

科目：統計學

系所組：統計資訊學系應用統計碩士班

1. 統計觀念解釋。

- (1) 以生活實例說明信賴度與信賴區間 (5 分)
- (2) 以生活實例說明型一誤差與型二誤差 (5 分)
- (3) 舉例說明大數法則常見的誤用 (5 分)
- (4) 無母數統計相較於母數統計適用時機 (5 分)

2. 從經驗中得知某郵局櫃台讓民眾等候服務的時間為均勻分配 $U(0, \theta)$ 。現隨機抽樣六位樣本，等候時間分別為 45 秒，1 分鐘，2 分鐘，3 分 15 秒，6 分鐘與 8 分鐘。請分別以動差法 (method of moments) 與最大概似法 (maximum likelihood method) 來估計參數 θ 的值 (10 分)；若以均方誤 (Mean Square Error) 的角度，您將會採用那一種估計法？(6 分)3. 某民調公司針對 1,200 位 20 歲以上民眾進行訪問，並發布兩黨民意調查結果：甲候選人為 38% 支持度，乙候選人為 34% 支持度，尚未決定為 28%。若考慮支持甲候選人的人數為 X ，支持乙候選人的人數為 Y ，則隨機變數 (X, Y) 很明顯無法視為二項分配，應採用三項分配 (Trinomial Distribution) 進行分析。

- (1) 請問此三項分配的參數 (parameter) 有哪些？寫出機率函數。(4 分)
- (2) 利用三項分配的共變異數 $Cov(X, Y) = -np_1p_2$ ，計算出樣本比例差的變異數 $Var(\hat{p}_1 - \hat{p}_2)$ 。(8 分)
- (3) 利用上述結果，列出甲候選人和乙候選人支持度之間差異的 95% 信賴區間 (列出算式即可)，並說明如何透過此區間判斷兩位候選人支持度是否有顯著的差異。(8 分)

4. 某線上遊戲設計每日可 4 連抽，假定 X 隨機變數為每位玩家可抽到稀有卡牌數且其分配是一個成功機率為 p 的二項分配 (binomial distribution)，其中 X 的可能值為 0, 1, 2, 3, 4。針對遊戲公司宣稱有 50% 的抽中率，考慮以下假設 $H_0: p = 0.5$ vs $H_1: p = 0.25$ ，給定以下 4 種檢定法：檢定法 A：如果 $X = 4$ ，則拒絕 H_0 ，反之則不拒絕 H_0 ，即拒絕域 (rejection region) 為集合 $\{4\}$ ；檢定法 B：如果 $X = 0$ ，則拒絕 H_0 ，反之則不拒絕 H_0 ，即拒絕域為集合 $\{0\}$ ；檢定法 C：如果 $X \leq 1$ ，則拒絕 H_0 ，反之則不拒絕 H_0 ，即拒絕域為集合 $\{0, 1\}$ ；檢定法 D：如果 $X = 0$ 或 $X = 4$ ，則拒絕 H_0 ，反之則不拒絕 H_0 ，即拒絕域為集合 $\{0, 4\}$ 。

二項分配機率表：

	0	1	2	3	4
$p = 0.25$	0.3164	0.4219	0.2109	0.0469	0.0039
$p = 0.5$	0.0625	0.2500	0.3750	0.2500	0.0625

設定顯著水準為 $\alpha = 0.0625$ ，請決定上述 4 種檢定法：

- (1) 那一個或那些檢定法是符合 α 之設定？(8 分)
- (2) 在符合設定下，哪個檢定有最大的檢定力？(8 分)

※注意：1. 考生須在「彌封答案卷」上作答。

2. 本試題紙空白部分可當稿紙使用，試題須隨答案卷繳回。

3. 考生於作答時可否使用計算機、法典、字典或其他資料或工具，以簡章之規定為準。

科目：統計學

系所組：統計資訊學系應用統計碩士班

5. 一位統計分析師受託分析33位銷售員，其週業績（Sales）和每週工作時數（Work）和年資（Experience）的關係。所建立的兩個分析結果如下：

【分析一】

Analysis of Variance

Source	Df	Sum of Squares	Mean Square	F-value	Pr > F
Model	1	95.234	95.234	0.911	0.347
Error	31	3241.312	104.558		
Total	32	3336.545			

Parameter Estimates

Variable	Df	Parameter Estimate	Standard Error	T-value	Pr > t
Intercept	1	144.775	8.520	16.992	< .001
Work	1	0.174	0.182	0.956	0.347

【分析二】

Analysis of Variance

Source	Df	Sum of Squares	Mean Square	F-value	Pr > F
Model	2	2186.595	1093.297	28.522	< .001
Error	30	1149.951	38.332		
Total	32	3336.545			

Parameter Estimates

Variable	Df	Parameter Estimate	Standard Error	T-value	Pr > t
Intercept	1	104.052	7.550	13.782	< .001
Work	1	0.435	0.116	3.750	0.001
Experience	1	1.573	0.213	7.385	< .001

- (1) 寫出分析一與分析二的迴歸方程式，並且計算 R^2 。(6分)
- (2) 比較分析一與分析二每週工作時數與業績之關係是否有矛盾？你會採用那個模式？(4分)
- (3) 將分析二之年資的單位由年改為月，估計值於標準誤是否會改變？並求改變後之值。(4分)
- (4) 除了每週工作時數(X_1 , Work)和年資(X_2 , Experience)以外，統計分析師也加入年齡(X_3)與性別(X_4)變項進行複迴歸分析，分析之前先分別以其中一個為因變數，而以其他三個為自變數時，所得到的判定係數 R^2 。請分別計算其變異數膨脹因子(Variance Inflation Factor, VIF)。並且可根據VIF值得到甚麼訊息？模型是否需要修正？(5分)

因變數	判定係數 R^2
X_1	0.06
X_2	0.94
X_3	0.24
X_4	0.56

※注意：1.考生須在「彌封答案卷」上作答。

2.本試題紙空白部分可當稿紙使用，試題須隨答案卷繳回。

3.考生於作答時可否使用計算機、法典、字典或其他資料或工具，以簡章之規定為準。

科目：統計學

系所組：統計資訊學系應用統計碩士班

- (5) 統計分析師進一步想進行模型檢驗，表1的第一欄是觀察值的順序，第二欄是殘差值，第三欄是標準化的殘差值，請以此判斷是否有異常點存在？請說明判斷準則。(3分) 表1第五欄是 Student化刪除殘差，第i個 R-Student殘差是在假定將資料中的第i個觀察值刪除，然後以剩下的n-1個觀察值來建立新的估計迴歸方程式而標準化獲得的 R-Student殘差值。請以此判斷是否有異常點存在？請說明判斷準則。(3分)。表1的最後一欄，第八欄是 DFFITS (Difference in Fits) 值。請以此判斷是否有可能的影響點存在？請說明判斷準則。(3分)

表1

Obs	Residual	Standardized Residual	Cook's D	R-Student	Hat Diag	Cov Ratio	DFFITS
1	0.521	0.423	-0.009	0.418	0.081	1.153	0.127
2	0.886	0.698	-0.011	0.692	0.015	1.050	0.116
3	-1.797	-1.430	0.021	-1.454	0.02	0.954	-0.313
4	0.370	0.288	-0.018	0.284	0.013	1.077	0.036
5	-0.997	-0.792	-0.01	-0.787	0.012	1.041	-0.160
6	-1.530	-1.214	0.004	-1.223	0.013	0.985	-0.242
7	1.470	1.159	0.004	1.166	0.013	0.987	0.199
8	-0.180	-0.147	-0.02	-0.145	0.025	1.097	-0.047
9	1.003	0.789	-0.009	0.784	0.012	1.037	0.127
10	-0.530	-0.423	-0.017	-0.418	0.013	1.073	-0.093
11	1.937	1.543	0.047	1.582	0.032	0.936	0.354
12	0.137	0.105	-0.02	0.103	0.032	1.105	0.009
13	1.003	0.789	-0.009	0.784	0.012	1.037	0.127
14	0.536	0.424	-0.015	0.418	0.030	1.090	0.081
15	-0.797	-0.634	-0.014	-0.628	0.012	1.057	-0.131
16	-0.263	-0.212	-0.019	-0.209	0.010	1.079	-0.054
17	-2.063	-1.652	0.053	-1.699	0.032	0.921	-0.412
18	-1.830	-1.478	0.058	-1.506	0.048	0.974	-0.422
19	0.120	0.090	-0.020	0.089	0.011	1.081	-0.001
20	2.336	1.924	0.216	2.021	0.091	0.906	0.701
21	0.603	0.502	0.002	0.496	0.121	1.201	0.189
22	0.470	0.368	-0.017	0.362	0.013	1.073	0.051
23	0.770	0.607	-0.012	0.600	0.019	1.063	0.105
24	-0.863	-0.717	0.011	-0.711	0.094	1.148	-0.268
25	-0.830	-0.661	-0.013	-0.655	0.013	1.055	-0.137
26	1.470	1.159	0.004	1.166	0.013	0.987	0.199
27	0.470	0.368	-0.017	0.362	0.013	1.073	0.051
28	-3.565	-3.457	3.254	-4.319	0.337	0.616	-3.220
29	0.069	0.053	-0.020	0.052	0.066	1.148	0.002
30	0.237	0.183	-0.019	0.180	0.010	1.078	0.015
31	-1.297	-1.033	0.001	-1.034	0.020	1.020	-0.227
32	1.303	1.045	0.020	1.047	0.046	1.041	0.263
33	0.170	0.133	-0.019	0.131	0.048	1.124	0.021

※注意：1.考生須在「彌封答案卷」上作答。

2.本試題紙空白部分可當稿紙使用，試題須隨答案卷繳回。

3.考生於作答時可否使用計算機、法典、字典或其他資料或工具，以簡章之規定為準。