Review of Causal Inference and Modeling[1]

Mehmet Altinpinar (19ma57@queensu.ca)
Julian M. Ortiz (julian.ortiz@queensu.ca)

Abstract

A cause and effect relationship is one of the universal constant realities and it can be observed in all aspects of life including physics, biology, medicine, psychology, engineering, management, law, statistics, economics etc. Fundamentally, everything happens for a reason in the universe. In more technical terms, an action, phenomenon, or condition is given rise by a “cause” and such cause can have different effects that can be traceable by establishing causal relationships. Causal inference is a comprehensive process of analyzing the causal relationship between a cause and effect by considering a change in the cause to draw conclusions about the alternative effects. Over the last few decades, many researchers have carried out studies on causal inference for different fields such as epidemiology, computer sciences, statistics and economics and social sciences, but no study has been conducted for the mining industry, yet. Causal relationships can be established for several different facets of the mining industry including exploration, development, production, mineral processing, etc.

In this paper, a review of Causal Inference is given in general terms and a preliminary study is performed to understand the causal relationship behind molybdenum recovery by considering the influencing factors, the causes.

1. Introduction

By definition, causality, which is also referred to as “causation” or “cause and effect”, means the connection between two events or states such that one produces or brings about the other; where one is the cause and the other its effect. This basic definition of causality sounds simple, but when it comes to the detail, this definition is just like a droplet in the entire ocean. A great number of papers have been published in the last couple of decades on different topics by utilizing the concept of causality. Plato was probably the first philosopher that mentioned causality, and he said in Timaeus 28a that “everything that becomes or changes must do so owing to some cause; for nothing can come to be without a cause”.

In 1986, Paul Holland published a seminal paper with the title of Statistics and Causal Inference. Our understanding of causal inference has since increased several folds, due primarily to advances in three areas:

1. Nonparametric structural equations.
2. Graphical models.
3. Symbiosis between counterfactual and graphical methods (Pearl, 2003).

[1] Cite as: Altinpinar M, Ortiz JM (2020) Review of Causal Inference and Modeling, Predictive Geometallurgy and Geostatistics Lab, Queen’s University, Annual Report 2020, paper 2020-08, 130-146.

In this paper Causal Inference is discussed in general terms by including basic principles, graphical models, difference between correlation and causation, counterfactuals, and Statistical Inference and Causal Inference are compared. Finally, a causal relationship is developed for a simple example in the mining industry.

2. Causal Inference

How do we know when we have found a "cause" of something? In order to answer this question, we need to focus on understanding what "causality" actually is. Basically, causality is the claim that "if something is changed, some specific outcome is going to happen". There are two basic types of cause and effect: Simple and Complex.

In Simple Cause and Effect, an action directly causes one immediate result. For example, when you press the brake pedal of your car, you "cause" it to slow down and eventually stop. If you hadn't had pressed the brake pedal, your car would have kept going.

In the Complex Cause and Effect, there are multiple causes over time with delayed and unclear effects. For example, an investment decision you make today will affect how you will live in your retirement.

Causality deals with when a policy is manipulated, what's going to happen to outcomes. With causal inference, one can accurately find out how changes in policy create changes in real world outcomes.

2.1. Basic Definitions

A variable in causal inference is a characteristic of the "Unit of Analysis" in the dataset and unit of analysis is where the analyses are being performed, such as:

- Countries
- City blocks
- People
- Aqueous environment
- Mineral deposits
- etc.

A population is the collected set of all the "units of analysis" or in other words, the collection of all these units is called the Population, such as:

- Population of countries
- Population of city blocks
- Population of People
- Fish species in Lake Ontario
- Precious metal content in a mineral deposit
- etc.

There are two main types of variables to be measured within the unit of analysis, namely: policy/treatment variable, i.e. the parent, and outcome variable, i.e. the child. Policy variable is the characteristic that will be used to change the outcome variable and outcome variable is the characteristic to be affected/changed. For example:

1. If the outcome variable is the level of prosperity in your retirement, then the policy/treatment variable might be whether you do more investment today or not.
2. If the outcome variable is the number of deaths in traffic accidents in a country, policy/treatment variable might be a measure of how strong the traffic rules are.
3. If the outcome variable is the metal recovery in a mineral processing plant, policy/treatment variable might be, *inter alia*, metal content of the run of mine ore fed to the processing plant.

It is important to figure out how to measure these characteristics. It is relatively easy to measure when the variables are quantitative. For example, in the third item above, both the policy and outcome variables can be quantified by using appropriate assessment methods and some judgments can be made based on these quantifications. When qualitative variables are in consideration, however, a method should be proposed to measure these kinds of variables. For example, in the first item above, the policy variable is quantitative as one could quantify the amount of investment by assessing it in terms of its unit, e.g. currency, immovable property, retirement plan, etc. On the other hand, the outcome variable is a qualitative variable and a method should be developed to measure it. Prosperity is a subjective concept and therefore, it changes from person to person. It is still possible to define different levels to describe prosperity based on the results of a survey to be conducted by involving retired people.

2.2. Describing Data

After measuring variables (policy and outcome variables), describing data is the first step of data analysis and the followings are the starting points:

- What is the average value of a variable? What are the largest and smallest values?
    - For an outcome variable, for example metal recovery, we want to know what's the average, the largest and smallest recovery values obtained during the entire mineral processing.
    - We also want to learn such values, i.e. the minimum, maximum, average values for a policy variable, for example the metal content of the run of mine ore fed to the processing plant.

Causality refers to how an intervention to change the policy variable affects the outcome variable. Variation in the policy variable in the data must be observed to learn about causality. Values of policy variable cannot stay constant across all units, which means, if there is zero variation in the policy variable in the data, then it would be difficult to learn about causality because there is no change in the policy variable. Suppose that the metal content of the run of mine ore fed to the processing plant had the same value. That would mean that there is zero variation in metal content in the data and that the throughput would have the same value at the end of each shift. This would not be realistic since it is impossible to feed the ore with the same metal content during a shift. Fluctuation will certainly be observed in the metal content of the ore fed and since there will be a variation in the data, there will be a causal relationship.

Statistically, variation is measured by using two numbers: the variance and the standard deviation. The way to interpret them is that the larger numbers mean there is more variation. Learning about causality is to compare units with different values of the policy variable. For example, comparing an ore that has high metal content with the one that has low metal content will be related to variation. Consequently, no comparison can be made if there is no variation in the policy variable.

2.3. Correlation vs. Causation

So far, univariate (one variable) descriptions of the data were discussed in the previous sections. For example, for the variable metal content, average, maximum and minimum as well as variation were discussed. Multivariate Description of the data, on the other hand, shows the relationships between multiple variables. For example: metal content of the ore fed and metal recovery obtained at the end of the mineral processing can be correlated under three options as follows:

- If the metal recovery is high when the metal content is high, then metal recovery and metal content are positively correlated.
- If the metal recovery is low when the metal content is high, then metal recovery and metal content are negatively correlated.
- If there is no relationship between metal recovery and metal content, then they are said to be uncorrelated.

Things may differ when it comes to the comparison of correlation and causation. If the outcome variable and policy variable are correlated in the data, then that does not necessarily imply that the policy variable has a causal effect on outcomes. Therefore, correlation does not imply causation. Let's explain this by using the following statement: Data show that people who hurry tend to be late to their meetings. Don't hurry, or you'll be late! The data can be correct and these two variables can be positively correlated to each other but this does not mean that there is a causal relationship between hurrying up and being late. The cause behind being late must be something else like, waking up late, spending too much time for breakfast etc. Figure 1 shows the true model of this statement.

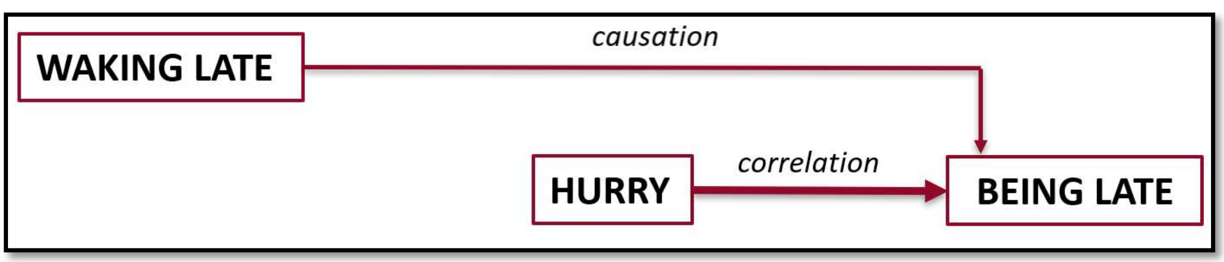


Figure 1 - Correlation vs. Causation

Though, causal relationship might still exist, but the information provided here is not sufficient for establishing a causal relationship between hurrying up and being late. On the other hand, if no correlation is observed after applying appropriate causal inference methods, then there is no causal effect. In some cases, correlation is highly suggestive of causation, in some other cases, however, one should be careful as it may not imply causality at all.

2.4. Counterfactuals

The outcome of a change in the policy variable is the one that happens when all the conditions point it out as the outcome. If something is changed in during the process, the outcome would obviously be different.

Assume that John, who just graduated from high school, makes applications to Queen's University and Western University for his post secondary education and that he gets acceptance from both universities and that he chooses to go to Queen's University. His future life, at least his professional life, will largely be dependant on this choice. When he graduates from university, he will look for a job, he will presumably find one and will pursue his professional career based on his education at Queen's University. In this example, the choice of going to Queen's University is the policy variable and John's professional career is the outcome variable. If the policy variable was the choice of going to Western University, the outcome, his professional career, would be different. The choice of going to Western University is the counterfactual in our example. He might also have applied to several more universities and in that case, all choices for those universities would be the counterfactuals of his actual choice. This means that one of the potential outcomes will become an actual outcome and the other potential outcomes will become counterfactual outcomes, which would have occurred if something different happened. If he had chosen to go to Western University, he would have found a job in London, Ontario. This kind of statement - an "if" statement in which the "if" portion is untrue or unrealized - is known as a counterfactual (Pearl, 2016) and counterfactuals are irrefutable by definition.

Causality deals with comparing actual outcomes with counterfactual outcomes. These kinds of comparisons will tell us about causal effects, and causality can be defined as the difference between actual outcomes and counterfactual outcomes. Since counterfactuals are not actually observed because they are what would have happened, causal inference methods refer to various assumptions to let us observe the unobservable - the counterfactual – outcome and in this way, an appropriate comparison can be made.

2.5. Statistical vs. Causal Inference

Statistical Inference involves the followings:

- Samples
- Hypothesis
- Confidence Intervals
- p-values
- Standard errors
- t-tests
- Statistical significance
- The normal distribution
- etc.

These are used to compute some numbers in the population and in practice it is not possible and/or feasible to observe every single individual (or sample).

Statistical inference is about computing statistics using the data we have to learn about the data that we don't have. We may not observe every single individual in a population, so we are going to compute the average for just the samples we observe and hope that it is close to the true average for the population. However, we could have gathered more data because all of the topics above are about quantifying how close we are to the truth. If we were given more data and in principle, if we had a very, very large amount of data, none of these topics would be important. In other words, when you have lots of observations, you usually don't need statistical inference.

In contrast to statistical inference, causal inference is used to learn about counterfactuals that are the data we can never observe. In causal inference, regardless of how much data we have, there is always a potential of having a problem about learning what we want to learn. A clearer perspective can be achieved by obtaining more and more data in statistical inference problems, but this will not be the case for causal inference problems and this is the main difference between statistical inference and causal inference.

Causal inference is often called identification analysis, because its goal is to reveal what we can learn and identify from the data we have. With causal inference, we are looking to identify causal effects. For this purpose:

1. The first step in causal analysis is to check whether we can learn and identify anything from the data. Is this data good enough for getting some idea about the causal inference?
2. The second step is to do some statistical inference to get some estimations about causal effects as we are not going to have the whole amount of data.
3. The final step is to quantify the estimated effects. Since we do not possess all the data we could possibly have gotten, our estimate is not going to be exactly the true causal effects but we still need to quantify the uncertainty in the estimated effects and how far we are from the reality.

3. GRAPHICAL MODELS

In Causal Inference, "causal graphs" or "directed acyclic graphs (DAG's)" are used to identify which variables to control for. They depict our assumptions about the relationship between the variables and they also help us figuring out the relationships among the variables, whether the variables are independent, dependent or conditionally dependent.

3.1. Basic Definitions About Graphs

3.1.1 Types of Graphs, Links and Nodes

There are two kinds of graphs, namely:

1. Undirected graphs, where the variables are associated but the causal direction is unknown. Figure 2 shows a typical undirected graph where X and Y are the variables (nodes) and the line between them is called as the link or edge.

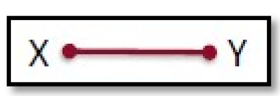


Figure 2 - An undirected graph

2. Directed graphs, where the relationship between the variables can be observed by looking at the directions of the links between them. When all the links between variables are directed, the graph is called Directed Graph. Figure 3 shows a typical directed graph where X and Y are the variables (nodes) and the line between them is again called as the link or edge. But this time the causal relationship can be seen by looking at the direction of the link between the nodes. According to the direction in Figure 3, X affects Y.

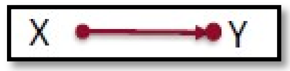


Figure 3 - A directed graph

3.1.2 Path

Figure 4 shows a path, which is a way to get from one node to another, travelling along the links. There are two paths from X to Z: X → Y → Z and X → Y → W → Z

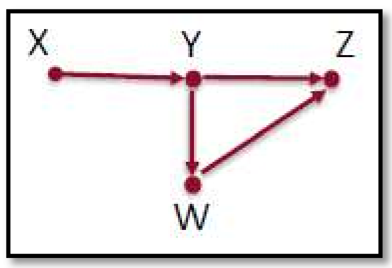


Figure 4 - Paths between the nodes

3.1.3 Directed Acyclic Graph (DAG)

If a directed graph has no undirected paths at all and if there is no cycle among the nodes, then it is called as a Directed Acyclic Graph (DAG). Figure 5 shows a typical DAG as well as the non-DAG conditions.

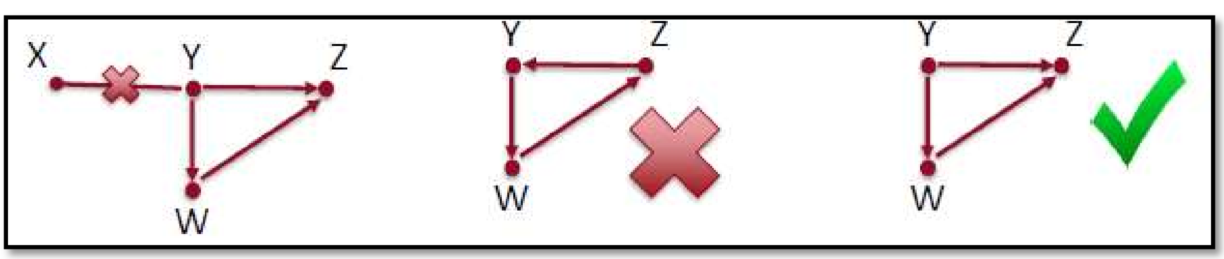


Figure 5 - DAG and non-DAG conditions

3.1.4 Parents, Children, Ancestors and Descendants

The node that a directed edge starts from is called the parent of the node that the edge goes into; conversely, the node that the edge goes into is the child of the node it comes from. If two nodes are connected by a directed path, then the first node is the ancestor of every node on the path, and every node on the path is the descendant of the first node (Pearl et. al., 2016). Followings are the example statements based on the graph shown in Figure 6.

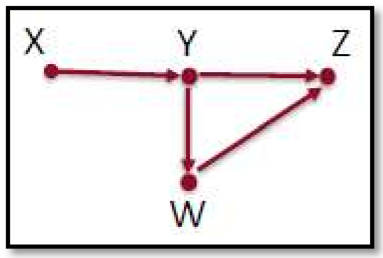


Figure 6 - Parents, Children, Ancestors and Descendants

- X is Y's parent
- W is a child of Y
- Z is a descendant of X
- X, Y and W are ancestors of Z
- Z has 2 parents, Y and W

3.2. Structural Causal Models

In order to deal rigorously with questions of causality, we must have a way of formally setting down our assumptions about the causal story behind a data set. To do so, we introduce the concept of the structural causal model, or SCM, which is a way of describing the relevant features of the world and how they interact with each other. Specifically, a structural causal model describes how nature assigns values to variables of interest (Pearl et. al., 2016).

An SCM consists of two sets of variables, "U" and "V", and a set of functions "f" that assigns each variable in "V" a value based on the values of the other variables in the model.

For the purpose of causality:

- A variable X is a direct cause of a variable Y, if it appears in the function that assigns value to Y.
- X is a cause of Y if it is either a direct or any type of cause of Y.
- The variables, U, are called exogenous variables which are external to the model and we are not dealing with how they are caused.
- The variables in V are endogenous and every endogenous variable in a model is a descendant of at least one exogenous variable.
- Exogenous variables cannot be descendants of any other variables; they have no ancestors and are represented as root nodes in graphs. (Pearl et. al., 2016).

3.3. Graphical Models and Their Applications

3.3.1 Chains

Three nodes and two edges, with one edge directed into and one edge directed out of the middle variable is called a chain. A typical chain is shown Figure 7 with the exogenous variables.

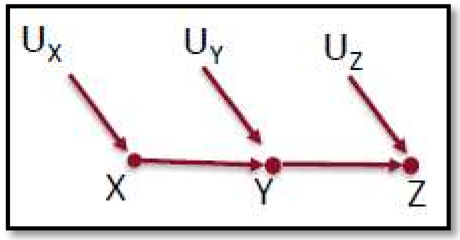


Figure 7 - Example of a chain graph

Example of a chain graph:

- X is a light switch
- Y is the state of an associated electrical circuit
- Z is the state of a light bulb

The functions determining the relationship among these variables are as follows:

$$V = \{X, Y, Z\}, U = \{U_X, U_Y, U_Z\}, F = \{f_X, f_Y, f_Z\}$$

$$f_X : X = U_X$$

$$f_Y : Y = \begin{cases} \text{Closed IF } (X = \text{Up AND } U_Y = 0) \text{ OR } (X = \text{Down AND } U_Y = 1) \\ \text{Open otherwise} \end{cases}$$

$$f_Z : Z = \begin{cases} \text{On IF } (Y = \text{Closed AND } U_Z = 0) \text{ OR } (Y = \text{Open AND } U_Z = 1) \\ \text{Off otherwise} \end{cases}$$

The independencies shared and the dependencies that are likely shared by these functions are:

1. *Z and Y are dependent*
   For some z, y, $P(Z = z|Y = y) \neq P(Z = z)$
2. *Y and X are dependent*
   For some y, x, $P(Y = y|X = x) \neq P(Y = y)$
3. *Z and X are likely dependent*
   For some z, x, $P(Z = z|X = x) \neq P(Z = z)$
4. *Z and X are independent, conditional on Y*
   For all x, y, z, $P(Z = z|X = x, Y = y) = P(Z = z|Y = y)$

3.3.2 Forks

Three nodes, with two arrows emanating from the middle variable is called a fork. A typical fork is shown Figure 8 with the exogenous variables.

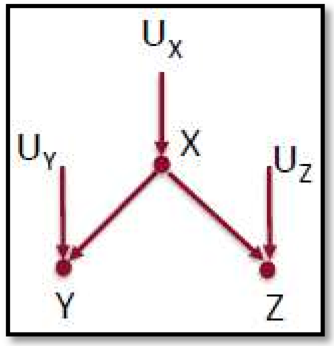


Figure 8 - Example of a fork graph

Example of a fork graph:

- X is a light switch
- Y is the state of the first light bulb
- Z is the state of the second light bulb

The functions determining the relationship among these variables are as follows:

$$V = \{X, Y, Z\}, U = \{U_X, U_Y, U_Z\}, F = \{f_X, f_Y, f_Z\}$$

$$f_X : X = U_X$$

$$f_Y : Y = \begin{cases} \text{On IF } (X = \text{Up AND } U_Y = 0) \text{ OR } (X = \text{Down AND } U_Y = 1) \\ \text{Off otherwise} \end{cases}$$

$$f_Z : Z = \begin{cases} \text{On IF } (X = \text{Up AND } U_Z = 0) \text{ OR } (X = \text{Down AND } U_Z = 1) \\ \text{Off otherwise} \end{cases}$$

The following independencies and the dependencies are valid for this fork graph:

1. *X and Y are dependent.*
   For some $x, y, P(X = x|Y = y) \neq P(X = x)$
2. *X and Z are dependent.*
   For some $x, z, P(X = x|Z = z) \neq P(X = x)$
3. *Z and Y are likely dependent.*
   For some $z, y, P(Z = z|Y = y) \neq P(Z = z)$
4. *Y and Z are independent, conditional on X.*
   For all $4x, y, z, P(Y = y|Z = z, X = x) = P(Y = y|X = x)$

3.3.3 Colliders

Three nodes, one node receives links from two other nodes is called a collider. A typical collider is shown Figure 9 with the exogenous variables.

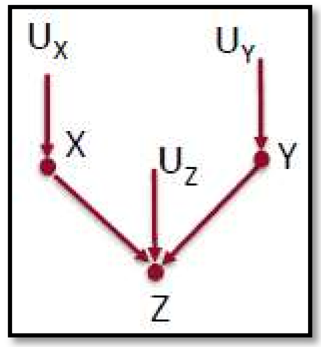


Figure 9 - Example of a collider

Example of a collider:

- X is a toss of a fair coin (simultaneous with Y)
- Y is a toss of a fair coin (simultaneous with X)
- Z is the bell that rings whenever at least one of the coins lands on heads

If we know that Coin 1 landed on heads, it tells us nothing about the outcome of Coin 2, due to their independence. But suppose that we hear the bell ring and then we learn that Coin 1 landed on tails. We now know that Coin 2 must have landed on heads. Similarly, if we assume that we've heard the bell ring, the probability that Coin 1 landed on heads changes if we learn that Coin 2 also landed on heads.

The following independencies and the dependencies are valid for this collider graph:

1. ***X and Z are dependent.***
   For some x, z, $P(X = x|Z = z) \neq P(X = x)$
2. ***Y and Z are dependent.***
   For some y, z, $P(Y = y|Z = z) \neq P(Y = y)$
3. ***X and Y are independent.***
   For all x, y, $P(X = x|Y = y) = P(X = x)$
4. ***X and Y are dependent conditional on Z.***
   For some x, y, z, $P(X = x|Y = y, Z = z) \neq P(X = x|Z = z)$

3.3.4 Blocking

Paths can be blocked by conditioning on nodes on the path. If we condition on Y (), we block the path from X to Z in Figure 10.

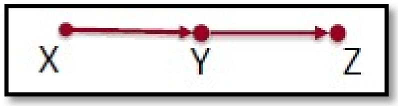


Figure 10 - Blocking a path

Let's assume that

- X is a light switch
- Y is the state of an associated electrical circuit
- Z is the state of a light bulb

If we condition on Y by giving a value, for example, 'broken', X and Z become independent.

Figure 11 shows the blocking situations of the graphs that contain a chain, a fork and a collider.

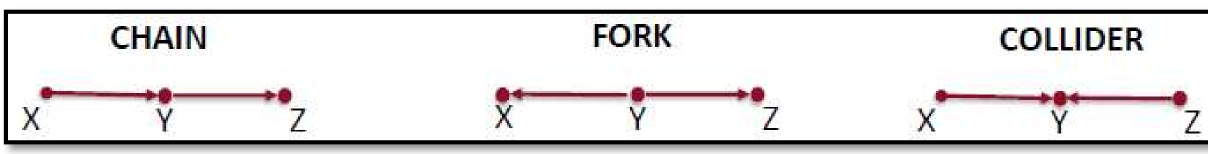


Figure 11 - Blocking situations of chains, forks and colliders

1. **For the chain:** If we condition on Y (the node in the middle of a chain), we block the path from X to Z.
2. **For the fork:** If we condition on Y (the node in the middle of a fork), we block the path from X to Z.
3. **For the collider:** If we condition on Y (the node in the middle of a collider), we unblock the path from X to Z. Colliders are blocked by nature and conditioning them unblocks the path.

3.3.5 d-Separation and d-Connection

Causal models are generally not as simple as the cases we have examined so far. Specifically, it is rare for a graphical model to consist of a single path between variables. In most graphical models, pairs of variables will have multiple possible paths connecting them, and each such a path will traverse a variety of chains, forks, and colliders. The question remains whether there is a criterion or process that can be applied to a graphical causal model of any complexity in order to predict dependencies that are shared by all data sets generated by that graph. There is, indeed, such a process: d-separation (the d stands for "directional"), which allows us to determine, for any pair of nodes, whether the nodes are d-connected, meaning there exists a connecting path between them, or d-separated, meaning there exists no such path. When we say that a pair of nodes are d-separated, we mean that the variables they represent are definitely independent; when we say that a pair of nodes are d-connected, we mean that they are possibly, or most likely, dependent (Pearl et. al., 2016).

Following statements explain d-separation and d-connection situations between the two nodes X and Y:

1. Two nodes X and Y are **d-separated** if **every** path between them is **blocked**, which means the path becomes **inactive.**
2. X and Y are d-separated when there are **no active paths** between them.
3. X and Y are d-connected if there is **any active path** between them
4. A path is active when **all the nodes on the path are active**.
5. X and Y **are d-separated** by the set {Z} **if and only if** they **are not d-connected** by the set {Z}, vice-versa.

In order to further analyse d-separation and d-connection let's consider the graph shown in Figure 12 and answer the following questions.

Q1. Are X and Y d-separated by the empty set { }?
Q2. Are X and Y d-separated by {Z1}?
Q3. Are X and Y d-separated by {Z2}?
Q4. Are X and Y d-separated by {Z1, Z2, Z3}?

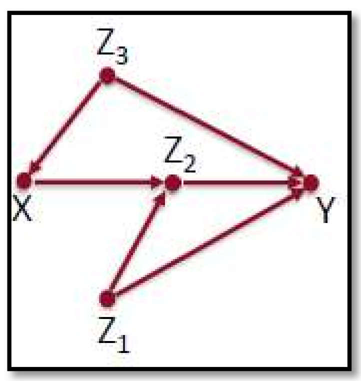


Figure 12 - d-Separation and d-Connection

All of the questions are related to X and Y nodes and in order to answer the questions from Q1 to Q4, we need to determine the paths from X to Y. Figure 13 shows all the possible paths from X to Y.

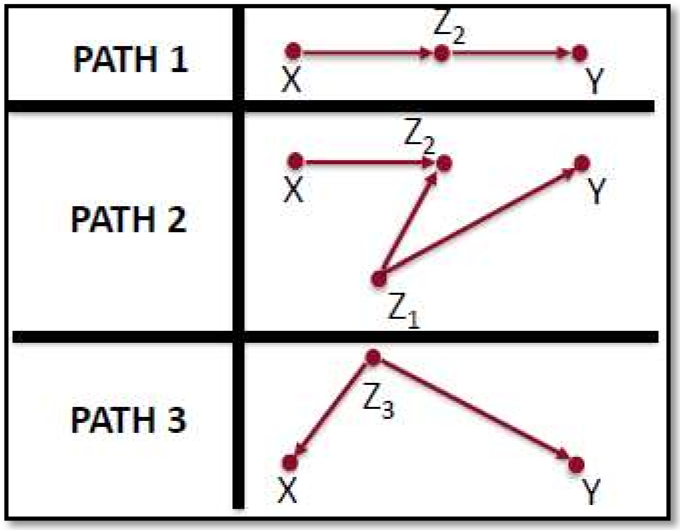


Figure 13 - Paths from X to Y

Now let's answer the questions.

Q1. Are X and Y d-separated by the empty set { }?

Here, empty set means there is no conditioning on any of the nodes and the graph is considered "as is". Therefore, the paths shown in Figure 13 are valid. Remember the statements given in 3.3.5.

2. Two nodes (X and Y) are d-separated when there are no active paths between them and

   If a graph is not d-separated then it has to be d-connected based on the statement of:

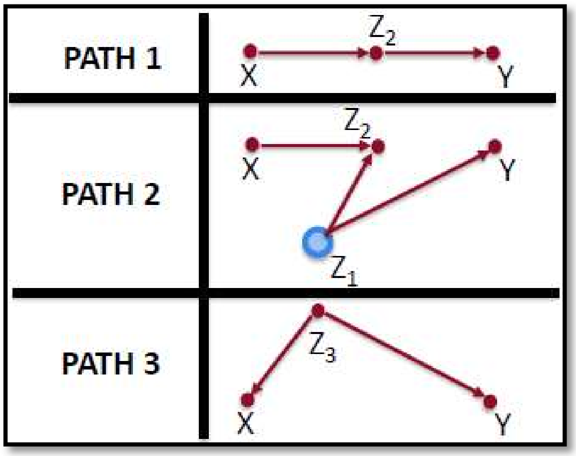
5. X and Y are d-separated by the set {Z} if and only if they are not d-connected by the set {Z}, vice-versa.

The following tables explains the relationship between the nodes, X and Y:

PATH	ACTIVE OR INACTIVE
Path 1: X → Z_2 → Y	Active
Path 2: X → Z_2 ← Z_1 → Y	Inactive
Path 3: X ←Z_3 → Y	Active
X d-separated Y | { }	No
X d-connected Y | { }	Yes

Q2. Are X and Y d-separated by {Z1}?

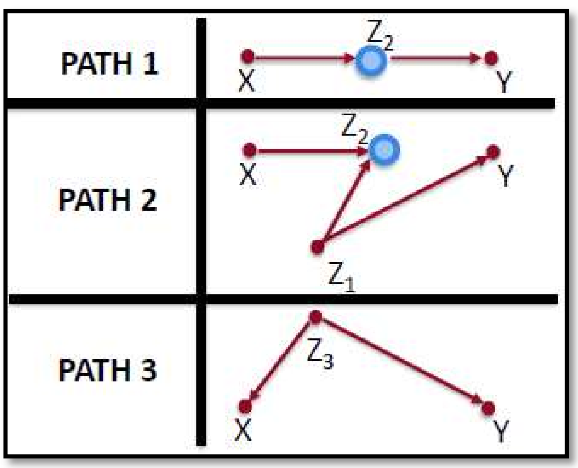
Here, we condition on Z1, which is a fork for the Path 2, and it is blocked when it is conditioned, and the new paths and the relationship between the nodes, X and Y, are shown below based on the 2. Statement given in 3.3.4.



PATH	ACTIVE OR INACTIVE
Path 1: X → Z_2 → Y	Active
Path 2: X → Z_2 ← Z_1 → Y	Inactive
Path 3: X ←Z_3 → Y	Active
X d-separated Y | { }	**No**
X d-connected Y | { }	**Yes**

Q3. Are X and Y d-separated by {Z2}?

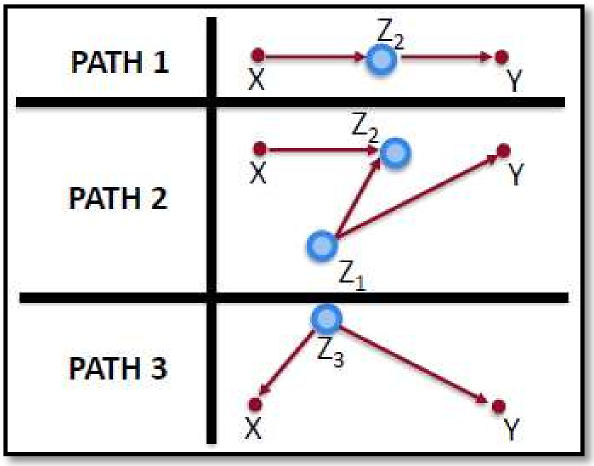
Here, we condition on Z2, which is a chain for the Path 1 and a collider for the Path 2. It is blocked on the Path 1 but it is unblocked on the Path 2. The new paths and the relationship between the nodes, X and Y, are shown below.



PATH	ACTIVE OR INACTIVE
Path 1: X → Z_2 → Y	Inactive
Path 2: X → Z_2 ← Z_1 → Y	Active
Path 3: X ←Z_3 → Y	Active
X d-separated Y | { }	**No**
X d-connected Y | { }	**Yes**

Q4. Are X and Y d-separated by {Z1, Z2, Z3}?

Here, we condition on Z1, Z2 and Z3 and the new paths and the relationship between the nodes, X and Y, are shown below.



PATH	ACTIVE OR INACTIVE
Path 1: X → Z_2 → Y	Inactive
Path 2: X → Z_2 ← Z_1 → Y	Inactive
Path 3: X ← Z_3 → Y	Inactive
X d-separated Y | { }	**Yes**
X d-connected Y | { }	**No**

4. CAUSAL INFERENCE IN MINING

Based on the literature survey performed within the scope of this paper, there has been no application of causal inference in mining industry. An attempt will be made to develop a causal relationship for a mineral processing application. The graph shown in Figure 14 is a draft of the causal relationship developed for a mineral processing plant where copper and molybdenum are being separated and recovered. U's are the exogeneous variables, which are considered as the root causes without having an ancestors, Cu and Mo are copper and molybdenum concentrations and RecCu and RecMo are the recovery of copper and recovery of molybdenum, respectively.

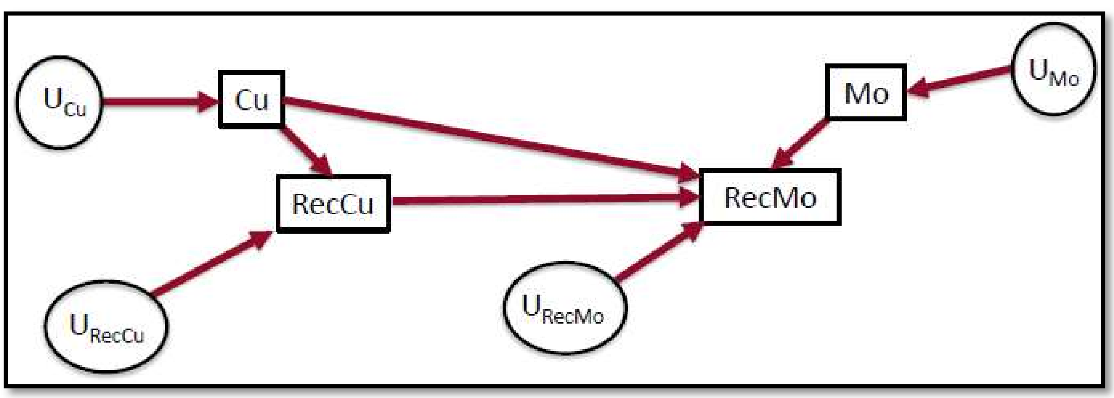


Figure 14 - Causal inference for Cu and Mo recoveries

This model implies that recovery of Cu only depends on the Cu content, while recovery of Mo depends on both the Mo content and the Cu content, as well as its recovery. In practice, there are many other factors that would influence the recoveries, particularly, mineralogical associations and other processing conditions. A graphical model such as the one presented above could be tested with data and modified to reflect the causal relationships that really exist in this process.

The graph is part of the preliminary study and still in progress.

5. Summary

In this paper, a review of Causal Inference is presented in general terms together with the basic definitions, the concept of counterfactuals and the difference between correlation and causation are discussed, fundamental principles of graphical models are explained by using simple graphical models, Statistical Inference and Causal Inference are compared and a causal relationship at an initial level is established for molybdenum recovery by developing a graphical model that involves influencing factors.

6. References

"What is causality? Definition and meaning" (n.d.). In Business-Dictionary.com. Retrieved from http://www.businessdictionary.com/

Holland, P.W., 1986. Journal of the American Statistical Association, Volume 81, Issue 396 pp. 945 - 960.

Hulswit, M., 2004. A Short History of 'Causation'. S.E.E.D. Journal (Semiosis, Evolution, Energy and Development, vol. 4, iss. 3, pp. 16-42.

Pearl, J., Jewell, P.N., Glymour, M., 2016. Causal Inference in Statistics: A Primer.

Pearl, J., 2003. Statistics and Causal Inference: A Review. Sociedad de Estadística e Investigación Operativa Test (2003) Vol. 12, No. 2, pp. 101–165.

Røysland, K., 2012. Counterfactual Analyses with Graphical Models Based on Local Independence. The Annals of Statistics, Vol. 40, No. 4, pp. 2162-2194