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:

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

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:

$$P(S|A) = \frac{3 \times 1 \times 2 \times 1 \times 3 \times 1 \times 1}{20^7} = \frac{18}{20^7}$$

$$P(S|B) = \frac{2 \times 2 \times 2 \times 2 \times 2 \times 1 \times 2}{20^7} = \frac{64}{20^7}$$

$$P(S|C) = \frac{1}{20^7}$$

(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 $x = 1\,\mathrm{cm}$ to $x = 20\,\mathrm{cm}$. N decays are observed at locations $\{x_1, \ldots, x_N\}$. What is λ?

Solution:

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

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

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

then

$$P(\lambda | \{x_1, \ldots, x_N\}) = \frac{\frac{1}{\lambda} e^{-\frac{x_1}{\lambda}} \times \frac{1}{\lambda} e^{-\frac{x_2}{\lambda}} \times \cdots \times \frac{1}{\lambda} e^{-\frac{x_N}{\lambda}}}{(Z(\lambda))^N} P(\lambda)$$
$$= \frac{e^{-\sum_1^N \frac{x_N}{\lambda}} P(\lambda)}{(\lambda Z(\lambda))^N}$$

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.

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

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 $P(p_a \mid \mathbf{s} = \mathtt{aba}, F = 3)$.
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,

$$P(p_a \mid \mathbf{s} = \mathtt{aba}, F = 3) = \frac{p_a^2(1 - p_a)}{\frac{2!}{4!}} = 12p_a^2(1 - p_a)$$

$$P(p_a \mid \mathbf{s} = \mathtt{bbb}, F = 3) = \frac{(1 - p_a)^3}{\frac{3!}{4!}} = 4(1 - p_a)^3$$

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

The most probable value is

$$\frac{d}{dp_a}(12p_a^2(1 - p_a)) = 0$$

$$\Longleftrightarrow 2p_a - 3p_a^2 = 0$$

$$\Longleftrightarrow p_a = \frac{2}{3} \quad \text{(check second derivative and boundary to verify this is the maximimum)}$$

$$\frac{d}{dp_a}(4(1 - p_a)^3) = 0$$

$$\Longleftrightarrow 3(1 - p_a)^2 = 0$$

$$\Longleftrightarrow p_a = 1$$

By checking boundary, $P(0) = 4, P(1) = 0$ so maximum at $p_a = 0$

The mean value is

$$\int_0^1 (12p_a^2(1 - p_a))p_a dp_a = \frac{3}{5}$$

$$\int_0^1 (4(1 - p_a)^3)p_a dp_a = \frac{1}{5}$$

4

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,

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

$$= \log \frac{F_a! F_b!}{(F_a + F_b + 1)!} - \log p_0^{F_a}(1 - p_0)^{F_b}$$

$$= \log F_a! + \log F_b! - \log(F_a + F_b + 1)! - F_a \log p_0 - F_b \log(1 - p_0)$$

$$= \log F_a! + \log F_b! - \log(F + 1) - \log F! - F_a \log p_0 - F_b \log(1 - p_0) \quad (\because F = F_a + F_b)$$

$$\approx F_a \log F_a - F_a + \frac{1}{2}\log(2\pi F_a) + F_b \log F_b - F_b + \frac{1}{2}\log(2\pi F_b) - \log F$$

$$\quad - F \log F + F - \frac{1}{2}\log(2\pi F) \quad (\because \text{Stirling's approximation and } F \approx F + 1 \text{ for large F})$$

$$= F_a \log \frac{F_a}{p_0} + F_b \log \frac{F_b}{1 - p_0} - F \log F + \frac{1}{2}\log(2\pi \frac{F_a F_b}{F})$$

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

$$F_a \log \frac{F_a}{p_0} + F_b \log \frac{F_b}{1 - p_0} - F \log F - \log F + \frac{1}{2}\log(2\pi \frac{F_a F_b}{F})$$

$$= F \left(f_a \log \frac{f_a}{p_0} + f_b \log \frac{f_b}{1 - p_0} \right) - \frac{1}{2}\log F + \frac{1}{2}\log(2\pi f_a f_b)$$

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. f_a, f_b are constant and can be ignored so the second term will dominate. The log evidence ratio grows at most as $\log F$ (negatively).

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 3.7

Putting your sampling theory hat on, assuming F_a has not yet been measured, compute a plausible range that the log evidence ratio might lie in, as a function of F and the true value of p_a, and sketch it as a function of F for $p_a = p_0 = 1/6$, $p_a = 0.25$, and $p_a = 1/2$. [Hint: sketch the log evidence as a function of the random variable F_a and work out the mean and standard deviation of F_a.]

Solution:

From 3.6, for $p_a = p_0 = 1/6$, the curve will asymptotically become a negative log curve, while with $p_a = 0.25$, and $p_a = 1/2$ the graph will asymptotically be a straight line with $p_a = 1/2$ having a steeper gradient.

Exercise 3.8

The three doors, normal rules.
On a game show, a contestant is told the rules as follows:

> There are three doors, labelled 1, 2, 3. A single prize has been hidden behind one of them. You get to select one door. Initially your chosen door will *not* be opened. Instead, the gameshow host will open one of the other two doors, and *he will do so in such a way as not to reveal the prize*. For example, if you first choose door 1, he will then open one of doors 2 and 3, and it is guaranteed that he will choose which one to open so that the prize will not be revealed.
>
> At this point, you will be given a fresh choice of door: you can either stick with your first choice, or you can switch to the other closed door. All the doors will then be opened and you will receive whatever is behind your final choice of door.

Imagine that the contestant chooses door 1 first; then the gameshow host opens door 3, revealing nothing behind the door, as promised. Should the contestant (a) stick with door 1, or (b) switch to door 2, or (c) does it make no difference?

Solution:

The contestant should (b) switch to door 2. let "this" be the event that the prize was in the first choice of door, "other" be the event that the prize was in the other unrevelead door, "evi" the revealed door being that specific door. Comparin the likelihood ratio.

$$\frac{P(\text{evi}|\text{this})}{P(\text{evi}|\text{other})} = \frac{\frac{1}{2}}{1} = \frac{1}{2}$$

so there is a double chance that the prize is in the other door.

> ### Exercise 3.9
>
> **The three doors, earthquake scenario.**
> Imagine that the game happens again and just as the gameshow host is about to open one of the doors a violent earthquake rattles the building and one of the three doors flies open. It happens to be door 3, and it happens not to have the prize behind it. The contestant had initially chosen door 1. Repositioning his toupée, the host suggests, 'OK, since you chose door 1 initially, door 3 is a valid door for me to open, according to the rules of the game; I'll let door 3 stay open. Let's carry on as if nothing happened.'
> Should the contestant stick with door 1, or switch to door 2, or does it make no difference? Assume that the prize was placed randomly, that the gameshow host does not know where it is, and that the door flew open because its latch was broken by the earthquake.
> [A similar alternative scenario is a gameshow whose *confused host* forgets the rules, and where the prize is, and opens one of the unchosen doors at random. He opens door 3, and the prize is not revealed. Should the contestant choose what's behind door 1 or door 2? Does the optimal decision for the contestant depend on the contestant's beliefs about whether the gameshow host is confused or not?]

Solution:

The likelihood ratio will be

$$\frac{P(\text{evi}|\text{this})}{P(\text{evi}|\text{other})} = 1$$

with the assumption that all doors have equal probability of breaking (symmetry) and independence from the player decision, so there will be no difference.

Exercise 3.10

Another example in which the emphasis is not on priors. You visit a family whose three children are all at the local school. You don't know anything about the sexes of the children. While walking clumsily round the home, you stumble through one of the three unlabelled bedroom doors that you know belong, one each, to the three children, and find that the bedroom contains girlie stuff in sufficient quantities to convince you that the child who lives in that bedroom is a girl. Later, you sneak a look at a letter addressed to the parents, which reads 'From the Headmaster: we are sending this letter to all parents who have male children at the school to inform them about the following boyish matters...'.

These two sources of evidence establish that at least one of the three children is a girl, and that at least one of the children is a boy. What are the probabilities that there are (a) two girls and one boy; (b) two boys and one girl?

Solution:

The bedroom evidence is a sampling evidence, while the letter evidence is a logical evidence. So,

$$P(GGB|\text{evi}) = \frac{P(\text{evi}|GGB)P(GGB)}{P(evi)}$$

$$P(BBG|\text{evi}) = \frac{P(\text{evi}|BBG)P(BBG)}{P(evi)}$$

We can simplify this to use the ratio of the two and then normalize it. $P(GGB) = P(BBG)$ so

$$P(GGB|\text{evi}) \propto P(\text{evi}|GGB) = \frac{2}{3}$$

$$P(BBG|\text{evi}) \propto P(\text{evi}|BBG) = \frac{1}{3}$$

This is already normalized so

$$P(GGB|\text{evi}) = \frac{2}{3}$$

$$P(BBG|\text{evi}) = \frac{1}{3}$$

Exercise 3.11

Mrs S is found stabbed in her family garden. Mr S behaves strangely after her death and is considered as a suspect. On investigation of police and social records it is found that Mr S had beaten up his wife on at least nine previous occasions. The prosecution advances this data as evidence in favour of the hypothesis that Mr S is guilty of the murder. 'Ah no,' says Mr S's highly paid lawyer, 'statistically, only one in a thousand wife-beaters actually goes on to murder his wife.[1] So the wife-beating is not strong evidence at all. In fact, given the wife-beating evidence alone, it's extremely unlikely that he would be the murderer of his wife – only a 1/1000 chance. You should therefore find him innocent.'

Is the lawyer right to imply that the history of wife-beating does not point to Mr S's being the murderer? Or is the lawyer a slimy trickster? If the latter, what is wrong with his argument?

Solution:

$$P(S_{\text{guilty}}|\text{beaten}, \text{murdered}) = \frac{P(\text{beaten}|S_{\text{guilty}}, \text{murdered})P(S_{\text{guilty}})}{P(\text{beaten}|\text{murdered})}$$

We must compare $P(\text{beaten}|S_{\text{guilty}}, \text{murdered})$ against $P(\text{beaten}|\text{murdered})$ to really conclude how the evidence affects how likely Mr S is guilty.

In fact, $P(\text{beaten}|S_{\text{guilty}}, \text{murdered})$ will likely be higher, so the lawyer is a slimy trickster. In other words, he is stating the probability of a wife-beater murdering the wife, whereas we need the probability of the wife-beating husband murdering the wife given the wife is murdered.

Exercise 3.12

A bag contains one counter, known to be either white or black. A white counter is put in, the bag is shaken, and a counter is drawn out, which proves to be white. What is now the chance of drawing a white counter? [Notice that the state of the bag, after the operations, is exactly identical to its state before.]

Solution:

$$P(W_2|W_1) = P(WW|W_1)$$
$$= \frac{P(W_1|WW)P(WW)}{P(W_1)}$$
$$= \frac{P(W_1|WW)P(WW)}{P(W_1|WW)P(WW) + P(W_1|WB)P(WB)}$$
$$= \frac{1 \times \frac{1}{2}}{1 \times \frac{1}{2} + \frac{1}{2} \times \frac{1}{2}}$$
$$= \frac{2}{3}$$

This is due to survivor bias.

Exercise 3.13

You move into a new house; the phone is connected, and you're pretty sure that the phone number is 740511, but not as sure as you would like to be. As an experiment, you pick up the phone and dial 740511; you obtain a 'busy' signal. Are you now more sure of your phone number? If so, how much?

Solution:

Let N be that the number is indeed correct, and E the evidence (the phone was busy).

$$P(N|E) = \frac{P(E|N)P(N)}{P(E)}$$
$$= \frac{P(N)}{P(E)}$$

Now, $P(E) = P(E|N)P(N) + P(E|\bar{N})P(\bar{N})$. If we say we think $P(N) = 0.9$ and $P(E|\bar{N}) = \frac{1}{1000}$ (considering unused phone numbers and phones that are not busy), then we have

$$P(E) = 0.9 + \frac{0.1}{10000} = 0.90001$$

so we now have confidence of 0.9999.

Exercise 3.14

In a game, two coins are tossed. If either of the coins comes up heads, you have won a prize. To claim the prize, you must point to one of your coins that is a head and say 'look, that coin's a head, I've won'. You watch Fred play the game. He tosses the two coins, and he

---

Solution:

Let D be the case Fred has double coins head, and W the case Fred won the game.

$$P(D|W) = \frac{P(W|D)P(D)}{P(W)} = \frac{1 \cdot \frac{1}{4}}{\frac{3}{4}} = \frac{1}{3}$$

Exercise 3.15

A statistical statement appeared in *The Guardian* on Friday January 4, 2002:

> When spun on edge 250 times, a Belgian one-euro coin came up heads 140 times and tails 110. 'It looks very suspicious to me', said Barry Blight, a statistics lecturer at the London School of Economics. 'If the coin were unbiased the chance of getting a result as extreme as that would be less than 7%'.

But *do* these data give evidence that the coin is biased rather than fair? [Hint: see equation (3.22).]

Solution:

If we assume the prior to be uniform,

$$\frac{P(H_1|s, F)}{P(H_0|s, F)} = \frac{140!110!}{251!} \Big/ \frac{1}{2}^{250} \quad (\because (3.22))$$

$$= 0.477$$

which still favours the unbiased case.
Even if we assume the extreme case, where H_1 prior to be 1 at exactly $f = \frac{140}{250}$, we get

$$\frac{P(H_1|s, F)}{P(H_0|s, F)} = 2^{250} f^{140} (1-f)^{110} = 6.08$$

where the hypothesis that there is bias is only six times more likely than the unbiased case.