

1. 某母群體(population)裏有四個單位 (elementary units). 這些單位的某特徵值 X 分別是 1, 2, 5, 10. 假設 X 的平均值是我們要估計的母群體參數. 假使採樣的樣本數(sample size)是一個單位. 在單位採樣的機率為表一的情況下
- (a)請您建議一個無偏差的估計值 (an unbiased estimator). (5 分)
- (b)在單位採樣的機率為表二的情況下, 請您建議一個無偏差的估計值. (10 分)
- (c)請您討論(a)和(b)的結果. (5 分)

表一: 各單位的採樣機率.

x	1	2	5	10
單位採樣機率	0.25	0.25	0.25	0.25

表二: 各單位的採樣機率.

x	1	2	5	10
單位採樣機率	0.40	0.10	0.20	0.30

2. 假設某母群體的特徵值 X 其平均值是 0. 表三所列乃從母群體重覆抽樣 30 次 (simple random sampling with replacement), 每次樣本數大小是 10, 所得到的 30 個對 X 的平均值所做的 95%信賴區間 (confidence intervals).
- (a) 請問在這 30 次重覆抽樣中實際得到的信心水準(confidence level)是多少? (10 分)
- (b) 請由(a)的結果解釋 95%信賴區間的含意. (10 分)

(背面仍有題目,請繼續作答)

表三: 95%信賴區間

樣本平均數	下限	上限
-0.014	-1.034	0.206
0.191	-0.429	0.811
-0.227	-0.847	0.393
-0.1	-0.720	0.520
0.117	-0.503	0.737
0.23	-0.389	0.850
0.01	-0.610	0.630
0.087	-0.533	0.707
-0.622	-1.242	-0.002
0.056	-0.564	0.676
-0.251	-0.871	0.369
0.097	-0.523	0.727
-0.391	-1.011	0.229
-0.272	-0.892	0.348
-0.224	-0.844	0.396
-0.64	-1.260	-0.020
0.091	-0.529	0.711
0.165	-0.455	0.785
-0.247	-0.867	0.373
-0.131	-0.751	0.489
-0.135	-0.755	0.485
-0.161	-0.781	0.459
-0.063	-0.683	0.557
0.164	-0.456	0.784
0.052	-0.568	0.672
0.09	-0.530	0.710
0.071	-0.549	0.691
0.404	-0.216	1.024
-0.224	-0.844	0.396
-0.15	-0.770	0.470

3. 假設我們要檢定母群體的某特徵值的平均值(μ)是否為 0 (null hypothesis)

或者是為 10^{-10} (alternative hypothesis). 如果採樣數的大小是 10^4 ,

(a) 請問樣本平均數大於多少的情況下, 你會拒絕 null hypothesis at

significance level 0.05 (i.e. type I error probability of 5%)? (5分)

標準差 (σ) = 1

提示: (I) 請用中間極限定理 (central limit theorem).

(II) The upper 5 percentile of a standard normal distribution is 1.645.

(b) 請問(a)的檢定規則 (decision rule) statistical power 是多少? (5分)

(c) 請根據(b)的答案討論假設檢定(hypothesis testing)在採樣數很大的情況下所產生的問題. (10分)

表四: 危險因子 A 和 B 對某疾病發生的危險比.

	B1	B2
A1	1	4
A2	4	7

表五: 危險因子 A 和 B 對某疾病發生的自然對數危險比.

	B1	B2
A1	0	1.39
A2	1.39	1.95

4. 假設兩危險因子 A 和 B 對某疾病發生的危險比(=得病機率/沒得病機率)

影響如表四. A 和 B 各有兩個 levels, (A1, A2) 和 (B1, B2).

(a) 請問 A 和 B 對疾病發生的危險比影響是否存在交互作用 (interaction).

也就是說在不同 level of A 下, B1 and B2 的影響是否不同? (5分)

(b) 將危險比取自然對數(natural logarithm) (見表五). 請問在自然對數

scale 上, A 和 B 是否存在交互作用 (interaction)? (5分)

(c) 請解釋(a)和 (b)答案的差異在資料分析上所隱含的意義. (10分)

5. 表六和表七各別為變數 X1-X5 和 Y1-Y5 之相關係數矩陣(correlation matrix).

(a) 請問何者適宜做因素分析 (factor analysis). 也就是說那一組變數彼此

相關性可用較少的共同因子(constructs)來描述? 為什麼? (5 分)

(b) 請問適宜做因素分析者, 可找出幾個共同因子? 為什麼? (5 分)

表六: 變數 X1-X5 之相關係數矩陣.

	X1	X2	X3	X4	X5
X1	1				
X2	0.10	1			
X3	0.97	0.15	1		
X4	0.44	0.69	0.53	1	
X5	0.02	0.86	0.12	0.78	1

表七: 變數 Y1-Y5 之相關係數矩陣.

	Y1	Y2	Y3	Y4	Y5
Y1	1				
Y2	0.00	1			
Y3	0.00	0.00	1		
Y4	0.00	0.00	0.00	1	
Y5	0.00	0.00	0.00	0.00	1

6. 當比較兩組(two parallel groups)平均值是否相等, 很自然地是看兩組平均值的差是否為 0 (例如 two-sample student-*t* test). 當比較兩組以上平均值是否有差異, 請問如何檢定所有組別平均值是否相等? 請針對 the *F*-test for one-way ANOVA 和 Tukey's honestly significant difference (HSD) test 所用之方法發表您的看法. (10 分)