

Выбраковка результатов химического анализа

Решение вопроса: резко отличающиеся результаты – это грубая погрешность (промах) или нет, и их необходимо использовать при общей статистической обработке?

- **Уровень значимости** – это максимальная вероятность того, что ошибка превзойдет некое предельное (критическое) значение $\pm\Delta x_{кр}$, т.е. такое значение, что появление этой ошибки можно рассматривать как следствие значимой (неслучайной) причины.

$$\beta = 1 - P$$

Уровень значимости (%) показывает, сколько раз в каждых ста испытаниях мы рискуем ошибиться, принимая случайное событие за значимое.

Q-критерий

1. располагают результаты анализа в ряд по нарастанию $x_1 < x_2 < x_3 < \dots < x_n$
2. Вызывают сомнения x_{\min} и x_{\max}
3. Вычисляют величину $Q_{\text{эксп}} = \frac{(x_2 - x_1)}{(x_n - x_1)}$ ил $Q_{\text{эксп}} = \frac{(x_n - x_{n-1})}{(x_n - x_1)}$
4. Из таблиц для заданных числа измерений n и доверительной вероятности P находят $Q_{\text{табл}}$
5. Если $Q_{\text{эксп}} < Q_{\text{табл}}$, то результат сохраняют – это не промах!

T-критерий

Прием, аналогичный расчету доверительных интервалов

1. Рассчитываем \bar{x} с учетом сомнительных значений
2. Рассчитываем стандартное отклонение S
3. Задаем уровень \bar{x} значимости $\beta = 1 - P = 1 - 2\alpha_{ст}$
4. Вводим $T_{кр} = |x_{кр} - \bar{x}| / S = \Delta x_{кр} / S$
или $\Delta x_{кр} = T_{кр} \cdot S$
5. Находим в таблице критический параметр $T_{кр}$ при заданных \bar{x} и β
6. Запишем $\pm \Delta x = \pm T_{кр} S$
7. Трактовка: в полученном интервале с заданной доверительной вероятностью могут находиться все значения, кроме промахов!

Сравнение дисперсий

Задача: сравнение с точки зрения воспроизводимости результатов анализа

- методик определения компонента в пробе;
- работы различных лабораторий по одной и той же методике;
- полученных на разных приборах;
- При работе на приборе в различных диапазонах измерений.

Рассматривают:

Дисперсии не совпадают, но несовпадение носит случайный характер, т.к. они характеризуют одну и ту же генеральную совокупность с генеральной дисперсией σ^2 . Но S_1^2 и S_2^2 могут существенно отличаться.

Вопрос: является ли различие двух дисперсий случайным (и выборки можно объединять) или значимым (выборки

F-критерий

Основан на распределении Фишера

1. Имеют две нормально распределенные выборочные совокупности размером n_1 и n_2 с выборочными дисперсиями S_1 и S_2 и степенями свободы $f_1 = n_1 - 1$, $f_2 = n_2 - 1$
2. Рассчитывают $F = S_1^2 / S_2^2$ (чтобы $F > 1$)
3. Находят $F_{кр}$ в таблице при заданных параметрах
4. Сравнивают: Если $F < F_{кр}$, анализы равнозначны, и выборки можно обрабатывать совместно.

Критерий Батлера

Используют, если выборок больше двух

ВЫБОРКА 1. $f_1=n-1, S_1^2$

ВЫБОРКА 2. $f_2=n-1, S_2^2$

ВЫБОРКА k. $f_k=n-1, S_k^2$

Рассчитывают средневзвешенную дисперсию

$$S_{nk}^2 = \frac{\sum_{i=1}^k f_i S_i^2}{n} = \frac{\sum_{i=1}^k f_i S_i^2}{\sum_{i=1}^k n}$$

Число степеней свободы совокупной выборки

$$f_{nk} = \sum_{i=1}^k f_i$$

Рассчитывают

$$B = 2.3(f_{nk} \lg S_{nk}^2 - \sum_{i=1}^k f_i \lg S_i^2) \quad C = 1 + \frac{1}{3(k-1)} \left(\sum_{i=1}^k \frac{1}{f_i} - \frac{1}{f_{nk}} \right)$$

Если $B/C < \chi^2$, дисперсии однородны, выборки объединяют

Критерий Батлера с удовлетворительной точностью применим для выборок с $n \geq 6$

Критерий Кохрана

- Если объемы выборок равны

$$n_1 = n_2 = n_3 = \dots = n_k$$

Рассчитываю: $G = S_{max}^2 / (S_1^2 + S_2^2 + S_3^2 + \dots + S_n^2)$

Находят в таблицах $G_{кр}$ при $\beta=0,05$ и $\beta=0,01$ и $f=k-1$

Если $G < G_{кр}$, то выборки однородны и их можно объединять

Сравнение средних. t-критерий Стьюдента

- Анализ одного образца

Серия А. n_A, S_A^2, \bar{x}_A

Серия В. n_B, S_B^2, \bar{x}_B

Пусть по F-критерию они значимо не отличаются.

Вопрос: значимо ли различие средних?

Используют t-критерий Стьюдента

Рассчитываю:
$$t_{AB} = \frac{\bar{x}_A - \bar{x}_B}{S_{AB}} \sqrt{\frac{n_A \cdot n_B}{n_A + n_B}}$$

Где средневзвешенная дисперсия
$$S_{AB}^2 = \frac{(n_A - 1)S_A^2 + (n_B - 1)S_B^2}{n_A + n_B - 2}$$

Выборочные средние \bar{x}_A и \bar{x}_B различаются значимо, если $t_{AB} > t_{p,f}$ для принятой доверительной вероятности p и числа степеней свободы объединенной выборки $f = n_A + n_B - 2$.

Подтверждение «Нуль-гипотезы»

- Совпадает ли найденное экспериментально с истинным значением μ ?

\bar{x}

Оценка предела обнаружения

- Предел обнаружения – это минимальное количество m_{\min} (или концентрация C_{\min}) определяемого компонента, которое может быть обнаружено с заданной достаточно высокой доверительной вероятностью (0,95 или 0,99).
- Предел обнаружения в единицах аналитического сигнала – это минимальный сигнал Y_{\min} , который можно с уверенностью отличить от сигнала холостой пробы (фона) $Y_{\text{фона}}$.
- Связь C_{\min} с Y_{\min} через коэффициент инструментальной чувствительности

$$C_{\min} = (Y_{\min} - Y_{\text{фона}}) / S_{y/c}$$

Как выбрать Y_{\min} ?

- Ошибки I рода – ошибка «неоткрытия» - принятие сигнала определяемого компонента за сигнал фона – ошибка пропуска аналитического сигнала определяемого компонента.
- Ошибки II рода – ошибки «переоткрытия» - принятие сигнала фона за сигнал компонента - т.е. обнаружение компонента в пробе, когда его нет.
- Уровень дискриминации сигнала Y_d или уровень выбраковки сигнала делит все сигналы на две части:
 - $Y < Y_d$ – сигнал фона
 - $Y > Y_d$ – сигнал пробы

Если принять, что сигнал фона и сигнал пробы распределены по одному закону, их стандартные отклонения приблизительно равны, то можно принять

$$\underbrace{Y_{\min} - Y_d}_{\text{I рода}} = \underbrace{Y_d - Y_{\text{фона}}}_{\text{II рода}}$$

Принято $Y_{\min} - Y_d = Y_d - Y_{\text{фона}} = 3\sigma_{\text{фона}}$, тогда $Y_{\min} - Y_{\text{фона}} = 6\sigma_{\text{фона}}$

Отсюда $C_{\min} = \Delta Y_{\min} / S_{Y/c} = 6\sigma_{\text{фона}} / S_{Y/c}$

- Если принять для рассмотрения нормальный закон распределения, то вероятности ошибок I и II рода очень малы

$$P_I = (Y_{\min} < Y_{д}) = P(Y_{\min} < Y_{\min} - 3\sigma_{Y_{\min}})$$

$$P_{II} = (Y_{\phi} > Y_{д}) = P(Y_{\phi} > Y_{\phi} + 3\sigma_{Y_{\phi}})$$

0.0014

- Если использовать неравенство Чебышева (при 3σ), то вероятности выше

$$P_I = (Y_{\min} < Y_{д})$$

$$P_{II} = (Y_{\phi} > Y_{д})$$

0,055