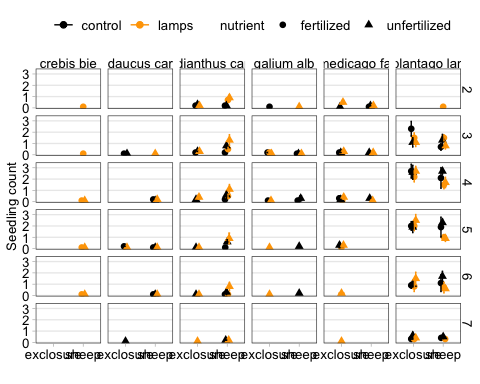
edivalo-seedlings-toothpicks

## Setup and data structuring

Loads libraries, sets up custom theme for ggplot, and reads in data (code not printed)

## Visualization plots

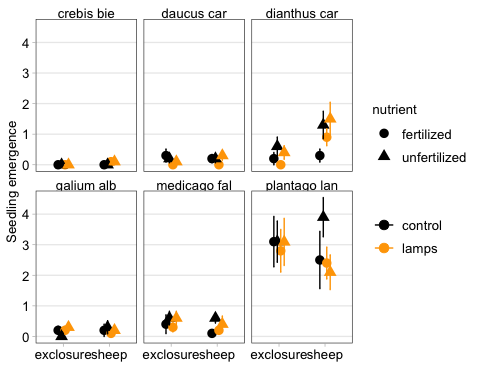
### All species through time

Taking a first look at our data, just the abundance counts of the six species through time. Months are listed on the right side (2-February to 7-July). Seedling abundance was generally highest in March-April-May, and the most common species were Plantago lanceolata and Dianthus. 

## Emergence

Data re-structuring to look at emergence rather than abundance. We also want to add in zeros for the ones that are missing for the analyses, we’re using a complete plot list for this. It is multiplied to have all the species in each plot.

### Emergence Plot

Plotting emergence by species and treatment 

### Emergence Analysis

Separate model for each species, only looking at dianthus and plantago. Currently using negative binomial, seem to be somewhat underdispersed

## Family: nbinom2 ( log )  
## Formula: emerged ~ grazing \* light + nutrient \* light + (1 | block)  
## Data: emerg.pla  
##   
## AIC BIC logLik deviance df.resid   
## 356.8 376.0 -170.4 340.8 73   
##   
## Random effects:  
##   
## Conditional model:  
## Groups Name Variance Std.Dev.  
## block (Intercept) 0.04437 0.2106   
## Number of obs: 81, groups: block, 10  
##   
## Overdispersion parameter for nbinom2 family (): 4.25   
##   
## Conditional model:  
## Estimate Std. Error z value Pr(>|z|)   
## (Intercept) 0.98529 0.22003 4.478 7.53e-06 \*\*\*  
## grazingsheep 0.02303 0.23757 0.097 0.923   
## lightlamps 0.08609 0.29769 0.289 0.772   
## nutrientunfertilized 0.24358 0.23898 1.019 0.308   
## grazingsheep:lightlamps -0.30254 0.34446 -0.878 0.380   
## lightlamps:nutrientunfertilized -0.24816 0.34463 -0.720 0.471   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

## [1] 0.8916475

## Family: nbinom2 ( log )  
## Formula: emerged ~ grazing \* light + nutrient \* light + (1 | block)  
## Data: emerg.dia  
##   
## AIC BIC logLik deviance df.resid   
## 172.2 191.3 -78.1 156.2 72   
##   
## Random effects:  
##   
## Conditional model:  
## Groups Name Variance Std.Dev.  
## block (Intercept) 0.06986 0.2643   
## Number of obs: 80, groups: block, 10  
##   
## Overdispersion parameter for nbinom2 family (): 2.09   
##   
## Conditional model:  
## Estimate Std. Error z value Pr(>|z|)   
## (Intercept) -1.8196 0.5922 -3.073 0.00212 \*\*  
## grazingsheep 0.6882 0.5036 1.367 0.17172   
## lightlamps -0.3490 0.8587 -0.406 0.68441   
## nutrientunfertilized 1.3300 0.5553 2.395 0.01663 \*   
## grazingsheep:lightlamps 1.1539 0.7794 1.480 0.13875   
## lightlamps:nutrientunfertilized -0.4995 0.7432 -0.672 0.50153   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

## [1] 0.9367955

## Survival

Survival structuring

Below is the code Mia had. Some of this wasn’t working or didn’t make sense to me, so I ended up taking a different approach. I’m keeping this here in case reverse changes need to be made :-)

Questions: 1. Are the days based on the actual days we sampled? 2. Cleaner way to do the sampling time etc without all the ifelse statements (ex matching to a vector) 3. Right now the “event” column is just the opposite of the “alive” column. I think it’s supposed to be the first time a seedling dies, which takes a little more work. But I need to check with this package, I think it’s different from the ones I used before (survival + survminer)

I think we want to limit this to the toothpicks where seedlings did emerge, not all of them. So I’m going to start with the mark.emergence data set, and then increase it to have the data needed for each month. We can use the tidyr function crossing() to expand it to the size we need, and then fill in the number column by merging with the mark.seedlings dataset and setting the NA’s to 0 (when a seedling wasn’t observed).

surv.full <- data.frame(month = 2:7) %>%   
 crossing(mark.emergence) %>%   
 left\_join(mark.seedlings) %>%  
 replace\_na(list(number = 0))

## Joining, by = c("month", "block", "plot", "grazing", "nutrient", "light", "species", "id..1.6.")

surv.full$alive <- ifelse(surv.full$number > 0, 1, 0)  
  
  
# this doesn't use the full time intervals the way Mia has it set up  
surv.full.2 <- surv.full %>%  
 group\_by(block, plot, grazing, nutrient, light, species, id..1.6.) %>%  
 dplyr::summarise(  
 dead = 1 - alive[month == max(month)], # at the end dead (1) or alive (0)  
 month.emerged = min(month[alive > 0]),  
 month.died = max(month[alive > 0]),  
 months.alive = (month.died - month.emerged) + 1, # not sure if we want the +1 or not  
 months.alive.2 = sum(alive)  
 )

## `summarise()` regrouping output by 'block', 'plot', 'grazing', 'nutrient', 'light', 'species' (override with `.groups` argument)

surv.pla <- filter(surv.full.2, species == 'plantago lan')  
surv.dia <- filter(surv.full.2, species == 'dianthus car')  
  
surv.obj.pla <- Surv(surv.pla$months.alive, surv.pla$dead, type = 'right')  
surv.obj.dia <- Surv(surv.dia$months.alive, surv.dia$dead, type = 'right')

Running models for the survival using the coxme package.

Here I used right-censored models (number of months alive, whether or not they’re currently dead). Mia uses a different survival structure, with day of year (could be month or day, that part isn’t super important) and event of dying, but this doesn’t include information on when they germinated?

The Anova part doesn’t seem to be working? But summary() has p-values so I’m not sure why we need it?

# Mia: We might need to add another level of hierachy in the random term?  
  
# Dianthus--doesn't run if 3-way interaction is used  
m.surv.dia <- coxme(surv.obj.dia ~ grazing\*light + nutrient\*light + (1|block/plot),   
 data = surv.dia)  
summary(m.surv.dia) # nothing sigificant

## Cox mixed-effects model fit by maximum likelihood  
## Data: surv.dia  
## events, n = 39, 43  
## Iterations= 6 34   
## NULL Integrated Fitted  
## Log-likelihood -118.355 -114.5461 -107.0574  
##   
## Chisq df p AIC BIC  
## Integrated loglik 7.62 7.00 0.367500 -6.38 -18.03  
## Penalized loglik 22.60 10.26 0.014084 2.08 -14.99  
##   
## Model: surv.obj.dia ~ grazing \* light + nutrient \* light + (1 | block/plot)   
## Fixed coefficients  
## coef exp(coef) se(coef) z p  
## grazingsheep -0.68489808 0.5041416 0.5788259 -1.18 0.24  
## lightlamps -0.96610281 0.3805633 1.1188204 -0.86 0.39  
## nutrientunfertilized -0.31635172 0.7288031 0.6504605 -0.49 0.63  
## grazingsheep:lightlamps 0.90306973 2.4671650 0.9673624 0.93 0.35  
## lightlamps:nutrientunfertilized -0.06840758 0.9338798 0.8317730 -0.08 0.93  
##   
## Random effects  
## Group Variable Std Dev Variance   
## block/plot (Intercept) 0.019926841 0.000397079  
## block (Intercept) 0.745547139 0.555840536

#Anova(m.surv.dia, type = "II", test.statistic = "Chisq")  
  
# Plantago  
m.surv.pla <- coxme(surv.obj.pla ~ grazing\*light + nutrient\*light + (1|block/plot),   
 data = surv.pla)  
summary(m.surv.pla) # nothing significant

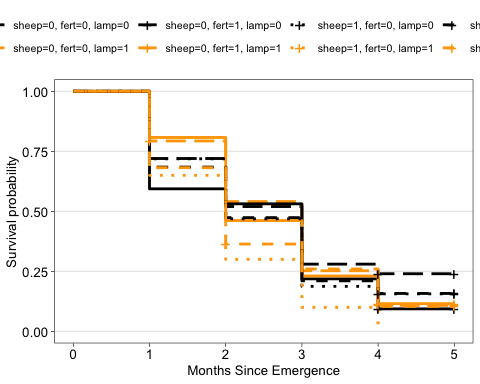
## Cox mixed-effects model fit by maximum likelihood  
## Data: surv.pla  
## events, n = 178, 205  
## Iterations= 15 79   
## NULL Integrated Fitted  
## Log-likelihood -822.2637 -819.1259 -802.7092  
##   
## Chisq df p AIC BIC  
## Integrated loglik 6.28 7.00 0.5079700 -7.72 -30.00  
## Penalized loglik 39.11 18.54 0.0035129 2.02 -56.99  
##   
## Model: surv.obj.pla ~ grazing \* light + nutrient \* light + (1 | block/plot)   
## Fixed coefficients  
## coef exp(coef) se(coef) z p  
## grazingsheep 0.12921494 1.1379347 0.2557812 0.51 0.61  
## lightlamps -0.05140463 0.9498942 0.3018635 -0.17 0.86  
## nutrientunfertilized -0.12727551 0.8804911 0.2555850 -0.50 0.62  
## grazingsheep:lightlamps 0.22623402 1.2538691 0.3630595 0.62 0.53  
## lightlamps:nutrientunfertilized 0.22432345 1.2514757 0.3623255 0.62 0.54  
##   
## Random effects  
## Group Variable Std Dev Variance   
## block/plot (Intercept) 3.502410e-01 1.226688e-01  
## block (Intercept) 9.076584e-03 8.238438e-05

# Anova(m.surv.pla, type = "II", test.statistic = "Chisq") this doesn't run?

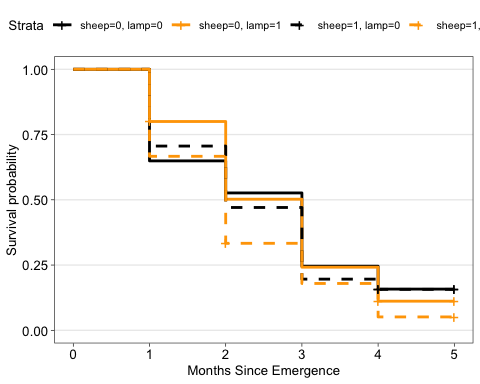
### Survival plots

Trying to take a look at the survival patterns, just for plantago. Plotting package doesn’t take random effects (ex block) into account

# this just makes the legend of the plot easier to see  
surv.pla.2 <- surv.pla  
surv.pla.2$sheep <- ifelse(surv.pla$grazing=='sheep', 1, 0)  
surv.pla.2$fert <- ifelse(surv.pla$nutrient=='fertilized', 1, 0)  
surv.pla.2$lamp <- ifelse(surv.pla$light=='lamps', 1, 0)  
  
surv.fit.pla <- survfit(surv.obj.pla ~ sheep + fert + lamp,   
 data = surv.pla.2)  
ggsurvplot(surv.fit.pla,   
 palette=c('black','orange','black','orange',  
 'black','orange','black','orange'),   
 linetype = c('longdash', 'longdash','solid', 'solid',   
 'dotted','dotted', 'dashed', 'dashed'),   
 ggtheme = theme\_cw(),   
 xlab='Months Since Emergence')

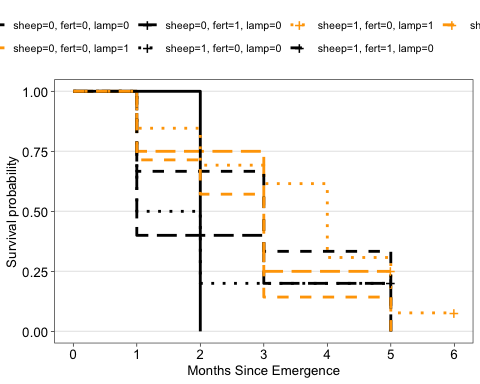


# just grazing and sheep  
surv.fit.pla.2 <- survfit(surv.obj.pla ~ sheep + lamp,   
 data = surv.pla.2)  
ggsurvplot(surv.fit.pla.2,   
 palette=c('black','orange','black','orange'),   
 linetype = c('solid', 'solid', 'dashed', 'dashed'),   
 ggtheme = theme\_cw(),   
 xlab='Months Since Emergence')

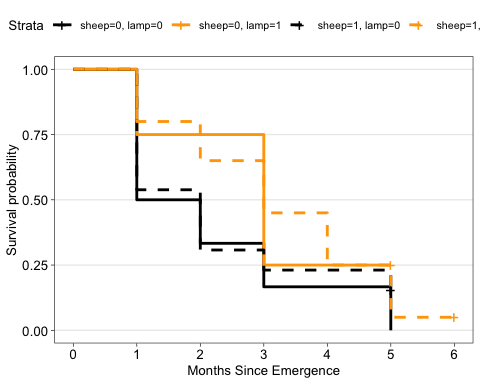


Same for dianthus

# this just makes the legend of the plot easier to see  
surv.dia.2 <- surv.dia  
surv.dia.2$sheep <- ifelse(surv.dia$grazing=='sheep', 1, 0)  
surv.dia.2$fert <- ifelse(surv.dia$nutrient=='fertilized', 1, 0)  
surv.dia.2$lamp <- ifelse(surv.dia$light=='lamps', 1, 0)  
  
surv.fit.dia <- survfit(surv.obj.dia ~ sheep + fert + lamp,   
 data = surv.dia.2)  
  
## apparently there were no dianthus seedlings in one of the conditions  
## (no sheep, fertilized, lamps)  
ggsurvplot(surv.fit.dia,   
 palette=c('black','orange','black',  
 'black','orange','black','orange'),   
 linetype = c('longdash', 'longdash','solid',   
 'dotted','dotted', 'dashed', 'dashed'),   
 ggtheme = theme\_cw(),   
 xlab='Months Since Emergence')



# just grazing and sheep  
# this makes it look like lamps may have had an effect  
# I think looking at a continuous spectrum of light could be a good option moving forward>  
surv.fit.dia.2 <- survfit(surv.obj.dia ~ sheep + lamp,   
 data = surv.dia.2)  
ggsurvplot(surv.fit.dia.2,   
 palette=c('black','orange','black','orange'),   
 linetype = c('solid', 'solid', 'dashed', 'dashed'),   
 ggtheme = theme\_cw(),   
 xlab='Months Since Emergence')



## Next steps

Might be worth taking another look at the survival data, especially if Mia can get it set up the way she was thinking to run it using the data structuring work I did. However, right now it looks like nothing is significant for the survival patterns, and similarly not much for emergence.

Survival plots of Dianthus make it look like light may have had a positive effect on survival in the beginning. I think looking at a continuous spectrum of light might be one way to look at this?