

# Análise de Repetibilidade e Reprodutibilidade para Dados Atributo

(<http://www.statistical.com.br/artigos.asp>)

Carlos Hugo Domenech

## Em 50 palavras ou menos

- A etapa Medir é uma das mais frustrantes do DMAIC pela baixa qualidade das medições. Este problema é ainda maior para dados tipo atributo.
- A validação de dados tipo atributo pode ser feita realizando um estudo no qual se testa a capacidade dos analistas usando indicadores kappa ou Kendall.

Se agradece a colaboração de Geraldo Nilton e Juliano Nicácio da Santanense pela ajuda para a publicação do trabalho sobre inspeção de tecidos e a André Toriello da Motorola pela autorização da menção do projeto realizado nessa empresa que obteve economias significativas ao melhorar um processo de inspeção por atributos.

Histo DOE, nosso colega Black Belt está novamente na ativa. Desta vez tem que solucionar um problema da empresa de tecidos Santanense. A gerência da empresa está preocupada porque a venda de rolos de tecidos tintos está gerando reclamações de clientes e indenizações. Uma parte considerável desta perda poderia ser evitada se a inspeção final do tecido fosse eficiente.

Embora o Seis Sigma recomenda fazer melhorias no processo e não na inspeção final, a empresa decidiu atacar de forma agressiva os custos das falhas que chegam ao cliente para conter a “sangria” perdida pelas indenizações. Há outros projetos em paralelo para a prevenção dos defeitos.

Note que o mesmo problema aparece em qualquer tipo de inspeção subjetiva realizada pelo homem (visão, audição, olfato, gosto e tato). Alguns exemplos destacam-se a seguir:

- Quando os inspetores analisam o display de um telefone digital, podem não enxergar alguns pixels inativos. Na Motorola, um projeto Seis Sigma conduzido por André Toriello na área de inspeção, conseguiu fazer melhorias que permitiram economizar milhares de dólares.
- Os inspetores de bagagens das máquinas de raios X podem deixar de perceber a passagem de itens proibidos. Juran and Godfrey (1998) mencionam o caso de um auditor que em um aeroporto de USA conseguiu passar uma granada de mão pelo sistema de monitorização.
- Na área de serviços, diversos técnicos de serviços de alarme que são chamados para resolver problemas de ativação falsa do sistema poderiam apresentar concordância baixa entre si, sobre a causa da ativação. Embora não seja evidente, aqui também temos um sistema de medição em que o “aparelho” é o técnico e ele decide qual é o defeito do sistema.
- Em software são feitas inspeções de linhas de código dos programas fontes. Para aumentar a eficiência os inspetores trabalham em pares (“peer review”) e se eles não estão bem “calibrados”, a taxa de defeitos pode superar os 40%.

Quando Histo DOE começou a estudar o problema da inspeção de tecidos ele não podia acreditar que defeitos de tamanho superior a 23 cm pudessem ter escapado ao processo de inspeção 100% e ter sido detectados somente pelo cliente. Porém ele já tem bastante experiência com sistemas de medição com dados contínuos e após inúmeros estudos de repetibilidade e reprodutibilidade (RR), tem já uma estatística própria que diz que: em quase todo tipo de empresa, entre 50-100% dos sistemas de medição não são apropriados!

Ele sabe também que a etapa Medição é uma das mais complicadas do ciclo DMAIC. Agora ele vai aplicar um estudo RR para dados atributo. A primeira atividade que ele teve com a equipe foi uma visita às mesas de inspeção, nas quais os rolos de tecido, de aproximadamente 1,6 m de largura são inspecionados a uma velocidade em torno de 30-40 m/min. Ele percebeu que a luminosidade não era adequada, o lugar era muito barulhento para que os inspetores (80 em vários turnos), pudessem se concentrar. Além disso viu que os inspetores ficavam 8 horas seguidas (com uma hora de intervalo para almoço) olhando o tecido passar! Ele pensou “agora entendo como podem escapar defeitos tão grandes”.

### ***A estratégia da equipe de Histo DOE***

Foi desenhado o estudo da Figura 1:

- Seis amostras com representação de todos os tipos de defeitos foram inspecionadas...
- duas vezes (em momentos diferentes) por 9 funcionários. A velocidade de inspeção foi similar à usada normalmente.
- Um supervisor considerado especialista no assunto, inspecionou os 6 rolos em condições ideais de iluminação e com velocidade reduzida. Isto permitiu criar o “mapa de defeitos” dos 6 rolos.

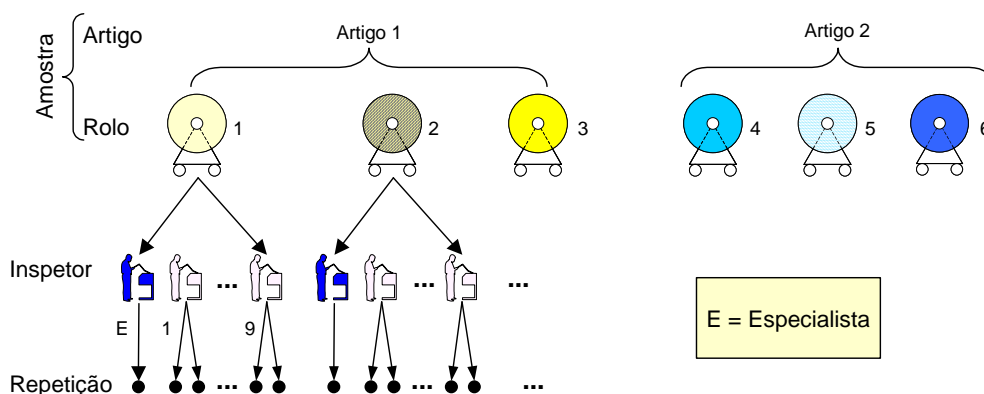


Figura 1 – Esquema do estudo RR atributo planejado pela equipe de Histo DOE

### ***Problemas da estratégia (nunca faltam!):***

- Histo DOE se defrontou com o seguinte problema: não posso inspecionar os mesmos rolos mais de uma vez porque os inspetores vão perceber. A chave deste tipo de estudo (e também dos estudos RR com dados contínuos) é que seja feito de maneira “cega”. Os funcionários não devem saber que estão sendo avaliados.
- Como fazer que os funcionários anotem os defeitos em uma folha separada do sistema normal? Além disso quando eles detectam um defeito tipo IV (de um tamanho superior a 23 cm), param o rolo e cortam o pedaço defeituoso. No ensaio os inspetores não poderiam fazer isso! Eles perceberiam que estão sendo observados e mudariam seu nível de atenção.
- Histo DOE assumiu que existia esse risco, mas seguiu em frente. Escolheram que os testes seriam feitos no período da tarde (o pessoal estaria mais cansado).

### Resultados:

As variáveis medidas em um estudo RR atributo podem ser de dois tipos:

- Atributo nominal: registra-se o tipo de defeito em uma escala qualitativa que não é passível de ordenação. Exemplo: fio grosso no tecido, manchas.
- Atributo ordinal: o defeito registra-se usando uma variável ordinal. Exemplo: defeito de tamanho I, II, III e IV (como no exemplo dos tecidos), ou contando o número de defeitos de um certo tipo.

Neste caso o estudo se estruturou de forma que eles registrassem 4 variáveis: D1, D2, D3 e D4 que representavam o número de defeitos de tipo I, II, III e IV respectivamente. As variáveis E1, E2, E3 e E4 indicavam o número de defeitos registrado pelo especialista. Note que o valor do especialista não muda quando um inspetor diferente analisa o mesmo rolo.

Na Tabela 1 apresenta-se uma parte dos dados, somente para o defeito 4. Os dados completos do estudo estão no link <http://www.statistical.com.br/admin/arquivos/RRSantanense.xls>. Como o defeito 4 é o mais evidente, esperava-se que houvesse uma correlação boa entre as medidas dos inspetores e do especialista. Na Figura 2 foi feito um diagrama de dispersão dos dados do especialista versus os dados dos inspetores. Idealmente os resultados deveriam ficar agrupados em torno da reta a 45°, ou seja quando o especialista declarou 2 defeitos, os inspetores deveriam ter encontrado também 2 defeitos. É evidente que a associação é muito fraca. Observe que em geral os inspetores encontraram menos defeitos que o especialista (maior concentração de pontos abaixo da reta): são os defeitos que geram as reclamações!

Tabela 1 – Resultados parciais do ensaio RR (somente dados dos inspetores 1 e 2)

Amostra	Inspetor	Repetição	D4	E4
1	1	1 <sup>a</sup>	1	9
2	1	1 <sup>a</sup>	3	7
3	1	1 <sup>a</sup>	1	4
4	1	1 <sup>a</sup>	1	3
5	1	1 <sup>a</sup>	0	2
6	1	1 <sup>a</sup>	0	3
1	1	2 <sup>a</sup>	1	9
2	1	2 <sup>a</sup>	5	7
3	1	2 <sup>a</sup>	1	4
4	1	2 <sup>a</sup>	2	3
5	1	2 <sup>a</sup>	1	2
6	1	2 <sup>a</sup>	0	3
1	2	1 <sup>a</sup>	10	9
2	2	1 <sup>a</sup>	8	7
3	2	1 <sup>a</sup>	3	4
4	2	1 <sup>a</sup>	2	3
5	2	1 <sup>a</sup>	2	2
6	2	1 <sup>a</sup>	2	3
1	2	2 <sup>a</sup>	6	9
2	2	2 <sup>a</sup>	5	7
3	2	2 <sup>a</sup>	4	4
4	2	2 <sup>a</sup>	2	3
5	2	2 <sup>a</sup>	3	2
6	2	2 <sup>a</sup>	3	3

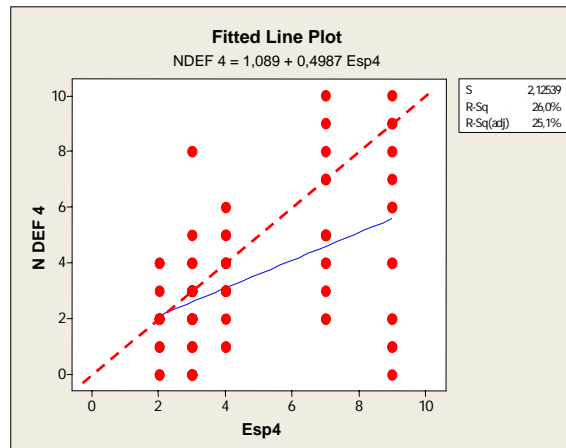


Figura 2 – Diagrama de dispersão entre as medidas dos inspetores e do especialista

Os dados foram analisados no software Minitab 14.12 (procedimento Stat\Quality Tools\Attribute Gage R&R Study). A partir deste tipo de experimento é possível extrair as seguintes conclusões:

- Grau de associação (atributo ordinal) ou concordância (atributo nominal) entre as medidas repetidas de cada inspetor. Análogo ao conceito de repetibilidade em um estudo RR com dados contínuos.
- Grau de associação ou concordância entre as medidas realizadas por diferentes analistas. Análogo ao conceito de reprodutibilidade em um estudo RR com dados contínuos.
- Grau de associação ou concordância entre cada analista e o especialista. Dá uma idéia da “exatidão” dos analistas; isto é, como estão “aferridos” em relação ao padrão (especialista).

As métricas calculadas para cada tipo de variável (nominal e ordinal) e as conclusões que se extraem a partir delas estão na Tabela 2. Valores dos coeficientes próximos de zero indicam pouca concordância ou associação entre as medidas. Valores próximos de 1 indicam boa concordância ou associação, enquanto que valores grandes e negativos – embora não serem comuns – podem indicar interpretação errônea dos defeitos pelos inspetores.

Tabela 2 – Métricas do estudo RR atributo

Variável atributo	Coefficiente	Conclusão
Nominal	Concordância kappa ( $\kappa$ )	Se $\kappa > 0,9$ ⇒ Ótimo
		Se $0,7 < \kappa < 0,9$ ⇒ Aceitável (a melhorar)
		Se $\kappa < 0,70$ ⇒ Medida não adequada
Ordinal	Correlação Kendall ( $\tau$ )	Se $\tau > 0,95$ ⇒ Ótimo
		Se $0,8 < \tau < 0,95$ ⇒ Aceitável (a melhorar)
		Se $\tau < 0,80$ ⇒ Medida não adequada

Histo DOE obteve os resultados da Tabela 3 para o projeto da Santanense.

Tabela 3 – Coeficientes de correlação de Kendall

Inspetor	Dentro de Inspetor (repe)	Entre Inspetores (Repro)	Inspetores vs Especialista (Exatidão)	Todos vs Especialista (Exatidão)
1	0,91	0,52	0,36	0,58
2	0,98		0,89	
3	1,00		0,93	
4	0,72		0,67	
5	0,35		-0,18	
6	0,78		0,72	
7	0,81		0,57	
8	0,97		0,93	
9	0,30		0,35	

### Conclusões do RR de inspeção:

- Dos nove inspetores que participaram, somente 3 (33%) foram considerados dentro de um padrão aceitável. A maioria dos inspetores apresentou pouca correlação com eles mesmos, com os outros inspetores ou com o especialista.
- A partir destes dados a equipe coordenada pelo Black Belt iniciou uma série de ações:
  - » Exame oftalmológico: vários funcionários apresentaram problemas de visão, sendo que alguns precisavam de óculos com urgência, problemas na identificação das cores verde e vermelho e dois apresentaram daltonismo.
  - » Foram idealizados treinamentos do tipo gincana ou quebra-cabeças para estimular a atenção e concentração. Vários funcionários apresentaram problemas em testes de rapidez e qualidade (reconhecimento de defeitos).
  - » Foi utilizada a norma NBR-5413 para adequar a luminosidade nos locais de inspeção. Consultou-se a norma NR-17 para determinar períodos adequados de descanso.
  - » Foi diminuída a velocidade de inspeção para 10-20 m/min.
- Foram desenhadas auditorias periódicas para avaliar o estado do sistema (atividade importante da etapa Controle do ciclo DMAIC). Para se ter uma idéia da melhoria, um diagrama de dispersão construído com dados destas auditorias apresentou os resultados da Figura 3. O “grau de explicação” ( $r^2$ , máximo 100%) passou de 27% na Figura 2 para 91% na Figura 3.
- Todos estes resultados contribuíram para melhorar de forma significativa o processo de inspeção e reduzir a porcentagem de defeitos encontrados pelos clientes.

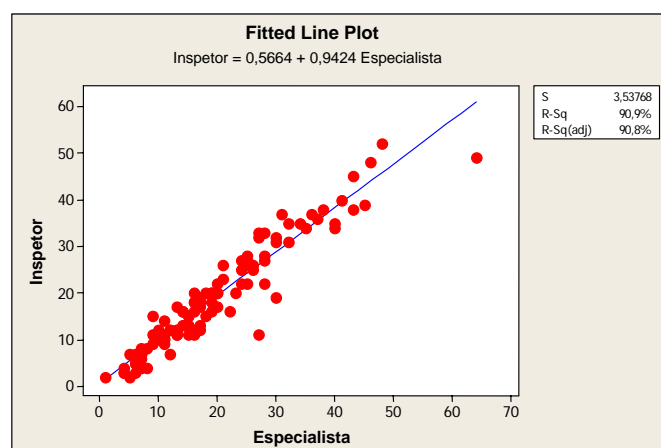


Figura 3 – Dados levantados nas auditorias. Correlação entre Inspetores e Especialista

**Conclusões sobre os estudos RR para dados atributo:**

- Os estudos RR para dados atributo são uma ferramenta fundamental quando se trabalha em processos de inspeção e similares.
- Em muitos casos os donos dos processos nem sequer percebem da falta de qualidade de seus sistemas de medição. Quando se trabalha com dados atributo, isto pode ser particularmente sério como se viu no exemplo da empresa de tecidos ou nas melhorias realizadas pela Motorola que permitiu uma economia de milhares de dólares.

**REFERÊNCIAS BIBLIOGRÁFICAS:**

- Futrell, David (1995). When quality is a matter of taste, use reliability indexes. Quality Progress, May, p. 81-86.
- Juran and Godfrey (1998). Juran's Quality Handbook. 5. ed. New York: McGraw-Hill. ISBN: 007034003X.
- NBR-5413 – Iluminância de interiores. 1992.
- NR-17 – Ergonomia. Ministério do Trabalho. 1990.

**Carlos Domenech** é diretor da M. I. Domenech, empresa de métodos avançados e soluções Seis Sigma. Mestre em estatística (IME/USP). Trabalhou durante 10 anos na Rhodia. Certificado pela Motorola (Texas) como instrutor de Black Belt. Especialização em DOE no Centro da Tecnologia da DuPont - Wilmington. [mi.domenech@statistical.com.br](mailto:mi.domenech@statistical.com.br).