1 The Allocation of Inspection Resources

Owen Jones, Andrew P. Robinson, Martin Shield and Jessica Sibley

1.1 Introduction

Inspection is carried out by biosecurity protection authorities to detect and exclude biosecurity contaminations, by customs services to intercept illegal weapons and drugs, by taxation organisations to verify taxation returns and by environmental protection authorities to determine the levels of pollutants in public goods. In this chapter, we focus on inspections performed by regulators to ensure that a process complies with regulations. Our specific interest is border inspections for biosecurity contaminations.

We define inspection as the examination of a unit to determine whether or not it is compliant with relevant regulations. In the present context, an inspection will determine whether the unit contains biosecurity risk material. A typical unit could be an international passenger, a sea container or a pallet of goods. Inspection usually involves examining the unit and any accompanying packaging, and depending on the nature of the unit, inspection may also involve the examination of a sample taken from the unit. For example, the inspection of a consignment of oranges might focus on a random sample of 600 oranges, and the inspection of a consignment of coffee beans might focus on one or more samples of coffee beans extracted from the container by means of a probe.

We will assume that units arrive sequentially and that there is no logical demarcation in the flow of arrivals that could be used to define a collection of units to serve as a basis for structuring an inspection system. Therefore, although traditional methods may be used to determine the procedure for sampling from a unit such as a container, they are not appropriate for deciding how many or which units to inspect. Rather, as each unit arrives, a decision must be made on whether or not to inspect it. We will suppose that our inspection criteria are updated after every N-th arrival, for some fixed N. In particular, after every N-th arrival, we update our estimate of the non-compliance rate and adjust the frequency of our inspections accordingly.

The frequency of inspections is determined by three requirements: to intercept non-compliant units, to estimate the contamination level and to deter maleficent agents. We will assume that the only data that we have on the non-compliance rate of arriving units are the results of previous inspections. Moreover, we do not wish to use data from more than N past arrivals because we want our estimates to be current. Thus the frequency of inspections directly affects the quality of our estimates. When contaminated units are identified, they are destroyed, treated or re-exported so that they do not present a biosecurity threat. The specific action taken will not affect our analysis here. Finally, although it is clear that when inspections are well publicised and the penalties for infraction are sufficient, the knowledge of inspection may influence the behaviour of importers; we make no attempt to model this feedback here.

The frequency of inspections will generally increase with the estimated rate of contamination, although see Cannon (2009) and Press (2009). The frequency of inspections will therefore be low when the estimated rate is negligible but should not be allowed to decrease so much that it becomes impossible to detect an important increase in the contamination frequency within a reasonable time frame. When the sampling rate is low, detecting a contaminated unit can cause a spike in the estimate of the contamination rate that may misleadingly portend a change in the baseline rate. A further consideration is that, assuming we update our estimated contamination rate after every N-th arrival, our inspection regime should allow for a rapid increase in the inspection frequency if there is an important increase in the number of non-compliant units detected. A brief review of inspection resource allocation strategies can be found in Robinson et al. (2011); see also Cannon (2009).

Robinson et al. (2008, 2011) developed the import risk inspection sampling (IRIS) algorithm with the goal of determining an inspection level that reflects the joint needs to intercept non-compliant units and maintaining adequate estimates of contamination levels. The IRIS algorithm allows the manager to choose the length N of the review period, but does not allow changing the inspection frequency between the review periods if there is evidence of an increase in the contamination frequency. In this chapter, we show how to combine the IRIS algorithm with the different sampling or alert modes used by Dodge (1943) and Dodge and Torrey (1951) in the continuous sampling plan (CSP) and its variants. The combined algorithm retains the convenience of regular review periods while including mechanisms to trigger periods of high-frequency inspections.

This chapter is structured as follows. We develop a conceptual framework for the inspection process in Section 1.2. We review and extend the IRIS algorithm in Section 1.3. We introduce the CSP in Section 1.4 and discuss how to combine it with IRIS and why this might be useful. In Sections 1.2, 1.3 and 1.4, we assume that we are acting on a single homogeneous pathway of units. In Section 1.5 we consider the problem of pathways that are too small to get adequate estimates of the contamination level, and suggest a way of combining pathways using our IRIS–CSP hybrid algorithm. We then test our approach using a simulation experiment based on inspection data.

1.2 Conceptual Framework

In this section we present a conceptual framework for the inspection process that we will use to describe our inspection algorithms. Here, we use the vocabulary and context most suited to biosecurity inspection, but the principles are quite general.

We define an inspection *unit* as the entity upon which inspection is performed. Diverse kinds of units are of interest, and the method for inspection of each depends on the characteristics of the unit. Examples of units include people, consignments of imported goods and containers of commodities. Inspection units are analogous to sampling units in sampling theory.

We define a *pathway* as a sequence of units that are deemed to be similar by the inspectorate and over which the inspectorate has some regulatory authority. Defining a pathway as a collection of like units is subjective because there are numerous hierarchical levels of collections of units. For example, a pathway could comprise all air passengers, all passengers who arrive from a certain departure point or all passengers who have been out of the country for more than six weeks. Similarly, in the case of imported coffee beans, a pathway could comprise all consignments of coffee beans, all consignments from a certain supplier, all consignments from a specific country, all consignments to a particular importer or any combination of these. Pathways are analogous to infinite populations in sampling theory.

We assume that inspection of a unit yields a binary result: the unit is deemed to be contaminated (non-compliant) or not contaminated (compliant). We also suppose that the *effectiveness*, w, of inspections is known and constant for any given pathway. This means that a non-compliant unit that is inspected will be detected with a known probability w. In general, this probability will not be known and must be estimated using a procedure called an *endpoint survey*.

Consider the k-th unit that arrives at the inspection point from a given pathway. We define the *approach rate*, p_k, as the probability that the unit is non-compliant. The p_k is indexed by k because, in general, we allow it to change over time, although in practice we expect any change to be gradual, perhaps with occasional jumps. We define the *sampling rate*, s_k, as the probability that the k-th unit is inspected, and we define the *leakage rate*, r_k, as the probability that the unit is non-compliant and allowed past the inspection point. Thus, $r_k = (1 - ws_k)\, p_k$. Broadly speaking, our goal is to choose a value of s_k that is as small as possible while keeping r_k at an acceptable level. An important feature of the IRIS algorithm is that when determining s_k, it specifically makes allowance for uncertainty in our estimate of r_k. It is also important to know how quickly the sampling rate increases when there is an increase in p_k, to which end we incorporate the CSP methodology.

Our definition of the leakage rate gives the probability that a unit arriving at the inspection point is non-compliant but still gets through. We could also consider the probability that a unit that leaves the inspection point is non-compliant (the post-inspection leakage rate). These probabilities will be the same in the case of rectifying inspections, in which detected non-compliant units are made compliant and then released. In the case of non-rectifying inspections, the post-inspection leakage rate will always be higher than the leakage rate. However, when p_k is small, which is often the case, the two will be close because the proportion of units that are rejected will be small.

We collect units into groups of sizes $N_1, N_2, \ldots$, where the first N_2 units to arrive are considered to belong to group 1, the next N_2 to group 2, and so on. We assume that the group sizes are known in advance. The groups could be natural groupings, such as containers on a ship or passengers on an aeroplane, or they could be units that arrive during a given time period, say every three months. In the latter case, the group sizes can only be estimated ahead of time and will not be known for certain. The accuracy of these estimates is not of particular importance when we look at the performance of our inspection algorithms (see Section 1.5.1).

To measure performance we use the long-run average leakage rate, either theoretical or estimated. For a given approach rate p, the average outgoing quality, AOQ(p), is defined as the long-run average leakage rate when the approach rate $p_k = p$ is constant, under the assumption that units are independent and inspections are perfectly accurate ($w = 1$).[1] For a specific data sample, the estimated long-run average leakage rate is the sample outgoing quality (SOQ). Note that by long-run average leakage rate we mean the proportion of the given pathway that is non-compliant and undetected.

1.3 The IRIS Algorithm

Throughout this section, we assume that the approach rate $p_k = p$ is constant. The IRIS algorithm is an ad hoc procedure designed to ensure that the leakage rate is kept below a set level with a given probability as long as the approach rate does not increase. Even when p is very small, we inspect frequently enough that our estimate of p remains acceptably accurate.

Suppose that in the first block of N_1 units there were n_1 inspections that found x_1 non-compliant units, giving us a point-estimate for p of $\hat{p}_1 = x_1 / (wn_1)$. Our aim is to choose n_2, the number of units to inspect from the next block of size N_2.

We start by adding a positive bias to $\hat{p}_1$ to allow for error and uncertainty in our estimate. Let $\hat{p}_* = \hat{p}_1 + \varepsilon$ be our biased estimate. Next, suppose that we sample n_2 units from the second block of N_2 units and find X_2 non-compliant units. Let $p_2 = X_2 / (wn_2)$ be the estimate of p obtained from these inspections, then $\mathrm{E}P_2 = p$ and $\mathrm{Var}P_2 = p(1 - wp) / (wn_2)$. Given these,[2] we adopt the following model for p using a beta distribution:

$$P \sim \mathrm{beta}(\hat{p}_* w' n_2 + 0.5, (1 - \hat{p}_*) w' n_2 + 0.5) \quad \text{where} \quad w' = \frac{(1 - \hat{p}_*) w}{1 - w \hat{p}_*}. \qquad (1.1)$$

[1] The term average outgoing quality was first used by Dodge (1943), who also used the average outgoing quality limit, AOQL = $\max_p$AOQ(p), to give an overall measure of the effectiveness of an inspection policy. Lieberman (1953) went a step further and proposed the unrestricted average outgoing quality limit, UAOQL, which is an upper bound for the long-run average leakage rate for any sequence of p_k, not just constant sequences.

[2] We are treating the sample units as independent and identically distributed observations and not as a sample from a finite population of size N_2. This is because we are estimating the long-run approach rate, not just the approach rate for the second sampling period.

Under this model, conditional on $\hat{p}_*$, P has mean

$$\frac{\hat{p}_* + \delta}{1+2\delta} \quad \text{where} \quad \delta = \frac{1}{2w'n_2} \tag{1.2}$$

and variance.

$$\frac{(\hat{p}_* + \delta)(1-\hat{p}_* + \delta)}{w'n_2(1+2\delta)^2(1+4\delta)} = \frac{\hat{p}_*(1-\hat{p}_*)}{w'n_2} + \delta' = \frac{\hat{p}_*(1-w\hat{p}_*)}{wn_2} + \delta' \quad \text{where} \quad \delta' = O(\delta). \tag{1.3}$$

The δ term is included so that, even if $\hat{p}_*$ is very small, P has a mean and variance bounded away from 0.

Given P, our estimate of the leakage rate, r, for the next block is $R = (1-ws)P$, where $s = n_2 / N_2$ is the proportion of the next block to be sampled. We take as our (positively biased) point estimate of r the $100(1-\alpha)\%$ point of R, where α is specified by the manager, for example 0.10. That is, if betainv is the inverse of the beta density,

$$\hat{r}_2 = (1-ws) \quad \text{betain v}(1-\alpha, \hat{p}_* w'n_2 + 0.5, (1-\hat{p}_*)w'n_2 + 0.5). \tag{1.4}$$

This construction allows the manager to apply a level of surety to the estimate, providing a platform for risk-averse inspection strategies if the consequences of failure are large. Writing $s = n_2 / N_2$ we see that by putting $\hat{r}_2 = r$, where r is our target leakage rate, we get an equation for n_2. Equation 1.4 is easily solved numerically by using a root-finding algorithm.

When the IRIS algorithm was originally introduced by Robinson et al. (2008, 2011), they suggested that ε, the bias added to $\hat{p}_1$ to get $\hat{p}_*$, should be such that $\hat{p}_*$ corresponds to a percentage point from a beta distribution with mean approximately $\hat{p}_1$ and variance proportional to $1/n_2$. However, if $\hat{p}_*$ depends on n_2, then Eq. 1.4 and $\hat{p}_*$ need to be solved iteratively. That is, we choose a $\hat{p}_*$ to start then solve Eq. 1.4 to get n_2, which gives us a new $\hat{p}_*$. Using this $\hat{p}_*$, we solve Eq. 1.4 again to get a new n_2 and thus a new $\hat{p}_*$. We continue until $\hat{p}_*$ and n_2 converge. We have included ε in our description because it is present in the original IRIS algorithm. However, the algorithm already includes a mechanism to deal with the uncertainty in our estimates, namely the α in Eq. 1.4. In practice, adding ε to $\hat{p}_1$ does not add a great deal to the robustness of the method and we now suggest that it can be omitted.

1.3.1 Bayes–IRIS

Although the IRIS algorithm produces reasonable sampling rates in operational settings (Robinson et al., 2011), the ad hoc nature of the algorithm makes it difficult to justify theoretically. In the remainder of this section we use a Bayesian approach to derive an analogous algorithm from first principles. We call the resulting algorithm Bayes–IRIS, and although it results in a rather different equation for n_2, it produces solutions similar to those of the IRIS algorithm in many operational settings.

We suppose that the first review period has just finished and we are planning for the second review period. In the first review period, we sampled n_1 out of N_1 units and found x_1 non-compliant units. Our goal is to choose n_2, the number of units to sample from the next N_2, so that the leakage rate is kept below a threshold r with probability $1-\alpha$. We will suppose initially that inspections are error free, that is, $w=1$.

As before, we start with an estimate of p. We use a Bayesian approach so that our estimate takes the form of a distribution. We deliberately choose not to use any information from before the first review period when estimating p. This is because we want our estimate to be current and capable of responding quickly to changes in the approach rate. Let P_0 be a distribution that represents our estimate at the start of the first review period based on no information. In Bayesian terminology, P_0 is called a non-informative prior. We use the usual choice of non-informative prior for a probability, the beta(0.5, 0.5) distribution.[3]

$$P_0 \sim \text{beta}(0.5, 0.5). \tag{1.5}$$

At the end of the first review period, we update our distribution for p based on the observed number of compliant and non-compliant units. We call this P_1 (the posterior distribution), and standard calculations give us

$$P_1 \sim \text{beta}(x_1 + 0.5, n_1 - x_1 + 0.5). \tag{1.6}$$

Now suppose that we take a sample of size n_2 from the N_2 units that arrive during the second review period. Let X_2 be the number of non-compliant units in that sample. If we knew p, then X_2 would have a $\text{binom}(n_2, p)$ distribution. Instead, using our distribution P_1 for p, we obtain the distribution of X_2 by integrating the binomial distribution over the possible values of p. The resulting distribution is known as the beta-binomial. We write $X_2 \sim \text{beta-binom}(n_2, x_1 + 0.5,\ n_1 - x_1 + 0.5)$, and we have

$$\mathrm{P}(X_2 = x_2) = \binom{n_2}{x_2} \frac{\beta(x_2 + x_1 + 0.5, n_2 - x_2 + n_1 - x_1 + 0.5)}{\beta(x_1 + 0.5, n_1 - x_1 + 0.5)}, \tag{1.7}$$

where $\beta(\text{a}, \text{b})$ is the beta function evaluated at (a, b).

Given X_2, the leakage rate is $R_2 = (1 - n_2 / N_2) X_2 / n_2 = \left(\frac{1}{n_2} - \frac{1}{N_2}\right) X_2$, and requiring $\mathrm{P}(R_2 > r) \le \alpha$ is equivalent to requiring $\mathrm{P}\left(X_2 > \frac{r n_2 N_2}{N_2 - n_2}\right) \le \alpha$. Our sample size for the second sampling period is the smallest n_2 for which

[3] Note that some authors such as Tuyl et al. (2009) argue that beta(1, 1) is a better choice (the uniform prior). However, the beta (0.5, 0.5) prior, which is an example of a Jeffreys prior, is still the most commonly used. Practically, the difference is apparent only when we have a very small sample size, n_1, in which case the Jeffreys prior favours extreme probabilities (closer to 0 or 1) more than the uniform prior does.

$$\sum_{x_2=0}^{\lfloor r n_2 N_2/(N_2-n_2)\rfloor} \binom{n_2}{x_2} \frac{\beta(x_2+x_1+0.5, n_2-x_2+n_1-x_1+0.5)}{\beta(x_1+0.5, n_1-x_1+0.5)} \geq 1-\alpha, \tag{1.8}$$

where the truncated brackets around the upper limit of the sum mean to round down to the next integer.

1.3.2 Bayes–IRIS with Imperfect Inspections

When dealing with imperfect inspections, the Bayesian analysis in Section 1.3.1 becomes more complicated. As before, we use $P_0 \sim \text{beta}(0.5, 0.5)$ as a prior for p at the start of the first review period. In addition, we suppose that w, the probability of successfully identifying a non-compliant unit being inspected, has the following prior distribution that is independent of P_0

$$1-W_0 \sim \text{beta}(a_w, b_w). \tag{1.9}$$

At the end of the first review period, having observed x_1 non-compliant units from n_1 inspected units, p and w have the following joint posterior density (Gaba & Winkler, 1992):

$$f_{P_1,W_1}(p,w|x_1,n_1) = \sum_{y=0}^{n_1-x_1} c_y\, f_\beta(p; n_1-y+0.5, y+0.5) \times f_\beta(1-w; n_1-x_1-y+a_w, x_1+b_w), \tag{1.10}$$

where $f_\beta(\cdot; a, b)$ is the beta(a,b) density, $c_y = a_y / \sum_{z=0}^{n_1-x_1} a_z$, and

$$a_y = \binom{n_1-x_1}{y} \beta(n_1-y+0.5, y+0.5)\beta(n_1-x_1-y+a_w, x_1+b_w). \tag{1.11}$$

In the sum, we can interpret y as the true number of compliant units from the n_1 that were sampled.

In the case where p is small and w is known exactly, the posterior of p is approximately gamma distributed (Johnson & Gastwirth, 1991):

$$P_1 \approx \text{gamma}(x_1+0.5, w(n_1-x_1)-0.5). \tag{1.12}$$

Given a distribution for P_1, we can again obtain a distribution for X_2 by integrating the binomial distribution over the possible values of p. Again, by fixing w and supposing p to be small, we get

$$\text{P}(X_2 = x_2) \approx \binom{n_2}{x_2} \frac{(w(n_1-x_1)-0.5)^{x_1+0.5}}{(n_2-x_2+w(n_1-x_1)-0.5)^{x_1+x_2+0.5}} \frac{\Gamma(x_1+x_2+0.5)}{\Gamma(x_1+0.5)}, \tag{1.13}$$

where $\Gamma(a)$ is the gamma function evaluated at a. (Note that this is not a true distribution because summing the right-hand side over $x_2 = 0, \ldots, n_2$ does not give 1. The approximation is, nonetheless, reasonable for small x_2.) Putting

$R_2 = (1 - n_2 / N_2) X_2 / n_2$ and requiring $\mathrm{P}(R_2 > r) \le \alpha$, we can calculate n_2 as before. Our sample size for the second sampling period is the smallest n_2 for which

$$\sum_{x_2=0}^{\lfloor r n_2 N_2/(N_2-n_2) \rfloor} \binom{n_2}{x_2} \frac{(w(n_1-x_1)-0.5)^{x_1+0.5}}{(n_2-x_2+w(n_1-x_1)-0.5)^{x_1+x_2+0.5}} \frac{\Gamma(x_1+x_2+0.5)}{\Gamma(x_1+0.5)} \ge 1-\alpha. \quad (1.14)$$

1.4 The CSP Algorithm

The IRIS algorithm allows for periodic updating of the sampling rate, and in particular makes sure that the sampling rate does not drop too low when few non-compliant units are detected, but it does not respond quickly to a sudden increase in the non-compliance rate. In contrast, the CSP is designed to increase the sampling rate quickly if a cluster of non-compliant units is detected, and then reduce it again if the non-compliance proves to be short lived. The CSP was introduced by Dodge (1943) and later extended by Dodge and Torrey (1951) and Govindaraju and Kandasamy (2000). We present a general description of the CSP that covers most schemes, including the multilevel plans of Lieberman and Solomon (1955).

We suppose that we have $K \ge 2$ states that represent how alert we are to non-compliant units, with state 1 the least alert and state K the most alert. For each state k, we have a sampling rate f_k, a window length g_k (also called a clearance number), and compliance numbers c_k^+ and c_k^- that are used to determine when to change to a different alert level. If a unit arrives while we are in state k, we will inspect it with probability f_k. If we are in state k and c_k^+ or more of the previous g_k items inspected in state k are non-compliant then we increase the alert level (by one or more levels). If c_k^- or fewer of the previous g_k items inspected in state k have been non-compliant then we decrease the alert level (by one or more levels). We can increase the alert level after only c_k^+ inspections, but we need at least g_k before we can decrease it. Lieberman and Solomon (1955) restrict themselves to the case where $c_k^- = 0$, and suppose that changes in state are by just one level at a time.

In Tables 1.1 to 1.3 we give details for some CSP algorithms. Here, the *Up destination* is the state you move to when increasing the alert level and the *Down destination* is the state you move to when decreasing the alert level. Values for the AOQ are taken from Stephens (1995) and give the theoretical long-run average leakage rate. Here, $q = 1 - p$.

When applying CSP-1, CSP-2 or CSP-3, we need to choose a sampling rate, f, and one or more window sizes. The usual approach is to start with an acceptable leakage rate, r, and a range of plausible approach rates, $[p^-, p^+]$. Using the AOQ, we can then get a set of potential parameters. For example, for CSP-3 we have

$$S = \{(f, g_c, g_a) \colon \mathrm{AOQ}(p) \le r \quad \text{for all} \quad p \in [p^-, p^+]\}.$$

We can then choose parameters from S according to some secondary consideration such as minimising f or g_c. Unfortunately, this approach is very much dependent

Table 1.1. CSP-1 algorithm

$$\mathrm{AOQ}(p) = \frac{p(1-f)q^{g}}{f(1-q^{g})+q^{g}}$$

Alertness state	Sampling rate	Window size	Up threshold	Up destination	Down threshold	Down destination
2 (census)	1	g			0	1
1 (sampling)	f	1	1	2		

From Dodge (1943).

Table 1.2. CSP-2 algorithm

$$\mathrm{AOQ}(p) = \frac{p(1-f)q^{g_c}(2-q^{g_a})}{f(1-q^{g_c})(1-q^{g_a})+q^{g_c}(2-q^{g_a})}$$

Alertness state	Sampling rate	Window size	Up threshold	Up destination	Down threshold	Down destination
3 (census)	1	g_c			0	1
2 (alert)	f	g_a	1	3	0	1
1 (sampling)	f	1	1	2		

From Dodge and Torrey (1951).

Table 1.3. CSP-3 algorithm

$$AOQ(p) = \frac{p(1-f)q^{g_c}(1+q^{4}(1-q^{g_a}))}{f(1-q^{g_c})(1-q^{g_a+4})+q^{g_c}(1-q^{4}(1-q^{g_a}))+4fpq^{g_c}}$$

Alertness state	Sampling rate	Window size	Up threshold	Up destination	Down threshold	Down destination
4 (census)	1	g_c			0	1
3 (limbo)	1	4	1	4	0	2
2 (alert)	f	g_a	1	4	0	1
1 (sampling)	f	1	1	3		

From Dodge and Torrey (1951).

on the value of p^{+} and can result in values of f that are too small if p^{+} is small or values of f that are too large if p^{+} is too large. If f is too small then the algorithm is too slow to respond to changes in p and we have no guarantee of the statistical value of the information gained from our inspections. If f is too large then we waste resources through unnecessary sampling.

1.4.1 Combining IRIS and CSP

Our response to the problem of parameter selection for the CSP algorithm is to combine it with the IRIS algorithm. At the end of each review period, we choose f using IRIS to achieve a given leakage rate, r, with a confidence of $100(1-\alpha)\%$. Given f, we then choose the window sizes g, or g_a and g_c according to secondary considerations, which may be operational.

As before, we suppose that units arrive in blocks or review periods of size N_1, N_2, and so forth. Suppose that during period 1 we used a CSP algorithm with base sampling rate f_1. (At any time k, the actual sampling rate, S_k, will be either f_1 or 1.) Given that we observed x_1 non-compliant units out of n_1 units inspected during period 1, we can estimate n_2 using IRIS as described in Section 1.3. The base sampling rate for period 2 is then $f_2 = n_2 / N_2$.

For example, suppose that we wish to combine CSP-1 and IRIS to determine an inspection algorithm for a given pathway. If in the previous batch of $N_1 = 1000$ units, $n_1 = 500$ were inspected and $x_1 = 1$ non-compliant units were found. The goal is to choose n_2, the number of units to inspect from the next batch of size $N_2 = 1000$. We assume that we want the prediction distribution of the leakage rate to be lower than 1% with probability 0.95 and that the inspection effectiveness, w, is known to be 0.9. Solving Eq. 1.4 for n_2 yields a sampling rate of $s = n_2 / N_2 = 0.479$, which we round to 0.5. To choose the window length g for the CSP-1 algorithm, in the absence of any other criteria, we can use the formula for the AOQ given in Table 1.1. [Graphs of this function can be found in Dodge (1943).] Using the point estimate $p = x_1 / n_1 = 1/500$ and $f = s = 0.5$ from the preceding, we can choose g to achieve the desired AOQ. For example, for an AOQ of less than 0.095%, the clearance number (window length) would be $g = 0.5$.

Alternatively, in some circumstances the clearance number can be interpreted directly as a burden on the importer, representing a period of intense scrutiny during which the importer needs to demonstrate proper compliance. Given this interpretation, the magnitude could be chosen to reflect expert opinion.

The IRIS algorithm is well suited to a slowly changing approach rate, with reviews at fixed points in time. It is not designed to continually monitor for a sudden increase in the approach rate, and it doesn't have an automatic reaction if this occurs. There is no need to monitor for a decrease in the approach rate under IRIS; we just wait until the next review point.

CSP algorithms provide an immediate measured response to any increase in the approach rate. CSP algorithms enable us to increase the sampling rate temporarily when there is a suspicion that the approach rate has increased, and then reduce it if there is not a problem. Where the CSP algorithms have problems, however, is in the choice of parameters f and g (or g_a and g_c). Using the IRIS algorithm to choose f means that we can choose g (or g_a and g_c) safe in the knowledge that we have already controlled the expected leakage rate and how large it could reasonably be. We also know that our overall sampling rate will be large enough to ensure that we will continue to have a good estimate of the approach rate. In the example given