TSC1 Notes

October 15, 2014

1 Simple Linear Example

Consider a simple driver-response system:

$$
\begin{aligned}
x_t &= \sin(t) \\
y_t &= x_{t-1} + \eta_t \\
&= \sin(t-1) + \eta_t \\
&= \sin(t)\cos(1) - \cos(t)\sin(1) + \eta_t
\end{aligned}
$$

with $\eta_t \sim \mathcal{N}(0,1)$. Define

$$\delta x_t \equiv \frac{dx}{dt} \approx \frac{\Delta x}{\Delta t} = x_t - x_{t-1}$$

and

$$\delta y_t \equiv \frac{dy}{dt} \approx \frac{\Delta y}{\Delta t} = y_t - y_{t-1} \quad .$$

It follows that

$$
\begin{aligned}
\delta x_t &= \sin(t) - \sin(t-1) \\
&= \sin(t) - (\sin(t)\cos(1) - \cos(t)\sin(1)) \\
&= \sin(t)\,(1-\cos(1)) + \sin(1)\cos(t) \\
&\equiv \kappa_1 \sin(t) + \kappa_2 \cos(t)
\end{aligned}
$$

with $\kappa_1 = 1 - \cos(1)$ and $\kappa_2 = \sin(1)$, and

$$
\begin{aligned}
\delta y_t &= x_{t-1} + \eta_t - x_{t-2} - \eta_{t-1} \\
&= \sin(t-1) - \sin(t-2) + \eta_t - \eta_{t-1} \\
&= (\sin(t)\cos(1) - \cos(t)\sin(1)) - (\sin(t)\cos(2) - \cos(t)\sin(2)) + \eta'_t \\
&= \sin(t)\,(\cos(1) - \cos(2)) + \cos(t)\,(\sin(2) - \sin(1)) + \eta'_t \\
&\equiv k_1 \sin(t) + k_2 \cos(t) + \eta'_t
\end{aligned}
$$

with $\eta'_t = \eta_t - \eta_{t-1} \sim \mathcal{N}(0,2)$[1], $k_1 = (\cos(1) - \cos(2))$ and $k_2 = (\sin(2) - \sin(1))$.

The main idea of Local Impulse Response (LIR) causality inference is to use a subsetting procedure on some (or all) of the four time series x_t, y_t, δx_t, and δy_t to determine the causality of the system.

Consider a few extreme points in the driver cycle, e.g. $t = n\pi$ with $n = 0, 1, 2, 3, 4, \ldots$. The driver values are

$$x_{n\pi} = \sin(n\pi) = 0 \quad ,$$

the response values are

$$y_{n\pi} = \sin(n\pi)\cos(1) - \cos(n\pi)\sin(1) + \eta_{n\pi} = \eta_{n\pi} + (-1)^n \sin(1) \quad ,$$

the local change in the driver values are

$$\delta x_{n\pi} = \kappa_1 \sin(n\pi) + \kappa_2 \cos(n\pi) = (-1)^n \kappa_2 \quad ,$$

and the local change in the response values are

$$\delta y_{n\pi} = k_1 \sin(n\pi) + k_2 \cos(n\pi) + \eta'_{n\pi} = \eta'_{n\pi} + (-1)^n k_2 \quad .$$

Consider $t = n\pi/2$ with $n = 1, 2, 3, 4, \ldots$. The driver values are

$$x_{n\frac{\pi}{2}} = \sin\left(n\frac{\pi}{2}\right) = (-1)^n \quad ,$$

[1]The difference of two normal distributions with means μ_1 and μ_2 and variances σ_1^2 and σ_2^2 is another normal distribution with mean $\mu_1 - \mu_2$ and variance $\sigma_1^2 + \sigma_2^2$.

the response values are

$$y_{n\frac{\pi}{2}} = \sin\left(n\frac{\pi}{2}\right)\cos(1) - \cos\left(n\frac{\pi}{2}\right)\sin(1) + \eta_{n\frac{\pi}{2}} = \eta_{n\frac{\pi}{2}} + (-1)^n\cos(1) \ \ ,$$

the local change in thd driver values are

$$\delta x_{n\frac{\pi}{2}} = \kappa_1\sin\left(n\frac{\pi}{2}\right) + \kappa_2\cos\left(n\frac{\pi}{2}\right) = (-1)^n\kappa_1 \ \ ,$$

and the local change in the response values are

$$\delta y_{n\frac{\pi}{2}} = k_1\sin\left(n\frac{\pi}{2}\right) + k_2\cos\left(n\frac{\pi}{2}\right) + \eta'_{n\frac{\pi}{2}} = \eta'_{n\frac{\pi}{2}} + (-1)^n k_1 \ \ .$$

2 Subsetting for LIR variance

Consider $\mathbf{X} = \{x_t\}$ and $\mathbf{Y} = \{y_t\}$ given $t \in [0, 4\pi]$. Let L be the library length of $\mathbf{X}$ and $\mathbf{Y}$. These systems have five points where $t = n\pi$ for $n = 0, 1, 2, 3$ and 4. Thus, an m-binned histogram of $\mathbf{X}$, where $m \geq L$, would have a bin, b_0, centered at $x_t = 0$ that contains the five points $\mathbf{X}_0 = \{x_{n\pi}\}$. (If $m < L$, then $\mathbf{b}_0$ would contain at least five points but the total number of points in $\mathbf{b}_0$ would be a function of the total number of bins, assuming bins of equal sizes.)

Consider the set of time steps $\mathbf{T} = \{t = n\pi\}\ \forall n = 0, 1, 2, 3, 4$ for which the values in $\mathbf{X_0}$ are achieved. The local impulses immediately preceding $\mathbf{b}_0$ are $\delta\mathbf{X_T} = \{\delta x_{n\pi}\}$, which also contains five points. However, those five points would not appear in a single bin of an m-binned histogram of $\{\delta x_t\}$ (given $m \geq L$). The set $\delta\mathbf{X_T}$ would actually be split into two separate bins in such a histogram, one for the three points equal to κ_2 and one for the two points equal to $-\kappa_2$. Thus, the time steps associated all of the points in a given histogram bin of a given time series, e.g. $\mathbf{b}_0$, do not necessarily correspond to points in a single histogram bin of different (though related) times series, e.g. $\delta\mathbf{X_T}$. This idea is straightforward but it is the basic idea underlying the subsetting method for calculating the LIR variance.

The subsetting is premised on the following:

- The local temporal response causally depends on the local temporal change in the driver; e.g. y_t causally depends on δx_t
- The local temporal response causally depends on the immediately preceding response; e.g. y_t causally depends on y_{t-1}
- The local temporal response does not causally depend on the immediately preceding driver except through the local temporal change in the driver; e.g. y_t does not causally depend on x_{t-1} except through δx_t

Thus, the subsetting procedure is as follows:

1. Create an m-binned histogram of the response signal $\mathbf{R} = \{r_t\}$.
2. Given m bins $\mathbf{b}_i$ where i denotes the center of the mth bin, create an m'-binned histogram of the change in the driver signal $\delta\mathbf{D} == \{\delta d_t\} = \{d_t - d_{t-1}\}$ at the time steps $\tau = \{t \mid r_t \in \mathbf{b}_i\}$.
3. Given m' bins $\mathbf{b'}_j$ where j denotes the center of the m'th bin, find the variance of the response at the time steps immediately following (i.e. $t+1$) the time steps $\tau' = \{t \mid \delta d_t \in \mathbf{b}_j\}$.

As an example, consider $\mathbf{R} = \mathbf{X}$ and $\mathbf{D} = \mathbf{Y}$. An m-binned histogram of $\mathbf{X}$ would lead to a bin centered at zero, b_0, that contains at least five points evaluated at $t = n\pi\ \forall n = 0, 1, 2, 3, 4$. Thus, the change in the driver signal, i.e. $\{\delta y_t\}$, evaluated at τ contains at least the five points $\eta'_{n\pi} + (-1)^n k_2$. The m'-binned histogram of step 2 would split these five points among different bins (depending on both the sign of k_2 and the value of $\eta'_{n\pi}$ for a given n). If each of the points is placed into a bin alone, then the variance calculations of step 3 become $var\left(\sin(n\pi + 1)\right)\ \forall n = 0, 1, 2, 3, 4$, which is five zeros because the variance of a single point is zero. Suppose all five points are placed into a single bin. The variance calculation of step 3 then becomes

$$
\begin{aligned}
var\left(\{\sin(n\pi+1)\,|n=0,1,2,3,4\}\right) &= var\left(\{\sin(n\pi)\cos(1)+\cos(n\pi)\sin(1)\,|n=0,1,2,3,4\}\right) && (1)\\
&= var\left(\{(-1)^n\kappa_2\,|n=0,1,2,3,4\}\right) && (2)\\
&= var\left(\{\kappa_2,-\kappa_2,\kappa_2,-\kappa_2,\kappa_2\}\right) && (3)\\
&= \frac{1}{5}\left((\kappa_2-\mu)^2+(-\kappa_2-\mu)^2+(\kappa_2-\mu)^2+(-\kappa_2-\mu)^2+(\kappa_2-\mu)^2\right) && (4)\\
&= \frac{1}{5}\left(\frac{16}{25}\kappa_2^2+\frac{36}{25}\kappa_2^2+\frac{16}{25}\kappa_2^2+\frac{36}{25}\kappa_2^2+\frac{16}{25}\kappa_2^2\right) && (5)\\
&= \frac{1}{5}\frac{120}{25}\kappa_2^2 && (6)\\
&= \frac{24}{25}\kappa_2^2 && (7)
\end{aligned}
$$

where $\mu = \kappa_2/5$. Thus, the LIR variance depends strongly on the number of bins used to construct the histograms in steps 1 and 2.

In this particular example, $\mathbf{Y}$ is known to be the response and $\mathbf{X}$ is known to be the driver. It may be assumed that our assignment of $\mathbf{R}=\mathbf{X}$ and $\mathbf{D}=\mathbf{Y}$ may proven "false" by comparing the variances given this assignment and it's complement, i.e. $LIRvar\,|\mathbf{R}=\mathbf{X},\mathbf{D}=\mathbf{Y}$ and $LIRvar\,|\mathbf{R}=\mathbf{Y},\mathbf{D}=\mathbf{X}$. It may be assumed that the lower LIR variance is indicative of a stronger causal inference, i.e. if $LIRvar\,|\mathbf{R}=\mathbf{X},\mathbf{D}=\mathbf{Y}$ is greater than $LIRvar\,|\mathbf{R}=\mathbf{Y},\mathbf{D}=\mathbf{X}$, then it may be assumed $\mathbf{R}=\mathbf{Y},\mathbf{D}=\mathbf{X}$ is the more "correct" assignment. Notice, however, that we have already shown that the incorrect assignment of $\mathbf{R}=\mathbf{X}$ and $\mathbf{D}=\mathbf{Y}$ can lead to an LIR variance of variance. The correct assignment of $\mathbf{R}=\mathbf{Y}$ and $\mathbf{D}=\mathbf{X}$ cannot lead to an LIR variance less than zero (variances are nonnegative). Thus, it seems that comparing LIR variances is not a robust method for causal inference.

3 LIR Approach to Probabilistic Causality

Probabilistic causality is centered on the definition that a cause C is said to *cause* (or *drive*) an effect E if

$$P\left(E|C\right) > P\left(E|\bar{C}\right) \quad ,$$

i.e. C causes E if the probability of E given C is higher than the probability of E given not C. The LIR causal inference method involves using e.g. $\{x_t\}$, $\{y_t\}$, $\{\delta x_t\}$, and $\{\delta y_t\}$ to determine the direction of causal influence in a system of two times series $\{x_t\}$ and $\{y_t\}$. It follows that applying LIR causal inference to probabilistic causality involves evaluating the above inequality given e.g. $C=\{x_t\}$, $\{y_t\}$, $\{\delta x_t\}$, or $\{\delta y_t\}$ and $E=\{x_t\}$, $\{y_t\}$, $\{\delta x_t\}$, or $\{\delta y_t\}$ given $E\neq C$ and for different temporal offsets.

The conditional probabilities are estimated using histograms of the time series data as follows:

$$P\left(E|C\right) \approx \frac{1}{L}H\left(E|C\right) = \frac{H\left(E\cap C\right)}{H(C)}$$

where $H(A)$ is an m-binned histogram of A, L is the library length of the E and C time series (which are assumed to be the same length). Similarly,

$$P\left(E|\bar{C}\right) \approx \frac{1}{L}H\left(E|\bar{C}\right) = \frac{H\left(E\cap\bar{C}\right)}{H(\bar{C})} \quad .$$

Define the *causal penchant*

$$\rho_{EC} = P\left(E|C\right) - P\left(E|\bar{C}\right) \approx \frac{H\left(E\cap C\right)}{H(C)} - \frac{H\left(E\cap\bar{C}\right)}{H(\bar{C})} \quad .$$

If C causes E, then $\rho_{EC} > 0$. Otherwise, i.e. $\rho_{EC} \leq 0$, the causal influence of the system is *undefined*(?). Causal influence between a pair of time series is accomplished by comparing penchants.