บทที่ 6

สรุป

6.1 การใช้วิธีที่เสนอมากับชีวิตจริง (real life)

วิทยานิพนธ์เรื่องนี้เป็นการนำเอาวิชาสถิติมาช่วยแก้ปัญหาในชีวิตประจำวันที่เป็นชีวิตจริงและมีผลเกี่ยวข้องกับคนจำนวนมาก ซึ่งล้วนเป็นคนสำคัญของชาติในอนาคต

6.2 ข้อดีของวิธีที่เสนอมา

จะเห็นว่าเมื่อได้ statistical model ที่เหมาะสมแล้วจะมีข้อดีคือ

ก. ประหยัดเวลา ค่าใช้จ่ายและตัดปัญหายุ่งยากต่าง ๆ ในการสอบคัดเลือก

ข. ได้คนดีที่เหมาะสมไปพัฒนาประเทศ สอดคล้องกับนโยบายรัฐบาล

ค. ได้ทราบผลที่แท้จริงว่าคะแนน ม.ศ. 5 วิชาใดบ้างที่มีผลต่อการศึกษาในแต่ละคณะในสถาบันอุดมศึกษาทำให้สามารถปรับปรุงระดับการศึกษาในชั้นมัธยมศึกษาได้ถูกต้อง

ง. นักเรียนในชั้นมัธยมศึกษาจะเพิ่มความสนใจกับบทเรียนมาก เพราะผลการเรียน หมายถึงอนาคตที่จะได้เข้าศึกษาต่อในสถาบันอุดมศึกษาหรือไม่ ทำให้ความรู้ของนักเรียนดีขึ้น เป็นการยกระดับการศึกษาของชาติในทางอ้อม

จ. ตัดปัญหายุ่งยากในการสอบใหม่จะเป็นโดยการย้ายคณะหรือสอบใหม่เพราะเข้าไม่ได้ในปีก่อนก็ตาม

6.3 สรุปการศึกษานี้

จากการศึกษานี้จะเห็นว่าสามารถนำทฤษฎีทางสถิติมาประยุกต์ได้ เป็นประโยชน์อย่างยิ่งแก่การศึกษาของประเทศ

ภาคผนวกเชิงคณิตศาสตร์ (Mathematical Appendix)

1. ต้องการ minimize $\Sigma(y - y_e)^2$ wrt. a,b เมื่อ $y_2 = a + bx$

$$\Rightarrow f(a, b) = \Sigma(y - a - bx)^2$$

$$\frac{\partial f(a, b)}{\partial a} = \partial\Sigma(y - a - bx)^2/\partial a$$

$$= -2(y - a - bx) = 0 \quad \ldots\ldots(1)$$

$$\frac{\partial f(a, b)}{\partial b} = \frac{\partial\Sigma(y - a - bx)^2}{b}$$

$$= -2\Sigma x(y - a - bx) = 0 \quad \ldots\ldots(2)$$

จาก (1), $\Sigma(y - a - bx) = 0$

$$\Sigma y - an - b\Sigma x = 0$$

$$\Rightarrow an + bx = \Sigma y \quad \ldots\ldots\ldots\ldots(3)$$

จาก (2) $\Sigma x(y - a - bx) = 0$

$$\Sigma xy - a\Sigma x - b\Sigma x^2 = 0$$

$$a\Sigma x + b\Sigma x^2 = \Sigma xy \quad \ldots\ldots\ldots(4)$$

จะได้ว่า a,b จะเป็น Solution ของสมการ normal equation

$$an + bx = \Sigma y$$

$$a\Sigma x + b\Sigma x^2 = \Sigma xy$$

2. ในกรณีที่ A เป็น symmetric matrix จะได้ว่า

$$\frac{\partial(x'A\,x)}{\partial\,x} = 2Ax$$

พิสูจน์

$$x'A\,x = \{x_1, x_2, \ldots, x_n\} \begin{bmatrix} a_{11} & a_{12} \cdots a_{1n} \\ a_{12} & a_{22} \cdots a_{2n} \\ \vdots & \vdots \quad \vdots \\ a_{1n} & a_{2n} \quad a_{nn} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$$

เทอมที่ (1, 1) ของ X'A จะเป็น

$$= (a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n)$$

เทอมที่ (1, 2) ของ X'A จะเป็น

$$= (a_{12}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n)$$

เทอมที่ (1, n) ของ X'A จะเป็น

$$= (a_{1n}x_1 + a_{2n}x_2 + \cdots + a_{nn}x_n)$$

$$\therefore X'AX = (a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n)\,x_1$$
$$+ (a_{12}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n)\,x_2$$
$$+ \cdots\cdots\cdots$$
$$+ (a_{1n}x_1 + a_{2n}x_2 + \cdots + a_{nn}x_n)\,x_n$$
$$= a_{11}x_1^2 + a_{12}x_1x_2 + \cdots + a_{1n}x_1x_n$$
$$+ a_{12}x_1x_2 + a_{22}x_2^2 + \cdots + a_{2n}x_2x_n$$
$$+ \cdots\cdots\cdots$$
$$+ a_{1n}x_1x_n + a_{2n}x_2x_n + \cdots + a_{nn}x_n^2$$
$$= a_{11}x_1^2 + 2a_{12}x_1x_2 + 2a_{13}x_1x_3 + \cdots + 2a_{1n}x_1x_n$$
$$+ a_{22}x_2^2 + 2a_{23}x_2x_3 + \cdots + 2a_{2n}x_2x_n$$
$$+ \cdots\cdots\cdots$$
$$+ a_{nn}x_n^2$$

$$\frac{\partial (X'AX)}{\partial x_1} = 2(a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n)$$

$$\frac{\partial (X'AX)}{\partial x_2} = 2(a_{12}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n)$$

$\cdots\cdots\cdots$

$$\frac{\partial (X'AX)}{\partial x_n} = 2(a_{1n}x_1 + a_{2n}x_2 + \cdots + a_{nn}x_n)$$

$$\Rightarrow \frac{\partial (X'AX)}{\partial x} = 2AX$$

Q.E.D.

```
C       MAIN PROGRAM
        DIMENSION XBAR(40),STD(40),D(40),RY(40),ISAVE(40),B(40),
       1          SB(40),T(40),W(40)
        DIMENSION RX(1600)
        DIMENSION R(820)
        DIMENSION ANS(10)
      1 FORMAT(A4,A2,I5,2I2)
      2 FORMAT(24H1MULTIPLE REGRESSION    A4,A2//6X,14HSELECTION      I2//
       1)
      3 FORMAT(9HCVARIABLE,5X,4HMEAN,6X,8HSTANDARD,6X,11HCORRELATION,4X,10
       1HREGRESSION,4X,10HSTD. ERROR,5X,8HCOMPUTED/6H   NO.,18X,9HDEVIATIO
       2N,7X,6HX VS Y,7X,11HCOEFFICIENT,3X,12HOF REG.COEF.,3X,7HT VALUE)
      4 FORMAT(1H ,I4,6F14.5)
      5 FORMAT(10H DEPENDENT)
      6 FORMAT(1HO/10H INTERCEPT,13X,F13.5//23H MULTIPLE CORRELATION  ,F13
       1.5//23H STD. ERROR OF ESTIMATE,F13.5//)
      7 FORMAT(1HO,21X,39HANALYSIS OF VARIANCE FOR THE REGRESSION//5X,19HS
       1OURCE OF VARIATION,7X,7HDEGREES,7X,6HSUM OF,10X,4HMEAN,12X,7HF VAL
       2UE/ 30X,10HOF FREEDOM,4X,7HSQUARES,9X,7HSQUARES)
      8 FORMAT(30H ATTRIBUTABLE TO REGRESSION   ,I6,3F16.5/30H DEVIATION F
       1ROM REGRESSION    ,I6,2F16.5)
      9 FORMAT(1H ,5X,5HTOTAL,19X,I6,F16.5)
     10 FORMAT(36I2)
     11 FORMAT(1H ,15X,18HTABLE OF RESIDUALS//9H CASE NO.,5X,7HY VALUE,5X,
       110HY ESTIMATE,6X,8HRESIDUAL)
     12 FORMAT(1H ,I6,F15.5,2F14.5)
     13 FORMAT(53H1NUMBER OF SELECTIONS NOT SPECIFIED.  JOB TERMINATED.)
     14 FORMAT(52HOTHE MATRIX IS SINGULAR.  THIS SELECTION IS SKIPPED.)
    100 READ (1,1,END=950) PR,PR1,N,M,NS
        REWIND 5
        IO=0
        X=0.0
        CALL CORRE (N,M,IO,X,XBAR,STD,RX,R,D,B,T)
        REWIND 5
        IF(NS) 108,108,109
    108 WRITE(3,13)
        GO TO 300
    109 DO 200 I=1,NS
        WRITE(3,2) PR,PR1,I
        READ(1,10) NRESI,NDEP,K,(ISAVE(J),J=1,K)
        CALL CRDER (M,R,NDEP,K,ISAVE,RX,RY)
        CALL MINV (RX,K,DET,B,T)
        IF(DET)112,110,112
    110 WRITE(3,14)
        GO TO 200
    112 CALL MULTR (N,K,XBAR,STD,D,RX,RY,ISAVE,B,SB,T,ANS)
        MM=K+1
        WRITE(3,3)
        DO 115 J=1,K
        L=ISAVE(J)
    115 WRITE(3,4) L,XBAR(L),STD(L),RY(J),B(J),SB(J),T(J)
        WRITE(3,5)
        L=ISAVE(MM)
        WRITE(3,4) L,XBAR(L),STD(L)
```

```
      WRITE(3,6) ANS(1),ANS(2),ANS(3)
      WRITE(3,7)
      L=ANS(8)
      WRITE(3,8) K,ANS(4),ANS(6),ANS(10),L,ANS(7),ANS(9)
      L=N-1
      SUM=ANS(4)+ANS(7)
      WRITE(3,9) L,SUM
      IF(NRESI) 200,200,120
  120 WRITE(3,2) PR,PR1,I
      WRITE(3,11)
      MM=ISAVE(K+1)
      DO 140 II=1,N
      READ(05) (W(J),J=1,M)
      SUM=ANS(1)
      DO 130 J=1,K
      L=ISAVE(J)
  130 SUM=SUM+W(L)*B(J)
      RESI=W(MM)-SUM
  140 WRITE (3,12) II,W(MM),SUM,RESI
      REWIND 5
  200 CONTINUE
      GO TO 100
  300 CONTINUE
  950 STOP
      END
```

```
      SUBROUTINE ORDER(M,R,NDEP,K,ISAVE,RX,RY)
      DIMENSION R(1),ISAVE(1),RX(1),RY(1)
      MM=0
      DO 130 J=1,K
      L2=ISAVE(J)
      IF(NDEP-L2)122,123,123
  122 L=NDEP+(L2*L2-L2)/2
      GO TO 125
  123 L=L2+(NDEP*NDEP-NDEP)/2
  125 RY(J)=R(L)
       DO 130 I=1,K
      L1=ISAVE(I)
      IF(L1-L2)127,128,128
  127 L=L1+(L2*L2-L2)/2
      GO TO 129
  128 L=L2+(L1*L1-L1)/2
  129 MM=MM+1
  130 RX(MM)=R(L)
      ISAVE(K+1)=NDEP
      RETURN
      END
```

```
      SUBROUTINE CORRE (N,M,IC,X,XBAR,STD,RX,R,B,D,T)
      DIMENSION X(1),XBAR(1),STD(1),RX(1),R(1),B(1),D(1),T(1)
      DO 100 J=1,M
      B(J)=0.0
  100 T(J)=0.0
      K=(M*M+M)/2
      DO 102 I=1,K
  102 R(I)=0.0
      FN=N
      L=0
      IF(IC)105,127,105
  105 DO 108 J=1,M
      DO 107 I=1,N
      L=L+1
  107 T(J)=T(J)+X(L)
      XBAR(J)=T(J)
  108 T(J)=T(J)/FN
      DO 115 I=1,N
      JK=0
      L=I-N
      DO 110 J=1,M
      L=L+N
      D(J)=X(L)-T(J)
  110 B(J)=B(J)+D(J)
      DO 115 J=1,M
      DO 115 K=1,J
      JK=JK+1
  115 R(JK)=R(JK)+D(J)*D(K)
      GO TO 205
  127 IF(N-M) 130,130,135
  130 KK=N
      GO TO 137
  135 KK=M
  137 DO 140 I=1,KK
      CALL DATA (M,D)
      DO 140 J=1,M
      T(J)=T(J)+D(J)
      L=L+1
  140 RX(L)=D(J)
      FKK=KK
      DO 150 J=1,M
      XBAR(J)=T(J)
  150 T(J)=T(J)/FKK
      L=0
      DO 180 I=1,KK
      JK=0
      DO 170 J=1,M
      L=L+1
  170 D(J)=RX(L)-T(J)
      DO 180 J=1,M
      B(J)=B(J)+D(J)
      DO 180 K=1,J
      JK=JK+1
  180 R(JK)=R(JK)+D(J)*D(K)
```

```
      IF(N-KK) 205,205,185
  185 KK=N-KK
      DO 200 I=1,KK
      JK=0
      CALL DATA(M,D)
      DO 190 J=1,M
      XBAR(J)=XBAR(J)+D(J)
      D(J)=D(J)-T(J)
  190 B(J)=B(J)+D(J)
      DO 200 J=1,M
      DO 200 K=1,J
      JK=JK+1
  200 R(JK)=R(JK)+D(J)*D(K)
  205 JK=0
      DO 210 J=1,M
      XBAR(J)=XBAR(J)/FN
      DO 210 K=1,J
      JK=JK+1
  210 R(JK)=R(JK)-B(J)*B(K)/FN
      JK=0
      DO 220 J=1,M
      JK=JK+J
  220 STD(J)=SQRT(ABS(R(JK)))
      DO 230 J=1,M
      DO 230 K=J,M
      JK=J+(K*K-K)/2
      L=M*(J-1)+K
      RX(L)=R(JK)
      L=M*(K-1)+J
      RX(L)=R(JK)
      IF(STD(J)*STD(K)) 225,222,225
  222 R(JK)=0.0
      GO TO 230
  225 R(JK)=R(JK)/(STD(J)*STD(K))
  230 CONTINUE
      FN=SQRT(FN-1.0)
      DO 240 J=1,M
  240 STD(J)=STD(J)/FN
      L=-M
      DO 250 I=1,M
      L=L+M+1
  250 B(I)=RX(L)
      RETURN
      END
```

```
      SUBROUTINE MINV(A,N,D,L,M)
      DIMENSION A(1),L(1),M(1)
      D=1.0
      NK=-N
      DO 80 K=1,N
      NK=NK+N
      L(K)=K
      M(K)=K
      KK=NK+K
      BIGA=A(KK)
      DO 20 J=K,N
      IZ=N*(J-1)
      DO 20 I=K,N
      IJ=IZ+I
   10 IF(ABS(BIGA)-ABS(A(IJ)))15,20,20
   15 BIGA=A(IJ)
      L(K)=I
      M(K)=J
   20 CONTINUE
      J=L(K)
      IF(J-K) 35,35,25
   25 KI=K-N
      DO 30 I=1,N
      KI=KI+N
      HOLD=-A(KI)
      JI=KI-K+J
      A(KI)=A(JI)
   30 A(JI)=HOLD
   35 I=M(K)
      IF(I-K) 45,45,38
   38 JP=N*(I-1)
      DO 40 J=1,N
      JK=NK+J
      JI=JP+J
      HOLD=-A(JK)
      A(JK)=A(JI)
   40 A(JI)=HOLD
   45 IF(BIGA) 48,46,48
   46 D=0.0
      RETURN
   48 DO 55 I=1,N
      IF(I-K) 50,55,50
   50 IK=NK+I
      A(IK)=A(IK)/(-BIGA)
   55 CONTINUE
      DO 65 I=1,N
      IK=NK+I
      HOLD=A(IK)
      IJ=I-N
      DO 65 J=1,N
      IJ=IJ+N
      IF(I-K) 60,65,60
   60 IF(J-K) 62,65,62
   62 KJ=IJ-I+K
```

```
      A(IJ)=HOLD*A(KJ)+A(IJ)
   65 CONTINUE
      KJ=K-N
      DO 75 J=1,N
      KJ=KJ+N
      IF(J-K) 70,75,70
   70 A(KJ)=A(KJ)/BIGA
   75 CONTINUE
      D=D*BIGA
      A(KK)=1.0/BIGA
   80 CONTINUE
      K=N
  100 K=(K-1)
      IF(K)150,150,105
  105 I=L(K)
      IF(I-K)120,120,108
  108 JQ=N*(K-1)
      JR=N*(I-1)
      DO 110 J=1,N
      JK=JQ+J
      HOLD=A(JK)
      JI=JR+J
      A(JK)=-A(JI)
  110 A(JI)=HOLD
  120 J=M(K)
      IF(J-K)100,100,125
  125 KI=K-N
      DO 130 I=1,N
      KI=KI+N
      HOLD=A(KI)
      JI=KI-K+J
      A(KI)=-A(JI)
  130 A(JI)=HOLD
      GO TO 100
  150 RETURN
      END


      SUBROUTINE DATA(M,D)
      DIMENSION D(1)
    1 FORMAT(24F3.0/2F3.0/14F3.0)
      READ(1,1) (D(I),I=1,M)
C     SAD,1952-69,MON.4/10/1971
      D(1) =D(1)-40.75*D(24)-0.23*D(3)
     1       +0.24*(ABS(D(29)+D(30)+D(31)+D(32)-800.0))
     2       -0.38*D(30)+0.22*D(35)
     2       +0.22*D(35)
      D(5)=D(5)/1000
      D(6)=D(5)**2
      D(7)=D(5)**3
      D(8)=D(5)**4
      WRITE(5)  (D(I),I=1,M)
      RETURN
      END
```

```
      SUBROUTINE MULTR (N,K,XBAR,STD,D,RX,RY,ISAVE,B,SB,T,ANS)
      DIMENSION XBAR(1),STD(1),D(1),RX(1),RY(1),ISAVE(1),B(1),SB(1),
     1          T(1),ANS(1)
      MM=K+1
      DO 100 J=1,K
  100 B(J)=0.0
      DO 110 J=1,K
      L1=K*(J-1)
      DO 110 I=1,K
      L=L1+I
  110 B(J)=B(J)*RY(I)*RX(L)
      RM=0.0
      BO=0.0
      L1=ISAVE(MM)
      DO 120 I=1,K
      RM=RM+B(I)*RY(I)
      L=ISAVE(I)
      B(I)=B(I)*(STD(L1)/STD(L))
  120 BO=BO+B(I)*XBAR(L)
      BO=XBAR(L1)-BO
      SSAR=RM*D(L1)
  122 RM=SQRT(ABS(RM))
      SSDR=D(L1)-SSAR
      FN=N-K-1
      SY=SSDR/FN
      DO 130 J=1,K
      L1=K*(J-1)+J
      L=ISAVE(J)
  125 SB(J)=SQRT(ABS((RX(L1)/D(L))*SY))
  130 T(J)=B(J)/SB(J)
  135 SY=SQRT(ABS(SY))
      FK=K
      SSARM=SSAR/FK
      SSDRM=SSDR/FN
      F=SSARM/SSDRM
      ANS(1)=BO
      ANS(2)=RM
      ANS(3)=SY
      ANS(4)=SSAR
      ANS(5)=FK
      ANS(6)=SSARM
      ANS(7)=SSDR
      ANS(8)=FN
      ANS(9)=SSDRM
      ANS(10)=F
      RETURN
      END
```

Program นี้ input data เป็น X_{ij}

$i = 1, 2, \ldots, n$ เป็นจำนวน observations

$j = 1, 2, \ldots, m$ เป็นจำนวน variables

1) $$S_{jk} = \sum_{i=1}^{n} (X_{ij} - T_j)(X_{ik} - T_k) - \frac{\sum_{i=1}^{n} (X_{ij} - T_j) \sum_{i=1}^{n} (X_{ik} - T_k)}{n} \quad \ldots\ldots\ldots(1)$$

เมื่อ S_{jk} = sums of cross-products of deviations

$j = 1, 2, \ldots, m$; $k = 1, 2, \ldots, m$

2) $$T_j = \frac{\sum_{i=1}^{m} X_{ij}}{m} \quad \ldots\ldots\ldots(2)$$

เมื่อ T_j

3) $$\bar{X}_j = \frac{\sum_{i=1}^{n} X_{ij}}{n} \quad \ldots\ldots\ldots(3)$$

เมื่อ $\bar{X}_j$ = ค่าเฉลี่ย

$j = 1, 2, \ldots, m$

4) $$r_{jk} = \frac{S_{jk}}{\sqrt{S_{jj}}\ \sqrt{S_{kk}}} \quad \ldots\ldots\ldots(4)$$

r_{jk} = correlation coefficients

$j = 1, 2, \ldots, m$; $k = 1, 2, \ldots, m$

5) $S_j = \frac{\sqrt{S_{jj}}}{\sqrt{n-1}}$ .....(5)

S_j = standard deviations

j = 1, 2, ..., m

6) $b_j = \beta_j - \frac{S_y}{S_j}$ .....(6)

b_j = regression coefficients

β_j = beta weights

$= \sum_{i=1}^{k} r_{iy} \cdot r_{ij}^{-1}$

r_{iy} = intercorrelation of i^{th} independent variable with dependent variable

r_{ij}^{-1} = inverse of intercorrelation r_{ij}

$i = j = 1, 2, ..., k$ เป็น independent variables

S_y = standard deviation ของ dependent variable

S_j = standard deviation ของ j^{th} independent variable

j = 1, 2, ..., k

7) $b_o = \bar{Y} - \sum_{j=1}^{k} b_j \cdot \bar{X}_j$ .....(7)

b_o = intercept

$\bar{Y}$ = mean of dependent variable

$\bar{X}_j$ = mean of j^{th} independent variable

8) $$R = \sqrt{\sum_{i=1}^{k} \beta_i r_{ij}} \qquad \text{.....(8)}$$

R = multiple correlation coefficient

9) $$SSAR = R^2 . D_{yy} \qquad \text{.....(9)}$$

$SSAR$ = sum of squares attributable to the regression

D_{yy} = sum of squares of deviations from mean for dependent variable

10) $$SSDR = D_{yy} - SSAR \qquad \text{.....(10)}$$

$SSDR$ = sum of squares of deviations from the regression

11) $$F = \frac{SSAR/k}{SSDR/(n-k-1)} \qquad \text{.....(11)}$$

F = F-value

12) $$S^2_{y.12...k} = \frac{SSDR}{n-k-1} \qquad \text{.....(12)}$$

$S^2_{y.12...k}$ = variance of estimate

n = number of observations

13) $$S_{b_j} = \frac{r^{-1}_{ij}}{D_{jj}} . S^2_{y.12...k} \qquad \text{.....(13)}$$

S_{b_j} = standard deviations of regression coefficients

D_{jj} = sum of squares of deviations from mean for j^{th} independent variable

14) $$t_j = \frac{b_j}{S_{b_j}} \qquad \text{.....(14)}$$

t_j = computed t

$j = 1, 2, \ldots, k$