

Aprenda Java e Processo de software

Baseado no Personal Software Process (PSP) para Engenheiros - Parte 1

Programa 3

1. Requisitos do software

Regressão linear é uma maneira de ajustar de forma ótima uma linha para um conjunto de dados. A linha da regressão linear é a aquela em que a distância de todos os pontos para a linha é minimizada. A equação pode ser escrita como $y = \beta_0 + \beta_1 x$.

Considerando um conjunto de valores x (variável independente, geralmente uma medida facilmente obtida e anterior ao que se deseja estimar) e y (variável dependente, relativa ao resultado desejado), o cálculo dos coeficientes β_0 e β_1 é realizado com as seguintes fórmulas:

$$\beta_1 = \frac{\left(\sum_{i=1}^n x_i y_i \right) - (n x_{avg} y_{avg})}{\left(\sum_{i=1}^n x_i^2 \right) - (n x_{avg}^2)}$$

$$\beta_0 = y_{avg} - \beta_1 x_{avg}$$

Na Figura 1, o melhor ajuste para a linha de regressão linear possui os parâmetros $\beta_0 = -4,0389$ e $\beta_1 = 0,1681$.

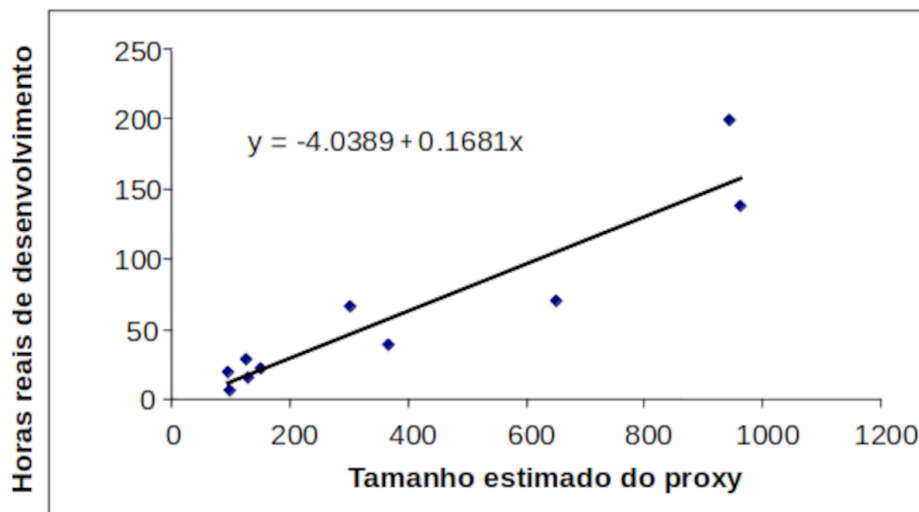


Figura 1. Gráfico da função da regressão linear referente à estimativa de esforço para desenvolvimento de software considerando o tamanho estimado do proxy em LOC.

No PSP, a regressão linear é utilizada para estimar o tamanho de um programa e o esforço necessário para desenvolver um programa. O método, denominado PSP PROBE, utiliza os parâmetros da regressão linear para fazer melhores previsões do tamanho e do tempo baseado nos dados históricos. Finalmente, o PSP PROBE possui quatro métodos internos para estimativa: A, B, C e D.

Por exemplo, observando a Figura 1, quantas horas você acha que são necessárias para desenvolver um programa com um proxy de tamanho estimado em 500 LOC? Utilizando o método A do PROBE para o cálculo do tempo, a estimativa seria calculada pela fórmula $y = \beta_0 + \beta_1 500$, ou seja, teríamos uma estimativa de 80,011 horas.

Os métodos PROBE A e B diferem apenas quanto aos dados históricos (valores x) utilizados para calcular os parâmetros da regressão linear. No método PROBE A, o tamanho estimado do proxy é utilizado como valores de x. No método PROBE B, o tamanho planejado de adições e modificações é utilizado como valores de x.

Os métodos PROBE para tamanho e tempo diferem apenas quanto aos dados históricos (valores y) utilizados para calcular os parâmetros de regressão linear. Para obter valores melhores de estimativa, as medidas de tamanho dos elementos efetivamente adicionados e removidos são utilizadas como valores de y (por exemplo, medidas de LOC adicionadas e removidas). Similarmente, para realizar previsões de tempo, valores efetivos de tempo de desenvolvimento são utilizados como valores de y.

Dado histórico utilizado		Valores de x	Valores de y
Estimativa de tamanho	PROBE A	Tamanho estimado do proxy	Tamanho efetivo de adições e modificações
	PROBE B	Tamanho planejado de adições e modificações	Tamanho efetivo de adições e modificações
Estimativa de tempo	PROBE A	Tamanho estimado do proxy	Tempo efetivo de desenvolvimento
	PROBE B	Tamanho planejado de adições e modificações	Tempo efetivo de desenvolvimento

Para analisar a qualidade do modelo de regressão linear, utiliza-se medidas de correlação. O cálculo da correlação determina a relação entre dois conjuntos de dados numéricos. Ela pode ser calculada com as seguintes fórmulas:

$$r_{x,y} = \frac{n \left(\sum_{i=1}^n x_i y_i \right) - \left(\sum_{i=1}^n x_i \right) \left(\sum_{i=1}^n y_i \right)}{\sqrt{\left[n \left(\sum_{i=1}^n x_i^2 \right) - \left(\sum_{i=1}^n x_i \right)^2 \right] \left[n \left(\sum_{i=1}^n y_i^2 \right) - \left(\sum_{i=1}^n y_i \right)^2 \right]}}$$

$$r^2 = r * r$$

em que:

- \sum é o símbolo do somatório,
- i é um índice para o número n ,
- x e y são dois conjuntos pareados de dados,
- n é um número de item contido em cada conjunto x e y ,
- x_{avg} é a média aritmética dos valores no conjunto x ,
- y_{avg} é a média aritmética dos valores no conjunto y .

A correlação $r_{x,y}$ pode variar de +1 até -1, sendo que:

- Resultados próximos de +1 indicam uma relação fortemente positiva, ou seja, se x aumenta, y também aumenta.
- Resultados próximo de -1 indicam uma relação fortemente negativa, ou seja, se x aumenta, y diminui.
- Resultados próximos de 0 indicam que não existe relação entre os conjuntos de dados.

A correlação é utilizada no PSP para julgar a qualidade da relação linear em vários conjuntos de dados históricos que são utilizados para o planejamento. Por exemplo, o método PROBE considera a relação entre o tamanho estimado do proxy e o tempo efetivo (PROBE A) ou a relação entre o tamanho de adições e modificações planejadas e o tempo efetivo (PROBE B). Para este propósito, consideraremos os valores da relação $r_{x,y}$ ao quadrado, ou seja, r^2 , interpretando-o conforme a tabela a seguir:

Se é r^2	a relação é
$0,9 \leq r^2$	preditiva; utilize-a com muita confiança
$0,7 \leq r^2 \leq 0,9$	forte e pode ser utilizada para planejamento
$0,5 \leq r^2 \leq 0,7$	adequada para planejamento, mas utilize-a com cautela
$r^2 < 0,5$	não é adequada para fins de planejamento

Observe que correlação não implica causa e efeito. Uma correlação forte pode ser coincidência. Por exemplo: De 1840 até 1960, nenhum presidente dos EUA eleito em um ano terminado em 0 sobreviveu durante seu mandato. Coincidência ou correlação? Muitas correlações acidentais podem ser encontradas ao processar dados históricos. Por isso, para usar uma correlação, você deve entender a relação causa-efeito do processo.

Estabelecido esses conceitos, construa um programa para:

- calcular os parâmetros de regressão linear β_0 e β_1 e os coeficientes de correlação $r_{x,y}$ e r^2 para um conjunto de n pares de dados,
- dada uma estimativa x_k , calcular uma previsão melhorada y_k em que $y_k = \beta_0 + \beta_1 x_k$
- melhorar o programa de lista encadeada desenvolvida no Programa 1 para armazenar os n pares de dados, em que cada registro armazena dois números reais.

Teste adequadamente o seu programa. A Tabela 1 contém históricos de tamanhos estimados, planejados e efetivos para dez programas. Os resultados esperados são apresentados na Tabela 2. No mínimo, teste o programa com os seguintes casos de teste:

- Caso de teste 1: Calcule os parâmetros de regressão e os coeficientes de correlação considerando os valores de tamanho estimado do proxy e de tamanho efetivo de adições e modificações apresentados na Tabela 1. Calcule a quantidade estimada de linhas adicionadas e modificadas considerando um proxy de tamanho estimado $x_k = 386$.
- Caso de teste 2: Calcule os parâmetros de regressão e os coeficientes de correlação considerando os valores de tamanho estimado do proxy e de tempo efetivo de desenvolvimento apresentados na Tabela 1. Calcule o tempo estimado considerando um proxy de tamanho estimado $x_k = 386$.
- Caso de teste 3: Calcule os parâmetros de regressão e os coeficientes de correlação considerando os valores de tamanho planejado de adições e modificações e de tamanho efetivo de adições e modificações apresentados na Tabela 1. Calcule o tamanho estimado das adições e modificações considerando um proxy de tamanho estimado $x_k = 386$.
- Caso de teste 4: Calcule os parâmetros de regressão e os coeficientes de correlação considerando os valores de tamanho planejado das adições e modificações e de tempo efetivo de desenvolvimento apresentados na Tabela 1. Calcule o tempo estimado considerando um proxy de tamanho estimado $x_k = 386$.

Número do programa	Tamanho estimado do proxy	Tamanho planejado de adições e modificações	Tamanho efetivo das adições e modificações	Tempo efetivo de desenvolvimento (em horas)
1	130	163	186	15,0
2	650	765	699	69,9
3	99	141	132	6,5
4	150	166	272	22,4
5	128	137	291	28,4
6	302	355	331	65,9
7	95	136	199	19,4
8	945	1206	1890	198,7
9	368	433	788	38,8
10	961	1130	1601	138,2

Tabela 1. Dados históricos sobre o desenvolvimento de dez programas.

Caso de teste	Valores esperados					Valores obtidos				
	β_0	β_1	$r_{x,y}$	r^2	y_k	β_0	β_1	$r_{x,y}$	r^2	y_k
1	-22,55	1,7279	0,9545	0,9111	644,429					
2	-4,039	0,1681	0,9333	0,8711	60,858					
3	-23,92	1,43097	0,9631	0,9276	528,4294					
4	-4,604	0,140164	0,9480	0,8988	49,4994					

Tabela 2. Dados previstos para o desenvolvimento de um programa considerando um proxy de 386 LOC.

2. Processo de software: Visão geral

2.1. Propósito

Guiar o desenvolvimento de programas modulares.

2.2. Critérios de entrada

- Descrição do problema.
- Formulário de Resumo de Planejamento de Projeto.
- Registro de Tempo.
- Registro de Erros.
- Padrão de estilo de codificação.

2.3. Atividades

- **Planejar**
 - Produzir ou obter requisitos.
 - Estimar o tempo necessário para desenvolvimento.
 - Preencher os dados de planejamento no formulário de Resumo de Planejamento de Projeto, **incluindo o tamanho previsto para as partes do software.**
 - Completar o Registro de Tempo.
- **Desenvolver**
 - Projetar o programa.
 - Implementar o projeto.
 - Compilar o programa, consertar e registrar todos os erros encontrados.
 - Testar o programa, **usando testes automatizados e critérios de teste**, consertar e registrar todos os erros encontrados.
 - Completar o Registro de Tempo.
- **Encerrar**

- Completar o formulário de Resumo de Planejamento de Projeto com os dados de tempo efetivo, erros e tamanho, **incluindo o tamanho efetivo das partes do software.**

2.4. Critérios de saída

- Programa rigorosamente testado, **considerando a satisfação de critérios de teste referente a com comandos e desvios.**
- Formulário de Resumo de Planejamento de Projeto preenchido com dados estimados e efetivos.
- Registro de Tempo completo.
- Registro de Erros completo.

3. Processo de software: Planejar

3.1. Propósito

Guiar o processo de planejamento do PSP.

3.2. Critérios de entrada

- Descrição do problema.
- Formulário de Resumo de Planejamento de Projeto.
- Registro de Tempo.

3.3. Atividades

- **Analisar requisitos**
 - Produzir ou obter os requisitos para o programa.
 - Garantir que os requisitos estão claros e sem ambiguidade.
 - Resolver qualquer questão quanto aos requisitos.
- **Definir partes do software**
 - **Desenhar o software, organizando-o em partes para fins de planejamento e análise.**
 - **Preencher os dados de tamanho das partes projetadas.**
- **Estimar recursos**
 - Fazer sua melhor estimativa do tempo necessário para desenvolver o programa.
 - Preencher os dados de tempo estimado/planejado no formulário de Resumo de Planejamento de Projeto.
 - Usando como guia os dados totais de tempo consumido por fase do projeto (To Date %), distribuir o tempo total de desenvolvimento para cada fase do projeto do programa atual.

3.4. Critérios de saída

- Requisitos documentados.
- Formulário de Resumo de Planejamento de Projeto preenchido com as estimativas de tempo de desenvolvimento **e o tamanho previsto para as partes do software.**
- Registro de Tempo completo.
- Registro de Erros completo.

4. Processo de software: Desenvolver

4.1. Propósito

Guiar o desenvolvimento de programas pequenos.

4.2. Critérios de entrada

- Requisitos.
- Registro de Tempo.
- Registro de Erros.
- Padrão de estilo de codificação.

4.3. Atividades

- **Desenhar / Projetar**
 - Revisar os requisitos e produzir um projeto para atendê-los.
 - Registrar no registro de erros qualquer erro encontrado nos requisitos.
 - Registrar o tempo no Registro de Tempo.
- **Programar**
 - Implementar o projeto, obedecendo o padrão de estilo de codificação.
 - Registrar no Registro de Erros qualquer erro encontrado nos requisitos e no projeto.
 - Registrar o tempo no Registro de Tempo.
- **Compilar**
 - Compilar o programa até que não existam mais erros de compilação.
 - Corrigir todos os erros encontrados.
 - Registrar erros no Registro de Erros.
 - Registrar o tempo no Registro de Tempo.
- **Testar**
 - **Criar casos de teste, considerando critérios de teste de software.**
 - Testar o programa até que os casos de teste executem sem erros **e satisfaça os critérios de teste.**
 - Corrigir todos os erros encontrados.
 - Registrar erros no Registro de Erros.
 - Registrar o tempo no Registro de Tempo.

4.4. Critérios de saída

- Programa rigorosamente testado, **considerando a satisfação de critérios de teste referente a com comandos e desvios**, e implementado de acordo com o padrão de estilo de codificação.
- Registro de Tempo completo.
- Registro de Erros completo.

5. Processo de software: Encerrar

5.1. Propósito

Guiar o processo de encerramento do PSP.

5.2. Critérios de entrada

- Descrição do problema e requisitos.
- Formulário de Resumo de Planejamento de Projeto **com dados de tempo de desenvolvimento do programa.**
- Registro de Tempo.
- Registro de Erros.
- Programa testado e executável **que está em conformidade com o padrão de estilo de codificação.**

5.3. Atividades

- **Garantir a consistência de dados de erros**
 - Verificar que todos os erros encontrados em cada fase foram realmente registrados.
 - Verificar se os dados de cada erro no Registro de Erros estão corretos e completos.
 - Verificar se a quantidade de erros inseridos e removidos em cada fase é razoável e correta.
 - Usando o melhor possível sua memória, corrigir qualquer dado omissos ou incorreto sobre os erros.
- **Garantir a consistência de dados de tempo**
 - Revisar o Registro de Tempo, procurando por erros e omissões.
 - Usando o melhor possível sua memória, corrigir qualquer dado omissos ou incorreto sobre os registros de tempo.
- **Medir tamanho do programa**
 - Medir o tamanho do programa completo **e de suas partes.**
 - Determinar o tamanho do código base, apagado, adicionado, reutilizado (que não seja o código base) e total.
 - Preencher os dados de tamanho no Formulário de Resumo de Planejamento de Projeto.

5.4. Critérios de saída

- Programa rigorosamente testado, **considerando a satisfação de critérios de teste referente a comandos e desvios**, e implementado de acordo com o padrão de estilo de codificação.
- Formulário de Resumo de Planejamento de Projeto completo. **incluindo o tamanho efetivo das partes do software.**
- Registro de Tempo completo.
- Registro de Erros completo.

6. Entrega

Quando você completar a fase de encerramento, envie os dados do pacote da tarefa, código fonte e resultados de teste para o professor.

O pacote de tarefa deve conter os seguintes a seguir, na ordem apresentada:

- Formulário de Resumo de Planejamento de Projeto,
- Registro de Tempo,
- Registro de Erros,
- Listagem do código fonte do Programa 3,
- Resultados dos testes.