

Blatt 6 (10 Punkte)

Abgabe durch Hochladen (nur PDF-Format bzw. Python-Code) auf der eCampus-Seite bis
Sonntag, 19.05.2024, 12:00 Uhr, in Gruppen von 3 Personen.

Aufgabe 6.1: Hintergrundevidenzen

(1)

Beweisen Sie die folgende Verallgemeinerung der Bayes' Regel, bei der e eine Hintergrundevidenz darstellt. Es sollte klar sein, dass ein Beweis nicht aus dem Rechnen eines Beispiels folgt! Wenn Sie Definitionen aus der Vorlesung verwenden, geben Sie dies an der jeweiligen Stelle an.

$$P(Y|X, e) = \frac{P(X | Y, e) \cdot P(Y | e)}{P(X | e)}$$

Aufgabe 6.2: Bedingte Wahrscheinlichkeit

(1 + 0 = 1)

Durch Information aus der Beobachtung statistisch abhängiger Zufallsvariablen ändern sich auch die Annahmen über die nicht beobachtete Zufallsvariable. Geben Sie zu folgenden Problemen die a priori Wahrscheinlichkeitsverteilungen und die bedingte Verteilung an.

Beispiel für die Darstellung der unbedingten Wahrscheinlichkeitsverteilung für die Aussage X = "Das Ergebnis eines Wurfes mit einem Würfel ist 4": $P(X) = \langle \frac{1}{6}, \frac{5}{6} \rangle$

- a) $P(A)$: Eine Person hat am 02.06. Geburtstag.
 $P(B)$: Eine Person ist Sternzeichen Zwilling (22.05.-21.06.).
 $P(A | b)$: Eine Person, die Zwilling ist, hat am 02.06. Geburtstag.

Setzen Sie die Zahl der Tage eines Jahres mit 365 Tagen an.

- b) Besprechen Sie gemeinsam **in der Übungsgruppe**:
 $P(C)$: Das Erstgeborene einer Familie mit zwei Kindern ist ein Mädchen.
 $P(D)$: Mindestens ein Kind ist ein Junge.
 $P(C | d)$: Das Erstgeborene ist ein Mädchen, gegeben mindestens ein Kind ist ein Junge.

Aufgabe 6.3: Rechnen mit Verbundverteilungen

(0,5 + 0,5 + 1 = 2)

Betrachten Sie die folgende Tabelle mit Verbundwahrscheinlichkeiten:

	übergewichtig		¬übergewichtig	
	raucher	¬ raucher	raucher	¬raucher
älter75	0,8%	2,0%	1%	2,2%
¬älter75	14%	15%	34%	31%

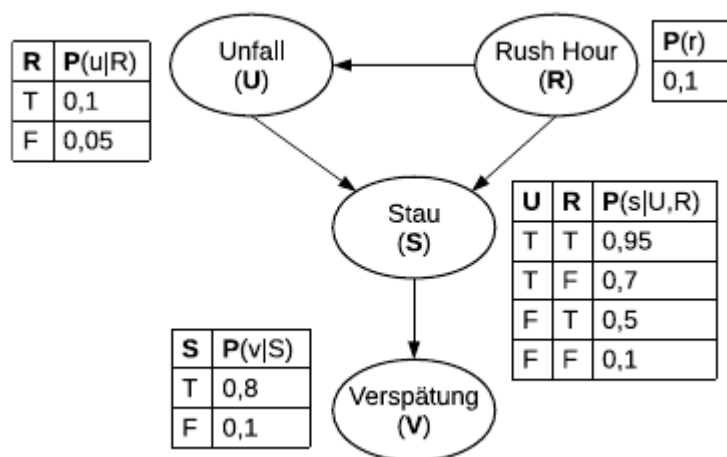
Berechnen Sie mit Hilfe der Tabelle folgende Wahrscheinlichkeitsverteilungen:

- $P(\text{Älter75})$, $P(\text{Raucher})$ und $P(\text{Übergewichtig})$
- $P(\text{Übergewichtig} \mid \neg \text{raucher})$
- $P(\text{Raucher} \mid \text{älter75} \vee \text{übergewichtig})$

Aufgabe 6.4: Bayessche Netzwerke: Verspätung

(0,5 + 2,5 = 3)

Gegeben sei folgendes Bayessche Netzwerk:



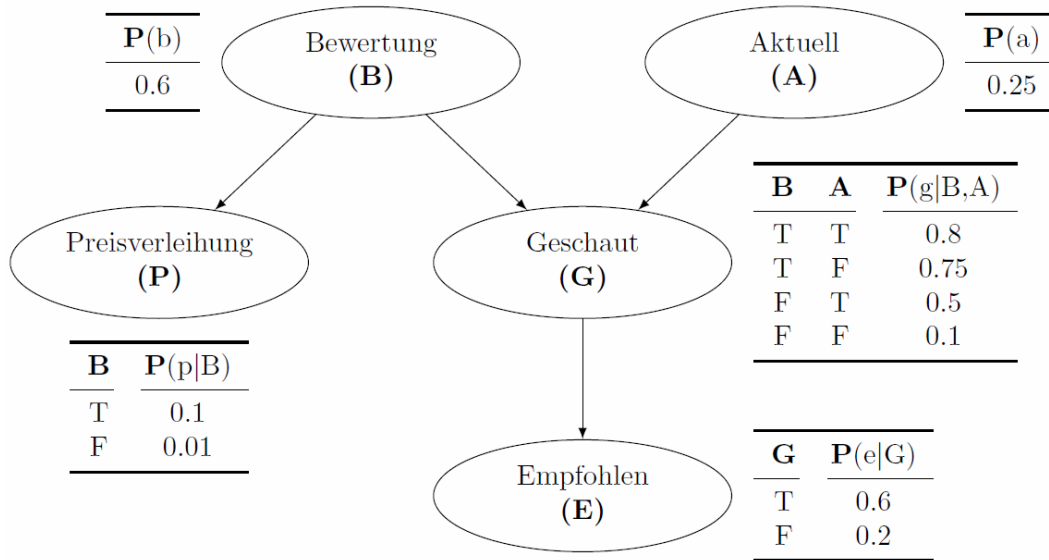
- Geben Sie die der Netztopologie entsprechende, in bedingte Wahrscheinlichkeiten faktorierte Schreibweise für $P(V, S, U, R)$ an. Berechnen Sie mit Hilfe der CPTs die Wahrscheinlichkeit für $P(\neg v, s, u, r)$.
- Berechnen Sie mittels exakter Inferenz durch Aufzählen die Anfrageverteilung *Verspätung* gegeben die Evidenzen *Unfall* und *Rush Hour*, d.h. $P(V|u, r)$. Es ist also eine **vollständige Fallanalyse** durchzuführen (s. Vorl. 10, Folien 26 ff. sowie Vorl. 11, Folien 18 - 20)! Nutzen Sie für einfacheres Rechnen eine Normierungskonstante (s. Vorl. 10, Folie 27).

Beachten Sie, dass von Aufgaben 6.5 und 6.6 nur eine auszusuchen ist, die abgegeben werden kann/muss. Es wird dementsprechend nur eine Aufgabe bepunktet korrigiert. Bei Abgabe beider Aufgaben bitte kenntlich machen, welche bewertet werden soll. Erfolgt in diesem Fall keine solche Kennzeichnung, wird Aufg. 6.5 bewertet.

Aufgabe 6.5: Sampling Methoden

(entweder hier 1 + 2 = 3)

Sei das folgende Bayessche Netzwerk, das sich auf Filme bezieht, gegeben:



Der Algorithmus PRIOR-SAMPLE erzeugt die folgenden Stichproben, die für Aufgabenteile a) und b) zu nutzen sind:

- | | | |
|-----------------------------------|--|--|
| 1. $\neg b$ a $\neg p$ g $\neg e$ | 5. b $\neg a$ $\neg p$ g e | 9. b a p $\neg g$ $\neg e$ |
| 2. b $\neg a$ p $\neg g$ $\neg e$ | 6. $\neg b$ $\neg a$ $\neg p$ g $\neg e$ | 10. $\neg b$ $\neg a$ $\neg p$ $\neg g$ e |
| 3. b $\neg a$ $\neg p$ g e | 7. b a $\neg p$ $\neg g$ e | 11. b a $\neg p$ g e |
| 4. $\neg b$ a $\neg p$ g $\neg e$ | 8. b $\neg a$ $\neg p$ g $\neg e$ | 12. $\neg b$ $\neg a$ $\neg p$ $\neg g$ $\neg e$ |

a) Bestimmen Sie unter Anwendung von Algorithmus REJECTION-SAMPLING die folgenden Wahrscheinlichkeiten:

- $P(e|g)$
- $P(e|b, g)$

b) Bestimmen Sie nun dieselben Wahrscheinlichkeiten unter Anwendung von LIKELIHOOD-WEIGHTING. Dokumentieren Sie dabei für jede Query zuerst die Gewichte der betrachteten Stichproben (natürlich nur solche, die konsistent zur Evidenz sind) und berechnen Sie anschließend die Wahrscheinlichkeiten.

Aufgabe 6.6: Programmieraufgabe: Rasensprenger

(oder hier 3)

Vorbereitung: Laden Sie das ZIP-Archiv *bayes.zip* herunter von unserer eCampus-Seite unter Kursunterlagen >> Python und AIMA Python >> Bayesian Networks. Das ZIP-Archiv enthält den Ordner *bayes* mit dem Skript *prior_sampling.py* sowie dem vorgegebenen Bayes-sches Netzwerk *gras_wet.bayes*. Fügen Sie den Ordner *bayes* auf der gleichen Ebene wie den Ordner *aima* ein.

Aufgabe:

- a) Erstellen Sie auf der Basis des gegebenen Skripts *prior_sampling.py* ein neues Skript *reject_sampling.py*, welches das Rejection-Sampling umsetzt.
- b) Schätzen Sie nun mit Hilfe Ihres neu erstellten Skripts *reject_sampling.py* die Verteilung $P(\text{Gras_wet} | \text{Cloudy} = \text{true})$ über 100 Iterationen (diese Einstellungen können Sie über den Properties-Dialog der GUI einstellen).

Die Punkte werden für die Lösung von Teil a) vergeben. Es werden jedoch keine Punkte für Teil a) vergeben ohne vernünftigen Abschluss durch zumindest den einen in Teil b) geforderten Testlauf und die entsprechende Antwort auf Teil b).