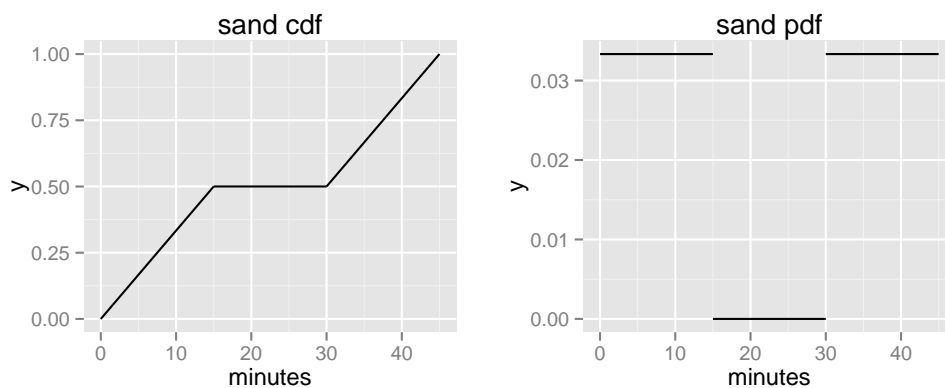# Assignment 2

Eric Mittman

January 28, 2016

2.2. (a) A container of sand containing buried object of interest is being poured out. When half of the sand has been poured, the pouring stops to allow search for the object. The "time to failure" is the time at which the object is emitted from the container.

(b)

(c) Since half the sand has been poured, defining $t_{.5}$ to be the 15 minutes, the first time at which half of the sand has been poured is sensible because that is the probability that the object is in the poured sand at 15 minutes.

An alternative definition would be that the pth quantile is the largest value of t such that $F(t) \leq p$. This would lead to $t_{.5} = 30$ minutes.
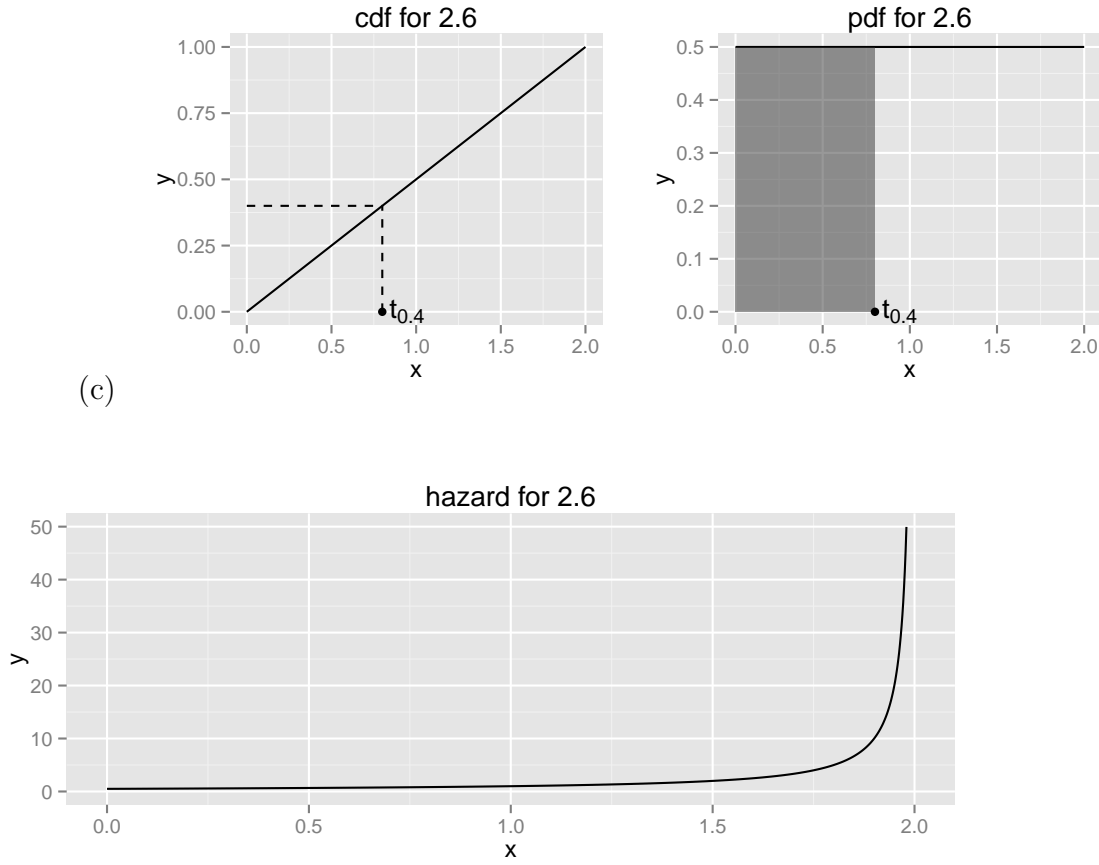
2.6. (a)

$$f(t) = 1/2, \quad t \in (0, 2)$$
$$h(t) = 1/(2 - t), \quad t \in (0, 2)$$

(b)

$$1 - \exp\left[-\int_0^t \frac{1}{2-x} dx\right] = 1 - \exp\left[\int_2^{2-t} \frac{1}{u} du\right] = 1 - \exp\left[\log\left(\frac{2-t}{2}\right)\right] = t/2 = F(t)$$

cdf for 2.6

pdf for 2.6

hazard for 2.6

Figure 1: The hazard function goes to infinity at $x = 2$ because as the cdf reaches 1 at 2. As the remaining time to fail decreases to zero, then the conditional rate of failure must increase without bound.

(d)

(e)

$$F(t) = p \iff t = 2p \implies t_{.4} = .8$$

(f) Since the pdf is constant over the support of T,

$$Pr(.1 < T \le .2) = P(.8 < T \le .9) = 1/2 \cdot .1 = .05$$

(g)

$$Pr(.1 < T \le .2 | T > .1) = \frac{.05}{1 - F(.1)} = \frac{.05}{.95} = \frac{1}{19}$$

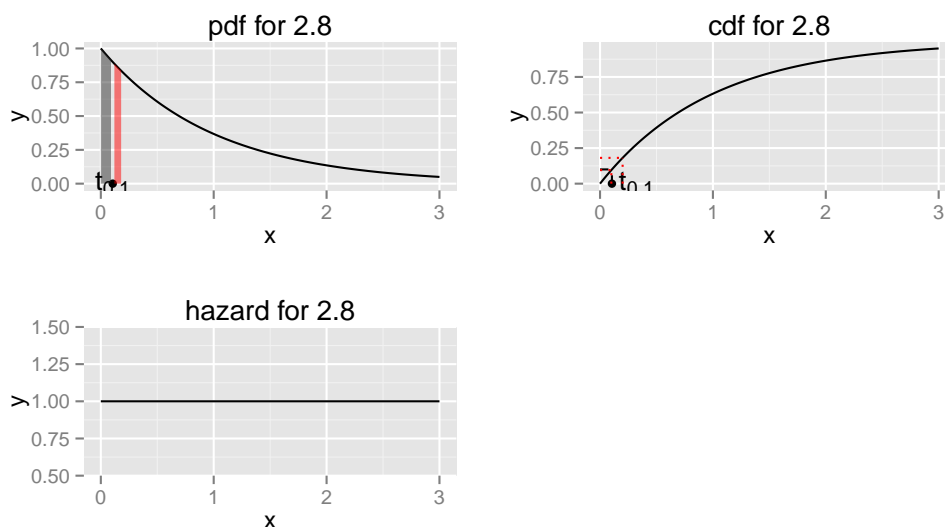$$Pr(.8 < T \le .9 | T > .8) = \frac{.05}{1 - F(.8)} = \frac{.05}{.6} = \frac{1}{12}$$

These are the same as what we would get by the approximation in (2.1):

$$h(.1) \times .1 = 1/19, \quad h(.8) \times .1 = 1/12$$

(h) The approximation is exact in this case because the density is constant. If the pdf was decreasing on $(t, t + \triangle t)$, then (2.1) would be biased high. It would be biased low if the pdf was increasing on the interval.

2.8 (a)

$$f(t) = e^{-t}, \quad h(t) = \frac{e^{-t}}{e^{-t}} = 1$$

(b)

(c)

$$F(t_p) = p \iff 1 - p = \exp(-t_p) \iff t_p = -log(1 - p)$$

So $t_{.1} = -log(.9) \approx .10536$.

(d)

$$F(.2) - F(.1) = e^{-.1} - e^{-.2} \approx 0.08611$$

Using (2.1) we get

$$h(.1) \times .1 \times (1 - F(.1)) = .1 \times e^{-.1} = 0.09048,$$

which is an overapproximation, as discussed in ex. 2.6.

2.18(d) From (2.6),

$$1 - p_i = Pr(T \leq t_{i-1} \vee T > t_i | T > t_{i-1}) = Pr(T > t_i | T > t_{i-1})$$

Since $t_0 = 0$ and the support of T is $(0, \infty)$, $1 - p_1 = Pr(T > t_1)$. It follows from the definition of conditional probability that

$$\prod_{j=1}^{i} (1 - p_j) = Pr(T > t_1, ..., T > t_i) = Pr(T > t_i) = S(t_i)$$

2.15 (a) Selecting the battery at random is fair.

(b) Selecting the battery with the most running time is fair.

(c) Selecting the battery with the least running time is fair.

(d) Selecting the battery with the lowest measured capacity is unfair.

   The last choice is unfair, because this battery is likely to fail sooner than other batteries. This strategy would bias analysis toward larger failure times.

   Among the fair choices, selecting the oldest battery would be reasonable if infant mortality is of primary interest. Selecting the newest battery would be better if wear out is a focus of the experiment. Selecting a battery at random would lead to a wider variety of censoring times and might be a good choice if both wear out and infant mortality are interesting.

3.1 (a)
$$\hat{F}(75) = 15/23 \approx 0.65217$$

(b)
$$L = \left\{ 1 + \frac{(23 - 15 + 1) \times 1.960116}{15} \right\}^{-1} = 0.45956$$

$$U = \left\{ 1 + \frac{23 - 15}{(15 + 1) \times 2.183212} \right\}^{-1} = 0.81366$$

(c)
$$L = 15/23 - 1.644854 \times \sqrt{(\frac{15/23 \times 8/23}{23})} = 0.48882$$

$$U = 15/23 + 1.644854 \times \sqrt{(\frac{15/23 \times 8/23}{23})} = 0.815565$$

(d) The intervals are similar, with the primary difference being that conservative estimate is a bit wider overall, reaching a bit further to the low side. There seems to be enough data in this instance to consider the normal approximation to be adequate. However, given that the conservative estimate is close, why not use it since it guarantees the nominal coverage?