Math 170E

Lecture Notes Section 3.4 [*,†]

Weibull, Gompertz, and mixed-type RVs

Instructor: Swee Hong Chan

NOTE: Materials that appear in the textbook but do not appear in the lecture notes might still be tested. Please send me an email if you find typos.

---

*Version date: Friday 5th February, 2021, 15:08.

†This notes is based on March Boedihardjo's and Jamie Haddock's notes from the past quarters, and I would like to thank them for their generosity. "*Nanos gigantum humeris insidentes* (I am but a dwarf standing on the shoulders of giants)".

1 Weibull RVs: definition

The **Weibull random variable** X with parameter α and β is the continuous random variable with the cumulative distribution function

$$F(x) \quad = \quad P(X \leq x) \quad = \quad \begin{cases} 1 - \exp\left[-\left(\frac{x}{\beta}\right)^{\alpha}\right] & \text{if } x \geq 0; \\ 0 & \text{if } x < 0. \end{cases}$$

Equivalently, we have

$$P(X > x) \quad = \quad \exp\left[-\left(\frac{x}{\beta}\right)^{\alpha}\right] \qquad \text{for } x \geq 0.$$

Weibull random variable X is **usually** the time it takes until an individual dies, until a product is broken, etc.

2 Weibull RVs: Example

A laptop has life expectancy X given by Weibull distribution with $\alpha = 2$ and $\beta = 10$. Compute the probability

1. The laptop lives less than five months.

2. The laptop dies this month, given that it is 5 months old.

3. The laptop dies this month, given that it is 7 years old.

Answer: For the first part,

$$P(X \leq 5) = 1 - \exp\left[-\left(\frac{5}{10} \right)^2 \right] = 1 - e^{-1/4} \approx 0.2212.$$

For the second and third part,

$$
\begin{aligned}
P(X \leq 6 \mid X \geq 5) &= \frac{P(5 \leq X \leq 6)}{P(X \geq 5)} \\
&= \frac{\left(\left(1 - \exp\left[-\left(\frac{6}{10}\right)^2\right]\right) - \left(1 - \exp\left[-\left(\frac{5}{10}\right)^2\right]\right)\right)}{\left(\exp\left[-\left(\frac{5}{10}\right)^2\right]\right)} \\
&\approx 0.104. \\
P(X \leq 85 \mid X \geq 84) &= \frac{P(84 \leq X \leq 85)}{P(X \geq 84)} \\
&= \frac{\left(\left(1 - \exp\left[-\left(\frac{85}{10}\right)^2\right]\right) - \left(1 - \exp\left[-\left(\frac{84}{10}\right)^2\right]\right)\right)}{\left(\exp\left[-\left(\frac{84}{10}\right)^2\right]\right)}
\end{aligned}
$$

$\approx 0.8154.$

Hence, a 7 year-old laptop is much more likely to die than 5-month-old laptop. This is what we expect in real life (older laptops are more likely to die due to wear-and-tear).

3 Weibull RVs: pdf, mean, variance

Theorem 1. *The random variable X with Weibull distribution with parameter α, β has probability density function*

$$f(x) = \frac{\alpha x^{\alpha-1}}{\beta^{\alpha}} \exp\left[-\left(\frac{x}{\beta}\right)^{\alpha}\right].$$

The mean and the variance is given by

$$\begin{aligned} \mu &= \beta \Gamma\left(1 + \frac{1}{\alpha}\right); \\ \sigma^2 &= \beta^2 \left[\Gamma\left(1 + \frac{2}{\alpha}\right) - \left[\Gamma\left(1 + \frac{1}{\alpha}\right)\right]^2\right], \end{aligned}$$

where Γ is the gamma function.

Please check the textbook for proofs.

4 Gompertz RVs: Definition

The random variable X with **Gompertz distribution** and parameter a and b is the continuous RV given by

$$\begin{aligned} S &= [0, \infty); \\ f(x) &= ae^{bx} \exp\left[-\frac{a}{b}e^{bx} + \frac{a}{b}\right] \qquad \text{for } x \geq 0. \end{aligned}$$

The cumulative distribution function is given by

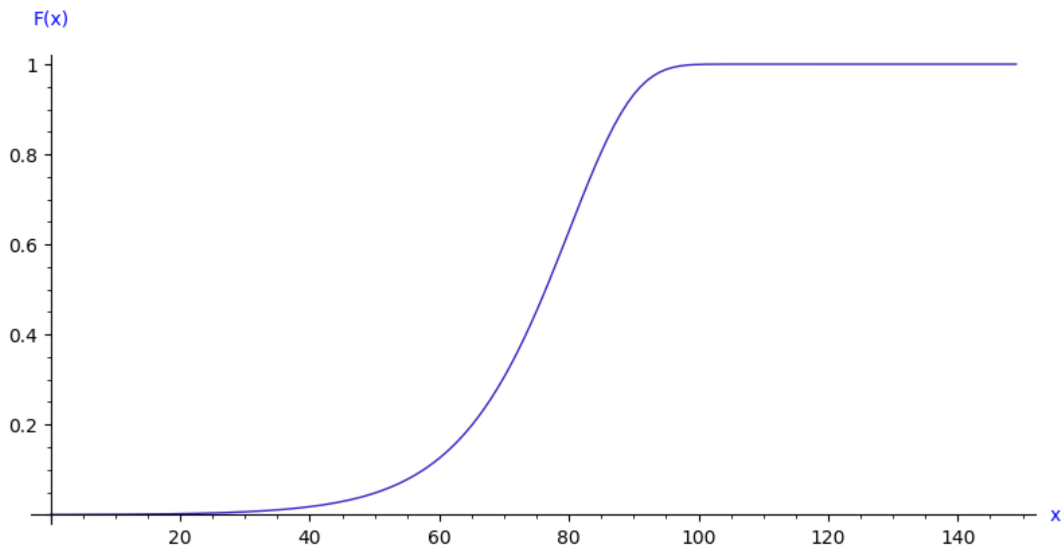
$$F(x) = P(X \leq x) = \begin{cases} 1 - \exp\left[-\frac{a}{b}e^{bx} + \frac{a}{b}\right] & \text{if } x \geq 0; \\ 0 & \text{if } x < 0. \end{cases}$$

Equivalently, we have

$$P(X > x) = \exp\left[-\frac{a}{b}e^{bx} + \frac{a}{b}\right] \qquad \text{for } x \geq 0.$$

5 Gompertz RVs: Example

A mortal human's length of life is modeled by Gompertz distribution with $a = 1/30000$ and $b = 1/10$. Here is a figure for the cumulative distribution function.



This **intuitively** tells us that it is very unlikely to live past 105 years old.

Let's check the figure of the probability density function to get a better picture.

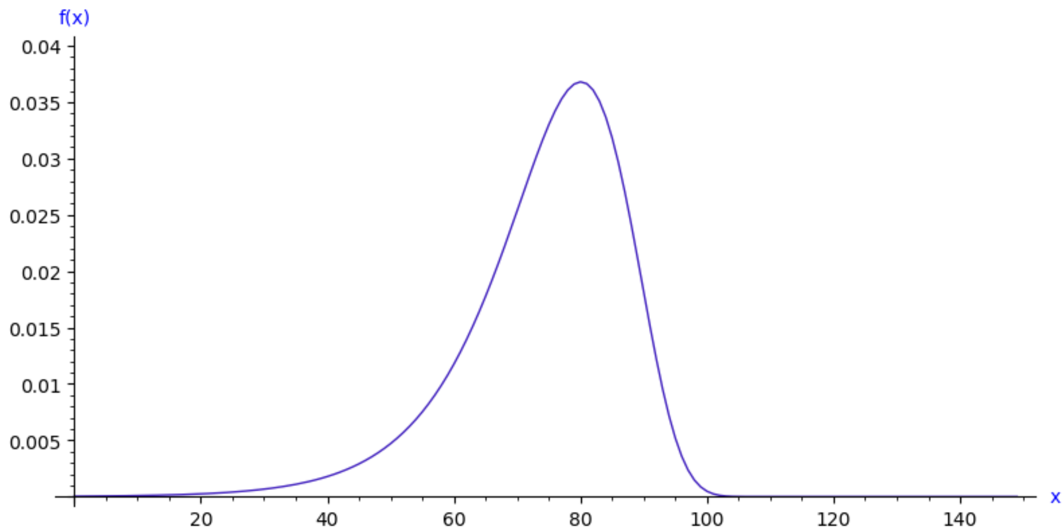


Figure 1: The probability density function of a human's length of life given by Gompertz distribution.

Intuitively,this tells us that most humans won't die before they are 20 years old.

As they grow older they are more likely to die.

Most people will probably die between the age of 60 and 90 years old.

This is what we would expect for human's life expectancy in real life.

6 Weibull RVs vs Gompertz RVs

The **Weibull RVs** has failure rate given by **polynomial function** $\left(\frac{x}{\beta}\right)^{\alpha}$.

The **Gompertz RVs** has failure rate given by **exponential function** ae^{bx}.

In particular, **Gompertz RVs** is often used to model **human's lifespan**, as exponential failure rate is more suitable for that purpose.

7 Mixed type RVs: Example 1

Consider the following game: A fair coin is tossed.

- If the outcome is head, the player is assigned to Santa Claus, who gives the player 2 USD.
- If the outcome is tail, the player is assigned to Uncle Scrooge McDuck, who picks a random number from uniform distribution on $[0, 1]$, and then gives the player this amount of money.

Let X be the amount of money received by the player. Compute the support and the cdf of X.

8 Mixed type RVs: Answer

The support and the cdf of X are given by

$$S = [0, 1] \cup \{2\};$$

$$F(x) = \begin{cases} 0, & \text{if } x < 0; \\ \frac{x}{2}, & \text{if } 0 \le x < 1; \\ \frac{1}{2}, & \text{if } 1 \le x < 2; \\ 1, & \text{if } 2 \le x. \end{cases}$$

This RV is **not discrete** since the support is uncountable.

This RV is **not continuous** since the cdf is not continuous (see picture).

9 Mixed type RVs: Definition

A random variable is of **mixed type** if it is neither discrete RV nor continuous RV.

Note: The **cumulative distribution function** $F : (-\infty, \infty) \to [0, 1]$ is always defined as

$$F(x) \quad := \quad P(X \leq x),$$

regardless if X is discrete, continuous, or mixed.

Note: The cumulative distribution function of a mixed type random variable will always be discontinuous.

Note: It does not make sense to talk about probability mass function or probability density function anymore, as these two notions are special to discrete and continuous type, respectively.

Note: In order to find the expectation of a mixed type random variable, we need to combine both Riemann sums and integrals.

10 Mixed type RVs: Example 2

Compute the mean and the variance of X from the previous example.

Answer: Let Y_1 be the coin toss in the beginning, and let Y_2 be the uniform random variable of Uncle Scrooge McDuck. Then the mean is

$$
\begin{aligned}
\mu &= E[X] \\
&= E[X \text{ when } Y_1 \text{ is head}] + E[X \text{ when } Y_1 \text{ is tail}] \\
&= E[2 \text{ when } Y_1 \text{ is head}] + E[Y_2 \text{ when } Y_1 \text{ is tail}] \\
&= E[2]\, P(Y_1 \text{ is head}) \;+\; E[Y_2]\, P(Y_1 \text{ is tail}) \\
&= \frac{1}{2}(2) + \frac{1}{2}\int_{-\infty}^{\infty} x\, f(x)\, dx \\
&= 1 + \frac{1}{2}\int_0^1 x\, dx \quad = \quad 1 + \frac{1}{2}\left(\frac{1}{2}\right) \quad = \quad \frac{5}{4}.
\end{aligned}
$$

The second moment is

$$
\begin{aligned}
E[X^2] \quad &= \quad E[X^2 \text{ when } Y_1 \text{ is head}] + E[X^2 \text{ when } Y_1 \text{ is tail}] \\
&= \quad E[2^2 \text{ when } Y_1 \text{ is head}] + E[Y_2^2 \text{ when } Y_1 \text{ is tail}] \\
&= \quad E[2^2]\, P(Y_1 \text{ is head}) \; + \; E[Y_2^2]\, P(Y_1 \text{ is tail}) \\
&= \quad \frac{1}{2}\,(2^2) \; + \; \frac{1}{2}\int_{-\infty}^{\infty} x^2\, f(x)\, dx \\
&= \quad 2 + \frac{1}{2}\int_0^1 x^2\, dx \\
&= \quad 2 + \frac{1}{2}\left(\frac{1}{3}\right) \quad = \quad \frac{13}{6}.
\end{aligned}
$$

The variance is then equal to

$$
\sigma^2 \quad = \quad E[X^2] - (E[X])^2 \quad = \quad \frac{13}{6} - \left(\frac{5}{4}\right)^2.
$$

11 Mixed type RVs: Example 3

Friendly Instructor Factory has suffered some damages from a recent forest fire.

Let X be the loss of money due to the forest fire in millions of dollars, which is an exponential random variable with mean 1.

Fortunately, the wise CEO has insured the factory before the forest fire.

Suppose that the insurance will cover losses between 2 million USD and 10 million USD.

Calculate the expected reimbursement paid by the insurance.

12 Mixed type RVs: Answer 3

Let Y be the reimbursement paid by the insurance. Then

$$Y = \begin{cases} 0 & \text{if } X < 2; \\ X & \text{if } 2 \le X \le 10; \\ 10 & \text{if } X > 10. \end{cases}$$

The mean is then equal to

$$\begin{aligned} & E[Y] \\ = \quad & E[0 \text{ when } X < 2] + E[X \text{ when } 2 \le X \le 10] \\ & + E[10 \text{ when } X > 10] \\ = \quad & (0) + \left(\int_2^{10} x e^{-x}\, dx \right) + \left(\int_{10}^{\infty} 10\, e^{-x}\, dx \right) \\ = \quad & \left[-x e^{-x} - e^{-x} \right]_2^{10} + \left[-10\, e^{-x} \right]_{10}^{\infty} \qquad \text{(how?)} \\ = \quad & (-11 e^{-10} + 3 e^{-2}) + 10\, e^{-10}. \end{aligned}$$

This phenomenon where the random variable is cut when it is below and above a certain threshold is known as **censoring**, which is often used in insurance industry or in experiments to exclude unlikely scenarios.

13 Random variables: Another example

Let X be a continuous random variable with

$$\text{support } [0, \infty); \qquad \text{pdf } f(x), \qquad \text{cdf } F(x).$$

Let N be a fixed nonnegative real number, and let Y be a random variable given by

$$Y = \begin{cases} 3N & \text{if } X > N; \\ 4X - N & \text{if } X \leq N. \end{cases}$$

Show that $E[Y]$ is maximized when $F(N) = \frac{3}{4}$.

14 Random variables: Answer

We first compute $E[Y]$: We have

$$
\begin{aligned}
E[Y] \quad &= \quad E[Y; X \le N] + E[Y; X > N] \\
&= \quad E[4X - N; X \le N] + E[3N; X > N] \\
&= \int_0^N (4x - N) f(x)\, dx \,+\, \int_N^\infty 3N f(x)\, dx. \\
&= \quad 4\int_0^N x f(x)\, dx \,-\, N\int_0^N f(x)\, dx \,+\, 3N\int_N^\infty f(x)\, dx \\
&= \quad 4\int_0^N x f(x)\, dx \,-\, N F(N) \,+\, 3N(1 - F(N)) \\
&= \quad 4\int_0^N x f(x)\, dx \,+\, 3N \,-\, 4N F(N).
\end{aligned}
$$

We now take the derivative of $E[Y]$:

$$\begin{aligned}
&\frac{\partial}{dN}E[Y] \\
= \quad & \frac{\partial}{dN}4\int_0^N xf(x)\,dx + \frac{\partial}{dN}3N - \frac{\partial}{dN}4NF(N).
\end{aligned}$$

Now note that, by the fundamental theorem of Calculus:

$$\frac{\partial}{dN}4\int_0^N xf(x)\,dx \quad = \quad 4\,N\,f(N).$$

Also note that, by the chain rule:

$$\frac{\partial}{dN}4NF(N) \;=\; 4F(N) \;+\; 4N\frac{\partial}{dN}F(N) \;=\; 4F(N) + 4Nf(N).$$

Hence we have

$$\begin{aligned}
\frac{\partial}{dN}E[Y] \quad &= \quad (4Nf(N)) \;+\; (3) \;-\; (4F(N) + 4Nf(N)) \\
&= \quad 3 - 4F(N).
\end{aligned}$$

We now take the second derivative of $E[Y]$:

$$\frac{\partial^2}{(dN)^2} E[Y] = \frac{\partial}{dN} 3 - 4F(N) = -4f(N) \leq 0.$$

Hence the first derivative of $E[Y]$ is equal to 0 when $N = F^{-1}(\frac{3}{4})$, and the second derivative of $E[Y]$ is always nonnegative.

Hence we conclude that $N = F^{-1}(\frac{3}{4})$ is the maximizer for the function $E[Y]$.