



UMA AVALIAÇÃO DO RECURSO PLANILHA PARA O ENSINO DE PROBABILIDADE

Lori Viali, Dr. (*)

Pontifícia Universidade Católica do Rio Grande do Sul

Av. Ipiranga, 6681 – 90619-900 Porto Alegre – RS

Fone (51) 320 3531 – Fax (51) 320 3631

viali@mat.pucrs.br

Sítio: <http://www.mat.pucrs.br/~viali/>

1. INTRODUÇÃO

As planilhas, notadamente o Excel, vão se firmando cada vez mais como um recurso instrucional em laboratórios de Estatística. Além dos recursos típicos elas oferecem um grande número de funções estatísticas e probabilísticas, testes de hipóteses, se bem que bastante limitados. Duas vantagens da planilha são a sua grande base instalada e seu preço relativamente barato, aliado ao fato de ser possível programá-la e, desta forma, realizar tarefas não previstas inicialmente. Além disso, o paradigma da planilha é conhecido por boa parte dos alunos, desta forma, diminuindo o tempo gasto na aprendizagem da mecânica de uma nova ferramenta de software. Isto é verdadeiro, é óbvio, se estivermos nos referindo ao mesmo software. Só que com o leque de opções oferecido hoje o problema poderá ser a diversidade de diferentes softwares conhecidos pelos alunos. Felizmente ou infelizmente, nesta área, parece que existe uma tendência, quase que universal, em favor do software da Microsoft a planilha Excel. Isto é confirmado pelo grande número de empresas e mesmo pessoas isoladamente que escrevem extensões para este programa. Pelos artigos na literatura técnica, pelos exemplos e exercícios disponíveis na Internet e até pelas listas de discussões, exclusivamente dedicadas a este software.

(*) Prof. Adjunto do Departamento de Estatística da Universidade Federal do Rio Grande do Sul

A grande vantagem da planilha é que ela é programável e dinâmica. Só que, como, ela não foi projetada para servir como recurso pedagógico e, especificamente para lecionar estatística, torna-se necessário fazer adaptações e estas adaptações não são tão simples. Esta dificuldade pode ser atestada pelo grande número de adendos (programas auxiliares) construídos com a finalidade de torná-la adequada ao ensino e que objetivam suprir carências de técnicas e procedimentos que ela não apresenta originalmente.

O principal objetivo deste trabalho é fornecer um panorama dos recursos probabilísticos que a planilha apresenta, mostrando as dificuldades que serão enfrentadas por professores e alunos que tenham a intenção de utilizá-la como recurso didático para o ensino de estatística e principalmente de probabilidade. Como consequência são fornecidas opções para corrigir os vários problemas de nomenclatura e terminologia na atribuição de nomes de funções e parâmetros destas funções.

É feita uma análise criteriosa das explicações nas janelas de edição das funções probabilísticas e na ajuda on-line destas mesmas funções apontando-se os erros e os conceitos confusos e mostrando uma opção para a melhoria destas explicações.

2. HISTÓRICO

A idéia da planilha é debitada a Dan Bricklin, um aluno do programa de pós-graduação de Harvard que precisava preparar uma planilha para um estudo de caso e tinha duas alternativas: fazê-lo manualmente ou usar um programa de mainframe pouco apropriado. Bricklin imaginou que deveria existir uma opção melhor e criou um programa em que as pessoas podiam visualizar uma planilha assim que fosse sendo criada. Sua metáfora foi a de um quadro-negro eletrônico e um giz eletrônico em uma aula. No verão de 1978 ele tinha programado a primeira versão operacional do seu conceito que era uma matriz de entrada com cinco colunas e vinte linhas. Como esta sua primeira versão não era muito amigável ele solicitou a um colega que expandisse e melhorasse o programa. No verão ele se juntou a outro colega e os três formaram a empresa Software Art, que em abril de 1979 começou a comercializar a *VisiCalc (Visible Calculator)*. O software fez sucesso quase imediato e vendeu aproximadamente um milhão de cópias durante seu tempo de vida e incentivou muitas pessoas da área de negócios a comprar um computador pessoal (POWER, 1998).

A planilha incentivou o desenvolvimento do microcomputador, pois antes do seu lançamento não existia um programa que justificasse realmente a compra de um micro. Em 1979 a planilha tornou-se disponível para o público através da plataforma Apple II e o impacto foi quase imediato, pois muitas pessoas compravam o micro apenas para usar este software.

A decisão de Bricklin de não patentear a VisiCalc serviu por um lado para incentivar e melhorar o desenvolvimento do software e por outro lado mostrou-se um desastre para sua empresa. A criação da Lotus 1-2-3 com mais recursos e com uma interface mais amigável provocou a queda nas vendas da VisiCalc que chegou a ter um faturamento, em 1982, de 11 milhões de dólares e 125 empregados, mas falhou no melhoramento do software e foi ultrapassada pela concorrente. Em 1984 as vendas da planilha tinham caído para apenas 2500 unidades por mês e a empresa acabou sendo vendida em 1985 para a Lotus Software.

3. A PLANILHA EXCEL

A liderança, na área, é exercida atualmente pela planilha Excel da empresa Microsoft. Esta planilha foi lançada em 1987 numa versão desenvolvida originalmente para os computadores Macintosh. A primeira versão para Windows foi rotulada como dois para corresponder à versão para Mac. Em 1990 foi lançada a versão três que incluía barras de ferramentas, capacidade para desenhar, suportes para programas adicionais e gráficos em 3 dimensões. A versão quatro, a primeira realmente popular, foi lançada em 1992. Em 1993 foi lançada a versão cinco com melhorias expressivas como planilhas múltiplas e suporte para a linguagem Visual Basic. Em 1995 foi lançada a versão sete da planilha, conhecida como Excel 95 e a primeira versão em 32 bits. A versão seis da planilha não existe, pois a Microsoft resolveu remunerar seus produtos para escritório de modo que todos eles tivessem a mesma versão. A versão oito foi lançada em 1997 e ficou conhecida como Excel 97 e apresentou uma nova interface para o desenvolvimento de aplicações em Visual Basic. A versão nove foi lançada em 2000. A partir desta versão as alterações foram apenas cosméticas, como ocorreu com a versão para o sistema operacional Windows XP (MICROSOFT, 2003).

4. OS RECURSOS PROBABILÍSTICOS

O Excel pode ser utilizado no ensino de Probabilidade, no entanto, este software apresenta uma série de deficiências que não inviabilizam seu uso, mas

comprometem e reduzem seu desempenho como recurso instrucional. Os problemas variam e incluem uma tradução mal feita e uma estrutura espaguete dos conceitos estatísticos e probabilísticos. Além disso, existe ainda falta de uniformidade de notação das funções e seus parâmetros, omissão de funções importantes, a forma desestruturada em que os parâmetros das funções são considerados e por fim a redação confusa e muitas vezes errada mesmo das explicações de conceitos e modelos (discretos e contínuos) probabilísticos apresentados.

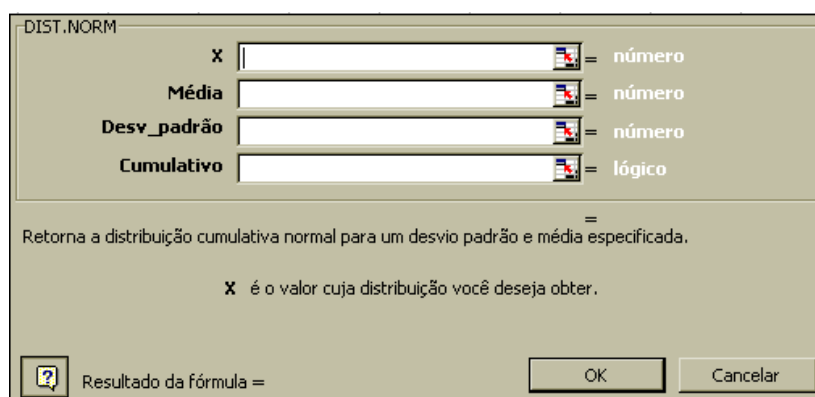
Para ilustrar o que estamos mencionando acima, considere-se a figura 01 que exemplifica um erro e uma meia verdade. Observe-se que na explicação da função está escrito: *retorna a distribuição cumulativa normal para um desvio padrão e média especificada*. Na realidade, isto só vai ocorrer se o parâmetro "Cumulativo" assumir o valor "verdadeiro" ou "um". Caso contrário, o retorno será o valor da função densidade de probabilidade (e não da função acumulada) no ponto "x", isto é, $y = f(x)$. Além disso, o que é retornado é um valor da função de distribuição $F(x)$ e não a "distribuição", que é o conjunto dos pontos $(x, f(x))$. Logo abaixo existe mais uma frase que agora contém um erro. Está escrito "**x** é o valor cuja distribuição você deseja obter". Isto é um equívoco, pois a distribuição é o conjunto de todos os valores $(x, f(x))$, então o correto seria dizer "**x** é o valor em que a função vai ser calculada (se cumulativo = 0) ou cuja área à esquerda será avaliada (se cumulativo = 1)". Obter a distribuição de uma variável aleatória é de fato determinar, a partir de um experimento (fenômeno), o conjunto de todos os valores que a variável pode assumir bem como suas respectivas probabilidades ou densidades.

Essencialmente os recursos da planilha, que são de interesse para o ensino de probabilidade estão limitados ao cálculo de probabilidades ou densidades de valores de variáveis aleatórias discretas e contínuas ou ainda ao cálculo de funções de distribuição (integrais) no caso das variáveis aleatórias contínuas. Num curso de probabilidade básica são tradicionalmente lecionadas algumas distribuições de probabilidades discretas (Bernoulli, Binomial, Hipergeométrica, Geométrica e Poisson) e outras tantas contínuas (Uniforme, Exponencial, Normal, Weibull, Gama, Beta, Student, Qui-quadrado, Snedcor). Algumas destas funções não são integráveis e outras apresentam expressões que tornam o cálculo manual, ou mesmo através de calculadoras, bastante trabalhoso ou tedioso. Desta forma a planilha apresenta vantagens sobre métodos manuais (que devem se valer de tabelas) ou mesmo calculadoras que ou não apresentam

as funções ou dependem de bibliotecas que exigem um conhecimento sofisticado para serem utilizadas.

Figura 01 - Caixa de diálogo para avaliar valores da distribuição normal

As distribuições acima não estão listadas de forma única e separadas como seria de se esperar. Elas estão misturadas no recurso "colar função" com as demais funções estatísticas e alguns casos dar-se conta de que elas existem não é tão simples em virtude da estrutura desordenada da listagem destas funções. A figura 02 ilustra a janela "colar função" e o início da lista destas funções.



5. AS DISTRIBUIÇÕES DISCRETAS

Das variáveis aleatórias discretas normalmente utilizadas em um curso de probabilidade básica a planilha Excel apresenta as seguintes: distribuição binomial negativa ou de Pascal - DIST.BIN.NEG() - distribuição Hipergeométrica - DIST.HIPERGEOM() - a distribuição binomial - DISTRBINOM() - e a distribuição de Poisson - POISSON(). Repare aqui o problema com a notação destas funções. A menos que o usuário tenha sido alertado ou tenha alguma intuição ele dificilmente perceberia que existe uma POISSON() e não uma DIST.POISSON() como seria o esperado. É comum que os alunos reclamem que não encontraram esta função ao tentar resolver os exercícios propostos.

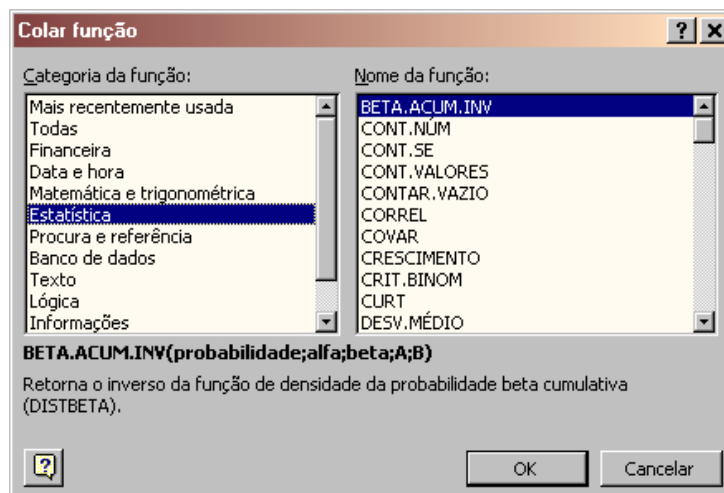


Figura 02 – A janela “colar função” com a categoria “funções estatísticas”

Este não é o problema mais sério aqui. Aparentemente não existe nenhuma lógica na atribuição do parâmetro cumulativo a estas distribuições. As funções Binomial e Poisson apresentam e as outras duas não. Também não existe uma uniformidade na atribuição de parâmetros e nomes a eles. Existe uma mistura de terminologias probabilísticas com estatísticas e outras peculiares exclusivamente a planilha. É um problema adicional para alunos, que já apresentam dificuldades de lidar com a notação matemática tradicional, se defrontarem com atribuições de nomes e símbolos totalmente esdrúxulos e fora de contexto a cada uma das funções da planilha.

Tabela 01 - Funções de probabilidade (discretas) apresentadas na planilha Excel

Função	Notação Excel	Ajuda - Funções estatísticas
Binomial Negativa ou Pascal	DIST.BIN.NEG	Retorna a distribuição binomial negativa.
Hipergeométrica	DIST.HIPERGEOM	Retorna a distribuição Hipergeométrica
Binomial	DISTRBINOM	Retorna a probabilidade da distribuição binomial do termo individual
Poisson	POISSON	Retorna a distribuição Poisson.

Como ilustração observe-se tabela um que relaciona as quatro distribuições discretas que a planilha apresenta. Uma falta a ser notada é a da distribuição Geométrica. Note-se que na atribuição dos nomes na lista geral das funções, além do problema com a Poisson já mencionado, as duas primeiras são referidas como “DIST.” com ponto, mas a terceira é “DISTR” com o “R” no lugar do ponto. Já no “Ajuda” (terceira coluna da tabela um) três delas são explicadas como “retorna a distribuição”, mas para Binomial a explicação é diferente e colocado que “retorna a probabilidade da distribuição binomial do termo individual” confundindo mais do explicando. A rigor

nenhuma das explicações do ajuda são, de fato, de grande valia, pois não são tecnicamente corretas, conforme já observado.

Como um exemplo da falta de critérios na atribuição da simbologia e terminologia, observe-se a última linha da tabela dois, que considera a distribuição de Poisson. Note-se que na janela de execução está escrito "Retorna a distribuição Poisson - X é o número de eventos". A primeira parte da frase já foi comentada. Vai-se comentar a declaração "X é o número de eventos". Esta terminologia está equivocada e em desacordo com a bibliografia. O correto aqui seria "x" (minúsculo) e não "X" maiúsculo, pois se trata de um valor da variável e não da variável em si. Segundo "x" é um valor da variável e não "o número de eventos", terminologia que é utilizada com outro propósito. O uso do valor "x" minúsculo deveria ser utilizado para todas as funções, mas na Binomial já aparece uma nova nomenclatura. É criado agora o "Núm_s" que recebe a explicação "é o número de tentativas bem sucedidas". Isto até pode ser verdade, mas é uma situação específica e não uma notação tradicional e consistente.

Tabela 02 - Funções de probabilidade (discretas) apresentadas na planilha Excel

Função	Explicação na janela de execução
Binomial Negativa	Retorna a distribuição binomial negativa, a probabilidade de que ocorrerá 'Núm_f' de falhas antes de 'Núm_s' de sucessos' com probabilidade 'Probabilidade_s' de um sucesso - Núm_f é o número de falhas.
Hipergeométrica	Retorna a distribuição Hipergeométrica - Exemplo_s é o número de sucessos em uma amostra.
Binomial	Retorna a probabilidade da distribuição binomial do termo individual - Núm_s é o número de tentativas bem sucedidas.
Poisson	Retorna a distribuição Poisson - X é o número de eventos.

6. AS DISTRIBUIÇÕES CONTÍNUAS

As distribuições probabilísticas contínuas são em número de dez se considerarmos que a distribuição normal padrão seja uma destas distribuições. As seguintes funções densidade de probabilidade são consideradas: Beta, Exponencial, F, Gama, Log-normal, Normal, Normal Padrão, t, χ^2 , e Weibull, conforme tabela três.

Quatro destas funções seguem o esquema da figura um, fornecendo um parâmetro lógico, denominado de "Cumulativo" que se for "falso ou 0", a função retornará para o valor (abscissa) x o valor $y = f(x)$, isto é, a ordenada, ou ainda a densidade no ponto "x". Se o valor do parâmetro "Cumulativo" for "verdadeiro ou 1", o

retorno deverá ser a $F(x) = \int_{-\infty}^x f(u)du$, ou a área abaixo (à esquerda) do ponto x , ou ainda, a Função de Distribuição Acumulada no ponto " x ". Esta forma de apresentação é bastante conveniente, pois com a mesma função é possível obter o desenho (fazer o diagrama) da curva se a opção "Cumulativo" for "0" e obter as probabilidades (áreas sob a curva) se a opção "Cumulativo" for "1". Além de uniformizar e padronizar todas as distribuições torna-se mais didático, pois facilita o entendimento. O problema é que o Excel não segue esta linha, parecendo que fez questão de confundir o usuário, através da adoção de uma opção diferente para cada situação.

Tabela 03 - Funções densidades contínuas apresentadas na planilha Excel

Nome	Notação Excel	Ajuda - Funções Estatísticas
Beta	DISTBETA	Retorna a função densidade da probabilidade beta acumulativa.
Exponencial	DISTEXPON	Retorna a distribuição exponencial.
Gama	DISTGAMA	Retorna a distribuição gama.
Log-Normal	DIST.LOGNORMAL	Retorna a distribuição log-normal cumulativa.
Normal	DIST.NORM	Retorna a distribuição cumulativa normal.
Normal Padrão	DIST.NORMP	Retorna a distribuição cumulativa normal padrão.
Qui-Quadrado	DIST.QUI	Retorna probabilidade uni-caudal da distribuição qui-quadrada.
Snedekor	DISTF	Retorna a distribuição da probabilidade F.
Student	DISTT	Retorna a distribuição t de Student.
Weibull	WEIBULL	Retorna a distribuição de Weibull.

Já na atribuição ou tradução do nome das funções nota-se o pouco caso com a nomenclatura, pois enquanto quase todas as funções são precedidas pela inicial DIST (de distribuição) a Weibull é referenciada pelo nome, assim se o usuário não for informado deste fato, dificilmente ele desconfiará da existência desta função, pois ela aparece no final da lista, já que todas as funções tanto estatísticas quanto probabilísticas estão misturadas. Além disso nove delas seguem o esquema Dist + Início do Nome, só que algumas possuem um ponto separando os dois itens enquanto outras não. Isto pode parecer irrelevante a princípio, mas é sem dúvida um ponto desfavorável sobre um software que apresenta uma notação consistente.

As demais funções, também, não seguem um padrão. A distribuição normal reduzida só retorna a função acumulada (área à esquerda de " x "), isto é, $\Phi(z)$. A distribuição qui-quadrado retorna $1 - F(x)$, isto é, a área à direita de " x ". Só que na função é informado que: *retorna a probabilidade uni-caudal da distribuição qui-quadrada*, sem especificar qual das caudas "direita" ou "esquerda". Se o usuário seguir a lógica da distribuição

normal concluirá que o valor retornado é $F(x)$, quando, de fato, o que é retornado é $1 - F(x)$.

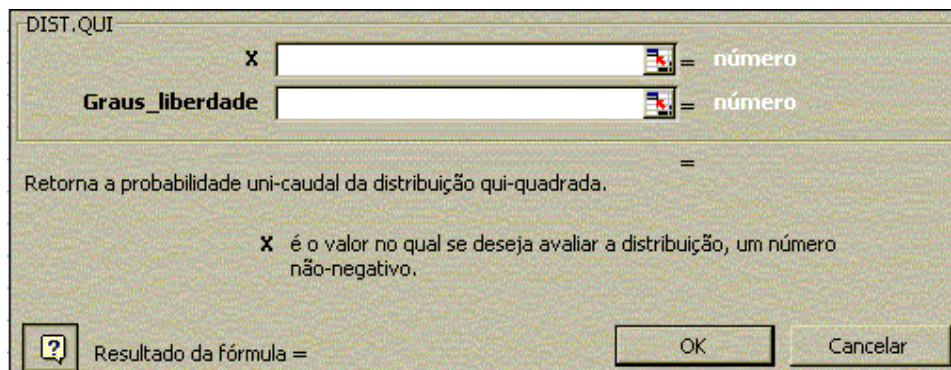


Figura 03 - Caixa de diálogo para avaliar valores da curva qui-quadrado

A distribuição F segue o mesmo caminho, tendo como retorno o valor da cauda da direita, isto é, $1 - F(x)$. Neste caso a informação contida na janela de edição da função é um pouco mais imaginativa, pois informa que: *retorna a distribuição de probabilidade F (grau de diversidade/variedade) para dois conjuntos de dados*. Aqui a liberalidade com a informalidade foi exacerbada. Retornar uma distribuição já seria um abuso de linguagem. O problema é que a explicação ficou ininteligível e é difícil imaginar o que de fato pretendiam dizer. Pode-se apenas acrescentar que o que deveria ser escrito é que o retorno é " $1 - F(x)$ " ou simplesmente a área à direita de " x ", fornecidos os parâmetros (da distribuição F , não da função do Excel F): "grau de liberdade do numerador" (aqui fantasiosamente chamado de "grau de diversidade" e "grau de liberdade do denominador" apelidado de "grau de variedade"). Se alguma coisa dá para inferir da pseudo-explicação dada é que confundiram "a necessidade de especificar dois parâmetros para a distribuição F " com o "retorno da distribuição F para dois conjuntos de dados"

Se o usuário ficar confuso e recorrer à ajuda (*help*) não vai ter muito sucesso, pois vai encontrar que "DISTF é calculada como $DISTF = P(F < x)$, onde F é uma variável randômica que possui uma distribuição F ", ou seja, está sendo colocado que a área fornecida é $F(x) = P(X < x)$, quando de fato é " $1 - F(x) = P(X > x)$ ". Em resumo a "ajuda" ao invés de corrigir o erro anterior acaba reforçando-o.

Na distribuição " t " (de Student) o parâmetro "Cumulativo", didático e útil, é substituído pelo parâmetro "Caudas" com a explicação de que se "Caudas = 1" o retorno será "distribuição uni-caudal" e "Caudas = 2" o retorno será "distribuição bi-caudal".

Quer dizer, ao invés de o software fornecer a $y = f(x)$ que seria útil para desenhar a curva e ilustrar para os alunos o seu formato os programadores acharam mais útil fazer uma simples multiplicação por 2, pois, por exemplo, se o valor digitado for $x = 1,96$ e o grau de liberdade for, digamos 100 com o parâmetro “caudas = 1”, o retorno será $F(1,96) = 5\%$, se o parâmetro caudas for alterado para “caudas = 2” o retorno, mantendo os demais valores, será $2.F(1,96) = 10\%$, que é um claro desperdício de recursos, pois esta multiplicação por dois foi colocada no lugar do parâmetro “cumulativo” que permitiria traçar o diagrama da distribuição o que seria bem mais proveitoso

Tabela 04 - Funções densidades contínuas apresentadas na planilha Excel

Função	Explicação na janela de execução
Beta	Retorna a função densidade da probabilidade beta acumulativa - X é o valor entre A e B no qual se avalia a função.
Exponencial	Retorna a distribuição exponencial. Consulte Ajuda para as equações utilizadas - X é o valor da função, um número não-negativo.
Gama	Retorna a distribuição gama. Consulte Ajuda para as equações utilizadas - X é o valor na qual se deseja avaliar a distribuição, um número não-negativo.
Log-Normal	Retorna a distribuição log-normal cumulativa de x, onde $\ln(x)$ é normalmente distribuído com parâmetros 'Média' e 'Desv-padrão' - X é o valor no qual se avalia a função, um número positivo.
Normal	Retorna a distribuição cumulativa normal para um desvio padrão e média especificada - x é o valor cuja distribuição você deseja obter.
Normal Padrão	Retorna a distribuição cumulativa normal padrão (possui uma média zero e um desvio padrão 1) - Z é o valor cuja distribuição você deseja obter.
Qui-Quadrado	Retorna probabilidade uni-caudal da distribuição qui-quadrada - x é o valor na qual se deseja avaliar a distribuição, um número não-negativo.
Snedekor	Retorna a distribuição da probabilidade F (grau de diversidade/variedade) para dois conjuntos de dados - X é o valor da função, um número não-negativo.
Student	Retorna a distribuição t de Student - X é o valor numérico em que se avalia a distribuição.
Weibull	Retorna a distribuição Weibull. Consulte Ajuda para as equações utilizadas - X é o valor na qual se avalia a função, um número não-negativo.

Além disso, se for digitado, por exemplo, o valor -2, grau de liberdade igual a 30, e por exemplo, caudas = 1, o resultado deveria ser aproximadamente $0,0273 = 2,73\%$ que é a área à esquerda do ponto -2 na curva t. O que, de fato, é retornado, neste caso é uma mensagem de erro (#NÚM!). Percebe-se, então, que a planilha não fornece as áreas da distribuição para valores negativos da distribuição t, o que é além de incompreensível, antididático. O retorno da área para valores negativos serviria para ressaltar que a variável pode assumir valores negativos, já que as tabelas, por razões óbvias, apresentam apenas valores positivos.

A distribuição log-normal retorna $F(x) = P(X \leq x)$ apenas. Esta é uma das poucas que não apresenta erro na explicação, pelo menos não do mesmo vulto dos demais, pois é afirmado que a função "Retorna a distribuição log-normal cumulativa de x , onde $\ln(x)$ é normalmente distribuído com "Média" e "Des_padrão".

A distribuição Beta, definida apenas no intervalo $[0; 1]$, recebeu novos parâmetros (A e B), que por omissão assumem os valores $A = 0$ e $B = 1$, além dos parâmetros "alfa" e "beta" normais. Estes parâmetros foram acrescentados para permitir que a variável X varie num intervalo diferente de zero a um. O problema é que este procedimento é exclusivo da planilha, já que na literatura a distribuição é definida no intervalo $[0; 1]$.

7. AS INVERSAS DAS FUNÇÕES CONTÍNUAS

A planilha apresenta um total de dez distribuições contínuas (considerando a normal genérica e a normal padrão como duas distintas). Entretanto, apenas sete foram contempladas com algoritmos de inversão. A Beta, a Exponencial e a Weibull não mereceram algoritmos de inversão. As tabelas cinco e seis apresentam a lista das inversas das funções que podem ser encontradas na planilha. Nem todas as funções de probabilidade inversa tem a mesma importância na inferência, sendo que as principais são a Normal, a t, a Qui-quadrado e a F. No entanto por coerência seria justo esperar que todas as que foram incluídas na planilha tivessem também a sua inversa considerada. Não parece que o problema tenha sido por dificuldade de implementação de algoritmos, mas sim por puro descuido mesmo.

Tabela 05 - Funções inversas de variáveis contínuas apresentadas na planilha Excel

Nome	Notação Excel	Ajuda - Funções Estatísticas
F inversa	INVF	Retorna o inverso da distribuição da probabilidade F.
Gama inversa	INVGAMA	Retorna o inverso da distribuição cumulativa gama.
Log-normal inversa	INVLOG	Retorna o inverso de uma distribuição log-normal.
Normal Inversa	INV.NORM	Retorna o inverso da distribuição cumulativa normal.
Normal padrão inversa	INV.NORMP	Retorna o inverso da distribuição cumulativa normal padrão.
Qui-quadrado inversa	INV.QUI	Retorna o inverso da probabilidade uni-caudal da distribuição qui-quadrada.
T inversa	INVT	Retorna o inverso da distribuição t de Student.

Como pode ser observado na tabela das sete funções inversas disponíveis em três são mencionadas as distribuições acumuladas, isto é, $F(x)$, enquanto nas demais é

necessário supor que está sendo invertido é a $F(x)$. Assim para o primeiro caso $F(x) = P(X \leq x) = \alpha$, então o inverso desta situação seria $F^{-1}[F(x)] = F^{-1}[P(X \leq x)]$ ou $x = F^{-1}[P(X \leq x)] = F^{-1}(\alpha)$. Ou seja, dada uma probabilidade $\alpha = P(X \leq x)$ a função inversa fornece o valor de "x" que satisfaz esta condição. Assim se a função for a Normal padrão tem-se que $\Phi(1) = P(Z \leq 1) = 0,8413 = 84,13\%$, então $\Phi^{-1}[\Phi(1)] = \Phi^{-1}[P(Z \leq 1)] = \Phi^{-1}(0,8413) = 1$.

Isto é o que se esperaria que ocorresse. Entretanto, este nem sempre é o caso. Assim se quisermos o inverso da distribuição t, somente iremos obter o inverso de um valor absoluto, isto é, das áreas das duas caudas somadas, o que é um contra-senso. Isto é, a função $INVT(95\%, 1) = 0,0787$, não é invertida por $DISTT(0,0787; 1; \mathbf{1}) = 47,50\%$, mas apenas por $DISTT(0,0787; 1; \mathbf{2}) = 95\%$. Isto se diferencia de todas as demais operações de inversão que são sempre unilaterais. Imagine-se o efeito disto em usuários inexperientes, pois mesmo, alguém já com boa prática pode, facilmente, ser conduzido ao erro.

Tabela 06 - Funções inversas de variáveis contínuas apresentadas na planilha Excel

Nome	Explicação na janela de execução
F inversa	Retorna o inverso da distribuição de probabilidades F: se $p = DISTF(x, \dots)$, então $INVF(p, \dots) = x$.
Gama inversa	Retorna o inverso da distribuição cumulativa gama: se $p = DISTGAMA(x, \dots)$, então $INVGAMA(p, \dots) = x$.
Log-normal inversa	Retorna o inverso da distribuição log-normal cumulativa de x, onde $\ln(x)$ é normalmente distribuída com parâmetros Média e Dev_ padrão.
Normal Inversa	Retorna o inverso da distribuição cumulativa normal para a média e o desvio padrão especificados.
Normal padrão inversa	Retorna o inverso da distribuição cumulativa normal padrão (possui uma média zero e um desvio padrão 1).
Qui-quadrado inversa	Retorna o inverso da probabilidade uni-caudal da distribuição qui-quadrada.
T inversa	Retorna o inverso da distribuição t de Student.

Quanto à distribuição Qui-Quadrado, ela de fato apresenta uma operação de inversão, mas, aqui, o truque é diferente, pois a área fornecida, ao contrário da FD (à esquerda do valor x) é a da direita de x, também levando invariavelmente a confusões. O mesmo problema ocorre com a função F. Assim se realizarmos no Excel a seguinte operação: $INV.QUI(95\%; 1) = 0,0039$. Este valor (0,0039) não é como seria o esperado o valor que reparte a distribuição em 95% dos valores à esquerda e 5% à direita e sim o contrário. Para alguém experiente o erro fica evidente, mas para os alunos que

normalmente apresentam muitas dificuldades para entender estes conceitos, ele passará facilmente sem ser notado.

8. VIRTUDES DA PLANILHA EXCEL

A principal virtude da utilização da planilha Excel no ensino de Estatística está na interface bem conhecida pelos alunos e na sua grande base instalada, isto é, dificilmente algum computador que utilize o sistema operacional Windows deixará de ter instalado o pacote Office que inclui a planilha. Para lecionar estatística básica para a maioria dos cursos universitários os recursos existentes na planilha são suficientes. O fato de a planilha aceitar a digitação de fórmulas também ajuda, pois as omissões podem ser corrigidas em alguns casos. A facilidade de aprender a manuseá-la com razoável competência é outro ponto a favor.

O fato do paradigma da planilha ser conhecido por boa parte dos alunos dispensa, o professor de Probabilidade e de Estatística, do ônus de ter que ensinar um novo recurso computacional, o que de fato, não é o seu objetivo e geralmente nem isto é possível. Os conteúdos das disciplinas de Probabilidade e Estatística são extensos e o número de horas disponíveis não é suficiente para cobrir todo o conteúdo, então é quase impraticável se gastar uma ou duas semanas (se isto for suficiente) para ensinar um novo software (VIALI, 2002). Aqueles alunos que, ainda, não a conhecem, normalmente, não reagem negativamente ao ter que aprendê-la, pois sabem que, cedo ou tarde eles, terão que fazer isto por uma imposição do mercado de trabalho, o mesmo já não se daria com um software específico.

9. CONCLUSÃO

Lecionar Estatística, como de resto disciplinas da área de Matemática, é uma tarefa difícil por vários motivos. Geralmente estas disciplinas envolvem abstrações que não encontram suporte em recursos visuais. Os alunos apresentam, quando não uma aversão completa pela disciplina, um desinteresse que dificulta tanto o relacionamento professor aluno quanto a aprendizagem. Melhorar este quadro, tornando o ensino mais visual, mais atrativo e mais abrangente, envolvendo perspectivas globais pode ser feito. Uma das maneiras para se atingir este objetivo é mudar substancialmente a maneira de apresentar a disciplina de probabilidade e estatística, através da eliminação do trabalho manual improdutivo que pode ser obtido através do uso da planilha. O ganho de tempo

será utilizado para melhorar a apresentação e discussão dos conceitos e para reforçar pontos obscuros ou difíceis.

Obviamente as falhas apontadas aqui não desmerecem as muitas virtudes deste recurso computacional. Mas se fossem corrigidas ou minimizadas certamente tornariam o recurso não apenas mais prazeroso de ser utilizado com esta finalidade, mas o tornariam, ainda, mais produtivo e eficiente.

A planilha pela sua universalidade pode cumprir com êxito o papel de um bom recurso didático no ensino de estatística e principalmente de probabilidade. Para tanto, bastaria que os projetistas, tradutores e programadores envolvidos tivessem tido um pouco mais de cuidado com as funções específicas. Espera-se que com este trabalho possamos estar contribuindo para que isto venha a acontecer numa próxima versão.

Palavras-chave: O computador no ensino; ensino de probabilidade, ensino através de planilhas.

10. REFERÊNCIAS

MICROSOFT. *Microsoft Excel, Ajuda On-line*.

POWER, Daniel. *A Brief History of Spreadsheets*. [on-line]

<http://dss.cba.uni.edu/dss/sshistory.html>

VIALI, Lóri. Using Spreadsheets And Simulation To Enhance Teaching Probability And Statistics To Engineering Students. *ICEE 2002 (International Conference on Engineering Education)*. Manchester, England. August 18 to 22/2002.