

# ESTUDO COMPARATIVO DE GRÁFICOS DE PROBABILIDADE NORMAL PARA ANÁLISE DE EXPERIMENTOS FATORIAIS NÃO REPLICADOS

Manasses Pereira Nóbrega<sup>1,2</sup>

*e-mail:* manassespereira@uern.br

Carla Almeida Vivacqua<sup>2</sup>

*e-mail:* vivacqua@ccet.ufrn.br

<sup>1</sup>Universidade do Estado do Rio Grande do Norte, Departamento de Matemática e Estatística, CAJIM  
CEP 59770-000, Patu - RN - Brasil

<sup>2</sup>Universidade Federal do Rio Grande do Norte, Departamento de Estatística, CCET, PPGMAE  
CEP 59072-970, Natal - RN - Brasil

**Resumo.** Os experimentos fatoriais  $2^k$  não replicados são muito úteis em várias aplicações, especialmente na experimentação industrial. Uma das razões de seu tão vasto uso é a redução de custos associada a tais experimentos. No entanto, sem replicação, não é possível obter uma estimativa direta da variabilidade do erro. Um método comum para se avaliar a significância dos efeitos é utilizar os gráficos normal ou semi-normal. Muitos pesquisadores usam o gráfico normal, ao passo que outros preferem o gráfico semi-normal. A escolha entre estes dois gráficos parece ser uma questão subjetiva. Neste artigo, apresentamos um estudo realizado para comparar estas duas técnicas gráficas. Pretendemos verificar em que situações um gráfico poderia ser melhor que o outro. Usamos simulações e estudos de caso para avaliar as habilidades de ambos os gráficos em identificar efeitos significativos, detectar valores discrepantes e identificar parcelas subdivididas inadvertidas em experimentos fatoriais  $2^k$  não replicados. Mostramos que essas simulações podem prover orientações potencialmente úteis aos pesquisadores ao interpretarem os resultados de um experimento. Discutimos, também, as vantagens e limitações de cada procedimento.

**Palavras Chave:** fatoriais com dois níveis; gráfico normal; gráfico semi-normal; efeitos significativos; parcelas subdivididas; valores discrepantes.

**Abstract.** Unreplicated two-level factorial designs are very useful in many applications, especially in industrial experimentation. One of the reasons of their wide spread use is the cost savings associated with such designs. However, without replication, a direct estimation of error variance is not possible. One common method to assess the significance of effects is to use normal or half-normal probability plots. Some experimenters use the normal plot, while others prefer the half-normal plot. The choice between these two plots seems to be a subjective issue. In this paper we present a study carried out to compare these two graphical techniques. We intend to verify in what situations one plot could be better than the other. We use simulation and case studies to evaluate the abilities of both graphical techniques to identify significant effects, to detect outliers, and to identify inadvertent split-plotting in unreplicated two-level factorial designs. We show that these simulations can provide potentially useful insights to practitioners when interpreting results from an experiment. We also discuss the advantages and limitations of each procedure.

**Key Words:** two-level factorials; normal plot; half-normal plot; significant effects; split-plot; outliers.

# 1 Introdução

Experimentos fatoriais são muito utilizados, especialmente na indústria (OLGUIN; FEARN, 1997). Estes tipos de planos experimentais são mais vantajosos que experimentos com um fator de cada vez, visto que podem detectar e estimar interações entre os fatores. Por se tratarem de uma classe extremamente importante, os experimentos fatoriais  $2^k$ , caso em que todos os fatores possuem dois níveis, destacam-se dentre os planos fatoriais (MONTGOMERY, 2001). Estes planos permitem uma análise estatística bastante simplificada e formam a base para outros planos úteis, tais como os experimentos de triagem fatorial ou experimentos exploratórios (*factor screening experiments*).

Na prática, quanto maior o número de fatores considerados maior será a quantidade de provas necessárias para a execução de um experimento, tornando-o dispendioso. Além disso, realizar replicações dos tratamentos pode ser inviável, considerando as limitações de recursos e de tempo. Sendo assim, são utilizados os experimentos fatoriais  $2^k$  não replicados, que são experimentos com uma única unidade experimental para cada tratamento. Mas, sem replicação, não é possível obter uma estimativa direta da variabilidade do erro para julgar a significância dos efeitos. Neste caso, dois procedimentos que podem ajudar na avaliação da magnitude dos efeitos são as análises feitas a partir do gráfico de probabilidade normal (*full-normal plot*) ou do gráfico de probabilidade semi-normal (*half-normal plot*).

A idéia de um gráfico de probabilidade é comparar duas distribuições de probabilidade: a função de distribuição da amostra (também chamada de distribuição empírica) com uma função de distribuição teórica (também chamada de distribuição hipotetizada). Quando a função teórica escolhida é a distribuição normal, o gráfico é chamado de gráfico de probabilidade normal. Neste gráfico, os pares ordenados são representados pelos valores observados na amostra e seus respectivos quantis na distribuição normal. Se o conjunto de pontos marcados segue um padrão linear, então, há fortes evidências para se concluir que a distribuição normal serve como um modelo plausível para explicar a população que produziu a amostra. Nesse caso, o intercepto da reta e sua inclinação indicam, respectivamente, a média e o desvio padrão da população da qual a amostra foi obtida (JOHNSON; BHATTACHARYYA, 1992; UPTON; COOK, 2006). Embora sejam procedimentos estatísticos informais, uma vez que sua interpretação é visual e subjetiva, de acordo com Wilk e Gnanadesikan (1964, 1968), os gráficos de probabilidade normal têm o seu valor, pois são úteis para verificar a normalidade das amostras, detectar a existência de valores discrepantes nos dados, julgar a significância dos efeitos nos planos fatoriais, como também identificar restrições na aleatorização.

Em 1959, Cuthbert Daniel desenvolveu o que ele chamou de *half-normal plot* ou gráfico semi-normal, como uma nova ferramenta estatística para analisar os experimentos fatoriais  $2^k$  (completos e fracionados) não replicados. A idéia central deste gráfico é semelhante àquela apresentada para o gráfico normal: comparar a distribuição da amostra com uma distribuição teórica que, neste caso, é a distribuição semi-normal. Neste gráfico, os pares ordenados são representados pelos módulos dos valores observados na amostra e seus respectivos quantis na distribuição semi-normal. Sua construção, definição e propriedades são apresentadas em seu artigo de origem, publicado em novembro de 1959 na revista *Technometrics* (DANIEL, 1959). Segundo Daniel (1959), o gráfico semi-normal pode ser utilizado para identificar efeitos significativos, estimar a medida da variância do erro experimental, detectar a presença de valores discrepantes (*outliers*) nos dados e ainda identificar restrições na aleatorização, como nos experimentos em parcelas subdivididas (*split-plot*).

## 2 A Controvérsia

Apesar de ter sido o idealizador e grande apoiador do gráfico semi-normal para a análise de experimentos fatoriais não replicados, Cuthber Daniel acaba mudando de idéia. Com a publicação de seu livro *Applications of Statistics to Industrial Experimentation* em 1976, Daniel deixou clara sua preferência pelo gráfico normal: todos os gráficos utilizados no livro para as análises dos fatoriais não replicados são gráficos normais. O motivo de sua nova escolha é o fato de, segundo o próprio Daniel, o gráfico normal ser mais útil em identificar algumas importantes características, tais como a presença de valores discrepantes nos dados e sua influência no cálculo dos efeitos; a avaliação da significância dos efeitos quando estes forem pequenos ou em número considerável; e também a avaliação da violação da suposição de homocedasticidade. Daniel também percebeu que o gráfico semi-normal era insensível a tais características no sentido de esconder tais situações.

Esta mudança deu início a uma controvérsia sobre o uso dos gráficos de probabilidade normal na análise dos planos fatoriais não replicados. Ao passo que alguns pesquisadores optam pelo gráfico semi-normal, outros utilizam o gráfico normal. Um considerável aumento pela preferência do gráfico normal deve-se ao artigo de Box e Meyer (1986). Ali os autores utilizam este gráfico com os mesmos objetivos do gráfico semi-normal: avaliar a significância dos efeitos e estimar o erro em experimentos fatoriais não replicados. Este artigo ainda é citado hoje como referência e justificativa para o uso do gráfico normal nos experimentos fatoriais. Apesar disso, o gráfico semi-normal ainda é a ferramenta mais utilizada para a avaliação da significância dos efeitos nos arranjos fatoriais não replicados.

Como se já não bastasse a controvérsia criada pela mudança de opinião Daniel em 1976, Mark J. Anderson (ver [www.statease.com/news/faqalert6-09.html](http://www.statease.com/news/faqalert6-09.html)) afirma algo realmente intrigante: Daniel “*reversed himself back*” e veio a favorecer novamente o gráfico semi-normal! Embora não se tenha encontrado outra referência que comprovasse, ou não, esta segunda mudança de opinião de Daniel, esta situação levanta algumas questões: Qual dos dois procedimentos é o “melhor”? Qual deles usar? Em que situações? Com que justificativa? Com isso, a controvérsia havia começado. Além da subjetividade própria das técnicas gráficas, o uso de tais ferramentas também é subjetivo na maioria das vezes, dependendo apenas da preferência daquele que a usa.

Desta controvérsia surgiu a motivação deste trabalho. Ora, o fato de o próprio inventor do gráfico semi-normal ter mudado de idéia algumas vezes sobre o desempenho de seu método, por si só é um bom motivo para se analisar mais profundamente essa questão. Uma vez que não há registro de procedimento formal ou teste estatístico que indique “qual delas é melhor”, pretendemos contribuir com critérios mais objetivos, se possível, para ajudar na tomada de decisão sobre qual dos dois métodos gráficos escolher. Para isso, usamos simulações (estudo comparativo experimental) e estudos de caso para avaliar as habilidades de ambos os gráficos em identificar efeitos significativos, detectar valores discrepantes e identificar parcelas subdivididas inadvertidas (*inadvertent plot-splitting*) em experimentos fatoriais  $2^k$  não replicados.

Enfim, aos cinquenta anos de publicação do artigo de Daniel sobre o gráfico semi-normal, este trabalho busca verificar as seguintes questões: Um gráfico pode ser melhor que o outro? Em que situações? Que informações um gráfico acrescenta à análise do experimento que possam complementar aquelas fornecidas pelo outro gráfico? Quais as restrições no uso de cada gráfico? O conhecimento construído pelas respostas dadas a estas e a outras questões com certeza contribuirá para o avanço da Estatística como um todo, e em particular, para a área de Planejamento de Experimentos, que é a base desta pesquisa.

### 3 Metodologia

Inicialmente é realizado um extenso levantamento bibliográfico. O objetivo é encontrar artigos nos quais o gráfico normal ou o semi-normal tenham sido utilizados para a análise de experimentos, a fim de delinear ou quantificar o uso desses dois gráficos nos trabalhos científicos, como também verificar quais as justificativas apresentadas para o uso dos mesmos. Uma vez terminado esse trabalho de pesquisa bibliográfica, situações experimentais (reais e simuladas) de experimentos fatoriais não replicados são construídas e, nestas, os dois métodos gráficos são utilizados e comparados, com o objetivo de analisar algumas características de seu desempenho e comportamento. Inicialmente, os cenários realizados são escolhidos dentre os artigos pesquisados. Nesta fase inicial, os critérios para a comparação dos gráficos são as percepções do analista, sua sensibilidade de captar características distintas nos gráficos por meio da análise visual dos mesmos, tendo por base os dados e os resultados descritos nos artigos de origem destes experimentos. Os resultados obtidos e as recomendações feitas nessa fase inicial fornecem as indicações para a metodologia a ser adotada no próximo passo do trabalho: a construção de simulações de cenários experimentais de planos fatoriais não replicados para realizar um estudo comparativo dos gráficos normal e semi-normal. Nesta segunda fase, os critérios usados para a comparação dos gráficos são métodos objetivos, isto é, testes e procedimentos formais, com resultados analíticos, a fim de que a confrontação gráfica não se dê apenas subjetivamente, como a que ocorre na fase inicial do estudo. Este artigo concentra-se apenas nesta segunda fase da pesquisa: o estudo comparativo experimental de simulação.

Os cenários experimentais construídos para as simulações têm por base as indicações do levantamento bibliográfico realizado. Estes encontram-se divididos em três grupos. Cada grupo representa o aspecto sob o qual o desempenho dos gráficos é avaliado: (1) o julgamento da significância dos efeitos dos fatores; (2) a capacidade de detectar a presença de valores discrepantes (*outliers*) nos dados; (3) a identificação de restrições na aleatorização: o caso das parcelas subdivididas inadvertidas (*inadvertent plot-splitting*). Estes aspectos são escolhidos, pois, segundo a proposta de Daniel (1959), o gráfico semi-normal é uma ferramenta útil para estas três situações.

Inicialmente, para cada cenário apresentado são construídas as duas formas gráficas: o gráfico normal e o gráfico semi-normal. O objetivo é comparar o desempenho destas duas técnicas sob os três aspectos mencionados no parágrafo anterior. Busca-se avaliar em que condições e sob quais aspectos uma técnica gráfica é mais útil do que outra; sendo este “mais útil” no sentido de fornecer informações mais precisas, claras e evidentes das características dos dados. Todos os cenários são executados no pacote estatístico livre R versão R 2.10.0, e envolvem dados de experimentos fatoriais  $2^k$  e  $2^{k-p}$  não replicados. Nos concentraremos nos fatoriais cujas provas são 8, 16, 32 e 64 por serem os mais utilizados em experimentação (DANIEL, 1976; DONG, 1993; LAWSON, 2008). Vale destacar que inicialmente foi avaliada a influência do número  $n$  de simulações usadas para cada cenário. Para tal avaliação, usou-se como valores  $n = 5000$  e  $n = 10000$ . Alguns casos gerais foram considerados, e não houve evidências de diferenças significativas entre  $n = 5000$  e  $n = 10000$ . Portanto, visando a redução de tempo e de esforços computacionais, é utilizado  $n = 5000$  como o número de simulações para cada cenário apresentado e discutido neste artigo.

Embora as interpretações de instrumentos gráficos sejam subjetivas, o intuito deste trabalho é fazer com que pelo menos a escolha de qual técnica usar seja a menos subjetiva possível. Os resultados apresentados e as conclusões feitas a partir destes cenários servem de subsídio teórico para que um investigador, tendo em mente seu plano experimental e a configuração dos seus dados, saiba que método gráfico seria mais apropriado para a consecução de seus objetivos. Busca-se também contribuir para a controversa questão que envolve o uso do gráfico normal e

do gráfico semi-normal. Vale ressaltar que nenhum teste formal para a análise de experimentos fatoriais não replicados é fornecido aqui. O trabalho concentra-se no uso das técnicas gráficas em tais experimentos. Tais técnicas são ferramentas não formais e, portanto, subjetivas nas informações que fornecem; devendo ser utilizadas, sempre que possível, em associação com algum teste.

## 4 Resultados

Esta seção trata da apresentação e análise dos resultados das simulações. Ela está dividida em três subseções, em que cada uma delas representa uma das condições sob a qual as técnicas gráficas são comparadas e avaliadas. Ao final de cada subseção, a título de ilustração de algumas das características comentadas, é apresentado um exemplo de experimento real, extraído de artigos selecionados no levantamento bibliográfico. Em tais exemplos, os dois gráficos estão sendo comparados e aparecerão juntos numa mesma figura, lado a lado, com o gráfico semi-normal à esquerda e o gráfico normal à direita.

### 4.1 Julgamento da Significância dos Efeitos

O principal uso dos gráficos de probabilidade nos experimentos fatoriais não replicados é avaliar a significância dos efeitos (BOX; HUNTER; HUNTER, 2005; MONTGOMERY, 2001). Tanto no gráfico semi-normal quanto no gráfico normal, um efeito significativo é identificado como sendo aquele ponto que se distancia mais acentuadamente da reta pela qual os demais pontos se ajustam (BOX; MEYER, 1986; DANIEL, 1959). Como a interpretação gráfica é subjetiva, surgem alguns questionamentos: “O que seria uma distância acentuada da reta? Um ponto distante da reta é significativamente distante a ponto de ser considerado um efeito real?” Muitas pessoas podem dar respostas diferentes ao interpretarem um mesmo gráfico (CLEVELAND, 1987), uma vez que, neste caso, a visualização e interpretação de “distante” é algo particular.

Nesta subseção, portanto, os dois métodos gráficos são comparados com o objetivo de responder à seguinte pergunta: “Embora a interpretação gráfica seja subjetiva, em qual dos dois métodos o julgamento dos efeitos é mais fácil?” Em outras palavras: “Em qual gráfico a visualização de um efeito real é mais clara? Qual deles apresenta uma distância maior, em relação à reta, do ponto que representa um efeito significativo, tornando mais clara e fácil sua identificação?” O **método objetivo** utilizado para a comparação dos gráficos é a distância vertical dos efeitos significativos à reta pela qual os demais pontos se ajustam. Por tal critério de comparação, o gráfico “melhor” é aquele que apresentar os maiores valores para esta distância, significando que a identificação visual dos efeitos significativos, em tal gráfico, é mais clara, uma vez que estes se apresentam mais nitidamente distantes da reta.

Dos resultados das simulações, nota-se que o gráfico semi-normal apresenta, em média, distâncias maiores do que o gráfico normal. Sendo assim, o gráfico semi-normal é a técnica mais útil em auxiliar o investigador no seu julgamento sobre a significância dos efeitos dos fatores. Uma outra característica observada é que, à medida que aumenta a quantidade de efeitos significativos, os valores das distâncias calculadas, para ambos os gráficos, também aumentam, assim como aumenta também o número de efeitos com distâncias semelhantes; o que torna difícil a interpretação visual do julgamento destes efeitos: uma vez que muitos efeitos estão, aproximadamente, a uma mesma distância da reta, fica difícil ao investigador discernir, apenas pela avaliação visual, qual deles é realmente significativo. Isso indica que a habilidade das técnicas gráficas em julgar tais efeitos é comprometida. A situação “ideal” para o uso dos gráficos de probabilidade é aquela

em que menos de 20% dos efeitos são significantes (DANIEL, 1967; DONG, 1993). A Figura 1 ilustra bem essas afirmações.

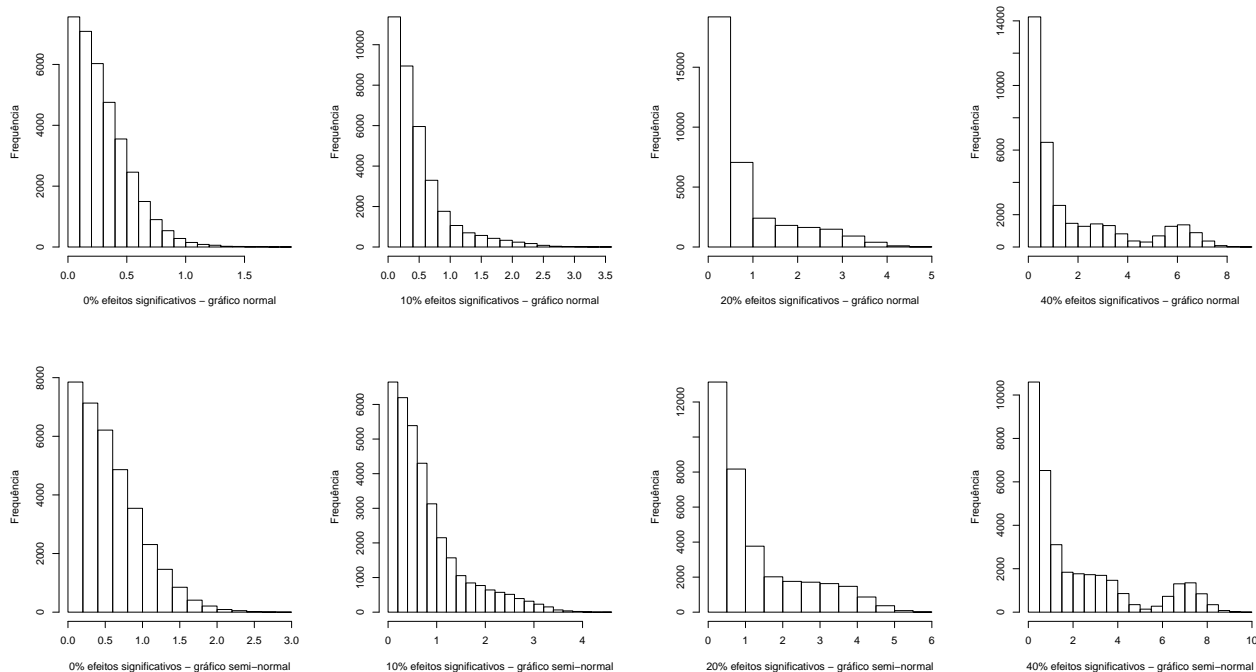


Figura 1: Comparação das distâncias nos gráficos de probabilidade. Trata-se de simulações de planos fatoriais  $2^3$ .

Com relação ao tamanho dos efeitos, nota-se que os efeitos  $\pm 3\sigma_{ef}$ <sup>1</sup> são, entre aqueles usados nas simulações, os mais difíceis de julgar, pois suas distâncias à reta são as menores, e quanto maior o número de provas (especialmente  $2^5$  e  $2^6$ ), mais difícil é essa identificação. Portanto, pode-se concluir que os gráficos de probabilidade aqui considerados são úteis em identificar efeitos significantes apenas se estes forem poucos (menor ou igual a 20%) e grandes (preferencialmente de tamanhos a partir de  $\pm 10\sigma_{ef}$ ).

**Exemplo 1 (Daniel, 1976):** Trata-se de um fatorial  $2^5$ . Os efeitos  $B$  e  $ABD$  são julgados significantes. Pela Figura 2, nota-se que tal julgamento é bem mais perceptível no gráfico semi-normal. Este exemplo também ilustra uma outra situação interessante. Aqui acontece algo que Daniel (1976) chama a atenção em seu livro: quando o valor absoluto de dois efeitos principais e de sua interação é aproximadamente o mesmo, então seus efeitos produzem a mesma resposta média. Para o caso em questão temos  $|B| = 0,20 \approx |D| = 0,17 \approx |BD| = 0,195$ . Esta situação é chamada de “critical mix” (BISGAARD; VIVACQUA; PINHO, 2004) ou mistura crítica<sup>2</sup>, sendo mais claramente percebida no gráfico semi-normal, pois os valores absolutos dos efeitos ficam quase na mesma altura. Esta é uma outra vantagem do gráfico semi-normal.

<sup>1</sup>Neste artigo,  $\sigma_{ef}$  refere-se ao desvio padrão dos efeitos, o qual, para os fatoriais não replicados, é dado por  $\sigma_{ef} = \frac{2 \cdot \sigma_{obs}}{\sqrt{2^k}}$  (Montgomery, 2001); em que  $\sigma_{obs}$  é o desvio padrão das observações que, por uma questão de simplicidade computacional e facilidade nas interpretações, nos cenários aqui apresentados,  $\sigma_{obs} = 1$ .

<sup>2</sup>O caso da mistura crítica não foi objeto de estudo deste trabalho.

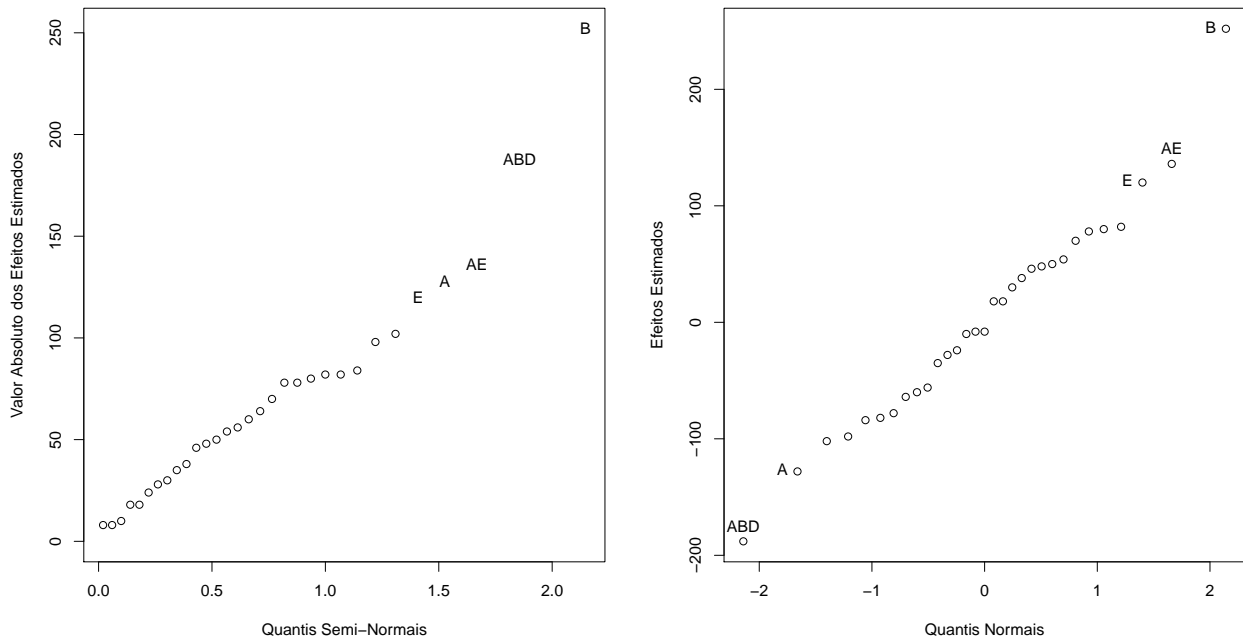


Figura 2: Gráficos de Probabilidade para o Exemplo 1.

## 4.2 Detecção de Valores Discrepantes (*outliers*) nos Dados

Valores discrepantes (*outliers*) são observações que não se ajustam ao padrão dos valores produzidos pelo resto dos dados; podem ser erros, ou então observações precisas, porém inesperadas, que lançam nova luz sobre o fenômeno em estudo; ou podem ser simplesmente manifestações da variabilidade inerente aos dados (STEFANSKY, 1971). Daniel (1959) acreditava que os valores discrepantes ou o que ele se refere como valores ruins (*bad values*) são o maior perigo de experimentos fatoriais não replicados, uma vez que eles “inflacionam” os contrastes, podendo fazer com que os efeitos reais não sejam detectados. Ele diz que os valores discrepantes “danificam” o experimento mais que os valores perdidos ou ausentes, pois não são facilmente identificáveis; e comenta também que sua ocorrência não é tão rara quanto se pensa.

É fácil entender a influência de um valor discrepante sobre a estimativa dos efeitos. Considere um fatorial  $2^k$  e assumamos que a  $i$ -ésima resposta,  $y_i$ , seja um valor discrepante, a qual vale agora  $y_i + \Delta$  em vez de apenas  $y_i$ . Todos os efeitos são então modificados por  $\frac{\Delta}{2^{(k-1)}}$  ou  $-\frac{\Delta}{2^{(k-1)}}$ , dependendo do sinal que aparece antes da resposta  $y_i$ , isto é, aqueles efeitos que possuem um sinal de “+” na linha do tratamento em que se encontra o valor discrepante serão  $\frac{\Delta}{2^{(k-1)}}$  unidades maiores; ao passo que aqueles efeitos que possuem o sinal de “-”, serão  $\frac{\Delta}{2^{(k-1)}}$  unidades menores. Por conta disso, a presença de valores discrepantes nos dados gera uma configuração padrão característica, específica para cada um dos dois tipos gráficos aqui considerados. No caso do gráfico normal, se  $\Delta$  é suficientemente grande, alguns efeitos são deslocados para a direita do gráfico e outros para a esquerda. Os pontos próximos a zero parecem seguir duas diferentes retas paralelas ao invés de uma, com os valores positivos sobre uma reta e os valores negativos sobre outra. Isto produz uma abertura ou uma lacuna (*gap*) no meio do gráfico; um espaço

aberto em torno da origem (BOX, 1990). Já no gráfico semi-normal, de acordo com Daniel (1960), a presença de um valor discrepante provoca um aumento no valor absoluto da média dos efeitos, fazendo com que o gráfico semi-normal tenha poucos efeitos próximos de zero. Desta forma, a reta traçada pela maioria dos pontos do gráfico não parte da origem, como deveria ser, mas aparece deslocada. Por isso, o **método objetivo** utilizado para a comparação gráfica é o tamanho dos *gaps*. A técnica gráfica terá o melhor desempenho em detectar a presença de valores discrepantes nos dados quanto maior for o tamanho do *gap* formado em tal gráfico, vindo assim a facilitar a identificação visual de tal situação. Portanto, ao observar esta forma padrão no gráfico, o pesquisador será avisado sobre a presença de um valor discrepante.

#### 4.2.1 Um Valor Discrepante Presente

Nas simulações em que um único valor discrepante estava presente, verificou-se que o *gap* formado no gráfico normal é significativamente maior que aquele formado no gráfico semi-normal; indicando que há fortes evidências para se crer que a identificação de valores discrepantes nos dados será mais clara no gráfico normal, uma vez que a percepção visual do *gap* é mais evidente. Notou-se também que, os sinais de “+” e “-” nos valores discrepantes não influenciam o tamanho dos *gaps*. Estes são afetados apenas pelo valor absoluto do tamanho do valor discrepante. Os valores negativos produzem resultados semelhantes aos valores positivos.

Quanto a posição em que se encontra o valor discrepante, os dois gráficos apresentam características distintas. No caso do gráfico semi-normal, o tamanho do *gap* permanece aproximadamente o mesmo, independentemente da posição do valor discrepante; só mudando de tamanho quando o valor discrepante muda seu tamanho também. No gráfico normal, a única posição que altera o tamanho do *gap* é quando o valor discrepante encontra-se na observação associada ao tratamento em que todos os fatores estão em seu nível “+”. Quando isso ocorre, há uma diminuição no tamanho do *gap*. Nota-se, então, que o gráfico semi-normal é sensível ao tamanho do valor discrepante; já o gráfico normal é sensível tanto ao tamanho quanto à posição do valor discrepante. Em relação ao tamanho, verificou-se que quanto maior for o valor discrepante, maior será o tamanho do *gap*. Isso evidencia que os gráficos de probabilidade são úteis em detectar grandes valores discrepantes (DANIEL, 1959, 1960; LAWSON, 2008). Pequenos valores discrepantes tornam a identificação gráfica difícil e ainda mais subjetiva.

Outra importante característica observada nas simulações é o fato de que os experimentos com 32 e 64 provas são mais robustos à presença de valores discrepantes (LAWSON, 2008). A forma gráfica da distribuição de seus *gaps* é bastante parecida à forma assumida quando nenhum valor discrepante está presente. Isso indica que a identificação de valores discrepantes nos gráficos de probabilidade construídos para tais planos experimentais pode ser mais difícil, uma vez que os *gaps* de tais gráficos têm tamanhos relativamente menores. Esta característica mostra que, o fato dum gráfico de probabilidade não evidenciar o característico *gap* que indica a presença de valores discrepantes, não significa que eles estejam ausentes. Pelo menos duas possibilidades podem ser consideradas em tal situação: ou o valor discrepante é significativamente pequeno, ou o experimento é robusto à sua presença.

No que diz respeito à influência do valor discrepante sobre a estimativa dos efeitos, uma característica observada é que, embora os sinais “+” e “-” dos valores discrepantes não apresentem diferenças significativas na influência do tamanho dos *gaps*, percebe-se que eles interferem na estimativa dos efeitos. Seja  $\Delta$  o tamanho do valor discrepante. Considere, inicialmente, o caso em que o valor discrepante está presente na observação associada ao tratamento em que todos os fatores estão em seu nível “+”. Se  $\Delta$  for positivo, então todos os efeitos estimados aumentam  $\frac{\Delta}{2^{(k-1)}}$ . Caso  $\Delta$  seja negativo, então todos os efeitos diminuem  $\frac{\Delta}{2^{(k-1)}}$ . Agora, se o valor discre-



pante estiver em qualquer outra posição, então metade dos efeitos é aumentada  $\frac{\Delta}{2^{(k-1)}}$  e a outra metade diminui  $\frac{\Delta}{2^{(k-1)}}$ . A influência dos sinais neste caso é a seguinte: os efeitos que aumentam quando o valor discrepante é positivo, diminuem quando este é negativo, e vice-versa.

Percebeu-se também que, em média, o tamanho do *gap* do gráfico normal é aproximadamente quatro vezes o tamanho do *gap* do gráfico semi-normal nos casos em que o valor discrepante é pequeno ou o número de provas é superior a 16. Essa razão tende a diminuir à medida que o tamanho do valor discrepante aumenta, mas permanece aproximadamente constante na presença de pequenos valores discrepantes ou nos experimentos robustos. A Figura 3, a Figura 4 e a Figura 5 ilustram algumas das conclusões mencionadas nesta Seção.

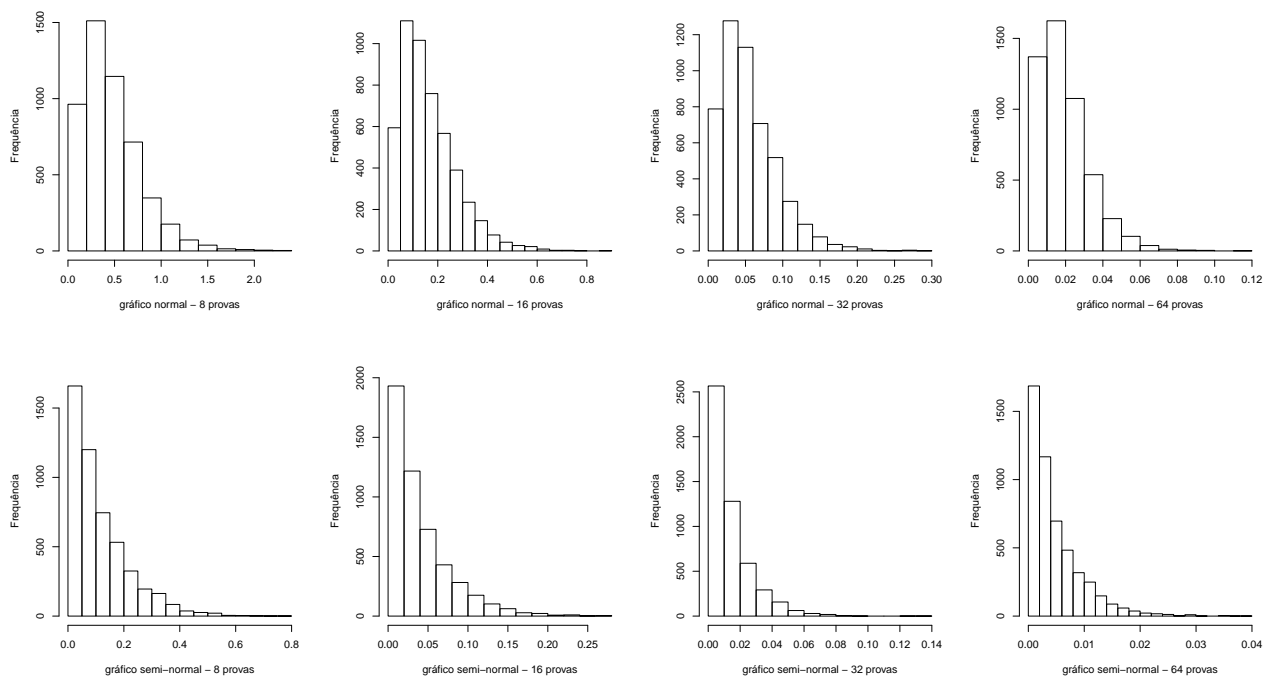


Figura 3: Comparação dos *gaps* nos gráficos de probabilidade sem a presença de valores discrepantes.

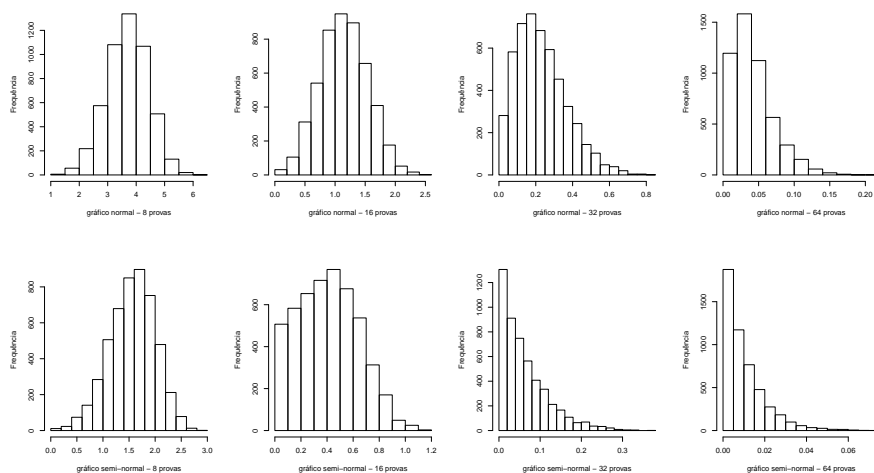


Figura 4: Comparação dos *gaps* nos gráficos de probabilidade quando o *outlier* é igual a 10. O *outlier* encontra-se no tratamento (1).

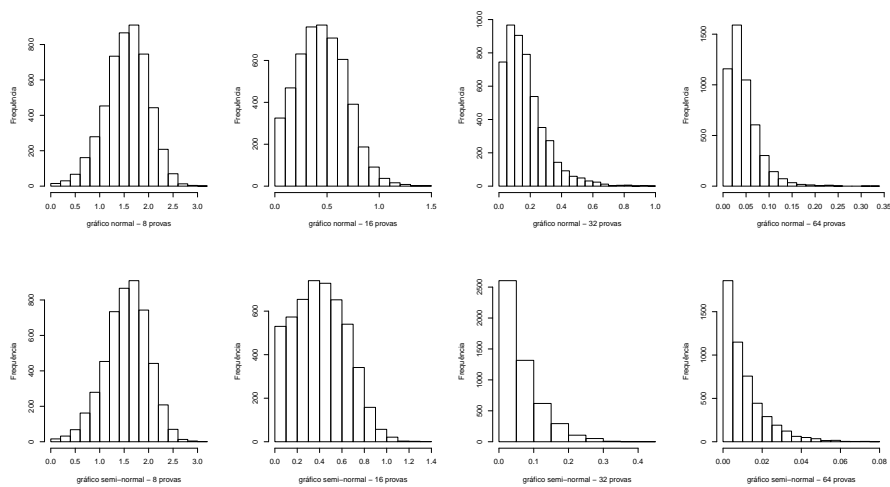


Figura 5: Comparação dos *gaps* nos gráficos de probabilidade quando o *outlier* é igual a 10. O *outlier* encontra-se no tratamento em que todos os fatores estão em seu nível “+”.

**Exemplo 2 (Goupy, 2006):** Desta feita temos um fatorial fracionado  $2^{6-2}$  cujos geradores são:  $E = ABC$ ;  $F = BCD$ . Pelos gráficos mostrados na Figura 6, é nítida a presença de valores discrepantes nos dados: o tamanho dos *gaps* é extremamente significativo. Neste caso, há um valor discrepante de tamanho  $10\sigma_{obs}$ , conforme especificado no artigo em questão.

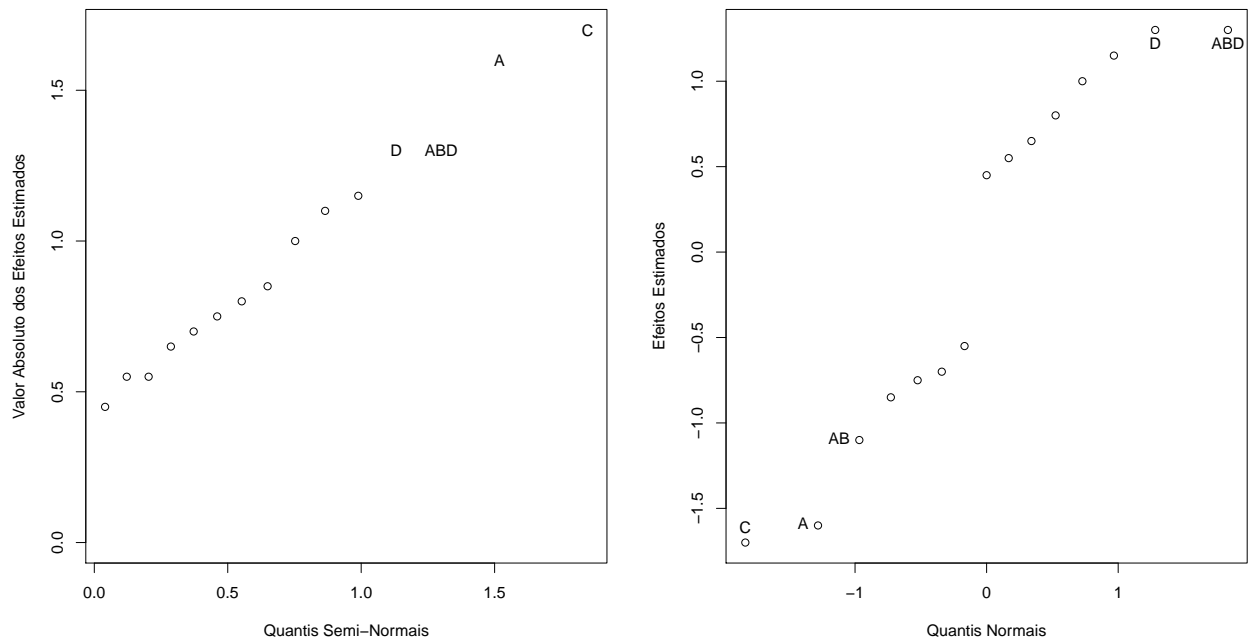


Figura 6: Gráficos de Probabilidade para o Exemplo 2.

**Exemplo 3 (Lawson, 2008):** Trata-se de um fatorial  $2^4$ . Pela Figura 7 pode-se perceber que o *gap* no gráfico normal é mais claramente identificável.

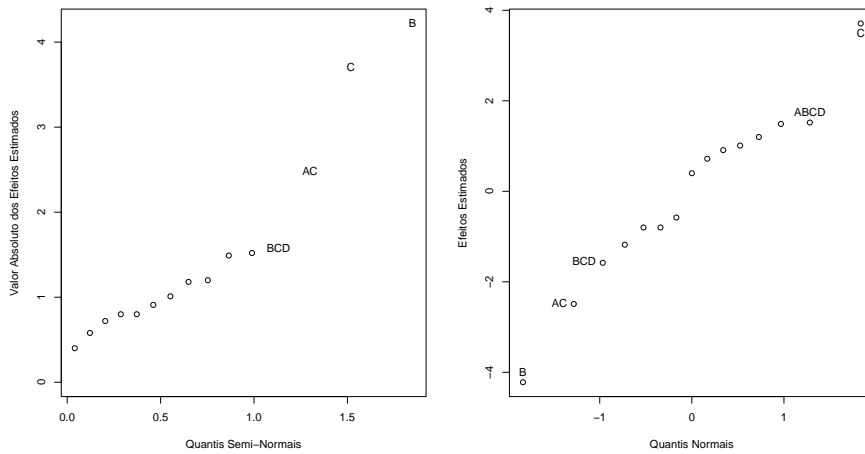


Figura 7: Gráficos de Probabilidade para o Exemplo 3.

**Exemplo 4 (Lawson, 2008):** Trata-se de um fatorial fracionado  $2^{5-1}$ . Embora o autor tenha detectado um valor discrepante nos dados pelo uso de seu método; tal conclusão não é óbvia pela análise dos gráficos da Figura 8.

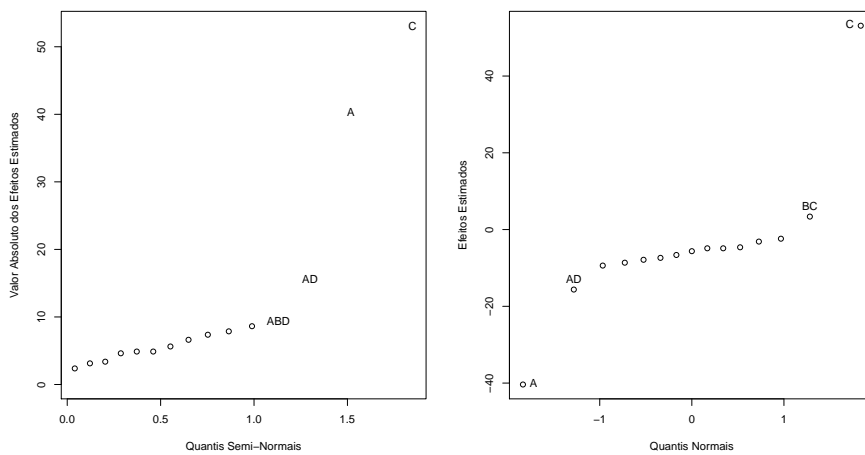


Figura 8: Gráficos de Probabilidade para o Exemplo 4.

### 4.2.2 Dois Valores Discrepantes Presentes

Nos casos em que dois valores discrepantes estão presentes nos dados, nota-se que, de um modo geral, o tamanho dos *gaps* é consideravelmente menor quando comparado com o caso de um único valor discrepante. A redução no tamanho dos *gaps* torna mais difícil a identificação visual da presença destes valores discrepantes por meio dos gráficos de probabilidade. Isso é uma indicação de que tais gráficos podem não ser úteis na identificação de mais de um valor discrepante nos dados (BOX, 1990; LAWSON, 2008). Apenas nas situações em que o tamanho de tais valores é grande, a identificação pode ser percebida mais claramente. Apesar disso, usando apenas os gráficos normal ou semi-normal, não é possível identificar se há mais de um valor discrepante, uma vez que apenas um único *gap* continua aparecendo, com tamanhos semelhantes ao caso de apenas um valor discrepante presente.

Assim como aconteceu no caso anterior, para todas as simulações, o *gap* formado no gráfico normal mostra-se significativamente maior que o formado no gráfico semi-normal. Isso reforça a indicação de que, no gráfico normal, a percepção da presença de valores atípicos pode ser mais clara. No caso do gráfico semi-normal algumas particularidades podem ser destacadas: mantém a característica, já observada, de que seu *gap* não é influenciado ou modificado se o valor discrepante estiver na observação associada ao tratamento “+++”; quando os dois valores discrepantes têm a mesma magnitude, independente do tamanho deles, o tamanho do *gap* é aproximadamente o mesmo. Considerando o gráfico normal, nota-se que: mesmo quando os valores discrepantes têm a mesma magnitude, o *gap* não preserva o mesmo valor, o que pode ser explicado pelo fato, já observado anteriormente, do tamanho do *gap* ser reduzido quando o valor discrepante encontra-se associado ao tratamento “+++”.

Com relação à influência de dois valores discrepantes, digamos  $\Delta_1$  e  $\Delta_2$ , sobre a estimativa dos efeitos, algumas particularidades foram observadas. Quando os dois têm a mesma magnitude: (1) Se nenhum deles estiver na posição “+++”, tem-se: metade dos efeitos permanece inalterada e, dos restantes, metade aumenta  $\frac{\Delta_1+\Delta_2}{2^{k-1}}$  e a outra metade diminui  $\frac{\Delta_1+\Delta_2}{2^{k-1}}$ ; (2) Se um deles estiver na posição “+++”, então metade dos efeitos estimados permanecem inalterados, e a outra metade aumenta  $\frac{\Delta_1+\Delta_2}{2^{k-1}}$ . Quando os dois têm magnitudes diferentes: (1) Se nenhum deles estiver na posição “+++”, tem-se: 25% aumentam  $\frac{\Delta_1+\Delta_2}{2^{k-1}}$ , 25% diminuem  $\frac{\Delta_1+\Delta_2}{2^{k-1}}$ , 25% aumentam  $\frac{\Delta_1-\Delta_2}{2^{k-1}}$  e 25% diminuem  $\frac{\Delta_1-\Delta_2}{2^{k-1}}$ ; (2) Se um deles estiver na posição “+++”, então, metade dos efeitos aumenta  $\frac{\Delta_1+\Delta_2}{2^{k-1}}$  e a outra metade aumenta  $\frac{\Delta_1-\Delta_2}{2^{k-1}}$ .

Do exposto nesta seção, pode-se concluir que, de um modo geral, os gráficos de probabilidade são úteis em detectar a presença de um único valor discrepante, sendo este de grande magnitude (preferencialmente de tamanhos a partir de  $10\sigma_{obs}$ ).

### 4.3 Identificação de Restrições na Aleatorização (*inadvertent plot-splitting*)

É comum em planejamento de experimentos a presença de fatores que são mais caros, consomem mais tempo ou são mais difíceis de mudar que outros. Isto torna inviável a execução do experimento sob a forma de um completamente aleatorizado, fazendo com que o investigador economize seus esforços por minimizar os custos dos fatores difíceis de mudar, criando então uma restrição na aleatorização (KULAHCI; BISGARRD, 2005). Esta situação dá origem aos experimentos *split-plot*, comumente conhecidos como experimentos em parcelas subdivididas.

As primeiras referências aos experimentos em parcelas subdivididas foram aplicadas no contexto da agricultura e apareceram na literatura por volta da década de 30, com os trabalhos de Fisher, que conduziu experimentos agrícolas com fertilizantes e culturas (ARVIDSSON et al., 2001). O lote inteiro (*whole plot*) de terra era subdividido ou parcelado (*split*) em sub-lotes

(*sub-plots*) ou sub-unidades, para que os diferentes fertilizantes fossem aplicados. Em seguida, os diferentes tipos de culturas poderiam então ser cultivados nestas diferentes parcelas. Daí a origem do nome *split-plot* ou parcelas subdivididas. Nestes planejamentos, os fatores difíceis de mudar são chamados de *whole-plot factors* ou fatores da parcela, e os mais fáceis de mudar de *sub-plot factors* ou fatores da subparcela.

Os planos em parcelas subdivididas apresentam restrições na aleatorização do experimento, sendo uma opção prática de planejamento nos casos em que não é possível ou razoável realizar o experimento numa ordem completamente aleatória. Quando um experimento fatorial não replicado tem restrições na aleatorização, Daniel (1959) sugeriu que fossem construídos dois gráficos semi-normais, um para os efeitos da parcela (*whole plot*) e outro para os efeitos da subparcela (*sub-plot*). A razão para isso é comentada no artigo: quando o gráfico é construído com todos os efeitos juntos, os dois tipos de erro (o erro associado à parcela e o erro associado à subparcela) estão misturados, o que acaba afetando a forma do gráfico e, conseqüentemente, o julgamento dos efeitos. Segundo Daniel, a forma curvada com que os pontos são dispostos no gráfico semi-normal é um indicativo de que há restrições na aleatorização do experimento, o que serve de alerta para o investigador sobre a presença de um plano em parcelas subdivididas. É um erro comum apresentar um experimento como completamente aleatorizado quando, na verdade, se trata de um planejamento em parcelas subdivididas. A proposta de Daniel é que o gráfico semi-normal identifique esta situação e oriente o investigador a corrigi-la. A esta situação ele se referiu como *inadvertent plot-splitting* ou parcelas subdivididas inadvertidas, isto é, embora tenha realizado o experimento em parcelas subdivididas, o pesquisador inadvertidamente o analisa de outra forma, por não conhecer tal plano experimental.

Søren Bisgaard, Howard Fuller e Ernesto Barrios (1995) aplicam para o gráfico normal o mesmo uso que Daniel indicou para o gráfico semi-normal no diz respeito a situação de parcelas subdivididas inadvertidas comentada no parágrafo anterior. A sugestão dos autores é calcular os efeitos em dois grupos separados: aqueles afetados pela variância do erro da parcela e aqueles afetados apenas pela variância do erro das subparcelas, e então construir dois gráficos normais, um para cada grupo, a fim de avaliar a significância dos efeitos. Isto é útil para não esconder efeitos significativos, nem tampouco para julgar como ativo um efeito que na verdade é inativo. Quando é feito apenas um único gráfico para todos os efeitos do experimento em parcelas subdivididas, a distribuição misturada tem caudas mais pesadas, fazendo com que a curva traçada no gráfico normal tenha uma forma se  $S$ .

Por serem tão comumente usados na pesquisa industrial, por serem os mais convincentes e em muitos casos, os únicos possíveis (DANIEL, 1976); os planejamentos em parcelas subdivididas não poderiam deixar de ser considerados neste trabalho. Apesar desta ser uma importante situação em que os gráficos de probabilidade são utilizados para a análise dos experimentos fatoriais não replicados, nenhum método objetivo para a comparação dos gráficos em tais situações é usado nas simulações aqui apresentadas. Motivo: embora algumas técnicas e métodos tenham sido elaborados e testados durante o andamento deste artigo, ainda não se obteve êxito em encontrar algum que pudesse ser útil para este estudo experimental comparativo. Apesar disso, algumas conclusões podem ser construídas a partir do levantamento bibliográfico realizado. O gráfico normal mostra-se a ferramenta mais útil e de melhor identificação visual em relação ao gráfico semi-normal. A sua característica forma de  $S$ , padrão para os casos dos experimentos com parcelas subdivididas, mostrou-se evidente nos experimentos encontrados nos artigos, tornando mais clara a detecção desta restrição na aleatorização. No caso do gráfico semi-normal, não foi percebida uma forma padrão para esta situação que ajudasse a identificar o plano em parcelas subdivididas conforme Daniel (1959) afirmou. Embora se note, é verdade, uma certa curvatura

no gráfico semi-normal nestes casos, não ficou muito claro que tipo de forma este gráfico assume nestas circunstâncias, ao contrário do que já se espera do gráfico normal.

**Exemplo 5 (Montgomery; Kulahci; Almimi, 2007):** Trata-se de um fatorial  $2^5$ . Pela Figura 9, a nítida forma de  $S$  no gráfico normal é um indicativo da ocorrência de parcelas subdivididas.

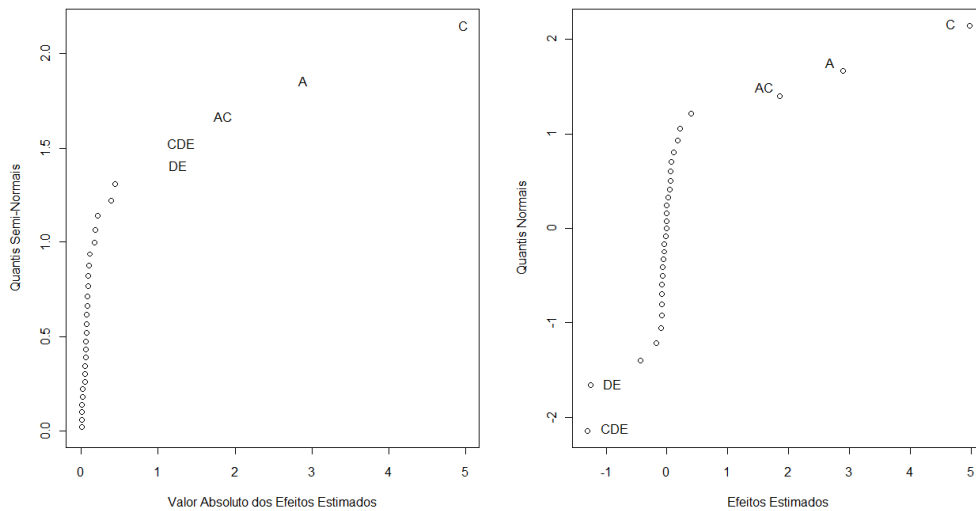


Figura 9: Gráficos de Probabilidade para o Exemplo 5.

## 5 Considerações Finais

Neste trabalho realizamos um estudo comparativo experimental de gráficos de probabilidade normal para análise de experimentos fatoriais não replicados. Os resultados obtidos fornecem orientações potencialmente úteis aos pesquisadores ao interpretarem os resultados de um experimento. Por exemplo, verificamos que não há um procedimento absolutamente melhor que outro. As duas técnicas gráficas apresentaram vantagens e limitações, dependendo do caso considerado. A decisão de qual gráfico usar dependerá da situação que se deseja analisar no experimento. Neste artigo foram consideradas três situações: (1) **Julgamento da significância dos efeitos:** Nesse caso, o gráfico semi-normal mostrou-se melhor. **Justificativa:** A distância do efeito à reta é consideravelmente maior neste gráfico do que no gráfico normal, facilitando assim o seu julgamento. (2) **Detecção de valores discrepantes:** Para essa situação, o recomendado é o gráfico normal. **Justificativa:** O tamanho do *gap* característico para esses casos é consideravelmente maior em tal gráfico do que no gráfico semi-normal. Apesar disso, o gráfico semi-normal tem a vantagem de seu *gap* não depender da posição em que tais valores se encontram. (3)

**Identificação de restrições na aleatorização:** Embora não se tenha apresentado neste trabalho nenhum procedimento formal para comparar os dois gráficos no caso da identificação das parcelas subdivididas inadvertidas, pelos poucos artigos encontrados que tratam do assunto, a indicação é usar o gráfico normal. **Justificativa:** A forma de  $S$  característica do gráfico normal para essa situação é mais claramente perceptível.

Não apenas para o gráfico semi-normal, mas também para o gráfico normal, foram identificadas algumas limitações: estes são úteis apenas se a quantidade de efeitos significativos for pequena; além de poucos, os efeitos devem ter tamanhos consideráveis, afim de que possam ser visivelmente identificados pelos gráficos; sua capacidade de detectar valores discrepantes é comprometida quando há mais de um presente, ou quando sua magnitude é pequena. Além disso, tendo apenas como base alguns poucos artigos analisados, ainda não ficou muito evidente qual a configuração do gráfico semi-normal na presença de restrições na aleatorização.

Portanto, com o conhecimento adquirido através desta pesquisa, tanto por meio dos exemplos discutidos dos artigos, quanto pelos resultados das simulações realizadas, a recomendação feita é que, ao analisar experimentos fatoriais  $2^k$  e  $2^{k-p}$  não replicados com o uso das técnicas gráficas aqui consideradas, o pesquisador utilize o gráfico normal e o gráfico semi-normal simultaneamente, em vez de optar pela escolha de uma ou de outra técnica, pois, conforme discutido, cada uma delas possui vantagens e limitações que, juntas, podem contribuir para uma análise subjetiva mais completa dos dados.

## Referências Bibliográficas

ARVIDSSON, M.; KAMMERLIND, P.; HYNÉN, A.; BERGMAN, B. Identification of factors influencing dispersion in split-plot experiments. *Journal of Applied Statistics*, v.28. n. 3-4, 269-283, 2001.

BISGAARD, Søren; FULLER, Howard T., BARRIOS, Ernesto. Two-level factorials run as split-plot experiments. *Report Series in Quality and Productivity - University of Wisconsin-Madison*, n. 136, 1995.

BISGAARD, Søren; VIVACQUA, Carla Almeida; PINHO, André L. S. de. Quality Quandaries: Not All Models Are Polynomials! *Quality Engineering*, v. 17, n. 1, p. 181-186, 2004.

BOX, G. E. P. Finding bad values in factorial designs. *Report Series in Quality and Productivity - University of Wisconsin-Madison*, n. 57, 1990.

BOX, G. E. P.; HUNTER, J. S.; HUNTER, W. G. *Statistics for experimenters: design, innovation, and discovery*. 2. ed. Hoboken, N.J.: Wiley-Interscience, 2005.

BOX, G. E. P.; MEYER, R. D. An analysis for unreplicated fractional factorials. *Technometrics*, v. 28, n. 1, p. 11-18, 1986.

CLEVELAND, William S. Research in Statistical Graphics. *Journal of the American Statistical Association*, v. 82, n. 398, p. 419-423, 1987.

DANIEL, C. Use of half-normal plots in interpreting factorial two-level experiments. *Technometrics*, v. 1, n. 4, p. 311-341, 1959.

DANIEL, C. Locating Outliers in Factorial Experiments. *Technometrics*, v. 2, n. 2, p. 149-156, 1960.



- DANIEL, C. *Half-Normal Plots*. In: Encyclopedia of Statistical Sciences, edição de Samuel Kotz & Norman L. Johnson. v. 3. New York: J. Wiley & Sons, Inc., 1967.
- DANIEL, C. *Applications of statistics to industrial experimentation*. New York: J. Wiley & Sons, Inc., 1976.
- DONG, Fang. On the identification of active contrasts in unreplicated fractional factorials. *Statistica Sinica*, v. 3, p. 209-217, 1993.
- GOUPY, Jacques. Factorial experimental design: Detecting an outlier with the dynamic variable and the Daniel's diagram. *Chemometrics and Intelligent Laboratory Systems*, v. 80, p. 156-166, 2006.
- JOHNSON, Richard A.; BHATTACHARYYA, Gouri K. *Statistics: principles e methods*. 2. ed. New York: John Wiley & Sons, Inc., 1992.
- KULAHCI, Murat; BISGAARD, Søren. The use of Plackett-Burman designs to construct split-plot designs. *Technometrics*, v. 47, n. 4, p. 495-501, 2005.
- LAWSON, J. SAS macros for analysis of unreplicated  $2^k$  and  $2^{k-p}$  designs with a possible outlier. *Journal of Statistical Software*, v. 25, n. 13, p. 1-17, 2008.
- MONTGOMERY, Douglas C. *Design and analysis of experiments*. 5. ed. New York: John Wiley & Sons, Inc., 2001.
- MONTGOMERY, Douglas C.; KULAHCI, Murat; ALMIMI, Ashraf A. Estimation of missing observations in two-level split-plot designs. *Quality and Reliability Engineering International*, v. 24, p. 127-152, 2007.
- OLGUIN, J.; FEARN, T. A new look at half-normal plots for assessing the significance of contrasts for unreplicated factorials. *Applied Statistics*, v. 46, n. 4, p. 449-462, 1997.
- STEFANSKY, Wilhelmine. Rejecting outliers by maximum normed residual. *The Annals of Mathematical Statistics*, v. 42, n. 1, p. 35-45, 1971.
- UPTON, Graham J. G.; COOK, Ian. *A Dictionary of Statistics*. 2. ed. New York: Oxford University Press, 2006.
- WILK, M. B.; GNANADESIKAN, R. Graphical methods for internal comparisons in multi-response experiments. *The Annals of Mathematical Statistics*, v. 35, n. 2, pp. 613-631, 1964.
- WILK, M. B.; GNANADESIKAN, R. Probability plotting methods for the analysis of data. *Biometrika*, v. 55, n. 1, pp. 1-17, 1968.