Inference in regression: confidence intervals

ST552 Lecture 9

Charlotte Wickham

2019-01-28

Today

- Another view of F-tests
- Confidence intervals for single parameters
- Confidence intervals for linear combinations of parameters
- Confidence intervals for parameters jointly

Certain hypotheses of interest can be set up as competing models. A full model and a simpler model (nested in the full model). A.K.A testing models.

Identify the models of interest. Fit both. **Check fit of full model.** Find F-statistic, and answer questions of interest.

Assuming the regression model:

$$Y = X\beta + \epsilon, \quad \epsilon \sim N(0, \sigma^2 I)$$

Consider the hypotheses:

$$H_0 : \underset{k \times p}{K^T} \underset{p \times 1}{\beta} = m$$

$$H_1 : K^T \beta \neq m$$

where $K_{k \times p}^T$ matrix with rank(K) = k.

$K_{p \times k}$

Then under the null hypothesis,

$$F = \frac{\left((K^T\beta - m)^T \left(K^T (X^T X)^{-1} K \right)^{-1} (K^T\beta - m) \right) / k}{RSS/(n-p)} \sim F_{k, n-p}$$

(Don't memorise for ST552, maybe for comps)

This alternative is equivalent to the model testing setup we considered. Every null hypothesis of the form $K^T \beta = m$ is comparing a full and reduced model and vice versa.

For example, consider

$$K = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}_{p \times 1}, \quad m = 0$$

$$K^T$$
$$(0 \, \ldots \, 1 \, \ldots \, 0) \begin{pmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_{p-1} \end{pmatrix} = \beta_{i-1}$$

where the 1 in K occurs in the ith row.

What is the null hypothesis being tested?

$$H_0 : K^T \beta = m$$

$$H_0 : \beta_{i-1} = 0$$

5

Your turn

What are K and m for exercises 1 and 5 from the handout from last time? HW#4

The t-test for an individual parameter can be flipped around to give $100(1 - \alpha)\%$ confidence intervals of the form

$$\hat{\beta}_j \pm t_{n-p}^{(\alpha/2)} \text{SE}\left(\hat{\beta}_j\right)$$

(handwritten annotation above superscript: $1 - \alpha/2$)

(Remember $\text{SE}\left(\hat{\beta}_j\right)$ is coming from the diagonal entry of the estimated variance-covariance matrix.)

Dataset comes from a study of blood coagulation times. 24 animals were randomly assigned to four different diets and the samples were taken in a random order.
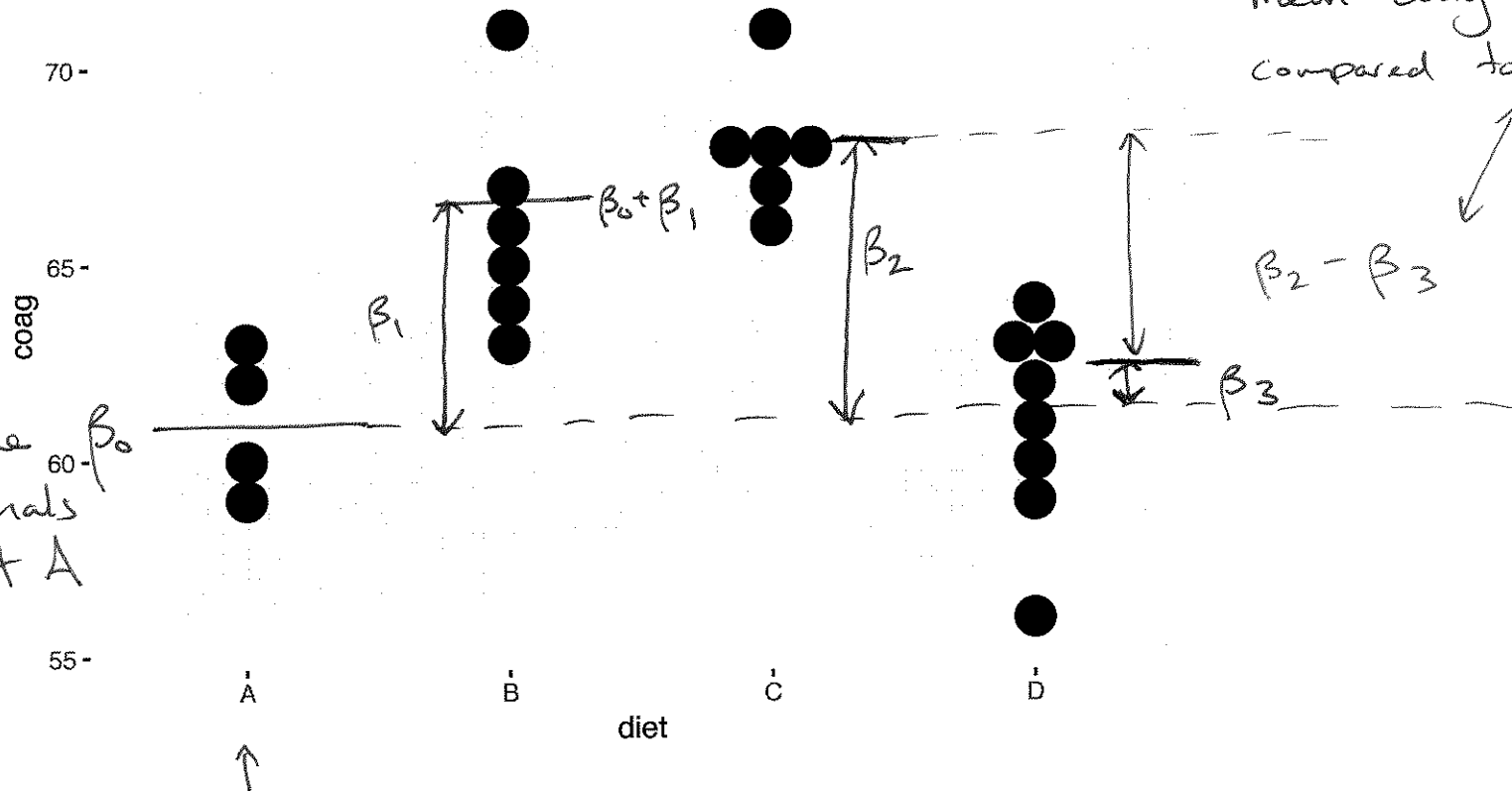
Consider the model:

$$\text{Coagulation time (s)}_i = \beta_0 + \beta_1 1\{\text{Diet B}\}_i$$
$$+ \beta_2 1\{\text{Diet C}\}_i + \beta_3 1\{\text{Diet D}\}_i + \epsilon_i$$

$$1\{\text{Diet B}\}_i = \begin{cases} 1 & \text{if animal } i \text{ recieved Diet B,} \\ 0 & \text{otherwise} \end{cases}$$

```
data(coagulation, package = "faraway")
ggplot(coagulation, aes(diet, coag)) +
  geom_dotplot(binaxis = "y", stackdir = "center", binwidth = 1)
```

```
fit <- lm(coag ~ diet, data = coagulation)
broom::tidy(fit) %>%
  knitr::kable(digits = 2)
```

term	estimate	std.error	statistic	p.value
(Intercept)	61	1.18	51.55	0
dietB	5	1.53	3.27	0
dietC	7	1.53	4.58	0
dietD	0	1.45	0.00	1

Find a 95% CI for β_0?

$$t_{n-p}^{(0.975)} = t_{20}^{(0.975)} = 2.09$$

β_1?

β_0: $\quad 61 \pm 2.09(1.18) = (58.53, 63.47)$

β_1: $(1.8 \quad , \quad 8.2)$

10

In R:

```
broom::tidy(fit, conf.int = TRUE)
# OR
(cis <- confint(fit))
```

Similarly, confidence intervals for a **linear combination** of the parameters, $c^T\beta$ where $c_{p\times 1}$, can be formed with

$$1\times p \quad p\times 1$$

$$c^T\hat{\beta} \pm t_{n-p}^{(\alpha/2)}\sqrt{\hat{\sigma}^2 c^T(X^TX)^{-1}c}$$

With the coagulation example

$$\text{Coagulation time (s)}_i = \beta_0 + \beta_1 1\{\text{Diet B}\}_i$$
$$+ \beta_2 1\{\text{Diet C}\}_i + \beta_3 1\{\text{Diet D}\}_i + \epsilon_i$$

What is c for the linear combination $\beta_0 - \beta_1$?

Find $c^T(X^TX)^{-1}c$.

$c^T \beta$

$$c = \begin{pmatrix} 1 \\ -1 \\ 0 \\ 0 \end{pmatrix}$$

$$\begin{pmatrix} c_1 & c_2 & c_3 & c_4 \end{pmatrix} \begin{pmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \\ \beta_3 \end{pmatrix} = \beta_0 - \beta_1$$

```
X <- model.matrix(fit)
round(solve(t(X) %*% X), 2)
```

```
##             (Intercept) dietB dietC dietD
## (Intercept)    0.25  -0.25 -0.25 -0.25
## dietB         -0.25   0.42  0.25  0.25
## dietC         -0.25   0.25  0.42  0.25
## dietD         -0.25   0.25  0.25  0.37
```

$$\begin{pmatrix} 1 & -1 & 0 & 0 \end{pmatrix} \begin{pmatrix} 0.25 & -0.25 & -0.25 & -0.25 \\ -0.25 & 0.42 & 0.25 & 0.25 \\ -0.25 & 0.25 & 0.42 & 0.25 \\ -0.25 & 0.25 & 0.25 & 0.37 \end{pmatrix} \begin{pmatrix} 1 \\ -1 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 0.5 & -0.67 \end{pmatrix} \begin{pmatrix} 1 \\ -1 \\ 0 \\ 0 \end{pmatrix}$$

$$Var\left(\hat{\beta}_0 - \hat{\beta}_1\right) = Var\left(\hat{\beta}_0\right) + Var\left(\hat{\beta}_1\right) - 2Cov\left(\hat{\beta}_0, \hat{\beta}_1\right) \quad \text{lot's of algebra}$$

$$= \hat{\sigma}\left(0.25 + 0.42 - 2(-0.25)\right)$$

13

With probability $1 - \alpha$

A joint $100(1 - \alpha)\%$ confidence for the vector β can be formed using,

$$(\hat{\beta} - \beta)^T X^T X (\hat{\beta} - \beta) \leq p\hat{\sigma}^2 F_{p,n-p}^{(\alpha)}$$
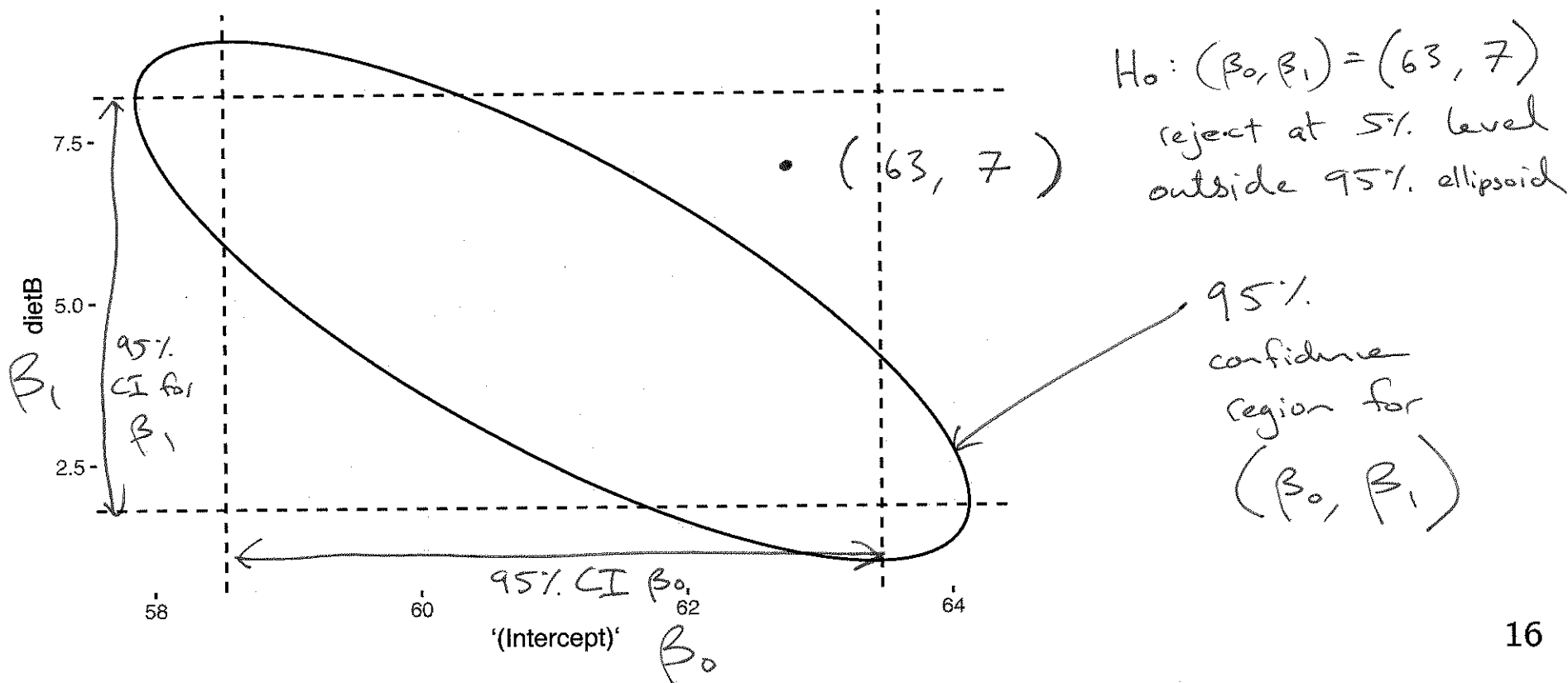
$$(\hat{\beta_0} - \beta_0) \leq \hat{\sigma} t_{n-p}^{(1-\alpha)}$$

and results in p-dimensional ellipsoids (very hard to visualise, but essential for communicating joint uncertainty when the parameters are correlated).

For example, (β_0, β_1) in

Coagulation time $(s)_i = \beta_0 + \beta_1 1\{\text{Diet B}\}_i$
$$+ \beta_2 1\{\text{Diet C}\}_i + \beta_3 1\{\text{Diet D}\}_i + \epsilon_i$$

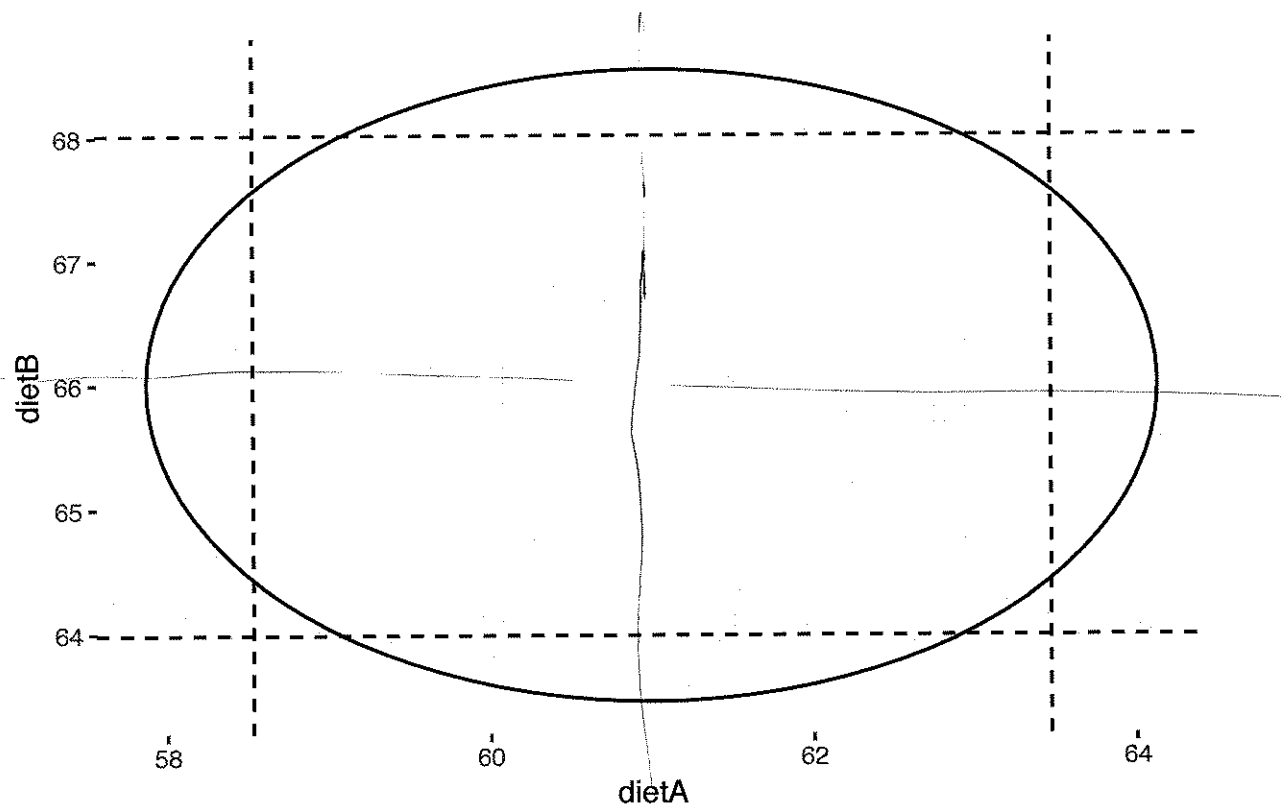$H_0 : (\beta_0, \beta_1) = (63, 7)$
reject at 5% level
outside 95% ellipsoid

$(63, 7)$

95%
confidence
region for
(β_0, β_1)

95%
CI for
β_1

β_1

dietB

95% CI β_0

'(Intercept)' β_0

16

Compare to γ_0 and γ_1 in this parameterization:

$$\text{Coagulation time (s)}_i = \gamma_0 1\{\text{Diet A}\}_i + \gamma_1 1\{\text{Diet B}\}_i$$
$$+ \gamma_2 1\{\text{Diet C}\}_i + \gamma_3 1\{\text{Diet D}\}_i + \epsilon_i$$

```
fit_nointercept <- lm(coag ~ diet - 1, data = coagulation)
```

$$(X^T X)^{-1}$$
$$= \begin{pmatrix} \cdot & O \\ O & \end{pmatrix}$$

17