Research Article

Open Access

Optimal threshold for static 99R

Abstract

The Static 99R is an actuarial instrument that is widely used to assess the sexual recidivism risk of sex offenders. It is frequently applied in jurisdictions as a decision-making tool for release or indefinite admission to a psychiatric hospital within the jail of sex offenders. The decision to release or retain a criminal depends solely on the total score which is considered as the only independent variable. In our study, two models of Static 99R are considered: the 5-year high risk model and the 10-year high risk model. To identify the most appropriate threshold, we performed four independent methods. These are: the point closest-to-(0,1), the concordance probability (CZ), the index of union (IU), and the plot of sensitivity versus specificity. Remarkably, all four methods yielded identical results. For the 5-year high risk model, the optimal threshold is 0.184, which corresponds to a cut-off score of 5. Consequently, a score of 5 or higher implies that the offender is very likely to recidivate. Similarly, for the 10-year high risk, the optimal threshold is 0.293 which corresponds also to a cut-off score of 5.

Keywords: logistic regression, Static 99R, recidivism rate, sex offender, risk assessment

YingYin Chang,[1] Kamel Rekab,[1] Majid Bani-Yaghoub,[1] Matthew Mueller[2]
[1]Division of Computing, Analytics and Mathematics, University of Missouri-Kansas City, USA
[2]MGM Law LLC, USA

Correspondence: YingYin Chang, Division of Computing, Analytics and Mathematics, University of Missouri-Kansas City, USA, Email ycx59@mail.umkc.edu

Abbreviations: CZ, concordance probability; IU, the index of union; ROC, receiver operating characteristic; AUC, area under the curve

Introduction

The Static 99R (www.saarna.org)[1–5] has been used in court proceedings as the primary actuarial instrument to predict the risk of sexual recidivism of sex offenders. It consists of 10 static variables that are derived from various factors related to demographic information, criminal history, and victim information. Adult male sex offenders assessed through Static 99R receive scores ranging from -3 to 11 which are subsequently categorized into five risk level: very low, below average, average, above average, and well-above average. Static 99R used the total score of sexual offenders as the only independent variable. In the context of this study, our primary goal of the present work was to define an optimal threshold (cut-off score), employing four independent methodologies outlined in subsequent sections. It is worth noting that Stat 999R does not differentiate between individual sex offenders who have the same total score.

Materials and methods

Sexual recidivism data

We obtained the Static 99R total scores from the Static-99R coding rules.[6,7] A summary of the observed data for 5-year and 10-year high risk sexual recidivism rates can be found in appendices A and B, respectively. For 5-year high, a sample size of 860 was used, resulting in 164 recidivists; for 10-year high, a sample size of 350 was used, resulting in 98 recidivists. These summarized data were used to replicate the original data. To illustrate within the 5-year high data (Appendix A) there were 21 sex offenders with the total score of -1, one of whom recidivated. Utilizing this information, we generated a column consisting of 21 entries assigned the score of -1, while the second column featured all zeros except for one entry marked with a value of 1. By doing so, we replicated the entire dataset for both 5-year and 10-year high risk.

Simple binary logistic model

In a population of n sexual offenders assume that r individuals will recidivate and $n - r$ will not. Then the proportional response is $\pi = r / n$ and the odds are defined by $Odds = \pi / (1 - \pi)$. When the logistic regression model is fitted, estimates of π are denoted by $\hat{\pi}$. The logit transformation

$$\ln\left(\frac{\hat{\pi}}{1 - \hat{\pi}}\right) = b_0 + b_1 x \qquad (1)$$

where $\hat{\pi}$ is the expected proportional response, b_0 is intercept, b_1 is slope and x is the total Static 99R score of each sex offender.

Methods for finding the optimal threshold

a. The point closest-to-(0,1).[8,9] It is known as the distance between a point of the ROC curve and an ideal point (0,1) representing zero false positive and perfect sensitivity. The optimal threshold was determined by choosing the threshold that minimizes the distance between (FP, TP) and (0,1).

$$d = \sqrt{\left(FP - 0\right)^2 + \left(TP - 1\right)^2} \qquad (2)$$

b. Concordance Probability method (CZ). This method was proposed by Liu X,[10] which defines the optimal cut-point as the point maximizing the product of sensitivity and specificity.

$$CZ = Sen * Spe \qquad (3)$$

c. Index of Union (IU). It was proposed by Ilker Unal,[11] the cut-point which minimizes the IU(c) will be the "optimal "cut-point value.

$$IU(c) = \left(\left|Se(c) - AUC\right| + \left|Sp(c) - AUC\right|\right) \qquad (4)$$

The optimal cut-off found by this method meets two conditions: (1) sensitivity and specificity obtained at this cut-point should be close to AUC value; (2) the difference between sensitivity and specificity obtained at this cut-point should be minimum.

d. The Plot of sensitivity and specificity versus each possible cut-point. According to Hosmer D & Lemeshow,[12] "One might select a point that maximizes both sensitivity and specificity. The "optimal" choice for a cut-off point will be approximately where the sensitivity and specificity curve cross".

Results

Utilizing the data for 5-year high risk and 10-year high risk, we replicated the logistic models for Static 99R. Subsequently, we

conducted an ROC analysis on the constructed logistic model and generated a table containing the coordinates of the ROC Curve. The Tables 1 & 2 presents the (sensitivity) and (1–specificity) values of the ROC curve at various cut-off points, which are represented as the predicted probability. By performing four independent methods, we determined that the optimal threshold for 5-year high risk is 0.184 as shown in Table 3 & Figure 1. The optimal threshold of 0.184 corresponds to a cut-off score of 5. Similarly, for 10-year high risk we identified the optimal threshold is 0.293, as shown in Table 4 & Figure 2. The optimal threshold of 0.293 corresponds to a cut-off score of 5.

5-Year high risk logistic model and optimal threshold (Tables 1–3 & Figure 1)

Table 1 Variables in the 5-year logistic model

		B	S.E.	Wald	df	Sig.	Exp(B)
Step 1[a]	Score	0.23	0.041	31.721	1	0	1.258
	Constant	-2.527	0.226	125.052	1	0	0.08

[a]Variable(s) entered on step 1: score.

Table 2 Variables in the 10-year logistic model

		B	S.E.	Wald	df	Sig.	Exp(B)
Step [a]	score	0.233	0.06	15.128	1	0	1.262
	Constant	-1.929	0.293	43.253	1	0	0.145

[a]Variable(s) entered on step 1: score.

Table 3 Sensitivity, 1- Specificity, Distance (0,1), Sen*Spe, IU, |Sen-Sep|, PPV, NPV, and ACC at Stat-99R cut-points

Positive if greater than or equal to[a]	Sen	1 - Spe	Distance	Sen*Spe	IU	|TP-TN|	PPV	NPV	ACC
0.0000000	1.000	1.000	1.000	0.000	1.000	1.000	0.192	-	0.192
0.0668520	0.994	0.971	0.971	0.029	0.965	0.965	0.196	0.952	0.215
0.0826749	0.988	0.932	0.932	0.067	0.919	0.919	0.201	0.959	0.245
0.1018302	0.957	0.846	0.847	0.147	0.803	0.803	0.212	0.938	0.308
0.1248141	0.890	0.771	0.779	0.203	0.661	0.661	0.216	0.898	0.356
0.1520999	0.829	0.636	0.659	0.302	0.465	0.465	0.237	0.900	0.454
0.1840867	0.646	0.459	0.579	0.349	0.105	0.105	0.251	0.865	0.562
0.2210352	0.476	0.292	0.599	0.337	0.232	0.232	0.28	0.85	0.664
0.2629964	0.293	0.158	0.724	0.247	0.549	0.549	0.306	0.833	0.736
0.3097423	0.152	0.067	0.851	0.142	0.781	0.781	0.352	0.822	0.783
0.3607171	0.067	0.022	0.933	0.066	0.911	0.911	0.423	0.815	0.803
0.4150261	0.030	0.004	0.970	0.030	0.966	0.966	0.625	0.812	0.810
1.0000000	0.000	0.000	1.000	0.000	1.000	1.000	-	0.808	0.808

Note: PPV, positive predictive value; NPV, negative predictive value; ACC, accuracy (proportion correctly classified.).

10-Year logistic model of static-99R and optimal threshold (Table 4 & Figure 2)

Table 4 Sensitivity, 1- Specificity, Distance (0,1), Sen*Spe, IU, |Sen-Sep|, PPV, NPV, and ACC at Stat-99R cut-points

Positive if greater than or equal to[a]	Sen	1 - Spe	Distance	Sen*Spe	IU	|TP-TN|	PPV	NPV	ACC
0.0000000	1.000	1.000	1.000	0	1	1.000	0.279	-	0.279
0.1150187	0.989	0.951	0.951	0.048	0.9406	0.941	0.287	0.923	0.311
0.1408593	0.979	0.894	0.895	0.103	0.8732	0.873	0.297	0.929	0.349
0.1713720	0.937	0.776	0.779	0.209	0.7132	0.713	0.318	0.902	0.422
0.2068914	0.853	0.720	0.734	0.239	0.5721	0.572	0.314	0.831	0.440
0.2475602	0.811	0.581	0.611	0.330	0.3918	0.392	0.350	0.851	0.528
0.2932531	0.568	0.370	0.568	0.358	0.0814	0.062	0.372	0.791	0.613
0.3435152	0.347	0.199	0.682	0.278	0.4534	0.453	0.402	0.761	0.674
0.3975338	0.179	0.106	0.828	0.160	0.7153	0.715	0.395	0.738	0.695
0.4541615	0.063	0.049	0.938	0.060	0.888	0.888	0.333	0.724	0.704
1.0000000	0.000	0.000	1.000	0.000	1.000	1.000	-	0.721	0.721

Note: PPV, positive predictive value; NPV, negative predictive value; ACC, accuracy (proportion correctly classified).
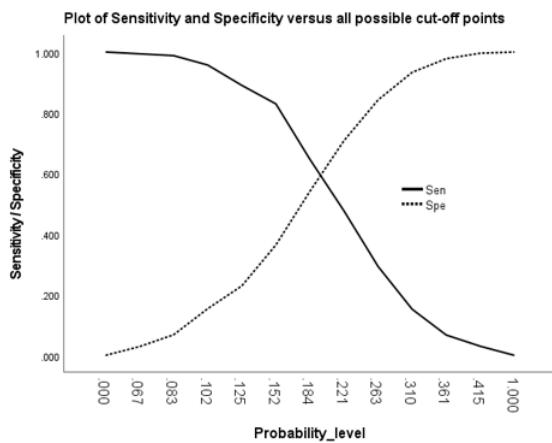
Figure 1 Plot of sensitivity and specificity versus possible cut-points based on Table 2.
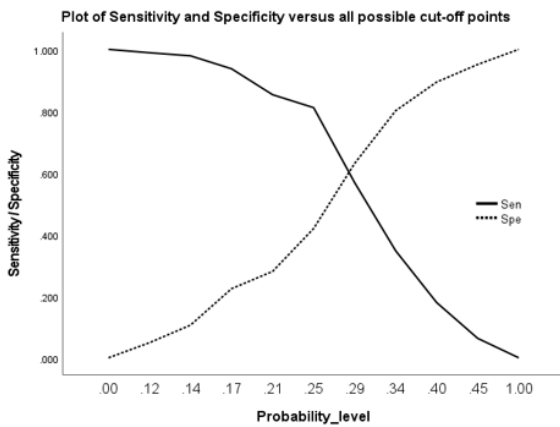

Figure 2 Plot of sensitivity and specificity versus possible cut-points based on Table 4.

Discussion

The Static 99R has been administered in many countries including the United States. It is utilized by psychiatrists or psychologists as part of their clinical evaluation of sex offenders to determine whether the sex offender is likely to recidivate. This study presents the four independent methods: the point closest-to-(0,1), the concordance probability (CZ), the index of union (IU), and the plot of sensitivity versus specificity to find the optimal threshold that classifies most of the individuals correctly and provides the diagnosis (recidivate or not). Remarkably, all four methods yielded identical results. For the 5-year high risk, our findings indicated that the optimal threshold is 0.184, corresponding to a cut-off score of 5. Therefore, if an offender receives a score of 5 or higher, implies that the offender is very likely to recidivate. Similarly, for the 10-year high risk, the optimal threshold is determined to be 0.293, corresponding also to a cut-off score of 5. Therefore, once again, a score of 5 or above implies a high likelihood of recidivism. It should be noted that although all four methods produced similar results, "the point closest to (0,1)" is the most preferred one since we want to minimize the probability of false positive and maximize the probability of true positives.

Conclusion

In our study, all four methods produced identical results for both models. It suggests a high level of consistency and agreement in determining the optimal threshold for the Static 99R. This consistency reinforces the reliability and validity of the findings. The thresholds determined in our study provide valuable guidance for professionals in making informed decisions regarding treatment, supervision, and intervention strategies. By incorporating these thresholds into their decision-making processes, professionals can adopt proactive measures to reduce the potential for future reoffending and enhance overall public safety. It should be noted that one deficiency of Static 99R is the fact that it does not differentiate between individual sex offenders who have the same total score. We suggest that a multiple binary logistics regression with ten independent variables will produce a more meaningful statistical model than the simple logistic regression with the total score as the only one independent variable.

Appendix

Appendix A: data and logistic model for 5 years high

Appendix B: data and logistic model for 10 years high

Appendix A Observed and estimated 5-year sexual recidivism rates for Static-99R: high risk/need sample

Score	Recidivists/total	Observed recidivism rate (%)	Predicted recidivism rate[1]		95% CI	
	Fixed follow-up		**Logistic regression estimates**			
-3	0/1	0				
-2	0/5	0				
-1	1/21	4.8	5.6	(5.97)	3.5	9.1
0	1/28	3.6	7.2	(7.4)	4.7	10.7
1	5/64	7.8	9.0	(9.14)	6.4	12.5
2	11/63	17.5	11.3	(11.23)	8.6	14.6
3	10/103	9.7	14.0	(13.73)	11.3	17.2
4	30/152	19.7	17.3	(16.69)	14.5	20.5
5	28/143	19.6	21.2	(20.13)	18.0	24.8
6	30/122	24.6	25.7	(24.08)	21.5	30.3
7	23/86	26.7	30.7	(28.52)	25.1	37.0
8	14/45	31.1	36.3	(33.43)	28.8	44.5
9	6/18	33.3	42.2	(38.72)	32.6	52.5
10	5/8	62.5	48.4	(44.29)	36.6	60.5
11	0/1	0.0				
Total	164/860	19.1				

[1]The values inside the parentheses are obtained from our replicated logistic regression model

Appendix B Observed and estimated 10-year sexual recidivism rates for Static-99R: high risk/need sample

	Fixed follow-up		Logistic regression estimates			
Score	Recidivists/total	Observed recidivism rate (%)	Predicted recidivism rate[1]		95% CI	
-3	0/1	0.0				
-2	0/5	0.0				
-1	1/21	4.8	5.6	(5.97)	3.5	9.10
0	1/28	3.6	7.2	(7.40)	4.7	10.7
1	5/64	7.8	9.0	(9.14)	6.4	12.5
2	11/63	17.5	11.3	(11.23)	8.6	14.6
3	10/103	9.7	14.0	(13.73)	11.3	17.2
4	30/152	19.7	17.3	(16.69)	14.5	20.5
5	28/143	19.6	21.2	(20.13)	18.0	24.8
6	30/122	24.6	25.7	(24.08)	21.5	30.3
7	23/86	26.7	30.7	(28.52)	25.1	37.0
8	14/45	31.1	36.3	(33.43)	28.8	44.5
9	6/18	33.3	42.2	(38.72)	32.6	52.5
10	5/8	62.5	48.4	(44.29)	36.6	60.5
11	0/1	0.0				
Total	164/860	19.1				

[1] The values inside the parentheses are obtained from our replicated logistic regression model

Acknowledgements

Conflicts of interest

Funding

References

1. Gonçalves LC, Gerth J, Rossegger A, et al. Predictive validity of the Static-99 and Static-99R in Switzerland. *Sexual Abuse J Research Treat*. 2020;32(2):203–219.

2. Hanson RK, Lunetta A, Phenix A, et al. The field validity of Static-99/R sex offender risk assessment tool in California. *J Threat Assess Manag*. 2014;1(2):102.

3. Lee SC, Hanson RK, Yoon JS. Predictive validity of Static-99R among 8,207 men convicted of sexual crimes in South Korea: a prospective field study. *Sexual Abuse*. 2022;10790632221139173.

4. Smallbone S, Rallings M. Short-term predictive validity of the Static-99 and Static-99-R for indigenous and nonindigenous Australian sexual offenders. *Sexual Abuse*. 2013;25(3):302–316.

5. Helmus LM, Lee SC, Phenix A. et al. Static-99R.

6. Phenix A, Fernandez Y, Harris AJ, et al. Static-99R coding rules, revised-2016. Public Safety Canada. 2016.

7. Static-99R Coding Rules Revised, 2016.

8. Pepe MS. The statistical evaluation of medical tests for classification and prediction. USA, Oxford University Press; 2003.

9. Perkins NJ, Schisterman EF. The inconsistency of "optimal" cut-points obtained using two criteria based on the receiver operating characteristic curve. *Am J Epidemiol*. 2006;163(7):670–675.

10. Liu X. Classification accuracy and cut point selection. *Stat Med*. 2012;31(23):2676–2686.

11. Unal I. Defining an optimal cut-point value in ROC analysis: an alternative approach. *Comput Math Methods Med*. 2017;2017:3762651.

12. Hosmer DW, Lemeshow S, Sturdivant RX. 2013. Applied logistic regression 3rd Ed.