Chapters 4: Prediction

Confidence Intervals for Predictions

- Given new predictors, x_0, what is the predicted response?

$$\hat{y}_0 = x_0^T \hat{\beta}$$

- Two types of predictions:
  - Prediction of a future observation
  - Prediction of the future mean response
- Prediction intervals vs. confidence intervals

Confidence Intervals for Predictions Ctd

For a future observation: $y_{new} = x_{new}^T \beta + \epsilon$

$$\hat{y}_0 \pm t_{n-(p+1)}^{(\alpha/2)} \hat{\sigma} \sqrt{1 + x_0^T (X^T X)^{-1} x_0}$$

For the future mean response: $E(y_{new}) = E(x_{new}^T \beta) + E(\epsilon)$

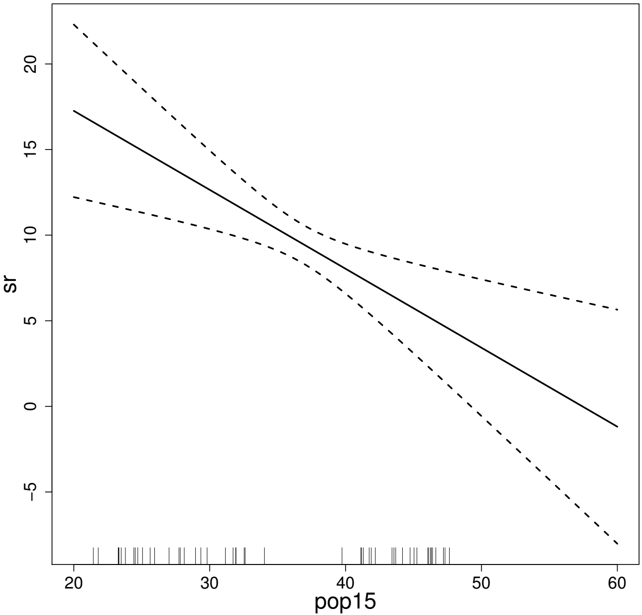
$$\hat{y}_0 \pm t_{n-(p+1)}^{(\alpha/2)} \hat{\sigma} \sqrt{x_0^T (X^T X)^{-1} x_0}$$

Savings Example

```
> result <- lm(sr ~ pop15 + pop75 + dpi + ddpi, savings)
## Convenient way to compute CI and PI
> x0 <- data.frame(pop15=35, pop75=2, dpi=1000, ddpi=4)
> predict(result, x0, interval="confidence")
fit      lwr      upr
[1,] 10.34321 9.093452 11.59297
> predict(result, x0, interval="prediction")
fit      lwr      upr
[1,] 10.34321 2.582946 18.10347
```

```
## Generate a sequence of points
> grid <- seq(20, 60, 1)
> pred <- predict(result, data.frame(pop15=grid, pop75=2,
dpi=1000, ddpi=4), interval="confidence")
## Plot a matrix
> matplot(grid, pred, lty=c(1,2,2), col=1, type="l",
xlab="pop15", ylab="sr")
> rug(savings$pop15)
```

Prediction Band Plot



Prediction Band Plot

Why is prediction band getting wider?

$$\hat{y}_0 \pm t_{n-(p+1)}^{(\alpha/2)} \hat{\sigma} \sqrt{1 + x_0^T (X^T X)^{-1} x_0}$$

$$\hat{y}_0 \pm t_{n-(p+1)}^{(\alpha/2)} \hat{\sigma} \sqrt{1 + a + 2bx_0 + cx_0^2}$$

where

$$(X^T X)^{-1} = \begin{pmatrix} a & b \\ b & c \end{pmatrix}.$$

```
## Data
age = c(rep(20, 10), rep(50, 10))
ht = c( rep(170, 5), rep(180, 5), rep(170, 5), rep(180, 5) )
wt = 0.5 * age + 0.8 * ht + round( rnorm(20, 0, 15), 0 )
data = data.frame(age, ht, wt)
xtabs(wt~age+ht,data)/5
```

age | ht	170	180
20	140.8	165.4
50	171.6	151.2

- Can you use this table to predict the weight for a age of 35 and a height of 170 cm?

```
## Data
age = c(rep(20, 10), rep(50, 10))
ht = c( rep(170, 5), rep(180, 5), rep(170, 5), rep(180, 5) )
wt = 0.5 * age + 0.8 * ht + round( rnorm(20, 0, 15), 0 )
data = data.frame(age, ht, wt)
xtabs(wt~age+ht,data)/5
```

age | ht	170	180
20	150.6	164.2
50	174.0	186.4

- Can you use this table to predict the weight for a age of 35 and a height of 170 cm?

```
## Data
age = c(rep(20, 10), rep(50, 10))
ht = c( rep(170, 5), rep(180, 5), rep(170, 5), rep(180, 5) )
wt = 0.5 * age + 0.8 * ht + round( rnorm(20, 0, 15), 0 )
data = data.frame(age, ht, wt)
xtabs(wt~age+ht,data)/5
```

age | ht	170	180
20	150.6	164.2
50	174.0	186.4

- Can you use this table to predict the weight for a age of 35 and a height of 175 cm?

age | ht	170	180
20	150.6	164.2
50	174.0	186.4

- Can you use this table to predict the weight for a age of 35 and a height of 175 cm?

$$\frac{150.6 + 164.2 + 174.0 + 186.4}{4} = 168.8.$$

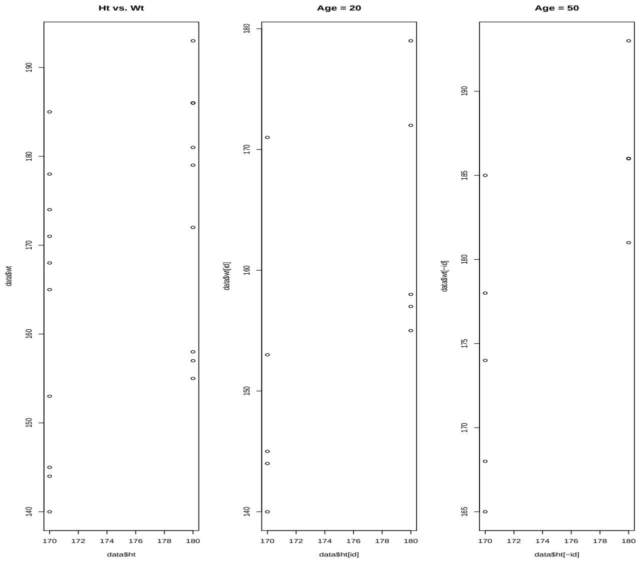
```
lm0 = lm(wt~age+ht,data)
summary(lm0)

Coefficients:
Estimate Std. Error t value Pr(>|t|)
(Intercept) -85.3000    71.0303  -1.201  0.24626
age           0.7600     0.1349   5.632 2.99e-05 ***
ht            1.3000     0.4048   3.211  0.00512 **
```

- Can you use this linear model to predict the weight for a age of 35 and a height of 175 cm?

$$\hat{Wt} = -85.30 + 0.76 \times 35 + 1.3 \times 175 = 168.8$$

```
id = which(data$age ==20)
par(mfrow = c(1,3))
plot(data$ht, data$wt)
plot(data$ht[id], data$wt[id] )
plot(data$ht[-id], data$wt[-id] )
```



Ht vs. Wt
data$wt
data$ht
Age = 20
data$wt[id]
data$ht[id]
Age = 50
data$wt[−id]
data$ht[−id]

```
lm0 = lm(wt~age+ht,data)
summary(lm0)

Coefficients:
Estimate Std. Error t value Pr(>|t|)
(Intercept) -85.3000    71.0303  -1.201  0.24626
age           0.7600     0.1349   5.632 2.99e-05 ***
ht            1.3000     0.4048   3.211  0.00512 **
```

- Can you use this linear model to predict the weight for a age of 35 and a height of 175 cm?

Interpreting Predictions

- True model and parameter values may never be known
- Concentrating on predicting future responses removes the need for interpretation of β
- Conceptually simpler, but need to worry about extrapolation

Extrapolation Example: Ht v.s. Wt

- Data: weights(lb) and heights (cm)
- Estimated from data:

$$\widehat{wt} = 34.61 + 0.82 \times ht$$

ht	185	195	184	175	184	169	182	181	175	178	161	157
wt	184	198	178	171	188	172	189	191	174	186	172	164

- New Obs: $ht = 220$ or $ht = 145$.