统计分析简介与数量方法的基础

01
CHAPTER

1-1 统计分析简介

在我们的生活中，存在着许许多多待解决的问题，有些问题是可以使用统计分析的方法来解决的。统计分析方法是以数学为基础，有严紧的逻辑和标准，需要遵循特定的规范，从确立目的、发展和选用题项、提出假设、进行抽样、数据的收集，分析和解释数据，得出结论，以提供数据给决策者作正确的决定。

在一般的统计课程中都会提到统计分析可分为描述性统计和统计推断。

描述性统计

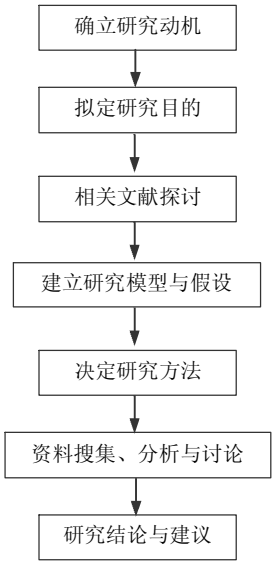
描述性统计是将研究中所得的数据加以整理、归类、简化或绘制成图和表，用来描述和归纳数据的特征（例如：人口变数统计），是最基本的统计方法。描述性统计主要提供资料的集中趋势、离散程度和相关强度，例如：平均值（$\overline{X}$）、标准差（σ）、相关系数（r）等。

统计推断

统计推断是用概率形式来决断资料之间是否存在某种关系及用样本统计值来推测总体的统计方法。统计推论包括假设检验和参数估计，最常用的方法有 Z 检验、t 检验、卡方检验等。

描述性统计和统计推断二者彼此息息相关，相辅相成，描述性统计是统计推断的基础，统计推断是描述性统计的进一步运用。一般的研究议题中，是采用描述性统计还是统计推断，需要依研究的目的而定，如果研究的目的是需要描述性统计的数据，则需使用描述性统计；若需要以样本信息来推断总体的情形，则需用统计推断。在社会科学的研究中，常常需要严谨的处理方式，描述性统计和统计推断经常是一起使用，来解决复杂的问题。

在社会科学的研究中，待解决的问题通常相当复杂，需要严谨的处理方式（研究流程）来解决复杂的问题，一般社会科学的研究流程有：确立研究动机→拟定研究目的→相关文献探讨→建立研究架构→决定研究方法→资料搜集与分析→研究结论与建议，我们整理一般的研究流程如下图：

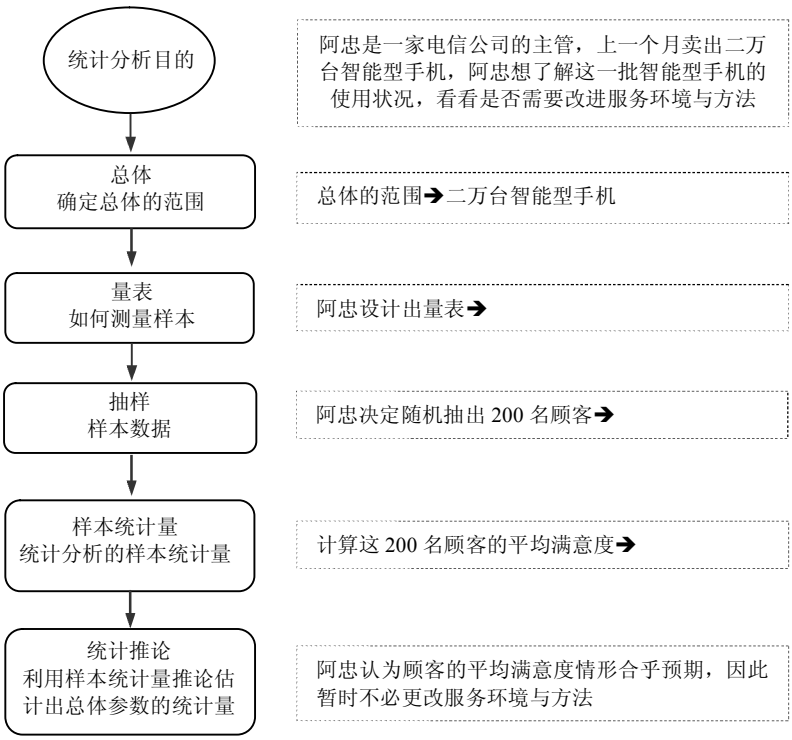


在我们一般生活或工作的社会中找出有意义的问题，形成我们的研究动机，在产生研究动机后，进而拟定研究目的。接着根据研究动机与目的来进行文献探讨，从文献探讨中建立观念性的研究架构，根据此架构决定所应使用的研究方法，包括问卷设计、数据分析工具的选择及分析方法的使用。在问卷回收期满结束后开始进行数据分析，以提供研究者进行讨论，最后作出研究结论及建议。

在社会科学的研究中，统计分析扮演的角色是十分重要的，统计分析也是社会科学研究中的一部分，在确定使用的研究方法和统计分析目的，统计分析首先需确定总体的范围，设计出量表（问卷设计），接着就是统计分析的工具选择及分析方法的使用，接着在问卷回收期满结束后开始进行数据分析，呈现出正确的数据分析结果。统计分析的实施步骤如下图：

统计分析的实施步骤：

在整个社会科学的研究中，牵涉到统计分析的部分相当广，包括理论、量表（问卷设计）、抽样（问卷发放和回收）、统计分析的基础统计学和常用的统计分析（多元分析或称为数量方法）。因此本章节将分别介绍统计分析简介与数量方法的基础，包括理论简介、量表简介、抽样简介、基础统计学和常用的统计分析（多元分析或称为数量方法）。



1-2 理论

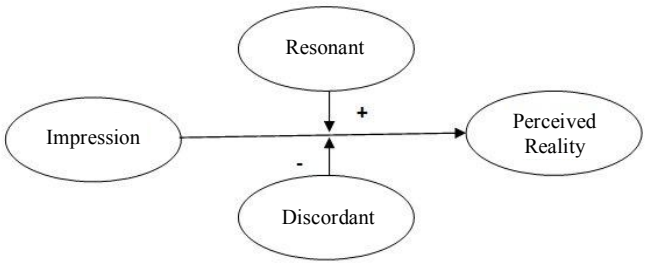
在我们一般生活中，理论是一组叙述或原理，用来解释事实或现象。在社会科学的研究中，理论是社会现象的系统观，也可以是一个信念或原理，用来引导行动，协助了解或判断社会现象。在社会科学的研究里有实证性的研究，实证性的研究是以某个理论为基础，进行相关现象的验证，在学术上，可以延伸和拓展理论的应用，在实务上，可以提供实务工作者遵循的依据。理论在社会科学的研究流程中常常扮演着主导的角色，在确立研究动机→拟定研究目的时，常常会提到以什么观点探讨，相当多的研究就会使用一个或多个理论观点探讨，接着相关文献探讨中，就必须要整理“理论观点”的相关文献，建立研究的模型也常来自于理论，研究结论与建议也需要提到研究议题的理论贡献和学术上的意涵。理论在社会科学的研究中扮演着核心的角色，也就是在社会科学的研究中，了解的社会现象是由理论建构起系统观，因此，我们更应该好好地了解什么是理论？以及更多的理论。

什么是理论？有关理论的说明是多方面的，理论是一组叙述或原理，用来解释事实或现象，特别是可以重复的验证，可以用来预测或是已经被广泛接受的自然现象。理论是社会现象的系统观，也可以是一个信念或原理用来引导行动，协助了解或判断。理论使用的分类也是多方面的，2006

年 Gregor 在 MIS Quarterly 的一篇文章中，区分理论使用的五大分类，有：①theory for analyzing（理论用来分析）；②theory for explaining（理论用来解释）；③theory for predicting（理论用来预测）；④theory for explaining and predicting（理论用来解释和预测）；⑤theory for design and action（理论用来设计和行动）。理论与研究的关系，我们所做的研究可以是研究先于理论（探索性的研究）或理论先于研究（实证性的研究），探索性的研究是对于某些尚未了解的现象进行初步的探讨，以熟悉此现象，取得的结果作为后续研究的基础。实证性的研究是以某个理论为基础，进行相关现象的验证，在学术上，可以延伸和拓展理论的应用，在实务上，可以提供实务工作者遵循的依据。理论在各个学科，例如：教育学、艺术学、体育学、图书信息学、心理学、法学、政治学、经济学、社会学、传播学、人类学、教育学、管理学（人力资源、组织行为、战略管理、医务管理、生管、交管、营销、资管、数量方法与项目研究应用）等，都扮演着相当重要的角色，我们整理常见的理论如后。

1-2-1 印象管理理论

印象管理理论（Theory of Impression Management）是由 Erving Goffman 于 1959 年所提出的，如下图。印象管理理论解释了在复杂人际互动和事实背后的动机。也说明每个人都会配合情境，运用适合的策略呈现自己。

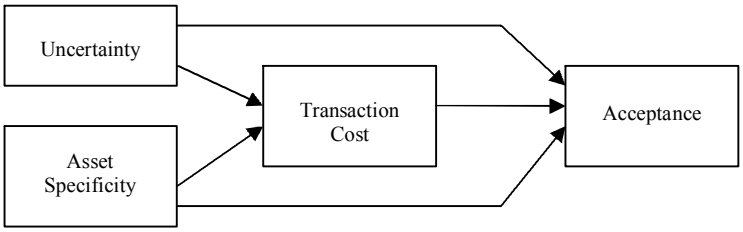


相关资料：

- Dillard, C., Browning, L.D., Sitkin, S.B., and Sutcliffe, K.M. 2000. “Impression Management and the Use of Procedures at the Ritz-Carlton: Moral Standards and Dramaturgical Discipline,” *Communication Studies* (51:4), pp. 404-414.
- Giacalone, R.A., and Rosenfeld, P. 1989. *Impression Management in the Organization,* Hillsdale, NJ: Lawrence Erlbaum Associates.
- Giacalone, R.A., and Rosenfeld, P. 1991. *Applied Impression Management*, Newbury Park, CA: Sage.
- Goffman, E. 1959. *The Presentation of Self in Everyday Life,* New York, NY: Doubleday.
- Schlenker, B.R. 1980. *Impression Management: The Self-Concept, Social Identity, and Interpersonal Relations,* Monterey, CA: Brooks/Cole Publishing Co.

1-2-2 交易成本理论

交易成本理论（Transaction Cost Theory）是由诺贝尔经济学奖得主科斯（Coase，1937）所提出，交易成本是指“当交易行为发生时，所随同产生的各项成本”。然而不同的交易往往就涉及不同种类的交易成本，例如：Williamson（1975）提出的交易成本包含：搜寻成本、信息成本、议价成本、决策成本、监督交易进行的成本、违约成本。使用交易成本理论的研究模型如下：



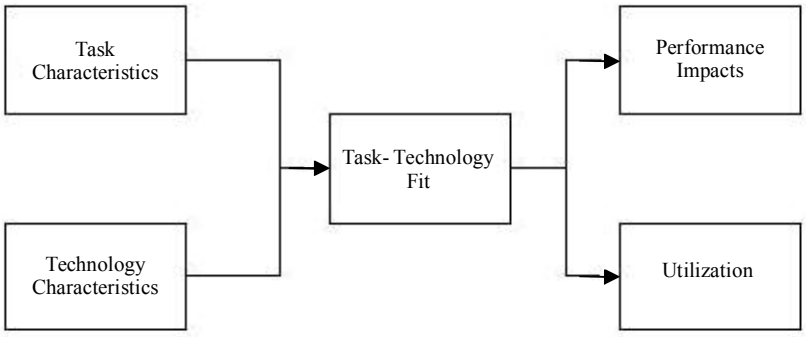
Source：Liang, T.P., and Huang, J.S. / Decision Support Systems (1998)

相关资料：

- Coase, R.H. 1937. “The nature of the firm,” *Economica, New Series* (4:16), pp. 386-405.
- Liang, T.P., and Huang, J.S. 1998. “An Empirical Study on Consumer Acceptance of Products on Electronic Markets: A Transaction Cost Model,” *Decision Support Systems* (24:1), pp. 29-43.
- Oliver, W. 1975. *Markets and hierarchies: Analysis and antitrust implications*, New York, NY: Free Press.

1-2-3 任务、技术匹配理论

任务、技术匹配理论（Task Technology Fit Theory）由 Goodhue 与 Thompson 于 1995 年提出任务、技术匹配理论，如下图。



Source：Goodhue and Thompson (1995)

任务、技术匹配理论认为 IT 可以正向地影响使用者个人的绩效，并且可以结合任务的能力，很容易被使用者所使用。

相关资料：

- Goodhue, D.L. 1995. “Understanding user evaluations of information systems,” *Management Science* (41:12), pp. 1827-1844.
- Goodhue, D.L., and Thompson, R.L. 1995 “Task-technology fit and individual performance,” *MIS Quarterly* (19:2), pp. 213-236.
- Zigurs, I., and Buckland, B.K. 1998. “A theory of task/technology fit and group support systems effectiveness,” *MIS Quarterly* (22:3), pp. 313-334.

1-2-4 长尾理论

2004 年 10 月，《Wired》杂志主编 Chris Anderson 首次提出长尾理论（The long tail），以简单的图表解释了电子商务的利基所在。通路只要够大，不是主流的商品（例如：需求量小的商品）“总销量”也能够和主流的、需求量大的商品销量竞争。在 Internet 上的实例就是 Amazon 和 Google，因此，长尾理论使得电子商务拥有一个具有说服力的理论基础。

1-2-5 制度理论

制度理论（Institutional Theory）由 Selznick（1948）、Dimaggio 和 Powell（1983）所倡导，制度理论的观点认为，组织在面对的制度环境（Institutional Environments）下会采取某一种组织结构设计或某项措施是为了取得所需的资源及组织内部成员和外部社会的支持，期望获得组织生存的合法性（legitimacy）。

相关资料：

- Selznick, P. 1948. “Foundations of the Theory of Organizations,” *American Sociological Review* (13), pp. 25-35.
- DiMaggio, P.J., and Powell, W.W. 1983. “The iron cage revisited: Institutional isomorphism and collective rationality in organizational fields,” *American Sociological Review* (48:2), pp. 147-160.

1-2-6 服务质量理论

服务质量理论（Service Quality，SERVQUAL）是由 Parasuraman、Berry 和 Zeithaml 为主要倡导学者，于 1985 年提出服务质量的概念模型。服务质量理论 SERVQUAL 是用来衡量服务质量，

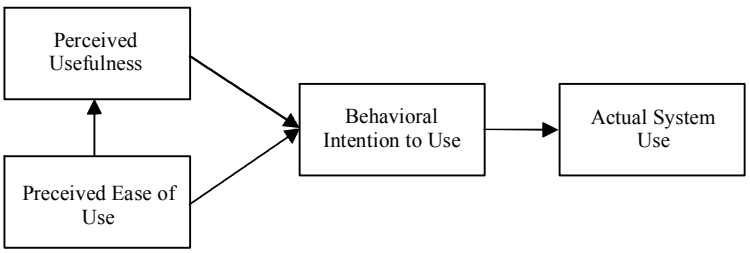
其定义是“认知服务质量”为顾客的“期望”与“感受”之间的差距。Parasuraman et al.（1988）提出服务质量的五个构念有：①Tangibles（有形性）：physical facilities, equipment, staff appearance, etc；②Reliability（可靠性）：ability to perform service dependably and accurately；③Responsiveness（反应度）：willingness to help and respond to customer need；④Assurance（信赖感）：ability of staff to inspire confidence and trust；⑤Empathy（关怀度）： the extent to which caring individualized service is given。服务质量理论（SERVQUAL, Service Quality）除了大量应用于营销领域外，目前有涉及到服务的各个领域，大多都认同以 SERVQUAL 做为衡量服务质量的考虑。

相关资料：

- Parasuraman, A., Berry, L.L., and Zeithaml, V.A. 1985. “A Conceptual Model of Service Quality and Its Implications for Future Research,” *Journal of Marketing* (49: 4), pp. 41-50.
- Parasuraman, A., Berry, L.L., and Zeithaml, V.A. 1988. “SERVQUAL: A Multiple-Item Scale For Measuring Consumer Perceptions of Service Quality,” *Journal of Retailing* (64:1), pp. 12-40.
- Parasuraman, A., Berry, L.L. and Zeithaml, V.A. 1991. “Refinement and Reassessment of the SERVQUAL Scale,” *Journal of Retailing* (67:4), pp. 420-450.

1-2-7 技术接受模型

技术接受模型（Technology Acceptance Model，TAM）是 Davis 于 1986 年所提出的，Davis 使用理性行为理论（Theory of Reasoned Action，TRA）和计划行为理论（Theory of Planned Behavior，TPB）模型为基础，发展成 TAM 技术接受模型，用来研究用户接受信息科技（Information Technology，IT）的影响因素。TAM 认为影响使用者的使用意向主要有认知有用性（Perceived Usefulness，PU）及认知易用性（Perceived Ease of Use，PEOU）这两个构念，再通过使用态度（Attitude Towards）及使用意向（Behavioral Intention to Use）的影响，进而实际地使用 IT。



Source：Davis et al. (1989), Venkatesh et al. (2003)

相关资料：

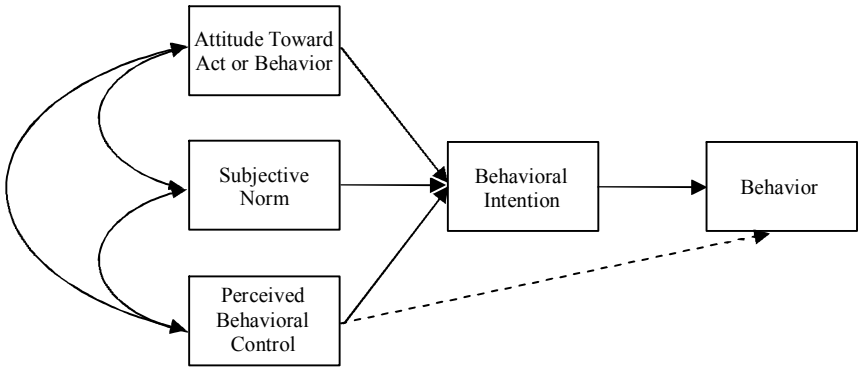
- Davis, F.D. 1986. “A technology acceptance model for empirically testing new end-user information systems: Theory and results,” Doctoral dissertation, Sloan School of Management, Massachusetts

Institute of Technology.

- Davis, F.D. 1989. “Perceived usefulness, perceived ease of use, and user acceptance of information technology,” *MIS Quarterly* (13:3), pp. 319-339.
- Venkatesh, V., Morris, M.G., Davis, G.B., and Davis, F.D. 2003. “User acceptance of information technology: Toward a unified view,” *MIS Quarterly* (27:3), pp. 425-478.

1-2-8 计划行为理论

计划行为理论（The Theory of Planned Behavior，TPB）为美国心理学家 Ajzen 所倡导，是用来预测行为的重要理论。计划行为理论（Ajzen 1985, 1991）指出“行为意向（behavior intention, BI）”是个人对从事某项行为（Behavior，B）的意愿，是预测行为最好的指标。意向由三个构念所组成：①对该行为所持的态度（Attitude Toward the Behavior，AT）；②主观规范（Subjective Norm，SN）；③感知行为控制（Perceived Behavioral Control，PBC）。计划行为理论的模型如下：



Source：Ajzen (1991)

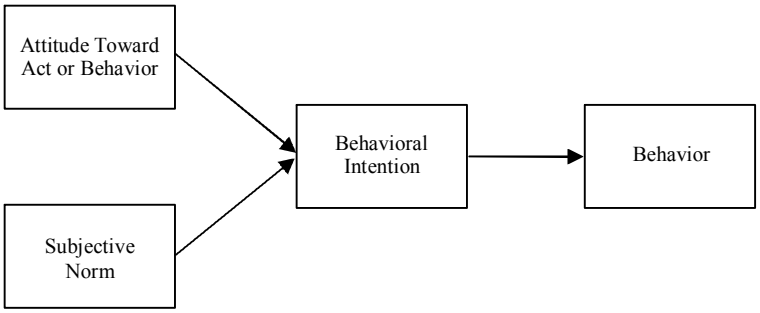
- Ajzen, I. 1985. “From intentions to actions: A theory of planned behavior,” In *Springer series in social psychology,* J. Kuhl, and J. Beckmann (eds.), Berlin: Springer. pp. 11-39.
- Ajzen, I. 1991. “The theory of planned behavior,” Organizational Behavior and Human Decision Processes (50:2), pp. 179-211.

计划行为理论就是想要分析出行为与心理之间的关系，进而从研究结果中找出可以影响行为的因素。

1-2-9 理性行为理论

理性行为理论（Theory of Reasoned Action，TRA）是由 Fishbein 和 Ajzen 于 1967 年所提出预测个人行为态度意向的理论。理性行为理论认为行为意向（Behavioral Intention）会受到“态度”

及“主观规范”所影响。态度是指个人对行为的想法，主观规范是指社会习俗、他人意见或压力。理性行为理论的模型如下：



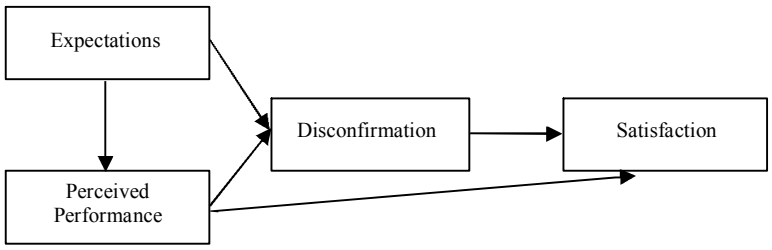
Source: Fishbein, M., and Ajzen, I. (1975).

相关资料：

- Ajzen, I., and Fishbein, M. 1973. “Attitudinal and normative variables as predictors of specific behavior,” *Journal of Personality and Social Psychology* (27:1), pp. 41-57.
- Fishbein, M. 1967. “Readings in attitude theory and measurement,” *Attitude and the prediction of behavior*, M. Fishbein (ed.), New York: Wiley. pp. 477-492.
- Fishbein, M., and Ajzen, I. 1975. *Belief, attitude, intention, and behavior : An introduction to theory and research,* Reading, MA: Addison-Wesley.

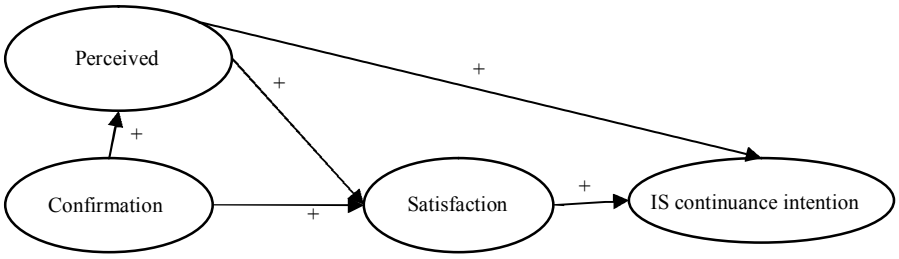
1-2-10 期望确认理论

期望确认理论（Expectation Confirmation Theory，ECT）最早由 Oliver（1977, 1980）提出的，是一般研究消费者满意度的基础模型，概念为消费者在购买前的预期及实际购买后的绩效，两者比较之后，有正向确认、负向确认，最后产生满意程度上的差异，如下图。



Source: Oliver (1977, 1980)

Bhattacherjee 修正了 ECT，提出“持续使用 IS 意向模型”，使其符合信息系统的情境，如下图。



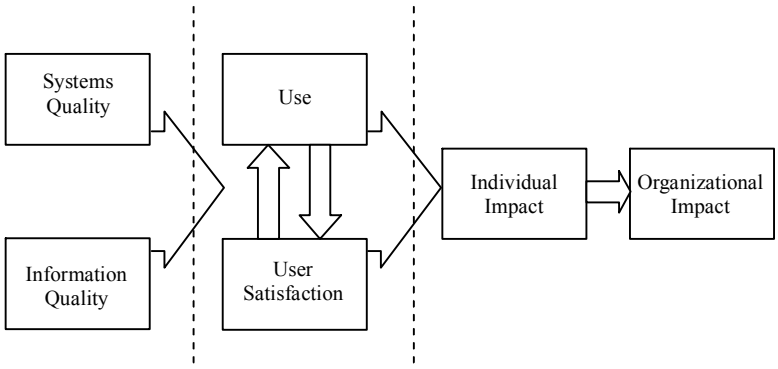
数据源：Bhattacherjee（2001）持续使用 IS 意向

相关资料：

- Oliver R.L. 1977. “Effect of Expectation and Disconfirmation on Post exposure Product Evaluations - an Alternative Interpretation,” *Journal of Applied Psychology* (62:4), pp. 480.
- Oliver R.L. 1980. “A Cognitive Model of the Antecedents and Consequences of Satisfaction Decisions,” *Journal of Marketing Research* (17:3), pp. 460.
- Spreng, R.A., MacKenzie, S.B., and Olshavsky, R.W.1996. “A reexamination of the determinants of consumer satisfaction,” *Journal of Marketing* (60:3), pp. 15.
- Bhattacherjee, A. 2001. “Understanding information systems continuance: An expectation- confirmation model,” *MIS Quarterly* (25:3), pp. 351.

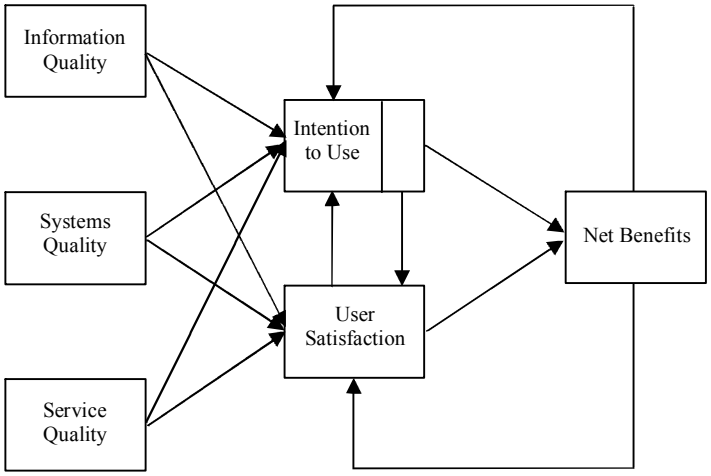
1-2-11 信息系统成功模型

DeLone and McLean 于 1992 年提出信息系统成功模型（Information Systems Success Model，也叫 Delone and McLean IS Success Model）。信息系统成功模型有六大构念：系统质量（Systems Quality）、信息质量（Information Quality）、使用（Use）与使用者满意（User Satisfaction）、个人的影响（Individual Impact）与组织的影响（Organizational Impact），如下图：



Source: Information Systems Success Model (DeLone & McLean, 1992)

DeLone and McLean（2003）回顾十年（1993 年到 2002 年中期）期刊，研究者提出了更新的模型如下图：



Source: Updated Information Systems Success Model (DeLone & McLean 2002, 2003)

相关资料：

- DeLone, W.H., and McLean, E.R. 1992. "Information Systems Success: The Quest for the Dependent Variable," *Information Systems Research* (3:1), pp 60-95.
- DeLone, W.H., and McLean, E.R. 2003. "The DeLone and McLean Model of Information Systems Success: A Ten-Year Update," *Journal of Management Information Systems* (19:4), pp. 9-30.

1-2-12 资源依赖理论

资源依赖理论（Resource Dependency Theory，RDT）是 Pfeffer and Salancik 于 1978 年所提出。资源依赖理论是指组织在一个开放性的社会系统和不确定性的环境下，无法自给自足，组织为了求生存需要依赖外部资源的供给，并适时提供资源给外部组织，而与外部环境不断地互动，组织才能持续生存下去。

相关资料：

- Pfeffer, J., and Salancik, G. 1978. The external control of organizations: A resource dependence perspective, New York: Harper & Row.
- Ulrich, D., and Barney, J.B. 1984. “Perspectives in organizations: Resource dependence, efficiency, and population,” *Academy of Management Review* (9:3), pp. 471.

1-2-13 资源基础理论

资源基础理论（Resource-Based Theory）是以 Penrose、Wernerfelt、Barney 等为主要倡导学者，资源基础理论的先驱是 Penrose（1959），在其《The theory of the growth of the firm》（企业的成长理论）一书中，提到企业为获取利润，不仅要拥有优越的资源，更要有有效地利用这些资源的能力，来追求企业成长。Wernerfelt（1984）延续 Penrose 的论点，在其“企业的资源基础观点”文章中，首先提出资源基础观点（Resource-Based View，RBV），以“资源观点”取代“产品观点”来分析企业。Barney（1986）则延续 Wernerfelt 所提出的观点，认为不同的企业对于不同的策略资源，所产生的价值也不相同，所以企业的绩效不只来自产品市场的竞争，也由企业不同的资源所产生，因此企业进行战略规划时，应先分析本身所具备的各种具有竞争优势的资源，例如：有价值的资源、稀有性的资源、不可模仿性的资源、不可替代性的资源。对于资源基础理论（Resource-Based Theory），不同的学者或许会有不同的见解与看法，但共同的最终目的在于探讨企业如何获取最大利益。

相关资料：

- Barney, J.B. 1986a. “Strategic factor markets: Expectations, luck and business strategy,” *Management Science* (32), pp. 1512-1514.
- Barney, J.B. 1986b. “Organizational culture: Can it be a source of sustained competitive advantage?” *Academy of Management Review* (11), pp. 656-665.
- Barney, J.B. 1986c. “Types of Competition and the Theory of Strategy: Toward an Integrative Framework,” *Academic of Management Review* (11), pp. 791-800.
- Penrose, E.T. 1959. *The Theory of the Growth of the Firm*, New York: Wiley.
- Wernerfelt, B. 1984. “A resource-based view of the firm,” *Strategic Management Journal* (5), pp. 171-180.

1-2-14 满意度

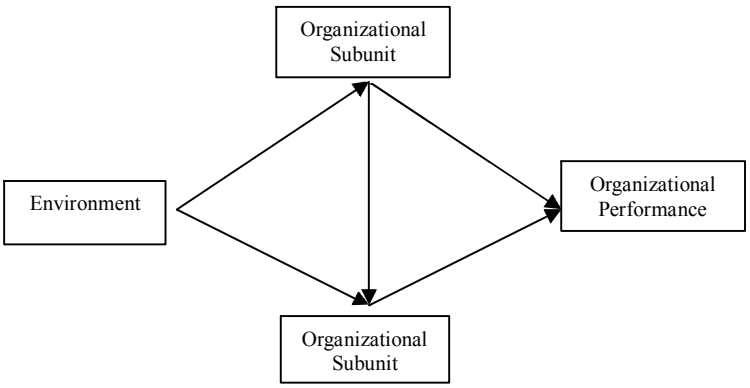
满意度（satisfaction）一般是指一个人感觉到愉快或失望的程度。由于对象的不同，使用的范围和方式也会有所不同，以顾客对产品为例子，Miller（1977）认为顾客满意度是由顾客的预期之程度和知觉之成效，二者交互作用所形成的程度。Kotler（1991）也认为顾客满意度的高低是取决于顾客感受的知觉价值和顾客的期望水平。以顾客对服务为例子，Hernon 等人（1999）认为建立顾客满意度应包含对接待人员的满意度和整体服务满意度两部分。以用户对信息系统为例子，Bailey and Pearson（1983）是通过文献研究、专家访问与访问调查等方式，归纳整理出 39 个问项（如正确性、及时性与人员的态度等）。以测量受访者对各问项相对信息需求的认知反应结果与强度。进而可以从研究结果中找出可以影响信息系统满意度的因素。对于一个组织而言，提供顾客满意的服务，是组织生存的必要条件之一，所以各行各业对于顾客满意度都相当的重视。

相关资料：

- Bailey, J.E., and Pearson, S.W. 1983. "Development of a tool for measuring and analyzing computer user satisfaction," *Management Science* (29), pp.530-545.
- Hernon, P.N., Danuta, A, and Altman, E. 1999. "Service Quality and Customer Satisfaction; an assessment and future direction," *The Journal of Academic Librarianship* (25), pp. 9-17.
- Kotler, P. 1991. *Marketing management: Analysis, planning, implementation, and control* (7th ed.), Englewood Cliffs, NJ: Prentice-Hall, pp. 455-459.
- Miller, J.A. 1977. "Studying Satisfaction Modifying Models, Eliciting Expectation, Posing Problem, and Meaningful Measurement," in *The Conceptualization of Consumer Satisfaction and Dissatisfaction,* H. Hunt (ed.), Cambridge: Marketing Science Institute.

1-2-15 权变理论

权变理论（Contingency Theory）是费德勒（Fiedler, 1964）于 1964 年所提出。一个简化的权变理论模型如下图：



A simplified model of contingency theory in organizational research

权变理论认为组织效能依赖于组织设计与其所面临的情境的适配度，特别强调情境因素的重要性。

相关资料：

- Fiedler, F.E. 1964. "A Contingency Model of Leadership Effectiveness," *Advances in Experimental Social Psychology* (1), New York: Academic Press. pp. 149-190.
- Weill, P., and Olson, M.H. 1989. "An Assessment of the Contingency Theory of Management Information Systems," *Journal of Management Information Systems* (6:1), pp. 63.

1-3 量表简介

研究人员想轻易地使用统计技术来分析资料时，大多都需要严谨的量表，但是量表如何测量数据呢？这就需要先了解数据的测量尺度，因此，本节将介绍数据的测量尺度和量表。

1-3-1 数据的测量尺度

社会科学的数据具有多种性质，因此，我们在测量这些数据时，就需要有不同的尺度，测量尺度（Scales of Measurement）是对资料给予适当的代表值，以作为统计运算的基础，一般我们常用的测量尺度有：名目尺度（Nominal Scale）、顺序尺度（Ordinal Scale）、区间尺度（Interval Scale）和比例尺度（Ratio Scale），分别介绍如下：

- 名目尺度（Nominal Scale）：名目尺度是用来处理分类的数据，也称为类别尺度（Ategorial Scale），在分类的数据中，都会以一个数字来代表一个类别，常用的范例如下：
  - 性别：0 代表男性，1 代表女性。
  - 婚姻： 0 代表未婚，1 代表已婚。
  - 企业规模：0 代表中小企业，1 代表大企业。
- 顺序尺度（Ordinal Scale）：顺序尺度用来处理有前后关系的数据，以表示高、低，好、坏，等级，这些数据可以给予大小不同的值，这些值只代表顺序，不代表差距有多大，也不代表有相同的距离，常用的范例如下：
  - 教育程度：0 代表小学，1 代表初中，2 代表高中，3 代表本科，4 代表研究生，5 代表博士。
  - 职位层级：0 代表实习生，1 代表职员，2 代表经理，3 代表总经理，4 代表董事长。
- 区间尺度（Interval Scale）：区间尺度用来处理有标准化的测量单位和相同距离尺度的数据，这些数据并无真正的零（无数据），常用的范例如下：
  - 温度：有-5°C、0°C、10°C 等。
  - 时间：有时、分、秒等。

区间尺度的大小是有意义的，数值之间的差距也有代表的意义，由于处理的是相同距离尺度的数据，所以也称为等距尺度或间距尺度。

- 比例尺度（Ratio Scale）：比例尺度用来处理其有标准化的测量单位和绝对零值的数据，绝对零值的意思是数值为零时，就代表无此数据，常用的比例尺度范例如下：
  - 年龄：1 岁、10 岁、20 岁、40 岁、60 岁、80 岁、100 岁。
  - 身高：20 公分、60 公分、100 公分、150 公分、200 公分。
  - 体重：20 公斤、40 公斤、60 公斤、80 公斤。

以上介绍数据的 4 种测量尺度，若是以涵盖的范围来比较则是名目尺度（类别）最小，接下来是顺序尺度、区间尺度，最大范围的是比例尺度，我们以下图来表示：

名目（类别）＜顺序＜区间＜比例

处理数据的范围越大 →

数据处理的范围越小，其处理的程度较不精确，例如：名目尺度的数据，反之，数据处理的范围越广，其处理的程度较精确，例如：比例尺度的资料最精确，至于我们该用何种测量尺度，则应视研究的数据形态而定，并非随意指定的，读者需要加以小心使用。

资料的分布

数据的各种分布情形代表的是测量而得到的各种特性，例如：我们调查某一家连锁超市的顾客消费行为，就可以知道顾客消费的大致情形，一般最常想了解的资料的分布情形有：资料的集中趋势（Measure of Central Tendency）和资料的离散程度（Measures of Dispersion），分别介绍如下：

- 资料的集中趋势：是在一组数据中，找出一个值，使得其他的数值会往它集中，常用的测量方法有众数、中位数、几何平均值、算术平均值和加权算术平均值，分别介绍如下：
  - 众数（mode）：计算出次数最多的观察值。
  - 中位数（median）：计算出位置排列在中央的数值。
  - 几何平均值（Geometric Mean）：用来处理等比级数的平均值，是观察值相乘 n 次就开 n 次根号，我们以几何平均值 G、观察值 X_1 、 X_2 、……、 X_n 为范例，数学式如下：

$$G = \sqrt[n]{X_1 X_2 X_3 \cdots X_n} = \prod_{i=1}^{n} (X_i)^{\frac{1}{n}}$$

  - 算术平均值（Arithmetic Mean）：将观察值加总，再除以观察值的个数，我们以算术平均值 $\overline{X}$，观察值 X_1 、 X_2 、……、 X_n 有 n 个为范例，数学式如下：

$$\overline{X} = \frac{X_1 + X_2 + \cdots + X_n}{n} = \frac{\sum_{i=1}^{n} X_i}{n}$$

  - 加权算术平均值（Weighted Arithmetic Mean）：在算术平均值中，每个观察值都是相同的比重，若是遇到观察值的重要程度不一样时，可以对每个观察值给以权重，再计算其平均值，我们以加权算术平均值 $\overline{X}_w$，观察值 X_1 、 X_2 、……、 X_n 有 n 个，权重（weight）：X_1 是 W_1，X_2 是 W_2， X_n 是 W_n 为范例，数学式如下：

$$\overline{X_w} = \frac{X_1W_1 + X_2W_2 + \cdots + X_nW_n}{W_1 + W_2 + \cdots + W_n} = \frac{\sum_{i=1}^{n} X_iW_i}{\sum_{i=1}^{n} W_i}$$

- 数据的离散程度：数据的离散程度是用来确认数据的集中趋势（例如：平均值）是否具有代表性的问题，资料的离散程度低时，平均值就具有高的代表性，反之，数据的离散程度高时，平均值会具有较低的数据代表性，常见的数据的离散程度测量有极差、方差和标准差，分别介绍如下：
  - 极差（range）：一组观察值中最大值与最小值的差，以 R 表示。

    R = 最大值 – 最小值

    R 越大，代表离散程度越大

    R 越小，代表离散程度越小
  - 方差：方差也称为平均平方离差（Mean Squared Deviation），是观察值与平均值离差（相减）的平方和，除以观察值的个数。我们以总体方差 σ^2（σ = sigma），总体观察值 X_1、X_2、……、X_n 有 n 个，总体平均值为 u，数学式如下：

$$\sigma^2 = \frac{\varepsilon(x-u)^2}{N}$$

    当方差的数据源是样本时，由于样本方差 S^2 失去一个自由度（Degree of Freedom，df），所以当样本方差分母为 $n-1$，x 为观察值，$\overline{x}$ 为平均值，样本方差为 S^2，数学式如下：

$$S^2 = \frac{\varepsilon(X-\overline{X})^2}{n-1}$$

  - 标准差：用来解释资料离散的情形，由于方差是平方值，不易解释，经由开根号后会得到标准差，标准差同样分为总体的标准差和样本的标准差，数学式如下：

    * σ 为总体的标准差：σ^2 为总体的方差

$$\sigma = \sqrt{\sigma^2}$$

    * S 为样本的标准差：S^2 为样本的方差

$$S = \sqrt{S^2}$$

（1）标准差会保留方差的所有特性，也易于解释，所以我们一般都用标准差作解释。

（2）标准差越大，代表资料越分散；标准差越小，代表资料越集中。

1-3-2 量表

在社会科学的研究分类中，可以分为质性研究和量化研究，无论是质性研究或量化研究都是可以加以测量的，只是测量的标准化程度不同。在质性研究中，研究者尽可能地搜集受访者的各种信息，因此，拥有最丰富的信息，但由于问项未经严谨处理，因此，标准化程度低，一般我们称这样的测量方式是使用非结构化的问项（Unstructured Questionnaire），非结构化的问项由于未设定明确的问项主轴内容，容易导致数据分析时，没有一定的方向可以遵循，因此，许多研究者会先设定问项的主轴，搜集受访者资料时，只分布一定范围内，采用非结构化的处理，这样的测量方式是使用半结构化的问项（Semi-structured Questionnaire），半结构化的问项仍倾向适用于质性研究，因为使用统计技术分析时，仍有相当的难度。研究人员想轻易地使用统计技术来分析资料时，大多都会严谨地先制定一个测量的工具——量表，具有一定的测量格式，以提供受访者自行填答（Self Reported），我们称这样的测量方式为结构化的问项（Structured Questionnaire），我们将结构化和非结构化的测量方式整理如下：

社会科学的研究

- 质性研究
  - 非结构化问项测量
  - 半结构化问项测量
- 量化研究 —— 结构化问项测量

在量化研究中，通常需要一个量表来进行问卷调查，量表是用一个以上的指标（indicator）来衡量待测物体或对象的特性，并且可以将此特性数值化，一般我们常用到的量表格式有 Thurstone 量表、Guttman 量表、语意差别量表（Semantic Differential Scale）和李克特量表（Likert Scale），分别介绍如下：

- Thurstone 量表：

  研究人员在编制 Thurstone 量表时，会先写好要测量的项目，交由专家来筛选项目（11 点的评分方式），我们对每一项计算其平均值和四分位数（Q Score），四分位数较高的题项代表一致性较差，适合被删除，然后选出较一致性的题项，进行后续的工作，由于 Thurstone 量表的编制较复杂，还有专家的主观问题，因此，现在较少使用，在一般的期刊论文中也较少看到。

- Guttman 量表：

  研究人员在编制 Guttman 量表时，会先将题项依一定的方向排列，也就是说，题项的内涵是由强到弱或由弱到强的方式编排，填答者依序作答时，会遇到转折的题项，在转折之前，填答者一致性的回答，就代表累积了多少分数，因此，也称 Guttman 量表为累积量表。

- 语意差别量表：

  研究人员在编制语意差别量表时，主要是使用形容词在语意上的差别，这些形容词常常是成对出现，语意经常是正好相反的，例如：快的－慢的、好的－坏的、重的－轻的、强的－弱的、忙的－闲的等，研究人员可以计算每个题项的平均值，还可以使用因子分析所取得的构念进行加总后，再进行后续的工作。

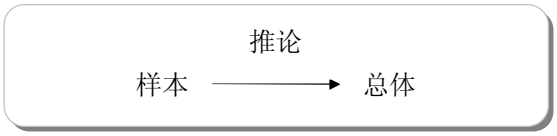
- 李克特量表（Likert Scale）：

  李克特量表在社会科学研究中，是最常出现的量表，广泛地应用在营销、组织行为、人力资源、学习科技、教育、财务管理、心理测验中等，特别适用于感受或态度上的衡量，例如：商业品牌代表着产品的销售好坏，李克特 5 点量表，可以使用数值 1 代表非常不同意，2 代表不同意，3 代表没有同意或不同意，4 代表同意，5 代表非常同意。李克特 5 点量表的数值与数值之间是等距的，经由因子分析后的构念可以加总，以得到一个加总计分值，形成一个构念可以由一个值代表，因此，李克特量表也是可加总量表的一种。

在量表的使用上，理论一直是很重要的因素，因为理论可以协助我们概念化测量的问题，凭借理论的协助，可以使我们发展或使用具有一致性的、有效的、可应用的量表。

1-4 抽样

为什么我们需要抽样（Sampling）？原因是总体太大，无法取得所有总体的数据，或是因为取得总体数据的成本太高，负担不起，因此，我们可以通过抽样取得的样本来推论总体，如下图：



抽样的好坏会直接影响推论的结果，也就是说，样本的正确性和准确性是相当重要的影响因素，才能代表（推论）总体，因此，我们进行抽样的目的是想获得具有代表性的样本，这样的样本才能代表总体，然而，存在社会现象的总体有很多种，必须针对不同的总体采取不同的抽样方法，才能获得代表性的样本，常见的抽样方式有简单随机抽样（Simple Random Sampling）、分层抽样（Stratifies Sampling）、整群抽样（Cluster Sampling）和便利抽样（Convenience Sampling），分别介绍如下：

- 简单随机抽样（Simple Random Sampling）：

  简单随机抽样会使总体的每一个单位被抽中的概率都一样，例如：A 班有 60 位学生，我们要从 A 班抽出 10 位学生参加拉拉队比赛，则可以使用简单随机抽样，常用的方式是使用随机数表（Random Number Table）来协助选出适当的样本。

  - 优点：当总体较小时，容易执行以取得适当的样本。

- 缺点：当总体较大时，总体的完整名册不易获得，造成抽样时的成本较大，执行起来很困难。
- 使用时机：（1）总体有完整的数据，可以进行编号。
  （2）总体的抽样单位差异小，才不会抽出偏误的样本。

■ 分层抽样（stratifies sampling）：

分层抽样是将总体依某个准则，区分成 N 个不重叠的组，我们称这些组为层（strata），先将总体区分成几个不同的层，之后再从每一层中分别抽取样本，最后将各层抽取的样本集合起来，成为我们所需要的总样本，这就是分层抽样。例如：B 系有博士班、硕士班和大学部，我们可以从这三个不同的层，抽出一定比例的人数参加“校长座谈”，这就是分层抽样。

- 优点：样本分布较平均，可以提高精确度，并且可以比较各层样本的差异做比较分析使用。
- 缺点：分层的特性若是没有考虑好，则会有抽样不均的情形，反而降低精确度。
- 使用时机：（1）总体的抽样单位差异较大时。
  （2）总体经分层后，层与层之间的变异较大，该层内的变异较小。

■ 整群抽样（随机抽样的一种）：

整群抽样是将总体分成几个群集（例如：部落或县、市、邻、里），经过随机选取群集后，只有选中的群集才进行抽出样本或进行普查，例如：研究人员想调查大专学生的生活支出时，可以从全国 150 所大专院校中，先随机抽出 15 所大专院校，再从这 15 所大专院校中，每所学校抽出 100 位学生当做样本，这就是整群抽样。

- 优点：可以大量降低抽样成本，容易实行。
- 缺点：容易发生抽样偏误，风险较高。
- 使用时机：群集与群集之间变异小、群集内的变异大时适用，刚好与分层抽样的使用时机相反。

■ 便利抽样（非随机抽样）：

便利抽样从字面上解释，是属于很方便进行抽样的方式，例如：街头访问、信息展的访问等。

- 优点：成本低，样本容易取得。
- 缺点：抽样取得的样本，缺乏代表性，所以较少使用。

在量化研究的抽样使用上，可以分为 pilot test 抽样和实测抽样，pilot test 抽样的目的是为了验证量表适切性和构念的正确性，实测抽样则是为了研究的结果所作的抽样，这两种抽样取得的样本，都要依赖后续章节数量方法的运算，得到我们需要的统计量，才能对研究的结果下结论。

1-5 统计分析的基础统计学

本节讨论的是在统计分析（多元分析或称为数量方法）中会用到的基础统计，方便读者理解统计分析的内涵，并不是讨论或介绍艰深的统计学。

1-5-1 描述性统计资料

一般基本的统计数据是要能描述数据的特性，例如：平均值（Mean）、中位数（Median）、众数（Mode）、标准差（Std deviation）、方差（Variance）等。以了解资料的集中趋势（Central Tendency）和离散情形（Dispersion）。常用的一般统计测量数如下：

- 百分位数值（Percentile Values）：
  - 四分位数（Quartiles），将数值排序后，分成四等份。
  - 自定义的几个相等分组（Cut point for equal groups）。
  - 百分位数（Percentile），将数值排序后，分成 100 等份，用来观察数据较大值或较小值百分比的分布情形。
- 集中趋势（Central Tendency）：
  - 平均值（Mean），将观察值加总，再除以观察值的个数，用来观察数据的平衡点，但是较容易受到极端值的影响。
  - 中位数（Median），计算出位置排列在中央的数值，适用于顺序数据或比例数据，较不易受到极端值的影响。
  - 众数（Mode），计算出次数最多的观察值，适用于类别资料（例如民意调查），不受到极端值的影响。
  - 总和（Sum），将观察值加总。
  - 分组的中间点的值（Values are group midpoints）。
- 离散情形（Dispersion）：
  - 标准差（Std deviation），将方差开根号，回归原始的单位，标准差越大代表资料越分散。
  - 方差（Variance），将观察值与平均值之差，平方后进行加总，再除以观察值的个数，方差越大代表资料越分散。
  - 极差（Range），将观察值中的最大值减去最小值。
  - 最小值（Minimum）。
  - 最大值（Maximum）。
  - 平均值的标准差（S.E. mean），S.E.越小，数据的可靠性越大。
- 分布情形（Distribution）：
  - 偏度（Skewness），数据分布的情形，以偏度来看除了正常的正态分布外，有可能是左偏或右偏的数据分布。
  - 峰度（Kurtosis），数据的分布，以峰度来看，除了正常的正态分布外，有可能是高狭峰态分布和低阔峰态分布。

1-5-2 概率分布

在社会科学中常用到随机抽样，其表示抽样是随着某个“概率”而产生的，若是我们可以知道某个概率的分布情形，便能够推算可能的结果，也就可以从样本推算总体，一般我们知道的概率分布，依照随机变量的不同可以分为“离散型概率分布”和“连续型概率分布”，整理如下：

概率分布

- 离散型概率分布：
  - 二项概率分布（Binomial Probability Distribution）：常用于每次测试会有成功或失败，1/2 的概率分布，例如：掷铜板。
  - 超几何概率分布（Hypergeometric Probability）：常用于抽出的样本不放回总体，计算抽出成功的次数分布，例如：乐透彩。
  - 泊松概率分布（Poisson Probability Distribution）：常用于一段时间内，随机发生的概率分布，例如：1 小时内到某家超市消费的人数。
- 连续型概率分布：
  - 正态分布（Normal Distribution）：数量方法中使用最多、最重要的分布，稍后介绍。
  - 均等分布（Uniform Distribution）：用在连续的一段时间内，其事件的分布是平均分布，例如：机械化生产的产品。
  - 指数分布（Exponential Distribution）：用来描述两次事件发生之间的等待时间。

1-5-3 正态分布

正态分布在统计学中，是相当重要的分布，其适用于相当多的自然科学和社会科学环境中，例如：人类的身高和体重大致上都呈现正态分布，其函数数学式如下：

$$f(x)=\frac{1}{\sqrt{2\pi\sigma}}\ e^{-(x-u)^2/2\sigma^2}\quad (-\infty<x<\infty)$$

$\pi = 3.14159$

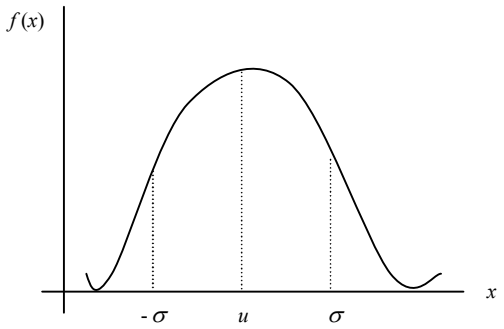
$\sigma =$ 标准差

$e = 2.71828$

$u =$ 平均值

$\infty =$ 无限大

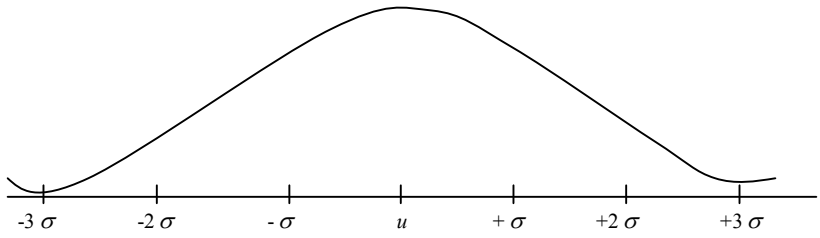
正态分布以图来表示如下：



正态分布的曲线最高点（最大值）是在平均值，平均值可以是正值、负值，也可以是零，图形以平均值为中心会呈现对称分布，整个正态分布。

所涵盖的面积总和等于 1，刚好等于正态随机变量的概率。

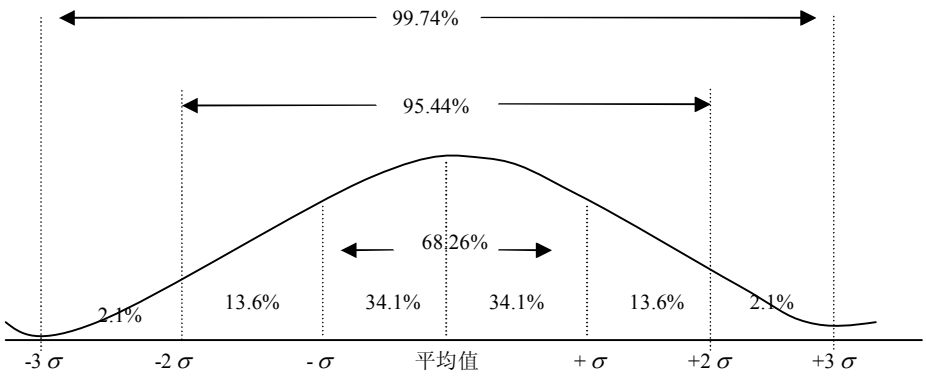
我们常用几个标准差来代表质量的好坏，其意义是指有多少的机会会落在可以控制的范围内，例如：1 个标准差，2 个标准差，3 个标准差，如下图：



u：平均值

σ：标准差

转换成数值后，如下图：



1 个标准差：有 68.26%的概率会落在离平均值±1 个标准差的范围内。

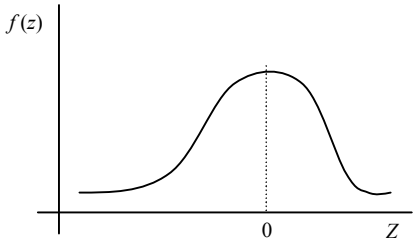
2 个标准差：有 95.44%的概率会落在离平均值 ± 2 个标准差的范围内。

3 个标准差：有 99.74%的概率会落在离平均值 ± 3 个标准差的范围内。

我们通常使用正态分布来进行统计推论，由样本来推论总体，这经常必须假设总体为正态分布下，我们经由样本的抽样分布，例如：*t* 分布、*F* 分布和卡方分布，才可以进行统计估计和假设检验，以推论结果是否如我们所预期的，所以正态分布在统计学中是相当重要的分布。

标准正态分布（*Z* 值表）

标准正态分布就是将正态分布进行标准化，使平均值=0，方差=1，标准差=1，以得到一个 *Z* score 的概率分布，如下图：



标准正态分布曲线下方的面积和为 1，也就是概率和为 1，以平均值 0 为中心，呈现对称分布，所以其概率表可以只给一边，也就是左边或右边概率表。

为什么需要标准正态分布呢？因为不同的正态分布，其平均值和标准差也有所不同，形成不一样的曲线，我们想要得到其区间估计的概率，得使用积分的方式，对于一般人而言，相当繁琐而且也不容易使用，于是研究人员通常会将常态随机变量（例如：A）标准化成标准常态随机变量 *Z*，经由查标准常态概率分布表（*Z* 值分布表）求得概率值。

1-5-4 决定样本数的大小（使用于总体平均值）

样本数的大小会影响我们估计的正确性，当样本越大误差较小，正确性较高，但是调查的成本也随样本数增大，反之，当样本数小，误差较大，正确性低，成本也较低。因此，在作抽样调查时，通常会考虑所需要的成本，另外，还有一个很重要的考虑因素是可容忍的误差，因为误差值直接影响正确性，我们可以通过可容忍的误差值，反推出所需要的样本数，由于抽样误差不超过 e 值，演变至估计总体平均值之样本数如下：

抽样误差 $\overline{X}-u \leqslant 1$

$$Z_{a/2}\frac{\sigma}{\sqrt{n}} \leqslant \mathrm{e}$$

$$\text{平方后 } (Z_{a/2})^2 \frac{\sigma^2}{n} \leqslant e^2$$

$$\text{移项后，样本大小 } n,\ n \geqslant \frac{(Z_{a/2})^2 \sigma^2}{e^2}$$

当总体方差未知时，使用样本方差 S^2 取代总体变异 σ^2

$$\text{总体样本数 } n \geqslant \frac{(Z_{a/2})^2 S^2}{e^2}$$

例如：有一家机械业的制造商，主要生产钻孔的轴承，目前正在生产一批轴承，质检人员随机抽样 16 个轴承，平均大小为 46 厘米，标准差为 5.7 厘米，在 95% 的置信区间，我们可以估算轴承可能的范围大小为 42.96～49.04 厘米，但是顾客要求在 95% 的置信区间时，估计误差不得超过 2 厘米，这时候，我们应该需要多少样本，才可以达到顾客的要求？

$$\text{估计误差 } \leqslant 2$$

$$\overline{X} - u \leqslant 2$$

$$Z_{a/2} \frac{\sigma}{\sqrt{n}} \leqslant 2$$

$$\text{平方，移项后 } n \geqslant \frac{(Z_{a/2})^2 \sigma^2}{2^2}$$

$$n \geqslant \frac{(1.96)^2 (5.7)^2}{2^2}$$

（总体方差未知，使用样本数 S^2 取代，总体方差 σ^2；S = 5.7）

$n \geqslant 31.2$

则 $n = 32$。

我们需要抽样达 32 个样本，才能达到顾客要求在 95% 的置信区间时，估计误差不超过 2 厘米的要求。

1-5-5 中心极限定理

当总体为正态分布时，不论我们的抽样数量是多少，样本平均值的抽样分布都为正态分布，问题是在很多情况下，总体的分布并非是正态分布，应怎样分辨？这时候，就要用到中心极限定理。

中心极限定理：不论总体是否为正态分布，抽样的样本数够大时，则样本平均值的抽样分布会趋近正态分布。

注意：中心极限定理只适用于大样本，至于大样本需要多大，才能称为大样本，则决定于总体的分布情形，一般建议样本≥30 时，样本平均值会趋近于正态分布。

我们整理总体为正态分布和不是正态分布时，样本大小的分布如下：

- 总体为正态分布：

  大样本：样本分布为正态分布；

小样本：样本分布为正态分布。

- 总体不是正态分布：

大样本：样本分布接近正态分布（中心极限定理）；

小样本：决定于总体的分布情形。

1-5-6 *t* 分布

t 分布适用于总体为正态分布，标准差σ未知，小样本的情形下，我们可以使用 *t* 分布求出平均值 *u* 的信赖区间，以便作统计推论，所以说，*t* 分布是使用于总体平均值的分布。

在一般的社会科学中，总体为正态分布，总体的标准差σ为未知，样本大小的处理方式不同，大样本仍呈现正态分布，小样本则不趋近正态分布，而是呈现自由度 *n*−1 的 *t* 分布，估计方式如下：

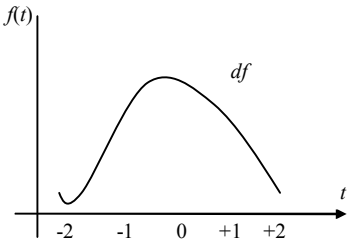
大样本：使用 $Z=\dfrac{\bar{x}-u}{s/\sqrt{n}}$ 做估计

小样本：使用 $t=\dfrac{\bar{x}-u}{s/\sqrt{n}}$（以样本标准差 *S* 代替总体标准差$\sigma$）

t 分布使用参数为自由度（*df*），自由度指的是在统计量中，随机变量可以自由变动的数目，*t* 分布的自由度为 *n*−1（*n* 为样本数）。

t 分布的特性

- *t* 分布是以平均值 0 为中心，呈现对称分布的情形，如下图：



- *t* 分布下的总面积等于 1。
- 当 *df* 趋近于无限大（∞）时，*t* 分布会近似标准正态分布。
- 在一般情形下，样本数≥30 时，我们会以标准正态分布取代 *t* 分布。

1-5-7 卡方分布（X^2 分布）

t 分布使用于总体平均值，而卡方分布则是使用于总体方差。在日常生活中，总体方差或标准差对于我们的生活用品有很大的影响，例如：我们使用的光盘（听音乐、存储数据、看影片），其中心直径大小的圆圈关系着能否播放的问题，也就是说，光盘片中心孔的方差不可以太大，它代表着光盘片的质量，可以利用抽样方式来检验其产品的质量，这时候，总体方差是未知的，我们则可以使用样本方差来进行估计，也就是说，使用卡方分布来推论总体方差。

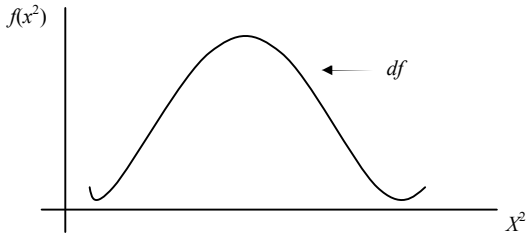
卡方分布的自由度为 *n*-1，自由度指的是在统计量中，随机变异可以自由变动的数目，从样本方差演变至卡方分布，如下：

样本 X_1、X_2、……、X_n，其方差 $S^2=\dfrac{\sum_{i=1}^{n}(X_i-\overline{X})^2}{n-1}$

$$S^2(n-1)=\sum_{i=1}^{n}(X_i-\overline{X})^2$$

$$\frac{S^2(n-1)}{\sigma^2}=\frac{\sum_{i=1}^{n}(X_i-\overline{X})^2}{\sigma^2} \sim \text{趋近 } X^2_{n-1}\text{（卡方分布）}$$

卡方分布会呈现大于等于 0 的分布，如下图：



卡方分布的特性：

- 卡方分布会随着自由度（*df*）的增加，呈现对称分布。
- 卡方分布的 *df* 趋近无限大∞时，可以由正态分布取代卡方分布。
- 卡方分布的平均值等于自由度 *df*，方差等于 2 倍的 *df*（自由度）。

1-5-8 *F* 分布

t 分布使用于总体平均值，卡方分布使用于总体方差，而 *F* 分布则使用于比较两个总体方差的

大小，也就是说，两个总体方差的比较，可以使用 F 分布来作估计和推论，使用的方法是用样本的方差比来推论总体方差是否相等。

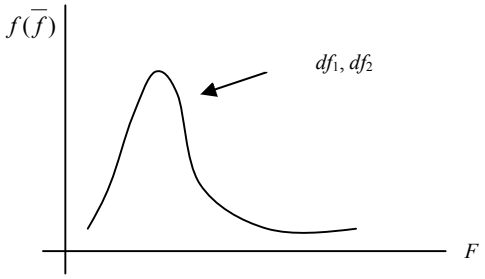
F 分布有个自由度，分别是由分子项的自由度决定 F 分布的第一个自由度，分母项的自由度决定 F 分布的第 2 个自由度，从样本方差演变至 F 分布如下：

$$\text{样本 A 的方差 } S_A^2 = \frac{\sum (X_A - \overline{X_A})^2}{n_A - 1}$$

$$\text{样本 B 的方差 } S_B^2 = \frac{\sum (X_B - \overline{X_B})^2}{n_B - 1}$$

$$\frac{S_A^2 / \sigma_1^2}{S_B^2 / \sigma_2^2} \sim \text{趋近} \quad F_{nA-1,nB-1} \quad (F \text{ 分布})$$

F 分布会呈现大于等于 0 的分布，如下图：



F 分布的特性

- F 分布是由 2 个自由度 df_1 和 df_2 所决定的。
- F 分布的倒数还是呈现 F 分布。

1-5-9 统计估计和假设检验

我们在前面章节介绍过，t 分布用来推论总体平均值 u，卡方（X^2）分布用来推论总体方差 σ^2 或标准差 σ，F 分布用来推论两个以上总体方差 σ^2 的比值或标准差 σ 的比值，我们整理如下表：

分布	推论
t	总体平均值 u
X^2	总体方差 σ^2 或标准差 σ
F	两个以上总体方差 σ^2 的比值或标准差 σ 的比值

- 统计估计

  统计估计是先设定分布的概率值，再推论总体的数值。

  例如：

  设定 t 分布的概率值，可以推论出总体的平均值 u 的值。

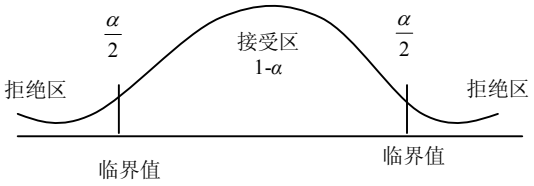
  设定 X^2 分布的概率值，可以推论出总体方差 σ^2 的值或标准差 σ 的值。

  设定 F 分布的概率值，可以推论出总体方差 σ^2 的比值或标准差 σ 的比值。

- 假设检验

  假设检验是先对总体的特性，提出一个假设（hypothesis），再利用抽样取得的样本统计量，来检验（test）总体特性是否符合提出的假设，以拒绝或接受此假设。

  我们进行推论总体时，都会使用到抽样的样本，利用样本的统计量（平均值、方差、标准差）推论总体可能的数量，抽样时，在概率分布图上是一个个的点，因此，称为点估计（Point Estimation），点估计随着样本的不同，变化较大，所以点估计的准确度较容易有问题，于是，将点估计的范围加大，形成一个总体会出现的区间，称之为区间估计（Interval Estimation），使用区间估计时，我们称总体可能会出现的区间为置信区间（Confidence Interval），以 1-α 表示，也就是接受区，α 代表显著水平，在 α 显著水平内的区域，我们称为拒绝区，如下图：



图的中间为置信区间，也就是接受区，两端为拒绝区，临界值是接受区与拒绝区的分界值，若是只有一端拒绝区，则使用 α，两端拒绝区时为 $\frac{\alpha}{2}$，拒绝区在两端时，我们称为双尾检验（Two-tailed Test），拒绝区只出现在右端时，称为右尾检验（Right-tailed Test），拒绝区只出现在左端时，称为左尾检验（Left-tailed Test），无论是使用右尾检验或是左尾检验，我们都称做单尾检验（One-tailed Test）。

在进行检验前，必须先对总体设定好两个假设（hypothesis），一个是零假设（Null Hypothesis），用 H_0 表示，另一个是对立假设（Alternative Hypothesis），用 H_1 表示。零假设是先设定总体的事件为真的，再进行检验，对立假设则是对总体的事件提出不同的假设，而使用样本的统计量做决策（Decision Rule）时，有 2 个情形，一个是会拒绝 H_0，另一个是不会拒绝 H_0，所以有四种组合（2 个假设和 2 个决策），其中有 2 种组合是正确的判断，另外 2 种组合是错误的判断。

使用样本的统计量做决策（Decision Rule）时，有可能产生判断错误，也就是说，总体 H_0 为真，而样本的决策为拒绝 H_0 时，我们称为类型 I 错误（Type I Error），用 α 表示类型 I 错误的概率，

另一个可能产生判断错误的情形，是当总体 H_0 为假（H_1 为真）时，而样本的决策为不拒绝 H_0，称为类型 II 错误（Type II Error），用 β 表示类型 II 错误的概率，整理如下：

		总体实际情形	
		H_0 为真	H_0 为假（H_1 为真）
样本统计量决策	拒绝 H_0	A	1-β
	不拒绝 H_0	1-α	B

我们希望得到情形是总体 H_0 为真，样本统计量决策为不拒绝 H_0，用 1-α 表示正确的概率，另一种情形是总体 H_0 为假（H_1 为真），样本统计量决策为拒绝 H_0，用 1-β 表示正确的概率，但是，偏偏在做假设检验时，类型 I 和类型 II 的判断错误很难避免，因此，我们只能想辨识降低错误的概率，也就是说，α 与 β 越小越好，不幸的是，α 变小时，β 会变大，反之亦然，β 变小时，α 也会变大，因此，我们通常是先固定好 α，决定临界值，再根据假设的情形进行两尾检验、左尾检验或右尾检验。

假设检验的步骤如下：

（1）设定两个假设（零假设和对立假设）。

（2）确认抽样的样本分布（例如：t 分布，X^2 分布），设定显著水平 α，一般设为 0.05，定出接受区和拒绝区。

（3）确定使用双尾、左尾或右尾检验，由显著水平 α、样本数、样本标准差，计算后经查表得到临界值。

（4）比较检验的统计量与临界值的大小，若是落在接受区，则接受零假设 H_0，若是落在拒绝区，则拒绝零假设 H_0，接受对立假设。

（5）由检验的结果——接受虚无假设或拒绝零假设（接受对立假设）进行讨论，并做成结论。

1-5-10 两个总体的估计与检验

在前面章节中，我们讨论的单一总体的统计估计和假设检验，若是要比较两个总体是否有差异，可以使用的方法如下：

- 两个独立总体平均值差的估计和检验——独立样本。
- 两个相关总体平均值差的估计和检验——配对样本。
- 两个总体变异比的估计和检验——F 分布做检验。

我们在进行两个总体的估计与检验时，同样地，需要抽出 2 个样本，以进行统计推论，若是两个样本来自于两个不相关的总体，则称为独立样本，若是两个样本来自于两个相关的总体，则称为配对样本（相依样本）。

1-5-11　三个（含）以上总体的估计与检验——方差分析

在进行三个（含）以上总体的估计与检验时，我们通常是比较三个（含）以上总体的平均值是否相等，也就是说，在比较多个总体平均值之间是否有差异（变异）时，会使用方差分析（Analysis of Variance），一般简称 ANOVA，方差分析除了用来检验三个（含）以上总体的平均值是否相同外，更常用来检验因子对因变量是否有影响，因此，若是进行单一因子对因变量的影响分析，就称为单因子方差分析（One-way Analysis of Variance），若是进行二因子对因变量的影响分析，就称为多变量方差分析（Multivariate Analysis of Variance，MANOVA），方差分析的 ANOVA 和 MANOVA 方法，在后面的章节中会有详细的介绍。

1-6　常用的统计分析

常用的统计分析（多变量分析或称为数量方法）有：

- 方差分析（平均值比较）（Analysis of Variance）
- 因子分析（Factor Analysis）
- 多元回归（Multiple Regression）
- 判别分析（Discriminant Analysis）
- 逻辑回归（Logic Regression）
- 单因子方差分析（Univariate Analysis of Variance，ANOVA）
- 多变量方差分析（Multivariate Analysis of Variance，MANOVA）
- 典型相关分析（Canonical Correlation Analysis）
- 联合分析（Conjoint Analysis）
- 结构方程模型（Structure Equation Model）

（详细内容，请参考后面各章节）

1-6-1　方差分析

方差分析（Analysis of Variance）适用于因变量是计量，自变量是非计量，方差分析可以分为单变量和多变量差异数分析，单一因变量（也称为单一准则变量）的计算，我们称为单变量方差分析（Univeriate Analysis of Variance，ANOVA），如下图：

$$Y = X_1 + X_2 + X_3 + \cdots + X_K$$

（计量）　（非计量，例如：名目）

多个因变量（也称为多个准则变量）的计算，我们称为多变量的方差分析（Multivariate Analysis of Variance，MANOVA），如下图：

$$Y_1+Y_2+Y_3+\cdots+Y_i = X_1+X_2+X_3+\cdots+X_k$$

（计量）　　（非计量，例如：名目）

方差分析的目的是要发掘多个类别的自变量对于单一或多个因变量的影响（之间的关系），检验的方式是比较平均值是否有显著的差异。

1-6-2 因子分析

因子分析（Factor Analysis）包含有：

- 主成分分析（Principal Component Analysis）
- 一般因子分析（Common Factor Analysis）

主成分分析是由 Pearson 于 1901 年发明，在 1993 年时由 Hotelling 加以发展和推广的分析方法。因子分析的目的是在压缩原始的一堆变量，形成较少的代表性变量，而且，这些代表性的变量具有最小的信息损失和保有最多原变量的信息（最大的方差）。简单地说，我们常用因子分析来去除不重要的变量，以形成少数的构念（dimensions），这些构念可以用来形成研究构念，或形成加总的尺度（Summated Scales），以方便后续的统计技术分析。

1-6-3 多元回归

多元回归（Multiple Regression）也称为复回归，适用于因变量和自变量都是计量的情形，如下图：

$$Y=X_1+X_2+X_3+\cdots+X_k$$

（计量）　　（计量）

多元回归的目的是用来预测当自变量 X 改变时，因变量 Y 会改变多少，计算的方式通常是使用最小平方法（Lease Square）来达成。

1-6-4 判别分析

判别分析（Discriminate Analysis）：在已知的样本分类，建立判别标准（判别函数），以判定新样本应归类于哪一群中。

判别分析适用于因变量是非计量，自变量是计量的情形，如下图：

$$Y = X_1+X_2+X_3+\cdots+X_k$$

（非计量，例如：名目）　　（计量）

判别分析的因变量最好是可以分为几组，在单一因变量下，可以分为 2 分法的性别（男生和女生）、多分法的薪资（高、中和低收入），判别分析的目的是要了解组别的差异和找到判别函数，用来判定单一受测者应该是归于哪一个的组别或群体。

1-6-5　逻辑回归

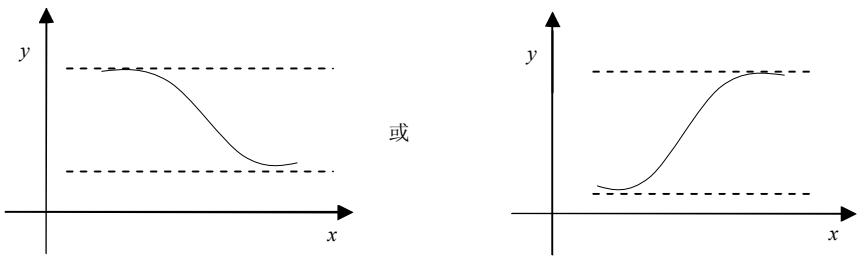
逻辑回归（Logic Regression）适用于因变量（Dependent Variable）为名义二分变量，自变量（Independent variable）为连续变量的情形，如下图：

$$Y = X_1 + X_2 + X_3 + \cdots$$

（名义二分变量）　　（连续变量）

- 逻辑回归、多元回归和判别分析之比较：

  逻辑回归和多元回归的差别是多元回归必须数据符合正态性分布，常用普通最小平方法（Ordinary Least Square）进行估计，而逻辑回归则是数据必须呈现 S 型的概率分布，也称为 Logic 分布，常用最大似然法（Maximum Likelihood Estimate，MLE）进行估计，如下图：



多元回归的因变量和自变量都是连续性的变量，逻辑回归的因变量是名义二分变量，自变量是连续变量。逻辑回归和判别分析的差异是，判别分析需要符合方差（variance）、协方差（covariance）相等，而逻辑回归较不受方差、协方差影响（Hair，1998），但是逻辑回归需要符合的是 S 型的 Logic 分布，逻辑回归和判别分析相同的是因变量为名义二分变量，自变量是连续变量。

1-6-6　单因子方差分析

自变量只有一个的方差分析，称为单因子方差分析（Univariate Analysis of Variance，ANOVA），

也就是 $y_1+y_2+\cdots=x$（y 可以是一个（含）以上，x 只有 1 个）

单因子方差分析有 2 种设计方式：独立样本和配对样本。

1．独立样本

受测者随机分派至不同组别，各组别的受测者没有任何关系，也称为完全随机化设计。

（1）各组人数相同：HSD 法，Newman-Keals 法。

（2）各组人数不同（或每次比较 2 个以上平均值时）：Scheffe 法。

2．配对样本：有两种情形

（1）重复量数：同一组受测者，重复接受多次（k）的测试以比较之间的差异。

（2）配对组法：选择一个与因变量有关控制配对条件完全相同，以比较 k 组受测者在因变量的差异。

方差分析的基本假设条件：

- 正态：直方图，偏度（skewness）和峰度（kurtosis），检验，改正（非正态可以透过数据转型来改正）。
- 线性：变量的散布图，检验，简单回归+ residual
- 方差同构型：1y，用 Levene 检验

  ≥2y 时，用 Box's M 检验

1-6-7 多变量方差分析

多变量方差分析（Multivariate Analysis of Variance，MANOVA）是 ANOVA 的延伸使用，用来作多个母群平均值比较的统计方法。

MANOVA 的 3 个基本假设与 Anova 相同都是共变量分析的基本假设有：

- 正态：直方图，偏度（skewness）和峰度（kurtosis），检验，改正（非正态可以透过数据转型来改正）
- 线性：变量的散布图，检验，简单回归+ residual
- 方差同构型：1y，用 Levene 检验

  ≥2y 时，用 Box's M 检验

MANOVA 可以指定两个或两个以上因变量的方差和共变量分析（针对单一因变量的方差分析，请使用 ANOVA），MANOVA 也可以分别对每个因变量进行检验（如同 ANOVA），问题是分开的个别检验无法处理因变量间的复（多个）共线性（Multicollinearity）问题，必须使用 MANOVA 才能处理。

1-6-8 典型相关

典型相关（Canonical Correlation）适用于依变数为计量或非计量，自变量也是计量或非计量的

情形，如下图：

$$Y_1+Y_2+Y_3+\cdots+Y_j = X_1+X_2+X_3+\cdots+X_k$$
（计量，非计量）　　（计量，非计量）

- 典型相关（Canonical Correlation）和回归（Regression）的不同：
  典型相关分析可以视为多元回归的延伸使用，多元回归的因变量（Y）只有一个，自变量（X）有多个，典型相关则是可以处理多个因变量和多个自变量。
  典型相关分析的目的是要找出因变量的线性组合和自变量的线性组合，这两个线性组合相关最大化，简单地说，就是找出 Y 这一组的线性组合、X 这一组的线性组合，这两个线性组合的最大化，换句话说，典型相关分析就是要求得一组的权重（Weight）以最大化因变量和自变量的相关。
- 典型相关和主成分的不同：
  主成分分析是处理一组变量内最大的萃取量，典型相关则是处理两组变量的关系最大化。
  适用于典型相关分析的数据是 2 组的变量，这二组变量拥有理论上的支持，一组为因变量，一组为自变量，经由分析所得到的典型相关，就可以用在很多的地方，因此，典型相关的目的，可以有下列几项：
  - 决定二组变量的关系强度。
  - 计算出因变量和自变量在线性关系最大化下的权重（Weight），另外的线性函数则会最大化，剩余的相关，并且和前面的线性组合是相互独立的。
  - 用来解释因变量和自变量关系存在的本质。

1-6-9　联合分析

联合分析（Conjoint Analysis）适用于因变量是计量或顺序，自变量是非计量，如下：

$$Y = X_1+X_2+X_3+\cdots+X_k$$
（计量或非计量）　（非计量，例如：名目）

联合分析是分析因子的效果，其目的是将受测者对受测体的整体评价予以分解，通过整体评价求出受测体因子的效用。

联合分析特别适用于了解客户的需求，针对新的产品或服务，我们可以将新的产品或服务分解成项组合，例如：手机分解成：品牌（2 种）、形状（2 种）和价格（3 种），如此一来，总共有 2*2*3=12 种组合，客户对这 12 种组合给予分数，最后再依据客户的整体评价求出各个组合的效用，以了解客户对于新产品的喜好。

1-6-10 结构方程模型

结构方程模型（Structural Equation Modeling，SEM）是一种统计的方法学，早期的发展与心理计量学和经济计量学息息相关，之后，逐渐受到社会学的重视，因为结构方程模型除了结合了因子分析和路径分析两大统计技术外，更是多用途的多变量分析技术。

在前面章节介绍的多变量分析技术，大都是处理单一关系的因变量和自变量，而 SEM 则是可以处理一组（二个或二个以上）关系的因变量和自变量，数学方程式如下：

$$Y_1 = X_{11} + X_{12} + \cdots + X_{1j}$$
$$Y_2 = X_{21} + X_{22} + \cdots + X_{2j}$$
$$\cdots\cdots$$
$$Y_i = X_{i1} + X_{i2} + \cdots + X_{ij}$$

（计量）　　（计量，非计量）

SEM 常用来指定和估计变量们的线性关系模型，结构方程模型也常用在因果模型、因果分析、同时间的方程模型、共变结构的分析、潜在变量路径分析和验证性的因子分析，研究中，SEM 可以用来处理相关的（可观察到的）变量或实验的变量，在一般情况下，大都使用在相关的变量。SEM 在相关的研究设计中，会使用横截面的（Cross-sectional designs）研究设计和长时间的（Longitudinal designs）研究设计。横截面的（Cross-sectional designs）研究设计，简单地说，就是取得一次的数据，例如：我们最常使用的方式，就是发一次问卷；而长时间的研究设计至少需要取得三次的数据，例如：对于相同的受测者，依时间的不同，发出了三次问卷。在应用方面，SEM 结构方程模型发展至今，已经应用到各种领域，列举如下：企业管理、信息管理、人力资源管理、健康医疗、社会学、心理学、经济学、宗教的研究、国际营销、消费者行为、通路的管理、广告、定价策略、满意度的调查。从以上的数据显示出 SEM 已经逐渐深入许多领域的研究了。

1-6-11 简易数量方法的记忆

口诀：ANOVA MANOVA 联合起来去区别典型的多元回归和 SEM（结构方程模型）。

我们整理如下表：

	因变量 (*y*)	自变量 (*x*)
ANOVA 单因子方差分析	计量	名目
MANOVA 多变量方差分析	计量	名目
联合分析（Conjoint Analysis）	计量，顺序	名目
判别分析（Discriminate Analysis）	名目	计量
典型相关分析（Canonical Correlation Analysis）	计量	计量

续表

	因变量（*y*）	自变量（*x*）
多元回归（Multiple Regression）	计量	计量
结构方程模型（Structure Equation Model）	计量	计量，非计量

在整个社会科学的研究中，牵动到统计分析的部分相当多，本章节已经分别介绍统计分析简介与数量方法的基础，包括理论简介、量表简介、抽样简介、基础统计学和常用的统计分析（多元分析或称为数量方法）。