

MỤC LỤC

BÀI I. GIỚI THIỆU MỘT SỐ MÔ HÌNH SINH TRƯỞNG	3
MÔ HÌNH GOMPERTZ	3
PHƯƠNG PHÁP TÌM CÁC THAM SỐ CỦA MÔ HÌNH GOMPERTZ.....	3
KHẢO SÁT HÀM GOMPERTZ.....	4
MÔ HÌNH SCHUMACHER.....	4
PHƯƠNG PHÁP TÌM CÁC THAM SỐ CỦA MÔ HÌNH SCHUMACHER	4
KHẢO SÁT HÀM SCHUMACHER	5
MÔ HÌNH KORF.....	6
XÁC ĐỊNH NHỮNG THAM SỐ CỦA MÔ HÌNH HỒI QUY PHI TUYẾN	6
THỦ TỤC HỒI QUY PHI TUYẾN (NONLINEAR REGRESSION).....	6
XÁC ĐỊNH NHỮNG THAM SỐ CỦA MÔ HÌNH PHI TUYẾN	9
VÍ DỤ XÁC ĐỊNH CÁC HỆ SỐ CỦA MÔ HÌNH HỒI QUY PHI TUYẾN	9
HÀM GOMPERTS	10
HÀM SCHUMACHER	11
BAI II. TÓM TẮT LÝ THUYẾT PHÂN CHIA CẤP ĐẤT	13
ĐỊNH NGHĨA	13
CHỈ TIÊU PHÂN CHIA CẤP ĐẤT	13
LẬP BIỂU CẤP ĐẤT	14
TUỔI CƠ SỞ	14
Chỉ số Cấp Đất	14
ĐƯỜNG CONG CHI THỊ CHO CẤP ĐẤT.....	14
CẤP ĐẤT TUYỆT ĐỐI VÀ CẤP ĐẤT TƯƠNG ĐỐI.....	14
Biểu Cấp Đất	14
XÁC ĐỊNH SỐ CẤP ĐẤT	15
PHƯƠNG PHÁP XÂY DỰNG ĐƯỜNG CONG CẤP ĐẤT	15
XÂY DỰNG ĐƯỜNG CONG CẤP ĐẤT BÌNH QUÂN CHUNG.....	15
PHƯƠNG PHÁP 1: HỒI QUY TUYẾN TÍNH	15
PHƯƠNG PHÁP 2: HỒI QUY PHI TUYẾN TÍNH	16

PHƯƠNG PHÁP 3: HỒI QUY PHÂN NHÓM	16
XÂY DỰNG ĐƯỜNG CONG CẤP ĐẤT BÌNH QUÂN CHO MỖI CẤP ĐẤT (H0I – A)	17
PHƯƠNG PHÁP PHÂN CHIA CẤP ĐẤT	17
PHƯƠNG PHÁP CỐ ĐỊNH ĐIỂM CHẶN, THAY ĐỔI ĐỘ DỐC THEO MỖI CẤP ĐẤT	17
PHƯƠNG PHÁP CỐ ĐỊNH ĐỘ DỐC, THAY ĐỔI ĐIỂM CHẶN THEO MỖI CẤP ĐẤT	17
PHƯƠNG PHÁP AFFILL.....	18
PHƯƠNG PHÁP XÂY DỰNG CÁC ĐƯỜNG CONG SINH TRƯỞNG CHIỀU CAO ĐỘC LẬP THEO TỪNG CẤP ĐẤT.....	18
PHƯƠNG PHÁP XÂY DỰNG CÁC ĐƯỜNG CONG CẤP ĐẤT THÔNG QUA SUẤT TĂNG TRƯỞNG	19
XÁC ĐỊNH CHỈ SỐ CẤP ĐẤT CHO LÂM PHẦN ĐIỀU TRA	20
TÓM TẮT PHƯƠNG PHÁP CHUNG LẬP BIỂU CẤP ĐẤT.....	20
THỰC HÀNH XÂY DỰNG BIỂU CẤP ĐẤT	21
ỨNG DỤNG MÔ HÌNH LOGISTIC TRONG SINH THÁI QXTV.....	31
QUAN HỆ GIỮA ĐỘ BẮT GẬP LOÀI CÂY GỖ VỚI CÁC YẾU TỐ MÔI TRƯỜNG	31
HỒI QUY LOGIT ĐA BIẾN	33
THỦ TỤC XỬ LÝ MÔ HÌNH LOGISTIC	34
<i>Đường cong sigmoid</i>	34
<i>Hồi quy Logit Gauss</i>	35
<i>Hồi quy Logit đa biến</i>	35
ỨNG DỤNG MÔ HÌNH POISSON TRONG SINH THÁI QXTV	36
QUAN HỆ GIỮA TỔ THÀNH CÂY GỖ VỚI CHIỀU CAO TÁN RỪNG.....	36
QUAN HỆ GIỮA TỔ THÀNH CÂY GỖ VỚI ĐỘ CAO ĐỊA HÌNH.....	37
XÁC ĐỊNH Tỷ Lệ CÂY GỖ QUÝ THEO CẤP ĐƯỜNG KÍNH.....	38
XÁC ĐỊNH Tỷ Lệ CÂY TÁI SINH THEO LỚP CHIỀU CAO.....	38
XÁC ĐỊNH Tỷ Lệ CÂY GỖ BỊ RỒNG RUỘT THEO CẤP ĐƯỜNG KÍNH	38
NHỮNG MÔ HÌNH POISSON CẦN PHÂN TÍCH	39

Bài I. GIỚI THIỆU MỘT SỐ MÔ HÌNH SINH TRƯỞNG

1.1. MỘT SỐ HÀM SINH TRƯỞNG THÔNG DỤNG

1.1.1. Mô hình Gompertz

$$Y = m \cdot \exp[-b \cdot \exp(-cT)]; \quad (1)$$

trong đó: T là tuổi cây; $m = Y_{\max}$ – chiều cao (hoặc đường kính, V...) thân cây lớn nhất ở tuổi thành thực.

+ **Phương pháp tìm các tham số của mô hình Gompertz**

+ **Cách thứ nhất**

$$\text{Đặt } Y/m = \exp[-b \cdot \exp(-cT)] \Rightarrow \ln(m/Y) = -b \cdot \exp(-cT)$$

$$\Rightarrow \ln(-\ln(Y/m)) = \ln b - cT$$

$$\text{Đặt } Y' = \ln(-\ln(Y/m)); A = \ln b; c = B.$$

$$\Rightarrow Y' = A + BT. \quad (2)$$

Giải phương trình (2) bằng phương pháp bình phương nhỏ nhất để tìm A và B. Khi biết $A = \ln b \Rightarrow b = e^A$. Từ đây có thể trở lại hàm phi tuyến.

+ **Cách thứ hai**

$$\text{Đặt } m/Y = \exp[-b \cdot \exp(-cT)] \Rightarrow \ln(m/Y) = -b \cdot \exp(-cT)$$

$$\Rightarrow \ln(\ln(m/Y)) = \ln b - cT$$

$$\text{Đặt } Y' = \ln(\ln(m/Y)); A = \ln b; c = B.$$

$$\Rightarrow Y' = A + BT. \quad (2)$$

Giải phương trình (2) bằng phương pháp bình phương nhỏ nhất để tìm A và B. Khi biết $A = \ln b \Rightarrow b = e^A$. Từ đây có thể trở lại hàm phi tuyến.

Lưu ý rằng, trong thực tế khi giải mô hình (1) bằng cách cho trước m và c, thì sai số hay $\sum(Y_{tn} - Y_{lt})^2$ vẫn còn rất lớn. Để đạt được tiêu chuẩn $\text{MinSum}(Y_{tn} - Y_{lt})^2$, cần phải xác định ba tham số m, b và c của mô hình (1) bằng phương pháp hồi quy tương quan phi tuyến tính. Thủ tục hồi quy tương quan phi tuyến tính được trình bày ở File (*Những hàm sinh trưởng*) thuộc nhóm [Đề cương nghiên cứu khoa học] và [Hướng dẫn sử dụng Statgraphics].

Khảo sát hàm Gompertz : $Y = m \cdot \exp(-b \cdot \exp(-cT))$

+ **Tốc độ sinh trưởng** (lượng tăng trưởng thường xuyên hàng năm CAI = Current Annual Increment hoặc PAI = Periodic Annual Increment):

$$ZY = Y' = mbc \cdot \exp(-cT - b \cdot \exp(-cT))$$

+ **Tốc độ sinh trưởng lớn nhất**

$$ZY_{\max} = Y'_{\max} = (mc)/e ; T = \ln(b)/c$$

+ **Điểm uốn** $Y'' = 0$ tại $T = \ln(b)/c$ và $Y = m/e$

+ **Tuổi đạt $ZY_{\max}$**

$$T_{(Y'_{\max})} = \ln(b)/c$$

+ **Lượng tăng trưởng trung bình hàng năm** (MAI = Mean Annual Increment):

$$\Delta Y = \frac{Y_T}{T} = \frac{m \cdot \exp(-b \cdot \exp(-cT))}{T}$$

$$\Delta Y = 0 \text{ tại } T = -\text{vô cùng}$$

+ **Suất tăng trưởng**

$$P_Y = 100 \cdot \frac{Y'}{Y} = 100 \cdot bc \cdot \exp(-cT)$$

1.1.2. Mô hình Schumacher

$$Y = m \cdot \exp(-b/T^c) \text{ hay } Y = m \cdot \exp(-bT^{-c}) \quad (3)$$

trong đó: T – tuổi rừng; $m = Y_{\max}$ – chiều cao (hoặc đường kính...) thân cây lớn nhất ở tuổi thành thực. Lưu ý: $c = 0,2 \div 2$; b luôn lấy giá trị âm.

+ **Phương pháp tìm các tham số của mô hình Schumacher**

+ **Cách thứ nhất.** Đặt $(Y/m) = \exp[(-bT^{-c})]$

$$\Rightarrow \ln(Y/m) = -bT^{-c} \Rightarrow \ln(-\ln(Y/m)) = \ln b - c \ln T.$$

$$\text{Đặt } Y' = \ln(-\ln(Y/m)); \ln b = A; c = B; \ln T = T'$$

$$\Rightarrow Y' = A + BT'. \quad (4)$$

Giải phương trình (4) bằng phương pháp bình phương nhỏ nhất sẽ tìm được A và B . Khi biết $A = \ln b \Rightarrow b = e^A$; $\ln T = T' \Rightarrow T = e^{T'}$. Từ đây có thể trở lại hàm phi tuyến.

Lưu ý rằng, giá trị m được chọn sao cho tổng bình phương các sai lệch giữa giá trị của biến y_i và giá trị ước lượng của biến y_i là nhỏ nhất, nghĩa là $\sum (y_i - \hat{y}_i)^2 = \min$.

+ **Cách thứ hai.** Từ hàm 3, lấy logarit tự nhiên hai vế:

$$\ln Y = \ln m - bT^c$$

$$\text{Đặt } \ln Y = Y'; \ln m = A; -b = B; 1/T^c = T'; c = 0,2 - 2,0.$$

$$\Rightarrow Y' = A + BT' \quad (5)$$

Lưu ý rằng, giá trị c được chọn sao cho $\sum (y_i - \hat{y}_i)^2 = \min$.

+ **Cách thứ ba.** Ba tham số m , b và c của mô hình (3) có thể được xác định bằng phương pháp hồi quy tương quan phi tuyến tính (Xem File *(Những hàm sinh trưởng)* thuộc nhóm ‘Đề cương nghiên cứu khoa học’ và ‘Hướng dẫn sử dụng Statgraphics’).

Khảo sát hàm Schumacher

$$Y = m \cdot \exp(-b/T^c) \text{ hay } Y = m \cdot \exp(-bT^{-c})$$

+ **Tốc độ sinh trưởng** (lượng tăng trưởng thường xuyên hàng năm):

$$ZY = Y' = mbcT^{-(c+1)} \exp(-bT^{-c})$$

+ **Tốc độ sinh trưởng lớn nhất**

$$ZY_{\max} = Y'_{\max} = (m \cdot (bc)^{-(1/c)}) \cdot (e/1+c)^{-((c+1)/c)}$$

Nếu tính cho M lâm phần, thì $ZY_{\max}$ là năng suất tối đa.

+ **Tuổi đạt $ZY_{\max}$**

$$T_{(Y'_{\max})} = (bc/(1+c))^{1/c}$$

$$\text{Tại } Y = m \cdot \exp(-(c+1)/c)$$

+ **Lượng tăng trưởng trung bình hàng năm**

$$\Delta Y = \frac{Y_T}{T} = \frac{me^{-(b/T^c)}}{T}$$

+ **Lượng tăng trưởng trung bình hàng năm lớn nhất**

$$\Delta Y_{\max} = m(bce)^{-(1/c)}$$

+ **Tuổi đạt $\Delta Y_{\max}$** (với M_{sf} là tuổi thành thực số lượng)

$$T_{\max} = (cb)^{1/c}$$

+ **Suất tăng trưởng**

$$P_Y = 100 \cdot \frac{Y'}{Y} = 100 \cdot bcT^{c-1}$$

+ **Điểm uốn** tại $T = ((c \cdot b)/(c+1))^{(1/c)}$ và $Y = m \cdot 2.7182^{(-(c+1)/c)}$

Vì hàm Schumacher nhận $Y = 0$ tại $X = 0$, nên nó biểu thị tốt quy luật biến đổi của D, H, M lâm phần theo tuổi.

1.1.3. Mô hình Korf

$$Y = Y_{\max} \exp(-C_1 T^{-C_2}) \quad (6)$$

trong đó: T – tuổi rừng; $Y_{\max}$ – chiều cao (hoặc đường kính) thân cây lớn nhất ở tuổi thành thực. Các tham số của phương trình Korf có quan hệ với nhau như sau: $Y_{\max} = e^{C_1}$, $C_1 = b(n-1)$, $C_2 = n-1$.

Giải: Đặt $(Y/Y_{\max}) = \exp(-C_1 T^{-C_2}) \Rightarrow -\ln(Y/Y_{\max}) = C_1 T^{-C_2}$

$$\Rightarrow \ln(-\ln(Y/Y_{\max})) = \ln C_1 - C_2 \ln T$$

$$\Rightarrow \text{Đặt } \ln(-\ln(Y/Y_{\max})) = Y'; \ln T = X, \ln C_1 = A, C_2 = B$$

$$\Rightarrow Y = A + BX. \quad (7)$$

Giải phương trình (7) bằng phương pháp bình phương nhỏ nhất sẽ tìm được A và B . Khi biết $A \Rightarrow C_1 = 2.7182^A$; $C_2 = B$. Từ đây có thể trở lại hàm phi tuyến.

Để làm phù hợp những hàm số trên đây với dãy số liệu quan sát, yêu cầu phải xác định được những tham số của mô hình sao cho “Tổng sai lệch bình phương giữa những trị số quan sát (y_i) và trị số ước lượng (y_i^*) là nhỏ nhất, nghĩa là $\min S(y_i - y_i^*)^2$.”

Các tham số của những mô hình phi tuyến có thể xác định bằng hai phương pháp: (1) hồi quy tuyến tính (Linear Regression), (2) hồi quy phi tuyến tính (Nonlinear Regression). Thủ tục hồi quy tuyến tính đã được giới thiệu trong nhiều tài liệu thống kê toán học. Phần dưới đây giới thiệu phương pháp xác định các tham số của mô hình phi tuyến bằng thủ tục hồi quy phi tuyến (Nonlinear Regression).

1.2. XÁC ĐỊNH NHỮNG THAM SỐ CỦA MÔ HÌNH HỒI QUY PHI TUYẾN

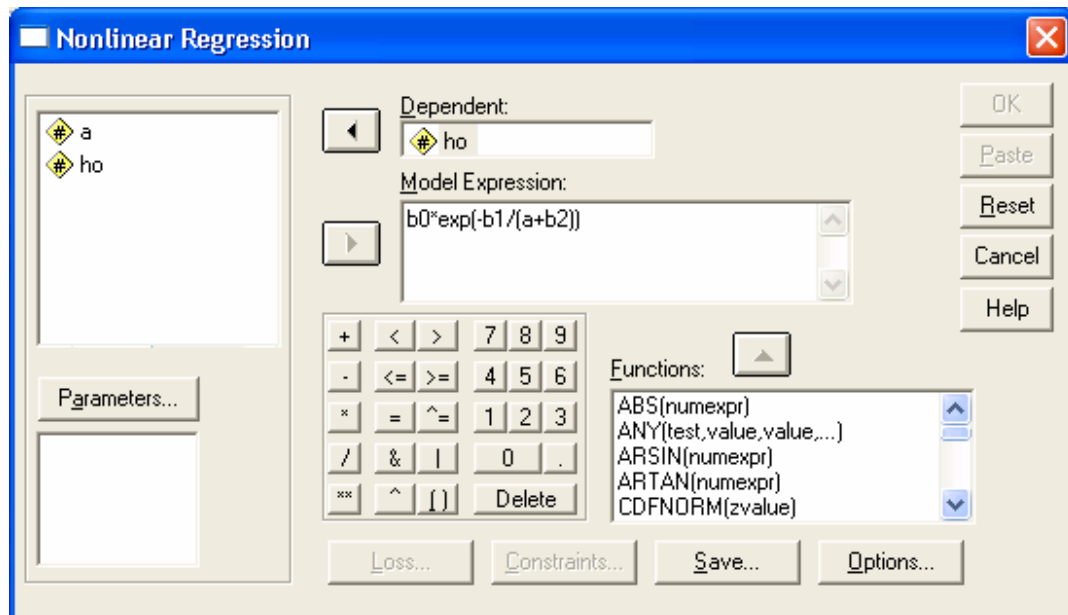
Thủ tục hồi quy phi tuyến trong phần mềm thống kê SPSS (10.0) cho phép tìm các hệ số của mô hình phi tuyến mà không cần phải thực hiện bước tuyến tính hoá các mô hình phi tuyến.

1.2.1. Thủ tục hồi quy phi tuyến (Nonlinear Regression)

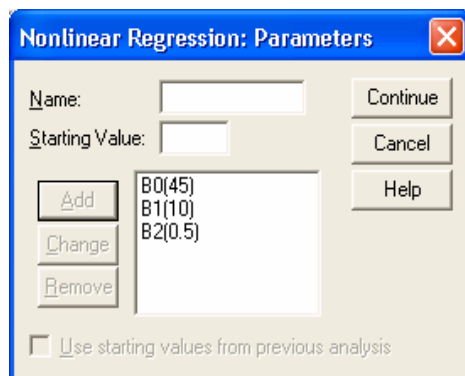
Trước hết vào đường lệnh hồi quy phi tuyến:

Analyze\Regression\Nonlinear.

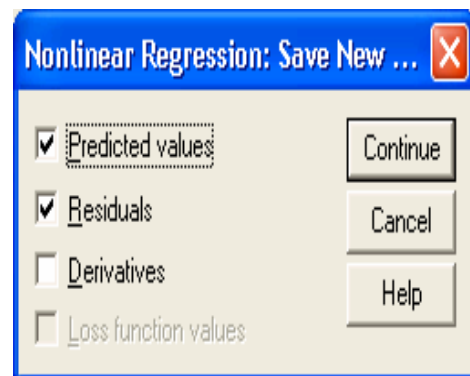
Kể đến khai báo những điều kiện của mô hình phi tuyến như sau:



Hình 1. Cửa sổ khai báo hàm phi tuyến



Hình 2. Khai báo tham số ban đầu



Hình 3. Chọn giá trị dự đoán và sai lệch

- Biến phụ thuộc Y. Biến này được khai báo trong cửa sổ Dependent bằng cách chọn Y và chuyển vào cửa sổ Dependent (hình 1).
- Dạng mô hình được chọn để làm phù hợp với dãy số liệu. Dạng mô hình được ghi trong cửa sổ Model expression (hình 1).
- Những tham số giả định ban đầu của mô hình. Những tham số này được khai báo trong cửa sổ Parameters (hình 2). Lưu ý chọn Add sau mỗi lần khai báo một

tham số để chương trình ghi nhận. Nếu khai báo sai hoặc có thay đổi tham số, thì chọn Remove và khai báo lại.

- Chọn những giá trị dự đoán (Predicted value) và những sai lệch (Residuals) trong thanh Save (hình 3).
- Cuối cùng chọn OK.

Kết quả báo cáo những thống kê tóm tắt của hồi quy phi tuyến như phân tích ANOVA, hệ số tương quan và các tham số của mô hình.

Lưu ý rằng, việc khai báo những giá trị ban đầu của các tham số của mô hình phi tuyến là rất quan trọng, nhưng việc này cũng rất khó khăn. Bởi vì, nếu khai báo gần đúng những giá trị ban đầu của các tham số, thì thủ tục hồi quy phi tuyến cho phép dự đoán những giá trị của biến phụ thuộc chính xác hơn [nghĩa là $\min S(y_i - y_i^*)^2$ nhỏ hơn] so với thủ tục hồi quy tuyến tính. Ngược lại, nếu những giá trị ban đầu của các tham số được khai báo không đúng hoặc chênh lệch quá lớn so với giá trị gần đúng nhất, thì thủ tục hồi quy phi tuyến hoặc báo lỗi hoặc báo cáo những giá trị dự đoán với sai lệch rất lớn.

Mục đích phân tích hồi quy trong lâm học:

- (1) Tìm được một hàm số phản ánh đúng quy luật biến đổi của các nhân tố điều tra.
- (2) Tìm được một hàm số để mô tả đúng mối quan hệ giữa các nhân tố điều tra với các biến môi trường.
- (3) Tìm được một hàm số để làm phù hợp số liệu và sử dụng nó để dự đoán sao cho tổng sai lệch bình phương nhỏ nhất, nghĩa là $\min S(y_i - y_i^*)^2$.

Nếu mục đích đặt ra là tìm được một hàm phi tuyến để làm phù hợp số liệu với sai lệch dự đoán nhỏ nhất, bạn có thể sử dụng thủ tục hồi quy phi tuyến. Tuy vậy, như đã nói ở trên, mục đích này chỉ có thể đạt được khi chúng ta khai báo gần đúng nhất những giá trị ban đầu của các tham số. Vậy bằng cách nào có thể khai báo gần đúng nhất những giá trị ban đầu của các tham số của hàm phi tuyến? Đây là một công việc rất khó khăn, nhất là đối với những ai thiếu kinh nghiệm. Dưới đây trình bày cách khai báo gần đúng nhất những giá trị ban đầu của các tham số của mô hình phi tuyến.

1.2.2. Xác định những tham số của mô hình phi tuyến

Bước 1. Chọn những hàm phi tuyến để làm phù hợp với số liệu đã cho. Công việc này có thể thực hiện bằng nhiều cách khác nhau, chẳng hạn như dựa vào lý thuyết và kinh nghiệm, sử dụng thủ tục so sánh các hàm hồi quy trong Statgraphics và SPSS, vẽ đồ thị...

Bước 2. Biến đổi hàm phi tuyến sang dạng tuyến tính và tìm những hệ số của hàm phi tuyến bằng thủ tục hồi quy tuyến tính.

Bước 3. Sử dụng những giá trị của những tham số vừa tìm được ở bước 2 để khai báo trong cửa sổ các tham số (parameters) của thủ tục hồi quy phi tuyến.

Cần lưu ý rằng, khi những tham số được giả định bằng những giá trị ban đầu khác nhau, thì các hàm phi tuyến sẽ nhận những hệ số khác nhau và tổng sai lệch bình phương cũng khác nhau. Do đó, để nhận được những tham số của hàm phi tuyến hợp lý nhất, bạn cần chú ý đến điều kiện “ $\min S(y_i - y_i^*)^2$ ”. Để thỏa mãn điều kiện $\min S(y_i - y_i^*)^2$, bạn cần phải khai báo lại một vài lần những giá trị ban đầu của các tham số của hàm phi tuyến. Mỗi lần khai báo lại, bạn chỉ cần thay đổi một vài đơn vị phần mười của mỗi tham số. Thủ tục sẽ dừng khi bạn đã nhận được $\min S(y_i - y_i^*)^2$.

Ví dụ xác định các hệ số của mô hình hồi quy phi tuyến

Bảng 1 ghi lại những số liệu về sinh trưởng đường kính thân cây (D, cm) của loài Keo lá tràm (*Acasia auriculiformis*) từ 2 – 21 tuổi, mọc trên đất xám phù sa cổ ở miền Đông Nam Bộ. Từ số liệu này, yêu cầu đặt ra là làm phù hợp số liệu đã cho với hàm Gomperts và hàm Schumacher để mô tả mối quan hệ giữa đường kính thân cây (D, cm) với tuổi cây (A, năm). Để đạt được yêu cầu này, dưới đây thử sử dụng hai thủ tục hồi quy tuyến tính và hồi quy phi tuyến để xác định các tham số của hàm Gomperts và hàm Schumacher.

Bảng 1. Sinh trưởng đường kính thân cây (D, cm) của loài

Keo lá tràm (*Acasia auriculiformis*) từ 2 – 21 tuổi

A, năm	2	3	4	5	6	7	8	9	10
D, cm	4.73	6.19	7.51	8.84	10.12	11.31	12.26	13.29	14.21
A, năm	11	12	13	14	15	16	17	18	19
D, cm	15.16	16.01	16.76	17.5	18.28	19.17	20	20.61	21.47
A, năm	20	21							
D, cm	22.3	23.1							

Giải:

Hàm Gomperts có dạng:

$$Y = Y_{\max} \cdot \exp(-b_0 \cdot \exp(-b_1 \cdot A))$$

trong đó:

$Y_{\max}$ = D lớn nhất ở tuổi thành thực, cm;

Exp = cơ số Neper = 2,7182;

b_0 và b_1 = những tham số của mô hình cần tìm;

A = biến số tuổi cây, năm.

Bước 1. Xác định các tham số của mô hình Gomperts bằng thủ tục hồi quy tuyến tính.

Trước hết, giả định $Y_{\max} = 70$ cm và bằng thủ tục tuyến tính hoá, đã xác định được hai tham số của hàm Gomperts như sau: $b_0 = -2.57685$; $b_1 = 0.04284$. Như vậy, mô hình biến đổi đường kính thân cây theo tuổi có dạng:

$$D_{1.3} = 70 \cdot \exp(-2.57685 \cdot \exp(-0.04284 \cdot A)) \quad (1)$$

Tổng sai lệch bình phương của mô hình (1) là 17,542.

Bước 2. Xác định các tham số của hàm Gomperts bằng thủ tục hồi quy phi tuyến.

Trước hết, chuyển dữ liệu D và A vào bảng tính SPSS 10.0. Kế đến mở đường dẫn vào thủ tục hồi quy phi tuyến: Analyze > Regression > Nonlinear. Tiếp theo chọn Y và chuyển vào khoảng Dependent. Sau đó khai báo mô hình Gomperts trong khoảng Model expression theo dạng: $Y_{\max} \cdot \exp(-b_0 \cdot \exp(-b_1 \cdot A))$. Cuối cùng khai báo những giá trị ban đầu của các tham số của hàm Gomperts trong thanh Parameters. Để tìm các tham số của hàm Gomperts và so sánh với cách xác định các tham số bằng thủ tục tuyến tính hoá, ở đây chúng ta thử ba giả định: (1) $Y_{\max} = 70$, $b_0 = -3.0$ và $b_1 = -0.3$; (2) $Y_{\max} = 70$, $b_0 = 2.0$ và $b_1 = 0.2$; (3) $Y_{\max} = 70$, $b_0 = 2.5$ và $b_1 = 0.04$ (những tham số gần giống với thủ tục tuyến tính).

Sau khi thực hiện thủ tục hồi quy phi tuyến, kết quả nhận được các tham số $Y_{\max}$, b_0 , b_1 và tổng bình phương sai lệch như sau:

✓ Đối với trường hợp 1, $Y_{\max} = 27.3938$, $b_0 = 2.0172$, $b_1 = 0.1111$, $r = 0.996$.

Tổng sai lệch bình phương nhỏ nhất là 0.3172.

- ✓ Đối với trường hợp 2, $Y_{\max} = 27.39375$, $b_0 = 2.01713$, $b_1 = 0.11111$, $r = 0.996$.
Tổng sai lệch bình phương nhỏ nhất là 0.3301.
- ✓ Đối với trường hợp 3, $Y_{\max} = 27.39377$, $b_0 = 2.01713$, $b_1 = 0.11111$, $r = 0.996$.
Tổng sai lệch bình phương nhỏ nhất là 2.1430.

So sánh những kết quả tính toán trên đây cho thấy, khi sử dụng thủ tục hồi quy phi tuyến tính để làm phù hợp số liệu với hàm Gomperts, thì tổng sai lệch bình phương ở trường hợp 1 và 2 (tương ứng 0.3172 và 0.3301) nhỏ hơn so với thủ tục hồi quy tuyến tính (17.542) tương ứng là 55 và 53 lần.

Hàm Schumacher có dạng:

$$Y = Y_{\max} * \exp(-b_0 * A^{-b_1})$$

trong đó:

Y = biến số phụ thuộc;

$Y_{\max}$ = giá trị lớn nhất của biến phụ thuộc;

Exp = cơ số Neper = 2,7182;

b_0 và b_1 = những tham số của mô hình;

A = tuổi cây.

Bước 1. Xác định các tham số của hàm Schumacher bằng thủ tục hồi quy tuyến tính

Trước hết, giả định $Y_{\max} = 70$ cm và bằng thủ tục tuyến tính hoá, đã xác định được hai tham số của hàm Schumacher như sau: $b_0 = -3.76679$; $b_1 = -0.38444$. Như vậy, mô hình biến đổi đường kính thân cây theo tuổi có dạng:

$$D_{1,3} = 70 * \exp(-3.76679 * A^{-0.38444}) \quad (2)$$

Tổng sai lệch bình phương của theo mô hình (2) là 6,67.

Bước 2. Xác định các tham số của hàm Schumacher bằng thủ tục hồi quy phi tính.

Để tìm các tham số của hàm Schumacher và so sánh với cách xác định các tham số bằng thủ tục tuyến tính hoá, ở đây thử ba giả định: (1) $Y_{\max} = 70$, $b_0 = 4.0$ và $b_1 = 0.4$ (những tham số gần giống với thủ tục tuyến tính); (2) $Y_{\max} = 70$, $b_0 = -3.0$ và $b_1 = -0.3$; (3) $Y_{\max} = 70$, $b_0 = 2.0$ và $b_1 = 0.2$.

Sau khi thực hiện thủ tục hồi quy phi tuyến, kết quả nhận được các tham số $Y_{\max}$, b_0 , b_1 và tổng bình phương sai lệch như sau:

- ✓ Đối với trường hợp 1, $Y_{\max} = 1159.0284$, $b_0 = -6.2011$, $b_1 = 0.1494$. Tổng sai lệch bình phương nhỏ nhất là 0.622.
- ✓ Đối với trường hợp 2, $Y_{\max} = 815.7010$, $b_0 = -5.8746$, $b_1 = 0.116236$. Tổng sai lệch bình phương nhỏ nhất là 0.729.
- ✓ Đối với trường hợp 3, $Y_{\max} = 638.5037$, $b_0 = -5.6512$, $b_1 = 0.1726$. Tổng sai lệch bình phương nhỏ nhất là 0.842.

So sánh những kết quả tính toán trên đây cho thấy, khi sử dụng thủ tục hồi quy phi tuyến tính để làm phù hợp số liệu với hàm Schumacher, thì tổng sai lệch bình phương ở trường hợp 1 và 2 (tương ứng 0.622 và 0.729) nhỏ hơn so với thủ tục hồi quy tuyến tính (6.67) tương ứng là 11 và 9 lần.

Những lưu ý

- ✓ Khi những tham số được giả định bằng những giá trị ban đầu khác nhau, thì các hàm phi tuyến sẽ nhận những hệ số khác nhau.
- ✓ Ngoài ra, đối với hai hàm Gomperts và Schumacher, nếu cố định hệ số b_0 và b_1 và thay đổi giá trị $Y_{\max}$ lớn dần, thì tổng sai lệch bình phương sẽ nhận giá trị nhỏ dần.
- ✓ Vì thế, để nhận được hàm phi tuyến mô tả tốt mối quan hệ giữa hai biến Y và X với tổng sai lệch bình phương nhỏ nhất, bạn có thể giả định những giá trị ban đầu của các tham số bằng những giá trị khác nhau. Thủ tục sẽ dừng khi mô hình ước lượng nhận được những giá trị dự đoán với sai số thoả mãn những điều kiện đặt ra.

Bài II. TÓM TẮT LÝ THUYẾT PHÂN CHIA CẤP ĐẤT

2.1. ĐỊNH NGHĨA. Cấp đất là chỉ tiêu biểu thị năng suất tiềm năng của lập địa.

- ✓ Cấp đất được sử dụng để phân chia các đơn vị dự đoán sản lượng và xác định hệ thống các biện pháp kinh doanh.
- ✓ Cấp đất được xây dựng cho quần thụ thuần loài đồng tuổi.
- ✓ Đối với rừng tự nhiên khác loài khác tuổi, cấp đất biểu thị cấp năng suất tương ứng với từng cấp chiều cao.

2.2. CHỈ TIÊU PHÂN CHIA CẤP ĐẤT (rừng thuần loài đồng tuổi)

+ **Yêu cầu:**

- ✓ Phản ánh tốt sự phù hợp của điều kiện lập địa đối với sinh trưởng của loài cây.
- ✓ Quan hệ chặt chẽ với trữ lượng quần thụ.
- ✓ Về cơ bản độc lập với mật độ.
- ✓ Không chịu ảnh hưởng của tia thưa.
- ✓ Xác định nhanh tại hiện trường.

+ **Chỉ tiêu phân chia cấp đất:** chiều cao.

- ✓ Quan hệ chặt với trữ lượng.
- ✓ Dễ xác định.
- ✓ Chịu ảnh hưởng của lập địa.
- ✓ Ít chịu ảnh hưởng của biện pháp kinh doanh.
- ✓ $H < H_d < H_g < H_z < H_0$

Ở đây: H = chiều cao bình quân thân cây; H_d = chiều cao thân cây tương ứng với đường kính bình quân lâm phần; H_g = chiều cao tương ứng với cây có tiết diện ngang bình quân; H_z = chiều cao Lorey; H_0 = chiều cao tầng trội.

- Khi rừng không bị tia thưa nhân tạo, có thể sử dụng H bình quân lâm phần để phân chia cấp đất.
- Khi rừng đã bị tia thưa nhân tạo, người ta sử dụng H bình quân của 20% số cây thuộc tầng trội (H_0) để phân chia cấp đất.

2.3. LẬP BIỂU CẤP ĐẤT

2.3.1. Một số khái niệm

(a) Tuổi cơ sở (A_0). Đó là tuổi được sử dụng để xác định số cấp đất cần phân chia cho mỗi loài, phạm vi biến động H và chỉ số cấp đất. Tuổi A_0 được chọn ở thời điểm sinh trưởng chiều cao đã khá ổn định. Nói chung, A_0 được chọn ở tuổi càng cao càng tốt. Đối với cây mọc nhanh, chu kỳ kinh doanh 10 - 20 năm, $A_0 = 8$ đến 16 năm.

Nguyên tắc chọn tuổi cơ sở A_0 :

- ✓ Tuổi thu được nhiều thông tin nhất.
- ✓ Dễ xác định phạm vi biến động H .

(b) Chỉ số cấp đất (S). Đó là chiều cao (H_0) cho trước tương ứng với từng cấp đất tại tuổi A_0 . Chỉ số cấp đất (S) thường lấy giá trị chẵn. Tùy theo sự biến đổi H_0 tại A_0 mà khoảng cách S_i lớn hay nhỏ.

(c) Đường cong cấp đất hay đường cong chỉ thị cho cấp đất. Đó là đường cong biểu thị quá trình sinh trưởng chiều cao bình quân tầng trội ($H_{20\%}$, H_{100} , H_g) theo tuổi tương ứng với từng cấp đất.

Nguyên tắc xây dựng đường cong cấp đất:

- ✓ Đi qua giá trị H_0 cho trước tại A_0 .
- ✓ Xuất phát từ gốc tọa độ.

(d) Cấp đất tuyệt đối và cấp đất tương đối. Cấp đất tuyệt đối là tên gọi cấp đất theo chỉ số cấp đất. Cấp đất tương đối là tên gọi cấp đất theo chữ số La Mã: I...V.

(e) Biểu cấp đất. Đó là một loại biểu ghi lại những giá trị chiều cao ($H_{0\text{dưới}}$, $H_{0\text{giữa}}$, $H_{0\text{trên}}$) của từng tuổi hay cấp tuổi tương ứng với những cấp đất khác nhau.

(g) Biểu đồ cấp đất. Đó là biểu đồ mô tả (biểu diễn) quá trình sinh trưởng chiều cao H_0 ($H_{0\text{dưới}}$, $H_{0\text{giữa}}$, $H_{0\text{trên}}$) của từng loài cây tương ứng với những cấp đất khác nhau.

(h) Cách sử dụng biểu cấp đất

- ✓ Xác định A và H_0 cho quần tiêu thụ.
- ✓ Điểm giao nhau giữa H_0 với A cho biết cấp đất.

2.3.2. XÁC ĐỊNH SỐ CẤP ĐẤT

- ✓ Vẽ các đường cong sinh trưởng chiều cao thực nghiệm lên biểu đồ.
- ✓ Xác định phạm vi biến động chiều cao tại A_i (giới hạn trên và dưới).
- ✓ Xác định A_0 . Đó là tuổi mà sinh trưởng chiều cao có khuynh hướng ổn định nhất. A_0 lấy giá trị càng cao càng tốt.
- ✓ Căn cứ phạm vi biến động H_0 tại A_0 để phân chia số cấp đất.
- ✓ Số cấp đất biến đổi từ 3 – 5.
- ✓ Xác định chỉ số cấp đất (S_i). Để xác định S_i , trước hết tính quan hệ $H_0 - A$ tương ứng với những cấp đất dự đoán trước. Sau đó, dựa vào H_0 tại A_0 để chọn S_i thích hợp. Giá trị S_i thường được chọn là những số chẵn.

2.3.3. PHƯƠNG PHÁP XÂY DỰNG ĐƯỜNG CONG CẤP ĐẤT

2.3.3.1. Nguyên tắc chung

- ✓ Trước hết, xây dựng những đường cong cấp đất bằng nhiều phương pháp khác nhau. Chẳng hạn phương pháp đồ thị, mô hình thống kê $H_0 - A \dots$
- ✓ Sau đó, chọn phương pháp thích hợp nhất.

Một phương pháp là thích hợp khi:

- ✓ Đại diện tốt cho quy luật sinh trưởng chiều cao của các cấp đất.
- ✓ Xuất phát từ gốc tọa độ.
- ✓ Độ dốc của đường cong tăng dần từ cấp đất xấu (V) đến tốt (I).
- ✓ Mỗi đường cong phải đi qua một điểm cho trước tại A_0 (đó là chỉ số cấp đất (S_i) cho trước tại A_0).

2.3.3.2. Phương pháp xây dựng đường cong cấp đất

(1) Xây dựng đường cong cấp đất bình quân chung

Đường cong cấp đất bình quân chung là đường cong mô tả quy luật sinh trưởng H_0 bình quân của tất cả các lâm phần trên những cấp đất khác nhau.

Có ba phương pháp lập đường cong cấp đất bình quân chung:

+ Phương pháp 1: Hồi quy tuyến tính

- ✓ Chọn mô hình thích hợp theo tiêu chuẩn $\min \sum (Y_i - Y_{\text{ilt}})^2$ để mô tả $H_0 - A$.
- ✓ Biến đổi mô hình về dạng tuyến tính: $y = a + bx$.
- ✓ Tìm các hệ số hồi quy bằng phương pháp bình phương nhỏ nhất.
- ✓ Chuyển phương trình tuyến tính về dạng chính tắc.

- ✓ Lập biểu và vẽ biểu đồ đường cong cấp đất bình quân chung.

+ Phương pháp 2: Hồi quy phi tuyến tính

- ✓ Chọn mô hình thích hợp theo tiêu chuẩn $\min \sum (Y_i - Y_{ilt})^2$ để mô tả $H_0 - A$.
- ✓ Tìm các hệ số hồi quy bằng phương pháp phi tuyến tính.
- ✓ Lập biểu và vẽ biểu đồ đường cong cấp đất bình quân chung.

+ Phương pháp 3: Hồi quy phân nhóm

Để mô tả quan hệ $H_0 - A$ bình quân cho tất cả các cấp đất, có thể sử dụng phương pháp hồi quy phân nhóm của Alders (1980). Thủ tục thực hiện như sau:

- ✓ Phân chia sơ bộ chiều cao H_0 của các lâm phần vào i cấp đất (nhóm).
- ✓ Chọn dạng phương trình thích hợp để mô tả quan hệ $H_0 - A$.
- ✓ Chuyển mô hình về dạng tuyến tính và tính các tham số của mô hình bằng phương pháp bình phương nhỏ nhất.
- ✓ Xác định a, b theo công thức:

$$a_i = \frac{\sum_{i=1}^m (\sum Y_i) - \sum_{i=1}^m \left\{ \frac{\sum X_i \sum X_i Y_i}{\sum X_i^2} \right\}}{\sum_{i=1}^m n_i - \sum_{i=1}^m \left\{ \frac{(\sum X_i)^2}{\sum X_i^2} \right\}}$$

$$b_i = \frac{\sum_{i=1}^m (\sum X_i Y_i) - \left\{ \frac{\sum X_i \sum Y_i}{n_i} \right\}}{\sum_{i=1}^m \left\{ \sum_{i=1}^m X_i^2 - \frac{(\sum X_i)^2}{\sum n_i} \right\}}$$

Trong đó :

- + m = số nhóm (số cấp đất)
- + n_i = số cặp H_0/A của nhóm i
- + $\sum X_i$ = tổng các giá trị của x thuộc nhóm i
- + $\sum X_i^2$ = tổng các giá trị của x_i^2 thuộc nhóm i
- + $\sum Y_i^2$ = tổng các giá trị của y_i^2 thuộc nhóm i
- + $\sum X_i Y_i$ = tổng tích x và y thuộc nhóm i

(2) Xây dựng đường cong cấp đất bình quân cho mỗi cấp đất ($H_{0i} - A$)

Trình tự thực hiện như sau

- ✓ Lập bảng quan hệ $H_0 - A$ theo từng cấp đất.
- ✓ Chọn dạng phương trình thích hợp theo $\min \sum (Y_i - Y_{\text{ilt}})^2$.
- ✓ Tính các tham số của mô hình bằng phương pháp bình phương nhỏ nhất.

2.3.4. PHƯƠNG PHÁP PHÂN CHIA CẤP ĐẤT

2.3.4.1. Phương pháp cố định điểm chặn, thay đổi độ dốc theo mỗi cấp đất

Từ mô hình $Y = a + bX$, cố định a , thay đổi b theo số cấp đất như sau:

Đối với cấp đất i , hàm hồi quy ước lượng chiều cao H_0 có dạng:

$$y_i = a + b_i x \quad (10)$$

$$b_i = (y_i - a)/x \quad (11)$$

Sau đó thay y_i bằng chỉ số cấp đất (S_i) và X bằng tuổi A_0 vào (11), có thể xác định được b_i cho mỗi cấp đất. Ví dụ:

+ **Đối với hàm Schumacher**

$$b_i = [\ln(S_i) - a] * A_0^{-c} \quad (12)$$

+ **Đối với hàm Gompertz**

$$b_i = (\ln S_i - a)/A_0 \quad (13)$$

- ✓ **Ưu điểm.** Vì b_i thay đổi, nên độ dốc của các đường cong sinh trưởng chiều cao cũng tăng dần từ cấp đất thấp đến cấp đất cao. Ngoài ra, đường cong cấp đất đi qua gốc tọa độ.
- ✓ **Nhược điểm.** Vì a cố định, nên tham số m (hay $Y_{\max}$) không thay đổi. Ngoài ra, giá trị $Y = Y_{\max}$ ở A rất cao.
- ✓ **Điều kiện ứng dụng.** Xây dựng đường cong sinh trưởng H_0 chung cho mỗi cấp đất. Chú ý: (1) A_0 càng cao càng tốt để cho khoảng cách giữa các đường sinh trưởng H_0 đạt cao nhất sau tuổi khai thác chính. (2) Dùng phương pháp hồi quy phân nhóm.

2.3.4.2. Phương pháp cố định độ dốc, thay đổi điểm chặn theo mỗi cấp đất

Từ mô hình $Y = a + bX$, cố định b , thay đổi a theo số cấp đất như sau:

+ Đối với cấp đất i , hàm hồi quy ước lượng có dạng:

$$y_i = a_i + bx \quad (14)$$

$$a_i = y_i - bx \quad (15)$$

Sau đó thay y_i bằng chỉ số cấp đất (S_i) và thay X tại tuổi A_0 vào (15), có thể xác định được a_i cho mỗi cấp đất. Ví dụ:

+ **Đối với hàm Schumacher**

$$a_i = \ln(S_i) + b \cdot A_0^{-c} \quad (16)$$

+ **Đối với hàm Gompertz**

$$a_i = \ln(S_i) - b \cdot A_0 \quad (17)$$

2.3.4.3. Phương pháp Affill

Trình tự thực hiện như sau

+ Trước hết, chuyển $y_i = a + bx$ về dạng phi tuyến (gốc)

$$H = f(A) \quad (18)$$

+ Kế đến, xác định H_0 lý thuyết tại A_0

$$H_{(A_0)} = F(A_0) \quad (19)$$

+ Tiếp theo, xác định phương trình trung gian

$$F_0 = f(A)/H_{(A_0)} = (18)/(19) \quad (20)$$

+ Tiếp đến, xác định mô hình hồi quy ước lượng cho từng cấp đất bằng cách nhân F_0 lần lượt với các chỉ số cấp đất (S_i) tại A_0 , nghĩa là:

$$F_i(A) = F_0 \cdot S_i \quad (21)$$

+ Lập biểu và vẽ biểu đồ cấp đất.

Nhược điểm của phương pháp 2 và 3. Vì b cố định, nên độ dốc của các đường sinh trưởng ở các cấp đất không đổi. Do đó, tốc độ sinh trưởng cao nhất ($Z_{h_{max}}$) đạt ở cùng một tuổi. Điều đó không phản ánh đúng quy luật sinh trưởng H_0 của từng cấp đất.

2.3.4.4. Phương pháp xây dựng các đường cong sinh trưởng chiều cao độc lập theo từng cấp đất

(a) Sử dụng chiều cao thực nghiệm bình quân

+ Mỗi cấp đất lập một đường cong sinh trưởng H_0 riêng biệt.

+ Nhược điểm: (1) Khoảng cách giữa các đường sinh trưởng ở mỗi tuổi có thể không đều nhau; (2) Đôi khi hai đường cong cấp đất có thể cắt nhau.

(b) Sử dụng hai cặp giá trị H/A .

Ở mỗi cấp đất, chọn hai cặp giá trị H/A như sau:

+ Cặp H/A thứ nhất lấy tại H_0/A_0 .
 + Cặp H/A thứ hai lấy tại H_0^*/A_0^* . Tuổi A^* được chọn sao cho dễ xác định chiều cao ở mỗi cấp đất.

+ Sau đó xác định các tham số a, b cho mỗi cấp đất như sau:

$$b = (H_2 - H_1)/(A_2 - A_1) \quad (22)$$

$$a = H_2 - bA_2 = H_1 - bA_1 \quad (23)$$

+ **Đổi với hàm Schumacher**

$$b_i = \frac{\ln H_2 - \ln H_1}{1/A_2^c - 1/A_1^c} \quad (24)$$

$$a_i = \ln H_2 - b_i * 1/A_2^c = \ln H_1 - b_i * 1/A_1^c \quad (25)$$

$$H_i = a_i + b_i * 1/A^c$$

Khi biết a_i và b_i , biến đổi về hàm Schumacher như sau:

$$H_0 = m * \exp(-b_i * 1/A^c)$$

hoặc:

$$H_0 = \exp(a_i) * \exp(-b_i * 1/A^c)$$

2.3.4.5. Phương pháp xây dựng các đường cong cấp đất thông qua suất tăng trưởng

Thủ tục như sau:

+ Sơ bộ phân chia các lâm phần thành i cấp đất ($i = 1 \dots 5$)
 + Xây dựng phương trình sinh trưởng H_{0i} cho từng cấp đất: $F(t)$
 + Tính tốc độ sinh trưởng (lượng tăng thường xuyên) cho từng cấp đất
 $Z_h = F(t)'$

+ Tính suất tăng trưởng H_0 cho mỗi cấp đất: $P_H = (Z_h/H) * 100$

+ Từ S_i cho trước ở tuổi A_0 , tính H_0 ở các tuổi A như sau:

$$\text{Khi } A < A_0: H_{(A-1)} = H_A \left(1 - \frac{P_{HA}}{100}\right) \quad (26)$$

$$\text{Khi } A > A_0: H_{A+1} = \frac{H_A}{1 - \frac{P_{HA+1}}{100}} \quad (27)$$

2.3.5. XÁC ĐỊNH CHỈ SỐ CẤP ĐẤT CHO LÂM PHẦN ĐIỀU TRA

Giả sử đường cong cấp đất lập theo hàm Schumacher.

+ Khi a cố định

Từ phương trình (11) và (12) có thể xác định được chỉ số cấp đất cho lâm phần bất kỳ theo công thức:

$$\ln(S_A) = a + [\ln(H_{0A}) - a](A/A_0)^c \quad (28)$$

Trong công thức (28), H_0 là chiều cao của lâm phần tương ứng với tuổi A. Ví dụ: $A_0 = 20$, $A = 16$, $H_{0(16)} = 15$ m. $\ln(S) = a + [\ln(15) - a]*(16/20)^c$.

+ Khi b cố định

Từ phương trình (11) và (16) có thể xác định được chỉ số cấp đất cho lâm phần bất kỳ theo công thức:

$$\ln S_A = \ln(H_0) + b(1/A^c - 1/A_0^c) \quad (29)$$

Ở đây, H_0 = chiều cao lâm phần tại tuổi A; A_0 = tuổi cơ sở.

TÓM TẮT PHƯƠNG PHÁP CHUNG LẬP BIỂU CẤP ĐẤT

1. Xác định đối tượng lập biểu cấp đất: loài cây; vùng hay địa phương.
2. Thu thập mẫu và phân chia sơ bộ số cấp đất. Ở bước này trước hết chọn tuổi cơ sở (A_0 , năm) và xác định biến động H_0 (m) tại tuổi A_0 . Kế đến, từ biến động H_0 ở tuổi cơ sở (A_0 , năm), xác định số cấp đất. Một cách khác là tính quan hệ $H_0 - A$; kế đến tìm khoảng tin cậy và khoảng dự đoán cho từng cấp đất; sau đó xác định các chỉ số (S_i , m) cho từng cấp đất.
3. Chọn mô hình mô tả quan hệ $H_0 - A$.
4. Xây dựng những mô hình $H_0 - A$ cho từng cấp đất.
5. Xây dựng mô hình $H_0 - A$ bình quân chung cho các cấp đất.
6. Xây dựng mô hình $H_0 - A$ ở giữa các cấp đất và $H_0 - A$ ở ranh giới các cấp đất.
7. Kiểm nghiệm tính phù hợp của các hàm hồi quy biểu thị cho mỗi cấp đất. Để giải quyết vấn đề này, cần chọn một số lâm phần không tham gia lập biểu để kiểm tra. Thủ tục kiểm tra như sau: (2) Lập $H_0 - A$ cho các lâm phần kiểm tra; (2) Tuyến tính hoá các hàm hồi quy $H_0 - A$; (3) Kiểm tra sự sai khác về độ dốc (tham số b) của các mô hình kiểm tra so với mô hình $H_0 - A$ trung bình của các cấp đất.
8. Cuối cùng, lập biểu cấp đất và vẽ đồ thị mô tả các đường cong cấp đất.

2.4. THỰC HÀNH XÂY DỰNG BIỂU CẤP ĐẤT

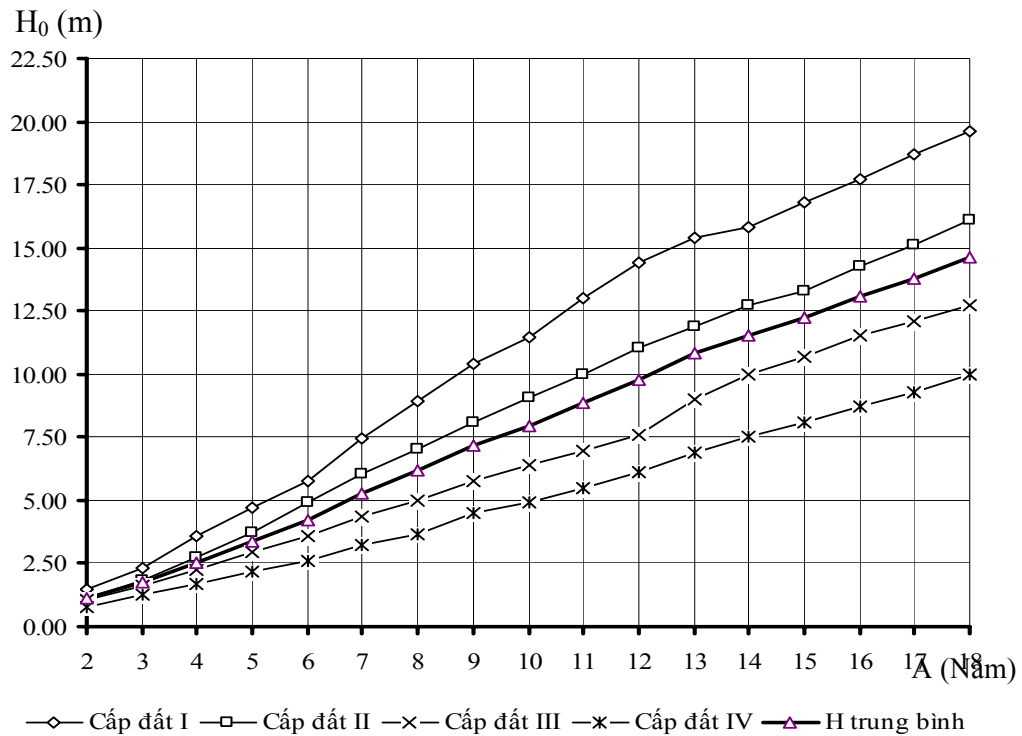
Bước 1. Tập hợp và chỉnh lý số liệu

Bảng 1 ghi lại chiều cao của tầng trệt (H_0 , m) của 65 lâm phần Keo lá tràm từ 2-18 tuổi ở khu vực miền Đông Nam Bộ. Thông qua điều tra, đã phân chia sơ bộ H_0 ở mỗi tuổi vào 4 cấp đất khác nhau.

Bảng 1. Chiều cao H_0 (m) của 65 lâm phần Keo lá tràm từ 2-18 tuổi được phân chia sơ bộ theo 4 cấp đất

TT	A(Năm)	H (m)	Cấp đất	TT	A(Năm)	H (m)	Cấp đất	TT	A(Năm)	H (m)	Cấp đất
1	2	1.48	1	24	7	4.35	3	46	13	15.40	1
2	2	1.10	2	25	7	3.21	4	47	13	11.90	2
3	2	1.04	3	26	8	8.96	1	48	13	9.02	3
4	2	0.77	4	27	8	7.00	2	49	13	6.90	4
5	3	2.30	1	28	8	5.00	3	50	14	15.84	1
6	3	1.84	2	29	8	3.67	4	51	14	12.73	2
7	3	1.61	3	30	9	10.43	1	52	14	10.00	3
8	3	1.27	4	31	9	8.11	2	53	14	7.50	4
9	4	3.56	1	32	9	5.77	3	54	15	16.80	1
10	4	2.74	2	33	9	4.50	4	55	15	13.30	2
11	4	2.25	3	34	10	11.43	1	56	15	10.70	3
12	4	1.67	4	35	10	9.09	2	57	15	8.10	4
13	5	4.68	1	36	10	6.39	3	58	16	17.70	1
14	5	3.75	2	37	10	4.90	4	59	16	14.30	2
15	5	2.94	3	38	11	13.00	1	60	16	11.50	3
16	5	2.18	4	39	11	9.96	2	61	17	18.70	1
17	6	5.76	1	40	11	6.98	3	62	17	15.10	2
18	6	4.91	2	41	11	5.46	4	63	17	12.10	3
19	6	3.61	3	42	12	14.43	1	64	18	19.64	1
20	6	2.63	4	43	12	11.07	2	65	18	12.75	3
21	7	7.43	1	44	12	7.60	3				
22	7	6.02	2	45	12	6.12	4				

Nguồn: Vũ Tiên Minh (2005)



Hình 1. Đồ thị mô tả $H_0 - A$ thực nghiệm của rừng Keo lá tràm

Bước 2. Chọn mô hình mô tả quan hệ H_0-A

Để mô tả quan hệ giữa $H_0 - A$, ở đây chọn mô hình Schumacher theo dạng:

$$H_0 = m \cdot \exp(-b/A^c) \text{ hay } \ln(H) = \ln(m) - b/A^c \quad (1)$$

Bước 3. Xây dựng mô hình $H_{0bq}-A$ cho mỗi cấp đất

Từ số liệu ở **bảng 1**, khi cho $c = 0,2$, bằng phương pháp biến giả, có thể xác định được các tham số của các mô hình biểu diễn quan hệ $H_{0bq} - A$ của 4 cấp đất và mô hình H_0-A bình quân chung cho các cấp đất (**bảng 2**). Từ số liệu của **bảng 2**, có thể xây dựng bốn mô hình mô tả quan hệ giữa $H_0 - A$ của rừng Keo lá tràm tương ứng với 4 cấp đất như sau:

$$\ln H_{0(I)} = 7.92635 - 8.74344/A^{0.2} \quad (2)$$

$$\ln H_{0(II)} = 7.77828 - 8.87132/A^{0.2} \quad (3)$$

$$\ln H_{0(III)} = 7.16274 - 8.33327/A^{0.2} \quad (4)$$

$$\ln H_{0(IV)} = 6.80355 - 8.23060/A^{0.2} \quad (5)$$

Từ các mô hình (2–5), có thể biến đổi về dạng mô hình Schumacher chính tắc như sau:

$$H_{0(I)} = 2768.6 * \exp(-8.74344/A^{0.2}) \quad (6)$$

$$H_{0(II)} = 2387.6 * \exp(-8.87132/A^{0.2}) \quad (7)$$

$$H_{0(III)} = 1290.2 * \exp(-8.33327/A^{0.2}) \quad (8)$$

$$H_{0(IV)} = 900.9 * \exp(-8.23060/A^{0.2}) \quad (9)$$

Bước 4. Xây dựng mô hình H_{0bq} -A cho bốn cấp đất

Mô hình H_0 -A bình quân cho bốn cấp đất có dạng:

$$\ln(H_{0bq}) = 7.58563 - 8.77137/A^{0.2} \quad (10)$$

$$\text{Hay } H_{0(bq)} = 1969.2 * \exp(-8.77137/A^{0.2}) \quad (11)$$

Bảng 2. Phân tích hồi quy tương quan giữa H_0 với A theo 4 cấp đất khác nhau

Multiple Regression Analysis

Parameter	Estimate	Standard Error	T Statistic	P-Value
CONSTANT	7.92635	0.0987489	80.2677	0.0000
A	-8.74344	0.14904	-58.6652	0.0000
Cap dat=2	-0.148062	0.143212	-1.03387	0.3056
Cap dat=3	-0.763613	0.139652	-5.46797	0.0000
Cap dat=4	-1.1228	0.152197	-7.37726	0.0000
A*Cap dat=2	-0.127879	0.215159	-0.594345	0.5547
A*Cap dat=3	0.410171	0.210774	1.94602	0.0567
A*Cap dat=4	0.512838	0.226091	2.26828	0.0272

Analysis of Variance

Source	Sum of Squares	Df	Mean Square	F-Ratio	P-Value
Model	41.3255	7	5.90365	2107.82	0.0000
Residual	0.156847	56	0.00280083		
Total (Corr.)	41.4824	63			

R-Squared = 99.6219 percent

R-Squared (adjusted for d.f.) = 99.5746 percent

Standard Error of Est. = 0.0529229

So sánh các mô hình (2 – 5) cho thấy, điểm chặn và độ dốc của bốn mô hình đều khác nhau rất lớn ở mức ý nghĩa $P < 0.01$ (Bảng 3).

Bảng 3. So sánh điểm chặn và độ dốc của các mô hình

Further ANOVA for Variables in the Order Fitted

Source	Sum of Squares	Df	Mean Square	F-Ratio	P-Value
A	36.0859	1	36.0859	12883.99	0.0000
Intercepts	5.2071	3	1.7357	619.71	0.0000
Slopes	0.0325087	3	0.0108362	3.87	0.0138
Model	41.3255	7			

Từ mô hình (6-11), có thể xác định được H_{0bq} tương ứng với các cấp đất (Bảng 4).

Bảng 4. Chiều cao bình quân tầng trội của rừng Keo lá tràm tương ứng với bốn cấp đất

A (Năm)	Chiều cao H_0 (m) tương ứng với cấp đất:				
	I	II	III	IV	Bình quân
2	1.4	1.1	0.9	0.7	1.0
3	2.5	1.9	1.6	1.2	1.7
4	3.7	2.9	2.3	1.8	2.6
5	4.9	3.9	3.1	2.3	3.4
6	6.1	4.8	3.8	2.9	4.3
7	7.4	5.9	4.6	3.4	5.2
8	8.7	6.9	5.3	3.9	6.0
9	9.9	7.9	6.0	4.5	6.9
10	11.1	8.9	6.7	5.0	7.8
11	12.3	9.8	7.4	5.5	8.6
12	13.6	10.8	8.1	6.0	9.5
13	14.8	11.8	8.8	6.5	10.3
14	15.9	12.7	9.5	7.0	11.1
15	17.1	13.7	10.1	7.5	12.0
16	18.3	14.6	10.8	8.0	12.8
17	19.4	15.6	11.4	8.4	13.6
18	20.5	16.5	12.0	8.9	14.4

Bước 5. Chọn tuổi cơ sở và xác định chỉ số cấp đất tại tuổi cơ sở

Tuổi cơ sở (A_0 , năm) được chọn là 15 năm. Từ số liệu của **bảng 4** cho thấy, tại tuổi cơ sở (15 năm) các giá trị H_{0bq} của 4 cấp đất từ IV đến I tương ứng là 7,5m, 10,1m, 13,7m và 17,1m. Do đó, chỉ số cấp đất (S_i , m) của 4 cấp đất tại tuổi cơ sở có thể được chọn tương ứng là 8m, 11m, 14m và 17m.

Bước 6. Xây dựng những mô hình chiều cao ở giữa các cấp đất và ranh giới giữa các cấp đất

6.1. Phương pháp cố định điểm chặn và thay đổi độ dốc

Khi cố định điểm chặn (tham số a), thì độ dốc (tham số b) của 4 đường cong ở giữa các cấp đất được xác định theo công thức:

$$b_i = [\ln(S_i) - a] * A_0^c \quad (12)$$

Ở công thức (12), giá trị a là điểm chặn chung của 4 cấp đất ($a = 7,58563$); S_i là chỉ số H_0 của 4 cấp đất tại tuổi A_0 ($S_{IV} = 8m$, $S_{III} = 11m$, $S_{II} = 14m$ và $S_I = 17m$). Từ đó có thể xác định độ dốc của 4 đường cong cấp đất như sau:

$$b_I = (\ln(17) - 7.58563) \cdot 15^{0.2} = -8,16832 \quad (13)$$

$$b_{II} = (\ln(14) - 7.58563) \cdot 15^{0.2} = -8,50203 \quad (14)$$

$$b_{III} = (\ln(11) - 7.58563) \cdot 15^{0.2} = -8,91653 \quad (15)$$

$$b_{IV} = (\ln(8) - 7.58563) \cdot 15^{0.2} = -9,46388 \quad (16)$$

Từ **mô hình (11) và (13-16)**, có thể xây dựng 4 đường cong ở giữa các cấp đất như sau:

$$H_{0(I)} = 1969.2 \cdot \exp(-8.16832/A^{0.2}) \quad (17)$$

$$H_{0(II)} = 1969.2 \cdot \exp(-8.50203/A^{0.2}) \quad (18)$$

$$H_{0(III)} = 1969.2 \cdot \exp(-8.91653/A^{0.2}) \quad (19)$$

$$H_{0(IV)} = 1969.2 \cdot \exp(-9.46388/A^{0.2}) \quad (20)$$

Với 4 cấp đất, cần phải xây dựng 5 mô hình biểu thị ranh giới H_0 giữa các cấp đất. Từ **mô hình 10**, $a = 7,58563$; còn S_i ở ranh giới giữa các cấp đất tại tuổi A_0 là $S_{IVdưới} = 6,5m$, $S_{III-IV} = 9,5m$, $S_{II-III} = 12,5m$, $S_{I-II} = 15,5m$ và $S_{Itrên} = 18,5m$. Sau khi tính b_i cho từng cấp đất, có thể xây dựng được 5 mô hình chiều cao ở ranh giới giữa các cấp đất như sau:

$$H_{0(I)} = 1969.2 \cdot \exp(-8.02299/A^{0.2}) \quad (21)$$

$$H_{0(II-I)} = 1969.2 \cdot \exp(-8.32709/A^{0.2}) \quad (22)$$

$$H_{0(III-II)} = 1969.2 \cdot \exp(-8.69682/A^{0.2}) \quad (23)$$

$$H_{0(IV-III)} = 1969.2 \cdot \exp(-9.16851/A^{0.2}) \quad (24)$$

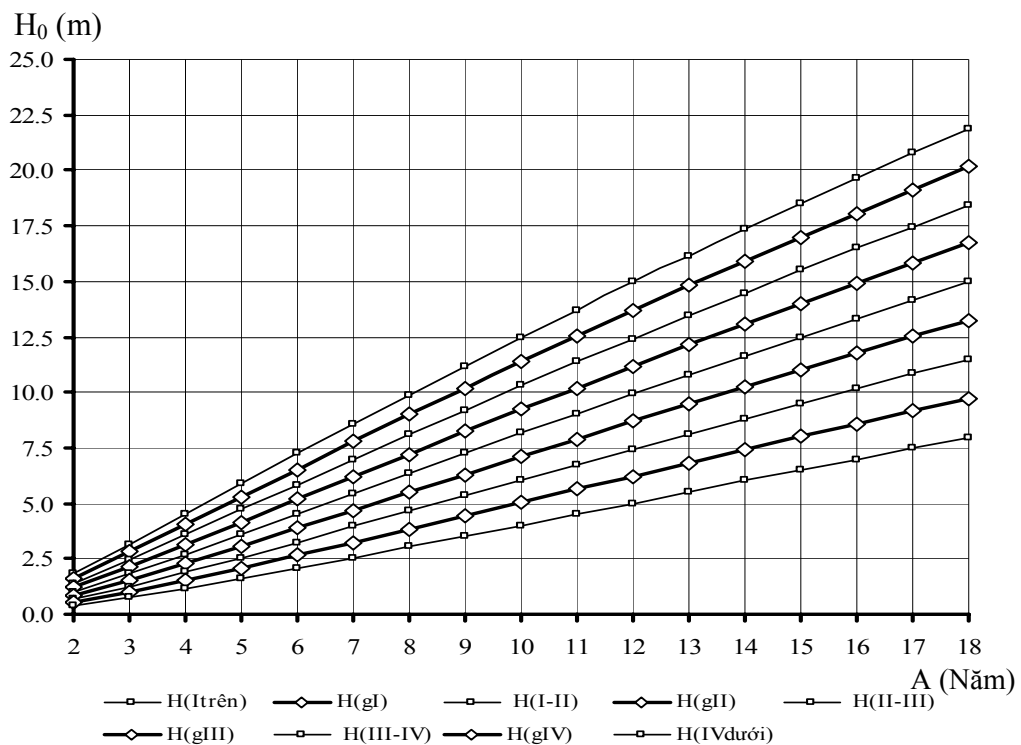
$$H_{0(IV)} = 1969.2 \cdot \exp(-9.82077/A^{0.2}) \quad (25)$$

Từ các mô hình (17-25), có thể xây dựng biểu cấp đất cho rừng Keo lá tràm ở miền Đông Nam Bộ (**Bảng 5; Hình 2**).

Bảng 5. Biểu cấp đất rừng Keo lá tràm ở khu vực miền Đông Nam Bộ

(Xây dựng theo phương pháp cố định điểm chặn và thay đổi độ dốc)

A (Năm)	H _(l trên)	H _(gl)	H(I-II)	H _(glII)	H(II-III)	H _(glIII)	H(III-IV)	H _(glIV)	H _(IV dưới)
2	1.8	1.6	1.4	1.2	1.0	0.8	0.7	0.5	0.4
3	3.1	2.8	2.5	2.1	1.8	1.5	1.3	1.0	0.7
4	4.5	4.0	3.6	3.1	2.7	2.3	1.9	1.5	1.2
5	5.9	5.3	4.7	4.2	3.6	3.1	2.6	2.1	1.6
6	7.2	6.5	5.8	5.2	4.5	3.9	3.2	2.6	2.1
7	8.6	7.8	7.0	6.2	5.4	4.7	3.9	3.2	2.5
8	9.9	9.0	8.1	7.2	6.3	5.5	4.6	3.8	3.0
9	11.2	10.2	9.2	8.2	7.3	6.3	5.4	4.4	3.5
10	12.5	11.4	10.3	9.2	8.2	7.1	6.1	5.0	4.0
11	13.7	12.5	11.4	10.2	9.0	7.9	6.8	5.6	4.5
12	14.9	13.7	12.4	11.2	9.9	8.7	7.4	6.2	5.0
13	16.2	14.8	13.5	12.1	10.8	9.5	8.1	6.8	5.5
14	17.3	15.9	14.5	13.1	11.6	10.2	8.8	7.4	6.0
15	18.5	17.0	15.5	14.0	12.5	11.0	9.5	8.0	6.5
16	19.6	18.1	16.5	14.9	13.3	11.8	10.2	8.6	7.0
17	20.8	19.1	17.5	15.8	14.2	12.5	10.8	9.2	7.5
18	21.9	20.1	18.4	16.7	15.0	13.2	11.5	9.7	8.0



Hình 2. Đồ thị mô tả 4 cấp đất của rừng Keo lá tràm ở miền Đông Nam Bộ

(Xây dựng theo phương pháp cố định điểm chặn và thay đổi độ dốc)

6.2. Phương pháp cố định độ dốc và thay đổi điểm chặn

Khi cố định độ dốc (tham số b), thì điểm chặn (tham số a) của 4 đường cong ở giữa các cấp đất được xác định theo công thức:

$$a_i = \ln(S_i) - b/A_0^c \quad (26)$$

Ở công thức (26), giá trị b là độ dốc chung của 4 cấp đất ($b = -8.77137$); S_i là chỉ số H_0 của 4 cấp đất tại tuổi A_0 ($S_{IV} = 8m$, $S_{III} = 11m$, $S_{II} = 14m$ và $S_I = 17m$). Từ đó có thể xác định điểm chặn của 4 đường cong cấp đất như sau:

$$a_I = \ln(17) + 8.77137/15^{0.2} = 7.9365 \quad (27)$$

$$a_{II} = (\ln(14) + 8.77137/15^{0.2} = 7.7423 \quad (28)$$

$$a_{III} = (\ln(11) + 8.77137/15^{0.2} = 7.5012 \quad (29)$$

$$a_{IV} = (\ln(8) + 8.77137/15^{0.2} = 7.1827 \quad (30)$$

Từ **mô hình (10) và (27-30)**, có thể xây dựng 4 đường cong ở giữa các cấp đất như sau:

$$H_{0(I)} = 2796.86 * 2.7182^{(-8.77137/A^{0.2})} \quad (31)$$

$$H_{0(II)} = 2303.31 * 2.7182^{(-8.77137/A^{0.2})} \quad (32)$$

$$H_{0(III)} = 1809.76 * 2.7182^{(-8.77137/A^{0.2})} \quad (33)$$

$$H_{0(IV)} = 1316.20 * 2.7182^{(-8.77137/A^{0.2})} \quad (34)$$

Năm mô hình biểu thị ranh giới H_0 giữa hai cấp đất kế cận có dạng:

$$H_{0(IV \text{ dưới})} = 1069.42 * \exp(-8.77137/A^{0.2}) \quad (35)$$

$$H_{0(IV-III)} = 1562.98 * \exp(-8.77137/A^{0.2}) \quad (36)$$

$$H_{0(III-II)} = 2056.53 * \exp(-8.77137/A^{0.2}) \quad (37)$$

$$H_{0(II-I)} = 2550.08 * \exp(-8.77137/A^{0.2}) \quad (38)$$

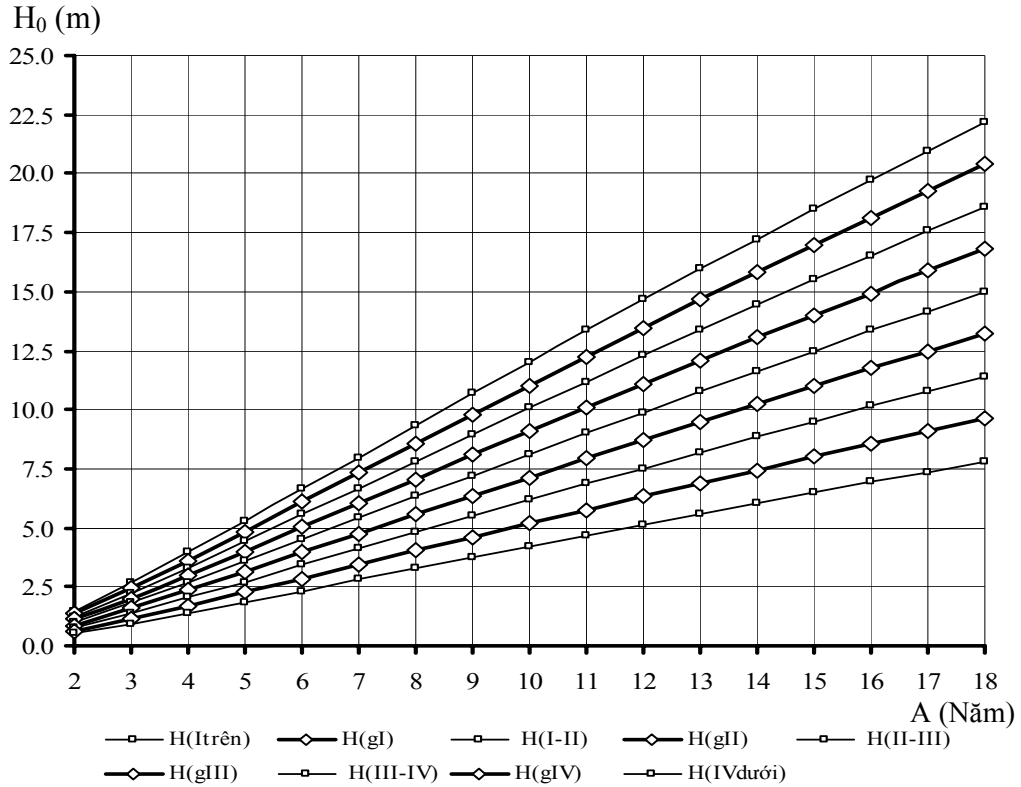
$$H_{0(I \text{ trên})} = 3043.63 * \exp(-8.77137/A^{0.2}) \quad (39)$$

Khi thay thế A vào mô hình 31-39, có thể xây dựng biểu cấp đất cho rừng Keo lá tràm ở miền Đông Nam Bộ (**Bảng 6; Hình 3**).

Bảng 6. Biểu cấp đất rừng Keo lá tràm ở khu vực miền Đông Nam Bộ

(Xây dựng theo phương pháp cố định độ dốc và thay đổi điểm chặn)

A (Năm)	H _(trên)	H _(gl)	H(I-II)	H _(glI)	H(II-III)	H _(glII)	H(III-IV)	H _(glIV)	H _(IVdưới)
2	1.5	1.4	1.2	1.1	1.0	0.9	0.8	0.6	0.5
3	2.7	2.4	2.2	2.0	1.8	1.6	1.4	1.2	0.9
4	3.9	3.6	3.3	3.0	2.7	2.3	2.0	1.7	1.4
5	5.3	4.9	4.4	4.0	3.6	3.1	2.7	2.3	1.9
6	6.6	6.1	5.6	5.0	4.5	3.9	3.4	2.9	2.3
7	8.0	7.3	6.7	6.0	5.4	4.7	4.1	3.5	2.8
8	9.3	8.6	7.8	7.1	6.3	5.6	4.8	4.0	3.3
9	10.7	9.8	9.0	8.1	7.2	6.4	5.5	4.6	3.8
10	12.0	11.0	10.1	9.1	8.1	7.1	6.2	5.2	4.2
11	13.3	12.3	11.2	10.1	9.0	7.9	6.9	5.8	4.7
12	14.7	13.5	12.3	11.1	9.9	8.7	7.5	6.3	5.1
13	15.9	14.7	13.4	12.1	10.8	9.5	8.2	6.9	5.6
14	17.2	15.8	14.4	13.0	11.6	10.2	8.8	7.5	6.1
15	18.5	17.0	15.5	14.0	12.5	11.0	9.5	8.0	6.5
16	19.7	18.1	16.5	14.9	13.3	11.7	10.1	8.5	6.9
17	21.0	19.3	17.6	15.9	14.2	12.5	10.8	9.1	7.4
18	22.2	20.4	18.6	16.8	15.0	13.2	11.4	9.6	7.8



Hình 3. Đồ thị mô tả cấp đất rừng Keo lá tràm ở miền Đông Nam Bộ

(Xây dựng theo phương pháp cố định độ dốc và thay đổi điểm chặn)

6.3. Phân chia cấp đất theo phương pháp Affill

- Trước hết, từ mô hình 11, xác định giá trị H_{lt} tại tuổi cơ sở ($A_0 = 15$ tuổi). Kết quả nhận được:

$$i. \quad H_{0(15)} = 1969.2 * \exp(-8.77137/15^{0.2}) = 12m.$$

- Kế đến, chia phương trình (11) cho chiều cao ở tuổi cơ sở (12m), ta có:

$$ii. \quad F_0 = [1969.2 * \exp(-8.77137/15^{0.2})]/12$$

$$iii. \quad F_0 = 164.10 * \exp(-8.77137/A^{0.2}) \quad (40)$$

- Tiếp theo, xác định 4 phương trình sinh trưởng chiều cao bình quân của mỗi cấp đất bằng cách nhân mô hình (40) với các chỉ số cấp đất tại tuổi cơ sở ($S_i = 17m$ – cấp đất I, $14m$ – cấp đất II, $11m$ – cấp đất III, $8m$ – cấp đất IV). Kết quả nhận được 4 phương trình sinh trưởng chiều cao bình quân của mỗi cấp đất như sau:

$$H_{0(I)} = 2789.7 * \exp(-8.77137/A^{0.2}) \quad (41)$$

$$H_{0(II)} = 2297.4 * \exp(-8.77137/A^{0.2}) \quad (42)$$

$$H_{0(III)} = 1805.1 * \exp(-8.77137/A^{0.2}) \quad (43)$$

$$H_{0(IV)} = 1312.8 * \exp(-8.77137/A^{0.2}) \quad (44)$$

- Tiếp theo, xác định mô hình H_0 -A ở ranh giới giữa hai cấp đất kế cận bằng cách nhân phương trình (40) với $H_{0IVdưới} = 6,5m$, $H_{0III-IV} = 9,5m$, $H_{0II-III} = 12,5m$, $H_{0I-II} = 15,5m$ và $H_{0trên} = 18,5m$. Kết quả nhận được 5 mô hình H_0 -A ở ranh giới giữa các cấp đất như sau:

$$H_{0(I)} = 3035.9 * \exp(-8.77137/A^{0.2}) \quad (45)$$

$$H_{0(II-I)} = 2543.6 * \exp(-8.77137/A^{0.2}) \quad (46)$$

$$H_{0(III-II)} = 2051.3 * \exp(-8.77137/A^{0.2}) \quad (47)$$

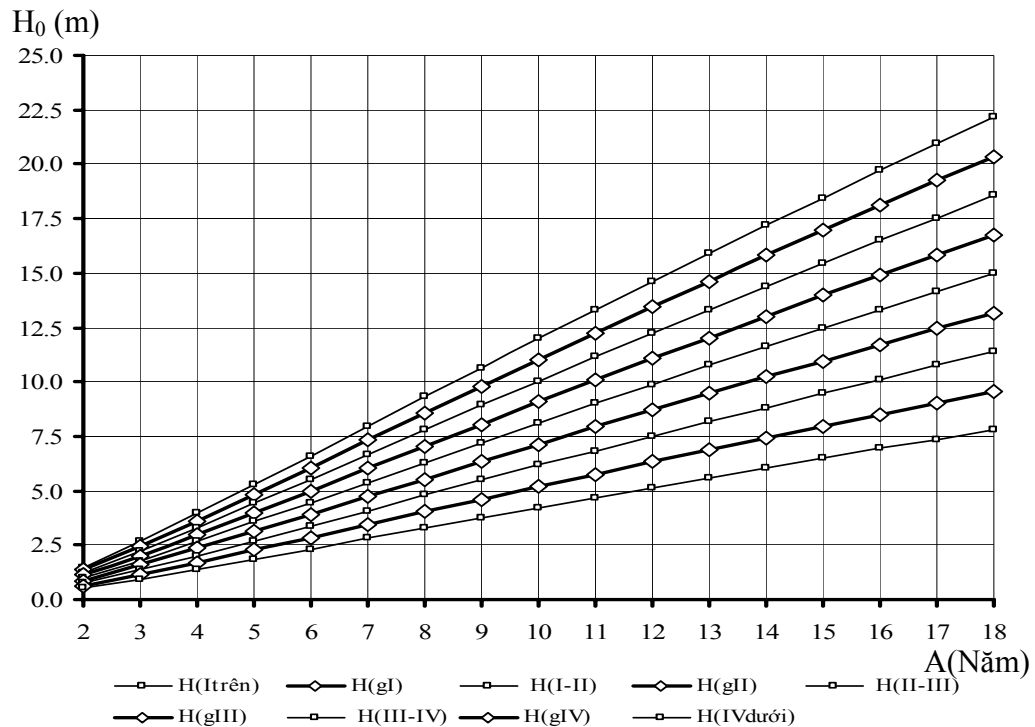
$$H_{0(IV-III)} = 1559.0 * \exp(-8.77137/A^{0.2}) \quad (48)$$

$$H_{0(IV)} = 1066.7 * \exp(-8.77137/A^{0.2}) \quad (49)$$

- Cuối cùng, từ các phương trình biểu thị chiều cao giữa các cấp đất (41-44) và chiều cao thuộc ranh giới giữa hai cấp đất kế cận (45 – 49), tính H_{lt} cho từng cấp đất. Sau đó lập biểu và vẽ đồ thị biểu diễn các cấp đất (**Bảng 7; Hình 4**).

Bảng 7. Biểu cấp đất rừng Keo lá tràm ở khu vực miền Đông Nam Bộ
(Xây dựng theo phương pháp Affill)

A (Năm)	$H_{(Itrên)}$	$H_{(gl)}$	$H_{(I-II)}$	$H_{(gII)}$	$H_{(II-III)}$	$H_{(gIII)}$	$H_{(III-IV)}$	$H_{(gIV)}$	$H_{(IVdưới)}$
2	1.5	1.3	1.2	1.1	1.0	0.9	0.8	0.6	0.5
3	2.7	2.4	2.2	2.0	1.8	1.6	1.4	1.1	0.9
4	3.9	3.6	3.3	3.0	2.7	2.3	2.0	1.7	1.4
5	5.3	4.8	4.4	4.0	3.6	3.1	2.7	2.3	1.8
6	6.6	6.1	5.5	5.0	4.5	3.9	3.4	2.9	2.3
7	8.0	7.3	6.7	6.0	5.4	4.7	4.1	3.4	2.8
8	9.3	8.6	7.8	7.0	6.3	5.5	4.8	4.0	3.3
9	10.7	9.8	8.9	8.1	7.2	6.3	5.5	4.6	3.7
10	12.0	11.0	10.0	9.1	8.1	7.1	6.2	5.2	4.2
11	13.3	12.2	11.2	10.1	9.0	7.9	6.8	5.8	4.7
12	14.6	13.4	12.2	11.1	9.9	8.7	7.5	6.3	5.1
13	15.9	14.6	13.3	12.0	10.7	9.5	8.2	6.9	5.6
14	17.2	15.8	14.4	13.0	11.6	10.2	8.8	7.4	6.0
15	18.4	17.0	15.5	14.0	12.5	11.0	9.5	8.0	6.5
16	19.7	18.1	16.5	14.9	13.3	11.7	10.1	8.5	6.9
17	20.9	19.2	17.5	15.8	14.1	12.4	10.7	9.1	7.4
18	22.1	20.4	18.6	16.8	15.0	13.2	11.4	9.6	7.8



Hình 4. Đồ thị mô tả cấp đất rừng Keo lá tràm ở miền Đông Nam Bộ
(Xây dựng theo phương pháp Affill)

Bài IV. ỨNG DỤNG MÔ HÌNH LOGISTIC TRONG SINH THÁI QUAN XÃ

4.1. MÔ HÌNH LOGISTIC

4.1.1. Quan hệ giữa độ bắt gặp loài cây gỗ với các yếu tố môi trường

Nội dung này chỉ nghiên cứu cho một số loài cây quý hiếm như gỗ đỏ, giáng hương, cẩm lai, sao đen, dầu rái, dầu song năng, vên vên... Cách làm như sau: (1) Bố trí những tuyến cắt ngang qua trạng thái rừng III_{A3} – III_B; (2) trên mỗi tuyến cứ sau 50 – 100 m lại bố trí 1 điểm đo (không cần bố trí ô tiêu chuẩn); (3) tại mỗi điểm đo, xác định loài nghiên cứu có mặt hay không (quy định có mặt khi loài đó có D ≥ 10 cm) bằng hai dấu hiệu: có = 1, không có = 0; (4) dùng máy đo độ ẩm đất (%) và pH_{H2O}, xác định độ cao địa hình, loại đất... Ở đây so sánh độ bắt gặp một số loài theo hai loại đất: đất bazan – đất đỏ vàng trên phiến sét, hoặc đất bazan – đất xám trên granit. Số lượng điểm đo đếm trên mỗi loại đất ít nhất là 150. Từ số liệu trên đây, chúng ta có thể xây dựng một số mô hình logistic để tính toán xác suất bắt gặp loài cây như sau:

a. Mô hình 1. Mô tả ảnh hưởng của trạng thái rừng đến độ bắt gặp loài cây quan tâm nào đó.

Giả sử chỉ xem xét ảnh hưởng của 2 trạng thái rừng ($X_1 = 1$ – rừng III_{A3} và 0 – rừng III_{A2}) đến độ bắt gặp loài, chúng ta có thể xây dựng phương trình logit như sau:

$$\text{logit(rừng)} = b_0 + b_1 X_1 \text{ hay } \ln(p/1-p) = b_0 + b_1 X_1 \quad (1.1)$$

$$\text{hay } p = \frac{\exp(b_0 + b_1 X_1)}{1 + \exp(b_0 + b_1 X_1)}$$

Để thu được biểu thức đối với tỷ lệ sai khác từ mô hình logistic, chúng ta phải so sánh sự sai khác giữa hai trạng thái. Ở ví dụ trên có hai trạng thái rừng III_{A3} ($X_1 = 1$) và rừng III_{A2} ($X_1 = 0$). Từ đó chúng ta có thể viết

$$\ln[\text{odds(trạng thái rừng III}_{A3})] = b_0 + b_1 * 1 = b_0 + b_1$$

$$\ln[\text{odds(trạng thái rừng III}_{A2})] = b_0 + b_1 * 0 = b_0$$

$$OR_{(PC-NPC)} = \text{odds(III}_{A3})/\text{odds(III}_{A2}) = \frac{e^{(b_0 + b_1)}}{e^{b_0}} = e^{b_1}$$

Như vậy, nếu chỉ sử dụng một biến dự đoán nhận hai giá trị 0 và 1, thì việc so sánh tỷ lệ sai khác của hai trạng thái rừng được thực hiện bằng cách lấy số mũ cơ số neper của hệ số của biến dự đoán trong mô hình.

b. Mô hình 2. Mô tả ảnh hưởng của 2 trạng thái rừng ($X_1 = 1$ – rừng III_{A3} và 0 – rừng III_{A2}) và độ pH (X_2) đến độ bắt gặp loài

Chúng ta có thể xây dựng phương trình logit như sau:

$$\text{logit}(p/1-p) = b_0 + b_1X_1 + b_2X_2 \quad (1.2)$$

$$\text{hay } p = \frac{\exp(b_0 + b_1X_1 + b_2X_2)}{1 + \exp(b_0 + b_1X_1 + b_2X_2)}$$

Khi so sánh tỷ lệ sai khác về độ bắt gặp loài giữa hai trạng thái rừng dựa trên một tập hợp các biến dự đoán X_1 và X_2 , thì nhóm rừng III_{A3} (nhóm A) bao gồm các biến dự đoán $X_A = (X_{A1}, X_{A2})$, còn nhóm III_{A2} (nhóm B) bao gồm các biến dự đoán $X_B = (X_{B1}, X_{B2})$. Ở mô hình này, khi sử dụng hai biến dự đoán ($k = 2$), trong đó $X_1 = \text{trạng thái rừng}$ ($1 = \text{III}_{A3}$, $0 = \text{III}_{A2}$), $X_2 = \text{pH}$, thì tỷ lệ sai khác giữa hai trạng thái rừng được tính theo công thức:

$$\begin{aligned} \text{OR}_{(X_A-X_B)} &= \text{odds}(X_A)/\text{odds}(X_B) \\ &= \exp(b_0 + \sum_{j=1}^k b_j X_{Aj}) / \exp(b_0 + \sum_{j=1}^k b_j X_{Bj}) \\ \text{hay } \text{OR}_{(X_A-X_B)} &= \exp[\sum_{j=1}^k (X_{Aj} - X_{Bj}) b_j] \end{aligned} \quad (1.3)$$

Từ **biểu thức 1.3**, chúng ta có thể so sánh độ bắt gặp loài ở nhóm rừng III_{A3} (1) và rừng III_{A2} (0), độ pH = 5. Chúng ta viết $X_A = (1, 5)$, $X_B = (0, 5)$. Từ **biểu thức 1.3**, chúng ta tính tỷ lệ sai khác giữa hai nhóm rừng như sau:

$$\text{OR}_{(X_A-X_B)} = \exp(1 - 0)b_1 + (5 - 5)b_2 = \exp(b_1)$$

Nếu so sánh $X_A = (1, 6)$, $X_B = (0, 4)$, chúng ta tính tỷ lệ sai khác giữa hai nhóm rừng như sau:

$$\text{OR}_{(X_A-X_B)} = \exp(1-0)b_1 + (6 - 4)b_2 = \exp(-b_1) + 2b_2.$$

Những so sánh khác cũng được thực hiện tương tự như trên.

c. Mô hình 3. Mô tả ảnh hưởng của 2 trạng thái rừng ($X_1 = 1$ – rừng III_{A3} và 0 – rừng III_{A2}), độ pH (X_2) và độ ẩm đất (X_3) đến độ bắt gặp loài

Chúng ta có thể xây dựng phương trình logit như sau:

$$\text{logit}(p/1-p) = b_0 + b_1X_1 + b_2X_2 + b_3X_3 \quad (1.4)$$

$$\text{hay } p = \frac{\exp(b_0 + b_1X_1 + b_2X_2 + b_3X_3)}{1 + \exp(b_0 + b_1X_1 + b_2X_2 + b_3X_3)}$$

Đến đây thực hiện những so sánh tỷ lệ sai khác giữa hai nhóm rừng. Hai nhóm so sánh với nhau dựa trên một tập hợp các biến dự đoán X_1, X_2 và X_3 . Nhóm III_{A3} (nhóm A) bao gồm các biến dự đoán $X_A = (X_{A1}, X_{A2}$ và $X_{A3})$, còn nhóm B bao gồm các biến dự đoán $X_B = (X_{B1}, X_{B2}$ và $X_{B3})$. Ở mô hình này có ba biến dự đoán ($k = 3$), trong đó $X_1 =$ trạng thái rừng ($1 = III_{A3}, 0 = III_{A2}$), $X_2 =$ pH, $X_3 =$ độ ẩm đất, chúng ta có thể xây dựng những so sánh tỷ lệ sai khác giữa hai loại rừng (III_{A3} và III_{A2}) như sau:

$$\begin{aligned} OR_{(X_A-X_B)} &= odds(X_A)/odds(X_B) \\ &= \exp(b_0 + \sum_{j=1}^k b_j X_A) / \exp(b_0 + \sum_{j=1}^k b_j X_B) \\ \text{hay } OR_{(X_A-X_B)} &= \exp[\sum_{j=1}^k (X_A - X_B)b_j] \end{aligned} \quad (1.5)$$

Từ **biểu thức 1.5**, chúng ta có thể so sánh độ bắt gặp loài ở nhóm rừng III_{A3} (1) và rừng III_{A2} (0), độ pH = 5, độ ẩm = 80%. Chúng ta viết $X_A = (1, 5, 80)$, $X_B = (0, 5, 80)$. Từ **biểu thức 1.5**, chúng ta tính tỷ lệ sai khác giữa hai nhóm rừng như sau:

$$OR_{(X_A-X_B)} = \exp(1-0)b_1 + (5-5)b_2 + (80-80)b_3 = \exp(b_1)$$

Nếu so sánh $X_A = (1, 6, 80)$, $X_B = (0, 4, 80)$, chúng ta tính tỷ lệ sai khác giữa hai nhóm rừng như sau:

$$OR_{(X_A-X_B)} = \exp(1-0)b_1 + (6-4)b_2 + (80-80)b_3 = \exp(-b_1) + 2b_2.$$

Nếu so sánh $X_A = (1, 6, 80)$, $X_B = (0, 4, 70)$, chúng ta tính tỷ lệ sai khác giữa hai nhóm rừng như sau:

$$OR_{(X_A-X_B)} = \exp(1-0)b_1 + (6-4)b_2 + (80-70)b_3 = \exp(-b_1) + 2b_2 + 10b_3$$

Những so sánh khác cũng được thực hiện tương tự như trên.

4.1.2. Hồi quy logit đa biến

Giả sử nghiên cứu độ bắt gặp của sao đen, dầu rái, dầu song nằng, chò chai, vên vên, gõ đỏ... trong quan hệ với độ tàn che, pH, độ ẩm...

Cách thức đo đạc và tính toán như sau (**xem thêm mục 1.1.1**):

Xây dựng những tuyến (ngang 2 m, dài tùy ý) cắt ngang qua các trạng thái rừng. Trên mỗi tuyến, cứ sau 20 – 50 m bố trí 1 điểm đo. Tại điểm đo này (2*2m),

thống kê sự bắt gặp loài cây quan tâm (bắt gặp dù chỉ 1 cá thể - kí hiệu 1, không = 0) và đo độ tàn che (tầng tán cây gỗ, cây bụi) và độ che phủ của thảm cỏ (theo Druze), pH, độ ẩm...Số lượng điểm đo đặc như trên ước tính khoảng 200 – 300 điểm. Cách tính toán giống như mục 1.1.1.

4.2. THỦ TỤC XỬ LÝ MÔ HÌNH LOGISTIC

Bước 1. Nhập số liệu như bảng dưới đây:

Ví dụ độ bắt gặp loài cây tùy thuộc loại đất và môi trường

Tt	Độ bắt gặp	Đất	pH	Âm
1	1	1	6.0	70
2	0	1	3.5	78
3	0	1	4.0	72
4	1	1	5.5	80
...				
1	1	0	5.5	72
2	1	0	4.5	71
3	0	0	6.0	72
4	1	0	7.0	80
Ghi chú: Đất bazan = 1, đất khác bazan = 0				

Ví dụ độ bắt gặp loài cây tùy thuộc trạng thái rừng và môi trường

Tt	Độ bắt gặp	Code (III _A)	pH	Âm
1	1	1	6.0	70
2	0	1	3.5	78
3	0	1	4.0	72
4	1	1	5.5	80
...				

Theo cách trên, chúng ta có thể suy diễn cho nhiều đối tượng khác

Bước 2. Gọi chương trình như sau đây:

+ Đường cong sigmoid

Mô hình bậc 1: $\log(p/1-p) = b_0 + b_1 \cdot X$ (a)

Thủ tục

- **Bước 1.** Special > Advanced regression > logistic regression. Đến đây chọn P đưa vào Dependent, chọn X đưa vào quantitative. Sau đó chọn OK. Kết quả cho ta mô hình sigmoid (a).

- **Bước 2.** Đặt con trỏ tại cửa sổ Estimate regression model, nhấp chuột phải và chọn những thủ tục cần tính khác.

+ Hồi quy Logit Gauss

$$\text{Mô hình bậc 2: } \log(p/1-p) = b_0 + b_1 * X + b_2 * X^2 \quad (b)$$

Thủ tục

- **Bước 1.** Special > Advanced regression > logistic regression. Đến đây chọn P đưa vào Dependent, chọn X đưa vào quantitative. Sau đó chọn OK.
- **Bước 2.** Đặt con trỏ tại cửa sổ Estimate regression model, nhấp chuột phải và chọn Analys Option, rồi chọn second Order (bậc 2). Kết quả cho ta mô hình logit Gauss (b).
- **Bước 3.** Mở Tabular Option và chọn những thủ tục khác

+ Hồi quy Logit đa biến

$$\text{Mô hình 1: } \log(p/1-p) = b_0 + b_1 * X_1 + b_2 X_2 \quad (c)$$

$$\text{Mô hình 2: } \log(p/1-p) = b_0 + b_1 * X_1 + b_2 X_1^2 + b_3 X_2 + b_4 X_2^2 \quad (d)$$

Thủ tục

- **Bước 1.** Special > Advanced regression > logistic regression. Đến đây chọn P đưa vào Dependent, chọn các X_i đưa vào quantitative. Sau đó chọn OK.
- **Bước 2.** Mở Tabular Option và chọn những thủ tục khác

Bài V. ỨNG DỤNG MÔ HÌNH POISSON TRONG SINH THÁI QUAN XÃ

5.1. MÔ HÌNH POISSON

5.1.1. Quan hệ giữa tổ thành cây gỗ với chiều cao tán rừng

Chiều cao tán rừng được chia theo cấp, mỗi cấp từ 2-3 m. Ở mỗi trạng thái rừng, bố trí 3 - 5 ô tiêu chuẩn 0,1 – 0,2 ha. Trong mỗi ô tiêu chuẩn, thống kê những nội dung sau đây: tên loài, đo D, H của từng cây gỗ lớn ($D > 8-10$ cm). Sau đó xây dựng mô hình mô tả tổ thành loài cây ở mỗi cấp H như sau:

Để phân tích hồi quy Poisson, chúng ta gọi Y là biến phụ thuộc. Ở đây Y là số cây của một loài nghiên cứu nào đó thu được từ một nhóm cây được mô tả bằng tập biến dự đoán $X_1, X_2, \dots, X_k$.

Đối với cấp chiều cao i ($i = 1, 2, \dots, n$), chúng ta gọi:

- Y_i là số cá thể của một loài quan tâm nào đó;
- l_i là tổng số cây của mỗi cấp chiều cao được nghiên cứu tổ thành;
- X_i = tập hợp các giá trị $X_1, X_2, \dots, X_n$ của cấp chiều cao i;
- $b = (b_0, b_1, \dots, b_n)$ là tập hợp những tham số chưa biết;
- $\lambda(X_i, b)$ là hàm số đặc trưng đối với X_i và b, nghĩa là $\lambda(X_i, b) = \exp(b_0 + \sum_{i=1}^n b_i X_i)$.

Hàm $\lambda(X_i, b)$ mô tả tổ thành loài cây ở cấp chiều cao i. Khi đó số cá thể của loài quan tâm ở cấp chiều cao i là

$$E(Y_i) = \mu_i = l_i \lambda(X_i, b), \text{ với } i = 1, 2, \dots, n; \quad (2.1)$$

Khi Y_i có phân bố Poisson với trung bình μ_i , thì

$$P(Y_i; \mu_i) = \frac{\mu_i^{y_i} e^{-\mu_i}}{Y_i!}, \quad i = 1, 2, \dots, n. \quad (2.2)$$

$$\text{Hay } P(Y_i; b) = \frac{l_i \lambda(X_i, b)^{y_i} e^{-l_i \lambda(X_i, b)}}{Y_i!} \quad (2.3)$$

Trong **công thức 2.3**, $Y_i = 0, 1, \dots, \infty$ và $i = 1, 2, \dots, n$.

Mô hình biểu diễn ảnh hưởng của cấp chiều cao ($i = 1, 2, \dots, n$) đến tổ thành loài cây nào đó như sau:

$$\lambda(X_i, b) = \exp(b_0 + \sum_{i=1}^n b_i X_i) \quad (2.4)$$

Nếu muốn so sánh tổ thành trong từng cấp chiều cao ($i = 1, 2, \dots, n$) ở hai địa phương ($j = 0$ và $j = 1$) hay hai trạng thái rừng bằng mô hình hồi quy Poisson, chúng ta xây dựng **mô hình 1** như sau:

$$E(Y_{ij}) = \mu_{ij} = l_{ij} * \lambda_{ij}$$

$$\text{với } \ln(\lambda_{ij}) = \alpha + \sum_{k=1}^7 \alpha_k U_k + bE \quad (2.5)$$

$$\text{hay } \lambda_{ij} = \exp(\alpha + \sum_{k=1}^7 \alpha_k U_k + bE)$$

Trong **công thức 2.5**, U_k là biến giả nhận hai giá trị 0 và 1 để biểu thị cấp chiều cao, còn E là biến giả nhận giá trị 0 và 1 để mô tả địa phương (vùng $A = 0$, vùng $B = 1$) hay trạng thái rừng. Các tham số của **mô hình 2.5** được xác định bằng nguyên lý hợp lý tối đa. Tính phù hợp của mô hình hồi quy Poisson cũng được kiểm định bằng thống kê χ^2 .

Nói chung, việc xác định các tham số (α_0, α_i và b) của **mô hình 2.5** được thực hiện dễ dàng bằng những chương trình máy tính.

Từ mô hình Poisson, chúng ta có thể tính được tỷ lệ tổ thành ở tất cả các cấp chiều cao theo công thức:

$$RR_i = \frac{\lambda_{i1}}{\lambda_{i0}} = e^b. \quad (2.6)$$

Khoảng tin cậy 95% của mẫu lớn đối với e^b được tính theo công thức:

$$\exp[b \pm S_{e(Sb^*)}] \quad (2.7)$$

trong đó b^* là giá trị ước lượng của hệ số b ; $S_{e(Sb^*)}$ là sai số chuẩn của hệ số ước lượng b^* .

Việc kiểm định giả thuyết $H_0: b = 0$, đối thuyết $H_A: b \neq 0$ có thể được thực hiện dựa trên thống kê Wald:

$$Z = \frac{b^* - 0}{S_{e(Sb^*)}} \quad (2.8)$$

ở đây Z có phân bố $N(0,1)$ dưới giả thuyết $H_0: b = 0$.

5.1.2. Quan hệ giữa tổ thành cây gỗ với độ cao địa hình

Địa hình được chia theo cấp, mỗi cấp từ 50 – 100 m/cấp. Trên mỗi cấp địa hình, bố trí 3 - 5 ô tiêu chuẩn 0,1 – 0,2 ha. Trong mỗi ô tiêu chuẩn, thống kê những nội dung sau đây: tên loài, đo D , H của từng cây gỗ lớn ($D > 8-10$ cm). Sau đó mô tả tổ thành loài cây theo mô hình: $\lambda(X_i, b) = \exp(b_0 + \sum_{i=1}^n b_i X_i)$. Nếu muốn so sánh tổ thành trong từng cấp chiều cao ($i = 1, 2, \dots, n$) ở hai địa phương ($j = 0$ và $j = 1$)

hay hai trạng thái rừng bằng mô hình hồi quy Poisson, chúng ta xây dựng **mô hình 1** như sau:

$$E(Y_{ij}) = \mu_{ij} = l_{ij} * \lambda_{ij}$$

$$\text{với } \ln(\lambda_{ij}) = \alpha + \sum_{k=1}^7 \alpha_k U_k + bE \text{ hay } \lambda_{ij} = \text{Exp}(\alpha + \sum_{k=1}^7 \alpha_k U_k + bE)$$

5.1.3. Xác định tỷ lệ cây gỗ quý theo cấp đường kính

Để giải quyết nội dung này, hãy phân chia đường kính thân cây thành cấp, mỗi cấp 4 – 8 cm. Trong mỗi cấp D thống kê tên loài cây, trong đó lưu ý đến những cây gỗ quý. Yêu cầu về lý thuyết: mỗi cấp kính và cấp chiều cao phải có ít nhất 100 cá thể (càng nhiều càng chính xác). Cách tính toán giống như **mục 2.1.2**

5.1.4. Xác định tỷ lệ cây tái sinh của những loài cây gỗ quý theo lớp chiều cao

Để làm rõ nội dung này, trước hết đo đặc cây tái sinh theo cấp H; mỗi cấp cần có hàng ngàn cây. Ở mỗi cấp chiều cao, hãy chú ý đến số lượng cây gỗ quý hiếm. Cách tính toán giống như **mục 2.1.2**.

5.1.5. Xác định tỷ lệ cây gỗ bị rồng ruột theo cấp đường kính

Để minh họa tính ứng dụng của phân tích hồi quy Poisson, chúng ta xem xét số liệu so sánh số cá thể của loài sao đen thuộc những cấp kính khác nhau bị bệnh rồng ruột ở vùng A và vùng B (**bảng 2.1**).

Bảng 2.1. So sánh số cây sao đen bị bệnh rồng ruột ở vùng A và B

TT	Cấp D, cm	Vùng A			Vùng B			RRi
		Y_{iA}	(l_{iA})	λ_{iA}	Y_{iB}	(l_{iB})	λ_{iB}	
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
1	10-20	2	5.500	0,00036	5	2.500	0,00200	5,5
2	21-30	4	6.000	0,00067	8	4.500	0,00178	2,7
3	31-40	5	8.500	0,00059	12	12.000	0,00100	1,7
4	41-50	8	10.000	0,00080	20	6.850	0,00292	3,6
5	51-60	10	9.500	0,00105	25	5.800	0,00431	4,1
6	61-70	11	5.500	0,00200	28	4.500	0,00622	3,1
7	71-80	12	4.550	0,00264	22	3.000	0,00733	2,8
8	> 81	6	10.000	0,00060	15	8.500	0,00176	2,9

(*) Nhóm A được sử dụng làm đối tượng so sánh

Ở ví dụ này, biến phụ thuộc Y là số cây sao đen bị bệnh rồng ruột. Vì có tám cấp kính và 2 địa phương được xem xét, chúng ta kí hiệu Y_{ij} là số cây sao đen ở cấp kính i thuộc địa phương j; trong đó $i = 1, 2, \dots, 8$; $j = 0$ (vùng A), $j = 1$ (vùng B). Chúng ta sử dụng l_{ij} để chỉ kích thước quần thể sao đen được nghiên cứu bệnh rồng ruột ở cấp kính i thuộc vùng j.

Mục đích của nghiên cứu này là xác định tỷ lệ hay xác suất phát triển bệnh trĩ ruột ở những cây sao đen thuộc những cấp kính khác nhau mọc ở vùng A có lớn hơn vùng B hay không.

NHỮNG MÔ HÌNH POISSON CẦN PHÂN TÍCH

Mô hình 1. Phân tích tỷ lệ bệnh trĩ ruột tùy thuộc vào cấp kính và địa phương. Mô hình có dạng:

$$\ln(\lambda_{ij}) = \alpha + \sum_{k=1}^7 \alpha_k U_k + bE \quad (a)$$

$$\text{hay } \lambda_{ij} = \text{Exp}(\alpha + \sum_{k=1}^7 \alpha_k U_k + bE)$$

Ở đây E = 1 và 0 = chỉ đối tượng so sánh; k = 1 – 7 (nhóm cấp kính)
 Từ ví dụ ở **bảng 2.1**, chúng ta nhận được kết quả sau đây (**bảng 2.2 – 2.5**):

Bảng 2.2. Mô hình hồi quy Poisson ước lượng

Tham số	Ước lượng	Sai số chuẩn	RRi ^(*)
Hằng số	-7,46934	0,24666	
Vùng	1,14918	0,15785	3,16
Cấp kính: 1	-0,08693	0,43698	0,92
Cấp kính: 2	0,04080	0,36189	1,04
Cấp kính: 3	-0,44180	0,32655	0,64
Cấp kính: 4	0,44010	0,28876	1,55
Cấp kính: 5	0,79180	0,27625	2,21
Cấp kính: 6	1,24452	0,27067	3,47
Cấp kính: 7	1,44770	0,27767	4,25

(*) Nhóm A được sử dụng làm đối tượng so sánh

Bảng 2.3. Phân tích biến động¹

Nguồn	Biến động	df	P
Mô hình	126,871	8	0,0000
Sai lệch	2,5459	7	0,9236
Tổng	129,417	15	

Bảng 2.4. Kiểm định Likelihood Ratio

Yếu tố	χ^2	df	P
Vùng	58,69	1	0,0000
Cấp kính	75,72	7	0,0000

¹ Deviance

Bảng 2.5. Kết quả dự đoán số cây sao đen bị bệnh rồng ruột ở vùng A và B

TT	Cấp D, cm	Vùng A			Vùng B			RRi ^(*)
		Y _{iA}	(I _{iA})	λ _{iA}	Y _{iB}	(I _{iB})	λ _{iB}	
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
1	10-20	2,9	5.500	0,00052	4,1	2.500	0,00165	3,2
2	21-30	3,5	6.000	0,00055	8,4	4.500	0,00187	3,4
3	31-40	3,1	8.500	0,00037	13,9	12.000	0,00116	3,2
4	41-50	8,9	10.000	0,00089	19,1	6.850	0,00280	3,2
5	51-60	12,0	9.500	0,00126	23,0	5.800	0,00397	3,2
6	61-70	10,9	5.500	0,00198	28,1	4.500	0,00625	3,2
7	71-80	11,0	4.550	0,00243	23,0	3.000	0,00766	3,2
8	> 81	5,7	10.000	0,00057	15,3	8.500	0,00180	3,2

(*) Nhóm A được sử dụng làm đối tượng so sánh

Thủ tục xử lý mô hình 1: $\ln(\lambda_{ij}) = \alpha + \sum_{k=1}^7 \alpha_k U_k + bE$

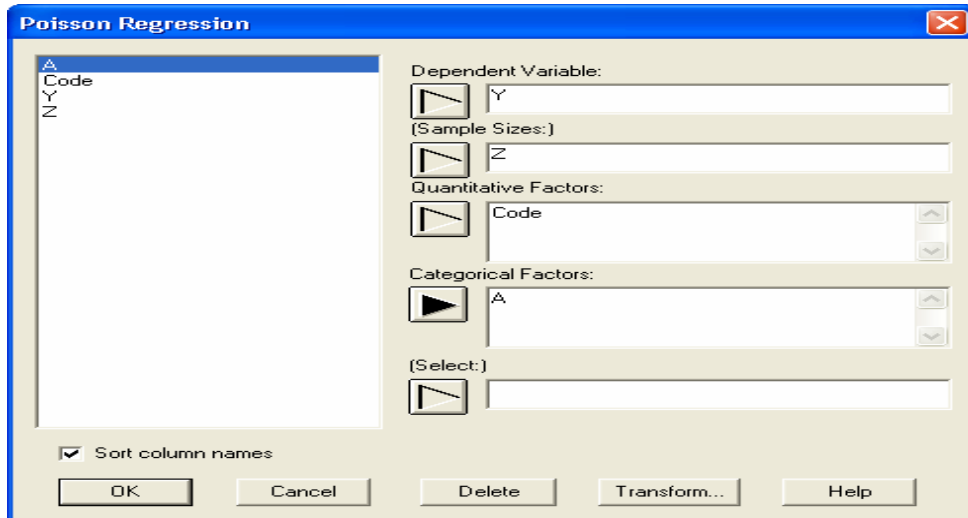
Lưu ý rằng, kết quả báo cáo $\lambda_{ij} = \text{Exp}(\alpha + \sum_{k=1}^7 \alpha_k U_k + bE)$

(1) Nhập biến Y, kích thước mẫu (Z), cấp kính (D) hay lớp tuổi (A), mã vùng (Code = 0 vùng A, code = 1 vùng B) vào bốn cột khác nhau như sau:

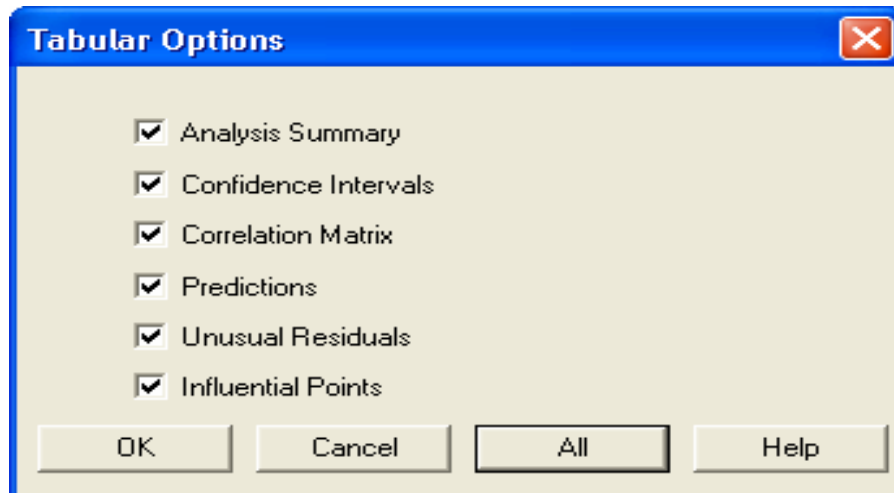
A	Z	Y	Code
1	5500	2	0
2	65250	4	0
3	86215	5	0
4	100000	8	0
5	95155	10	0
6	55720	11	0
7	38850	12	0
8	10320	6	0
1	18350	5	1
2	145200	8	1
3	120375	12	1
4	115350	20	1
5	85005	25	1
6	65930	28	1
7	30005	22	1
8	8500	15	1

(2) Gọi Special > Advanced Regression > Poisson Regression > OK

(3) Đưa Y vào Dependent, Z vào Sample Sizes, Code (E = vùng) vào quantitative Factors, D vào Categorical Factors > OK.



(4) Mở Tabular Options, chọn All để xác định khoảng tin cậy của các hệ số, ma trận tương quan, dự đoán tần số trong mỗi lớp tuổi.



(5) Để dự đoán tần số bệnh (Y) của từng lớp tuổi, trước hết chọn cửa sổ Predictions For Y, kế đến đặt con trỏ chuột ở trang này và chọn chuột phải, sau đó chọn Pane Options và All values.

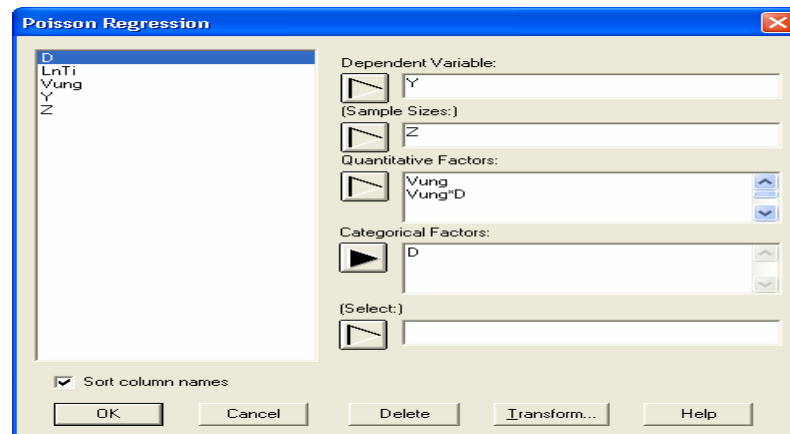
Mô hình 2. Phân tích ảnh hưởng tương tác giữa cấp kính và địa phương đến tỷ lệ bệnh rỗng ruột. Mô hình có dạng:

$$\ln(\lambda_{ij}) = \alpha + \sum_{k=1}^7 \alpha_k U_k + bE + \sum_{k=1}^7 \delta_k (U_k * E) \quad (b)$$

Ở đây E = 1 và 0 = chỉ đối tượng so sánh; k = 1 – 7 (nhóm cấp kính). Khi gọi chương trình lưu ý biến đổi số liệu để loại bỏ cộng tuyến tính.

Thủ tục gọi mô hình 2

- ✓ Nhập biến Y, kích thước mẫu (Z), lớp tuổi (D, A hay U_k), mã vùng (Code) vào bốn cột khác nhau.
- ✓ Gọi Special > Advanced Regression > Poisson Regression > OK
- ✓ Đưa Y vào Dependent, Z vào Sample Sizes, Code (vùng) và Vùng*D vào quantitative Factors, D vào Categorical > OK (Hình)



- ✓ Mở Tabular Options, chọn All để xác định khoảng tin cậy của các hệ số, ma trận tương quan, dự đoán tần số trong mỗi lớp tuổi.
- ✓ Để dự đoán tần số bệnh (Y) của từng lớp tuổi, chọn cửa sổ Predictions For Y, sau đó đặt con trỏ chuột ở trang này và chọn chuột phải, sau đó chọn Pane Options và All values.

Mô hình 3. Mô hình ba trả lời một số câu hỏi sau đây:

Mô hình 3a. Nếu tỷ lệ cây bị bệnh rồng ruột ở những cấp kính khác nhau của hai vùng là không khác nhau, thì có cần phải xem xét ảnh hưởng của cấp kính hay không? Để trả lời câu hỏi này, chúng ta xây dựng mô hình sau đây:

$$\ln(\lambda_{ij}) = \alpha + bE \quad (c)$$

Ở đây E = 1 và 0 = chỉ đối tượng so sánh.

Thủ tục xử lý mô hình 3a: $\ln(\lambda_{ij}) = \alpha + bE$

Lưu ý rằng, kết quả báo cáo $\lambda_{ij} = \text{Exp}(\alpha + bE)$

- (1) Gọi Special > Advanced Regression > Poisson Regression > OK

- (2) **Gọi chương trình theo cách 1.** Đưa Y vào Dependent, Z vào Sample Sizes, Code (vùng) vào quantitative Factors, sau đó chọn OK.
- (3) **Gọi chương trình theo cách 2.** Trước hết đưa Y vào Dependent, Z vào Sample Sizes, Code (vùng) vào Categorical Factors, D vào quantitative Factors > OK. Sau đó tại cửa sổ Poisson Regression, chọn Analysis options. Kế đến chọn Exclude, sau đó nhấp chuột hai lần vào yếu tố D để bỏ D, rồi chọn OK.

Mô hình 3b. Nếu tỷ lệ cây bị bệnh rỗng ruột ở những địa phương khác nhau là không khác nhau, thì có cần phải xem xét ảnh hưởng của địa phương hay không? Nói cách khác, **mô hình 3b** chỉ mô tả ảnh hưởng của cấp kính. Để trả lời câu hỏi này, chúng ta xây dựng mô hình sau đây:

$$\ln(\lambda_{ij}) = \alpha + \sum_{k=1}^7 \alpha_k U_k \quad (d)$$

Ở đây $k = 1 - 7$ (nhóm cấp kính).

Thủ tục xử lý mô hình 3b: $\ln(\lambda_{ij}) = \alpha + \sum_{k=1}^7 \alpha_k U_k$
 Lưu ý rằng, kết quả báo cáo $\lambda_{ij} = \text{Exp}(\alpha + \sum_{k=1}^7 \alpha_k U_k)$

- (1) Gọi Special > Advanced Regression > Poisson Regression > OK
- (2) **Gọi chương trình theo cách 1.** Đưa Y vào Dependent, Z vào Sample Sizes, D vào quantitative Factors, sau đó chọn OK.
- (3) **Gọi chương trình theo cách 2.** Trước hết đưa Y vào Dependent, Z vào Sample Sizes, Code (vùng) vào quantitative Factors, D vào Categorical Factors > OK. Sau đó tại cửa sổ Poisson Regression bạn chọn Analysis options. Kế đến chọn Exclude, sau đó nhấp chuột hai lần vào yếu tố E để bỏ E, rồi chọn OK.

Mô hình 4. Xử lý mô hình hồi quy Poisson theo một cách khác. Trong trường hợp này, yếu tố cấp đường kính được xử lý như là biến đo khoảng. Cách thức xử lý như sau:

+ Trước hết biến D_i được biến đổi như sau:

$$D_i = \frac{[\text{Giá trị giữa của cấp đường kính thứ } i] - D_{\min}}{35}, \text{ với } i = 1, 2, \dots, 8.$$

+ Kế đến lấy $\ln D_i$ và $\ln \lambda_{ij}$. Từ ví dụ ở **bảng 2.1**, chúng ta biến đổi thành **bảng 2.6**.

+ Sau đó xây dựng mô hình chỉ bao gồm ảnh hưởng tuyến tính của yếu tố đường kính, không có tương tác giữa đường kính với vùng (E). Nói một cách cụ thể, chúng ta xem xét **mô hình 4** sau đây:

$$\begin{aligned} \ln \hat{\lambda}_{ij} &= \alpha + \theta * \ln D_i + bE \\ \text{hay } \hat{\lambda}_{ij} &= \text{Exp}(\alpha + \theta * \ln D_i + bE) \end{aligned} \quad (e)$$

Để giải mô hình 4, chúng ta biến đổi số liệu như **bảng 2.6**. Kết quả xử lý được ghi lại ở **bảng 2.7-2.9**.

Bảng 2.6. Giá trị $\ln\lambda_{i0}$ và $\ln\lambda_{i1}$ được biến đổi từ số liệu ở **bảng 2.1**

Cấp D, cm	Giữa cấp D	Di	$\ln Di$	$\ln\lambda_{i0}$	$\ln\lambda_{i1}$
(1)	(2)	(3)	(4)	(5)	(6)
10-20	15	0,14286	-1,9459	-7,9194	-6,2146
21-30	25	0,42857	-0,8473	-7,3132	-6,3324
31-40	35	0,71429	-0,3365	-7,4384	-6,9078
41-50	45	<u>1,00000</u>	<u>0,0000</u>	-7,1309	-5,8363
51-60	55	1,28571	0,2513	-6,8565	-5,4467
61-70	65	1,57143	0,4520	-6,2146	-5,0796
71-80	75	1,85714	0,6190	-5,938	-4,9153
> 81	85	2,14286	0,7621	-7,4186	-6,3398

Bảng 2.7.. Mô hình hồi quy Poisson ước lượng

Tham số	Ước lượng	Sai số chuẩn	RRi
Hằng số	-6,97112	0,13327	
Vùng	1,06251	0,15700	2,89361
$\ln T_i$	0,50885	0,12541	1,66337

Bảng 2.8. Phân tích biến động²

Nguồn	Biến động	df	P
Mô hình	70,0166	2	0,0000
Sai lệch	59,4001	13	0,0000
Tổng		15	

Bảng 2.9. Kiểm định Likelihood Ratio

Yếu tố	χ^2	df	P
Vùng	50,632	1	0,0000
Cấp kính	18,866	1	0,0000

Từ kết quả ở **bảng 2.7**, chúng ta có mô hình hồi quy ước lượng như sau:

$$\hat{\lambda}_{ij} = \exp(-6,97112 + 1,06251 \cdot \text{vùng} + 0,50885 \cdot \ln T_i)$$

$$\ln \hat{\lambda}_{ij} = (-6,97112 + 1,06251 \cdot \text{vùng} + 0,50885 \cdot \ln T_i)$$

Kết quả phân tích ở **bảng 2.8** chỉ ra mô hình 4 là phù hợp ($P < 0,01$). Do đó, ước lượng tỷ lệ sai khác về bệnh trĩ ruột ở loài sao đen mọc ở hai vùng sinh thái

² Deviance

khác nhau là $RR_i = e^{b^*} = e^{1,06251} = 2,89$. Khoảng tin cậy 95% của $e^{b^*} = \exp[b^* \pm 1,96 * Se(s_{b^*})] = \exp(1,06251 \pm 1,96 * 0,157) = (2,13; 3,94)$. Như vậy, sử dụng mô hình 4 cũng dẫn đến câu trả lời tương tự như mô hình 1, nhưng độ tin cậy cao hơn (bởi vì khoảng tin cậy của mô hình 4 hẹp hơn). So với mô hình 1, mô hình 4 đơn giản hơn vì chứa ít tham số hơn. Ngoài ra, mô hình 4 có hiệu lực cao hơn mô hình 1, bởi vì mô hình 4 chỉ mô tả ảnh hưởng tuyến tính của tuổi. Vì thế, mô hình 4 được chọn để mô tả tỷ lệ bệnh rối ruột ở loài sao đen mọc ở hai địa phương khác nhau.

Thủ tục xử lý mô hình 4

Lưu ý rằng, kết quả báo cáo $\lambda_{ij}^* = \text{Exp}(\alpha + \theta * \ln D_i + bE)$

- (1) Nhập biến Y, kích thước mẫu (Z), cấp kính (D_i), $\ln D_i$ (biến này đã biến đổi) và mã vùng (Code) vào bốn cột khác nhau.
- (2) Gọi Special > Advanced Regression > Poisson Regression > OK
- (3) Đưa Y vào Dependent, Z vào Sample Sizes, Code (E = vùng) và $\ln D_i$ vào quantitative Factors và OK.
- (4) Mở Tabular Options, chọn All để xác định khoảng tin cậy của các hệ số, ma trận tương quan, dự đoán tần số trong mỗi lớp tuổi.
- (5) Để dự đoán tần số bệnh (Y) của từng lớp tuổi, trước hết chọn cửa sổ Predictions For Y, kế đến đặt con trỏ chuột ở trang này và chọn chuột phải, sau đó chọn Pane Options và All values.

THỦ TỤC GỌI HÀM HỒI QUY POISSON

Sử dụng Statgraphics Plus Version 5.1

Ứng dụng 1. Xác định mô hình biểu thị tỷ lệ rối ruột tùy theo cấp kính và vùng

Mô hình có dạng : $\ln(\lambda_{ij}) = \alpha + \sum_{k=1}^7 \alpha_k U_k + bE$

Thực hành

Bài tập lấy từ bảng 24.1 hoặc 7.6

Bảng 7.6 ghi lại số liệu điều tra tỷ lệ bệnh rối ruột của sao đen tùy thuộc vào cấp D và hai vùng sinh thái khác nhau.

Bảng 7.6. So sánh số cây sao đen bị bệnh rối ruột ở vùng A và B

TT	Cấp D, cm	Vùng A			Vùng B			RRi
		Y_{iA}	(I_{iA})	λ_{iA}	Y_{iB}	(I_{iB})	λ_{iB}	
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
1	10-20	2	5.500	0,00036	5	2.500	0,00200	5,5
2	21-30	4	6.000	0,00067	8	4.500	0,00178	2,7
3	31-40	5	8.500	0,00059	12	12.000	0,00100	1,7
4	41-50	8	10.000	0,00080	20	6.850	0,00292	3,6
5	51-60	10	9.500	0,00105	25	5.800	0,00431	4,1
6	61-70	11	5.500	0,00200	28	4.500	0,00622	3,1
7	71-80	12	4.550	0,00264	22	3.000	0,00733	2,8
8	> 81	6	10.000	0,00060	15	8.500	0,00176	2,9

(*) Nhóm A được sử dụng làm đối tượng so sánh

Giải thích số liệu bảng 7.6

- ✓ Cột 1 = cấp D
- ✓ Cột 2 & 6 = số cây bị bệnh rối ruột
- ✓ Cột 4 & 7 = dung lượng mẫu quan sát ở mỗi cấp D
- ✓ Cột 5 & 8 = cột 3/cột 4 và cột 6/cột 7
- ✓ Cột 9 = [(cột 6/cột 7)/(cột 3/cột 4)] = cột 8/cột 5.

Mô hình 1: Tỷ lệ rối ruột tùy theo cấp kính và vùng

Những nội dung cần tính toán:

Nội dung 1. Xây dựng mô hình biểu thị tỷ lệ rối ruột của sao đen tùy theo cấp kính và vùng sinh thái.

Mô hình tổng quát có dạng: $\ln(\lambda_{ij}) = \alpha + \sum_{k=1}^7 \alpha_k U_k + bE$

Thủ tục xây dựng mô hình (Statgraphics Plus 5.1)

1. Nhập biến Y, kích thước mẫu (Z), lớp tuổi (A), mã vùng (Code = E) vào bốn cột khác nhau.
2. Gọi Special > Advanced Regression > Poisson Regression > OK
3. Đưa Y vào Dependent, Z vào Sample Sizes, Code (vùng = E) vào quantitative Factors, A vào Categorical Factors > OK.

Kết quả xác định được mô hình Poisson như sau (Bảng 7.7):

Bảng 7.7. Mô hình hồi quy Poisson ước lượng

Tham số	Ước lượng	Sai số chuẩn	RRi ^(*)
Hằng số	-7,46934	0,24666	
Vùng	1,14918	0,15785	3,16
Cấp kính: 1	-0,08693	0,43698	0,92
Cấp kính: 2	0,04080	0,36189	1,04
Cấp kính: 3	-0,44180	0,32655	0,64
Cấp kính: 4	0,44010	0,28876	1,55
Cấp kính: 5	0,79180	0,27625	2,21
Cấp kính: 6	1,24452	0,27067	3,47
Cấp kính: 7	1,44770	0,27767	4,25

(*) Nhóm A được sử dụng làm đối tượng so sánh

$\lambda_{ij} = \exp(-7,46934 + 1,14918 \cdot \text{vùng} - 0,08693 \cdot D=1 + 0,0408 \cdot D=2 - 0,4418 \cdot D=3 + 0,4401 \cdot D=4 + 0,7918 \cdot D=5 + 1,2445 \cdot D=6 + 1,4477 \cdot D=7)$ (mô hình 1)

Nội dung 2. Dự đoán tỷ lệ cây sao đen bị bệnh rỗng ruột tùy theo cấp D ở vùng A và B. Tỷ lệ rỗng ruột ở sao đen được tính từ công thức 7.37 như sau:

$$\lambda_{ij} = \text{Exp}(\alpha + \sum_{k=1}^7 \alpha_k U_k + bE) \quad (7.37)$$

Khi E (vùng A) = 0, tỷ lệ cây bị rỗng ruột ở các cấp kính được tính theo công thức:

$$\text{Cấp kính 1: } \lambda_{10} = \exp(\alpha + \alpha_1 \cdot 1)$$

$$\text{Cấp kính 2: } \lambda_{20} = \exp(\alpha + \alpha_2 \cdot 1)$$

$$\text{Cấp kính 3: } \lambda_{30} = \exp(\alpha + \alpha_3 \cdot 1)$$

$$\text{Cấp kính 4: } \lambda_{40} = \exp(\alpha + \alpha_4 \cdot 1)$$

$$\text{Cấp kính 5: } \lambda_{50} = \exp(\alpha + \alpha_5 \cdot 1)$$

$$\text{Cấp kính 6: } \lambda_{60} = \exp(\alpha + \alpha_6 \cdot 1)$$

$$\text{Cấp kính 7: } \lambda_{70} = \exp(\alpha + \alpha_7 \cdot 1)$$

$$\text{Cấp kính 8: } \lambda_{80} = \exp(\alpha)$$

Khi E (vùng B) = 1, tỷ lệ cây bị rỗng ruột ở các cấp kính được tính theo công thức:

- Cấp kính 1: $\lambda_{11} = \exp(\alpha + \alpha_1 * 1 + b)$
 Cấp kính 2: $\lambda_{21} = \exp(\alpha + \alpha_2 * 1 + b)$
 Cấp kính 3: $\lambda_{31} = \exp(\alpha + \alpha_3 * 1 + b)$
 Cấp kính 4: $\lambda_{41} = \exp(\alpha + \alpha_4 * 1 + b)$
 Cấp kính 5: $\lambda_{51} = \exp(\alpha + \alpha_5 * 1 + b)$
 Cấp kính 6: $\lambda_{61} = \exp(\alpha + \alpha_6 * 1 + b)$
 Cấp kính 7: $\lambda_{71} = \exp(\alpha + \alpha_7 * 1 + b)$
 Cấp kính 8: $\lambda_{81} = \exp(\alpha + b)$

Áp dụng vào ví dụ này, từ mô hình 1, ta có:

+ Khi $E = 0$ (vùng A), từ mô hình 1 ta có:

- Cấp kính 1. $\lambda_{1A} = \exp(-7,46934 - 0,08693 * 1) = \exp(-7,55627)$
 Cấp kính 2. $\lambda_{2A} = \exp(-7,46934 + 0,0408 * 1) = \exp(-7,51014)$
 Cấp kính 3. $\lambda_{3A} = \exp(-7,46934 - 0,4418 * 1) = \exp(-7,91114)$
 Cấp kính 4. $\lambda_{4A} = \exp(-7,46934 + 0,4401 * 1) = \exp(-7,02924)$
 Cấp kính 5. $\lambda_{5A} = \exp(-7,46934 + 0,7918 * 1) = \exp(-6,67754)$
 Cấp kính 6. $\lambda_{6A} = \exp(-7,46934 + 1,2445 * 1) = \exp(-6,22484)$
 Cấp kính 7. $\lambda_{7A} = \exp(-7,46934 + 1,4477 * 1) = \exp(-6,02164)$
 Cấp kính 8. $\lambda_{8A} = \exp(-7,46934)$

+ Khi $E = 1$ (vùng B), từ mô hình 1 ta có:

- Cấp kính 1. $\lambda_{1B} = \exp(-7,46934 + 1,14918 * 1 - 0,08693 * 1)$
 Cấp kính 2. $\lambda_{2B} = \exp(-7,46934 + 1,14918 * 1 + 0,0408 * 1)$
 Cấp kính 3. $\lambda_{3B} = \exp(-7,46934 + 1,14918 * 1 - 0,4418 * 1)$
 Cấp kính 4. $\lambda_{4B} = \exp(-7,46934 + 1,14918 * 1 + 0,4401 * 1)$
 Cấp kính 5. $\lambda_{5B} = \exp(-7,46934 + 1,14918 * 1 + 0,7918 * 1)$
 Cấp kính 6. $\lambda_{6B} = \exp(-7,46934 + 1,14918 * 1 + 1,2445 * 1)$
 Cấp kính 7. $\lambda_{7B} = \exp(-7,46934 + 1,14918 * 1 + 1,4477 * 1)$
 Cấp kính 8. $\lambda_{8B} = \exp(-7,46934 + 1,14918 * 1)$

Lưu ý: Để dự đoán tần số bệnh (Y) của từng lớp tuổi hay cấp kính, trước hết chọn cửa sổ Predictions For Y, kế đến đặt con trỏ chuột ở trang này và chọn chuột phải, sau đó chọn Pane Options và All values.

Nội dung 3. Tính tần số cây bị rồng ruột ở các cấp kính

Số cây bị rồng ruột ở các cấp kính được tính theo công thức:

$$E(Y_{ij}) = \mu_{ij} = l_{ij} * \lambda_{ij} \quad (a)$$

Từ (a), ta có các giá trị Y_{iA} ở cột 3 = cột 4 * cột 5. Tương tự, các giá trị Y_{iB} ở cột 6 = cột 7 * cột 8. Tất cả những tính toán này được ghi lại ở **bảng 7.10**.

Lưu ý rằng, để dự đoán tần số bệnh (Y) của từng lớp tuổi hay cấp kính ở bảng 7.10, sau khi gọi đến bước 3, ta chọn cửa sổ Predictions For Y. Kế đến đặt con trỏ chuột ở trang này và chọn chuột phải, sau đó chọn Pane Options và All values.

Bảng 7.10. Kết quả dự đoán số cây sao đen bị bệnh rỗng ruột ở vùng A và B

TT	Cấp D, cm	Vùng A			Vùng B			RRi ^(*)
		Y _{iA}	(I _{iA})	λ _{iA}	Y _{iB}	(I _{iB})	λ _{iB}	
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
1	10-20	2,9	5.500	0,00052	4,1	2.500	0,00165	3,2
2	21-30	3,5	6.000	0,00055	8,4	4.500	0,00187	3,4
3	31-40	3,1	8.500	0,00037	13,9	12.000	0,00116	3,2
4	41-50	8,9	10.000	0,00089	19,1	6.850	0,00280	3,2
5	51-60	12,0	9.500	0,00126	23,0	5.800	0,00397	3,2
6	61-70	10,9	5.500	0,00198	28,1	4.500	0,00625	3,2
7	71-80	11,0	4.550	0,00243	23,0	3.000	0,00766	3,2
8	> 81	5,7	10.000	0,00057	15,3	8.500	0,00180	3,2

(*) Nhóm A được sử dụng làm đối tượng so sánh

Nội dung 4. Ước lượng chung về tỷ lệ sao đen bị rỗng ruột ở hai vùng sinh thái.

Từ mô hình 1 cho thấy, bệnh rỗng ruột ở sao đen phụ thuộc vào cả vùng sinh thái và cấp tuổi hay cấp D. Trong trường hợp này, hệ số ước lượng biểu thị ảnh hưởng của vùng sinh thái đến tỷ lệ cây sao đen bị bệnh rỗng ruột là $b^{\wedge} = 1,14918$, với sai số $Se(s_{b^{\wedge}}) = 0,15785$. Giả thuyết $H_0: b = 0$ bị bác bỏ, bởi vì $Z = 7,28$ với $P < 0,01$.

Do đó, ước lượng tỷ lệ sai khác về bệnh rỗng ruột ở sao đen giữa hai vùng là $RR_i = e^{b^{\wedge}} = e^{1,14918} = 3,16$.

Khoảng tin cậy 95% của $e^{b^{\wedge}} = \exp[b^{\wedge} \pm 1,96 * Se(s_{b^{\wedge}})] = \exp(1,14918 \pm 1,96 * 0,15785) = (2,32; 4,30)$.

Nội dung 5. Nếu tỷ lệ cây bị bệnh rỗng ruột ở những cấp kính khác nhau của hai vùng là không khác nhau rõ rệt, thì mô hình biểu thị tỷ lệ rỗng ruột của sao đen ở hai vùng sinh thái có dạng (sách gọi là mô hình 3a):

$$\ln(\lambda_{ij}) = \alpha + bE \text{ hay } \lambda_{ij} = \exp(\alpha + bE) \quad (7.47)$$

Lưu ý rằng, nếu bỏ yếu tố cấp đường kính, chúng ta cần phải xem xét hệ số $b^{\wedge}$ hoặc $e^{b^{\wedge}}$ thay đổi có ý nghĩa hay không. Nói một cách cụ thể, khi bỏ thành phần tuổi hay cấp D trong mô hình 1 (nghĩa là $\sum_{k=1}^7 \alpha_k U_k$), chúng ta cần phải xác định hệ số ước lượng của E trong mô hình 7.47 thay đổi có ý nghĩa hay không.

Khi xử lý với mô hình 1, hệ số của E là 1,14918, do đó $e^{1,14918} = 3,16$.

Nếu xử lý mô hình 3 theo số liệu của **bảng 7.6**, chúng ta sẽ nhận được kết quả sau đây:

$$\ln(\lambda_j) = -6,9341 + 1,0678 * \text{vùng} \text{ hay } \lambda_j = \exp(-6,9341 + 1,0678 * \text{vùng})$$

Từ kết quả ở mô hình 3 cho thấy, $\hat{b} = 1,0678$, $Se(s_{\hat{b}}) = 0,1570$ và $RRi = e^{\hat{b}} = e^{1,0678} = 2,91$. Khoảng tin cậy 95% của $e^{\hat{b}} = \exp[\hat{b} \pm 1,96 * Se(s_{\hat{b}})] = \exp(1,0678 \pm 1,96 * 0,1570) = (2,14; 3,95)$. Như vậy, so với mô hình 1, hệ số $\hat{b}$ ở mô hình 3 thay đổi đáng kể. Điều đó cho thấy, để dự đoán tỷ lệ bệnh rỗng ruột ở sao đen chúng ta không thể bỏ qua yếu tố cấp đường kính. Ở đây chúng ta xử lý cấp đường kính như là biến số thực bằng việc sử dụng 7 dạng thức $\alpha_k U_k$ ($k = 1, 2, \dots, 7$) cộng với hệ số α trong mô hình 1 để phản ánh 8 cấp kính.

Thủ tục xử lý mô hình 3: $\ln(\lambda_j) = \alpha + bE$ hay $\lambda_j = \text{Exp}(\alpha + bE)$

(4) Gọi Special > Advanced Regression > Poisson Regression > OK

(5) **Gọi chương trình theo cách 1.** Đưa Y vào Dependent, Z vào Sample Sizes, Code (vùng) vào Categorical Factors, sau đó chọn OK.

(6) **Gọi chương trình theo cách 2.** Trước hết đưa Y vào Dependent, Z vào Sample Sizes, Code (vùng) vào quantitative Factors, D vào Categorical Factors > OK. Sau đó tại cửa sổ Poisson Regression bạn chọn Analysis options. Kế đến chọn Exclude, sau đó nhấp chuột hai lần vào yếu tố D để bỏ D, rồi chọn OK.

(7) **Để dự đoán tần số bệnh (Y) của từng lớp tuổi** hay cấp kính, trước hết chọn cửa sổ Predictions For Y, kế đến đặt con trỏ chuột ở trang này và chọn chuột phải, sau đó chọn Pane Options và All values.

Mô hình 3b. Nếu tỷ lệ sao đen bị bệnh rỗng ruột ở những địa phương khác nhau là không khác nhau, thì có cần phải xem xét ảnh hưởng của địa phương hay không? Nói cách khác, **mô hình 3b** chỉ mô tả ảnh hưởng của cấp kính. Để trả lời câu hỏi này, chúng ta xây dựng mô hình sau đây:

$$\ln(\lambda_{ij}) = \alpha + \sum_{k=1}^7 \alpha_k U_k \quad (d)$$

Ở đây $k = 1 - 7$ (nhóm cấp kính).

Thủ tục xử lý mô hình 3b: $\ln(\lambda_{ij}) = \alpha + \sum_{k=1}^7 \alpha_k U_k$
Lưu ý rằng, kết quả báo cáo $\lambda_{ij} = \text{Exp}(\alpha + \sum_{k=1}^7 \alpha_k U_k)$

(4) Gọi Special > Advanced Regression > Poisson Regression > OK

(5) **Gọi chương trình theo cách 1.** Đưa Y vào Dependent, Z vào Sample Sizes, D vào quantitative Factors, sau đó chọn OK.

(6) **Gọi chương trình theo cách 2.** Trước hết đưa Y vào Dependent, Z vào Sample Sizes, Code (vùng) vào quantitative Factors, D vào Categorical Factors > OK. Sau đó tại cửa sổ Poisson Regression bạn chọn Analysis options. Kế đến chọn Exclude, sau đó nhấp chuột hai lần vào yếu tố E để bỏ E, rồi chọn OK.

Nếu xử lý mô hình 3b theo số liệu của **bảng 7.6**, chúng ta sẽ nhận được kết quả sau đây:

$$\lambda_{ij} = \exp(-6.781 - 0.26028*D=1 + 0.00677*D=2 - 0.31396*D=3 + 0.38110*D=4 + 0.70074*D=5 + 1.23422*D=6 + 1.37806*D=7) \quad (\text{mô hình 3b})$$

Từ mô hình (3b), có thể dự đoán tỷ lệ sao đen bị rỗng ruột ở các cấp kính như sau:

Cấp D	Tỷ lệ rỗng ruột (λ_i)
1	$\lambda_1 = \exp(-6.781 - 0.26028*1)$
2	$\lambda_2 = \exp(-6.781 + 0.00677*1)$
3	$\lambda_3 = \exp(-6.781 - 0.31396*1)$
4	$\lambda_4 = \exp(-6.781 + 0.38110*1)$
5	$\lambda_5 = \exp(-6.781 + 0.70074*1)$
6	$\lambda_6 = \exp(-6.781 + 1.23422*1)$
7	$\lambda_7 = \exp(-6.781 + 1.37806*1)$
8	$\lambda_8 = \exp(-6.781)$