CH1--[學習筆記] 統計學：基本概念 Statistics

統計學 (statistics) 是一種由資料 (data) 萃取出資訊 (information) 的方法。

細分為：描述統計學 (descriptive statistics)、推論統計學 (inferential statistics)。

描述統計學 (descriptive statistics)將資料加以組織、總結、呈現，使其令人容易瞭解其中所蘊含的意義。

細分為：圖形技巧 (graphical techniques)、數值技巧 (numerical techniques)。

推論統計學 (inferential statistics)根據樣本資料 (sample)，推論母群體 (population) 的特性。

統計學主要概念母群體 (population)

想要研究的所有對象所組成的集合。

樣本 (sample)

母群體 (population) 的子集合。

參數 (parameter, 母數)

關於母群體的描述性量測(descriptive measurement)。

舉例，對於母群體的資料，求得的算術平均數、標準差是一種參數。

統計量 (statistic)

關於樣本的描述性量測(descriptive measurement)。

舉例，對於樣本資料，求得的算術平均數、標準差是一種統計量。

統計推論 (statistical inference)

根據樣本資料 (sample)，估計、預測、決定母群體 (population) 的特性。

因為樣本資料的資訊小於母群體的資訊，這樣的推論無法完全準確。

推論的可靠度量度包含信心水平 (confidence level)、顯著性差異 (signifance level)。

統計推論有2種作法:

(1) 估計 (estimation)

(2) 假設檢定 (hypothesis testing)。

CH3--[學習筆記] 統計學：數值描述技巧 Numerical Descriptive Techniques

中心位置量數 (Measures of Central Location)

算數平均數 (Arithmetic Mean, Average, or Mean)

Population Mean

μ: 唸作mu.

xi: 母群體的第i筆資料.

N: 母群體的資料個數 (population size).

Sample Mean

xi: 取樣的第i筆資料.

n: 取樣的資料個數 (sample size).

算數平均數適用於區間資料。

中位數 (Median)

把資料加以排序，落在中間的數值，即為中位數。

若有偶數筆資料，則中位數是落在中間的2筆資料的平均。

中位數的意義在於有1半的資料<中位數；另1半的資料>中位數。

和算數平均數相比，中位數的好處是，不會受到極端值的影響。

舉例來說，有5名同學成績排序為: 50, 60, 70, 80, 90，則中位數為70。有1半的同學<70分；另1半的同學>70分。

中位數適用於區間資料、序位資料。

眾數 (Mode)

資料中出現次數最多的數值，即為眾數。

眾數可能不只一個。

眾數適用於區間資料、序位資料、類別資料。

幾何平均數 (Geometric Mean)

適用於找出隨著時間變化的變數的成長率或改變的速率。

幾何平均數適用於區間資料。

變異量數 (Measures of Variability)適用於區間資料。

全距 (Range)

Range = 資料中的最大值 - 資料中的最小值

優點: 簡單.

缺點: 只考慮到資料中的2筆數值，包含的資訊有限。

變異數 (Variance)

Population Variance

σ2: 唸作sigma squared.

Sample Variance (corrected for the mean)

算式中，xi-μ取平方的原因，是因為我們想觀察的是距離，需要避免正負數值互相抵銷。

變異數的單位是資料的單位的平方。舉例來說，若資料的單位是秒，變異數的單位則是秒2。

變異數適用於比較相同型態變數的兩組以上的資料。

標準差 (Standard Deviation)

Population Standard Deviation

Sample Standard Deviation

標準差的單位和資料的單位相同。

根據 Chebysheff's Theorem，資料中至少有比例為 1 - (1/k2) 的觀測落在距離算術平均數的k個標準差之內 (其中k>1)：

1. 至少有75%的觀測落在距離算術平均數的2個標準差之內。(k=2)

2. 至少有88.9%的觀測落在距離算術平均數的3個標準差之內。(k=3)

如果資料的histogram是鐘形 (bell shaped)，可以使用下列經驗法則(Empirical Rule)來解讀標準差的意涵：

1. 大約68%的觀測落在距離算術平均數的1個標準差之內。

2. 大約95%的觀測落在距離算術平均數的2個標準差之內。

3. 大約99.7%的觀測落在距離算術平均數的3個標準差之內。

變易係數 (Coefficient of Variation)

Population Coefficient of Variation

Sample Coefficient of Variation

相對位置量數 (Measures of Relative Standing)適用於區間資料、序位資料。

百分位數 (Percentile)

第P百分位數(Pth percentile): 是一個數值，其中P%的資料<該數值，而(100-P)%的資料>該數值。

第25百分位數，又稱為第1四分位數(first quartile, Q1)。

第50百分位數(50th percentile)就是中位數(median)，又稱為第2四分位數(second quartile, Q2)。

第75百分位數，又稱為第3四分位數(third quartile, Q3)。

百分位數的近似位置(Location of a Percentile)

LP: the location of the Pth percentile

四分位距 (Interquartile Range)

四分位距 = Q3 - Q1

線性關係量數 (Measures of Linear Relationship)描述2個變數之間的關聯性，適用於區間資料。

共變異數 (Covariance)

Population Covariance

Sample Covariance

一般來說，當2個變數移動的方向相同，共變異數會是較大的正數；

當2個變數移動的方向相反，共變異數會是較大的負數；

當2個變數關聯性較低，共變異數會是較小的數值。

相關係數 (Coefficient of Correlation)

Population Coefficient of Correlation

ρ: 唸作rho.

-1 <= ρ <= +1

Sample Coefficient of Correlation

-1 <= r <= +1

當相關係數接近+1，表示2個變數之間有正向的線性關係，散佈圖(scatter diagram)呈現接近正斜率的直線；

當相關係數接近-1，表示2個變數之間有負向的線性關係，散佈圖(scatter diagram)呈現接近負斜率的直線；

當相關係數接近0，表示2個變數之間沒有線性關係。

有相關性不代表有因果性(Correlation does not imply causation)。

最小平方法 (Least Squares Method)

對於變數X, Y，可得出一條直線，其中下列數值是最小的：

該直線方程式如下：

決定係數 (Coefficient of Determination)

決定係數 R2 = 相關係數的平方

決定係數的意函是相依變數(dependent variable)中多少比例的變異量(variation)和獨立變數(independent variable)有關。

舉例來說，R2=.88，表示88%的變異量和獨立變數有關。

CH4--[學習筆記] 統計學：資料搜集及取樣 Sampling

資料搜集的方式 Methods of Collecting Data

直接觀測 Direct Observation

可能只能找到相關性，而非因果關係。

實驗 Experiments

設計周延的實驗，會比直接觀測，更能真正找到因果關係。

調查 Surveys

回應率若太低，統計分析結果的可信度會大打折扣。

面談 Personal Interview

回應率 (response rate)較高，但成本較高。

電話訪談 Telephone Interview

回應率較低，但成本較低。

問卷調查 Self-Administered Survey

回應率最低，且因為可能誤解問題使得錯誤的回應比例較高。

取樣 Sampling

目標母群體 (target population): 想要進行推論的母群體(population)。

舉例來說，電台的聽眾。

樣本 (sampled population): 母群體中實際被取樣的集合。

舉例來說，參與call-in民調的電台聽眾。

自我選擇取樣 (self-selected sampling)

舉例來說，參與call-in民調的電台聽眾，因為對於該議題感興趣而參與調查，這是聽眾的自我選擇。

這種取樣方式幾乎總是有偏差的。

取樣計劃 Sampling Plans

簡易隨機取樣 Simple Random Sampling

每個樣本被選取的機率是相同的。

分層隨機取樣 Stratified Random Sampling

先對母群體加以分層，然後對於每一層 (stratum) 進行簡易隨機取樣。

分層方式，舉例來說，對於人口資料，以性別區分、以年齡區分、以職業區分、以收入區分。

好處是可以獲得比較多的資訊。舉例來說，不同收入層級的人對於增稅的想法，因為增稅對於不同收入層級的人有不同的影響。

群集取樣 Cluster Sampling

對於群集進行簡易隨機取樣。

舉例來說，要得知台北市的平均家庭年收入，可以將每一區（如中山區）視為一群集，然後對每一群集，進行簡易隨機取樣。

成本較低，但誤差較大。

取樣誤差 Sampling Errors

統計量和參數之間的差距。

降低取樣誤差唯一的方式是增加樣本數(sample size)。

非取樣誤差 Nonsampling Errors

資料取得誤差 (errors in data acquisition)

取得錯誤的資料、取得資料的過程中出錯。

無回應誤差 (nonresponse error)

因為取樣的對象沒有回應，導致的偏誤。

選擇偏差 (selection bias)

取樣計劃中，某些成員並沒有被包含進去。

舉例來說，在電台節目中使用call-in進行取樣，沒有電話的人，或沒有收聽廣播的人，並沒有被包含進來。

CH5--[學習筆記] 統計學：機率 Probability

隨機實驗 Random Experiment

會導致得到幾種可能的結果 (outcome) 之一的行動或程序。

舉例：擲銅板，可能的結果：正面、反面。

樣本空間 Sample Space

隨機實驗的樣本空間包含所有可能的結果。這些結果必須是 exhaustive 及 mutually exclusive。

樣本空間表示為 S = {O1, O2, ..., Ok}

舉例：擲骰子的樣本空間 S = {1, 2, 3, 4, 5, 6}

機率的要求 Requirement of Probabilities

1. 0<=P(Oi)<=1, 其中 1<=i<=k, P(Oi)代表outcome i發生的機率。

2. P(O1) + P(O2) + ... + P(Ok) = 1

簡單事件 Simple Event

樣本空間內的一個結果。

事件 Event

樣本空間內的一個或多個簡單事件的集合。

事件的機率 Probability of Events

事件的機率等於「構成事件的簡單事件的機率的總和」。

舉例：擲骰子，點數>3的機率 = 點數為4的機率 + 點數為5的機率 + 點數為6的機率。

事件A和B的交集 Intersection of Events A and B

事件A和事件B都發生的情況，表示為「A and B」。

聯合機率 (joint probability)

兩個以上的事件都發生的機率，稱為聯合機率。

邊際機率 (marginal probability)

在有兩個以上的事件的樣本空間中，若僅考慮某一事件個別發生的機率，稱為邊際機率。

舉例：

其中 P(A1), P(A2), P(B1), P(B2) 為邊際機率。

條件機率 (conditional probability)

在事件B發生的條件下，事件A發生的機率。表示為 P(A|B)，其中「|」唸作given。

P(A|B) = P(A and B) / P(B)

可從集合的概念來理解。

以上例而言，當車主為白領的條件下，購買進口車的機率，P(B2|A2) = P(B2 and A2) / P(A2) = 0.3 / 0.4 = 0.75

獨立事件 Independent Events

若 P(A|B) = P(A) 或 P(B|A) = P(B)，則事件A和事件B稱為獨立事件。

事件A和B的聯集 Union of Events A and B

事件A或事件B發生的情況，表示為「A or B」。

互補法則 Complement Rule

P(Ac) = 1 - P(A)

其中, P(A): 事件A發生的機率, P(Ac): 事件A不發生的機率。

乘法法則 Multiplication Rule

P(A and B) = P(B)P(A|B)

P(A and B) = P(A)P(B|A)

可由條件機率推導出來。

獨立事件的乘法法則 Multiplication Rule for Independent Events

P(A and B) = P(A)P(B)

由乘法法則，加上獨立事件的定義，就可以很容易推導出來。

加法法則 Addition Rule

P(A or B) = P(A) + P(B) - P(A and B)

以集合的概念來看(畫有交集2個圈圈分別代表A及B集合)，就可以很容易瞭解。

互斥事件的加法法則 Addition Rule for Mutually Exclusive Events

P(A or B) = P(A) + P(B)

因為互斥事件沒有交集 P(A and B) = 0

CH6--[學習筆記] 統計學：隨機變數及離散型機率分佈 Random Variables and Discrete Probability Distributions

隨機變數 Random Variable

為隨機實驗 (random experiment) 的每一個結果 (outcome) 指定相對應的數字的函數或規則。

舉例來說，擲兩顆骰子，我們可以定義隨機變數 X 是兩顆骰子的點數的和，X 的值域為 2, 3, 4, ..., 12。

隨機變數(X)的某個值(x)發生的機率表示為P(X=x)或P(x)。

離散型隨機變數 Discrete Random Variable

值域為「有限」，或「無限且與自然數有一對一的對應」。

連續型隨機變數 Continuous Random Variable

值域為「某一區間」或「區間的集合」內的所有數值。

機率分佈 Probability Distribution

描述隨機變數的「值」及其「相對應的機率」的表格、公式或圖。

離散型隨機變數的機率分佈的要求 Requirements for a Distribution of a Discrete Random Variable

1. 0 <= P(x) <= 1 for all x

2.

母體平均, 又稱為期望值 Population Mean, or Expected Value

母體變異數 Population Variance

期望值定律 Laws of Expected Value

1. E(c) = c

2. E(X + c) = E(X) + c

3. E(cX) = cE(X)

變異數定律 Laws of Variance

1. V(c) = 0

2. V(X + c) = V(X)

3. V(cX) = c sup2 V(X)

雙隨機變數的機率分佈 Bivariate Distributions

離散型雙隨機變數的機率分佈的要求 Requirements for a Discrete Bivariate Distribution

1. 0 <= P(x,y) <= 1 for all pairs of values (x,y)

2.

共變異數 Covariance

雙隨機變數的和的期望值定律 Laws of Expected Value of the Sum of Two Variables

E(X + Y) = E(X) + E(Y)

雙隨機變數的和的變異數定律 Laws of Variance of the Sum of Two Variables

V(X + Y) = V(X) + V(Y) + 2COV(X,Y)

二項式實驗 Binomial Experiment1. 包含固定次數(表示為 n)的實驗。

2. 每次的實驗都有可能出現兩種結果。標示為成功或失敗。

3. 成功的機率是p；失敗的機率是1-p。

4. 每次的實驗彼此是獨立的。也就是說，某次實驗的結果，不會影響到其他次實驗的結果。

舉例，擲銅板100次。

舉例，在只有2位候選人的情況下，調查5000位選名會投給哪位候選人。

若滿足上述2,3,4的條件，則稱每次的實驗為白努力程序(Bernoulli process)。

二項式隨機變數 Binomial Random Variable

n次二項式實驗中成功的次數。

二項式機率分佈 Binomial Probability Distribution

在實驗n次的二項式實驗中，若成功的機率為p，成功x次的機率為

for x=0,1,2,...,n

二項式隨機變數的算術平均數 Mean of a Binomial Random Variable

二項式隨機變數的變異數 Variance of a Binomial Random Variable

Poisson 實驗1. 在任意區間(interval)內成功的次數和在其他區間內成功的次數是彼此獨立的。

2. 對於所有相同大小的區間，在任一區間內成功的機率都相同。

3. 在某一區間內成功的機率和該區間的大小成正比。

4. 當區間變小，超過1次成功的機率會趨近於0。

舉例，1小時內抵達公車站的車次。

舉例，1天內在某個區段的高速公路發生事故的次數(同時包含時間區間及空間區間)。

Poisson 隨機變數

在 Poisson 實驗中，在指定的區間內成功的次數。

Poisson 機率分佈 Poisson Probability Distribution

x=0,1,2,...

μ: 在區間中成功次數的平均值

e: 自然對數的基底 (2.71828)

Poisson 隨機變數的變異數 Variance of a Poisson Random Variable

CH7--[學習筆記] 統計學：連續型機率分佈 Continuous Probability Distributions

機率密度函式 Probability Density Function

機率密度函式的要求 Requirements for a Probability Density Function

1. f(x) >= 0 for all x between a and b

2. f(x)曲線在 a 和 b 之間的總面積 = 1.0

其中 f(x) 為一機率密度函式，其定義域為 a<=x<=b

對於連續型隨機變數，在單一點發生的機率是0，因此只能估計在一區間內的機率，而在一區間內的機率，就是機率密度函式在該區間內的面積。

均勻分佈 Uniform Distribution

均勻機率密度函式 Uniform Probability Density Function

其中 a<=x<=b

常態分佈 Normal Distribution

常態機率密度函式 Normal Density Function

其中

-∞ < x < ∞

e = 2.71828...

π = 3.14159...

Excel 函數

若要計算累積常態機率(cumulative normal probabilities) P(X<x)，Excel 函數為 NORMDIST(x, μ, σ,

True)。

其中 True 代表要計算 cumulative normal probability, False 代表要計算 normal density function。

若要計算x，使得 P(X<x) = A，Excel 函數為 NORMINV(A, μ, σ)。

舉例，一投資標的之報酬率為常態分佈，平均值為10%，標準差為5%。

賠錢的機率為

其中Z為標準常態隨機變數(standard normal random variable, μ=0, σ=1)。

CH8--[學習筆記] 取樣分佈 Sampling Distributions

取樣分佈 Sampling Distributions對於特定樣本數 (sample size) N，所有可能的取樣之統計量（舉例來說，算術平均數）的機率分佈。

算術平均數的取樣分佈　Sampling Distribution of the Mean

若母群(X)為常態分佈(算術平均數為μ，標準差為σ)，

則對於任意樣本數，所有可能的取樣的算術平均數()亦為常態分佈，

且算術平均數為μ，標準差為。

可以看出來，若樣本數愈大，算術平均數的變異數及標準差會越小。

舉例來說，若台灣人口(設2000萬人)的體重呈常態分佈，則任意調查N個人，所有可能的取樣的體重的算術平均數，也是呈常態分佈。

若N=1，所有可能的取樣為2000萬筆體重資料。

中央極限定理 Central Limit Theorem

從任何一個母群(算術平均數為μ，標準差為σ)中取大小為N之樣本，

當N足夠大時，取樣的算術平均數會接近常態分佈，

且算術平均數為μ，標準差為。

一般來說，N=30，可視為「足夠大」。是否足夠大，視母群的分佈近似常態分佈的程度。

舉例來說，電機系主任宣稱電機系畢業生的月薪平均為NT$80,000，標準差為$10,000。

某甲想要驗證這個說法的真實性，調查了25位電機系畢業生，發現月薪平均為NT$75,000。

某甲想要瞭解，樣本數為25，月薪平均<=NT$75,000的機率。

X: 母群的月薪。

: 所有可能的取樣的算術平均數，其中樣本數為25。

我們想計算的是 P(eq1 < 75,000)

機率很低，所以可以認為電機系主任的說法很可能是不符合事實的。

CH9--[學習筆記] 統計學：估計 Estimation

估計 Estimation根據樣本的統計量，決定母群體的參數的近似值。

舉例來說，根據樣本的算術平均數，估計母群體的算術平均數。

估計式: estimator。舉例，樣本的算術平均數。

估計量: estimate。

點估計式 Point Estimator

使用單一的值(單點)來估計未知的母群體參數，以導出對於母群體的推論。

舉例，為了瞭解台北市民的平均月收入(假設有260萬人)，挑選其中1000人，計算月收入的算術平均數，假設為60000，

點估計會推論台北市民的平均月收入為60000。

缺點:

1. 估計完全正確的可能性是0。因為連續型隨機變數在任何一點發生的機率是0。

2. 無法知道估計量和參數有多接近。

3. 當樣本數愈大，估計量預計應該會愈準確，但點估計無法反應這個特性。

區間估計式 Interval Estimator

使用區間來估計未知的母群體參數，以導出對於母群體的推論。

舉例，為了瞭解台北市民的平均月收入(假設有260萬人)，挑選其中1000人，紀錄月收入的分佈，區間估計會推論台北市民的平均月收入在50000到70000之間。

不偏估計 Unbiased Estimator

若一估計式的期望值等於母群體參數，則該估計式稱為不偏估計式。

意為，若對母群體進行無限次數的取樣，求得的估計值，會等於母群體參數。

舉例來說，樣本的算術平均數是母群體算術平均數的不偏估計。

一致性 Consistency

隨著樣本數變大，若估計量和母群體參數的差異隨之變小，則稱該不偏估計具有一致性。

前述差異的量測值為變異數（或標準差）。

舉例來說，樣本的算術平均數()是母群體的算術平均數(μ)的一致性估計式，因為的標準差為，當樣本數n愈大，的變異數變小，有愈多的樣本算術平均數會接近μ。

相對有效性 Relative Efficiency

對於2個不偏估計，變異數較小的稱為相對較有效 (relatively more efficient)。

舉例來說，對於母群體算術平均數的估計，樣本算術平均數比樣本中位數相對有效。

區間估計

母群體標準差為已知時，估計母群體算術平均數的方式

這個例子雖然不實際，但因為很簡單，有助於瞭解區間估計的概念。

假設母群體算術平均數為μ，標準差為σ。

假設母群體算術平均數是未知，而標準差是已知，我們想要估計母群體的算術平均數。

我們會隨機取n個樣本，並計算樣本算術平均數為 。

根據中央極限定理，如果X是常態分佈，或如果X不是常態分佈但n足夠大，也會是常態分佈。

因此隨機變數Z

會是標準常態分佈。

可推導出

最後一個式子意義為，當重複從母群體中進行取樣， , 區間包含母群體算術平均數μ的機率是1-α。

其中Zα的意義為某一個z值，使得標準常態分佈曲線在該z值右方的面積為α，換句話說，P(Z>Zα) = α

1-α稱為信賴水準(confidence level)。

舉例來說，信賴水準為95%，意義為，若取樣多次(每次取樣數為n)，其中95%的取樣，會使得取樣的算術平均數落在信賴區間內。但實際上，通常只會取樣一次，意義為該次取樣的算術平均數落在信賴區間內的機率是95%。

稱為下信賴界限(lower confidence limit, LCL).

稱為上信賴界限(upper confidence limit, UCL).

通常信賴區間估計式(confidence interal estimator)表示為

估計誤差

估計量和母群體參數之間的差距。

根據前述的例子，可推導為下式，其中-μ是估計誤差。

也就是說， 的機率是1-α

[學習筆記] 統計學：假設檢定 Hypothesis Testing

假設檢定 Hypothesis Testing

以統計方法進行決策的過程中，會提出兩個假設：

H0: null hypothesis (虛無假設)。

H1: alternative or research hypothesis (對立假設、研究假設)。

把想要檢定的假設定為 H1，H0 則為其相反之假設。

首先，假設 null hypothesis 為真。據此進行推論。

可能的結論：

(i) 有足夠的統計證據可推論 alternative hypohesis 為真 (rejecting the null hypothesis in favor of the alternative)。

(ii) 沒有足夠的統計證據可推論 alternative hypohesis 為真 (not rejecting the null hypothesis in favor of the alternative)。

假設檢定可能犯的錯誤：

Type I error (第一型錯誤): reject a true null hypothesis. P(Type I error) = α. α 又稱為 significance level (顯著水準)。

Type II error (第二型錯誤): don't reject a false null hypothesis. P(Type II error) = β.

已知母群體標準差檢定母群體數算平均數

從範例較容易瞭解假設檢定的概念，舉例如下，便利商店經理根據財務分析，認為若顧客平均每次消費金額高於$170，發行NFC卡將可以獲利。

假設每次消費金額是常態分佈，標準差為$65。

現以400人進行取樣，發現樣本算術平均數(sample mean)為$178。

便利商店經理是否能夠推論發行NFC卡可以獲利？

我們想要檢定的假設是

H1: μ>170

因此，null hypothesis為

H0: μ<=170

但若我們設定 H0: μ=170，也可以達到和上式相同的結論，而這樣的設定的好處是我們可以直接以母群體算術平均數為μ來進行計算，因此實際上我們設定的null

hypothesis會只取等式的部份，以此例為

H0: μ=170

主要有2種方式可以進行假設檢定:

1. rejection region method

2. p-value approach

Rejection Region

若檢定統計量(test statistic，舉例來說，樣本算術平均數)落在 rejection region，我們會決定 reject the null hypothesis in favor of the alternative。

見上圖，以此例而言，rejection region為，其中為樣本算術平均數。

根據 Type I error 的定義，可推導出

α = P(rejecting H0 given that H0 is true)

= P( given that H0 is true)

如果便利商店經理設定 α 為 5%，則 zα = 1.645，因此

因此 rejection region 為 > 175.34

因為取樣得到的樣本算術平均數是178，落在 rejection region，我們 reject the null hypothesis，有足夠的證據可推論 alternative hypohesis: μ>170 為真。

p-Value Approach

p-value是在假設 null hypothesis 為 true 的前提下，觀察到檢定統計量 (test statistic) 比取樣得到的值更極端的機率。

以此例而言，

根據取樣分佈，當母群體算術平均數(population mean)為170時，我們觀察到樣本算術平均數大於178的機率是0.0069，因為這樣的機率很低，我們懷疑假設 null hypothesis 為 true

的前提，因此我們 reject the null hypothesis，而推論 alternative hypothesis 為真。

p-value要多小，才適合推論 alternative hypothesis為真？

這取決於犯下Type I 及 Type II錯誤的成本，若成本很高，會需要較低的值，才推論 alternative hypothesis為真。

p-value < 0.01: there is overwhelming evidence to infer that the alternative hypothesis is true. The test is highly significant.

0.01 < p-value < 0.05: there is strong evidence to infer that the alternative hypothesis is true. The test is significant.

0.05 < p-value < 0.10: there is weak evidence to infer that the alternative hypothesis is true. The test is not statistically significant.

0.10 < p-value: there is no evidence to infer that the alternative hypothesis is true. The test is not statistically significant.

計算 Type II Error 的機率

根據 Type II error 的定義，以前例而言，可推導出

β = P( < 175.34, given that the null hypothesis is false)

以前例而言，若顧客平均每次消費金額(μ)高於$180，發行NFC卡的獲利會很高使得便利商店經理不願意犯 Type II error，因此

β = P( < 175.34, given that μ=180)

意即如果母群體算術平均數實際上是180，錯誤地 not reject the null hypothesis 的機率是 0.0764。

「犯下Type I error的機率」和「Type II error的機率」的關聯性

若試著降低犯下 Type I error 的機率 (α)，犯下 Type II error 的機率 (β) 將會升高。

要如何取捨犯下 Type I error 及 Type II error 的機率，取決於犯下 Type I error 及 Type II error

導致的代價。

樣本數的影響

若增加樣本數(sample size)，可以在不改變 α 的情況下，降低 β。

樣本數愈大，代表資訊愈完整，犯錯的機率會降低，作出的判斷品質會提高。

決定 alternative hypothesis 的方式

以前例而言，若「決定發行NFC卡但實際上無法獲利」的代價(e.g. 若發行NFC卡但無法獲利會賠1個資本額)比「決定不發行NFC卡但實際上可以獲利」嚴重，

因為我們想要避免犯下代價較高的錯誤，我們會把目標設定為證實發NFC卡可以獲利，因此假設會安排如下：

H0: μ = 170

H1: μ > 170

反之，若「決定不發行NFC卡但實際上可以獲利」的代價(e.g.

若不發行NFC卡會少賺1個資本額)較嚴重，我們會把目標設定為證實發NFC卡無法獲利，因此假設會安排如下：

H0: μ = 170

H1: μ < 170