

Supporting Materials for Change Point Testing in Logistic Regression Models with Interaction Term

Youyi Fong¹, Chongzhi Di², Sallie Permar³

¹Vaccine and Infectious Disease Division and Public Health Sciences Division

Fred Hutchinson Cancer Research Center,

²Public Health Sciences Division, Fred Hutchinson Cancer Research Center

³Duke Human Vaccine Institute, Duke University Medical Center

Simulation results for covariates with skewed distributions

To examine the performance of the tests when the covariate distributions are skewed, we perform an additional simulation study. Z is simulated from a Gamma distribution with shape 2.5 and rate 1. A histogram of Z from a simulated dataset is shown in Figure S1. X is simulated from the same Gamma distribution, but shifted and scaled to have mean 4.7 and standard deviation 1.6 to be comparable to the normal simulation scenarios.

The power under the simulation scenario in Section 4.1 is shown in Table S1. The sizes of the tests are 4.7, 5.0, 4.9, 5.0 and 5.2, respectively. Comparing these results to the normal covariates, we see that the power of the tests is similar.

The power under the simulation scenario in Section 4.2 with $\beta_1 = \log(0.67)$ is shown in Table S2. The sizes of the tests are 4.3, 3.6, 1.9, 5.1, 3.5 and 3.4, respectively. Comparing these results to the normal covariates, we see that the proposed test LR_{\max}^{MC} is less powerful, but still useful under skewed covariate distribution. For example, when $\beta_2 = \log(0.4)$ and the true threshold is at 21% and 79% quantiles, the power under normal ($\rho = 0$) covariate distribution is 60.2% and 53.7%, respectively;

and the power under Gamma covariate distribution are 53.8% and 40.2%, respectively.

The same trend is also true for other tests.

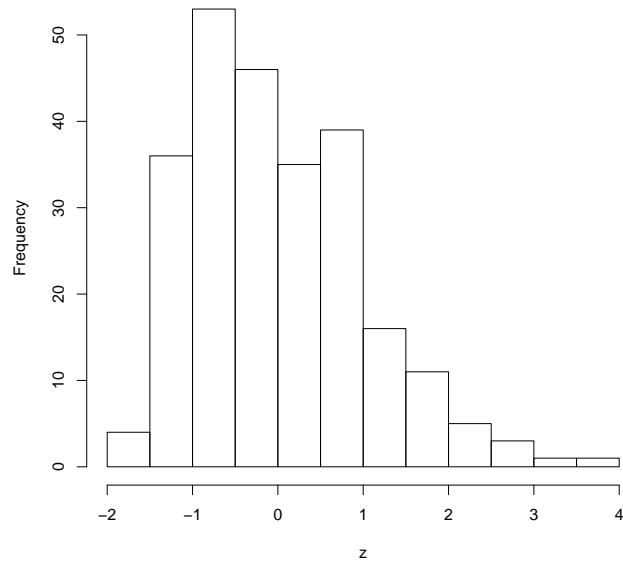


Figure S1: Histogram of Z from a Gamma distribution.

Threshold	0.21	0.36	0.46	0.58	0.79
<i>OR=0.8</i>					
trichotomized	6.4	9.2	9.4	9.7	5.8
dichotomized	7.3	11.1	12.2	9.0	5.6
T_{\max} ($M = 10$)	8.6	9.7	9.6	8.6	6.2
T_{\max} ($M = 50$)	9.0	10.3	9.9	8.8	6.3
T_{\max} ($M = 100$)	8.8	10.3	9.8	8.8	6.4
<i>OR=0.6</i>					
trichotomized	16.7	27.9	26.2	28.0	11.6
dichotomized	13.5	36.9	43.6	23.9	10.2
T_{\max} ($M = 10$)	22.5	31.5	32.1	26.1	14.4
T_{\max} ($M = 50$)	24.8	32.6	32.8	27.0	15.3
T_{\max} ($M = 100$)	24.6	32.4	32.8	27.1	15.6
<i>OR=0.4</i>					
trichotomized	47.0	73.6	66.2	75.5	26.4
dichotomized	33.7	82.5	85.7	60.3	19.4
T_{\max} ($M = 10$)	64.5	79.8	78.0	71.5	41.5
T_{\max} ($M = 50$)	67.7	81.2	78.6	72.6	43.9
T_{\max} ($M = 100$)	67.3	81.0	78.5	72.8	44.2

Table S1: Powers of hypothesis testing procedures for logistic regression models with change point variable as a main effect only. The covariates are simulated from Gamma distributions. The threshold values are given in quantiles of the distribution of X . OR: odds ratios for the main effect.

Threshold	0.21	0.36	0.46	0.58	0.79
<i>OR</i> =(0.67, 0.8)					
dichotomized	8.6	21.8	24.2	14.1	6.8
T_{\max}	14.5	19.8	18.9	16.6	9.7
LR_{\max}^{PGC}	6.5	9.9	10.2	9.7	5.3
LR_{\max}^{MC}	14.2	19.2	19.4	17.3	11.9
two-sided T_{\max}^w	12.2	17.4	16.7	14.6	8.6
one-sided T_{\max}^w	14.1	19.8	19.3	17.2	9.2
<i>OR</i> =(0.67, 0.6)					
dichotomized	13.0	39.8	44.9	24.6	8.8
T_{\max}	14.3	19.2	18.1	15.6	8.8
LR_{\max}^{PGC}	14.4	22.9	22.1	19.7	10.3
LR_{\max}^{MC}	24.7	37.9	35.3	31.5	19.6
two-sided T_{\max}^w	24.8	34.2	31.0	26.7	13.6
one-sided T_{\max}^w	27.6	38.2	34.9	29.9	16.0
<i>OR</i> =(0.67, 0.4)					
dichotomized	27.3	73.2	78.5	46.1	13.8
T_{\max}	12.5	13.7	13.2	11.7	7.1
LR_{\max}^{PGC}	38.5	55.0	55.0	47.6	24.9
LR_{\max}^{MC}	53.8	70.6	70.3	64.7	40.2
two-sided T_{\max}^w	57.8	67.7	65.1	55.5	28.5
one-sided T_{\max}^w	60.0	71.1	69.2	60.3	32.6

Table S2: Powers of hypothesis testing procedures for logistic regression models with change point variable both as a main effect and as part of an interaction term. The covariates are simulated from Gamma distributions. The threshold values are given in quantiles of the distribution of X . OR: odds ratios for the main and interaction effect.