《統計學概要》

註：$t_{.05}(9)=1.833$, $t_{.025}(9)=2.262$, $t_{.05}(10)=1.812$, $t_{.025}(10)=2.228$,
$F_{.05}(3,50)=2.79$, $F_{.05}(4,50)=2.56$, $F_{.025}(3,50)=3.39$, $F_{.025}(4,50)=3.05$,
$\Phi(1)=0.8413$, $\Phi(1.5)=0.9332$, $\Phi(2.0)=0.9772$, Φ是標準常態分配的CDF,
$z_{.05}=1.645$, $z_{.01}=2.33$
所有假設檢定問題，皆需正確寫出虛無假設、對立假設、檢定統計量、拒絕域、檢定結果與結論。

一、假設我們有下列聯合機率：

	A_1	A_2	A_3
B_1	.15	.20	.10
B_2	.25	.25	.05

(一)計算$P(A_1 \cup B_2)$。（5分）
(二)計算$P(B_1 \mid A_2)$。（5分）

試題評析	本題涉及交集聯集機率及條件機率之基礎運算。
考點命中	《高點・高上統計學講義第一回》，趙治勳編撰，第三章3.3節。

答：

(一) $P(A_1 \cup B_2) = P(A_1) + P(B_2) - P(A_1 \cap B_2) = 0.4 + 0.55 - 0.25 = 0.7$

(二) $P(B_1 \mid A_2) = \dfrac{P(A_2 \cap B_1)}{P(A_2)} = \dfrac{0.2}{0.45} = 0.444$

二、已知一組（n = 200）資料的分布圖呈現鐘形分配，平均數為60，標準差為10。
(一)約有多少觀察值介於40與80之間。（5分）
(二)約有多少觀察值超過75。（5分）

試題評析	本題涉及單邊柴比雪夫不等式。
考點命中	《高點・高上統計學講義第一回》，趙治勳編撰，頁70。

答：

(一)由經驗法則, $P(40 < X < 80) = P(60-(2)(10) < X < 60+(2)(10)) \approx 0.95$
約有 $200 \times 0.95 = 190$ 個觀察值介於40與80之間

(二)由單邊柴比雪夫,

$$P(X \ge 75) = P(X \ge 60+15) \le \frac{10^2}{10^2+15^2} = 0.3077$$

至多有 $200 \times 0.3077 = 61.54 \approx 61$ 個觀察值超過75

三、三年甲班數學學期成績服從常態分配，平均數70，標準差10。求下列各值：
(一)隨機抽取1位同學，其數學成績超過80的機率。（5分）
(二)隨機抽取25位同學，其數學平均成績超過74的機率。（5分）
(三)數學老師打算當掉成績在最後5%的同學，則同學要有多少分才不會被當。（5分）

答：

令 X 表數學學期成績，$X \sim N(70,10^2)$

(一) $P(X>80)=P(Z>1)=0.1587$

(二) $X_1,X_2,\cdots,X_{25}\overset{iid}{\sim}N(70,10^2)$，$\overline{X}\sim N(70,\frac{10^2}{25})$

$P(\overline{X}>74)=P(Z>2)=0.0228$

四、下列是10位學生的身高資料：

172　168　164　170　176　160　154　179　160　166

(一)求中位數。(4分)

(二)求四分位距（InterQuartile Range）。(7分)

(三)求第90百分位數。(4分)

試題評析	本題涉及敘述統計學中分位數。
考點命中	《高點・高上統計學講義第一回》，趙治勳編撰，頁15。

答：

排序: 154　160　160　164　166　168　170　172　176　179

(一)中位數爲區間(166,168)內所有值,以167代表之

(二) $IQR=Q_3-Q_1$=172-160=12

(三)第90百分位數爲9.5

五、為了瞭解黃色壘球在晚上的比賽中是否比白色壘球較容易看見，分別記錄了6場用黃色壘球與6場用白色壘球比賽時失誤的次數。

黃色	5	2	6	7	2	5
白色	7	6	8	5	9	11

(一)請問在.05顯著水準下，我們是否可以推論用黃色壘球比賽時的平均失誤次數比用白色壘球比賽時少？

(二)在執行(一)的檢定時，我們需要那些假設條件？(5分)

試題評析	本題涉及兩獨立母體平均數差之假設檢定。
考點命中	《高點・高上統計學講義第二回》，趙治勳編撰，頁69。

答：

(一)

令 X_1,X_2 分別代表黃色壘球與白色壘球之失誤次數

母體：$X_1\sim N(\mu_1,\sigma_1^2)\perp X_2\sim N(\mu_2,\sigma_2^2)$

樣本：$X_{11},X_{12},\cdots,X_{16}\overset{iid}{\sim}N(\mu_1,\sigma_1^2)$, $X_{21},X_{22},\cdots,X_{26}\overset{iid}{\sim}N(\mu_2,\sigma_2^2)$

點估計：$\overline{X_1}-\overline{X_2}\sim N(\mu_1-\mu_2,\frac{\sigma_1^2}{6}+\frac{\sigma_2^2}{6})$

$H_0:\mu_1-\mu_2\ge 0$ vs $H_0:\mu_1-\mu_2<0$

T.S.：$T=\dfrac{\overline{X_1}-\overline{X_2}-(0)}{\sqrt{\dfrac{S_1^2}{6}+\dfrac{S_2^2}{6}}}\sim t_{(v=10)}$

$$其中\ v=\frac{(\frac{S_1^2}{n_1}+\frac{S_2^2}{n_2})^2}{\frac{(\frac{S_1^2}{n_1})^2}{n_1-1}+\frac{(\frac{S_2^2}{n_2})^2}{n_2-1}}=\frac{(\frac{2.0736^2}{6}+\frac{2.1602^2}{6})^2}{\frac{(\frac{2.0736^2}{6})^2}{6-1}+\frac{(\frac{2.1602^2}{6})^2}{6-1}}=9.983\approx 10$$

R.R.: Reject H_0 at $\alpha=0.05$ if $T^*<-t_{0.05(10)}=-1.812$

$\because T^*=-2.591 \qquad \therefore reject H_0$

我們有足夠證據去推論黃色壘球之失誤次數少於白色壘球

(二)

假設兩獨立常態母體且隨機樣本

六、考慮一單因子變異數分析，經整理獲得下列資料：

	處	理		
統計量	1	2	3	4
n	10	14	11	18
$\bar{x}$	30	35	33	40
S^2	10	10	10	10

(一)請完成ANOVA表。(20分)

(二)請問在.05顯著水準下，我們是否可以推論4種處理的平均值間存在差異？(10分)

試題評析	本題涉及一因子變異數分析。
考點命中	《高上統計學講義第二回》，趙治勳編撰，頁108例(2)。

答：

(一)

$SSE=\sum(n_i-1)S_i^2=9(10)+13(10)+10(10)+17(10)=490$

$SSR=\sum n_i(\bar{X}_{i\bullet}-\bar{\bar{X}})^2=10(30-35.34)^2+14(35-35.34)^2+11(33-35.34)^2+18(40-35.34)^2$

$=737.8868$

其中 $\bar{\bar{X}}=\frac{(10)(30)+(14)(35)+(11)(33)+(18)(40)}{10+14+11+18}=35.34$

$SST=SSR+SSE=1227.8868$

ANOVA TABLE

source	SS	d.f.	MS	F
trt	737.8868	3	245.962	24.5962
Error	490	49	10	
Total	1227.8868	52		

(二)

假設模型 $X_{ij}=\mu+\alpha_i+\varepsilon_{ij},\varepsilon_{ij}\overset{iid}{\sim}N(0,\sigma^2)$

$H_0:\mu_1=\mu_2=\mu_3=\mu_4$ vs H_1:至少一個$\mu_i\neq\mu_j$

T.S.: $F=\frac{MSR}{MSE}\sim F_{(3,49)}$

R.R.： Reject H_0 at $\alpha = 0.05$ if $F^* > F_{0.05(3,49)} = 2.7939$ (考卷上給的查表值有誤)

$\because F^* = 24.5962 \qquad \therefore rejectH_0$

結論：我們有足夠的統計證據推論四種處理之平均數存在差異