Q1 (11 marks)

(a) Consider following probability density function (p.d.f.) of a random variable T:

$$f(t) = (\lambda + \gamma t)\ \exp\ -[\lambda t + \frac{\gamma t^2}{2}]\quad (t \geq 0).$$

(i) Find the cumulative distribution function F(t) and the Survival function S(t). (2)

(ii) Find the hazard rate h(t) and plot it. (1)

(iii) Find the cumulative hazard rate function H(t). Compute H(t) at t = 3 when $\lambda = 2, \gamma = 3$. (2)

(iv) Find the median M of the distribution. Compute M when $\lambda = 2, \gamma = 3$ (3)

(b) The survival times (in months) from diagnosis until death of 12 hemophiliacs with AIDS were: 2, 3, 6, 6, 7, 10, 15, 15, 16, 27, 30, and 32. Compute and plot the estimated survival function, the p.d.f., hazard function and the median. (3)

Q2. (8 marks)

The remission times (in months) of 8 patients with a certain disease are: 6,6,6,7,10,13,16,22,23,6+, 9+, 10+, 11+, 17+, 19+, 20+, 25+, 32+, 32+,34+, 35+. The corresponding Kaplan-Meier estimates $\hat{S}(t)$ of the survival function $S(t)$ are:

Kaplan-Meier Estimates

Time	Number at Risk	Number Failed	Survival Probability	Standard Error	95.0% Normal CI Lower	Upper
6	21	3	0.857143	0.076360	0.707479	1.00000
7	17	1	0.806723	0.086935	0.636333	0.97711
10	15	1	0.752941	0.096350	0.564099	0.94178
13	12	1	0.690196	0.106815	0.480843	0.89955
16	11	1	0.627451	0.114054	0.403910	0.85099
22	7	1	0.537815	0.128234	0.286482	0.78915
23	6	1	0.448179	0.134591	0.184385	0.71197

(i) Estimate the median survival time. (1)

(ii) Estimate the mean survival time. (4)

(iii)Construct a 95% confidence interval for the survival function at 12 months. (3)

Q3. (8 marks)

Suppose the remission times (in months) of 12 patients with blood cancer are as follows:

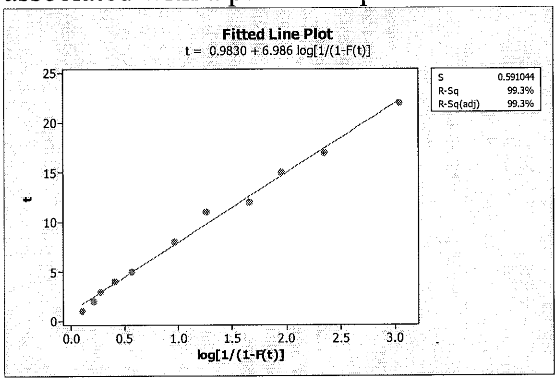
Treatment group	24	30	42	15+	40+	42+
Control group	10	26	28	30	41	12+

(a) For each of the data sets above, assuming a two-parameter exponential model compute the estimates of the hazard rate λ and the "guarantee time" G. (3)

(b)Use the log-rank test based on normal approximations to test the claim that the treatment is effective. Use 5% level of significance. (5)

Q4. (10 marks)

(a) For a given set of survival data, the following computer output is a graphical plot associated with a particular parametric model.



(i) What is the name of this plot? (0.5)

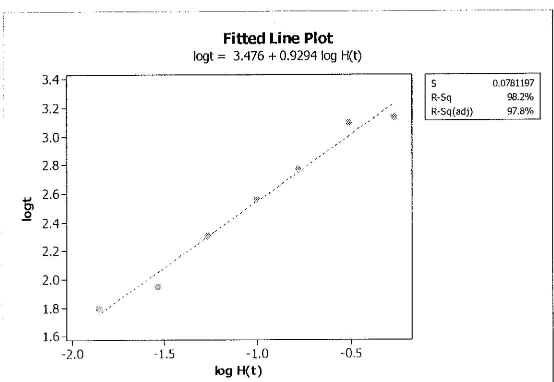
(ii) What parametric model would you associate with the given data? (0.5)

(iii) Is the parametric model obtained suitable for the data? Why? (0.5)

(iv) Obtain estimates of the parameters of the model. (1)

(v) Estimate the survival function at time t = 40 for this parametric model. (1)

(b) For a given set of survival data, the following computer output is a graphical plot associated with a specific parametric model.



(i) What is the name of this plot? (0.5)

(ii) What parametric model would you associate with the given data? (0.5)

(iii) Is the parametric model obtained suitable for the data? Why? (0.5)

(iv) Obtain estimates of the parameters of the model. (1)

(v) Estimate the survival function at time t = 40 for this parametric model. (1)

(iii) Estimate the probability of surviving between t =10 and t = 12. (1)

(iv) Estimate the inter-quartile range. (1)

(v) Estimate the mean and standard deviation of the survival times. (1)

Q5. (5 marks)

Suppose that the survival time (in years) of two groups of leukemia patients follows the Weibull distribution. A random sample of 46 patients from each group was studied. The maximum likelihood estimates obtained from the two groups are, respectively

$\hat{\gamma}_1 = 4.3, \ \hat{\lambda}_1 = 1.7, \ \hat{\gamma}_2 = 3.9, \ \hat{\lambda}_2 = 2.3$

At 5% significance level, test the hypothesis that the two groups have the same Weibull distribution.

Q6. (8 marks)

A data set contained the survival times (in months) of 65 patients and includes measurements on each patient for the following five covariates:

X1 = Logarithm of a blood urea nitrogen measurement at diagnosis,
X2 = Hemoglobin measurement at diagnosis,
X3 = Age at diagnosis, x4 = Sex: 0 for male and 1 for female,
X5 = Serum calcium measurement at diagnosis.

Some of the data were right-censored. The data has been analyzed using **Weibull Regression Model.** A summary of the analysis using MINITAB are given in the table below:

Regression Table

Predictor	Coef	Standard Error	Z	P	95.0% Normal CI Lower	Upper
Intercept	4.56741	1.44738	3.16	0.002	1.73060	7.40423
x1	-1.63230	0.517331	-3.16	0.002	-2.64625	-0.618350
x2	0.118322	0.0537559	2.20	0.028	0.0129622	0.223682
x3	0.0187257	0.0140084	1.34	0.181	-0.0087303	0.0461818
x4	0.0356253	0.273780	0.13	0.896	-0.500974	0.572225
x5	-0.121190	0.0883431	-1.37	0.170	-0.294339	0.0519597
Shape	1.13426	0.121312			0.919760	1.39879

Log-Likelihood = -207.322

(i) Which of the five regression variables have significant effect on survival time? Explain your answers. (2)

(ii) Is the **Exponential Regression Model** likely to give a good fit to the data? Why? (2)

(iii) Estimate the probability of surviving between 36 and 48 months for a person having $x_1 = 1, x_2 = 1, x_3 = 1$. (2)

(iv) Estimate the relative hazard of females to that of males. Explain your answers. (2)

Formula Sheet:

$H(t) = \int_0^t h(y)dy$, Slope $= 1/\hat{\lambda}$, y-intercept $= \ln 1/\hat{\lambda}$, slope $= 1/\hat{\gamma}$,

$S(t) = e^{-(\lambda t)^{\gamma}}$, $h(t) = \gamma\lambda(\lambda t)^{\gamma-1}$, $Q_1 = \frac{1}{\lambda}(\ln 4/3)^{1/\gamma}$ $Q_2 = \frac{1}{\lambda}(\ln 2)^{1/\gamma}$ $Q_3 = \frac{1}{\lambda}(\ln 4)^{1/\gamma}$,

$\hat{\mu}$ = Area under KME $\hat{S}(t)$,

$Var(\hat{\mu}) = \sum_r A_r^2 / \{(n-r)(n-r+1)\}$, $\hat{\mu} \pm z_{\alpha/2} \frac{\hat{\mu}}{\sqrt{r}}$,

$Var[\hat{S}(t)] \cong [\hat{S}(t)]^2 \sum_r 1/\{n-r)(n-r+1)$,

$S = \sum_{i=1}^{n2} w_i$, $Var(S) = \frac{n_1 n_2}{(n_1+n_2)(n_1+n_2+1)} \sum_{i=1}^{n1+n2} w_i^2$, $\chi^2 = \frac{(O_1-E_1)^2}{E_1} + \frac{(O_2-E_2)^2}{E_2}$,

$G = \frac{\hat{\gamma}_1 + \hat{\gamma}_2}{2} \ln\left(\hat{\lambda}_2 / \hat{\lambda}_1\right)$, $h(t;\lambda_i,\gamma) = \lambda_i \gamma \, t^{\gamma-1}$, $\lambda_i = \exp[-(a_0 + a_1 x_1 + \ldots + a_p x_p)/\sigma]$, $\sigma = 1/\gamma$

For n = 46 the 95% point in Table B-12 is 1.314,
For n = 46 the 95% point in Table B-13 is 0.376.

$z_{0.01} = 2.326$, $z_{0.05} = 1.645$, $z_{0.025} = 1.96$, $z_{0.005} = 2.58$,
$\chi_{1,0.05} = 3.84146$, $\chi_{1,0.025} = 5.02389$, $\chi_{1,0.01} = 6.63490$,
$\chi_{2,0.05} = 5.99147$, $\chi_{2,0.025} = 7.37776$, $\chi_{2,0.01} = 9.21034$,
$\chi_{3,0.05} = 7.81473$, $\chi_{3,0.025} = 9.34840$, $\chi_{3,0.01} = 11.3449$.