

[題 1]. (10 分)

我們想研究三種基因型民眾之「糖化血色素」值 (HbA1c)，於是自全台隨機抽取若干民眾，檢測其 HbA1c (單位：%)，資料如下表所示。如已知 HbA1c 不呈常態分布，請說明該如何檢定三組之「糖化血色素」值有無統計上之顯著差異，顯著水準設為 0.05。請先寫出虛無假說與對立假說，並寫出建構檢定統計量之想法概念，不需列出詳細公式。如有自設符號請定義。

基因型 AA	基因型 Aa	基因型 aa
6.3	5.3	5.6
5.4	6.2	5.6
4.6	6.3	5.5
5.6	5.6	5.7
6.9	5.6	5.7
5.6	5.8	5.7
5.9	6.0	5.4
6.9	6.2	5.7
6.4	5.9	6.0
6.0	5.4	5.4
6.9	5.1	
6.0	5.5	
5.4		
5.9		
6.1		

[題 2]. (20 分) 請回答以下四小題，每小題 5 分：

- (1) 請問何謂「最大概似估計值」(maximum likelihood estimate, 簡寫為 MLE)? 並舉實例說明之。
- (2) 何謂「中央極限定理」(central limit theorem)? 並舉實例說明之。
- (3) 何謂「統計 p 值」(p -value)? 請分別針對「雙尾檢定」與「單尾檢定」說明之。
- (4) 何謂「顯著水準」(significance level)? 與「統計 p 值」之差異為何?

[題 3]. (10 分)

請問是否所有的「最大概似估計值」皆為「不偏估計值」(unbiased estimate)，如是請證明，如否請舉反例並證明之。

[題 4]. (10 分)

某班有 10 人，期末考後發現僅有 2 人及格，請問及格比例 p 的 95% 信賴區間如何推導? 請由二項分布為基礎推導(可不用計算出最後結果)，由於僅有 10 人，本題不適合使用「中央極限定理」。

見背面

[題 5]. (50 分，3 小題)

為減低新冠肺炎重症(S)發生，某藥物臨床試驗針對 500 位輕症(M)病人進行新藥物(N)與一般標準治療(U)比較，經隨機分派至兩組(約 1:1 比例)，在療程後統計各組個案進展至重症(S)與維持在輕症(M)的人數，結果發現如下表一：

表一：組別與症狀結果 2x2 聯列表

		組別		合計
		N	U	
果 症 狀 結 果	S	27	50	77
	M	218	205	423
	合計	245	255	500

以下是做為下列統計分析假定、情境及相關資料。

- (a) 假設重症發生是稀少事件，
- (b) 效益以減少重症發生相對勝算比(Odds Ratio, OR)表示，
- (c) 假設 $\ln(OR)$ (相對勝算比的自然對數轉換)遵循常態分佈，其平均值及變異數估計值分別為

$$\hat{\theta} = \ln(\overline{OR}) = \ln \left[\frac{(a_1 + \frac{1}{2})(a_4 + \frac{1}{2})}{(a_2 + \frac{1}{2})(a_3 + \frac{1}{2})} \right]$$

$$Var(\hat{\theta}) = \frac{1}{(a_1 + \frac{1}{2})} + \frac{1}{(a_2 + \frac{1}{2})} + \frac{1}{(a_3 + \frac{1}{2})} + \frac{1}{(a_4 + \frac{1}{2})}$$

其中 a_1, a_2, a_3 及 a_4 是以上 2x2 聯列表中 4 個儲存格的數值，如下表二所示。

表二：2x2 聯列表儲存格

		組別		
		N	U	
果 重 症 結 果	結	S	a_1	a_2
	果	M	a_3	a_4

- (d) 標準常態分佈表見附表。

請回答以下三小題：

- (1) (15 分) 利用(a)-(c)計算表一 OR 之估計值及 95%信賴區間，並寫出統計檢定假說，依 P value 計算結果說明是否可以推翻虛無假說 ($\alpha=0.05$)。

接次頁

(2) (15 分) 若隨機變數 X 是來自此臨床試驗之母群體效益參數 θ [$\ln(\text{OR})$]，其分佈為常態分佈，表示為

$$X \sim N\left(\theta, \frac{\sigma^2}{n}\right)$$

在 H_0 (兩組間效益相等) 之下，假設標準差 $(\sigma)=2$ ，利用抽樣分配理論，令 k 為此樣本之估計值，以 (a) 及 (c) 證明 $\text{Var}(\hat{\theta}) \cong \frac{2}{a_1} \cong \frac{4}{k}$ ，且解釋 k 在此臨床試驗稀少疾病假設下之意義為何。再利用表一 2×2 聯列表資料，估計常態分佈參數 θ ，並在標準差 (σ) 已知為 2 之下，估計 k 之值。

(3) (20 分) 若過去先前研究也報導相同研究設計但不同受試者研究的結果，其效益 (OR) 的估計值為 0.60 (95% 信賴區間：0.21-1.00)，假設 $\ln(\text{OR})$ 也遵循常態分佈且表示為 $N\left(\mu_0, \frac{\sigma^2}{n_0}\right)$ (視為事前分佈)，同樣在標準差 (σ) 已知為 2 之下，根據此樣本資料所得到之平均值及標準誤 (由 95% 信賴區間推估) 求 n_0 之值，並寫出其事前常態分佈。再將題 (2) 所得分佈建立常態概似函數 (Likelihood)，利用貝氏常態分佈共軛對，依題 (2) 所得之 k 及上述 n_0 求其事後分佈。並依貝氏觀點寫出統計檢定式，並求效益 $\theta[\ln(\text{OR})]$ 小於 0 之機率是否大於 95% (依標準常態分佈， $\Phi(1.96)=0.95$)。

附表：標準常態分佈表

z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
3.9	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
3.8	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999
3.7	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999
3.6	0.9999	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998
3.5	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998
3.4	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997
3.3	0.9995	0.9995	0.9995	0.9996	0.9996	0.9996	0.9996	0.9996	0.9996	0.9997
3.2	0.9993	0.9993	0.9994	0.9994	0.9994	0.9994	0.9994	0.9995	0.9995	0.9995
3.1	0.9990	0.9991	0.9991	0.9991	0.9992	0.9992	0.9992	0.9992	0.9993	0.9993
3.0	0.9987	0.9987	0.9987	0.9988	0.9988	0.9988	0.9989	0.9989	0.9989	0.9990
2.9	0.9981	0.9982	0.9982	0.9983	0.9983	0.9984	0.9984	0.9985	0.9985	0.9986
2.8	0.9974	0.9975	0.9976	0.9977	0.9977	0.9978	0.9978	0.9979	0.9980	0.9981
2.7	0.9965	0.9966	0.9967	0.9968	0.9969	0.9970	0.9971	0.9972	0.9973	0.9974
2.6	0.9953	0.9955	0.9956	0.9957	0.9959	0.9960	0.9961	0.9962	0.9963	0.9964
2.5	0.9938	0.9940	0.9941	0.9943	0.9945	0.9946	0.9948	0.9949	0.9951	0.9952
2.4	0.9918	0.9920	0.9922	0.9925	0.9927	0.9929	0.9931	0.9932	0.9934	0.9936
2.3	0.9893	0.9896	0.9898	0.9901	0.9904	0.9906	0.9909	0.9911	0.9913	0.9916
2.2	0.9861	0.9864	0.9868	0.9871	0.9875	0.9878	0.9881	0.9884	0.9887	0.9890
2.1	0.9821	0.9826	0.9830	0.9834	0.9838	0.9842	0.9846	0.9850	0.9854	0.9857
2.0	0.9772	0.9778	0.9783	0.9788	0.9793	0.9798	0.9803	0.9808	0.9812	0.9817
1.9	0.9713	0.9719	0.9726	0.9732	0.9738	0.9744	0.9750	0.9756	0.9761	0.9767
1.8	0.9641	0.9649	0.9656	0.9664	0.9671	0.9678	0.9685	0.9693	0.9699	0.9706
1.7	0.9554	0.9564	0.9573	0.9582	0.9591	0.9599	0.9608	0.9616	0.9625	0.9633
1.6	0.9452	0.9463	0.9474	0.9486	0.9495	0.9505	0.9515	0.9525	0.9535	0.9545
1.5	0.9332	0.9345	0.9357	0.9370	0.9382	0.9394	0.9406	0.9418	0.9429	0.9441
1.4	0.9192	0.9207	0.9222	0.9236	0.9251	0.9265	0.9279	0.9292	0.9306	0.9319
1.3	0.9032	0.9049	0.9066	0.9082	0.9099	0.9115	0.9131	0.9147	0.9162	0.9177
1.2	0.8849	0.8869	0.8888	0.8907	0.8925	0.8944	0.8962	0.8980	0.8997	0.9015
1.1	0.8643	0.8665	0.8686	0.8708	0.8729	0.8749	0.8770	0.8790	0.8810	0.8830
1.0	0.8413	0.8438	0.8461	0.8485	0.8508	0.8531	0.8554	0.8577	0.8599	0.8621
0.9	0.8159	0.8186	0.8212	0.8238	0.8264	0.8289	0.8315	0.8340	0.8365	0.8389
0.8	0.7981	0.7910	0.7939	0.7967	0.7995	0.8023	0.8051	0.8078	0.8106	0.8133
0.7	0.7580	0.7611	0.7642	0.7673	0.7704	0.7734	0.7764	0.7794	0.7823	0.7852
0.6	0.7257	0.7291	0.7324	0.7357	0.7389	0.7422	0.7454	0.7486	0.7517	0.7549
0.5	0.6915	0.6950	0.6985	0.7019	0.7054	0.7088	0.7123	0.7157	0.7190	0.7224
0.4	0.6554	0.6591	0.6628	0.6664	0.6700	0.6736	0.6772	0.6808	0.6844	0.6879
0.3	0.6179	0.6217	0.6255	0.6293	0.6331	0.6368	0.6405	0.6443	0.6480	0.6517
0.2	0.5793	0.5832	0.5871	0.5910	0.5948	0.5987	0.6025	0.6064	0.6103	0.6141
0.1	0.5398	0.5438	0.5478	0.5517	0.5557	0.5596	0.5635	0.5675	0.5714	0.5753
0.0	0.5000	0.5040	0.5080	0.5120	0.5160	0.5199	0.5239	0.5279	0.5319	0.5359