

checking Ben's BCI species selection estimates

```
rm(list=ls())
```

preamble

This checks Bens frequency independent selection estimates with linear regression

libraries

data

```
a<-readRDS("reproductives_names_stan.rds")
b<-readRDS("selection_reproductives.rds")
write.csv(b, "July5selectionestimates.csv")
#a<-fread("BCI_data_forben_reproductivesonly.csv")
z<-fread("generation_time_imputed.csv")
```

Sort out Ben's beaten data

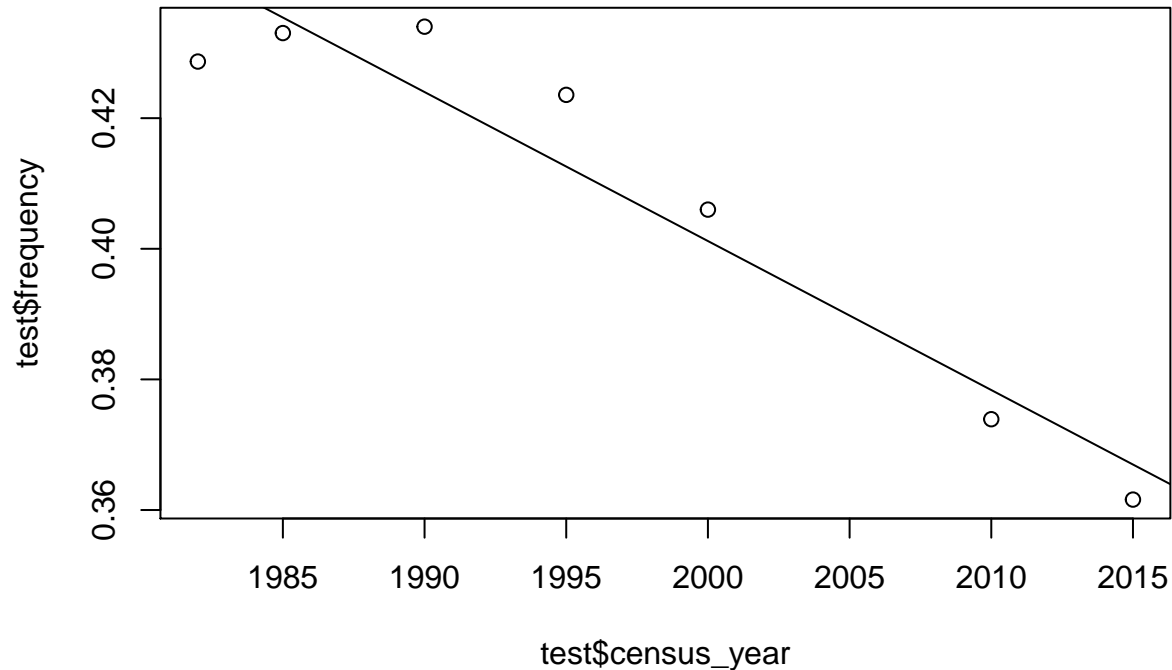
```
a1<-as.data.frame(a$counts)
years<-c(1982, 1985, 1990, 1995, 2000, 2010,2015)
names(a1)<-years
a1$species<-a$species_names
```

Melt and get frequencies

```
a2<-melt(a1, id.var=("species"))
names(a2)<-c("species", "census_year", "N")
a2<-a2%>%
  group_by(census_year)%>%
  mutate(N_year=sum(N))
a2$frequency<-a2$N/a2$N_year
a2<- a2[order(-a2$frequency),]
a2$census_year<-as.numeric(as.character(a2$census_year))
```

get a linear model for one species

```
test<-subset(a2, species=="Heisteria concinna")
test<- test[order(test$census_year),]
m<-lm(frequency~census_year, data=test)
plot(test$census_year, test$frequency)
abline(m)
```



linear model estimates for all species

```
m.all<-a2%>%
group_by(species) %>%
do(fit = lm(frequency ~ census_year, data = .))
m.all.2<-as.data.frame(tidy(m.all, fit))
head(m.all.2)
```

```
##           species      term      estimate  std.error  statistic
## 1  Abarema macradenia (Intercept) -7.047366e-04 2.775632e-04 -2.53901310
## 2  Abarema macradenia census_year  3.598229e-07 1.390077e-07  2.58851141
## 3  Acalypha diversifolia (Intercept) -6.236592e-03 1.989263e-01 -0.03135126
## 4  Acalypha diversifolia census_year  6.895815e-06 9.962520e-05  0.06921758
## 5  Acalypha macrostachya (Intercept) -2.001764e-03 9.588431e-03 -0.20876870
## 6  Acalypha macrostachya census_year  1.290284e-06 4.802025e-06  0.26869588
##           p.value
## 1 0.05195621
## 2 0.04892496
## 3 0.97620238
## 4 0.94749938
## 5 0.84286733
## 6 0.79889793
```

```
m.all.3<-subset(m.all.2, term=="census_year")
head(m.all.3)
```

```
##           species      term      estimate  std.error
## 2  Abarema macradenia census_year  3.598229e-07 1.390077e-07
## 4  Acalypha diversifolia census_year  6.895815e-06 9.962520e-05
## 6  Acalypha macrostachya census_year  1.290284e-06 4.802025e-06
## 8  Adelia triloba census_year -1.457146e-05 2.942773e-07
## 10 Aegiphila panamensis census_year -1.479791e-05 1.753940e-06
## 12 Alchornea costaricensis census_year  6.399350e-06 2.135826e-06
##           statistic      p.value
```

```
## 2    2.58851141 4.892496e-02
## 4    0.06921758 9.474994e-01
## 6    0.26869588 7.988979e-01
## 8   -49.51610140 6.348567e-08
## 10   -8.43695464 3.838230e-04
## 12    2.99619454 3.023118e-02
```

```
subset(m.all.3, species=="Heisteria concinna")
```

```
##           species      term  estimate  std.error statistic
## 214 Heisteria concinna census_year -0.00228145 0.0003233001 -7.056756
##           p.value
## 214 0.0008832835
```

```
coef(m)
```

```
## (Intercept) census_year
## 4.96408613 -0.00228145
```

the census_year term agrees with the single species linear regression

get selection estimates

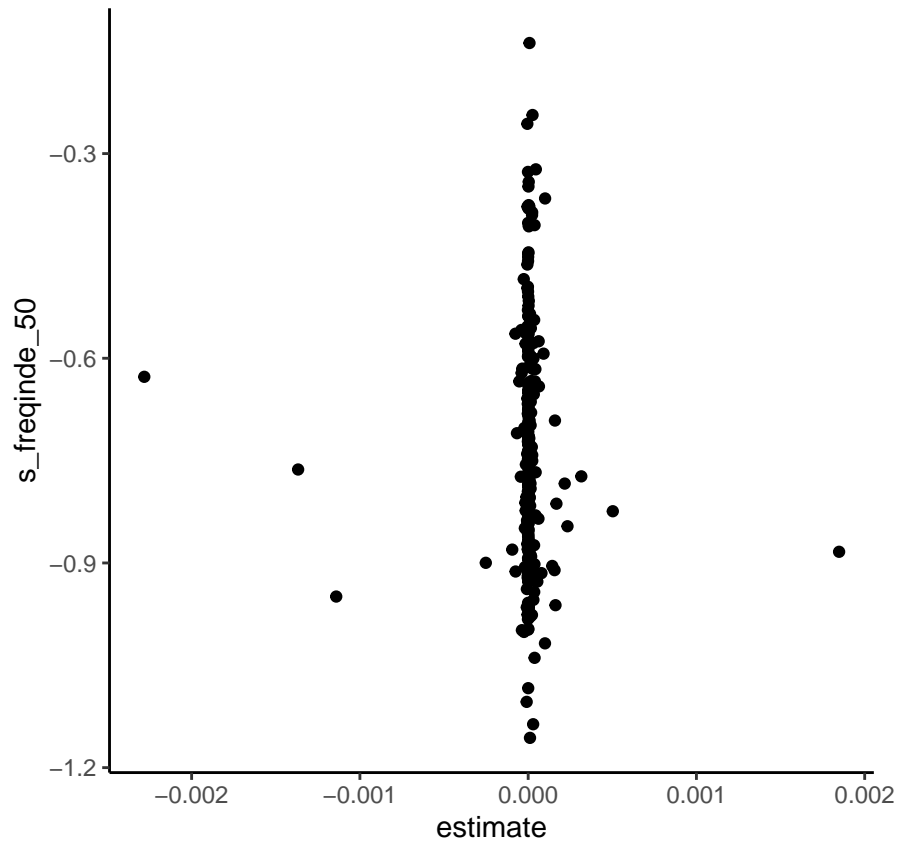
```
b<- b[order(-b$s_freqinde_50),]
```

get selection estimates and linear coefficients and merge

```
b2<-b[c("species", "s_freqinde_50")]
m.all.4<-m.all.3[c("species", "estimate")]
b3<-merge(b2, m.all.4, by="species")
```

plot

```
ggplot(b3, aes(x=estimate, y=s_freqinde_50))+
geom_point()+
theme_classic()+
theme(aspect.ratio=1)
```



That does not look good at all. Seem to confirm my impression that the selection estimates have gotten scrambled.