Assignment 3

Katie Beidler

February 1, 2015

1) Using the tree's dataset develop seperate models for the following species: "Acer rubrum", "Pinus strobus", "Abies fraseri." For each species which of the available explanatory variables seems to be most strongly correlated to the cover of that tree.

```
## subsetting the data--------------------------------
acer = subset(trees, subset = spcode == 'ACERRUB')
pinus = subset(trees, subset = spcode == 'PINUSTR')
aibes = subset(trees, subset = spcode == 'ABIEFRA')
## Making Models for Acer rubrum (red maple) ----------
# now to perform some simple linear regression models to determine which explanatory
variables are significantly related to cover (abundance of red maple trees)
acer_cover_v_elev = summary(lm(cover ~ elev, data = acer))
acer_cover_v_tci = summary(lm(cover ~ tci, data = acer))
acer_cover_v_streamdist = summary(lm (cover ~ streamdist, data = acer))
acer_cover_v_disturb = summary(lm(cover ~ disturb, data = acer))
acer_cover_v_beers = summary(lm(cover ~ beers, data = acer))
```

Of the explanatory variables: elevation, stream distance, disturbance and beers are significantly related to red maple tree cover- meaning that for each of these variables we reject the null that the slope is equal to 0. Beers seems to be most strongly related (P< 0.01) - even though only 1.3% of the variability in cover is explained by its linear relationship with beers.

```
# now to fit a multiple linear regression model using the significnat variables noted
above
acer_mod1 = lm (cover ~ elev + streamdist + disturb + beers, data = acer)
summary(acer_mod1)

##
## Call:
## lm(formula = cover ~ elev + streamdist + disturb + beers, data = acer)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -4.8108 -1.2356  0.3249  1.4053  5.1759
##
## Coefficients:
##                 Estimate Std. Error t value Pr(>|t|)
## (Intercept)    5.6280083  0.3549667  15.855  < 2e-16 ***
## elev          -0.0007187  0.0002952  -2.435  0.01513 *
## streamdist     0.0013420  0.0004481   2.995  0.00283 **
## disturbLT-SEL  0.1976155  0.2098562   0.942  0.34665
## disturbSETTLE  0.1855700  0.2608056   0.712  0.47697
## disturbVIRGIN  0.3656379  0.2419422   1.511  0.13112
## beers         -0.3033203  0.1042541  -2.909  0.00372 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 2.015 on 787 degrees of freedom
```

```
##   (1 observation deleted due to missingness)
## Multiple R-squared:  0.03412,	Adjusted R-squared:  0.02675 
## F-statistic: 4.633 on 6 and 787 DF,  p-value: 0.0001225
```

```r
# it seems as though disturbance has become less important with the inclusion of the other variables- in model 2 we will leave out disturbance
acer_mod2 = lm(cover ~ elev + streamdist + beers, data = acer)
summary(acer_mod2)
```

```
## 
## Call:
## lm(formula = cover ~ elev + streamdist + beers, data = acer)
## 
## Residuals:
##     Min      1Q  Median      3Q     Max 
## -4.9221 -1.3520  0.3404  1.4280  5.1992 
## 
## Coefficients:
##               Estimate Std. Error t value Pr(>|t|)    
## (Intercept)  5.8124275  0.2332516  24.919  < 2e-16 ***
## elev        -0.0007294  0.0002430  -3.001  0.00278 ** 
## streamdist   0.0014011  0.0004400   3.184  0.00151 ** 
## beers       -0.3087878  0.1025541  -3.011  0.00269 ** 
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
## 
## Residual standard error: 2.015 on 790 degrees of freedom
##   (1 observation deleted due to missingness)
## Multiple R-squared:  0.03124,	Adjusted R-squared:  0.02756 
## F-statistic: 8.492 on 3 and 790 DF,  p-value: 1.479e-05
```

```r
# the model now includes all signifcant terms and the adjusted R2 value went from 0.027 to 0.028- but does this make the model a better fit?
AIC(acer_mod1)
```

```
## [1] 3375.096
```

```r
AIC(acer_mod2)
```

```
## [1] 3371.457
```

The AIC is lower for model 2 and the AICs differ by more than 2-- meaning that model 2 explains more variability in red maple tree cover.

```r
# now to test for colinearity or interactions among the explanatory variables
acer_mod2 = lm(cover ~ elev + streamdist + beers, data = acer)
acer_mod_full = update(acer_mod2, ~ . + elev*streamdist*beers)
summary(acer_mod_full)
```

```
## 
## Call:
## lm(formula = cover ~ elev + streamdist + beers + elev:streamdist + 
##     elev:beers + streamdist:beers + elev:streamdist:beers, data = acer)
## 
## Residuals:
##     Min      1Q  Median      3Q     Max 
## -4.9816 -1.2736  0.2741  1.4115  5.2378 
## 
```

```
## Coefficients:
##                       Estimate Std. Error t value Pr(>|t|)    
## (Intercept)          4.143e+00  5.805e-01   7.137 2.17e-12 ***
## elev                 1.392e-03  7.171e-04   1.940 0.052677 .  
## streamdist           3.094e-03  2.124e-03   1.457 0.145568    
## beers                1.270e+00  4.714e-01   2.693 0.007224 ** 
## elev:streamdist     -2.516e-06  2.400e-06  -1.048 0.294870    
## elev:beers          -1.987e-03  5.623e-04  -3.534 0.000433 ***
## streamdist:beers    -1.109e-03  1.856e-03  -0.597 0.550439    
## elev:streamdist:beers  1.895e-06  1.939e-06   0.977 0.328732    
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
## 
## Residual standard error: 1.992 on 786 degrees of freedom
##   (1 observation deleted due to missingness)
## Multiple R-squared:  0.05754,    Adjusted R-squared:  0.04915 
## F-statistic: 6.856 on 7 and 786 DF,  p-value: 6.306e-08

# there seems to be a significant interaction between elevation and beers-- which makes sense given that the beers calculation depends on position on the mountain- better include this interaction in the model
acer_mod2_inter = lm(cover ~ elev + streamdist + beers + elev:beers,  data = acer)
summary(acer_mod2_inter)

## 
## Call:
## lm(formula = cover ~ elev + streamdist + beers + elev:beers, 
##     data = acer)
## 
## Residuals:
##    Min     1Q Median     3Q    Max 
## -4.988 -1.267  0.296  1.384  5.251 
## 
## Coefficients:
##               Estimate Std. Error t value Pr(>|t|)    
## (Intercept)  4.5695120  0.3599542  12.695  < 2e-16 ***
## elev         0.0007075  0.0003998   1.770 0.077166 .  
## streamdist   0.0015465  0.0004360   3.547 0.000412 ***
## beers        0.9492128  0.2976353   3.189 0.001483 ** 
## elev:beers  -0.0014565  0.0003240  -4.495 7.99e-06 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
## 
## Residual standard error: 1.99 on 789 degrees of freedom
##   (1 observation deleted due to missingness)
## Multiple R-squared:  0.05543,    Adjusted R-squared:  0.05064 
## F-statistic: 11.58 on 4 and 789 DF,  p-value: 3.88e-09

AIC(acer_mod2)

## [1] 3371.457

AIC(acer_mod2_inter) # this seems to be the best model for red maple

## [1] 3353.378

## Making Models for Pinus strobus (white pine) --------------------
pinus_cover_v_elev = summary(lm(cover ~ elev, data = pinus))
```

```r
pinus_cover_v_tci = summary(lm(cover ~ tci, data = pinus))
pinsu_cover_v_streamdist = summary(lm (cover ~ streamdist, data = pinus))
pinus_cover_v_disturb = summary(lm(cover ~ disturb, data = pinus))
pinus_cover_v_beers = summary(lm(cover ~ beers, data = pinus))
```

Of the explanatory variables: elevation, stream distance, disturbance and beers are significantly related to Fraser fir tree cover. Elevation, stream distance and beers seem to be most strongly related ($P< 0.001$) - elevation explains ~29%, stream distance explains ~11%, and beers explains 12% of the variation in fir tree cover.

```r
# now to fit a multiple linear regression model using the significnat variables noted above
pinus_mod1 = lm (cover ~ elev + streamdist + disturb, data = pinus)
summary(pinus_mod1)
```

```
##
## Call:
## lm(formula = cover ~ elev + streamdist + disturb, data = pinus)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -3.6264 -1.1994 -0.3109  1.0671  5.3344
##
## Coefficients:
##                 Estimate Std. Error t value Pr(>|t|)
## (Intercept)    5.1433047  0.4986161  10.315  < 2e-16 ***
## elev          -0.0027723  0.0005517  -5.025 8.14e-07 ***
## streamdist    -0.0012814  0.0005905  -2.170   0.0307 *
## disturbLT-SEL  0.7248595  0.3614152   2.006   0.0457 *
## disturbSETTLE  0.6862825  0.3863891   1.776   0.0766 .
## disturbVIRGIN -0.0947739  0.5611094  -0.169   0.8660
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1.737 on 342 degrees of freedom
##   (1 observation deleted due to missingness)
## Multiple R-squared:  0.1558, Adjusted R-squared:  0.1435
## F-statistic: 12.63 on 5 and 342 DF,  p-value: 2.884e-11
```

```r
# this model explains ~ 14% of the variability in cover, which is better than elevation alone.
# now to test for colinearity or interactions among the explanatory variables
pinus_mod1 = lm (cover ~ elev + streamdist + disturb, data = pinus)
pinus_mod_full = update(pinus_mod1, ~ . + elev*streamdist*disturb)
summary(pinus_mod_full)
```

```
##
## Call:
## lm(formula = cover ~ elev + streamdist + disturb + elev:streamdist +
##     elev:disturb + streamdist:disturb + elev:streamdist:disturb,
##     data = pinus)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -3.5972 -1.1653 -0.2424  1.0194  5.3316
##
## Coefficients:
##                                 Estimate Std. Error t value Pr(>|t|)
```

```
## (Intercept)                    4.137e+00  2.399e+00   1.724   0.0856 .
## elev                          -1.319e-03  4.114e-03  -0.321   0.7487
## streamdist                     1.742e-03  9.592e-03   0.182   0.8560
## disturbLT-SEL                  1.528e+00  2.540e+00   0.601   0.5479
## disturbSETTLE                  2.761e-01  2.848e+00   0.097   0.9228
## disturbVIRGIN                  2.884e+00  5.725e+00   0.504   0.6148
## elev:streamdist               -4.020e-06  1.590e-05  -0.253   0.8005
## elev:disturbLT-SEL            -1.066e-03  4.329e-03  -0.246   0.8057
## elev:disturbSETTLE             1.148e-03  4.843e-03   0.237   0.8128
## elev:disturbVIRGIN            -2.837e-03  6.418e-03  -0.442   0.6588
## streamdist:disturbLT-SEL      -1.801e-03  9.965e-03  -0.181   0.8567
## streamdist:disturbSETTLE       2.469e-02  1.753e-02   1.409   0.1598
## streamdist:disturbVIRGIN      -1.761e-02  1.874e-02  -0.939   0.3482
## elev:streamdist:disturbLT-SEL  1.909e-06  1.645e-05   0.116   0.9077
## elev:streamdist:disturbSETTLE -4.587e-05  3.065e-05  -1.496   0.1355
## elev:streamdist:disturbVIRGIN  1.568e-05  2.076e-05   0.756   0.4505
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1.747 on 332 degrees of freedom
##   (1 observation deleted due to missingness)
## Multiple R-squared:  0.1706, Adjusted R-squared:  0.1331
## F-statistic: 4.553 on 15 and 332 DF,  p-value: 6.775e-08
```

```
# no interactions- pinus_mod1 appears to be the best model
## Making Models for Abies fraseri (Fraser fir) -------------------
aibes_cover_v_elev = summary(lm(cover ~ elev, data = aibes))
aibes_cover_v_tci = summary(lm(cover ~ tci, data = aibes))
aibes_cover_v_streamdist = summary(lm (cover ~ streamdist, data = aibes))
aibes_cover_v_disturb = summary(lm(cover ~ disturb, data = aibes))
aibes_cover_v_beers = summary(lm(cover ~ beers, data = aibes))
```

Of the explanatory variables: elevation, stream distance and disturbance are significantly related to white pine tree cover. Elevation and stream distance seem to be most strongly related (P< 0.001) - elevation explains ~13% of the variation in tree cover and ~ 3.6% of the variability in cover is explained by stream distance.

```
# now to fit a multiple linear regression model using the significant variables noted above
aibes_mod1 = lm (cover ~ elev + streamdist + disturb + beers, data = aibes)
summary(aibes_mod1)
```

```
##
## Call:
## lm(formula = cover ~ elev + streamdist + disturb + beers, data = aibes)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -4.0960 -1.0512  0.0961  1.5607  3.4713
##
## Coefficients:
##                Estimate Std. Error t value Pr(>|t|)
## (Intercept)   -6.425600   2.033318  -3.160 0.002116 **
## elev           0.005290   0.001342   3.943 0.000154 ***
## streamdist     0.001243   0.001134   1.097 0.275616
## disturbLT-SEL  0.982458   1.049059   0.937 0.351383
## disturbSETTLE -0.450739   1.371354  -0.329 0.743120
## disturbVIRGIN  0.811883   0.615495   1.319 0.190315
```

```
## beers          0.624853   0.273522   2.284 0.024570 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1.778 on 95 degrees of freedom
## Multiple R-squared:  0.3688, Adjusted R-squared:  0.3289
## F-statistic: 9.251 on 6 and 95 DF,  p-value: 5.642e-08
```

```r
# it seems as though disturbance has become less important with the inclusion of the other variables- in model 2 we will leave out disturbance
aibes_mod2 = lm(cover ~ elev + streamdist + beers, data = aibes)
summary(aibes_mod2)
```

```
##
## Call:
## lm(formula = cover ~ elev + streamdist + beers, data = aibes)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -4.0699 -1.1728 -0.0034  1.4601  3.6200
##
## Coefficients:
##               Estimate Std. Error t value Pr(>|t|)
## (Intercept) -6.1613508  1.9875839  -3.100  0.00253 **
## elev         0.0055571  0.0013045   4.260  4.7e-05 ***
## streamdist   0.0009135  0.0011109   0.822  0.41288
## beers        0.7348161  0.2605144   2.821  0.00580 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1.775 on 98 degrees of freedom
## Multiple R-squared:  0.3512, Adjusted R-squared:  0.3313
## F-statistic: 17.68 on 3 and 98 DF,  p-value: 2.982e-09
```

```r
# now stream distance is no longer significant- let's leave it out and see what happens
aibes_mod3 = lm(cover ~ elev + beers, data = aibes)
summary(aibes_mod3)
```

```
##
## Call:
## lm(formula = cover ~ elev + beers, data = aibes)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -4.1397 -1.2127  0.0178  1.3664  3.6883
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept) -6.833995   1.808507  -3.779 0.000269 ***
## elev         0.006181   0.001059   5.835 6.79e-08 ***
## beers        0.705383   0.257622   2.738 0.007329 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1.772 on 99 degrees of freedom
## Multiple R-squared:  0.3467, Adjusted R-squared:  0.3335
## F-statistic: 26.27 on 2 and 99 DF,  p-value: 7.031e-10
```

```
AIC(aibes_mod1)

## [1] 415.6468

AIC(aibes_mod2)

## [1] 412.4506

AIC(aibes_mod3)

## [1] 411.1521

# both model 2 and 3 explain ~33% of the variability in tree cover- the AIC is lower for
model 3, but only differs from model 2 by 1-- it is hard to say which is better, but
models 2 & 3 are better than model 1.
```

2) From the tree data construct a new species richness variable wich summarizes how many unique species occur in each plot. Summarize this richness variable using the summary() function. Hint: the function tapply() could be helpful in this case.

```
summary(as.numeric(table((trees$plotID))))

##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   1.000   4.000   8.000   9.595  11.500  62.000

# on average there are ~9.5 trees in each plot
richness = tapply(trees$species, trees$plotID,function(x) length(unique(x)))
# this should tell us how many unique species are in each plot
summary(richness)

##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   1.000   4.000   8.000   7.913  11.000  21.000

# on average there are ~7.9 unique species in each plot
```

3) What kind of a variable is richness (continuous, discerte, categorical)?

Species richness is a discrete variable (counts).

4) For each of the unique plot id's extract the enviornmental information available for that plot

```
plot_levels = unique(trees$plotID)
# 935 levels or 935 unique plot IDs
dim(trees)

## [1] 8971   13

# 8971 rows- some plots are sampled more than once- as noted in the metadata
plot_duplicates_rm = subset(trees, !duplicated(plotID))
# now the duplicate plot IDs should be removed and the subsetted data frame should have
935 rows
dim(plot_duplicates_rm)

## [1] 935  13

#making a new data frame with the columns containing the env info and richness for each
plot
df <- data.frame(plotID=names(richness), richness =richness)
rownames(df) <- NULL
```

```r
env_info = plot_duplicates_rm[ ,cbind(1,7,8,9,10,11,12,13)]
rownames(env_info) <- NULL
env_info_plots <- merge(env_info, df, by="plotID")
```

5) **Construct a model of richness using the glm() function. Use the stepAIC() function to carry out a forward and also a backward stepwise selection of a best fitting model. Compare the results of this to the classic function step() Which model appears to be the best according to each approach? Why do you think this approach could be considered "dangerous" or potentially misleading?**

```r
rich_glm = glm(richness ~ elev + disturb + beers + elev*disturb*beers, data =
env_info_plots, family=poisson)
summary(rich_glm)

##
## Call:
## glm(formula = richness ~ elev + disturb + beers + elev * disturb *
##     beers, family = poisson, data = env_info_plots)
##
## Deviance Residuals:
##     Min       1Q   Median       3Q      Max
## -4.0285  -0.8098  -0.0411   0.6763   3.6054
##
## Coefficients:
##                               Estimate Std. Error z value Pr(>|z|)
## (Intercept)                  3.274e+00  1.558e-01  21.013  < 2e-16 ***
## elev                        -1.146e-03  1.408e-04  -8.142 3.88e-16 ***
## disturbLT-SEL               -3.955e-01  1.833e-01  -2.157 0.030993 *
## disturbSETTLE               -8.922e-01  2.227e-01  -4.006 6.17e-05 ***
## disturbVIRGIN                7.181e-01  2.586e-01   2.776 0.005500 **
## beers                       -3.000e-01  1.382e-01  -2.171 0.029942 *
## elev:disturbLT-SEL           4.372e-04  1.902e-04   2.298 0.021562 *
## elev:disturbSETTLE           9.104e-04  2.760e-04   3.298 0.000973 ***
## elev:disturbVIRGIN          -6.309e-04  2.190e-04  -2.881 0.003964 **
## elev:beers                   2.283e-04  1.244e-04   1.835 0.066471 .
## disturbLT-SEL:beers          3.264e-01  1.563e-01   2.088 0.036764 *
## disturbSETTLE:beers          1.550e-01  2.029e-01   0.764 0.444759
## disturbVIRGIN:beers          1.154e-01  2.160e-01   0.535 0.592963
## elev:disturbLT-SEL:beers    -3.364e-04  1.553e-04  -2.167 0.030263 *
## elev:disturbSETTLE:beers    -7.865e-05  2.563e-04  -0.307 0.758959
## elev:disturbVIRGIN:beers    -8.195e-05  1.828e-04  -0.448 0.653924
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for poisson family taken to be 1)
##
##     Null deviance: 2175.7  on 934  degrees of freedom
## Residual deviance: 1191.8  on 919  degrees of freedom
## AIC: 4737.4
##
## Number of Fisher Scoring iterations: 4

library(MASS)
stepAIC(rich_glm, direction = "both")

## Start:  AIC=4737.36
## richness ~ elev + disturb + beers + elev * disturb * beers
##
```

```
##                        Df Deviance    AIC
## - elev:disturb:beers  3   1197.5 4737.0
## <none>                    1191.8 4737.4
##
## Step:  AIC=4737.01
## richness ~ elev + disturb + beers + elev:disturb + elev:beers +
##     disturb:beers
##
##                        Df Deviance    AIC
## - disturb:beers       3   1198.1 4731.6
## - elev:beers          1   1198.2 4735.7
## <none>                    1197.5 4737.0
## + elev:disturb:beers  3   1191.8 4737.4
## - elev:disturb        3   1276.2 4809.8
##
## Step:  AIC=4731.61
## richness ~ elev + disturb + beers + elev:disturb + elev:beers
##
##                  Df Deviance    AIC
## - elev:beers     1   1199.3 4730.9
## <none>               1198.1 4731.6
## + disturb:beers  3   1197.5 4737.0
## - elev:disturb   3   1276.8 4804.4
##
## Step:  AIC=4730.87
## richness ~ elev + disturb + beers + elev:disturb
##
##                  Df Deviance    AIC
## <none>               1199.3 4730.9
## + elev:beers     1   1198.1 4731.6
## - beers          1   1206.1 4735.6
## + disturb:beers  3   1198.2 4735.7
## - elev:disturb   3   1277.9 4803.4
```

```
##
## Call:  glm(formula = richness ~ elev + disturb + beers + elev:disturb,
##     family = poisson, data = env_info_plots)
##
## Coefficients:
##        (Intercept)                 elev       disturbLT-SEL
##          3.027e+00           -9.289e-04          -6.183e-02
##      disturbSETTLE        disturbVIRGIN               beers
##         -7.234e-01            8.250e-01          -4.430e-02
## elev:disturbLT-SEL  elev:disturbSETTLE  elev:disturbVIRGIN
##          8.474e-05            8.101e-04          -6.993e-04
##
## Degrees of Freedom: 934 Total (i.e. Null);  926 Residual
## Null Deviance:       2176
## Residual Deviance: 1199  AIC: 4731
```

```
step(rich_glm)
```

```
## Start:  AIC=4737.36
## richness ~ elev + disturb + beers + elev * disturb * beers
##
##                        Df Deviance    AIC
```

```
## - elev:disturb:beers  3   1197.5 4737.0
## <none>                     1191.8 4737.4
##
## Step:  AIC=4737.01
## richness ~ elev + disturb + beers + elev:disturb + elev:beers +
##     disturb:beers
##
##                  Df Deviance    AIC
## - disturb:beers  3   1198.1 4731.6
## - elev:beers     1   1198.2 4735.7
## <none>                1197.5 4737.0
## - elev:disturb   3   1276.2 4809.8
##
## Step:  AIC=4731.61
## richness ~ elev + disturb + beers + elev:disturb + elev:beers
##
##                 Df Deviance    AIC
## - elev:beers    1   1199.3 4730.9
## <none>               1198.1 4731.6
## - elev:disturb  3   1276.8 4804.4
##
## Step:  AIC=4730.87
## richness ~ elev + disturb + beers + elev:disturb
##
##                 Df Deviance    AIC
## <none>               1199.3 4730.9
## - beers         1   1206.1 4735.6
## - elev:disturb  3   1277.9 4803.4

##
## Call:  glm(formula = richness ~ elev + disturb + beers + elev:disturb,
##     family = poisson, data = env_info_plots)
##
## Coefficients:
##       (Intercept)                elev       disturbLT-SEL
##         3.027e+00          -9.289e-04          -6.183e-02
##      disturbSETTLE       disturbVIRGIN               beers
##        -7.234e-01           8.250e-01          -4.430e-02
## elev:disturbLT-SEL  elev:disturbSETTLE  elev:disturbVIRGIN
##         8.474e-05           8.101e-04          -6.993e-04
##
## Degrees of Freedom: 934 Total (i.e. Null);  926 Residual
## Null Deviance:       2176
## Residual Deviance: 1199  AIC: 4731
```

According to both the the stepAIC() and step() approaches, the model: richness ~ elev + disturb + beers + elev:disturb, is the best- although, I am not sure how the approaches differ. However, choosing models using this approach could be considrered dangerous or misleading, as this approach may lead to overfitting. As the number of terms in the model increases, so does the proportion of variance explained by the model (R^2 gets closer to 1), but you may be modeling the random noise in your data set--diminishing the predictive power of your model. Multicollinearity can also be a problem, if predictor variables are correlated, it may be hard to say what is driving the effect on the response variable.