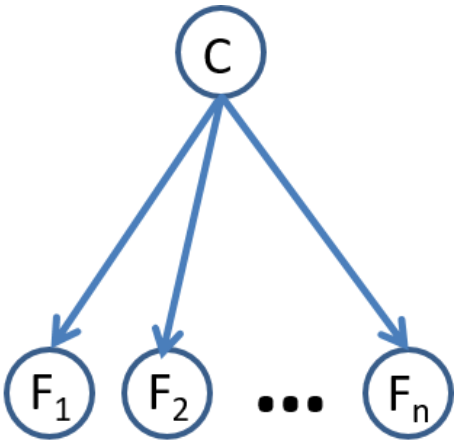
Started on	Tuesday, 23 January 2024, 9:25 AM
State	Finished
Completed on	Tuesday, 23 January 2024, 10:54 AM
Time taken	1 hour 29 mins
Grade	**44.00** out of 75.00 (**58.67**%)

Information

Distributions and BNs

Consider the following Bayesian network:



Information

Which of the following expressions (if any) correctly describe the full joint distribution represented by this network?

Question 1

Correct

Mark 1.00 out of 1.00

$P(C) \sum_i P(F_i \mid C)$

- True
- False ✔

The correct answer is 'False'.

Question 2

Correct

Mark 1.00 out of 1.00

$P(C)\prod_i P(F_i)$

- True
- False ✔

The correct answer is 'False'.

Question 3

Incorrect

Mark 0.00 out of 1.00

$P(F_1,\ldots,F_n,C)$

- True
- False ✖

The correct answer is 'True'.

Question 4

Correct

Mark 1.00 out of 1.00

$\prod_i P(F_i \mid F_{i-1})$

- True
- False ✔

The correct answer is 'False'.

Question 5

Correct

Mark 1.00 out of 1.00

$P(C)P(C \mid F_1,\ldots,F_n)$

- True
- False ✔

The correct answer is 'False'.

Question 6

Correct

Mark 1.00 out of 1.00

$P(C)\prod_i P(F_i \mid F_{i-1})$

- True
- False ✔

The correct answer is 'False'.

Information

From now on, suppose all the variables in the above model are Boolean. Which of the following statements are correct?

Question 7

Correct

Mark 1.00 out of 1.00

If we reversed the direction of all arrows, the full joint distribution represented by the network would remain the same, if we keep the distribution tables unchanged

- True
- False ✔

The correct answer is 'False'.

Question 8

Correct

Mark 1.00 out of 1.00

If we reversed the direction of all arrows, the number of non-redundant parameters would grow to a number of \(n \cdot 2n \)

- True
- False ✔

The correct answer is 'False'.

Question 9

Correct

Mark 1.00 out of 1.00

If we remove edge $C \rightarrow F_1$, C and F_2 become conditionally independent, given F_1

- True
- False ✔

The correct answer is 'False'.

Question 10

Incorrect

Mark 0.00 out of 1.00

Adding an edge $F_1 \rightarrow F_2$ would lead to an illegal (cyclic) graph structure

- True ✖
- False

The correct answer is 'False'.

Question 11

Incorrect

Mark 0.00 out of 1.00

If we remove the edge $C \rightarrow F_1$, then $(C, F_2,...,F_n \sim \bot \sim F_1)$

- True
- False ✖

The correct answer is 'True'.

Question 12

Incorrect

Mark 0.00 out of 1.00

If we reversed the direction of all arrows, the number of non-redundant parameters would grow to a number of 2^n+n

- True
- False ✖

The correct answer is 'True'.

Question 13

Correct

Mark 1.00 out of 1.00

If we removed the edge $C \rightarrow F_1$, the number of non-redundant parameters in the network would be halved

- True
- False ✔

The correct answer is 'False'.

Question 14

Incorrect

Mark 0.00 out of 1.00

Adding an edge $F_1 \rightarrow F_2$ would double the number of non-redundant parameters in F_2's probability table

- True
- False ✖

The correct answer is 'True'.

Information

Now consider having to learn the parameters of this network from a set of training examples $\mathcal{D}=\{\mathbf{x}_1,...,\mathbf{x}_M\}$, in a maximum-likelihood (ML) way. Which of the following statements are true?

Question 15

Incorrect

Mark 0.00 out of 1.00

Each training example is a vector with n components

- True ✖
- False

The correct answer is 'False'.

Question 16

Correct

Mark 1.00 out of 1.00

When estimating $P(F_i \mid C)$ for some variable F_i, the distributions of the other variables $F_j, j \neq i$ in $\mathcal{D}$ are irrelevant

- True ✔
- False

The correct answer is 'True'.

Question 17

Incorrect

Mark 0.00 out of 1.00

If all the training examples in $\mathcal{D}$ were identical and *only* contained false values for all variables, all parameters estimated for the model will be zero

- True ✖
- False

The correct answer is 'False'.

Question 18

Correct

Mark 1.00 out of 1.00

The estimate $\hat{P}(C=\mathit{true})$ for the parentless variable C will always be 1

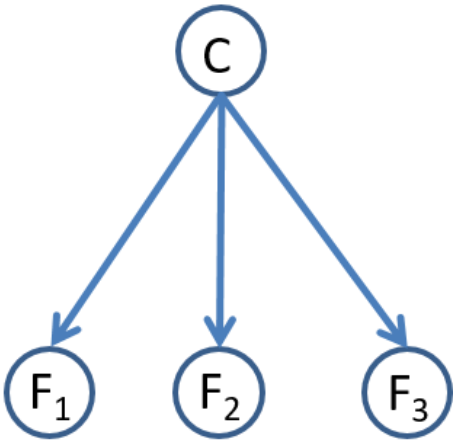
- True
- False ✔

The correct answer is 'False'.

Information

Inference BNs

Consider the following simple Bayesian network model:



with all variables being binary with two values t and f, and the following (Conditional) Probability Tables (where the three variables $F_1 \ldots F_3$ have identical tables):

$P(C)$:	t	f
	0.5	0.5

$P(F_i \mid C)$:	t	f
t	0.5	0.5
f	0.0	1.0

Information

Calculate the following probabilities (where c stands for $C = t$, $\neg f_i$ for $F_i = f$ etc). **Enter all decimal places**.

Question 19

Correct

Mark 1.00 out of 1.00

$P(\neg c, \neg f_1, \neg f_2, \neg f_3)$

Answer: 0.5 ✔

The correct answer is: 0.5

Question 20

Incorrect

Mark 0.00 out of 1.00

$P(\neg c, f_1, f_2, f_3)$

Answer: 0.0625 ✖

The correct answer is: 0

Question 21

Incorrect

Mark 0.00 out of 1.00

$P(\neg f_1, \neg f_2, \neg f_3)$

Answer: 0.125 ✖

The correct answer is: 0.5625

Question 22

Correct

Mark 1.00 out of 1.00

$P(c, \neg f_1, \neg f_2, \neg f_3)$

Answer: 0.0625 ✔

The correct answer is: 0.0625

Question 23

Correct

Mark 1.00 out of 1.00

How many topological orderings are there for this graph structure?

- 1
- 2
- 3
- 6 ✔
- 7
- 12

The correct answer is:
6

Information

Suppose we have drawn the following set of samples from the Bayesian network above:

C	F1	F2	F3
t	t	f	f
t	t	f	f
t	t	t	f
t	t	t	t
t	t	t	t
f	f	f	f
f	f	f	f
t	f	f	f
t	f	f	f
t	f	f	f

What estimates will be get by Rejection Sampling for the following queries:

Question 24

Incorrect

Mark 0.00 out of 1.00

$P(c \mid \neg f_1, \neg f_2, \neg f_3)$: 0.33 ✖

$P(\neg c \mid \neg f_1, \neg f_2, \neg f_3)$: 0.66 ✖

Question 25

Incorrect

Mark 0.00 out of 1.00

$P(c \mid \neg f_1)$: 0.66 ✗

$P(\neg c \mid \neg f_1)$: 0.33 ✗

Question 26

Correct

Mark 1.00 out of 1.00

$P(c \mid f_1)$: 1 ✓

$P(\neg c \mid f_1)$: 0 ✓

Information

Consider the query $P(C \mid f_1)$. What is the correct weight for the following samples, when we do likelihood weighting?

Question 27

Incorrect

Mark 0.00 out of 1.00

$\neg c, f_1, \neg f_2, \neg f_3$

Answer: 0.5 ✗

The correct answer is: 0

Question 28

Correct

Mark 1.00 out of 1.00

c, f_1, f_2, f_3

Answer: 0.5 ✓

The correct answer is: 0.5

Question 29

Correct

Mark 1.00 out of 1.00

$c, f_1, f_2, \neg f_3$

Answer: 0.5 ✔

The correct answer is: 0.5

Information

Consider the query $P(C \mid f_1)$. Which of the following (if any) is a correct Gibbs Resampling Distribution for variable C?

Question 30

Incorrect

Mark 0.00 out of 1.00

$1/Z \times P(C \mid F_1, F_2, F_3)$

- True
- False ✖

The correct answer is 'True'.

Question 31

Correct

Mark 1.00 out of 1.00

$1/Z \times P(C)P(F_1 \mid C)P(F_2 \mid C)P(F_3 \mid C)$

- True ✔
- False

The correct answer is 'True'.

Question 32

Correct

Mark 1.00 out of 1.00

$1/Z \times P(C)P(F_1 \mid C)$

- True
- False ✔

The correct answer is 'False'.

Question 33

Correct

Mark 1.00 out of 1.00

$1/Z \times P(F_1 \mid C)$

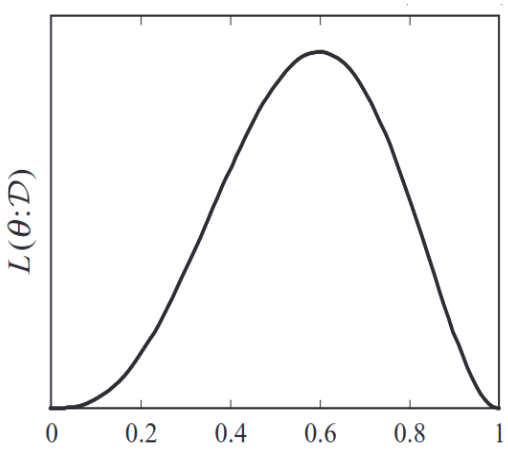
- True
- False ✔

The correct answer is 'False'.

Information

Parameter Learning

Consider the example, from the lecture slides, of learning the parameter $\theta = P(heads)$ for our thumbtack, where we had the following likelihood function $L(\theta, \mathcal{D})$ based on a set $\mathcal{D}$ of $N = 5$ observations, of which $k = 3$ were heads, and $(N - k) = 2$ were tails:



Information

Which of the following statements are true (about **this specific likelihood function**, relating to **this specific dataset**):

Question 34

Incorrect

Mark 0.00 out of 1.00

$L(0 : \mathcal{D}) = 0$

- True
- False ✖

The correct answer is 'True'.

Question 35

Correct

Mark 1.00 out of 1.00

$L(0.4 : \mathcal{D}) = L(0.6 : \mathcal{D}) - 2$

- True
- False ✔

The correct answer is 'False'.

Question 36

Correct

Mark 1.00 out of 1.00

$L(0.6 : \mathcal{D}) = 3/5$

- True
- False ✔

The correct answer is 'False'.

Question 37

Correct

Mark 1.00 out of 1.00

$L(0.6 : \mathcal{D}) = P(\mathcal{D} \mid \theta = 0.6)$

- True ✔
- False

The correct answer is 'True'.

Information

Which of the following statements are true **in general** (i.e., about any possible likelihood function over one parameter θ, and any dataset):

Question 38

Correct

Mark 1.00 out of 1.00

The likelihood for any θ cannot be > 1.0

- True
- False ✔

The correct answer is 'False'.

Question 39

Incorrect

Mark 0.00 out of 1.00

The likelihood for $\theta = 0$ and $\theta = 1$ is always 0

- True ✖
- False

The correct answer is 'False'.

Question 40

Correct

Mark 1.00 out of 1.00

The likelihood function is a probability distribution over θ

- True
- False ✔

The correct answer is 'False'.

Question 41

Correct

Mark 1.00 out of 1.00

A likelihood function can only have one maximum

- True
- False ✔

The correct answer is 'False'.

Information

Assume we want to do Bayesian parameter estimation for some parameter θ, and assume we use a uniform distribution over θ as the prior, which gives the same probability to all values of θ. Which of the following statements are correct:

Question 42

Correct

Mark 1.00 out of 1.00

$P(\theta \mid \mathcal{D}) = P(\mathcal{D} \mid \theta)$

- True
- False ✔

The correct answer is 'False'.

Question 43

Correct

Mark 1.00 out of 1.00

The posterior will be proportional to the likelihood

- True ✔
- False

The correct answer is 'True'.

Question 44

Correct

Mark 1.00 out of 1.00

The posterior mean estimate can be different from the MAP estimate

- True ✔
- False

The correct answer is 'True'.

Information

Model Fitting and Structure Learning

Consider a world with three Boolean variables A, B, C with full joint distribution $P(A, B, C)$, and four potential Bayesian network graph structures over these variables:

$$
\begin{array}{lccc}
(1) & A & B & C \\
(2) & A \longrightarrow & B & \longrightarrow C \\
(3) & A & B & \longleftarrow C \\
(4) & A \longrightarrow & B & \longleftarrow C
\end{array}
$$

Also, suppose we have a training set $\mathcal{D}$ of 100 examples sampled i.i.d. from the above joint distribution. Which of the following statements are true (always assuming that a network graph is parametrised with its ML parameters on $\mathcal{D}$)?

Question 45

Incorrect

Mark 0.00 out of 1.00

The description length of a model depends not on the number of parent relations, but on the size of its parameters (probabilities in the CPDs)

- True ✖
- False

The correct answer is 'False'.

Question 46

Correct

Mark 1.00 out of 1.00

If we replaced the model dimension term in the BIC score with any constant value > 0, a learning algorithm using this score would always learn the model (1), regardless of the data in $\mathcal{D}$

- True
- False ✔

The correct answer is 'False'.

Question 47

Incorrect

Mark 0.00 out of 1.00

The BIC score will always prefer structure (1) to model (4), because it requires strictly fewer parameters

- True ✖
- False

The correct answer is 'False'.

Question 48

Correct

Mark 1.00 out of 1.00

Models (3) and (4) will always have the same likelihood on $\mathcal{D}$

- True
- False ✔

The correct answer is 'False'.

Question 49

Incorrect

Mark 0.00 out of 1.00

Model (1) can never have a higher likelihood than Model (2) on $\mathcal{D}$.

- True
- False ✖

The correct answer is 'True'.

Question 50

Incorrect

Mark 0.00 out of 1.00

Model structure (2) is preferable compared to (4) because it models in the direction of causality.

- True ✖
- False

The correct answer is 'False'.

Question 51

Correct

Mark 1.00 out of 1.00

If all four models had the same likelihood relative to $\mathcal{D}$, the BIC score would prefer model

- 2
- 1 ✔
- 4
- 3

The correct answer is:
1

Information

Temporal Models / HMMs

Consider a HMM defined over a discrete state variable $S = \{s_1, \ldots, s_N\}$ and a discrete observation variable $O = \{o_1, \ldots, o_M\}$. Let matrix $\mathbf{A}$ be the state transition matrix, matrix $\mathbf{B}$ the observation model, and Π the initial state distribution model, as defined in the lecture slides.

Information

Give the following quantities:

Question 52

Incorrect

Mark 0.00 out of 1.00

Give the dimension of $\mathbf{B}$

rows: N ✗ , columns: M ✗

Question 53

Incorrect

Mark 0.00 out of 1.00

Give the dimension of $P(o_2 \mid S)$

numbers: M ✗

Question 54

Correct

Mark 1.00 out of 1.00

Give the dimension of $P(S \mid o_1)$

numbers: N ✓

Question 55

Correct

Mark 1.00 out of 1.00

Give the dimension of $\mathbf{A}$

rows: N ✓ , columns: N ✓

Question 56

Correct

Mark 1.00 out of 1.00

Give the dimension of $P(O \mid s_N)$

numbers: M ✔

Question 57

Correct

Mark 1.00 out of 1.00

Give the dimension of Π

rows: N ✔ , columns: 1 ✔

Information

Assume the above HMM is unrolled to a length $T = 2$ (call the resulting network HMM2). Which of the following are valid topological orderings w.r.t. HMM2?

Question 58

Incorrect

Mark 0.00 out of 1.00

$S^{(0)}, S^{(1)}, O^{(1)}, S^{(2)}, O^{(2)}$

- True
- False ✖

The correct answer is 'True'.

Question 59

Incorrect

Mark 0.00 out of 1.00

$\mathbf{S}^{(0:2)}, \mathbf{O}^{(1:2)}$

- True
- False ✖

The correct answer is 'True'.

Question 60

Incorrect

Mark 0.00 out of 1.00

$S^{(0)}, O^{(1)}, S^{(1)}, O^{(2)}, S^{(2)}$

- True ✖
- False

The correct answer is 'False'.

Information

Which of the following statements are correct for any HMM?

Question 61

Correct

Mark 1.00 out of 1.00

Given an observation sequence $\mathbf{O} = \mathbf{o}^{(1:t)}$, $P(S^{(t+1)} \mid \mathbf{O})$ can be predicted by running the Filtering Algorithm, followed by the Prediction Algorithm.

- True ✔
- False

The correct answer is 'True'.

Question 62

Correct

Mark 1.00 out of 1.00

Given an observation sequence $\mathbf{O} = \mathbf{o}^{(1:t)}$, one can also ignore the observations completely and still make a prediction for $P(S^{(t)})$

- True ✔
- False

The correct answer is 'True'.

Question 63

Correct

Mark 1.00 out of 1.00

The message p in the prediction algorithm always sums to 1

- True ✔
- False

The correct answer is 'True'.

Question 64

Incorrect

Mark 0.00 out of 1.00

Smoothing could also be done by unrolling the HMM to the observation sequence length T and then calculating all distributions $P(S^{(t)})$ via Inference by Enumeration

- True
- False ✖

The correct answer is 'True'.

Information

Linear Gaussian Models

Consider the Kalman Filter model of a bicycle riding along a one-dimensional road, as introduced in the lecture slides, with random variables P (position), V (velocity), Z (GPS values).

Information

Which of the following distributions (if any) does the Kalman Filtering Algorithm calculate?

Question 65

Incorrect

Mark 0.00 out of 1.00

$P(P^{(t)}, V^{(t)} \mid \mathbf{z}^{(1:t)})$

- True
- False ✖

The correct answer is 'True'.

Question 66

Correct

Mark 1.00 out of 1.00

$P(P^{(t)}, V^{(t)} \mid z^{(t)}, p^{(t-1)}, v^{(t-1)})$

- True
- False ✔

The correct answer is 'False'.

Question 67

Incorrect

Mark 0.00 out of 1.00

$P(P^{(t)}, V^{(t)} \mid P^{(t-1)}, V^{(t-1)})$

- True ✖
- False

The correct answer is 'False'.

Question 68

Correct

Mark 1.00 out of 1.00

$P(P^{(t)}, V^{(t)} \mid \mathbf{z}^{(1:t)}, p^{(t-1)}, v^{(t-1)})$

- True
- False ✔

The correct answer is 'False'.

Question 69

Incorrect

Mark 0.00 out of 1.00

$P(P^{(t)}, V^{(t)} \mid z^{(t)})$

- True ✖
- False

The correct answer is 'False'.

Information

Which of the following assumptions about the bicycle's movement are encoded in the specific transition matrix $\mathbf{A}$ and observation model $\mathbf{B}$ that we specified in the lecture slides:

Question 70

Incorrect

Mark 0.00 out of 1.00

that GPS reading $Z^{(t)}$ and velocity $V^{(t)}$ at the same time point are independent

- True ✖
- False

The correct answer is 'False'.

Question 71

Correct

Mark 1.00 out of 1.00

that the probability of slowing down by a certain amount is independent of the current velocity

- True ✔
- False

The correct answer is 'True'.

Question 72

Correct

Mark 1.00 out of 1.00

that the GPS sensor has the same inaccuracy everywhere

- True ✔
- False

The correct answer is 'True'.

Question 73

Incorrect

Mark 0.00 out of 1.00

that the distribution over possible GPS values at time t depends only on $p^{(t)}$

- True
- False ✖

The correct answer is 'True'.

Question 74

Incorrect

Mark 0.00 out of 1.00

that the bicycle's velocity will tend to stay the same from moment to moment

- True
- False ✖

The correct answer is 'True'.

Question 75

Correct

Mark 1.00 out of 1.00

that the GPS sensor readings give us no information about velocity \(V \)

- True
- False ✔

The correct answer is 'False'.