Intelligent Systems
Exercises Block 2, Chapter 1
Probabilistic Reasoning

Escuela Técnica Superior de Informática
Dep. de Sistemas Informáticos y Computación
Universitat Politècnica de València

26 de octubre de 2016

Questions

1 $\boxed{\text{B}}$ (Exam January 15, 2014) Given the joint probability of two random variables X and Y, the conditional probability $P(Y = y \mid X = x)$ can be calculated as:

A) $P(y \mid x) \;=\; 1 \;/\; P(x, y)$
B) $P(y \mid x) \;=\; P(x, y) \;/\; \sum_{y'} P(x, y')$
C) $P(y \mid x) \;=\; \sum_{x'} P(x', y) \;/\; \sum_{y'} P(x, y')$
D) $P(y \mid x) \;=\; \sum_{x'} P(x', y) \cdot \sum_{y'} P(x, y')$

2 $\boxed{\text{A}}$ (Exam January 15, 2014) In a binary decision problem ($D = \{0, 1\}$), let y be a fact or data and $d^\star(y) = 0$ be the decision of optimal classification (minimum error classification). Indicate which of the following expressions is ***not correct*** to determine the minimum probability of error for y:

A) $P_\star(\text{error} \mid Y = y) \;=\; 1 - P(D = 1 \mid Y = y)$
B) $P_\star(\text{error} \mid Y = y) \;=\; 1 - P(D = 0 \mid Y = y)$
C) $P_\star(\text{error} \mid Y = y) \;=\; P(D = 1 \mid Y = y)$
D) $P_\star(\text{error} \mid Y = y) \;=\; 1 - \text{máx}_d\, P(D = d \mid Y = y)$

3 $\boxed{\text{D}}$ (Exam January 15, 2014) In a differential diagnosis between *Flu* and *Cold*, we know that the relative occurrence of *Flu* with respect to *Cold* is 30 %. We know the following distribution of fever values in Celsius degrees:

$t(^oC)$	36	37	38	39	40
$P(T = t \mid D = \text{FLU})$	0.05	0.10	0.20	0.30	0.35
$P(T = t \mid D = \text{COLD})$	0.10	0.30	0.40	0.15	0.05

The conditional (posterior) probability that a patient has *Flu* given he has a fever of 38^oC is:

A) greater than 0.8
B) lower than 0.1
C) between 0.3 and 0.6
D) lower than the probability that the patient has *Cold* with the same fever of 38^oC

4 $\boxed{\text{D}}$ (Exam January 28, 2014) In a classification experiment with 300 test samples, 15 wrong decisions were found. With a 95 % of confidence, we can affirm that the true probability of error is:

A) $P(\text{error}) = 5\,\% \pm 0.3\,\%$
B) $P(\text{error}) = 0.05 \pm 0.3$
C) $P(\text{error}) = 0.05$, exactly
D) $P(\text{error}) = 0.05 \pm 0.03$

$$0.05 \pm 1.96\sqrt{\frac{0.05\,0.95}{300}} = 0.05 \pm 0.03 \quad (5\,\% \pm 3\,\%)$$

5 $\boxed{\text{B}}$ (Exam January 28, 2014) In a differential diagnosis between *Flu* and *Cold*, we know that the relative occurrence of *Flu* with respect to *Cold* is 30 %. We know the following distribution of fever values in Celsius degrees:

$t(^oC)$	36	37	38	39	40
$P(T = t \mid D = \text{FLU})$	0.05	0.10	0.20	0.30	0.35
$P(T = t \mid D = \text{COLD})$	0.10	0.30	0.40	0.15	0.05

$$P(\text{FLU} \mid 37) = \frac{\frac{30}{130}\,0.10}{\frac{30}{130}\,0.10 + \frac{100}{130}\,0.30} = \frac{1}{11}$$

The most probable diagnosis for a patient that has a fever of 37^oC is:

A) *Flu*
B) *Cold*
C) There is a tie between the two diagnosis.
D) The given probabilities are not correct because they don't sum up 1; therefore, a diagnosis cannot be made.

6 $\boxed{\text{C}}$ (January 13, 2015) Regarding the Bayes' rule, which of the following expressions is ***not correct***?

A) $P(x \mid y) \;=\; \dfrac{P(y, x)}{\sum_z P(y \mid z)\; P(z)}$

B) $P(x \mid y) \;=\; \dfrac{P(x, y)}{\sum_z P(y, z)}$

C) $P(x \mid y) \;=\; \dfrac{\sum_z P(x, z)}{P(y)}$

D) $P(x \mid y) \;=\; \dfrac{P(y \mid x)\; P(x)}{P(y)}$

7 $\boxed{\text{B}}$ (January 13, 2015) The commercial assessment of the 300 movies screened in a cinema over the last year was *success* for 120 movies and *failure* for the rest of the movies. We know the distribution of the movie genres given its commercial assessment:

g	Romance	Comedy	Intrigue
$P(G = g \mid A =$ Success$)$	0.30	0.35	0.35
$P(G = g \mid A =$ Failure$)$	0.20	0.50	0.30

Which is the most probable commercial assessment for an intrigue film?

A) *Success*
B) *Failure* $\quad P(V = \text{Fracaso} \mid \text{G}=\text{Intriga}) = 0.5625$
C) Both commercial assessments have the same probability
D) It is impossible to determine the prediction with the available data

8 $\boxed{\text{D}}$ (January 13, 2015) In a classification problem in three classes, $C = \{a, b, c\}$, let y be a fact or data. The decision of optimal classification for y is class a with a posterior probability of 0.40. Which of the following assertions is ***incorrect***?

A) $P(C = a \mid Y = y) \;\leq\; P(C = b \mid Y = y) + P(C = c \mid Y = y)$
B) $P_\star(\text{error} \mid Y = y) \;=\; P(C = b \mid Y = y) + P(C = c \mid Y = y)$
C) $P_\star(\text{error} \mid Y = y) \;=\; 1 - P(C = a \mid Y = y)$
D) $P_\star(\text{error} \mid Y = y) \;=\; 1 - \text{máx}_{d \in \{b,c\}} P(C = d \mid Y = y)$

9 $\boxed{\text{D}}$ (January 26, 2015) Let X, Y and Z be three random variables. X and Y are *conditionally independent* given Z if and only if

$$P(X = x, Y = y \mid Z = z) \;=\; P(X = x \mid Z = z)\, P(Y = y \mid Z = z) \qquad \text{for all } x,\ y \text{ y } z.$$

If the above condition holds, $P(Z = z \mid X = x, Y = y)$ can be calculated as ...:

A) $P(Z = z \mid X = x, Y = y) \;=\; \dfrac{P(X = x, Y = y, Z = z)}{P(X = x, Y = y)}$

B) $P(Z = z \mid X = x, Y = y) \;=\; \dfrac{P(Z = z)\, P(X = x, Y = y \mid Z = z)}{P(X = x, Y = y)}$

C) $P(Z = z \mid X = x, Y = y) \;=\; \dfrac{P(Z = z)\, P(X = x \mid Z = z)\, P(Y = y \mid Z = z)}{P(X = x, Y = y)}$

D) The three above answers are all correct to calculate $P(Z = z \mid X = x, Y = y)$.

10 $\boxed{\text{C}}$ (January 26, 2015) Let be a classification problem in three classes, $C = \{a, b, c\}$, where the number of samples of class a is 100, the number of samples of class b is 100, and the number of samples of class c is 100, and let y be a fact or data. The decision of optimal classification for y is class a with a posterior probability of 0.50. Which of the following assertions is ***correct***?

A) $P(C = a \mid Y = y) \;>\; P(C = b \mid Y = y) + P(C = c \mid Y = y)$

B) $P(Y=y \mid C=a) \;=\; \dfrac{0.5\, P(C=a)}{P(Y=y)}$

C) $P(Y=y \mid C=a) \;=\; P(Y=y \mid C=b) + P(Y=y \mid C=c)$

D) None of the above.

11 $\boxed{\text{D}}$ (January, 2016) Which of the following expressions is **CORRECT**?

A) $P(x \mid y) \;=\; \dfrac{1}{P(z)} \sum_x P(x,y,z)$

B) $P(x \mid y) \;=\; \dfrac{1}{P(z)} \sum_z P(x,y,z)$

C) $P(x \mid y) \;=\; \dfrac{1}{P(y)} \sum_x P(x,y,z)$

D) $P(x \mid y) \;=\; \dfrac{1}{P(y)} \sum_z P(x,y,z)$

12 $\boxed{\text{A}}$ (January, 2016) A physician knows that:

- The meningitis disease causes neck stiffness in the 70 % of the cases.
- The prior probability that a patient suffers from meningitis is 1 / 100 000.
- The prior probability that a patient has neck stiffness is 1 %.

Based on the above knowledge, the probability P that a patient who has neck stiffness suffers from meningitis is:

A) $0.000 \leq P < 0.001$ $\quad P = P(m \mid r) = \frac{P(m)\, P(r|m)}{P(r)} = \frac{1/100\,000 \cdot 70/100}{1/100} = 0.0007$

B) $0.001 \leq P < 0.002$

C) $0.002 \leq P < 0.003$

D) $0.003 \leq P$

13 $\boxed{\text{D}}$ (January 2016) Which of the following assertions is **TRUE**?

A) $P(x,y) \;=\; \sum_z P(x)\, P(y)\, P(z).$

B) $P(x,y) \;=\; \sum_z P(x)\, P(y \mid z).$

C) $P(x,y) \;=\; \sum_z P(x \mid z)\, P(y \mid z)\, P(z).$

D) $P(x,y) \;=\; \sum_z P(x,y \mid z)\, P(z).$ $\quad P(x,y) = \sum_z P(x,y,z) = \sum_z P(x,y \mid z)\, P(z)$

14 $\boxed{\text{A}}$ (January 2016) An entomologist discovers a rare subspecies of beetle, due to the pattern of his back. In this rare subspecies, 98 % of the specimen have this pattern. In the common subspecies, 5 % of the specimen have this pattern. The rare subspecies represents 0.1 % of the population. The probability P that a beetle with the pattern of his back belongs to the rare subspecies is:

A) $0.00 \leq P < 0.05.$ $\; P = P(r \mid p) = \frac{P(r)\, P(p|r)}{P(p)} = \frac{P(r)\, P(p|r)}{P(r)\, P(p|r) + P(c)\, P(p|c)} = \frac{1/1000 \cdot 98/100}{1/1000 \cdot 98/100 + 999/1000 \cdot 5/100} = \frac{98}{5093} = 0.0192$

B) $0.05 \leq P < 0.10.$

C) $0.10 \leq P < 0.20.$

D) $0.20 \leq P.$

15 $\boxed{\text{C}}$ (January 2016) Let x be an object (represented with a feature vector or string of symbols) that we want to classify in one among C possible classes. Indicate which of the following expressions **DOES NOT** classify x by minimum classification error:

A) $c(x) \;=\; \underset{c=1,\ldots,C}{\arg\max}\; \log_2 p(c \mid x)$

B) $c(x) \;=\; \underset{c=1,\ldots,C}{\arg\max}\; \log_{10} p(c \mid x)$

C) $c(x) \;=\; \underset{c=1,\ldots,C}{\arg\max}\; a\, p(c \mid x) + b$ $\quad$ being a and b two real constants

D) $c(x) \;=\; \underset{c=1,\ldots,C}{\arg\max}\; p(c \mid x)^3$

Problems

1. (Exam November 26, 2012) In order to design a differential diagnosis between Flu and Cold, we use the histograms of body temperatures (fever) of a sample of patients who had these diseases. From the histograms, we get the following distribution of fever values in Celsius degrees:

$f(^oC)$	36	37	38	39	40
$P(F=f \mid D=\text{FLU})$	0.05	0.10	0.20	0.30	0.35
$P(F=f \mid D=\text{COLD})$	0.10	0.30	0.40	0.15	0.05

knowing that the relative occurrence of flu compared to cold is 30 % (i.e., $P(D=\text{FLU})=0.30$), calculate:

a) The conditional (posterior) probability that a patient has flu given he has a fever of 39^oC

b) The most probable diagnosis for this patient and the probability of error

c) The probabilities of diagnosis FLU and COLD $\forall f \in \{36,37,38,39,40\}$, and the minimum average probability of error ($P_\star(error)$) for a diagnosis system designed with the above observations.

Solution

a)

$$P(D=\text{FLU}) \;=\; 0.3; \qquad P(D=\text{COLD}) \;=\; 1-0.3 \;=\; 0.7$$

$$P(D=\text{FLU} \mid T=39) \;=\; \frac{P(D=\text{FLU})P(T=39 \mid D=\text{FLU})}{P(T=39)}$$

$$P(T=39) \;=\; P(D=\text{FLU})P(T=39 \mid D=\text{FLU}) \;+\; P(D=\text{COLD})P(T=39 \mid D=\text{COLD}) \;=$$
$$0.3\cdot 0.3+0.7\cdot 0.15 \;=\; 0.195$$

$$P(D=\text{FLU} \mid T=39) \;=\; \frac{0.3\cdot 0.3}{0.195} \;=\; 0.462$$

b)

$$P(D=\text{COLD} \mid T=39) \;=\; \frac{0.7\cdot 0.15}{0.195} \;=\; 0.538$$

Most probable diagnosis:

$$d^\star(T=39) \;=\; \operatorname*{arg\,max}_{d\in\{\text{FLU},\text{COLD}\}} P(D=d \mid T=39) \;=\; \text{COLD}$$

Probability that COLD is a wrong diagnosis for para $t=39$:

$$P_\star(\text{error} \mid T=39) \;=\; 1-\text{máx}(P(D=\text{FLU} \mid T=39), P(D=\text{COLD} \mid T=39)) \;=\; 1-\text{máx}(0.462,0.538) \;=\; 0.462$$

c) Repeating the same calculations for $t \in \{36,37,38,40\}$:

$t(^oC)$	36	37	38	39	40
$P(T=t)$	0.085	0.240	0.340	0.195	0.140
$P(D=\text{FLU} \mid T=t)$	0.176	0.125	0.176	0.462	0.750
$P(D=\text{COLD} \mid T=t)$	0.824	0.875	0.824	0.538	0.250
$P_\star(\text{error} \mid t)$	0.176	0.125	0.176	0.462	0.250

$$P_\star(\text{error}) \;=\; \sum_{t=36}^{40} P_\star(\text{error} \mid T=t)P(T=t) \;=$$
$$0.176\cdot 0.085+0.125\cdot 0.240+0.176\cdot 0.340+0.462\cdot 0.195+0.250\cdot 0.140 \;=\; .230$$

2. **Exercise Irish Flowers, slide #22 of Chapter 1**

petal sizes in cm^2

$P(x \mid c)$	<1	1	2	3	4	5	6	7	8	9	10	>10
SETO	0.90	0.10	0	0	0	0	0	0	0	0	0	0
VERS	0	0	0	0.20	0.30	0.32	0.12	0.06	0	0	0	0
VIRG	0	0	0	0	0	0	0.08	0.12	0.24	0.14	0.20	0.22

Solution

Same prior probabilities

Assuming that the three classes have the same probability, calculate:

a) The conditional (posterior) probabilities $P(c \mid x)$, $c \in \{\text{Seto},\text{Vers},\text{Virg}\}$, for a flower whose petal size is $x = 7\,\text{cm}^2$

$$P(\text{SETO} \mid 7) = \frac{P(7 \mid \text{SETO})P(\text{SETO})}{P(7)} = \frac{0 * 0.33}{P(7)} = 0$$

$$P(\text{VERS} \mid 7) = \frac{P(7 \mid \text{VERS})P(\text{VERS})}{P(7)} = \frac{0.06 * 0.33}{P(7)} = \frac{0.06 * 0.33}{0.06} = \frac{0.02}{0.06} = 0.33$$

$$P(\text{VIRG} \mid 7) = \frac{P(7 \mid \text{VIRG})P(\text{VIRG})}{P(7)} = \frac{0.12 * 0.33}{P(7)} = \frac{0.12 * 0.33}{0.06} = \frac{0.04}{0.06} = 0.67$$

$P(7) = P(7 \mid \text{SETO}) * P(\text{SETO}) + P(7 \mid \text{VERS}) * P(\text{VERS}) + P(7 \mid \text{VIRG}) * P(\text{VIRG}) = 0 + 0.06 * 0.33 + 0.12 * 0.33 = 0.0198 + 0.0396 = 0.06$

b) The decision of optimal classification for this flower and the probability of taking a wrong decision.
Best decision for this flower is class Virginica. The probability of taking a wrong decision is 0.33 $(1 - \text{máx}_{d \in \mathcal{D}} P(d \mid x) = 1 - 0.67 = 0.33)$. $P(error \mid 7) = 0.33$.

c) The best decision and probability of error for petals $1, 2, \ldots, 10$ cm^2
For petal = 1cm

$$P(\text{SETO} \mid 1) = \frac{P(1 \mid \text{SETO})P(\text{SETO})}{P(1)} = \frac{0.10 * 0.33}{0.033} = \frac{0.033}{0.033} = 1$$

$$P(\text{VERS} \mid 1) = \frac{P(1 \mid \text{VERS})P(\text{VERS})}{P(1)} = \frac{0 * 0.33}{P(1)} = \frac{0}{0.033} = 0$$

$$P(\text{VIRG} \mid 1) = \frac{P(1 \mid \text{VIRG})P(\text{VIRG})}{P(1)} = \frac{0 * 0.33}{P(1)} = \frac{0}{0.033} = 0$$

$P(1) = P(1 \mid \text{SETO}) * P(\text{SETO}) + P(1 \mid \text{VERS}) * P(\text{VERS}) + P(1 \mid \text{VIRG}) * P(\text{VIRG}) = 0.10 * 0.33 + 0 * 0.33 + 0 * 0.33 = 0.033$

This case is clear. The best decision is the only one possible: Setosa. So the probability of error is 0. $P(error \mid 1) = 0$.

For petal = 2cm

$$P(\text{SETO} \mid 2) = \frac{P(2 \mid \text{SETO})P(\text{SETO})}{P(2)} = \frac{0 * 0.33}{0} = \frac{0}{0} = 0$$

$$P(\text{VERS} \mid 2) = \frac{P(2 \mid \text{VERS})P(\text{VERS})}{P(2)} = \frac{0 * 0.33}{P(2)} = \frac{0}{0} = 0$$

$$P(\text{VIRG} \mid 2) = \frac{P(2 \mid \text{VIRG})P(\text{VIRG})}{P(2)} = \frac{0 * 0.33}{P(2)} = \frac{0}{0} = 0$$

$P(2) = P(2 \mid \text{SETO}) * P(\text{SETO}) + P(2 \mid \text{VERS}) * P(\text{VERS}) + P(2 \mid \text{VIRG}) * P(\text{VIRG}) = 0 * 0.33 + 0 * 0.33 + 0 * 0.33 = 0$

This case is clear. There is no best decision possible so the probability of error is 1. $P(error \mid 2) = 1$.

For petal = 3cm

$$P(\text{SETO} \mid 3) = \frac{P(3 \mid \text{SETO})P(\text{SETO})}{P(3)} = \frac{0 * 0.33}{0.033} = \frac{0}{0.067} = 0$$

$$P(\text{VERS} \mid 3) = \frac{P(3 \mid \text{VERS})P(\text{VERS})}{P(3)} = \frac{0.20 * 0.33}{P(3)} = \frac{0.067}{0.067} = 1$$

$$P(\text{VIRG} \mid 3) = \frac{P(3 \mid \text{VIRG})P(\text{VIRG})}{P(3)} = \frac{0 * 0.33}{P(3)} = \frac{0}{0.067} = 0$$

$P(3) = P(3 \mid \text{SETO}) * P(\text{SETO}) + P(3 \mid \text{VERS}) * P(\text{VERS}) + P(3 \mid \text{VIRG}) * P(\text{VIRG}) = 0 * 0.33 + 0.20 * 0.33 + 0 * 0.33 = 0.067$

This case is clear. The best decision is the only one possible: Versicolor. So the probability of error is 0. $P(error \mid 3) = 0$.

For petal = 4cm

$$P(\text{SETO} \mid 4) = \frac{P(4 \mid \text{SETO})P(\text{SETO})}{P(4)} = \frac{0 * 0.33}{0.033} = \frac{0}{0.01} = 0$$

$$P(\text{VERS} \mid 4) = \frac{P(4 \mid \text{VERS})P(\text{VERS})}{P(4)} = \frac{0.30 * 0.33}{P(4)} = \frac{0.01}{0.01} = 1$$

$$P(\text{VIRG} \mid 4) = \frac{P(4 \mid \text{VIRG})P(\text{VIRG})}{P(4)} = \frac{0 * 0.33}{P(4)} = \frac{0}{0.06} = 0$$

$P(4) = P(4 \mid \text{SETO}) * P(\text{SETO}) + P(4 \mid \text{VERS}) * P(\text{VERS}) + P(4 \mid \text{VIRG}) * P(\text{VIRG}) = 0 * 0.33 + 0.30 * 0.33 + 0 * 0.33 = 0.01$

This case is clear. The best decision is the only one possible: Versicolor. So the probability of error is 0. $P(error \mid 4) = 0$.

For petal = 5cm

$$P(\text{SETO} \mid 5) = \frac{P(5 \mid \text{SETO})P(\text{SETO})}{P(5)} = \frac{0 * 0.33}{0.033} = \frac{0}{0.11} = 0$$

$$P(\text{VERS} \mid 5) = \frac{P(5 \mid \text{VERS})P(\text{VERS})}{P(5)} = \frac{0.33 * 0.33}{P(5)} = \frac{0.11}{0.11} = 1$$

$$P(\text{VIRG} \mid 5) = \frac{P(5 \mid \text{VIRG})P(\text{VIRG})}{P(5)} = \frac{0 * 0.33}{P(5)} = \frac{0}{0.11} = 0$$

$P(5) = P(5 \mid \text{SETO}) * P(\text{SETO}) + P(5 \mid \text{VERS}) * P(\text{VERS}) + P(5 \mid \text{VIRG}) * P(\text{VIRG}) = 0 * 0.33 + 0.32 * 0.33 + 0 * 0.33 = 0.11$

This case is clear. The best decision is the only one possible: Versicolor. So the probability of error is 0. $P(error \mid 5) = 0$.

For petal = 6cm

$$P(\text{SETO} \mid 6) = \frac{P(6 \mid \text{SETO})P(\text{SETO})}{P(6)} = \frac{0 * 0.33}{P(6)} = \frac{0}{0.066} = 0$$

$$P(\text{VERS} \mid 6) = \frac{P(6 \mid \text{VERS})P(\text{VERS})}{P(6)} = \frac{0.12 * 0.33}{P(6)} = \frac{0.0396}{0.066} = 0.60$$

$$P(\text{VIRG} \mid 6) = \frac{P(6 \mid \text{VIRG})P(\text{VIRG})}{P(6)} = \frac{0.08 * 0.33}{P(6)} = \frac{0.0264}{0.066} = 0.40$$

$P(6) = P(6 \mid \text{SETO}) * P(\text{SETO}) + P(6 \mid \text{VERS}) * P(\text{VERS}) + P(6 \mid \text{VIRG}) * P(\text{VIRG}) = 0 * 0.33 + 0.12 * 0.33 + 0.08 * 0.33 = 0.0396 + 0.0264 = 0.066$

The best decision is Versicolor. So the probability of error is 0.40. $P(error \mid 6) = 0.40$.

For petal = 8cm

$$P(\text{SETO} \mid 8) = \frac{P(8 \mid \text{SETO})P(\text{SETO})}{P(8)} = \frac{0 * 0.33}{P(8)} = \frac{0}{0.08} = 0$$

$$P(\text{VERS} \mid 8) = \frac{P(8 \mid \text{VERS})P(\text{VERS})}{P(8)} = \frac{0 * 0.33}{P(8)} = \frac{0}{0.08} = 0$$

$$P(\textsc{virg} \mid 8) = \frac{P(8 \mid \textsc{virg})P(\textsc{virg})}{P(8)} = \frac{0.24 * 0.33}{P(8)} = \frac{0.08}{0.08} = 1$$

$P(8) = P(8 \mid \textsc{seto}) * P(\textsc{seto}) + P(8 \mid \textsc{vers}) * P(\textsc{vers}) + P(8 \mid \textsc{virg}) * P(\textsc{virg}) = 0*0.33+0*0.33+0.24*0.33 = 0.0792$

This case is clear. The best decision is the only one possible: Virginica. So the probability of error is 0. $P(error \mid 8) = 0$.

For petal = 9cm

$$P(\textsc{seto} \mid 9) = \frac{P(9 \mid \textsc{seto})P(\textsc{seto})}{P(9)} = \frac{0 * 0.33}{P(9)} = \frac{0}{0.08} = 0$$

$$P(\textsc{vers} \mid 9) = \frac{P(9 \mid \textsc{vers})P(\textsc{vers})}{P(9)} = \frac{0 * 0.33}{P(9)} = \frac{0}{0.08} = 0$$

$$P(\textsc{virg} \mid 9) = \frac{P(9 \mid \textsc{virg})P(\textsc{virg})}{P(9)} = \frac{0.14 * 0.33}{P(9)} = \frac{0.0462}{0.0462} = 1$$

$P(9) = P(9 \mid \textsc{seto}) * P(\textsc{seto}) + P(9 \mid \textsc{vers}) * P(\textsc{vers}) + P(9 \mid \textsc{virg}) * P(\textsc{virg}) = 0*0.33+0*0.33+0.14*0.33 = 0.0462$

This case is clear. The best decision is the only one possible: Virginica. So the probability of error is 0. $P(error \mid 9) = 0$.

For petal = 10cm

$$P(\textsc{seto} \mid 10) = \frac{P(10 \mid \textsc{seto})P(\textsc{seto})}{P(10)} = \frac{0 * 0.33}{P(10)} = \frac{0}{0.08} = 0$$

$$P(\textsc{vers} \mid 10) = \frac{P(10 \mid \textsc{vers})P(\textsc{vers})}{P(10)} = \frac{0 * 0.33}{P(10)} = \frac{0}{0.08} = 0.$$

$$P(\textsc{virg} \mid 10) = \frac{P(10 \mid \textsc{virg})P(\textsc{virg})}{P(10)} = \frac{0.20 * 0.33}{P(10)} = \frac{0.066}{0.066} = 1$$

$P(10) = P(10 \mid \textsc{seto}) * P(\textsc{seto}) + P(10 \mid \textsc{vers}) * P(\textsc{vers}) + P(10 \mid \textsc{virg}) * P(\textsc{virg}) = 0*0.33+0*0.33+0.20*0.33 = 0.066$

This case is clear. The best decision is the only one possible: Virginica. So the probability of error is 0. $P(error \mid 10) = 0$.

d) The minimum probability of error for any iris flower; that is, $P_\star(\text{error})$
$P_\star(\text{error}) = \sum_x P(error \mid x)P(x) = P(error \mid < 1)P(< 1) +$
$P(error \mid 1)P(1) + P(error \mid 2)P(2) + \ldots + P(error \mid 10)P(10) +$
$P(error \mid > 10)P(> 10) = 0 + 0 + 0 + \ldots + 0.40 * 0.066 + 0.33 * 0.06 + 0 + 0 + 0 + 0 = 0.0264 + 0.0198 = 0.0462 \approx 0.05 = 5\,\%$ error

Section e): Different prior probabilities

Repeat the same calculations assuming that the prior probabilities are:
$P(\textsc{Seto}) = 0.3,\ P(\textsc{Vers}) = 0.5,\ P(\textsc{Virg}) = 0.2$

a) The conditional (posterior) probabilities $P(c \mid x)$, $c \in \{\textsc{Seto}, \textsc{Vers}, \textsc{Virg}\}$, for a flower whose petal size is $x = 7\,\text{cm}^2$

$$P(\textsc{seto} \mid 7) = \frac{P(7 \mid \textsc{seto})P(\textsc{seto})}{P(7)} = \frac{0 * 0.30}{P(7)} = 0$$

$$P(\textsc{vers} \mid 7) = \frac{P(7 \mid \textsc{vers})P(\textsc{vers})}{P(7)} = \frac{0.06 * 0.50}{P(7)} = \frac{0.06 * 0.50}{0.054} = \frac{0.03}{0.054} = 0.56$$

$$P(\text{VIRG} \mid 7) = \frac{P(7 \mid \text{VIRG})P(\text{VIRG})}{P(7)} = \frac{0.12 * 0.20}{P(7)} = \frac{0.12 * 0.20}{0.054} = \frac{0.024}{0.054} = 0.44$$

$P(7) = P(7 \mid \text{SETO}) * P(\text{SETO}) + P(7 \mid \text{VERS}) * P(\text{VERS}) + P(7 \mid \text{VIRG}) * P(\text{VIRG}) = 0 + 0.06 * 0.50 + 0.12 * 0.20 = 0.03 + 0.024 = 0.054$

b) The decision of optimal classification for this flower and the probability of taking a wrong decision.
Best decision for this flower is class Versicolor. The probability of taking a wrong decision is 0.44 ($1 - \text{máx}_{d \in \mathcal{D}} P(d \mid x) = 1 - 0.56 = 0.44$). $P(error \mid 7) = 0.44$.

c) For the exercise in section c) and d), we only need to compute the values for petal size 6 and 7, which are the petal sizes which condition the value of the minimum average probability of error (minimum probability of error for any iris flower) as the values for the rest of petal sizes is 0.

$$P(\text{SETO} \mid 7) = \frac{P(7 \mid \text{SETO})P(\text{SETO})}{P(7)} = \frac{0 * 0.3}{P(7)} = 0$$

$$P(\text{VERS} \mid 7) = \frac{P(7 \mid \text{VERS})P(\text{VERS})}{P(7)} = \frac{0.06 * 0.5}{P(7)} = \frac{0.06 * 0.5}{0.054} = \frac{0.03}{0.054} = 0.56$$

$$P(\text{VIRG} \mid 7) = \frac{P(7 \mid \text{VIRG})P(\text{VIRG})}{P(7)} = \frac{0.12 * 0.2}{P(7)} = \frac{0.12 * 0.2}{0.054} = \frac{0.024}{0.054} = 0.44$$

$P(7) = P(7 \mid \text{SETO}) * P(\text{SETO}) + P(7 \mid \text{VERS}) * P(\text{VERS}) + P(7 \mid \text{VIRG}) * P(\text{VIRG}) = 0 + 0.06 * 0.50 + 0.12 * 0.20 = 0.03 + 0.024 = 0.054$

$$P(\text{SETO} \mid 6) = \frac{P(6 \mid \text{SETO})P(\text{SETO})}{P(6)} = \frac{0 * 0.30}{P(6)} = 0$$

$$P(\text{VERS} \mid 6) = \frac{P(6 \mid \text{VERS})P(\text{VERS})}{P(6)} = \frac{0.12 * 0.50}{P(6)} = \frac{0.06}{0.076} = 0.79$$

$$P(\text{VIRG} \mid 6) = \frac{P(6 \mid \text{VIRG})P(\text{VIRG})}{P(6)} = \frac{0.08 * 0.20}{P(6)} = \frac{0.016}{0.076} = 0.21$$

$P(6) = P(6 \mid \text{SETO}) * P(\text{SETO}) + P(6 \mid \text{VERS}) * P(\text{VERS}) + P(6 \mid \text{VIRG}) * P(\text{VIRG}) = 0 + 0.12 * 0.50 + 0.08 * 0.20 = 0.06 + 0.016 = 0.076$

d) The minimum probability of error for any iris flower; that is, $P_\star(\text{error})$
$P_\star(\text{error}) = \sum_x P(error \mid x)P(x) = P(error \mid < 1)P(< 1) +$
$P(error \mid 1)P(1) + P(error \mid 2)P(2) + \ldots + P(error \mid 10)P(10) +$
$P(error \mid > 10)P(> 10) = 0+0+0+\ldots+0.21*0.076+0.44*0.054+0+0+0+0 = 0.016+0.024 = 0.04 \rightarrow 4\,\%$ error