STA221

Neil Montgomery

Last edited: 2017-02-08 15:00

some test info

Bring a calculator. Suggested methodology:

1. Somehow get a calculator if you don't have one.

some test info

Bring a calculator. Suggested methodology:

1. Somehow get a calculator if you don't have one.
2. Bring it to the test.

some test info

Bring a calculator. Suggested methodology:

1. Somehow get a calculator if you don't have one.
2. Bring it to the test.

some test info

Bring a calculator. Suggested methodology:

1. Somehow get a calculator if you don't have one.
2. Bring it to the test.

This is a two-step process. If you have one and don't bring it, you won't be able to use it.

some test info

Bring a calculator. Suggested methodology:

1. Somehow get a calculator if you don't have one.
2. Bring it to the test.

This is a two-step process. If you have one and don't bring it, you won't be able to use it.

I tease because there is always someone who forgets.

some test info

Bring a calculator. Suggested methodology:

1. Somehow get a calculator if you don't have one.
2. Bring it to the test.

This is a two-step process. If you have one and don't bring it, you won't be able to use it.

I tease because there is always someone who forgets.

Time and location: Monday, February 13, 15:20 to 16:50, in this room BA1160.

some test info

Bring a calculator. Suggested methodology:

1. Somehow get a calculator if you don't have one.
2. Bring it to the test.

This is a two-step process. If you have one and don't bring it, you won't be able to use it.

I tease because there is always someone who forgets.

Time and location: Monday, February 13, 15:20 to 16:50, in this room BA1160.

There will be a formula sheet that comes with the test, along with χ^2 and t probability tables. I will publish these in advance so you will know what to expect.

estimation and prediction with regression models

estimate the mean response at a new x value

Suppose you want to estimate the mean "response" at some new x_ν (may or may not be one of the original x's.)

estimate the mean response at a new x value

Suppose you want to estimate the mean "response" at some new x_ν (may or may not be one of the original x's.)

The *true value* for the mean response is:

$$\mu_\nu = \beta_0 + \beta_1 x_\nu$$

estimate the mean response at a new x value

Suppose you want to estimate the mean "response" at some new x_ν (may or may not be one of the original x's.)

The *true value* for the mean response is:

$$\mu_\nu = \beta_0 + \beta_1 x_\nu$$

What's the "obvious" best guess using the data?

$$\hat{\mu}_\nu = b_0 + b_1 x_\nu$$

estimate the mean response—with confidence

A confidence interval will be as usual based on:

$$\frac{\hat{\mu}_\nu - \mu_\nu}{s.e.(\hat{\mu}_\nu - \mu_\nu)} \sim t_{n-2}$$

estimate the mean response—with confidence

A confidence interval will be as usual based on:

$$\frac{\hat{\mu}_\nu - \mu_\nu}{s.e.(\hat{\mu}_\nu - \mu_\nu)} \sim t_{n-2}$$

The standard error of $\hat{\mu}_\nu - \mu_\nu$ is:

$$s_e\sqrt{\frac{1}{n} + \frac{(x_\nu - \overline{x})^2}{S_{xx}}}$$

estimate the mean response—with confidence

A confidence interval will be as usual based on:

$$\frac{\hat{\mu}_\nu - \mu_\nu}{s.e.(\hat{\mu}_\nu - \mu_\nu)} \sim t_{n-2}$$

The standard error of $\hat{\mu}_\nu - \mu_\nu$ is:

$$s_e\sqrt{\frac{1}{n} + \frac{(x_\nu - \overline{x})^2}{S_{xx}}}$$

So the 95% C.I. for the mean response at x_ν will be:

$$\hat{\mu}_\nu \pm t^*_{n-2} s_e\sqrt{\frac{1}{n} + \frac{(x_\nu - \overline{x})^2}{S_{xx}}}$$

weight model example

Let's make a 95% CI for the mean response at a weight of $x_\nu = 200$ pounds. Here's the R output:

```
##
## Coefficients:
##              Estimate Std. Error t value    Pr(>|t|)
## (Intercept) -14.69314    2.76045  -5.323 0.000000229
## Weight        0.18938    0.01533  12.357     < 2e-16
##
## Residual standard error: 6.538 on 248 degrees of freedom
## Multiple R-squared:  0.3811, Adjusted R-squared:  0.3786
## F-statistic: 152.7 on 1 and 248 DF,  p-value: < 2.2e-16
```

Weight at 200 example

Weight at 200 example

The estimate will be -14.69 $+$ 0.19(200) $=$ 23.18.

Weight at 200 example

The estimate will be -14.69 $+$ 0.19(200) $=$ 23.18.

$s_e = 6.54$

Weight at 200 example

The estimate will be -14.69 $+$ 0.19(200) $=$ 23.18.

$s_e = 6.54$

We would need to be given $\overline{x}$, which in this case is the sample average of the `Weight` variable. This is: 178.08.

Weight at 200 example

The estimate will be -14.69 + 0.19(200) = 23.18.

$s_e = 6.54$

We would need to be given $\overline{x}$, which in this case is the sample average of the `Weight` variable. This is: 178.08.

We also need S_{xx}, which is 181998.49. Could we have determined that from the output given?

Weight at 200 example

The estimate will be -14.69 + 0.19(200) = 23.18.

$s_e = 6.54$

We would need to be given $\overline{x}$, which in this case is the sample average of the `Weight` variable. This is: 178.08.

We also need S_{xx}, which is 181998.49. Could we have determined that from the output given?

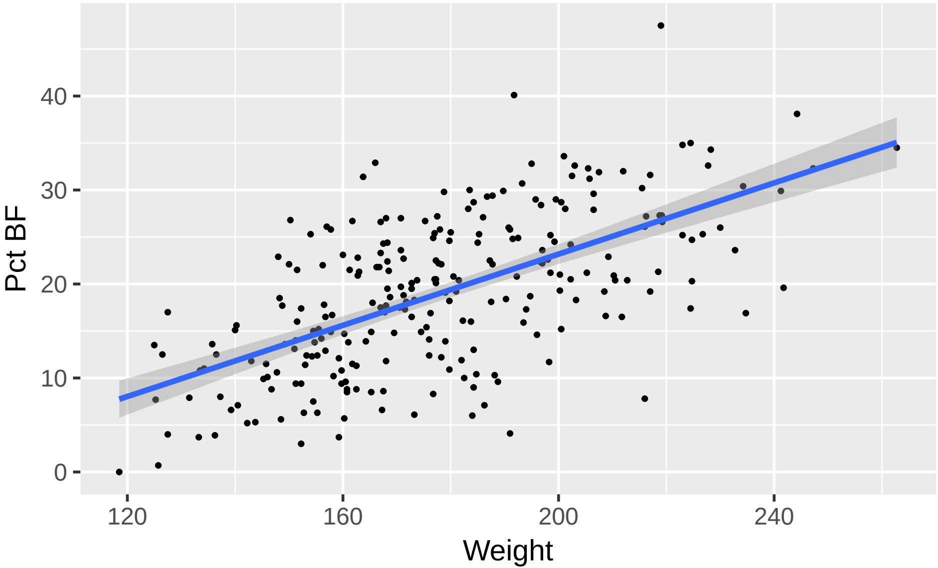
So the 95% CI for the mean `Pct BF` at `Weight`=200 is:

$$23.18 \pm 1.97 \cdot 6.54 \sqrt{\frac{1}{250} + \frac{(200 - 178.08)^2}{181998.49}}$$

or:

$$(22.13, 24.23)$$

picture of 95% CI for mean response - weight model



predict y at a new x value

Suppose you want to predict what y might be at some new x_ν (may or may not be one of the original x's.)

predict y at a new x value

Suppose you want to predict what y might be at some new x_ν (may or may not be one of the original x's.)

There is no *true* value. We are predicting something random (and un-knowable)—not estimating something fixed (but unknown.)

predict y at a new x value

Suppose you want to predict what y might be at some new x_ν (may or may not be one of the original x's.)

There is no *true* value. We are predicting something random (and un-knowable)—not estimating something fixed (but unknown.)

What's the "obvious" best guess using the data?

$$\hat{y}_\nu = b_0 + b_1 x_\nu$$

The *same* guess as the estimate for μ_ν.

predict y at a new x value

Suppose you want to predict what y might be at some new x_ν (may or may not be one of the original x's.)

There is no *true* value. We are predicting something random (and un-knowable)—not estimating something fixed (but unknown.)

What's the "obvious" best guess using the data?

$$\hat{y}_\nu = b_0 + b_1 x_\nu$$

The *same* guess as the estimate for μ_ν.

The variation inherent in such a prediction is different.

predict a new value—with confidence

A prediction interval will be based on, similar to a confidence interval:

$$\frac{\hat{y}_\nu - y_\nu}{s.e.(\hat{y}_\nu - y_\nu)} \sim t_{n-2}$$

predict a new value—with confidence

A prediction interval will be based on, similar to a confidence interval:

$$\frac{\hat{y}_\nu - y_\nu}{s.e.(\hat{y}_\nu - y_\nu)} \sim t_{n-2}$$

The standard error of $\hat{y}_\nu - y_\nu$ is:

$$s_e\sqrt{1 + \frac{1}{n} + \frac{(x_\nu - \overline{x})^2}{S_{xx}}}$$

predict at `Weight` of 200

predict at `Weight` of 200

The prediction will be (also) -14.69 $+$ 0.19(200) $=$ 23.18.

predict at `Weight` of 200

The prediction will be (also) -14.69 + 0.19(200) = 23.18.

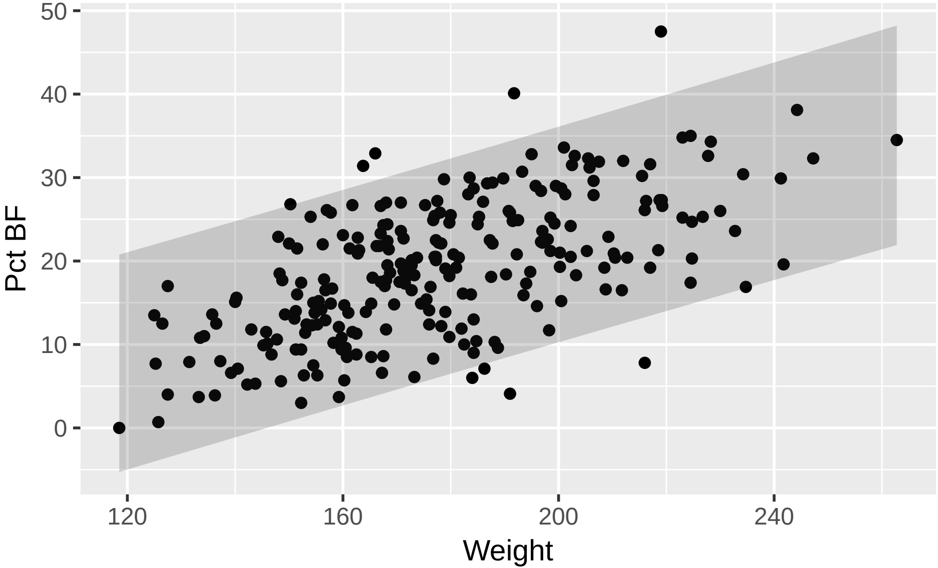
But the 95% "prediction interval" will be quite a bit wider:

$$23.18 \pm 1.97 \cdot 6.54\sqrt{1 + \frac{1}{250} + \frac{(200 - 178.08)^2}{181998.49}}$$

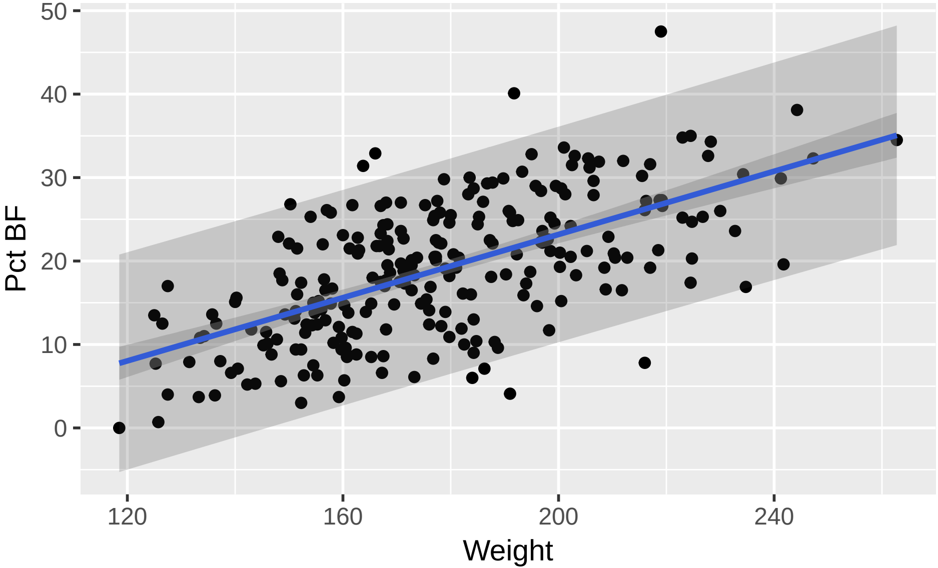
or:

$$(10.26, 36.1)$$

picture of 95% PI weight model



picture of both intervals



“you can never be too rich or too thin” - The Duchess of Windsor

```
## # A tibble: 250 × 15
##   `Pct BF`   Age Weight Height  Neck Chest Abdomen waist   Hip Thigh
##      <dbl> <int>  <dbl>  <dbl> <dbl> <dbl>   <dbl> <dbl> <dbl> <dbl>
## 1      0.0    40    118     68    34    79      69    27    85    47
## 2      0.7    35    126     66    34    91      75    30    89    50
## 3      3.0    35    152     68    37    92      82    32    93    55
## 4      3.7    27    159     72    36    90      80    31    96    55
## 5      3.7    27    133     65    36    94      74    29    88    50
## # ... with 245 more rows, and 5 more variables: Knee <dbl>,
## #   Ankle <dbl>, Bicep <dbl>, Forearm <dbl>, Wrist <dbl>
```