《統計學概要》

試題評析	本年度考題雖然有50分的基本觀念題，但是第一題(四)、第三題(三)難度超出統計學概要的範圍，再加上第三題(二)無解、第二題(四)需要轉個彎，經建類科考生書寫起來應該會覺得稍微困難。預估一般考生成績大概落在50-60之間，若能做對70分則足以達到上榜的水準。
考點命中	(一)《高點統計講義》第一回，秦大成編撰，頁46例題6。 (二) (三)《高點迴歸講義》第二回，秦大成編撰，頁107例題1。 (四)《高點迴歸講義》第三回，秦大成編撰，頁81實例說明。

一、某工廠僱有男女員工2,000人，其中男性員工1,500人，餘者為女性員工。廠長欲考察該廠男性員工之薪資水準，抽取100名男性員工，發現其薪資分配為：

每月薪資（千元）	員工人數
4～6	5
6～8	20
8～10	30
10～12	30
12～14	15

廠長認為「性別歧視」可能是造成男女員工薪資差異的主要因素，故聘請李教授協助研究分析，其設立迴歸模型如下：

$W_i = \alpha + \beta S_i + \varepsilon_i$

式中，W代表個人薪資，S為代表性別之虛擬變數（Dummy Variable），男性為0，女性為1。且另抽取100名女性員工，得到平均薪資為9,000元，標準差為100元。

試問：

（一）求男性員工之平均數及中位數。（8分）

（二）α, β 代表何種含義？（8分）

（三）試求 α, β 之估計值。（8分）

（四）若某研究生將李教授所設定之模型改為 $W_i = \alpha + \beta F_i + \gamma M_i + \varepsilon_i$，其中F表女性資料，M為男性數據，請問如此設定會發生何問題？（6分）

答：

(一)

每月薪資(千元)	組中點	男性員工人數
4-6	5	5
6-8	7	20
8-10	9	30
10-12	11	30
12-14	13	15

男性員工：

(1)平均數 $\bar{X} = \frac{\sum_{i=1}^{k} x_i f_i}{n} = \frac{960}{100} = 9.6$（千元）

(2)中位數 $\mathrm{Me} = L_3 + \frac{\frac{n}{2} - C_2}{f_3} \times h = 8 + \frac{50-25}{30} \times 2 \doteq 9.6667$（千元）

(二)β：代表女、男性之間個人平均薪資差值

α：代表男性個人平均薪資

(三)$\begin{cases}\widehat{W}_{S=0}=\hat{\alpha}+(\hat{\beta}\cdot 0)=9.6\\ \widehat{W}_{S=1}=\hat{\alpha}+(\hat{\beta}\cdot 1)=9\end{cases} \Rightarrow \begin{cases}\hat{\alpha}=9.6\\ \hat{\beta}=-0.6\end{cases}$

(四)$\because F+M=1 \Rightarrow$ F與M兩個變數之間完全共線

$\therefore$ OLSE：$\hat{F}$與$\hat{M}$無限解，且 $Var(\hat{F})\to\infty$，$Var(\hat{M})\to\infty$

使得迴歸模型完全沒有解釋能力

二、某品牌手機廣告宣稱其手機的待機時間平均為10天，及變異數為4天左右。若消基會從中抽出64支，以便檢驗該公司的廣告是否真實？

（一）請建立檢定假設。（10分）

（二）若在$\alpha=0.025$下，消基會發現平均待機時間約為9天，試問該品牌手機的待機時間是否符合標準？（10分）

（三）求平均待機時間為9天的檢定力。（10分）

（四）如果廠商希望9.5天能符合規定，則廠商可建議消基會最多選擇多少的顯著水準（α）去檢定？（10分）

考點命中	《高點統計講義》第五回，秦大成編撰，頁60例題1。

答：

(一)$\begin{cases}H_0：\mu\geq 10\\ H_1：\mu<10\end{cases}$

(二)$n=64\geq 30$ 採C.L.T.查Z表

$\because Z=\frac{\bar{X}-\mu_0}{\sigma/\sqrt{n}}=\frac{9-10}{2/\sqrt{64}}=-4<-Z_{0.025}=-1.96$

$\therefore$ reject H_0，有充分證據顯示該品牌手機待機時間不符合標準

(三)$C=\left\{\bar{X}\,\middle|\,\bar{X}<\mu_0-Z_{0.025}\cdot\frac{\sigma}{\sqrt{n}}=10-1.96\cdot\frac{2}{\sqrt{64}}=9.51\right\}$

power $=P(\bar{X}<9.51 \mid \mu=\mu_1=9\in H_1)$

$=P\left(Z<\frac{9.51-9}{2/\sqrt{64}}\right)=P(Z<2.04)=0.9793$

(四)$\because Z=\frac{\bar{X}-\mu_0}{\sigma/\sqrt{n}}=\frac{9.5-10}{2/\sqrt{64}}=-2\geq -Z_{\alpha}$

當 $Z_{\alpha}=2$ 時 $\Rightarrow \alpha=1-0.9772=0.0228$

$\therefore$ 最多只能選擇 $\alpha=0.0228$

三、設X為一服從二項分配（Binomial(n, p)）隨機變數，且已知$E(X)=1.2$，$Var(X)=0.84$。

（一）試求n與p。（10分）

（二）試問需做多少次貝努利（Bernoulli）試驗才可使其標準差等於期望值之百分之七十？

（10分）

（三）另有一與X互為獨立之二項分配Y～Binomial(3, 0.4)，若令 Z＝X＋Y ，試求 P(Z＝1) ？（10分）

考點命中	(一) (二)《高點統計講義》第三回，秦大成編撰，頁8例題4。 (三)《高點統計講義》第二回，秦大成編撰，頁98例題1。

答：

(一)$\begin{cases} EX = np = 1.2 \\ V(X) = np(1-p) = 0.84 \end{cases} \Rightarrow \begin{cases} p = 0.3 \\ n = 4 \end{cases}$

(二)$\sigma = 0.7 \cdot EX \Rightarrow \sqrt{np(1-p)} = 0.7 \cdot (np)$

$p = 0.3$

$\Rightarrow \sqrt{n \cdot 0.3 \cdot (1-0.3)} = 0.7 \cdot (n \cdot 0.3)$

$\Rightarrow 0.21n(0.21n-1) = 0 \quad \Rightarrow n = 0$(不合) 或 $n = \dfrac{1}{0.21} \notin$ **整數**

∴ **本題無解**

(三)$X \sim Bin(n=4, p=0.3)$， $Y \sim Bin(n=3, p=0.4)$

雖然X，Y獨立，但P不相等 ∴無法做二項式加成性：$X + Y \sim Bin(n, p)$

因此，本題必須用變數變換求$P(Z=1)$，已超出統計學概要的範圍

$f(x,y) = f(x) \cdot f(y) = (C^4_x 0.3^x \cdot 0.7^{4-x})(C^3_y 0.4^y \cdot 0.6^{3-y})$

$x = 0,1,2,3,4 \quad , \quad y = 0,1,2,3$

$z = x + y \quad , z = 0,1,2,3,4,5,6,7 \quad \Rightarrow y = z - x$

$f(x,z) = f_{XY}(x, z-x) = (C^4_x 0.3^x \cdot 0.7^{4-x})(C^3_{z-x} 0.4^{z-x} \cdot 0.6^{3-z+x})$

Z	(x , y)	(x ,z)
0	(0 ,0)	(0 ,0)
1	(0 ,1) (1,0)	(0 ,1) (1,1)
2	(0 ,2) (1,1) (2,0)	(0 ,2) (1,2) (2,2)
3	(0 ,3) (1,2) (2,1) (3,0)	(0 ,3) (1,3) (2,3) (3,3)
4	(1,3) (2,2) (3,1) (4,0)	(1,4) (2,4) (3,4) (4,4)
5	(2,3) (3,2) (4,1)	(2,5) (3,5) (4,5)
6	(3,3) (4,2)	(3,6) (4,6)
7	(4,3)	(4,7)

$P(Z=1) = f_Z(1) = \sum_{x=0}^{1} f_{XZ}(x, 1)$

$= \sum_{x=0}^{1} (C^4_x 0.3^x \cdot 0.7^{4-x})(C^3_{1-x} 0.4^{1-x} \cdot 0.6^{3-1+x})$

$= (C^4_0 0.3^0 \cdot 0.7^{4-0})(C^3_{1-0} 0.4^{1-0} \cdot 0.6^{3-1+0})$

$+ (C^4_1 0.3^1 \cdot 0.7^{4-1})(C^3_{1-1} 0.4^{1-1} \cdot 0.6^{3-1+1})$

$= (0.2401 \times 0.432) + (0.4116 \times 0.216) = 0.1926$