. df	Resid. Dev	Test	Df	LR stat.	Pr(Chi)
## 1	A + B + C	822	558.2678				
## 2	B * C + A	819	557.1330	1 vs 2	3	1.134887	0.7686596

```
model <- multinom(flight_case ~ mass_per + sex_c + wing2body_c, data = df)
model_table = calculate_P3(model)
```

weights: 20 (12 variable)

initial value 385.389832

iter 10 value 286.740091

iter 20 value 280.436850

iter 30 value 279.437125

iter 40 value 279.174660

iter 50 value 279.134087

final value 279.133921

converged

##

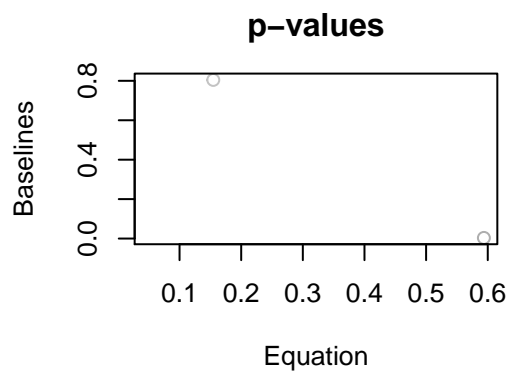
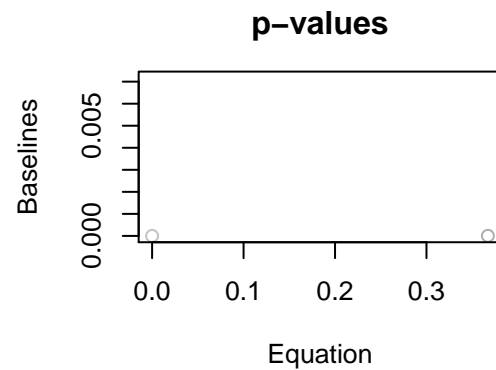
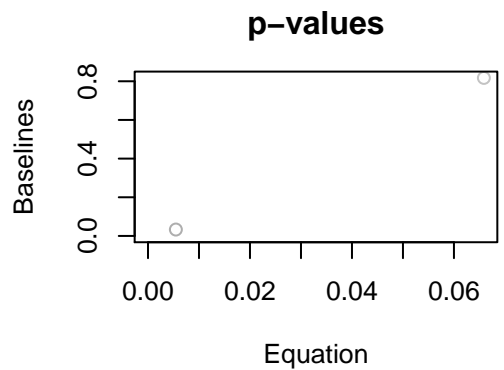
AIC: 582.2678

##	(Intercept)	mass %	sex	wing2body	DF	SEi	SE1	SE2	SE3	zi
## -1	-0.935	0.041	-0.571	23.739	12	0.243	0.010	0.212	12.059	-3.854

```
## 1      -8.177 -0.005 -6.954      -6.595 12 0.187 0.025 0.187 18.786 -43.767
## 2       0.201  0.018 -0.760      28.094 12 0.172 0.008 0.166  9.718  1.173
##      z1      z2      z3      waldi wald1      wald2 wald3 Pi>|z| P1>|z| P2>|z|
## -1  4.254  -2.698  1.969      14.850 18.096      7.278 3.875  0.000  0.000  0.007
## 1  -0.215 -37.102 -0.351 1915.510  0.046 1376.550 0.123  0.000  0.830  0.000
## 2   2.141 -4.590  2.891      1.375  4.585      21.071 8.357  0.241  0.032  0.000
##      P3>|z|
## -1  0.049
## 1   0.726
## 2   0.004
```

```
run_multinom_model = function(d) {
  m <- multinom(flight_case ~ mass_per + sex_c + wing2body_c, trace=FALSE, data = d)
  model_table = calculate_P3(m, print_table=FALSE)
  return(model_table)
}
```

```
# ML's below are all the same
par(mfrow=c(2,2))
MASS_PER_ML = get_significant_models(19) # mass%
SEX_ML = get_significant_models(20) # sex
WING2BODY_ML = get_significant_models(21) # wing2body
```



```
#library(knitr) # table formatting
#library(dplyr) # data manipulation
#library(ggformula) # ggplot plotting
#library(cowplot) # ggplot helper functions to arrange multi-panel figures
```