19. Continuous probability distributions: the fundamental theorem again

In Lecture 8, we introduced the concept of probability density function f for a continuous random variable X; f is nonnegative, and total area under its graph is 1. In this lecture, we assume that X is also nonnegative (but we will relax this assumption in Lecture 28). Then the p.d.f. is defined on $[0, \infty)$ with

$$f(x) \geq 0, \quad 0 \leq x < \infty \qquad (19.1a)$$

and

$$\text{Int}(f,[0,\infty)) = \text{Area}(f,[0,\infty)) = \int_0^\infty f(x)\,dx = 1. \qquad (19.1b)$$

Note what this implies: no matter how far you go to the right, the area under the graph of f remains precisely 1, which can happen only if $f(x) \to 0$ as $x \to \infty$. Sometimes this condition is satisfied because there exists some b such that $f(x) = 0$ for $x > b$; then $\text{Int}(f, [b, \infty)) = 0$, and (1b) reduces to $\text{Int}(f, [0, b]) = 1$. At other times, however, $f(x) \to 0$ as $x \to \infty$ despite f(x) being positive (if mostly very small) throughout $[0, \infty)$. Then $\text{Int}(f, [0, \infty))$ is interpreted to mean the limit of $\text{Int}(f, [0, K])$ as $K \to \infty$ (and is precisely 1). More generally, if there exists no b such that $f(x) = 0$ for $x > b$, then $\text{Int}(f, [a, \infty))$ is called an **improper integral** and is interpreted to mean the limit as $K \to \infty$ of $\text{Int}(f, [a, K])$. For this limit to exist, however, i.e., for the integral to **converge**, it is not enough for f(x) to approach zero as $x \to \infty$; rather, f must approach zero sufficiently rapidly to prevent the enclosed area from growing without bound. We discuss improper integrals more fully in Lecture 27. Meanwhile, we finesse the issue by always choosing f to guarantee convergence.

Conversely, any function f that satisfies (1) is the p.d.f. of a random variable distributed over $[0, \infty)$. For example, the function f defined by

$$f(x) = \begin{cases} A + \frac{1}{4}\{1-3A\}x & \text{if } 0 \leq x < 2 \\ \frac{4(1-A)}{x^3} & \text{if } 2 \leq x < \infty \end{cases} \qquad (19.2)$$

is a p.d.f. if $0 < A < 1$ because (1) is then satisfied (see Exercise 1). We used this p.d.f. with $A = 0.768$ in Lecture 15 to model survival of melanoma patients. According to this model, for example, a patient survives between 1 and 3 years with probability

$$\int_1^3 f(x)\,dx = \frac{47-29A}{72} = 0.343; \qquad (19.3)$$

see Exercise 2.[1]

From Lecture 10, the cumulative distribution function of X is defined in terms of its p.d.f. If F is the c.d.f., then F is defined on $[0, \infty)$ by

$$F(t) = \text{Prob}(0 \leq X \leq t) = \text{Area}(f,[0,t]) = \int_0^t f(x)\,dx. \qquad (19.4)$$

As t increases from 0 to ∞, more and more of the area under the graph of f is accounted for, so that F(t) increases from 0 to 1. That is, F must satisfy

---

[1] Note, however, that according to Table 5.3, the same probability is (48+23)/256 = 0.277. The discrepancy is due to the error in the model, which Figure 15.5 reveals to be greatest on [1, 3].

$$F(0) \;=\; 0 \qquad (19.5a)$$
$$F \text{ is nondecreasing} \qquad (19.5b)$$
$$\lim_{t\to\infty} F(t) \;=\; 1 \qquad (19.5c)$$

or, equivalently,

$$F(0) \;=\; 0 \qquad (19.6a)$$
$$F'(t) \geq 0, \quad 0 \leq t < \infty \qquad (19.6b)$$
$$F(\infty) \;=\; 1. \qquad (19.6c)$$

Conversely, any F that satisfies (5) or (6) is the c.d.f. of a random variable on $[0, \infty)$. For example, the function F defined by

$$F(t) \;=\; \begin{cases} At \;+\; \dfrac{1}{8}\{1-3A\}t^2 & \text{if } 0 \leq t \leq 2 \\ 1 - \dfrac{2(1-A)}{t^2} & \text{if } 2 \leq t < \infty, \end{cases} \qquad (19.7)$$

is a c.d.f. if $0 < A < 1$ because (5) is then satisfied (see Exercise 1). We used this c.d.f. with $A = 0.768$ in Lecture 15 to model survival of melanoma patients. For example, a patient survives between 1 and 3 years with probability

$$F(3) \;-\; F(1) \;=\; \frac{1}{9}\{7+2A\} \;-\; \frac{1}{8}\{1+5A\} \;=\; 0.343; \qquad (19.8)$$

see Exercise 2. A more versatile example of a c.d.f. involves the exponential function. In Lecture 7 we showed that R defined by

$$R(x) \;=\; \exp(Ax^m) \qquad (19.9)$$

is strictly increasing on $[0, \infty)$; see (7.23). So F defined on $[0, \infty)$ by

$$F(x) \;=\; 1 \;-\; \frac{1}{R(x)} \;=\; 1 \;-\; \frac{1}{\exp(Ax^m)} \qquad (19.10)$$

is a legitimate c.d.f. for a continuous random variable. Why? First, because $R(0) = 1$, we have $F(0) = 1 - 1 = 0$. Second, because $R(x)$ increases with x, $1/R(x)$ decreases with x, and thus $1 - 1/R(x)$ increases with x, so that F is nondecreasing. Third, because R is strictly increasing, $1/R(x)$ approaches zero as $x \to \infty$, and so $F(\infty) = 1 - 0 = 1$. Thus (5) is satisfied.

Now, in Lectures 8 and 10, we first defined the p.d.f and then used (4) to deduce the c.d.f. But the fundamental theorem tells us that $F(t) = \text{Int}(f, [0, t])$ implies $f(t) = F'(t)$. So another way to specify a distribution is to define F first and then use $f = F'$ to deduce the p.d.f. In other words, because

$$f(t) \;=\; F'(t) \quad \Leftrightarrow \quad F(t) \;=\; \int_0^t f(x)\,dx, \qquad (19.11)$$

either f or F completely specifies a continuous distribution.[2]

In fact, when fitting a distribution to a sample, the method of choice is to fit the data to the c.d.f. (by, e.g., the method of Lectures 10 and 15), and then use (11) to deduce the p.d.f. Consider, for example, the data in Table 1. It shows year of death for 545 male

---

[2] Nevertheless, it is traditional to characterize a distribution in terms of properties of its p.d.f. For example, if f is piecewise-constant or **piecewise-uniform**, then F is piecewise-linear, but we describe the distribution as piecewise-uniform; see Exercise 3. Similarly, the distribution defined in Exercise 4 is piecewise-linear.

prairie dogs living in South Dakota between 1975 and 1989 (Hoogland, 1995, p. 396). Let us define a continous random variable X by

$$X = \text{AGE AT DEATH OF PRAIRIE DOG} \tag{19.12}$$

chosen randomly from Hoogland's data, and a sequence $\{P_n\}$ by

$$P_n = \text{Prob}(X \le n). \tag{19.13}$$

Then, from Table 1, $P_0 = 0$ and

$$P_1 = \tfrac{312}{545} \quad P_2 = \tfrac{413}{545} \quad P_3 = \tfrac{487}{545}$$
$$P_4 = \tfrac{522}{545} \quad P_5 = \tfrac{537}{545} \quad P_6 = 1 \tag{19.14}$$

The sequence $\{P_n\}$ is graphed in Figure 1(b).

YEAR OF DEATH	FREQUENCY	YEAR OF DEATH	FREQUENCY
1	312	5	12
2	101	6	8
3	74	≥ 7	0
4	35		

Table 19.1 Prairie dog lifespans

We think of $\{P_n\}$ as sampling the function F defined by $F(x) = \text{Prob}(X \le x) =$ Prob(PRAIRIE DOG LIVES AT MOST x YEARS) at integer values of x. If F is a perfect model, then

$$F(n) = P_n \tag{19.15}$$

for all n (in other words, when $\{P_n\}$ and F are graphed together, the dots all lie on the curve).[3] A measure of the extent to which this constraint is violated is the sum of squared errors

$$\Delta = \sum_{n=1}^{6} \{F(n) - P_n\}^2. \tag{19.16}$$

The smaller the value of Δ, the better the fit of the c.d.f. So we try to make Δ as small as possible.

For example, we will be able to show in Lecture 20 that F defined by (10) is concave down if m = 1 but has an inflection point if m ≥ 2. From Figure 1, however, the prairie-dog c.d.f. is evidently concave down, and so we will fit the data to the c.d.f. defined by

$$F(x) = 1 - \frac{1}{\exp(Ax)}, \tag{19.17}$$

i.e., (10) with m = 1. From (14) and (16), the sum of squared errors is

$$\begin{aligned} \Delta &= \sum_{n=1}^{6} \{F(n) - P_n\}^2 = \sum_{n=1}^{6} \left\{1 - \frac{1}{\exp(An)} - P_n\right\}^2 \\ &= \left(\tfrac{233}{545} - \tfrac{1}{\exp(A)}\right)^2 + \left(\tfrac{132}{545} - \tfrac{1}{\exp(2A)}\right)^2 + \left(\tfrac{58}{545} - \tfrac{1}{\exp(3A)}\right)^2 \\ &\quad + \left(\tfrac{23}{545} - \tfrac{1}{\exp(4A)}\right)^2 + \left(\tfrac{8}{545} - \tfrac{1}{\exp(5A)}\right)^2 + \left(\tfrac{1}{\exp(6A)}\right)^2. \end{aligned} \tag{19.18}$$

---

[3] Note, however, that the converse of this statement is false: if $F(n) = P_n$ for all n, then it does not follow that F is a perfect model. See Exercise 3.

Note that Δ depends only on A. Although it does not reduce to a simple expression like (15.38), it is just as easy for a computer to plot. The graph of Δ versus A is shown in Figure 2, which reveals that the error is least where A = 0.778 and

$$\Delta = \sum_{n=1}^{6}\{F(n)-P_n\}^2 = 0.220\times10^{-2}. \qquad (19.19)$$

So we choose F defined by

$$F(x) = 1 - \frac{1}{\exp(0.778x)} \qquad (19.20)$$

to model the prairie-dog distribution.

As a further illustration of this method, we now choose F to model the distribution of leaf thickness in *Dicerandra linearifolia* (Lectures 5 and 8). The relevant data from Lecture 5 are reproduced as Table 2. Because thicknesses are rounded to the nearest sixtieth of a millimeter, any thickness between 7/120 = 0.0583 mm and 9/120 = 0.075 m was recorded as 8/120 = 0.067 mm, whereas any thickness between 9/120 mm and 11/120 = 0.0916 mm was recorded as 10/120 = 0.083 mm, etc. Thus, if X is the thickness of a leaf selected at random, Table 2 implies $\mathrm{Prob}(X \le 0.0583) = 0$, $\mathrm{Prob}(X \le 0.075) = 9/489 = 0.0184$, $\mathrm{Prob}(0.075 \le X \le 0.0916) = \{9+5\}/489 = 14/489 = 0.0286$, and so on. It is therefore convenient to define a sequence $\{x_n\}$ by $\{x_0, x_1, x_2,\ldots, x_{12}\} = \{0.0583, 0.075, 0.0916, \ldots, 0.2583\}$ or

$$x_n = \tfrac{1}{120}\{2n+7\}, \quad 0 \le n \le 12 \qquad (19.21)$$

and to redefine P_n by

$$P_n = \mathrm{Prob}(X \le x_n), \qquad (19.22)$$

instead of (15). Then we regard F as fitting the data perfectly if

$$F(x_n) = P_n, \quad 0 \le n \le 12 \qquad (19.23)$$

and our measure of the extent to which this constraint is violated becomes

$$\Delta = \sum_{n=0}^{12}\{F(x_n)-P_n\}^2, \qquad (19.24)$$

in place of (16).

THICKNESS (mm)	FREQUENCY	THICKNESS (mm)	FREQUENCY
0.067	9	0.167	118
0.083	5	0.183	17
0.1	28	0.2	9
0.117	45	0.233	1
0.133	165	0.25	2
0.15	90		

Table 19.2 Leaf thicknesses in *Dicerandra linearifolia*

From (22) and Table 2, the sequence $\{P_n\}$ is defined on [0...12] by

$$P_0 = 0 \quad P_1 = \tfrac{3}{163} \quad P_2 = \tfrac{14}{489} \quad P_3 = \tfrac{14}{163} \quad P_4 = \tfrac{29}{163} \quad P_5 = \tfrac{84}{163} \quad P_6 = \tfrac{114}{163}$$
$$P_7 = \tfrac{460}{489} \quad P_8 = \tfrac{159}{163} \quad P_9 = \tfrac{162}{163} \quad P_{10} = \tfrac{162}{163} \quad P_{11} = \tfrac{487}{489} \quad P_{12} = 1. \qquad (19.25)$$

It is plotted in Figure 3(b), from which it appears that there is an inflection point between n = 4 and n = 6, i.e., between $x_4 = 0.125$ mm and $x_6 = 0.1553$ mm. We therefore require $m \geq 2$ in (10). Suppose we take m = 7. Then, replacing A by B^7 in (10), we have

$$F(x) \quad = \quad 1 \; - \; \frac{1}{\exp(\{Bx\}^7)}. \qquad (19.26)$$

Which value of B shall we choose? From (21) and (24), the sum of squared errors is

$$\Delta \quad = \quad \sum_{n=0}^{12} \left\{ 1 - \frac{1}{\exp(\{(2n+7)B/120\}^7)} - P_n \right\}^2, \qquad (19.27)$$

which is plotted against B in Figure 4(a). Because Δ is least for B = 6.5735, we choose

$$F(x) \quad = \quad 1 \; - \; \frac{1}{\exp(\{6.5735x\}^7)} \quad = \quad 1 \; - \; \frac{1}{\exp(530388x^7)} \qquad (19.28)$$

to model the data; see Figure 3(b). Note the inflection point where x = 0.149[4].

m	LEAST SUM OF SQUARED ERRORS (Δ)	B AT MINIMUM
2	0.442	6.311
3	0.188	6.389
4	0.0796	6.453
5	0.0318	6.504
6	1.23×10^{-2}	6.543
7	0.75×10^{-2}	6.574
8	1.07×10^{-2}	6.599
9	1.84×10^{-2}	6.621
10	2.85×10^{-2}	6.641

Table 19.3 Least sum of squared errors when fitting (29) to the leaf-thickness data

But why choose m = 7 to begin with? If you were to repeat the above exercise with

$$F(x) \quad = \quad 1 \; - \; \frac{1}{\exp(\{Bx\}^m)} \qquad (19.29)$$

for different m and in each case calculate Δ, then you would find that Δ decreases with m for $2 \leq m \leq 7$ but increases again for $m \geq 8$; see Table 3. So m = 7 is optimal.

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

4 See (20.32), from which the inflection point is at 0.978/B = 0.149.

m	LEAST SUM OF SQUARED ERRORS (Δ)	B AT MINIMUM
2	0.164	0.05513
3	0.0482	0.05563
4	1.32×10^{-2}	0.05587
5	0.787×10^{-2}	0.05607
6	1.42×10^{-2}	0.05624
7	2.5×10^{-2}	0.05639
8	3.72×10^{-2}	0.05649

Table 19.5 Least sum of squared errors when fitting (27) to minnow size (above base length)

As our third and final illustration of the method, we now choose F to model the distribution of size above base length in D'Arcy Thompson's minnows. The relevant data from Lecture 10 are reproduced as Table 4. Now, in place of (21), the sequence $\{x_n\}$ is defined by $\{x_0, x_1, x_2,..., x_9\} = \{0.5, 3.5, 6.5, ... , 27.5\}$ or

$$x_n = \tfrac{1}{2}\{6n+1\}, \quad 0 \le n \le 9, \qquad (19.30)$$

and in place of (25) the sequence $\{P_n\}$ is defined by

$$\begin{aligned} &P_0 = 0 \quad P_1 = \tfrac{1}{733} \quad P_2 = \tfrac{23}{733} \quad P_3 = \tfrac{75}{733} \quad P_4 = \tfrac{142}{733} \\ &P_5 = \tfrac{256}{733} \quad P_6 = \tfrac{513}{733} \quad P_7 = \tfrac{690}{733} \quad P_8 = \tfrac{731}{733} \quad P_9 = 1. \end{aligned} \qquad (19.31)$$

The sequence $\{P_n\}$ plotted in Figure 5(b), from which it appears that F should have an inflection point between n = 5 and n = 6, i.e., between $x_5 = 15.5$ mm and $x_6 = 18.5$ mm. We again require $m \ge 2$, but this time the optimal value of m turns out to be m = 5; see Table 5. From (29)-(30), the sum of squared errors is

$$\Delta = \sum_{n=0}^{9}\left\{1 - \frac{1}{\exp(\{(6n+1)B/2\}^5)} - P_n\right\}^2, \qquad (19.32)$$

which is plotted against B in Figure 4(b). Because Δ is least for B = 0.05607, we choose

$$F(x) = 1 - \frac{1}{\exp(\{0.05607x\}^5)} = 1 - \frac{1}{\exp(5.5399x^5 \cdot 10^{-7})} \qquad (19.33)$$

to model the data; see Figure 5(b). Note the inflection point where x = 17.1.[5]

Also plotted, in Figures 1(a), 3(a) and 5(a), are the probability density functions of the fitted distributions for prairie-dog life span, leaf thickness and minnow size. In all three cases, from (10) and (11), the p.d.f. is defined by

$$f(x) = F'(x) = \frac{d}{dx}\left\{1 - \frac{1}{R(x)}\right\} = -\frac{d}{dx}\left\{\frac{1}{R(x)}\right\}. \qquad (19.34)$$

We can simplify this expression by using the product rule; because

$$1 = R(x)\cdot\frac{1}{R(x)},$$

we have

$$0 = \frac{d}{dx}\{1\} = \frac{d}{dx}\left\{R(x)\cdot\frac{1}{R(x)}\right\} = R'(x)\cdot\frac{1}{R(x)} + R(x)\frac{d}{dx}\left\{\frac{1}{R(x)}\right\}, \qquad (19.35)$$

---

5 Again, see (20.32), from which the inflection point is at 0.956/B = 17.1.

from which (34) yields

$$f(x) \quad = \quad -\frac{d}{dx}\left\{\frac{1}{R(x)}\right\} \quad = \quad \frac{R'(x)}{\{R(x)\}^2}. \qquad (19.36)$$

Thus we know f(x) if we know $R'(x)$. But R is a composition. So immediately we ask, how do we find the derivative of a composition? To that we will turn our attention in Lecture 20.

Meanwhile, there are a number of distributions for which your knowledge of derivatives is already sufficient to obtain the p.d.f. from the c.d.f. by using $f = F'$. For example, you can readily verify that (7) implies (2). Other examples appear in Exercises 3, 4, 7 and 14.

Reference

Hoogland, J.L. (1995) The Black-Tailed Prairie Dog. University of Chicago Press

Exercises 19

19.1 (i) Show that (2) implies Int(f, [0, 2]) = (1+A)/2 and Int(f, [2, ∞]) = (1–A)/2. Hence verify that f is a p.d.f. when $0 < A < 1$, i.e., that (1a) and (1b) are satisfied.
(ii) Verify that F defined by (7) is a c.d.f. when $0 < A < 1$, i.e., (5) is satisfied.

19.2 Verify (3) and (8).

19.3 Show that the piecewise-linear join F defined by

$$F(x) = \begin{cases} \frac{312}{545}x & \text{if } 0 \le x < 1 \\ \frac{101x + 211}{545} & \text{if } 1 \le x < 2 \\ \frac{74x + 265}{545} & \text{if } 2 \le x < 3 \\ \frac{35x + 382}{545} & \text{if } 3 \le x < 4 \\ \frac{3(5x + 154)}{545} & \text{if } 4 \le x < 5 \\ \frac{8x + 497}{545} & \text{if } 5 \le x < 6 \\ 1 & \text{if } 6 \le x < \infty \end{cases}$$

is a cumulative distribution function, and find the associated probability density function. Verify that F satisfies (13) exactly for every value of n. Does F yield a better model of prairie-dog survival than the c.d.f. in Figure 1(b)? Why or why not?

19.4 Show that the piecewise-quadratic join F of seven components defined by

$$4360F(x) = \begin{cases} 2496x & \text{if } 0 \le x \le \frac{1}{2} \\ -211 + 3340x - 844x^2 & \text{if } \frac{1}{2} \le x \le \frac{3}{2} \\ 1445 + 1132x - 108x^2 & \text{if } \frac{3}{2} \le x \le \frac{5}{2} \\ 1145 + 1372x - 156x^2 & \text{if } \frac{5}{2} \le x \le \frac{7}{2} \\ 2076 + 840x - 80x^2 & \text{if } \frac{7}{2} \le x \le \frac{9}{2} \\ 3129 + 372x - 28x^2 & \text{if } \frac{9}{2} \le x \le \frac{11}{2} \\ 3976 + 64x & \text{if } \frac{11}{2} \le x \le 6 \\ 4360 & \text{if } 6 \le x < \infty \end{cases}$$

is a cumulative distribution function, and find the corresponding probability density function. Is F smooth? Show that if F is used to model the prairie-dog survival data in Table 1, then the sum of squared errors is $\Delta = 0.248 \times 10^{-2}$.

19.5 Verify Table 8.2 (by using a calculator).

19.6 A function f is defined on [0, ∞) by

$$f(x) = \begin{cases} 4x / b^2 & \text{if } 0 \le x \le b/2 \\ 4(b-x)/b^2 & \text{if } b/2 \le x \le b \\ 0 & \text{if } b \le x < \infty \end{cases}$$

where b is an arbitrary positive constant.
(i) Show that f defines a legitimate probability density function on [0, ∞)
(ii) Find an explicit expression for its cumulative distribution function.

19.7 Is the function defined by

$$F(x) = \begin{cases} 1 + b(x-a)^3 & \text{if } 0 \le x \le a \\ 1 & \text{if } a \le x < \infty \end{cases}$$

a legitimate c.d.f for (i) all values of b (ii) no values of b or (iii) some values of b? Elucidate. If F is indeed a c.d.f., then find its p.d.f.

19.8 Is the function defined by

$$F(x) = \begin{cases} \frac{1}{6}x^3 - x^2 + \frac{19}{12}x & \text{if } 0 \le x \le 4 \\ 1 & \text{if } 4 \le x < \infty \end{cases}$$

a legitimate c.d.f? Why or why not?

19.9 It is known that the probability density function f defined on $[0, \infty)$ by

$$f(t) = \begin{cases} 2t^5 & \text{if } 0 \le t < 1 \\ \alpha & \text{if } 1 \le t < \frac{7}{6} \\ 9 - \beta t & \text{if } \frac{7}{6} \le t < \frac{3}{2} \\ 0 & \text{if } \frac{3}{2} \le t < \infty \end{cases}$$

is a continuous function. What are therefore the values of α and β? Obtain an explicit formula for the associated cumulative distribution function F on $[0, \infty)$.

19.10 It is known that the probability density function f defined on $[0, \infty)$ by

$$f(t) = \begin{cases} t^3 & \text{if } 0 \le t < 1 \\ 1 & \text{if } 1 \le t < \alpha \\ 2(\beta - t) & \text{if } \alpha \le t < \beta \\ 0 & \text{if } \beta \le t < \infty \end{cases}$$

is a continuous function. What are therefore the values of α and β? Obtain an explicit formula for the associated cumulative distribution function F on $[0, \infty)$.

19.11 A smooth function f, defined on $[0, 3]$ by

$$f(t) = \begin{cases} At - Bt^2 & \text{if } 0 \le t \le 2 \\ C(3-t)^2 & \text{if } 2 \le t \le 3 \end{cases}$$

satisfies

$$\int_0^3 f(t)\,dt = 1$$

(i) Find A, B and C
(ii) Find all local extrema of f on $[0, 3]$
(iii) Is f a probability density function? Why, or why not?
(iv) Crudely sketch the graphs of f, f' and f'', clearly indicating
  (A) the unique global maximum of f
  (B) the unique inflection point of f

19.12 A smooth function f, defined on [0, 5] by

$$f(t) \;=\; \begin{cases} At(16-t^2) & \text{if } 0 \le t < 3 \\ Bt - Ct^2 & \text{if } 3 \le t \le 5 \end{cases}$$

satisfies

$$\int_0^5 f(t)\,dt \;=\; 1.$$

(i) Find A, B and C
(ii) Find all local extrema of f on [0, 5], if any.
(iii) Find all inflection points of f on [0, 5], if any.
(iv) Is f a probability density function? Why, or why not?
(v) Crudely sketch the graph of f, indicating both its global maximum and global minimum on [0, 5].

19.13 If q denotes a nondecreasing function on [0, b] satisfying q(b) = 1, show that F defined on [0, ∞) by

$$F(x) \;=\; \begin{cases} \frac{x}{b}\left\{1 + \left(1 - \frac{x}{b}\right)q(x)\right\} & \text{if } 0 \le x < b \\ 1 & \text{if } b \le x < \infty \end{cases}$$

is a legitimate continuous c.d.f.

19.14 A function F is defined on [0, ∞) by

$$F(t) \;=\; \begin{cases} At(c-t) \;+\; \frac{\theta + Ac^2}{\theta+1}\left(\frac{t}{c}\right) & \text{if } 0 \le t \le c \\ 1 \;-\; \frac{1-Ac^2}{\theta+1}\left(\frac{c}{t}\right)^{\theta} & \text{if } c \le t < \infty \end{cases}$$

where θ is a positive integer.
(i) When is F a cumulative distribution function?
(ii) What is then the associated probability density function f?
(iii) Is F smooth?
(iv) Is f smooth?

19.15 A smooth function f, defined on [0, ∞] by

$$f(t) \;=\; \begin{cases} 1 - At(t-1)^2 & \text{if } 0 \le t < 1 \\ \frac{Bt - C}{t^5} & \text{if } 1 \le t \le \infty \end{cases}$$

satisfies

$$\int_0^\infty f(t)\,dt \;=\; 1.$$

(i) Find A, B and C
(ii) Find all local extrema of f on [0, ∞], if any.
(iii) Find all inflection points of f on [0, ∞], if any. Where is f concave down?
(iv) Is f a probability density function? Why, or why not?
(v) Crudely sketch the graph of f, indicating both global extrema.

Answers and Hints for Selected Exercises

19.1 From (2) and (12.25) with $a = 0$, $b = 2$, $k = A$, $q = (1–3A)/4$, $u(x) = 1$ and $v(x) = x$,

$$\int_0^2 f(x)dx = \int_0^2 \left\{A + \frac{1}{4}\{1-3A\}x\right\}dx$$
$$= A\int_0^2 1\,dx + \frac{1}{4}\{1-3A\}\int_0^2 x\,dx$$
$$= A(2-0) + \frac{1}{4}\{1-3A\}\cdot\frac{1}{2}\left(2^2-0^2\right) = \frac{1+A}{2}$$

From (2), (12.25) with $a = 2$, $b = \infty$, $k = 4(1–A)$, $q = 0$, $u(x) = x^{-3}$ and Table 18.1,

$$\int_2^\infty f(x)dx = \int_2^\infty \left\{\frac{4(1-A)}{x^3}\right\}dx = 4(1-A)\int_2^\infty x^{-3}\,dx$$
$$= 4(1-A)\int_2^\infty \frac{d}{dx}\left\{-\frac{1}{2}x^{-2}\right\}dx = 4(1-A)\int_2^\infty V'(x)dx,$$

where we have defined V on $[2, \infty)$ by

$$V(x) = -\frac{1}{2x^2}.$$

Hence, by the fundamental theorem, i.e., (18.20), with $a = 2$ and $t \to \infty$, we have

$$\int_2^\infty f(x)dx = 4(1-A)\int_2^\infty V'(x)dx = 4(1-A)\{V(\infty)-V(2)\}$$
$$= 4(1-A)\left\{0 - \left(-\frac{1}{8}\right)\right\} = \frac{1-A}{2},$$

because $1/2x^2$ approaches zero as x approaches infinity.

19.3 The p.d.f. is f defined by

$$f(x) = \begin{cases} \frac{312}{545} & \text{if } 0 \le x < 1 \\ \frac{101}{545} & \text{if } 1 \le x < 2 \\ \frac{74}{545} & \text{if } 2 \le x < 3 \\ \frac{7}{109} & \text{if } 3 \le x < 4 \\ \frac{3}{109} & \text{if } 4 \le x < 5 \\ \frac{8}{545} & \text{if } 5 \le x < 6 \\ 0 & \text{if } 6 \le x < \infty. \end{cases}$$

19.4 The p.d.f. is f defined by

$$f(x) = \begin{cases} \frac{312}{545} & \text{if } 0 < x < \frac{1}{2} \\ \frac{835-422x}{1090} & \text{if } \frac{1}{2} < x < \frac{3}{2} \\ \frac{283-54x}{1090} & \text{if } \frac{3}{2} < x < \frac{5}{2} \\ \frac{343-78x}{1090} & \text{if } \frac{5}{2} < x < \frac{7}{2} \\ \frac{21-4x}{109} & \text{if } \frac{7}{2} < x < \frac{9}{2} \\ \frac{93-14x}{1090} & \text{if } \frac{9}{2} < x < \frac{11}{2} \\ \frac{8}{545} & \text{if } \frac{11}{2} < x < 6 \end{cases}$$

19.6 (i) Clearly $f(x) \geq 0$ for all $x \in [0, \infty)$. So f is a p.d.f. if Int(f, [0, b]) = 1. Because f is piecewise-linear and continuous with f(0) = 0 = f(b) and mode b/2, Area(f, [0, b]) is the area of a triangle with base b and height f(m) = f(b/2) = 2/b. So Int(f, [0, b]) = Area(f, [0, b]) = (1/2) · b · 2/b = 1.

(ii) For $t \leq b/2$, we have F(t) =

$$\int_0^t f(x)dx = \int_0^t \frac{4x}{b^2}\,dx = \frac{4}{b^2}\int_0^t x\,dx = \frac{4}{b^2}\left.\frac{x^2}{2}\right|_0^t = \frac{4}{b^2}\left(\frac{t^2}{2} - \frac{0^2}{2}\right) = \frac{2t^2}{b^2}.$$

Note that F(m) = F(b/2) = 1/2. Thus, for $b/2 \leq t \leq b$, we have

$$F(t) = \int_0^t f(x)dx = \int_0^{b/2} f(x)dx + \int_{b/2}^t f(x)dx = F(b/2) + \int_{b/2}^t \frac{4(b-x)}{b^2}dx$$

$$= \frac{1}{2} + \frac{4}{b^2}\int_{b/2}^t (b-x)\,dx = \frac{1}{2} + \frac{4}{b^2}\left\{b\int_{b/2}^t 1\,dx - \int_{b/2}^t x\,dx\right\}$$

$$= \frac{1}{2} + \frac{4}{b^2}\left\{b(t-b/2) - \frac{1}{2}\left\{t^2 - \left(\tfrac{b}{2}\right)^2\right\}\right\} = \frac{4bt - b^2 - 2t^2}{b^2}$$

after simplification. Note that F(b) = 1. Thus the c.d.f. is given by

$$F(t) = \begin{cases} 2t^2/b^2 & \text{if } 0 \leq t \leq b/2 \\ (4bt - b^2 - 2t^2)/b^2 & \text{if } b/2 \leq t \leq b \\ 1 & \text{if } b \leq t < \infty \end{cases}$$

19.7 F(0) = 0 implies that F is a c.d.f only if $b = a^{-3}$. Then, because

$$\frac{d}{dx}\{(x-a)^3\} = \frac{d}{dx}\{x^3 - 3ax^2 + 3a^2x - a^3\} = 3x^2 - 6ax^2 + 3a^2 - 0 = 3(x-a)^2,$$

we have

$$f(x) = \begin{cases} 3(a-x)^2/a^3 & \text{if } 0 \leq x < a \\ 0 & \text{if } a \leq x < \infty \end{cases}$$

Note that the method used here to obtain $f = F'$ is inefficient; it will be superseded in the following lecture.

19.8 No, because there exists a subdomain on which F is decreasing.

19.9 From f(1–) = f(1+) and f(7/6 –) = f(7/6 +) we obtain $\alpha = 2$ and $\beta = 6$. So

$$f(x) = \begin{cases} 2x^5 & \text{if } 0 \le x < 1 \\ 2 & \text{if } 1 \le x < \frac{7}{6} \\ 9 - 6x & \text{if } \frac{7}{6} \le x < \frac{3}{2} \\ 0 & \text{if } \frac{3}{2} \le x < \infty \end{cases} \Rightarrow F(t) = \int_0^t f(x)dx = \begin{cases} \frac{1}{3}t^6 & \text{if } 0 \le t < 1 \\ 2t - \frac{5}{3} & \text{if } 1 \le t < \frac{7}{6} \\ 9t - 3t^2 - \frac{69}{12} & \text{if } \frac{7}{6} \le t < \frac{3}{2} \\ 1 & \text{if } \frac{3}{2} \le t < \infty \end{cases}$$

19.10 Because f is a p.d.f. with f(t) = 0 on $[\beta, \infty)$, Int(f, $[0, \infty)$) = 1 $\Rightarrow$

$$\int_0^1 t^3\,dt \;+\; \int_1^\alpha 1\,dt \;+\; \int_\alpha^\beta 2(\beta - t)\,dt \;=\; 1.$$

So, on using the fundamental theorem, we have

$$\left.\frac{t^4}{4}\right|_0^1 \;+\; t\Big|_1^\alpha \;+\; \left.\left\{-(\beta - t)^2\right\}\right|_\alpha^\beta \;=\; 1,$$

implying

$$\frac{1}{4} \;+\; \alpha - 1 \;-\; (\beta - \beta)^2 \;+\; (\beta - \alpha)^2 \;=\; 1$$

or $\alpha + (\beta - \alpha)^2 = 7/4$. Because f is continuous, $f(\alpha-) = f(\alpha+)$, implying $1 = 2(\beta - \alpha)$. So $\beta - \alpha = 1/2$, implying $\alpha + (1/2)^2 = 7/4$ or $\alpha = 3/2$. Then $\beta = \alpha + 1/2 = 2$. So

$$f(x) = \begin{cases} x^3 & \text{if } 0 \le x < 1 \\ 1 & \text{if } 1 \le x < \frac{3}{2} \\ 2(2 - x) & \text{if } \frac{3}{2} \le x < 2 \\ 0 & \text{if } 2 \le x < \infty \end{cases} \Rightarrow F(t) = \int_0^t f(x)dx = \begin{cases} \frac{1}{4}t^4 & \text{if } 0 \le t < 1 \\ t - \frac{3}{4} & \text{if } 1 \le t < \frac{3}{2} \\ 4t - t^2 - 3 & \text{if } \frac{3}{2} \le t < 2 \\ 1 & \text{if } 2 \le t < \infty \end{cases}$$

19.11 (i) On [0, 3],

$$f(t) \;=\; \begin{cases} At - Bt^2 & \text{if } 0 \le t < 2 \\ C(3 - t)^2 & \text{if } 2 \le t \le 3 \end{cases}$$

implies

$$f'(t) \;=\; \begin{cases} A - 2Bt & \text{if } 0 \le t < 2 \\ -2C(3 - t) & \text{if } 2 \le t \le 3 \end{cases}$$

Because f is continuous, f(2-) = f(2+) $\Rightarrow$ 2A – 4B = C; because f is smooth, $f'(2\text{-}) = f'(2+)$ $\Rightarrow$ A – 4B = –2C. Subtracting to eliminate 4B, we get A = 3C. Therefore B = (2A–C)/4 = 5C/4. So

$$f(t) \;=\; \frac{C}{4}\begin{cases} 12t - 5t^2 & \text{if } 0 \le t < 2 \\ 4(3 - t)^2 & \text{if } 2 \le t \le 3 \end{cases}$$

But Int(f, [0, 3]) = 1. Therefore,

$$\frac{C}{4}\left\{\int_0^2 (12t - 5t^2)\,dt \;+\; \int_2^3 4(3 - t)^2\,dt\right\} \;=\; 1,$$

implying

$$\frac{C}{4}\left\{\int_0^2 \frac{d}{dt}(6t^2 - \tfrac{5}{3}t^3)\ dt \quad + \quad \int_2^3 -\tfrac{4}{3}(3-t)^3\ dt\right\} \quad = \quad 1$$

and, by the fundamental theorem,

$$\frac{C}{4}\left\{6t^2 - \tfrac{5}{3}t^3\Big|_0^2 \quad + \quad -\tfrac{4}{3}(3-t)^3\Big|_2^3\right\} \quad = \quad 1$$

or

$$\frac{C}{4}\left\{6\cdot 2^2 - \tfrac{5}{3}2^3 - 0 \quad + \quad \{-\tfrac{4}{3}0^3 - (-\tfrac{4}{3}1^3)\}\right\} \quad = \quad 1$$

$$\Rightarrow \quad \frac{C}{4}\left\{\frac{32}{3} \quad + \quad \frac{4}{3}\right\} \quad = \quad 1 \quad \Rightarrow \quad C \quad = \quad \frac{1}{3}$$

(ii) From above,

$$f'(t) \quad = \quad \begin{cases} 1 - 5t/6 & \text{if } 0 \le t < 2 \\ \tfrac{2}{3}(t-3) & \text{if } 2 \le t \le 3 \end{cases}$$

On [2, 3), because $t-3 < 0$, we have $f'(t) < 0$. On [0, 2], $f'(t) > 0$ if $0 \le t < 6/5$ but $f'(t) < 0$ if $6/5 < t \le 2$. So $f'(t) > 0$ on [0, 6/5) but $f'(t) < 0$ on (6/5, 3), implying a local maximum where $t = 6/5$.

(iii) Yes, f is a p.d.f. because Int(f, [0, 5]) = 1 *and*

$$f(t) \quad = \quad \begin{cases} t(1 - 5t/12) & \text{if } 0 \le t < 2 \\ \tfrac{1}{3}(3-t)^2 & \text{if } 2 \le t \le 3 \end{cases}$$

is nonnegative on [0, 3].

(iv) The graph of f is concave down on [0, 2] and concave up on [2, 3]; f increases from $f(0) = 0$ to its global maximum $f(6/5) = 3/5$ and then decreases again to $f(3) = 0$. So there are two global minimizers. The graph of f' is piecewise linear. It decreases from $f'(0) = 1$ to $f'(2) = -2/3$ and then increases again to $f'(3) = 0$. The graph of f'' is piecewise constant (or, if you prefer, piecewise uniform); it jumps from $-5/6$ to $2/3$ at $t = 2$, the unique inflection point (either because f'' discontinuously changes sign, or because f' has a local minimum). Go to http://www.math.fsu.edu/~mm-g/QuizBank/mac3311.f96.html (Problem #2) to see the graphs.

19.12 (i) On [0, 5],

$$f(t) \quad = \quad \begin{cases} A(16t - t^3) & \text{if } 0 \le t < 3 \\ Bt - Ct^2 & \text{if } 3 \le t \le 5 \end{cases}$$

implies

$$f'(t) \quad = \quad \begin{cases} A(16 - 3t^2) & \text{if } 0 \le t < 3 \\ B - 2Ct & \text{if } 3 \le t \le 5 \end{cases}$$

Because f is continuous, $f(3-) = f(3+) \Rightarrow 21A = 3B - 9C \Rightarrow 7A = B - 3C$; because f is smooth, $f'(3-) = f'(3+) \Rightarrow -11A = B - 6C$. Subtracting to eliminate B, we get $18A = 3C$, or $A = C/6$. Then $B = 7A + 3C = 25C/6$. So

$$f(t) = \frac{C}{6}\begin{cases} 16t - t^3 & \text{if } 0 \le t < 3 \\ 25t - 6t^2 & \text{if } 3 \le t \le 5 \end{cases}$$

But Int(f, [0, 5]) = 1. Therefore,

$$\frac{C}{6}\left\{\int_0^3 (16t - t^3)dt + \int_3^5 (25t - 6t^2)dt\right\} = 1,$$

implying

$$\frac{C}{6}\left\{\int_0^3 \frac{d}{dt}(8t^2 - \tfrac{1}{4}t^4)\ dt + \int_3^5 (\tfrac{25}{2}t^2 - 2t^3)\ dt\right\} = 1$$

and, by the fundamental theorem,

$$\frac{C}{6}\left\{8t^2 - \tfrac{1}{4}t^4\Big|_0^3 + \tfrac{25}{2}t^2 - 2t^3\Big|_3^5\right\} = 1$$

or

$$\frac{C}{6}\left\{8\cdot 3^2 - \tfrac{1}{4}3^4 - 0 + \tfrac{25}{2}5^2 - 2\cdot 5^3 - (\tfrac{25}{2}3^2 - 2\cdot 3^3)\right\} = 1$$

$$\Rightarrow \frac{C}{6}\left\{3^2\left(8 - \frac{9}{4}\right) + 5^2\left(\frac{25}{2} - 10\right) - 3^2\left(\frac{25}{2} - 6\right)\right\} = 1$$

$$\Rightarrow \frac{C}{6}\left\{\frac{207}{4} + \frac{125}{2} - \frac{117}{2}\right\} = 1 \Rightarrow C = \frac{24}{223}$$

(ii) From above,

$$f'(t) = \frac{4}{223}\begin{cases} 16 - 3t^2 & \text{if } 0 \le t < 3 \\ 25 - 12t & \text{if } 3 \le t \le 5 \end{cases}$$

On [3, 5], because $25 - 12t \le 25 - 36 = -11$, we have $f'(t) < 0$. On [0, 3], if we define $t^* = \sqrt{16/3} = 4/\sqrt{3}$, then

$$f'(t) = \frac{4}{223}(16 - 3t^2) = \frac{12}{223}\left\{\frac{16}{3} - t^2\right\} = \frac{12}{223}\left\{(t^*)^2 - t^2\right\} = \frac{12}{223}(t^* - t)(t^* + t)$$

So $f'(t) > 0$ if $0 \le t < t^*$ but $f'(t) < 0$ if $t^* < t \le 5$, implying a local maximum $f(t^*) = 512/\{669\sqrt{3}\} = 0.442$ where $t = t^* = 2.309$.

(iii) From above,

$$f''(t) = \frac{4}{223}\begin{cases} -6t & \text{if } 0 \le t < 3 \\ -12 & \text{if } 3 \le t \le 5 \end{cases}$$

So $f''(t)$ does not change sign on [0, 5]. Hence no inflection points.

(iv) f is a p.d.f. if Int(f, [0, 5]) = 1 and $f(t) \ge 0$ on [0, 5]. The first condition is satisfied, but not the second, because on [3, 5] we have $f(t) = 4t(25-6t)/223$, which is negative for $25/6 < t \le 5$.

(v) The global maximum and minimum are $f(t^*) = 512/\{669\sqrt{3}\} = 0.442$ and $f(5) = -100/223 = -0.448$, respectively.

19.13 First show that

$$F'(x) = \begin{cases} \frac{1}{b}\left\{1+\left(1-\frac{2x}{b}\right)q(x)+\left(1-\frac{x}{b}\right)x\,q'(x)\right\} & \text{if } 0 \le x < b \\ 0 & \text{if } b \le x < \infty \end{cases}$$

Then use $q(x) \le q(b) = 1$, $q'(x) \ge 0$.

19.14 (i) When $A\,c^2 \le 1$ (if $A\,c^2 > 1$ then $F(t) > 1$).

(ii) See (24.2).

(iii) Yes, because $F(c-) = F(c+) = (\theta + A\,c^2)/(\theta+1)$, $f(c-) = f(c+) = \theta(1-A\,c^2)/(\theta+1)c$ and $f = F'$.

(iv) Only if $f'(c-) = f'(c+)$, i.e., if $-2A = -\theta(1 - A\,c^2)/c^2$, or $A\,c^2 = \theta/(\theta+2)$.

19.15 (i) From

$$f(t) = \begin{cases} 1 + A(2t^2 - t^3 - t) & \text{if } 0 \le t < 1 \\ \frac{B}{t^4} - \frac{C}{t^5} & \text{if } 1 \le t \le \infty \end{cases}$$

we have

$$f'(t) = \begin{cases} A(4t - 3t^2 - 1) & \text{if } 0 \le t < 1 \\ -\frac{4B}{t^5} + \frac{5C}{t^6} & \text{if } 1 \le t \le \infty \end{cases}$$

on using our rule for the derivative of a sum of multiples (in conjunction with Table 18.1). Because f is continuous, $f(1-) = f(1+) \Rightarrow 1 = B - C$; and because f is smooth, $f'(1-) = f'(1+) \Rightarrow 0 = 5C - 4B$. So $B = 5$ and $C = 4$, implying

$$\int_1^\infty f(t)\,dt = \int_1^\infty \left\{\frac{5}{t^4} - \frac{4}{t^5}\right\}dt = 5\int_1^\infty t^{-4}\,dt - 4\int_1^\infty t^{-5}\,dt$$

$$= 5\int_1^\infty \frac{d}{dt}\left\{-\frac{1}{3}t^{-3}\right\}dt - 4\int_1^\infty \frac{d}{dt}\left\{-\frac{1}{4}t^{-4}\right\}dt$$

$$= 5\left(-\frac{1}{3}t^{-3}\right)\Big|_1^\infty - 4\left(-\frac{1}{4}t^{-4}\right)\Big|_1^\infty = 5\left\{0-\left(-\frac{1}{3}\right)\right\} - 4\left\{0-\left(-\frac{1}{4}\right)\right\} = \frac{2}{3}$$

But $\mathrm{Int}(f, [0, 1]) + \mathrm{Int}(f, [1, \infty)) = \mathrm{Int}(f, [0, \infty)) = 1$. So $\mathrm{Int}(f, [0, 1]) = 1 - 2/3 = 1/3$. That is,

$$\frac{1}{3} = \int_0^1 \left\{1 + A(2t^2 - t^3 - t)\right\}dt = \int_0^1 1\,dt + A\int_0^1 \{2t^2 - t^3 - t\}dt$$

$$= 1 - 0 + A\int_0^1 \frac{d}{dt}\left\{\tfrac{2}{3}t^3 - \tfrac{1}{4}t^4 - \tfrac{1}{2}t^2\right\}dt$$

$$= 1 + A\left(\tfrac{2}{3}t^3 - \tfrac{1}{4}t^4 - \tfrac{1}{2}t^2\right)\Big|_0^1 = 1 + A\left(\tfrac{2}{3} - \tfrac{1}{4} - \tfrac{1}{2} - 0\right) = 1 - \frac{A}{12},$$

implying $A = 8$. In sum, $A = 8$, $B = 5$ and $C = 4$.

(ii) Now

$$f'(t) \;=\; \begin{cases} 8(3t-1)(1-t) & \text{if } 0 \le t < 1 \\ \dfrac{20(1-t)}{t^6} & \text{if } 1 \le t \le \infty \end{cases}$$

So $f'(t)$ is negative on [0, 1/3), positive on (1/3, 1) and negative on (1, ∞), with $f'(1/3) = 0 = f'(1)$. This means there is a local (also global) minimum where t = 1/3, and a local (also global) maximum where t = 1.

(iii) From

$$f''(t) \;=\; \begin{cases} A(4-6t) & \text{if } 0 \le t < 1 \\ \dfrac{20B}{t^6} - \dfrac{30C}{t^7} & \text{if } 1 \le t \le \infty \end{cases} \;=\; \begin{cases} 16(2-3t) & \text{if } 0 \le t < 1 \\ \dfrac{20(5t-6)}{t^7} & \text{if } 1 \le t \le \infty \end{cases}$$

we see that there are two inflection points, one where t = 2/3 and another where t = 6/5 (there is no inflection point where t = 1 because f'' does not change sign as it drops from $f''(1-) = -16$ to $f''(1+) = -20$). Because f'' is positive on [0, 2/3), negative on (2/3, 1) and (1, 6/5) and positive on (6/5, ∞), f is concave down on (2/3, 6/5) and concave up elsewhere.

(iv) No, f is a p.d.f. because although Int(f, [0, ∞]) = 1,

$$f(t) \;=\; \begin{cases} 1 + A(2t^2 - t^3 - t) & \text{if } 0 \le t < 1 \\ \dfrac{B}{t^4} - \dfrac{C}{t^5} & \text{if } 1 \le t \le \infty \end{cases} \;=\; \begin{cases} (1-2t)(1-6t+4t^2) & \text{if } 0 \le t < 1 \\ \dfrac{5t-4}{t^5} & \text{if } 1 \le t \le \infty \end{cases}$$

is negative on the interval [c, 1/2], where $c = (6-2\sqrt{5})/8 \approx 0.191$.

(v) Go to **http://www.math.fsu.edu/~mm-g/QuizBank/MAC2311.f97/Answers/assC4.gif** for the graph. The global maxima and minima are f(1) = 1 and f(1/3) = –0.185, respectively.