tscausal: time series causal inference

Motivation

The objective of the *tscausal* package is to provide an easy to use framework to evaluate any time series based forecast distribution, proxying for the counterfactual distribution, and generate an automatic set of diagnostics on the significance of an intervention. It is heavily influenced by the CausalImpact package described in Brodersen et al. [2015], and from which it borrows parts of the diagnostics and output narrative.

Unlike *CausalImpact*, the package is agnostic to any particular time series model, but it does require that the forecast distribution conform to certain requirements, namely that it inherits class "tsmodel.distribution". Any matrix of size $N \times h$, representing the N simulated points for each forecast period $t \in h$, where $t > T$ (the intervention date), can be coerced to this class. It is assumed that the user can and has trained a model on the pre-intervention period and our code repository provides for a wide array of models, with common calling conventions and the ability to generate frequentist simulated or Bayesian predictive distributions, all of which inherit class "tsmodel.distribution".

In the absence of very good controls (external regressors), most of the time series models we use are founded on autoregressive type building blocks for the motion dynamics which impart the typical widening of the prediction intervals as the forecast horizon grows, commensurate with future uncertainty.

Evaluation of Significance

Pointwise Impact

Given an intervention date T, and a post-intervention period under consideration $T+1, \ldots, T+h$, then the causal pointwise effect $\phi_{t \forall t \in T+1, \ldots, T+h}^{(s)}$ given a draw s from the predictive distribution is:

$$\phi_t^{(s)} = y_t - \hat{y}_t^{(s)}$$

where y represents the outcome variable and $\hat{y}$ the forecasted (or counterfactual) variable. Because we have a distributional forecast, representing the uncertainty about the forecasted variable $\hat{y}$ at each date in the horizon $T+1, \ldots, T+h$, and indexed by draws $(s)_1, \ldots, (s)_N$, we can generate an empirical distribution of any function applied to the distribution (e.g. summation, differences etc). Given a significance level $a\%$, then the $100 - a\%$ confidence interval can be evaluated by making use of the empirical quantile function from this distribution[1]

Cumulative Impact

The cumulative impact is derived from the pointwise impact as :

$$\Phi_t^{(s)} = \sum_{i=T+1}^{t} \phi_i^{(s)}$$

---

[1]Where the empirical quantile is defined as :

$$F_n^{-1}(p) = \inf \{x : F_n(x) \geqslant p\} = X_{n(i)}$$

where i is chosen such that $\frac{i-1}{n} < p \leqslant \frac{i}{n}$ and $X_{n(i)}, \ldots, X_{n(n)}$ are the order statistics of the sample i.e. such that $X_{n(1)} \leqslant \ldots \leqslant X_{n(n)}$ and $\left(X_{n(1)}, \ldots, X_{n(n)}\right)$ is a permutation of the sample $(X_1, \ldots, X_n)$.

which is a useful quantity when the outcome represents a flow variable (e.g. revenue, signups), measured over an interval of time, and not interpretable when it represents a stock variable (e.g. inventory, subscribers).

Mean Effect

Given an intervention date T, and a post-intervention period $T+1, \ldots, T+h$, then the average effect given a draw s from the predictive distribution is:

$$\alpha^{(s)} = \frac{1}{t-T} \sum_{i=T+1}^{T+h} y_i - \hat{y}_i^{(s)}$$

Mean Relative Effect

A related measure, showing the relative effect of the intervention can be defined :

$$r^{(s)} = \frac{\alpha^{(s)}}{\mu_{\hat{y}}}$$

where $\mu_{\bar{y}}$ is the mean of the forecasted variable $\forall t > T$ and $\forall s$.

Evaluation of Significance

Given a signficance level a, then under the null hypothesis that the intervention was not significant, we fail to reject if $F_n^{-1}(\alpha/2) < 0$ && $F_n^{-1}(1-\alpha/2) > 0$, and reject otherwise. For instance, if the effect was positive with a lower quantile value above zero then we would reject the null and conclude that the effect was positive and significant. If the effect was negative with an upper quantile value less than zero we would also reject the null and conclude that the effect was negative and significant.

Another statistic which can be used to evaluate whether the results could have occurred by chance is given by the following:

$$\frac{1}{N} \sum_{j=1}^{N} \left[\left(\sum_{i=T+1}^{T+h} \phi_t^{(j)} \right) \geqslant 0 \right]$$

which measures the percent of times the total net effect for the entire forecast horizon and across all draws of the distribution was positive.[2]

Package Features

The package has one main method called **tscausal** which takes as arguments a *tsmodel.distribution* object, the outcome variable (as xts object which includes both the pre and post intervention data), an optional xts vector (or *tsmodel.distribution*) of the in-sample fitted values and the significance level for hypothesis testing **alpha**, as shown below:

```
## function (object, actual, fitted = NULL, alpha = 0.05, ...)
## NULL
```

Additionally, once the causal object has been created, there are methods for plotting (*plot*), printing of results (*print*) as well as a reporting (*tsreport*) which can be used to automatically generate a pdf, doc or html report with full evaluation and graphics. The next section provides a demonstration.

---

[2]For negative net effects, reverse the direction of the inequality, or to capture all cases use:

$$\min \left\{ \frac{1}{N} \sum_{j=1}^{N} \left[\left(\sum_{i=T+1}^{T+h} \phi_t^{(j)} \right) \geqslant 0 \right], \frac{1}{N} \sum_{j=1}^{N} \left[\left(\sum_{i=T+1}^{T+h} \phi_t^{(j)} \right) \leqslant 0 \right] \right\}$$

Package Demo

We use the priceunits dataset, and in particular the units series, and add an artificial structural break during 1999-03 which could represent for instance a large drop in prices, some large substitution effect or a succesful marketing campaign.

```r
data(priceunits, package = "tsdatasets")
priceunits <- priceunits[,2]
priceunits["1999-03/"] <- priceunits["1999-03/"] + 10
train <- "/1999-02"
test = "1999-03/1999-08"
y_train = priceunits[train]
y_test  = priceunits[test]
spec <- ets_modelspec(y_train, frequency = 12, lambda = 0, model = "AAN")
mod <- estimate(spec)
```

```r
summary(mod)
```

```
##
## ETS Model [ AAN ]
##
## Param  Description                       Est[Value]
## -------------------------------------------------
##  alpha State[Level-coef]                   0.32527
##  beta  State[Slope-coef]                   0.10566
##  l0    State[Level-init]                   1.35278
##  b0    State[Slope-init]                   0.06322
##  sigma Observation[Standard Deviation] 0.08240
##
##  AIC    BIC    AICc
##  -40.17 -28.32 -39.35
##
##  MAPE   MASE   MSLRE  BIAS
##  0.0692 0.9563 0.0067 0.0054
```

```r
prediction <- predict(mod, h = nrow(y_test), nsim = 5000)
```

Note that the prediction object has a number of slots, one of which is called **distribution** and inherits the following classes : tsets.distribution, tsmodel.distribution. All our packages have a predictable and common set of slots in the returned prediction objects.

The final step is to pass the forecast distribution to the **tscausal** method and generate a report with a narrative explaining the results and a decision on whether to reject or not the null of no intervention effect.

```r
cause <- tscausal(prediction$distribution, actual = rbind(y_train, y_test),
                  fitted = fitted(mod))
print(cause)
```

```
## Posterior inference {tscausal}
##
##                           Average          Cumulative
## Actual                    37.33            224.00
## Prediction (s.d.)         27.87 (2.629)    167.21 (15.776)
## 95% CI                    [23, 33]         [139, 200]
##
## Absolute effect (s.d.)    9.464 (2.629)    56.786 (15.776)
## 95% CI                    [3.9, 14]        [23.5, 85]
##
## Relative effect (s.d.)    33.96% (9.434%)  33.96% (9.434%)
## 95% CI                    [14%, 51%]       [14%, 51%]
##
## Posterior tail-area probability p:   0.0014
## Posterior prob. of a causal effect:  99.86003%
```

```r
tsreport(cause)
```

During the post-intervention period, the response variable had an average value of approx. 37.3333 . By contrast, in the absence of an intervention, we would have expected an average response of 27.87 . The 95% interval of this counterfactual prediction is [23.1667 , 33.4127]. Subtracting this prediction from the observed response yields an estimate of the causal effect the intervention had on the response variable. This effect is 9.4643 with a 95% interval of [3.9206 , 14.1666]. For a discussion of the significance of this effect, see below. Summing up the individual data points during the post-intervention period (which can only sometimes be meaningfully interpreted), the response variable had an overall value of 224.0000 . By contrast, had the intervention not taken place, we would have expected a sum of 167.21 . The 95% interval of this prediction is [139.0004 , 200.4763].

The above results are given in terms of absolute numbers. In relative terms, the response variable showed an increase of +34% . The 95% interval of this percentage is [+14% , +51%].

This means that the positive effect observed during the intervention period is statistically significant and unlikely to be due to random fluctuations. It should be noted, however, that the question of whether this increase also bears substantive significance can only be answered by comparing the absolute effect (9.4643) to the original goal of the underlying intervention.

The probability of obtaining this effect by chance is very small (Bayesian one-sided tail-area probability p = 0.001). This means the causal effect can be considered statistically significant. NULL

The *tsreport* method has options for printing to screen or generating a pdf, html or docx (with option to use a template as well). Example usage is provided below:

```r
tsreport(cause, type = "pdf", output_dir = "~/tmp",
         args = list(model = "ETS[AAN]", frequency = 12, name = "Units"))
tsreport(cause, type = "html", output_dir = "~/tmp",
         args = list(model = "ETS[AAN]", frequency = 12, name = "Units"))
tsreport(cause, type = "doc", doc_template = "~/tmp/mytemplate.docx",
         output_dir = "~/tmp",
         args = list(model = "ETS[AAN]", frequency = 12, name = "Units"))
```
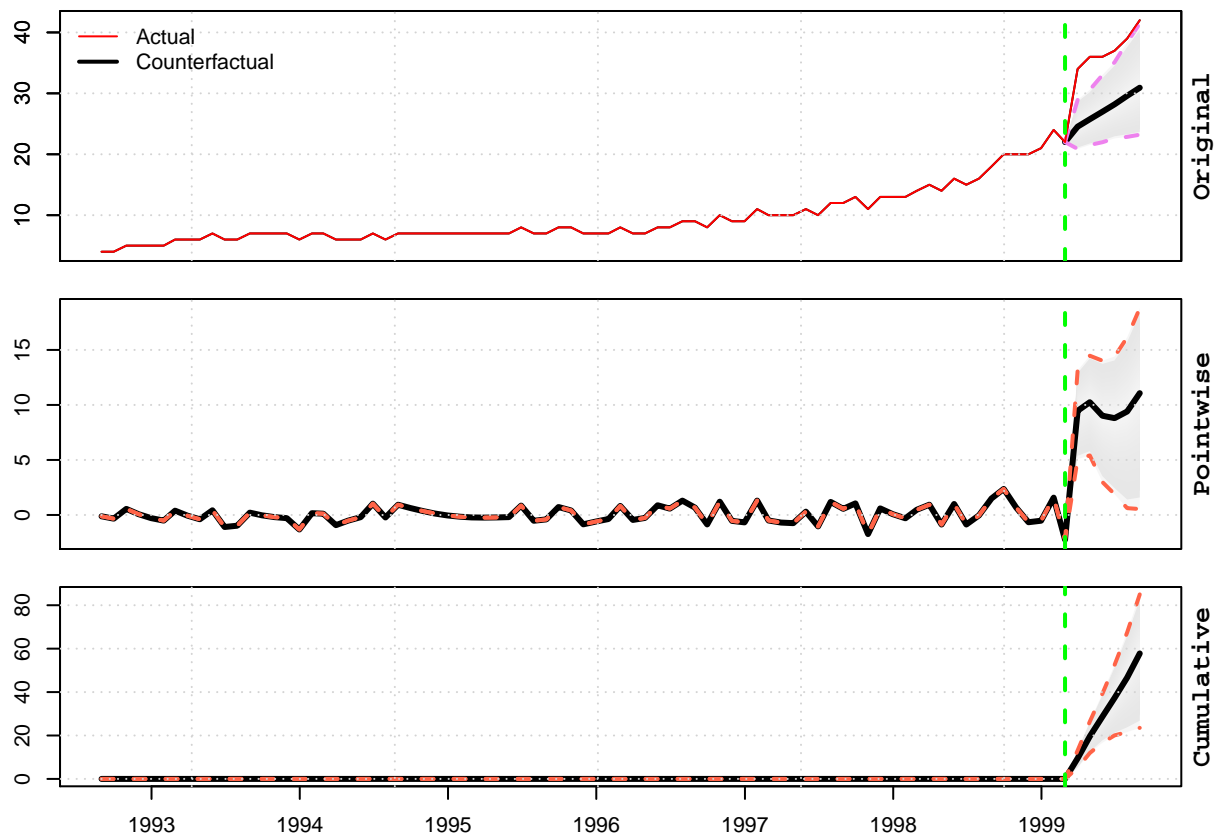
References

Kay H Brodersen, Fabian Gallusser, Jim Koehler, Nicolas Remy, Steven L Scott, et al. Inferring causal impact using bayesian structural time-series models. *The Annals of Applied Statistics*, 9(1):247–274, 2015.