※ 注意：請於試卷上「非選擇題作答區」標明題號並依序作答。

**是非題**（每題2分：O 或 X）

1. 統計考驗的顯著水準乃指虛無假設為真的機率。
2. (弱)大數法則探討的是樣本平均數的收斂性質。
3. 先把兩變項標準化之後再算共變數，或是直接算兩變項的共變數，所得結果相同。
4. 先把兩變項標準化之後再算相關係數，或是直接算兩變項的相關係數，所得結果相同。
5. 先把兩變項標準化之後再算迴歸係數，或是直接算兩變項的迴歸係數，所得結果相同。
6. 簡單迴歸中，雙側的相關係數檢定 ($H_0$: $\rho = 0$) 與雙側的斜率檢定 ($H_0$: $\beta = 0$) 是等價的。
7. Spearman's correlation coefficient 與 Kendall's tau 皆是根據排序資料算得。
8. Logistic 迴歸適用於當 response（或稱 predicted）variable 為次序變項時。
9. Matched-sample case 中的 Wilcoxon's test 與 sign test 相比，多考量了排序大小之值。
10. 估計量的「一致性」指的是當樣本數趨近於無限大時，估計量會收斂至真實參數值。

**單選題**（每題4分）

11. 有些參加心理系轉學考的學生會來本系旁聽心統。某機構調查發現，有來本系旁聽心統的學生中，有 90% 考上本系；而沒來本系旁聽心統的學生，則只有 2% 考上本系。根據歷年資料，已知有來本系旁聽心統的學生佔總考生的 3%。現有一學生轉學考進了本系，請問她有來本系旁聽心統的機率是多少？
(A) $(0.9)(0.03)$ (B) $\frac{(0.9)(0.03)}{(0.9)(0.03)+(0.1)(0.98)}$ (C) $\frac{(0.9)(0.03)}{(0.9)(0.03)+(0.1)(0.97)}$
(D) $\frac{(0.9)(0.03)}{(0.9)(0.03)+(0.02)(0.98)}$ (E) $\frac{(0.9)(0.03)}{(0.9)(0.03)+(0.02)(0.97)}$

12. 我們隨機抽樣 30 位學生問他們身高，得一平均值。然後將每位學生身高值減去平均值，再將之平方，最後累加。想像你重覆此實驗 10000 次，得一分配。請問此分配與何種分配有密切關係？
(A) 卡方分配 (B) $t$ 分配 (C) 常態分配 (D) $F$ 分配 (E) 二項分配

13. 考慮一 $n = 150$，$p = 0.4$ 的二項分配。請問此分配 0.16 quantile 的值介於下列何區間？
(A) 10 至 20 (B) 20 至 30 (C) 30 至 40 (D) 40 至 50 (E) 50 至 60

14. 令 $X_i$ ($i = 1, \ldots, n$) 為從常態分配（平均數為 $\mu$，變異數為 $\sigma^2$）抽出的 $n$ 個獨立樣本。請問 $\mathrm{Cov}(X_i, \bar{X})/\sigma^2$ 之值為何？（註：$\bar{X} = \frac{\sum X_i}{n}$）
(A) 0 (B) 1 (C) $n$ (D) $1/n$ (E) $1/n^2$

15. 考慮一 population level 之模型方程式：$\mu_{jk} = \mu + \alpha_j + \beta_k + (\alpha\beta)_{jk}$，其中 $\mu_{jk}$ 為 cell means，$\mu$ 為總平均值，$\alpha_j$ 代表 Factor A 的 treatment effect，$\beta_k$ 代表 Factor B 的 treatment effect，$(\alpha\beta)_{jk}$ 為交互作用項。若 A、B 各有兩 levels ($A_1$, $A_2$ 及 $B_1$, $B_2$)，且其 cell means 值如下：

| A \ B | $B_1$ | $B_2$ |
|---|---|---|
| $A_1$ | 1 | 3 |
| $A_2$ | 5 | 3 |

請問 $(\alpha\beta)_{12}$ 值為何？
(A) 1 (B) 2 (C) 3 (D) 4 (E) 5

見背面

16. 你將一筆實驗資料根據某離散分配假設，做單側考驗得一 $p$-value。請問以下式子何者為真？（註：P(X|Y) 代表在 Y 條件下，X 之機率。）
(A) $p$-value ≤ P(資料|$H_0$)　(B) $p$-value ≥ P(資料|$H_0$)　(C) $p$-value = P(資料|$H_0$) P($H_0$)
(D) $p$-value ≤ 2(1−P(資料|$H_0$))　(E) $p$-value ≥ 2(1−P(資料|$H_0$))

17. 考慮二因子受試者間設計及其 ANOVA。已知 Factor A 的 degrees of freedom (df) 為 3，交互作用項的 df 為 3，殘差的 df 為 24。請問此實驗總共有多少位受試者參與？
(A) 34　(B) 33　(C) 32　(D) 31　(E) 30

考慮一筆資料 $(x_1, y_1), \ldots, (x_n, y_n)$ 及其迴歸式：$Y_i = \alpha + \beta x_i + \varepsilon_i$，其中$\varepsilon_i$為誤差項。請回答以下三題。（註：$\bar{x} = \frac{\sum x_i}{n}$，$\bar{Y} = \frac{\sum Y_i}{n}$，$\hat{Y}_i$ 代表 $Y_i$ 的估計值。）

18. 簡單迴歸的最小平方法是要 minimize 以下哪個式子？
(A) $\sum(Y_i - (\alpha + \beta x_i + \varepsilon_i))^2$　(B) $\sum(Y_i - (\alpha + \beta x_i))^2$　(C) $\sum(\bar{Y} - (\alpha + \beta x_i + \varepsilon_i))^2$
(D) $\sum(\bar{Y} - (\alpha + \beta x_i))^2$　(E) $\sum(Y_i - \bar{Y} - (\alpha + \beta x_i + \varepsilon_i))^2$

19. 簡單迴歸常提及 SSTotal = SSError + SSRegression。請問 SSRegression 的定義為何？
(A) $\sum(Y_i - \hat{Y}_i)^2$　(B) $\sum(Y_i - \bar{Y})^2$　(C) $\sum(\bar{Y} - \hat{Y}_i)^2$　(D) $\sum(Y_i - \hat{Y}_i - \bar{Y})^2$　(E) $\sum(\hat{Y}_i - Y_i - \bar{Y})^2$

20. 以下哪個式子 (直觀上) 最能用來解釋 regression to the mean？（註：$r$為樣本相關係數）
(A) SSRegression = $r^2$ SSTotal　(B) $\hat{\beta} = \frac{\sum[(x_i-\bar{x})(Y_i-\bar{Y})]}{\sum(x_i-\bar{x})^2}$　(C) $r = \frac{\sum[(x_i-\bar{x})(Y_i-\bar{Y})]}{\sqrt{\sum(x_i-\bar{x})^2}\sqrt{\sum(Y_i-\bar{Y})^2}}$
(D) $\frac{\hat{Y}_i-\bar{Y}}{\sqrt{\frac{\sum(Y_i-\bar{Y})^2}{n-1}}} = r\frac{x_i-\bar{x}}{\sqrt{\frac{\sum(x_i-\bar{x})^2}{n-1}}}$　(E) $1 - r^2$ = SSError/SSTotal

**複選題**（每題 4 分）

21. 令 $F$ 檢定之臨界值為 $F_{\alpha;m,n}$（$m$為分子自由度，$n$為分母自由度，$\alpha$為第一類型錯誤率）。請問 $t$ 檢定之臨界值 $t_{\alpha;5}$（5為自由度，$\alpha$為第一類型錯誤率）等於以下何者？
(A) $\sqrt{F_{\alpha;1,5}}$　(B) $\sqrt{F_{2\alpha;1,5}}$　(C) $\sqrt{\frac{1}{F_{1-\alpha;1,5}}}$　(D) $\sqrt{\frac{1}{F_{1-2\alpha;1,5}}}$　(E) $\sqrt{\frac{1}{F_{1-2\alpha;5,1}}}$

22. 我們根據 two-sample $t$考驗分析資料，得一平均值之 99% 信賴區間 [-0.21, -0.05]。請問以下解讀何者有誤？
(A) 真實平均值有 99% 機率座落於此區間。
(B) 我們有 99% 的信心真實平均值是介於 -0.21 與 -0.05。
(C) 若我們重覆此實驗 100 次，大概有 99 次真實平均值會座落於此區間。
(D) 真實平均值為 0 的機率是 1%。
(E) 虛無假設被拒絕的機率為 0.01。

23. 統計檢定力受到以下哪些因素影響？
(A) 效果量　(B) 殘差　(C) 樣本數　(D) 第一類型錯誤率　(E) 虛無假設被拒機率

接次頁

24. 令$\widehat{\theta_m}$為根據 $m$ 個樣本所得之參數$\theta$的估計量。以下敘述何者為真？（註：E(X)、Var(X)分別代表 X 的期望值及變異數。）
(A) $E[(\theta-\widehat{\theta_m})^2]=\mathrm{Var}(\widehat{\theta_m})$　(B) $E[(\theta-\widehat{\theta_m})^2]\geq\mathrm{Var}(\widehat{\theta_m})$　(C) $E[(\theta-\widehat{\theta_m})^2]\leq\mathrm{Var}(\widehat{\theta_m})$
(D) $E(\widehat{\theta_m})-\theta=0$　(E) $E[(\theta-\widehat{\theta_m})^2]$可用來評量估計量的「有效性」

25. 考慮一因子受試者間設計及其 ANOVA，共有 4 個 treatment levels，每一 level 有 15 位受試者參與。已知 4 組的樣本變異數各為 10, 7, 8, 10；另外，4 組的平均值各為 1, 2, 4, 5。請據此算出 SSWG 及 SSBG。（註：SSWG 代表 within-group sum of squared deviation；SSBG 代表 between-group sum of squared deviation）
(A) SSWG = (14)(35)/4　(B) SSWG = (14)(35)　(C) SSWG = 35
(D) SSBG = (15)(10)　(E) SSBG = 10

26. 下列何者可 (直接) 套用中央極限定理解釋？
(A) 樣本數夠大時，某樣本平均數的分配近似於常態分配。
(B) 樣本數夠大時，某樣本變異數的分配近似於常態分配。
(C) 樣本數夠大時，某樣本平均數除上樣本標準差而得之分配近似於標準化常態分配。
(D) 當 degrees of freedom 夠大時，$t$ 分配近似於標準化常態分配。
(E) 當 degrees of freedom 夠大時，卡方分配近似於常態分配。

令 $Y_{ij}=\mu+\alpha_j+\epsilon_{ij}$（其中$Y_{ij}$為觀察值，μ為總平均值，$\alpha_j$為 treatment effect，$\epsilon_{ij}$為誤差項）為一因子受試者間設計的 ANOVA (fixed-effects) 模型方程式。請回答以下兩題。

27. 以下哪些屬模型假設？（註：$N(m, v^2)$代表平均數為$m$、變異數為$v^2$之常態分配。）
(A) $\alpha_j\sim N(0,\sigma_\alpha{}^2)$　(B) $\sum_{j=1}^{J}\alpha_j=0$　(C) For all $j$, $\sum_{i=1}^{I}\epsilon_{ij}=0$　(D) $\epsilon_{ij}\sim N(0,\sigma_\varepsilon{}^2)$
(E) $\sum_{j=1}^{J}\sum_{i=1}^{I}\epsilon_{ij}=0$

28. 以下哪些在虛無假設為真的情況才成立？（MSWG 代表 within-group mean square；MSBG 代表 between-group mean square）
(A) $\sum_{j=1}^{J}\alpha_j^2=0$　(B) $\mathrm{Var}(\alpha_j)=1$　(C) $\mathrm{Var}(\epsilon_{ij})=1$　(D) E(MSWG) = E(MSBG)
(E) MSBG/MSWG 為 (ordinary) $F$ 分配

令 $Y_{ijk}=\mu+\alpha_j+\beta_k+\alpha\beta_{jk}+\epsilon_{i(jk)}$為兩因子受試者間設計的 ANOVA (fixed-effects) 模型方程式，其中$Y_{ijk}$為觀察值，μ為總平均值，$\alpha_j$代表 Factor A 的 treatment effect，$\beta_k$代表 Factor B 的 treatment effect，$\alpha\beta_{jk}$為交互作用項，$\epsilon_{i(jk)}$為誤差項。請回答以下兩題。

29. 以下哪些為隨機變數？
(A) $Y_{ijk}$　(B) $\alpha_j$　(C) $\beta_k$　(D) $\alpha\beta_{jk}$　(E) $\epsilon_{i(jk)}$

30. 以下哪些屬模型假設？
(A) $\mathrm{Var}(\epsilon_{i(jk)})=1$　(B) $\sum_{j=1}^{J}\alpha_j=0$　(C) $\sum_{k=1}^{K}\beta_k=0$
(D) For all $j$ & $k$, $\sum_{i=1}^{I}\epsilon_{i(jk)}=0$　(E) $\sum_{j=1}^{J}\sum_{k=1}^{K}\alpha\beta_{jk}=0$

試題隨卷繳回