

羅吉斯回歸模型的應用：內孤立子的傳導影響

陳宗豪¹ 楊顯爵² 陳震遠^{3*} 陳震武⁴

¹樹德科技大學企業管理系

²高雄第一科技大學風險管理與保險系

³屏東教育大學資訊科學系(原任教：永達技術學院資訊管理系)

⁴樹德科技大學運籌管理系

摘 要

本研究旨在運用二元羅吉斯回歸模型(binary logistic regression model)及順序羅吉斯回歸模型(cumulative logistic regression model)分析海底山脊高度對孤立內波(internal solitary wave)振幅透射率之影響程度，並且試圖藉尋找穩定且精確的回歸模式，來預測海底山脊高度對孤立內波的影響程度。我們以所建構之兩種羅吉斯回歸模式進行比較其模型擬合程度及預測率高低。結果顯示：順序羅吉斯模型模型擬合資料解釋程度與預測能力檢定優於二元羅吉斯模型，由於二者差異有限，且順序模型之個別預測準確率差異很大，另由樣本型態呈現明顯低障礙物高度出現高透射率，而高障礙物則較常導致於低透射率分佈，所呈現之現象以二元分類較為穩定簡潔且符合其樣本特性。

關鍵詞：二元羅吉斯回歸模型；順序羅吉斯回歸模型；孤立內波；振幅透射率

Application of Logistic Regression Model: Propagation Effect on Internal Soliton

T. H. Chen¹, H. C. P. Yang², C. Y. Chen^{3*} and C. W. Chen⁴

¹ Department of Business Administration, Shu-Te University

² Department of Risk Management and Insurance, National Kaohsiung First University of Science and Technology

³ Department of Computer Science, National Pingtung University of Education, No. 4-18, Ming Shen Rd., Pingtung 90003, Taiwan (previously teaching in Department of Management Information System, Yung-Ta Institute of Technology and Commerce)

⁴ Department of Logistics Management, Shu-Te University

Abstract

A laboratory experiment on internal solitary wave propagation over a submarine ridge was performed. Analytical methods, binary and cumulative logistic regression model included, were appropriately employed to analyze amplitude-based transmission rate, defined as ratio of incident wave amplitude and transmission wave amplitude. We concluded that the prediction comprehension by the cumulative logistic regression model was better than the binary logistic regression model. Furthermore, the binary logistics is more stable and suitable to describe the characteristics of samples.

Keywords: binary logistic regression model; cumulative logistic regression model; internal solitary wave; amplitude-based transmission rate.

文稿收件日期 95.11.8; 文稿修正後接受日期 98.4.24;*通訊作者

Manuscript received November 8, 2006; revised April 24, 2009;*Corresponding author

一、前言

內波能使上下層流體造成巨大擾動，致使海底營養鹽傳輸，促進浮游生物的成長；內波上下層流體的異向流速，使屹立於海洋中的結構物遭受破壞；內波數十米高度的振幅，有足夠的能量擾動大陸斜坡的汙染拋棄物，造成近岸的擴散；水下潛艦可以藉著內波對聲納的散射而受到庇護。

可知內波在海洋環境、工程、生物、軍事方面都有深遠的影響。在內波現場調查研究的重點研究區域，分別以中國海域的台灣東北角及南海海域為主。南海(South China Sea, SCS)位處西太平洋西南隅，它是半封閉邊緣海域，東北部海盆鄰台灣海峽，東邊的呂宋海峽寬約300km，是與太平洋相接的海道。南海的孤立內波(Internal solitary waves)的活動相當頻繁，且相較於其他區域的內波運動規模較為大。

Ebbesmeyer and Romea(1992)推測內波發生區域是位於呂宋海峽 Batan 與 Sabtang 兩島間僅4km

寬的水道(圖1)，孤立內波於此區域形成後，再穿越南海北部向西行進至東沙島附近的陸棚區。

現場的量測工作受到極多限制，並且無法將許多控制條件逐步釐清。有鑑於此，實驗室的擬現場研究開始實施，Chen et al. (2007a, 2007b)模擬大陸斜坡地形，探討各種角度斜坡對孤立內波傳遞的影響，並且以統計回歸方法，檢測內波振幅及能量的反射影響(Chen et al., 2007)；在海底變化地形方面，Chen (2007)控制不同的海底地形，調查不同海脊地形對孤立內波演化的影響，Chen (2007)進而建立一個適當的碎波法則。

本篇文章延伸該研究重心，整合海底變化地形方面研究的量測數據，以羅吉斯回歸模型，解釋海脊高度影響內波振幅透射率程度，並說明模型之預測能力。

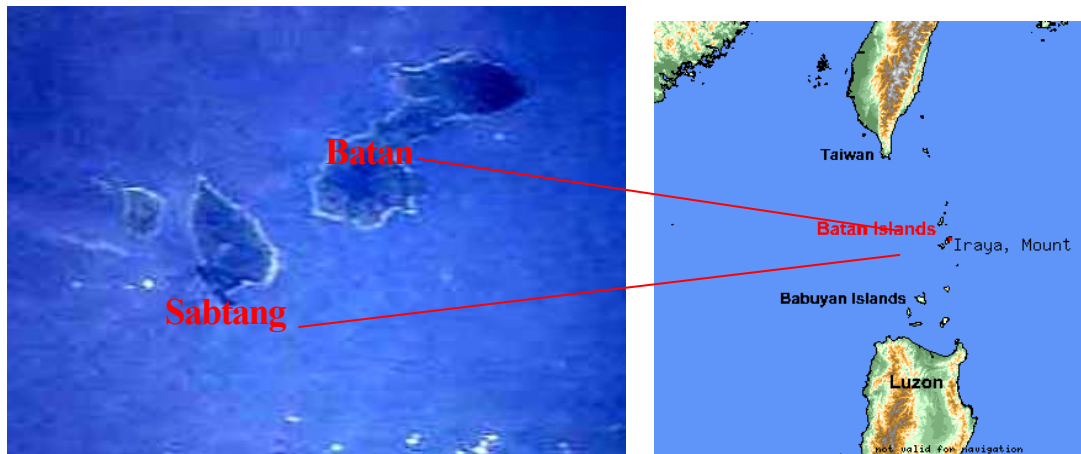


圖1 呂宋海峽 Batan 與 Sabtang 兩島衛星空照及地理位置圖

二、研究架構

本研究的數據來自實驗室的調查工作，使用的內波水槽，長度 12 公尺、寬度 0.5 公尺、高度 0.7 公尺，詳細介紹內容可以參考 Chen et al. (2007)。研究焦點著重於障礙物之高度對振幅反射率及振幅透射率影響，從圖 2 可知，當障礙物之高度在 10~15cm 時，實驗資料較集中於高內波透射率（資料較集左上方），隨著障礙物高度升高至 20cm 時，透射程度則變為分散上下兩部，然而障礙物高度再升高至 25~30cm 時，已明顯發現實驗資料較集中於低內波透射率（資料較集中右下方），該實驗資料分佈形態顯示可能區分為高低二類透射率或高中低三類透射率之情況，符合類別資料分析方式。故推測該實驗資料因變數 Y 可能呈現二元(binary)強透射程度及弱透射程度兩類，亦或是順序(ordered) 強透射程度、中透射程度及弱透射程度三類。

由於羅吉斯回歸模型常用於分析二類或三類依變數資料受到一種或多種解釋變數影響之統計模型方法，而其分析特質在於類別型態之反應變數，最常見為二元反應變數 (Walker, 2002)，例如金融危機的發生或不發生、信用風險之違約或未違約及產物保險投保之續保等，故擬將透射程度依 0.5 區格分為強透射程度及弱透射程度兩類，採用二元羅吉斯模型，抑或區分 3 段分為強透射程度、中透射程度及弱透射程度三類，採用順序羅吉斯模型，先探討實驗內波透射率的模型擬合程度，進而選擇合適模型及解釋應用，最後以較合適模型進行預測。

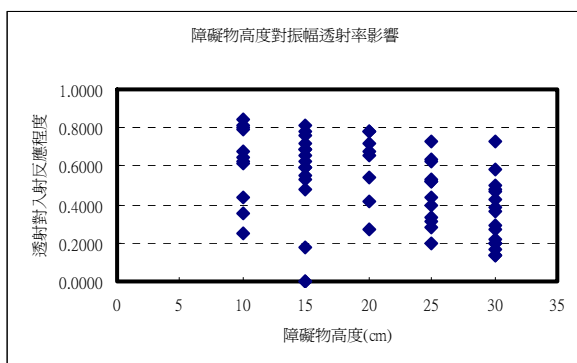


圖 2 障礙物高度對透射反應程度之影響

2.1 羅吉斯回歸背景

羅吉斯資料可能為類別資料最一般化型式，而羅吉斯回歸早在 1950 年即是應用於生物統計有用之工具 (Myers, 2002)。羅吉斯回歸分析是一種分析二分類依變數資料受到一種或多種解釋變數影響之統計模型方法，而其分析特質在於在於類別型態之反應變數，最常見為二元反應變數 (Walker, 2002)。然而典型羅吉斯回歸分析模型雖為二元反應，但已可以擴充應用到順序反應變項 (Stokes, 1995)。

當分析變數之間的關係時，若有一個屬質的反應變數，對應一個或多個解釋變數的情況，就適合採用羅吉斯回歸來進行分析 (林及陳, 2004)，羅吉斯回歸的因變數為二分變數，各自變數可以是連續變數，也可以是離散變數，還可以是虛擬變數 (Agresti, 1996; Christensen, 1997; Twisk, 2003; 王及郭, 2004)，故本研究障礙物之高度影響透射波振幅對入射波振幅反應係數之影響資料有著一個屬質且為二分變數的反應變數 (強透射程度及弱透射程度)，而解釋變數有連續、離散、及虛擬變數，符合中外學者一致適合採用羅吉斯回歸分析其資料的看法，且羅吉斯方程式機率預估公式作為預測之資料分析。

2.2 二元羅吉斯回歸模型

羅吉斯回歸模型表示式，可寫成

$$p = \frac{e^{f(x)}}{1 + e^{f(x)}}$$

$$f(x) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_k x_k \quad (1)$$

羅吉斯的關鍵來自於勝算率 ($p/1-p$) 的觀念，令 p 為成功事件的機率， $1-p$ 為失敗事件的機率。勝算率為 $p/(1-p) = e^{f(x)}$ ，兩邊取自然對數後可得

$$\ln\left(\frac{p}{1-p}\right) = f(x) \quad (2)$$

$$= \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_k x_k$$

回歸分析是描述一個應變數與多個預測變數之間的關係式，在羅吉斯回歸分析中因變數與自變數之關係呈現如圖 3 所示之 S 曲線型式，即因素 X 在開始時，成功比例 P 緩慢 (例如圖示 $X=0\sim4$) 由 0 上升，而因素在中間階段 (例如圖示 $X=5\sim6$) 時，成功比例迅速上升，而因素水準高於某個水準後 (例如圖示 $X=7\sim10$)， P 又緩慢上升到 1。

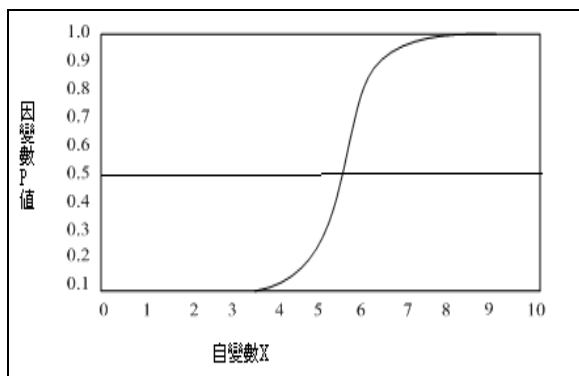


圖 3 羅吉斯回歸曲線圖

2.3 順序羅吉斯回歸模型

羅吉斯回歸模型常用於二元反應變數，然而亦可以應用於多分類反應變數。當多分類反應變數成次序關係時，應該採用順序羅吉斯回歸模型 (cumulative logistic regression model)(王及郭，2004)。順序羅吉斯模型假設依變數有 K 個反應情況，若這些反應情況為順序型態，則反應變數被稱為順序模式。通常被應用於疾病之衡量，例如，不痛、些微痛、很痛等三類(SAS Institute Inc.,1995)。

模型所面對之累積機率(The cumulative probabilities)為

$$\begin{aligned}
 F_{ij} &= p(Y \leq j) \\
 &= \sum_{m=1}^j p_{im} = p_{i1} + p_{i2} + \dots + p_{ij} \quad (3) \\
 \sum_{j=1}^k p_{ij} &= 1
 \end{aligned}$$

令 p_{ij} 為個別解釋變數對依變數的第 j 個情況之機率值

$p_{ij} = p(Y_i = j) = p(\text{individual } i \text{ responds in category } j)$
依變數情況為順序型態排列： $j=1, \dots, J$.

F_{ij} ：為小於等於第 j 個順序情況之累積機率。

$F_{ij} = p(Y_i \leq j) = p[\text{individual } i \text{ responds in category } j \text{ or worse}]$

$$\begin{aligned}
 \text{logit}[p(Y \leq j)] &= \ln \left[\frac{p(Y \leq j)}{1 - p(Y \leq j)} \right] \\
 &= \ln \left[\frac{F_{ij}}{1 - F_{ij}} \right] \quad (4)
 \end{aligned}$$

$$\begin{aligned}
 &= \ln \left[\frac{p_{i1} + p_{i2} + \dots + p_{ij}}{p_{ij+1} + \dots + p_{iJ}} \right] \\
 &= \alpha_j + \beta X_i \dots j = 1, 2, \dots, J - 1 \\
 \beta X_i &= \beta_1 x_{i1} + \beta_2 x_{i2} + \dots + \beta_k x_{ik} \quad (5)
 \end{aligned}$$

故各個累積依變數情況之羅吉優勢比如下
其中， P 為紊流動能生成率(Turbulent Kinetic Energy Production Rate)； $\mu_t = C_\mu k^2 / \varepsilon$ 為紊流黏滯係數(Turbulent Viscosity)。各係數值為 $C_\mu=0.09$ ； $C_{\varepsilon 1}=1.44$ ； $C_{\varepsilon 2}=1.44$ 。

$$\begin{aligned}
 &\ln \left[\frac{F_{i1}}{1 - F_{i1}} \right] \\
 &= \alpha_1 + \beta X_i, \dots \ln \left[\frac{F_{i2}}{1 - F_{i2}} \right] \quad (6) \\
 &= \alpha_2 + \beta X_i
 \end{aligned}$$

三、分析結果

從圖 2 所得資訊進行兩種可能模型之探討：二元羅吉斯模型與順序羅吉斯模型，探討障礙物高度對透射反應程度之影響，其所得結果分析如下：

3.1 二元羅吉斯模型

3.1.1 描述性統計

由描述性統計資料得知實驗數據共 63 組，屬於強透射程度(用 1 數值表示)有 35 組，弱透射程度(用 0 數值表示)有 28 組。

3.1.2 模型擬合資料程度

由表 1 可看出該羅吉斯回歸模型與實驗資料擬合程度 Max-rescaled R-Square 為 0.2282 即可解釋單一障礙物之高度影響射波振幅對入射波振幅反應係數之影響程度為 22.82%之變異因素，並且採用三種檢定 Likelihood Ratio、Score、Wald 檢定皆顯示 P 值皆小於 0.01，表示自變數障礙物之高度之係數 beta 非常顯著大於 0。

表1 模型的擬合統計

R-Square: 0.1704 Max-rescaled R-Square: 0.2282			
Testing Global Null Hypothesis: BETA=0			
Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	11.7706	1	0.0006
Score	11.1778	1	0.0008
Wald	10.0323	1	0.0015

3.1.3 羅吉斯回歸模型參數估計值

依照表 2 羅吉斯回歸方程式為

$$\ln\left(\frac{p}{1-p}\right) = f(x) \quad (7)$$

$$= -2.9058 + 0.1286x_1$$

表2 最大概似估計分析

Parameter	DF	Standard		Wald	
		Estimate	Error	Chi-Square	Pr > ChiSq
Intercept	1	-2.9058	0.9062	10.2815	0.0013
障礙高度	1	0.1286	0.0406	10.0323	0.0015

3.1.4 羅吉斯方程式應用

假設欲使 $p(x)$ 機率 > 0.5 ，則

$$p = \frac{e^{f(x)}}{1 + e^{f(x)}}$$

$$= \frac{e^{-2.9058 + 0.1286x}}{1 + e^{-2.9058 + 0.1286x}} > 0.5$$

$$\Rightarrow e^{-2.9058 + 0.1286x} > 1$$

$$\Rightarrow x > 22.60 \text{ cm} \quad (8)$$

即表示當障礙高度大於22.60公分時，為透射波振幅對入射波振幅反應係數為弱透射程度之機率大於50%。

3.1.5 羅吉斯回歸模型之預測能力

回歸模型之預測能力，顯示如表 3，依據模型所示反應變數之預測機率我們可得到 Somer's 數值為 0.477，另外 gamma's 數值為 0.569，兩種方法之數值越大，表示預測效果越佳。

表 3 羅吉斯回歸模型之預測能力

Association of Predicted Probabilities and Observed Responses			
Percent Concordant	65.7	Somers' D	0.477
Percent Discordant	18.1	Gamma	0.569

交叉分類表之詳細介紹請參考陳等人 (2006)指出，本研究所建構之羅吉斯回歸模型進行預測，結果在弱透射件案件正確預測率達 67.9%，另在強透射 35 件案件正確預測率達 74.3%，體正確預測率達 71.4%。

3.2 順序羅吉斯回歸模型

3.2.1 描述性統計

該模型以弱透射率為最低分類標準，由描述性統計資料得知實驗數據共 63 組，屬於強透射程度(用 3 數值表示)有 4 組，中透射程度(用 2 數值表示)有 38 組，弱透射程度(用 1 數值表示)有 21 組。

3.2.2 檢測順序回歸模型之適合與否

欲檢測順序回歸模型之適合與否，採透過檢定模型是否具備比利優勢均衡之虛無假設，假設若模型合適，則其模型可區分兩條二元羅吉斯回歸線(合併 1 及 2、另獨立 3，或合併 2 及 3、另獨立 1)，且兩條二元回歸應為平行。

然而，Peterson and Harrell(1990) 顯示這個檢測假設是非常違反傳統，他們認為可使用來檢測當 p -value 較大時，則無法拒絕假設該模型是合適有效的。由表 4 得知 p -value 為 0.1364，故很清楚沒有證據去拒絕假設該模型是合適有效的。

表4 檢測比例優勢假設Table

Score Test for the Proportional Odds Assumption		
Chi-Square	DF	Pr > ChiSq
2.2180	1	0.1364

3.2.3 模型擬合資料程度

由表 5 可看出該羅吉斯回歸模型與實驗資料擬合程度 Max-rescaled R-Square 為 0.2335 即可解釋單一障礙物之高度影響射波振幅對入射波振幅反應係數之影響程度為 23.35%之變異因素，可看出順序羅吉斯模型模型擬合資料程度優於二元羅吉斯模型之 22.82%。並且

採用三種檢定 Likelihood Ratio、Score、Wald 檢定皆顯示 P 值皆小於 0.01，表示自變數障礙物之高度之係數 beta 非常顯著大於 0。

就判斷兩個模型之改善程度差異，可以運用兩個模型之間的對數概似值成以 -2 的結果之差值(簡寫為 -2LL)，該檢定統計量稱為概似比(Likelihood Ratio)(Aldrich & Nelso, 1984; Greene, 1990; Long, 1997)。從概似比得到之模型改善數值，亦相同可看出順序羅吉斯模型模型擬合改善資料程度為 13.3149 高於二元羅吉斯模型之 11.7706。

表5 模型的擬合統計

R-Square: 0.1905 Max-rescaled R-Square: 0.2335			
Testing Global Null Hypothesis: BETA=0			
Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	13.3149	1	0.0003
Score	11.8082	1	0.0006
Wald	11.4853	1	0.0007

3.2.4 順序羅吉斯回歸模型參數估計值

依照表 6 可得之順序羅吉斯回歸模型之回歸方程式

\hat{p}_1 定義為弱透射率的機率， \hat{p}_2 定義為中透射率的機率， \hat{p}_3 定義為強透射率的機率，而兩條平行回歸線方程式為：

$$\begin{aligned} \text{logit}(\hat{p}_1) &= \ln\left(\frac{\hat{p}_1}{1-\hat{p}_1}\right) \\ &= \alpha_1 + \beta x = -3.8001 + 0.1418x \\ \text{logit}(\hat{p}_1 + \hat{p}_2) & \\ &= \ln\left[\frac{\hat{p}_1 + \hat{p}_2}{1-(\hat{p}_1 + \hat{p}_2)}\right] \\ &= \alpha_2 + \beta x = 0.1740 + 0.1418x \end{aligned} \quad (9)$$

從正的障礙物高度係數可看出隨著高度之增加透射程度機率越差。再將此兩條羅吉斯回歸線轉換為原始之 \hat{p}_1 、 \hat{p}_2 、 \hat{p}_3 機率值如下

$$\begin{aligned} \hat{p}_1 &= \frac{e^{\alpha_1 + \beta x}}{1 + e^{\alpha_1 + \beta x}} = \frac{e^{-3.8001 + 0.1418x}}{1 + e^{-3.8001 + 0.1418x}} \\ \hat{p}_2 &= \frac{e^{0.1740 + 0.1418x}}{1 + e^{0.1740 + 0.1418x}} - \hat{p}_1 \\ &= \frac{e^{0.1740 + 0.1418x}}{1 + e^{0.1740 + 0.1418x}} - \frac{e^{-3.8001 + 0.1418x}}{1 + e^{-3.8001 + 0.1418x}} \\ \hat{p}_3 &= 1 - (\hat{p}_1 + \hat{p}_2) \\ &= 1 - \frac{e^{0.1740 + 0.1418x}}{1 + e^{0.1740 + 0.1418x}} \end{aligned} \quad (10)$$

表6 最大可能度評估分析

Parameter	DF	Estimate	Standard Error	Wald	
				Chi-Square	Pr > ChiSq
Intercept1	1	-3.8001	0.9977	14.5078	0.0001
Intercept2	1	0.1740	0.8372	0.0432	0.8354
Blockage	1	0.1418	0.0418	11.4853	0.0007

3.2.5 順序羅吉斯模型之應用

假設若障礙高度為 30 公分，代入公式(10)即能得知為弱透射程度 $\hat{p}_1(30)$ 之機率 61.1%；中透射程度 $\hat{p}_2(30)$ 之機率為 37.66%；強透射程度 $\hat{p}_3(30)$ 之機率為 1.18%，如此分類所產生之機率較二元羅吉斯回歸分類較為仔細且清楚。若欲滿足使 $\hat{p}_1(x)$ 機率 > 0.5，則

$$\begin{aligned} \hat{p}_1(x) &= \frac{e^{\alpha_1 + \beta x}}{1 + e^{\alpha_1 + \beta x}} \\ &= \frac{e^{-3.8001 + 0.1418x}}{1 + e^{-3.8001 + 0.1418x}} > 0.5 \\ \Rightarrow e^{-3.8001 + 0.1418x} &> 1 \\ \Rightarrow x &> 26.80 \text{ cm} \end{aligned} \quad (11)$$

即表示當障礙高度大於 26.80 公分時，為透射波振幅對入射波振幅反應係數為弱透射程度之機率大於 50%。

3.2.6 順序羅吉斯回歸模型之預測能力

回歸模型之預測能力，顯示如表 7，依據表 7 得知模型所示反應變數之預測機率 Somer's 數值為 0.479，另外 gamma's 數值為

0.586，若數值越大，則表示預測效果越佳。可看出順序羅吉斯模型預測能力高於二元羅吉斯模型之 Somer's D 數值為 0.477，及 gamma's 數值為 0.569。

表7 順序羅吉斯回歸模型之預測能力

Percent Concordant	64.8	Somers' D	0.479
Percent Discordant	16.9	Gamma	0.586

3.2.7 順序羅吉斯回歸模型之個別機率圖

點虛線為弱透射率的機率函數：

$$\hat{p}_1(x) = \frac{e^{f(x)}}{1+e^{f(x)}} \quad (12)$$

$$= \frac{e^{\alpha_1+\beta x}}{1+e^{\alpha_1+\beta x}} = \frac{e^{-3.8001+0.1418x}}{1+e^{-3.8001+0.1418x}}$$

虛線為中透射率的機率函數：

$$\hat{p}_2 = \frac{e^{0.1740+0.1418x}}{1+e^{0.1740+0.1418x}} - \hat{p}_1 \quad (13)$$

$$= \frac{e^{0.1740+0.1418x}}{1+e^{0.1740+0.1418x}} - \frac{e^{-3.8001+0.1418x}}{1+e^{-3.8001+0.1418x}}$$

實線為強透射率的機率函數：

$$\hat{p}_3 = 1 - (\hat{p}_1 + \hat{p}_2) = 1 - \frac{e^{0.1740+0.1418x}}{1+e^{0.1740+0.1418x}} \quad (14)$$

由圖 4 可看出反應變數三個情況各有其機率預測函數，且各機率預測函數及圖形斜率皆不同，隨著障礙物高度之增加使得強透射率的機率 \hat{p}_3 (實線) 逐漸下降，而弱透射率的機率 \hat{p}_1 (點虛線) 則相對上升，然而當障礙物高度之增加至較高時，可看見中透射率的機率 \hat{p}_2 (虛線) 已低於弱透射率的機率 \hat{p}_1 (點虛線) 了。

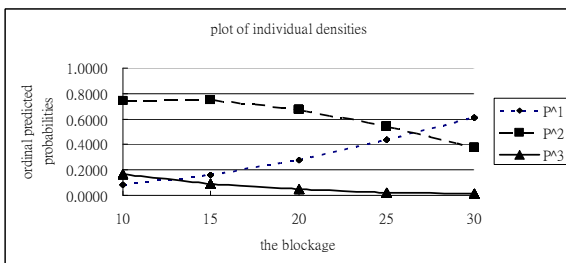


圖 4 順序羅吉斯回歸模型之個別機率圖

藉由表 8 交叉分類表指出，本研究所建構之順序羅吉斯回歸模型進行預測，結果在弱透射 21 件案件正確預測率達 28.6%，中透射件 38 案件正確預測率達 84.2%，另在強透射 4

件案件正確預測率達 0%，整體正確預測率達 60.3%。可見結果顯示：順序羅吉斯模型模型擬合資料解釋程度與預測能力檢定雖然數值皆高於二元羅吉斯模型，由於二者差異有限，且順序模型之個別預測準確率差異很大，另由樣本型態呈現明顯低障礙物高度出現高透射率，而高障礙物則較常導致於低透射率分佈，所呈現之現象以二元分類較為穩定簡潔且符合其樣本特性。

表8 順序羅吉斯回歸模型之預測準確度

種類	實際	正確	錯誤	預測正確率
弱透射率的機率預測	21	6	15	28.6%
中透射率的機率預測	38	32	6	84.2%
強透射率的機率預測	4	0	4	0.0%
總預測率	63	38	25	60.3%

3.2.8 二元羅吉斯回歸模型與順序羅吉斯回歸模型之預測機率函數圖

由圖 5 可比較二元羅吉斯回歸模型與順序羅吉斯回歸模型之機率預測函數圖形，可看出二元羅吉斯回歸模型之弱透射率的機率 p (虛線) 與順序羅吉斯回歸模型之弱透射率的機率 \hat{p}_1 (實線) 平行，並且在障礙物高度之任何一點上之弱透射率的機率皆高於順序羅吉斯回歸模型，另外順序羅吉斯回歸模型之中加弱透射率 ($\hat{p}_1 + \hat{p}_2$) 的機率線 (點虛線) 必然高於單獨之弱透射率的機率線。

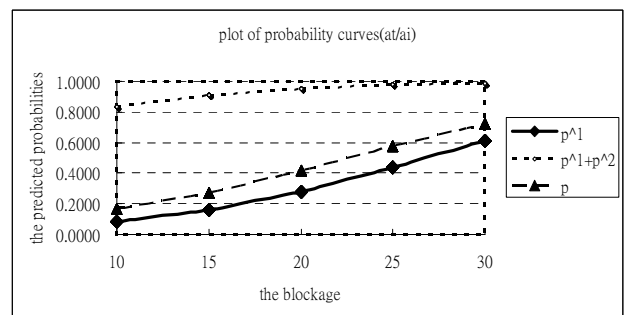


圖 5 二元羅吉斯回歸模型與順序羅吉斯回歸模型之預測機率函數比較

點虛線為順序羅吉斯回歸模型之弱透射率與中透射率相加的機率函數：

$$\begin{aligned}
 & (\hat{p}_1 + \hat{p}_2) \\
 &= \frac{e^{\alpha_2 + \beta x}}{1 + e^{\alpha_2 + \beta x}} \quad (\text{順序羅吉斯反應函數}) \\
 &= \frac{e^{0.1740 + 0.1418x}}{1 + e^{0.1740 + 0.1418x}}
 \end{aligned} \tag{15}$$

虛線為二元羅吉斯回歸模型之弱透射率的機率函數：

$$\begin{aligned}
 p(x) &= \frac{e^{f(x)}}{1 + e^{f(x)}} \quad (\text{二元羅吉斯反應函數}) \\
 &= \frac{e^{-2.9058 + 0.1286x}}{1 + e^{-2.9058 + 0.1286x}}
 \end{aligned} \tag{16}$$

實線為順序羅吉斯回歸模型之弱透射率的機率函數：

$$\begin{aligned}
 \hat{p}_1 &= \frac{e^{\alpha_1 + \beta x}}{1 + e^{\alpha_1 + \beta x}} \quad (\text{順序羅吉斯反應函數}) \\
 &= \frac{e^{-3.8001 + 0.1418x}}{1 + e^{-3.8001 + 0.1418x}}
 \end{aligned} \tag{17}$$

四、結論與建議

從本次實驗數據，控制障礙物高度對內波透射率影響，經羅吉斯回歸模型之分析得知：

- (1) 檢定障礙物高度係數 beta 非常顯著大於 0，故對透射程度之強弱有顯著之影響程度。
- (2) 運用所得出之二元羅吉斯回歸模型之回歸方程式，詳細求得當障礙高度須大於 22.60 公分時，透射波振幅對入射波振幅反應係數為弱透射程度之機率大於 50%，若以順序羅吉斯回歸模型檢驗障礙物高度對振幅透射率，當以強中弱三種形式分類顯著特性，回歸模型推測山脊高度須大於 26.80 公分，弱透射程度之機率始大於 50%。
- (3) 以所建構之兩種羅吉斯回歸模式進行比較其模型擬合程度及預測率高低，結果顯示：順序羅吉斯回歸模型之模型擬合資料解釋程度 Max-rescaled R-Square 為 23.35% 高於二元羅吉斯回歸模型之 22.82%。
- (4) 以本研究所建構之二元羅吉斯回歸模型進行預測，結果在實際弱透射案件正確預測率達 67.9%，強透射案件正確預測率達

74.3%，整體正確預測率達 71.4%。而順序羅吉斯模型反應變數之預測機率 Somer's 數值為 0.479，另外 gamma's 數值為 0.586，若數值越大，則表示預測效果越佳。可看出預測能力高於二元羅吉斯模型之 Somer's 數值為 0.477，及 gamma's 數值為 0.569。

- (5) 本研究所建構之順序羅吉斯回歸模型進行預測，結果在弱透射 21 件案件正確預測率達 28.6%，中透射件 38 案件正確預測率達 84.2%，另在強透射 4 件案件正確預測率達 0%，整體正確預測率達 60.3%。可見結果顯示：順序羅吉斯模型模型擬合資料解釋程度與預測能力檢定雖然數值皆高於二元羅吉斯模型，由於二者差異有限，且順序模型之個別預測準確率差異很大，另由樣本型態呈現明顯低障礙物高度出現高透射率，而高障礙物則較常導致於低透射率分佈，所呈現之現象以二元分類較為穩定簡潔且符合其樣本特性。

內波現場的量測，至今仍採取固定位置測量方式，測量的數據僅能提供單點的時間連續鹽度剖面或壓力分布，無法突破全斷面測量的限制，再者，現場的工作受到氣候及環境因素，測量資料往往無法提供合理的解釋，本研究以實驗室試驗方式，控制許多重要的物理因子，探討單一物裡條件對運動機制的影響，從實驗資料的擷取、整理、分析，皆以系統化的方式進行，文中藉由羅吉斯回歸模型，找尋適當的預測方程式，研究成果將可提供海洋科學研究領域，在內波動力研究的啟示。

誌謝

本文承行政院國科會提供經費補助(計畫編號 NSC 96-2628-E-366-004MY2, NSC 96-2628-E-132-001MY2, NSC 97-2410-H-004-107)，同時感謝三位審察者的寶貴建議，讓本文能夠修改更加完善，特此表達深忱謝意。

參考文獻

- [1] 陳震遠、楊顯爵、陳震武、陳宗豪，羅吉斯回歸模型應用於系統分析與設計，科研月刊，總第 24 期，頁 13-15，2006 年 12 月。
- [2] 王濟川，郭志剛，羅吉斯迴歸模型-方法及應用，台北，五南圖書出版股份有限公司

- 司，2004。
- [3] Agresti, A. An introduction to categorical data analysis. John Wiley & Sons, Inc., 1996.
- [4] Aldrich, J and Nelson, FD. Linear probability, logit, and probit models. Newbury Park, CA: Sage Publications, 1984.
- [5] Chen, C.W., Chen, C.Y., Yang, H.C.P., and Chen, T.H., "Analysis of experimental data on internal waves with statistical method," Engineering Computations. Vol. 24(2), pp. 116-150, 2007.
- [6] Chen, C.Y., Hsu, J.R.C., Chen, H.H., Kuo, C.F., and Cheng, M.H., "Laboratory observations on internal solitary wave evolution on steep uniform slopes," Ocean Eng. Vol. 34(1), pp. 157-170, 2007a.
- [7] Chen, C.Y., Hsu, J.R.C., Cheng, M.H., Chen, H.H., and Kuo, C.F., "An investigation on internal solitary waves in a two-layer fluid: propagation and reflection from steep slopes," Ocean Eng. Vol. 34(1), pp. 171-184, 2007b.
- [8] Chen, C.Y., Hsu, J.R.C., Kuo, C.F., Chen, H.H., and Cheng, M.H., "Laboratory observations on internal solitary wave evolution over a submarine ridge," China Ocean Eng. Vol. 20(1), pp. 61-72, 2006.
- [9] Chen, C.Y., "An experimental study of stratified mixing caused by internal solitary waves in a two-layered fluid system over variable seabed topography," Ocean Eng. Vol. 34(14-15), pp. 1995-2008, 2007.
- [10] Chesser, D.L., "Prediction loan noncompliance," Journal of Commercial Bank Lending. Vol. 56(8), pp. 28-38, 1974.
- [11] Christensen, R. Log-Linear models and logistic regression. Springer texts in statistics. 1997.
- [12] Ebbesmeyer, C.C. and Romea, R.D. Final design parameters for solitons at selected locations in South China Sea final and supplementary reports prepared for Amoco Production Company, 209, 1992.
- [13] Greene, W. H., Econometric analysis. New York: Macmillan, 1990.
- [14] Long, JS. Regression models for categorical and limited dependent variables. Thousand Oaks, California: Sage Publications, 1997.
- [15] Myers, R. H., Montgomery, D. C., Vining, G. G., Generalized linear models with applications in engineering and sciences. John Wiley & Sons, Inc., 2002.
- [16] Peterson, B. and Harrell, F., "Partial proportional odds models for ordinal response variables," Applied Statistics. Vol. 39, pp. 205-217, 1997.
- [17] SAS Institute., Logistic regression examples using the SAS system, version 6 ,first edition, Cary, NC:SAS Institute Inc., 1995.
- [18] Stokes, M. E., Davis, C. S., Koch, G. G.,Categorical Data analysis using the SAS system, Cary, NC:SAS institute Inc.,1995.
- [19] Twisk, JWR. Applied longitudinal data analysis for epidemiology: a practical guide. the Press Syndicate of the University of Cambridge, 2003.
- [20] Walker, Glenn a. Common statistical methods for clinical research with SAS examples, second edition. Cary, NC:SAS Institute Inc., 2002.

陳宗豪等
羅吉斯回歸模型的應用：內孤立子的傳遞影響