

Crítica à definição de intervalo de confiança, predição e campo de arbítrio na NBR

14.653-02

Luiz Fernando Palin Droubi*

Carlos Augusto Zilli†

Willian Zonato‡

Norberto Hochheim§

14/08/2019

1 ESTUDOS DE CASO

1.1 Dados reais de mercado

Neste estudo são utilizados os dados disponíveis em Hochheim (2015, pp. 21–22). O modelo utilizado também foi o obtido no trabalho citado. Para este estudo de caso foi utilizado o software R 3.5.1 (R CORE TEAM, 2019).

Um resumo dos coeficientes e estatísticas do modelo pode ser visto na tabela 1.

Como o modelo foi ajustado com a transformação logarítmica natural, os valores ajustados devem ser retransformados para a escala original. Aqui se adotará a transformação pela mediana da distribuição lognormal, como em Hochheim (2015).

Na escala log, o valor central obtido foi de 13,78. Pode-se mostrar que o erro-padrão do ajuste ($s.e.(\hat{Y})$) para o bem-avaliando (obtido de acordo com a expressão $\sqrt{x_0' \hat{\sigma}^2 (X'X)^{-1} x_0}$) é de $\approx 0,03$ (também na escala log). O valor de $t_{41}^{10\%}$ é ≈ -1.3025434 e $t_{41}^{10\%}$ é ≈ 1.3025434 . Substituindo-se nas expressões abaixo, pode-se calcular os limites do IC para a média, já que é sabido que o intervalo de confiança @80% foi calculado com a distribuição t com 41 graus de liberdade:

$$\hat{Y}_{sup} = \hat{Y} + t_{41}^{10\%} s.e.(\hat{Y}) \text{ e } \hat{Y}_{inf} = \hat{Y} + t_{41}^{90\%} s.e.(\hat{Y})$$

Ou seja:

$$\hat{Y}_{sup} = 13,78 + 1,30.0,03 \text{ e } \hat{Y}_{inf} = 13,78 - 1,30.0,03$$

Caso se opte por arbitrar o valor da estimativa central (R\$961.660,64) e o intervalo de confiança, a avaliação intervalar fica limitada pelo último, ou seja, entre R\$924.768,13 e R\$1.000.024,94 e o campo de arbítrio ([R\$817.411,55; R\$1.105.909,74]) não pode ser utilizado.

*SPU/SC, luiz.droubi@planejamento.gov.br

†UFSC, carloszilli@gmail.com

‡SPU/SC, willian.zonato@planejamento.gov.br

§UFSC, hochheim@gmail.com

Tabela 1: Coeficientes e estatísticas do modelo adotado.

	Estimate	CI (lower)	CI (upper)	Std. Error	t value	Pr(> t)	
(Intercept)	13.564	13.264	13.864	0.230	58.847	<0.001	***
area_total	0.001	0.001	0.002	0.000	5.113	<0.001	***
quartos	0.164	0.118	0.210	0.035	4.626	<0.001	***
suites	0.061	0.017	0.105	0.034	1.810	0.078	.
garagens	0.209	0.165	0.252	0.033	6.247	<0.001	***
log(dist_b_mar)	-0.141	-0.176	-0.105	0.027	-5.174	<0.001	***
rec(padrao)	-0.563	-0.700	-0.426	0.105	-5.360	<0.001	***

No entanto, caso seja arbitrado o valor de R\$1.100.000,00 para o bem-avaliando, como em Hochheim (2015, pp. 73–74), os novos limites intervalares ficam entre R\$1.060.973,75 e R\$1.105.909,74.

Transformando-se estes valores para a escala log-natural, tem-se que $13,87 \leq 13,91 \leq 15,84$.

Ora, se sabe-se o valor do erro-padrão naquele ponto e o valor central do IC, pode-se calcular o valor de t para os novos valores:

$$t_{inf} = 3,27$$

$$t_{arbitrado} = 4,48$$

$$t_{sup} = 68,81$$

Ora, de posse dos t calculados acima, pode-se calcular as suas probabilidades de ocorrência: 0,001084, 0,00002977 e 0,00. Ou seja, de acordo com o IC originalmente calculado, o valor arbitrado, o valor inferior e o valor calculado tem probabilidade baixíssima de ocorrer.

Isso, contudo, apenas significa que os valores inferior, arbitrado e superior encontram-se muito acima do intervalo de confiança para a *média*, ou seja, significa apenas que os valores citados, “arbitrados” são superiores à média do mercado.

Se, por outro lado, considera-se o intervalo de predição:

$$\hat{Y}_{inf} = 13,59 \quad \hat{Y}_{sup} = 13,96$$

Na escala original, $802.017,63 \leq 961.660,64 \leq 1.153.080,88$.

Porém, neste caso o IP estaria limitado pelo campo de arbítrio, pois este está no intervalo [R\$ 817.411,55, R\$ 1.105.909,74].

Da mesma maneira que foi feito para o IC, pode-se calcular a que probabilidade está associado o campo de arbítrio em relação ao IP:

$$t_{inf-CA} = -1,17$$

$$t_{sup-CA} = 1,00$$

Ou seja, de acordo com o IP, o limite inferior do CA encontra-se no percentil 12,50% e o limite superior do CA no percentil 83,90%, de maneira que, portanto, o C.A. não pode ser inteiramente utilizado, apesar de seus limites se encontrarem em percentis confortáveis do IP.

A fórmula para o cálculo do intervalo de predição pode ser escrita da seguinte forma:

$$\hat{y}_0 \pm t_{n-k}^{\alpha/2} \sqrt{\sigma^2 + s.e.^2(\hat{y}_o)}$$

Deve-se notar que, fixado o valor central \hat{y}_0 e a amplitude do intervalo, pode-se calcular o valor do termo $\sqrt{\sigma^2 + s.e.^2(\hat{y}_o)}$ em função do t especificado. No limite, sabe-se que a distribuição t aproxima-se assintoticamente da distribuição normal, que será adotada aqui preliminarmente, por simplicidade.

1.2 Dados randômicos

Imagine-se que em um determinado modelo previamente ajustado, foi estimado o valor de um bem em R\$1.000.000,00. Imagine-se então que o modelo possui graus de liberdade suficiente para que o intervalo de predição possa ser assumido normalmente distribuído. Isto implica que o termo $\sqrt{\sigma^2 + s.e.^2(\hat{y}_o)}$ possui os seguintes valores, dependendo da amplitude do IP calculado:

Amplitude IP	$\sqrt{\sigma^2 + s.e.^2(\hat{y}_o)}$	Semi- amplitude do CA	Z_{inf}	Z_{sup}	p_{inf}	p_{sup}
30%	117.045,62	150.000	-1,28	1,28	10,00%	90,00%
40%	156.060,83	150.000	-0,96	0,96	16,80%	83,20%
50%	195.076,04	150.000	-0,77	0,77	22,10%	77,90%

O mesmo pode ser feito para a distribuição t com 40 graus de liberdade, por exemplo:

Amplitude IP	$\sqrt{\sigma^2 + s.e.^2(\hat{y}_o)}$	Semi- amplitude do CA	t_{inf}	t_{sup}	p_{inf}	p_{sup}
30%	115.112,15	150.000	-1,30	1,30	10,00%	90,00%
40%	153.482,87	150.000	-0,98	0,98	16,70%	83,30%
50%	191.853,58	150.000	-0,78	0,78	21,90%	78,10%

Ou ainda com 10 graus de liberdade:

Amplitude IP	$\sqrt{\sigma^2 + s.e.^2(\hat{y}_o)}$	Semi- amplitude do CA	t_{inf}	t_{sup}	p_{inf}	p_{sup}
30%	109.314,81	150.000	-1,37	1,37	10,00%	90,00%
40%	145.753,09	150.000	-1,03	1,03	16,40%	83,60%
50%	182.191,36	150.000	-0,82	0,82	21,50%	78,50%

Pode-se notar nas tabelas acima que a influência do número de graus de liberdade existe, mas é moderada, caso $(T - k) \geq 10$. Em todos os casos, a variação mais relevante é a das diferentes amplitudes dos intervalos de predição, *i.e.*, importa mais a variabilidade populacional.

Já o impacto da variabilidade do mercado pode ser vista graficamente na figura 1, onde

os pontos e a linha em vermelho representam o valor ajustado e os limites do Campo de Arbítrio, a área em azul claro representa o intervalo de predição @ 80%, a reta azul é a reta de regressão em relação aos dados amostrais e, finalmente, a área cinza representa o intervalo de confiança para esta regressão, para vários cenários hipotéticos.

Primeiramente, foi imaginado um cenário hipotético de mercado equilibrado, onde, por acaso, o Campo de Arbítrio se ajusta perfeitamente à variabilidade do mercado e poderia ser utilizado sem restrições. No segundo cenário, apenas foi considerado um intervalo mais amplo de validade para o modelo, porém com mesma variabilidade. No terceiro cenário, chamado de “Parado”, foi considerado um mercado com baixíssima variabilidade, para ilustrar o fato que, nestes casos, o Campo de Arbítrio pode superar os limites populacionais em ambos os sentidos. Finalmente, no quarto cenário, foi imaginado um mercado aquecido, com alta variabilidade, onde os limites do Campo de Arbítrio de $\pm 15\%$ não chegam nem à metade do intervalo de predição.

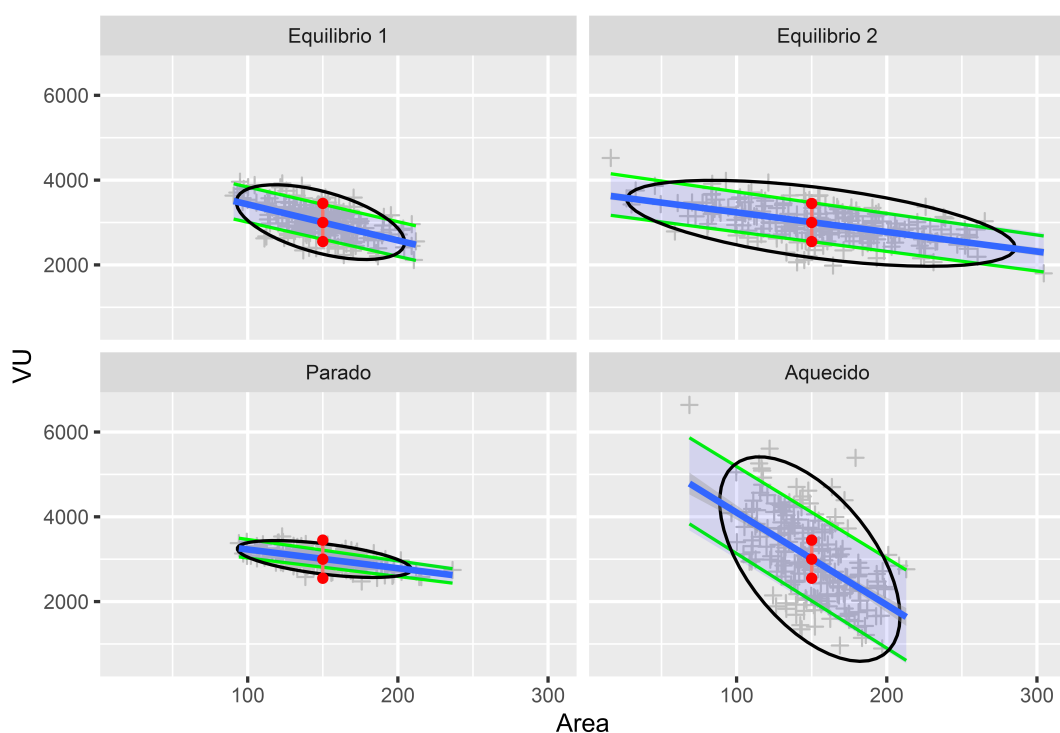


Figura 1: Comparação do campo de arbítrio em diferentes realidades mercadológicas.

Como se pode notar, é possível que o campo de arbítrio seja um bom parâmetro para o arbítrio do valor de mercado de um imóvel, porém isso só ocorrerá se, por acaso, a variabilidade do mercado for da mesma magnitude da amplitude do campo de arbítrio.

Se, contudo, o mercado tiver menos variabilidade a utilização dos limites do C.A. estará sub ou superavaliando o bem-avaliando para aquele mercado. Já no caso de um mercado com maior variabilidade, a utilização dos limites do C.A. pode ser insuficiente para dar ao avaliador a possibilidade de arbitrar valores nas faixas mais distantes da média do mercado, apesar destes imóveis serem parte da realidade do mercado.

Os coeficientes de determinação dos vários cenários estudados podem ser vistos na tabela abaixo:

13,8155

$$\hat{\sigma} =$$

0,1003

Então, com este modelo, o avaliador teria obtido os seguintes valores com o seu software de avaliação:

CA/IP	Moda	Mediana	Média
	990.000,00	1.000.000,00	1.005.037,82
CA_{inf}	841.500,00	850.000,00	854.282,14
CA_{sup}	1.138.500,00	1.150.000,00	1.155.793,49
IP_{inf}	870.639,20	879.433,53	883.863,96
IP_{sup}	1.125.724,64	1.137.095,60	1.142.824,08

Por simplicidade, aqui foi admitido que o erro na estimativa de \hat{y} é zero ($se(\hat{y}) = 0$). Com esta hipótese, o cálculo do Intervalo de Predição ficou resumido a:

$$IP = \hat{Y} \pm Z_{90}\hat{\sigma}$$

Pode ser calculado a que percentil da distribuição lognormal corresponde cada valor adotado. Oos resultados podem ser vistos na tabela abaixo:

CA/IP	Moda	Mediana	Média
	46,00%	50,00%	52,00%
CA_{inf}	4,26%	5,25%	5,81%
CA_{sup}	90,20%	91,80%	92,60%
IP_{inf}	8,35%	10,00%	10,90%
IP_{sup}	88,10%	90,00%	90,90%

Pode-se notar que a adoção da estimativa pela moda e do CA inferior corresponde ao percentil 4,26%, ou seja, um percentil muito baixo da distribuição encontrada pelo modelo.

No outro extremo, está a adoção do limite superior do CA e da estimativa pela média, encontrando-se num percentil extremamente alto da distribuição encontrada (92,60%).

Mesmo para a mediana, neste caso, a utilização dos valores extremos do CA leva a percentis muito extremos da distribuição de probabilidades (5,25% e 91,80%).

Enquanto isto, a adoção da estimativa pela mediana e a utilização do IP leva aos percentis que consideram-se adequados, ou seja, os percentis 10% e 90%.

1.3.1 Avaliação Intervalar segundo a atual NBR14.653-02

Segundo a atual NBR14.653-02, o avaliador teria várias alternativas para a definição da estimativa final, dentre as quais elencam-se:

1. Adoção da estimativa pela moda e do limite inferior do CA:

Se o avaliador optasse pela moda, adotando o limite inferior do CA, o intervalo final da avaliação, segundo a atual normativa, seria: [841.500,00; 977.224,64]. Ou seja, fazendo-se esta escolha, o intervalo final obtido estaria entre os percentis 4,26% e 40,90%.

2. Adoção da mediana, com limite inferior do CA:

Se o avaliador adotasse a mediana como estimativa de tendência central (lembre-se que para o cálculo da mediana não se utiliza o valor de $\hat{\sigma}$), e o limite inferior do CA, o avaliador teria como intervalo final: [850.000,00; 987.095,60]. Ou seja, fazendo-se esta escolha, o intervalo final obtido estaria entre os percentis 5,25% e 44,80%.

3. Simples adoção da mediana:

Se o avaliador adotasse simplesmente a mediana, teria como intervalo final da avaliação: [879.433,53; 1.137.095,60]. Ou seja, fazendo-se esta escolha, o intervalo final obtido estaria entre os percentis 5,25% e 91,80%.

4. Adoção da mediana, com limite superior do CA:

Se o avaliador adotasse a mediana como estimativa de tendência central e o limite superior do CA, o avaliador teria como intervalo final: [1.029.433,53; 1.150.000,00]. Ou seja, fazendo-se esta escolha, o intervalo final obtido estaria entre os percentis 61,40% e 91,80%.

5. Adoção da média, com limite superior do CA:

Finalmente, se o avaliador optasse pela média e o intervalo superior do CA, ele teria como intervalo final: [1.034.619,63; 1.155.793,49]. Ou seja, fazendo-se esta escolha, o intervalo final obtido estaria entre os percentis 63,30% e 92,60%.

2 CONCLUSÃO

A NBR14.653-01 (2001) e NBR14.653-02 (2011) fariam melhor em definir valor médio ($\mu(Y|X = t)$), intervalo de confiança para a média ($IC(\mu(t))$), valor arbitrado (\hat{Y}) e intervalo de predição para os valores arbitrados.

O avaliador deveria ter o poder de arbitrar valores para os bem-avaliandos dentro do intervalo de predição! Claro, pois o intervalo de confiança é para a média, mas o avaliador deveria ter condições de dizer se o bem-avaliando encontra-se acima ou abaixo dela. Por outro lado, o avaliador nunca deveria ter o poder de arbitrar valores para a média fora do IC! Isto é absurdo, pois foi feito uma inferência para a estimação da média e seu intervalo de confiança, mas a NBR14.653-2 permite ao avaliador arbitrar que a média encontra-se fora do IC! Isto não faz qualquer sentido científico.

O Campo de Arbítrio do avaliador não é um bom parâmetro para a arbitragem de valores, dado que o mesmo é constante ($\pm 15\%$), independente da variabilidade do mercado. Ora, como se mostrou no estudo de caso, com um modelo muito bem ajustado, causa espécie que o valor do CA fique restrito aos percentis 12,50% e 83,90% do IP. Desta maneira se está restringindo desnecessariamente o avaliador naquele intervalo, sendo que o avaliador poderia ter maior flexibilidade. Imagine-se o caso de um modelo menos ajustado, onde o IP seja de aproximadamente 50%: pode ocorrer de o C.A. ficar entre percentis muito próximos à média, ou seja, sabendo-se

que há uma grande variabilidade no mercado, restringe-se o avaliador àqueles 15%, que muitas vezes é pouco. Por que não se restringir o avaliador aos percentis 10% e 90% do IP?

A NBR14.653-02 tem um problema conceitual: ela mistura o conceito de valor médio e intervalo de confiança para a média, com o conceito de valor ajustado e intervalo de predição.

O valor médio e o intervalo de confiança para a média é de pouco interesse para o avaliador: estes conceitos tem valor mais explicativo. Pode-se utilizar o valor médio e seu intervalo de confiança para responder a seguinte pergunta: qual o valor médio de um apartamento com tais características naquele mercado? Para responder a esta pergunta de uma maneira qualitativa, deve-se utilizar o conceito de intervalo de confiança: um apartamento com tais características, naquela localidade, vale em média R\$3.000,00/m² \pm R\$500,00 ao nível de confiança de 80%. A simples resposta que um apartamento de tais características vale em média R\$3.000,00/m² não é uma boa resposta, pois não descreve a variabilidade do mercado. O IC dá uma idéia da variabilidade em torno da média.

Existem vários ramos da ciência que se valem destes conceitos, como é o caso da medicina: “o modelo comprovou que pessoas com IMC entre 25 e 30 vivem em média 10 anos a mais do que pessoas com IMC>35, ao nível de 90%”. Apesar de ser uma média para a população, dificilmente alguém poderia dizer quantos anos viverá uma determinada pessoa com IMC igual ou superior a 35, contudo, pois a variabilidade é muito maior.

Na Engenharia de Avaliações, contudo, apenas raramente se está interessado nestas funções descritivas da inferência clássica. Mais frequentemente se está interessado em conhecer, à partir de uma amostra, o valor para um novo elemento do mercado, não o valor médio do mercado de bens com aquelas características. Por isso o avaliador é obrigado a fazer uma vistoria no bem-avaliando, já que então ele pode intuir se o imóvel estará acima ou abaixo da média para aquele mercado. Se fosse de outra maneira, não seria necessário ao avaliador realizar a vistoria, bastaria ele dizer que em média tal imóvel vale R\$3.000,00 o metro quadrado \pm R\$500,00 @80%.

REFERÊNCIAS

ABNT. **NBR 14653-1: Avaliação de bens – parte 1: Procedimentos gerais**. Rio de Janeiro: Associação Brasileira de Normas Técnicas, 2001.

ABNT. **NBR 14653-2: Avaliação de bens – parte 2: Imóveis urbanos**. Rio de Janeiro: Associação Brasileira de Normas Técnicas, 2011.

HOCHHEIM, N. **Engenharia de avaliações - módulo básico**. Florianópolis: IBAPE - SC, 2015.

R CORE TEAM. **R: A language and environment for statistical computing**. Vienna, Austria: R Foundation for Statistical Computing, 2019.