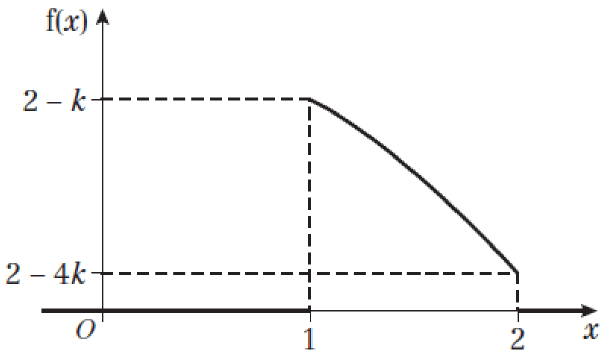
Exam-style practice: AS level

1 a As k is positive, between $(1, 2-k)$ and $(1, 2-4k)$ f(x) is part of a quadratic with a negative x^2 coefficient; otherwise f(x) is 0. The sketch of the function is:



b The mode of a continuous random variable is the value of x for which the probability distribution function is a maximum. So the mode is the value of x at the maximum of f(x), i.e. the highest point of the graph. In this case, from the graph, the maximum occurs at $x = 1$. Thus the mode of X is 1.

c The area under the probability distribution function curve must equal 1, so $\int_1^2 (2-kx^2)\,dx = 1$:

$$\int_1^2 (2-kx^2)\,dx = \left[2x - \frac{kx^3}{3}\right]_1^2 = \left[4 - \frac{8k}{3}\right] - \left[2 - \frac{k}{3}\right] = 2 - \frac{7k}{3}$$

So $2 - \frac{7k}{3} = 1 \Rightarrow \frac{7k}{3} = 1 \Rightarrow k = \frac{3}{7}$

d In order to find $E(X)$, use $E(X) = \int_{-\infty}^{\infty} x\,f(x)\,dx$

So $E(X) = \int_1^2 x\left(2 - \frac{3}{7}x^2\right)dx = \int_1^2 2x - \frac{3}{7}x^3\,dx$

$$= \left[x^2 - \frac{3x^4}{28}\right]_1^2 = 4 - \frac{12}{7} - \left(1 - \frac{3}{28}\right)$$

$$= \frac{16}{7} - \frac{25}{28} = \frac{64-25}{28} = \frac{39}{28} = 1.393 \text{ (4 s.f.)}$$

e The cumulative distribution function, F(x), for $1 \leqslant x \leqslant 2$ is given by:

$$F(x) = \int_1^x 2 - \frac{3t^2}{7}\,dt = \left[2t - \frac{t^3}{7}\right]_1^x = 2x - \frac{x^3}{7} - \left(2 - \frac{1}{7}\right) = 2x - \frac{x^3}{7} - \frac{13}{7}$$

$$\text{So } F(x) = \begin{cases} 0 & x < 1 \\ 2x - \frac{x^3}{7} - \frac{13}{7} & 1 \leqslant x \leqslant 2 \\ 1 & x > 2 \end{cases}$$

1 f The median is the value, m, such that $\mathrm{F}(m)=0.5$. So from part **e**:

$$\mathrm{F}(m)=2m-\frac{m^3}{7}-\frac{13}{7}=0.5$$

$\Rightarrow 28m-2m^3-26=7$ multiplying both sides by 14

$\Rightarrow 2m^3-28m+33=0$

g As the mode (1) < median (1.357) < mean (1.395), the distribution of X is positively skewed. This can also be deduced from the graph in part **a**.

2 a $\dfrac{1}{b-a}=\dfrac{1}{40-(0)}=\dfrac{1}{40}$

So the probability distribution function is

$$\mathrm{f}(x)=\begin{cases}\dfrac{1}{40} & 0\leqslant x\leqslant 40\\ 0 & \text{otherwise}\end{cases}$$

b $\mathrm{E}(x)=\dfrac{b+a}{2}=\dfrac{(40+0)}{2}=20$

c $\mathrm{Var}(x)=\dfrac{(b-a)^2}{12}=\dfrac{(40-0)^2}{12}=\dfrac{1600}{12}=\dfrac{400}{3}=133.3 \text{ (4 s.f.)}$

d $\mathrm{P}(15\leqslant X\leqslant 30)=\dfrac{1}{40}(30-15)=\dfrac{15}{40}=\dfrac{3}{8}=0.375$

e $\mathrm{P}(X\leqslant 25\mid X>15)=\dfrac{\mathrm{P}(X>15)\cap \mathrm{P}(X\leq 25)}{\mathrm{P}(X>15)}=\dfrac{\mathrm{P}(15<X\leqslant 25)}{\mathrm{P}(X>15)}=\dfrac{\mathrm{P}(15<X\leqslant 25)}{\mathrm{P}(15<X\leqslant 40)}$

So $\mathrm{P}(X\leqslant 25\mid X>15)=\dfrac{\frac{1}{40}(25-15)}{\frac{1}{40}(40-15)}=\dfrac{10}{25}=\dfrac{2}{5}=0.4$

3 a The table shows the ranks and d and d^2 for each pair of ranks:

Student	x	y	x_{rank}	y_{rank}	d	d^2
A	21	40	2	2	0	0
B	5	31	10	5	5	25
C	13	23	8	9	–1	1
D	16	27	6	7	–1	1
E	12	26	9	8	1	1
F	15	38	7	3	4	16
G	18	21	4	10	–6	36
H	28	47	1	1	0	0
I	17	33	5	4	1	1
J	19	30	3	6	–3	9

$$\sum d^2 = 90$$

$$r_s = 1 - \frac{6\sum d^2}{n(n^2-1)} = 1 - \frac{6\times 90}{10(10^2-1)} = 0.455 \text{ (3 s.f.)}$$

b $H_0: \rho = 0, \; H_1: \rho \neq 0$

Sample size = 10

Significance level = 0.05 (so 0.025 in each tail)

The critical value for r_s for a 0.025 significance level with a sample size of 10 is $r_s = 0.6485$

As $0.455 < 0.6485$, accept H_0. There is not sufficient evidence at the 5% significance level of an association between French and Spanish scores.

c $H_0: \rho = 0, \; H_1: \rho > 0$

Sample size = 10

Significance level = 0.05

The critical value for r for a 0.05 significance level with a sample size of 10 is $r = 0.5494$

As $0.568 > 0.5494$, r lies within the critical region, so reject H_0. There is evidence at the 5% that there is a positive correlation between the French and Spanish scores.

d Spearman's rank correlation coefficient does not use the actual data, just the ranks.

4 To find the RSS for coffee find S_{tc}.

$$S_{tc} = \sum tc - \frac{\sum t \sum c}{n} = 561.3 - \frac{88\times 32.3}{5} = -7.18.$$

Now we have enough information to calculate the RSS of coffee

$$\text{So RSS}_c = S_{cc} - \frac{(S_{tc})^2}{S_{tt}} = 0.852 - \frac{(-7.18)^2}{65.2} = 0.0613 \text{ (3 s.f.)}$$

Since $0.0524 < 0.0613$, (i.e. the RSS of ice-cream is less than the RSS of coffee), it can be concluded that ice-cream sales are more likely to have a linear model.