International Journal of Social Science, Education and Humanities
ISSN Print: 3104-4239
ISSN Online: 3104-4247
CD ISSN: 3104-4255
https://doi.org/10.63313/IJSSEH.2001

Understanding Counseling Attrition: Insights from Logistic Regression

Brendon Lee Gore*. Hu Yue
School of Science, Zhejiang University of Science and Technology, Hangzhou, China
Email: gorebrendolee@gmail.com

How to cite this paper: Gore, B. L., & Yue, H. (2025). Understanding Counseling Attrition: Insights from Logistic Regression. International Journal of Social Science, Education and Humanities, 1(2), 1-13. ISSN Print: 3104-4239, ISSN Online: 3104-4247.
https://doi.org/10.63313/IJSSEH.2001
Published: 2025-09-25

Abstract

This study investigates the predictors of early termination among clients receiving counseling services at a community mental health centre. Early termi-nation, or dropout, poses significant challenges to the efficacy of therapeutic interventions, often leading to incomplete treatment and unaddressed mental health needs. Understanding the factors that contribute to this phenomenon is crucial for improving client retention and optimizing therapeutic outcomes. To explore these factors, a sample of clients who had engaged in counseling at the mental health centre was selected for the study. Each participant was inter-viewed, and their responses were meticulously recorded to ensure the accuracy and integrity of the data. The interview process was designed to encourage hon-est disclosure, enabling the collection of reliable information regarding their experiences and reasons for either continuing or discontinuing therapy. The study employed a binary logistic regression model to analyse the likelihood of early termination among these clients. The dependent variable in the model was the occurrence of early termination, operationalized as a binary outcome where '1' indicated early termination and '0' signified continuation of counseling until the intended completion. The group that did not terminate early was used as the reference category, while those who terminated early represented the target group for analysis. Two independent variables were identified as potential pre-dictors of early termination: symptom severity and avoidance of disclosure. Symptom severity ('sympsev') refers to the intensity of the clients' psychologi-cal symptoms, which was hypothesized to influence their likelihood of remain-ing in therapy. Avoidance of disclosure ('avdiscl') measures the extent to which clients were reluctant to share personal information during counseling sessions, which could hinder the therapeutic process and contribute to premature drop-out. Both predictors were treated as continuous variables within the model. The logistic regression analysis aimed to determine the relative contribution of these predictors to the likelihood of early termination. By calculating the odds ratios for each predictor, the study assessed how variations in symptom severity and avoidance of disclosure influenced the probability of a client terminating therapy prematurely. The results of this analysis provide valuable insights into the dynamics of client retention in counseling settings

Keywords

Early Termination; Symptom Severity; odds ratio; Avoidance of Disclosure; The Hosmer & Lemeshow

1. Introduction

Binary logistic regression analysis is a statistical technique used to model the relationship between a binary dependent variable (also called the outcome or response variable) and one or more independent variables (also known as predictors or covariates). This type of regression is used to determine the likelihood of an event happening in various fields, including healthcare, social science, and business. According to [1], binary logistic regression can be used to estimate the likelihood of a binary outcome variable, such as the likelihood of a patient having a certain disease, based on one or several independent variables, such as age, gender, and health status. The model is called 'logistic' because it uses the logarithm of the odds of the event occurring, instead of the probability itself. The logistic regression model assumes a linear relationship between the independent variables and the logarithm of the odds of the outcome variable [2]. Model estimates coefficients for independent variables in order to find out the magnitude and direction of the effect of these variables on the outcome variable.

The coefficients of independent variables can now be used to predict the likelihood of an outcome variable, given certain values of the independent variables. One of the advantages of binary logistic regression is having ability to handle both categorical and continuous independent variables [3]. Fundamentally, the researcher is addressing the question, "What is the probability that a given case falls into one of two categories on the dependent variable, given the predictors in the model?". As one might be inclined to ask why we don't use standard ordinary least squares regression (OLS) instead of BLR, OLS regression assumes (a) a linear relationship between the independent variables and the dependent variable, (b) the residuals are normally distributed, and (c) the residuals exhibit constant variance (Pituch & Stevens, 2016) [4]. All three assumptions are violated if the outcome variable in an OLS model is binary. And pivoting off (a), the relationship between one or more predictors and the probability of a target outcome is inherently non-linear as probabilities are bounded at 0 and 1. When modeling a binary outcome using OLS regression, the estimation of model parameters ignores this boundedness, which has the notable effect of producing predicted probabilities that fall outside the 0-1 range. BLR estimates regression parameters by considering the fact that probabilities are bounded and 0 and 1. It also does not assume that residuals are normally distributed and exhibit constant variance.

2. Literature Review

2.1. Model Estimation

Unlike OLS regression, BLR uses maximum likelihood (ML) to estimate model parameters. Maximum likelihood estimation is an iterative process aimed at arriving at population (parameter) values that most likely produced the observed (sample) data. Generally, this estimation approach assumes large samples and, aside from issues of power, smaller sample sizes can create problems with model convergence and estimation of model parameters. [Side note: With smaller samples, Exact logistic regression or Firth procedure using Penalized Maximum likelihood can be used. Unfortunately, these options are not commonly available in statistics programs.] From our sample, the model will be predicting given the independent factors of avoidance of disclosure and symptom severity and be able to display the table outlining the odds ratio pertaining to each predictor and for us to be able to draw the conclusion.

2.2. Evaluation of Model fit

Evaluation of model fit in binary logistic regression can be done using various measures such as goodness-of-fit tests, pseudo R-squared values, and receiver operating characteristic (ROC) curves. The Goodness-of-fit tests: tests assess the overall fit of the model by comparing the observed frequencies with the expected frequencies based on the model. With the Hosmer-Lemeshow test [5] being the most commonly used goodness-of-fit test, which divides the data into groups based on predicted probabilities and compares the observed and expected frequencies within each group. A significant result indicates poor model fit. Pseudo R-squared values: these are measures of how well the model explains the variability in the data. The most commonly used pseudo R-squared values are Nagelkerke's R-squared [6], Cox and Snell R-squared [7], and McFadden's R-squared [8]. These values range from 0 to 1, with higher values indicating better model fit. ROC curves: these plots show the trade-off between sensitivity (true positive rate) and specificity (true negative rate) for different classification thresholds.

The area under the curve (AUC) [9] is a commonly used measure of model fit, with values ranging from 0.5 (random guessing) to 1 (perfect classification). With that being said, it is important to note that none of these measures alone can fully capture the performance of a logistic regression model, and a combination of these measures should be used to assess the overall fit of the model.

3. Methodology

This study involved 45 participants who provided an honest assessment of their likelihood to terminate their counseling contract early. Termination, in this context, refers to the cancellation of a client's counseling membership at a Mental Health Center and serves as the study's dependent variable, which is binary, encompassing two possible outcomes: early termination or continuation of counseling. The study examined independent variables such as avoidance of disclosure and symptom se-

verity. Avoidance of disclosure (avdiscl) in therapy occurs when a client hesitates or refuses to share certain experiences or information with their therapist, often due to factors like fear of rejection or judgment, feelings of shame or guilt, or a lack of trust in the therapist. Symptom severity (sympsev) refers to the extent to which a particular symptom or set of symptoms affects a person's quality of life and daily functioning. In mental health treatment, tracking the severity of symptoms is crucial for assessing the progress and effectiveness of the therapeutic intervention.

3.1. The Model

When we have a binary outcome, our desire is to predict the probability (likelihood) of membership in a target outcome Prob (Y=1; terminate early), given what we know from a set of predictors. Unfortunately, we cannot model this relationship directly using standard OLS regression, since the relationship between the predictors and probability that Y=1 is inherently non-linear (following an S-shaped curve; logistic curve). In a logistic regression, it addresses this problem by "linearizing" the relation using a logit link function. The dependent variable that we are directly predicting, therefore, are logits (a mathematical transformation of probabilities). This changes our interpretation of intercepts and slopes in our regression model since the we are speaking in terms of logits rather than probabilities. The logistic regression prediction equation can be expressed as:

$$\begin{aligned}\text{logit(p)} &= \ln\left(\frac{\text{P}}{1-\text{P}}\right) \\ &= \mathbf{B_0} + \boldsymbol{B_1} avdiscl + \boldsymbol{B_2} sympsev \end{aligned} \quad (1)$$

$$\pi(x_i) = \frac{e^{\mathbf{B_0}+\boldsymbol{B_1} avdiscl+\boldsymbol{B_2} sympsev}}{1+e^{\mathbf{B_0}+\boldsymbol{B_1} avdiscl+\boldsymbol{B_2} sympsev}} \quad (2)$$

Where logit(p) is our dependent variable terminate, B_0 is a constant term and I'ts calculated value from variables in the equation table is **-1.139,** B_1 is the coefficient of the predictor ***avdiscl*** and its calculated value from variables in the equation table is **.356**, B_2 is the coefficient of the predictor ***sympsev*** and its calculated value from variables in the equation table is **-.317**.

3.2. Assumption Checking

Before conducting a binary logistic regression analysis, it is essential to ensure that certain assumptions are met to guarantee the validity and reliability of the model. If all these assumptions are satisfied, we can proceed with analysing the data using the binary logistic regression model. These assumptions include:

a) *The dependent/response variable is binary or* dichotomous.

Dependent Variable Encoding

Original Value	Internal Value
.00	0
1.00	1

Figure 1. Dependent Variable Encoding

Logistic regression assumes that the response variable only takes on two possible outcomes. From Figure 1, we can clearly see that we have two outcomes coded 0 and 1 for 'do not terminate early' and 'terminate early' respectfully as our only two outcomes, therefore our model meets this assumption.

b) The Observations are Independent

This assumption states that the dataset observations should be independent of each other. The observations should not be related to each other or emerge from repeated measurements of the same individual type. In our data, all our observations of each client are independent from each other.

c) Little or no multicollinearity between the predictor/explanatory variables [10].

Coefficients[a]

Model		Collinearity Statistics	
		Tolerance	VIF
1	avdiscl	.960	1.042
	sympsev	.960	1.042

a. Dependent Variable: terminate

Figure 2. Collinearity Statistics

In In binary logistic regression, multicollinearity refers to the presence of high correlations among the predictor or explanatory variables. The assumption of "little or no multicollinearity" implies that the variables included in the regression model are not excessively correlated with one another. As shown in Figure 2, the Collinearity Statistics indicate that all tolerance values exceed 0.1, which suggests that this assumption is not violated. Additionally, the Variance Inflation Factor (VIF) values for all predictors are below 10, further confirming the absence of multicollinearity in our dataset.

Multicollinearity can introduce significant issues in binary logistic regression, such as unstable or unreliable estimates of regression coefficients, inflated standard errors, and challenges in interpreting the coefficients. Therefore, it is crucial to check for multicollinearity before fitting the binary logistic regression model to ensure the accuracy and interpretability of the results.

d) The sample size is sufficiently large (at least 10-20 observations per independent

variable) [11].

Logistic regression assumes that the dataset's sample size is sufficiently large to draw valid conclusions from the fitted model. A common guideline is to have at least 10 cases for the least frequent outcome for each explanatory variable. As illustrated in Figure 3, the dataset comprises a total sample size of 45 clients, which is considered adequate for ensuring the robustness of the logistic regression analysis.

e) There are no outliers

Logistic regression assumes the absence of extreme outliers or influential observations within the dataset. To assess this, we used Mahalanobis distance to identify potential outliers. The results indicate that no outliers are present, as the minimum Mahalanobis distance value exceeds 0.001, confirming that the assumption of no extreme outliers has been satisfied. Having verified that our data meets all necessary assumptions for binary logistic regression, we are now prepared to proceed with the analysis.

Case Processing Summary

Unweighted Cases[a]		N	Percent
Selected Cases	Included in Analysis	45	100.0
	Missing Cases	0	.0
	Total	45	100.0
Unselected Cases		0	.0
Total		45	100.0

a. If weight is in effect, see classification table for the total number of cases.

Figure 3. Case Processing Summary

4. Analysis and Result Discussion

Case Processing Summary

Unweighted Cases[a]		N	Percent
Selected Cases	Included in Analysis	45	100.0
	Missing Cases	0	.0
	Total	45	100.0
Unselected Cases		0	.0
Total		45	100.0

a. If weight is in effect, see classification table for the total number of cases.

Figure 4. Case Processing Summary

The first section of the output shows Case Processing Summary Highlighting the cases included in the analysis. In this study we have a total of 45 respondents as shown in Figure 4.

Dependent Variable Encoding

Original Value	Internal Value
.00	0
1.00	1

Figure 5. Dependent Variable Encoding

The Dependent variable encoding shows the coding for the criterion variable. In this case those who will choose to not terminate early(0.00) are classified as 0 whereas those that choose to terminate early(1.00) are classified as 1.(see Figure 5)

Block 0: Beginning Block

Classification Table[a,b]

			Predicted		
			terminate		Percentage Correct
	Observed		.00	1.00	
Step 0	terminate	.00	25	0	100.0
		1.00	20	0	.0
	Overall Percentage				55.6

a. Constant is included in the model.

b. The cut value is .500

Variables in the Equation

		B	S.E.	Wald	df	Sig.	Exp(B)
Step 0	Constant	-.223	.300	.553	1	.457	.800

Figure 6. Block 0: Beginning Block

The next section of the output, headed Block 0, as shown on Figure 6, is the result of the analysis without any of our independent variables used in the model. Therefore, this will serve as a baseline later for comparing the model with our predictor variable included. he output is presented in blocks. Block 0 contains the results from a null (i.e., intercept-only) model. The next portion of the output is headed as Block 1. This part of the output is our main focus when interpreting results, as it is based on our regression model including our predictor(s).

The Omnibus Tests of Model Coefficients (refer to Figure 8) present the results from the likelihood ratio chi-square tests, which assess whether a model that includes the full set of predictors significantly improves the fit compared to the null (inter-

cept-only) model. Essentially, this test serves as an omnibus evaluation of the null hypothesis, positing that the regression slopes for all predictors in the model are equal to zero (Pituch & Stevens, 2016).4]. The results shown here indicate that the model fits the data significantly better than a null model, $\chi^2(2)=26.174$, $p<0.001$.

Omnibus Tests of Model Coefficients

		Chi-square	df	Sig.
Step 1	Step	26.174	2	.000
	Block	26.174	2	.000
	Model	26.174	2	.000

Figure 7. Omnibus Tests of Model Coefficients

The Omnibus Tests of Model Coefficients (refer to Figure 8) present the results from the likelihood ratio chi-square tests, which assess whether a model that includes the full set of predictors significantly improves the fit compared to the null (intercept-only) model. Essentially, this test serves as an omnibus evaluation of the null hypothesis, positing that the regression slopes for all predictors in the model are equal to zero (Pituch & Stevens, 2016).4]. The results shown here indicate that the model fits the data significantly better than a null model, $\chi^2(2)=26.174$, $p<0.001$.

Model Summary

Step	-2 Log likelihood	Cox & Snell R Square	Nagelkerke R Square
1	35.653[a]	.441	.590

a. Estimation terminated at iteration number 6 because parameter estimates changed by less than .001.

Figure 8. Model Summary

The Model Summary (see Figure 8) presents the -2 Log Likelihood (-2LL) and two "pseudo-R-square" measures. These pseudo-R-squares serve as descriptive indices, providing insight into the overall fit of the model, though they are not calculated in the same manner as the R-square in OLS regression. The -2 Log Likelihood, also known as model deviance, measures the discrepancy between the model and the observed data; values closer to zero indicate a better fit, while higher values suggest a poorer fit. The likelihood ratio (LR) chi-square value, also shown in Figure 8, repre-

sents the difference in deviances (-2LL) between the full model, which includes all predictors, and the reduced model, which contains only the intercept. This difference helps assess whether the inclusion of the predictors significantly improves the model fit. The deviance for the intercept-only model can be computed using the LR chi-square and the deviance of the full model.

$$\boldsymbol{Deviance(\ null\ model)} = \mathbf{35.653} + \mathbf{26.174} = \mathbf{61.827} \quad (3)$$

The Cox & Snell and Nagelkerke R-squares are considered 'pseudo-R-square' values because they are not calculated in the same manner as the R-square in Ordinary Least Squares (OLS) regression. In OLS regression, the R-square is interpreted as the proportion of variation in the dependent variable that is explained by the predictors. However, in the context of logistic regression, these pseudo-R-square values are intended to reflect the proportionate improvement in model fit relative to the intercept-only model. They offer a way to gauge how much the inclusion of predictors has enhanced the model's explanatory power.

Below is the computation of Cox & Snell pseudo-R square (see Figure 9). Unlike R-square in OLS regression, the upper bound is less than 1.

Model Summary

Step	-2 Log likelihood	Cox & Snell R Square	Nagelkerke R Square
1	35.653[a]	.441	.590

a. Estimation terminated at iteration number 6 because parameter estimates changed by less than .001.

Figure 9. Model Summary on Cox & Snell R Square Calculation

$$\boldsymbol{CS} = \mathbf{1} - \boldsymbol{exp}\left(\frac{\boldsymbol{deviance_{full}} - \boldsymbol{deviance_{null}}}{\boldsymbol{n}}\right)$$

$$= \mathbf{1} - \boldsymbol{exp}\left(\frac{\mathbf{35.653} - \mathbf{61.827}}{\mathbf{45}}\right) = \mathbf{1} - \boldsymbol{e}^{-.5816} = .\mathbf{441} \quad (4)$$

* The 'exp' in the first two expressions above are equivalents to the 'e' in third expression. This was done to make the equation more readable. Nagelkerke provided an adjustment to Cox & Snell providing an index ranging from 0 to 1. Below is the computation of the Nagelkerke pseudo-R-square.

Model Summary

Step	-2 Log likelihood	Cox & Snell R Square	Nagelkerke R Square
1	35.653[a]	.441	.590

a. Estimation terminated at iteration number 6 because parameter estimates changed by less than .001.

Figure 10. Model Summary on Nagelkerke pseudo-R-square

$$Nagelkerke = \frac{CS}{1 - exp\left(-\frac{deviance_{null}}{n}\right)}$$

$$= \frac{.441}{1 - exp\left(-\frac{61.827}{45}\right)} = .590 \qquad (5)$$

* These versions of the Cox & Snell and Nagelkerke pseudo-R-squares are provided by Field (2018) [12]. The -2*Log likelihood (also referred to as "model deviance" [4] is most useful for comparing competing models, particularly because it is distributed as chi-square [12].

Hosmer and Lemeshow Test

Step	Chi-square	df	Sig.
1	2.024	7	.958

Figure 11. Hosmer and Lemeshow Test

The Hosmer & Lemeshow test is another test that can be used to evaluate global fit. A non-significant test result (see Figure 11; p=.958) is an indicator of good model fit.

Classification Table[a]

			Predicted		
			terminate		Percentage Correct
	Observed		.00	1.00	
Step 1	terminate	.00	21	4	84.0
		1.00	4	16	80.0
	Overall Percentage				82.2

a. The cut value is .500

Figure 12. Classification Table

The classification table provides the frequencies and percentages reflecting the degree to which the model correctly and incorrectly predicts category membership on the dependent variable. On Figure 12, we see that 100%×(21/(21+4))=84% of cases that were observed not terminate early were correctly predicted (by the model) to not terminate early. Of the 20 cases observed to terminate early, 100%×(16/(4+16))=80% were correctly predicted by the model to terminate early. [As you can see, sensitivity refers to accuracy of the model in predicting target group membership, whereas specificity refers to the accuracy of a model to predict non-target group membership.] The overall classification accuracy based on the model was 100%×((21+16)/45)=82%. The 'Estimate' column in the output contains the regression coefficients for each predictor. These coefficients represent the predicted change in the log odds of belonging to the target group (as opposed to the reference group on the dependent variable) for each one-unit increase in the predictor, while controlling for the other predictors in the model. It's important to note that a common misconception is that the regression coefficient indicates the predicted change in the probability of target group membership per unit increase in the predictor. This is incorrect.

Variables in the Equation

								95% C.I.for EXP(B)	
		B	S.E.	Wald	df	Sig.	Exp(B)	Lower	Upper
Step 1[a]	sympsev	-.317	.123	6.582	1	.010	.729	.572	.928
	avdiscl	.356	.118	9.148	1	.002	1.427	1.133	1.798
	Constant	-1.139	1.500	.577	1	.448	.320		

a. Variable(s) entered on step 1: sympsev, avdiscl.

Figure 13. Variables in the Equation

The coefficient actually reflects the predicted change in log odds, not the probability, for each unit increase in the predictor. You can generally interpret a positive regression coefficient as indicating that, loosely speaking, the probability of falling into the target group increases as the predictor variable increases. Conversely, a negative coefficient suggests that the probability of belonging to the target group decreases with an increase in the predictor. If the regression coefficient equals 0, it indicates that changes in the predictor have no effect on the probability of being in the target group. The Odds Ratio (OR) column presents values that reflect the multiplicative change in odds for each one-unit increase in a predictor. Typically, an odds ratio greater than 1 indicates that as the predictor increases, there is a higher probability of the case falling into the target group on the dependent variable. An odds ratio less than 1 suggests a decreasing probability of being in the target group as the predictor increases. If the OR equals 1, this signifies that changes in the predictor do not affect the proba-

bility of being in the target group. The 95% confidence interval for the Odds ratio can also be used to test the observed OR to determine if it is significantly different from the null OR of 1.0. If 1.0 falls between the lower and upper bound for a given interval, then the computed odds ratio is not significantly different from 1.0 (indicating no change as a function of the predictor). On diagram (see Figure 14), Avoidance of disclosure(avdiscl) is a positive and significant (b=.356, s.e.=.118, p=.002) predictor of the probability of early termination, with the Odds Ratio indicating that for every one unit increase on this predictor the odds of early termination change by a factor of 1.427 (meaning the odds are increasing).

Variables in the Equation

								95% C.I.for EXP(B)	
		B	S.E.	Wald	df	Sig.	Exp(B)	Lower	Upper
Step 1[a]	sympsev	-.317	.123	6.582	1	.010	.729	.572	.928
	avdiscl	.356	.118	9.148	1	.002	1.427	1.133	1.798
	Constant	-1.139	1.500	.577	1	.448	.320		

a. Variable(s) entered on step 1: sympsev, avdiscl.

Figure 14. Variables in the Equation Interpretation

Symptom severity(sympsev) is a negative and significant (b=-.317, s.e.=.123, p=.010) predictor of the probability of early termination. The Odds Ratio indicates that for every one-unit increment on the predictor, the odds of terminating increase by a factor of .729 (meaning that the odds are decreasing).

5. Conclusion

Binary Logistic Regression was used to test the likelihood of early termination from counseling from a sample of clients from a mental health centre. A preliminary analysis suggested that the assumption of multicollinearity was met (tolerance=.96). The model was statistically significant, x^2 (2, N=45) =26.174, p<.001 indicating that the model fits the data significantly better than a null model and can model the likelihood of early termination correctly.

	B	**SE**	**Wald**	**df**	**p**	**OR**	**CI for OR**	
							LL	UL
avdiscl	0.36	0.12	9.15	1	0.002	1.43	1.13	1.80
sympsev	-0.32	0.12	6.58	1	0.010	0.73	0.57	0.93
Constant	**-1.14**	**1.50**	**0.58**	**1**	**0.448**	**0.32**		

The model explained between 44.1% (Cox and Snell R square) and 59% Nagelkerke R square) of the variance in the dependent variable and correctly classified 82.2% of the cases. As shown in the Table 1 below, symptom severity and avoidance of disclosure both contributes to the model. As summarized in Figure 15, the regression analysis indicates that avoidance of disclosure (avdiscl) is both positive and significant. Specifically, the odds of a client terminating early due to avoidance of disclosure are

1.43 times higher than the odds of not terminating early, with a 95% confidence interval (CI) ranging from 1.13 to 1.8. On the other hand, symptom severity (sympsev) is negative and significant, with the odds of early termination due to symptom severity being 0.73 times lower than the odds of not terminating early, accompanied by a 95% CI of 0.57 to 0.93.

References

[1] Hosmer Jr, D. W., Lemeshow, S., & Sturdivant, R. X. (2013). Applied logistic regression. John Wiley & Sons.

[2] Allison, P. D. (2012). Logistic regression using SAS: theory and application. SAS Institute.

[3] Menard, S. (2002). Applied logistic regression analysis (Vol. 106). Sage.

[4] Pituch, K.A., & Stevens, J.A. (2016). Applied multivariate statistics for the social sciences (6th ed). New York: Routledge.

[5] Hosmer, D.W., Hosmer, T. and Lemeshow, S. (1980) A Goodness-of-Fit Tests for the Multi-ple Logistic Regression Model. Communications in Statistics, 10, 1043-1069.

https://doi.org/10.1080/03610928008827941

[6] Nagelkerke, N.J.D. (1991) A Note on a General Definition of the Coefficient of Determination. Biometrika, 78, 691-692.

https://doi.org/10.1093/biomet/78.3.691

[7] Cox, D. R., & Snell, E. J. (1989). Analysis of binary data (2nd ed.). Chapman and Hall.

[8] McFadden, D. (1974) Conditional Logit Analysis of Qualitative Choice Behavior. In: Za-rembka, P., Ed., Economic Theory and Mathematical Economics, Academic Press, New York, NY, 105-142.

[9] Fawcett, T. (2006). An introduction to ROC analysis. Pattern Recognition Letters, 27(8), 861-874. https://doi.org/10.1016/j.patrec.2005.10.010

[10] Hair Jr, J. F., Black, W. C., Babin, B. J., & Anderson, R. E. (2014). Multivariate data analysis (7th ed.). Pearson Education Limited.

[11] Peduzzi, P., Concato, J., Kemper, E., Holford, T. R., & Feinstein, A. R. (1996). A simulation study of the number of events per variable in logistic regression analysis. Journal of clinical epidemiology, 49(12), 1373–1379. https://doi.org/10.1016/s0895-4356(96)00236-3

[12] Field, A. (2018). Discovering statistics using IBM SPSS statistics (5th ed). Los Angeles: Sage.