



UNIVERSIDADE ESTADUAL DA PARAÍBA
CENTRO de Ciências e Tecnologia
Departamento de Estatística

Joel Pereira Costa

**Estudo de uma ação da saúde no município de Aroeiras – PB
Através da Estatística Descritiva e Inferencial**

Campina Grande
Dezembro de 2013

Joel Pereira Costa

**Estudo de uma ação da saúde no município de Aroeiras – PB
Através da Estatística Descritiva e Inferencial**

Trabalho de Conclusão de Curso apresentado ao curso de Bacharelado em Estatística do Departamento de Estatística do Centro de Ciências e Tecnologia da Universidade Estadual da Paraíba em cumprimento às exigências legais para obtenção do título de bacharel em Estatística.

Orientador:

Prof. Ms. Pedro Cesar Pereira Coelho

Campina Grande – PB

Dezembro de 2013

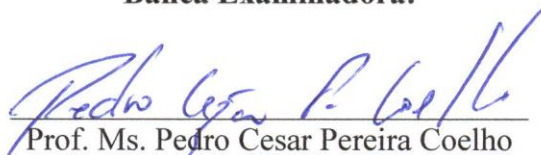
Joel Pereira Costa

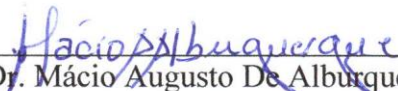
**Estudo de uma ação da saúde no município de Aroeiras – PB
Através da Estatística Descritiva e Inferencial**


Trabalho de Conclusão de Curso apresentado ao curso de Bacharelado em Estatística do Departamento de Estatística do Centro de Ciências e Tecnologia da Universidade Estadual da Paraíba em cumprimento às exigências legais para obtenção do título de bacharel em Estatística.

Aprovado em: 12/12/13

Banca Examinadora:


Prof. Ms. Pedro Cesar Pereira Coelho
Departamento de Matemática e Estatística
- CCT/UEPB
Orientador


Prof. Dr. Mácio Augusto De Albuquerque
-CCT/UEPB
Examinador


Prof. Dr. Vandenberg Lopes Vieira
-CCT/UFCG
Examinador

C837e Costa, Joel Pereira.
Estudo de uma ação da saúde no município de Aroeiras -PB através da estatística descritiva e inferencial [manuscrito] / Joel Pereira Costa. – 2013.

22 f. : il.

Trabalho de Conclusão de Curso (Graduação em Estatística) – Universidade Estadual da Paraíba, Centro de Ciências e Tecnologia, 2013.

“Orientação: Prof. Dr. Pedro Cesar Pereira Coelho, Departamento de Estatística”.

1. Estatística descritiva. 2. Níveis de glicemia. 3. Teste de hipótese. I. Título.

21. ed. CDD 310

Agradecimentos

Ao Professor Orientador, Prof. Ms. Pedro Cesar Pereira Coelho, pela orientação e dedicação e (paciência) a este trabalho e pelo apoio e colaboração durante a orientação da pesquisa.

A minha família Nilda (esposa), filha01 (filha), filha02 (filha), filha03(filha), que sempre me apoiaram e incentivaram e por estarem sempre presentes em todas as fases da minha vida.

Ao departamento de Estatística do CCT/UEPB, pela estrutura e dedicação ao curso de Bacharelado em Estatística.

Aos funcionários do departamento de Estatística, Agnaldo e Mercias por todo serviço prestado durante o curso.

E a todos os amigos que conquistei durante os anos de curso, que me incentivaram e vibraram comigo a cada conquista alcançada durante o curso.

Resumo

Este trabalho teve como objetivo principal aplicar a estatística descritiva, bem como a inferencial através da construção de intervalos de confiança e teste de hipótese para a associação entre determinadas variáveis: Índice de Massa corpórea (IMC) e glicemia, relacionadas à pacientes no Município de Aroeiras – PB. Selecionou-se uma amostragem probabilística; “*Amostragem aleatória*”, 126 prontuários, de um total de 430 pacientes examinados em um evento de saúde no município, onde foram observadas as variáveis: glicemia, idade, peso, sexo e IMC. Alguns dos resultados mostraram que: a média de glicemia dos pacientes foi de 104,29 mg/dL; a grande maioria dos entrevistados foram de pessoas entre 20 a 50 anos, com uma frequência de 83 pacientes, e a proporção de pacientes do sexo feminino foi 78%. Constatou-se usando teste de correlação de pearson que não existe relação entre as variáveis de glicemia e o IMC.

O Município de Aroeiras localiza-se na microrregião Aroeiras e na mesorregião do Agreste Paraibano e possui uma Área territorial de aproximadamente 374,697 Km². Este trabalho teve como principal objetivo aplicar a estatística descritiva e inferencial no estudo de pacientes no Município de Aroeiras – PB.

Palavras chaves- Estatística Descritiva, Níveis de Glicemia, Teste de hipótese

Abstract

This work had as main objective to apply the descriptive statistics inferential as well as through the construction of confidence intervals and hypothesis testing for the association between certain variables: body mass index (BMI) and glucose, related to patients in the municipality of howler monkey – PB. Selected using a probability sampling; "Random sampling", 126 charts, out of a total of 430 patients examined in a health event in the city, where variables were observed: glycemia, age, weight, sex, and BMI. Some of the results showed that: the average blood glucose of patients was of 104.29 mg/dL; the vast majority of respondents were of people aged 20 to 50 years, with a frequency of 83 patients, and the proportion of female patients was 78%. It was found using the pearson correlation test that there is no relationship between the variables of blood glucose and BMI.

The Municipality of Aroeiras located in the micro and meso Aroeiras Paraíba arid and has a land area of about 374.697 km². This study aimed to apply the descriptive and inferential statistics in the study of patients in Município de Aroeiras - PB.

Key Words - Descriptive Statistics, Blood Sugar Levels, Test of Hypothesis

Sumário

1 INTRODUÇÃO.....	pag. 8
2 FUNDAMENTAÇÃO TEÓRICA.....	pag. 9
2.1 ESTATÍSTICA DESCRITIVA.....	pag. 9
2.2 MEDIDAS DE POSIÇÃO.....	pag. 9
2.3 MEDIDAS DE DISPERSÃO.....	pag. 11
3 MATERIAL E MÉTODOS.....	pag. 13
4 CALCULO DO TAMANHO DA AMOSTRA.....	pag. 14
5 RESULTADOS E DISCUSSÕES.....	pag. 16
6 CONCLUSÃO.....	pag. 21
7 REFERENCIAS BIBLIOGRÁFICAS.....	pag. 22

Dedicatória

Dedico este trabalho de conclusão de curso a todos que sempre me incentivaram a concluir meu curso. Em especial minha esposa Nilda, minhas filhas e mestres. Sinto-me muito feliz e é com muito carinho, que faço esta dedicação.

Lista de tabelas

TABELA 1-	Peso dos pacientes.....	17
TABELA 2	Glicemia dos pacientes.....	18

Lista de gráficos

GRÁFICO 1 -	Sexo dos pacientes.....	16
GRÁFICO 2 -	Idade dos pacientes.....	16
GRÁFICO 3 -	IMC dos pacientes.....	17
GRÁFICO 4 -	Gráfico de dispersão com valores discrepantes.....	20
GRÁFICO 5 -	Gráfico de dispersão.....	20

1 Introdução

O estudo de variáveis ligadas à saúde como peso, glicemia, índice de massa corporal (IMC) dentre outras é de grande importância para a qualidade de vida de uma população. Conhecer as características destas variáveis é fundamental para planejamentos adequados para melhoria da população e tentar minimizar os problemas causados por complicações resultantes do descontrole de algumas dessas variáveis. Outra questão a ser considerada é que também é preciso conhecer algumas informações desta população através de variáveis como idade, sexo e etc... que permitem saber algumas características fundamentais a qualquer estudo.

Diante dessas necessidades, recorre-se muitas vezes a estatística descritiva no sentido de buscar informações numéricas que auxiliem no planejamento de diversas áreas do conhecimento científico.

A **glicemia** é a concentração de glicose no sangue ou mais precisamente no plasma. O nosso corpo transforma alguns dos hidratos de carbono ingeridos em glicose e a glicemia é o nível de glicose presente no nosso sangue. Ou seja, quando comemos muito, a glicemia aumenta, ao passo que quando comemos pouco, esta mantém-se baixa.

Mede-se a glicemia através da confirmação dos sinais e sintomas clássicos da glicemia em jejum (exame de sangue onde são verificadas as taxas de glicose no sangue) e do teste padronizado de tolerância à glicose (TTG).

Estes critérios diagnosticados estão baseados nas recomendações da comunidade médico-científica atual:

- **Normal:** Abaixo de 110 mg/dL
- **Intolerância à glicose:** jejum de 111 a 125 mg/dL; 2 horas após 75g de glicose: de 141 a 199 mg/dL
- **hDiabetes melitus:** jejum maior que 126 mg/dL; 2 horas após 75g de glicose: maior que 200 mg/dL

Além da insulina, diabéticos podem controlar a glicemia através de dietas específicas e prática de exercícios físicos, pois, a prática regular de exercício físico aumenta a ação da insulina, fazendo com que a glicose saia da corrente sanguínea, diminuindo, conseqüentemente, a glicemia. Atualmente tornou-se comum a "contagem de carboidratos" dos alimentos através do "índice glicêmico", um indicador de qualidade do carboidrato quanto à sua habilidade em aumentar e/ou influenciar a glicemia.

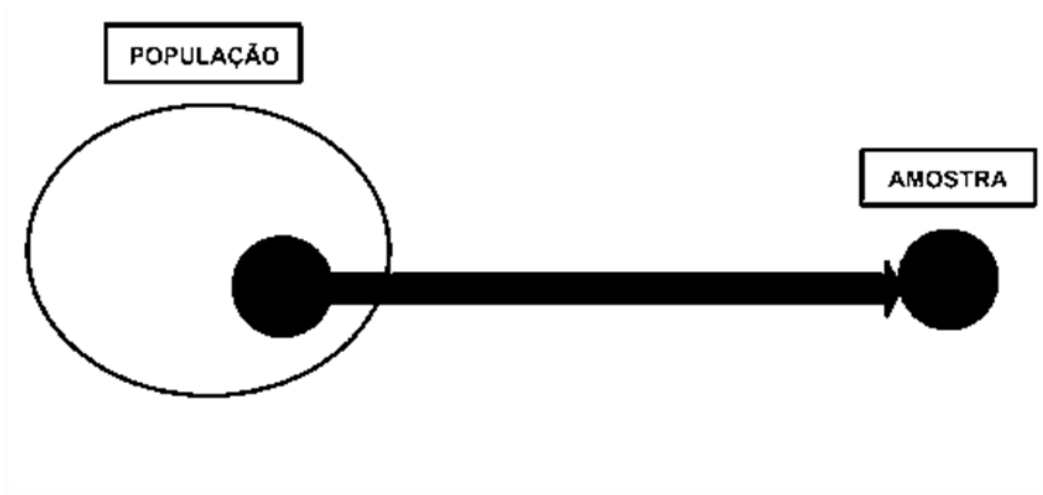
2 Fundamentação Teórica

2.1 Estatística Descritiva

A Estatística é um conjunto de técnicas e métodos de pesquisa que entre outros tópicos envolve o planejamento do experimento a ser realizado, a coleta qualificada dos dados, a inferência, o processamento, a análise e a disseminação das informações.

Na Estatística trabalhamos com dados, os quais podem ser obtidos por meio de uma amostra da população em estudo:

- **População:** Conjunto de elementos que tem pelo menos uma característica em comum. Esta característica deve delimitar corretamente quais são os conceitos da população.
- **Amostra:** subconjunto de elementos de uma população, que são representativos para estudar a característica de interesse da população. A seleção dos elementos que irão compor a amostra pode ser feita de várias maneiras e irá depender do conhecimento que se tem da população e da quantidade de recursos disponíveis.



2.2 Medidas de posição

São as estatísticas que representam uma série de dados orientando-nos quanto à posição da distribuição em relação ao eixo horizontal do gráfico da curva de frequência. As medidas de posições mais importantes são média aritmética, mediana e moda. Usaremos as seguintes notações:

- X : valor de cada indivíduo da amostra.
- \bar{X} : média amostral.
- n : tamanho amostral.

Média populacional

A média populacional é calculada somando-se todos os valores da população e dividindo o resultado pelo total de elementos da população. Numa população de \$ N \$ elementos, a média populacional é dada por:

$$\mu = \frac{X_1 + \dots + X_N}{N}$$

Media amostral

A média amostral, aritmética, ou simplesmente média, é calculada somando-se os valores das observações da amostra e dividindo-se o resultado pelo número de valores. Assim, a média amostral é dada por

$$\bar{X} = \frac{X_1 + \dots + X_n}{n}$$

Mediana

Para calcular a mediana devemos, em primeiro lugar, ordenar os dados do menor para o maior valor. Se o número de observações for ímpar, a mediana será a observação central. Se o número de observações for par, a mediana será a média aritmética das duas observações centrais.

Notação: \tilde{X}

Moda

A moda de um conjunto de valores é o valor que apresenta a maior frequência.

2.3 - Medidas de dispersão

Dispersão é sinônimo de variação ou variabilidade. Para medir a dispersão são usadas mais frequentemente duas medidas: a amplitude e o desvio padrão.

Amplitude

A amplitude é definida como sendo a diferença entre o maior e o menor valor do conjunto de dados. Denotaremos a amplitude por R .

Para definirmos desvio padrão é necessário definir variância. A notação mais comumente usada é:

- S^2 : variância amostral.
- σ^2 : variância populacional.
- s : desvio padrão amostral.
- σ : desvio padrão populacional.

Variância populacional

A variância de uma população $\{x_1, \dots, x_N\}$ de N elementos é a medida de dispersão definida como a média do quadrado dos desvios dos elementos em relação à média populacional μ . Ou seja, a variância populacional é dada por:

$$\sigma^2 = \sum_{i=1}^N \frac{(x_i - \mu)^2}{N}$$

Variância amostral

A variância de uma amostra $\{x_1, \dots, x_n\}$ de n elementos é definida como a soma dos quadrados dos desvios de elementos em relação à sua média \bar{x} dividido por $(n-1)$. Ou seja, a variância amostral é dada por:

$$s^2 = \sum_{i=1}^n \frac{(x_i - \bar{x})^2}{n - 1}$$

Desvio padrão populacional

O desvio padrão populacional de um conjunto de dados é igual à raiz quadrada da variância populacional. Desta forma, o desvio padrão populacional é dado por:

$$\sigma = \sqrt{\sigma^2} = \sqrt{\sum_{i=1}^N \frac{(x_i - \mu)^2}{N}}$$

Desvio padrão amostral

O desvio padrão amostral de um conjunto de dados é igual à raiz quadrada da variância amostral. Desta forma, o desvio padrão amostral é dado por:

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

Coefficiente de Variação

O coeficiente de variação é uma medida relativa de variabilidade. É independente da unidade de medida utilizada, sendo que a unidade dos dados observados pode ser diferente que seu valor não será alterado.

A mesma é usada para analisar a dispersão em termos relativos a seu valor médio quando duas ou mais séries de valores apresentam unidades de medida diferentes

$$CV = \frac{\text{Desvio Padrão}}{\text{Média}} \times 100$$

3 Material e métodos

A pesquisa foi realizada no município de Aroeiras na mesorregião do agreste paraibano com uma população de aproximadamente 19.082 habitantes onde foi utilizada uma amostragem probabilística “*Amostragem aleatória*”, ou seja, foram selecionados 126 prontuários de um total 430 pacientes examinados em uma ação da saúde no município, onde foram observadas as variáveis: glicemia, idade, peso, sexo e IMC. A coleta foi realizada entre os dias 10 e 11 de outubro de 2012, onde se criou o banco de dados com informações de 126 pessoas. Nas análises, utilizou-se de ferramentas da estatística descritiva, teste de hipótese e intervalo de confiança (Bussab e Morettin, 2007) no sentido de estudar e identificar informações importantes no estudo destas variáveis, juntamente com software IBM – SPSS versão 20 que possibilitaram obter as informações necessárias de maneira rápida e prática.

4 Cálculo do Tamanho da Amostra

Para o cálculo do tamanho amostral utilizamos a estimação da Média Populacional de uma Variável Aleatória Quantitativa através da Média Amostral – Cálculo de amostra para população Finita.

O tamanho amostral foi obtido considerando-se:

- Margem de erro de 2,2%;
- Confiabilidade de 95,0%;
- Tamanho populacional igual ao número de 430 prontuários“ *Não estamos trabalhando com toda a população devido ao banco de dados irregular que se dispõe para o estudo*”.

A forma de cálculo utilizada para o tamanho da amostra para cada item foi:

$$n_0 = \frac{z^2 \cdot \sigma^2}{\varepsilon^2}$$

Onde:

n = Tamanho da amostra

z = Valor para a probabilidade escolhida

σ = Desvio padrão da população estimado previamente

ε = Erro

Para calcular a amostra(n) foram utilizados os seguintes dados:

$Z = 1,96$, considerando uma probabilidade de 95%

$\sigma = 15\%$ - Considerando que esse desvio vem da população em estudo.

$$E = 2,2\%$$

Assim, o tamanho da amostra válida seria de 178 prontuários médicos, porém, após verificar o fator de correção para amostras finitas.

Dada pela notação $\frac{n_0}{N} \leq 0,022$ observamos;

Verificação:

$$\frac{178}{430} = 0,41, \text{ e desta forma sendo } 0,41 > 0,022 \text{ devemos usar o correção } n = \frac{n_0}{1 + \frac{n_0}{N}}$$

$$\text{Substituindo, pelos nossos valores, chegamos: } n = \frac{178}{1 + \frac{178}{430}} \cong 126$$

Ou seja, o tamanho mínimo da amostra necessária para validar a pesquisa, com erro máximo de amostragem igual a 2,2% e um nível de confiança de 95% é de aproximadamente **126 prontuários**. Com isso, conclui-se que o tamanho da amostra é suficiente.

5 Resultados e discussão

Com o propósito de saber a proporção referente ao sexo dos pacientes, construiu-se a **Gráfico 1** abaixo:

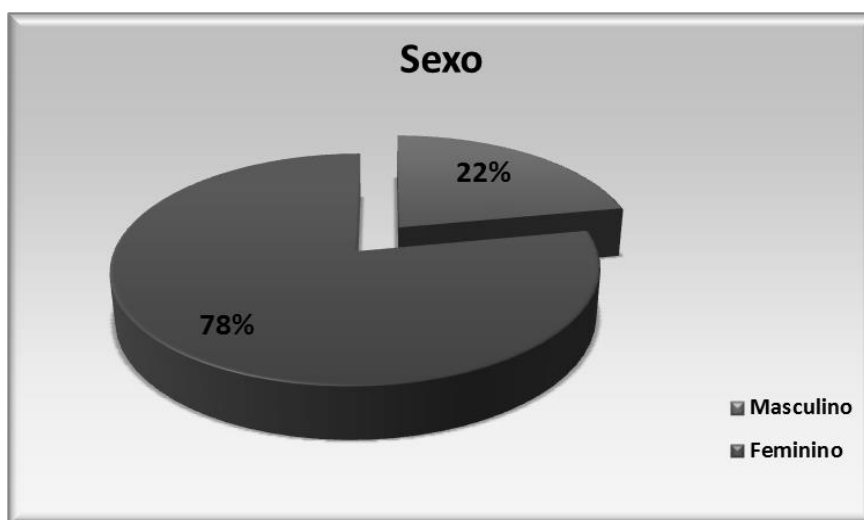


Gráfico 1: Proporção do sexo dos pacientes no Município de Aroeiras– PB

Por meio da figura acima, percebe-se a alta proporção de pacientes do sexo feminino (78%), enquanto que masculino foi de 22%.

A **Gráfico 2** abaixo mostra a distribuição de frequência da idade dos pacientes, distribuídos em 9 classes.

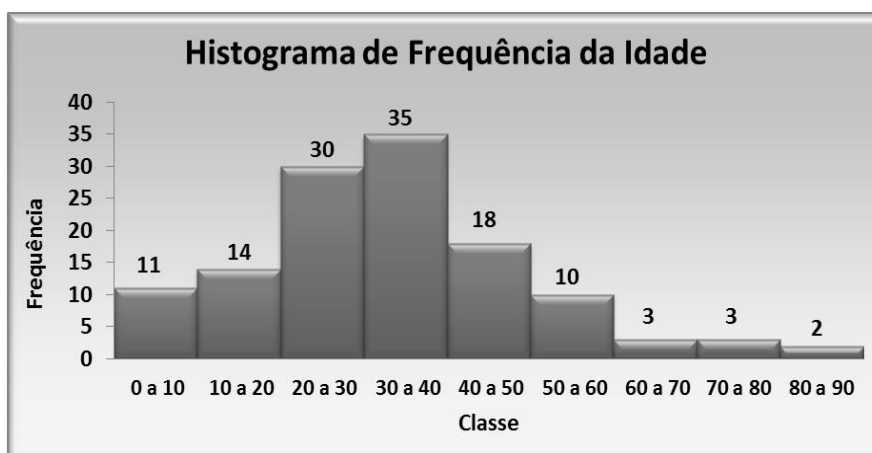


Gráfico 2: Histograma de frequência da idade dos pacientes no Município de Aroeiras– PB

Através desta figura, nota-se que a grande maioria dos entrevistados foi de pessoas entre 20 e 50 anos, com uma frequência de 83 pacientes. Já a classe que apresentou maior frequência dentre todas foi à de 30 a 40 anos com 35 pacientes, enquanto que a classe de menor frequência foi à de 80 a 90 anos, representada por apenas 2 pacientes.

Uma das preocupações nos dias de hoje é o peso. Então objetivando identificar o comportamento da variável peso destes pacientes, foi construída a **Tabela 1** abaixo contendo um resumo estatístico desta variável.

Tabela 1: Resumo estatístico da variável peso dos pacientes no Município de Aroeiras– PB

Resumo Estatístico	Peso (Kg)
Média	61,84
Mediana	65,15
Desvio-Padrão	19,04
Máximo	97,1
Mínimo	39
Amplitude total	88,1
CV	30,78%
IC	61,84±3,38
Tamanho da amostra	126

Pode-se perceber que diante uma amostra composta por 126 pacientes, a média do peso foi de 61,84 Kg, próximo da mediana (65,15 Kg). O valor máximo foi de 97,1 Kg e o mínimo de 9 Kg, sendo este mínimo atribuído à presença de crianças nesta pesquisa.

Uma outra característica que nos dá informações importantes para orientar uma pessoa vir a ter uma melhor qualidade de vida é o IMC (Índice de Massa Corpórea) o qual começamos analisar através de um histograma.

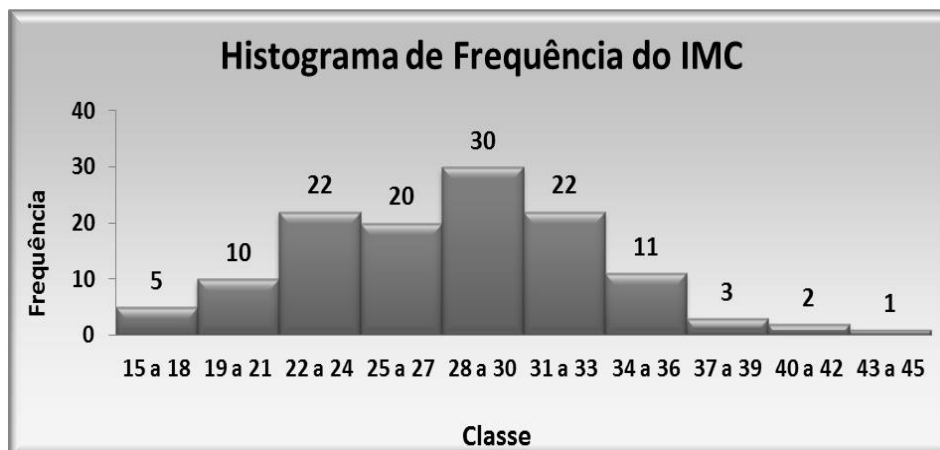


Gráfico 3: Histograma de frequência do IMC dos pacientes no Município de Aroeiras– PB

Segundo (Wikipédia, 2012) o IMC é uma medida internacional usada para calcular se uma pessoa está ou não no peso considerado ideal. Diante deste histograma contendo 10 classes, percebe-se que a maior parte dos pacientes apresentou índice entre 21 e 33, com frequência de 94 pacientes. Já a maior classe observada foi de 27 a 30, com frequência de 30 pacientes, ou seja, pacientes que já podem ser considerados com sobrepeso.

A **Tabela 2** a seguir mostra o resumo estatístico da variável glicemia dos pacientes no Município de Aroeiras.

Tabela 2: Resumo estatístico da variável glicemia dos pacientes de Aroeiras– PB

Resumo Estatístico	Glicemia (mg/dL)
Média	104,29
Mediana	96
Desvio-Padrão	46,86
Máximo	452
Mínimo	64
Amplitude Total	388
CV	44,93%
IC	104,29 ±8,2
Tamanho da Amostra	126

Através desta tabela, pode-se perceber que a média de glicemia dos pacientes foi de 104,29 mg/dL com uma variabilidade de 46,86 mg/dL indicada pelo desvio-padrão. Com relação aos valores extremos desta amostra, o máximo valor apresentado foi de 452 mg/dL e o mínimo de 64 mg/dL.

Portanto, conclui-se que o uso da estatística descritiva foi de fundamental importância no estudo de pacientes no município de Aroeiras.

Intervalo de Confiança:

Com a intenção de verificar um parâmetro populacional verdadeiro para a glicemia foi obtido intervalos de confiança de 95% para a média de glicemia. Para a obtenção do intervalo se utilizou a seguinte notação e dados:

$$\text{Notação} - \left(\bar{x} - 1.96 \times \frac{s}{\sqrt{n}} , \bar{x} + 1.96 \times \frac{s}{\sqrt{n}} \right).$$

$$\text{Média}(\bar{X}) = 104,29$$

$$\text{Desvio Padrão (s)} = 46,86$$

$$Z = 95\% \Rightarrow 1,96$$

$$n = 126$$

Construção do intervalo

$$\left[104,29 - 1,96 \times \frac{46,86}{\sqrt{126}} ; 104,29 + 1,96 \times \frac{46,86}{\sqrt{126}} \right] \text{ Assim intervalo encontrado foi: } [96,09; 112,49]$$

Diante das informações podemos afirmar que o nível glicêmico médio da população estudada está contido, com 95% de confiança, no intervalo [96;112]mg/dL. Evidenciando possíveis riscos de saúde na população estudada.

Teste de Hipótese:

Segundo dados da OMS (Organização Mundial de Saúde) a glicemia está em muitos casos associada ao IMC (índice massa corpórea), ou seja, quanto maior o valor dessa variável aumenta a possibilidade de uma glicemia alterada.

Uma hipótese estatística é uma afirmativa a respeito de um parâmetro de uma distribuição de probabilidade.

H_0 : É chamada de hipótese nula e H_1 chamada hipótese alternativa.

Vias de regra, a hipótese nula é feita com base no comportamento passado enquanto a alternativa é formulada em função de alterações.

No ambiente atual de melhoria continua é fácil entender a importância dos testes de hipótese: eles permitem confirmar a eficácia ou não das medidas de melhoria adotadas.

Relação entre o IMC e a Glicemia:

H_0 : Não existe associação do IMC nos valores da Glicemia

H_1 : Existe associação do IMC nos valores da Glicemia

Gráfico 4: Gráfico de dispersão

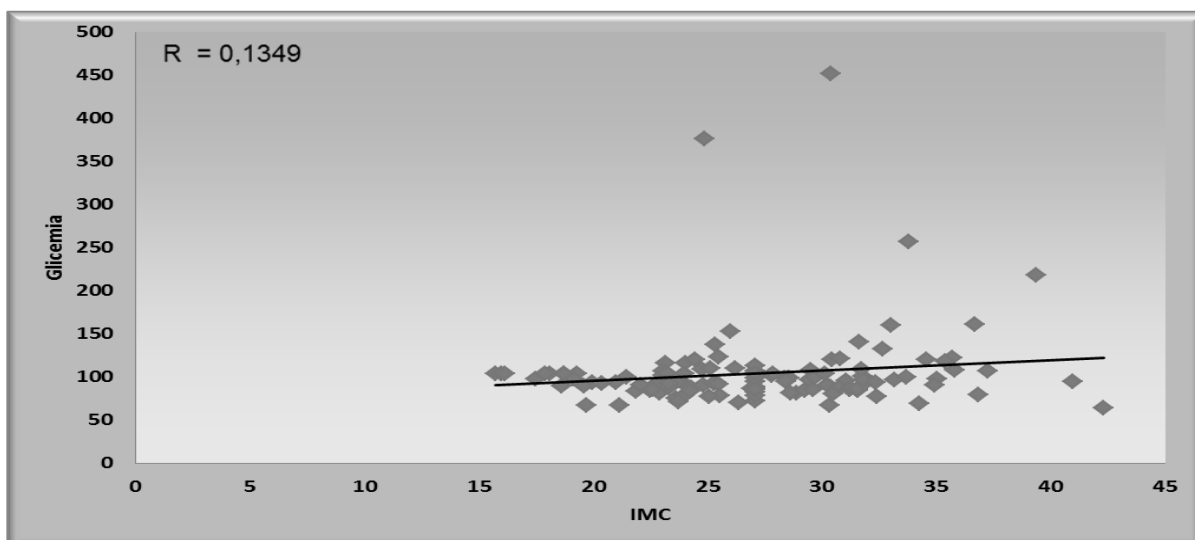


Gráfico 4: Gráfico de dispersão entre o IMC e a Glicemia com valores discrepantes

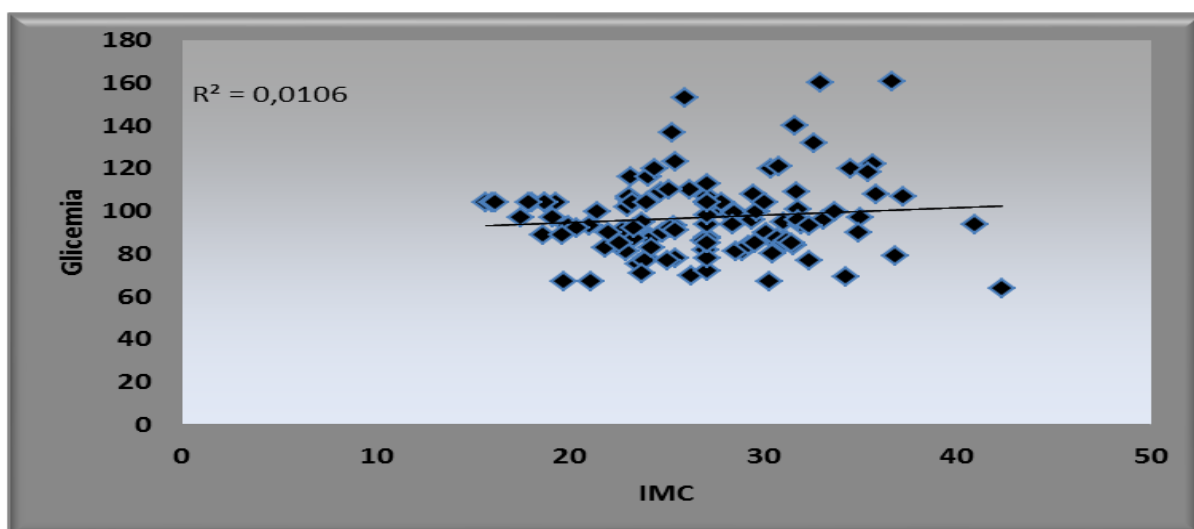


Gráfico 5: Gráfico de dispersão entre o IMC e a Glicemia sem os valores discrepantes

Analisando os dados verificamos que não existe uma associação entre os valores do IMC em relação aos valores da Glicemia, ou seja, encontramos nos dados evidências para rejeitarmos H_0 : Não existe influência do IMC nos valores da Glicemia.

6 Conclusão

- A média de glicemia dos pacientes foi de 104,29 mg/dL com uma variabilidade de 46,86 mg/dL indicada pelo desvio-padrão;
- A maioria dos entrevistados foi de pessoas de 20 a 50 anos, com uma frequência de 83 pacientes;
- A média do peso foi de 61,84 Kg, próximo da mediana (65,15 Kg);
- A alta proporção de pacientes do sexo feminino (78%), enquanto que masculino foi de 22%;
- A maior classe no histograma para o IMC foi de 27 a 30, composto de 30 pacientes, ou seja, considerados com sobrepeso;
- O teste de hipótese através para correlação de Pearson revelou que não existe associação do IMC nos níveis glicêmicos dos pacientes em questão..
- E o intervalo de confiança mostrou valores relativamente altos de glicemia que podem afetar diretamente a saúde da população de estudo.

7 Referências

BUSSAB, W. de O.; MORETTIN, P. A. **Estatística Básica**. 5. ed. São Paulo: Editora Saraiva, 2007. 526 p;

Índice de massa corporal. In: Wikipédia: a enciclopédia livre. Disponível em <http://pt.wikipedia.org/wiki/%C3%8Dndice_de_massa_corporal>. Acesso em: 26 nov. 2010.

Interpretação do P valor - <http://www.portalaction.com.br/539-512-c%C3%A1culo-e-interpreta%C3%A7%C3%A3o-do-p-valor>

Magalhães, M. N. & LIMA, A. C. P. (2009) Noções de Probabilidade e Estatística. 7a ed. São Paulo: Edusp.

COCHRAN, Willian G – Técnicas de Amostragem – Editora Fundo de cultura, Rio de Janeiro, 1965.

Índice de Glicemia <http://www.diabetes.org.br/indice-glicemico/212-o-que-e-indice-glicemico>