

## 1. Fitting a Logistic Regression Model

- a. Submit the program **pmlr00d01.sas**. Create a macro variable to store  $\pi_1$ , the proportion of responders in the population. This value is 0.05. Then fit a logistic regression model using the **pva** data set with **TARGET\_B** as the target variable and **PEP\_STAR**, **RECENT\_AVG\_GIFT\_AMT**, and **FREQUENCY\_STATUS\_97NK** as the input variables. Model the probability that the target variable equals 1 and request 95% profile likelihood confidence intervals for the odds ratio. Use the SCORE statement to append the predicted probability to the data, correcting for oversampling. Create effect plots with confidence bands for the three input variables and create an odds ratio plot with a horizontal orientation and display the statistics.
- 1) Interpret the c statistic.
  - 2) Interpret the odds ratio for **PEP\_STAR**.
  - 3) Interpret the effect plot for **RECENT\_AVG\_GIFT\_AMT**.

```
%let ex_pi1=0.05;

proc logistic data=pva plots(only)=(effect(clband x=(pep_star
recent_avg_gift_amt frequency_status_97nk))
oddsratio (type=horizontalstat));
model target_b(event='1')=pep_star recent_avg_gift_amt
frequency_status_97nk / clodds=pl;
score data=pva out=scopva priorevent=&ex_pi1;
run;
```

# SAS 시스템

## The LOGISTIC Procedure

Model Information	
Data Set	WORK.PVA
Response Variable	TARGET_B
Number of Response Levels	2
Model	binary logit
Optimization Technique	Fisher's scoring

Number of Observations Read	19372
Number of Observations Used	19372

Response Profile		
Ordered Value	TARGET_B	Total Frequency
1	0	14529
2	1	4843

Probability modeled is TARGET\_B=1.

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

Model Fit Statistics		
Criterion	Intercept Only	Intercept and Covariates
AIC	21789.113	21351.668
SC	21796.984	21383.154
-2 Log L	21787.113	21343.668

Testing Global Null Hypothesis: BETA=0				
Test	Chi-Square	DF	Pr > ChiSq	
Likelihood Ratio	443.4451	3	<.0001	
Score	448.9127	3	<.0001	
Wald	439.1624	3	<.0001	

Analysis of Maximum Likelihood Estimates					
Parameter	DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept	1	-1.6241	0.0584	773.5119	<.0001
PEP_STAR	1	0.3078	0.0361	72.8485	<.0001
RECENT_AVG_GIFT_AMT	1	-0.00555	0.00204	7.3829	0.0066
FREQUENCY_STATUS_97N	1	0.2147	0.0168	163.3538	<.0001

# SAS 시스템

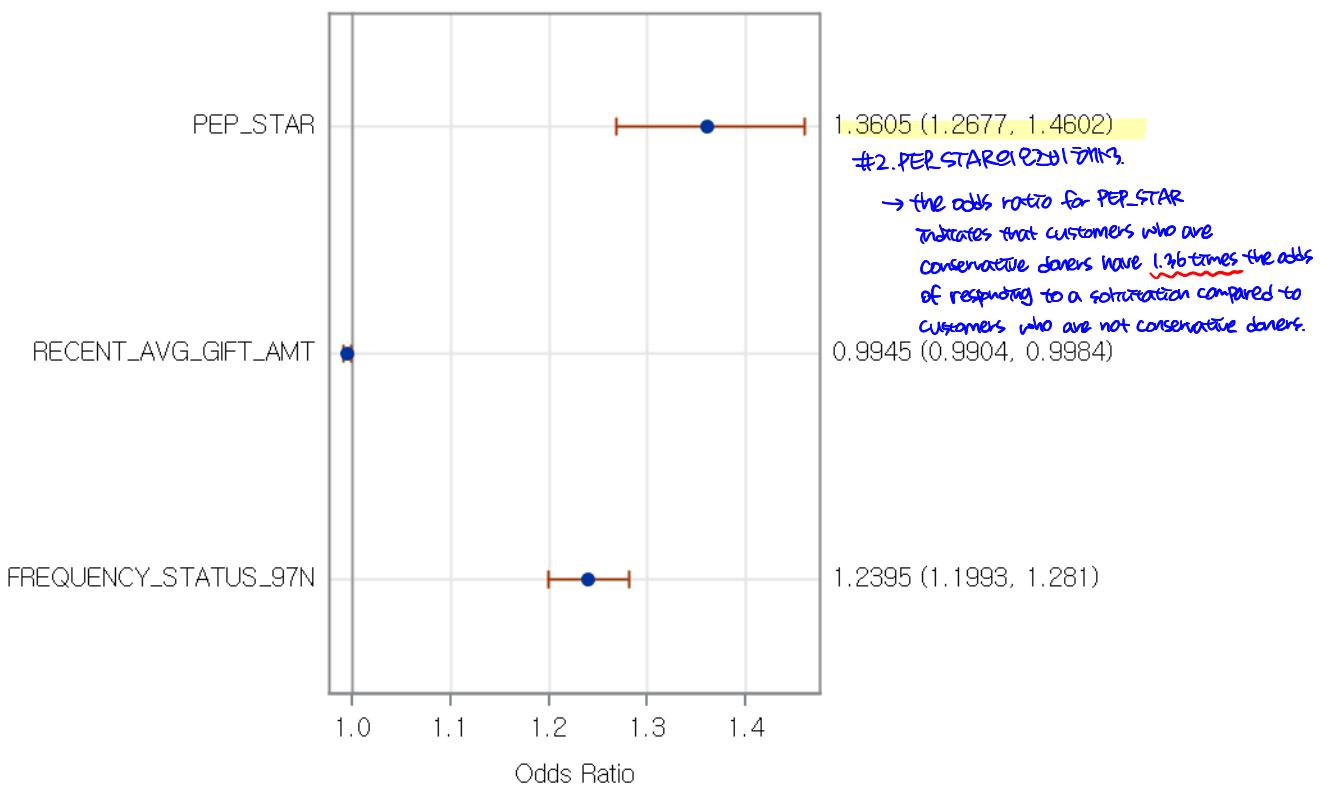
## The LOGISTIC Procedure

Association of Predicted Probabilities and Observed Responses			
Percent Concordant	59.5	Somers' D	0.201
Percent Discordant	39.4	Gamma	0.203
Percent Tied	1.1	Tau-a	0.075
Pairs	70363947	c	0.601

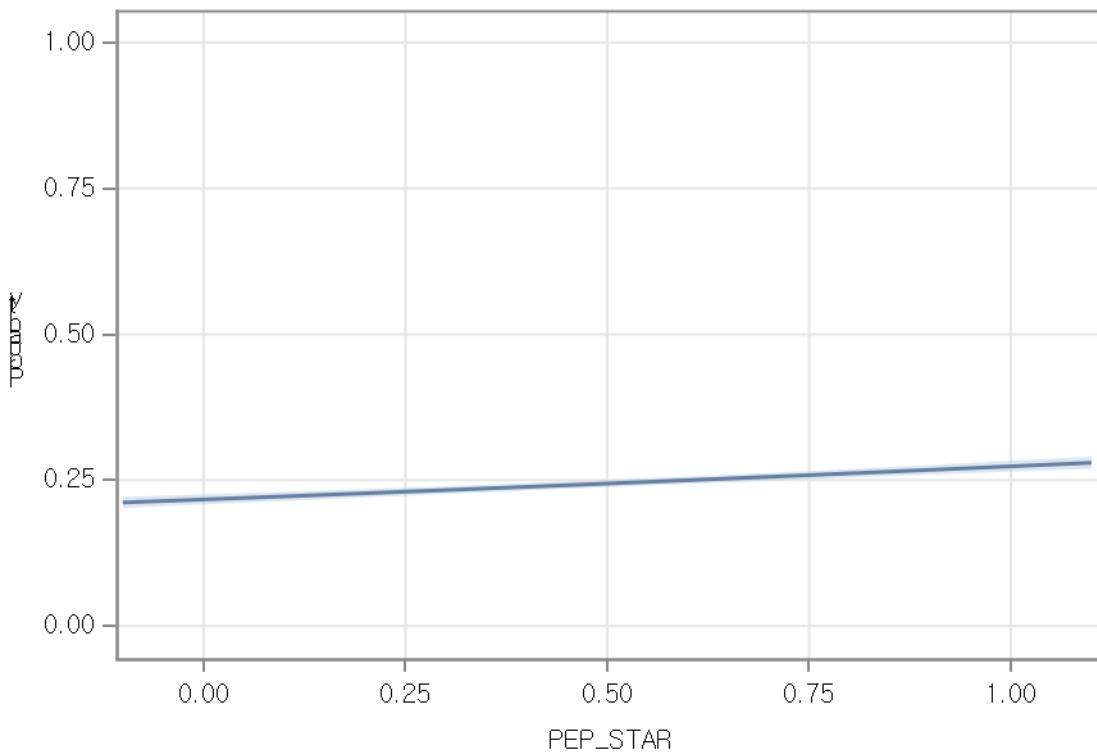
#1. C 등급입니다  
→ National veterans' organization에  
가장 많은 고객이 그룹에 있는 고객이 더 많음  
예측 확률은 같습니다.

Odds Ratio Estimates and Profile-Likelihood Confidence Intervals				
Effect	Unit	Estimate	95% Confidence Limits	
PEP_STAR	1.0000	1.360	1.268	1.460
RECENT_AVG_GIFT_AMT	1.0000	0.994	0.990	0.998
FREQUENCY_STATUS_97N	1.0000	1.239	1.199	1.281

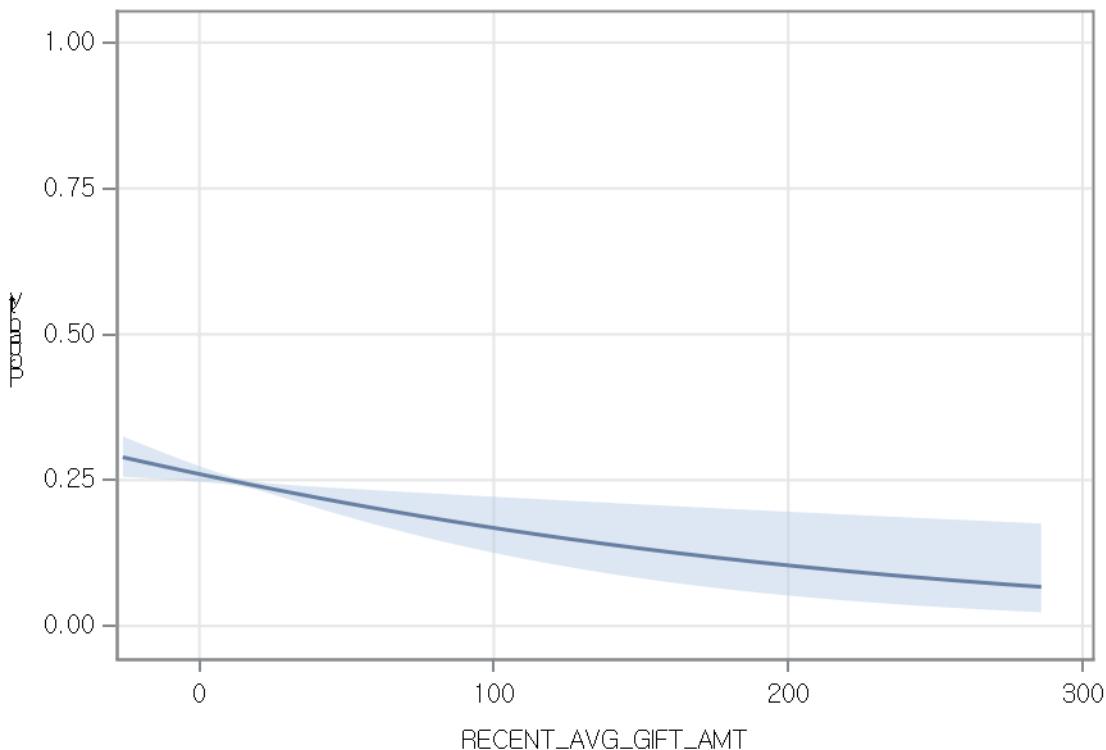
### Odds Ratios with 95% Profile-Likelihood Confidence Limits



**Predicted Probabilities for TARGET\_B=1 with 95% Confidence Limits**  
At RECENT\_AVG\_GIFT\_AMT=15.37 FREQUENCY\_STATUS\_97NK=1.984



**Predicted Probabilities for TARGET\_B=1 with 95% Confidence Limits**  
At PEP\_STAR=0.504 FREQUENCY\_STATUS\_97NK=1.984



#3. RECENT\_AVG\_GIFT\_AMT 단계

→ the effect plot for RECENT\_AVG\_GIFT\_AMT shows a negative relationship between the average donation amount to promotions since June (1994) and the predicted probabilities of responding to the solicitation in June of (PP).

**Predicted Probabilities for TARGET\_B=1 with 95% Confidence Limits**

At PEP\_STAR=0.504 RECENT\_AVG\_GIFT\_AMT=15.37

