Sample Problems
Lecture 1: Probability Concepts

Problems prepared by: F. Jalayer and H. Ebrahimian

1. The earthquake early warning

Suppose that we have installed a real-time earthquake warning system which issues an alarm when a significant earthquake takes place. The probability that a significant earthquake takes place is around 2×10^{-3}. However, our earthquake real-time warning system is subject to a 15% error called the *false alarm* error: That is, the probability that there is an alarm issued given that it was not followed by a significant earthquake is 0.15 and the probability that no alarm is issued given that it is followed by a significant earthquake is 0.15.

- **What is the probability that no significant earthquake takes place?**

  $EQ \equiv$ a significant earthquake takes place

  $$\Rightarrow P\left(\overline{EQ}\middle|\mathbf{I}\right)=1-P\left(EQ\middle|\mathbf{I}\right)=1-2\times10^{-3}=998\times10^{-3}$$

- **What is the probability that a real-time alarm is issued given that an earthquake takes place?**

  $A \equiv$ an alarm is issued by the system

  $$\Rightarrow P\left(A\middle|EQ,\mathbf{I}\right)=1-P\left(\overline{A}\middle|EQ,\mathbf{I}\right)=1-0.15=0.85$$

- **What is the probability that our system issues an alarm and that an earthquake takes place?**

  $$P\left(A\cdot EQ\middle|\mathbf{I}\right)=P\left(EQ\middle|\mathbf{I}\right)P\left(A\middle|EQ,\mathbf{I}\right)=(2\times10^{-3})\times0.85=1.7\times10^{-3}$$

- **What is the probability that our system issues an alarm and that no earthquake takes place?**

$$P\left(A\cdot\overline{EQ}\middle|\mathbf{I}\right)=P\left(\overline{EQ}\middle|\mathbf{I}\right)P\left(A\middle|\overline{EQ},\mathbf{I}\right)=(998\times10^{-3})\times0.15\cong0.15$$

- **What is the probability that our system issues an alarm? (hint: consider using the total probability theorem)**

$$P\left(A\middle|\mathbf{I}\right)=P\left(A\cdot(EQ+\overline{EQ})\middle|\mathbf{I}\right)=P\left(A\cdot EQ\middle|\mathbf{I}\right)+P\left(A\cdot\overline{EQ}\middle|\mathbf{I}\right)=1.7\times10^{-3}+0.15=0.1517$$

- **Suppose that our device has issued a warning. What is the probability that a significant earthquake takes place?**

$$P\left(EQ\middle|A,\mathbf{I}\right)=\frac{P\left(A\cdot EQ\middle|\mathbf{I}\right)}{P\left(A\middle|\mathbf{I}\right)}=\frac{1.7\times10^{-3}}{0.1517}\cong0.01$$

2. Identification of Flood Prone Area

An approximate way for identifying the flood-prone (FP) areas is to use a method called the topographic wetness index (TWI). This method identifies the FP areas as those areas with (*TWI*>*t*) where *t* is a given threshold. A more accurate way of calculating the flood prone areas consists of calculating the inundation profile. In this method, the inundated zones (IN) are identified as those with the flood profile $h > 0$. According to the TWI method, in a given area, there is 5% probability that the zone is flood prone. Moreover, the probability that a zone is not inundated (IN) given that it is indicated as flood prone FP is equal to 5% and the probability that a zone is not flood prone (FP) given that it is indicated as inundated IN is equal to 2%.

- **What is the probability that a given area is not flood prone?**

$$P\left(\overline{FP}\middle|\mathbf{I}\right)=1-P\left(FP\middle|\mathbf{I}\right)=(1-0.05)=0.95$$

- **What is the probability that a given area is inundated given that it is indicated as flood prone?**

$$P\left(IN\middle|FP,\mathbf{I}\right)=1-P\left(\overline{IN}\middle|FP,\mathbf{I}\right)=1-0.05=0.95$$

- **What is the probability that a given area is indicated both as flood prone and inundated?**

$$P\left(FP\cdot IN\middle|\mathbf{I}\right)=P\left(IN\middle|FP,\mathbf{I}\right)P\left(FP\middle|\mathbf{I}\right)=0.95\times0.05=0.0475$$

- **What is the probability that a given area is not inundated but (and) it is flood prone?**

$$P\left(\overline{IN}\cdot FP\middle|\mathbf{I}\right)=P\left(\overline{IN}\middle|FP,\mathbf{I}\right)P\left(FP\middle|\mathbf{I}\right)=0.05\times0.05=0.0025$$

- **Suppose that the probability that a zone is not inundated given that it is not flood prone is equal to 0.90. What is the probability that a given zone is inundated? (hint: consider using the total probability theorem)**

$$P\left(\overline{IN}\middle|\overline{FP},\mathbf{I}\right)=0.90$$

$$P\left(IN\middle|\mathbf{I}\right)=P\left(IN\cdot FP\middle|\mathbf{I}\right)+P\left(IN\cdot\overline{FP}\middle|\mathbf{I}\right)=P\left(IN\cdot FP\middle|\mathbf{I}\right)+P\left(IN\middle|\overline{FP},\mathbf{I}\right)P\left(\overline{FP}\middle|\mathbf{I}\right)$$
$$=0.0475+(1-0.90)\times 0.95=0.1425$$

3. Structural Damage Identification

The probability that a structural member develops fatigue crack is 10^{-4}. Suppose that a non-destructive inspection is performed on the structural member in order to check whether it has developed a fatigue crack. However, the inspection is subject to some error: That is, the probability that the inspection rejects (says that the member has developed fatigue crack) a member with no fatigue crack is 0.01 (given no fatigue crack), and the probability that the test accepts a member with a fatigue crack is 0.01 (given fatigue crack).

- **What is the probability that the member has fatigue crack?**

  Consider the following logical statements:

  $D \equiv$ inspection detects a member crack

  $Cr \equiv$ the crack exist

  $P\left(D\middle|\overline{Cr},\mathbf{I}\right)=0.01,\ P\left(\overline{D}\middle|Cr,\mathbf{I}\right)=0.01$

  $P\left(Cr\middle|\mathbf{I}\right)=10^{-4}$

- **What is the probability that the inspection rejects a member (given) with fatigue crack?**

  $P\left(D\middle|Cr,\mathbf{I}\right)=1-P\left(\overline{D}\middle|Cr,\mathbf{I}\right)=0.99$

- **What is the probability that the member has fatigue crack and it is also rejected by the inspection?**

  $P\left(Cr.D\middle|\mathbf{I}\right)=P\left(D\middle|Cr,\mathbf{I}\right)P\left(Cr\middle|\mathbf{I}\right)=(1-0.01)\times 10^{-4}\cong 10^{-4}$

- **What is the probability that the member has no fatigue crack and it is rejected by the inspection?**

  $P\left(\overline{Cr}.D\middle|\mathbf{I}\right)=P\left(D\middle|\overline{Cr},\mathbf{I}\right)P\left(\overline{Cr}\middle|\mathbf{I}\right)=0.01\times(1-10^{-4})\cong 0.01$

- **What is the probability that the inspection rejects the member?**

$$P\left(D\middle|\mathbf{I}\right)=P\left(D\cdot Cr\middle|\mathbf{I}\right)+P\left(D\cdot\overline{Cr}\middle|\mathbf{I}\right)=10^{-4}+0.01\cong 0.01$$

- **What is the probability that the inspection accepts the member?**

$$P\left(\overline{D}\middle|\mathbf{I}\right)=0.99$$

- **Imagine that the inspection has rejected the member. What is the probability that the member has developed a fatigue crack (that means with what probability you believe the test results?**

$$P\left(Cr\middle|D,\mathbf{I}\right)=\frac{P\left(Cr.D\middle|\mathbf{I}\right)}{P\left(D\middle|\mathbf{I}\right)}=\frac{10^{-4}}{0.01}=0.01$$

4. The weather forecasting problem

Suppose that during the month of October on average there is 10% probability that it rains heavily within a time interval of 24 hours. If there is 80% probability that the roof will leak due to heavy rain,

- **What is the probability that it rains heavily around noon today?**

  Consider the following logical statements:

  $R \equiv$ it rains heavily within a time interval of 24 hours

  $Lr \equiv$ the roof leak due to heavy rain

  $P\left(R\middle|\mathbf{I}\right)=0.10,\ P\left(Lr\middle|R,\mathbf{I}\right)=0.80$

  $P\left(R\middle|\mathbf{I}\right)=0.10$

- **What is the probability that it rains heavily around noon today and the roof leaks?**

  $P\left(Lr\cdot R\middle|\mathbf{I}\right)=P\left(Lr\middle|R,\mathbf{I}\right)P\left(R\middle|\mathbf{I}\right)=0.80\times 0.10=0.08$

- **What is the probability that it does not rain heavily around noon today however the roof leaks? Suppose that there is only 2% probability that the roof leaks when there has not been a heavy rain before.**

  $P\left(Lr\middle|\overline{R},\mathbf{I}\right)=0.02$

  $P\left(\overline{R}\cdot Lr\middle|\mathbf{I}\right)=P\left(Lr\middle|\overline{R},\mathbf{I}\right)P\left(\overline{R}\middle|\mathbf{I}\right)=0.02\times(1-0.1)=0.018$

- **What is the probability that that the roof will leak this afternoon?**

  $P\left(Lr\middle|\mathbf{I}\right)=P\left(Lr\cdot(R+\overline{R})\middle|\mathbf{I}\right)=P\left(Lr\cdot R\middle|\mathbf{I}\right)+P\left(Lr\cdot\overline{R}\middle|\mathbf{I}\right)=0.08+0.018=0.098$

- **Imagine that you have spent the whole day on the midterm exam and have not looked outside the window to see if it has rained. In the afternoon you notice that the roof of the classroom is leaking, what is the probability that it has rained earlier today?**

$$P\left(R\middle|Lr,\mathbf{I}\right)=\frac{P\left(Lr\cdot R\middle|\mathbf{I}\right)}{P\left(Lr\middle|\mathbf{I}\right)}=\frac{0.08}{0.098}\cong 0.8$$

5. Multiple Hypothesis Testing

Let the background information *X* to be the following statement:

X ≡ We have 11 automatic machines producing gadgets, which pour out of the machines into 11 boxes. This example corresponds to a very early state in the development of the gadgets, because 10 of the machines produce one in six defective. The 11th machine is even worse; it makes one in three defective. There is also the possibility that something goes entirely wrong in the machine that made our gadgets and it is turning out 99% defective. The output of each machine has been collected in an unlabeled box and stored in the warehouse.

The prior evidences (hypotheses) for various propositions of interest are as follows:

A ≡ we choose a bad batch 1/3 defective

B ≡ we choose a good batch 1/6 defective

C ≡ we choose a bad batch 99/100 defective

The prior probability of each hypothesis is defined as:

$$P\left(A|X\right)=\frac{1}{11}(1-10^{-6})$$

$$P\left(B|X\right)=\frac{10}{11}(1-10^{-6})$$

$$P\left(C|X\right)=10^{-6}$$

The data proposition D would stand for the statement that:

D ≡ *m* gadgets were tested and everyone was defective.

- **Calculate *P*(*D*|*AX*), P(*D*|*BX*), P(*D*|*CX*)**

  Considering *m* independent samples:

$$P\left(D|AX\right)=\left(\frac{1}{3}\right)^{m}$$

$$P\left(D|BX\right)=\left(\frac{1}{6}\right)^{m}$$

$$P\left(D\middle|CX\right)=\left(\frac{99}{100}\right)^m$$

- **Calculate *P*(*D*|*X*)**

$$\begin{aligned}P\left(D\middle|X\right)&=P\left(D\middle|AX\right)P\left(A\middle|X\right)+P\left(D\middle|BX\right)P\left(B\middle|X\right)+P\left(D\middle|CX\right)P\left(C\middle|X\right)\\&=\left(\frac{1}{3}\right)^m\times\frac{1}{11}(1-10^{-6})+\left(\frac{1}{6}\right)^m\times\frac{10}{11}(1-10^{-6})+\left(\frac{99}{100}\right)^m\times10^{-6}\\&\cong\left(\frac{1}{3}\right)^m\times\frac{1}{11}+\left(\frac{1}{6}\right)^m\times\frac{10}{11}+\left(\frac{99}{100}\right)^m\times10^{-6}\end{aligned}$$

- **Calculate *P*(*A*|*DX*), P(*B*|*DX*), P(*C*|*DX*)**

$$P\left(A\middle|DX\right)=\frac{P\left(D\middle|AX\right)P\left(A\middle|X\right)}{P\left(D\middle|X\right)}\cong\frac{\left(\frac{1}{3}\right)^m\times\frac{1}{11}}{\left(\frac{1}{3}\right)^m\times\frac{1}{11}+\left(\frac{1}{6}\right)^m\times\frac{10}{11}+\left(\frac{99}{100}\right)^m\times10^{-6}}$$

$$P\left(B\middle|DX\right)\cong\frac{\left(\frac{1}{6}\right)^m\times\frac{10}{11}}{\left(\frac{1}{3}\right)^m\times\frac{1}{11}+\left(\frac{1}{6}\right)^m\times\frac{10}{11}+\left(\frac{99}{100}\right)^m\times10^{-6}}$$

$$P\left(C\middle|DX\right)\cong\frac{\left(\frac{99}{100}\right)^m\times10^{-6}}{\left(\frac{1}{3}\right)^m\times\frac{1}{11}+\left(\frac{1}{6}\right)^m\times\frac{10}{11}+\left(\frac{99}{100}\right)^m\times10^{-6}}$$

- **Calculate the evidence *e*(*A*|*DX*), *e*(*B*|*DX*), *e*(*C*|*DX*)**

In order to calculate the evidence, let's start with the derivation of the evidence $e(A|DX)$:

$$P\left(A\middle|DX\right)=\frac{P\left(D\middle|AX\right)P\left(A\middle|X\right)}{P\left(D\middle|X\right)},\quad P\left(\overline{A}\middle|DX\right)=\frac{P\left(D\middle|\overline{A}X\right)P\left(\overline{A}\middle|X\right)}{P\left(D\middle|X\right)}$$

Take the ratio of both equations (note that the ratio of the probability that Hypothesis A is TRUE to the probability that it is FALSE, is called "*odds*" on the proposition A):

$$\frac{P\left(A\middle|DX\right)}{P\left(\overline{A}\middle|DX\right)}=\frac{P\left(D\middle|AX\right)}{P\left(D\middle|\overline{A}X\right)}\frac{P\left(A\middle|X\right)}{P\left(\overline{A}\middle|X\right)}$$

The evidence of hypothesis *A* given data *D* and background information *X* is defined as the logarithm in base 10 of ratio of the probability that the hypothesis A is TRUE to the probability that it is FALSE multiplied by 10 (the evidence is measured in *decibels*, *db*):

$$e\left(A\middle|DX\right)\triangleq 10\log_{10}\left(\frac{P(A|DX)}{P(\overline{A}|DX)}\right)$$

Thus, we have:

$$e\left(A\middle|DX\right)=10\log_{10}\left(\frac{P\left(A\middle|X\right)}{P\left(\overline{A}\middle|X\right)}\frac{P\left(D\middle|AX\right)}{P\left(D\middle|\overline{A}X\right)}\right)=e\left(A\middle|X\right)+10\log_{10}\left(\frac{P\left(D\middle|AX\right)}{P\left(D\middle|\overline{A}X\right)}\right)$$

where,

$$e\left(A\middle|X\right)=10\log_{10}\left(\frac{P\left(A\middle|X\right)}{P\left(\overline{A}\middle|X\right)}\right)=10\log_{10}\left(\frac{P\left(A\middle|X\right)}{1-P\left(A\middle|X\right)}\right)=10\log_{10}\left(\frac{1}{10}\right)=-10$$

In order to estimate $P\left(D\middle|\overline{A}X\right)$:

$$P\left(D\middle|\overline{A}X\right)=\frac{P\left(\overline{A}\middle|DX\right)P\left(D\middle|X\right)}{P\left(\overline{A}\middle|X\right)}$$

Since *A*, *B* and *C* are MECE hypothesis,

$$P\left(\overline{A}\middle|DX\right)=1-P\left(A\middle|DX\right)=P\left(B\middle|DX\right)+P\left(C\middle|DX\right)$$
$$P\left(\overline{A}\middle|X\right)=1-P\left(A\middle|X\right)$$

Hence,

$$P\left(D\middle|\overline{A}X\right)=\frac{P\left(B\middle|DX\right)P\left(D\middle|X\right)+P\left(C\middle|DX\right)P\left(D\middle|X\right)}{1-P\left(A\middle|X\right)}$$

$$=\frac{P\left(D\middle|BX\right)P\left(B\middle|X\right)+P\left(D\middle|CX\right)P\left(C\middle|X\right)}{1-P\left(A\middle|X\right)}$$

$$=\frac{\left(\frac{1}{6}\right)^m\times\frac{10}{11}+\left(\frac{99}{100}\right)^m\times10^{-6}}{1-\frac{1}{11}(1-10^{-6})}=\frac{\left(\frac{1}{6}\right)^m\times\frac{10}{11}+\left(\frac{99}{100}\right)^m\times10^{-6}}{\frac{10}{11}}$$

$$=\left(\frac{1}{6}\right)^m+\left(\frac{99}{100}\right)^m\frac{11}{10}\times10^{-6}$$

As a result,

$$e\left(A\middle|DX\right)=-10+10\log_{10}\left(\frac{\left(\frac{1}{3}\right)^m}{\left(\frac{1}{6}\right)^m+\left(\frac{99}{100}\right)^m\frac{11}{10}\times10^{-6}}\right)$$

Similarly, we have:

$$e\left(B\middle|DX\right)=e\left(B\middle|X\right)+10\log_{10}\left(\frac{P\left(D\middle|BX\right)}{P\left(D\middle|\overline{B}X\right)}\right)$$

where,

$$e\left(B\middle|X\right)=10\log_{10}\left(\frac{P\left(A\middle|X\right)}{1-P\left(B\middle|X\right)}\right)=10\log_{10}(10)=+10$$

$$P\left(D\middle|\overline{B}X\right)=\frac{P\left(D\middle|AX\right)P\left(A\middle|X\right)+P\left(D\middle|CX\right)P\left(C\middle|X\right)}{1-P\left(B\middle|X\right)}$$

$$=\frac{\left(\frac{1}{3}\right)^m\times\frac{1}{11}+\left(\frac{99}{100}\right)^m\times10^{-6}}{\frac{1}{11}}=\left(\frac{1}{3}\right)^m+\left(\frac{99}{100}\right)^m11\times10^{-6}$$

Therefore,

$$e\left(B\middle|DX\right)=+10+10\log_{10}\left(\frac{\left(\frac{1}{6}\right)^m}{\left(\frac{1}{3}\right)^m+\left(\frac{99}{100}\right)^m 11\times10^{-6}}\right)$$

Finally,

$$e\left(C\middle|DX\right)=e\left(C\middle|X\right)+10\log_{10}\left(\frac{P\left(D\middle|CX\right)}{P\left(D\middle|\overline{C}X\right)}\right)$$

where,

$$e\left(C\middle|X\right)=10\log_{10}\left(10^{-6}\right)=-60$$

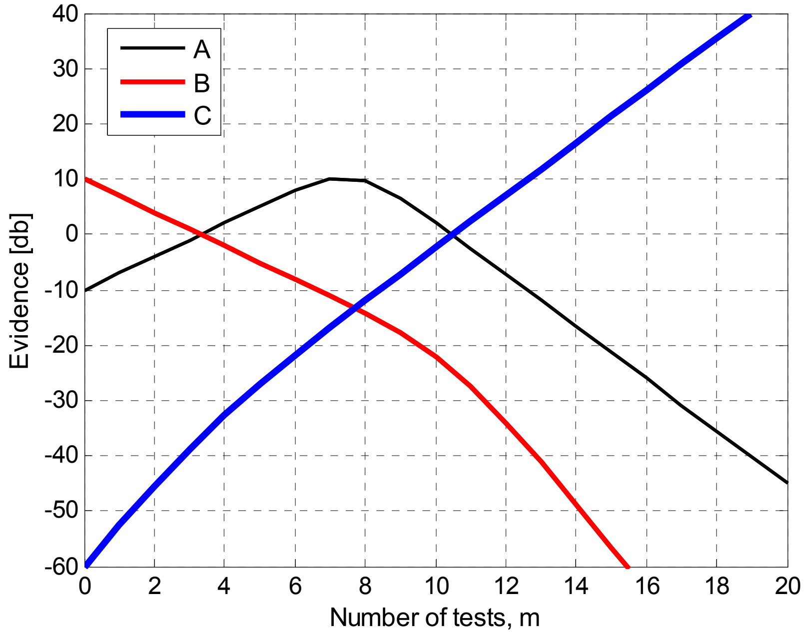
$$P\left(D\middle|\overline{C}X\right)=\frac{P\left(D\middle|AX\right)P\left(A\middle|X\right)+P\left(D\middle|BX\right)P\left(B\middle|X\right)}{1-P\left(C\middle|X\right)}$$

$$=\frac{\left(\frac{1}{3}\right)^m\times\frac{1}{11}+\left(\frac{1}{6}\right)^m\times\frac{10}{11}}{1}=\left(\frac{1}{3}\right)^m\times\frac{1}{11}+\left(\frac{1}{6}\right)^m\times\frac{10}{11}$$

Therefore,

$$e\left(C\middle|DX\right)=-60+10\log_{10}\left(\frac{\left(\frac{99}{100}\right)^m}{\left(\frac{1}{3}\right)^m\times\frac{1}{11}+\left(\frac{1}{6}\right)^m\times\frac{10}{11}}\right)$$

- **Plot in the same graph *e*(C|DX), *e*(A|DX) and *e*(B|DX) as a function of *m* number of tests (which are all defective).**



- **What is the number of defective results needed in order to accumulate 20 db of information in favor of hypothesis C?**

  According to the above figure, *m* should be around 15

6. Structural Damage Detection

It is known that the error associated with a specific destructive test for identification of a specific damage (defect) is equal to 5%. This means that the probability that the test detects the damage given that the defect does not exist is equal to 5%, and the probability of not damage detection given that the defect exists is also equal to 5%. It is also known that the probability of having damage is equal to 1%.

- **What is the probability that no damage (defect) exists?**

  Consider the following logical statements:

  $V \equiv$ the test detects the defect (damage)

  $D \equiv$ the damage exists

  $$P\left(\overline{D}\middle|\mathbf{I}\right)=1-P\left(D\middle|\mathbf{I}\right)=1-0.01=0.99$$

- **What is the probability that the test detects a defect and that the damage really exists?**

  $$P\left(VD\middle|\mathbf{I}\right)=P\left(V\middle|D\mathbf{I}\right)P\left(D\middle|\mathbf{I}\right)=\left[1-P\left(\overline{V}\middle|D\mathbf{I}\right)\right]P\left(D\middle|\mathbf{I}\right)=\left(1-0.05\right)\times 0.01=0.0095$$

- **What is the probability that the test does not detect a defect and that the defect does not exist?**

  $$P\left(\overline{V}\,\overline{D}\middle|\mathbf{I}\right)=P\left(\overline{V}\middle|\overline{D}\mathbf{I}\right)P\left(\overline{D}\middle|\mathbf{I}\right)=\left[1-P\left(V\middle|\overline{D}\mathbf{I}\right)\right]P\left(\overline{D}\middle|\mathbf{I}\right)=\left(1-0.05\right)\times(1-0.01)\cong 0.95$$

- **What is the probability that damage is detected by the test?**

  $$\begin{aligned}P\left(V\middle|\mathbf{I}\right)&=P\left(VD\middle|\mathbf{I}\right)+P\left(V\overline{D}\middle|\mathbf{I}\right)=P\left(VD\middle|\mathbf{I}\right)+P\left(V\middle|\overline{D}\mathbf{I}\right)\cdot P\left(\overline{D}\middle|\mathbf{I}\right)\\&=0.0095+0.05\times 0.99=0.059\end{aligned}$$

- **The result of a test is that a defect exists. What is the probability that the damage exists?**

$$P\left(D\middle|V\mathbf{I}\right)=\frac{P\left(VD\middle|\mathbf{I}\right)}{P\left(V\middle|\mathbf{I}\right)}=\frac{0.0095}{0.059}=0.161$$