STAT 417: Survival Analysis Methods
Homework 2: Due Wednesday 1-22-25 at 11:59 PM

Note: Unless a problem begins with "[Minitab]," show all mathematical steps. Minitab instructions (Graphing Functions of Survival Time and Computing Probabilities) are available in the Week 2 module on Canvas.

---

1. Bus Waiting Time (Uniform Distribution)

Suppose that T denotes the waiting time for a bus to arrive at a particular stop and follows a uniform distribution on the interval 0 to 10 minutes with the density function:

$$f(t) = \begin{cases} 1/15 & \text{for } 0 \le t \le 15 \\ 0 & \text{otherwise} \end{cases}$$

a) Derive the cumulative distribution function F(t). Be sure to include the values that t can take in your answer.

F(t) = **t / 15, 0 ≤ t ≤ 15**

0, t < 0

1, t > 15

b) Derive the survival function S(t). Be sure to include the values that t can take in your answer.

S(t) = 1 - F(t) = **1 - (t / 15), 0 ≤ t ≤ 15**

1, t < 0

0, t > 15

c) Find the probability that an individual will wait longer than 5 minutes for the bus to arrive.

P(t > 5) = S(5) = 1 - (5 / 15) = 2 / 3 = **0.6667**

---

2. Light Bulb Lifetime (Exponential Distribution)

The lifetime of light bulbs follows an exponential distribution with λ=1200 See notes for the form of the pdf).

a) Derive the cumulative distribution function F(t) (show your work).

$$F(t) = \int_{-\infty}^{t} f(y)\, dy = \int_{0}^{t} (1 / 1200) * \exp[-y / 1200]\, dy$$

$$= (1 / 1200) \int_{0}^{t} \exp(-y / 1200)\, dy$$

$$= (1 / 1200) * [-1200 * \exp(-y / 1200)]_{0}^{t}$$

$$= -\exp(-t / 1200) - (-\exp[-0 / 1200])$$

$$= \mathbf{1 - \exp(-t / 1200), t \geq 0}$$

b) Derive the survival function S(t).

$$S(t) = 1 - F(t) = \mathbf{\exp(-t / 1200),\ \ t \geq 0}$$

c) Compute the hazard function h(t).

$$h(t) = f(t) / S(t) = [(1 / 1200) * \exp(-t / 1200)] / \exp(-t / 1200)$$

$$= \mathbf{1 / 1200}$$

d) Find the probability that a randomly selected light bulb will last longer than 1300 hours.

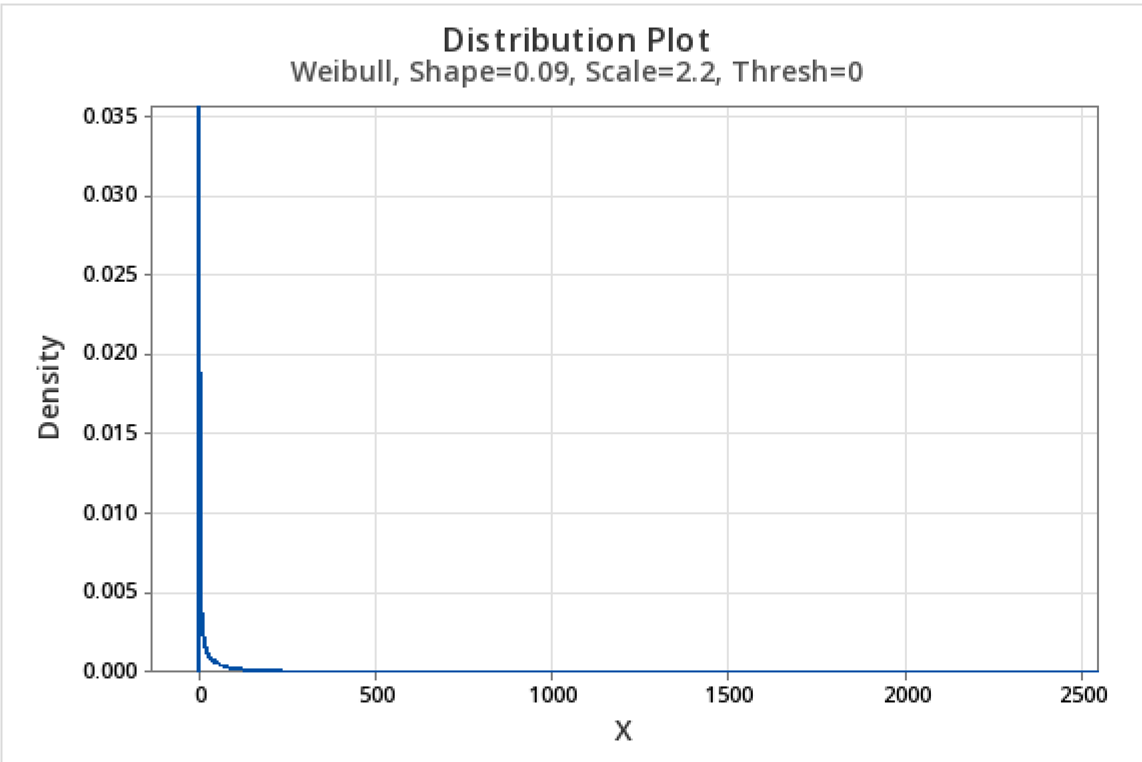
$$P(t > 1300) = S(1300) = \exp(-1300 / 1200) = \mathbf{0.3384}$$

---

3. Tumor Development (Weibull Distribution)

[Minitab] Suppose that the time in days to development of a tumor for rats exposed to a carcinogen follows a Weibull distribution with parameters λ=2.2 and β=0.09.

a) Graph the probability density function for the time-to-event random variable. Submit this graph with your assignment.



b) Find the probability that a rat will be tumor-free beyond 20 days.

$P(t > 20) = S(20) =$ **0.2953**

c) Find the probability that a rat develops a tumor between the 20th and 30th days.

$P(20 < t < 30) = F(30) - F(20) =$ **0.01309**

---

4. Bus Waiting Time (Hazard Function)

The waiting time for a bus to arrive at a particular stop has a uniform distribution on the interval 0 to 10 minutes with density function:

$$f(t) = \begin{cases} 1/10 & \text{for } 0 \leq t \leq 10 \\ 0 & \text{otherwise} \end{cases}$$

a) Find the hazard function h(t) using the expression h(t)= f(t) / S(t).

$h(t) = [1 / 10] / [1 - (t / 10)] = [1 / 10] / [(10 - t) / 10)]$

$= \mathbf{1 / (10 - t)}$

b) Find the cumulative hazard function H(t) using the expression $H(t) = \int_0^t h(y)dy$.

$H(t) = \int_0^t 1 / (10 - y)\, dy$

$u = 10 - y;\ du = - dy$

$= - \int_0^t 1 / u\, du$

$= - [\ln(u)]_0^t$

$= - [\ln(10 - t) - \ln(10)]$

$= \mathbf{-\ln[(10 - t) / 10],\ 0 \leq t \leq 10}$

c) Find the average time that an individual will wait for a bus at this stop.

$E(T) = \int_0^{10} t * f(t)\, dt = \int_0^{10} t / 10\, dt$

$= [t^2 / 20]_0^{10}$

$= (10^2 / 20) - (0 / 20)$

$= \mathbf{5}$

---

5. Battery Life (Pareto Distribution)

A random variable T that follows a Pareto distribution with parameters $\theta > 0$ and $\lambda > 0$ has the pdf given by:

$$\boxed{f(t) = \frac{\theta\lambda^\theta}{t^{\theta+1}}, \quad t \geq \lambda}$$

Suppose the battery life of an internal pacemaker (in years) follows a Pareto distribution with $\theta=2$ and $\lambda=4$. Note: when solving integrals, keep in mind that $t \geq \lambda$.

a) Derive the survival function S(t).

$f(t) = 2(4^2) / t^{2+1} = 32 / t^3, t \geq 4$

$S(t) = 1 - F(t)$

$F(t) = \int_4^t f(y)\, dy = \int_4^t 32 / y^3\, dy$

$= 32 \int_0^t 1 / y^3\, dy$

$= 32[1 / -2y^2]_4^t$

$= 32[1 / -2t^2 - (1 / -2(4^2))]$

$= 1 - (16 / t^2)$

$= 1 - [1 - (16 / t^2)]$

$= \mathbf{16 / t^2, t \geq 4}$

$\mathbf{1, \quad t < 4}$

b) Derive the cumulative hazard function H(t).

$H(t) = - \ln[S(t)] = - \ln(16 / t^2) = \mathbf{\ln(t^2 / 16), t \geq 4}$

$\mathbf{0, \quad t < 4}$

c) Find the mean time to battery failure. NOTE: Use the general definition for expectation to find E(T), and be careful about limits of integration.

$$E(T) = \int_4^\infty t * f(t)\, dt = \int_4^\infty 32 / t^2\, dt$$

$$= [-32 / t]_4^\infty$$

$$= (-32 / \infty) - (-32 / 4)$$

$$= 0 + 6$$

$$= \mathbf{8\ years}$$

d) What is the average remaining lifetime of batteries still working after 7 years?

$$mrl(7) = \int_7^\infty S(t)\, dt / S(7) = (\int_7^\infty 16 / t^2\, dt) / (16 / 7^2)$$

$$= [-16t^{-1}]_7^\infty / (16 / 7^2)$$

$$= (16 / 7) / (16 / 7^2)$$

$$= \mathbf{7\ years}$$

e) If the battery is scheduled to be replaced when only 5% of all batteries are still working, then find this time.

$$S(t_{95}) = 0.05$$

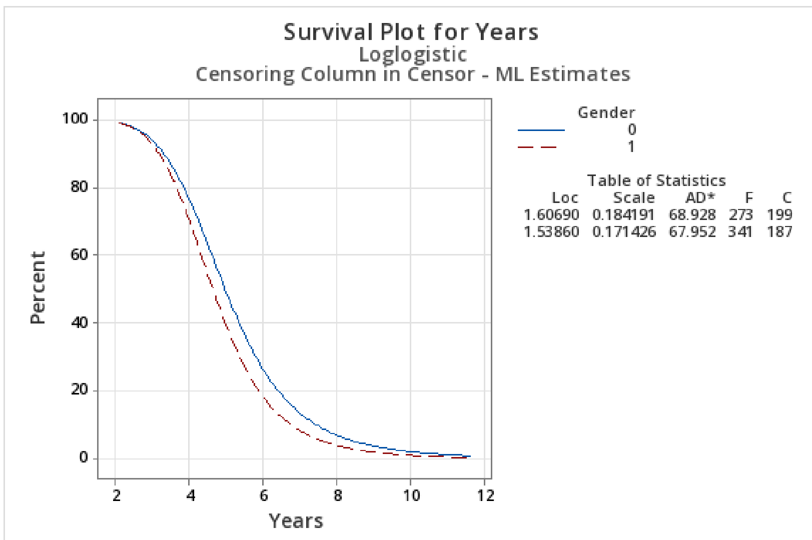
$$16 / t_{95}^2 = 0.05$$

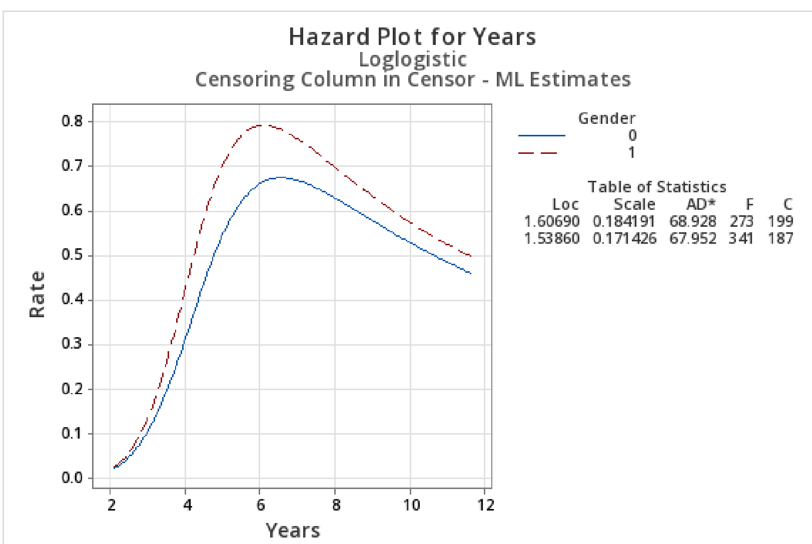
$$t_{95} = \sqrt{(320)} = \mathbf{17.8885\ years}$$

6. Graduation Time (Log Logistic Distribution)

[Minitab] The Minitab file "Graduate" in the Homework 2 folder contains the time (in years) that 1000 students (472 males and 528 females) took to graduate (obtain a bachelor's degree) from college (measured from the time they entered a post-secondary institution, i.e. either a junior college or four year degree granting institution). The Gender column contains the gender of each student, and Censor contains the values of the censoring status variable.

a) Assuming the time-to-event random variable follows a log-logistic distribution, construct plots of the survival functions for male and female students (on the same graph), and construct plots of the hazard functions for males and females (on the same graph). These graphs are to be submitted.





b) Examine the survival curves, and briefly compare the survival experiences of males and females.

> Male students tend to take longer to graduate from college (have a higher survival rate) than female students for any time t.

c) Examine the hazard curves for male and female students, and briefly compare their hazard experiences. In particular, note that the hazard curves decrease after some number of years - how would you explain why the hazard decreases?

> The hazard rate for female students is greater than that for male students for any time t. Both genders rates increase, however the hazard curve for females begins decreasing at around 6 years while that for males begins decreasing around 6.3 years. The hazard might begin decreasing after around 6 years because a student who is in college for this long tends to complete their degree rather than drop out.