Joshua Main-Smith
P 5.2
joshuamainsmith@ufl.edu
UFID: 7160-7655
April 12, 2020

12.10 The following data are given:

x	0	1	2	3	4	5	6
y	1	4	5	3	2	3	4

(a) Fit the cubic model $\mu_{Y|x} = \beta_0 + \beta_1 x + \beta_2 x^2 + \beta_3 x^3$.
(b) Predict Y when $x = 2$.

(a)

```
> x <- c(0, 1, 2, 3, 4, 5, 6)
> y <- c(1, 4, 5, 3, 2, 3, 4)
>
> tmp <- poly(x, 3, raw = TRUE)
>
> mod <- lm(y ~ tmp)
>
> mod

Call:
lm(formula = y ~ tmp)

Coefficients:
(Intercept)         tmp1         tmp2         tmp3
     1.0714       4.6032      -1.8452       0.1944
```

So, $\mu Y|x = 1.0714 + 4.6032x – 1.8452x^2 + 0.1944x^3$

(b)

```
> tmp <- 2
> predict(lm(y~poly(x, 3)), tmp)
       1        2        3        4        5        6        7
1.071429 4.023810 4.452381 3.523810 2.404762 2.261905 4.261905
```

So, when x = 2, Y = 4.452381

12.25 Using the data of Exercise 12.2 on page 450 and the estimate of σ^2 from Exercise 12.17, compute 95% confidence intervals for the predicted response and the mean response when $x_1 = 900$ and $x_2 = 1.00$.

```
Prediction response:
> y <- c(0.231,
+        0.107,
+        0.053,
+        0.129,
+        0.069,
```

```
+         0.030,
+         1.005,
+         0.559,
+         0.321,
+         2.948,
+         1.633,
+         0.934)
>
> x1 <- c(740,
+          740,
+          740,
+          805,
+          805,
+          805,
+          980,
+          980,
+          980,
+          1235,
+          1235,
+          1235)
>
> x2 <- c(1.10,
+          0.62,
+          0.31,
+          1.10,
+          0.62,
+          0.31,
+          1.10,
+          0.62,
+          0.31,
+          1.10,
+          0.62,
+          0.31)
>
>
>
> a <- predict(lm(y~poly(x1, 2, raw=TRUE)), tmp, interval="prediction", level
=0.95)
> b <- predict(lm(y~poly(x2, 2, raw=TRUE)), tmp, interval="prediction", level
=0.95)
> predict(lm(a~b), tmp, interval="prediction", level=0.95)
       fit        lwr      upr
1  0.66825 -0.6329387 1.969439
2  0.66825 -0.6329387 1.969439
3  0.66825 -0.6329387 1.969439
4  0.66825 -0.6329387 1.969439
5  0.66825 -0.6329387 1.969439
6  0.66825 -0.6329387 1.969439
7  0.66825 -0.6329387 1.969439
8  0.66825 -0.6329387 1.969439
9  0.66825 -0.6329387 1.969439
10 0.66825 -0.6329387 1.969439
11 0.66825 -0.6329387 1.969439
12 0.66825 -0.6329387 1.969439
```

Mean response:

```
> a <- predict(lm(y~poly(x1, 2, raw=TRUE)), tmp, interval="confidence", level
=0.95)
> b <- predict(lm(y~poly(x2, 2, raw=TRUE)), tmp, interval="confidence", level
=0.95)
> predict(lm(a~b), tmp, interval="confidence", level=0.95)
       fit        lwr      upr
```

```
1  0.66825 0.09414802 1.242352
2  0.66825 0.09414802 1.242352
3  0.66825 0.09414802 1.242352
4  0.66825 0.09414802 1.242352
5  0.66825 0.09414802 1.242352
6  0.66825 0.09414802 1.242352
7  0.66825 0.09414802 1.242352
8  0.66825 0.09414802 1.242352
9  0.66825 0.09414802 1.242352
10 0.66825 0.09414802 1.242352
11 0.66825 0.09414802 1.242352
12 0.66825 0.09414802 1.242352
```

12.46 A study was done to determine whether the gender of the credit card holder was an important factor in generating profit for a certain credit card company. The variables considered were income, the number of family members, and the gender of the card holder. The data are as follows:

Profit	Income	Gender	Family Members
157	45,000	M	1
−181	55,000	M	2
−253	45,800	M	4
158	38,000	M	3
75	75,000	M	4
202	99,750	M	4
−451	28,000	M	1
146	39,000	M	2
89	54,350	M	1
−357	32,500	M	1
522	36,750	F	1
78	42,500	F	3
5	34,250	F	2
−177	36,750	F	3
123	24,500	F	2
251	27,500	F	1
−56	18,000	F	1
453	24,500	F	1
288	88,750	F	1
−104	19,750	F	2

(a) Fit a linear regression model using the variables available. Based on the fitted model, would the company prefer male or female customers?

(b) Would you say that income was an important factor in explaining the variability in profit?

(a)

```
> p <- c(157,
+        -181,
+        -253,
+        158,
+        75,
+        202,
+        -451,
+        146,
+        89,
+        -357,
```

```
+         522,
+         78,
+         5,
+         -177,
+         123,
+         251,
+         -56,
+         453,
+         288,
+         -104)
>
>
> i <- c(45000,
+         55000,
+         45800,
+         38000,
+         75000,
+         99750,
+         28000,
+         39000,
+         54350,
+         32500,
+         36750,
+         42500,
+         34250,
+         36750,
+         24500,
+         27500,
+         18000,
+         24500,
+         88750,
+         19750)
>
>
> g <- c('M',
+         'M',
+         'M',
+         'M',
+         'M',
+         'M',
+         'M',
+         'M',
+         'M',
+         'M',
+         'F',
+         'F',
+         'F',
+         'F',
+         'F',
+         'F',
+         'F',
+         'F',
+         'F',
+         'F')
>
> f <- c(1,
+         2,
+         4,
+         3,
+         4,
+         4,
+         1,
+         2,
+         1,
```

```
+        1,
+        1,
+        3,
+        2,
+        3,
+        2,
+        1,
+        1,
+        1,
+        1,
+        2)
>
>
> lm(p ~ i + g + f)

Call:
lm(formula = p ~ i + g + f)

Coefficients:
(Intercept)            i           gM            f
  3.008e+01    5.433e-03   -2.367e+02   -4.924e+01
```
The company would prefer female customers.

(b)