

Counting by weighing: construction of two-sided confidence intervals

J. Peng¹, W. Liu², F. Bretz³, A.J. Hayter⁴

¹Department of Mathematics and Statistics

Acadia University, Canada

²S3RI and School of Mathematics

University of Southampton, SO17 1BJ, UK

³Novartis Pharma AG

Basel, 4002, Switzerland

⁴Dept of Statistics and Operations Technology

University of Denver, USA

August 2017, January 2018, April 2018

Abstract

Counting by weighing is widely used in industry and often more efficient than counting manually which is time consuming and prone to human errors especially when the number of items is large. Lower confidence bounds on the numbers of items in infinitely many future bags based on the weights of the bags have been proposed recently in Liu *et al.* (2016). These confidence bounds are constructed using the data from one calibration experiment and for different parameters (or numbers), but have the frequency interpretation similar to a usual confidence set for one parameter only. In this paper, the more challenging problem of constructing two-sided confidence intervals is studied. A simulation-based method for computing the critical constant is proposed. This method is proven to give the required critical constant when the number of simulations goes to infinity, and shown to be easily implemented on an ordinary computer to compute the critical constant accurately and quickly. The methodology is illustrated with a real data example.

Keywords: Confidence bounds; Confidence level; Confidence set; Counting by weighing; Statistical inference; Statistical simulation.

1 Introduction

Our motivating problem comes from banking industry. Clients bring in bags of one-penny coins to exchange for bank notes or one- or two-pound coins. In order to pay the clients, the company needs to know the numbers of coins in the bags. To count the number of coins in a bag manually is not only prone to human errors but also time consuming (and so expensive), especially when the number is large. An efficient alternative is counting by weighing. The bags can be weighed and the (net) weight W_i grams of a bag can be used to infer the number n_i of coins in the bag for infinitely many future bags $i = 1, 2, \dots$. Counting by weighing is a well-known practice used by banks, and also other industries to count, for example, the numbers of plant seeds, bolts/nails, printed labels or medical tablets etc (cf. Nelson, 1983, Guttman and Menzefricke, 1986, Yu, 1989, Mullennix, 1990, Nickerson, 1993, 2003, 2008, Arntzen *et al.*, 1994, Ridout and Suntheralingam, 1997, Ridout and Roberts, 1997). While we focus on counting one-penny coins in this paper, the methodology developed can clearly be applied to counting other items.

As in Liu *et al.* (2016), it is assumed throughout this paper that the weight of the coins in a bag can be weighed accurately, and the weight of a one-penny coin U has a normal distribution $N(\mu, \sigma^2)$ for some μ and σ^2 . Due to the randomness in U , the exact number of one-penny coins in a bag with weight W_i cannot be pin-pointed. The company is more interested in a lower confidence bound on the exact number n_i of one-penny coins in a bag with weight W_i , which can be used as a conservative estimation of n_i . This problem has been solved in Liu *et al.* (2016). From the clients or neutral point of view, however, a two-sided confidence interval for n_i is more interesting and the focus of this paper.

The two-sided confidence intervals are more challenging to construct than the one-sided

confidence bounds as the boundary of the region for which the probability needs to be computed is no longer explicitly available. The novelty of this paper is that a simulation-based method for computing the required critical constant is proposed, which is proved to give the required critical constant when the number of simulations goes to infinity and can be easily implemented on a computer using R software. Similar to the lower confidence bounds, the two-sided confidence intervals are for infinitely many unknown parameters n_i ($i = 1, 2, \dots$) based on just one calibration data set, but still have a frequentist interpretation similar to a usual confidence interval for one unknown parameter only.

The layout of the paper is as follows. The simple situation where μ and σ^2 are assumed to be known is considered in Section 2. The more realistic situation where both μ and σ^2 are unknown parameters is studied in Section 3. An example is given in Section 4. Section 5 contains conclusions and discussion. Finally some mathematical details are provided in the appendix

2 Known μ and σ^2

Assume the values of μ and σ^2 are known in this section. This simple situation helps to motivate and understand the method developed in Section 3 for the more realistic situation that the values of μ and σ^2 are unknown. It is clear that distribution $(W_i - n_i\mu)/\sqrt{n_i\sigma^2} \sim N(0, 1)$. We construct the confidence set for n_i in (1) by using Neyman's (1937) method of inverting a family of acceptance sets for testing $H_0 : n_i = n$:

$$\mathcal{C}(W_i) = \{n : |W_i - n\mu|/\sqrt{n\sigma^2} < c\} \quad (1)$$

where c is a suitably chosen critical constant whose determination is discussed next.

We require that, among the infinitely many confidence sets $\mathcal{C}(W_i)$ for possibly different parameters n_i ($i = 1, 2, \dots$), at least proportion β will contain the true n_i for the pre-specified β (close to one), that is,

$$\liminf_{N \rightarrow \infty} \frac{1}{N} \sum_{i=1}^N I_{\{n_i \in \mathcal{C}(W_i)\}} \geq \beta \quad (2)$$

where I_A denotes the indicator function of set A and so $\frac{1}{N} \sum_{i=1}^N I_{\{n_i \in \mathcal{C}(W_i)\}}$ is the proportion among the N confidence sets $\mathcal{C}(W_i)$ that contain the true n_i . Note that

$$\lim_{N \rightarrow \infty} \frac{1}{N} \sum_{i=1}^N I_{\{n_i \in \mathcal{C}(W_i)\}} = \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{i=1}^N P\{n_i \in \mathcal{C}(W_i)\} = \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{i=1}^N (2\Phi(c) - 1) = 2\Phi(c) - 1,$$

where the first equality above follows from the classical strong law of large numbers (cf. Chow and Teicher, 1978, pp.333), and the second from the definition of $\mathcal{C}(W_i)$ in (1). Hence c is set to be $z_{(\beta+1)/2}$, the $(\beta+1)/2$ quantile of the $N(0, 1)$ distribution, in order to guarantee the property in (2) with equality.

Since $\mathcal{C}(W_i)$ is a standard β level confidence set for n_i for each i , even though the confidence sets $\mathcal{C}(W_i)$ are for possibly different parameters n_i ($i = 1, 2, \dots$), the interpretation of the property in (2) is similar to coverage frequency of a standard confidence set.

Straightforward manipulation shows that the confidence set $\mathcal{C}(W_i)$ in (1) is given by all the natural numbers contained in the interval

$$\left[\frac{2\mu W_i + c^2 \sigma^2 - c\sigma \sqrt{c^2 \sigma^2 + 4\mu W_i}}{2\mu^2}, \frac{2\mu W_i + c^2 \sigma^2 + c\sigma \sqrt{c^2 \sigma^2 + 4\mu W_i}}{2\mu^2} \right]. \quad (3)$$

Note that a sensible point estimator of n_i is W_i/μ .

3 Unknown μ and σ^2

In this section the more realistic situation, where both the values of μ and σ^2 are unknown and so need to be estimated independently of the future weights W_i , is considered. For this the company carries out the following calibration experiment to estimate μ and σ : weigh the j th bag known to contain m_j one-penny coins and record the weight W_{0j} , $j = 1, \dots, k$. Assume $k \geq 2$ in order to be able to estimate σ^2 . This includes the special case that each bag contains only one one-penny coin and so W_{0j} , $j = 1, \dots, k$ are the weights of a random sample of k one-penny coins. From Liu *et al.* (2016), the calibration data $\mathcal{E} = \{(m_j, W_{0j}) : j = 1, \dots, k\}$ can be used to estimate μ and σ^2 in the following way:

$$\begin{aligned}\hat{\mu} &= r(W_{01} + \dots + W_{0k}) \sim N(\mu, r\sigma^2) \quad \text{with} \quad r = \frac{1}{m_1 + \dots + m_k} \\ \hat{\sigma}^2 &= \frac{1}{\nu} \sum_{j=1}^k \left(\frac{W_{0j} - m_j \hat{\mu}}{\sqrt{m_j}} \right)^2 \sim \sigma^2 \chi_\nu^2 / \nu \quad \text{with} \quad \nu = k - 1\end{aligned}$$

and $\hat{\mu}$ and $\hat{\sigma}^2$ are independent.

Due to independence of the weight of a future bag W_i and the calibration data \mathcal{E} , we have

$$\frac{W_i - n_i \hat{\mu}}{\hat{\sigma} \sqrt{n_i + r n_i^2}} \sim t_\nu$$

Using Neyman's (1937) method again, we construct

$$\mathcal{C}(W_i) = \left\{ n : \frac{|W_i - n \hat{\mu}|}{\hat{\sigma} \sqrt{n + r n^2}} < c \right\}, \quad i = 1, 2, \dots \quad (4)$$

as the confidence set for n_i , where c is a critical constant whose determination is considered below.

As in Section 2, we require that the proportion of all the future confidence sets $\mathcal{C}(W_i)$ ($i = 1, 2, \dots$) that include the true n_i is no less than the pre-specified β , that is,

$$\liminf_{N \rightarrow \infty} \frac{1}{N} \sum_{i=1}^N I_{\{n_i \in \mathcal{C}(W_i)\}} \geq \beta. \quad (5)$$

An argument similar to that in Liu *et al.* (2016) shows that a sufficient condition for guaranteeing (5) is

$$\inf_{n_i \in \mathcal{N}} E_{W_i|\mathcal{E}} I_{\{n_i \in \mathcal{C}(W_i)\}} \geq \beta \quad (6)$$

where $E_{W_i|\mathcal{E}}$ denotes the conditional expectation with respect to the random variable W_i conditioning on the calibration data \mathcal{E} (or, equivalently, $\hat{\mu}$ and $\hat{\sigma}$), and \mathcal{N} denotes the set of natural numbers known *a priori* to contain all the future n_i values. It is sensible to assume that all the future n_i values are in a known range $[n_l, n_u]$ even though the individual future n_i values are unknown. For example, it is sensible to assume all future n_i 's must be at least n_l , $n_l = 5$ say, since it is highly unlikely that someone would bring in less than n_l 1p coins to exchange for bank notes or £1 or £2 coins. Similarly, one can easily set a conservative upper limit n_u from the capacity of the weighing scale. Hence we assume all the future n_i values are contained in the known $\mathcal{N} = \{n : n_l \leq n \leq n_u\}$ in the rest of this paper.

Now straightforward manipulation, using the definition of $\mathcal{C}(W_i)$ in (4), show that

$$\inf_{n_i \in \mathcal{N}} E_{W_i|\mathcal{E}} I_{\{n_i \in \mathcal{C}(W_i)\}} = \inf_{n_i \in \mathcal{N}} \Psi \left(\sqrt{rn_i} Z, c\sqrt{1+rn_i} X \right) \quad (7)$$

where $\Psi(a, b) = \Phi(a + b) - \Phi(a - b)$, $Z = (\hat{\mu} - \mu)/\sqrt{r\sigma^2} \sim N(0, 1)$, $X = \hat{\sigma}/\sigma \sim \sqrt{\chi_\nu^2/\nu}$, and Z and X are independent. Note that the last expression in (7) depends on the random variables Z and X (via $\hat{\mu}$ and $\hat{\sigma}$), and so the condition in (6) cannot be guaranteed for all Z and X . For example, if Z is large and X is small then the expression in (7) can be smaller than 0.5 and so β . Hence we guarantee (6) with a large probability $1 - \alpha$ with respect to the randomness in Z and X :

$$P_{\mathcal{E}} \left\{ \inf_{n_i \in \mathcal{N}} \Psi \left(\sqrt{rn_i} Z, c\sqrt{1+rn_i} X \right) \geq \beta \right\} \geq 1 - \alpha. \quad (8)$$

This in turn guarantees that

$$P_{\mathcal{E}} \left\{ \liminf_{N \rightarrow \infty} \frac{1}{N} \sum_{i=1}^N I_{\{n_i \in \mathcal{C}(W_i)\}} \geq \beta \right\} \geq 1 - \alpha. \quad (9)$$

The interpretation of this statement is that, based on one set of calibration data \mathcal{E} , one constructs confidence sets $\mathcal{C}(W_i)$ for n_i for all future bags ($i = 1, 2, \dots$) and claims that at least β proportion of these confidence sets do contain the true n_i . Then we are $1 - \alpha$ confident with respect to the randomness of the calibration data \mathcal{E} that the claim is correct.

We now consider how to compute the critical constant c so that the probability in (8) is equal to $1 - \alpha$. For the one-sided case considered in Liu *et al.* (2016), numerical quadrature is used to compute the probability and then a searching algorithm is used to compute the required critical constant. The two-sided case in this paper is more challenging than the one-sided case since the region of (Z, X) specified in the probability sign in (8) does not have a simple expression and so the computation of the probability using numerical quadrature is not straightforward. Hence we propose the following Monte Carlo simulation method to find the critical constant c .

In each simulation indexed by s , first generate independent $Z_s \sim N(0, 1)$ and $X_s \sim \sqrt{\chi_\nu^2/\nu}$. Then compute $c = c_s$ such that

$$\inf_{n_i \in \mathcal{N}} \Psi \left(\sqrt{rn_i} Z_s, c\sqrt{1 + rn_i} X_s \right) = \beta. \quad (10)$$

Since the infimum in the expression above is over the finite set \mathcal{N} , the expression on the left hand side of the equality in (10) is easy to compute for any given $c > 0$. Note further that this expression is strictly increasing in $c > 0$ and hence the c that solves the equation in (10) is easy to find and denoted as c_s . Repeat the simulation process S times to get c_1, c_2, \dots, c_S . Sort the c_s 's as $c_{[1]} \leq \dots \leq c_{[S]}$ and use $c_{[\langle(1-\alpha)S\rangle]}$ as c , where $\langle a \rangle$ denotes the integer part of number a .

It is shown in the appendix that $c_{[\langle(1-\alpha)S\rangle]} \rightarrow c$ almost surely (a.s.) as $S \rightarrow \infty$. Hence $c_{[\langle(1-\alpha)S\rangle]}$ can be as close to the exact critical constant c as required by using a sufficiently

large S . An R program has been written to implement this computation and is available from the authors. For the example given in Section 4, $S = 1,000,000$ is used. Based on our experiments with different random seeds, the critical value $c_{[(1-\alpha)S]}$ based on $S = 1,000,000$ is accurate to two decimal places at least and so more than sufficient for practical purpose. It is noteworthy that a similar simulation method can be devised to find the critical constant for the lower confidence bounds in Liu *et al.* (2016). See Section 4 for more information on numerical results.

It is straightforward to show that the confidence set $\mathcal{C}(W_i)$ in (4) is given by the interval $[L(W_i), U(W_i)]$ where

$$L(W_i) = \begin{cases} \frac{2\hat{\mu}W_i + c^2\hat{\sigma}^2 - c\hat{\sigma}\sqrt{c^2\hat{\sigma}^2 + 4\hat{\mu}W_i + 4rW_i^2}}{2(\hat{\mu}^2 - rc^2\hat{\sigma}^2)} & \text{if } \hat{\mu}^2 - rc^2\hat{\sigma}^2 \neq 0 \\ \frac{W_i^2}{2\hat{\mu}W_i + c^2\hat{\sigma}^2} & \text{otherwise} \end{cases} \quad (11)$$

$$U(W_i) = \begin{cases} \frac{2\hat{\mu}W_i + c^2\hat{\sigma}^2 + c\hat{\sigma}\sqrt{c^2\hat{\sigma}^2 + 4\hat{\mu}W_i + 4rW_i^2}}{2(\hat{\mu}^2 - rc^2\hat{\sigma}^2)} & \text{if } \hat{\mu}^2 - rc^2\hat{\sigma}^2 \neq 0 \\ \frac{W_i^2}{2\hat{\mu}W_i + c^2\hat{\sigma}^2} & \text{otherwise} \end{cases} \quad (12)$$

Note that $P\{\hat{\mu}^2 - rc^2\hat{\sigma}^2 = 0\} = 0$ since $\hat{\mu}$ and $\hat{\sigma}^2$ are continuous random variables. Incorporating the *priori* information that $n_l \leq n_i \leq n_u$, the final confidence set is given by all the natural numbers in the interval

$$\mathcal{C}(W_i) = [n_l, n_u] \cap [L(W_i), U(W_i)]. \quad (13)$$

4 An example

Consider the example given in Liu *et al.* (2016). In the calibration experiment, eleven bags of one-penny coins of known numbers $(m_1, \dots, m_{11}) = (100, 200, \dots, 1100)$ have been weighed to give the corresponding weights (grams) $(W_{0,1}, \dots, W_{0,11}) = (356.97, 716.67, 1060.96, 1427.10,$

1781.28, 2130.23, 2489.94, 2849.74, 3223.32, 3575.76, 3907.94). The data give $\hat{\mu} = 3.564$ with $r = 1/6600$ and $\hat{\sigma}^2 = 0.101$ with degrees of freedom $\nu = 10$.

Let us assume that the number of coins in each future bags n_i is known *a priori* to be between $n_l = 100$ and $n_u = 1400$, and set $\beta = 0.95$ and $\alpha = 0.05$. Then the critical constant in (4) is computed by our R program (using $S=1,000,000$ simulations for all the computations in this section) to be $c = 3.232$, which is larger than the critical constant 2.833 for the lower confidence bounds given in Liu *et al.* (2016) as expected. Now for any future bag with weight W_i , one can use the formula in (13) to compute the confidence interval $\mathcal{C}(W_i)$ on n_i . For example, if $W_i = 4000$ then $\mathcal{C}(W_i) = [1112.1, 1132.9]$ while the lower confidence bound is 1118.9 and the point estimate is 1122.5; if $W_i = 4600$ then $\mathcal{C}(W_i) = [1279.5, 1302.2]$ while the lower confidence bound is 1286.8 and the point estimate is 1290.8. These results show that the confidence interval $\mathcal{C}(W_i)$ is quite tight around the point estimate. It is clearly advantageous to use the confidence interval instead of the point estimate due to the confidence statement (9) associated with the confidence interval. The confidence interval provides the extra information on the plausible upper bound on n_i , which is of interest to the clients but not available from the lower confidence bound given in Liu *et al.* (2016).

If we treat 3.564 and 0.101 as the known values of μ and σ^2 , respectively, then the formula in (3) can be used to compute the confidence interval $\mathcal{C}(W_i)$ on n_i , ignoring the value of α . For example, if $W_i = 4000$ then $\mathcal{C}(W_i) = [1116.6, 1128.3]$ while the lower confidence bound is 1117.5 and the point estimate is still 1122.5; if $W_i = 4600$ then $\mathcal{C}(W_i) = [1284.6, 1297.1]$ while the lower confidence bound is 1285.6 and the point estimate is still 1290.8. The differences between the two confidence intervals in (3) and (11) are small in this example.

To get some idea on how sensitive the critical constant c and the confidence interval $\mathcal{C}(W_i)$

in (11) are to the bounds n_l and n_u , we have computed c , $\mathcal{C}(400)$ and $\mathcal{C}(4000)$ for various (n_l, n_u) . From the results given in Table 1, one can see that both c and the confidence intervals $\mathcal{C}(400)$ and $\mathcal{C}(4000)$ are not sensitive to the small changes in (n_l, n_u) . Hence one may prefer to use more conservative bounds (n_l, n_u) , for example $(n_l, n_u) = (50, 1500)$ instead of $(n_l, n_u) = (100, 1400)$, to make sure that all the future n_i are in the interval $[n_l, n_u]$.

Table 1: Critical constant c , $\mathcal{C}(4000)$ and $\mathcal{C}(400)$ for various (n_l, n_u)

(n_l, n_u)	(100, 1400)	(50, 1400)	(150, 1400)	(100, 1300)	(100, 1500)	(50, 1500)
c	3.2320	3.2287	3.2335	3.2258	3.2380	3.2351
$\mathcal{C}(4000)$	[1112.05, 1132.96]	[1112.06, 1132.95]	[1112.05, 1132.97]	[1112.07, 1132.94]	[1112.04, 1132.98]	[1112.04, 1132.98]
$\mathcal{C}(400)$	[109.21, 115.37]	[109.21, 115.37]	[109.20, 115.37]	[109.20, 115.37]	[109.20, 115.38]	[109.20, 115.37]

To assess the accuracy of the critical constants c given in Table 1, we have tried different random seeds in our R program. For example, when $(n_l, n_u) = (100, 1400)$, the critical constants from different random seeds are 3.2320, 3.2317, 3.2351, 3.2350, 3.2311, 3.2347, 3.2337. This indicates that the critical constants based on one million simulations are accurate to two decimal places at least. Of course one can always increase the number of simulations to achieve higher accuracy, but at the cost of longer computation time. Note that only one critical constant needs to be computed for use with infinitely many future bags, and it takes about sixty five minutes to compute one critical constant using one million simulations on an ordinary Window's PC (Core(TM2) Due CPU P8400@2.26GHz).

We have also written an R program using the simulation method proposed in this paper to compute the critical constant c for the lower confidence bounds in Liu *et al.* (2016). For example, when $(n_l, n_u) = (100, 1400)$, $\beta = 0.95$ and $\alpha = 0.05$, the critical constant $c = 2.833$

in Liu *et al.* (2016) is computed to be 2.8336 from our R program based on one million simulations. This agrees with the observation above that the critical constant based on one million simulations is accurate to at least two decimal places.

5 Conclusions

Construction of two-sided confidence intervals for infinitely many future n_i 's based on the data \mathcal{E} collected from only one calibration experiment is considered in this paper. Two-sided confidence intervals are more interesting, from the clients or neutral point of view at least, than the lower confidence bounds, which are of interest mainly to the company.

A simulation-based method is proposed for the computation of the critical constant c , from which the confidence intervals can be easily calculated. We have proved that this method provides the accurate critical constant when the number of simulations goes to infinity and demonstrated that this method allows c to be computed accurately and quickly on an ordinary PC. This method can be easily adapted to compute the critical constant for the lower confidence bounds in Liu *et al.* (2016). But it is not clear how the computation method in Liu *et al.* (2016) can be used to compute the critical constant for the confidence intervals in this paper.

As in Liu *et al.* (2016), the two sources of randomness in the calibration data \mathcal{E} and in a future weights W_i have been treated differently, and the confidence intervals have the interpretation that, with confidence level $1 - \alpha$ about the randomness in \mathcal{E} , the proportion of the confidence intervals that contain the true n_i 's is at least β .

If the two sources of randomness are treated on equal footing then a standard confidence

interval for one n_i is given by (4) with $c = t_{\nu, (1+\beta)/2}$, the $(1 + \beta)/2$ quantile of the t_ν distribution. This confidence interval has the following coverage frequency interpretation. Do one calibration experiment to collect \mathcal{E} and measure the weight W_i of one future bag, based on which the confidence set $C(W_i)$ for this one n_i is constructed; then the frequency of a large number of confidence sets that contain the corresponding n_i 's is approximately β . Note that one calibration data set \mathcal{E} is used only once with one future W_i to produce one confidence set in this construction method. This is clearly different from the real situation in which the data \mathcal{E} from one calibration experiment is used repeatedly for inferences for infinitely many future n_i values.

Finally, several problems warrant further research, which include the optimal design of the calibration experiment which not only produces good confidence intervals in a certain sense but also reduces the cost and possible human errors in counting the coins manually in the calibration experiment, and possibly non-normality of the weight of a one-penny coin. Based on the specific calibration data \mathcal{E} given in Section 4 at least, the confidence sets $C(W_i)$ contain more than one natural number (and so the counting from these confidence sets is not “accurate”). This may be because we are unlucky in getting the \mathcal{E} which overestimate the unknown σ^2 . It is clear that if $\sigma^2 = 0$ then counting by weighing is accurate. It is interesting to study how small σ^2 should be in order that the confidence sets $C(W_i)$ contain only one natural number with a pre-specified large chance. Furthermore, it is interesting to investigate the possibility of generalizing the methods of this paper and Liu *et al.* (2016) to deal with situations in which the bags contain two (or more) different types of coins.

6 Appendix: Mathematical details

Note that the exact critical constant c satisfies

$$P_{\mathcal{E}} \left\{ \inf_{n_i \in \mathcal{N}} \Psi \left(\sqrt{rn_i} Z, c\sqrt{1+rn_i} X \right) \geq \beta \right\} = 1 - \alpha. \quad (14)$$

To show that $c_{[\langle(1-\alpha)S\rangle]} \rightarrow c$ a.s. as $S \rightarrow \infty$, it suffices to show that, $\forall \epsilon > 0$, we have

$P\{|c_{[\langle(1-\alpha)S\rangle]} - c| > \epsilon \text{ i.o.}\} = 0$, which is implied by

$$\lim_{k \rightarrow \infty} \sum_{S=k}^{\infty} \left(P \left\{ c_{[\langle(1-\alpha)S\rangle]} - c > \epsilon \right\} + P \left\{ c_{[\langle(1-\alpha)S\rangle]} - c < -\epsilon \right\} \right) = 0; \quad (15)$$

see, for example, Chow and Teicher (1978, pp. 41-42, Lemmas 1 and 2).

Now from the definition of $c_{[\langle(1-\alpha)S\rangle]}$ we have

$$\begin{aligned} & P \left\{ c_{[\langle(1-\alpha)S\rangle]} - c < -\epsilon \right\} \\ &= P \left\{ c_{[\langle(1-\alpha)S\rangle]} < c - \epsilon \right\} \\ &= P \left\{ \sum_{j=1}^S I_{\left\{ \inf_{n_i \in \mathcal{N}} \Psi \left(\sqrt{rn_i} Z_j, (c-\epsilon)\sqrt{1+rn_i} X_j \right) \geq \beta \right\}} > \langle(1-\alpha)S\rangle \right\} \\ &= P \left\{ \frac{1}{S} \sum_{j=1}^S (Y_j - E(Y_j)) > \frac{1}{S} (\langle(1-\alpha)S\rangle - SE(Y_1)) \right\} \end{aligned} \quad (16)$$

where $Y_j = I_{\left\{ \inf_{n_i \in \mathcal{N}} \Psi \left(\sqrt{rn_i} Z_j, (c-\epsilon)\sqrt{1+rn_i} X_j \right) \geq \beta \right\}}$. Note that Y_1, Y_2, \dots are i.i.d. Bernoulli random variables with success probability

$$p = E(Y_j) = P \left\{ \inf_{n_i \in \mathcal{N}} \Psi \left(\sqrt{rn_i} Z_j, (c-\epsilon)\sqrt{1+rn_i} X_j \right) \geq \beta \right\} = (1-\alpha) - \delta$$

for some $\delta = \delta_\epsilon > 0$ since $P \left\{ \inf_{n_i \in \mathcal{N}} \Psi \left(\sqrt{rn_i} Z_j, c\sqrt{1+rn_i} X_j \right) \geq \beta \right\} = (1-\alpha)$ from (14).

Hence, for all sufficiently large S , we have $(\langle(1-\alpha)S\rangle - SE(Y_1))/S > \delta/2$ and so the probability in (16) is no more than

$$P \left\{ \frac{1}{S} \sum_{j=1}^S (Y_j - E(Y_j)) > \delta/2 \right\}$$

$$\leq E \left(\sum_{j=1}^S (Y_j - E(Y_j)) \right)^4 / (S\delta/2)^4 \quad (17)$$

$$= Sp(1-p) (3Sp(1-p) - 6p(1-p) + 1) / (S\delta/2)^4 \quad (18)$$

where the inequality in (17) follows directly from the Markov inequality (cf. Chow and Teicher, 1978, pp.88), and the equality in (18) follows from direct calculation of the fourth central moment of the binomial random variable $\sum_{j=1}^S Y_j$ (cf. Chow and Teicher, 1978, p.41). From the upper bound in (18), it is clear that

$$\lim_{k \rightarrow \infty} \sum_{S=k}^{\infty} P \left\{ c_{[\langle (1-\alpha)S \rangle]} - c < -\epsilon \right\} = 0. \quad (19)$$

A similar argument shows that

$$\lim_{k \rightarrow \infty} \sum_{S=k}^{\infty} P \left\{ c_{[\langle (1-\alpha)S \rangle]} - c > \epsilon \right\} = 0. \quad (20)$$

The combination of (19-20) establishes (15) and so completes the proof.

References

- Arntzen, F.K., Mulder, J.G. and Visser, J.H.M. (1994). Rapid and nondestructive assessment of the number of eggs in cysts of potato cyst nematodes by weighing. *Fundamental and Applied Nematology*, 17, 299-301.
- Chow, Y.S. and Teicher, H. (1978). *Probability Theory: Independence, Interchangeability, Martingales*. Springer-Verlag, New York.
- Guttman, I. and Menzefricke, U. (1986). Counting by weighing: an approach using renewal theory. *Journal of the American Statistical Association*, 81, 129-131.
- Liu, W., Han, Y., Bretz, F., Wan, F. and Yang, P. (2016). Counting by weighing: know your numbers with confidence. *J. Royal Statistical Society (C)*, 65(4), 641-648.

- Mullennix, P. (1990). Accurate estimation of count totals using ratio estimators. *Journal of Quality Technology*, 22, 277-282.
- Nelson, P.R. (1983). Counting a large number of items using the weight of a small sample. *Technometrics*, 25, 173-177.
- Nickerson, D.M. (1993). Another look at counting by weighing. *Commun. Statist. – Simula.*, 22, 323-343.
- Nickerson, D.M. (2003). Addendum to counting by weighing and the seed testing problem. *Annals of Applied Biology*, 143, 371-374.
- Nickerson, D.M. (2008). Counting by weighing and its effect on comparing population proportions *Statistics & Probability Letters*, 78, 2321-2326.
- Neyman, J. (1937). Outline of a theory of statistical estimation based on the classical theory of probability. *Philosophical Transaction of the Royal Society of London, Series A, Mathematical and Physical Sciences*, 236, 333-380.
- Ridout, M.S. and Roberts, S.J. (1997). Improving quality control procedures for seed-borne pathogens by testing several samples of seeds. *Seed Science and Technology*, 25, 195-202.
- Ridout, M.S., Suntheralingam, N. (1997). Counting by weighing and its effect on seed testing procedures. *Annals of Applied Biology*, 130, 179-185.
- Yu, K.F. (1989). On a problem of counting by weighing. *The Canadian Journal of Statistics*, 17, 377-389.