

Section 2.4 Intervals for Predictions

Load needed packages.

```
library(Stat2Data)
library(mosaic)
```

Create a dataframe for **AccordPrice** and look at the structure of the data.

```
data("AccordPrice")
str(AccordPrice)
```

```
## 'data.frame':  30 obs. of  3 variables:
## $ Age      : int  7 4 4 7 9 1 18 2 2 5 ...
## $ Price    : num  12 17.9 15.7 12.5 9.5 21.5 3.5 22.8 26.8 13.6 ...
## $ Mileage  : num  74.9 53 79.1 50.1 62 4.8 89.4 20.8 4.8 48.3 ...
```

Find the least-squares regression line.

```
regmodel=lm(Price~Mileage, data=AccordPrice)
summary(regmodel)
```

```
##
## Call:
## lm(formula = Price ~ Mileage, data = AccordPrice)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -6.5984 -1.8169 -0.4148  1.4502  6.5655
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)  20.8096     0.9529   21.84 < 2e-16 ***
## Mileage      -0.1198     0.0141   -8.50 3.06e-09 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 3.085 on 28 degrees of freedom
## Multiple R-squared:  0.7207, Adjusted R-squared:  0.7107
## F-statistic: 72.25 on 1 and 28 DF,  p-value: 3.055e-09
```

EXAMPLE 2.6 Intervals for Accord prices

```
newx=data.frame(Mileage=50)
fit=predict(regmodel,newdata=newx)
fit
```

```
##          1
## 14.81902
```

Now, let's get our prediction interval.

```
predict.lm(regmodel, newx, interval='prediction')
```

```
##          fit      lwr      upr
## 1 14.81902  8.393807 21.24422
```

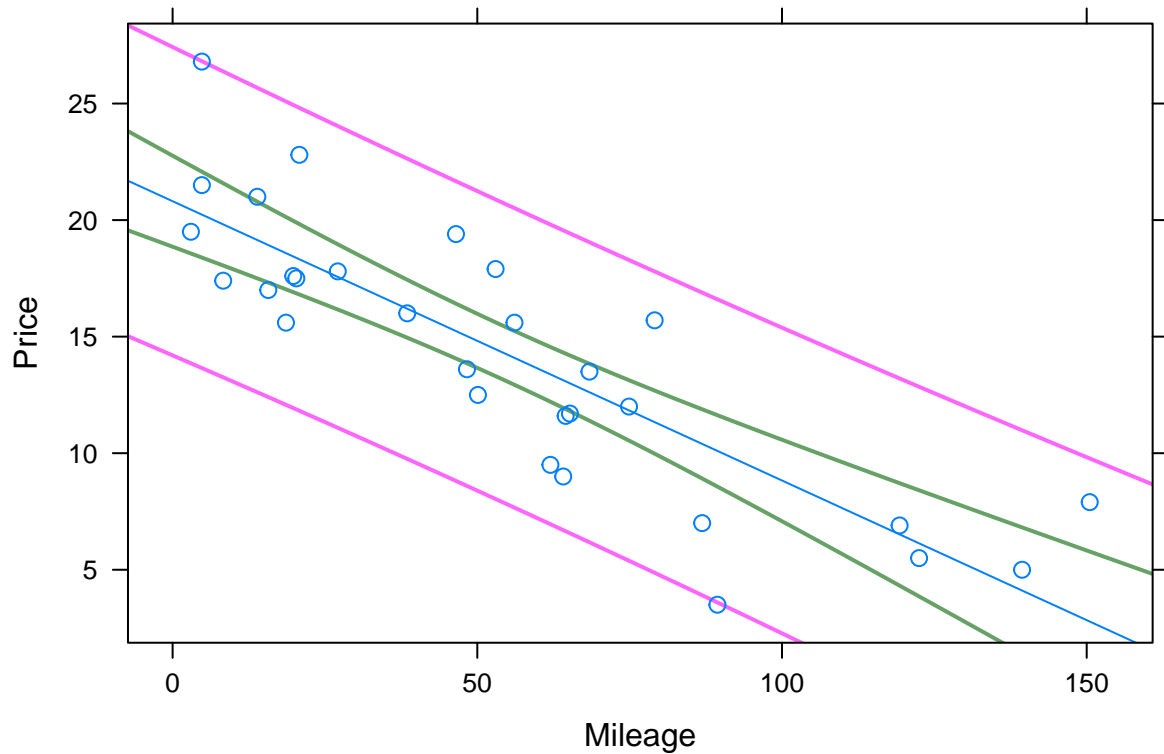
To get a 95% confidence interval for the mean of all accords with 50,000 miles, we simply change our option to confidence.

```
predict.lm(regmodel, newx, interval='confidence')
```

```
##          fit      lwr      upr
## 1 14.81902 13.65787 15.98016
```

To add the confidence intervals and prediction intervals to our scatterplot, we can use the `xyplot()` function in `mosaic`.

```
xyplot(Price~Mileage, panel=panel.lmbands, cex=1, band.lwd=2, data=AccordPrice)
```



Alternative Solutions

There are many ways to create scatterplots with prediction and confidence intervals included. Below you will find three additional ways. (This is one of the benefits of R. Once you get comfortable with the coding, there are lots of different ways to do the same task.)

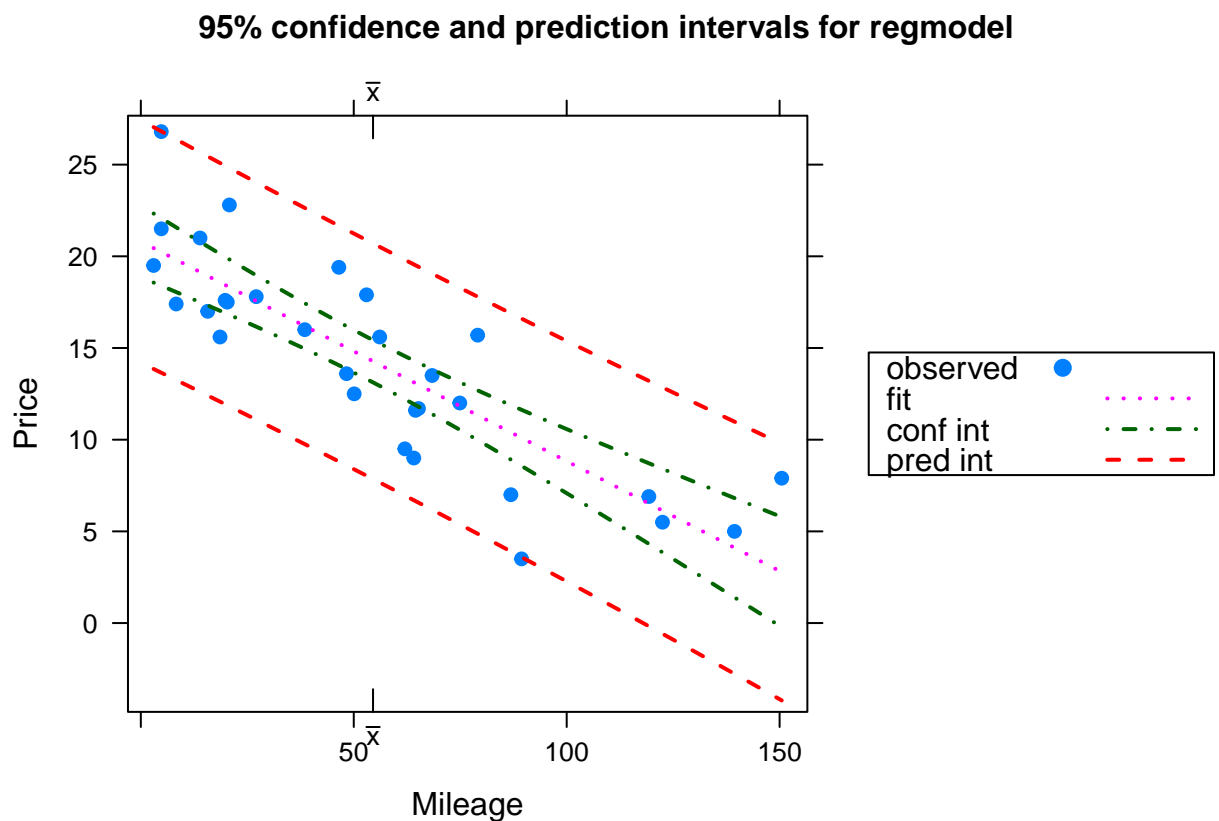
Load an additional package.

Note: The HH library requires 20+ additional packages.

```
library(HH)
```

To add the confidence intervals and prediction intervals to our scatterplot, we can use the `ci.plot()` function.

```
ci.plot(regmodel)
```



You can always do things the long way too.

```
regmodel = lm(Price~Mileage, data=AccordPrice)
#Get a list of lots of possible Mileages
xp = seq(min(AccordPrice$Mileage),max(AccordPrice$Mileage), length = 100)
newx=data.frame(Mileage=xp)
#get the Prediction and confidence intervals for all those Mileage values
p1 = predict(regmodel, newx, interval = 'confidence')
p2 = predict(regmodel, newx, interval = 'prediction')

plot(Price~Mileage, data = AccordPrice, pch = 16, ylim = range(p2))
```

```
abline(regmodel,lwd=2)
lines(xp, p1[,2], lty = 2,lwd="3", col="red"); lines(xp, p1[,3], lty = 2, lwd="3", col="red")
lines(xp, p2[,2], lty = 3, lwd="3", col="green"); lines(xp, p2[,3], lty = 3, lwd="3", col="green")
```

