P8110 Applied Regression II - Homework 8

Name: Xuange Liang **UNI:** xl3493 **Date:** December 3, 2025

Question 1

Fit a GEE model with temperature as outcome and time, treatment, and their interactions as covariates. Write the mean response of the GEE model and treat time as a categorical variable.

SAS Code

```
PROC GENMOD DATA=tempdata;
    CLASS ID time(REF='0') treatment(REF='0') / PARAM=REF;
    MODEL temp = time treatment time*treatment / DIST=NORMAL;
    REPEATED SUBJECT=ID / TYPE=CS CORRW;
RUN;
```

SAS Output

GEE Fit Criteria	
QIC	1702.3998
QICu	1703.0000

Figure 1: GEE Fit Criteria

Analysis Of GEE Parameter Estimates							
Empirical Standard Error Estimates							
Parameter			Estimate	Standard Error	95% Confidence Limits	Z	Pr > |Z|
Intercept			100.4902	0.1256	100.2440 100.7364	800.07	<.0001
time	2		-0.3216	0.0718	-0.4623 -0.1808	-4.48	<.0001
time	4		-0.3627	0.0907	-0.5404 -0.1849	-4.00	<.0001
time	8		-0.4503	0.1107	-0.6674 -0.2333	-4.07	<.0001
treatment	1		-0.1287	0.1900	-0.5011 0.2437	-0.68	0.4982
time*treatment	2	1	-0.5449	0.1106	-0.7617 -0.3281	-4.93	<.0001
time*treatment	4	1	-1.0883	0.1454	-1.3733 -0.8034	-7.49	<.0001
time*treatment	8	1	-1.3135	0.1609	-1.6288 -0.9982	-8.16	<.0001

Figure 2: Analysis Of GEE Parameter Estimates

Mean Response Model

The mean response of the GEE model treating time as a categorical variable is:

$$E(Y_{ij}) = \mu_{ij} = \beta_0 + \beta_1 \cdot I(\text{time}_{ij} = 2) + \beta_2 \cdot I(\text{time}_{ij} = 4) + \beta_3 \cdot I(\text{time}_{ij} = 8)$$

$$+\beta_4 \cdot I(\text{treatment}_i = B) + \beta_5 \cdot I(\text{time}_{ij} = 2) \cdot I(\text{treatment}_i = B)$$

$$+\beta_6 \cdot I(\text{time}_{ij} = 4) \cdot I(\text{treatment}_i = B) + \beta_7 \cdot I(\text{time}_{ij} = 8) \cdot I(\text{treatment}_i = B)$$

Where: - Y_{ij} = temperature for patient i at time point j - $I(\cdot)$ = indicator function - β_0 = mean temperature at baseline (time=0) for Treatment A - $\beta_1, \beta_2, \beta_3$ = effects of time 2, 4, 8 (vs baseline) for Treatment A - β_4 = effect of Treatment B vs A at baseline - $\beta_5, \beta_6, \beta_7$ = interaction effects (difference in time effects between treatments)

Question 2

Try different working correlation structures (CS, AR(1), and UN) for the GEE model in (1). Which model yields the best QIC value? Show the SAS code and relevant SAS output.

SAS Code

```
/* Compound Symmetry (CS) / Exchangeable */
PROC GENMOD DATA=tempdata;
    CLASS ID time(REF='0') treatment(REF='0') / PARAM=REF;
    MODEL temp = time treatment time*treatment / DIST=NORMAL;
    REPEATED SUBJECT=ID / TYPE=CS CORRW;
RUN;

/* AR(1) */
PROC GENMOD DATA=tempdata;
    CLASS ID time(REF='0') treatment(REF='0') / PARAM=REF;
    MODEL temp = time treatment time*treatment / DIST=NORMAL;
    REPEATED SUBJECT=ID / TYPE=AR(1) CORRW;
RUN;

/* Unstructured */
PROC GENMOD DATA=tempdata;
    CLASS ID time(REF='0') treatment(REF='0') / PARAM=REF;
    MODEL temp = time treatment time*treatment / DIST=NORMAL;
    REPEATED SUBJECT=ID / TYPE=UN CORRW;
RUN;
```

SAS Output - QIC Comparison

Compound Symmetry (CS)

AR(1)

GEE Fit Criteria	
QIC	1702.3998
QICu	1703.0000

Figure 3: GEE Fit Criteria - CS

Analysis Of GEE Parameter Estimates								
Empirical Standard Error Estimates								
Parameter			Estimate	Standard Error	95% Confidence Limits		Z	Pr > |Z|
Intercept			100.4902	0.1256	100.2440	100.7364	800.07	<.0001
time	2		-0.3216	0.0718	-0.4623	-0.1808	-4.48	<.0001
time	4		-0.3627	0.0907	-0.5404	-0.1849	-4.00	<.0001
time	8		-0.4503	0.1107	-0.6674	-0.2333	-4.07	<.0001
treatment	1		-0.1287	0.1900	-0.5011	0.2437	-0.68	0.4982
time*treatment	2	1	-0.5449	0.1106	-0.7617	-0.3281	-4.93	<.0001
time*treatment	4	1	-1.0883	0.1454	-1.3733	-0.8034	-7.49	<.0001
time*treatment	8	1	-1.3135	0.1609	-1.6288	-0.9982	-8.16	<.0001

Figure 4: Analysis Of GEE Parameter Estimates - CS

GEE Fit Criteria	
QIC	1702.2653
QICu	1703.0000

Analysis Of GEE Parameter Estimates								
Empirical Standard Error Estimates								
Parameter			Estimate	Standard Error	95% Confidence Limits		Z	Pr > |Z|
Intercept			100.4902	0.1256	100.2440	100.7364	800.07	<.0001
time	2		-0.3286	0.0722	-0.4700	-0.1871	-4.55	<.0001
time	4		-0.3752	0.0912	-0.5540	-0.1964	-4.11	<.0001
time	8		-0.4324	0.1113	-0.6506	-0.2142	-3.88	0.0001
treatment	1		-0.1287	0.1900	-0.5011	0.2437	-0.68	0.4982
time*treatment	2	1	-0.5395	0.1111	-0.7572	-0.3218	-4.86	<.0001
time*treatment	4	1	-1.0850	0.1453	-1.3698	-0.8003	-7.47	<.0001
time*treatment	8	1	-1.2962	0.1615	-1.6127	-0.9798	-8.03	<.0001

Figure 5: GEE Results - AR(1)

GEE Fit Criteria	
QIC	1702.2859
QICu	1703.0000

Analysis Of GEE Parameter Estimates								
Empirical Standard Error Estimates								
Parameter			Estimate	Standard Error	95% Confidence Limits		Z	Pr > |Z|
Intercept			100.4902	0.1256	100.2440	100.7364	800.07	<.0001
time	2		-0.3209	0.0723	-0.4627	-0.1792	-4.44	<.0001
time	4		-0.4030	0.0925	-0.5843	-0.2218	-4.36	<.0001
time	8		-0.4382	0.1108	-0.6553	-0.2211	-3.96	<.0001
treatment	1		-0.1287	0.1900	-0.5011	0.2437	-0.68	0.4982
time*treatment	2	1	-0.5539	0.1115	-0.7725	-0.3354	-4.97	<.0001
time*treatment	4	1	-1.0646	0.1461	-1.3509	-0.7782	-7.29	<.0001
time*treatment	8	1	-1.3090	0.1604	-1.6233	-0.9947	-8.16	<.0001

Figure 6: GEE Results - Unstructured

Unstructured

QIC Comparison Table

Correlation Structure	QIC	QICu
Compound Symmetry (CS)	1702.40	1703.00
AR(1)	**1702.27**	**1703.00**
Unstructured	1702.29	1703.00

Answer

Comparing the QIC values from the three correlation structures:

- CS: QIC = 1702.40
- AR(1): QIC = 1702.27
- UN: QIC = 1702.29

The **AR(1)** model has the lowest QIC value (1702.27), so it is the best model. I will use the AR(1) correlation structure for the following questions.

Question 3

Use the model selected in (2) to test whether the trajectory of temperature over time is different between the two treatments. Write down the hypothesis, test statistic, p-value, and conclusion.

SAS Code

```
PROC GENMOD DATA=tempdata;
    CLASS ID time(REF='0') treatment(REF='0') / PARAM=REF;
    MODEL temp = time treatment time*treatment / DIST=NORMAL;
    REPEATED SUBJECT=ID / TYPE=AR(1) CORRW;
    /* Joint test for interaction terms */
    CONTRAST 'Treatment x Time Interaction'
        time*treatment 1 0 0,
        time*treatment 0 1 0,
        time*treatment 0 0 1 / WALD;
RUN;
```

SAS Output

GEE Fit Criteria	
QIC	1702.2653
QICu	1703.0000

Analysis Of GEE Parameter Estimates								
Empirical Standard Error Estimates								
Parameter			Estimate	Standard Error	95% Confidence Limits		Z	Pr > |Z|
Intercept			100.4902	0.1256	100.2440	100.7364	800.07	<.0001
time	2		-0.3286	0.0722	-0.4700	-0.1871	-4.55	<.0001
time	4		-0.3752	0.0912	-0.5540	-0.1964	-4.11	<.0001
time	8		-0.4324	0.1113	-0.6506	-0.2142	-3.88	0.0001
treatment	1		-0.1287	0.1900	-0.5011	0.2437	-0.68	0.4982
time*treatment	2	1	-0.5395	0.1111	-0.7572	-0.3218	-4.86	<.0001
time*treatment	4	1	-1.0850	0.1453	-1.3698	-0.8003	-7.47	<.0001
time*treatment	8	1	-1.2962	0.1615	-1.6127	-0.9798	-8.03	<.0001

Contrast Results for GEE Analysis				
Contrast	DF	Chi-Square	Pr > ChiSq	Type
Treatment x Time Interaction	3	69.37	<.0001	Wald

Figure 7: Contrast Results for GEE Analysis

Answer

Hypothesis:

- H_0: $\beta_5 = \beta_6 = \beta_7 = 0$ (no treatment-by-time interaction)
- H_a: At least one $\beta_j \neq 0$ for $j = 5, 6, 7$

Test Statistic: Wald χ^2 = 69.37, df = 3

P-value: < 0.0001

Conclusion: Since p < 0.0001 < 0.05, we reject H_0. There is a significant difference in temperature trajectory over time between the two treatment groups.

Question 4

Use the model selected in (2) to estimate the mean temperature change from baseline to two hours after entry into study for patients in treatment A group and those in treatment B group, respectively.

SAS Code

```
PROC GENMOD DATA=tempdata;
    CLASS ID time(REF='0') treatment(REF='0') / PARAM=REF;
    MODEL temp = time treatment time*treatment / DIST=NORMAL;
    REPEATED SUBJECT=ID / TYPE=AR(1) CORRW;

    /* Treatment A: Change from time=0 to time=2 = beta1 */
    ESTIMATE 'Change 0->2hr, Treatment A' time 1 0 0;

    /* Treatment B: Change from time=0 to time=2 = beta1 + beta5 */
    ESTIMATE 'Change 0->2hr, Treatment B' time 1 0 0 time*treatment 1 0 0;
RUN;
```

SAS Output

Calculation Explanation

For Treatment A (treatment = 0): - Mean at time 0: $E(Y|time = 0, trt = A) = \beta_0$ - Mean at time 2: $E(Y|time = 2, trt = A) = \beta_0 + \beta_1$ - **Change from baseline to time 2 =** β_1 **= -0.329**

For Treatment B (treatment = 1): - Mean at time 0: $E(Y|time = 0, trt = B) = \beta_0 + \beta_4$ - Mean at time 2: $E(Y|time = 2, trt = B) = \beta_0 + \beta_1 + \beta_4 + \beta_5$ - **Change from baseline to time 2 =** $\beta_1 + \beta_5$ **= -0.329 + (-0.540) = -0.868**

Answer

Treatment A:

- Mean temp change from baseline to 2 hours = β_1 = **-0.329** (SE = 0.0722, 95% CI: -0.470, -0.187)

Treatment B:

- Mean temp change from baseline to 2 hours = $\beta_1 + \beta_5$ = -0.329 + (-0.540) = **-0.868** (SE = 0.0844, 95% CI: -1.034, -0.703)

Both groups show a decrease in temperature, but Treatment B shows a larger decrease.

GEE Fit Criteria	
QIC	1702.2653
QICu	1703.0000

Analysis Of GEE Parameter Estimates

Empirical Standard Error Estimates

Parameter			Estimate	Standard Error	95% Confidence Limits		Z	Pr > |Z|
Intercept			100.4902	0.1256	100.2440	100.7364	800.07	<.0001
time	2		-0.3286	0.0722	-0.4700	-0.1871	-4.55	<.0001
time	4		-0.3752	0.0912	-0.5540	-0.1964	-4.11	<.0001
time	8		-0.4324	0.1113	-0.6506	-0.2142	-3.88	0.0001
treatment	1		-0.1287	0.1900	-0.5011	0.2437	-0.68	0.4982
time*treatment	2	1	-0.5395	0.1111	-0.7572	-0.3218	-4.86	<.0001
time*treatment	4	1	-1.0850	0.1453	-1.3698	-0.8003	-7.47	<.0001
time*treatment	8	1	-1.2962	0.1615	-1.6127	-0.9798	-8.03	<.0001

Contrast Estimate Results

Label	Mean Estimate	Mean Confidence Limits		L'Beta Estimate	Standard Error	Alpha	L'Beta Confidence Limits		Chi-Square	Pr > ChiSq
Change 0->2hr, Treatment A	-0.3286	-0.4700	-0.1871	-0.3286	0.0722	0.05	-0.4700	-0.1871	20.72	<.0001
Exp(Change 0->2hr, Treatment A)				0.7200	0.0520	0.05	0.6250	0.8294		
Change 0->2hr, Treatment B	-0.8680	-1.0335	-0.7026	-0.8680	0.0844	0.05	-1.0335	-0.7026	105.70	<.0001
Exp(Change 0->2hr, Treatment B)				0.4198	0.0354	0.05	0.3558	0.4953		

Figure 8: Contrast Estimate Results - Q4

Question 5

Calculate the difference of the two estimates in (4). Denote the difference as DIFF. Which β coefficient does DIFF represent? Interpret this β coefficient.

SAS Code

```
PROC GENMOD DATA=tempdata;
    CLASS ID time(REF='0') treatment(REF='0') / PARAM=REF;
    MODEL temp = time treatment time*treatment / DIST=NORMAL;
    REPEATED SUBJECT=ID / TYPE=AR(1) CORRW;

    /* DIFF = Change_B - Change_A = (beta1 + beta5) - beta1 = beta5 */
    ESTIMATE 'DIFF: Change_B - Change_A' time*treatment 1 0 0;
RUN;
```

SAS Output

Answer

DIFF Calculation:

$$\text{DIFF} = \text{Change}_B - \text{Change}_A = (\beta_1 + \beta_5) - \beta_1 = \beta_5 = -0.540$$

Which β does DIFF represent?

GEE Fit Criteria	
QIC	1702.2653
QICu	1703.0000

Analysis Of GEE Parameter Estimates

Empirical Standard Error Estimates

Parameter			Estimate	Standard Error	95% Confidence Limits		Z	Pr > |Z|
Intercept			100.4902	0.1256	100.2440	100.7364	800.07	<.0001
time	2		-0.3286	0.0722	-0.4700	-0.1871	-4.55	<.0001
time	4		-0.3752	0.0912	-0.5540	-0.1964	-4.11	<.0001
time	8		-0.4324	0.1113	-0.6506	-0.2142	-3.88	0.0001
treatment	1		-0.1287	0.1900	-0.5011	0.2437	-0.68	0.4982
time*treatment	2	1	-0.5395	0.1111	-0.7572	-0.3218	-4.86	<.0001
time*treatment	4	1	-1.0850	0.1453	-1.3698	-0.8003	-7.47	<.0001
time*treatment	8	1	-1.2962	0.1615	-1.6127	-0.9798	-8.03	<.0001

Contrast Estimate Results

Label	Mean Estimate	Mean Confidence Limits		L'Beta Estimate	Standard Error	Alpha	L'Beta Confidence Limits		Chi-Square	Pr > ChiSq
DIFF: Change_B - Change_A	-0.5395	-0.7572	-0.3218	-0.5395	0.1111	0.05	-0.7572	-0.3218	23.59	<.0001

Figure 9: Contrast Estimate Results - Q5

DIFF = β_5, the interaction coefficient for time=2 and treatment=B.

Interpretation:

$\beta_5 = -0.540$ means that the temperature change from baseline to 2 hours is 0.540 degrees **lower** (greater decrease) for Treatment B compared to Treatment A. The 95% CI (-0.757, -0.322) does not include 0, so this difference is statistically significant ($p < 0.0001$).
