Universidade Estadual da Paraíba
Centro de Ciências e Tecnologia
Departamento de Estatística

Maria do Socorro Felix da Silva

Construção da distribuição de qui-quadrado através da soma dos quadrados de normais padrão

Campina Grande

Junho de 2014

Maria do Socorro Felix da Silva

Construção da distribuição de qui-quadrado através da soma dos quadrados de normais padrão

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:
Gustavo Henrique Esteves

Campina Grande
Junho de 2014

S586c Silva, Maria do Socorro Felix da.
Construção da distribuição de qui-quadrado através da soma dos quadrados de normais padrão [manuscrito] / Maria do Socorro Felix da Silva. - 2014.
36 p. : il. color.

Digitado.
Trabalho de Conclusão de Curso (Graduação em Estatística) - Universidade Estadual da Paraíba, Centro de Ciências e Tecnologia, 2014.
"Orientação: Prof. Dr. Gustavo Henrique Esteves, Departamento de Estatística".

1. Probabilidade. 2. Distribuições. 3. Simulação. I. Título.
21. ed. CDD 519.5

Maria do Socorro Felix da Silva

Construção da distribuição de qui-quadrado através da soma dos quadrados de normais padrão

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: / /

Banca Examinadora:

Prof. Gustavo Henrique Esteves
Universidade Estadual da Paraíba - DE/CCT
Orientador

Prof. João Gil de Luna
Universidade Estadual da Paraíba - DE/CCT
Examinador

Prof. Divanilda Maia Esteves
Universidade Estadual da Paraíba - DE/CCT
Examinador

Dedicatória

Dedico esta monografia a minha família, pelo incentivo e colaboração, principalmente nos momentos de dificuldade, pela fé e confiança demonstrada.

Aos meus amigos pelo apoio incondicional

Aos professores da UEPB pelo simples fato de estarem dispostos a ensinar.

Ao orientador pela paciência, pela força e principalmente pelo carinho demonstrado no decorrer do trabalho.

Enfim a todos que de alguma forma tornaram este caminho mais simples de ser cursado. Valeu a pena toda distância, todo sofrimento, todas as renúncias.

Agradecimentos

Agradeço a Deus pela chance de estar concretizando este trabalho, o que seria de mim sem a fé que eu tenho nele.

Aos meus pais, irmãos e a toda minha família que, com muito carinho e apoio, não mediram esforços para que eu chegasse ate esta etapa de minha vida.

A todos os professores do curso, que foram tão importantes na minha vida acadêmica e no desenvolvimento desta monografia.

A meu orientador por estar sempre disposto a ajudar.

Aos meus colegas em especial Juliana, Eveline e Moniclaudia pelas palavras amigas nas horas difíceis, pelo auxilio nos trabalhos e dificuldades e principalmente por estarem comigo nesta caminhada tornando-a mais fácil e agradável.

Resumo

A medida da incerteza é tratada mediante técnicas e métodos que se fundamentam na teoria de probabilidade. Essa teoria procura quantificar a incerteza existente em determinada situação. A estatística se preocupa em organizar, descrever, analisar e interpretar as distribuições a partir de uma observação ou de um experimento. Conhecer algumas das distribuições é de extrema importância, pois as mesmas contribuem para toda teoria estatística. Podemos utilizar recursos computacionais, em particular simulações para inferir distribuições amostrais, na teoria de estatística existem vários resultados que podem ser ilustrados via simulação, que nos permite uma melhor visualização e compreensão dos resultados. Um exemplo disso é a construção da distribuição qui-quadrado através da soma dos quadrados da distribuição normal padrão, e para ilustrar este resultado usamos simulações, utilizando o programa *R Core Team* na versão 3.0.2, por meio do *interface* RStudio na versão 0.97.551. No entanto, para se ter uma melhor compreensão dos resultados foi feito o histograma acrescentado a curva da distribuição e o seu respectivo *qq-plot*.

Palavra-chave: Probabilidade, Distribuições, Simulação.

Abstract

The uncertainty measure is treated by techniques and methods that are based on probability theory. This theory seeks to quantify the uncertainty in a given situation. Statistics mainlt tocus on organization, visualization, analyzes and interpretation of distributions from an observed or experimental data. So, the understanding about some known distributions is of great importance because they contribute to the whole statistical theory. We can use computational methods, in particular data simulation to infer about sampling distributions. In statistical theory there are several results that can be illustrated via simulation, which allows us to better visualize and understand the results. As an example, we have the chi-square distribution definition as the sum of the squared standard normal distributions, and we use simulations to illustrate this result, using the R Core Team program version 3.0.2 and RStudio interface version 0.97.551. However, to get a better view of the results we also did histograms added with the theoretical density and their respective qq -plot.

Keywords: Probability, Distribution, Simulation.

Sumário

Lista de Figuras

1 Introdução

Todos os dias somos confrontados com situações, que nos conduzem a utilizar, intuitivamente, a noção de probabilidade. Nos mais variados aspecto da nossa vida está presente a incerteza. Uma forma de defini-la seria dizer que a probabilidade quantifica a margem de sucesso ou fracasso de um acontecimento. Ou seja, a probabilidade é responsável pelos estudos do comportamento dos fenômenos aleatórios.

Uma variável aleatória é uma função do espaço amostral Ω em $\mathbb{R}$ para cada resultado de um experimento, podendo ser discretas assumindo valores finitos ou enumeráveis, ou contínua assumindo inúmeros valores num intervalo de números reais e é medida numa escala contínua. Quando conhecemos todos os valores de uma variável aleatória juntamente com suas respectivas probabilidades, temos uma distribuição de probabilidades.

A distribuição dos dados pode ser representada por um histograma de probabilidades que nos permite visualizar a forma da distribuição. A esperança, a variância e o desvio padrão traduzem outras características.

Uma distribuição de probabilidades é um modelo matemático que relaciona certo valor da variável em estudo com a sua probabilidade de ocorrência. Há dois tipos de distribuições de probabilidades que são as distribuições discretas que descrevem quantidades aleatórias e podem assumir valores e os valores são finitos, e as distribuições contínuas que representam quantidades aleatórias contínuas e podem tomar qualquer valor dentro de um intervalo especificado dos números reais.

Existem várias distribuições de probabilidades discretas e contínuas. Dentre estas citamos as distribuições discretas binomial e poisson e as distribuições contínuas uniforme, normal e qui-quadrado.

A teoria de estatística ressalta que vários resultados podem ser ilustrados via simulação, das quais se podem utilizar recursos computacionais e em particular simulações para inferir distribuições amostrais. Sendo assim, a simulação é uma técnica estatística que utiliza programas computacionais para simular processos reais de interesse.

O presente trabalho tem por objetivo construir a distribuição qui-quadrado através da soma dos quadrados da distribuição normal padrão, desde a demonstração do resultado teórico até a ilustração deste resultado usando simulação. Neste caso, ilustrar o resultado usando simulações para mostrar a relação entre a distribuição normal padrão e a qui-quadrado com três simulações alternando seus graus de liberdade em 1, 5 e 15 respectivamente da distribuição qui-quadrado. No entanto, para se ter uma melhor compreensão e visualização dos resultados faremos o histograma acrescentando a curva da distribuição e seu respectivo *qq-plot*.

2 Fundamentação Teórica

2.1 Marco Histórico

A probabilidade surgiu do interesse do homem em estudar os fenômenos que envolviam determinadas possibilidades. Alguns indícios alegam que o surgimento da teoria de probabilidades teve início com os jogos de azar difundidos na Idade Média.

O desenvolvimento das teorias de probabilidades e os avanços dos cálculos probabilísticos devem ser atribuídos a vários matemáticos. Por exemplo, os algebristas italianos Pacioli, Cardano e Tartaglia (séc. XVI), podem ser responsáveis pelas primeiras considerações matemáticas acerca dos jogos e das apostas. Com o aprofundamento dos estudos, outros matemáticos contribuíram para a sintetização de uma ferramenta muito utilizada cotidianamente. Dentre os mais importantes, podemos citar: Blaise Pascal (1623-1662), Pierre de Fermat (1601-1655), Jacob Bernoulli (1654-1705), Pierre Simon Laplace (1749-1827), Carl Friedrich Gauss(1777-1855) e Siméon Denis Poisson (1781-1840) (NOE, 2009).

Atualmente, os estudos relacionados às probabilidades são utilizados em diversas situações, pois possuem axiomas, teoremas e definições bem contundentes. Seus estudos dizem respeito à equidade dos jogos e dos respectivos prêmios, tendo sua principal aplicação a Estatística Indutiva, na acepção de amostra, extensão dos resultados à população e na previsão de acontecimentos futuros (NOE, 2009).

A probabilidade é um conceito filosófico e matemático que permite a quantificação da incerteza, admitindo que ela seja aferida, analisada e usada para realizar previsões ou para a orientação de intervenções. É aquilo que torna possível lidar de forma racional com problemas envolvendo o imprevisível.

A teoria de probabilidades admite que se calcule a chance da ocorrência de algum evento de interesse em um experimento aleatório.

Os experimentos aleatórios são aqueles que repetidos em iguais condições, podem fornecer resultados diferentes, ou seja, são resultados explicados ao acaso. Se lançarmos uma moeda ao chão para observamos a face que ficou para cima, o resultado é imprevisível, pois tanto pode sair cara como coroa. Neste caso o seu espaço amostral será representado por: $\Omega = \{\text{cara, coroa}\}$.

O espaço amostral é o conjunto de todos os resultados possíveis de um experimento aleatório, representado pela letra grega maiúscula Ω (ômega). A cada evento S que é definido por subconjuntos do espaço amostral, associa-se um número real representado por $P(S)$ e denominado probabilidade de S.

2.1.1 Axiomas de Kolmogorov

A Probabilidade é uma função que pode receber um evento qualquer e retornar um número real entre 0 e 1. Então, quanto mais próximo de 0 for o valor, mais difícil é para o evento acontecer. E quanto mais próximo de 1, mais provável é a ocorrência de um evento. A teoria de probabilidades possui os seguintes axiomas.

- A Probabilidade de um evento é sempre um número entre 0 e 1. Ou seja:

  $0 \leq P(S) \leq 1$

- A Probabilidade de ocorrer algo dentro do Espaço Amostral é 1 que é entendido como a ocorrência do evento certo. Ou seja:

  $P(\Omega) = 1$

- Se A e B forem eventos mutualmente excludentes, $P(A \cup B) = P(A) + P(B)$.

  Observe-se que, para qualquer n finito, se $A_1, A_2, ..., A_n$ forem mutualmente excludentes,

$$P\left(\bigcup_{i=1}^{n} A_i\right) = \sum_{i=1}^{n} P(A_i).$$

2.1.2 Propriedades de uma medida de probabilidade

A teoria de probabilidades consiste em utilizar a intuição humana para estudar os fenômenos do nosso cotidiano de trabalho. Para isso, utiliza-se o princípio básico do aprendizado humano que é a ideia de experimentos. Os teoremas de base de probabilidades podem ser demonstrados a partir dos axiomas de probabilidades e da teoria de conjuntos. A seguir serão apresentadas algumas propriedades acerca de probabilidades.

- A probabilidade de ocorrer um conjunto vazio pertencente ao conjunto de pontos amostrais é sempre zero. Este é um evento impossível.

- A probabilidade de qualquer subconjunto pertencente ao espaço amostral pode ser calculada através da soma da probabilidade de seus elementos. Ou seja, uma probabilidade, $P(\omega)$, para cada ponto amostral, de tal sorte que seja possível encontrar a probabilidade $P(A)$ de qualquer subconjunto A de Ω. De modo geral, se A for qualquer evento de Ω para espaços amostrais finitos, então

$$P(A) = \sum_j P(\omega_j)$$

onde a soma é estendida a todos os pontos amostrais $\omega_j \in A$.

- A probabilidade que ocorra um evento é igual a 1 menos a probabilidade de ocorrer o evento complementar. Ou seja: $P(A) = 1 - P(\bar{A})$. Isto indica, por exemplo, que a probabilidade de ocorrer o número 3 ou 6 no lançamento de um dado é igual a $1 - (P(1) + P(2) + P(4) + P(5))$.

- A probabilidade de ocorrer a união de dois subconjuntos é a mesma da soma das probabilidades de ambos menos a probabilidade dos dois ocorrerem juntos. Ou seja: $P(A \cup B) = P(A) + P(B) - P(A \cap B)$. Por exemplo, no lançamento de dois dados a probabilidade que ocorram os resultados 2 ou 4 é a mesma de que ocorra o resultado 2 em um dos dados mais a probabilidade de ocorrer o resultado 4 em um dos dados menos a probabilidade de que ocorram os dois juntos.

2.2 Variáveis Aleatórias

Muitos experimentos aleatórios determinam resultados não numéricos. Antes de analisá-los, é apropriado transformar seus resultados em números, o que é feito através da variável aleatória, que é uma função matemática que associa um valor numérico a cada ponto do espaço amostral. No entanto, as variáveis aleatórias são variáveis numéricas as quais iremos associar modelos probabilísticos. Observaremos que uma variável aleatória tem um número para cada resultado de um experimento e que uma distribuição de probabilidades associa uma probabilidade a cada resultado numérico de uma variável aleatória (CORREA, 2003).

Uma variável aleatória (v.a.) é uma função que atribui um valor numérico a cada resultado individual de uma experiência aleatória. Ou seja: $X : \Omega \to \mathbb{R}$

Então, se X associa um resultado a um número, temos que X é uma função cujo domínio é o conjunto de resultados e cuja imagem é o conjunto dos números reais.

2.2.1 Variáveis aleatórias discretas e contínuas

De acordo com Portnoi (2005), as variáveis aleatórias (v.a.) podem ser classificadas em discretas e contínuas:

Diz-se que uma variável aleatória é discreta se todos os seus valores podem ser postos em listas, e estes valores pertencem a um conjunto finito ou infinito, que seja enumerável. Por exemplo, temos: número de chegadas a uma fila, número de caras em uma jogada de duas moedas, resultado da jogada de um dado.

Uma variável aleatória é contínua se os seus valores não podem ser postos em listas, o que seria um conjunto não enumerável, mas pode assumir um número infinito de valores em um intervalo limitado ou não. Por exemplo, temos intervalo de tempo entre chegadas, altura e o peso de pessoas em uma sala.

Seja X uma variável aleatória discreta assumindo valores em um conjunto. A função de distribuição de probabilidades de X é uma $f_X(x)$ que associa, a cada valor possível x de X a sua probabilidade.

$f_X(x) = P(X = x)$.

Esta função apresenta as seguintes propriedades:

$0 \leq f_X(x) \leq 1 \qquad \forall x \in \mathbb{R}$,

$\sum_{x_i} f_X(x_i) = 1$.

Uma variável aleatória X diz-se continua se existe uma função $f_X(x)$ não negativa ($f_X(x) \geq 0, \qquad \forall x \in \mathbb{R}$) e integrável com

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

tal que: $P(a < X \leq b) = \int_a^b f_X(x)dx, \qquad \forall a, b \in \mathbb{R} : a \leq b$.

A função f_X é a função densidade de probabilidade (f.d.p.) da variável aleatória X, e sua área total delimitada pela curva da função densidade e pelo eixo das abscissas é igual a 1. Para uma variável aleatória contínua a probabilidade $X = x_0$ é sempre igual a 0, isto é, $P(X = x_0) = 0$ para qualquer que seja o ponto x_0.

2.2.2 Esperança, variância e desvio padrão

A esperança de uma variável aleatória é, por definição, a média de sua distribuição de probabilidades. Ou seja, dada uma variável aleatória X sobre um espaço amostral, a esperança de X é o valor médio assumido por X. Esperança de uma variável aleatória X é definida por

$$\mu_x = E(X) = \begin{cases} \sum x_i p(x_i) & \text{se } X \text{ é discreta,} \\ \\ \int_{-\infty}^{\infty} x f(x) dx & \text{se } X \text{ é contínua com densidade } f. \end{cases}$$

Sejam X e Y variáveis aleatórias e a, b e c constantes reais. Então, as principais propriedades da esperança matemática de uma variável aleatória são:

- $E(c) = c$
- $E(cX) = cE(X)$;
- $E(aX + bY) = aE(X) + bE(Y)$;
- Se $X(\omega) \leq Y(\omega) \quad \forall \omega \in \Omega$, então $\quad E(X) \leq E(Y)$

Variância é a medida que dá o grau de dispersão de probabilidade em torno da média. A variância de uma v.a. X é definida por:

$$V(X) = \sigma_x^2 = E[X - E(X)]^2 = E[X - \mu_x]^2$$

- Para $Xv.a.d. : V(X) = \sum_x (x - \mu_x)^2 p(x)$
- Para $Xv.a.c. : V(X) = \int_{-\infty}^{\infty} (x - \mu_x)^2 f(x) dx$

Uma fórmula mais simples para calcular a variância é:

$$V(X) = E(X^2) - [E(X)]^2$$

Em que:

- Para $Xv.a.d. : E(X^2) = \sum_i x_i^2 p(x_i)$
- Para $Xv.a.c. : E(X^2) = \int_{-\infty}^{\infty} x^2 f(x) dx$

Seja X uma variável aleatória e a, b e c constantes reais. Então, as principais propriedades da variância de uma variável aleatória são:

- $Var(X) = E(X^2) - [E(X)]^2$
- $Var(c) = 0$
- $Var(aX + b) = a^2 Var(X)$

O Desvio Padrão $DP(X)$ é a raiz quadrada da variância.

$DP(X) = \sqrt{Var(X)}$

2.3 Distribuições de Probabilidades

Quando aplicamos a estatística na resolução de problemas administrativos, verificamos que muitos problemas apresentam as mesmas características o que nos permite estabelecer um modelo teórico para determinação da solução de problemas. Os componentes principais de um modelo estatístico teórico são:

- Os possíveis valores que a variável X pode assumir;
- A função de probabilidade associada a variável aleatória X;
- O valor esperado da variável aleatória X;
- A variância e o desvio padrão da variável aleatória X.

Existem dois tipos de modelos de distribuições teóricos que correspondem a diferentes tipos de dados ou variáveis aleatórias: a distribuição discreta e a distribuição contínua.

2.3.1 Distribuições de probabilidades discretas

A distribuição discreta descreve quantidades aleatórias, ou seja, dados de interesse que podem assumir valores particulares e os valores são finitos. Por exemplo, uma variável aleátoria discreta pode assumir somente os valores 0 e 1, ou qualquer outro inteiro não negativo, etc. Existem várias distribuições discretas, algumas delas serão mostradas abaixo.

Distribuição Binomial

Uma variável aleatória tem distribuição binomial quando o experimento ao qual está relacionada apresenta dois resultados (sucesso ou fracasso). Este modelo fundamenta-se nas seguintes hipóteses (MORETTIN, 1999):

- n experimentos independentes e do mesmo tipo são realizadas;
- cada experimento admite dois resultados - Sucesso ou Fracasso;
- a probabilidade de sucesso em cada experimento é p e de fracasso $1 - p = q$

Definimos a variável aleatória X: n$^{\underline{o}}$ de sucessos nos n experimentos. Nestas condições dizemos que X tem distribuição binomial com parâmetros n e p cuja função de probabilidade é dada por:

$$P(X = x) = \begin{pmatrix} n \\ x \end{pmatrix} p^x (1-p)^{n-x}, \qquad x = 0, 1, 2, ..., n$$

Notação: $X \sim Bin(n; p)$.

É importante observar que uma variável aleatória com distribuição Binomial(n, p) pode ser escrita como a soma de n variáveis aleatórias independentes com distribuição Bernoulli(p) (MARTINS, 2005).

Se $X \sim Bin(n; p)$, em que $0 < p < 1$, no entanto, a medida que x vai de 0 a n, $P(X = x)$ primeiro cresce depois decresce, onde atinge seu valor máximo quando x é o maior inteiro menor ou igual a $(n + 1)p$ (LEBENSZTAYN; COLETTI, 2008).

A média e a variância de uma v.a. binomial, com parâmetros n e p são dados, respectivamente, por

$$E(X) = np,$$
$$Var(X) = npq.$$

$$\text{Em que:} \begin{cases} n = & \text{número de repetições do experimento;} \\ p = & \text{probabilidade de sucesso.} \end{cases}$$

Distribuição de Poisson

A distribuição de Poisson é empregada em experimentos, nos quais não se está interessado no número de sucessos obtidos em n tentativas, mas sim no número de sucessos ocorridos durante um intervalo contínuo, podendo ser um intervalo de tempo, espaço,etc. Como por exemplo:

- o número de suicídios ocorridos em uma cidade durante um ano;
- o número de acidentes automobilísticos ocorridos numa rodovia em um mês;
- defeitos por unidade $(m^2, m, etc.)$ por peça fabricada;
- números de erros tipográficos por página, em um material impresso;
- número de usuários de computador ligados à internet;
- número de carros que passam por um cruzamento por minuto, durante certa hora do dia.

A distribuição de Poisson exige que:

- a variável aleatória X seja o número de ocorrências de um evento em um intervalo;
- as ocorrências sejam aleatórias;
- as ocorrências sejam independentes uma das outras;
- as ocorrências tenham a mesma probabilidade sobre o intervalo considerado.

A variável aleatória discreta X tem distribuição de Poisson com parâmetro λ se sua função de distribuição de probabilidade é dada por

$$P(X = x) = \frac{e^{-\lambda}\lambda^x}{x!} \qquad\qquad x = 0, 1, 2, ...$$

Utilizamos a notação $X \sim$ Po(λ) ou $X \sim$ Poisson(λ). O parâmetro λ representa o número médio de eventos ocorrendo no intervalo considerado.
Se X tiver distribuição Poisson com parâmetro λ, então $E(X) = \lambda$ e $Var(X) = \lambda$
Observa-se que uma variável aleatória de Poisson apresenta a particularidade de ter seu valor esperado igual a sua variância.

Uma distribuição de Poisson difere de uma distribuição binomial nestes aspectos fundamentais:

- A distribuição binomial é afetada pelo tamanho da amostra n e pela probabilidade p, enquanto que a distribuição de Poisson é afetada apenas pela média;

- Na distribuição binomial, os valores possíveis da variável aleatória X são $0, 1, 2, ..., n$, mas a distribuição de Poisson têm os valores de X de $0, 1, 2, ...$, sem qualquer limite superior.

2.3.2 Distribuição de probabilidades contínuas

A distribuição contínua representa quantidades aleatórias contínuas que podem tomar qualquer valor dentro de um intervalo explicitado dos números reais. Por exemplo, uma variável aleatória contínua deve ser definida entre os números 0 e 1, ou números reais não negativos, ou para algumas distribuições, qualquer número real. Existem várias distribuições contínuas, algumas delas serão mostradas a seguir.

Distribuição Uniforme

Seja X uma variável aleatória contínua que pode tomar todos os valores num intervalo $[a, b]$. Se a probabilidade de a variável assumir valores num subintervalo for a mesma que para qualquer outro subintervalo de mesmo comprimento teremos então uma distribuição uniforme. A função de densidade de probabilidade de uma variável aleatória contínua deste tipo será:

$$f(x) = \begin{cases} \frac{1}{(b-a)}, & \text{para} \quad a \leq x \leq b; \\ 0, & \text{para qualquer outro valor.} \end{cases}$$

Dize-se que X é uniformemente distribuída no intervalo $[a, b]$, e denota-se por, $X \sim U(a, b)$.

O valor esperado e a variância de uma variável aleatória uniformemente distribuída no intervalo $[a, b]$ são respectivamente

$$E(X) = \frac{a+b}{2}$$
$$Var(X) = \frac{(b-a)^2}{12}$$

A função de distribuição acumulada da uniforme pode ser facilmente avaliada e, vale:

$$F(x) = P(X = x) = \int_{-\infty}^{\infty} \frac{1}{b-a} du = \begin{cases} 0, \text{se} & x < a \\ \frac{x-a}{b-a}, \text{se} & a \leq x \leq b \\ 1, \text{se} & x \geq b \end{cases}$$

Distribuição Normal

A distribuição normal também conhecida como distribuição gaussiana é a mais importante das distribuições de probabilidade contínuas na estatística. Sua importância se deve a vários fatores, entre eles podemos citar o Teorema Central do Limite, o qual é um resultado fundamental em aplicações práticas e teóricas, pois ele garante que mesmo que os dados não sejam distribuídos segundo uma normal a média dos dados converge para uma distribuição normal conforme o número de dados aumenta.

É notório que quando se efetuam repetidas mensurações de determinada grandeza com um aparelho equilibrado, não se chega ao mesmo resultado todas as vezes, obtém-se ao contrário um conjunto de valores que oscilam, de modo aproximadamente simétrico, em torno do verdadeiro valor. Construindo-se o histograma desses valores, obtém-se uma figura com forma aproximadamente simétrica.

Uma variável aleatória X tem distribuição normal se a sua função densidade de probabilidade for da seguinte forma

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} exp\left[-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right], \qquad -\infty \leq x \leq \infty$$

Em que μ e σ^2 são os parâmetros da distribuição normal, tais que $-\infty \leq \mu \leq \infty$ e $\sigma > 0$.

A seguinte notação é empregada: $X \sim N(\mu, \sigma^2)$ se, e somente se, a distribuição de probabilidade da variável aleatória X for dada pela equação acima.

Sua representação gráfica apresenta forma de sino, conforme ilustra a Figura 1. O seu posicionamento em relação ao eixo das ordenadas e seu achatamento vai ser determinado pelos parâmetros μ e σ^2, respectivamente.

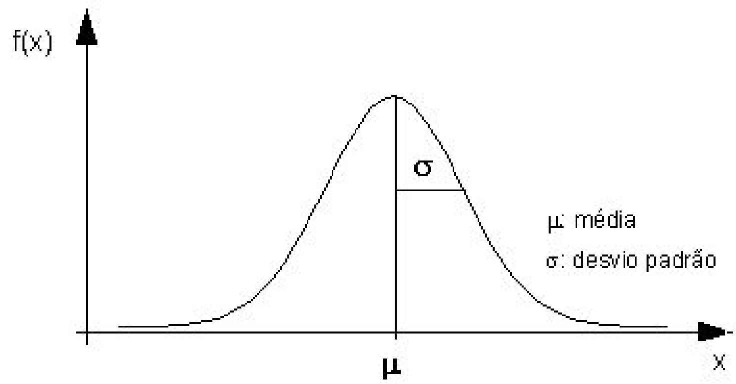

Figura 1: Gráfico da função normal.

A função de distribuição acumulada é dada por:

$$F(x) = P(X \leq x) = \int_{-\infty}^{x} f(x)dx = \int_{-\infty}^{x} \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\frac{(x-\mu)^2}{\sigma^2}} dx$$

A $f(x)$ apresenta as seguintes propriedades:

- possui um ponto de máximo para $x = \mu$;
- tem dois pontos de inflexão cujas abscissas valem $\mu + \sigma$ e $\mu - \sigma$;
- é simetrica em relação a $x = \mu$, e ainda, a média μ é igual a moda, M_o que é igual a mediana M_d;
- tende a zero quando x tende para $\pm\infty$ (assintótica em relação ao eixo x).

Se a variável aleatória normal tiver $\mu = 0$ e $\sigma^2 = 1$, dizemos que ela tem distribuição normal padronizada (também chamada de distribuição normal reduzida). Para obter tal distribuição, isto é, quando se tem uma variável X com distribuição normal com média μ diferente de zero e desvio padrão diferente de um, deve-se reduzi-la a uma variável Z, efetuando o seguinte cálculo:

$$Z = \frac{x - \mu}{\sigma}$$

Assim, a distribuição passa a ter média $\mu = 0$ e desvio padrão $\sigma = 1$. Comumente denotamos esta variável aleatória padronizada de Z, isto é, diz-se que $Z \sim N(0, 1)$. Percebe-se pelo gráfico da curva normal padronizada ilustrado na Figura 2, que aproximadamente toda a probabilidade está dentro de ± 3 a partir da média.

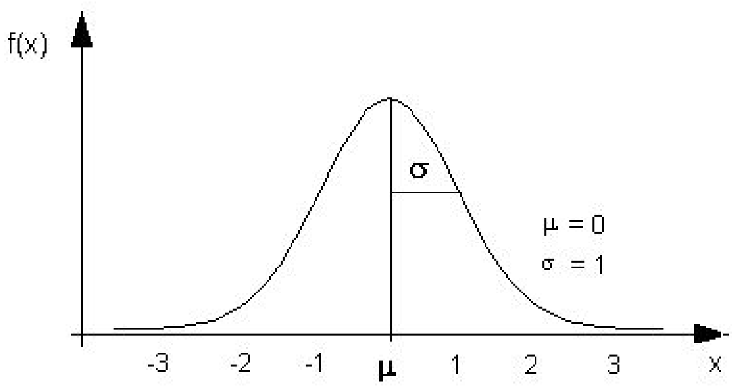

Figura 2: Gráfico da função normal padronizada.

Para média zero e variância unitária, define-se a seguinte distribuição normal padrão:

$$f_Z(z) = N(z; 0, 1) = \frac{1}{\sqrt{2\pi}} exp\left(-\frac{z^2}{2}\right)$$

Se X for uma variável aleatória contínua com distribuição normal, a média e a variância são respectivamente:

$$E(X) = \mu$$
$$Var(X) = \sigma^2$$

Distribuição Qui-Quadrado

Uma v.a. contínua X, com valores positivos, tem uma ditribuição qui-quadrado com v graus de liberdade, se sua função densidade for dada por

$$f(x;v) = \begin{cases} \frac{1}{\Gamma(v/2)2^{v/2}} x^{v/2-1} e^{-x/2}, & x > 0 \\ \\ 0, & x < 0. \end{cases}$$

A notação simplificada para representar uma v.a. X que segue a distribuição qui-quadrado é $X \sim \chi^2_{(v)}$.

A distribuição de qui-quadrado possui diversas aplicações importantes a inferência estatística, e em virtude de sua importância, está tabulada para diferentes valores do parâmetro v. Deste modo, pode-se achar na tabela, o valor denotado por χ^2_α que satisfaça $P(Z \leq \chi^2_\alpha) = \alpha, 0 < \alpha < 1$ (MEYER, 1983).

A distribuição de qui-quadrado apresenta as seguintes características.

- Curva positiva e não simétrica;
- Se $X \sim \chi^2_v$ então: $E(X) = v$, $Var(X) = 2v$ e $M_X(t) = (\frac{1}{1-2t})^{v/2}$;
- A soma de funções com distribuição qui-quadrado independentes tem distribuição qui-quadrado com graus de liberdade igual a soma dos graus de liberdade dessas distribuições, isto é: $X_i \sim \chi^2_{v_i}, i = 1, 2, ..., n \Rightarrow \sum_{i=1}^n X_i \sim \chi^2_{(\sum v_i)}$;
- O seu aspecto gráfico depende do parâmetro v conforme ilustra a Figura 3.

Sua função densidade de probabilidade está definida apenas para valores não-negativos de x e, assim como a t de student, depende dos graus de liberdade v.

O teste qui-quadrado formulado por Karl Pearson no início do século XX, é um teste não paramétrico, ou seja, não depende dos parâmetros populacionais, como média e variância, e que é muito utilizado em muitas situações práticas no contexto de Estatística Aplicada.

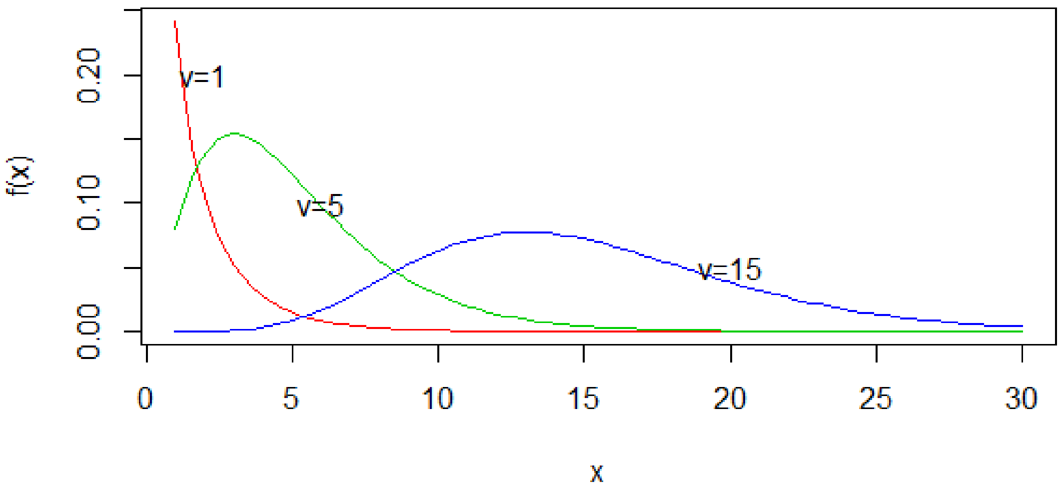

Figura 3: Gráfico da distribuição de qui-quadrado com 1, 5 e 15 $g.l.$.

Se destina a encontrar um valor da dispersão para duas variáveis nominais, e avaliar a associação existente entre variáveis qualitativas.

Para aplicar o teste as seguintes suposições precisam ser satisfeitas:

- Os dois grupos são independentes.
- Os itens de cada grupo são selecionados aleatoriamente.
- As observações devem ser frequências ou contagens.
- Cada observação pertence a uma e somente uma categoria.
- A amostra deve ser relativamente grande

A estatística do teste qui-quadrado para uma única amostra é dada por:

$$X^2 = \sum_{i=1}^{k} \frac{(o_i - e_i)^2}{e_i}$$

Em que:
o_i é a frequência observada na classe i,
e_i é a frequência esperada no i-ésimo intervalo de classe,
k é o número de categorias (ou classe) consideradas na distribuição probabilística sob investigação.

Há duas hipóteses a serem testadas (ou contrastadas):

- Hipótese nula: não existe diferença entre as frequências (contagens) dos grupos. Portanto, não há associação entre os grupos.
- Hipótese alternativa: existe diferença entre as frequências. Portanto, há associação entre os grupos

Faz-se necessário avaliar a estatística de teste a partir dos dados da amostra e comparar com quantil superior da distribuição de qui-quadrado com v $g.l.$ ao nível de significância α, (em geral $\alpha = 0,05$). A regra de decisão será rejeitar H_0 ao nível de significância α, se, e somente se, $X^2 = \chi^2_{[v,\alpha]}$.

Quanto maiores forem as diferenças entre as frequências esperadas e observadas maior será o valor da X^2 concordando com a rejeição de H_0 .

2.4 Função Geradora de Momentos

Seja X, uma variável aleatória discreta, com distribuição de probabilidade $p(x_i) = P(X = x_i)$, $i = 1, 2, ...$ A função M_X, denominada função geradora de momentos de X, é definida por (MEYER, 1983)

$$M_X(t) = \sum_{i=1}^{\infty} e^{tx_i} p(x_i).$$

A função geradora de momentos, fgm, de uma variável aleatória contínua X com f.d.p. f definida em $\mathbb{R}$, é dada por (AZEVEDO, 2007)

$$M_X(t) = \int_{-\infty}^{\infty} e^{tx} f(x) dx,$$

existindo desde $\int_{-\infty}^{\infty} \mid e^{tx} f(x) \mid dx < \infty$.

Se existir, $M_X(t)$ é real e positiva para todo t num intervalo $(-t_0, t_0)$, com $t_0 > 0$ sendo diferenciável em $t = 0$. Note que $M_X(0) = 1$. $M_X(t)$ pode não existir para $t \neq 0$.

Em qualquer dos casos, o discreto ou o contínuo, $M_X(t)$ é apenas o valor esperado de e^{tX}. Por isso, pode-se combinar as expressões acima e escrever (MEYER, 1983)

$$M_X(t) = \mathbb{E}(e^{tX}).$$

$M_X(t)$ é o valor que a função M_X toma para a variável (real) t. A notação, que indica a dependência de X, é empregada caso deseja-se considerar duas variáveis aleatórias, X e Y, e depois investigar a função geradora de momentos de cada uma delas, a saber M_X e M_Y.

Se X tem função geradora de Momentos $M_X(t)$, então

$$\frac{\partial^x}{\partial t^x} M_X(t)\Big|_{t=0} = \mathbb{E}(X^x).$$

Demonstração: Se M_X existe, então $M_X(t)$ é diferenciavel até a ordem $x(>0)$ em $t = 0$.

$$\frac{\partial^x}{\partial t^x} M_X(t)\Big|_{t=0} = \frac{\partial^x}{\partial t^x}\mathbb{E}[e^{tX}]\Big|_{t=0} = \mathbb{E}\left[\frac{\partial^x}{\partial t^x} e^{tX}\Big|_{t=0}\right] = \mathbb{E}(X^x).$$

No caso de X ser uma variável aleatória simples, essa é a demonstração, pois a esperança é finita e podemos derivar dentro da soma. No caso geral, há que se justificar a derivada dentro da esperança. Este passo será omitido porque envolve Teorema de Medida (ROLLA, 2013).

2.5 Obtenção da distribuição qui-quadrado a partir de normais padrão

A distribuição qui-quadrado pode ser interpretado de duas formas: como sendo um caso particular da distribuição gama, ou como a soma de normais padronizadas ao quadrado. Assumindo-se que $X_i \sim N(0,1)$, então $\sum_{j=1}^{r} X_j^2 \sim \chi_r^2$. Esse fato será demonstrado no **Teorema** 2.4.1.

Segundo Bussab e Morettin (2002), o quadrado de uma variável aleatória com distribuição normal padrão é uma variável aleátoria com distribuição qui-quadrado com 1 $g.l.$ Ou seja, uma variável aleatória $\chi^2_{(v)}$ pode ser vista como a soma de v $v.a.$ independentes com distribuição normal padrão ao quadrado.

Uma variável aleatória contínua X tem distribuição qui-quadrado com v graus de liberdade se sua função densidade for dada por:

$$f(x) = \frac{1}{2^{\frac{v}{2}}\Gamma(\frac{v}{2})} x^{(\frac{v}{2}-1)} exp(-\frac{x}{2}); \qquad v > 0, \qquad x > 0.$$

em que

$$\Gamma(w) = \int_0^\infty x^{w-1} e^{-x} dx, \qquad w > 0.$$

Denotamos $X \sim \chi^2_v$.

Nota-se pelo gráfico da distribuição qui-quadrado, conforme ilustra a Figura 4, que ela é assimétrica à direita e x só assume valores positivos. Isto porque trata-se da soma de normais ao quadrado, logo só pode ser positiva.

A distribuição qui-quadrado possui diversas aplicações na inferência estatística, e, como o caso da normal, existem tabelas para obter probabilidades.

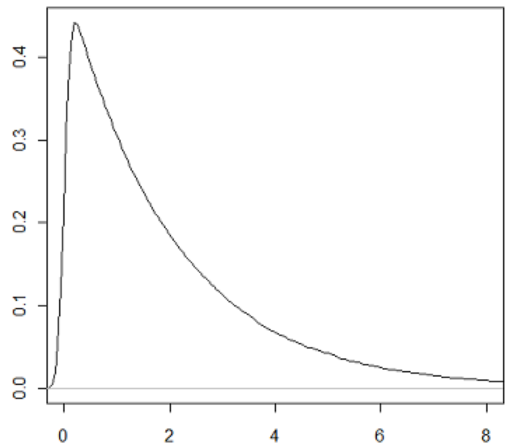

Figura 4: Gráfico da função qui-quadrado com 2 graus de liberdade.

Para entender a ideia de graus de liberdade, considera-se um conjunto de dados qualquer. Graus de liberdade é o número de valores deste conjunto de dados que podem variar após terem sido impostos certas restrições a todos os valores.

Teorema 2.5.1 *Sejam $X_i \sim N(0,1)$ variáveis aleátorias independentes, com $i = 1, 2, ..., v$ e $v \in \mathbb{N}$. Então $Z = \sum_{i=1}^{v}(X_i)^2$ segue uma distribuição qui-quadrado com v graus de liberdade (ACTION, 2011).*

Vamos demonstrar este teorema via a função geradora de momentos. Como as variáveis aleátorias $X_i, ..., X_n$ são independentes, temos que

$$M_Z(t) = \mathbb{E}[e^{tZ}] = \mathbb{E}\left(e^{t\left(\sum_{i=1}^{v}(X_i)\right)}\right) = \mathbb{E}\left(\prod_{i=1}^{v} e^{t(X_i)^2}\right) = \prod_{i=1}^{v} \mathbb{E}\left(e^{t(X_i)^2}\right).$$

Mas

$$\mathbb{E}\left(e^{t(X_i)^2}\right) = \int_{-\infty}^{\infty} e^{tx_i^2}\left(\frac{1}{\sqrt{2\pi}}\right)e^{-(1/2)x_i^2}dx_i = \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}}e^{-(1/2)(1-2t)x_i^2}dx_i$$

e, portanto,

$$\mathbb{E}\Big(e^{t(X_i)^2}\Big) = \frac{1}{\sqrt{1-2t}}\int_{-\infty}^{\infty}\frac{\sqrt{1-2t}}{\sqrt{2\pi}}e^{-\frac{1}{2}(1-2t)x_i^2}dx_i = \Big(\frac{1}{1-2t}\Big)^{\frac{1}{2}}$$

Em que a última integral é igual a 1, pois trata-se justamente de uma normal com média 0 e variância 1/(1 - 2t). Portanto,

$$\mathbb{E}(e^{tZ}) = \prod_{i=1}^{v}\mathbb{E}\big(e^{t(X_i)^2}\big) = \prod_{i=1}^{v}\frac{1}{\sqrt{1-2t}} = \Big(\frac{1}{1-2t}\Big)^{v/2}.$$

correspondente a função geradora de momentos da distribuição qui-quadrado com v graus de liberdade.

Como visto no **Teorema** 2.4.1, se X é uma variável aleatória com distribuição qui-quadrado com v graus de liberdades, sua função geradora de momentos é dado por:

$$M_X(t) = \mathbb{E}\big(e^{tX}\big) = \Big(\frac{1}{1-2t}\Big)^{v/2}$$

Desta forma, temos que

$$M_X'(t) = v\Big(\frac{1}{1-2t}\Big)^{\frac{v+2}{2}}$$

e

$$M_X''(t) = (v^2+2v)\Big(\frac{1}{1-2t}\Big)^{\frac{v+4}{2}}.$$

Pontanto, pode-se calcular o valor esperado e a variância da variável X. De fato, temos que

$$E(X) = M_X'(0) = v\Big(\frac{1}{1}\Big)^{\frac{v+2}{2}} = v$$

e

$$Var(X) = E(X^2) - E^2(X) = M_X''(0) - M_X'^2(0)$$

de onde concluí-se que

$$Var(X) = (v^2+2v)\Big(\frac{1}{1}\Big)^{\frac{v+4}{2}} - v^2 = v^2 + 2v - v^2 = 2v.$$

3 Aplicação

Na teoria de estatística existem vários resultados que podem ser ilustrados via simulação, o que possibilita a compreensão e visualização de conceitos e resultados. Podemos utilizar recursos computacionais e em particular simulações para inferir distribuições amostrais de quantidades de interesse.

A simulação é um processo computacional que utiliza números aleatórios para produzir resultados. Assim seu uso é apenas um ponto de partida, pois estas são especialmente úteis para explorar situações onde resultados teóricos não são conhecidos ou não podem ser obtidos.

Uma variável aleatória qui-quadrado pode ser vista como a soma de normais padrão ao quadrado, independentes. Para ilustrar esta relação entre a distribuição normal padrão e a qui-quadrado, faremos simulações alternando os graus de liberdade em 1, 5 e 15 respectivamente da distribuição qui-quadrado, utilizando o programa *R Core Team* na versão 3.0.2, por meio do *interface* RStudio na versão 0.97.551. E, para uma melhor visualização destes resultados construiremos o histograma dos valores com a curva da distribuição e o seu respectivo *qq-plot*.

Resultado 1: Se $Z \sim N(0,1)$ então $Z^2 \sim \chi^2_{(1)}$.

Inicialmente para ilustrar este resultado gera-se números aleatórios para que os resultados possam ser reproduzidos baseados na distribuição definida. O nome da função é iniciado pela letra r mais o nome da distribuição (ex: `rnorm`). Neste contexto, começaremos gerando uma amostra de 10000 números da distribuição normal padrão. A seguir faremos um histograma dos dados obtidos e sobrepor a curva da distribuição, utizando os comandos abaixo e o resultado está no gráfico da esquerda da Figura 5.

```
z <- rnorm(10000)
hist(z, prob=T)
curve(dnorm(x), -4, 4, col=2, add=T)
```

É evidente que, para fazer a comparação do histograma e da curva da distribuição normal padrão é necessário que o histograma seja de frequências relativas e para isto usamos o argumento `prob=T`.

Estudaremos então, o comportamento do quadrado da variável. O gráfico da direita da Figura 5 mostra o histograma dos quadrados dos valores da amostra e a curva da distribuição de $\chi^2_{(1)}$.

```
hist(z^2, prob=T)
curve(dchisq(x, df=1),0 , 10, col=2, add=T)
```

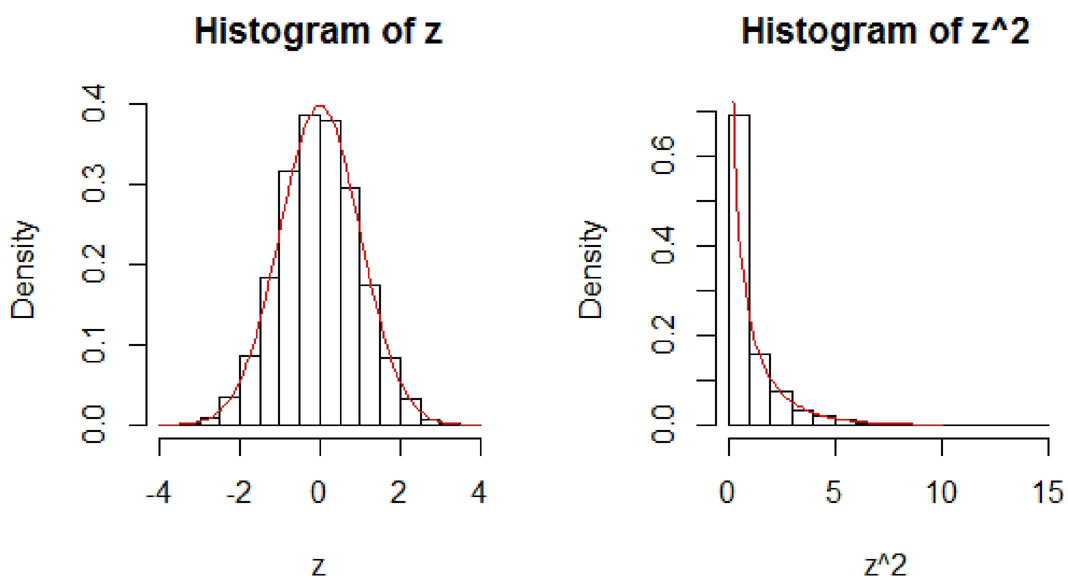

Figura 5: Histograma das amostras e a curva da distribuição normal padrão (esquerda) e o histograma dos valores ao quadrado com a curva da distribuição $\chi^2_{(1)}$ (direita).

Outra forma de comparar distribuições é através dos quantis destas distribuições e para isto utiliza-se o *qq-plot*, que é um gráfico de dados ordenados contra os quantis esperados de uma certa distribuição. Quanto mais próximo os pontos estiverem da bissetriz do primeiro quadrante, mais próximos os dados observados estão da distribuição considerada. Portanto, para se fazer o *qq-plot* segue-se os seguintes passos.

- obter os dados;
- obter os quantis da distribuição;
- fazer um gráfico dos dados ordenados contra os quantis da distribuição.

No entanto, vamos considerar como dados os quadrados da amostra da normal padrão obtida acima. Depois obtemos os quantis da distribuição χ^2 usando a função `qchisq` em um conjunto de probabilidades geradas pela função `ppoints`. E, por fim usa-se a função *qq-plot* para obter o gráfico na Figura 6, adicionando neste gráfico a bissetriz do primeiro quadrante para facilitar a avaliação do ajuste. Ilustraremos isto no comando abaixo.

```
quantis <- qchisq(ppoints(length(z)),df=1)
qqplot(quantis, z^2)
abline(0,1,col=3)
```

Nota-se que o comando `qchisq(ppoints(length(2)))` citado acima está concatenando 3 comandos e calcula os quantis da χ^2 a partir de uma sequência de valores de probabilidades gerada por `ppoints`. O número de elementos desta sequência deve ser igual ao número de dados e por isto usou-se `length(z)`.

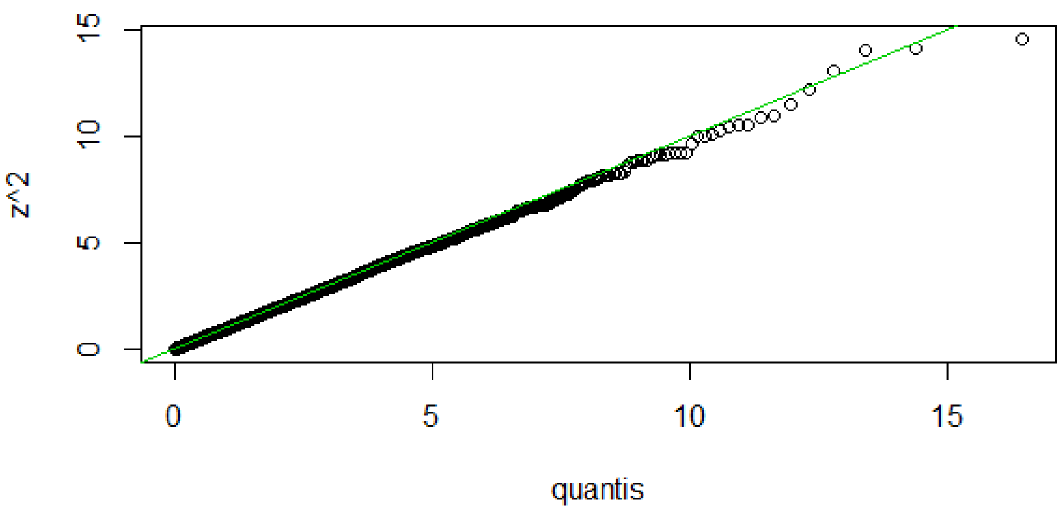

Figura 6: Comparando dados e quantis da χ^2 utilizando o *qq-plot*.

Resultado 2: Se $Z_1, ..., Z_5 \sim N(0,1)$, então $\sum_{i=1}^{5} Z_i^2 \sim \chi^2_{(5)}$.

Para ilustrar este resultado geramos 10000 amostras de 5 elementos cada da distribuição normal padrão, elevando os valores ao quadrado e, para cada amostra, somar os quadrados dos cinco números. Na Figura 7 mostramos no gráfico à esquerda, o histograma dos valores obtidos com a curva da distribuição esperada e no da direita o *qq-plot* para a distribuição $\chi^2_{(5)}$.

```
z <- matrix(rnorm(50000), nc=5)
sz2 <- apply(z^2, 1, sum)
par(mfrow=c(1,2))
hist(sz2, prob=T)
curve(dchisq(x, df=5), 0, 50, col=2, add=T)
qqplot(qchisq(ppoints(length(sz2)), df=5), sz2)
abline(0,1,col=3)
```

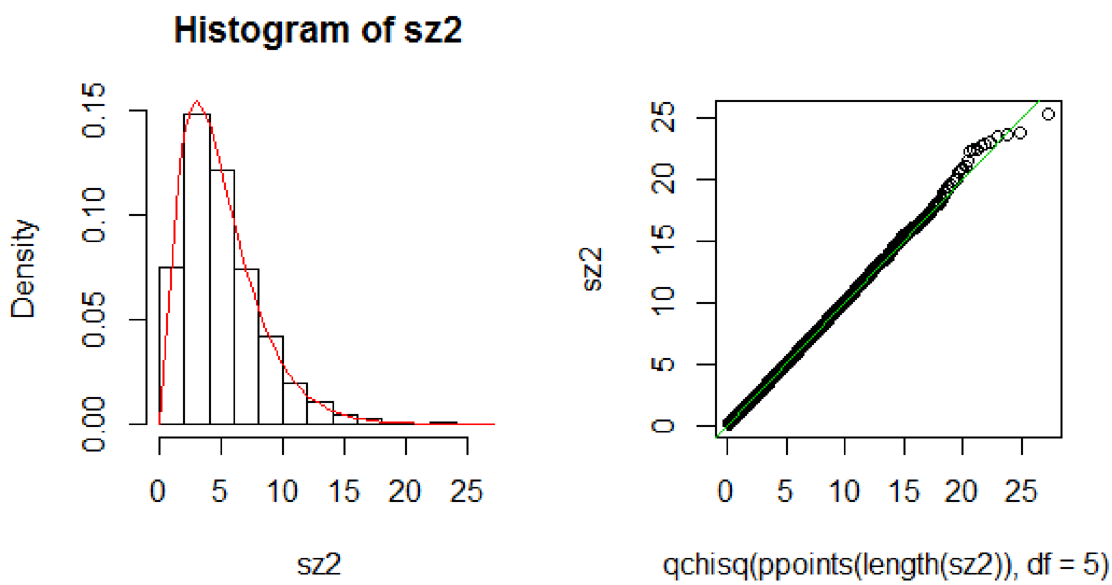

Figura 7: Histograma de uma amostra da soma dos quadrados de cinco valores da distribuição normal padrão e a curva da distribuição de $\chi^2_{(5)}$ (esquerda) e o respectivo *qq-plot*.

Resultado 3: Se $Z_1, Z_2, ..., Z_{15} \sim N(0,1)$ então, $\sum_{i=1}^{15} Z_i^2 \sim \chi^2_{(15)}$

Ilustraremos este resultado gerando 10000 amostras de 15 elementos cada da distribuição normal padrão, eleva-se os valores ao quadrado e, para cada amostra, soma os quadrados dos quinze números. O gráfico mostra o histograma dos valores obtidos com a curva da distribuição esperada (esquerda) e o *qq-plot* para a distribuição $\chi^2_{(15)}$ ilustrado na Figura 8.

```
z <- matrix(rnorm(150000), nc=15)
sz3 <- apply(z^2, 1, sum)
par(mfrow=c(1,2))
hist(sz3, prob=T)
```

```
curve(dchisq(x, df=15), 0, 150, col=2, add=T)
qqplot(qchisq(ppoints(length(sz3)), df=15), sz3)
abline(0,1,col=3)
```

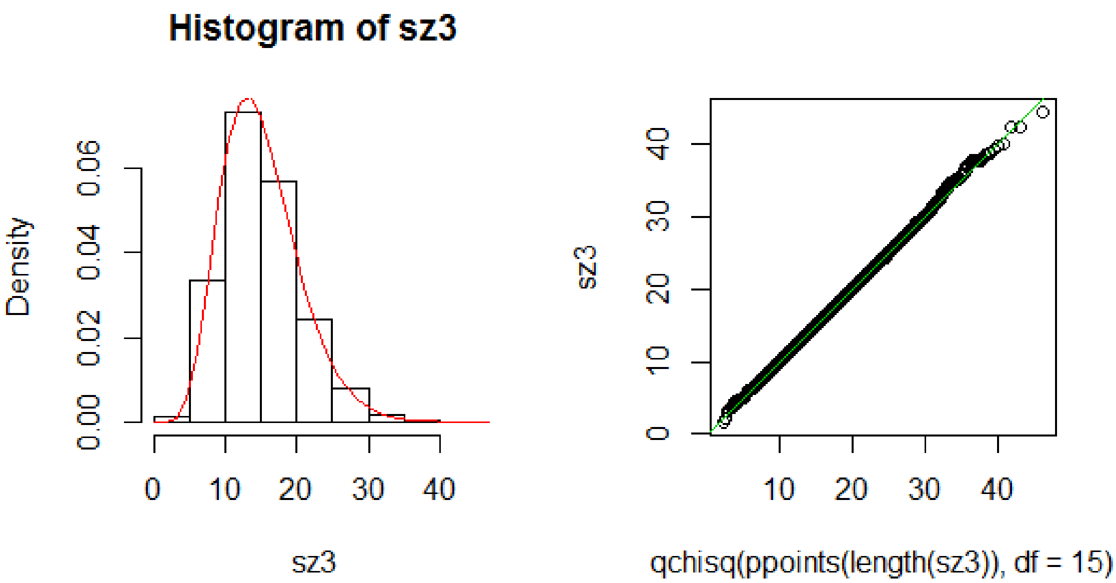

Figura 8: Histograma de uma amostra da soma dos quadrados de quinze valores da distribuição normal padrão e a curva da distribuição de $\chi^2_{(15)}$ (esquerda) e o respectivo *qq-plot*.

4 Conclusão

A ideia de probabilidade desempenha papel importante em muitas situações que envolvem uma tomada de decisão e a utilizamos em muitos momentos de nossa vida, nos proporciona um modo de medir a incerteza e de como aplicar a matemática para resolver problemas reais. Muitas vezes profissionais da área de exatas fazem uso de modelos probabilísticos para representar situações reais ou então para descrever um experimento aleatório. Para isso, indica-se o ensino das noções probabilísticas a partir de um conjunto de regras e métodos que visam à descoberta por meio de proposição de problemas concretos e da realização de experimentos reais ou simulados.

A distribuição normal é uma das mais importantes na estatística porque um grande número de fenômenos aleatórios podem ser aproximados por essa distribuição, a suposição de normalidade nos dados também é muito frequente na maioria dos modelos estatísticos conhecidos e usados.

Já a distribuição qui-quadrado é uma distribuição contínua de grande importância para a inferência estatística. É uma distribuição que surge no contexto de somas de quadrados de variáveis aleatórias com distribuição normal. É uma das mais utilizadas em estatística inferencial principalmente para realizar testes qui-quadrado. Estes testes servem para avaliar quantitativamente a relação entre o resultado de um experimento e a distribuição esperada para o fenômeno. Muitos outros testes de hipóteses também usam a distribuição qui-quadrado.

Geralmente usamos distribuições para cálculo de probabilidade associados a resultados de interesse. No entanto, existem os recursos computacionais e em especial simulações para inferir tais distribuições amostrais. Um dos passos principais no desenvolvimento de uma simulação é a geração de números aleatórios, para que os valores assumidos pelas variáveis aleatórias gere o modelo simulado. E assim, temos a oportunidade de relacionar a teoria com a prática, dar mais significado ao aprendizado e argumentar com base em fatos.

Elaborar este trabalho demostrou ser uma experiência muito interessante, e descobrimos fatos acerca da teoria de probabilidades que a conseguiram tornar objeto de grande curiosidade. Em seu desenvolvimento buscamos aprofundar nossos conhecimentos acerca do tema probabilidade e mostrar através de simulação a geração de distribuições de probabilidades, a maior parte dos algoritmos geradores desses modelos encontra-se na literatura de estatística computacional e simulação, Porém é difícil encontrar uma obra que as reúna. Daí a importância deste trabalho, o que constitui uma fonte para trabalhos futuros.

Referências

ACTION, P. *Distribuição qui-quadrado.* 2011. Disponível em: <http://www.portalaction.com.br/content/63-ditribuição qui-quadrado>. Acesso em: 20/01/2014.

AZEVEDO, C. C. *Notas de Estatística e Matemática.* [S.l.], 2007.

BUSSAB, W. O.; MORETTIN, P. A. *Estatística Básica.* 5. ed. São Paulo: Saraiva, 2002.

CORREA, S. M. B. B. *Probabilidade e Estatística.* 2. ed. Belo Horizonte: Puc Minas Virtual, 2003.

LEBENSZTAYN Élcio; COLETTI, C. F. *Probabilidade: Teoria e exercícios.* [S.l.], 2008.

MARTINS, M. E. G. *Introdução à Probabilidade e à Estatística.* [S.l.], 2005.

MEYER, P. L. *Probabilidade: aplicações à estatística.* 2. ed. São Paulo: LTC, 1983.

MORETTIN, L. G. *Estatística básica-probabilidade.* 7. ed. São Paulo: Makron Books, 1999.

NOE, M. *História da Probabilidade.* 2009. Disponível em: <http://www.brasilescola.com/matematica/historia-probabilidade.htm>. Acesso em: 23/09/2013.

PORTNOI, M. Probabilidade, variáveis aleatórias, distribuições de probabilidade e geração aleatória. *Universidade Salvador*, p. 30, 2005.

ROLLA, L. T. *Introdução à Probabilidade.* [S.l.], 2013.