

## Elementary hypothesis testing

← Back to Chapters

### Exercise 4.1

Given

$$P(D_1 \dots D_m | H_i X) = \prod_j P(D_j | H_i X)$$

and

$$P(D_1 \dots D_m | \overline{H}_i X) = \prod_j P(D_j | \overline{H}_i X)$$

for  $1 \leq i \leq n$  and  $n > 2$  show that for any fixed  $i$  at most one of

$$\frac{P(D_j | H_i X)}{P(D_j | \overline{H}_i X)}$$

is not equal to 1.

Proof:

Firstly, we claim that the case of 2 pieces of data implies the general result.

Indeed, independence assumptions of all  $D_j$  together imply analogous pairwise independence for any pair  $D_k$  and  $D_l$ . So, assuming the case with two data pieces is solved, we have, for a fixed  $i$  and for any pair of  $k, l$ , either  $\frac{P(D_k | H_i X)}{P(D_k | \overline{H}_i X)} = 1$ , or  $\frac{P(D_l | H_i X)}{P(D_l | \overline{H}_i X)} = 1$  (or both), so of the whole set of  $\frac{P(D_j | H_i X)}{P(D_j | \overline{H}_i X)}$  at most one is not equal to 1, as wanted.

So it is enough to solve the case of only two data sets,  $D_1$  and  $D_2$ .

We will denote probability density/mass function of  $D_1$  under hypothesis  $H_i X$  by  $V_i$  and that of  $D_2$  by  $U_i$ . We will also denote probability density/mass function of  $D_1$  under hypothesis  $\overline{H}_i X$  by  $V_i^c$  and that of  $D_2$  by  $U_i^c$ , though we will not use these until the very end.

Remark 1: The proof actually works for arbitrary (not necessarily discrete or continuous) real-valued random variables, one just has to say that  $V$ s and  $U$ s are CDFs instead of PMFs/PDFs. The reason for that is, firstly, that independence of two random variables can be equivalently written as joint CDF being a product or as joint PMF/PDF being a product, and, secondly, that equality of CDFs is equivalent to equality of PMFs/PDFs. We work with PDFs/PMFs out of a strange esthetic choice, rather than for any other reason.

We will denote value of  $V_i$  at some  $D_1 = x_1$  by  $v_{i1}$  (and at  $D_1 = x_2$  by  $v_{i2}$ ). Similarly the values of  $U$  at  $D_2 = y_1$  will be denoted by  $u_{i1}$ .

We will use independence of  $D_1$  and  $D_2$  conditional on various hypotheses to prove independence under other hypotheses. Note that independence always means  $P(D_1 = x, D_2 = y|H) = P(D_1 = x|H)P(D_2 = y|H)$  and can be checked by checking this for (arbitrary) specific values.

With this in mind, we pick any pair of possible value pairs  $D_1 = x_1, D_2 = y_1$  and  $D_1 = x_2, D_2 = y_2$ , fixed from now on, and form vectors  $v_i = \begin{bmatrix} v_{i1} \\ v_{i2} \end{bmatrix}$  and  $u_i = \begin{bmatrix} u_{i1} \\ u_{i2} \end{bmatrix}$

Then independence of  $D_1$  and  $D_2$  (conditional on  $H_i X$ ) says that the joint distribution of  $P(D_1 = x, D_2 = y|H_i X)$  is a product of distributions of  $D_1$  and  $D_2$  and is given (at  $x = x_1, x_2$  and  $y = y_1, y_2$ ) by the matrix

$$\begin{aligned} M_i &= v_i u_i^T = \begin{bmatrix} v_{i1} \\ v_{i2} \end{bmatrix} \begin{bmatrix} u_{i1} & u_{i2} \end{bmatrix} = \\ &= \begin{bmatrix} v_{i1}u_{i1} & v_{i1}u_{i2} \\ v_{i2}u_{i1} & v_{i2}u_{i2} \end{bmatrix} \end{aligned}$$

It follows from this that the (joint) probability matrix of  $D_1 D_2$  (again, at  $x = x_1, x_2$  and  $y = y_1, y_2$ ) conditional on  $\overline{H_i} X$  is obtained by taking all the matrices of  $H_j$  with  $j \neq i$  weighing them by (prior) probabilities  $h_j$  of  $H_j$  and adding them (and then dividing by the sum of the weights, but this is an overall normalizing factor which will not be important for us). That is, the matrix is proportional to

$$\begin{aligned} \sum_{j \neq i} h_j M_j &= \sum_{j \neq i} h_j v_j u_j^T = \\ &= \begin{bmatrix} \sum_{j \neq i} h_j v_{j1} u_{j1} & \sum_{j \neq i} h_j v_{j1} u_{j2} \\ \sum_{j \neq i} h_j v_{j2} u_{j1} & \sum_{j \neq i} h_j v_{j2} u_{j2} \end{bmatrix} \end{aligned}$$

Now the assumption that  $D_1$  and  $D_2$  are independent conditional on  $\overline{H_i}$  means this matrix is also a product of “marginal” matrices of  $D_1|\overline{H_i} X$  and  $D_2|\overline{H_i} X$ , i.e. is of rank 1. This means that it has determinant 0.

**Three hypothesis.** Let's start with the case of only 3 hypothesis  $H_1, H_2, H_3$ . Start with  $i = 3$ .

####Lemma: A sum  $M$  of two rank 1 matrices  $M = h_1 v_1 u_1^T + h_2 v_2 u_2^T$  can only be rank 1 if either  $v_1$  and  $v_2$  are linearly dependent or  $u_1$  and  $u_2$  are linearly dependent.

A “conceptual” proof is as follows: Consider the image of  $M$ . Vectors  $M(u_1)$  and  $M(u_2)$  are both linear combinations of  $h_1 v_1$  and  $h_2 v_2$ , so, to get that the image of  $M$  is the span of  $v_1$  and  $v_2$ , it is enough that the matrix of coefficients  $G = \begin{pmatrix} u_1^T u_1 & u_1^T u_2 \\ u_2^T u_1 & u_2^T u_2 \end{pmatrix}$  is invertible. But  $G$  is the Gramian of  $u_1, u_2$  and is invertible precisely when  $u_1$  and  $u_2$  are linearly independent (its determinant is the square of the area of the parallelogram spanned by  $u_1$  and  $u_2$ , as you can easily verify). In that case (of independent  $u$ s), the rank of  $M$  is the dimension of the span of  $v_1, v_2$  and if  $v_1, v_2$  were independent, it would be 2. So if rank of  $M$  is below 2, then either  $u$ s or  $v$ s are dependent, as wanted.

Remark 2: Those familiar with tensors may realize that we use metric in which  $u_1$  and  $u_2$  are orthonormal (that’s the inverse of  $G$ ) to “raise and index” and go from a bilinear form encoded by  $M$  to a linear map, whose range is then the span of  $v$ s.

Remark 3: Alternatively, for those who don’t like linear algebra, a computational proof of Lemma 1 is as follows: A (non-zero) 2 by 2 matrix has rank one when its determinant is zero. Writing this out in our case we get:

$$\begin{aligned} & [h_1 v_{11} u_{11} + h_2 v_{21} u_{21}] [h_1 v_{12} u_{12} + h_2 v_{22} u_{22}] = \\ & [h_1 v_{11} u_{12} + h_2 v_{21} u_{22}] [h_1 v_{12} u_{11} + h_2 v_{22} u_{21}] \end{aligned}$$

Additively canceling  $h_1^2 v_{11} v_{12} u_{11} u_{12}$  and  $h_2^2 v_{21} v_{22} u_{21} u_{22}$  and then dividing by  $h_1 h_2$  we have

$$\begin{aligned} & v_{11} u_{11} v_{22} u_{22} + v_{21} u_{21} v_{12} u_{12} = \\ & v_{11} u_{12} v_{22} u_{21} + v_{21} u_{22} v_{12} u_{11} \end{aligned}$$

or

$$(v_{11} v_{22} - v_{12} v_{21})(u_{11} u_{22} - u_{12} u_{21}) = 0,$$

so either  $v_1$  and  $v_2$  are linearly dependent, or  $u_1$  and  $u_2$  are.

Continuing with the case of three hypothesis, recall  $v_i$  and  $u_i$  were likelihood/probability vectors of  $D_1$  taking values  $x_1, x_2$  and  $D_2$  taking values  $y_1, y_2$  (under hypothesis  $H_i X$ ).

Observe that a pair of non-zero functions  $V_1$  and  $V_2$  such that  $(V_1(x_1), V_1(x_2))$  is always proportional to  $(V_2(x_1), V_2(x_2))$  are “globally” proportional meaning  $V_1 = kV_2$  (take any  $x$  with  $V_2(x) \neq 0$  and make  $k = V_1(x)/V_2(x)$ ).

If distributions of  $D_1$  under  $H_1X$  and  $H_2X$  are different, then they are also not proportional. By the previous paragraph, this implies that there will be two values  $x_1$  and  $x_2$  giving unproportional  $v_1$  and  $v_2$ . Then for arbitrary pair of values  $y_1, y_2$  of  $D_2$  the corresponding vectors  $u_1$  and  $u_2$  are proportional, so, again, by the previous paragraph, the whole probability mass/density functions  $U_1$  and  $U_2$  of  $D_2$  under  $H_1X$  and  $H_2X$  are proportional, ergo equal.

So either  $V_1 = V_2$  or  $U_1 = U_2$ .

Now, in the same way as we just did for  $i = 3$ , from  $i = 1$  and  $i = 2$  we get that (either  $V_2 = V_3$  or  $U_2 = U_3$ ) and (either  $V_1 = V_3$  or  $U_1 = U_3$ ). Since we have 3 equalities and only two types of distributions ( $U$  and  $V$ ), either the  $V$ s are equal twice, and  $V_1 = V_2 = V_3$ , or the  $U$ s are (and  $U_1 = U_2 = U_3$ ). Correspondingly either  $D_1$  has same distribution under all 3 hypothesis, and then  $\frac{P(D_1|H_iX)}{P(D_1|\bar{H}_iX)}$  are all equal to 1, or  $D_2$  does (and then all  $\frac{P(D_2|H_iX)}{P(D_2|\bar{H}_iX)}$  are equal to 1). In either case, we get what we want.

This completes the case of 3 hypothesis.

**The general case.** To get the extension to more than 3 hypothesis we use the following approach. As we mentioned before, a 2 by 2 matrix is of rank at most 1 if its determinant is zero. So we need some an efficient way of telling when the determinant of 2 by 2 matrix is zero.

Remark 4: More generally, a matrix is of rank at most 1 if all 2 by 2 minors have determinant zero i.e. all  $M_{(i,j),(k,l)}^{\wedge 2} = M_{ik}M_{jl} - M_{il}M_{jk}$  are zero. In tensor analysis, these are the entries of the second exterior power  $M^{\wedge 2}$  of  $M$ . When dimension is 2 there is only one minor, and the  $M^{\wedge 2}$  is a scalar, equal to  $\det M$ . So in dimensions above 2, we can formulate everything that follows in terms of determinants.

We will use the following property of 2D determinants. If  $M$  and  $N$  are 2 by 2 matrices then

$$D(M, N) := \frac{1}{2}(\det(M + N) - \det M - \det N)$$

is symmetric and bilinear in  $M, N$ . This means

$$D(M, N) = D(N, M)$$

and

$$D(M_1 + M_2, N) = D(M_1, N) + D(M_2, N)$$

(and hence the same for second variable). Indeed, one computes

$$D(M, N) = \frac{1}{2}(M_{11}N_{22} + M_{22}N_{11} - M_{12}N_{21} - M_{21}N_{12})$$

and the resulting formula is linear in  $M$  and in  $N$ , i.e. bilinear.

Observe that  $D(M, M) = \det M$ . We then have, by induction on the number of summands,

$$\det(\sum M_i) = D(\sum M_i, \sum M_j) = \sum_{i,j} D(M_i, M_j)$$

Remark 5: We also have  $D(\lambda M, N) = \lambda D(M, N)$ , as usual in bilinearity, but we don't need this.

Remark 6: In higher (possibly) dimensions, and using tensor language, we are saying that taking second exterior power, which is quadratic in the matrix input, is a restriction of a symmetric bilinear operation (on two inputs),  $(M \wedge N)(\vec{a} \wedge \vec{b}) = \frac{1}{2}[(M\vec{a}) \wedge (N\vec{b}) + (N\vec{a}) \wedge (M\vec{b})]$

Now we can apply this to our problem. Let  $M_i = h_i v_i u_i^T$  and  $N_i = \sum_{j \neq i} M_j$ , and  $M = M_i + N_i = \sum_j M_j$ .

Our assumptions are that all  $M_i$  and  $N_j$  are rank 1 (i.e. have zero determinant). We now show that  $M$  has rank 1 (i.e. has zero determinant).

To that end we write

$$\det M = \sum_{j,k} D(M_j, M_k)$$

We want to see that this is zero. We know

$$0 = \det(N_i) = \sum_{j \neq i, k \neq i} D(M_j, M_k)$$

Summing over  $i$  we get (taking note that each  $D(M_l, M_l)$  will appear  $n - 1$  times, while those  $D(M_j, M_k)$  with  $j \neq k$  will appear only  $n - 2$  times):

$$\sum_l D(M_l, M_l) + (n - 2) \sum_{j,k} D(M_j, M_k) = 0$$

So, since  $D(M_l, M_l) = 0$ , as long as  $n \neq 2$  we have what we want.

This gives  $M = vu^T$ . Going back to  $M = M_i + N_i$  we again see two rank one matrices add up to a rank one matrix. We conclude, just as in the case of 3 hypothesis, that for each specific  $i$ , either  $V_i^c = V_i$  and hence  $\frac{P(D_1|H_i X)}{P(D_1|\bar{H}_i X)} = 1$

OR

$U_i^c = U_i$  and hence  $\frac{P(D_2|H_iX)}{P(D_2|\bar{H}_iX)} = 1$ . This is exactly what we wanted to prove.

#### Some remarks on sequential vs. batch updates.

(See formula 4.11; the discussion is in the context of section 4.4 and formula 4.44.)

We are considering a batch of widgets from a single machine. We have three hypothesis,  $A$  – the machine has failure rate  $\frac{1}{3}$  (prior probability  $\frac{1}{11}$ ),  $B$  – the machine has failure rate  $\frac{1}{6}$  (prior probability  $\frac{10}{11}$ ), and  $C$  – the machine has failure rate  $\frac{99}{100}$  (prior probability  $\frac{1}{10^6}$ ).

Prior odds ratio is then (to a very good approximation)  $\frac{10^6}{11} : \frac{10^7}{11} : 1$ .

The update rule for arbitrary data tells us how to compute the posterior odds  $A$  vs  $B$  vs  $C$ : one takes the prior odds and multiplies each number by the likelihood of data under corresponding hypothesis. For data of  $m$  bad widgets the likelihoods are  $\frac{1}{3^m}, \frac{1}{6^m}, \frac{99^m}{100^m}$ . This gives posterior odds:

$$\frac{10^6}{11} \frac{1}{3^m} : \frac{10^7}{11} \frac{1}{6^m} : \frac{1}{10^6} \frac{99^m}{100^m}.$$

Note that since the individual draws are independent **conditional on one specific hypothesis**, the likelihoods for  $m$  widgets are each a product of individual likelihoods of draws, and we see very clearly that the same posterior odds will be obtained by updating on all the data of  $m$  bad widgets, or by splitting  $m = m_1 + m_2$  and updating first to odds ratio

$$\frac{10^6}{11} \frac{1}{3^{m_1}} : \frac{10^7}{11} \frac{1}{6^{m_1}} : \frac{1}{10^6} \frac{99^{m_1}}{100^{m_1}},$$

using that as new prior odds and then updating to the final posterior odds

$$\frac{10^6}{11} \frac{1}{3^{m_1}} \frac{1}{3^{m_2}} : \frac{10^7}{11} \frac{1}{6^{m_1}} \frac{1}{6^{m_2}} : \frac{1}{10^6} \frac{99^{m_1}}{100^{m_1}} \frac{99^{m_2}}{100^{m_2}}$$

getting the same result.

It is only when merging  $A$  and  $B$  into a single hypothesis  $\bar{C} = A + B$  that we have trouble: individual draws are not independent under  $\bar{C}$ . To illustrate, consider  $m = 2$ . Firstly, the prior odds of  $C : \bar{C}$  are

$$C_0 : \bar{C} = C_0 : A_0 + B_0 = 1 : \frac{10^6}{11} + \frac{10^7}{11} = 1 : 10^6$$

The likelihood of the first widget being bad under  $C$  is as before  $\frac{99}{100}$ . The likelihood of it being bad under  $\bar{C} = A + B$  is trickier:  $B$  is 10 times more likely

than  $A$ , so the probability of a bad widget under  $A + B$  is weighted average of probabilities under  $A$  and  $B$  with the one under  $B$  getting ten times more weight. This is simply the “total probability formula”

$$\begin{aligned} P(D_1|A+B) &= P(D_1|A)P(A|A+B) + P(D_1|B)P(B|A+B) = \\ &= \frac{1}{3} \frac{1}{11} + \frac{1}{6} \frac{10}{11} = \frac{2}{11} \end{aligned}$$

After the first widget is drawn, the posterior odds are prior odds times likelihood:

$$C_1 : \bar{C}_1 = 1 \frac{99}{100} : 10^6 \frac{2}{11}$$

We can check it's the same result as one obtained by first updating the odds of  $A : B : C$  and then summing the ones for  $A$  and  $B$ :

$$C_1 : A_1 + B_1 = \frac{99^1}{100^1} : \frac{10^6}{11} \frac{1}{3^1} + \frac{10^7}{11} \frac{1}{6^1} = \frac{99}{100} : \frac{2 \cdot 10^6}{11}$$

Now **if** the likelihood of second bad widget under  $\bar{C}$  was again  $P(D_2|A+B) = \frac{2}{11}$  we would get the **WRONG**  $C_2 : \bar{C}_2 = \frac{99}{100} \frac{99}{100} : \frac{2 \cdot 10^6}{11} \frac{2}{11}$ .

However, it is not -  $D_2$  is not independent of  $D_1$  under  $\bar{C}$  - having learned that the first draw was defective we now think it is relatively more likely that the batch came from  $A$  rather than  $B$ , as follows. The odds ratios after  $D_1$  of  $A_1 : B_1 = 1\frac{1}{3} : \frac{1}{6}10 = 1 : 5$ , so  $P(A|(A+B)D_1) = \frac{1}{6}$  and  $P(B|(A+B)D_1) = \frac{5}{6}$ . This means that the second widget is more likely to come from  $A$  rather than  $B$  and is more likely to be bad:

$$\begin{aligned} P(D_2|(A+B)D_1) &= \\ P(D_2|AD_1)P(A|(A+B)D_1) + P(D_2|BD_1)P(B|(A+B)D_1) &= \\ = \frac{1}{3} \frac{1}{6} + \frac{1}{6} \frac{5}{6} = \frac{7}{36} > \frac{2}{11} \end{aligned}$$

This means that seeing the second bad widget is giving somewhat less evidence for  $C$  over  $\bar{C}$  than seeing the first one did (some evidence in  $D_2$  is “already” in  $D_1$ ).

We can now finish the computation for the second update:

$$C_2 : \overline{C}_2 = \frac{99}{100} \frac{99}{100} : \frac{2 \cdot 10^6}{11} \frac{7}{36} = \frac{99^2}{100^2} : \frac{14 \cdot 10^6}{11 \cdot 6^2}$$

and compare it to the “batch” calculation:

$$C_2 : A_2 + B_2 = \frac{99^2}{100^2} : \frac{10^6}{11} \frac{1}{3^2} + \frac{10^7}{11} \frac{1}{6^2} = \frac{99^2}{100^2} : \frac{14 \cdot 10^6}{11 \cdot 6^2}.$$

**Most of Exercise 4.2** Suppose prior odds of hypothesis  $H_i$  are  $s_i$ , and the probability of good widget under  $H_i$  is  $p_i$  and of a bad one is  $q_i = 1 - p_i$ .

For  $n_0$  bad widgets and  $n_1$  good ones the posterior odds are

$$\dots : L_i = s_i p_i^{n_0} q_i^{n_1} : \dots$$

If the fraction of bad widgets is  $f$  then we have  $n_0 = fn$ ,  $n_1 = (1 - f)n$  and posterior odds

$$\dots : L_i = s_i (p_i^{1-f})^n (q_i^f)^n : \dots$$

For two different hypothesis the ratio  $L_i : L_j$  does not go to zero or infinity precisely when  $p_i^{1-f} q_i^f = p_j^{1-f} q_j^f$ . Solving for  $f$  we get:

$$\frac{p_i}{p_j} = \left( \frac{q_j p_i}{q_i p_j} \right)^f$$

$$\ln \frac{p_i}{p_j} = f \ln \frac{q_j p_i}{q_i p_j}$$

$$f = \frac{\ln \frac{p_i}{p_j}}{\ln \frac{q_j}{q_i} + \ln \frac{p_i}{p_j}}$$

$$f = \frac{1}{\frac{\ln \frac{q_j}{p_j}}{\frac{q_i}{p_i}} + 1}$$

For the case in section 4.4 we have  $p_C = 1/100$ ,  $p_A = 2/3$ ,  $q_C = 99/100$ ,  $q_A = 1/3$ ,

$$f = \frac{1}{\frac{\ln \frac{q_j}{p_j}}{\frac{q_i}{p_i}} + 1} = \frac{1}{\frac{\ln \frac{297}{100}}{\ln \frac{200}{3}} + 1} = 0.7941552598429126932775507544006687123649252334748242$$

**Some of Exercise 4.3** We now consider the case of 4 hypothesis  $A, B, C$  as before and  $F$ . We remark that since we will study the evolution of our beliefs in the scenario where about 1/4 of widgets are defective, neither  $C$  nor the hypothetical  $E$  from the first part of this exercise under which most widgets are good are of any importance, so we can, in fact just study how  $F$  fares against  $A$  and  $B$ .



The likelihoods of data of  $n/4$  bad widgets and  $3n/4$  good ones under  $A : B : F$  are

$$\left[ \left( \frac{1}{3} \right)^{\frac{1}{4}} \left( \frac{2}{3} \right)^{\frac{3}{4}} \right]^n : \left[ \left( \frac{1}{6} \right)^{\frac{1}{4}} \left( \frac{5}{6} \right)^{\frac{3}{4}} \right]^n : \left[ \left( \frac{1}{4} \right)^{\frac{1}{4}} \left( \frac{3}{4} \right)^{\frac{3}{4}} \right]^n$$

$$0.5606^n : 0.5573^n : 0.5699^n$$

So the evidence for  $F$  is (in Jaynesian db)  $10(\log_{10} 0.5699^n - \log_{10}(0.5606^n + 0.5573^n))$  which for large  $n$  is  $n \cdot 10 \log_{10} \frac{0.5699}{0.5606} \approx n \cdot 0.0714556 \approx \frac{n}{14}$ . So we would need about 14 samples per decibel of evidence in favor of  $F$ , i.e. multiple hundreds before something like 20db of evidence for  $F$  over  $A$  and  $B$  can be collected.

**Exercise 4.4** This problem is simplified if we assume that  $A$  and  $B$  are the only hypothesis under consideration. Then evidence for  $A$  is evidence against  $B$  and vice versa. In addition, the individual widget tests are independent both under  $A = \overline{B}$  and  $B = \overline{A}$ .

So each negative test gives  $10 \log_{10} \frac{1/3}{1/6} = 3$  db evidence for  $A$ . If  $B$  is true, the probability of this is  $\frac{1}{6}$ .

And a positive gives  $10 \log_{10} \frac{2/3}{5/6} = -0.97$  db evidence against  $A$ . If  $B$  is true, the probability of this is  $\frac{5}{6}$ .

We now indeed have a random walk. We start at 0 db of evidence, and then at each step we either, with probability  $\frac{1}{6}$  take a step of size 3, or, with probability  $\frac{5}{6}$  take a step of size  $-0.97$ .

We now ask: 1) what is the probability of ever getting to 20 db? 2) how long will it take to get  $-20$  db?

This is an exercise in “Wald theory” a la A. Wald, “Sequential analysis” or W. Feller, “An introduction to probability theory and its applications”, Vol 1 Ch. XIV., Vol 2 Ch.XII (Abraham Wald is the statistician of the “holes you see in the returning bombers are in the areas where a bomber can take a hit and still come back, it’s all the other areas you should worry about, you dolts!” fame; someone really should make a movie about his life.)

One gets a simplification by rounding  $-0.97$  to  $-1$ , for then the possible amounts of evidence are all integers, and the random walk is on  $\mathbb{Z}$ . Moreover, the resulting random walk is “skip-free to the left” - one can not move left by more than one unit at a time. There are a more or less standard approaches for analyzing such a walk.

(Without simplifying exact solution becomes intractable, but one can get various approximations. We derive only very basic such approximation at the end this write-up, but it is discussed by Wald and in Vol II of Feller.)

**Probability of getting 20 db for A.** Let  $p_i$  be the probability that the walk will ever reach the state  $i$  decibel.

Firstly, the expected evidence is  $3\frac{1}{6} + (-1)\frac{5}{6} = -\frac{1}{3}$  is negative, so by the strong law of large numbers the walk goes off to  $-\infty$  with probability 1. Since the only way to reach it starting from 0 is to go through all negative values,  $p(i) = 1$  for all  $i \leq 0$ .

In addition, this means that, if we denote by  $X_j$  the set of all walks that reach state  $i$  but don't reach state  $i + 1$  then  $X = \cup_{j \geq 0} X_j$  has probability 1 (each walk that converges to  $-\infty$  will belong to one of the  $X_i$ ). This means that  $p_i = P(\cup_{j \geq i-1} X_j)$  converges to zero as  $i \rightarrow \infty$ .

The above two paragraphs give boundary conditions for  $p_i$ . Now let's find a relation between different  $p_i$ s (for positive  $i$ ). Consider the first step of the walk. Either, with probability  $\frac{1}{6}$  it is 3 to the right, and so what remains is to reach  $i - 3$  steps to the right, probability of which is  $p_{i-3}$ . Or, with probability  $\frac{5}{6}$  it is 1 to the left, and so now the walk needs to reach  $i + 1$  steps to the right, probability of which is  $p_{i+1}$ . So

$$p_i = \frac{1}{6}p_{i-3} + \frac{5}{6}p_{i+1}$$

The first few are  $p_1 = \frac{1}{6} + \frac{5}{6}p_2$ ,  $p_2 = \frac{1}{6} + \frac{5}{6}p_3$ ,  $p_3 = \frac{1}{6} + \frac{5}{6}p_4$ ,  $p_4 = \frac{1}{6}p_1 + \frac{5}{6}p_5$  etc.

Remark:  $\frac{1}{6}p_{-3} + \frac{5}{6}p_1 = \frac{1}{6} + \frac{5}{6}p_1$  is the probability of returning to 0 **after** the first step, and so in **not** equal to  $p_0 = 1$ .

There is a standard way to solve such recurrence relations, and even several variations on it.

**In a more direct way**, one starts with  $p_i = r^i$  to get characteristic equation

$$r^3 = \frac{1}{6} + \frac{5}{6}r^4.$$

This has roots  $r_1 = 1$ ,  $r_2 \approx 0.8$ ,  $r_{3,4} \approx -0.3 \pm 0.4i$  (so  $|r_{3,4}| \approx 0.5$ ). A general solution is a linear combination of these,  $p_i = c_1r_1^i + c_2r_2^i + c_3r_3^i + c_4r_4^i$ . The boundary condition  $\lim_{i \rightarrow \infty} p_i = 0$  means  $c_1 = 0$ , and the other conditions  $p_0 = p_{-1} = p_{-2} = 1$  give

$$\begin{pmatrix} 1 & 1 & 1 \\ r_2^{-1} & r_3^{-1} & r_4^{-1} \\ r_2^{-2} & r_3^{-2} & r_4^{-2} \end{pmatrix} \begin{pmatrix} c_2 \\ c_3 \\ c_4 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$$

$$\begin{pmatrix} c_2 \\ c_3 \\ c_4 \end{pmatrix} = \begin{pmatrix} 1 & 1 & 1 \\ r_2^{-1} & r_3^{-1} & r_4^{-1} \\ r_2^{-2} & r_3^{-2} & r_4^{-2} \end{pmatrix}^{-1} \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$$

I seem to get  $c_4 = 0.07492831 - 0.0259995i$ ,  $c_3 = 0.07492831 + 0.0259995i$ ,  $c_2 = 0.85014338$ . This means that we can ignore the smaller modulus complex roots, and write  $p_i \approx 0.85 \cdot (0.7823728)^i$ . For  $i = 20$  we have

$$p_{20} \approx 0.006276157.$$

This can probably be described as “negligibly small”.

**Alternatively**, another way (Feller, Chapter XI.1 and XI.4, particularly Example b) is to define  $Q(s) = \sum_{i=-2}^{\infty} p_i s^i$  and observe that starting from

$$p_i = \frac{1}{6}p_{i-3} + \frac{5}{6}p_{i+1}$$

multiplying by  $s^i$ , and summing over  $i \geq 1$  and summing, using  $p_{-2} = p_{-1} = p_0 = 1$  and  $p_1 = p$  gives:

$$Q(s) - 1 - s^{-1} - s^{-2} = \frac{1}{6}s^3Q(s) + \frac{5}{6}\left(\frac{1}{s}(Q(s) - sp - 1 - s^{-1} - s^{-2})\right)$$

$$\frac{5}{6}(ps^3 + s^2 + s + 1) - (s^3 + s^2 + s) = Q(s)s^2\left(\frac{1}{6}s^4 - s + \frac{5}{6}\right)$$

$$Q(s)s^2 = \frac{\frac{5}{6}(ps^3 + s^2 + s + 1) - (s^3 + s^2 + s)}{\frac{1}{6}s^4 - s + \frac{5}{6}}$$

This means that  $Q(s)s^2 = \frac{U(s)}{V(s)}$  is a rational function and has partial fraction expansion  $\sum_{j=1}^4 \frac{a_k}{s_k - s}$  with  $s_k$  the roots of  $V(s)$  (luckily in our case the roots of  $V$  are distinct) and

$$a_k = \frac{-U(s_k)}{V'(s_k)}$$

and

$$p_{i-2} = \frac{a_1}{s_1^{i+1}} + \frac{a_2}{s_2^{i+1}} + \frac{a_3}{s_3^{i+1}} + \frac{a_4}{s_4^{i+1}}.$$

We note that the roots  $s_k$  of  $V(s)$  are the inverses of the roots  $r_k$  in our previous “direct approach”, and the formula above is thus identical with the one we got before.

One of the roots of  $V(s)$  is 1, but the corresponding coefficient  $c_1$  must be 0 otherwise the limit  $p_i$  will not be zero. This gives  $U(1) = 0$  meaning  $\frac{5}{6}(p+3) - 3 = 0$ , or  $p = \frac{3}{5}$ .

The next smallest in absolute value root is  $s_2 \approx 1.278$ . The corresponding coefficient

$$a_2 = \frac{-U(s_2)}{V'(s_2)} \approx \frac{0.696}{1.225424} \approx 1.775215$$

giving

$$p_{20} \approx \frac{1.775215}{(1.2781)^{23}} \approx 0.0062772128$$

as before (within precision of calculations). We could compute  $a_3$  and  $a_4$  to make sure they are not huge and ignoring the other terms is justified, but we have already seen this in the previous method, so we omit this calculation.

**Remark:** The fact that one can use these two methods to compute the coefficients  $c_k$  in the expansion of  $p_i$  means that the inverse of the Vandermonde matrix is computable via partial fractions expansion. This has been noted in the literature.

**Time to get 20 db for  $B$ . For a shortcut skip to the bottom “Wald identities”.**

Since the the walk is skip-free on the left, to get 20 db for  $B$  one has to first get 19 db, and for that one needs to first get 18 db etc. So we start by getting a distribution of times to get 1 db for  $B$ . Let  $T$  be the (random) time until the walk hits  $-1$ . We define the formal power series

$$f(x) = E(x^T) = xP(T=1) + x^2P(T=2) + \dots$$

encoding the distribution of  $T$ . We have  $f(1) = 1$  and  $f(0) = 0$ . We also have  $f'(1) = E(T)$  and  $f''(1) = E(T(T-1))$  and so on.

Remark: This annoying  $T^2 - T$  as opposed to just  $T$  can be avoided by changing variables  $x = e^u$  (giving moment generating function of  $T$ ), or, even better,  $x = e^{iu}$  (giving characteristic function of  $T$ ). This is in some ways nicer, but we stick to generating function version and its polynomial look and feel.

Then the distribution of (random) time to hit  $-k$  is given by writing  $T_k = Z_1 + Z_2 + \dots + Z_k$  with  $Z_i$  i.i.d. with same distribution as  $T$ , so that the formal power series encoding the distribution of  $T_k$  is

$$f_k(x) = E(x^{T_k}) = E(x^{\sum Z_i}) = E(\prod x^{Z_i}) = \prod E(x^{Z_i}) = f(x)^k$$

Now consider  $T$  again, and condition on the first step of the walk. With probability  $\frac{5}{6}$  it is to the left and  $T = 1$ . With probability  $\frac{1}{6}$  it is to the right and then  $T$  is distributed as encoded by  $f_4(x) = xf(x)^4$  (the extra  $x$  is for the one extra time beat that the first step used). So overall  $T$ , being a mixture of these two possibilities, is distributed as  $\frac{5}{6}x + \frac{1}{6}f(x)^4$ . Hence

$$f(x) = \frac{5}{6}x + \frac{1}{6}xf(x)^4$$

One could write  $f(x) = c_1x + c_2x^2 + \dots$  and plug in to get

$$c_1x + c_2x^2 + \dots = \frac{5}{6}x + \frac{1}{6}x(c_1x + c_2x^2 + \dots)^4$$

which can be rewritten as a “upper triangular” system of equations to be solved one after another, getting  $c_1 = \frac{5}{6}, c_2 = c_3 = c_4 = 0, c_5 = \frac{1}{6}c_1^4 = \frac{5^4}{6^5}, c_6 = \frac{1}{6}4c_1^3c_2 = 0$  etc. Then, of course it is  $f(x)^{20}$  that gives the distribution of times to hit 20 db that we are after.

Abandoning the pursuit of full distribution of  $T_{20}$  we instead compute its first and second moments.

First we deal with the moments of  $T$ . Differentiating the equation for  $f$  we have

$$f' = \frac{5}{6} + \frac{1}{6}f^4 + \frac{1}{6}x4f'f^3$$

Since  $f(1) = 1$  we get  $f'(1) = 1 + \frac{4}{6}f'(1)$ ,  $f'(1) = 3$ .

We have  $E(T_{20}) = (f^{20})'(1) = 20f'(1)f(1)^{19} = 20 \times 3 = 60$ . This tells us the expected time to get 20 db for  $B$  is 60 samples.

We will now find the variance of this estimate.

Differentiating again

$$f'' = \frac{1}{6}4f'f^3 + \frac{1}{6}4f'f^3 + \frac{1}{6}x4f''f^3 + \frac{1}{6}x4f'3f'f^2$$

from which  $f''(1) = 2 + 2 + \frac{2}{3}f''(1) + 18$ ,  $f''(1) = 66$ . Thus  $E(T^2) = 69$  and  $Var(T) = E(T^2) - E(T)^2 = 60$ , and standard deviation is  $\sqrt{60} \approx 7.75$

Remark 1: This low expectation and high variance are due to the high probability of  $T = 1$ ; conditional on  $T \neq 1$  the expectation of  $T$  is  $(3 - \frac{5}{6})/\frac{1}{6} = 13$  (which coincides with the (unconditional) expectation of  $T_4$  plus 1, i.e.  $3 \cdot 4 + 1 = 13$ ). The conditional variance is the variance of  $T_4$  and is 240, so the standard deviation is  $\approx 15.5$  (see below). This is still indicating high probability of low values and a “fat tail”, but a bit less so than the unconditional case.

On the other hand

$$(f^{20})'' = (20f^{19}f')' = 20(19(f')^2f^{18} + f^{19}f'') = 20(19 \cdot 9 + 66) = 4740$$

so  $E(T_{20}^2) = 4740 + 60 = 4800$  and  $Var(T) = 4800 - 60^2 = 1200$ , and standard deviation  $\sqrt{1200} \approx 34.64$ , which is more reasonable. It indicates that we may get lucky and get 20 db evidence much sooner than 60 tests, but may get unlucky and have to wait substantially longer, maybe as long as 130 trials or even more. One could compute further moments using the same method (with a longer calculation) to get more of an idea.

For general  $k$  we get  $E(T_k^2) = k((k-1)9 + 66) + 3k$ ,  $Var(T_k) = 60k$

#### Alternative solution using Wald identities.

We have in general  $E(T_k) = 3k$ . Can be deduced much more quickly from the general First Wald Identity: if  $T$  is a stopping time with respect to i.i.d.  $Z_i$ s with  $E(T)$  finite, and  $X = \sum_{i=1}^T Z_i$  then

$$E(X) = E(T)E(Z_1).$$

In our case, assuming  $E(T_k)$  is finite, we have  $E_{Z_1} = 3\frac{1}{6} - \frac{5}{6} = -\frac{1}{3}$ ,  $X = -k$  (always), so  $E(T_k) = 3k$ .

We still need to check  $E(T_k) < \infty$ . This follows from (a generalization) of our analysis of the random walk on  $\mathbb{Z}$  but is also part of Theorem 2 Ch. XII.2 in Volume 2 of Feller’s book.

Wald’s second identity (under same assumption  $E(T)$  finite) is

$$E[(X - E(Z_i)T)^2] = Var(Z_i)E(T).$$

For us  $X = -k$ ,  $E(Z_i) = -\frac{1}{3}$ ,  $E(T_k) = 3k$  and

$$Var Z_i = \frac{1}{6}(3 + \frac{1}{3})^2 + \frac{5}{6}(-1 + \frac{1}{3})^2 = 100/54 + 20/54 = 120/54$$

Plugging in, this gives:

$$k^2 - 2(-k)(-1/3)3k + 1/9E(T_k^2) = 120/54 \cdot 3k$$

$$-9k^2 + E(T_k^2) = 60k$$

$$E(T_k^2) = 9k^2 + 60k$$

$$Var(T_k) = 9k^2 + 60k - 9k^2 = 60k$$

as before.

One advantage of using this method is that it can handle the unapproximated problem as well.

Indeed, if we keep the leftward step 0.97 instead of making it 1,  $E(Z_1)$  changes a bit to  $-\frac{1}{3} + \frac{5}{6}0.03 = \frac{-200+15}{600} = -\frac{37}{120}$ , and the  $E(X)$  is no longer  $-k$ , but is between  $-k$  and  $-k - 0.97$ , so we get an estimate  $\frac{120}{54}k \leq E(T_k) < \frac{120}{54}(k + 0.97)$ . Similar estimate can be found for the variance of  $T_k$  as well.