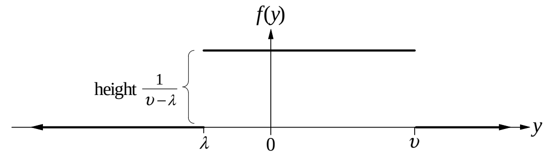
Figure 5.11. PROBABILITY MODELLING: Continuous Uniform Distributions

The continuous uniform distribution can be used (as the word *uniform* implies) to model the shape of a data set which has (approximately) the *same* frequency at *all* values; *i.e.*., a 'flat' or 'rectangular' distribution. *Un*like the normal distribution, the continuous uniform distribution is *not* often useful in practice as a probability model for the shape of a data distribution, although its *discrete* analogue is widely used as the basis for introductory probability calculations.

1. Shape: The p.d.f. of the continuous uniform distribution is *rectangular* in shape; its equation and graph (for the random variable Y) are:

$$f(y) = \begin{cases} \dfrac{1}{\upsilon - \lambda} & ; \quad \lambda < y \le \upsilon \\ 0 & ; \quad \text{otherwise} \end{cases} \qquad \text{-----(5.11.1)}$$



NOTES: 1. The notation is λ (lower case Greek lambda) for the *lower* value of the random variable at which the p.d.f. becomes non-zero, and υ (lower case Greek upsilon) for the corresponding *upper* value; elsewhere, you may see other symbols (*e.g.*, a and b) instead of λ and υ.

- If the random variable Y has a continuous uniform distribution on $(\lambda, \upsilon]$, we write: $Y \sim U(\lambda, \upsilon]$.
- The p.d.f. graph above is shown with λ negative and υ positive, but both could be negative or both positive, provided only that $\upsilon > \lambda$.

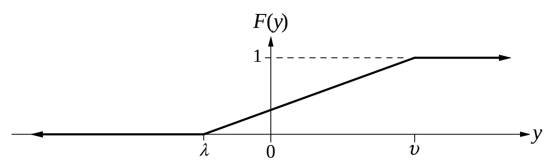
2. Like the normal distribution, the continuous uniform distribution has two parameters but, *un*like μ and σ, λ and υ are *not* the mean and standard deviation of the distribution; these two characteristics of the continuous uniform distribution are discussed below.

3. We see (from geometrical considerations) from either equation (5.11.1) or the graph given above for the p.d.f. that, as required for *any* p.d.f., the area under the continuous uniform distribution is 1 (the *normalization* requirement).

4. The p.d.f. of the continuous uniform distribution also reminds us of some properties (summarised in Table 5.9.1 at the bottom of the fourth side of Figure 5.9) of all p..d.f.s; *e.g.*, its domain is $(-\infty, \infty)$, its range is $f(y) \ge 0$ [here, specifically: $0 \le f(y) \le 1/(\upsilon - \lambda)$], and $f(-\infty) = 0$, $f(\infty) = 0$.

2. C.d.f.: The equation and graph of the (cumulative) distribution function for the continuous uniform distribution on $(\lambda, \upsilon]$ (for the random variable Y) are:

$$F(y) = \begin{cases} 0 & ; \quad y \le \lambda \\ \dfrac{y - \lambda}{\upsilon - \lambda} & ; \quad \lambda < y \le \upsilon \\ 1 & ; \quad y > \upsilon \end{cases} \qquad \text{-----(5.11.2)}$$



NOTE: 5. The c.d.f. of the continuous uniform distribution, like its p.d.f., reminds us of properties common to *all* continuous random variables; *e.g.*, the domain of the c.d.f. is $(-\infty, \infty)$, its range is $0 \le F(y) \le 1$, and $F(-\infty) = 0$, $F(\infty) = 1$. Also remember that $F(y) = \Pr(Y \le y)$.

- Also recall [from the fourth side (page 5.24) of Figure 5.9] that the c.d.f. is the *integral* of the p.d.f. and the p.d.f. is the *derivative* of the c.d.f.

3. Mean: The mean of the continuous uniform distribution on $(\lambda, \upsilon]$ (for the random variable Y) is given by:

$$\mu \equiv \mu_Y \equiv E(Y) = \int_{-\infty}^{\infty} y \cdot f(y)\,dy = \int_{\lambda}^{\upsilon} y \cdot \frac{1}{\upsilon - \lambda}\,dy = \tfrac{1}{2}(\upsilon^2 - \lambda^2) \cdot \frac{1}{\upsilon - \lambda} = \tfrac{1}{2}(\upsilon + \lambda). \qquad \text{-----(5.11.3)}$$

NOTE: 6. This result, derived by integration from the definition of the mean, could also be obtained using the symmetry of the continuous uniform p.d.f. and the fact that its centre is at $\tfrac{1}{2}(\upsilon + \lambda)$, half way along the interval where the p.d.f. is non-zero.

- By symmetry, $\tfrac{1}{2}(\upsilon + \lambda)$ is also the *median* of the continuous uniform distribution. -----(5.11.4)

4. S.d.: To find the standard deviation of the continuous uniform distribution, we must first find $E(Y^2)$ as follows:

(*continued overleaf*)

4. S.d.: (cont.)

$$E(Y^2) = \int_{-\infty}^{\infty} y^2 \bullet f(y)dy = \int_{\lambda}^{\upsilon} y^2 \bullet \frac{1}{\upsilon-\lambda}dy = \tfrac{1}{3}(\upsilon^3-\lambda^3)\bullet\frac{1}{\upsilon-\lambda} = \tfrac{1}{3}(\upsilon^2+\upsilon\lambda+\lambda^2);$$

hence: $$\sigma \equiv \sigma_Y \equiv s.d.(Y) = \sqrt{E[Y-E(Y)]^2} \equiv \sqrt{E(Y^2)-[E(Y)]^2} = \tfrac{1}{2\sqrt{3}}(\upsilon-\lambda). \qquad \text{-----(5.11.5)}$$

NOTES: 7. We see from equations (5.11.3) overleaf and (5.11.5) above for $E(Y)$ and $s.d.(Y)$ that the parameters, υ and λ, of the continuous uniform distribution determine *both* its mean and standard devation; however, it is only our choice of parameterization for the continuous uniform distribution that makes its mean and standard deviation *appear* to be dependent. Like the normal distribution, its values for centre and spread are *un*related.

8. The *variance* of the continuous uniform distribution is $\tfrac{1}{12}(\upsilon-\lambda)^2$.

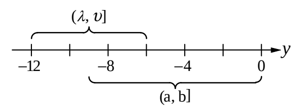
5. Probability: If the random variable $Y \sim U(\lambda, \upsilon]$, then:

$$\Pr(a < Y \le b) = \int_a^b f(y)dy = \int_a^b \frac{1}{\upsilon-\lambda}dy = \frac{b-a}{\upsilon-\lambda} = \frac{\text{length of interval } (a, b]}{\text{length of interval } (\lambda, \upsilon]}; \qquad \text{-----(5.11.6)}$$

e.g., if $(\lambda, \upsilon]$ is (–1, 9], the probability Y lies in *any* interval of length 3 within $(\lambda, \upsilon]$ is $\tfrac{3}{10} = 0.3$.

If (a, b] lies partly *out*side $(\lambda, \upsilon]$, $\Pr(a < Y \le b)$ is the length of the interval (a, b] which is *in*side $(\lambda, \upsilon]$, divided by the length of the interval $(\lambda, \upsilon]$.

e.g., if $(\lambda, \upsilon]$ is (–12, –6] and (a, b] is (–9, 0],
$\Pr(-9 < Y \le 0) \equiv \Pr(-9 < Y \le -6) = \tfrac{1}{2}$ [because $\Pr(-6 < Y \le 0) = 0$].



NOTES: 9. The expression (5.11.6) above for $\Pr(a < Y \le b)$ for the $U(\lambda, \upsilon]$ distribution shows its *interquartile range (IQR)* is $\tfrac{1}{2}(\upsilon-\lambda)$ [which is about 70% *larger* than the standard deviation]. -----(5.11.7)

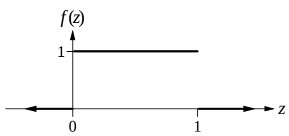
10. The expression (5.11.6) above for $\Pr(a < Y \le b)$, together with the results (5.11.3) and (5.11.5) for the mean and standard deviation of the continuous uniform distribution on $(\lambda, \upsilon]$, show that the interval $\mu \pm \sigma$ contains $1/\sqrt{3} \simeq 0.5774$ (*i.e.*, about 58%) of the area under this p.d.f.

- This result reminds us the *68–95–99.7 rule* (*e.g.*, see Figure 5.3) is *only* for the normal distribution.

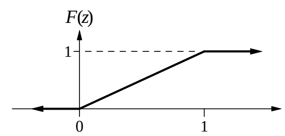
6. APPENDIX: The Continuous Uniform Distribution on (0, 1]

A notable special case of the $U(\lambda, \upsilon]$ distribution is the $U(0, 1]$ distribution; its p.d.f. and c.d.f., for the random variable Z, are:

$$f(z) = \begin{cases} 1 & ; \quad 0 < z \le 1 \\ 0 & ; \quad \text{otherwise} \end{cases}$$



$$F(z) = \begin{cases} 0 & ; \quad z \le 0 \\ z & ; \quad 0 < z \le 1 \\ 1 & ; \quad z > 1 \end{cases}$$



The mean and the median of the $U(0, 1]$ distribution are both $\tfrac{1}{2}$, its standard deviation is $\tfrac{1}{2\sqrt{3}}$ and its IQR is $\tfrac{1}{2}$. -----(5.11.8)

1 In Note 3 overleaf on page 5.29, justify the statement that the area under the p.d.f. of the $U(\lambda, \upsilon]$ distribution is 1.

2 Referring to Note 5 overleaf on page 5.29, obtain the c.d.f. of the $U(\lambda, \upsilon]$ distribution from its p.d.f. by integration, and its p.d.f. from its c.d.f. by differentiation.

3 Referring to Note 6 overleaf on page 5.29, obtain the result (5.11.4) for the *median* of the $U(\lambda, \upsilon]$ distribution from both its p.d.f. and its c.d.f.

4 For the result (5.11.5) given above for the standard deviation of the $U(\lambda, \upsilon]$ distribution, show the details of the algebraic manipulations for $E(Y^2)$ and $s.d.(Y)$.

5 Verify the result (5.11.6) for $\Pr(a < Y \le b)$ given above for the $U(\lambda, \upsilon]$ distribution.

6 In Note 9 above, verify the result (5.11.7) for the IQR of the $U(\lambda, \upsilon]$ distribution.

7 Referring to Note 10 above, find the areas under the p.d.f. of the $U(\lambda, \upsilon]$ distribution in the intervals $\mu \pm 2\sigma$ and $\mu \pm 3\sigma$.
- Compare and contrast the results for all *three* intervals with those for the normal distribution.

8 Verify the results (5.11.8) given above in the Appendix for the mean and standard deviation of the $U(0, 1]$ distribution.