Problem 6.2

How to read an external data file into R:

```
> shoot=read.table("http://educ.jmu.edu/~chen3lx/math321/
                    shoot.txt",header=TRUE)
```

use read.table to read a text file (columns separated by space) and use read.csv to read a csv file (columns separated by comma). The part inside " " gives the directory of the data file, header=TRUE if the first line gives the names of variables. header=TRUE is not needed if first line starts with data values.

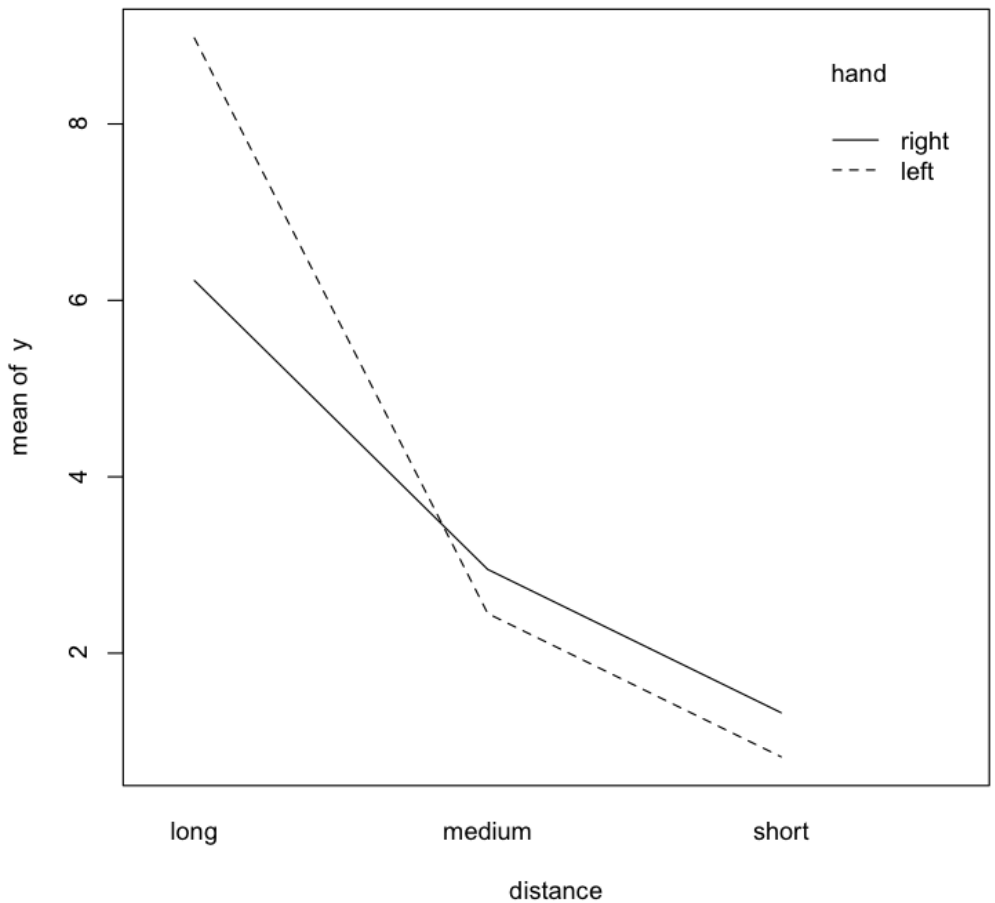
y	distance	hand
0.00	short	left
1.5	short	left
0.0	short	left
0.625	short	left
2.00	short	left
3.375	short	right
0.375	short	right
2.125	short	right
0.250	short	right
0.50	short	right
3.50	medium	left
3.25	medium	left
0.125	medium	left
3.25	medium	left
2.125	medium	left
1.00	medium	right
4.875	medium	right
1.00	medium	right
3.25	medium	right
4.625	medium	right
13.25	long	left
7.00	long	left
8.125	long	left
7.75	long	left
8.75	long	left
3.125	long	right
1.125	long	right
14.375	long	right

Interaction plot

```
> interaction.plot(shoot$distance, shoot$hand, shoot$y)
```

or

```
> with(shoot, interaction.plot(distance,hand,y))
```



Call a variable in a data set

You can skip the factor step in R if the labels for a factor are words as in the shoot data set. If you want to use factor function, use datafile$variable to refer to a variable in a data set.

```
shoot$hand=factor(shoot$factor)
shoot$distance=factor(shoot$distance)
```

ANOVA table

```
> shoot=read.table("http://educ.jmu.edu/~chen3lx/math321/
                    shoot.txt",header=TRUE)
> output=aov(y~distance*hand,data=shoot)
> summary(output)
              Df Sum Sq Mean Sq F value  Pr(>F)
distance       2 230.75  115.38  15.738 4.3e-05 ***
hand           1   2.55    2.55   0.348   0.561
distance:hand  2  17.60    8.80   1.201   0.318
Residuals     24 175.94    7.33
```

Drop the interaction term in the model since it is insignificant:

```
> output2=aov(y~distance+hand,data=shoot)
> summary(output2)
             Df Sum Sq Mean Sq F value  Pr(>F)
distance      2 230.75  115.38  15.499 3.7e-05 ***
hand          1   2.55    2.55   0.343   0.563
Residuals    26 193.55    7.44
```

Fit a simpler model

We can further drop the hand factor from the model:

```
> output3=aov(y~distance,data=shoot)
> summary(output3)
             Df Sum Sq Mean Sq F value   Pr(>F)
distance      2  230.8  115.38   15.89 2.75e-05 ***
Residuals    27  196.1    7.26

> TukeyHSD(output3)
  Tukey multiple comparisons of means
    95% family-wise confidence level

Fit: aov(formula = y ~ distance, data = shoot)
$distance
                 diff       lwr       upr     p adj
medium-long  -4.900 -7.888282 -1.911718 0.0010508
short-long   -6.525 -9.513282 -3.536718 0.0000292
short-medium -1.625 -4.613282  1.363282 0.3816362
```

Prob 6.4

The anova table shows there is significant interaction effect at .10 significance level.

```
> butter = read.table("/Users/lchen/Sites/math321/
                butter.txt",header=TRUE)
> butter$brand=factor(butter$brand)
> butter$cookmethod=factor(butter$cookmethod)
> output = aov(y~brand*cookmethod,data=butter)
> anova(output)

Analysis of Variance Table

Response: y
                  Df  Sum Sq Mean Sq F value    Pr(>F)
brand              2  2683.0  1341.5  6.0931   0.01492 *
cookmethod         1 11806.7 11806.7 53.6263 9.185e-06 ***
brand:cookmethod   2  1470.8   735.4  3.3401   0.07027 .
Residuals         12  2642.0   220.2
```

Study factor effect with significant interactions

When the interaction effect is significant, we need to study the effect of one factor at the fixed level(s) of another factor.
e.g., compare the effect of *cookmethod* (stove vs oven) fixing factor *brand* at level *lakes*.

or compare the effect of *brand* (lakes vs value, lakes vs cabot, value vs cabot) fixing *cookmethod at level stove*.

	stove	oven
lakes	173,125,158	166,179,203
cabot	97,110, 126	177,197,183
value	150,154,157	206,195,205

Butter example

Get all the treatment means as below:

	stove	oven
lakes	152.00	182.67
cabot	111.00	185.67
value	153.67	202.00

e.g., compare oven vs stove at brand level $=$ lakes:
$182.67 - 152.00 \pm 2.179 * \sqrt{220.2} * \sqrt{\frac{1}{3} + \frac{1}{3}} = 30.67 \pm 26.40 =$
$(4.27, 57.07)$.
$R : qt(0.975, 12)$ note $m = 1$ here as we make only one comparison so there is no need for adjustment.

Compare the effect of the factor brand

Compare the effect of brand fixing cookmethod = stove.
$t = 3$
critical value in the CI: 2.668

```
> qtukey(0.95,3,12)/sqrt(2)
[1] 2.667864
```

margin of error $= 2.668 * \sqrt{220.2} * \sqrt{\frac{1}{3} + \frac{1}{3}} = 32.32.$

lakes - cabot: $152.00 - 111.00 \pm 32.32 = (8.68, 73.32)$

lakes - value: $152 - 153.67 \pm 32.32 = (-33.99, 30.65).$

cabot - value: $111.00 - 153.67 \pm 32.32 = (-74.99, -10.35).$

Pairwise comparisons with significant interacitons

```
> install.packages("emmeans")
> library(emmeans)
> output=lm(y~brand*cookmethod,data=butter)
> confint(emmeans(output,pairwise~brand|cookmethod),
          level=0.95)
```

brand|cookmethod compares the three brands at the fixed level of cookmethod, while cookmethod| brand compares the two cookmethods at the fixed level of brand. Next pages shows pairwise comparisons of brands fixing cookmethod.

```
> confint(emmeans(output,pairwise~cookmethod|brand),
           level=0.95)
```

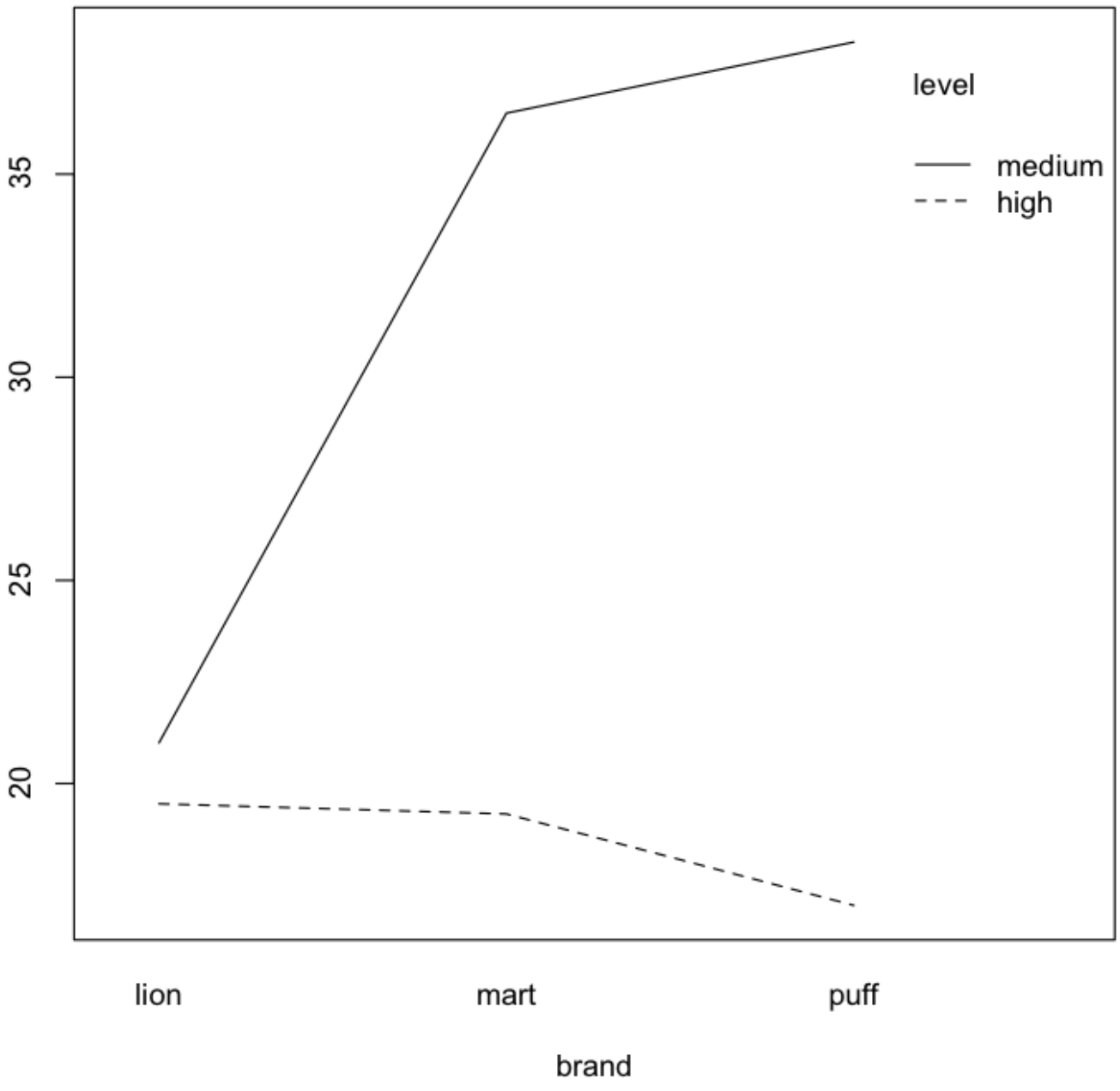
Tukey CIs

```
cookmethod = oven:
 contrast         estimate       SE df  lower.CL   upper.CL
 cabot - lakes    3.000000 12.11519 12 -29.32167  35.321669
 cabot - value -16.333333 12.11519 12 -48.65500  15.988336
 lakes - value -19.333333 12.11519 12 -51.65500  12.988336

cookmethod = stove:
 contrast         estimate       SE df  lower.CL   upper.CL
 cabot - lakes -41.000000 12.11519 12 -73.32167  -8.678331
 cabot - value -42.666667 12.11519 12 -74.98834 -10.344997
 lakes - value  -1.666667 12.11519 12 -33.98834  30.655003

Confidence level used: 0.95
Conf-level adjustment: tukey method for comparing a family
of 3 estimates
```

Time Order	Brand	Level	Amount of Time(seconds)
2	Food Lion	Medium	16
8			37
14			15
19			16
1	Food Lion	High	19
11			18
18			18
23			23
3	Jet Puff	Medium	39
10			38
17			39
20			37
6	Jet Puff	High	16
9			17
15			18
21			17
4	WalMart	Medium	15
12			44
16			44
22			43
5	WalMart	High	16
7			19
13			22
24			20


level
medium
high
35
30
25
20
lion
mart
puff
brand