

104年公務人員初等考試試題

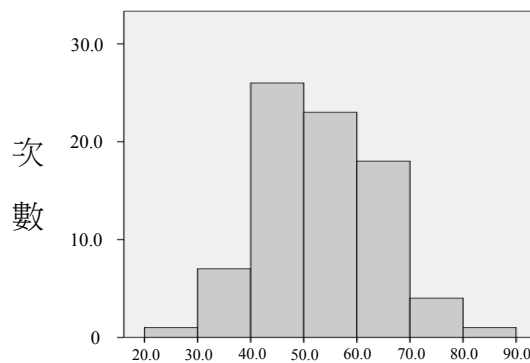
等 別：初等考試
類 科：統計
科 目：統計學大意

考試時間：1 小時

座號：_____

- ※注意：(一)本試題為單一選擇題，請選出一個正確或最適當的答案，複選作答者，該題不予計分。
(二)本科目共 40 題，每題 2.5 分，須用 2B 鉛筆 在試卡上依題號清楚劃記，於本試題上作答者，不予計分。
(三)可以使用電子計算器。

- 製作盒形圖 (box plot) 時，必須使用下列那一個統計量數？
(A)平均數 (mean) (B)中位數 (median)
(C)眾數 (mode) (D)第 90 百分位數 (90-th percentile)
- 某次統計學測驗成績的平均分數是 60，標準差為 10，則下列敘述何者正確？
(A)大約有 95% 的學生成績落於(40, 80) (B)大約有 75% 的學生成績落於(40, 80)
(C)至少有 95% 的學生成績落於(40, 80) (D)至少有 75% 的學生成績落於(40, 80)
- 一個包含正反面的公正的銅板投擲 3 次，現定義 $A = \{\text{出現正面奇數次}\}$ ， $B = \{\text{至少出現正面 1 次}\}$ ，則 $P(A \cup B)$ 機率值為多少？
(A)0.875 (B)0.75 (C)0.625 (D)0.5
- 若 E_1, E_2, E_3 為互補且互相排斥事件， $P(E_1) = 0.65$ ， $P(E_2) = 0.15$ ， $P(E_3) = 0.20$ ；另有事件 A 和 E_1, E_2, E_3 互相獨立，若 $P(A) = 0.4$ ，則 $P(E_1|A)$ 機率值為多少？
(A)0.80 (B)0.65 (C)0.35 (D)0.20
- 以隨機抽樣從兩獨立母體各抽出 100 個樣本單位，母體一標準差為 3，母體二標準差為 4，試求出兩樣本平均數差的標準差為多少？
(A)0.7 (B)0.6 (C)0.5 (D)0.4
- 某公司的員工能力測驗成績的直方圖如下圖，若以 \bar{x} 代表所有員工成績的平均，則員工能力測驗成績小於 \bar{x} 的比例為何？



- (A)小於 50% (B)大約 50% (C)大於 50% (D)不能確定
- 在一所學校將進行一個研究計畫，而該校有 1000 位學生，100 位教員與 200 位職員，若計畫訪問 100 位學生，10 位教員與 20 位職員，則最適合使用的機率抽樣方法為何？
(A)簡單隨機抽樣 (simple random sampling) (B)分層隨機抽樣 (stratified random sampling)
(C)便利抽樣 (convenience sampling) (D)判斷抽樣 (judgment sampling)
- 關於中央極限定理 (central limit theorem)，下列敘述何者正確？
(A)樣本大小 (sample size) 至少 30
(B)樣本平均等於母體平均
(C)樣本平均的變異數會隨著樣本數的增加而減少
(D)母體平均的近似分配是常態分配 (normal distribution)
- 某公司客服部平均每一分鐘接到 5 通電話，若以 Y 代表每一分鐘的電話數，且 Y 服從卜瓦松分配 (Poisson distribution)，則兩通電話之間隔時間 X 的機率分配為何？
(A)平均數為 5，且變異數為 5 的常態數分配
(B)平均數為 1/5，且變異數為 1/5 的常態分配
(C)平均數為 5 的指數分配 (Exponential distribution)
(D)平均數為 1/5 的指數分配
- 某實驗成功機率為 0.5，若獨立進行此實驗 15 次，且實驗成功次數服從二項分配，則實驗成功次數的眾數為多少？
(A)7 (B)8 (C)7, 8 (D)8, 9

- 11 由樣本資料得到常態母體平均數的 95%信賴區間 (confidence interval) 為 (100, 120)，若由同一筆資料得到母體平均數的 90%信賴區間為(a, b)，則下列敘述何者正確？
 (A) $100 < a < b < 120$ (B) $100 < a < 120 < b$ (C) $a < 100 < 120 < b$ (D) $a < 100 < b < 120$
- 12 隨機樣本 X_1, X_2, \dots, X_n ，其母體平均數 = μ ，若 $\bar{X} = \sum_{i=1}^n X_i/n$ ，下列關於 \bar{X} 的敘述，何者錯誤？
 (A) \bar{X} 是 μ 的不偏估計量 (unbiased estimator)
 (B) \bar{X} 是 μ 的點估計量 (point estimator)
 (C) \bar{X} 是 μ 的動差法估計量 (method-of-moments estimator)
 (D) \bar{X} 與 μ 的區間估計量之寬度有關
- 13 某研究者進行假設檢定得到的 p 值 (p-value) 是 0.02，若設定顯著水準 $\alpha = 0.05$ ，則下列有關該研究可能發生之錯誤的敘述，何者正確？
 (A) 僅可能發生第一類型錯誤 (Type I error)
 (B) 僅可能發生第二類型錯誤 (Type II error)
 (C) 可能同時發生第一類型和第二類型錯誤
 (D) 僅可能發生第一類型或第二類型錯誤之一，但不能確定何者會發生
- 14 黑色袋子中有 5 個彈珠，可能為透明或花色，若自袋中拿出 3 顆彈珠皆為花色，請以最大概似法 (maximum likelihood method) 估計黑色袋中花色彈珠的數量為多少？
 (A) 2 (B) 3 (C) 4 (D) 5
- 15 檢定力 (power) 指的是下列那一種敘述的機率？
 (A) 拒絕了錯誤的虛無假設 (null hypothesis) (B) 接受了正確的虛無假設
 (C) 拒絕了正確的虛無假設 (D) 接受了錯誤的虛無假設
- 16 某品牌罐裝牛奶的市場占有率是 20%，為了提升其市場占有率，該品牌的行銷部門提出一個促銷計畫，為了確認該促銷計畫是否確實提升其市場占有率，若以 p 代表市場占有率，則下列有關假設的敘述何者正確？
 (A) $H_0: p \leq 0.2$ ， $H_1: p > 0.2$ (B) $H_0: p < 0.2$ ， $H_1: p \geq 0.2$
 (C) $H_0: p \geq 0.2$ ， $H_1: p < 0.2$ (D) $H_0: p > 0.2$ ， $H_1: p \leq 0.2$
- 17 關於等距尺度 (interval scale)，下列敘述何者錯誤？
 (A) 具有順序尺度 (ordinal scale) 的性質 (B) 一定是數值 (numeric)
 (C) 可做加、減、乘、除的運算 (D) 沒有絕對零點
- 18 減重中心推出新快速減重專案，由於參與此專案成員體重差異極大，為了解其減重成效，試問以下述何種統計分析方法最適當？
 (A) 單一樣本 t 檢定 (B) 兩獨立樣本 t 檢定，假設變異數相等
 (C) 兩獨立樣本 t 檢定，假設變異數不相等 (D) 成對樣本 t 檢定
- 19 二因子變異數分析 (two-way ANOVA) 的統計模型為：

$$y_{ijk} = \mu + \alpha_i + \beta_j + (\alpha\beta)_{ij} + \varepsilon_{ijk}, i=1,2,\dots,a, j=1,2,\dots,b, k=1,2,\dots,n$$
 此模型之交互作用項 $(\alpha\beta)_{ij}$ 的自由度為何？
 (A) $a+b$ (B) $a+b-2$ (C) $a \times b$ (D) $(a-1) \times (b-1)$
- 20 假設由兩個常態母體 $N(\mu_1, \sigma^2)$ 和 $N(\mu_2, \sigma^2)$ 分別獨立抽出樣本大小為 n_1, n_2 的隨機樣本，得到樣本平均數和樣本標準差分別為： \bar{X}_1, \bar{X}_2 和 S_1, S_2 ，若母體標準差 σ^2 未知，則下列關於檢定 $H_0: \mu_1 = \mu_2$ ， $H_1: \mu_1 \neq \mu_2$ 的敘述何者錯誤？
 (A) $\bar{X}_1 - \bar{X}_2$ 是服從常態的分配
 (B) $\text{Var}(\bar{X}_1 - \bar{X}_2) = \frac{\sigma^2}{n_1} + \frac{\sigma^2}{n_2}$
 (C) 可以使用兩獨立樣本 t 檢定，其檢定統計量為 $\frac{\bar{X}_1 - \bar{X}_2}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$
 (D) 可以使用變異數分析法的 F 檢定，其分子與分母的自由度為 1, n_1+n_2-2

- 21 研究者要了解 4 種不同品牌的汽油之效能是否有所差異而提出一個研究計畫，由於研究者認為不同廠牌的汽車會有不同的效能，因此計畫選取 5 個不同廠牌的汽車各一輛，每一輛汽車都使用 4 種品牌的汽油各測試一次，並且紀錄每一次每 1 公升汽油所跑的里程數，則下列何者是最適合的設計方式？
(A)完全隨機化設計 (completely randomized design) (B)隨機化區集設計 (randomized block design)
(C)二因子設計 (two-factor design) (D)獨立樣本設計 (independent samples design)
- 22 下表是一個隨機化區集設計 (randomized block design) 的變異數分析 (ANOVA) 表，則下列何者最接近檢定不同處理間的平均是否完全相等的 F 值？

變異來源	平方和 (Sum of Squares)	自由度 (degree of freedom)	均方和 (Mean Square)	F
處理間 (Between Treatments)	153	3		
區集間 (Between Blocks)		4	48	
誤差 (Error)	120	12	10	

- (A)1.06 (B)1.28 (C)4.8 (D)5.1
- 23 隨機抽樣 10 個樣本單位，其特性值為 0, 1, -1, 4, -3, 6, -5, 10, -10, 15，試求其四分位距 (interquartile range, IQR) 為多少？
(A)9 (B)10 (C)11 (D)12
- 24 甲、乙兩家上市公司前 30 天股價平均數為 20 元及 100 元，則欲比較此兩家上市公司股價變動分散情形，則以何統計量最適當？
(A)變異係數 (coefficient of variation) (B)標準差 (standard deviation)
(C)全距 (range) (D)變異數 (variance)
- 25 某組數據呈現平均數為 10，變異數為 4 的常態分配，依據經驗法則 (empirical rule)，則數據小於 6 的機率約為多少？
(A)0.0013 (B)0.0026 (C)0.0114 (D)0.0228
- 26 假設兩變數間的相關係數為 r ，且估計迴歸方程式 (regression equation) 為： $\hat{y} = \hat{\beta}_0 + \hat{\beta}_1 x$ ，則下列敘述何者錯誤？
(A) $r = 0$ ，則 $\hat{\beta}_1 = 0$ (B) r 與 $\hat{\beta}_1$ 有相同的符號
(C) r 愈大，則 $\hat{\beta}_1$ 愈大 (D) 判定係數 (coefficient of determination) = r^2
- 27 關於使用卡方分配作獨立性檢定，下列敘述何者不恰當？
(A)可檢定兩母體兩變數間是否相互獨立
(B)其自由度為 (列個數-1) 乘以 (行個數-1)
(C)計算各格子卡方貢獻值之分母應使用理論次數
(D)各格子理論次數等於所對應之列總和乘以行總和除以樣本數
- 28 研究者想用適合度檢定法檢查下列資料是否符合卜瓦松分配的假設，若不考慮合併資料，則該檢定的自由度為何？

觀察值	≤ 1	2	3	4	5	6	7	8	≥ 9
次數	12	12	14	20	24	24	18	14	8

- (A)9 (B)8 (C)7 (D)6
- 29 根據下列各月份的時間數列值 (time series value)，使用 3 個月移動平均 (moving average) 預測，則六月份的預測值應為何？

月份	一月	二月	三月	四月	五月	六月
時間數列值	39	44	56	59	74	86

- (A)73 (B)63 (C)53 (D)46.3
- 30 以 19 個時間點的數列資料估計得到的線性趨勢為： $T_t = 34 + 1.5t$ ，則預測第 20 個時間點的趨勢值 (trend value) 為何？
(A)34 (B)61 (C)62.5 (D)64

- 31 假設 $P(A|B)=0.05$ ， $P(A)=0.15$ ， $P(B)=0.6$ ，則 $P(A^c|B^c)$ 最接近下列那一個數值？
 (A)0.05 (B)0.70 (C)0.85 (D)0.95
- 32 某一位發展心理學者想了解在不同年齡（5 歲、8 歲、11 歲）的兒童，對於某種特定音樂的喜好是否有所不同，若以 p_{ij} 代表聯合機率（joint probability）， $p_{j|i}$ 代表第 j 行在第 i 列的條件機率（conditional probability）（如下表所示），則本研究的虛無假設為何？

年齡	喜歡	不喜歡
5 歲	$p_{11}(p_{1 1})$	$p_{12}(p_{2 1})$
8 歲	$p_{21}(p_{1 2})$	$p_{22}(p_{2 2})$
11 歲	$p_{31}(p_{1 3})$	$p_{32}(p_{2 3})$

- (A) $p_{11} = p_{21} = p_{31}$ ，且 $p_{12} = p_{22} = p_{32}$ (B) $p_{11} = p_{12}$ ， $p_{21} = p_{22}$ ，且 $p_{31} = p_{32}$
 (C) $p_{1|1} = p_{1|2} = p_{1|3}$ ，且 $p_{2|1} = p_{2|2} = p_{2|3}$ (D) $p_{1|1} = p_{2|1}$ ， $p_{1|2} = p_{2|2}$ ，且 $p_{1|3} = p_{2|3}$
- 33 一因子變異數分析（one-way ANOVA）的統計模型： $y_{ij} = \mu_i + \varepsilon_{ij}$ ， $i=1,2,\dots,k$ ， $j=1,2,\dots,n$ ， $\varepsilon_{ij} \sim N(0, \sigma^2)$ ，若 $\bar{y}_i = \frac{\sum_{j=1}^n y_{ij}}{n}$ ， $\bar{y}_+ = \frac{\sum_{i=1}^k \sum_{j=1}^n y_{ij}}{nk}$ ， $MSB = \frac{\sum_{i=1}^k n(\bar{y}_i - \bar{y}_+)^2}{k-1}$ ， $MSE = \frac{\sum_{i=1}^k \sum_{j=1}^n (y_{ij} - \bar{y}_i)^2}{nk-k}$ ，則 $E(MSB)$ 與 $E(MSE)$ 的關係為何？

- (A) $E(MSB) = E(MSE)$ (B) $E(MSB) \geq E(MSE)$
 (C) $E(MSB) \leq E(MSE)$ (D) 不能確定 $E(MSB)$ 與 $E(MSE)$ 的大小關係
- 34 簡單線性迴歸關係為： $y = \beta_0 + \beta_1 x + \varepsilon$ ，若 ε 具常態分配，則使用於 β_1 之區間估計的機率分配為何？
 (A) 常態分配 (B) T 分配 (C) F 分配 (D) 卡方分配
- 35 簡單線性迴歸關係為： $y = \beta_0 + \beta_1 x + \varepsilon$ ，且 $e = y - \hat{y}$ ，則下列關於殘差分析的敘述何者錯誤？
 (A) 若 e 的絕對值大於 3，則可懷疑是離群點（outlier）
 (B) 若 e 的常態機率圖不是一直線，則可能 ε 並非服從常態分配
 (C) 若 e 與 x 的散佈圖（scatter plot）呈現喇叭狀，則可能 ε 的變異數並不相等
 (D) 若 e 與 x 的散佈圖（scatter plot）呈現曲線狀，則可能兩變數間並非線性關係
- 36 以隨機抽樣從兩獨立母體進行抽樣，若關於兩母體的比例皆未知，估計比例之差的容許偏差（margin of error）為 0.025，使用顯著水準 0.10，假設從兩獨立母體抽出的樣本數相同，則總共需抽出多少樣本？（註： $Z_{0.10}=1.28$ ， $Z_{0.05}=1.645$ ）
 (A) 656 (B) 1084 (C) 2622 (D) 4330

- 37 某個抽菸與 3 種疾病之關聯性的研究結果如下表，則下列檢定抽菸與疾病是否有關連的敘述何者錯誤？

疾病	沒抽菸	有抽菸	總和
肺氣腫	20	60	80
心血管疾病	70	40	110
癌症	30	80	110
總和	120	180	300

- (A) H_0 ：抽菸與疾病有關， H_1 ：抽菸與疾病無關 (B)（癌症，沒抽菸）的期望個數 = 44
 (C) 檢定統計量的值 = 40.53 (D) 卡方檢定的自由度 = 2
- 38 以 N ， n 和 p 分別代表母體大小，樣本大小和母體比例， \hat{p} 代表樣本比例，則下列敘述何者是正確的？
 (A) $E(\hat{p}) = p$ (B) $\text{Var}(\hat{p}) = \frac{p(1-p)}{n}$
 (C) \hat{p} 服從二項分配（Binomial distribution） (D) $n \geq 30$ 時， \hat{p} 近似常態分配
- 39 某變異數分析（ANOVA）之實驗設計包含一因子，A 分為 A1, A2, A3, A4 及 A5 五個水準，今以隨機抽樣抽取 16 個樣本單位，經由一因子變異數分析，得到因子平均變動（MSF, Mean Square for Factors）= 10，F 值 = 5；試問判定係數（R-Sq, R^2 ）為多少？
 (A) 0.323 (B) 0.355 (C) 0.600 (D) 0.645
- 40 承上題，試求誤差項變動（SSE, Sum of Squares for Error）為多少？
 (A) 11 (B) 20 (C) 22 (D) 40