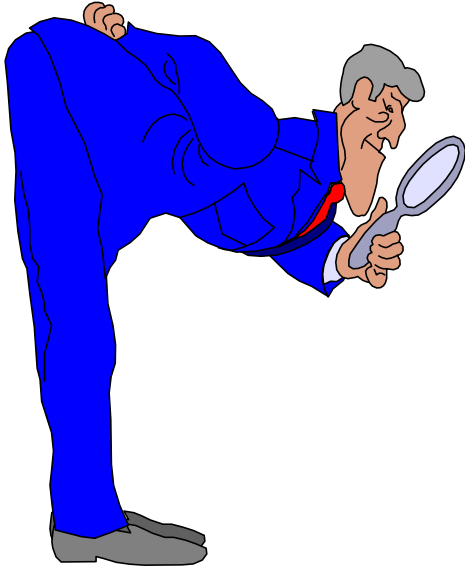


# O QUE É E COMO FUNCIONA O CREDIT SCORING – PARTE I



- *A utilização de escores na avaliação de crédito*
- *Como montar um plano de amostragem para o credit scoring?*
- *Como escolher as variáveis no modelo de credit scoring?*
- *Como classificar os indivíduos em uma análise de crédito através do credit scoring?*

**Autor: Leopoldo Costa Junior (6961406@fgvsp.br)**

- Graduado em Administração de Empresas pela Escola de Administração de Empresas de São Paulo – FGV.
- Mestrando em Economia de Empresas pela Escola de Administração de Empresas de São Paulo – FGV.
- Atualmente é consultor financeiro na área de formulação de estratégia para reequacionamento financeiro, projeção de cenários futuros em modelo de simulação e análise da situação financeira e de mercado da empresa.

## ÍNDICE

	PÁG
◆ APRESENTAÇÃO	3
◆ O QUE É CREDIT SCORING?	4
◆ A UTILIZAÇÃO DE ESCORES	5
◆ COMO MONTAR O PLANO DE AMOSTRAGEM	7
◆ COMO CLASSIFICAR OS INDIVÍDUOS EM UMA ANÁLISE DE CRÉDITO	11

## APRESENTAÇÃO

A utilização de escores é uma prática muito comum hoje em dia.

As instituições financeiras utilizam escores para avaliar os seus clientes e, eventualmente, concedem benefícios em função da pontuação obtida. Os analistas de mercado também podem utilizar escores para avaliar o desempenho de empresas.

Dentro dessas aplicações, vamos ver como funciona, em linhas gerais, a utilização de escores na avaliação de crédito em massa, o chamado **credit scoring**.

A nossa intenção neste **Up-To-Date®** é dar uma boa idéia de como funciona e como se monta e utiliza o escore de crédito.

Evidentemente, a implantação e definição de um escore de crédito deve ser deixada a cargo de um especialista.

## O QUE É CREDIT SCORING?

A concessão de crédito em massa via CDC (Crédito Direto ao Consumidor) é uma das alternativas possíveis para se alavancar um negócio permitindo a realização de vendas que não ocorreriam porque o consumidor não dispõe dos recursos necessários para adquirir o bem.

Um dos inconvenientes de se conceder crédito é que o risco de não recebimento é maior do que no pagamento à vista.

Se a análise de crédito não for bem feita, o lucro obtido com uma venda adicional pode ser totalmente comprometido com as despesas de cobrança de um mau pagador ou com a perda do crédito.

Neste contexto, uma das técnicas usualmente empregadas na análise de crédito é justamente o **credit scoring**.

O **credit scoring** é uma das aplicações mais conhecidas da análise discriminante para a área financeira.

Análise discriminante é um conjunto de técnicas estatísticas que permitem obter uma regra para classificar indivíduos em um conjunto de grupos previamente determinados.

No caso da concessão de crédito em massa, um escore de crédito pretende determinar se um certo cliente é merecedor de crédito a partir da análise de certos dados do seu cadastro (sexo, idade, estado civil, renda etc.).

## A UTILIZAÇÃO DE ESCORES

Dentro do processo de análise de crédito, o score de crédito é normalmente empregado após a análise inicial do cadastro de cliente. Nessa análise inicial são eliminados os clientes que:

- ✓ Tenham os nomes negativados em algum serviço de proteção ao crédito;
- ✓ Sejam menores de idade;
- ✓ Tenham renda inferior a um certo mínimo etc.

Após essa análise inicial, o score de crédito é usado para se determinar se o cliente deve ou não merecer crédito. Na sua forma mais simples o score de crédito procura discriminar os clientes em dois grupos (bons e maus pagadores) a partir da análise de uma amostra de clientes. As razões para se selecionar uma amostra e não toda a população de clientes são as seguintes:

- ✓ Coletar, conferir e processar os dados de toda a população de clientes pode ser muito dispendioso ou demorado;
- ✓ Obter o cadastro de toda a população é impossível;
- ✓ Podemos obter uma boa precisão trabalhando apenas com uma parte dessa população (uma amostra) etc.

Para se construir um escore de crédito, é necessário elaborar um plano de amostragem, ou seja, deve-se determinar:

- ✓ A população a ser estudada (os nossos atuais clientes ou os clientes potenciais do negócio);
- ✓ As variáveis a serem utilizadas (informações sobre os clientes que serão empregadas);
- ✓ O método de escolha da amostra (a maneira como escolheremos, entre toda a população, aquela parcela a ser analisada);
- ✓ O tamanho da amostra (quantos indivíduos devem compô-la).

## COMO MONTAR O PLANO DE AMOSTRAGEM

O primeiro passo para elaboração de um plano de amostragem é definir claramente a população a ser estudada.

É necessário que tenhamos acesso às informações a respeito desses clientes. Quando essa população é composta por clientes dos quais dispomos de cadastro, esse trabalho é imensamente facilitado. Caso contrário, essas informações devem ser obtidas de outra forma (cadastro de um negócio semelhante, por exemplo).

Se isso não for possível, o credit scoring não pode ser utilizado.

O segundo passo do plano de amostragem é a escolha das variáveis que vão ser utilizadas.

A escolha das variáveis deve se pautar por dois princípios: a **relevância** e a **praticidade**.

**Relevância** significa que a variável escolhida é útil para se discriminar o bom e o mau pagador.

**Praticidade** significa que devemos ter acesso a essas informações sem muitas dificuldades.

De nada adianta achar que uma certa variável é muito boa se não se pode medi-la ou ter acesso a ela.

A consulta a alguém com experiência na área deve ser usada como ponto de partida na escolha de variáveis.

Um analista de crédito acha que certas informações são importantes a partir da sua experiência prática. Ele pode afirmar, por exemplo, que:

- ✓ As mulheres são melhores pagadoras que os homens;
- ✓ Os casados são melhores pagadores que os solteiros;
- ✓ Os mais pobres são melhores pagadores que os mais ricos etc.

A análise estatística que vai ser realizada com os dados obtidos vai procurar determinar até que ponto esses palpites têm algum fundamento ou são mero preconceito.

Uma advertência: O fato de haver uma diferença estatisticamente significativa com relação ao sexo, estado civil etc. não significa que as diferenças observadas sejam importantes. Uma diferença estatisticamente significativa pode ser algo muito pequeno como 1%.

Uma vez definidas a população e as variáveis, devemos adotar um método de amostragem. Em geral, o melhor método de escolher uma amostra é o aleatório, ou seja, a seleção ao acaso de um certo número de clientes do nosso cadastro.

No caso do score de crédito, em particular, usamos normalmente uma variação do método aleatório que é o método aleatório estratificado, normalmente com estratos de tamanhos iguais.

O método aleatório estratificado consiste em selecionar uma amostra aleatória de cada um dos grupos que estamos interessados.

A qualificação "de tamanhos iguais" indica que o número de indivíduos selecionados em cada grupo é igual.



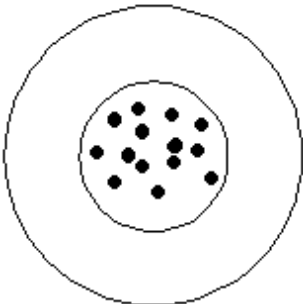
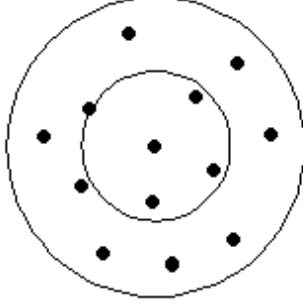
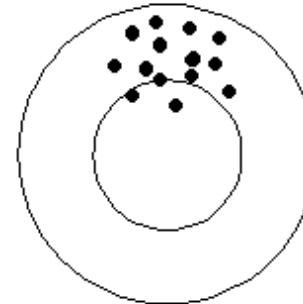
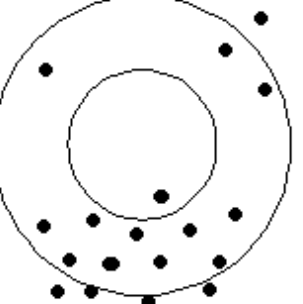
No nosso caso iremos selecionar uma certa quantidade de clientes bons pagadores e a mesma quantidade de clientes maus pagadores, não importando que os clientes maus pagadores sejam apenas 10% do cadastro, por exemplo.

Uma vez decidido o método de amostragem deve-se determinar o tamanho dessa amostra.

Como não estamos trabalhando com a população, mas com uma amostra, em geral não dispomos dos valores populacionais (parâmetros) das variáveis selecionadas, devemos, portanto, estimar esses valores.

As características desejáveis desses estimadores são duas: eles devem ser **não enviesados** (ou seja, os valores estimados devem estar em torno dos valores reais) e **precisos** (ou seja, os valores estimados devem estar concentrados em torno de um ponto). Essas características são resumidas na tabela abaixo.

**Tabela 1 - Representação dos conceitos de viés e precisão**

	Alta precisão	Baixa precisão
Não enviesado		
Enviesado		

Os estimadores tendem a ser **não enviesados** se a amostra for representativa da população da qual ela foi retirada.

Essa condição é normalmente satisfeita se o método de amostragem for aleatório.

A **precisão** tende a ser garantida pelo tamanho da amostra.

Quanto maior a amostra, mais precisos se tornam os estimadores.

Esse é um ponto importante. A precisão depende do tamanho da amostra e não do tamanho da população (desde que esta última seja grande).

Vamos explicar de novo em uma linguagem bem simples.

Para sabermos se uma sopa está salgada basta que experimentemos uma colherada dessa sopa, não importa se tirada de uma panela ou de um enorme caldeirão. Uma amostra de 1.000 indivíduos vai nos dar uma precisão tão boa para uma população de 50 mil habitantes como para uma população de 160 milhões.

Dados o erro padrão da estimativa (o valor absoluto da diferença admitida entre estimativa e parâmetro) e o nível de confiança desejado (a probabilidade de o intervalo definido como (estimativa - erro, estimativa + erro) conter o parâmetro estimado), o analista terá condições de determinar o tamanho da amostra adequado para cada caso.

Uma vez determinado o tamanho da amostra, estamos prontos para executar o plano de amostragem e obter os dados que iremos utilizar para elaborar o escore de crédito.

O nosso próximo passo é descobrir uma função discriminante que consiga separar adequadamente os dois grupos de indivíduos que temos, os bons e maus pagadores.

## COMO CLASSIFICAR OS INDIVÍDUOS EM UMA ANÁLISE DE CRÉDITO

A obtenção de uma regra de classificação para os indivíduos em uma análise de crédito é feita, em geral, com base em uma combinação linear das variáveis que selecionamos anteriormente.

Essa combinação é feita usando-se um método estatístico para analisar a relação entre essas variáveis, que chamamos de independentes, e uma variável dependente, que é o fato do indivíduo ser bom ou mau pagador.

Esse método estatístico chama-se regressão linear.

A idéia da regressão linear é construir um modelo tal que  $Y=f(x_1, x_2, \dots, x_n) = a + b_1 \cdot x_1 + b_2 \cdot x_2 + \dots + b_n \cdot x_n$ , no qual  $x_1, x_2, \dots, x_n$  são as variáveis que selecionamos anteriormente e que julgamos ser capazes de nos dizer se aquele cliente é bom pagador ( $Y=1$ ) ou não ( $Y=0$ ).

Os coeficientes de cada uma das variáveis independentes, ou, no nosso caso, o peso atribuído a cada variável é determinado pela regressão linear.

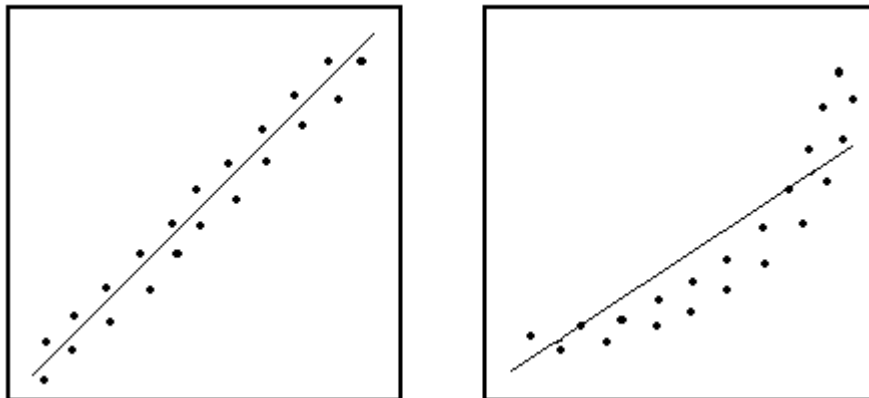
No caso de uma regressão simples (uma variável dependente e uma independente) esses pesos são determinados ao se encontrar uma reta que minimize o erro quadrático.

No caso de uma regressão linear múltipla (uma variável dependente e duas ou mais variáveis independentes) esses pesos são determinados da mesma forma, mas nesse caso a figura a ser encontrada é um plano ou hiperplano. Maiores detalhes podem ser obtidos em um bom livro de Estatística.

Cabe aqui uma advertência. Apesar de ser a mais empregada na prática por sua simplicidade, a função discriminante linear nem sempre produz uma regra eficiente. Isso pode acontecer porque:

- ✓ Não existe relação entre as variáveis que escolhemos;
- ✓ A relação linear existe mas ocorre por mero acaso, como associar a cor dos olhos com a qualidade do cliente, por exemplo;
- ✓ A relação que estamos procurando, apesar de existir, não é linear (vide um exemplo na figura abaixo). Nesses casos teremos que usar técnicas mais sofisticadas, que fogem ao escopo desse trabalho.

**Figura 1 - Relação linear e não linear**



Existem diversos pacotes estatísticos que geram essas funções. O Excel, inclusive, dispõe da função PROJ.LIN que permite fazer regressões lineares simples e múltiplas. No **Up-To-Date® 84** vamos construir um exemplo bem simples para mostrar como se pode calcular um escore de crédito usando-se o Excel.