

# plan

Christos Dimitrakakis

November 20, 2019

## Contents

<b>1</b>	<b>Artificial Intelligence, Science and Society</b>	<b>2</b>
1.1	Course content . . . . .	3
1.2	Prerequisites . . . . .	3
1.2.1	Essential . . . . .	3
1.2.2	Recommended . . . . .	3
1.3	Learning outcomes . . . . .	3
1.4	Prerequisites . . . . .	4
1.5	Course content . . . . .	4
1.6	Examination . . . . .	5
1.6.1	A . . . . .	5
1.6.2	B . . . . .	6
1.6.3	C . . . . .	7
1.6.4	D . . . . .	7
1.6.5	E . . . . .	8
1.6.6	F . . . . .	8
1.7	Motivation . . . . .	8
<b>2</b>	<b>Schedule</b>	<b>9</b>
2.1	2019 . . . . .	9
<b>3</b>	<b>Lecture plan</b>	<b>11</b>
3.1	Module 1: Reproducibility . . . . .	11
3.1.1	kNN ACTIVITY . . . . .	11
3.1.2	Reproducibility ACTIVITY . . . . .	11
3.1.3	Modelling TUTORIAL . . . . .	11
3.2	Module 2: Decision problems . . . . .	12
3.2.1	Decision Problems . . . . .	12

3.2.2	Project Introduction: Credit risk for mortgages. TU-TORIAL . . . . .	12
3.3	Week 3: Privacy . . . . .	12
3.4	Week 4: Fairness . . . . .	12
3.5	Project start: Experiment design for Medical Diagnostics [Aim: Reproducibility, Safety] PROJECT . . . . .	13
3.6	Week 5: Clustering . . . . .	13
3.7	Week 6: The web and recommendation systems . . . . .	13
3.7.1	Online data collection. Optimal stopping (expensive labels) A/B Testing, Bandit Problems. . . . .	13
3.7.2	Markov decision processes and Dynamic Programming (active learning and experiment design more generally) . . . . .	13
3.8	Optional. Safety: Risk-Sensitive Decision Making . . . . .	13
<b>4</b>	<b>Exam subjects</b>	<b>14</b>
4.1	Reproducibility . . . . .	14
4.1.1	Possible answer . . . . .	14
4.2	Conditional probability and Bayesian inference . . . . .	15
4.2.1	Possible answer . . . . .	15
4.3	Utility . . . . .	16
4.3.1	Possible answer . . . . .	16
4.4	Differential privacy . . . . .	17
4.4.1	Possible answer . . . . .	17
4.5	Graphical models . . . . .	17
4.5.1	Possible answer . . . . .	17
4.6	Conditional independence . . . . .	18
4.6.1	Possible answer . . . . .	18
4.7	Fairness . . . . .	18
4.8	Causality . . . . .	20
4.8.1	Possible answer . . . . .	21
4.9	Markov decision processes and experiment design . . . . .	21
4.9.1	Possible answer . . . . .	22

## 1 Artificial Intelligence, Science and Society

- ECTS Credits: 6.
- Expected development time: 400 hours (200 for the lecture notes and 200 for the assignments, tutorials and projects)

- Teaching time: 264-384 hours (48 contact, 96 preparation, 120-240 corrections for 20-40 students)

## 1.1 Course content

Classic approaches in data analysis are use a static procedure for both collecting and processing data. Modern approaches deal with the adaptive procedures which in practice almost always are used.

In this course you will learn how to design systems that adaptively collect and process data in order to make decisions autonomously or in collaboration with humans.

The course applies core principles from machine learning, artificial intelligence and databases to real-world problems in safety, reproducibility, causal reasoning, privacy and fairness.

## 1.2 Prerequisites

### 1.2.1 Essential

- Mathematics R1+R2
- Python programming (e.g. IN1900 – Introduction to Programming with Scientific Applications).

### 1.2.2 Recommended

- Elementary knowledge of probability and statistics (STK1000/STK1100)
- Elementary calculus and linear algebra (MAT1100 or MAT1110)

## 1.3 Learning outcomes

There are two types of learning outcomes. Firstly, those that are the core of the course, and secondly methodologies that are used as part of the course.

Core learning outcomes:

1. Ensuring reproducibility in both science and AI development.
2. Recognising privacy issues and be able to mitigate them using appropriate formalisms.
3. Mitigating issues with potential fairness and discrimination when algorithms are applied at scale.

4. Performing inference when there are causal elements.
5. Developing adaptive experimental design protocols for online and scientific applications.
6. Understanding when it is possible to provide performance guarantees for AI algorithms.

AI learning outcomes:

1. Understanding how to use data for learning, estimation and testing to create reproducible research.
2. Understanding Bayesian inference and decision theory and being able to describe dependencies with graphical models.
3. Understanding neural networks and how to apply stochastic optimisation algorithms.
4. Understanding and using differential privacy as a formalism.
5. Understanding causal inference, interventions and counterfactuals.
6. Understanding the recommendation problem in terms of both modelling and decision making.

## **1.4 Prerequisites**

## **1.5 Course content**

The course is split in 6 modules, which should be taken in sequence.

Module 1. Reproducibility: bootstrapping, Bayesian inference, decision problems, false discovery, confidence bounds. Module 2. Privacy: Databases, k-anonymity, graphical models, differential privacy. Module 3. Fairness: Decision diagrams, conditional independence, meritocracy, discrimination. Module 4. The web: Recommendation systems, clustering, latent variable models. Module 5. Causality: Interventions and counterfactuals. Module 6. Adaptive experiment design: Bandit problems, stochastic optimisation, Markov decision processes, dynamic programming.

## 1.6 Examination

There are 2 projects (formally take-home exams), split into 3 parts each. Each one takes 2-4 hours and is partly done in a tutorial session.

Each question is weighted equally in each home exam, so that by correctly answering the elementary parts of each question, students can be guaranteed a passing grade. Each exam counts for 40% of the score. A final exam is also given by the students. This counts for 20% of the final score.

Criteria for full marks in each part of the exam are the following.

1. Documenting of the work in a way that enables reproduction.
2. Technical correctness of their analysis.
3. Demonstrating that they have understood the assumptions underlying their analysis.
4. Addressing issues of reproducibility in research.
5. Addressing ethical questions where applicable, and if not, clearly explain why they are not.
6. Consulting additional resources beyond the source material with proper citations.

The follow marking guidelines are what one would expect from students attaining each grade.

### 1.6.1 A

1. Submission of a detailed report from which one can definitely reconstruct their work without referring to their code. There should be no ambiguities in the described methodology. Well-documented code where design decisions are explained.
2. Extensive analysis and discussion. Technical correctness of their analysis. Nearly error-free implementation.
3. The report should detail what models are used and what the assumptions are behind them. The conclusions of the should include appropriate caveats. When the problem includes simple decision making, the optimality metric should be well-defined and justified. Simiarly, when well-defined optimality criteria should given for the experiment

design, when necessary. The design should be (to some degree of approximation, depending on problem complexity) optimal according to this criteria.

4. Appropriate methods to measure reproducibility. Use of cross-validation or hold-out sets to measure performance. Use of an unbiased methodology for algorithm, model or parameter selection. Appropriate reporting of a confidence level (e.g. using bootstrapping) in their analytical results. Relevant assumptions are mentioned when required.
5. When dealing with data relating to humans, privacy and/or fairness should be addressed. A formal definition of privacy and/or should be selected, and the resulting policy should be examined.
6. The report contains some independent thinking, or includes additional resources beyond the source material with proper citations. The students go beyond their way to research material and implement methods not discussed in the course.

### **1.6.2 B**

1. Submission of a report from which one can plausibly reconstruct their work without referring to their code. There should be no major ambiguities in the described methodology.
2. Technical correctness of their analysis, with a good discussion. Possibly minor errors in the implementation.
3. The report should detail what models are used, as well as the optimality criteria, including for the experiment design. The conclusions of the report must contain appropriate caveats.
4. Use of cross-validation or hold-out sets to measure performance. Use of an unbiased methodology for algorithm, model or parameter selection.
5. When dealing with data relating to humans, privacy and/or fairness should be addressed. While an analysis of this issue may not be performed, there is a substantial discussion of the issue that clearly shows understanding by the student.
6. The report contains some independent thinking, or the students mention other methods beyond the source material, with proper citations, but do not further investigate them.

### 1.6.3 C

1. Submission of a report from which one can partially reconstruct most of their work without referring to their code. There might be some ambiguities in parts of the described methodology.
2. Technical correctness of their analysis, with an adequate discussion. Some errors in a part of the implementation.
3. The report should detail what models are used, as well as the optimality criteria and the choice of experiment design. Analysis caveats are not included.
4. Either use of cross-validation or hold-out sets to measure performance, or use of an unbiased methodology for algorithm, model or parameter selection - but in a possibly inconsistent manner.
5. When dealing with data relating to humans, privacy and/or fairness are addressed superficially.
6. There is little mention of methods beyond the source material or independent thinking.

### 1.6.4 D

1. Submission of a report from which one can partially reconstruct most of their work without referring to their code. There might be serious ambiguities in parts of the described methodology.
2. Technical correctness of their analysis with limited discussion. Possibly major errors in a part of the implementation.
3. The report should detail what models are used, as well as the optimality criteria. Analysis caveats are not included.
4. Either use of cross-validation or hold-out sets to measure performance, or use of an unbiased methodology for algorithm, model or parameter selection - but in a possibly inconsistent manner.
5. When dealing with data relating to humans, privacy and/or fairness are addressed superficially or not at all.
6. There is little mention of methods beyond the source material or independent thinking.

### **1.6.5 E**

1. Submission of a report from which one can obtain a high-level idea of their work without referring to their code. There might be serious ambiguities in all of the described methodology.
2. Technical correctness of their analysis with very little discussion. Possibly major errors in only a part of the implementation.
3. The report might mention what models are used or the optimality criteria, but not in sufficient detail and caveats are not mentioned.
4. Use of cross-validation or hold-out sets to simultaneously measure performance and optimise hyperparameters, but possibly in a way that introduces some bias.
5. When dealing with data relating to humans, privacy and/or fairness are addressed superficially or not at all.
6. There is no mention of methods beyond the source material or independent thinking.

### **1.6.6 F**

1. The report does not adequately explain their work.
2. There is very little discussion and major parts of the analysis are technically incorrect, or there are errors in the implementation.
3. The models used might be mentioned, but not any other details.
4. There is no effort to ensure reproducibility or robustness.
5. When applicable: Privacy and fairness are not mentioned.
6. There is no mention of methods beyond the source material or independent thinking.

## **1.7 Motivation**

Algorithms from Artificial Intelligence are becoming ever more complicated and are used in manifold ways in today's society: from prosaic applications like web advertising to scientific research. Their indiscriminate use creates many externalities that can be, however, precisely quantified and mitigated against.



The purpose of this course is to familiarise students with societal and scientific effects due to the use of artificial intelligence at scale. It will equip students with all the requisite knowledge to apply state-of-the-art machine learning tools to a problem, while recognising potential pit-falls. The focus of the course is not on explaining a large set of models. It uses three basic types of models for illustration: k nearest-neighbour, neural networks and probabilistic graphical models, with an emphasis on the latter for interpretability and the first for lab work. It is instead on the issues of reproducibility, data collection and experiment design, privacy, fairness and safety when applying machine learning algorithms. For that reason, we will cover technical topics not typically covered in an AI course: false discovery rates, differential privacy, fairness, causality and risk. Some familiarity with machine learning concepts and artificial intelligence is expected, but not necessary.

## **2 Schedule**

### **2.1 2019**

21 Aug	L1. Reproducibility, kNN	Christos
22 Aug	L2. Classification, Decision Problems, Project Overview	Christos
29 Aug	A1. Python, scikitlearn, classification, holdouts, overfitting	Dirk
29 Aug	A2. Bootstrapping, XV, project #1 introduction	Dirk
30 Aug	Mini-assignment	
4 Sep	L3. Decision Problems, Classification, Neural Networks, SGD	Christos
5 Sep	L4. Bayesian inference tutorial; neural networks	Christos
12 Sep	A3. Compare kNN/MLP, discover interesting features	Dirk
12 Sep	A4. Project Lab	Dirk
18 Sep	Project 1 1st Deadline	
18 Sep	L5. Databases, anonymity, privacy	Christos
19 Sep	L6. Differential privacy	Christos
26 Sep	A5. DB tutorial/distributed computing	Dirk
26 Sep	A6. Project DP tutorial: Laplace mechanism	Dirk
2 Oct	Project 1 2nd Deadline	
2 Oct	L7. Fairness and graphical models	Christos
3 Oct	L8. Estimating conditional independence	Christos
10 Oct	A7. Production ML: SageMaker/Pipelines	Dirk
10 Oct	A8. Project: fairness	Dirk
16 Oct	Project 1 Final Deadline	
16 Oct	L9. Recommendation systems [can be skipped?]	Christos
17 Oct	L10. Latent variables and importance sampling	Christos
24 Oct	A9. Restful APIs	Dirk
24 Oct	A10. An example latent variable model?	Dirk
30 Oct	L11. Causality	Christos
31 Oct	L12. Interventions and Counterfactuals	Christos
7 Nov	A11. Causality lab	Dirk
7 Oct	A12. Causality lab	Dirk
13 Nov	L13. Bandit problems	Christos
14 Nov	L14. Experiment design	Christos
20 Nov	A13. Experiment design lab	Dirk
21 Nov	A14. Experiment design lab	Dirk
2 Dec	Exam: 9AM Lessart Lesesal A Eilert Sundts hus, A-blokka	
11 Dec	Project 2 Deadline	

1. kNN, Reproducibility
2. Bayesian Inference, Decision Problems, Hypothesis Testing
3. Neural Networks, Stochastic Gradient Descent
4. Databases, k-anonymity, differential privacy

5. Fairness, Graphical models
6. Recommendation systems, latent variables, importance sampling
7. Causality, interventions, counterfactuals
8. Bandit problems and experiment design
9. Markov decision processes
10. Reinforcement learning

### 3 Lecture plan

#### 3.1 Module 1: Reproducibility

Machine learning as science: hypotheses, experiments and conclusions. kNN example: What is classification? What is clustering? Making sure you formalise the problem.

1. Reproducibility
2. KNN.
3. Bootstrapping

##### 3.1.1 kNN

**ACTIVITY**

##### 3.1.2 Reproducibility

**ACTIVITY**

##### 3.1.3 Modelling

**TUTORIAL**

1. Linear models
2. Neural networks
3. Confidence and  $p$ -values
4. Naive Bayes: Model mismatch
5.  $p$ -values, cross-validation and model mismatch

The purpose of this lecture is to familiarise students with all the decisions made from the beginning to the end of the data science process, and with the possible externalities when an algorithm is applied to real data.

## **3.2 Module 2: Decision problems**

1. Decision hierarchies
2. Bayesian inference
3. Optimisation and SGD.

### **3.2.1 Decision Problems**

### **3.2.2 Project Introduction: Credit risk for mortgages. TUTORIAL**

## **3.3 Week 3: Privacy**

1. Privacy in databases.
2. k-anonymity.
3. Differential Privacy.
4. The Random Response Mechanism.
5. Laplace Mechanism.
6. Exponential mechanism.

The purpose of this lecture is to introduce the students to basic database concepts, as well as to privacy problems that can occur when allowing access to a database to a third party.

## **3.4 Week 4: Fairness**

1. Graphical Models.
2. Fairness as independence.
3. Decision diagrams.
4. Fairness as smoothness.
5. Fairness as meritocracy.
6. Bayesian notions of fairness.

### **3.5 Project start: Experiment design for Medical Diagnostics [Aim: Reproducibility, Safety] PROJECT**

### **3.6 Week 5: Clustering**

Unstructured databases. Clustering / Anomaly detection.

The purpose of this lecture is to talk about non-matrix data, like graphs, and make a link to graphical models and simple problems like anomaly detection.

DNA testing and HMMs.

Here we talk more about unstructured data, in this case about DNA data.

### **3.7 Week 6: The web and recommendation systems**

Web data, ontologies, crawling. Knowledge representation.

This is web-structured data, which typically has some meta-information.

Matrix Factorisation / LDA: Recommendation systems I (user similarity)

This lecture introduces analysis of text data, and an application to recommendation systems.

#### **3.7.1 Online data collection. Optimal stopping (expensive labels) A/B Testing, Bandit Problems.**

This lecture introduces the concept of online data collection, rather than going through existing data. The applications considered are manual labelling via AMT or advertising.

#### **3.7.2 Markov decision processes and Dynamic Programming (active learning and experiment design more generally)**

The optimal data collection procedure can be formalised as an MDP, and this is explained here.

### **3.8 Optional. Safety: Risk-Sensitive Decision Making**

Sometimes we are risk averse... what do we mean by this, and what algorithms can we use? When we have developed an algorithm, how sure can we be that it works well in the real world?

## 4 Exam subjects

Here are some example questions for the exam. Answers can range from simple one-liners to relatively complex designs. Half of the points will come from 10 1-point questions and the remaining from 2 or 3 2-5-point questions.

### 4.1 Reproducibility

You are given a set of clinical data  $x_1, \dots, x_T$  with associated labels  $y_1, \dots, y_T$ , where  $y_t \in \{0, 1\}$  indicates whether a patient has a disease. Each point  $x_t$  is decomposable into  $n$  features  $x_{t,1}, \dots, x_{t,n}$ . Discuss how you can use a classification algorithm that estimates  $\hat{P}(y|x)$  from the data in order to discover predictive features, and how you can validate your findings in a reproducible manner.

#### 4.1.1 Possible answer

(Many approaches are possible, the main thing I want to see is that you can validate your findings)

From a statistical point of view, we want to see the strength of the dependence between an individual feature (or set of features) and the data. The strictest possible test is to see whether or not the labels are completely independent of a feature  $i$  given the remaining features, i.e. we want to check that

$$y_t \perp x_{t,i} \mid x_{t,-i} \quad x_{t,-i} \triangleq x_{t,1}, \dots, x_{t,i-1}, x_{t,i+1}, x_{t,n}$$

However this check is possibly too strict.

If this is the case, then  $P(y_t \mid x_t) = P(y_t \mid x_{t,-i})$ . One possible method is to fit the classification model of choice  $\mu = \hat{P}(y_t \mid x_t)$  and a sequence of models  $\mu_i = \hat{P}(y_t \mid x_{t,-i})$  on a subset  $D_1$  of the dataset. Consequently, we can measure the likelihood of models on the remaining data  $D_2$ , so that we obtain

$$\ell(\mu) = \prod_{t \in D_2} \hat{P}(y_t \mid x_t), \quad \ell(\mu_i) = \prod_{t \in D_2} \hat{P}(y_t \mid x_{t,-i}).$$

We may then consider all features  $i$  with  $\ell(\mu_i) < \ell(\mu)$  to be redundant. However, this may not be the case for two reasons:

1. If individually redundant features are correlated, then removing all of them may be difficult. For that reason, we may want to also test the performance of models which remove combinations of features.

2. Since probably no feature is completely useless, one reason for the apparent lack of predictive ability of some features maybe the amount of data we have. In the limit, if  $y_t \perp x_{t,i} \mid x_{t,-i}$  then our estimators will satisfy  $\hat{P}(y_t \mid x_t) = \hat{P}(y_t \mid x_{t,-i})$ . However, it is hard to verify this condition when the amount of data is little. Conversely, with a lot of data, even weakly dependent features will not satisfy independence.

## 4.2 Conditional probability and Bayesian inference

A prosecutor claims that the defendant is guilty because they have found DNA matching them on the scene of the crime. He claims that DNA testing has a false positive rate of one in a million ( $10^{-6}$ ). While this is indeed evidence for the prosecution, it does not mean that the probability that the defendant is innocent is  $10^{-6}$ . What other information would you need to calculate the probability of the defendant being guilty given the evidence, and how would you incorporate it?

### 4.2.1 Possible answer

Let us define the fact that the defendant committed a crime as  $C$  and the converse as  $\neg C$ . Let us also denote the event that a test is positive as  $T$ . Let us also define the case where the DNA being tested is the one being compared to as  $M$ . Then the information we have is

$$\mathbb{P}(T \mid \neg M) = 10^{-6} \quad (1)$$

$$T \text{ is true} \quad (2)$$

In order to predict whether somebody has actually committed the crime given the information, we must calculate  $\mathbb{P}(C \mid T)$ . This means we must calculate the following

$$\mathbb{P}(C \mid T) = \mathbb{P}(C \mid M) \mathbb{P}(M \mid T) + \mathbb{P}(C \mid \neg M) \mathbb{P}(\neg M \mid T) \quad (3)$$

$$= \mathbb{P}(C \mid M)[1 - \mathbb{P}(\neg M \mid T) + \mathbb{P}(C \mid \neg M) \mathbb{P}(\neg M \mid T)] \quad (4)$$

$$= \mathbb{P}(C \mid M)[1 - \mathbb{P}(T \mid \neg M) \mathbb{P}(\neg M) / \mathbb{P}(T) + \mathbb{P}(C \mid \neg M) \mathbb{P}(T \mid \neg M) \mathbb{P}(\neg M) / \mathbb{P}(T)], \quad \mathbb{P}(T) = \mathbb{P}(T \mid M) \mathbb{P}(M) + \mathbb{P}(T \mid \neg M) \mathbb{P}(\neg M) \quad (5)$$

As you can see, we are missing four important quantities.

- $\mathbb{P}(M)$ , the *a priori* probability that this is the defendant's DNA
- $\mathbb{P}(T \mid M)$  the probability of a test being positive if the DNA fragments come from the same person.

- $\mathbb{P}(C \mid M)$ , the probability that the defendant committed the crime if the DNA was really theirs.
- $\mathbb{P}(C \mid \neg M)$ , the probability that the defendant committed the crime if the DNA was not theirs.

So the false positive rate is far from sufficient evidence for a conviction and must be combined with other evidence.

### 4.3 Utility

If  $X$  is our set of rewards, our utility function is  $U : X \rightarrow \mathbb{R}$  and we prefer reward  $a$  to  $b$  (and write  $a >^* b$ ) iff  $U(a) > U(b)$ , then our preferences are transitive. Give an example of a preference relation  $>^*$  among objects so that transitivity can be violated, e.g when  $X = \mathbb{R}^2$ . In that case, we cannot create a utility function that will satisfy the same relation. Back your example with a thought experiment.

#### 4.3.1 Possible answer

A simple example is when  $U : \mathbb{R}^2 \rightarrow \mathbb{R}$ , with rewards having two attributes. Then we might prefer  $a$  to  $b$  if  $a_1 > b_1 + \epsilon$ , but if  $|a_1 - b_1| < \epsilon$  then we prefer  $a$  to  $b$  if  $a_2 > b_2$ . An example is if the first attribute is the IQ score of a job candidate and the second attribute their years of experience. We might prefer a brighter candidate as long as they are clearly much better (as IQ scores are fiddly), otherwise we will prefer the ones that have more experience. As an example, consider three candidates

Id	IQ	XP
a	120	5
b	130	4
c	140	3

In this example, we can set  $\epsilon = 15$  so we prefer a candidate if he has at least an IQ score 15 points higher than another. Due to this, we have  $a >^* c$ . However, as  $a$  and  $b$  have similar IQs we prefer  $a$  to  $b$ , i.e.  $b >^* a$  and similarly  $c >^* b$ . If transitivity held, then we'd have  $c >^* a$ , which we don't.

Note that if we mapped these to a utility function, i.e.  $U(a) = a_1 + a_2$ , we will always get a transitive relation.



## 4.4 Differential privacy

Consider a system where we obtain data  $x_1, \dots, x_n$  from individuals, where  $x_t \in X$  corresponds to data from a single individual. Consider a mechanism that, from this data, publishes an output  $a_1, \dots, a_n$  by partitioning  $X$  in two sets,  $A, B$  so that  $a_t = 1$  if  $x_t \in A$  and 0 otherwise. Is the mechanism  $\pi(a|x)$   $\epsilon$ -differentially private? If so, with what value of  $\epsilon$ ?

### 4.4.1 Possible answer

In general, DP algorithms must be stochastic, so that this algorithm cannot satisfy DP at all.

In more detail, differential privacy requires that  $\pi(a | x) \leq \pi(a | x')e^\epsilon$  for some  $\epsilon$  for any neighbouring  $x, x'$ . Consider a dataset where the  $t$ -th person has  $x_t \in A$ . Then  $a_t = 1$ . Consider a neighbouring dataset where  $x'_t \notin A$ . Then  $a_t = 0$  w.p. 1, so  $a_t = 1$  has probability 0.

$$\begin{aligned}\pi(a | x) &= \prod_i \pi(a_i | x_i) = \pi(a_t | x_t) \prod_{i \neq t} \pi(a_i | x_i) \\ \pi(a | x') &= \prod_i \pi(a_i | x_i) = [1 - \pi(a_t | x_t)] \prod_{i \neq t} \pi(a_i | x_i)\end{aligned}$$

Dividing the two, we get

$$\pi(a | x) = \pi(a | x')\pi(a_t | x_t)/[1 - \pi(a_t | x_t)].$$

However, the ratio on the right is not bounded (i.e. it can be  $\infty$ ), hence there is no DP.

## 4.5 Graphical models

A patient is coming to the doctor complaining of chest pains. The doctor recommends that the patient undergoes EEG examination in order to diagnose the patient's underlying condition and observes the result. Describe appropriate decision variables and random variables corresponding to this problem and draw a graphical model detailing their relationship.

### 4.5.1 Possible answer

Variables:

- C: Chest pain

- H: Underlying health condition
- P: Doctor policy
- X: examination decision
- Y: test result.

```
{H}->(C)
|      |
v      v
(Y)<-(X)<-[P]
```

[ ] indicates decision variables, ( ) observed random variables, { } latent variables

## 4.6 Conditional independence

Consider four random variables  $w, x, y, z$  with the following properties: (a)  $w$  is independent of  $x$  given  $y$  and  $z$ , (b) it is not completely independent of  $x$ . Draw a graphical model that satisfies them.

### 4.6.1 Possible answer

(a) means that there is no path from  $x$  to  $w$  given  $y, z$  (b) means that there is some path from  $x$  to  $w$ .

So a graphical model representing this is:

```
(z)--\
^      |
|      v
(x)   (w)
|      ^
v      |
(y)--/
```

## 4.7 Fairness

Consider a decision problem where a decision maker (DM) takes actions affecting a set of individuals. Let the DM's action be  $a \in A$ . This action results in an outcome  $y \in Y$ , also depending on the underlying characteristics  $x$  of the population and has conditional distribution  $P(y | x, a)$ . Assume that the DM has a utility function  $U : A \times Y \rightarrow \mathbb{R}$ .

1. Complete the following formula to show how the DM would maximise expected utility, assuming she observes  $x$ :

$$\max_a \mathbb{E}[U \mid a, x]$$

Note that  $\mathbb{E}[U \mid a, x] = \sum_y U(a, y)P(y, \mid x, a)$ .

1. Assume each individual  $i$  also receives some utility from the DM's actions. This is specified through a collection of utility functions  $v_i : A \times Y \rightarrow \mathbb{R}$ . Two typical definitions of fairness from social choice theory concentrate on maximising a *social welfare* function that depends on the utilities of the whole population. There are two typical such functions (a) The (expected) total utility of the population (b) The (expected) utility of the worst-off member of the population.

Formalise those definitions within our framework.

(a) Can be described as  $V = \sum_i v_i$ . Then the objective of the decision maker would be to find an  $a$  maximising

$$\mathbb{E} \left[ \sum_i v_i \mid a, x \right] = \sum_y P(y \mid a, x) \sum_i v_i(a, y)$$

(b) can be described as  $V = \min_i v_i$ . Similarly

$$\mathbb{E} \left[ \sum_i v_i \mid a, x \right] = \sum_y P(y \mid a, x) \min_i v_i(a, y)$$

1. Describe a method whereby the DM can trade-off maximising her own utility and social welfare. Under which conditions do the two objectives coincide?

A simple idea is to combine the social welfare linearly with the DM's utility. Then we can try to maximise

$$\mathbb{E}[(1 - \alpha)U + \alpha V \mid x, a].$$

The two objectives obviously coincide when  $U = V$ . However, any utility function  $U$  which has the same maximum as  $V$  is compatible with social welfare.

## 4.8 Causality

Patients arrive at a hospital and receive a treatment that depends on their symptoms. The first table shows how many people receive each treatment. Assume that the number of people with each symptom is representative of the population.

Applications	Symptom 1	Symptom 2
Treatment A	20	90
Treatment B	180	10

Table 1: Number of treatment applications

The second table describes the number of people that were cured after the treatment was applied.

Cured	Symptom 1	Symptom 2
Treatment A	15	60
Treatment B	90	4

Table 2: Effect of treatment

1. Draw a graphical model with the following four variables:

- $\pi$ : Treatment policy
- $x_t$ : Symptoms
- $a_t$ : Treatment
- $y_t$ : Treatment effect

1. What would the expected curing rate of a policy that uniformly randomly assigned treatments have been? (It is OK to provide a simple point estimate)
1. Given the above data, what would be the treatment policy  $\hat{\pi}^*$  that maximises the curing rate? and how much would the curing rate of  $\hat{\pi}^*$  be?
2. Is there some reason why the original policy  $\pi$  would be preferred to  $\hat{\pi}^*$ ?

#### 4.8.1 Possible answer

1. Note that typically the symptoms and treatment effect depend on an underlying medical condition, but the question did not ask about this.

$$\begin{array}{c}
 [\pi] \text{ ---> } (a_t) \\
 \quad \quad \quad \wedge \quad \backslash \\
 \quad \quad \quad | \quad \quad \backslash (y_t) \\
 \quad \quad \quad | \quad \quad / \\
 \quad \quad \quad | \quad \quad / \\
 \quad \quad \quad (x_t)
 \end{array}$$

1. For S1, Treatment A works  $15/20=3/4$  and B:  $90/180=1/2$ . Randomly assigning treatments:  $3/8+1/4 = (3+2)/8 = 5/8$

For S2, Treatment B works  $60/90=2/3$  and B:  $4/10=2/5$ . Randomly assigning treatments:  $1/3+1/5 = (3+5)/15 = 8/15$  S1 has 200 patients and S2 has 100 patients, so  $2/3$  of people have S1. So the overall treatment rate would have been  $5/8 * 2/3 + 8/15*1/3 = 10 / 24 + 8 / 45 \sim 5 / 12 + 2 / 11 \sim 7 / 12$

1. It appears that Treatment A always works best, i.e.  $3/4$  of the time and  $1/2$  for each symptom.

So the overall curing rate based on the data would be  $3/4 * 2/3 + 1/2*1/3 = 6/12 + 1/6 = 3/6+1/6 = 4/6=2/3$ .

1. Firstly, there could be hidden medical or financial costs. One treatment might be more expensive than the other, or may have more side-effects. In addition, one type of symptoms might be less acute or life-threatening than the other, thus requiring less aggressive treatment. Secondly, the new policy always uses the same treatment, and this means that we do not get information about the effectiveness of alternative treatments. This may be important in the initial stages of executing a treatment.

#### 4.9 Markov decision processes and experiment design

Consider a Markov decision process with two actions  $A = \{0,1\}$  and three states  $S = \{0,1,2\}$ , with a horizon  $T = 2$ , with starting state  $s_1 = 10$  and the following transition distribution:

$$P(s_t = 0 \mid s_t = 0, a_t = 0) = 1 \quad P(s_t = 1 \mid s_t = 0, a_t = 1) = 0.8 \\ P(s_t = 2 \mid s_t = 0, a_t = 1) = 0.2$$

We also receive a deterministic reward:

$$r_t = \begin{cases} 0 & s_t = 0 \\ 1 & s_t = 1 \\ -1 & s_t = 2 \end{cases}$$

Since  $T = 2$ , the MDP ends after we take the first action, observe  $S_2$  and obtain  $r_2$ . Our goal is to maximise

$$\mathbb{E} \sum_{t=1}^2 r_t.$$

What is the optimal policy for achieving that?

#### 4.9.1 Possible answer

We always start in state 1. Taking action 0, we end up in state 1 again, with reward 0. So  $\mathbb{E}[\sum_{t=1}^2 r_t \mid a_1 = 0] = 0 + 0$ .

Taking action 1, we end up in state 2 w.p 0.2 and state 1 w.p. 0.8. So  $\mathbb{E}[\sum_{t=1}^2 r_t \mid a_1 = 1] = 1 \times 0.8 - 1 \times 0.2 = 0.6$

So it is better to take action 1 in state 0.