Q1

Solutions for Fixed Effects

Effect	Estimate	Standard Error	DF	t Value	Pr > |t|
Intercept	3.5687	0.4442	361	8.03	<.0001
time1	-0.3628	0.2708	724	-1.34	0.1808
time2	-0.4316	0.2703	724	-1.60	0.1107
male	-0.3818	0.4429	724	-0.86	0.3889

Covariance Parameter Estimates

Cov Parm	Subject	Estimate	Standard Error
Intercept	id	9.5543	2.0156

Q1. $\text{logit}\{P[Y=1|u]\} = u_i + \beta t + r x_i$; $t = 1, 2, 3.$

$$\hat{u}_i = (1+0.346 \times 9.5543)^{-\frac{1}{2}} \times 3.587 = 1.7198$$

$$\hat{\beta}_1 = (1+0.346 \times 9.5543)^{-\frac{1}{2}} \times (-0.3628) = -0.1748$$

$$\hat{\beta}_2 = (1+0.346 \times 9.5543)^{-\frac{1}{2}} \times (-0.4316) = -0.2080$$

$$\hat{r} = (1+0.346 \times 9.5543)^{-\frac{1}{2}} \times (-0.3818) = -0.1840.$$

Analysis Of GEE Parameter Estimates

Empirical Standard Error Estimates

Parameter	Estimate	Standard Error	95% Confidence Limits		Z	Pr > |Z|
Intercept	1.6497	0.1859	1.2853	2.0141	8.87	<.0001
time1	-0.1814	0.1451	-0.4659	0.1030	-1.25	0.2112
time2	-0.2154	0.1317	-0.4736	0.0428	-1.64	0.1020
male	-0.1731	0.2180	-0.6004	0.2542	-0.79	0.4271

Compared to our results to the output above, the results are really close to each other.

Q2

Covariance Parameter Estimates			
Cov Parm	Subject	Estimate	Standard Error
Intercept	id	0.4589	0.8318

Solutions for Fixed Effects						
Effect	y	Estimate	Standard Error	DF	t Value	Pr > |t|
Intercept	1	-0.9809	0.4244	117	-2.31	0.0226
Intercept	2	0.9810	0.4245	117	2.31	0.0226
time		1.9225	0.6612	116	2.91	0.0044
trt		-2.9382	0.5938	116	-4.95	<.0001
time*trt		1.0973	0.7225	116	1.52	0.1316

Model: $\text{logit}\{P[Y \le k \mid b]\} = a_k + b_i + \underbrace{\beta_1 I}_{\text{time}} + \underbrace{\beta_2 trt_i}_{\text{trt}} + \underbrace{\beta_3 I \times trt_i}_{\text{time* trt}}$.

which $i = 1, 2, 3, 4, \ldots, 239$; $j = 1, 2$; $k = 1, 2$. I is $I(j=2)$.

From the output below, $\hat{\beta}_1 = 1.9225$ $\therefore e^{\hat{\beta}_1} = e^{1.9225} = 6.8380$.

$\therefore$ The odds ratio of placebo group to the standard (baseline) group is 6.8380 on falling asleep time. and the p-value here is 0.0044 which is smaller than 0.05, so the effect of is significant ($\hat{\beta}_1 \neq 0$).

$\hat{\beta}_1 + \hat{\beta}_3 = 3.0198$ $\therefore e^{\hat{\beta}_1 + \hat{\beta}_3} = 20.4872$ $\therefore$ The odds ratio of the treated groups to the standard (baseline) group on falling asleep shorter time is 20.4872

Q3

(a)

Parameter		DF	Estimate
Intercept		0	0.0000
id	1	0	0.4000
id	2	0	0.8000
id	3	0	0.2000
id	4	0	0.6000
id	5	0	0.6000
id	6	0	1.0000
id	7	0	0.8000
id	8	0	0.4000
id	9	0	0.6000
id	10	0	0.2000
Scale		0	1.0000

(a)

From the output above. probability from 1 to 10 is. 0.4, 0.8, 0.2, 0.6,
0.6, 1.0, 0.8, 0.4, 0.6, 0.2.

(b)

Solutions for Fixed Effects

Effect	Estimate	Standard Error	DF	t Value	Pr > |t|
Intercept	0.2589	0.3452	9	0.75	0.4724

Effect	Subject	Estimate
Intercept	id 1	-0.1839
Intercept	id 2	0.2660
Intercept	id 3	-0.4074
Intercept	id 4	0.04000
Intercept	id 5	0.04000
Intercept	id 6	0.4955
Intercept	id 7	0.2660
Intercept	id 8	-0.1839
Intercept	id 9	0.04000
Intercept	id 10	-0.4074

(b). Model: $\text{logit}(\pi_i) = \alpha + u_i$; $u \sim N(0, \sigma^2)$.

From the output above, $\hat{\alpha} = 0.2589$. $u_1 \sim u_{10}$ is -0.1839, 0.2660, -0.4074,

fix ~~Random~~ effect ?

0.04, 0.04, 0.4955, 0.2660, -0.1839, 0.04, -0.4074.

(c)

id	p	pihat
1	0.4	0.51874
2	0.8	0.62830
3	0.2	0.46296
4	0.6	0.57418
5	0.6	0.57418
6	1.0	0.68014
7	0.8	0.62830
8	0.4	0.51874
9	0.6	0.57418
10	0.2	0.46296

(c) From the output above, estimate probability from 1 to 10 is 0.51874, 0.62830, 0.46296, 0.57418, 0.68014, 0.62830, 0.51874, 0.57418. 0.46296, which means the estimate probability of Lead.

(d)

(d) The answer from b is better. As we know, the random sample has really unstable used for estimate population when sample size is small. Therefore, if we used the random effect and fixed effect when sample size is small, the estimate accuracy is higher.

(e)

(e)

Part A

Average Absolute Distance $= \sum_{i=1}^{10} |0.5 - \pi_i| = 2.2$

Part C

Output from part C

Average Absolute Distance $= \sum_{i=1}^{10} |0.5 - \pi_i| = 0.77084.$

So as we can see here, the average absolute distance is smaller when we used with part c is smaller compare to the part A. It is because the variance of the estimate is smaller in part c compared to the data in part a.

So the estimate is more close to the true value.

Q4

Q4

(a). From the output from the question. $\hat{\beta_1} = 4.2227$. $\hat{\beta_3} = -0.7751$

∴ $\hat{\beta_1} - \hat{\beta_3} = 4.9978$ ∴ $e^{\hat{\beta_1} - \hat{\beta_3}} = 148.0870$.

∴ The odds ratio of cigarettes (Yes) ~~is~~ ~~#~~ to marijuana (Yes) is 148.0870.

(b). The value of $\hat{\sigma}$ implys the difference between subjects. So the large value indicates a big / huge variation. between subjects

(c). As we know, u_i follows a normal distribution. so if the value is positive, it indicates a big random effect. If the value is too large, it may over the scope, as an outlier.