

“Aplicação da Regressão linear múltipla no estudo da tensão arterial nas gestantes”

Application of Multiple Linear Regression in the Study of Blood Pressure in Pregnant Women

Makambo Kapusa Bo Eustache ^{1*}

¹ Phd. Professor assistente. Instituto Politécnico de Saurimo. makambostch87@gmail.com ORCID: <https://orcid.org/0000-0003-3709-6117>

* Autor para correspondência: makambostch87@gmail.com

RESUMO

O presente artigo focaliza-se nas gestantes, utentes do Centro Médico Dom Eliseu Maria Corolli, objetivando apresentar alguns resultados sobre as variáveis que influenciam a tensão arterial delas. Neste estudo de caso foram recolhidos dados por observação directa, inquérito e entrevista tanto ao pessoal de saúde, quanto aos utentes. O estudo foi feito numa abordagem estatística com análise de regressão linear para medir os efeitos de variáveis que incidem na tensão arterial. Com um enfoque maior na análise de regressão linear múltipla que se distingue da regressão linear simples quando incorporadas mais do que uma variável independente no modelo, debruçou-se da Análise de Variância, salientando o teste F para a validação das hipóteses. Os cálculos foram feitos com auxílio do MICROSOFT Excel usando a ferramenta de análises de dados para averiguar os valores do nível descritivo p da tensão arterial sistólica e diastólica, a fim de revelar o relacionamento e o grau de correlação entre as variáveis.

Palavras-chave: Regressão Linear Múltipla, Estimadores, Testes de Hipóteses, Tensão arterial, Sistólica, Diastólica.

ABSTRACT

This article focuses on the effects of variables affecting arterial tensile strength in pregnant women at the Centro Médico Dom Eliseu Maria Corolli. Data were collected through direct observation, interview, and interviews with both the patient and the nurses. The study used a statistical approach with linear regression analysis to measure the effects of variables on arterial tensile strength. The analysis used a more advanced approach, excluding ANOVA and using the F test for hypothesis validation. The calculations were performed using Microsoft Excel and data analysis tools to determine the values of the p-level of systolic and diastolic arterial tensile strength, revealing the relationship and correlation between variables.

Keywords: *multiple linear regression, estimators, hypothesis tests, systolic and diastolic blood pressure.*

INTRODUÇÃO

A análise de regressão, uma técnica estatística para examinar e modelar a relação entre variáveis, foca no estudo de uma variável dependente em relação a uma ou mais variáveis, as variáveis explicativas, com o objetivo de estimar e/ou prever a mediana da dependente (da população) ou o valor médio em termos de valores conhecidos ou fixos (repetidos) das explicativas (Ferreira, 2013, p.1). É uma das mais utilizadas na análise de dados, com o objetivo de estimar os parâmetros desconhecidos do modelo. No campo da saúde, há inúmeras aplicações da análise de regressão para determinar as variáveis que influenciam a tensão arterial em gestantes (alta ou baixa).

Existem vários métodos para estimar estes parâmetros e, este artigo tem um enfoque nos métodos dos mínimos quadrados.

Problemática

Em Luanda, como nos demais Centros hospitalares e maternidades do País, vivenciam-se situações desastrosas devido ao estado sanitário das mães, sobretudo gestantes. Segundo dados estatísticos da Associação das Gestantes e Obstetras de Angola (AGOA) nas gestantes um dos motivos de partos com efeitos ou consequência que levam a mortes tanto aos bebés e/ou mães além das doenças crónicas stress ou traumatismo pós-parto, a tensão arterial é apontada como um dos fatores que favorece tais casos, conforme afirmado por Mário Fernando da Sociedade Angolana de Doenças Cardiovasculares no Jornal Angop (Disponível em <https://www.angop.com>).

Durante a gravidez as mulheres por certas situações vivenciadas como: estilos de vida inadequados, alimentação não saudável, sedentarismo, stress e outros problemas que podem desenvolver tensão arterial nas gestantes, que esta por sua vez segundo a Organização Mundial da Saúde (OMS)¹ é uma doença que faz com que as gestantes e o feto correm muitos riscos. As gestantes devem regularmente medir a sua tensão arterial para evitar certas complicações como a pré-eclâmpsia ou eclâmpsia que é caracterizada por crises convulsivas e o seu prognóstico é grave para a mãe e para a criança. Surge geralmente no final da gravidez, durante o trabalho de parto, e por vezes depois do parto.

Devido aos elevados números de mortes nas maternidades por causa de hipertensão ou hipotensão nas gestantes levou nos ao seguinte problema que se resume na necessidade de analisar as variáveis que influenciam a tensão arterial nas gestantes

O problema acima destacado assenta-se na Regressão linear múltipla aplicada a Tensão arterial nas gestantes do Centro Médico “Dom Eliseu maria Corolli”, no Bairro Simone / Comuna de Camama no Município de Talatona - Luanda. Portanto, neste artigo objetiva estudar a relação entre tensão arterial e as variáveis que a influenciem. .

Conhecendo as variáveis fortes que influenciam a sermos hipertensos ou hipotensos, as gestantes vão adotar novas medidas que o lhes possibilitarão a não desenvolver esta patologia ou mesmo controlar melhor a sua tensão arterial. Sendo que, estabelecendo o grau de correlação entre a tensão arterial das gestantes, é possível se evitar danos colaterais a saúde pública a nível dos utentes do Centro Médico Dom Eliseu Maria Corrolli.

Segundo a OMS (Organização Mundial da Saúde)², tem se notado que as doenças cardiovasculares tem sido um problema de saúde pública em muitos países, e a tensão arterial tem sido também uma das patologias que causa as doenças cardiovasculares, razão pela qual levou-nos a trabalhar com as mulheres gestantes, por estas serem consideradas como sendo um grupo de risco, interessou-nos a fazer a simbiose entre a estatística que é uma das áreas conhecidas como teoria das tomadas das decisões e a saúde.

¹ OMS (2021) "Relatório lançado pela OMS detalha o impacto devastador da hipertensão e as formas de combatê-la", publicado aos 19 de setembro de 2023** e tem o título:

² Relatório "OMS revela: mais de 700 milhões de hipertensos estão sem tratamento" (2021)

A análise de regressão linear surgiu a partir da necessidade de cientistas descobrirem, através do cálculo de probabilidades se algumas características físicas e psicológicas, entre membros de uma mesma família poderiam estar associados de forma que fosse possível explicá-las através de um modelo matemático. Nesta pesquisa, foi estudada a estimação dos parâmetros de regressão linear múltipla, o coeficiente de determinação, testes de hipóteses, e somas de quadrados.

Para exemplificar os procedimentos teóricos de análise de regressão múltipla, foi utilizado dados referentes a tensão arterial nas gestantes do Centro Médico Dom Eliseu Maria Corolli durante o mês julho de 2020.

Sendo a tensão arterial um problema de saúde pública, que na qual por meio dela podemos desenvolver algumas patologias como: Doenças cardíacas, AVC (derrame cerebral), insuficiência renal crônica, aneurismas e lesões nos vasos sanguíneos dos olhos etc. E atendendo aos elevados problemas que a sociedade enfrenta, todos estamos propensos a adquirir a mesma patologia, razão pela qual com este trabalho pretendemos que as gestantes de uma maneira em geral adotem hábitos de medição regular da tensão arterial. Porque quanto mais cedo for diagnosticado melhor será controlada. Assim se estabelece o grau de correlação entre a tensão arterial em gestantes é possível evitar danos colaterais e melhorar a saúde pública a nível do Centro Médico Dom Eliseu Maria Corolli.

Apresentação do Centro Médico Dom Eliseu Maria Corolli

O Centro Médico Dom Eliseu Maria Corolli sita no bairro Simeone Mucune, comuna da Camama, ao oeste do topo do Cemitério do Camama, comparticipada com o governo, construído no ano de 2003, iniciando o seu atendimento ao público no dia 22 de fevereiro de 2005, geograficamente pertencente ao município do Talatona, Distrito do Camama. O centro atende os seguintes serviços: Pediatria, Medicina, Consulta pré-natal (CPN), Puericultura, programa alargado de vacinação (PAV), Laboratório e Farmácia.

A unidade sanitária funciona com 32 recursos humanos distribuídos da seguinte forma:

Tabela.1

Dados estatísticos do 1º semestre de 2020.

Médico	Enfermeiros	Estatístico	Serviços gerais	Auxiliar de Diagnóstico Terapeuta
1	15	1	1	4
Cozinheira	Segurança	Auxiliar de Limpeza	Técnico de gestão e recursos humanos	
1	6	2	1	

Tabela 2.

Realizou-se 11.553 Consultas, fazendo uma média diária de 76 consultas distribuídos da seguinte maneira:

Nº	Denominação	Número de consultas
1	Medicina	2.854
2	Pediatria	2.905
3	Puericultura	3.856
4	C.P.N	402
5	Vacinas em crianças, adultos e gestantes.	8.126
Exames laboratoriais estão distribuídos da seguinte maneira:		
6	Hematologia	850
7	Grupo sanguíneo	367
8	Serologia	2037
9	Urinas	794

10	Vacinas em crianças, adultos e gestantes.	11.245.
11	Fezes	13
12	Gota espessa	5452
13	VS	1124
14	BK	0
15	Falciformação	12
16	Outras	596

Conceitos gerais

1. Regressão Linear Múltipla (RLM)

O modelo de regressão múltipla com p variáveis é definido da seguinte forma

$$Y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \dots + \beta_p x_{ip} + \varepsilon_i, i = 1, \dots, n, ()$$

Em que:

y_i representa o valor da variável resposta na observação i , $i = 1, \dots, n$;

$x_{i1}, x_{i2}, \dots, x_{ip}$, $i = 1, \dots, n$ são valores da i -ésima observação das p variáveis explicativas, (constantes conhecidas);

$\beta_0, \beta_1, \beta_2, \dots, \beta_p$ são os parâmetros ou coeficientes de regressão.

e_i $i = 1, \dots, n$ Correspondem aos erros aleatórios

Os parâmetros b_j , $j = 1, \dots, p$, representam a média esperada na variável resposta, y , quando a variável x_j , $j = 1, \dots, p$, sofre um acréscimo unitário, enquanto todas as outras variáveis x_k , $k \neq j$, são mantidas constantes. Por esse motivo os b_j , $j = 1, \dots, p$, são chamados de coeficientes parciais. O parâmetro β_0 corresponde ao intercepto do plano de regressão. Se a abrangência do modelo incluir $x_j = 0$, $j = 1, \dots, p$, então β_0 será a média y de nesse ponto. Caso contrário não existe interpretação prática para β_0 .

Vamos considerar o caso particular do modelo de regressão linear múltipla com duas variáveis explicativas x_1 e x_2 . Assim, o modelo será definido por

$$y_i = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \varepsilon \quad (1)$$

1.1 Coeficiente de determinação e coeficiente determinação ajustado

O Coeficiente de determinação é igual ao quadrado do coeficiente de correlação de Pearson, que agora poderá ser reescrito da seguinte forma

$$R^2 = \frac{\text{Variação explicada}}{\text{Variação Total}} = \frac{SQR}{SQT} = 1 - \frac{SQE}{SQT} \quad , \text{ com } 0 \leq R^2 \leq 1, \quad (12)$$

Este coeficiente, que é uma medida da proporção da variação da variável resposta Y explicada pela equação de regressão, é usado para quantificar a capacidade explicativa do modelo. Deve-se ter cuidado, no entanto, ao facto de que $R^2 \cong 1$ não significa que o modelo de regressão providencia um bom ajustamento aos dados, dado que a adição de uma variável aumenta sempre o valor deste coeficiente (mesmo que tenha muito pouco poder explicativo sobre a variável resposta). Assim sendo, é preferível utilizar o coeficiente de determinação ajustado, que é uma medida ajustada do coeficiente de determinação e que é “penalizada” quando são adicionadas variáveis pouco explicativas.

O coeficiente de determinação ajustado é definido por:

$$R^2_a = 1 - \left(\frac{n-1}{n-(p+1)} \right) (1 - R^2) \quad (13)$$

Ao contrário do coeficiente de determinação, o coeficiente de determinação ajustado, não aumenta sempre quando adicionamos uma nova variável. Aliás, se adicionarmos variáveis com pouco poder explicativo este tende a decrescer. (Rodrigues, 2012. Pág. 32)

1.2 Teste F para significância da equação de regressão linear múltipla

Aqui testa-se a existência da regressão linear no modelo, fazendo-se uso da estatística F

$$F = \frac{QMReg}{QMRes}$$

Utilizado para testar as hipóteses a seguir:

$$\begin{cases} H_0: \beta_1 = \beta_2 = \beta_3 = \dots = \beta_i = 0 \\ H_1: \text{pelo menos } \beta_i \neq 0, i = 1, 2, \dots, n \end{cases}$$

Se $F_{calculado}$ for maior que $F_{tabelado}$, onde F tem distribuição F de Snedecor com $p - 1$ e $n - p$ graus de liberdade, então rejeita-se H_0 e pode se afirmar que, ao nível α de significância, pelo menos um $\beta_i \neq 0$, portanto pode se dizer que existe regressão linear entre as variáveis do modelo.

Teste F para as partes de um modelo de regressão linear.

A influência de uma variável explicativa no modelo de regressão linear múltipla pode ser determinada pelo teste F parcial. Dessa forma, avalia-se contribuição de uma variável explicativa para a soma dos quadrados devido a regressão, depois que todas as outras variáveis independentes foram incluídas no modelo. Assim, influência desta variável x_i do modelo, para soma de quadrados da regressão será estimada por:

$$SQReg(x_i) = SQR_{(total)} - SQReg_{(total \text{ } x_i)}$$

As hipóteses que serão testados são:

H_0 : A variável x_i não melhora significativamente o modelo

H_1 : A variável x_i melhora significativamente o modelo

Uma forma que nos permite mostrar teste é o $F_{calculado}$ dado pela expressão (Ferreira, 2013, Pág.37)

$$F_c = \frac{SQRegx_i}{QMRes}$$

DESENVOLVIMENTO

2.1. Metodologia aplicada

Segundo Severino (2000), existem vários tipos de pesquisa: pesquisa bibliográfica, pesquisa de campo, pesquisa experimental, pesquisa documental, pesquisa histórica, etc.

A pesquisa que se projeta, em prol ao alcance dos objetivos, será abordada numa perspectiva quantitativa, pois serão expressos números para análise e interpretação. Segundo Silva (2012, p.28) “Reconhecem-se os dados enfatizando os números que permitem verificar a ocorrência ou não das consequências, e daí a aceitação ou não das hipóteses. Neste momento recorre-se a estatística ou outras técnicas matemáticas”.

Para entender o funcionamento de análise de regressão, um levantamento bibliográfico (estudo sistematizado desenvolvido com base a material público em livros, revistas, jornais, redes eletrônicas) que explica os diferentes tipos de modelos de regressão. Quanto aos objetivos de pesquisa, teve carácter exploratório. Para Silva (2012, p.59), “Pesquisa exploratória tem como objetivo proporcionar maior familiaridade com o problema, para torna-lo mais explícito ou para construir hipóteses”; o que permitiu analisar e identificar os fatores que influenciam o problema.

2.2. Situação real do Centro Dom Eliseu Maria Corelli

O Centro, nos primeiros cinco (5) meses, atendeu no ano de 2020, 335 pacientes em consultas pré-natais e o nosso estudo, ou seja, os levantamentos de dados foram feitos no mês a seguir, isto é, os dados em estudo levantaram-se no período de um mês, estamos a falar do mês de julho, daí termos a percepção de que mensalmente em média o centro atendeu 67 pacientes. Onde na qual conseguimos obter o número da população gestante do Simeone e outros bairros periféricos que ocorrem na unidade sanitária em referência, então $N = 67$ pacientes. Com base a fórmula utilizada conseguimos então obter o tamanho da amostra $n = 50$, onde considerou se a um nível de significância de 7% ($\alpha = 0.07$) e um intervalo de confiança a 93%. Quanto aos testes de hipótese, utilizamos O nível descritivo p, também conhecido como p-value.

Critérios de inclusão

- Pacientes gestantes que fizeram as consultas no período de julho, nas segundas, quartas e sextas feiras de 2020.
- Gestantes que aceitaram participar na pesquisa, previa explicação dos objetivos.

2.2. Dimensionamento da amostra

Com base em M. E., Martins (2005), considerou-se a fórmula simplificada para determinação do tamanho da amostra (n) com base na estimativa da proporção populacional é $n = \frac{N \cdot 0,5 \cdot (1-0,5) \cdot 2^2}{(N-1) \cdot e^2 + 0,5 \cdot (1-0,5) \cdot 2^2} = \frac{N}{(N-1) \cdot e^2 + 1}$, com $n_0 = \frac{1}{e^2}$. Considerando que para tamanhos de amostras “razoáveis” ou tamanhos grandes, resultados obtidos com N-1 se aproximam dos resultados obtidos com N, em que a fórmula fica $n = \frac{N \cdot n_0}{N + n_0}$, com: N- Representa o tamanho da população; e - representa o limite do erro de estimativa; n_0 - Representa o a primeira estimativa da amostra; n- amostra de população.

Fundamentando nos argumentos anteriores em que se considerou N=67, populações que inclui as gestantes moradoras tanto do Bairro Simeone e dos outros bairros periféricos do município do Talatona, e = 0,07 do erro de estimativa da amostra dos pacientes gestantes e n_0 - a primeira estimativa da amostra,

n-tamanho de amostra;

$$n_0 = \frac{1}{(0,07)^2} = \frac{1}{0,0049} = 204,0816$$

$$n = \frac{67 \times 204,0816}{67 + 204,0816} = 50,44 \cong 50$$

Quanto aos procedimentos de seleção dos indivíduos da amostra, aplicamos a aleatoriedade para nos conceder uma amostra sistemática com base a listagem das fichas de consulta do centro.

RESULTADOS E DISCUSSÃO

O estudo coletou dados por meio de entrevistas baseadas na idade, altura, índice de massa corporal, atividades físicas e uso de tabaco e álcool para aumentar a tensão arterial. A amostra consistiu de 50 participantes de consultas pré-natais, selecionados aleatoriamente em julho de 2020. O estudo teve como objetivo analisar as variáveis que influenciam a tensão arterial. Os dados foram analisados utilizando variáveis descritivas como idade, altura, peso, atividades físicas, uso de tabaco e uso de álcool. Os resultados mostraram que a mediana da resposta da variável dependente para sístole e diástole foi de 130,68 mmHg, enquanto a mediana da resposta da variável independente foi de 25,8, altura 79,43, peso 62,846, atividades físicas 0,24, uso de tabaco 0,02 e uso de álcool 0,4. O estudo encontrou que a idade média variava de 16 a 42 anos, e as mesmas variáveis foram divididas em classes de acordo com o modelo estatístico utilizado pelo Ministério da Saúde para consultas pré-natais.

Tabela 3

Variáveis clínicas em relação a sístole

x_1	x_2	x_3	x_4	x_5	x_6	y_1
Idade	Altura	Peso	Act. Físicas	Uso tabaco	Uso do álcool	Sistólica
21	1,65	65	0	0	1	100
27	1,65	85,5	0	0	0	220
18	1,67	57,4	0	0	1	160
19	1,58	59	0	0	0	140
20	1,54	57,2	1	0	1	180
18	1,54	52,5	0	0	1	200
21	1,57	32,5	0	0	1	140

34	1,55	55,5	1	0	0	170
30	1,59	76	0	0	0	160
22	1,55	55	1	1	0	150
29	1,63	41,1	1	0	1	180
22	1,52	58,6	0	0	0	80
32	1,53	54,5	0	0	0	140
38	1,64	54,5	0	0	1	140
18	1,48	56,6	0	0	0	132
23	1,56	54,5	0	0	1	160
16	1,62	87	0	0	0	160
40	1,64	58,8	1	0	1	160
38	1,52	57,7	0	0	0	120
24	1,68	58	0	0	0	100
17	1,58	68,1	1	0	1	180
22	1,55	66	0	0	1	160
22	1,57	74,5	0	0	0	140
37	1,56	63	1	0	0	120
20	1,6	59	0	0	0	140
25	1,57	66,5	0	0	0	160
20	1,6	74,5	1	0	0	140
42	1,55	63	0	0	0	140
27	1,61	74,5	0	0	1	96
16	1,65	63	1	0	1	119
32	1,63	59	0	0	0	101
29	1,65	66,5	0	0	1	98
20	1,55	60	0	0	0	89
27	1,59	68	0	0	0	89
24	1,67	62	0	0	0	99
21	1,5	42,3	0	0	0	105
24	1,62	66,7	0	0	1	115
28	1,67	81	0	0	0	114
37	1,6	69,3	0	0	0	117
32	1,5	48	0	0	1	129
16	1,58	72	0	0	1	126
17	1,5	52	1	0	0	129
35	1,64	71,5	1	0	0	113
16	1,6	59	0	0	0	111
33	1,51	62	1	0	0	104
25	1,7	71	0	0	1	98
26	1,53	81	0	0	0	119
33	1,62	65	0	0	0	89
19	1,72	80	0	0	1	98
27	1,5	57	0	0	1	104

Tabela 4.

Variáveis clínicas em relação a diástole

x_1	x_2	x_3	x_4	x_5	x_6	y_2
idade	Altura	Peso	Act. Físicas	Uso do tabaco	Uso do álcool	Diastólica
21	1,65	65	0	0	1	60
27	1,65	85,5	0	0	0	120
18	1,67	57,4	0	0	1	80
19	1,58	59	0	0	0	60
20	1,54	57,2	1	0	1	140
18	1,54	52,5	0	0	1	160
21	1,57	32,5	0	0	1	80
34	1,55	55,5	1	0	0	100
30	1,59	76	0	0	0	60
22	1,55	55	1	1	0	160
29	1,63	41,1	1	0	1	80
22	1,52	58,6	0	0	0	180
32	1,53	54,5	0	0	0	60
38	1,64	54,5	0	0	1	100
18	1,48	56,6	0	0	0	180
23	1,56	54,5	0	0	1	160
16	1,62	87	0	0	0	80
40	1,64	58,8	1	0	1	120
38	1,52	57,7	0	0	0	60
24	1,68	58	0	0	0	70
17	1,58	68,1	1	0	1	150
22	1,55	66	0	0	1	80
22	1,57	74,5	0	0	0	80
37	1,56	63	1	0	0	100
20	1,6	59	0	0	0	180
25	1,57	66,5	0	0	0	100
20	1,6	74,5	1	0	0	80
42	1,55	63	0	0	0	120
27	1,61	74,5	0	0	1	120
16	1,65	63	1	0	1	80
32	1,63	59	0	0	0	120
29	1,65	66,5	0	0	1	60
20	1,55	60	0	0	0	90
27	1,59	68	0	0	0	100
24	1,67	62	0	0	0	80
21	1,5	42,3	0	0	0	70
24	1,62	66,7	0	0	1	120
28	1,67	81	0	0	0	80
37	1,6	69,3	0	0	0	80
32	1,5	48	0	0	1	80

16	1,58	72	0	0	1	60
17	1,5	52	1	0	0	120
35	1,64	71,5	1	0	0	100
16	1,6	59	0	0	0	180
33	1,51	62	1	0	0	60
25	1,7	71	0	0	1	80
26	1,53	81	0	0	0	150
33	1,62	65	0	0	0	80
19	1,72	80	0	0	1	60
27	1,5	57	0	0	1	120

As duas tabelas acima podem ser resumidas na seguinte estatística descritiva relativamente aos dados da Sístole e Diástole, gráfico 1 e 2.

Gráfico 1

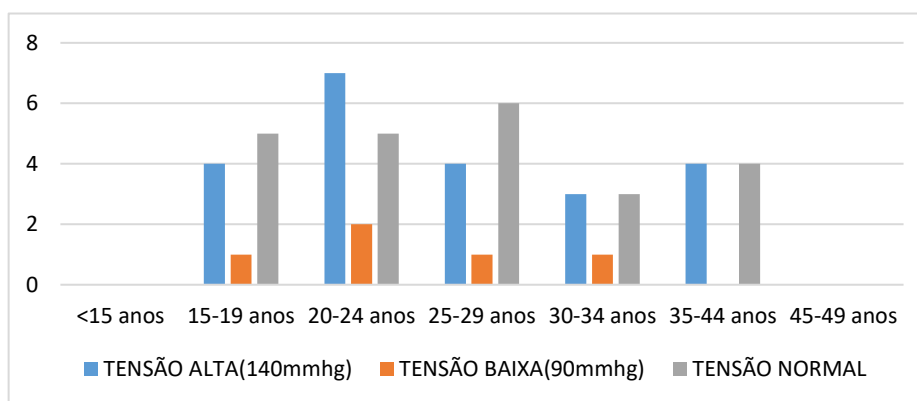
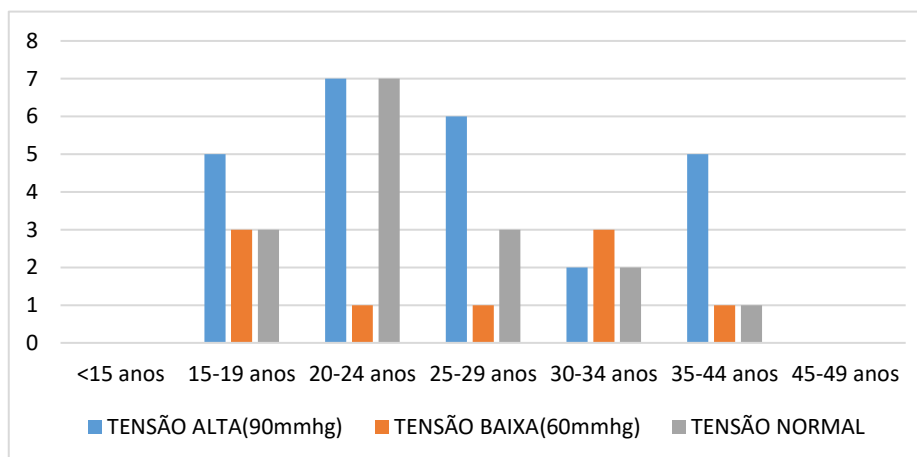


Gráfico 2.



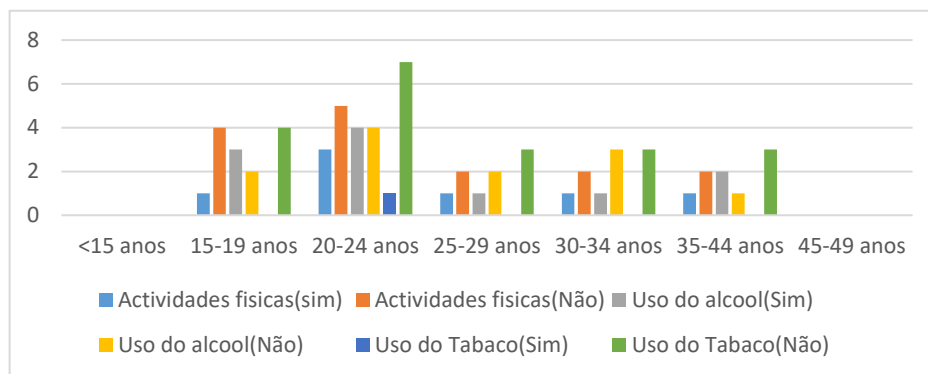
Os gráficos mostram a prevalência da tensão arterial por grupo etário em relação à sístole e diástole. No grupo de 15-19 anos, 4 gestantes têm hipertensão arterial, 1 tem hipertensão arterial baixa e 5 têm pressão arterial normal. No grupo de 20-24 anos, 7 têm hipertensão arterial, 2 têm hipotensão arterial e 5 têm pressão arterial normal. No grupo de 30-34 anos, três têm hipertensão arterial, três têm pressão arterial normal e um tem hipotensão arterial.

O estudo não encontrou registros de idade gestacional em relação à diástole, com 5 gestantes apresentando hipertensão diastólica, 3 apresentando hipotensão arterial e 3 apresentando pressão

arterial normal. Entre 20-24 anos, 7 tinham tensão arterial diastólica alta, 6 tinham tensão arterial baixa, 3 tinham tensão arterial normal, e 2 tinham tensão arterial alta. Entre 35-44 anos, 5 tinham hipertensão arterial, 1 tinha hipotensão arterial, e nenhuma gestante tinha hipertensão e hipotensão arterial diastólica.

Gráfico 3.

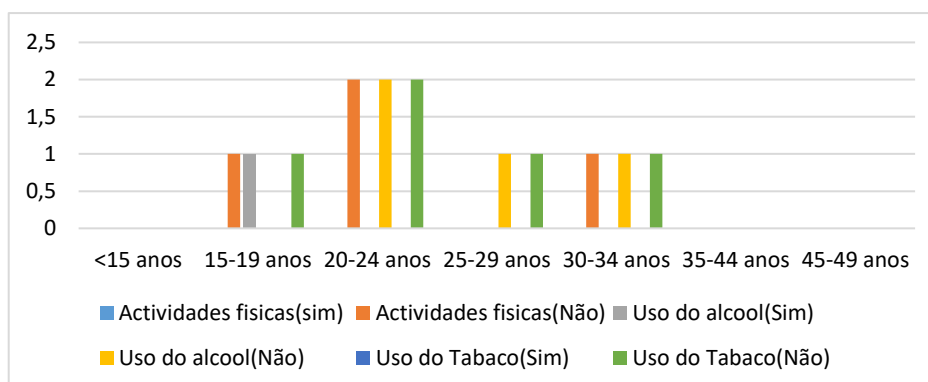
Prevalência da tensão arterial por grupo etário em relação a diástole.



Os resultados apresentam que o grupo etário com maior índice elevado de gestantes com tensão arterial alta é dos 20-24 anos e as gestantes que não fazem o uso do tabaco estão em alta.

Gráfico 4.

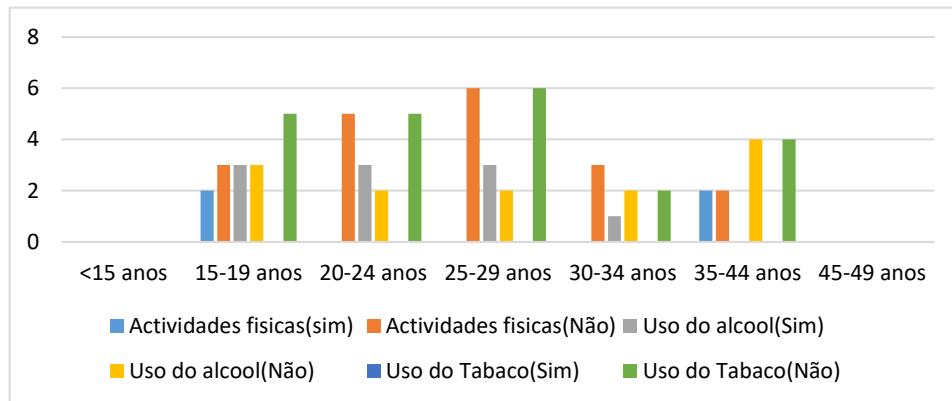
Prevalência da tensão arterial alta sistólica em função das variáveis qualitativa.



Quanto ao gráfico abaixo, temos que o maior grupo etário que apresenta a tensão baixa é dos 20-24 anos, dentre elas a que não praticam actividade físicas.

Gráfico 5.

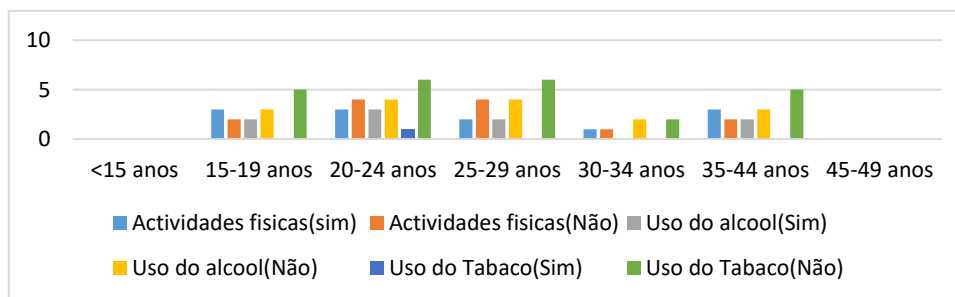
Prevalência da tensão arterial baixa sistólica em função de cada variável qualitativa.



Apresenta-se um registro um pouco equilibrado. Quanto ao grupo etário com maior número de tensão arterial normal. Há um registro com números mais acima ao do grupo etário dos 25-29 anos, com variável em alta as gestantes que não praticam atividades físicas e que não fazem o uso do tabaco

Gráfico 6

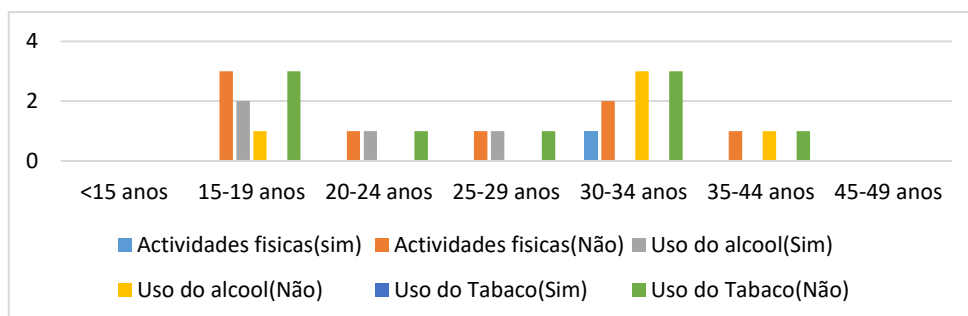
Prevalência da tensão arterial normal sistólica em função de cada Variável qualitativa



Apresenta-nos o grupo etário com maior índice de tensão arterial alta diastólica, dos 20-24 anos com e com uma prevalência das gestantes que não usam tabaco.

Gráfico 7.

Prevalência da tensão arterial alta diastólica em função de cada variável qualitativa.



Prevalência da tensão arterial baixa diastólica em função de cada variável qualitativa

Analisando o gráfico, consegue-se observar um registro de números elevados de tensão arterial baixa nos grupos etários dos 15-19 anos, dos 25-29 anos, dos 30-34 e 35-44 anos, com maior realce as gestantes que não praticam atividades físicas, e as que não usam tabaco

Prevalência da tensão arterial normal diastólica em função de cada variável

O estudo descobriu que o grupo com o maior número de gestantes com pressão arterial normal era o de 20-24 anos. O grupo com pressão arterial alta e baixa foi de 32% do grupo, com 18% tendo pressão arterial alta, 2% tendo pressão arterial baixa e 12% tendo pressão arterial normal. A maioria das gestantes com idades entre 25 e 2 anos tinha pressão arterial alta, com alturas medianas de 1,58 cm e pesos de 61,48 kg. A maioria das gestantes com idades entre 23-24 anos tinha pressão arterial alta, com alturas medianas de 1,55 cm e pesos de 62,2 kg. A maioria das gestantes com 24 anos ou mais tinha pressão arterial alta, com alturas medianas de 1,61 cm e pesos de 62,83 kg.

3- APLICAÇÃO DA REGRESSÃO LINEAR MÚLTIPLA A TENSÃO ARTERIAL EM GESTANTES DO CENTRO MÉDICO DOM ELISEU MARIA COROLLI

1 – Modelo de regressão linear múltipla

Com vista a perceber os fatores que influenciam a tensão arterial nas gestantes, foi realizada uma análise de regressão linear múltipla. Analisaremos inicialmente a relação sistólica com cada uma das variáveis individualmente, através do gráfico de dispersão. Gráfico de dispersão permite decidir empiricamente (Rodrigues, 2015, pág.7):

a)-Se um relacionamento entre as variáveis x e y é assumido.

b)-Se o grau de relacionamento linear entre as variáveis é forte ou fraco.

Para uma análise minuciosa em relação a correlação individualmente entre as variáveis, utilizaremos a seguinte fórmula com as suas condições:

$$r = \frac{\sum_{i=1}^n x_i y_i - n \bar{x} \bar{y}}{\sqrt{(\sum_{i=1}^n x_i^2 - n \bar{x}^2)(\sum_{i=1}^n y_i^2 - n \bar{y}^2)}}$$

A correlação está limitada em $-1 \leq r \leq 1$

$r = 1$; relação linear perfeita (e positiva) entre x e y

$r = 0$; inexistência de relação linear entre x e y

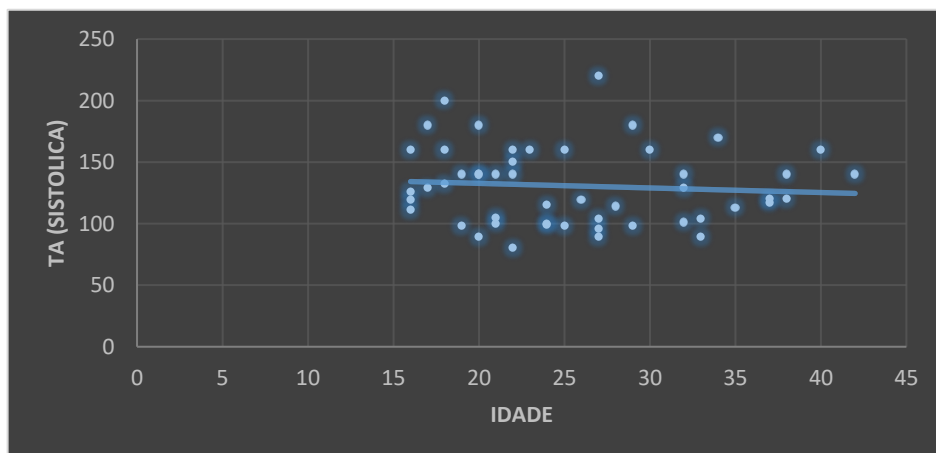
$r = -1$; Relação linear perfeita (e negativa) entre x e y

$r > 0$; Relação linear positiva entre x e y

$r < 0$; Relação linear negativa entre x e y

Gráfico 8

Sistólica x Idade

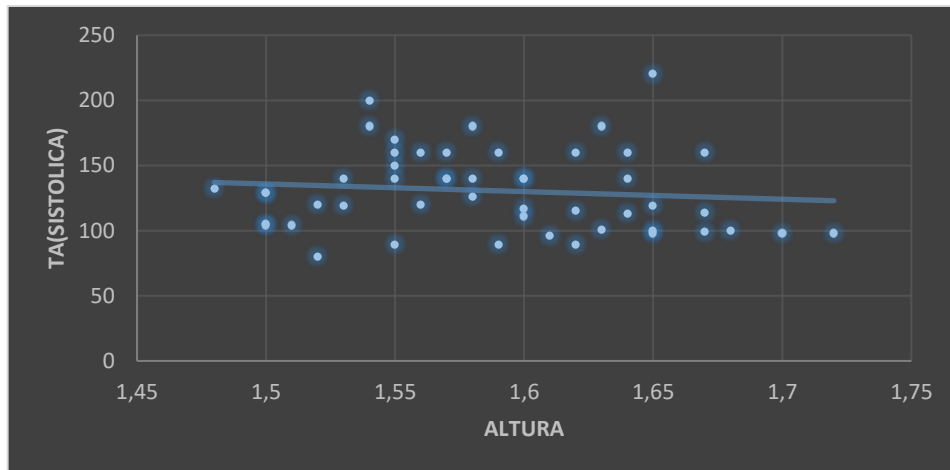


De facto, há uma relação linear descendente da esquerda para a direita, isto revela uma ligação entre dois conjuntos de dados, a medida que a idade aumenta a tensão arterial sistólica baixa, ou seja, a

gestante com maior índice de massa corporal tem tendência de adquirir uma hipotensão, isto é, do lado esquerdo do coração bombeia sangue rico em oxigénio por todo o corpo, o lado esquerdo bombeia sangue a uma pressão maior do que o lado direito, essa pressão chama-se tensão sistólica). Existe uma correlação negativa.

Gráfico 9

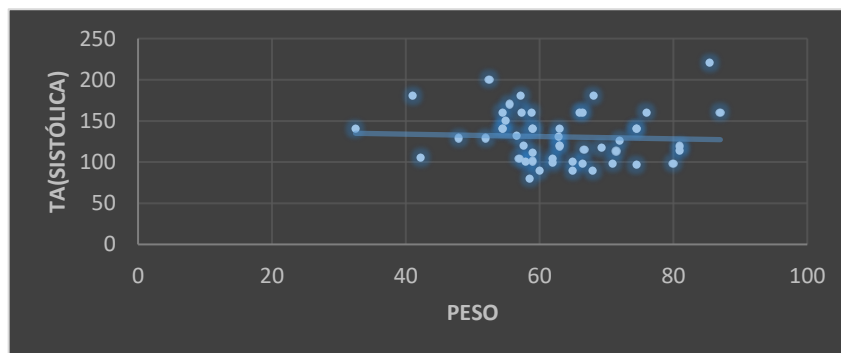
Sistólica x Altura



Há uma relação linear descendente da esquerda para a direita, o gráfico de dispersão mostra nos que as gestantes com maior altura a sua tensão arterial sistólica tende a baixar. Verificamos que existe uma correlação negativa.

Gráfico 10

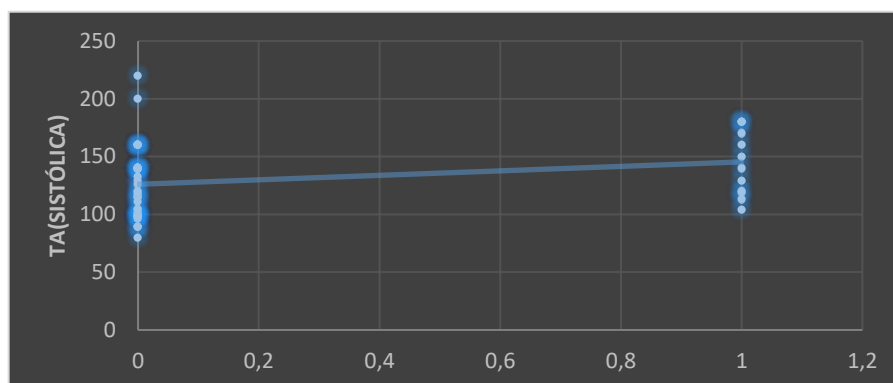
Sistólica x Peso



Há uma relação linear descendente da esquerda para a direita, o gráfico de dispersão mostra nos que as gestantes com maior peso a sua tensão arterial sistólica tende a baixar. Ou seja, quanto maior é o índice de massa corporal menor é a tensão arterial sistólica, as gestantes adquirem hipotensão arterial sistólica. Verificamos que também existe uma correlação negativa.

Gráfico 11

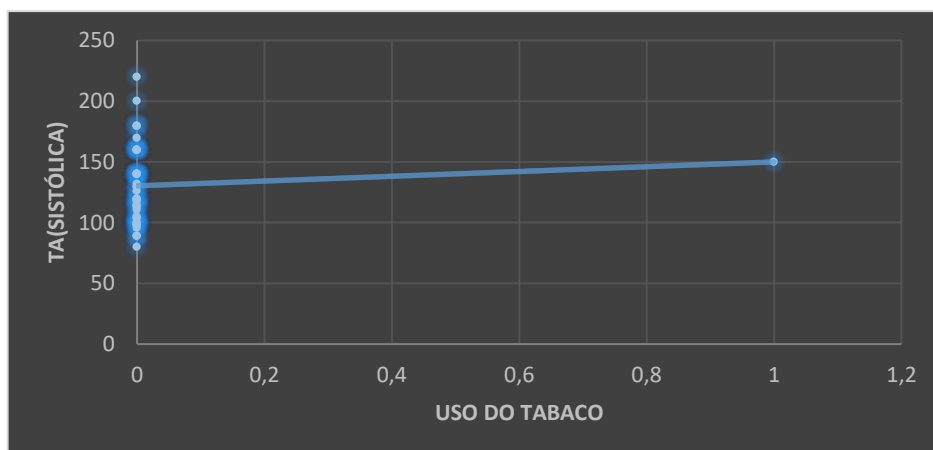
Sistólica x Atividades físicas.



O gráfico mostra-nos que não existe uma relação linear entre a prática de atividades físicas e a tensão arterial sistólica e a sua correlação é negativa.

Gráfico 12

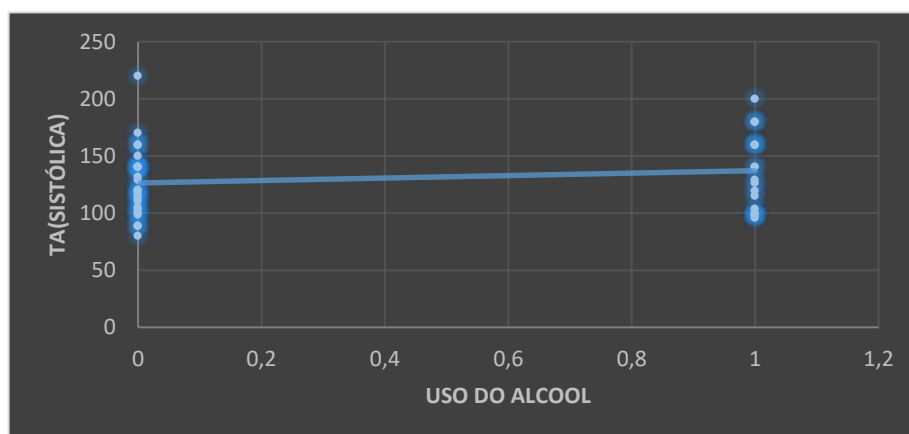
Sistólica x Uso do tabaco.



O Gráfico mostra nos que não existe nenhuma relação linear entre o uso do tabaco e tensão arterial sistólica, portanto o uso do tabaco pouco ou nada influencia para a tensão arterial nas gestantes em estudo. Portanto não existe correlação entre ambos.

Gráfico 13

Sistólica x Uso do Álcool



O gráfico mostra-nos que não existe uma relação linear entre o uso do álcool e tensão arterial sistólica, e sua correlação é negativa. Utilizando a ferramenta do Excel, Análise de Dados, pode-se verificar o R^2 , R^2 ajustado, teste F de significância global e os coeficientes β_0 , β_1 , β_2 , β_3 , β_4 , β_5 , e β_6 da equação de Regressão Múltipla.

O teste F global de significância deve ser menor que 0,07, mas este modelo não tem nenhuma influência significativa na tensão arterial sistólica gestacional. O teste individual para variáveis relacionadas à tensão arterial sistólica mostra que idade, altitude, peso, atividades físicas, uso de tabaco e uso de álcool têm todos um valor inferior a 0,07, indicando que não influenciam a tensão arterial sistólica gestacional. O valor ajustado de R^2 é 10,70%, explicando apenas 10,70% da variação total na tensão arterial sistólica. O valor ajustado de R^2 é -1,75%, tornando o modelo ineficaz para estimar a tensão arterial sistólica gestacional. A equação estimada é:

As variáveis idades, altura, peso, atividades físicas, uso de tabaco e uso de álcool do modelo não influenciam a tensão arterial sistólica em gestantes, e as variáveis do modelo explicam 10,70% da variabilidade em tensão arterial.

As variáveis idades, altura, peso, atividades físicas, uso de tabaco e uso de álcool do modelo não influenciam a tensão arterial sistólica em gestantes, e as variáveis do modelo explicam 10,70% da variabilidade em tensão arterial.

. Sendo assim teremos a seguinte equação estimada:

y_1 = Tensão arterial sistólica

X_1 = Idade, X_2 = Altura, X_3 = Peso, X_4 = Atividades físicas, X_5 = Uso do tabaco e

X_6 = Uso do álcool

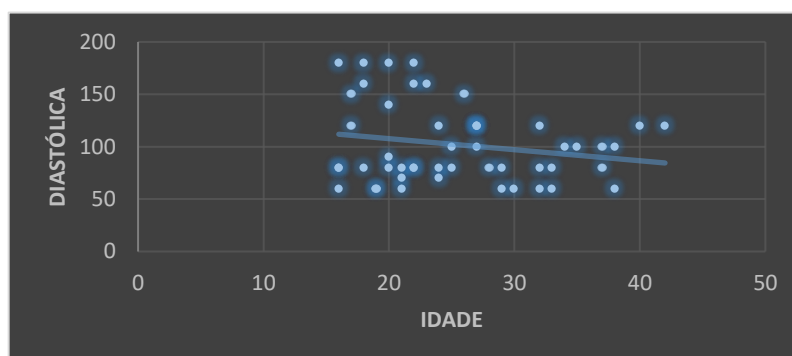
Então, a equação de regressão múltipla desse modelo para tensão arterial sistólica é dada por:

$$\hat{Y}_i = 222,457290 - 0,20430863X_1 - 84,7377912X_2 - 0,55574760X_3 + 12,668768X_4 + 20,146189X_5 + 16,483593X_6$$

A mesma análise que se fez para a sistólica relativamente ao diagrama de dispersão, onde se fez análise individualmente de cada variável, faremo-lo também para a diastólica.

Gráfico 14

Diastólica x idade



O teste F global de significância deve ser menor que 0,07, mas este modelo não tem nenhuma influência significativa na tensão arterial diastólica gestante. O teste individual para variáveis relacionadas à tensão arterial diastólica mostra que idade, peso, atividades físicas, uso de tabaco e uso de álcool têm todos um valor menor que 0,07, o que significa que não influenciam a tensão arterial gestacional. A altura do modelo só influencia a tensão arterial diastólica, pois o valor escolhido é 0,066518. O valor do R^2 ajustado é 17,36%, explicando apenas 17,36% da variação total na tensão arterial diastólica. O valor ajustado de R^2 é 5,83%, indicando que o modelo não é útil para estimar a

tensão arterial diastólica em gestantes. O estudo conclui que as variáveis que afetam a tensão arterial diastólica não são significativas, e apenas a altura afeta a tensão arterial em gestantes.

y_2 = Tensão arterial sistólica

Então, a equação de Regressão Múltipla desse modelo é dada por:

$$\hat{Y}_2 = 427,194 - 193,2X_1 - 0,08082X_2 - 2,89107X_3 + 2,89107X_4 + 46,8813X_5 + 0,841584X_6$$

CONCLUSÕES

O estudo apresentou os pressupostos do modelo de regressão linear e o método dos mínimos quadráticos, o que levou à estimação das variáveis implicadas. Ao utilizar a inferência para os parâmetros, utilizou o nível descritivo ou p-value, abordando, portanto, o modelo linear múltiplo para o estudo da tensão arterial alta nas gestantes do Centro Médico Dom Eliseu Maria Corolli em Luanda / Angola. Validando suposições sobre erros do modelo, o estudo analisou o coeficiente de determinação e a relação linear entre variáveis sistólicas e diastólicas usando gráficos dispersivos para o estudo de correlação.

O estudo encontrou uma relação linear entre idades, altura, atividades físicas e pressão arterial sistêmica, com uma correlação negativa entre elas. No entanto, não houve uma relação linear entre o uso de tabaco e o consumo de álcool entre as gestantes. As variáveis - idade, peso, altura, atividades físicas, uso de tabaco e peso - não influenciaram significativamente a pressão arterial sistêmica em gestantes. Apesar das atividades físicas, a ausência do uso de tabaco ou consumo de álcool resultou em pressão arterial sistêmica alta ou baixa. O estudo descobriu que apenas a altura influencia significativamente a resistência arterial sistólica e diastólica, com apenas 10,70% para a sistole e 17,36% para a diástole sendo explicados pela resistência sistólica e diastólica, e 89,3% para a sistole e 82,64% para a diástole sendo explicados por variáveis latentes ou não elencadas. O estresse, que pode ser causado por vários fatores, como insatisfação no trabalho, altas temperaturas, dor de garganta, histórico familiar, gestação acima de 35 anos e idade gestacional relacionada à gravidez, também é um fator significativo.

AGRADECIMENTOS

Agrademos aos gestores e pessoal do Centro Médico Dom Eliseu Maria Corolli, no Bairro Simeone Mucune, Comuna da Camama, Município de Talatona em Luanda / Angola por ter permitido a colheita de dados. E de igual modo, agradecemos todos os participantes ao inquérito pelo tempo disponibilizado, em especial os estudantes do Curso de Matemática da Faculdade de Ciências da Universidade Agostinho Neto: .Filo Miguel e Helena Alfonso.

BIBLIOGRAFIA

- Almeida, Humberto Moreira (2014) Análise de Regressão linear múltipla com estudos relacionados a horas de máquinas paradas na linha de produção de uma indústria dos calçados, Campinas grande, agosto de 2014.
- Castro, V. D CAR, M.R. (2020) O Cotidiano da vida de hipertensos: Mudanças, restrições e reações. USP Junho DE 2020
- Fernando, Mário SOCIEDADE ANGOLANA DE DOENÇAS CARDIOVASCULARES, disponível em: [http:// WWW.ANGOP.COM](http://WWW.ANGOP.COM), acesso em 2025.03.20.
- Ferreira, M. C. C. dos S. (2013). Modelos de regressão: uma aplicação em medicina dentária [Dissertação de mestrado, Universidade Aberta]. Recuperado de

<https://www.repositorioaberto.uab.pt/bitstream/10400.2/2371/1/TESE%20Cristina%20Ferreira.pdf>

- Firmino, Daniel Alvarez (Sd) Regressão linear simples e Múltipla: Quando usar, como fazer no EXCEL e como analisar.
- Hoffman, Rodolfo (2016) Análise de regressão uma introdução a Econometria. 4ª edição
- Leite, Maria Madalena Januário (2010). Educação Januário. Educação em saúde: desafios para uma prática inovadora/ Maria Madalena Januário Leite, Cláudio Pedro, Heloísa Helena Ciqueto Pires. 1ª Edição. São Caetano do Sul, SP: Difusão Editora. (Serie educação em saúde)
- Lopes, Mislaine Casagrande de Lima & Marcon, Sonia Silva (2009). A hipertensão e a família: a necessidade do cuidado familiar. Ver. Esc. Enferm USP [online.]. Vol. 43.
- Malfatti, Carlos Ricardo Maneck e ASSUNÇÃO (2011). Ari Nunes. Hipertensão arterial e diabetes na estratégia de saúde da família: uma análise da frequência de acompanhamento pela equipe de saúde da família. Ciência saúde colectiva[online]. Vol 16.
- Malta, Deborah Carvalho et al. (2006). a construção de vigilância e prevenção das doenças crônicas não transmissíveis no contexto do sistema único e saúde. Epidemiol, Ser. Saúde [online]. Vol. 15.
- Mancini, Márcio (1998), Prevalência de Hipertensão Arterial, Arquivo Brás. Cardiol, Porto Alegre, volume 70 (nº 2), 111-114. Métodos de avaliação da obesidade e alguns.
- Martins, Maria Eugénia (2005). Introdução a Probabilidade e a estatística com complementos de Excel.
- Pierin, Ângela M.G (2013) Hipertensão arterial: uma proposta para o cuidar. São Paulo. Revista amrigs, (Sd) Conselho Federal de Medicina, porto alegre
- Revista Brasileira De Hipertensão (2010). Rio de janeiro: Sociedade Brasileira de Cardiologia, V.17
- Rodrigues, Lígia Henriques (2015). Regressão linear simples, estatística básica, 8ª edição.
- Rodrigues, Maria Cristina Antunes (2012). Modelo de Regressão e suas aplicações. Covilhã.
- Severino, A. J. (2000). Metodologia científica: teoria da ciência e prática da pesquisa. São Paulo: Atlas.
- Silva, José da. (2012) Metodologia Científica: Fundamentos e Aplicações. 2. ed. São Paulo: Atlas, 2012.
- Toledo, G. L., Fonseca, J. S. da, & Martins, G. D. A. (2001). Estatística aplicada (2ª ed.). São Paulo, SP: Atlas.

Sínteses curricular dos autores.

Makambo Eustache é docente efetivo do Instituto Politécnico de Saurimo, unidade orgânica da Universidade Lueji-A-N'Konde, onde leciona várias cadeiras ligadas às Matemáticas tais como Análise Matemática, Teoria das Funções, Matemática Discreta, Métodos Estatísticos. Foi docente convidado no Departamento de Matemáticas da Faculdade de Ciências Naturais da Universidade Agostinho Neto. Já publicou vários artigos científicos em muitas revistas científicas, com muitas citações. Da orientação de um dos inúmeros trabalhos de fim de curso de licenciatura no Departamento de Matemática da UAN/ FCN, elaborou-se o presente estudo.

ORCID: <https://orcid.org/0000-0003-3709-6117>