CRÍTICA À AVALIAÇÃO INTERVALAR NA NBR 14.653-2

RESUMO

Em estatística são amplamente conhecidos e utilizados dois tipos de intervalos, quais sejam: o intervalo de confiança e o intervalo de predição. Enquanto o intervalo de confiança está relacionado à probabilidade de um intervalo de valores conter um parâmetro populacional, por exemplo a média da população, mas não somente (podem ser criados intervalos para a mediana ou para outros percentis da população, através da regressão quantílica), o intervalo de predição está associado à probabilidade de um intervalo conter o valor da variável resposta de um novo dado da população, não constante da amostra utilizada para a confecção do modelo. A NBR 14.653-2, em seu anexo A, traz considerações sobre o uso de intervalos para a avaliação intervalar, apontando os casos em que se utiliza o intervalo de confiança (IC) ou intervalo de predição (IP), além das limitações de uso destes intervalos, em função do campo de arbítrio do avaliador. Nesse estudo pretende-se demonstrar que este procedimento de avaliação intervalar contido no anexo A da NBR 14.653-2 não possui coerência científica, levando os avaliadores a confusão quanto a utilização de um ou outro tipo de intervalo, além de possibilitar ou limitar, injustificadamente, a arbitragem de valor do bem-avaliando, segundo critérios científicos.

Palavras-Chave: *Avaliação Intervalar; Campo de Arbítrio; Intervalo de Confiança; Intervalo de Predição.*

1 INTRODUÇÃO

A avaliação de imóveis baseia-se em fatos e acontecimentos que influenciam, em cada momento, o resultado final do valor do imóvel, convindo, sempre que possível, não se ater a um único aspecto da questão mas, pelo contrário, considerar simultaneamente os fatores custo e utilidade. Avaliar é, pois, estimar o valor de mercado de um ou mais interesses identificados em uma parcela específica de um imóvel, em um determinado momento (ABUNAHMAN, 2008).

Quando se avalia um imóvel, se busca estimar o valor de mercado dele e, segundo Dantas (2012), essa estimação pode ser pontual ou intervalar.

Conforme Gujarati e Porter (2011, p. 128), na estatística, a confiabilidade de um estimador pontual é medida por seu erro padrão. Em vez de tomarmos como base apenas a estimativa pontual, podemos construir um intervalo em torno de um estimador pontual, por exemplo, de dois ou três erros padrão de cada lado do estimador pontual, de modo que esse intervalo tenha, por exemplo, 95% de probabilidade de incluir o verdadeiro valor do parâmetro. Essa é, de certa forma, a ideia que está por trás do que se chama de estimação intervalar.

Em avaliação de imóveis, a estimação será pontual quando o resultado da estimativa é um único valor, como quando se utiliza a média aritmética da amostra para estimar a média da população. Com base na média e no desvio padrão amostral faz-se a estimativa intervalar, que corresponde a um conjunto de pontos dentro de determinados limites, no qual se pode afirmar com determinada probabilidade que a média da população está nele contido (DANTAS, 2012).

A norma de avaliação de imóveis, NBR 14.653-2, no seu anexo A, traz, em seu item A.10, considerações importantes sobre a chamada avaliação intervalar, citando que o objetivo da avaliação intervalar é estabelecer, quando solicitado pelo contratante, um intervalo de valores admissíveis em torno da estimativa de tendência central, podendo ser na forma de intervalo de confiança ou de predição.

Contudo, entre os engenheiros e arquitetos avaliadores de imóveis, as definições, as diferenças e as formas de aplicação de intervalo de confiança e intervalo de predição em avaliações tem gerado dúvidas e polêmicas, principalmente por não se ter uma normatização clara quanto à viabilidade de utilização de um ou outro intervalo. Há de se destacar que a própria literatura técnica de estatística utiliza, para o que a NBR 14.653 chama de intervalo de predição, por vezes, uma nomenclatura diferente.

Esse estudo objetiva trazer um pouco de luz a esse assunto na Engenharia de Avaliações, buscando auxiliar na compreensão de cada um desses intervalos e na apropriada aplicação destes na avaliação de bens.

2 REVISÃO DA LITERATURA

2.1 AVALIAÇÃO POR REGRESSÃO LINEAR

Dado um conjunto de características de um bem avaliando $\mathbf{x}_0$, o valor de um imóvel pode ser calculado, segundo a teoria da inferência estatística, através de (1):

$$y_0 = \mathbf{x}_0^T \boldsymbol{\beta} + \boldsymbol{\varepsilon} \qquad (01)$$

Onde $\boldsymbol{\beta}$ é o valor real (da população) dos coeficientes de regressão linear, ou seja, $\boldsymbol{\beta}$ é uma função apenas da população. Como normalmente não se tem acesso a todos os dados da população, não se pode ter acesso ao valor real de $\boldsymbol{\beta}$, o valor dos coeficientes de regressão deve ser estimado segundo uma amostra desta população. O vetor dos coeficientes estimados assim obtido, representado por $\hat{\boldsymbol{\beta}}$, pode ser utilizado para a estimativa de tendência central para bem-avaliando, através da expressão $\hat{y}_0 = \mathbf{x}_0 \hat{\boldsymbol{\beta}}$, já que o valor esperado de $\boldsymbol{\varepsilon}$ é zero, *i.e.*, $E(\boldsymbol{\varepsilon}) = 0$ (FARAWAY, 2004, p. 41).

Nessa perspectiva, ao se trabalhar com modelos de regressão linear, surge o que chamamos de estimativa intervalar, que são métodos estatísticos que usam dados de uma amostra para produzir intervalos de valores que provavelmente contêm o parâmetro populacional de interesse. Dos diferentes tipos de intervalos estatísticos, os intervalos de confiança são os mais conhecidos. No entanto, certos tipos de análises e situações exigem outros tipos de intervalos que forneçam informações diferentes. Nos capítulos seguintes será discorrido sobre intervalo de confiança e intervalo de predição, muito utilizados em engenharia de avaliações.

2.1.1 INTERVALO DE CONFIANÇA

Um intervalo de confiança é um intervalo de valores, derivado das estatísticas da amostra, que provavelmente contém o valor de um parâmetro populacional desconhecido. Devido à sua natureza aleatória, é improvável que duas amostras de uma determinada população produzam estimativas idênticas.

Caso se considere diferentes amostras, todas retiradas da mesma população, pode-se produzir estimativas e intervalos de confiança ligeiramente diferentes.

Nesse sentido, Meeker *et al* (2017) ensinam que, se se coleta diversas amostras aleatórias da mesma população e calcula-se uma estimativa e um intervalo de confiança para cada uma dessas amostras, uma determinada porcentagem das faixas conterá o parâmetro desconhecido dessa população.

Essa porcentagem é o que, em estatística, chamamos de nível de confiança.

Os cálculos do intervalo de confiança a partir dos dados da amostra acabarão por produzir um intervalo de valores que provavelmente contém o parâmetro da população. Se, por exemplo, o intervalo de confiança para a média amostral for IC = [6, 9], muito provavelmente a média da população estará entre 6 e 9, inclusive.

Por exemplo, sobre um intervalo de confiança de 90% para a média, pode-se afirmar que se este processo de amostragem se repetir infinitamente, 90% dos intervalos de confiança construídos irão conter a média real da população.

Se se retira 10 amostras aleatórias de determinada população e com elas se calculam 10 intervalos de confiança para a média populacional, um nível de confiança de 90% indica que deveria se esperar que 9 dos 10 intervalos de confiança incluíssem a média da população. O procedimento de intervalo de confiança é útil porque produz intervalos que geralmente contêm o parâmetro.

A figura 01 mostra essa situação, onde a linha vermelha tracejada representa a média populacional, os pontos azuis as médias de cada uma das 10 amostras extraídas da população, e as linhas verdes os respectivos intervalos de confiança para a média, de cada amostra coletada. Observa-se na figura 01 abaixo que somente um dos intervalos de confiança não possui a média populacional.

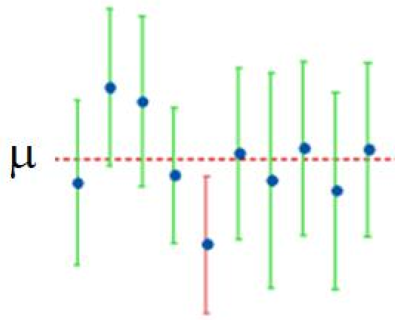

Figura 01 - Intervalos de confiança para a μ a partir de dez amostras e 90% de confiança.

É fácil e intuitivo perceber que, à medida que se extrai amostras aleatórias cada vez maiores da mesma população, os intervalos de confiança para determinado parâmetro populacional tendem a se tornar cada vez mais estreitos.

Além disso, à medida que se aumenta o nível de confiança para uma determinada amostra, digamos de 80% a 95%, o intervalo se torna mais amplo.

Importante destacar que um intervalo de confiança para a média não diz nada sobre como ocorre a dispersão de dados da população.

Pode-se utilizar intervalos de confiança para se produzir intervalos para todos os tipos de parâmetros populacionais. Na maioria das vezes, usam-se intervalos de confiança atrelados à média ou ao desvio padrão, mas também é possível obtê-los para coeficientes da regressão, para proporções ou para taxas de ocorrência.

Segundo Dantas (2012), definido o modelo de regressão linear múltipla adequado, está o engenheiro de avaliações interessado em utilizá-lo para fazer estimativas acerca do bem avaliando, ou de um paradigma, que possui características x_{01}, x_{02}, x_{03},..., x_{0k}. Como $\hat{y}_i$ é um estimador não tendencioso para

$E(y_i)$, é suficiente que se substituam no modelo as características do bem a avaliar, obtendo-se a estimativa pontual correspondente. A equação é:

$$\hat{y}_0 = \hat{\beta}_0 + \hat{\beta}_1 x_{01} + ... + \hat{\beta}_k x_{0k} \quad (02)$$

Uma vez que o erro aleatório possui distribuição normal, a distribuição t de Student terá ($n - k - 1$) graus de liberdade. Constrói-se então um intervalo de confiança a um nível ($1 - \alpha$) para $\hat{y}_0$ na forma da equação 03 abaixo:

$$[\hat{y}_0 - t_{(1-\alpha/2;n-k-1)} \cdot se(\hat{y}_0); \ \ \hat{y}_0 + t_{(1-\alpha/2;n-k-1)} \cdot se(\hat{y}_0)] \quad (03)$$

Na equação, $se(\hat{y}_0)$ refere-se ao erro padrão do estimador (lembre-se que $se(\hat{y}_0) = se(\mathbf{x}_0^T \hat{\boldsymbol{\beta}})$). Para o cálculo de $se(\hat{y}_0)$ pode-se utilizar, na forma matricial, a equação (04), que pressupõe satisfeitas as hipóteses de ausência de heteroscedasticidade e autocorrelação:

$$se(\hat{y}_0) = \sqrt{\mathrm{var}(\mathbf{x}_0^T \hat{\boldsymbol{\beta}})} = \sqrt{\hat{\sigma}^2 \mathbf{x}_0^T \left(\mathbf{X}^T \mathbf{X}\right)^{-1} \mathbf{x}_0} = \hat{\sigma} \sqrt{\mathbf{x}_0^T \left(\mathbf{X}^T \mathbf{X}\right)^{-1} \mathbf{x}_0} \quad (04)$$

Onde:
$\hat{\sigma}^2$ é a variância estimada pelo modelo
$\mathbf{x}_0$ é o vetor das características do bem-avaliando
$\mathbf{X}$ é a matriz das variáveis independentes das amostras do modelo
$\left(\mathbf{X}^T \mathbf{X}\right)^{-1}$ é a matriz inversa da matriz $\mathbf{X}^T \mathbf{X}$
$\mathbf{x}_0^T$ é o vetor transposto do vetor $\mathbf{x}_h$

Esse intervalo pode ser escrito na forma da equação (05), ao nível de confiança de 80% (NBR 14.653-2).

$$IC = \hat{y}_0 \pm t_{(0,90;\ n-k-1)} \cdot \hat{\sigma} \sqrt{\mathbf{x}_0^T \left(\mathbf{X}^T \mathbf{X}\right)^{-1} \mathbf{x}_0} \quad (05)$$

Onde:
$\hat{y}_0$ é o valor estimado na equação
$t_{(0,90;\ n-k-1)}$ é o valor da variável t de Student a 90% para $n-k-1$ graus de liberdade
$\hat{\sigma}$ é o desvio-padrão estimado da população

2.1.2 INTERVALO DE PREDIÇÃO

Pode-se, após ajustar um modelo de regressão, obter intervalos de predição. Os intervalos de predição preveem o valor da variável dependente (explicada), dados ajustes específicos efetuados nas variáveis independentes (explicativas).

Segundo Meeker *et al* (2017), um intervalo de previsão para uma única observação futura é um intervalo que conterá, com um grau de confiança especificado, uma observação selecionada aleatoriamente a partir de uma distribuição.

Nesse caso, um intervalo de predição é uma estimativa de um intervalo em que uma observação futura irá cair, com certo nível de confiança, dadas as observações que já foram realizadas. Em um intervalo de predição de 90%, por exemplo, pode-se afirmar que, se se repetisse o processo de amostragem infinitamente, 90% dos intervalos de predição construídos conteriam a nova observação.

Sendo assim, um intervalo de predição é um intervalo que provavelmente contém o valor da variável dependente para uma única nova observação, dados valores específicos das variáveis independentes. Com esse tipo de intervalo, se está prevendo intervalos para observações individuais em vez do valor médio.

Os intervalos de predição são normalmente utilizados para fazer previsões a partir de um modelo de regressão, em que uma quantidade está sendo prevista. Um exemplo da apresentação de um intervalo de predição é o seguinte:

Dada uma predição de 'y' dado 'x', há uma probabilidade de 95% de que o intervalo 'a' para 'b' cubra o resultado verdadeiro (MEEKER *et al*, 2017).

Assim, os intervalos de predição devem considerar tanto a incerteza em saber o valor da média da população, quanto a dispersão dos dados. Conforme James *et al* (2013), os intervalos de predição serão sempre mais amplos do que os intervalos de confiança, pois eles são os responsáveis pela incerteza associada ao erro.

Com isso, tem-se que o intervalo de confiança quantifica a incerteza em uma variável populacional estimada, como a média ou o desvio padrão, e o intervalo de predição quantifica a incerteza em uma única observação estimada da população.

A forma com que se determina o intervalo de predição é bastante semelhante à forma como se determina o intervalo de confiança. A diferença está no fato de que, no cálculo dos limites do intervalo de predição, deve-se levar em conta não apenas a variância do estimador de mínimos quadrados, mas também a variância da população, obtendo-se assim, um intervalo mais amplo.

Novamente, uma vez que o erro aleatório possui distribuição normal, a distribuição t de Student terá ($n - k - 1$) graus de liberdade. Constrói-se então um intervalo de predição a um nível ($1 - \alpha$) para $\hat{y}_0$ na forma do intervalo 06 abaixo:

$$[\hat{y}_0 - t_{(1-\alpha/2;n-k-1)} \cdot \sqrt{\mathrm{var}(\mathbf{x}_0^T\hat{\boldsymbol{\beta}} + \boldsymbol{\varepsilon})};\;\; \hat{y}_0 + t_{(1-\alpha/2;n-k-1)} \cdot \sqrt{\mathrm{var}(\mathbf{x}_0^T\hat{\boldsymbol{\beta}} + \boldsymbol{\varepsilon})}] \tag{06}$$

Satisfeita a suposição de que os erros são independentes, então $\mathrm{var}(\mathbf{x}_0^T\hat{\boldsymbol{\beta}} + \boldsymbol{\varepsilon}) = \mathrm{var}(\mathbf{x}_0^T\hat{\boldsymbol{\beta}}) + \mathrm{var}(\boldsymbol{\varepsilon})$. Ora, como a variância do termo de erro é um parâmetro da população (o desvio-padrão da população), este também deve ser estimado: $\mathrm{var}(\boldsymbol{\varepsilon}) = \hat{\sigma}^2$

O intervalo de predição pode ser escrito, então, na forma da equação (07), ao nível de confiança de 80% (NBR 14.653-2).

$$IP = \hat{y}_0 \pm t_{(0{,}90;\; n-k-1)} \cdot \sqrt{\hat{\sigma}^2 \mathbf{x}_0^T \left(\mathbf{X}^T\mathbf{X}\right)^{-1} \mathbf{x}_0 + \hat{\sigma}^2} \Leftrightarrow \tag{07}$$

$$IP = \hat{y}_0 \pm t_{(0{,}90;\; n-k-1)} \cdot \hat{\sigma}\sqrt{\mathbf{x}_0^T \left(\mathbf{X}^T\mathbf{X}\right)^{-1} \mathbf{x}_0 + 1}$$

2.1.3 CAMPO DE ARBÍTRIO DO AVALIADOR

A definição do campo de arbítrio do avaliador já foi alterada em várias revisões normativas e tem sido objeto de estudo de diversos autores da área (GIANNAKOS *et al.*(1996); GIANNAKOS; LEÃO (1998); MACEDO *et al*. (2001)).

A NBR 14.653-1 (2019) define campo de arbítrio em seu item 3.8, p. 3:

> **campo de arbítrio:** Intervalo de variação no entorno do estimador pontual adotado na avaliação, dentro do qual pode-se arbitrar o valor do bem, desde que justificado pela existência de características próprias não contempladas no modelo.

Em termos estatísticos, isto equivale a "corrigir" o modelo devido a presença de viés, ou seja, devido a variáveis (características próprias) omitidas na modelagem. Em estatística, este tipo de viés é chamado de viés de modelagem, ou *model bias,* ou seja, quanto maior o número de variáveis explicativas, mais rico o modelo, menor será o viés, porém será maior a variância (MATLOFF, 2017, p. 25). Já a parte 2 da NBR 14.653 (2011), na mesma filosofia da parte 1, estabelece que:

> **8.2.1.5.2** O campo de arbítrio pode ser utilizado quando variáveis relevantes para a avaliação do imóvel não tiverem sido contempladas no modelo, por escassez de dados de mercado, por inexistência de fatores de homogeneização aplicáveis ou porque essas variáveis não se apresentaram estatisticamente significantes em modelos de regressão, **desde que a amplitude de até mais ou menos 15 % seja suficiente para absorver as influências não consideradas e que os ajustes sejam justificados** (ABNT, 2011, p. 17, *grifo nosso*)**.**

E faz uma ressalva:

> **8.2.1.5.3** Quando a amplitude do campo de arbítrio não for suficiente para absorver as influências não consideradas, o modelo é insuficiente para que a avaliação possa atingir o grau mínimo de fundamentação no método comparativo direto de dados de mercado e esse fato deve ser consignado no laudo.

De maneira que, subentende-se, o fator limitante último de uma avaliação, ou seja, o fator limitante que estará em voga quando todos os outros fatores tiverem sido descartados, é o Campo de Arbítrio do Avaliador.

2.2 AVALIAÇÃO INTERVALAR

Segundo a NBR 14.653-1 (2019), a identificação do valor de um bem deve ser efetuada segundo a metodologia que melhor se aplique ao mercado de inserção do bem e a partir do tratamento dos dados de mercado, permitindo-se indicar a faixa de variação de preços do mercado admitida como tolerável em relação ao valor final, desde que indicada a probabilidade associada.

No contexto da expressão "faixa de variação de preços do mercado", indicando a probabilidade a ela associada, a norma de avaliações faz referência à chamada avaliação intervalar, que é novamente retomada em sua parte 2 - anexo A, no que diz respeito à regressão linear.

A avaliação intervalar tem como objetivo estabelecer, quando solicitado pelo contratante, um intervalo de valores admissíveis em torno da estimativa de tendência central ou do valor arbitrado (NBR 14.653-2, 2011).

O intervalo de valores admissíveis citado pode ser na forma de intervalo de confiança ou na forma de intervalo de predição. A seguir serão vistas as respectivas definições. Conforme a NBR 14.653-2 (2011) sobre avaliação de bens – parte 2: imóveis urbanos, em seu tópico sobre termos e definições, define:

> **Intervalo de Confiança**: intervalo de valores dentro do qual está contido o parâmetro populacional com determinada confiança (ABNT, 2011).

Essa definição é bastante precisa, ainda que sucinta, para quem conhece os conceitos de inferência estatística. Em geral, na engenharia de avaliações, o parâmetro populacional que se busca estimar é o valor de mercado, representado pela média, em uma população em que o desvio padrão é desconhecido.

Nesse caso, o intervalo de confiança irá definir os limites inferior e superior do intervalo de valores dentro do qual, invariavelmente, estará o valor de mercado onde o avaliador pesquisou os dados de mercado, com certo grau de confiança.

Importante destacar que o intervalo de confiança é definido tendo como referência os dados de uma amostra coletada pelo avaliador. Sendo assim, o intervalo de confiança é o intervalo onde a média populacional estará, considerando a confiança predeterminada no estudo. A NBR 14.653-2 (2011) para imóveis urbanos define como nível de confiança o valor de 80%.

O intervalo de confiança é importante na engenharia de avaliações uma vez que é ele quem determina o grau de precisão que será atingido na avaliação, sendo que esse grau de precisão é determinado em função da amplitude do intervalo de confiança, conforme o item 9.2.3 da NBR 14.653-2 (2011).

A parte 2 da NBR 14.653-2 (2011), no tópico de termos e definições, define:

> **Intervalo de Predição:** estimativa de um intervalo de valores, a partir de dados de mercado observados, dentro do qual novos dados do mesmo contexto estarão contidos, com determinada probabilidade (ABNT, 2011).

Conforme a NBR 14.653-2 (2011), quando for adotada a estimativa de tendência central para a avaliação, o intervalo de valores admissíveis deve estar limitado simultaneamente: A) ao intervalo de predição ou ao intervalo de confiança de 80 % para a estimativa de tendência central; B) ao campo de arbítrio. Ver a esse respeito a figura 02.

Ainda segundo a norma, o intervalo de confiança e de predição são utilizados também no arbitramento do valor pontual da avaliação. Nesse caso, quando for adotado o valor arbitrado, o intervalo de valores admissíveis deve estar limitado simultaneamente: A) ao intervalo em torno do valor arbitrado com amplitude igual à do intervalo de predição ou ao intervalo de confiança de 80% para a estimativa de tendência central; B) ao campo de arbítrio em torno da estimativa de tendência central. Ver a esse respeito a figura 03.

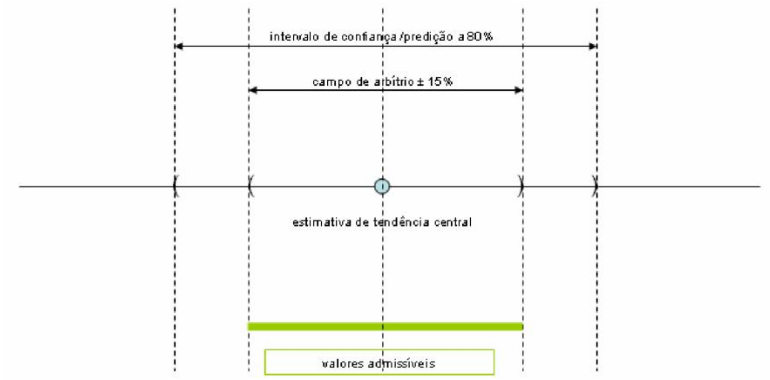

Figura 02 - Valores admissíveis quando for adotada a estimativa de tendência central. Fonte: NBR 14.653-2.

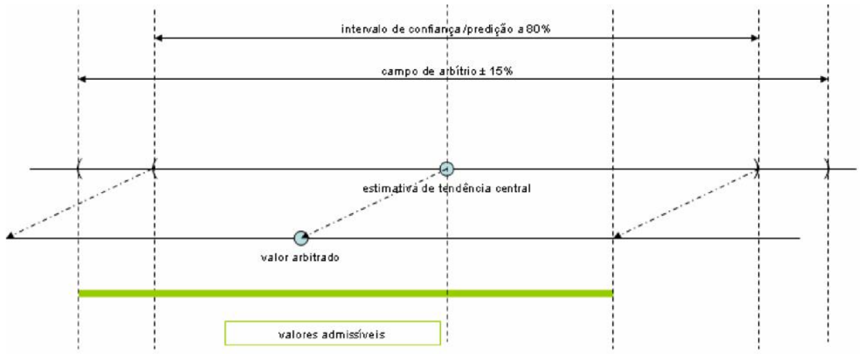

Figura 03 - Valores admissíveis quando for adotado o valor arbitrado. Fonte: NBR 14.653-2.

2.3 INTERVALO DE CONFIANÇA OU DE PREDIÇÃO: QUAL UTILIZAR?

Um aspecto polêmico envolvendo esse tema é que, segundo a NBR 14.653-2 (2011), o intervalo de confiança será utilizado se o objetivo for estimar o valor de mercado. Se o objetivo for estimar preços, utiliza-se o intervalo de predição.

É importante, então, definir as expressões *valor de mercado* e *preço*:

> **Preço:** é uma expressão monetária que define uma transação de um bem, de seu fruto, de um direito, ou da expectativa da sua transação (NBR 14.653-1, 2019, p.6). O engenheiro Abunahman (2008) define preço como a quantia paga pelo comprador ao vendedor.
>
> **Valor de Mercado:** quantia mais provável pela qual se negociaria voluntariamente e conscientemente um bem, em uma data de referência, dentro das condições do mercado vigente (NBR 14.653-1, 2019, p.7). Já para Abunahman (2008), valor de mercado é aquele encontrado por um vendedor desejoso de vender mas não forçado e um comprador desejoso de comprar mas também não forçado, tendo ambos pleno conhecimento das condições de compra e venda e da utilidade da propriedade.

Na literatura estatística, no entanto, encontra-se algo ligeiramente diferente. Segundo Faraway (2004, p. 41-42), se o que se deseja conhecer é o valor médio que os imóveis com as características x_{01}, x_{02}, ..., x_{0k} assumem, então apenas a variância do estimador, no caso, o estimador dos mínimos quadrados ordinários, é levada em consideração para o cálculo do intervalo, que neste caso denomina-se intervalo de confiança. Em contrapartida, se o que se deseja saber é o valor de um imóvel em específico (e não o valor médio de um imóvel com aquelas características), para o cálculo do seu intervalo de variação, deve-se levar em consideração não apenas a variância de estimação do modelo, mas também a variância populacional, que neste caso é chamado de intervalo de predição (FARAWAY, 2004, pp. 41–42).

Segundo Faraway (2004, p. 41), na maioria das vezes se está interessado no intervalo de predição, que é chamado de predição de um valor futuro, enquanto o outro caso, chamado de predição da resposta média, é menos requerido na prática.

Matematicamente, a diferença entre o intervalo de confiança e o intervalo de predição pode ser melhor compreendido entendendo-se a origem da cada um: enquanto o intervalo de confiança é obtido do cálculo da variância da estimação ($\operatorname{var}(\mathbf{x}_0^T\hat{\boldsymbol{\beta}})$), o intervalo de predição leva em consideração não apenas a variância da estimação, mas também a variância do mercado ($\operatorname{var}(\mathbf{x}_0^T\hat{\boldsymbol{\beta}}+\boldsymbol{\varepsilon})$).

Em suma, os valores dos $\hat{\beta}_i$ estimados pelo modelo de regressão possuem uma dada variância, pois se trabalha apenas com uma pequena amostragem da população para inferi-los: modificada a amostra, mudam-se os valores dos $\hat{\beta}_i$. Para o cálculo do intervalo de confiança, apenas esta variância é levada em conta. Esta variância pode ser reduzida com o simples aumento do número de dados amostrais, sendo que ela tende a zero quando o número de dados amostrais tende ao número de elementos da população (em outras palavras, se tivéssemos acesso a todos os dados da população, por conseguinte teríamos como calcular o valor real de $\boldsymbol{\beta}$, e a amplitude do *IC* seria zero, ou seja, a média calculada seria exata).

No entanto, mesmo que conhecêssemos toda a população, ou seja, mesmo que o valor real de $\boldsymbol{\beta}$ fosse exatamente conhecido, ainda assim estaria presente a variância da população em si (σ^2), que no caso da Engenharia de Avaliações, é a variância do mercado, que deve ser levada em conta, se o que se pretende é estimar o valor de um bem em particular, e não o valor médio dos bens com características $\mathbf{x}_0$ iguais ao do bem-avaliando.

2.4 PROBLEMAS DA ATUAL NORMATIZAÇÃO BRASILEIRA

A figura 04 demonstra graficamente, em duas dimensões (valor unitário x área), o efeito da aplicação da atual normatização em diferentes cenários. Na figura, as elipses representam uma população em particular. A reta verde representa a reta de regressão obtida de uma determinada amostra dessa população. Em azul podem ser vistos os limites dos intervalos de confiança @80%. Em preto, os limites do intervalo de predição. A reta vermelha no centro dos gráficos representa a amplitude do campo de arbítrio, centralizado em relação à estimativa de tendência central.

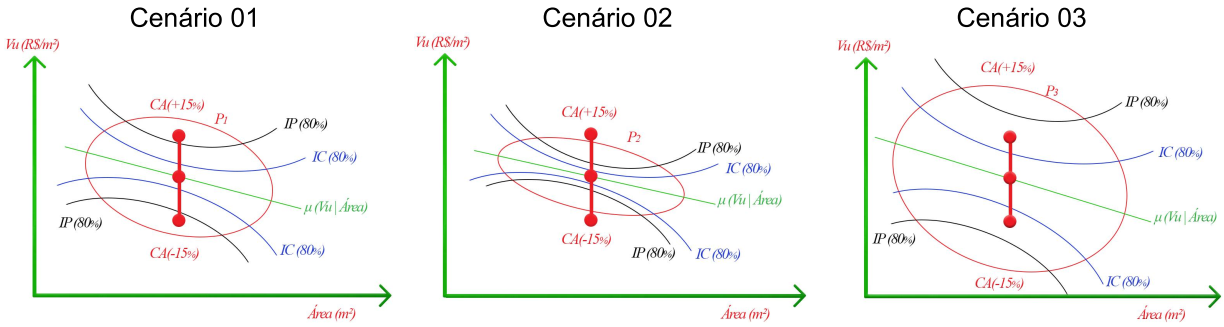

Figura 04 - Representação gráfica da aplicação do campo de arbítrio em diferentes mercados.

Pode-se notar que, apenas por acaso os limites do campo de arbítrio podem se encontrar dentro dos limites da população (Cenário 01). Mas, a depender da variância da população, os limites do campo de arbítrio pode estar além dos limites populacionais (Cenário 02), ou ainda muito aquém (Cenário 02).

Como a atual normatização permite ao avaliador a adoção de qualquer valor dentro da amplitude do campo de arbítrio ($\pm 15\%$), pode-se estar permitindo ao avaliador o arbitramento de valores irreais, ou ainda limitando desnecessariamente o arbítrio do mesmo.

3 ESTUDOS DE CASOS

3.1 ESTUDO COM DADOS DE MERCADO

Neste estudo são utilizados 50 dados de mercado obtidos em 2015 na região central de Florianópolis (Anexo A). Foi realizada uma análise descritiva dos dados e gerado um modelo de regressão linear atendendo a todos os pressupostos básicos. Para este estudo de caso foi utilizado o software R 3.6.0 (R CORE TEAM, 2019).

Um resumo dos coeficientes e estatísticas do modelo pode ser visto na tabela 1. Como o modelo ajustado tem a transformação logarítmica natural, os valores ajustados devem ser re-transformados para a escala original. Aqui se adotará a transformação pela mediana da distribuição log-normal.

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	***
ln(dist_b_mar)	0.141	0.176	0.105	0.027	5.174	<0.001	***
Padrao^{-1}	0.563	0.700	0.426	0.105	5.360	<0.001	***

Na escala log, o valor central obtido foi de 13,7764. Pode-se mostrar que o erro padrão do ajuste $s.e.(\hat{y})$ para o bem avaliando, obtido de acordo com a expressão $\sqrt{\hat{\sigma}^2 x'_0 (X'X)^{-1} x_0}$ é de ≈ 0,0300 (também na escala log). O valor de $t_{41}^{10\%}$ é aproximadamente igual a –1,3025 e o de $t_{41}^{90\%}$ é ≈ 1,3025. 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 (foram utilizadas 6 variáveis explicativas e 2 dados foram identificados como *outliers*):

$$\hat{y}_{\sup} = \hat{y} + t_{41}^{10\%} s.e.(\hat{y}) \quad \text{e} \quad \hat{y}_{\inf} = \hat{y} + t_{41}^{90\%} s.e.(\hat{y}) \tag{08}$$

Ou seja: $\hat{y}_{\sup} = 13{,}7764 + 1{,}3025 \cdot 0{,}0300$ e $\hat{y}_{\inf} = 13{,}7764 - 1{,}3025 \cdot 0{,}0300$

Caso se opte por adotar 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, pois o campo de arbítrio [R$ 817.411,55; R$ 1.105.909,74] não pode ser inteiramente utilizado.

No entanto, caso seja arbitrado, por exemplo, o valor de R$ 1.100.000,00 para o bem avaliando, 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 13,8747, 13,9108 e 13,9162.

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

$$t_{inf} = 3{,}15 \qquad t_{arbitrado} = 4{,}36 \qquad t_{sup} = 4{,}53$$

De posse dos t calculados acima, pode-se calcular as suas probabilidades de ocorrência: 0,15%; 0,004% e 0,002%. Ou seja, de acordo com o *IC* originalmente calculado, o valor arbitrado, o valor inferior e o valor superior do “novo intervalo”, tem baixíssima probabilidade de ocorrência.

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 ($\hat{Y}_{\inf} = 13{,}5949$ e $\hat{Y}_{\sup} = 13{,}9579$, na escala original, 802.017,63 ≤ 961.660,64 ≤ 1.153.080,88), 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*:

$$t_{inf\_CA} = -1{,}17 \qquad t_{sup\_CA} = 1{,}00$$

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 *CA* não pode ser inteiramente utilizado, apesar de seus limites se encontrarem em percentis confortáveis do *IP,* definido pela própria norma.

3.2 ESTUDO COM DADOS RANDÔMICOS

3.2.1 Impacto do Número de Graus de Liberdade do Modelo

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

$$IP = \hat{y}_0 \pm t_{n-k}^{\alpha/2} \sqrt{\hat{\sigma}^2 + s.e.^2(\hat{y}_0)} \qquad (09)$$

Deve-se notar que, fixado o valor central $\hat{y}_0$ e a amplitude do intervalo, podese calcular o valor do termo $\sqrt{\hat{\sigma}^2 + s.e.^2(\hat{y}_0)}$ 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.

Imagine-se que, em um determinado modelo previamente ajustado, foi estimado o valor de um bem, adotando-se a estimativa de tendência central, o valor de R$ 1.000.000,00.

Imagine-se então que o modelo possui graus de liberdade suficientes para que o intervalo de predição possa ser assumido normalmente distribuído. Isto implica que, assumindo diferentes valores para a amplitude do IP, o valor do termo $\sqrt{\hat{\sigma}^2 + s.e.^2(\hat{y}_0)}$ pode ser calculado, conforme expressão abaixo, onde $\overline{IP}/2$ é a semi-amplitude do IP considerado:

$$\sqrt{\hat{\sigma}^2 + s.e.^2(\hat{y}_0)} = \frac{\overline{IP}/2 * 1.000.000 - 1.000.000}{p_{normal,10\%}} = \frac{1.000.000(\overline{IP}/2 - 1)}{1,2815}$$

De posse dos valores dos termos para as amplitudes de 30, 40 e 50%, podem ser calculados os *z-scores* dos limites inferior e superior do campo de arbítrio:

$$Z_{\inf} = \frac{(850.000 - 1.100.000)}{\sqrt{\sigma^2 + se^2(\hat{y}_0)}}; Z_{\sup} = \frac{(1.150.000 - 1.100.000)}{\sqrt{\sigma^2 + se^2(\hat{y}_0)}}$$

A partir dos valores dos *z-scores* assim calculados, podem ser obtidos os seus respectivos percentis de probabilidade de ocorrência. Os resultados estão resumidos na tabela 2.

Tabela 2 - Percentis do *CA* considerando *IP* com distribuição normal.

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%

Na tabela acima, as colunas $Z_{\inf}$ e $Z_{\sup}$ mostram os quantis, ou *z-scores*, da distribuição normal para os limites inferior e superior do *CA*, enquanto as colunas $p_{\inf}$ e $p_{\sup}$ mostram os percentis correspondentes.

O mesmo pode ser feito para distribuição *t* com 40 graus de liberdade, o que pode ser visto na tabela 3. Neste caso, o termo de erro deve ser assim calculado:

$$\sqrt{\hat{\sigma}^2 + s.e.^2(\hat{y}_0)} = \frac{\overline{IP}/2*1.000.000 - 1.000.000}{t^{90}_{40,10\%}} = \frac{1.000.000(\overline{IP}/2-1)}{1,3031}$$

E os valores dos *t* correspondetes ao campo de arbítrio:

$$t^{40}_{\inf} = \frac{(850.000 - 1.100.000)}{\sqrt{\sigma^2 + se^2(\hat{y}_0)}}; t^{40}_{\sup} = \frac{(1.150.000 - 1.100.000)}{\sqrt{\sigma^2 + se^2(\hat{y}_0)}}$$

Tabela 3 - Percentis do *CA* considerando *IP* com distribuição *t* e 40 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%	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%

O mesmo ainda pode ser feito para distribuição *t* com 10 graus de liberdade, o que pode ser visto na tabela 4.

Tabela 4 - Percentis do *CA* considerando *IP* com distribuição t e 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 (n – k – 1) ≥ 10. Em todos os casos, a variação mais relevante é a das diferentes amplitudes dos intervalos de predição, *i.e.*, importa mais a variabilidade do mercado, ou variabilidade populacional.

3.2.2 Impacto da Variabilidade do Mercado

Para ilustrar qual o impacto da variabilidade do mercado nos procedimentos de avaliação intervalar foram criados quatro conjuntos de 200 dados randômicos normais multivariados cada, com valor unitário médio de R$ 3.000,00 e área média de 150 m^2, com diferentes variâncias. Posteriormente, para cada conjunto de dados criado, foi ajustada uma reta de regressão e calculados os intervalos de predição @80% e de confiança também @80%. Os resultados podem ser vistos graficamente na figura 05, onde os pontos e linhas 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 área cinza representa o intervalo de confiança @80%, as retas em verde representam as retas de regressão quantílica para os quantis 10% e 90% e a reta azul é a reta de regressão, em vários cenários hipotéticos. As elipses desenhadas são elipses de confiança ao nível de 95%, representando aproximadamente as populações.

Primeiramente, foi imaginado um cenário hipotético de mercado equilibrado, onde, por acaso, o campo de arbítrio CA 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 CA 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 ± 15% não chegam nem à metade do intervalo de predição.

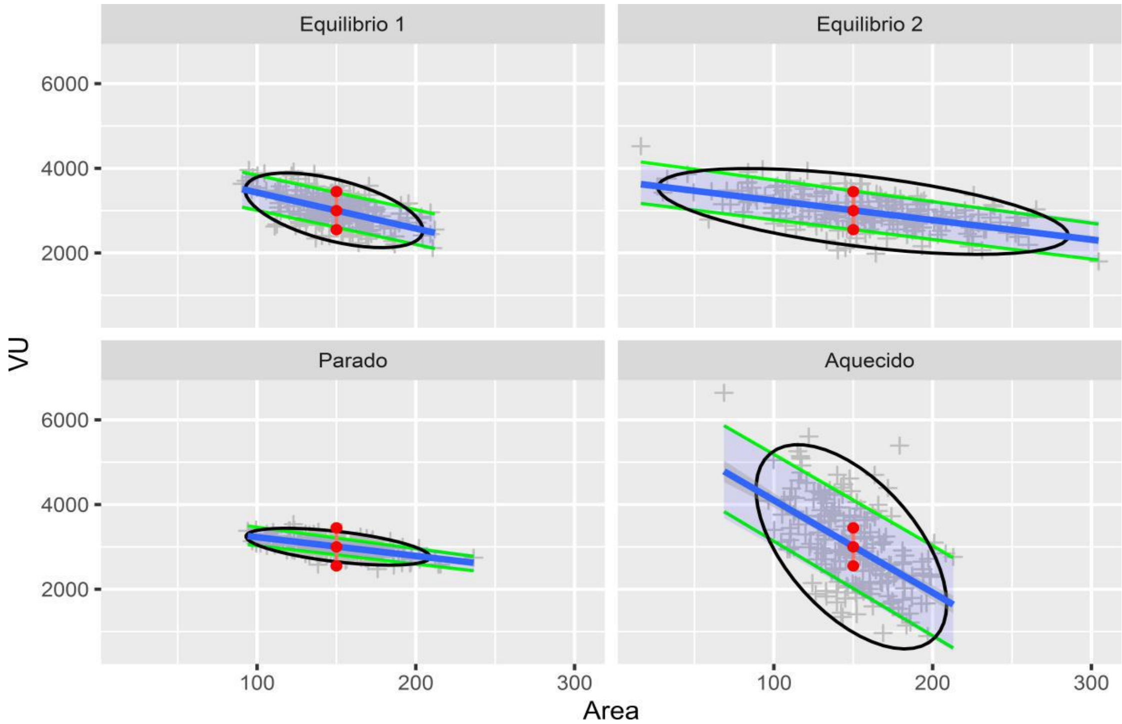

Figura 05 - Comparação do campo de arbítrio em diferentes realidades mercadológicas.

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 *CA* estará sub ou super-avaliando o bem avaliando para aquele mercado.

Já no caso de um mercado com maior variabilidade, a utilização dos limites do *CA* 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.

Tabela 5 - valores da estimativa central, limites do *IP* e R^2 para os dados simulados

Cenário	Estimativa Central	IP_{inf}	IP_{sup}	Amplitude IP	R^2_{adj}
Mercado Equilibrado 1	3.010,78	2.633,27	3.388,29	25,10%	0,25
Mercado Equilibrado 2	2.945,85	2.592,27	3.299,44	24,00%	0,44
Mercado Parado	2.987,87	2.803,33	3.172,40	12,40%	0,46
Mercado Aquecido	3.072,36	2.095,38	4.049,34	63,60%	0,43

A tabela acima mostra com mais precisão a magnitude do problema: efetuouse a estimativa para os valores centrais (*Área = 150 m²*) com o intervalo de predição @80%. Nos dois primeiros cenários, os limites do *IP* e *CA* são praticamente equiva-lentes. Porém, no terceiro cenário o campo de arbítrio ficaria muito além do *IP*, enquanto no quarto cenário o *CA* seria insuficiente para chegar aos extremos do *IP*.

Finalmente, para averiguar a pertinência da adoção do *IC* para a avaliação intervalar e para a aferição do Grau de Precisão, foi elaborada a tabela abaixo:

Tabela 6 - valores da estimativa central, limites do *IC* e R^2 para os dados simulados

Cenário	Estimativa Central	IC_{inf}	IC_{sup}	Amplitude IC	R^2_{adj}
Mercado Equilibrado 1	3.010,78	2.957,16	3.064,40	3,56%	0,25
Mercado Equilibrado 2	2.945,85	2.896,32	2.995,39	3,36%	0,44
Mercado Parado	2.987,87	2.961,99	3.013,74	1,73%	0,46
Mercado Aquecido	3.072,36	2.935,33	3.209,39	8,92%	0,43

Como se pode notar, em nenhum dos cenários o campo de arbítrio *CA* seria menos restritivo do que o intervalo de confiança *IC*, mesmo no cenário de mercado aquecido. É interessante notar também que a amplitude do *IC* e do *IP* pouco tem a ver também com o grau de ajuste dos modelos: mesmo para baixos valores de R^2, a amplitude dos *IC*s é muito baixa, enquanto a amplitude dos *IP*s varia drasticamente de acordo com a variância da população, como queríamos demonstrar.

3.3 ESTUDO COM DADOS LOGNORMAIS

Imagine-se um modelo de regressão linear onde a variável resposta tenha sido transformada com a função logaritmo natural e que, com este modelo, tenham sido obtidas as seguintes estimativas para o bem avaliando:

$$\ln(\hat{y}) = 13{,}8155 \quad \text{e} \quad \hat{\sigma} = 0{,}1003$$

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

Tabela 7 – Estimativas do campo de arbítrio e intervalo de predição.

CA ou 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, 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. Os resultados podem ser vistos na tabela abaixo:

Tabela 8 – estimativas do campo de arbítrio e intervalo de predição

CA ou 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 combinada com a adoção do limite inferior do CA corresponde ao percentil 4,26%, ou seja, um percentil muito baixo da distribuição da população.

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 se consideram adequados, ou seja, os percentis 10% e 90%.

3.3.1 Avaliação Intervalar Segundo a NBR14.653-2

Segundo a atual NBR 14.653-2, o avaliador teria várias alternativas para a definição da estimativa final, dentre as quais se elencam as alternativas I a V abaixo, cujos valores dos limites inferior e superior do intervalo de valores admissíveis segundo a atual normativa estão resumidas na tabela 9 abaixo:

Tabela 9 – Intervalos de valores admissíveis

Caso	Limites (R$)		Percentis (%)	
	Inferior	Superior	Inferior	Superior
I	841.500,00	977.224,64	4,26	40,90
II	850.000,00	987.095,60	5,25	44,80
III	879.433,53	987.095,60	10,00	90,00
IV	1.029.433,53	1.150.000,00	61,40	91,80
V	1.034.619,63	1.155.793,49	63,30	92,60

I. A adoção da estimativa pela moda e do limite inferior do CA:

Se o avaliador optasse pela estimação pela moda e, fazendo uso do campo de arbítrio, arbitrasse o limite inferior do CA (R$841.500,00), o intervalo final de valores admissíveis da sua avaliação, segundo a atual normativa, seria: [841.500,00; 977.224,64]. Ou seja, fazendo-se esta escolha, o intervalo final de valores admissíveis obtido estaria entre os percentis 4,26% e 40,90%.

II. A 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%.

III. A 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 10% e 90%.

IV. A 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%.

V. A 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%.

4 CONSIDERAÇÕES FINAIS

A NBR 14.653-2 (2011) faria melhor em distinguir valor médio, $\mu(Y \mid X = t)$ e intervalo de confiança para a média $IC(\mu(t))$, de valor ajustado/arbitrado ($\hat{Y}$) e intervalo de predição para os valores ajustados/arbitrados.

A amplitude do IC está intimamente relacionada com o número de dados: quanto maior o número de dados, ou seja, quanto mais representativa é a amostra, menor a amplitude do IC. A amplitude do IC também está relacionada ao número de regressores, através do *tradeoff* entre viés e variância.

O IC não contempla, porém, a variabilidade do mercado. Duas populações com médias iguais e diferentes variâncias podem apresentar ICs de mesma amplitude (ou ainda o IC da população de maior variância pode ter menor amplitude do que o IC da população de menor variância), a depender das amostragens.

Já o IP leva em conta não apenas a variância do estimador, mas também a variância estimada da população, de maneira que o IP não depende apenas da precisão do modelo ajustado. Por mais ajustado que seja um modelo, ele terá IP de grande amplitude, se a variância da população for alta.

O avaliador deveria ter o poder, então, de arbitrar valores para os bens-avaliandos dentro do IP, independente de outras restrições normativas. Claro, pois o IC é para a resposta média, mas o avaliador deveria ter condições de dizer se o bem-avaliando encontra-se acima ou abaixo dela, a depender da “existência de características próprias não contempladas no modelo”.

Por outro lado, o avaliador nunca deveria ter o poder de arbitrar valores para a média fora do IC, pois foi feita uma inferência para a estimação da média e seu intervalo de confiança. A NBR 14.653-2, porém, permite ao avaliador arbitrar que a média encontra-se fora do IC. Isto não faz qualquer sentido científico.

Quanto ao campo de arbítrio do avaliador, este não é um bom parâmetro para a arbitragem de valores, dado que o mesmo é constante ($\pm$15%), independente da variabilidade do mercado. O que a NBR 14.653-2 propõe com o conceito de campo de arbítrio, em termos estatísticos, é absurdo, pois o campo de arbítrio limita todas as avaliações de bens para os valores médios do mercado, mais ou menos 15%, como se não houvesse imóveis fora da média do mercado, e como se todos os mercados tivessem variâncias iguais. Ou seja, é como pretender que se possa estabelecer, por meio de norma, a variância de uma população.

Pela análise dos dados lognormais, fica ainda mais visível o quão absurdas são as regras de arbítrio e de avaliação intervalar na atual normativa.

Por um lado, a norma dá ao avaliador a liberdade de escolher a medida de tendência central que melhor entender e a liberdade de ainda aplicar o campo de arbítrio de mais ou menos 15% como bem entender para o estabelecimento do valor central. Por outro lado, estabelece regras para a definição dos intervalos finais que não fazem qualquer sentido técnico e/ou científico. Os intervalos finais obtidos podem ser feitos tão extremos quanto se deseja, a depender da combinação de escolha de medida de tendência central e localização da avaliação no campo de arbítrio. Pior ainda é o fato de que os percentis correspondentes dos intervalos finais não ficam claros no laudo, pois a NBR 14.653-2 ainda estabelece que: “A.10.2 No caso da utilização do valor arbitrado, este fato deve ser citado e não será calculada a probabilidade associada ao intervalo”.

A NBR 14.653-2 tem um problema conceitual: ela mistura, através da avaliação intervalar, os conceitos de valor médio e intervalo de confiança para a média, com os conceitos de valor ajustado/arbitrado e intervalo de predição. O valor médio e o intervalo de confiança para a média são, geralmente, de pouco interesse para o avaliador: estes conceitos tem valor mais explicativo. O conceito de média e IC tem sentido para responder perguntas do tipo: *quanto, em média, uma garagem adicional acrescenta em valor a um imóvel num determinado mercado?* De posse do valor do coeficiente de regressão da variável garagem e do seu IC, pode-se dizer que uma garagem adicional, em média, acrescenta, digamos, R$ 30.000 (R$ 20.000 a R$ 40.000, IC @80%) ao valor de um imóvel.

Existem vários ramos da ciência que se valem destes conceitos, como é o caso da medicina, por exemplo: "o modelo comprovou que pessoas com IMC (índice de massa corpórea) entre 25 e 30 vivem, em média, 10 anos a mais do que pessoas com IMC > 35 (8 - 12 , ao nível de 95%)". Isto dá embasamento aos médicos para recomendar que a população em geral deve manter o peso sob controle. Mas, apesar de ser uma média para a população, dificilmente alguém poderia dizer, baseado em um modelo deste tipo, quantos anos a mais viverá uma determinada pessoa com dado *IMC*, pois a variabilidade é muito grande.

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 da população, não o valor médio de mercado de bens com aquelas características. Por isso o avaliador é obrigado a fazer uma vistoria no bemavaliando, 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%.

Em suma, entende-se que o objetivo do avaliador na avaliação de um determinado imóvel deveria ser estabelecer um valor arbitrado para o bem-avaliando, de acordo com um modelo de regressão linear, acima ou abaixo da média do mercado, em função das características não contempladas no mesmo modelo, deixando claro em que percentil do IP o avaliador enquadrou o bem-avaliando. O que a NBR 14.653-2 deveria fixar, talvez, seriam os percentis máximos e mínimos do IP passíveis de utilização, ou melhor, qual seria o nível de confiança aceitável para uma determinada avaliação, a depender da precisão que se requer, não uma porcentagem fixa em torno da média, como faz o campo de arbítrio. Além disso, seria possível estabelecer um intervalo de valores possíveis, em função dos percentis do IP, por exemplo, considerar que um imóvel encontra-se acima da média do mercado, entre os percentis 80% e 90% do IP. Isto possibilitaria a definição de um valor arbitrado e um intervalo justo de valores em torno deste valor arbitrado.

Lembrando, por fim, que nada impediria o avaliador de utilizar os valores médios, caso este considere não haver motivos para depreciar ou valorizar o imóvel em específico no mercado em que se encontra, de acordo com o seu modelo.

REFERÊNCIAS BIBLIOGRÁFICAS

ABNT. ASSOCIAÇÃO BRASILEIRA DE NORMAS TÉCNICAS. **NBR 14653-1:** Avaliação de Bens. Parte 1: Procedimentos Gerais. Rio de Janeiro, 2019.

ABNT. ASSOCIAÇÃO BRASILEIRA DE NORMAS TÉCNICAS. **NBR 14653-2:** Avaliação de Bens. Parte 2: Imóveis Urbanos. Rio de Janeiro, 2011.

ABUNAHMAN, Sérgio Antônio. **Curso básico de engenharia legal e de avaliações**. 4. ed. São Paulo: Pini, 2008.

DANTAS, Rubens Alves. **Engenharia de avaliações:** uma introdução à metodologia científica. 3. ed. São Paulo: Pini, 2012. 255 p.

FARAWAY, J. **Linear models with R.** Taylor & Francis, 2004.

GIANNAKOS, I. B. S.; SILVEIRA, L.B.; LEÃO, M.L. Os limites de confiança e o arbítrio do avaliador. In: IX Congresso Brasileiro de Avaliações e Perícias. **Anais ...** pp. 290-295, 1997. São Paulo: IBAPE.

GIANNAKOS, I. B. S., LEÃO, M. L. Proposta de alteração da NB 502/89: Campo de Arbítrio. In: X Congresso Brasileiro de Engenharia de Avaliações e Perícias. **Anais...**. 1999. Fortaleza: IBAPE.

GUJARATI, D. N; PORTER, D. C. **Econometria básica.** 5. ed. Porto Alegre: AMGH Bookman, 2011.

JAMES, Gareth; WITTEN, Daniela; HASTIE, Trevor; TIBSHIRANI, Robert. **An introduction to statistical learning:** with applications in R. 7. ed. Springer, 2017.

MACEDO, L.F.R; PACHECO, M.P.; EINSFELD, R.A. Uma formulação racional para determinação do nível de confiança em avaliações de imóveis. In: XI Congresso Brasileiro de Avaliações e Perícias. **Anais...**. 2001. Guarapari: IBAPE.

MEEKER, W.Q.; HAHN, G.J; ESCOBAR, L. A. **Statistical Intervals:** a guide for practitioners and researchers. 2. ed. Wiley, 2017. 648 p.

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

ANEXO A – DADOS DE MERCADO UTILIZADOS NO ESTUDO

IDENT	Valor_Total	Area_Total	Quartos	Suites	Garagens	Dist_B_Mar	Padrão
AP_01	1.060.000,00	350,00	3	1	2	720	2
AP_02	510.000,00	136,56	3	1	1	665	2
AP_03	780.000,00	164,77	3	1	2	415	2
AP_04	550.000,00	174,58	3	1	1	320	2
AP_05	850.000,00	123,01	3	1	3	895	3
AP_06	300.000,00	89,83	2	0	1	645	1
AP_07	750.000,00	174,00	2	1	2	860	3
AP_08	650.000,00	123,00	3	1	1	745	3
AP_09	620.000,00	121,00	3	1	1	745	3
AP_10	740.000,00	109,00	3	1	1	300	2
AP_11	770.000,00	170,00	3	1	2	590	2
AP_12	680.000,00	141,00	3	1	1	290	2
AP_13	850.000,00	174,00	3	1	1	465	2
AP_14	420.000,00	105,00	3	1	0	60	1
AP_15	547.000,00	128,00	3	1	1	745	3
AP_16	1.600.000,00	163,00	4	2	2	90	3
AP_17	1.320.000,00	230,00	3	1	2	215	3
AP_18	615.000,00	108,00	3	1	1	745	3
AP_19	705.000,00	174,00	2	1	2	900	3
AP_20	418.000,00	85,00	1	0	1	620	3
AP_21	270.000,00	71,00	2	0	0	1380	1
AP_22	418.000,00	100,00	1	1	1	620	3
AP_23	650.000,00	90,00	2	1	1	215	3
AP_24	700.000,00	161,00	2	1	2	500	3
AP_25	680.000,00	174,00	2	1	2	860	3
AP_26	420.000,00	76,00	2	1	1	700	1
AP_27	195.000,00	48,00	1	0	0	730	1
AP_28	290.000,00	66,00	1	0	1	745	1
AP_29	272.000,00	50,00	1	0	1	1430	1
AP_30	430.000,00	61,00	2	0	1	170	1
AP_31	895.000,00	109,00	3	1	1	530	2
AP_32	450.000,00	89,00	2	0	1	745	2
AP_33	1.950.000,00	393	3	1	3	550	3
AP_34	2.150.000,00	578	3	2	3	260	3
AP_35	940.000,00	182	3	1	2	200	2
AP_36	1.400.000,00	262	4	1	1	60	3
AP_37	1.090.000,00	205,00	3	0	3	465	2
AP_38	1.272.000,00	196	3	3	2	610	3
AP_39	2.800.000,00	463	3	3	3	590	3
AP_40	1.796.000,00	273	3	3	4	140	2
AP_41	1.400.000,00	330	4	2	2	655	3
AP_42	3.000.000,00	533	4	3	4	427	3
AP_43	1.200.000,00	221	3	3	2	607	3
AP_44	800.000,00	220	3	1	1	1000	2
AP_45	950.000,00	127	2	1	1	60	2
AP_46	2.061.000,00	362,00	3	3	4	310	3
AP_47	1.326.000,00	315,00	3	3	3	600	3
AP_48	850.000,00	151,00	3	1	2	660	2
AP_49	1.650.000,00	246,00	3	3	3	307	3
AP_50	650.000,00	159,72	3	1	1	120	2