靜宜大學企管系 97 學年第 3 學期

『統計學』講義

任課老師：陳欣得

2009 年 6 月

第零章　統計學概論

2006 年 9 月 18 日 最後修改

0.1 統計學的定義
0.2 敘述統計學與推論統計學
0.3 測量尺度
0.4 資料、資訊與因果關係

0.1　統計學的定義

統計學（*Statistics*）

蒐集、組織、呈現、分析與詮釋資料的科學，其目的在協助作更好的決策。

統計學作**蒐集**、**組織**、**呈現**、**分析**與**詮釋**等操作。

統計學操作的對象是**資料**（*data*）。

統計學的目的是**協助決策**。

範例 0.1　統計學如何協助決策

決策者需要有資訊才能最適當的決策。例如，行銷人員需要知道推出的新產品受歡迎的程度，政治人物需要知道選民支持率。■

0.2　敘述統計學與推論統計學

統計學以操作資料的動作分成兩類：

敘述統計學（*Descriptive Statistics*）、**推論統計學**（*Inferential Statistics*）。

只作蒐集、組織、呈現資料之操作即可達到目的者爲**敘述統計學**；

除以上三者外，還需作分析、詮釋資料者爲推論統計學。

母體（*Population*）與**樣本**（*Sample*）

研究對象的全體稱爲**母體**；

有特殊目的、使用特定方法挑選出來之母體的部分集合稱爲**樣本**。

母體資料經蒐集、組織、呈現後即可達到瞭解研究對象的目的，此爲**敘述統計學**；
樣本資料除蒐集、組織、呈現外，還需作分析、詮釋才能瞭解研究對象，此爲**推論統計學**。

範例 0.2 母體與樣本

我們有興趣的研究對象稱爲母體。如果我們關心企管系同學的學業表現，那麼企管系所有的同學是母體；但是如果關心的是整個靜宜大學同學們的學業表現，則企管系同學只是樣本。

一般而言，我們真正接觸、測量的對象是樣本。例如，在超市前的訪談對象是樣本，電話訪談的對象也是樣本；前者的母體可能是所有的消費者，後者的母體會是所有的投票公民。 ■

0.3 測量尺度

特徵（*Characters*）、**屬性**（*Attributes*）與**變數**（*Variables*）

用以描述具體研究對象的工具。

概念（*Concept*）、**構念**（*Construct*）、**構面**（*Dimensions*）與**變數**（*Variables*）

用以描述抽象研究對象的工具。

測量（*Measurement*）：對某特定對象，賦予某變數一個內容（數值）。

範例 0.3 測量

執行測量之前需確定三樣東西：(1)測量對象，(2)測量特徵，(3)測量工具。以體重計量張三的身高，或以 IQ 量表測量李四的智商，這兩句話中都包含上述的三樣東西。這三樣東西中，測量工具的花樣最多。例如就測量體重而言，除體重計外，人腦的判斷也是成用的測量工具。不是常聽到，這人看起來有六、七十公斤，或這人是肥胖身材等等。 ■

四種**測量尺度**（*Levels of Measurement*）

(1)名目尺度（Nominal Level）

(2)順序尺度（Ordinal Level）

(3)等距尺度（Interval Level）

(4)比例尺度（Ratio Level）

資料（*data*）：

資料是測量的產物。

資料一定與某變數相連結，而一筆資料裡應該會有很多個數值。

測量尺度是對資料的描述，不是對變數的描述。

測量尺度的意義：確定資料可以使用的哪些數學運算作整理。

名目尺度資料—只可計數（算出現幾次）

順序尺度資料—可以排序

等距尺度資料—可以作加、減運算

比例尺度資料—可以作比例（倍數）的運算

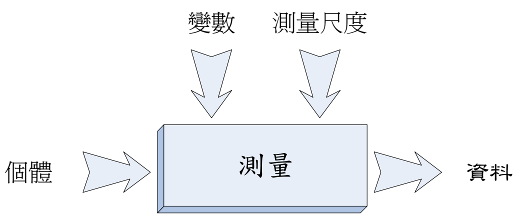
範例 0.4　與測量工具

測量尺度的不同是由測量工具的不同所產生。例如，用體重計測量體重可得到 75.6 公斤的比例尺度資料，用人腦判斷則只能有肥胖、中等、纖瘦等順序尺度資料。 ■

質性資料（*Qualitative Data*）：名目尺度資料或順序尺度資料；

量化資料（*Quantitative Data*）：等距尺度資料或比例尺度資料（可以作加、減計算）。

離散資料（*Discrete Data*）、**連續資料**（*Continuous Data*）



範例 0.5　質性資料與量化資料

質性資料、量化資料的分類，主要在於能否作數學的加、減運算；量化資料可作加減運算，數學工具比較可以幫得上忙，因而從資料中可以獲得比較多的訊息。

另外，質性、量化是對資料的形容，不是對變數的形容。例如下列資料中

ID	性別	身高	體重	血型
1	1	高挑	73	A
2	1	高挑	65	O
3	0	中等	50	O
4	1	矮小	53	B
5	0	高挑	52	A
6	0	中等	45	AB
7	1	矮小	58	O

身高是質性資料（順序尺度），但一般身高計量出來的 170 公分、168.5 公分等是量化資料。亦即，說身高是質性或量化都是不適當的。

最後，質性資料比量化資料容易得到。還是以測量身高爲例，量化資料非有正式的身高計不可，但質性資料只要人腦眼看判斷即可。 ■

互斥且周延（*mutually exclusive and exhaustive*）

測量儀器產生的測量解果必須有互斥且周延的特性。

互斥（*Mutually Exclusive*） 只有一種結果

周延（*Exhaustive*） 一定有結果

範例 0.6 互斥且周延

互斥且周延的限制，是希望測量出來會恰有一個結果（不多不少就是一個結果）。就某特徵對企管系同學作測量，其結果爲『一、二、三、四』之一是合理的，如果出現結果爲『一、二、三、A、B』之一就不合理了。前者的合理在於滿足互斥且周延的精神。事實上如嚴格考究，若有延畢生則『一、二、三、四』就不周延了。解決之道很簡單，加上一個『其他』項就可以了。 ■

0.4 資料、資訊與因果關係

資料（*data*）

測量的結果。

資訊（*information*）

資料分析、詮釋後的結果。

有用的資料爲資訊。

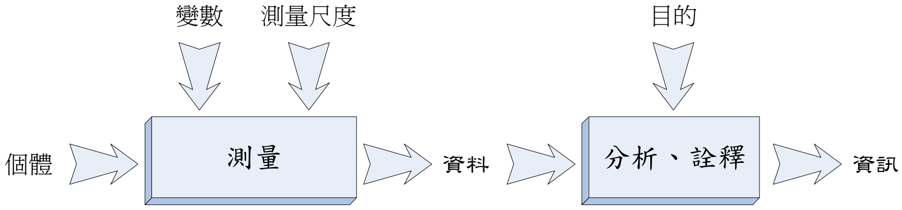
資料有用與否一定涉及研究目的。

研究目的

會影響到變數、測量尺度、樣本的選擇（以蒐集資料）

會影響到組織、呈現資料之方法的選擇（以傳達某種資訊）

會影響到分析、詮釋資料的立場、角度



統計關係（*statistical association*）與因果關係（*causation*）

範例 0.7　統計關係與因果關係

統計可以告訴我們兩個現象之間有同步變化的關連。例如，新書在架上的陳列數量與銷售量之間呈正向關係——陳列數量多者銷售量亦多，潮水水位高低與路上交通擁塞成度也呈正向關係。我們知道後者沒有因果關係，也就是說，控制交通擁塞並不會影響潮水水位。前者是有因果關係，但是哪一個是因呢？陳列數量多造成銷售量高，還是銷售量高引致陳列數量多？ ■

第一章　敘述統計學

2007 年 10 月 30 日 最後修改

1.1　原始資料

1.2　統計表

1.3　統計圖

1.4　統計量值

1.5　一些經驗法則

1.1　原始資料

下表是測量 34 個體（items）之 7 個變數的原始資料：

編號	性別	年齡	學歷	年資	職位	城市	月薪
1	男	36	4	10	5	台北	44,200
2	男	23	1	3	1	台中	16,600
3	女	30	3	4	3	高雄	30,100
4	女	23	1	1	1	高雄	16,400
5	女	31	2	1	1	台北	20,200
6	女	37	3	6	3	高雄	29,600
7	男	27	2	7	4	高雄	34,800
8	男	37	1	4	2	台北	26,100
9	男	34	4	9	5	台北	43,200
10	男	32	2	6	3	台北	32,000
11	男	34	3	4	2	台北	23,300
12	女	27	3	7	3	高雄	29,800
13	男	36	1	13	5	台北	43,100
14	女	30	3	11	5	台北	43,600
15	男	23	3	4	3	台中	28,800
16	男	35	1	7	3	高雄	27,100
17	男	27	3	5	3	台北	29,600
18	男	23	1	3	2	台中	21,500
19	女	28	2	6	3	台中	28,800
20	女	28	2	2	2	台中	22,200
21	男	26	3	7	4	高雄	36,700
22	女	28	3	7	3	台北	32,100
23	男	25	4	3	3	高雄	30,300
24	男	27	4	4	4	高雄	36,800
25	男	30	2	10	4	高雄	34,100
26	男	26	3	3	2	高雄	21,800
27	男	34	3	5	3	台北	31,100
28	男	23	2	3	2	台北	23,000
29	女	23	3	1	2	台北	23,800
30	女	24	3	5	3	高雄	28,200
31	女	23	3	2	2	高雄	24,100
32	女	31	1	9	4	台中	35,100
33	女	33	2	6	2	高雄	21,200
34	女	30	4	2	2	台中	23,200

原始資料（*source data*）

每欄（行）表示一變數，每列表示一個測量的個體。

測量尺度在原始資料中並不一定看得出來。

若只有一個變數，原始資料會如下表（質性資料）

台北	台中	高雄	高雄	台北	高雄
高雄	台北	台北	台北	台北	高雄
台北	台北	台中	高雄	台北	台中
台中	台中	高雄	台北	高雄	高雄
高雄	高雄	台北	台北	台北	高雄
高雄	台中	高雄	台中		

或（量化資料）

36	23	30	23	31	37
27	37	34	32	34	27
36	30	23	35	27	23
28	28	26	28	25	27
30	26	34	23	23	24
23	31	33	30		

分組（*grouping*）

原始資料的觀察個數（列數）一般會超過人類可以輕鬆處理的 5~9 項。

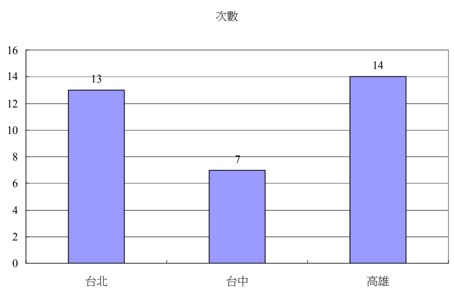
分組是處理大量資料的典型作法。

統計表（*charts*）與統計圖（*graphs*）

統計表類似原始資料表，只是給的是分組後的資料，而且內容是組出現次數。

統計圖是將統計表的訊息（各組出現次數）以圖形來表示。

城市	次數
台北	13
台中	7
高雄	14
合計	34



統計量（*statistic*）

另一種處理大量資料的作法爲用一個數值來表示。

這個數值稱爲統計量。（嚴格而言應爲統計量值。）

最常見的統計量為平均數、變異數、極大值、極小值等。

1.2 統計表

製作統計表有三個動作：

(1)分組；(2)計數；(3)整理成統計表。

其中最重要的是分組。

分組

分組的原則：互斥且周延。

組數最好不要超過 9 組，一般建議 5 至 7 組。

組數的決定也與觀測值的個數有關，觀測個數少則不適合分太多組。

質性資料有自然的分組，但仍須注意是否需合併以免過多組數。

量化資料需人工分組。

量化資料分組程序：

(1)決定全距（range），R

$$R = x_{\max} - x_{\min}$$
$$(\text{全距} = \text{極大值} - \text{極小值})$$

(2)決定組數

(3)決定組距（class width），w

$$w = \frac{R}{n}$$
$$\left(\text{組距} = \frac{\text{全距}}{\text{組數}}\right)$$

(4)寫出組限（class limits），ℓ_i 、 u_i

$$\ell_i = x_{\min} + (i-1) \times w$$
$$u_i = x_{\min} + i \times w$$
$$(\ell_1 = x_{\min}, \quad \ell_{i+1} = u_i, i > 1)$$

(5)寫出分組準則

第一組：$x_{\min} \le x \le u_1$

第 i 組：$\ell_i < x \le u_i \qquad (i \ne 1)$

各組組距不一定要相等。

順序尺度以上資料之分組應依次序（由大而小、或由小而大）排列。

範例 1.1　定組限

就以下 200 個測量值的資料：

56	56	29	38	59	42	47	51	67	26
60	59	75	34	45	29	56	67	65	22
38	39	54	41	20	49	77	46	52	74
57	35	54	46	54	33	68	70	37	85
45	50	40	65	47	44	77	58	47	84
60	80	48	61	14	35	35	65	52	51
61	54	44	46	46	58	68	72	39	63
39	38	49	36	58	38	55	53	54	31
54	27	53	81	30	58	49	54	57	59
32	65	55	32	58	43	44	66	69	37
51	47	27	60	68	56	50	33	42	64
43	50	59	70	50	43	42	37	40	59
59	56	61	53	30	65	52	24	63	40
48	34	43	61	41	41	65	46	73	17
42	37	24	63	40	27	46	56	54	42
59	49	38	62	51	59	37	55	59	53
51	48	58	43	70	66	50	77	49	55
11	43	44	66	38	62	26	59	39	62
34	42	65	42	27	43	30	35	42	51
72	32	42	77	53	46	60	24	59	34

極小值與極大值分別為 $x_{\min} = 11$、$x_{\max} = 85$，全距為

$$R = 85 - 11 = 74$$

我們決定分成 $n = 8$ 組，組距應為

$$w = \frac{R}{n} = \frac{74}{8} = 9.25$$

為了符合一般習慣，將組距、第一組的下限作以下調整：

$$w = 10, \quad \ell_1 = x_{\min} \approx 10$$

則各組上下限依次為

$$u_1 = \ell_2 = 20, \quad u_2 = \ell_3 = 30, \quad \ldots \quad u_8 = 90,$$

各組條件寫出如下表

$$10 \le x \le 20$$
$$20 < x \le 30$$
$$\vdots$$
$$80 < x \le 90$$

如果取得的是離散資料（整數數值），則第一組以後可以寫成

$11 \le x \le 20, \quad 21 \le x \le 30, \quad \ldots \quad 81 \le x \le 90$ 。 ■

統計表分類：

(1)次數分配表（*frequency tables*）

(2)相對次數分配表（*relative frequency tables*）

(3)累計次數分配表（*cumulative frequency tables*）、累計相對次數分配表

(4)列聯表（*contingency tables*、交叉表）

次數分配表

城市	次數
台北	13
台中	7
高雄	14
合計	34

相對次數分配表、累計次數分配表

城市	次數	相對次數
台北	13	13/34
台中	7	7/34
高雄	14	14/34
合計	34	1

年資	次數	相對次數	累計次數(以下)	累計次數(以上)
1~2	6	6/34	6	34
3~5	13	13/34	19	28
6~9	11	11/34	30	15
10 以上	4	4/34	34	4
	34	1		

只有順序尺度以上資料才可以有累計次數分配表（為什麼？）。

相對次數分配表有助於不同資料間的比較。

相對次數分配表是推論統計的基礎，需多注意。

列聯表

列聯表（*contingency table*，**交叉表**）：將兩個變數的次數分配列於同一個統計表。

列聯表中的**邊際次數**（*marginal frequency*）即個別變數的次數分配。

年資＼城市	台北	台中	高雄	邊際次數
1~2	2	2	2	6
3~5	5	3	5	13
6~9	3	2	6	11
10 以上	3	0	1	4
邊際次數	13	7	14	34

1.3　統計圖

以原始資料為基礎的統計圖：

(1)莖葉圖 —— 分組整理資料並呈現

(2)點圖 —— 分組整理資料並呈現

(3)散佈圖 —— 呈現兩變數的變化關係

以統計表為基礎的統計圖：

(1)長條圖、直方圖 —— 比較分組間的次數大小

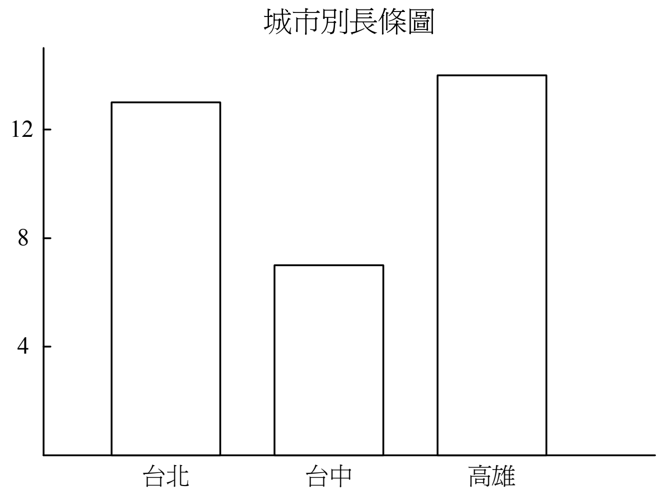
(2)圓餅圖 —— 比較分組次數占整體的比例

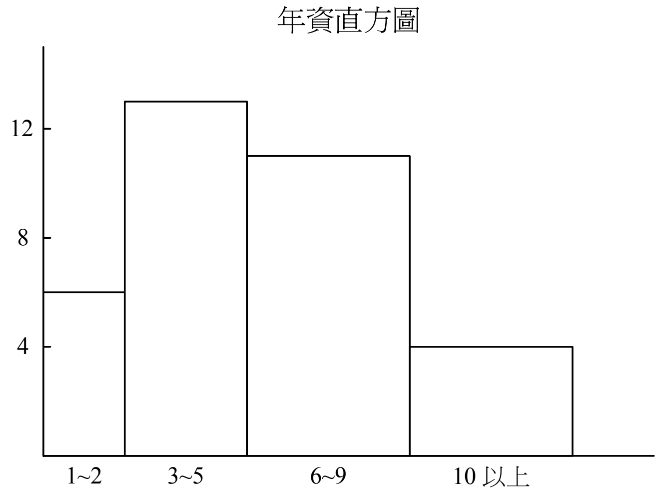
(3)折線圖、肩形圖 —— 比較組間次數的變化趨勢

長條圖、直方圖

長條圖（*bar chart*）：用於質性資料

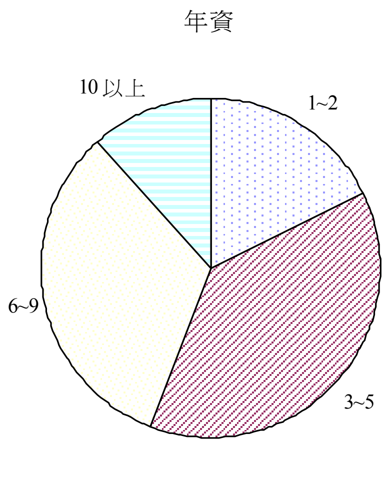
直方圖（*histogram*）：用於量化資料





圓餅圖

圓餅圖（*pie chart*）：用於相對次數分配表

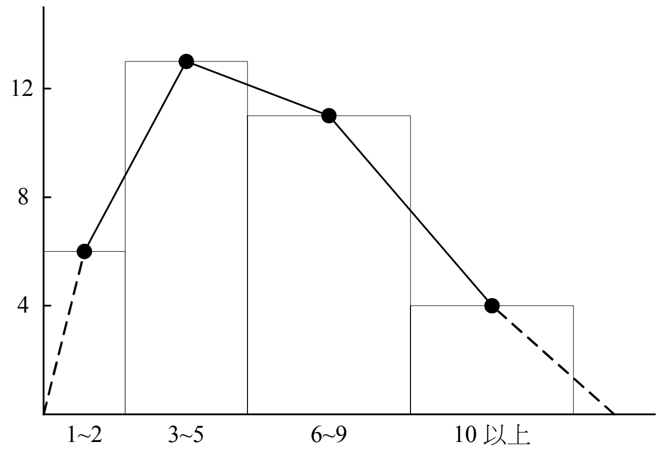


折線圖、肩形圖

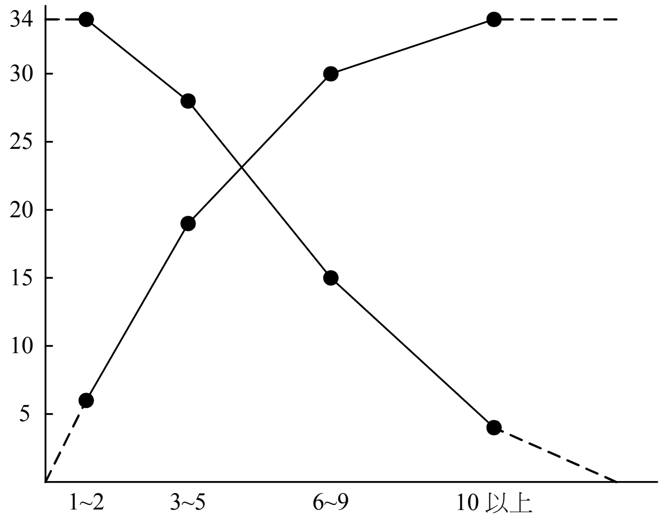
折線圖：用於順序尺度以上之資料

肩形圖（*ogive*）：用於累計次數表（累計相對次數表）

年資折線圖



年資肩形圖



莖葉圖

莖葉圖（*stem-and-leaf display*）：保留原始資料

範例 1.2　莖葉圖

例 1 資料前 5 筆畫入莖葉圖如下：

10			
20	9		
30	8		
40			
50	6	6	9
60			
70			
80			

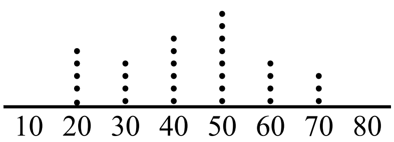
其中第一行稱爲莖，其餘爲葉。圖中表示五十幾的有 3 筆，分別是 56、56、59。前 30 筆填入的結果如下：

10								
20	9	6	9	2	0			
30	8	4	8	9				
40	2	7	5	1	9	6		
50	6	6	9	1	9	6	4	2
60	7	0	7	5				
70	5	7	4					
80								

莖葉圖只適用於小量資料，200 筆太多了。 ■

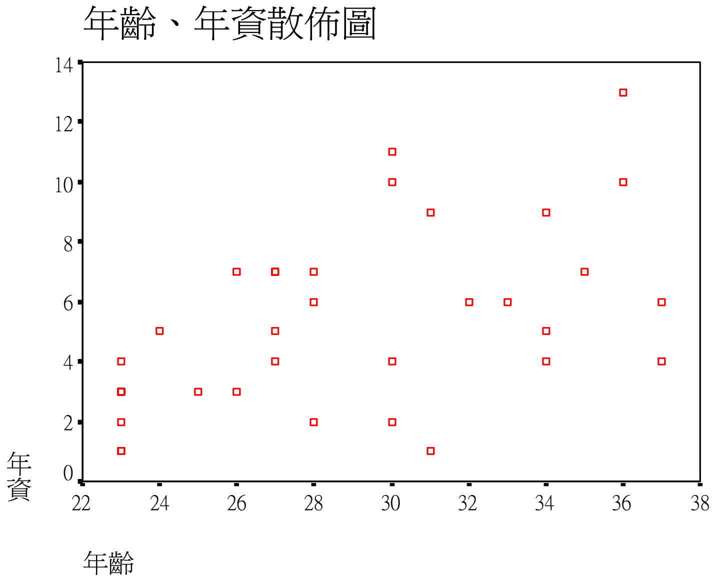
點圖

點圖（*dot plot*）：適用於觀測數量大的資料



散佈圖

散佈圖（*scatter diagram*）：瞭解兩量化資料爲正相關或負相關。



1.4　統計量值

順序尺度資料的統計量值：

(1)**中位數**（*median*，M_e）

至少有 50%的數值小於等於M_e，且最少有 50%的數值小於M_e。

(2)*眾數*（*mode*，M_o）

出現次數最多的數值，可能有一個以上的眾數。

(3)**四分位數**（*quartiles*，Q_1、Q_2、Q_3）

至少有 25%的數值小於等於Q_1，且最少有 25%的數值小於Q_1。
Q_1稱爲第一四分位數，Q_3稱爲第三四分位數。
$Q_2 = M_e$

(4)**百分位數**（*percentiles*，P_1、P_{10}、P_{90}、P_{99}）

至少有$n\%$的數值小於等於P_n，且最少有$n\%$的數值小於P_n。
$P_{25} = Q_1$、$P_{50} = Q_2 = M_e$、$P_{75} = Q_3$

(5)極值

量化資料的統計值：

(1)位置測量值：平均數（*mean*，μ）

$$\mu = \frac{x_1 + x_2 + \cdots + x_N}{N} = \frac{\Sigma x_i}{N}$$

樣本平均數： $\overline{x} = \frac{x_1 + x_2 + \cdots + x_n}{n} = \frac{\Sigma x_i}{n}$

(2)離散測量值：標準差（*standard deviation*，σ）、變異數（*variance*，σ^2）

$$\sigma^2 = \frac{\Sigma\left(x_i - \mu\right)^2}{N} = \frac{\Sigma x_i^2 - N \times \mu^2}{N}$$

樣本變異數： $s^2 = \frac{\Sigma\left(x_i - \overline{x}\right)^2}{n-1} = \frac{\Sigma x_i^2 - n \times \overline{x}^2}{n-1}$

(3)變異係數（*coefficient of variance*，*CV*）

$$CV = \frac{\sigma}{\mu} \times 100\% \quad 或 \quad CV = \frac{s}{\overline{x}} \times 100\%$$

(4)偏態（*skewness*，α_3）

$$\alpha_3 = \frac{\Sigma\left(x_i - \mu\right)^3 / N}{\sigma^3}$$

$\alpha_3 < 0$ 的情況稱爲左偏（*negatively skewed*）；

$\alpha_3 > 0$ 的情況稱爲右偏（*positively skewed*）。

偏態係數（*coefficient of skewness*）

$$S_k = \frac{3\left(\mu - M_e\right)}{\sigma}$$

$S_k < 0$ 稱爲左偏、$S_k > 0$ 爲右偏。

(5)峰態（*kurtosis*，α_4）

$$\alpha_4 = \frac{\Sigma\left(x_i - \mu\right)^4 / N}{\sigma^4}$$

$\alpha_4 < 3$ 稱爲低闊峰（*platy-kurtosis*）、$\alpha_4 > 3$ 稱爲高狹峰（*lepto-kurtosis*）

兩變數間相關性的統計量值：

(1)共變數（*covariance*，σ_{xy}）

$$\sigma_{xy}=\frac{\Sigma(x_i-\mu_x)(y_i-\mu_y)}{N}=\frac{\Sigma x_i y_i-N\times\mu_x\times\mu_y}{N}$$

樣本共變數：$s_{xy}=\frac{\Sigma(x_i-\bar{x})(y_i-\bar{y})}{n-1}=\frac{\Sigma x_i y_i-n\times\bar{x}\times\bar{y}}{n-1}=\frac{\Sigma x_i y_i-(\Sigma x_i\Sigma y_i)/n}{n-1}$

(2)相關係數（*correlation coefficients*，ρ）

$$\rho=\frac{\sigma_{xy}}{\sigma_x\sigma_y}=\frac{\Sigma(x_i-\mu_x)(y_i-\mu_y)}{\sqrt{\Sigma(x_i-\mu_x)^2}\sqrt{\Sigma(y_i-\mu_y)^2}}=\frac{\Sigma x_i y_i-N\times\mu_x\times\mu_y}{\sqrt{\Sigma x_i^2-N\times\mu_x^2}\sqrt{\Sigma y_i^2-N\times\mu_y^2}}$$

樣本相關係數：

$$r=\frac{s_{xy}}{s_x s_y}=\frac{\Sigma x_i y_i-n\times\bar{x}\times\bar{y}}{\sqrt{\Sigma x_i^2-n\times\bar{x}^2}\sqrt{\Sigma y_i^2-n\times\bar{y}^2}}=\frac{\Sigma x_i y_i-(\Sigma x_i\Sigma y_i)/n}{\sqrt{\Sigma x_i^2-(\Sigma x_i)^2/n}\sqrt{\Sigma y_i^2-(\Sigma y_i)^2/n}}$$

$-1\le\rho\le1$，$\rho=-1$時稱完全負相關、$\rho=1$時稱完全正相關。

第 *n* 百分位數（P_n）的求解步驟

(1)求所在位置的名次（令總觀察值數目爲 N）

$$i=N\times n\%+0.5$$

(2)找第 i 名的數值即爲 P_n

(a)未分組資料

(a1)需要報告觀察值

$$P_n=x_{round(i)}$$

(a2)不需要報告觀察值

$$P_n=\frac{x_I+x_{I+1}}{2}\quad\text{或}\quad P_n=x_i\ (i\text{ 爲整數})$$

其中，I 爲 i 之整數部分。

(b)分組資料

$$P_n=\ell_k+\frac{i-0.5-N_k}{n_k}\times w_k=\ell_k+\frac{n\%\times N-N_k}{n_k}\times w_k$$

其中，ℓ_k 、 n_k 、 N_k 、 w_k 爲所在組的下限、該組個數、該組前累計個數、組寬

第 *n* 百分位數（P_n）的求解步驟（課本的作法）

(1)求所在位置的名次（令總觀察值數目爲 N）

$$i = N \times n\%$$

(2)找第 i 名的數值即爲 P_n

(a)未分組資料

(a1)若 i 爲整數

$$P_n = \frac{x_i + x_{i+1}}{2}$$

(a2)若 i 不是整數

$$P_n = x_{I+1}$$

其中，I 爲 i 之整數部分。

(b)分組資料

$$P_n = \ell_k + \frac{i - N_k}{n_k} \times w_k = \ell_k + \frac{n\% \times N - N_k}{n_k} \times w_k$$

其中，ℓ_k 、 n_k 、 N_k 、 w_k 爲所在組的下限、該組個數、該組前累計個數、組寬

第 *n* 百分位數（P_n）的求解步驟（**Excel** 內建函數的作法）

(1)求所在位置的名次（令總觀察值數目爲 N）

$$i = (N - 1) \times n\% + 1$$

(2)找第 i 名的數值即爲 P_n

(a)未分組資料

(a1)若 i 爲整數

$$P_n = x_i$$

(a2)若 i 不是整數

$$P_n = x_I + R \times (x_{I+1} - x_I)$$

其中，I 爲 i 之整數部分，R 爲 i 之小數部分。

(b)分組資料（Excel 未提供此部分解答）

範例 1.3　未分組資料之 P_{10}

就下列資料：

10								
20	9	6	9	2	0			
30	8	4	8	9				
40	2	7	5	1	9	6		
50	6	6	9	1	9	6	4	2
60	7	0	7	5				
70	5	7	4					
80								

求 P_{10}：

$N = 30$ ，$i = \frac{1}{10} \times 30 + 0.5 = 3.5$ ，則

(需給觀察值) $P_{10} = x_{round(3.5)} = x_4 = 29$

(不需觀察值) $P_n = \dfrac{x_3 + x_4}{2} = \dfrac{26 + 29}{2} = 27.5$

求 Q_1 、 Q_3：

$N = 30$ ，$i_1 = \frac{1}{4} \times 30 + 0.5 = 8$ 、 $i_2 = \frac{2}{4} \times 30 + 0.5 = 15.5$ 、 $i_3 = \frac{3}{4} \times 30 + 0.5 = 23$ ，則

$Q_1 = x_8 = 38$ 、 $M_e = Q_2 = x_{15.5} = \dfrac{x_{15} + x_{16}}{2} = \dfrac{49 + 51}{2} = 50$ 、 $Q_3 = x_{23} = 59$

$IQR = Q_3 - Q_1 = 59 - 38 = 21$

IQR 稱爲**四分位數距**。

（課本作法）

$N = 30$ ，$i_1 = \frac{1}{4} \times 30 = 7.5$ 、 $i_2 = \frac{2}{4} \times 30 = 15$ 、 $i_3 = \frac{3}{4} \times 30 = 22.5$ ，則

$Q_1 = x_8 = 38$ 、 $M_e = Q_2 = \dfrac{x_{15} + x_{16}}{2} = \dfrac{49 + 51}{2} = 50$ 、 $Q_3 = x_{23} = 59$

$IQR = Q_3 - Q_1 = 59 - 38 = 21$

（Excel 作法）

$N = 30$ ，$i_1 = \frac{1}{4} \times 29 + 1 = 8.25$ 、 $i_2 = \frac{2}{4} \times 29 + 1 = 15.5$ 、 $i_3 = \frac{3}{4} \times 29 + 1 = 22.75$ ，則

$$Q_1 = x_8 + 0.25\times(x_9 - x_8) = 38 + 0.25\times(39-38) = 38.25$$

$$M_e = Q_2 = x_{15} + 0.5\times(x_{16} - x_{15}) = 49 + 0.5\times(51-49) = 50$$

$$Q_3 = x_{12} + 0.75\times(x_{23} - x_{22}) = 59 + 0.75\times(59-59) = 59$$

$$IQR = Q_3 - Q_1 = 59 - 38.25 = 20.75$$ ■

範例 1.4　分組資料之 P_{10}

就下列資料：

組別	次數
20≤X<30	4
30≤X<40	6
40≤X<50	12
50≤X<60	14
60≤X<70	7
70≤X≤80	7
總和	50

求 P_{10}：

$N = 50$ ，$i = \frac{1}{10}\times 50 + 0.5 = 5.5$ ，則

$$k = 2,\quad \ell_2 = 30,\quad n_2 = 6,\quad N_2 = 4,\quad w_2 = 10$$

$$P_{10} = \ell_k + \frac{i - 0.5 - N_k}{n_k}\times w_k = 30 + \frac{5.5 - 0.5 - 4}{6}\times 10 = 31.67$$

求 Q_1 、 Q_3 ：

$N = 50$ ，$i_1 = \frac{1}{4}\times 50 + 0.5 = 13$ 、 $i_2 = \frac{2}{4}\times 50 + 0.5 = 25.5$ 、 $i_3 = \frac{3}{4}\times 50 + 0.5 = 38$ ，

則

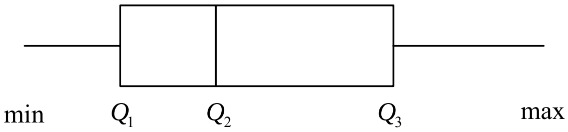
$$Q_1 = x_{13} = \ell_3 + \frac{i_1 - 0.5 - N_3}{n_3}\times w_3 = 40 + \frac{13 - 0.5 - 10}{12}\times 10 = 42.08$$

$$M_e = Q_2 = x_{25.5} = \ell_4 + \frac{i_2 - 0.5 - N_4}{n_4}\times w_4 = 50 + \frac{25.5 - 0.5 - 22}{14}\times 10 = 52.14$$

$$Q_3 = x_{38} = \ell_5 + \frac{i_3 - 0.5 - N_5}{n_5}\times w_5 = 60 + \frac{38 - 0.5 - 36}{7}\times 10 = 62.14$$

$$IQR = Q_3 - Q_1 = 62.14 - 40.08 = 22.06$$ ■

盒鬚圖（*box chart*）：將 min、 Q_1 、 Q_2 、 Q_3 、max 化在同一圖上。



範例 1.5　盒鬚圖

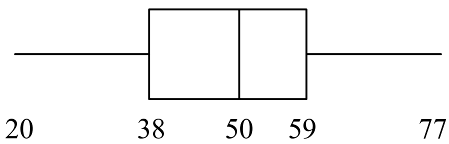
就下列資料：

10								
20	9	6	9	2	0			
30	8	4	8	9				
40	2	7	5	1	9	6		
50	6	6	9	1	9	6	4	2
60	7	0	7	5				
70	5	7	4					
80								

$x_{\min} = 20$ 、 $x_{\max} = 77$

$Q_1 = x_8 = 38$ 、 $M_e = Q_2 = x_{15.5} = \dfrac{x_{15} + x_{16}}{2} = \dfrac{49+51}{2} = 50$ 、 $Q_3 = x_{23} = 59$

其盒鬚圖如下：



由盒鬚圖得知這些數值整體而言呈對稱分配，但有一點左偏。 ■

動差（*moments*）

(1)**零動差**（*zero moments*）

$$n\text{ 級零動差} = \frac{\Sigma x_i^n}{N}$$

(2)**主動差**（*principle moments*）

$$n\text{ 級主動差} = M_n = \frac{\Sigma\left(x_i - \mu\right)^n}{N} = \frac{\Sigma\left(x_i - \mu\right)^n}{N}$$

$$M_2 = \frac{\Sigma\left(x_i - \mu\right)^2}{N} = \frac{\Sigma x_i^2 - \left(\Sigma x_i\right)^2 / N}{N}$$

$$M_1 = 0$$

$$M_2 = \sigma^2$$

$$M_3 = \alpha_3 \times \sigma^3$$

$$M_4 = \alpha_4 \times \sigma^4$$

範例 1.6　未分組資料之變異數

就下列資料：

x	12	14	22	16	34	16	23	21

計算工作表如下

x	x^2
12	144
14	196
22	484
16	256
34	1,156
16	256
23	529
21	441
158	3,462

$$\sum(x-\bar{x})^2 = \Sigma x^2 - \frac{(\Sigma x)^2}{n}$$

$$\text{母體變異數 } \sigma_X^2 = \frac{\Sigma x^2 - \frac{(\Sigma x)^2}{n}}{n} = \frac{3462 - \frac{158^2}{8}}{8} = \frac{341.5}{8} = 42.69$$

$$\text{樣本變異數 } s_X^2 = \frac{\Sigma x^2 - \frac{(\Sigma x)^2}{n}}{n-1} = \frac{3462 - \frac{158^2}{8}}{7} = \frac{341.5}{7} = 48.79$$ ■

範例 1.7　分組資料之變異數

就下列資料：

組別	次數
20≤X<30	4
30≤X<40	6
40≤X<50	12
50≤X<60	14
60≤X<70	7
70≤X≤80	7
總和	50

計算工作表如下

組別	代表數	次數	累計次數	X	X^2
20≤X<30	25	4	4	100	2,500
30≤X<40	35	6	10	210	7,350
40≤X<50	45	12	22	540	24,300
50≤X<60	55	14	36	770	42,350
60≤X<70	65	7	43	455	29,575
70≤X≤80	75	7	50	525	39,375
總和		50		2,600	145,450

$$\mu=\frac{\Sigma x}{N}=\frac{2600}{50}=52$$

$$\sigma^2=\frac{\Sigma x^2-(\Sigma x)^2/N}{N}=\frac{145,450-2,600^2/50}{50}=205$$

$$\sigma=\sqrt{205}=14.32$$

$$CV=\frac{\sigma}{\mu}=\frac{14.32}{52}=0.2754=27.54\%$$

■

範例 1.8　分組資料之相關係數

就下列資料：

x	12	14	22	16	34	16	23	21
y	47	51	75	53	90	43	67	64

計算工作表如下

x	*y*	x^2	y^2	*xy*
12	47	144	2,209	564
14	51	196	2,601	714
22	75	484	5,625	1,650
16	53	256	2,809	848
34	90	1,156	8,100	3,060
16	43	256	1,849	688
23	67	529	4,489	1,541
21	64	441	4,096	1,344
158	490	3,462	31,778	10,409

$$\sum\left(X-\overline{X}\right)\left(Y-\overline{Y}\right)=\Sigma XY-\frac{\Sigma X\times\Sigma Y}{n}$$

$$s_x=\sqrt{\frac{\Sigma x^2-\frac{(\Sigma x)^2}{n}}{n-1}}=\sqrt{\frac{3,462-\frac{158\times158}{8}}{8-1}}=\sqrt{48.79}=6.98$$

$$s_y=\sqrt{\frac{\Sigma y^2-\frac{(\Sigma y)^2}{n}}{n-1}}=\sqrt{\frac{31,778-\frac{490\times 490}{8}}{8-1}}=\sqrt{252.21}=15.88$$

$$s_{xy}=\frac{\Sigma xy-\frac{\Sigma x\times\Sigma y}{n}}{n-1}=\frac{10,409-\frac{158\times 490}{8}}{8-1}=\frac{731.5}{7}=104.50$$

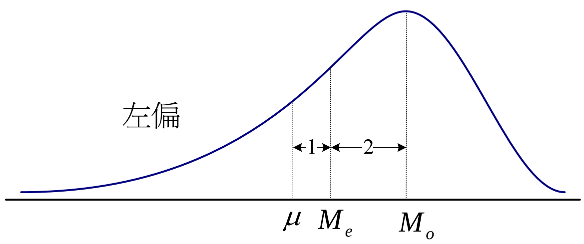
$$r=\frac{s_{xy}}{s_x s_y}=\frac{104.50}{6.98\times 15.88}=0.94$$ ■

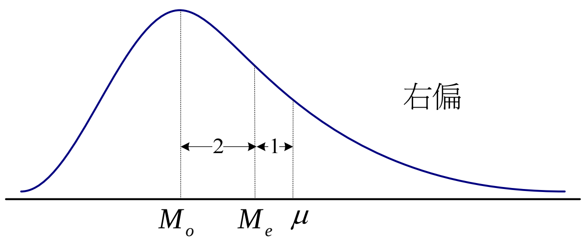
平均數、中位數、眾數的關係

皮爾森公式：

(1)中位數一定介於平均數與眾數之間

(2)$|$眾數$-$中位數$|=2|$平均數$-$中位數$|$





範例 1.9　皮爾森公式

已知$\mu=20$、$M_e=16$，求M_o。

【解】

由 $M_e<\mu$ 知此資料為右偏，且$M_o<M_e=16$，

因 $|M_o-M_e|=2|\mu-M_e|=2(20-16)=8$，

故$M_o=M_e-8=16-8=8$ ■

1.5　一些經驗法則

z **分數**（*z-score*）

$$z_i = \frac{x_i - \mu}{\sigma} \qquad \text{或} \qquad z_i = \frac{x_i - \bar{x}}{s}$$

範例 1.10　z 分數

求以下數值的 z 分數：

6	9	11	8	3

【解】

計算平均數與標準差，工作表如下：

x	x^2	z-score
6	36	-0.46
9	81	0.52
11	121	1.18
8	64	0.20
3	9	-1.44
37	311	

其中

$$\Sigma x = 37, \quad \Sigma x^2 = 311, \quad n = 5$$

故

$$\bar{x} = \frac{\Sigma x}{n} = \frac{37}{5} = 7.4, \quad s^2 = \frac{\Sigma x^2 - (\Sigma x)^2 / n}{n-1} = \frac{311 - 37^2/5}{5-1} = 9.3, \quad s = \sqrt{9.3} = 3.05$$

$$z_1 = \frac{x_1 - \bar{x}}{s} = \frac{6 - 7.4}{3.05} = -0.46$$

$$z_2 = \frac{x_2 - \bar{x}}{s} = \frac{9 - 7.4}{3.05} = 0.52$$

Z 分數小於零表示該觀測值小於平均數；z 分數之絕對值越大，表示該觀測值離平均數越遠。 $z = 1.18$ 表示該觀測值大於平均數有 1.18 個標準差。 ■

柴比雪夫定理（*Chebyshev's theory*）

一組觀察值中，至少有$1-\frac{1}{k^2}$比例的觀察值，落在距離平均數k個標準差之內。($k \ge 1$)

以機率符號表示則爲

$$P\left(|x-\mu| \le k\sigma\right) \ \ge \ 1-\frac{1}{k^2}$$

範例 1.11　柴比雪夫定理

已知企管系 100 名學生，統計學的平均分數爲 60 分，變異數爲 16；

(a)請估計至少有多少人分數落在 55 分到 65 分之間；

(b)請找出至少有 89 個學生在內的區間；

(c)請估計至少有多少人分數落在 55 分到 75 分之間。

【解】

(a)

$$k=\frac{|55-60|}{\sqrt{16}}=\frac{|65-60|}{\sqrt{16}}=\frac{5}{4} \ \Rightarrow \ P\left(55 \le X \le 65\right)=1-\frac{1}{\left(\frac{5}{4}\right)^2}=\frac{9}{25}$$ 。

至少有 $\frac{9}{25}\times 100=36$ 位同學。

(b)

機率$=\frac{89}{100}=1-\frac{1}{k^2} \ \Rightarrow \ k=3.015$

範圍爲 $\left\{60-k\times\sqrt{16} \le X \le 60+k\times\sqrt{16}\right\}=\left\{48 \le X \le 72\right\}$

(c)

$$k_1=\frac{|55-60|}{\sqrt{16}}=\frac{5}{4}\ ,\ k_2=\frac{|75-60|}{\sqrt{16}}=\frac{15}{4}$$

$$\Rightarrow \ P\left(55 \le X \le 75\right)=\frac{1}{2}\times\left[\left(1-\frac{1}{\left(\frac{5}{4}\right)^2}\right)+\left(1-\frac{1}{\left(\frac{15}{4}\right)^2}\right)\right]=\frac{1}{2}\times\frac{81+209}{225}=\frac{29}{45}$$ ■

經驗法則（*the empirical rule*）

若資料呈鐘形（單峰、對稱）分佈，則

(1)約 68% 的觀察值落在離平均數 1 個標準差內；

(2)約 95% 的觀察值落在離平均數 2 個標準差內；

(3)約 99.7% 的觀察值落在離平均數 3 個標準差內；

(4)約 $1-0.1^k$ 的觀察值落在離平均 k（$k \ge 4$）個標準差內。

範例 1.12　經驗法則

已知企管系 100 名學生，統計學的分數呈常態的鐘形分配，其平均分數爲 60 分，變異數爲 16；

(a)請估計至少有多少人分數落在 56 分到 64 分之間。

(b)請找出至少有 95 個學生在內的區間。

【解】

$\mu = 60,\ \sigma = \sqrt{16} = 4$

(a)

$$z = \frac{|56-60|}{\sqrt{16}} = \frac{|64-60|}{\sqrt{16}} = 1 \ \Rightarrow\ P(56 \le X \le 64) = 68\%$$

至少有 $68\% \times 100 = 68$ 位同學。

(b)

$$\text{機率} = \frac{95}{100} = 95\% \ \Rightarrow\ z = 2$$

$$\text{範圍爲 } \left\{60 - z \times \sqrt{16} \le X \le 60 + z \times \sqrt{16}\right\} = \{52 \le X \le 68\}$$ ■

離群值（*outliner*）

離平均數太遠，以致於被認爲不屬於該群資料的數值。

離群值的判斷

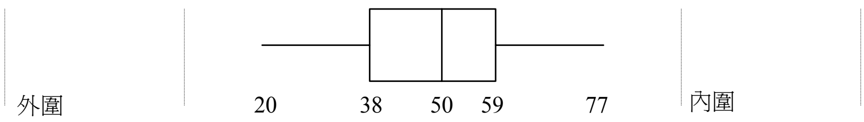
(1)z 分數法

若 $|z_i| > 3$，則 x_i 可視爲離群值。

(2)盒鬚圖法

若 $Q_1 - x_i > 1.5 \times IQR$ 或 $x_i - Q_3 > 1.5 \times IQR$，則 x_i 認定爲**懷疑離群值**；

若 $Q_1 - x_i > 3\times IQR$ 或 $x_i - Q_3 > 3\times IQR$，則 x_i 認定爲**確定離群值**。



範例 1.13　離群值

就下列資料：

96	105	85	74	110	90	103	106	102	99

請判斷是否有離群資料。

【解】

96	105	85	74	110	90	103	106	102	99

平均數	97.0
樣本標準差	11.1
max	110
min	74

計算結果：$\bar{x} = 97.0,\ s = 11.1,\ \max = 110,\ \min = 74$

下界限值 $= 97 - 3\times 11.1 = 63.7$, 上界限值 $= 97 + 3\times 11.1 = 130.3$

沒有任何離群值。 ■

範例 1.14　離群值

就下列資料：

21	37	29	39	26	42	29	25	33	25
65	24	26	26	27	23	29	25	26	44

(a)請分別計算懷疑之離群值的上、下界限值（內圍值）；

(b)請分別計算確定之離群值的上、下界限值（外圍值）；

(c)請問本組資料有沒有確定之離群值、或懷疑之離群值。

【解】

21	37	29	39	26	42	29	25	33	25
65	24	26	26	27	23	29	25	26	44

Q1=	25.0	mean =	31.1
Q2=	26.5	median =	26.5
Q3=	34.0	mode =	26
IQR=	9.0		
MAX=	65.0		
MIN=	21.0		

(a)

懷疑離群值之上、下限分別離 Q_3、Q_1 1.5 個 IQR ：

$$Q_3 + 1.5 \times IQR = 34 + 1.5 \times 9 = 47.5$$

$$Q_1 - 1.5 \times IQR = 25 - 1.5 \times 9 = 11.5$$

(b)

確定離群值之上、下限分別離 Q_3、Q_1 3 個 IQR ：

$$Q_3 + 3 \times IQR = 34 + 3 \times 9 = 61$$

$$Q_1 - 3 \times IQR = 25 - 3 \times 9 = -2$$

(c)本組資料有確定離群值 65。 ■

第二章　機率概論

2006 年 10 月 7 日 最後修改

2.1　相對次數與機率

2.2　樣本空間、事件與隨機變數

2.3　抽樣與樣本空間

2.4　事件的性質與計數法則

2.5　貝氏定理

2.1　相對次數與機率

相對次數大的組別出現的機會比較大，我們以相對次數作爲該組別出現的機率。

城市	次數	相對次數
台北	13	13/34
台中	7	7/34
高雄	14	14/34
合計	34	1

機率（*probability*）

客觀相對次數，機率大表示該類別組成份子之數目比較多，因此比較容易出現。

可能性（*possibility*）

主觀相退次數，可能性大表示決策者選擇該組的傾向比較大。

範例 2.1　機率與可能性

陳老師在拍賣網站購買某特定商品的可能性是 30%，

根據以往經驗，陳老師在網路上購物被騙的機率是 10%；

所有進入某特定商品拍賣網頁的瀏覽者中，出價的機率是 5%。

機率的分類

古典機率（*classical probability*）

已知母體次數分配表，相對次數即該類別（事件）出現的機率。

經驗機率（*empirical probability*）

只知樣本次數分配表，相對次數即該類別（事件）出現的機率。

客觀機率（*objective probability*）

有真正次數分配表爲依據的機率，如古典機率、經驗機率等。

主觀機率（*subjective probability*）

沒有完整次數分配表爲依據，主要由人類心智判斷、產生的機率。

範例 2.2　機率與可能性

牛蛇對戰中，沒有牛勝的**古典機率**，因爲未來的對戰情況尚未發生；
參考本球季的紀錄，牛勝的機率爲 $\frac{10}{16}$ （**經驗機率**）；

基於牛奪冠，興隆超市會全館打折，陳老師希望牛勝的機率是 90%（**主觀機率**）；

考慮兩對的紀錄與狀況，陳老師認爲牛勝的機率是 70%（**什麼機率？**）。

量化資料的機率

質性資料不能作加法運算，在數學無法介入的情況下，能作的事並不多。

量化資料的相對次數（機率）分配表如果可以用數學公式呈現，則大有作爲。

可以寫成量化資料相對次數（機率）分配表的變數稱爲**隨機變數**（*random variables*）。

請瞭解，統計表（相對次數分配表）中的一個分組，即是隨機變數之每一個對應值。

範例 2.3　機率分配與隨機變數

以下是量化資料的相對次數分配表，其中相對次數已經直接寫成機率 $P(x)$：

x	$P(x)$
0	1/16
1	4/16
2	6/16
3	4/16
4	1/16
	1

上表可以寫成函數形式

$$P(x)=\frac{24}{x!(4-x)!}\times\left(\frac{1}{2}\right)^4=\frac{3}{2}\frac{1}{x!(4-x)!},\quad x=0,1,\ldots,4$$

x 爲一個有 5 個可能數值（5 組）的隨機變數。

2.2 樣本空間、事件與隨機變數

在機率的應用裡，我們有興趣的分組不見得與樣本特徵相關的自然分組相同。

範例 2.4 樣本特徵與分組

一個骰子有六個面，若以 $x, x=1,\ldots,6$ 代表出現那面的點數，則擲一個骰子的機率分配表如下：

x	$p(x)$
1	1/6
2	1/6
3	1/6
4	1/6
5	1/6
6	1/6

然而，擲骰子遊戲每次不止擲一個骰子。最常見的是擲四個骰子，扣除兩個相同點數的骰子，以其他兩個骰子的點數和之個位數爲其結果。例如，擲出 $(2,3,2,4)$ 其結果爲 7，而擲出 $(3,3,5,6)$ 的結果爲 1。若四個骰子的點數都不相同，則需重擲。爲了可以簡單分析，假設我們只擲兩個骰子，令 y 爲其結果，則 y 的分配表如下：

y	$p(y)$
0	3/36
1	2/36
2	2/36
3	2/36
4	3/36
5	4/36
6	5/36
7	6/36
8	5/36
9	4/36

爲了增加遊戲的趣味性，我們會擲不同個數的骰子，也會有不同結果的規則；但是不變的是擲一個骰子的點數分配一定如 x，而其他的變化一定以 x 爲基礎計算而來。

樣本空間（*sample space*，Ω）

樣本空間是我們有興趣之個體所成的集合，即相對次數分配表中的母體。

樣本空間是一個**集合**（*set*）。

樣本空間的每一**元素**（個體）是我們作測量的單位。

隨機變數（*random variables*，RV）

隨機變數爲建立於樣本空間的實數值函數。

隨機變數是一個函數（不是一個變數！）。

隨機變數的定義域是樣本空間，值域是實數。

隨機變數的定義讓相對次數分配表的分組一定是量化資料。

實驗（*experiment*）

實驗爲對樣本空間之元素作挑選。

出象（*outcome*）

實驗的結果稱爲出象，也就是對實驗挑選的元素作測量的結果。

事件（*events*）

事件爲樣本空間的部分集合。

對母體而言，事件是有某種共同特徵的個體所組成的集合。

我們以其共同特徵來指稱某事件。

對相對次數分配表（機率分配表）而言，事件是一個分組。

在敘述統計學中，事件大多以隨機變數的某個範圍來描述。

事件空間（*event space*，E）

所有事件所成的集合稱爲事件空間。

機率（*probability*）

建立於事件空間，值域爲$\left[0,1\right]$區間的函數。

機率是事件的特徵。

機率公理

函數$f\left(\bullet\right)$要成爲一機率函數，需滿足下列三條件：

(1) $f\left(\varnothing\right)=0$，$f\left(\Omega\right)=1$；

(2) $0\le f\left(A\right)\le 1, \forall A\in E$；

(3) $A\subseteq B \;\Rightarrow\; f\left(A\right)\le f\left(B\right)$。

機率空間（*probability space*）

樣本空間、事件空間與機率函數，$\left(f(\bullet),E,\Omega\right)$，之組合。

注意隨機變數在機率空間中的角色。

範例 2.5 機率空間

在擲兩個骰子的範例中，所有可能結果的集合爲

$$\Omega=\left\{\left(x_1,x_2\right)\middle|\,x_1=1,\ldots,6,\ x_2=1,\ldots,6\right\}$$
$$=\left\{(1,1),(1,2),\ldots,(1,6),(2,1),\ldots,(2,6),(3,1),\ldots,(4,1),\ldots,(5,1),\ldots,(6,1),\ldots,(6,6)\right\}$$

事件空間爲

$$E=\left\{E_y\middle|\,y=0,1,\ldots,9\right\}$$
$$=\left\{E_0,E_1,\ldots,E_9\right\}$$

其中

$$E_y=\left\{\left(x_1,x_2\right)\middle|\,x_1+x_2=y \text{ 或 } x_1+x_2=10+y\right\}$$
$$E_0=\left\{(4,6),(5,5),(6,4)\right\}$$
$$E_1=\left\{(5,6),(6,5)\right\}$$
$$\vdots$$
$$E_9=\left\{(4,5),(5,4)\right\}$$

機率函數爲

$$P(y=0)=P(E_0)=\tfrac{3}{36}$$
$$P(y=1)=P(E_1)=\tfrac{2}{36}$$
$$\vdots$$
$$P(y=9)=P(E_9)=\tfrac{4}{36}$$

離散隨機變數（*discrete random variables*）

以整數或整數之部分集合爲值域的隨機變數。

連續隨機變數（*continuous random variables*）

以實數或實數之部分集合爲值域的隨機變數。

【注意】將相對次數分配表對照、比喻爲機率空間有兩個差異點：

(1)前者單一個體間的機率完全相等，後者則沒有這種限制；

(2)前者可以是質性資料，後者則必然是量化資料。

2.3 抽樣與樣本空間

樣本空間的元素可以代表單一個體，也可以代表一組個體。
後者在統計上的應用比較重要。

抽樣（*sampling*）

抽樣為從母體中選出一組個體（母體的部分集合）。

樣本（*sample*）

樣本為抽樣所選出來的那組個體。
樣本是母體的部分集合；
大多情況，『一個樣本』中有多於一個的個體；因此，稱為『一組樣本』比較恰當。
有時（比較不嚴謹），樣本會指一組樣本中的一個元素。

樣本空間（*sample space*）

所有樣本所成的集合。

範例 2.6 樣本空間

假設母體為四個元素所成的集合

$$P=\{a,b,c,d\}$$

而且每一組樣本為母體中之 2 個相異元素所組成，則樣本空間為

$$\Omega=\{(a,b),(a,c),(a,d),(b,c),(b,d),(c,d)\}$$

又令隨機變數 X 的定義為

$$X(a,b)=3,\ X(a,c)=4,\ X(a,d)=5,\ X(b,c)=5,\ X(b,d)=6,\ X(c,d)=7$$

則事件 $\{X=5\}$ 的機率為

$$P(X=5)=P(\{(a,d),(b,c)\})=\frac{n\{(a,d),(b,c)\}}{n(\Omega)}=\frac{2}{6}$$ 。

範例 2.7 樣本空間

一班 50 位同學，如果每一元素代表一個同學，則樣本空間有 50 個元素；

如果每一元素代表 2 位一組同學，則樣本空間有 $\frac{50\times 49}{2}=1,225$ 個元素；
如果每一元素代表 3 位一組同學，則樣本空間有 $\frac{50\times 49\times 48}{3\times 2}=20,600$ 個元素。

排列與組合

排列與組合是兩種用以計算樣本空間之元素個數的技術。
因爲有不同的抽樣方法，故有各種不同的計算花樣。

（階乘）定義

$$x!=\begin{cases}1 & x=0\\ x\times(x-1)! & x=1,2,\ldots\end{cases}$$

$$0!=1$$
$$1!=1$$
$$4!=4\times 3\times 2\times 1=24$$
$$\frac{7!}{4!}=7\times 6\times 5$$

排列（*permutation*）

從 n 個不同的物件中取 m 個排成一列（次序有關係），則有 P_m^n 種排法

$$P_m^n=\underbrace{n\times(n-1)\times\cdots\times(n-m+1)}_{m個}=\frac{n!}{(n-m)!}$$

組合（*combination*）

從 n 個不同的物件中取出 m 個成一組（次序沒有關係），則有 C_m^n 種組法

$$C_m^n=\frac{P_m^n}{m!}=\frac{n\times(n-1)\times\cdots\times(n-m+1)}{m!}=\frac{n!}{m!(n-m)!}$$

$$C_m^n=C_{n-m}^n$$
$$C_m^n=\frac{n}{m}\times C_{m-1}^{n-1}$$
$$C_m^n=C_m^{n-1}+C_{m-1}^{n-1}$$

重複排列

從 n 個不同的物件選取 m 次（選後放回，次序有關），則有 $n^m=\underbrace{n\times\cdots\times n}_{m個}$ 種選法。

重複組合

從 n 個不同的物件選取 m 次（選後放回，次序無關），則有 H_m^n 種選法

$$H_m^n = C_m^{n+m-1} = \frac{(n+m-1)!}{(n-1)!m!}$$

二項式定理與多項式定理

$$(x+y)^n = \sum_{m=0}^{n} C_m^n x^m y^{n-m}$$

$$(x+y+z)^n = \sum_{\substack{p+q+r=n \\ p,q,r=0,\ldots,n}} \frac{n!}{p!q!r!} x^p y^q z^r$$

範例 2.8　樣本空間元素個數

梭哈遊戲爲從一副撲克牌（52 張）中取 5 張爲一手牌，然後以手牌中花樣出現的機率大小來決定勝負。梭哈遊戲手牌的所有可能組合個數爲

$$C_5^{52} = \frac{52!}{5! \times 47!} = 2,598,960$$

大樂透從 49 個號碼中選 6 個，因此頭彩是以下所有組合之一

$$C_6^{49} = 13,983,816$$

樂透彩從 38 個號碼中選 6 個，一定要中頭彩的話，只要買

$$C_6^{38} = 2,760,681$$

張彩券就可以了。四星彩正彩由四個阿拉伯數字組成（可重複、次序有關），一共有

$$10^4 = 10000$$

種組合。

加法律與乘法律

若抽樣需分幾個階段進行，計算樣本空間元素個數會稍複雜。
同一階段中，不同狀況的次數應加總計算，是爲**加法律**；
不同階段間，應將各階段的次數相乘以得樣本空間元素個數，是爲**乘法律**。

範例 2.9　加法律與乘法律

假設從甲地到乙地的交通工具有 3 種火車，4 種公車，

(1)則從甲地到乙地一共有$3+4=7$種方式（加法律）；

(2)來回一共有$(3+4)\times(3+4)=49$種方式（階段內加法律、階段間乘法律）；

(3)若甲地到乙地需先搭火車、然後再搭公車，則一共有$3\times4=12$種方式（乘法律）。

範例 2.10　加法律與乘法律

排列公式的邏輯：分 m 個階段選出物件，第 1 階段有 n 個可以挑，第 2 階段有$n-1$個可以挑選，直到第 m 階段剩下$n-m+1$個可以挑，因此共有

$$P_m^n = n\times(n-1)\times\cdots\times(n-m+1)$$

種挑選法（乘法律）。

範例 2.11　加法律與乘法律

五張同花的梭哈牌有幾種？

分兩個階段決定五張同花牌：第一階段決定花色，第二階段決定數字。前者有 4 種花色，後者爲 13 個數字中選 5 個，因此

$$\text{五張同花的手牌數目} = 4\times C_5^{13} = 4\times1,287 = 5,148$$

富豪（full-house，三張同數字配上另兩張同數字）的梭哈牌有幾種？

分兩個大階段進行：先挑出三張同數字，再挑出另兩張同數字。每個大階段又分成兩個小階段：先決定三條（一對）的數字，再由該數字的 4 張牌中挑出 3 張（或 2 張）。因此

$$\text{富豪的手牌數目} = \left[C_1^{13}\times C_3^4\right]\times\left[C_1^{12}\times C_2^4\right] = 13\times4\times12\times6 = 3,744$$

順子的梭哈牌有幾種？

分兩個階段進行：先決定順子的最小數字（A 到 10 共 10 個），然後決定各個數字的花色（每個數字都可以有 4 種花色）。因此

$$\text{順子的手牌數目} = 10\times\left(4^5\right) = 10,240$$

2.4 事件的性質與計數法則

事件 A 的機率（相對次數）:

$$P(A)=\frac{n(A)}{n(\Omega)}$$

其中，Ω爲樣本空間 ，$A\subseteq\Omega$，
$n(\Omega)$爲樣本空間元素個數，$n(A)$爲集合（事件）A之元素個數 。

範例 2.12 事件的機率

分別求梭哈牌中取得富豪、同花、順子的機率？

令 A_1、A_2、A_3 分別爲取得富豪、同花、順子的事件。由之前範例的討論，得知

$$P(A_1)=\frac{n(A_1)}{n(\Omega)}=\frac{3,744}{2,598,960}$$

$$P(A_2)=\frac{n(A_2)}{n(\Omega)}=\frac{5,148}{2,598,960}$$

$$P(A_2)=\frac{n(A_2)}{n(\Omega)}=\frac{10,240}{2,598,960}$$

交集事件的機率

聯合機率（*joint probability*）

$P(A\cap B)$爲 A、B 兩事件交集的機率，也稱爲聯合機率。

互斥（*exclusive*）

兩事件 A、B，若 $A\cap B=\varnothing$，則稱爲此兩事件互斥。

互斥事件的聯合機率爲零。

$$\begin{aligned} A\text{與}B\text{互斥} \quad &\equiv \quad A\cap B=\varnothing \\ &\equiv \quad P(A\cap B)=0 \end{aligned}$$

範例 2.13　聯合機率與列聯表

就下列列聯表：

年資＼城市	台北	台中	高雄	邊際次數
1~2	2	2	2	6
3~5	5	3	5	13
6~9	3	2	6	11
10 以上	3	0	1	4
邊際次數	13	7	14	34

令 A_1、A_2、A_3 分別表示住台北、台中、高雄的事件，B_1、B_2、B_3、B_4 分別表示四種不同年資的事件。則

$$P(A_2 \cap A_3) = P(\text{住台中且住高雄}) = P(\varnothing) = 0$$

$$P(A_3 \cap B_3) = P(\text{住高雄且年資6到9年}) = \frac{6}{34}$$

因分組有互斥且周延的要求，故 A_1、A_2、A_3 相互間互斥，B_1、B_2、B_3、B_4 亦同。

條件機率

條件機率（*conditional probability*）

$P(A|B)$ 爲事件 B 發生的條件下，事件 A 的機率，其定義如下

$$P(A|B) = \frac{n(A \cap B)}{n(B)} = \frac{P(A \cap B)}{P(B)}$$

若條件事件爲樣本空間Ω，則條件機率即爲其絕對機率

$$P(A|\Omega) = \frac{n(A \cap \Omega)}{n(\Omega)} = \frac{n(A)}{n(\Omega)} = P(A)$$

注意：$P(A|B)$ 是有關事件 A 的機率；$P(B|A)$ 是有關事件 B 的機率。

獨立（*independent*）

若 $P(A|B) = P(A)$，則稱 A、B 兩事件獨立。

$$\begin{aligned} A\text{與}B\text{獨立} \quad &\equiv \quad P(A|B) = P(A) \\ &\equiv \quad P(B|A) = P(B) \\ &\equiv \quad P(A \cap B) = P(A)P(B) \end{aligned}$$

邊際機率（*margin probability*）

若 $B_1, B_2, \ldots, B_n$ 互斥且周延，則

$$P(B_1)P(A|B_1)+P(B_2)P(A|B_2)+\cdots+P(B_n)P(A|B_n)=P(A|\Omega)=P(A)$$

爲 A 之邊際機率。

範例 2.14　條件機率、邊際機率與列聯表

就下列列聯表：

年資＼城市	台北	台中	高雄	邊際次數
1~2	2	2	2	6
3~5	5	3	5	13
6~9	3	2	6	11
10 以上	3	0	1	4
邊際次數	13	7	14	34

令 B_1、B_2、B_3 分別表示住台北、台中、高雄的事件，A_1、A_2、A_3、A_4 分別表示四種不同年資的事件。則

$$P(A_3|B_3)=\frac{n(A_3\cap B_3)}{n(B_3)}=\frac{6}{14}$$　（只看 B_3 = 高雄 那一行）

$$P(B_3|A_3)=\frac{n(B_3\cap A_3)}{n(A_3)}=\frac{6}{11}$$　（只看 $A_3 = 6\sim 9$ 歲那一列）

$$P(B_3)=\frac{14}{34}$$　（只看最後一行的邊際次數）

$$P(A_3)=\frac{11}{34}$$　（只看最後一列的邊際次數）

聯集事件的機率

因

$$n(A\cup B)=n(A)+n(B)-n(A\cap B)$$

故

$$P(A\cup B)=P(A)+P(B)-P(A\cap B)$$

補集與機率

$$\overline{A\cap B}=\bar{A}\cup\bar{B} \qquad \overline{A\cup B}=\bar{A}\cap\bar{B}$$

$$P\left(\bar{A}\right)=1-P(A)$$

$$P\left(\overline{A\cap B}\right)=P\left(\bar{A}\cup\bar{B}\right)=1-P(A\cap B) \qquad P\left(\overline{A\cup B}\right)=P\left(\bar{A}\cap\bar{B}\right)=1-P(A\cup B)$$

機率加法法則

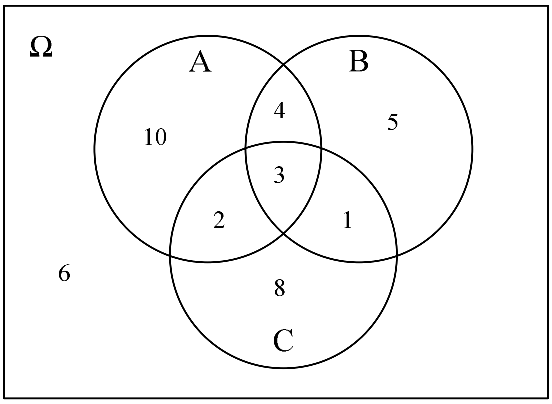
若 A 與 B **互斥**，則 $P(A\cap B)=0$ ，因此 $P(A\cup B)=P(A)+P(B)$

機率乘法法則

若 A 與 B **獨立**，則 $P(A|B)=P(A)$ 、$P(B|A)=P(B)$ ，因此 $P(A\cap B)=P(A)P(B)$

範例 2.15　事件的機率

就下列資料：



直接計算可得：

$$n(A)=19,\quad n(B)=13,\quad n(C)=14,\quad n(\Omega)=39$$

以下爲事件機率的計算：

$$P(A)=P(A|\Omega)=\frac{n(A)}{n(\Omega)}=\frac{19}{39}$$

$$P(A|B)=\frac{n(A\cap B)}{n(B)}=\frac{7}{13}$$

$$P(A|B\cap C)=\frac{n(A\cap B\cap C)}{n(B\cap C)}=\frac{3}{4}$$

$$P\left(A\cap C\middle|B\cup C\right)=\frac{n\left(\left(A\cap C\right)\cap\left(B\cup C\right)\right)}{n\left(B\cup C\right)}=\frac{n\left(A\cap C\right)}{n\left(B\cup C\right)}=\frac{5}{23}$$

2.5 貝氏定理

分割（*partition*）：

若 $A_1, A_2, \ldots, A_n$ 互斥且周延，則稱 $A_1, A_2, \ldots, A_n$ 為 Ω 的一組分割；
也稱 $A_1, A_2, \ldots, A_n$ 分割 Ω。

次數分配表、列聯表中的分組必須是母體（或樣本空間）的一組分割。

總機率法則

若 $B_1, B_2, \ldots, B_n$ 為 Ω 的一組分割，則

$$\begin{aligned}P(A)&=P(B_1\cap A)+P(B_2\cap A)+\cdots+P(B_n\cap A)\\&=P(B_1)P(A|B_1)+P(B_2)P(A|B_2)+\cdots+P(B_n)P(A|B_n)\end{aligned}$$

貝氏定理（Bayes' Rule）

若 $B_1, B_2, \ldots, B_n$ 為 Ω 的一組分割，且 $P(A)\neq 0$，則

$$P\left(B_j\middle|A\right)=\frac{P\left(B_j\cap A\right)}{P(A)}=\frac{P(B_j)P(A|B_j)}{P(B_1)P(A|B_1)+P(B_2)P(A|B_2)+\cdots+P(B_n)P(A|B_n)}$$

注意：貝氏定理等號左邊是 A 條件下 B_j 的機率，右邊則是 A 的條件機率與 B_j 的機率。

事前機率（*prior probability*）與**事後機率**（*posterior probability*）：

貝氏定理中等號右邊的 $P\left(A\middle|B_j\right)$ 稱為事前機率，左邊的 $P\left(B_j\middle|A\right)$ 則稱為事後機率。

貝氏定理一定涉及兩種分割，一個為已知的條件機率（事前機率，$A_1, A_2, \ldots, A_n$ 之條件機率），另一個為需要求解的事後機率（$B_1, B_2, \ldots, B_n$ 之條件機率）。後者之分割的絕對機率則為已知（$B_1, B_2, \ldots, B_n$ 之機率）。

範例 2.16 貝氏定理

某公司某產品由三個工廠生產，產量分別為 20%、30%、50%；三工廠的該產品的不良率分別為 2%、3%、1%。若發現不良品，計算該不良品由第一個工廠生產的機率。

【解】

兩種分割，

事前機率分組：『工廠一、工廠二、工廠三』

事後機率分組：『正常品、不良品』

已知

$$P\left(B_1\right)=0.2,\ P\left(B_2\right)=0.3,\ P\left(B_3\right)=0.5$$
$$P\left(A_2\middle|B_1\right)=0.02,\ P\left(A_2\middle|B_2\right)=0.03,\ P\left(A_2\middle|B_3\right)=0.01$$

整理成以下工作表

事前機率 $P(行_j \mid 列_i)$、$P(列_i)$

	正常品	不良品	邊際機率
工廠一	98.00%	2.00%	20%
工廠二	97.00%	3.00%	30%
工廠三	99.00%	1.00%	50%

作以下兩工作表的計算

聯合機率 $P(列_i \cap 行_j)$

	正常品	不良品
工廠一	0.1960	0.0040
工廠二	0.2910	0.0090
工廠三	0.4950	0.0050
邊際機率	0.9820	0.0180

其中，$98\% \times 20\% = 0.1960$ 、 $97\% \times 30\% = 0.2910$ 。

事後機率 $P(列_i \mid 行_j)$、$P(行_j)$

	正常品	不良品
工廠一	0.1996	0.2222
工廠二	0.2963	0.5000
工廠三	0.5041	0.2778
邊際機率	0.9820	0.0180

其中，$\dfrac{0.1960}{0.9820} = 0.1996$ 、 $\dfrac{0.0040}{0.0180} = 0.2222$ 。

不良品由第一個工廠生產的機率爲 $P\left(\text{工廠一}\middle|\text{不良品}\right)=0.2222$ 。

第三章　機率分配

2006 年 10 月 12 日 最後修改

3.1　機率分配概論
3.2　離散隨機變數之機率分配
3.3　連續隨機變數之機率分配
3.4　機率表
3.5　計算隨機變數的期望值與變異數
3.6　動差與動差母函數

3.1　機率分配概論

機率分配（*probability distribution*）

機率分配是隨機變數的機率函數。

機率分配是一個函數，其定義域是實數值，值域是機率（$\left[0,1\right]$）。

機率分配可以想像爲相對次數分配表，只是這裡分組是隨機變數的值。（所以，會有很多，或無限多個分組。）

以下是一個機率分配：

x	*P(x)*
0	1/16
1	4/16
2	6/16
3	4/16
4	1/16
	1

其中， $P\left(x\right)=C_x^4\times\left(\frac{1}{2}\right)^4,\quad x=0,1,\ldots,4$

期望值（*expectation value*），**變異數**（*variance*）

機率分配的期望值如同相對次數分配表的平均數，變異數亦同。

隨機變數 X 的期望值與變異數分別寫成 $E\left(X\right)$ 、$V\left(X\right)$ 。

$$E\left(X\right)=\mu_x=\sum_{\forall x}xP\left(x\right)$$

$$V\left(X\right)=E\left[\left(x-\mu_x\right)^2\right]=\sum_{\forall x}\left(x-\mu_x\right)^2P\left(x\right)=\sum x^2P\left(x\right)-\mu_x^2$$

或

$$E(X)=\int_x xP(x)=\int_x xf(x)dx$$
$$V(X)=E\left[(x-\mu_x)^2\right]=\int_x (x-\mu_x)^2 f(x)dx=\int_x x^2 f(x)dx-\mu_x^2$$

其中，$P(x)$爲離散隨機變數X的機率函數，$f(x)$是連續隨機變數X的機率函數。

共變數（*covariance*）

如同敘述統計中的共變數，以$\operatorname{cov}(X,Y)$表示隨機變數X、Y的共變數。

$$\operatorname{cov}(X,Y)=E\left[(x-\mu_x)(y-\mu_y)\right]=\sum_{\forall x,y} xyP(x,y)-\mu_x\mu_y$$

或

$$\operatorname{cov}(X,Y)=E\left[(x-\mu_x\mu_x)(y-\mu_y)\right]=\iint_{x,y} xyf(x,y)dxdy-\mu_x\mu_y$$

期望值、變異數、共變數的希臘字母符號

$$\mu_X=E(X)$$
$$\sigma_X^2=V(X)=E(X^2)-\left[E(X)\right]^2$$
$$\sigma_{X,Y}=\operatorname{cov}(X,Y)=E(XY)-E(X)E(Y)$$

$\Sigma\bullet$ 、 $E(\bullet)$ 、 $V(\bullet)$ 、 $\operatorname{cov}(\bullet,\bullet)$ 的操作

$$\Sigma x_i=x_1+x_2+\cdots+x_n$$
$$\Sigma af(x_i)=af(x_1)+af(x_2)+\cdots+af(x_n)=a\Sigma f(x_i)$$
$$\Sigma\left[f(x_i)+g(x_i)\right]=\Sigma f(x_i)+\Sigma g(x_i)$$

$$E(a)=a$$
$$E(aX)=aE(X)$$
$$E\left[f(X)+g(X)\right]=E\left[f(X)\right]+E\left[g(X)\right]$$
$$E(aX+bY)=aE(X)+bE(Y)$$
$$E\left[(X-\mu_X)^2\right]=E\left[X^2-2\mu_X X+\mu_X^2\right]=E(X^2)-2\mu_X E(X)+\mu_X^2=E(X^2)-\mu_X^2$$
$$E\left[(X-\mu_X)(Y-\mu_Y)\right]=E(XY)-\mu_X\mu_Y$$

$$V(X)=E\left[(X-\mu_X)^2\right]$$
$$V(a)=0$$
$$V(aX)=a^2V(X)$$
$$V\left[f(X)+g(X)\right]=V(f+g)=V(f)+V(g)+2E\left[(f-\mu_f)(g-\mu_g)\right]$$
$$V(aX+bY)=a^2V(X)+b^2V(Y)+2ab\,\mathrm{cov}(X,Y)$$

$$\mathrm{cov}(X,Y)=E\left[(X-\mu_X)(Y-\mu_Y)\right]$$
$$\mathrm{cov}(a,X)=0$$
$$\mathrm{cov}(aX,bY)=ab\,\mathrm{cov}(X,Y)$$
$$\mathrm{cov}(X+Y,Z)=\mathrm{cov}(X,Z)+\mathrm{cov}(Y,Z)$$

範例 3.1　$E(\bullet)$ 的操作

已知$E(X)=8$、$E(Y)=6$，求$E(3X-2Y)$。

【解】

$E(3X-2Y)=3E(X)-2E(Y)=3\times 8-2\times 6=12$

範例 3.2　$V(\bullet)$ 的操作

已知$V(X)=8$、$V(Y)=6$，且 X 與 Y 互相獨立，求$V(3X-2Y)$。

【解】

X 與 Y 互相獨立隱含 $\mathrm{cov}(X,Y)=0$，因此

$$V(3X-2Y)=3^2V(X)+(-2)^2V(Y)=9\times 8+4\times 6=96$$

3.2　離散隨機變數之機率分配

討論機率分配應注意事項：

(1)有哪些參數，參數的意義爲何

(2)隨機變數的意義爲何

(3)分配的機率函數（機率密度函數）爲何

(4)隨機變數的期望值、變異數爲何

我們將介紹 7 個離散機率分配：

(1)白努力分配

(2)二項分配

(3)超幾何分配

(4)卜瓦松分配

(5)多項分配

(6)負二項分配

(7)幾何分配

白努力分配

白努力分配（*Bernoulli distribution*）

白努力分配又稱爲**點二項分配**（*point binomial distribution*）。

白努力分配是所有機率分配的基礎。

參數：每次實驗成功的機率 p，每個樣本作 1 次實驗

隨機變數：樣本中有 x 次成功

機率函數、期望值與變異數：

$$P(x) = f(x; p) = p^x (1-p)^{1-x}, \qquad x \in \{0,1\},\ 0 < p < 1$$
$$E(x) = p$$
$$V(x) = p(1-p)$$

二項分配

二項分配（*binomial distribution*）

二項分配是複數樣本的白努力分配。

參數：每次實驗成功的機率 p，每個樣本作 n 次實驗

隨機變數：樣本中有 x 次成功

機率函數：

$$P(x) = f(x; n, p) = C_x^n p^x (1-p)^{n-x}, \qquad x \in \{0,1,\ldots,n\},\ 0 < p < 1$$
$$E(x) = np$$
$$V(x) = np(1-p)$$

二項分配與白努力分配的關係

若 $X_1, X_2, \ldots, X_n$ 為相同且互相獨立的白努力分配，令

$$X = X_1 + X_2 + \cdots + X_n$$

則 X 為二項分配。

範例 3.3　二項分配

若有 10%的網友謊報其性別，則 10 個網友中，

(1)恰好有 2 位謊報性別的機率為何？

(2)沒有人謊報的機率？

(3)2 位以上謊報的機率為何？

【解】

$p = 10\% = 0.1,\ n = 10,$ 二項分配

(1)

$$P(x=2) = C_2^{10} \times 0.1^2 \times 0.9^8 = \frac{10\times 9}{2!} \times 0.01 \times 0.43047 = 0.1937$$

(2)

$$P(x=0) = C_0^{10} \times 0.1^0 \times 0.9^{10} = 1\times 1\times 0.3487 = 0.3487$$

(3)

$$P(x>2) = 1 - P(x=0) - P(x=1) - P(x=2) = 1 - 0.3487 - 0.3874 - 0.1937 = 0.0702$$

超幾何分配

超幾何分配（*hyper-geometric distribution*）

如同二項分配，只是成功機率在抽取樣本的過程會變動。

參數：一共有 N 個球，其中有 S 球為『成功』的球，每次（每樣本）抽出 n 個球

隨機變數：樣本中有 x 個『成功』球

機率函數：

$$P(x)=f(x;n,N,S)=\frac{C_x^S C_{n-x}^{N-S}}{C_n^N},\qquad x\in\{0,1,\ldots,n\},\ 0\le S\le N$$

$$E(x)=n\frac{S}{N}$$

$$V(x)=n\frac{S}{N}\left(1-\frac{S}{N}\right)\frac{N-n}{N-1}$$

超幾何分配與二項分配的關係

抽取後『不放回』與抽取後『放回』：

超幾何分配是『不放回』的情況。

若抽取後『放回』，則抽到『成功』的機率都是 $p=\frac{S}{N}$，故與二項分配相同。

當 $N\to\infty$ 的時候：

若 $N\to\infty$，則每次抽到『成功』的機率都是 $p=\frac{S}{N}$，故與二項分配相同。

$$\begin{aligned}
\lim_{N\to\infty}\frac{C_x^S C_{n-x}^{N-S}}{C_n^N} &= \lim_{N\to\infty}\frac{S!}{x!(S-x)!}\times\frac{(N-S)!}{(n-x)!(N-S-n+x)!}\times\frac{n!(N-n)!}{N!}\\
&= \lim_{N\to\infty}\frac{n!}{x!(n-x)!}\times\frac{\overbrace{S(S-1)\cdots}^{x\text{ 項}}\times\overbrace{(N-S)(N-S-1)\cdots}^{n-x\text{ 項}}}{\underbrace{N(N-1)\cdots}_{n\text{ 項}}}\\
&= \frac{n!}{x!(n-x)!}\times\frac{S^x(N-S)^{n-x}}{N^n}\\
&= C_x^n\left(\frac{S}{N}\right)^x\left(1-\frac{S}{N}\right)^{n-x}
\end{aligned}$$

超幾何分配可視爲 $p=\frac{S}{N}$ 的二項分配，只是變異數需加上修正項 $\frac{N-n}{N-1}$：

$$E(x)=np=n\frac{S}{N},\qquad V(x)=np(1-p)\frac{N-n}{N-1}=n\frac{S}{N}\left(1-\frac{S}{N}\right)\frac{N-n}{N-1}$$

範例 3.4 超幾何分配

某一 12 人的旅行團中，有 8 個人睡覺會打鼾，若 3 人一個房間，則某一房間

(1)恰好有 2 人會打鼾的機率爲何？

(2)沒有人打鼾的機率？

【解】

$N=12,\ S=8,\ n=3,$ 超幾何分配

(1)

$$P(x=2)=\frac{C_2^8 C_1^4}{C_3^{12}}=\frac{\frac{8\times7}{2!}\times\frac{4}{1}}{\frac{12\times11\times10}{3!}}=\frac{8\times7\times2}{2\times11\times10}=\frac{28}{55}$$

(2)

$$P(x=0)=C_0^{10}\times0.1^0\times0.9^{10}=1\times1\times0.3487=0.3487$$

卜瓦松分配

卜瓦松分配（*Poisson distribution*）

卜瓦松分配爲每組樣本中實驗次數非常多（$n\to\infty$）的二項分配。

參數：每組樣本的平均成功次數爲λ

隨機變數：樣本中有 x 次成功

機率函數：

$$P(x)=f(x;\lambda)=\frac{\lambda^x e^{-\lambda}}{x!},\qquad x\in\{0,1,\ldots\},\ \lambda>0$$
$$E(x)=\lambda$$
$$V(x)=\lambda$$

卜瓦松分配與二項分配的關係

二項分配的極限狀況：

$n\to\infty$ 、 $p\to0$ 、且 $np=\lambda$ （λ爲常數）

在極限狀況下，二項分配會趨近於卜瓦松分配。

$$\begin{aligned}\lim_{n\to\infty}C_x^n p^x(1-p)^{n-x} &= \lim_{n\to\infty}\frac{n!}{x!(n-x)!}\left(\frac{\lambda}{n}\right)^x\left(1-\frac{\lambda}{n}\right)^{n-x}\\ &= \lim_{n\to\infty}\frac{\lambda^x}{x!}\frac{\overbrace{n\times(n-1)\times\cdots\times(n-x+1)}^{x\text{ 項}}}{n^x}\left(1-\frac{\lambda}{n}\right)^n\left(1-\frac{\lambda}{n}\right)^{-x}\\ &= \frac{\lambda^x}{x!}e^{-\lambda}\end{aligned}$$

其中

$$\lim_{n\to\infty}\frac{\overbrace{n\times(n-1)\times\cdots\times(n-x+1)}^{x\text{ 項}}}{n^x}=1$$

$$\lim_{n\to\infty}\left(1-\frac{\lambda}{n}\right)^n=e^{-\lambda}$$

$$\lim_{n\to\infty}\left(1-\frac{\lambda}{n}\right)^{-x}=1$$

範例 3.5　卜瓦松分配

若 0.15%的車子有保颱風洪水險，則 100 輛泡水車中，

(1)恰好有 2 輛有保洪水險的機率為何？

(2)沒有車輛保洪水險的機率？

【解】

$p=0.15\%=0.0015,\ n=100,$ 二項分配

(1)

$$P(x=2)=C_2^{100}\times0.0015^2\times0.9985^{98}=\cdots$$

(2)

$$P(x=0)=C_0^{100}\times0.0015^0\times0.9985^{100}=\cdots$$

或

$\lambda=np=100\times0.15\%=0.15,$ 卜瓦松分配

(1)

$$P(x=2)=\frac{0.15^2}{2!}e^{-0.15}=0.01125\times0.8607=0.00968$$

(2)

$$P(x=0)=\frac{0.15^0}{0!}e^{-0.15}=e^{-0.15}=0.8607$$

範例 3.6　卜瓦松分配

某打字員平均每兩頁出現一次錯誤，則該員在 5 頁的文稿中，

(1)恰好有 2 個錯誤機率為何？

(2)都沒有錯誤的機率？

【解】

$\lambda=\frac{1}{2}\times5=2.5,$ 卜瓦松分配

(1)

$$P(x=2)=\frac{2.5^2}{2!}e^{-2.5}=3.125\times 0.0821=0.2565$$

(2)

$$P(x=0)=\frac{2.5^0}{0!}e^{-2.5}=e^{-2.5}=0.0821$$

多項分配

多項分配（*multinomial distribution*）

多項分配是二項分配的延伸，每次實驗有兩個以上的出象。

參數：每次實驗有 m 種出象，出現機率分別為 $p_1, p_2, \ldots, p_m$ ，每個樣本作 n 次實驗

隨機變數：每組樣本中第 i 種出象出現有 x_i 次， $i=1,2,\ldots,m$

機率函數：

$$P(x_1,x_2,x_3)=f(x_1,x_2,x_3;n,p_1,p_2,p_3)=\frac{n!}{x_1!x_2!x_3!}p_1^{x_1}p_2^{x_2}p_3^{x_3}$$

$$E(x_1)=np_1,\qquad E(x_2)=np_2,\qquad E(x_3)=np_3$$

$$V(x_1)=np_1(1-p_1),\quad V(x_2)=np_2(1-p_2),\quad V(x_3)=np_3(1-p_3)$$

其中

$$x_1,x_2,x_3\in\{0,1,\ldots,n\},\quad x_1+x_2+x_3=n,\quad 0<p_1,p_2,p_3<1,\quad p_1+p_2+p_3=1$$

負二項分配

負二項分配（*negative binomial distribution*）

負二項分配與二項分配互為對偶關。

二項分配以每組樣本中之實驗次數為參數，成功次數為隨機變數；

負二項分配以每組樣本中之實驗次數為隨機變數，成功次數為參數。

（最後一次實驗的出象一定是成功）

參數：每次實驗成功的機率 p，每組樣本中有 r 次成功實驗

隨機變數：樣本中有 x 次實驗

機率函數：

$$P(x)=f(x;r,p)=C_{r-1}^{x-1}p^r(1-p)^{x-r},\qquad x\in\{r,r+1,\ldots\},\ 0<p<1$$

$$E(x)=\frac{r}{p}$$

$$V(x)=\frac{r(1-p)}{p^2}$$

【注意】另一種隨機變數的定義：第 r 次成功之前作了 x 次實驗。

幾何分配

幾何分配（*geometric distribution*）

幾何分配爲負二項分配的特殊狀況（第一次成功）。

幾何分配與白努力分配互爲對偶關係。

參數：每次實驗成功的機率 p，每組樣本中有 1 次成功實驗

隨機變數：樣本中有 x 次實驗

機率函數：

$$P(x)=f(x;p)=p(1-p)^{x-1},\qquad x\in\{1,2,\ldots\},\ 0<p<1$$

$$E(x)=\frac{1}{p}$$

$$V(x)=\frac{1-p}{p^2}$$

範例 3.7 幾何分配與負二項分配

假設在某廟裡擲出『笑杯』的機率爲 0.4，則連續擲杯 5 次後，

(1)出現第一個笑杯的機率爲何？

(2)出現第三個笑杯的機率？

【解】

(1)

$p=0.4,$ 幾何分配

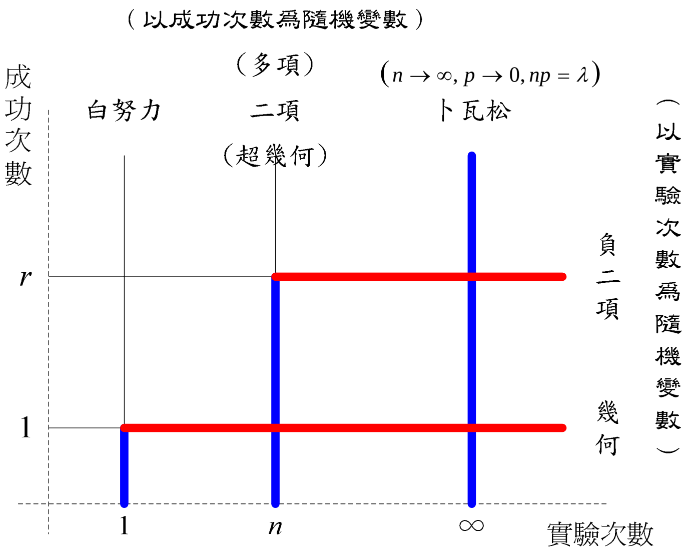
$$P(x=5)=0.4\times0.6^4=0.0518$$

(2)

$p=0.4,\ n=3,$ 負二項分配

$$P(x=5)=C_2^4\times0.4^3\times0.6^2=6\times0.064\times0.36=0.1382$$

離散分配間之關係



3.3　連續隨機變數之機率分配

我們將介紹 3 個連續機率分配：

(1)均等分配

(2)指數分配

(3)常態分配

機率函數與機率密度函數

機率密度函數（*probability density function*）：

若連續隨機變數 X 之機率密度函數為 $f(x)$，則

機率函數 $P(x) = f(x)dx$；

累計機率函數 $F(a) = P(x \le a) = \int_{-\infty}^{a} f(x)dx$。

均等分配

均等分配（*uniform distribution*）

每個點出現的機率均相等稱為均等分配。

參數與隨機變數：$x \in [a,b]$ 且出現的機率皆相等

機率密度函數：

$$f(x)=f(x;a,b)=\frac{1}{b-a},\qquad a\le x\le b$$

$$E(x)=\frac{a+b}{2}$$

$$V(x)=\frac{(b-a)^2}{12}$$

範例 3.8　均等分配

公車每 30 分鐘來一班，假設乘客不看時刻表，隨機來等車，則

(1)平均等候時間爲何？

(2)等候時間的標準差爲何？

(3)等候超過 10 分鐘的機率？

【解】

$a=0,\ b=30,$ 均等分配

(1)

$$E(x)=\frac{a+b}{2}=\frac{0+30}{2}=15$$

(2)

$$\sigma=\sqrt{V(X)}=\sqrt{\frac{(b-a)^2}{12}}=\sqrt{\frac{30^2}{12}}=8.66$$

(3)

$$P(x>10)=\int_{10}^{30}\frac{1}{30}dx=\frac{1}{30}x\Big|_{10}^{30}=\frac{20}{30}=\frac{2}{3}$$

指數分配

指數分配（*exponential distribution*）

指數分配與卜瓦松分配互爲對偶。

兩者有相同的參數：每個樣本的平均成功次數爲 λ。

兩者隨機變數互爲對偶：

卜瓦松分配：以樣本的成功次數爲隨機變數，

指數分配：以兩成功實驗間的時間間隔爲隨機變數。

（成功次數多則時間間隔小，反之亦然。）

參數：單位時間（每個樣本）的平均成功次數爲 λ

隨機變數：經過時間 x 後產生第一次成功

機率函數：

$$f(x)=f(x;\lambda)=\lambda e^{-\lambda x}, \qquad x>0,\lambda>0$$
$$E(x)=\frac{1}{\lambda}$$
$$V(x)=\left(\frac{1}{\lambda}\right)^2$$
$$F(a)=P(x\le a)=1-e^{-\lambda a}$$

指數分配的另一種表示法：

參數：平均兩次成功間的時間間隔爲 β（即 $\lambda=1/\beta$）

隨機變數：經過時間 x 後產生第一次成功

機率函數：

$$f(x)=f(x;\beta)=\frac{1}{\beta}e^{-\frac{1}{\beta}x}, \qquad x>0,\beta>0$$
$$E(x)=\beta$$
$$V(x)=\beta^2$$
$$F(a)=P(x\le a)=1-e^{-\frac{1}{\beta}a}$$

範例 3.9 指數分配

某路段平均每年發生 10 件車禍，則

(1)平均多久發生一次車禍？

(2)某個月內沒有發生任何車禍的機率爲何？

【解】

$\lambda=10$, 指數分配

(1)

$$E(x)=\frac{1}{\lambda}=\frac{1}{10}\ (\text{年})$$

(2)

$$P(x>\text{一個月})=P\left(x>\tfrac{1}{12}\text{年}\right)=1-F\left(\tfrac{1}{12}\right)=e^{-10\times\frac{1}{12}}=0.4346$$

常態分配

常態分配（*normal distribution*）

常態分配爲抽樣分配的共同極限分配。

參數：平均數 μ ，標準差 σ

機率密度函數：

$$f(x)=f(x;\mu,\sigma)=\frac{1}{\sigma\sqrt{2\pi}}e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2},\qquad x\in R,\ \sigma>0,\ \mu\in R$$

$$E(x)=\mu$$

$$V(x)=\sigma^2$$

標準常態分配（*normal distribution*）

標準常態分配爲平均數 $\mu=0$ ，標準差 $\sigma=1$ 的常態分配。

$$f(z)=\frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}z^2},\qquad z\in R$$

$$E(z)=0$$

$$V(z)=1$$

將常態分配標準化

若 X 爲參數 μ 、σ 的常態分配，Z 爲標準常態分配，則

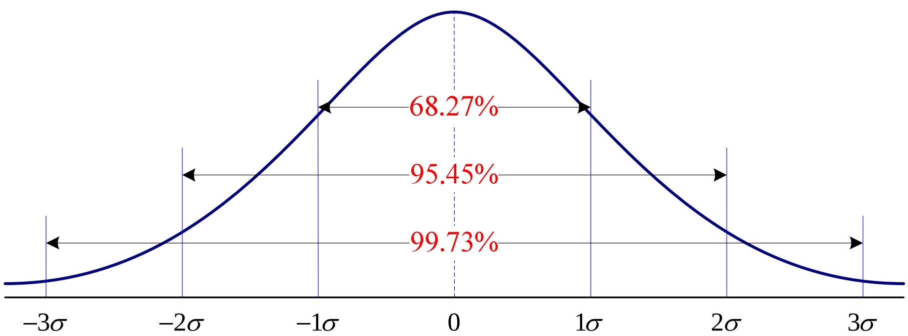
$$z=\frac{x-\mu}{\sigma} \quad 或 \quad x=\mu+z\sigma$$

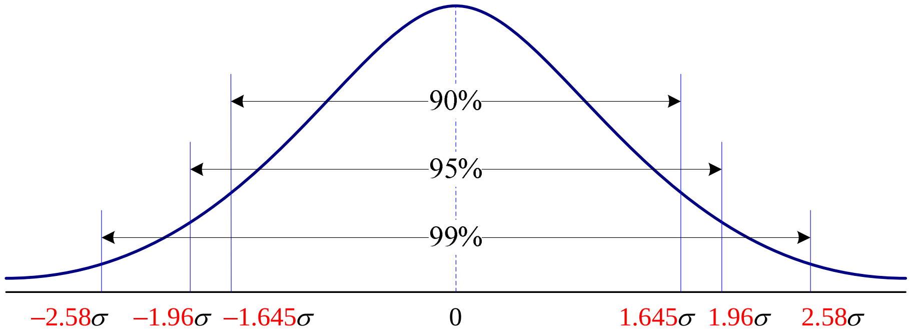
有關常標準分配的幾個重要數值：

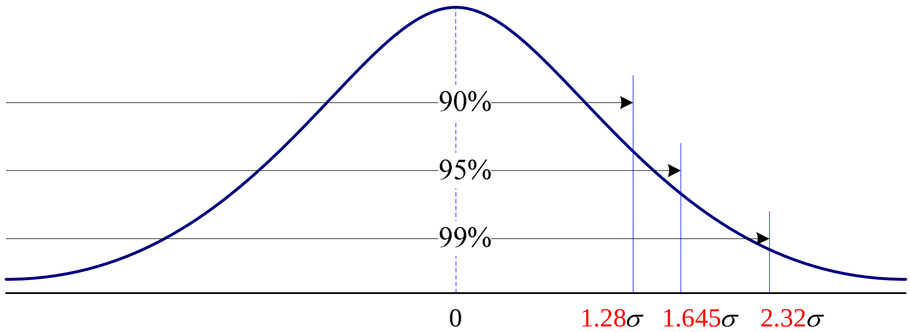
$$P(-1\le z\le 1)=0.6827=68.27\% \qquad P(-1.645\le z\le 1.645)=0.9 \qquad P(z\le 1.282)=0.90$$

$$P(-2\le z\le 2)=0.9545=95.45\% \qquad P(-1.96\le z\le 1.960)=0.95 \qquad P(z\le 1.645)=0.95$$

$$P(-3\le z\le 3)=0.9973=99.73\% \qquad P(-2.58\le z\le 2.576)=0.99 \qquad P(z\le 2.326)=0.99$$







範例 3.10　常態分配

已知 x 爲 $\mu=12$ 、 $\sigma=4$ 的常態分配，則

(1)求 $P(8\le X\le 16)$ ；

(2)求 a 使 $P(X\le a)=0.9$ 。

【解】

(1)

$$z_1^*=\frac{8-\mu}{\sigma}=\frac{8-12}{4}=-1,\ z_2^*=\frac{16-\mu}{\sigma}=\frac{16-12}{4}=1$$

$$P(8\le X\le 16)=P\left(z_1^*\le z\le z_2^*\right)=P(-1\le z\le 1)=0.6827$$

(2)

$$P(z\le 1.282)=0.9 \quad \Rightarrow \quad z^*=1.282$$

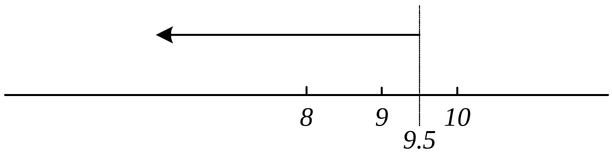
$$a=\mu+z^*\times\sigma=12+1.282\times 4=17.128$$

常態分配是二項分配的極限分配

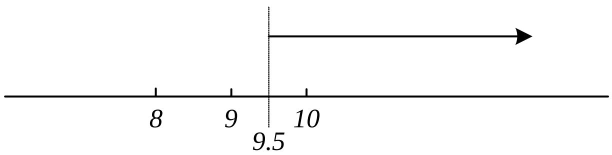
若 n 越大，則二項分配（參數 n、p）與參數 $\mu=np$ 、 $\sigma=\sqrt{np(1-p)}$ 的常態分配越接近。

以常態分配（連續分配）近似二項分配（離散分配）時，特別注意修正項$\pm\frac{1}{2}$ ：

$$P_{\text{離散分配}}(x=a)=P_{\text{連續分配}}\left(a-\tfrac{1}{2}\le x\le a+\tfrac{1}{2}\right)$$
$$P_{\text{離散分配}}(x\le a)=P_{\text{連續分配}}\left(x\le a+\tfrac{1}{2}\right),\quad P_{\text{離散分配}}(x\ge a)=P_{\text{連續分配}}\left(x\ge a-\tfrac{1}{2}\right)$$
$$P_{\text{離散分配}}(x<a)=P_{\text{連續分配}}\left(x\le a-\tfrac{1}{2}\right),\quad P_{\text{離散分配}}(x>a)=P_{\text{連續分配}}\left(x\ge a+\tfrac{1}{2}\right)$$



$$P_{\text{連續分配}}(x\le 9.5)=P_{\text{離散分配}}(x\le 9)=P_{\text{離散分配}}(x<10)$$



$$P_{\text{連續分配}}(x\ge 9.5)=P_{\text{離散分配}}(x\ge 10)=P_{\text{離散分配}}(x>9)$$

範例 3.11　以常態分配近似二項分配

已知 x 爲 $n=100$ 、 $p=0.2$ 的二項分配，y 爲 $\mu=20$ 、 $\sigma=4$ 的常態分配，請計算

(1) $P(19\le X\le 21)$；

(2) $P(19\le Y\le 21)$；

(3) $P(18.5\le Y\le 21.5)$。

【解】

$$E(X)=np=20,\quad Var(X)=np(1-p)=16,\quad \sigma_X=\sqrt{np(1-p)}=4$$

(1)

$$P(x=a)=\frac{100!}{a!\times(100-a)!}\times 0.2^a\times 0.8^{100-a}$$
$$P(19\le x\le 21)=P(x=19)+P(x=20)+P(x=21)=0.0981+0.0993+0.0946=0.2919$$

(2)

$$z_1^*=\frac{19-20}{4}=-0.25,\quad z_2^*=\frac{21-20}{4}=0.25$$
$$P(19\le Y\le 21)=P(-0.25\le z\le 0.25)=0.1974$$

(3)

$$z_1^* = \frac{18.5-20}{4} = -0.375, \quad z_2^* = \frac{21.5-20}{4} = 0.375$$

$$P(18.5 \le Y \le 21.5) = P(-0.375 \le z \le 0.375) = 0.2923$$

3.4　機率表

兩種方法計算隨機變數的機率：

(1)直接由機率函數計算；

(2)查機率表。

四種形式的機率：

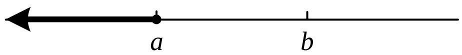
(1)左尾，$P(x \le a) = \alpha$ ，以下累計機率；

(2)右尾，$P(x \ge b) = \alpha$ ，以上累計機率；

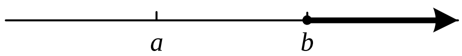
(3)區間，$P(a \le x \le b) = \alpha$ ；

(4)雙尾，$P(x \le a \text{ 或 } x \ge b) = \alpha$ 。

其中，a、b 稱爲**臨界值**，α 爲**機率**。

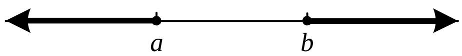


左尾：$P(x \le a) = \alpha$



右尾：$P(x \ge b) = \alpha$

區間：$P(a \le x \le b) = \alpha$



雙尾：$P(x \le a \text{ 或 } x \ge b) = \alpha$

查機率表注意事項：

(1)機率表的內容可能是機率，也可能是臨界值。

(2)內容是機率時，表索引爲臨界值；而內容是臨界值時，表索引爲機率。

(3)必須看清楚機率的形式是**左尾**或**右尾**（離散變數時也可能是**區間**）。

二項分配與卜瓦松機率表

*二項*分配**左尾**機率表

x \ p	0.05	0.1	0.15	0.2	0.25
	n = 10				
0	0.5987	0.3487	0.1969	0.1074	0.0563
1	0.9139	0.7361	0.5443	0.3758	0.2440
2	0.9885	0.9298	0.8202	0.6778	0.5256
3	0.9990	0.9872	0.9500	0.8791	0.7759
4	0.9999	0.9984	0.9901	0.9672	0.9219
5	1.0000	0.9999	0.9986	0.9936	0.9803
6	1.0000	1.0000	0.9999	0.9991	0.9965
7	1.0000	1.0000	1.0000	0.9999	0.9996
8	1.0000	1.0000	1.0000	1.0000	1.0000
9	1.0000	1.0000	1.0000	1.0000	1.0000
10	1.0000	1.0000	1.0000	1.0000	1.0000
11					
12					
13					
14					

*卜瓦松項*分配**左尾**機率表

x	0.15	0.2	2	2.5	3
	λ				
0	0.8607	0.8187	0.1353	0.0821	0.0498
1	0.9898	0.9825	0.4060	0.2873	0.1991
2	0.9995	0.9989	0.6767	0.5438	0.4232
3	1.0000	0.9999	0.8571	0.7576	0.6472
4	1.0000	1.0000	0.9473	0.8912	0.8153
5	1.0000	1.0000	0.9834	0.9580	0.9161
6	1.0000	1.0000	0.9955	0.9858	0.9665
7	1.0000	1.0000	0.9989	0.9958	0.9881
8	1.0000	1.0000	0.9998	0.9989	0.9962
9	1.0000	1.0000	1.0000	0.9997	0.9989
10	1.0000	1.0000	1.0000	0.9999	0.9997
11	1.0000	1.0000	1.0000	1.0000	0.9999
12	1.0000	1.0000	1.0000	1.0000	1.0000
13	1.0000	1.0000	1.0000	1.0000	1.0000
14	1.0000	1.0000	1.0000	1.0000	1.0000

範例 3.12　查表求二項分配的機率

假設 x 爲 $n=10$ 、 $p=0.1$ 的二項分配，請計算

(1) $P(x\le 2)$ ；(2) $P(x\le 1)$ ；(3) $P(x=2)$ ；(4) $P(x>2)$ 。

【解】

(1) $P(x\le 2)=0.9298$

(2) $P(x\le 1)=0.7361$

(3) $P(x=2)=P(x\le 2)-P(x\le 1)=0.9298-0.7361=0.1937$

(4) $P(x>2)=1-P(x\le 2)=1-0.9298=0.0702$

範例 3.13　查表求卜瓦松分配的機率

假設 x 爲 $\lambda=2$ 卜瓦松分配，請計算

(1) $P(x\le 2)$ ；(2) $P(x\le 1)$ ；(3) $P(x=2)$ ；(4) $P(x>2)$ 。

【解】

(1) $P(x\le 2)=0.6767$

(2) $P(x\le 1)=0.4060$

(3) $P(x=2)=P(x\le 2)-P(x\le 1)=0.6767-0.4060=0.2707$

(4) $P(x>2)=1-P(x\le 2)=1-0.6767=0.3233$

標準常態機率表

z 分配機率表

z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.0000	0.0040	0.0080	0.0120	0.0160	0.0199	0.0239	0.0279	0.0319	0.0359
0.1	0.0398	0.0438	0.0478	0.0517	0.0557	0.0596	0.0636	0.0675	0.0714	0.0753
0.2	0.0793	0.0832	0.0871	0.0910	0.0948	0.0987	0.1026	0.1064	0.1103	0.1141
0.3	0.1179	0.1217	0.1255	0.1293	0.1331	0.1368	0.1406	0.1443	0.1480	0.1517
0.4	0.1554	0.1591	0.1628	0.1664	0.1700	0.1736	0.1772	0.1808	0.1844	0.1879
0.5	0.1915	0.1950	0.1985	0.2019	0.2054	0.2088	0.2123	0.2157	0.2190	0.2224
z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.6	0.2257	0.2291	0.2324	0.2357	0.2389	0.2422	0.2454	0.2486	0.2517	0.2549
0.7	0.2580	0.2611	0.2642	0.2673	0.2704	0.2734	0.2764	0.2794	0.2823	0.2852
0.8	0.2881	0.2910	0.2939	0.2967	0.2995	0.3023	0.3051	0.3078	0.3106	0.3133
0.9	0.3159	0.3186	0.3212	0.3238	0.3264	0.3289	0.3315	0.3340	0.3365	0.3389
1.0	0.3413	0.3438	0.3461	0.3485	0.3508	0.3531	0.3554	0.3577	0.3599	0.3621
1.1	0.3643	0.3665	0.3686	0.3708	0.3729	0.3749	0.3770	0.3790	0.3810	0.3830

範例 3.14 查表求標準常態分配的機率

假設 z 爲標準常態分配，請計算

(a) $P(0\le z\le 1)$；(b) $P(z\le 1)$；(c) $P(-1\le z\le 0)$；(d) $P(-1\le z\le 1)$；

(e) $P(z\le -0.82)$；(f) $P(z\le -0.82)$；(g) $P(z\ge 0.26)$。

【解】

z 分配是一個以 $z=0$ 爲中心的對稱分配，上表給的是 $P(0\le z\le z^*)$ 的機率。

令 $\alpha_b = P(0\le z\le b)$，則

$$P(z\le b)=P(z\le 0)+P(0\le z\le b)=0.5+\alpha_b$$

$$P(z\ge -b)=P(z\le b)=0.5+\alpha_b$$

$$P(-b\le z\le 0)=P(0\le z\le b)=\alpha_b$$

$$P(-b\le z\le b)=P(-b\le z\le 0)+P(0\le z\le b)=2\times\alpha_b$$

$$P(z\le -b)=P(z\le 0)-P(-b\le z\le 0)=0.5-\alpha_b$$

$$P(z\ge b)=P(z\le -b)=0.5-\alpha_b$$

$$(a)\ z^*=1,\text{查表得}\alpha_{z^*=1}=0.3413,\quad P(0\le z\le 1)=\alpha_{z^*=1}=0.3413$$

$$(b)\ z^*=1,\quad P(z\le 1)=0.5+\alpha_{z^*=1}=0.8413$$

$$(c)\ z^*=|-1|=1,\quad P(-1\le z\le 0)=\alpha_{z^*=1}=0.3413$$

$$(d)\ z^*=|-1|=1,\quad P(-1\le z\le 1)=2\times\alpha_{z^*=1}=0.6826$$

$$(e)\ z^*=|-0.82|=0.82,\quad P(z\le -0.82)=0.5-\alpha_{z^*=0.82}=0.5-0.2939=0.2061$$

$$(f)\ z^*=0.26,\quad P(z\ge 0.26)=0.5-\alpha_{z^*=0.26}=0.5-0.1026=0.3974$$

範例 3.15　查表求標準常態分配的臨界值

假設 z 爲標準常態分配，請求臨界值 b：
(a) $P(z \le b) = 0.22$；(b) $P(z \le b) = 0.84$；(c) $P(z \ge b) = 0.18$；(d) $P(z \ge b) = 0.64$。

【解】

$(a)\ P(z \le b) = 0.22 \Rightarrow \alpha_{z^*} = 0.5 - 0.22 = 0.28 \Rightarrow z^* = 0.77 \Rightarrow b = -z^* = -0.77$

$(b)\ P(z \le b) = 0.84 \Rightarrow \alpha_{z^*} = 0.84 - 0.5 = 0.34 \Rightarrow z^* = 0.995 \Rightarrow b = z^* = 0.995$

$(c)\ P(z \ge b) = 0.18 \Rightarrow \alpha_{z^*} = 0.5 - 0.18 = 0.32 \Rightarrow z^* = 0.915 \Rightarrow b = z^* = 0.915$

$(d)\ P(z \ge b) = 0.64 \Rightarrow \alpha_{z^*} = 0.64 - 0.5 = 0.14 \Rightarrow z^* = 0.36 \Rightarrow b = -z^* = -0.36$

範例 3.16　查表求常態分配的機率、臨界值

(a)標準常態分配，左尾，臨界值爲 1，求 α；
(b)標準常態分配，左尾，臨界值爲 -0.82，求 α；
(c)標準常態分配，右尾，臨界值爲 0.26，求 α；
(d)標準常態分配，左尾，$\alpha = 0.22$，求臨界值；
(e)標準常態分配，左尾，$\alpha = 0.84$，求臨界值；
(f)標準常態分配，右尾，$\alpha = 0.64$，求臨界值。

【解】

(a)標準常態分配，左尾，臨界值爲 1，求 α $\Rightarrow$ $\alpha = P(z \le 1) = 0.8413$

(b)標準常態分配，左尾，臨界值爲 -0.82，求 α $\Rightarrow$ $\alpha = P(z \le -0.82) = 0.2061$

(c)標準常態分配，右尾，臨界值爲 0.26，求 α $\Rightarrow$ $\alpha = P(z \ge 0.26) = 0.3974$

(d)標準常態分配，左尾，$\alpha = 0.22$，求臨界值 $\Rightarrow$ $P(z \le b) = 0.22 \Rightarrow b = -0.77$

(e)標準常態分配，左尾，$\alpha = 0.84$，求臨界值 $\Rightarrow$ $P(z \le b) = 0.84 \Rightarrow b = 0.995$

(f)標準常態分配，右尾，$\alpha = 0.64$，求臨界值 $\Rightarrow$ $P(z \ge b) = 0.64 \Rightarrow b = -0.36$

3.5　計算隨機變數的期望值與變異數

積分基本操作

(1)多項函數的不定積分：

$$\int x^a dx = \frac{1}{a+1} x^{a+1}, \quad a \neq -1$$

(2)指數函數的不定積分：

$$\int e^{ax} dx = \frac{1}{a} e^{ax}, \quad a \neq 0$$

(3)定積分：

$$\int_{\alpha}^{\beta} x^a dx = \left(\frac{1}{a+1} x^{a+1} \right) \Big|_{\alpha}^{\beta} = \frac{1}{a+1} \alpha^{a+1} - \frac{1}{a+1} \beta^{a+1}, \quad a \neq -1, \alpha < \beta$$

(4)函數和的積分：

$$\int \left[f(x) + g(x) \right] dx = \int f(x) dx + \int g(x) dx$$

(5)常數積的積分：

$$\int \left[af(x) \right] dx = a \int f(x) dx$$

(6)函數積的積分（部分積分）：

$$\int f(x) g'(x) dx = f(x) g(x) - \int g(x) f'(x) dx$$

或

$$\int fdg = fg - \int gdf$$

範例 3.17　基本積分計算

請計算下列積分：

(a) $\int_{x=0}^{x=3} \frac{1}{3} x^2 dx$

(b) $\int_0^{\infty} \lambda e^{-\lambda x} dx$

(c) $\int_{y=0}^{y=3} \left(\frac{2}{3} - ay \right) dy$

(d) $\int_0^{\infty} x \lambda e^{-\lambda x} dx$ 。

【解】

(a)

$$\int_{x=0}^{x=3}\frac{1}{3}x^2dx=\frac{1}{3}\times\frac{1}{3}x^3\Big|_{x=0}^{x=3}=\frac{1}{9}\times 27=3$$

(b)

$$\int_0^{\infty}\lambda e^{-\lambda x}dx=\lambda\tfrac{1}{-\lambda}e^{-\lambda x}\Big|_0^{\infty}=1$$

(c)

$$\int_{y=0}^{y=3}\left(\tfrac{2}{3}-ay\right)dy=\int_{y=0}^{y=3}\tfrac{2}{3}dy-\int_{y=0}^{y=3}aydy=\left(\tfrac{2}{3}y-a\times\tfrac{1}{2}y^2\right)\Big|_{y=0}^{y=3}=2-\tfrac{9}{2}a$$

(d)

令 $f=x, dg=\lambda e^{-\lambda x}dx \quad\Rightarrow\quad df=dx, g=-e^{-\lambda x}$

$$\int_0^{\infty}x\lambda e^{-\lambda x}dx=\int_0^{\infty}fdg=fg\Big|_0^{\infty}-\int_0^{\infty}gdf=-xe^{-\lambda x}\Big|_0^{\infty}+\int_0^{\infty}e^{-\lambda x}dx=-\frac{1}{\lambda}e^{-\lambda x}\Big|_0^{\infty}=\frac{1}{\lambda}$$

範例 3.18　以積分計算隨機變數之期望值、變異數

隨機變數 X 的機率密度函數如下：

$$f(x)=\frac{1}{3},\ \ 0\le x\le 3$$

計算 $E(X)$ 、 $Var(X)$ 。

【解】

計算期望值

$$E(X)=\int_{x=0}^{x=3}xf(x)dx=\int_{x=0}^{x=3}x\frac{1}{3}dx=\frac{x^2}{6}\Big|_{x=0}^{x=3}=\frac{9}{6}=\frac{3}{2}$$

計算變異數

$$Var(X)=\int_{x=0}^{x=3}\left(x-E(x)\right)^2 f(x)dx=\int_{x=0}^{x=3}\left(x-\frac{3}{2}\right)^2\frac{1}{3}dx=\frac{3}{4}$$

或

$$Var(X)=\int_{x=0}^{x=3}x^2 f(x)dx-\left(E(x)\right)^2=\int_{x=0}^{x=3}\frac{1}{3}x^2dx-\left(\frac{3}{2}\right)^2=3-\frac{9}{4}=\frac{3}{4}$$

範例 3.19　以積分計算隨機變數之期望值、變異數

隨機變數 Y 的機率密度函數如下：

$$f(y)=\tfrac{2}{3}-ay,\ \ 0\le y\le 3$$

計算 $E(Y)$ 、 $Var(Y)$ 。

【解】

求未知參數 a：

$$\int_{y=0}^{y=3} f(y)dy=1 \ \Rightarrow\ \int_{y=0}^{y=3}\left(\tfrac{2}{3}-ay\right)dy=\left(\tfrac{2}{3}y-a\frac{y^2}{2}\right)\Bigg|_{y=0}^{y=3}=2-\tfrac{9}{2}a=1 \ \ \Rightarrow\ \ a=\frac{2}{9}$$

計算期望值

$$E(Y)=\int_{y=0}^{y=3} yf(y)dy=\int_{y=0}^{y=3} y\left(\tfrac{2}{3}-\tfrac{2}{9}y\right)dy=\left(\tfrac{1}{3}y^2-\tfrac{2}{27}y^3\right)\Big|_{y=0}^{y=3}=1$$

計算變異數

$$Var(Y)=\int_{y=0}^{y=3}\left(y-E(y)\right)^2 f(y)dy=\int_{y=0}^{y=3}\left(y-1\right)^2\left(\tfrac{2}{3}-\tfrac{2}{9}y\right)dy=-\frac{9}{2}+10-7+2=\frac{1}{2}$$

或

$$Var(Y)=\int_{y=0}^{y=3} y^2 f(y)dy-\left(E(y)\right)^2=\int_{y=0}^{y=3} y^2\left(\tfrac{2}{3}-\tfrac{2}{9}y\right)dy-1^2=\left(6-\frac{9}{2}\right)-1=\frac{1}{2}$$

範例 3.20　計算條件機率、邊際機率

隨機變數 X 、 Y 的聯合機率密度函數如下：

$$\begin{aligned} f(x,y)&=x+ay, && 0<x<1, 0<y<1\\ &=0 && \text{otherwise}\end{aligned}$$

計算 $f(x)$ 、 $f(y)$ 、 $V(X)$ 、 $\operatorname{cov}(X,Y)$ 。

【解】

求未知參數 a：

$$\int_{y=0}^{1}\int_{x=0}^{1}\left(x+ay\right)dxdy=\int_{y=0}^{1}\left(\tfrac{1}{2}+ay\right)dy=\tfrac{1}{2}+\tfrac{a}{2}=1 \ \ \Rightarrow\ \ a=1$$

$$f(x)=\int_{y=0}^{1}\left(x+y\right)dy=x+\tfrac{1}{2}$$

$f(y)=\int_{x=0}^{1}(x+y)dx=\frac{1}{2}+y$

因 $f(x,y)=x+y\neq f(x)f(y)=\left(x+\frac{1}{2}\right)\left(\frac{1}{2}+y\right)$，故 X 、Y 兩者不獨立。

$$E(x)=E(y)=\int_0^1 y\left(\tfrac{1}{2}+y\right)dy=\frac{1}{2}\times\frac{1}{2}+\frac{1}{3}=\frac{7}{12}$$

$$E(x^2)=E(y^2)=\int_0^1 y^2\left(\tfrac{1}{2}+y\right)dy=\frac{1}{2}\times\frac{1}{4}+\frac{1}{4}=\frac{5}{12}$$

$$E(xy)=\int_0^1\int_0^1 xy(x+y)dxdy=\int_0^1\left(\frac{y}{3}+\frac{y^2}{2}\right)dy=\frac{1}{3}\times\frac{1}{2}+\frac{1}{2}\times\frac{1}{3}=\frac{1}{3}$$

$$V(X)=E(x^2)-\left[E(x)\right]^2=\frac{5}{12}-\frac{7}{12}\times\frac{7}{12}=\frac{11}{144}$$

$$\operatorname{cov}(X,Y)=E(xy)-E(x)E(y)=\frac{1}{3}-\frac{7}{12}\times\frac{7}{12}=-\frac{1}{144}$$

3.6 動差與動差母函數

動差（*moments*）或**零動差**（*zero moments*）

$$n \text{ 級零動差} = \mu_n' = M_n' = E\left(x^n\right)$$

中央動差（*central moments*）或**主動差**（*principle moments*）

$$n \text{ 級主動差} = \mu_n = M_n = E\left[\left(x-\mu\right)^n\right]$$

動差、中央動差、平均數、與變異數之關係

中央動差直接定義 σ^2 、 α_3 、 α_4 等有意義的係數。

零動差比較容易計算。

$$\begin{aligned}
M_2 &= E\left(x-\mu\right)^2 = E\left(x^2\right) - 2\mu E\left(x\right) + \mu^2 \\
M_3 &= E\left(x-\mu\right)^3 = E\left(x^3\right) - 3\mu E\left(x^2\right) + 3\mu^2 E\left(x\right) - \mu^3 \\
&\vdots \\
M_n &= E\left(x-\mu\right)^n = \sum_{i=0}^{n} C_i^n \left(-\mu\right)^i E\left(x^{n-i}\right)
\end{aligned}$$

$$\begin{aligned}
\mu &= E\left(x\right) \\
\sigma^2 &= M_2 = E\left(x^2\right) - \mu^2 \\
\alpha_3 &= \frac{M_3}{\sigma^3}
\end{aligned}$$

動差母函數（*moment generating function*，**動差生成函數**）

由動差母函數可以很容易計算各級動差（零動差）。

$$m\left(t\right) = E\left(e^{tx}\right)$$

範例 3.21　由動差母函數產生各級動差

e^{tx} 的泰勒展開式為

$$e^{tx} = 1 + tx + \frac{\left(tx\right)^2}{2!} + \frac{\left(tx\right)^3}{3!} + \frac{\left(tx\right)^4}{4!} + \cdots$$

即

$$m(t)=E\left(e^{tx}\right)=1+tE(x)+\frac{t^2}{2!}E\left(x^2\right)+\frac{t^3}{3!}E\left(x^3\right)+\frac{t^4}{4!}E\left(x^4\right)+\cdots$$

也就是

$$\begin{aligned}
\left.\frac{d}{dt}m(t)\right]_{t=0} &= \left.E(x)+tE\left(x^2\right)+\frac{t^2}{2!}E\left(x^3\right)+\cdots\right]_{t=0}=E(x)\\
\left.\frac{d^2}{dt^2}m(t)\right]_{t=0} &= \left.E\left(x^2\right)+tE\left(x^3\right)+\frac{t^2}{2!}E\left(x^4\right)+\cdots\right]_{t=0}=E\left(x^2\right)\\
&\vdots\\
\left.\frac{d^k}{dt^k}m(t)\right]_{t=0} &= \left.E\left(x^k\right)+tE\left(x^{k+1}\right)+\frac{t^2}{2!}E\left(x^{k+2}\right)+\cdots\right]_{t=0}=E\left(x^k\right)
\end{aligned}$$

範例 3.22 由動差母函數求期望值、變異數

若隨機變數 x 的動差母函數爲

$$m(t)=e^{\lambda\left(e^t-1\right)}$$

請計算 $E(x)$ 、 $Var(x)$ 。

【解】

由 $m(t)$ 求一、二級動差，過程如下

$$\begin{aligned}
E(x)&=m^{(1)}(0)=\left.e^{\lambda\left(e^t-1\right)}\lambda e^t\right]_{t=0}=\lambda\\
E\left(x^2\right)&=m^{(2)}(0)=\left.\left\{e^{\lambda\left(e^t-1\right)}\left(\lambda e^t\right)^2+e^{\lambda\left(e^t-1\right)}\lambda e^t\right\}\right]_{t=0}=\lambda^2+\lambda
\end{aligned}$$

則

$$Var(x)=E\left(x^2\right)-\left[E(x)\right]^2=\lambda^2+\lambda-\lambda^2=\lambda$$

範例 3.23 白努力分配之機率函數驗證、動差母函數、期望值、變異數

白努力分配的機率函數爲

$$P(x)=f(x;p)=p^x(1-p)^{1-x},\qquad x\in\{0,1\},\ 0<p<1$$

驗證機率函數

$$\sum_{x=0}^{1} p^x (1-p)^{1-x} \;=\; (1-p)+p \;=\; 1$$

其 $m(t)$ 的求解過程如下

$$m(t)=E\left(e^{tx}\right)=\sum_{x=0}^{1} e^{tx} p^x (1-p)^{1-x}=(1-p)+e^t p=pe^t+(1-p)$$

期望值

$$E(x)=m^{(1)}(0)=pe^t\Big]_{t=0}=p$$

變異數

$$E\left(x^2\right)=m^{(2)}(0)=pe^t\Big]_{t=0}=p$$

$$Var(x)=E\left(x^2\right)-\left[E(x)\right]^2=p-p^2=p(1-p)$$

範例 3.24 二項分配之機率函數驗證、動差母函數、期望值、變異數

二項分配的機率函數爲

$$P(x)=f(x;n,p)=C_x^n p^x (1-p)^{n-x}, \qquad x\in\{0,1,\ldots,n\},\ 0<p<1$$

驗證機率函數

$$\sum_{x=0}^{n} C_x^n p^x (1-p)^{n-x}=\left[p+(1-p)\right]^n=1 \qquad \text{（二項式定理）}$$

其 $m(t)$ 的求解過程如下

$$m(t)=E\left(e^{tx}\right)=\sum_{x=0}^{n} e^{tx} C_x^n p^x (1-p)^{n-x}=\sum_{x=0}^{n} C_x^n \left(pe^t\right)^x (1-p)^{n-x}=\left[pe^t+(1-p)\right]^n$$

期望值

$$E(x)=m^{(1)}(0)=n\left[pe^t+(1-p)\right]^{n-1} pe^t\Big]_{t=0}=np$$

變異數

$$E\left(x^2\right)=m^{(2)}(0)$$
$$=n(n-1)\left[pe^t+(1-p)\right]^{n-2}\left(pe^t\right)^2+n\left[pe^t+(1-p)\right]^{n-1}pe^t\Big]_{t=0}=n(n-1)p^2+np$$
$$Var(x)=E\left(x^2\right)-\left[E(x)\right]^2=n(n-1)p^2+np-(np)^2=np(1-p)$$

範例 3.25　動差母函數之性質

令 $X_1,X_2,\ldots,X_n$ 爲一致且互相獨立的分配，其動差母函數皆爲 $m(t)$，若

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

則 Y 之動差母函數爲

$$E\left(e^{ty}\right)=E\left(e^{t(x_1+x_2+\cdots+x_n)}\right)=E\left(e^{tx_1}e^{tx_2}\cdots e^{tx_n}\right)=E\left(e^{tx_1}\right)E\left(e^{tx_2}\right)\cdots E\left(e^{tx_n}\right)=\left[m(t)\right]^n$$

例如，白努力分配的動差爲

$$m(t)=pe^t+(1-p)$$

則二項分配的動差爲

$$\left[m(t)\right]^n=\left[pe^t+(1-p)\right]^n$$

範例 3.26　幾何分配之機率函數驗證、動差母函數、期望值、變異數

幾何分配的機率函數爲

$$P(x)=f(x;p)=p(1-p)^{x-1},\qquad x\in\{1,2,\ldots\},\ 0<p<1$$

驗證機率函數

$$\sum_{x=1}^{\infty}p(1-p)^{x-1}=p\sum_{x=1}^{\infty}(1-p)^{x-1}=p\times\frac{1}{1-(1-p)}=1\qquad\left(\sum_{n=0}^{\infty}x^n=\frac{1}{1-x}\right)$$

其 $m(t)$ 的求解過程如下

$$m(t)=E\left(e^{tx}\right)=\sum_{x=0}^{1}e^{tx}p(1-p)^{x-1}=pe^t\sum_{x=0}^{1}\left[e^t(1-p)\right]^{x-1}=\frac{pe^t}{1-e^t(1-p)}=\frac{pe^t}{1-(1-p)e^t}$$

期望值

$$E(x)=m^{(1)}(0)=\left.\frac{pe^t}{1-(1-p)e^t}+\frac{pe^t\times(1-p)e^t}{\left[1-(1-p)e^t\right]^2}\right]_{t=0}=\left.\frac{pe^t}{\left[1-(1-p)e^t\right]^2}\right]_{t=0}=\frac{1}{p}$$

變異數

$$E(x^2)=m^{(2)}(0)$$

$$=\left.\frac{pe^t}{\left[1-(1-p)e^t\right]^2}+\frac{pe^t\times2(1-p)e^t}{\left[1-(1-p)e^t\right]^3}\right]_{t=0}=\left.\frac{pe^t\times\left[1+(1-p)e^t\right]}{\left[1-(1-p)e^t\right]^3}\right]_{t=0}=\frac{2-p}{p^2}$$

$$Var(x)=E(x^2)-\left[E(x)\right]^2=\frac{2-p}{p^2}-\left(\frac{1}{p}\right)^2=\frac{1-p}{p^2}=\frac{1}{p}\times\frac{1-p}{p}$$

範例 3.27 負二項分配之機率函數驗證、動差母函數、期望值、變異數

負二項分配的機率函數爲

$$P(x)=f(x;r,p)=C_{r-1}^{x-1}p^r(1-p)^{x-r},\qquad x\in\{r,r+1,\ldots\},\ 0<p<1$$

$$E(x)=\frac{r}{p}$$

$$V(x)=\frac{r(1-p)}{p^2}$$

驗證機率函數

$$\sum_{x=r}^{\infty}C_{r-1}^{x-1}p^r(1-p)^{x-r}=p^n\sum_{x=r}^{\infty}\frac{\overbrace{(x-1)(x-2)\cdots}^{r-1\text{ 項}}}{(r-1)!}(1-p)^{x-r}=p^r\times\frac{1}{(r-1)!}\frac{(r-1)!}{\left[1-(1-p)\right]^r}=1$$

其 $m(t)$ 可由幾何分配的動機母函數寫出

$$m(t)=E(e^{tx})=\left[\frac{pe^t}{1-(1-p)e^t}\right]^r$$

期望值

$$E(x)=m^{(1)}(0)=r\left.\left[\frac{pe^t}{1-(1-p)e^t}\right]^{r-1}\frac{pe^t}{\left[1-(1-p)e^t\right]^2}\right]_{t=0}=\frac{r}{p}$$

變異數

$$E\left(x^2\right)=m^{(2)}(0)=r(r-1)\left[\frac{pe^t}{1-(1-p)e^t}\right]^{r-2}\left(\frac{pe^t}{\left[1-(1-p)e^t\right]^2}\right)^2\Bigg]_{t=0}$$

$$+r\left[\frac{pe^t}{1-(1-p)e^t}\right]^{r-1}\frac{pe^t\times\left[1+(1-p)e^t\right]}{\left[1-(1-p)e^t\right]^3}\Bigg]_{t=0}$$

$$=\frac{r(r-1)}{p^2}+\frac{r(2-p)}{p^2}=\frac{r(r+1-p)}{p^2}$$

$$Var(x)=E\left(x^2\right)-\left[E(x)\right]^2=\frac{r(r+1-p)}{p^2}-\left(\frac{r}{p}\right)^2=\frac{r(1-p)}{p^2}=r\times\frac{1}{p}\times\frac{1-p}{p}$$

範例 3.28　卜瓦松分配之機率函數驗證、動差母函數、期望值、變異數

卜瓦松分配的機率函數為

$$P(x)=f(x;\lambda)=\frac{\lambda^x e^{-\lambda}}{x!},\qquad x\in\{0,1,\ldots\},\ \lambda>0$$

驗證機率函數

$$\sum_{x=0}^{\infty}\frac{\lambda^x}{x!}e^{-\lambda}\ =\ e^{-\lambda}\sum_{x=0}^{\infty}\frac{\lambda^x}{x!}\ =\ e^{-\lambda}\times e^{\lambda}\ =\ 1\qquad\qquad \left(e^{\lambda}=\sum_{x=0}^{\infty}\frac{\lambda^x}{x!}\right)$$

其 $m(t)$ 的求解過程如下

$$m(t)=E\left(e^{tx}\right)=\sum_{x=0}^{\infty}e^{tx}\frac{\lambda^x e^{-\lambda}}{x!}=e^{-\lambda}\sum_{x=0}^{\infty}\frac{\left(\lambda e^t\right)^x}{x!}=e^{-\lambda}e^{\lambda e^t}=e^{\lambda\left(e^t-1\right)}$$

期望值

$$E(x)=m^{(1)}(0)=e^{\lambda\left(e^t-1\right)}\lambda e^t\Big]_{t=0}=\lambda$$

變異數

$$E\left(x^2\right)=m^{(2)}(0)=\left\{e^{\lambda\left(e^t-1\right)}\left(\lambda e^t\right)^2+e^{\lambda\left(e^t-1\right)}\lambda e^t\right\}\Big]_{t=0}=\lambda^2+\lambda$$

$$Var(x)=E\left(x^2\right)-\left[E(x)\right]^2=\lambda^2+\lambda-\lambda^2=\lambda$$

範例 3.29　指數分配之機率函數驗證、動差母函數、期望值、變異數

指數分配的機率密度函數爲

$$f(x)=f(x;\lambda)=\lambda e^{-\lambda x},\qquad x>0,\lambda>0$$

驗證機率密度函數

$$\int_0^\infty \lambda e^{-\lambda x}dx=\lambda\frac{1}{-\lambda}e^{-\lambda x}\Big|_0^\infty=1$$

其 $m(t)$ 的求解過程如下

$$m(t)=E\left(e^{tx}\right)=\int_{x=0}^\infty e^{tx}\lambda e^{-\lambda x}dx=\lambda\int_0^\infty e^{(t-\lambda)x}dx=\lambda\frac{1}{t-\lambda}e^{(t-\lambda)x}\Big]_0^\infty=\frac{\lambda}{\lambda-t}$$

期望值

$$E(x)=m^{(1)}(0)=\frac{\lambda}{(\lambda-t)^2}\Bigg]_{t=0}=\frac{1}{\lambda}$$

變異數

$$E\left(x^2\right)=m^{(2)}(0)=\frac{2\lambda}{(\lambda-t)^3}\Bigg]_{t=0}=\frac{2}{\lambda^2}$$

$$Var(x)=E\left(x^2\right)-\left[E(x)\right]^2=\frac{2}{\lambda^2}-\left(\frac{1}{\lambda}\right)^2=\frac{1}{\lambda^2}$$

附錄：各分配期望值與變異數之計算過程（直接積分）

白努力分配

機率函數： $f(x;p)=p^x(1-p)^{1-x},\ x\in\{0,1\},\ 0<p<1$

$$\sum_{x=0}^{1} p^x(1-p)^{1-x} = (1-p)+p = 1$$

$$E(X) = \sum_{x=0}^{1} xp^x(1-p)^{1-x} = 0+p = p$$

$$V(X) = \sum_{x=0}^{1} x^2p^x(1-p)^{1-x}-\big(E(X)\big)^2 = p-p^2 = p(1-p)$$

二項分配

機率函數： $f(x;p,n)=C_x^n p^x(1-p)^{n-x},\ x\in\{0,1,\ldots,n\},\ 0<p<1$

$$\sum_{x=0}^{n} C_x^n p^x(1-p)^{n-x}=\big[p+(1-p)\big]^n=1$$

$$E(X)=\sum_{x=0}^{n} xC_x^n p^x(1-p)^{n-x}=np\big(p+(1-p)\big)^{n-1}=np$$

$$\begin{aligned}V(X)&=E\left(X^2\right)-\left(E(X)\right)^2=E\big(X(X-1)\big)+E(X)-\left(E(X)\right)^2\\&=\sum_{x=1}^{n}x(x-1)\frac{n!}{x!(n-x)!}p^x(1-p)^{n-x}+np-(np)^2\\&=n(n-1)p^2+np-(np)^2=np(1-p)\end{aligned}$$

$$\boxed{\begin{aligned}&(p+q)^n=\sum_{x=0}^{n}C_x^n p^x q^{n-x}\\&\sum_{x=0}^{n}xC_x^n p^x q^{n-x}=\sum_{x=0}^{n}x\frac{n}{x}pC_{x-1}^{n-1}p^{x-1}q^{n-x}=np(p+q)^{n-1}\\&\sum_{x=0}^{n}x(x-1)C_x^n p^x q^{n-x}=n(n-1)p^2(p+q)^{n-2}\end{aligned}}$$

超幾何分配

機率函數： $f(x;N,S,n)=\dfrac{C_x^S C_{n-x}^{N-S}}{C_n^N},\ x\in\{0,1,\ldots,n\},\ 0<S<N$

$$\sum_{x=0}^{n}\frac{C_x^S C_{n-x}^{N-S}}{C_n^N} = \frac{1}{C_n^N}\sum_{x=0}^{n}C_x^S C_{n-x}^{N-S} = \frac{1}{C_n^N}\times C_{x+(n-x)}^{S+(N-S)} = 1$$

$$E(X)=\sum_{x=0}^{n}x\frac{C_x^S C_{n-x}^{N-S}}{C_n^N}=\frac{1}{C_n^N}\sum_{x=0}^{n}xC_x^S C_{n-x}^{N-S}=\frac{1}{\frac{N}{n}C_{n-1}^{N-1}}\times SC_{n-1}^{N-1}=n\frac{S}{N}$$

$$V(X)=E\left(X^2\right)-\left(E(X)\right)^2 = E\left(X(X-1)\right)+E(X)-\left(E(X)\right)^2$$

$$=\frac{1}{C_n^N}\sum_{x=0}^{n}x(x-1)C_x^S C_{n-x}^{N-S}+n\frac{S}{N}-\left(n\frac{S}{N}\right)^2$$

$$= \frac{n(n-1)}{N(N-1)}\times S(S-1)+n\frac{S}{N}-\left(n\frac{S}{N}\right)^2$$

$$=n\frac{S}{N}\left(\frac{(n-1)S}{(N-1)}+1+\frac{nS}{N}\right) = n\frac{S}{N}\times\frac{(n-1)SN+(N-1)N+nS(N-1)}{N(N-1)}$$

$$=n\frac{S}{N}\times\frac{(N-S)(N-n)}{N(N-1)} = n\frac{S}{N}\left(1-\frac{S}{N}\right)\times\frac{(N-n)}{(N-1)}$$

$$\boxed{\begin{aligned}
&\sum_{x=0}^{n}C_x^M C_{n-x}^N=C_n^{M+N}\\
&C_n^N=\frac{N!}{n!(N-n)!}=\frac{N}{n}\times\frac{(N-1)!}{(n-1)!(N-n)!}=\frac{N}{n}C_{n-1}^{N-1}\\
&\sum_{x=0}^{n}xC_x^M C_{n-x}^N=M\sum_{x=1}^{n}C_{x-1}^{M-1}C_{n-x}^N=MC_{n-1}^{M+N-1}\\
&\sum_{x=0}^{n}x(x-1)C_x^M C_{n-x}^N=M(M-1)C_{n-2}^{M+N-2}
\end{aligned}}$$

卜瓦松分配

機率函數： $f(x;\lambda)=\dfrac{\lambda^x}{x!}e^{-\lambda},\ x\in\{0,1,\ldots\},\ \lambda>0$

$$\sum_{x=0}^{\infty}\frac{\lambda^x}{x!}e^{-\lambda} = e^{-\lambda}\sum_{x=0}^{\infty}\frac{\lambda^x}{x!} = e^{-\lambda}\times e^{\lambda} = 1$$

$$E(X)=\sum_{x=0}^{n}x\frac{\lambda^x}{x!}e^{-\lambda}=e^{-\lambda}\times\lambda\sum_{x=1}^{\infty}\frac{\lambda^{x-1}}{(x-1)!}=\lambda$$

$$V(X)=E(X^2)-(E(X))^2 \;=\; E(X(X-1))+E(X)-(E(X))^2$$
$$=\sum_{x=0}^{n} x(x-1)\frac{\lambda^x}{x!}e^{-\lambda}+\lambda-\lambda^2$$
$$=\lambda^2\sum_{x=2}^{n}\frac{\lambda^{x-2}}{(x-2)!}e^{-\lambda}+\lambda-\lambda^2$$
$$=\lambda$$

$$\boxed{e^x=1+x+\frac{1}{2}x^2+\cdots+\frac{1}{n!}x^n+\cdots=\sum_{n=0}^{\infty}\frac{1}{n!}x^n}$$

幾何分配

機率函數： $f(x;p)=p(1-p)^{x-1},\ x\in\{1,2,\ldots\},\ 0<p<1$

$$\sum_{x=1}^{\infty}p(1-p)^{x-1} \;=\; p\sum_{x=1}^{\infty}(1-p)^{x-1} \;=\; p\times\frac{1}{1-(1-p)} \;=\; 1$$

$$E(X)=\sum_{x=1}^{\infty}xp(1-p)^{x-1}=p\sum_{x=1}^{\infty}x(1-p)^{x-1}=p\times\frac{1}{\left[1-(1-p)\right]^2}=\frac{1}{p}$$

$$V(X)=E(X(X-1))+E(X)-(E(X))^2$$
$$=\sum_{x=1}^{\infty}x(x-1)p(1-p)^{x-1}+\frac{1}{p}-\frac{1}{p^2}=p(1-p)\sum_{x=2}^{\infty}x(x-1)(1-p)^{x-2}+\frac{1}{p}-\frac{1}{p^2}$$
$$=p(1-p)\times\frac{2}{\left[1-(1-p)\right]^3}+\frac{1}{p}-\frac{1}{p^2} \;=\; \frac{2(1-p)}{p^2}+\frac{p}{p^2}-\frac{1}{p^2}$$
$$=\frac{1-p}{p^2}$$

$$\boxed{\begin{aligned}&\sum_{x=0}^{\infty}q^x=\sum_{x=1}^{\infty}q^{x-1}=1+q+q^2+\cdots=\frac{1}{1-q}\\&\frac{d}{dq}\left(\sum_{x=0}^{\infty}q^x\right)=\sum_{x=1}^{\infty}xq^{x-1}=\frac{1}{(1-q)^2}\\&\frac{d}{dq^2}\left(\sum_{x=0}^{\infty}q^x\right)=\sum_{x=2}^{\infty}x(x-1)q^{x-2}=\frac{2}{(1-q)^3}\end{aligned}}$$

負二項分配

機率函數： $f(x;n,p)=C_{n-1}^{x-1}p^n(1-p)^{x-n}$, $x\in\{n,n+1,\ldots\}$, $0<p<1$

$$\sum_{x=n}^{\infty}C_{n-1}^{x-1}p^n(1-p)^{x-n}=p^n\sum_{x=n}^{\infty}\frac{\overbrace{(x-1)(x-2)\cdots}^{n-1\text{ 項}}}{(n-1)!}(1-p)^{x-n}=p^n\times\frac{1}{(n-1)!}\frac{(n-1)!}{\left[1-(1-p)\right]^n}=1$$

$$E(X)=\sum_{x=n}^{\infty}xC_{n-1}^{x-1}p^n(1-p)^{x-n}=\frac{p^n}{(n-1)!}\sum_{x=n}^{\infty}x\times\overbrace{(x-1)(x-2)\cdots}^{n-1\text{ 項}}\times(1-p)^{x-n}$$

$$=\frac{p^n}{(n-1)!}\times\frac{n!}{\left[1-(1-p)\right]^{n+1}}=\frac{n}{p}$$

$$V(X)=E\left(X(X-n)\right)+nE(X)-\left(E(X)\right)^2$$

$$=\sum_{x=n}^{\infty}x(x-n)C_{n-1}^{x-1}p^n(1-p)^{x-n}+\frac{n^2}{p}-\frac{n^2}{p^2}$$

$$=\frac{p^n(1-p)}{(n-1)!}\sum_{x=n}^{\infty}x(x-n)\times\overbrace{(x-1)(x-2)\cdots}^{n-1\text{ 項}}\times(1-p)^{x-n-1}+\frac{n^2}{p}-\frac{n^2}{p^2}$$

$$=\frac{p^n(1-p)}{(n-1)!}\times\frac{(n+1)!}{\left[1-(1-p)\right]^{n+2}}+\frac{n^2}{p}-\frac{n^2}{p^2}$$

$$=\frac{n(n+1)(1-p)}{p^2}+\frac{n^2p}{p^2}-\frac{n^2}{p^2}=\frac{n(1-p)}{p^2}$$

$$\boxed{\begin{aligned}&\sum_{x=0}^{\infty}q^x=\sum_{x=1}^{\infty}q^{x-1}=1+q+q^2+\cdots=\frac{1}{1-q}\\&\frac{d}{dq^{n-1}}\left(\sum_{x=1}^{\infty}q^{x-1}\right)=\sum_{x=n}^{\infty}\overbrace{(x-1)(x-2)\cdots}^{n-1\text{項}}q^{x-n}=\frac{(n-1)!}{(1-q)^n}\\&\frac{d}{dq^n}\left(\sum_{x=0}^{\infty}q^x\right)=\sum_{x=n}^{\infty}\overbrace{x(x-1)\cdots}^{n\text{ 項}}q^{x-n}=\frac{n!}{(1-q)^{n+1}}\end{aligned}}$$

均等分配

機率密度函數： $f(x;a,b)=\dfrac{1}{b-a},\ a\le x\le b$

$$\int_a^b \frac{1}{b-a}dx=\left.\frac{x}{b-a}\right|_a^b=\frac{b-a}{b-a}=1$$

$$E(x)=\int_a^b \frac{x}{b-a}dx=\left.\frac{1}{2}\frac{x^2}{b-a}\right|_a^b=\frac{1}{2}\frac{b^2-a^2}{b-a}=\frac{1}{2}(b+a)$$

$$V(x)=E\left(x^2\right)-\left(E(x)\right)^2=\int_a^b \frac{x^2}{b-a}dx-\left[\frac{(b+a)}{2}\right]^2$$

$$=\frac{1}{3}\frac{b^3-a^3}{b-a}-\left[\frac{(b+a)}{2}\right]^2=\frac{(b-a)^2}{12}$$

指數分配

機率密度函數： $f(x;\lambda)=\lambda e^{-\lambda x},\ \ x>0,\lambda>0$

$$\int_0^\infty \lambda e^{-\lambda x}dx=\left.\lambda\tfrac{1}{-\lambda}e^{-\lambda x}\right|_0^\infty=1$$

$$E(x)=\int_0^\infty x\lambda e^{-\lambda x}dx=\tfrac{1}{\lambda}$$

$$V(x)=\int_0^\infty x^2\lambda e^{-\lambda x}dx-\left(\tfrac{1}{\lambda}\right)^2=\left(\tfrac{1}{\lambda}\right)^2$$

$$P(x>a)=1-F(a)=\int_a^\infty \lambda e^{-\lambda x}dx=\left.\lambda\tfrac{1}{-\lambda}e^{-\lambda x}\right|_a^\infty=e^{-\lambda a}$$

第四章　抽樣與抽樣分配

2006 年 8 月 9 日 最後修改

4.1　抽樣與抽樣方法

4.2　抽樣分配概論

4.3　常見的抽樣分配

4.4　中央極限定理

4.1　抽樣與抽樣方法

母體（*population*）：我們有興趣的研究對象，一般是由許多個體或所組成的集合。

樣本（*sample*）：母體的部分集合。

我們有興趣的是母體，但是實際測量、研究的是樣本。

我們希望經由樣本提供的資訊來推測母體的狀況（推論統計）。

需要樣本的理由：

普查的成本高（時間、金錢）

無法作普查（無法掌握每一個體）

有時測量會破壞樣本

抽樣（*sampling*）：由母體中挑選出樣本的動作。

抽樣的考量：

代表性（正確反應母體的狀況）

抽樣成本（金錢、時間、方便性）

研究母體（*study population*，**抽樣母體**）：以抽樣的目的而對母體的另一種定義。

真正可以被選取之個體的集合爲研究母體。

【例如】以電話訪問選民的投票傾向，**母體**是所有合格選民，**研究母體**（**抽樣母體**）則是有登錄電話號碼的選民。

抽樣方法

抽樣方法分成兩大類：

(1)隨機抽樣方法

(2)非隨機抽樣方法

隨機抽樣（*random sampling*）：知道母體中特定個體被抽中之機率。

隨機抽樣可以計算抽出某特定樣本的機率。

非隨機抽樣（*non-random sampling*）：不涉及機率的抽樣方法。

隨機抽樣方法：

(1)**簡單隨機抽樣**（*Simple Random Sampling*）

每個元素被抽中的機率皆相等。

(2)**分層隨機抽樣**（*Stratified Random Sampling*）

有自然分組，各組內分別作簡單隨機抽樣。

適用組間差異大的情況。

(3)**叢集抽樣**（*Cluster Sampling*）

有自然分組，以組爲單位作簡單隨機抽樣（抽中哪些組）。

適用組間差異小的情況。

(4)**系統抽樣法**（*Systematic Sampling*）

有系統分組後，（以簡單隨機抽樣法）隨機抽取一組。

隨機抽樣方法：

(1)立意抽樣法

(2)滾雪球抽樣法

(3)偶遇抽樣法

(4)定額抽樣法

抽樣誤差

母體參數（*population parameters*）：母體的特徵，如 μ 、 σ 。

樣本統計量（*sample statistic*）：樣本的特徵，如 $\overline{x}$ 、 s 。

樣本統計量是一個樣本的函數，應寫成

$$\overline{x} = \overline{x}(x_1, x_2, \dots, x_n) = \frac{x_1 + x_2 + \cdots + x_n}{n}$$

$$s = s(x_1, x_2, \dots, x_n) = \sqrt{\frac{(x_1 - \overline{x})^2 + (x_2 - \overline{x})^2 + \cdots + (x_n - \overline{x})^2}{n-1}}$$

樣本統計量隨樣本的不同而改變，母體參數則爲定值。

樣本統計量也是隨機變數

既然 $x_1, x_2, \dots, x_n$ 都是隨機變數，$\overline{x}(x_1, x_2, \dots, x_n)$ 、$s(x_1, x_2, \dots, x_n)$ 當然也是隨機變數。

抽樣誤差（*sampling error*）：樣本統計量與母體參數間的差距。

抽樣誤差，$\varepsilon = \overline{x} - \mu = \overline{x}(x_1, x_2, \dots, x_n) - \mu$ ，也是隨機變數。

4.2 抽樣分配概論

抽樣分配（*sampling distributions*）：樣本統計量的分配。

最常見的樣本統計量爲：

$$\overline{x} = \overline{x}(x_1, x_2, \dots, x_n) = \frac{x_1 + x_2 + \cdots + x_n}{n}$$

$$s^2 = s^2(x_1, x_2, \dots, x_n) = \frac{(x_1 - \overline{x})^2 + (x_2 - \overline{x})^2 + \cdots + (x_n - \overline{x})^2}{n-1}$$

基本樣本統計量：

$$T = T(x_1, x_2, \dots, x_n) = x_1 + x_2 + \cdots + x_n$$

$$U = U(x_1, x_2, \dots, x_n) = x_1^2 + x_2^2 + \cdots + x_n^2$$

一致且獨立分配（*Identical and Independent Distribution, i.i.d.*）：

若 $x_1, x_2, \dots, x_n$ 有相同的分配，而且互相獨立，則稱 $x_1, x_2, \dots, x_n$ 爲一致且獨立。

樣本需 i.i.d. 才容易計算其抽樣分配。

【例】若 $x_1, x_2, \dots, x_n$ 皆爲參數 p 的白努力分配，且互相獨立，則

$$T = x_1 + x_2 + \cdots + x_n$$

爲參數 n、p 的二項分配。

【例題 1】設 $Z = X + Y$ ，其中 X、Y 的機率分配分別如下：

P(X,Y)

$X \setminus Y$	1	2	3
1	3/32	3/32	4/32
2	3/32	4/32	3/32
3	4/32	5/32	3/32

求 Z 的分配。

【解】

X	Y	$Z=X+Y$	$P(Z)=P(X+Y)$
1	1	2	3/32
1	2	3	3/32
1	3	4	4/32
2	1	3	3/32
2	2	4	4/32
2	3	5	3/32
3	1	4	4/32
3	2	5	5/32
3	3	6	3/32

$\Rightarrow$

Z	$P(Z)$
2	3/32
3	6/32
4	12/32
5	8/32
6	3/32

【例題 2】設 $Z = X + Y$ ，其中 X、Y 的機率分配分別如下：

X	$P(X)$
3	1/3
5	2/3

Y	$P(Y)$
2	1/2
4	1/2

X、Y 相互獨立，求 Z 的分配。

【解】

$P(X,Y)$

$X \backslash Y$	2	4
3	1/6	1/6
5	2/6	2/6

X	Y	$Z=X+Y$	$P(Z)$
3	2	5	1/6
3	4	7	1/6
5	2	7	2/6
5	4	9	2/6

$\Rightarrow$

Z	$P(Z)$
5	1/6
7	3/6
9	2/6

【例題 3】設 $Z = X_1 + X_2 + X_3$ ，其中 X_1, X_2, X_3 的機率分配分別皆如下：

X	$P(X)$
0	1/3
1	2/3

X_1, X_2, X_3 相互獨立，求 Z 的分配。

【解】

X_1	X_2	X_3	$Z=X_1+X_2+X_3$	$P(Z)$
0	0	0	0	1/27
0	0	1	1	2/27
0	1	0	1	2/27
0	1	1	2	4/27
1	0	0	1	2/27
1	0	1	2	4/27
1	1	0	2	4/27
1	1	1	3	8/27

$\Rightarrow$

Z	$P(Z)$
0	1/27
1	6/27
2	12/27
3	8/27

計算統計量的分配

給定一組隨機變數 $x_1, x_2, \ldots, x_n$ ，則求取統計量 $f\left(x_1, x_2, \ldots, x_n\right)$ 的分配有三種方法：

(1)累計分配函數法（*Cumulative-distribution-function Technique*）

(2)動差母函數法（*Moment-generating-function Technique*）

(3)變數變換法（*Transformations*）

兩個簡單的結果：

若 X,Y 的聯合機率函數為 $P(X,Y)$，且 $Z=X+Y$ 則

$$P(Z)=\sum_{X\text{ 範圍}} P(X, Z-X)=\sum_{Y\text{ 範圍}} P(Z-Y, Y)$$

若 X,Y 的聯合機率密度函數為 $f(x,y)$，且 $Z=X+Y$ 則

$$f(z)=\int_{x\text{ 範圍}} f(x, z-x)\,dx=\int_{y\text{ 範圍}} f(z-y, y)\,dy$$

【例題 4】若 $f(x,y)=xy,\ 0\le x\le 1,\ 0\le y\le 2$ ，且令 $z=x+y$ ，求

(a) $f(x)$ 、(b) $f(y)$ 、(c) $f(x|y)$ 、(d) $f(z)$ 。

【解】

$$(a)\, f(x)=\int_{y=-\infty}^{\infty} f(x,y)\,dy=\int_{y=0}^{2} xy\,dy = x\times \tfrac{1}{2}y^2\Big|_0^2 = 2x$$

$$(b)\, f(y)=\int_{x=-\infty}^{\infty} f(x,y)\,dx=\int_{x=0}^{1} xy\,dx = y\times \tfrac{1}{2}x^2\Big|_0^1 = \tfrac{1}{2}y$$

$$(c)\, f(x|y)=\frac{f(x,y)}{f(y)}=\frac{xy}{\frac{1}{2}y}=2x$$

$$(d)\, f(z)=\int_{x=-\infty}^{\infty} f(x,z-x)\,dx$$

$$=\begin{cases}\int_{x=0}^{1} x(z-x)\,dx=\frac{1}{2}z-\frac{1}{3} & 0\le z\le 1\\ \int_{x=0}^{z-1} x(z-x)\,dx=\frac{1}{2}z(z-1)^2-\frac{1}{3}(z-1)^3 & 1<z\le 2\\ \int_{x=z-2}^{1} x(z-x)\,dx=\frac{1}{2}z\left[1-(z-2)^2\right]-\frac{1}{3}\left[1-(z-2)^3\right] & 2<z\le 3\end{cases}$$

計算統計量的期望值與變異數

定理：

令 $Z=X_1+X_2+\cdots+X_n$ ，若 $X_1,X_2,\ldots,X_n$ 互相獨立，則

$$E(Z)=E(X_1)+E(X_2)+\cdots+E(X_n)$$

$$Var(Z)=Var(X_1)+Var(X_2)+\cdots+Var(X_n)$$

定理：

對任一隨機變數 X，與任一實數值常數 $r\in R$ ，則

$$E(rX)=rE(X)$$

$$Var(rX)=r^2Var(X)$$

定理：

令 $Z=\dfrac{X_1+X_2+\cdots+X_n}{n}$ ，若 $X_1,X_2,\ldots,X_n$ 互相獨立，則

$$E(Z)=\frac{E(X_1)+E(X_2)+\cdots+E(X_n)}{n}$$

$$Var(Z)=\frac{Var(X_1)+Var(X_2)+\cdots+Var(X_n)}{n^2}$$

定理：

令 $Z=\dfrac{X_1+X_2+\cdots+X_n}{n}$ ，若 $X_1,X_2,\ldots,X_n$ 互相獨立，且

$$E(X_1)=E(X_2)=\cdots=E(X_n)=\mu$$

$$Var(X_1)=Var(X_2)=\cdots=Var(X_n)=\sigma^2$$

則

$$E(Z)=\frac{E(X_1)+E(X_2)+\cdots+E(X_n)}{n}=\frac{n\mu}{n}=\mu$$

$$Var(Z)=\frac{Var(X_1)+Var(X_2)+\cdots+Var(X_n)}{n^2}=\frac{n\sigma^2}{n^2}=\frac{\sigma^2}{n}$$

【例題 5】設 $Z=X+Y$，其中

X 爲 $n=2$、$p=0.5$ 的二項分配，Y 爲 $n=3$、$p=0.3$ 的二項分配，

且 X、Y 相互獨立，計算 $E(Z)$、$Var(Z)$。

【解】

$$E(Z)=E(X+Y)=E(X)+E(Y)=2\times 0.5+3\times 0.3=1.9$$

$$\begin{aligned}Var(Z)&=Var(X+Y)=Var(X)+Var(Y)\\&=2\times 0.5\times(1-0.5)+3\times 0.3\times(1-0.3)=1.13\end{aligned}$$

【例題 6】設 $Z=X+Y$，其中 X、Y 爲相互獨立的標準常態分配，計算 $E(Z)$、$Var(Z)$。

【解】

$$E(Z)=E(X)+E(Y)=0+0=0$$

$$Var(Z)=Var(X)+Var(Y)=1^2+1^2=2$$

【例題 7】設 $Z=X_1+X_2+X_3$，

X_1 爲 $\lambda=2$ 的卜瓦松分配，X_2 爲 $\lambda=1$ 的卜瓦松分配，X_3 爲 $\lambda=0.5$ 的卜瓦松分配，

計算 $E(Z)$、$Var(Z)$。

【解】

$$E(Z)=E(X_1)+E(X_2)+E(X_3)=2+1+0.5=3.5$$

$$Var(Z)=Var(X_1)+Var(X_2)+Var(X_3)=2+1+0.5=3.5$$

【例題 8】設 $Z=X_1+X_2+X_3+X_4+X_5$，

$X_1,X_2,\ldots,X_5$ 分別爲 $n=10$、$p=0.4$，相互獨立的二項分配，

計算 $E(Z)$、$Var(Z)$。

【解】

$$E(Z)=5\times E(X)=5\times n\times p=5\times 10\times 0.4=20$$

$$Var(Z)=5\times Var(X)=5\times n\times p\times(1-p)=5\times 10\times 0.4\times 0.6=12$$

4.3 常見的抽樣分配

四個基本的抽樣分配：

(1) z 分配

標準常態分配，對稱、鐘形，$\mu=0$，$\sigma^2=1$

(2) χ^2 分配

卡方分配，右偏，$\mu=k$，$\sigma^2=2k$（自由度$=k$）

(3) F 分配

右偏分配，$\mu=\dfrac{n}{n-2}$，$\sigma^2=\dfrac{2n^2(m+n-2)}{m(m-2)^2(n-4)}$（自由度$=(m,n)$）

(4) t 分配

對稱、鐘形分配，$\mu=0$，$\sigma^2=\dfrac{k}{k-2}$（自由度$=k$）

> 若 $X_1,X_2,\ldots,X_n$ 皆為標準常態分配，且相互獨立，
> 則 $Y=X_1^2+X_2^2+\cdots+X_n^2$ 為自由度 n 的卡方分配 $\chi^2_{df=n}$。

> 若 X_1、X_2 皆為自由度 n_1、n_2 的卡方分配，且相互獨立，
> 則 $Y=\dfrac{X_1/n_1}{X_2/n_2}$ 為自由度 (n_1,n_2) 的 **F 分配** $F_{df=(n_1,n_2)}$。

> 若 X_1 為標準常態分配，X_2 為自由度 n 的卡方分配，且相互獨立，
> 則 $Y=\sqrt{\dfrac{X_1^2}{X_2/n}}$ 為自由度 n 的 **t 分配** $t_{df=n}$。

一些有用的結果：

> 若 $X_1\sim N(\mu_1,\sigma_1)$、$X_2\sim N(\mu_2,\sigma_2)$，且相互獨立，則
> $Y=X_1+X_2 \sim N\left(\mu_1+\mu_2,\sqrt{\sigma_1^2+\sigma_2^2}\right)$

> 若 $X_1,X_2,\ldots,X_n\sim N(\mu,\sigma)$，且相互獨立，則

$$T = X_1 + X_2 + \cdots + X_n \sim N\left(n\mu, \sqrt{n\sigma^2}\right)$$

$$\bar{X} = \frac{X_1 + X_2 + \cdots + X_n}{n} \sim N\left(\mu, \sqrt{\frac{\sigma^2}{n}}\right)$$

$$Y = \frac{\bar{X} - \mu}{\sigma/\sqrt{n}} \sim \text{z 分配}$$

若 $X_1, X_2, \ldots, X_n \sim N(\mu, \sigma)$，且相互獨立，則

$$Y = \frac{(n-1)s^2}{\sigma^2} \sim \chi^2_{df=n-1}$$

其中

$$s^2 = \frac{\left(X_1 - \bar{X}\right)^2 + \left(X_2 - \bar{X}\right)^2 + \cdots + \left(X_n - \bar{X}\right)^2}{n-1},\ \bar{X} = \frac{X_1 + X_2 + \cdots + X_n}{n}$$

若 $X_1, X_2, \ldots, X_{n_1}, Y_1, Y_2, \ldots, Y_{n_2} \sim N(\mu, \sigma)$，且相互獨立，則

$$Z = \frac{s_1^2}{s_2^2} \sim F_{df=(n_1-1, n_2-1)}$$

其中

$$s_1^2 = \frac{\left(X_1 - \bar{X}\right)^2 + \left(X_2 - \bar{X}\right)^2 + \cdots + \left(X_{n_1} - \bar{X}\right)^2}{n_1 - 1},\ \bar{X} = \frac{X_1 + X_2 + \cdots + X_{n_1}}{n_1}$$

$$s_2^2 = \frac{\left(Y_1 - \bar{Y}\right)^2 + \left(Y_2 - \bar{Y}\right)^2 + \cdots + \left(Y_{n_2} - \bar{Y}\right)^2}{n_2 - 1},\ \bar{Y} = \frac{Y_1 + Y_2 + \cdots + Y_{n_2}}{n_2}$$

若 $X_1, X_2, \ldots, X_n \sim N(\mu, \sigma)$，且相互獨立，則

$$Y = \frac{\bar{X} - \mu}{s/\sqrt{n}} \sim t_{df=n-1}$$

其中

$$s = \sqrt{\frac{\left(X_1 - \bar{X}\right)^2 + \left(X_2 - \bar{X}\right)^2 + \cdots + \left(X_n - \bar{X}\right)^2}{n-1}},\ \bar{X} = \frac{X_1 + X_2 + \cdots + X_n}{n}$$

【例題 9】設 $Z = X_1 + X_2 + \cdots + X_{10}$，其中 $X_1, X_2, \ldots, X_{10}$ 爲相互獨立的標準常態分配，

(1)右尾檢定，臨界值爲 5，請計算 α 值。

(2)右尾檢定，$\alpha = 0.2$，請計算臨界值。

【解】

Z 成爲 $\mu = 10 \times 0 = 0$ 、 $\sigma = \sqrt{10 \times 1^2} = \sqrt{10}$ 的常態分配。

(1)計算 $\alpha = P(Z \ge 5) = 0.0569$，

(2)計算 a 使得 $P(Z \ge a) = 0.2 \Rightarrow a = 2.6614$。

【例題 10】設 $Z = \dfrac{X_1 + X_2 + \cdots + X_{10}}{10}$，其中 $X_1, X_2, \ldots, X_{10}$ 為相互獨立的標準常態分配，

(1)右尾檢定，臨界值為 1，請計算 α 值。

(2)右尾檢定，$\alpha = 0.2$，請計算臨界值。

【解】

Z 成為 $\mu = 10 \times 0/10 = 0$、$\sigma = \sqrt{10 \times 1^2/10^2} = \sqrt{10}/10$ 的常態分配。

(1)計算 $\alpha = P(Z \ge 1) = 0.0008$，

(2)計算 a 使得 $P(Z \ge a) = 0.2 \Rightarrow a = 0.2661$。

【例題 11】設 $Z = X_1^2 + X_2^2 + \cdots + X_{10}^2$，其中 $X_1, X_2, \ldots, X_{10}$ 為相互獨立的標準常態分配，

(1)右尾檢定，臨界值為 15，請計算 α 值。

(2)右尾檢定，$\alpha = 0.1$，請計算臨界值。

【解】

Z 成為自由度 10 的 χ^2 分配。

(1)計算 $\alpha = P\left(Z = \chi^2_{df=10} \ge 15\right) = 0.1321$，

(2)以及計算 a 使得 $P\left(Z = \chi^2_{df=10} \ge a\right) = 0.1 \Rightarrow a = 15.9872$。

【例題 12】設 $Z = \dfrac{\left(X_1^2 + X_2^2 + X_3^2\right)/3}{\left(Y_1^2 + Y_2^2\right)/2}$，其中 X_1, X_2, X_3, Y_1, Y_2 為相互獨立的標準常態分配，

(1)右尾檢定，臨界值為 20，請計算 α 值。

(2)右尾檢定，$\alpha = 0.1$，請計算臨界值。

【解】

Z 成為自由度 $(3,2)$ 的 F 分配。

(1)計算 $\alpha = P\left(Z = F_{df=3,2} \ge 20\right) = 0.0480$，

(2)以及計算 a 使得 $P\left(Z = F_{df=3,2} \ge a\right) = 0.1 \Rightarrow a = 9.1618$。

【例題 13】設 $Z = \sqrt{\dfrac{X^2}{\left(Y_1^2 + Y_2^2 + Y_3^2\right)/3}}$，其中 X, Y_1, Y_2, Y_3 為相互獨立的標準常態分配，

(1)右尾檢定，臨界值為 3，請計算 α 值。

(2)右尾檢定，$\alpha = 0.1$，請計算臨界值。

【解】

Z 成為自由度3的 t 分配。

(1)計算 $\alpha = P\left(Z = t_{df=3} \ge 3\right) = 0.0288$，

(2)以及計算 a 使得 $P\left(Z = t_{df=3} \ge a\right) = 0.1 \Rightarrow a = 1.6377$。

【例題 14】設 $Z = \dfrac{\left(X_1 - \frac{X_1+X_2+X_3}{3}\right)^2 + \left(X_2 - \frac{X_1+X_2+X_3}{3}\right)^2 + \left(X_3 - \frac{X_1+X_2+X_3}{3}\right)^2}{\sigma^2}$，其中 X_1, X_2, X_3 分別為平均數 μ、標準差 σ，相互獨立的常態分配，

(1)右尾檢定，臨界值為 9，請計算 α 值。

(2)右尾檢定，$\alpha = 0.1$，請計算臨界值。

【解】

Z 成為自由度 $3-1=2$ 的 χ^2 分配。

(1)計算 $\alpha = P\left(Z = \chi^2_{df=2} \ge 9\right) = 0.0111$，

(2)以及計算 a 使得 $P\left(Z = \chi^2_{df=2} \ge a\right) = 0.1 \Rightarrow a = 4.6052$。

四個抽樣分配間的關係

符號：

z_α 表示累計機率為 α 之 z 分配臨界值：$P\left(x \le z_\alpha\right) = \alpha$

$\chi^2_{\alpha,\, df}$ 表示累計機率為 α、自由度 df 之 χ^2 分配臨界值：$P\left(x \le \chi^2_{\alpha,\, df}\right) = \alpha$

$F_{\alpha,\, (n_1,n_2)}$ 表示累計機率為 α、自由度 (n_1,n_2) 之 F 分配臨界值：$P\left(x \le F_{\alpha,\, (n_1,n_2)}\right) = \alpha$

$t_{\alpha,\, df}$ 表示累計機率為 α、自由度 df 之 t 分配臨界值：$P\left(x \le t_{\alpha,\, df}\right) = \alpha$

χ^2 分配的變化

自由度 $df = 1$ 時，成為 z 分配的平方：$\chi^2_{\alpha,\, df=1} = z^2_{\frac{\alpha}{2}}$

自由度 $df = \infty$ 時，$\dfrac{\chi^2_{\alpha,\, df}}{df}$ 為常數：$\lim\limits_{n\to\infty} \dfrac{\chi^2_{\alpha,\, df=n}}{n} = 1$

F 分配的變化

自由度 $df_1 = \left(n_1, n_2\right)$ 與 $df_2 = \left(n_2, n_1\right)$ 的 F 分配互為倒數：$F_{\alpha,\, df=(n_1,n_2)} \times F_{1-\alpha,\, df=(n_2,n_1)} = 1$

自由度 $df = \left(n_1, \infty\right)$ 時，成為 χ^2 分配除分子自由度：$F_{\alpha,\, df=(n_1,\infty)} = \dfrac{\chi^2_{\alpha,\, df=n_1}}{n_1}$

自由度 $df=(1,n_2)$ 時，成為 t 分配的平方：$F_{\alpha,\,df=(1,n_2)}=\dfrac{z_{\frac{\alpha}{2}}^2}{\chi^2_{\alpha,\,df=n_2}\big/n_2}=t^2_{\frac{\alpha}{2},\,df=n_2}$

t 分配的變化

自由度 $df=\infty$ 時，成為 z 分配：$t_{\alpha,\,df=\infty}=z_\alpha$

F 可由 F 分配得到：$F_{\alpha,\,df=(n_1,n_2)}=\dfrac{1}{F_{1-\alpha,\,df=(n_2,n_1)}}$

χ^2 可由 F 分配得到：$\chi^2_{\alpha,\,df=n}=n\times F_{\alpha,\,df=(n,\infty)}$

t 可由 F 分配得到：$t_{\alpha,\,df=n}=\sqrt{F_{2\alpha,\,df=(1,n)}}$

z 可由 F 、χ^2 或 t 分配得到：$z_\alpha=\sqrt{F_{2\alpha,\,df=(1,\infty)}}=\sqrt{\chi^2_{2\alpha,\,df=1}}=t_{\alpha,\,df=\infty}$

【例題 15】就以下 F 分配機率表，

請計算：(1) $F_{\alpha=0.88,df=(16,8)}$；(2) $\chi^2_{\alpha=0.06,df=12}$；(3) $t_{\alpha=0.06,df=16}$；(4) $z_{\alpha=0.03}$。

F 分配**右尾**機率表

n2 \ n1	α = 0.03				α = 0.06				α = 0.12			
	1	8	12	16	1	8	12	16	1	8	12	16
4	10.8744	8.1061	7.9072	7.8038	6.7640	5.4216	5.3131	5.2563	3.8858	3.5140	3.4692	3.4452
8	6.9370	4.1556	3.9442	3.8322	4.7924	3.2019	3.0664	2.9939	3.0284	2.3866	2.3149	2.2757
12	6.0550	3.3309	3.1163	3.0010	4.3116	2.6845	2.5403	2.4619	2.8018	2.0939	2.0126	1.9674
16	5.6718	2.9810	2.7641	2.6463	4.0966	2.4561	2.3070	2.2250	2.6977	1.9596	1.8729	1.8241
20	5.4580	2.7883	2.5695	2.4498	3.9749	2.3277	2.1753	2.0907	2.6379	1.8824	1.7923	1.7410
9999	4.7106	2.1275	1.8964	1.7632	3.5382	1.8704	1.7003	1.6006	2.4177	1.5968	1.4889	1.4238

【解】

(1) $F_{\alpha=0.88,df=(16,8)}=\dfrac{1}{F_{\alpha=0.12,df=(8,16)}}=\dfrac{1}{1.9596}=0.5103$

(2) $\chi^2_{\alpha=0.06,df=12}=12\times F_{\alpha=0.06,df=(12,\infty)}=12\times1.7003=20.4036$

(3) $t_{\alpha=0.06,df=16}=\sqrt{F_{\alpha=0.06\times2,df=(1,16)}}=\sqrt{2.6977}=1.6425$

(4) $z_{\alpha=0.03}=t_{\alpha=0.03,df=\infty}=\sqrt{F_{\alpha=0.06,df=(1,\infty)}}=\sqrt{3.5382}=1.8810$

抽樣分配的查表練習題目

四種機率（左尾、右尾、區間、雙尾）中，區間又稱為**信賴區間**（*confidence intervals*），

左尾、右尾、雙尾有時會稱為**拒絕區域**（*region of rejection*）。

一般拒絕區域不會包含等式：$P(X<a)$ 、$P(X>b)$ 、$P(X<a$ 或 $X>b)$。

左尾、右尾、雙尾的機率又稱為**顯著水準**（*significance level*），以 α 表示，如 $P(x\le a)=\alpha$ ，

左尾、右尾、雙尾的機率有時也稱爲 **p 值**（*p-value*）。

信賴區間的機率又稱爲**信賴度**（*level of confidence*），以$1-\alpha$ 表示，如$P(a \le x \le b)=1-\alpha$ 。

信賴區間的兩**臨界值**（*confidence limits*）一般都以平均數對稱點互相對稱。

【例題 16】隨機變數 X 爲標準常態分配：

(a)計算信賴區間$\{-1 \le X \le 1\}$的信賴度；

(b)計算左尾$\{X \le -2\}$的顯著水準；

(c)計算右尾$\{X \ge 3\}$的顯著水準；

(d)左尾檢定，$\alpha = 0.05$，求臨界值；

(e)右尾檢定，$\alpha = 0.05$，求臨界值；

(f)雙尾檢定，$\alpha = 0.05$，求臨界值。

【解】

$(a)\,1-\alpha = P(-1 \le z \le 1) = 0.6827$

$(b)\,\alpha = P(z \le -2) = 0.0228$

$(c)\,\alpha = P(z \ge 3) = 0.0013$

$(d)\,P(z \le a) = 0.05 \quad \Rightarrow \quad a = -1.645$

$(e)\,P(z \ge a) = 0.05 \quad \Rightarrow \quad a = 1.645$

$(f)\,P(-a \le z \le a) = 0.05 \quad \Rightarrow \quad a = 1.96$

【例題 17】隨機變數 X 爲 $\mu = 4$ ，$\sigma = 2$ 的常態分配：

(a)計算信賴區間$\{1 \le X \le 7\}$的信賴度；

(b)求信賴度$1-\alpha = 0.9$的信賴區間；

(c)左尾檢定，臨界值爲-1，求 p 值；

(d)右尾檢定，臨界值爲9 ，求 p 值；

(e)左尾檢定，$\alpha = 0.01$，求臨界值；

(f)右尾檢定，$\alpha = 0.01$，求臨界值；

(g)雙尾檢定，$\alpha = 0.01$，求臨界值。

【解】

$(a)\,z^*_{x=1} = \frac{1-4}{2} = -\frac{3}{2},\, z^*_{x=7} = \frac{7-4}{2} = \frac{3}{2}, \quad 1-\alpha = P\left(-\frac{3}{2} \le z \le \frac{3}{2}\right) = 0.8664$

$(b)\,1-\alpha = P(-z^* \le z \le z^*) = 0.9 \quad \Rightarrow \quad z^* = 1.645 \quad \Rightarrow \quad$ 臨界值$= 4 \pm 1.645 \times 2$

$\Rightarrow \quad a = 0.71, b = 7.29$

$(c)\ z^*_{x=-1}=\frac{-1-4}{2}=-\frac{5}{2},\quad p=P\left(z\le -\frac{5}{2}\right)=0.0062$

$(d)\ z^*_{x=9}=\frac{9-4}{2}=\frac{5}{2},\quad p=P\left(z\ge \frac{5}{2}\right)=0.0062$

$(e)\ \alpha=P\left(z\le z^*\right)=0.01 \quad\Rightarrow\quad z^*=-2.326 \quad\Rightarrow\quad a=4-2.326\times 2=-0.652$

$(f)\ \alpha=P\left(z\ge z^*\right)=0.01 \quad\Rightarrow\quad z^*=2.326 \quad\Rightarrow\quad b=4+2.326\times 2=8.652$

$(g)\ \alpha=P\left(z\le -z^* \text{ 或 } z\le z^*\right)=0.01 \quad\Rightarrow\quad z^*=2.576 \quad\Rightarrow\quad \text{臨界值}=4\pm 2.576\times 2$

$\Rightarrow\quad a=-1.152, b=9.152$

【例題 18】設 $Z=X+Y$ ，其中 X、Y 相互獨立，且

X 為 $\mu=1$ 、 $\sigma=3$ 的常態分配，Y 為 $\mu=2$ 、 $\sigma=4$ 的常態分配，

(a)計算信賴區間 $\{1\le Z\le 7\}$ 的信賴度；

(b)求信賴度 $1-\alpha=0.9$ 的信賴區間；

(c)左尾檢定，臨界值為 -1，求 p 值；

(d)右尾檢定，臨界值為 9 ，求 p 值；

(e)左尾檢定， $\alpha=0.01$，求臨界值、拒絕區域；

(f)右尾檢定， $\alpha=0.01$，求臨界值、拒絕區域；

(g)雙尾檢定， $\alpha=0.01$，求臨界值、拒絕區域。

【解】

Z 成為 $\mu=1+2=3$ 、 $\sigma=\sqrt{3^2+4^2}=5$ 的常態分配。

$(a)\ 1-\alpha=0.4436$ $\quad (b)\ \{-5.22\le Z\le 11.22\}$ $\quad (c)\ p=0.2119$

$(d)\ p=0.1151$ $\quad (e)\ \{Z<-8.63\}$ $\quad (f)\ \{Z>14.63\}$

$(g)\ \{Z<-9.88 \text{ 或 } Z>15.88\}$

【例題 19】設 $Z=X_1+X_2+\cdots+X_9$ ，其中 $X_1,X_2,\ldots,X_9$ 為相互獨立的標準常態分配，

(a)計算信賴區間 $\{1\le Z\le 7\}$ 的信賴度；

(b)求信賴度 $1-\alpha=0.9$ 的信賴區間；

(c)左尾檢定，臨界值為 -1，求 p 值；

(d)右尾檢定，臨界值為 9 ，求 p 值；

(e)左尾檢定， $\alpha=0.05$ ，求臨界值、拒絕區域；

(f)右尾檢定， $\alpha=0.05$ ，求臨界值、拒絕區域；

(g)雙尾檢定， $\alpha=0.05$ ，求臨界值、拒絕區域。

【解】

Z 成為 $\mu=9\times 0=0$ 、 $\sigma=\sqrt{9\times 1^2}=3$ 的常態分配。

$(a)\ 1-\alpha = 0.3596$ $(b)\ \{-4.93 \le Z \le 4.93\}$ $(c)\ p = 0.3694$
$(d)\ p = 0.0013$ $(e)\ \{Z < -4.93\}$ $(f)\ \{Z > 4.93\}$
$(g)\ \{Z < -5.88$ 或 $Z > 5.88\}$

【例題 20】設 $Z = \dfrac{X_1 + X_2 + \cdots + X_{16}}{16}$，其中 $X_1, X_2, \ldots, X_{16}$ 為相互獨立的標準常態分配，

(a)計算信賴區間 $\{-0.5 \le Z \le 0.5\}$ 的信賴度；

(b)求信賴度 $1-\alpha = 0.9$ 的信賴區間；

(c)左尾檢定，臨界值為 -1，求 p 值；

(d)右尾檢定，臨界值為 1，求 p 值；

(e)左尾檢定，$\alpha = 0.05$，求臨界值、拒絕區域；

(f)右尾檢定，$\alpha = 0.05$，求臨界值、拒絕區域；

(g)雙尾檢定，$\alpha = 0.05$，求臨界值、拒絕區域。

【解】

Z 成為 $\mu = \dfrac{16 \times 0}{16} = 0$ 、 $\sigma = \sqrt{\dfrac{16 \times 1^2}{16^2}} = \dfrac{1}{\sqrt{16}} = \dfrac{1}{4}$ 的常態分配。

$(a)\ 1-\alpha = 0.9545$ $(b)\ \{-0.4112 \le Z \le 0.4112\}$ $(c)\ p = 0.00003$
$(d)\ p = 0.00003$ $(e)\ \{Z < -0.4112\}$ $(f)\ \{Z > 0.4112\}$
$(g)\ \{Z < -0.4900$ 或 $Z > 0.4900\}$

【例題 21】設 $Z = X_1^2 + X_2^2 + \cdots + X_{10}^2$，其中 $X_1, X_2, \ldots, X_{10}$ 為相互獨立的標準常態分配，

(a)計算信賴區間 $\{2 \le Z \le 18\}$ 的信賴度；

(b)求信賴度 $1-\alpha = 0.9$ 的信賴區間；

(c)左尾檢定，臨界值為 2，求 p 值；

(d)右尾檢定，臨界值為 18，求 p 值；

(e)左尾檢定，$\alpha = 0.05$，求臨界值、拒絕區域；

(f)右尾檢定，$\alpha = 0.05$，求臨界值、拒絕區域；

(g)雙尾檢定，$\alpha = 0.05$，求臨界值、拒絕區域。

【解】

Z 成為自由度 10 的 χ^2 分配。

$(a)\ 1-\alpha = 0.9413$ $(b)\ \{3.94 \le Z \le 18.31\}$ $(c)\ p = 0.0037$
$(d)\ p = 0.0550$ $(e)\ \{Z < 3.94\}$ $(f)\ \{Z > 18.31\}$
$(g)\ \{Z < 3.25$ 或 $Z > 20.48\}$

【例題 22】設 $Z=\dfrac{\left(X_1^2+X_2^2+X_3^2\right)/3}{\left(Y_1^2+Y_2^2\right)/2}$ ，其中 X_1,X_2,X_3,Y_1,Y_2 爲相互獨立的標準常態分配，

(a)計算信賴區間 $\{2\le Z\le 18\}$ 的信賴度；

(b)求信賴度 $1-\alpha=0.9$ 的信賴區間；

(c)左尾檢定，臨界值爲 2 ，求 p 值；

(d)右尾檢定，臨界值爲 18 ，求 p 值；

(e)左尾檢定， $\alpha=0.05$ ，求臨界值、拒絕區域；

(f)右尾檢定， $\alpha=0.05$ ，求臨界值、拒絕區域；

(g)雙尾檢定， $\alpha=0.05$ ，求臨界值、拒絕區域。

【解】

Z 成爲自由度 $(3,2)$ 的 F 分配。

$(a)\ 1-\alpha=0.2974$　　$(b)\ \{0.105\le Z\le 19.16\}$　　$(c)\ p=0.6495$

$(d)\ p=0.0531$　　$(e)\ \{Z<0.105\}$　　$(f)\ \{Z>19.16\}$

$(g)\ \{Z<0.062$ 或 $Z>39.17\}$

【例題 23】設 $Z=\sqrt{\dfrac{X^2}{\left(Y_1^2+Y_2^2+Y_3^2\right)/3}}$ ，其中 X,Y_1,Y_2,Y_3 爲相互獨立的標準常態分配，

(a)計算信賴區間 $\{-2\le Z\le 2\}$ 的信賴度；

(b)求信賴度 $1-\alpha=0.9$ 的信賴區間；

(c)左尾檢定，臨界值爲 -3 ，求 p 值；

(d)右尾檢定，臨界值爲 3 ，求 p 值；

(e)左尾檢定， $\alpha=0.05$ ，求臨界值、拒絕區域；

(f)右尾檢定， $\alpha=0.05$ ，求臨界值、拒絕區域；

(g)雙尾檢定， $\alpha=0.05$ ，求臨界值、拒絕區域。

【解】

Z 成爲自由度 3 的 t 分配。

$(a)\ 1-\alpha=0.8607$　　$(b)\ \{-2.35\le Z\le 2.35\}$　　$(c)\ p=0.0288$

$(d)\ p=0.0288$　　$(e)\ \{Z<-2.35\}$　　$(f)\ \{Z>2.35\}$

$(g)\ \{Z<-3.182$ 或 $Z>3.182\}$

【例題 24】設 $Z=\dfrac{\left(X_1-\frac{X_1+X_2+X_3}{3}\right)^2+\left(X_2-\frac{X_1+X_2+X_3}{3}\right)^2+\left(X_3-\frac{X_1+X_2+X_3}{3}\right)^2}{\sigma^2}$ ，其中

X_1,X_2,X_3 分別爲平均數 μ 、標準差 σ ，相互獨立的常態分配，

(a)計算信賴區間$\{1 \le Z \le 6\}$的信賴度；

(b)求信賴度$1-\alpha = 0.9$的信賴區間；

(c)左尾檢定，臨界值為1，求p值；

(d)右尾檢定，臨界值為6，求p值；

(e)左尾檢定，$\alpha = 0.05$，求臨界值、拒絕區域；

(f)右尾檢定，$\alpha = 0.05$，求臨界值、拒絕區域；

(g)雙尾檢定，$\alpha = 0.05$，求臨界值、拒絕區域。

【解】

Z成為自由度$3-1=2$的χ^2分配。

$(a)\,1-\alpha = 0.6897$　$(b)\,\{0.352 \le Z \le 7.815\}$　$(c)\ p = 0.1987$

$(d)\ p = 0.1116$　$(e)\,\{Z < 0.352\}$　$(f)\,\{Z > 7.815\}$

$(g)\,\{Z < 0.216$ 或 $Z > 9.348\}$

【例題 25】設$Z = \dfrac{Y-\mu}{\sqrt{s^2}}$，其中

$$s^2 = \frac{\left(X_1 - \frac{X_1+X_2+X_3}{3}\right)^2 + \left(X_2 - \frac{X_1+X_2+X_3}{3}\right)^2 + \left(X_3 - \frac{X_1+X_2+X_3}{3}\right)^2}{3-1},$$

X_1, X_2, X_3, Y分別為平均數μ、標準差σ，相互獨立的常態分配，

(a)計算信賴區間$\{-2 \le Z \le 2\}$的信賴度；

(b)求信賴度$1-\alpha = 0.9$的信賴區間；

(c)左尾檢定，臨界值為-3，求p值；

(d)右尾檢定，臨界值為3，求p值；

(e)左尾檢定，$\alpha = 0.05$，求臨界值、拒絕區域；

(f)右尾檢定，$\alpha = 0.05$，求臨界值、拒絕區域；

(g)雙尾檢定，$\alpha = 0.05$，求臨界值、拒絕區域。

【解】

Z成為自由度$3-1=2$的t分配。

$(a)\,1-\alpha = 0.8165$　$(b)\,\{-2.92 \le Z \le 2.92\}$　$(c)\ p = 0.0918$

$(d)\ p = 0.0918$　$(e)\,\{Z < -2.92\}$　$(f)\,\{Z > 2.92\}$

$(g)\,\{Z < -4.30$ 或 $Z > 4.30\}$

【例題 26】設$Z = \dfrac{s_1^2}{s_2^2}$，其中

$$s_1^2 = \frac{\left(X_1 - \frac{X_1+X_2+X_3}{3}\right)^2 + \left(X_2 - \frac{X_1+X_2+X_3}{3}\right)^2 + \left(X_3 - \frac{X_1+X_2+X_3}{3}\right)^2}{3-1}$$

$$s_2^2 = \frac{\left(Y_1 - \frac{Y_1+Y_2+Y_3}{3}\right)^2 + \left(Y_2 - \frac{Y_1+Y_2+Y_3}{3}\right)^2 + \left(Y_3 - \frac{Y_1+Y_2+Y_3}{3}\right)^2}{3-1}$$

$X_1, X_2, X_3, Y_1, Y_2, Y_3$分別爲平均數$\mu$、標準差$\sigma$，相互獨立的常態分配，

(a)計算信賴區間$\{2 \le Z \le 18\}$的信賴度；

(b)求信賴度$1-\alpha = 0.9$的信賴區間；

(c)左尾檢定，臨界值爲2，求p值；

(d)右尾檢定，臨界值爲18，求p值；

(e)左尾檢定，$\alpha = 0.05$，求臨界值、拒絕區域；

(f)右尾檢定，$\alpha = 0.05$，求臨界值、拒絕區域；

(g)雙尾檢定，$\alpha = 0.05$，求臨界值、拒絕區域。

【解】

Z成爲自由度$(2,2)$的F分配。

$(a)\ 1-\alpha = 0.2807$　　$(b)\ \{0.053 \le Z \le 19.00\}$　　$(c)\ p = 0.6667$

$(d)\ p = 0.0526$　　$(e)\ \{Z < 0.053\}$　　$(f)\ \{Z > 19.00\}$

$(g)\ \{Z < 0.026$ 或 $Z > 39.00\}$

4.4　中央極限定理

若$X_1, X_2, \ldots, X_n$爲 i.i.d.，來自平均數μ、標準差σ的母體，令

$$\bar{X} = \frac{X_1 + X_2 + \cdots + X_n}{n}$$

則當$n \to \infty$時，

$$\bar{X} \sim N\left(\mu, \frac{\sigma}{\sqrt{n}}\right)$$

$$\frac{\bar{X} - \mu}{\sigma/\sqrt{n}} \sim N(0,1)$$

第五章　估計與信賴區間

2007 年 3 月 16 日 最後修改

5.1　估計概論
5.2　估計量的分配
5.3　信賴度、信賴區間與最大容忍誤差
5.4　優良估計量的評估準則
5.5　產生估計量的方法

5.1　估計概論

估計（*estimate*）：以樣本的結果推測母體參數的過程。

有資料依據、有方法的推測是估計，沒有根據的就是猜測。

最需要被估計的母體參數為平均數與標準差。

點估計（*point estimate*）：推測母體參數就是某個數值稱為點估計。

由樣本資料產生點估計數值的公式稱為估計量（*estimator*）。

將特定樣本資料帶入估計量所得的數值稱為估計值（*estimate*）。

【例如】假設樣本資料為 $x_1, x_2, \ldots, x_n$，最常用的母體平均數 μ 的估計量為

$$\bar{x} = \frac{x_1 + x_2 + \cdots + x_n}{n}$$

而最常用的母體標準差 σ 的估計量為

$$s = \sqrt{\frac{(x_1 - \bar{x})^2 + (x_2 - \bar{x})^2 + \cdots + (x_n - \bar{x})^2}{n-1}} = \sqrt{\frac{(x_1^2 + x_2^2 + \cdots + x_n^2) - n\bar{x}^2}{n-1}}$$

估計量是隨機變數的函數，因此估計量也是一個隨機變數。

估計區間（*estimate interval*）：推測母體參數應落入某區間，該區間稱為估計區間。

母體參數落入估計區間的機率與估計區間的大小成正向關係。

若知道估計量的分配，則可以計算母體參數落入某估計區間的機率。

估計區間也稱為信賴區間（*confidence interval*）。

該區間的機率稱為信賴度（*degree of confidence*），或信心水準（*level of confidence*）。

信賴度或信心水準一般以百分比來表示，如 90%信心水準、信賴度為 95%。

【例題 1】某樣本資料爲$\{10,8,14,9,12,7\}$，(a)請分別以$\bar{x}$、s^2作母體μ與σ^2的點估計；(b)若母體爲常態分配（$\sigma=2.6$），且估計區間爲$[8,12]$，請計算信賴度；(c)請計算信賴度爲 95%的信賴區間。

【解】

$$(a)\ \bar{x}=\frac{10+8+14+9+12+7}{6}=10$$

$$s^2=\frac{0^2+2^2+4^2+1^2+2^2+3^2}{6-1}=6.8$$

μ的點估計值爲 10，σ^2的點估計值爲 6.8。

$$(b)\ \bar{x}\text{ 爲 }\mu_{\bar{x}}=10,\ \sigma_{\bar{x}}=\frac{2.6}{\sqrt{6}}=1.06\text{ 的常態分配}$$

$$z_1^*=\frac{8-10}{1.06}=-1.89,\ z_2^*=\frac{12-10}{1.06}=1.89$$

$$P(8\le\bar{x}\le12)=P(-1.89\le z\le1.89)=0.9412=94.12\%$$

$(\bar{x}$ 爲 $\mu_{\bar{x}}=10,\ \sigma_{\bar{x}}=1.06$ 的常態分配，區間$=\{8\le\bar{x}\le12\}$，求機率。$)$

$$(c)\ 1-\alpha=P(-z^*\le z\le z^*)=95\%\quad\Rightarrow\quad z^*=1.96$$

$$\text{臨界值}=\mu_{\bar{x}}\pm z^*\sigma_{\bar{x}}=10\pm1.96\times1.06\quad\Rightarrow\quad\text{臨界值}=7.92\text{ 或 }12.08$$

$$\text{信賴區間}=\{7.92\le\bar{x}\le12.08\}$$

$(\bar{x}$ 爲 $\mu=10,\ \sigma_{\bar{x}}=1.06$ 的常態分配，$1-\alpha=95\%$，求信賴區間。$)$

5.2 估計量的分配

最常用的估計量爲$\bar{x}$與s^2：

$$\bar{x}=\frac{x_1+x_2+\cdots+x_n}{n}$$

$$s^2=\frac{(x_1-\bar{x})^2+(x_2-\bar{x})^2+\cdots+(x_n-\bar{x})^2}{n-1}$$

其中，$x_1,x_2,\ldots,x_n$互相獨立，來自參數爲μ、σ的常態母體。

(1)樣本平均數$\bar{x}$的分配：

(a) 母體標準差σ已知

$$\frac{\bar{x}-\mu_{\bar{x}}}{\sigma_{\bar{x}}}=\frac{\bar{x}-\mu}{\sigma/\sqrt{n}}\text{ 爲 } z \text{ 分配}$$

(b) 母體標準差 σ 未知，以樣本標準差 s 代替

$$\frac{\bar{x}-\mu_{\bar{x}}}{\sigma_{\bar{x}}}=\frac{\bar{x}-\mu}{s/\sqrt{n}}$$ 爲自由度 $n-1$ 的 t 分配

(2)樣本比例 $\bar{p}$ 的分配：

$$\frac{\bar{p}-\mu_{\bar{p}}}{\sigma_{\bar{p}}}=\frac{\bar{p}-p}{\sqrt{p(1-p)}/\sqrt{n}}$$ 爲 z 分配

(3)樣本變異數 s^2 的分配（母體標準差 σ 已知）：

$$\frac{(x_1-\bar{x})^2+\cdots+(x_n-\bar{x})^2}{\sigma^2}=\frac{(n-1)s^2}{\sigma^2}$$ 爲自由度 $n-1$ 的 χ^2 分配

若有兩組樣本：

$x_1,x_2,\ldots,x_{n_1}$ 互相獨立，來自參數爲 μ_1 、 σ_1 的常態母體；

$y_1,y_2,\ldots,y_{n_2}$ 互相獨立，來自參數爲 μ_2 、 σ_2 的常態母體；

$$\bar{x}=\frac{x_1+x_2+\cdots+x_{n_1}}{n_1},\quad s_1^2=\frac{(x_1-\bar{x})^2+(x_2-\bar{x})^2+\cdots+(x_{n_1}-\bar{x})^2}{n_1-1}$$

$$\bar{y}=\frac{y_1+y_2+\cdots+y_{n_2}}{n_2},\quad s_2^2=\frac{(y_1-\bar{y})^2+(y_2-\bar{y})^2+\cdots+(y_{n_2}-\bar{y})^2}{n_2-1}$$

(4)樣本平均數差 $\bar{x}-\bar{y}$ 的分配：

(a) 母體標準差 σ_1、σ_2 已知

$$\frac{(\bar{x}-\bar{y})-\mu_{\bar{x}-\bar{y}}}{\sigma_{\bar{x}-\bar{y}}}=\frac{(\bar{x}-\bar{y})-(\mu_1-\mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}}}$$ 爲 z 分配

(b) 母體標準差 σ_1、σ_2 未知且不等，以樣本標準差 s_1、s_2 代替

$$\frac{(\bar{x}-\bar{y})-\mu_{\bar{x}-\bar{y}}}{\sigma_{\bar{x}-\bar{y}}}=\frac{(\bar{x}-\bar{y})-(\mu_1-\mu_2)}{\sqrt{\frac{s_1^2}{n_1}+\frac{s_2^2}{n_2}}}$$ 爲自由度 n_p 的 t 分配

其中 n_p 由 $$\frac{\left(\frac{s_1^2}{n_1}+\frac{s_2^2}{n_2}\right)^2}{n_p}=\frac{\left(\frac{s_1^2}{n_1}\right)^2}{n_1-1}+\frac{\left(\frac{s_2^2}{n_2}\right)^2}{n_2-1}$$ 求得

(c) 母體標準差 σ_1、σ_2 未知但相等，以聯合估計量 s_p 代替

$$\frac{(\bar{x}-\bar{y})-\mu_{\bar{x}-\bar{y}}}{\sigma_{\bar{x}-\bar{y}}}=\frac{(\bar{x}-\bar{y})-(\mu_1-\mu_2)}{\sqrt{\frac{s_p^2}{n_1}+\frac{s_p^2}{n_2}}}$$ 為自由度 n_1+n_2-2 的 t 分配

其中 $s_p^2=\dfrac{(n_1-1)s_1^2+(n_2-1)s_2^2}{n_1+n_2-2}$

(5)樣本比例差 $\bar{p}_1-\bar{p}_2$ 的分配：

(a) 母體比例 p_1、p_2 未知且不等 $(p_1\neq p_2)$

$$\frac{(\bar{p}_1-\bar{p}_2)-\mu_{\bar{p}_1-\bar{p}_2}}{\sigma_{\bar{p}_1-\bar{p}_2}}=\frac{(\bar{p}_1-\bar{p}_2)-(p_1-p_2)}{\sqrt{\frac{\bar{p}_1(1-\bar{p}_1)}{n_1}+\frac{\bar{p}_2(1-\bar{p}_2)}{n_2}}}$$ 為自由度 n_p 的 t 分配

其中 n_p 由 $$\frac{\left(\frac{\bar{p}_1(1-\bar{p}_1)}{n_1}+\frac{\bar{p}_2(1-\bar{p}_2)}{n_2}\right)^2}{n_p}=\frac{\left(\frac{\bar{p}_1(1-\bar{p}_1)}{n_1}\right)^2}{n_1-1}+\frac{\left(\frac{\bar{p}_2(1-\bar{p}_2)}{n_2}\right)^2}{n_2-1}$$ 求得

(b) 母體比例 p_1、p_2 未知但相等 $(p_1=p_2)$

$$\frac{(\bar{p}_1-\bar{p}_2)-\mu_{\bar{p}_1-\bar{p}_2}}{\sigma_{\bar{p}_1-\bar{p}_2}}=\frac{(\bar{p}_1-\bar{p}_2)}{\sqrt{\frac{p_p(1-p_p)}{n_1+n_2-2}}}$$ 為自由度 n_1+n_2-2 的 t 分配

其中 $p_p=\dfrac{n_1\bar{p}_1+n_2\bar{p}_2}{n_1+n_2}$

(6)樣本變異數商 $\dfrac{s_1^2}{s_2^2}$ 的分配（母體標準差 $\sigma_1=\sigma_2=\sigma$ ）：

$$\frac{s_1^2}{s_2^2}=\frac{s_1^2/\sigma_1^2}{s_2^2/\sigma_2^2}=\frac{\chi_1^2/(n_1-1)}{\chi_2^2/(n_2-1)}$$ 為自由度 (n_1-1,n_2-1) 的 F 分配

5.3 信賴度、信賴區間與最大容忍誤差

信賴區間（*confidence interval*）：估計量的區間。

信賴度（*degree of confidence*，$1-\alpha$ ），、**信賴係數**（*confidence coefficient*）：

估計量之信賴區間的機率。

最大容忍誤差（maximum allowable error，ε ）：

信賴區間之臨界值與估計值（估計量之期望值）間的距離。

估計量$\overline{x}$是對稱分配，因此才有最大容忍誤差。

(1)估計量$\overline{x}$的最大容忍誤差：

(a) 母體標準差σ已知：$\dfrac{\overline{x}-\mu_{\overline{x}}}{\sigma_{\overline{x}}}$爲 z 分配

$$\varepsilon = z_{\frac{\alpha}{2}} \times \sigma_{\overline{x}} = z_{\frac{\alpha}{2}} \times \frac{\sigma}{\sqrt{n}}$$

其中$z_{\frac{\alpha}{2}}$爲機率α之雙尾臨界值，1-α爲信賴度

(b) 母體標準差σ未知：$\dfrac{\overline{x}-\mu_{\overline{x}}}{\sigma_{\overline{x}}}$爲 t 分配

$$\varepsilon = t_{\frac{\alpha}{2}} \times \sigma_{\overline{x}} = t_{\frac{\alpha}{2}, df=n-1} \times \frac{s}{\sqrt{n}}$$

其中$t_{\frac{\alpha}{2}}$爲機率α之雙尾臨界值，1-α爲信賴度

(2)估計量$\overline{p}$的最大容忍誤差：

(a) 不放回情況：$\dfrac{\overline{p}-\mu_{\overline{p}}}{\sigma_{\overline{p}}}$爲 z 分配

$$\varepsilon = z_{\frac{\alpha}{2}} \times \sigma_{\overline{p}} = z_{\frac{\alpha}{2}} \times \frac{\sqrt{p(1-p)}}{\sqrt{n}}$$

(b) 放回情況（有限母體）：$\dfrac{\overline{p}-\mu_{\overline{p}}}{\sigma_{\overline{p}}}$爲 z 分配

$$\varepsilon = z_{\frac{\alpha}{2}} \times \sigma_{\overline{p}} = z_{\frac{\alpha}{2}} \times \frac{\sqrt{p(1-p)}}{\sqrt{n}} \sqrt{\frac{N-n}{N-1}}$$

(3)估計量$\overline{x}-\overline{y}$的最大容忍誤差：

(a) 母體標準差σ_1, σ_2已知：$\dfrac{(\overline{x}-\overline{y})-\mu_{\overline{x}-\overline{y}}}{\sigma_{\overline{x}-\overline{y}}}$爲 z 分配

$$\varepsilon = z_{\frac{\alpha}{2}} \times \sigma_{\overline{x}} = z_{\frac{\alpha}{2}} \times \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}$$

(b) 母體標準差σ_1, σ_2未知，但不等：$\dfrac{(\overline{x}-\overline{y})-\mu_{\overline{x}-\overline{y}}}{\sigma_{\overline{x}-\overline{y}}}$爲 t 分配

$$\varepsilon = t_{\frac{\alpha}{2}} \times \sigma_{\overline{x}} = t_{\frac{\alpha}{2}, df=n_p} \times \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}$$

(c) 母體標準差 σ_1, σ_2 未知，但相等： $\dfrac{(\overline{x}-\overline{y})-\mu_{\overline{x}-\overline{y}}}{\sigma_{\overline{x}-\overline{y}}}$ 為 t 分配

$$\varepsilon = t_{\frac{\alpha}{2}} \times \sigma_{\overline{x}} = t_{\frac{\alpha}{2}, df=n_1+n_2-2} \times \sqrt{\frac{s_p^2}{n_1}+\frac{s_p^2}{n_2}}$$

(4)估計量 $\overline{p}_1 - \overline{p}_2$ 的最大容忍誤差：

(a) 母體比例 p_1、p_2 未知且不等： $\dfrac{(\overline{p}_1-\overline{p}_2)-\mu_{\overline{p}_1-\overline{p}_2}}{\sigma_{\overline{p}_1-\overline{p}_2}}$ 為 t 分配

$$\varepsilon = t_{\frac{\alpha}{2}} \times \sigma_{\overline{p}_1-\overline{p}_2} = t_{\frac{\alpha}{2}, df=n_p} \times \sqrt{\frac{\overline{p}_1(1-\overline{p}_1)}{n_1}+\frac{\overline{p}_2(1-\overline{p}_2)}{n_2}}$$

(b) 母體比例 p_1、p_2 未知但相等： $\dfrac{(\overline{p}_1-\overline{p}_2)-\mu_{\overline{p}_1-\overline{p}_2}}{\sigma_{\overline{p}_1-\overline{p}_2}}$ 為 t 分配

$$\varepsilon = t_{\frac{\alpha}{2}} \times \sigma_{\overline{p}_1-\overline{p}_2} = t_{\frac{\alpha}{2}, df=n_1+n_2-2} \times \sqrt{\frac{p_p(1-p_p)}{n_1+n_2-2}}$$

其中 $p_p = \dfrac{n_1\overline{p}_1 + n_2\overline{p}_2}{n_1+n_2}$

最大容忍誤差 ε 與信賴度 $1-\alpha$ 、母體標準差 σ 、樣本數 n 有關。

大樣本情況：

因 $\lim\limits_{n\to\infty} t_{\frac{\alpha}{2}, df=n-1} = z_{\frac{\alpha}{2}}$ ，故若自由度太大以致查表找不到 $t_{\frac{\alpha}{2}, df=n-1}$ ，則以 $z_{\frac{\alpha}{2}}$ 代替。

由最大容忍誤差寫出信賴區間：

$$\text{信賴區間} = C.I. = \left\{\mu_{\overline{x}} - \varepsilon \le \overline{x} \le \mu_{\overline{x}} + \varepsilon\right\}$$

【例題 2】某校想瞭解該校同學每週用於讀書的時間多寡，經抽取 49 位同學，初步統計發現每週讀書時間的樣本平均數為 24 小時，樣本標準差為 4 小時。請以 95%的信賴度估計該校同學每週的讀書時間為何？

【解】

$n=49,\ \overline{x}=24,\ s=4 \quad \Rightarrow \quad$ 令 $\mu=\overline{x}=24, \sigma=s=4$

$1-\alpha=95\% \quad \Rightarrow \quad t_{\frac{\alpha}{2}, df=49-1} \doteq z_{\frac{\alpha}{2}} = 1.96$

$\Rightarrow \varepsilon = t_{\frac{\alpha}{2}, df=48}\dfrac{s}{\sqrt{n}} = 1.96 \times \dfrac{4}{\sqrt{49}} = 1.12$

$\Rightarrow$ 信賴區間 $=\left\{\mu-\varepsilon \le \overline{x} \le \mu+\varepsilon\right\} = \left\{22.88 \le \overline{x} \le 25.12\right\}$

【例題 3】調查 500 位科技新貴，其中有 175 人想在宜蘭置產，準備退休後住那裡。請以 98%的信心水準估計有意到宜蘭退休之科技新貴的比例。

【解】

$$\overline{p}=\frac{175}{500}=0.35 \quad \rightarrow \quad p=0.35$$

$$1-\alpha=98\% \quad \rightarrow \quad z_{\frac{\alpha}{2}}=2.33$$

$$\Rightarrow \varepsilon=z_{\frac{\alpha}{2}}\frac{\sqrt{p(1-p)}}{\sqrt{n}}=2.33\times\frac{\sqrt{0.35\times(1-0.35)}}{\sqrt{500}}=0.0497$$

$$\Rightarrow C.I.=\left\{p-\varepsilon\le\overline{p}\le p+\varepsilon\right\}=\left\{0.35-0.0497\le\overline{p}\le 0.35+0.0497\right\}$$

【例題 4】某學生人數 500 人的學校想瞭解該校同學每週用於讀書的時間多寡，經抽取 49 位同學，初步統計發現每週讀書時間的樣本平均數爲 24 小時，樣本標準差爲 4 小時。請以 95%的信賴度估計該校同學每週的讀書時間爲何？

【解】

$$n=49,\ \ \overline{x}=24,\ \ s=4,\ \ N=500 \quad \Rightarrow \quad 令 \ \ \mu=\overline{x}=24, \sigma=s=4$$

$$1-\alpha=95\% \quad \Rightarrow \quad t_{\frac{\alpha}{2},df=49-1}\doteq z_{\frac{\alpha}{2}}=1.96$$

$$\Rightarrow \varepsilon\approx z_{\frac{\alpha}{2}}\frac{s}{\sqrt{n}}\sqrt{\frac{N-n}{N-1}}=1.96\times\frac{4}{\sqrt{49}}\times\sqrt{\frac{500-49}{500-1}}=1.0648$$

$$\Rightarrow C.I.=\left\{\mu-\varepsilon\le\overline{x}\le\mu+\varepsilon\right\}=\left\{24-1.0648\le\overline{x}\le 24+1.0648\right\}$$

決定樣本數

最大容忍誤差 ε 、信賴度 $1-\alpha$ 、母體參數 σ 或 p、與樣本數 n 有以下的關係：

(a) 統計量 $\overline{x}$，母體標準差 σ 已知：

$$\varepsilon=z_{\frac{\alpha}{2}}\times\frac{\sigma}{\sqrt{n}} \quad \Rightarrow \quad n=\frac{z_{\frac{\alpha}{2}}^2\sigma^2}{\varepsilon^2}$$

(b) 統計量 $\overline{p}$，母體比例 p 已知：

$$\varepsilon=z_{\frac{\alpha}{2}}\times\frac{\sqrt{p(1-p)}}{\sqrt{n}} \quad \Rightarrow \quad n=\frac{z_{\frac{\alpha}{2}}^2 p(1-p)}{\varepsilon^2}$$

【例題 5】某市調單位想瞭解單人住家每月的電費，假設他們要求信心水準 99%以上，最大容忍誤差$5 以下，且根據以往的研究經驗，該類住家每月電費的標準差爲$20。請計算該研究的最低樣本數。

【解】

$\varepsilon = 5, \sigma = 20, 1-\alpha = 99\% \quad \rightarrow \quad z_{\frac{\alpha}{2}} = 2.58$

$\Rightarrow n = z_{\frac{\alpha}{2}}^2 \dfrac{\sigma^2}{\varepsilon^2} = 2.58^2 \times \dfrac{20^2}{5^2} \approx 107$

5.4 優良估計量的評估準則

四個常用的評估準則：

不偏性（Unbiased）、有效性（Efficiency）、一致性（Consistency）、充分性（Sufficiency）

不偏估計量（*Unbiased Estimators*）：估計量的期望值等於其所估計之母體參數

令 $\hat{\theta}$ 為 θ 的估計量，若

$$E\left(\hat{\theta}\right) = \theta$$

則稱 $\hat{\theta}$ 為 θ 的**不偏估計量**。

有效估計量（*Efficient Estimators*）：

在不偏估計的前提下，變異數小稱為有效性

兩種有效性：**相對有效性**（*Relative Efficiency*）、**絕對有效性**（*Absolute Efficiency*）

令 $\hat{\theta}_1$ 、$\hat{\theta}_2$ 皆為 θ 的不偏估計量，且 $V\left(\hat{\theta}_1\right)$ 、$V\left(\hat{\theta}_2\right)$ 分別為其變異數，定義

$$\text{eff}\left(\hat{\theta}_1, \hat{\theta}_2\right) = \frac{V\left(\hat{\theta}_2\right)}{V\left(\hat{\theta}_1\right)}$$

為、之**相對有效性比值**（*Ratio of Relative Efficiency*）。若

$$\text{eff}\left(\hat{\theta}_1, \hat{\theta}_2\right) > 1$$

則表示 $\hat{\theta}_1$ 相對有效於 $\hat{\theta}_2$ 。

一致估計量（*Consistent Estimators*）：樣本數夠大的話，估計誤差會很小

令 $\hat{\theta}_n$ 爲 θ 在樣本數 n 的估計量，對任意 $\varepsilon > 0$，若

$$\lim_{n\to\infty} P\left(\left|\hat{\theta}_n - \theta\right| \le \varepsilon\right) = 1$$

則稱 $\hat{\theta}_n$ 爲 θ 的一致估計量。

令 $\hat{\theta}_n$ 爲 θ 的不偏估計量，且

$$\lim_{n\to\infty} V\left(\hat{\theta}_n\right) = 0$$

則稱 $\hat{\theta}_n$ 爲 θ 的一致估計量。

充分估計量（*Sufficient Estimators*）：估計量反應了所有母體參數的資訊

令 $\hat{\theta} = \hat{\theta}\left(x_1,\ldots,x_n\right)$ 爲 θ 的不偏估計量，若

$P\left(x_1,\ldots,x_n \middle| \hat{\theta} = a\right)$ 不是 θ 的函數，

則稱 $\hat{\theta}$ 爲 θ 的**充分估計量**。

令 $\hat{\theta} = \hat{\theta}\left(x_1,\ldots,x_n\right)$ 爲 θ 的充分估計量，若且唯若 $L\left(x_1,\ldots,x_n \middle| \theta\right) = g\left(\hat{\theta},\theta\right) \times h\left(x_1,\ldots,x_n\right)$ 其中 $L\left(x_1,\ldots,x_n \middle| \theta\right) = f\left(x_1 \middle| \theta\right) \cdots f\left(x_n \middle| \theta\right)$ 爲 $x_1,\ldots,x_n$ 的概似函數。

【例題 6】試證 $\overline{X}$ 、 s^2 分別爲 μ 、 σ^2 的不偏估計量。

【解】

令 $X_1,\ldots,X_n$ 爲互相獨立，參數 μ 、 σ 的常態分配，

$$\overline{X} = \frac{X_1 + \cdots + X_n}{n} \text{ ， } s^2 = \frac{\left(X_1 - \overline{X}\right)^2 + \cdots + \left(X_n - \overline{X}\right)^2}{n-1}$$

則

$$E\left[\overline{X}\right] = E\left(\frac{X_1 + \cdots + X_n}{n}\right) = \frac{E\left(X_1\right) + \cdots + E\left(X_n\right)}{n} = \frac{\mu + \cdots + \mu}{n} = \mu$$

故 $\overline{X}$ 爲 μ 的不偏估計量。又

$$E\left[s^2\right]=E\left(\frac{\left(X_1-\bar{X}\right)^2+\cdots+\left(X_n-\bar{X}\right)^2}{n-1}\right)=\frac{E\left(X_1-\bar{X}\right)^2+\cdots+E\left(X_n-\bar{X}\right)^2}{n-1}$$
$$=\frac{\frac{n-1}{n}\sigma^2+\cdots+\frac{n-1}{n}\sigma^2}{n-1}=\sigma^2$$

其中

$$E\left(X_k-\bar{X}\right)^2=E\left(X_k^2\right)-2E\left(X_k\frac{\left(X_1+\cdots+X_n\right)}{n}\right)+E\left(\frac{\left(X_1+\cdots+X_n\right)^2}{n^2}\right)$$
$$=\left(\sigma^2+\mu^2\right)-\tfrac{2}{n}\left(\sigma^2+n\mu^2\right)+\tfrac{1}{n^2}\left(n\sigma^2+n^2\mu^2\right)=\tfrac{n-1}{n}\sigma^2$$
$$E\left(X_iX_j\right)=\begin{cases}\sigma^2+\mu^2 & \text{when } i=j\\ \mu^2 & \text{when } i\neq j\ \left(X_i,X_j\ \text{互相獨立}\right)\end{cases}$$

故 s^2 爲 σ^2 的不偏估計量。

【例題 7】令 $X_1,\ldots,X_n$ 爲互相獨立，參數 μ 、 σ 的常態分配，已知

$$Y_1=\left(X_1+\cdots+X_n\right)/n \text{ 、 } Y_2=\left(X_1+X_n\right)/2$$

皆爲 μ 的不偏估計量，是比較 Y_1 、 Y_2 的相對有效性。

【解】

因

$$\text{eff}\left(Y_1,Y_2\right)=\frac{V\left(Y_2\right)}{V\left(Y_1\right)}=\frac{\sigma^2/2}{\sigma^2/n}=\frac{n}{2}>1$$

其中

$$V\left(Y_1\right)=V\left(\frac{X_1+\cdots+X_n}{n}\right)=\frac{\sigma^2+\cdots+\sigma^2}{n^2}=\frac{\sigma^2}{n}$$
$$V\left(Y_2\right)=V\left(\frac{X_1+X_n}{2}\right)=\frac{\sigma^2+\sigma^2}{2^2}=\frac{\sigma^2}{2}$$

故 Y_1 相對有效於 Y_2 。

【例題 8】令 $X_1,\ldots,X_n$ 爲互相獨立，參數 μ 、 σ 的常態分配，已知

$$Y_1=\left(X_1+\cdots+X_n\right)/n \text{ 、 } Y_2=\left(X_1+X_n\right)/2$$

皆爲μ的不偏估計量，請檢查Y_1、Y_2是否爲之一致估計量。

【解】

因

$$\lim_{n\to\infty} V\left(Y_1\right)=\lim_{n\to\infty} V\left(\tfrac{X_1+\cdots+X_n}{n}\right)=\lim_{n\to\infty}\frac{\sigma^2}{n}=0$$

故Y_1爲μ的一致估計量；而

$$\lim_{n\to\infty} V\left(Y_2\right)=\lim_{n\to\infty} V\left(\tfrac{X_1+X_n}{2}\right)=\lim_{n\to\infty}\frac{\sigma^2}{2}=\frac{\sigma^2}{2}>0$$

故Y_2不是μ的一致估計量。

【例題 9】令$X_1,\ldots,X_n$爲互相獨立，參數p的白努力分配，已知

$$Y_1=\left(X_1+\cdots+X_n\right)/n \text{ 、 } Y_2=\left(X_1+X_n\right)/2$$

皆爲p的不偏估計量，請檢查Y_1、Y_2是否爲充分估計量。

【解】

（**機率密度函數法**）因

$$\begin{aligned}
&P\left(X_1=x_1,\ldots,X_n=x_n \middle| Y_1=\tfrac{y}{n}\right)\\
&=\left[p^{x_1+\cdots+x_n}\left(1-p\right)^{n-\left(x_1+\cdots+x_n\right)}\right]\Big/\left[C_y^n p^y\left(1-p\right)^{n-y}\right]\\
&=\left[p^y\left(1-p\right)^{n-y}\right]\Big/\left[C_y^n p^y\left(1-p\right)^{n-y}\right]\\
&=1\big/C_y^n \qquad \left(\text{不是一個 } p \text{ 的函數}\right)
\end{aligned}$$

故Y_1爲p的充分估計量；而

$$\begin{aligned}
&P\left(X_1=x_1,\ldots,X_n=x_n \middle| Y_2=\tfrac{y}{2}\right)\\
&=\left[p^{\left(x_2+\cdots+x_{n-1}\right)+y}\left(1-p\right)^{n-\left(x_2+\cdots+x_{n-1}\right)-y}\right]\Big/\left[C_y^2 p^y\left(1-p\right)^{2-y}\right]\\
&=\left[p^{\left(x_2+\cdots+x_{n-1}\right)}\left(1-p\right)^{n-2-\left(x_2+\cdots+x_{n-1}\right)}\right]\Big/C_y^2 \quad \text{（是 } p \text{ 的函數）}
\end{aligned}$$

故Y_2不是p的充分估計量。

（**概似函數法**）因

$$
\begin{aligned}
&L\left(X_1=x_1,\ldots,X_n=x_n \,\middle|\, p, Y_1=\tfrac{y}{n}\right)\\
&=\left[p^{x_1}(1-p)^{1-x_1}\right]\cdots\left[p^{x_n}(1-p)^{1-x_n}\right]\\
&=p^{y}(1-p)^{n-y} \quad (\text{可以將 } p \text{ 與 } x_1,\ldots,x_n \text{分離})
\end{aligned}
$$

故 Y_1 爲 p 的充分估計量；而

$$
\begin{aligned}
&L\left(X_1=x_1,\ldots,X_n=x_n \,\middle|\, p, Y_2=\tfrac{y}{2}\right)\\
&=\left[p^{x_1}(1-p)^{1-x_1}\right]\cdots\left[p^{x_n}(1-p)^{1-x_n}\right]\\
&=p^{x_2+\cdots+x_{n-1}+y}(1-p)^{n-(x_2+\cdots+x_{n-1})-y} \quad (\text{無法分離 } p \text{ 與 } x_1,\ldots,x_n)
\end{aligned}
$$

故 Y_2 不是 p 的充分估計量。

5.5 產生估計量的方法

有兩種產生估計量的方法：

最大概似法（*Maximum Likelihood*）：讓該組樣本的出現機率最大

動差法（*Methods of Moments*）：讓統計量的各級動差分別等於母體的相對動差

5.5.1 最大概似法

概似函數（*Likelihood Function*）：樣本發生的聯合機率函數（函數之變數爲母體參數）

令 $f(x)$ 爲母體的機率函數，$\theta_1,\ldots,\theta_k$ 爲其參數；又令 $x_1,x_2,\ldots,x_n$ 爲一組樣本，則

$$f\left(x_i \mid \theta_1,\ldots,\theta_k\right)$$

爲 $x_i, i=1,2,\ldots,n$ 發生的條件機率，而

$$f\left(x_1 \mid \theta_1,\ldots,\theta_k\right) f\left(x_2 \mid \theta_1,\ldots,\theta_k\right)\cdots f\left(x_n \mid \theta_1,\ldots,\theta_k\right)$$

爲 $x_1,x_2,\ldots,x_n$ 同時發生的聯合機率。令

$$L\left(\theta_1,\ldots,\theta_k\right)=f\left(x_1 \mid \theta_1,\ldots,\theta_k\right) f\left(x_2 \mid \theta_1,\ldots,\theta_k\right)\cdots f\left(x_n \mid \theta_1,\ldots,\theta_k\right)$$

則稱 $L\left(\theta_1,\ldots,\theta_k\right)$ 爲機率分配 $f(x)$ 在樣本 $x_1,x_2,\ldots,x_n$ 下的概似函數。

最大概似估計量（*Maximum Likelihood Estimators*）：

母體參數之估計量爲使概似函數值極大時之參數。

令 $L\left(\theta_1,\ldots,\theta_k\right)$ 爲概似函數，則

$$\begin{cases} \dfrac{\partial}{\partial\theta_1}L\left(\theta_1,\ldots,\theta_k\right)=0 \\ \qquad\vdots \\ \dfrac{\partial}{\partial\theta_k}L\left(\theta_1,\ldots,\theta_k\right)=0 \end{cases}$$

之 $\hat{\theta}_1,\ldots,\hat{\theta}_k$ 解爲 $\theta_1,\ldots,\theta_k$ 的最大概似估計量。

【例題 10】請寫出參數爲 n、p 之白努力分配的概似函數。

【解】

白努力分配的機率函數爲 $f\left(x\right)=p^x\left(1-p\right)^{1-x},x\in\left\{0,1\right\}$ ，參數 p 爲發生機率，則

$$\begin{aligned} L\left(p\right)&=\left[p^{x_1}\left(1-p\right)^{1-x_1}\right]\left[p^{x_2}\left(1-p\right)^{1-x_2}\right]\cdots\left[p^{x_n}\left(1-p\right)^{1-x_n}\right] \\ &=p^{x_1+x_2+\cdots+x_n}\left(1-p\right)^{n-\left(x_1+x_2+\cdots+x_n\right)} \end{aligned}$$

爲白努力分配之 p 的概似函數。

【例題 11】請寫出參數爲 p 的白努力分配之最大概似估計量。

【解】

白努力分配參數 p 的概似函數爲

$$L\left(p\right)=p^{x_1+x_2+\cdots+x_n}\left(1-p\right)^{n-\left(x_1+x_2+\cdots+x_n\right)}$$

令

$$L^*=\ln L\left(p\right)=\left(x_1+\cdots+x_n\right)\ln p+\left[n-\left(x_1+\cdots+x_n\right)\right]\ln\left(1-p\right)$$

則

$$\begin{aligned} &\frac{d}{dp}L^*=\frac{\left(x_1+\cdots+x_n\right)}{p}-\frac{n-\left(x_1+\cdots+x_n\right)}{1-p}=0 \\ &\Rightarrow\hat{p}=\frac{x_1+\cdots+x_n}{n} \end{aligned}$$

即 $\hat{p}=\dfrac{x_1+\cdots+x_n}{n}$ 爲 p 的最大概似估計量。

【例題 12】請寫出參數爲 μ 、 σ 的常態分配之最大概似估計量。

【解】

參數爲 μ 、 σ^2 之常態分配，其概似函數爲

$$L(\mu,\sigma)=\left[\tfrac{1}{\sigma\sqrt{2\pi}}e^{-\frac{1}{2}\left(\frac{x_1-\mu}{\sigma}\right)^2}\right]\cdots\left[\tfrac{1}{\sigma\sqrt{2\pi}}e^{-\frac{1}{2}\left(\frac{x_n-\mu}{\sigma}\right)^2}\right]=\left(\tfrac{1}{\sigma^2 2\pi}\right)^{\frac{n}{2}}e^{-\frac{1}{2\sigma^2}\Sigma(x_i-\mu)^2}$$

則

$$L^*=\ln L(\mu,\sigma)=-\tfrac{n}{2}\ln\sigma^2-\tfrac{n}{2}\ln 2\pi-\tfrac{1}{2\sigma^2}\Sigma(x_i-\mu)^2$$

$$\begin{cases}\dfrac{d}{d\mu}L^*=\tfrac{1}{2\sigma^2}\Sigma(x_i-\mu)^2=0\\ \dfrac{d}{d\sigma^2}L^*=-\tfrac{n}{2}\tfrac{1}{\sigma^2}+\tfrac{1}{2\sigma^4}\Sigma(x_i-\mu)^2=0\end{cases}$$

$$\Rightarrow\begin{cases}\hat{\mu}=\tfrac{1}{n}\Sigma x_i=\overline{x}\\ \hat{\sigma}^2=\tfrac{1}{n}\Sigma(x_i-\overline{x})^2\end{cases}$$

即 $\hat{\mu}=\frac{1}{n}\Sigma x_i=\overline{x}$ 、 $\hat{\sigma}^2=\frac{1}{n}\Sigma(x_i-\overline{x})^2$ 分別爲 μ 、 σ^2 的最大概似估計量。

5.5.2 動差法

r 級動差（r^{th} *Moments*）： $\mu_r'=E(X^r)$ 稱爲隨機變數 X 的 r 級動差

中央動差（r^{th} *Central Moments*）： $\mu_r=E\left[(X-\mu_X)^r\right]$ 稱爲隨機變數 X 的 r 級中央動差

動稱母函數（*Moment Generating Function*）：

$m(t)=E(e^{tX})$ 稱爲隨機變數 X 的動差母函數（或翻譯成動差生成函數）

動差母函數可以用來產生各級動差

$$m'(t)=\frac{d}{dt}m(t)=E(Xe^{tX})\Rightarrow m'(0)=E(Xe^{tX})\Big|_{t=0}=E(X)=\mu_1'$$

$$m''(t)=\frac{d^2}{dt^2}m(t)=E(X^2e^{tX})\Rightarrow m''(0)=E(X^2e^{tX})\Big|_{t=0}=E(X^2)=\mu_2'$$

$$\vdots$$

$$m^{(r)}(0)=E(X^r)=\mu_r'$$

以動差法找估計量：

(1)令樣本動差等於母體動差，有 k 個參數則寫出前 k 級動差；

(2)將母體動差以參數表示；

(3)解聯立方程式，各參數之解就是其估計量。

一、二級動差與 μ 、σ^2 的關係

$$\mu = E(X) = \mu_1'$$

$$\sigma^2 = E(X-\mu)^2 = E(X^2 - 2\mu X + \mu^2) = \mu_2' - (\mu_1')^2$$

$$E(X-\mu)^3 = E(X^3 - 3\mu X^2 + 3\mu^2 X - \mu^3) = \mu_3' - 3\mu\mu_2' + 3\mu^2\mu_1' - \mu^3$$

離散機率的動差母函數

分配名稱	機率密度函數	變數範圍	期望值	變異數	動差母函數
白努力	$f(x) = p^x(1-p)^{1-x}$	$x = 0,1$	p	$p(1-p)$	$pe^t + 1 - p$
二項	$f(x) = C_x^n p^x (1-p)^{1-x}$	$x = 0,1,\ldots,n$	np	$np(1-p)$	$(pe^t + 1 - p)^n$
超幾何	$f(x) = C_x^S C_{n-x}^{N-S} / C_n^N$	$x = 0,1,\ldots,n$	$n\frac{S}{N}$	$n\frac{S}{N}\frac{N-S}{N}\frac{N-n}{N-1}$	
卜瓦松	$f(x) = \lambda^x e^{-\lambda} / x!$	$x = 0,1,\ldots$	λ	λ	$e^{\lambda e^t} / e^{\lambda}$
幾何	$f(x) = p(1-p)^x$	$x = 0,1,\ldots$	$\frac{1-p}{p}$	$\frac{1-p}{p^2}$	$\frac{p}{1-(1-p)e^t}$
負二項	$f(x) = C_x^{r+x-1} p^r (1-p)^x$	$x = 0,1,\ldots$	$r\frac{1-p}{p}$	$r\frac{1-p}{p^2}$	$\left(\frac{p}{1-(1-p)e^t}\right)^r$

【例題 13】請寫出參數 n、p 之二項分配的動差母函數，並求其第一、二級動差。

【解】

參數 n、p 的二項分配，其機率函數爲

$$f(x) = C_x^n p^x (1-p)^{n-x} = \frac{n!}{x!(n-x)!} p^x (1-p)^{n-x}, \quad x = 0,1,\ldots,n$$

$$\sum_{x=0}^{n} C_x^n p^x (1-p)^{n-x} = [p + (1-p)]^n = 1$$

第一、二級動差分別爲

$$\mu_1' = E(X) = \sum_{x=0}^{n} x C_x^n p^x (1-p)^{n-x} = np \sum_{x=1}^{n} x C_{x-1}^{n-1} p^{x-1} (1-p)^{n-x} = np$$

$$\mu_2' = E(X^2) = E(X(X-1)) + E(X)$$

$$= \sum_{x=0}^{n} x(x-1) C_x^n p^x (1-p)^{n-x} + np = n(n-1)p^2 + np$$

動差母函數為

$$m(t)=E\left(e^{tx}\right)=\sum_{x=0}^{n} e^{tx} C_x^n p^x (1-p)^{n-x}$$
$$=\sum_{x=0}^{n} C_x^n \left(pe^t\right)^x (1-p)^{n-x}=\left(pe^t+1-p\right)^n$$

由動差母函數求一、二級動差的過程如下

$$m'(t)=\frac{d}{dt}m(t)=\frac{d}{dt}\left(pe^t+1-p\right)^n=n\left(pe^t+1-p\right)^{n-1}pe^t$$
$$m''(t)=\frac{d^2}{dt^2}m(t)=n(n-1)\left(pe^t+1-p\right)^{n-2}\left(pe^t\right)^2+n\left(pe^t+1-p\right)^{n-1}pe^t$$
$$\mu_1'=m'(0)=n(p+1-p)p=np$$
$$\mu_1''=m''(0)=n(n-1)(p+1-p)p^2+n(p+1-p)p=n(n-1)p^2+np$$

【例題 14】請以動差法求常態分配參數μ、σ的估計量。

【解】

令$X_1,\ldots,X_n$為參數μ、σ^2之常態分配的一組樣本，則

$$\begin{cases}\mu_1'=\mu=E(X)=\dfrac{x_1+\cdots+x_n}{n}\\[2ex]\mu_2'=\sigma^2+\mu^2=E\left(X^2\right)=\dfrac{x_1^2+\cdots+x_n^2}{n}\end{cases}$$
$$\Rightarrow\begin{cases}\hat{\mu}=\dfrac{x_1+\cdots+x_n}{n}=\overline{x}\\[2ex]\hat{\sigma}^2=\dfrac{x_1^2+\cdots+x_n^2}{n}-\mu^2=\dfrac{x_1^2+\cdots+x_n^2}{n}-\overline{x}^2=\dfrac{\left(x_1-\overline{x}\right)^2+\cdots+\left(x_n-\overline{x}\right)^2}{n}\end{cases}$$

故μ、σ^2之估計量分別為$\hat{\mu}=\overline{x}=\frac{1}{n}\Sigma x_i$、$\hat{\sigma}^2=\frac{1}{n}\Sigma\left(x_i-\overline{x}\right)^2$。

【例題 15】試求參數λ卜瓦松分配之動差母函數。

【解】

$$\sum_{x=0}^{\infty}\lambda^x e^{-\lambda}\big/x!=e^{-\lambda}\sum_{x=0}^{\infty}\lambda^x\big/x!=e^{-\lambda}e^{\lambda}=1$$
$$m(t)=E\left(e^{tx}\right)=\sum_{x=0}^{\infty}e^{tx}\lambda^x e^{-\lambda}\big/x!=e^{-\lambda}\sum_{x=0}^{\infty}\left(e^t\lambda\right)^x\big/x!=e^{-\lambda}e^{e^t\lambda}$$

【例題 16】試求參數 p 幾何分配之動差母函數。

【解】

幾何分配的機率函數寫成 $f(x)=p(1-p)^x,\ x=0,1,\ldots$，其中參數 p 爲成功機率，隨機變數 x 爲第一次成功之前的失敗次數。

$$\sum_{x=0}^{\infty} p(1-p)^x = p\sum_{x=0}^{\infty}(1-p)^x = p\times p^{-1} = 1$$

$$m(t)=\sum_{x=0}^{\infty} e^{tx} p(1-p)^x = p\sum_{x=0}^{\infty}\left[e^t(1-p)\right]^x = p\left[1-e^t(1-p)\right]^{-1}$$

【例題 17】試求負二項分配之動差母函數。

【解】

負二項分配，$f(x)=C_x^{r+x-1}p^r(1-p)^x,\ x=0,1,\ldots$，爲成功機率 p 的試驗，第 r 次成功前失敗 x 數。

$$\sum_{x=0}^{\infty} q^x = (1-q)^{-1}$$

$$\Rightarrow \begin{cases} \dfrac{d}{dq^{n-1}}\left(\displaystyle\sum_{x=0}^{\infty} q^x\right) = \displaystyle\sum_{x=n-1}^{\infty} \overbrace{x(x-1)\cdots}^{n-1\text{ 項}} q^{x-n+1} = (n-1)!\sum_{x=n-1}^{\infty} C_{n-1}^{x} q^{x-n+1} = (n-1)!\sum_{y=0}^{\infty} C_{n-1}^{y+n-1} q^y \\ \dfrac{d}{dq^{n-1}}(1-q)^{-1} = (n-1)!(1-q)^{-n} \end{cases}$$

$$\Rightarrow \sum_{y=0}^{\infty} C_{n-1}^{y+n-1} q^y = \sum_{y=0}^{\infty} C_{y}^{y+n-1} q^y = (1-q)^{-n}$$

$$\sum_{x=0}^{\infty} C_x^{r+x-1} p^r(1-p)^x = p^r\sum_{x=0}^{\infty} C_x^{r+x-1}(1-p)^x = p^r p^{-r} = 1$$

$$m(t)=\sum_{x=0}^{\infty} e^{tx} C_x^{r+x-1} p^r(1-p)^x = p^r\sum_{x=0}^{\infty} C_x^{r+x-1}\left[e^t(1-p)\right]^x = p^r\left[1-e^t(1-p)\right]^{-r}$$

第六章　假設檢定

2006 年 8 月 21 日 最後修改

6.1 假設檢定概論
6.2 檢定統計量
6.3 假設檢定的形式與步驟
6.4 單一樣本之假設檢定
6.5 兩組樣本之假設檢定
6.6 型 I 錯誤與型 II 錯誤
6.7 檢定力函數與作業曲線
6.8 相關係數的檢定

6.1　假設檢定概論

假設（*Hypothesis*）：一個對母體參數可判定真實與否的陳述

假設檢定（*Hypothesis Testing*）：以樣本檢測對母體參數之陳述是否真實的操作程序

虛無假設（*Null Hypothesis*）：用來檢定的陳述（寫成 H_0）

對立假設（*Alternate Hypothesis*）：虛無假設的否定陳述（寫成 H_1）

範例 6.1　虛無假設與對立假設

以下是有關母體平均數虛無假設：

$$H_0 : \mu \le 4$$

數學三一律告訴我們，對立假設應為

$$H_1 : \mu > 4$$

另外的虛無、對立假設如 $H_0 : \mu = 8 \Leftrightarrow H_1 : \mu \ne 8$ 、 $H_0 : \mu \ge 8 \Leftrightarrow H_1 : \mu < 8$ 。 ■

假設檢定的邏輯

虛無假設很難證明其為**真**，但只要有一個反證就可證明其為**偽**。

不可能樣本、拒絕區域

在虛無假設的條件下，出現**不可能出現的樣本**即可證明虛無假設爲僞。

不可能樣本：出現機率爲零的樣本

拒絕區域（*Reject Region*）：不可能樣本的區域（範圍）

顯著水準

機率小於某個臨界值就會被當成零，該臨界值稱爲顯著水準

顯著水準（*Level of Significance*）：顯然不是零的臨界值

顯著水準是一個機率值，習慣上以α來表示，如$\alpha = 0.05$

找到不可能樣本，然後推翻虛無假設才有意義，反之則沒有意義。
在統計裡，我們只有一次推翻虛無假設的機會。
另外，我們需要知道這個樣本出現的機率，如此才可以判斷其是否爲不可能樣本。
其者需要**抽樣分配**的知識，後者需要會計算**拒絕區域**（已知分配下給機率求臨界值）。

範例 6.2　（假設檢定的邏輯）

張三宣稱袋中的 100 個球都是白色的（**虛無假設**），若從袋中抽出紅球來（**不可能樣本**），則可證明張三的宣稱爲僞。反過來，即使抽出來的是白球（合理樣本），我們也不能證實『袋中都是白球』的虛無假設爲真（接受對立假設沒有意義）；要證實該陳述爲真的唯一方法，是把所有球都拿出來檢查（**支持虛無假設很困難**）。

抽出一個紅球後，張三改口說『袋中的 100 個球中，除了三個紅球之外，都是白色的』（**虛無假設**）。若第二次又抽出紅球來，這時有人會跳出來指責張三撒謊，因爲**不可能**連續抽出兩個紅球（出現**不可能樣本**）。事實上，連續兩次抽出紅球的機率大約 0.0006，並不是不可能，只是出面指責的人認爲 0.0006 與 0 沒有差異（0.0006 已經小於他的**顯著水準**）。不可否認的，這時還是有些神經比較大條的人會認爲，連續出現兩次紅球的機率是小一點，但是也不是不可能呀（他們的**顯著水準**比 0.0006 小）。如果第三次又抽出紅球呢？剛剛還耐得住性子的人也該翻臉了，因爲連續三次紅球的機率只有 0.000006。 ■

虛無假設與對立假設

統計虛無假設的形式如下

$$H_0: \text{母體參數} \begin{matrix} \geq \\ = \\ \leq \end{matrix} \text{常數}$$

其中 母體參數$\in \left\{\mu, p, \sigma^2, \quad \mu_x - \mu_y, p_x - p_y, \sigma_x^2 / \sigma_y^2\right\}$

範例 6.3 （常見之虛無假設與對立假設）

以下是常見的虛無假設與對立假設：

母體參數	右尾檢定	雙尾檢定	左尾檢定	兩組樣本
平均數	$H_0: \mu \leq a$ $H_1: \mu > a$	$H_0: \mu = a$ $H_1: \mu \neq a$	$H_0: \mu \geq a$ $H_1: \mu < a$	$H_0: \mu_1 - \mu_2 = 0$ $H_1: \mu_1 - \mu_2 \neq 0$
比例	$H_0: p \leq a$ $H_1: p > a$	$H_0: p = a$ $H_1: p \neq a$	$H_0: p \geq a$ $H_1: p < a$	$H_0: p_1 - p_2 = 0$ $H_1: p_1 - p_2 \neq 0$
變異數	$H_0: \sigma^2 \leq a$ $H_1: \sigma^2 > a$	$H_0: \sigma^2 = a$ $H_1: \sigma^2 \neq a$	$H_0: \sigma^2 \geq a$ $H_1: \sigma^2 < a$	$H_0: \sigma_1^2 / \sigma_2^2 = 1$ $H_1: \sigma_1^2 / \sigma_2^2 \neq 1$

其中 a 為一常數 ■

寫出虛無假設

假設檢定的兩種結果：

(1)出現不可能樣本→拒絕虛無假設→證實虛無假設的陳述為偽→接受對立假設

(2)沒有出現不可能樣本→無法拒絕虛無假設→無法證實虛無假設的陳述為偽

有學者這樣做，但我不喜歡（因為在我的認知，以下的推論是錯誤的）：

無法拒絕虛無假設 → 接受虛無假設 → 接受虛無假設的陳述為真

就假設檢定的邏輯，出現拒絕虛無假設的結果才有意義。

撰寫虛無假設的三個考量點：

(1)將被拒絕後（證實其為偽），關係人會採取行動的陳述列為虛無假設

(2)將被拒絕後（證實其為偽），後果比較嚴重的陳述列為虛無假設

(3)將看起來不正確的陳述列為虛無假設

範例 6.4 （寫出虛無假設與對立假設）

檢定某罐裝飲料的裝填量是否爲設定的 250cc，

(a)若關係人是顧客，則虛無假設 $H_0: \mu \geq 250$ ，拒絕虛無假設後顧客會有行動；

(b)若關係人是老闆，則虛無假設 $H_0: \mu \leq 250$ ，拒絕虛無假設後老闆會有行動；

(c)若關係人是廠長，則虛無假設 $H_0: \mu = 250$ ，拒絕虛無假設後廠長會有行動。

在法院裡，虛無假設是『被告無罪』，該假設證實爲僞的結果是被告需坐牢，反之，則只是原告不爽快而已。若訴訟雙方，一爲高官一爲平民，虛無假設是高官有理，道理同上，高官輸的結果比較嚴重（對法官、對社會都很嚴重）。

某罐裝飲料應該裝 250cc，初步非正式調查，平均裝填量不到 250cc，則虛無假設應該寫成 $H_0: \mu \geq 250$ 。 ■

6.2 檢定統計量

檢定統計量、樣本檢定統計量值

檢定統計量（*test statistic*）：用來檢定虛無假設的抽樣分配

常用的檢定統計量：z 分配、t 分配、 χ^2 分配、F 分配

樣本檢定統計量值：檢定統計量帶入樣本所求出的函數值

範例 6.5 （樣本檢定統計量值）

檢定 $H_0: \mu \leq \mu^*$ 常用的**檢定統計量**爲

$$t = \frac{\overline{x} - \mu^*}{\sigma / \sqrt{n}}$$

其中，μ^* 、σ 分別爲母體平均數與母體標準差。若已知 $\mu^* = 5$ 、 $\sigma = 2$ ，且樣本爲

$$\{x_1, x_2, x_3, x_4\} = \{5, 8, 4, 7\}$$

則**樣本檢定統計量值**爲

$$t = \frac{6-5}{2/\sqrt{4}} = 1 \quad \left(\overline{x} = \frac{5+8+4+7}{4} = 6\right)$$

■

範例 6.6 （檢定母體參數與檢定統計量的關係）

常見之檢定母體參數有三個： μ 、p、與 σ^2

常用的檢定統計量：z 分配、t 分配、 χ^2 分配、F 分配

這兩者的關係如下：

母體參數		檢定統計量		
平均數	μ	$z=\dfrac{\bar{X}-\mu}{\sigma/\sqrt{n}}$	或	$t_{df=n-1}=\dfrac{\bar{X}-\mu}{s/\sqrt{n}}$
比例	p	$z=\dfrac{\bar{p}-p}{\sqrt{p(1-p)}/\sqrt{n}}$		
變異數	σ^2	$\chi^2_{df=n-1}=\dfrac{(n-1)s^2}{\sigma^2}$		$F_{df=(n_1-1,n_2-1)}=\dfrac{s_1^2}{s_2^2}\dfrac{\sigma_2^2}{\sigma_1^2}$

■

範例 6.7　（虛無假設與拒絕區域的關係）

下表是虛無假設型態與拒絕區域（不可能樣本區域）型態之關係：

虛無假設	不可能樣本	檢定型態	拒絕區域型態
$H_0:\theta\le a$	在右	右尾檢定	$\{\bar{\theta}>u\}$
$H_0:\theta=a$	在兩端	雙尾檢定	$\{\bar{\theta}<\ell$ 或 $\bar{\theta}>u\}$
$H_0:\theta\ge a$	在左	左尾檢定	$\{\bar{\theta}<\ell\}$

其中θ爲母體參數、$\bar{\theta}$爲檢定統計量值、ℓ爲下臨界值、u爲上臨界值。 ■

顯著水準、拒絕區域、與 *p* 值

顯著水準與拒絕區域大小成正向關係：

顯著水準小，則拒絕區域也較小

基與保守原則，拒絕區域習慣不包含臨界值，如$R=\left\{\chi^*>9.21\right\}$或$R=\{z<-1.96\}$。

範例 6.8　（給顯著水準求拒絕區域）

若檢定統計量爲 z 分配，右尾檢定，已知顯著水準$\alpha=0.05$，則拒絕區域爲

$R=\{z>1.645\}$

又若雙尾檢定，已知顯著水準$\alpha=0.05$，則拒絕區域爲

$R=\{z<-1.96$ 或 $z>1.96\}$

■

p 值（p *value*）：不可能樣本區間（拒絕區域）的機率

決策法則（*Decision Rule*）：若檢定統計量值在拒絕區域內則拒絕 H_0，否則無法拒絕 H_0。

範例 6.9　（給拒絕區域求 *p* 值）

若檢定統計量爲 z 分配，右尾檢定，已知拒絕區域 $R=\{z>2.58\}$，則 p 值爲

$$p(z>2.58)=0.01$$

又若雙尾檢定，已知拒絕區域 $R=\{z<-1.645 \text{ 或 } z>1.645\}$，則 p 值爲

$$p(z<-1.645 \text{ 或 } z>1.645)=0.1$$

■

6.3　假設檢定的形式與步驟

三種檢定的形式：z 值法、p 值法、與信賴區間法

z 值法

假設檢定五步驟（***z*** **值法**）：

步驟一：寫出虛無假設（與對立假設）

步驟二：確定檢定統計量（z、t、χ^2、或 F）

步驟三：以顯著水準、檢定型式、與檢定統計量，求出拒絕區域

步驟四：計算樣本檢定統計量值、作假設檢定決策（若在拒絕區域則拒絕 H_0）

步驟五：寫假設檢定報告

範例 6.10　（步驟一：寫出虛無假設）

寫出虛無假設有兩個步驟：(1)決定母體參數；(2)決定左、右、或雙尾檢定。

第一個步驟不會有問題，第二個步驟有以下原則：

(1)拒絕虛無假設後必須有所行動；(2)拒絕虛無假設才有意義。

拒絕虛無假設後必須有所行動　　假設研究對象是罐裝飲的料裝填量是否正常。對顧客而言，裝填量太少就會有抗議行動，因此，不可能樣本在左端，應爲左尾檢定。對工程師而言，裝填量太多或太少都顯示機器設定有問題，必須檢修機器，因此應爲雙尾檢定。對老闆而言，裝填量太多會增加成本，因此應爲右尾檢定。

拒絕虛無假設才有意義　我們希望虛無假設容易被推翻，因此，如果樣本檢定統計量值偏高，則為右尾檢定；反之若統計量值偏低則為左尾檢定。 ■

範例 6.11　**（步驟一：寫出虛無假設）**

(1)平均減肥量為 10 磅，樣本平均數為 9 磅 ⇒ 左尾檢定。

(2)平均睡眠 7 小時，樣本平均睡眠時間為 6.8 小時 ⇒ 左尾檢定。

(3)Has the special additive **increased** the mean weight of the chickens? ⇒ 右尾檢定。

(4)Is there a **change** in the mean length of the bars? ⇒ 雙尾檢定。 ■

範例 6.12　**（母體平均數檢定，*z* 值法）**

某罐裝咖啡標示其咖啡因含量少於 20cc，今隨機抽取 9 罐此品牌咖啡作檢查，發現其平均咖啡因的含量為 26cc，標準差 8cc，請利用顯著水準$\alpha = 0.05$來檢定其標示是否為真？

【解】

（基本資料為 $\mu = 20, n = 9, \bar{x} = 26, s = 8, \alpha = 0.05$ ）

(1)虛無假設為$H_0 : \mu \le 20$ （右尾檢定）

(2)檢定統計量 $t = \dfrac{\bar{x} - \mu}{s/\sqrt{n}}$ 為自由度$n-1=8$ 的 t 分配

(3)右尾檢定、自由度 $n-1=8$ 的 t 分配、 $\alpha = 0.05$ ，拒絕區域為$R = \{t > 1.860\}$

(4)樣本檢定統計量為$\dfrac{26-20}{8/\sqrt{9}} = \dfrac{18}{8} = 2.25 \in R$

(5)拒絕虛無假設，有充分證據證實咖啡因含量高於 20cc。 ■

p 值法

假設檢定五步驟（*p* 值法）：

步驟一：寫出虛無假設（與對立假設）

步驟二：確定檢定統計量

步驟三：計算樣本檢定統計量值，寫出假定之**拒絕區域**

步驟四：以**拒絕區域**、檢定型式、與檢定統計量，求出 *p* 值

步驟五：寫假設檢定報告

範例 6.13　**（母體平均數檢定，*p* 值法）**

某罐裝咖啡標示其咖啡因含量少於 20cc，今隨機抽取 9 罐此品牌咖啡作檢查，發現其平均咖啡因的含量爲 26cc，標準差 8cc，請利用顯著水準$\alpha = 0.05$來檢定其標示是否爲真？

【解】

（基本資料爲 $\mu = 20, n = 9, \bar{x} = 26, s = 8, \alpha = 0.05$ ）

(1)虛無假設爲$H_0 : \mu \le 20$ （右尾檢定）

(2)檢定統計量 $t = \dfrac{\bar{x} - \mu}{s/\sqrt{n}}$ 爲自由度$n-1=8$ 的 t 分配

(3)樣本檢定統計量爲$\dfrac{26-20}{8/\sqrt{9}} = \dfrac{18}{8} = 2.25$

(4)右尾檢定、自由度 $n-1=8$ 的 t 分配、臨界值爲 2.25，求得 $p = 0.0273$

(5)若顯著水準高於 0.0273 則應該拒絕H_0，否則應接受H_1。 ■

z 值法與 *p* 值法的關係

z 值法：已有決策者的顯著水準，求不可能樣本區間（給機率求區間）

p 值法：沒有決策者的顯著水準，假定不可能樣本區間後求 *p* 值（給區間求機率）

p 值法中假定以**樣本檢定統計量值**爲臨界值的區間爲不可能樣本區間

信賴區間法

假設檢定五步驟（**信賴區間法**）：

步驟一：寫出虛無假設（與對立假設）

步驟二：確定檢定統計量（z、t、 χ^2 、或F）

步驟三：以顯著水準、檢定型式、與檢定統計量，求出**信賴區間**

步驟四：計算樣本檢定統計量值、作假設檢定決策（若在信賴區間則無法拒絕H_0）

步驟五：寫假設檢定報告

一般信賴區間法常用於雙尾檢定，其他形式的檢定比較少見，但也不是不可以。畢竟，將信賴區間視爲拒絕區域的補集合，則一切是那麼自然。

範例 6.14 （母體平均數檢定，信賴區間法）

某罐裝咖啡標示其咖啡因含量少於 20cc，今隨機抽取 9 罐此品牌咖啡作檢查，發現其平均咖啡因的含量爲 26cc，標準差 8cc，請利用信賴區間法（信賴度$1-\alpha = 95\%$ ）來檢定其標示是否爲真？

【解】

（基本資料爲 $\mu=20, n=9, \overline{x}=26, s=8, \alpha=0.05$ ）

(1)虛無假設爲 $H_0:\mu\le 20$ （右尾檢定）

(2)檢定統計量 $t=\dfrac{\overline{x}-\mu}{s/\sqrt{n}}$ 爲自由度 $n-1=8$ 的 t 分配

(3)右尾檢定、自由度 $n-1=8$ 的 t 分配、$1-\alpha=95\%$，信賴區間 $CI_t=\{t\le 1.860\}$

或者，$CI_{\overline{x}}=\left\{\overline{x}\le 20+1.860\times\dfrac{8}{\sqrt{9}}=24.96\right\}$、$CI_\mu=\left\{\mu\ge 26-1.860\times\dfrac{8}{\sqrt{9}}=21.04\right\}$

(4)樣本檢定統計量爲 $t=\dfrac{26-20}{8/\sqrt{9}}=\dfrac{18}{8}=2.25\notin CI_t$ ，拒絕虛無假設

或者，$\overline{x}=26\notin CI_{\overline{x}}$ 、 $\mu=20\notin CI_\mu$ ，拒絕虛無假設

(5)有充分證據證實咖啡因含量高於 20cc。 ■

6.4 單一樣本之假設檢定

單一樣本與平均有關（設相關統計量 Y）之檢定統計量有下列兩組：

$$z=\frac{Y-\mu_Y}{\sigma_Y},\quad t=\frac{Y-\mu_Y}{s_Y}$$

(a) $Y=\overline{x}$

$$\mu_{\overline{x}}=\mu,\quad \sigma_{\overline{x}}=\frac{\sigma}{\sqrt{n}}\quad\Rightarrow\quad z=\frac{\overline{x}-\mu}{\sigma/\sqrt{n}}$$

$$\mu_{\overline{x}}=\mu,\quad s_{\overline{x}}=\frac{s}{\sqrt{n}}\quad\Rightarrow\quad t=\frac{\overline{x}-\mu}{s/\sqrt{n}}$$

(b) $Y=\overline{p}$

$$\mu_{\overline{p}}=p,\quad \sigma_{\overline{p}}=\frac{\sqrt{p(1-p)}}{\sqrt{n}}\quad\Rightarrow\quad z=\frac{\overline{p}-p}{\sqrt{p(1-p)}/\sqrt{n}}$$

單一樣本與變異數有關之檢定統計量只有下列一組：

$$\chi^2=\frac{(n-1)s^2}{\sigma^2}$$

範例 6.15　（母體比例檢定）

某系宣稱有 5%的畢業生考上研究所，今隨機抽問 50 位當年度畢業生，其中有 1 人考上研究所，請作統計推論。

【解】

（基本資料爲 $p = 0.05, n = 50, \overline{p} = 1/50 = 0.02$ ）

(1)虛無假設爲 $H_0 : p \ge 0.05$ （左尾檢定）

(2)檢定統計量 $z = \dfrac{\overline{p} - p}{\sqrt{p(1-p)} \big/ \sqrt{n}}$ 爲 z 分配

(3)樣本檢定統計量爲 $z = \dfrac{0.02 - 0.05}{\sqrt{0.05 \times (1 - 0.05) / 50}} = -0.9740$

(4)左尾檢定、z 分配、臨界值爲 -0.9740 ，求得 $p = 0.1650$

(5)若顯著水準高於 0.1650 則應該拒絕 H_0 ，否則應接受 H_1 。 ■

範例 6.16　（母體變異數檢定）

A company claims that the standard deviation in their delivery time is less than 5 days. A sample of 27 past customers is taken. The average delivery time in the sample was 14 days with a standard deviation of 4.5 days. At 95% confidence, test the company's claim.

【解】

(z 值法； 基本資料：σ^2=25、雙尾、n=27、s^2=20.25、α=0.05)

(1) 虛無假設 Ho：$\sigma^2 = 25$（雙尾檢定）

(2) 檢定統計量 $(n-1)s^2/\sigma^2$ 爲 自由度 26 的 $\chi 2$ 分配

(3) 自由度 26 的 $\chi 2$ 分配，雙尾，α=0.05，求得拒絕區域
R = { $\chi 2 < 13.8439$ 或 $\chi 2 > 41.9232$ }

(4) 樣本檢定統計量值 $\chi 2 = 26 * 20.25 / 25 = 21.06$ 不屬於 拒絕區域 R

(5) 沒有充分理由來拒絕虛無假設 Ho

■

6.5　兩組樣本之假設檢定

兩組樣本與平均有關（設相關統計量 Y）之檢定統計量有下列兩組：

$$z = \frac{Y - \mu_Y}{\sigma_Y}, \quad t = \frac{Y - \mu_Y}{s_Y}$$

(a) $Y=\overline{x}_1-\overline{x}_2$

$$\mu_{\overline{x}_1-\overline{x}_2}=\mu_1-\mu_2,\quad \sigma_{\overline{x}_1-\overline{x}_2}=\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}}\quad\Rightarrow\quad z=\frac{(\overline{x}_1-\overline{x}_2)-(\mu_1-\mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}}}$$

（σ_1、σ_2 未知，且 $\sigma_1\neq\sigma_2$）

$$s_{\overline{x}_1-\overline{x}_2}=\sqrt{\frac{s_1^2}{n_1}+\frac{s_2^2}{n_2}}\quad\Rightarrow\quad t=\frac{(\overline{x}_1-\overline{x}_2)-(\mu_1-\mu_2)}{\sqrt{\frac{s_1^2}{n_1}+\frac{s_2^2}{n_2}}}$$

其中，自由度 df ： $\dfrac{\left(s_1^2/n_1+s_2^2/n_2\right)^2}{df}=\dfrac{\left(s_1^2/n_1\right)^2}{n_1-1}+\dfrac{\left(s_2^2/n_2\right)^2}{n_2-1}$

（σ_1、σ_2 未知，且 $\sigma_1=\sigma_2$）

$$s_{\overline{x}_1-\overline{x}_2}=\sqrt{\frac{s_p^2}{n_1}+\frac{s_p^2}{n_2}}\quad\Rightarrow\quad t=\frac{(\overline{x}_1-\overline{x}_2)-(\mu_1-\mu_2)}{\sqrt{\frac{s_p^2}{n_1}+\frac{s_p^2}{n_2}}}$$

其中聯合估計之樣本變異數爲 $s_P^2=\dfrac{(n_1-1)s_1^2+(n_2-1)s_2^2}{(n_1-1)+(n_2-1)}$

(b) $Y=\overline{p}_1-\overline{p}_2$

$$\sigma_{\overline{p}_1-\overline{p}_2}=\sqrt{\frac{p_1(1-p_1)}{n_1}+\frac{p_2(1-p_2)}{n_2}}\quad\Rightarrow\quad z=\frac{(\overline{p}_1-\overline{p}_2)-(p_1-p_2)}{\sqrt{\frac{p_1(1-p_1)}{n_1}+\frac{p_2(1-p_2)}{n_2}}}$$

（若 p_1、p_2 未知，$p_1\neq p_2$）

$$s_{\overline{p}_1-\overline{p}_2}=\sqrt{\frac{\overline{p}_1(1-\overline{p}_1)}{n_1}+\frac{\overline{p}_2(1-\overline{p}_2)}{n_2}}\quad\Rightarrow\quad t=\frac{(\overline{p}_1-\overline{p}_2)-(p_1-p_2)}{\sqrt{\frac{\overline{p}_1(1-\overline{p}_1)}{n_1}+\frac{\overline{p}_2(1-\overline{p}_2)}{n_2}}}$$

（若 p_1、p_2 未知，$p_1=p_2$）

$$s_{\overline{p}_1-\overline{p}_2}=\sqrt{\frac{\overline{p}_c(1-\overline{p}_c)}{n_1}+\frac{\overline{p}_c(1-\overline{p}_c)}{n_2}}\quad\Rightarrow\quad t=\frac{(\overline{p}_1-\overline{p}_2)-(p_1-p_2)}{\sqrt{\frac{\overline{p}_c(1-\overline{p}_c)}{n_1}+\frac{\overline{p}_c(1-\overline{p}_c)}{n_2}}}$$

其中聯合估計之樣本比例爲 $\overline{p}_c=\dfrac{n_1\overline{p}_1+n_2\overline{p}_2}{n_1+n_2}$

一般 n_1、n_2 都很大，不需要查 t 分配表（以 z 分配表代替）。

兩組樣本與變異數有關之檢定統計量只有下列一組：

$$F=\frac{\frac{s_1^2}{\sigma_1^2}}{\frac{s_2^2}{\sigma_2^2}}=\frac{s_1^2}{s_2^2}\frac{\sigma_2^2}{\sigma_1^2}$$

範例 6.17　（平均數差之檢定）

某保養品工廠生產某種神奇美容乳液，生產線設定每瓶裝填量爲 15cc。今該工廠新設立一條生產線，爲了驗證新生產線的效能，分別對兩生產線作隨機取樣檢查，結果如下：

舊生產線	新生產線
$n_1=64$	$n_2=36$
$\overline{x}_1=14.36$	$\overline{x}_2=15.02$
$s_1=2.4$	$s_2=2.5$

請問兩生產線的裝填量是否有差異。

【解】

（基本資料爲 $n_1=64,\overline{x}_1=14.36,s_1=2.4,\ \ n_2=36,\overline{x}_2=15.02,s_2=2.5,\ \ \alpha=0.05$ ）

（兩獨立樣本，母體變異數未知，且不知其是否相等）

(1)虛無假設爲 $H_0:\mu_1=\mu_2$ 或 $H_0:\mu_1-\mu_2=0$ （雙尾檢定）

(2)檢定統計量 $t=\dfrac{\left(\overline{X}_1-\overline{X}_2\right)-\mu_{\overline{X}_1-\overline{X}_2}}{s_{\overline{X}_1-\overline{X}_2}}=\dfrac{\overline{X}_1-\overline{X}_2}{\sqrt{s_1^2/n_1+s_2^2/n_2}}$ 爲自由度 $df=98$ 的 t 分配

其中，$df=\left\lfloor\dfrac{\left(s_1^2/n_1+s_2^2/n_2\right)^2}{\left(s_1^2/n_1\right)^2/(n_1-1)+\left(s_2^2/n_2\right)^2/(n_2-1)}\right\rfloor\approx 98$

(3)自由度 98 的 t 分配、雙尾檢定、$\alpha=0.05$，拒絕區域 $R=\{t<-1.984,\ t>1.984\}$

（若以大樣本看待，查 z 分配表，則拒絕區域 $R=\{z<-1.96,\ z>1.96\}$ ）

(4)樣本檢定統計量值 $\dfrac{14.36-15.02}{\sqrt{2.4^2/64+2.5^2/36}}=-1.2856\notin R$

(5)無法拒絕虛無假設：『沒有充分的證據懷疑，兩生產線的裝填量設定不同』。 ■

範例 6.18　（平均數差之檢定，母體變異數相等）

某保養品工廠生產某種神奇美容乳液，該工廠每天定期作品檢。以下是前後兩天品檢取樣檢查的結果：

第一天	第二天
$n_1 = 64$	$n_2 = 36$
$\overline{x_1} = 14.36$	$\overline{x_2} = 15.02$
$s_1 = 2.4$	$s_2 = 2.5$

請問這兩天生產線的裝填量是否有差異？

【解】

（基本資料爲 $n_1 = 64, \bar{x}_1 = 14.36, s_1 = 2.4,\ \ n_2 = 36, \bar{x}_2 = 15.02, s_2 = 2.5,\ \ \alpha = 0.05$ ）

（兩獨立樣本，母體變異數未知，相同生產線，故假設母體變異數相等）

(1)虛無假設爲 $H_0 : \mu_1 = \mu_2$ 或 $H_0 : \mu_1 - \mu_2 = 0$ （雙尾檢定）

(2)檢定統計量 $t = \dfrac{\left(\bar{X}_1 - \bar{X}_2\right) - \mu_{\bar{X}_1 - \bar{X}_2}}{s_{\bar{X}_1 - \bar{X}_2}} = \dfrac{\bar{X}_1 - \bar{X}_2}{s_P \sqrt{1/n_1 + 1/n_2}}$ 爲自由度 $df = 98$ 的 t 分配

其中，$df = n_1 + n_2 - 2 = 98,\ \ s_P = \sqrt{\dfrac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}}$

(3)自由度 98 的 t 分配、雙尾檢定、 $\alpha = 0.05$ ，拒絕區域 $R = \{t < -1.984,\ t > 1.984\}$

(4)樣本檢定統計量值 $\dfrac{14.36 - 15.02}{\sqrt{\dfrac{63 \times 2.4^2 + 35 \times 2.5^2}{64 + 36 - 2}} \sqrt{\left(\dfrac{1}{64} + \dfrac{1}{36}\right)}} = -1.300 \notin R$

(5)無法拒絕虛無假設：『沒有充分的證據懷疑，兩生產線的裝填量設定不同』。 ■

範例 6.19 （**比例之檢定**）

某兩系分別調查其畢業生考研究所的情況，第一個系隨機抽問 20 人，其中有 3 人考上研究所，第二個系抽問 12 人，其中有 1 人考上研究所，請問這兩者考上研究所學生的比例是否有差異？

【解】

（基本資料爲 $n_1 = 20, \bar{p}_1 = 3/20,\ \ n_2 = 12, \bar{p}_2 = 1/12,\ \ \alpha = 0.05$ ）

（兩獨立樣本，母體變異數未知，由虛無假設得知其相等，）

(1)虛無假設爲 $H_0 : p_1 = p_2$ （雙尾檢定）

(2)檢定統計量 $t = \dfrac{\bar{p}_1 - \bar{p}_2}{\sqrt{\bar{p}_c (1 - \bar{p}_c)} \sqrt{1/n_1 + 1/n_2}}$ 爲自由度 $df = 30$ 的 t 分配

其中，$df = n_1 + n_2 - 2 = 30,\ \ \bar{p}_c = \dfrac{n_1 \bar{p}_1 + n_2 \bar{p}_2}{n_1 + n_2} = \dfrac{3 + 1}{20 + 12} = \dfrac{1}{8}$

(3)自由度 30 的 t 分配、雙尾檢定、 $\alpha = 0.05$ ，拒絕區域 $R = \{t < -2.0423,\ t > 2.0423\}$

(4)樣本檢定統計量值 $\dfrac{\dfrac{3}{20}-\dfrac{1}{12}}{\sqrt{\dfrac{1}{8}\times\dfrac{7}{8}}\sqrt{\dfrac{1}{20}+\dfrac{1}{12}}}=\dfrac{0.0667}{0.1208}=0.5521\notin R$

(5)無法拒絕虛無假設：『沒有充分的證據顯示，兩系的錄取率有差異』。 ■

範例 6.20　（變異數之檢定）

某保養品工廠生產某種神奇美容乳液，生產線設定每瓶裝填量為 15cc。今該工廠新設立一條生產線，為了驗證新生產線的效能，分別對兩生產線作隨機取樣檢查，結果如下：

舊生產線	新生產線
$n_1=64$	$n_2=36$
$\overline{x_1}=14.36$	$\overline{x_2}=15.02$
$s_1=2.4$	$s_2=2.5$

請問這兩天生產線的裝填量之變異數是否有差異？

【解】

（基本資料為 $n_1=64, \overline{x}_1=14.36, s_1=2.4,\ n_2=36, \overline{x}_2=15.02, s_2=2.5,\ \alpha=0.05$）

(1)虛無假設為 $H_0:\sigma_1^2=\sigma_2^2$ 或 $H_0:\sigma_1^2/\sigma_2^2=1$ （雙尾檢定）

(2)檢定統計量 $F=\dfrac{s_2^2}{s_1^2}\dfrac{\sigma_1^2}{\sigma_2^2}=\dfrac{s_2^2}{s_1^2}$ 為自由度 $df=(35,63)$ 的 F 分配

其中，$df=(n_2-1,n_1-1)=(35,63)$

（注意：我們習慣把變異數大的擺分子，因 $s_2^2>s_1^2$，故 s_2^2 在分子。）

(3)自由度 $(35,63)$ 的 F 分配、雙尾檢定、$\alpha=0.05$，

拒絕區域 $R=\{F<0.5393,\ F>1.7637\}$

(4)樣本檢定統計量值 $\dfrac{2.5^2}{2.4^2}=1.085\notin R$

(5)無法拒絕虛無假設：『沒有充分證據懷疑，兩生產線裝填量之變異數不同』。 ■

範例 6.21　（母體變異數比例檢定）

The following information was obtained from two independent random samples representing populations A and B:

	Population A	Population B
Sample Size	18	21
Sample Mean	200	230
Sample Variance	40	90

If you were to test for the equality of the two sample means, would you need to pool the variances? Why or why not? Use a 0.05 level of significance.

【解】

（基本資料 $\sigma_A^2/\sigma_B^2 = 1, n_A = 18, s_A^2 = 40, n_B = 21, s_B^2 = 90, \alpha = 0.05$ ）

(1)虛無假設為 $H_0 : \sigma_A^2/\sigma_B^2 = 1$（雙尾檢定）

(2)檢定統計量 $\dfrac{s_A^2/\sigma_A^2}{s_B^2/\sigma_B^2} = \dfrac{s_A^2}{s_B^2} \times \dfrac{\sigma_B^2}{\sigma_A^2}$ 為自由度 $(n_A - 1, n_B - 1) = (17, 20)$ 的 F 分配

(3)雙尾檢定、自由度 $(17,20)$ 的 F 分配、$\alpha = 0.05$，拒絕區域為

$R = \{F < 0.38$ 或 $F > 2.52\}$

(4)樣本檢定統計量為 $\dfrac{40}{90} \times 1 = 0.4444 \notin R$

(5)無法拒絕虛無假設，兩母體變異數應視為相等，因此需聯合估計樣本變異數。 ■

6.6 型 I 錯誤與型 II 錯誤

型 I 錯誤（*Type I Error*）：拒絕 H_0 所產生的錯誤

型 II 錯誤（*Type II Error*）：沒有拒絕 H_0（接受 H_0 ）所產生的錯誤

型 I 錯誤的機率：α （ H_0 為真時才有型 I 錯誤）

型 II 錯誤的機率：β （ H_0 為偽時才有型 II 錯誤）

令 $P(\bullet)$ 是**真實分配的機率函數**，則 $\beta = 1 - P($拒絕區域$)$

計算型 II 錯誤 β 值的步驟

步驟一：寫出虛無假設（與對立假設）

步驟二：確定檢定統計量（z、t、χ^2 、或 F）

步驟三：以顯著水準、檢定型式、與檢定統計量，求出**拒絕區域**

步驟四：以新檢定統計量（新母體參數）轉換拒絕區域之臨界值

步驟五：以**相反檢定型式**、新拒絕區域、求出**機率**（ β **值**）

其中

前三步驟與假設檢定完全相同；

相反檢定型式：右尾→左尾、左尾→右尾、雙尾→信賴區間；

檢定統計量：涉及兩個（母體參數不同）檢定統計量。

臨界值的轉換公式

$$\begin{cases} z = \dfrac{Y^* - \mu_Y}{\sigma_Y} \\ z' = \dfrac{Y^* - \mu'_Y}{\sigma_Y} \end{cases} \Rightarrow \quad z' = \frac{(\mu_Y + z \times \sigma_Y) - \mu'_Y}{\sigma_Y} = z + \frac{\mu_Y - \mu'_Y}{\sigma_Y}$$

或

$$\begin{cases} t = \dfrac{Y^* - \mu_Y}{s_Y} \\ t' = \dfrac{Y^* - \mu'_Y}{s_Y} \end{cases} \Rightarrow \quad t' = \frac{(\mu_Y + t \times s_Y) - \mu'_Y}{s_Y} = t + \frac{\mu_Y - \mu'_Y}{s_Y}$$

其中，μ'_Y 、z' 分別爲新母體參數與新臨界值。最常見到的是

$$z' = z + \frac{\mu - \mu'}{\sigma/\sqrt{n}}$$

範例 6.22　（計算型 II 錯誤）

某罐裝咖啡標示其咖啡因含量少於 20cc，今隨機抽取 9 罐此品牌咖啡作檢查，發現其平均咖啡因的含量爲 26cc，標準差 8cc，在用顯著水準 $\alpha = 0.05$ 來檢定的場合，若真正的咖啡因含量爲 30cc 下的 β 值。

【解】

（基本資料爲 $\mu = 20, n = 9, \overline{x} = 26, s = 8, \alpha = 0.05, \mu' = 30$ ）

(1)虛無假設爲 $H_0 : \mu \le 20$ （右尾檢定）

(2)檢定統計量 $t = \dfrac{\overline{x} - \mu}{s/\sqrt{n}}$ 爲自由度 $n-1=8$ 的 t 分配

(3)右尾檢定、自由度 $n-1=8$ 的 t 分配、$\alpha = 0.05$，拒絕區域爲 $R = \{t > 1.860\}$

(4)轉換臨界值 $t' = t + \dfrac{\mu - \mu'}{s/\sqrt{n}} = 1.860 + \dfrac{20-30}{8/\sqrt{9}} = -1.891$

(5)左尾檢定、自由度 $n-1=8$ 的 t 分配、臨界值 $t' = -1.891$，求得 $\beta = 0.9523$。 ■

6.7　檢定力函數與作業曲線

檢定力（*Power of Test*）：不犯型 II 錯誤的機率，$1 - \beta = P(\text{拒絕區域})$

檢定力越高表示 H_0 爲僞時越不會犯錯

檢定力函數（*Power Function*）：表示$1-\beta$與μ'（真實的μ）之關係的函數

作業曲線（*Operation Characteristic Curve*）：表示β與μ'之關係的曲線（函數）

右尾檢定情況下（不可能樣本在右），β與μ'有反向關係

左尾檢定情況下（不可能樣本在左），β與μ'有正向關係

雙尾檢定情況下（不可能樣本在兩端），β與μ'爲中央高（臨界值位置）兩端低

範例 6.23　（β與μ'的關係）

因有以下關係

$$1-\beta=P(R=\text{拒絕區域})\quad\text{或}\quad\beta=1-P(R)$$

得知β與R有反向關係（$1-\beta$與R有反向關係）。

右尾檢定：R在右邊$\Rightarrow R$與μ'有正向關係$\Rightarrow\beta$與μ'有反向關係；

左尾檢定：R在左邊$\Rightarrow R$與μ'有反向關係$\Rightarrow\beta$與μ'有正向關係；

雙尾檢定：R在μ兩邊$\Rightarrow\mu'$離μ越遠R越大$\Rightarrow\beta$中央高兩頭低。 ■

6.8　相關係數的檢定

虛無假設$H_0:\rho=0$時，

$$\text{檢定統計量}\quad\frac{r}{\sqrt{(1-r^2)/(n-2)}}\quad\text{爲自由度}\ df=n-2\ \text{的}\ t\ \text{分配}$$

虛無假設$H_0:\rho=\rho_0$時，

$$\text{檢定統計量}\quad\frac{Z_r-Z_\rho}{1/(n-3)}\quad\text{爲}\ z\ \text{分配（}n>30\text{，大樣本時）}$$

其中，$Z_r=\frac{1}{2}\ln\frac{1+r}{1-r}$，$Z_\rho=\frac{1}{2}\ln\frac{1+\rho_0}{1-\rho_0}$，$Z_r\sim N\left(Z_\rho,\frac{1}{n-3}\right)$

範例 6.24　（相關係數的檢定・92-政大-財管）

4. (20) Two personnel evaluation techniques are available. The first requires a two-hour test interview session, while the second can be completed in less than an hour. A high correlation between test scores would indicate that the second test could replace the two-hour test and hence save time and money. Following is data that gives the scores on Test I (x) and Test II (y) for 15 applicants.

x	75	89	60	71	92	105	55	87	73	77	84	91	75	82	76
y	38	56	35	45	59	70	31	52	48	41	51	58	45	49	47

a. Find the coefficient of correlation, r.

b. Interpret r.

c. Test to see if the correlation between x and y is significant at $\alpha = 0.01$.

【解】

(a)

計算原始資料：

x	75	89	60	71	92	105	55	87	73	77	84	91	75	82	76	1192
y	38	56	35	45	59	70	31	52	48	41	51	58	45	49	47	725
x^2	5625	7921	3600	5041	8464	11025	3025	7569	5329	5929	7056	8281	5625	6724	5776	96990
y^2	1444	3136	1225	2025	3481	4900	961	2704	2304	1681	2601	3364	2025	2401	2209	36461
xy	2850	4984	2100	3195	5428	7350	1705	4524	3504	3157	4284	5278	3375	4018	3572	59324

$$r = \frac{\Sigma xy - \frac{(\Sigma x)(\Sigma y)}{n}}{\sqrt{\Sigma x^2 - \frac{(\Sigma x)^2}{n}}\sqrt{\Sigma y^2 - \frac{(\Sigma y)^2}{n}}} = \frac{59324 - \frac{1192\times 725}{15}}{\sqrt{96990 - \frac{1192^2}{15}}\sqrt{36461 - \frac{725^2}{15}}} = 0.9539$$

(b)

x、y 之間有高度正相關

(c)

基本資料：$n = 15$ 、 $r = 0.9539$

(1)虛無假設 $H_0 : \rho = 0$

(2)檢定統計量 $t = \dfrac{r - \rho}{\sqrt{(1 - r^2)/(n-2)}}$ ，自由度 $n - 2 = 15 - 2 = 13$

(3)雙尾，自由度 13 之 t 分配，$\alpha = 0.01$，求得拒絕區域 $R = \{|t| > 3.0123\}$

(4)樣本檢定統計量 $t^* = \dfrac{0.9539}{\sqrt{(1 - 0.9539^2)/(15-2)}} = 11.4597 \in R$ ，拒絕虛無假設

(5)相關係數顯然不為零 ■

附錄：常用統計表

z 分配**左尾**機率表

z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.5000	0.5040	0.5080	0.5120	0.5160	0.5199	0.5239	0.5279	0.5319	0.5359
0.1	0.5398	0.5438	0.5478	0.5517	0.5557	0.5596	0.5636	0.5675	0.5714	0.5753
0.2	0.5793	0.5832	0.5871	0.5910	0.5948	0.5987	0.6026	0.6064	0.6103	0.6141
0.3	0.6179	0.6217	0.6255	0.6293	0.6331	0.6368	0.6406	0.6443	0.6480	0.6517
0.4	0.6554	0.6591	0.6628	0.6664	0.6700	0.6736	0.6772	0.6808	0.6844	0.6879
0.5	0.6915	0.6950	0.6985	0.7019	0.7054	0.7088	0.7123	0.7157	0.7190	0.7224
0.6	0.7257	0.7291	0.7324	0.7357	0.7389	0.7422	0.7454	0.7486	0.7517	0.7549
0.7	0.7580	0.7611	0.7642	0.7673	0.7704	0.7734	0.7764	0.7794	0.7823	0.7852
0.8	0.7881	0.7910	0.7939	0.7967	0.7995	0.8023	0.8051	0.8078	0.8106	0.8133
0.9	0.8159	0.8186	0.8212	0.8238	0.8264	0.8289	0.8315	0.8340	0.8365	0.8389
1.0	0.8413	0.8438	0.8461	0.8485	0.8508	0.8531	0.8554	0.8577	0.8599	0.8621
1.1	0.8643	0.8665	0.8686	0.8708	0.8729	0.8749	0.8770	0.8790	0.8810	0.8830
1.2	0.8849	0.8869	0.8888	0.8907	0.8925	0.8944	0.8962	0.8980	0.8997	0.9015
1.3	0.9032	0.9049	0.9066	0.9082	0.9099	0.9115	0.9131	0.9147	0.9162	0.9177
1.4	0.9192	0.9207	0.9222	0.9236	0.9251	0.9265	0.9279	0.9292	0.9306	0.9319
1.5	0.9332	0.9345	0.9357	0.9370	0.9382	0.9394	0.9406	0.9418	0.9429	0.9441
1.6	0.9452	0.9463	0.9474	0.9484	0.9495	0.9505	0.9515	0.9525	0.9535	0.9545
1.7	0.9554	0.9564	0.9573	0.9582	0.9591	0.9599	0.9608	0.9616	0.9625	0.9633
1.8	0.9641	0.9649	0.9656	0.9664	0.9671	0.9678	0.9686	0.9693	0.9699	0.9706
1.9	0.9713	0.9719	0.9726	0.9732	0.9738	0.9744	0.9750	0.9756	0.9761	0.9767
2.0	0.9772	0.9778	0.9783	0.9788	0.9793	0.9798	0.9803	0.9808	0.9812	0.9817
2.1	0.9821	0.9826	0.9830	0.9834	0.9838	0.9842	0.9846	0.9850	0.9854	0.9857
2.2	0.9861	0.9864	0.9868	0.9871	0.9875	0.9878	0.9881	0.9884	0.9887	0.9890
2.3	0.9893	0.9896	0.9898	0.9901	0.9904	0.9906	0.9909	0.9911	0.9913	0.9916
2.4	0.9918	0.9920	0.9922	0.9925	0.9927	0.9929	0.9931	0.9932	0.9934	0.9936
2.5	0.9938	0.9940	0.9941	0.9943	0.9945	0.9946	0.9948	0.9949	0.9951	0.9952
2.6	0.9953	0.9955	0.9956	0.9957	0.9959	0.9960	0.9961	0.9962	0.9963	0.9964
2.7	0.9965	0.9966	0.9967	0.9968	0.9969	0.9970	0.9971	0.9972	0.9973	0.9974
2.8	0.9974	0.9975	0.9976	0.9977	0.9977	0.9978	0.9979	0.9979	0.9980	0.9981
2.9	0.9981	0.9982	0.9982	0.9983	0.9984	0.9984	0.9985	0.9985	0.9986	0.9986
3.0	0.9987	0.9987	0.9987	0.9988	0.9988	0.9989	0.9989	0.9989	0.9990	0.9990
3.1	0.9990	0.9991	0.9991	0.9991	0.9992	0.9992	0.9992	0.9992	0.9993	0.9993
3.2	0.9993	0.9993	0.9994	0.9994	0.9994	0.9994	0.9994	0.9995	0.9995	0.9995
3.3	0.9995	0.9995	0.9995	0.9996	0.9996	0.9996	0.9996	0.9996	0.9996	0.9997
3.4	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9998

z 分配機率表

z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.0000	0.0040	0.0080	0.0120	0.0160	0.0199	0.0239	0.0279	0.0319	0.0359
0.1	0.0398	0.0438	0.0478	0.0517	0.0557	0.0596	0.0636	0.0675	0.0714	0.0753
0.2	0.0793	0.0832	0.0871	0.0910	0.0948	0.0987	0.1026	0.1064	0.1103	0.1141
0.3	0.1179	0.1217	0.1255	0.1293	0.1331	0.1368	0.1406	0.1443	0.1480	0.1517
0.4	0.1554	0.1591	0.1628	0.1664	0.1700	0.1736	0.1772	0.1808	0.1844	0.1879
0.5	0.1915	0.1950	0.1985	0.2019	0.2054	0.2088	0.2123	0.2157	0.2190	0.2224
0.6	0.2257	0.2291	0.2324	0.2357	0.2389	0.2422	0.2454	0.2486	0.2517	0.2549
0.7	0.2580	0.2611	0.2642	0.2673	0.2704	0.2734	0.2764	0.2794	0.2823	0.2852
0.8	0.2881	0.2910	0.2939	0.2967	0.2995	0.3023	0.3051	0.3078	0.3106	0.3133
0.9	0.3159	0.3186	0.3212	0.3238	0.3264	0.3289	0.3315	0.3340	0.3365	0.3389
1.0	0.3413	0.3438	0.3461	0.3485	0.3508	0.3531	0.3554	0.3577	0.3599	0.3621
1.1	0.3643	0.3665	0.3686	0.3708	0.3729	0.3749	0.3770	0.3790	0.3810	0.3830
1.2	0.3849	0.3869	0.3888	0.3907	0.3925	0.3944	0.3962	0.3980	0.3997	0.4015
1.3	0.4032	0.4049	0.4066	0.4082	0.4099	0.4115	0.4131	0.4147	0.4162	0.4177
1.4	0.4192	0.4207	0.4222	0.4236	0.4251	0.4265	0.4279	0.4292	0.4306	0.4319
1.5	0.4332	0.4345	0.4357	0.4370	0.4382	0.4394	0.4406	0.4418	0.4429	0.4441
1.6	0.4452	0.4463	0.4474	0.4484	0.4495	0.4505	0.4515	0.4525	0.4535	0.4545
1.7	0.4554	0.4564	0.4573	0.4582	0.4591	0.4599	0.4608	0.4616	0.4625	0.4633
1.8	0.4641	0.4649	0.4656	0.4664	0.4671	0.4678	0.4686	0.4693	0.4699	0.4706
1.9	0.4713	0.4719	0.4726	0.4732	0.4738	0.4744	0.4750	0.4756	0.4761	0.4767
2.0	0.4772	0.4778	0.4783	0.4788	0.4793	0.4798	0.4803	0.4808	0.4812	0.4817
2.1	0.4821	0.4826	0.4830	0.4834	0.4838	0.4842	0.4846	0.4850	0.4854	0.4857
2.2	0.4861	0.4864	0.4868	0.4871	0.4875	0.4878	0.4881	0.4884	0.4887	0.4890
2.3	0.4893	0.4896	0.4898	0.4901	0.4904	0.4906	0.4909	0.4911	0.4913	0.4916
2.4	0.4918	0.4920	0.4922	0.4925	0.4927	0.4929	0.4931	0.4932	0.4934	0.4936
2.5	0.4938	0.4940	0.4941	0.4943	0.4945	0.4946	0.4948	0.4949	0.4951	0.4952
2.6	0.4953	0.4955	0.4956	0.4957	0.4959	0.4960	0.4961	0.4962	0.4963	0.4964
2.7	0.4965	0.4966	0.4967	0.4968	0.4969	0.4970	0.4971	0.4972	0.4973	0.4974
2.8	0.4974	0.4975	0.4976	0.4977	0.4977	0.4978	0.4979	0.4979	0.4980	0.4981
2.9	0.4981	0.4982	0.4982	0.4983	0.4984	0.4984	0.4985	0.4985	0.4986	0.4986
3.0	0.4987	0.4987	0.4987	0.4988	0.4988	0.4989	0.4989	0.4989	0.4990	0.4990
3.1	0.4990	0.4991	0.4991	0.4991	0.4992	0.4992	0.4992	0.4992	0.4993	0.4993
3.2	0.4993	0.4993	0.4994	0.4994	0.4994	0.4994	0.4994	0.4995	0.4995	0.4995
3.3	0.4995	0.4995	0.4995	0.4996	0.4996	0.4996	0.4996	0.4996	0.4996	0.4997
3.4	0.4997	0.4997	0.4997	0.4997	0.4997	0.4997	0.4997	0.4997	0.4997	0.4998

t 分配**右尾**臨界值表

自由度	α				
	0.005	0.01	0.025	0.05	0.1
1	63.6567	31.8205	12.7062	6.3138	3.0777
2	9.9248	6.9646	4.3027	2.9200	1.8856
3	5.8409	4.5407	3.1824	2.3534	1.6377
4	4.6041	3.7469	2.7764	2.1318	1.5332
5	4.0321	3.3649	2.5706	2.0150	1.4759
6	3.7074	3.1427	2.4469	1.9432	1.4398
7	3.4995	2.9980	2.3646	1.8946	1.4149
8	3.3554	2.8965	2.3060	1.8595	1.3968
9	3.2498	2.8214	2.2622	1.8331	1.3830
10	3.1693	2.7638	2.2281	1.8125	1.3722
11	3.1058	2.7181	2.2010	1.7959	1.3634
12	3.0545	2.6810	2.1788	1.7823	1.3562
13	3.0123	2.6503	2.1604	1.7709	1.3502
14	2.9768	2.6245	2.1448	1.7613	1.3450
15	2.9467	2.6025	2.1314	1.7531	1.3406
16	2.9208	2.5835	2.1199	1.7459	1.3368
17	2.8982	2.5669	2.1098	1.7396	1.3334
18	2.8784	2.5524	2.1009	1.7341	1.3304
19	2.8609	2.5395	2.0930	1.7291	1.3277
20	2.8453	2.5280	2.0860	1.7247	1.3253
21	2.8314	2.5176	2.0796	1.7207	1.3232
22	2.8188	2.5083	2.0739	1.7171	1.3212
23	2.8073	2.4999	2.0687	1.7139	1.3195
24	2.7969	2.4922	2.0639	1.7109	1.3178
25	2.7874	2.4851	2.0595	1.7081	1.3163
26	2.7787	2.4786	2.0555	1.7056	1.3150
27	2.7707	2.4727	2.0518	1.7033	1.3137
28	2.7633	2.4671	2.0484	1.7011	1.3125
29	2.7564	2.4620	2.0452	1.6991	1.3114
30	2.7500	2.4573	2.0423	1.6973	1.3104
35	2.7238	2.4377	2.0301	1.6896	1.3062
40	2.7045	2.4233	2.0211	1.6839	1.3031
50	2.6778	2.4033	2.0086	1.6759	1.2987
60	2.6603	2.3901	2.0003	1.6706	1.2958
100	2.6259	2.3642	1.9840	1.6602	1.2901
999999	2.5758	2.3264	1.9600	1.6449	1.2816

χ^2 分配**右尾**臨界值表

自由度	α					α				
	0.1	0.05	0.025	0.01	0.005	0.995	0.99	0.975	0.95	0.9
1	2.7055	3.8415	5.0239	6.6349	7.8794	0.0000	0.0002	0.0010	0.0039	0.0158
2	4.6052	5.9915	7.3778	9.2103	10.5966	0.0100	0.0201	0.0506	0.1026	0.2107
3	6.2514	7.8147	9.3484	11.3449	12.8382	0.0717	0.1148	0.2158	0.3518	0.5844
4	7.7794	9.4877	11.1433	13.2767	14.8603	0.2070	0.2971	0.4844	0.7107	1.0636
5	9.2364	11.0705	12.8325	15.0863	16.7496	0.4117	0.5543	0.8312	1.1455	1.6103
6	10.6446	12.5916	14.4494	16.8119	18.5476	0.6757	0.8721	1.2373	1.6354	2.2041
7	12.0170	14.0671	16.0128	18.4753	20.2777	0.9893	1.2390	1.6899	2.1673	2.8331
8	13.3616	15.5073	17.5345	20.0902	21.9550	1.3444	1.6465	2.1797	2.7326	3.4895
9	14.6837	16.9190	19.0228	21.6660	23.5894	1.7349	2.0879	2.7004	3.3251	4.1682
10	15.9872	18.3070	20.4832	23.2093	25.1882	2.1559	2.5582	3.2470	3.9403	4.8652
11	17.2750	19.6751	21.9200	24.7250	26.7568	2.6032	3.0535	3.8157	4.5748	5.5778
12	18.5493	21.0261	23.3367	26.2170	28.2995	3.0738	3.5706	4.4038	5.2260	6.3038
13	19.8119	22.3620	24.7356	27.6882	29.8195	3.5650	4.1069	5.0088	5.8919	7.0415
14	21.0641	23.6848	26.1189	29.1412	31.3193	4.0747	4.6604	5.6287	6.5706	7.7895
15	22.3071	24.9958	27.4884	30.5779	32.8013	4.6009	5.2293	6.2621	7.2609	8.5468
16	23.5418	26.2962	28.8454	31.9999	34.2672	5.1422	5.8122	6.9077	7.9616	9.3122
17	24.7690	27.5871	30.1910	33.4087	35.7185	5.6972	6.4078	7.5642	8.6718	10.0852
18	25.9894	28.8693	31.5264	34.8053	37.1565	6.2648	7.0149	8.2307	9.3905	10.8649
19	27.2036	30.1435	32.8523	36.1909	38.5823	6.8440	7.6327	8.9065	10.1170	11.6509
20	28.4120	31.4104	34.1696	37.5662	39.9968	7.4338	8.2604	9.5908	10.8508	12.4426
21	29.6151	32.6706	35.4789	38.9322	41.4011	8.0337	8.8972	10.2829	11.5913	13.2396
22	30.8133	33.9244	36.7807	40.2894	42.7957	8.6427	9.5425	10.9823	12.3380	14.0415
23	32.0069	35.1725	38.0756	41.6384	44.1813	9.2604	10.1957	11.6886	13.0905	14.8480
24	33.1962	36.4150	39.3641	42.9798	45.5585	9.8862	10.8564	12.4012	13.8484	15.6587
25	34.3816	37.6525	40.6465	44.3141	46.9279	10.5197	11.5240	13.1197	14.6114	16.4734
26	35.5632	38.8851	41.9232	45.6417	48.2899	11.1602	12.1981	13.8439	15.3792	17.2919
27	36.7412	40.1133	43.1945	46.9629	49.6449	11.8076	12.8785	14.5734	16.1514	18.1139
28	37.9159	41.3371	44.4608	48.2782	50.9934	12.4613	13.5647	15.3079	16.9279	18.9392
29	39.0875	42.5570	45.7223	49.5879	52.3356	13.1211	14.2565	16.0471	17.7084	19.7677
30	40.2560	43.7730	46.9792	50.8922	53.6720	13.7867	14.9535	16.7908	18.4927	20.5992
31	41.4217	44.9853	48.2319	52.1914	55.0027	14.4578	15.6555	17.5387	19.2806	21.4336
32	42.5847	46.1943	49.4804	53.4858	56.3281	15.1340	16.3622	18.2908	20.0719	22.2706
33	43.7452	47.3999	50.7251	54.7755	57.6484	15.8153	17.0735	19.0467	20.8665	23.1102
34	44.9032	48.6024	51.9660	56.0609	58.9639	16.5013	17.7891	19.8063	21.6643	23.9523
35	46.0588	49.8018	53.2033	57.3421	60.2748	17.1918	18.5089	20.5694	22.4650	24.7967

F 分配**右尾**臨界值表

n2 \ n1	α = 0.05																			
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
1	161.45	199.50	215.71	224.58	230.16	233.99	236.77	238.88	240.54	241.88	242.98	243.91	244.69	245.36	245.95	246.46	246.92	247.32	247.69	248.01
2	18.513	19.000	19.164	19.247	19.296	19.330	19.353	19.371	19.385	19.396	19.405	19.413	19.419	19.424	19.429	19.433	19.437	19.440	19.443	19.446
3	10.128	9.5521	9.2766	9.1172	9.0135	8.9406	8.8867	8.8452	8.8123	8.7855	8.7633	8.7446	8.7287	8.7149	8.7029	8.6923	8.6829	8.6745	8.6670	8.6602
4	7.7086	6.9443	6.5914	6.3882	6.2561	6.1631	6.0942	6.0410	5.9988	5.9644	5.9358	5.9117	5.8911	5.8733	5.8578	5.8441	5.8320	5.8211	5.8114	5.8025
5	6.6079	5.7861	5.4095	5.1922	5.0503	4.9503	4.8759	4.8183	4.7725	4.7351	4.7040	4.6777	4.6552	4.6358	4.6188	4.6038	4.5904	4.5785	4.5678	4.5581
6	5.9874	5.1433	4.7571	4.5337	4.3874	4.2839	4.2067	4.1468	4.0990	4.0600	4.0274	3.9999	3.9764	3.9559	3.9381	3.9223	3.9083	3.8957	3.8844	3.8742
7	5.5914	4.7374	4.3468	4.1203	3.9715	3.8660	3.7870	3.7257	3.6767	3.6365	3.6030	3.5747	3.5503	3.5292	3.5107	3.4944	3.4799	3.4669	3.4551	3.4445
8	5.3177	4.4590	4.0662	3.8379	3.6875	3.5806	3.5005	3.4381	3.3881	3.3472	3.3130	3.2839	3.2590	3.2374	3.2184	3.2016	3.1867	3.1733	3.1613	3.1503
9	5.1174	4.2565	3.8625	3.6331	3.4817	3.3738	3.2927	3.2296	3.1789	3.1373	3.1025	3.0729	3.0475	3.0255	3.0061	2.9890	2.9737	2.9600	2.9477	2.9365
10	4.9646	4.1028	3.7083	3.4780	3.3258	3.2172	3.1355	3.0717	3.0204	2.9782	2.9430	2.9130	2.8872	2.8647	2.8450	2.8276	2.8120	2.7980	2.7854	2.7740
11	4.8443	3.9823	3.5874	3.3567	3.2039	3.0946	3.0123	2.9480	2.8962	2.8536	2.8179	2.7876	2.7614	2.7386	2.7186	2.7009	2.6851	2.6709	2.6581	2.6464
12	4.7472	3.8853	3.4903	3.2592	3.1059	2.9961	2.9134	2.8486	2.7964	2.7534	2.7173	2.6866	2.6602	2.6371	2.6169	2.5989	2.5828	2.5684	2.5554	2.5436
13	4.6672	3.8056	3.4105	3.1791	3.0254	2.9153	2.8321	2.7669	2.7144	2.6710	2.6347	2.6037	2.5769	2.5536	2.5331	2.5149	2.4987	2.4841	2.4709	2.4589
14	4.6001	3.7389	3.3439	3.1122	2.9582	2.8477	2.7642	2.6987	2.6458	2.6022	2.5655	2.5342	2.5073	2.4837	2.4630	2.4446	2.4282	2.4134	2.4000	2.3879
15	4.5431	3.6823	3.2874	3.0556	2.9013	2.7905	2.7066	2.6408	2.5876	2.5437	2.5068	2.4753	2.4481	2.4244	2.4034	2.3849	2.3683	2.3533	2.3398	2.3275
16	4.4940	3.6337	3.2389	3.0069	2.8524	2.7413	2.6572	2.5911	2.5377	2.4935	2.4564	2.4247	2.3973	2.3733	2.3522	2.3335	2.3167	2.3016	2.2880	2.2756
17	4.4513	3.5915	3.1968	2.9647	2.8100	2.6987	2.6143	2.5480	2.4943	2.4499	2.4126	2.3807	2.3531	2.3290	2.3077	2.2888	2.2719	2.2567	2.2429	2.2304
18	4.4139	3.5546	3.1599	2.9277	2.7729	2.6613	2.5767	2.5102	2.4563	2.4117	2.3742	2.3421	2.3143	2.2900	2.2686	2.2496	2.2325	2.2172	2.2033	2.1906
19	4.3807	3.5219	3.1274	2.8951	2.7401	2.6283	2.5435	2.4768	2.4227	2.3779	2.3402	2.3080	2.2800	2.2556	2.2341	2.2149	2.1977	2.1823	2.1683	2.1555
20	4.3512	3.4928	3.0984	2.8661	2.7109	2.5990	2.5140	2.4471	2.3928	2.3479	2.3100	2.2776	2.2495	2.2250	2.2033	2.1840	2.1667	2.1511	2.1370	2.1242
21	4.3248	3.4668	3.0725	2.8401	2.6848	2.5727	2.4876	2.4205	2.3660	2.3210	2.2829	2.2504	2.2222	2.1975	2.1757	2.1563	2.1389	2.1232	2.1090	2.0960
22	4.3009	3.4434	3.0491	2.8167	2.6613	2.5491	2.4638	2.3965	2.3419	2.2967	2.2585	2.2258	2.1975	2.1727	2.1508	2.1313	2.1138	2.0980	2.0837	2.0707
23	4.2793	3.4221	3.0280	2.7955	2.6400	2.5277	2.4422	2.3748	2.3201	2.2747	2.2364	2.2036	2.1752	2.1502	2.1282	2.1086	2.0910	2.0751	2.0608	2.0476
24	4.2597	3.4028	3.0088	2.7763	2.6207	2.5082	2.4226	2.3551	2.3002	2.2547	2.2163	2.1834	2.1548	2.1298	2.1077	2.0880	2.0703	2.0543	2.0399	2.0267
25	4.2417	3.3852	2.9912	2.7587	2.6030	2.4904	2.4047	2.3371	2.2821	2.2365	2.1979	2.1649	2.1362	2.1111	2.0889	2.0691	2.0513	2.0353	2.0207	2.0075
26	4.2252	3.3690	2.9752	2.7426	2.5868	2.4741	2.3883	2.3205	2.2655	2.2197	2.1811	2.1479	2.1192	2.0939	2.0716	2.0518	2.0339	2.0178	2.0032	1.9898
27	4.2100	3.3541	2.9604	2.7278	2.5719	2.4591	2.3732	2.3053	2.2501	2.2043	2.1655	2.1323	2.1035	2.0781	2.0558	2.0358	2.0179	2.0017	1.9870	1.9736
28	4.1960	3.3404	2.9467	2.7141	2.5581	2.4453	2.3593	2.2913	2.2360	2.1900	2.1512	2.1179	2.0889	2.0635	2.0411	2.0210	2.0030	1.9868	1.9720	1.9586
29	4.1830	3.3277	2.9340	2.7014	2.5454	2.4324	2.3463	2.2783	2.2229	2.1768	2.1379	2.1045	2.0755	2.0500	2.0275	2.0073	1.9893	1.9730	1.9581	1.9446
30	4.1709	3.3158	2.9223	2.6896	2.5336	2.4205	2.3343	2.2662	2.2107	2.1646	2.1256	2.0921	2.0630	2.0374	2.0148	1.9946	1.9765	1.9601	1.9452	1.9317
999999	3.8415	2.9957	2.6049	2.3719	2.2141	2.0986	2.0096	1.9384	1.8799	1.8307	1.7887	1.7522	1.7202	1.6918	1.6664	1.6435	1.6228	1.6039	1.5865	1.5705

F 分配**右尾**臨界值表

n2 \ n1	α = 0.01																			
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
1	4052.2	4999.5	5403.4	5624.6	5763.6	5859.0	5928.4	5981.1	6022.5	6055.8	6083.3	6106.3	6125.9	6142.7	6157.3	6170.1	6181.4	6191.5	6200.6	6208.7
2	98.503	99.000	99.166	99.249	99.299	99.333	99.356	99.374	99.388	99.399	99.408	99.416	99.422	99.428	99.433	99.437	99.440	99.444	99.447	99.449
3	34.116	30.817	29.457	28.710	28.237	27.911	27.672	27.489	27.345	27.229	27.133	27.052	26.983	26.924	26.872	26.827	26.787	26.751	26.719	26.690
4	21.198	18.000	16.694	15.977	15.522	15.207	14.976	14.799	14.659	14.546	14.452	14.374	14.307	14.249	14.198	14.154	14.115	14.080	14.048	14.020
5	16.258	13.274	12.060	11.392	10.967	10.672	10.456	10.289	10.158	10.051	9.9626	9.8883	9.8248	9.7700	9.7222	9.6802	9.6429	9.6096	9.5797	9.5526
6	13.745	10.925	9.7795	9.1483	8.7459	8.4661	8.2600	8.1017	7.9761	7.8741	7.7896	7.7183	7.6575	7.6049	7.5590	7.5186	7.4827	7.4507	7.4219	7.3958
7	12.246	9.5466	8.4513	7.8466	7.4604	7.1914	6.9928	6.8400	6.7188	6.6201	6.5382	6.4691	6.4100	6.3590	6.3143	6.2750	6.2401	6.2089	6.1808	6.1554
8	11.259	8.6491	7.5910	7.0061	6.6318	6.3707	6.1776	6.0289	5.9106	5.8143	5.7343	5.6667	5.6089	5.5589	5.5151	5.4766	5.4423	5.4116	5.3840	5.3591
9	10.561	8.0215	6.9919	6.4221	6.0569	5.8018	5.6129	5.4671	5.3511	5.2565	5.1779	5.1114	5.0545	5.0052	4.9621	4.9240	4.8902	4.8599	4.8327	4.8080
10	10.044	7.5594	6.5523	5.9943	5.6363	5.3858	5.2001	5.0567	4.9424	4.8491	4.7715	4.7059	4.6496	4.6008	4.5581	4.5204	4.4869	4.4569	4.4299	4.4054
11	9.6460	7.2057	6.2167	5.6683	5.3160	5.0692	4.8861	4.7445	4.6315	4.5393	4.4624	4.3974	4.3416	4.2932	4.2509	4.2134	4.1801	4.1503	4.1234	4.0990
12	9.3302	6.9266	5.9525	5.4120	5.0643	4.8206	4.6395	4.4994	4.3875	4.2961	4.2198	4.1553	4.0999	4.0518	4.0096	3.9724	3.9392	3.9095	3.8827	3.8584
13	9.0738	6.7010	5.7394	5.2053	4.8616	4.6204	4.4410	4.3021	4.1911	4.1003	4.0245	3.9603	3.9052	3.8573	3.8154	3.7783	3.7452	3.7156	3.6888	3.6646
14	8.8616	6.5149	5.5639	5.0354	4.6950	4.4558	4.2779	4.1399	4.0297	3.9394	3.8640	3.8001	3.7452	3.6975	3.6557	3.6187	3.5857	3.5561	3.5294	3.5052
15	8.6831	6.3589	5.4170	4.8932	4.5556	4.3183	4.1415	4.0045	3.8948	3.8049	3.7299	3.6662	3.6115	3.5639	3.5222	3.4852	3.4523	3.4228	3.3961	3.3719
16	8.5310	6.2262	5.2922	4.7726	4.4374	4.2016	4.0259	3.8896	3.7804	3.6909	3.6162	3.5527	3.4981	3.4506	3.4089	3.3720	3.3391	3.3096	3.2829	3.2587
17	8.3997	6.1121	5.1850	4.6690	4.3359	4.1015	3.9267	3.7910	3.6822	3.5931	3.5185	3.4552	3.4007	3.3533	3.3117	3.2748	3.2419	3.2124	3.1857	3.1615
18	8.2854	6.0129	5.0919	4.5790	4.2479	4.0146	3.8406	3.7054	3.5971	3.5082	3.4338	3.3706	3.3162	3.2689	3.2273	3.1904	3.1575	3.1280	3.1013	3.0771
19	8.1849	5.9259	5.0103	4.5003	4.1708	3.9386	3.7653	3.6305	3.5225	3.4338	3.3596	3.2965	3.2422	3.1949	3.1533	3.1165	3.0836	3.0541	3.0274	3.0031
20	8.0960	5.8489	4.9382	4.4307	4.1027	3.8714	3.6987	3.5644	3.4567	3.3682	3.2941	3.2311	3.1769	3.1296	3.0880	3.0512	3.0183	2.9887	2.9620	2.9377
21	8.0166	5.7804	4.8740	4.3688	4.0421	3.8117	3.6396	3.5056	3.3981	3.3098	3.2359	3.1730	3.1187	3.0715	3.0300	2.9931	2.9602	2.9306	2.9039	2.8796
22	7.9454	5.7190	4.8166	4.3134	3.9880	3.7583	3.5867	3.4530	3.3458	3.2576	3.1837	3.1209	3.0667	3.0195	2.9779	2.9411	2.9082	2.8786	2.8518	2.8274
23	7.8811	5.6637	4.7649	4.2636	3.9392	3.7102	3.5390	3.4057	3.2986	3.2106	3.1368	3.0740	3.0199	2.9727	2.9311	2.8943	2.8613	2.8317	2.8049	2.7805
24	7.8229	5.6136	4.7181	4.2184	3.8951	3.6667	3.4959	3.3629	3.2560	3.1681	3.0944	3.0316	2.9775	2.9303	2.8887	2.8519	2.8189	2.7892	2.7624	2.7380
25	7.7698	5.5680	4.6755	4.1774	3.8550	3.6272	3.4568	3.3239	3.2172	3.1294	3.0558	2.9931	2.9389	2.8917	2.8502	2.8133	2.7803	2.7506	2.7238	2.6993
26	7.7213	5.5263	4.6366	4.1400	3.8183	3.5911	3.4210	3.2884	3.1818	3.0941	3.0205	2.9578	2.9038	2.8566	2.8150	2.7781	2.7451	2.7153	2.6885	2.6640
27	7.6767	5.4881	4.6009	4.1056	3.7848	3.5580	3.3882	3.2558	3.1494	3.0618	2.9882	2.9256	2.8715	2.8243	2.7827	2.7458	2.7127	2.6830	2.6561	2.6316
28	7.6356	5.4529	4.5681	4.0740	3.7539	3.5276	3.3581	3.2259	3.1195	3.0320	2.9585	2.8959	2.8418	2.7946	2.7530	2.7160	2.6830	2.6532	2.6263	2.6017
29	7.5977	5.4204	4.5378	4.0449	3.7254	3.4995	3.3303	3.1982	3.0920	3.0045	2.9311	2.8685	2.8144	2.7672	2.7256	2.6886	2.6555	2.6257	2.5987	2.5742
30	7.5625	5.3903	4.5097	4.0179	3.6990	3.4735	3.3045	3.1726	3.0665	2.9791	2.9057	2.8431	2.7890	2.7418	2.7002	2.6632	2.6301	2.6003	2.5732	2.5487
999999	6.6349	4.6052	3.7816	3.3192	3.0173	2.8020	2.6393	2.5113	2.4074	2.3209	2.2477	2.1848	2.1299	2.0815	2.0385	2.0000	1.9652	1.9336	1.9048	1.8783

靜宜大學企管系『統計學』講義—微積分基本概念

2006 年 10 月 25 日 最後修改

本講義的目標讀者是正在學習統計的同學，目的在於讓讀者瞭解微分與積分在統計上的應用，希望同學讀得懂微積分符號並會實際操作。

一、基本函數的性質

學習微積分必須瞭解數學函數。這裡只介紹三個最常用的函數：多項函數、指數函數與、對數函數。初級微積分還會介紹三角函數。三角函數比較複雜，而且在初等統計學中用的並不多，因此本講義將它們省略掉。

多項函數可以寫成

$$x^n \quad \text{其中} \quad x\text{爲變數}, \; n \in R$$

原本 x^n 表示將 x 連乘 n 次，廣義後，n 不是整數也沒有關係。請區別多項式與多項函數，多項式是將幾個多項函數加（減）在一起，以下是幾個多項函數的多項式：

$$x^5 + 3x^2 + 4$$
$$2x^4 + (x+5)^3 + 6x^2$$

以上兩者都是三項的多項式，其中第二個多項式的第二項是以多項式爲底的多項函數。

指數函數的變數在指數，以 a 爲底的指數函數可以寫成

$$a^x \text{ 或 } a^{f(x)} \quad \text{其中} \quad a \in R, \; x\text{爲變數}, \; f(x)\text{爲任意函數}$$

指數函數中有一個稱爲自然數的特殊底，我們以 e 來代表。e 是一個無理數，其數值大略如下：

$$e = 2.71828182845905$$

以自然數爲底的指數函數 e^x 有下列兩個定義：

e^x 的定義	$e^x = \lim_{n\to\infty}\left(1+\frac{x}{n}\right)^n$ $e^x = \sum_{n=0}^{\infty}\frac{1}{n!}x^n = 1 + x + \frac{1}{2!}x^2 + \cdots + \frac{1}{n!}x^n + \cdots$

指數函數有下列特性：

$$a^x \times a^y = a^{x+y}$$
$$a^x \times b^x = (ab)^x$$

對數函數是指數函數的反函數，以 a 爲底的對數函數可以寫成

$\log_a x$ 或 $\log_a f(x)$ 其中 $a \in R$, x爲變數, $f(x)$爲任意函數

若以自然數爲底，則寫成

$\ln x = \log_e x$ 或 $\ln f(x) = \log_e f(x)$ 其中 x爲變數, $f(x)$爲任意函數

對數函數有以下的性質：

$$\log\left(f(x)\times g(x)\right) = \log f(x) + \log g(x)$$
$$\log_a x = \frac{\log x}{\log a}$$

其中我們沒有特地標明對數的底爲何，表示任何底都成立；另外，上面的第二式稱爲對數函數的換底公式，它可以將原有的底換成任何適當的底。例如我們常希望換成以自然數爲底的對數：$\log_a x = \frac{\ln x}{\ln a} = \frac{1}{\ln a}\ln x$ 。

對數函數與指數函數互爲反函數，因此有下列性質：

$$\log_a a^x = x$$
$$\ln e^x = x$$
$$\log a^x = x\log a$$
$$a^{\log_a x} = x$$
$$e^{\ln x} = x$$
$$a^{\log x} = x^{\log a}$$

$$a^x = e^{\ln a^x} = e^{x\ln a}$$

以下是一些指數函數與對數函數的計算例題：

$$\ln 5x = \ln 5 + \ln x$$
$$\ln x^3 = 3\ln x$$
$$\ln\left(6x^4y^2\right) = \ln 6 + 4\ln x + 2\ln y$$
$$\ln 7^x = x\ln 7$$
$$\log_8 x = \frac{\ln x}{\ln 8} = \frac{\ln x}{3\ln 2}$$
$$e^{2+x+y} = e^2 e^x e^y$$

二、微分基本公式

以下是三個基本函數（多項函數、指數函數、對數函數）的微分基本公式：

微分基本公式一 $\frac{d}{dx}\left[x^n\right] = \frac{d\left(x^n\right)}{dx} = nx^{n-1}$

微分基本公式二 $\frac{d}{dx}\left[e^x\right] = \frac{d\left(e^x\right)}{dx} = e^x$

微分基本公式三 $\frac{d}{dx}\left[\ln x\right] = \frac{d\left(\ln x\right)}{dx} = \frac{1}{x}$

如果指數函數與對數函數不以自然數（e）爲底，則先換底後再微分，其結果如下：

$$\frac{d}{dx}\left[a^x\right] = \frac{d}{dx}\left[e^{x\ln a}\right] = \ln a \times a^x$$

$$\frac{d}{dx}\left[\log_a x\right] = \frac{d}{dx}\left[\frac{\ln x}{\ln a}\right] = \frac{1}{\ln a} \times \frac{1}{x}$$

以下是幾個計算例題：

$$\frac{d}{dx}\left[x^5\right] = 5x^4$$

$$\frac{d}{dx}\left[x^{-2}\right]=-2x^{-3}$$
$$\frac{d}{dx}\left[\frac{1}{x^{7}}\right]=\frac{d}{dx}\left[x^{-7}\right]=-7x^{-8}=\frac{-7}{x^{8}}$$
$$\frac{d}{dx}\left[5^{x}\right]=\ln 5\times 5^{x}$$
$$\frac{d}{dx}\left[\log_{8}x\right]=\frac{1}{\ln 8}\times\frac{1}{x}$$

以上微分符號也可以作以下的寫法：

微分基本公式一	$\frac{d\left(x^{n}\right)}{dx}=nx^{n-1} \quad\Rightarrow\quad d\left(x^{n}\right)=nx^{n-1}dx$
微分基本公式二	$\frac{d\left(e^{x}\right)}{dx}=e^{x} \quad\Rightarrow\quad d\left(e^{x}\right)=e^{x}dx$
微分基本公式三	$\frac{d\left(\ln x\right)}{dx}=\frac{1}{x} \quad\Rightarrow\quad d\left(\ln x\right)=\frac{1}{x}dx$

這種表示法是把 $d\left(x^{n}\right)$、$d\left(e^{x}\right)$、$d\left(\ln x\right)$、dx 等微分操作當成一個基本算數單元，這有助於複雜微分式的瞭解。以下是上面例題用這種符號改寫後的結果。

$$d\left(x^{5}\right)=5x^{4}dx$$
$$d\left(x^{-2}\right)=-2x^{-3}dx$$
$$d\left(\frac{1}{x^{7}}\right)=d\left(x^{-7}\right)=-7x^{-8}dx=\frac{-7}{x^{8}}dx$$
$$d\left(5^{x}\right)=\ln 5\times 5^{x}dx$$
$$d\left(\log_{8}x\right)=\frac{1}{\ln 8}\times\frac{1}{x}dx$$

基本公式一至三呈現只有一個基本函數的微分情況，以下四個基本公式處理基本函數組合的微分。這些公式中，$f\left(x\right)$、$g\left(x\right)$ 爲任意基本函數或其組合，c 爲任意實數。

微分基本公式四	$d\left[cf\left(x\right)\right]=c\times d\left[f\left(x\right)\right] \quad\Rightarrow\quad d\left(c\times f\right)=c\times d\left(f\right)$
微分基本公式五	$d\left[f\left(x\right)+g\left(x\right)\right]=d\left[f\left(x\right)\right]+d\left[g\left(x\right)\right] \quad\Rightarrow\quad d\left(f+g\right)=d\left(f\right)+d\left(g\right)$

微分基本公式六 $d\left[f\times g\right]=d\left(f\right)\times g+f\times d\left(g\right)$

微分基本公式七 $\dfrac{df}{dx}=\dfrac{df}{dg}\times\dfrac{dg}{dx}$

其中，基本公式七稱爲微分的 *chain rule*。這些公式也可以延伸至兩個以上函數的情況，例如，令 $h(x)$ 也是一個基本函數或其組合，則

$$d\left(f+g+h\right)=d\left(f\right)+d\left(g\right)+d\left(h\right)$$
$$d\left(f\times g\times h\right)=d\left(f\right)\times g\times h+f\times d\left(g\right)\times h+f\times g\times d\left(h\right)$$

以下是基本公式四、五的練習題：

$$d\left(6x^3\right)=6\times d\left(x^3\right)=6\times 3x^2dx=18x^2dx$$
$$d\left(2e^x+5\ln x\right)=2d\left(e^x\right)+5d\left(\ln x\right)=\left(2e^x+\frac{5}{x}\right)dx$$

以下是基本公式六的練習題：

$$d\left(x^4e^x\right)=d\left(x^4\right)e^x+x^4d\left(e^x\right)=4x^3e^xdx+x^4e^xdx=\left(4x^3e^x+x^4e^x\right)dx$$
$$\frac{d\left(x^4\ln x\right)}{dx}=\frac{d\left(x^4\right)}{dx}\ln x+x^4\frac{d\left(\ln x\right)}{dx}=4x^3\ln x+x^4\frac{1}{x}=4x^3\ln x+x^3$$
$$\frac{d\left[e^x\left(x^3+x^2\right)\right]}{dx}=\frac{d\left(e^x\right)}{dx}\left(x^3+x^2\right)+e^x\frac{d\left(x^3+x^2\right)}{dx}=e^x\left(x^3+x^2\right)+e^x\left(3x^2+2x\right)$$

以下是基本公式七的練習題：

$$\frac{d\left[\left(x^3+x\right)^5\right]}{dx}=\frac{d\left[\left(x^3+x\right)^5\right]}{d\left(x^3+x\right)}\times\frac{d\left(x^3+x\right)}{dx}=5\left(x^3+x\right)^4\times\left(3x^2+1\right)$$
$$\frac{d\left[\left(x^3+x\right)^5\right]}{dx}=\frac{d\left[y^5\right]}{dy}\times\frac{dy}{dx}=5y^4\times\frac{d\left(x^3+x\right)}{dx}=5\left(x^3+x\right)^4\times\left(3x^2+1\right)$$
$$\frac{d\left(e^{3x}\right)}{dx}=\frac{d\left(e^{3x}\right)}{d\left(3x\right)}\times\frac{d\left(3x\right)}{dx}=e^{3x}\times 3=3e^{3x}$$
$$\frac{d\left(e^{3x}\right)}{dx}=\frac{d\left(e^y\right)}{dy}\times\frac{dy}{dx}=e^y\times\frac{d\left(3x\right)}{dx}=e^{3x}\times 3=3e^{3x}$$

$$\frac{d\left(e^{\left(x^3+x\right)^5}\right)}{dx}=\frac{d\left(e^{\left(x^3+x\right)^5}\right)}{d\left[\left(x^3+x\right)^5\right]}\times\frac{d\left[\left(x^3+x\right)^5\right]}{d\left(x^3+x\right)}\times\frac{d\left(x^3+x\right)}{dx}=e^{\left(x^3+x\right)^5}\times 5\left(x^3+x\right)^4\times\left(3x^2+1\right)$$

$$\frac{d\left(e^{\left(x^3+x\right)^5}\right)}{dx}=\frac{d\left(e^y\right)}{dy}\times\frac{dy}{dx}=e^y\times\frac{d\left[\left(x^3+x\right)^5\right]}{dx}=e^y\times\frac{d\left(z^5\right)}{dz}\times\frac{dz}{dx}$$

$$=e^y\times 5z^4\times\frac{d\left(x^3+x\right)}{dx}=e^{\left(x^3+x\right)^5}\times 5\left(x^3+x\right)^4\times\left(3x^2+1\right)$$

以下是綜合練習題：

$$\frac{d\left[x^2e^{x^2}\right]}{dx}=\frac{d\left(x^2\right)}{dx}e^{x^2}+x^2\frac{d\left(e^{x^2}\right)}{dx}=2xe^{x^2}+x^2e^{x^2}\left(2x\right)=2xe^{x^2}+2x^3e^{x^2}$$

$$\frac{d\left(\frac{1}{1-x}\right)}{dx}=\frac{d\left[\left(1-x\right)^{-1}\right]}{dx}=\frac{d\left[\left(1-x\right)^{-1}\right]}{d\left(1-x\right)}\times\frac{d\left(1-x\right)}{dx}=-\left(1-x\right)^{-2}\times\left(-1\right)=\frac{1}{\left(1-x\right)^2}$$

三、積分基本公式

原則上積分是微分的逆向操作，以下可以視爲積分的基本定義：

$$\int d\left[f\left(x\right)\right]=f\left(x\right)$$

$$\int_{x=a}^{x=b}d\left[f\left(x\right)\right]=f\left(x\right)\Big|_{x=a}^{x=b}=f\left(b\right)-f\left(a\right)$$

其中，第一式沒有指定變數 x 的範圍，稱爲『不定積分』。

相對於微分基本公式一、二、三，我們有以下積分基本公式一、二、三：

積分基本公式一 $\int nx^{n-1}dx=\int d\left(x^n\right)=x^n \quad\Rightarrow\quad \int x^n dx=\frac{1}{n+1}x^{n+1}$

積分基本公式二 $\int e^x dx=\int d\left(e^x\right)=e^x$

積分基本公式三 $\int\frac{1}{x}dx=\int d\left(\ln x\right)=\ln x$

以下是基本公式一、二、三的練習題：

$$\int x^5 dx = \frac{1}{6}x^6$$

$$\int_3^4 x^2 dx = \frac{1}{3}x^3\Big|_3^4 = \frac{1}{3}\left(4^3 - 3^3\right) = \frac{37}{3}$$

$$\int_4^\infty \frac{1}{x^3}dx = \int_4^\infty x^{-3}dx = \frac{1}{-2}x^{-2}\Big|_4^\infty = \frac{1}{-2}\left(0 - 4^{-2}\right) = \frac{1}{32}$$

$$\int_0^1 e^x dx = e^x\Big|_0^1 = e^1 - e^0 = e - 1$$

$$\int_1^2 \frac{1}{x}dx = \ln x\Big|_1^2 = \ln 2 - \ln 1 = \ln 2$$

相對於微分基本公式四、五，我們有以下積分基本公式四、五：

積分基本公式四	$\int c\times f\left(x\right)dx = c\times\int f\left(x\right)dx$
積分基本公式五	$\int\left[f\left(x\right)+g\left(x\right)\right]dx = \int f\left(x\right)dx + \int g\left(x\right)dx$

以下是基本公式四、五的練習題：

$$\int 3x^5 dx = 3\int x^5 dx = 3\times\frac{1}{6}x^6 = \frac{1}{2}x^6$$

$$\int_3^4\left(3x^5 + x^2\right)dx = \left(\frac{1}{2}x^6 + \frac{1}{3}x^3\right)\Bigg|_3^4 = \frac{1}{2}\left(4^6 - 3^6\right) + \frac{1}{3}\left(4^3 - 3^3\right) = \frac{3367}{2} + \frac{37}{3}$$

相對於微分基本公式六的積分稱爲『**部分積分**』,其道理請參考下列數學式的推導：

$$d\left[f\times g\right] = d\left(f\right)\times g + f\times d\left(g\right)$$
$$\Rightarrow f\times d\left(g\right) = d\left[f\times g\right] - d\left(f\right)\times g$$
$$\Rightarrow \int f\times d\left(g\right) = \int d\left[f\times g\right] - \int g\times d\left(f\right) = f\times g - \int g\times d\left(f\right)$$

令 $f'\left(x\right) = \frac{d}{dx}\left[f\left(x\right)\right],\ g'\left(x\right) = \frac{d}{dx}\left[g\left(x\right)\right]$，則有下列公式

積分基本公式六	$\int f\left(x\right)\times g'\left(x\right)dx = f\left(x\right)\times g\left(x\right) - \int g\left(x\right)\times f'\left(x\right)dx$

以下是基本公式六（部分積分）的練習題：

$$\int x\lambda e^{-\lambda x}dx = ?$$

令 $f=x,\ \ g'=\lambda e^{-\lambda x} \quad \Rightarrow \quad f'=\dfrac{d}{dx}\left[x\right]=1,\ \ g=\int \lambda e^{-\lambda x}dx=-e^{-\lambda x}$

$\int x\lambda e^{-\lambda x}dx=f\,g-\int g\,f'dx=-xe^{-\lambda x}-\int\left(-e^{-\lambda x}\right)dx=-xe^{-\lambda x}-\frac{1}{\lambda}e^{-\lambda x}$

以下是另一個基本公式六（部分積分）的練習題：

$\int_0^\infty x^2\lambda e^{-\lambda x}dx=?$

令 $f=x^2,\ \ g'=\lambda e^{-\lambda x} \quad \Rightarrow \quad f'=\dfrac{d}{dx}\left[x^2\right]=2x,\ \ g=\int \lambda e^{-\lambda x}dx=-e^{-\lambda x}$

$$\begin{aligned}\int x^2\lambda e^{-\lambda x}dx&=-x^2e^{-\lambda x}-\int\left(-2xe^{-\lambda x}\right)dx=-x^2e^{-\lambda x}+\tfrac{2}{\lambda}\int x\lambda e^{-\lambda x}dx\\&=-x^2e^{-\lambda x}+\tfrac{2}{\lambda}\left(-xe^{-\lambda x}-\tfrac{1}{\lambda}e^{-\lambda x}\right)=\left(-x^2-\tfrac{2}{\lambda}x-\tfrac{2}{\lambda^2}\right)e^{-\lambda x}\end{aligned}$$

因此 $\int_0^\infty x^2\lambda e^{-\lambda x}dx=\left(-x^2-\frac{2}{\lambda}x-\frac{2}{\lambda^2}\right)e^{-\lambda x}\Big|_0^\infty=\frac{2}{\lambda^2}$

相對於微分基本公式七的積分稱爲『**變數變換**』：

積分基本公式七 $\int f\left(g\right)\times g'\left(x\right)dx=\int f\left(g\right)dg$

以下是基本公式七（變數變換）的練習題：

$\int 2x\left(x^2-3\right)^4dx=?$

令 $g=x^2-3$ 則 $dg=2xdx,\ \left(x^2-3\right)^4=g^4$

$\Rightarrow \int 2x\left(x^2-3\right)^4dx=\int\left(x^2-3\right)^4 2xdx=\int g^4dg=\frac{1}{5}g^5=\frac{1}{5}\left(x^2-3\right)^5$

如果指定積分變數的上、下限，經變數變換後，上下限需要跟著變換：

$\int_{x=2}^{x=3}2x\left(x^2-3\right)^4dx=?$

令 $g=x^2-3$ 則 $dg=2xdx,\ \left(x^2-3\right)^4=g^4$

$\Rightarrow \int_{x=2}^{x=3}2x\left(x^2-3\right)^4dx=\frac{1}{5}\left(x^2-3\right)^5\Big|_{x=2}^{x=3}=\frac{1}{5}\left(6^5-1^5\right)$

或者 $\int_{x=2}^{x=3}2x\left(x^2-3\right)^4dx=\int_{g=2^2-3=1}^{g=3^2-3=6}g^4dg=\frac{1}{5}g^5\Big|_{g=1}^{g=6}=\frac{1}{5}\left(6^5-1^5\right)$

以下是另一個基本公式七（變數變換）的練習題：

$$\int_{-\infty}^{\infty} z\frac{1}{\sqrt{2\pi}}e^{-\frac{z^2}{2}}dz=?$$

令 $g=e^{-\frac{z^2}{2}}$ 則 $dg=-ze^{-\frac{z^2}{2}}dz$

$$\Rightarrow \int_{-\infty}^{\infty} z\frac{1}{\sqrt{2\pi}}e^{-\frac{z^2}{2}}dz=-\frac{1}{\sqrt{2\pi}}\int_{z=-\infty}^{z=\infty}dg=-g\Big|_{z=-\infty}^{z=\infty}=-e^{-\frac{z^2}{2}}\Big|_{z=-\infty}^{z=\infty}=0-0=0$$

四、多重微分與多重積分

兩個變數以上微分或積分稱爲多重微分或多重積分。

多重微分的原則與單一變數微分的原則一樣，唯一要注意的地方是，當對某一變數微分的時候，將其他變數視爲常數。另外，我們會以『$\frac{\partial}{\partial x}\left[f(x,y)\right]$』（**偏微分**）來表示其微分。

以下是幾個多重微分的例題：

$$\frac{\partial}{\partial x}\left(y^2x+yx^2\right)=\frac{\partial}{\partial x}\left(y^2x\right)+\frac{\partial}{\partial x}\left(yx^2\right)=y^2+2yx$$

$$\frac{\partial}{\partial y}\left(y^2x+yx^2\right)=\frac{\partial}{\partial y}\left(y^2x\right)+\frac{\partial}{\partial y}\left(yx^2\right)=2yx+x^2$$

$$\frac{\partial^2}{\partial x^2}\left(y^2x+yx^2\right)=\frac{\partial}{\partial x}\left[\frac{\partial}{\partial x}\left(y^2x+yx^2\right)\right]=\frac{\partial}{\partial x}\left(y^2+2yx\right)=2y$$

$$\frac{\partial^2}{\partial y^2}\left(y^2x+yx^2\right)=\frac{\partial}{\partial y}\left[\frac{\partial}{\partial y}\left(y^2x+yx^2\right)\right]=\frac{\partial}{\partial y}\left(2yx+x^2\right)=2x$$

$$\frac{\partial^2}{\partial x\partial y}\left(y^2x+yx^2\right)=\frac{\partial}{\partial x}\left[\frac{\partial}{\partial y}\left(y^2x+yx^2\right)\right]=\frac{\partial}{\partial x}\left(2yx+x^2\right)=2y+2x$$

$$\frac{\partial^2}{\partial y\partial x}\left(y^2x+yx^2\right)=\frac{\partial}{\partial y}\left[\frac{\partial}{\partial x}\left(y^2x+yx^2\right)\right]=\frac{\partial}{\partial y}\left(y^2+2yx\right)=2y+2x$$

相對於偏微分，當我們寫『$d\left[f(x,y)\right]$』時，表示的是**全微分**：

全微分 $d\left[f(x,y)\right]=\frac{\partial}{\partial x}\left[f(x,y)\right]dx+\frac{\partial}{\partial y}\left[f(x,y)\right]dy$

例如：

$$d\left(y^2x+yx^2\right)=\frac{\partial}{\partial x}\left(y^2x+yx^2\right)dx+\frac{\partial}{\partial y}\left(y^2x+yx^2\right)dy=\left(y^2+2yx\right)dx+\left(2yx+x^2\right)dy$$

如同多重微分與單一變數微分的關係，多重積分的原則與大部分單一變數積分的原

則一樣，只是當對某一變數積分的時候，將其他變數視爲常數。

以下是幾個多重積分的例題：

$$\int_y\int_x 6xy^2dxdy=\int_y 3y^2\left(\int_x 2xdx\right)dy=\int_y 3y^2x^2dy=x^2\int_y 3y^2dy=x^2y^3$$

$$\int_y\int_x 6xy^2dxdy=\int_x\int_y 6xy^2dydx=\int_x 2x\left(\int_y 3y^2dy\right)dx=\int_x 2xy^3dx=y^3\int_x 2xdx=y^3x^2$$

$$\int_0^2\int_1^y 6xy^2dxdy=\int_0^2 3y^2\left(\int_1^y 2xdx\right)dy=\int_0^2 3y^2\left(x^2\Big|_1^y\right)dy=\int_0^2 3y^2\left(y^2-1\right)dy$$

$$=\int_0^2\left(3y^4-3y^2\right)dy=\left(\frac{3}{5}y^5-y^3\right)\Bigg|_0^2=\frac{96}{5}-8=\frac{56}{5}$$

$$\int_0^1\int_0^1(x+y)(x-y)^2\,dxdy=\int_0^1\int_0^1\left(x^3-x^2y-xy^2+y^3\right)dxdy$$

$$=\int_0^1\left(\tfrac{1}{4}-\tfrac{1}{3}y-\tfrac{1}{2}y^2+y^3\right)dy=\tfrac{1}{4}-\tfrac{1}{3}\times\tfrac{1}{2}-\tfrac{1}{2}\times\tfrac{1}{3}+\tfrac{1}{4}=\tfrac{1}{6}$$

多重積分在作變數變換時，其操作比較複雜。若 x、y 是原變數，u、v 是轉換以後的變數，將原變數寫成新變數的函數，則我們有以下的轉換規則。

雙變數之變數轉換　令 $x=x(u,v),\ y=y(u,v),\ J=\begin{vmatrix}\frac{\partial x}{\partial u} & \frac{\partial x}{\partial v}\\ \frac{\partial y}{\partial u} & \frac{\partial y}{\partial v}\end{vmatrix}$, 則 $dxdy=|J|dudv$

上面的$\begin{vmatrix}\frac{\partial x}{\partial u} & \frac{\partial x}{\partial v}\\ \frac{\partial y}{\partial u} & \frac{\partial y}{\partial v}\end{vmatrix}$稱爲 Jacobian，該 Jacobian 也可以寫成$\begin{vmatrix}\frac{\partial x}{\partial u} & \frac{\partial y}{\partial u}\\ \frac{\partial x}{\partial v} & \frac{\partial y}{\partial v}\end{vmatrix}$。

以下是多重變數變換的例題，請特別注意積分界限的變換：

$$\int_0^1\int_0^1(x+y)(x-y)^2\,dxdy=?$$

令 $u=x+y,\ v=x-y,$ 即 $x=\dfrac{u+v}{2},\ y=\dfrac{u-v}{2}$

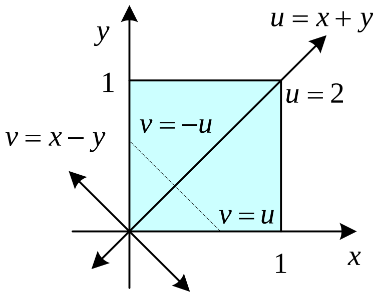
則 $\begin{vmatrix}\frac{\partial x}{\partial u} & \frac{\partial x}{\partial v}\\ \frac{\partial y}{\partial u} & \frac{\partial y}{\partial v}\end{vmatrix}=\begin{vmatrix}\frac{1}{2} & \frac{1}{2}\\ \frac{1}{2} & \frac{-1}{2}\end{vmatrix}=-\frac{1}{2}\ \Rightarrow\ dxdy=\left|-\frac{1}{2}\right|dudv=\frac{1}{2}dudv$

$$\Rightarrow\int_0^1\int_0^1(x+y)(x-y)^2\,dxdy=\int_0^1\int_{-u}^{u}\left(\tfrac{1}{2}uv^2\right)dvdu+\int_1^2\int_{-(2-u)}^{(2-u)}\left(\tfrac{1}{2}uv^2\right)dvdu$$

$$=\frac{1}{2}\int_0^1 u\left(\tfrac{1}{3}v^3\Big|_{-u}^{u}\right)du+\frac{1}{2}\int_1^2 u\left(\tfrac{1}{3}v^3\Big|_{-(2-u)}^{(2-u)}\right)du$$

$$=\frac{1}{6}\int_0^1 2u^4du+\frac{1}{6}\int_1^2 2u(2-u)^3\,du=\frac{1}{3}\int_0^1 u^4du+\frac{2}{3}\int_1^2(2-u)^3\,du-\frac{1}{3}\int_1^2(2-u)^4\,du$$

$$=\frac{1}{3}\left(\tfrac{1}{5}u^5\right)\Big|_0^1+\frac{2}{3}\left(\tfrac{-1}{4}(2-u)^4\right)\Big|_1^2-\frac{1}{3}\left(\tfrac{-1}{5}(2-u)^5\right)\Big|_1^2=\frac{1}{3}\times\frac{1}{5}+\frac{2}{3}\times\frac{1}{4}-\frac{1}{3}\times\frac{1}{5}=\frac{1}{6}$$



以下例題用於計算標準常態分配的機率和是否爲 1：

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} e^{-\frac{x^2+y^2}{2}} dxdy = ?$$

令 $x = r\sin\theta,\ y = r\cos\theta$ 則 $J = \begin{vmatrix} \frac{\partial x}{\partial r} & \frac{\partial y}{\partial r} \\ \frac{\partial x}{\partial \theta} & \frac{\partial y}{\partial \theta} \end{vmatrix} = \begin{vmatrix} \sin\theta & \cos\theta \\ r\cos\theta & -r\sin\theta \end{vmatrix} = -r,\ dxdy = |J|\,drd\theta = rdrd\theta$

$$\Rightarrow \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} e^{-\frac{x^2+y^2}{2}} dxdy = \int_0^{\infty}\int_0^{2\pi} e^{-\frac{r^2}{2}} r d\theta dr = \int_0^{\infty} re^{-\frac{r^2}{2}}\left(\int_0^{2\pi} d\theta\right)dr = 2\pi\int_0^{\infty} re^{-\frac{r^2}{2}} dr = 2\pi$$

五、微積分在統計上的應用

微分在初等統計上的應用，主要有兩個地方：(1)將累計機率函數轉換成機率密度函數或機率函數；(2)產生母體參數之估計量，包括最大概似估計法與動差法。積分的應用主要在機率的計算，包括條件機率、期望值、變異數、共變數、與相關係數等等；另外，積分技術也用於尋找抽樣分配的機率密度函數。

累計機率函數與機率密度函數

令 $f(x)$ 、 $F(x)$ 分別爲隨機變數 X 的機率密度函數與累計機率函數，則

$$F(X) = P(x \le X) = \int_{-\infty}^{X} f(x)dx$$
$$f(x) = \frac{d}{dx}\left[F(x)\right]$$

指數分配的機率密度函數爲

$$f(x) = \lambda e^{-\lambda x}, \quad x > 0$$

則其累計機率函數爲

$$F(X) = P(x \le X) = \int_0^X \lambda e^{-\lambda x} dx = -e^{-\lambda x}\Big|_0^X = 1 - e^{-\lambda X}, \quad X > 0$$

某隨機變數 Y 的累計機率函數爲

$$F(y)=\begin{cases}0, & y<0\\ y^2, & 0\le y\le 1\\ 1, & y>1\end{cases}$$

則其機率密度函數爲

$$f(y)=\frac{d}{dy}F(y)=\begin{cases}\dfrac{d}{dy}0=0, & y<0\\ \dfrac{d}{dy}y^2=2y, & 0\le y<1\\ \dfrac{d}{dy}1=0, & y>1\end{cases}$$

請注意 $f(y)$ 在 $y=1$ 的位置爲不連續，因爲 $F(y)$ 在該點並不平滑。

期望值、變異數、與動差母函數

令 $f(x)$ 爲隨機變數 X 的機率密度函，則

$$\int_{-\infty}^{\infty} f(x)dx=1$$
$$E(X)=\int_{-\infty}^{\infty} xf(x)dx$$
$$E(X^2)=\int_{-\infty}^{\infty} x^2 f(x)dx$$
$$\text{var}(X)=E(X^2)-\left[E(X)\right]^2=\int_{-\infty}^{\infty} x^2 f(x)dx-\left[\int_{-\infty}^{\infty} xf(x)dx\right]^2$$
$$m(t)=E\left(e^{tx}\right)=\int_{-\infty}^{\infty} e^{tx} f(x)dx$$

假設隨機變數 X 的機率密度函如下

$$f(x)=\begin{cases}kx, & 0\le x<2\\ k(4-x), & 2\le x\le 4\\ 0, & \text{otherwise}\end{cases}$$

先求 k 值如下

$$\int f(x)dx=1 \quad\Rightarrow\quad \int_0^2 kxdx+\int_2^4 k(4-x)dx=4k=1 \quad\Rightarrow\quad k=\frac{1}{4}$$

則

$$E(X)=\int xf(x)dx=\int_0^2 \frac{1}{4}x^2dx+\int_2^4 \frac{1}{4}x(4-x)dx=\left.\frac{x^3}{12}\right|_0^2+\left.\left(\frac{x^2}{2}-\frac{x^3}{12}\right)\right|_2^4=2$$

$$E(X^2)=\int x^2 f(x)dx=\int_0^2 \frac{1}{4}x^3dx+\int_2^4 \frac{1}{4}x^2(4-x)dx=\left.\frac{x^4}{16}\right|_0^2+\left.\left(\frac{x^3}{3}-\frac{x^4}{16}\right)\right|_2^4=\frac{14}{3}$$

$$\operatorname{var}(X)=E(X^2)-\left[E(X)\right]^2=\frac{14}{3}-(2)^2=\frac{2}{3}$$

$$m(t)=E\left(e^{tx}\right)=\int_{-\infty}^{\infty} e^{tx}f(x)dx$$

$$=\int_0^2 \frac{1}{4}xe^{tx}dx+\int_2^4 \frac{1}{4}(4-x)e^{tx}dx=\frac{1+(-1+2t)e^{2t}}{4t^2}+\frac{\left(-1+e^{2t}-2t\right)e^{2t}}{4t^2}=\frac{\left(-1+e^{2t}\right)^2}{4t^2}$$

條件機率、共變數、與相關係數

令 $f(x,y)$ 爲隨機變數 X 、 Y 的聯合機率密度函，則

$$f(x)=\int_{-\infty}^{\infty} f(x,y)dy,\qquad f(y)=\int_{-\infty}^{\infty} f(x,y)dx$$

$$f(x|Y=y)=\frac{f(x,y)}{f(y)},\qquad f(y|X=x)=\frac{f(x,y)}{f(x)}$$

$$E(X|Y=y)=\int_{-\infty}^{\infty} xf(x|Y=y)dx,\qquad E(Y|X=x)=\int_{-\infty}^{\infty} yf(y|X=x)dy$$

$$E(XY)=\int_{x=-\infty}^{\infty}\int_{y=-\infty}^{\infty} xyf(x,y)dydx$$

$$\operatorname{cov}(X,Y)=E(XY)-E(X)E(Y)$$

$$\rho=\frac{\operatorname{cov}(X,Y)}{\sqrt{\operatorname{var}(X)}\sqrt{\operatorname{var}(Y)}}=\frac{E(XY)-E(X)E(Y)}{\sqrt{E(X^2)-\left[E(X)\right]^2}\sqrt{E(Y^2)-\left[E(Y)\right]^2}}$$

【例題】假設有以下聯合機率密度函數

$$f(x,y)=x+y,\quad 0<x<1, 0<y<1$$

則計算其相關係數的過程如下

$$E(X)=\int_{y=0}^{1}\int_{x=0}^{1} x(x+y)dxdy=\frac{7}{12},\quad E(Y)=E(X)=\frac{7}{12}$$

$$E(X^2)=\int_{y=0}^{1}\int_{x=0}^{1} x^2(x+y)dxdy=\frac{5}{12},\quad E(Y^2)=E(X^2)=\frac{5}{12}$$

$$E(XY)=\int_{y=0}^{1}\int_{x=0}^{1} xy(x+y)dxdy=\frac{1}{3}$$

$$\text{var}(X)=E(X^2)-[E(X)]^2=\frac{5}{12}-\frac{7}{12}\times\frac{7}{12}=\frac{11}{144}$$
$$\text{cov}(X,Y)=E(XY)-E(X)E(Y)=\frac{1}{3}-\frac{7}{12}\times\frac{7}{12}=-\frac{1}{144}$$
$$\rho=\frac{\text{cov}(X,Y)}{\sqrt{\text{var}(X)}\sqrt{\text{var}(Y)}}=\frac{-1/144}{\sqrt{11/144}\sqrt{11/144}}=-\frac{1}{11}$$

邊際機率與條件機率的計算如下

$$f(y)=\int_{x=0}^{1}(x+y)\,dx=\frac{1}{2}+y$$
$$f(x|Y=y)=\frac{f(x,y)}{f(y)}=\frac{x+y}{\frac{1}{2}+y}$$
$$E(X|Y=y)=\int_0^1 x\frac{x+y}{\frac{1}{2}+y}dx=\frac{\frac{1}{3}+\frac{1}{2}y}{\frac{1}{2}+y}=\frac{2+3y}{3+6y}$$
$$E(X)=\int E(X|Y=y)f(y)\,dy=\int_0^1\frac{\frac{1}{3}+\frac{1}{2}y}{\frac{1}{2}+y}\times\left(\frac{1}{2}+y\right)dy=\int_0^1\left(\frac{1}{3}+\frac{1}{2}y\right)dx=\frac{1}{3}+\frac{1}{4}=\frac{7}{12}$$

從動差母函數求期望值、變異數

令 $m(t)$ 為隨機變數 X 的動差母函數，則

$$E(X)=\frac{d}{dt}m(t=0)$$
$$E(X^2)=\frac{d^2}{dt^2}m(t=0)$$

【**例題**】Gamma 分配的動差母函數為

$$m(t)=(1-\beta t)^{-\alpha}$$

則

$$\frac{d}{dt}m(t)=\frac{d}{dt}(1-\beta t)^{-\alpha}=\alpha\beta(1-\beta t)^{-\alpha-1}$$
$$\frac{d^2}{dt^2}m(t)=\frac{d}{dt}\alpha\beta(1-\beta t)^{-\alpha-1}=\alpha(\alpha+1)\beta^2(1-\beta t)^{-\alpha-2}$$
$$E(X)=\frac{d}{dt}m(t=0)=\alpha\beta$$

$$E\left(X^2\right)=\frac{d^2}{dt^2}m\left(t=0\right)=\alpha\left(\alpha+1\right)\beta^2$$
$$\operatorname{var}\left(X\right)=E\left(X^2\right)-\left[E\left(X\right)\right]^2=\alpha\left(\alpha+1\right)\beta^2-\left(\alpha\beta\right)^2=\alpha\beta^2$$

以最大概似法找估計量

若隨機變數 X 爲平均數 μ 、標準差 σ 之常態分配，則其機率密度函數爲

$$f\left(x\right)=\tfrac{1}{\sigma\sqrt{2\pi}}e^{-\frac{1}{2}\left(\frac{x_1-\mu}{\sigma}\right)^2}$$

令 $x_1,x_2,\ldots,x_n$ 爲一組相互獨立，接來自上述常態分配的樣本，則其概似函數爲

$$L\left(\mu,\sigma\right)=f\left(x_1\right)\cdots f\left(x_n\right)=\left[\tfrac{1}{\sigma\sqrt{2\pi}}e^{-\frac{1}{2}\left(\frac{x_1-\mu}{\sigma}\right)^2}\right]\cdots\left[\tfrac{1}{\sigma\sqrt{2\pi}}e^{-\frac{1}{2}\left(\frac{x_n-\mu}{\sigma}\right)^2}\right]=\left(\tfrac{1}{\sigma^2 2\pi}\right)^{\frac{n}{2}}e^{-\frac{1}{2\sigma^2}\Sigma\left(x_i-\mu\right)^2}$$

則以下是求最大概似估計量的過程：

$$L^*=\ln L\left(\mu,\sigma\right)=-\frac{n}{2}\ln\sigma^2-\frac{n}{2}\ln 2\pi-\frac{1}{2\sigma^2}\Sigma\left(x_i-\mu\right)^2$$
$$\begin{cases}\dfrac{d}{d\mu}L^*=\dfrac{1}{2\sigma^2}\Sigma\left(x_i-\mu\right)^2=0\\ \dfrac{d}{d\sigma^2}L^*=\dfrac{n}{2}\dfrac{1}{\sigma^2}+\dfrac{1}{2\sigma^2}\Sigma\left(x_i-\mu\right)^2=0\end{cases}\Rightarrow\begin{cases}\hat{\mu}=\dfrac{1}{n}\Sigma x_i=\overline{x}\\ \hat{\sigma}^2=\dfrac{1}{n}\Sigma\left(x_i-\overline{x}\right)^2\end{cases}$$

即 $\hat{\mu}=\frac{1}{n}\Sigma x_i=\overline{x}$ 、 $\hat{\sigma}^2=\frac{1}{n}\Sigma\left(x_i-\overline{x}\right)^2$ 分別爲 μ 、 σ^2 的最大概似估計量。

以動差法找估計量

若隨機變數 X 爲平均數 μ 、標準差 σ 之常態分配，則動差母函數爲

$$m\left(t\right)=\exp\left(\mu t+\frac{\sigma^2t^2}{2}\right)$$

第一、二級動差分別爲

$$E\left(X\right)=\frac{d}{dt}m\left(t=0\right)=\left(\mu+\sigma^2t\right)\exp\left(\mu t+\frac{\sigma^2t^2}{2}\right)\Bigg|_{t=0}=\mu$$

$$E\left(X^2\right)=\frac{d^2}{dt^2}m\left(t=0\right)=\left[\sigma^2+\left(\mu+\sigma^2 t\right)^2\right]\exp\left(\mu t+\frac{\sigma^2 t^2}{2}\right)\Bigg|_{t=0}=\sigma^2+\mu^2$$

令 $x_1,x_2,\ldots,x_n$ 爲一組相互獨立，接來自上述常態分配的樣本，則樣本第一、二級動差分別爲母體第一、二級動差：

$$\begin{cases}\dfrac{x_1+x_2+\cdots+x_n}{n}=\mu\\ \dfrac{x_1^2+x_2^2+\cdots+x_n^2}{n}=\sigma^2+\mu^2\end{cases}$$

上述聯立方程式的解即爲參數估計量 $\hat{\mu}$ 、 $\hat{\sigma}^2$ ；

$$\begin{cases}\hat{\mu}=\dfrac{x_1+x_2+\cdots+x_n}{n}=\overline{x}\\ \hat{\sigma}^2=\dfrac{x_1^2+x_2^2+\cdots+x_n^2}{n}-\overline{x}^2=\dfrac{\left(x_1-\overline{x}\right)^2+\left(x_2-\overline{x}\right)^2+\cdots+\left(x_n-\overline{x}\right)^2}{n}\end{cases}$$

透過累計機率函數找新分配的機率密度函數

若 $f_{X,Y}\left(x,y\right)$ 爲隨機變數 X 、 Y 的聯合機率密度函數，且令 $Z=X+Y$ ，則 Z 的累計機率函數爲

$$\begin{aligned}F_Z\left(z\right)&=P\left(Z\le z\right)=P\left(X+Y\le z\right)\\&=\iint\limits_{x+y\le z}f_{X,Y}\left(x,y\right)dxdy=\int_{-\infty}^{\infty}\left[\int_{-\infty}^{z-x}f_{X,Y}\left(x,y\right)dy\right]dx=\int_{-\infty}^{\infty}\left[\int_{-\infty}^{u}f_{X,Y}\left(x,u-x\right)du\right]dx\end{aligned}$$

其中最後一個步驟爲令 $y=u-x$ 作變數變換的結果。因此

$$\begin{aligned}f_Z\left(z\right)&=\frac{d}{dz}F_Z\left(z\right)=\frac{d}{dz}\int_{-\infty}^{\infty}\left[\int_{-\infty}^{u}f_{X,Y}\left(x,u-x\right)du\right]dx\\&=\frac{d}{dz}\int_{-\infty}^{u}\left[\int_{-\infty}^{\infty}f_{X,Y}\left(x,u-z\right)dx\right]du=\int_{-\infty}^{\infty}f_{X,Y}\left(x,z-x\right)dx\end{aligned}$$

亦即

$$\boxed{f_{Z=X+Y}\left(z\right)=\int_{-\infty}^{\infty}f_{X,Y}\left(x,z-x\right)dx=\int_{-\infty}^{\infty}f_{X,Y}\left(z-y,y\right)dy}$$

【**例題**】若隨機變數 X 、Y 的聯合機率密度函數為

$$f(x,y)=8xy, \qquad 0\le x\le 1, 0\le y\le x$$

則 $Z=X+Y$ 之機率密度函數如下

$$\begin{aligned} f_Z(z) &= \int_{-\infty}^{\infty} f_{X,Y}(x,z-x)\,dx \\ &= \int_{\frac{1}{2}z}^{1} 8x(z-x)\,dx = 4z\left(1-\tfrac{1}{4}z^2\right)-\tfrac{8}{3}\left(1-\tfrac{1}{8}z^3\right) = -\tfrac{8}{3}+4z-\tfrac{2}{3}z^3, \qquad 0\le z\le 2 \end{aligned}$$

其中，$y=z-x\le x \Rightarrow x\ge \frac{1}{2}z$ 。

若 $f_X(x)$、$f_Y(y)$ 分別為獨立隨機變數 X 、Y 的機率密度函數，且令 $Z=X+Y$ ，則 Z 的累計機率函數為

$$F_Z(z)=P(Z\le z)=P(X+Y\le z)=\int_{-\infty}^{\infty} P(x+Y\le z)\, f_X(x)\,dx=\int_{-\infty}^{\infty} F(z-x)\, f_X(x)\,dx$$

因此

$$\begin{aligned} f_Z(z) &= \frac{d}{dz}F_Z(z)=\frac{d}{dz}\int_{-\infty}^{\infty} F(z-x)\, f_X(x)\,dx \\ &= \int_{-\infty}^{\infty}\left[\frac{d}{dz}F(z-x)\right] f_X(x)\,dx=\int_{-\infty}^{\infty} f_Y(z-x)\, f_X(x)\,dx \end{aligned}$$

同理可得以下結果

$$\boxed{f_{Z=X+Y}(z)=\int_{-\infty}^{\infty} f_Y(z-x)\, f_X(x)\,dx=\int_{-\infty}^{\infty} f_X(z-y)\, f_Y(y)\,dy}$$

我們可以這樣理解，X 、Y 獨立意味著

$$f_{X,Y}(x,y)=f_X(x)\,f_Y(y)$$

因此

$$f_{Z=X+Y}(z)=\int_{-\infty}^{\infty} f_{X,Y}(x,z-x)\,dx=\int_{-\infty}^{\infty} f_Y(z-x)\,f_X(x)\,dx$$
$$f_{Z=X+Y}(z)=\int_{-\infty}^{\infty} f_{X,Y}(z-y,y)\,dy=\int_{-\infty}^{\infty} f_X(z-y)\,f_Y(y)\,dy$$

【例題】隨機變數 X 、Y 皆為參數λ的指數分配且互相獨立：

$$f_X(x)=\lambda e^{-\lambda x},\quad x>0, \qquad\qquad f_Y(y)=\lambda e^{-\lambda y},\quad y>0$$

則 $Z=X+Y$ 之機率密度函數如下

$$\begin{aligned} f_{Z=X+Y}(z) &= \int_{-\infty}^{\infty} f_Y(z-x)\,f_X(x)\,dx \\ &= \int_0^z \lambda e^{-\lambda(z-x)}\lambda e^{-\lambda x}\,dx=\int_0^z \lambda^2 e^{-\lambda z}\,dx=\lambda^2 z e^{-\lambda z}, \qquad z>0 \end{aligned}$$

其中， $y=z-x>0 \Rightarrow x<z$ 。