機械所設計組 R12522615 王邑安 HW_ID:2

Introduction to statistical control and optimization

Final Exam

1.

Three design factors in this experiment:

code	paper height(mm)	width-height ratio	leg length(mm)
-1	148.5	0.4	8
1	210	0.5	10

2.

Unwanted noises:

I. The material property of papers may vary in each frog.
II. The humidity in air will affect the elasticity of paper.
III. The angle between the finger and a frog will vary in each jumping test.
IV. The friction coefficient between frog and finger may change because of sweat on hand.
V. The finger's force and placement on a frog will vary in each jumping test.

3.

- X1: paper height
- X2: width-height ratio
- X3: leg length

2^3 factorial experiment:

test	x1	x2	x3
1	-1	-1	-1
2	1	-1	-1
3	-1	1	-1
4	1	1	-1
5	-1	-1	1
6	1	-1	1
7	-1	1	1
8	1	1	1

In my experiment, I will create three identical frogs for each combination of factors. To ensure randomization, I will randomize the order of the jumping trials instead of following a fixed sequence for each run. Since there are three runs, each frog will perform three jumps.

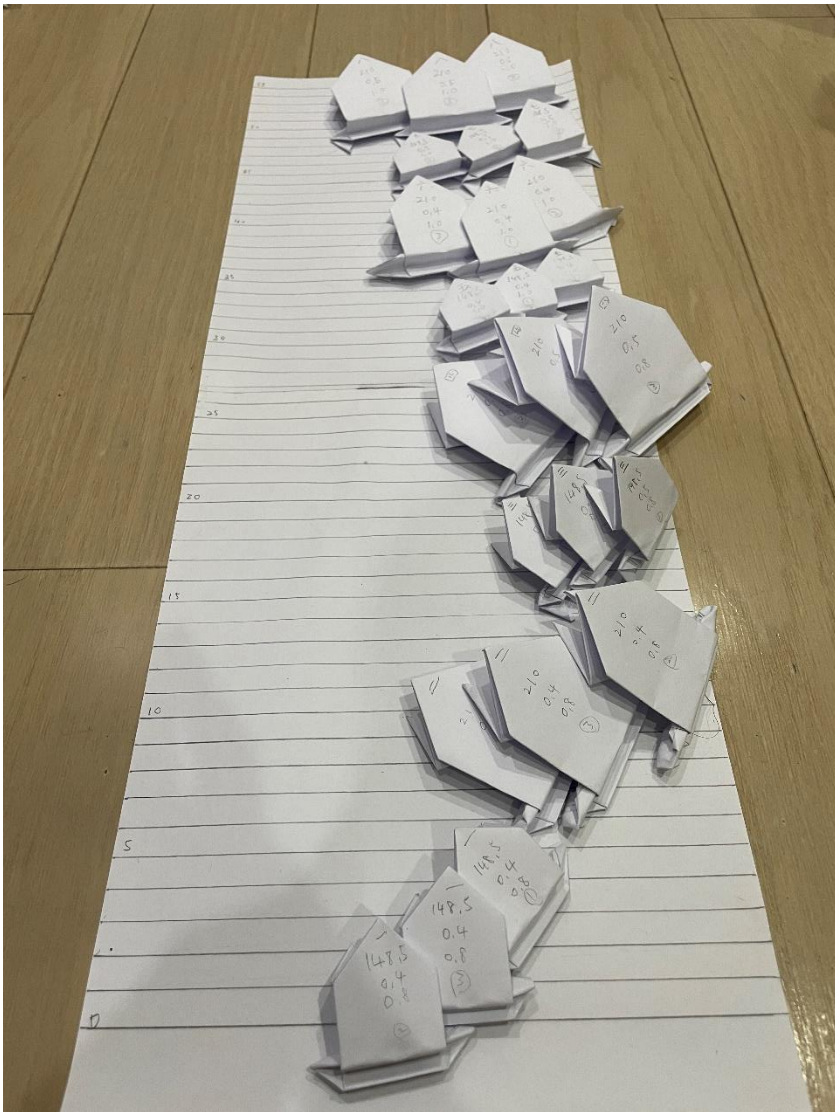
Experiment table:

				run1						run2						run3					
test	paper height(mm)	width-height ratio	leg length(mm)	replication1	order	replication2	order	replication3	order	replication1	order	replication2	order	replication3	order	replication1	order	replication2	order	replication3	order
1	148.5	0.4	8																		
2	210	0.4	8																		
3	148.5	0.5	8																		
4	210	0.5	8																		
5	148.5	0.4	10																		
6	210	0.4	10																		
7	148.5	0.5	10																		
8	210	0.5	10																		

How to handle the noises:

- To prevent I using my fingernails to press paper frogs, I always cut my nails before I do experiment.
- Do more replication and randomize the experimental trials to make noises more uniformly and more randomly distributed.

4.



5.

Experiment details:

- Day: 6/10 Monday
- time: 11:00am~12:30pm
- place: home
- witness: 王邑宇 (reason: he is my brother)

measurement method:

- Measurement plate: Design a plate with markings ranging from 0 to 55 cm (as the picture in 4.)
- Align frogs: Position each paper frog so that its rear edge aligns with the zero line on the measurement plate before each jump.
- Execute a jump: Use fingertip to press down on the back of the frog and release it to initiate the jump
- Valid jumps: If the frog rolls or bounces twice or more, disregard the result and perform the jump again.
- Distance measurement: For each valid jump, measure the shortest distance between the zero line and the landing point of the frog. This distance is the result for that trial.

experiment result:

				run1						run2						run3					
test	paper height(mm)	width-height ratio	leg length(mm)	replication1	order	replication2	order	replication3	order	replication1	order	replication2	order	replication3	order	replication1	order	replication2	order	replication3	order
1	148.5	0.4	8	255	23	240	11	280	10	300	24	310	14	300	19	285	11	300	21	360	4
2	210	0.4	8	200	17	220	15	150	12	205	4	155	2	165	11	230	15	200	16	200	23
3	148.5	0.5	8	150	18	190	24	220	7	175	21	150	6	195	17	160	6	190	5	205	20
4	210	0.5	8	170	4	160	3	150	9	210	23	170	12	155	9	160	24	185	3	200	2
5	148.5	0.4	10	300	21	230	13	275	2	290	10	285	18	295	16	345	22	270	7	385	12
6	210	0.4	10	140	16	200	6	145	20	150	7	150	1	140	3	145	18	185	17	135	10
7	148.5	0.5	10	175	22	230	14	245	19	210	8	210	13	210	5	230	9	245	13	275	19
8	210	0.5	10	180	1	200	5	200	8	175	20	250	15	175	22	205	8	205	14	205	1

6.

Parameters of this experiment: $m = 8, n = 9 \Rightarrow N = 72$

Average and sample variance in each test: $\bar{y}_i = \sum_{j=1}^{n} \frac{y_j}{n},\ s_i^2 = \frac{\sum_{j=1}^{n}(y_j - \bar{y})^2}{n-1}$

				replicate 1			replicate 2			replicate 3				
test	x1	x2	x3	run1	run2	run3	run1	run2	run3	run1	run2	run3	y_bar	s_i^2
1	-1	-1	-1	255	300	285	240	310	300	280	300	360	292.2222	1175.694
2	1	-1	-1	200	205	230	220	155	200	150	165	200	191.6667	806.25
3	-1	1	-1	150	175	160	190	150	190	220	195	205	181.6667	606.25
4	1	1	-1	170	210	160	160	170	185	150	155	200	173.3333	431.25
5	-1	-1	1	300	290	345	230	285	270	275	295	385	297.2222	1994.444
6	1	-1	1	140	150	145	200	150	185	145	140	135	154.4444	502.7778
7	-1	1	1	175	210	230	230	210	245	245	210	275	225.5556	815.2778
8	1	1	1	180	175	205	200	250	205	200	175	205	199.4444	527.7778

Effect coefficient:

E1	E2	E3	E12	E23	E13	E123
-69.4444	-38.8889	9.444444	52.22222	25.55556	-15	6.111111111

Assuming all effects are null effect: $H_0: E_i = 0\,,\ H_1: E_i \neq 0\,,\ \alpha = 0.05$

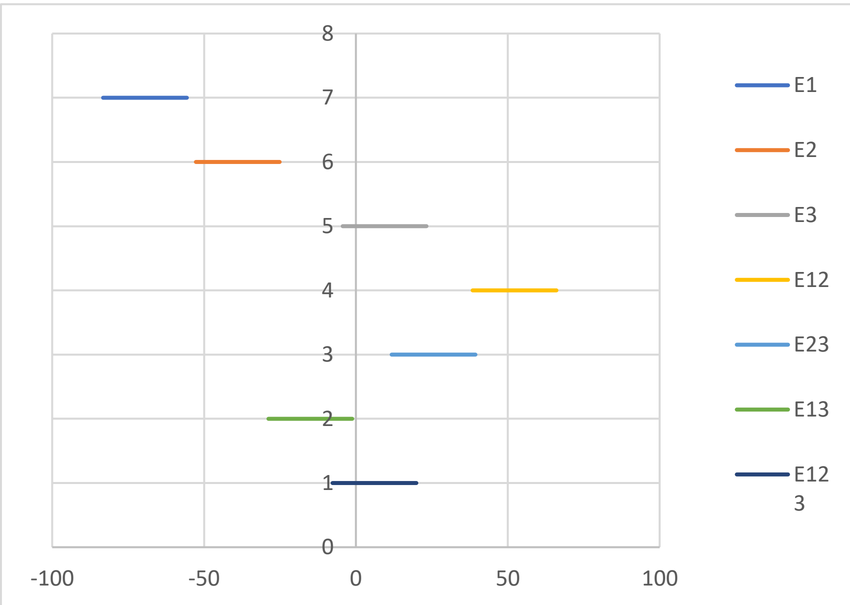
Degree of freedom for t-test: $\nu = m(n-1) = 8 \times (9-1) = 64$

$$t_{E_i} = \frac{E_i - 0}{s_{effect}}, \qquad t_{E_i} < t_{64,0.025} \text{ or } t_{E_i} > t_{64,0.975} \Rightarrow \text{reject } H_0$$

$$\text{where } t_{64,0.025} = -1.997729654,\ t_{64,0.975} = 1.997729654$$

$$s_{effect}^2 = \widehat{Var}(E) = \frac{4s_p^2}{mn} = 47.63695988, \qquad s_p^2 = \frac{\sum_i^m s_i^2}{m} = 857.4652778$$

		t-test	p-value	t_64,0.025	t_64,0.975
E1	-69.4444	-10.0616	8.28E-15	-1.99773	1.997729654
E2	-38.8889	-5.63448	4.23E-07	-1.99773	1.997729654
E3	9.444444	1.368373	0.17598	-1.99773	1.997729654
E12	52.22222	7.566296	1.88E-10	-1.99773	1.997729654
E23	25.55556	3.702656	0.000446	-1.99773	1.997729654
E13	-15	-2.1733	0.033465	-1.99773	1.997729654
E123	6.111111	0.885418	0.379246	-1.99773	1.997729654



$E_1, E_2, E_{12}, E_{23}, E_{13}$ reject H_0, they are significant factors.

$$\hat{y} = b_0 + b_1x_1 + b_2x_2 + b_{12}x_1x_2 + b_{23}x_2x_3 + b_{13}x_1x_3$$

$= 214.4444444 - 34.72222222x_1 - 19.44444444x_2 + 26.11111111x_1x_2 + 12.77777778x_2x_3 - 7.5x_1x_3$

7.

source of variation	SS	DOF	MS	F	$F_{0.05}$
E1	86805.56	1	86805.56	100.2381	3.986269
E2	27222.22	1	27222.22	31.43468	3.986269
E12	49088.89	1	49088.89	56.68507	3.986269
E23	11755.56	1	11755.56	13.57465	3.986269
E13	4050	1	4050	4.676711	3.986269
error	57155.56	66	865.9933		
total	236077.8	71			

$$R^2 = \frac{SSR}{SST} = \frac{SS_{E_1} + SS_{E_2} + SS_{E_{12}} + SS_{E_{23}} + SS_{E_{13}}}{SST} = 0.757895232$$

$$\text{adjusted } R^2 = 1 - \frac{SSE/dof}{SST/dof} = \frac{57155.55556/66}{236077.7778/71} = 0.739553962$$

8.

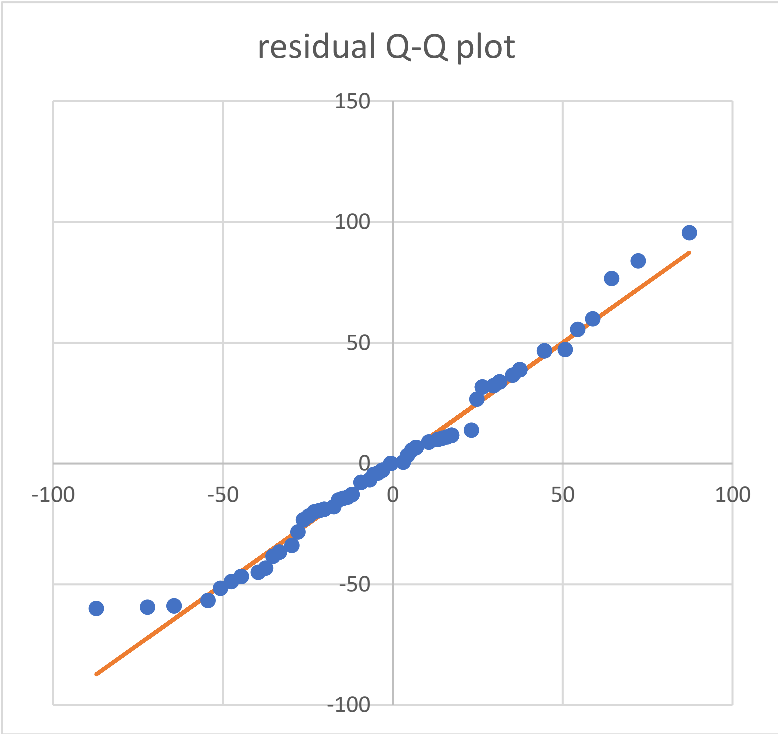
Suppose that residuals are normal distributed. The parameters of normal distribution are $\bar{e}$ and σ_e:

$$\bar{e} = \sum_{i=1}^{8} \sum_{j=1}^{9} \frac{e_{ij}}{72} = -5.92119E - 15$$

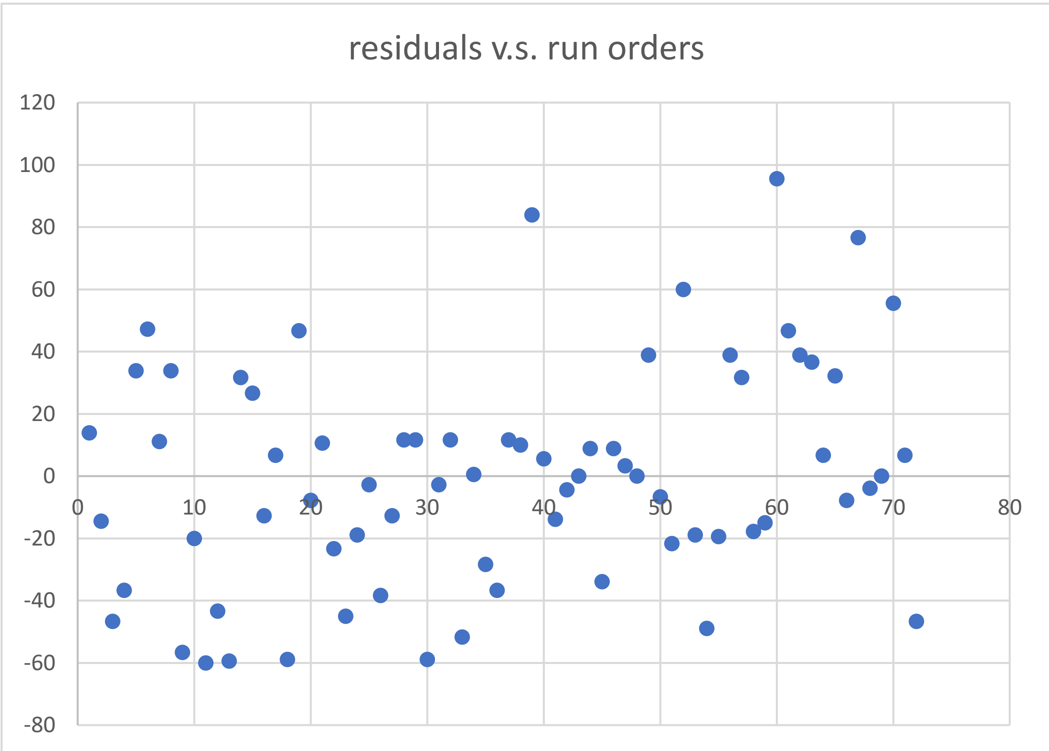
$$\sigma_e = \sum_{i}^{8} \sum_{j}^{9} \frac{\left(e_{ij} - \bar{e}\right)^2}{72 - 1} = 35.4757516$$

- Rank each residual from 1~72
- The c.d.f. of $i\ th$ effect is: $\hat{p}_i = (i - 0.5)/72$
- The normal distribution value of $i\ th$ effect: $N^{-1}(\hat{p}_i; \bar{e}, \sigma_e)$

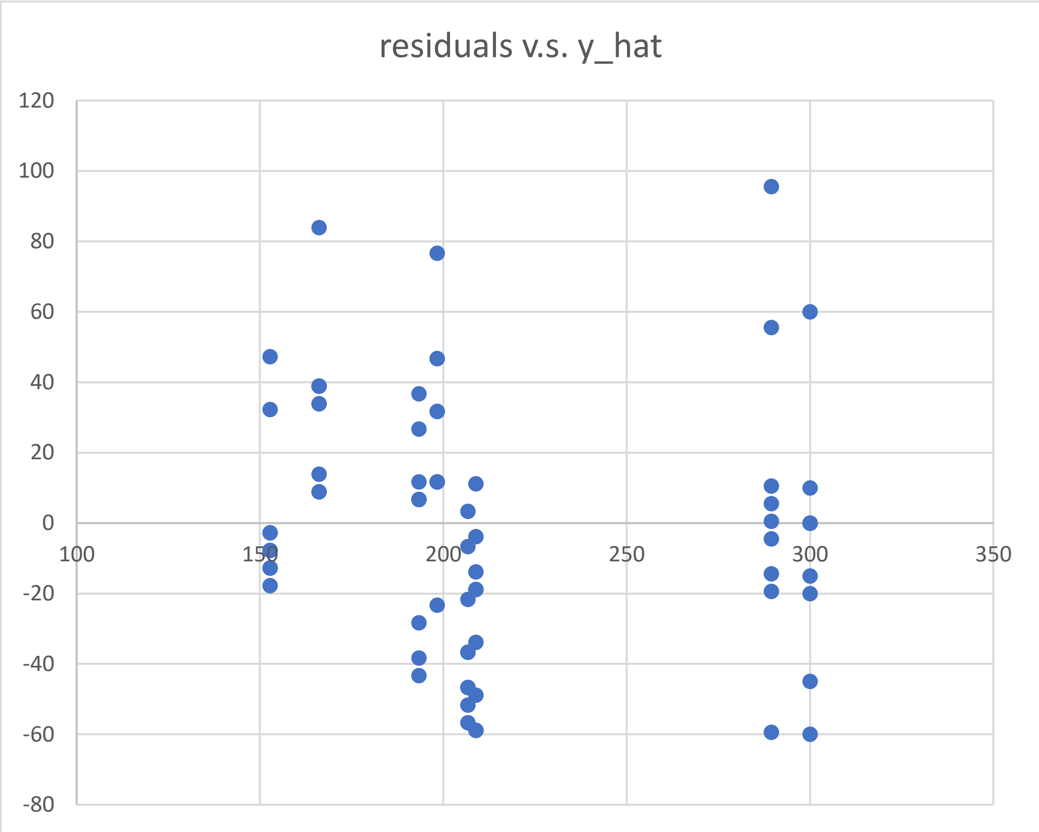
Residual Q-Q plot:



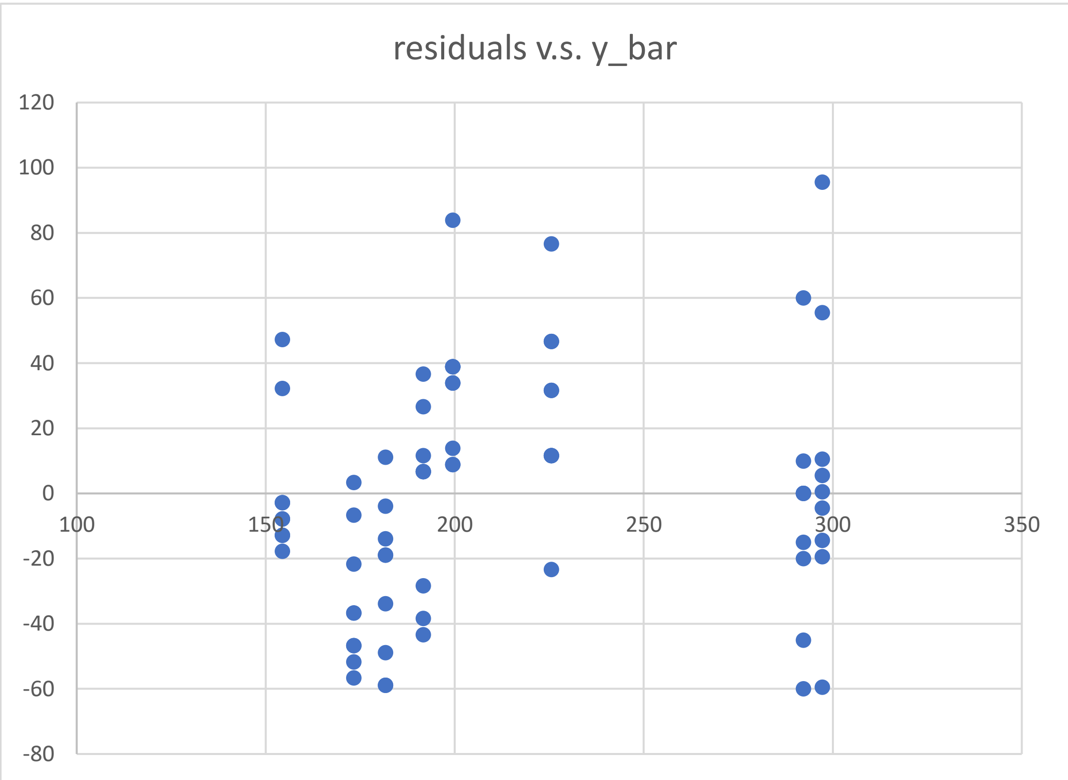
Plot the residuals vs. run orders:



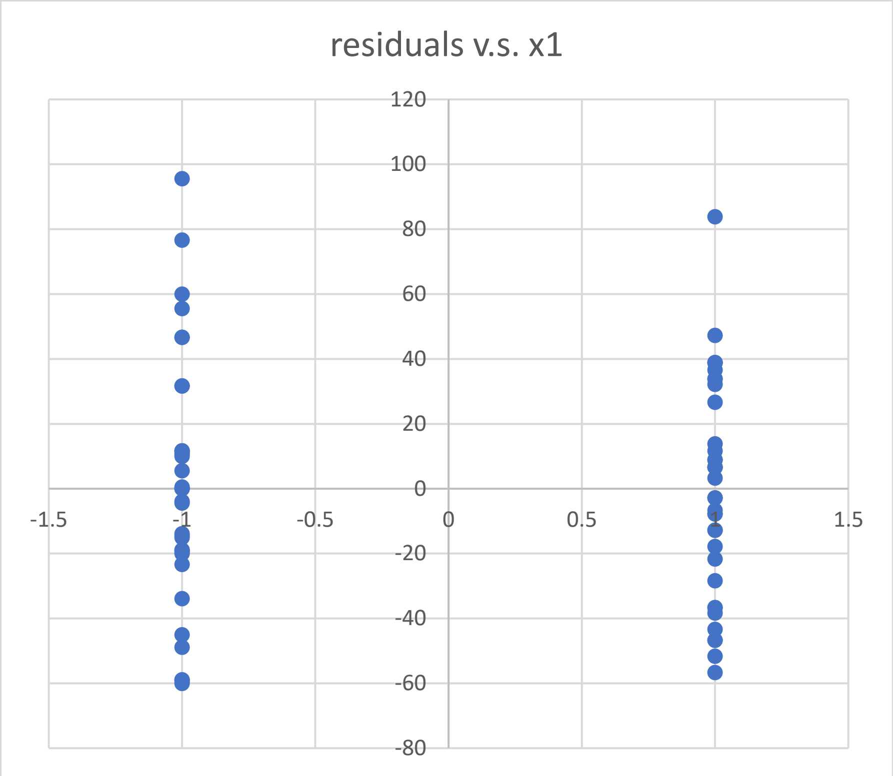
Plot the residuals vs. $\hat{y}$:



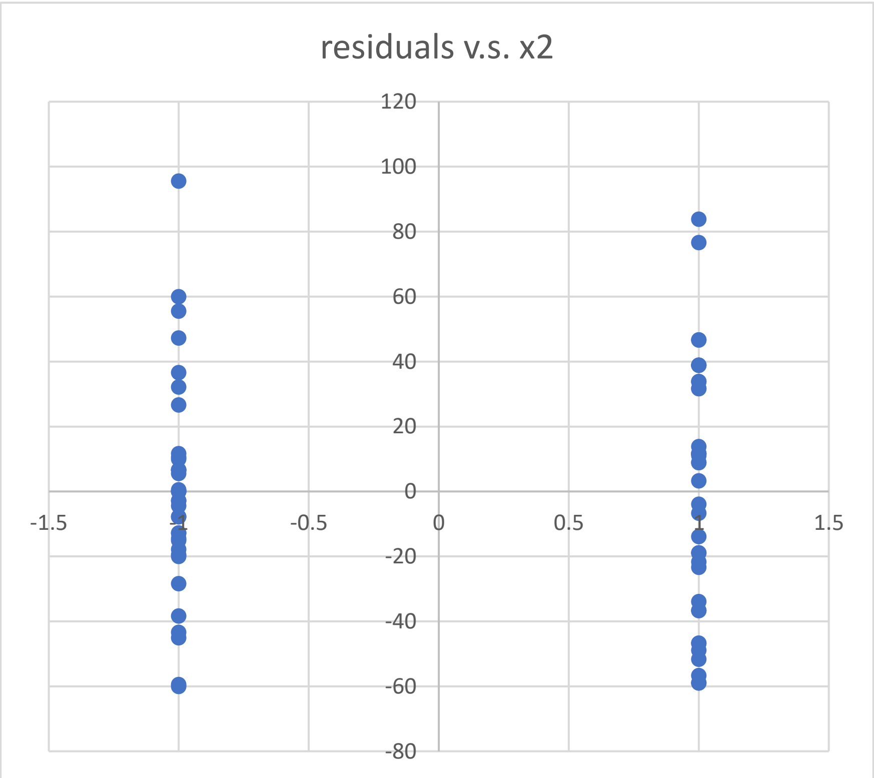
Plot the residuals vs. $\bar{y}$:



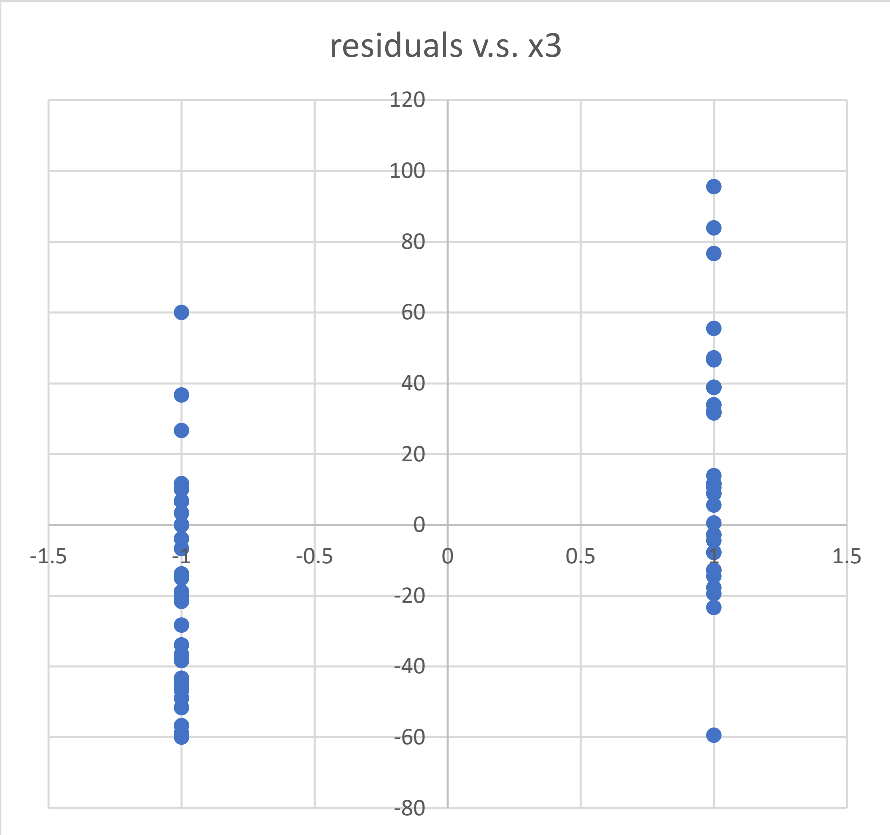
Plot the residuals vs. x1:



Plot the residuals vs. x2:



Plot the residuals vs. x3:



Bartlett's test ($\alpha = 0.05$):

$$H_0: \sigma_1^2 = \sigma_2^2 = \cdots = \sigma_m^2 = \sigma_\epsilon^2$$

$$H_1\text{:at least one } \sigma_a^2 \neq \sigma_b^2 \text{ , } a \neq b$$

Reject H_0 if $\mathcal{X}_{calc}^2 > \mathcal{X}_{m-1,\alpha}^2 = \mathcal{X}_{7,0.05}^2$:

$$\mathcal{X}_{calc}^2 = \frac{M}{c}$$

$$M = (N-m)\ln s_p^2 - \sum_{i=1}^{m}(n_i - 1)\ln s_i^2 \text{ , } s_p^2 = \sum_{i}^{m}\frac{(n_i-1)s_i^2}{N-m}$$

$$c = 1 + \frac{1}{3(m-1)}[\left(\sum_{i=1}^{m}\frac{1}{n_i - 1}\right) - \frac{1}{N-m}]$$

$$\Rightarrow \mathcal{X}_{calc}^2 = 7.682931678 \text{ , } \mathcal{X}_{7,0.05}^2 = 14.06714045$$

$$\Rightarrow \mathcal{X}_{calc}^2 < \mathcal{X}_{m-1,\alpha}^2 \text{ , } accept\ H_0$$

In the residual plots, variability of the residuals tends to increase as the predicted values ($\hat{y}$) or the actual average values ($\bar{y}$) rise, with this effect being particularly pronounced for the two highest values of $\hat{y}$ and $\bar{y}$. This may suggest that the variation of the test 1 and test 5 is too high.

9.

SN ratio: Larger-the-best

$$\eta = -10\log_{10}\left[\frac{1}{n}\sum_{i=1}^{n}\left(\frac{1}{X_i}\right)^2\right]$$

				replicate 1			replicate 2			replicate 3			
test	x1	x2	x3	run1	run2	run3	run1	run2	run3	run1	run2	run3	SN
1	-1	-1	-1	255	300	285	240	310	300	280	300	360	49.15775
2	1	-1	-1	200	205	230	220	155	200	150	165	200	45.37338
3	-1	1	-1	150	175	160	190	150	190	220	195	205	44.96722
4	1	1	-1	170	210	160	160	170	185	150	155	200	44.62473
5	-1	-1	1	300	290	345	230	285	270	275	295	385	49.21638
6	1	-1	1	140	150	145	200	150	185	145	140	135	43.57489
7	-1	1	1	175	210	230	230	210	245	245	210	275	46.87
8	1	1	1	180	175	205	200	250	205	200	175	205	45.85851

Effect coefficient:

E1	E2	E3	E12	E23	E13	E123
-2.694963152	-1.250482833	0.349173936	2.017970648	1.219103481	-0.631532694	0.297030052

Assuming all effects are null effects, and estimated standard error by the lowest three effects. $H_0: E_i = 0$, $H_1: E_i \neq 0$, $\alpha = 0.05$

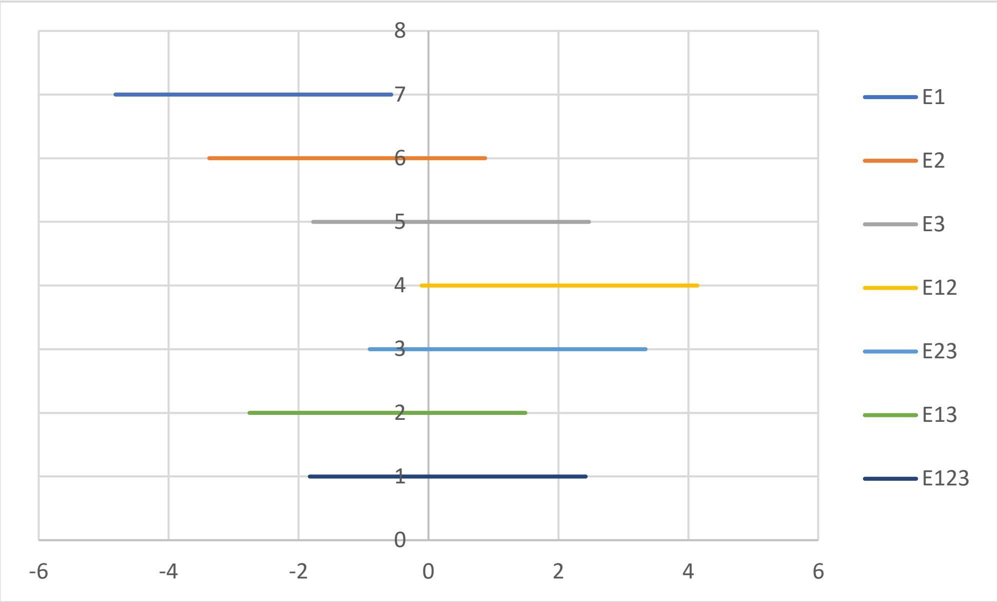
E_{13} and E_{123} are relatively small, therefore, they are used for testing the statistical significance of main effects and larger interaction effects.

$$s_{effect}^2 = \frac{(E_{13}-0)^2 + (E_{123}-0)^2}{2}$$

$$t_{E_i} = \frac{E_i - 0}{s_{effect}}, \qquad t_{E_i} < t_{2,0.025} \text{ or } t_{E_i} > t_{2,0.975} \Rightarrow \text{reject } H_0$$

$$\text{where } t_{2,0.025} = -4.30265273,\ t_{2,0.975} = 4.30265273$$

	effect	t-test	p-value	t_2,0.025	t_2,0.975
E1	-2.694963152	-5.461053371	0.031934	-4.30265	4.302653
E2	-1.250482833	-2.533969152	0.126788	-4.30265	4.302653
E3	0.349173936	0.707563479	0.552555	-4.30265	4.302653
E12	2.017970648	4.089200775	0.054923	-4.30265	4.302653
E23	1.219103481	2.470382266	0.132146	-4.30265	4.302653
E13	-0.631532694	-1.279733174	0.329027	-4.30265	4.302653
E123	0.297030052	0.601899497	0.608386	-4.30265	4.302653



It shows that only E_1 is significant effect to the SN ratio. Consequently, the predictive model of SN ratio is:

$$\hat{\eta} = b_0 + b_1 x_1 = 46.20535766 - 1.347481576 x_1$$

10.

Based on the predictive model of jumping distance of paper frogs, a best combination of three factor can be find:

$$\hat{y} = 214.4444444 - 34.72222222 x_1 - 19.44444444 x_2 + 26.11111111 x_1 x_2 + 12.77777778 x_2 x_3 - 7.5 x_1 x_3$$

x1	x2	x3	$\hat{y}$	x1	x2	x3	$\hat{y}$
-1	-1	-1	300	-1	-1	1	289.4444
1	-1	-1	193.3333	1	-1	1	152.7778
-1	1	-1	183.3333	-1	1	1	223.8889
1	1	-1	181.1111	1	1	1	191.6667

The optimum is 300mm when the combination of three factors is $(x_1, x_2, x_3) = (-1, -1, -1)$

On the other hand, the predictive model of SN ratio is:

$$\hat{\eta} = 46.20535766 - 1.347481576 x_1$$

x1	$\hat{\eta}$
-1	47.55283923
1	44.85787608

It shows that when $x_1 = -1$, the jumping distance SN ratio of paper frogs is largest.

As a result, there is no conflict between the predictive model of jumping distance and the predictive model of the SN ratio for jumping distance. The estimated $\hat{\eta}$ indicates that the SN ratio performs best when $x_1 = -1$. Similarly, the estimated $\hat{y}$ shows that the optimal jumping distance occurs when $(x_1, x_2, x_3) = (-1, -1, -1)$.

$$\hat{y}|_{x_1=-1, x_2=-1, x_3=-1}$$
$$= 214.4444444 - 34.72222222x_1 - 19.44444444x_2 + 26.11111111x_1x_2 + 12.77777778x_2x_3 - 7.5x_1x_3$$
$$= 300$$
$$\hat{\eta}|_{x_1=-1} = 46.20535766 - 1.347481576x_1 = -47.55283923$$

11.

Make three optimum frogs and measure their jumping distance:

verify			
	replication1	replication2	replication3
run1 distance(mm)	270	330	360
run2 distance(mm)	290	275	340
run3 distance(mm)	355	310	330
average	317.7777778	SN	49.90908236

In the verification, the average jumping distance of the frogs at the optimal conditions was 317.78 mm. This value is higher than the most of the jumping distance tested in the previous experiment, suggesting that the regression model shows the tendency of the jumping distance in optimum factor combination. However, the average jumping distance of paper frogs in this verification is slightly larger than the predictive distance(300mm). This indicates that the regression model is likely to underestimate the performance of paper frogs. Moreover, the SN ratio of jumping distance in the verification is 49.90908236. This value is relatively high comparing to the previous experiment results. This means that the predictive model shows the trend of SN ratio of paper frogs' jumping distance. Nevertheless, the SN ratio value in verification still has some difference comparing to the value of the predictive model (-47.55283923).

12.

To measure error, the test of center point (test 15) should be executed more than once.

test	x1	x2	x3
1	-1	-1	-1
2	1	-1	-1
3	-1	1	-1
4	1	1	-1
5	-1	-1	1
6	1	-1	1
7	-1	1	1
8	1	1	1
9	-1.682	0	0
10	1.682	0	0
11	0	-1.682	0
12	0	1.682	0
13	0	0	-1.682
14	0	0	1.682
15	0	0	0
Measure error			
15--1	0	0	0
15--2	0	0	0
15--3	0	0	0
15--4	0	0	0

13.

Experiment details:

- Day: 6/10 Monday
- time: 1:30pm~2:30pm
- place: home
- witness: 王邑宇 (reason: he is my brother)

measurement method: same method in (5)

experiment result:

				run1						run2						run3					
test	paper height(mm)	width-height ratio	leg length(mm)	replication1	order	replication2	order	replication3	order	replication1	order	replication2	order	replication3	order	replication1	order	replication2	order	replication3	order
1	148.5	0.4	8	255	23	240	11	280	10	300	24	310	14	300	19	285	11	300	21	360	4
2	210	0.4	8	200	17	220	15	150	12	205	4	155	2	165	11	230	15	200	16	200	23
3	148.5	0.5	8	150	18	190	24	220	7	175	21	150	6	195	17	160	6	190	5	205	20
4	210	0.5	8	170	4	160	3	150	9	210	23	170	12	155	9	160	24	185	3	200	2
5	148.5	0.4	10	300	21	230	13	275	2	290	10	285	18	295	16	345	22	270	7	385	12
6	210	0.4	10	140	16	200	6	145	20	150	7	150	1	140	3	145	18	185	17	135	10
7	148.5	0.5	10	175	22	230	14	245	19	210	8	210	13	210	5	230	9	245	13	275	19
8	210	0.5	10	180	1	200	5	200	8	175	20	250	15	175	22	205	8	205	14	205	1
9	127.5285	0.45	9	210	37	145	31	165	35	220	45	160	29	175	39	215	31	150	43	175	39
10	230.9715	0.45	9	195	26	170	38	155	25	205	44	190	40	165	42	215	26	165	42	170	33
11	179.25	0.3659	9	205	32	175	36	225	45	185	38	185	27	180	35	180	37	225	30	205	44
12	179.25	0.5341	9	200	42	225	34	240	30	190	37	195	31	235	30	205	29	220	45	220	36
13	179.25	0.45	7.318	205	29	210	28	175	27	190	33	230	36	155	43	220	27	200	25	170	35
14	179.25	0.45	10.682	195	44	230	39	225	41	220	32	245	26	230	25	210	34	245	40	240	28
15	179.25	0.45	9	250	43	210	40	200	33	230	28	260	41	200	34	260	41	225	32	220	38

measure error (x1,x2,x3)= (0,0,0)									
test orde	replication1			replication2			replication3		
test1	220	220	215	205	260	250	200	255	245
test2	205	245	200	235	255	250	230	230	220
test3	215	245	185	250	210	205	200	215	200
test4	235	230	205	255	250	225	200	205	210

14.

Parameters of this experiment: $m = 15, n = 9 \Rightarrow N = 171$

Average and sample variance in each test: $\bar{y}_i = \sum_{j=1}^{n} \frac{y_j}{n}$, $s_i^2 = \frac{\sum_{j=1}^{n}(y_j - \bar{y})^2}{n-1}$

				replicate 1			replicate 2			replicate 3				
test	x1	x2	x3	run1	run2	run3	run1	run2	run3	run1	run2	run3	y_bar	s_i^2
1	-1	-1	-1	255	300	285	240	310	300	280	300	360	292.2222	1175.694
2	1	-1	-1	200	205	230	220	155	200	150	165	200	191.6667	806.25
3	-1	1	-1	150	175	160	190	150	190	220	195	205	181.6667	606.25
4	1	1	-1	170	210	160	160	170	185	150	155	200	173.3333	431.25
5	-1	-1	1	300	290	345	230	285	270	275	295	385	297.2222	1994.444
6	1	-1	1	140	150	145	200	150	185	145	140	135	154.4444	502.7778
7	-1	1	1	175	210	230	230	210	245	245	210	275	225.5556	815.2778
8	1	1	1	180	175	205	200	250	205	200	175	205	199.4444	527.7778
9	-1.682	0	0	210	220	215	145	160	150	165	175	175	179.4444	815.2778
10	1.682	0	0	195	205	215	170	190	165	155	165	170	181.1111	429.8611
11	0	-1.682	0	205	185	180	175	185	225	225	180	205	196.1111	379.8611
12	0	1.682	0	200	190	205	225	195	220	240	235	220	214.4444	315.2778
13	0	0	-1.682	205	190	220	210	230	200	175	155	170	195	606.25
14	0	0	1.682	195	220	210	230	245	245	225	230	240	226.6667	275
15	0	0	0	250	230	260	210	260	225	200	200	220	228.3333	562.5
15--1	0	0	0	220	220	215	205	260	250	200	255	245	230	512.5

15--2	0	0	0	205	245	200	235	255	250	230	230	220	230	362.5
15--3	0	0	0	215	245	185	250	210	205	200	215	200	213.8889	448.6111
15--4	0	0	0	235	230	205	255	250	225	200	205	210	223.8889	411.1111

Check the correlation coefficient between each factor (including quadratic terms):

	x1	x2	x3	x1x2	x2x3	x1x3	x1x2x3	x1^2	x2^2	x3^2	y
x1	1										
x2	0	1									
x3	0	0	1								
x1x2	0	0	0	1							
x2x3	0	0	0	0	1						
x1x3	0	0	0	0	0	1					
x1x2x3	0	0	0	0	0	0	1				
x1^2	0	0	0	0	0	0	0	1			
x2^2	0	0	0	0	0	0	0	-0.12814	1		
x3^2	0	0	0	0	0	0	0	-0.12814	-0.12814	1	
y	-0.39964	-0.18126	0.132315	0.396679	0.194119	-0.11394	0.04642	-0.23466	-0.03296	0.011866	1

The result shows that there is little correlation between each factor, therefore, it can be considered as no multicollinearity effect among this regression models:

$$\hat{y} = b_0 + b_1x_1 + b_2x_2 + b_3x_3 + b_{12}x_1x_2 + b_{23}x_2x_3 + b_{13}x_1x_3 + b_{123}x_1x_2x_3 + b_{1^2}x_1^2 + b_{2^2}x_2^2 + b_{3^2}x_3^2$$

Regression analysis results:

摘要輸出

迴歸統計	
R 的倍數	0.692915
R 平方	0.480131
調整的 R 平方	0.447639
標準誤	31.83754
觀察值個數	171

ANOVA

	自由度	SS	MS	F	顯著值
迴歸	10	149783.7	14978.37	14.77697	2.02E-18
殘差	160	162180.6	1013.629		

總和	170	311964.3

	係數	標準誤	t 統計	P-值	下限 95%	上限 95%
截距	224.557	4.740528	47.36962	4.68E-96	215.1949	233.9191
x1	-20.1325	2.871579	-7.01095	6.28E-11	-25.8036	-14.4614
x2	-9.1314	2.871579	-3.17992	0.001769	-14.8025	-3.46031
x3	6.665651	2.871579	2.321249	0.021535	0.994565	12.33674
x1x2	26.11111	3.75209	6.959084	8.33E-11	18.7011	33.52112
x2x3	12.77778	3.75209	3.405509	0.000835	5.367769	20.18779
x1x3	-7.5	3.75209	-1.99889	0.047314	-14.91	-0.08999
x1x2x3	3.055556	3.75209	0.814361	0.416649	-4.35445	10.46556
x1^2	-12.2111	2.871912	-4.2519	3.59E-05	-17.8828	-6.53933
x2^2	-3.37442	2.871912	-1.17497	0.241751	-9.04616	2.297324
x3^2	-1.41072	2.871912	-0.49121	0.62395	-7.08246	4.261025

First, the table shows that $E_1, E_2, E_3, E_{12}, E_{23}, E_{13}, E_{1^2}$ reject the null hypothesis with $\alpha = 0.05$. Therefore, they are significant effects. Second, look closely on the ANOVA table, the p-value(“顯著值”) of F-test is significantly small. This indicates that the observed differences between tests are not caused by noise. Finally, the R^2 and adjusted R^2 are 0.480131 and 0.447639 respectively. This implies that the proportion of the variation explained by model is almost half.

15.

Based on the first regression analysis table, a predictive regression model of the paper frogs’ jumping distance can be constructed:

$$\hat{y} = b_0 + b_1x_1 + b_2x_2 + b_3x_3 + b_{12}x_1x_2 + b_{23}x_2x_3 + b_{13}x_1x_3 + b_{1^2}x_1^2$$

Final regression analysis:

摘要輸出

迴歸統計	
R 的倍數	0.687861
R 平方	0.473153
調整的 R 平方	0.450528
標準誤	31.75417
觀察值個數	171

ANOVA

	自由度	SS	MS	F	顯著值
迴歸	7	147607	21086.71	20.91256	6.29E-20
殘差	163	164357.4	1008.327		
總和	170	311964.3			

	係數	標準誤	t 統計	P-值	下限 95%	上限 95%
截距	220.6764	3.158602	69.86522	2.1E-123	214.4393	226.9134
x1	-20.1325	2.864059	-7.02935	5.39E-11	-25.7879	-14.477
x2	-9.1314	2.864059	-3.18827	0.001716	-14.7868	-3.47596
x3	6.665651	2.864059	2.327344	0.021177	1.010209	12.32109
x1x2	26.11111	3.742265	6.977355	7.17E-11	18.72154	33.50068
x2x3	12.77778	3.742265	3.41445	0.000807	5.388209	20.16735
x1x3	-7.5	3.742265	-2.00413	0.046712	-14.8896	-0.11043
x1^2	-11.5979	2.809925	-4.12747	5.84E-05	-17.1464	-6.04935

The analysis table shows that all effects in the predictive model are significant. The F-test value in ANOVA table is still large and the p-value of F-test is relatively small, implying that the observed differences between vary factors are statistically significant and not merely due to random noise. The R^2 and adjusted R^2 are 0.473153 and 0.450528, respectively. This suggests that almost half of the variation could be explained by this model.

Compare to the predictive model built in (6), several similarities and differences are observed.

Predictive model in (6):

$$\begin{aligned}\hat{y} &= b_0 + b_1x_1 + b_2x_2 + b_{12}x_1x_2 + b_{23}x_2x_3 + b_{13}x_1x_3 \\ &= 214.4444444 - 34.72222222x_1 - 19.44444444x_2 \\ &\quad + 26.11111111x_1x_2 + 12.77777778x_2x_3 - 7.5x_1x_3\end{aligned}$$

Predictive model in (15):

$$\begin{aligned}\hat{y} &= b_0 + b_1x_1 + b_2x_2 + b_3x_3 + b_{12}x_1x_2 + b_{23}x_2x_3 + b_{13}x_1x_3 + b_{1^2}x_1^2 \\ &= 220.6763865 - 20.13248291x_1 - 9.131397298x_2 \\ &\quad + 6.665650756x_3 + 26.11111111x_1x_2 + 12.77777778x_2x_3 \\ &\quad - 7.5x_1x_3 - 11.59789296x_1^2\end{aligned}$$

- Similarity:
  - ✓ Both models identify x_1(length of papers), x_2(width ratio of papers),

x_1x_2(interaction effect between length and width ratio), x_2x_3(interaction effect between width ratio and leg length), and x_1x_3(interaction effect between length and leg length) as significant effects. This suggests that these main effects and interactions effects are crucial in determining the jumping distance of the paper frogs.

- Difference:
  - ✓ The current model includes x_3(length of the paper frog's legs) as a significant factor. This highlights that the leg length of the paper frog plays a vital role in influencing the jumping distance, which was not directly considered in the previous model.
  - ✓ The current model considers the quadratic term of x_1, indicating a nonlinear relationship between frogs' jumping distance and papers length.
  - ✓ Comparing to the previous model, the coefficients for x_1 and x_2 are significantly reduced in the current model. This suggests that, while x_1 and x_2 remain important, other factors or interactions might play a more dominant role in determining the jumping distance in the current context.

16.

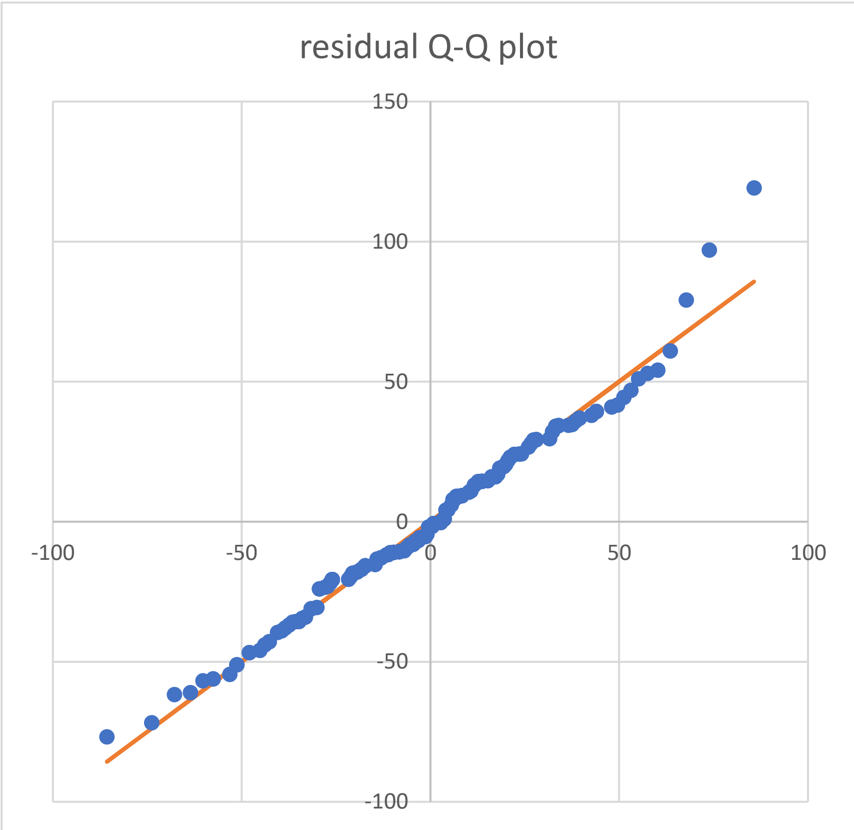
Suppose that residuals are normal distributed. The parameters of normal distribution are $\bar{e}$ and σ_e:

$$\bar{e} = \sum_{i=1}^{19}\sum_{j=1}^{9}\frac{e_{ij}}{171} = -1.26319E-15$$

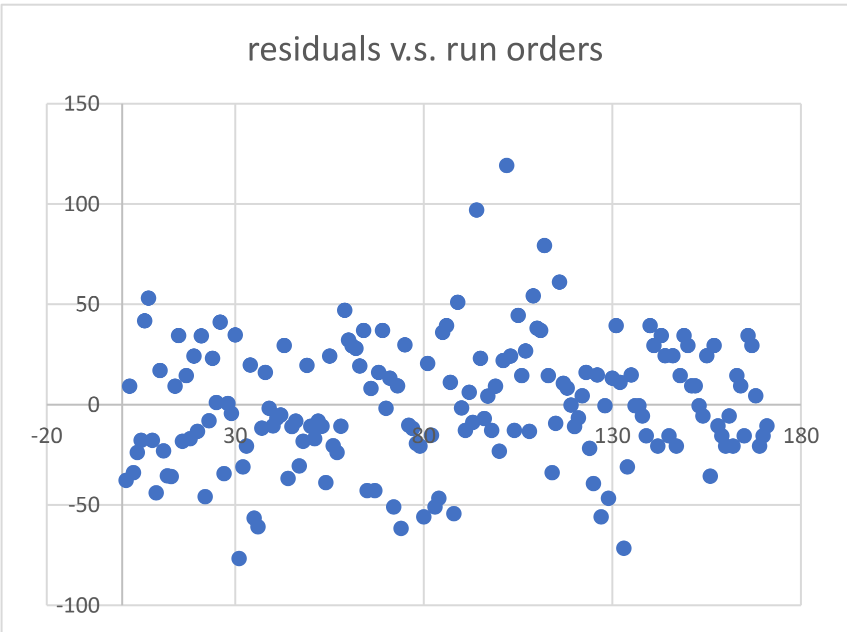
$$\sigma_e = \sum_{i}^{19}\sum_{j}^{9}\frac{\left(e_{ij}-\bar{e}\right)^2}{171-1} = 33.69912197$$

- Rank each residual from 1~171
- The c.d.f. of $i\ th$ effect is: $\hat{p}_i = (i-0.5)/171$
- The normal distribution value of $i\ th$ effect: $N^{-1}(\hat{p}_i;\bar{e},\sigma_e)$

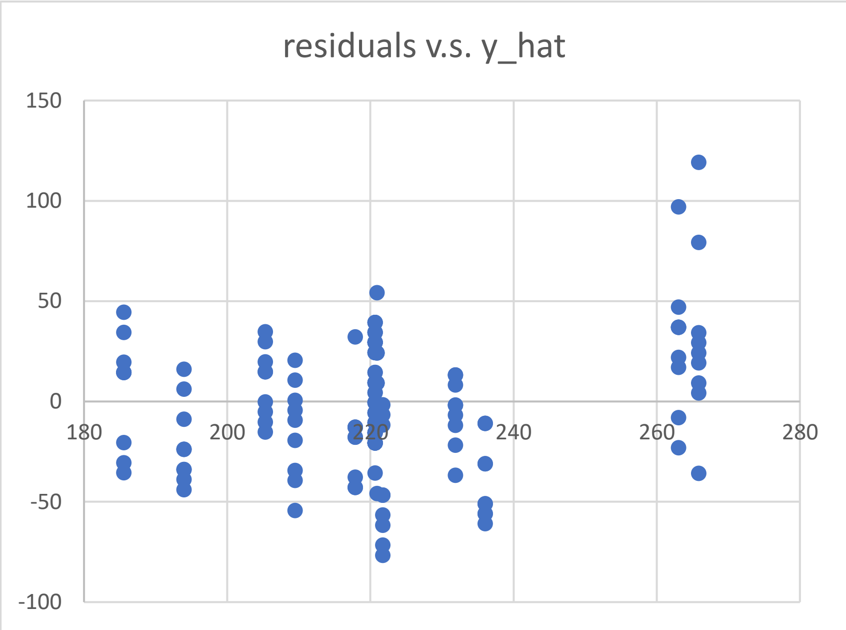
Residual Q-Q plot:



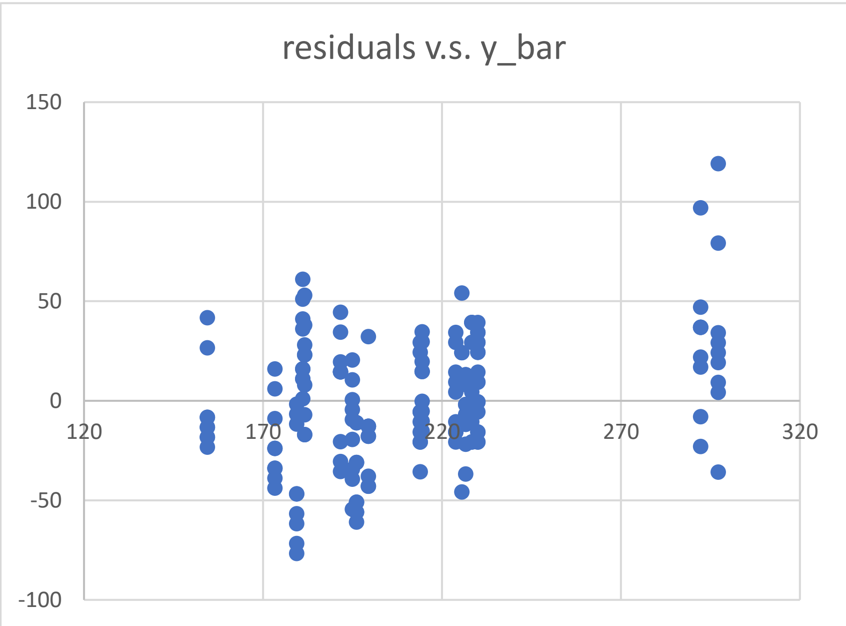
Plot the residuals vs. run orders:



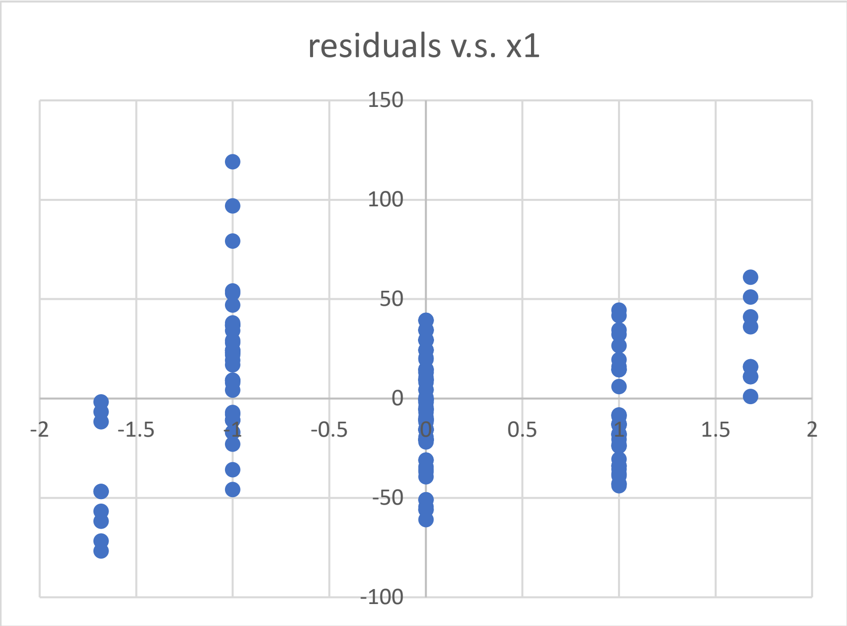
Plot the residuals vs. $\hat{y}$:



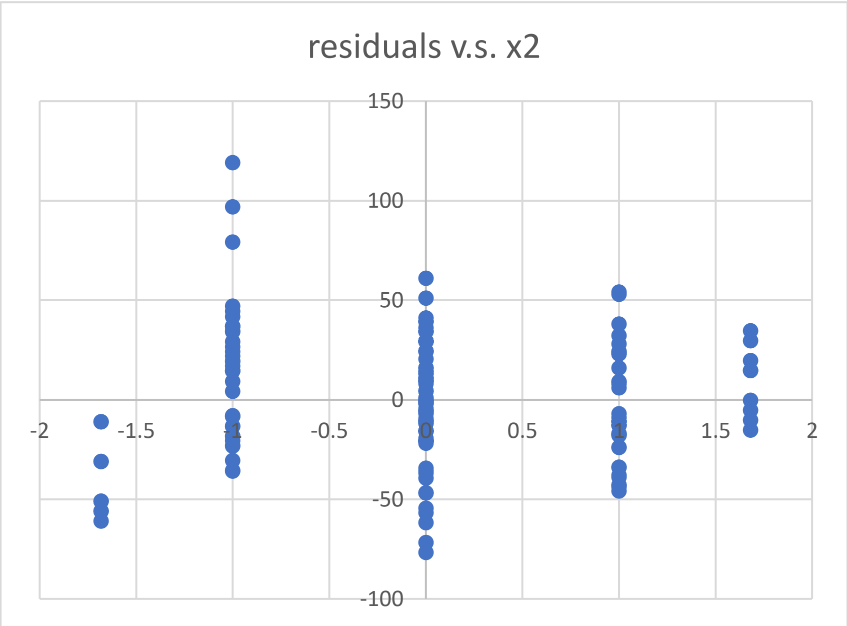
Plot the residuals vs. $\bar{y}$:



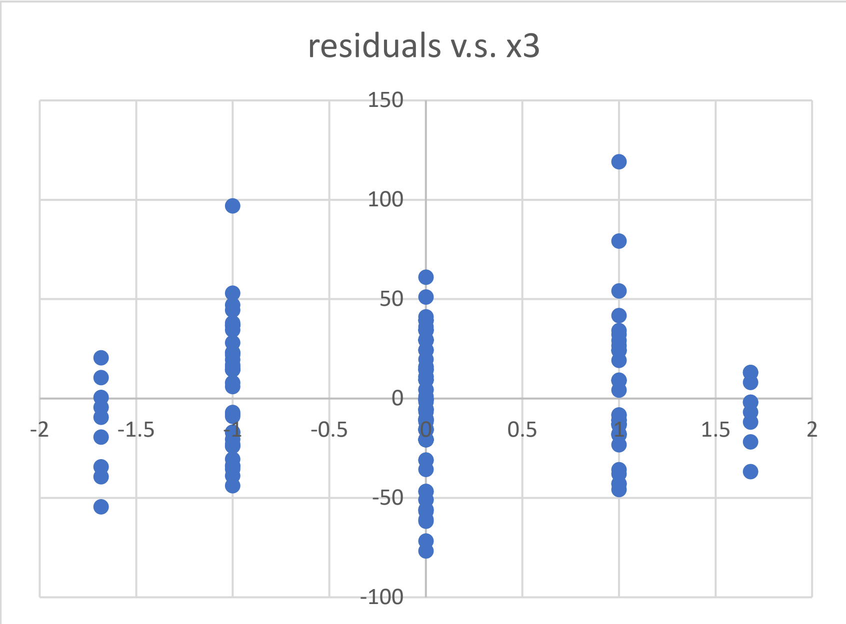
Plot the residuals vs. x1:



Plot the residuals vs. x2:



Plot the residuals vs. x3:



Problem observation:

- In the residual plots, variability of the residuals tends to increase as the predicted values ($\hat{y}$) or the actual average values ($\bar{y}$) rise, with this effect being particularly pronounced for the two highest values of $\hat{y}$ and $\bar{y}$.
- In the “residual vs. x1”, “residual vs. x2”, and “residual vs. x3” plots, the residuals show greater variability around the center of the graphs compared to the two sides.

17.

SN ratio: Larger-the-best

$$\eta = -10\log_{10}\left[\frac{1}{n}\sum_{i=1}^{n}\left(\frac{1}{X_i}\right)^2\right]$$

				replicate 1			replicate 2			replicate 3			
test	x1	x2	x3	run1	run2	run3	run1	run2	run3	run1	run2	run3	SN
1	-1	-1	-1	255	300	285	240	310	300	280	300	360	49.15775
2	1	-1	-1	200	205	230	220	155	200	150	165	200	45.37338
3	-1	1	-1	150	175	160	190	150	190	220	195	205	44.96722
4	1	1	-1	170	210	160	160	170	185	150	155	200	44.62473
5	-1	-1	1	300	290	345	230	285	270	275	295	385	49.21638
6	1	-1	1	140	150	145	200	150	185	145	140	135	43.57489
7	-1	1	1	175	210	230	230	210	245	245	210	275	46.87
8	1	1	1	180	175	205	200	250	205	200	175	205	45.85851
9	-1.682	0	0	210	220	215	145	160	150	165	175	175	44.79915
10	1.682	0	0	195	205	215	170	190	165	155	165	170	45.01357

11	0	-1.682	0	205	185	180	175	185	225	225	180	205	45.74185
12	0	1.682	0	200	190	205	225	195	220	240	235	220	46.5465
13	0	0	-1.682	205	190	220	210	230	200	175	155	170	45.60465
14	0	0	1.682	195	220	210	230	245	245	225	230	240	47.04125
15	0	0	0	250	230	260	210	260	225	200	200	220	47.0485
15--1	0	0	0	220	220	215	205	260	250	200	255	245	47.12242
15--2	0	0	0	205	245	200	235	255	250	230	230	220	47.15169
15--3	0	0	0	215	245	185	250	210	205	200	215	200	46.49718
15--4	0	0	0	235	230	205	255	250	225	200	205	210	46.90821

Check the correlation coefficient between each factor (including quadratic terms):

	x1	x2	x3	x1x2	x2x3	x1x3	x1x2x3	x1^2	x2^2	x3^2	SN
x1	1										
x2	0	1									
x3	0	0	1								
x1x2	0	0	0	1							
x2x3	0	0	0	0	1						
x1x3	0	0	0	0	0	1					
x1x2x3	0	0	0	0	0	0	1				
x1^2	0	0	0	0	0	0	0	1			
x2^2	0	0	0	0	0	0	0	-0.12814	1		
x3^2	0	0	0	0	0	0	0	-0.12814	-0.12814	1	
SN	-0.45901	-0.16073	0.16798	0.464637	0.280698	-0.14541	0.068391	-0.35546	-0.05274	-0.00902	1

The result shows that there is little correlation between each factor, therefore, it can be considered as no multicollinearity effect among this regression models:

$$\hat{\eta} = b_0 + b_1x_1 + b_2x_2 + b_3x_3 + b_{12}x_1x_2 + b_{23}x_2x_3 + b_{13}x_1x_3 + b_{123}x_1x_2x_3 + b_{1^2}x_1^2 + b_{2^2}x_2^2 + b_{3^2}x_3^2$$

Regression analysis results:

摘要輸出

迴歸統計	
R 的倍數	0.852273
R 平方	0.726369
調整的 R 平方	0.38433

標準誤	1.135935
觀察值個數	19

ANOVA

	自由度	SS	MS	F	顯著值
迴歸	10	27.40242	2.740242	2.123645	0.149064
殘差	8	10.32279	1.290349		
總和	18	37.72521			

	係數	標準誤	t 統計	P-值	下限 95%	上限 95%
截距	46.92693	0.507413	92.48262	2.09E-13	45.75683	48.09703
x1	-0.76285	0.307366	-2.48189	0.037999	-1.47164	-0.05406
x2	-0.26713	0.307366	-0.86909	0.410114	-0.97592	0.44166
x3	0.279175	0.307366	0.908282	0.39027	-0.42961	0.987963
x1x2	1.008985	0.401614	2.512328	0.036239	0.082862	1.935108
x2x3	0.609552	0.401614	1.517756	0.167554	-0.31657	1.535675
x1x3	-0.31577	0.401614	-0.78624	0.454381	-1.24189	0.610357
x1x2x3	0.148515	0.401614	0.369796	0.721132	-0.77761	1.074638
x1^2	-0.61766	0.307402	-2.00931	0.079363	-1.32653	0.091205
x2^2	-0.18014	0.307402	-0.586	0.574033	-0.88901	0.528733
x3^2	-0.11695	0.307402	-0.38044	0.713525	-0.82582	0.591923

With $\alpha = 0.05$, assuming all main effects, interaction effects, and quadratic effects are null. The table shows that only E_1 and E_{12} reject H_0 and could be considered as significant effects. Look closely on the ANOVA table, the p-value of F-test is 0.149064, which is relatively high. This suggests that the mean square (MS) between tests is similar to the MS of error and the difference between tests could be caused by noises and therefore can't explained by the regression model. In this analysis, R^2 is 0.726369, indicating that approximately 72.64% of the variation in jumping distance is explained by the model. However, adjusted R^2 is significantly lower at 0.38433. The low adjusted R^2 suggests that the model may overestimate its explanatory power and may not be able to capture the variation of jumping distance SN ratio.

18.

Base on the previous regression analysis, the final regression model to predict the SN ratio of jumping distance is:

$$\hat{\eta} = d_0 + d_1 x_1 + d_{12} x_1 x_2$$

The final regression analysis:

摘要輸出

迴歸統計	
R 的倍數	0.653128
R 平方	0.426577
調整的 R 平方	0.354899
標準誤	1.162769
觀察值個數	19

ANOVA

	自由度	SS	MS	F	顯著值
迴歸	2	16.0927	8.046349	5.9513	0.01169
殘差	16	21.63251	1.352032		
總和	18	37.72521			

	係數	標準誤	t 統計	P-值	下限 95%	上限 95%
截距	46.26936	0.266758	173.451	1.25E-27	45.70386	46.83486
x1	-0.76285	0.314627	-2.42462	0.027535	-1.42983	-0.09587
x1x2	1.008985	0.411101	2.454349	0.025946	0.13749	1.88048

First, this table shows that all terms in the final regression model are significant. Look closely on the ANOVA table, the p-value of F-test is 0.01169. It is considerably lower than the previous analysis. This indicates that current model has higher explanatory power on variation between tests. However, the R^2 and adjusted R^2 are 0.426577 and 0.354899, respectively, meaning that the regression model can only explain part of variability of jumping distance SN ratio, and remains over half of variation unexplained.

Compare to the predictive model construct in (9):

Regression model in (9):

$$\hat{\eta} = b_0 + b_1 x_1 = 46.20535766 - 1.347481576 x_1$$

Regression model in (18):

$$\hat{\eta} = d_0 + d_1 x_1 + d_{12} x_1 x_2$$
$$= 46.26935882 - 0.762850192 x_1 + 1.008985324 x_1 x_2$$

- Similarity:
  - ✓ Both models identify x_1(length of papers) as significant effect. This

suggests that the effect of x_1 are crucial in determining the jumping distance of the paper frogs.

- Difference:
  - ✓ The previous model only considers the direct linear impact of x_1 on SN ratio of paper frogs' jumping distance. On the other hand, the current model incorporates the interaction effect between x_1 and x_2 into consideration, providing a deeper understanding of how the combined influence of paper length and width affects the outcome.
  - ✓ In the predictive model in (9), x_1 has a stronger negative effect on SN ratio with a coefficient of −1.347481576. However, the main effect in the current model is less negative ($d_1 = -0.762850192$), and has a interaction effect with x_2. This suggesting that the impact of x_1 on SN ratio is depended on x_2.

19.

There are two objective functions:

- For the paper frog's jumping distance:

$$\begin{aligned}\max \hat{y}(x_1, x_2, x_3) &= 220.6763865 - 20.13248291x_1 - 9.131397298x_2 \\ &+ 6.665650756x_3 + 26.11111111x_1x_2 + 12.77777778x_2x_3 \\ &- 7.5x_1x_3 - 11.59789296x_1^2\end{aligned}$$

$$subjected\ to \begin{cases} -1 \le x_1 \le 1 \\ -1 \le x_2 \le 1 \\ -1 \le x_3 \le 1 \end{cases}$$

- For the SN ratio of paper frog's jumping distance

$$\max \hat{\eta}(x_1, x_2) = 46.26935882 - 0.762850192x_1 + 1.008985324x_1x_2$$

$$subjected\ to \begin{cases} -1 \le x_1 \le 1 \\ -1 \le x_2 \le 1 \end{cases}$$

In the first objective function, an optimum $\hat{y}$ can be find when $(x_1, x_2, x_3) = (-1, -1, 1)$:

$$\hat{y}(-1, -1, 1) = 265.8413578$$

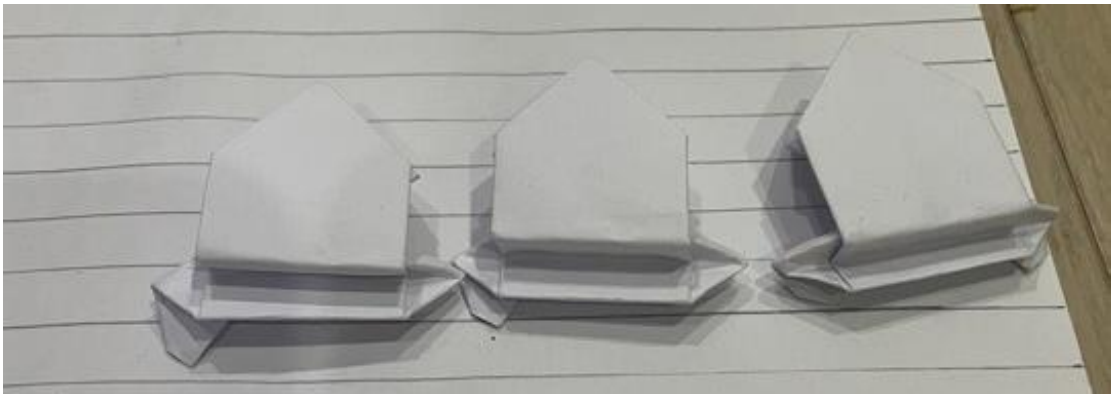
In the second objective function, an optimum $\hat{\eta}$ can be find when $(x_1, x_2) = (-1, -1)$:

$$\hat{\eta}(-1, -1) = 48.04119433$$

As a result, the both predictive models in (15) and (18) suggests that the paper frog will have a best jumping performance ($\text{distance} = 265.8413578mm$) when $(x_1, x_2, x_3) = (-1, -1, 1)$.

20.

Make three optimum frogs and measure their jumping distance:



verify			
	replication1	replication2	replication3
run1 distance(mm)	290	265	310
run2 distance(mm)	260	275	270
run3 distance(mm)	300	295	280
average	282.777778	SN	48.9876096

The average jumping distance of paper frogs in this verification is relatively larger than the most of the results in the previous experiment. This indicates that the regression model accurately captures the underlying trend of the paper frogs' jumping performance. Nonetheless, the average distance result in this verification (282.777778mm) is slightly higher than the predictive jumping distance (265.8413578mm) in the regression model built in (15), showing that this predictive model may underestimate the performance of reality paper frog. In addition, the SN ratio of jumping distance in the verification is quite large comparing to the previous experiment results, suggesting that the regression model have the ability to predict the tendency of paper frogs' jumping distance SN ratio. Moreover, the SN value in the verification is 48.9876096, and the maximize SN ratio in the model is 48.04119433. Although there is a slightly difference, this shows that the predictive model indeed reflects the SN ratio in the reality.

Compare the result in (20) to the result in (11):

- The regression models of jumping distance of paper frogs in (11) and (20) shows the different optimum setting. The current one predicts that paper frogs will have the best performance (with distance =265.8413578mm) when $(x_1, x_2, x_3) = (-1, -1, 1)$. While the previous one predicts that when $(x_1, x_2, x_3) = (-1, -1, -1)$, paper frog will jump farthest (distance=300mm). Both regression models show the trend that shorter paper length and width will have a positive impact on frogs' jumping performance. But they have different point of views on the effect of frogs' leg length.

In addition, the predictive model of SN ratio in (11) only considers the effect of paper length. In contrast, the current regression model of SN ratio considers not only the main effect of x_1 but also the interaction effect between paper length and width ratio. However, both regression models exclude the impact of the length of paper frogs' leg.

This problem may due to the too small difference of the experiment parameters. provided the variation of the leg length and the width of the paper could be enlarged, this problem will be likely to be solved.

- Although both models underestimate the result in the verifications, the current model of jumping distance tends to predicts a lower optimum comparing to the model in (11).