**Chapter 10**

**Causal Control Charts**

# [H1] Learning Objectives

**[INSERT NL]**

1. Define the difference between association and causation
2. List causal assumptions
3. Simulate data so alternative explanations are ruled out
4. Create causal control chart
**[END NL]**

# [H1] Key Concepts

**[INSERT BL]**

* Causal reasoning
* Comprehensive list of explanations for data
* Association among data
* Sequence between events
* Counterfactual
* Mediation
* Simulated data
* Switch method
* Balancing data
* Strata

**[END BL]**

# [H1] Chapter at a Glance

At its core, performance improvement requires causal inference. The cause of poor performance needs to be identified and remedied. Remedying variables that are only associated with—not the cause of—better outcomes would create change but not improvement. Control charts were designed to help improvement teams focus on the causes of adverse outcomes. Analysts will deem a special cause to exist if the observed event is outside of three standard deviations of historical or risk-adjusted patterns (Amin 2001). All other events are attributed to random variation in the underlying processes. Despite clear causal interpretation of the control chart, these methods do not use causal analysis and therefore could be misleading. This chapter demonstrates how to construct control charts.

# [H1] Assumptions of Causal Claims

|  |
| --- |
| **[INSERT BOX]**Five Assumptions of Causality1. *Association.* Causes are associated with effects.
2. *Sequence*. Causes must precede effects.
3. *Mechanism*. Causes must have a mechanism that leads to the effect.
4. *Counterfactual*. Effects should not happen if causes are not present.
5. *Comprehensiveness*. All relevant causes must be examined.

**[END BOX]** |

Several assumptions are necessary for making a causal claim (Pearl 2009; see “Five Assumptions of Causality” for a summary). The first and obvious assumption is that the cause should lead to the effect—alternately stated, there is an *association* between cause and effect. The expectation is that one would see a change in the process after an intervention.

The second assumption is that causes must occur before effects. Because a control chart is based on change over time, it is relatively easy to see that the effect of intervention follows the intervention, instead of preceding it.

The third assumption is that there should be a clear mechanism that connects causes to effects. In most control charts, the mechanism is not shown, but implied.

The fourth assumption is that the causal impact of an intervention is calculated by comparing the effect when the cause is present to the effect when the cause is absent. In observation data used typically for constructing control charts, one can never be sure what would have happened if a particular intervention was not made. The post-intervention effects most always reflect the effect of the cause, and it is difficult to assess what would have happened if the intervention had not been made. Because these situations cannot be observed directly, they are referred to as *counterfactual*; this fourth assumption is known as the *counterfactual assumption*.

Finally, a fifth assumption has to be made that all relevant causes are measured and available. Existing approaches to control charts verify the association, sequence, and perhaps mechanism assumptions, but not the counterfactual assumption. The causal control chart directly tests the appropriateness of three of the assumptions: association, sequence, and counterfactual. The mechanism assumption is left to the imagination of the reader, and the assumption that all relevant causes are measured and available is not tested—it is assumed that over time, as more information about the factors that affect the process become available, causal control charts become increasingly accurate.

In any causal analysis, an assumption is made that all relevant variables are measured. In this chapter, we assume that all relevant differences among cases and controls are measured as covariates. This is an important assumption that is not testable in the data. We can never be sure that all relevant variables have been measured and are available. One way is to include all known causes. This means that the improvement team should examine the published literature and talk to experts to make sure that all known causes are included.

Like all scientific inquiries, every analysis is suspect. No analysis is complete, and some future investigator could do a better job of including all relevant variables. Similarly, causal analyses are suspect until additional variables that could change the findings are specified. This lack of completeness does not mean that we should avoid causal analysis. Improvement, like science, is a cumulative and iterative effort, and over time, the analyst will get better at including all relevant variables. Each investigator adds a new set of variables and continues to build on the work of others. After each investigation, a more complete set of causes is investigated.

# [H1] Attributable Risk

Managers often compare the risk of different events occurring. To do so, they need to calculate the hazard rates associated with different events. A *hazard rate* provides the conditional probability of an adverse event occurring, if it has not occurred to date. It can be calculated from probability density and cumulative distribution functions. Recall the definition of a *probability density function* (introduced in chapter 8): the probability that the event of interest will occur at period *t*, calculated as

**[INSERT EQUATION]**

.

**[END EQUATION]**

The *cumulative distribution function* is the probability that the event would occur prior to or at time *t*:

**[INSERT EQUATION]**

.

**[END EQUATION]**

To calculate the hazard rate, analysts introduce the survival function. This is the probability of surviving beyond period *t*. The survival function is the complement of the cumulative distribution function and can be calculated as

**[INSERT EQUATION]**

.

**[END EQUATION]**

Now, we are ready to calculate the hazard rate. The hazard rate, *h*, is calculated as the ratio of the probability density function divided by the survival function, so

**[INSERT EQUATION]**

.

**[END EQUATION]**

When the sentinel event is rare, the survival function is near 1, and the hazard rate and probability distribution function are nearly equal.

### [H1] Example: Fires in the Operating Room

An example will demonstrate the calculation of a hazard rate from probability density functions (see exhibit 10.1 for the calculations). Suppose fires will occur in a hospital’s surgical units at the rate of 0.10 (i.e., there is a fire in 10 percent of surgeries) for the first two years, but because of planned changes in oxygen equipment supplied to surgical suites, the rate will drop to 0.01 thereafter. Note that in this example, failure is defined as a fire and survival is not having a fire. The probability of failure any time during the five years does add up to 1.0, meaning that at least one fire will occur in the next five years. Given this distribution, we want to understand the probability that if fire had not broken out for last two years, it would occur in the third year. More briefly, if we are in year 2 with no history of fire, what is the hazard associated with fire next year?

**[INSERT EXHIBIT]**

**Exhibit 10.1** Calculating Hazard Associated with Fire in Surgical Rooms

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| *Year* | *Probability Density* | *Cumulative Distribution* | *Survival* | *Hazard* |
| 1 | 0.10 | 0.10 | 1.00 | 0.10 |
| 2 | 0.10 | 0.20 | 0.90 | 0.11 |
| 3 | 0.01 | 0.21 | 0.80 | 0.01 |
| 4 | 0.01 | 0.22 | 0.79 | 0.01 |
| 5 | 0.01 | 0.23 | 0.78 | 0.01 |
| 5+ | 0.77 | 1.00 | 0.77 | 1.00 |

**[END EXHIBIT]**

The survival function is 1 minus the prior year’s cumulative distribution. The probability of surviving at the start of year 1 is 1. The probability of surviving until the start of year 2, in this example, is .90 (or 90 percent), as 10 percent of surgeries will have had fire in the prior year. The probability of no fire until the start of year 3 is 80 percent, as 20 percent will have had fire in the first two years while the organization was waiting for better equipment. Note that the survival probability decreases with time.

Now we can calculate the hazard rate. This is the probability of the event occurring this year if it has not occurred in the prior years. This is calculated as the ratio of the probability density function and the survival function:

**[INSERT EQUATION]**

**[END EQUATION]**

Note that the hazard rate drops in year 3, the year we expect to have better equipment. If we have survived up to this point without a fire, we now face the lower probability of fire associated with better equipment. The hazard rate for the first two years is more than 10 percent but afterward drops to 1 percent.

If we assume that various causes are independent of each other, the hazard rate from all sources can be calculated as the sum of the hazard rate from each source. This is a very helpful concept because it allows us to calculate the total impact of multiple causes. If *H* symbolizes the combined hazard rate and *h*i shows the hazard associated with cause *i*, then the hazard rate of the combination can be calculated as

**[INSERT EQUATION]**

.

**[END EQUATION]**

For example, under the assumption of independence, the hazard rate for a medication error could be the sum of hazards resulting from a fatigued nurse and the hazard associated with a physician’s poorly written prescription note, written as

**[INSERT EQUATION]**

**[END EQUATION]**

The combined hazard risk can also be used to calculate the attributable risk related to a specific cause. The attributable risk (AR) related to a cause is calculated as the ratio of its source- specific hazard rate and the hazard rate from all sources in this way:

**[INSERT EQUATION]**

.

**[END EQUATION]**

An example can help demonstrate the concept of attributable risk. If the hazard rate for a medication error caused by a fatigued nurse is 1 in 1,000, the hazard rate for a medication error caused by an illegible prescription order is 2 in 1,000, and these two are all the causes of a medication error, then the risk attributable to the nurse is

**[INSERT EQUATION]**

.

**[END EQUATION]**

In this case, the ratio of the nurse’s hazard rate and the total hazard rate is 33 percent. In contrast, 66 percent of the medication error risk is attributable to the physician’s poor prescription-writing skills. The most probable cause of the medication error is the physician’s illegible writing. In this fashion, relative attribution can be made to separate risks.

Correct attribution, of course, is important to finding the right solutions. If a hospital tries to solve medication errors caused by fatigue, but not by poor writing, it will not be solving its most important problem. Of course, the best action is to address all causes, but if leaders cannot and must set a priority, the analysis shows how to focus on key causes that are more responsible for the observed sentinel events.

# [H1] Causal Analysis in the Context of Control Charts

Two types of variation are detected by traditional control charts. The first type of variation is the result of random variation. *Random variation* is natural variation that is internal to the process. The second type of variation has a special or assignable cause. *Special* or *assignable* causes are variances in outcomes of the process that can be traced to a source but are not part of the process (Bass 2007). For example, one can examine whether hospital losses result from Congressional changes in hospital reimbursement. For another example, one can examine whether increased waiting times in emergency rooms result from natural or human-made disasters. Assignable causes arise from factors external to the system that are not part of the normal process flow.

The key difference between a causal control chart and a traditional control chart is the construction of the counterfactual group. A counterfactual group is an artificially constructed group of patients who have the same features as the observed cases except for the presence of the known special or assignable causes. In the counterfactual group, the analyst considers what might be observed if known causes were not present. This process is speculative, but it is an organized speculation wherein the effects of known causes are removed.

In recent years, a number of authors have proposed new ways to risk-adjust control charts—in principle, creating counterfactual groups (Alemi and Oliver 2001; Alemi and Sullivan 2001; Cockings, Cook, and Iqbal 2006; Cook, Coory, and Webster 2011; Cook et al. 2008; Hart et al. 2004). The method used to create a counterfactual group relies on data balancing. In data balancing, the intervention (cases) and nonintervention group (controls) are weighted so that they do not differ in their risks (also referred to as *covariates*). One way to balance the data is through propensity scoring (where the intervention is predicted from various covariates). Another way is through stratification (where the impact of intervention is examined within combinations, or strata, of covariates).

If patients are stratified based on the covariates, then stratification controls the influence of known causes, and data in stratified groups provide estimates for the impact of the intervention independent of the covariates. In short, stratification can be used to remove the effects of known causes and speculate about what would have occurred if it were not for these known causes. This chapter focuses on stratification methods for balancing data when constructing control charts.

# [H1] Methods

The counterfactual model for causal analysis of observational data can be traced to a series of articles by statistician Donald B. Rubin (1974, 1977, 1978). It also has roots in econometric models (Heckman 1979, 2008), probability models (Pearl 2009), and philosophy (Paul, Hall, and Collins 2004). The core concept behind this method is to construct artificially a control group that would resemble the pattern of known causes in the intervention cases. Then, the comparison of the cases and the counterfactual control group can provide an estimate of the causal effect of the intervention independent of known causes.

In causal control charts, the same principals are followed. In these charts, cases are the observations that have the intervention. Controls are observations made before the intervention. Counterfactual controls are adjusted controls if they had the same rate of known causes as in the cases. Control limits are derived by identifying the 95 percent or 99 percent values in the counterfactual group. In this fashion, a causal control chart compares cases with the intervention to simulated controls without the intervention but with the same set of known causes. Except for the intervention, the controls are similar to cases in other measured aspects.

Suppose that we are interested in the impact of intervention *x* on outcome *y*. Follow this topic by examining the data in exhibit 10.2. For simplicity, assume that we have two outcomes: *y* = 1 or *y* = 0. Any time the intervention is present, we call it a case (*x* = 1); any time the intervention is absent, we call it a control (*x* = 0). In some charts, controls precede cases (i.e., controls occur prior to the intervention and cases follow the change). The data for both cases and controls are divided into *k* strata, and each stratum represents a combination of known causes of the outcome, which we will refer to as covariates. These causes co-occur with the intervention (*x*), and thus their impact on outcome is confounded. The purpose of the analysis is to remove the confounding through balancing the data, to display the relationship between *x* and *y* visually, and to calculate the unconfounded impact of *x* on *y*.

**[INSET EXHIBIT]**

**Exhibit 10.2** Observations of Cases and Controls over Time

|  |  |  |
| --- | --- | --- |
| *Strata of Known Causes* | *Outcome y = 1* | *Outcome y = 0* |
| *Cases x = 1* | *Controls x = 0* | *Cases x = 1* | *Controls x = 0* |
| 1 | *a*1*t* | *c*1*t* | *b*1*t* | *d*1*t* |
| 2 | *a*2*t* | *c*2*t* | *b*2*t* | *d*2*t* |
| … | … | … | … | … |
| *i* | *ait* | *cit* | *bit* | *dit* |
| … | … | … | … | … |
| *k* | *akt* | *ckt* | *bkt* | *dkt* |

*Note*: and cases and controls do not occur at same time (i.e., at time *t*). Either or .

**[END EXHIBIT]**

To balance the data, Alemi, ElRafey, and Avramovic (2016) recommend weighing controls so that the rate of known causes among the intervention and nonintervention group is the same at any point in time. A set of weights that ensures known causes occur in controls at the same rate as in cases at time *t* is given as

**[INSERT EQUATION]**

**[END EQUATION]**

In these formulas, the index *i* indicates the strata composed of a combination of known causes. The index *t* indicates the period. In other words, cases continue to be weighted as before, but controls are weighted so that in each stratum, they occur at the same rate as the cases. These weights provide a way of removing the effect of the known causes. After weights are applied, the only variation that remains is the difference in the intervention and nonintervention groups.

The upper control limit (UCL) and lower control limit (LCL) are estimated using the data from weighted controls and the equation

**[INSERT EQUATION]**

**[END EQUATION]**

In this equation, is the average probability of excessive boarding among the weighted controls.

## **[H1] A Simulated Example of Emergency Department Delays**

To demonstrate the procedure, this chapter shows how it can be applied to analysis of excess boarding times at an emergency department (ED). *Excess boarding time* refers to patients waiting for the next step in their care (e.g., admission to the hospital or discharge to home) in excess of six hours from arrival at the ED. Exhibit 10.3 shows hypothetical admissions data. These data were adapted from boarding experiences reported elsewhere (Kheirbek et al. 2015).

Our hospital’s excessive boarding time may result from multiple causes, including backups at the imaging department or unavailability of hospital beds. These were known causes of delay. To reduce boarding time, the hospital hired additional staff. Hiring another person allows the ED to work more efficiently, but if the organization is experiencing high hospital occupancy or if imaging is backed up, adding more staff may not reduce the backup problem.

One question to answer is whether the new hire has, in fact, led to reduced boarding time. Before we can answer this question, we need to remove the effect of the other two known causes. This is done by dividing the data into strata, calculating the frequency of cases and controls in each stratum, and weighting the controls so they occur at the same rate as cases. The first step in such an analysis is to construct four strata from the possible combinations of the alternative explanations:

**[INSERT NL]**

1. Neither imaging backup nor high occupancy
2. Only imaging backup
3. Only high occupancy
4. Both imaging backup and high occupancy
**[END NL]**

These strata are used to report the number of patients and the number with excessive boarding time (see exhibit 10.3). The first 13 months are pre-intervention, and these data are used as controls for the post-intervention cases. For example, we see that in month 1 through month 13, when there was imaging backup, 7 out of 17 patients had excessive boarding. In contrast, in the first month after hiring the new employee (i.e., month 14), 1 out of 4 patients had excessive boarding when there was also an imaging backup.

**[INSERT EXHIBIT]**

 **Exhibit 10.3** Number of Patients with Excessive Boarding (Total Number of Patients)

|  |  |  |
| --- | --- | --- |
|   | *Average in Prehiring Months 1–13: Controls* | *Post-Hiring: Cases* |
| *Month 14* | *Month 15* | *Month 16* | *Month 17* | *Month 18* | *Month 19* | *Month 20* | *Month 21* | *Month 22* | *Month 23* | *Month 24* | *Month 25* |
| Neither backups | 0.92 (12.69) | 2 (10) | 0 (9) | 1 (23) | 0 (29) | 1 (10) | 1 (24) | 0 (6) | 1 (25) | 0 (2) | 1 (26) | 1 (27) | 0 (12) |
| Imaging backup | 6.92 (17.08) | 1 (4) | 3 (5) | 3 (8) | 10 (13) | 0 (1) | 5 (6) | 7 (23) | 5 (18) | 14 (15) | 7 (19) | 9 (14) | 1 (2) |
| High occupancy | 2.85 (6.69) | 2 (6) | 7 (15) | 1 (2) | 0 (2) | 5 (17) | 6 (13) | 7 (21) | 7 (9) | 4 (24) | 5 (17) | 11 (35) | 8 (15) |
| Both backups | 11 (13.92) | 1 (9) | 12 (25) | 4 (10) | 7 (10) | 8 (17) | 9 (18) | 5 (22) | 6 (10) | 6 (27) | 8 (19) | 9 (16) | 4 (15) |

**[END EXHIBIT]**

The calculation of control limits for the causal control charts occurs in several steps. In step 1, the average number of patients with known causes in pre-intervention periods is calculated. In exhibit 10.3, these averages are shown in the first column. In step 2, weights are calculated to normalize the total number of cases and controls. For example, weights for imaging backup in month 14 are calculated as

**[INSERT EQUATION]**

.

**[END EQUATION]**

In step 3, weights are applied to calculate the weighted number of controls and weighted number of patients with excessive boarding. The weighted number of controls is calculated as

**[INSERT EQUATION]**

**[END EQUATION]**

For example, the weighted number of controls for month 14 is calculated as 0.79 × 12.69 + 0.23 × 17.08 + 0.90 × 6.69 + 0.65 × 13.92 = 29. Note that the weighted number of controls in month 14 is the same as the total number of cases in month 14. This is by design—guaranteed by the way we composed the weights. The weights are also applied to the number of patients with excessive boarding in controls. This is done as follows:

**[INSERT EQUATION]**

**[END EQUATION]**

For example, using weights for month 14, the number of control patients with excessive boarding is calculated as

**[INSERT EQUATION]**

0.790.92 + 0.236.92 + 0.902.85 + 0.6511.00 = 12.01.

**[END EQUATION]**

In the third and final step, the proportion of weighted controls with excessive boarding is used to estimate the UCL and LCL. For example, the proportion of weighted controls with excessive boarding can be found with the formula

**[INSERT EQUATION]**

**[END EQUATION]**

In month 14, as the equation is

**[INSERT EQUATION]**

**[END EQUATION]**

Using this proportion, we can calculate the UCL and LCL for month 14 as

**[INSERT EQUATION]**

**[END EQUATION]**

The completed set of calculations is shown in exhibit 10.4, although for ease of viewing some of the values are rounded. As mentioned earlier, the weighting procedure ensures that each of the strata occur at the same rate in the post-intervention periods. Therefore, the interpretation of these procedures is that the effects of known causes have been removed through the weighting procedure; what is shown in the control chart is the impact of intervention when known causes occur at the same rates as in the pre- and post-intervention periods.

**[INSERT EXHIBIT]**

**Exhibit 10.4** Calculation of Weighted Control Limits

|  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | *Month 14* | *Month 15* | *Month 16* | *Month 17* | *Month 18* | *Month 19* | *Month 20* | *Month 21* | *Month 22* | *Month 23* | *Month 24* | *Month 25* |
| Observed rates | Neither backups | 0.20 | 0.00 | 0.04 | 0.00 | 0.10 | 0.04 | 0.00 | 0.04 | 0.00 | 0.04 | 0.04 | 0.00 |
| Imaging backup | 0.25 | 0.60 | 0.38 | 0.77 | 0.00 | 0.83 | 0.30 | 0.28 | 0.93 | 0.37 | 0.64 | 0.50 |
| High occupancy | 0.33 | 0.47 | 0.50 | 0.00 | 0.29 | 0.46 | 0.33 | 0.78 | 0.17 | 0.29 | 0.31 | 0.53 |
| Both backups | 0.11 | 0.48 | 0.40 | 0.70 | 0.47 | 0.50 | 0.23 | 0.60 | 0.22 | 0.42 | 0.56 | 0.27 |
| Observed excessive boarding | 0.21 | 0.41 | 0.21 | 0.31 | 0.31 | 0.34 | 0.26 | 0.31 | 0.35 | 0.26 | 0.33 | 0.30 |
| Weights | Neither backups | 0.79 | 0.71 | 1.81 | 2.28 | 0.79 | 1.89 | 0.47 | 1.97 | 0.16 | 2.05 | 2.13 | 0.95 |
| Imaging backup | 0.23 | 0.29 | 0.47 | 0.76 | 0.06 | 0.35 | 1.35 | 1.05 | 0.88 | 1.11 | 0.82 | 0.12 |
| High occupancy | 0.90 | 2.24 | 0.30 | 0.30 | 2.54 | 1.94 | 3.14 | 1.34 | 3.59 | 2.54 | 5.23 | 2.24 |
| Both backups | 0.65 | 1.80 | 0.72 | 0.72 | 1.22 | 1.29 | 1.58 | 0.72 | 1.94 | 1.36 | 1.15 | 1.08 |
| Weighted control limits | Weighted rate | 0.41 | 0.53 | 0.32 | 0.30 | 0.48 | 0.39 | 0.50 | 0.34 | 0.56 | 0.39 | 0.38 | 0.45 |
| Weighted upper control limit | 0.69 | 0.74 | 0.53 | 0.49 | 0.71 | 0.58 | 0.68 | 0.52 | 0.74 | 0.56 | 0.53 | 0.68 |
| Weighted lower control limit | 0.14 | 0.33 | 0.10 | 0.11 | 0.26 | 0.20 | 0.32 | 0.16 | 0.37 | 0.23 | 0.23 | 0.23 |

**[END EXHIBIT]**

 A counterfactual control chart can be displayed in similar fashion to any chart: on the *x*‑axis, time is given; the *y*-axis displays the outcome of interest. The weighted UCL and LCL are drawn so that observations outside the control limits have a low probability of occurring—typically less than 5 or 1 percent. Exhibit 10.5 shows an example of how a causal control chart is displayed. In this exhibit, we see the effect of hiring on boarding time while controlling for differences in imaging and hospital occupancy backups. The balanced (weighted) UCL and LCL are shown in solid red lines. The observed rate of excessive boarding after hiring is shown with markers showing the rate in different months. If we consider the weighted control limits, the reduction in excessive boarding time occurred only in month 20 and in no other months. These data say that after controlling for imaging and hospital backups, the hiring of the clinician had a temporary effect. In short, we now know that the hiring did not change much.

**[INSERT EXHIBIT; Make upper red line dashes and label it UCL. Make lower red line dots and label it LCL. Make blue line solid black, and make the “squares” that mark the data points diamonds (per other graphs). Label the solid black line “Observed excessive boarding.” Make axis labels and numbers rom, not bold. Delete legend.]**

**Exhibit 10.5** Causal Control Chart for Impact of Hiring New Staff on Excessive Boarding

*Note*: Imaging and hospital bed backup occur at the same rate before and after the hiring.

## **[END EXHIBIT]**

## **[H1] Application to Stock Market Prices**

 Alemi and Coffin (2018) applied causal control charts to the analysis of stock market prices. They examined the impact of the 2016 election of Donald Trump on stock prices for Humana and other health insurance companies. Humana was one of the participants in Barack Obama’s healthcare reform. During the campaign, Trump had promised to repeal and replace the reforms. Therefore, his election should have led to changes in evaluation of the stock. Because stock prices are affected by a host of variables, we needed to remove the effect of other known causes (e.g., general change in economy or changes in the healthcare index funds). At the time, stock prices were rising, and attributing the rise to the election might have been an error. We used causal control charts to remove the effect of the known causes and examine the impact of the election in isolation. We defined known causes based on two index funds: one measuring general stock prices and the other measuring changes in healthcare stock prices. These two index funds measured the historical trend in stock prices. We balanced the data so the effects of these known causes were removed. Next, we examined the effect of the election on the prices of the stock of the insurance companies, including Humana. The analysis allowed these investigators to see the impact of Trump’s victory, independent of trends in general and healthcare-specific stock prices.

#  [H1] Summary

This chapter has laid out how competing causes can be statistically controlled for, allowing the effect of intervention on outcome of care to be isolated.

# [H1] Supplemental Resources

A problem set, solutions to problems, multimedia presentations, SQL code, and other related materials are on the course website.

# [H1] References

Alemi, F., and T. Coffin. 2018. “Causal Control Charts: Application to Assessing Impact of Trump’s Election on Insurance Stock Prices.” *Transaction on Business and Economics* 15: 259–72.

Alemi, F., A. ElRafey, and I. Avramovic. 2016. “Covariate Balancing Through Naturally Occurring Strata.” *Health Services Research* 53 (1).

Alemi, F., and D. Oliver. 2001. “Tutorial on Risk-Adjusted P-charts.” *Quality Management in Healthcare* 10 (1): 1–9.

Alemi, F., and T. Sullivan. 2001. “Tutorial on Risk Adjusted X-bar Charts: Applications to Measurement of Diabetes Control.” *Quality Management in Healthcare* 9 (3): 57–65.

Amin, S. G. 2001. “Control Charts 101: A Guide to Health Care Applications.” *Quality Management in Healthcare* 9 (3): 1–27.

Bass, I. 2007. *Six Sigma Statistics with Excel and Minitab*. New York: McGraw-Hill Professional.

Cockings, J., D. Cook, and R. Iqbal. 2006. “Process Monitoring in Intensive Care with the Use of Cumulative Expected minus Observed Mortality and Risk-Adjusted P Charts.” *Critical Care* 10 (1): R28.

Cook, D., M. Coory, and R. Webster. 2011. “Exponentially Weighted Moving Average Charts to Compare Observed and Expected Values for Monitoring Risk-Adjusted Hospital Indicators.” *BMJ Quality and Safety* 20 (6): 469–74.

Cook, D. A., G. Duke, G. K. Hart, D. Pilcher, and D. Mullany. 2008. “Review of the Application of Risk-Adjusted Charts to Analyze Mortality Outcomes in Critical Care.” *Critical Care and Resuscitation* 10 (3): 239–51.

Hart, M. K., J. W. Robertson, R. F. Hart, and K. Y. Lee. 2004. “Application of Variables Control Charts to Risk-Adjusted Time-Ordered Healthcare Data.” *Quality Management in Healthcare* (13): 2: 99–119.

Heckman, J. J. 2008. “Econometric Causality.” *International Statistical Review* 76 (1): 1–27.

———. 1979. “Sample Selection Bias as a Specification Error.” *Econometrica* 47 (1): 153–61.

Kheirbek, R. E., S. Beygi, M. Zargoush, F. Alemi, A. W. Smith, R. D. Fletcher, P. N. Seton, and B. A. Hawkins. 2015. “Causal Analysis of Emergency Department Delays.” *Quality Management in Healthcare* 24 (3): 162–66.

Paul, L. A., E. J. Hall, and J. D. Collins. 2004. *Causation and Counterfactuals*. Cambridge, Mass.: MIT Press.

Pearl, J. 2009. “Causal Inference in Statistics: An Overview.” *Statistical Surveys* 3: 96–146.

Rubin, D. B. 1978. “Bayesian Inference for Causal Effects: The Role of Randomization.” *The Annals of Statistics* 6 (1): 34–58.

———. 1977. “Assignment to Treatment Group on the Basis of a Covariate.” *Journal of Educational Statistics* 2 (1): 1–26.

———. 1974. “Estimating Causal Effects of Treatments in Randomized and Nonrandomized Studies.” *Journal of Educational Psychology* 66 (5): 688–701.