WRITING STATISTICS

DR JOSH HODGE – JHODGE@IC.AC.UK

BASIC STRUCTURE OF WRITING STATISTICS (METHODS TO DISCUSSION)

Methods

- Research question/statement
- Hypothesis
- Statistical test
- Details about the response and explanatory variable(s)
  - Response variable
  - Fixed effects/factors
  - Interactions
  - Random effects
  - Any transformations? Why?
- Assumptions

Results

- Descriptive statistics paragraph
- The main result in answering the research question
- Statistical test
- Parameter estimates, degrees of freedom, test statistics and p-values
- Supporting table(s) and figure(s)
- Primary interpretation
  - This suggests…

Discussion

- Main results
- Comparison with existing literature
- Final interpretation
- Future study/implications for conservation etc.

EXAMPLE FROM A PAPER

Statistical analyses

Before averaging each feeding behavior trait over the test period, repeatability estimates for the feeding behavior traits were calculated as:

$$t = \frac{\sigma_b^2}{(\sigma_b^2 + \sigma_w^2)}$$

where t denotes repeatability, σ_b^2 denotes the between-animal variance, and σ_w^2 denotes the within-animal variance. Variance components were estimated in linear mixed models using PROC MIXED (SAS v9.4, SAS Institute Inc., Cary, NC, USA) where the fixed effects were those as described later; animal within batch was included as a random effect. Repeatability estimates were not calculated for the time interval traits. After the feeding behavior traits were averaged across the test period, the factors associated with each of the feeding behavior, performance, and efficiency traits were quantified using univariate linear mixed models also in SAS 9.4. Fixed effects considered in all models were age in months at the end of test, sex (steer or heifer), heterosis coefficient class, recombination loss coefficient class, the number of animals in a pen (4, 5, or 6), dam parity (1, 2, 3, 4, ≥5, and missing), and animal breed proportion fitted as a series of linear covariates, with a separate covariate in each model for each breed. Breeds included in all analyses were Holstein-Friesian, Aberdeen-Angus, Hereford, Shorthorn, Limousin, Charolais, Simmental, Aubrac, Parthenaise, Saler, Blonde d'Aquitaine, and Belgian-Blue. Pen nested within batch was included as a random effect in all models. The residuals from the mixed models for all traits were retained and used to estimate partial Spearman's rank correlations among and between the feeding behavior, performance, and efficiency traits. In a separate analysis with MEI as the dependent variable, the traits reflecting feeding time per day, feed events per day, total mealtime per day, and meals per day were progressively added to a model already including MBW, ADG, UFD and two-way interactions between UFD with both ADG and MBW to investigate their marginal contribution to explaining the variability in MEI.

Results

Raw descriptive statistics and repeatability estimates for the feeding behavior traits are presented in Table 1. On average, the cattle in the present study ate for 142.36 min per day, with a feeding frequency of 34.41 feed events per day; the average time per feed event was 4.88 min. Of all of the traditional feeding behavior traits, energy intake per feed event had the largest inter-animal variability (coefficient of variation [CV] = 38.1%), while the number of feed events per meal had the largest CV of 36.9% of the meal behavior traits. Repeatability estimates for the traditional feeding behavior traits at the day level were larger than the repeatability estimates of the feed event-level feeding behaviors. Similarly, the day-level meal behavior repeatability estimates were greater than the repeatability estimates for the meal-level traits. Interestingly, the within-pen feeding order after the feed was refreshed was moderately repeatable indicating that cattle were somewhat consistent in the order they went to feed; feeding order repeatability estimates were also similar between heifers and steers (repeatability estimates of 0.40 and 0.37 for heifers and steers, respectively).

Table 1. Raw means, standard deviations (SD), and repeatability estimates (t) for the traditional feeding behavior traits, feeding order, and the meal behavior traits[1]

Trait	Mean	SD	t
Feed events per day	34.41	12.30	0.57
Feeding time per day, min	142.36	25.07	0.36
Feeding rate, MJ/min	1.09	0.23	0.48
Feed event duration, min	4.88	1.78	0.13
Energy intake per feed event, MJ	5.08	1.94	0.12
Feeding order			0.39
Mean time between feed events, min	42.31	15.16	
Meals per day	8.13	1.12	0.21
Total meal time per day, min	228.51	38.19	0.32
Feed events per meal	4.28	1.58	0.06
Energy intake per meal, MJ	18.50	3.94	0.03
Meal duration, min	17.87	4.16	0.05
Mean time between meals, min	151.55	23.91	
Mean time between feed events within a meal, min	3.57	1.04	

[1]Means and SD were not calculated for feeding order nor were repeatability estimates calculated for the time interval traits.

Correlations between meal behavior traits and both performance and efficiency traits

In general, cattle that consumed more energy per day also had a longer total mealtime per day, had a greater energy intake per meal, and had longer duration individual meals with more feed events within a meal (Table 6). Both heavier cattle and faster-growing cattle consumed more energy per meal. The meal feeding behavior of cattle differed depending on whether the animal was deemed to be efficient or not. Across all feed efficiency traits, more efficient cattle generally had more meals per day albeit the correlations between the efficiency traits and meals per day were weak; the more efficient animals also had a shorter total mealtime per day.

Table 6. Partial Spearman correlations between the meal behavior traits with performance and efficiency traits[1]

Trait	MEI	ADG	MBW	UFD	ECR	REI	REI_U	RG	RG_U	RIG	RIG_U
Meals per day	−0.07	0.05	−0.01	0.05	−0.06	−0.08	−0.09	0.06	0.07	0.09	0.10
Total meal time per day	0.17	0.06	−0.01	0.21	0.08	0.22	0.21	−0.01	−0.01	−0.13	−0.14
Feed events per meal	0.31	−0.02	−0.05	0.11	0.23	0.40	0.43	−0.10	−0.15	−0.31	−0.36
Energy intake per meal	0.66	0.20	0.31	0.13	0.22	0.54	0.55	−0.02	−0.04	−0.35	−0.37
Meal duration	0.24	0.09	0.07	0.12	0.05	0.22	0.19	0.03	0.04	−0.12	−0.10
Mean time between meals	0.04	−0.05	0.02	−0.07	0.05	0.03	0.05	−0.06	−0.06	−0.05	−0.07
Mean time between feed events within a meal	−0.32	−0.05	−0.04	−0.12	−0.19	−0.37	−0.38	0.07	0.09	0.28	0.30

[1]Spearman correlations ≤ |0.08| were not different ($P > 0.05$) from zero.

METHODS EXAMPLE

```
Linear mixed model fit by REML ['lmerMod']
Formula:
NegPerChick ~ FoodTreatment + SexParent + ArrivalTime.z + (1 |
    Nest)
   Data: Owls

REML criterion at convergence: 2192.9

Scaled residuals:
    Min      1Q  Median      3Q     Max
-1.8549 -0.6781 -0.2187  0.5226  4.9416

Random effects:
 Groups   Name        Variance Std.Dev.
 Nest     (Intercept) 0.2599   0.5098
 Residual             2.1292   1.4592
Number of obs: 599, groups:  Nest, 27

Fixed effects:
                      Estimate Std. Error t value
(Intercept)            1.96304    0.15685  12.515
FoodTreatmentSatiated -0.85689    0.12554  -6.826
SexParentMale          0.08019    0.13316   0.602
ArrivalTime.z         -0.37922    0.06155  -6.161
```

LINEAR MODELS- METHODS

To examine how satiation, parental sex and arrival time influences owl nestling calls, a linear mixed model was fitted. The response variable was number of calls per nestling in a single visitation by a parent and there were three explanatory variables: a two-level fixed factor of sex (male and female), a two-level fixed factor of food treatment (deprived and satiated) and a continuous variable of arrival time (in seconds) that was z-standardized. Arrival time was z-standardized to improve interpretability of the model. All explanatory variables were fitted additively as each variable was thought to be independent. Nest was fitted as a random effect to account for non-independence between nest visits due to similar environmental conditions and parental care. The assumptions of normality of residuals, homogeneity of variance and outliers were examined through diagnostic plots and no violations were detected. Statistical significance of fixed effects was judged consider when the *t*-value was < -1.96 or >1.96.

- Research question/statement
- Statistical test
- Details and justification about model structure.
- Assumptions

RESULTS - DESCRIPTIVE STATISTICS

- Every results section should start with descriptive statistics paragraph
- Important characteristics:
  - Defines the sample size (n) and important sub-samples
  - Provides measures of central tendency and spread (depending on the data type – normal, Poisson or binomial) for the variables in the data.

"Data from 599 owl nest visits was collected for 27 independent nests. Of these visits, three hundred and twenty were to food deprived chicks and 279 to satiated chicks. Parents arrived on average 24.76 seconds after chick calls with a standard deviation of 1.91 seconds, and mean chick calls were 1.56 calls per chick (SD=1.61). Males visited nests more frequently (354 visits) than females (245 visits) to broods ranging from one to seven chicks with a median of four."

LINEAR MODELS– RESULTS

```
Linear mixed model fit by REML ['lmerMod']
Formula:
NegPerChick ~ FoodTreatment + SexParent + ArrivalTime.z + (1 |
    Nest)
   Data: Owls

REML criterion at convergence: 2192.9

Scaled residuals:
    Min      1Q  Median      3Q     Max
-1.8549 -0.6781 -0.2187  0.5226  4.9416

Random effects:
 Groups   Name        Variance Std.Dev.
 Nest     (Intercept) 0.2599   0.5098
 Residual             2.1292   1.4592
Number of obs: 599, groups:  Nest, 27

Fixed effects:
                      Estimate Std. Error t value
(Intercept)            1.96304    0.15685  12.515
FoodTreatmentSatiated -0.85689    0.12554  -6.826
SexParentMale          0.08019    0.13316   0.602
ArrivalTime.z         -0.37922    0.06155  -6.161
```

	NegPerChick		
Predictors	*Estimates*	*CI*	*p*
(Intercept)	1.96	1.65 – 2.27	**<0.001**
FoodTreatment [Satiated]	-0.86	-1.10 – -0.61	**<0.001**
SexParent [Male]	0.08	-0.18 – 0.34	0.547
ArrivalTime z	-0.38	-0.50 – -0.26	**<0.001**
Random Effects			
σ^2	2.13		
τ_{00} Nest	0.26		
ICC	0.11		
N Nest	27		
Observations	599		
Marginal R^2 / Conditional R^2	0.117 / 0.213		

LINEAR MODELS– RESULTS TABLES

Fixed Effects		
Coefficient	Estimate+SE	t-value
Intercept	1.96±0.16	12.52
Satiated	-0.86±0.13	-6.83
Male	0.08±0.13	0.60
z-ArrivalTime	-0.38±0.06	-6.16
Random Effects		
Nest	0.26	
Residual	2.13	

Table-1: Coefficients from the linear mixed model examining the relationship between calls per chick and food treatment, parental sex and arrival time. Coefficients (β) are presented alongside their standard errors (SE). The random intercept for NestID reflects variability attributable to nest-specific factors. Significance was assumed when $t < -1.96$ or > 1.96.

LINEAR MODELS– RESULTS

Table-1: Coefficients from the linear mixed model examining the relationship between calls per chick and food treatment, parental sex and arrival time. Coefficients (β) are presented alongside their standard errors (SE). The random intercept for NestID (variance estimates) reflects variability attributable to nest-specific factors. Significance was assumed when *t*< -1.96 or > 1.96.

Fixed Effects		
Coefficient	Estimate+SE	t-value
Intercept	1.96±0.16	12.52
Satiated	-0.86±0.13	-6.83
Male	0.08±0.13	0.60
z-ArrivalTime	-0.38±0.06	-6.16
Random Effects		
Nest	0.26	
Residual	2.13	

Owl chicks had fewer calls when satiated and their parents took longer to arrive. A linear mixed model found significant effects of food treatment and arrival time, but no effect of parental sex (Table-1). Satiated owl chicks were found to have 0.86 fewer calls per minute than their food deprived counterparts (SE=0.13, t=-6.83). Parental arrival time found for a one standard deviation increase in arrival time calls per chick decreased by 0.38 (SE=0.06, t=-6.16). Nest was found to explain 10.88% of variation in call per chick. The results suggest hungry owl chicks of attentive parents (shorter arrival time) call more often.

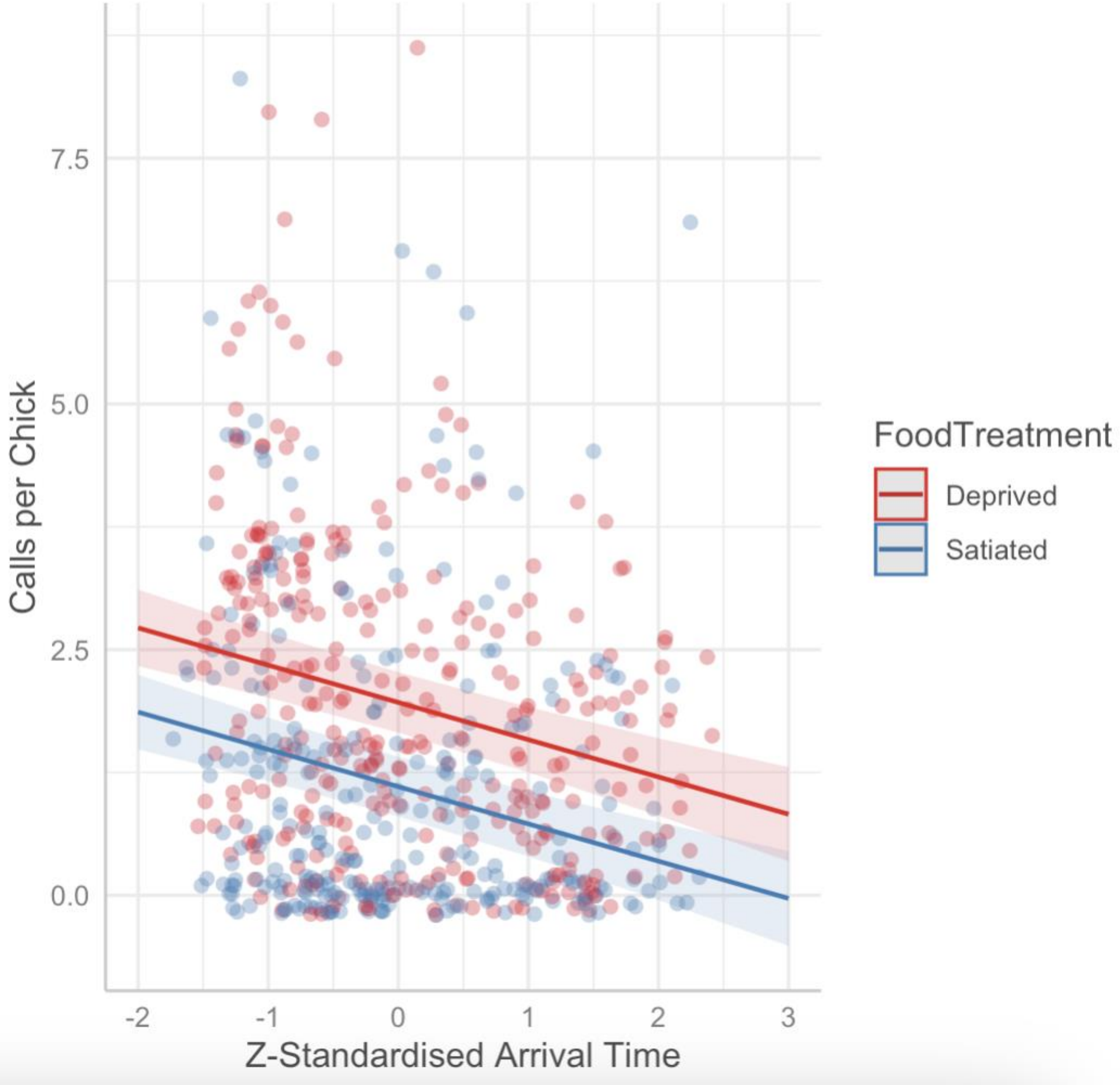
LINEAR MODELS– RESULTS

Owl chicks had fewer calls when satiated and their parents took longer to arrive. A linear mixed model found significant effects of food treatment and arrival time, but no effect of parental sex (Table-1). Satiated owl chicks were found to have 0.86 fewer calls per minute than their food deprived counterparts (SE=0.13, t=-6.83). Parental arrival time found for a one standard deviation increase in arrival time calls per chick decreased by 0.38 (SE=0.06, t=-6.16). Nest was found to explain 10.88% of variation in call per chick. The results suggest hungry owl chicks of attentive parents (shorter arrival time) call more often.

- The main result in answering the research question(s)
- Statistical test
- Parameter estimates, degrees of freedom, test statistics and p-values
- Primary interpretation

LINEAR MODELS– RESULTS

Figure-1: Relationship between calls per chick and z-standardised arrival time. Each point denotes a separate visit classified as deprived (red) or satiated (blue). Regression lines with 95% confidence intervals are plotted according to linear mixed model coefficients from Table-1.



SUMMARY – MAIN POINTS

- Read papers! Some report badly…
- Key points to report:
  - Descriptive statistics (N, measures of central tendency and/or spread)
  - **Justifications** for the inferential statistics (To compare… to investigate…)
  - Parameter estimates (measures of central tendency and/or spread, intercepts and slopes)
  - Test statistics (F- and t-statistics)
  - Degrees of freedom
  - p-values (where appropriate)
- Don’t forget the ecology! You need an interpretation sentence:
  - This suggests….

SUGGESTED LAYOUT OF ASSIGNMENT

Title

"Research Question"
Hypothesis(ses):

- XXX

Statistical Analyses
XXX

Results
XXXDescriptive statisticsXXX

XXXMain result(s)XXX

1

Table-1:XXX

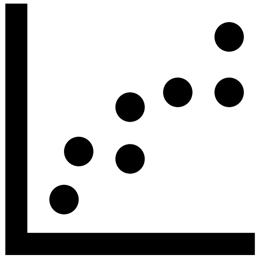


Figure-1:XXX

- No background
- No methods (apart from statistical analyses)
  - You might need to give brief details of data wrangling if you have constructed a new variable from the data
- No discussion
- No references
- You have a max of 3 pages!