Estimation of Mixing Index and Contact Number by Coordination Number Sampling in an Incompletely Mixed State

Hisakazu Shindo
Department of Computer Science, Faculty of Engineering, Yamanashi University*

Abstract

A statistical model of the distribution of the contact number for a binary mixture in an incompletely mixed state is introduced. The model is derived from the beta-binomial model proposed by Yoshizawa and Shindo3). The mixing index based on the contact number is defined and described by only one parameter in this model.

The precision of the estimation of the mixing index, based on the contact number by coordination number sampling in an incompletely mixed state, has been derived and simulated on a computer.

1. Introduction

Many of conventional solid mixes indicates their mixing indexes based on the variance of target components among spot samples. However, the variance is macroscopic and cannot provide any definite physical meaning, though it is very important statistically. Thus, it is difficult to get a definite meaning of variance when a mixing index is applied to the microscopic analysis of a reaction rate in a solid reaction.

From the above viewpoint, Akao, Fan and others have proposed the representation of a mixing index with a contact number. The purpose of this is to estimate the number of contacts in a mixed substance to represent the mixing index by a sampling method, called coordination number sampling, which measures the number of target component particles in contact with specific component particles extracted randomly. The mixing index based on the contact number, which is both microscopic and geometric, can provide a definite physical meaning.

Yoshizawa and Shindo3) have proposed a model based on the β-binomial model to describe any mixed state of a binary mixture and investigated the precision of the estimation of the Lacey’s mixing index4) in an incompletely mixed state as well as the relationship of the model with the Polya urn scheme5).

We found the distribution of the contact number in an incompletely mixed state, up until now still unknown, with the same model configuration as that of Yoshizawa and Shindo, discussed about an estimation method for the population contact number and investigated the validity of the theory through computer simulation. The following is a description of our findings.

2. Distribution of contact number and its mixing index

2.1 Completely segregated state and completely random state

Now, we consider a mixture of a two-dimensional regular array consisting of two kinds of equal diameter particles, \( A_0 \) and \( A_1 \). Their population concentration is assumed to be \( X_0 \) and \( X_1 (=1 - X_0) \), respectively. A particle extracted randomly from the mixture is called a sample particle and the total number of particles in contact with the sample particle is called the total coordination number \( n^* \). Total coordination number \( n^* \), which is determined by

* 4-3-11, Takeda, Kofu-shi, 400
TEL. 0552-52-1111
† This report was originally printed in J. Soc. Powder Technology, Japan, 24, 521-526 (1987) in Japanese, before being translated into English with the permission of the editorial committee of the Soc. Powder Technology, Japan.
Fig. 1 Illustration of a coordination number sampling of size $n^* = 6$

**Fig. 2** Distribution of the contact number and the mixing index $M$ with the total coordination number $n^* = 4$ and the population concentration $X_1 = 0.5$
In secondary moment, contact number mixing index \( M \) is based on the expected value or primary moment, as known from Eq. (7).

2.3 Incompletely mixed state

In an incompletely mixed state, the local concentration of the target particles is largely non-uniform and the concentration of the entire mixture is not capable of being uniform. Thus, it should be considered that the local concentration itself is a random variable following a certain distribution. Therefore, we assume the random variable \( z \) of the local concentration and the probability density function \( f(z) \), which has the \( \beta \) density with parameters \( \tau \) and \( \bar{X}_1 \) as expressed by:

\[
f(z) = \frac{z^{\tau \bar{x}_1-1}(1-z)^{\tau(1-\bar{x}_1)-1}}{B(\tau \bar{x}_1, \tau(1-\bar{x}_1))}
\]  

where \( \tau \) is the positive parameter and \( B(\cdot, \cdot) \) is the \( \beta \) function.

Now, under the condition that local concentration \( Z \) is known, the probability that the sample particle is \( A_0 \) and the \( y \) particles in the \( n^* \) particles around the sample particle are \( A_1 \) particles is expressed by the equation,

\[
P_r [Y = y, A_0 | z] = \binom{n^*}{y} z^y (1-z)^{n^*-y+1}
\]  

where \( Y \) is the random variable against \( y \). Therefore, \( P_r [Y = y, A_0] \) can be expressed by the equation below as the marginal probability of the simultaneous probabilities in Eqs. (8) and (9):

\[
P_r [Y = y, A_0] = \int_0^1 P_r [Y = y, A_0 | z] f(z) \, dz
\]

This gives us the probability that the \( y \) particles are in contact with the sample particle which is assumed to be an \( A_0 \) particle to be expressed by the following equation:

\[
P_r [Y = y | A_0] = P_r [Y = y, A_0] / P_r [A_0]
\]

where \( P_r [Y = y, A_0] = \binom{n^*}{y} z^y (1-z)^{n^*-y+1} \),

\[
\times \frac{z^{\tau \bar{x}_1-1}(1-z)^{\tau(1-\bar{x}_1)-1}}{B(\tau \bar{x}_1, \tau(1-\bar{x}_1))} \, dz
\]

and \( y = 0, 1, 2, \ldots, n^* \)

As a result of \( P_r [Y = y | A_0] = P_r [C_{1(o)} = y] \), the distribution presented by Eq. (11) indicates the contact number distribution in an incompletely mixed state.

In contact number distribution Eq. (11), \( i \)-th factorial moment \( \mu_{(i)} \) of \( C_{1(o)} \) can be:

\[
\mu_{(i)} = n^* \bar{x}_1 \tau / (\tau + 1)
\]

and variance \( \sigma^2 \) of the contact number \( C_{1(o)} \) in an incompletely mixed state to be determined by the following equations, respectively:

\[
\mu = \mu_{(1)} = n^* \bar{x}_1 \tau / (\tau + 1)
\]

\[
\sigma^2 = \mu_{(2)} - \mu_{(1)}^2
\]

The limits of the expected value and variance when parameter \( \tau \) is made close to zero or infinity are,

\[
\lim_{\tau \to 0} \mu = 0 = \mu_s \quad \lim_{\tau \to 0} \sigma^2 = 0 = \sigma_s^2
\]

\[
\lim_{\tau \to \infty} \mu = n^* \bar{x}_1 = \mu_r \quad \lim_{\tau \to \infty} \sigma^2 = n^* \bar{x}_1 (1-\bar{x}_1) = \sigma_r^2
\]

The changes when parameter \( \tau \) is altered from zero to infinity corresponds to the status that the distribution of contact number \( C_{1(o)} \) moves gradually from the one-point distribution of the completely segregated state in Eq. (1) to the binomial distribution of the completely random state in Eq. (4), as shown in Fig. 2.

Substituting the expected value (13) of contact number \( C_{1(o)} \) in an incompletely mixed state for the mixture defining Eq. (7), we come up with the following equation associating mixture \( M \) with parameter:
This relationship coincides with the relationship between Lacey’s mixing index and the parameter \( r \) in the \( \beta \)-binomial model proposed by Yoshizawa and Shindo.

### 3. Estimation of mixing index and its accuracy

We now consider the problem that the mixture is estimated with \( k \) coordination number samples of size \( n^* \) from the mixture including target particles of population concentration \( X_1 \). In the mixing problem, \( X_1 \) is assumed to be already known because it is often given when the mixture is supplied. This causes the denominator of the mixture defining Eq. (7) to be constant. At last, we may estimate the contact number through the equation,

\[
k_c = \frac{\sum_{i=1}^{k} c_i}{k}
\]

then substitute it for Eq. (7).

From Eqs. (13) and (14), expected value \( E(\hat{c}) \) and variance \( V(\hat{c}) \) of the estimate of the contact number \( \hat{c} \) can be expressed as follows:

\[
E(\hat{c}) = n^* \bar{X}_1 \frac{\tau}{(\tau + 1)}
\]

\[
V(\hat{c}) = \frac{1}{k} \cdot \frac{\tau}{\tau + 1} \cdot n^* \bar{X}_1 \left( 1 - \frac{\tau}{\tau + 1} \bar{X}_1 \right) \times \frac{\tau + 1 + n^*}{\tau + 2}
\]

Equation (19) indicates that \( \hat{c} \) is an unbiased estimator. Depending on the above equations, standard error \( D(\hat{c}) \) and relative standard error \( RSE(\hat{c}) \) of the estimate \( \hat{c} \) can be expressed by the following equations, respectively:

\[
D(\hat{c}) = \sqrt{V(\hat{c})} = \left[ \frac{1}{k} \cdot \frac{\tau}{\tau + 1} \cdot n^* \bar{X}_1 \times \left( 1 - \frac{\tau}{\tau + 1} \bar{X}_1 \right) \frac{\tau + 1 + n^*}{\tau + 2} \right]^{1/2}
\]

\[
RSE(\hat{c}) = \frac{D(\hat{c})}{E(\hat{c})} = \left[ \frac{(\tau + 1 - \tau \bar{X}_1)\left( (\tau + 1 + n^*) \bar{X}_1 \right)}{kn^* \bar{X}_1 \tau(\tau + 2)} \right]^{1/2}
\]

To find the number of samples \( k \) satisfying the specified relative standard error \( RSE(\hat{c}) \) in the sampling design, we may use the following equation, which results from solving Eq. (22) regarding \( k \):

\[
k = \frac{(\tau + 1 - \tau \bar{X}_1)\left( (\tau + 1 + n^*) \bar{X}_1 \right)}{n^* \bar{X}_1 \tau(\tau + 2)\left( RSE(\hat{c}) \right)^2}
\]

Equations (21), (22), and (23) are equations to find the standard error, relative standard error and the number of samples \( k \) of estimate \( \hat{c} \) in the completely random state. They are presented as the limits when parameter \( \tau \) is made close to infinite, as shown below:

\[
D(\hat{c}) = \left[ n^* \bar{X}_1 (1 - \bar{X}_1) \right]^{1/2}
\]

\[
RSE(\hat{c}) = \left[ \frac{(1 - \bar{X}_1) / kn^* \bar{X}_1}{1 - \bar{X}_1} \right]^{1/2}
\]

\[
k = \left( 1 - \bar{X}_1 / n^* \bar{X}_1 \right) RSE(\hat{c})^2
\]

These equations can be obtained by estimate Eq. (5) and variance Eq. (6) of the binomial distribution Eq. (4), which is the distribution of the contact number in the completely random state.

### 4. Results and discussions of computer simulation

To determine the validity of estimate accuracy Eqs. (22) and (25), we carried out computer simulation. As shown in Fig. 2, the distribution of the contact number is a discrete distribution. Using the following recurrent equation which satisfied between adjacent probabilities,

\[
P_j(C_1(o) = c_1(o) + 1) = P_{j-1}(C_1(o) = c_1(o)) \cdot \frac{c_1(o) + \tau \bar{X}_1}{c_1(o) + 1} \cdot \frac{n^* - c_1(o)}{n^* - c_1(o) + \tau (1 - \bar{X}_1)}
\]

we found the accumulated probability and generated random numbers obeying Eq. (11) with the reverse conversion method.

We made a simulation using the following procedure:

(i) First mixture \( M \), population concentration \( \bar{X}_1 \) and sample size \( n^* \) were specified.

(ii) Next, the number of samples \( k \) to a certain value were set.

(iii) Then, assuming the \( j \)-th sample in the \( i \)-th count to be \( c_{ij} \), mean contact number \( \hat{c}_i \) \((i = 1, \cdots, N)\) with the following equation was estimated:

\[
\hat{c}_i = \frac{\sum_{j=1}^{k} c_{ij}}{k}
\]

(iv) Next, average \( \bar{c} \) and standard error \( D(\hat{c}) \) of
the mean contact number of $N$ particles using the following respective equations were calculated:

$$\tilde{c} = \frac{1}{N} \sum_{i=1}^{N} \hat{c}_i$$  \hspace{1cm} (29)

$$D(\hat{c}) = \left\{ \frac{1}{N} \sum_{i=1}^{N} (\hat{c}_i - \frac{1}{N} \sum_{i=1}^{N} \hat{c}_i)^2 / (N - 1) \right\}^{1/2}$$  \hspace{1cm} (30)

(v) Then, relative standard error $RSE(c_{i(0)})$ by $D(\hat{c})/\tilde{c}$ could be estimated.

Figure 3 (a) plots the relationship of relative standard error $RSE(c_{i(0)})$ with the number of samples $k$ on log-log graphs in the completely random state with total coordination number $n^* = 4$ and population concentration $\bar{X}_1 = 0.1$ or 0.5. The solid lines indicate the theoretical values calculated from Eq. (25), which are just suited to the simulation results. Since the lines with population concentration $\bar{X}_1 = 0.5$ are always drawn below those with $\bar{X}_1 = 0.1$, the relative standard error is smaller to have higher estimate precision as the population concentration $\bar{X}_1$ is closer to 0.5.

As a sample result in an incompletely mixed state, Fig. 3 (b) shows the same relationship as in Fig. 3 (a) with mixture $M = 0.5$. The solid lines in the figure are the theoretical values found from Eq. (22). The relative standard error is larger to be less estimate precision in comparison with Fig. 3 (a).

Actually, it is better to estimate the contact number by the coordination number sampling for an appropriate number of samples prior to mixing. Figure 4 shows the relationship of the relative standard error with mixture for population concentration $\bar{X}_1 = 0.1$ and 0.5 with the number of samples $k$ fixed to 20 and total coordination number $n^* = 4$. This indicates that the relative standard error is decreased gradually to make the estimate precision higher prior to mixing.

In designing sampling operation, it is easy to use Figs. 3(a) and 3(b), while it can be designed by using Eqs. (23) and (26). Table 1 shows the results of calculating the number of samples using Eqs. (23) and (26) to estimate the contact number with the relative standard error of 10% in two cases of the population concentration $\bar{X}_1$ of 0.1 and 0.5 with the total coordination number $n^* = 4$. This also indicates that the estimate precision is higher and the number of samples may be less as population concentration $\bar{X}_1$ becomes closer to 0.5 and the mixture becomes closer to the completely random state.

5. Conclusion

Akao and others have already proposed a method to indicate mixture using the contact number. However, with this method, we could not discuss the estimate precision because it did not clarify the distribution of the contact number in an incompletely mixed state.

This report introduced the theoretical distri-
bution of the contact number in an incom­
pletely mixed state through the similar prere­
quisite as the β-binomial model proposed by
Yoshizawa and Shindo, investigated the esti­
mate precision of the contact number based on
the model and indicated examples of the sam­
ping designs.

In estimation of the contact number, spot
sampling can be used instead of coordination
number sampling. The spot sampling method
has already been reported.

Acknowledgments

The information in this report was created
during the research work for the "Study on
Solids Mixing", a Japan-U.S. joint scientific co­
operation project of the Japan Society for the
Promotion of Science carried out for two years
from 1978. The author would like to thank
Prof. Y. Akao of Yamanashi University (who is
now Professor of Tamagawa University) and
Prof. T. Yoshizawa of Yamanashi University
(who is now Professor of the University of
Tsukuba), Japanese representatives of the pro­
ject for their guidance. The author also would
like to thank Prof. L.T. Fan of the Kansas
State University, Asst. Prof. F.S. Lai, and Mr.
J.R. Too for their useful discussions and help
in creating this report.

Nomenclature

\( A_0, A_1 \) : kinds of particles

\( B \) : beta function

\( C_{i(0)}, \hat{C}_{i(0)} \): random variable denoting the
contact number and its value

\( c_{ij} \): the \( j \)-th sample used in the \( i \)-th esti­
mentation

\( \hat{c} \): estimate of the contact number

\( \hat{c}_i \): the \( i \)-th estimate of the contact
number

\( \bar{c} \): mean of the \( N \) estimates

\( D \): standard error

\( D(\bar{c}) \): standard error of the \( N \) estimates

\( E \): expectation operator

\( f(z) \): probability density function of \( Z \)

\( k \): number of samples

\( M \): mixing index

\( N \): number of iterations of the
simulation

\( n^* \): total coordination number or
sample size

\( n^*(i) \): factorial function

\( RSE \): relative standard error

\( V \): variance operator

\( \bar{X}_0, \bar{X}_1 \): population concentration of
each kind of particles

\( Y, y \): random variable denoting the
number of particles within a co­
ordination number sample and
its value

\( Z, z \): random variable denoting the
local concentration and its value

\( \mu, \sigma^2 \): expectation and variance of the
contact number in an arbitrarily
mixed state

\( \mu_r, \sigma_r^2 \): ditto in the completely random
state

\( \mu_t, \sigma_t^2 \): ditto in the completely segregated
state

\( \mu(i) \): the \( i \)-th factorial moment

\( \tau \): parameter

References

1) Fan, L.T., S.J. Chen and C.A. Watson: *Ind. Eng.
Chem., 62, [7], 53 (1970).
2) Akao, Y., H. Kunisawa, L.T. Fan, F.S. Lai and R.H.
Wang: *Powder Technol., 15, 267 (1976).
3) Yoshizawa, T. and H. Shindo: *Journal of the Japan
Society for Quality Control, 7, 3 (1977).
4) Lacey, P. M. C.: *J. Appl. Chem.*, 4, 257 (1954).
5) Johnson, N. L. and S. Kotz: "Distributions in Statistics: Discrete Distributions", Houghton Mifflin, Boston (1969).
6) Naylor, T. H., J. L. Balintfy, D. S. Burdick, and K. Chu: "Computer Simulation Techniques (2nd Ed.)", John Wiley & Son (1967).
7) Shindo, H., et al.: *Powder Technol.*, 21, 105 (1978).