《統計學概要(經建)》

一、張三期末考考三科（微積分、統計學及英文），他微積分及格的機率為0.7，統計學及英文及格的機率分別為0.8及0.6。已知張三在此三科的表現互為獨立，請計算出張三在以下三種情況下的機率：

(一)此三科考試都不及格的機率。（5分）

(二)只有一科不及格的機率。（10分）

(三)已知只有一科不及格的情況下，此不及格科目為英文的機率。（10分）

試題評析	本題是在考事件機率之計算，相關題型考古題甚多，且講義中也有收錄相關題型，考生只要小心計算，獲得高分應該不難。
考點命中	《高點‧高上統計學講義》第一回，趙治勳編撰，第四章。

答：

令C,S,E分別表微積分、統計學與英文為成績及格

(一) $P(C^c \cap S^c \cap E^c) = P(C^c)P(S^c)P(E^c) = 0.3 \times 0.2 \times 0.4 = 0.024$

(二) $P(C^c \cap S \cap E) + P(C \cap S^c \cap E) + P(C \cap S \cap E^c)$

$= P(C^c)P(S)P(E) + P(C)P(S^c)P(E) + P(C)P(S)P(E^c)$

$= 0.3 \times 0.8 \times 0.6 + 0.7 \times 0.2 \times 0.6 + 0.7 \times 0.8 \times 0.4 = 0.452$

(三) $\dfrac{P(C \cap S \cap E^c)}{P(C^c \cap S \cap E) + P(C \cap S^c \cap E) + P(C \cap S \cap E^c)} = \dfrac{0.7 \times 0.8 \times 0.4}{0.452} = 0.4956$

二、某樣本抽自母體變異數為36之常態分配，研究人員想檢定H_0：$\mu \le 70$，H_a：$\mu > 70$。試問：

(一)若希望此檢定所犯的型Ⅰ過誤為0.01且型Ⅱ過誤為0.2，則所抽的樣本數大小應為何？（5分）

(二)若樣本數為100，研究人員決定以下列決策法則進行檢定：如果樣本平均數小於等於71.4，則接受H_0；如果樣本平均數大於71.4，則拒絕H_0。請計算出此檢定的型Ⅰ過誤。（10分）

(三)續(二)，如果實際的母體平均數為72.1，則所犯的型Ⅱ過誤為何？（10分）

試題評析	本題是單一母體平均數之假設檢定下，計算樣本數與型I型II誤差，相關題型考古題甚多，且講義中也有收錄相關題型，只是出題老師未提供相關抽樣分配之查表值，考生只要在答案卷上寫下計算公式，應該就會給分。
考點命中	《高點‧高上統計學講義》第五回，趙治勳編撰，第十一章例18與例19。

答：

(一) 母體：$X \sim N(\mu, \sigma^2)$

樣本：$X_1, X_2, \cdots, X_n \overset{iid}{\sim} N(\mu, \sigma^2)$

點估計：$\bar{X} \sim N(\mu, \frac{\sigma^2}{n} = \frac{36}{n})$

$H_0: \mu \le 70$ vs $H_1: \mu > 70$

$$max\,\alpha = 0.01 = P(\bar{X} > c \mid \mu = 70) = P(Z > \frac{c - 70}{\sqrt{36/n}})$$

$$\Rightarrow \frac{c-70}{\sqrt{36/n}} = z_{0.01} = 2.326 \text{ ----------(1)}$$

$$\beta = 0.2 = P(\bar{X} \le c \mid \mu = 72.1) = P(Z \le \frac{c-72.1}{\sqrt{36/n}})$$

$$\Rightarrow \frac{c-72.1}{\sqrt{36/n}} = -z_{0.2} = -0.8416 \text{ ---------(2)}$$

以上(1)(2)解聯立方程可得 $n = \frac{(2.326+0.8416)^2 \times 36}{(72.1-70)^2} = 81.9077 \approx 82$

因此，樣本數應該至少為 82 個

(二) 母體： $X \sim N(\mu, \sigma^2)$

樣本： $X_1, X_2, \cdots, X_{100} \overset{iid}{\sim} N(\mu, \sigma^2)$

點估計： $\bar{X} \sim N(\mu, \frac{\sigma^2}{n} = \frac{36}{100})$

$H_0 : \mu \le 70$ vs $H_1 : \mu > 70$

$$\max \alpha = P(\bar{X} > 71.4 \mid \mu = 70) = P(Z > \frac{71.4-70}{\sqrt{36/100}}) = P(Z > 2.33) = 0.01$$

(三) $\beta = P(\bar{X} \le 71.4 \mid \mu = 72.1) = P(Z \le \frac{71.4-72.1}{\sqrt{36/100}}) = P(Z \le -1.17) = 0.121$

三、某連鎖服飾公司想知道甲分店的來客購買率是否大於乙分店的來客購買率，業務部研究人員隨機分別自甲分店及乙分店選取150名及100名入店之客人，且分別有118名及72名客人購買產品。在10%的顯著水準下，使用此樣本資料幫助業務部研究人員進行檢定。請使用：

(一)信賴區間法。（10分）

(二)P值法。（10分）

試題評析	本題是兩獨立母體成功比例差之假設檢定，相關題型考古題甚多，且講義中也有收錄相關題型，只是出題老師未提供相關抽樣分配之查表值，考生只要在答案卷上寫下計算公式，應該就會給分。
考點命中	《高點・高上統計學講義》第五回，趙治勳編撰，第十一章例16。

答：

令 X_1, X_2 分別表甲分店與乙分店客人有購買產品

母體: $X_1 \sim Ber(p_1) \perp X_2 \sim Ber(p_2)$ （假設 $X_1 \perp X_2$）

樣本: $X_{11}, X_{12}, \cdots, X_{1150} \overset{iid}{\sim} Ber(p_1)$ $X_{21}, X_{22}, \cdots, X_{2100} \overset{iid}{\sim} Ber(p_2)$

點估計: $\hat{p}_1 - \hat{p}_2 \underset{by C.L.T.}{\sim} N(p_1 - p_2, \frac{p_1(1-p_1)}{150} + \frac{p_2(1-p_2)}{100})$

$H_0 : p_1 \le p_2$ vs $H_1 : p_1 > p_2$

T.S.: $Z = \frac{(\hat{p}_1 - \hat{p}_2) - (0)}{\sqrt{\hat{p}(1-\hat{p})(\frac{1}{150}+\frac{1}{100})}} \underset{by C.L.T.}{\sim} N(0,1)$ 其中 $\hat{p} = \frac{118+72}{150+100} = 0.76$

(一) R.R.: Reject H_0 at $\alpha=0.1$ if 0 未落在$((\hat{p}_1-\hat{p}_2)-z_{0.1}\sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{150}+\frac{\hat{p}_2(1-\hat{p}_2)}{100}}$，1)

$$\because((\hat{p}_1-\hat{p}_2)-z_{0.1}\sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{150}+\frac{\hat{p}_2(1-\hat{p}_2)}{100}},1)=(-0.005112,1)$$

$\therefore$ don't reject H_0

我們沒有足夠證據去推論甲分店之客人平均購買率大於乙分店

(二) R.R.: Reject H_0 at $\alpha=0.1$ if $\alpha>p-value$

$$\therefore p-value=P(Z>\frac{(\hat{p}_1-\hat{p}_2)-0}{\sqrt{\hat{p}(1-\hat{p})(\frac{1}{150}+\frac{1}{100})}})=P(Z>1.21)=0.113$$

$\therefore$ don't reject H_0

我們沒有足夠證據去推論甲分店之客人平均購買率大於乙分店

四、以下為蒐集來的六天累積雨量（毫米）及日照時間（小時）：

雨量（mm）	1.47	2.17	2.45	1.94	1.74	1.26
日照時間（小時）	2.8	7.6	5.5	1.8	5.7	1.1

(一)若欲以日照時間預測雨量，請求出最小平方直線，並在散佈圖中繪出此直線。（10分）
(二)在顯著水準為0.05情況下，檢定迴歸斜率是否不為零。（10分）
(三)試求出雨量與日照時間的樣本相關係數，並利用樣本相關係數檢定此二變數是否具有顯著關係（顯著水準為0.05）。（10分）

試題評析	本題是簡迴歸之基本計算題，相關題型考古題甚多，且講義中也有收錄相關題型，只是出題老師未提供相關抽樣分配之查表值，考生只要在答案卷上寫下計算公式，應該就會給分。
考點命中	《高點・高上統計學講義》第六回，趙治勳編撰，第十三章。

答：

令 X 表日照時間(小時)， Y 表雨量(mm)

假設：$Y_i=\beta_0+\beta_1X_i+\varepsilon_i,\varepsilon_i\overset{iid}{\sim}N(0,\sigma^2),i=1,2,\cdots,6$

$$SS_X=\sum_{i=1}^{n}X_i^2-\frac{(\sum_{i=1}^{n}X_i)^2}{n}=32.7483，\quad SS_Y=\sum_{i=1}^{n}Y_i^2-\frac{(\sum_{i=1}^{n}Y_i)^2}{n}=0.9743$$

$$SS_{XY}=\sum_{i=1}^{n}X_iY_i-\frac{(\sum_{i=1}^{n}X_i)(\sum_{i=1}^{n}Y_i)}{n}=3.8398$$

(一) $\hat{\beta}_1=\frac{SS_{XY}}{SS_X}=0.1173 \qquad \hat{\beta}_0=\overline{Y}-\hat{\beta}\overline{X}=-1.3594$

$\therefore\hat{y}=1.3594+0.1173x$

(二) $H_0:\beta_1=0$ vs $H_1:\beta_1\neq0$

T.S.：$T=\frac{\hat{\beta}_1-0}{S(\hat{\beta}_1)}\sim t_{(6-2=4)}$

R.R.：Reject H_0 at $\alpha = 0.05$ if $|T^*| > t_{(4)0.025} = 2.776$

$\because |T^*| = |\dfrac{0.1173\text{-}0}{0.06323}| = 1.8551$ $\therefore$ don't reject H_0

其中 $S(\hat{\beta}_1) = \sqrt{\dfrac{MSE}{SS_X}} = \sqrt{\dfrac{\frac{SST-SSR}{n-2}}{SS_X}} = \sqrt{\dfrac{\frac{SS_Y - \hat{\beta}_1^2 SS_X}{n-2}}{SS_X}} = 0.06323$

我們沒有足夠證據去推論斜率不為零

(三) $R^2 = \dfrac{SSR}{SST} = \dfrac{\hat{\beta}_1^2 SS_X}{SS_Y} = 0.4625$

$r_{XY} = \pm\sqrt{R^2} = +\sqrt{0.4625} = 0.6801$ ($\because \hat{\beta}_1 > 0$ 故取 $+$ 號)

$H_0 : \rho_{XY} = 0$ vs $H_1 : \rho_{XY} \neq 0$

T.S.：$T = \dfrac{r_{XY}\sqrt{n-2}}{\sqrt{1-r_{XY}^2}} \sim t_{(6-2=4)}$

R.R.：Reject H_0 at $\alpha = 0.05$ if $|T^*| > t_{(4)0.025} = 2.776$

$\because |T^*| = |\dfrac{0.6801\sqrt{6-2}}{\sqrt{1-(0.6801)^2}}| = 1.855$ $\therefore$ don't reject H_0

我們沒有足夠證據去推論雨量與日照時間具有線性相關