

科目：統計學

系所組：統計資訊學系應用統計碩士班

1. 出國旅遊已成各國觀光的休閒安排之一。假設有一結果調查 500 位非亞洲人口和 500 位亞洲人口為樣本，會選擇的旅遊區域比例如表，試問在顯著水準 0.05 下，非亞洲旅遊者選擇在風景區與遊樂園區的比例差異區間。(四捨五入至小數點第二位)

	非亞洲人	亞洲人
風景區	0.52	0.40
遊樂園區	0.28	0.50
沒有限制	0.20	0.10

(a) 選擇風景區的比例的真正差異為何? (10%)

(b) 亞洲人選擇風景區與遊樂園區比例的真正差異為何? (10%)

請計算(a)與(b)比例真正差異的信賴區間。 $Z_{0.05} = 1.645, Z_{0.025} = 1.96$

2. 假設有兩個班級(甲、乙)修習統計學課程，甲班有 60 位學生，乙班有 40 位學生。令 μ_1 與 μ_2 分別是甲班和乙班的母體分數平均數， σ_1 與 σ_2 分別是甲班和乙班的母體分數標準差。現在有一學期分數其樣本平均數與樣本標準差分別為 $\bar{X}_1 = 84, S_1 = 4$ 與 $\bar{X}_2 = 77, S_2 = 6$ 。在顯著水準 0.1 下，

(a) 請建構適當的區間估計，判斷 $\mu_1 = \mu_2$ 是否成立?並做結論。(10%)

(b) 請建構適當的假設檢定，判斷 $\sigma_1^2 = \sigma_2^2$ 是否成立?並做結論。(10%)

$$F_{0.95,60,40} = 1.6373, F_{0.95,40,60} = 1.5943, F_{0.9,60,40} = 1.4672, F_{0.9,40,60} = 1.4373$$

3. 假設 X_1, X_2, \dots, X_n 為樣本數 n 且滿足以下機率密度函數的隨機變數，其機率密度函數

$$f(x) = \frac{1}{\theta} e^{-\frac{x}{\theta}}, x > 0, \theta > 0$$

(a) 請計算 X 的期望值(expected value)、變異數(variance)、動差母函數(moment generating function)? (9%)

(b) 假設 $\theta = 2$ ，請計算 $P(X > 5 | X > 2) = ?$ (10%)

(c) 請計算隨機變數 X 的最大概似估計值(maximum likelihood estimator) (10%)

(a、b、c 問題，均請列出計算過程，未列者不予計分)

4. 某校在有限資料下為探討大學生統計成績的影響因素，隨機蒐集 200 位學生並利用迴歸分析方法得到以下迴歸估計方程式為 $Y = 35.8 + 0.15X_1 + 0.72X_2 - 8.3D_1 + 5.3D_2 + 4.5D_3$ ，其中

Y = 大學生統計學成績

X_1 = 學生智力測驗分數

X_2 = 大學生微積分成績

$D_1 = 1$ 表示為男生

= 0 表示為女生

$D_2 = 1$ 表示為文、管理學院學生

= 0 表示為其他學院學生

$D_3 = 1$ 表示為理學院學生

= 0 表示為其他學院學生

※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。

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

迴歸系數的變異數共變數矩陣(variance-covariance matrix)為

	常數	X_1	X_2	D_1	D_2	D_3
常數	15.36	8.76	3.35	5.56	3.09	1.89
X_1		0.09	0.05	2.14	1.67	1.08
X_2			0.04	1.25	-2.35	1.45
D_1				4.25	3.35	-1.95
D_2					5.56	3.56
D_3						2.25

在顯著水準 0.05 下

- 進行迴歸分析需要滿足那三個基本條件? (6%)
- 就此迴歸方程式，那些因素會影響到大學生統計學成績。(10%)
- 各學院大學生的統計學成績有無顯著差異? (10%)
- 假設此迴歸分析結果判定係數為 0.65，其代表意義為何? (5%)

※注意：1.考生須在「彌封答案卷」上作答。

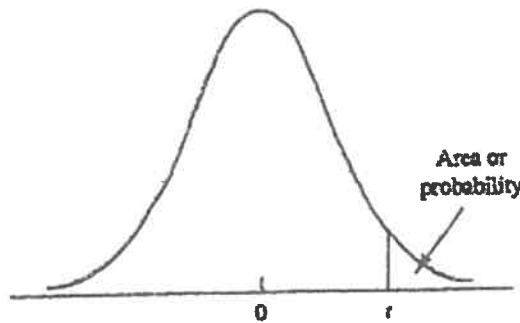
2.本試題紙空白部分可當稿紙使用。

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

t DISTRIBUTION



Entries in the table give t values for an area or probability in the upper tail of the t distribution. For example, with 10 degrees of freedom and a .05 area in the upper tail, $t_{.05} = 1.812$.

Degrees of Freedom	Area in Upper Tail				
	.10	.05	.025	.01	.005
1	3.078	6.314	12.706	31.821	63.657
2	1.886	2.920	4.303	6.965	9.925
3	1.638	2.353	3.182	4.541	5.841
4	1.533	2.132	2.776	3.747	4.604
5	1.476	2.015	2.571	3.365	4.032
6	1.440	1.943	2.447	3.143	3.707
7	1.415	1.895	2.365	2.998	3.499
8	1.397	1.860	2.306	2.896	3.355
9	1.383	1.833	2.262	2.821	3.250
10	1.372	1.812	2.228	2.764	3.169
11	1.363	1.796	2.201	2.718	3.106
12	1.356	1.782	2.179	2.681	3.055
13	1.350	1.771	2.160	2.650	3.012
14	1.345	1.761	2.145	2.624	2.977
15	1.341	1.753	2.131	2.602	2.947
16	1.337	1.746	2.120	2.583	2.921
17	1.333	1.740	2.110	2.567	2.898
18	1.330	1.734	2.101	2.552	2.878
19	1.328	1.729	2.093	2.539	2.861
20	1.325	1.725	2.086	2.528	2.845
21	1.323	1.721	2.080	2.518	2.831
22	1.321	1.717	2.074	2.508	2.819
23	1.319	1.714	2.069	2.500	2.807
24	1.318	1.711	2.064	2.492	2.797
25	1.316	1.708	2.060	2.485	2.787
26	1.315	1.706	2.056	2.479	2.779
27	1.314	1.703	2.052	2.473	2.771
28	1.313	1.701	2.048	2.467	2.763
29	1.311	1.699	2.045	2.462	2.756
30	1.310	1.697	2.042	2.457	2.750
40	1.303	1.694	2.021	2.423	2.704
60	1.296	1.671	2.003	2.390	2.660
120	1.289	1.658	1.989	2.358	2.617
∞	1.282	1.645	1.960	2.326	2.576

This table is reprinted by permission of Oxford University Press on behalf of The Biometrika Trustees from Table 12, Percentage Points of the t Distribution, by E. S. Pearson and H. O. Hartley, *Biometrika Tables for Statisticians*, Vol. 1, 3rd ed., 1966.

- ※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。
 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：計算機概論

系所組：統計資訊學系應用統計碩士班

一、請說明資料庫相關概念：(30%)

- (1) 何謂主鍵?
- (2) 何謂外鍵?
- (3) 為何需要正規化?
- (4) 何謂第一正規化(1NF)? 請舉例說明
- (5) 何謂第二正規化(2NF)? 請舉例說明
- (6) 何謂第三正規化(3NF)? 請舉例說明

二、請寫出下列名詞解釋(40%)

1. Domain name system
2. Blockchain
3. Decision tree
4. Clustering
5. Machine learning
6. P2P
7. O2O
8. Association rules
9. Artificial intelligence
10. SQL

三、請就您所了解的技術或實務案例儘可能地描述何謂大數據(Big Data)?(10%)

四、請使用您熟悉的任何一種程式語言,寫一個利用 For 迴圈加總 1~100 偶數總和的程式(10%)

五、請使用您熟悉的任何一種程式語言,利用 while 迴圈改寫上一題的程式(10%)

※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。

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：微積分

系所組：統計資訊學系應用統計碩士班

1. (本題共 20%)在時間序列分析中，其中一個被廣泛應用的技巧是 Z-transform (“Time Series Analysis” by James D. Hamilton; or “Time Series Analysis” by Henrik Madsen)；而 Z-transform 的計算基礎則是令 $z = e^{-j\omega}$ ， $j = \sqrt{-1}$ 來完成對時間序列的分析計算。若 $X(j\omega) = \int_{-\infty}^{\infty} x(t)e^{-j\omega t} dt$ ， $x(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} X(j\omega)e^{j\omega t} d\omega$ ，與 $e^{j\theta} = \cos\theta + j\sin\theta$ (尤拉公式，請把 θ 當成 ω 即可)，請問：
- A. (10%) $\frac{d}{dt} x(t)$ 與 $X(j\omega)$ 之間的關係是什麼？
- B. (10%) 若 $x(t) = \sin\left(\frac{\pi}{3}t - \frac{\pi}{2}\right)$ ，請問 $X(j\omega)$ 在 $\omega = \frac{\pi}{2}$ 與 $\omega = \frac{\pi}{3}$ 時的值分別是什麼？(需要兩個問題都選對才給分)
- (a) 0, (b) 1 或無窮大, (c) $\frac{\pi}{2}$, (d) $\frac{\pi}{3}$, (e) $\sin\left(\frac{\pi}{3}\right)$, (f) $\cos\left(\frac{\pi}{3}\right)$, (g) -1 或負無窮大,

2. (10%) 令 X 為 Cauchy random variable，並服從以下機率密度函數：

$$f(x) = \frac{1}{\pi(1+x^2)}, \quad -\infty < x < \infty.$$

請問 X 的期望值為何？

3. (本題共 20%)請畫出 $f(x) = \frac{54(x^2-1)}{x^2+27}$ 的圖形(10%)，並標記極值(5%)、反曲點(2%)、漸近線(3%)。

若無極值、反曲點、漸近線時，需請註明。

(以上為第一頁)

※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。

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：微積分

系所組：統計資訊學系應用統計碩士班

(以下為第二頁)

4. (本題共 40%)請計算以下各積分式的值。

A. (10%) $\int_0^2 (3x^2 + x - 4)dx$

B. (10%) $\int_2^3 \frac{1}{(x+1)^2} dx$

C. (10%) $\int_{-3}^{12} \sqrt{x+4} dx$

D. (10%) $\int_3^4 x^2 dx$

5. (10%) 請計算 $\lim_{x \rightarrow -1} \left(\frac{3x^3 - 3x^2 - 6x}{x^2 + x} \right)$ 的極限值。

(共 5 大題，以下空白)

※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。

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。