

# Uma Metodologia para Desenvolvimento de Modelos de Desempenho de Processos para Gerência Quantitativa de Projetos de Software

Mariano Montoni<sup>1</sup>, Marcos Kalinowski<sup>1</sup>, Peter Lupo<sup>1</sup>, José Fortuna Abrantes<sup>1</sup>,  
Analia Irigoyen Ferreiro Ferreira<sup>1,2</sup>, Ana Regina Rocha<sup>1</sup>

<sup>1</sup> COPPE/UFRJ - Universidade Federal do Rio de Janeiro  
Caixa Postal 68511 – CEP 21945-970 – Rio de Janeiro, Brasil

{mmontoni, mkali, jfa, darocha}@cos.ufrj.br

<sup>2</sup> BL Informática Ltda.

Av. Visconde do Rio Branco 305/8º andar - Niterói - RJ - CEP 24020-002

analia@blnet.com

**Abstract.** *Statistical process control has been used in order to control the performance of software processes. However, this technique is limited to control the value of a metric in a current project, compared to the same metric considering the process performance. In this paper we present a quantitative project management methodology combining statistical process control and statistical performance models to allow predicting behaviors for different scenarios to support managerial decisions. The methodology was developed during the application of those techniques in a software organization to control and relate the Quality Assurance and Verification processes.*

**Resumo.** *O controle estatístico de processos tem sido utilizado para controlar o desempenho de processos de software. Esta técnica se limita a controlar o valor de uma métrica em um projeto atual em relação a esta mesma métrica considerando o desempenho do processo. Este artigo apresenta uma metodologia para a gerência quantitativa de projetos de software combinando o controle estatístico de processos com modelos estatísticos de desempenho possibilitando que possam ser feitas previsões sobre diferentes cenários, apoiando a tomada de decisões gerenciais. A metodologia foi desenvolvida durante a aplicação destas técnicas em uma organização de software para controlar e relacionar os processos Garantia da Qualidade e Verificação.*

## 1. Introdução

De acordo com Florac e Carleton (1999) embora profissionais da área de software entendam que dados podem ser úteis, eles raramente os utilizam durante seu trabalho. Entre os motivos para a não utilização de dados podemos destacar dois: (1) dificuldade em entender e controlar os dados coletados; e (2) dificuldade em construir modelos estatísticos que permitam utilizar os dados coletados para apoiar decisões que visem a melhoria da qualidade e a redução de custos de desenvolvimento.

Considerando o primeiro destes motivos, o controle estatístico de processos, conceito inicialmente utilizado na área de manufatura (Shewhart, 1931), têm se

mostrado uma abordagem eficiente para, baseado em dados, entender e controlar o comportamento de processos executados em empresas. Ele permite estabelecer mecanismos para detectar causas atribuíveis para variações no processo e controlar a implementação de melhorias com o objetivo de alcançar metas viáveis de qualidade, tempo e custo.

Florac e Carleton (1999) descrevem como o controle estatístico de processos pode ser aplicado para apoiar a melhoria de processos de software. Em relação ao segundo objetivo, por sua vez, uma metodologia para a análise dos dados e construção de modelos estatísticos pode ser encontrada em (Maxwell, 2002). Tais modelos têm sido utilizados por organizações desenvolvedoras de software com o intuito de permitir previsões considerando diferentes cenários, apoiando a tomada de decisões gerenciais que visem o aumento da qualidade dos produtos desenvolvidos, diminuindo tempo e custo de execução dos processos.

Neste artigo apresentamos uma metodologia para desenvolvimento de modelos de desempenho de processos para gerência quantitativa de projetos de software. Essa metodologia foi aplicada no contexto da empresa BL Informática, uma organização desenvolvedora de software. Essa organização foi recentemente avaliada no nível de maturidade 3 do modelo CMMI (*Capability Maturity Model Integration*). Para alcançar esse nível a organização investiu nos últimos 2 anos na implementação de processos de gerência de projetos e processos de engenharia de software.

A aplicação da metodologia apresentada neste trabalho resulta em ferramentas para apoiar a gerência quantitativa de projetos de software dessa empresa. Estas ferramentas são constituídas por gráficos de controle e um modelo estatístico de desempenho, permitindo o gerente de projeto realizar estimativas mais precisas para garantir a qualidade dos produtos desenvolvidos ao longo do ciclo de vida do software e assim atingir as metas de desempenho e qualidade dos processos da organização.

O restante do artigo está organizado da seguinte forma. Na seção 2 são apresentados trabalhos relacionados que fundamentam a metodologia do estudo apresentado neste artigo. Na seção 3 é descrita a aplicação da metodologia em uma organização de desenvolvimento de software. Na seção 4 são discutidos os resultados obtidos do estudo. Na seção 5 são apresentadas conclusões e discutidas perspectivas futuras do trabalho apresentado.

## **2. Fundamentação teórica**

O controle estatístico de processos (Shewhart, 1931) é um método que permite monitorar processos visualmente, por meio do uso de gráficos de controle aplicados sobre coletas frequentes de dados. Estes gráficos se pautam na teoria de que as variações inerentes ao processo variam dentro de limites previsíveis, fazendo uso de cálculos estatísticos para obter o comportamento esperado do processo e os limites (superior e inferior) aceitáveis para sua variação, a partir de dados históricos. De acordo com Wheeler (1999) o controle estatístico de processos é um método eficiente para detectar mudanças e identificar causas atribuíveis que afetam a qualidade do processo. Uma causa atribuível é uma causa que poderia ter sido prevenida, não tendo sua origem em um fenômeno natural e inerente ao processo (Florac e Carleton, 1999). O potencial

de custo benefício fica evidente, uma vez que torna possível controlar a qualidade do processo durante o desenvolvimento e não após a entrega do produto.

Fairley (1999) cita o controle estatístico de processos como ferramenta útil para controlar também processos no contexto de desenvolvimento de software. Um exemplo de sua aplicação para controlar o processo de detecção de defeitos por meio de revisões por pares se encontra descrito em (Hong et al., 1999). Wang et al. (2006) apresentam uma abordagem baseada em análise causal e no controle estatístico de processos para estabelecer uma baseline do desempenho de processos de software. A importância do controle estatístico de processos no desenvolvimento de software é ainda evidenciada pelas exigências referentes ao mesmo no processo Gerência Quantitativa de Projetos do nível B do MPS.BR (MPS.BR, 2006) e 4 do CMMI (Crissis et al., 2006). Uma descrição de como aplicá-lo neste contexto pode ser encontrada em (Florac e Carleton, 1999).

Entretanto, o controle estatístico de processos se limita a controlar o valor de uma métrica em um projeto atual, em relação a esta mesma métrica considerando o desempenho do processo. Assim, por si só, não permite entender o relacionamento entre diferentes métricas nem previsões para diferentes cenários de desenvolvimento. É neste quesito que modelos estatísticos de desempenho podem complementar o controle estatístico de processos, permitindo descrever o relacionamento entre diferentes métricas (Maxwell, 2002). Assim, modelos estatísticos de desempenho podem ser utilizados como ferramenta para a previsão de métricas ainda não obtidas a partir de métricas atuais. Uma das formas de obter um modelo estatístico de desempenho é por meio da análise de regressão, relacionando variáveis dependentes com variáveis independentes (Maxwell, 2002).

Fenton (1999), realizou uma análise crítica a respeito de modelos de estimativa para defeitos de software. Ele destaca a importância de utilizar modelos que considerem a realidade da organização e processos bem documentados, que foram efetivamente seguidos. Os modelos estatísticos de desempenho são construídos a partir de métricas da organização, refletindo a realidade destas. Em relação aos processos bem documentados e efetivamente seguidos, é importante destacar que a precisão dos modelos pode aumentar quando são exigidos processos institucionalizados e estatisticamente controlados, assegurando que os valores das métricas utilizados na concepção do modelo representam o desempenho do processo da organização e não variações originadas por causas atribuíveis em projetos específicos. Maxwell (2002) destaca ainda a importância de verificar a distribuição das métricas antes de dar início à análise de regressão. Sabe-se ainda que a acurácia das previsões por modelos estatísticos de desempenho normalmente não é significativamente menor quando o número de métricas é reduzido (3-6 ao invés de 30) (Munson e Khoshgoftaar, 1990). O motivo para tal é que muitas métricas são colineares, capturando a mesma informação básica. Isto destaca a importância da análise de correlação entre as diferentes métricas.

Com base nesta fundamentação teórica a seção seguinte descreve a aplicação da metodologia para gerar ferramentas de gerência quantitativa de projetos de software focando nos processos Garantia da Qualidade e Verificação de produtos de software em uma organização desenvolvedora de software, certificada no nível 3 do CMMI, possuindo estes processos bem definidos e institucionalizados. Pudemos verificar ainda

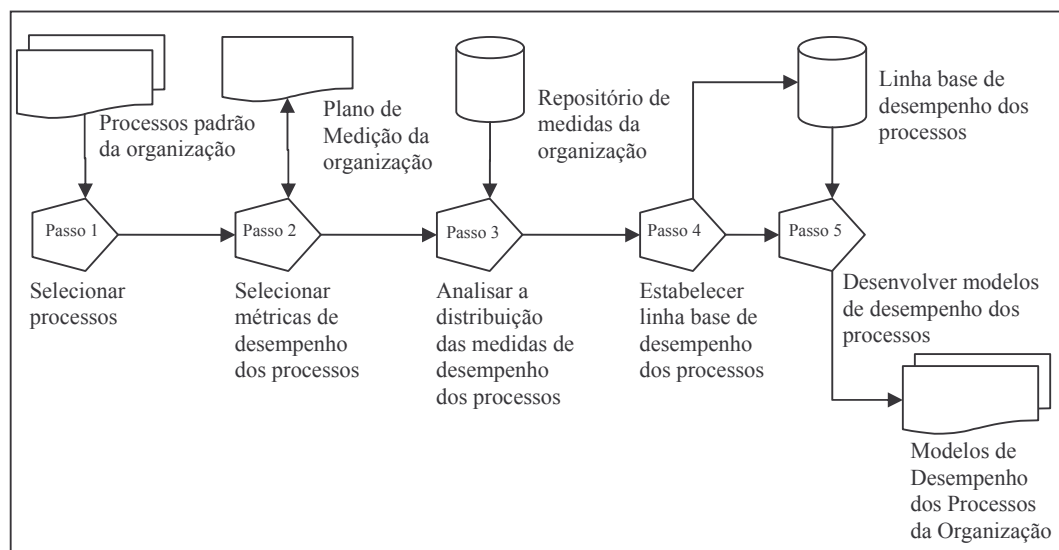
que as medidas utilizadas para a concepção das ferramentas apresentavam distribuição normal e os subprocessos selecionados se encontravam estatisticamente controlados.

### 3. Metodologia

A metodologia apresentada neste trabalho foi usada como base para realizar um estudo na BL Informática. Esse estudo visa desenvolver mecanismos para gerenciar estatisticamente processos críticos executados nos projetos de desenvolvimento de produtos de software da organização.

Após os processos terem sido implementados e implantados na organização, percebeu-se que ocorriam desvios freqüentes nas estimativas de esforço e prazo dos projetos devido a um alto índice de retrabalho associado a problemas técnicos encontrados em produtos de software de diversos tipos, por exemplo, especificações de requisitos, planos e relatórios de testes, etc. Esses problemas são detectados nos projetos a partir da execução de atividades de avaliação técnica de revisão por pares que constitui parte do processo Verificação implantado na organização. Dessa forma, o objetivo da organização é melhorar a estimativa dos projetos de desenvolvimento de software por meio de uma maior exatidão na previsão do esforço necessário para corrigir problemas detectados nas atividades de avaliação técnica de revisão por pares. Assim, melhores estimativas poderão ser geradas acarretando menos desvios de esforço, prazo e custo nos projetos de desenvolvimento de software.

A metodologia para desenvolvimento de modelos de desempenho de processos para gerência quantitativa de projetos de Software é constituída de um conjunto de passos, ilustrados na Figura 1 e descritos a seguir.



**Figura 1. Processo de Desenvolvimento de Modelos de Desempenho de Processos para Gerência Quantitativa de Projetos de Software.**

#### *Passo 1 – Selecionar processos*

O primeiro passo da metodologia adotada no estudo é selecionar os processos para avaliação de estabilidade. A seleção destes processos deve ser realizada com base nos

processos padrões institucionalizados na organização a partir de critérios objetivos, por exemplo, devem ser selecionados processos que se relacionam com objetivos de negócio ou que possuem dados históricos relevantes.

O primeiro processo escolhido foi o processo Verificação, mais especificamente o subprocesso contendo as atividades de avaliação técnica de revisão por pares e o segundo processo escolhido foi o processo Garantia da Qualidade, mais especificamente o subprocesso contendo as atividades de avaliação de garantia de qualidade.

Estes subprocessos foram escolhidos, pois se acredita que possa haver uma correlação significativa entre a quantidade de problemas encontrados nas atividades de avaliação de garantia de qualidade e a quantidade de problemas encontrados nas atividades de avaliação técnica de revisão por pares. Se essa relação realmente existir, será possível desenvolver um modelo de desempenho para prever a quantidade de problemas encontrados nas atividades de avaliação técnica de revisão por pares a partir do resultado das avaliações de garantia de qualidade do produto.

Esse modelo será útil para os gerentes de projeto durante as atividades de monitoração do desempenho do projeto. Por exemplo, após a avaliação de garantia de qualidade de um produto de software específico, o gerente de projeto poderá replanejar o projeto se for previsto um número grande de problemas a serem detectados em uma atividade de avaliação técnica de revisão por pares de um produto de software específico.

### ***Passo 2 – Selecionar métricas de desempenho dos processos***

Após a seleção dos processos para serem avaliados quanto à sua estabilidade, o segundo passo da metodologia é identificar as métricas do processo que descrevem o desempenho do processo, também chamadas de variáveis estatísticas. Para derivar essas variáveis foi adotado o paradigma GQM (*Goal-Question-Metric*). Esse paradigma define que métricas devem ser definidas de forma objetiva para responder questões associadas a objetivos de medição da organização (Solingen e Berghout, 1999). A tabela 1 apresenta o resultado da aplicação dessa abordagem para identificar as variáveis estatísticas que serão coletadas para analisar o desempenho de processos de software em questão.

**Tabela 1. Aplicação do GQM para identificar métricas de desempenho de processos**

<b>Objetivo</b>	<b>Questão</b>	<b>Métrica</b>
Aumentar o controle do desempenho das atividades de avaliação de garantia de qualidade e das atividades de avaliação técnica de revisão por pares.	Qual o número de problemas identificados na atividade de avaliação de garantia de qualidade?	Número de problemas identificados na atividade de avaliação de garantia de qualidade (ProblemasGQ)
	Qual o número de problemas identificados na atividade de avaliação técnica de revisão por pares?	Número de problemas identificados na atividade de avaliação técnica de revisão por pares (ProblemasAP)

A seleção das métricas de desempenho dos processos deve ser realizada com base no Plano de Medição institucionalizado na organização, ou seja, deve-se procurar

identificar métricas que já estejam sendo coletadas na organização para agilizar o estabelecimento da linha base de desempenho dos processos (próximo passo da metodologia apresentada). Se o Plano de Medição da organização não contiver as métricas desejadas para medir o desempenho dos processos, então as métricas deverão ser especificadas, incorporadas no Plano de Medição e coletadas nos projetos.

Considerando as métricas da Tabela 1, o objetivo do modelo a ser gerado pela metodologia é permitir que a métrica ProblemasAP possa ser calculada a partir da métrica ProblemasGQ. Para orientar o desenvolvimento desse modelo, foi elaborada uma hipótese que pudesse ser testada por meio de coleta e análise estatística de dados gerados a partir da execução de atividades de avaliação de garantia de qualidade e atividades de avaliação técnica de revisão por pares. Inicialmente foi elaborada a seguinte hipótese a ser testada:

- $H_0$ : Quanto maior o número de problemas identificados nas atividades de avaliação de garantia da qualidade, *menor* é o número de problemas identificados nas atividades de avaliação técnica de revisão por pares.

Essa hipótese foi definida com base na percepção dos autores de que após um produto de software ter passado por uma avaliação de qualidade e terem sido removidos os problemas identificados nessa avaliação, o produto terá menos problemas técnicos visto que um certo esforço já foi dedicado na remoção de problemas em um momento prévio. Essa hipótese poderá ser testada como sendo verdadeira se for identificado um modelo de desempenho estatisticamente significativo que relaciona *negativamente* a variável ProblemasGQ com a variável ProblemasAP.

### ***Passo 3 – Analisar a distribuição das medidas de desempenho dos processos***

Para construir o modelo de desempenho de processo, o terceiro passo da metodologia do estudo é analisar a distribuição das medidas de desempenho dos processos com base nas métricas selecionadas armazenadas no repositório de medidas da organização. Assim, foram analisados 16 relatórios de avaliação de 4 projetos de desenvolvimento de software. O Anexo A apresenta os dados coletados a partir dos relatórios dos projetos. O tamanho dos projetos selecionados são apresentadas na tabela a seguir:

**Tabela 2. Tamanho dos projetos de desenvolvimento de software analisados**

<b>Projeto</b>	<b>Tamanho Real (Pontos de Função)</b>
P1	41
P2	89
P3	54,55
P5	205,43

A partir da análise dos relatórios de avaliação desses projetos foram coletadas as métricas ProblemasGQ e ProblemasAP com o propósito de desenvolver os gráficos de controle e o modelo de desempenho de processo descritos na seção anterior. Para cada relatório de avaliação, foram identificadas medidas da métrica ProblemasGQ relacionadas a outras medidas da métrica ProblemasAP. Esses dados foram analisados para garantir que fazem parte de uma mesma distribuição. Esta análise foi realizada usando o teste estatístico chamado Anderson–Darling (D’Agostino e Stephens, 1986).

Este teste é usado para testar se uma amostra de dados corresponde a uma distribuição específica. Valores Anderson–Darling pequenos indicam que existe uma melhor adequação entre os dados e a distribuição específica. Este método é uma modificação do teste Kolmogorov-Smirnov (K-S) e provê mais peso nas extremidades da distribuição que o teste K-S. O teste de normalidade Anderson–Darling gera um gráfico de probabilidade de normalidade de dados e realiza um teste de hipótese para examinar se os dados observados seguem ou não uma distribuição normal. Para o teste de normalidade, as hipóteses são:

- $H_0$ : Os dados seguem uma distribuição normal.
- $H_1$ : Os dados não seguem uma distribuição normal.

O teste estatístico Anderson–Darling é definido como:

$$A^2 = -N - (1/N) \sum (2i - 1) (\ln F(Y_i) + \ln(1 - F(Y_{N+1-i})))$$

onde:

F é a função de distribuição acumulativa da distribuição normal, e  $Y_i$  são as observações ordenadas.

O teste Anderson-Darling utiliza a distribuição específica no cálculo de valores críticos. Isto tem uma vantagem de permitir um teste mais sensível e a desvantagem de que valores críticos devem ser calculados para cada distribuição. A tabela a seguir apresenta os valores críticos do teste Anderson-Darling para a distribuição normal (D'Agostino e Stephens, 1986). O valor na primeira coluna (confiabilidade do teste) é equivalente a  $100(1 - \alpha)$ , sendo que  $\alpha$  é o nível de significância estatística (coluna 2). O teste Anderson-Darling não afirma que os dados seguem uma distribuição normal. Esse teste apenas diz quando os dados tornam pouco provável que a distribuição não seja normal. Portanto, a hipótese nula é rejeitada no nível de significância estatística específico se o valor do teste Anderson-Darling é maior que o valor crítico na coluna 3.

**Tabela 3. Valores críticos do teste Anderson-Darling para a distribuição normal**

Confiabilidade do Teste	$\alpha$	Valor Crítico
90 %	0,1	0,631
95 %	0,05	0,752
97,5 %	0,025	0,873
99 %	0,01	1,035

Os gráficos de normalidades das métricas ProblemasGQ e ProblemasAP normalizadas pelo tamanho dos projetos gerados a partir da aplicação do teste estatístico Anderson-Darling são apresentadas nas figuras 2 e 3, respectivamente. A normalização das medidas foi realizada para minimizar distorções decorrentes da grande variação de complexidade dos projetos de desenvolvimento de software da organização. Os gráficos de normalidades exibem os valores de p-value do teste estatístico Anderson-Darling (p-value < 0,005). Esses valores indicam que em níveis de significância estatística  $\alpha$  maiores que 0,005 há evidências de que os dados não seguem uma distribuição normal.

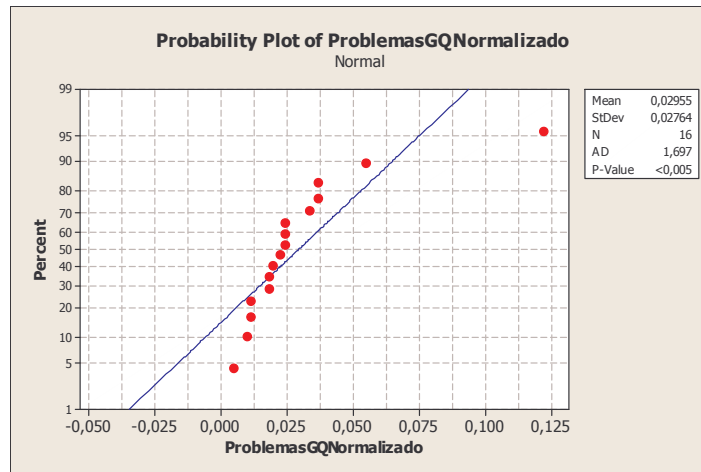


Figura 2. Gráfico de normalidade da métrica ProblemasGQ Normalizada.

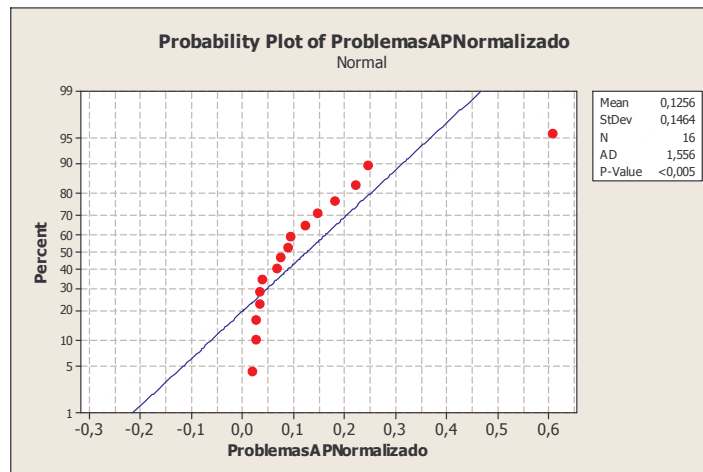
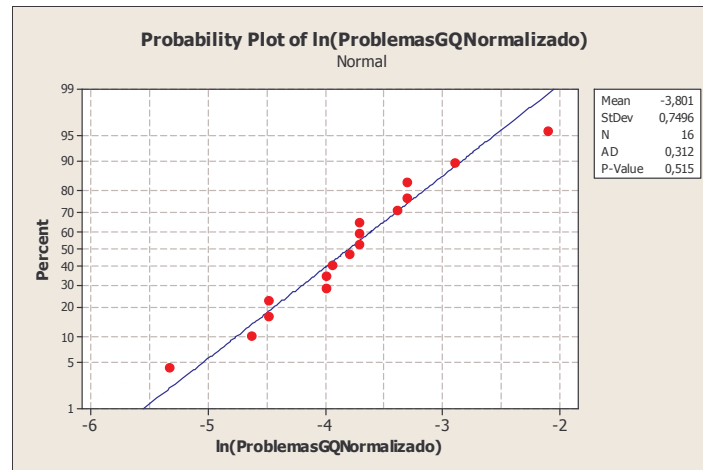


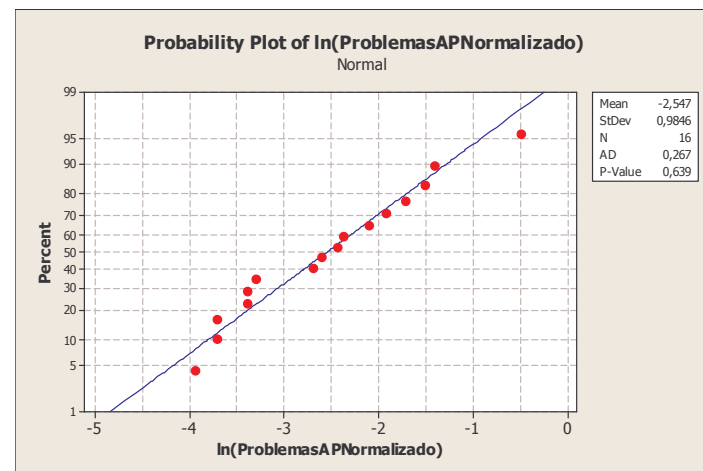
Figura 3. Gráfico de normalidade da métrica ProblemasAP Normalizada.

Considerando que o número de problemas encontrados em uma avaliação de um artefato não seguem uma distribuição normal, uma segunda análise dos dados foi realizada para verificar a normalidade dos dados coletados. Segundo Maxwell (2002) é típico que projetos de desenvolvimento de software gerem dados que não são normalmente distribuídos devido à diversidade dos projetos. Maxwell (2002) sugere que para aproximar os dados coletados em contextos diferentes de desenvolvimento de software a uma distribuição normal é necessário transformar as métricas por meio do cálculo do logaritmo natural das medidas. O logaritmo natural torna valores muito grandes em valores menores e dessa forma os dados são mais aproximados. As figuras 4 e 5 apresentam os gráficos de normalidades do logaritmo natural das métricas ProblemasGQ e ProblemasAP normalizadas pelo tamanho dos projetos.





**Figura 4. Gráfico de normalidade do logaritmo natural da métrica ProblemasGQ Normalizada.**



**Figura 5. Gráfico de normalidade do logaritmo natural da métrica ProblemasAP Normalizada.**

O gráfico da Figura 4 apresenta o valor de p-value do teste estatístico Anderson-Darling igual a 0,515. Esse valor indica que em níveis de significância estatística  $\alpha$  maiores que 0,515 há evidências de que os dados não seguem uma distribuição normal. Isto significa que existe probabilidade baixa de que os dados coletados da métrica ProblemasGQ não seguem uma distribuição normal. O gráfico da Figura 5 apresenta o valor de p-value do teste estatístico Anderson-Darling igual a 0,639. Desta forma, igualmente à análise da métrica ProblemasGQ, o teste Anderson-Darling indica que existe probabilidade baixa de que os dados coletados da métrica ProblemasAP não seguem uma distribuição normal. Portanto, podemos concluir que é altamente provável que os dados coletados fazem parte de uma distribuição normal.

Após a verificação da normalidade da distribuição das amostras utilizando-se do teste Anderson-Darling, foram gerados histogramas do logaritmo natural das medidas coletadas normalizadas com base no tamanho real do projeto medido em pontos de

função. As figuras 6 e 7 apresentam os histogramas dos logaritmos naturais das variáveis ProblemasGQ e ProblemasAP normalizadas. Nos gráficos são apresentados também os valores de média e desvios padrões das métricas de desempenho de processos.

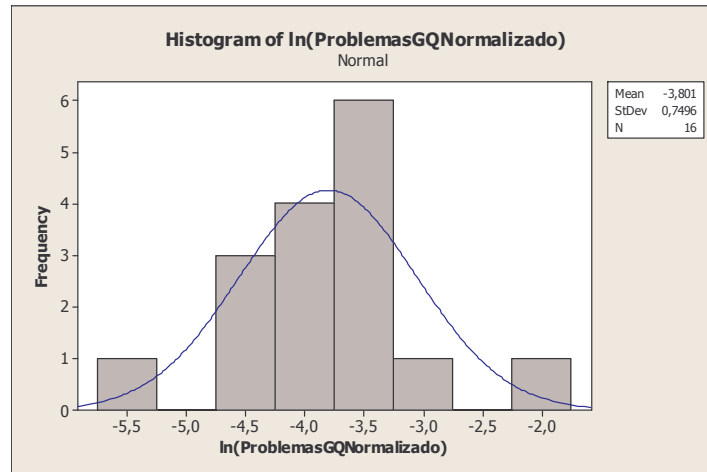


Figura 6. Histograma de distribuição normal do logaritmo natural da métrica ProblemasGQ Normalizada.

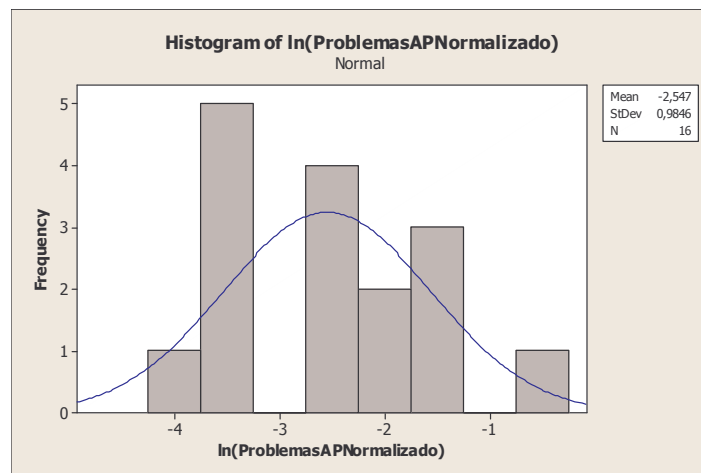


Figura 7. Histograma de distribuição normal do logaritmo natural da métrica ProblemasAP Normalizada.

#### ***Passo 4 – Estabelecer linha base de desempenho dos processos***

O próximo passo da metodologia é estabelecer uma linha base de desempenho dos processos. Para realizar este passo, deve-se primeiro analisar a estabilidade dos subprocessos selecionados. Este passo deve ser executado somente se foi identificado no passo anterior que as medidas coletadas representam uma mesma distribuição, ou seja, os dados representam um comportamento único de desempenho do processo e portanto modelos de desempenho precisos podem ser derivados com base nesses dados.

Para analisar a estabilidade dos processos, foram utilizados gráficos de controle *Individuals and Moving Range* (XmR charts). Para o cálculo das faixas de operação e dos limites, um padrão comumente utilizado em gráficos de controle foi seguido, no qual o limite superior se encontra três desvios padrão acima da média e o inferior três desvios padrão abaixo da média. Para analisar a estabilidade de um processo, os gráficos de controle devem ser examinados para detectar padrões que indicam comportamento incomum. Valores que fogem dos limites e padrões não usuais sugerem que causas atribuíveis existem, ou seja, as variações detectadas nos processos não são consideradas naturais do processo e investigações sobre as causas dessas variações devem ser realizadas. Vários tipos de testes podem ser aplicados para analisar a estabilidade de um processo, a seguir são apresentados 4 tipos comuns desses testes (Wheeler, 1999):

- Teste 1: Um ponto único está fora dos limites de controle 3-sigma.
- Teste 2: Pelo menos dois de três valores sucessivos estão no mesmo lado e mais de duas unidades de sigma de distância da linha central.
- Teste 3: Pelo menos quatro de cinco valores sucessivos estão no mesmo lado e mais de uma unidade de sigma de distância da linha central.
- Teste 4: Pelo menos oito valores sucessivos estão no mesmo lado da linha central.

Os gráficos de controle XmR para as métricas selecionadas (logaritmo natural das métricas ProblemasGQ e ProblemasAP normalizadas pelo tamanho do projeto) são apresentados nas figuras 8 e 9. Os gráficos apresentam também os valores de média e limites superiores e inferiores para controle estatístico.

Aplicando os testes de estabilidade de processos nos gráficos de controle das figuras 8 e 9 podemos concluir que os processos analisados são estáveis. Portanto, a média e os limites naturais calculados para os processos em questão podem ser usados para estabelecer os objetivos de desempenho de processos e qualidade da organização. Isto significa que nos próximos projetos desenvolvidos na organização, pode-se verificar se os processos Verificação e Garantia de Qualidade (mais especificamente as atividades de avaliação técnica e avaliação da garantia da qualidade) se comportam conforme os objetivos de desempenho de processos definidos pela organização com base na variação inerente dos processos calculada a partir de dados históricos.

O conjunto de dados históricos usados para definir o comportamento natural dos processos constitui a linha base de desempenho de processos da organização. Periodicamente, esta linha base deve ser revista e atualizada para ajustar os objetivos de qualidade da organização com base em mudanças introduzidas nos processos que resultam em alterações no comportamento e desempenho dos processos da organização.

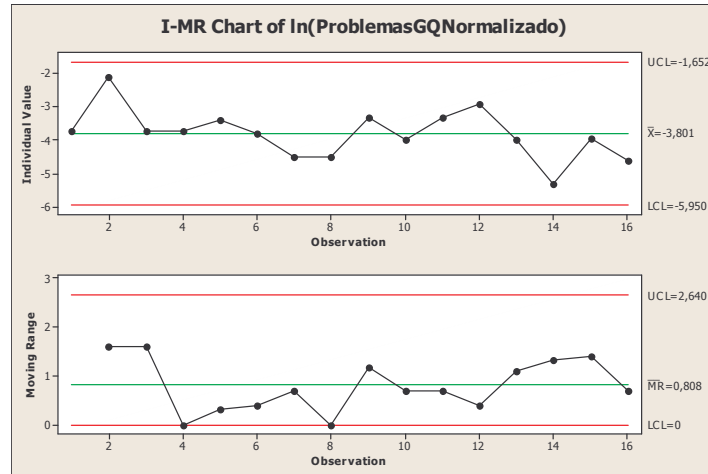


Figura 8. Gráfico de análise de estabilidade da variável ProblemasGQ.

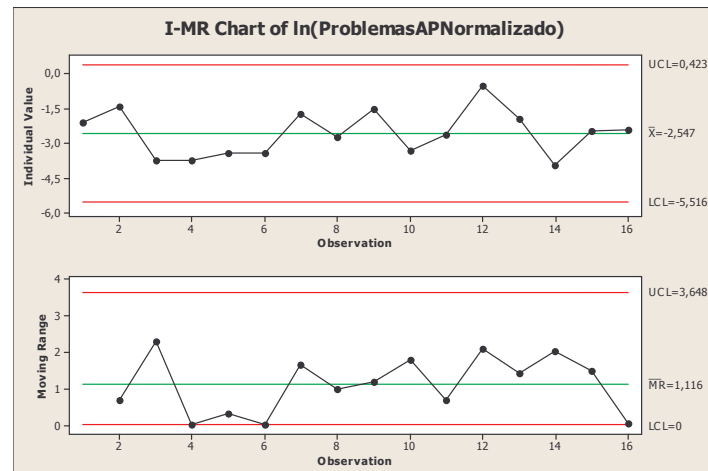


Figura 9. Gráfico de análise de estabilidade da variável ProblemasAP.

### *Passo 5 – Desenvolver modelos de desempenho dos processos*

Após verificar que as medidas coletadas fazem parte da mesma distribuição de dados e representam processos estáveis, o passo seguinte na metodologia é desenvolver o modelo de desempenho. No caso apresentado neste trabalho, o modelo de desempenho foi desenvolvido para relacionar as métricas ProblemasGQ e ProblemasAP normalizadas pelo tamanho real dos projetos de desenvolvimento software.

Supondo distribuição normal dos dados e estabilidade dos processos, a linha base de desempenho de processos estabelecida no passo anterior foi usada no desenvolvimento do modelo estatístico de desempenho por meio da análise de correlação (utilizando a correlação de Pearson) e de regressão dos dados. Para desenvolver os modelos de desempenho é necessário identificar a variável dependente e as variáveis independentes a serem consideradas na elaboração dos modelos. No estudo em questão, foi identificada a variável ProblemasGQ Normalizada como sendo a variável independente e a variável ProblemasAP Normalizada como dependente.

A partir da aplicação da técnica de análise de regressão na linha base de desempenho dos processos estabelecida, derivou-se a seguinte equação de regressão:

$$\text{ProblemasAPNormalizado} = 0,0502 + 2,55 \text{ ProblemasGQNormalizado}$$

Pode-se observar na equação de regressão derivada que existe uma correlação positiva entre o número de problemas identificados nas atividades de avaliação de garantia da qualidade e o número de problemas identificados nas atividades de avaliação técnica de revisão por pares. Esta correlação indica que o número de problemas encontrados nas atividades de avaliação técnica de revisão por pares cresce proporcionalmente à medida que cresce o número de problemas encontrados nas atividades de avaliação de garantia da qualidade. Esta relação serve de indicador da qualidade dos produtos desenvolvidos, ou seja, se o produto apresenta muitos problemas sob o ponto de vista da qualidade, então é natural que esse mesmo produto também tenha muitos problemas sob o ponto de vista de conteúdo.

A partir do modelo de desempenho obtido, podemos perceber que a hipótese  $H_0$  definida no passo 2 da metodologia é negada. Assim a hipótese verificada como verdadeira é a seguinte:

- $H_1$ : Quanto maior o número de problemas identificados nas atividades de avaliação de garantia da qualidade, *maior* é o número de problemas identificados nas atividades de avaliação técnica de revisão por pares.

É importante ressaltar que foram utilizadas apenas 16 observações para o cálculo do modelo e que o nível de significância é bastante alto, mostrando um grau de certeza de aproximadamente 95% (p-value igual a 0,059). Além do mais, se observou um valor de correlação de Pearson de 0,482. Uma outra informação gerada a partir da aplicação da técnica de análise de regressão é que o número de problemas identificados nas atividades de avaliação de garantia da qualidade é responsável em 23,2% no resultado do número de problemas identificados nas atividades de avaliação técnica de revisão por pares. Estes dados são bastante satisfatórios visto que o estudo tem caráter exploratório e que o modelo foi desenvolvido com um número reduzido de observações.

#### 4. Discussão

Por meio do modelo de desempenho de processo derivado, o gerente de projeto pode monitorar quantitativamente os projetos de desenvolvimento de software a partir da predição da qualidade do produto sob o ponto de vista do seu conteúdo e assim melhorar suas estimativas a respeito da previsão do esforço necessário para corrigir problemas detectados nas atividades de avaliação técnica de revisão por pares. Por exemplo, se forem encontrados poucos problemas nas atividades de avaliação de garantia da qualidade, então é um indicador de que a qualidade do produto está adequada e estimativas mais realistas podem ser obtidas a respeito do esforço, prazo e custo das atividades de resolução de problemas detectados nas avaliações técnicas de revisão por pares.

No entanto, para atingir esse propósito adequadamente, outros modelos de desempenho devem ser definidos procurando correlacionar, por exemplo, o número de

problemas identificados nas atividades de avaliação técnica de produtos com o esforço necessário para corrigir esses problemas. Esta necessidade irá motivar os próximos ciclos de melhoria de processos da organização possibilitando o refinamento da linha base de desempenho dos processos e o desenvolvimento de novos modelos de desempenho.

## 5. Conclusão

Este trabalho apresentou uma metodologia para desenvolvimento de modelos de desempenho de processos para gerência quantitativa de projetos de software. Foi apresentada também a aplicação desta metodologia na BL Informática, uma empresa brasileira de desenvolvimento de software se preparando para atingir altos níveis de maturidade dos seus processos. Esta metodologia se apresentou adequada para derivar modelos de desempenho. No entanto, estudos mais profundos devem ser realizados para avaliar se os modelos de desempenho estão sendo utilizados adequadamente pelos gerentes de projeto dessa organização e se os objetivos de desempenho dos processos e de qualidade estabelecidos estão sendo atingidos. Além do mais, outros modelos de desempenho de processos devem ser explorados para possibilitar uma melhor gerência quantitativa dos projetos de software da organização.

A metodologia apresentada neste trabalho é um ponto importante a ser considerado por organizações que desejam atingir altos níveis de maturidade. Por exemplo, a metodologia pode ser usada como base por organizações que desejam implementar processos de análise de desempenho do processo organizacional definidos no nível 4 do CMMI (Crissis et al., 2006) e no nível B do MPS.BR (MPS.BR, 2006).

## Referências

- Crissis, M.B., Konrad, M., Shrum, S. (2006), *CMMI – Guidelines for Process Integration and Product Improvement (2nd Edition)*, SEI Series in Software Engineering, Addison-Wesley Professional.
- D’Agostino, Stephens (1986), *Goodness-Of-Fit Techniques*, Marcel-Dekker, New York, Table 4.7, p.123. All of Chapter 4, pp.97-193, deals with goodness-of-fit tests based on empirical distribution function (EDF) statistics
- Fairley, R.E. (1999), “Managing by the Numbers: A tutorial on quantitative measurement and control of software projects”, in: *International Conference on Software Engineering, ICSE’99*, Los Angeles, CA.
- Fenton, N.E. (1999), “A Critique of Software Defect Prediction Models”, *IEEE Transactions on Software Engineering*, vol 25, no 5.
- Florac, W.A., Carleton, A.D. (1999), *Measuring the Software Process – Statistical Process Control for Software Process Improvement*, Addison-Wesley.
- Hong, G.Y., Xie, M., Shanmugan, P. (1999), A Statistical Method For Controlling Software Defect Detection Process, *J. on Comp. and Industrial Engineering*, vol. 37.
- Maxwell, K.D. (2002), *Applied Statistics for Software Managers*, Software Quality Institute Series, Prentice-Hall.

MPS.BR – Melhoria de Processo do Software Brasileiro, Guia Geral v1.1, (2006), Maio, [http://www.softex.br/mpsbr/\\_guias/MPS.BR\\_Guia\\_Geral\\_V1.1.pdf](http://www.softex.br/mpsbr/_guias/MPS.BR_Guia_Geral_V1.1.pdf).

Munson, J.C., Khoshgoftaar, T.M. (1990) “Regression Modelling of Software Quality: An Empirical Investigation,” *Information and Software Technology*, vol. 32, no. 2, pp. 106-114, 1990.

Shewhart, W.A. (1931), *Economic Control of Quality of Manufactured Products*, Van Nostrand, New York.

Solingen, R., Berghout, E. (1999), *The Goal/Question/Metric Method – A Practical Guide for Quality Improvement of Software Development*, McGrawHill.

Wang, Q., Jiang, N., Gou, L., Liu, X., Li, M., Wang, Y. (2006), BSR: A Statistic-based Approach for Establishing and Refining Software Process Performance Baseline, in: *Int. Conf. on Soft. Eng., ICSE’06*, ACM Press, New York, NY, USA, pp. 585-594.

Wheeler, Donald J. (1999). *Understanding Variation: The Key to Managing Chaos*, 2nd Edition. SPC Press.

## Anexo A - Dados dos Projetos

**Tabela 4. Dados de Avaliações nos Projetos de Desenvolvimento de Software**

Projeto	ID	ProblemasGQ	ProblemasAP	ProblemasGQ Normalizado	ProblemasAP Normalizado
P1	A3	1	5	0,024390	0,121951
P1	A5	5	10	0,121951	0,243902
P1	A9	1	1	0,024390	0,024390
P1	A11	1	1	0,024390	0,024390
P2	A13	3	3	0,033708	0,033708
P2	A14	2	3	0,022472	0,033708
P2	A16	1	16	0,011236	0,179775
P2	A19	1	6	0,011236	0,067416
P3	A25	2	12	0,036664	0,219982
P3	A26	1	2	0,018332	0,036664
P3	A27	2	4	0,036664	0,073327
P3	A30	3	33	0,054995	0,604950
P3	A31	1	8	0,018332	0,146654
P5	A42	1	4	0,004868	0,019471
P5	A47	4	18	0,019471	0,087621
P5	A49	2	19	0,009736	0,092489