

LACK-OF-FIT TESTING WHEN REPLICATES ARE NOT AVAILABLE





目錄



🔍 **1. GROUPWISE REGRESSION**

🔍 **2. 比較**

🔍 **3. 結論**

GROUPWISE REGRESSION

建立在一樣的分組條件下，對組內做出不同的計算處理，進而獲得組內的RESIDUAL SUM OF SQUARE，來建立新的檢定統計量，並與前面所提到F1 與 F2做比較。



GREEN'S PROCEDURE

初始擬合模型： $E(Y) = X\beta$

模型增補：在每個分組內部向擬合模型增加一個 x 的獨立多項式。

替代模型假設： $E(Y) = X_1v + \eta$

假設 M 組以成立下，以一個 q 階多項式去近似， $E(Y) = X\beta$ ；
 $E(Y)$ 每個組別內的應變量期望值 X 為一個設計矩陣 $N \times P$

GREEN'S PROCEDURE

$$GMSE_W' = \frac{\sum_1^M SSE_j}{N - r} \quad Q' = \frac{MSE_B}{GMSE_W}$$




BREIMAN AND MEISEL

數據劃分成子區域 R_1, R_2, \dots, R_M ，並在每個子區域中進行簡單線性回歸。通過將子區域進一步分割成兩個部分，通過在兩個新的子區域中分別進行線性回歸可以獲得更好的擬合。如果F值顯著，就用新的子區域替代原始子區域。如果在M次嘗試中沒有顯著的F值產生，那麼子區域將保持不分割。



BREIMAN AND MEISEL

$$BM_W = \frac{\sum_1^M SSE_j}{N} \quad Q = \frac{MSE_B}{BM_W}$$


檢定統計量

$$Q' = \frac{MSE_B}{GMSE_W}$$

這是基於 Green 使用 q 次多項式的方法，假設組的形成使得真實關係可以透過每個組中的超曲面近似

$$Q = \frac{MSE_B}{BM_W}$$

此統計量根據 Breiman and Meisel (1976) 所提出的 data-splitting algorithm

比較

$$F2 = \frac{MSE_B}{MSE_{W'}}$$

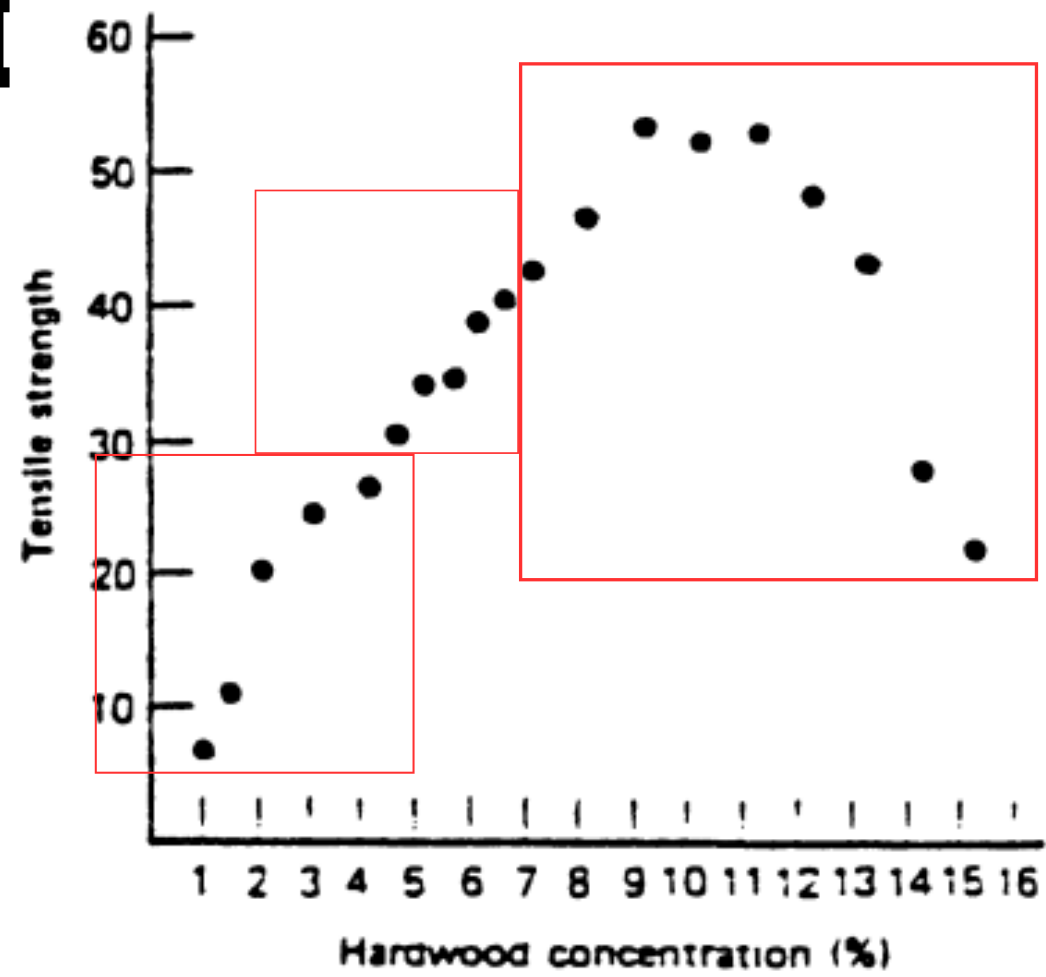
$$Q = \frac{MSE_B}{BM_W}$$

$$Q' = \frac{MSE_B}{GMSE_W}$$

比較分組



I



II

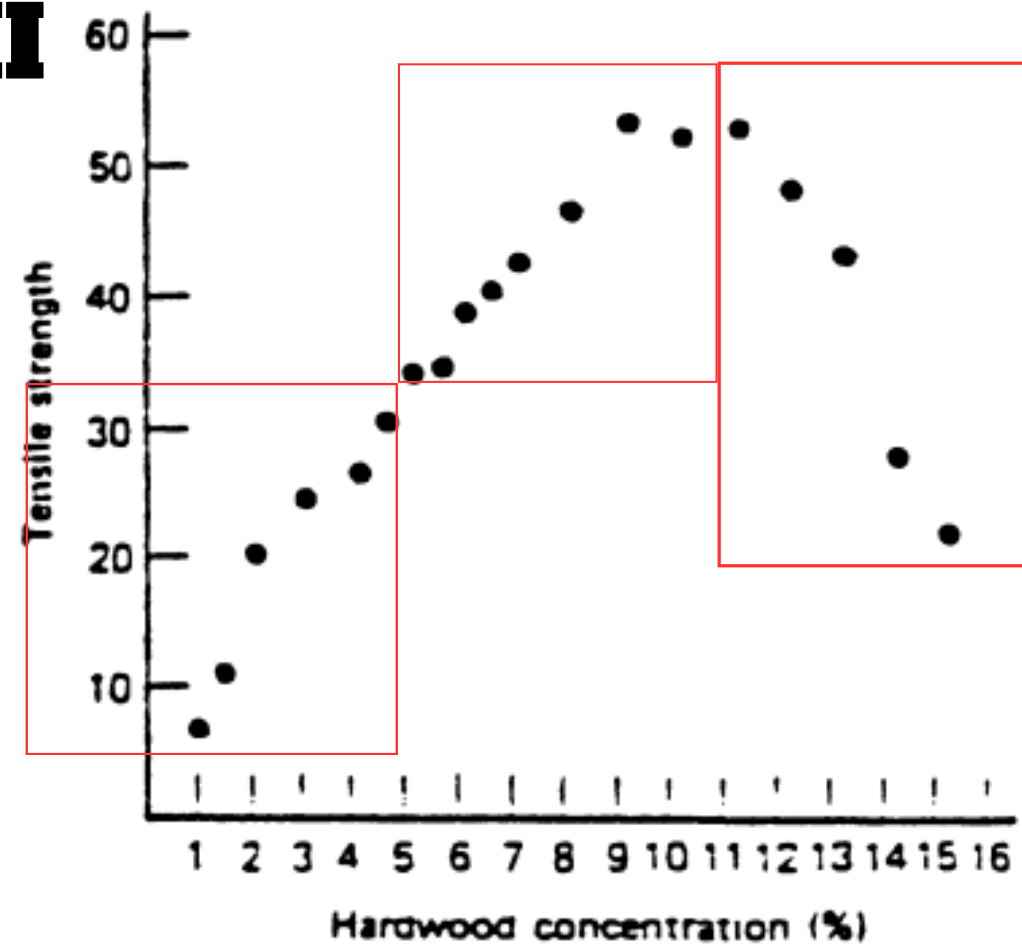


Table 6. Tensile-Strength Data: F Ratios

Grouping	F_1	df	Q	F_2	df	Q'	df
I	2.9949	(1, 15)	7.3138	5.0042	(1, 13)	57.9713	(1, 10)
II	10.6400	(1, 15)	209.0500	143.0358	(1, 13)	198.2443	(1, 10)

分組一

5%的顯著水準下 1%的顯著水準下

F_1 不拒絕 H_0

F_1 不拒絕 H_0

F_2 拒絕 H_0

F_2 不拒絕 H_0

Q 拒絕 H_0

Q 不拒絕 H_0

Q' 拒絕 H_0

Q' 拒絕 H_0

分組二

5%的顯著水準下 1%的顯著水準下

F_1 拒絕 H_0

F_1 拒絕 H_0

F_2 拒絕 H_0

F_2 拒絕 H_0

Q 拒絕 H_0

Q 拒絕 H_0

Q' 拒絕 H_0

Q' 拒絕 H_0

CONCLUSION

雖然Q'在比較中視表現最好的，我們面臨著選擇 M 和 q 的艱難決定。另外，對於涉及不止一個預測變數的模型，GREEN'S PROCEDURE需要大量的觀察值。

作者建議在沒有重複觀察值得情況下，使用Breiman 和 Meisel (1976) 提出的演算法，適合根據數據中存在的各種斜率來分離群。

$$Q = \frac{MSE_B}{BM_W}$$

THANK YOU

Presentation by Alexander Aronowitz

