HW Question Week 1

STA2101F 2021

Amy Farrow

17 September 2021

(a) Fit the linear model `resist ~ x1 + x2 + x3 + x4`. Extract the X matrix using the `model.matrix` function. How have the levels of the factors been coded?

Model 1:

```
##             Estimate Std. Error t value  Pr(>|t|)
## (Intercept)  236.781     14.769 16.0322 5.645e-09
## x1+           25.762     13.210  1.9502 0.0770849
## x2+          -69.887     13.210 -5.2906 0.0002561
## x3+           43.587     13.210  3.2996 0.0070828
## x4+          -14.488     13.210 -1.0967 0.2961929
##
## n = 16, p = 5, Residual SE = 26.41970, R-Squared = 0.8
```

X Matrix:

(Intercept)	x1+	x2+	x3+	x4+
1	0	0	0	0
1	1	0	0	0
1	0	1	0	0
1	1	1	0	0
1	0	0	1	0
1	1	0	1	0
1	0	1	1	0
1	1	1	1	0
1	0	0	0	1
1	1	0	0	1
1	0	1	0	1
1	1	1	0	1
1	0	0	1	1
1	1	0	1	1
1	0	1	1	1
1	1	1	1	1

- is coded as 0 and + is coded as 1.

(b) Compute the correlation between the columns of the X matrix. Why are there some missing values?

Correlation Matrix:

	(Intercept)	x1+	x2+	x3+	x4+
(Intercept)	1	NA	NA	NA	NA
x1+	NA	1	0	0	0
x2+	NA	0	1	0	0
x3+	NA	0	0	1	0
x4+	NA	0	0	0	1

The standard deviation of the Intercept column in X is zero, so the correlation coefficient between it and all other columns is not calculable.

(c) What difference in resistance is expected when moving from the low to the high level of `x1`?

From part (a), we can expect an increase of 25.762 when moving from the low to the high level of `x1`.

(d) Refit the model without `x4` and examine the regression coefficients and standard errors. What stayed the same and what changed? How is this related to the correlation matrix of X?

Model 2:

```
##             Estimate Std. Error t value  Pr(>|t|)
## (Intercept)  229.538     13.321 17.2312 7.883e-10
## x1+           25.762     13.321  1.9340 0.0770472
## x2+          -69.887     13.321 -5.2464 0.0002056
## x3+           43.587     13.321  3.2721 0.0066773
##
## n = 16, p = 4, Residual SE = 26.64201, R-Squared = 0.78
```

Comparing to part (a), the estimated coefficients for `x1`, `x2`, and `x3` do not change. This is expected because the correlation matrix of X shows zero correlation between each pair of `xi`.

The standard errors, t-values, and p-values for `xi` change slightly (increasing, magnitude decreasing, and decreasing, respectively). The intercept estimate decreases slightly, and the corresponding standard error decreases, t-value increases in magnitude, and p-value decreases.