Estimation of in vitro bactericidal potency based on colony counting method

Máté Szalai,[1*] **and Péter Kevei**[1]

[1]*Bolyai Institute, University of Szeged, Aradi vértanúk tere 1, 6720 Szeged, Hungary*

Abstract: To model the growth of a bacterial population in the presence of antibiotics we use the stochastic model from Bogdanov et al. [2]. We assume that bacterial cells either die or duplicate, with probabilities $p_0(c)$ and $p_2(c)$, where $p_2(c) = 1/(1 + \alpha c^{\beta})$ for some α, β, where c stands for the antibiotic concentration. Using measurements based on colony counting method we obtain weakly consistent, asymptotically normal estimator both for (α, β) and for the minimal inhibitory concentration (MIC), a relevant parameter in pharmacology.

Keywords: Galton–Watson process, extinction probability, asymptotically normal estimator, MIC

AMS subject classification: 92C40, 60J80

1 Introduction

The correct estimation of bactericidal potency is a critical issue for the safe and proper use of antibiotics. In Bogdanov et al. [2] we worked out a Bienaymé–Galton–Watson branching model for the growth of the bacterial population, and we obtained weakly consistent asymptotically normal estimators for the relevant parameters when for the biological measurements quantitative PCR (qPCR) method is used. In [2] we found that the 2-parameter model fits very well to real biological data. In the present note we provide an estimator under the same model assumptions but for different biological data: we assume that the experimental data was obtained using colony counting method. The qPCR method measures the total bacterial genom, which is the total number of *dead and alive* bacterial cells multiplied by a constant. On the other hand, colony counting gives an estimator for the extinction probability. The basic experiment is the following. Originally, x_0 bacterial cells (e.g. *Escherichia coli*) are inoculated onto agar plates containing a series of antibiotic concentration, and after the incubation period all the viable colonies are enumerated, see e.g. Liu et al. [1].

As in [2] we assume that the bacterial population is homogeneous, in particular, there is no resistant type. Long-term evolution of bacterial populations with both *resistant* and *susceptible* types was investigated in several papers using deterministic models, see Svara and Rankin [4], Paterson et al. [3], and the references therein. Closest to our model is the deterministic model given by Liu et al. [1], where the biological measurements were obtained by colony counting. In [1] a deterministic expression for the number of colony forming units was obtained in terms of the antibiotic concentration.

*Corresponding author: szalaim@math.u-szeged.hu

Next we describe the mathematical model. We consider a simple Galton–Watson branching process where each bacterium either dies (leaves no offspring) or divides (leaves 2 offsprings) with respective concentration dependent probabilities $p_0 = p_0(c)$ and $p_2 = p_2(c) = 1 - p_0(c)$. Let $f(s) = f_c(s) = p_0 + p_2 s^2$ denote the offspring generating function and $m = m(c) = 2p_2(c)$ the offspring mean if the antibiotic concentration is c. The process starts with a single ancestor $X_{0;c} = 1$, and

$$X_{n+1;c} = \sum_{i=1}^{X_{n;c}} \xi_{i;c}^{(n)},$$

where $\{\xi_c, \xi_{i;c}^{(n)} : i \geq 1, n \geq 1\}$ are independent and identically distributed (iid) random variables with generating function f_c. We further assume that the offspring distribution is given by

$$p_2(c) = \frac{1}{1 + \alpha c^\beta}, \tag{1}$$

where $\alpha > 0$, $\beta > 0$ are unknown parameters. Note that as $m = 2p_2$ this is the same assumption as in [2]. Under this model the minimal inhibitory concentration (MIC), the smallest antibiotic concentration preventing bacterial growth, is the smallest c for which $m(c) = 1$, that is $\alpha^{-1/\beta}$.

If $m \leq 1$ then the process dies out almost surely, while if the process is supercritical, i.e. $m > 1$ then the probability of extinction is the smaller root of $f_c(q) = q$, which is in our setup

$$q(c) = \begin{cases} \frac{1-p_2(c)}{p_2(c)}, & \text{if } p_2(c) > 1/2, \\ 1, & \text{if } p_2(c) \leq 1/2. \end{cases} \tag{2}$$

2 Estimation of the parameters

Assume that the initial number of bacterial cells is x_0, that is we observe x_0 independent copies of the Galton–Watson process $(X_{n;c})$. Then the number Y_c of living colonies has binomial distribution with parameters x_0 and $1 - q(c)$. Therefore, the natural estimator for $q(c)$ is $\widehat{q}(c) = 1 - \frac{Y_c}{x_0}$. The law of large numbers and the central limit theorem implies that $\widehat{q}(c)$ is a weakly consistent estimator and as $x_0 \to \infty$

$$\frac{\sqrt{x_0}}{\sqrt{q(c)(1-q(c))}} \left(\widehat{q}(c) - q(c)\right) \xrightarrow{\mathcal{D}} \mathcal{N}(0,1), \tag{3}$$

where $\xrightarrow{\mathcal{D}}$ stands for convergence in distribution.

From (2) we see that we can estimate $p_2(c)$ only if $q(c) < 1$, or equivalently $m(c) > 1$, in which case

$$\widehat{p_2}(c) = \frac{1}{1 + \widehat{q}(c)}. \tag{4}$$

We assume that the offspring mean as a function of c satisfies (1) for some unknown parameters $\alpha > 0$, $\beta > 0$. Rewriting (1)

$$\log \alpha + \beta \log c = \log\left(\frac{1}{p_2(c)} - 1\right).$$

Assume that we have measurements for $K \geq 2$ different concentrations $c_1 < c_2 < \ldots < c_K$, such that $m(c_K) > 1$. As in (4), we obtain the estimator $\widehat{p}_2(c_i)$ at different concentrations, from which, using simple least squares estimator we obtain the estimator

$$\widehat{\beta} = \frac{K\sum_{i=1}^K f_i\ell_i - \sum_{i=1}^K f_i L_1}{KL_2 - L_1^2},$$
$$\widehat{\alpha} = \exp\left\{\frac{\sum_{i=1}^K f_i - \widehat{\beta}L_1}{K}\right\},$$

where to ease notation we write

$$f_i = \log\left(\frac{1}{\widehat{p}_2(c_i)} - 1\right), \quad \ell_i = \log c_i,$$

and $L_1 = \sum_{i=1}^K \ell_i$, $L_2 = \sum_{i=1}^K \ell_i^2$. By the Cauchy–Schwarz inequality the denominator of $\widehat{\beta}$ is strictly positive for $K \geq 2$.

Under the assumption (1) the MIC equals $\vartheta = \alpha^{-1/\beta}$, therefore its natural estimator is

$$\widehat{\vartheta} = \widehat{\alpha}^{-1/\widehat{\beta}}.$$

Using (3), as in [2] we can prove that these estimators are asymptotically normal. Introduce the notation

$$k_i = \frac{p_2(c_i)}{1 - p_2(c_i)}\sqrt{q(c_i)(1 - q(c_i))}, \quad i = 1, 2, \ldots, K.$$

Proposition 1. *Assume that $c_1 < \ldots < c_K$ are given concentrations such that $m(c_K) > 1$. Then as $x_0 \to \infty$, $\widehat{\alpha}, \widehat{\beta}$, and $\widehat{\vartheta}$ are weakly consistent estimators of the corresponding quantities. Furthermore, as $x_0 \to \infty$*

$$\sqrt{x_0}(\widehat{\alpha} - \alpha, \widehat{\beta} - \beta) \xrightarrow{\mathcal{D}} (U, V),$$

where (U, V) is a two-dimensional normal random vector with mean 0 and covariance matrix $\begin{pmatrix} \sigma_\alpha^2 & \sigma_{\alpha\beta} \\ \sigma_{\alpha\beta} & \sigma_\beta^2 \end{pmatrix}$, where

$$\sigma_\alpha^2 = \frac{\alpha^2}{(KL_2 - L_1^2)^2}\sum_{i=1}^K k_i^2(L_2 - L_1\ell_i)^2,$$
$$\sigma_{\alpha\beta} = \frac{\alpha}{(KL_2 - L_1^2)^2}\sum_{i=1}^K k_i^2(K\ell_i - L_1)(L_2 - L_1\ell_i),$$
$$\sigma_\beta^2 = \frac{1}{(KL_2 - L_1^2)^2}\sum_{i=1}^K k_i^2\,(K\ell_i - L_1)^2\,,$$

and $\sqrt{x_0}(\widehat{\vartheta} - \vartheta) \xrightarrow{\mathcal{D}} \mathcal{N}(0, \sigma_\vartheta^2)$ as $x_0 \to \infty$, with

$$\sigma_\vartheta^2 = \frac{\vartheta^2\,(\log\alpha)^2}{\beta^2(KL_2 - L_1^2)^2}\sum_{i=1}^K k_i^2\left(\frac{L_2 - L_1\ell_i}{\log\alpha} - \frac{K\ell_i - L_1}{\beta}\right)^2.$$

3 Simulation study

If $m(c_K) > 1$, then regardless of the fixed values $\mathbf{c} = (c_1, \ldots, c_K)$ the estimate $(\widehat{\alpha}, \widehat{\beta})$ is weakly consistent and asymptotically normal as $x_0 \to \infty$. However, the asymptotic variances in Proposition 1 do depend on the specific choice of $K \geq 2$ and the values $c_1 < \ldots < c_K$. Intuitively, it is clear that we should choose values for the concentrations where the derivative of m is large, that is m is close to 1, see Figure 1.

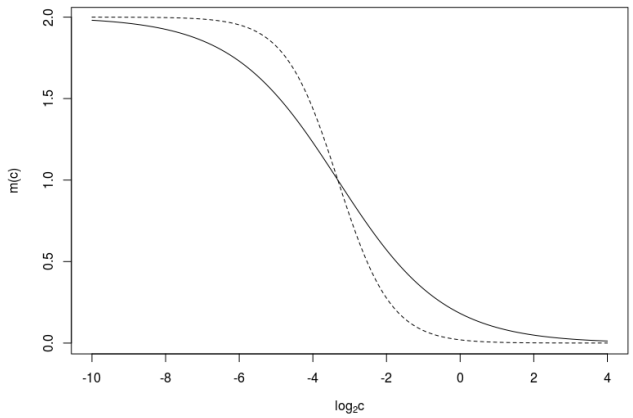

Figure 1: $m(c)$ in a logarithmic scale (solid $(\alpha, \beta) = (10, 1)$, dashed $(\alpha, \beta) = (100, 2)$)

As in [2] we compare two rather different biologically relevant scenarios: $(\alpha, \beta) = (10, 1)$ and $(\alpha, \beta) = (100, 2)$. In Figure 1 we see the mean function for these two cases. Note that in both cases $\vartheta = 0.1$. Table 1 contains the theoretical variances given in Proposition 1 for different choices of the concentrations. For the steeper function ($(\alpha, \beta) = (100, 2)$) the variances of α and β are significantly larger, however the variance of the MIC is of the same order. We also see that a wrong choice of the concentrations might result much larger variations. For $\mathbf{c}_3$ all the concentrations are small, the antibiotic does not have any effect, so we cannot make a good estimate from observations at these concentrations.

concentrations	σ^2_{10}	σ^2_1	$\sigma^2_{0.1}$	σ^2_{100}	σ^2_2	$\sigma^2_{0.1}$
$\mathbf{c}_1 = (2^{-7}, 2^{-4})$	2424	2.87	0.015	$2.98 \cdot 10^6$	38	0.0027
$\mathbf{c}_2 = (2^{-5}, 2^{-4.5}, 2^{-3.4})$	875	1.36	0.0016	$3.54 \cdot 10^5$	5.5	0.0014
$\mathbf{c}_3 = (2^{-9}, 2^{-8}, 2^{-7})$	$8.99 \cdot 10^4$	32	2.89	$3.84 \cdot 10^8$	1448	29

Table 1: Asymptotic variances for $(\alpha, \beta) = (10, 1)$ and $(\alpha, \beta) = (100, 2)$.

Choosing the right antibiotic concentration is important to get a good estimate. The larger variances above are not surprising, because in the present setup the estimator for the mean $m(c)$ works only for supercritical processes, that is for those c, for which $m(c) > 1$. That is we can sample only from the upper part of the mean function $m(c)$ in Figure 1. This is in sharp contrast to the situation treated in [2], where the total number of dead and alive bacteria was counted, and the estimator for the mean works for any c.

x_0	$\overline{\alpha}$	$\overline{\beta}$	$\overline{\vartheta}$	$\widehat{\sigma}^2_\alpha$	$\widehat{\sigma}_{\alpha,\beta}$	$\widehat{\sigma}^2_\beta$	$\widehat{\sigma}^2_\vartheta$
50	11.25	1.01	0.101	1464	41	1.3	0.002
100	10.79	1.01	0.1004	1349	43	1.48	0.0019
300	10.23	1.003	0.1	981	36.2	1.44	0.0018
500	10.17	1.003	0.1	931	34.9	1.34	0.0016
∞	10	1	0.1	875	34	1.36	0.0017

Table 2: Empirical mean and variances for $(\alpha, \beta) = (10, 1)$.

With $\alpha = 10$, $\beta = 1$ and concentration vector $\mathbf{c}_2$ we simulate the process as follows. For a given concentration c_k, $k = 1, \ldots, K$, we calculate $p_2(c_k)$ from (1). From each measurement we calculate the estimation $(\widehat{\alpha}, \widehat{\beta})$ as described in (2). We simulated the measurements 1000 times. The resulting means and empirical variances of $\sqrt{x_0}(\widehat{\alpha}-\alpha, \widehat{\beta}-\beta)$ and $\sqrt{x_0}(\widehat{\vartheta}-\vartheta)$ are given in Table 2. We see that even for small initial number of bacteria the empirical variances are close to the theoretical counterparts.

Bibliography

[1] Y. Q. Liu, Y. Z. Zhang, and P. J. Gao. *Novel concentration-killing curve method for estimation of bactericidal potency of antibiotics in an in vitro dynamic model.* Antimicrobial. Agents and chemotherapy 48(10):3884-3891, 2004.

[2] A. Bogdanov, P. Kevei, M. Szalai, and D. Virok. *Stochastic modeling of in vitro bactericidal potency.* https://arxiv.org/abs/2104.11525, 2021.

[3] I.K. Paterson, A. Hoyle, G. Ochoa, C. Baker-Austin, and N.G.H. Taylor. *Optimising Antibiotic Usage to Treat Bacterial Infections.* Sci. Rep. 6(37853), 2016.

[4] Fabian Svara, and Daniel J. Rankin. *The evolution of plasmid-carried antibiotic resistance.* BMC Evolutionary Biology, 11(130), 2011.