Journal of Risk Model Validation **14**(4), 89–100
DOI: 10.21314/JRMV.2020.227

Copyright Infopro Digital Limited 2021. All rights reserved. You may share using our article tools. This article may be printed for the sole use of the Authorised User (named subscriber), as outlined in our terms and conditions. https://www.infopro-insight.com/termsconditions/insight-subscriptions

Research Paper

Statistical properties of the population stability index

Bilal Yurdakul[1] and Joshua Naranjo[2]

[1] Enterprise Risk Management & Analytics, USAA, 10750 McDermott Freeway, San Antonio, TX 78288, USA; email: bilalyurdakul@gmail.com

[2] Department of Statistics, Western Michigan University, 1903 W Michigan Avenue, Kalamazoo, MI 49008, USA; email: joshua.naranjo@wmich.edu

(Received August 14, 2019; revised May 13, 2020; accepted May 13, 2020)

ABSTRACT

The population stability index (PSI) is a widely used statistic that measures how much a variable has shifted over time. A high PSI may alert the business to a change in the characteristics of a population. This shift may require investigation and possibly a model update. PSI is commonly used among banks to measure the shift between model development data and current data. Banks may face additional risks if models are used without proper validation. The incorrect use of PSI may bring unexpected risks for these institutions. However, there are not many studies about the statistical properties of PSI. In practice, the following "rule of thumb" is used: PSI < 0.10 means a "little change", $0.10 \leqslant$ PSI < 0.25 means a "moderate change" and $0.25 \leqslant$ PSI means a "significant change, action required". These benchmarks are used without reference to statistical type I or type II error rates. This paper aims to fill the gap by providing statistical properties of the PSI and some recommendations on its use.

Keywords: population stability; divergence; model validation; model monitoring; model risk; credit risk.

Corresponding author: B. Yurdakul

Print ISSN 1753-9579 | Online ISSN 1753-9587
© 2021 Infopro Digital Risk (IP) Limited

1 INTRODUCTION

The population stability index (PSI) is a measure of the population stability between two population samples. Suppose we are interested in comparing the distribution of credit scores for a base year, say 2017, and a target year, say 2019. Did the distribution of credit scores remain the same or did it change? The PSI is calculated by classifying the credit scores into B bins and comparing the multinomial frequencies over the two years. For example, credit scores are classified into seven bins in Table 1, with the highest scores in grade A and the lowest scores in grade G. The table is based on data freely available from LendingClub (2018). The bin boundaries were determined by LendingClub.

Let n be the sample size for the base population and m be the sample size for the target population. Let $\hat{p}_i = x_i/n$ be the relative frequency of the ith grade for the base year and $\hat{q}_i = y_i/m$ be the relative frequency of the ith grade for the target year. The PSI is computed as

$$\mathrm{PSI} = \sum_{i=1}^{B} (\hat{p}_i - \hat{q}_i)(\ln \hat{p}_i - \ln \hat{q}_i). \tag{1.1}$$

In general, base year data refers to the development data for model validation, and PSI gives a measure of the change in distribution between current and development data. A high PSI indicates a large shift, which then makes the model questionable for use. Even though the PSI is a widely used tool to measure population stability, its distributional properties are not well known. In this paper, we discuss the statistical properties of the PSI. Knowing these properties allows the derivation of, among other things, benchmarks with known rates of false significance.

2 LITERATURE REVIEW

The PSI is discussed in several books on credit scoring or risk analysis (see, for example, Anderson 2007; Baesens 2016; Lewis 1994; Siddiqi 2017). Recommendations for industry best practice are made by Pruitt (2010), along with the implementation of the PSI in SAS ENTERPRISE MINER. There is a patent issued for a machine that calculates the PSI by Liu *et al* (2009). There are also papers and manuals by governmental regulatory bodies that do not discuss the PSI directly (BGFRS–OCC 2011; Federal Deposit Insurance Corporation 2007; Office of the Comptroller of the Currency 2016), but discuss the stability of population as an ongoing monitoring requirement. Most of the literature talks about practical use, including rule-of-thumb recommendations. Anderson (2007) points out that the rule of thumb is used as a traffic light approach in the industry.

TABLE 1 PSI calculation for distribution of credit scores.

Grade (i)	Base ($\hat{p}_i$)	Target ($\hat{q}_i$)	$\hat{p}_i - \hat{q}_i$	ln($\hat{p}_i$) – ln($\hat{q}_i$)	Product
A	0.253	0.177	0.077	0.357	0.027
B	0.302	0.262	0.040	0.142	0.006
C	0.204	0.285	–0.081	–0.334	0.027
D	0.134	0.158	–0.024	–0.165	0.004
E	0.072	0.088	–0.016	–0.201	0.003
F	0.026	0.025	0.001	0.039	0.000
G	0.008	0.006	0.002	0.288	0.001
Total	1.000	1.000			PSI = 0.068

The PSI can be written as some form of Kullback–Leibler (KL) divergence (Kullback 1978; Kullback and Leibler 1951). KL divergence is well studied and can be found in Wu and Olson (2010), Lin (2017), Gottschalk (2016) and Thomas (2009). Some of the books mentioned above make the connection between the PSI and KL divergence. In particular, Anderson (2007) and Thomas *et al* (2017) acknowledge that the PSI is a χ^2-based statistic and discuss the connection between the PSI and KL divergence. However, the rule of thumb is still lacking discussion in the literature.

Let $p(x)$ and $q(x)$ be two distributions of a discrete random variable X. The x divergence of $q(x)$ from $p(x)$ is

$$D_{\text{KL}}(q(x) \mid p(x)) = E_p\left(\ln \frac{p(X)}{q(X)}\right) = \sum_{i=1}^{B} p(x_i) \ln \frac{p(x_i)}{q(x_i)}. \tag{2.1}$$

We might think of $p(x)$ as the true distribution and of $q(x)$ as the model distribution so that D_{KL} represents some sort of loss due to using the wrong distribution. Even though D_{KL} measures the divergence of $q(x)$ from $p(x)$, it is technically not a distance measure because the definition is not symmetric, ie, $D_{\text{KL}}(q(x) \mid p(x)) \neq D_{\text{KL}}(p(x) \mid q(x))$. However, we can easily obtain a symmetric measure of divergence by defining

$$\begin{aligned} D^*(p,q) &= D_{\text{KL}}(q \mid p) + D_{\text{KL}}(p \mid q) \\ &= \sum p(x_i) \ln \frac{p(x_i)}{q(x_i)} + \sum q(x_i) \ln \frac{q(x_i)}{p(x_i)} \\ &= \sum p(x_i) \ln \frac{p(x_i)}{q(x_i)} - \sum q(x_i) \ln \frac{p(x_i)}{q(x_i)} \\ &= \sum (p(x_i) - q(x_i))(\ln p(x_i) - \ln q(x_i)), \end{aligned}$$

which brings us to the formula for the PSI.

Although KL divergence is not symmetric, the PSI is symmetric under the assumption that cutoff values are predetermined. However, in practice that is not true, since the cutoff points correspond to the percentiles of the base population that ultimately determine the PSI. If the base and target populations switch roles, then this changes the cutoff points and consequently the PSI. This is an additional layer of complexity in the use of the PSI in practice.

Kullback (1978, Chapter 6) tackles the problem for multinomial distributions by comparing cell percentages between two populations. Kullback defined $J = (1/n + 1/m)^{-1} \times$PSI. His definition uses the notation $J(1,2)$, defined as the sum of so-called information:

$$J(1,2) = I(1,2) + I(2,1). \tag{2.2}$$

Kullback called J the divergence between H_1 and H_2, two simple statistical hypotheses. H_1 and H_2 basically assume different population percentages. Since Kullback considered only multinomial distribution, he did not have to consider cutoff values for binning. However, per the current use of the PSI, the cell counts or percentages are created from binning the underlying distribution in the development data. The bins are basically dependent on the base population's distribution.

3 STATISTICAL PROPERTIES OF THE POPULATION STABILITY INDEX

Let $\boldsymbol{X} = (X_1, X_2, \dots, X_B)$ be a multinomial random variable with parameters n and $\boldsymbol{p}$, where $\boldsymbol{p} = (p_1, p_2, \dots, p_B)$ and $p_i > 0$ for $i = 1, \dots, B$. In our motivating example, n is the total number of base-year credit scores, and X_i is the number that falls in the ith grade. Similarly, let $\boldsymbol{Y} = (Y_1, Y_2, \dots, Y_B)$ be a multinomial random variable with parameters m and $\boldsymbol{q} = (q_1, q_2, \dots, q_B)$ with $q_i > 0$ for $i = 1, \dots, B$. The following theorem provides a polynomial approximation of the PSI.

THEOREM 3.1 *Let $\hat{p}_i = X_i/n$, where $(X_1, \dots, X_B)$ is multinomial with parameters n and $(p_1, \dots, p_B)$. Let $\hat{q}_i = Y_i/m$, where $(Y_1, \dots, Y_B)$ is multinomial with parameters m and $(q_1, \dots, q_B)$. Then,*

$$\text{PSI} = \text{PSI}^* + O_p(n^{-3/2}) + O_p(m^{-3/2}), \tag{3.1}$$

where

$$\text{PSI}^* = \sum_{i=1}^{B} (\hat{p}_i - \hat{q}_i) \left[\ln p_i - \ln q_i + \frac{\hat{p}_i - p_i}{p_i} - \frac{\hat{q}_i - q_i}{q_i} - \frac{(\hat{p}_i - p_i)^2}{2p_i^2} + \frac{(\hat{q}_i - q_i)^2}{2q_i^2} \right].$$

PROOF By Taylor's expansion,

$$\ln(\hat{p}_i) = \ln(p_i) + \frac{\hat{p}_i - p_i}{p_i} - \frac{(\hat{p}_i - p_i)^2}{2p_i^2} + R_n,$$

where R_n consists of higher-order terms. Now $\sqrt{n}(\hat{p}_i - p_i)$ is asymptotically normal, so $(\hat{p}_i - p_i)$ is $O_p(n^{-1/2})$. Thus, $(\hat{p}_i - p_i)^3$ and the higher-order terms are $O_p(n^{-3/2})$. By similarly expanding $\ln(\hat{q}_i)$ and substituting it into $(\ln \hat{p}_i - \ln \hat{q}_i)$, the result follows. □

The polynomial approximation simplifies considerably when the two multinomials have equal probability vectors.

THEOREM 3.2 *If $p_i = q_i$, $i = 1, \dots, B$, and the assumptions of Theorem 3.1 hold, then*

$$\text{PSI} = \sum_{i=1}^{B} \frac{(\hat{p}_i - \hat{q}_i)^2}{p_i} + o_p(n^{-1}) + o_p(m^{-1}). \tag{3.2}$$

PROOF First, note that $\ln p_i - \ln q_i = 0$. Second,

$$\frac{\hat{p}_i - p_i}{p_i} - \frac{\hat{q}_i - q_i}{q_i} = \frac{\hat{p}_i - \hat{q}_i}{p_i}.$$

Finally,

$$(\hat{p}_i - \hat{q}_i)\frac{(\hat{p}_i - p_i)^2}{2p_i^2} = [(\hat{p}_i - p_i) - (\hat{q}_i - q_i)]\frac{(\hat{p}_i - p_i)^2}{2p_i^2} = o_p(n^{-1}),$$

since $(\hat{p}_i - p_i)$ and $(\hat{q}_i - q_i)$ are $o_p(1)$ and $(\hat{p}_i - p_i)^2$ is $O_p(n^{-1})$. □

Let

$$\text{PSI}^{**} = \sum_{i=1}^{B} \frac{(\hat{p}_i - \hat{q}_i)^2}{p_i}.$$

We will derive the properties of this quadratic approximation. First, we write it in matrix form, ie,

$$\text{PSI}^{**} = \sum_{i=1}^{B} \frac{(\hat{p}_i - \hat{q}_i)^2}{p_i} = V'AV, \tag{3.3}$$

where

$$V = \begin{bmatrix} \hat{p}_1 - \hat{q}_1 \\ \hat{p}_2 - \hat{q}_2 \\ \vdots \\ \hat{p}_B - \hat{q}_B \end{bmatrix} \quad \text{and} \quad A = \begin{bmatrix} \frac{1}{p_1} & 0 & \cdots & 0 \\ 0 & \frac{1}{p_2} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \frac{1}{p_B} \end{bmatrix}.$$

THEOREM 3.3 *Under the assumptions of Theorem 3.2,* $(1/n + 1/m)^{-1}\mathrm{PSI}^{**}$ *has an approximate* χ^2 *distribution with* $B - 1$ *degrees of freedom.*

PROOF Note that

$$\left(\frac{1}{n} + \frac{1}{m}\right)^{-1} \mathrm{PSI}^{**} = V' A^* V,$$

where

$$A^* = \left(\frac{1}{n} + \frac{1}{m}\right)^{-1} A.$$

It is known from multinomial theory that $(\hat{p}_1, \dots, \hat{p}_B)$ is asymptotically multivariate normal, and hence V is approximately normal for large n and m (see, for example, Agresti 2007). To show that $V' A^* V$ has a chi-square distribution, a sufficient condition is that $A^* \Sigma$ is idempotent, where Σ is the variance–covariance matrix of V (see, for example, Stapleton 1995). By the independence of $\hat{p}_i$ and $\hat{q}_i$, we have

$$\begin{aligned}\mathrm{var}(\hat{p}_i - \hat{q}_i) &= \frac{p_i(1 - p_i)}{n} + \frac{q_i(1 - q_i)}{m} \\ &= p_i(1 - p_i)\left(\frac{1}{n} + \frac{1}{m}\right),\end{aligned}$$

since $p_i = q_i$. Also,

$$\mathrm{cov}(\hat{p}_i - \hat{q}_i, \hat{p}_j - \hat{q}_j) = \mathrm{cov}(\hat{p}_i, \hat{p}_j) - \mathrm{cov}(\hat{p}_i, \hat{q}_j) - \mathrm{cov}(\hat{q}_i, \hat{p}_j) + \mathrm{cov}(\hat{q}_i, \hat{q}_j),$$

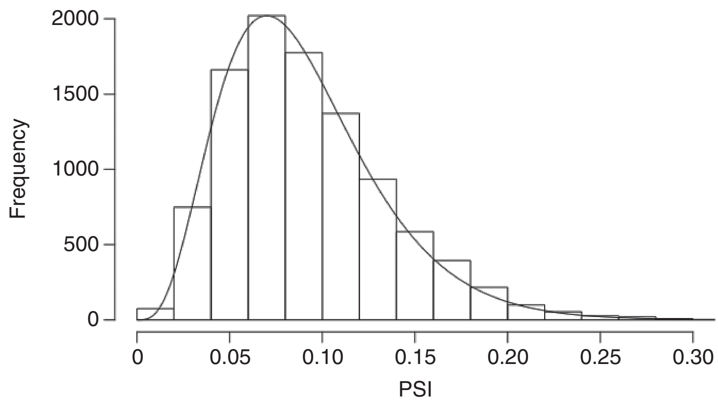
which equals

$$\frac{-p_i p_j}{n} - 0 + 0 + \frac{-q_i q_j}{m} = -p_i p_j \left(\frac{1}{n} + \frac{1}{m}\right).$$

(For details on the covariance between multinomial frequencies, see, for example, Cochran (1977).) Then,

$$A^* \Sigma = \begin{bmatrix} \frac{1}{p_1} & 0 & \cdots & 0 \\ 0 & \frac{1}{p_2} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & & \frac{1}{p_B} \end{bmatrix} \begin{bmatrix} p_1(1 - p_1) & -p_1 p_2 & \cdots & -p_1 p_n \\ -p_2 p_1 & p_2(1 - p_2) & \cdots & -p_2 p_B \\ \vdots & \vdots & \ddots & \vdots \\ -p_n p_1 & -p_n p_2 & \cdots & p_B(1 - p_B) \end{bmatrix},$$

FIGURE 1 Histogram of 10 000 simulated PSI values and overlay of chi-square density.



which can be shown to be idempotent (Yurdakul 2018). Finally, the degrees of freedom are given by

$$\operatorname{tr}(A^* \Sigma) = \operatorname{tr}\begin{bmatrix} 1-p_1 & -p_2 & \cdots & -p_B \\ -p_1 & 1-p_2 & \cdots & -p_b \\ \vdots & \vdots & \ddots & \vdots \\ -p_1 & -p_2 & \cdots & 1-p_B \end{bmatrix} = B-1.$$

□

Theorems 3.1–3.3 (from Yurdakul (2018)) suggest that PSI has a distribution that is approximated by $(1/n + 1/m)$ times a χ^2 random variable with $B-1$ degrees of freedom. We ran simulations to confirm this. For example, we used the `rmultinom()` function in R to generate multinomial counts of $n = 400$ observations falling into $B = 10$ categories with equal probabilities $p_i = 0.10$ for $i = 1, \dots, 10$. The second sample was generated with $m = 400$ and similar probabilities $q_i = 0.10$ for all i. The PSI value (1.1) for comparing the two samples was calculated. Figure 1 shows the histogram of 10 000 simulated values of PSI. The density of $(1/n + 1/m)\chi^2_9$ is overlaid for comparison.

4 BENCHMARKS FOR THE POPULATION STABILITY INDEX

The PSI is used to measure how much a distribution has shifted between two populations or over time. A typical rule of thumb (Lewis 1994) for the extent to which a distribution has shifted is the following: PSI < 0.10 means a "little change",

TABLE 2 PSI benchmarks for $B = 10$ and $\alpha = 0.05$.

	m					
n	**100**	**200**	**400**	**600**	**800**	**1000**
100	0.338	0.254	0.211	0.197	0.190	0.186
200	0.254	0.169	0.127	0.113	0.106	0.102
400	0.211	0.127	0.085	0.070	0.063	0.059
600	0.197	0.113	0.070	0.056	0.049	0.045
800	0.190	0.106	0.063	0.049	0.042	0.038
1000	0.186	0.102	0.059	0.045	0.038	0.034

$0.10 \leqslant \text{PSI} < 0.25$ means a "moderate change", and $0.25 \leqslant \text{PSI}$ means a "significant change". However, the statistical properties of these benchmarks are unknown. For example, how frequently will the PSI exceed 0.10 when in truth there has been no population shift? The asymptotic distribution of the PSI allows us to use benchmarks that control the type I error rate. Consider formally testing

$$\mathrm{H}_0: \qquad p_i* = q_i \quad i = 1, \dots, B$$

versus

$$\mathrm{H}_1: \qquad p_i \neq q_i \quad \text{for at least one } i.$$

A test with type I error rate α would reject H_0 and declare a change if

$$\text{PSI} > \left(\frac{1}{n} + \frac{1}{m}\right)\chi^2_{\alpha,B-1}, \tag{4.1}$$

where $\chi^2_{\alpha,B-1}$ is the upper α percentile of the χ^2 distribution with $B - 1$ degrees of freedom. Table 2 provides values of the benchmark (4.1) for $B = 10$ and $\alpha = 0.05$. The benchmarks change considerably with n and m.

When $B = 10$, the rule-of-thumb $\text{PSI} > 0.25$ seems reasonable for sample sizes n and m between 100 and 200, but it is too conservative for larger sample sizes. The benchmark also changes with the number of bins B. Table 3 provides the benchmark values (4.1) for $B = 20$ and $\alpha = 0.05$. Additional tables can be found in Yurdakul (2018).

Exactly how liberal or conservative are the 0.10 and 0.25 rules of thumb? We conducted a simulation study to assess rates of rejection under both the null and alternative hypotheses. We generated n observations from the $N(\mu_1 = 0, \sigma^2 = 64)$ and let $(\hat{p}_1, \dots, \hat{p}_{10})$ be the observed proportions that fall into the $B = 10$ population deciles. A second sample of m observations was generated from $N(\mu_2, \sigma^2 =$

TABLE 3 PSI benchmarks for $B = 20$ and $\alpha = 0.05$.

	m					
n	100	200	400	600	800	1000
100	0.603	0.452	0.377	0.352	0.339	0.332
200	0.452	0.301	0.226	0.201	0.188	0.181
400	0.377	0.226	0.151	0.126	0.113	0.106
600	0.352	0.201	0.126	0.100	0.088	0.080
800	0.339	0.188	0.113	0.088	0.075	0.068
1000	0.332	0.181	0.106	0.080	0.068	0.060

TABLE 4 Power comparison of PSI $> 0.10, 0.25$ and $(1/n + 1/m)\chi^2_{0.05}$ when $B = 10$ and the mean of target population has shifted by 0, $\frac{1}{4}$ and $\frac{1}{2}$ standard deviations.

Benchmark		$\mu_2 = \mu_1$			$\mu_2 = \mu_1 + (\sigma/4)$			$\mu_2 = \mu_1 + (\sigma/2)$		
m	n	0.10	0.25	χ^2	0.10	0.25	χ^2	0.10	0.25	χ^2
100	100	0.849	0.233	0.076	0.943	0.459	0.218	0.997	0.883	0.711
100	200	0.691	0.074	0.069	0.890	0.270	0.258	0.996	0.835	0.826
100	400	0.560	0.029	0.066	0.834	0.176	0.295	0.997	0.787	0.887
200	200	0.369	0.004	0.057	0.775	0.085	0.360	0.997	0.769	0.954
200	400	0.160	0.000	0.060	0.673	0.028	0.473	0.997	0.720	0.990
400	400	0.020	0.000	0.051	0.513	0.004	0.671	0.999	0.669	0.999

64) and we let $(\hat{q}_1, \dots, \hat{q}_{10})$ be the observed proportions falling into the deciles of $N(0, \sigma^2 = 64)$. We used three values – $\mu_2 = 0, 2, 4$ – representing mean shifts of 0, $\frac{1}{4}$ and $\frac{1}{2}$ standard deviations, respectively. The percentage of times that PSI was rejected (out of 10 000 runs) is presented in Table 4.

Under the null case where $\mu_2 = \mu_1$, the type I error rate for PSI > 0.10 is too large, except when both sample sizes exceed 400. In contrast, the type I error rate for PSI > 0.25 is too small when both sample sizes exceed 200, so the test is too conservative. The chi-square benchmark maintains a reasonable type I error, close to $\alpha = 0.05$.

Under the alternative cases where the two population means are not the same, the power of the PSI test based on the chi-square benchmark increases with the sample size. However, the tests PSI > 0.10 and PSI > 0.25 have powers that actually decrease with sample size (or type II error rates that increase with sample size). This counterintuitive behavior can be explained by Theorem 3.3, which states that PSI has expected value $(1/n+1/m)(B-1)$ and variance $2(1/n+1/m)^2(B-1)$. Both mean

TABLE 5 χ^2- and normal-based benchmarks for $m = n = 100$ and $\alpha = 0.05$.

	B			
	5	**10**	**15**	**20**
χ^2-based	0.19	0.34	0.47	0.60
z-based	0.17	0.32	0.450	0.58

and variance reduce to zero as n and m increase, which means that PSI converges to a point mass at zero.

Since the chi-square distribution is approximately normal for moderately large degrees of freedom, the percentile $\chi^2_{\alpha,B-1}$ in (4.1) may be approximated by $(B - 1) + z_\alpha\sqrt{2(B-1)}$. This results in

$$\text{PSI} > \left(\frac{1}{n} + \frac{1}{m}\right)(B - 1 + z_\alpha\sqrt{2(B-1)}). \tag{4.2}$$

Table 5 shows the values of chi-square and normal-based PSI benchmarks in (4.1) and (4.2).

5 CONCLUSION

The PSI is used to measure the divergence of frequency distributions between two samples or in time. In practice, it is used to measure the divergence between a development data set of a particular model and the current population the model is used on. It is used as a validation tool to show the applicability of the model to the current population. In its current use in industry, the PSI is compared with the rule-of-thumb benchmarks that have no known properties. In this paper, we provide guidance by deriving the asymptotic distribution of the PSI. Users may calculate benchmarks depending on the desired level of significance α. Besides α, the benchmarks depend on the number of bins B, and the base and target sample sizes n and m, respectively. Instead of the fixed rules of thumb 0.10 and 0.25, we suggest the following benchmarks for PSI, after deciding on a desired significance level α:

$$\text{PSI} > \chi^2_{\alpha,B-1} \times \left(\frac{1}{n} + \frac{1}{m}\right), \tag{5.1}$$

$$\text{PSI} > \left(\frac{1}{n} + \frac{1}{m}\right)(B - 1 + z_\alpha\sqrt{2(B-1)}). \tag{5.2}$$

The practitioner may choose $\alpha = 0.10, 0.05, 0.01$ or 0.005, depending on the acceptable level of risk the institution or practitioner assumes.

DECLARATION OF INTEREST

The authors report no conflicts of interest. The authors alone are responsible for the content and writing of the paper.

REFERENCES

Agresti, A. (2007). *An Introduction to Categorical Data Analysis*. Wiley-Interscience.

Anderson, R. (2007). *The Credit Scoring Toolkit: Theory and Practice for Retail Credit Risk Management and Decision Automation*. Oxford University Press.

Baesens, S. H., and Rosch, D. (2016). *Credit Risk Analytics*. SAS Institute/Wiley (https://doi.org/10.1002/9781119449560).

BGFRS–OCC (2011). Supervisory guidance on model risk management contents. SR Letter 11-7, Board of Governors of the Federal Reserve System and Office of the Comptroller of the Currency, Washington, DC. URL: https://www.federalreserve.gov/supervisionreg/srletters/sr1107a1.pdf.

Cochran, W. (1977). *Sampling Techniques*. Wiley.

Federal Deposit Insurance Corporation (2007). Scoring and modeling. In *Credit Card Activities Manual*. Federal Deposit Insurance Corporation, Arlington, VA. URL: https://www.fdic.gov/regulations/examinations/credit_card/pdf_version/ch8.pdf.

Gottschalk, S. (2016). Entropy and credit risk in highly correlated markets. Preprint (arXiv:1604.07042 [q-fin.RM]).

Kullback, S. (1978). *Information Theory and Statistics*. Peter Smith, Gloucester, MA.

Kullback, S., and Leibler, R. (1951). On information and sufficiency. *Annals of Mathematical Statistics* **22**(1), 79–86 (https://doi.org/10.1214/aoms/1177729694).

LendingClub (2018). LendingClub historical loan issuance data. URL: https://www.lendingclub.com/.

Lewis, E. M. (1994). *An Introduction to Credit Scoring*. Athena Press, London.

Lin, A. Z. (2017). Examining distributional shifts by using population stability index (PSI) for model validation and diagnosis. Working Paper. URL: https://www.lexjansen.com/wuss/2017/47_Final_Paper_PDF.pdf.

Liu, H., Parise, G. F., and Ware, L. R. (2009). Machine-based population stability. United States Patent 8326575, US Patent Office, Washington, DC.

Office of the Comptroller of the Currency (2016). Sensitivity to market risk installment lending. Comptroller's Handbook Collection, Office of the Comptroller of the Currency, Washington, DC. URL: https://bit.ly/389js6J.

Pruitt, R. (2010). The applied use of population stability index (PSI) in SAS Enterprise Miner. Poster, SAS Global Forum 2010. URL: http://support.sas.com/resources/papers/proceedings10/288-2010.pdf.

Siddiqi, N. (2017). *Intelligent Credit Scoring*. Wiley (https://doi.org/10.1002/9781119282396).

Stapleton, J. H. (ed) (1995). *Linear Statistical Models*. Wiley (https://doi.org/10.1002/9780470316924).

Thomas, L. C. (2009). *Consumer Credit Models: Pricing, Profit, and Portfolios*. Oxford University Press.

Thomas, L. C., Crook, J. N., and Edelman, D. B. (2017). *Credit Scoring and Its Applications*, 2nd edn. Society for Industrial and Applied Mathematics, Philadelphia, PA (https://doi.org/10.1137/1.9781611974560).

Wu, D., and Olson, D. (2010). Enterprise risk management: coping with model risk in a large bank. *Journal of the Operational Research Society* **61**(2), 179–190 (https://doi.org/10.1057/jors.2008.144).

Yurdakul, B. (2018). Statistical properties of population stability index. Graduate Dissertation, University of Michigan. URL: https://scholarworks.wmich.edu/dissertations/3208.