How to calculate the 95% confidence interval for the slope in a linear regression model in R

Ask Question

Here is an exercise from Introductory Statistics with R:

With the rmr data set, plot metabolic rate versus body weight. Fit a linear regression model to the relation. According to the fitted model, what is the predicted metabolic rate for a body weight of 70 kg? Give a 95% confidence interval for the slope of the line.

rmr data set is in the 'ISwR' package. It looks like this:

```
> rmr
   body.weight metabolic.rate
1         49.9           1079
2         50.8           1146
3         51.8           1115
4         52.6           1161
5         57.6           1325
6         61.4           1351
7         62.3           1402
8         64.9           1365
9         43.1            870
10        48.1           1372
11        52.2           1132
12        53.5           1172
13        55.0           1034
14        55.0           1155
15        56.0           1392
16        57.8           1090
17        59.0            982
18        59.0           1178
19        59.2           1342
20        59.5           1027
21        60.0           1316
22        62.1           1574
23        64.9           1526
24        66.0           1268
25        66.4           1205
26        72.8           1382
27        74.8           1273
28        77.1           1439
29        82.0           1536
30        82.0           1151
```

```
37        103.0           1382
38        104.5           1414
39        107.7           1473
40        110.2           2074
41        122.0           1777
42        123.1           1640
43        125.2           1630
44        143.3           1708
```

I know how to calculate the predicted y at a given x but how can I calculate the confidence interval for the slope?

r statistics linear-regression confidence-interval

edited Mar 2 '13 at 22:19
NPE
337k 57 721 859

asked Mar 2 '13 at 22:09
Yu Fu
555 1 5 13

5 have you tried `?confint` ? – Arun Mar 2 '13 at 22:14

1 Answer

Let's fit the model:

```
> library(ISwR)
> fit <- lm(metabolic.rate ~ body.weig
> summary(fit)

Call:
lm(formula = metabolic.rate ~ body.wei

Residuals:
    Min      1Q  Median      3Q     Ma
-245.74 -113.99  -32.05  104.96  484.8

Coefficients:
            Estimate Std. Error t valu
(Intercept) 811.2267    76.9755  10.53
body.weight   7.0595     0.9776   7.22
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.0

Residual standard error: 157.9 on 42 d
Multiple R-squared: 0.5539, Adjusted R
F-statistic: 52.15 on 1 and 42 DF,  p-
```

This can be computed using `confint` :

```
> confint(fit, 'body.weight', level=0.
                2.5 % 97.5 %
body.weight 5.086656 9.0324
```

edited Mar 3 '13 at 8:40

answered Mar 2 '13 at 22:15

NPE
337k 57 721 859

---

13 This is equivalent to the following: `coef=summary(fit)$coefficients[2,1] err=summary(fit)$coefficients[2,2] coef + c(-1,1)*err*qt(0.975, 42) [1] 5.086656 9.032400` : it's the estimated coefficient +- qt(1-alpha/2, df) standard errors – ds440 Mar 2 '13 at 22:59

1 Thank you, NPE! So estimated coefficient +/- two standard errors is an approximation and the latter method provides a accurate way to calculate the confidence interval, right? – Yu Fu Mar 2 '13 at 23:09

5 Yes, the two SE is a good ballpark: if the linear model assumptions are correct then it will follow a T distribution so as sample size increases it approaches ~1.96, for smaller samples it is higher. – ds440 Mar 2 '13 at 23:21

1 @NPE: Are you assuming that a Gaussian pdf for the slope values? If this hypothesis does not hold you could use bootstrap methods. – jpcgandre Sep 12 '14 at 11:53