社會科學統計方法（下）

九十五學年度第二學期　　　　　　　　授課人：黃　紀
上課時間：週二下午 1：10 至 3：00＋實習　　研究室：綜合院館 271668 or 270507
導生時間：週四、五下午 4：10 至 5：00　　電　話：分機 51668 or 2938-4514
e-mail: chihuang@nccu.edu.tw

黃紀教授政治地緣資料庫 網頁：http://tegis.nccu.edu.tw/

一、 課程宗旨

　　本課程延續「社會科學統計方法（上）」，繼續介紹社會科學研究中常用的統計方法，但重心逐漸移到**當依變數爲類別變數**（即 **categorical responses**, 包括二分或多分之 nominal 及 ordinal variables）**及非負值整數之 count variable** 時，所適用之推廣線型模型（generalized linear models, 簡稱 GLM），包括勝算對數模型 logit（含 binary and multinomial logit, conditional logit, nested logit, ordinal logit, sequential logit 等）,機率單元模型 probit, 與 卜瓦松模型 Poisson models 等。這些模型與其他常見之統計模型的適用時機，詳見下表。

常見之統計模型

		自變數	
		全是類別變數	**至少有一個整數或連續變數**
依變數	**二分**	2×c×… 行列表分析；機率單元(probit)模型、勝算對數(logit)模型	機率單元模型、成長曲線(logistic)迴歸
	無序多分	r×c×… 行列表分析；多項(multinomial)之機率單元模型、勝算對數模型	多項之機率單元模型、成長曲線迴歸
	有序多分	r×c×… 行列表分析；有序多分之機率單元模型、勝算對數模型	有序多分之機率單元模型、成長曲線迴歸
	整數	*對數線型（loglinear）模型；卜瓦松（Poisson）迴歸及其延伸	卜瓦松迴歸及其延伸
	連續	變異數分析(ANOVA)；線型或非線型迴歸	共變數分析(ANCOVA)；線型或非線型迴歸

　　*註：嚴格說來，對數線型模型並不區分自變數與依變數，而是以行列表細格內之頻數分佈爲解釋對象，並以組成行列表的所有變數及其互動，作爲解釋變數。

本表引自 黃紀:〈質變數之計量分析〉，載謝復生、盛杏湲主編:《政治學的範圍與方法》，台北：五南圖書出版公司，民國 89 年，頁 387-411。

　　Quantitative analysis is *a highly interactive and iterative process*: you have to constantly shift back and forth among mathematical statistical theory, substantive political theories and knowledge, and your hard-earned empirical data.　This process can be tedious and sometimes downright frustrating.　So you should reserve ample time for yourself to complete data checking, transformation and analysis.

The computer statistical package used in this class will be **STATA 9. SAS 9** (Statistical Analysis System) for Windows will be referred to when necessary. Both of them are powerful and popular programs and thus are highly recommended for analyzing categorical response data. Modules of Regression and Advanced Models of **SPSS 12.0** or higher can also handle most of the models covered in this class. Freeware such as ₤EM is available but has a rather steep learning curve to climb.

二、必備課本及工具

*Agresti, Alan. 1996. *An Introduction to Categorical Data Analysis*. New York: Wiley. (hereafter, Agresti)

Agresti, Alan, and Barbara Finlay. 1997. *Statistical Methods for the Social Sciences*, 3rd ed. NJ: Prentice Hall. (Hereafter, Agresti & Finlay)（雙葉書廊代理）

Hagle, Timothy M. 1996. *Basic Math for Social Scientists: Problems and Solutions*. Thousand Oaks, CA: Sage. (QASS# 109)

*Hamilton, Lawrence C. 2004. *Statistics with Stata: Updated for Version 8.* Pacific Grove, CA: Duxbury Press. (hereafter, Hamilton)

*Long, J. Scott. 1997. *Regression Models for Categorical and Limited Dependent Variables.* Thousand Oaks, CA: Sage. (hereafter, Long) （中譯本：鄭旭智 等譯，2002,《類別與受限依變項的迴歸統計模式》,台北：弘智。）

*Long, J. Scott, and Jeremy Freese. 2006. *Regression Models for Categorical Dependent Variables Using Stata,* 2nd Edition. College Station, TX: Stata Press. (hereafter, Long & Freese)

方世榮，2005,《統計學導論》五版，台北：華泰。（以下簡稱 方世榮）

王保進，2006,《中文視窗版 SPSS 與行為科學研究》台北：心理出版社。（以下簡稱 王保進）

必備工具：計算機(只須具備四則運算+－×÷ 、開根號√、乘冪x^y、對數log、簡單記憶M之功能)。

參考課本：

*Agresti, Alan. 2002. *Categorical Data Analysis,* 2nd edition. New York: Wiley. （雙葉代理）

*Eliason, Scott R. 1993. *Maximum Likelihood Estimation: Logic and Practice.* Newbury Park, CA: Sage. (QASS#96) (hereafter, Eliason) （雙葉代理）

*Hosmer, David W., and Stanley Lemeshow. 2000. *Applied Logistic Regression,* 2nd edition. New York: Wiley. (hereafter, Hosmer & Lemeshow) （華泰代理）

*King, Gary. 1998. *Unifying Political Methodology: The Likelihood Theory of Statistical Inference*. Ann Arbor: Michigan University Press.

*McCullagh, P., and J.A. Nelder. 1989. *Generalized Linear Models*, 2nd edition. London: Chapman and Hall.

*Rabe-Hesketh, Sophia, and Brian Everitt. 2006. *A Handbook of Statistical Analyses Using Stata*, 4th edition. Boca Raton: Chapman & Hall/CRC.

SPSS Inc. 2003. *SPSS Regression Models 12.0*. Chicago: SPSS Inc.

Stata Press. 2005. *Stata Reference Manuals: Release 9.0*. College Station, TX: Stata Press.

三、指定作業

(1) 作業、練習

Besides the required 社會科學統計方法實習，You should expect frequent in-class exercises and homework. If you find any difficulty in solving the problems, please raise your questions in the class. Otherwise, I shall simply assume that you have done your homework.

(2) 期末報告

You are required to write a complete research paper **by carrying out the proposal that you turned in last semester in my class of** 社會科學統計方法（上）. Besides reviewing the literature, you should formulate your hypotheses, build a data set (or selecting an appropriate second-hand data set) and then test hypotheses with the methods that you learn in the class, and interpret your empirical findings.

Attach your original term paper (with my comments) in 社會科學統計方法（上） as well as SPSS/Stata computer printout as appendices to your paper. **Any paper without these appendices will receive a score of zero.** Needless to say, **you must be the main person who compiles the data, runs the SPSS program, does the analysis and interprets the results. Failure to observe this rule is considered cheating.**

❖❖ Papers must be written in a scholarly style with footnotes (or endnotes) and references **in author-date system.** They must be typed in a word processor and printed with a laserjet printer. (※Correct style in author-date system will account for 20% of your term paper grade.)❖❖

(3) 期中考試

An in-class close-book midterm exam is scheduled on May 2, 2007.

※注意：考試時，只可攜帶不超出＋－×÷√ xʸ **log M** 功能之計算機。

(4) 出席、回答抽問

　　　Due to the formal nature of statistical techniques, three quarters of the class time will be used for lectures and the rest for discussions.　In order to know if you actually complete all required readings and reflect upon them before coming to the class, however, I shall also ask you questions from the reading assignments during my lectures.

四、計分方式

- 作業、練習、小考　　10%
- 期中考試　　　　　　40%
- 期末報告　　　　　　45%
- 出席、回答抽問　　　5%

　　　I reserve the right to administer pop quizzes during lectures.　Each quiz will be counted as an exercise.　Unexcused missed exams, exercises, and pop quizzes will receive a score of zero. Also, **be aware that the *minimum* punishment for either cheating or plagiarism is a zero for that assignment.**

五、課程內容及指定閱讀

　　　This is a tentative schedule for the semester.　Unless we fall behind or move through some sections more quickly than expected, your reading assignments will be as indicated here. The only optional part is the reading marked as "for further references" (「參考」).　If adjustments are required along the way, I will announce them in class.

Week 1　　　Introduction to Statistical Software Stata

> Hamilton, pp. 1-9, Chapters 2 and 3;
> Long & Freese, Chapters 1, 2, and 3.

CROSS-TABULAR ANALYSIS

Week 2　Association between Two Categorical Variables: Two-Way Contingency Tables

> Agresti & Finlay, Chapter 8;
> Agresti, Chapter 2;
> Hamilton, pp. 130-136;
> 王保進，第七章;

方世榮，第 11 章。

實例—黃紀，2001，〈一致與分裂投票：方法論之探討〉，《人文及社會科學集刊》，13(5): 541-574。（讀前兩節）

Week 3 **Multi-Way Contingency Tables**

Agresti, Chapter 3;

Agresti & Finlay, Chapter 10;

Hamilton, pp. 130-136;

實例—吳重禮、湯京平、黃紀，1999，〈我國「政治功效意識」測量之初探〉，《選舉研究》，6(2): 23-44.

黃紀，2005，〈投票穩定與變遷之分析方法：定群類別資料之馬可夫鍊模型〉，《選舉研究》，12(1): 117-150。（讀前四節）

LINEAR REGRESSION FOR CONTINUOUS DEPENDENT VARIABLES

Week 4 **Correlation and Simple Linear Regression**

Agresti & Finlay, Chapter 9;

王保進，第六、九章；

方世榮，第 13 章之 13.1-13.7（pp. 518-560）。

Week 5 **Multiple Linear Regression**

Agresti & Finlay, Chapters 11 (skip sections 11.7-11.8) and 14;

Hamilton, Chapter 6;

方世榮，第 13 章之 13.9（pp. 565-586）；

複習：王保進，第九章；

4/3 迎春活動週

Week 7 **Multiple Linear Regression with Categorical Independent Variables**

Agresti & Finlay, Chapter 13.

實例—Huang, Chi.（黃紀）1992. "Leadership Change and Government Size in East Asian Authoritarian Regimes." Chapter 7 in Cal Clark and Steve Chan eds. *The Evolving Pacific Basin in the Global Political Economy.* Boulder, CO: Lynne Rienner.

A DIGRESSION TO BASIC MATH

Week 8   Introduction to Matrix Algebra and Basic Calculus

Hagle, 1996.

Long, Chapter2.

Week 9   Discrete Probability Distributions and Maximum Likelihood Principle

Agresti, Chapter 1;
Eliason, pp. 1-62.

5/2   **期 中 考 試**

GENERALIZED LINEAR MODELS (GLM)

Week 11   GLM: A Generalization of Linear Regression

Agresti, Chapter 4;
黃紀，2000：〈質變數之計量分析〉，載 謝復生、盛杏湲主編：《政治學的範圍
　　與方法》， 台北：五南圖書出版公司，頁 387-411。
Hamilton, pp. 261-265;
實例一：Huang, Chi（黃紀）, Woosang Kim and Samuel Wu（吳秀光）. 1992.
　　"Conflicts Across the Taiwan Strait: 1951-1978." *Issues and Studies* 28(6):
　　35-58.
實例二：Shields, Todd, and Chi Huang（黃紀）. 1997. "Executive Vetoes: Testing
　　Presidency- Versus President-Centered Perspectives of Presidential
　　Behavior." *American Politics Quarterly* 25(4): 431-457.
參考：McCullagh & Nelder, 1989.

Week 12   Logistic Regression for Binary Dependent Variable (I): Introduction

Agresti, Chapter 5 (skip 5.6 and 5.7);
Long, pp. 34-61;
Long & Freese, Chapter 4;
Hamilton, pp. 262-278;
　　黃紀：〈附錄：Probit, Logit, Logistic Regression 等詞之起源、意義與翻譯〉；
實例：Huth, Paul, Christopher Gelpi and D. Scott Bennett. 1993. "The Escalation of
　　Great Power Militarized Disputes: Testing Rational Deterrence Theory and
　　Structural Realism." *American Political Science Review* 87: 609-623.

Week 13 Logistic Regression for Binary Dependent Variable (II): Tests & Interpretation

Long, pp. 61-83, Chapter 4;
Huang, Chi（黃紀）, and Todd Shields. 1994. "Modeling and Interpreting Interactions in Logit Analysis." 《選舉研究》，1(1): 171-196;
Huang, Chi（黃紀）, and Todd Shields. 2000. "Interpretation of Interaction Effects in Logit and Probit Analysis: Reconsidering the Relationship between Registration Laws, Education, and Voter Turnout." *American Politics Quarterly* 28(1): 80-95.

Week 14 Ordinal Logit Models

Agresti, Sections 8.2 and 8.3 of Chapter 8;
Long, Chapter 5;
Long & Freese, Chapter 5
Hamilton, pp. 278-280;
實例：Huang, Chi.(黃紀) 2004. "Explaining Referendum Voting Choices in Taiwan." *Issues & Studies* 40(3/4): 316-333.（a sequential logit model for ordinal response data）

Week 15 Multinomial Logit (MNL) Model and Conditional Multinomial Logit (CML) Model

Agresti, Section 8.1 of Chapter 8;
Long, pp. 148-186;
Long & Freese, Chapter 6 & pp.293-313;
Hamilton, pp. 230-236;
實例：徐火炎，1995：〈『李登輝情結』與省市長選舉的投票行為：一項政治心理學的分析〉，《選舉研究》，2(2): 1-36。

Week 16 Nested Logit (NL) Model and Multinomial Probit (MNP) Model

Maddala, 1983, pp. 61-64 and 67-70;
Long & Freese, pp. 313-328.

實例：Alvarez, R. Michael, and Jonathan Nagler. 1998. "When Politics and Models Collide: Estimating Models fo Multiparty Elections." *American Journal of Political Science* 42(1): 55-96.
參考：Hausman, Jerry A. and David A. Wise. 1978. "A Conditional Probit Model for Qualitative Choice: Discrete Decisions Recognizing Interdependence and Heterogeneous Preferences." *Econometrica* 46(2): 403-426

6/19 端午節

6/26 期末課堂報告

7/2 ＊＊繳期末報告＊＊