

- General intro to Bayesian methods
 - Deciphering Enigma messages
 - Air France 447
 - Body Mass Index
- Recap of probability
 - Frequency, classical and subjective interpretations
 - Limitations
 - Rules: conditional probability, total probability, Bayes
 - Examples: blood test, **car diagnosis**, multiple choice exam

Example 1.3

Suppose that your car suffers from two intermittent problems, one caused by a fault in the engine (θ_1) and the other due to a fault in the gearbox (θ_2).

These occur with probabilities 0.4 and 0.6 respectively.

When examined your car exhibits one of the following symptoms

x_1 : overheating only,

x_2 : irregular traction only,

x_3 : both symptoms.

Example 1.3

Suppose it is known in the garage trade that these symptoms occur with probabilities that depend on the fault.

The probabilities $\Pr(X = x|\theta)$ are given in Table 1.1.

	O/H	I/T	Both
	x_1	x_2	x_3
θ_1 : fault in engine	0.1	0.4	0.5
θ_2 : fault in gearbox	0.5	0.3	0.2

Construct a diagnostic rule for these symptoms and determine the probability of misdiagnosis.

Solution to Example 1.3 (1/3)

First we must calculate the **posterior probabilities** $\Pr(\theta_1|x)$ and $\Pr(\theta_2|x)$ for $x = x_1, x_2, x_3$.

Since θ_1 and θ_2 form a partition, we can use Bayes' Theorem as follows.

We have

$$\begin{aligned}\Pr(\theta_1|x_1) &= \frac{\Pr(X = x_1|\theta_1)\Pr(\theta_1)}{\Pr(X = x_1|\theta_1)\Pr(\theta_1) + \Pr(X = x_1|\theta_2)\Pr(\theta_2)} \\&= \frac{0.1 \times 0.4}{0.1 \times 0.4 + 0.5 \times 0.6} \\&= \frac{4}{34} = 0.118.\end{aligned}$$

Solution to Example 1.3 (1/3)

First we must calculate the **posterior probabilities** $\Pr(\theta_1|x)$ and $\Pr(\theta_2|x)$ for $x = x_1, x_2, x_3$.

Since θ_1 and θ_2 form a partition, we can use Bayes' Theorem as follows.

We have

$$\begin{aligned}\Pr(\theta_1|x_1) &= \frac{\Pr(X = x_1|\theta_1)\Pr(\theta_1)}{\Pr(X = x_1|\theta_1)\Pr(\theta_1) + \Pr(X = x_1|\theta_2)\Pr(\theta_2)} \\ &= \frac{0.1 \times 0.4}{0.1 \times 0.4 + 0.5 \times 0.6} \\ &= \frac{4}{34} = 0.118.\end{aligned}$$

Solution to Example 1.3 (1/3)

First we must calculate the **posterior probabilities** $\Pr(\theta_1|x)$ and $\Pr(\theta_2|x)$ for $x = x_1, x_2, x_3$.

Since θ_1 and θ_2 form a partition, we can use Bayes' Theorem as follows.

We have

$$\begin{aligned}\Pr(\theta_1|x_1) &= \frac{\Pr(X = x_1|\theta_1)\Pr(\theta_1)}{\Pr(X = x_1|\theta_1)\Pr(\theta_1) + \Pr(X = x_1|\theta_2)\Pr(\theta_2)} \\ &= \frac{0.1 \times 0.4}{0.1 \times 0.4 + 0.5 \times 0.6} \\ &= \frac{4}{34} = 0.118.\end{aligned}$$

Solution to Example 1.3 (2/3)

Also,

$$\begin{aligned}\Pr(\theta_1|x_2) &= \frac{\Pr(X = x_2|\theta_1)\Pr(\theta_1)}{\Pr(X = x_2|\theta_1)\Pr(\theta_1) + \Pr(X = x_2|\theta_2)\Pr(\theta_2)} \\&= \frac{0.4 \times 0.4}{0.4 \times 0.4 + 0.3 \times 0.6} \\&= \frac{16}{34} = 0.471.\end{aligned}$$

Solution to Example 1.3 (2/3)

Also,

$$\begin{aligned}\Pr(\theta_1|x_2) &= \frac{\Pr(X = x_2|\theta_1)\Pr(\theta_1)}{\Pr(X = x_2|\theta_1)\Pr(\theta_1) + \Pr(X = x_2|\theta_2)\Pr(\theta_2)} \\&= \frac{0.4 \times 0.4}{0.4 \times 0.4 + 0.3 \times 0.6} \\&= \frac{16}{34} = 0.471.\end{aligned}$$

Solution to Example 1.3 (3/3)

And

$$\begin{aligned}\Pr(\theta_1|x_3) &= \frac{\Pr(X = x_3|\theta_1)\Pr(\theta_1)}{\Pr(X = x_3|\theta_1)\Pr(\theta_1) + \Pr(X = x_3|\theta_2)\Pr(\theta_2)} \\ &= \frac{0.5 \times 0.4}{0.5 \times 0.4 + 0.2 \times 0.6} \\ &= \frac{20}{32} = 0.625.\end{aligned}$$

Also, $\Pr(\theta_2|x_i) = 1 - \Pr(\theta_1|x_i)$, $i = 1, 2, 3$, and so we obtain the **posterior distributions** $\Pr(\theta|x)$ given in Table 1.2.

And

$$\begin{aligned}\Pr(\theta_1|x_3) &= \frac{\Pr(X = x_3|\theta_1)\Pr(\theta_1)}{\Pr(X = x_3|\theta_1)\Pr(\theta_1) + \Pr(X = x_3|\theta_2)\Pr(\theta_2)} \\ &= \frac{0.5 \times 0.4}{0.5 \times 0.4 + 0.2 \times 0.6} \\ &= \frac{20}{32} = 0.625.\end{aligned}$$

Also, $\Pr(\theta_2|x_i) = 1 - \Pr(\theta_1|x_i)$, $i = 1, 2, 3$, and so we obtain the **posterior distributions** $\Pr(\theta|x)$ given in Table 1.2.

Example 1.3

	O/H x_1	I/T x_2	Both x_3
θ_1 : fault in engine	0.118	0.471	0.625
θ_2 : fault in gearbox	0.882	0.529	0.375

This table is very informative. For example, it shows that if both symptoms (x_3) are observed, then the probability that the fault is in the engine (θ_1) changes from 0.4 to 0.625.

In terms of **odds**:

$$\text{Prior odds : } \frac{\Pr(\theta_1)}{\Pr(\theta_2)} = \frac{0.4}{0.6} = \frac{2}{3} \quad \text{or 3:2 in favour of } \theta_2$$

$$\text{Posterior odds : } \frac{\Pr(\theta_1|x_3)}{\Pr(\theta_2|x_3)} = \frac{0.625}{0.375} = \frac{5}{3} \quad \text{or 5:3 in favour of } \theta_1.$$

Example 1.3

	O/H x_1	I/T x_2	Both x_3
θ_1 : fault in engine	0.118	0.471	0.625
θ_2 : fault in gearbox	0.882	0.529	0.375

This table is very informative. For example, it shows that if both symptoms (x_3) are observed, then the probability that the fault is in the engine (θ_1) changes from 0.4 to 0.625.

In terms of **odds**:

$$\text{Prior odds : } \frac{\Pr(\theta_1)}{\Pr(\theta_2)} = \frac{0.4}{0.6} = \frac{2}{3} \quad \text{or 3:2 in favour of } \theta_2$$

$$\text{Posterior odds : } \frac{\Pr(\theta_1|x_3)}{\Pr(\theta_2|x_3)} = \frac{0.625}{0.375} = \frac{5}{3} \quad \text{or 5:3 in favour of } \theta_1.$$

Example 1.3

	O/H x_1	I/T x_2	Both x_3
θ_1 : fault in engine	0.118	0.471	0.625
θ_2 : fault in gearbox	0.882	0.529	0.375

This table is very informative. For example, it shows that if both symptoms (x_3) are observed, then the probability that the fault is in the engine (θ_1) changes from 0.4 to 0.625.

In terms of **odds**:

$$\text{Prior odds : } \frac{\Pr(\theta_1)}{\Pr(\theta_2)} = \frac{0.4}{0.6} = \frac{2}{3} \quad \text{or 3:2 in favour of } \theta_2$$

$$\text{Posterior odds : } \frac{\Pr(\theta_1|x_3)}{\Pr(\theta_2|x_3)} = \frac{0.625}{0.375} = \frac{5}{3} \quad \text{or 5:3 in favour of } \theta_1.$$

Example 1.3

	O/H x_1	I/T x_2	Both x_3
θ_1 : fault in engine	0.118	0.471	0.625
θ_2 : fault in gearbox	0.882	0.529	0.375

This table is very informative. For example, it shows that if both symptoms (x_3) are observed, then the probability that the fault is in the engine (θ_1) changes from 0.4 to 0.625.

In terms of **odds**:

$$\text{Prior odds : } \frac{\Pr(\theta_1)}{\Pr(\theta_2)} = \frac{0.4}{0.6} = \frac{2}{3} \quad \text{or 3:2 in favour of } \theta_2$$

$$\text{Posterior odds : } \frac{\Pr(\theta_1|x_3)}{\Pr(\theta_2|x_3)} = \frac{0.625}{0.375} = \frac{5}{3} \quad \text{or 5:3 in favour of } \theta_1.$$

Example 1.3

We are now in a position to design our **diagnostic rule**.

This is simply a rule which diagnoses a symptom (x) as being due to some particular fault (θ).

Consider first that we observe overheating only (x_1). The posterior probabilities are in favour of declaring the fault as in the gearbox (θ_2) since $\Pr(\theta_2|x_1) > \Pr(\theta_1|x_1)$.

In the same way, we can determine the most likely diagnosis having observed irregular traction only (x_2) and both symptoms (x_3), giving the diagnostic rule in Table 1.3.

Symptom	Diagnosis
overheating only (x_1)	fault in gearbox (θ_2)
irregular traction only (x_2)	fault in gearbox (θ_2)
both symptoms (x_3)	fault in engine (θ_1)

Example 1.3

We are now in a position to design our **diagnostic rule**.

This is simply a rule which diagnoses a symptom (x) as being due to some particular fault (θ).

Consider first that we observe overheating only (x_1). The posterior probabilities are in favour of declaring the fault as in the gearbox (θ_2) since $\Pr(\theta_2|x_1) > \Pr(\theta_1|x_1)$.

In the same way, we can determine the most likely diagnosis having observed irregular traction only (x_2) and both symptoms (x_3), giving the diagnostic rule in Table 1.3.

Symptom	Diagnosis
overheating only (x_1)	fault in gearbox (θ_2)
irregular traction only (x_2)	fault in gearbox (θ_2)
both symptoms (x_3)	fault in engine (θ_1)

Example 1.3

We are now in a position to design our **diagnostic rule**.

This is simply a rule which diagnoses a symptom (x) as being due to some particular fault (θ).

Consider first that we observe overheating only (x_1). The posterior probabilities are in favour of declaring the fault as in the gearbox (θ_2) since $\Pr(\theta_2|x_1) > \Pr(\theta_1|x_1)$.

In the same way, we can determine the most likely diagnosis having observed irregular traction only (x_2) and both symptoms (x_3), giving the diagnostic rule in Table 1.3.

Symptom	Diagnosis
overheating only (x_1)	fault in gearbox (θ_2)
irregular traction only (x_2)	fault in gearbox (θ_2)
both symptoms (x_3)	fault in engine (θ_1)

Example 1.3

We are now in a position to design our **diagnostic rule**.

This is simply a rule which diagnoses a symptom (x) as being due to some particular fault (θ).

Consider first that we observe overheating only (x_1). The posterior probabilities are in favour of declaring the fault as in the gearbox (θ_2) since $\Pr(\theta_2|x_1) > \Pr(\theta_1|x_1)$.

In the same way, we can determine the most likely diagnosis having observed irregular traction only (x_2) and both symptoms (x_3), giving the diagnostic rule in Table 1.3.

Symptom	Diagnosis
overheating only (x_1)	fault in gearbox (θ_2)
irregular traction only (x_2)	fault in gearbox (θ_2)
both symptoms (x_3)	fault in engine (θ_1)

Example 1.3

We are now in a position to design our **diagnostic rule**.

This is simply a rule which diagnoses a symptom (x) as being due to some particular fault (θ).

Consider first that we observe overheating only (x_1). The posterior probabilities are in favour of declaring the fault as in the gearbox (θ_2) since $\Pr(\theta_2|x_1) > \Pr(\theta_1|x_1)$.

In the same way, we can determine the most likely diagnosis having observed irregular traction only (x_2) and both symptoms (x_3), giving the diagnostic rule in Table 1.3.

Symptom	Diagnosis
overheating only (x_1)	fault in gearbox (θ_2)
irregular traction only (x_2)	fault in gearbox (θ_2)
both symptoms (x_3)	fault in engine (θ_1)

Solution to Example 1.3 (1/1)

If we were to carry out this diagnostic rule repeatedly then the probability of misdiagnosing a fault is

$$\Pr(\text{Misdiagnosis})$$

$$= \Pr(\theta_1, X = x_1) + \Pr(\theta_1, X = x_2) + \Pr(\theta_2, X = x_3)$$

$$= \Pr(X = x_1 | \theta_1) \Pr(\theta_1) + \Pr(X = x_2 | \theta_1) \Pr(\theta_1) \\ + \Pr(X = x_3 | \theta_2) \Pr(\theta_2)$$

$$= (0.1 \times 0.4) + (0.4 \times 0.4) + (0.2 \times 0.6)$$

$$= 0.32.$$

Therefore, in repeated use of this rule, around a third of the diagnoses will be wrong.

Solution to Example 1.3 (1/1)

If we were to carry out this diagnostic rule repeatedly then the probability of misdiagnosing a fault is

$$\Pr(\text{Misdiagnosis})$$

$$= \Pr(\theta_1, X = x_1) + \Pr(\theta_1, X = x_2) + \Pr(\theta_2, X = x_3)$$

$$= \Pr(X = x_1|\theta_1)\Pr(\theta_1) + \Pr(X = x_2|\theta_1)\Pr(\theta_1) \\ + \Pr(X = x_3|\theta_2)\Pr(\theta_2)$$

$$= (0.1 \times 0.4) + (0.4 \times 0.4) + (0.2 \times 0.6)$$

$$= 0.32.$$

Therefore, in repeated use of this rule, around a third of the diagnoses will be wrong.

Solution to Example 1.3 (1/1)

If we were to carry out this diagnostic rule repeatedly then the probability of misdiagnosing a fault is

$$\Pr(\text{Misdiagnosis})$$

$$= \Pr(\theta_1, X = x_1) + \Pr(\theta_1, X = x_2) + \Pr(\theta_2, X = x_3)$$

$$\begin{aligned} = \Pr(X = x_1|\theta_1)\Pr(\theta_1) + \Pr(X = x_2|\theta_1)\Pr(\theta_1) \\ + \Pr(X = x_3|\theta_2)\Pr(\theta_2) \end{aligned}$$

$$= (0.1 \times 0.4) + (0.4 \times 0.4) + (0.2 \times 0.6)$$

$$= 0.32.$$

Therefore, in repeated use of this rule, around a third of the diagnoses will be wrong.

Example 1.4

A student sits a multiple choice exam in which there are m alternative answers to each question.

The student either knows the answer (with probability θ) or guesses randomly (with probability $1 - \theta$).

What is the probability that the student actually knew the answer to a question they answered correctly?

Solution to Example 1.4 (1/2)

Let

C = student answers question correctly

K = student knows answer.

We require $\Pr(K|C)$. From the question we have

- $\Pr(K) = \theta$
- $\Pr(K^c) = 1 - \theta$
- $\Pr(C|K) = 1$ and
- $\Pr(C|K^c) = \frac{1}{m}$.

Solution to Example 1.4 (1/2)

Let

C = student answers question correctly

K = student knows answer.

We require $\Pr(K|C)$. From the question we have

- $\Pr(K) = \theta$
- $\Pr(K^c) = 1 - \theta$
- $\Pr(C|K) = 1$ and
- $\Pr(C|K^c) = \frac{1}{m}$.

Solution to Example 1.4 (1/2)

Let

C = student answers question correctly

K = student knows answer.

We require $\Pr(K|C)$. From the question we have

- $\Pr(K) = \theta$
- $\Pr(K^c) = 1 - \theta$
- $\Pr(C|K) = 1$ and
- $\Pr(C|K^c) = \frac{1}{m}$.

Solution to Example 1.4 (1/2)

Let

C = student answers question correctly

K = student knows answer.

We require $\Pr(K|C)$. From the question we have

- $\Pr(K) = \theta$
- $\Pr(K^c) = 1 - \theta$
- $\Pr(C|K) = 1$ and
- $\Pr(C|K^c) = \frac{1}{m}$.

Solution to Example 1.4 (1/2)

Let

C = student answers question correctly

K = student knows answer.

We require $\Pr(K|C)$. From the question we have

- $\Pr(K) = \theta$
- $\Pr(K^c) = 1 - \theta$
- $\Pr(C|K) = 1$ and
- $\Pr(C|K^c) = \frac{1}{m}$.

Solution to Example 1.4 (2/2)

Also, K and K^c form a partition of $\mathcal{S}$. Therefore, by Bayes' Theorem,

$$\begin{aligned}\Pr(K|C) &= \frac{\Pr(C|K)\Pr(K)}{\Pr(C|K)\Pr(K) + \Pr(C|K^c)\Pr(K^c)} \\&= \frac{1 \times \theta}{1 \times \theta + \frac{1}{m} \times (1 - \theta)} \\&= \frac{\theta}{\frac{m\theta + (1 - \theta)}{m}} \\&= \frac{m\theta}{1 + (m - 1)\theta}.\end{aligned}$$

Solution to Example 1.4 (2/2)

Also, K and K^c form a partition of $\mathcal{S}$. Therefore, by Bayes' Theorem,

$$\begin{aligned}\Pr(K|C) &= \frac{\Pr(C|K)\Pr(K)}{\Pr(C|K)\Pr(K) + \Pr(C|K^c)\Pr(K^c)} \\&= \frac{1 \times \theta}{1 \times \theta + \frac{1}{m} \times (1 - \theta)} \\&= \frac{\theta}{\frac{m\theta + (1 - \theta)}{m}} \\&= \frac{m\theta}{1 + (m - 1)\theta}.\end{aligned}$$

Solution to Example 1.4 (2/2)

Also, K and K^c form a partition of $\mathcal{S}$. Therefore, by Bayes' Theorem,

$$\begin{aligned}\Pr(K|C) &= \frac{\Pr(C|K)\Pr(K)}{\Pr(C|K)\Pr(K) + \Pr(C|K^c)\Pr(K^c)} \\&= \frac{1 \times \theta}{1 \times \theta + \frac{1}{m} \times (1 - \theta)} \\&= \frac{\theta}{\frac{m\theta + (1 - \theta)}{m}} \\&= \frac{m\theta}{1 + (m - 1)\theta}.\end{aligned}$$

Solution to Example 1.4 (2/2)

Also, K and K^c form a partition of $\mathcal{S}$. Therefore, by Bayes' Theorem,

$$\begin{aligned}\Pr(K|C) &= \frac{\Pr(C|K)\Pr(K)}{\Pr(C|K)\Pr(K) + \Pr(C|K^c)\Pr(K^c)} \\&= \frac{1 \times \theta}{1 \times \theta + \frac{1}{m} \times (1 - \theta)} \\&= \frac{\theta}{\frac{m\theta + (1 - \theta)}{m}} \\&= \frac{m\theta}{1 + (m - 1)\theta}.\end{aligned}$$

Solution to Example 1.4 (2/2)

Also, K and K^c form a partition of $\mathcal{S}$. Therefore, by Bayes' Theorem,

$$\begin{aligned}\Pr(K|C) &= \frac{\Pr(C|K)\Pr(K)}{\Pr(C|K)\Pr(K) + \Pr(C|K^c)\Pr(K^c)} \\&= \frac{1 \times \theta}{1 \times \theta + \frac{1}{m} \times (1 - \theta)} \\&= \frac{\theta}{\frac{m\theta + (1 - \theta)}{m}} \\&= \frac{m\theta}{1 + (m - 1)\theta}.\end{aligned}$$

Example 1.4

Suppose that there are $m = 5$ alternative answers for each question. Then

$$\Pr(K|C) = \frac{5\theta}{1 + 4\theta}.$$

We can see the effect of observing a correct answer on our belief that the student actually knows the answer by calculating $\Pr(K|C)$ for various θ – see Table 1.4.

Example 1.4

Suppose that there are $m = 5$ alternative answers for each question. Then

$$\Pr(K|C) = \frac{5\theta}{1 + 4\theta}.$$

We can see the effect of observing a correct answer on our belief that the student actually knows the answer by calculating $\Pr(K|C)$ for various θ – see Table 1.4.

Example 1.4

Suppose that there are $m = 5$ alternative answers for each question. Then

$$\Pr(K|C) = \frac{5\theta}{1 + 4\theta}.$$

We can see the effect of observing a correct answer on our belief that the student actually knows the answer by calculating $\Pr(K|C)$ for various θ – see Table 1.4.

Example 1.4

$\Pr(K)$ $= \theta$	$\Pr(K C)$ $= 5\theta/(1 + 4\theta)$
0.0	0.000
0.1	0.357
0.2	0.556
0.3	0.682
0.4	0.769
0.5	0.833
0.6	0.882
0.7	0.921
0.8	0.952
0.9	0.978
1.0	1.000

Example 1.4

