1. 衡量資料時有哪些尺度可用來衡量資料的數值？

解

四種尺度：

①名目尺度衡量類別資料。

②順序尺度衡量順序資料。

③區間尺度衡量無絕對“0”點的數值資料。

④比例尺度衡量有絕對“0”點的數值資料。

2. 試說明以下變數之測量方法屬於何種測量尺度？

①紅色（0），白色（2），藍色（4）

②非常滿意（2），滿意（1），普通（0），不滿意（-1），非常不滿意（-2）

③智商（IQ 表示）④薪資（以元爲單位）

解

①名目尺度。②順序尺度。③區間尺度。④比例尺度。

3. 描述資料特性的統計測量數（statistical measurement）主要包括哪些？請簡單說明之。

解

描述資料特性的統計測量數主要包括中央趨勢的衡量，分散度的衡量，以及偏度，峰度的衡量，這四種測量數均可描述資料的特性。

4. 某專科學校工管系一年級學生共 120 人，分成男、女兩組，男生 80 人平均身高爲 172 公分，標準差爲 7；女生 40 人平均身高爲 164 公分，標準差爲 6。試問：

①男生或女生的身高較一致？

②求全班之平均身高及標準差。

解

①男生的變異係數$CV=\dfrac{7}{172}=0.0407$

女生的變異係數$CV=\dfrac{6}{164}=0.036$

故知女生的身高較一致。

②全班的平均身高=$\dfrac{80\times172+40\times164}{120}=169.33$

標準差=$\sqrt{\dfrac{1}{120}\left[80(7^2+172^2)+40(6^2+164^2)\right]-169.33^2}=7.747$

5. 設台中市醫生和律師去年全年所得的分配情形如下（單位：萬元）：

職業	人數	平均所得	中位數	眾數	標準差

醫　生	240	200	150	120	50
律　師	160	150	120	100	50

① 計算該市從事這二種職業的人的總平均所得。

② 計算二種職業所得的標準差。

③ 哪一種職業的所得差異較大？

解

①總平均所得$=\dfrac{240\times200+160\times150}{240+160}=180$（萬元）

②變異數$=\dfrac{(240-1)50^2+(160-1)50^2+240(200-180)^2+160(150-80)^2}{240+160-1}=3095$

標準差$=\sqrt{3095}=55.63$萬元

③醫生的變異係數$CV=\dfrac{50}{200}=0.25$

律師的變異係數$CV=\dfrac{50}{150}=\dfrac{1}{3}=0.33$

故知律師的所得差異較大。

6. 暑假打工賺取零用錢，其工作情形如下表所示，試計算其每月總收入及平均每小時薪水。

工作種類	家教	學校工讀	速食店	便利商店	補習班	KTV
每月工作時數	8	10	10	16	4	12
時薪　（元）	350	80	100	90	350	110

資料來源：虛擬。

解

每月總收入 = (8×350+…+12×110) = 8,760

每月總工作時數 = 8+10+10+16+4+12 = 60

平均每小時薪水 = 8,760/60=146 (元/小時)

7. 政院主計處調查顯示，台灣地區十五歲以上通勤通學民眾平日通勤通學的時間按地區分如下表所示，試求台灣地區十五歲以上通勤通學民眾平日平均花多少時間通勤通學？

地區別	北部	中部	東部	南部
占總人數比例　（%）	45.30	23.72	2.33	28.65
通勤通學時間　（小時）	1.15	1.00	0.58	0.50

資料來源：《中華民國八十九年台灣地區社會發展趨勢調查報告》，行政院主計處。

解

平均通勤通學時間

$= (45.30\% \times 1.15)+(23.72\% \times 1.00)+(2.33\% \times 0.58)+(28.65\% \times 0.50) = 0.86$ (小時)

8. 一般皆認爲六月是最適宜結婚的月份。依據統計資料，T 市每年約有 23,500 對新人結婚，其中大約有 2,200 對在六月完婚。試分別依下列各機率理論計算出一對新人在六月結婚的機率。
   ①相對次數機率理論
   ②主觀機率理論
   ③先驗機率理論

解

①依據統計資料，在六月結婚的機率 $= \dfrac{2,200}{23,500} = 0.0936$ 。

②依主觀的想法，各月有各月的好處，在哪一個月結婚並無差別，所以在六月結婚的機率 $= 1/12 = 0.0833$ 。

③先驗機率論假設選擇每一天結婚的機率均相等，所以在六月結婚的機率 $= 30/365 = 0.0822$ 。

9. A公司生產的烘乾機每月市場的需求量變化很大。根據過去數年的統計資料，該公司生產的烘乾機每月市場需求的機率分配如下：

需求台數	200	300	400	500
機率	0.1	0.3	0.4	0.2

①若該公司每月的生產量等於每月需求量的期望值，則該公司每月生產多少台烘乾機？

②若每台烘乾機的生產成本爲3,000元，售價爲5,000元，且該公司於上個月賣出340台烘乾機，試問該公司上月份的盈虧爲何？

解

① $200 \times 0.1 + 300 \times 0.3 + 400 \times 0.4 + 500 \times 0.2 = 370$

該公司每月生產370台烘乾機

②該公司上個月的成本 $= 370 \times 3,000 = 1,110,000$

該公司上個月的收入 $= 340 \times 5,000 = 1,700,000$

該公司上個月賺了 $1,700,000 - 1,110,000 = 590,000$ (元)

10. 全校600人性向測驗的平均數75分，標準差5分，試求：
   ①成績在65分與85分之間約有若干人？
   ②成績在62.5分與87.5分之間約有若干人？

解

① $P(65 \le X \le 85) = P(|X-75| < 10) = P(|X-75| \le 2\sigma_X) \ge 1 - 1/4 = 3/4$

因此，$600\times\frac{3}{4}=450$ 。

②

$$P(62.5\le X<87.5)=P(|X-75|<12.5)=P(|X-75|\le 2.5\sigma_X)\ge 1-1/25^2=0.84$$

因此，$600\times 0.84\approx 504$ 即成績在65分與85分之間者至少有450人；62.5分與87.5分之間者至少有504人。

11. 設普通影印機的壽命X爲一常態分配，平均值爲6年，標準差爲2年，製造商於保證期內會免費修理影印機的任何毛病。
①如果保證期爲3年，製造商將免費修理多少的影印機？
②如果製造商只願免費修理5%的影印機，則保證期必須定爲多久？

解

① $P(X\le 3)=P(Z\le -1.5)=0.5-0.4332=0.0668$若保證期爲3年，製造商將必須免費修理6.68%的影印機。

②設保證期爲X年，$P(X\le x)=P(Z\le\frac{x-6}{2})=0.05=P(Z\le -1.645)$ 。

$\Rightarrow\frac{x-6}{2}=-1.645\Rightarrow x=2.71$，保證期約爲2.71年。

12. 爲了檢驗某款迷你車的耗油量，經測試1公升的油料所能行駛的里程數6次，分別是17.2、16.5、17.5、17.7、16.1、15.9公里。若假設里程數爲常態分配，試求該款車1公升油料平均所行駛之里程數的95%信賴區間。

解

$$\overline{X}=\frac{17.2+16.5+17.5+17.7+16.1+15.9}{6}=16.82$$

$$S^2=\frac{\sum X_i^2-n\overline{X}^2}{n-1}=\frac{1699.65-1697.47}{5}=0.44\ ,\ S=\sqrt{S^2}=0.66$$ 。

故μ的信賴區間爲：

$$16.82\pm t_{5,0.025}\frac{0.66}{\sqrt{6}}=16.82\pm 2.571\times 0.27=16.82\pm 0.69$$

即$16.12\le\mu\le 17.51$

13. 從台北市隨機抽取500個人，詢問是否贊成週休二日制，結果有312名贊成。試求台北市贊成週休二日制的比率95%信賴區間。

解

$\hat{p}=\frac{312}{500}=0.624$，$p$的信賴區間爲：

$$0.624 - Z_{0.025}\sqrt{\frac{0.624(1-0.624)}{500}} \le p \le Z_{0.025}\sqrt{\frac{0.624(1-0.624)}{500}}$$

$$0.624 - 0.042 \le p \le 0.624 + 0.042$$

$$0.582 \le p \le 0.666$$

14. 某茶葉製造公司欲了解其在市場的佔有率，乃在市場上進行抽樣調查。假設該公司要求樣本比例與母體之誤差不能超過0.01，且有95%的信賴度，則樣本數應爲何？

解

p未知，故以$p = 1/2$代入，

可解得$n = \dfrac{(Z_{\frac{\alpha}{2}})^2 P(1-P)}{(0.01)^2} = \dfrac{(1.96)^2 \times \frac{1}{4}}{(0.01)^2} \approx 9,604$

故至少應選取9,604個樣本點。

15. 爲了檢驗某款迷你車的耗油量，經測試1公升的油料所能行駛的里程數6次，分別是17.2、16.5、17.5、17.7、16.1、15.9公里。若假設里程數爲常態分配，試求該款車1公升油料平均所行駛之里程數的95%信賴區間。

解

$$\overline{X} = \frac{17.2+16.5+17.5+17.7+16.1+15.9}{6} = 16.82$$

$$S^2 = \frac{\sum X_i^2 - n\overline{X}^2}{n-1} = \frac{1699.65-1697.47}{5} = 0.44 \text{ ， } S = \sqrt{S^2} = 0.66 \text{ 。}$$

故μ的信賴區間爲：

$$16.82 \pm t_{5,0.025}\frac{0.66}{\sqrt{6}} = 16.82 \pm 2.571 \times 0.27 = 16.82 \pm 0.69$$

即$16.12 \le \mu \le 17.51$

16. 假設顧客在娛樂場所所消耗的時間T爲一常態分配，每位客人待在此娛樂場所的平均時間及標準差均爲60分鐘。該娛樂場所的消費額C與T有如下之函數關係：$C = 600 + 5T + 0.2T^2$。試問每位顧客消費額的期望值爲何？

解

$$T \sim N(60,\ \ 3,600) \Rightarrow E(T) = 60, E(T^2) = 3,600 + 3,600 = 7,200$$

$$E(C) = 600 + 5E(T) + 0.2E(T^2) = 600 + 300 + 1,440 = 2,340$$

17. 根據一項調查，高雄市的上班族中，30%的人平常不做任何運動。現若隨機抽取高雄市400名上班族，試以常態連續校正法求其中至少有125名上班族平常不做任何運動的機率。

解

令X表不做運動的上班族的人數，則X的分配近似$N(120,84)$。

$$P(X \geq 125) = P(Z \geq \frac{124.5-120}{\sqrt{84}}) = P(Z \geq 0.49) = 0.3121$$

18. 從台北市隨機抽取500個人，詢問是否贊成週休二日制，結果有312名贊成。試求台北市贊成週休二日制的比率95%信賴區間。

解

$\hat{p} = \frac{312}{500} = 0.624$，$p$的信賴區間為：

$$0.624 - Z_{0.025}\sqrt{\frac{0.624(1-0.624)}{500}} \leq p \leq Z_{0.025}\sqrt{\frac{0.624(1-0.624)}{500}}$$

$$0.624 - 0.042 \leq p \leq 0.624 + 0.042$$

$$0.582 \leq p \leq 0.666$$

19. 台灣有意自美國引入愛國者飛彈，該型飛彈素以攔截地對地飛彈著名。今在一次實彈演習中，40顆愛國者型飛彈成功攔截了28顆的地對地飛彈。

①請問成功攔截機率之估計值。

②又成功攔截機率的90%信賴區間。

解

① $\hat{P} = \frac{28}{40} = 0.7$

② $0.7 - 1.645\sqrt{\frac{0.7 \times 0.3}{40}} \leq p \leq 0.7 + 1.645\sqrt{\frac{0.7 \times 0.3}{40}}$

$$0.581 \leq p \leq 0.819$$

20. 試討論影響信賴區間可靠度與精確度的因素。

解

信賴水準越大、樣本數越小、標準差越大都會使信賴區間越大，精確度越低，可靠度越高。

21. 針對中共軍事演習，我方必須嚴陣以待，以防其假戲真做。設有一雷達監視員，當雷達上有不明飛行物體，他必須在下兩項中做一決定：

H_0：一切良好，只是偶然干擾而已

H_1：有敵機來襲

①「錯誤警報」為何種誤差？（Type Ⅰ or Type Ⅱ）；其機率表示為（填α，β）。

②「疏忽而未放警報」為　　　誤差，其機率表示為　　（填α, β）。

③「寧可錯放警報」爲　　　增加而　　　減少　（填α,β）。

解

①型 I 錯誤(type I error)、α 。
②型 II 錯誤(type II error)、β 。
③α 、β 。

22. 陳教授懷疑學生家庭作業互相抄襲，因此家庭作業成績的標準差很小。根據以往的經驗，標準差約爲10分，且作業成績呈常態分配。上個星期該班共有30位學生繳交作業，作業成績的標準差爲6分，試問在顯著水準爲5%時，這位教授的懷疑是否獲得證實？

解

令σ表家庭作業的標準差，欲檢定：$H_0:\sigma \geq 10$，$H_1:\sigma < 10$。

檢定統計量：$\chi^2 = \dfrac{(30-1)6^2}{10^2} = 10.44$，臨界值：$\chi^2_{29,0.95} = 17.7084$

因檢定統計量小於臨界值，故拒絕虛無假設，該教授的懷疑獲得證實。

23. 何謂變異數分析？爲何變異數分析的檢定是採右尾檢定？

解

變異數分析是檢定兩個母體或多個母體均數是否相等的統計方法，亦是檢定因子對依變數是否有影響的統計方法。變異數分析採右尾檢定的理由是因所檢定的對立假設 $H_1:\mu_i$不全等，就是檢定 $H_1:E(MSF)>\sigma^2$，因此爲了使檢定力最大，其決策規則採右尾檢定。

24. 某知名百貨公司爲了解顧客的消費行為以決定其產品促銷的方案，紀錄了顧客在該公司消費時所使用的信用卡種類（包括 VISA、MasterCard 和 JCB）和每筆信用卡交易的金額。若該公司欲比較三種信用卡的平均每筆交易金額是否有差異，試問此一研究的各要素爲何？
①反應變數（response variable），②因子（factor），
③處理（treatment），④實驗單位（experimental unit）

解

①反應變數爲每筆信用卡交易的金額。
②因子爲信用卡的種類。
③本實驗有 3 個處理，爲 VISA、MasterCard 和 JCB。
④實驗單位爲在該百貨公司利用信用卡購物的顧客。

25. 在簡單迴歸分析中請說明「樣本判定係數」之意義？若爲 0 則其意義爲何？

解

樣本判定係數是指迴歸方程式解釋的變異佔總變異的比例，亦即代表迴歸方程式的釋能力。若判定係數爲0，表示迴歸方程式無解釋能力。

26. 一唱片公司欲知打歌費用(十萬元)(X)與唱片銷售量(千張)(Y)之間的關係，乃從其所發行的唱片中隨機抽選了10張，得如下的資料：

$\Sigma X = 28$ ，$\Sigma X^2 = 303.4$ ，$\Sigma Y = 75$ ，$\Sigma Y^2 = 598.5$ ，$\Sigma XY = 237$ 。

①試求迴歸直線 $\hat{Y} = \hat{\alpha} + \hat{\beta}X$ 。

②是否打歌費花得愈多，唱片的銷售量就愈高($\alpha = 5\%$)？

解

①$\hat{\beta} = \dfrac{\sum XY - n\overline{X}\,\overline{Y}}{\sum X^2 - n\overline{X}^2} = \dfrac{237 - 10(2.8)(7.5)}{303.4 - 10(2.8)^2} = \dfrac{27}{225} = 0.12$

$\hat{\alpha} = \overline{Y} - \hat{\beta}\overline{X} = 7.5 - 0.12(2.8) = 7.164$，迴歸直線：$\hat{Y} = 7.164 + 0.12X$

②欲檢定 $H_0{:}\beta \le 0,\quad H_1{:}\beta > 0$

$SSE = \sum y^2 - \hat{\beta}\sum xy = (\sum Y^2 - n\overline{Y}^2) - \hat{\beta}(\sum XY - n\overline{X}\,\overline{Y})$

$= [598.5 - 10(7.5)^2] - 0.12[237 - 10(2.8)(7.5)] = 36 - 3.24 = 32.76$

$S_{\hat{\beta}} = \sqrt{\dfrac{SSE\ /(n-2)}{\sum X^2 - n\overline{X}^2}} = 0.135$，$t = \dfrac{0.12 - 0}{0.135} = 0.89 < t_{8,0.05} = 1.86$ 。

拒絕虛無假設，打歌費花得愈多，唱片銷售量未必就愈高。

27. 某五星級大飯店的住屋率(%)(X)與每天每間客房的成本(元)(Y)如下：

X	100	75	65	55	50
Y	2000	2500	2800	3200	4000

①試求迴歸直線 $\hat{Y} = \hat{\alpha} + \hat{\beta}X$ 。

②試檢定此迴歸直線的斜率是否爲零($\alpha = 5\%$)？

解

① $\overline{X} = 69$ ，$\overline{Y} = 2{,}900$ ，$\sum X^2 = 25{,}375$ ，$\sum Y^2 = 44{,}330{,}000$ ，

$\sum XY = 945{,}500$

$\hat{\beta} = \dfrac{\sum XY - n\overline{X}\,\overline{Y}}{\sum X^2 - n\overline{X}^2} = \dfrac{945500 - 5(69)(2900)}{25375 - 5(69)^2} = \dfrac{-55000}{1570} = -35.0318$

$\hat{\alpha} = \overline{Y} - \hat{\beta}\overline{X} = 2900 - (-35.0318)(69) = 5317.1942$ 。

迴歸直線爲 $\hat{Y} = 5317.1942 - 35.0318X$

②欲檢定：$H_0{:}\beta = 0,\quad H_1{:}\beta \ne 0$

計算得

$SSE=\sum(Y-\hat{Y})^2=\sum y^2-\hat{\beta}^2\sum x^2=2280000-1926746.3=353253.7$

，

$$S_{\hat{\beta}}=\sqrt{\frac{SSE/(n-2)}{\sum X^2-n\overline{X}^2}}=\sqrt{\frac{353253.7/3}{1,570}}=8.660$$

因$t=\dfrac{-35.0318-0}{8.660}=-4.045<-t_{3,0.025}=-3.182$，故拒絕虛無假設，斜率不爲零。

28. 某人建立一個迴歸模型進行實証研究，該迴歸模型如下：

$Y=\beta_0+\beta_1X_1+\beta_2X_2+\beta_3X_3+\beta_4X_4+\varepsilon$

已知樣本數爲30，且 $SSE=0.35$， $R^2=0.94$，試回答一列問題：

①是否可以從 SSE 和 R^2 判斷該迴歸模型的解釋能力？請說明理由。

② 此一迴歸模型是否能用來預測 Y？並檢定 $H_0:\beta_1=\beta_2=\beta_3=\beta_4=0$ 和 $H_1:\beta_i$不全爲0。（$\alpha=5\%$）

解

①可以。$R^2=0.94$ 很接近 1，表示該迴歸模型的解釋能力很強。但因爲不知道 Y 的單位爲何，故無法從 SSE 判斷迴歸模型的解釋能力。

②可以。因爲本題的檢定量爲

$$F=\frac{R^2/k}{(1-R^2)/n-k-1}=\frac{0.94/4}{(1-0.94)/30-4-1}=97.92$$

，查表得臨界值 $F_{4,25,0.05}=2.76$。因爲 F 檢定量 97.92 大於臨界值 2.76，因此在 $\alpha=5\%$ 顯著水準下，拒絕 H_0。此即表示，該迴歸模型可接受。

29. 在複迴歸模型中，當我們使用 F 檢定統計量去檢定 $H_0:\beta_i$全爲0 和 $H_1:\beta_i$不全爲0 時，得到檢定結果爲拒絕 $H_0:\beta_i$全爲0，即 β_i 不全爲 0，試問：

①此一檢定結果是否表示該迴歸模型具有最佳的解釋能力？

②此一檢定結果是否表示所有的自變數（X）對於依變數（Y）均具有重要的解釋能力？

③根據此一檢定結果，請說明你認爲最適當的結論。

解

①否。不一定爲最佳。

②否。由檢定結果僅可得知該迴歸模型包括了部份重要的解釋變數，但同時也可能有部分重要的解釋變數並未在該模型之中。

③由檢定結果可知，至少有一個解釋變數對於該模型具解釋能力。

30. 張快樂利用OLS估計得迴歸方程式如下：

$\hat{Y}=16.5+2.1X_1+50.0X_2$ ， $R^2=0.333,\ \ n=28,\ \ F=18.75$

SD (10.0) (20.0)

試回答下列問題：

①請檢定 X_1 之迴歸係數。 $H_0:\beta_1=1,\ \ H_1:\beta_1>1(\alpha=5\%)$

②請檢定 X_2 之迴歸係數。 $H_0:\beta_2=0,\ \ H_1:\beta_2\neq 0(\alpha=5\%)$

③請檢定 $H_0:\beta_1=\beta_2=0,\ \ H_1:H_0$不爲真 $(\alpha=5\%)$ 。

解

①檢定 $H_0:\beta_1=1$ ； $H_1:\beta_1>1$

$$t=\frac{2.1-1}{0.5}=2.2>t_{25,0.05}=1.708$$

因此拒絕 H_0 ， X_1 的迴歸係數大於1。

②檢定 $H_0:\beta_2=0$ ； $H_1:\beta_2\neq 0$

$t=\dfrac{50}{20}=2.5>t_{25,0.025}=2.060$ 。 因此拒絕 H_0 ， X_2 對Y有影響。

③檢定 $H_0:\beta_1=\beta_2=0$ ； $H_1:H_0$不真

$F=18.75>F_{2,25,0.05}=3.39$

因此拒絕 H_0 ， X_1 與 X_2 對Y有聯合影響力。