

# $\rho$ -GNF: A Copula-based Sensitivity Analysis to Unobserved Confounding Using Normalizing Flows

**Sourabh Balgi**

*STIMA, IDA, Linköping University, Sweden*

SOURABH.BALGI@LIU.SE

**Jose M. Peña**

*STIMA, IDA, Linköping University, Sweden*

JOSE.M.PENA@LIU.SE

**Adel Daoud**

*IAS, IEI, Linköping University, Sweden*

ADEL.DAOU@LIU.SE

**Editors:** J.H.P. Kwisthout & S. Renooij

## Abstract

We propose a novel sensitivity analysis to unobserved confounding in observational studies using copulas and normalizing flows. Using the idea of interventional equivalence of structural causal models, we develop  $\rho$ -GNF ( $\rho$ -graphical normalizing flow), where  $\rho \in [-1, +1]$  is a bounded sensitivity parameter. This parameter represents the back-door non-causal association due to unobserved confounding, and which is encoded with a Gaussian copula. In other words, the  $\rho$ -GNF enables scholars to estimate the average causal effect (ACE) as a function of  $\rho$ , while accounting for various assumed strengths of the unobserved confounding. The output of the  $\rho$ -GNF is what we denote as the  $\rho_{curve}$  that provides the bounds for the ACE given an interval of assumed  $\rho$  values. In particular, the  $\rho_{curve}$  enables scholars to identify the confounding strength required to nullify the ACE, similar to other sensitivity analysis methods (e.g., the E-value). Leveraging on experiments from simulated and real-world data, we show the benefits of  $\rho$ -GNF. One benefit is that the  $\rho$ -GNF uses a Gaussian copula to encode the distribution of the unobserved causes, which is commonly used in many applied settings. This distributional assumption produces narrower ACE bounds compared to other popular sensitivity analysis methods.

**Keywords:** Sensitivity analysis; unconfoundedness; structural causal model; normalizing flow; Gaussian copula.

## 1. Introduction

Epidemiologists, sociologists, economists, and other applied scientists, often leverage randomized controlled trials (RCTs), as RCTs provide the safest methodological route to disentangle cause and effect. RCTs are the *gold-standard* as they require the least assumptions (Wright, 1921; Fisher, 1936; Cox, 1958; Imbens and Rubin, 2015). Most importantly, by randomizing which experimental subjects (e.g., people, villages, schools) should take the treatment and which subjects should abstain, an RCT provides unconditional ignorability or exchangeability or unconfoundedness. *Unconfoundedness* implies no unobserved confounders in the causal system of interest: no unobserved common causes of the treatment and the outcome. When unconfoundedness is satisfied, scholars can calculate the causal effect of interest from collected data; that means that the causal quantity is identified (Robins, 1986; Rubin, 1990; Robins and Hernán, 2008; Pearl, 2009b; Hernán and Robins, 2009). However,

despite the importance of the RCT design, it remains infeasible for a slew of applied settings. It may be costly to implement (e.g., testing a population-wide medicine); it may be unethical (e.g., testing a new drug); or, it may be impractical to implement (e.g., testing a social policy across the world). Therefore, applied researchers often rely on observational data – which are often secondary data sources with no treatment randomization and where the experimenter had no control over the data generating process. Yet when using observational data, scholars make themselves susceptible for failing to satisfy the unconfoundedness assumption, even when some confounders are observed.

Because the unconfoundedness assumption is so critical and at the same time untestable in observational studies (Rubin, 1990; D’Onofrio et al., 2020), methodologists (statisticians, computer scientist, and others) have developed various frameworks for stress testing how causal effect estimates change under varying strength of unconfoundedness failure. These sort of tests are named *sensitivity analysis* (Schlesselman, 1978; Manski, 1990; Imbens, 2003; Brumback et al., 2004; VanderWeele and Arah, 2011), also known as *bias analysis* in epidemiology (Cornfield et al., 1959; Cochran and Rubin, 1973; Rothman et al., 2008; Lash et al., 2009). Nonetheless, existing sensitivity analysis frameworks are limited in at least three ways, and our proposed  $\rho$ -GNF improve on these limitations, thereby moving the state-of-the-art forward.

First, with the idea of interventional equivalence of structural causal models (SCMs), we propose a deep-learning method for sensitivity analysis based on the graphical normalizing flow (GNF) (Wehenkel and Louppe, 2021), causal graphical normalizing flow (c-GNF) (Baldi et al., 2022a) and Causal Normalizing Flow (CNF) (Javaloy et al., 2023), because of GNF’s attractive properties of non-linearity and invertibility for counterfactual inference and the similarities to the most well studied and popular elliptical copula, i.e., Gaussian copula. Hence, we aptly name the model  $\rho$ -GNF, where  $\rho \in [-1, +1]$  is the bounded sensitivity parameter of the Gaussian copula that controls the degree of unconfoundedness between the observed treatment and outcome. Unlike most sensitivity analysis methods where the sensitivity parameters are unbounded and difficult to specify and interpret,  $\rho$  is bounded in the range  $[-1, +1]$  and it represents the non-causal dependence between the treatment and outcome. Second, we show that with interventional equivalence,  $\rho$ -GNF enables us to estimate the interventional causal effects such as average causal effect (ACE) as a function of  $\rho$ . We call this the  $\rho_{curve}$ , and this curve enables us to identify the ACE bounds and analyze the confounding strength required to explain away the causal effect. Thus, the  $\rho_{curve}$  enables us to provide bounds for the causal effect given a specific interval of  $\rho \in [-1, +1]$  that the domain expert considers appropriate. We also define  $\rho_{value}$  as the value of  $\rho$  that explains away the causal effect and demonstrate its similarities to the widely used E-value (VanderWeele and Ding, 2017). Third, we empirically demonstrate sharp and narrower bounds compared to the widely popular assumption-free bounds through simulated as well as real-world experiments. Unlike existing sensitivity analysis methods that provide bounds either only for discrete or continuous outcomes,  $\rho$ -GNF accommodates both discrete and continuous outcomes.

Our work proceeds as follows. After briefly reviewing related literature in Section 1.1, we define our notation and the main objective of ACE estimation, using the idea of interventional equivalence of SCMs with bivariate Gaussian copula in Section 2.1. We perform the identification and estimation of the ACE and analyse the sensitivity to the different de-

gree of unconfoundedness in Section 2.2-2.3, via the sensitivity parameter  $\rho$ . In Section 3, we present our results with simulated as well as real-world data under different settings of outcome variable (i.e., continuous or binary or categorical), and compare them with the popular assumption-free (AF) bounds. Finally, in Section 4, we conclude with discussing the key contributions of our  $\rho$ -GNF method in encouraging the use of sensitivity analysis when working with non-randomized observational data.

### 1.1. Background and Related Work

The sensitivity analysis literature can be roughly categorized into two streams: (i) identify the *bounds* of the causal effect as functions of some sensitivity parameters that encode the strength of the unobserved confounders (Robins, 1989; Manski, 1990; VanderWeele and Arah, 2011; Ding and VanderWeele, 2016; Sjölander, 2020; Sjölander and Hössjer, 2021; Peña, 2022); and (ii) identify how large the influence of the unobserved confounders needs to be to *explain-away* the causal effect (Imbens, 2003; VanderWeele and Ding, 2017; Veitch and Zaveri, 2020; Sjölander and Greenland, 2022). While Robins (1989); Manski (1990) provide assumption-free (AF) bounds of the causal effect for binary outcome, more recent methods such as Ilse et al. (2021) extend the bounds to categorical outcomes. Other methods provide bounds as functions of sensitivity parameters to be tuned by a domain expert (Sjölander, 2020; Sjölander and Hössjer, 2021; Peña, 2022). Cinelli et al. (2019); Cinelli and Hazlett (2020) study sensitivity to unobserved confounding in a linear SCM setting. In contrast to the *bounds* stream, there are methods that fall under the *explain-away* stream. For example, Imbens (2003); VanderWeele and Ding (2017) reason on the lines of the minimum strength of the unmeasured confounder that is needed, conditional on the measured confounders, to explain-away the estimated causal effect. Similar to Imbens (2003) and the E-value (VanderWeele and Ding, 2017), the recently developed *Austen plots* (Veitch and Zaveri, 2020) identify the influence of the confounding needed to explain a specific amount of bias in the causal effect estimate.

While there exists a wide spectrum of sensitivity analysis methods with unique advantages, they are not without limiting assumptions. For example, Robins (1989); Manski (1990); Sjölander (2020); Sjölander and Hössjer (2021); Peña (2022) require the outcome variable being binary, and other methods assume a specific type of parametric model, e.g., Cinelli et al. (2019); Cinelli and Hazlett (2020) assume a linear parametric model. While some methods offer sharp ACE bounds (Robins, 1989; Manski, 1990; Sjölander, 2020; Sjölander and Hössjer, 2021; Peña, 2022), method such as VanderWeele and Ding (2017) may result in wider bounds than the AF bounds, as shown by Ioannidis et al. (2019) and Sjölander (2020). Some methods are exclusively suited for specific causal estimands, e.g., ACE or conditional ACE (CACE) or mediation effects (Tchetgen and Shpitser, 2012; Lindmark et al., 2018). While Robins (1989); Manski (1990) offer no sensitivity parameters that can explain a certain causal effect, VanderWeele and Ding (2017); Veitch and Zaveri (2020) offer multiple parameters that are unbounded and hard to specify for the domain analyst (Ioannidis et al., 2019).

To summarize, even though there exists several sensitivity analysis frameworks, there is still a lack of unifying method that is flexible enough that can suit many different types of observational data, with easy-to-use sensitivity parameters, and that can be applied to

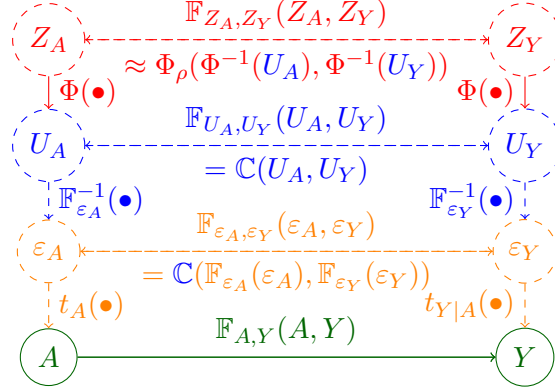


Figure 1: Observationally and interventionally equivalent SCM representations from Eqs. (1)-(4). The solid-lined nodes indicate observed treatment/outcome variables and the dash-lined nodes indicate unobserved cause/noise variables. The green, orange, blue and red lines respectively indicate the desired front-door causal association  $A \rightarrow Y$ , non-causal association due to “effective” noises  $\varepsilon_X$ , non-causal association due to uniform marginal noises  $U_X \in [0, 1]$ , non-causal association due to Gaussian marginal noises  $Z_X$ .

not only binary outcome variables but also categorical and continuous outcomes. Moreover, it’s imperative to establish a method enabling researchers to specify distributional assumptions regarding the unobserved causes within the causal system under study. Such assumptions enable tighter ACE bounds, enhancing the certainty of outcomes. Our  $\rho$ -GNF method targets all these lacks. Our method use deep neural networks, allowing for maximum flexibility and non-linearity.  $\rho$ -GNF provides a single, bounded sensitivity parameter  $\rho \in [-1, +1]$  that is easily interpreted as the measure of non-causal association due to unobserved confounders. Thus, our method enhances an applied researcher’s causal toolbox, and shows how deep learning can further causal inference.

## 2. Notation and Problem Definition

Let us consider the standard *canonical representation* (in Chapter 3.4 of Peters et al. (2017)) for the structural causal model (SCM) (Wright, 1921; Haavelmo, 1943; Goldberger, 1972; Fienberg and Duncan, 1975; Tabar et al., 2022), where  $A$  is the treatment (or cause) and  $Y$  is the outcome (or effect), and  $\varepsilon_A$  and  $\varepsilon_Y$  are their respective unobserved causes such that

$$A := t_A(\varepsilon_A) , Y := t_Y(A, \varepsilon_Y) = t_{Y|A}(\varepsilon_Y) , (\varepsilon_A, \varepsilon_Y) \sim \mathbb{F}_{\varepsilon_A, \varepsilon_Y}(\varepsilon_A, \varepsilon_Y) , \quad (1)$$

where  $(\varepsilon_A, \varepsilon_Y)$  follows the joint CDF  $\mathbb{F}_{\varepsilon_A, \varepsilon_Y}(\varepsilon_A, \varepsilon_Y)$  and is not limited to the usual assumptions of standard normal or uniform random variables as in most common SCM definitions. Using the universality of the uniform (also know as the probability integral transform) (Angus, 1994), the noise variables  $\varepsilon_A$  and  $\varepsilon_Y$  of the SCM in Eq. (1) can equivalently be written in terms of uniform variables  $U_A$  and  $U_Y$  in the interval  $[0, 1]$  resulting in the SCM for Figure 1 (orange+blue) as

$$A := t_A(\mathbb{F}_{\varepsilon_A}^{-1}(U_A)) , Y := t_{Y|A}(\mathbb{F}_{\varepsilon_Y}^{-1}(U_Y)) , (U_A, U_Y) \sim \mathbb{F}_{U_A, U_Y}(U_A, U_Y) , \quad (2)$$

where  $\mathbb{F}_{\varepsilon_A}$  and  $\mathbb{F}_{\varepsilon_Y}$  respectively denote the marginal CDFs of  $\varepsilon_A$  and  $\varepsilon_Y$ , and  $(U_A, U_Y)$  follows the joint CDF  $\mathbb{F}_{U_A, U_Y}(U_A, U_Y)$  with uniform marginals in  $[0,1]$ . From the universality of the uniform, Eq. (2) can further be simplified in terms of  $\mathbb{F}_A$  and  $\mathbb{F}_{Y|A}$  that denote the marginal CDFs of  $A$  and  $Y$  conditioned on  $A$ , respectively as below.

$$A := \mathbb{F}_A^{-1}(U_A) , Y := \mathbb{F}_{Y|A}^{-1}(U_Y) , (U_A, U_Y) \sim \mathbb{F}_{U_A, U_Y}(U_A, U_Y) . \quad (3)$$

Further, we represent  $U_A$  and  $U_Y$  in Eq. (3) as transformation of standard normal variables using the CDF of the standard normal  $\Phi$ , as shown in Figure 1 (orange+blue+red), as

$$A := \mathbb{F}_A^{-1}(\Phi(Z_A)) , Y := \mathbb{F}_{Y|A}^{-1}(\Phi(Z_Y)) , (Z_A, Z_Y) \sim \mathbb{F}_{Z_A, Z_Y}(Z_A, Z_Y) , \quad (4)$$

where  $(Z_A, Z_Y)$  follows the joint CDF  $\mathbb{F}_{Z_A, Z_Y}(Z_A, Z_Y)$  with standard normal marginals. Eqs. (1)-(4) represent observationally and interventionally equivalent SCMs from Figure 1, but with different unobserved noises and corresponding joint CDFs. *Observational equivalence* means that the researcher specified model yields the same distribution as observed by the true (nature or God) data-generating process  $\mathbb{F}_{A, Y}(A, Y)$ ; *interventional equivalence* means that the researcher specified model follows the same interventional distribution  $\mathbb{F}_Y(Y|do(a))$  and  $\mathbb{F}_A(A|do(y))$  (Mooij et al., 2016). Since these noises (distributions) are unknown or unobserved, the completeness of *do*-calculus (Tian and Pearl, 2002; Huang and Valertorta, 2006; Pearl, 2012) states that the causal effects of interest are not identifiable from any of the equivalent SCMs in Eqs. (1)-(4), without further assumptions. To achieve causal effect/estimand identification and bounds, in the subsequent sections, we propose to use a bivariate Gaussian copula, to model the back-door non-causal noise distributions so that the causal effects can be parametrically estimated using deep-neural-network-inspired normalizing flows trained only on observational data.

## 2.1. Representing the Back-Door Non-Causal Association using a Bivariate Gaussian Copula

A copula is a multivariate distribution function defined on the unit hypercube with uniform marginals (Sklar, 1959, 1973). As the name suggests, a copula ‘ties’ or ‘links’ or ‘couples’ a multidimensional joint distribution to its marginals (Nelsen, 2007). From the result of Sklar’s Theorem, we have that the bivariate joint CDF  $\mathbb{F}_{U_A, U_Y}(U_A, U_Y)$  in Eqs. (2) and (3) with uniform marginals in  $[0,1]$  can be represented using a bivariate copula  $\mathbb{C}(U_A, U_Y)$ .

$$\mathbb{F}_{U_A, U_Y}(U_A, U_Y) = \mathbb{C}(U_A, U_Y) . \quad (5a)$$

$$\mathbb{F}_{\varepsilon_A, \varepsilon_Y}(\varepsilon_A, \varepsilon_Y) = \mathbb{C}(\mathbb{F}_{\varepsilon_A}(\varepsilon_A), \mathbb{F}_{\varepsilon_Y}(\varepsilon_Y)) . \quad (5b)$$

$$\mathbb{F}_{Z_A, Z_Y}(Z_A, Z_Y) = \mathbb{C}(\Phi(Z_A), \Phi(Z_Y)) . \quad (5c)$$

Eqs. (5b) and (5c) follow from the scale-invariance property of copula  $\mathbb{C}(U_A, U_Y)$  to strictly increasing transformations/CDFs  $\mathbb{F}_{\varepsilon_A}$ ,  $\mathbb{F}_{\varepsilon_Y}$ , and  $\Phi$ . The unknown copula  $\mathbb{C}(U_A, U_Y)$  in Figure 1 essentially models the non-causal back-door association, and the degree of the non-causal association between  $U_A$  and  $U_Y$  may be quantified using measures of association such as Spearman (1987, 2010)’s  $\rho_S$  or Kendall (1938)’s  $\tau_K$ , where  $\rho_S, \tau_K \in [-1, +1]$ . From the scale-invariance property of  $\rho_S$  and  $\tau_K$  to the strictly increasing transformations  $\mathbb{F}_{\varepsilon_A}$ ,  $\mathbb{F}_{\varepsilon_Y}$ , and  $\Phi$ , this measure of association between  $U_A$  and  $U_Y$  is the same between  $\varepsilon_A$  and  $\varepsilon_Y$ ,

and between  $Z_A$  and  $Z_Y$ , i.e.,  $\rho_S(\varepsilon_A, \varepsilon_Y) = \rho_S(U_A, U_Y) = \rho_S(Z_A, Z_Y) = \rho_{S_C}$ . In other words, the orange, blue and red back-door paths in Figure 1 induce the same measure of non-causal association  $\rho_{S_C}$  that can be represented using the copula  $\mathbb{C}(U_A, U_Y)$ . This result intuitively follows as the SCMs are observationally and interventionally equivalent, i.e., same measures of total observed and causal associations are expected, implying the non-causal associations also to be the same. As discussed above, for identifying the causal effect of interest, it is sufficient if we observe/known/hypothesize the noises or the copula to adjust for the non-causal back-door path. However, the copula  $\mathbb{C}(U_A, U_Y)$ , although uniquely exists because of the strictly increasing continuous transformations  $\mathbb{F}_{\varepsilon_A}$ ,  $\mathbb{F}_{\varepsilon_Y}$ , and  $\Phi$ , remains unknown and cannot be estimated from observational data as the noises are unobserved.

Since the copula  $\mathbb{C}(U_A, U_Y)$  is unknown and unlearnable, it is inevitable to make assumptions about it to achieve causal effect identification. Specifically,  $\mathbb{C}(U_A, U_Y)$  may be chosen from any of the vast families of copulas such as Archimedean copulas (Ling et al., 2020) (Clayton, Frank, Gumbel, etc.), elliptical, or empirical copulas (Nelsen, 2007; Salvadori et al., 2007; Durante and Sempi, 2016; Benali et al., 2021). Recent works such as Zheng et al. (2021, 2022) have proposed sensitivity analysis with one of the most well studied and used elliptical copula, namely the Gaussian copula, but without the use of normalizing flows. In our current work, we present our analysis by assuming and approximating the unknown copula  $\mathbb{C}(U_A, U_Y)$  with the Gaussian copula, while extending the normalizing flows with monotonic transformers that are recently shown to be universal non-linear SCM approximators (Huang et al., 2018; Wehenkel and Louppe, 2019, 2021; Balgi et al., 2022a).

The Gaussian copula has been widely used in the fields of quantitative finance (Cherubini et al., 2004; Salmon, 2009; MacKenzie and Spears, 2014), hydrology research (Renard and Lang, 2007; Zhang and Singh, 2019), logistics (Kumar, 2019), astronomy (Takeuchi, 2010), and similar fields (Nelsen, 2007; Salvadori et al., 2007; Durante and Sempi, 2016). This is one of the several motivations for the particular selection of the Gaussian copula. The assumption and approximation of the unknown copula  $\mathbb{C}(U_A, U_Y)$  in Eqs. (5a)-(5c) with a Gaussian copula  $\Phi_\rho(\Phi^{-1}(U_A), \Phi^{-1}(U_Y))$  achieves causal effect identification as the back-door non-causal association between  $Z_A$  and  $Z_Y$  may be identified as  $\mathbb{F}_{Z_A, Z_Y}(Z_A, Z_Y) = \mathbb{C}(\Phi(Z_A), \Phi(Z_Y)) \approx \Phi_\rho(Z_A, Z_Y)$ , where  $\rho \in [-1, +1]$  is the Pearson’s correlation between  $Z_A$  and  $Z_Y$ . As the copula and the noises  $\varepsilon_X$  and  $U_X$  in Figure 1 may exhibit non-linear dependence, it is more appropriate to equivalently represent the linear Pearson’s correlation parameter  $\rho$  in the Gaussian copula  $\Phi_\rho$  in terms of the non-linear measure of association  $\rho_{S_C}$  (or  $\tau_{K_C}$ ) using the following results for bivariate Gaussian copula (Kruskal, 1958; Meyer, 2013). This measure of back-door non-causal association  $\rho_{S_C}$  due to the Gaussian copula assumption equates to the Gaussian copula parameter  $\rho$  as  $\rho = 2 \sin(\pi \rho_{S_C} / 6)$  (i.e.,  $\rho_{S_C} \approx \rho$ ). With the Gaussian copula  $\Phi_\rho$  assumption in Eq. (5a)-(5c), the back-door non-causal association in Eqs. (1)-(4) is known, signifying that the causal effects are now identifiable under the assumed copula. Moreover, they can be estimated from a given observational dataset  $\{(A^\ell, Y^\ell)\}_{\ell=1}^{N_{train}}$  to train a parametric model as the proposed  $\rho$ -GNF in Figure 1 by rewriting Eq. (4) as

$$\rho\text{-GNF} :: A := \mathbb{T}_A^{-1}(Z_A; \theta_A) , Y := \mathbb{T}_{Y|A}^{-1}(Z_Y; \theta_Y) , (Z_A, Z_Y) \sim \Phi_\rho(Z_A, Z_Y) , \quad (6)$$

where  $\mathbb{T}_A(\bullet; \theta_A)$  and  $\mathbb{T}_{Y|A}(\bullet; \theta_Y)$  represent monotonic transformations parameterized by deep neural networks  $\theta = (\theta_A, \theta_Y)$  using the integration-based unconstrained monotonic neu-

ral network (UMNN) transformer (Wehenkel and Louppe, 2019). The conditioning of  $Y$  on its parent variable  $A$  in  $\mathbb{T}_{Y|A}(\bullet; \theta_Y)$  is done with the graphical conditioner from graphical normalizing flow (GNF) (Wehenkel and Louppe, 2021) using the directed acyclic graph (DAG)  $A \rightarrow Y$  assumed in Figure 1. We aptly refer to our model with the Gaussian copula assumption in Eq. (6) as  $\rho$ -GNF due to the similarities to GNF and c-GNF that propose normalizing flows for observational density estimation and causal inference, but not for sensitivity analysis (herein lies our novelty over GNF or c-GNF). As any normalizing flow, the UMNN transformers and graphical conditioners are trained by maximizing the log-likelihood of the observational training dataset  $\{(A^\ell, Y^\ell)\}_{\ell=1}^N$  (Wehenkel and Louppe, 2021; Balgi et al., 2022a), for a fixed  $\rho$ .

In principle, any copula maybe assumed in place of Gaussian copula in Eqs. (5a)-(5c). The Gaussian copula assumption, similar to normalizing flows (Tabak and Vanden-Eijnden, 2010; Tabak and Turner, 2013; Rezende and Mohamed, 2015; Papamakarios et al., 2017, 2021; Kobyzev et al., 2021), facilitates efficient computation of the log-likelihood of the observational training dataset. Thus, enabling **computationally efficient training** of  $\rho$ -GNF in Eq. (6). The Gaussian copula assumption further enables sampling  $(Z_A, Z_Y)$  from  $\mathbb{F}_{Z_A, Z_Y}(Z_A, Z_Y) \approx \Phi_\rho(Z_A, Z_Y)$  efficiently for the estimation of Monte-Carlo expectation in Eq. (7c), thus enabling **computationally efficient inference**. Most importantly, the Gaussian copula assumption provides a single bounded sensitivity parameter  $\rho \in [-1, +1]$  for sensitivity analysis that can be used to control/model/block/adjust the back-door non-causal association under which the ACE is identifiable. Thus, enabling **simple and efficient sensitivity analysis**. Our subsequent experiments and results show that the Gaussian copula assumption works well empirically, as Ilse et al. (2021) also observe for multiple unobserved confounders, the joint distribution of the observed variables becomes increasingly Gaussian due to the central limit theorem.

## 2.2. Sensitivity Analysis and Estimation of the Causal Estimand, i.e., ACE

The main objective is to estimate the average causal effect (ACE), which can be expressed as

$$ACE_\rho = \mathbf{E}[Y_1 - Y_0] = \mathbf{E}[Y_1] - \mathbf{E}[Y_0] \quad , \quad (7)$$

where  $Y_a$  denotes the potential outcome under the intervention  $A := a$ . In practise, the estimation of the ACE is done by Monte-Carlo expectation estimation by drawing the samples from the interventional distributions to approximate  $\mathbf{E}[Y_1]$  and  $\mathbf{E}[Y_0]$  as indicated below in Eqs. (7a)-(7c), after having trained the  $\rho$ -GNF for a specific measure of unobserved confounding  $\rho$  on the given observational dataset. *The First Law of Causal Inference* (Pearl, 1999, 2009a,b; Pearl and Mackenzie, 2018) provides three steps, i.e, abduction, action and prediction, to estimate  $\mathbf{E}[Y_a]$  and thus the  $ACE_\rho$  in Eq. (7).

$$Z_Y^\ell = \mathbb{T}_{Y|A^\ell}(Y^\ell; \theta_Y) \quad \forall \ell \in \{1, \dots, N\} \quad . \quad (7a)$$

$$Y_a^\ell = \mathbb{T}_{Y|a}^{-1}(Z_Y^\ell; \theta_Y) \quad \forall \ell \in \{1, \dots, N\} \quad . \quad (7b)$$

$$ACE_\rho = \mathbf{E}[Y_1] - \mathbf{E}[Y_0] \approx \frac{\sum_{\ell=1}^N Y_1^\ell}{N} - \frac{\sum_{\ell=1}^N Y_0^\ell}{N} \quad . \quad (7c)$$

### 2.3. Similarities to E-value

VanderWeele and Ding (2017) propose the E-value as the minimum strength of association, on the risk ratio scale, that an unmeasured confounder would need to have with both the treatment and the outcome to fully explain away a specific causal treatment–outcome association, conditional on the measured covariates. A large E-value implies that considerable unmeasured confounding would be needed to explain away an effect estimate. A small E-value implies that little unmeasured confounding would be needed to explain away an effect estimate. Similar to the E-value, we propose the  $\rho_{value}$  which represents the Gaussian copula parameter value that explains away the causal association between the observed treatment  $A$  and the observed outcome  $Y$ . In other words, setting the Gaussian copula parameter  $\rho = \rho_{value}$  results in  $ACE_\rho = 0$ , i.e.,  $\mathbf{E}[Y_1] = \mathbf{E}[Y_0]$  in Eqs. (7) and (7c), i.e., the potential outcomes are independent of the treatments/interventions. This implies that the strictly increasing transformation  $\mathbb{T}_{Y|A}^{-1}$  modeling  $Y$  in Eq. (6) is independent of  $A$ , i.e., we have  $Y = \mathbb{T}_Y^{-1}(Z_Y)$ .

From the scale-invariance property of Spearman’s correlation to strictly increasing transformations  $\mathbb{T}_A^{-1}$  and  $\mathbb{T}_Y^{-1}$ , we have  $\rho_S(Z_A, Z_Y) = \rho_{S_C} = \rho_S(A, Y) = \rho_{S_{Obs}}$ , where  $\rho_{S_{Obs}}$  represents the observed Spearman’s correlation between the observational data  $A$  and  $Y$ . For a given observational dataset  $A$  and  $Y$ , if the total observed association  $\rho_{S_{Obs}}$  between  $A$  and  $Y$  is completely due to the non-causal association  $\rho_{S_C}$ , there can be no contribution from the causal association, i.e.,  $ACE_\rho = 0 \Leftrightarrow \rho = \rho_{value}$  where we have  $\rho_{value} = 2 \sin(\pi \rho_{S_{Obs}}/6)$ . In Section 3.1, we also experimentally verify that  $\rho = \rho_{value} \Leftrightarrow ACE_\rho = 0$ . Hence, the E-value and  $\rho_{value}$  represent the same concept, rather interpreted in different scales. Similar to the E-value that can be computed directly from the observational dataset (Mathur et al., 2018), the  $\rho_{value}$  is also identified as a single, bounded intuitive parameter that can be computed directly from Spearman’s rho  $\rho_{S_{Obs}}$  between  $A$  and  $Y$ , i.e.,  $\rho_{value} = 2 \sin(\pi \rho_{S_{Obs}}/6)$ . This further shows the similarities of the E-value and  $\rho_{value}$ . Since the total observed association between  $A$  and  $Y$  remains constant for a given observational dataset, the  $ACE_\rho$  varies with  $\rho \in [-1, +1]$ . We aptly refer this sensitivity plot of  $ACE_\rho$  to  $\rho$  for a given observational dataset as the  $\rho_{curve}$ .

The  $\rho_{value}$  equips an analyst/domain-expert to determine the sign of the ACE (i.e., whether the treatment is harmful or beneficial), which is arguably the most important part of the ACE and one of the ultimate goals of causal inference. Specifically, suppose the domain expert hypothesizes a measure of confounding in the interval  $[\rho_{min}, \rho_{max}]$ . Thus, the  $\rho_{curve}$  enables us to bound the true ACE to the narrower interval  $[ACE_{\rho_{max}}, ACE_{\rho_{min}}]$ , which may in turn help us identify the most important insight of the causal inference, i.e., the sign of the true ACE. In particular, if  $\rho_{min} > \rho_{value}$  (or  $\rho_{max} < \rho_{value}$ ) then we may conclude that the true ACE is negative (or positive), as observed from the  $\rho_{curve}$  in Figures 2(a), 3, and 4. Furthermore, unlike E-value that only indicates the strength of the association due to unobserved confounding,  $\rho_{value}$  indicates both strength and the sign of the association due to unobserved confounding, i.e., *positive* or *negative* association between  $A$  and  $Y$ .



### 3. Experimental Setup, Results, Analysis and Discussion

Our  $\rho$ -GNF is implemented in PyTorch (Paszke et al., 2017)<sup>1</sup> by adapting the baseline code of UMNN (Wehenkel and Louppe, 2019) and GNF (Wehenkel and Louppe, 2021)<sup>2</sup>. As normalizing flows are developed for continuous variables, we use the Gaussian dequantization trick from c-GNF (Balgi et al., 2022a) to model discrete variables into  $\rho$ -GNF. For the experiments, we present two different settings: (i) simulated dataset with continuous outcomes in Section 3.1, and (ii) simulated dataset with binary outcomes in Section 3.2. We additionally present the experiments using a real-world dataset with categorical outcomes analysing the impact of the IMF (International Monetary Fund) program on the degree of child poverty in the Global-South region by denoting the outcome as the total degree of child poverty ranging in 0-7 classes.

In practise, the  $\rho_{curve}$  for a given observational dataset is obtained by varying the values of  $\rho \in [-1, +1]$ , i.e., we select  $\rho = \{-0.99, -0.8, -0.6, -0.4, -0.2, 0.0, +0.2, +0.4, +0.6, +0.8, +0.99\}$  to train  $\rho$ -GNFs. We estimate the respective  $ACE_\rho$  from Eqs. (7)-(7c) to plot the  $\rho_{curve}$  as interpolation of the points  $(\rho, ACE_\rho)$  shown in Figures 2, 3 and 4. Our empirical ACE bounds are obtained as the infimum and supremum of the  $\rho_{curve}$ , i.e.,  $\inf[ACE_\rho] \leq ACE_{true} \leq \sup[ACE_\rho]$ . We identify the  $\rho_{value}$  that explains away the causal association for the benefit of domain expert/analyst.

#### 3.1. Experiments with Continuous Outcomes

$SCM_{\alpha,\beta,\delta}$	$\alpha$	$\beta$	$\delta$	$\rho_{P_{Obs}}$	$\rho_{true}$	$ACE_{true}$	$\rho \Leftrightarrow ACE_\rho$
$SCM_1$	0.2	-0.6	0.72	-0.55	-0.71	0.2	-0.71 $\Leftrightarrow$ 0.2
$SCM_2$	0.0	-0.4	0.52	-0.55	-0.55	0.0	-0.55 $\Leftrightarrow$ 0.0
$SCM_3$	-0.2	-0.2	0.40	-0.55	-0.32	-0.2	-0.32 $\Leftrightarrow$ -0.2
$SCM_4$	0.2	0.2	0.40	0.55	0.32	0.2	0.32 $\Leftrightarrow$ 0.2
$SCM_5$	0.0	0.4	0.52	0.55	0.55	0.0	0.55 $\Leftrightarrow$ 0.0
$SCM_6$	-0.2	0.6	0.72	0.55	0.71	-0.2	0.71 $\Leftrightarrow$ -0.2

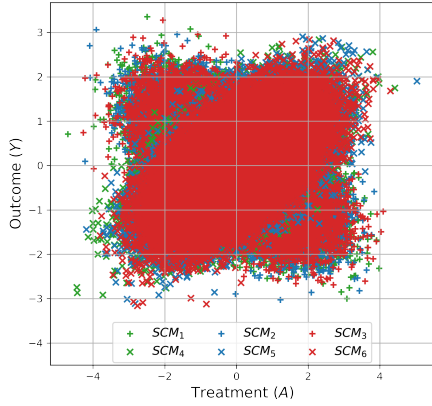
Table 1: Six observationally and/or interventionally dissimilar SCMs with different mixtures of total observed association ( $\rho_{P_{Obs}}$ ), non-causal association ( $\rho_{true}$ ), and causal association ( $ACE_{true}$ ).

In our first set of simulated experiments, we consider the SCM with continuous treatment  $A$  and outcome  $Y$  variables by Hoover (2006). This is a well-studied SCM from economics and econometrics, and is defined as below. The true causal, non-causal, and total observed (causal+non-causal) associations in the SCM are known and are used for verification purposes.

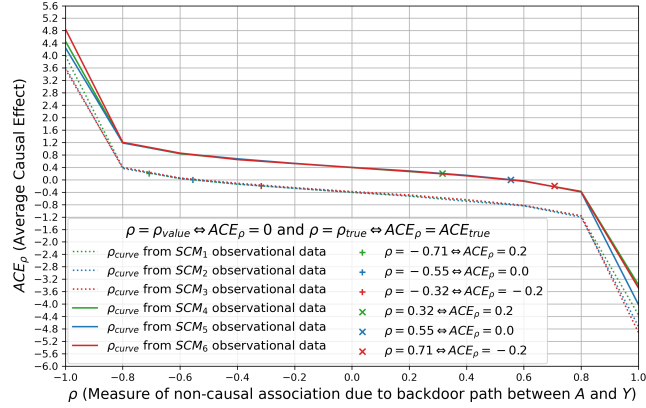
$$SCM_{\alpha,\beta,\delta} :: A := \varepsilon_A, \quad Y := \alpha A + \varepsilon_Y, \quad \begin{bmatrix} \varepsilon_A \\ \varepsilon_Y \end{bmatrix} \sim \mathcal{N} \left( \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 & \beta \\ \beta & \delta \end{bmatrix} \right). \quad (8)$$

1. The  $\rho$ -GNF code is available at <https://github.com/sobalgi/rhoGNF>.

2. GNF: <https://github.com/AWehenkel/Graphical-Normalizing-Flows>, UMNN: <https://github.com/AWehenkel/UMNN>.



(a) Scatter plots of the six observational datasets from the respective SCMs.



(b) Six  $\rho_{curve}$  obtained using observational dataset from six SCMs in Table 1.

Figure 2: Observationally equivalent SCMs, e.g.,  $\{SCM_1, SCM_2, SCM_3\}$  and  $\{SCM_4, SCM_5, SCM_6\}$ , show similar scatter plots, and  $\rho_{P_{Obs}}$ , as well as similar  $\rho_{curve}$ , confirming respective observational equivalences.

For a given set of SCM parameters  $\alpha, \beta, \delta$  in Eq. (8), we have the corresponding Pearson’s correlation  $\rho_{P_{Obs}} = \frac{\sigma_{A,Y}}{\sigma_A \sigma_Y}$ ,  $\sigma_A^2 = 1$ ,  $\sigma_Y^2 = \alpha^2 + \delta + 2\alpha\beta$ ,  $\sigma_{A,Y} = \alpha + \beta$ ,  $\rho_{true} = \beta/\delta$  and  $ACE_{true} = \alpha$  as tabulated in Table 1. Figure 2(a) indicates the scatter plot of each of the six observational datasets from the SCM in Eq. (8) as parameterized in Table 1. The SCMs that are observationally equivalent with the same  $\rho_{P_{Obs}}$  present similar observational data distributions even though these data distributions have been generated from interventionally dissimilar SCMs. Figure 2(b) shows observationally equivalent datasets result in equivalent  $\rho_{curve}$ , e.g.,  $\rho_{curve}$  of  $\{SCM_1, SCM_2, SCM_3\}$  are similar since they are observationally equivalent as verifiable in the scatter plots of their respective observational dataset from Figure 2(a). From Figure 2(b) and Table 1, we empirically observe the results  $\rho = \rho_{value}$  corresponds to  $ACE_\rho = 0$  and  $\rho = \rho_{true}$  corresponds to  $ACE_\rho = ACE_{true}$  as presented in Section 2.3. In summary, the main contribution of our method is the estimation of ACE as a function of the sensitivity parameter  $\rho$  that indicates the degree of the confounding, thus distinguishing observationally equivalent SCMs based on the assumed degree of unconfoundedness.

The SCM in Eq. (8) used to illustrate  $\rho$ -GNF with continuous outcomes may be considered a simple case, i.e. a linear SCM similar to the ones studied in Cinelli et al. (2019); Cinelli and Hazlett (2020). However, as the  $\rho$ -GNF is parameterized to learn arbitrary non-linear monotonic functions using UMNN transformers (Wehenkel and Louppe, 2019), our work is a generalization of the sensitivity analysis for simple linear SCMs to complex non-linear SCMs, as we demonstrate in the subsequent experiments with binary and categorical outcomes in Section 3.2.

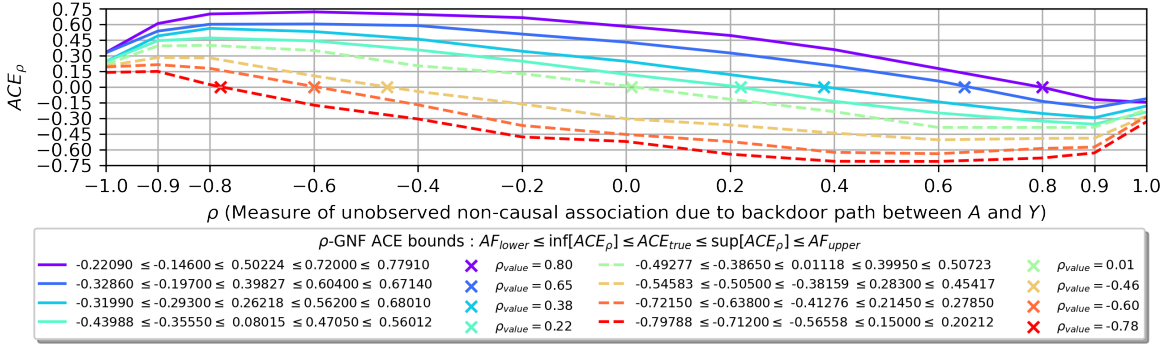


Figure 3: Eight  $\rho_{curve}$  each corresponding to a randomly generated dataset from a randomly parameterized DGP, and their respective  $\rho_{value}$ . The true ACE identified using *do*-calculus is indeed included for all the  $\rho_{curve}$  with our empirical bounds narrower than the AF bounds.

### 3.2. Experiments with Discrete (Binary/Categorical) Outcomes

In the second set of experiments, we consider multiple (twenty) randomly generated observational datasets from randomly sampled data generating processes as proposed in [Sjölander \(2020\)](#); [Sjölander and Hössjer \(2021\)](#); [Peña \(2022\)](#), where all variables are binary. As proposed, we randomly sample the parameters  $\{\mathbb{P}(U), \mathbb{P}(A|U), \mathbb{P}(Y|A, U)\}$  of the binary data generating process (DGP) from the uniform distribution in  $[0,1]$ , where  $U$  denotes the confounder of  $A$  and  $Y$  which is intentionally hidden during training to simulate unobserved confounding. For this simple binary outcome/treatment/confounder case, the gold-standard AF bounds of a constant width of 1 ( $= p_1 + p_0$ ) with 100% certainty of including the true ACE are presented in [Robins \(1989\)](#); [Manski \(1990\)](#) as

$$AF \text{ ACE bounds : } AF_{lower} = q_1 p_1 - q_0 p_0 - p_1 \leq ACE_{true} \leq AF_{upper} = q_1 p_1 - q_0 p_0 + p_0 \quad , \quad (9)$$

where  $p_a = \mathbb{P}(A=a)$  and  $q_a = \mathbb{P}(Y=1|A=a)$  are estimated using observational data. The  $\rho_{curve}$  in Figure 3 does include the true ACE in all of the randomly generated DGPs. Further, we obtain a narrower bound of  $0.83 \pm 0.03$  (i.e.,  $\approx 14\%$  narrower than the AF bounds) such that our empirical bounds lie within the AF bounds, unlike the bounds by [VanderWeele and Ding \(2017\)](#) which may be wider than the AF bounds, as shown in [Sjölander \(2020\)](#). The former comes as no surprise: Since our bounds are empirically obtained particularly assuming the Gaussian copula, they must lie within the AF bounds. Figure 3 verifies it, i.e.,

$$\rho\text{-GNF ACE bounds : } AF_{lower} \leq \inf[ACE_\rho] \leq ACE_{true} \leq \sup[ACE_\rho] \leq AF_{upper} \quad .(10)$$

Subsequently, we also extend the experiment from binary outcome to categorical outcome using a real-world dataset with categorical outcomes analysing the IMF (International Monetary Fund) program impact on the degree of child poverty in the Global-South region with the outcome as the total degree of child poverty ranging in 0-7 classes ([Baldi et al., 2022b](#)). Unlike the binary outcomes, the AF bounds are not available for the non-binary outcomes such as the categorical total degree of child poverty (degree 0: no poverty to degree 7: severe poverty). Since the degree of child poverty is formulated as the sum of

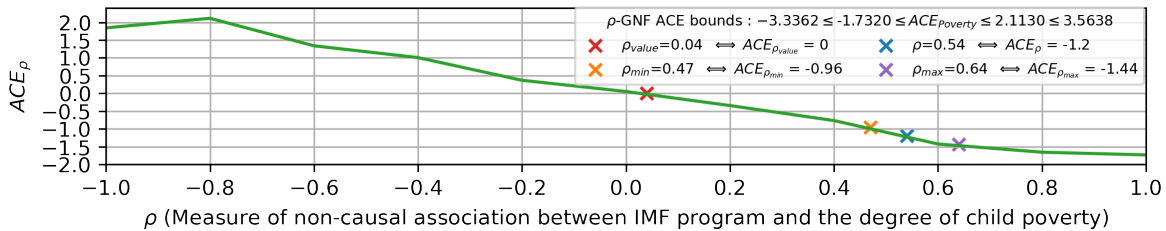


Figure 4:  $\rho_{curve}$  for the total degree of the child poverty along with the bounds from Eq. (10). The point  $(\rho, ACE_{\rho})=(0.54, -1.2)$  corresponds to the ACE estimate from [Balgı et al. \(2022b\)](#) obtained with observed confounders and unconfoundedness as well as the ACE interval  $-1.2 \pm 0.24$  respectively corresponds to  $[+0.47, +0.64]$  interval of  $\rho$ .

seven binary individual dimension of child poverty, we may identify the AF bounds for each binary individual dimension of child poverty and extend the AF bounds to the total degree of child poverty. From Eq. (9) and the observational dataset of the seven binary individual dimensions of child poverty, the assumption-free lower and upper bounds are identified as, (i) education:  $[-0.4843, 0.5157]$ , (ii) health:  $[-0.5009, 0.4991]$ , (iii) information:  $[-0.4893, 0.5107]$ , (iv) malnutrition:  $[-0.5111, 0.4889]$ , (v) sanitization:  $[-0.5012, 0.4988]$ , (vi) shelter:  $[-0.4756, 0.5244]$ , (vii) water:  $[-0.4738, 0.5262]$ .

Since, the total degree of child poverty is the sum of these seven binary individual dimensions of child poverty, the lower and upper assumption-free bounds total degree of child poverty can be identified as sum of the respective lower and upper bounds. Thus, obtain the AF bounds for the total degree of child poverty as  $[-3.3362, +3.5638]$  as indicated in the  $\rho_{curve}$  in Figure 4. Note that since the binary AF bounds are of width 1, and the total degree of child poverty is a sum of seven binary variables, the AF bounds for the total degree of the child poverty is of width 7. From Figure 4, our ACE bounds  $[-1.73, +2.11]$  are verified empirically with a narrower width of 3.84, i.e., a reduction of the bounds by 45.1% under the Gaussian copula assumption.

#### 4. Conclusion

We proposed a novel copula-based approach for sensitivity analysis termed  $\rho$ -GNF to bound the causal effect in the risk difference scale where  $\rho \in [-1, +1]$  is a bounded sensitivity parameter representing the unobserved back-door non-causal association between the observed treatment and outcome. Under the Gaussian copula assumption, we showed that  $\rho$ -GNF enabled us to estimate the causal effect as a function of  $\rho$  in the form of  $\rho_{curve}$ . The  $\rho_{curve}$  enabled us to identify narrower empirical bounds in contrast to the wider AF bounds, irrespective of discrete or continuous outcome variables. We identified  $\rho_{value}$  as the measure of the unobserved confounding that explains away the causal effect and presented the similarities to the E-value with experimental validation. Further, the  $\rho_{curve}$  enabled us to provide finer bounds for the causal effect given an interval of  $\rho$  values that the domain expert considers appropriate to identify the sign of the ACE, thus deducing if the treatment is beneficial or harmful. The adaptability of  $\rho$ -GNF for both discrete and continuous outcomes should encourage the use of sensitivity analysis when working with non-randomized observational data to draw causal conclusions.

## References

- J. E. Angus. [The Probability Integral Transform and Related Results](#). *SIAM review*, 36(4): 652–654, 1994.
- S. Balgi, J. M. Peña, and A. Daoud. [Personalized Public Policy Analysis in Social Sciences Using Causal-Graphical Normalizing Flows](#). In *AAAI Conference on Artificial Intelligence (AAAI)*, pages 11810–11818, 2022a.
- S. Balgi, J. M. Peña, and A. Daoud. [Counterfactual Analysis of the Impact of the IMF Program on Child Poverty in the Global-South Region using Causal-Graphical Normalizing Flows](#). *arXiv preprint arXiv:2202.09391*, 2022b.
- F. Benali, D. Bodénès, N. Labroche, and C. de Runz. [MTCopula: Synthetic Complex Data Generation Using Copula](#). In *International Workshop on Design, Optimization, Languages and Analytical Processing of Big Data (DOLAP)*, pages 51–60, 2021.
- B. A. Brumback, M. A. Hernán, S. J. Haneuse, and J. M. Robins. [Sensitivity Analyses for Unmeasured Confounding Assuming a Marginal Structural Model for Repeated Measures](#). *Statistics in Medicine*, 23(5):749–767, 2004.
- U. Cherubini, E. Luciano, and W. Vecchiato. *Copula Methods in Finance*. John Wiley & Sons, 2004.
- C. Cinelli and C. Hazlett. [Making Sense of Sensitivity: Extending Omitted Variable Bias](#). *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 82(1):39–67, 2020.
- C. Cinelli, D. Kumor, B. Chen, J. Pearl, and E. Bareinboim. [Sensitivity Analysis of Linear Structural Causal Models](#). In *International Conference on Machine Learning (ICML)*, pages 1252–1261, 2019.
- W. G. Cochran and D. B. Rubin. [Controlling Bias in Observational Studies: A Review](#). *Sankhyā: The Indian Journal of Statistics, Series A*, pages 417–446, 1973.
- J. Cornfield, W. Haenszel, E. C. Hammond, A. M. Lilienfeld, M. B. Shimkin, and E. L. Wynder. [Smoking and Lung Cancer: Recent Evidence and a Discussion of Some Questions](#). *Journal of the National Cancer Institute (JNCI)*, 22(1):173–203, 1959.
- D. R. Cox. *Planning of Experiments*. New York: Wiley, 1958.
- P. Ding and T. J. VanderWeele. [Sensitivity Analysis Without Assumptions](#). *Epidemiology (Cambridge, Mass.)*, 27(3):368, 2016.
- B. M. D’Onofrio, A. Sjölander, B. B. Lahey, P. Lichtenstein, and A. S. Öberg. [Accounting for confounding in observational studies](#). *Annual Review of Clinical Psychology*, 16:25–48, 2020.
- F. Durante and C. Sempi. *Principles of Copula Theory*, volume 474. CRC press Boca Raton, 2016.

- S. E. Fienberg and O. D. Duncan. **Introduction to Structural Equation Models**. *Journal of the American Statistical Association (JASA)*, 72:485, 1975.
- R. A. Fisher. **Design of Experiments**. *British Medical Journal (BMJ)*, 1:554–554, 1936.
- A. S. Goldberger. **Structural Equation Methods in the Social Sciences**. *Econometrica*, 40: 979–1001, 1972.
- T. Haavelmo. **The Statistical Implications of a System of Simultaneous Equations**. *Econometrica*, 11:1–12, 1943.
- M. A. Hernán and J. M. Robins. *Causal Inference: What If*. Boca Raton: Chapman & Hall/CRC, 2009.
- K. D. Hoover. **Causality in Economics and Econometrics**. *SSRN eLibrary*, 2006.
- C. Huang, D. Krueger, A. Lacoste, and A. C. Courville. **Neural Autoregressive Flows**. In *International Conference on Machine Learning (ICML)*, pages 2083–2092, 2018.
- Y. Huang and M. Valtorta. **Pearl’s Calculus of Intervention Is Complete**. In *Uncertainty in Artificial Intelligence (UAI)*, 2006.
- M. Ilse, P. Forré, M. Welling, and J. M. Mooij. **Combining Interventional and Observational Data Using Causal Reductions**. *arXiv preprint arXiv:2103.04786*, 2021.
- G. W. Imbens. **Sensitivity to Exogeneity Assumptions in Program Evaluation**. *American Economic Review*, 93(2):126–132, 2003.
- G. W. Imbens and D. B. Rubin. *Causal Inference for Statistics, Social, and Biomedical Sciences: An Introduction*. Cambridge University Press, USA, 2015.
- J. Ioannidis, Y. Tan, and M. Blum. **Limitations and Misinterpretations of E-values for Sensitivity Analyses of Observational Studies**. *Internal Medicine*, 170(2):108–111, 2019.
- A. Javaloy, P. Sánchez-Martín, and I. Valera. **Causal Normalizing Flows: From Theory to Practice**. In *Neural Information Processing Systems (NeurIPS)*, 2023.
- M. G. Kendall. **A New Measure of Rank Correlation**. *Biometrika*, 30(1/2):81–93, 1938.
- I. Kobyzev, S. Prince, and M. Brubaker. **Normalizing Flows: An Introduction and Review of Current Methods**. *IEEE Transactions on Pattern Analysis and Machine Intelligence (TPAMI)*, 43(11):3964–3979, 2021.
- W. H. Kruskal. **Ordinal Measures of Association**. *Journal of the American Statistical Association (JASA)*, 53(284):814–861, 1958.
- P. Kumar. **Copula Functions and Applications in Engineering**. In *Logistics, Supply Chain and Financial Predictive Analytics*, pages 195–209. Springer, 2019.
- T. L. Lash, M. P. Fox, A. K. Fink, et al. *Applying Quantitative Bias Analysis to Epidemiologic Data*. Springer, 2009.

- A. Lindmark, X. de Luna, and M. Eriksson. **Sensitivity Analysis for Unobserved Confounding of Direct and Indirect Effects Using Uncertainty Intervals**. *Statistics in Medicine*, 37(10):1744–1762, 2018.
- C. K. Ling, F. Fang, and J. Z. Kolter. **Deep Archimedean Copulas**. *Neural Information Processing Systems (NeurIPS)*, 33:1535–1545, 2020.
- D. MacKenzie and T. Spears. **The Formula That Killed Wall Street: The Gaussian Copula and Modelling Practices in Investment Banking**. *Social Studies of Science*, 44(3):393–417, 2014.
- C. F. Manski. **Nonparametric Bounds on Treatment Effects**. *American Economic Review*, 80(2):319–323, 1990.
- M. B. Mathur, P. Ding, C. A. Riddell, and T. J. VanderWeele. **Website and R Package for Computing E-values**. *Epidemiology*, 29(5):e45, 2018.
- C. Meyer. **The Bivariate Normal Copula**. *Communications in Statistics-Theory and Methods*, 42(13):2402–2422, 2013.
- J. M. Mooij, J. Peters, D. Janzing, J. Zscheischler, and B. Schölkopf. **Distinguishing Cause From Effect Using Observational Data: Methods and Benchmarks**. *Journal of Machine Learning Research (JMLR)*, 17(1):1103–1204, 2016.
- R. B. Nelsen. *An Introduction to Copulas*. Springer Science & Business Media, 2007.
- G. Papamakarios, I. Murray, and T. Pavlakou. **Masked Autoregressive Flow for Density Estimation**. In *Neural Information Processing Systems (NeurIPS)*, pages 2338–2347, 2017.
- G. Papamakarios, E. Nalisnick, D. J. Rezende, S. Mohamed, and B. Lakshminarayanan. **Normalizing Flows for Probabilistic Modeling and Inference**. *Journal of Machine Learning Research (JMLR)*, 22(57):1–64, 2021.
- A. Paszke, S. Gross, S. Chintala, G. Chanan, E. Yang, Z. DeVito, Z. Lin, A. Desmaison, L. Antiga, and A. Lerer. **Automatic Differentiation in PyTorch**. *NeurIPS Workshops*, 2017.
- J. Pearl. **Probabilities of Causation: Three Counterfactual Interpretations and Their Identification**. *Synthese*, 121(1):93–149, 1999.
- J. Pearl. **Causal Inference in Statistics: An Overview**. *Statistics Surveys*, 3:96–146, 2009a.
- J. Pearl. *Causality: Models, Reasoning and Inference*. Cambridge University Press, USA, 2009b.
- J. Pearl. **The Do-Calculus Revisited**. In *Uncertainty in Artificial Intelligence (UAI)*, page 3–11, 2012.
- J. Pearl and D. Mackenzie. *The Book of Why: The New Science of Cause and Effect*. Basic Books, Inc., 2018.

- J. M. Peña. **Simple Yet Sharp Sensitivity Analysis for Unmeasured Confounding.** *Journal of Causal Inference (JCI)*, 10(1):1–17, 2022.
- J. Peters, D. Janzing, and B. Schölkopf. *Elements of Causal Inference: Foundations and Learning Algorithms.* The MIT Press, 2017.
- B. Renard and M. Lang. **Use of a Gaussian Copula for Multivariate Extreme Value Analysis: Some Case Studies in Hydrology.** *Water Resources*, 30(4):897–912, 2007.
- D. Rezende and S. Mohamed. **Variational Inference with Normalizing Flows.** In *International Conference on Machine Learning (ICML)*, pages 1530–1538, 2015.
- J. M. Robins. **A New Approach to Causal Inference in Mortality Studies with a Sustained Exposure Period—Application to Control of the Healthy Worker Survivor Effect.** *Mathematical Modelling*, 7(9):1393–1512, 1986.
- J. M. Robins. **The Analysis of Randomized and Non-randomized AIDS Treatment Trials Using a New Approach to Causal Inference in Longitudinal Studies.** *Health Service Research Methodology: A Focus on AIDS*, pages 113–159, 1989.
- J. M. Robins and M. A. Hernán. **Estimation of the Causal Effects of Time-varying Exposure.** *Longitudinal Data Analysis (LDA)*, pages 553–599, 2008.
- K. J. Rothman, S. Greenland, T. L. Lash, et al. *Modern Epidemiology*, volume 3. Wolters Kluwer Health/Lippincott Williams & Wilkins Philadelphia, 2008.
- D. B. Rubin. **Formal Mode of Statistical Inference for Causal Effects.** *Journal of Statistical Planning and Inference (JSPI)*, 25(3):279–292, 1990.
- F. Salmon. **Recipe for Disaster: The Formula That Killed Wall Street.** *Wired Magazine*, 17(3):17–03, 2009.
- G. Salvadori, C. De Michele, N. T. Kottegoda, and R. Rosso. *Extremes in Nature: An Approach Using Copulas*, volume 56. Springer Science & Business Media, 2007.
- J. J. Schlesselman. **Assessing Effects of Confounding Variables.** *American Journal of Epidemiology*, 108(1):3–8, 1978.
- A. Sjölander. **A Note on a Sensitivity Analysis for Unmeasured Confounding, and the Related E-value.** *Journal of Causal Inference (JCI)*, 8(1):229–248, 2020.
- A. Sjölander and S. Greenland. **Are E-values Too Optimistic or Too Pessimistic? Both and Neither!** *International Journal of Epidemiology*, 2022.
- A. Sjölander and O. Hössjer. **Novel Bounds for Causal Effects Based on Sensitivity Parameters on the Risk Difference Scale.** *Journal of Causal Inference (JCI)*, 9(1):190–210, 2021.
- A. Sklar. **Fonctions De Répartition à N Dimensions Et Leurs Marges.** *Publ. inst. statist. univ. Paris*, 8:229–231, 1959.



- A. Sklar. **Random Variables, Joint Distribution Functions, and Copulas.** *Kybernetika*, 9(6): 449–460, 1973.
- C. Spearman. **The Proof and Measurement of Association Between Two Things.** *The American Journal of Psychology*, 100(3/4):441–471, 1987.
- C. Spearman. **The Proof and Measurement of Association Between Two Things.** *International Journal of Epidemiology*, 39(5):1137–1150, 2010.
- E. Tabak and E. Vanden-Eijnden. **Density Estimation by Dual Ascent of the Log-Likelihood.** *Communications in Mathematical Sciences*, 8:217–233, 2010.
- E. G. Tabak and C. V. Turner. **A Family of Nonparametric Density Estimation Algorithms.** *Communications on Pure and Applied Mathematics*, 66:145–164, 2013.
- M. Tabar, W. Jung, A. Yadav, O. W. Chavez, A. Flores, and D. Lee. **Forecasting the Number of Tenants At-Risk of Formal Eviction: A Machine Learning Approach to Inform Public Policy.** In *International Joint Conference on Artificial Intelligence (IJCAI)*, pages 5178–5184, 2022.
- T. T. Takeuchi. **Constructing a Bivariate Distribution Function With Given Marginals and Correlation: Application to the Galaxy Luminosity Function.** *Monthly Notices of the Royal Astronomical Society*, 406(3):1830–1840, 2010.
- E. J. T. Tchetgen and I. Shpitser. **Semiparametric Theory for Causal Mediation Analysis: Efficiency Bounds, Multiple Robustness, and Sensitivity Analysis.** *Annals of Statistics*, 40(3):1816, 2012.
- J. Tian and J. Pearl. **A General Identification Condition for Causal Effects.** In *AAAI Conference on Artificial Intelligence (AAAI)*, page 567–573, 2002.
- T. J. VanderWeele and O. A. Arah. **Bias Formulas for Sensitivity Analysis of Unmeasured Confounding for General Outcomes, Treatments, and Confounders.** *Epidemiology*, pages 42–52, 2011.
- T. J. VanderWeele and P. Ding. **Sensitivity Analysis in Observational Research: Introducing the E-value.** *Internal Medicine*, 167(4):268–274, 2017.
- V. Veitch and A. Zaveri. **Sense and Sensitivity Analysis: Simple Post-hoc Analysis of Bias Due to Unobserved Confounding.** *Neural Information Processing Systems (NeurIPS)*, 33: 10999–11009, 2020.
- A. Wehenkel and G. Louppe. **Unconstrained Monotonic Neural Networks.** In *Neural Information Processing Systems (NeurIPS)*, pages 1545–1555, 2019.
- A. Wehenkel and G. Louppe. **Graphical Normalizing Flows.** In *International Conference on Artificial Intelligence and Statistics (AISTATS)*, pages 37–45, 2021.
- S. Wright. **Correlation and Causation.** *Journal of Agricultural Research (JAR)*, 20:557–585, 1921.

- L. Zhang and V. P. Singh. *Copulas and Their Applications in Water Resources Engineering*. Cambridge University Press, 2019.
- J. Zheng, A. D'Amour, and A. Franks. *Copula-based Sensitivity Analysis for Multi-Treatment Causal Inference with Unobserved Confounding*. *arXiv preprint arXiv:2102.09412*, 2021.
- J. Zheng, J. Wu, A. D'Amour, and A. Franks. *Sensitivity to Unobserved Confounding in Studies with Factor-structured Outcomes*. *arXiv preprint arXiv:2208.06552*, 2022.