830-phase4

Yuqing Zhang

2021/8/16

PHASE4: Confirmation

Objective:

After finding the optimal operating conditions, we hope to confirm that the "optimal" factor settings $\boldsymbol{x}^*$ are indeed "optimal", and that the optimized response $\hat{y}(\boldsymbol{x}^*)$ is reproducible.

In this phase, I used the **comparative probability metrics** to do the analysis. At a given location x (the combination of factor levels), our interest lies in comparing estimates of the response surface as determined by

- the response surface experiment, $\hat{y}(\boldsymbol{x})$
- the confirmation experiment, $\hat{y}_c(\boldsymbol{x}) = \frac{1}{n_c} \sum_{l=1}^{n_c} y_l$

Our interest lies in characterizing the difference $\hat{y}_c(\boldsymbol{x}) - \hat{y}(\boldsymbol{x})$

The **commensurate probability** is used to define how much better, or at least not worse, our confirmation results are relative to what we predicted from the response surface. It can be defined as $\theta_C(\boldsymbol{x}) = \Pr\left(\hat{y}_c(\boldsymbol{x}) - \hat{y}(\boldsymbol{x}) < \delta\right)$

The approach has the advantage of emphasizing practical importance over statistical significance. It can take the practitioner's input, so we choose it instead of an interval approach.

Data & Experiments

We have identified the optimal process conditions in phase III to be $(x_1, x_2) = (70, 77)$ yielding a predicted browse time of 10.29217 minutes with a 95% prediction interval given by $(10.1151, 10.46921)$.

$n_c = 100$ confirmation runs were performed at this optimum to validate the predicted response. The confirmation runs suggest an average browse time of 10.4663 could be achieved. If we define that 2% is the threshold that we can be satisfied with our result, then using the commensurate probability we can quantify the likelihood of the confirmation results being within $2\%(\delta = 0.20584)$ of the predicted response.

The probability that $(x_1, x_2) = (70, 77)$ is the optimal combination, is estimated to be $\hat{\theta}_C = 0.5607475$. Given the uncertainty in the response surface model and the variation in confirmation observations, this figure is not high enough to confirm that the result can be in agreement with the fitted response surface.

We collected a very small region data using the x values close to what we observed in the optimal values in Phase III ($\hat{x_1} = 70, \hat{x_2} = 77$, with preview type is TT). Here we chose $x_1 \in \{65, 70, 75\}, x_2 \in \{76, 77, 78\}$. In total, 9 experimental conditions are chosen.

We also explored other points around the optimal point in phase 3. The result can be seen in the table :

prev length	match score	average browse time	commensurate probability
65	76	10.94052	0.4936506
65	77	10.97416	0.3654684
65	78	11.0143	0.2725913
70	76	10.60671	0.5214829
70	77	10.5224	0.6401777
70	78	10.58257	0.5196920
75	76	10.03334	0.9999142
75	77	10.08804	0.9998020
75	78	10.32053	0.9869881

From the table we can see for preview length = 75, the commensurate probability is almost one. This also makes sense because in phase III, our optimal point is 71.54665 and we rounded down to 70.

So overall, we hope to change the optimum to $x_1 = 75, x_2 = 76$.

Conclusion

Given the uncertainty in the response surface model and the variation in confirmation observations, it is relatively unlikely that the confirmation results generated by our estimated optimum in phase III can be in the agreement with the fitted response surface.

By further analysis, we hope to change the optimum to:

- preview length = 75
- matching score = 76
- preview type = TT