Решение задачи

Денисов Никита

18 февраля 2022 г.

Задача

Прибор для выявления брака на фабрике имеет вероятность ошибки 5% (и первого и второго рода), процент брака составляет 5% от всего объёма выпускаемой продукции.

1) Какая вероятность того, что мы выявили брак, если прибор выдал положительный результат - "продукция бракованная"?

2) Почему же в жизни все-таки используют такие приборы? Что можно было бы изменить в процедуре поиска брака, не меняя точности прибора, так, чтобы вероятность из первого вопроса P(брак|"+") выросла?

3) Какое правило можно сформулировать для выбора точности прибора, зная процент бракованной продукции?

$$p = 0.05 \text{ — вероятность брака.}$$
$$q_0 = 0.05 \text{ — вероятность ложноположительного срабатывания прибора.}$$
$$q_1 = 0.05 \text{ — вероятность ложноотрицательно срабатывания прибора.}$$

1. Хотим посчитать $P(\text{Брак} \mid +)$. Воспользуемся формулой Байеса:
   $P(\text{Брак} \mid +) = \frac{P(+ \mid \text{Брак}) \cdot P(\text{Брак})}{P(+)}$. $P(+)$ посчитаем по формуле полной вероятности:

   $P(+) = p(1 - q_1) + (1 - p)q_0 = 0.05 \cdot 0.95 + 0.95 \cdot 0.05$

   $P(+ \mid \text{Брак}) = 1 - q_1 = 0.95$

   Итого, получаем:
   $P(\text{Брак} \mid +) = \frac{P(+ \mid \text{Брак}) \cdot p}{P(+)} = \frac{0.95 \cdot 0.05}{2 \cdot 0.95 \cdot 0.05} = \frac{1}{2}$

2. Такие приборы в жизни используют, чтобы отметать хотя бы какую-то часть бракованных изеделий. Но отвергаются также и изделия без брака. Это все равно лучше, чем допускать все изделия, ведь брак может привести к печальным последствиям.
   Если точность прибора изменить нельзя, то можно попробовать запускать прибор несколько раз на каждой продукции. Если наша цель более точно определять бракованная ли деталь, то стоит k раз запускать прибор и в зависимости от количества положительных срабатываний определять бракованная ли деталь.
   Пусть мы сделали k запусков на одной продукции и получили q положительных результатов. Определим положительный результат прибора, если $q \geq t(k)$. Тогда посмотрим:

   $P(\text{Брак} \mid +) = \frac{P(+ \mid \text{Брак}) \cdot p}{P(+)}$
   $P(+) = p \cdot \sum_{q=t(k)}^{k} \binom{k}{q}(1 - q_1)^q \cdot q_1^{k-q} + (1 - p) \cdot \sum_{q=t(k)}^{k} \binom{k}{q} q_0^q (1 - q_0)^{k-q}$ — формула полной вероятности и перебрали количество положительных срабатываний, которые нас устраивают.

$P(+ \mid \text{Брак}) = \sum_{q=t(k)}^{k} \binom{k}{q} (1-q_1)^q \cdot q_1^{k-q}$

Хотим $P(\text{Брак} \mid +) = \frac{A}{A+B} \geq x > \frac{1}{2}$, где
$A = p \cdot \sum_{q=t(k)}^{k} \binom{k}{q} (1-q_1)^q \cdot q_1^{k-q}$,
$B = (1-p) \cdot \sum_{q=t(k)}^{k} \binom{k}{q} q_0^q (1-q_0)^{k-q}$

Написал скрипт на Python и проверил какие $k, t(k)$ нам подходят. Перебрал k до 10 и t вплоть до k и подсчитал искомую величину: $\frac{A}{A+B}$

Скрипт выдал результат, что уже при $k = 2, t(k) = 2$ величина $\frac{A}{A+B}$ оказывается равна 0.95 — довольно-таки хорошо
Если нужна вероятность больше, то при $k = 4, t = 4$ вероятность оказывается равной: 0.99985, а при $k = 8, t = 8$: 0.999999998.

3. Теперь зная p определим какие q_0, q_1 нужны для прибора чтобы получить $P(\text{Брак} \mid +) \geq x$.
   Мы уже выяснили:
   $P(+) = p(1-q_1) + (1-p)q_0$
   $P(\text{Брак} \mid +) = \frac{P(+ \mid \text{Брак}) \cdot p}{P(+)} = \frac{p(1-q_1)}{p(1-q_1)+(1-p)q_0} \geq x$. Выразим q_0 и определим для него границу при фиксированном q_1:
   $\frac{p(1-q_1)(1-x)}{(1-p)x} \geq q_0$
   Можем и наоборот, выразить q_1 и определить границу для него при фиксированном q_1:
   $p - pq_1 \geq -xpq_1 + px + xq_0 - xpq_0 \Leftrightarrow -q_1 p(1-x) \geq p(x-1) + q_0 x(1-p) \Leftrightarrow -q_1 \geq \frac{p(x-1)+q_0 x(1-p)}{p(1-x)} \Leftrightarrow q_1 \leq \frac{p(1-x)-q_0 x(1-p)}{p(1-x)}$

   Зная эти границы можно экспериментальным путем определить подходящие q_0, q_1 по заданному x, например, написав скрипт.