Estudo Dirigido

Adrian Giovanne[2], Gustavo Mendel[3], João Vitor[3], Lucas Lima[3],
Paulo Mascarenhas[1], Romildo Andrade[1] e Tiago Avelino[1].
Engenharia Elétrica[1], Engenharia Mecânica[2], Ciência da Computação[3]
MAT236-T07
Professora: Renata Esquivel

Resumo: O reconhecimento e a aplicação de probabilidade e distribuição de probabilidade são imprescindíveis em diversas áreas da ciência, tecnologia e negócios. A probabilidade é a medida da chance de um evento ocorrer, enquanto a distribuição de probabilidade é a descrição matemática da probabilidade de todas as possíveis ocorrências de um evento aleatório. Compreender esses conceitos é crucial para a modelagem estatística e para a tomada de decisões informadas em situações incertas. Neste estudo dirigido, fomos requisitados a pesquisar por artigos científicos que utilizem de métodos estatísticos para que possamos explorar os principais tipos de distribuições de probabilidade e como foram aplicadas. O artigo escolhido possui o seguinte título: "Fluxo de potência trifásico probabilístico para redes de distribuição usando o método de estimação por pontos".

1. Introdução

À princípio, o artigo começa introduzindo ao leitor a importância do fluxo de potência para a obtenção de dados valiosos de um determinado sistema elétrico. É apresentado que os métodos de solução do fluxo de potência são modelados de forma determinística, ou seja, só admitem valores fixos, que é algo que não condiz com a realidade, haja vista que em sistemas elétricos reais, esses dados de entrada estão sujeitos à incertezas, tais como erros nas medições, imprecisão nos cálculos, incerteza na distribuição das cargas pelas fases, etc. Desse modo, para considerar essas incertezas no cálculo do fluxo de potência, foram propostas várias metodologias, que em suma, são divididas em três grandes grupos: Métodos Analíticos, Métodos de Simulação e Métodos Híbridos. Os analíticos são computacionalmente mais efetivos, mas requerem hipóteses matemáticas. Em relação aos métodos de simulação, o autor cita o Método de Monte Carlo (SMC) - que inclusive é levado em consideração na comparação dos resultados - e o define como um procedimento aleatório, que permite determinar probabilisticamente os possíveis valores das variáveis de interesse em determinado instante. Ou seja, a SMC é aplicada para considerar a incerteza natural das cargas dos sistemas elétricos, o que faz com que as variáveis de interesse sejam modeladas através de funções de distribuição de probabilidade.

Assim, o estudo dirigido visa apontar os métodos utilizados pelos autores, e relacioná-los com as abordagens apresentadas em sala de aula. Nesse sentido, foi selecionado um artigo científico dentro de uma área correlata a todos os membros da equipe. O artigo aborda o problema do fluxo de potência em sistemas trifásicos desbalanceados, buscando estimar a magnitude e a fase da potência nas linhas de transmissão através do método da estimação por pontos. Este método busca determinar as FDP (funções de distribuição de probabilidade) das grandezas desejadas no problema. Ademais, uma vez obtidas as funções de probabilidade desejadas, e computados as respectivas médias e desvios padrões, é possível realizar outras estimações também vistas em sala de aula, como distribuição pela normal, Poisson, dentre outras.

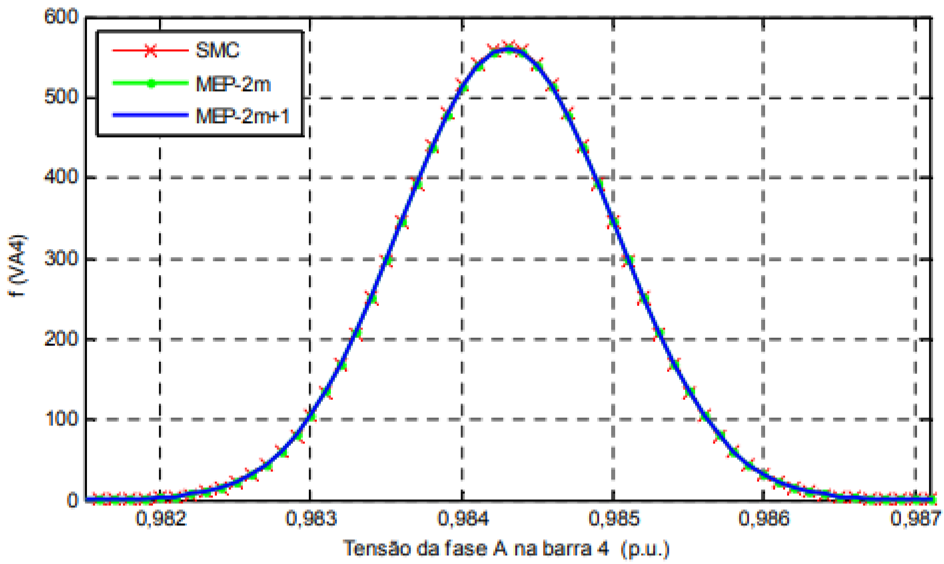
2. Método Estatístico

O método, ou ferramental estatístico empregado na solução do problema, foi o método da estimação por pontos, também conhecido como MEP. Os métodos de MEP foram desenvolvidos devido às dificuldades inerentes à determinação das FDP das variáveis estocásticas. O MEP empregado, é um caso particular do método de estimação por pontos proposto em (Hong, 1998), que visa calcular os momentos estatísticos das variáveis aleatórias em questão. Tal método obteve sucesso ao resolver o problema do fluxo de potência probabilístico em sistemas de transmissão de pequeno porte, assim como também obteve sucesso ao resolver o problema do fluxo de potência ótimo probabilístico (Verbic et al., 2006).

Uma estimativa pontual de uma característica da população é um único número que é baseado em dados amostrais e representa um valor plausível da característica de interesse. A estimativa pontual é obtida selecionando-se, primeiro, uma estatística adequada.

Algumas vezes, mais de uma estatística pode ser usada para obter uma estimativa pontual de uma característica da população. Em termos gerais, a estatística utilizada deve ser aquela que tende a produzir uma estimativa acurada, isto é, uma estimativa perto do valor da característica de população. Informações sobre a precisão da estimativa de uma determinada estatística é fornecida pela distribuição de amostragem da estatística, como no exemplo abaixo:

Figura 1: Função Distribuição de Probabilidade.



(a) FDP

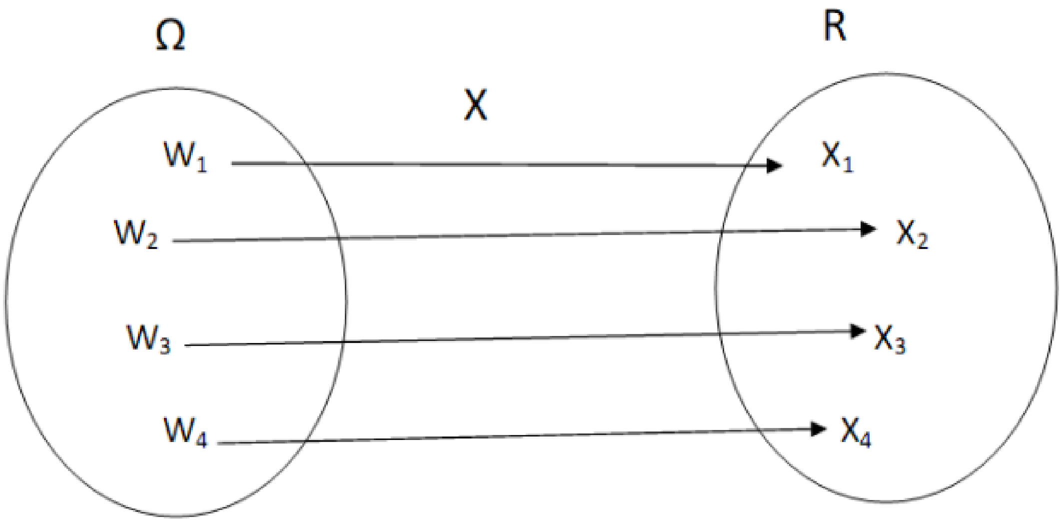
Fonte: Artigo base para o estudo.

3. Variável Aleatória e Distribuição de Probabilidade

Definições básicas são importantes para o desenvolvimento dos métodos estatísticos mais rebuscados utilizados no artigo. Desse modo, pudemos perceber que variáveis aleatórias e distribuição de probabilidade - assuntos que foram vistos na primeira unidade - são conceitos recorrentes nesse tipo de pesquisa.

Assim, começamos definindo Variável Aleatória: Sejam E um experimento e Ω um espaço amostral associado ao experimento. Uma função X que associe a cada elemento Wi ∈ Ω um número real, X(Wi), é denominada variável aleatória.

Figura 2: Definição variável aleatória.



Fonte: Apostila.

Seja X uma variável aleatória. Se X assume valores em um conjunto finito ou infinito enumerável, então X é denominada variável aleatória discreta.

Seja X uma variável aleatória. Se X assume valores em um conjunto infinito não enumerável, então X é denominada variável aleatória contínua.

Seja X uma variável aleatória discreta que assume os valores x_1, x_2, ... ,x_n. A distribuição de probabilidades de X é o conjunto de pares de valores que associa a cada valor da variável x_i a probabilidade $P(X = x_i)$:

$$(x_1, P(X = x_1)),(x_2, P(X = x_2)), \ldots,(x_n, P(X = x_n))$$

De maneira que

$$\sum_{i=1}^{\infty} P(X = x_i) = 1$$

$$P(X = x) = p(x) \geq 0$$

Seja X uma variável aleatória contínua. A distribuição de probabilidade é dada na forma de uma função, chamada de densidade de probabilidade e denotada por f(x).

Uma função de densidade de probabilidade (fdp) satisfaz as seguintes condições:

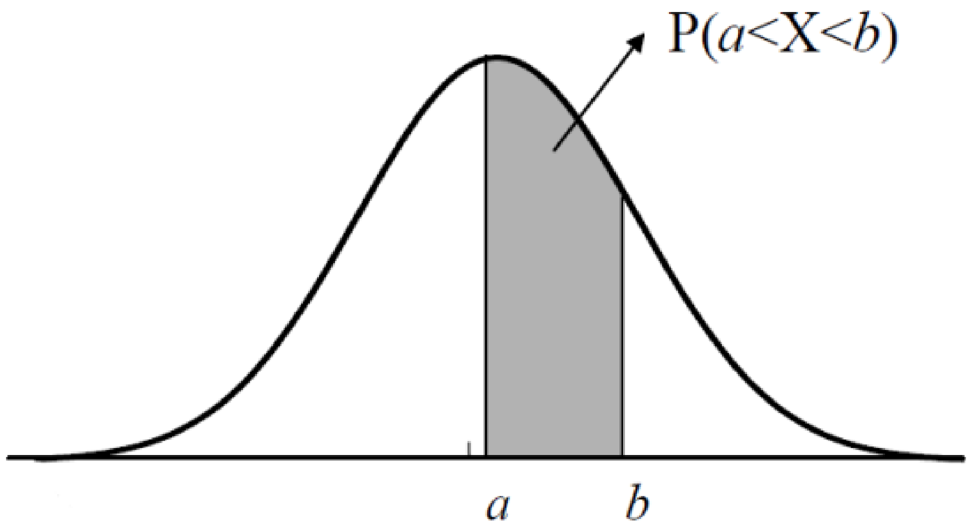
$$f(x) \geq 0, \ \forall \ x \in R$$

$$\int_{-\infty}^{+\infty} f(x)dx = 1$$

A função densidade, por definição, possui área sob a curva limitada pelo eixo x igual a 1 e a probabilidade de X tomar um valor entre a e b é obtida calculando-se a área compreendida entre esses dois valores. Isto é, para qualquer a < b em R.

$$P(a < X < b) = \int_{a}^{b} f(x)dx$$

Figura 3: Função densidade.



Fonte: Apostila.

Observações importantes para uma variável aleatória contínua:
Qualquer valor especificado de X tem probabilidade zero, isto é, $P(X = xi) = 0$, pois:

$$P(X = x_i) = \int_{xi}^{xi} f(x)dx = 0$$

Assim, as probabilidades abaixo serão todas iguais, se X for uma variável aleatória contínua:

$$P(a \leq X \leq b) = P(a \leq X < b) = P(a < X \leq b) = P(a < X < b)$$

Uma vez definida a função densidade de probabilidade das variáveis abordadas nos problemas. Vários modelos de aproximações de distribuições estão ao dispor a depender da natureza da variável aleatória, seja esta contínua ou discreta, binária ou real, quantitativa ou qualitativa (não abordada neste momento). No problema, definindo uma média e uma variância, é possível modelar curvas normais, aproximações binomiais, modelos exponenciais de Poisson e parâmetros.

4. O problema do fluxo de potência probabilístico

O problema pode ser modelado matematicamente por dois conjuntos de equações

$$Y = G(X)$$
$$Z = F(X)$$

Sendo:

Y: Vetor da inserção de potências

Z: Vetor das variáveis de saída

X: Vetor das variáveis de estado

G,F: Funções do fluxo de potência

A solução desse sistema para o i-ésimo termo de Z, pode ser escrito da seguinte forma:

$$Z_i = F_i(p_1, p_2, \ldots, p_l, ..., p_m)$$

Sendo p_l um parâmetro que pode representar a inserção de potência ou uma impedância do sistema. No artigo, considera-se p_l somente considera a incerteza na inserção de potências.

5. O MEP para 2 pontos

Dado que para este problema, apenas as variáveis incertas são as inserções de potência, então o vetor $W = (p_1, p_2, \ldots, p_l, ..., p_m)$ tem apenas elementos das incertezas das inserções de potência. Cada elemento p_l possui uma FDP f_{pl}. Para dois pontos, são selecionadas duas concentrações ou pontos os quais possuem a máxima probabilidade de de ocorrência ($p_{l,1}$ e $p_{l,2}$), cujos valores podem ser estimados como:

$$p_{l,k} = \mu_{pl} + \xi_{l,k}\sigma_{pl}$$

Sendo μ_{pl} e σ_{pl} a média e o desvio padrão da FDP $f_{pl.}$ O valor de ξ pode ser calculado através dos coeficientes de assimetria, mas a noção desse conteúdo não contribui o escopo deste

estudo dirigido. O fato é, uma vez determinados os pontos $p_{l,1}$ e $p_{l,2}$, podemos estimar os valores das funções $Z_i(l,1)$ e $Z_i(l,2)$, executando um fluxo de potência para cada funções.
É necessário ainda, calcular os momentos estatísticos das entradas e saídas. Para as entradas, os momentos são calculados como:

$$w_{l,k} = (1/m) * (-1)^k * (\xi_{l,3-k}/\varsigma_l)$$

onde m é o número de entradas e ς é determinado através dos coeficientes de assimetria.
Uma vez computado o peso das entradas, é possível estimar as esperanças das saídas, como:

$$E(Z_i^j) \cong \sum_{l=1}^{m} \sum_{k=1}^{3} w_{l,k} \times [Z_i(l,k)]^j$$

Assim, computando $\mu = E(Z)$ e $E(Z^2)$, é trivial calcular $\sigma^2 = Var(Z) = E(Z^2) - (E(Z))^2$. Uma vez obtidos os parâmetros de média e variância, podemos estimar FDP de uma curva normal para a predição dos resultados. É possível estimar ainda a função de densidade acumulada FDA. Onde:

$$N(\mu, \sigma^2) = \frac{1}{\sqrt{2\pi\sigma}} exp(\frac{-1}{2} (\frac{x-\mu}{\sigma})^2)$$

A integral dessa função é não elementar, porém é possível estimá-la para pontos arbitrários ao invés de resolvê-la analiticamente. Os valores padronizados são tabelados para facilidade de cálculo e agilidade no tratamento dos dados.

6. Auto avaliação

A seguir, temos a tabela de auto avaliação, requisitada pela professora. Nesta tabela, cada participante do grupo deverá avaliar a si mesmo e aos seus colegas de acordo com sua percepção da participação dos mesmos na concepção e realização do trabalho. Essa avaliação será feita numa escala de cinco a dez, e cada linha representa o conjunto de notas que um dos integrantes do grupo atribuiu para os demais.

Tabela 1: Tabela de auto avaliação.

	Avaliados						
Avaliadores	**Paulo**	**Tiago**	**Romildo**	**João**	**Lucas**	**Gustavo**	**Adrian**
Paulo	10	10	10	10	10	10	10
Tiago	10	10	10	10	10	10	10
Romildo	10	10	10	10	10	10	10
João	10	10	10	10	10	10	10
Lucas	10	10	10	10	10	10	10
Gustavo	10	10	10	10	10	9	10
Adrian	10	10	10	10	10	10	10

Fonte: Elaborada pelos autores.

7. Conclusão

Pudemos concluir, portanto, que a estatística é de extrema importância em diversas áreas da ciência. Os métodos estatísticos fornecem ferramentas para coletar, analisar e interpretar dados, permitindo que pesquisadores e profissionais tomem decisões informadas em situações incertas. Além disso, a estatística ajuda a identificar padrões e tendências nos dados, permitindo que sejam feitas previsões e projeções sobre o futuro. Sem esses métodos, muitas pesquisas seriam impossíveis ou imprecisas, e muitas decisões importantes seriam baseadas em intuição ou suposições infundadas. Desse modo, percebemos o quão valiosa é a estatística como uma ferramenta para aqueles que buscam entender o mundo ao seu redor e tomar decisões com base em dados confiáveis.

Ademais, especificamente a respeito sobre o método dos pontos, nossa equipe foi capaz de notar como ele não é viável como a única forma de estimar um valor. Devido a sua natureza pontual, o MEP é mais propenso a erros do que outros tipos de estimativas, dando ênfase a estimativas intervalares, que usam de um intervalo de dados para calcular de maneira mais correta uma esperança para a estimativa.

Também é preciso mencionar que o valor de uma estimativa pontual depende de qual amostra foi selecionada, de todas as amostras possíveis, então a escolha de amostras adequadas se torna ainda mais crucial neste tipo de estimativa do que em outras.

Ao trazer estas descobertas para o contexto dos assuntos vistos durante as aulas, nossa equipe pode notar como a utilização de variáveis contínuas em intervalos ao invés de variáveis discretas e pontos pode fornecer dados mais corretos e fáceis de analisar, visto que um intervalo pode ser estudado e manipulado muito mais do que um ponto.

Além disso, o estudo do artigo também nos mostrou como é rara a utilização de um único método para tratar dados e como é valorosa a utilização de diversos métodos, visto que isso possibilita a comparação entre métodos, o que facilita a identificação de erros e pode ajudar com o tratamento de dados.

8. Referências

[1] - GALLEGO, Luis; ECHEVERRI, Mauricio; PADILHA-FELTRIN, Antonio. Fluxo de potência trifásico probabilístico para redes de distribuição usando o método de estimação por pontos. **Revista Controle & Automação**, Vol. 23, p.180-189, Março, 2012. Disponível em: https://www.scielo.br/j/ca/a/L8VDF4kZJNt66byYCkQ3MJb/?lang=pt&format=pdf

[2] - AZEVEDO, Caio. Métodos de estimação (pontual). **Aula UNICAMP**. Disponível em: https://www.ime.unicamp.br/~cnaber/aula_Met_Estim_Mest_2S_2019.pdf

[3] - Hong, H. P. (1998). An efficient point estimated method for probabilistic analysis, Reliability Engineering and System Safety, Barking 59(3): 261–267.

[4] - SILVA, Giovana; MORAES, Lia; CASTRO, Rosana; FIACCONE, Rosemeire. Notas de aula MAT236 - Métodos Estatísticos - 1ª Unidade. **Apostila UFBA.**