

# 《統計學》

參考值： $z_{0.01} = 2.33, z_{0.05} = 1.645, z_{0.1} = 1.28, \chi_{2,0.05}^2 = 5.99, \chi_{3,0.05}^2 = 7.81,$   
 $\chi_{4,0.05}^2 = 9.49, \chi_{5,0.05}^2 = 11.07, \chi_{2,0.1}^2 = 4.61, F_{6,18,0.025} = 3.22, F_{6,18,0.05} = 2.66,$   
 $F_{5,18,0.025} = 3.38, F_{5,18,0.05} = 2.77$

一、假設金牌輪胎的使用壽命（單位：月）服從韋伯（Weibull）分配，機率密度函數為

$$f(x) = \frac{k}{\lambda} \left(\frac{x}{\lambda}\right)^{k-1} e^{-\left(\frac{x}{\lambda}\right)^k}, x \geq 0$$

其中 $\lambda = 120$ 是比例參數（scale parameter）， $k = 5$ 是形狀參數（shape parameter）。廠商擬訂出一個保固期限 $T$ （單位：月），使得僅有約0.05%的輪胎之壽命低於 $T$ 。利用伽瑪（Gamma）函數的定義與性質：

$$\Gamma(z) = \int_0^{\infty} x^{z-1} e^{-x} dx,$$

$\Gamma(z+1) = z\Gamma(z)$ 與 $\Gamma(0.2) = 4.59$ ，回答以下問題：（每小題10分，共20分）

（一）試問金牌輪胎之平均壽命為幾個月？

（二）試問保固期限 $T$ 應該訂為多少？

試題評析	本題是隨機變數期望值與事件機率之計算題型，其中Weibull分配於講義中已經有介紹過，並有說明其期望值之計算技巧，考生只要熟讀上課講義，滿分不難。
考點命中	《高點·高上統計學講義》第二回，趙治勳編撰，第七章第五節。

答：

$$X \sim f_X(x) = \frac{5}{120^5} x^{5-1} e^{-\frac{1}{120^5} x^5}, 0 < x < \infty$$

$$(一) E(X) = \int_0^{\infty} x \frac{5}{120^5} x^{5-1} e^{-\frac{1}{120^5} x^5} dx \quad \text{令 } u = x^5 \Rightarrow du = 5x^4 dx$$

$$= \frac{1}{120^5} \int_0^{\infty} u^{\frac{1}{5}+1-1} e^{-\frac{1}{120^5} u} du = \frac{1}{120^5} \frac{\Gamma(\frac{1}{5}+1)}{\left(\frac{1}{120^5}\right)^{\frac{1}{5}}} = 120 \times \frac{1}{5} \times 4.59 = 110.16 \text{ (月)}$$

$$(二) \text{ 令 } P(X \leq T) = \int_0^T \frac{5}{120^5} x^{5-1} e^{-\frac{1}{120^5} x^5} dx \quad \text{令 } u = x^5 \Rightarrow du = 5x^4 dx$$

$$= \frac{1}{120^5} \int_0^{T^5} e^{-\frac{1}{120^5} u} du = 1 - e^{-\left(\frac{T}{120}\right)^5} \leq 0.0005$$

$$\Rightarrow T \leq 120 \sqrt[5]{-\ln(0.9995)} = 26.242$$

故保固期可以訂為26個月

二、某部門調查員工每人上個月上網購物消費金額。假設共有100名員工，上個月平均網購消費為3000元，標準差500元。

（一）若網購消費金額之分配近似常態，試問上個月網購介於2500元至3500元的員工大約幾

人？(10分)

(二)若網購消費金額之分配近似常態，試問上個月網購高於4000元的員工大約幾人？(5分)

(三)若網購消費金額之分配為右偏，而你上個月網購消費金額為3000元，試問多數員工的網購消費金額比你高或低？為什麼？(5分)

試題評析	本題屬於常態分配事件機率之計算題型，其中查表值沒有清楚提供，然而考生找接近的或合理的代替即可，滿分不難。
考點命中	《高點·高上統計學講義》第二回，趙治勳編撰，第七章第三節。

**答：**

令  $X$  表員工上個月上網購物之消費金額(元)

(一)  $X \sim N(3000, 500^2)$

$$P(2500 < X < 3500) = P(-1 < Z < 1) \approx 0.2$$

[注意:從考卷所提供的查表值去找接近的]

故約有  $100 \times 0.2 = 200$  員工上個月上網購物之消費金額介於2500至3500元。

(二)  $X \sim N(3000, 500^2)$

$$P(X > 4000) = P(Z > 2) \approx 0.02 \quad [\text{注意:已知 } z_{0.025} = 1.96]$$

故約有  $100 \times 0.02 = 2$  員工上個月上網購物之消費金額高於4000元

(三)

右偏分配時  $Mode < Median < Mean$ ，因為我的消費金額為3000元剛好是平均數，可以預期多數員工的網購消費金額比我低。

三、林主任想分析其部門承接業務之件次是否因不同時段而有不同。根據下表業務件次資料，回答以下問題：

11-12am	1-2pm	2-3pm
27	13	24

(一)在0.1的顯著水準下，檢定以上三個時段的承接件次是否相同？(10分)

(二)執行(一)之檢定時，需對母體作何假設？(5分)

試題評析	本題屬於卡方適合度之計算題型，考古題常出現，考生拿到滿分不難。
考點命中	《高點·高上統計學講義》第三回，趙治勳編撰，第十三章。

**答：**

(一)卡方適合度檢定

	1	2	3
$O_i$	27	13	24
$E_i$	$21\frac{1}{3}$	$21\frac{1}{3}$	$21\frac{1}{3}$

$$H_0: p_1 = p_2 = p_3 = \frac{1}{3} \quad \text{vs} \quad H_1: \text{至少一個 } p_i \neq p_j, i \neq j = 1, 2, 3$$

$$\text{T.S.: } \chi^2 = \sum_{i=1}^3 \frac{(O_i - E_i)^2}{E_i} \sim \chi_{(3-1=2)}^2$$

R.R.: Reject  $H_0$  at  $\alpha = 0.1$  if  $\chi^{2*} > \chi_{(2),0.1}^2 = 4.61$

$\therefore \chi^{2*} = 5.094 \quad \therefore \text{reject } H_0$

結論：我們有足夠證據去推論三個時段的承接件次不盡相同的。

(二)卡方適合度檢定為無母數統計方法之一，對於母體分配不須做任何假設。

四、W車輛測試中心擬研究ABCDEF六個不同品牌的汽車是否耗油程度有所差別。今隨機抽取此六個品牌的汽車各四輛，分別測試其使用一公升汽油所行駛之公里數，得到各品牌四輛車的平均行駛距離分別為4.6、5.3、4.4、4.7、4.8與6.2公里，誤差平方和（sum of squares for error）為2.25。

(一)試做出此實驗之變異數分析表。（10分）

(二)在0.05的顯著水準之下，檢定不同品牌汽車的耗油程度有無差別。（5分）

<b>試題評析</b>	本題屬於one-way ANOVA之計算題型，考古題常出現，考生拿到滿分不難。
<b>考點命中</b>	《高點·高上統計學講義》第三回，趙治勳編撰，第十二章。

**答：**

$$(一) SSR = \sum_i n_i (\bar{Y}_i - \bar{\bar{Y}})^2 = 4[(4.6-5)^2 + \dots + (6.2-5)^2] = 8.72$$

$$\text{其中 } \bar{\bar{Y}} = \frac{4.6+5.3+\dots+6.2}{6} = 5$$

ANOVA TABLE				
source	SS	d.f.	MS	F
品牌	8.72	5	1.744	$F^* = 13.952$
Error	2.25	18	0.125	
Total	10.97	23		

(二) 假設模型： $Y_{ij} = \mu + \alpha_i + \varepsilon_{ij}, \varepsilon_{ij} \stackrel{iid}{\sim} N(0, \sigma^2), i = 1, 2, \dots, 6, j = 1, 2, 3, 4$

$H_0: \mu_1 = \mu_2 = \dots = \mu_6$  vs  $H_1: \text{至少一個 } \mu_i \neq \mu_j, i \neq j = 1, 2, \dots, 6$

$$\text{T.S.: } F = \frac{MSR}{MSE} \sim F_{(5,18)}$$

R.R.: Reject  $H_0$  at  $\alpha = 0.05$  if  $F^* > F_{(5,18),0.05} = 2.77$

$\therefore F^* = 13.952 \quad \therefore \text{reject } H_0$

結論：我們有足夠證據去推論不同品牌汽車耗油程度是有差異的。

五、鄭主任使用線上適性化平台來訓練其部屬的軟體操作能力，該平台依學習者能力，適性調整每個人在線上接受訓練之時間，當能力達到精熟程度時，該訓練即停止。以下是10名部屬在訓練前之能力評估分數（0~200分，分數越高代表能力越強），以及達到精熟程度所需要之練習時間（分鐘）。（每小題10分，共30分）

能力 181 33 13 34 47 11 20 91 74 164  
 時間 50 135 165 138 122 184 155 80 88 52

- (一) 試求出「能力」(y) 對「時間」(x) 之線性迴歸方程式，並計算判定係數 (coefficient of determination)。
- (二) 試說明對「能力」作何種轉換 (指數轉換： $z = e^y$ ，或自然對數轉換： $z = \log_e y$ )，可以使轉換後之變數與「時間」之線性相關程度提高？先畫出原始資料中「能力」對「時間」之散布圖 (scatter plot)，並計算二者之相關係數，再與轉換後之散布圖與相關係數做比較。
- (三) 以轉換後之變數對「時間」作線性迴歸，得出迴歸方程式及判定係數，並說明該判定係數之意義。

試題評析	本題屬於簡單線性迴歸分析之計算題型，考古題常出現，只是畫散佈圖與計算過程有些繁瑣，考生計算需要小心謹慎，但拿到滿分不難。
考點命中	《迴歸分析申論題完全制霸》，高點文化出版，趙治勳編著，第二章。

答：

假設模型： $Y_i = \beta_0 + \beta_1 X_i + \varepsilon_i, \varepsilon_i \sim N(0, \sigma^2), i = 1, 2, \dots, 10$

$$\sum X_i = 1169, \sum X_i^2 = 156607, \sum Y_i = 667, \sum Y_i^2 = 78197, \sum X_i Y_i = 53470$$

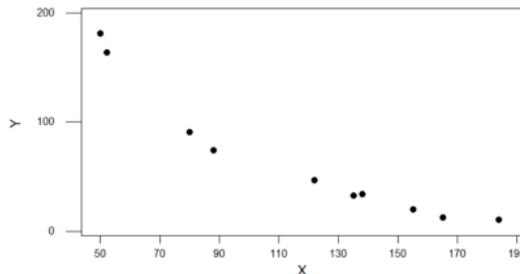
$$SS_X = 19950.9, SS_Y = 33708.1, SS_{XY} = -24502.3$$

$$(一) \hat{\beta}_1 = \frac{SS_{XY}}{SS_X} = -1.228 \quad \hat{\beta}_0 = \bar{Y} - \hat{\beta}_1 \bar{X} = 210.2532$$

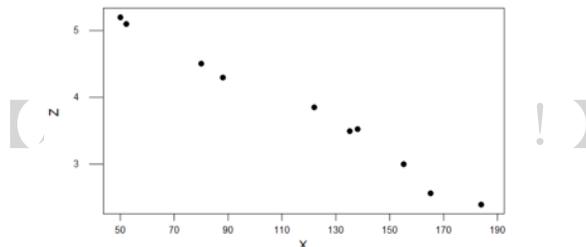
$$\therefore \hat{y} = 210.2532 - 1.228x$$

$$R^2 = \frac{SSR}{SST} = \frac{\hat{\beta}_1^2 SS_X}{SS_Y} = 0.8925$$

(二) 下圖為原始資料之散佈圖，相關係數  $r_{XY} = (\text{sign} \hat{\beta}_1) \sqrt{R^2} = -0.9447$



取自然對數轉換：令  $Z = \ln Y$ ，下圖為轉換後資料之散佈圖



$$\sum X_i = 1169, \sum X_i^2 = 156607, \sum Z_i = 37.945, \sum Z_i^2 = 152.693, \sum X_i Z_i = 4021.952$$

$$SS_X = 19950.9, SS_Z = 8.711, SS_{XZ} = -413.8185$$

$$r_{XZ} = \frac{SS_{XZ}}{\sqrt{SS_X} \sqrt{SS_Z}} = -0.993$$

因為  $|r_{XZ}| > |r_{XY}|$ ，故經由對數轉換後有提高線性相關程度

(三) 假設模型： $Z_i = \alpha_0 + \alpha_1 X_i + \eta_i, \eta_i \stackrel{iid}{\sim} N(0, \sigma^2), i = 1, 2, \dots, 10$

$$\hat{\alpha}_1 = \frac{SS_{XZ}}{SS_X} = -0.0207 \quad \hat{\alpha}_0 = \bar{Z} - \hat{\alpha}_1 \bar{X} = 6.21433$$

$$\therefore \hat{z} = 6.21433 - 0.0207x$$

$$R^2 = \frac{SSR}{SST} = \frac{\hat{\alpha}_1^2 SS_X}{SS_Z} = 0.9814 \text{ 表示考慮之自變數}(X)\text{與模型可以解釋}Z\text{之總變異達}98.14\%$$

【版權所有，重製必究！】