

8-31-2021

Bayesian Sensitivity-Specificity and ROC Analysis for Finding Key Drivers

Stan Lipovetsky

GfK North America, stan.lipovetsky@gmail.com

Michael W. Conklin

GfK North America, mike.conklin@gfk.com

Follow this and additional works at: <https://digitalcommons.wayne.edu/jmasm>



Part of the [Applied Statistics Commons](#), [Social and Behavioral Sciences Commons](#), and the [Statistical Theory Commons](#)

Recommended Citation

Lipovetsky, S., & Conklin, M. W. (2020). Bayesian sensitivity-specificity and ROC analysis for finding key drivers. *Journal of Modern Applied Statistical Methods*, 19(1), eP3023. doi: [10.22237/jmasm/1619481960](https://doi.org/10.22237/jmasm/1619481960)

This Regular Article is brought to you for free and open access by the Open Access Journals at DigitalCommons@WayneState. It has been accepted for inclusion in Journal of Modern Applied Statistical Methods by an authorized editor of DigitalCommons@WayneState.

Bayesian Sensitivity-Specificity and ROC Analysis for Finding Key Drivers

Stan Lipovetsky
GfK North America
Minneapolis, MN

Michael W. Conklin
GfK North America
Minneapolis, MN

Finding key drivers in regression modeling via Bayesian Sensitivity-Specificity and Receiver Operating Characteristic is suggested, and clearly interpretable results are obtained. Numerical comparisons with other techniques show that this methodology can be useful in practical statistical modeling and analysis helping to researchers and managers in making meaningful decisions.

Keywords: Predictor importance, Bayesian sensitivity and specificity, receiver operating characteristic, key drivers

Introduction

Finding key drivers of the outcome variability and using them to predict a possible lift in the dependent variable of interest is a common problem of the statistical modeling and analysis in marketing research. Estimation of the individual predictors' importance is often considered via partitioning of the measure of the variability – the coefficient of multiple determination R^2 – to the so-called net effects, and this problem has been described in numerous works (for instance, Kruskal & Majors, 1989; Lipovetsky & Conklin, 2001; Johnson & Lebreton, 2004; Gromping, 2007). Measures of predictor importance based on an orthonormal matrix approximation to data are given in Green et al. (1978), Johnson (2000), and Lipovetsky and Conklin (2015), and other approaches are used as well for finding key drivers and predictors' values for a desirable lift in the outcome variable (for example, Budescu & Azen, 2004; Conklin et al., 2004; Liakhovitski et al., 2010; Buschken et al., 2013; Gray, 2014; Lipovetsky, 2017).

The current paper considers detection of the key drivers based on Bayesian Sensitivity-Specificity in the Receiver Operating Characteristic – (BASS-ROC).

Bayesian approach has been applied to a wide range of problems (for instance, Gregory, 2010; Shemyakin & Kniazev, 2017), and Sensitivity-Specificity in ROC curve is a well-known tool helping in various research projects (for example, Axelsson, 2000; Lipovetsky, 2015). The consideration is performed for the binary dependent variable y and binary predictors often used in applied research, and the continuous variables can also be dichotomized around some thresholds. This approach yields more clearly interpretable and adequate results for the key driver identification, as will be discussed below. Numerical comparisons with other techniques support the conclusion that the BASS-ROC tool can be very useful in practical statistical modeling and analysis helping managers to find meaningful decisions.

It is important to emphasize that finding the key drivers differs from the problem of selection of the main predictors and reducing the number of predictors in regression model to a smaller set. In the key driver analysis, the researchers want to keep and use all the available regressors as tools for efficient managerial decisions. For example, in the analysis of relations of clients with a bank, the managers need to identify the main possible drivers of influence and pursue the most beneficial directions of improving customer satisfaction using all available tools.

The paper is structured as follows: Bayesian sensitivity and specificity are described, these are applied to the key driver problem, and the final section summarizes.

Bayesian Sensitivity and Specificity Analysis

Let us briefly describe the characteristics of Sensitivity and Specificity used in evaluating the performance of classification tests. Suppose we have a binary dependent variable y and binary predictors x_1, x_2, \dots, x_n (where n is their number). Sensitivity, or the True Positive rate (TP), is defined by the proportion of the happened event ($y = 1$) correctly identified by an applied statistical tool (e.g., in prediction by regression), and it is complementary to the False Negative rate (FN, which corresponds to the Type II error). Specificity, or the True Negative rate (TN), is defined by the proportion of the non-happened cases ($y = 0$) which are correctly identified as such by the same applied statistical tool, and it is complementary to the False Positive rate (FP, corresponding to the Type I error). All these occurrences can be summarized in the 2 by 2 cross-table presented in [Figure 1](#).

BAYESIAN SENSITIVITY-SPECIFICITY, ROC, AND KEY DRIVERS

		Cause, or observed value	
		y=1	y=0
Effect, or outcome value	y=1	True Positive	False Positive
	y=0	False Negative	True Negative

Figure 1. Contingency table for a binary classification

A comprehensive description on this topic is given, for instance, in Cook (2018) or in Wikipedia (Sensitivity and specificity, n.d.).

Let us consider an explicit example in terms of “Intrusion-Alarm” problem described in Axelsson (2000). For a detection system, let us denote the Intrusion and Alarm events I and A . In general terms, they correspond to Cause and Effect, respectively, so I corresponds to the columns, and A to the rows in Figure 1. A detection system should have such properties:

- a) A high detection rate, in other words high True Positive (TP) rate, or Sensitivity: if there is intrusion, there will be alarm, so we should be worried about it. It is estimated as the conditional probability $TP = P(A | I)$. It also corresponds to a low False Negative rate (FN), that is $P(\neg A | I)$, i.e. a rare absence of alarm when intrusion happened.
- b) A high True Negative rate, or Specificity: if there is no intrusion, there will be no alarm, so no worries. It is estimated as the conditional probability $TN = P(\neg A | \neg I)$, where we denote the non-events as no-Intrusion = $\neg I$, and no-Alarm = $\neg A$. It also means a low False Positive rate (FP), i.e., a rare false alarm in absence of intrusion, $P(A | \neg I)$.

Using the counts in the cells of contingency table in Figure 1, the sensitivity is defined as the empirical conditional probability

$$\text{Sensitivity} = P(A | I) = \frac{\text{number of TP}}{\text{number of TP} + \text{number of FN}} \quad (1)$$

and the specificity is defined as another empirical conditional probability

$$\text{Specificity} = P(-A | -I) = \frac{\text{number of TN}}{\text{number of TN} + \text{number of FP}} \quad (2)$$

An effective detection system can be considered from the opposite point of view of the so-called Bayesian detection rates, which means that judging by the Alarm we should be able to recognize is there an actual Intrusion or not. It means the detection system should have the following features as the sufficient and necessary conditions:

- a) Alarm really indicates that there is intrusion, $P(I | A)$, which is called the Bayesian detection rate;
- b) Absence of Alarm means that there is no Intrusion, $P(-I | -A)$, so nothing to worry about.

Applying Bayes' theorem to calculate these probabilities, we can associate the corresponding quantities with the Bayesian Sensitivity (BSens)

$$\text{BSens} = P(I | A) = \frac{P(A | I)P(I)}{P(A)} = \frac{P(A | I)P(I)}{P(A | I)P(I) + P(A | -I)P(-I)} \quad (3)$$

and the Bayesian Specificity (BSpec)

$$\begin{aligned} \text{BSpec} = P(-I | -A) &= \frac{P(-A | -I)P(-I)}{P(-A)} \\ &= \frac{P(-A | -I)P(-I)}{P(-A | -I)P(-I) + P(-A | I)P(I)} \end{aligned} \quad (4)$$

These quantities can be presented using the Receiver Operating Characteristic, or ROC curve, plotted as Sensitivity versus 1-Specificity, or True Positive versus False Positive rates.

Methodology of BASS-ROC Key Drivers Analysis

The problem of identification of the key drivers is usually considered via the derived importance of the predictors in a regression model, when the variables with the maximum impact on the outcome dependent variable (judged by the net effects, or other indices of individual predictor's contribution to the model) are interpreted as key drivers. For a constructed model, the prediction by it estimates the probability of the event happened ($y = 1$) due to the various combinations of the predictors' values.

In the current approach, subject to the condition that the event happened ($y = 1$), we consider the question: what is the probability of each predictor contribution to it? To answer it, we apply the new tool of the Bayesian Sensitivity-Specificity analysis (3)-(4) and compare the importance of the predictors plotting them on the plane of the Bayesian sensitivity versus one minus Bayesian specificity, or True Positive versus False Positive Bayesian rates.

To present the new methodology in the explicit form, let us describe each of its steps by the numerical example on a real marketing research data of a bank's credit card usage. The data was elicited from 170 respondents by the dependent variable of the card used or not ($y = 1$ or 0, respectively) is modeled by sixteen other binary predictors named in Table 1.

With two states of the outcome y_i and a set of predictors x_{ij} ($i = 1, 2, \dots, N$ – the number of observations; $j = 1, 2, \dots, n$ – the number of independent variables), the binary logistical regression model is

$$p(y_i | z_i) = \frac{1}{1 + \exp(-z_i)}, \quad (5)$$

where z is a linear aggregate of the predictors:

$$z_i = a_0 + a_1x_{i1} + a_2x_{i2} + \dots + a_nx_{in}. \quad (6)$$

The intercept a_0 and parameters a_j are usually estimated by the maximum likelihood criterion. Table 1 presents the results of logistic modeling.

In Table 1 we see the pair correlations of y with each x in the first numerical column and the coefficients of logistic regression in the next column. In addition, there is a column of the percent of predictors contribution to the quality of the model fit estimated via the so-called Shapley Value (or SV, described in detail in Lipovetsky & Conklin, 2001), and their ranks corresponding to the importance of

key drivers. We see that the correlations are weak, and the logit model is of a low quality of fit (its residual deviance 177.42 is too close to the null deviance is 190.09, so the pseudo- R^2 as an analogue of the coefficient of multiple determination equals only 6.7%). But SV yields rather robust estimations of the net effects, by which we can judge by the first ranks about the key drivers (several main of those are defined in the order $x_3, x_5, x_{11}, x_8, x_1$, etc.).

With the found parameters of the logistic regression, the prediction of probability $p(y_i | z_i)$ of choice y_i in any i^{th} point of predictor values are made by the model (5)-(6). In the case of rare events (for example, a small number of 1s in the binary outcome), correction of the logit intercept is needed, otherwise most of the predicted values will be skewed to the side of the more frequent event (of 0s). For making the cross-table as in Figure 1, it is important to identify predictions of 1 and 0 by the continuous probability obtained by the logistic regression. It means we need to decide where we should set the threshold for rounding probabilities to 1 or to 0 – should it be at the middle of the logit curve with the probability 0.5, or should the threshold be shifted to a larger or smaller level. Such an adjustment can be performed by a procedure of shifting in the intercept of the logistic regression described in King and Zeng (2001).

Table 1. Correlations, logit parameters, predictor contributions, and their ranks

	Predictors	r(y, x)	Logit	SV %	Rank
	Intercept		0.790		
x_1	Easy	0.082	0.563	9.470	5
x_2	Simple	-0.013	-0.195	3.433	9
x_3	Frictionless	-0.061	-0.910	18.846	1
x_4	Protects	0.006	-0.373	3.316	10
x_5	Private	0.116	0.826	17.288	2
x_6	Liability	0.016	-0.414	2.158	11
x_7	Relevant	0.031	0.069	0.722	16
x_8	Customize	0.081	0.638	10.066	4
x_9	Personalized	0.048	-0.192	0.863	14
x_{10}	Compatible	0.083	0.432	6.626	7
x_{11}	Accepted	0.096	0.596	10.604	3
x_{12}	Limits	-0.026	-0.733	8.801	6
x_{13}	Standard	0.034	-0.289	2.093	12
x_{14}	Control	0.036	0.333	1.876	13
x_{15}	Impressed	0.056	0.255	3.438	8
x_{16}	Intrigued	0.016	0.210	0.816	15

BAYESIAN SENSITIVITY-SPECIFICITY, ROC, AND KEY DRIVERS

Table 2. ROC sliding threshold for finding the optimum predictions by the logit model

Threshold	Hit rate	Sensitivity	Specificity	1-Specificity	Distance	Rank
0.479	0.771	0.977	0.143	0.857	0.857	19
0.501	0.776	0.977	0.167	0.833	0.834	15
0.523	0.771	0.969	0.167	0.833	0.834	16
0.545	0.765	0.961	0.167	0.833	0.834	17
0.566	0.771	0.961	0.190	0.810	0.810	14
0.588	0.771	0.953	0.214	0.786	0.787	12
0.610	0.765	0.945	0.214	0.786	0.788	13
0.632	0.771	0.938	0.262	0.738	0.741	9
0.654	0.741	0.898	0.262	0.738	0.745	10
0.675	0.706	0.836	0.310	0.690	0.710	8
0.697	0.676	0.766	0.405	0.595	0.640	6
0.719	0.647	0.703	0.476	0.524	0.602	3
0.741	0.624	0.664	0.500	0.500	0.602	4
0.763	0.594	0.586	0.619	0.381	0.563	1
0.784	0.571	0.539	0.667	0.333	0.569	2
0.806	0.494	0.406	0.762	0.238	0.640	5
0.828	0.471	0.344	0.857	0.143	0.672	7
0.850	0.394	0.227	0.905	0.095	0.779	11
0.871	0.365	0.164	0.976	0.024	0.836	18
0.893	0.312	0.086	1.000	0.000	0.914	20

We suggest another way of the model improvement based also on the application of an ROC curve. In order to perform it, we use a sliding threshold added to the intercept of the logistic regression, then find the cross-table (Figure 1) of the predicted versus the observed values of the dependent variable, calculate sensitivity and specificity. The results for several values of the threshold are shown in Table 2.

Table 2 also presents the hit rate (trace on the diagonal divided by the total number of observations), values of sensitivity (1), specificity (2), and one minus specificity as the False Positive rate. As is suggested in the signal detection theory and ROC analysis, we look for the Sensitivity to be close to 1, and 1-Specificity (FP rate) close to 0, which can be estimated by the Distance to the North-West corner in the coordinates of ROC-curve

$$D = \sqrt{(1 - \text{Sensitivity})^2 + (1 - \text{Specificity})^2} . \quad (7)$$

Distance values are also shown in Table 2, with the corresponding ranks where the rank 1 defines the minimum distance, and the corresponding value of the threshold probability in this row equals $p_{\text{threshold}} = 0.763$.

The value z^* corresponding to this optimum threshold can be found from (5) in the linear link function as follows:

$$z^* = \ln \frac{p_{\text{threshold}}}{1 - p_{\text{threshold}}} . \quad (8)$$

Then we can find such a shift Δz to the value z^* (8) that

$$z^* + \Delta z = 0 , \quad (9)$$

with the probability in the center of the logistic curve matches $p = 0.5$. This shift equals

$$\Delta z = -z^* = \ln \frac{1 - p_{\text{threshold}}}{p_{\text{threshold}}} . \quad (10)$$

The constant Δz (10) can be added to the intercept in the aggregate (6)

$$a_0^{\text{adj}} = a_0 + \Delta z , \quad (11)$$

and the adjusted intercept a_0^{adj} (11) can be used in predictions by logit (5) with the guaranteed maximum for sensitivity and specificity due to the ROC optimum in (7).

In our example with $p_{\text{threshold}} = 0.763$, the intercept of the logit model shown in Table 1 as 0.790 can be corrected to the value -0.377 which yields the optimum predictions corresponding to the minimum of both False Positive and Negative rates, or minimum of both Type I and Type II errors. It is interesting to note that the optimum threshold defined in terms of the best combination of TP and TN rates does not necessarily mean the maximum for the hit rate value seen in Table 2. The results of Table 2 are also illustrated by the ROC curve shown in Figure 2.

BAYESIAN SENSITIVITY-SPECIFICITY, ROC, AND KEY DRIVERS

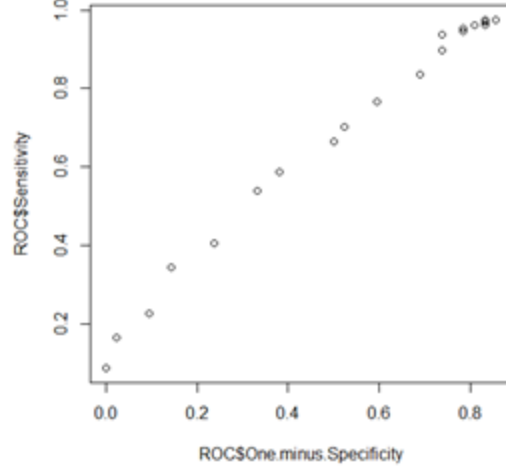


Figure 2. ROC curve for finding an optimum threshold in logit model adjustment

Using adjusted logit, the Bayesian Sensitivity (3) for many predictors can be presented as follows:

$$\text{BSens}_{ij} = P(x_{ij} = 1 | y_i = 1) = \frac{p(y_i = 1 | z_i(x_{ij} \equiv 1)) \cdot p(x_j = 1)}{\sum_{k=1}^n p(y_i = 1 | z_i(x_{ik} \equiv 1)) \cdot p(x_k = 1)}, \quad (12)$$

where the first term in the numerator $p(y_i = 1 | z_i(x_{ij} \equiv 1))$ (called likelihood in the Bayesian formula) is defined by the prediction by the formula $p(y_i | z_i)$ (5) with the predictors' values taken from each i^{th} row for the aggregate (6) but with the x_j put to one. The second term in the numerator (12) is the prior sample probability $p(x_j = 1)$ of each predictor, and the denominator normalizes all the j^{th} shares to the total equal one for each i^{th} observation.

Similarly, the Bayesian Specificity (4) for many predictors is given as follows:

$$\text{BSpec}_{ij} = P(x_{ij} = 0 | y_i = 0) = \frac{p(y_i = 0 | z_i(x_{ij} \equiv 0)) \cdot p(x_j = 0)}{\sum_{k=1}^n p(y_i = 0 | z_i(x_{ik} \equiv 0)) \cdot p(x_k = 0)}, \quad (13)$$

where the first term in the numerator $p(y_i = 0 | z_i(x_{ij} \equiv 0))$ is defined as $1 - p(y_i | z_i)$ with the prediction $p(y_i | z_i)$ made by the formula (5) with the predictors' values taken from each i^{th} row for the aggregate (6) but with the x_j put to zero. The second term in the numerator (13) is the prior sample probability of $x_j = 0$ which can be defined as $p(x_j = 0) = 1 - p(x_j = 1)$ for each predictor and the denominator normalizes all j^{th} shares to the total equal one in each observation.

By the obtained values (12)-(13) in i^{th} observation, we find the distances (similar to (7)) to estimating closeness of Bayesian sensitivity and specificity of the predictors to their optimum:

$$D_{ij} = \sqrt{(1 - \text{BSens}_{ij})^2 + (1 - \text{BSpec}_{ij})^2} . \quad (14)$$

Averaging Bayesian probabilities (12)-(13) across observations, then applying the formula (14) and ranking the distance yields the key driver results on the aggregate level. Table 3 presents the results obtained by the same data on the aggregated level: Bayesian sensitivity (12) and specificity (13), distance (14), and the corresponding ranks. The first ranks identify the key drivers in the order x_5, x_{11}, x_1, x_8 , etc.

Matching the key drivers order obtained in regression modeling (Table 1) with those from BASS-ROC identification (Table 3), we see that the main key drivers are the same with the striking exception of the first driver x_3 found by the regression model. This predictor has the negative correlation with the outcome (see in Table 1), thus, although it increases variability of the predictors' aggregate and helps to yield a better fit to the dependent variable, but its individual values 0 and 1 correspond rather to the values 1 and 0 of the outcome variable. Therefore, the Bayesian sensitivity-specificity approach reveals its weak connection with the outcome values.

An example of ranking for the predictors on the individual respondents' level is presented in Table 4 for arbitrarily chosen ten observations out of total sample of 170 observations. We see that the main key driver x_5 can also be on the second or third place, or the next important driver x_{11} can take from the first to the fifth place, or the worst predictor x_3 can occupy the places from 13th to 16th, being not always the last by importance, and so on. The last column in Table 4 shows the standard deviation of the ranks of each predictor among all 170 respondents, and we see that the mean deviation is mostly about one step in ranks, and sometimes even two steps. Therefore, there exists a variability of the key drivers for each

BAYESIAN SENSITIVITY-SPECIFICITY, ROC, AND KEY DRIVERS

individual response. This result could be very useful if an aim of research requires finding key drivers specifically for individual respondents.

Table 3. Bayesian Sensitivity and Specificity, distance to optimum, and ranks

Predictors	BSens or TPRate	BSpec or TNRate	Distance	Rank	
X ₁	Easy	0.089	0.056	1.311	3
X ₂	Simple	0.067	0.052	1.330	10
X ₃	Frictionless	0.055	0.045	1.344	16
X ₄	Protects	0.073	0.043	1.332	11
X ₅	Private	0.091	0.064	1.305	1
X ₆	Liability	0.047	0.062	1.337	14
X ₇	Relevant	0.048	0.073	1.329	8
X ₈	Customize	0.058	0.081	1.316	4
X ₉	Personalized	0.067	0.052	1.330	9
X ₁₀	Compatible	0.071	0.067	1.317	5
X ₁₁	Accepted	0.095	0.053	1.310	2
X ₁₂	Limits	0.045	0.057	1.342	15
X ₁₃	Standard	0.064	0.051	1.333	12
X ₁₄	Control	0.063	0.069	1.321	6
X ₁₅	Impressed	0.043	0.081	1.327	7
X ₁₆	Intrigued	0.023	0.093	1.333	13

Table 4. Ranks of key drivers on individual respondents' level

Observation	1	2	3	4	5	6	7	8	9	10	stdev	
X ₁	Easy	3	3	4	3	4	4	3	4	3	2	1.0
X ₂	Simple	10	11	9	10	10	8	8	10	10	9	1.1
X ₃	Frictionless	16	13	16	15	15	16	16	14	15	15	0.7
X ₄	Protects	13	15	11	11	11	12	12	8	11	11	1.4
X ₅	Private	2	1	1	1	2	3	2	1	1	1	0.7
X ₆	Liability	14	14	13	14	14	13	13	15	14	16	0.6
X ₇	Relevant	8	8	10	8	8	9	10	13	13	8	0.8
X ₈	Customize	5	4	2	4	1	1	5	2	2	3	1.6
X ₉	Personalized	9	10	8	9	9	7	7	9	9	13	1.2
X ₁₀	Compatible	4	5	5	5	5	6	4	5	6	4	0.8
X ₁₁	Accepted	1	2	3	2	3	2	1	3	4	5	0.9
X ₁₂	Limits	15	16	15	16	16	15	15	16	16	14	0.7
X ₁₃	Standard	11	12	12	13	12	11	11	11	12	12	0.8
X ₁₄	Control	6	6	7	6	6	5	6	6	7	6	0.5
X ₁₅	Impressed	7	7	6	7	7	10	9	7	5	7	1.0
X ₁₆	Intrigued	12	9	14	12	13	14	14	12	8	10	1.9

Summary

A possibility to find key drivers in Bayesian sensitivity and specificity approach with the corresponding ROC analysis has been investigated. The proposed methods and methodology are useful for practical application in numerous problems in marketing research. The obtained results demonstrate that finding key drivers can be achieved very easily in comparison with commonly used complicated techniques and this approach does not require special software. The developed technique can be very useful for finding managerial solutions, particularly, when sets of the key drivers should be specifically focused on individual features of the respondents.

Acknowledgements

We are thankful to three reviewers for their comments improving the paper.

References

- Axelsson, S. (2000). The base-rate fallacy and the difficulty of intrusion detection. *ACM Transactions on Information and System Security*, 3(3), 186-205. doi: 10.1145/357830.357849
- Budescu, D.V., & Azen, R. (2004). Beyond global measures of relative importance: Some insights from dominance analysis. *Organizational Research Methods*, 7(3), 341-350. doi: 10.1177/1094428104267049
- Buschken, J., Otter, T., & Allenby, G. M. (2013). The dimensionality of customer satisfaction survey responses and implications for driver analysis. *Marketing Science*, 32(4), 533-553. doi: 10.1287/mksc.2013.0779
- Conklin, M., Powaga, K., & Lipovetsky, S. (2004). Customer satisfaction analysis: Identification of key drivers. *European Journal of Operational Research*, 154(3), 819-827. doi: 10.1016/S0377-2217(02)00877-9
- Cook, J. D. (2018, September 6). Accuracy, precision, and recall. *John D. Cook Consulting*. Retrieved from <https://www.johndcook.com/blog/2018/09/06/accuracy-precision-and-recall/>
- Gray, K. (2014, February 14). Best practices for key driver analysis. *Quirks' e-Newsletter*. Retrieved from <https://www.quirks.com/articles/best-practices-for-key-driver-analysis>

BAYESIAN SENSITIVITY-SPECIFICITY, ROC, AND KEY DRIVERS

- Green, P. E., Carroll, J. D., & DeSarbo, W. S. (1978). A new measure of predictor variable importance in multiple regression. *Journal of Marketing Research*, 15(3), 356-360. doi: 10.1177/002224377801500305
- Gregory, P. (2010). *Bayesian logical data analysis for the physical sciences: A comparative approach with Mathematica support*. Cambridge, NY: Cambridge University Press.
- Gromping, U. (2007). Estimators of relative importance in linear regression based on variance decomposition. *The American Statistician*, 61(2), 139-147. doi: 10.1198/000313007X188252
- Johnson, J. W. (2000). A heuristic method for estimating the relative weight of predictor variables in multiple regression. *Multivariate Behavioral Research*, 35(1), 1-19. doi: 10.1207/S15327906MBR3501_1
- Johnson, J. W., & Lebreton, J. M. (2004). History and use of relative importance indices in organizational research. *Organizational Research Methods*, 7(3), 238-257. doi: 10.1177/1094428104266510
- King, G., & Zeng, L. (2001). Logistic regression in rare events data. *Political Analysis*, 9(2), 137-163. doi: 10.1093/oxfordjournals.pan.a004868
- Kruskal, W. H., & Majors, R. (1989). Concepts of relative importance in recent scientific literature. *The American Statistician*, 43(1), 2-6. doi: 10.1080/00031305.1989.10475596
- Liakhovitski, D., Bryukhov, Y., & Conklin, M. (2010). Relative importance of predictors: Comparison of random forest with Johnson's relative weights. *Model Assisted Statistics and Applications*, 5(4), 235-249. doi: 10.3233/MAS-2010-0172
- Lipovetsky, S. (2015). Trinomial response modeling in one logit regression. *Annals of Data Science*, 2(2) 157-163. doi: 10.1007/s40745-015-0039-7
- Lipovetsky, S. (2017). Prediction of percent change in linear regression by correlated variables. *Journal of Modern Applied Statistical Methods*, 16(2), 347-358. doi: 10.22237/jmasm/1509495480
- Lipovetsky, S., & Conklin, M. (2001). Analysis of regression in game theory approach. *Applied Stochastic Models in Business and Industry*, 17(4), 319-330. doi: 10.1002/asmb.446
- Lipovetsky, S., & Conklin, M. (2015). Predictor relative importance and matching regression parameters. *Journal of Applied Statistics*, 42(5), 1017-1031. doi: 10.1080/02664763.2014.994480

Sensitivity and specificity. (n.d.). In *Wikipedia*. Retrieved from https://en.wikipedia.org/wiki/Sensitivity_and_specificity

Shemyakin, A., & Kniazev, A. (2017). *Introduction to Bayesian estimation and copula models of dependence*. Hoboken, NJ: Wiley.