# Appendix 1: consistency of the corrected estimator of the net benefit

For simplicity, we only consider one right-censored outcome and we drop the subscript relative to the outcome.

The U-statistic theory (e.g. Chapter 2 in U-statistics Theory and Practice by A.J. Lee, 1990, CRC) gives that for two independent samples and that are iid realization of, respectively, X and Y then the 2-sample U statistic based on the kernel f:

is an unbiased estimator of .

* with complete data

If we denote **,** we have that

is an unbiased estimator of .

* with right-censored data and the Gehan’s scoring rule

Denoting:

the Gehan’s scoring rule can be written as:

So the estimator

obtained with the Gehan’s scoring rule is a consistent estimator of

Using the assumption of random censoring and:

which is in general different from .

Considering now:

As shown in the article:

This converges towards:

So our estimator is consistent.

* with right-censored data and the Peron’s scoring rule

Consider first the case where the survival curves are known. Then if our survival model is correct

So

Is a consistent estimator of

i.e. of

which equals .

When we the survival curves are not fully known but estimated

The bias is mostly a function of (since has expectation 0). It is due to the fact that we use a downward biased estimators of and (denoted and in the article).

We once more consider:

where denotes the score obtained with the Peron’s scoring rule. As before

We introduce the empirical average of the estimated uninformative score ( over the pairs. Using that we get:

We don’t have a formal expression for the bias of but assuming that it is downward biased by a factor , then the corrected estimator is unbiased. This is for instance the case when or .

# Appendix 2: consistency of the estimator of the net benefit in presence of two outcomes

In presence of two outcomes we can express the net benefit as:

The estimator of the net benefit in absence of censoring is where:

**Peron’s scoring rule**

To simplify we assume that the survival are known. See appendix 1 for a discussion of how the bias induced by using lower bounds to estimate the scores in the case of 1 outcome can be removed by using the correction at the pair level.

* Same outcome with different thresholds:

The final net benefit should be the same as using only the smallest threshold. Indeed:

We can check that this is the case using our software:

> library(BuyseTest) ## BuyseTest version 1.7.7

> set.seed(10)

> dt <- simBuyseTest(25)

> BuyseTest(Treatment ~ tte(eventtime, censoring = status, threshold = 1) + tte(eventtime, censoring = status, threshold = 0),

data = dt, scoring.rule = "Peron", trace = 0, method.inference = "none")

endpoint threshold delta Delta

eventtime 1 0.4398 0.4398

eventtime 1e-12 -0.0343 0.4055

> BuyseTest(Treatment ~ tte(eventtime, censoring = status, threshold = 0),

data = dt, scoring.rule = "Peron", trace = 0, method.inference = "none")

endpoint threshold delta Delta

eventtime 1e-12 0.4055 0.4055

* Two distinct outcomes (with known survival curves)

We consider:

where:

So it essentially remains to compare the expectation of and .

If the two outcomes are independent then

=]

=].

Therefore and have the same expectation.

If the two outcomes are not independent then:

So we essentially assume that