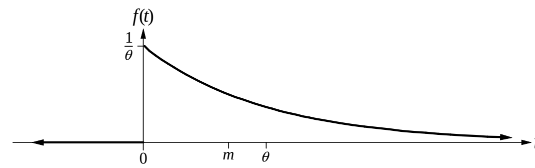
Figure 5.12. PROBABILITY MODELLING: Exponential Distributions

The exponential distribution can be used to model data sets whose distribution shape is (approximately) that of an *exponential decay.* Such a shape may arise with data sets consisting of lifetimes to failure; we have seen an example in the Boeing aircraft air conditioning system failure times in Figure 3.14.

1. Shape: The shape of the p.d.f. of the exponential distribution is that of an *exponential decay*; its equation and graph (for the random variable T) are:

$$f(t) = \begin{cases} \frac{1}{\theta}e^{-t/\theta} & ; \quad t > 0 \\ 0 & ; \quad \text{otherwise} \end{cases} \qquad \text{-----(5.12.1)}$$



NOTES: 1. Because the exponential distribution is used to model failure times, a common choice for the random variable is T; the distribution has *one* parameter, often denoted θ.

- The p.d.f. *reaches* 0 on the right only at infinity.
- If the random variable T has an exponential distribution with parameter θ, we write: $T \sim Exp(\theta)$.
- The (continuous) exponential distribution is related to a *discrete* distribution called the *Poisson distribution*, which we learn about in Part 10 of STAT 221; the parameter of the Poisson distribution is often denoted λ, and is the *reciprocal* of θ used here. In situations where the *connection* between these two distributions needs to be emphasized, you may see the exponential distribution parameterized with λ in place of $1/\theta$.

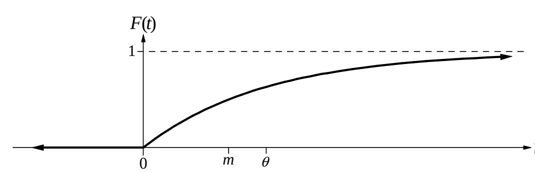
2. *Un*like the two-parameter normal and continuous uniform distributions, the exponential distribution has only *one* parameter which, as shown below, is *both* its mean *and* its standard deviation.

3. It can be shown by integration that, as required for *any* p.d.f., the area under the exponential distribution is 1 (the *normalization* requirement).

4. The p.d.f. of the exponential distribution also reminds us of some properties [summarised in Table 5.9.1 at the bottom of the fourth side (page 5.24) of Figure 5.9] of all p.d.f.s; *e.g.*, its domain is $(-\infty, \infty)$, its range is $f(t) \geq 0$ [here, specifically: $0 \leq f(t) \leq 1/\theta$], and $f(-\infty) = 0$, $f(\infty) = 0$.

2. C.d.f.: The equation and graph of the (cumulative) distribution function for the exponential distribution with parameter θ are:

$$F(t) = \begin{cases} 0 & ; \quad t \leq 0 \\ 1 - e^{-t/\theta} & ; \quad t > 0 \end{cases} \qquad \text{-----(5.12.2)}$$



NOTE: 5. The c.d.f. of the exponential 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 $0 \leq F(t) \leq 1$, and $F(-\infty) = 0$, $F(\infty) = 1$. Also remember that $F(t) = \Pr(T \leq t)$.

- The c.d.f. *reaches* 1 on the right only at infinity.
- 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 exponential distribution with parameter θ is given by:

$$\mu \equiv \mu_T \equiv E(T) = \int_{-\infty}^{\infty} t \cdot f(t)dt = \int_0^{\infty} t \cdot \tfrac{1}{\theta}e^{-t/\theta}dt = \theta. \qquad \text{-----(5.12.3)}$$

NOTES: 6. This calculation justifies the statement in Note 2 above that the parameter θ of the exponential distribution is its mean.

7. The *median* (m, say) of T is given by: $F(m) = 0.5$, which yields: $m = \theta \ln 2 \simeq 0.693\theta$; -----(5.12.4)
i.e., for the exponential distribution, the median is about 70% of the mean;
both m and θ are marked on the two diagrams at the right above.

(*continued overleaf*)

4. S.d.: To find the standard deviation of the exponential distribution, we first find $E(T^2)$:

$$E(T^2) = \int_{-\infty}^{\infty} t^2 \cdot f(t)dt = \int_0^{\infty} t^2 \cdot \tfrac{1}{\theta} e^{-t/\theta}dt = 2\theta^2; \qquad \text{hence:} \quad \sigma \equiv \sigma_T \equiv s.d.(T) = \sqrt{E(T^2) - [E(T)]^2} = \theta. \qquad \text{-----(5.12.5)}$$

NOTES: 8. This calculation justifies the statement in Note 2 overleaf on page 5.31 that the parameter θ of the exponential distribution is its standard deviation.

- The exponential distribution is unusual in that its mean and standard deviation are *equal* to each other and are both the (one) parameter of the distribution (here, θ).
  - A noteworthy contrast among the three named continuous models we have considered – the normal, continuous uniform and exponential distributions – is the difference in the degree of dependence of their respective means and standard deviations; these two characteristics are *in*dependent for the normal and continuous uniform distributions but *dependent* (*i.e.*, equal) for the exponential distribution.

9. The *variance* of the exponential distribution is θ^2.

5. Probability: If the random variable $T \sim Exp(\theta)$, then assuming b > a > 0:

$$\Pr(a < T \leq b) = \int_a^b f(t)dt = \int_a^b \tfrac{1}{\theta} e^{-t/\theta}dt = e^{-a/\theta} - e^{-b/\theta}. \qquad \text{-----(5.12.6)}$$

If a ≤ 0 and b > 0, $\Pr(a < T \leq b)$ is obtained by integrating over the interval (0, b] [because $\Pr(T \leq 0) \equiv 0$];

e.g., if the random variable $T \sim Exp(\theta = 2.5)$: $\Pr(-4 < T \leq 6) = \int_{-4}^{6} f(t)dt = \int_0^6 \tfrac{1}{\theta} e^{-t/\theta}dt = 1 - e^{-2.4} \simeq 0.9093.$

NOTES: 10. Probabilities can also be found as a difference in c.d.f values: $\Pr(a < T \leq b) = F(b) - F(a)$.

11. The first and third quartiles ($Q_{.25}$ and $Q_{.75}$, say) of the exponential distribution with parameter θ are given by: $F(Q_{.25}) = \frac{1}{4}$ and $F(Q_{.75}) = \frac{3}{4}$;
from these expressions, we find that: $IQR = \theta \ln 3 \simeq 1.099\theta$; -----(5.12.7)
[which is about 10% *larger* than the standard deviation].

12. The expression (5.12.6) above for $\Pr(a < T \leq b)$, together with the results (5.12.3) and (5.12.5) for the mean and standard deviation of the exponential distribution with parameter θ, show that the interval $\mu \pm \sigma$ contains $1 - e^{-2} \simeq 0.8647$ (*i.e.*, nearly 87%) of the area under this p.d.f.

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

1 In Note 3 overleaf on page 5.31,, justify the statement that the area under the p.d.f. of the $Exp(\theta)$ distribution is 1.

2 Referring to Note 5 overleaf on page 5.31, obtain the c.d.f. of the $Exp(\theta)$ distribution from its p.d.f. by integration, and its p.d.f. from its c.d.f. by differentiation.

3 For the result (5.12.3) given for the *mean* of the $Exp(\theta)$ distribution near the bottom of the page overleaf on page 5.31, show the details of the algebraic manipulations.

4 Referring to Note 7 overleaf on page 5.31, obtain the result (5.12.4) given for the *median* of the $Exp(\theta)$ distribution from its *p*.d.f.

- Explain briefly the reason(s) for the *relative* positions of the mean and median of the exponential distribution.

5 For the result (5.12.5) given above for the standard deviation of the $Exp(\theta)$ distribution, show the details of the algebraic manipulations for $E(T^2)$ and $s.d.(T)$.

6 Verify the result (5.12.6) for $\Pr(a < T \leq b)$ given above for the $Exp(\theta)$ distribution.

- Show how the *same* result is obtained from the *c*.d.f. (Note 10).
- Verify the result for $\Pr(-4 < T \leq 6)$ given above for the $Exp(2.5)$ distribution.

7 In Note 11 above, verify the result (5.12.7) for the IQR of the $Exp(\theta)$ distribution.

- Obtain the *same* result by integrating the *p*.d.f.

8 Referring to Note 12 above, find the areas under the p.d.f. of the $Exp(\theta)$ 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 and continuous uniform distributions.