Statistische Inferenz und Kausalität

Interventionen

-

14.07.2023

Inhaltsverzeichnis

Einführung

Ziel der kausalen Inferenz: Wirkungen mögliche Ursachen zuordnen.

Ziel der kausalen Inferenz: Wirkungen mögliche Ursachen zuordnen.

Die *kausale Hierarchie* (*causal hierarchy*) ist gegeben durch.

(i) Assoziation

(ii) Intervention

(iii) Kontrafaktisches

Inhaltsverzeichnis

Definition 4.1 (Bedingte Wahrscheinlichkeit und bedingte Erwartung)

Seien X und Y messbare Zufallsvariablen mit Werten x und y. Sei weiter $\mathbb{P}(Y = y) > 0$. Dann ist die *bedingte Wahrscheinlichkeit* von $X = x$ unter der Bedingung $Y = y$ definiert durch

$$\mathbb{P}(X = x \mid Y = y) := \frac{\mathbb{P}(X = x, Y = y)}{\mathbb{P}(Y = y)} \; .$$

Bedingte Wahrscheinlichkeit und bedingte Erwartung

Definition 4.1 (Bedingte Wahrscheinlichkeit und bedingte Erwartung)

Seien X und Y messbare Zufallsvariablen mit Werten x und y. Sei weiter $\mathbb{P}(Y = y) > 0$. Dann ist die *bedingte Wahrscheinlichkeit* von $X = x$ unter der Bedingung $Y = y$ definiert durch

$$\mathbb{P}(X = x \mid Y = y) := \frac{\mathbb{P}(X = x, Y = y)}{\mathbb{P}(Y = y)} \ .$$

Seien $x_1, \ldots, x_n$ alle Werte von X. Dann ist die *bedingte Erwartung* von X unter der Bedingung $Y = y$ definiert durch

$$\mathbb{E}(X \mid Y = y) := \sum_{i=1}^{n} x_i \mathbb{P}(X = x_i \mid Y = y) \ .$$

Definition 4.2 (Stochastische Unabhängigkeit)

Seien X und Y messbare Zufallsvariablen. X und Y heißen *stochastische unabhängig*, wenn für alle Kombinationen von Werten x von X und y von Y gilt

$$\mathbb{P}(X = x, Y = y) = \mathbb{P}(X = x) \cdot \mathbb{P}(Y = y) \ .$$

Stochastische Unabhängigkeit

Definition 4.2 (Stochastische Unabhängigkeit)

Seien X und Y messbare Zufallsvariablen. X und Y heißen *stochastische unabhängig*, wenn für alle Kombinationen von Werten x von X und y von Y gilt

$$\mathbb{P}(X = x, Y = y) = \mathbb{P}(X = x) \cdot \mathbb{P}(Y = y) \ .$$

Sind X und Y unabhängig, so gilt

$$\mathbb{P}(X = x \mid Y = y) = \mathbb{P}(X = x) \ .$$

Satz von der totalen Wahrscheinlichkeit

Satz 4.5 (Satz von der totalen Wahrscheinlichkeit)

Seien X und Y messbare Zufallsvariablen. Sei x ein Wert von X und $y_1, \ldots, y_n$ alle Werte von Y. Dann gilt folgende Zerlegung

$$\mathbb{P}(X = x) = \sum_{i=1}^{n} \mathbb{P}(X = x \mid Y = y_i) \cdot \mathbb{P}(Y = y_i) \ .$$

Definition 4.6 (Kausal- und Korrelationsgraph)

Ein *Kausalgraph* (*SCM*) ist ein gerichteter, azyklischer Graph. Die Knoten des Graphen beschreiben die betrachteten Variablen und die Pfeile dazwischen die Richtung des kausalen Zusammenhangs.

Kausal- und Korrelationsgraph

Definition 4.6 (Kausal- und Korrelationsgraph)

Ein *Kausalgraph* (SCM) ist ein gerichteter, azyklischer Graph. Die Knoten des Graphen beschreiben die betrachteten Variablen und die Pfeile dazwischen die Richtung des kausalen Zusammenhangs.
Ein *Korrelationsgraph* ist ein ungerichteter Graph. Die Knoten des Graphen beschreiben wieder die betrachteten Variablen und die Linien dazwischen die Assoziationen der Variablen.

Exogene und endogene Variablen

Definition 4.9 (Exogene und endogene Variablen)

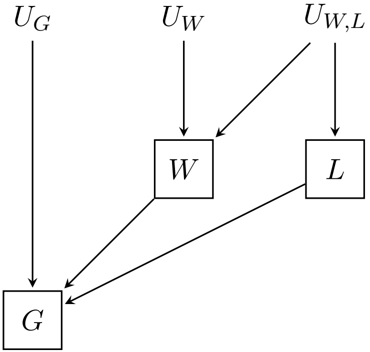
Exogene Variablen (*exogenous variables*) sind Variablen, deren Werte außerhalb eines gegebenen Modells bestimmt werden, womit darauf kein Einfluss genommen werden kann. Sie werden als gegeben und unabhängig von allen anderen Variablen angenommen. Bezeichnet werden die Variablen mit $U_X, U_Y, U_Z, \ldots$ und ihre Werte mit $u_X, u_Y, u_Z, \ldots$.

Definition 4.9 (Exogene und endogene Variablen)

Exogene Variablen (*exogenous variables*) sind Variablen, deren Werte außerhalb eines gegebenen Modells bestimmt werden, womit darauf kein Einfluss genommen werden kann. Sie werden als gegeben und unabhängig von allen anderen Variablen angenommen. Bezeichnet werden die Variablen mit $U_X, U_Y, U_Z, \ldots$ und ihre Werte mit $u_X, u_Y, u_Z, \ldots$.
Endogene Variablen (*endogenous variables*) sind Variablen, die durch Zustände, die im Modell stattfindet und durch die exogenen Variablen bestimmt werden. Bezeichnet werden die Variablen mit $X, Y, Z, \ldots$ und ihre Werte mit $x, y, z, \ldots$.
Wir setzen voraus, dass wir endogene Variablen also Funktionen von anderen Variablen.

$$\begin{aligned} W &= \text{Wasser} \\ L &= \text{Licht} \\ G &= \text{Größer} \end{aligned}$$

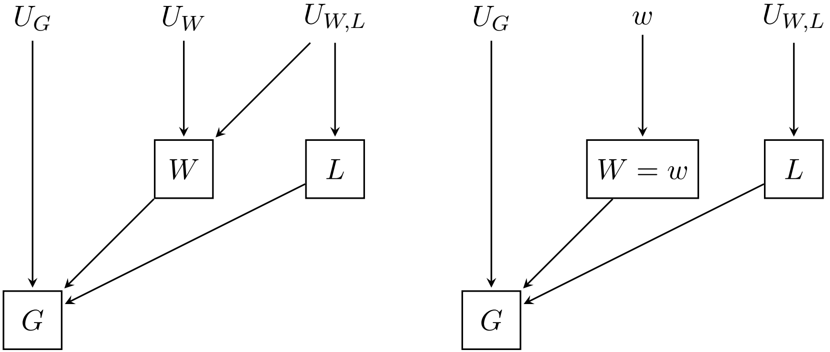
$$G = f_G(W, L, U_G)$$



Pflanzenwachstum

$$
\begin{aligned}
W =&\ \text{Wasser}\\
L =&\ \text{Licht}\\
G =&\ \text{Größer}
\end{aligned}
$$

$$G = f_G(W, L, U_G)$$



Inhaltsverzeichnis

Definition 4.12 Kausalität

Seien X und Y Zufallsvariablen. Wir sagen X *verursacht* Y (*cause*), wenn die Werte von Y in irgendeiner Weise von den Werten von X abhängen.

Definition 4.12 Kausalität

Seien X und Y Zufallsvariablen. Wir sagen X *verursacht* Y (*cause*), wenn die Werte von Y in irgendeiner Weise von den Werten von X abhängen.

$$X \longrightarrow Y$$

Definition 4.16 Korrelation

Seien X und Y Zufallsvariablen. Wir sagen X und Y sind *korreliert* (*correlated*), wenn ihre *Kovarianz* nicht verschwindet. Es soll also gelten

$$\mathrm{Cov}(X,Y) = \mathbb{E}(XY) - \mathbb{E}(X)\mathbb{E}(Y) \neq 0 \ .$$

Man sagt auch X und Y sind *assoziiert*.

Korrelation

Definition 4.16 Korrelation

Seien X und Y Zufallsvariablen. Wir sagen X und Y sind *korreliert* (*correlated*), wenn ihre *Kovarianz* nicht verschwindet. Es soll also gelten

$$\mathrm{Cov}(X,Y) = \mathbb{E}(XY) - \mathbb{E}(X)\mathbb{E}(Y) \neq 0 \ .$$

Man sagt auch X und Y sind *assoziiert.*

$$X \quad\text{———}\quad Y$$

Kausalprinzip

Seien X und Y assoziierte Variablen. Dann gilt das *Kausalprinzip* und somit eine der folgenden Aussagen.

(i) X folgt aus Y

(ii) Y folgt aus X

(iii) Es gibt ein Z mit X folgt aus Z und Y folgt aus Z

Kausalprinzip

Seien X und Y assoziierte Variablen. Dann gilt das *Kausalprinzip* und somit eine der folgenden Aussagen.

(i) X folgt aus Y

(ii) Y folgt aus X

(iii) Es gibt ein Z mit X folgt aus Z und Y folgt aus Z

$$X \longleftarrow Y \qquad Y \longleftarrow X$$

$$\begin{tikzcd} & Z \arrow[dl] \arrow[dr] & \\ X & & Y \end{tikzcd}$$

Eisverkauf

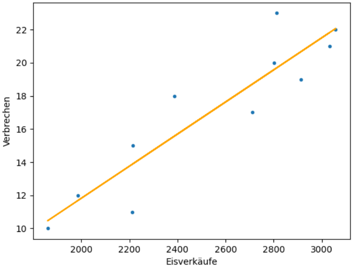
Sei E die Anzahl der Eisverkäufe und D die Anzahl der Taschendiebstähle.

E	1860	1985	2211	2215	2712	2387	2801	2912	3058	3032	2812
D	10	12	11	15	17	18	20	19	22	21	23

Eisverkauf

Sei E die Anzahl der Eisverkäufe und D die Anzahl der Taschendiebstähle.

E	1860	1985	2211	2215	2712	2387	2801	2912	3058	3032	2812
D	10	12	11	15	17	18	20	19	22	21	23



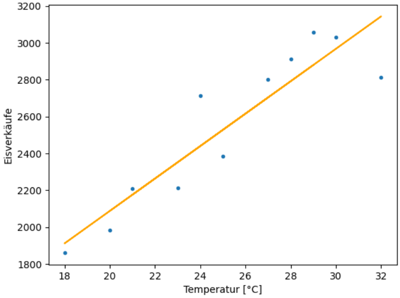
Eisverkauf

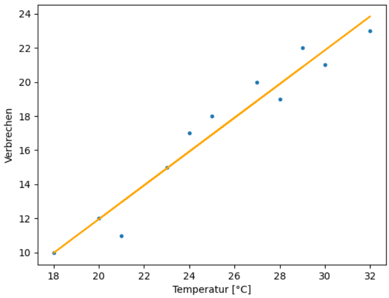
Sei T die Temperatur.

T	18	20	21	23	24	25	27	28	29	30	32
E	1860	1985	2211	2215	2712	2387	2801	2912	3058	3032	2812
D	10	12	11	15	17	18	20	19	22	21	23

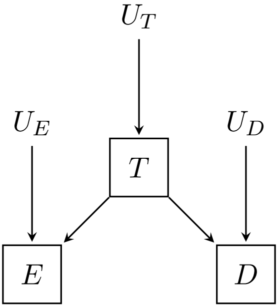
Sei T die Temperatur.

T	18	20	21	23	24	25	27	28	29	30	32
E	1860	1985	2211	2215	2712	2387	2801	2912	3058	3032	2812
D	10	12	11	15	17	18	20	19	22	21	23





Eisverkauf



Inhaltsverzeichnis

Definition 4.19 (Intervention)

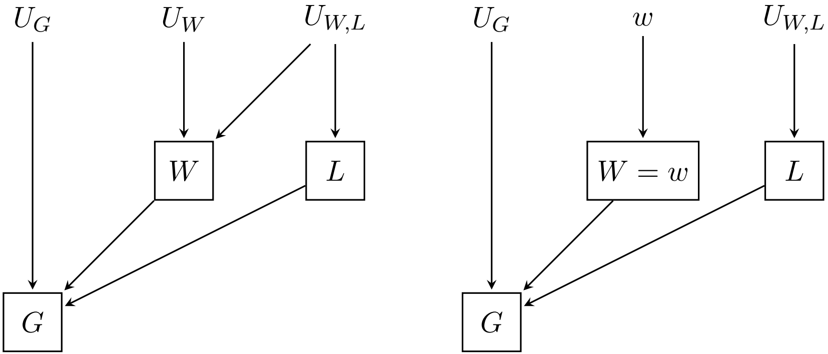
Sei X eine Zufallsvariable und $x \in \mathrm{Im}(X)$ einer ihrer Werte. Eine *Intervention* ist eine Maßnahme, um den Wert der Zufallsvariable auf x zu fixieren. Dies schreiben wir als $\mathrm{do}(X = x)$.

Definition 4.19 (Intervention)

Sei X eine Zufallsvariable und $x \in \mathrm{Im}(X)$ einer ihrer Werte. Eine *Intervention* ist eine Maßnahme, um den Wert der Zufallsvariable auf x zu fixieren. Dies schreiben wir als $\mathrm{do}(X = x)$.

Lokalität: Intervention hat keine Nebenwirkungen.

Pflanzenwachstum



Präintervention

E	1860	1985	2211	2215	2712	2387	2801	2912	3058	3032	2812
D	10	12	11	15	17	18	20	19	22	21	23

$$\mathbb{E}(D \mid E \leq 2500) \approx 13.2$$

Eisverkauf

Präintervention

E	1860	1985	2211	2215	2712	2387	2801	2912	3058	3032	2812
D	10	12	11	15	17	18	20	19	22	21	23

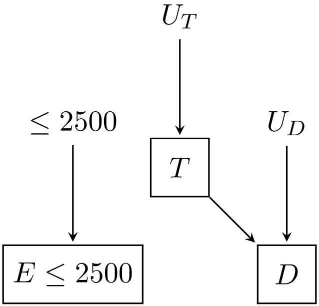
$$\mathbb{E}(D \mid E \leq 2500) \approx 13.2$$

Postintervention

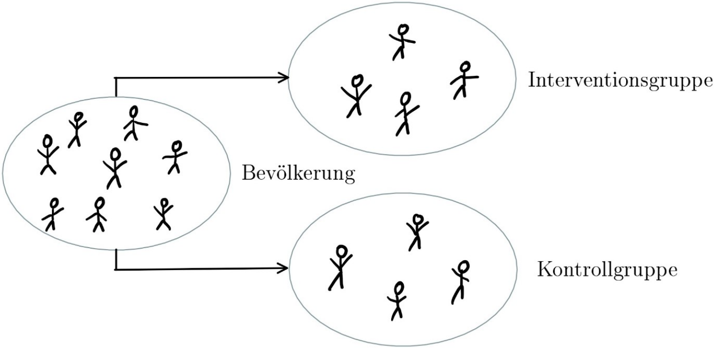
E	1864	1992	2221	2315	2500	2387	2500	2500	2500	2500	2500
D	10	11	10	15	18	18	19	20	21	21	24

$$\mathbb{E}(D \mid \mathrm{do}(E \leq 2500)) \approx 17.1$$

Eisverkauf



Randomisierte kontrollierte Studie



Unterschied Konditionierung und Intervention

Seien X und Y Zufallsvariablen und y ein Wert von Y.
$\mathbb{P}(X = \cdot \mid Y = y)$ ist die Verteilung von X unter der Teilmenge der Individuen mit $Y = y$.

Unterschied Konditionierung und Intervention

Seien X und Y Zufallsvariablen und y ein Wert von Y.
$\mathbb{P}(X = \cdot \mid Y = y)$ ist die Verteilung von X unter der Teilmenge der Individuen mit $Y = y$.

$\mathbb{P}(X = \cdot \mid \mathrm{do}(Y = y))$ ist die Verteilung von X unter einer Bevölkerung, die ausschließlich $Y = y$ besitzt.

Inhaltsverzeichnis

Definition 4.25 (Durchschnittlicher kausaler Effekt)

Seien X und Y Variablen und X sogar *binär*, habe also nur die Werte 0 und 1. Wir definieren den *durchschnittlichen kausalen Effekt von X auf Y* (ACE) (*average causal effect, causal effect difference*) durch

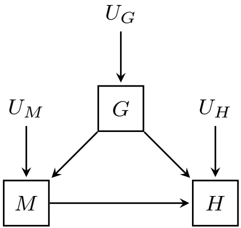
$$\mathrm{ACE}(X \to Y) := \mathbb{P}(Y = y \mid \mathrm{do}(X = 1)) - \mathbb{P}(Y = y \mid \mathrm{do}(X = 0)) .$$

Beschreibe M die Gabe des Medikaments, G das Geschlecht und H die Heilung.

	$M = 0$	$M = 1$
$G = 0$	55 von 80 (69%)	192 von 263 (73%)
$G = 1$	234 von 270 (87%)	81 von 87 (93%)
Gesamt	289 von 350 (83%)	273 von 350 (78%)

Beschreibe M die Gabe des Medikaments, G das Geschlecht und H die Heilung.

	$M = 0$	$M = 1$
$G = 0$	55 von 80 (69%)	192 von 263 (73%)
$G = 1$	234 von 270 (87%)	81 von 87 (93%)
Gesamt	289 von 350 (83%)	273 von 350 (78%)



Naiv

$$\mathbb{P}(H = 1 \mid M = 1) - \mathbb{P}(H = 1 \mid M = 0) = -5\% \ .$$

Naiv

$$\mathbb{P}(H = 1 \mid M = 1) - \mathbb{P}(H = 1 \mid M = 0) = -5\% \ .$$

Paradox, weil

$$\mathbb{P}(H = 1 \mid G = 0, M = 1) > \mathbb{P}(H = 1 \mid G = 0, M = 0)$$

und

$$\mathbb{P}(H = 1 \mid G = 1, M = 1) > \mathbb{P}(H = 1 \mid G = 1, M = 0) \ .$$

Simpson Paradoxon

Naiv

$$\mathbb{P}(H = 1 \mid M = 1) - \mathbb{P}(H = 1 \mid M = 0) = -5\% \ .$$

Paradox, weil

$$\mathbb{P}(H = 1 \mid G = 0, M = 1) > \mathbb{P}(H = 1 \mid G = 0, M = 0)$$

und

$$\mathbb{P}(H = 1 \mid G = 1, M = 1) > \mathbb{P}(H = 1 \mid G = 1, M = 0) \ .$$

Daher

$$\mathrm{ACE}(M \to H) = \mathbb{P}(H = 1 \mid \mathrm{do}(M = 1)) - \mathbb{P}(H = 1 \mid \mathrm{do}(M = 0)) \ .$$

Intervention zur Auflösung von $G \to M$.

$$
\begin{tikzcd}
 & U_G \arrow[d] & \\
m \arrow[d] & \boxed{G} \arrow[dr] & U_H \arrow[d] \\
\boxed{M = m} \arrow[rr] & & \boxed{H}
\end{tikzcd}
$$

Simpson Paradoxon

Die noch zu zeigende Adjustierungsformel liefert

$$\begin{aligned}\mathbb{P}(H=1 \mid \mathrm{do}(M=1)) &= \mathbb{P}(H=1 \mid M=1, G=0)\mathbb{P}(G=0)\\ &+ \mathbb{P}(H=1 \mid M=1, G=1)\mathbb{P}(G=1)\\ &= 73\% \cdot \frac{263+80}{700} + 93\% \cdot \frac{87+270}{700}\\ &= 35.77\% + 47.43\% = 83.2\%\end{aligned}$$

und

$$\begin{aligned}\mathbb{P}(H=1 \mid \mathrm{do}(M=0)) &= \mathbb{P}(D=1 \mid M=0, G=0)\mathbb{P}(G=0)\\ &+ \mathbb{P}(H=1 \mid M=0, G=1)\mathbb{P}(G=1)\\ &= 69\% \cdot \frac{263+80}{700} + 87\% \cdot \frac{87+270}{700}\\ &= 33.81\% + 44.37\% = 78.18\%\ .\end{aligned}$$

Simpson Paradoxon

Daraus erhalten wir

$$\begin{aligned}\mathrm{ACE}(M \to H) &= \mathbb{P}(H = 1 \mid \mathrm{do}(M = 1)) \\ &\quad - \mathbb{P}(H = 1 \mid \mathrm{do}(M = 0)) \\ &= 5.02\% \ .\end{aligned}$$

Entspricht der Intuition.

Satz 4.27 (Adjustierungsformel)

Seien X, Y und Z Zufallsvariablen, sodass Z eine Ursache für X und Y ist. Weiter sei Y noch eine Folge von X. Außerdem sei x ein Wert von X und y ein Wert von Y. Nach der *Adjustierungsformel* (*adjustment formula*) gilt dann

$$\mathbb{P}(Y = y \mid \mathrm{do}(X = x)) = \sum_{z \in \mathrm{Im}(Z)} \mathbb{P}(Y = y \mid X = x, Z = z) \cdot \mathbb{P}(Z = z) \; .$$

Blutdruck

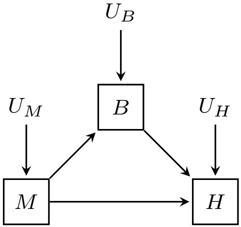
Beschreibe M die Gabe des Medikaments, B den Blutdruck nach Gabe des Medikaments / Placebos und H die Heilung.

	$M = 0$	$M = 1$
$B = 0$	81 von 87 (93%)	234 von 270 (87%)
$B = 1$	192 von 263 (73%)	55 von 80 (69%)
Gesamt	273 von 350 (78%)	289 von 350 (83%)

Blutdruck

Beschreibe M die Gabe des Medikaments, B den Blutdruck nach Gabe des Medikaments / Placebos und H die Heilung.

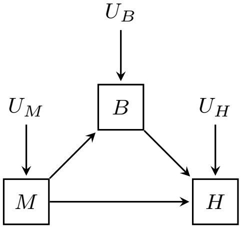
	$M = 0$	$M = 1$
$B = 0$	81 von 87 (93%)	234 von 270 (87%)
$B = 1$	192 von 263 (73%)	55 von 80 (69%)
Gesamt	273 von 350 (78%)	289 von 350 (83%)



Blutdruck

Beschreibe M die Gabe des Medikaments, B den Blutdruck nach Gabe des Medikaments / Placebos und H die Heilung.

	$M = 0$	$M = 1$
$B = 0$	81 von 87 (93%)	234 von 270 (87%)
$B = 1$	192 von 263 (73%)	55 von 80 (69%)
Gesamt	273 von 350 (78%)	289 von 350 (83%)



Da M bis auf U_M keine Eltern hat, gilt

$$\mathbb{P}(H = h \mid \mathrm{do}(M = m)) = \mathbb{P}(H = h \mid M = m) \ .$$

Inhaltsverzeichnis

Satz 4.30 (Die Regel vom kausalen Effekt)

Sei ein Kausalgraph gegeben und $\mathrm{Pa}(X)$ der Zufallsvektor der Eltern von X, sowie Y eine weiter Zufallsvariable. Dann gilt für den kausalen Effekt von X auf Y

$$\mathbb{P}(Y = y \mid \mathrm{do}(X = x)) = \sum_{z \in \mathrm{Im}(\mathrm{Pa}(X))} \mathbb{P}(Y = y \mid X = x, \mathrm{Pa}(X) = z) \cdot \mathbb{P}(\mathrm{Pa}(X) = z) \ .$$

Hierbei seien y und x Werte von Y respektive X und z ein mögliche Kombination von Werten des Vektors $\mathrm{Pa}(X)$.

Beschreibe M die Gabe des Medikaments, G das Geschlecht, B den Blutdruck nach Gabe des Medikaments / Placebos und H die Heilung.

Beschreibe M die Gabe des Medikaments, G das Geschlecht, B den Blutdruck nach Gabe des Medikaments / Placebos und H die Heilung.

$$
\begin{tikzcd}
U_G \arrow[r] & G \arrow[r] \arrow[d] \arrow[dr] & B \arrow[dl] & U_B \arrow[l] \\
U_H \arrow[r] & H & M \arrow[l] \arrow[u] & U_M \arrow[l]
\end{tikzcd}
$$

Blutdruck und Geschlecht

	$M = 0$	$M = 1$
$G = 0$	128 von 500 (26%)	372 von 500 (74%)
$G = 1$	52 von 500 (10%)	448 von 500 (90%)

$$\begin{aligned}
\mathbb{P}(H = 1 \mid B = 0, G = 0, M = 0) &= 15\% \\
\mathbb{P}(H = 1 \mid B = 0, G = 1, M = 0) &= 16\% \\
\mathbb{P}(H = 1 \mid B = 0, G = 0, M = 1) &= 53\% \\
\mathbb{P}(H = 1 \mid B = 0, G = 1, M = 1) &= 42\% \;.
\end{aligned}$$

Blutdruck und Geschlecht

	$M = 0$	$M = 1$
$G = 0$	128 von 500 (26%)	372 von 500 (74%)
$G = 1$	52 von 500 (10%)	448 von 500 (90%)

$$\begin{aligned}
\mathbb{P}(H = 1 \mid B = 0, G = 0, M = 0) &= 15\% \\
\mathbb{P}(H = 1 \mid B = 0, G = 1, M = 0) &= 16\% \\
\mathbb{P}(H = 1 \mid B = 0, G = 0, M = 1) &= 53\% \\
\mathbb{P}(H = 1 \mid B = 0, G = 1, M = 1) &= 42\% \; .
\end{aligned}$$

Mit $\mathrm{Pa}(B) = \{G, M, U_B\}$ folgt mit der Regel vom kausalen Effekt

$$\begin{aligned}
\mathbb{P}(H = 1 \mid \mathrm{do}(B = 0)) &= \sum_{z \in \{0,1\}^2} \mathbb{P}(H = 1 \mid B = 0, G = g, M = m) \cdot \mathbb{P}(G = g, M = m) \\
&\approx 83\% \; .
\end{aligned}$$

Poissonverteilung

Sei

$$
\begin{aligned}
U_V &\sim \text{Poiss}(5)\\
U_W &\sim \text{Poiss}(5)\\
U_X &\sim \text{Poiss}(3)\\
U_Y &\sim \text{Poiss}(1)\\
U_Z &\sim \text{Poiss}(1)
\end{aligned}
$$

und

$$
\begin{aligned}
V &:= U_V\\
W &:= U_W\\
X &:= V + U_X\\
Y &:= V + W + X + U_Y\\
Z &:= X + Y + U_Z \ .
\end{aligned}
$$

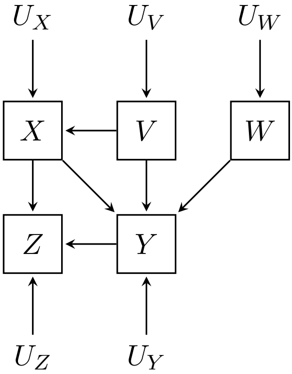
Poissonverteilung

Sei

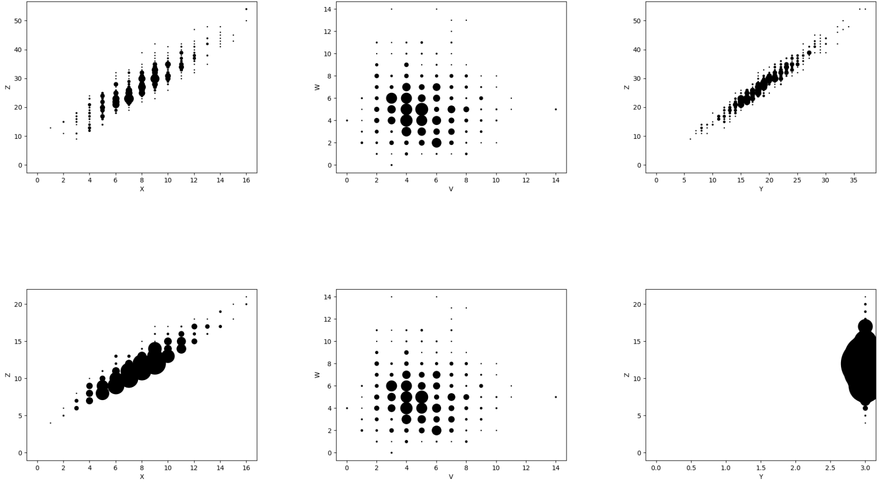
$$\begin{aligned} U_V &\sim \text{Poiss}(5) \\ U_W &\sim \text{Poiss}(5) \\ U_X &\sim \text{Poiss}(3) \\ U_Y &\sim \text{Poiss}(1) \\ U_Z &\sim \text{Poiss}(1) \end{aligned}$$

und

$$\begin{aligned} V &:= U_V \\ W &:= U_W \\ X &:= V + U_X \\ Y &:= V + W + X + U_Y \\ Z &:= X + Y + U_Z \ . \end{aligned}$$



Poissonverteilung



Poissonverteilung

Die Regel vom kausalen Effekt liefert

$$\mathbb{P}(Z = z \mid \mathrm{do}(Y = y)) = \sum_{(v,w,x)\in\mathbb{N}_0^3} \mathbb{P}(Z = z \mid Y = y, V = v, W = w, X = x) \cdot \mathbb{P}(V = v, W = w, X = x)\ .$$

Nach einer Simulation für $Z = 10$ und $Y = 3$ ist dies ungefähr 12%.

Normalverteilung

Sei

$$
\begin{aligned}
U_V &\sim \mathcal{N}(0,1)\\
U_W &\sim \mathcal{N}(0,2)\\
U_X &\sim \mathcal{N}(0,1)\\
U_Y &\sim \mathcal{N}(0,0.1)\\
U_Z &\sim \mathcal{N}(0,0.2)
\end{aligned}
$$

und

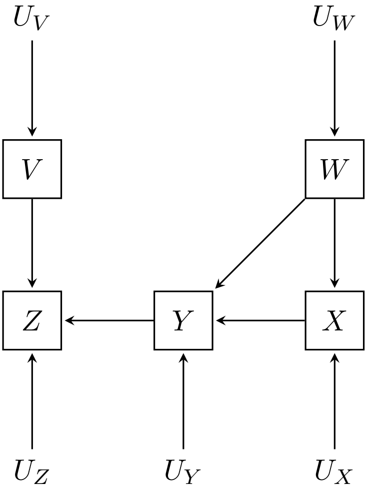
$$
\begin{aligned}
V &:= U_V\\
W &:= U_W\\
X &:= W + U_X\\
Y &:= W + X + U_Y\\
Z &:= V + Y + U_Z \ .
\end{aligned}
$$

Sei

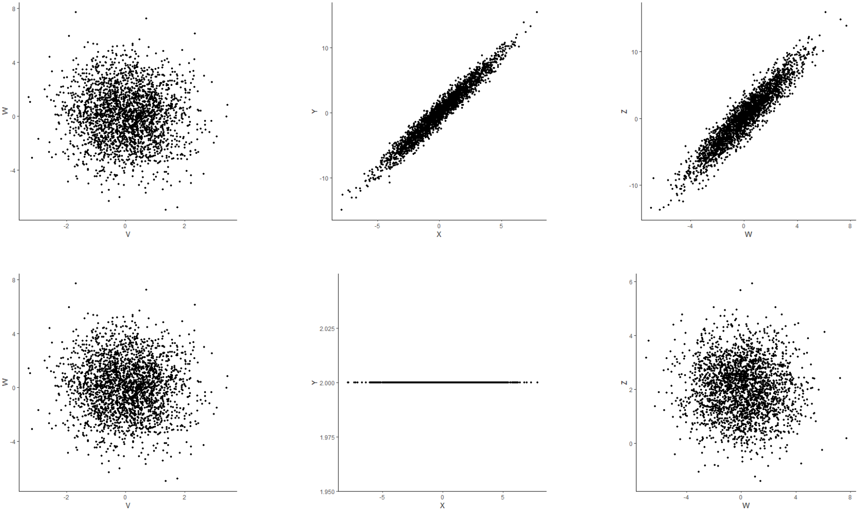
$$\begin{aligned} U_V &\sim \mathcal{N}(0,1) \\ U_W &\sim \mathcal{N}(0,2) \\ U_X &\sim \mathcal{N}(0,1) \\ U_Y &\sim \mathcal{N}(0,0.1) \\ U_Z &\sim \mathcal{N}(0,0.2) \end{aligned}$$

und

$$\begin{aligned} V &:= U_V \\ W &:= U_W \\ X &:= W + U_X \\ Y &:= W + X + U_Y \\ Z &:= V + Y + U_Z \ . \end{aligned}$$



Normalverteilung



Normalverteilung

Man könnte hier jetzt simulieren

$$\mathbb{P}(Z \in [0,1] \mid \mathrm{do}(Y = 2)) \approx 14\% \ .$$

Inhaltsverzeichnis

Satz 4.36 (Produktdekomposition)

Sei ein Kausalgraph gegeben und $X_1, \ldots, X_n$ alle Variablen in diesem Modell, sowie $x_1, \ldots, x_n$ jeweils Werte dieser Zufallsvariablen. Es gilt die folgende *Produktdekomposition*

$$\mathbb{P}(X_1 = x_1, \ldots, X_n = x_n) = \prod_{i=1}^{n} \mathbb{P}(X_i = x_i \mid \mathrm{Pa}(X_i) = \mathrm{pa}(X_i)) \; .$$

Hierbei sei $\mathrm{pa}(X_i)$ konsistent mit $x_1, \ldots, x_n$.

Abgeschnittene Produktregel

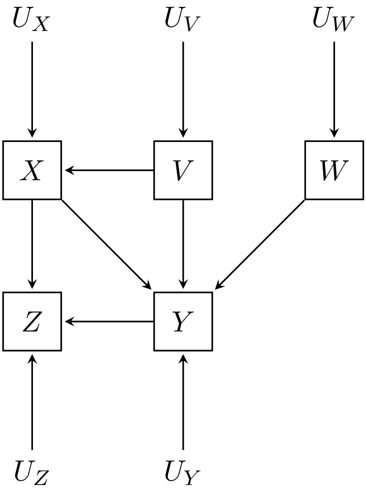
Satz 4.35 (Abgeschnittene Produktregel)

Sei ein Kausalgraph gegeben und $X_1, \dots, X_n$ alle Variablen in diesem Modell, sowie $x_1, \dots, x_n$ jeweils Werte dieser Zufallsvariablen. Sei X eine Zufallsvektor mit Variablen aus dem Modell. Ist x ein Wert des Zufallsvaktors, so gilt die *abgeschnittene Produktregel*

$$\mathbb{P}(X_1 = x_1, \dots, X_n = x_n \mid \mathrm{do}(X = x)) = \prod_{i \in I} \mathbb{P}(X_i = x_i \mid \mathrm{Pa}(X_i) = \mathrm{pa}(X_i)) \ .$$

Hierbei sei $I \subset \{1, \dots, n\}$ die Menge aller Indizes der Variablen, die nicht in X sind. Weiter sei x und $\mathrm{pa}(X_i)$ konsistent mit $x_1, \dots, x_n$.
Ist dies nicht konsistent, so verschwindet die Interventionswahrscheinlichkeit.

Poissonverteilung



Inhaltsverzeichnis

Probleme

(i) Lokalität
(ii) Nicht-manipulierbare Variablen
(iii) Azyklizität
(iv) Interpretation der Assoziationen
(v) Kausalgraph
(vi) Ungemessene Variablen
(viii) Verstetigung