Chapter 4 SAS 프로그램 입력 및 출력결과

```
/********************************
         Chapter 4: SAS Programs
********************************/
/*read data from the path of dataset*/
Data dat;
infile "C:/Users/HSY/Desktop/DBP.csv" delimiter="," firstobs=2;
input Subject TRT$ DBP1 DBP2 DBP3 DBP4 DBP5 Age Sex$;
/*create a column diff*/
diff= DBP5-DBP1;
RUN;
proc print data=dat;
run;
```

OBS	Subject	TRT	DBP1	DBP2	DBP3	DBP4	DBP5	Age	Sex	diff
1	1	A	114	115	113	109	105	43	F	-9
2	2	A	116	113	112	103	101	51	M	-15
3	3	A	119	115	113	104	98	48	F	-21
4	4	A	115	113	112	109	101	42	F	-14
5	5	A	116	112	107	104	105	49	M	-11
6	6	A	117	112	113	104	102	47	M	-15
7	7	A	118	111	100	109	99	50	F	-19
8	8	A	120	115	113	102	102	61	M	-18
9	9	A	114	112	113	109	103	43	M	-11
10	10	A	115	113	108	106	97	51	M	-18
11	11	A	117	112	110	109	101	47	F	-16
12	12	A	116	115	113	109	102	45	M	-14
13	13	A	119	117	110	106	104	54	F	-15
14	14	A	118	115	113	102	99	52	M	-19
15	15	A	115	112	108	105	102	42	M	-13
16	16	A	114	111	111	107	100	44	F	-14
17	17	A	117	114	110	108	102	48	M	-15
18	18	A	120	115	113	107	103	63	F	-17
19	19	A	114	113	109	104	100	41	M	-14
20	20	A	117	115	113	109	101	51	M	-16
21	21	B	114	115	113	111	113	39	M	-1
22	22	B	116	114	114	109	110	40	F	-6
23	23	B	114	115	113	111	109	39	F	-5
24	24	B	114	115	113	114	115	38	M	1
25	25	B	116	113	113	109	109	39	F	-7
26	26	B	114	115	114	111	110	41	M	-4
27	27	B	119	118	118	117	115	56	F	-4
28	28	B	118	117	117	116	112	56	M	-6
29	29	B	114	113	113	109	108	38	M	-6
30	30	B	120	115	113	113	113	57	M	-7
31	31	B	117	115	113	114	115	47	F	-2
32	32	B	118	114	112	109	110	48	M	-8
33	33	B	121	119	117	114	115	61	F	-6
34	34	B	116	115	116	114	111	49	M	-5
35	35	B	118	118	113	113	112	52	M	-6
36	36	B	119	115	115	114	111	55	F	-8
37	37	B	116	114	113	109	109	45	F	-7
38	38	B	116	115	114	114	112	42	M	-4
39	39	B	117	115	113	114	115	49	F	-2
40	40	B	118	114	114	114	115	50	F	-3

```
/************************************************
                Section 4.3.1.1
************************************************/
/*test whether the DBP means are different*/
PROC TTEST data= dat;
class TRT;
var DBP1;
RUN;
```

```
                        The TTEST Procedure

                        Variable:  DBP1

        TRT             N       Mean    Std Dev    Std Err    Minimum    Maximum

        A              20      116.6     1.9861     0.4441      114.0      120.0
        B              20      116.8     2.1244     0.4750      114.0      121.0
        Diff (1-2)            -0.2000     2.0564     0.6503

    TRT         Method          Mean      95% CL Mean        Std Dev    95% CL Std Dev

    A                          116.6     115.6    117.5       1.9861    1.5104    2.9009
    B                          116.8     115.8    117.7       2.1244    1.6156    3.1029
    Diff (1-2)  Pooled        -0.2000   -1.5165   1.1165       2.0564    1.6806    2.6503
    Diff (1-2)  Satterthwaite -0.2000   -1.5167   1.1167

            Method          Variances        DF     t Value    Pr > |t|

            Pooled          Equal            38      -0.31      0.7601
            Satterthwaite   Unequal      37.829      -0.31      0.7601

                        Equality of Variances

            Method      Num DF    Den DF    F Value    Pr > F

            Folded F       19        19      1.14      0.7722
```

```
/*make 2 by 2 table using variable "Sex"*/
PROC FREQ data= dat;
tables Trt*Sex / out= freqs ;
RUN;
proc print data=freqs;
run;
```

```
                        FREQ 프로시저

                        테이블:TRT * Sex

            TRT         Sex

            빈도      |
            백분율    |
            행 백분율 |
            칼럼 백분율|F      |M       |    합계
            -----------+--------+--------+
            A          |     8 |     12 |     20
                       | 20.00 | 30.00 |  50.00
                       | 40.00 | 60.00 |
                       | 44.44 | 54.55 |
            -----------+--------+--------+
            B          |    10 |     10 |     20
                       | 25.00 | 25.00 |  50.00
                       | 50.00 | 50.00 |
                       | 55.56 | 45.45 |
            -----------+--------+--------+
            합계             18       22      40
                          45.00    55.00  100.00
```

OBS	TRT	Sex	COUNT	PERCENT
1	A	F	8	20
2	A	M	12	30
3	B	F	10	25
4	B	M	10	25

```
PROC TRANSPOSE data= freqs out= SexbyTrt(drop=_:);
   id  Sex;
   var count;
   by  Trt;
RUN;
/*print the table*/
PROC PRINT data= SexbyTrt;
RUN;
```

OBS	TRT	F	M
1	A	8	12
2	B	10	10

```
/*test equality of proportions of 2 treatment groups
using Pearson's Chi squares*/
PROC FREQ data= freqs;
weight count;
tables Trt*Sex/ chisq;
RUN;
```

FREQ 프로시저

테이블:TRT * Sex

```
TRT         Sex

빈도       |
백분율     |
행 백분율  |
칼럼 백분율|F       |M       |      합계
-----------+--------+--------+
A          |      8 |     12 |     20
           |  20.00 |  30.00 |  50.00
           |  40.00 |  60.00 |
           |  44.44 |  54.55 |
-----------+--------+--------+
B          |     10 |     10 |     20
           |  25.00 |  25.00 |  50.00
           |  50.00 |  50.00 |
           |  55.56 |  45.45 |
-----------+--------+--------+
합계             18       22       40
              45.00    55.00   100.00
```

TRT * Sex 테이블에 대한 통계량

통계량	자유도	값	Prob
카이제곱	1	0.4040	0.5250
우도비 카이제곱	1	0.4048	0.5246
연속성 수정 카이제곱	1	0.1010	0.7506
Mantel-Haenszel 카이제곱	1	0.3939	0.5302
파이 계수		-0.1005	
우발성 계수		0.1000	
크래머의 V		-0.1005	

Fisher의 정확 검정

(1,1) 셀 빈도(F)   8
하단측 p값 Pr <= F   0.3756
상단측 p값 Pr >= F   0.8297

테이블 확률 (P)   0.2053
양측 p값 Pr <= P   0.7512

표본 크기 = 40

```
/*Fit the main effect model on "Sex" and "Age"*/
PROC GLM data= dat;
class Sex(ref="F");
model DBP1= Sex Age / solution;
RUN;
```

                            The GLM Procedure

                        Class Level Information

                    Class   Levels   Values

                    Sex   2   M F

                    Number of Observations Read   40
                    Number of Observations Used   40

    Dependent Variable: DBP1

                                        Sum of
            Source   DF   Squares   Mean Square   F Value   Pr > F

            Model   2   128.0280063   64.0140031   71.62   <.0001

            Error   37   33.0719937   0.8938377

            Corrected Total   39   161.1000000

                    R-Square   Coeff Var   Root MSE   DBP1 Mean

                    0.794711   0.810484   0.945430   116.6500

            Source   DF   Type I SS   Mean Square   F Value   Pr > F

            Sex   1   8.7363636   8.7363636   9.77   0.0034
            Age   1   119.2916426   119.2916426   133.46   <.0001

            Source   DF   Type III SS   Mean Square   F Value   Pr > F

            Sex   1   4.0532688   4.0532688   4.53   0.0399
            Age   1   119.2916426   119.2916426   133.46   <.0001

                                        Standard
            Parameter   Estimate   Error   t Value   Pr > |t|

            Intercept   104.3833726 B   1.12875520   92.48   <.0001
            Sex   M   -0.6422010 B   0.30157658   -2.13   0.0399
            Sex   F   0.0000000 B   .   .   .
            Age   0.2638753   0.02284142   11.55   <.0001

NOTE: The X'X matrix has been found to be singular, and a generalized inverse was used to solve the normal
equations. Terms whose estimates are followed by the letter 'B' are not uniquely estimable.
```

```
/*********************************************
        Section 4.3.1.2
*********************************************/
/*stepwise model selection */
PROC GLMSELECT data= dat;
class Trt Sex;
model diff= Trt|Sex|Age
/ selection= stepwise(select= AIC) stats= all;
RUN;
```

The GLMSELECT Procedure

Data Set	WORK.DAT
Dependent Variable	diff
Selection Method	Stepwise
Select Criterion	AIC
Stop Criterion	AIC
Effect Hierarchy Enforced	None

Number of Observations Read	40
Number of Observations Used	40

Class Level Information

Class	Levels	Values
TRT	2	A B
Sex	2	F M

Dimensions

Number of Effects	8
Number of Parameters	18

Stepwise Selection Summary

Step	Effect Entered	Effect Removed	Number Effects In	Number Parms In	Model R-Square	Adjusted R-Square	AIC	AICC	BIC	CP
0	Intercept		1	1	0.0000	0.0000	185.0544	185.3787	142.0692	184.3130
1	Age*TRT		2	3	0.8393	0.8307*	115.9144*	117.0572*	76.6053*	1.7155*

* Optimal Value of Criterion

Stepwise Selection Summary

Step	Effect Entered	Effect Removed	SBC	PRESS	ASE	F Value	Pr > F
0	Intercept		144.7433	1430.6377	34.0000	0.00	1.0000
1	Age*TRT		78.9810*	252.4546*	5.4622	96.65	<.0001

* Optimal Value of Criterion

Selection stopped at a local minimum of the AIC criterion.

Stop Details

Candidate For	Effect	Candidate AIC		Compare AIC
Entry	Sex	117.6498	>	115.9144
Removal	Age*TRT	185.0544	>	115.9144

Selected Model

The selected model is the model at the last step (Step 1).

Effects: Intercept Age*TRT

Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Value
Model	2	1141.51043	570.75521	96.65
Error	37	218.48957	5.90512	
Corrected Total	39	1360.00000		

Root MSE	2.43005
Dependent Mean	-10.00000
R-Square	0.8393
Adj R-Sq	0.8307
AIC	115.91435
AICC	117.05721
BIC	76.60535
C(p)	1.71549
PRESS	252.45464
SBC	78.98099
ASE	5.46224

Parameter Estimates

Parameter		DF	Estimate	Standard Error	t Value
Intercept		1	-0.538700	2.834116	-0.19
Age*TRT	A	1	-0.301210	0.058563	-5.14
Age*TRT	B	1	-0.091047	0.059960	-1.52

```
/*fit the reduced model*/
PROC GLM data= dat;
class TRT(ref="A");
model diff= Age Trt / solution;
RUN;
```

The GLM Procedure

Class Level Information

Class	Levels	Values
TRT	2	B A

Number of Observations Read   40
Number of Observations Used   40

Dependent Variable: diff

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	2	1132.669217	566.334608	92.18	<.0001
Error	37	227.330783	6.144075		
Corrected Total	39	1360.000000			

R-Square	Coeff Var	Root MSE	diff Mean
0.832845	-24.78725	2.478725	-10.00000

Source	DF	Type I SS	Mean Square	F Value	Pr > F
Age	1	120.488476	120.488476	19.61	<.0001
TRT	1	1012.180741	1012.180741	164.74	<.0001

Source	DF	Type III SS	Mean Square	F Value	Pr > F
Age	1	51.069217	51.069217	8.31	0.0065
TRT	1	1012.180741	1012.180741	164.74	<.0001

Parameter		Estimate	Standard Error	t Value	Pr > |t|
Intercept		-6.78085500 B	2.97236119	-2.28	0.0284
Age		-0.17323344	0.06008698	-2.88	0.0065
TRT	B	10.13148817 B	0.78935519	12.84	<.0001
TRT	A	0.00000000 B	.	.	.

NOTE: The X'X matrix has been found to be singular, and a generalized inverse was used to solve the normal
equations. Terms whose estimates are followed by the letter 'B' are not uniquely estimable.

```
/******************************
Section 4.3.1.3: MACNOVA for Treatment Difference
*************************************/
/* create the data*/
data mdat;
set dat;
diff2to1 = DBP2-DBP1;
diff3to1 = DBP3-DBP1;
diff4to1 = DBP4-DBP1;
diff5to1 = DBP5-DBP1;
run;
proc print data=mdat;
run;
```

OBS	Subject	TRT	DBP1	DBP2	DBP3	DBP4	DBP5	Age	Sex	diff	diff2to1	diff3to1	diff4to1	diff5to1
1	1	A	114	115	113	109	105	43	F	-9	1	-1	-5	-9
2	2	A	116	113	112	103	101	51	M	-15	-3	-4	-13	-15
3	3	A	119	115	113	104	98	48	F	-21	-4	-6	-15	-21
4	4	A	115	113	112	109	101	42	F	-14	-2	-3	-6	-14
5	5	A	116	112	107	104	105	49	M	-11	-4	-9	-12	-11
6	6	A	117	112	113	104	102	47	M	-15	-5	-4	-13	-15
7	7	A	118	111	100	109	99	50	F	-19	-7	-18	-9	-19
8	8	A	120	115	113	102	102	61	M	-18	-5	-7	-18	-18
9	9	A	114	112	113	109	103	43	M	-11	-2	-1	-5	-11
10	10	A	115	113	108	106	97	51	M	-18	-2	-7	-9	-18
11	11	A	117	112	110	109	101	47	F	-16	-5	-7	-8	-16
12	12	A	116	115	113	109	102	45	M	-14	-1	-3	-7	-14
13	13	A	119	117	110	106	104	54	F	-15	-2	-9	-13	-15
14	14	A	118	115	113	102	99	52	M	-19	-3	-5	-16	-19
15	15	A	115	112	108	105	102	42	M	-13	-3	-7	-10	-13
16	16	A	114	111	111	107	100	44	F	-14	-3	-3	-7	-14
17	17	A	117	114	110	108	102	48	M	-15	-3	-7	-9	-15
18	18	A	120	115	113	107	103	63	F	-17	-5	-7	-13	-17
19	19	A	114	113	109	104	100	41	M	-14	-1	-5	-10	-14
20	20	A	117	115	113	109	101	51	M	-16	-2	-4	-8	-16
21	21	B	114	115	113	111	113	39	M	-1	1	-1	-3	-1
22	22	B	116	114	114	109	110	40	F	-6	-2	-2	-7	-6
23	23	B	114	115	113	111	109	39	F	-5	1	-1	-3	-5
24	24	B	114	115	113	114	115	38	M	1	1	-1	0	1
25	25	B	116	113	113	109	109	39	F	-7	-3	-3	-7	-7
26	26	B	114	115	114	111	110	41	M	-4	1	0	-3	-4
27	27	B	119	118	118	117	115	56	F	-4	-1	-1	-2	-4
28	28	B	118	117	117	116	112	56	M	-6	-1	-1	-2	-6
29	29	B	114	113	113	109	108	38	M	-6	-1	-1	-5	-6
30	30	B	120	115	113	113	113	57	M	-7	-5	-7	-7	-7
31	31	B	117	115	113	114	115	47	F	-2	-2	-4	-3	-2
32	32	B	118	114	112	109	110	48	M	-8	-4	-6	-9	-8
33	33	B	121	119	117	114	115	61	F	-6	-2	-4	-7	-6
34	34	B	116	115	116	114	111	49	M	-5	-1	0	-2	-5
35	35	B	118	118	113	113	112	52	M	-6	0	-5	-5	-6
36	36	B	119	115	115	114	111	55	F	-8	-4	-4	-5	-8
37	37	B	116	114	113	109	109	45	F	-7	-2	-3	-7	-7
38	38	B	116	115	114	114	112	42	M	-4	-1	-2	-2	-4
39	39	B	117	115	113	114	115	49	F	-2	-2	-4	-3	-2
40	40	B	118	114	114	114	115	50	F	-3	-4	-4	-4	-3

```
/* manova using glm*/
PROC glm data= mdat;
class TRT;
model diff2to1 diff3to1 diff4to1 diff5to1= TRT Age/ss3;
contrast '1 vs 2' TRT 1 -1;
manova h=_all_;
RUN;
```

The GLM Procedure

Class Level Information

Class	Levels	Values
TRT	2	A B

Number of Observations Read 40
Number of Observations Used 40

Dependent Variable: diff2to1

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	2	51.6829719	25.8414860	9.89	0.0004
Error	37	96.7170281	2.6139737		
Corrected Total	39	148.4000000			

R-Square	Coeff Var	Root MSE	diff2to1 Mean
0.348268	-70.29473	1.616779	-2.300000

Source	DF	Type III SS	Mean Square	F Value	Pr > F
TRT	1	16.58847237	16.58847237	6.35	0.0162
Age	1	29.18297194	29.18297194	11.16	0.0019

Contrast	DF	Contrast SS	Mean Square	F Value	Pr > F
1 vs 2	1	16.58847237	16.58847237	6.35	0.0162

Dependent Variable: diff3to1

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	2	148.3384479	74.1692240	9.61	0.0004
Error	37	285.6365521	7.7199068		
Corrected Total	39	433.9750000			

R-Square	Coeff Var	Root MSE	diff3to1 Mean
0.341813	-64.99350	2.778472	-4.275000

Source	DF	Type III SS	Mean Square	F Value	Pr > F
TRT	1	82.16915840	82.16915840	10.64	0.0024
Age	1	49.11344792	49.11344792	6.36	0.0161

Contrast	DF	Contrast SS	Mean Square	F Value	Pr > F
1 vs 2	1	82.16915840	82.16915840	10.64	0.0024

Dependent Variable: diff4to1

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	2	426.7762068	213.3881034	26.35	<.0001
Error	37	299.6237932	8.0979404		
Corrected Total	39	726.4000000			

R-Square	Coeff Var	Root MSE	diff4to1 Mean
0.587522	-38.98203	2.845688	-7.300000

Source	DF	Type III SS	Mean Square	F Value	Pr > F
TRT	1	319.5860981	319.5860981	39.47	<.0001
Age	1	66.7762068	66.7762068	8.25	0.0067

Contrast	DF	Contrast SS	Mean Square	F Value	Pr > F
1 vs 2	1	319.5860981	319.5860981	39.47	<.0001

Dependent Variable: diff5to1

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	2	1132.669217	566.334608	92.18	<.0001
Error	37	227.330783	6.144075		
Corrected Total	39	1360.000000			

R-Square	Coeff Var	Root MSE	diff5to1 Mean
0.832845	-24.78725	2.478725	-10.00000

Source	DF	Type III SS	Mean Square	F Value	Pr > F
TRT	1	1012.180741	1012.180741	164.74	<.0001
Age	1	51.069217	51.069217	8.31	0.0065

Contrast	DF	Contrast SS	Mean Square	F Value	Pr > F
1 vs 2	1	1012.180741	1012.180741	164.74	<.0001

Multivariate Analysis of Variance

Characteristic Roots and Vectors of: E Inverse * H, where
H = Type III SSCP Matrix for TRT
E = Error SSCP Matrix

Characteristic Root	Percent	Characteristic Vector V'EV=1 diff2to1	diff3to1	diff4to1	diff5to1
4.82671479	100.00	-0.03557795	0.00667843	-0.00217534	0.07292943
0.00000000	0.00	-0.08658748	0.07395376	0.00030295	-0.01015642
0.00000000	0.00	0.09807889	0.01036329	0.00047048	-0.01577305
0.00000000	0.00	0.02813182	-0.00095893	-0.07179300	0.03701285

MANOVA Test Criteria and Exact F Statistics for the Hypothesis of No Overall TRT Effect
H = Type III SSCP Matrix for TRT
E = Error SSCP Matrix

S=1   M=1   N=16

Statistic	Value	F Value	Num DF	Den DF	Pr > F
Wilks' Lambda	0.17162330	41.03	4	34	<.0001
Pillai's Trace	0.82837670	41.03	4	34	<.0001
Hotelling-Lawley Trace	4.82671479	41.03	4	34	<.0001
Roy's Greatest Root	4.82671479	41.03	4	34	<.0001

Characteristic Roots and Vectors of: E Inverse * H, where
H = Type III SSCP Matrix for Age
E = Error SSCP Matrix

Characteristic Root	Percent	Characteristic Vector V'EV=1 diff2to1	diff3to1	diff4to1	diff5to1
0.40347759	100.00	0.05099586	0.00980365	0.01912575	0.01885157
0.00000000	0.00	-0.08952795	0.07405131	-0.00232116	-0.00228788
0.00000000	0.00	-0.04144328	-0.00585332	-0.03905817	0.08173113
0.00000000	0.00	-0.08270836	-0.00284742	0.05711880	0.00000000

MANOVA Test Criteria and Exact F Statistics for the Hypothesis of No Overall Age Effect
H = Type III SSCP Matrix for Age
E = Error SSCP Matrix

S=1   M=1   N=16

Statistic	Value	F Value	Num DF	Den DF	Pr > F
Wilks' Lambda	0.71251583	3.43	4	34	0.0185
Pillai's Trace	0.28748417	3.43	4	34	0.0185
Hotelling-Lawley Trace	0.40347759	3.43	4	34	0.0185
Roy's Greatest Root	0.40347759	3.43	4	34	0.0185

Characteristic Roots and Vectors of: E Inverse * H, where
H = Contrast SSCP Matrix for 1 vs 2
E = Error SSCP Matrix

Characteristic Root	Percent	Characteristic Vector V'EV=1 diff2to1	diff3to1	diff4to1	diff5to1
4.82671479	100.00	-0.03557795	0.00667843	-0.00217534	0.07292943
0.00000000	0.00	-0.08658748	0.07395376	0.00030295	-0.01015642

```
                    Characteristic Roots and Vectors of: E Inverse * H, where
                            H = Contrast SSCP Matrix for 1 vs 2
                                  E = Error SSCP Matrix

        Characteristic                 Characteristic Vector  V'EV=1
                Root      Percent       diff2to1        diff3to1         diff4to1         diff5to1

          0.00000000       0.00        0.09807889      0.01036329       0.00047048      -0.01577305
          0.00000000       0.00        0.02813182     -0.00095893      -0.07179300       0.03701285


          MANOVA Test Criteria and Exact F Statistics for the Hypothesis of No Overall 1 vs 2 Effect
                            H = Contrast SSCP Matrix for 1 vs 2
                                  E = Error SSCP Matrix

                                     S=1    M=1    N=16

             Statistic                     Value    F Value   Num DF    Den DF    Pr > F

             Wilks' Lambda            0.17162330      41.03        4        34    <.0001
             Pillai's Trace           0.82837670      41.03        4        34    <.0001
             Hotelling-Lawley Trace   4.82671479      41.03        4        34    <.0001
             Roy's Greatest Root      4.82671479      41.03        4        34    <.0001
```

```
/*********************************************
              Section 4.3.2
**********************************************/
/*read data from the path of dataset*/
Data dat1;
infile "C:/Users/HSY/Desktop/betablocker.csv" delimiter="," firstobs=2;
input Deaths Total Center Treatment$;
RUN;
proc print data=dat1;
run;
```

OBS	Deaths	Total	Center	Treatment
1	3	39	1	Control
2	14	116	2	Control
3	11	93	3	Control
4	127	1520	4	Control
5	27	365	5	Control
6	6	52	6	Control
7	152	939	7	Control
8	48	471	8	Control
9	37	282	9	Control
10	188	1921	10	Control
11	52	583	11	Control
12	47	266	12	Control
13	16	293	13	Control
14	45	883	14	Control
15	31	147	15	Control
16	38	213	16	Control
17	12	122	17	Control
18	6	154	18	Control
19	3	134	19	Control
20	40	218	20	Control
21	43	364	21	Control
22	39	674	22	Control
23	3	38	1	Treated
24	7	114	2	Treated
25	5	69	3	Treated
26	102	1533	4	Treated
27	28	355	5	Treated
28	4	59	6	Treated
29	98	945	7	Treated
30	60	632	8	Treated
31	25	278	9	Treated
32	138	1916	10	Treated
33	64	873	11	Treated
34	45	263	12	Treated
35	9	291	13	Treated
36	57	858	14	Treated
37	25	154	15	Treated
38	33	207	16	Treated
39	28	251	17	Treated
40	8	151	18	Treated
41	6	174	19	Treated
42	32	209	20	Treated
43	27	391	21	Treated
44	22	680	22	Treated

```
/*fit the logistic regression model*/
PROC GENMOD data= dat1;
class Center(ref="1") Treatment(ref="Control");
model Deaths/Total= Center Treatment/dist= binomial link= logit;
RUN;
```

The GENMOD Procedure

Model Information

Data Set	WORK.DAT1
Distribution	Binomial
Link Function	Logit
Response Variable (Events)	Deaths
Response Variable (Trials)	Total

Number of Observations Read	44
Number of Observations Used	44
Number of Events	1811
Number of Trials	20290

Class Level Information

Class	Levels	Values
Center	22	2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 1
Treatment	2	Treated Control

Response Profile

Ordered Value	Binary Outcome	Total Frequency
1	Event	1811
2	Nonevent	18479

Criteria For Assessing Goodness Of Fit

Criterion	DF	Value	Value/DF
Deviance	21	23.6206	1.1248
Scaled Deviance	21	23.6206	1.1248
Pearson Chi-Square	21	23.5593	1.1219
Scaled Pearson X2	21	23.5593	1.1219
Log Likelihood		-5948.8017	
Full Log Likelihood		-120.5256	
AIC (smaller is better)		287.0511	
AICC (smaller is better)		342.2511	
BIC (smaller is better)		328.0875	

Algorithm converged.

Analysis Of Maximum Likelihood Parameter Estimates

Parameter		DF	Estimate	Standard Error	Wald 95% Confidence Limits		Wald Chi-Square	Pr > ChiSq
Intercept		1	-2.3493	0.4260	-3.1842	-1.5144	30.42	<.0001
Center	2	1	0.1739	0.4832	-0.7731	1.1209	0.13	0.7189
Center	3	1	0.2428	0.5004	-0.7380	1.2236	0.24	0.6275
Center	4	1	-0.0391	0.4309	-0.8837	0.8055	0.01	0.9278
Center	5	1	-0.0217	0.4480	-0.8997	0.8563	0.00	0.9614
Center	6	1	0.1685	0.5395	-0.8888	1.2259	0.10	0.7548
Center	7	1	0.5966	0.4308	-0.2477	1.4410	1.92	0.1661
Center	8	1	0.2715	0.4373	-0.5857	1.1286	0.39	0.5348
Center	9	1	0.3888	0.4462	-0.4859	1.2634	0.76	0.3836
Center	10	1	0.0958	0.4293	-0.7457	0.9373	0.05	0.8234
Center	11	1	0.0520	0.4363	-0.8032	0.9072	0.01	0.9051
Center	12	1	0.9153	0.4406	0.0517	1.7790	4.32	0.0378
Center	13	1	-0.6357	0.4720	-1.5608	0.2894	1.81	0.1780
Center	14	1	-0.3065	0.4375	-1.1639	0.5510	0.49	0.4836
Center	15	1	1.0016	0.4505	0.1186	1.8846	4.94	0.0262
Center	16	1	0.8799	0.4449	0.0079	1.7519	3.91	0.0480
Center	17	1	0.3997	0.4573	-0.4966	1.2959	0.76	0.3821
Center	18	1	-0.5635	0.5059	-1.5549	0.4280	1.24	0.2653
Center	19	1	-1.0144	0.5436	-2.0798	0.0510	3.48	0.0620
Center	20	1	0.8759	0.4447	0.0044	1.7474	3.88	0.0488
Center	21	1	0.1966	0.4436	-0.6728	1.0659	0.20	0.6576
Center	22	1	-0.5812	0.4451	-1.4537	0.2912	1.70	0.1917
Center	1	0	0.0000	0.0000	0.0000	0.0000	.	.
Treatment	Treated	1	-0.2610	0.0499	-0.3589	-0.1631	27.30	<.0001
Treatment	Control	0	0.0000	0.0000	0.0000	0.0000	.	.
Scale		0	1.0000	0.0000	1.0000	1.0000		

NOTE: The scale parameter was held fixed.

```
/*disperson of parameter*/
/*check table "Criteria For Assessing Goodness Of Fit"
line "pearson chi sqaure*/

/*adjust model fitting with estimate of dispersion*/
PROC GENMOD data= dat1;
class Center(ref="1") Treatment(ref="Control");
model Deaths/Total= Center Treatment
/dist= binomial link= logit scale= pearson;
RUN;
```

The GENMOD Procedure

Model Information

Data Set	WORK.DAT1
Distribution	Binomial
Link Function	Logit
Response Variable (Events)	Deaths
Response Variable (Trials)	Total

Number of Observations Read	44
Number of Observations Used	44
Number of Events	1811
Number of Trials	20290

Class Level Information

Class	Levels	Values
Center	22	2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 1
Treatment	2	Treated Control

Response Profile

Ordered Value	Binary Outcome	Total Frequency
1	Event	1811
2	Nonevent	18479

Criteria For Assessing Goodness Of Fit

Criterion	DF	Value	Value/DF
Deviance	21	23.6206	1.1248
Scaled Deviance	21	21.0546	1.0026
Pearson Chi-Square	21	23.5593	1.1219
Scaled Pearson X2	21	21.0000	1.0000
Log Likelihood		-5302.5729	
Full Log Likelihood		-120.5256	
AIC (smaller is better)		287.0511	
AICC (smaller is better)		342.2511	
BIC (smaller is better)		328.0875	

Algorithm converged.

Analysis Of Maximum Likelihood Parameter Estimates

Parameter		DF	Estimate	Standard Error	Wald 95% Confidence Limits		Wald Chi-Square	Pr > ChiSq
Intercept		1	-2.3493	0.4512	-3.2336	-1.4650	27.11	<.0001
Center	2	1	0.1739	0.5118	-0.8291	1.1769	0.12	0.7340
Center	3	1	0.2428	0.5300	-0.7961	1.2817	0.21	0.6469
Center	4	1	-0.0391	0.4564	-0.9337	0.8555	0.01	0.9318
Center	5	1	-0.0217	0.4745	-0.9516	0.9083	0.00	0.9636
Center	6	1	0.1685	0.5714	-0.9514	1.2884	0.09	0.7681
Center	7	1	0.5966	0.4563	-0.2977	1.4910	1.71	0.1910
Center	8	1	0.2715	0.4632	-0.6364	1.1794	0.34	0.5578
Center	9	1	0.3888	0.4727	-0.5376	1.3152	0.68	0.4108
Center	10	1	0.0958	0.4547	-0.7955	0.9871	0.04	0.8331
Center	11	1	0.0520	0.4621	-0.8538	0.9578	0.01	0.9104
Center	12	1	0.9153	0.4667	0.0006	1.8301	3.85	0.0498
Center	13	1	-0.6357	0.4999	-1.6156	0.3442	1.62	0.2035
Center	14	1	-0.3065	0.4634	-1.2147	0.6018	0.44	0.5084
Center	15	1	1.0016	0.4772	0.0663	1.9368	4.41	0.0358
Center	16	1	0.8799	0.4713	-0.0438	1.8036	3.49	0.0619
Center	17	1	0.3997	0.4843	-0.5496	1.3490	0.68	0.4093
Center	18	1	-0.5635	0.5358	-1.6136	0.4867	1.11	0.2930
Center	19	1	-1.0144	0.5758	-2.1429	0.1140	3.10	0.0781
Center	20	1	0.8759	0.4710	-0.0471	1.7990	3.46	0.0629
Center	21	1	0.1966	0.4698	-0.7242	1.1174	0.18	0.6756
Center	22	1	-0.5812	0.4715	-1.5053	0.3429	1.52	0.2177
Center	1	0	0.0000	0.0000	0.0000	0.0000	.	.
Treatment	Treated	1	-0.2610	0.0529	-0.3647	-0.1573	24.34	<.0001
Treatment	Control	0	0.0000	0.0000	0.0000	0.0000	.	.
Scale		0	1.0592	0.0000	1.0592	1.0592		

NOTE: The scale parameter was estimated by the square root of Pearson's Chi-Square/DOF.

```
/*fit quasi-likelihood for binomial data*/
PROC GLIMMIX data= dat1;
class Center(ref="1") Treatment(ref="Control");
model Deaths/Total= Center Treatment
              / link=logit dist=binomial solution;
/*specify the overdispersion parameter*/
random _residual_;
RUN;
```

<div align="center">

The GLIMMIX Procedure

Model Information

Data Set	WORK.DAT1
Response Variable (Events)	Deaths
Response Variable (Trials)	Total
Response Distribution	Binomial
Link Function	Logit
Variance Function	Default
Variance Matrix	Diagonal
Estimation Technique	Maximum Likelihood
Degrees of Freedom Method	Residual

Class Level Information

Class	Levels	Values
Center	22	2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 1
Treatment	2	Treated Control

Number of Observations Read	44
Number of Observations Used	44
Number of Events	1811
Number of Trials	20290

Dimensions

Covariance Parameters	1
Columns in X	25
Columns in Z	0
Subjects (Blocks in V)	1
Max Obs per Subject	44

Optimization Information

Optimization Technique	Newton-Raphson
Parameters in Optimization	23
Lower Boundaries	0
Upper Boundaries	0
Fixed Effects	Not Profiled

Iteration History

Iteration	Restarts	Evaluations	Objective Function	Change	Max Gradient
0	0	4	120.62135	.	9.701853
1	0	3	120.52557325	0.09577675	0.082238
2	0	3	120.52555502	0.00001824	0.000016
3	0	3	120.52555502	0.00000000	1.55E-12

Convergence criterion (GCONV=1E-8) satisfied.

Fit Statistics

-2 Log Likelihood	241.05
AIC (smaller is better)	287.05
AICC (smaller is better)	342.25
BIC (smaller is better)	328.09
CAIC (smaller is better)	351.09
HQIC (smaller is better)	302.27
Pearson Chi-Square	23.56
Pearson Chi-Square / DF	1.12

Parameter Estimates

Effect	Treatment	Center	Estimate	Standard Error	DF	t Value	Pr > |t|
Intercept			-2.3493	0.4512	21	-5.21	<.0001
Center		2	0.1739	0.5118	21	0.34	0.7374
Center		3	0.2428	0.5300	21	0.46	0.6516
Center		4	-0.03907	0.4564	21	-0.09	0.9326
Center		5	-0.02168	0.4745	21	-0.05	0.9640
Center		6	0.1685	0.5714	21	0.29	0.7710
Center		7	0.5966	0.4563	21	1.31	0.2052
Center		8	0.2715	0.4632	21	0.59	0.5641
Center		9	0.3888	0.4727	21	0.82	0.4200
Center		10	0.09580	0.4547	21	0.21	0.8352
Center		11	0.05200	0.4621	21	0.11	0.9115
Center		12	0.9153	0.4667	21	1.96	0.0632
Center		13	-0.6357	0.4999	21	-1.27	0.2174
Center		14	-0.3065	0.4634	21	-0.66	0.5156
Center		15	1.0016	0.4772	21	2.10	0.0481
Center		16	0.8799	0.4713	21	1.87	0.0759
Center		17	0.3997	0.4843	21	0.83	0.4185
Center		18	-0.5635	0.5358	21	-1.05	0.3049
Center		19	-1.0144	0.5758	21	-1.76	0.0926
Center		20	0.8759	0.4710	21	1.86	0.0770
Center		21	0.1966	0.4698	21	0.42	0.6799
Center		22	-0.5812	0.4715	21	-1.23	0.2313
Center		1	0	.	.	.	.
Treatment	Treated		-0.2610	0.05290	21	-4.93	<.0001
Treatment	Control		0	.	.	.	.
Residual			1.1219	.	.	.	.

</div>

```
                    Type III Tests of Fixed Effects

                          Num    Den
             Effect       DF     DF    F Value   Pr > F

             Center       21     21     11.78    <.0001
             Treatment     1     21     24.34    <.0001
```

```
/*Section 4.3.3*/
/*read data from the path of dataset*/
Data dat2;
infile "C:/Users/HSY/Desktop/polyps.csv" delimiter="," firstobs=2;
input number treat$ age;
RUN;
proc print data=dat2;
run;
```

```
                    OBS    number    treat      age

                      1      63      placebo     20
                      2       2      drug        16
                      3      28      placebo     18
                      4      17      drug        22
                      5      61      placebo     13
                      6       1      drug        23
                      7       7      placebo     34
                      8      15      placebo     50
                      9      44      placebo     19
                     10      25      drug        17
                     11       3      drug        23
                     12      28      placebo     22
                     13      10      placebo     30
                     14      40      placebo     27
                     15      33      drug        23
                     16      46      placebo     22
                     17      50      placebo     34
                     18       3      drug        23
                     19       1      drug        22
                     20       4      drug        42
```

```
/*fit poisson regression model*/
PROC GENMOD data= dat2;
class treat(ref="placebo");
model number= age|treat / dist= poisson link= log;
RUN;
```

       The GENMOD Procedure

                              Model Information

                        Data Set   WORK.DAT2
                        Distribution   Poisson
                        Link Function   Log
                        Dependent Variable   number

                   Number of Observations Read   20
                   Number of Observations Used   20

                         Class Level Information

                   Class   Levels   Values

                   treat   2   drug placebo

                     Criteria For Assessing Goodness Of Fit

                  Criterion   DF   Value   Value/DF

                  Deviance   16   179.4906   11.2182
                  Scaled Deviance   16   179.4906   11.2182
                  Pearson Chi-Square   16   182.0059   11.3754
                  Scaled Pearson X2   16   182.0059   11.3754
                  Log Likelihood   1148.2295
                  Full Log Likelihood   -133.9133
                  AIC (smaller is better)   275.8266
                  AICC (smaller is better)   278.4933
                  BIC (smaller is better)   279.8095

             Algorithm converged.

                  Analysis Of Maximum Likelihood Parameter Estimates

                                         Standard   Wald 95% Confidence   Wald
         Parameter   DF   Estimate   Error   Limits   Chi-Square Pr > ChiSq

         Intercept   1   4.5191   0.1534   4.2185   4.8197   868.30   <.0001
         age   1   -0.0384   0.0062 -0.0506   -0.0262   37.81   <.0001
         treat   drug   1   -1.2573   0.4716 -2.1816   -0.3329   7.11   0.0077
         treat   placebo 0   0.0000   0.0000   0.0000   0.0000   .   .
         age*treat   drug   1   -0.0046   0.0208 -0.0454   0.0362   0.05   0.8240
         age*treat   placebo 0   0.0000   0.0000   0.0000   0.0000   .   .
         Scale   0   1.0000   0.0000   1.0000   1.0000

NOTE: The scale parameter was held fixed.
```

```
/*disperson of parameter*/
/*check table "Criteria For Assessing Goodness Of Fit"
line "pearson chi sqaure*/

/*adjust model fitting with estimate of dispersion*/
PROC GENMOD data= dat2;
class treat(ref="placebo");
model number= age|treat / dist= poisson link= log
                          scale= PEARSON;
RUN;
```

The GENMOD Procedure

Model Information

Data Set	WORK.DAT2
Distribution	Poisson
Link Function	Log
Dependent Variable	number

Number of Observations Read   20
Number of Observations Used   20

Class Level Information

Class	Levels	Values
treat	2	drug placebo

Criteria For Assessing Goodness Of Fit

Criterion	DF	Value	Value/DF
Deviance	16	179.4906	11.2182
Scaled Deviance	16	15.7789	0.9862
Pearson Chi-Square	16	182.0059	11.3754
Scaled Pearson X2	16	16.0000	1.0000
Log Likelihood		100.9400	
Full Log Likelihood		-133.9133	
AIC (smaller is better)		275.8266	
AICC (smaller is better)		278.4933	
BIC (smaller is better)		279.8095	

Algorithm converged.

Analysis Of Maximum Likelihood Parameter Estimates

Parameter		DF	Estimate	Standard Error	Wald 95% Confidence Limits		Wald Chi-Square	Pr > ChiSq
Intercept		1	4.5191	0.5173	3.5053	5.5329	76.33	<.0001
age		1	-0.0384	0.0211	-0.0797	0.0029	3.32	0.0683
treat	drug	1	-1.2573	1.5907	-4.3749	1.8604	0.62	0.4293
treat	placebo	0	0.0000	0.0000	0.0000	0.0000	.	.
age*treat	drug	1	-0.0046	0.0702	-0.1423	0.1330	0.00	0.9474
age*treat	placebo	0	0.0000	0.0000	0.0000	0.0000	.	.
Scale		0	3.3727	0.0000	3.3727	3.3727		

NOTE: The scale parameter was estimated by the square root of Pearson's Chi-Square/DOF.

```
/*refit the model without interaction*/
PROC GENMOD data= dat2;
class treat(ref="placebo");
model number= age treat / dist= poisson link= log
                                scale= PEARSON;
RUN;
```

The GENMOD Procedure

Model Information

Data Set	WORK.DAT2
Distribution	Poisson
Link Function	Log
Dependent Variable	number

Number of Observations Read   20
Number of Observations Used   20

Class Level Information

Class	Levels	Values
treat	2	drug placebo

Criteria For Assessing Goodness Of Fit

Criterion	DF	Value	Value/DF
Deviance	17	179.5408	10.5612
Scaled Deviance	17	16.7360	0.9845
Pearson Chi-Square	17	182.3731	10.7278
Scaled Pearson X2	17	17.0000	1.0000
Log Likelihood		107.0304	
Full Log Likelihood		-133.9384	
AIC (smaller is better)		273.8768	
AICC (smaller is better)		275.3768	
BIC (smaller is better)		276.8640	

Algorithm converged.

Analysis Of Maximum Likelihood Parameter Estimates

Parameter		DF	Estimate	Standard Error	Wald 95% Confidence Limits		Wald Chi-Square	Pr > ChiSq
Intercept		1	4.5290	0.4811	3.5862	5.4719	88.64	<.0001
age		1	-0.0388	0.0195	-0.0771	-0.0006	3.96	0.0465
treat	drug	1	-1.3591	0.3853	-2.1143	-0.6039	12.44	0.0004
treat	placebo	0	0.0000	0.0000	0.0000	0.0000	.	.
Scale		0	3.2753	0.0000	3.2753	3.2753		

NOTE: The scale parameter was estimated by the square root of Pearson's Chi-Square/DOF.

```
/*disperson of parameter*/
/*check table "Criteria For Assessing Goodness Of Fit"
line "pearson chi sqaure*/

/*fit the quasi poisson*/
PROC GLIMMIX data= dat2;
class treat(ref="placebo");
model number= age treat / dist= poisson link= log solution;
random _residual_;
RUN;
```

The GLIMMIX Procedure

Model Information

Data Set	WORK.DAT2
Response Variable	number
Response Distribution	Poisson
Link Function	Log
Variance Function	Default
Variance Matrix	Diagonal
Estimation Technique	Maximum Likelihood
Degrees of Freedom Method	Residual

Class Level Information

Class	Levels	Values
treat	2	drug placebo

Number of Observations Read	20
Number of Observations Used	20

Dimensions

Covariance Parameters	1
Columns in X	4
Columns in Z	0
Subjects (Blocks in V)	1
Max Obs per Subject	20

Optimization Information

Optimization Technique	Newton-Raphson
Parameters in Optimization	3
Lower Boundaries	0
Upper Boundaries	0
Fixed Effects	Not Profiled

Iteration History

Iteration	Restarts	Evaluations	Objective Function	Change	Max Gradient
0	0	4	150.45214679	.	99.74975
1	0	3	134.54270883	15.90943796	12.75127
2	0	3	133.94007739	0.60263144	0.563253
3	0	3	133.93841261	0.00166478	0.001658
4	0	3	133.93841259	0.00000002	1.533E-8

Convergence criterion (GCONV=1E-8) satisfied.

Fit Statistics

-2 Log Likelihood	267.88
AIC (smaller is better)	273.88
AICC (smaller is better)	275.38
BIC (smaller is better)	276.86
CAIC (smaller is better)	279.86
HQIC (smaller is better)	274.46
Pearson Chi-Square	182.37
Pearson Chi-Square / DF	10.73

Parameter Estimates

Effect	treat	Estimate	Standard Error	DF	t Value	Pr > |t|
Intercept		4.5290	0.4811	17	9.41	<.0001
age		-0.03883	0.01951	17	-1.99	0.0628
treat	drug	-1.3591	0.3853	17	-3.53	0.0026
treat	placebo	0	.	.	.	.
Residual		10.7278	.	.	.	.

Type III Tests of Fixed Effects

Effect	Num DF	Den DF	F Value	Pr > F
age	1	17	3.96	0.0628
treat	1	17	12.44	0.0026

```
/*fit the negative-binomial model*/
PROC GENMOD data= dat2;
class treat(ref="placebo");
model number= age treat / dist= negbin;
RUN;

/*parameter estimate: k=1/dispersion: 1/0.5816= 1.7194*/
```

The GENMOD Procedure

Model Information

Data Set	WORK.DAT2
Distribution	Negative Binomial
Link Function	Log
Dependent Variable	number

Number of Observations Read	20
Number of Observations Used	20

Class Level Information

Class	Levels	Values
treat	2	drug placebo

Criteria For Assessing Goodness Of Fit

Criterion	DF	Value	Value/DF
Deviance	17	22.0022	1.2942
Scaled Deviance	17	22.0022	1.2942
Pearson Chi-Square	17	19.7505	1.1618
Scaled Pearson X2	17	19.7505	1.1618
Log Likelihood		1203.7028	
Full Log Likelihood		-78.4400	
AIC (smaller is better)		164.8800	
AICC (smaller is better)		167.5467	
BIC (smaller is better)		168.8629	

Algorithm converged.

Analysis Of Maximum Likelihood Parameter Estimates

Parameter		DF	Estimate	Standard Error	Wald 95% Confidence Limits		Wald Chi-Square	Pr > ChiSq
Intercept		1	4.5260	0.5918	3.3660	5.6860	58.48	<.0001
age		1	-0.0386	0.0207	-0.0791	0.0020	3.47	0.0625
treat	drug	1	-1.3681	0.3701	-2.0936	-0.6427	13.66	0.0002
treat	placebo	0	0.0000	0.0000	0.0000	0.0000	.	.
Dispersion		1	0.5816	0.2052	0.2913	1.1612		

NOTE: The negative binomial dispersion parameter was estimated by maximum likelihood.