社會科學統計方法（下）

九十九學年度第二學期　　授課人：黃 紀
上課時間: 週二下午 1：10 至 4：00＋實習　　研究室：綜合院館 271668 or 270507
導生時間: 週五下午 3：10 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/
台灣選舉與民主化調查 TEDS 網頁：http://www.tedsnet.org/

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**, 包括二分類或多分類之 nominal 及 ordinal variables）**及非負值整數之計次變數 count variable** 時，所適用之推廣線型模型（generalized linear models, 簡稱 GLM），包括**勝算之對數模型 logit**（含 binary and multinomial logit, conditional logit, nested logit, ordinal logit, sequential logit 等）,**機率單元模型 probit**, 與 **泊松模型 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

依變數 \ 自變數	**全是類別變數**	**至少有一個整數或連續變數**
二分類	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
無序多分類	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
有序多分類	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； 泊松（Poisson）迴歸及其延伸	泊松迴歸及其延伸
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聯合次數分佈爲解釋對象，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 11.** It is a powerful and popular program and thus is highly recommended for analyzing categorical response data. Modules of Regression and Advanced Models of **SPSS 18.0** (**PASW Statistics 18.0**) or higher can also handle most of the models covered in this class. Freeware such as **ℓEM** or **R** is available but has a rather steep learning curve to climb.

二、必備課本及工具（打*者為本學期新增課本）

*Agresti, Alan. 2007. *An Introduction to Categorical Data Analysis*, 2nd edition. New York: Wiley. (hereafter, Agresti) （華泰代理）

Agresti, Alan, and Barbara Finlay. 2009. *Statistical Methods for the Social Sciences*, 4th ed. NJ: Prentice Hall. (Hereafter, Agresti & Finlay)（雙葉書廊代理）

Fox, John. 2009. *A Mathematical Primer for Social Scientists*. Thousand Oaks, CA: Sage. (QASS# 159) (Hereafter, Fox)

*Hamilton, Lawrence C. 2009. *Statistics with Stata: Updated for Version 10.* 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、張文賢，2010，《統計學導論》第六版，台北：華泰。(以下簡稱 方世榮)

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)。

必玩網頁：免費又好玩！

1. Internet Glossary of Statistical Term:

http://www.animatedsoftware.com./elearning/Statistics%20Explained/glossary/se_glossary.html

2. Probability Calculator: http://www.stat.berkeley.edu/~stark/Java/Html/ProbCalc.htm

3. Sampling Distribution: http://onlinestatbook.com/stat_sim/sampling_dist/index.html

參考課本：

Agresti, Alan. 2002. *Categorical Data Analysis*, 2nd edition. Hoboken: Wiley.

Box-Steffensmeier, Janet M., Henry E. Brady, and David Collier. eds. 2008. *Oxford Handbook of Political Methodology*. Oxford: Oxford University Press.

*Cameron, A. Colin, and Pravin K. Trivedi. 2010. *Microeconometrics Using Stata,* Revised Edition. College Station: Stata Press.

*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.

Menard, Scott. 2010. *Logistic Regression: From Introductory to Advanced Concepts and Applications.* Thousand Oaks: Sage.

Pollock, Philip H., III. 2010. *A Stata Companion to Political Analysis,* 2nd edition. Washington, DC: CQ Press.

StataCorp. 2009. *Stata Reference Manuals: Release 11*. 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. **報告之體例，一律依照《選舉研究》第 14 卷第 1 期開始採用之〈論文撰稿用例〉，見《選舉研究》，14(1): 155-165。❖❖**

(3) 期中、期末考試

Midterm and final exams are scheduled on April 19 and June 21, 2011.

※注意：考試時，只可攜帶不超出+ − × ÷ √ x^y 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 15%
- 期中考試 20%
- 期末考試 25%
- 期末報告 35%
- 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 Inference from Two Samples & Intro. to Stata

Agresti & Finlay, Chapter 7 (skip sections 7.5-7.7); 方世榮，第 10 章；
Hamilton, Chapters 1, 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, Sections 2.1-2.6 of Chapter 2;
Hamilton, pp. 130-133;
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, Sections 2.7 of Chapter 2;
Agresti & Finlay, Chapter 10;
Hamilton, pp. 133-136;
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 and Section 12.3 (pp. 379-381);
方世榮，第 13 章之 13.1-13.7（pp. 497-538）。

Week 5 Multiple Linear Regression

Agresti & Finlay, Chapter 11, Section 12.5 (pp. 386-390), and Chapter 14;
Hamilton, Chapter 6;
方世榮，第 13 章之 13.9（pp. 543-559）。

Week 6 Multiple Linear Regression with Categorical Independent Variables

Agresti & Finlay, Chapter 13.

A DIGRESSION TO BASIC MATH

Week 7 **4/5 清明節（放假）**

Week 8 **Matrix Algebra, Discrete Probability Distributions, and Maximum Likelihood Principle**

Fox, Chapters 1,2, & 3;
Agresti, Chapter 1;
Long, Chapter 2;
參考：Eliason, pp. 1-62.

4/19 ****期中考試****

GENERALIZED LINEAR MODELS (GLM)

Week 10 **GLM: A Generalization of Linear Regression**

Agresti, Chapter 3;
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。
實例：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;
Glasgow, Garrett, and R. Michael Alvarez. 2008. "Discrete Choice Methods." Chapter 22 in *Oxford Handbook of Political Methodology*.

Week 11 **Logistic Regression for Binary Dependent Variable (I): Introduction**

Agresti, Chapter 4;
Long, pp. 34-61;
Long & Freese, Chapter 4;
Hamilton, pp. 262-278;
黃紀，2000，〈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 12 **Logistic Regression for Binary Dependent Variable (II): Tests & Interpretation**

Agresti, Chapter 5;
Long, pp. 61-83, Chapter 4;
Hamilton, pp. 313-317;
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 13 Ordinal Logit Models

Agresti, Sections 6.2 to 6.4 of Chapter 6;
Long, Chapter 5;
Long & Freese, Chapter 5
Hamilton, pp. 278-280;
實例：Huang, Chi.(黃紀) 2008. “Referendum and Democracy: the Experience of Taiwan.” Pp. 121-134 in *Democratization in Taiwan: Challenges in Transformation*, eds. Philip Paolino and Jim Meernik. Hants, U.K.: Ashgate.（a sequential logit model for ordinal response data）

Week 14 Multinomial Logit (MNL) Model and Conditional Multinomial Logit (CML) Model

Agresti, Section 6.1 of Chapter 6;
Long, pp. 148-186;
Long & Freese, Chapter 6 & pp.293-313;
Hamilton, pp. 280-287;
實例：黃紀、王鼎銘、郭銘峰，2008，〈「混合選制」下選民之一致與分裂投票：1996 年日本眾議員選舉自民黨選票之分析〉,《選舉研究》, 15(2): 1-35。

Week 15 Nested Logit (NL) Model and Multinomial Probit (MNP) Model

Cameron & Trivedi. 2009. *Microeconometrics Using Stata*, pp. 496-507;
Long & Freese, pp. 313-338;
實例：Alvarez, R. Michael, and Jonathan Nagler. 1998. “When Politics and Models Collide: Estimating Models of Multiparty Elections.” *American Journal of Political Science* 42(1): 55-96.
黃紀、王德育，2009，〈2008 年立委選舉對總統選舉的影響：鐘擺效應？西瓜效應？〉，載於 陳陸輝、游清鑫、黃紀 編《2008 年總統選舉：論二次政黨輪替之關鍵選舉》，台北：五南書局。

Week 16 Limited Continuous Dependent Variables: Explicit and Incidental Selection Problems

Long, Chapter 7.
參考：Heckman, James J. 1976. “The Common Structure of Statistical Models of Truncation, Sample Selection and Limited Dependent Variables and a

Simple Estimator for Such Models." *Annals of Economic and Social Measurement* 5(4): 475-492.

Week 17 **期末課堂報告**

6/21 ****期末考試****

6/27 ****繳期末報告****

附錄：希臘字母表 Greek Alphabet

LOWER CASE 小寫	CAPITAL 大寫	NAME 唸法
α	Α	alpha
β	Β	beta
γ	Γ	gamma
δ	Δ	delta
ε	Ε	epsilon
ζ	Ζ	zeta
η	Η	eta
θ	Θ	theta
ι	Ι	iota
κ	Κ	kappa
λ	Λ	lambda
μ	Μ	mu
ν	Ν	nu
ξ	Ξ	xi
o	Ο	omicron
π	Π	pi
ρ	Ρ	rho
σ	Σ	sigma
τ	Τ	tau
υ	Υ	upsilon
φ	Φ	phi
χ	Χ	chi
ϕ	Ψ	psi
ω	Ω	omega