Exercise 3.1

A die is selected at random from two twenty-faced dice on which the symbols 1–10 are written with nonuniform frequency as follows.

Symbol	1	2	3	4	5	6	7	8	9	10
Number of faces of die A	6	4	3	2	1	1	1	1	1	0
Number of faces of die B	3	3	2	2	2	2	2	2	1	1

The randomly chosen die is rolled 7 times, with the following outcomes:

$$5, 3, 9, 3, 8, 4, 7.$$

What is the probability that the die is die A?

Solution:

$$\begin{aligned}
P(S|A) &= \frac{1}{20} \times \frac{3}{20} \times \frac{1}{20} \times \frac{3}{20} \times \frac{1}{20} \times \frac{2}{20} \times \frac{1}{20} \\
&= \frac{18}{20^7} \\
P(S|B) &= \frac{2}{20} \times \frac{2}{20} \times \frac{1}{20} \times \frac{2}{20} \times \frac{2}{20} \times \frac{2}{20} \times \frac{2}{20} \\
&= \frac{64}{20^7} \\
P(A|S) &= \frac{\frac{18}{20^7} \times \frac{1}{2}}{\frac{18}{20^7} \times \frac{1}{2} + \frac{64}{20^7} \times \frac{1}{2}} \\
&= \frac{9}{41}
\end{aligned}$$

Exercise 3.2

Assume that there is a third twenty-faced die, die C, on which the symbols 1–20 are written once each. As above, one of the three dice is selected at random and rolled 7 times, giving the outcomes:

$$3, 5, 4, 8, 3, 9, 7.$$

What is the probability that the die is (a) die A, (b) die B, (c) die C?

Solution:

$$
\begin{aligned}
P(S|A) &= \frac{3 \times 1 \times 2 \times 1 \times 3 \times 1 \times 1}{20^7} = \frac{18}{20^7} \\
P(S|B) &= \frac{2 \times 2 \times 2 \times 2 \times 2 \times 1 \times 2}{20^7} = \frac{64}{20^7} \\
P(S|C) &= \frac{1}{20^7}
\end{aligned}
$$

(i)

$$P(A|S) = \frac{18}{18 + 64 + 1} = \frac{18}{83}$$

(ii)

$$P(B|S) = \frac{64}{18 + 64 + 1} = \frac{64}{83}$$

(iii)

$$P(C|S) = \frac{1}{18 + 64 + 1} = \frac{1}{83}$$

Exercise 3.3

Inferring a decay constant
Unstable particles are emitted from a source and decay at a distance x, a real number that has an exponential probability distribution with characteristic length λ. Decay events can be observed only if they occur in a window extending from $x = 1\,\text{cm}$ to $x = 20\,\text{cm}$. N decays are observed at locations $\{x_1, \ldots, x_N\}$. What is λ?

Solution:

$$
\begin{aligned}
P(\lambda|\{x_1, \ldots, x_N\}) &= \frac{P(\{x_1, \ldots, x_N\}|\lambda)P(\lambda)}{P(\{x_1, \ldots, x_N\})} \\
&= \frac{P(x_1|\lambda)P(x_2|\lambda)\ldots P(x_N|\lambda)P(\lambda)}{P(\{x_1, \ldots, x_N\})}
\end{aligned}
$$

Now,

$$P(x|\lambda) \begin{cases} \frac{1}{\lambda} e^{-\frac{x}{\lambda}} & \text{if } 1 \leq x \leq 20 \\ 0 & \text{otherwise} \end{cases}$$

and if we let

$$
\begin{aligned}
Z(\lambda) - \int_1^{20} \frac{1}{\lambda} e^{-\frac{x}{\lambda}} dx \\
&= \left[-e^{-\frac{x}{\lambda}}\right]_1^2 0 \\
&= e^{-\frac{1}{\lambda}} - e^{-\frac{20}{\lambda}}
\end{aligned}
$$

then

$$
\begin{aligned}
P(\lambda|\{x_1, \ldots, x_N\}) &= \frac{\frac{1}{\lambda} e^{-\frac{x_1}{\lambda}} \times \frac{1}{\lambda} e^{-\frac{x_2}{\lambda}} \times \cdots \times \frac{1}{\lambda} e^{-\frac{x_N}{\lambda}}}{(Z(\lambda))^N} P(\lambda) \\
&= \frac{e^{-\Sigma_1^N \frac{x_N}{\lambda}} P(\lambda)}{(\lambda Z(\lambda))^N}
\end{aligned}
$$

Exercise 3.4

Forensic evidence

Two people have left traces of their own blood at the scene of a crime. A suspect, Oliver, is tested and found to have type ‘O’ blood. The blood groups of the two traces are found to be of type ‘O’ (a common type in the local population, having frequency 60%) and of type ‘AB’ (a rare type, with frequency 1%). Do these data (type ‘O’ and ‘AB’ blood were found at scene) give evidence in favour of the proposition that Oliver was one of the two people present at the crime?

Solution:

Let O be the case that Oliver is the criminal, and e the evidence.

$$
\begin{aligned}
P(O|e) &= \frac{P(e|O)P(O)}{P(e)} \\
&= \frac{0.01P(o)}{2 \times 0.01 \times 0.6} \\
&= \frac{P(O)}{1.2}
\end{aligned}
$$

so no, the posterior is lower than the prior, so the evidence is not in favour of the proposition that Oliver was one of the two people present at the crime.

Exercise 3.5

Sketch the posterior probability Sketch the posterior probability $P(p_\mathrm{a} \mid \mathbf{s} = \texttt{aba}, F = 3)$.
What is the most probable value of p_a (i.e., the value that maximizes the posterior probability density)? What is the mean value of p_a under this distribution?

Solution:

Using (3.11) and (3.12) from the textbook,

$$
\begin{aligned}
P(p_\mathrm{a} \mid \mathbf{s} = \texttt{aba}, F = 3) &= \frac{p_a^2(1-p_a)}{\frac{2!}{4!}} = 12p_a^2(1-p_a) \\
P(p_\mathrm{a} \mid \mathbf{s} = \texttt{bbb}, F = 3) &= \frac{(1-p_a)^3}{\frac{3!}{4!}} = 4(1-p_a)^3
\end{aligned}
$$

The graph is available in the following link: www.desmos.com/calculator/falltenjzo

Exercise 3.6

Show that after F tosses have taken place, the biggest value that the log evidence ratio

$$\log \frac{P(\mathbf{s}|F,\mathcal{H}_1)}{P(\mathbf{s}|F,\mathcal{H}_0)} \tag{3.23}$$

can have scales *linearly* with F if $\mathcal{H}_1$ is more probable, but the log evidence in favour of $\mathcal{H}_0$ can grow at most as $\log F$.

Solution:

From (3.12) and (3.20) in textbook,

$$
\begin{aligned}
\log \frac{P(\mathbf{s}|F,\mathcal{H}_1)}{P(\mathbf{s}|F,\mathcal{H}_0)} &= \log \frac{F_a!F_b!}{(F_a+F_b+1)!} - \log p_0^{F_a}(1-p_0)^{F_b} \\
&= \log F_a! + \log F_b! - \log(F_a+F_b+1)! - F_a \log p_0 - F_b \log(1-p_0) \\
&\approx F_a \log F_a - F_a + \frac{1}{2}\log(2\pi F_a) + F_b \log F_b - F_b + \frac{1}{2}\log(2\pi F_b) \\
&\quad - F\log F + F - \frac{1}{2}\log(2\pi F) \\
&\quad (\because \text{Stirling's approximation and } F \approx F_a + F_b + 1 \text{ for large F}) \\
&= F_a \log\frac{F_a}{p_0} + F_b \log\frac{F_b}{1-p_0} - F\log F + \frac{1}{2}\log(8\pi^3 \frac{F_aF_b}{F})
\end{aligned}
$$

Now, let $f_a = \frac{F_a}{F}$, $f_b = \frac{F_b}{F}$ then

$$\begin{aligned} & F_a \log \frac{F_a}{p_0} + F_b \log \frac{F_b}{1-p_0} - F \log F + \frac{1}{2} \log(8\pi^3 \frac{F_a F_b}{F}) \\ =& F\left(f_a \log \frac{f_a}{p_0} + f_b \log \frac{f_b}{1-p_0} \right) + \frac{1}{2} \log(8\pi^3 F f_a f_b) \end{aligned}$$

When $\mathcal{H}_1$ is more probable, $f_a \neq p_0$ so $f_a \log \frac{f_a}{p_0} + f_b \log \frac{f_b}{1-p_0}$ becomes a constant, hence log evidence ratio scales linearly with F.
When $\mathcal{H}_0$ is more probable, $f_a \to p_0$ so $f_a \log \frac{f_a}{p_0} + f_b \log \frac{f_b}{1-p_0}$ becomes 0, and the second term will dominate. f_a, f_b are constant hence log evidence ratio grows at most as $\log F$.

Note $f_a \log \frac{f_a}{p_0} + f_b \log \frac{f_b}{1-p_0} = KL(Q||P)$ where Q is the observed distribution and P is the null hypothesis distribution!

Exercise

Solution: