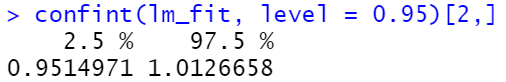
# EXERCISE 1

## a)

## Because 1 is in the interval of B1, so it's plausible for B1=1



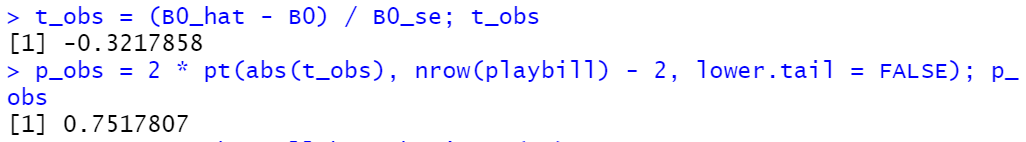
Dựa vào kết quả mô hình tuyến bên dưới ta tìm được 95% khoảng tin cậy cho giá trị B1: [0.9515 ; 1.0127]

Nhận thấy giá trị 1 thuộc khoảng tin cậy của B1, vì vậy có thể xem B1 = 1.

## b)

## We accept the null hypothesis, t(16)=-0.32, p=0.75

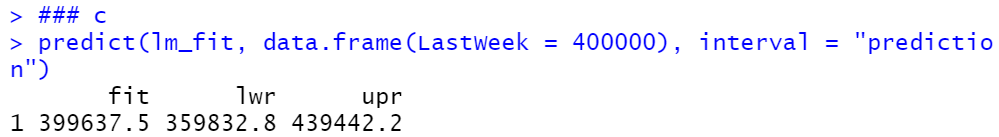
Với p=0.75, chúng ta không đủ cơ sở để bác bỏ giả thuyết không.



## c)

## Because $450000 is out of the prediction interval so it's not feasible.

Dễ nhận thấy 450000 không thuộc khoảng dự đoán [359832.8, 439442.2] của mô hình, nên nó không là giá trị phù hợp.



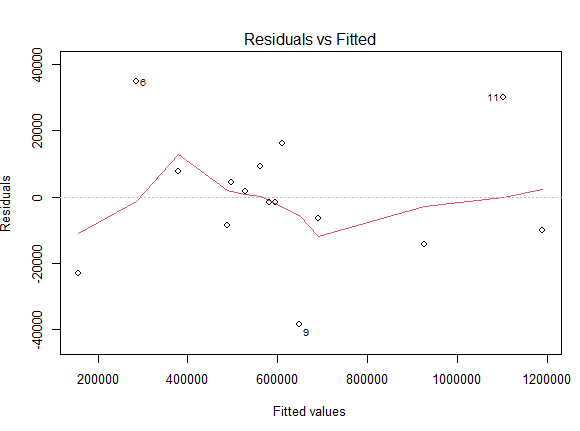
## d)

### Following the Residuals vs Fitted plot, we'll see that the prediction rule is quite reasonable, however, there are at least three bad values outside the LR line, therefore, we do need more data for analyzing or another model to have more clues to reduce the risky probability.

Dựa vào biểu đồ Residuals vs Fitted, ta thấy rằng chiến thuật dự đoán doanh số tuần hiện tại so với tuần trước là một chiến lược khá hợp lý.

Tuy nhiên, có ít nhất 3 điểm bị lệch khỏi đường hồi quy.

Vì vậy, họ cần thêm dữ liệu để có cái nhìn khách quan hơn hoặc dùng một mô hình khác có thể tạo ra quy tắc dự đoán tốt hơn, giảm thiểu rủi ro.

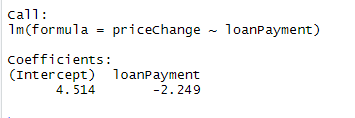


# EXERCISE 2

## a ) Find a 95% confidence interval for the slope of the regression model, b1 . On the basis of this confidence interval decide whether there is evidence of a significant negative linear association.



Dựa vào kết quả mô hình tuyến bên dưới ta tìm được 95% khoảng tin cậy cho giá trị B1: [-4.16:-0.33]



Ta có dạng tổng quát của mô hình hồi quy tuyến tính được cho là: y = 4.514 - 2.249x

(trong đó x: Tỷ lệ cho vay cầm cố, y: Phần trăm thay đổi trong giá trung bình) nên y và x có mối tương quan nghịch hay x tăng thì y giảm và ngược lại.

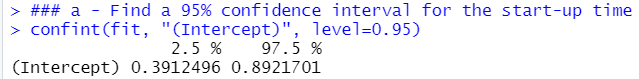
## b) Use the fitted regression model to estimate E ( Y | X =4). Find a 95% confidence interval for E ( Y | X =4). Is 0% a feasible value for E ( Y | X =4)? Give a reason to support your answer.

## 

1. Dựa vào mô hình tổng quát mối liên hệ giữa 2 biến x và y ta tính được giá trị ước lượng của y tại x = 4 là y = -4.48
2. Tương tự ta tính được 95% độ tin cậy của y khi x = 4 nằm trong khoảng: [-6.6:-2.3]
3. Vì 95% độ tin cậy của y khi x = 4 nằm trong khoảng: [-6.6:-2.3] nên 0% không phải là giá trị mong đợi phù hợp khi x = 4

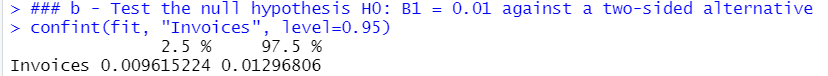
# EXERCISE 3

## a) Find a 95% confidence interval for the start-up time, i.e., b0.



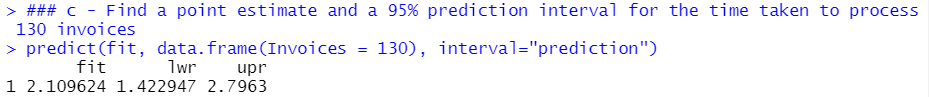
Dựa vào kết quả mô hình bên dưới ta tìm được 95% khoảng tin cậy cho giá trị B0: [0.3912;0.8922]

## b) Suppose that a best practice benchmark for the average processing time for an additional invoice is 0.01 hours (or 0.6 minutes). Test the null hypothesis H0: B1 = 0.01 against a two-sided alternative. Interpret your result.



B1 = 0.01 nằm trong khoảng tin cậy nên chưa đủ cơ sở để bác bỏ H0 với mức ý nghĩa 5%

## c) Find a point estimate and a 95% prediction interval for the time taken to process 130 invoices.



Dựa vào mô hình dự đoán, thời gian cần để xử lý 130 hóa đơn là 2.1 giờ với khoảng tin cậy 95% là [1.4; 2.8] hours.