初等考試
五等特考

2010年

統計學大意測驗題庫

經典試題解析

宏典文化
greatbooks.com.tw

九十八年公務人員初等考試試題

科別：統計　　科目：統計學大意

1.（　）為瞭解國小學童的零用錢使用狀況，針對臺北縣某國小進行抽樣，其方式為自該學校一至六年級，每個年級隨機抽取30位小朋友。此抽樣方式為何？ (A) 系統隨機抽樣（systematic random sampling）(B)簡單隨機抽樣（simple random sampling）(C)分層隨機抽樣（stratified random sampling） (D)群集隨機抽樣（cluster random sampling）

2.（　）承上題，若依此調查結果而說「該國小學生每週平均零用錢約為新臺幣40元」，此稱為何？ (A) 母體參數 (B)統計推論 (C)隨機抽樣 (D)變異數分析

3.（　）若X是一個隨機變數，其均數為μ_x，標準差為σ_x；Y亦為一隨機變數，其均數為μ_y，標準差為σ_y。則X+Y的均數為何？ (A)$\mu_x+\mu_y$ (B) $\mu_x/\sigma_x+\mu_y/\sigma_y$ (C) $\mu_x+\mu_y$，且必須X與Y是互相獨立的 (D) $\mu_x/\sigma_x+\mu_y/\sigma_y$，且必須X與Y是互相獨立的

4.（　）某班級有50位同學，其統計學期末考的全班平均成績為75分，標準差10分。現因某題目出錯，整題送分，若每位同學的成績因此各加5分，則新的全班成績的標準差為何？ (A)10/50 (B)100/50 (C)5 (D)10

5.（　）設A與B表示同一隨機實驗之任兩個事件，則下列敘述何者正確？ (A)若$P(A\cap B)=0$，則事件A與B為獨立事件 (B)若$P(A\cap B)=P(A)P(B)$，$P(B\cap C)=P(B)P(C)$，$P(A\cap C)=P(A)P(C)$，則事件A、B、C為獨立事件 (C)若A與B為獨立事件，則$P(B\mid A)=P(A\mid B)$ (D)$P(A)=P(A\cap B)+P(A\cap B^C)$， B^C為B之補集

6.（　）某市政府之市政查詢專線電話，在上班日早上9點至11點之間，其接到電話的次數服從卜瓦松（Poisson）分配，其平均數為每15分鐘3.5通。則在某個星期一早上9點至11點之間，該專線電話平均會接到幾通電話： (A)3.5 (B)14 (C)28 (D)7

7.（　）某市政府之市政查詢專線電話，在上班日早上9點至11點之間，其接到電話的次數服從卜瓦松（Poisson）分配，其平均數為每15分鐘3.5通。則在某個星期一早上9點至9點15分之間，未接到任何一通電話的機率為： (A)0.0009 (B)0.0151 (C)0.0302 (D)0.2333

8.（　）某市場單位為衡量其磅秤的精準度，將一個已知是一公克的物品重複衡量多次。若該磅秤的測量值為常態分配，其具有未知均數μ及標準差0.01公克。在95%信心水準下，欲得到μ與樣本平均數間的誤差在±0.0001，則需要測量多少次？ (A)384 (B)9604 (C)25252 (D)38416

9. (　) 在統計假設檢定的問題中，下列何者為足夠的證據，得以拒絕虛無假設？ (A)使用較小的顯著水準 (B)使用較大的顯著水準 (C)由資料計算得到較小的p值（p-value） (D)由資料計算得到較大的p值（p-value）

10. (　) 隨機抽取某產業400位中階主管，由此樣本資料得到其平均年紀為43.7歲，標準差4.2歲。試估計該產業中階主管平均年紀之95%信賴區間為何？ (A)43.7 ± 0.41 (B)43.7 ± 8.23 (C)43.7 ± 0.02 (D)43.7 ± 0.35

11. (　) 若其他條件不變時，當信心水準由95%改為90%，其所對應的信賴區間的寬度會如何改變？ (A)變寬 (B)變窄 (C)維持一致 (D)視抽樣母體而決定其區間變寬或變窄

12. (　) 以下何者的敘述不適合t分配？ (A)為一連續分配 (B)為一對稱分配 (C)較標準常態分配更為平坦 (D)較標準常態分配更為右偏

13. (　) 以下有關抽樣誤差（sampling error）的敘述，何者是正確的？ (A)永遠為正數 (B)永遠為負數 (C)樣本統計量與母體參數之間的差 (D)母體參數的一個值

14. (　) 當下列何條件成立時，常態分配可用來近似二項分配？ (A)np至少大於25 (B)np及$n(1-p)$ 兩者皆至少大於5 (C)$np(1-p)$ 至少大於5 (D)只有當Z分數（Z-score）大於5時

15. (　) 以下何者為二項分配與超幾何分配之差異？ (A)超幾何分配必須是大樣本 (B)二項分配的均數與變異數必須相等 (C)對超幾何分配而言，其每次試驗的成功機率不需相等 (D)對二項分配而言，每次試驗不需是獨立的

16. (　) 以下為隨機抽取自某大型企業的員工所得的訊息：

	男性	女性
樣本數	64	36
樣本月平均收入 (NT$1,000)	44	41
母體變異數	128	72

則該公司男女員工的平均月收入差的95%信賴區間為何？ (A)(-0.30,6.30) (B)(0.23,5.77) (C)(-1.96,1.96) (D)(-0.92,6.92)

17. (　) 以下為某縣隨機抽取去年交通違規罰單繳款人，其中選擇自動櫃員機（ATM）轉帳方式繳款的統計：

	40歲（含）以下	40歲以上
抽樣人數	500	600
以ATM轉帳	180	150

若欲檢定40歲（含）以下及40歲以上兩者間，以ATM轉帳繳款方式的比例差是否相等。請計算此檢定之檢定統計量的值為何？ (A)7.96 (B)0.24 (C)2.96 (D)3.96

18. (　) 隨機抽取某藥廠所生產的20毫升瓶裝感冒藥水20瓶，得到其容量的樣本標準差為0.4毫升。設藥水容量之母體分配為常態，請計算該藥廠所生產的瓶裝感冒藥水的容量之母體變異數的95%信賴區間為何？ (A)(0.2313，0.8533) (B)(0.2224，0.7924) (C)(0.0889，0.3169) (D)(0.0925，0.3413)

19. (　) 以下對間斷隨機變數X及其機率分配函數f(x)的敘述何者正確？ (A)X的值永遠為非負的 (B)有些f(x)可以是負的，只要$\sum_x f(x)=1$ (C) f(x)必須是大於或等於0且$\sum_x f(x)=1$ (D) f(x)必須是單調遞增函數

20. (　) 某國小足球隊單場進球數的機率分配如下：

進球數	0	1	2	3	4
機率	0.05	0.15	0.35	0.30	0.15

該足球隊平均單場進球數為何？ (A)0 (B)1 (C)2 (D)2.35

21. (　) 某地政單位櫃臺處理完成一位民眾臨櫃之事務所需的時間服從均勻分配，其時間介於6至10分鐘。則該單位平均處理一位民眾臨櫃所需的時間為多少分鐘？ (A)16 (B)2 (C)8 (D)4

22. (　) 若隨機變數X為具有均數10，標準差5的常態分配。則X等於10的機率為何？ (A)1 (B)0.5 (C)0 (D)0.95

23. (　) 若隨機變數X的機率分配函數如下：$f(x)=\dfrac{1}{10}e^{-x/10}$　$x\geq 0$，則X的平均數為何？ (A)0.1 (B)10 (C)100 (D)1000

24. (　) 若商學碩士畢業生的月起薪服從常態分配，其均數為新臺幣40,000元，標準差為新臺幣5,000元。則隨機抽取一個商學碩士畢業生，其月起薪大於30,000元的機率為何？ (A)0.4772 (B)0.9772 (C)0.0228 (D)0.5000

25. (　) 當以下何條件成立時，一樣本統計量為母體參數的不偏（unbiased）估計？ (A)該樣本統計量的期望值等於0 (B)該樣本統計量的變異數等於0 (C)該樣本統計量的期望值等於母體參數 (D)該樣本統計量的變異數等於母體參數

26. (　) 若國內成年人贊成減稅的比例為80%，隨機抽取400位成年人，會得到贊成的比例大於0.83的機率為何？ (A)0.4332 (B)0.9332 (C)0.0668 (D)0.5668

27. (　) 當樣本數增加時，樣本平均數的變異數的增減變化為何？ (A)增加 (B)減少 (C)不變 (D)其增加或減少之改變，視抽樣母體之不同而有所不同

28. (　) 當我們配適一簡單線性迴歸模型$Y=\beta_0+\beta_1X+\varepsilon$至一組資料（$X_i, Y_i$），i=1, 2,..., n，ε為隨機誤差項；並以最小平方法估計其迴歸參數β_0及β_1。則其所有殘差的和：(A)必等於0 (B)必大於0 (C)必小於0 (D)不必然大於、小於或等於0

29.（　）承上題，若β_1的95%信賴區間估計包含0，則在$=\alpha$ 0.05時：(A)不會拒絕H_0：$\beta_1=0$，當H_1：$\beta_1\neq0$ (B)會拒絕H_0：$\beta_1=0$，當H_1：$\beta_1\neq0$ (C)會拒絕H_0：$\beta_1=0$，當H_1：$\beta_1>0$ (D)會拒絕H_0：$\beta_1=0$，當H_1：$\beta_1<0$

30.（　）迴歸分析中，共線性（multicollinearity）的現象是指：(A)解釋變數（explanatory variable）之間是相關的 (B)解釋變數與時間有關 (C)解釋變數與反應變數（response variable）之間是相關的 (D)殘差不具有固定變異數

31.（　）若Y為迴歸分析中的反應變數（response variable），$\overline{Y}$為其平均數，$\hat{Y}$為對應的配適值(fitted value)。最小平方法估計意指為何？ (A) $\Sigma(Y-\hat{Y})^2=0$ (B) $\Sigma(Y-\overline{Y})^2$的極大化 (C) $\Sigma(Y-\hat{Y})^2=0$的極小化(D) $\Sigma(Y-\overline{Y})^2$的極小化

32.（　）線性迴歸分析中，估計的標準誤（standard error）是：(A)與反應變數有相同的單位 (B)所有殘差和的平均 (C)永遠介於-1與+1之間 (D)永遠與迴歸線的斜率具相同正負符號

33.（　）若已知$\sum X=39$，，$\sum X^2=209$，$\sum Y=1092$，$\sum Y^2=128940$，$\sum XY=4903$，n=10，則X與Y的相關係數估計等於：(A)0.752 (B)0.867 (C)0 (D)1

34.（　）若將100名基層公務人員依其職務（分三類）及性別分類，製作成一列聯表（contingency table）。則對此列聯表做卡方（χ^2）獨立性檢定時，其所需自由度為何？ (A)2 (B)3 (C)4 (D)99

35.（　）對一列聯表做卡方（χ^2）獨立性檢定時，其觀測次數（observed frequency）的和與期望次數（expected frequency）的和：(A)必須至少30 (B)永遠相等 (C)永遠小於5% (D)永遠小於5

36.（　）進行變異數分析（analysis of variance）時，不需要那些假設？ (A)母體為常態分配 (B)母體變異數相等 (C)母體均數相等 (D)樣本是互相獨立的

37.（　）變異數分析（analysis of variance）的均方誤（mean square error）項是指：(A)共同母體均數的估計值 (B)共同母體變異數的估計值 (C)處理的變異 (D)樣本標準差的估計值

38.（　）考慮一包含15個解釋變數及200個觀測值的複迴歸分析，若得到SSTotal=800，SSE = 240，則其判定係數（coefficient of determination）為何？ (A)0.300 (B)0.192 (C)0.500 (D)0.700

39.（　）若以下之估計迴歸式是得自30個觀測值：$\hat{Y}=35+4.2X_1-3.1X_2+8.2X_3+2.4X_4$，同時亦計算得到SSReg = 700 及SSE = 100。則欲檢定上述模型之F檢定的檢定統計值為何？ (A)43.75 (B)0.875 (C)50.19 (D)7.00

40.（　）於變異數分析（ANOVA）的實驗設計，設母體分配符合常態分配，考慮以下的檢定：

H_0：$\mu_1=\mu_2=\mu_3=\mu_4$

H_1：至少有一個均數是不相等的

且經計算得到以下訊息：

SS(處理) = 6,750　，　SSE = 8,000　，　$n_1=n_2=n_3=n_4=5$為各組樣本數。

則檢定統計值為何？ (A)0.22 (B)0.84 (C)4.22 (D)4.5

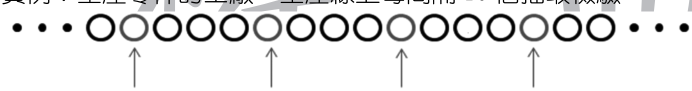
<table>
<tr><td>科目</td><td colspan="3">統計學大意　(試題代號：2508)</td><td>題數</td><td>40</td></tr>
<tr><td>題序</td><td>01 — 10</td><td>11 — 20</td><td>21 — 30</td><td colspan="2">31 — 40</td></tr>
<tr><td>答案</td><td>CBADDCCDCA</td><td>BDCBCDDDCD</td><td>CCBBCCBAAA</td><td colspan="2">CABABCBDAD</td></tr>
<tr><td>備註</td><td colspan="5">第 30 題答 A 或 D 或 A D 者均給分。</td></tr>
</table>

解析

1.（C）

(A) 系統隨機抽樣（systematic random sampling）: 從已排序的母體中，隨機抽出第一個元素，以此元素為起點，每間隔 k 個抽取下一個元素。

實例：生產零件的工廠，生產線上每間隔 10 個抽取檢驗。

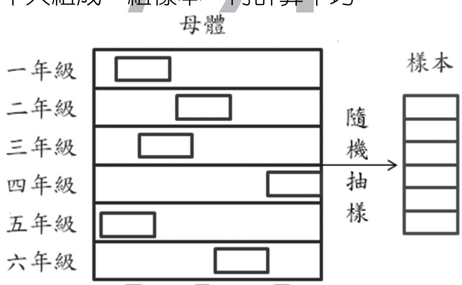


(B) 簡單隨機抽樣（simple random sampling）: 從母體中隨機抽取若干個元素，每個元素被抽取機率皆相同。

實例：大樂透，從 49 個號碼隨機抽取 6 個號碼和 1 個特別號。

(C) 分層隨機抽樣（stratified random sampling）: 將母體依某特徵分為若干群（層），每群都要互斥，在從各群中隨機抽出若干個組成一組樣本。

實例：此題，想要知道小學生的每月零用金，從每一個年級（即為一個層或群）抽出若干人組成一組樣本，再計算平均。



(D) 部落抽樣（cluster sampling）: 又稱群集抽樣、集體抽樣，根據母體的某特徵(如區域)，把母體分成若干個群集，隨機抽取若干個群集，並針對抽取群體中的每一個普查。注意！層和群集的不同在於群集間的變異小，群集內的變異大，而層則相反。

實例：電訪調查某產品市民的滿意度，從台北市 12 個行政區中，隨機選出 3 個行政區，再針對這 3 個行政區每位市民做調查。

2.（B）

依照資料(樣本)的特性進行推論(母體)，這是統計推論的定義。此題即可利用「該國小學生每週平均零用錢約為新台幣 40 元」這個樣本數據，來推論母體的均數 μ。

3.（A）

利用期望值「線性」的特性：$E(X+Y)=E(X)+E(Y)=\mu_x+\mu_y$。注意期望值即是均數的概念。例如某班某次段考國文平均 90 分，數學平均 85 分，所以兩科平均總分為 90+85=175 分。

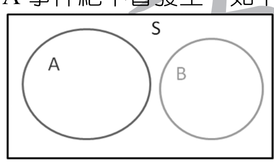
4.（D）

標準差表示資料的分散度，所以資料平移後，每筆資料移動等量的距離，並不會改變資料的分散程度。可由樣本標準差定義來解釋：$S_x=\sqrt{\dfrac{\sum_{i=1}^{n}(x_i-\bar{x})^2}{n-1}}$，若把資料平移，亦即把每一筆資料 x_i 加上一個常數 C，令 $y_i=x_i+C$，則有 $\bar{y}=\bar{x}+C$ 且有 $S_y=\sqrt{\dfrac{\sum_{i=1}^{n}(y_i-\bar{y})^2}{n-1}}$

$=\sqrt{\dfrac{\sum_{i=1}^{n}[(x_i+\cancel{c})-(\bar{x}+\cancel{c})]^2}{n-1}}=\sqrt{\dfrac{\sum_{i=1}^{n}(x_i-\bar{x})^2}{n-1}}$，所以知道資料的平移會使得標準差不改變。此題中加 5 分即是資料平移 5 分。

5.（D）

(A)若 A 和 B 為互斥事件，表示 A 事件發生時，B 事件絕不會發生，同理，B 事件發生時，A 事件絕不會發生，如下圖。



<=A 和 B 互斥（表示互不相交，沒有重疊，P(A∩B)=0）

若 A 與 B 為獨立事件（表示互不影響），則 P(A∩B)=P(A)P(B)。

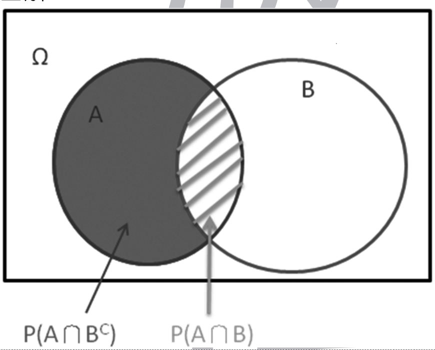
(B) 兩兩獨立不保證三個事件的獨立！若 A,B,C 三事件獨立，則必須滿足：$P(A\cap B)=P(A)P(B)$，$P(B\cap C)=P(B)P(C)$，$P(A\cap C)=P(A)P(C)$ 以及 $P(A\cap B\cap C)=P(A)P(B)P(C)$，所以(B)選項並不一定會成立。

(C) 若 A,B 為獨立事件，則 $P(A\cap B)=P(A)P(B)$，因此 $P(A|B)=P(A\cap B)/P(B)=P(A)P(B)/P(B)=P(A)$，表示不論 A 事件有無發生，都不會影響到 B 事件發生的機率，同理，$P(B|A)=P(B)$。

(D) 利用集合論中的分配律：$A=A\cap\Omega=A\cap(B\cup B^c)=(A\cap B)\cup(A\cap B^c)$

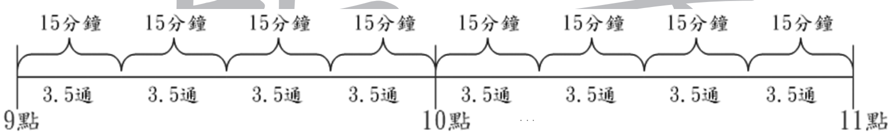
其中 $A\cap B$ 與 $A\cap B^c$ 互斥，故有 $p(A)=P(A\cap B)+P(A\cap B^c)$。

圖解：



6.（C）

因為卜瓦松分配的期望值λ會和時間的長度成正比。題目告訴我們「15分鐘的均數為3.5通」，所以 9 點到 11 點 2 個小時(為 15 分鐘的 8 倍)的期望值為 $3.5\times8=28$，即平均接到 28 通。

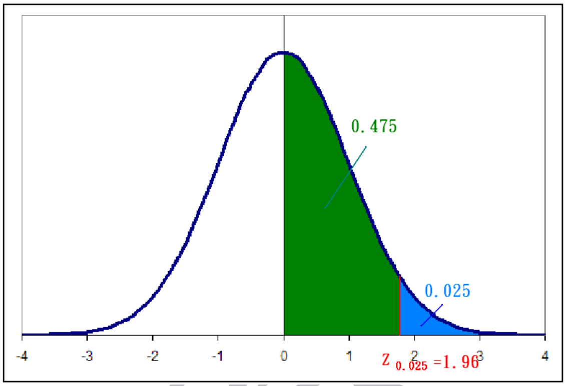


7.（C）

同樣利用卜瓦松分配的期望值λ會和時間的長度成正比，所以早上 9 點至 9 點 15 分之間，共 15 分鐘，這段時間內平均會接到 3.5 通，即 $\lambda=3.5$，而卜瓦松分配的機率函數為 $f(x)=\frac{e^{-\lambda}\lambda^x}{x!}$，隨機變數 X 表示發生次數，故本題未接到任何一通電話的機率為 $f(0)=\frac{e^{-3.5}3.5^0}{0!}=e^{-3.5}\approx0.0302$。

8.（D）

此題 σ^2 已知，可用以下的信賴係數(1-α)%的信賴區間來估計 μ：$\left(\overline{X}-Z_{\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}},\overline{X}+Z_{\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}}\right)$，其估計誤差為 $\pm Z_{\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}}$，本題要求估計誤差在 ±0.0001，即 $Z_{\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}}=0.0001$，$Z_{\frac{0.05}{2}}\frac{0.01}{\sqrt{n}}=0.0001$，$n=38416$，故需要測量 38416 次。其中 $Z_{0.025}=1.96$。

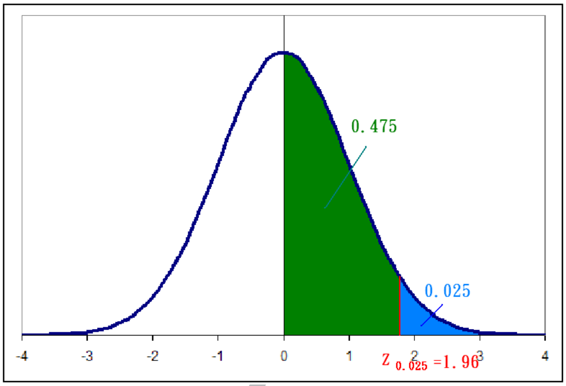


9.（C）

一般來說，顯著水準是檢定一開始就設定的，與是否拒絕虛無假設是無關的。此外，當 p-值 (p-value)越小，虛無假設越應該拒絕，這表示虛無假設越不可能成立。

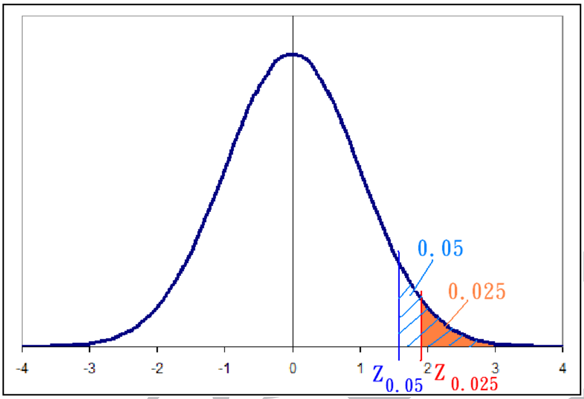
10.（A）

由於題目未提及年齡的分配為何，但當樣本數 n 夠大時，可用常態分配近似，則其 μ 的信賴係數(1-α)%的信賴區間為$\left(\overline{X}-Z_{\frac{\alpha}{2}}\frac{S}{\sqrt{n}},\overline{X}+Z_{\frac{\alpha}{2}}\frac{S}{\sqrt{n}}\right)$，所以可得到 $43.7\pm1.96\frac{4.2}{\sqrt{400}}=43.7\pm0.4116$ $\approx43.7\pm0.41$。其中 $Z_{0.025}=1.96$。



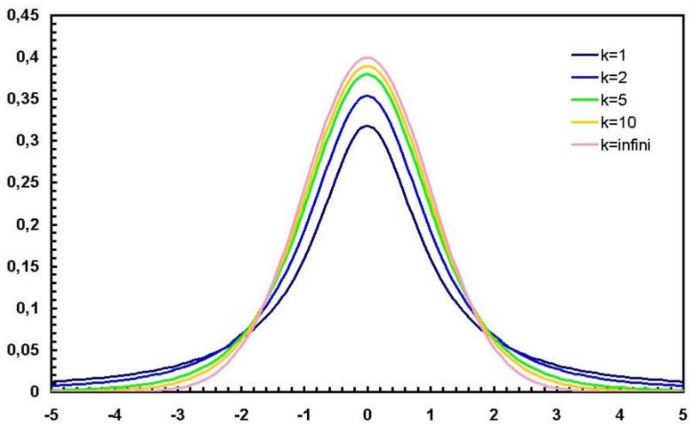
11.（B）

假設母體變異數 σ^2 已知，可用以下的信賴係數(1-α)%的信賴區間來估計 μ ：$\left(\overline{X}-Z_{\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}},\overline{X}+Z_{\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}}\right)$，區間寬度為 $2*Z_{\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}}$，所以當信賴水準由 95%改為 90%時，$Z_{0.025}>Z_{0.05}$，在其他條件不變下，區間寬度會變窄。



12.（D）

t 分配為一連續分配且對稱，當自由度愈大時，愈接近常態分配，從下圖可知 t 分配較常態分配平坦。



圖片來源：http://en.wikipedia.org/wiki/Student%27s_t-distribution，圖中 k 表示自由度。

13.（C）

用來估計母體參數的信賴區間為（樣本統計量 ± 抽樣誤差），以估計 μ 的信賴區間為例，假設母體變異數 σ^2 已知，可用以下的信賴係數(1-α)%的信賴區間來估計 μ ：$\left(\overline{X}-Z_{\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}},\overline{X}+Z_{\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}}\right)$，抽樣誤差為$Z_{\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}}$，所以抽樣誤差為樣本統計量與母體參數之間的差。

14.（B）

當 $np > 5$ 且 $n(1-p) > 5$ 時，才可用常態分配近似二項分配。

15.（C）

(A) 不一定需要大樣本。

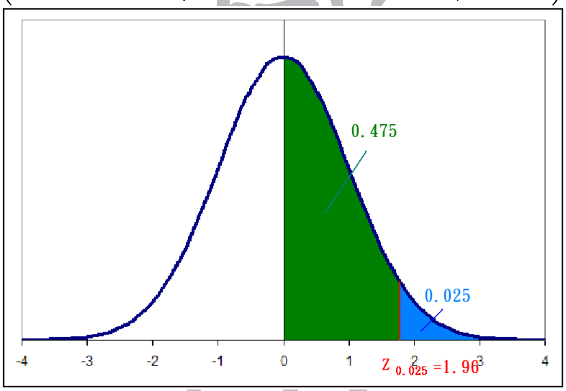
(B) 二項分配的均數為 np，變異數為 np(1-p)。

(C) 對超幾何分配而言，其每次試驗的成功機率不需相等，因為超幾何分配為「不放回」的試驗，所以每次抽取後，其樣本空間會改變，因此每次試驗的成功機率不一定相同。

(D) 對二項分配而言，每次試驗需是獨立的。

16.（D）

此題中，男性、女性為兩獨立個體，但未提及母體分配是否為常態，可因樣本數皆大於 30，且母體變異數已知，故利用兩獨立母體平均數差 $\mu_1 - \mu_2$ 的區間估計公式：$\left(\overline{x_1} - \overline{x_2} - Z_{\frac{\alpha}{2}}\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}, \overline{x_1} - \overline{x_2} + Z_{\frac{\alpha}{2}}\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}\right)$，所以此題男女員工的平均月收入差的 95%信賴區間為 $\left(44 - 41 - Z_{\frac{0.05}{2}}\sqrt{\frac{128}{64} + \frac{72}{36}}, 44 - 41 + Z_{\frac{0.05}{2}}\sqrt{\frac{128}{64} + \frac{72}{36}}\right)$ ≈ (-0.92, 6.92)。其中 $Z_{0.025} = 1.96$。



17.（D）

	40 歲（含）以下	40 歲以上
抽樣人數	n_1=500	n_2=600
以 ATM 轉帳	x_1=180	x_2=150
樣本比例	$\hat{p}_1$=180/500=0.36	$\hat{p}_2$=150/600=0.25

H_0：$p_1 - p_2 = 0$

H_1：$p_1 - p_2 \neq 0$

因為樣本數皆大於 30，可用 Z 分配。

檢定統計量 $Z = \frac{\hat{p}_1 - \hat{p}_2}{\sqrt{\frac{\hat{p}(1-\hat{p})}{n_1} + \frac{\hat{p}(1-\hat{p})}{n_2}}} = \frac{0.36 - 0.25}{\sqrt{\frac{0.3(1-0.3)}{500} + \frac{0.3(1-0.3)}{600}}} \approx 3.96$。

其中 $\hat{p} = \frac{x_1 + x_2}{n_1 + n_2} = \frac{180+150}{500+600} = 0.3$ 為合併比例。

18.（D）

母體為常態且未知母體平均數，故其母體變異數信賴係數(1-α)%的信賴區間為

$\left(\frac{(n-1)s^2}{\chi^2_{\frac{\alpha}{2}}(n-1)},\frac{(n-1)s^2}{\chi^2_{1-\frac{\alpha}{2}}(n-1)}\right)$，因此瓶裝感冒藥水的容量之母體變異數的 95%信賴區間為

$\left(\frac{(20-1)0.4^2}{\chi^2_{\frac{0.05}{2}}(20-1)},\frac{(20-1)0.4^2}{\chi^2_{1-\frac{0.05}{2}}(20-1)}\right)$ ≈ (0.0925, 0.3413)。

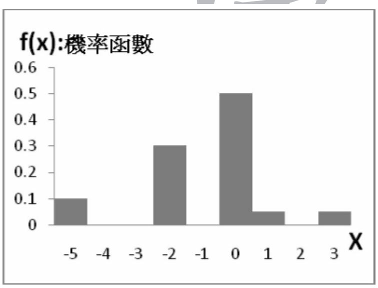
其中$\chi^2_{0.025}(19)$ = 32.8523，$\chi^2_{0.975}(19)$ = 8.90655。

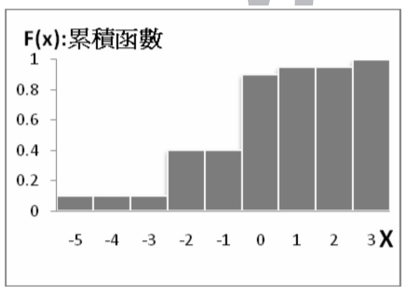
19.（C）

(A) 間斷隨機變數 X 可以是正數。

(B) (C)在間斷型分配中，機率函數 f(x)即式機率值，因此 f(x)必須為正數，且機率總和要為 1，即$\sum_x f(x) = 1$。

(D) 累積機率函數 F(x)才必須要是單調遞增函數。





20.（D）

利用求離散型期望值公式：E(X) = $\sum_{-\infty}^{\infty} xf(x)$。

進球數 x	0	1	2	3	4
機率 f(x)	0.05	0.15	0.35	0.30	0.15
xf(x)	0	0.15	0.70	0.90	0.60

=>$\sum_{-\infty}^{\infty} xf(x)$ = 0+0.15+0.70+0.90+0.60 = 2.35。

21.（C）

計算平均處理一位民眾臨櫃所需的時間，即計算期望值，而均勻分配的期望值為(a+b)/2，此題 a = 6，b = 10，故期望值為 8。

22.（C）

因為常態分配為連續型分配，在連續型分配中，單點機率值為 0，所以 X 等於 10 的機率為 0。

23.（B）

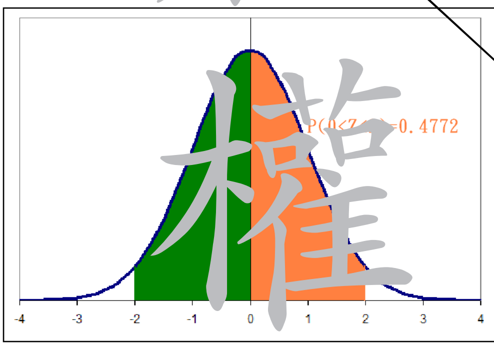
解法一： $E(X) = \int_{-\infty}^{\infty} xf(x)dx = \int_{0}^{\infty} x * \frac{1}{10} e^{-x/10}dx = -xe^{-x/10} - 10e^{-x/10}\Big|_{x=0}^{\infty}$ = 10。

解法二： $f(x) = \frac{1}{10} e^{-x/10}$, x≥0 和指數分配的機率函數 $f(x) = \lambda e^{-\lambda x}$, x≥0 相同，因此 X 為指數分配，故其期望值 E(X) = 1/ λ = 1/(1/10)= 10。

24.（B）

令商學碩士畢業生的月起薪為隨機變數 X，根據題意 E(X)=40000，V(X)=5000^2，要算出隨機抽取一個商學碩士畢業生，其月起薪大於 30,000 元的機率，即計算 P(X > 30000)。

$P(X > 30000) = P\left(\frac{X-E(X)}{\sqrt{V(X)}} > \frac{30000-40000}{5000}\right)$ = P(Z > -2) = **P(-2 < Z < 0)** + P(0 < Z <∞) = **P(0 < Z < 2)** + 0.5 = 0.4772+0.5 = 0.9772。



常態分配 $\xrightarrow[\text{標準化}]{}$ 標準常態分配（標準常態分配為常態分配之特例，其 μ 為 0，σ 為 1）

$X \sim N(\mu, \sigma^2) \xrightarrow[\text{標準化}]{} Z = \frac{X-\mu}{\sigma} \sim N(0,1)$

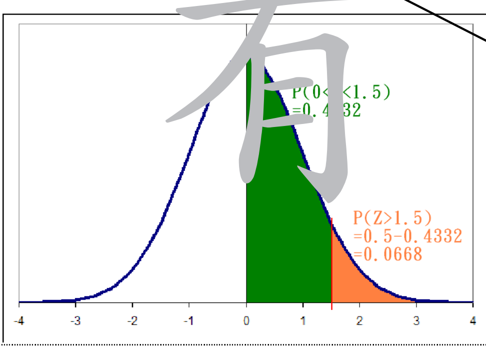
25.（C）

一個估計量是不偏的（unbiased），即 $E(\hat{\theta}) = \theta$，其中$\hat{\theta}$為樣本統計量，用來估計母體參數 θ 。

26.（C）

根據題意，隨機抽取 400 位成年人，即樣本數 n=400，因為樣本數 n=400 夠大，二項分配可用常態分配近似之，令樣本贊成的比例為$\hat{p}$，$E(\hat{p})$= p = 80%，$V(\hat{p})$= p(1-p)/n = 0.0004，要算出贊成的比例大於 0.83 的機率，即計算 $P(\hat{p} > 0.83)$。

$P(\hat{p} > 0.83) = P\left(\frac{\hat{p}-E(\hat{p})}{\sqrt{V(\hat{p})}} > \frac{0.83-0.8}{\sqrt{0.0004}}\right)$ = P(Z > 1.5) = 0.0668。



常態分配 $\xrightarrow[\text{標準化}]{}$ 標準常態分配（標準常態分配為常態分配之特例，其 μ 為 0，σ 為 1）

$X \sim N(\mu, \sigma^2) \xrightarrow[\text{標準化}]{} Z = \frac{X-\mu}{\sigma} \sim N(0,1)$

27.（B）

不論母題的分配為何，當母體的平均數為 μ，標準差為 σ，其樣本平均數的變異數 $V(\overline{X}) = \sigma^2/n$，故樣本數 n 增加時，樣本平均數的變異數減少。

28.（A）

簡單線性迴歸模型 $Y=\beta_0+\beta_1 X+\varepsilon$，利用最小平方法估計其迴歸參數 β_0 及 β_1，求得樣本迴歸線 $\hat{Y}=\hat{\beta}_0+\hat{\beta}_1 X$，此線必過$(\overline{X}, \overline{Y})$，即$\overline{Y}=\hat{\beta}_0+\hat{\beta}_1\overline{X}$，而所有殘差的和為$\sum_{i=0}^{n}[Y_i-(\hat{\beta}_0+\hat{\beta}_1 X_i)]=\sum_{i=0}^{n}Y_i-n\hat{\beta}_0-\hat{\beta}_1\sum_{i=0}^{n}X_i=n\left(\frac{\sum_{i=0}^{n}Y_i}{n}-\hat{\beta}_0-\hat{\beta}_1\frac{\sum_{i=0}^{n}X_i}{n}\right)=n(\overline{Y}-\hat{\beta}_0-\hat{\beta}_1\overline{X})=n*0=0$。

29.（A）

若 β_1 的 95%信賴區間估計包含 0，則代表 $\alpha=0.05$，β_1 的 95%信賴區間包含 0，β_1 有可能等於 0，因此不會拒絕 $H_0: \beta_1=0$，當 $H_1: \beta_1\neq 0$。

30.（A）

迴歸分析中，共線性（multicollinearity）的現象是指解釋變數（explanatory variable）（即自變數）之間是有高度相關性。

31.（C）

最小平方法求得樣本迴歸線$\hat{Y}=\hat{\beta}_0+\hat{\beta}_1 X$，即是用$\sum(Y-\hat{Y})^2$ 的極小化。

32.（A）

估計值的標準誤 $s=\sqrt{MSE}=\sqrt{\dfrac{\sum\left(Y_i-\hat{Y}_i\right)^2}{n-2}}$

(A)與反應變數 Y 有相同的單位。

(B)所有殘差的平方和除以 n-2，再開根號。

(C) (D)$\sqrt{MSE}\geq 0$。

33.（B）

樣本相關係數 $r=\dfrac{\sum_{i=0}^{n}X_iY_i-n\overline{X}*\overline{Y}}{\sqrt{\sum_{i=0}^{n}X_i^2-n\overline{X}^2}\sqrt{\sum_{i=0}^{n}Y_i^2-n\overline{Y}^2}}=\dfrac{\sum_{i=0}^{n}X_iY_i-n\frac{\sum_{i=0}^{n}X_i}{n}*\frac{\sum_{i=0}^{n}Y_i}{n}}{\sqrt{\sum_{i=0}^{n}X_i^2-n\left(\frac{\sum_{i=0}^{n}X_i}{n}\right)^2}\sqrt{\sum_{i=0}^{n}Y_i^2-n\left(\frac{\sum_{i=0}^{n}Y_i}{n}\right)^2}}=$

$\dfrac{4903-10*\frac{39}{10}*\frac{1092}{10}}{\sqrt{209-10(\frac{39}{10})^2}\sqrt{128940-10(\frac{1092}{10})^2}}\approx 0.867$。

34.（A）

職務分三類，性別有兩種，所以自由度為(3-1)*(2-1)=2。

35.（B）

對一列聯表做卡方（χ^2）獨立性檢定時，其觀測次數（observed frequency）的和與期望次數（expected frequency）的和必須相等。

36.（C）

變異數分析即是檢定三個以上的母體平均數是否相等，事先假設(A)母體為常態分配 (B)母體變異數相等 (D)樣本是互相獨立的，但不需假設 (C)母體均數相等，因為這是現在需要檢定的。

37.（B）

變異數分析（analysis of variance）的均方誤（mean square error）項是 MSE，而 $E(MSE) = \sigma^2$，即共同母體變異數的估計值。

38.（D）

複判定係數（multiple coefficient of determination）為
$R^2 = \frac{SSR}{SSTO} = 1 - \frac{SSE}{SSTO}$，此題 $R^2 = 1 - \frac{240}{800} = 0.7$。

39.（A）

來源	平方和(SS)	自由度	均方和(MS)	F 值
迴歸	SSR	k	MSR= SSR/(k)	F=MSR/MSE
誤差	SSE	n-k-1	MSE=SSE/(n-k-1)	
總變異	SSTO	n-1		

對照本題，n = 30，k = 4，

來源	平方和(SS)	自由度	均方和(MS)	F 值
迴歸	700	4	MSR= SSR/(k)=700/4=175	F=MSR/MSE=175/4=**43.75**
誤差	100	25	MSE=SSE/(n-k-1)=100/25=4	
總變異	800	29		

40.（B）

	平方和(SS)	自由度	均方和(MS)	F 值
組間變異(處理)	SSTR	k-1	MSTR= SSTR/(k-1)	F=MSTR/MSE
組內變異(誤差)	SSE	n-k	MSE=SSE/(n-k)	
總變異	SSTO	n-1		

對照本題，n = n1+n2+n3+n4 = 20，k = 4，

變異來源	平方和(SS)	自由度	均方和(MS)	F 值
組間變異(處理)	6750	3	MSTR= SSTR/(k-1)=6750/3=2250	F =MSTR/MSE =2250/500 =4.5
組內變異(誤差)	8000	16	MSE=SSE/(n-k)=8000/16=500	
總變異	14750	19		