Part III: Multiple linear regression

L15 Type II SS

1. Review: Partial F-test

   (1) Scheme
   For model $y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \epsilon$

   $H_0: \beta_1 = \beta_3 = 0$ vs $H_a: \beta_1 \neq 0$ or $\beta_3 \neq 0$
   Test Statistic: $F = \frac{MSH}{MSE}$
   Reject H_0 if $F > F_\alpha(2, n-4)$ | P-value: $P(F(2,\, n-4) > F_{ob})$

   (2) ANOVA

Source	SS	DF	MS	F	$Pr > F$
Hypothesis	SSH	2	MSH	MSH/MSE	$P(F(2, n-4) > F_{ob})$
Error	SSE	n-4	MSE		
Reduced model Error	SSE$_r$	n-2			

proc reg; model y=x1 x2 x3; run;	$\Longrightarrow$	SSE, DF=n-4, MSE
proc reg; model y=x2; run;	$\Longrightarrow$	SSE$_r$, DF=n-2

   $SSH = SSE_r - SSE$, $q = (n-2) - (n-4) = 2$ of parameters in H_0.

   (3) SAS
   can produce ANOVA table for the test without fitting two models

```
proc reg;
    model y=x1 x2 x3;
    test x1=0, x3=0;
    run;
```

$\Longrightarrow$

	MS	DF	F	$Pr > F$
Numerator	MSH	2	MSH/MSE	$P(F(2, n-4) > F_{ob})$
Denominator	MSE	n-4		

2. Inference on β_i
   Consider the model in (1) of 1.

   (1) t-tests
   Consider lower-sided H_a.

   $H_0: \beta_2 \geq 3$ versus $H_a: \beta_2 < 3$
   Test statistic: $t = \frac{\widehat{\beta}_2 - 3}{S_{\widehat{\beta}_2}}$
   Reject H_0 if $t < -t_\alpha(n-4)$ | P-value: $P(t(n-4) < t_{ob})$

   (2) t-intervals
   Consider lower-sided CI.
   $\left(-\infty,\, \widehat{\beta}_2 + t_{0.05}(n-4) S_{\widehat{\beta}_2}\right)$ is a 95% lower-sided CI for β_2.

(3) Tool
proc reg; model y=x1 x2 x3/clb alpha=0.10; run;

β_i	$\widehat{\beta}_i$	$S_{\widehat{\beta}_i}$	$t = \frac{\widehat{\beta}_i}{S_{\widehat{\beta}_i}}$	$2P(t(n-p) > \lvert t_{ob}\rvert)$	L	U
β_0	$\widehat{\beta}_0$	$S_{\widehat{\beta}_0}$	$t = \frac{\widehat{\beta}_0}{S_{\widehat{\beta}_0}}$	$2P(t(n-p) > \lvert t_{ob}\rvert)$	L_0	U_0
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
β_3	$\widehat{\beta}_3$	$S_{\widehat{\beta}_3}$	$t = \frac{\widehat{\beta}_3}{S_{\widehat{\beta}_3}}$	$2P(t(n-p) > \lvert t_{ob}\rvert)$	L_3	U_3

3. Type II SS
Consider the model in (1) of 1.

(1) Test schemes

$H_0:\ \beta_2 = 0$ versus $H_a:\ \beta_2 \neq 0$
Test Statistic: $F = \frac{SSH}{MSE}$
Reject H_0 if $F > F_\alpha(1,\, n-4)$ | p-value: $P(F(1,\, n-4) > F_{ob})$.

$SSH = SSE_r - SSE$ can be obtained by fitting data to full and reduced models or SSH, MSE, F and p can be obtained by "test" statements.

(2) Type II SS

```
proc reg;
   model y=x1 x2 x3/ss2;
   run;
```

will add a column Type II SS to parameter table to include SSII_0, SSII_1, SSII_2 and SSII_3.

$$\begin{aligned} SSII_2 &= SS(\beta_2|\beta_0,\, \beta_1,\, \beta_3) = SSE(\beta_0,\, \beta_1,\, \beta_3) - SSE(\beta_0,\, \beta_1,\, \beta_2,\, \beta_3) \\ &= SSE_r - SSE = SSH \text{ for } H_0:\ \beta_2 = 0. \end{aligned}$$

Thus in F-test in (1) of 3, $F = \frac{SSII_2}{MSE}$ where both numerator and denominator are in SAS output.

Ex1: For model $y = \beta_1 x_1 + \beta_2 x_2 + \epsilon$, $\epsilon \sim N(0,\, \sigma^2)$

proc reg; model y=x1 x2 x3/noint ss2; run;

produces MSE, n-p, SSII_1, SSII_2 and SSII_3 needed in

$H_0:\ \beta_i = 0$ vs $H_a:\ \beta_i \neq 0$
Test statistic: $F = \frac{SSII_i}{MSE}$
Reject H_0 if $F > F_\alpha(1,\, n-p)$ | p-value: $P(F(1,\, n-p) > F_{ob})$

Ex2: Testing $H_0:\ \beta_i = 0$ using $F = \frac{SSII_i}{MSE}$ produced p-value: p_i
The contribution from X_i is greater or equal to that from $X_j \Longleftrightarrow p_i \leq p_j$
$\Longleftrightarrow \frac{SSII_i}{MSE} \geq \frac{SSII_j}{MSE} \Longleftrightarrow SSII_i \geq SSII_j$

L16: General F-tests in regression

$H_0: A\beta = b$ versus $H_a: A\beta \neq b$ Test statistic: $F = \frac{(A\widehat{\beta}-b)'[A(X'X)^{-1}A']^{-1}(A\widehat{\beta}-b)/q}{MSE}$ Reject H_0 if $F > F_\alpha(q, n-p)$ | P-value: $P(F(q, n-p) > F_{ob})$

1. Hypotheses

   (1) q equations
   In $H_0: A\beta = b$, A is a $q \times p$ matrix, $\beta \in R^p$ and $b \in R^q$. Thus H_0 gives q equations.
   $$\begin{cases} a_{11}\beta_0 + \beta_{12}\beta_1 + \cdots + a_{1p}\beta_{p-1} & = & b_1 \\ \vdots & \vdots & \vdots \\ a_{q1}\beta_0 + a_{q2}\beta_1 + \cdots + a_{qp}\beta_{p-1} & = & b_q \end{cases}.$$

   (2) Consistency and linearly independent rows of A
   It is assumed that the equations are consistent, i.e., the equations have solutions.
   It is also assumed that the rows of A are linearly independent.

   Ex1: For $y = \beta_1 x_1 + \beta_2 x_2 + \epsilon$ with $H_0: A\beta = b$ where $A = \begin{pmatrix} 1 & 2 \\ 3 & 6 \end{pmatrix}$ with linearly dependent rows. If $b = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$, then system is inconsistent; if $b = \begin{pmatrix} 1 \\ 3 \end{pmatrix}$, then H_0 is reduced to contain $\beta_1 + 2\beta_2 = 1$ only.

   Ex2: $H_0: A\beta = b$ covers all cases we studied so far. For example with model $y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \epsilon$,

   $H_0: \beta_i = 0$ for all $i = 1, 2, 3$ is $A\beta = b$ where $A = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$ and $b = 0 \in R^3$.

   $H_0: \beta_1 = \beta_3 = 0$ is $A\beta = b$ with $A = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$ and $b = 0 \in R^2$.

   $H_0: \beta_2 = 0$ is $A\beta = 0$ with $A = (0, 0, 1, 0)$ and $b = 0$.

2. Test statistic

   (1) LRT statistic
   Let $SSH = SSE_r - SSE$. Then SSH is an SS with DF=(DF of SSE$_r$)-(DF of SSE)=q.
   Let MSH= $\frac{SSH}{q}$ and $F = \frac{MSH}{MSE}$. Then this F is a LRT.

   Proof. $\Lambda = \frac{\max[L(\beta,\sigma^2)]}{\max[L(\beta,\sigma^2)\colon H_0]} = \left(\frac{SSE_r}{SSE}\right)^{n/2} = \left(\frac{MSH}{MSE} \cdot \frac{q}{n-p} + 1\right)^{n/2} = \left(F \cdot \frac{q}{n-p} + 1\right)^{n/2}$
   is an increasing function of $F = \frac{MSH}{MSE}$. Thus F is a LRT statistic.

   (2) A specific form: $F = \frac{(A\widehat{\beta}-b)'[A(X'X)^{-1}A']^{-1}(A\widehat{\beta}-b)/q}{MSE}$
   It can be shown that $SSH = SSE_r - SSE = (A\widehat{\beta} - b)'[A(X'X)^{-1}A']^{-1}(A\widehat{\beta} - b)$.
   Thus $F = \frac{(A\widehat{\beta}-b)'[A(X'X)^{-1}A']^{-1}(A\widehat{\beta}-b)/q}{MSE}$ is a more specific form of F.

   Ex2: $H_0: \beta_i = 0$ can be written as $e_i'\beta = 0$. Thus
   $$\begin{aligned} (A\widehat{\beta} - b)[A(X'X)^{-1}A']^{-1}(A\widehat{\beta}_i - b) &= (e_i'\widehat{\beta} - 0)[e_i'(X'X)^{-1}e_i]^{-1}(e_i'\widehat{\beta} - 0) \\ &= \frac{\widehat{\beta}_i^2}{e_i'(X'X)^{-1}e_i}. \end{aligned}$$

With $q=1$, $F=\frac{\widehat{\beta}_i^2}{MSE\, e_i'(X'X)^{-1}e_i} = \left(\frac{\widehat{\beta}_i}{S_{\widehat{\beta}_i}}\right)^2 = t^2$.

3. Implementation

(1) Full model and reduced model approach

For model $y=\beta_0+\beta_1x_1+\beta_2x_2+\beta_3x_3+\epsilon$, $\epsilon\sim N(0,\sigma^2)$.

> $H_0: \beta_1=\beta_3,\ \beta_1+\beta_3=2\beta_2$
> H_a: At least one equation in H_0 is false
> Test statistic: $F=\frac{MSH}{MSE}$
> p-value: $P(F(2,n-4)>F_{ob})$

(i) Statistics from full model `proc reg; model y=x1 x2 x3; run;`

$\Longrightarrow SSE=97.6579,\ DF=8,\ MSE=12.2072.$

(ii) Reduced model

$H_0: \begin{pmatrix}1 & 2 & -1 & 0\\ 0 & 0 & 1 & -1\end{pmatrix}\beta=\begin{pmatrix}1\\2\end{pmatrix} \Longleftrightarrow \begin{cases}\beta_0+2\beta_1-\beta_0=1\\ \beta_2-\beta_3=2\end{cases} \Longleftrightarrow \begin{cases}\beta_0=1+\beta_2-2\beta_1\\ \beta_3=\beta_2-2\end{cases}$

Model under H_0 becomes

$$\begin{aligned} y &= \beta_0+\beta_1x_1+\beta_2x_2+\beta_3x_3+\epsilon\\ &= (1+\beta_2-2\beta_1)+\beta_1x_1+\beta_2x_2+(\beta_2-2)x_3+\epsilon\\ &= \beta_1(x_1-2)+\beta_2(x_2+x_3+1)+1-2x_3+\epsilon, \end{aligned}$$

i.e., $y+2x_3-1=\beta_1(x_1-2)+\beta_2(x_2+x_3+1)+\epsilon$, $\epsilon\sim N(0,\sigma^2)$.

(iii) Statistics from the reduced model

```
data b; set a;
  ny=y+2*x3-1; nx1=x1-2; nx2=x2+x3+1;
proc reg;
   model ny=nx1 nx2/noint;
   run;
```

$\Longrightarrow SSE_r=638.13031,\ DF=10.$

(iv) With results from (1) and (3),

Source	DF	SS	MS	F	p
Hypothesis	2	540.47242	270.23621	22.7314	0.001
Error	8	97.6579	12.20724		
Error (R)	10	638.13031			

(2) Use "test"

```
proc reg;
   model y=x1 x2 x3;
   test intercept=1+x2-2*x1, x3=x2-2;
   run;
```

$\Longrightarrow$

	DF	MS	P	p
N	2	270.23621	22.7314	0.001
D	8	12.20724		