Fig. 3.5 Prediction

MP

2015-06-04

```
> setwd("D:/Dropbox/R/2015-NUS/Session-2/(a) Data Modelling - Basics/Figure 3.5/Prediction")
```

```
> Dataset <-
+   read.table("D:/Dropbox/R/2015-NUS/Session-2/(a) Data Modelling - Basics/Figure 3.5/Prediction/Table 3.1 Sales-Advertising.csv",
+    header=TRUE, sep=",", na.strings="NA", dec=".", strip.white=TRUE)
```

```
> x <- Dataset$ADVT
```

```
> x
```

```
 [1]  9.5 10.1  9.4 11.6 10.3  9.5 11.2  9.0 11.0  8.4 11.2  8.8 11.9  9.8
[15] 11.2  9.6  9.3 10.0  8.4 10.5 11.8  9.0 10.2 11.8 10.4
```

```
> y <- Dataset$SALES
```

```
> y
```

```
 [1] 145.1 128.3 121.3 134.4 106.5 111.5 132.7 126.9 151.0 123.3 154.6
[12] 108.0 159.3 136.3 111.4 133.6 137.0 112.9 122.1 140.5 141.5  88.5
[23] 127.7 130.7 122.3
```

```
> predict(lm(y ~ x))
```

```
       1        2        3        4        5        6        7        8
123.3581 127.8745 122.6054 139.1653 129.3799 123.3581 136.1544 119.5945
       9       10       11       12       13       14       15       16
134.6490 115.0781 136.1544 118.0890 141.4235 125.6163 136.1544 124.1108
      17       18       19       20       21       22       23       24
121.8527 127.1217 115.0781 130.8854 140.6708 119.5945 128.6272 140.6708
      25
130.1326
```

```
> new <- data.frame(x = seq(8, 12, 0.25))
```

```
> new
```

```
       x
1   8.00
2   8.25
3   8.50
4   8.75
5   9.00
6   9.25
7   9.50
8   9.75
9  10.00
10 10.25
11 10.50
12 10.75
13 11.00
14 11.25
15 11.50
16 11.75
17 12.00
```

```
> predict(lm(y ~ x), new, se.fit = TRUE)
```

```
$fit
       1        2        3        4        5        6        7        8
112.0672 113.9491 115.8309 117.7127 119.5945 121.4763 123.3581 125.2399
       9       10       11       12       13       14       15       16
127.1217 129.0036 130.8854 132.7672 134.6490 136.5308 138.4126 140.2944
      17
142.1762

$se.fit
       1        2        3        4        5        6        7        8
6.584129 5.976631 5.387833 4.824588 4.296955 3.819721 3.414086 3.108210
       9       10       11       12       13       14       15       16
2.933467 2.913548 3.051485 3.327707 3.711466 4.173200 4.689934 5.245440
      17
5.828643

$df
[1] 23

$residual.scale
[1] 14.51066
```

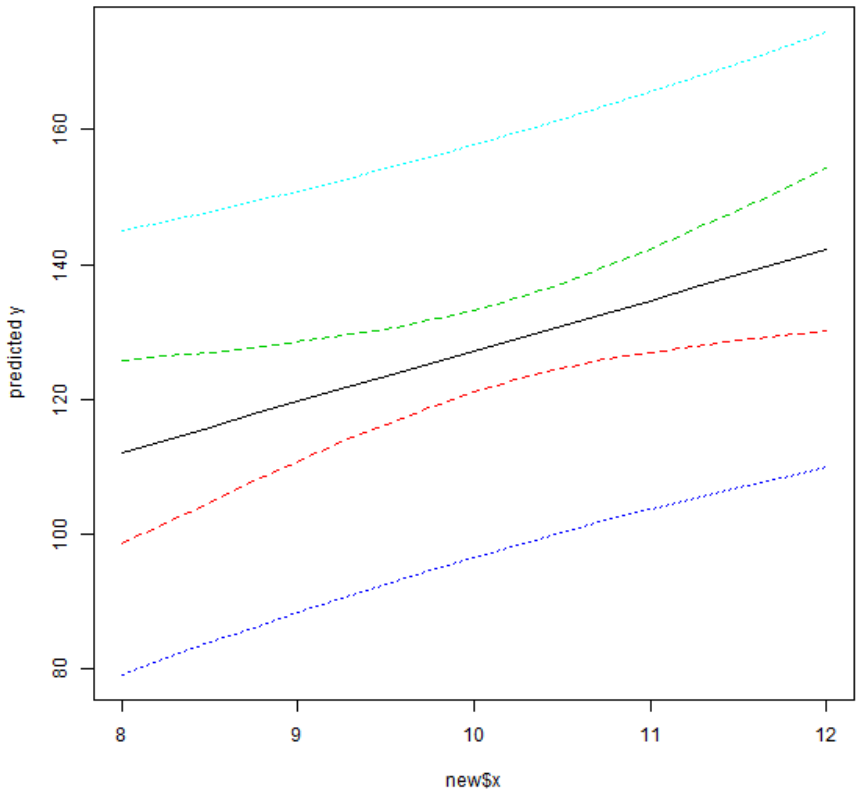
```
> pred.w.plim <- predict(lm(y ~ x), new, interval = "prediction")
```

```
> pred.w.clim <- predict(lm(y ~ x), new, interval = "confidence")
```

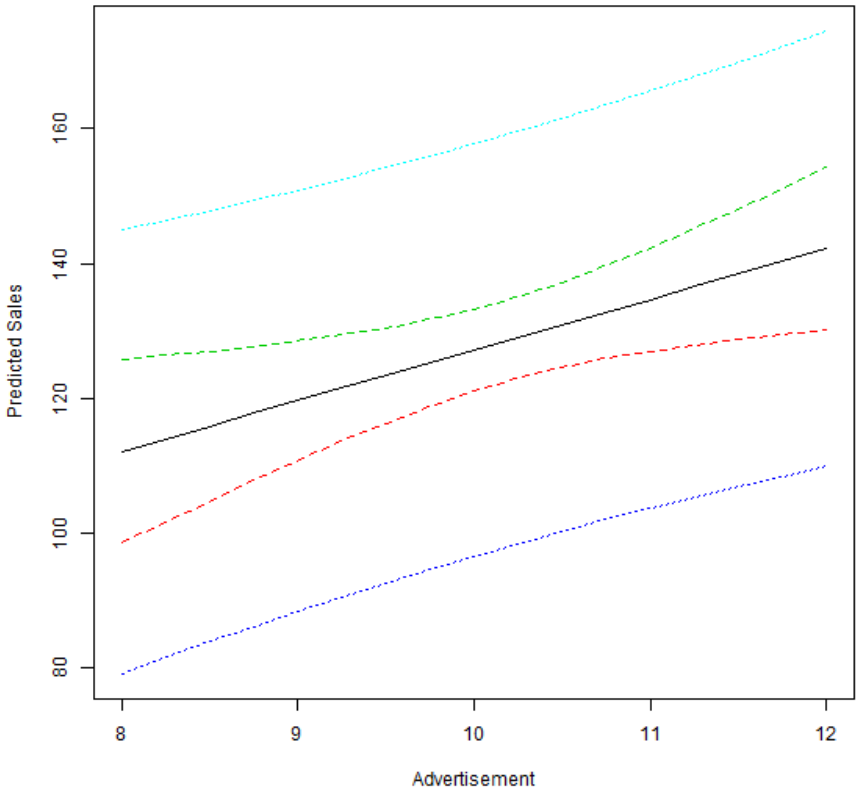
```
> new$x
```

```
 [1]  8.00  8.25  8.50  8.75  9.00  9.25  9.50  9.75 10.00 10.25 10.50
[12] 10.75 11.00 11.25 11.50 11.75 12.00
```

```
> matplot(new$x, cbind(pred.w.clim, pred.w.plim[,-1]),lty = c(1,2,2,3,3), type = "l", ylab =
+   "predicted y")
```




```
> matplot(new$x, cbind(pred.w.clim, pred.w.plim[,-1]),lty = c(1,2,2,3,3), type = "l", ylab =
+   "Predicted Sales", xlab = "Advertisement")
```




```
> RegModel.1 <- lm(SALES~ADVT, data=Dataset)
> summary(RegModel.1)
```

```
Call:
lm(formula = SALES ~ ADVT, data = Dataset)

Residuals:
     Min       1Q   Median       3Q      Max
-31.0945  -9.9708   0.4255   9.6146  21.7419

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)   51.849     27.990   1.852   0.0768 .
ADVT           7.527      2.741   2.746   0.0115 *
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 14.51 on 23 degrees of freedom
Multiple R-squared:  0.2469,    Adjusted R-squared:  0.2142
F-statistic:  7.54 on 1 and 23 DF,  p-value: 0.01151
```

```
> new.Point <- data.frame(ADVT=9.2)
```

```
> new.Point
```

```
  ADVT
1  9.2
```

```
> predict(RegModel.1,new.Point)
```

```
       1
121.0999
```

```
> predict(RegModel.1,new.Point, int="p")
```

```
       fit     lwr      upr
1 121.0999 90.0116 152.1883
```

```
> predict(RegModel.1,new.Point, int="c")
```

```
       fit     lwr      upr
1 121.0999 113.011 129.1889
```