Midterm *2019. 4. 17*

(Written)

1. (30%) For independent binomial sampling with response

$$Y_i \sim B(m_i, p_i), i = 1, 2, \cdots, n,$$

please write down the IRLS estimates with respect to the following two link functions.

(a) Logistic link function: $logit(p_i) = \beta_0 + \beta_1 x_i$.

(b) Probit link function: $\Phi^{-1}(p_i) = \beta_0 + \beta_1 x_i$.

Please specify the weight matrix W_t **and** z_t **explicitly.**

2. (20%) Let $\mathbf{X} = (X_1, X_2) \sim Multinomial(n, p_1, p_2)$ **have the multinomial distribution, i.e., the probability distribution being**

$$f(x_1, x_2) = \frac{n!}{x_1!\, x_2!\, (n - x_1 - x_2)!} p_1^{x_1} p_2^{x_2} (1 - p_1 - p_2)^{n - x_1 - x_2}$$

(a) Find the moment generating function for $\mathbf{X} = (X_1, X_2)$.

(b) Show that $E(X_i) = np_i, i = 1, 2,$ **and** $Cov(X_1, X_2) = -np_1 p_2$.

(Computer)

1. **(50%) For problem 1 in written part, Please write two R programs for (a) and (b), respectively, to carry out the computation. In the program, the inputs are the response and covariates while the output is the IRLS estimate. Please simulate data to test the program you write and compare the results with the ones using the R function "glm". Comment on the results.**

2. (40%) The following table gives frequencies for whether a boy scout, delinquency, and socioeconomic status.

		Delinquent (crime)	
Socioeconomic Status	Boy Scout	Yes	No
Low	Yes	10	40
	No	40	160
Medium	Yes	18	132
	No	18	132
High	Yes	8	192
	No	2	48

(a) Please analyze the data using both *logit and probit* links and make conclusions.

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

3. (40%) The following data concern the number of car accidents.

Age	Gender	Region	Number of car accidents
20 – 34	Male	A	288
20 – 34	Female	A	60
20 – 34	Male	B	90
20 – 34	Female	B	19
20 – 34	Male	C	226
20 – 34	Female	C	88
20 – 34	Male	D	285
20 – 34	Female	D	110
35 – 44	Male	A	224
35 – 44	Female	A	35
35 – 44	Male	B	96
35 – 44	Female	B	12
35 – 44	Male	C	189
35 – 44	Female	C	44
35 – 44	Male	D	225
35 – 44	Female	D	53
> 44	Male	A	337
> 44	Female	A	70
> 44	Male	B	124
> 44	Female	B	17
> 44	Male	C	156
> 44	Female	C	70
> 44	Male	D	324
> 44	Female	D	60

Assume that the numbers of the accidents are Poisson variables Y.

(a) Please find the estimate of σ^2, the variance of the Poisson variable Y, and then analyze the data based on the estimate of σ^2. Please make conclusions.

(b) Please give the residual plots, including the one for Pearson residuals, the one for deviance residuals, and the one for Anscombe residuals.