INFO-F422 - Statistical foundations of machine learning - 202122

Commencé le	dimanche 5 juin 2022, 11:23
État	Terminé
Terminé le	dimanche 5 juin 2022, 12:00
Temps mis	36 min 18 s
Note	**6,83** sur 20,00 (**34**%)

Navigation du test

1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16

Afficher une page à la fois
Terminer la relecture

Question **1**
Non répondue
Noté sur 1,00
Marquer la question

Let $\mathbf{x}$ and $\mathbf{y}$ two discrete dependent r.v.s.
The domain of $\mathbf{x}$ is $\{-1,0,1\}$ and its probability distribution is uniform (i.e. 1/3 for all values). Let $\mathbf{y}=\mathbf{x}^2$.

Compute first

$E[\mathbf{xy}] =$ [ ]

Then show that the two variables are uncorrelated yet dependent

While x takes three values, y may take only two values: 0 and 1.
The variable z=xy may then take three values: -1, 0, 1
Let us see the probability of the three possible values

P(z=-1)= P(x=-1,y=1)=P(y=1 | x=-1) P(x=-1)=1*1/3
P(z=0)= P(x=1,y=0)+P(x=-1,y=0)+P(x=0,y=0)=0+0+1/3
P(z=1)=P(x=1,y=1)=1/3

E[z]=-1/3+0+1/3=0

Question **2**
Terminé
Note de 1,00 sur 1,00
Marquer la question

If two events are independent then they have the same probability.

- a. YES
- b. NO

No. The property of independence concerns only the equality P(A,B)=P(A)*P(B)

La réponse correcte est :
NO

Question **3**
Terminé
Note de 1,00 sur 1,00
Marquer la question

Let $\mathbf{y}$ and $\mathbf{x}$ have the same distribution. Is the variance of $\mathbf{x}-\mathbf{y}$ necessarily zero?

- a. YES
- b. NO

No. Var{x-y}=Var{x}+Var{y}-2Cov{x,y}. So all depends on the nature of x and y

La réponse correcte est :
NO

Question **4**
Terminé
Note de 1,00 sur 1,00
Marquer la question

Let $\mathbf{y}$ and $\mathbf{x}$ be two binary random variables. If $\mathbf{x}$ and $\mathbf{y}$ are dependent , if $P(\mathbf{y}=1|\mathbf{x}=1)>0.5$ is also necessarily $P(\mathbf{y}=1)>0.5$?

- a. YES
- b. NO

No. This is not required by the notion of dependency.

La réponse correcte est :
NO

Question **5**
Terminé
Note de 1,00 sur 1,00
Marquer la question

Let $\mathbf{y}$ and $\mathbf{x}$ be two binary random variables. If $\mathbf{x}$ and $\mathbf{y}$ are dependent $P(\mathbf{y}=1|\mathbf{x}=1)$ is always different $P(\mathbf{y}=1|\mathbf{x}=0)$

- a. YES
- b. NO

No. The dependency property does not impose such requirement.

La réponse correcte est :
NO

Question **6**
Terminé
Note de 0,00 sur 1,00
Marquer la question

Are two conditionally independent random variables necessarily dependent?

- a. YES
- b. NO

No. Conditional independence does not imply dependence.

La réponse correcte est :
NO

Question **7**
Terminé
Note de 1,00 sur 1,00
Marquer la question

Let $\mathbf{x}$ and $\mathbf{y}$ two discrete independent r.v.

such that

$P_{\mathbf{x}}(-1)=0.1, P_{\mathbf{x}}(0)=0.8, P_{\mathbf{x}}(1)=0.1$

and

$P_{\mathbf{y}}(1)=0.1, P_{\mathbf{y}}(2)=0.8, P_{\mathbf{y}}(3)=0.1$

Compute

$E[\mathbf{x}] =$ 0

$E[\mathbf{y}] =$ 2

If $\mathbf{z}=\mathbf{x}+\mathbf{y}$ compute numerically

$E[\mathbf{z}] =$ 2 and verify

that $E[\mathbf{z}]=E[\mathbf{x}]+E[\mathbf{y}]$

E[x]=-1*0.1+0.1=0
E[y]=0.1+1.6+0.3=2

The domain of z is: {0, 1, 2, 3, 4} and the probability function takes the values P={0.01, 0.16,0.66,0.16,0.01}

E[z]=0.16+0.66*2+0.16*3+0.01*4=2

Question **8**
Terminé
Note de 1,00 sur 1,00
Marquer la question

Let $\mathbf{y}$ and $\mathbf{x}$ be two binary random variables. If $\mathbf{x}$ and $\mathbf{y}$ are dependent , is $P(\mathbf{y}=1|\mathbf{x}=1)$ always different from $P(\mathbf{y}=1)$?

- a. YES
- b. NO

Not necessarily. It could be that $P(\mathbf{y}=1|\mathbf{x}=1)=P(\mathbf{y}=1)$ but $P(\mathbf{y}=1|\mathbf{x}=0)\neq P(\mathbf{y}=1)$

La réponse correcte est :
NO

Question **9**
Terminé
Note de 0,00 sur 1,00
Marquer la question

Are two independent random variables always conditionally independent?

- a. YES
- b. NO

No. Independence does not imply conditional independence.

La réponse correcte est :
NO

Question **10**
Terminé
Note de 0,00 sur 1,00
Marquer la question

Consider four binary random variables $\mathbf{x}_1,\mathbf{x}_2,\mathbf{x}_3,\mathbf{y}$ such that the variables $\mathbf{x}_i$ are conditionally independent given $\mathbf{y}$.

How many parameters are needed to specify its distribution? (Use the chain rule)

- a. 6
- b. 5
- c. 12
- d. 3
- e. 15
- f. 7
- g. 8
- h. 10
- i. 11
- j. Infinity

P(x1,x2,x3,y)=P(x1|x2,x3,y)P(x2|x3,y)P(x3|y) P(y)= P(x1|y)P(x2|y) P(x3|y) P(y)

7= 2+2+2+1

La réponse correcte est :
7

Question **11**
Terminé
Note de 0,50 sur 1,00
Marquer la question

If the event $\mathcal{E}_1 \subset \mathcal{E}_2$ and $P(\mathcal{E}_1)>0$, $P(\mathcal{E}_2)=0.5$ then $P(\mathcal{E}_1|\mathcal{E}_2)$ is

- a. 1
- b. <1
- c. 0
- d. $\geq P(\mathcal{E}_1)$

$> P(\mathcal{E}_1)$ and <1 since the intersection of $\mathcal{E}_1$ and $\mathcal{E}_2$ is $\mathcal{E}_1$

Les réponses correctes sont :
<1,
\(\ge P({\mathcal E}_1)\)

Question **12**
Non répondue
Noté sur 1,00
Marquer la question

Consider four binary random variables $\mathbf{x},\mathbf{y},\mathbf{z},\mathbf{w}$ such that

- $\mathbf{y}$ is conditionally independent of $\mathbf{z}$ given $\mathbf{w}$
- $\mathbf{z}$ is independent of $\mathbf{w}$

How many parameters are needed to specify its distribution ? (Use the chain rule)

- a. 6
- b. 5
- c. 12
- d. 3
- e. 15
- f. 7
- g. 8
- h. 10
- i. 11
- j. Infinity

P(x,y,z,w)=P(x|y,z,w)P(y|z,w)P(z|w) P(w)= P(x|y,z,w)P(y|w)P(z) P(w)

12= 8+2+1+1

La réponse correcte est :
12

Question **13**
Non répondue
Noté sur 2,00
Marquer la question

An important probabiity result is the Jensen's inequality
$f(E[\mathbf{x}]) \geq E[f(\mathbf{x})]$ if the function f is concave.
Suppose that $\mathbf{x} \sim Nor(1,\sigma^2=1)$ and $f(x)=-x^2$
The student should compute by using $R=10000$ trials of a Monte Carlo simulation:

$f(E[\mathbf{x}]) =$ [ ]

$E[f(\mathbf{x})] =$ [ ]

and verify that the inequality is satisfied.

```
set.seed(0)
R=10000
# number of MC trials
f<-function(x){
 -x^2
}

FX=NULL
X=NULL

a=1
b=1

for ( r in 1:R){

 x=rnorm(1,a,b)
 fx=f(x)

 FX=c(FX,fx)
 X=c(X,x)

}
muX=mean(X)
cat("E[f(x)]=",mean(FX))
cat("\n f(E[x])=",f(muX))
```

Question **14**
Non répondue
Noté sur 2,00
Marquer la question

Let us consider two discrete r.v.s.
The variable $\mathbf{x} \in \mathcal{X}=\{NEG,POS\}$ and
$\mathbf{y} \in \mathcal{Y}=\{GOOD,NEUTRAL,BAD\}$
Let

- $P(\mathbf{x}=NEG)=P(\mathbf{x}=POS)$
- $P(\mathbf{y}=GOOD|\mathbf{x}=NEG))=P(\mathbf{y}=NEUTRAL|\mathbf{x}=NEG))=0.2$
- $P(\mathbf{y}=GOOD|\mathbf{x}=POS))=P(\mathbf{y}=BAD|\mathbf{x}=POS))=0.2$

Compute the following entropy terms (using the natural logarithm)

$H[\mathbf{x}] =$ [ ]

$H[\mathbf{y}] =$ [ ]

$H[\mathbf{y}|\mathbf{x}=POS] ==$ [ ]

By using the law of total probability
P(y=GOOD)=P(y=GOOD| x=NEG) P(x=NEG)+ P(y=GOOD| x=POS) P(x=POS)=0.2*0.5+0.2*0.5=0.2
P(y=NEUTRAL)=P(y=NEUTRAL| x=NEG) P(x=NEG)+ P(y=NEUTRAL| x=POS) P(x=POS)=0.2*0.5+0.6*0.5=0.4
P(y=BAD)=P(y=BAD| x=NEG) P(x=NEG)+ P(y=BAD| x=POS) P(x=POS)=0.6*0.5+0.2*0.5=0.4
H[y]= 1.05

P(y=GOOD| x=POS)=0.2
P(y=NEUTRAL| x=POS)=0.6
P(y=BAD| x=POS)=0.2
H[y| x=POS]= 1.05

Question **15**
Non répondue
Noté sur 2,00
Marquer la question

Let us consider three binary random variables $\mathbf{x}_1,\mathbf{x}_2,\mathbf{x}_3$ and their joint distribution
How many parameters have to be set to define their joint distribution if

1. we have no additional information [ ]
2. the three variables are independent [ ]
3. $\mathbf{x}_1$ is conditionally independent of $\mathbf{x}_3$ given $\mathbf{x}_2$: [ ] Hint: use the chain rule
4. $\mathbf{x}_1$ is independent of the other two variables [ ] Hint: use the chain rule

1. 8-1=7 parameters have to be defined since the sample space is made of 8 experimental outcomes and the probabilistic constraint applies
2. 1*3=3 since P(x1,x2,x3)=P(x1)P(x2)P(x3) and for each of the three terms one parameter is enough
3. 2+2+1=5 By using the chain rule and the conditional independence we obtain P(x1,x2,x3)=P(x1| x2,x3)P(x2| x3) P(x3) = P(x1| x2)P(x2| x3) P(x3)
4. 1+2+1=4 By using the chain rule and the conditional independence we obtain P(x1,x2,x3)=P(x1| x2,x3)P(x2| x3) P(x3) = P(x1)P(x2| x3) P(x3)

Question **16**
Terminé
Note de 0,33 sur 2,00
Marquer la question

Consider the data frame **Q1.G2.D** containing the joint probability distribution of the 4 categorical binary random variables $\mathbf{x}_1$, $\mathbf{x}_2$, $\mathbf{x}_3$ and $\mathbf{y}$ each taking two possible values: *yes* and *no*.
The student should compute:

P($\mathbf{y}$=no | $\mathbf{x}_1$=yes, $\mathbf{x}_2$=no, $\mathbf{x}_3$=no) = 0.6923077

P($\mathbf{x}_3$=no | $\mathbf{x}_1$=yes, $\mathbf{y}$=no) = [ ]

P($\mathbf{x}_2$=no | $\mathbf{x}_1$=yes, $\mathbf{y}$=no) = [ ]

P($\mathbf{x}_1$=no | $\mathbf{x}_2$=no, $\mathbf{x}_3$=no) = [ ]

Are the two variables **y** and $\mathbf{x}_1$ independent?

- Yes
- No

La réponse correcte est : No

Explain why Prob {E1 ∩ E2} = Prob {E1, E2} = Prob {E1} Prob {E2}

Are the two variables $\mathbf{x}_3$ and $\mathbf{x}_1$ conditionally independent given **y**=no?

- Yes
- No

La réponse correcte est : Yes

Explain why [ ]

Terminer la relecture

« ACTIVITÉ PRÉCÉDENTE Probability cheat sheet

ACTIVITÉ SUIVANTE Video 2/3/21 (1st part) »

Aller à...