

Preparativos para Atender aos Requisitos de Modelagem do Novo Acordo da Basileia

Parte 2: Validação do Modelo

Este artigo, o segundo de uma série de quatro, debate algumas abordagens à validação de modelos estatísticos, tal como exigida pelo Novo Acordo de Capital.

Como foi dito no primeiro artigo desta série, segundo a atual escola de pensamento, no que concerne aos modelos de Probabilidade de Inadimplência (PI) e Perda em Caso de Inadimplência (PCI) a que se referem os Documentos de Consulta da Basileia, os bancos devem dispor de modelos separados para os devedores e para as linhas. O modelo do devedor deve prever a PI — normalmente inadimplência superior a 90 dias ou em execução, falência, baixa, busca e apreensão ou reestruturação. Os modelos de linha devem prever a perda em caso de inadimplência (PCI), ou 1 menos a taxa de recuperação.

No artigo anterior, a *regressão logística* foi o método recomendado para construir modelos de PI. Essa técnica estatística emprega um conjunto de variáveis cujos valores atuais podem prever a pro-

Jeffrey S. Morrison

Preparing for Basel II Modeling Requirements

Part 2: Model Validation

This article, the second in a four-part series, discusses some approaches to the validation of statistical models as required by the new Capital Accord.

As mentioned in the first article in this series, the current school of thought surrounding the probability of default (PD) and loss given default (LGD) models mentioned in Basel Consultative Papers is that banks should have separate models for the obligor and the facility. The obligor model should predict the PD — usually 90-plus days delinquent or in foreclosure, bankruptcy, charge-off, repossession, or restructuring. Models on the facility side should predict the loss given default (LGD) or 1 minus the recovery rate.

In the initial article, logistic regression was the approach recommended for building PD models. This statistical technique uses a

set of explanatory variables whose values today would hopefully predict a loan's probability of default sometime over the next 12 months. On the LGD side, the approach recommended was to use either linear regression or tobit regression to estimate the model.

Directives from Basel II

Paramount to using the advanced approach as specified in the Basel II Capital Accord is a focus on model validation. The New Basel Capital Accord, published January 2001, includes the following:

302. Banks must have a robust system in place to validate the accuracy and consistency of rating systems, processes, and the estimation of PDs. A bank must demonstrate to its supervisor that the internal validation process enables it to assess the performance of internal rating and risk quantification systems consistently and meaningfully.

305. The process cycle of model validation must also include:

- > Ongoing periodic monitoring of model performance, including evaluation and rigorous statistical testing of the dynamic stability of the model and its key coefficients;
- > Identifying and documenting fixed relationships in the model that are no longer appropriate;
- > Periodic testing of model outputs on an

habilidade de inadimplência de um empréstimo nos próximos 12 meses. Quanto à PCI, a abordagem recomendada foi o uso de *regressão linear* ou *regressão tobit* para estimativa do modelo.

Diretrizes do Novo Acordo da Basileia

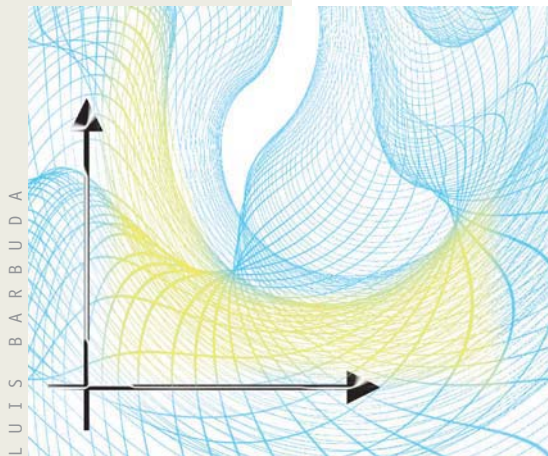
É fundamental, para o uso da abordagem avançada tal como especificada no Novo Acordo da Basileia, manter o foco sobre a validação dos modelos. O Novo Acordo, publicado em janeiro de 2001, inclui as seguintes disposições:

302. Os bancos devem possuir um sistema robusto para validar a precisão e a coerência dos sistemas e processos de *rating* e da estimativa de PIs. Cada instituição necessita demonstrar aos seus supervisores que o processo interno

de validação possibilita estimar o desempenho dos sistemas internos de *rating* e quantificação de risco de forma consistente e significativa.

305. O ciclo do processo de validação do modelo deve incluir, ainda:

- > Monitoramento constante do desempenho do modelo, inclusive avaliação e rigorosos testes estatísticos da estabilidade dinâmica do modelo e de seus principais coeficientes;
- > Identificação e documentação de relações fixas do modelo que não sejam mais apropriadas;
- > Teste periódico dos resultados do modelo, a intervalos no mínimo anuais; e
- > Um rigoroso processo de controle de mudan-



ças, estipulando os procedimentos a serem seguidos antes de se realizar qualquer alteração do modelo em resposta aos resultados da validação.

Até agora, o *Novo Acordo de Capital da Basileia* não indica especificações ou padrões, quanto ao processo de validação.

Introdução à Validação

A validação inclui questões ligadas à qualidade dos dados, à documentação, à análise de sensibilidade, à especificação do modelo, à concepção das amostras, ao desempenho em testes estatísticos e ao desenvolvimento de medidas de precisão do modelo. Sem subestimar a importância das demais áreas, por causa das limitações de espaço, o presente artigo se concentrará na quantificação de medidas de precisão. Assim, a validação do modelo refere-se simplesmente à verificação da precisão do modelo de um banco, por um determinado período de tempo. Quantos empréstimos entraram efetivamente em inadimplência durante o ano e quais eram suas probabilidades previstas de inadimplência? Se a maioria de suas inadimplências tinha probabilidade prevista próxima a 10%, é possível que seu modelo não tivesse um bom desempenho.

A validação não é apenas parte das exigências do Novo Acordo, como também é fundamental para qualquer processo de desenvolvimento de mode-

annual basis, at a minimum; and

> A rigorous change control process, which stipulates the procedures that must be followed prior to making changes in the model in response to validation outcomes.

As of yet, The New Basel Capital Accord does not give specifics or standards related to the validation process.

*É fundamental
manter o foco na
validação do
modelo.*

*To conserve
a focus
on model
validation is
paramount.*

Introduction to Validations

Validation includes issues of data quality, documentation, sensitivity analysis, model specification, sample design, the performance of statistical tests, and the development of measures for model accuracy. Although not minimizing the importance of these other areas, for brevity's sake the remainder of this article will focus on quantifying accuracy measures. In this light, model validation simply refers to checking the accuracy of your

model over some specific period of time. How many loans actually went into default during the year and what did their predicted default probabilities look like? If most of your defaults had a predicted probability of default near 10%, then your model may be doing a poor job.

Not only is the validation part of Basel requirements, it is central to any model

development process, regardless of its application. Even econometric forecasting models – models developed using aggregate data with economic time series – are validated for accuracy. Credit-scoring models are also validated for accuracy. Because modeling is indeed an art, statistical algorithms are developed and redeveloped until a formulation is found that reflects the most accurate results and makes the most business sense.

Validations can be done in a variety of ways, ranging from the simple to the complex:

1. Performing the validation only on your model development sample.

2. Performing the validation on a sample of accounts that were not used to develop the model, but were taken from the same period of time.

3. Performing the validation on a single holdout sample from time periods outside your model development window.

4. Performing a step-through simulation process across multiple time periods while recalibrating the model.

The first approach is the most straightforward and is typically performed as the model is developed. Here, the same data that was used for estimating the model is used for validation. Although this type of validation tends to overstate the model's predictive ability, it may be necessary if there are a limited number of defaults available for model building purposes.

If there are sufficient defaults available, the second method is preferred. A random

*los, independentemente da aplicação. Até mesmo os modelos econométricos preditivos — modelos desenvolvidos com o uso de dados agregados por meio de séries temporais econômicas — são submetidos à validação de precisão. Também se valida a precisão dos modelos de *credit scoring*. Como a modelagem é, antes de tudo, uma arte, os algoritmos estatísticos são desenvolvidos até que se encontre uma formulação que proporcione os resultados mais precisos e seja o mais lógico sob o ponto de vista dos negócios.*

As validações podem ser realizadas de diversas maneiras, partindo das mais simples até chegar às mais complexas:

1. Realizar a validação apenas na amostra de desenvolvimento do modelo.

2. Realizar a validação numa amostra de contas que não entrou no desenvolvimento do modelo, mas foi colhida no mesmo intervalo de tempo.

3. Realizar a validação em uma única amostra de reserva de intervalos de tempo diferentes da janela de desenvolvimento do modelo.

4. Realizar um processo de simulação por etapas, em diferentes intervalos de tempo ao mesmo tempo em que se calibra o modelo.

A primeira abordagem é a mais direta e costuma ser empregada durante o desenvolvimento do modelo. Nesse caso, os mesmos dados que foram utilizados para estimar o modelo são empregados para fins de validação. Embora esse tipo de validação tenda a superestimar a capacidade de predição do modelo, isso pode ser necessário se a quantidade de inadimplências disponíveis para construir o modelo for limitada.

Havendo disponibilidade suficiente de inadimplências, dá-se preferência ao segundo modelo. Re-

serva-se uma amostra aleatória de dados durante a estimativa do modelo; o segundo método adota os dados reservados pelo modelo e calcula seus valores previstos para fins de validação. Esse método é geralmente utilizado para validar uma grande variedade de modelos e dá suporte aos estatísticos, no que se refere à escolha do melhor modelo.

Os métodos restantes são mais avançados — não porque suas técnicas sejam necessariamente mais complicadas, mas por requerer um histórico mais aprofundado das inadimplências. A terceira abordagem reserva dados de períodos anteriores para fins de validação, a fim de determinar se o nível de precisão se mantém de ano para ano. Isso serve como indicação do grau de estabilidade do modelo ao longo do tempo. A quarta abordagem é uma combinação de validações e calibrações do modelo. Ela pretende simular o desenvolvimento do modelo e de sua capacidade de prever ao longo do tempo, uma vez que os modelos costumam ser revistos anualmente, à medida que novas inadimplências são acumuladas e acrescentadas ao processo.

Validação do Modelo do Devedor: Probabilidade de Inadimplência

Vamos admitir que se utilize o segundo método de validação para verificar a precisão de um modelo de PI e se reserve uma amostra de contas que não foram utilizadas no desenvolvimento do modelo. Quantas inadimplências serão necessárias? De maneira geral, é preciso dispor de centenas de inadimplências para testar corretamente o modelo — quanto maior o número, melhor. Como há mais inadimplências disponíveis

sample of data is held out from the model estimation; the second method runs the hold-out data against the model to compute its predicted values for validation purposes. This method is widely used for validating a variety of different models and serves as an aid to the statistician in selecting the best model.

The remaining methods are more advanced — not because their techniques are necessarily more complicated but because they require a greater depth of default history. The third approach holds out data for validation from prior periods to see if the level of accuracy remains the same from year to year. This is an indication of how stable your model may be over time. The fourth approach is a combination of validations and model recalibrations. The idea is to simulate model development and its predictiveness over time given that model revisions are done annually as new defaults are accumulated and added to the process.

Validating the Obligor Model: Probability of Default

Let's assume you use the second validation method to evaluate the accuracy of a PD model and hold out a sample of accounts that were not used in model development. So how many defaults do you need? Generally speaking, hundreds of defaults are necessary to properly test the model — the more, the better. Since more defaults are available in credit card portfolios, thousands of defaults are typically used in validations.

To perform a validation, you need predicted default probabilities from your model and a default status indicator showing whether the account defaulted or not. With this information handy, you can easily calculate two measures of model accuracy where all you need to be able to do is sort and add. Following is a step-by-step guide for PD validation.

Step 1: Create your holdout sample, if available.

Step 2: Code your default indicator. Since you built your model with a default indicator of "1" if the loan defaulted and "0" otherwise, make sure your default indicator in your holdout sample is coded the same way.

Step 3: Sort your holdout sample. Sort the data from highest to lowest, based on the probability of default. If your data set is small enough, you could even do this in Excel.

Step 4: Record minimum and maximum probabilities in each 5% bucket.

Step 5: Add some numbers together. Now start totaling the data into buckets at 5% intervals from the top down. Produce the following columns by bucket:

> Number of defaults.

> Number of nondefaults.

> Step 6: Calculate cumulative percentages by bucket:

> Cumulative number of defaults.

> Cumulative number of nondefaults.

Figure 1 shows this process for a fictional model in which the number of defaults was totaled into 20 buckets, each representing

nas carteiras de cartões de crédito, é comum usar milhares de inadimplências nas validações.

Para realizar uma validação, é preciso ter as probabilidades de inadimplência previstas pelo modelo e um indicador de estado de inadimplência para apontar se a conta está ou não inadimplente. Com essas informações em mãos, é fácil calcular duas medidas de precisão do modelo, bastando, para tanto, saber classificar e somar. Segue um guia passo a passo de validação de PI.

Passo 1: Crie sua amostra de reserva, caso haja.

Passo 2: Codifique o indicador de inadimplência. Como o modelo foi construído com indicador "1" em caso de inadimplência e indicador "0" ao contrário, certifique-se de que o indicador de inadimplência da amostra de reserva esteja codificado da mesma maneira.

Passo 3: Classifique a amostra de reserva. Organize os dados de cima para baixo, com base na probabilidade de inadimplência. Se o conjunto de dados não for muito grande, isso pode até ser realizado em Excel.

Passo 4: Registre as probabilidades mínima e máxima em cada intervalo de 5%.

Passo 5: Comece a somar os valores. Inicie a totalização dos dados em intervalos de 5% de cima para baixo. Produza, para cada intervalo, as seguintes colunas:

> Número de inadimplências.

> Número de adimplências.

> Passo 6: Calcule as porcentagens acumuladas por intervalo:

> Número acumulado de inadimplências.

> Número acumulado de adimplências.

A Figura 1 representa o processo para um modelo fictício em que o número de inadimplências

foi totalizado em 20 intervalos, cada um representando cerca de 5% das contas. Quando esse procedimento é aplicado a um modelo preciso, a maioria dos inadimplentes deve se concentrar nos primeiros intervalos. Da mesma forma, a maioria dos adimplentes deve ficar nos intervalos inferiores. Isso indica a capacidade que o modelo tem de distinguir entre inadimplentes e adimplentes.

Passo 7: Identificar as porcentagens de inadimplentes e adimplentes nos dois intervalos

about 5% of the accounts. When this procedure is applied to an accurate model, the majority of defaulters should be accumulated in the earlier buckets. Likewise, the nondefaulters should be found toward the bottom buckets. This shows us the power the model has in distinguishing between defaulters and nondefaulters.

Step 7: Identify percentages of defaulters and nondefaulters in the top two buckets. Note that the column F labeled "Cumulative % Defaults" in-

Figura 1

A	B	C	D	E	F	G	H
Intervalo de 5%	Probabilidade Mínima	Probabilidade Máxima	Número de Inadimplências	Número de adimplências	Porcentagem Acumulada de Inadimplências	Porcentagem Acumulada de Adimplências	Diferença entre as Porcentagens. Acumuladas
1	0,987637	0,998625	33	4	10,1852	0,9390	9,2
2	0,961737	0,987637	35	3	20,9877	1,6432	19,3
3	0,932949	0,961737	36	1	32,0988	1,8779	30,2
4	0,897801	0,932949	36	2	43,2099	2,3474	40,9
5	0,821660	0,897801	30	7	52,4691	3,9906	48,5
6	0,813351	0,821660	38	0	64,1975	3,9906	60,2
7	0,545989	0,813351	28	9	72,8395	6,0133	66,7
8	0,530848	0,545989	24	14	80,2469	9,3897	70,9
9	0,398955	0,519828	22	15	87,0370	12,9108	74,1
10	0,314351	0,398955	8	30	89,5062	19,9531	69,6
11	0,192926	0,295930	5	33	91,0494	27,6995	63,3
12	0,132092	0,192926	3	34	91,9753	35,6808	56,3
13	0,117058	0,132092	5	33	93,5185	43,4272	50,1
14	0,099934	0,117058	0	37	93,5185	51,1127	41,4
15	0,099934	0,099934	0	38	93,5185	61,0329	32,5
16	0,090868	0,099934	7	30	95,6790	68,0751	27,6
17	0,090868	0,090868	0	38	95,6790	76,9953	18,7
18	0,069882	0,090868	12	25	99,3827	82,8638	16,5
19	0,045697	0,057391	2	36	100,0000	91,3146	8,7
20	0,036294	0,045697	0	37	100,0000	100,0000	0

Figure 1

A	B	C	D	E	F	G	H
5% Bucket	Min Probability	Max Probability	# Defaults	# of Non Defaults	Cumulative % Defaults	Cumulative % Nondefaults	Difference in % Cumulatives
1	0.987637	0.998625	33	4	10.1852	0.9390	9.2
2	0.961737	0.987637	35	3	20.9877	1.6432	19.3
3	0.932949	0.961737	36	1	32.0988	1.8779	30.2
4	0.897801	0.932949	36	2	43.2099	2.3474	40.9
5	0.821660	0.897801	30	7	52.4691	3.9906	48.5
6	0.813351	0.821660	38	0	64.1975	3.9906	60.2
7	0.545989	0.813351	28	9	72.8395	6.0133	66.7
8	0.530848	0.545989	24	14	80.2469	9.3897	70.9
9	0.398955	0.519828	22	15	87.0370	12.9108	74.1
10	0.314351	0.398955	8	30	89.5062	19.9531	69.6
11	0.192926	0.295930	5	33	91.0494	27.6995	63.3
12	0.132092	0.192926	3	34	91.9753	35.6808	56.3
13	0.117058	0.132092	5	33	93.5185	43.4272	50.1
14	0.099934	0.117058	0	37	93.5185	51.1127	41.4
15	0.099934	0.099934	0	38	93.5185	61.0329	32.5
16	0.090868	0.099934	7	30	95.6790	68.0751	27.6
17	0.090868	0.090868	0	38	95.6790	76.9953	18.7
18	0.069882	0.090868	12	25	99.3827	82.8638	16.5
19	0.045697	0.057391	2	36	100.0000	91.3146	8.7
20	0.036294	0.045697	0	37	100.0000	100.0000	0

dicates that 20.98% of the total defaulters in the holdout sample were identified in the top 10% (two buckets) of the sorted list. The bigger these numbers, the better. Note only 1.64% of the nondefaulters were found, as shown in column G labeled “Cumulative % Nondefaults.” These measures, reflecting the accuracy of the model for the top 10% of the data, can serve as an excellent way to validate competing models.

Step 8: Calculate the K-S value. Another measure of accuracy can be computed from columns F and G. This value, called K-S, is simply the maximum difference between these two columns of numbers. The K-S value can range between 0 and 100, with 100 implying the model does a perfect job in predicting defaults or separating the two populations. In general, the higher the K-S, the better the model. The place where the maximum occurs is that point of maximum separation. As shown in Column H, the K-S value in this example is 74.1 and occurs in the 9th bucket.

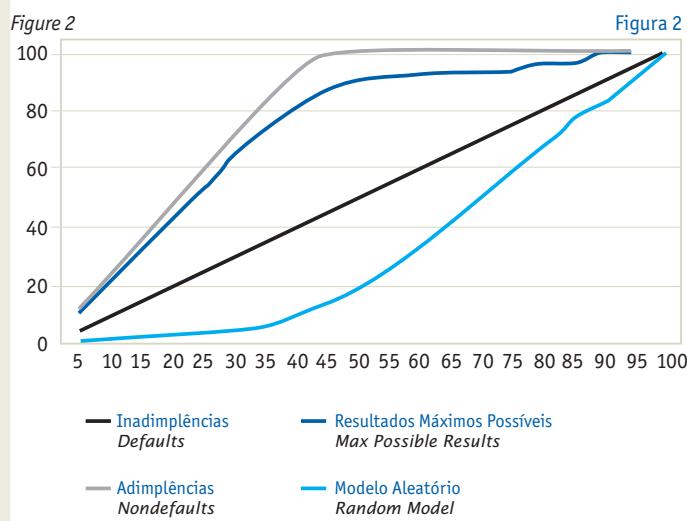
Step 9: Produce a graph. This is done by simply graphing columns F and G. A graphi-

superiores. Observe que a coluna F, intitulada “Porcentagem Acumulada de Inadimplências”, indica que 20,98% do total de inadimplentes da amostra de reserva foram identificados nos 10% (dois intervalos) mais elevados da lista classificada. Quanto mais elevado esse valor, melhor. Observe que apenas 1,64% dos adimplentes foram ali colocados, como mostra a coluna G, intitulada “Porcentagem Acumulada de Adimplências”. Essas medidas, que refletem a precisão do modelo para os primeiros 10% dos dados, podem ser uma excelente maneira de validar modelos concorrentes.

Passo 8: Calcule o valor K-S. Outra medida de precisão pode ser calculada a partir das colunas F e G. Esse valor, chamado de K-S, é simplesmente a mínima diferença entre essas duas colunas de números. O valor K-S pode variar entre 0 e 100, sendo que 100 implica um modelo perfeito na previsão de inadimplências ou na separação das duas populações. De maneira geral, quanto maior o K-S, melhor o modelo. O ponto de ocorrência do máximo é o ponto de maior separação. Como se vê na Coluna H, o valor K-S desse exemplo é de 74,1 e ocorre no 9º intervalo.

Passo 9: Produza um gráfico. Isso é feito simplesmente traçando um gráfico das colunas F e G. Uma representação gráfica dessa tabela, como mostra a Figura 2, é conhecida por diversos nomes, como, por exemplo, curva de potência ou gráfico de elevação.

O eixo vertical é a porcentagem acumulada de inadimplentes ou adimplentes contados ou identificados. O eixo horizontal reflete o ponto em que se



está na lista classificada. Em outras palavras, um valor de 30 no eixo horizontal significa que foram examinados os 30% superiores dos dados de validação. A linha cinzenta do topo reflete os resultados de um modelo teoricamente perfeito que preveja corretamente todas as inadimplências, ou seja, o máximo a que se pode chegar. A linha azul escuro mostra a porcentagem acumulada de inadimplentes com base no modelo estimado (coluna F), enquanto a linha azul mais clara mostra a porcentagem acumulada de não inadimplentes (coluna G).

A linha negra do centro reflete um modelo simplório que identifica os inadimplentes de maneira simplesmente aleatória. Em outras palavras, esse “modelo” aleatório não tem qualquer conteúdo em termos de informação de predição. É por vezes usado como *benchmark* para fins de comparação entre modelos concorrentes.

Para ter qualquer valor, um modelo estatístico deve apresentar desempenho melhor do que uma adivinhação aleatória sobre quem viria a se tornar inadimplente. O valor K-S de 74,1 é a distância vertical entre as linhas de cor azul escuro e azul claro.

Embora essas abordagens sejam frequentemente utilizadas para avaliar a precisão associada a modelos de inadimplência, há outras que também podem ser úteis.

cal depiction of this table, as seen in Figure 2, goes by a variety of names, such as power curve or lift chart.

The vertical axis is the cumulative percentage of defaulters or nondefaulters counted or identified. The horizontal axis reflects how far down the sorted list you are. In other words, a

value of 30 on the horizontal axis means that you have examined the upper 30% of the validation data. The grey line at the top reflects the results of a theoretically perfect model that correctly predicts all the defaults. That’s the best you can do. The dark blue line shows the cumulative percentage of defaulters based on your estimated model (column F) while the lighter blue line shows the cumulative percentage of nondefaulters (column G).

The black line in the middle reflects a naïve model that identifies defaulters simply at random. In other words, this random “model” has no predictive information content. It is sometimes used as a benchmark when comparing competing models. For a statistical model to have any value at all, it must perform better than a random guess at who would default. The K-S value of 74.1 is the vertical distance between the dark blue and lighter blue lines.

*Índices de entropia
informacional
são medidas de
precisão.*

*Information
entropy ratios
are accuracy
measures.*

Although these approaches are commonly used in evaluating accuracy associated with default models, other approaches may also be helpful.

Classification errors. Assuming some probability cutoff value, an account can be classified into one of two buckets – defaults or nondefaults. For example, if the estimated probability of default is greater than or equal to, say, 50%, the account might be assumed to default. If the account's estimated probability is less than 50%, then we have a nondefault. By comparing these classification results to our historical data, we can determine an overall classification error rate as well as the number of false positives (Type I error) and false negatives (Type II error). A false positive is mistakenly predicting an account will default when it actually was a nondefault.

Likewise, a false negative is mistakenly predicting an account will not default when it, in fact, actually did. The distribution of false positives and false negatives produced by the model can have substantial cost implications when applied to an entire portfolio.

Information entropy ratios. These are accuracy measures where two states of uncer-

Erros de classificação. Admitindo um valor de corte de qualquer probabilidade, uma conta pode ser classificada em um de dois grupos — inadimplências ou adimplências. Por exemplo, se a probabilidade estimada de inadimplência for maior ou igual a 50%, digamos, pode-se presumir que ela se tornará inadimplente. Se sua

probabilidade estimada for inferior a 50%, então temos uma adimplência. Comparando esses resultados de classificação com nossos dados históricos, podemos determinar uma taxa geral de erro de classificação, sem falar no número de falsos positivos (erro do Tipo I) e falsos negativos (erro do Tipo II).

Um falso positivo surge quando se prevê erroneamente que uma conta se torne inadimplente quando, na verdade, revelou-se não inadimplente. Da mesma forma, um falso negativo surge quando se prevê erroneamente que uma conta não se tornará inadimplente quando, na verdade, isso ocorreu. A distribuição de falsos

positivos e falsos negativos produzida pelo modelo pode ter implicações substanciais quanto aos custos, quando aplicada à carteira como um todo.

Índices de entropia informacional. São medidas de precisão que comparam dois estados de incerteza. Na segunda nota, no final do artigo, há maiores detalhes.

Hundreds of defaults are necessary to test the model.

Centenas de inadimplências são necessárias para testar o modelo.

Validação dos Modelos de Linha: PCI

O processo de validação do modelo de PCI é um pouco diferente do usado para o modelo de PI. No modelo de PI há duas subpopulações — inadimplentes e adimplentes. No modelo de PCI, só são usadas as informações de inadimplentes. São necessários três itens para validar os modelos de PCI: (1) recuperação efetiva em dólares, (2) inadimplência em dólares e (3) taxa prevista de recuperação. Eis um guia passo a passo para validar seu modelo de PCI:

1. Escolha o intervalo de tempo da validação.
2. Reuna os dados de reserva, caso haja.
3. Calcule a taxa efetiva de recuperação, com base na inadimplência em dólares e no total recuperado em dólares.
4. Calcule o Quadrado do Erro Médio (*Mean Squared Error* – MSE). Essa será a medida de precisão da validação:

$$MSE = S(\% \text{ Efetiva} - \% \text{ Prevista})^2 / N - 1$$

Onde N representa o número de observações da amostra de validação. Trata-se, simplesmente, do quadrado da diferença entre as taxas de perda efetiva e previstas somadas e divididas pelo tamanho da amostra menos 1. Quanto menor o MSE, mais preciso o modelo, em igualdade das demais condições. Também podem ser calculadas outras medidas de precisão, como a Raiz Quadrada do Erro Médio (*root mean squared error* – RMSE), desvio absoluto médio (*mean absolute deviation* – MAD) e erro percentual absoluto médio (*mean absolute percent error* – MAPE) — cujas fórmulas

are compared. For further details, see the second footnote at the end of this article.

Validating Facility Models: LGD

The validation process for the LGD model is a little different from that for the PD model. In the PD model, there are two subpopulations — defaulters and nondefaulters. In the LGD model, only defaulted information is used. For validating LGD models, three items are required: (1) actual dollars recovered, (2) defaulted dollars, and (3) predicted recovery rate. Here is a step-by-step guideline for validating your LGD model:

1. Select the period of time for validation.
2. Put together holdout data if available.
3. Compute actual recovery rate from defaulted dollars and total dollars recovered.

4. Compute Mean Squared Error (MSE). This is your measure of validation accuracy:

$$MSE = S(\text{Actual}\% - \text{Predicted}\%)^2 / N - 1$$

Where N represents the number of observations in the validation sample. This is just the squared difference between the actual and predicted loss rates, summed up and then divided by the sample size, less one. The lower the MSE, the more accurate the model, all other things remaining equal. Additional measures of accuracy can also be computed, such as root mean squared error (RMSE), mean absolute deviation (MAD), and mean absolute percent error (MAPE) — formulas for

which can be found in any forecasting textbook. In the fictional example shown as Figure 3, LGD validation data is shown for 10 loans with a recovery rate ranging from 0 to 100% (Columns A, B, and C will have a minimum value of 0% and a maximum value of 100%). Column D is column C minus column B, then we square that answer for each loan. Then we sum column D and divide by 9. Loss given default (LGD, column B) is equal to one (100%) minus the recovery rate.

Figure 3

Mean Squared Error

Actual Recovery Rate	Actual LGD Rate	Predicted LGD Rate	MSE Calculations
A	B	C	D
32.4	67.6	55	158.76
54.5	45.5	33.3	148.84
87.3	12.7	66.4	2883.69
22.3	77.7	17.8	3588.01
43.4	56.6	33.2	547.56
0	100	76.9	533.61
0	100	65.8	1169.64
1.7	98.3	12.7	7327.36
3.8	96.2	95.1	1.21
0	100	87.7	151.29
MSE =			1834.44

Summary

From a validation perspective, how good is good enough? Unfortunately, there is no magic answer. On the obligor side, a validation model using a payment history and delinquency information could produce K-S values of 50 or more. At a bare minimum, it should certainly perform better than predicting defaults randomly. As research has shown, it's hard to make LGD accurate because of the difficulty in obtaining good predictive explanatory variables and the wide variation in col-

las podem ser encontradas em qualquer livro-texto de previsão. No exemplo fictício da Figura 3, são mostrados dados de validação de PCI de 10 empréstimos com taxa de recuperação entre 0 e 100% (as colunas A, B e C terão valor mínimo de 0% e máximo de 100%). A coluna D é o valor da coluna C menos o da coluna B, elevado ao quadrado, somado e dividido por 9. A perda em caso de inadimplência (PCI, coluna B) é igual a 1 (100%) menos a taxa de recuperação.

Figura 3

Quadrado do Erro Médio

Taxa efetiva de recuperação	Taxa Efetiva PCI	Taxa Prevista de PCI	Cálculo do MSE
A	B	C	D
32,4	67,6	55	158,76
54,5	45,5	33,3	148,84
87,3	12,7	66,4	2883,69
22,3	77,7	17,8	3588,01
43,4	56,6	33,2	547,56
0	100	76,9	533,61
0	100	65,8	1169,64
1,7	98,3	12,7	7327,36
3,8	96,2	95,1	1,21
0	100	87,7	151,29
MSE =			1834,44

Sumário

Do ponto de vista da validação, o que é suficientemente recomendável? Infelizmente, não há uma resposta definitiva. Da parte do devedor, um modelo de validação que use histórico de pagamentos e informações sobre inadimplência deve produzir valores de K-S de 50 ou mais. No mínimo, deve ter desempenho de predição melhor do que um modelo aleatório. Como demonstram as pesquisas, é difícil chegar à precisão de PCI por causa da dificuldade de obtenção de boas variáveis explicativas e da ampla varia-

ção dos esforços de cobrança dependendo do tipo de garantia. De maneira geral, é preciso lembrar, no que se refere à PI, quanto maior o K-S, melhor. Mas, no caso da PCI, vence quem tiver o menor MSE.

Assim sendo, com que frequência é preciso submeter o modelo a um novo desenvolvimento? Isso depende da disponibilidade das informações de inadimplência referentes ao modelo atual e da estabilidade de sua carteira. Se só houver um número marginal de inadimplências disponível para o desenvolvimento do modelo, maiores informações sobre elas poderiam aprimorar a precisão da validação em grau suficiente para justificar um novo desenvolvimento. Por outro lado, se as características subjacentes de sua carteira mudarem a cada ano, então essa seria a periodicidade recomendada dos novos desenvolvimentos do modelo. Como o Acordo da Basileia exige, de uma maneira ou de outra, a produção periódica de estudos de validação, uma abordagem sábia poderia avaliar novos modelos concorrentes a cada ano e comparar seus resultados de validação com os do modelo em vigor — e que vença o melhor.

O Novo Acordo da Basileia menciona palavras como “processo” ou “sistema” cerca de 275 vezes em suas 139 páginas. Isso implica ênfase a uma abordagem sistemática aos aspectos analíticos dos modelos — uma abordagem que integra os requisitos de programação de diversas fontes a métodos e procedimentos padronizados. O terceiro artigo desta série apresentará o desenvolvimento, pelo SunTrust, de uma plataforma analítica que integra esses requisitos numa interface assemelhada ao Windows, permitindo desenvolver sofisticados modelos, validá-los e documentá-

lection efforts across collaterals. In general, remember that on the PD side, the higher the K-S, the better, but as for LGD, the one with the lowest MSE wins.

So, how often should your model be redeveloped? That depends on the availability of default information on your current model and the stability of your portfolio. If you had only a marginal number of defaults available for model development, then additional default information could improve validation accuracy enough to prompt redevelopment. On the other hand, if the underlying characteristics of your portfolio change from year to year, then model redevelopment would be recommended. Since it is required by Base l to produce validation studies on a periodic basis anyway, perhaps a wise approach might be to evaluate newer competing models each year and compare their validation results to your existing model — winner takes all.

The New Basel Capital Accord references words like process or system about 275 times in its 139 pages. This implies a special emphasis on the need to develop a systematic approach to model analytics — an approach that integrates programming requirements from a variety of sources with standardized methods and procedures. The third article in this series will present the development of an analytics platform by SunTrust that integrates these requirements into a Windows-like interface, allowing sophisticated statistical models to be

developed, validated, and documented quickly and efficiently.

Notes

1 *Credit Risk Modeling – Design and Applications*, edited by Elizabeth Mays, 1998, Glenlake Publishing Company.

2 SOBEHART, KEENAN and STEI, “Benchmarking Quantitative Default Risk Models: A Validation Methodology,” *Moody’s Investor Services Global Credit Research*, March 2000.

©2003 RMA. Jeff Morrison is vice president, Credit Metrics – PRISM Team, at SunTrust Banks Inc., Atlanta, Georgia. Contact Morrison at Jeff.Morrison@suntrust.com

RMA - Risk Management Association is an international association of financial services professionals. For membership information, e-mail acauley@rmahq.org; to subscribe to *The RMA Journal*, visit www.rmahq.org/Ed_Opps/pubs/journalad.htm

los com agilidade e eficiência.

Notas

1 *Credit Risk Modeling – Design and Applications*, editado por Elizabeth Mays, 1998, Glenlake Publishing Company.

2 SOBEHART, KEENAN and STEI, “Benchmarking Quantitative Inadimplência Risk Models: A Validation Methodology,” *Moody’s Investor Services Global Credit Research*, março de 2000.

©2003 RMA. Jeff Morrison é vice-presidente da Credit Metrics – Equipe PRISM, do SunTrust Banks Inc., Atlanta, Georgia. Os contatos com Morrison podem ser feitos pelo endereço eletrônico Jeff.Morrison@suntrust.com

A RMA - Risk Management Association é uma associação internacional de serviços financeiros profissionais. Para informações, e-mail acauley@rmahq.org; Para assinar *The RMA Journal* visite o site www.rmahq.org/Ed_Opps/pubs/journalad.htm