Question 1

(a) As X = 1, $logit\{\pi(x=1,z)\} = \beta x + \beta_k^Z$; X=0, $logit\{\pi(x=0,z)\} = \beta_k^Z$;

So $\beta = logit\{\pi(x=1,z)\} - logit\{\pi(x=0,z)\}$, $e^{\beta} = \frac{\pi(x=1,z)/\{1-\pi(x=1,z)\}}{\pi(x=0,z)/\{1-\pi(x=0,z)\}}$

From the equation above, we can see the odds-ratio between X and Y are not decided by Z, so it implies the common odds-ratio between X and Y across Z = 1,2,3,4.

(b)

Analysis Of Maximum Likelihood Parameter Estimates								
Parameter		**DF**	**Estimate**	**Standard Error**	**Likelihood Ratio 95% Confidence Limits**		**Wald Chi-Square**	**Pr > ChiSq**
Intercept		0	0.0000	0.0000	0.0000	0.0000	.	.
center	1	1	-1.3583	0.5619	-2.5450	-0.3138	5.84	0.0156
center	2	1	-1.0172	0.4569	-1.9644	-0.1547	4.96	0.0260
center	3	1	-0.3715	0.4886	-1.3527	0.5856	0.58	0.4470
center	4	1	-0.4670	0.4723	-1.4212	0.4522	0.98	0.3228
trt		1	0.9340	0.4343	0.0935	1.8039	4.62	0.0315
Scale		0	1.0000	0.0000	1.0000	1.0000		

From the output above, we can see the estimate of $\hat{\beta}$ is 0.9340; $\hat{\beta}_1^z$ is -1.3583, $\hat{\beta}_2^z$ is -1.0172, $\hat{\beta}_3^z$ is -0.3715, $\hat{\beta}_4^z$ is -0.4670.

$e^{\hat{\beta}}$ interpretation: For each center, the odds of success under new treatments are about 2.54 times of odds of success under old treatments.

95% CI for $e^{\hat{\beta}}$: $(e^{0.0935}, e^{1.8039}) = (1.0980, 6.0733)$

(c) As $\mu_{11k} = \frac{n_{1+k}n_{+1k}}{n_{++k}}$ and $var = \frac{n_{1+k}n_{2+k}n_{+1k}n_{+2k}}{n_{++k}^2(n_{++k}-1)}$

$$n_{1+1} = 10, n_{+11} = 6, n_{2+1} = 10, n_{+21} = 14, n_{++1} = 20, \mu_{111} = 3, var = 1.105$$

$$n_{1+2} = 12, n_{+12} = 10, n_{2+2} = 16, n_{+22} = 18, n_{++2} = 28, \mu_{112} = 4.29, var = 1.633$$

$$n_{1+3} = 14, n_{+13} = 13, n_{2+3} = 10, n_{+23} = 11, n_{++3} = 24, \mu_{112} = 7.58, var = 1.511$$

$$n_{1+4} = 12, n_{+14} = 12, n_{2+4} = 12, n_{+24} = 12, n_{++4} = 24, \mu_{112} = 6, var = 1.565$$

$$\chi^2 = \frac{\{(4-3)+(6-4.29)+(8-7.58)+(8-6)\}^2}{(1.105+1.633+1.511+1.565)} \approx 4.526$$

And as $\chi^2_{0.05,1} = 3.841$ so our test statistic is bigger than $\chi^2_{0.05,1}$ at 0.05 level. Therefore, we reject H_0, X and Y are not independent when conditional on Z.

(d)

Criteria For Assessing Goodness Of Fit			
Criterion	DF	Value	Value/DF
Deviance	3	0.9201	0.3067
Scaled Deviance	3	0.9201	0.3067
Pearson Chi-Square	3	0.9263	0.3088
Scaled Pearson X2	3	0.9263	0.3088
Log Likelihood		-61.2778	
Full Log Likelihood		-11.8429	
AIC (smaller is better)		33.6857	
AICC (smaller is better)		63.6857	
BIC (smaller is better)		34.0829	

As $\chi^2_{0.05,3} = 7.815$, Our Pearson Chi-square test statistic is 0.9263 which is smaller than 7.815, so we fail to reject H_0, our model fits the data adequately.

(e)

Exact Conditional Tests				
			p-Value	
Effect	Test	Statistic	Exact	Mid
trt	Score	4.5279	0.0388	0.0301
	Probability	0.0173	0.0388	0.0301

The test statistic here is 4.5279. The exact p-value here is 0.0388 which is smaller than 0.05, so we reject H_0, X and Y are not independent conditionally on Z at the significance level of 0.05.

(f)

Analysis of Conditional Maximum Likelihood Estimates					
Parameter	DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
trt	1	0.8939	0.4243	4.4377	0.0352

$$logit\{\pi(x)\} = 0.8939x$$

From the conditional fit the estimate of $\hat{\beta}$ is 0.8939.

Testing Global Null Hypothesis: BETA=0			
Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	4.5515	1	0.0329
Score	4.5279	1	0.0333
Wald	4.4377	1	0.0352

From the likelihood Ratio Test, the p-value is 0.0329 which is smaller than 0.05, so we reject H_0, the beta is not equal to 0, so X is not independent to Y as conditionally on Z at significance level of 0.05.

(g) Exact conditional inference – Exact CMH (Cochran–Mantel–Haenszel) test.

SAS Codes for question 1:

```
data question1b;
input center trt S F;
n =S + F;
datalines;
1 1 4 6
1 0 2 8
2 1 6 6
2 0 4 12
3 1 8 6
3 0 5 5
4 1 8 4
4 0 4 8
;

proc genmod data = question1b;
class center;
model S/n = center trt / dist = bin link=logit type3 lcri noint;
run;

proc logistic data=question1b;
class center / param=ref;
model S/n = center trt;
exact trt;
run;
```

```
proc logistic data = question1b;
class center;
model S/n = trt;
strata center;
run;
```

Question 2

Y	$\hat{\pi}$	$\hat{Y}_{0.3^-}$	$\hat{Y}_{0.4^-}$	$\hat{Y}_{0.5^-}$	$\hat{Y}_{0.6^-}$	$\hat{Y}_{0.7^-}$	$\hat{Y}_{0.8^-}$	$\hat{Y}_{0.8^+}$
1	0.8	1	1	1	1	1	1	0
1	0.6	1	1	1	1	0	0	0
1	0.4	1	1	0	0	0	0	0
0	0.7	1	1	1	1	1	0	0
0	0.5	1	1	1	0	0	0	0
0	0.4	1	1	0	0	0	0	0
0	0.3	1	0	0	0	0	0	0

Corresponding Classification Tables:

Y	1	0
1	3	0
0	4	0

se = 1
sp = 0

	1	2
	0	4

se = $\frac{1}{3}$
sp = 1

3	0
3	1

se = 1
sp = $\frac{1}{4}$

0	3
0	4

se = 0
sp = 1

2	1
2	2

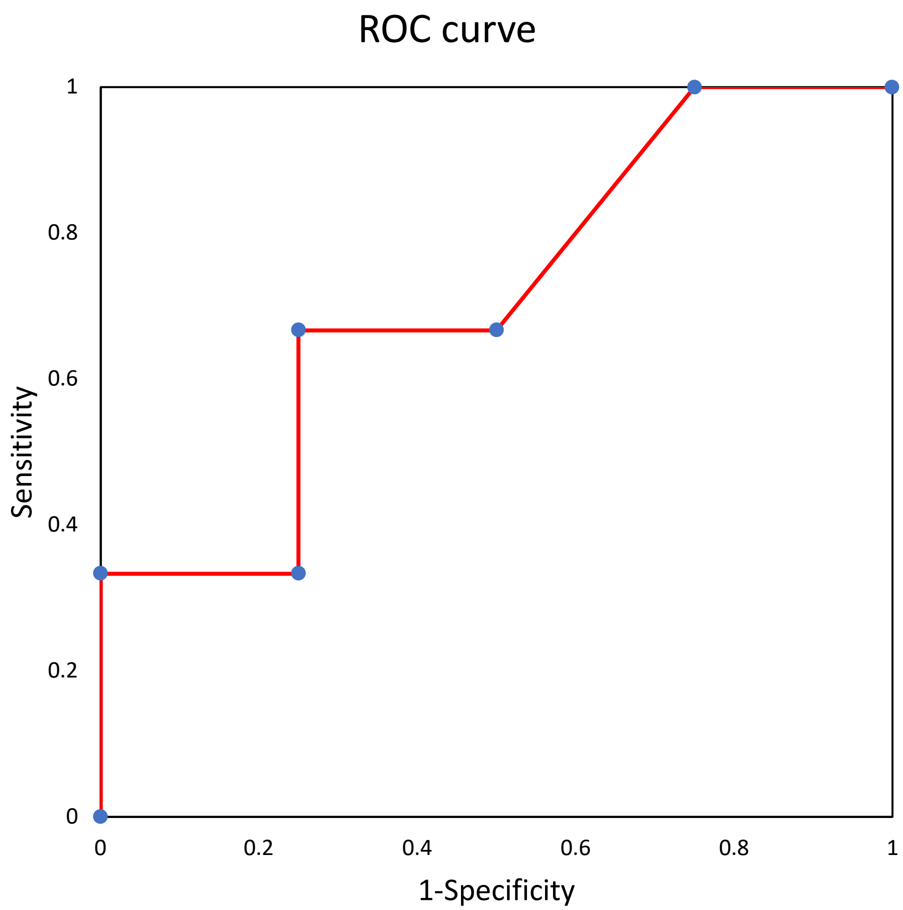
se = $\frac{2}{3}$
sp = $\frac{1}{2}$

2	1
1	3

se = $\frac{2}{3}$
sp = $\frac{3}{4}$

1	2
1	3

se = $\frac{1}{3}$
sp = $\frac{3}{4}$



AUC = 8.5/12 = 0.7083: Proportion of concordant pairs in $(Y_i, \hat{\pi}_i)$ among all pairs with different outcome Y_i plus half of number of ties with different outcomes.

8.5 = # of concordant pairs (8) + 0.5 * # of ties in $\hat{\pi}_i$ with different outcomes.

12 = # of pairs with different outcomes.

Question 3

(a)

Analysis of Maximum Likelihood Estimates						
Parameter		DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept	1	1	-0.3471	0.2994	1.3446	0.2462
Intercept	2	1	0.8607	0.3073	7.8442	0.0051
x		1	0.8978	0.3296	7.4191	0.0065
z		1	-0.3781	0.3265	1.3407	0.2469

$logit\{\tau(x,z)\} = \hat{\alpha}_j + 0.8978x - 0.3781z \quad j = 1,2 \quad \hat{\alpha}_1 = -0.3471 \; \hat{\alpha}_2 = 0.8607$

(x: method, z: gender)

(b)

Odds Ratio Estimates			
Effect	Point Estimate	95% Wald Confidence Limits	
x	2.454	1.286	4.683
z	0.685	0.361	1.299

The estimate of odds ratio of recovery complete between the treatment and placebo for patients with same gender is 2.454, 95% confidence interval is (1.286, 4.683).

(c)

Deviance and Pearson Goodness-of-Fit Statistics				
Criterion	Value	DF	Value/DF	Pr > ChiSq
Deviance	0.4599	4	0.1150	0.9773
Pearson	0.4652	4	0.1163	0.9768

Deviance of this model is 0.4599 and degrees of freedom is 4.

Degrees of freedom = (I-1)(J-1) – dim(x) = (4-1)*(3-1)-2 = 4

As $\chi^2_{4,0.05} = 9.488$ here, our value is 0.4652 which is smaller than 9.488, so we fail to reject H_0. The model fits the data well.

(d)

Score Test for the Proportional Odds Assumption		
Chi-Square	DF	Pr > ChiSq
0.1941	2	0.9075

Score test statistic is 0.1941.

Degrees of freedom = (J-2)*dim(x) = (3-2)*2=2

Alternative model: $\log\left\{\frac{\tau_j(x)}{1-\tau_j(x)}\right\} = \hat{\alpha}_j + \hat{\beta}_j x + \hat{\gamma}_j z$

As p -value here is 0.9075 > 0.05, we fail to reject H0, so our null hypothesis model fits the data well.

(e) $\tau_1(x) = \frac{e^{-0.3471+0.8978-0.3781}}{1+e^{-0.3471+0.8978-0.3781}} = 0.5430$ $\tau_2(x) = \frac{e^{0.8607+0.8978-0.3781}}{1+e^{0.8607+0.8978-0.3781}} = 0.7991$

$\pi_1(x) = \tau_1(x) = 0.5430$

$\pi_2(x) = \tau_2(x) - \tau_1(x) = 0.7991 - 0.5430 = 0.2561$

$\pi_3(x) = 1 - \pi_1(x) - \pi_2(x) = 0.2009$

(f)

Deviance and Pearson Goodness-of-Fit Statistics				
Criterion	Value	DF	Value/DF	Pr > ChiSq
Deviance	0.3495	2	0.1747	0.8397
Pearson	0.3494	2	0.1747	0.8397

Deviance is 0.3495, and degrees of freedom is 2.

Df = (J-1)*[(I-1)-# of x's] = (3-1)*[(4-1)-2]=2

As $\chi^2_{4,0.05} = 5.991$ is bigger than our deviance which is 0.3495, so we fail to reject H0, the model fits the data well.

Analysis of Maximum Likelihood Estimates						
Parameter	y	DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept	2	1	-0.2670	0.3731	0.5122	0.4742
Intercept	3	1	-0.3569	0.3847	0.8610	0.3535
X	2	1	-0.7426	0.4228	3.0854	0.0790
X	3	1	-1.1121	0.4372	6.4696	0.0110
z	2	1	0.1428	0.4176	0.1169	0.7324
z	3	1	0.5123	0.4360	1.3808	0.2400

From the output above, $\log\left(\frac{\hat{\pi}_2}{\hat{\pi}_1}\right) = -0.2670 - 0.7426x + 0.1428z$

$\log\left(\frac{\hat{\pi}_3}{\hat{\pi}_1}\right) = -0.3569 - 1.1121x + 0.5123z$

For all males with receiving treatments, x = z =1;

So $\hat{\pi}_1(x) = \frac{1}{1+e^{-0.2670-0.7426+0.1428}+e^{-0.3569-1.1121+0.5123}} = 0.5542$

$\hat{\pi}_2(x) = \frac{e^{-0.2670-0.7426+0.1428}}{1+e^{-0.2670-0.7426+0.1428}+e^{-0.3569-1.1121+0.5123}} = 0.2329$

$\hat{\pi}_3(x) = \frac{e^{-0.3569-1.1121+0.5123}}{1+e^{-0.2670-0.7426+0.1428}+e^{-0.3569-1.1121+0.5123}} = 0.2129$

Compared to above probabilities calculated above, cell probabilities for male receiving treatments are similar to what we calculated above.

SAS codes for question 3:

```
data question3a;
input gender $ X y1-y3;
z = (gender='Male');
datalines;
Male 1 22 10 8
Male 0 10 8 12
Female 1 24 8 6
Female 0 12 10 8
;

data question3b;
set question3a;
array temp {3} y1-y3;
do y=1 to 3;
count = temp(y);
output;
end;
run;
```

```
proc logistic data = question3b;
freq count;
model y = x z / aggregate scale = none;
run;

proc logistic data=question3b;
freq count;
model y (ref='1')= x z/ link=glogit aggregate scale=none;
run;
```