Causal Inference: Homework 1

Pakhomov Daniil

September 26, 2016

1 Task 2.2.2

1.1 Problem definition

Which confidence interval is wider, for the mean of X_1 or the mean of X_2. Why do you think that is?

What substantive conclusions can you draw from the coefficient values of the regression model you fit? Depression was evaluated by a questionaire where a lower value is better.

1.2 Solution

The confidence interval is wider for the mean of X_2 because the diagonal values of the Σ represent variance of the random variable. Since variance of the X_2 is greater than X_1, we conclude that confidence interval is wider by definition of variance.

The coefficient of the regression model show that the post depression level is very strongly associated with the depression level before the experiment. Financial problems are also associated with high level of stress. It is important to notice that this is just an association that doesn't mean causation.

2 Task 1

2.1 Problem definition

Give examples for the Reichenbach's Common Cause Principle of two variable A and B where A and B are associated.

2.2 Solution

A causes B example: Let B represent the fact that the majority of people in a certain region wear winter clothes and A - the temperature goes below a certain temperature (like 0 degrees celsius).

Let A and B represent the fact if two employees (A and B) were late. And variable C should represent if the train arrived on time. Two employees take this train to work and if the train is late, it causes both of them to be late.

For the example of A and B that are associated but doesn't comply with Reichenbach's Common Cause Principle, we can take any example from the spurious correlations website.

3 Task 2

3.1 Problem definition

Consider two versions of Pearl's example. We record the recovery rates of 700 patients who were given access to the drug. A total of 350 patients chose to take the drug and 350 did not. In version 2 of the experiment some patients complained of a side effect, and recovery rates seem different in those who complained.

Would you give the drug in version 1 of the problem? In version 2 of the problem? Explain.

3.2 Solution

Let us use the formula:

$$P(Y = y|do(X = x)) = \sum_z P(Y = y|X = x, Z = z)P(Z = z)$$

Where $X = 1$ stands for the patient taking the drug, $Z = 1$ stands for the patient being male, $Y = 1$ for the patient recovering.

Applying the formula gives us:

$$P(Y = 1|do(X = 1)) = 0.832$$

And

$$P(Y = 1|do(X = 0)) = 0.7818$$

Computing ACE we get:

$$ACE = P(Y = 1|do(X = 1)) - P(Y = 1|do(X = 0)) = 0.832 - 0.7818 = 0.0502$$

This shows positive advantage of taking the drug.

In case of side effect, we can perform the same experiment but taking Z variable to be side effect. Most probably the side effect wasn't caused by the drug as also people from the placebo group were affected. We can interpret this in a way that a certain amount of participants were randomly affected by disease which we can represent using Z variable. The computation of ACE will be the same and the drug will positively affect the patients.

4 Task 3

4.1 Problem definition

Why do you think Simpson's paradox is a paradox?

4.2 Solution

Like other paradoxes, it only appears to be a paradox because of incorrect assumptions, incomplete or misguided information, or a lack of understanding a particular concept.

According to one example at wikipedia, the paradox is caused when the percentage is provided but not the ratio. So one worker can be seem to work more productive than another given only percentage for each day, but be worse in the final productivity computation. This occurs because only percentage was reported for each day and the final statistics and no the actual ratio.

The paradoxical elements disappear when causal relations are brought into consideration.

5 Task 4

5.1 Problem definition

Derive the maximum likelihood estimate of λ for exponential distribution.

Derive expectation of the exponential distribution.

5.2 Solution

Probability that we want to maximize:

$$p(x_1, \ldots, x_n) = \prod_{i=1}^{n} \lambda e^{(-\lambda x_i)}$$

Log-likelihood:

$$ln(\prod_{i=1}^{n} \lambda e^{(-\lambda x_i)}) = nln(\lambda) + \sum_{i=1}^{n} ln(e^{-\lambda x_i}) = nln(\lambda) - \sum_{i=1}^{n} \lambda x_i$$

Derivative of log-likelihood relative to λ and equated to zero:

$$\frac{n}{\lambda} = \sum_{i=1}^{n} x_i \implies \lambda = \frac{n}{\sum_{i=1}^{n} x_i}$$

Expectation of exponential distribution can be derived by solving an integral:

$$E[x] = \int_{-\infty}^{+\infty} e^{-\lambda x} x dx$$

Integral is solved by parts taking $u = x$ and $dv = e^{-\lambda x} dx$ and taking a limit:

$$E[x] = \frac{1}{\lambda^2}$$

6 Task 5

6.1 Problem definition

Show that conditional independence satisfies the semi-graphoid axioms.

6.2 Solution

6.2.1 Symmetry

$$X \perp\!\!\!\perp Y|Z \implies p(X, Y|Z) = p(X|Z)p(Y|Z) = p(Y|Z)p(X|Z) \implies Y \perp\!\!\!\perp X|Z$$

6.2.2 Decomposition

To prove:

$$W \cup X \perp\!\!\!\perp Y|Z \implies X \perp\!\!\!\perp Y|Z$$

The proof:

$$\begin{aligned}
&p(w, x, y|z) = p(w, x|z)p(y|z) \\
&p(w, x, y, z) = \frac{p(w, x, z)p(y, z)}{p(z)} \\
&p(x, y, z) = \sum_w p(w = w, x, y, z) = \sum_w \frac{p(w = w, x, z)p(y, z)}{p(z)} = \frac{p(y, z)}{p(z)} p(x, z) \\
&p(x, y|z) = \frac{p(x, y, z)}{p(z)} = \frac{p(y, z)p(x, z)}{p(z)p(z)} = p(x|z)p(y|z) \implies X \perp\!\!\!\perp Y|Z
\end{aligned}$$

6.2.3 Contraction

To prove:

$$(W \perp\!\!\!\perp X|Z) \wedge (W \perp\!\!\!\perp Y|Z \cup X) \implies W \perp\!\!\!\perp X \cup Y|Z$$

The proof:

$$\begin{aligned}
&p(w, y|x, z) = p(w|x, z)p(y|x, z) \\
&p(w, y, x, z) = \frac{p(w, x, z)p(y, x, z)}{p(x, z)} \\
&p(w, x|z) = p(w|z)p(x|z) \\
&p(w, x, z) = \frac{p(w, z)p(x, z)}{p(z)} \\
&p(w, y, x, z) = \frac{p(w, z)p(x, z)p(y, x, z)}{p(x, z)p(z)} = \frac{p(w, z)p(y, x, z)}{p(z)} \\
&p(w, y, x|z) = p(w|z)p(y, x|z) \implies W \perp\!\!\!\perp X \cup Y|Z
\end{aligned}$$

6.2.4 Weak Union

To prove:

$$W \cup X \perp\!\!\!\perp Y | Z \implies X \perp\!\!\!\perp Y | Z \cup W$$

The proof:

$$W \cup X \perp\!\!\!\perp Y | Z \implies p(x, y, w|z) = p(w, x|z)p(y|z)$$
$$p(x, y|z, w) = \frac{p(x, y, w|z)}{p(w|z)} = \frac{p(w, x|z)p(y|z)}{p(w|z)}$$
$$\frac{p(w, x|z)p(y|z)}{p(w|z)} = p(y|w, z)p(y|z) = p(y|w, z)p(y|w, z) \implies X \perp\!\!\!\perp Y | Z \cup W$$