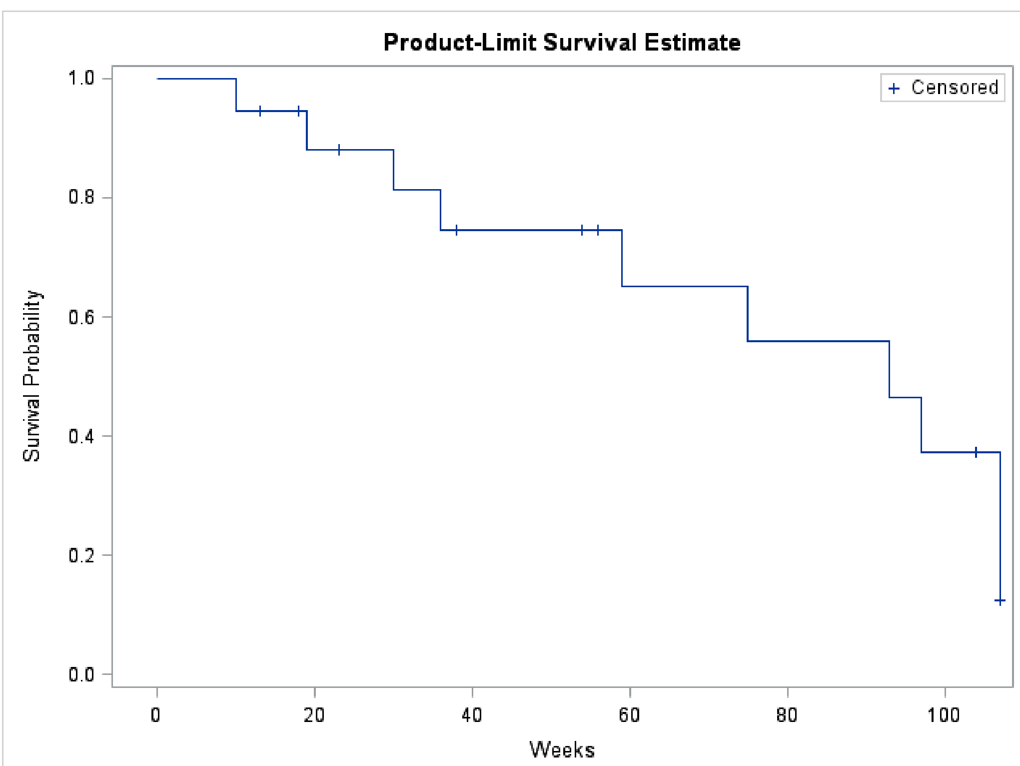
BST 665 - Survival Analysis
Homework 1 - Solution

1. A.

Survival Time (t_i)	Number at Risk (n_i)	Deaths (d_i)	$\hat{S}(t)$
10	18	1	17/18 = 0.944
13	17	0	0.944 · 1 = 0.944
18	16	0	0.944 · 1 = 0.944
19	15	1	0.944 · 14/15 = 0.881
23	14	0	0.881· 1 = 0.881
30	13	1	0.881· 12/13 = 0.814
36	12	1	0.814 · 11/12 = 0.746
38	11	0	0.746 · 1 = 0.746
54	10	0	0.746 · 1 = 0.746
56	9	0	0.746 · 1 = 0.746
59	8	1	0.746 · 7/8 = 0.653
75	7	1	0.653 · 6/7 = 0.559
93	6	1	0.559 · 5/6 = 0.466
97	5	1	0.466 · 4/5 = 0.373
104	4	0	0.373 · 1 = 0.373
107	3	2	0.373 · 1/3 = 0.124



B. The median time to discontinuation is

$$\hat{t}_{50} = min\{t : \hat{S}(t) \leq 0.50\}$$

For these subjects, the first (or shortest) survival time where the Kaplan-Meier estimate is less than or equal to 0.50 is 93 weeks. So the median time to discontinuation is 93 weeks.

C. The gynecologist would like to know what proportion of women will use the IUD for at least a year (i.e., 52 weeks). So, we need to estimate $P(T > 52) = S(52)$. From our Kaplan-Meier estimate, $\hat{S}(52) = 0.746$. So we estimate that 74.6% of women will use the IUD for at least one year.

2. A. The survival function for T is

$$S(t) = 1 - F(t) = 1 - \frac{t^3}{1+t^3} = \frac{1}{1+t^3}$$

B. The pdf of T is

$$f(t) = \frac{d}{dt}F(t) = \frac{d}{dt}\frac{t^3}{1+t^3} = \frac{(1+t^3)\cdot 3t^2 - t^3 \cdot 3t^2}{(1+t^3)^2} = \frac{3t^2}{(1+t^3)^2}$$

So, the hazard function for T is

$$h(t) = \frac{f(t)}{S(t)} = \frac{\frac{3t^2}{(1+t^3)^2}}{\frac{1}{1+t^3}} = \frac{3t^2}{1+t^3}$$

C. The median survival time is the value of t such that S(t) = 0.5. So to find the median survival time, we need to solve the following equation:

$$S(t) = \frac{1}{1+t^3} = 0.5$$

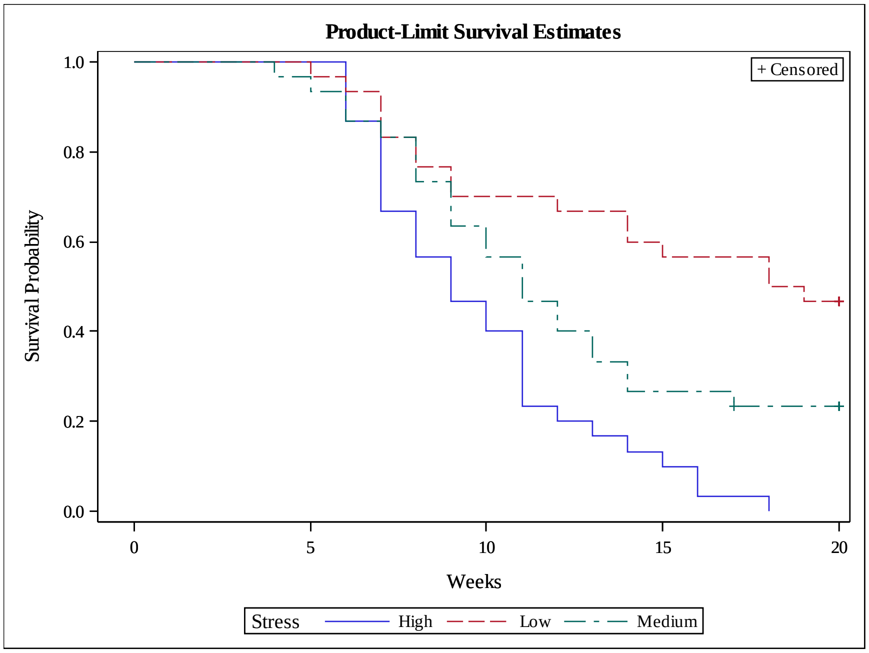
$$1 = 0.5 + 0.5t^3$$

$$t^3 = 1$$

$$t = 1$$

So the median survival time is 1.

3. A.



The rats exposed to the high-stress environment showed the worst survival, while rats exposed to the low-stress environment had the best survival.

B. The median survival times for each of the three environments are shown in the table below.

Diet Group	Median Survival Time
Low-Stress	18.0 weeks*
Medium-Stress	11.0 weeks
High-Stress	9.0 weeks

* 18.5 weeks also acceptable

C. Let $S_L(t)$, $S_M(t)$, and $S_H(t)$ be the survival functions for rats exposed to the low-stress, medium-stress, and high-stress environments, respectively. We will use the log-rank test to test the null hypothesis H_0: $S_L(t) = S_M(t) = S_H(t)$ against the alternative hypothesis H_1: At least two survival functions are different. We will use $\alpha = 0.05$ as the significance level for this test.

The test statistic for the log-rank test is Q = 20.330. The p-value is $P(\chi^2_2 \geq 20.330) < 0.0001$, so we reject H_0. We conclude that there is evidence that the survival functions for the three environments are not the same.

D. Ninety rats were exposed to one of three environments: low-stress (n = 30), medium-stress (n = 30), or high-stress (n = 30). The rats were then injected with tumor cells and observed for up to 20 weeks. All of the rats exposed to the high-stress environment developed tumors; for the medium-stress and low-stress environments, 77% and 53% of rats developed tumors, respectively. The Kaplan-Meier estimate for the tumor-free time for each environment is shown in Figure 1. The median time to tumor development was 9, 11, and 18 weeks for the high-stress, medium-stress, and low-stress environments, respectively. Environment was significantly associated with tumor development ($p < 0.001$ from the log-rank test).