War	Dates	Powers involved
Thirty Years' War	1618-1648	6 out of 7
Dutch War of Louis XIV	1672-1678	6 out of 7
War of the League of Augsburg	1688-1697	5 out of 7
War of the Spanish Succession	1701-1713	5 out of 6
War of the Austrian Succession	1739-1748	6 out of 6
Seven Years' War	1755-1763	6 out of 6
French Revolutionary and Napoleonic Wars	1792-1815	6 out of 6
World War I	1914-1918	8 out of 8
World War II	1939-1945	7 out of 7

Table 7.1: General wars

7 Stochastic Models of War Onset and Duration

If "general wars" are defined as "wars involving at least two-thirds of the Great Powers and an intensity exceeding 1,000 battle deaths per million population" (Levy, 1983, p. 75), then there were nine such wars during the last half-millennium of the 20th century, as indicated in Table 7.1. The Thirty Years' War was the first general war following the Treaty of Venice in 1495. Thus the times elapsing between the *onset* of one war and the next war onset are 123, 54, 16, 13, 38, 16, 37, 122 and 25 years, respectively. What can we say about the random processes that underlie these numbers?

To explore this issue further, let time to war onset be the random variable T, let G be its cumulative distribution function or cdf and let g be its probability density function or pdf. That is, define

$$G(t) \;=\; \text{Prob}(T \leq t) \;=\; \int_0^t g(\tau)\, d\tau \tag{7.1}$$

on $[0, \infty)$ with $g(t) \geq 0$ on $(0, \infty)$, so that

$$g(t) \;=\; G'(t) \tag{7.2}$$

(where a prime denotes differentiation with respect to argument),

$$G(0) \;=\; 0 \tag{7.3}$$

and

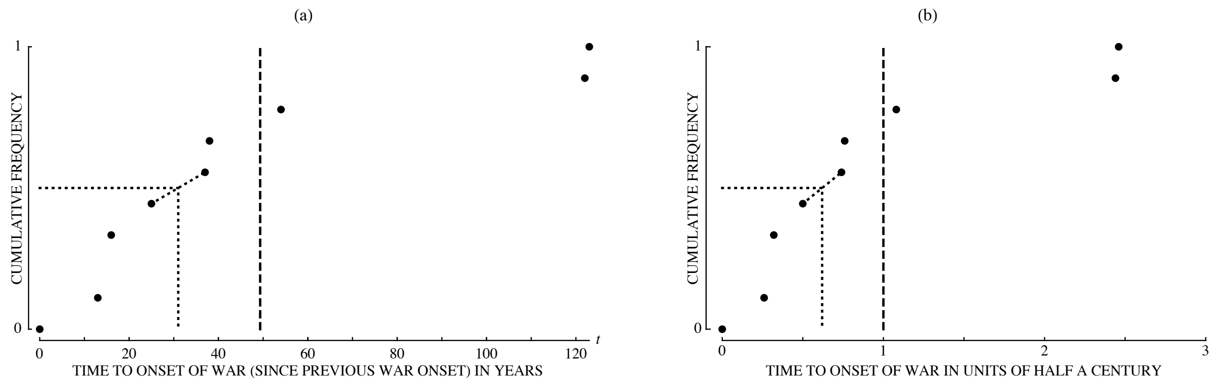
$$G(\infty) \;=\; 1. \tag{7.4}$$

We denote the mean, variance and median of this distribution by μ, σ^2 and m, respectively. That is,

$$\mu \;=\; \int\limits_0^\infty g(t)\, dt, \tag{7.5}$$

$$\sigma^2 \;=\; \int\limits_0^\infty (t-\mu)^2 g(t)\, dt \;=\; \int\limits_0^\infty t^2 g(t)\, dt - \mu^2 \tag{7.6}$$

Figure 7.1: Empirical cdf of times to war onset for general wars during the last half-millennium of the 20th century. In (a), which corresponds to Figure 2.4 of Cioffi-Revilla (1998, p. 58), time is measured in years; in (b), time is measured in units of half a century. Dashed lines indicates the mean, dotted lines the median.



and

$$G(m) \;=\; \tfrac{1}{2}. \tag{7.7}$$

Of the observed times to war onset in Table 7.1, the numbers that fail to exceed 13, 16, 25, 37, 38, 54, 122 and 123 are 1, 3, 4, 5, 6, 7, 8 and 9, respectively. Thus the proportions of onset times failing to exceed 13, 16, 25, 37, 38, 54, 122 and 123 are $\frac{1}{9}$, $\frac{1}{3}$, $\frac{4}{9}$, $\frac{5}{9}$, $\frac{2}{3}$, $\frac{7}{9}$, $\frac{8}{9}$ and 1 respectively. This—admittedly rather small—sample of observations generates an empirical cdf, which (is necessarily discrete and) is plotted in Figure 7.1.

Let us briefly digress to note that Figure 7.1 indicates the median time of onset for a general war to be 31 years. If the i-th observation in our sample of 9 onset times is denoted by t_i, then the sample mean and variance are $\bar{t} = \sum_{i=1}^{9} t_i = \frac{148}{3}$ and $s^2 = \frac{1}{8}\sum_{i=1}^{9}(t-\bar{t})^2 = 1893$, respectively.[1] Thus the mean significantly exceeds the median: the odds in favor of general war become greater than even almost two decades before war is expected. According to Cioffi-Revilla (1998, p. 67), this observation alone "can explain why highly devastating wars, such as general world wars, are said to arrive "by surprise," as many have said of World War I."

Whatever (continuous) distribution T may have, we would expect its cdf to lie close in some sense to the data points in Figure 7.1. What sort of random process might give rise to such a cdf? If we think of war as a failure of the political system, then T represents time until failure, g is its pdf and G is its cdf. We can therefore draw on reliability or survival analysis—a well established body of theory—to explore the question we have just raised. Here a central concept is the hazard rate, $h(t)$, which is the probability per unit time that a system will fail at time t, given that it has not failed before time t. The probability that the system has not failed by time t is $1-G(t)$. What is the probability that the system will fail within an infinitesimal interval of length δt containing t? On the one

[1]And yield unbiased estimates of the mean μ and variance σ^2 of the statistical population from which the sample has been drawn.

hand, it is $h(t)\delta t + o(\delta t)$.[2] On the other hand, the probability of failure within this interval is $g(t)\delta t + o(\delta t)$, and so the probability of failure within this interval, given that failure has not yet occurred, must be $\{g(t)\delta t + o(\delta t)\}/\{1 - G(t)\}$. These two expressions are equal by definition. That is,

$$h(t)\delta t + o(\delta t) \;=\; \frac{g(t)\delta t + o(\delta t)}{1 - G(t)}$$

Dividing by δt and taking the limit as $\delta t \to 0$ yields

$$h(t) \;=\; \frac{g(t)}{1 - G(t)} \tag{7.8}$$

In view of (7.2), we have

$$-h(t) \;=\; \frac{-G'(t)}{1 - G(t)} \;=\; \frac{d}{dt}\Big\{\ln(1 - G(t))\Big\}, \tag{7.9}$$

from which—in view of (7.3)—it follows more or less immediately that

$$G(t) \;=\; 1 - e^{-\int_0^t h(\tau)\, d\tau} \tag{7.10}$$

with

$$g(t) \;=\; h(t)e^{-\int_0^t h(\tau)\, d\tau} \tag{7.11}$$

by (7.2). Note that, because we assume $h(t) > 0$ for all $t > 0$, (7.4) must hold.

Different kinds of risk can now be conceptualized in terms of different h, and (7.10) can be used to deduce the associated failure distribution. For example, if h is constant, then it follows from (7.5) and (7.11) that T has an exponential distribution with mean $\frac{1}{h}$. A significantly larger class of hazard forces is encapsulated by

$$h(t) \;=\; at^b, \qquad a > 0, b > -1 \tag{7.12}$$

for which (7.11) yields

$$g(t) \;=\; at^b e^{-at^{b+1}/(b+1)}. \tag{7.13}$$

Cioffi-Revilla (1998, p. 118) refers to a as the intensity parameter and to b as the adversity parameter of the hazard force, and he allows b to take any real value. I prefer, however, to exclude the possibility that $b \in (-\infty, 1]$, partly because $b > -1$ more than suffices for our purposes, but mainly because (7.13) can then be aligned with a known probability distribution, namely, the Weibull distribution with cdf G defined on $[0, \infty)$ by

$$G(t) \;=\; 1 - e^{-(t/s)^\theta} \tag{7.14}$$

and pdf g defined on $(0, \infty)$ by

$$g(t) \;=\; \frac{\theta t^{\theta - 1} e^{-(t/s)^\theta}}{s^\theta} \tag{7.15}$$

[2]Where $o(\delta t)$ means terms so small that they tend to zero with δt even after division by δt: $\lim\limits_{\delta t \to 0} \frac{o(\delta t)}{\delta t} = 0$.

Type of hazard force	Adversity parameter	Shape parameter
Declining	$-1 < b < 0$	$0 < \theta < 1$
Constant	$b = 0$	$\theta = 1$
Increasing but decelerating	$0 < b < 1$	$1 < \theta < 2$
Increasing linearly, or steadily	$b = 1$	$\theta = 2$
Increasing and accelerating	$b > 1$	$\theta > 2$

Table 7.2: Types of hazard force

The two parameters, s and θ, both positive, are known as the scale parameter and the shape parameter, respectively.[3] Straightforward application of (7.5)–(7.6) reveals that this distribution has mean

$$\mu \;=\; \Gamma(1+1/\theta)\, s \;=\; \frac{1}{\theta}\,\Gamma(1/\theta)s, \tag{7.16}$$

variance

$$\sigma^2 \;=\; \frac{\big(2\theta\Gamma(2/\theta) - \{\Gamma(1/\theta)\}^2\big)s^2}{\theta^2} \tag{7.17}$$

and hence coefficient of variation

$$\kappa \;=\; \frac{\sigma}{\mu} \;=\; \sqrt{\frac{2\theta\Gamma(2/\theta)}{\{\Gamma(1/\theta)\}^2} - 1}. \tag{7.18}$$

Also, by (7.7), the median is

$$m \;=\; \{\ln(2)\}^{1/\theta}\, s. \tag{7.19}$$

Note that g has a singularity at $t = 0$ for $0 < \theta < 1$. Thus although the distribution is always unimodal, the mode is finite only for $\theta \geq 1$, in which case the mode is

$$t^* \;=\; \left(1 - \frac{1}{\theta}\right)^{\frac{1}{\theta}} s, \tag{7.20}$$

(it being easily verified that $g'(t^*) = 0$, $g''(t^*) < 0$ if $\theta > 1$ and $g'(t^*) < 0$ for all $t \geq 0$ if $\theta = 1$). Comparison of (7.13) with (7.15) now indicates that our alignment requires

$$a \;=\; \frac{\theta}{s^\theta}, \qquad b \;=\; \theta - 1 \tag{7.21a}$$

or

$$s \;=\; \left(\frac{1+b}{a}\right)^{\frac{1}{1+b}}, \qquad \theta \;=\; 1+b. \tag{7.21b}$$

Five different types of hazard force can now be distinguished by different values of the parameters b and θ, as indicated in Table 7.2 and illustrated by Figure 7.2. These categories are distinguished solely by the shape parameter of the associated Weibull distribution, and so its scale parameter plays only a passive role—it is merely a representative time

[3] See, e.g., McCool (2012, p. 73).

Figure 7.2: Shapes of normalized Weibull distributions corresponding to different types of hazard force (Table 7.2); κ is the coefficient of variation, and $\kappa_c = \sqrt{4/\pi - 1} \approx 0.5227$ corresponds to $\theta = 2$. (a) Constant (green) or declining (red). (b) Increasing in a steady (green) or decelerating (red) manner. (c) Increasing and accelerating.

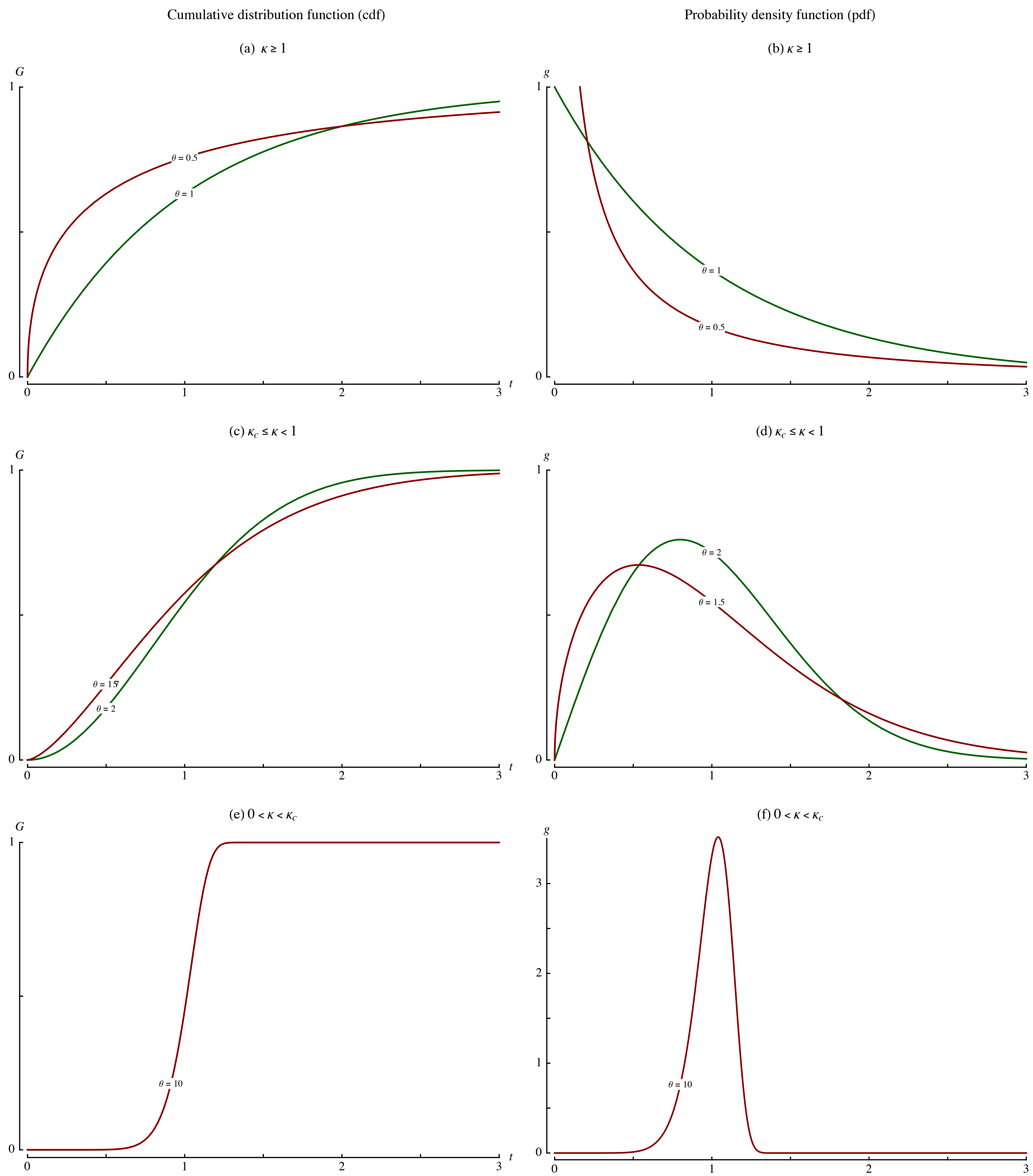
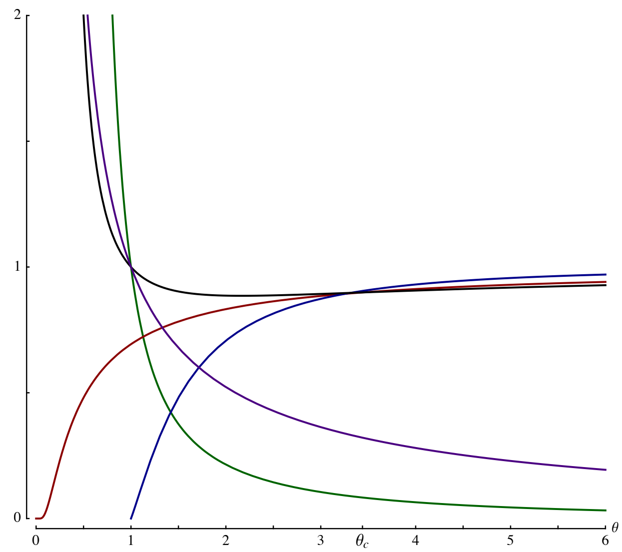


Figure 7.3: Mean $\mu = \Gamma(1/\theta)/\theta$ (black), variance $\sigma^2 = 2\Gamma(2/\theta)/\theta - \{\Gamma(1/\theta)/\theta\}^2$ (green), median $m = \{\ln(2)\}^{1/\theta}$ (red), mode $t^* = (1 - 1/\theta)^{1/\theta}$ (blue) and coefficient of variation $\kappa = \sigma/\mu$ (indigo) of a normalized Weibull distribution as a function of its shape parameter θ. Note that μ, σ^2, κ all approach ∞ as $\theta \to 0$ and μ, m, t^* all approach 1 as $\theta \to \infty$ (and $\sigma^2 \to 0$). Note also that the mean exceeds the median for $\theta < \theta_c$ but the median exceeds the mean for $\theta > \theta_c$, where $\theta_c \approx 3.44$.



scale. Indeed we can dispense with s entirely by defining dimensionless time $\hat{t} = t/s$, which in effect uses s as the unit of time. Then the random variable $\hat{T} = T/s$ has cdf $G(\hat{t}) = 1 - e^{-\hat{t}^\theta}$ with pdf $g(\hat{t}) = \theta\hat{t}^{\theta-1}e^{-\hat{t}^\theta}$. Now that we have switched to dimensionless time, however, there is little to be gained by using $\hat{t}$ in place of t. So from now on we consider only the distribution with cdf

$$G(t) \;=\; 1 - e^{-t^\theta} \tag{7.22a}$$

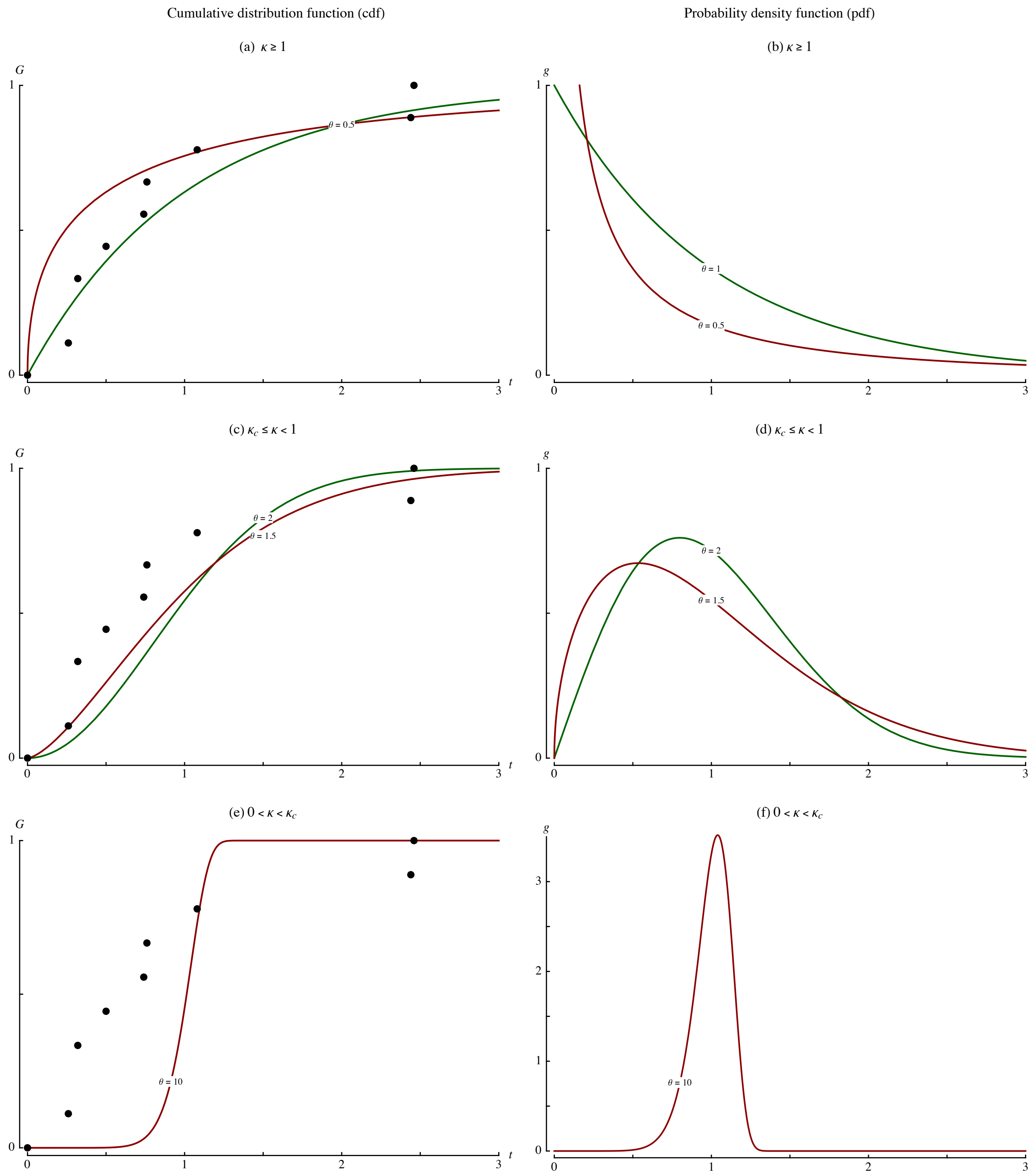
and pdf

$$g(t) \;=\; \theta t^{\theta-1}e^{-t^\theta} \tag{7.22b}$$

for which $s = 1$ in (7.14)–(7.17) and (7.19). We refer to (7.22) as the normalized Weibull distribution, and it is the distribution plotted in Figure 7.2. Its mean, variance, median, mode and coefficient of variation are plotted against θ in Figure 7.3.

As noted above, whatever distribution T may have, we would expect its cdf to lie close in some sense to the data points in Figure 7.1. Let us therefore compare Figure 7.1 with Figure 7.2, as in Figure 7.4. Figure 7.4(e) would appear to indicate strongly that the hazard forces are not increasing and accelerating. Moreover, Figures 7.4(a) and 7.4(c) suggest that the hazard forces are not even increasing, but rather are declining. In support of this interpretation, from Figure 7.1 we have already seen that the mean of the observations exceeds the median by about 59%. The normalized Weibull distribution whose mean exceeds its median by 59% has $\theta \approx 0.9278$, which falls squarely inside the declining regime (where

Figure 7.4: Comparison of Figure 7.2 with observed data in Figure 7.1.



$0 < \theta < 1$). On the other hand, we have also seen that the sample variance and mean are $s^2 = 1893$ and $\bar{t} = \frac{148}{3}$, respectively, and might therefore expect $s/\bar{t} = 3\sqrt{1893}/148 \approx 0.8819$ to be a reasonable estimate of $\kappa = \sigma/\mu$.[4] The corresponding normalized Weibull distribution has $\theta \approx 1.136$, which falls squarely inside the increasing but decelerating regime (where $1 < \theta < 2$), leading Cioffi-Revilla (1998, p. 65) to argue that the "onset of general world war during the past five hundred years has therefore been hypoexponential"—by which he means that the variance of T is lower than if it were exponentially distributed. Which of these two interpretations—if either—is correct? Certainly, they cannot both be correct!

I find it surprising that Cioffi-Revilla prefers to base his conclusion on the coefficient of variation rather than a comparison of mean and median, because he appears to stress the importance of the latter by arguing that when $\mu > m$,

> political behavior occurs in a way that is in some sense overdue according to the probabilistic odds alone, although not according to the historical record

and, conversely, when $\mu < m$,

> political political behavior occurs in a way that is premature by the odds alone, although not by the historical experience ... These concepts provide new insights and tools for understanding political uncertainty (Cioffi-Revilla, 1998, p. 67).

Nevertheless, I suspect that the answer to the question above is neither. If we were to join the dots in Figure 7.1 with a smoooth curve, sensibly (without trying to force the curve through the last dot), then we would see three inflexion points, which suggests that the underlying random process may have a bimodal distribution that cannot be modelled as a Weibull distribution, so that the class of hazard forces in Table 7.1 is not sufficiently general to encapsulate the data in Table 7.2.

Regardless, with further verbal and graphical reasoning, Cioffi-Revilla (1998) attempts to tease out further insights towards a theory of war onset and other political uncertainty from little more than the above analysis and the observations in Figure 7.1. For example, he argues that "the force or pressure for general world war" appears to remain "relatively weak and approximately constant" for a very long time, "thereafter escalating to much greater intensity in just a few years" (Cioffi-Revilla, 1998, p. 81). Again I find his argument surprising, since it would appear to contradict that the value of the adversity parameter b associated with the red curves in Figure 7.5 is less than 1, and hence corresponds to increasing but decelerating hazard forces by Table 7.2. Thus the most significant point may well be that his findings "also suggest new questions" (Cioffi-Revilla, 1998, p. 81).

Although the Weibull distribution appears to be inadequate as a model of the onset of war, the duration of war is a different matter. In this regard, two preliminary remarks are in order. First, although the two-parameter distribution defined by (7.14) or (7.15) is the most commonly used version of the Weibull distribution, there is also a less widely used three-parameter version with cdf G defined on $[l, \infty)$ by

$$G(t) \;=\; 1 - e^{-(\{t-l\}/s)^{\theta}} \tag{7.23}$$

[4] Albeit not an unbiased estimate, a well known result in statistics.

and pdf g defined on (l, ∞) by

$$g(t) \;=\; \frac{\theta(t-l)^{\theta-1}e^{-(\{t-l\}/s)^{\theta}}}{s^{\theta}} \tag{7.24}$$

where l is known as the location parameter. Second, $h(t)$ in (7.8) is fundamentally merely the probability per unit time of the next occurrence of a certain event, given that the event has not yet happened. We called h the hazard rate because an outbreak of war is a negative outcome, but the fundamental definition of h would be unchanged if T were instead the time until the next occurrence of a positive event, such as the end of a war. We might then prefer to give h a different name—for example, the hope rate—but its relationship to g and G would be unchanged. Hence if war duration has the distribution defined by (7.23) or (7.24), then it follows immediately from (7.8) that the hope rate—the probability per unit time at time t that the war will end—is

$$h(t) \;=\; \theta(t-l)^{\theta-1}/s^{\theta} \;=\; a(t-l)^{b} \tag{7.25}$$

by (7.21).

That—for all practical purposes—war duration does indeed have a three-parameter Weibull distribution was established empirically by Horvath (1968, p. 24), who in effect used data from wars between 1820 and 1949 to estimate $\theta \approx 0.66$, $s \approx (0.608)^{1/\theta} \approx 15$ and $l \approx 0.05$ (years). By (7.21), $b \approx -0.34$. Hence b is negative. The significance of this fact was noted by Cioffi-Revilla (1989, p. 571): because $b < 0$ makes h a decreasing function, it means that wars have a declining propensity to terminate, once they have begun—at least for wars between 1820 and 1949. But it also seems to be true of more recent war, especially civil war (Fearon and Laitin, 2003; Hironaka, 2005), and squares with Bellany's (1999) contention that prolongation and stalemate are the default state of modern war (as discussed in Lecture 6).

If any of you is sufficiently interested, perhaps an exploration of Cioffi-Revilla's "new questions" could form the basis of a more in-depth investigation leading to an end-of-term presentation. For the rest of us, however, this is as far as we go for now!

Figure 7.5: Comparison of observations in Figure 7.1 with cdf (left) and pdf (right) of the normalized Weibull distributions for which $\theta \approx 0.9278$, $\kappa = 1.079 > 1$ (green) and $\theta \approx 1.136$, $\kappa = 0.8819 < 1$ (red).