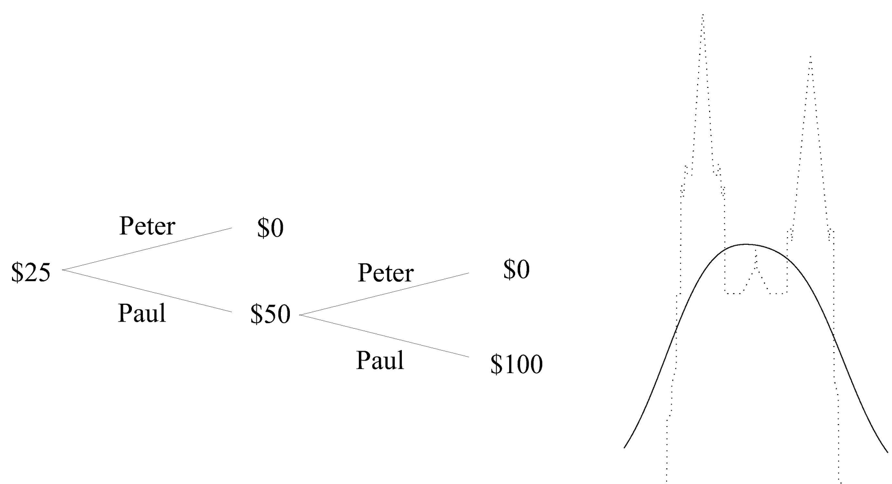
Confidence intervals for causal effects in sequential decision making

Vladimir Vovk and Ruodu Wang



The Game-Theoretic Probability and Finance Project

Working Paper #68

June 17, 2026

Project web site:
`http://www.probabilityandfinance.com`

Abstract

We derive confidence intervals and confidence sequences for causal effects in situations where the back-door criterion is applicable. Our tightest confidence intervals hold in the standard setting where the training data consists of IID observations over a system described by a given causal diagram. When interventions are allowed to depend on the past data, our confidence intervals become wider and involve a term coming from the law of the iterated logarithm, even where the number of observations is known in advance. In the sequential setting where the number of observations is not given, our confidence intervals, arranged into a confidence sequence for causal effects, involve more iterated logarithm terms and become even wider.

Contents

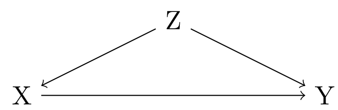


Figure 1: The basic causal graph of this paper

1 Introduction

A major limitation of many results in causal inference is that they assume, implicitly or explicitly, IID (independent and identically distributed) observations over a causal system. This limitation is shared by [13], where we derive prediction sets in situations covered by the back-door and front-door criteria [5, Theorems 3.3.2 and 3.3.4] (one section of [13] goes beyond the IID picture but only slightly). It was noted in [12] that the limitation can be easily overcome by applying limit theorems of probability theory, but the results there (such as [12, Theorem 1]) are asymptotic. In this paper we derive finite-sample confidence intervals in natural non-IID settings.

We start in Sect. 2 by giving the definition of the causal effect adapted to the back-door criterion. In Sect. 3 we consider the simplest IID setting giving the narrowest confidence intervals. Limitations of this setting are pointed out in, e.g., [12, Sect. 1] and [5, end of Sect. 3.6.1].

In Sect. 4 we drop the assumption of IID data, which leads to the appearance of an iterated logarithm term in our confidence intervals. This is developed further in Sect. 5; there we make our confidence intervals anytime-valid, which leads to further iterated logarithm terms.

This paper was motivated by the difficulty of applying the methods developed in [13] to the most natural setting of sequential causal inference (which we called the "strong interpretation" of causal diagrams). The methods of this paper are completely different, and we are targeting confidence intervals rather than prediction sets, which were targeted in [13]. In Sect. 6 we briefly discuss derivation of prediction sets from our confidence intervals, but this is likely to lead to much more conservative prediction sets in the IID setting as compared with [13].

Finally, Sect. 7 concludes and lists some directions of further research.

2 Causal effect

Our running example will be the causal diagram in Figure 1, which we now use to explain our notation (in which we follow mainly Pearl [5]). The variables, such as X, Y, and Z, in our causal diagrams will always range over finite sets denoted by the corresponding boldface letters, such as $\mathbf{X}$, $\mathbf{Y}$, and $\mathbf{Z}$, and called their *domains* (equipped with the discrete σ-algebras). However, when talking specifically about the example in Figure 1, we will usually assume that

$\mathbf{X} = \mathbf{Y} = \mathbf{Z} = \{0, 1\}$, so that X, Y, Z are the indicator functions of events.

Let P be a positive probability measure on $\mathbf{X} \times \mathbf{Y} \times \mathbf{Z}$ (generating the random variables X, Y, and Z). Suppose it factorizes according to Figure 1:

$$P(X = x, Y = y, Z = z) \\ = P(Z = z)P(X = x \mid Z = z)P(Y = y \mid X = x, Z = z) \quad (1)$$

for all $x \in \mathbf{X}$, $y \in \mathbf{Y}$, and $z \in \mathbf{Z}$. It will be very convenient to use Pearl's [5, Sects. 1.1.4 and 1.1.5] convention and abbreviate $P(X = x)$ to $P(x)$, $P(Y = y)$ to $P(y)$, etc.; such abbreviated notation will also be used when we have, say, $\tilde{x}$ or x' in place of x. We will also often omit mentioning that $x \in \mathbf{X}$, $y \in \mathbf{Y}$, etc. With this convention, we can rewrite (1) as

$$P(x, y, z) = P(z)P(x \mid z)P(y \mid x, z).$$

Let $\tilde{x} \in \mathbf{X}$. We use Pearl's [5] notation $\mathrm{do}(X = \tilde{x})$, usually abbreviated to $\mathrm{do}(\tilde{x})$, to signify setting X to $\tilde{x}$ (we will define what this means formally only in specific contexts). Let us define, in the context of Figure 1, the *causal effect* of X on Y as

$$P(y \mid \mathrm{do}(\tilde{x})) := \sum_z P(y \mid \tilde{x}, z)P(z). \quad (2)$$

(The definition in [5, Sect. 3.2] is more general, but it is not our focus in this paper and we will use simpler *ad hoc* definitions.) The interpretation of (2) (and of causal effects in general) is that it is the probability of $Y = y$ in the mutilated causal model in which the arrow from Z to X in Figure 1 has been removed and X has been set to $\tilde{x}$.

The decomposition (1) and these notational conventions generalize to any directed acyclic graph (dag), and Figure 1 can be generalized to the following *back-door criterion*, which is stated in terms of "blocking", as defined in [5, Definition 1.2.3]. If X, Y, and Z are disjoint non-empty sets of variables in a dag, Z is said to satisfy the back-door criterion relative to (X, Y) if, for any $X' \in X$ and $Y' \in Y$,

- no vertex in Z is a descendant of X', and
- Z blocks every *backdoor path* from X' to Y', i.e., every path between X' and Y' that contains an arrow into X'

[5, Definition 3.3.1]. If the back-door criterion is satisfied, the *causal effect* can still be defined as (2) [5, Theorem 3.3.2]. However, now the summing $\sum_z$ over z in (2) means summing over all possible values of the variables in Z:

$$\sum_z := \sum_{z_1 \in \mathbf{Z}^1} \cdots \sum_{z_k \in \mathbf{Z}^k},$$

where $\mathbf{Z}^i$ is the domain of Z^i, $i = 1, \ldots, k$, and $Z^1, \ldots, Z^k$ are the elements of Z, $Z = \{Z^1, \ldots, Z^k\}$. Moreover, $\tilde{x}$ should specify the values for all variables in X. In the theorems below we will use the notation

$$|\mathbf{Z}| := \left|\mathbf{Z}^1\right| \ldots \left|\mathbf{Z}^k\right|. \quad (3)$$

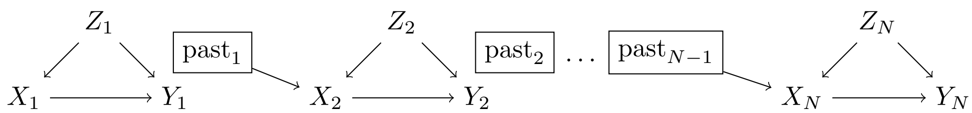


Figure 2: The repeated causal graph

It is very important (but does not concern us in this paper) that some of the variables in the causal dag may be unobservable; it's fine as long as these variables do not enter expressions for causal effects, such as (2).

3 The IID setting

We start from the most standard setting where the observations before intervention are IID; see, e.g., [5, the beginning of Sect. 2.2]. Informally, Nature possesses stable causal mechanisms that are organized in the form of a graphical structure. In causal calculus, the structure is a known dag, and the stable causal mechanisms are unknown probability distributions of the variables in the vertices of the dag given their parents. The available observations are generated in the IID fashion. (The IID nature of the observations usually stays implicit, and it appears that in Pearl's book [5] it is made explicitly only in [5, the end of Sect. 3.6.1].)

Figure 2 shows N repetitions of the causal system represented in Figure 1. Let us ignore the cells labelled past_i, $i \in \{1, \dots, N-1\}$, for now (formally, we are assuming that these variables take a fixed known value). This composite causal diagram then represents N IID observations over the base diagram of Figure 1. In this section we are interested in Figure 2 with an arbitrary dag as the base diagram, not necessarily the one in Figure 1. This IID picture will give the narrowest confidence intervals out of those derived in this paper.

The following theorem treats the case of the back-door criterion, and it is proved (as all other results in Sects 3–5) in Appendix A. A confidence interval $[m-h, m+h]$ will be represented in terms of its *midpoint* m and *half-width* h. The number N of repetitions is fixed, and we set

$$\#xy := |\{n \in [N] : (X_n, Y_n) = (x, y)\}| \tag{4}$$

(i.e., $\#xy$ is the number of times $(X_n, Y_n) = (x, y)$); the analogous notation will be used for sequences other than xy, such as xyz and z. See (3) for the definition of $|\mathbf{Z}|$.

Theorem 1. *Let $\delta > 0$, $\tilde{x} \in \mathbf{X}$, and $y \in \mathbf{Y}$. The following is a $(1-\delta)$-confidence interval for the parameter $P(y \mid \mathrm{do}(\tilde{x}))$ defined by (2): the midpoint is*

$$\sum_z \hat{p}(y \mid \tilde{x}, z)\hat{p}(z) \tag{5}$$

and the half-width is

$$|\mathbf{Z}|\sqrt{\frac{\ln\frac{4|\mathbf{Z}|}{\delta}}{2N}}+\sum_{z}\sqrt{\frac{\ln\frac{4|\mathbf{Z}|}{\delta}}{2\#\tilde{x}z}}, \tag{6}$$

where $\hat{p}(y \mid \tilde{x}, z) := \#\tilde{x}yz/\#\tilde{x}z$ *is the standard estimate for* $P(y \mid \tilde{x}, z)$ *and* $\hat{p}(z) := \#z/N$ *is the standard estimate for* $P(z)$*. (The half-width* (6) *is understood to be* ∞ *when* $\#\tilde{x}z = 0$*.)*

In the case of Figure 1 with binary variables, we can replace (6) by

$$2\sqrt{\frac{\ln\frac{6}{\delta}}{2N}}+\sum_{z\in\{0,1\}}\sqrt{\frac{\ln\frac{6}{\delta}}{2\#\tilde{x}z}} \tag{7}$$

(although this does not quite follow from (6)).

4 The adaptive setting with a fixed horizon

Under the *strong interpretation* of Figure 2, considered in this section, each box past_n stands for the whole past, including the variables X_i, Y_i, and Z_i, $i \in [n]$. Now each X_{n+1}, $n \in [N-1]$, has incoming arrows (not shown explicitly in the figure) from all X_i, Y_i, and Z_i, $i \in [n]$. As before, we allow repetition of any dag in Figure 2, not just the one in Figure 1.

Our interpretation of Figure 2 is that X is a decision that has Y as its result. The decision at step $n+1$ may depend on the past decisions and past values of Y and Z. In other words, the decision maker has access to all past observations. In the case of Figure 1, all Z_n are independent of the past and identically distributed; all Y_n have the same distribution given X_n and Z_n, and they are conditionally independent of the past.

For an integer $n \ge 2$, let $\llbracket n \rrbracket$ be the largest integer of the form 2^k, $k \in \{1, 2, \dots\}$, satisfying $2^k \le n$ (and for $n < 2$, $\llbracket n \rrbracket$ is defined as, say, 1). We let lb stand for binary logarithm $\log_2$, and we will often use it in the context of $\text{lb}\llbracket n \rrbracket = \lfloor \text{lb}\, n \rfloor$ for a positive integer n.

It will be useful to extend the notation (4) and set, e.g.,

$$\#_m xy := \left|\{n \in [m] : (X_n, Y_n) = (x, y)\}\right|,$$

where m may be different from N. Earlier we defined standard estimates such as $\hat{p}(y \mid \tilde{x}, z) := \#\tilde{x}yz/\#\tilde{x}z$ (and we will refrain from defining $\hat{p}$ in other similar contexts in the following theorems). We will also need the modification of $\hat{p}(y \mid \tilde{x}, z)$ defined by

$$\hat{\hat{p}}(y \mid \tilde{x}, z) := \frac{\left|\{n \in [N] : \#_n\tilde{x}z \le \llbracket \#\tilde{x}z \rrbracket, X_n = \tilde{x}, Y_n = y, Z_n = z\}\right|}{\llbracket \#\tilde{x}z \rrbracket}. \tag{8}$$

In words, $\hat{\hat{p}}(y \mid \tilde{x}, z)$ is the fraction of the first $\llbracket \#\tilde{x}z \rrbracket$ observations with $X_n = \tilde{x}$ and $Z_n = z$ for which $Y_n = y$. It is also an estimate of $P(y \mid \tilde{x}, z)$, but it might not use all the available data (however, it uses at least one half of the relevant observations). We will also use the notation $\hat{\hat{p}}$ in other contexts, such as $\hat{\hat{p}}(z \mid \tilde{x})$.

Now we have to replace the Hoeffding inequality used in our proof of Theorem 1 in Appendix A by the law of the iterated logarithm, and so we will get our first iterated logarithm term.

Theorem 2. *Fix $\delta > 0$, $\tilde{x}$, and y; the time horizon N is also fixed. Under the strong interpretation, the following is a $(1-\delta)$-confidence interval for the parameter $P(y \mid \mathrm{do}(\tilde{x}))$ defined by (2): the midpoint is*

$$\sum_z \hat{p}(z)\hat{\hat{p}}(y \mid \tilde{x}, z) \tag{9}$$

and the half-width is

$$|\mathbf{Z}| \sqrt{\frac{\ln \frac{4|\mathbf{Z}|}{\delta}}{2N}} + \sum_z \sqrt{\frac{2 \ln \mathrm{lb} \llbracket \# \tilde{x} z \rrbracket + \ln \frac{6.6|\mathbf{Z}|}{\delta}}{2 \llbracket \# \tilde{x} z \rrbracket}}. \tag{10}$$

The case $\#\tilde{x}z \in \{0, 1\}$ in (10) requires special treatment; namely, we set $\ln \mathrm{lb}\, 1 := \infty$, and so (10) is interpreted as ∞ unless $\#\tilde{x}z \geq 2$ for all z.

For the binary case of Figure 1, we can replace (10) by

$$2\sqrt{\frac{\ln \frac{6}{\delta}}{2N}} + \sum_{z \in \{0,1\}} \sqrt{\frac{2 \ln \mathrm{lb} \llbracket \# \tilde{x} z \rrbracket + \ln \frac{10}{\delta}}{2 \llbracket \# \tilde{x} z \rrbracket}}, \tag{11}$$

similarly to (7).

5 The anytime-valid adaptive setting

Theorem 2 can be easily extended to the setting in which the time horizon N is not fixed in advance. Now we would like our results to be anytime valid, with N ranging over the positive integers $\{1, 2, \dots\}$. Since N is variable, now we will write $\hat{\hat{p}}_N(y \mid \tilde{x}, z)$ in place of $\hat{\hat{p}}(y \mid \tilde{x}, z)$ defined by (8) and add the lower index N in other similar places; in particular, $\hat{\hat{p}}_N(z) := \#_{\llbracket N \rrbracket} z / \llbracket N \rrbracket$. Remember that a $(1-\delta)$-*confidence sequence* is a sequence of confidence intervals whose intersection covers the true parameter value with probability at least $1 - \delta$.

Theorem 3. *Let $\delta > 0$, $\tilde{x} \in \mathbf{X}$, and $y \in \mathbf{Y}$. The following is a $(1-\delta)$-confidence sequence for the parameter $P(y \mid \mathrm{do}(\tilde{x}))$ defined by (2): the midpoint is*

$$\sum_z \hat{\hat{p}}_N(z)\hat{\hat{p}}_N(y \mid \tilde{x}, z)$$

and the half-width is

$$|\mathbf{Z}| \sqrt{\frac{2 \ln \mathrm{lb} \llbracket N \rrbracket + \ln \frac{6.6|\mathbf{Z}|}{\delta}}{2 \llbracket N \rrbracket}} + \sum_z \sqrt{\frac{2 \ln \mathrm{lb} \llbracket \#_N \tilde{x} z \rrbracket + \ln \frac{6.6|\mathbf{Z}|}{\delta}}{2 \llbracket \#_N \tilde{x} z \rrbracket}}. \tag{12}$$

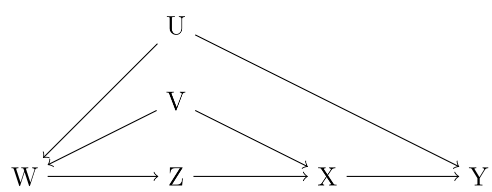


Figure 3: The napkin graph

In the binary case of Figure 1, (12) can be slightly strengthened to

$$2\sqrt{\frac{2\ln\mathrm{lb}\llbracket N\rrbracket + \ln\frac{10}{\delta}}{2\llbracket N\rrbracket}} + \sum_{z\in\{0,1\}}\sqrt{\frac{2\ln\mathrm{lb}\llbracket \#_N\tilde{x}z\rrbracket + \ln\frac{10}{\delta}}{2\llbracket \#_N\tilde{x}z\rrbracket}}.$$

Theorem 2–3 may be considered to be finite-sample analogues of Theorem 1 in [12]. Their characteristic feature is the presence of iterated logarithm terms, which are unavoidable and are especially prominent under the strong interpretation.

Remark 4. In this paper we only discuss, outside of this remark, causal effects that are representable as arithmetic expressions involving only two arithmetic operations, plus and multiplication (minus could be added for free but is not useful). There are, however, situations in which the causal effect is given in a form involving division, such as the *napkin graph*, shown in Figure 3. It has the following expression for the causal effect of X on Y:

$$P(y \mid \mathrm{do}(\tilde{x})) := \frac{\sum_w P(y, \tilde{x} \mid z, w)P(w)}{\sum_w P(\tilde{x} \mid z, w)P(w)} \tag{13}$$

(see [2, Figure 2(c) and (1)]). The expression (13) is a ratio, and our methods do not work for it. However, we can still compute the left and right end-points of the overall confidence interval by combining the left and right end-points of the constituent confidence intervals.

An interesting feature of the expression (13) is that its right-hand side involves z but does not really depend on it, as is clear from its left-hand side; this is an instance of so-called Verma constraints [1]. The presence of z on the right-hand side of (13) is in a certain sense inevitable; formally, Z is a "trapdoor variable" as defined in [2, Definition 3] and explained in [2, Sect. 2.3].

6 Applications to prediction sets

Theorems 1–3 provide confidence intervals for causal effects, whereas in [13] we were interested in prediction sets for Y. In order to discuss connections between our results here and the strong interpretation in [13], in this section we will state a corollary of the toy version of Theorem 2 for the binary case of Figure 1, with (10) replaced by (11), giving prediction sets; similar corollaries can be easily deduced from Theorems 1–3 as well. We consider the strong interpretation of Figure 1, as in Sect. 4, with (X_n, Y_n, Z_n), $n \in [N]$, complemented by another observation Y with the probabilities of $Y = y$, $y \in \mathbf{Y}$, given by the right-hand side of (2) for a fixed $\tilde{x}$. Remember that $\mathbf{X} = \mathbf{Y} = \mathbf{Z} = \{0, 1\}$.

Corollary 5. *Fix $\delta > 0$, N, and $\tilde{x} \in \mathbf{X}$. Then*

$$\Gamma := \left\{ y \in \mathbf{Y} : \sum_{z \in \{0,1\}} \hat{p}(z)\hat{\hat{p}}(y \mid \tilde{x}, z) + 2\sqrt{\frac{\ln \frac{12}{\delta}}{2N}} \right.$$
$$\left. + \sum_{z \in \{0,1\}} \sqrt{\frac{2 \ln \operatorname{lb}[\![\#\tilde{x}z]\!] + \ln \frac{20}{\delta}}{2[\![\#\tilde{x}z]\!]}} > \frac{\delta}{2} \right\}$$

is a $(1 - \delta)$-prediction set.

Proof. We are required to prove that $Y \notin \Gamma$ with probability at most δ. Let us fix $y \in \mathbf{Y}$ and prove that the probability of the conjunction of $Y = y$ and $y \notin \Gamma$ is at most $\delta/2$. We will use the confidence interval (9) $\pm$ (11) with $\delta/2$ in place of δ. Consider two cases:

- Suppose the right-hand-side of (2) exceeds $\delta/2$. If $y \notin \Gamma$, the right end-point (9) + (11) (with δ replaced by $\delta/2$) of the confidence interval is at most $\delta/2$, and the probability of this is at most $\delta/2$ by the definition of a confidence interval.
- Otherwise, $Y = y$ with probability at most $\delta/2$ by the definition of Y. □

This procedure is sub-optimal for several reasons. One of them is that Hoeffding's inequality applied to the Bernoulli model can be greatly improved when the probability of error is small; see, e.g., Vapnik's [9, Sects. 4.2 and 4.4] use of multiplicative Chernoff inequalities (in what he calls optimistic and pessimistic settings of learning problems).

Analogous corollaries of Theorem 1 become more comparable with the results that we obtain in [13]. However, the prediction sets derived in [13] are based on e-values (are "e-prediction sets") whereas the prediction sets here are traditional ones. We expect that methods of this paper lead to much looser results, but their advantage is that they also work for the strong interpretation.

7 Conclusion

In this paper we derive confidence intervals and confidence sequences for causal effects in IID and sequential non-IID settings. These are some directions of further research:

- The results of this paper only provide upper bounds for achievable widths of confidence intervals, and they need to be complemented by lower bounds. (It is easy to check that a $\log\log$ term is unavoidable even in the adaptive setting of Figure 1 with $\mathbf{X} = \mathbf{Y} = \{0, 1\}$ and $\mathbf{Z} = \{0\}$.)
- This paper is based on the standard measure-theoretic probability [3,7,8]. To make our results as strong as possible, we could try and present them in the language of game-theoretic probability [6] (this was listed as direction of further research already in [12, Sect. 5]).

References

[1] Rohit Bhattacharya and Razieh Nabi. On testability of the front-door model via Verma constraints. In *Proceedings of the 38th Conference on Uncertainty in Artificial Intelligence (UAI 2022)*, volume 180 of *Proceedings of Machine Learning Research*, pages 202–212, 2022.

[2] Jouni Helske, Santtu Tikka, and Juha Karvanen. Estimation of causal effects with small data in the presence of trapdoor variables. *Journal of the Royal Statistical Society A*, 184:1030–1051, 2021.

[3] Andrei N. Kolmogorov. *Grundbegriffe der Wahrscheinlichkeitsrechnung.* Springer, Berlin, 1933. English translation: *Foundations of the Theory of Probability.* Chelsea, New York, 1950.

[4] Masashi Okamoto. Some inequalities relating to the partial sum of binomial probabilities. *Annals of the Institute of Statistical Mathematics*, 10:29–35, 1959.

[5] Judea Pearl. *Causality: Models, Reasoning, and Inference.* Cambridge, Cambridge University Press, second edition, 2009.

[6] Glenn Shafer and Vladimir Vovk. *Game-Theoretic Foundations for Probability and Finance.* Wiley, Hoboken, NJ, 2019.

[7] Albert N. Shiryaev. *Probability-1.* Springer, New York, third edition, 2016.

[8] Albert N. Shiryaev. *Probability-2.* Springer, New York, third edition, 2019.

[9] Vladimir N. Vapnik. *Statistical Learning Theory.* Wiley, New York, 1998.

[10] Jean Ville. *Etude critique de la notion de collectif.* Gauthier-Villars, Paris, 1939.

[11] Václav Voráček. Treatment of statistical estimation problems in randomized smoothing for adversarial robustness. In *NeurIPS 2024*, 2024.

[12] Vladimir Vovk. Another semantics for Pearl's action calculus. In Alex Gammerman, editor, *Computational Learning and Probabilistic Reasoning*, chapter 7, pages 127–146. Wiley, New York, 1996.

[13] Vladimir Vovk and Ruodu Wang. Conformal e-prediction in the presence of confounding. Technical Report arXiv:2603.11134 [math.ST], arXiv.org e-Print archive, March 2026. For the latest version, see alrw.net, Working Paper 46.

A Proofs

In the following two propositions we consider an IID Bernoulli sequence $\xi_1, \xi_2, \ldots$ with probability of success p and use the notation $\hat{p}_n := \frac{1}{n}\sum_{i=1}^n \xi_i$ for the standard estimate of p. We start from a standard confidence interval for the probability of success given by Hoeffding's inequality.

Proposition 6. *For a fixed n and $\delta > 0$,*

$$I := \left\{ p : |p - \hat{p}_n| < \sqrt{\frac{\ln \frac{2}{\delta}}{2n}} \right\} \tag{14}$$

is a $(1-\delta)$-confidence interval for p, in the sense of

$$\mathbb{P}\left(p \in I\right) \geq 1 - \delta.$$

Proof. By Hoeffding's inequality (or Okamoto's earlier result [4, Theorem 1]), for all $c > 0$,

$$\mathbb{P}\left(|p - \hat{p}_n| \geq c\right) \leq 2\exp(-2c^2 n),$$

which gives the confidence interval (14). □

We will also need the following confidence sequence.

Proposition 7. *For each $\delta > 0$,*

$$I_n := \left\{ p : \left|p - \hat{p}_{\llbracket n \rrbracket}\right| < \sqrt{\frac{2 \ln \operatorname{lb}\llbracket n \rrbracket + \ln \frac{3.3}{\delta}}{2\llbracket n \rrbracket}} \right\} \tag{15}$$

is a $(1-\delta)$-confidence sequence, i.e.,

$$\mathbb{P}\left(\forall n : p \in I_n\right) \geq 1 - \delta. \tag{16}$$

Proof. Similar confidence sequences can be obtained using Ville's [10] method of continuous mixtures of test martingales or its discrete analogue [6, Sect. 5.1], but we will model our proof on [11, Sect. E]. Fix $p \in [0, 1]$.

Since $\zeta(2) = \pi^2/6$, we can split the significance level δ into the series $\delta = \sum_{k=1}^{\infty} \delta_k$, where

$$\delta_k = \frac{6}{\pi^2 k^2}\delta.$$

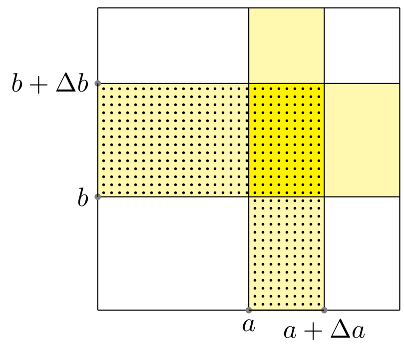


Figure 4: Illustration of an inequality.

Applying (14) with $n_k = 2^k$ in place of n and δ_k in place of δ gives

$$\mathbb{P}\left(|p - \hat{p}_{n_k}| \geq \sqrt{\frac{\ln \frac{\pi^2 k^2}{3\delta}}{2n_k}} \right) \leq \delta_k. \tag{17}$$

Finally, (17) implies (16) since $\pi^2/3 \approx 3.29 < 3.3$. (This argument works for $n \geq 2$; otherwise, the inequality in (16) is trivial since our convention, introduced in Sect. 4, is that $\ln \operatorname{lb} 1 := \infty$.) $\square$

Next we need a simple result from interval arithmetic. We are only interested in subintervals of $[0,1]$. Let $c \pm \Delta c$, where $c \in \mathbb{R}$ and $\Delta c \geq 0$, stand for the interval

$$c \pm \Delta c := [c - \Delta c, c + \Delta c] \cap [0,1].$$

For a binary operation $*$ on the reals (we are mostly interested in addition and multiplication), we define its result on intervals pointwise:

$$I_1 * I_2 := \{p_1 * p_2 : p_1 \in I_1, p_2 \in I_2\} \cap [0,1].$$

Lemma 8. *For any two intervals $a \pm \Delta a$ and $b \pm \Delta b$,*

$$(a \pm \Delta a) + (b \pm \Delta b) \subseteq (a + b) \pm (\Delta a + \Delta b), \tag{18}$$
$$(a \pm \Delta a) \times (b \pm \Delta b) \subseteq (a \times b) \pm (\Delta a + \Delta b). \tag{19}$$

Proof. The inclusion (18) is obvious, so we will only prove (19). The latter inclusion reduces to the conjunction of two inequalities:

$$(a - \Delta a)(b - \Delta b) \geq ab - (\Delta a + \Delta b), \tag{20}$$
$$((a + \Delta a) \wedge 1)((b + \Delta b) \wedge 1) \leq ab + (\Delta a + \Delta b). \tag{21}$$

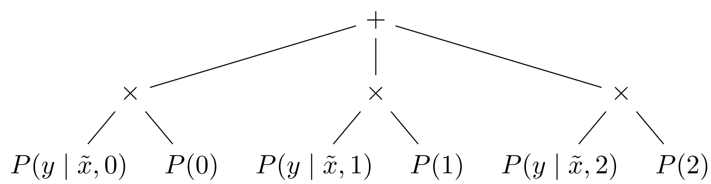


Figure 5: The binary tree representing the polynomial expression (2) over the formal variables $P(y \mid \tilde{x}, z)$ and $P(z)$ in the case of $\mathbf{Z} = \{0, 1, 2\}$

The inequality (20) is obvious, and in (21) we can assume, without loss of generality, $a + \Delta a \leq 1$ and $b + \Delta b \leq 1$, which reduces it to

$$(a + \Delta a)(b + \Delta b) - ab \leq \Delta a + \Delta b. \tag{22}$$

The last inequality is illustrated in Figure 4: the dotted area of the plot represents the left-hand side of (22), and the yellow area represents the right-hand side of (22) (with the darker yellow area counted twice). □

We will distinguish between (multivariate) polynomials and polynomial expressions. A *polynomial expression* is formed from a finite number of formal variables by repeatedly applying the operations of multiplication and addition; we do not allow constants (equivalently, our polynomials and polynomial expressions are over the field $\{0, 1\}$). A polynomial expression can be represented as a tree (let us call it a *polynomial tree*), not necessarily binary, such as Figure 5 in the case of the polynomial expression given by the right-hand side of (2) for $\mathbf{Z} = \{0, 1, 2\}$. A *polynomial* is an equivalence class of polynomial expressions where we do not distinguish polynomial expressions that reduce to each other by applying the usual laws of commutativity, associativity, and distributivity (associativity was already used implicitly when we allowed non-binary multiplications and additions in our trees). Without loss of generality we may assume that the operations at different levels of polynomial trees alternate (so at each level we have the same operation, "+" or "×", and perhaps some formal variables as leaf nodes; the operations in adjecent levels are different).

Corollary 9. *Let E be a polynomial expression involving m distinct formal variables. Define E^* to be the polynomial obtained from E by replacing all multiplications by additions. Then, for any intervals $a_i \pm \Delta a_i$, $i = 1, \ldots, m$,*

$$E(a_1 \pm \Delta a_1, \ldots, a_m \pm \Delta a_m) \subseteq E(a_1, \ldots, a_m) \pm E^*(\Delta a_i). \tag{23}$$

The expression $E^*(\Delta a_i)$ in (23) is, of course, the result of substituting Δa_i for the formal variables in E^* and evaluating the resulting expression.

Proof of Corollary 9. We proceed by induction on the height of E considered as polynomial tree. Repeatedly applying (18), we can extend it to any finite sums.

Similarly, repeatedly applying (19), we can extend it to any finite products. This gives the statement (23) when the height of E is 1. The inductive step is also provided by extensions of (18) and (19) to finite sums and products. $\square$

Proof of Theorem 1. In the proofs of Theorems 1–3 we will use the slightly informal notation exemplified by (5) ± (6) being the confidence interval with midpoint (5) and half-width (6).

We obtain the confidence interval (5) ± (6) for (2) (involving $2\,|\mathbf{Z}|$ probabilities) by combining the confidence intervals (14) for each of the $2\,|\mathbf{Z}|$ constituent probabilities. To ensure the overall confidence level $1-\delta$, we replace the δ in (14) by $\delta/(2\,|\mathbf{Z}|)$. By Corollary 9 applied to (2), we then indeed obtain the overall $(1-\delta)$-confidence interval with midpoint (5) and semi-width

$$\sum_z \left(\sqrt{\frac{\ln \frac{4|\mathbf{Z}|}{\delta}}{2N}} + \sqrt{\frac{\ln \frac{4|\mathbf{Z}|}{\delta}}{2\#\tilde{x}z}} \right), \tag{24}$$

i.e., (6). The two square roots in (24) are the half-widths of the confidence intervals for $P(z)$ and $P(y \mid \tilde{x}, z)$, respectively, given by Proposition 6. $\square$

To derive (7) for Figure 1 with binary variables, notice that in the binary case we only need confidence intervals for three constituent probabilities, since an interval estimate for $P(Z=0)$ gives one for $P(Z=1)$ and vice versa. This allows us to replace δ by $\delta/3$ rather than $\delta/4$ in (14).

Proof of Theorem 2. We apply Proposition 6 to estimating $P(z)$ and Proposition 7 to estimating $P(y \mid \tilde{x}, z)$ in (2). Since the total number of distinct formal variables in (2) is $2\,|\mathbf{Z}|$, we replace the δ in (14) and (15) by $\delta/(2\,|\mathbf{Z}|)$. Since (9) is obviously the midpoint of the confidence interval given by Corollary 9, we only check that (10) is its half-width.

Plugging the half-width of the confidence interval for $P(z)$ given by Proposition 6 and the half-width of the confidence interval for $P(y \mid \tilde{x}, z)$; given by Proposition 7 into the polynomial expression (2) with the multiplications replaced by additions, we obtain the overall half-width

$$\sum_z \left(\sqrt{\frac{\ln \frac{4|\mathbf{Z}|}{\delta}}{2N}} + \sqrt{\frac{2 \ln \mathrm{lb} \llbracket \#\tilde{x}z \rrbracket + \ln \frac{6.6|\mathbf{Z}|}{\delta}}{2\llbracket \#\tilde{x}z \rrbracket}} \right)$$

i.e., (10). $\square$

In the case of Figure 1 with binary variables, we obtain (11) if we again replace δ by $\delta/3$ rather than $\delta/4$ (and round up 9.9 to 10).

Proof of Theorem 3. The proof of Theorem 3 is analogous, and the only difference is that we use the confidence sequence (15) for estimating $P(z)$ as well. $\square$