10. Area as limit of a function sequence. D'Arcy Thompson's mini minnows

In this lecture, we combine our knowledge of function sequences (Lecture 7) and index functions (Lectures 8-9) to calculate area as the limit of a function sequence. In doing so, we explore further aspects of continuous distributions.

First, recall from Lecture 8 that if f is the probability density function or p.d.f. of a continuous random variable X, e.g., leaf thickness, then

$$\text{Prob}(a \le X \le b) \quad = \quad \text{Area}(f, [a, b]) \qquad (10.1)$$

is the probability that X lies between a and b. It appears at first glance that, in order to calculate this probability, you need to know an ordinary function (namely, f) on the whole of subdomain [a, b]. In fact, however, to calculate the above probability it is necessary to know an ordinary function only at the ends of the subdomain — but the ordinary function in question is not the p.d.f. Rather, it is the **cumulative distribution function** or **c.d.f**., i.e., F defined on $[0, \infty)$ by

$$F(t) \quad = \quad \text{Prob}(0 \le X \le t) \quad = \quad \text{Area}(f, [0, t]). \qquad (10.2)$$

For (9.17) implies that Area(f, [0, a]) + Area(f, [a, b]) = Area(f, [0, b]) whenever $0 \le a \le b$, and a slight rearrangement of this equation yields

$$\begin{aligned} \text{Prob}(a \le X \le b) \quad &= \quad \text{Area}(f, [a, b]) \\ &= \quad \text{Area}(f, [0, b]) \; - \; \text{Area}(f, [0, a]) \\ &= \quad F(b) \; - \; F(a) \qquad (10.3) \end{aligned}$$

on using (1)-(2). So the c.d.f. is an important function for calculating probabilities associated with X, and we would like to broach the task of constructing it; but unimodal distributions, like that of leaf thickness in Figure 8.4, are not well suited to this task.

SIZE (mm) ABOVE BASE LENGTH	1-3	4-6	7-9	10-12	13-15	16-18	19-21	22-24	25-27
NUMBER	1	22	52	67	114	257	177	41	2
PROBABILITY	0.00	0.03	0.07	0.09	0.16	0.35	0.24	0.06	0.00

Table 10.1 Frequency and probability of lengths above 12 mm in Thompson's catch of 733 minnows

Instead, therefore, our point of departure is the observation that not every distribution in nature is unimodal. Among exceptions are conditional distributions of size. To fix ideas, we take a look at D'Arcy Thompson's data on size in minnows. Thompson (1942, pp. 133-134) measured a catch of 733 minnows in 3 mm groupings. No minnow exceeded 39 mm in length, and no minnow was less than 12 mm. Accordingly, it is convenient for our purpose to take 12mm as base length and to consider variation in excess of that. Size then varies between 0 and 27 mm. That is, the set of all possible sizes or **sample space** becomes [0, 27], and Thompson's data become those in Table 1. The corresponding distribution is modelled by the p.d.f. (both solid and dashed) in Figure 1. It is clearly unimodal.

Imagine, however, that Figure 1 describes not Thompson's catch of minnows, but rather size distribution in a pond that is suddenly subjected to harvesting, by nets whose mesh size allows mini minnows to escape (because the target is a much larger species and big minnows are merely "by-catch"). For the sake of simplicity, suppose that every minnow whose length exceeds base length by more than 15 mm will be captured by the nets. The size distribution for these big minnows is a **conditional distribution**, because its sample space is [15, 27], which is merely a subset of the original sample space: size is now conditioned on exceeding 15mm. What is the big-minnow p.d.f? Is it the dashed curve in Figure 1? The answer is no, because the area beneath the dashed curve is only about 3/5, and the area beneath a p.d.f. must always be precisely 1, even if the distribution is a conditional one. To obtain the big-minnow p.d.f., we have to multiply the dashed p.d.f. by a factor of approximately 5/3, as indicated by the thicker dashed curve of Figure 2(b), where the thinner curve is identical to the dashed curve of Figure 1.

Similar considerations apply to mini minnows. Because no survivor is longer than 15 mm above base length, the mini-minnow distribution is also a conditional one, with sample space is [0, 15]. What is the mini-minnow p.d.f? Is it the solid curve in Figure 1? Again, the answer is no, because the area beneath the solid curve is only about 2/5. To obtain the mini-minnow p.d.f., we have to multiply the dashed p.d.f. by approximately 5/2, as indicated by the thicker solid curve of Figure 2(a), where the thinner curve is identical to the solid curve of Figure 1.

So far, so good. If we know a formula for the original p.d.f in Figure 1, then we can easily deduce a formula for the p.d.f. of the mini-minnow (conditional) distribution in Figure 2(a), and from there we can use (2) to construct the c.d.f. Unfortunately, however, we don't yet have a formula for the p.d.f. in Figure 1, and so we can't yet deduce a formula for the mini-minnow c.d.f. Now, a formula for the original p.d.f. appears in Lecture 19. But this formula is in terms of the exponential function, which we haven't yet studied in detail. Do we really want to wait until we can study the exponential function further before constructing an explicit model of the mini-minnow distribution? No, of course not; we are too impatient. So let us approximate the p.d.f. by using a function we are more familiar with.

Whichever function we use, it is clear from Figure 2(a) that it must be nonnegative, increasing and concave up. One such function is the polynomial f defined on [0, 15] by

$$f(x) = \alpha x + \beta x^2, \qquad (10.4)$$

where α and β are positive parameters. Note that if we also define functions g and h by

$$g(x) = x^2 \qquad (10.5)$$

and

$$h(x) = x. \qquad (10.6)$$

then (1) implies

$$f = \beta g + \alpha h. \qquad (10.7)$$

To some extent, we can vary the parameters α and β to yield the closest possible fit between f and Thompson's data. But we cannot vary these parameters at will, because the area under the curve must always be precisely one. That is, Area(f, [0, 15]) = 1 or, from (7),

$$\text{Area}(\beta g + \alpha h, [0,15]) = 1. \qquad (10.8)$$

From (9. 17) with $k = \beta$ and $q = \alpha$, (8) implies

$$\beta \cdot \mathrm{Area}(g,[0,15]) \ + \ \alpha \cdot \mathrm{Area}(h,[0,15]) = \ 1 \qquad (10.9)$$

and from (9.20) with a = 0 and t = 15 (or Figure 9.5) we have

$$\mathrm{Area}(h,[0,15]) \ = \ \frac{1}{2}15^2 - \frac{1}{2}0^2 \ = \ \frac{225}{2}. \qquad (10.10)$$

So, from (9)-(10),

$$\beta \ = \ \frac{2-225\alpha}{2 \cdot \mathrm{Area}(g,[0,15])}. \qquad (10.11)$$

But what is Area(g, [0, 15])?

For reasons to emerge in due course, it will save us labor in the long run if we calculate not Area(g, [0, 15]), but rather

$$G(t) \ = \ \mathrm{Area}(g,[0,t]) \qquad (10.12)$$

for arbitrary t satisfying $0 \le t \le 15$. We calculate G(t) as the limit of a sequence of approximating functions. In Figure 3, the shaded areas are all overestimates of the area under the graph of g between x = 0 and x = t, but each overestimate is closer to G(t) than the previous one. Let $G_n(t)$ be the n-th such approximation. Then because, from (5), the first shaded area is that of a right-angled triangle with base t and height $g(t) = t^2$, our first approximation is

$$G_1(t) \ = \ \frac{1}{2}t \cdot g(t) \ = \ \frac{1}{2}t \cdot t^2 \ = \ \frac{1}{2}t^3. \qquad (10.13)$$

The second shaded area is that of a triangle with base t/2 and height $g(t/2) = (t/2)^2$ PLUS that of a trapezium of width t/2, minimum height $g(t/2) = (t/2)^2$ and maximum height $g(t) = t^2$. Thus our second approximation is

$$\begin{aligned} G_2(t) \ &= \ \frac{1}{2}\frac{t}{2}g(t/2) \ + \ \frac{1}{2}\frac{t}{2}\{g(t/2)+g(t)\} \\ &= \ \frac{1}{4}t \cdot (t/2)^2 \ + \ \frac{1}{4}t\{(t/2)^2 + t^2\} \qquad (10.14) \\ &= \ \frac{1}{16}t^3 \ + \ \frac{5}{16}t^3 \ = \ \frac{3}{8}t^3. \end{aligned}$$

The third shaded area is that of a triangle with base t/3 and height $g(t/3) = (t/3)^2$ PLUS that of a trapezium of width t/3, minimum height $g(t/3) = (t/3)^2$ and maximum height $g(2t/3) = (2t/3)^2$ PLUS that of a trapezium of width t/3, minimum height g(2t/3) = $(2t/3)^2$ and maximum height $g(t) = t^2$. Thus our third approximation is

$$\begin{aligned} G_3(t) \ &= \ \frac{1}{2}\frac{t}{3}g(t/3) \ + \ \frac{1}{2}\frac{t}{3}\{g(t/3)+g(2t/3)\} \ + \ \frac{1}{2}\frac{t}{3}\{g(2t/3)+g(t)\} \\ &= \ \frac{1}{6}t \cdot (t/3)^2 \ + \ \frac{1}{6}t\{(t/3)^2 + (2t/3)^2\} \ + \ \frac{1}{6}t\{(2t/3)^2 + t^2\} \qquad (10.15) \\ &= \ \frac{1}{54}t^3 \ + \ \frac{5}{54}t^3 \ + \ \frac{13}{54}t^3 \ = \ \frac{19}{54}t^3. \end{aligned}$$

Notice that $t^3/2 > 3t^3/8 > 19t^3/54$: each successive approximation is smaller than the previous one. Whatever Area(g, [0, t]) might be, it cannot possibly exceed $19t^3/54$.

We could go on like this indefinitely, but it is time to proceed to the general case, i.e., to an expression for the n-th approximation. This is the sum of n trapeziums, each of width t/n. The base of the k-th such trapezium stretches from x = (k–1)t/n to x = kt/n. Thus the k-th trapezium has minimum height g((k–1)t/n) and maximum height g(kt/n). Its area is therefore 1/2 times t/n times {g((k–1)t/n) + g(kt/n)}. The total shaded area is obtained by summing over all such trapeziums, i.e., by summing over k from k = 1 to k = n. Thus

$$\begin{aligned} G_n(t) &= \sum_{k=1}^{n} \frac{1}{2}\frac{t}{n}\left\{g((k-1)t/n) + g(kt/n)\right\} \\ &= \frac{1}{2}\frac{t}{n}\sum_{k=1}^{n}\left\{((k-1)t/n)^2 + (kt/n)^2\right\} \\ &= \frac{1}{2}\frac{t}{n}\sum_{k=1}^{n}\frac{t^2}{n^2}\left\{(k-1)^2 + k^2\right\} \\ &= \frac{1}{2}\frac{t^3}{n^3}\sum_{k=1}^{n}\left\{(k-1)^2 + k^2\right\} \\ &= \frac{1}{2}\frac{t^3}{n^3}\left\{\sum_{k=1}^{n}(k-1)^2 + \sum_{k=1}^{n}k^2\right\}. \end{aligned} \tag{10.16}$$

But

$$\begin{aligned} \sum_{k=1}^{n}(k-1)^2 &= (1-1)^2 + (2-1)^2 + (3-1)^2 + \ldots + (n-1)^2 \\ &= 0^2 + 1^2 + 2^2 + \ldots + (n-1)^2 \\ &= 0^2 + \sum_{k=1}^{n-1}k^2 \\ &= \sum_{k=1}^{n-1}k^2, \end{aligned} \tag{10.17}$$

which reduces (16) to

$$G_n(t) = \frac{1}{2}\frac{t^3}{n^3}\left\{\sum_{k=1}^{n-1}k^2 + \sum_{k=1}^{n}k^2\right\}. \tag{10.18}$$

From (6.20), however, we have

$$\sum_{k=1}^{M}k^2 = \frac{1}{6}M(M+1)(2M+1) \tag{10.19}$$

for any positive integer M. Setting M = n yields

$$\sum_{k=1}^{n}k^2 = \frac{1}{6}n(n+1)(2n+1) = \frac{1}{6}n(2n^2+3n+1), \tag{10.20}$$

whereas setting M = n – 1 yields

$$\begin{aligned}\sum_{k=1}^{n-1} k^2 &= \frac{1}{6}(n-1)(n-1+1)(2\{n-1\}+1) = \frac{1}{6}(n-1)n(2n-1) \\ &= \frac{1}{6}n(n-1)(2n-1) = \frac{1}{6}n(2n^2-3n+1).\end{aligned} \tag{10.21}$$

Substituting from (20)-(21) into (18), we now have

$$\begin{aligned}G_n(t) &= \frac{1}{2}\frac{t^3}{n^3}\left\{\sum_{k=1}^{n} k^2 + \sum_{k=1}^{n-1} k^2\right\} \\ &= \frac{1}{2}\frac{t^3}{n^3}\left\{\frac{1}{6}n(2n^2+3n+1) + \frac{1}{6}n(2n^2-3n+1)\right\} \\ &= \frac{1}{2}\frac{t^3}{n^3}\cdot\frac{1}{6}n\left\{2n^2+3n+1 + 2n^2-3n+1\right\} \\ &= \frac{1}{12}\frac{t^3}{n^2}\left\{4n^2+2\right\} \\ &= \frac{1}{3}t^3\left\{\frac{4n^2+2}{4n^2}\right\} \\ &= \frac{1}{3}\left\{1+\frac{1}{2n^2}\right\}t^3.\end{aligned} \tag{10.22}$$

Hence the last three shaded areas in Figure 3 are $G_4(t) = 11\,t^3/32$, $G_5(t) = 17\,t^3/50$ and $G_6(t) = 73\,t^3/216$. Each successive approximation is smaller than its predecessor, but it still overestimates G(t), because all approximations are overestimates. As n gets larger and larger, however, $1/n^2$ gets smaller and smaller, until eventually its contribution to $G_n(t)$ is so small as to be thoroughly ignorable (Figure 4). Then, in the limit as $n \to \infty$, we obtain

$$G(t) = \lim_{n\to\infty} G_n(t) = \frac{1}{3}t^3. \tag{10.23}$$

That is, the area under the graph of g between x = 0 and x = t is

$$G(t) = \text{Area}(g,[0,t]) = \frac{1}{3}t^3. \tag{10.24}$$

In particular, G(15) = 1125. So (11)-(12) imply

$$\beta = \frac{2-225\alpha}{2G(15)} = \frac{2-225\alpha}{2250}, \tag{10.25}$$

from which (2), (5)-(7), (9.17), (9.20) and (25) yield

$$\begin{aligned}F(t) &= \text{Area}(f,[0,t]) \\ &= \text{Area}(\beta g + \alpha h,[0,t]) \\ &= \beta\cdot\text{Area}(g,[0,t]) + \alpha\cdot\text{Area}(h,[0,t]) \\ &= \beta\cdot G(t) + \alpha\cdot\tfrac{1}{2}t^2 \\ &= \beta\cdot\tfrac{1}{3}t^3 + \alpha\cdot\tfrac{1}{2}t^2 \\ &= \left(\frac{2-225\alpha}{2250}\right)\cdot\tfrac{1}{3}t^3 + \alpha\cdot\tfrac{1}{2}t^2\end{aligned} \tag{10.26a}$$

or

$$F(t) \quad = \quad \frac{1}{2}\alpha t^2 \quad + \quad \frac{1}{3}\left(\frac{1}{1125} - \frac{\alpha}{10}\right)t^3 \qquad (10.26b)$$

after simplification. Note that

$$F(15) = 1 \qquad (10.27)$$

for any α.

SIZE (mm) ABOVE BASE LENGTH	1-3	4-6	7-9	10-12	13-15
NUMBER	1	22	52	67	114
NUMBER OF THAT SIZE OR SMALLER	1	23	75	142	256

Table 10.2 Frequencies of sizes 0-15 mm (above base length) in smallest 256 minnows

t	OBSERVED y	PREDICTED y	SQUARED ERROR
3	1/256	$18\alpha/5 + 1/125$	$(3.6\alpha + 0.004094)^2$
6	23/256	$54\alpha/5 + 8/125$	$(10.8\alpha - 0.02584)^2$
9	75/256	$81\alpha/5 + 27/125$	$(16.2\alpha - 0.07697)^2$
12	71/128	$72\alpha/5 + 64/125$	$(14.4\alpha - 0.04269)^2$
15	1	1	0

Table 10.3 Squares of errors between predictions of (26) and observations in Table 2

SIZE (mm) ABOVE BASE LENGTH	1-3	4-6	7-9	10-12	13-15
NUMBER	1	22	52	67	114
OBSERVED (CONDITIONAL) PROBABILITY	0.004	0.086	0.203	0.262	0.445
PREDICTION FROM MODEL $f(x) = 0.003547x + 0.0005342x^2$	0.021	0.082	0.171	0.290	0.437

Table 10.4 Quadratic model of size distribution (above base length) in smallest 256 minnows

The value of α remains at our disposal, and we can choose it to yield the closest between the model and Thompson's data. From (26), the probability of size not exceeding 3 mm (above base length) is predicted to be

$$F(3) \quad = \quad \frac{1}{2}\alpha \cdot 3^2 \quad + \quad \frac{1}{3}\left(\frac{1}{1125} - \frac{\alpha}{10}\right)\cdot 3^3 \quad = \quad \frac{18\alpha}{5} + \frac{1}{125} \qquad (10.28)$$

but, according to Table 2, the observed probability is 1/256. So the discrepancy or error is

$$\frac{18\alpha}{5} \; + \; \frac{1}{125} \; - \; \frac{1}{256} \quad = \quad 3.6\alpha \; + \; 0.004094. \qquad (10.29)$$

to 4 s.f. Similarly, the probability of size not exceeding 6 mm is predicted to be

$$F(6) \quad = \quad \frac{1}{2}\alpha \cdot 6^2 \quad + \quad \frac{1}{3}\left(\frac{1}{1125} - \frac{\alpha}{10}\right)\cdot 6^3 \quad = \quad \frac{54\alpha}{5} + \frac{8}{125} \qquad (10.30)$$

but observed to be 23/256, so the discrepancy or error is

$$\frac{54\alpha}{5} + \frac{8}{125} - \frac{23}{256} = 10.8\alpha - 0.02584. \tag{10.31}$$

to 4 s.f. Continuing in this manner, we obtain the expressions in Table 3 for the squares of the discrepancies between the data in Table 2 and the model y = F(t) in (26). From the last column of Table 3, the sum of square errors is $(3.6\alpha + 0.004094)^2 + (10.8\alpha - 0.02584)^2 + (16.2\alpha - 0.07697)^2 + (14.4\alpha - 0.04269)^2$ or

$$599.4(\alpha - 0.003547)^2 + 0.0008906, \tag{10.32}$$

after simplication (Exercise 1). This sum of squared errors is clearly least when α = 0.003547, in which case β = 0.0005342, from (25). So, from (4), the best-fit p.d.f. is defined by

$$f(x) = 0.003547x + 0.0005342x^2. \tag{10.33}$$

It is graphed in Figure 5(a). The corresponding c.d.f. F, defined by

$$F(t) = 0.001773t^2 + 0.0001781t^3, \tag{10.34}$$

is graphed in Figure 5(b). This model is compared with the data in Table 4, where we see that the maximum discrepancy between model and data is 0.032 for sizes 7-9. Thus the fit is by no means a bad one, as indicated in Figure 5(b).

We will study continuous distributions further in Lecture 19. Meanwhile, be sure to attempt Exercises 3 and 4.

Reference

Thompson, D'Arcy W (1942). On Growth and Form. Cambridge University Press.

Exercises 10

10.1 Verify (32)

10.2 Use the method of this lecture in conjunction with the result

$$\sum_{n=1}^{M} n = \frac{1}{2}M(M+1)$$

obtained in Exercise 3.7 to verify that $\phi(x) = \alpha x$ implies Area(ϕ, [0, t]) = $\alpha t^2/2$.

10.3* The horn of a certain species of beetle never exceeds 5 mm in length. A coleopterist who staged 10 fights among pairs of males of this species observed the following horn lengths among winners:

HORN SIZE (mm)	0 – 1	1 – 2	2 – 3	3 – 4	4 – 5
FREQUENCY	1	1	2	2	4

The coleopterist believes that the probability density f defined on [0, 5] by

$$f(x) \;=\; \alpha + \beta x,$$

where α and β are positive parameters, is an adequate model of the distribution from which these winning horn sizes are drawn.

(i) What must be the value of β, in terms of α?
(ii) Find the cumulative distribution function, F.
(iii) If P_n denotes the proportion of horn sizes that are less than or equal to n among the sample, then a measure of the discrepancy between model and data is the sum of squared errors, in this case defined by

$$\Delta \;=\; \sum_{n=1}^{5}(F(n) - P_n)^2 \;=\; \sum_{n=1}^{4}(F(n) - P_n)^2$$

(because $F(5) = 1 = P_5$). Show that

$$\Delta \;=\; \frac{104}{25}\left(\alpha - \tfrac{7}{260}\right)^2 + \frac{7}{1300}\,.$$

(iv) What is the best model of horn size distribution among winners?

10.4 The coleopterist in Exercise 2 observed the following horn lengths among losers:

HORN SIZE (mm)	0 – 1	1 – 2	2 – 3	3 – 4	4 – 5
FREQUENCY	3	3	2	2	0

The coleopterist believes that the probability density g defined on [0, 5] by

$$g(x) \;=\; \rho - \sigma x,$$

where ρ and σ are positive parameters, is an adequate model of the distribution from which these losing horn sizes are drawn.

(i) What must be the value of σ, in terms of ρ?
(ii) Find the cumulative distribution function, F.
(iii) If P_n denotes the proportion of horn sizes that are less than or equal to n among this new sample, then the sum of squared errors remains as defined in Exercise 2. Show that now

$$\Delta \;=\; \frac{104}{25}\left(\rho - \tfrac{97}{260}\right)^2 + \frac{7}{1300}\,,$$

and hence that $f(x) = (97 - 18x)/260$ defines the best model of horn size distribution among losers.

Answers and Hints for Selected Exercises

10.3 (i) We require Area(f, [0, 5]) = 1. But Area(f, [0, 5]) is the area of a trapezium with base 5, minimum height α and maximum height $\alpha + 5\beta$. Thus $5(2\alpha + 5\beta)/2 = 1$, implying

$$\beta = \frac{2}{5}\left(\frac{1}{5} - \alpha\right).$$

(ii) F(x) = Area(f, [0, x]), which is the area of a trapezium with base x, minimum height α and maximum height $\alpha + \beta x$. So

$$F(x) = \frac{x}{2}(2\alpha + \beta x) = \alpha x + \frac{1}{5}\left(\frac{1}{5} - \alpha\right)x^2 .$$

(iii) Clearly, $P_1 = 1/10$, $P_2 = 2/10$, $P_3 = 4/10$ and $P_4 = 6/10$. Also,

$$F(1) = \frac{1}{25} + \frac{4\alpha}{5}, \quad F(2) = \frac{4}{25} + \frac{6\alpha}{5}, \quad F(3) = \frac{9}{25} + \frac{6\alpha}{5}, \quad F(4) = \frac{16}{25} + \frac{4\alpha}{5}.$$

So $F(1) - P_1 = 4\alpha/5 - 3/50$, $F(2) - P_2 = 6\alpha/5 - 1/25$, $F(3) = P_3$ and $F(4) - P_4 = 4\alpha/5 + 1/25$, implying

$$\Delta = \sum_{n=1}^{4}(F(n) - P_n)^2 = \left(\tfrac{4\alpha}{5} - \tfrac{3}{50}\right)^2 + 2\cdot\left(\tfrac{6\alpha}{5} - \tfrac{1}{25}\right)^2 + \left(\tfrac{4\alpha}{5} + \tfrac{1}{25}\right)^2$$

$$= \frac{10400\alpha^2 - 560\alpha + 21}{2500}$$

$$= \frac{104}{25}\left(\alpha - \tfrac{7}{260}\right)^2 + \frac{7}{1300}$$

after simplification.

(iv) $\alpha = 7/260$ minimizes Δ. Then $\beta = 9/130$, and $f(x) = (7 + 18x)/260$.