

М. А. Д р а н и ц ы н а, Т. В. З а х а р о в а (Москва, ВМК МГУ).
Свойства процедуры двух односторонних t-тестов в рамках модели с недостающими данными.

Недостающие данные могут приводить к систематическим ошибкам в клинических исследованиях (см. [1]). Возникновение и объем недостающих данных зависит от множества причин, в том числе от длительности эксперимента, таким образом, запись биологических и медицинских сигналов подвержена возникновению недостающих данных. Примером такого сигнала является фармакокинетическая кривая, отражающая зависимость концентрации некоторого соединения в цельной крови, сыворотке или плазме человека или животного после его введения в разных формуляциях от времени. В исследованиях фармакокинетики обычно получают такие записи для разных формуляций у одного и того же субъекта. Для сравнения таких сигналов (то есть формуляций) используют метрики на основе наблюдаемых данных (например площади под такой кривой, оцененную методом трапеций), для чего используют процедуру двух односторонних тестов Шуирмана (см. подробнее [2]). Гипотезу H_0 и альтернативу H_A согласно исследовательскому предположению об эквивалентности метрик, характеризующих такой сигнал, можно представить следующим образом:

$$\begin{aligned} H_0 : \mu_1 - \mu_2 &\leq \delta_{min} \text{ или } \mu_1 - \mu_2 \geq \delta_{max} \\ H_A : \delta_{min} &< \mu_1 - \mu_2 < \delta_{max}, \end{aligned} \quad (1)$$

где δ_{min} и δ_{max} — заданные пределы признания эквивалентности, а μ_1 и μ_2 — математические ожидания метрик сравниваемых формуляций (сигналов 1 и 2).

Для проверки гипотез используют два односторонних t-теста:

$$\begin{aligned} t_1 &= \frac{\bar{Y}_1 - \bar{Y}_2 - \delta_{min}}{\hat{\sigma}_d \sqrt{\frac{2}{n}}} > t(\alpha, 2n - 2) \\ t_2 &= \frac{\bar{Y}_1 - \bar{Y}_2 - \delta_{max}}{\sigma_d \sqrt{\frac{2}{n}}} < -t(\alpha, 2n - 2), \end{aligned} \quad (2)$$

где n — число записей сигналов 1 и 2 ($n_1 = n_2$), $t(\alpha, 2n - 2)$ — квантиль центрального t-распределения с $2n - 2$ степенями свободы, $\hat{\sigma}_d$ — обобщенная выборочная дисперсия разностей между записями сигналов 1 и 2 у субъектов.

Ранее было показано, что при использовании процедуры двух односторонних тестов величина ошибки I рода не превышает заданный уровень α . Предположим, что для сигнала 1 наблюдаются недостающие данные. Пусть q — уровень полноты данных, то есть доля данных, оставшаяся от изначальных данных, $0 < q \leq 1$, ($1 - q$ — доля пропущенных данных в выборке), тогда $\tilde{Y}_1 = \bar{Y}_1 + \ln(q)$ — выборочное среднее метрик для сигнала 1 при условии недостающих данных. Можно показать, что в рамках такой модели при использовании той же процедуры проверки гипотез величина ошибки I рода может превышать значение α (см. [3]).

Работа выполнена при поддержке РФФИ (проект 18-07-00252).

СПИСОК ЛИТЕРАТУРЫ

1. *EMA Committee for Medicinal Products for Human Use (CHMP) Guideline on Missing Data in Confirmatory Clinical Trials*, EMA/CPMP/EWP/1776/99 Rev. 1.
2. Драницына М. А., Захарова Т. В., Ниязов Р. Р. Свойства процедуры двух односторонних тестов для признания биоэквивалентности лекарственных препаратов. — Ремедиум. Журнал о рынке лекарств и медицинской технике, 2019, в. 3, с. 40–47.
3. Захарова Т. В., Тархов А. А. Оценка уровня значимости критерия Шуирманна для проверки гипотезы биоэквивалентности при наличии пропущенных данных. — Информатика и ее применения, 2019 (в печати).