Robert W. Allen MBB, PMP

www.robertwalterallen.com

Monte Carlo Analysis / Probabilistic Modeling to Optimize Surgery Scheduling

Summary

Scheduling customers for virtually any process or service (event) can be notoriously inaccurate and unreliable, especially when there are numerous factors affecting the duration of the event, or there are numerous event types. The resultant effect of scheduling difficulties includes wait time for the service provider (staff), wait time for the customers, and later than expected end-of-workday. In addition to wait time, either early or later-than-scheduled start or finish times can result in other forms of waste and non-value-added activity.

Optimized scheduling balances voice-of-the-customer (VOC) with the voice-of-the-process (VOP); achieving this balance reduces inefficiencies associated with not meeting needs of the customer and/or the business.

An example of this is hospital operating room surgery scheduling. The effect of surgery scheduling difficulties includes physician & staff wait-time (for non back-to-back procedures), physicians late out of surgery (at end-of-day), and excessive patient wait-time in surgery centers. In addition to these inefficiencies, starting a surgery prior-to, or after, the scheduled start time may also have process quality implications associated with it.

Background

The approach presented herein uses curve fitting of historical data (of the duration of the events scheduled), and Monte Carlo analysis to determine the cumulative effect of those distributions on subsequent scheduled events. (An “event” is a process with an inherent duration. In the above example, a surgical procedure is the event, with a scheduled duration and an actual duration once the event takes place. This approach can be applied to any event type, however.)

The goal is to apply the model to optimize our scheduled vs. actual event duration based on using a range of discrete percentiles (60th, 65th, 70th and 75th) of historical event times. This can also be described as a probabilistic approach: establishing an optimum percentile based on some probability of being early or late.

The optimum percentile is determined according to the model results. Note that, what’s considered ‘optimum’ is related to the risk associated with being either too ‘aggressive’ or too lenient with scheduling. Would we rather have customer, or staff, wait for an event to begin? What is the associated requirement for our desired range? Modeling results based on historical data helps us make that business decision and make scheduling adjustments accordingly.

Approach

The scheduling process (and data) includes who was performing the procedure (ie. providing the service for the customer), what type of procedure was being performed, the procedure duration (when) and the location the procedure was to take place. (Location was relevant since different procedure types took place in specialized operating rooms.)

For modeling any schedule, the data should be stratified to accommodate all critical inputs affecting the schedule or event durations. In other words, what do we need to know to create the schedule, and what else could we enter into the schedule or model (within reason) which would make our predictions more accurate?

Additional factors considered for any scheduling model might include information about the customer and associated factors affecting the event duration. Similarly, additional (in-house) factors affecting the scheduling could be added to the model. If historical data is available, it can be segregated in this manner and the schedule would be based on knowing this information accordingly. (A regression analysis to determine these factors may therefore be beneficial but is outside the scope of this paper. However, whether segregated or not, the effect of these sources of variation exists in the historical data on event durations.)

Curve fitting was performed based on historical data (durations) for each procedure type. Figure 1 shows a curve fit of a procedure which is a specific type (procedure code INHESLL) performed by a specific (fictitious) service provider (Dr. Ayala).

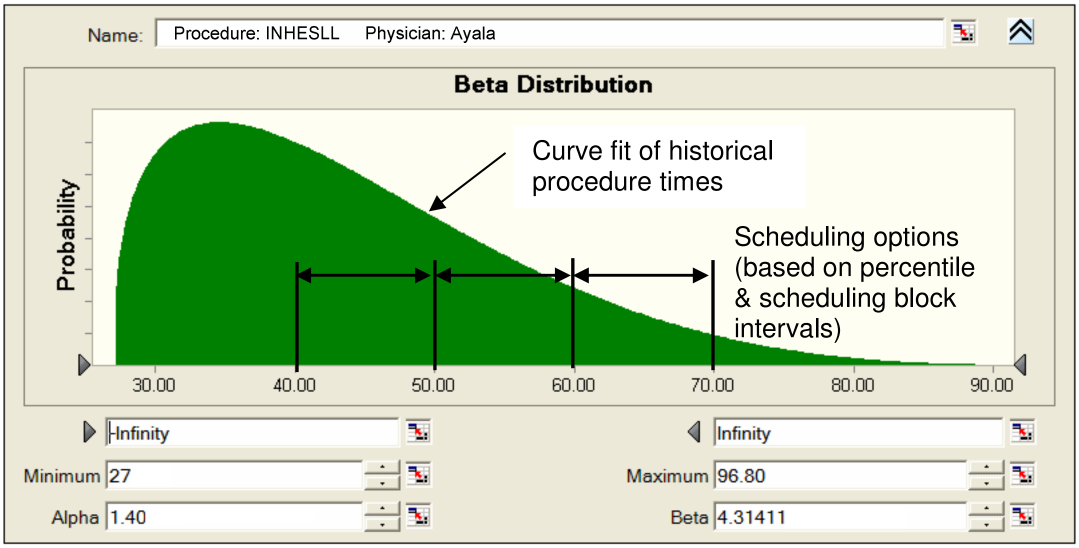

FIGURE 1
Sample of a Curve-Fitted Distribution

Recall these procedures are scheduled back-to-back, therefore, the likelihood of whether a procedure is completed on-time, or not, can be calculated by simply adding the distributions of back-to-back procedures and evaluating this distribution against the scheduled time. (Also, recall that completing a procedure on-time is synonymous with an on-time start of the subsequent procedure).

The approach is best illustrated by the following figures and examples. Table 1 shows a set of subsequent procedures and Figure 2 shows the corresponding series of curve-fit distributions. Note there are only two procedure types shown in Figure 2 up till 11:15am (matching the two event types in Table 1).

Operating Room #	Procedure Code	Physician	Scheduled Start Time
21	INHESLL	Ayala	8:15
21	INHESLL	Ayala	9:15
21	INHESLL	Ayala	10:15
21	INHAMYL	Ayala	11:15
21	INHAMYL	Ayala	12:30
21	INHAMYL	Ayala	13:45
21	VNHEYGL	Ayala	14:45

TABLE 1
Sample Set of Subsequent Procedures

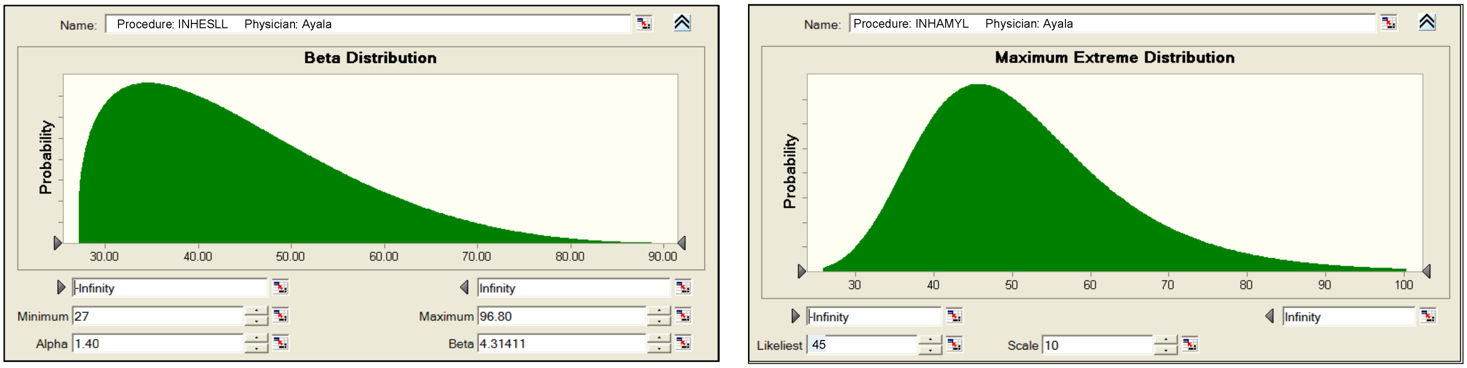

FIGURE 2
Fitted Curves Based on Historical Data
(for two procedure types: INHESLL and INHAMYL)

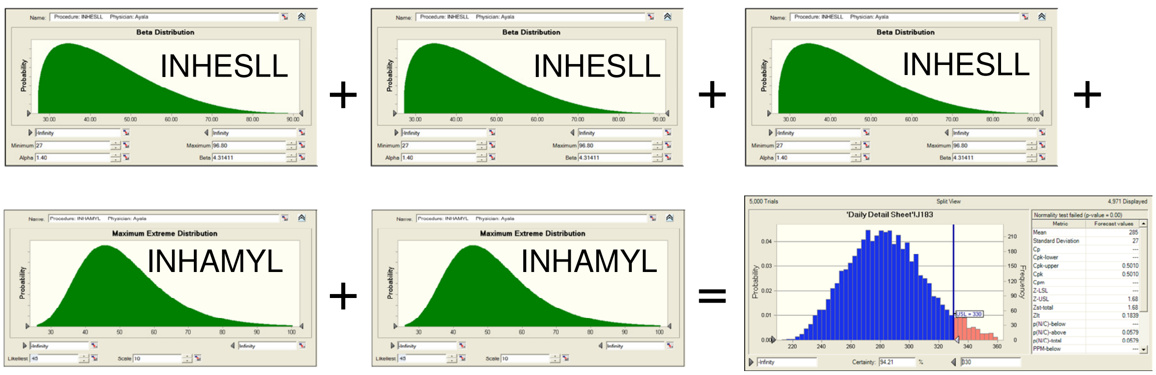

FIGURE 3
Cumulative Effect of 5 Procedures
(based on adding curve-fitted distributions)

Using Monte Carlo analysis, the distributions can be added to obtain a forecast for the probability of start time for the subsequent procedure as illustrated in Figure 3.

In addition to adding distributions to obtain a forecast, <u>scheduled</u> procedure durations (based on using the 75^{th} percentile point of each curve-fitted distribution, rounded to the nearest 10 minute interval on the operating room block-schedule) were added and input as the upper specification limit (USL).

The comparison between the forecast and the USL provides the certainty the subsequent procedure will start on time. In the example forecast shown in Figure 3, the subsequent procedure on-time probability for the procedure which occurs <u>after</u> the 5 procedures (or distributions) are added, is 94.21%.

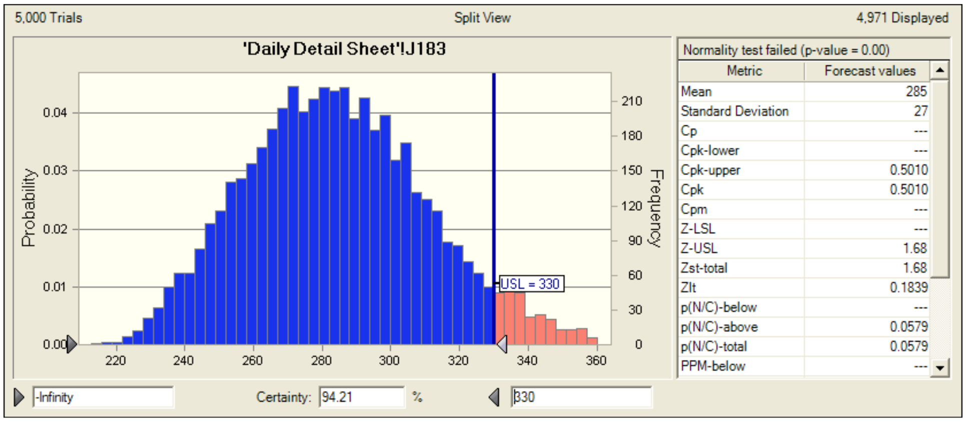

FIGURE 4
Example Forecast of Last Subsequent Procedure On-Time Start Performance
(320 minute cumulative schedule duration)

Daily Performance Chart

Using the above-reference modeling approach, the results for a day worth of subsequent procedure probabilities can now be graphed, as shown in Figure 5. The upward trend suggests over-scheduled durations (ie each subsequent procedure being <u>more and more likely</u> to begin on time because <u>too much</u> time is scheduled). Aside from the upward trend, overall likelihood of on-time procedure starts is high, resulting in staff wait time if the patient isn't ready.

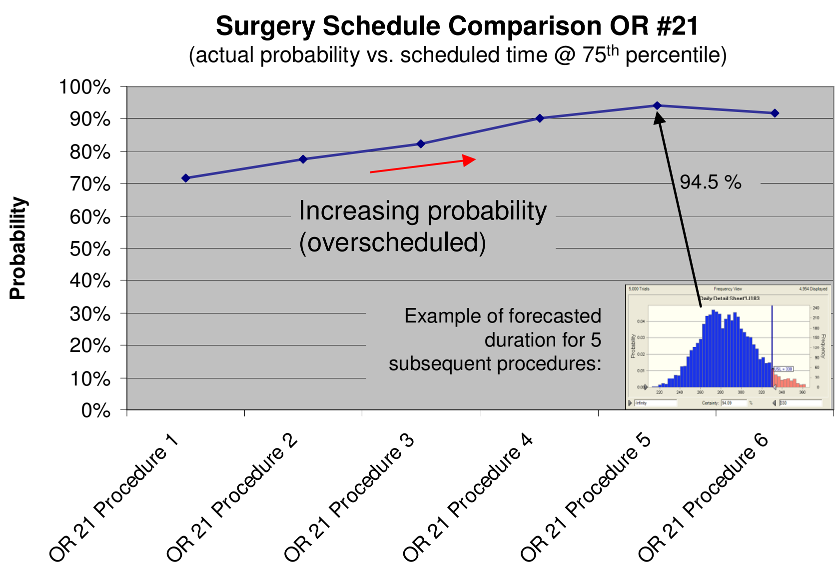

FIGURE 5
Subsequent Procedure Probability @ 75th Percentile Schedule Durations
(note upward trend)

Due to the overscheduling revealed at the 75th percentile point, we can analyze additional scheduling (percentile) points of each curve-fitted distribution (at 70th, 65th and 60th percentiles). Figure 6 is an example of the subsequent probabilities for the 60th percentile. The downward trend suggests under-scheduled durations (ie. each subsequent procedure being <u>less and less likely</u> to begin on time because <u>not enough</u> time is scheduled). Aside from the downward trend, overall likelihood of on-time procedure starts is low (only 30%), resulting in patient wait time.

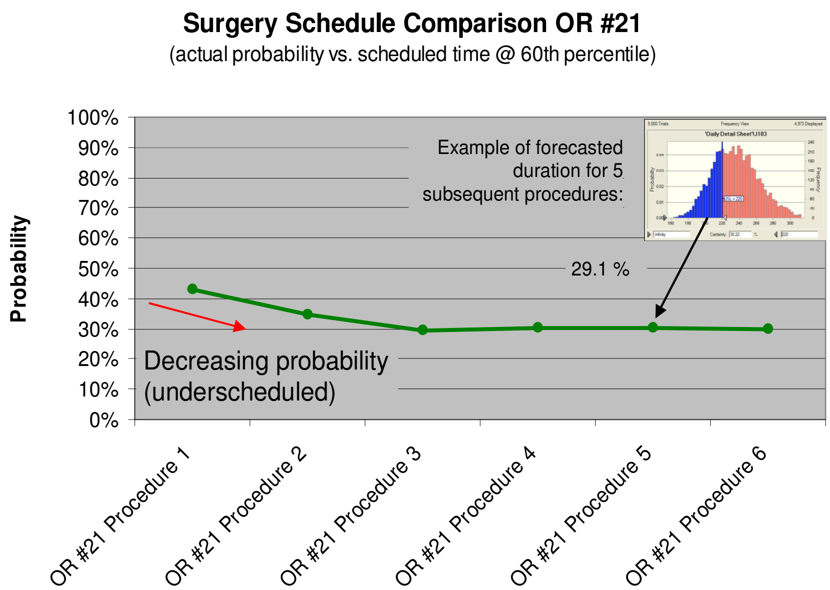

FIGURE 6
Subsequent Procedure Probability @ 60th Percentile Schedule Durations
(note downward trend)

Trend Graphical Analysis

The ideal trend (slope) would therefore be flat or neutral, where the ideal percentile results in neither under-scheduled nor over-scheduled durations. These trends were analyzed for 60th, 65th, 70th and 75th percentiles on over 30 subsequent procedure segments and mixtures of different procedure types. We can take advantage of this knowledge and graphically analyze the breakdown of slopes for subsequent procedure segments. Segments at each scheduling percentile were categorized as increasing, decreasing or neutral slopes and plotted as follows:

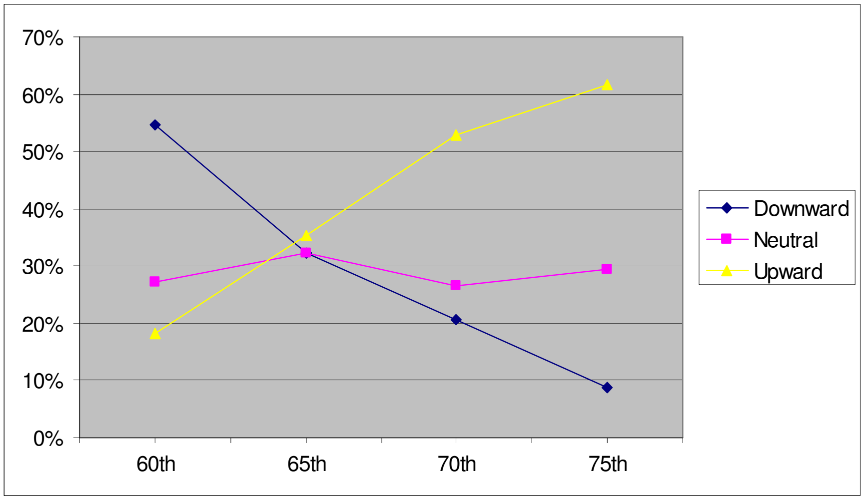

FIGURE 7
Percentage of Upward, Downward, or Neutral Trend Lines by Percentile

The graph shown in Figure 6 reveals the 60th – 65th percentile range as the optimum schedule duration percentile.

Accuracy Improvement Analysis

Another way in which we can establish the optimum schedule duration is to compare the improvement in accuracy from the currently scheduled durations for the last subsequent procedure of the day (which maximizes cumulative effects).

For each schedule duration percentile (60th, 65th, 70th and 75th), the future state (revised) schedule time was calculated and compared to the current state schedule time (for the time period analyzed). The graph provided in Figure 7 shows the peak average accuracy improvement at the 60th – 65th percentile. This value represents the average number of minutes per procedure the revised (future state) time is more closely scheduled to the actual time.

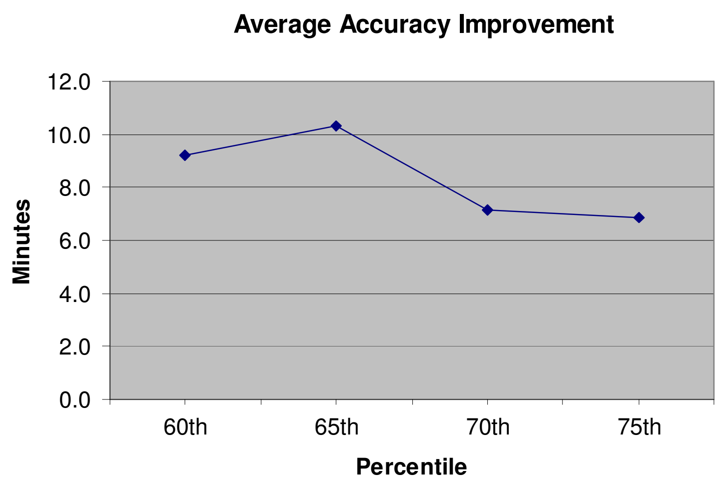

FIGURE 8
Average Accuracy Improvement Comparing Current vs. Future State Schedule

Probability Centering

Now that we've got a range of percentiles, a final consideration in optimizing schedule durations is to establish the probability that the last subsequent procedure will begin no earlier than 20 minutes prior to the schedule time, and no later than 10 minutes after the schedule time.

Table 2 compares the probabilities of the various percentile durations. The optimum probability centering is show to be at the 60th percentile. *The practical application of this business decision (accepting a downward trend in subsequent procedure probability) and centering of the probability curve assures us we aren't stretching procedures out till later in the day. This errs on the side of ensuring physician and staff wait time is avoided at the (lesser) expense of patient wait time.*

Percentile	Future State Probability (< 20 min)	Future State Probability (> 10 min)
60th	25.4%	65.2%
65th	37.3%	75.7%
70th	43.3%	80.1%
75th	50.9%	84.9%

TABLE 2
Average Future State Probability Centering by Percentile

Implementation & Control

Software scripts automatically provide event schedule time recommendations based on procedure code. Beta testing compared to the spreadsheet/model estimates to ensure implementation accuracy. Ongoing metrics based on the results (listed below) and implementation of delay codes was also tracked to ensure continuous improvement.

Results

Using the 60th percentile schedule duration, the revised schedule was 'superimposed' over actual procedure times over a three week period of time. The results of this analysis revealed the following.

- The probability of on-time starts was improved from 26% to 65%.
- A slight increase (9.5%) in probability of the last subsequent procedure starting more than 20 minutes early is also predicted. The analysis shows the probability is centered within these (-20, +10) targets such that staff & physician wait time is avoided.
- The cumulative reduction in patient wait time in the Surgery Center was estimated to be approximately 15 hours/week (97 patients affected). Also, Surgery Center staff would not be needed to care for patients waiting for this period of time.
- Physician wait time was calculated for situations where one physician is scheduled after another (ie. non back-to-back procedures). The cumulative reduction in physician wait-time was calculated to be approximately 3 hours/week (11 instances of physicians waiting for another physician to finish).
- Cumulative improvement in physician end-of-day accuracy was determined to be approximately 4 hours/week (based on 27 end-of-day procedures).
- This analysis is based on 27 subsequent procedure OR days and a total of 133 individual procedures thereby establishing a statistically significant sample size.
- Last subsequent procedures were under-scheduled by an average of 45 minutes per procedure (current state) vs. 28 minutes per procedure with the revised schedule (a 38% improvement).
- Further gains were identified by reducing the scheduling interval from 10 minutes to 5 minutes (resulting in further discrimination and improved accuracy). Scheduling software limitations prevented this, however.

Summary

The unique approach applied herein shows the importance of collecting accurate historical data, and effectiveness in using a statistics and simulations to implement scheduling estimates as accurate as possible.