Final *2019. 06. 19*

I. Written (80%)

1. (40%) Suppose we have the following data for the survival times of ovarian cancer patients: T: treatment group; C: control group.

Subject	Survival time	Censor indicator	Group	Age
I	156	1	T	66
II	1040	0	T	38
III	59	1	T	72
IV	421	0	C	53
V	329	1	T	43
VI	769	0	C	59

(a) Calculate the Kaplan-Meier estimate for the data and find the survival functions evaluated at 250.

(b) Fit the above data by the Weibull distribution with density function

$$f(t) = 2\lambda t \cdot exp(-\lambda t^2)$$

Find the MLE of λ and find the estimated survival function.

(c) With $\alpha = 0.05$, please perform the log-rank and Wilcoxon tests to test if there is the group effect.

(d) Suppose the variable Age is the variable of interest. Using proportional hazards model, derive the partial likelihood and find the score function.

2. (20%) Suppose the independent data $Y_1, Y_2, \cdots, Y_n$ have the mean μ_i and the variance function. $V_i(\mu_i)$

(a) If $\mu_i = \mu, V_i(\mu) = \mu^3$, find the quasi-likelihood function and maximized quasi-likelihood estimate for μ.

(b) If $\mu_i = \mu, V_i(\mu) = \mu(1-\mu)$, find the quasi-likelihood function and maximized quasi-likelihood estimate for μ.

3. (20%) Suppose $Y_1 \sim P(\mu_1)$ and $Y_2 \sim P(\mu_2)$ and we are interested in the ratio $\varphi = {\mu_1}/{\mu_2}$. Please find the conditional likelihood estimate.

II. Computer (120%)

1. (40%) For the data set in problem 1 of written part,
   (a) (20%) Find the survival estimates in (a) and (b) of the problem and plot the survival functions.
   (b) (10%) Test the group effect using Log-rank test, Wilcoxon test, and the proportional hazards models at $\alpha = 0.05$.
   (c) (10%) Fit the following proportional hazards models $\lambda_0(t) \cdot exp(\eta)$ and please make conclusions at $\alpha = 0.05$:
   - $\eta = \beta \cdot \text{Age}$
   - $\eta = \beta \cdot \text{Group}$

2. (30%) The following table refers to a prospective study of maternal drinking and congenital malformations. Observations were recorded on presence or absence of congenital organ malformations.

	Alcohol Consumption (Ave. no. Drinks per Day)				
Malformation	0	<1	$1-2$	$3-5$	$>=6$
Absent	17066	14464	788	126	37
Present	48	38	5	1	1

(a) (20%) Please use proportional odds model to analyze the above data and make conclusions with $\alpha = 0.05$.
(b) (10%) Please select the sensible models and make conclusions.

2. (50%) The following data are in many ways typical of social-science investigations, although the data concerns the behavior of lizards rather than humans.

			T					
			Early		Mid – day		Late	
S	D (in)	H (ft)	G	O	G	O	G	O
Sun	≤ 2	<5	20	2	8	1	4	4
		≥ 5	13	0	8	0	12	0
	>2	<5	8	3	4	1	5	3
		≥ 5	6	0	0	0	1	1
Shade	≤ 2	<5	34	11	69	20	18	10
		≥ 5	31	5	55	4	13	3
	>2	<5	17	15	60	32	8	8
		≥ 5	12	1	21	5	4	4

G: Grahami lizard; O: Opalinus lizard;

S: sunny/shady; H: perch height; D: perch diameter; T: time of day.

(a) (20%) Please analyze the data using both *logit and probit* links and make conclusions with $\alpha = 0.05$.

(b) (10%) Please select the sensible models and make conclusions.

(c) (10%) Please give the residual plots, including the one for Pearson residuals and the one for deviance residuals.

(d) (10%) Please use the quasi-likelihood method with $V(\mu) = \mu$ and $V(\mu) = \mu(1-\mu)$, respectively, and compare the results with the ones given in (a).