EXERCISE P.4

(a) The sum of squared deviations about the mean is usually more easily computed using this result.

$$\sum_{i=1}^{n}\left(x_i-\bar{x}\right)^2=\sum_{i=1}^{n}\left(x_i^2+\bar{x}^2-2x_i\bar{x}\right)=\left(\sum_{i=1}^{n}x_i^2\right)+\sum_{i=1}^{n}\left(\bar{x}^2\right)+\sum_{i=1}^{n}\left(-2x_i\bar{x}\right)=\left(\sum_{i=1}^{n}x_i^2\right)+n\bar{x}^2-2\bar{x}\sum_{i=1}^{n}x_i$$
$$=\left(\sum_{i=1}^{n}x_i^2\right)+n\bar{x}^2-2\bar{x}\left(\sum_{i=1}^{n}x_i\right)\frac{n}{n}=\left(\sum_{i=1}^{n}x_i^2\right)+n\bar{x}^2-2n\bar{x}^2=\left(\sum_{i=1}^{n}x_i^2\right)-n\bar{x}^2$$

(b) The sum of the crossproduct of, each variable minus its sample mean, is usually more easily computed using this result.

$$\sum_{i=1}^{n}\left(x_i-\bar{x}\right)\left(y_i-\bar{y}\right)=\sum_{i=1}^{n}\left(x_iy_i-x_i\bar{y}-\bar{x}y_i+\bar{x}\bar{y}\right)=\left(\sum_{i=1}^{n}x_iy_i\right)-\bar{y}\sum_{i=1}^{n}x_i-\bar{x}\sum_{i=1}^{n}y_i+n\bar{x}\bar{y}$$
$$=\left(\sum_{i=1}^{n}x_iy_i\right)-\bar{y}\sum_{i=1}^{n}x_i\left(\frac{n}{n}\right)-\bar{x}\sum_{i=1}^{n}y_i\left(\frac{n}{n}\right)+n\bar{x}\bar{y}=\left(\sum_{i=1}^{n}x_iy_i\right)-n\bar{x}\bar{y}-n\bar{x}\bar{y}+n\bar{x}\bar{y}$$
$$=\left(\sum_{i=1}^{n}x_iy_i\right)-n\bar{x}\bar{y}$$

(c) The sum of a variable minus its arithmetic mean (average) is always zero.

$$\sum\nolimits_{j=1}^{n}\left(x_j-\bar{x}\right)=\sum_{j=1}^{n}x_j-\sum_{j=1}^{n}\bar{x}=\sum_{j=1}^{n}x_j-n\bar{x}=\sum_{j=1}^{n}x_j-n\frac{\sum_{j=1}^{n}x_j}{n}=\sum_{j=1}^{n}x_j-\sum_{j=1}^{n}x_j=0$$

EXERCISE P.9

(a) The marginal distributions are

Political Party	Probability
Republican	0.45
Independent	0.15
Democrat	0.40

War Attitude	Probability
against	0.45
neutral	0.25
in favor	0.30

(b) The conditional probability is

$$P(INDEPENDENT \mid IN\ FAVOR) = \frac{P(INDEPENDENT \text{ and } IN\ FAVOR)}{P(IN\ FAVOR)} = \frac{0.05}{0.30} = 0.167$$

(c) They are not independent. For example

$$P(DEMOCRAT \text{ and } IN\ FAVOR) = 0 \neq P(DEMOCRAT) \times P(IN\ FAVOR) = 0.40 \times 0.30 = 0.12$$

(d) The probability distribution of *WAR* is

WAR	1	2	3
$f(WAR)$	0.45	0.25	0.30

The expected value of *WAR* is

$$E(WAR) = \sum_{war=1}^{3} war \times f(war) = 1(0.45) + 2(0.25) + 3(0.30) = 0.45 + 0.50 + 0.90 = 1.85$$

The expectation of WAR implies that on this scale the U.S. Population has an attitude between Against and Neutral.

To find the variance first calculate

$$E(WAR^2) = \sum_{war=1}^{3} war^2 \times f(war) = 1^2(0.45) + 2^2(0.25) + 3^2(0.30) = 0.45 + 1.00 + 2.70 = 4.15$$

Then

$$\operatorname{var}(WAR) = E(WAR^2) - [E(WAR)]^2 = 4.15 - (1.85)^2 = 4.15 - 3.4225 = 0.7275$$

Exercise P.9 (continued)

(e) The relation $CONTRIBUTIONS = 10 + 2 \times WAR$ implies the expected value

$$E(CONTRIBUTIONS) = E(10 + 2 \times WAR) = 10 + 2 \times E(WAR) = 10 + 2(1.85) = 13.7$$

The variance of *CONTRIBUTIONS* is

$$\text{var}(CONTRIBUTIONS) = \text{var}(10 + 2 \times WAR) = 2^2 \times \text{var}(WAR) = 2.91$$

Then,

$$\text{standard deviation}(CONTRIBUTIONS) = \sqrt{\text{var}(CONTRIBUTIONS)} = \sqrt{2.91} = 1.71$$

EXERCISE P.12

(a)

GRADE	4	3	2	1	0
f(*grade*)	0.13	0.22	0.35	0.20	0.10

(b)

$$E(GRADE) = \sum_{grade} grade \times f(grade) = 4(0.13) + 3(0.22) + 2(0.35) + 1(0.20) + 0(0.10)$$
$$= 0.52 + 0.66 + 0.70 + 0.20 = 2.08$$

$$E(GRADE^2) = \sum_{grade} grade^2 \times f(grade) = 4^2(0.13) + 3^2(0.22) + 2^2(0.35) + 1^2(0.20) + 0^2(0.10)$$
$$= 2.08 + 1.98 + 1.40 + 0.20 = 5.66$$

$$\text{var}(GRADE) = E(GRADE^2) - \left[E(GRADE)\right]^2 = 5.66 - (2.08)^2 = 1.3336$$

(c)

$$E(CLASS\_AVG) = E\left(\sum_{i=1}^{300} GRADE_i \Big/ 300\right) = (1/300)E\left(\sum_{i=1}^{300} GRADE_i\right) = (1/300)300E(GRADE)$$

$$E(CLASS\_AVG) = E(GRADE) = 2.08$$

$$\text{var}(CLASS\_AVG) = \text{var}\left(\sum_{i=1}^{300} GRADE_i \Big/ 300\right) = \text{var}\left((1/300)\sum_{i=1}^{300} GRADE_i\right) = (1/300)^2 \text{var}\left(\sum_{i=1}^{300} GRADE_i\right)$$

Since the *GRADE*'s are independent, the variance of their sum is the sum of their variances:

$$\text{var}(CLASS\_AVG) = (1/300)^2 (300 \text{var}(GRADE)) = (1/300)\text{var}(GRADE) = 0.0044$$

(d) The probability that, on a given Monday, more than 3 students are absent is

$$E(MAJORS) = E(50 + 10CLASS\_AVG) = 50 + 10E(CLASS\_AVG) = 50 + 10(2.08) = 70.08$$
$$\text{var}(MAJORS) = \text{var}(50 + 10CLASS\_AVG) = 10^2 \text{var}(CLASS\_AVG) = 100\text{var}(CLASS\_AVG) = 0.44$$

library(devtools)

library(PoEdata)

data(food)

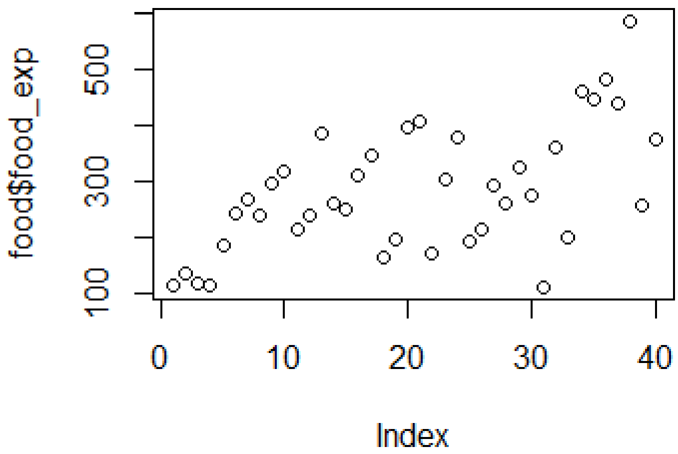
#b

#i

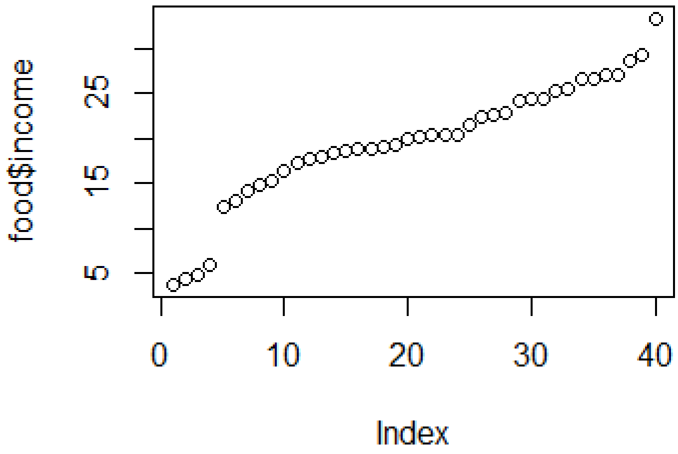
```
> head(food)
  food_exp income
1   115.22   3.69
2   135.98   4.39
3   119.34   4.75
4   114.96   6.03
5   187.05  12.47
6   243.92  12.98
> tail(food)
   food_exp income
35   447.76  26.70
36   482.55  27.14
37   438.29  27.16
38   587.66  28.62
39   257.95  29.40
40   375.73  33.40
```

#ii

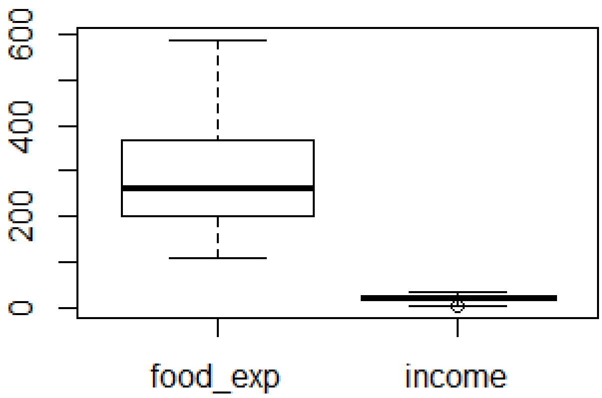
plot(food$food_exp)



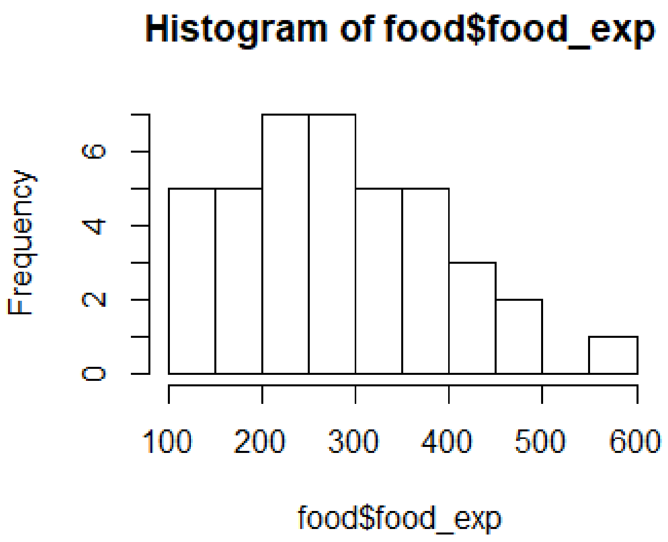
```
plot(food$income)
```

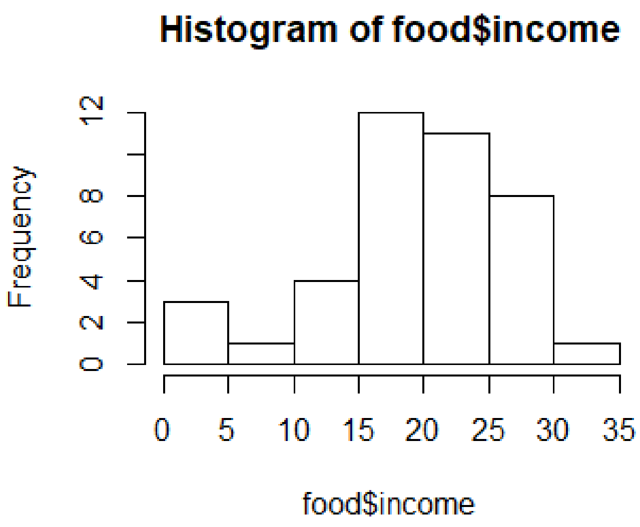



```
#iii
boxplot(food)
```

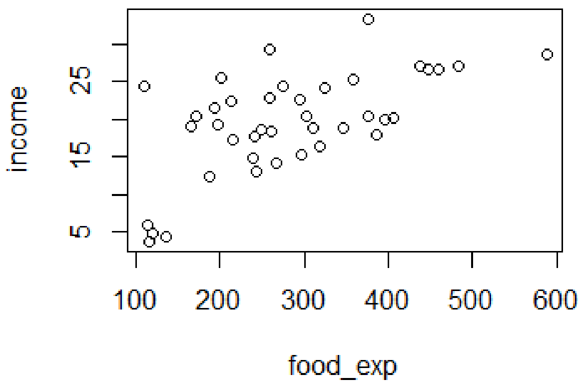


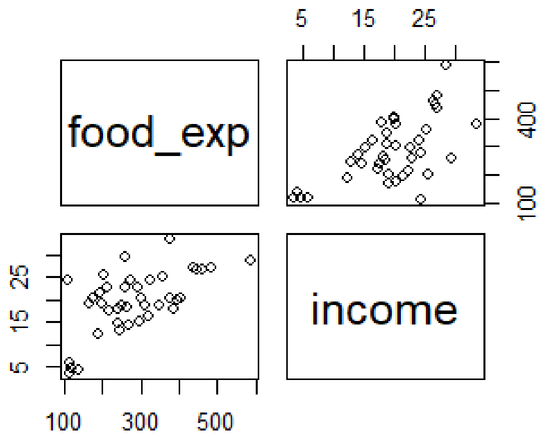
```
#iv
hist(food$food_exp)
hist(food$income)
```





#v
plot(food)
pairs(food)






```
> #i
> p=pbinom(q=40, size=100, prob=0.25)
> p
[1] 0.999676
> #ii
> p2=pnorm(40, mean=100*0.25, sd=sqrt(100*0.25*.75/100))
> p2
[1] 1
```