

## APRENDER FAZENDO: COMO TIRAR PROVEITO DO COMPUTADOR PARA MELHORAR A APRENDIZAGEM DA ESTATÍSTICA

Lorí Viali – PUCRS/UFRGS – viali@pucrs.br<sup>1</sup>

### INTRODUÇÃO

Existe uma inegável má vontade por parte de grande parte dos estudantes com disciplinas de Estatística ou Probabilidade. Não é incomum ouvir um estudante dizer que escolheu determinado curso pensando ter-se livrado da matemática e para sua surpresa lá está uma disciplina de Estatística no currículo. Disciplinas de Estatística costumam ser mais freqüentes, embora menos numerosas, nos currículos da maioria dos cursos universitários do que disciplinas de Matemática. São raros os cursos que não contemplem pelo menos uma disciplina envolvendo Estatística ou Probabilidade ou mesmo ambas.

Com a reforma universitária realizada e em alguns casos ainda sendo realizada houve um enxugamento nos currículos e várias disciplinas foram eliminadas. Uma das áreas mais atingidas foi a da Matemática. A maior redução, talvez tenha sido, nas Engenharias onde a Matemática dominava boa parte do currículo. A Estatística não sofreu muito nesse processo de adequação, pois sua presença nos currículos já era mínima. Mesmo assim alguns cursos de Engenharia que contemplavam seis horas semanais mudaram para uma disciplina de apenas quatro horas.

Outras áreas também apresentaram reduções. Cursos de Administração, Economia e Contábeis que tinham três ou até quatro disciplinas de quatro horas semanais acabaram reduzindo a carga a duas disciplinas de quatro horas semanais e alguns casos até menos. Outros cursos que praticamente não tinham mais o que reduzir, pois apresentavam apenas uma disciplina de quatro horas acabaram reduzindo-a a duas horas semanais.

Alguns cursos simplesmente eliminaram disciplinas da área de exatas. A Pedagogia em muitas universidades é um caso. Existiam disciplinas de matemática e de estatística nesses cursos. Hoje é raro encontrá-las. É fácil perceber o prejuízo advindo uma vez que os professores formados nesses cursos serão os alfabetizadores ou professores das séries iniciais. Esses professores serão os encarregados de ensinar as primeiras noções numéricas e, portanto, terão um papel decisivo no gosto ou não dos alunos pela

---

<sup>1</sup> Professor titular do Departamento de Estatística, Faculdade de Matemática da PUCRS.  
Professor adjunto do Departamento de Estatística, Instituto de Matemática da UFRGS.

matemática. Se um professor não sabe ou pouco sabe sobre matemática ou análise de dados é improvável que ele vá mostrar ou demonstrar aos seus alunos apreço pelos assuntos.

O número escasso de horas dessas disciplinas não é o maior problema. A dificuldade é que os conteúdos a serem desenvolvidos não acompanharam a redução da carga horária na mesma proporção. Em muitos casos eles ficaram inalterados quando o número de horas foi reduzido à metade ou menos. Não é incomum encontrarmos disciplinas de dois créditos que contemplem conteúdos envolvendo praticamente toda a Estatística Básica. Um exemplo típico é incluir Estatística Descritiva, Probabilidade, Amostragem, Estimação e Testes de Hipóteses para serem desenvolvidos em cerca de 30 encontros semestrais quando o sensato seria tratar apenas um desses tópicos para que a aprendizagem tivesse um mínimo de qualidade.

Essa não é uma situação típica de um único país. Levy, Evans e Zirger (2003, p. 3) colocam:

*“Estatística para Administradores é desenvolvida em dois créditos, em encontros semanais de 2 horas e 45 minutos durante cinco semanas e cobre os conceitos estatísticos fundamentais incluindo estatística descritiva, probabilidade, variáveis aleatórias, amostragem, estimação e teste de hipóteses.”*

Eles estão se referindo a uma disciplina desenvolvida num MBA da faculdade de administração da Universidade de Cincinnati. Por esta razão e ironizando a quantidade de tempo disponível acrescentam:

*“Por uma série de razões os responsáveis pelo programa da disciplina escolheram tomar por base a planilha para a implementação dos procedimentos computacionais. Dado o luxo relativo de 30 horas de contato com os alunos o ensino com a planilha facilita o manuseio dos dados e a sua análise.”*

Diante desse quadro o professor encontra-se numa posição difícil. Em geral a saída é desenvolver os conteúdos tipo receita, dando apenas notícias rápidas de cada tópico sem aprofundar, pois não existe espaço para discussões, reflexões ou mesmo qualquer prática.

Contudo não é obrigatório que a situação seja assim. Aqui se apresenta uma abordagem pesquisada por alguns anos de prática com diferentes cursos universitários onde têm sido utilizados recursos computacionais, especificamente a planilha, com relativo

sucesso. Obviamente o computador não é o Santo Graal do ensino. Não é possível resolver-se todos os problemas com a sua utilização, mesmo por que isso nem sempre é simples. A tecnologia resolveu alguns problemas, mas, por outro lado, criou outros. Entretanto gostando ou não quer queiramos ou não a tecnologia está aí e veio para ficar.

O computador permite que sejam abordados problemas reais, principalmente aqueles que envolvem grandes quantidades de dados, problemas esses que por razões óbvias não podem ser tratados nos livros didáticos ou manipulados com o recurso de uma calculadora.

Gandhi (2006, p. 2) coloca: “É reconhecido por muitos estatísticos que são necessárias mudanças na Educação Estatística” e dentre várias sugestões para essas mudanças acrescenta (p. 3): “Um laboratório de estatística deve ser parte de um curso de estatística”.

Um aluno que foi submetido em toda a sua vida escolar somente a problemas com respostas bem definidas, dados bem comportados e simplificações irrealistas terá grande dificuldade de se adaptar ao mundo não escolar ou ao mercado de trabalho.

Existem considerações e situações que sempre são evitadas ou mesmo suprimidas nos livros didáticos. Da mesma forma nas aulas desenvolvidas sem recursos tecnológicos os problemas são sempre irrealisticamente simplificados ou no caso da estatística o número de dados é bastante reduzido ou mesmo adaptado para que sejam adequados a uma exposição oral em que os recursos são apenas o giz, o quadro e a imaginação.

Obviamente a imaginação é um recurso valioso, mas ela pode ser estimulada, desenvolvida e ampliada com o tratamento de problemas reais ou quase. Grandes conjuntos de dados têm que ser tratados de forma diferente de que pequenos conjuntos. É necessário fazer uma redução prévia da quantidade de modos que consigamos perceber com o que estamos lidando. Convém salientar que isso é necessário porque nossos sentidos são limitados. Para tratar grandes massas de dados num computador não é necessário condensá-los ou resumi-los. Essa é uma necessidade essencialmente humana que é fruto das nossas limitações. Para podermos entender e comunicar resultados de grandes conjuntos de dados é necessário que eles sejam compactados ou resumidos. Dessa forma as tabelas e os gráficos são os maiores aliados nessa tarefa. Estatística Descritiva é essencialmente a organização, resumo e apresentação de conjuntos de dados. Dados esses que podem ser tanto qualitativos quanto quantitativos, pois o computador, ao contrário, do que muitos imaginam é uma máquina que processa texto e que precisa ser adaptada para processar números.

A abordagem gráfica é hoje uma linguagem universal. Qualquer mídia irá utilizá-los em menor ou maior quantidade. Portanto saber fazê-los e interpretá-los faz parte da alfabetização. É raro depararmos com um jornal, revista ou outras mídias impressas que não façam uso de gráficos e diagramas. Os próprios textos didáticos os utilizam em quantidade bem como a literatura científica. É difícil imaginar um profissional de qualquer área que não precise ter uma boa idéia de tabelas e gráficos.

Tratar esse assunto de forma manual ou apenas com recursos convencionais como quadro, giz e livro texto é continuar aquele ensino simplista e irreal que dará apenas um arremedo de realidade que um aluno ou profissional precisa atualmente. Hall (1995, p. 6) coloca: “Uma vantagem adicional das planilhas é a facilidade com que resultados gráficos podem ser obtidos. Os gráficos auxiliam os estudantes a entender os conceitos matemáticos e são também auxiliares na modelagem matemática.”

Não são muitas as disciplinas de Estatística que apresentam exclusivamente conteúdos de Estatística Descritiva. Um exemplo típico é a disciplina oferecida aos cursos de Engenharia que envolve Descritiva, Probabilidade e Inferência. Mesmo com a oportunidade de iniciar o curso com a Descritiva muitos professores preferem iniciar com Probabilidade. Essa não é a opinião do autor e nem a de Hall (1995, p. 5) que coloca:

*“A experiência desse autor mostra que a maioria dos alunos de Engenharia ou Ciências vão apreciar melhor a Estatística e a Probabilidade se o estudo for iniciado com uma análise de dados prática e evoluir pela Estatística até a Probabilidade, isto é, progredir do prático para o abstrato”.*

É raro o exemplo, em qualquer área, onde um pesquisador comece formulando uma teoria sem dados ou fatos. No entanto, no ensino de Estatística, que não deixa de fazer parte da metodologia da ciência, muitos professores começam um curso que envolve estatística apresentando sem qualquer preparação adicional os modelos probabilísticos que cientistas e pesquisadores levaram anos, às vezes, décadas para estabelecer.

## **1. RECURSOS**

A planilha oferece uma boa variedade de recursos para se organizar, resumir e apresentar conjuntos de dados (amostrais ou populacionais). A principal vantagem da planilha é a sua grande base instalada e seu preço relativamente barato (VIALI, 2004, p. 356).

Hall (1995, p. 6) coloca a respeito da abordagem proposta: “O curso tem uma abordagem prática pelo uso de um pacote computacional. Para um primeiro curso dedicado

a análise de dados o software específico não é essencial e qualquer boa planilha pode ser utilizada.” E acrescenta: “A planilhas são utilizadas em quase todo o comércio e a indústria e a habilidade em utilizá-las é uma competência profissional útil.”

Uma vez obtido o conjunto a primeira tarefa é resumi-lo de forma que se possa ter uma idéia de modelo, isto é, que tipo de conjunto temos em mãos. Para fazer isso é necessário considerar o tipo de variável sendo tratada. A variável pode ser qualitativa ou quantitativa. Se for uma variável qualitativa então pode-se utilizar a função CONT.SE(conjunto; critérios) para resumir o conjunto. Essa função também pode ser utilizada para contar valores (atributos) de uma variável qualitativa. Se ela for quantitativa a função adequada nesse caso é a FREQUÊNCIA(Matriz\_dados, Matriz\_bin). Para utilizá-la é necessário construir um novo conjunto de dados denominado de “Matriz\_bin” que seriam as classes. Os limites dessas classes devem ser colocados em células contíguas

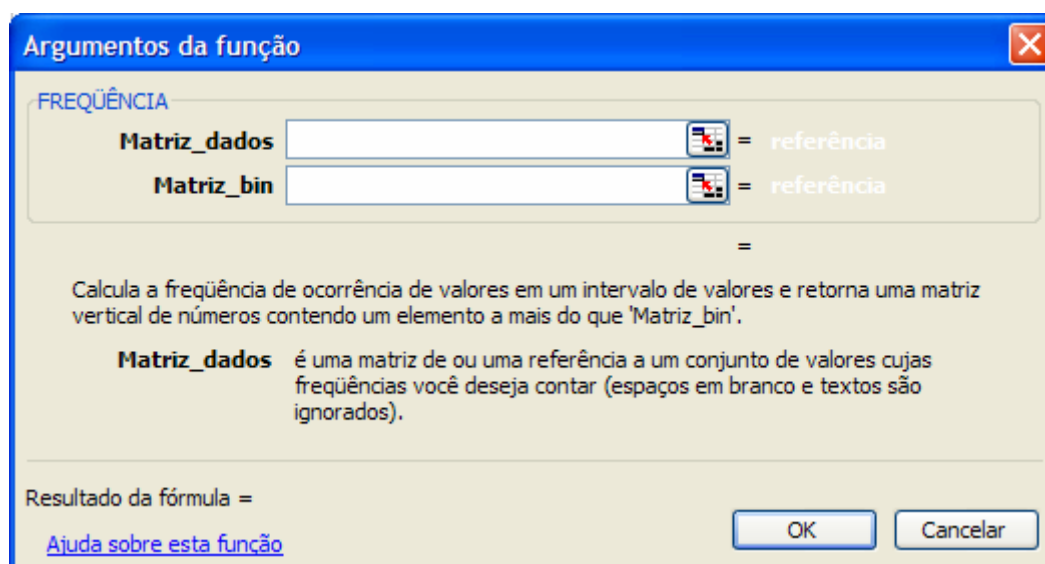


Figura 01 – Função para agrupar valores de variáveis discretas e qualitativas

Esse tipo de estratégia pode ser aplicada a qualquer área. Na de humanas, por exemplo, em um curso de pedagogia pode ser uma coleção de conceitos. Pode-se ilustrar considerando os resultados de uma avaliação de uma turma de digamos 50 alunos realizada por meio dos seguintes conceitos: ótimo, muito bom, bom, regular e insuficiente. Digitado esses resultados na planilha então rapidamente é possível fazer uma tabela com os resultados calculando-se então os percentuais em cada caso.

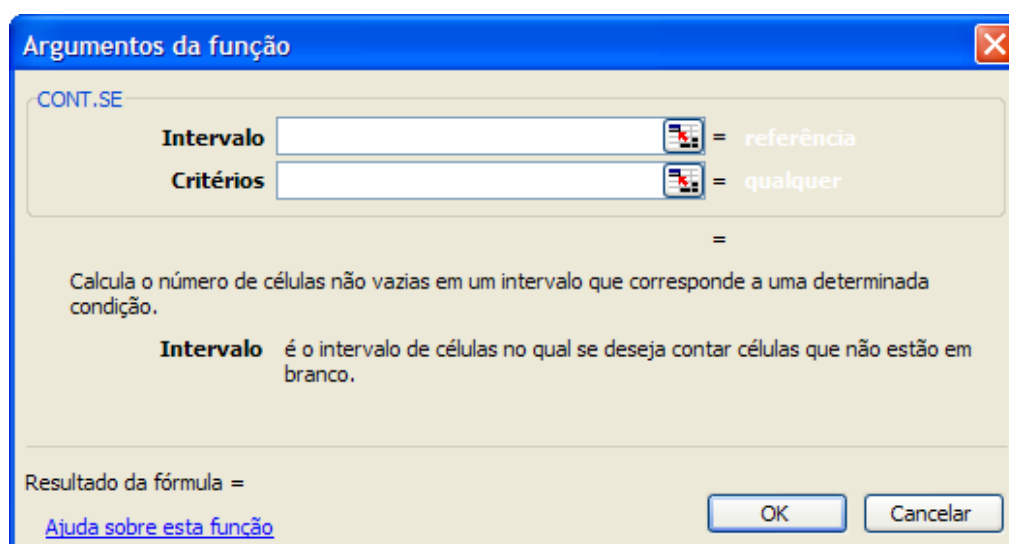


Figura 02 – Função para agrupar valores de variáveis contínuas

No item de menu "Ferramentas de análise" existe um recurso denominado de "Análise de Dados". Este recurso é de fato uma biblioteca que precisa ser ativada, pois ela está instalada só que não é ativada automaticamente. Para que ela se torne operacional é necessário clicar no item de menu "Ferramentas" subitem "Suplementos" e então marcar o suplemento (biblioteca) "Ferramentas de Análise". Somente depois de tomada esta providência o subitem "Análise de Dados" do item de menu "Ferramentas" estará disponível para uso.

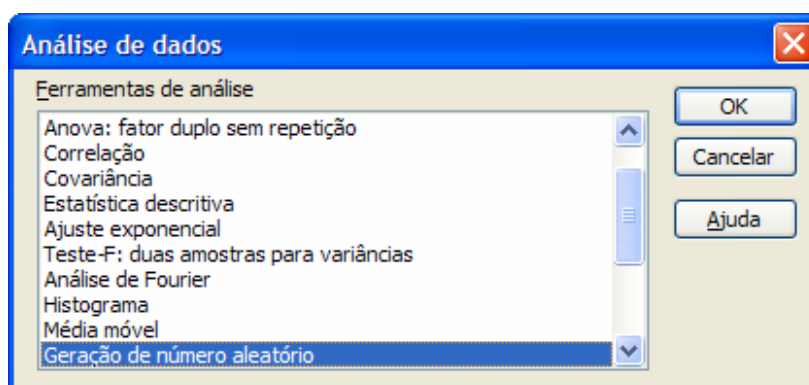


Figura 03 – A biblioteca "Ferramentas de Análise"

Esta biblioteca apresenta recursos para se trabalhar com boa parte da Estatística Básica. A desvantagem desta ferramenta é que os resultados obtidos de suas funções e procedimentos não são dinâmicos, isto é, atualizados automaticamente. Assim se o conjunto de dados de entrada for alterado o resultado não muda automaticamente como ocorre com as funções do subitem "Função" do item de menu "Inserir". Essa desvantagem é apenas aparente, pois se tivéssemos feito um levantamento de dados e digitados esses valores na planilha eles certamente não teriam essa característica. Assim esse recurso

poderá ser utilizado com proveito para se poupar tempo e os riscos desnecessários da coleta ou experimentação com dados. São poucas as situações em que um professor poderá tomar uma classe inteira de alunos para sair a campo coletando dados. Também são poucas as instituições de ensino que dispõe de laboratórios onde se possa colocar uma turma, em geral numerosa, para realizar experimentações de forma a obter dados reais. Trabalhar com dados reais é o ideal, mas nem sempre o ideal é possível, assim a geração ou obtenção de dados com o recurso do computador, além de praticamente não ter custo, pode gerar valores que estejam bastante próximos dos reais.

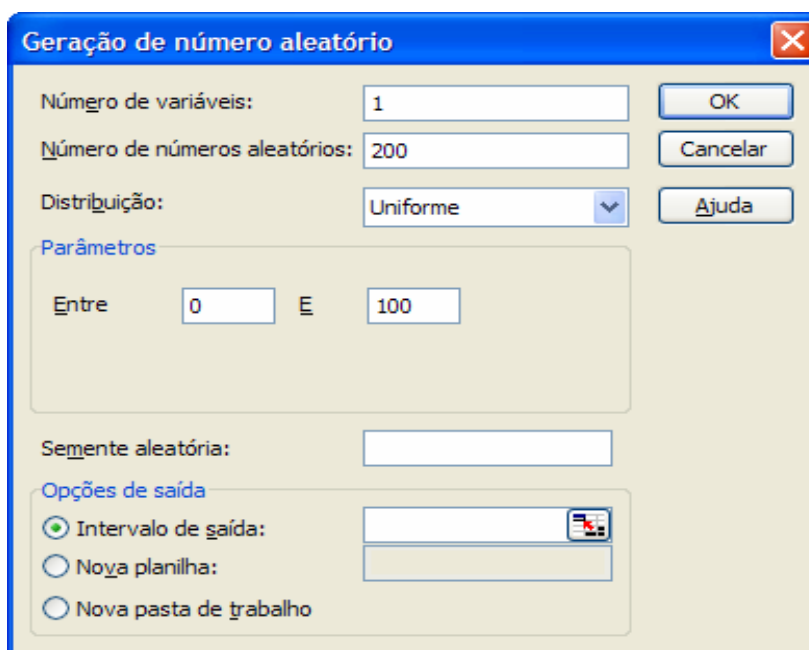


Figura 04 – O recurso "Geração de número aleatório" do menu "Ferramentas"

Na Figura quatro, a terceira linha, denominada de "Distribuição", apresenta um menu em cascata que fornece sete opções de simulações de distribuições. Destas, três são variáveis discretas e incluem a "Bernoulli", a "Binomial" e a "Poisson". Existe uma quarta opção denominada "Discreta" que não é um modelo probabilístico, mas uma opção que permite simular uma distribuição empírica fornecida pelo usuário. Além do recurso empírico e dos três modelos discretos existem ainda duas opções de variáveis contínuas sendo a primeira a uniforme e a segunda a distribuição normal. A opção "Padronizada" está colocada aqui por algum equívoco, pois ela de fato gera seqüências previsíveis de valores e não aleatórios como as demais.

Convém esclarecer que o termo aleatório não pode ser tomado ao pé da letra, quando se utiliza um recurso computacional uma vez que os valores gerados são feitos por algoritmos que são determinísticos e, portanto previsíveis. Alguns autores recomendam se falar em valores pseudo-aleatórios para os gerados dessa forma alegando que valores

verdadeiramente aleatórios só podem ser obtidos através de componentes mecânicos tipo uma roleta ou um bingo. Aqui vamos seguir a maioria dos autores que falam em “valores aleatórios” mesmo quando eles foram obtidos por via computacional.

A opção “Discreta” é apropriada para simular problemas reais cujo comportamento deva ser observado ao longo de um determinado tempo que geralmente o professor não dispõe. Aqui o professor poderá solicitar que cada aluno observe, por exemplo, uma determinada variável ao longo de um, dois ou mesmo três meses, anotando os resultados individualmente de forma a obter a distribuição empírica da variável. Por exemplo, o número de carros que passam por determinado cruzamento, o número de alunos que comparecem a aula a cada encontro, o número de dias com chuva ou nublados no período, etc. Aqui é possível fazer integração com outras disciplinas como Português observando o tamanho das palavras (número de letras) de um texto. Com a Biologia ou Ciências observando o tamanho de determinado inseto ou mesmo o número de insetos em determinada área. Com a Física avaliando temperaturas e outras grandezas. O professor de Educação Física pode calcular o número médio de acertos a uma cesta de basquete ou o número de chutes convertidas os gols.

Esses exemplos podem ser obtidos pelos alunos, anotados e depois tratados em aula com ou sem o recurso da planilha. Se a planilha for utilizada haverá a vantagem de não ser necessário se preocupar com o tamanho do conjunto, isto é, com o a quantidade de valores coletados de cada variável. Já se o tratamento for manual ou com o recurso da calculadora então será necessário limitar o número de observações tornando, o exemplo, talvez irrealista. Com conjuntos grandes o trabalho se tornará penoso e a adesão e entusiasmo que se possa conseguir com tal abordagem será desperdiçada na hora de se tratar os dados fazendo com que os alunos trabalhem horas sobre cálculos que uma máquina poderia executar em alguns segundos.

Os exemplos com poucas anotações ou resultados poderão ser utilizados como base empírica da distribuição discreta para gerar outros valores. Nesse caso será possível gerar quantos forem desejados. É conveniente gerar alguns milhares para mostrar que sem o recurso computacional não se teria como manejar esses dados. Mas que, agora, cada aluno pode fazer o trabalho que tomava o tempo de uma equipe inteira a poucos anos passados.

O número limitado de opções para gerar variáveis contínuas pode ser um problema. No entanto, a uniforme pode ser utilizada para gerar qualquer outro modelo



desejado. O modelo normal pode ser utilizado para representar muitas variáveis reais. Por exemplo, a altura, o peso, tamanhos de espécies, etc.

## 2. A ABORDAGEM

Um exemplo que tem sido utilizado com sucesso em vários cursos da área da área de exatas é a simulação de viagem entre dois locais. Por exemplo, de A para B numa distância de digamos 100 km. A velocidade dos carros, variável contínua, pode ser simulada pelo modelo normal. Pode ser utilizado o valor médio de 100 km com desvio padrão de 5 km, ou qualquer outro valor que for julgado adequado. O número de valores gerado pode ser escolhido convenientemente. Contudo convém utilizar 200 ou mais para que o formato do modelo obtido possa ficar aparente quando a representação gráfica for feita.

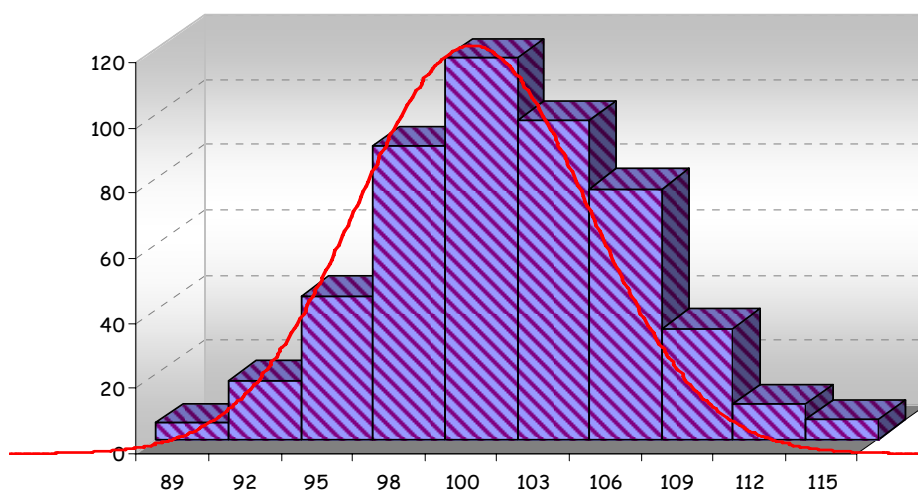


Figura 05 – Valores gerados e o modelo normal

O número de carros, variável discreta, que passam num determinado local, por exemplo, num posto de pedágio, pode ser simulada pela Poisson. Dependendo do trecho de estrada sendo considerado os valores dos parâmetros podem ser escolhidos em diálogo com os alunos. Pode-se estabelecer uma discussão produtiva sobre que taxa média de carros seria a mais adequada.

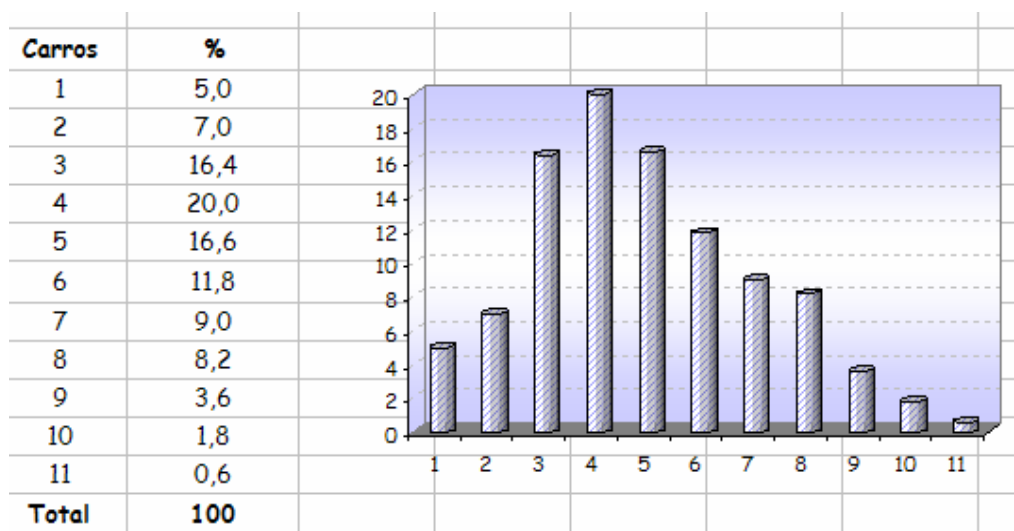


Figura 06 – Representação tabular e gráfica da variável “número de carros”

O sexo do motorista (homem, mulher) é outra situação que pode ser simulada com o auxílio do modelo Bernoulli. Aqui também pode-se fazer um levantamento sobre a opinião dos alunos sobre o que eles considerariam um percentual adequado de motoristas mulheres (ou homens). Esse tipo de variável é adequado para exemplificar a convergência média para o modelo normal. Pode-se, por exemplo, gerar certo número de valores dicotômicos, representá-los por “zero” (homem) e “um” (mulher) e agrupá-los em grupos de dois ou mais calculando a média de cada grupo que nesse caso será o percentual de homens (zeros) ou mulheres (uns) e representar esses resultados graficamente.

Para completar o leque de modelos é conveniente gerar pelo menos uma variável contínua além do modelo normal. Existe uma tendência nos textos didáticos de valorizar em demasia o modelo normal e o aluno acaba tendo a idéia de que ele é sempre adequado ou que é o único modelo existente. Assim uma opção é simular a variável “distância percorrida pelos carros a partir da cidade de origem” (A, por exemplo). Nesse caso, é conveniente a utilização do modelo uniforme. Carros que ainda não deixaram a cidade teriam distância percorrida  $a = \text{zero}$  e carros que tivessem chegado ao destino (cidade B) teriam distância percorrida igual a  $b = \text{distância real entre as duas cidades}$ . Aqui o conjunto de valores gerados vai ter uma representação gráfica plana, distinta da variável “velocidade”, representada pelo modelo normal que terá uma representação conforme figura seis.

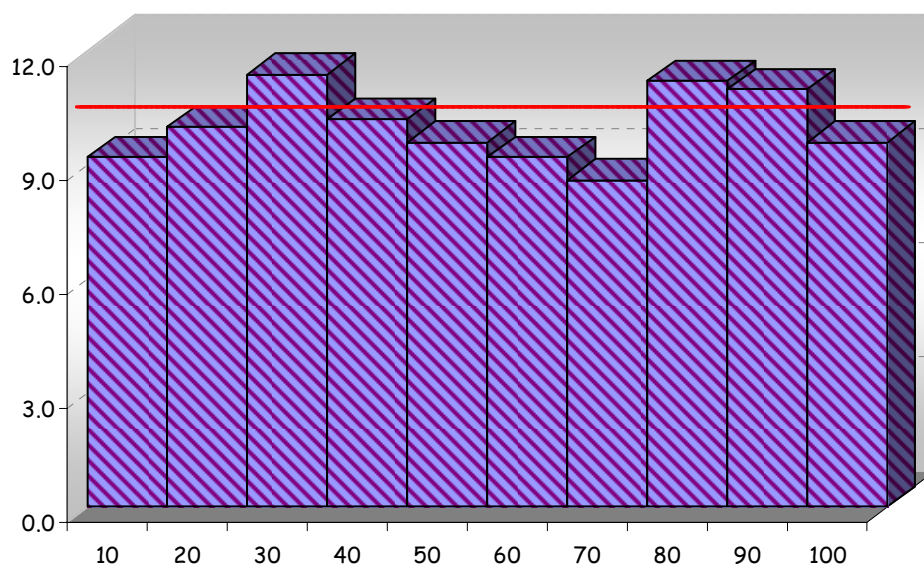


Figura 07 – Valores gerados e o modelo Uniforme

A maior vantagem nesse tipo de abordagem sobre a tradicional é que os alunos podem perceber, inferir ou mesmo aplicar outros tipos de conhecimento além dos puramente descritivos. Descrever dados é um objetivo em si, mas não é tudo. O que é valioso é o poder de se fazer inferências a partir dos dados sendo analisados. Se a análise e descrição dos dados forem às únicas tarefas não se estará indo além do que a máquina pode oferecer. A máquina é apenas um facilitador para o poder de análise, síntese e abstração que só o ser humano pode fazer. Do agrupamento dos dados é possível e desejável e altamente recomendável se fazer a representação gráfica. E a partir dela pode-se abstrair e partir para o mundo da imaginação, dos modelos, da teoria. E é só com teorias que nosso conhecimento avança. É modelando, estabelecendo relações, construindo explicações e fazendo abstrações sobre os dados e fatos observados de modo que eles façam sentido que se gera conhecimento novo. Sem os modelos e as teorias nosso conhecimento continuará no nível primário que foram obtidos, isto é, continuará ao nível da máquina e não se terá feito avanços.

### 3. CONCLUSÃO

Considerando que as reformas curriculares concentraram conteúdos em menos horas semanais torna-se necessárias novas abordagens para o ensino de Estatística. Nesse cenário não é possível deixar de considerar o computador como um aliado que pode servir como um recurso para o aumento da produtividade do ensino/aprendizagem com qualidade.

A planilha pode ser utilizada com proveito para se trabalhar com praticamente todos os conteúdos de estatística univariada e alguns tópicos de multivariada. Nesse

trabalho foram apresentadas algumas estratégias para se tirar proveito deste recurso computacional no ensino de Estatística Descritiva.

Apesar de o recurso ter completado um quarto de século e estar se popularizando no ensino de Matemática em todos os níveis, nota-se ainda que sua utilização no ensino de Estatística ainda é tímido. Basta olharmos qualquer periódico ou congresso para perceber isso.

O leque de recursos é expressivo na maioria dos pacotes estatísticos; no entanto, o que torna um pacote adequado ao ensino não é o que ela faz, mas sim como o faz. O aluno já tem dificuldades com o conteúdo em si e forçá-lo a aprender a manejar um software que ele provavelmente nunca mais irá utilizar, não é muito animador. A planilha com sua simplicidade e universalidade pode cumprir com êxito esse papel mesmo porque certamente ela fará parte da vida profissional da grande maioria dos estudantes (VIALI, 2004, p. 394).

Os PCNs já estão fazendo aniversário de uma década de publicação. Entretanto pouco ou quase nada tem sido visto de sua efetiva utilização no ensino de Estatística e Probabilidade. A análise de dados e o tratamento da incerteza continuam apenas como metas distantes, pois ao invés de avanços nos currículos universitários com o acréscimo desse tipo de conhecimento tem-se visto exatamente o contrário. Estamos caminhando mais uma vez na contramão das necessidades de formação de uma sociedade mais educada numericamente e digitalmente. O ensino continua sendo o último a refletir as mudanças sociais quando deveria ser o primeiro. Gandhi (2006, p. 4) conclui seu breve artigo declarando: “aprender a analisar dados e informação é uma tarefa inadiável para que o aluno de hoje prospere como profissional e como cidadão responsável.”

A insistência em submeter os alunos a grandes quantidades de cálculos manuais ou exercícios que pouco se aproximam da realidade faz que o já pouco interesse que eles possam ter pela disciplina seja rapidamente perdido contribuindo para que o ensino seja apenas ensino sem nenhuma ou quase nenhuma aprendizagem.

## REFERÊNCIAS

- GANDHI, Bodapati V. R. **Improving the Teaching of Statistics in Business Education: Lessons and Reflexions**. *ICOTS 7*, Salvador, Bahia, Brazil 2006.
- HALL, A. G. **A workshop approach using spreadsheets for the teaching of statistics and probability**. *Computers in Education*. v. 25, n. 1/2, p. 5-12, 1995.
- LEVY, Martin S., EVANS, James R., ZIRGER, B. J. **Web Enhanced Learning of Statistical Computational Tools**. OLN and ITEC-Ohio Present "The Convergence of Learning and Technology" - Windows on the Future 2003. Easton. March 3-4, 2003.
- MICROSOFT Office Excel 2003. Microsoft Corporation, 2003.
- VIALI, Lorí. **Utilizando Recursos Computacionais (Planilhas) no Ensino de Cálculo de Probabilidades (capítulo 13)**. In: *Disciplinas Matemáticas em Cursos Superiores: Reflexões, Relatos, Propostas*. Helena Noronha Cury (org.). Porto Alegre: EDIPUCRS. P. 351-95, 2004.