BÀI TẬP VỀ NHÀ

MÔN HỌC: HỒI QUY TUYẾN TÍNH

**Bài 2:** Bảng số liệu

|  |  |  |  |
| --- | --- | --- | --- |
| Số thứ tự | Độ bền dẻo  (Yi) | Độ dày vật liệu  (Xi1) | Mật độ vật liệu  (Xi2) |
| 1 | 37,8 | 4 | 4,0 |
| 2 | 22,5 | 4 | 3,6 |
| 3 | 17,1 | 3 | 3,1 |
| 4 | 10,8 | 2 | 3,2 |
| 5 | 7,2 | 1 | 3,0 |
| 6 | 42,3 | 6 | 3,8 |
| 7 | 30,2 | 4 | 3,8 |
| 8 | 19,4 | 4 | 2,9 |
| 9 | 14,8 | 1 | 3,8 |
| 10 | 9,5 | 1 | 2,8 |
| 11 | 32,4 | 3 | 3,4 |
| 12 | 21,6 | 4 | 2,8 |

***(Học viên sử dụng phần mềm R để thực hiện bài tập)***

**Câu hỏi 1:** Tìm 2 phương trình đường thẳng hồi quy và 1 phương trình siêu phẳng (Nếu có)?

*Trả lời:*

Nhập liệu R:

Y<-c(37.8, 22.5, 17.1, 10.8, 7.2, 42.3, 30.2, 19.4, 14.8, 9.5, 32.4, 21.6)

X1<-c(4, 4, 3, 2, 1, 6, 4, 4, 1, 1, 3, 4)

X2<-c(4.0, 3.6, 3.1, 3.2, 3.0, 3.8, 3.8, 2.9, 3.8, 2.8, 3.4, 2.8)

Ta xây dựng mô hình hồi quy như sau:

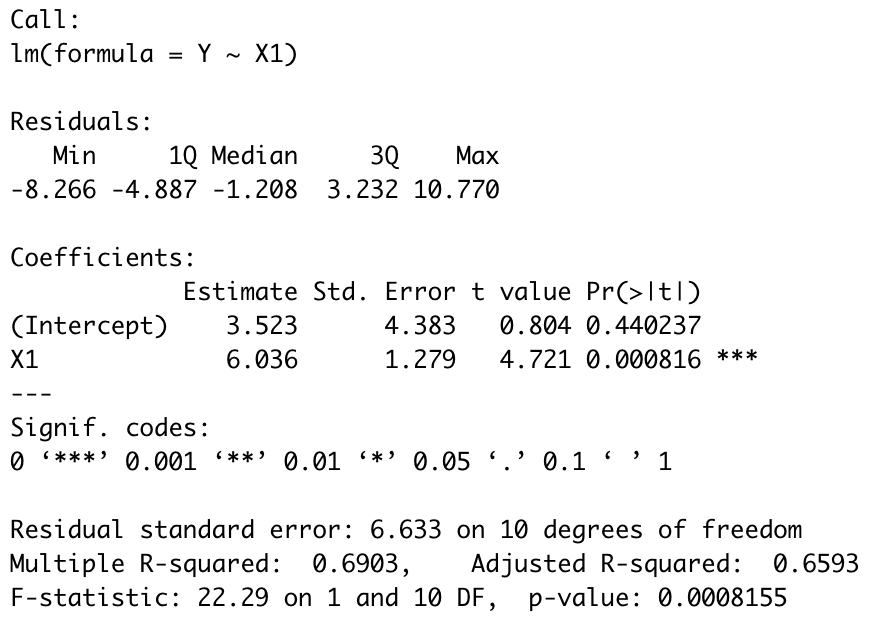
1. **Mô hình 1:**

Mô hình đường thẳng hồi quy:

*Code:*

Y1<-lm(Y~X1)

summary(Y1)



Kết quả:

và

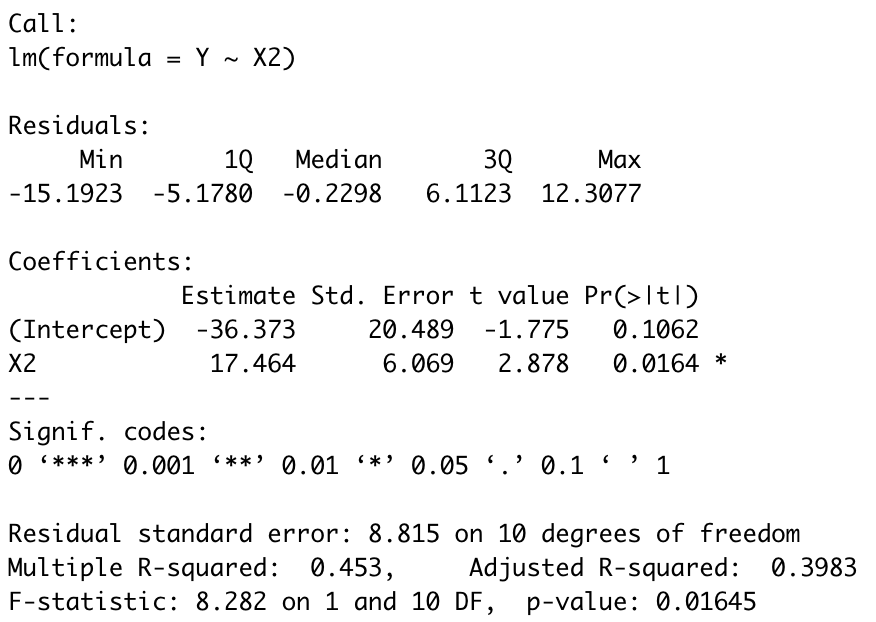
1. **Mô hình 2:**

Mô hình đường thẳng hồi quy:

*Code:*

Y2<-lm(Y~X2)

summary(Y2)



Kết quả:

và

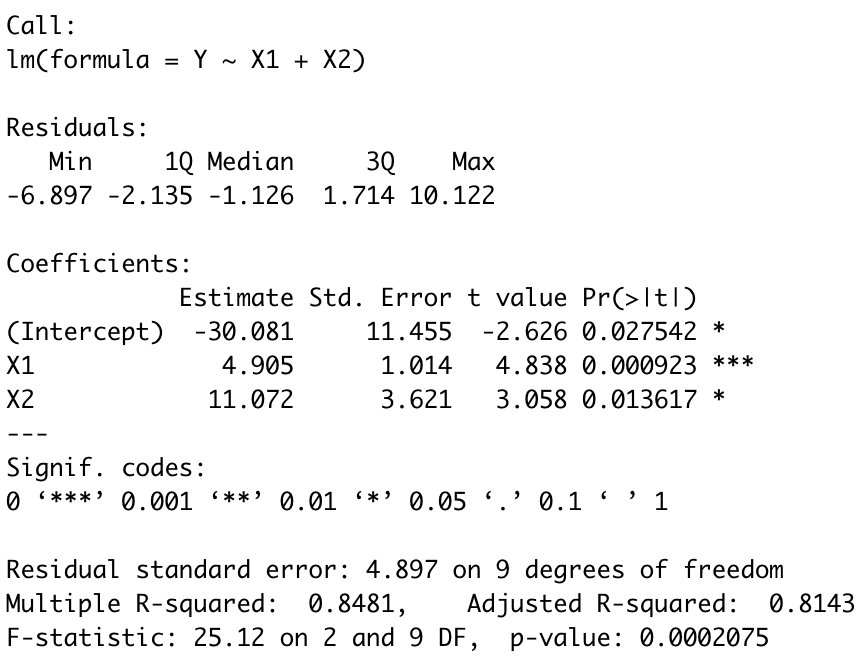
1. **Mô hình 3:**

Mô hình siêu phẳng hồi quy:

*Code:*

Y3<-lm(Y~X1+X2)

summary(Y3)



Kết quả:

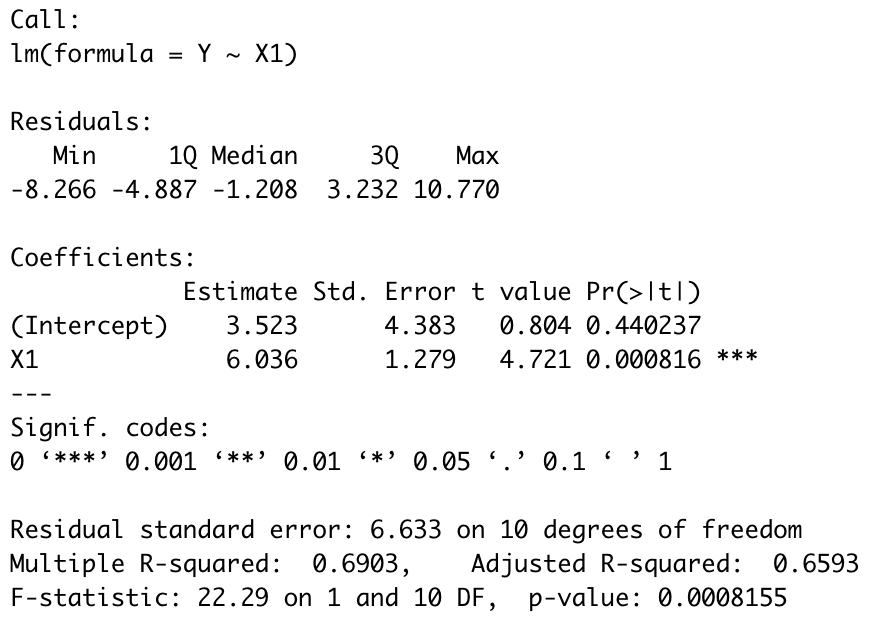
và và

**Câu hỏi 2:** Xác định tỉ lệ phần trăm sự biến thiên của biến phụ thuộc cho từng mô hình có thể có trên.

*Trả lời:*

1. **Mô hình 1:**

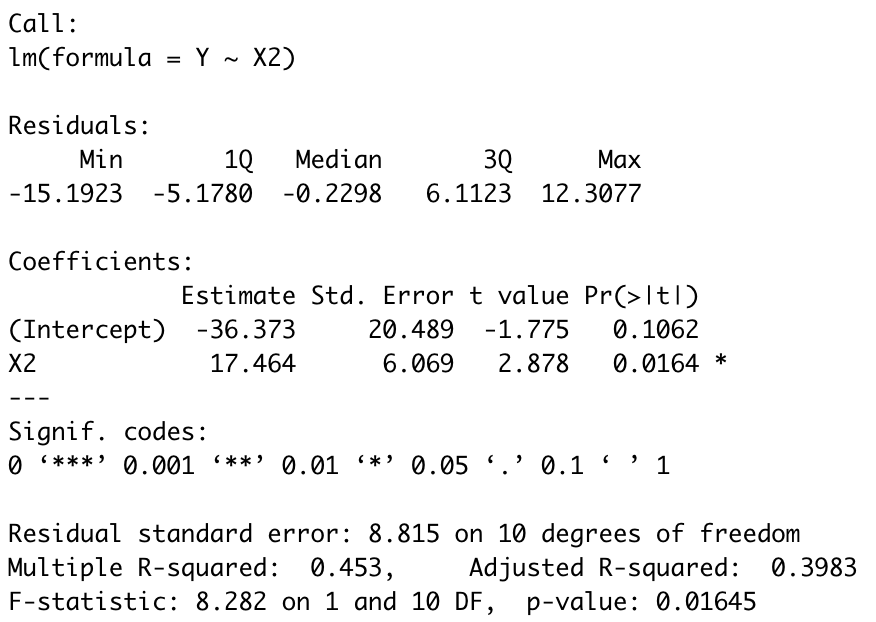
*Quan sát lại:*



* **Kết quả:** Dựa vào *hệ số xác định bội (Multiple R-squared)* là 0,6903. Vậy biến giải thích giải thích 69,03% sự biến thiên của độ dẻo vật liệu (Y) phụ thuộc vào độ dày vật liệu (X1).

1. **Mô hình 2:**

*Quan sát lại:*



* Kết quả: Dựa vào *hệ số xác định bội (Multiple R-squared)* là 0,453. Vậy biến giải thích giải thích 45,3% sự biến thiên của độ dẻo vật liệu (Y) phụ thuộc vào mật độ vật liệu (X2).

1. **Mô hình 3:**

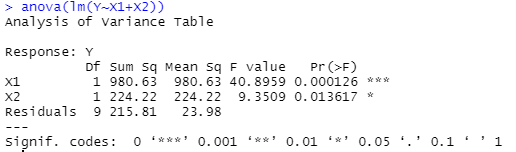
|  |
| --- |
| *Quay lại mô hình:* |

* Kết quả: Dựa vào *hệ số xác định bội (Multiple R-squared)* là 0,8481. Vậy biến giải thích giải thích 84,81% sự biến thiên của độ dẻo (Y) phụ thuộc vào độ dày và mật độ vật liệu .

**Câu hỏi 3:** Nếu chúng ta chỉ quan tâm đến cả hai biến giải thích, hãy lập bảng ANOVA?

*Trả lời:*

Bảng ANOVA

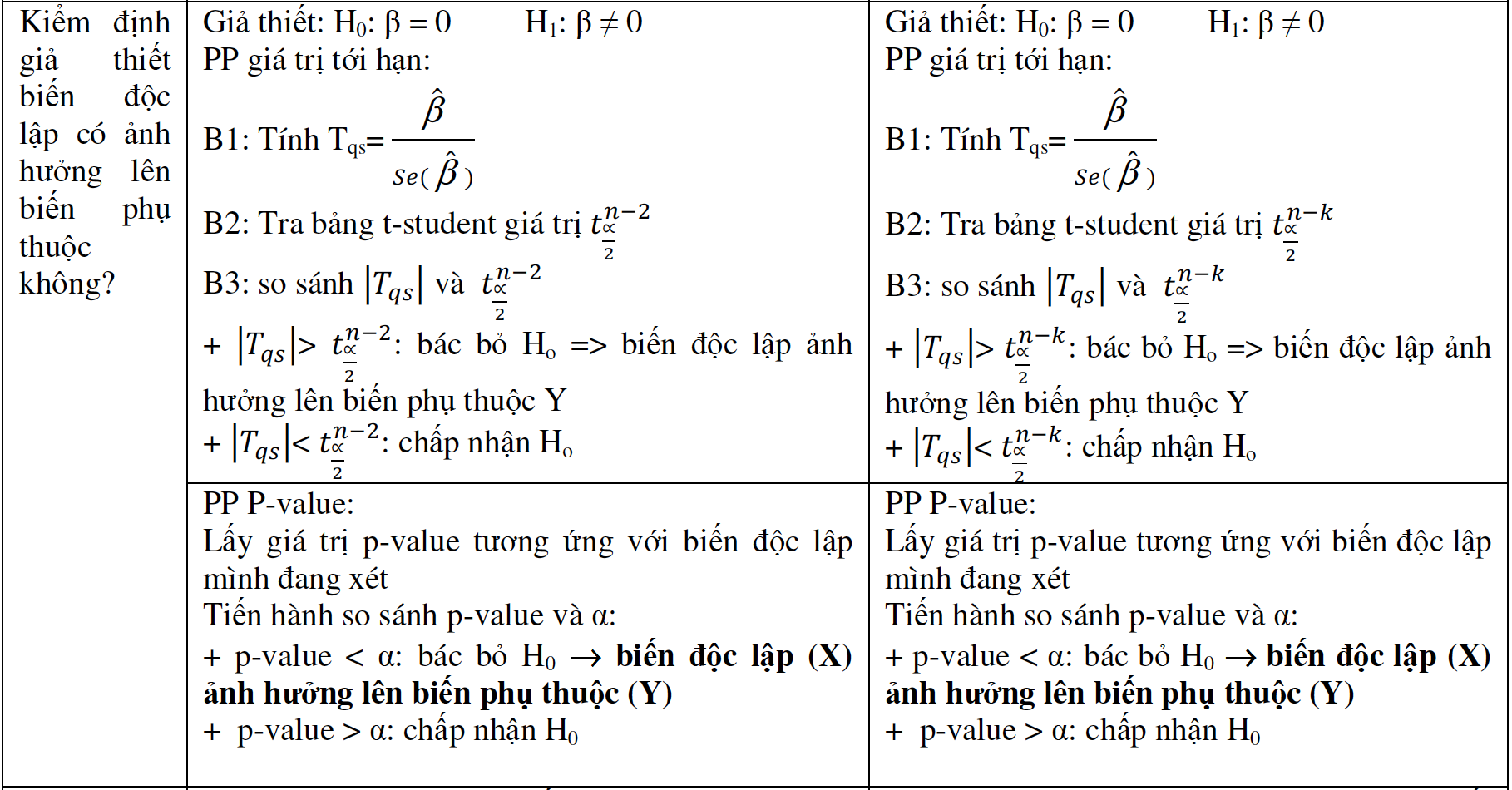


**Câu hỏi 4:** Kiểm định giả thiết sau với mức ý nghĩa 5%,

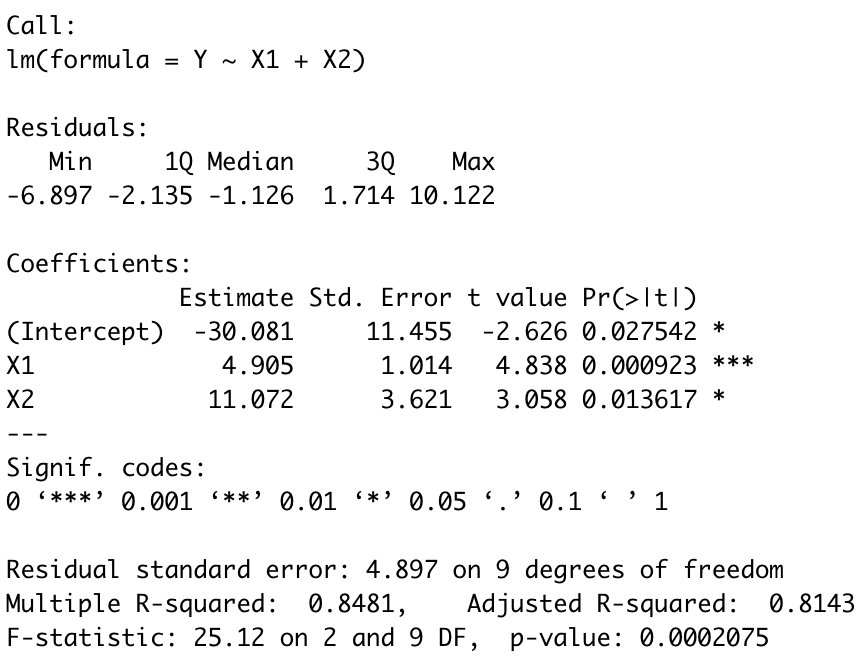
Vậy đối thuyết:

*Trả lời:*

*Nhắc lại kiến thức:*



Sử dụng P-Value trong mô hình hồi quy (3):

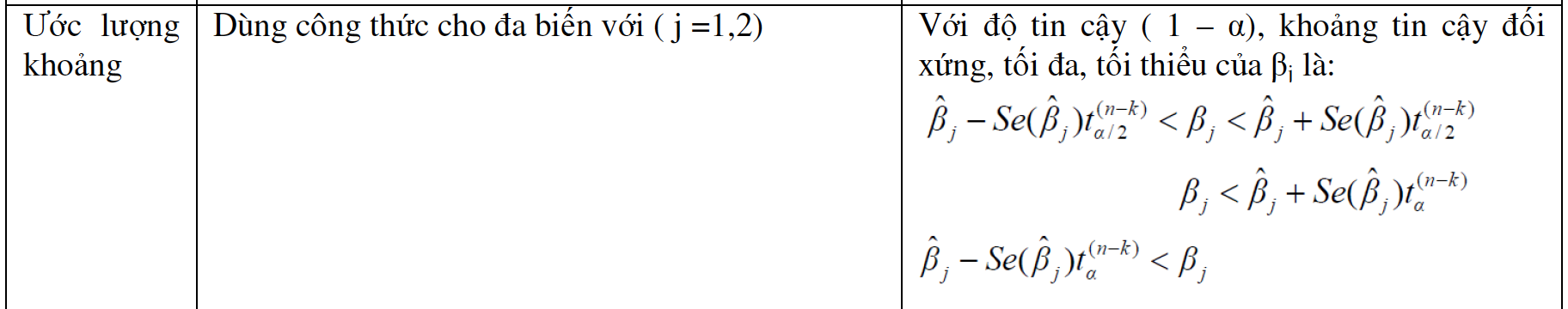


* P-value của X1 = 0,000923 < 0.05, hoặc
* P-value của X2 = 0,013617 < 0.05
* Nên ta bác bỏ giả thuyết
* Chấo nhận đối thuyết:

**Câu hỏi 5:** Xác định khoảng tin cậy với mức ý nghĩa 5% cho trong trường hợp mô hình chỉ có biến độc lập là độ dày của vật liệu.

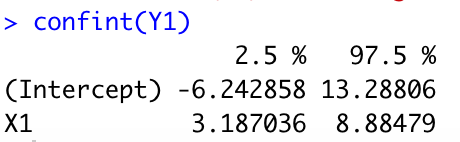
*Trả lời:*

*Nhắc lại kiến thức:*



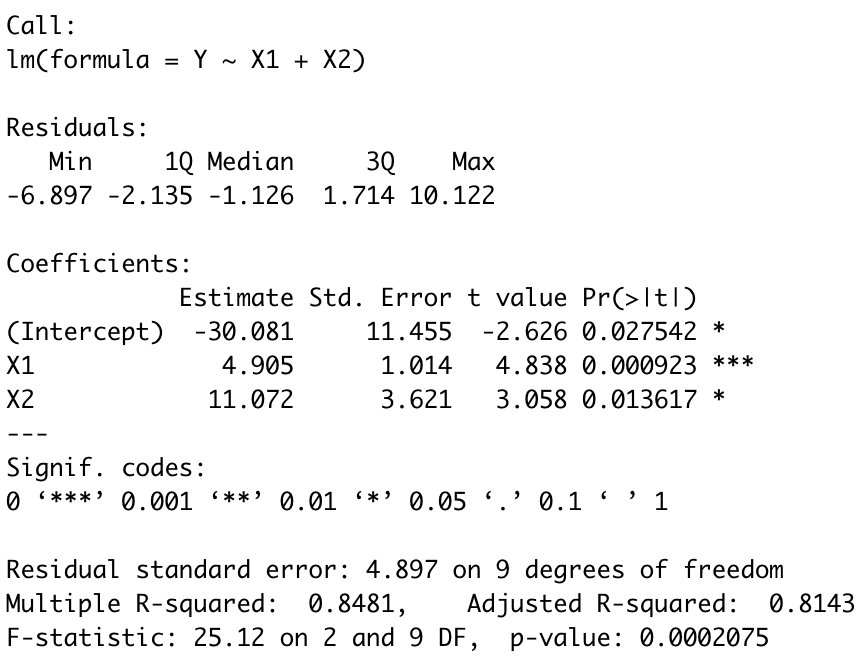
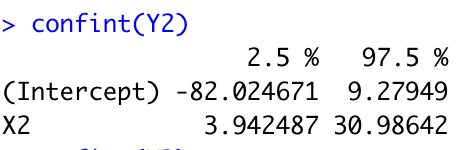
* Quan sát lại kết quả từ mô hình hồi quy:
* Khoảng tin cậy với mức ý nghĩa 5% của là:

Sử dụng code R, ta có kết quả:



**Câu hỏi 6:** Với khoảng tin cậy vừa tìm được ở câu 5, chúng ta có thể khẳng định rằng hồi quy tuyến tính là có ý nghĩa giữa mức độ bền dẻo của nhựa và độ dày của vật liệu và mật độ của vật liệu không? Chứng minh điều khẳng định của bạn.

*Trả lời:*

* Hồi quy tuyến tính là có ý nghĩa giữa mức độ bền dẻo của nhựa và độ dày của vật liệu và mật độ của vật liệu, Vì:
* Quan sát bảng kết quả cho mô hình hồi quy hai biến:
* 
* Ta thấy vẫn thuộc khoảng tin cậy dành cho ước lượng hệ số hồi quy của biến trong mô hình hồi quy đơn.
* Ta tính khoảng tin cậy 95% cho trong mô hình hồi quy đơn:
* Khoảng tin cậy với mức ý nghĩa 5% của là:
* Sử dụng code R, ta có kết quả:
* 
* Ta thấy vẫn thuộc khoảng tin cậy dành cho ước lượng hệ số hồi quy của biến trong mô hình hồi quy đơn.
* Do đó ta kết luận như trên.