Lab 4: ST 412/512

Brian Cervantes Alvarez

January 29, 2025

Table of contents

Inference on Multiple Linear Regression in R **2**

F-tests for Model Comparison . . . 3

Nested Model F-test . . . 3

Overall Model F-test . . . 4

Confidence & Prediction Intervals . . . 5

Model Setup . . . 5

Creating Prediction Grid . . . 5

Making Predictions . . . 6

Visualization Examples . . . 7

Exercises . . . 9

Summary Table . . . 10

Inference on Multiple Linear Regression in R

Lab Goals:

1. Perform F-tests for model comparison
2. Calculate and interpret confidence intervals for mean responses
3. Create prediction intervals for individual observations

F-tests for Model Comparison

Nested Model F-test

When comparing two nested models (where one is a simplified version of the other), we use an extra sum of squares F-test. In R:

```r
library(Sleuth3)

# Continuous intensity model
fit_continuous <- lm(Flowers ~ Intensity + Time, data = case0901)

# Categorical intensity model
fit_indicator <- lm(Flowers ~ factor(Intensity) + Time, data = case0901)

# Model comparison
anova(fit_continuous, fit_indicator)
```

```
Analysis of Variance Table

Model 1: Flowers ~ Intensity + Time
Model 2: Flowers ~ factor(Intensity) + Time
  Res.Df    RSS Df Sum of Sq      F Pr(>F)
1     21 871.24
2     17 767.47  4    103.76 0.5746 0.6848
```

Interpretation: A small p-value suggests the more complex model (with categorical intensity) provides significantly better fit.

Overall Model F-test

To test the overall significance of a model against the intercept-only model:

```r
# Fit full model
fit_full <- lm(Flowers ~ Intensity + Time, data = case0901)

# ANOVA table for overall significance
anova(fit_full)
```

```
Analysis of Variance Table

Response: Flowers
          Df  Sum Sq Mean Sq F value    Pr(>F)
Intensity  1 2579.75 2579.75  62.181 1.037e-07 ***
Time       1  886.95  886.95  21.379 0.0001464 ***
Residuals 21  871.24   41.49
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Key Components:

- **F-statistic:** Ratio of mean square regression to mean square residual
- **p-value:** Probability of observing this F-statistic if null hypothesis (no relationship) is true

```r
# Manual F-statistic calculation
F_stat <- (1733.35 / 41.49)
p_value <- pf(F_stat, 2, 21, lower.tail = FALSE)
cat("F-statistic:", F_stat, "\np-value:", p_value)
```

```
F-statistic: 41.77754
p-value: 4.788273e-08
```

Confidence & Prediction Intervals

Model Setup

We'll use the following model to predict log income:

```
# Load and prepare data
data(ex0923)
head(ex0923)
```

```
  Subject Gender   AFQT Educ Income2005
1       2 female  6.841   12       5500
2       6   male 99.393   16      65000
3       7   male 47.412   12      19000
4       8 female 44.022   14      36000
5       9   male 59.683   14      65000
6      13   male 72.313   16       8000
```

```
# Fit model with na handling
income_model <- lm(log(Income2005) ~ AFQT + Educ + Gender,
                   data = ex0923,
                   na.action = na.exclude)
```

Creating Prediction Grid

Generate combinations of predictor values for visualization:

```
pred_grid <- expand.grid(
  Educ = unique(ex0923$Educ),
  AFQT = seq(0, 100, 10),
  Gender = unique(ex0923$Gender)
)
head(pred_grid)
```

```
  Educ AFQT Gender
1   12    0 female
2   16    0 female
3   14    0 female
4   13    0 female
5   17    0 female
6   15    0 female
```

Making Predictions

Calculate estimated means and intervals:

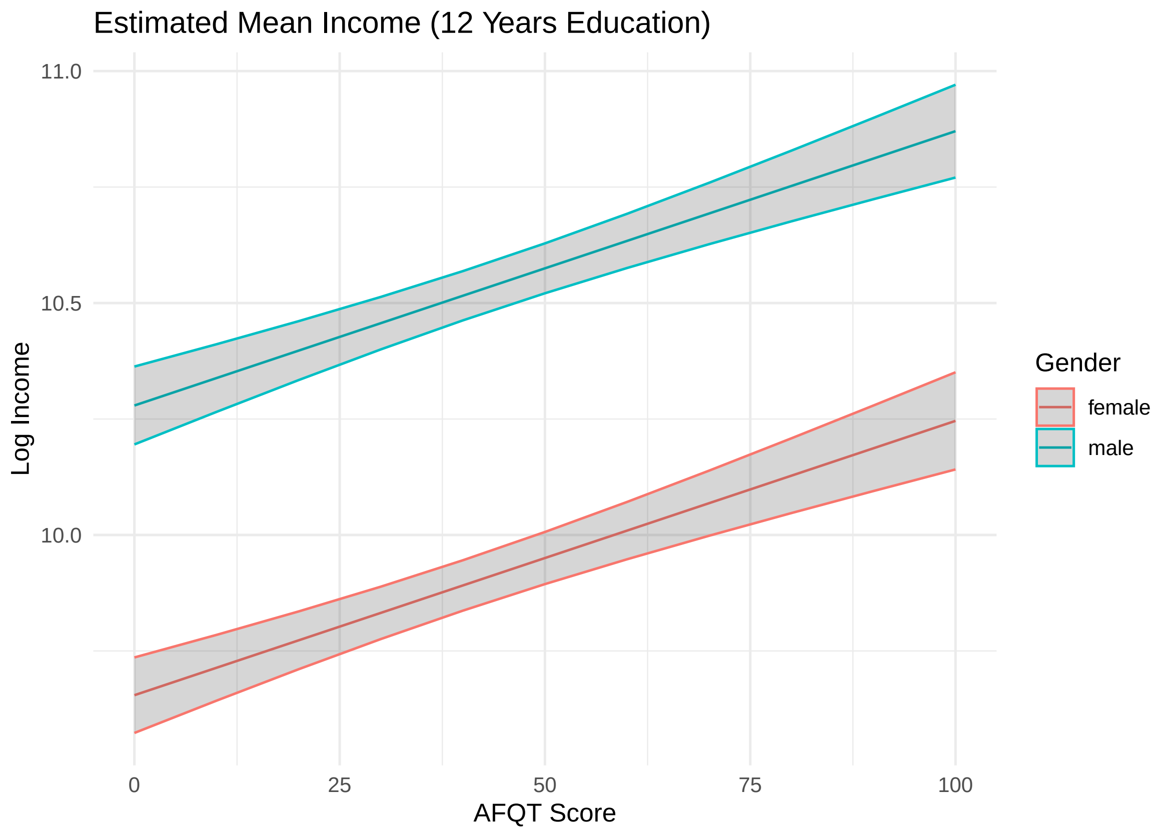
```
# Confidence intervals for mean response
conf_intervals <- predict(income_model,
                          newdata = pred_grid,
                          interval = "confidence")

# Prediction intervals for individual responses
pred_intervals <- predict(income_model,
                          newdata = pred_grid,
                          interval = "prediction")

# Combine results
results <- cbind(pred_grid,
                 conf_mean = conf_intervals[, 1],
                 conf_lwr = conf_intervals[, 2],
                 conf_upr = conf_intervals[, 3],
                 pred_lwr = pred_intervals[, 2],
                 pred_upr = pred_intervals[, 3])
head(results)
```

```
  Educ AFQT Gender conf_mean conf_lwr  conf_upr pred_lwr pred_upr
1   12    0 female  9.654619 9.573247  9.735991 7.954435 11.35480
2   16    0 female  9.962421 9.848133 10.076709 8.260344 11.66450
3   14    0 female  9.808520 9.715068  9.901972 8.107715 11.50932
4   13    0 female  9.731569 9.645545  9.817594 8.031156 11.43198
5   17    0 female 10.039372 9.912655 10.166089 8.336415 11.74233
6   15    0 female  9.885471 9.782415  9.988526 8.184111 11.58683
```

Visualization Examples

1. Mean Income by Education Level:

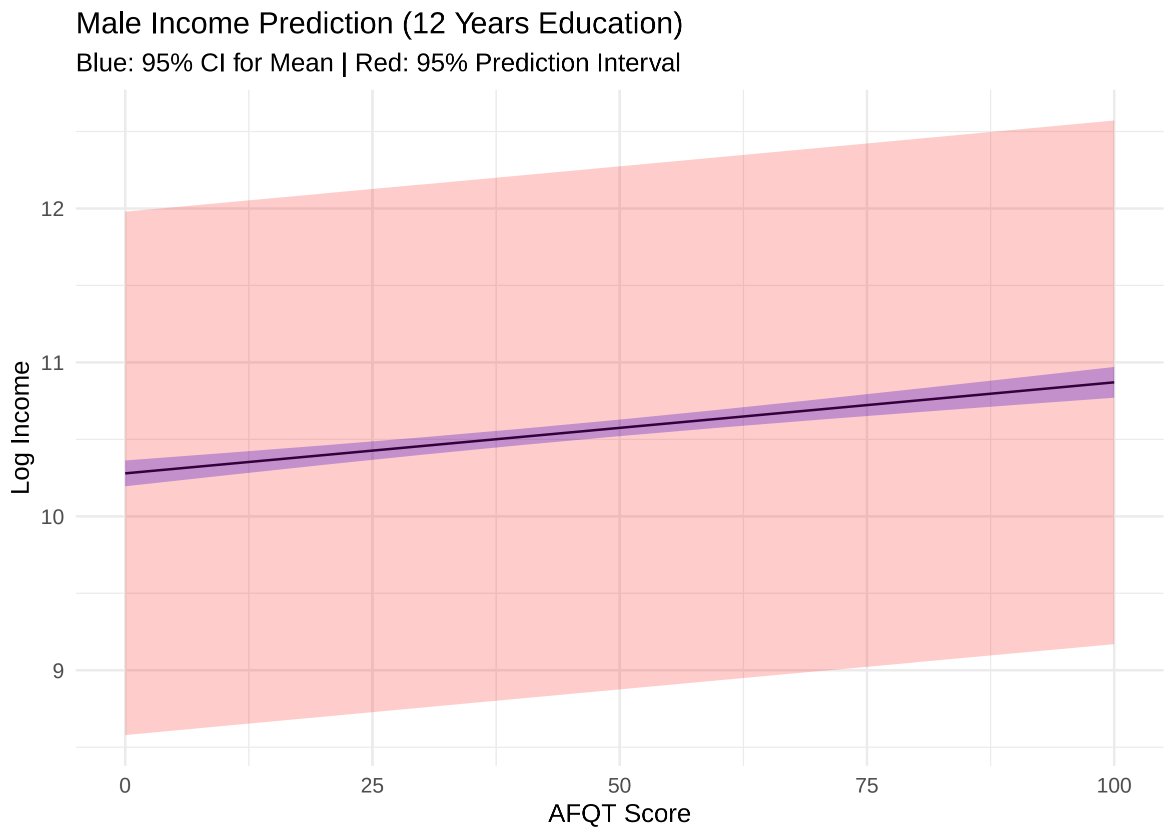
```r
library(ggplot2)

ggplot(subset(results, Educ == 12),
       aes(AFQT, conf_mean, color = Gender)) +
  geom_line() +
  geom_ribbon(aes(ymin = conf_lwr, ymax = conf_upr),
              alpha = 0.2) +
  labs(title = "Estimated Mean Income (12 Years Education)",
       y = "Log Income",
       x = "AFQT Score") +
  theme_minimal()
```



2. Prediction Intervals Comparison:

```r
ggplot(subset(results, Educ == 12 & Gender == "male"),
       aes(AFQT, conf_mean)) +
  geom_line() +
```

```r
  geom_ribbon(aes(ymin = conf_lwr, ymax = conf_upr),
              alpha = 0.3, fill = "blue") +
  geom_ribbon(aes(ymin = pred_lwr, ymax = pred_upr),
              alpha = 0.2, fill = "red") +
  labs(title = "Male Income Prediction (12 Years Education)",
       subtitle = "Blue: 95% CI for Mean | Red: 95% Prediction
         Interval",
       y = "Log Income",
       x = "AFQT Score") +
  theme_minimal()
```



Exercises

1. **Model Comparison:** Fit a model adding an `Educ:AFQT` interaction term. Compare it to our original model using an F-test. Interpret the results.
2. **Visualization:** Recreate the first plot with back-transformed values (original income scale instead of log scale). What changes in the interpretation?
3. **Interval Interpretation:** For a female with 16 years education and AFQT=85:
   - Calculate both confidence and prediction intervals
   - Explain the difference between these intervals in context

Summary Table

Concept	Purpose	R Function
Nested Model F-test	Compare model complexity	`anova(model1, model2)`
Confidence Interval	Estimate precision of mean response	`predict(..., interval = "confidence")`
Prediction Interval	Range for individual observations	`predict(..., interval = "prediction")`

Key Takeaways:

- F-tests help determine if additional predictors significantly improve model fit
- Confidence intervals quantify uncertainty about *mean* responses
- Prediction intervals account for both model uncertainty and individual variation