《統計學》

試題評析	這次地特三等經建行政的統計學，應該是近幾年國考統計學最簡單的一份，即使同學很快完成答題，也一定要再三檢查，把握好分數。題目有四大題，第一大題也是簡單二項機率分配應用，剩下的二到四題是基本題型，應用統計學5大重點就考了3大題。

一、某一箱子有 20 顆球，白球 10 顆，黑球 10 顆。今以取出放回的方式從箱內隨機抽取 3 球。令 X 代表白球的個數。

(一) 求 $X=2$ 的機率，即 $P(X=2)$。（5 分）

(二) 求 X 的動差母函數，即 $E[e^{tX}]$。（5 分）

(三) 求 $E[X^3]$。（5 分）

(四) 若隨機變數 Y 與 X 獨立且兩者有相同的機率分配，求 $E[(X-Y)^3]$。（10 分）

考點命中	《高點・高上統計學講義》第三回，4-1一維常用分配離散型，蘇建郎編撰，頁5。

答：

(一) $X\sim B(n=3,p=\frac{1}{2})$ ， $f_X(x)=C_x^3(\frac{1}{2})^3$, $x=0,1,2,3,$

$$P(X=2)=C_2^3(\frac{1}{2})^2(\frac{1}{2})=\frac{3}{8}$$

(二) $M_X(t)=E(e^{tX})=\sum_{x=0}^{3}e^{tx}f_X(x)=\sum_{x=0}^{3}e^{tx}C_x^3(\frac{1}{2})^x(\frac{1}{2})^{3-x}=\sum_{x=0}^{3}C_x^3(\frac{1}{2}e^t)^x(\frac{1}{2})^{3-x}=(\frac{1}{2}+\frac{1}{2}e^t)^3,t\in\mathbb{R}$

(三) $E(X^3)=0^3\times\frac{1}{8}+1^3\times\frac{3}{8}+2^3\times\frac{3}{8}+3^3\times\frac{1}{8}=\frac{54}{8}=6.75$

(四) $E[(X-Y)^3]=E(X^3-3X^2Y+3XY^2-Y^3)$

$$=E(X^3)-3E(X^2)E(Y)+3E(X)E(Y^2)-E(Y^3)=0$$

二、設隨機變數 X 服從常態分配，具有平均數 μ 未知，變異數 4，從該母體 X 取出一組隨機樣本，有 5 個觀察值，數值如下：2, 3, 4, 4, 2。現進行 H_0：$\mu=1.4$，H_a：$\mu>1.4$ 之檢定。

(一) 試問該組樣本（前述觀測值樣本）平均數所對應的觀測值的顯著水準（亦稱 p 值）？（10 分）

(二) 在型一誤差機率為 0.05 之下，試問接受或拒絕虛無假設 H_0？（5 分）

(三) 使用型一誤差機率為 0.05 的拒絕域，求 H_a：$\mu=2.645$ 的檢定力。（10 分）

考點命中	1.《高點・高上統計學講義》第五回，7-2檢定統計量法決策準則，蘇建郎編撰，頁30。 2.《高點・高上統計學講義》第五回，7-3臨界值法、樣本數決定，蘇建郎編撰，頁52。

答：

(一) $n=5$, $\bar{x}=3$, $\sigma^2=4$ ， $z^*=\dfrac{\bar{x}-\mu_0}{\sigma/\sqrt{n}}=\dfrac{3-1.4}{2/\sqrt{5}}=1.79$

$$p-value=P(Z>z^*\,|\,H_0)=P(Z>1.79)=0.0337$$

(二) $p-value=0.0337<0.05\ \Rightarrow\ reject\ H_0$

(三)拒絕域$\{\underset{\sim}{X} \mid \bar{X} > k\}$ $\quad k = 1.4 + 1.96\dfrac{2}{\sqrt{5}} = 3.153$

此檢定的拒絕域為$C = \{\underset{\sim}{x} \mid \bar{x} > 3.153\}$

檢定力$1-\beta = P_1(\bar{X} > 3.153) = P(\bar{X} > 3.153 \mid \mu = 2.645)$

$$= P(Z > \frac{3.153-2.645}{2/\sqrt{5}}) = P(Z > 0.57) = 0.2843$$

三、今欲比較某作物四種不同品種平均產量差異比較，分別在三個不同地區進行，每個地區種植順序完全隨機安排，記錄其產量，共得 12 筆資料，資料符合變異數分析（ANOVA）模型之假設條件，經由分析結果，摘錄於以下 ANOVA 表：

變異來源	df	SS	MS	F
品種	（A）	（D）	0.222	（G）
地區	（B）	60.667	（F）	（H）
殘差	（C）	（E）		
總計		72.667		

(一) 完成上述 ANOVA 表內（A）至（H）格的數值。（8 分）

(二) 在型一誤差$\alpha = 0.05$之下，試問該作物四種不同品種平均產量是否有顯著差異？（8 分）

(三) 在型一誤差$\alpha = 0.05$之下，試問該作物在三個不同地區的平均產量是否有顯著差異？（9 分）

考點命中	《高點・高上統計學講義》第七回，8-3 隨機集區設計，蘇建郎編撰，頁11。

答：

(一)

Source	*df*	*SS*	*MS*	*F*
品種	(A)3	(D)0.6666	0.222	(G)0.1175
地區	(B)2	60.667	(F)30.333	(H)16.0577
殘差	(C)6	(E)11.334	1.889	
總計	11	72.667		

(二)

①H_0:四個不同品種平均產量無差異

H_1:四個不同品種平均產量有差異

②由(一)ANOVA表可知，$F^* = 0.1175$

③$\alpha = 0.05$，$C = \{F : F > F_{0.05}(3,6) = 4.76\}$

$F^* = 0.1175 \notin C$，*don't reject* H_0

④在$\alpha = 0.05$情況下，我們沒有顯著證據說四個不同品種平均產量有差異。

(三)

①H_0:作物在三個不同地區平均產量無差異

H_1 : 作物在三個不同地區平均產量有差異

②由(一)ANOVA表可知， $F^* = 16.0577$

③ $\alpha = 0.05$ ， $C = \{F : F > F_{0.05}(2,6) = 5.14\}$

$F^* = 16.0577 \in C$ ， $reject\ H_0$

④在 $\alpha = 0.05$ 情況下，我們有顯著證據說作物在三個不同地區平均產量有差異。

四、考慮簡單線性迴歸反應變數模型 $y_i = \beta_0 + \beta_1 x_i + \varepsilon_i$ ，若誤差 ε_i 是常態分配，平均數是 0，標準差 σ 未知。解釋變數 x 與反應變數 y 的 5 個觀測值 $(x_1, y_1), \cdots, (x_5, y_5)$ ，經計算得到 $\sum_{i=1}^{5} x_i = 15$ ， $\sum_{i=1}^{5} y_i = 20$ ， $\sum_{i=1}^{5} x_i y_i = 69$ ， $\sum_{i=1}^{5} x_i^2 = 55$ ， $\sum_{i=1}^{5} y_i^2 = 90$ 。

(一) 求 x 與 y 的樣本相關係數。（8 分）

(二) 以最小平方法求 β_1 的估計值。（7 分）

(三) 若 β_1 估計式（estimator）的標準誤是 0.253，求 β_1 的 95%信賴區間。（10 分）

考點命中	《高點・高上統計學講義》第八回，9-2簡單線性迴歸分析，蘇建郎編撰，頁18。

答：

(一) $\bar{x} = 3$, $\bar{y} = 4$ ， $SS_{xx} = \sum (x_i - \bar{x})^2 = \sum x_i^2 - n\bar{x}^2 = 10$

$$SS_{xy} = \sum (x_i - \bar{x})(y_i - \bar{y}) = \sum x_i y_i - n\overline{xy} = 9$$

$$SS_{yy} = \sum (y_i - \bar{y})^2 = \sum y_i^2 - n\bar{y}^2 = 10$$

樣本相關係數 $r_{X,Y} = \dfrac{SS_{xx}}{\sqrt{SS_{xx} SS_{yy}}} = \dfrac{9}{\sqrt{10 \times 10}} = 0.9$

(二) $\hat{\beta}_1 = \dfrac{SS_{xy}}{SS_{xx}} = \dfrac{9}{10} = 0.9$

(三) $t_{0.025}(3) = 3.182$

故 $\hat{\beta}_1 \pm t_{0.025}(3) \times S(\hat{\beta}_1) \Rightarrow 0.9 \pm 3.182 \times 0.253 \Rightarrow (0.095, 1.705)$

為 β_1 之95%信賴區間。