

ANÁLISE BAYESIANA DE TOXIDEZ E INTERFERÊNCIA EM ENSAIOS DE DILUIÇÃO SERIADA

Andréa Cristiane dos Santos DELFINO¹
Maria Imaculada de Sousa SILVA²
Roberta Hilsdorf PICCOLI³
Júlio Sílvio de Sousa BUENO FILHO⁴

- RESUMO: O número mais provável (NMP) é o resultado da estimativa da densidade de microrganismos provenientes de diluições seriadas. O modelo padrão utilizado para obter tabelas destas estimativas, doravante chamado modelo usual, não contempla situações tais como toxidez e interferência. A análise bayesiana foi utilizada para inferir sobre estes efeitos. Usando o algoritmo Metropolis-Hastings, foi possível gerar cadeias para cada parâmetro dos modelos. A comparação entre os modelos foi feita por meio do fator de Bayes. Foi possível discriminar, entre modelos com parâmetros de interferência e toxidez, os quais são mais precisos que o modelo usual para certos resultados experimentais. Em situações experimentais em que o crescimento do microrganismo de interesse é inibido em baixas diluições, o modelo usual não é adequado e tem valores subestimados para o NMP. No entanto, em casos em que há crescimento de microrganismo em baixas diluições, o modelo usual é mais adequado e seus intervalos de credibilidade são menores que aqueles dos modelos mais complexos. A rotina implementada em R pode ser estendida para uma ampla gama de planos de diluição (número de diluições e número de tubos por diluição) e pode ser usada em substituição às tabelas dos laboratórios.
- PALAVRAS-CHAVE: Algoritmo Metropolis-Hastings; ensaios de diluição; inferência bayesiana; número mais provável.

¹Universidade Federal de São João Del-Rei – UFSJ, Departamento de Matemática e Estatística, CEP: 36307-352, São João Del-Rei, MG, Brasil. E-mail: andrea@ufsj.edu.br

²Universidade Federal de Uberlândia – UFU, Faculdade de Matemática, CEP: 38408-100, Uberlândia, MG, Brasil. E-mail: maria@famat.ufu.br

³Universidade Federal de Lavras – UFLA, Departamento de Ciência dos Alimentos, Caixa Postal 3037, CEP: 37200-000, Lavras, MG, Brasil. E-mail: rhpiccoli@dca.ufla.br

⁴Universidade Federal de Lavras – UFLA, Departamento de Ciências Exatas, Caixa Postal 3037, CEP: 37200-000, Lavras, MG, Brasil. E-mail: juliobuenof@gmail.com

1 Introdução

Na análise de alimentos processados, a presença de um determinado microrganismo é considerada um indicativo de contaminação pós-sanificação ou pós-processamento, e de que as práticas de higienização estão aquém dos padrões estabelecidos (Silva *et al.*, 1997). Desta forma, a utilização de técnicas de enumeração de microrganismos indicativos é rotineira.

A quantidade destes microrganismos determina o quão seguro um determinado alimento está para o consumo humano. Dentre os microrganismos indicativos, os mais comuns são os coliformes totais, e os coliformes termotolerantes ou fecais. Tais bactérias são indicadores de contaminação fecal. A quantificação desses microrganismos geralmente é realizada empregando-se a técnica do Número Mais Provável (NMP), embora outras sejam possíveis.

A técnica do NMP baseia-se em diluições seriadas que consistem na realização sucessiva de diluições, a partir de uma amostra inicial, onde se mede a concentração dos microrganismos indicadores. Na versão mais simples e mais utilizada deste tipo de ensaio, mede-se apenas a presença ou a ausência de microrganismos em um ou mais tubos das diversas diluições.

O procedimento laboratorial padrão do NMP, utiliza números restritos de diluições para cada amostra. Após a incubação, é observada a presença (sinal positivo) ou ausência (sinal negativo) de crescimento evidenciado por meio da formação de bolha.

De modo geral, nos procedimentos de diluição seriada, faz-se a homogeneização do material que será analisado em um diluente adequado. No caso de líquidos, eles são colocados em frascos com espaço suficiente para a agitação. Embora existam várias exceções, que dependem do material a ser analisado e da quantidade de microrganismos existentes, normalmente, utilizam-se 25 ml do material a ser analisado e 225 ml de diluente na homogeneização. A escolha do diluente depende do material de análise. Depois da homogeneização, transfere-se uma porção de 1 ml da amostra homogeneizada para um tubo contendo 9 ml de diluente. Esta primeira diluição é denotada por 10^{-1} . Como é um processo de diluições sucessivas, para a diluição 10^{-2} , é retirado 1 ml da diluição 10^{-1} , o qual é colocado em um tubo com 9 ml do diluente; para a diluição 10^{-3} , o procedimento é o mesmo e assim sucessivamente, sempre tirando 1 ml da diluição anterior e inoculando-se em 9 ml de diluente. Em outras palavras, pode-se dizer que, na primeira diluição, a quantidade inoculada é 10^{-1} g ou ml; na segunda diluição, a quantidade inoculada é 10^{-2} g ou ml e, na terceira diluição, a quantidade inoculada é 10^{-3} g ou ml. A mesma forma é utilizada para realizar as demais diluições após a terceira. A quantidade inoculada da amostra determina a qual diluição ela pertence, de acordo com o número de diluições utilizadas.

Após o processo de diluições sucessivas, realiza-se o teste presuntivo, que consiste em inocular uma série de três tubos de Durhan por diluição no meio de cultura adequado, como por exemplo, caldo Lauril Sulfato Triptose (LST). Este meio de cultura é utilizado para favorecer o crescimento do microrganismo que será

analisado. Em cada um desses três tubos será adicionado 1 ml da solução diluída a 10 ml de LST. Sendo assim, tem-se uma série nove tubos, ou seja, de três diluições e, para cada diluição, três tubos. Nesta fase, conta-se a quantidade de tubos que foram positivos, ou seja, que apresentaram crescimento (Silva *et al.*, 1997).

Supondo um ensaio experimental com 3 diluições e 3 tubos para cada diluição, um possível resultado seria (321). Este resultado significa que, na primeira diluição, houve crescimento de microrganismo em 3 tubos; na segunda diluição, apenas 2 tubos mostraram crescimento e, na terceira somente 1 tubo mostrou crescimento.

Ao interpretar os resultados obtidos nestes ensaios, usando modelo probabilístico hierárquico, calcula-se o estimador de máxima verossimilhança do número de microrganismos. O método de obtenção de tais estimativas a partir desses delineamentos ficou conhecido como o método do número mais provável, sendo este conceito introduzido por McCrady (1915) e, desde então, utilizado amplamente.

A estimativa do NMP é muito utilizada para verificação da qualidade do alimento. Para a interpretação dos resultados das análises microbiológicas de alimentos, a Agência Nacional de Vigilância Sanitária (Anvisa, 2001) dispõe de tabelas implementadas em seu regulamento, que indicam valores de tolerância máxima da ocorrência de determinados microrganismos em alimentos destinados ao consumo humano. Tais estimativas são obtidas pelo método da máxima verossimilhança.

Na literatura, as estimativas do NMP podem ser encontradas em tabelas que são usadas de forma rotineira em laboratórios de análises microbiológicas, considerando-se apenas três diluições sucessivas, e séries de 3, 5, 8 e 10 tubos. Entretanto, as tabelas são adequadas para um número limitado de resultados experimentais, contemplando situações em que, em baixas diluições, pode ou não ocorrer o crescimento de microrganismos. No entanto é importante ressaltar que estas tabelas foram confeccionadas considerando situação específica, ou seja, que haja crescimento de microrganismos em baixas diluições. Esta é a segunda suposição citada por Cochran (1950) para o cálculo do NMP. O autor também salienta que se o crescimento de microrganismos é inibido por algum fator, a estimativa do número mais provável terá seu valor subestimado. Sendo assim nas situações em que não há crescimento de microrganismo em baixas diluições, os valores das estimativas de microrganismos proveniente destas tabelas estão subestimados.

Blodgett e Garthright (1998), apresentam o modelo comumente utilizado para se obter tais estimativas, doravante denominado “modelo usual”. Este modelo probabilístico não se ajusta a resultados experimentais em que não há crescimento de microrganismos em baixas diluições. A inibição do crescimento de microrganismos pode estar associada a fatores como toxidez no meio de cultura e interferência por competição entre espécies, além de outros de difícil explicação.

Blodgett e Garthright (1998) sugerem então, modelos alternativos que seriam mais prováveis para alguns desses resultados experimentais, sendo estes o “modelo de toxidez” e o “modelo de interferência”. No entanto, não foram apresentados métodos integrados de estimação e avaliação do ajuste em tais modelos. Além disso, não foram encontradas na literatura estimativas de número mais provável

envolvendo os modelos descritos por Blodgett e Garthright (1998).

O objetivo deste trabalho foi desenvolver uma abordagem bayesiana para os modelos usual, de toxidez e de interferência. Isto permite estabelecer comparações sobre a qualidade dos ajustes destes modelos usando o fator de Bayes e decidir se o modelo usual pode ser empregado. Este tipo de informação, permite avaliar experimentos de uma ampla gama de planos de diluição, incluindo situações em que não se aplicam as tabelas usuais na avaliação de ensaios. A análise de dois resultados experimentais (um que favorece o modelo usual e outro em que tabelas do modelo usual não podem ser empregadas) é apresentada para fins de ilustração.

2 Material e métodos

2.1 Material experimental

São apresentados neste trabalho, dois possíveis resultados experimentais de diluições seriadas que ilustram situações comuns em procedimentos laboratoriais. Para tanto foram considerados experimentos com cinco diluições, e, para cada uma dessas, três tubos.

Os resultados experimentais utilizadas no ajuste dos modelos foram (32100) e (01300) sendo que as quantidades inoculadas das amostras foram 0,1; 0,01; 0,001; 0,0001 e 0,00001 diluições sucessivas de 1:10. O resultado experimental (32100) indica que houve crescimento de microorganismos nos três tubos da primeira diluição, em dois dos três tubos da segunda diluição, em um tubo da terceira diluição, e em nenhum tubo da quarta e quinta diluição. No resultado experimental (01300), houve crescimento em um dos tubos da segunda diluição, e nos três tubos da terceira diluição.

2.2 Metodologia

Nas próximas seções apresentam-se as descrições dos modelos usual, de toxidez e de interferência (Blodgett e Garthright, 1998). Uma generalização dos modelos, possibilitando contemplar a toxidez e a interferência em um único modelo, também pode ser de interesse para alguns resultados experimentais. Essa alternativa já foi desenvolvida em outro trabalho (Santos, 2008), e da forma como foi proposta, verificou-se forte dependência entre o modelo completo e os modelos de toxidez e de interferência. Portanto, na prática, os modelos em separado, que são o foco principal deste trabalho, podem ser suficientes para detectar possíveis desvios do modelo usual. A forma de implementar a inferência bayesiana e o processo de amostragem empregado são descritos a seguir.

2.2.1 Modelo usual

Seja um ensaio de diluição em que k denota o número de diluições seriadas; z_j , $j = 1, \dots, k$, é a quantidade inoculada da amostra na j -ésima diluição ($z_j > 0$),

λ é a concentração de microrganismos na substância original ($\lambda \geq 0$) e X é o número de microrganismos dentro de um tubo qualquer. Sendo $E(X|z_j) = \lambda z_j$ o número esperado de microrganismos dentro de um tubo qualquer da j -ésima diluição, determina-se, por meio da distribuição Poisson, a probabilidade de não ocorrer crescimento de microrganismos nesse tubo, qual seja,

$$P(X = 0) = \frac{e^{-\lambda z_j} (\lambda z_j)^0}{0!} = e^{-\lambda z_j}. \quad (1)$$

Seja y o número de tubos que apresentam crescimento na j -ésima diluição $Y = (y_1, y_2 \dots y_k)$. Então Y tem distribuição binomial, com probabilidade de sucesso dada por:

$$\pi_j = 1 - e^{-\lambda z_j}.$$

Considerando n_j tubos na diluição z_j , de acordo com a distribuição binomial a função de verossimilhança é definida por:

$$L(Y|\lambda) = \prod_{j=1}^k \binom{n_j}{y_j} (1 - e^{-\lambda z_j})^{y_j} (e^{-\lambda z_j})^{(n_j - y_j)}. \quad (2)$$

A expressão (2) é denominada “modelo usual”.

2.2.2 Modelo de toxidez

No modelo de toxidez considera-se que o meio de cultura inicialmente contenha alguma substância tóxica para o microrganismo de interesse, inibindo seu crescimento. Com as sucessivas diluições, a concentração desta substância diminuirá, criando situação propícia para o crescimento do microrganismo. Para representar o efeito tóxico, é associada a probabilidade T_j de que o crescimento do microrganismo seja inibido na j -ésima diluição, sendo $T_1 < T_2 < \dots < T_k$, $0 \leq T_j \leq 1$. Sendo assim, a probabilidade de não haver crescimento de microrganismo em um tubo qualquer é definida por:

$$P(X = 0) = e^{-\lambda z_j} \sum_{i=0}^{\infty} \frac{(\lambda z_j T_j)^i}{i!} = e^{-\lambda z_j} e^{\lambda z_j T_j} = e^{-\lambda z_j (1 - T_j)}. \quad (3)$$

A variável aleatória Y que denota o número de tubos que apresentam crescimento de microrganismos segue o modelo binomial e a função de verossimilhança deste modelo é dada por:

$$L(Y|\lambda, T_j) = \prod_{j=1}^k \binom{n_j}{y_j} (1 - e^{-\lambda z_j (1 - T_j)})^{y_j} \times (e^{-\lambda z_j (1 - T_j)})^{(n_j - y_j)}. \quad (4)$$

Pode-se notar que o que difere a expressão (2) da expressão (4) é a grandeza $(1 - T_j)$ multiplicando o expoente. Se T_j for igual a zero, a Equação 4 se reduz a Equação 2.

Diferentemente de Blodgett e Garthright (1998), que consideram o parâmetro referente à substância tóxica como um valor fixo, pretende-se estimar, ou ao menos inferir sobre a ocorrência, tal efeito.

2.2.3 Modelo de interferência

No modelo de interferência, assume-se que, no mesmo tubo, estejam crescendo duas espécies de microrganismos. Seja λ a concentração de microrganismos da espécie de interesse e $\lambda_i \geq 0$ a concentração da espécie que interfere no crescimento do microrganismo de interesse. Esta interferência pode ser predação, competição por nutrientes ou outro mecanismo.

A probabilidade de crescimento do microrganismo de interesse é a mesma do modelo usual e é dada por: $1 - e^{-\lambda z_j}$.

Por outro lado a probabilidade de não haver uma espécie interferindo é dada por:

$$P(X = 0) = \frac{e^{-\lambda_i z_j} (\lambda_i z_j)^0}{0!} = e^{-\lambda_i z_j}. \quad (5)$$

Supondo que as duas espécies sejam distribuídas independentemente, a probabilidade de crescimento do microrganismo de interesse em um tubo na j -ésima diluição é dada por:

$$e^{-\lambda_i z_j} (1 - e^{-\lambda z_j}). \quad (6)$$

Desta forma define-se a função de verossimilhança do modelo por:

$$L(Y|\lambda, \lambda_i) = \prod_{j=1}^k \binom{n_j}{y_j} [e^{-\lambda_i z_j} (1 - e^{-\lambda z_j})]^{y_j} \times [1 - e^{-\lambda_i z_j} (1 - e^{-\lambda z_j})]^{(n_j - y_j)}. \quad (7)$$

Se λ_i for igual a zero, a Equação 7 se reduz à expressão (2).

2.2.4 Métodos bayesianos

A inferência sobre todos os parâmetros de interesse, nos modelos apresentados, foram feitas usando aproximações das distribuições *a posteriori* conjuntas por meio de métodos Monte Carlo via cadeias de Markov (MCMC) (Gelman *et al.*, 2000), mais especificamente o algoritmo Metropolis-Hastings. O algoritmo e justificativas teóricas para seu uso, podem ser encontrados com maiores detalhes em Chib e Greenberg (1995) e em Gamerman (1997).

A distribuição *a priori* considerada para os parâmetros λ , dos três modelos, e λ_i , do modelo de interferência, foi a gama com parâmetros de forma e escala especificados a priori por (α, β) e (α_i, β_i) , respectivamente, também chamados

de hiperparâmetros conhecidos. Essas distribuições são então representadas respectivamente por:

$$P(\lambda) \propto \lambda^{\alpha-1} e^{-\frac{\lambda}{\beta}}. \quad (8)$$

$$P(\lambda_i) \propto \lambda_i^{\alpha_i-1} e^{-\frac{\lambda_i}{\beta_i}}. \quad (9)$$

Os hiperparâmetros desta distribuição foram continuamente atualizados com uma amostra dos últimos 1000 valores da cadeia da distribuição condicional completa a posteriori, recalculando-se os valores com base nas definições de média e variância da distribuição gama. Os valores iniciais para os hiperparâmetros foram 2 para o parâmetro de forma e $\frac{1}{100}$ para o parâmetro de escala, para todas as priors com distribuição gama, de todos os modelos.

A distribuição uniforme em intervalos $(0, T_{j-1})$ foi utilizada como distribuição *a priori* para o parâmetro T_j do modelo de toxidez, sendo que para T_0 foi utilizado o intervalo $(0, 1)$.

Os critérios utilizados para verificar se a cadeia é estacionária foram os diagnósticos de convergência de Raftery e Lewis (1992), de Heidelberger e Welch (1983), além da análise visual do traço da cadeia, que indica convergência quando se observa uma aparência aleatória e estacionária.

O diagnóstico de Raftery e Lewis (1992) baseia-se em obter, a partir de uma amostra piloto, estimativas do número de iterações necessárias para se obter a convergência, do número de iterações iniciais que devem ser descartadas (burn-in) e da distância mínima (k); k é o fator de dependência de uma iteração à outra para se obter uma amostra independente. Segundo os autores, se o fator de dependência, for maior que 5, pode-se concluir que a cadeia apresenta problemas de convergência. O Fator de dependência é uma medida da autocorrelação espacial e indica o alcance do processo autorregressivo, ou seja, observações precisam ser tomadas a este intervalo para serem consideradas não correlacionadas. O caso mais comum é o de o fator ser maior que um, indicando que há um processo autorregressivo. Neste caso, o tamanho efetivo da amostra final será menor que o tamanho real da amostra e o erro do processo Monte Carlo será maior que o esperado para uma amostragem aleatória simples.

Para testar a hipótese nula da estacionaridade de uma cadeia gerada usando o critério de Heidelberger e Welch (1983), dois testes simultâneos são utilizados. Se a hipótese nula é rejeitada para um dado valor, o teste é repetido depois de descartados os 10% valores iniciais da seqüência. Se a hipótese é novamente rejeitada, mais 10% dos valores iniciais são descartados e assim sucessivamente até serem descartados os 50% valores iniciais. Se a hipótese for novamente rejeitada, isso indica que é necessário um número maior de iterações. Caso contrário, o número inicial de iterações descartadas é indicado como o tamanho do “burn-in” Considerando-se que uma seqüência de valores amostrados passou no teste de estacionaridade, aplica-se, nesta parte, o teste de Half-Width para verificar se a média estimada está sendo calculada com uma acurácia pré-especificada. A convergência fica garantida quando se aceita a hipótese nula nos dois testes.

O erro Monte Carlo foi calculado usando a biblioteca CODA do R e está baseado na estimativa da série temporal da cadeia ajustada. Assim, o cálculo corrige para quaisquer faltas de convergência devidas a autocorrelação na cadeia.

A decisão sobre o modelo mais adequado para cada resultado experimental foi feita utilizando-se o Fator de Bayes, segundo Kass e Raftery (1995). No entanto, as aproximações para as expressões utilizadas foram feitas de acordo com Raftery e Newton (2007).

Uma interpretação do fator de Bayes, sugerida por Kass e Raftery (1995), é apresentada na Tabela 1.

Tabela 1 - Interpretação do Fator de Bayes

FB	Evidência a favor do modelo do numerador
< 1	Negativa (modelo do denominador)
$1 \vdash 3$	Fraca
$3 \vdash 20$	Positiva
$20 \vdash 150$	Forte
> 150	Muito Forte

2.2.5 Distribuição conjunta *a posteriori*

A distribuição conjunta de todos os parâmetros dos modelos *a posteriori*, ou seja, dado que foi observado o resultado experimental, foi obtida a partir do Teorema de Bayes. Supondo independência entre as distribuições *a priori*, isso é feito multiplicando-se a verossimilhança e todas as distribuições *a priori* correspondentes, para cada modelo.

A distribuição conjunta *a posteriori* do modelo usual, considerando a função de verossimilhança definida na Equação 2 e a distribuição *a priori* definida na Equação 8, é dada por:

$$\begin{aligned}
 P(\lambda|Y) &\propto L(Y|\lambda)P(\lambda) \\
 P(\lambda|Y) &\propto \prod_{j=1}^k \binom{n_j}{y_j} (1 - e^{-\lambda z_j})^{y_j} e^{-\lambda z_j (n_j - y_j)} \\
 &\times \lambda^{\alpha-1} e^{-\frac{\lambda}{\beta}}.
 \end{aligned} \tag{10}$$

No modelo de toxidez, têm-se os parâmetros λ , cuja distribuição *a priori* está definida na Equação 8, e T_j com distribuição *a priori* descrita na seção 2.2.4. Sendo assim, a distribuição conjunta *a posteriori* para este modelo é definida como:

$$\begin{aligned}
P(\lambda, T_j|Y) &\propto L(Y|\lambda, T_j)P(\lambda)P(T_j) \\
P(\lambda, T_j|Y) &\propto \prod_{j=1}^k \binom{n_j}{y_j} (1 - e^{-\lambda z_j(1-T_j)})^{y_j} \\
&\times (e^{-\lambda z_j(1-T_j)})^{(n_j-y_j)} \times \lambda^{\alpha-1} e^{-\frac{\lambda}{\beta}} \times \prod_{j=1}^k \frac{1}{T_j-1}. \quad (11)
\end{aligned}$$

A obtenção da distribuição conjunta *a posteriori* do modelo de interferência é análoga aos modelos anteriores, ou seja, dadas as *prioris* nas Equações 8 e 9 para λ e λ_i e a verossimilhança, tem-se a conjunta dada pela equação:

$$\begin{aligned}
P(\lambda, \lambda_i|Y) &\propto L(Y|\lambda, \lambda_i)P(\lambda)P(\lambda_i) \\
P(\lambda, \lambda_i|Y) &\propto \prod_{j=1}^k \binom{n_j}{y_j} [e^{-\lambda_i z_j} (1 - e^{-\lambda z_j})]^{y_j} \times [1 - e^{-\lambda_i z_j} (1 - e^{-\lambda z_j})]^{(n_j-y_j)} \\
&\times \lambda^{\alpha-1} e^{-\frac{\lambda}{\beta}} \times \lambda_i^{\alpha_i-1} e^{-\frac{\lambda_i}{\beta_i}}. \quad (12)
\end{aligned}$$

Não sendo possível tomar amostras diretamente das distribuições condicionais completas *a posteriori* para qualquer dos parâmetros dos modelos utilizamos algoritmo de Metropolis-Hastings. Neste algoritmo, a cada iteração um valor candidato para cada parâmetro é amostrado da distribuição candidata especificada. Esse novo valor é aceito com base na probabilidade de aceitação que pode ser calculada com o método geral apresentado em Chib e Greenberg (1995).

2.2.6 Processo de geração das cadeias de Markov

No processo de geração das cadeias foi obtida uma amostra final de tamanho 5000 para cada parâmetro, após ter sido feito um descarte de 50000 valores iniciais, e um salto de 100 iterações. Esses valores são muito superiores aqueles indicados pelo diagnóstico de convergência de Raftery e Lewis (1992), como forma de garantir que todos os parâmetros dos modelos atingissem a convergência, o que se justifica pela velocidade de obtenção dos resultados.

A quantidade de parâmetros T está associada à quantidade de diluições existentes no experimento. Sendo assim estimou-se uma probabilidade de toxidez para cada diluição do experimento estudado.

A distribuição utilizada como geradora de valores candidatos para os parâmetros λ e λ_i é a gama com hiperparâmetros (α^c, β^c) e (α_i^c, β_i^c) para os respectivos parâmetros. Os parâmetros desta distribuição foram continuamente atualizados com uma amostra dos últimos 1000 valores da cadeia gerada. Este processo consistiu em calcular valores de α^c e β^c , segundo as definições de média e variância da distribuição gama que são, respectivamente: $E(Y) = \alpha^c \beta^c$ e

$V(Y) = \alpha^c(\beta^c)^2$. Assim, tem-se uma distribuição geradora de valores candidatos dinâmica, enquanto a distribuição *a priori* destes parâmetros é estática.

A distribuição geradora de valores candidatos para T foi a uniforme, e o parâmetro desta distribuição para a primeira diluição é $(0, 1)$. Os demais parâmetros estarão entre zero e o valor da probabilidade de toxidez da diluição anterior. Para maior simplicidade no processo amostral estas distribuições são as mesmas das *prioris* correspondentes e podem ser definidas da seguinte maneira:

$$\begin{aligned} T_1^c &\sim U[0, 1] \\ T_2^c|T_1^c &\sim U[0, T_1^c] \\ &\vdots \\ T_j^c|T_1^c, T_2^c \dots T_{j-1}^c &\sim U[0, T_{j-1}^c]. \end{aligned}$$

A probabilidade conjunta destes parâmetros é dada por:

$$P(T_1^c, T_2^c, \dots, T_j^c) = 1 \times \frac{1}{T_1^c} \times \frac{1}{T_2^c} \times \dots \times \frac{1}{T_{j-1}^c} = \prod_{j=1}^k \frac{1}{T_{j-1}^c}.$$

Todos os métodos foram implementados em **R** (R Development Core Team, 2008). O código fonte pode ser conseguido com o primeiro autor.

A avaliação de convergência das cadeias geradas para cada parâmetro foi feita usando os pacotes BOA (“Bayesian Output Analysis”) e CODA (“Convergence Diagnostics and Output Analysis”) do *software* R.

3 Resultados e discussão

Nas análises realizadas há muitas situações em que os modelos de interferência e toxidez resultaram em mais prováveis que os modelo usual. Uma análise detalhada destas situações e a geração de tabelas alternativas seria possível, mas foge ao escopo deste trabalho. São apresentados resultados que ilustram as propriedades do ajuste quando o modelo usual foi melhor e quando outros modelos devem ser utilizados.

A convergência, a estimação dos parâmetros e a comparação entre os modelos ajustados foram feitas utilizando-se dois resultados experimentais que contemplam situações comuns em ensaios de microbiologia de alimentos, sendo um favorável ao modelo usual e outro em que não se podem utilizar as tabelas do modelo usual.

3.1 Convergência do processo de Markov e estimação do número mais provável para o resultado experimental (32100)

O resultado experimental (32100) representa a situação em que o crescimento do microrganismo não é inibido em baixas diluições. Neste caso espera-se que o modelo usual seja o mais adequado para ajustar este tipo de resultado experimental.

O resultado obtido pelo critério de convergência de Raftery e Lewis (1992) encontra-se na Tabela 2. Por meio deste observa-se que os valores do fator de dependência para os parâmetros foram menores que 5, condição suficiente para a garantia da convergência das cadeias dos parâmetros. Além disso, os dois testes feitos no critério de Heidelberg e Welch (1983) reforçaram este resultado.

Tabela 2 - Critério de convergência de Raftery e Lewis para os parâmetros dos modelos considerando o resultado experimental (32100)

Modelo	Parâmetros	Fator de dependência
toxidez	λ	0,9823
	T_1	1,0152
	T_2	1,0320
	T_3	1,0320
	T_4	1,0152
	T_5	0,9823
interferência	λ	0,9823
	λ_i	0,9823
usual	λ	1,0557

A média *a posteriori*, o intervalo de credibilidade e o erro de Monte Carlo para o resultado experimental analisado (32100) encontram-se na Tabela 3. As diferenças numéricas nas estimativas do NMP (média *a posteriori*) são menos acentuadas entre os modelos de interferência e usual. Além disso, nota-se que o parâmetro λ do modelo de toxidez fica superestimado em relação ao mesmo parâmetro considerando os demais modelos. A média do parâmetro λ_i no modelo de interferência quando comparada com a média do parâmetro λ é pequena. O resultado experimental utilizado, sugere a ausência de mecanismo que esteja inibindo o crescimento do microrganismo de interesse em baixas diluições. Sendo assim, pode-se explicar o baixo valor da média para o parâmetro λ_i . De modo geral, os intervalos de credibilidade foram amplos, ressaltando-se que a menor amplitude é observada no modelo usual e a maior no modelo de toxidez. O erro de Monte Carlo, que representa a variação decorrente da aproximação numérica da distribuição de interesse, apresentou valores relativamente baixos. Considerando o modelo de interferência, o intervalo de credibilidade para o parâmetro λ_i incluiu o valor zero, sugerindo que este seja estatisticamente nulo. Essa hipótese pode ser testada usando alternativamente ao teste de hipóteses, o fator de Bayes, uma vez que este permite rejeitar o modelo em que esse parâmetro seja diferente de zero. Esses resultados podem ser verificados na Tabela 4.

Os resultados das comparações dos modelos por meio do fator de Bayes encontram-se na Tabela 4.

Os resultados do fator de Bayes são interpretados segundo a sugestão de Kass e Raftery (1995), contida na Tabela 1. Verifica-se na Tabela 4, que a evidência a favor do modelo de toxidez, quando comparado com o de interferência, é positiva,

Tabela 3 - Estimativas das médias *a posteriori*, intervalos de credibilidade (I.C.95%, LI: limite inferior, LS: limite superior) dos parâmetros dos modelos e erro de Monte Carlo (Erro M.C.) para o resultado (32100)

Modelos	Parâmetros	Média <i>a posteriori</i>	I.C. 95%		Erro M.C.
			LI	LS	
toxidez	λ	221,8047	31,9538	457,3796	2,1526
	T_1	0,6773	0,3546	0,9774	0,0027
	T_2	0,4623	0,1403	0,7950	0,0022
	T_3	0,2881	0,0422	0,5590	0,0019
	T_4	0,1602	0,0043	0,3425	0,0013
	T_5	0,0804	0,0000	0,2263	0,0011
interferência	λ	146,5478	6,2758	340,0585	2,0094
	λ_i	2,8909	0,0000	8,7767	0,0496
usual	λ	138,1516	14,6721	311,6698	2,3693

significando que o modelo de toxidez é mais adequado aos dados que o modelo de interferência. A comparação do modelo de toxidez, com o usual, mostra evidência negativa a favor do modelo de toxidez, ou seja, o fator de Bayes indica o modelo do denominador como mais adequado a essa situação. A mesma interpretação é válida na comparação entre o modelo de interferência e o usual. Sendo assim, verifica-se que, conforme esperado, o modelo mais provável, dado o resultado experimental (32100) é o usual, ou seja, esse modelo é mais adequado para esses dados. De acordo com essa última comparação, conclui-se que a hipótese de que parâmetro λ_i seja nulo, não poderia ser rejeitada.

Tabela 4 - Fator de Bayes aproximado e grau de evidência em favor do modelo da margem esquerda, em comparação ao modelo da margem superior, considerando o resultado experimental (32100)

Numerador	Denominador	
	interferência	usual
toxidez	$5,1553 \times 10^{0P}$	$8,1615 \times 10^{-2N}$
interferência		$1,5831 \times 10^{-2N}$

P: Positiva; N: Negativa.

3.2 Convergência do processo de Markov e estimação do número mais provável para o resultado experimental (01300)

O resultado experimental (01300) representa a situação na qual não há crescimento de microrganismos em baixas diluições.

Observa-se, na da Tabela 5, que a convergência das cadeias dos parâmetros dos modelos não foi satisfeita pelo diagnóstico de Raftery e Lewis (1992) apenas para parâmetros T_3 , por apresentar fator de dependência maior que 5. Esse resultado poderia sugerir a necessidade de um número maior de iterações ou a adoção de um valor maior de thinning para este parâmetro. Quanto ao critério de Heidelberger e Welch (1983) apenas para o parâmetro λ do modelo de interferência não se atingiu convergência. Uma avaliação visual das cadeias mostra, no entanto, estacionariedade e os critérios podem ser relativizados pois o maior reflexo é na precisão das estimativas, o que não é o objetivo imediato desta investigação.

Tabela 5 - Critério de convergência de Raftery e Lewis para os parâmetros dos modelos considerando o resultado experimental (01300)

Modelo	Parâmetros	Fator de dependência
toxidez	λ	0,9986
	T_1	1,0320
	T_2	1,1580
	T_3	6,1145
	T_4	3,7896
	T_5	3,4949
interferência	λ	1,0843
	λ_i	0,9986
usual	λ	1,0387

Como salientado anteriormente, os gráficos do traço para todos os parâmetros (figuras não apresentadas) sugerem estacionariedade e convergência das cadeias. Dessa forma, pode-se dizer que a convergência foi confirmada para todos os parâmetros, por pelo menos dois dos critérios considerados. Note-se que o programa é no entanto bem rápido e que poderiam ser empregados valores maiores de “thinning”, bem como maior número de iterações para satisfazer aspectos de precisão e garantia de convergência não seriam problema.

Os resultados referentes à média *a posteriori*, seus respectivos intervalos de credibilidade e erros de Monte Carlo para os parâmetros dos modelos encontram-se na Tabela 6.

Verifica-se, por meio da Tabela 6, que a estimativa do NMP (média *a posteriori* para o parâmetro (λ), tiveram valores diferente nos modelos utilizados. No entanto, nota-se que esta diferença é menos acentuada entre os modelos de toxidez e de interferência. No modelo usual, o λ está subestimado em relação ao mesmo parâmetro nos demais modelos. Esta é a consequência da utilização deste modelo a resultados experimentais em que há inibição do crescimento de microrganismos em baixas diluições, citada por Cochran (1950). Este modelo é pouco adequado para este tipo de resultado experimental.

O valor da estimativa do parâmetro (λ_i), reforça a idéia de que há mecanismos

Tabela 6 - Estimativas das médias *a posteriori*, intervalos de credibilidade (I.C.95%, LI: limite inferior, LS: limite superior), dos parâmetros dos modelos e erro de Monte Carlo (Erro M.C.) para o resultado (01300)

Modelos	Parâmetros	Média <i>a posteriori</i>	I.C. 95%		Erro M.C.
			LI	LS	
toxidez	λ	220,8000	35,9692	459,3319	2,2812
	T_1	0,9841	0,9490	0,9999	0,0003
	T_2	0,6637	0,3204	0,9691	0,0037
	T_3	0,2798	0,0339	0,5693	0,0026
	T_4	0,1697	0,0043	0,4145	0,0021
	T_5	0,0851	0,0000	0,2644	0,0015
interferência	λ	293,3000	17,0000	715,2454	3,8614
	λ_i	25,9300	4,1746	54,3104	0,2261
usual	λ	15,14	2,0610	30,7885	0,1998

perturbadores do modelo usual agindo no ensaio. Além disso, o valor do parâmetro T_1 é também indicio de que potencialmente há substância tóxica na primeira diluição. Estes resultados são coerentes em relação ao modelo experimental utilizado pois, de fato, não há crescimento de microrganismos em baixas diluições, indicando a presença de algum fator estranho no meio de cultura. Os intervalos de credibilidade dos parâmetros dos modelos são amplos, mas quando compara-se os modelos de toxidez e o de interferência verifica-se, que a amplitude é menor no modelo de toxidez. De modo geral, os erros de Monte Carlo são pequenos.

Os resultados das comparações dos modelos por meio do fator de Bayes encontram-se na Tabela 7.

Tabela 7 - Fator de Bayes aproximado e grau de evidência em favor do modelo da margem esquerda, em comparação ao modelo da margem superior, considerando o resultado experimental (01300)

Numerador	Denominador	
	interferência	usual
toxidez	$1,6001 \times 10^{-1N}$	$7,6592 \times 10^4 MF$
interferência		$4,7867 \times 10^5 MF$

MF: Muito forte; N: Negativa.

Segundo a sugestão contida na Tabela 1, verifica-se na Tabela 7, que a evidência a favor do modelo de toxidez, quando comparado com o de interferência, é negativa, ou seja o método indica, como mais adequado aos dados, o modelo de interferência. Nas comparações dos modelo de toxidez e interferência com o modelo usual, observa-se que em ambos casos a evidência é muito forte a favor do modelo

que está no numerador. Sendo assim, pode-se dizer que o modelo de interferência é o mais adequado para o ajuste deste resultado experimental.

Em resumo, na situação experimental em que há crescimento de microrganismo em baixas diluições, o modelo usual mostrou-se mais provável. No caso em que no resultado experimental o crescimento do microrganismo de interesse é inibido em baixas diluições, o modelo usual não é adequado e tem valores subestimados para o NMP. Neste caso o modelo de interferência é mais adequado.

Tais resultados estão de acordo com o esperado e permitem dizer que a rotina pode ser empregada em substituição às tabelas que apenas contemplam o modelo usual. Além da possibilidade de estimar a toxidez e a interferência, o método considerado permite a construção do intervalo de credibilidade, em substituição ao intervalo de confiança simétrico, obtido assintoticamente usando aproximação normal.

É importante ressaltar que, embora tenham sido considerados apenas resultados experimentais envolvendo cinco diluições com três tubos para cada diluição, é possível obter estimativas para outras combinações de diluições e tubos. Essa é uma outra desvantagem do uso das tabelas, que consideram um número restrito de situações e delineamentos.

Conclusões

Os resultados permitem concluir que a rotina implementada é confiável e que pode ser estendida para uma ampla gama de planos de diluição, para os quais nem sempre é adequado o modelo usual. A utilização do Fator de Bayes viabilizou testar a existência de interferência ou toxidez, em prejuízo da utilização do modelo usual. O método proposto pode ser utilizado para verificar a possibilidade de se utilizar as tabelas de laboratório e, caso não se possa utilizá-las, empregar as estimativas do modelo de interferência ou do modelo de toxidez em substituição às das tabelas usuais.

DELFINO, A. C. S.; SILVA, M. I. S.; PICCOLI, R. H.; BUENO FILHO, J. S. S. Bayesian analysis of interfering and toxicity effects in dilution assays. *Rev. Bras. Biom.*, São Paulo, v.29, n.2, p.325-341, 2011.

■ **ABSTRACT:** *Estimation of microorganism densities by means of the Most Probable Number (MPN) is a technique introduced by McCrady (1915) to analyse serial dilution assays. The standard model used to generate MPN tables does not consider medium toxicity nor interference due to competitor microorganisms. In this work we aim to develop a Bayesian framework to analyze these phenomena. MCMC methods using Metropolis-Hastings algorithm were used to get posterior distributions given some experimental results. Convergence was monitored using graphical display and both Raftery e Lewis (1992) and Heidelberg e Welsh (1983) criteria. Model comparison was done using Bayes Factors. It was possible to sort out models with interfering and toxicant parameters that were more probable than standard model for some experimental results.*

When microorganism do not grow in initial dilutions, the standard model underestimates MPN. In the situations in which standard model is the most probable, MPN estimates from any model are similar, although standard model is the best with smaller credibility interval. A very flexible **R** routine was implemented. It can manage a wide range of dilution designs with more dilutions and more tubes per dilution and is a suitable tool for replacing standard tables in laboratory.

■ **KEYWORDS:** MCMC; bayesian inference; dilution assay; Metropolis-Hastings algorithm; microbiology; most probable number.

Referências

- AGÊNCIA NACIONAL DE VIGILÂNCIA SANITÁRIA. ANVISA. Resolução-RDC n. 12, de 2 de jan. de 2001. Disponível em: <http://www.anvisa.gov.br/legis/resol/12_01rdc.-htm>. Acesso em: 02 mar. 2007.
- BLODGETT, R. J.; GARTHRIGHT, W. E. Several MPN models for serial dilution with suppressed growth at low dilutions. *Food Microbiol.*, London, v.15, n.1, p.91-99, 1998. Disponível em: <<http://dx.doi.org/10.1006/fmic.1997.0144>>.
- CHIB, S.; GREENBERG, E. Understanding the Metropolis-Hasting algorithm. *Am. Stat.*, Alexandria, v.49, n.4, p.327-335, 1995.
- COCHRAN, W. G. Estimation of bacterial densities by means of the most probable number. *Biometrics*, Oxford, v.6, n.2, p.105-116, 1950.
- GAMERMAN, D. *Markov chain monte Carlo*: Stochastic simulation for bayesian inference. London: Chapman & Hall, 1997. 245p.
- GELMAN, A. CARLIN, J. B. STERN, H. S.; RUBIN, D. B. *Bayesian data analysis*. London: Chapman & Hall, 2000. 526p.
- HEIDELBERGER, P.; WELCH, P. Simulation run length control in the presence of an initial transient. *Oper. Res.*, Hanover, v.31, n.6, p.1109-44, 1983.
- KASS, R. E.; RAFTERY, A. E. Bayes factors and model uncertainty. *J. Am. Stat. Assoc.*, Schaumburg, v.90, p.773-775, 1995.
- McCRADY, M. H. The numerical interpretation of fermentation-tube results. *J. Infect. Dis.*, Oxford, v.17, n.1, p.183-212, 1915.
- RAFTERY, A. E.; LEWIS, S. How many iterations in the Gibbs sampler? In: BERNARDO, J. M.; BERGER, J. O.; DAWID, A. P.; SMITH, A. F. M. (Ed.). *Bayesian statistics 4*. Oxford: University Press, 1992. v.4, p.763-773.
- RAFTERY, A. E.; NEWTON, A. M. Estimating the integrated likelihood via posterior simulation using the harmonic mean identity. *Bayesian statistics*. Oxford: University Press, 2007. v.8, p.1-45.
- R Development Core Team (2008). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL <http://www.R-project.org>.

SANTOS, A. C. *Análise bayesiana de interferência e toxidez em ensaios de diluição serial*. 2008. 97f. Tese (Doutorado em Estatística e Experimentação Agropecuária) – Universidade Federal de Lavras, Lavras, 2008.

SILVA, N.; JUNQUEIRA, V. C. A.; SILVEIRA, N. F. A. *Manual de métodos de análise microbiológica de alimentos*. São Paulo: Varela, 1997. 295p.

Recebido em 22.11.2010.

Aprovado após revisão em 06.09.2011.