Midterm Exam R12522615 王邑安 ID:2

1.

i.

Assumption:

- IDD assumption in each term: Every winning set has identical probability to be chosen independently each term.
- IDD assumption in choosing a number each time: Every number has the identical probability to be chosen independently each time.

There are totally C_6^{49} outcomes in Big Lotto. To win the fourth prize, four of six numbers should be the winning numbers in the first set, and one of the rest numbers should be the special numbers. The probability of fourth prize will be:

$$p_0 = \frac{C_4^6 \times C_1^1 \times C_1^{42}}{C_6^{49}} = 4.50521E - 05$$

ii.

In maximum likelihood estimation, $\hat{p}_{MLE}$ will maximize $P(X_1 = x_1, X_2 = x_2, \cdots, X_{50} = x_{50})$.

$$P(X_1=x_1, X_2=x_2, \cdots X_{50}=x_{50}) = \prod_{i=1}^{50} P(X_i=x_i) \quad \Big| \text{ Binomial}: P(X=x)=\binom{n}{x}p^{x}(1-p)^{n-x}$$

$$\Rightarrow \prod_{i=1}^{50} P(X_i=x_i) = \prod_{i=1}^{50} C_{x_i}^{n_i} p^{x_i}(1-p)^{n_i-x_i} \Rightarrow f(p) = \ln\left[\prod_{i=1}^{50} C_{x_i}^{n_i} p^{x_i}(1-p)^{n_i-x_i}\right]$$

$$\Rightarrow f(p) = \sum_{i=1}^{50} \ln C_{x_i}^{n_i} + \left(\sum_{i=1}^{50} x_i\right)\cdot \ln p + \left[\sum_{i=1}^{50}(n_i-x_i)\right]\cdot \ln(1-p)$$

$$\Rightarrow \frac{df}{dp} = 0 + \sum_{i=1}^{50} x_i \times \frac{1}{p} + \sum_{i=1}^{50}(n_i-x_i)\times\frac{(-1)}{1-p} = 0$$

$$\Rightarrow \frac{\sum_{i=1}^{50} x_i(1-p) - \sum_{i=1}^{50}(n_i-x_i)p}{p(1-p)} = 0 \Rightarrow \sum_{i=1}^{50} x_i - \cancel{\sum_{i=1}^{50} x_i p} - \sum_{i=1}^{50} n_i p + \cancel{\sum_{i=1}^{50} x_i p} = 0$$

$$\Rightarrow \hat{p}_{MLE} = \frac{\sum_{i=1}^{50} x_i}{\sum_{i=1}^{50} n_i} \quad \#$$

$$E(\hat{p}_{MLE}) = E\left(\frac{\sum_{i=1}^{50} x_i}{\sum_{i=1}^{50} n_i}\right) = \frac{1}{\sum_{i=1}^{50} n_i}\cdot E\left(\sum_{i=1}^{50} X_i\right) = \frac{1}{\sum_{i=1}^{50} n_i}\cdot \sum_{i=1}^{50} E(X_i),\ E(X_i) = n_i p$$

$$\Rightarrow E(\hat{p}_{MLE}) = \frac{1}{\sum_{i=1}^{n} n_i}\cdot\sum_{i=1}^{n} n_i \cdot p = p \Rightarrow E(\hat{p}_{MLE}) - p = 0 \quad \therefore \hat{p}_{MLE} \text{ no biases}$$

By the derivation above, $\hat{p}_{MLE}$ could be written as the total fourth prize number $\sum_{i=1}^{50} x_i$ over the period divided by total trials over the period $\sum_{i=1}^{50} n_i$. Also, $\hat{p}_{MLE}$ could be proven that it has no biases by finding the expected value of

$\hat{p}_{MLE}$. Once $\hat{p}_{MLE}$ was founded, the standard error of it could be estimated by the following formula.

$$Var(\hat{p}_{MLE}) = Var\left(\frac{\sum_{i=1}^{50} x_i}{\sum_{i=1}^{50} n_i}\right) = \left(\frac{1}{\sum_{i=1}^{50} n_i}\right)^2 \cdot Var\left(\sum_{i=1}^{50} x_i\right) = \left(\frac{1}{\sum_{i=1}^{50} n_i}\right)^2 \cdot \sum_{i=1}^{50} Var(X_i),\ Var(X_i) = n_i p(1-p)$$

$$\Rightarrow Var(\hat{P}_{MLE}) = \left(\frac{1}{\sum_{i=1}^{50} n_i}\right)^2 \cdot \sum_{i=1}^{50} n_i p(1-p) = \frac{p(1-p)}{\sum_{i=1}^{50} n_i} \Rightarrow \sigma = \sqrt{Var} = \sqrt{\frac{p(1-p)}{\sum_{i=1}^{50} n_i}} \ \#$$

開獎日	注數	4 獎中獎數	開獎日	注數	4 獎中獎數
4 月 23 日	1890662	78	2 月 16 日	3766045	151
4 月 19 日	1825215	83	2 月 15 日	3946720	187
4 月 16 日	1875931	80	2 月 14 日	5440029	233
4 月 12 日	1898399	89	2 月 13 日	3807036	130
4 月 9 日	2217157	139	2 月 12 日	4978061	190
4 月 9 日	2023176	76	2 月 11 日	5421670	178
4 月 2 日	2044755	77	2 月 10 日	6228404	344
3 月 29 日	1962803	93	2 月 9 日	5782611	271
3 月 26 日	1973645	87	2 月 8 日	3673079	158
3 月 22 日	2018177	90	2 月 7 日	2942186	109
3 月 19 日	2057045	111	2 月 6 日	4308773	216
3 月 15 日	2052515	100	2 月 2 日	2179897	81
3 月 12 日	2945834	100	1 月 30 日	2135837	78
3 月 8 日	2651765	100	1 月 26 日	2048380	69
3 月 5 日	2693112	143	1 月 23 日	1859946	105
3 月 1 日	2397648	89	1 月 19 日	2058135	72
2 月 27 日	2285357	107	1 月 16 日	2091740	90
2 月 24 日	2814840	150	1 月 12 日	2094995	91
2 月 23 日	3095869	153	1 月 9 日	2088048	69
2 月 22 日	2739793	127	1 月 5 日	2025467	81
2 月 21 日	2772715	119	1 月 2 日	1883075	90
2 月 20 日	3122054	146	12 月 29 日	3156543	156
2 月 19 日	2963152	133	12 月 26 日	1956325	129
2 月 18 日	3034165	120	12 月 22 日	1848658	61
2 月 17 日	3296659	148	12 月 19 日	1924018	122

$$\sum_{i=1}^{50} x_i = 6199\,, \sum_{\mathrm{i}=1}^{50} \mathrm{n_i} = 140298121 \Rightarrow \hat{p}_{MLE} = 4.41845E-05$$

$$\sigma_{\hat{p}_{MLE}} = 5.61176E-07$$

The estimated $\hat{p}_{MLE}$ is slightly smaller than p_0. If the sample size is bigger, the difference between $\hat{p}_{MLE}$ and p_0 may become smaller.

iii.

Assume that the probability of winning a fourth prize follow the binomial distribution. To do the Hypothesis test with $\alpha = 0.1$, the lower bound could be constructed as the inverse of binomial distribution with $n = n_i$ and $c.d.f. = 0.05$. The upper bound could be constructed as the inverse of binomial distribution with $n = n_i$ and $c.d.f. = 0.95$. The result will reject H_0 if the numbers of winning fourth prize exceed the boundaries.

$$H_0: p = p_0 = 4.50521E - 05$$
$$H_1: \neq p_0 \neq 4.50521E - 05$$
$$lowerbound: B_{(0.05, n_i)}, upperbound: B_{(0.95, n_i)}$$

開獎日	注數	4 獎中獎數	p_0	Binom_0.05(lb)	Binom_0.95(ub)	H_0	p value
4 月 23 日	1890662	78	4.50521E-05	70	101	accept	0.474568822
4 月 19 日	1825215	83	4.50521E-05	68	97	accept	0.874390998
4 月 16 日	1875931	80	4.50521E-05	70	100	accept	0.673077487
4 月 12 日	1898399	89	4.50521E-05	71	101	accept	0.656843332
4 月 9 日	2217157	139	4.50521E-05	84	117	reject	0.000174714
4 月 9 日	2023176	76	4.50521E-05	76	107	accept	0.118647814
4 月 2 日	2044755	77	4.50521E-05	77	108	accept	0.12146387
3 月 29 日	1962803	93	4.50521E-05	73	104	accept	0.581028259
3 月 26 日	1973645	87	4.50521E-05	74	105	accept	0.894262818
3 月 22 日	2018177	90	4.50521E-05	76	107	accept	0.978547348
3 月 19 日	2057045	111	4.50521E-05	77	109	reject	0.055972509
3 月 15 日	2052515	100	4.50521E-05	77	109	accept	0.400624037
3 月 12 日	2945834	100	4.50521E-05	114	152	reject	0.003639961
3 月 8 日	2651765	100	4.50521E-05	102	138	reject	0.076904438
3 月 5 日	2693112	143	4.50521E-05	104	140	reject	0.048793296
3 月 1 日	2397648	89	4.50521E-05	91	125	reject	0.068719223
2 月 27 日	2285357	107	4.50521E-05	87	120	accept	0.64514132
2 月 24 日	2814840	150	4.50521E-05	109	146	reject	0.039697948
2 月 23 日	3095869	153	4.50521E-05	120	159	accept	0.23707133
2 月 22 日	2739793	127	4.50521E-05	105	142	accept	0.704695596
2 月 21 日	2772715	119	4.50521E-05	107	144	accept	0.636086578
2 月 20 日	3122054	146	4.50521E-05	121	160	accept	0.614639399
2 月 19 日	2963152	133	4.50521E-05	115	153	accept	0.988216535
2 月 18 日	3034165	120	4.50521E-05	118	156	accept	0.161670735

2月17日	3296659	148	4.50521E-05	129	169	accept	0.99047783
2月16日	3766045	151	4.50521E-05	149	191	accept	0.159173533
2月15日	3946720	187	4.50521E-05	156	200	accept	0.463682693
2月14日	5440029	233	4.50521E-05	220	271	accept	0.462209083
2月13日	3807036	130	4.50521E-05	150	193	reject	0.001123036
2月12日	4978061	190	4.50521E-05	200	249	reject	0.021210908
2月11日	5421670	178	4.50521E-05	219	270	reject	1.03308E-05
2月10日	6228404	344	4.50521E-05	253	308	reject	0.000222804
2月9日	5782611	271	4.50521E-05	234	287	accept	0.492755929
2月8日	3673079	158	4.50521E-05	145	187	accept	0.593736346
2月7日	2942186	109	4.50521E-05	114	152	reject	0.040364854
2月6日	4308773	216	4.50521E-05	171	217	accept	0.112135238
2月2日	2179897	81	4.50521E-05	82	115	reject	0.085379239
1月30日	2135837	78	4.50521E-05	80	113	reject	0.064385092
1月26日	2048380	69	4.50521E-05	77	108	reject	0.013783809
1月23日	1859946	105	4.50521E-05	69	99	reject	0.021672162
1月19日	2058135	72	4.50521E-05	77	109	reject	0.030277027
1月16日	2091740	90	4.50521E-05	79	110	accept	0.71118009
1月12日	2094995	91	4.50521E-05	79	111	accept	0.778620616
1月9日	2088048	69	4.50521E-05	78	110	reject	0.008297825
1月5日	2025467	81	4.50521E-05	76	107	accept	0.306737318
1月2日	1883075	90	4.50521E-05	70	100	accept	0.531204805
12月29日	3156543	156	4.50521E-05	123	162	accept	0.232900504
12月26日	1956325	129	4.50521E-05	73	104	reject	3.59799E-05
12月22日	1848658	61	4.50521E-05	69	99	reject	0.012938894
12月19日	1924018	122	4.50521E-05	72	102	reject	0.00027628

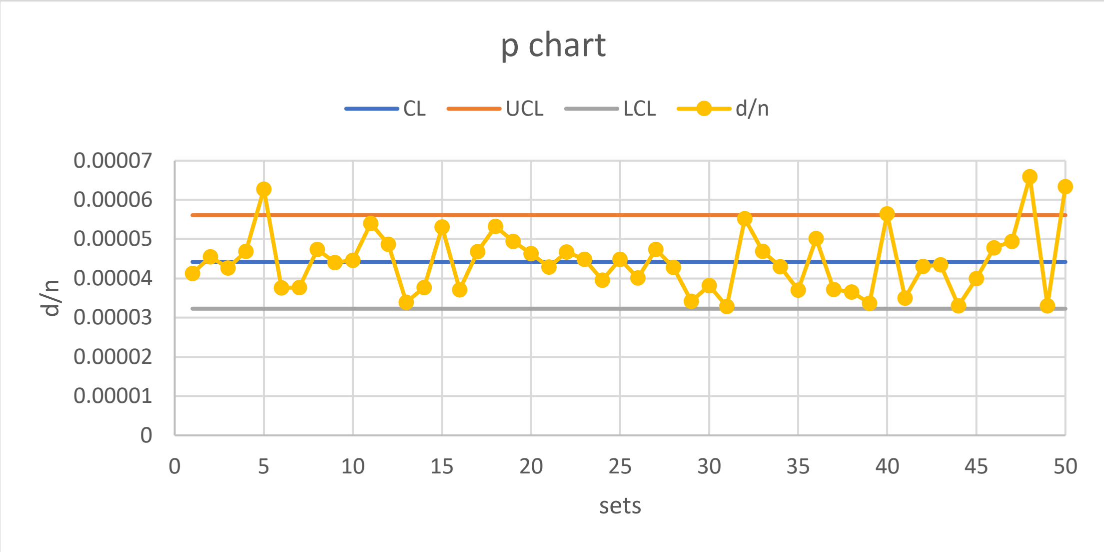
iv.

d_i is the number of winning fourth prize in a draw. n_i is the total trials in each draw. The following parameters can be derived as:

$$\hat{p} = \frac{\sum_{i=1}^{50} d_i}{\sum_{i=1}^{50} n_i}, \bar{n} = \frac{\sum_{i=1}^{50} n_i}{50}, \sigma_{\bar{n}} = \sqrt{\frac{\hat{p}(1-\hat{p})}{n_i}}$$

$$CL = \hat{p}\,, UCL = \hat{p} + 3 \times \sigma_{\bar{n}}\,, LCL = \hat{p} - 3 \times \sigma_{\bar{n}}$$

$\hat{p}$	$\bar{n}$	$\sigma_{\bar{n}}$	CL	UCL	LCL
4.41845E-05	2805962.42	3.96812E-06	4.42E-05	5.61E-05	3.23E-05



	$1 - B(n \times UCL; \bar{n}, \hat{p})$	$B(n \times LCL; \bar{n}, \hat{p})$
Binomial.dist	0.001850711	0.000831981

$$\alpha = 0.001850711 + 0.000831981 = 0.002682692$$

$$\Rightarrow ARL_0 = 1/\alpha = 372.7598392$$

v.

In the beginning, there has been the special number in a set. As a result, there are only five numbers we can choose. The outcome space is C_5^{48}. Four number should be the winning numbers. The rest of it should not be any winning number. The probability in this situation is:

$$p_1 = \frac{C_4^6 \times C_1^{42}}{C_6^{48}} = 0.000367925$$

vi.

The probability of type II error is $P_1(LCL \leq X \leq UCL)$. n is the trial number in each draw. p_1 is the shifted probability. ub and lb is the upper bound and lower bound constructed in (iii) respectively. And the following parameter is:

$$P_1(x) = \sum_{i=0}^{x} b(i; n, p_1)$$

$$\beta = P_1(ub) - P_1(lb)$$

the result:

開獎日	注數	4 獎中獎數	p_0	Binom_0.05(lb)	Binom_0.95(ub)	p_1	P_new(lb)	P_new(ub)	β
4 月 23 日	1890662	78	4.50521E-05	70	101	0.0004	6.1E-204	1.07E-175	1.1E-175
4 月 19 日	1825215	83	4.50521E-05	68	97	0.0004	1.6E-196	4.21E-170	4.2E-170
4 月 16 日	1875931	80	4.50521E-05	70	100	0.0004	8E-202	1.6E-174	1.6E-174
4 月 12 日	1898399	89	4.50521E-05	71	101	0.0004	4.6E-204	9.34E-177	9.3E-177

4 月 9 日	2217157	139	4.50521E-05	84	117	0.0004	5.9E-237	6.28E-207	6.3E-207
4 月 9 日	2023176	76	4.50521E-05	76	107	0.0004	5E-217	8.66E-189	8.7E-189
4 月 2 日	2044755	77	4.50521E-05	77	108	0.0004	3.9E-219	6.68E-191	6.7E-191
3 月 29 日	1962803	93	4.50521E-05	73	104	0.0004	2.5E-211	4.77E-183	4.8E-183
3 月 26 日	1973645	87	4.50521E-05	74	105	0.0004	6.8E-212	1.08E-183	1.1E-183
3 月 22 日	2018177	90	4.50521E-05	76	107	0.0004	2.6E-216	4.19E-188	4.2E-188
3 月 19 日	2057045	111	4.50521E-05	77	109	0.0004	6.7E-221	9.64E-192	9.6E-192
3 月 15 日	2052515	100	4.50521E-05	77	109	0.0004	3E-220	4.02E-191	4E-191
3 月 12 日	2945834	100	4.50521E-05	114	152	0.0004	0	3.09E-277	3.1E-277
3 月 8 日	2651765	100	4.50521E-05	102	138	0.0004	1.6E-281	9.36E-249	9.4E-249
3 月 5 日	2693112	143	4.50521E-05	104	140	0.0004	1.7E-285	9.86E-253	9.9E-253
3 月 1 日	2397648	89	4.50521E-05	91	125	0.0004	6.2E-256	6.57E-225	6.6E-225
2 月 27 日	2285357	107	4.50521E-05	87	120	0.0004	8.9E-244	9.66E-214	9.7E-214
2 月 24 日	2814840	150	4.50521E-05	109	146	0.0004	5E-298	2.39E-264	2.4E-264
2 月 23 日	3095869	153	4.50521E-05	120	159	0.0004	0	6.86E-292	6.9E-292
2 月 22 日	2739793	127	4.50521E-05	105	142	0.0004	3.4E-291	1.92E-257	1.9E-257
2 月 21 日	2772715	119	4.50521E-05	107	144	0.0004	6E-294	2.9E-260	2.9E-260
2 月 20 日	3122054	146	4.50521E-05	121	160	0.0004	0	1.23E-294	1.2E-294
2 月 19 日	2963152	133	4.50521E-05	115	153	0.0004	0	9.17E-279	9.2E-279
2 月 18 日	3034165	120	4.50521E-05	118	156	0.0004	0	5.75E-286	5.7E-286
2 月 17 日	3296659	148	4.50521E-05	129	169	0.0004	0	0	0
2 月 16 日	3766045	151	4.50521E-05	149	191	0.0004	0	0	0
2 月 15 日	3946720	187	4.50521E-05	156	200	0.0004	0	0	0
2 月 14 日	5440029	233	4.50521E-05	220	271	0.0004	0	0	0
2 月 13 日	3807036	130	4.50521E-05	150	193	0.0004	0	0	0
2 月 12 日	4978061	190	4.50521E-05	200	249	0.0004	0	0	0
2 月 11 日	5421670	178	4.50521E-05	219	270	0.0004	0	0	0
2 月 10 日	6228404	344	4.50521E-05	253	308	0.0004	0	0	0
2 月 9 日	5782611	271	4.50521E-05	234	287	0.0004	0	0	0
2 月 8 日	3673079	158	4.50521E-05	145	187	0.0004	0	0	0
2 月 7 日	2942186	109	4.50521E-05	114	152	0.0004	0	9.81E-277	9.8E-277
2 月 6 日	4308773	216	4.50521E-05	171	217	0.0004	0	0	0
2 月 2 日	2179897	81	4.50521E-05	82	115	0.0004	1.4E-233	1.64E-203	1.6E-203
1 月 30 日	2135837	78	4.50521E-05	80	113	0.0004	3.1E-229	3.66E-199	3.7E-199
1 月 26 日	2048380	69	4.50521E-05	77	108	0.0004	1.2E-219	2.13E-191	2.1E-191
1 月 23 日	1859946	105	4.50521E-05	69	99	0.0004	1.6E-200	3.56E-173	3.6E-173
1 月 19 日	2058135	72	4.50521E-05	77	109	0.0004	4.7E-221	6.84E-192	6.8E-192

1月16日	2091740	90	4.50521E-05	79	110	0.0004	6.7E-224	1.19E-195	1.2E-195
1月12日	2094995	91	4.50521E-05	79	111	0.0004	2.3E-224	2.96E-195	3E-195
1月9日	2088048	69	4.50521E-05	78	110	0.0004	2.3E-224	3.81E-195	3.8E-195
1月5日	2025467	81	4.50521E-05	76	107	0.0004	2.4E-217	4.21E-189	4.2E-189
1月2日	1883075	90	4.50521E-05	70	100	0.0004	7.5E-203	1.69E-175	1.7E-175
12月29日	3156543	156	4.50521E-05	123	162	0.0004	0	1.13E-297	1.1E-297
12月26日	1956325	129	4.50521E-05	73	104	0.0004	2.1E-210	3.67E-182	3.7E-182
12月22日	1848658	61	4.50521E-05	69	99	0.0004	6.7E-199	1.24E-171	1.2E-171
12月19日	1924018	122	4.50521E-05	72	102	0.0004	9.5E-207	2.02E-179	2E-179

vii.

The probability of type II error is $P_1(LCL \le X \le UCL)$. n is the trial number in each draw. p_1 is the shifted probability. And the following parameter is:

$$P_1(x) = \sum_{i=0}^{x} b(i; n, p_1)$$

$$\beta = P_1(\bar{n} \cdot UCL) - P_1(\bar{n} \cdot LCL), ARL_1 = 1/(1-\beta)$$

the result:

p_1	$P_1(UCL)$	$P_1(LCL)$	β	ARL_1
0.000368	5.7171E-254	0	5.7E-254	1

2.

i.

A chi-squared proportion test with $\alpha = 0.1$ is constructed to estimate the accident death per month in Taoyuan. The number n is 84 because there are 84 months from Jan.2011 to Dec.2017. The maximum death in a month is 17 and the minimum death is 4. As a result, 16 kinds of death number (~3,4,…,17,18~) had been tested. The Poisson parameter λ is 10.

$$H_0: possion\ distribution\ with\ \lambda = 10$$
$$H_1: not\ possion\ distribution\ with\ \lambda = 10$$

n	max death number	min death number	k
84	17	4	16

death number	x_i	frequency	p_i(poisson λ=10)	np_i	$(X_i - np_i)^2/np_i$
~3	0	0	0.010336051	0.868228	0.868228257
4	1	0.011905	0.018916637	1.588998	0.218325136
5	3	0.035714	0.037833275	3.177995	0.009969257

6	9	0.107143	0.063055458	5.296658	2.58931901
7	9	0.107143	0.090079226	7.566655	0.271517337
8	6	0.071429	0.112599032	9.458319	1.26449199
9	10	0.119048	0.125110036	10.50924	0.024676224
10	13	0.154762	0.125110036	10.50924	0.590325148
11	8	0.095238	0.113736396	9.553857	0.252722262
12	9	0.107143	0.09478033	7.961548	0.13544893
13	5	0.059524	0.072907946	6.124267	0.206388336
14	5	0.059524	0.052077104	4.374477	0.089445967
15	4	0.047619	0.03471807	2.916318	0.402688275
16	1	0.011905	0.021698794	1.822699	0.371335698
17	1	0.011905	0.012763996	1.072176	0.004858652
18~	0	0	0.014277614	1.19932	1.199319542

$$\Rightarrow \mathrm{C}^2 = \sum \frac{(x_i - np_i)^2}{np_i} = 8.499060021, \mathcal{X}^2_{0.1,15} = 22.30712958$$

$$\infty < 93.27018043 \Rightarrow accept\ H_0$$

The result shows that the accident death per month in Taoyuan over the period followed a Poisson distribution with $\lambda = 10$.

ii.

Suppose the accident death per month in Taoyuan followed a Poisson distribution. $p(x, \lambda)$ is the probability of x in Poisson distribution. The upper bound and lower bound in the Hypothesis test with $\alpha = 0.05$ can be constructed by the following:

$$P(x; \lambda) = cdf[p(x, \lambda)]$$

$$P(x_1; 10) > 0.025, P(x_1 - 1; 10) < 0.025 \Rightarrow x_1 = lower\ bound = 4$$

$$P(x_2; 10) > 0.975, P(x_2 - 1; 10) < 0.975 \Rightarrow x_2 = upper\ bound = 17$$

Since Poisson is a discrete distribution, the upper bound and lower bound in the Hypothesis test with $\lambda = \lambda_{10}, \alpha = 0.05$ are not really accurate. I use the p-value to do the Hypothesis test. If the p-value is smaller than α =0.05, the death number in that month reject H0.

$$H_0: \lambda = \lambda_0 = 10$$

$$H_1: \lambda \neq \lambda_0$$

$$p - value\ in\ each\ x_i: \begin{cases} P(x_i; 10) \times 2 & , P(x_i; 10) \leq 0.5 \\ \left(1 - P(x_i; 10)\right) \times 2 & , P(x_i; 10) > 0.5 \end{cases}$$

month	death in Taoyuan	p-value	H_0	month	death in Taoyuan	p-value	H_0
1-Jan	7	0.4404413	accept	1-Mar	18	0.01437301	reject
1-Feb	7	0.4404413	accept	1-Apr	16	0.05408322	accept
1-Mar	5	0.1341719	accept	1-May	18	0.01437301	reject
1-Apr	4	0.0585054	accept	1-Jun	7	0.44044129	accept
1-May	9	0.9158594	accept	1-Jul	15	0.09748081	accept
1-Jun	7	0.4404413	accept	1-Aug	14	0.16691695	accept
1-Jul	10	0.8339205	accept	1-Sep	15	0.09748081	accept
1-Aug	6	0.2602828	accept	1-Oct	13	0.27107115	accept
1-Sep	14	0.1669169	accept	1-Nov	18	0.01437301	reject
1-Oct	6	0.2602828	accept	1-Dec	22	0.00059147	reject
1-Nov	13	0.2710712	accept	1-Jan	18	0.01437301	reject
1-Dec	5	0.1341719	accept	1-Feb	10	0.8339205	accept
1-Jan	21	0.0013993	reject	1-Mar	24	9.3899E-05	reject
1-Feb	13	0.2710712	accept	1-Apr	13	0.27107115	accept
1-Mar	13	0.2710712	accept	1-May	12	0.41688705	accept
1-Apr	11	0.6064477	accept	1-Jun	13	0.27107115	accept
1-May	22	0.0005915	reject	1-Jul	19	0.00690868	reject
1-Jun	17	0.0285552	reject	1-Aug	9	0.91585943	accept
1-Jul	9	0.9158594	accept	1-Sep	15	0.09748081	accept
1-Aug	16	0.0540832	accept	1-Oct	24	9.3899E-05	reject
1-Sep	22	0.0005915	reject	1-Nov	14	0.16691695	accept
1-Oct	16	0.0540832	accept	1-Dec	12	0.41688705	accept
1-Nov	11	0.6064477	accept	1-Jan	18	0.01437301	reject
1-Dec	17	0.0285552	reject	1-Feb	15	0.09748081	accept
1-Jan	21	0.0013993	reject	1-Mar	23	0.00024024	reject
1-Feb	24	9.39E-05	reject				

iii.

The upper bound in the Hypothesis test is 17, and the lower bound is 4. The type Ⅱ error probability β is $P(4 \leq X \leq 17)$ in Poisson distribution with $\lambda = 13$.

$$\beta = P(17; 13) - P(4; 13) = 0.886724794$$

iv.

The parameter c_i is the accident death in Taoyuan cities per month. The parameter $k = 84$ because there are totally 84 months from Jan.2011 to Dec.2017. The $\hat{\lambda}$ is estimated by the following equation:

$$\hat{\lambda} = \frac{\sum_{i=1}^{k} c_i}{k}$$

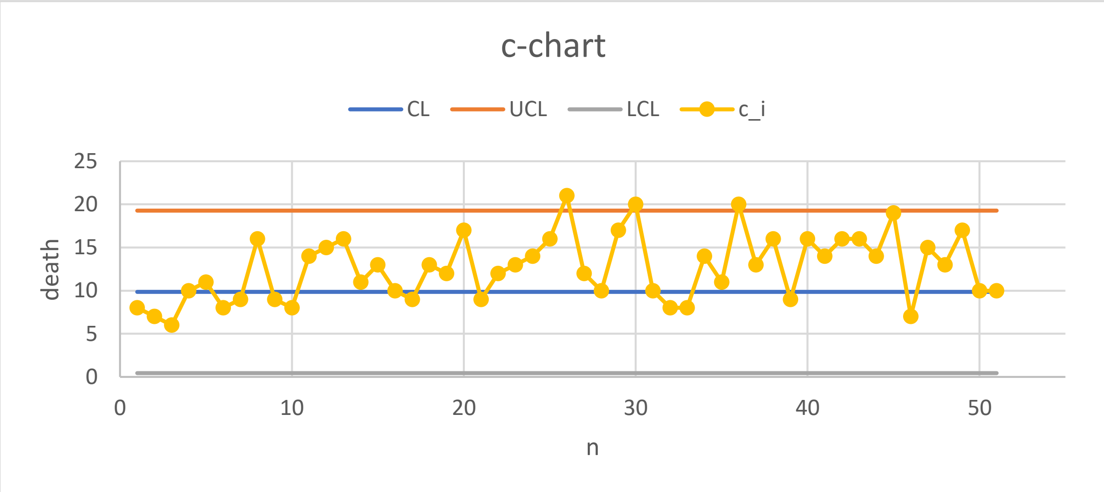
The standard error is the square root of $\hat{\lambda}$. Finally, CL, UCL, and LCL can be constructed:

$$\sigma = \sqrt{\hat{\lambda}}$$

$$CL = \hat{\lambda}\,, UCL = \hat{\lambda} + 3 \times \sigma, LCL = \hat{\lambda} - 3 \times \sigma$$

$\hat{\lambda}$	σ	CL	UCL	LCL
10	3.139609	9.857143	19.27597	0.438317

Use the central line and the two control limits to monitor the number of accident deaths from Jan.2018 to March 2022.



The probability of type Ⅰ error α is:

$$\alpha = P(LCL;\hat{\lambda}) + 1 - P(UCL;\hat{\lambda}), ARL_0 = 1/\alpha$$

where $p(x;\lambda)$ is the probability of x in Poisson with λ. And $P(x,\lambda) = c.d.f.[p(x;\lambda)]$.

$1 - P(UCL;\hat{\lambda})$	$P(LCL;\hat{\lambda})$	α	ARL_0
0.002954	5.24E-05	0.003007	332.5889

3.

i.

The average thicknesses and the standard error of right side and left side could be calculated by the 85-wafers sample in bottom zone.

average				
up	middle	down	left	right
349.7176	349.6471	356.1882	348.7412	357.0941

sample variance				
23.75266	15.15966	142.1308	138.1941	32.56246
standard error				
4.87367	3.893541	11.92186	11.7556	5.706353

The t-distribution could be used in this Hypothesis test. The DOF in t-distribution is $(85-1)\times 2$ because there are 85 numbers in a sample. The α is 0.02. Consequently, the result will reject H_0 if $t-test < t_{0.01,168}$ or $t-test > t_{0.99,168}$.

$$\alpha = 0.02\ , H_0: \mu_{right} = \mu_{left}\ , H_1: \mu_{right} \neq \mu_{left}$$

$$s_p = \sqrt{\frac{1}{2}s_{right}^2 + \frac{1}{2}s_{left}^2} = 9.240037409$$

$$\mathrm{t-test} = \frac{\mu_{right} - \mu_{left}}{s_p\sqrt{\frac{1}{84}+\frac{1}{84}}} = 5.85855267$$

$$t_{0.99,168} = 2.348748824\ , t_{0.01,168} = -2.348748824$$

$$t - test > t_{0.99,168} \Rightarrow reject\ H_0\ , p - value = 2.40101E-08$$

ii.

The proportion between two sample variances might be used in the Hypothesis test. The smaller on should be the denominator. F-test could be used to test whether the result reject H_0 or not. The result will accept H_0 if $F - test < F_{0.02,84,84}$.

$$\alpha = 0.02\ , H_0: \sigma_{middle}^2 = \sigma_{right}^2\ , H_1: \sigma_{middle}^2 < \sigma^2 right$$

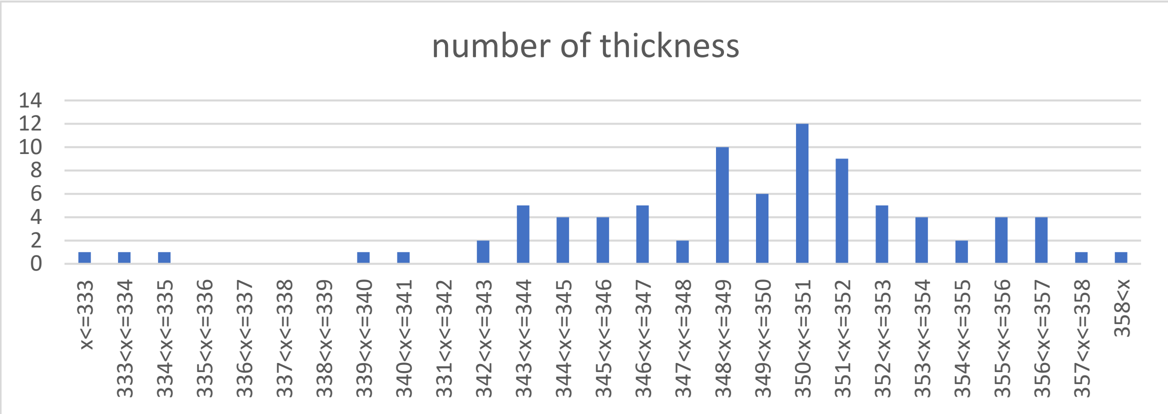
$$F - test = \frac{s_{right}^2}{s_{middle}^2} = 2.14796748$$

$$F_{0.02,84,84} = 1.570492538$$

$$F - test > F_{0.02,84,84} \Rightarrow reject\ H_0\ , p - value = 0.000278223$$

iii.

The maximum and minimum are considered as extreme values in those 85 numbers. Therefore, only the values between the second biggest number and the second smallest number are counted precisely in the histogram.



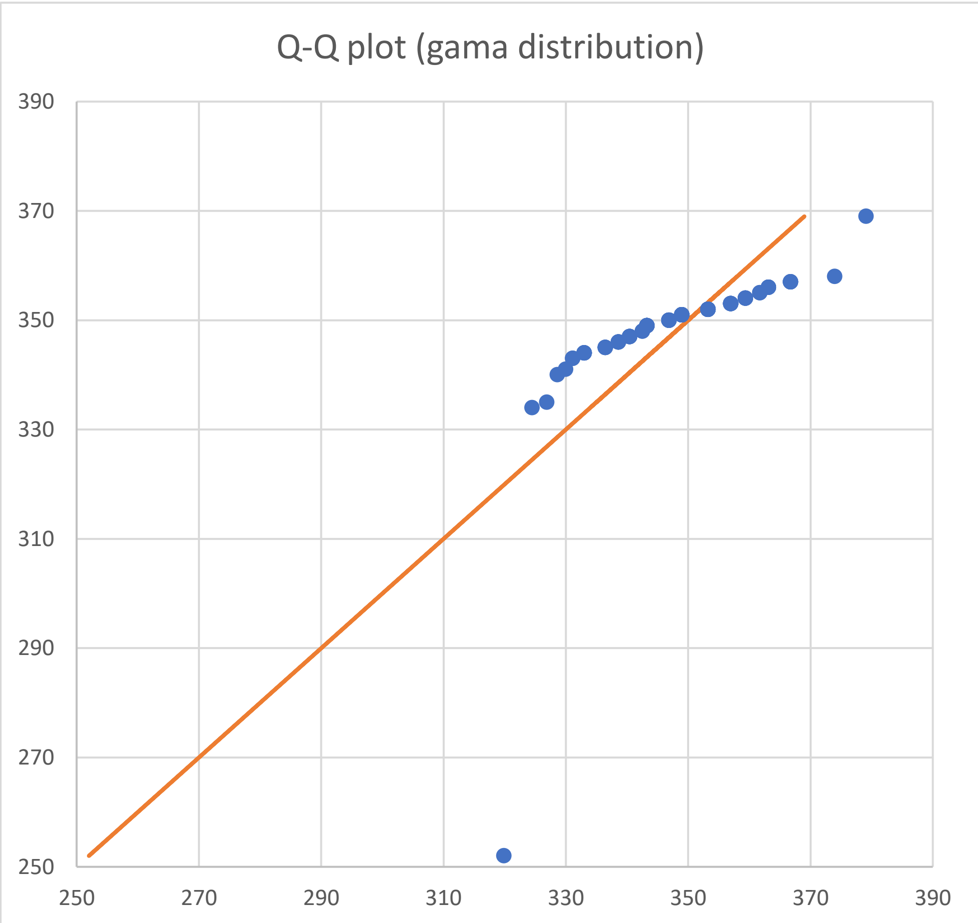
The data in the Q-Q plot has been rearranged from the smallest to the largest.

The {x , y} in Q-Q plot is $\{c.d.f.[(i-0.5)/n]\ ,\ i$th smallest sample observation}.

The Gamma distribution has been assumed in this data. The parameter of Gamma can be estimated by following:

$$by\ moment\ estimator: \alpha = \bar{X}^2/\sigma^2, \beta = \sigma^2/\bar{X}$$

average	sample variance	standard error	estimated α	estimated β
348.7412	138.1941	11.7556	880.0694	0.396266



To do the $\mathcal{X}^2$ proportion test, the appearance numbers of 27 kinds thickness value counted in the histogram must be compared with the probability of those value in Gamma distribution. The probability of those x_i is:

$$p_i(x_i) = P(x_i) - P(x_i - 1)$$

Where $P(x_i)$ is $c.d.f.[f(x_i;\alpha,\beta)]$, and $f(x_i;\alpha,\beta)$ follow Gamma distribution with α and β.

$$H_0\text{: Gamma distribution }, H_1\text{: not Gamma distribution}$$

thickness	number x_i	p_i(gamma.dist)	np_i	thickness	number x_i	p_i (gamma.dist)	np_i
x<=333	1	0.088757	7.544311	346<x<=347	5	0.033525	2.849623
333<x<=334	1	0.014932	1.269207	347<x<=348	2	0.033854	2.877624
334<x<=335	1	0.01664	1.414405	348<x<=349	10	0.033939	2.884836
335<x<=336	0	0.018399	1.563887	349<x<=350	6	0.033779	2.871222
336<x<=337	0	0.020185	1.715725	350<x<=351	12	0.033379	2.837191
337<x<=338	0	0.021974	1.867759	351<x<=352	9	0.032748	2.783587
338<x<=339	0	0.023737	2.017643	352<x<=353	5	0.031902	2.711644
339<x<=340	1	0.025446	2.162909	353<x<=354	4	0.030858	2.622949
340<x<=341	1	0.027071	2.301029	354<x<=355	2	0.02964	2.519379
331<x<=342	0	0.028582	2.429491	355<x<=356	4	0.028271	2.403041
342<x<=343	2	0.029951	2.545875	356<x<=357	4	0.026779	2.276195
343<x<=344	5	0.031152	2.647928	357<x<=358	1	0.02519	2.141192
344<x<=345	4	0.032161	2.733644	358<x	1	0.214193	18.20639
345<x<=346	4	0.032957	2.80132	k=27			

$$C^2 = \sum_{i=1}^{k} \frac{(x\_i - np\_i)^\wedge 2}{np\_i} = 108.437064$$

$$\mathcal{X}^2_{0.02,26} = 42.85583479$$

$$C^2 > \mathcal{X}^2_{0.02,26} \Rightarrow rejectH_0\ , p - value = 2.78251E - 12$$

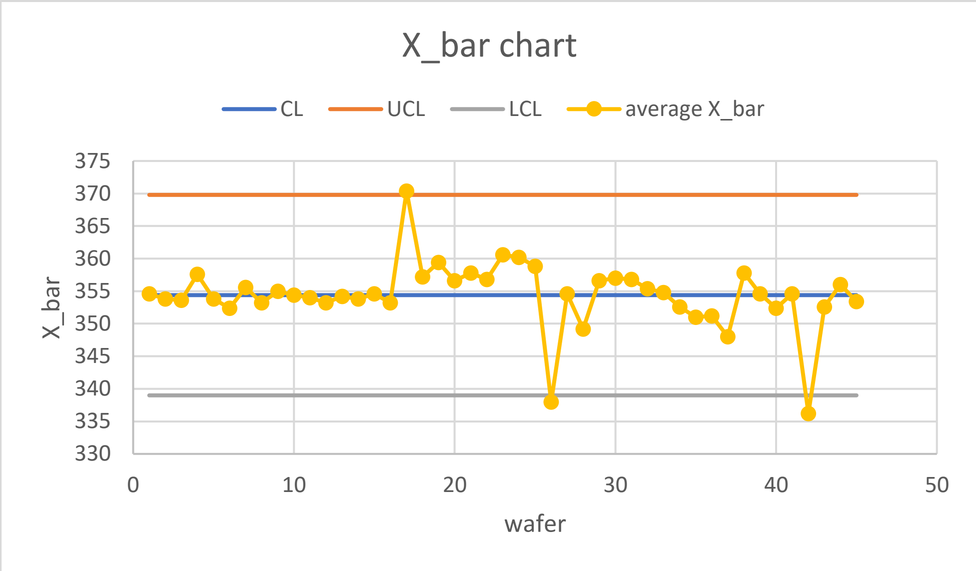
The result shows that the thickness wasn't follow the Gamma distribution.

iv.

Sample mean $\bar{\bar{X}}$ and standard error $s_{\bar{X}}$ are estimated from first 45 wafers. The control lines are constructed as following:

$$CL = \bar{\bar{X}}\ , UCL = \bar{\bar{X}} + 3\sigma\ , LCL = \bar{\bar{X}} - 3\sigma$$

$\bar{\bar{X}}$	sample variance	standard error σ	CL	UCL	LCL
354.3911	26.34174	5.13242	354.3911	369.7884	338.9939

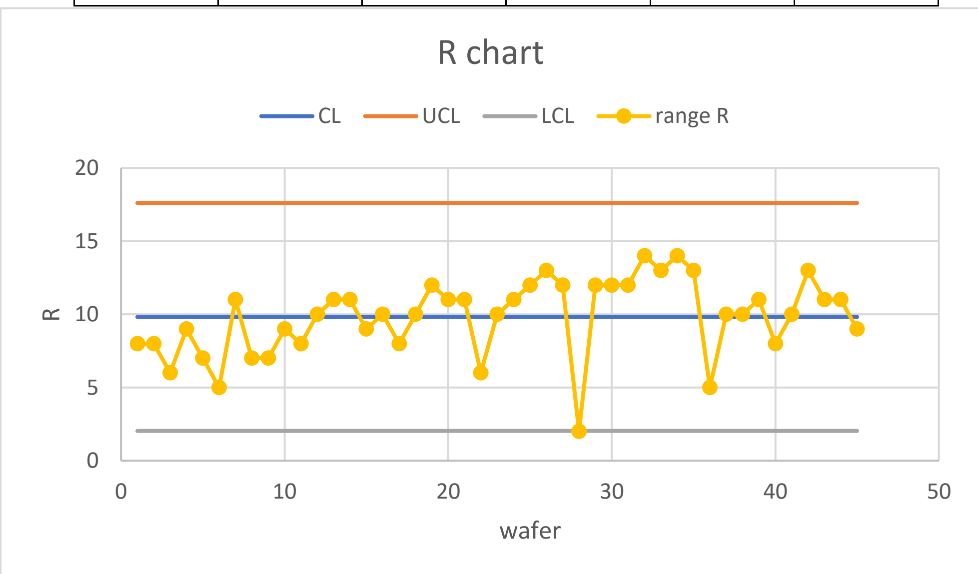


3 thickness values exceeded the boundary of $\bar{X}$ chart.

Sample mean of range $\bar{\bar{R}}$ and standard error $s_{\bar{R}}$ are estimated from first 45 wafers. The control lines are constructed as following:

$$CL = \bar{R}, UCL = \bar{R} + 3\sigma\ , LCL = \bar{R} - 3\sigma$$

$\bar{R}$	sample variance	standard error σ	CL	UCL	LCL
9.822222	6.740404	2.596229	9.822222	17.61091	2.033536

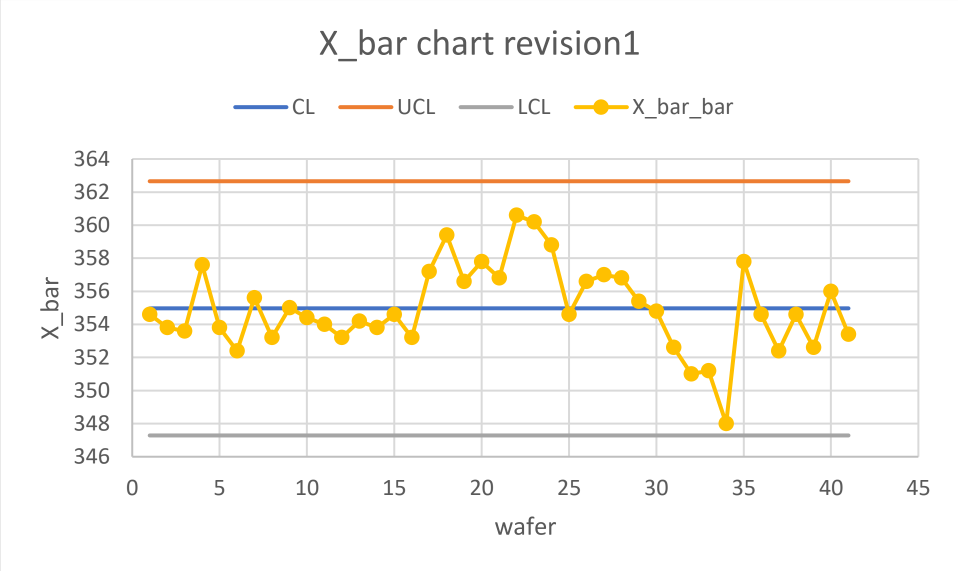


One R value exceeded the boundary of R chart.

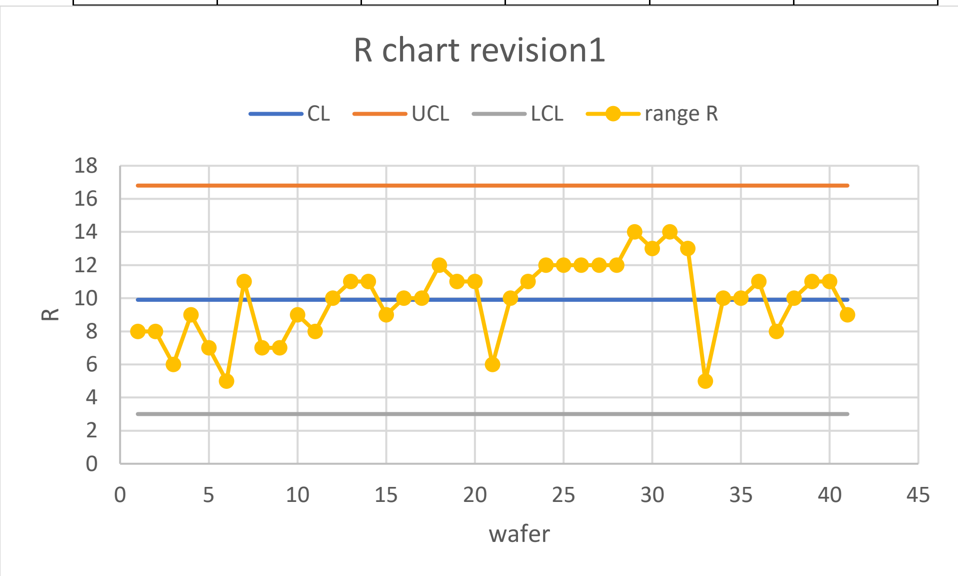
Remove the extreme wafers in the previous $\bar{X}$ chart and R chart, and construct a $\bar{X}$ chart and R chart again with the rest of 41 wafers.

$$\Rightarrow rivision$$

$\bar{\bar{X}}$	sample variance	standard error σ	CL	UCL	LCL
354.9707	6.561122	2.561469	354.9707	362.6551	347.2863

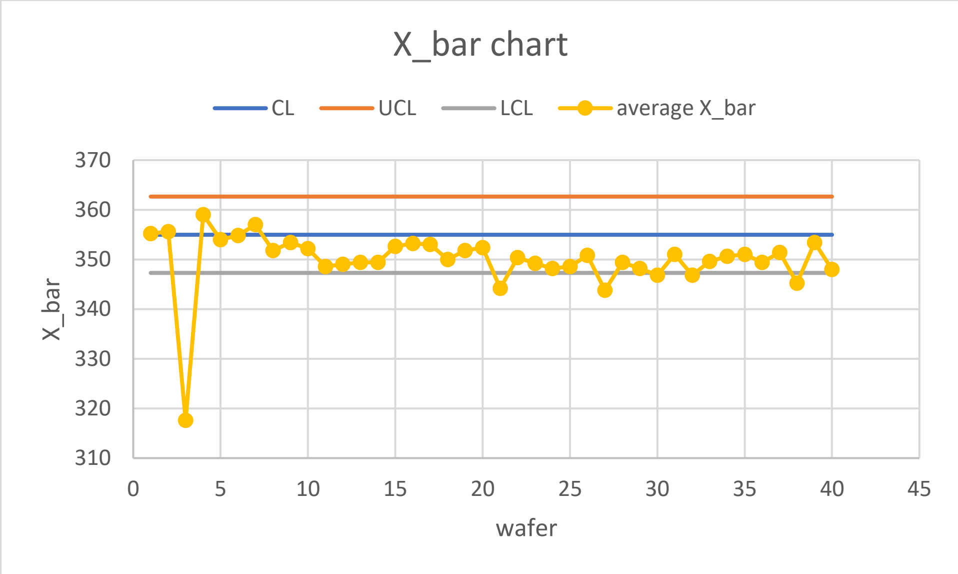


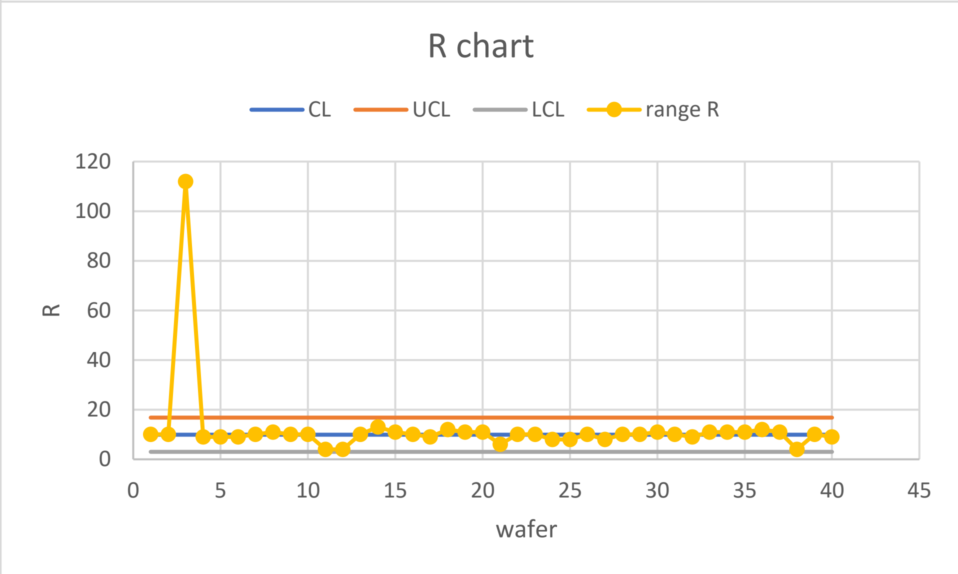
$\bar{R}$	sample variance	standard error σ	CL	UCL	LCL
9.902439	5.290244	2.300053	9.902439	16.8026	3.00228



The result shows that there isn't any value exceed the boundary in $\bar{X}$ chart and R chart. The CL, UCL, and LCL in both previous charts can be used to construct the Shewhart $\bar{X} - R$ control chart to monitor the last 40 wafers.

$$\Rightarrow monitor\ last\ 40\ wafers$$





It is obvious that there is a wafer out of range in the both charts. Most of the wafers' $\bar{\bar{X}}$ are lower than the control line, and some wafers' $\bar{\bar{X}}$ are lower than the control bound in $\bar{X}$ chart.

v.

In the sequential likelihood ratio test with $\alpha = 0.003, \beta = 0.2$, μ_0 is 350, standard error $\sigma_{\bar{X}}$ could be estimated by $s_{\bar{X}}$, $\Delta = \mu_1 - \mu_0$ is -1.6σ, and t could be 1 to 40. C_t must be constructed by the following equation. (C_t will be reset to zero if C_t exceed the boundaries of sequential likelihood ratio test.)

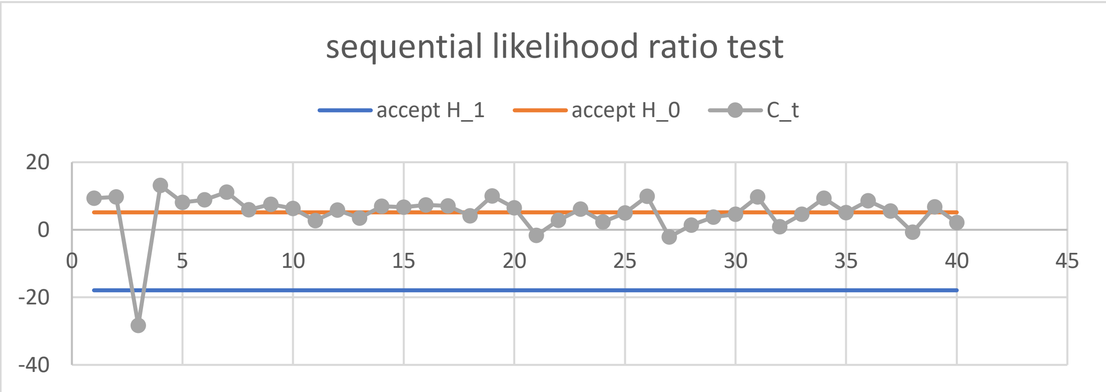
$$\triangle = \mu_1 - \mu_0 = -1.6\sigma_{\bar{X}} = -8.211872361$$

$$C_t = \sum_{i=1}^{t}\left(x_i - \frac{\mu_0 + \mu_1}{2}\right), \mu_1 < \mu_0$$

$$H_1: \mu = \mu_0 \rightarrow C_t \leq \frac{\sigma^2}{\Delta} ln\left(\frac{1-\beta}{\alpha}\right), \Delta = \mu_1 - \mu_0$$

$$H_0 : \mu = \mu_1 \rightarrow C_t \geq \frac{\sigma^2}{\Delta} ln\left(\frac{\beta}{1-\alpha}\right), \Delta = \mu_1 - \mu_0$$

average	sample variance	standard error σ	μ_0	μ_1	Δ	α	β	accept H_1 (if lower)	accept H_0(if higher)
354.3911	26.34174	5.13242	350	341.7881	-8.21187	0.003	0.2	-17.9186	5.153057

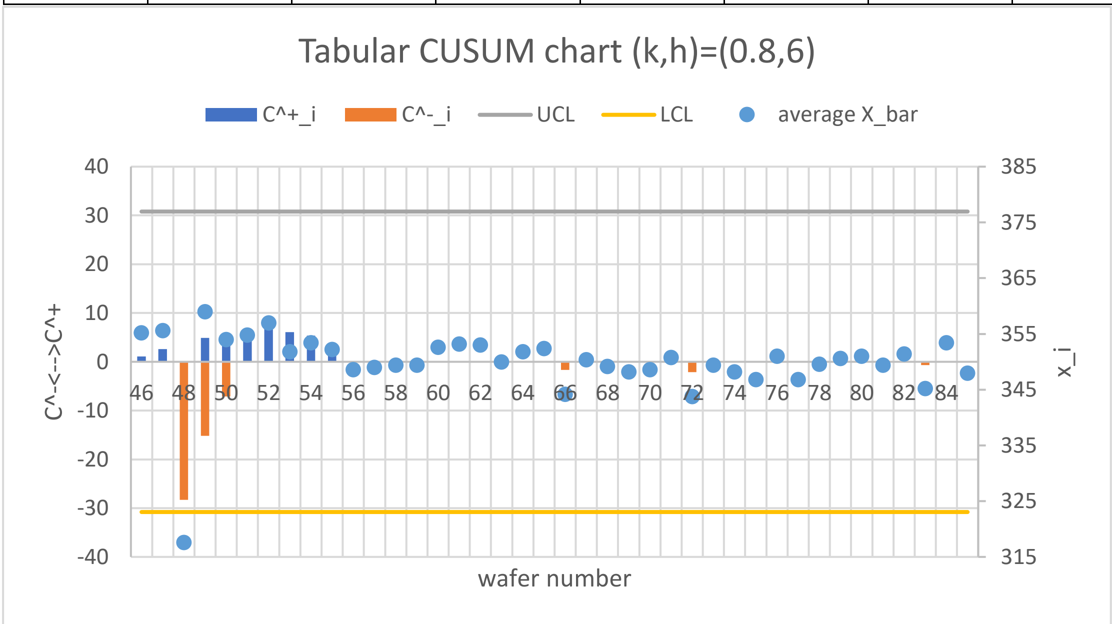


In the "graphical" Tabular CUSUM chart, $K = k\sigma, H = h\sigma. C^+$ and C^- can be constructed by the following equation:

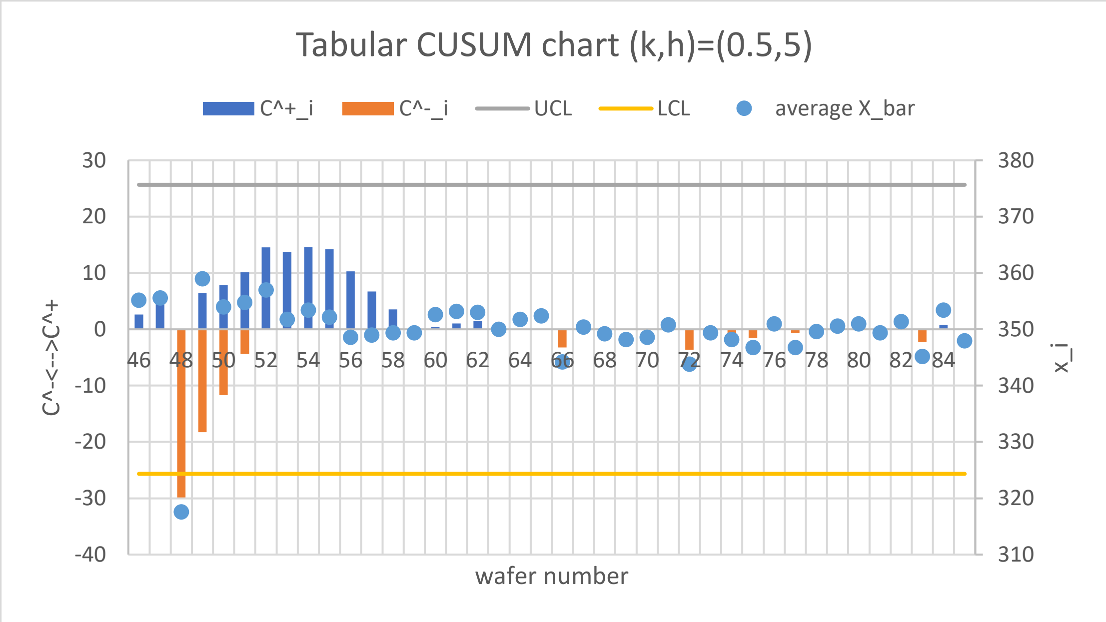
$$C_i^+ = \max[0, x_i - (\mu_0 + K) + C_{i-1}^+], C_0^+ = 0$$

$$C_i^- = \min[0, x_i - (\mu_0 - K) + C_{i-1}^-], C_0^+ = 0$$

average	sample variance	standard error σ	μ_0	k	K	h	H
354.3911	26.34174	5.13242	350	0.8	4.105936	6	30.79452



average	sample variance	standard error σ	μ_0	k	K	h	H
354.3911	26.34174	5.13242	350	0.5	2.56621	5	25.6621



Comparison:

Sequential likelihood ratio test shows that most of the C_t tend to close to H_0. It accepts H_0 and rejects H_1. Between wafers 45 and 58, there are some C^+ become positive in second Tabular CUSUM chart. Similar pattern can be found in the first Tabular CUSUM chart from wafer 45 to 55. However, the majority of both C^+ and C^- values are zero, suggesting that the process remains in control. There is an out -of-control signal in Tabular CUSUM chart with $(k, h) = (0.5,5)$.

The new process mean is : $\bar{\mu} = \mu_0 - K + \frac{C_3^-}{N^-} =$ 317.6 . Since the $N^- = 1$, and C^- begin to decrease dramatically, it seems that only wafer 48 out of control, not the total process.

vi.

In the optimal Tabular CUSUM chart, to compare the ARL_1 with different shifted means, k^* and h' should be fixed. Only δ is changed. h^* and ARL_1 can be found by following:

$$\mathrm{k}^* = \delta^*/2 = 1.6/2 = 0.8$$

$$ARL_0 = \frac{e^{2k^*h'} - 1 - 2k^*h'}{(2k^*)^2} = 500$$

$$\Rightarrow by \text{ "Goal seek" } in\ Excel, h' = 4.475605986, \qquad h^* = 3.309605986$$

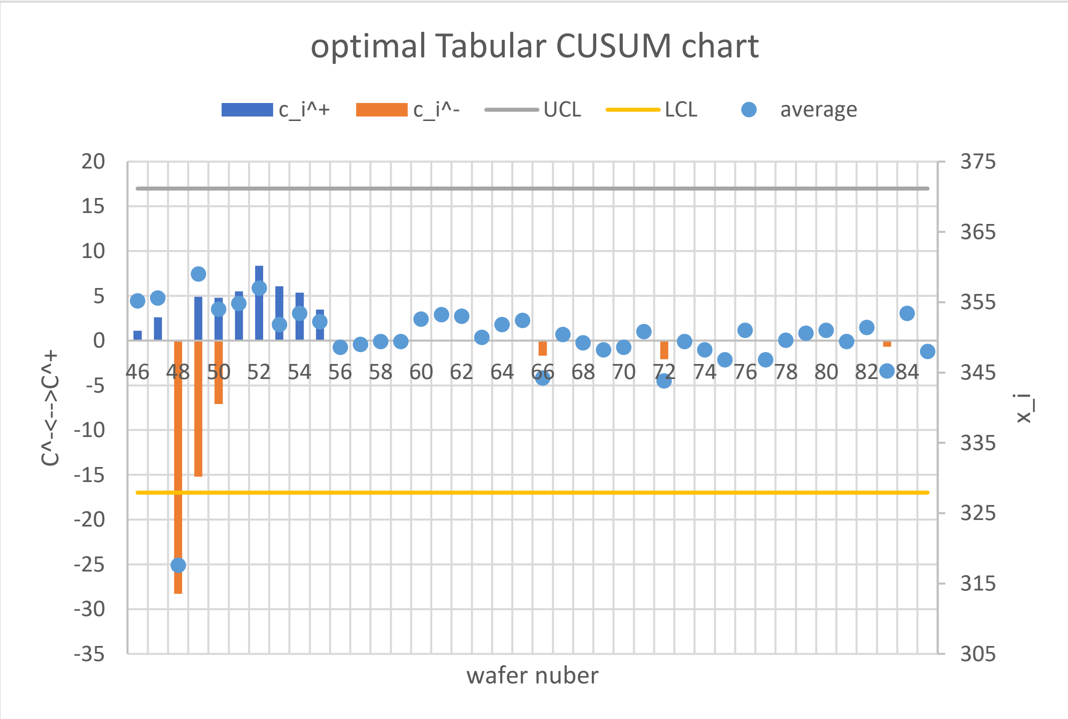
$$ARL_1^* = \left[\frac{2(\delta-k)^2}{e^{-2(\delta-k)h'}-1+2(\delta-k)h'} + \frac{2(\delta+k)^2}{e^{2(\delta+k)h'}-1-2(\delta+k)h'}\right]^{-1}$$

Comparison:

shift	δ	k*	h'	ARL_1
0.5σ	0.5	0.8	4.475606	60.87891
1.0σ	1	0.8	4.475606	11.96443
1.5σ	1.5	0.8	4.475606	5.375253
2.0σ	2	0.8	4.475606	3.382457

In the optimal Tabular CUSUM chart, $\sigma = 5.13242$ is estimated by the first 45 wafers, $\mu_0 = 350$, $K = k\sigma = 4.1059$, $H = 16.9863$. C^+ and C^- can be constructed by the following equation:

$$C_i^+ = \max[0, x_i - (\mu_0 + K) + C_{i-1}^+], C_0^+ = 0$$
$$C_i^- = \min[0, x_i - (\mu_0 - K) + C_{i-1}^-], C_0^+ = 0$$



There is an out-of-control signal in optimal Tabular CUSUM chart. The new process mean is : $\bar{\mu} = \mu_0 - K + \frac{C_3^-}{N^-} = 317.6$

vii.

In the optimal EWMA chart, to compare the ARL_1 with different shifted means,λ^* and L^* should be fixed.

$$\lambda^* \approx \frac{1.0234\delta^{*2}}{b - ln\, b}, where\ b = 2\ ln\ (1.0234\left(\frac{2}{\pi}\right)^{1/2} \delta^{*2} ARL_0$$

$$L^* \approx (b - ln\, b)^{1/2} - \lambda^*$$

g and w could be found by the following equation:

$$g = L\left(\frac{\lambda}{2\delta^2}\right)^{1/2}, w = L + 1.166(\delta\lambda)^{1/2} - \left(\frac{2\delta^2}{\lambda}\right)^{1/2}$$

ARL_1 could be estimated by different equations, depend on the value of g:

$$\text{if}(\text{g} < 1) \Rightarrow \text{ARL}_{1|\delta} \approx -\frac{1}{\lambda} ln(1-g) - \frac{g}{4(1-g)\delta^2} + \frac{3}{4}$$

$$if(g > 1\ and\ \delta \leq 1) \Rightarrow ARL_{1|\delta} \approx \frac{1}{\lambda w}[\phi(w)]^{-1}\Phi(w)$$

$$\phi(w) = n(w, 0,1), \Phi(w) = N(w, 0,1)$$

Comparison:

shift	δ	λ*	L*	g	w	ϕ(w)	Φ(w)	ARL_1
0.5σ	0.5	0.232431	3.12491	2.130589	2.055716	0.048223	0.980095	42.53612
1.0σ	1	0.232431	3.12491	1.065295	0.753675	0.300306	0.774478	14.72196
1.5σ	1.5	0.232431	3.12491	0.710196	-0.58667	0.33587	0.278711	5.806391
2.0σ	2	0.232431	3.12491	0.532647	-1.94685	0.059961	0.025776	3.95144

When $\delta = 0.5$, the ARL_1 in EWMA is smaller than CUSUM by about 18.343. This indicates that EWMA could be more sensitive to a slight shift in the mean compared to CUSUM. However, when $\delta = 1$, the ARL_1 in CUSUM drop sharply to 11.964 which is smaller than the ARL_1 in EWMA. Both ARL_1 in CUSUM and EWMA are small when mean shift become bigger.

In the optimal EWMA chart, Z_i could be calculated by the equation:

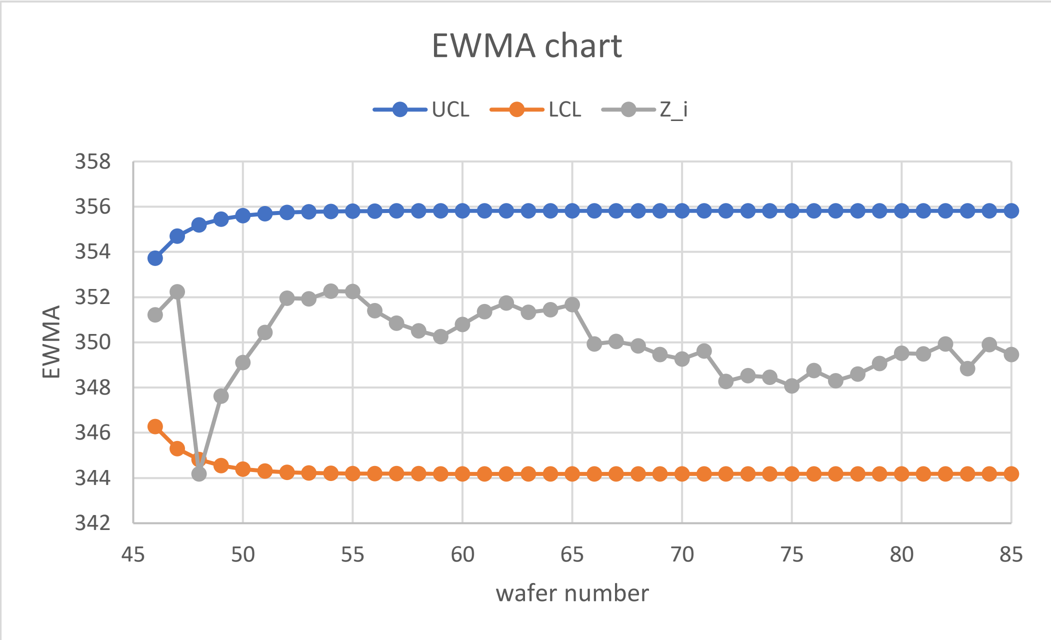
$$Z_i = \lambda x_i + (1-\lambda)Z_{i-1}$$

I suppose it is zero state, therefor, $Z_0 = \mu_0 = 350$.

The control limits are:

$$UCL = \mu_0 + L\sigma\sqrt{\frac{\lambda}{2-\lambda}[1-(1-\lambda)^{2i}]}$$

$$LCL = \mu_0 - L\sigma\sqrt{\frac{\lambda}{2-\lambda}[1-(1-\lambda)^{2i}]}$$



There is also an out-of-control signal in EWMA chart. Both chats indicate that wafer 48 dose not meet the requirement. Nevertheless, the majority of wafer's values tend to stay in the control limits in both charts, suggesting that the process remains in control.

4.

i.

To estimate all 425 thickness readings in the bottom zone, the $\bar{X}$ is the average of all 425 values, and the σ_X is the standard error of total 425 values. USL and LSL are 335 and 360 respectively.

C_p could be calculated by the following equation:

$$C_p = \frac{USL - LSL}{6\sigma_X}$$

C_{pk} could be calculated by the following equation:

$$C_{pk} = min\left[\frac{\bar{X} - LSL}{3\sigma_X}, \frac{USL - \bar{X}}{3\sigma_X}\right]$$

T is the target which is $\mu_0 = 350$, and μ is the sample average $\bar{X}$. $\tilde{\sigma}$ could found by following:

$$\tilde{\sigma} = \sqrt{\sigma_X^2 + (\mu - T)^2}$$

C_{pm} could be calculated by the following equation:

$$C_{pm} = \frac{USL - LSL}{6\tilde{\sigma}}$$

C_{pm}^* could be calculated by the following equation:

$$C_{pm}^{*} = \frac{min(USL - T, T - LSL)}{3\tilde{\sigma}}$$

Out-of-spec% could be estimated by the following equations:

$$\text{out-of-spec\%} = [P(X \geq USL) + P(X \leq LSL)] \times 100\%$$

$$Z_U = \frac{USL - \bar{X}}{\sigma_X}, Z_L = \frac{LSL - \bar{X}}{\sigma_X}$$

$\bar{X}$	μ_0	sample variance	standard error σ_x	$\tilde{\sigma}$	USL	LSL	Z_U	Z_L
352.2776	350	82.62084	9.089601	9.37062	360	335	0.849581	-1.90081

C_p	C_{pk}	C_{pm}	C_{pm}^{*}	out-of-spec%
0.458399	0.283194	0.444652	0.355722	22.64%

To improve the process to achieve overall $C_{pk} = 2.0$, sample average $\bar{X}$ should be reduced, and σ_X should be smaller. The following table present various combination of $\bar{X}$ and σ_X making the C_{pk} achieve 2.0.

$\bar{X}$	352.2776	351.0832	349.8888	348.6944	347.5
σ	1.286904	1.485969	1.685159	1.883921	2.083314
C_{pk}	2.000241	2.000213	2.000044	2.000365	2.000019

To improve the process to achieve overall $C_{pm} = 2.0$, $\tilde{\sigma}$ should become smaller. The equation $\tilde{\sigma} = \sqrt{\sigma_X^2 + (T - \bar{X})^2}$ indicates that σ_X should be reduced. On the other hand, $\bar{X}$ should be as close to target T as possible.

If $\mu - T = 0 \Rightarrow \tilde{\sigma} = \sigma_X \Rightarrow \sigma = 12.5$

ii.

To estimate the "within-wafer" reading, the $\bar{X}$ is the average of five thickness values in each wafer, and the σ_X is the standard error of five thickness values in each wafer. USL and LSL are 335 and 360 respectively.

C_p could be calculated by the following equation:

$$C_p = \frac{USL - LSL}{6\sigma_X}$$

C_{pk} could be calculated by the following equation:

$$C_{pk} = min\left[\frac{\bar{X} - LSL}{3\sigma_X}, \frac{USL - \bar{X}}{3\sigma_X}\right]$$

T is the target which is $\mu_0 = 350$, and μ is the sample average $\bar{X}$. $\tilde{\sigma}$ could found by following:

$$\tilde{\sigma} = \sqrt{\sigma_X^2 + (\mu - T)^2}$$

C_{pm} could be calculated by the following equation:

$$C_{pm} = \frac{USL - LSL}{6\tilde{\sigma}}$$

C_{pm}^* could be calculated by the following equation:

$$C_{pm}^* = \frac{min(USL - T, T - LSL)}{3\tilde{\sigma}}$$

Out-of-spec% could be estimated by the following equations:

$$\text{out-of-spec\%} = [P(X \geq USL) + P(X \leq LSL)] \times 100\%$$

$$Z_U = \frac{USL - \bar{X}}{\sigma_X}, Z_L = \frac{LSL - \bar{X}}{\sigma_X}$$

wafer number	X_bar	σ_X	σ^~	Z_U	Z_L
1	354.6	3.209361	5.608921	1.682578	-6.10713
2	353.8	3.114482	4.913247	1.9907	-6.03632
3	353.6	3.130495	4.770744	2.044405	-5.94155
4	357.6	3.847077	8.518216	0.62385	-5.87459
5	353.8	2.774887	4.705316	2.234325	-6.77505
6	352.4	2.19089	3.249615	3.46891	-7.94198
7	355.6	5.029911	7.527284	0.874767	-4.0955
8	353.2	3.563706	4.789572	1.908126	-5.10704
9	355	3.24037	5.958188	1.543033	-6.17213
10	354.4	3.435113	5.582114	1.630223	-5.64756
11	354	3	5	2	-6.33333
12	353.2	4.207137	5.28583	1.616301	-4.32598
13	354.2	5.357238	6.807349	1.082647	-3.58394
14	353.8	4.868265	6.175759	1.273554	-3.86175
15	354.6	4.929503	6.742403	1.095445	-3.97606
16	353.2	4.494441	5.517246	1.51298	-4.04945
17	370.4	3.507136	20.69928	-2.96538	-10.0937
18	357.2	4.147288	8.309031	0.67514	-5.3529
19	359.4	6.0663	11.18749	0.098907	-4.02222
20	356.6	5.412947	8.535807	0.628124	-3.99043
21	357.8	5.674504	9.645724	0.387699	-4.01797
22	356.8	3.03315	7.445804	1.055009	-7.18725
23	360.6	4.97996	11.71153	-0.12048	-5.1406
24	360.2	4.969909	11.34637	-0.04024	-5.07051

25	358.8	5.357238	10.30243	0.223996	-4.44259
26	338	4.949747	12.98075	4.444671	-0.60609
27	354.6	5.899152	7.480642	0.915386	-3.32251
28	349.2	0.83666	1.157584	12.90847	-16.9722
29	356.6	5.549775	8.623224	0.612637	-3.89205
30	357	5.09902	8.660254	0.588348	-4.31455
31	356.8	5.761944	8.912912	0.555368	-3.78345
32	355.4	6.580274	8.512344	0.699059	-3.10018
33	354.8	6.180615	7.825599	0.84134	-3.20356
34	352.6	6.426508	6.932532	1.151481	-2.73866
35	351	5.244044	5.338539	1.716233	-3.05108
36	351.2	1.923538	2.267157	4.574902	-8.42198
37	348	4.795832	5.196152	2.502173	-2.71069
38	357.8	4.816638	9.167333	0.45675	-4.73359
39	354.6	5.458938	7.138627	0.989203	-3.59044
40	352.4	3.286335	4.069398	2.312606	-5.29465
41	354.6	4.97996	6.779381	1.084346	-3.93577
42	336.2	4.816638	14.61643	4.941206	-0.24914
43	352.6	5.412947	6.004998	1.367093	-3.25146
44	356	5.522681	8.154753	0.724286	-3.8025
45	353.4	4.27785	5.46443	1.542831	-4.30123
46	355.2	4.868265	7.123202	0.985978	-4.14932
47	355.6	4.560702	7.222188	0.964764	-4.51685
48	317.6	55.55448	64.31221	0.763215	0.313206
49	359	4.582576	10.0995	0.218218	-5.23723
50	354	4.582576	6.082763	1.309307	-4.14614
51	354.8	4.764452	6.763135	1.091416	-4.15578
52	357	4.690416	8.42615	0.639602	-4.69042
53	351.8	5.263079	5.562374	1.558023	-3.19205
54	353.4	5.128353	6.153048	1.286963	-3.5879
55	352.2	4.868265	5.342284	1.602214	-3.53309
56	348.6	1.516575	2.063977	7.516937	-8.96757
57	349	1.581139	1.870829	6.957011	-8.85438
58	349.4	5.128353	5.163332	2.066941	-2.80792
59	349.4	6.14817	6.177378	1.72409	-2.34216
60	352.6	5.029911	5.662155	1.471199	-3.49907
61	353.2	4.32435	5.379591	1.572491	-4.20873

62	353	3.674235	4.743416	1.905159	-4.89898
63	350	5.244044	5.244044	1.906925	-2.86039
64	351.8	5.263079	5.562374	1.558023	-3.19205
65	352.4	5.176872	5.706137	1.468068	-3.3611
66	344.2	2.683282	6.390618	5.888312	-3.42864
67	350.4	4.722288	4.739198	2.032913	-3.26113
68	349.2	4.494441	4.565085	2.402968	-3.15946
69	348.2	4.38178	4.737088	2.692969	-3.01247
70	348.6	4.09878	4.331282	2.781315	-3.31806
71	350.8	4.868265	4.933559	1.88979	-3.24551
72	343.8	3.49285	7.116179	4.638047	-2.51943
73	349.4	5.176872	5.211526	2.047569	-2.7816
74	348.2	4.91935	5.23832	2.398691	-2.68328
75	346.8	5.263079	6.159545	2.508038	-2.24203
76	351	5.477226	5.567764	1.643168	-2.92119
77	346.8	4.38178	5.425864	3.012474	-2.69297
78	349.6	5.458938	5.473573	1.905133	-2.67451
79	350.6	5.412947	5.4461	1.736577	-2.88198
80	351	5.612486	5.700877	1.603567	-2.85079
81	349.4	5.366563	5.4	1.975193	-2.68328
82	351.4	5.59464	5.767148	1.537186	-2.93138
83	345.2	1.788854	5.122499	8.273452	-5.70197
84	353.4	4.722288	5.818935	1.397628	-3.89642
85	348	4.582576	5	2.618615	-2.83683

wafer number	C_p	C_pk	C_pm	C*_pm	out-of-spec%
1	1.298285	0.560859	0.742864	0.594291	0.046228429
2	1.337836	0.663567	0.848047	0.678438	0.023256939
3	1.330993	0.681468	0.873379	0.698703	0.020456778
4	1.083073	0.20795	0.489148	0.391318	0.266362955
5	1.501562	0.744775	0.885523	0.708419	0.012730851
6	1.901814	1.156303	1.282203	1.025762	0.000261288
7	0.828378	0.291589	0.553542	0.442833	0.190871397
8	1.169195	0.636042	0.869946	0.695956	0.028187629
9	1.285861	0.514344	0.699318	0.559454	0.061411324
10	1.212964	0.543408	0.746432	0.597145	0.051527193

11	1.388889	0.666667	0.833333	0.666667	0.022750132
12	0.990381	0.538767	0.788271	0.630617	0.053022203
13	0.777764	0.360882	0.612084	0.489667	0.139651703
14	0.855883	0.424518	0.674681	0.539745	0.101466972
15	0.845251	0.365148	0.617979	0.494384	0.136695872
16	0.927071	0.504327	0.755208	0.604166	0.065168021
17	1.188054	-0.98846	0.201295	0.161036	0.998488468
18	1.004673	0.225047	0.501462	0.40117	0.24979347
19	0.686855	0.032969	0.37244	0.297952	0.460634853
20	0.769759	0.209375	0.48814	0.390512	0.264994466
21	0.734279	0.129233	0.43197	0.345576	0.349148727
22	1.373709	0.35167	0.559599	0.447679	0.14571066
23	0.836687	-0.04016	0.355775	0.28462	0.547949824
24	0.838379	-0.01341	0.367225	0.29378	0.516050174
25	0.777764	0.074665	0.404435	0.323548	0.411384645
26	0.841794	0.202031	0.320988	0.25679	0.27223139
27	0.706316	0.305129	0.556993	0.445595	0.180440639
28	4.980119	4.302823	3.599452	2.879561	6.58965E-65
29	0.750781	0.204212	0.483191	0.386553	0.270107737
30	0.817151	0.196116	0.481125	0.3849	0.278157227
31	0.723136	0.185123	0.467487	0.373989	0.289398787
32	0.633206	0.23302	0.489485	0.391588	0.243224585
33	0.674151	0.280447	0.532441	0.425952	0.200757364
34	0.648356	0.383827	0.601031	0.480825	0.127851784
35	0.794552	0.572078	0.780488	0.624391	0.044199832
36	2.166147	1.524967	1.837838	1.47027	2.38221E-06
37	0.86881	0.834058	0.801875	0.6415	0.009528875
38	0.865057	0.15225	0.454512	0.36361	0.323926427
39	0.763274	0.329734	0.583679	0.466943	0.161446861
40	1.267876	0.770869	1.023902	0.819122	0.010372206
41	0.836687	0.361449	0.614609	0.491687	0.139147152
42	0.865057	0.083045	0.285067	0.228054	0.40162801
43	0.769759	0.455698	0.693866	0.555093	0.086372216
44	0.754465	0.241429	0.510949	0.40876	0.234516716
45	0.97401	0.514277	0.762507	0.610006	0.061444371
46	0.855883	0.328659	0.584943	0.467954	0.162088717
47	0.913602	0.321588	0.576926	0.461541	0.167334699

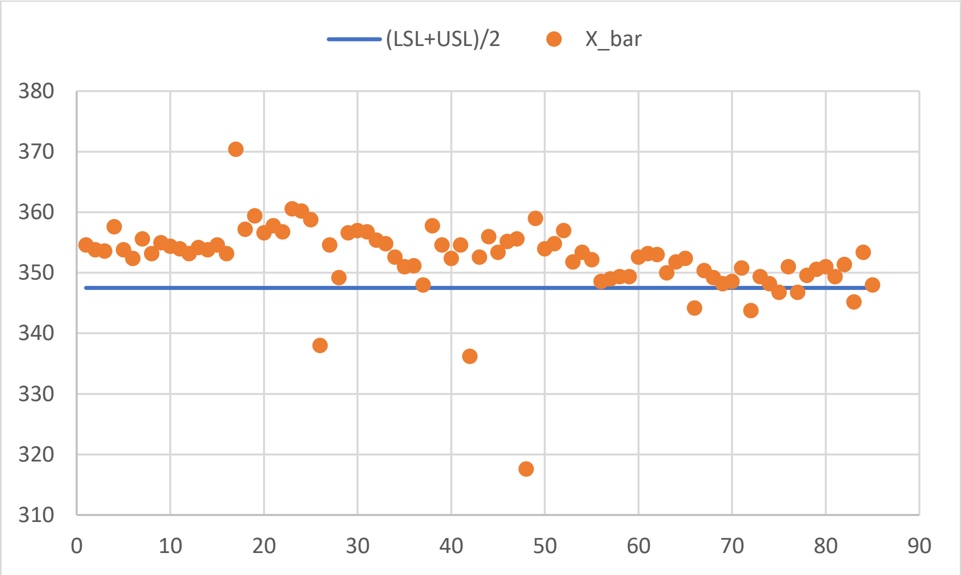
48	0.075001	-0.1044	0.064788	0.05183	0.845605599
49	0.909241	0.072739	0.412561	0.330049	0.413629755
50	0.909241	0.436436	0.684996	0.547997	0.095232038
51	0.874532	0.363805	0.616085	0.492868	0.137561092
52	0.888336	0.213201	0.494492	0.395594	0.261217006
53	0.791679	0.519341	0.749081	0.599265	0.060320194
54	0.812477	0.428988	0.677171	0.541737	0.09922027
55	0.855883	0.534071	0.779941	0.623953	0.054759554
56	2.747419	2.505646	2.018757	1.615005	2.80888E-14
57	2.635231	2.319004	2.227177	1.781742	1.73783E-12
58	0.812477	0.68898	0.806972	0.645578	0.021863013
59	0.677708	0.574697	0.674504	0.539603	0.051932029
60	0.828378	0.4904	0.73588	0.588704	0.07085208
61	0.963536	0.524164	0.774532	0.619626	0.057931228
62	1.134023	0.635053	0.87841	0.702728	0.028380205
63	0.794552	0.635642	0.794552	0.635642	0.030380755
64	0.791679	0.519341	0.749081	0.599265	0.060320194
65	0.804862	0.489356	0.730208	0.584166	0.071431025
66	1.552825	1.142879	0.651997	0.521598	0.000303311
67	0.882341	0.677638	0.879192	0.703354	0.021585498
68	0.927071	0.800989	0.912725	0.73018	0.008921608
69	0.950907	0.897656	0.879584	0.703667	0.004836578
70	1.016563	0.927105	0.961994	0.769595	0.003160182
71	0.855883	0.62993	0.844556	0.675645	0.029979197
72	1.192913	0.839811	0.58552	0.468416	0.005878964
73	0.804862	0.682523	0.79951	0.639608	0.023005704
74	0.846995	0.799564	0.79542	0.636336	0.011872073
75	0.791679	0.747345	0.676457	0.541165	0.018549786
76	0.760726	0.547723	0.748355	0.598684	0.051917626
77	0.950907	0.897656	0.767927	0.614341	0.004836578
78	0.763274	0.635044	0.761233	0.608987	0.032123312
79	0.769759	0.578859	0.765074	0.612059	0.043206863
80	0.742392	0.534522	0.730882	0.584705	0.056585277
81	0.776412	0.658398	0.771605	0.617284	0.027768282
82	0.74476	0.512395	0.722483	0.577986	0.063811258
83	2.329237	1.900658	0.813405	0.650724	5.92142E-09
84	0.882341	0.465876	0.716053	0.572843	0.081161259

85	0.909241	0.872872	0.833333	0.666667	0.006692555

To improve the process to achieve "within-wafer" $C_{pk} = 2.0$:

- The $\bar{X}$ in each wafer should be close to $\frac{LSL+USL}{2}$.
- The σ_X in each wafer should be as small as possible.

The following graph compare the $\bar{X}$ in each wafer with $\frac{LSL+USL}{2}$.



The graph indicates that most of the wafers' $\bar{X}$ is bigger than $\frac{USL+LSL}{2}$. The process may be adjusted to reduce the SiO2 thickness of wafers and lessen their standard errors, aiming to achieve "within-wafer" $C_{pk} = 2.0$.

To improve the process to achieve "within-wafer" $C_{pm} = 2.0$, the $\tilde{\sigma}$ in each wafer should be smaller. Since the equation $\tilde{\sigma} = \sqrt{\sigma_X^2 + (T - \bar{X})^2}$, there are some solutions:

- The $\bar{X}$ in each wafer should be close to T.
- The σ_X in each wafer should be as small as possible.