běrů testujeme hypothesu, že střední hodnota rozdílů je rovna nule.) Jiným zobecněním by byl necentrální α test, při kterém by však bylo nutno specifikovat formu distribuční funkce.

6. β-test. Máme-li dva nespárované nezávislé výběry $x_1, x_2, \ldots, x_n$ a $y_1, y_2, \ldots, y_m$, lze nulovou hypothesu, pravící, že všech $n + m$ hodnot bylo vybráno z téhož rozdělení, otestovat pomocí β-testu. Stačí zjistit počet inversí mezi oběma výběry, t. j. počet případů, že $y_i < y_j$ pro $i = 1, 2, \ldots, m$ a $j = 1, 2, \ldots, n$. Za předpokladu, že nulová hypothesa platí, bude mít tento počet inversí β-rozdělení, kde $N - n = m$. Dolní 5%-ní a 10%-ní kritické hodnoty pro n, $m \leq 10$ jsou pro β-test uvedeny v tabulkách 3. a 4. S tabelovanými kritickými hodnotami opět srovnáváme β vypočtené na základě těch inversí, kterých je méně.

7. γ-test. Mějme N nezávislých párů pozorování $(x_1, y_1), (x_2, y_2), \ldots, (x_N, y_N)$ z dvojrozměrného rozdělení. Nulovou hypotesu, že x a y jsou nezávislé, lze pak otestovat pomocí γ-testu. Stačí zjistit počet inversí v posloupnosti $y^{(1)}$, $y^{(2)}, \ldots, y^{(N)}$ vytvořené tak, že odpovídající hodnoty x_i tvoří rostoucí posloupnost $x^{(1)} < x^{(2)} < \ldots < x^{(N)}$. Za předpokladu, že nulová hypothesa platí, bude mít tento počet inversí γ-rozdělení.

LITERATURA

[1] *Wilcoxon F.*, Individual comparisons by ranking methods, Biometrics Bull. 1, 80—83 (1945).

[2] *H. B. Mann* and *D. R. Whitney*, On a test of whether one of two random variables is stochastically larger than the other, Ann. Math. Stat. vol. *18*, 50—61, (1947).

[3] *Van der Vaerden*, Order tests for the two sample problem and their power. Proceedings A, 55, 453—458 (1952);
Order tests for the two sample problem (second communication), Proceedings A, 56, 303—310; (1953).
Order tests for two sample problem (third communication), Proceedings A, 56, 311—316 (1953).

Резюме

НЕКОТОРЫЕ ПОРЯДКОВЫЕ РАСПРЕДЕЛЕНИЯ И ИХ ПРИЛОЖЕНИЯ

ЯРОСЛАВ ХАЕК (Jaroslav Hájek), Прага.

(Поступило в редакцию 24/III 1954 г.)

В предлагаемой работе выводятся три типа распределений и три соответствующих теста α, β, и γ; при выводе используются элементарные рассуж-

дения о множестве $\{1, 2, \ldots, N\}$ натуральных чисел $1, 2, \ldots, N$. Асимптотическая эффективность α-теста и β-теста равна 0,955.

α-распределение. Мы будем считать все подмножества множества $\{1, 2, \ldots, N\}$ элементарным событиямии. Пусть вероятность каждого элементарного событяия равна 2^{-N}. Сумма α всех чисел в каждом элементарном событии есть случайная величина, производящий полином которой $P_N(t)$ дан соотношением

$$P_N(t)\, 2^N = \prod_{n=1}^{N} (1 + t^n) .$$

Средняя величипа $E(\alpha)$, дисперсия $D^2(\alpha)$ и остальные характеристики даны формулами на стр. 19. α-распределение стремится к нормальному распределению, как непосредственно вытекает из теоремы Ляпунова.

γ-распределение. Рассмотрим все перестановки множества $\{1, 2, \ldots, N\}$, считая каждую из них элементарным событием. Пусть $(N!)^{-1}$ есть вероятность каждого такого события. Число γ всех инверсий в каждой перестановке является случайной величиной. Производящий полином $R_N(t)$ можно составить по уравнению (1).

β-распределение. Рассмотрим систему всех подмножеств, содержащих n элементов ($n \leqq N$) множества $\{1, 2, \ldots, N\}$, причем условимся считать каждое подмножество элементарным событием. Число β всех инверсий между подмножеством и дополнительным к нему множеством в множестве $\{1, 2, \ldots, N\}$ является случайной величиной. Под инверсией между двумя множествами мы подразумеваем каждую пару чисел (p, q), где $p > q$ и p принадлежит первому, а q — второму множеству.

Элементы каждой перестановки можно разделить на две группы; первая группа содержит первые n элементов перестановки, а вторая содержит остальные $N - n$ элементов. Итак, $\gamma = \gamma_1 + \gamma_2 + \beta$, где через $\gamma_1(\gamma_2)$ обозначено число всех инверсий в первой (второй) группе, а β означает число всех инверсий между обеими группами. Производящие полиномы γ, γ_1 и γ_2 даны уравнением (1). Отсюда следует, что производящий полином $Q_{N,n}(t)$ β-распределения удовлетворяет следующему соотношению:

$$R_N(t) = R_n(t)\, R_{N-n}(t)\, Q_{N,n}(t) .$$

Полином $Q_{N,n}(t)$ можно вычислить по формуле (2). Очевидно,

$$Q_{N,n}(t) = Q_{N,N-n}(t) ,$$

значит, β-распределение симметрично. Формулы для средней $E(\beta)$ и дисперсии $D^2(\beta)$ приведены на стр. 21 и 22.

Пусть теперь $N \to \infty$. Предположим, что n постоянно. Если в (2) заменить t выражением $e^{itN^{-1}}$, то получится характеристическая функция $\varphi(t)$ распределения βN^{-1}, откуда следует, что β-распределение стремится к распределению суммы n независимых ортогональных случайных величин. В том случае, когда как $n \to \infty$, так и $N \to \infty$, получим нормальное распределение.

α-тест. Пусть дано N пар независимых наблюдений (x_i, y_i), $i = 1, 2, \ldots, N$, при которых условия A влияли на наблюдения x_i, а условия B влияли на наблюдения y_i. Пусть, кроме того, дальнейшие условия C влияли на обе величины x_i и y_i одинаковым образом, но изменялись от пары к паре. Нашей задачей будет проверить нулевую гипотезу: каждая пара наблюдений представляет две величины, выбранные независимо друг от друга из одного и того же непрерывного распределения, которое, конечно, может меняться от пары к паре. Расположим абсолютные величины разностей $d_i = x_i - y_i$ в виде возрастающей последовательности

$$|d^{(1)}| < |d^{(2)}| < \ldots < |d^{(N)}|$$

и поставим в соответствие неблюдениям (x_i, y_i) $i = 1, 2, \ldots, N$, подмножество множества $\{1, 2, \ldots, N\}$, содержащее все числа k, для которых $d^{(k)} > 0$. Из нулевой гипотезы следует, что вероятности всех таких подмножеств одинаковы. Кроме того, сумма порядковых индексов положительных разностей, расположенных по абсолютной величине, имеет α-распределение. Следовательно, можно ставнить вычисленное α с табулированным критическим значением для соответствующей степени значительности. α-тест использован для анализа классического опыта Дарвина, опубликованного в „*The Design of Experiments*" и касающегося растений, возникших скрещиванием и самооплодотворением. В этом случае мы получим $\alpha = 24$. В таблице 2 для $N = 15$ и 5% мы находим значение 25,3. t-тест Стьюдента дает $t = 2{,}148$. Итак, оба теста дают практически одинаковые результаты.

α-тестом можно пользоваться даже в том случае, когда данные распределения не являются непрерывными.

β-тест. Пусть $x_1, x_2, \ldots, x_n$ и $y_1, y_2, \ldots, y_m$ — две независимые непарные выборки. Предположим, что справедлива нулевая гипотеза, что все $n + m$ значений были выбраны из одного и того же распределения. В таком случае можно воспользоваться β-тестом. Достаточно определить число инверсий между выборками. Это число обладает β-распределением, где $N - n = m$. Нижние 5%-ные и 10%-ные критические значения затабулированы.

γ-тест. Пусть $(x_1, y_1), \ldots, (x_N, y_N)$ — N друг от друга независимых пар наблюдений из двухмерного распределения. Если x и y независимы (нулевая гипотеза), то можно применить γ-тест. Расположим значения x_i в виде возрастающей последовательности $x^{(1)} < x^{(2)} < \ldots < x^{(N)}$ и предположим, что соответствующие значения y_i будут $y^{(1)}, y^{(2)}, \ldots, y^{(N)}$. Если справедлива нулевая гипотеза, то число инверсий в упомянутой последовательности обладает γ-распределением.