The presence of sprouted or diseased kernels in wheat can reduce the value of a wheat producers entire crop. It is important to identify these kernels after being harvested, prior to sale.
To this end, a study was conducted examining physical properties of a kernel: density, hardness, size, weight, and moisture content. Two different classes of wheat were considered, hard red winter (hrw) and soft red winter (srw). By visual inspection each kernel condition was classified as `Healthy, Sprout`, or `Scab`. The data is available in file `wheat.csv` on blackboard.

a) Fit a multinomial regression model to identify the properties affecting the kernel condition. Use `Healthy` as the base level for the response.

b) What predictors have a different effect on (all or some) kernel conditions?

c) Find CIs on each estimated coefficient.

d) Predict probabilities for each observed kernel condition.

e) Estimate odds ratios for a one standard deviation change in each predictor.

f) Find a CI on each odds ratio.

g) Fit a multinomial model with `density` as predictor.

h) Plot the probability curves from this multinomial model.

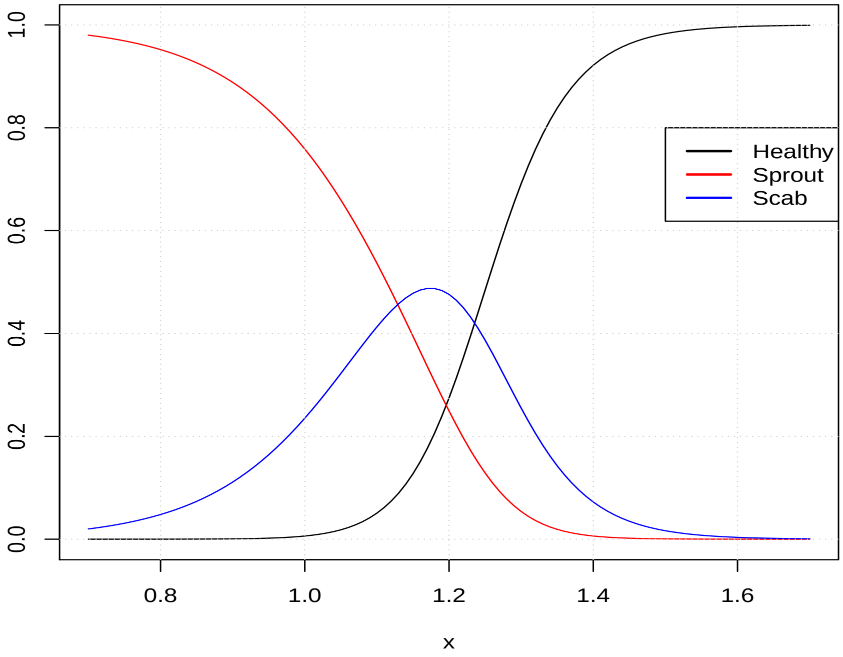
```
# labwheat.r

setwd("C:/Users/USC Guest/Downloads2")
wheat = read.csv("wheat.csv",header=T)
str(wheat)
#’data.frame’:   275 obs. of  7 variables:
# $ class   : Factor w/ 2 levels "hrw","srw": 1 1 1 1 1 1 1 1 1 1 ...
# $ density : num  1.35 1.29 1.23 1.34 1.26 ...
# $ hardness: num  60.3 56.1 44 53.8 44.4 ...
# $ size    : num  2.3 2.73 2.51 2.27 2.35 ...
# $ weight  : num  24.6 33.3 31.8 32.7 26.1 ...
# $ moisture: num  12 12.2 11.9 12.1 12.1 ...
# $ type    : Factor w/ 3 levels "Healthy","Scab",..: 1 1 1 1 1 1 1 1 1 1 ...

levels(wheat$type)  #Shows the 3 categories
# "Healthy" "Scab"    "Sprout"


# multinomial regression model
#==========================================================================

library(nnet)
m1 = multinom(type˜., wheat)
summary(m1)
# Coefficients:
#       (Intercept)   classsrw   density    hardness      size     weight    moisture
#Scab      30.54650 -0.6481277 -21.59715 -0.01590741 1.0691139 -0.2896482  0.10956505
#Sprout    19.16857 -0.2247384 -15.11667 -0.02102047 0.8756135 -0.0473169 -0.04299695

#Std. Errors:
#       (Intercept)  classsrw  density    hardness      size     weight  moisture
#Scab      4.289865 0.6630948 3.116174 0.010274587 0.7722862 0.06170252 0.1548407
#Sprout    3.767214 0.5009199 2.764306 0.008105748 0.5409317 0.03697493 0.1127188

#Residual Deviance: 384.2247
#AIC: 412.2247

# fitted equations

# log(pi-scab/pi-healthy) = 30.54650 -0.6481277 class -21.59715 density -0.01590741 hardness
                                     +1.0691139 size -0.2896482 weight + 0.10956505 moisture

# log(pi-sprout/pi-healthy) = 19.17 -0.2247384 class -15.11667 density -0.02102047 hardness
                                     + 0.8756135 size -0.0473169 weight -0.04299695 moisture
```

```
summary(m1,Wald=T)

#Coefficients:
#       (Intercept)   classsrw   density    hardness      size     weight    moisture
#Scab      30.54650 -0.6481277 -21.59715 -0.01590741 1.0691139 -0.2896482  0.10956505
#Sprout    19.16857 -0.2247384 -15.11667 -0.02102047 0.8756135 -0.0473169 -0.04299695

#Std. Errors:
#       (Intercept)  classsrw  density    hardness      size     weight  moisture
#Scab      4.289865 0.6630948 3.116174 0.010274587 0.7722862 0.06170252 0.1548407
#Sprout    3.767214 0.5009199 2.764306 0.008105748 0.5409317 0.03697493 0.1127188

#Value/SE (Wald statistics):
#       (Intercept)   classsrw   density  hardness     size    weight   moisture
#Scab      7.120620 -0.9774285 -6.930664 -1.548229 1.384349 -4.694269  0.7075983
#Sprout    5.088261 -0.4486513 -5.468523 -2.593279 1.618714 -1.279702 -0.3814532

#Residual Deviance: 384.2247
#AIC: 412.2247

# but no p-values shown, so try this way




# tests
sum.fit  = summary(m1)
test.stat = sum.fit$coefficients/sum.fit$standard.errors
p.value  = 2*(1-pnorm(q = abs(test.stat)))
test.stat
#       (Intercept)   classsrw   density  hardness     size    weight   moisture
#Scab      7.120620 -0.9774285 -6.930664 -1.548229 1.384349 -4.694269  0.7075983
#Sprout    5.088261 -0.4486513 -5.468523 -2.593279 1.618714 -1.279702 -0.3814532

p.value
#        (Intercept)  classsrw      density    hardness      size       weight  moisture
#Scab   1.074474e-12 0.3283570 4.188649e-12 0.121567269 0.1662515 2.675618e-06 0.4791947
#Sprout 3.613623e-07 0.6536832 4.538002e-08 0.009506554 0.1055089 2.006500e-01 0.7028670
round(p.value,3)
#       (Intercept) classsrw density hardness   size weight moisture
#Scab             0   0.3284       0   0.1216 0.1663 0.0000   0.4792
#Sprout           0   0.6537       0   0.0095 0.1055 0.2006   0.7029

# There is no evidence that wheat class, size, and moisture have different effects on kernel conditio
# There is evidence that hardness has some effect on kernel Sprout only
# There is evidence that weight has some effect on kernel Scab only
# Effects across all kernel conditions

library(car)
Anova(m1)

#Analysis of Deviance Table (Type II tests)
#Response: type
#         LR Chisq Df Pr(>Chisq)
#class       0.964  2     0.6175
#density    90.555  2  < 2.2e-16 ***
#hardness    7.074  2     0.0291 *
#size        3.211  2     0.2008
#weight     28.230  2  7.411e-07 ***
#moisture    1.193  2     0.5506

# density, hardness and weight have some effect on wheat kernel condition

# CIs on betas
conf.beta<-confint(m1)
# , , Scab
#                   2.5 %       97.5 %
#(Intercept)  22.13851497  38.954475222
#classsrw     -1.94776958   0.651514098
#density     -27.70474380 -15.489565975
#hardness     -0.03604523   0.004230411
#size         -0.44453927   2.582767006
#weight       -0.41058295  -0.168713512
#moisture     -0.19391723   0.413047326

#, , Sprout
#                   2.5 %      97.5 %
#(Intercept)  11.78496433 26.552173165
#classsrw     -1.20652328  0.757046542
#density     -20.53461137 -9.698731394
#hardness     -0.03690744 -0.005133494
#size         -0.18459306  1.935820104
#weight       -0.11978643  0.025152642
#moisture     -0.26392179  0.177927888
```

```
# predict probabilities

pi.hat = predict(m1, newdata = wheat, type = "probs")
head(pi.hat)
#     Healthy        Scab     Sprout
#1 0.8552110 0.046396827 0.09839221
#2 0.7492553 0.021572158 0.22917255
#3 0.5172800 0.068979903 0.41374011
#4 0.8982064 0.006740716 0.09505287
#5 0.5103245 0.176260796 0.31341473
#6 0.7924907 0.015304122 0.19220522


# Odds ratios for a c=1 unit sdev increase in each predictor
#================================================================================

summary(wheat)
# class        density          hardness           size            weight          moisture
# hrw:143   Min.   :0.7352   Min.   :-44.080   Min.   :0.5973   Min.   : 8.532   Min.   : 6.486   Hea
# srw:132   1st Qu.:1.1358   1st Qu.:  0.689   1st Qu.:1.8900   1st Qu.:21.982   1st Qu.: 9.540   Sca
#           Median :1.2126   Median : 24.465   Median :2.2303   Median :27.610   Median :11.909   Spr
#           Mean   :1.1885   Mean   : 25.564   Mean   :2.2047   Mean   :27.501   Mean   :11.192
#           3rd Qu.:1.2687   3rd Qu.: 45.606   3rd Qu.:2.5125   3rd Qu.:32.882   3rd Qu.:12.538
#           Max.   :1.6454   Max.   :111.934   Max.   :4.3100   Max.   :46.334   Max.   :14.514

sd.wheat = apply(wheat[,-c(1,7)],2,sd)
sd.wheat
#   density   hardness       size     weight   moisture
# 0.1313021 27.3561563  0.4906125  7.9154398  2.0332132

# coeffs
beta.hat2<-coefficients(m1)[1,2:7]
beta.hat2
#    classsrw      density     hardness         size       weight     moisture
# -0.64812774 -21.59715489  -0.01590741   1.06911387  -0.28964823   0.10956505

beta.hat3<-coefficients(m1)[2,2:7]
beta.hat3
#    classsrw      density     hardness         size       weight     moisture
# -0.22473837 -15.11667138  -0.02102047   0.87561352  -0.04731690  -0.04299695

# add column class
c.value = c(class=1,sd.wheat)
round(c.value,3)
#   class  density hardness     size   weight moisture
#   1.000    0.131   27.356    0.491    7.915    2.033
# Odds ratios (scab vs. healthy)
round(exp(c.value*beta.hat2),3)
#   class  density hardness     size   weight moisture
#   0.523    0.059    0.647    1.690    0.101    1.250
round(1/exp(c.value*beta.hat2),3)
#   class  density hardness     size   weight moisture
#   1.912   17.043    1.545    0.592    9.902    0.800

# scab vs. healthy
# Odds change by 0.059 for a 0.13 increase in density, holding other vars constant
# Odds change by 17.04 for a 0.13 decrease in density, holding other vars constant
# Odds change by  9.90 for a 7.92 decrease in weight, holding other vars constant

# Odds ratios (sprout vs. healthy)
round(exp(c.value*beta.hat3),3)
#   class  density hardness     size   weight moisture
#   0.799    0.137    0.563    1.537    0.688    0.916
round(1/exp(c.value*beta.hat3),3)
#   class  density hardness     size   weight moisture
#   1.252    7.278    1.777    0.651    1.454    1.091

# sprout vs. healthy
# Odds change by  7.28 for a 0.13 decrease in density, holding other vars constant
# Odds change by  1.45 for a 7.92 decrease in weight, holding other vars constant


# CIs for OR
ci.OR2<-exp(c.value*conf.beta[2:7,1:2,1])
round(ci.OR2,4)
#          2.5 % 97.5 %
#classsrw 0.1426 1.9184
#density  0.0263 0.1308
#hardness 0.3730 1.1227
#size     0.8040 3.5507
#weight   0.0388 0.2630
#moisture 0.6742 2.3159

ci.OR3<-exp(c.value*conf.beta[2:7,1:2,2])
round(ci.OR3,4)
#          2.5 % 97.5 %
#classsrw 0.2992 2.1320
#density  0.0675 0.2799
#hardness 0.3643 0.8690
#size     0.9134 2.5850
#weight   0.3875 1.2203
#moisture 0.5847 1.4359
```

```
# model with density
#================================================================================

m2 = multinom(type ~ density,wheat)
summary(m2)
# Coefficients:
#       (Intercept)   density
#Scab      29.37827 -24.56215
#Sprout    19.12165 -15.47633

#Std. Errors:
#       (Intercept)  density
#Scab      3.676892 3.017842
#Sprout    3.337092 2.691429

#Residual Deviance: 459.4246
#AIC: 467.4246

beta.hat = coefficients(m2)
beta.hat
#       (Intercept)   density
#Scab      29.37827 -24.56215
#Sprout    19.12165 -15.47633

# predict probabilities

pi.hat = predict(m2, newdata = wheat, type = "probs")
head(pi.hat)
#     Healthy       Scab    Sprout
#1 0.8366072 0.01943493 0.1439578
#2 0.6435285 0.06823514 0.2882363
#3 0.4134757 0.16296658 0.4235578
#4 0.8056325 0.02557888 0.1687886
#5 0.5240738 0.11162920 0.3642970
#6 0.6921854 0.05357109 0.2542435

b11 = beta.hat[1,1]
b12 = beta.hat[1,2]
b21 = beta.hat[2,1]
b22 = beta.hat[2,2]
```

```
# plot

f1=function(x){1/(1 + exp(b11 + b12*x) + exp(b21 + b22*x))}
f2=function(x){exp(b11 + b12*x)/(1 + exp(b11 + b12*x) + exp(b21 + b22*x))}
f3=function(x){exp(b21 + b22*x)/(1 + exp(b11 + b12*x) + exp(b21 + b22*x))}
curve(f1,0.7,1.7,ylab="")
curve(f2,0.7,1.7,col="red",add=T)
curve(f3,0.7,1.7,col="blue",add=T)
colors = c("black","red","blue")
labels = c("Healthy", "Sprout", "Scab")
legend(x=1.5,y=0.8,legend=labels,col=colors,lwd = c(2,2,2))
grid()
```


x
Healthy
Sprout
Scab