Extensions of Dorfman's Theory

Rui Santos, Dinis Pestana, and João Paulo Martins

Abstract

Economic impact of composite sampling is investigated in the realistic framework of tests with positive probability of false positive and of false negative results. Sensitivity and specificity when pooling samples are also discussed, using rarefaction as a framework.

1 Introduction

During the Second World War, Dorfman [2] used composite sampling in order to get a more efficient method of identifying syphilis infected soldiers: testing a group of n members is feasible and saves resources. A negative result of a compound test means that no one is infected and a positive result means that at least one of the members is infected. In this last case, individual tests to all the group soldiers would be mandatory, in order to identify all infected members. As the main goal is to minimize the number of expected tests to identify all the infected members, the main issue is to find the optimal group dimension n. The compound analyses allow us to save resources in a variety of problems aside from blood testing, cf. [1]. Besides, the optimal n depends on the researcher's main goal: to identify all the infected individuals (classification problem) or to estimate the prevalence rate (estimation problem).

R. Santos (✉) · J.P. Martins
School of Technology and Management, Polytechnic Institute of Leiria, CEAUL — Center of Statistics and Applications of University of Lisbon, Portugal
e-mail: rui.santos@ipleiria.pt; jpmartins@estg.ipleiria.pt

D. Pestana
Faculty of Sciences of Lisbon, University of Lisbon, CEAUL — Center of Statistics and Applications of University of Lisbon, Portugal
e-mail: dinis.pestana@fc.ul.pt

P.E. Oliveira et al. (eds.), *Recent Developments in Modeling and Applications in Statistics*, Studies in Theoretical and Applied Statistics,
DOI 10.1007/978-3-642-32419-2_19,

Dorfman's original idea was applied to qualitative analyses (presence or absence of the infection), without measuring any quantitative variable (antigenes or antibodies or bacteria counts, or proportion of specific cells, or weight or volume of some chemical compound). In his methodology, it was implicitly assumed that the blood tests had no errors. In the present work, we assume that each test may return a false negative or a false positive, i.e., the conditional probability of returning the true diagnostic in the presence or absence of the disease, known respectively as sensitivity and specificity, are essential operating characteristics of the test. Some previous works, concerned mainly with prevalence rate estimation, discuss misclassification issues. Nevertheless, most of these studies assume that pooling does not affect misclassification (cf. [8, 9, 12, 13]). Others studies, such as [7], provide simple models for modeling sensitivity without taking into account the number of infected members in the group. However, an infected member in the pooled sample can be excessively diluted and become undetectable in the compound test. Furthermore, the sensitivity and specificity of a compound test must depend of the number of infected individual on the group—the dilution effect, cf. [6]. Hierarchical models to capture the dilution effect in prevalence estimation were used in [14, 16], but they do not measure the probability of misclassification. Thus, the main goal of this work is to analyze the influence of dilution and rarefaction on sensitivity and specificity in the use of compound tests and Dorfman's classification methodology. Moreover, for low prevalence rates, it is shown that the dilution effect for just one (or two) infected element in the group is sufficient to evaluate misclassification in the use of pooled samples.

2 Dorfman's Theory

Let p denote the prevalence rate of the infection and the independent Bernoulli random variables X_i, with $i = 1, \ldots, N$, represent the presence $(X_i = 1)$ or absence $(X_i = 0)$ of the infection in the i-th population unit. A compound test on a sample of size n has negative result if none of those units is infected, i.e., $\sum_{i=1}^{n} X_i = 0$, and it has positive result if at least one element is infected, i.e., $\sum_{i=1}^{n} X_i \geq 1$. In this case individual tests have to be carried out to identify all infected members. In this chapter we follow Dorfman's procedure, although our methods can be applied to others methodologies, such as the ones proposed in [3,4,8,10] or [11]. Hence, the expected number of tests needed to identify infected units in the population using groups of n members is $\mathbb{E}[T] = N\left(\frac{1}{n} + 1 - (1-p)^n\right)$ (for simplicity, assume that $\frac{N}{n} \in \mathbb{N}$). Thus a simple quantification of the efficiency of compound versus individual tests is the relative cost $\mathbf{RC} = \frac{\mathbb{E}[T]}{N} = \frac{n+1}{n} - (1-p)^n$, $n \geq 2$, which can be used to find the most efficient value for n knowing the infection prevalence p. Dorfman [2] analyzed different prevalence rates and concluded that the compound test is more efficient if p is lower than approximately 0.3066 and the efficient value for n (represented by n^*) will decrease with p, without attaining the value $n = 2$, as described in Table 1. Observe that for $0.123 \leq p \leq 0.307$ we have $n^* = 3$

Table 1 Relative cost and n optimum for some prevalence rates p

p	0.31	0.30	0.13	0.12	0.07	0.05	0.025	0.01	0.005	10^{-3}	10^{-4}	10^{-5}	10^{-6}
n^*	1	3	3	4	4	5	7	11	15	32	100	317	1000
RC	1	0.990	0.675	0.650	0.502	0.426	0.305	0.196	0.139	0.063	0.020	0.006	0.002

Table 2 Joint probabilities of an individual test

	Individual test result		
Truth	X_i^+	X_i^-	
$X_i = 1$	$\varphi_s p$	$(1-\varphi_s)p$	p
$X_i = 0$	$(1-\varphi_e)q$	$\varphi_e q$	q
	$\varphi_s p + (1-\varphi_e)q$	$(1-\varphi_s)p + \varphi_e q$	1

and for $p > 0.307$ the compound test is not efficient ($n^* = 1$). For $p < 0.123$ the approximation $(1-p)^n \approx 1 - np$ (cf. [15]) provides the approximate optimal solution $n^* \approx p^{-0.5}$, whose maximum error is 1, with no significant influence on the relative cost. Thus $\mathbf{RC} \approx \frac{1}{n} + np$ and, for $n = n^*$, $\mathbb{E}[T] \approx 2N\sqrt{p}$. Finucan [3] proposes the smallest integer not less than $p^{-0.5} + 0.5$ for n^*. In all examples we will use the most efficient value for n, i.e. $n = n^*$.

3 The Inclusion of Errors in the Tests Results

Now consider that the tests results are subject to some sources of error. We shall denote $\varphi_s \in (0, 1]$ the sensitivity of one single test—probability of one positive test (X_i^+) in one infected sample ($X_i = 1$), i.e., $\mathbb{P}\left(X_i^+ | X_i = 1\right)$—and $\varphi_e \in (0, 1]$ the specificity of the test—probability of getting a negative result (X_i^-) from a not infected sample ($X_i = 0$), i.e., $\mathbb{P}\left(X_i^- | X_i = 0\right)$. Then, $1-\varphi_s$ will be the probability of one false negative test and $1-\varphi_e$ the probability of one false positive. Assuming that the results of the test and the condition of the tested units are independent, denoting $q = 1 - p$, we obtain the probabilities shown in Table 2.

3.1 Sensitivity and Specificity in Compound Tests

Let $I^{[n]} = \sum_{i=1}^{n} X_i$ represent the number of infected elements in a sample of size n and $I^{[i,n]} = \mathbb{P}(I^{[n]} = i) = \binom{n}{i} p^i q^{n-i}, i = 0, \ldots, n$. Let $X^{[+,n]}$ [resp. $X^{[-,n]}$] represent a positive [resp. negative] result on the compound test. The purpose of this section is to compare the sensitivity φ_s and the specificity φ_e of one simple test ($n = 1$) with the pooled sensitivity $\varphi_s^{[n]} = \mathbb{P}(X^{[+,n]} | I^{[n]} \geq 1)$ and pooled specificity $\varphi_e^{[n]} = \mathbb{P}(X^{[-,n]} | I^{[n]} = 0)$ of one compound test in a sample of size n.

Observe that the value of n does not affect the specificity; in fact, if we dissolve one cm^3 of blood from n not infected elements and then we choose, at random, one cm^3 from this mixture to test, then this sample of blood will not be infected, and

this framework is obviously equivalent to using just one cm^3 of blood from one non infected person. So $\varphi_e^{[n]} = \varphi_e$.

On the other hand, the sensitivity of the compound test $\varphi_s^{[n]}$ is a function of the number m of infected members in the group. Denote $\varphi_s^{[m,n]} = \mathbb{P}(X^{[+,n]}|I^{[n]} = m)$, and assume that $\varphi_s^{[n,n]} = \varphi_s$, i.e., that the sensitivity of the compound test when all members are infected is equal to the sensitivity of one simple test. Under these assumptions we shall get $0 < \varphi_s^{[1,n]} \leq \varphi_s^{[2,n]} \leq \cdots \leq \varphi_s^{[n,n]} = \varphi_s$, as a consequence of the dilution of the fluid and its rarefaction, and

$$\varphi_s^{[n]} = \frac{\sum_{j=1}^{n} \mathbb{P}\left(X^{[+,n]}|I^{[n]} = j\right) \mathbb{P}\left(I^{[n]} = j\right)}{\mathbb{P}\left(I^{[n]} \geq 1\right)} = \sum_{j=1}^{n} \varphi_s^{[j,n]} \frac{I^{[j,n]}}{1-q^n} = \sum_{j=1}^{n} \varphi_s^{[j,n]} \lambda_j, \tag{1}$$

where $\sum_{j=1}^{n} \lambda_j = 1$ and therefore $\varphi_s^{[n]}$ is a weighted mean of the sensitivities $\varphi_s^{[j,n]}$; so $\varphi_s^{[n]} \leq \varphi_s$ as a consequence of $\varphi_s^{[m,n]} \leq \varphi_s$, for every m. Hung and Swallow [6] applied a similar approach in order to analyze the robustness of prevalence estimation under misclassification concerned with the dilution effect. Nevertheless, they considered $\varphi_s = \varphi_e = 1$ (misclassification is only due the dilution effect) and proposed two different specific forms to $\varphi_s^{[j,n]}$ without providing general formulas for sensitivity and specificity in compound tests.

In fact, the expected number of infected members in the n elements, supposing $p \approx 0$ (thus $n^* \approx \frac{1}{\sqrt{p}}$), is $\frac{1}{n}$ and therefore $\lambda_i > \lambda_j$ to $i < j$. Moreover, the first weight λ_1 decreases with p but the remaining λ_i increase with p. For $p = 0.1$, $\lambda_1 \approx 0.85$ and $\lambda_2 \approx 0.14$; for $p = 0.01$ $\lambda_1 \approx 0.951$ and $\lambda_2 \approx 0.048$. In general, the $\varphi_s^{[j,n]}$ for $j > 2$ can be disregarded, and in most cases $\varphi_s^{[2,n]}$ can also be discarded (for $p = 0.001$, for instance, $\lambda_1 \approx 0.985$). So, the value $\varphi_s^{[1,n]}$ of the sensitivity of one compound test when just one of the n members is infected has a large impact on the total sensitivity of the compound test (without requiring any assumptions concerning the dilution effect). If the sensitivity $\varphi_s^{[1,n]}$ is low (compared with φ_s), we do not recommend using the compound test.

In order to compare $\varphi_s^{[m,n]}$ with φ_s we can use the weights $k_m^{[n]} = 1 - \varphi_s^{[m,n]} \varphi_s^{-1}$ where $1 > k_1^{[n]} \geq k_2^{[n]} \geq \cdots \geq k_n^{[n]} = 0$. Thus $\varphi_s^{[n]} = \varphi_s - \sum_{j=1}^{n} \lambda_j k_j^{[n]} \varphi_s$, and the difference between the sensitivity φ_s and the pooled sensitivity $\varphi_s^{[n]}$ is given by $\varphi_s - \varphi_s^{[n]} = \sum_{j=1}^{n} \lambda_j k_j^{[n]} \varphi_s$. So, when $p \approx 0$, we have $\varphi_s - \varphi_s^{[n]} \approx (\lambda_1 k_1^{[n]} + \lambda_2 k_2^{[n]}) \varphi_s$. We illustrate the use of those weights in Sect. 4.1 using the Poisson distribution.

In a compound test with n elements, the probability of each outcome is given in Table 3.

3.2 Sensitivity and Specificity in Dorfman's Methodology

Using Dorfman's methodology, for one infected person being identified by the test, the compound test must be positive (this happens with probability $\varphi_s^{[n]}$ which depends on $I^{[n]}$) and then, in the simple test, we also need to get a positive result

Table 3 Joint probabilities of a compound test

	Compound test result		
Truth	$X^{[+,n]}$	$X^{[-,n]}$	
$I^{[n]} \geq 1$	$\varphi_s^{[n]}\,(1-q^n)$	$\left(1-\varphi_s^{[n]}\right)(1-q^n)$	$1-q^n$
$I^{[n]} = 0$	$(1-\varphi_e)q^n$	$\varphi_e\, q^n$	q^n
	$\varphi_s^{[n]} - q^n\left(\varphi_s^{[n]} + \varphi_e - 1\right)$	$1-\varphi_s^{[n]} + q^n\left(\varphi_s^{[n]} + \varphi_e - 1\right)$	1

(with probability φ_s). Therefore, supposing that the tests results are independent, the sensitivity φ_{s_n} of the process will be given by

$$\begin{aligned}
\varphi_{s_n} &= \mathbb{P}\left(X_1^+ | X_1 = 1\right) = \sum_{i=0}^{n-1} \mathbb{P}\left(X_1^+ | X_1 = 1, I^{[n-1]} = i\right) \mathbb{P}\left(I^{[n-1]} = i\right) \\
&= \sum_{i=0}^{n-1} \mathbb{P}\left(X_1^+ | X_1 = 1\right) \mathbb{P}\left(X^{[+,n]} | I^{[n]} = i+1\right) I^{[i,n-1]} \\
&= \varphi_s \sum_{i=0}^{n-1} \varphi_s^{[i+1,n]} I^{[i,n-1]} \qquad (2)
\end{aligned}$$

and $\varphi_{s_n} \leq \varphi_s^2 \leq \varphi_s$, i.e., the sensitivity in Dorfman's methodology is smaller than the sensitivity of one simple test. There are two different possibilities for one not infected person being identified by the test: either the compound test is negative or the compound test is positive and the simple test is negative. So the specificity φ_{e_n} of the process will be given by

$$\begin{aligned}
\varphi_{e_n} &= \mathbb{P}\left(X_1^- | X_1 = 0\right) = \sum_{i=0}^{n-1} \mathbb{P}\left(X_1^- | X_1 = 0, I^{[n-1]} = i\right) \mathbb{P}\left(I^{[n-1]} = i\right) \\
&= \sum_{i=0}^{n-1} \left[\mathbb{P}\left(X_1^- | X_1 = 0\right) \mathbb{P}\left(X^{[+,n]} | I^{[n]} = i\right) + \mathbb{P}\left(X^{[-,n]} | I^{[n]} = i\right)\right] I^{[i,n-1]} \\
&= \left[\varphi_e + \varphi_e(1-\varphi_e)\right] q^{n-1} + \sum_{i=1}^{n-1} \left[\varphi_e \varphi_s^{[i,n]} + \left(1-\varphi_s^{[i,n]}\right)\right] I^{[i,n-1]} \\
&= 1 - (1-\varphi_e)\left[(1-\varphi_e)q^{n-1} + \sum_{i=1}^{n-1} \varphi_s^{[i,n]} I^{[i,n-1]}\right] \\
&= 1 - (1-\varphi_e)\xi \qquad (3)
\end{aligned}$$

where $\xi = \xi(\varphi_s^{[i,n]}, \varphi_e, p, n)$ is a weighted mean of $(1-\varphi_e)$ and $\varphi_s^{[i,n]}$, $i = 1, \ldots, n-1$, therefore $\xi \leq 1$ and consequently $\varphi_{e_n} \geq \varphi_e$ (the specificity in Dorfman's methodology is greater than in a simple test). If we do not include

the rarefaction factor (i.e., supposing $\varphi_s^{[m,n]} = \varphi_s, 1 \leq m \leq n$, cf. [8]), then $\varphi_{s_n} = \varphi_s^2 \leq \varphi_s$, which does not depend on n and corresponds to the maximum value for φ_{s_n} (best situation for φ_{s_n}). In this case the specificity φ_{e_n} is given by

$$\varphi_{e_n} = 1 - (1 - \varphi_e)\varphi_s - (1 - \varphi_e)(1 - \varphi_e - \varphi_s)q^{n-1}, \tag{4}$$

which depends on n, increases with q^{n-1} (for suitable values for φ_s and φ_e, i.e., $\varphi_s + \varphi_e > 1$) and corresponds to the minimum value for φ_{e_n} (worst case for φ_{e_n}). If we use the approximation $n \approx \frac{1}{\sqrt{p}}$, then $q^{n-1} \approx (1-p)^{\frac{1}{\sqrt{p}}-1}$ decreases with $p \in (0, 0.31)$ and, therefore, φ_{e_n} also decreases with p. Nevertheless, if φ_e is close to one, the value of $(1 - \varphi_e)(1 - \varphi_e - \varphi_s)$ will be very close to zero and, therefore, the value of p will not imply great changes in the specificity φ_{e_n}.

Hence, in this simplified situation, the gain of specificity using Dorfman's methodology is generally negligible, but the loss of sensitivity can be large. For instance, for $\varphi_s = 0.9$ we get $\varphi_{s_n} = 0.81$ and the probability of a false negative (a most serious error, as an infected person is not detected) increases 9 % taking into account the rarefaction factor, the loss would be even greater.

4 The Number of Bacteria

Let us consider the problem of detecting the presence of some type of bacteria to test the contamination of some substance, for instance of yogurt at an industrial unit.

First, let us assume that the contamination rate is equal to p. Thus, if we choose n yogurts at random to test, the number of contaminated ones is a random variable $I \frown \mathbf{B}(n, p)$. Let us assume further that the number of bacteria Y_i^* in one ml taken from the i-th yogurt, chosen at random, is zero if it is not contaminated. Otherwise, it is modeled by some random variable Y_i with discrete distribution $\mathbf{D}$ with parameters vector θ and support in $\mathbb{N}_0$, that is $Y_i \frown \mathbf{D}(\theta)$.

So, when pooling together one ml from n yogurts, the number of bacteria present in the pooled sample will be $B_n = \sum_{i=1}^n Y_i^* = \sum_{i=1}^I Y_i$. If the compound test uses one ml of the mixture of yogurts, as in the simple test, i.e., if we choose, at random, one ml from the n ml of the amalgamated sample, assuming that the mixture is homogeneous, the number of bacteria in one ml of this compound is $B_1 \frown \mathbf{B}\left(B_n, \frac{1}{n}\right)$ (we can get the same result by considering $B_1 = \sum_{i=1}^n W_i^* = \sum_{i=1}^I W_i$ where the random variable W_i is described by the hierarchic model $W_i \sim \mathbf{B}\left(T, \frac{1}{n}\right)$ where $T \sim \mathbf{D}(\theta)$). Hierarchical pooling models using continuous distributions were used in [14, 16] to develop procedures for HIV prevalence estimation with the dilution effect.

Table 4 exhibits the results for some commonly used count distributions, namely the main "basic count distributions" with extended Panjer's recursion, cf. [5].

Table 4 Example with some common count distributions

	$\mathbf{D}(\theta)$	$B_1\vert_{I=i}$
Poisson	$Y_i \sim \mathbf{P}(\lambda)$, $\lambda > 0$	$B_1\vert_{I=i} \sim \mathbf{P}\left(i\frac{\lambda}{n}\right)$
Binomial	$Y_i \sim \mathbf{B}(m, p_1)$	$B_1\vert_{I=i} \sim \mathbf{B}\left(im, \frac{p_1}{n}\right)$
Negative binomial	$Y_i \sim \mathbf{NB}(r, p_1)$, $y \in \mathbb{N}_0$ $\mathbb{P}(Y_i = y) = \binom{y+r-1}{r-1} p_1^r (1-p_1)^y$	$B_1\vert_{I=i} \sim \mathbf{NB}\left(ir, \frac{np_1}{np_1+1-p_1}\right)$
Logarithmic	$Y_i \sim \mathbf{Log}(\theta)$, $\theta \in (0,1)$, $\mathbb{P}(Y_i = y) = -\frac{1}{\ln(1-\theta)} \frac{\theta^y}{y}$ $y \in \mathbb{N}$	$f_{W_i}(x) = \begin{cases} \frac{\ln\left(1-\theta+\frac{\theta}{n}\right)}{\ln(1-\theta)} & x = 0 \\ \frac{-1}{x\ln(1-\theta)} \left(\frac{\frac{\theta}{n}}{1-\theta\left(1-\frac{1}{n}\right)}\right)^x & x \in \mathbb{N} \end{cases}$

4.1 Sensitivity and Specificity

For illustration purposes, our analyses will be restricted to the case of the number of bacteria in one contaminated yogurt being modeled by the Poisson distribution, i.e., $Y_i \sim \mathbf{P}(\lambda)$ (the others cases are analogous).

Let us suppose that the test never returns wrong results (it identifies always the presence of bacteria if any exists in the ml analyzed). Hence, the sensitivity is the probability of existing at least one bacterium in one ml from on infected yogurt, and therefore $\varphi_s = \mathbb{P}(Y_i > 0) = 1 - \mathrm{e}^{-\lambda}$; and the specificity will be equal to certitude (one not contaminated yogurt does not have any bacterium). Under these conditions, the number of bacteria Y^* in one ml is modeled by a zero-inflated Poisson distribution with probability mass function

$$f_{Y^*}(x) = \begin{cases} (1-p) + p\mathrm{e}^{-\lambda} & x = 0 \\ p\,\dfrac{\mathrm{e}^{-\lambda}\,\lambda^x}{x!} & x \in \mathbb{N}. \end{cases} \tag{5}$$

In the compound test, the number of bacteria in 1 ml of yogurt will be characterized by a Poisson distribution with parameter $m\frac{\lambda}{n}$, where $m \leq n$ represents the number of infected yogurts in the group. As the number m of contaminated yogurt is $m \sim \mathbf{B}(n, p)$, the number of bacteria Y_n^* in 1 ml of the pooled sample has probability function given by

$$f_{Y_n^*}(x) = \begin{cases} \displaystyle\sum_{i=0}^{n} I^{[i,n]} \mathrm{e}^{-i\frac{\lambda}{n}} & x = 0 \\ \displaystyle\sum_{i=1}^{n} I^{[i,n]}\,\frac{\mathrm{e}^{-i\frac{\lambda}{n}}\,\left(i\frac{\lambda}{n}\right)^x}{x!} & x \in \mathbb{N}, \end{cases} \tag{6}$$

that is equal to $f_{Y^*}(x)$ when $n = 1$. So, even in the compound test, we still have $\varphi_{e_n} = 1$ (if none of the yogurts is contaminated then the test is always negative, as we obtain in (3) using $\varphi_e = 1$) and, denoting by $\mathbf{C}$ the contaminated yogurts,

$$\varphi_{s_n} = \mathbb{P}\left(Y_n^* \geq 1 | Y_1 \in \mathbf{C}\right) \times \mathbb{P}\left(Y_1 \geq 1\right) = \sum_{i=0}^{n-1} I^{[i,n-1]} \left(1 - \mathrm{e}^{-(i+1)\frac{\lambda}{n}}\right) \left(1 - \mathrm{e}^{-\lambda}\right),$$

which corresponds to formula (2) with $\varphi_s^{[i+1,n]} = 1 - \mathrm{e}^{-(i+1)\frac{\lambda}{n}}$ and $\varphi_s = 1 - e^{-\lambda}$.

Applying the weights $k_m^{[n]}$ computed in Sect. 3.1, we get $k_m^{[n]} = 1 - \frac{1-e^{-\frac{m}{n}\lambda}}{1-e^{-\lambda}}$, and the difference $\varphi_s - \varphi_s^{[n]}$ is approximately $\left(\lambda_1 k_1^{[n]} + \lambda_2 k_2^{[n]}\right) \varphi_s$. If we consider, as an example, a prevalence rate of $p = 0.01$ ($n^* = 11$) and a parameter $\lambda = 10$, then $\varphi_s - \varphi_s^{[n]} \approx 0.39\varphi_s$, so the difference between the single test sensitivity and the pooled sensitivity is approximately 40 %!

We can also include in the model the probability of having an extra source of error associated with the test itself, specifically $1 - \varphi_{e_T}$ of false positive and probability $1 - \varphi_{s_T}$ of false negative, the resulting specificity is

$$\begin{aligned}
\varphi_{e_n} &= \varphi_{e_T} \mathbb{P}\left(Y_n^* = 0 | Y_1 \notin \mathbf{C}\right) + (1 - \varphi_{s_T}) \mathbb{P}\left(Y_n^* > 0 | Y_1 \notin \mathbf{C}\right) + \\
&+ \varphi_{e_T} \left[(1 - \varphi_{e_T}) \mathbb{P}\left(Y_n^* = 0 | Y_1 \notin \mathbf{C}\right) + \varphi_{s_T} \mathbb{P}\left(Y_n^* > 0 | Y_1 \notin \mathbf{C}\right)\right] \\
&= \varphi_{e_T} \varphi_0 + (1 - \varphi_{s_T})(1 - \varphi_0) + \varphi_{e_T} \left[(1 - \varphi_{e_T})\varphi_0 + \varphi_{s_T}(1 - \varphi_0)\right] \\
&= 1 - (1 - \varphi_{e_T})\varphi_{s_T} - (1 - \varphi_{e_T})(1 - \varphi_{e_T} - \varphi_{s_T})\varphi_0, \qquad (7)
\end{aligned}$$

where $\varphi_0 = \sum_{i=0}^{n-1} I^{[i,n-1]} \left(\mathrm{e}^{-i\frac{\lambda}{n}}\right)$. This formula is analogous to the one previously deduced in Sect. 3.2 (but with φ_0 instead of q^{n-1}). The sensitivity (where both the compound and single tests have to be positive) is

$$\begin{aligned}
\varphi_{s_n} &= \left[(1 - \varphi_{e_T}) \mathbb{P}\left(Y_n^* = 0 | Y_1 \in \mathbf{C}\right) + \varphi_{s_T} \mathbb{P}\left(Y_n^* > 0 | Y_1 \in \mathbf{C}\right)\right] \times \\
&\quad \times \left[\varphi_{s_T} \mathbb{P}\left(Y_1 \geq 1\right) + (1 - \varphi_{e_T}) \mathbb{P}\left(Y_1 = 0\right)\right] \\
&= \left[(1 - \varphi_{e_T})\varphi_1 + \varphi_{s_T}(1 - \varphi_1)\right] \left[\varphi_{s_T}\left(1 - \mathrm{e}^{-\lambda}\right) + (1 - \varphi_{e_T})\mathrm{e}^{-\lambda}\right] \\
&= \varphi_{s_T}^2 + \varphi_1 \mathrm{e}^{-\lambda}(1 - \varphi_{e_T} - \varphi_{s_T})^2 + \varphi_{s_T}(\varphi_1 + \mathrm{e}^{-\lambda})(1 - \varphi_{e_T} - \varphi_{s_T}) \qquad (8)
\end{aligned}$$

where $\varphi_1 = \sum_{i=0}^{n-1} I^{[i,n-1]} \left(\mathrm{e}^{-(i+1)\frac{\lambda}{n}}\right)$. This incorporates the probabilities of erroneous outcomes due to the operational characteristics of the test himself (where we obtain $\varphi_{s_T}^2$, as in the cases previously analyzed) and of sampling errors (in one contaminated group, we get no bacterium in the compound sample).

Table 5 Relative cost, sensitivity, and specificity for two subpopulations

p_1	0.0075	0.005	0.001	0.0001	0.005	0.005	0.005	0.005	0.001	0.001	0.001	0.001
ω_1	0.5	0.5	0.5	0.5	0.25	0.10	0.75	0.9	0.25	0.10	0.75	0.9
p_2	0.0125	0.015	0.019	0.0199	0.011(6)	0.010(5)	0.025	0.055	0.013	0.011	0.037	0.091
$\mathbf{RC}$	0.1939	0.1887	0.1650	0.1467	0.1928	0.1945	0.1806	0.1698	0.1823	0.1905	0.1393	0.1132
$\varphi_{e_n}^{[S]}$	0.9950	0.9950	0.9953	0.9953	0.9949	0.9950	0.9951	0.9953	0.9951	0.9950	0.9956	0.9961
$\varphi_{s_n}^{[S]}$	0.5936	0.5606	0.4810	0.4048	0.5757	0.5935	0.5386	0.5111	0.5506	0.5746	0.4026	0.3388

5 Dealing with Subpopulations

Let us consider that our population is divided into k groups (subpopulations) with weights $\omega_1, \omega_2, \ldots, \omega_k$, where $\sum_{i=1}^{k} \omega_i = 1$, and prevalence rates $p_1, p_2, \ldots, p_k$. If the presence of subpopulations in the compound tests is disregarded, the results for efficiency, sensitivity, and specificity are the same as the ones obtained with one population with prevalence rate given by $p = \sum_{i=1}^{k} \omega_i p_i$.

On the other hand, if this information is not disregarded in the Dorfman's methodology, the expected number of tests will be $\mathbb{E}[T_S] = \sum_{i=1}^{k} \omega_i N \min_{n_i} \left\{ \frac{n_i+1}{n_i} - (1-p_i)^{n_i}, 1 \right\}$, $n_i \geq 2$, where n_i is chosen in function of the p_i (cf. Table 1). The relative cost, $\mathbf{RC} = \frac{\mathbb{E}[T_S]}{N}$, is the weighted mean of the relatives cost of each subpopulation, cf. [3].

For the simplified case (without rarefaction), the sensitivity $\varphi_{s_n}^{[S]} = \varphi_s^2 = \varphi_{s_n}$. But the specificity is $\varphi_{e_n}^{[S]} = 1-(1-\varphi_e)\varphi_s - (1-\varphi_e)(1-\varphi_e-\varphi_s)\sum_{i=1}^{k} \omega_i q_i^{n_i-1}$. For $p \approx 0$, we get $\varphi_{e_n}^{[S]} \approx 1-(1-\varphi_e)\varphi_s - (1-\varphi_e)(1-\varphi_e-\varphi_s)\sum_{i=1}^{k} \omega_i (1-p_i)^{\frac{1}{\sqrt{p_i}}-1}$. As $f(p) = (1-p)^{\frac{1}{\sqrt{p}}-1}$ is a convex function for $p \in (0,1)$, using Jensen's inequality $\sum_{i=1}^{k} \omega_i f(p_i) \geq f\left(\sum_{i=1}^{k} \omega_i p_i\right)$, and therefore ($p = \sum_{i=1}^{k} \omega_i p_i$),

$$\sum_{i=1}^{k} \omega_i (1-p_i)^{\frac{1}{\sqrt{p_i}}-1} \geq \left(1 - \sum_{i=1}^{k} \omega_i p_i\right)^{\frac{1}{\sqrt{\sum_{i=1}^{k} \omega_i p_i}}-1} = (1-p)^{\frac{1}{\sqrt{p}}-1}. \quad (9)$$

Hence, if we analyze the subpopulation strata separately, the specificity increases.

As an example, consider the Poisson example with $\lambda = 10$, $p = 0.01$, $\varphi_{s_T} = \varphi_{e_T} = 0.95$. Working with the population as a whole, using (7) and (8), we have $\varphi_{e_n} = 0.9949$, $\varphi_{s_n} = 0.5780$ and $\mathbb{E}[T] = 0.1956N$ (i.e., $\mathbf{RC} = 0.1956$). Table 5 shows the results testing the two subpopulations separately (with $k = 2$). Testing the subpopulations separately is slightly more efficient than otherwise, but the loss of sensitivity should be analyzed previously.

6 Conclusion

The use of compound tests optimizes the number of expected tests and consequently saves resources. On the other hand, the possible loss of sensitivity and specificity can restrict the usefulness of this methodology. Although the pooled sensitivity and specificity are widely studied, the rarefaction factor is often disregarded. In this work we stress the importance of determining the $\varphi_s^{[1,n]}$ sensitivity to control the total pooled sensitivity. Therefore, the rarefaction factor is crucial in pooled sample analysis, but for low prevalence rates the knowledge of the sensitivity $\varphi_s^{[1,n]}$ can be a simple alternative to measure the overall misclassification, compared to the use of more general dilution models (cf. [6, 7]).

Acknowledgements The authors thank the referees for their very useful comments. This research has been supported by National Funds through FCT—Fundação para a Ciência e a Tecnologia, project PEst-OE/MAT/UI0006/2011, and PTDC/FEDER.

References

1. Boswell, M.T., Gore, S.D., Lovison, G., Patil, G.P.: Annotated bibliography of composite sampling, Part A: 1936–92. Environ. Ecol. Stat. **3**, 1–50 (1996)
2. Dorfman, R.: The detection of defective members in large populations. Ann. Math. Stat. **14**, 436–440 (1943)
3. Finucan, H.M.: The blood testing problem. Appl. Stat. J. Roy. Stat. C **13**, 43–50 (1964)
4. Gill, A., Gottlieb, D.: The identification of a set by successive intersections. In: Information and Control, pp. 20–25. Ellis Horwood, Chichester (1974)
5. Hess, K.T., Liewald, A., Schmidt, K.D.: An extension of Panjer's recursion. Astin Bull. **32**(2), 283–297 (2002)
6. Hung, M., Swallow, W.: Robustness of group testing in the estimation of proportions. Biometrics **55**, 231–237 (1999)
7. Hwang, F.K.: Group testing with a dilution effect. Biometrika **63**, 671–673 (1976)
8. Kim, H., Hudgens, M., Dreyfuss, J., Westreich, D., Pilcher, C.: Comparison of group testing algorithms for case identification in the presence of testing errors. Biometrics **63**, 1152–1163 (2007)
9. Liu, S.C., Chiang, K.S., Lin, C.H., Chung, W.C., Lin, S.H., Yang, T.C.: Cost analysis in choosing group size when group testing for Potato virus Y in the presence of classification errors. Ann. Appl. Biol. **159**, 491–502 (2011)
10. Sterret, A.: On the detection of defective members of large populations. Ann. Math. Stat. **28**, 1033–1036 (1957)
11. Sobel, M., Elashoff, R.: Group testing with a new goal, estimation. Biometrika **62**, 181–193 (1975)
12. Tu, X.M., Litvak, E., Pagano, M.: Studies of AIDS and HIV surveillance, Screening tests: Can we get more by doing less? Stat. Med. **13**, 1905–1919 (1994)
13. Tu, X.M., Litvak, E., Pagano, M.: On the informativeness and accuracy of pooled testing in estimating prevalence of a rare disease: Application to HIV screening. Biometrika **82**(2), 287–297 (1995)
14. Wein, L., Zenios, S.: Pooled testing for HIV screening: capturing the dilution effect. Oper. Res. **44**, 543–569 (1996)