Appendix

The probability of not detecting an illegal pack with a legal Codentify code

We employed a modified Bernoulli formula for independent repeated sampling without replenishment taking into account the increasing probability of detection after each draw.
The probability of not finding a duplicate code in a random sample of packs (i.e. the chance of missing a duplicate; outcome number 2) where all packs have replicated Codentify codes can be calculated using multinomial distribution formula:

$$\boldsymbol{P(no\ duplicates) = \frac{m^n \binom{k}{n}}{\binom{km}{n}}} \qquad \textbf{(A)}$$

where $\boldsymbol{n}$ is the number of items sampled, $\boldsymbol{k}$ is the number of unique digital codes that have been stolen, $\boldsymbol{m}$ is the number of times each digital code has been replicated, and

$\binom{k}{n}$ **= k!/[n! * (k-n)!]**

$\binom{km}{n}$ **= km!/[n! * (km-n)!]**

k! = k * (k – 1) * (k – 2) * (k – 3) * ….* 1
n! = n * (n – 1) * (n – 2) * (n – 3) * ….* 1
km! = km * (km – 1) * (km – 2) * (km – 3) * ….* 1
0! = 1

The legitimate packs whose codes were replicated need to be added to the pool of illegal packs since Codentify is deficient in the sense that it cannot determine which of the two packs with matching codes is legitimate.

The formula implies that if there is only one item sampled (**n=1; n! = 1**) or if all codes are harvested codes and used just once (**m=1;** $\binom{km}{n} = \binom{k}{n}$), the probability of missing a duplicate in a given batch of packs is one.

In practical terms, we choose the probability **P** (e.g. 0.9) and solve for **n** while knowing **N**, **m**, and **k**.

To take into account that packs with duplicated codes are mixed together with packs carrying legitimate code, the number of inspections is calculated as follow:

$$\boldsymbol{Number\ of\ inspections = n * \frac{1}{a}}$$

where $\boldsymbol{a}$ represents the estimated share of packs with stolen codes (plus the packs that were the source for code replication) in the distribution.

The probability of not detecting an illegal pack using a material-based security

The probability of not finding an illegal pack in a random sample of packs using a material-based security solution, or the probability of outcome number 3 or 4 are based on the hypergeometric distribution:

$$\boldsymbol{P(no\ illicit) = \frac{\binom{N(1-a)}{n}}{\binom{N}{n}}} \qquad \textbf{(B)}$$

Where **N** represents the number packs available for inspection, ***a*** represents the estimated share of illegal packs in the distribution, and

$\boldsymbol{\binom{N\,(1-a)}{n}}$ **= N(1-a)!/[n! * {N(1-a)-n}!]**

$\boldsymbol{\binom{N}{n}}$ **= N!/[n! * (N-n)!]**

n!, N!, (N-n)!, N(1-a)!, N(1-a)-n}! are factorials defined as above.