Записки научных
семинаров ПОМИ
Том 454, 2016 г.

М. И. Ревяков

РАНЖИРОВАНИЕ И СЕЛЕКЦИЯ ПОПУЛЯЦИЙ ПО ВЫБОРОЧНЫМ СРЕДНИМ

§1. Введение

В работе [1] было уделено внимание вопросу повышения достоверности статистических выводов при увеличении размера выборки. Эти выводы касались параметров семейств распределений и вышеуказанная проблема возникала тогда, когда семейства не обладали нетривиальными достаточными статистиками. В настоящей статье, как и в [1], предлагается оперировать выборочными средними в ряде ситуаций, так или иначе связанных с ранжированием и селекцией популяций.

Вопрос монотонного возрастания достоверности статистических выводов, в виде уменьшения риска с ростом числа наблюдений, развивается в нескольких направлениях. В теореме 1 по сравнению с [1] дополнительно рассматривается ситуация с неодинаковыми симметричными законами, в которых эти параметры фигурируют, а также удалось снять требование симметричности законов, если напротив, они принадлежат одному и тому же семейству с параметром сдвига. Перевод, согласно рекомендации [2], задачи ранжирования популяций на путь проверки гипотез дал возможность исследовать монотонное стремление к нулю ошибок первого и второго рода при бесконечном увеличении размерности выборочного среднего.

Следует отметить, что пункт 2.2 базируется на работе [3], в которой обобщается известный результат работы [4] об остроконечности случайных величин (с.в.), использованный в пункте 2.1 вслед за [1]. Это позволяет нам усреднять наблюдения над разнораспределёнными с.в. (следствие 2), связанными с варьированием параметра масштаба в двухпараметрическом семействе. В случае, когда среднеквадратическое отклонение подвержено постепенному изменению, а наблюдения разнесены во времени, тогда полезные приложения более всего следует

Ключевые слова: достоверность выводов, проверка гипотез, логарифмически вогнутые плотности, мажоризация, статистический критерий, неравенство Лоренца.

ожидать из макромира (пример 1). Дело в том, что именно там можно рассчитывать на то, что значения параметра сдвига, подлежащие сравнению, будут оставаться неизменными на протяжении длительного времени.

Поскольку мы рассматриваем статистические процедуры, обладающие симметрией, то оказывается, что в случае, когда речь идёт о *ранжировании* популяций, инвариантность возникает относительно группы перестановок (ср. [5,6]). Однако для *селекции* популяций пришлось привлечь группу преобразований $G = \{g_1, g_2\} : g_1 X = X,\ g_2 X = 2\theta_0 - X$, где θ_0 – уровень контроля (доказательство теоремы 4). Сочетание двух этих групп должно позволить в дальнейшем выйти на решение смешанных задач ранжирования с селекцией.

Попутно затрагивается вопрос о рандомизации. Показано, что для симметричных задач ранжирования и селекции популяций использование рандомизированных решающих процедур нецелесообразно, если число возможных решений конечно, а фигурирующие случайные величины непрерывны.

Наконец, привлечение известного неравенства Лоренца способствовало расширению возможности приложений теоретических исследований к оптимизации структуры электрических цепей. Тем более, если учесть, что пример 2 в принципе допускает распространение естественным образом полученных результатов с двухканальных цепей на многоканальные.

Статья состоит из 4 разделов (включая Введение).

В разделе 2 решается задача проверки гипотез, касающихся сравнения параметров распределений двух с.в. из одного семейства. Более простой вариант затрагивает семейства с параметром сдвига. Здесь установлены условия, при которых увеличение размера выборочного среднего приводит к монотонному уменьшению ошибок первого и второго рода. Последующее принципиальное усложнение заключается в рассмотрении двухпараметрических семейств распределений.

Раздел 3 посвящён селекции “хороших” популяций из семейств с параметром сдвига, когда разграничение задаётся контрольным фиксированным значением параметра.

В разделе 4 хорошо известное неравенство Лоренца используется в целях оптимального заполнения двухканальной электрической цепи усилителями на основе их тестирования.

§2. Проверка гипотез в рамках задачи ранжирования популяций

В работе [7] рассмотрена следующая задача. Пусть $X^{(1)}$ и $X^{(2)}$ – две независимые действительные с.в. с плотностями из одного семейства $f_\theta(x)$, функциональный вид которого известен, а величины параметров θ_1 и θ_2 неизвестны. Задача заключается в проверке гипотезы H: $\theta_1 < \theta_2$ относительно конкурирующей гипотезы H$'$: $\theta_1 > \theta_2$.

В [7] эта задача решалась методом последовательного анализа по критерию отношения вероятностей. В настоящем разделе мы собираемся осуществить проверку этих гипотез на основании сравнений статистик от фиксированного числа наблюдений (n) над с.в. $X^{(1)}$ и $X^{(2)}$. В предположении отсутствия нетривиальных достаточных статистик у семейства $f_\theta(x)$ мы хотим подобрать такие статистики, чтобы при оптимальном решении ошибки первого и второго рода (α_n и β_n) монотонно убывали с ростом n и стремились к 0 при $n \to \infty$.

2.1. Однопараметрическое семейство. Сначала мы ограничимся семействами с параметром сдвига, т.е. когда $f_\theta(x) = h(x-\theta)$. В качестве статистического критерия мы выбираем правило, по которому принимается гипотеза H, если $Z_n^{(1)} < Z_n^{(2)}$, где $Z_n^{(j)}$, $j = 1, 2$, – выборочные средние из n независимых наблюдений над с.в. $X^{(j)}$. В противном случае, когда $Z_n^{(1)} > Z_n^{(2)}$, принимается гипотеза H$'$. (В настоящей статье рассматриваются только статистические задачи, обладающие симметрией).

Пусть семейство $f_\theta(x)$ имеет монотонное отношение правдоподобия, т.е. является M.L.R.-семейством. Это означает, что $h(t)$ лог-вогнута [2, 9.2]. Кроме того, мы допускаем, что при $\theta_1 = \theta_2$ принятие любой гипотезы не является ошибкой. Теперь в соответствии с рекомендациями [2, 3.3] будем рассматривать нашу задачу проверки гипотез как проблему с решениями d_1 (принятие H: $\theta_1 < \theta_2$) и d_2 (принятие H$'$: $\theta_1 > \theta_2$) и с функцией потерь $L(\theta_1, \theta_2; d_j) = L_j(\theta_1, \theta_2)$. Для того, чтобы выбранный статистический критерий оказался инвариантным относительно группы перестановок двух координат, мы полагаем $L_2(\theta_1, \theta_2) = L_1(\theta_2, \theta_1)$, причём функция $L_1(\theta_1, \theta_2) = 0$ при $\theta_1 \leqslant \theta_2$, не убывает по θ_1 и не возрастает по θ_2.

Мы можем воспользоваться тем фактом (ср. [8]), что для $Z_n^{(j)} = \sum_{i=1}^{n} X_i^{(j)}/n$, $j = 1, 2$, плотность $h_n(z-\theta_j)$ является лог-вогнутой. Это означает, что она имеет M.L.R.-свойство. Таким образом, для всех n (нерандомизированный) статистический критерий, введённый в начале этого раздела, окажется о п т и м а л ь н ы м в том смысле, что он будет минимаксным и обладающим при любых (θ_1, θ_2) минимальным риском среди инвариантных критериев, основанных на выборочных средних [5, 6]. Риск (средние потери) при таком критерии есть

$$R_n(\theta_1,\theta_2)=L_2(\theta_1,\theta_2)\,\mathbf{P}\left[Z_n^{(1)}>Z_n^{(2)}\right] \text{ при } \theta_1<\theta_2,\; R_n(\theta_2,\theta_1)=R_n(\theta_1,\theta_2).$$

Покажем, что в нашей ситуации $R_n(\theta_1, \theta_2)$ при любых (θ_1, θ_2) с ростом n, монотонно не возрастая, стремится к 0. Иначе говоря, нужно установить, что при $\theta_1 < \theta_2$ ошибки $\alpha_n(\theta_1, \theta_2) = \beta_n(\theta_2, \theta_1) = \mathbf{P}\left[Z_n^{(1)} > Z_n^{(2)}\right]$ монотонно стремятся к 0.

Последующие рассуждения опираются на работу [4], в связи с чем мы обсудим некоторые аспекты мажоризации [9]. Говорят, что вектор $b = (b_1, \ldots, b_n)$ *мажорируется* вектором $a = (a_1, \ldots, a_n)$, и записывают $a \succ b$, если $\sum\limits_{i=1}^{n} a_i = \sum\limits_{i=1}^{n} b_i$ и $\sum\limits_{i=1}^{k} a_{[i]} \geqslant \sum\limits_{i=1}^{k} b_{[i]}$, $k = 1, 2, \ldots, n-1$, где $a_{[1]} \geqslant \ldots \geqslant a_{[n]}$, $b_{[1]} \geqslant \ldots \geqslant b_{[n]}$ – упорядоченные по убыванию (невозрастанию) компоненты векторов a и b соответственно.

Введём формальные обозначения:

$$P_1 = \mathbf{P}\left[\sum_{i=1}^{n} p'_i X_i^{(1)} < \sum_{i=1}^{n} p'_i X_i^{(2)}\right], \quad P_2 = \mathbf{P}\left[\sum_{i=1}^{n} p_i X_i^{(1)} < \sum_{i=1}^{n} p_i X_i^{(2)}\right].$$

Теорема 1. *Пусть $X_1^{(j)}, \ldots, X_n^{(j)}$, $j = 1, 2$, – независимые с.в. с лог-вогнутыми плотностями $l_j(x-\theta_j)$, причём $\theta_1 < \theta_2$, а $l_j(t) = l_j(-t)$ для всех t, если $l_1(t) \not\equiv l_2(t)$. Пусть далее $p \succ p'$, $p_i, p'_i \geqslant 0$, $i = 1, \ldots, n$, $\sum\limits_{i=1}^{n} p_i = \sum\limits_{i=1}^{n} p'_i = 1$. Тогда $P_1 \geqslant P_2$.*

Замечание. Это есть уточнение и дополнение к теореме 5 в [1]. Отсюда следует, в частности, что из условий теоремы 5, следствия 2 и теоремы 6 в [1] можно убрать требование симметричности плотностей.

Доказательство. Обозначим

$$\widetilde{F}_1(t) = \mathbf{P}\left[\sum_{i=1}^{n} p'_i \overset{o}{X}{}_i^{(1)} \leqslant t\right], \quad \widetilde{F}_2(t) = \mathbf{P}\left[\sum_{i=1}^{n} p_i \overset{o}{X}{}_i^{(1)} \leqslant t\right],$$

$$\widehat{F}_1(t) = \mathbf{P}\left[\sum_{i=1}^{n} p'_i \overset{o}{X}{}_i^{(2)} \leqslant t\right], \quad \widehat{F}_2(t) = \mathbf{P}\left[\sum_{i=1}^{n} p_i \overset{o}{X}{}_i^{(2)} \leqslant t\right],$$

и $\widetilde{f}_1(t) = \widetilde{F}'_1(t)$, $\widetilde{f}_2(t) = \widetilde{F}'_2(t)$, где $\overset{o}{X}{}_1^{(j)}, \ldots, \overset{o}{X}{}_n^{(j)}$ – независимые с.в. с плотностями $l_j(x)$. Тогда

$$\begin{aligned} P_k &= \int_{-\infty}^{\infty} \widehat{F}_k(t-\theta_1)\, \widetilde{f}_k(t-\theta_2)\, dt \\ &= \int_{-\infty}^{\infty} \widehat{F}_k(\vartheta + t)\, \widetilde{f}_k(t)\, dt = \mathbf{P}\left[\chi_{1k} - \chi_{2k} \leqslant \vartheta\right], \quad k = 1, 2, \end{aligned}$$

где $\chi_{j1} = \sum\limits_{i=1}^{n} p'_i \overset{o}{X}{}_i^{(j)}$ и $\chi_{j2} = \sum\limits_{i=1}^{n} p_i \overset{o}{X}{}_i^{(j)}$, $j = 1, 2$, и $\vartheta = \theta_2 - \theta_1 > 0$. Следовательно,

$$\chi_{11} - \chi_{21} = \sum_{i=1}^{n} p'_i \big(\overset{o}{X}{}_i^{(1)} - \overset{o}{X}{}_i^{(2)}\big)$$

и

$$\chi_{12} - \chi_{22} = \sum_{i=1}^{n} p_i \big(\overset{o}{X}{}_i^{(1)} - \overset{o}{X}{}_i^{(2)}\big).$$

Полагаем $Y_i = \overset{o}{X}{}_i^{(1)} - \overset{o}{X}{}_i^{(2)}$, $i = 1, \ldots, n$. Из [9, 18.B.1] легко следует, что Y_i лог-вогнута. Симметричность Y_i следует из симметричности $\overset{o}{X}{}_i^{(1)}$ и $\overset{o}{X}{}_i^{(2)}$, а в случае $l_1(t) \equiv l_2(t)$ – из того, что разность независимых одинаково распределенных (н.о.р.) величин симметрична. Поскольку Y_i, $i = 1, \ldots, n$, – н.о.р. величины, а $\vartheta \geqslant 0$, то теорема 2.3 в [4] означает, что $\mathbf{P}\big[\sum\limits_{i=1}^{n} p'_i Y_i \leqslant \vartheta\big] \geqslant \mathbf{P}\big[\sum\limits_{i=1}^{n} p_i Y_i \leqslant \vartheta\big]$, т.е. $P_1 \geqslant P_2$. □

Следующий результат является нашей целью.

Следствие 1. *Пусть $X_1^{(j)}, X_2^{(j)}, \ldots$, $j = 1, 2$, – независимые с.в. с лог-вогнутыми плотностями $l_j(x-\theta_j)$, причём $\theta_1 < \theta_2$, а $l_j(t) = l_j(-t)$ для всех t, если $l_1(t) \not\equiv l_2(t)$. Тогда $\mathbf{P}\,[Z_n^{(1)} < Z_n^{(2)}]$ не возрастает по n.*

Доказательство. Заметим, что $p = (1/n,\ 1/n, \ldots,\ 1/n,\ 0) \succ (1/(n+1),\ 1/(n+1), \ldots,\ 1/(n+1),\ 1/(n+1)) = p'$, где каждый вектор состоит из $n+1$ компонент. Теперь требуемый результат непосредственно вытекает из теоремы 1. $\square$

При наших требованиях симметричности статистических процедур следует воспользоваться следствием 1 при $l_1(t) \equiv l_2(t)$, т.е. когда конкурирующие с.в. принадлежат одному и тому же семейству с параметром сдвига. Тогда получаем, что ошибки $\alpha_n(\theta_1, \theta_2) = \beta_n(\theta_2, \theta_1) = \mathbf{P}\left[Z_n^{(1)} > Z_n^{(2)}\right]$ монотонно не возрастают с ростом n. Далее, учитывая, что с.в. с лог-вогнутой плотностью имеет все моменты [10] (в том числе первый), легко находим, что с.в. $W = X^{(1)} - X^{(2)}$ имеет среднее $\theta_1 - \theta_2 = -\vartheta < 0$. Тогда из теоремы Хинчина [11, 2.2, с. 3] следует, что $\mathbf{P}\left[Z_n^{(1)} - Z_n^{(2)} > 0\right]$ стремится к 0 при $n \to \infty$.

2.2. Двухпараметрическое семейство. Рассмотренная ниже ситуация опирается на следующий результат в [3], обобщающий теорему 2.3 в [4], которая применялась в 2.1.

Напомним, что с.в. W_1 превосходит W_2 в смысле отношения правдоподобия, с записью $W_1 \geqslant^{lr} W_2$, если $f_{W_1}(t)/f_{W_2}(t)$ является неубывающей функцией от t, где $f_{W_j}(t)$ представляет плотность W_j, $j = 1, 2$.

Теорема 2 ([3]). *Пусть $X_1, \ldots, X_n$ – независимые симметричные с.в. с лог-вогнутыми плотностями $f_1, \ldots, f_n$ соответственно. Если*

$$|X_1| \geqslant^{lr} \cdots \geqslant^{lr} |X_n|,$$

тогда

$$\mathbf{P}\left[\sum_{i=1}^{n} p_i' X_i \leqslant \vartheta\right] \geqslant \mathbf{P}\left[\sum_{i=1}^{n} p_i X_i \leqslant \vartheta\right],$$

при условии $\vartheta > 0,\ p \succ p',\ p_i, p_i' \geqslant 0,\ i = 1, \ldots, n,\ \sum_{i=1}^{n} p_i = \sum_{i=1}^{n} p_i' = 1.$

Рассмотрим двухпараметрическое семейство плотностей с параметрами сдвига и масштаба: $f_{\lambda,\theta}(t) = \lambda h(\lambda(t-\theta))$ и в нём n пар случайных величин $X_i^{(1)},\ X_i^{(2)},\ i = 1, \ldots, n$, с плотностями $\lambda_i h(\lambda_i(t-\theta_1))$ и $\lambda_i h(\lambda_i(t-\theta_2))$ соответственно, при неизвестных θ_1 и θ_2. Мы предполагаем, что наблюдения над $X_i^{(1)}$ и $X_i^{(2)}$ осуществляются при значениях параметра масштаба λ_i, о которых известно, что $0 < \lambda_1 \leqslant \lambda_2 \leqslant \ldots \leqslant \lambda_n$.

Прежде всего, обозначив $Z_n^{(j)} = \sum_{i=1}^{n} X_i^{(j)}/n$, $j = 1, 2$, легко получаем, что плотность с.в. $Z_n^{(j)}$ имеет вид $h_n(x - \theta_j)$ и к тому же она лог-вогнута. Это означает, согласно [5,6], что для всех n статистический критерий, введённый в начале раздела 2.1, является оптимальным.

Далее, нам понадобятся две леммы.

Лемма 1. *Если $X_i^{(1)}$, $X_i^{(2)}$ – н.о.р. случайные величины с плотностью $f_i(r) = \lambda_i h(\lambda_i r)$, которая является лог-вогнутой, тогда $W_i = X_i^{(1)} - X_i^{(2)}$ имеет плотность вида $\widehat{f}_i(t) = \lambda_i \widehat{h}(\lambda_i t)$, которая симметрична и лог-вогнута.*

Доказательство. Плотность W_i есть

$$\widehat{f}_i(t) = \int_{-\infty}^{\infty} f_i(t + r)\, f_i(r) dr = \lambda_i^2 \int_{-\infty}^{\infty} h(\lambda_i t + \lambda_i r)\, h(\lambda_i r) dr$$
$$= \lambda_i \int_{-\infty}^{\infty} h(\lambda_i t + s)\, h(s) ds = \lambda_i \widehat{h}(\lambda_i t).$$

Функция $\widehat{f}_i(t)$ симметрична как плотность разности н.о.р. случайных величин и лог-вогнута как свёртка лог-вогнутых функций. □

Лемма 2. *Пусть независимые случайные величины X и Y имеют плотности $\widehat{f}_1(t) = \lambda_1 \widehat{h}(\lambda_1 t)$ и $\widehat{f}_2(t) = \lambda_2 \widehat{h}(\lambda_2 t)$ соответственно, причём функция $\widehat{h}(\cdot)$ дифференцируема, симметрична и лог-вогнута. Тогда при $0 < \lambda_1 \leqslant \lambda_2$ имеем*

$$|X| \geqslant^{lr} |Y|.$$

Доказательство. В силу симметричности плотностей, согласно теореме 2, требуется показать, что $\widehat{f}_1(t)/\widehat{f}_2(t)$ не убывает по t при $t \geqslant 0$. Имеем, обозначая $z_1 = \lambda_1 t$ и $z_2 = \lambda_2 t$,

$$\left(\frac{\widehat{f}_1(t)}{\widehat{f}_2(t)}\right)_t' = \left(\frac{\lambda_1 \widehat{h}(z_1)}{\lambda_2 \widehat{h}(z_2)}\right)_t' = \frac{\lambda_1}{\lambda_2} \cdot \frac{\lambda_1 \widehat{h}'(z_1)\widehat{h}(z_2) - \lambda_2 \widehat{h}'(z_2)\widehat{h}(z_1)}{\widehat{h}^2(z_2)}.$$

Таким образом, остаётся доказать, что $\lambda_1 \frac{\widehat{h}'(z_1)}{\widehat{h}(z_1)} - \lambda_2 \frac{\widehat{h}'(z_2)}{\widehat{h}(z_2)} \geqslant 0$, т.е. что

$$\lambda_1 \left[\ln \widehat{h}(z)\right]_z' \Big|_{z=\lambda_1 t} - \lambda_2 \left[\ln \widehat{h}(z)\right]_z' \Big|_{z=\lambda_2 t} \geqslant 0.$$

Поскольку функция $\widehat{h}(z)$ лог-вогнута и $\widehat{h}'(z) \leqslant 0$ при $z \geqslant 0$, то

$$\left[\ln \widehat{h}(z)\right]'_z\Big|_{z=\lambda_2 t} \leqslant \left[\ln \widehat{h}(z)\right]'_z\Big|_{z=\lambda_1 t} \leqslant 0;$$

и учитывая, что $\lambda_1 \leqslant \lambda_2$, отсюда получаем требуемое неравенство. □

Теорема 3. *Пусть* $X_i^{(j)}$, $i = 1,\ldots,n$, $j = 1,2$, – *независимые с.в. с плотностями* $\lambda_i h(\lambda_i(t-\theta_j))$, *где* $h(t)$ – *дифференцируемая лог-вогнутая функция,* $\theta_1 < \theta_2$, $0 < \lambda_1 \leqslant \lambda_2 \leqslant \cdots \leqslant \lambda_n$. *Пусть далее* $p \succ p'$, $p_i, p_i' \geqslant 0$, $\sum\limits_{i=1}^n p_i = \sum\limits_{i=1}^n p_i' = 1$. *Тогда*

$$\mathbf{P}\left[\sum_{i=1}^n p_i' X_i^{(1)} < \sum_{i=1}^n p_i' X_i^{(2)}\right] \geqslant \mathbf{P}\left[\sum_{i=1}^n p_i X_i^{(1)} < \sum_{i=1}^n p_i X_i^{(2)}\right].$$

Доказательство аналогично доказательству теоремы 1, если учесть, что в нашем случае при каждом $i = 1,\ldots,n$ $\overset{o}{X}{}_i^{(1)}$ и $\overset{o}{X}{}_i^{(2)}$ одинаково распределены с плотностью $\lambda_i h(\lambda_i x)$, а в конце воспользоваться леммами 1 и 2 и теоремой 2 вместо теоремы 2.3 в [4]. □

Из теоремы 3 получим следствие 2 таким же образом, как из теоремы 1 получено следствие 1.

Следствие 2. *Пусть* $X_i^{(j)}$, $i = 1,2,\ldots$, $j = 1,2$, – *независимые с.в. с плотностями* $\lambda_i h(\lambda_i(t-\theta_j))$, *где* $h(t)$ – *дифференцируемая лог-вогнутая функция,* $\theta_1 < \theta_2$, $0 < \lambda_1 \leqslant \lambda_2 \leqslant \cdots \leqslant \lambda_n$. *Тогда* $\mathbf{P}\left[Z_n^{(1)} < Z_n^{(2)}\right]$ *не убывает по* n.

Вопрос о стремлении к 0 ошибок первого и второго рода решается как в конце пункта 2.1, только вместо теоремы Хинчина нужно применить теорему Чебышёва [11, 2.2., с. 3] с учётом того, что в нашем случае

$$\lim_{n\to\infty} \frac{1}{n^2}\sum_{i=1}^n \sigma_i^2 \leqslant \lim_{n\to\infty} \frac{\sigma_1^2}{n} = 0,$$

поскольку среднеквадратическое отклонение σ, существование которого обеспечивается лог-вогнутостью плотности, обратно пропорционально λ.

Пример 1. Рассмотрим ситуацию, когда в отношении двух небесных тел одного и того же типа возникает вопрос, у какого из них больше масса. Воспользуемся тем, что для таких объектов справедливо соотношение (эмпирический закон) "масса–светимость", которое означает, что существует достаточно жёсткая связь между этими величинами. Поскольку эта зависимость монотонная, то ошибка вывода о том, что масса одного объекта меньше массы другого объекта, равна ошибке соответствующего вывода относительно параметров их светимости θ_1 и θ_2.

Будем считать, что на полезные сигналы, принимаемые обсерваторией от этих объектов, накладывается помеха в виде нормально распределённой с.в. $N(0,\sigma)$ с плотностью

$$f_\sigma(x) = \frac{1}{\sigma\sqrt{2\pi}}\, e^{-x^2/(2\sigma^2)}.$$

Если в силу каких-либо технических или астрономических причин отсутствует возможность провести одномоментную серию наблюдений над случайными величинами $X^{(1)}$ и $X^{(2)}$, представляющими собой значения θ_1 и θ_2 с вышеуказанной аддитивной помехой, то их приходится проводить через некоторые промежутки времени. Таким образом, усреднение результатов наблюдений проходит в режиме дискретного процесса с моментами наблюдений $\tau_1, \tau_2, \dots$, $\tau_1 \leqslant \tau_2 \leqslant \dots$.

Предположим, что параллельно с этим процессом осуществляется непрерывное совершенствование процедуры подавления помехи, что приводит к уменьшению её среднеквадратического отклонения σ при сохранении вида закона, то есть $\sigma = \sigma(\tau)$ – невозрастающая функция. Тогда, согласно следствию 2, если для любого фиксированного n решение принимается на основании выборочных средних, то при

$$\sum_{i=1}^{n} X^{(1)}(\tau_i) < (>) \sum_{i=1}^{n} X^{(2)}(\tau_i)$$

следует сделать вывод, что $\theta_1 < (>)\ \theta_2$, поскольку $\sigma(\tau_1) \leqslant \sigma(\tau_2) \leqslant \dots \leqslant \sigma(\tau_n)$, а плотность $f_\sigma(x)$ – лог-вогнутая функция.

Далее, на основании этого решения, в силу закона "масса–светимость" делается соответствующий вывод о том, что масса первого небесного тела меньше (больше) массы второго небесного тела. При этом (равные) ошибки первого и второго рода монотонно стремятся к 0 при $n \to \infty$.

§3. Селекция популяций относительно контроля

Рассмотрим k независимых популяций $\pi_1, \pi_2, \dots, \pi_k$. Качество i-й популяции характеризуется вещественным параметром θ_i. Будем говорить, что i-я популяция позитивная (или хорошая), если $\theta_i > \theta_0$, и негативная (или плохая), если $\theta_i < \theta_0$, где θ_0 – заданное число.

Как и в [8], мы ограничимся распределениями с параметром сдвига. Пусть $X_1, \dots, X_k$ суть k наблюдений над независимыми с.в. с плотностями $h_1(x-\theta_1), \dots, h_k(x-\theta_k)$, соответственно. В качестве наблюдений $X_1, \dots, X_k$ могут фигурировать статистики (прежде всего достаточные статистики, если они существуют), представляющие каждую из k популяций.

Будем рассматривать общий статистический критерий в виде $\underline{D} = (D_1, D_2, \dots, D_k)$, где $D_i = D_i(x)$ – критерий (правило), сопоставляющий выборочному значению $X_i = x$ одно из двух решений: популяция π_i хорошая или популяция π_i плохая. При этом если решение принято правильное, то потери $L^{(i)}$ равны нулю, а если ошибочное, то они равны $L(|\theta_0 - \theta_i|)$, где $L(\tau)$ (> 0) не убывает при $\tau > 0$. Будем считать, что потери L от принятого общего решения равны сумме потерь от принятых решений по каждой популяции, то есть $L = \sum\limits_i L^{(i)}$. Тогда, полагая $\underline{\theta} = (\theta_0, \theta_1, \dots, \theta_k)$, $\underline{X} = (X_1, \dots, X_k)$, получим полный риск

$$\begin{aligned}
R(\underline{\theta}; \underline{D}) &= \mathbf{E}\, L(\underline{\theta}; \underline{D}(\underline{X})) \\
&= \int\limits_{-\infty}^{\infty} \int\limits_{-\infty}^{\infty} \dots \int\limits_{-\infty}^{\infty} \sum_{i=1}^{k} L^{(i)}(\theta_0, \theta_i; D_i(x_i)) \prod_{i=1}^{k} [h_i(x_i - \theta_i)\, dx_i] \\
&= \sum_{i=1}^{k} \int\limits_{-\infty}^{\infty} L^{(i)}(\theta_0, \theta_i; D_i(x))\, h_i(x - \theta_i)\, dx \\
&= \sum_{i=1}^{k} \mathbf{E}\, L^{(i)}(\theta_0, \theta_i; D_i(X_i)) = \sum_{i=1}^{k} R^{(i)}(\theta_0, \theta_i; D_i),
\end{aligned}$$

где $R^{(i)}$ – риск по i-й популяции.

Заметим, что

$$R^{(i)}(\theta_0, \theta_i; D_i) = \int\limits_{-\infty}^{\infty} L^{(i)}(\theta_0, \theta_i; D_i(x))\, h_i(x - \theta_i)\, dx, \tag{1}$$

и предположим, что для любого $i = 1, \ldots, k$ выполняется условие $h_i(t) = h_i(-t)$ для всех t, причём $h_i(t)$ – унимодальная функция.

В разделах 2 и 4 инвариантность возникает относительно группы перестановок двух координат точек выборочного пространства; в настоящем разделе в этой же роли выступает группа $G = G_1 \times G_2 \times \cdots \times G_k$, являющаяся прямым произведением групп преобразований $G_i = \{g_1, g_2\}: \ g_1 X_i = X_i$ и $g_2 X_i = 2\theta_0 - X_i, \ i = 1, \ldots, k$.

Возьмём произвольный элемент группы $G: \ g = (g_{j_1}, g_{j_2}, \ldots, g_{j_k})$, $j_i = 1$ или 2. Заметим, прежде всего, что в силу независимости X_i и симметричности $h_i(t)$ имеем

$$\prod_{i=1}^{k} h_i(g_{j_i} X_i - \overline{g}_{j_i}\theta_i) = \prod_{i=1}^{k} h_i(X_i - \theta_i).$$

Далее, чтобы мы могли рассчитывать на наличие решающих правил, инвариантных относительно группы G, должно быть $L(\overline{g}\,\underline{\theta}; g'\underline{D}(\underline{X})) = L(\underline{\theta}; \underline{D}(\underline{X}))$, т.е.

$$\sum_{i=1}^{k} L^{(i)}(\theta_0, \overline{g}_{j_i}\theta_i; g'_{j_i} D_i(X_i)) = \sum_{i=1}^{k} L^{(i)}(\theta_0, \theta_i; D_i(X_i)).$$

Покажем, что в нашей ситуации это условие выполняется. Действительно, обозначая решение отличное от D_i через $\overline{D}_i$, для $j_i = 2$ получаем

$$L^{(i)}(\theta_0, 2\theta_0 - \theta_i; \overline{D}_i(X_i)) = L^{(i)}(\theta_0, \theta_i; D_i(X_i)),$$

поскольку решение $\overline{D}_i(X_i)$ в левой части и решение $D_i(X_i)$ в правой части либо оба корректны, либо оба некорректны. Наконец, для решений должно выполняться соотношение

$$(g'_{j_1} D_1(X_1), \ldots, g'_{j_k} D_k(X_k)) = (D_1(g_{j_1} X_1), \ldots, D_k(g_{j_k} X_k)). \qquad (2)$$

Введём статистический критерий $\underline{D}^*$ с компонентами D_i^*: если $X_i < \theta_0$, то популяция признаётся плохой, если $X_i > \theta_0$, то популяция признаётся хорошей. Тогда i-я составляющая полной функции потерь, соответствующая критерию D_i^*, равна

$$L^{(i)}(\theta_0, \theta_i; D_i^*(X_i)) = \begin{cases} L(|\theta_0 - \theta_i|)\,, & (X_i - \theta_0)(\theta_0 - \theta_i) > 0 \\ 0, & \text{иначе}. \end{cases} \qquad (3)$$

$\underline{D}^*$ – инвариантный статистический критерий, поскольку (2) выполняется в силу того, что для $j_i = 2$ очевидно имеем $\overline{D}_i^*(X_i) = D_i^*(2\theta_0 - X_i)$.

Теорема 4. *Если плотности $h_i(t)$ являются чётными и унимодальными функциями, то критерий $\underline{D}^* = (D_1^*, D_2^*, \ldots, D_k^*)$ оказывается минимаксным и минимизирует риск $R(\underline{\theta}; \underline{D})$ среди инвариантных статистических критериев $\underline{D}$.*

Доказательство. Поскольку в нашем случае средний риск есть

$$r(\underline{\theta}, \underline{D}) = \frac{1}{2k} \sum_{i=1}^{k} \sum_{j=1}^{2} R^{(i)}(\theta_0, \overline{g}_j \theta_i; D_i)$$
$$= \frac{1}{2k} \sum_{i=1}^{k} \int_{-\infty}^{\infty} \sum_{j=1}^{2} L^{(i)}(\theta_0, \overline{g}_j \theta_i; D_i(x))\, h_i(x - \overline{g}_j \theta_i)\, dx,$$

то из леммы 1 и примера 1 в [6] следует, что достаточно установить для любых i и x неравенство

$$L^{(i)}(\theta_0, \theta_i; D_i^*(x))\, h_i(x - \theta_i) + L^{(i)}(\theta_0, 2\,\theta_0 - \theta_i; D_i^*(x))\, h_i(x - (2\,\theta_0 - \theta_i))$$
$$\leqslant L^{(i)}(\theta_0, \theta_i; D_i(x))\, h_i(x - \theta_i) + L^{(i)}(\theta_0, 2\,\theta_0 - \theta_i; D_i(x)\, h_i(x - (2\,\theta_0 - \theta_i)),$$

где D_i – произвольный критерий. Без ограничения общности, полагаем $(x - \theta_0)(\theta_0 - \theta_i) < 0$. Если решения $D_i^*(x)$ и $D_i(x)$ совпадают, то мы, очевидно, имеем равенство. В противном случае, первое и последнее слагаемые равны нулю, а $L^{(i)}(\theta_0, 2\theta_0 - \theta_i; D_i^*(x)) = L^{(i)}(\theta_0, \theta_i; D_i(x)) = L(|\theta_0 - \theta_i|)$, и остаётся убедиться, что для чётной унимодальной функции $h_i(t)$ имеем $h_i(x - (2\theta_0 - \theta_i)) \leqslant h_i(x - \theta_i)$ при $(x - \theta_0)(\theta_0 - \theta_i) < 0$. $\square$

Замечание. Утверждение теоремы справедливо и для *рандомизированных* процедур принятия решения. К этому приводят следующие рассуждения общего характера.

Действительно, если независимые с.в. X_i, $i = 1, \ldots, k$, имеют плотности $f_i(x_i, \theta_i)$, то для произвольной (рандомизированной) решающей функции $\psi(\underline{x})$, где $\underline{x} = (x_1, \ldots, x_k)$, риск равен

$$\rho(\psi, \underline{\theta}) = \int \sum_{\underline{D}} \psi_{\underline{D}}(\underline{x})\, L(\underline{\theta}; \underline{D}) \prod_{i=1}^{k} f_i(x_i, \theta_i)\, dx_1 \ldots dx_k,$$

где $\psi_{\underline{D}}(\underline{x})$ – вероятность принятия решения $\underline{D}$, когда значение наблюдения над $\underline{X}$ равно $\underline{x}$. Пусть решающая функция $\psi^0(\underline{x})$ такова, что $\psi^0_{\underline{D}}(\underline{x}) = 1$, если $\underline{D} = \underline{D}^*(\underline{x})$, и, соответственно, $\psi^0_{\underline{D}}(\underline{x}) = 0$, если

$\underline{D} \neq \underline{D}^*(\underline{x})$. Очевидно, что в нашем случае ψ^0 – инвариантная решающая функция, и значит, согласно [6], остаётся установить неравенство

$$\sum_g \rho(\psi^0, \overline{g}\underline{\theta}) \leqslant \sum_g \rho(\psi, \overline{g}\underline{\theta}). \tag{4}$$

Поскольку нами доказано, что для любых $\underline{D}, \underline{x}$ и $\underline{\theta}$ $\Phi_{\underline{D}^*}(\underline{x}, \underline{\theta}) \leqslant \Phi_{\underline{D}}(\underline{x}, \underline{\theta})$, где

$$\Phi_{\underline{D}}(\underline{x}, \underline{\theta}) = \sum_g L(\overline{g}\underline{\theta}; \underline{D}) \prod_{i=1}^{k} f_i(x_i, \overline{g}_{j_i}\theta_i),$$

то отсюда легко сделать заключение, что

$$\int \Phi_{\underline{D}^*}(\underline{x}, \underline{\theta})\, dx_1 \ldots dx_k \leqslant \int \sum_{\underline{D}} \psi_{\underline{D}}(\underline{x})\, \Phi_{\underline{D}}(\underline{x}, \underline{\theta})\, dx_1 \ldots dx_k.$$

А от этого неравенства прямо приходим к (4) за счёт перемены порядка суммирований и интегрирования. □

Заметим теперь, что из (1) и (3) следует

$$R^{(i)}(\theta_0, \theta_i; D_i^*) = L(|\theta_0 - \theta_i|)\, \mathbf{P}\, [(X_i - \theta_0)(\theta_0 - \theta_i) > 0].$$

Это, в частности, означает, что в случае $L(\tau) \equiv 1$ полная функция потерь L есть сумма “0–1”-функций потерь $L^{(i)}$ и, следовательно, полный риск

$$R(\underline{\theta}; \underline{D}^*) = N_1 + N_2,$$

где N_1 (N_2) – среднее число хороших (плохих) популяций, признанных плохими (хорошими). Если вместо X_i, $i = 1, \ldots, k$, оперировать выборочными средними размера n_i для популяции π_i, то принимая во внимание, что лог-вогнутые функции являются унимодальными, из следствия 2.4 в [4] вытекает

Теорема 5. *Если плотности $h_i(t)$ являются чётными и лог-вогнутыми функциями, то среднее число $(N_1 + N_2)$ популяций, в отношении которых по критерию $\underline{D}^*$ принято ошибочное решение, не возрастает по каждому n_i.*

Замечание. Рассуждая, как в конце п. 2.1, получим, что если одновременно все $n_i \to \infty$, то величина $N_1 + N_2$ монотонно стремится к 0.

§4. Перестановки Лоренцева типа с минимальным риском

Пусть функция $\varphi(r,s)$ имеет непрерывные вторые частные производные. Тогда из соотношения

$$\int\limits_a^b \int\limits_c^d \varphi''_{rs}(r,s)\,ds\,dr = \varphi(b,d) - \varphi(b,c) - \varphi(a,d) + \varphi(a,c) \qquad (5)$$

следует, что условие

$$\varphi''_{rs}(r,s) \geqslant 0 \qquad (6)$$

является необходимым и достаточным для неотрицательности правой части (5) при любых $a < b$ и $c < d$. Этот результат Лоренца [12] применим к функции $\varphi(r,s) = r^s$. Имеем

$$\varphi''_{rs}(r,s) = r^{s-1}\,(s \ln r + 1). \qquad (7)$$

Правая часть (7) положительна при $r \geqslant 1$ и $s \geqslant 1$, что позволяет нам оптимизировать одну практическую ситуацию.

Пример 2. Рассмотрим двухканальную электрическую цепь, содержащую разное (фиксированное) число усилителей A в каналах: $m1$ и $m2, m1 > m2$, и представленную на рис. 1, где S обозначает сумматор.

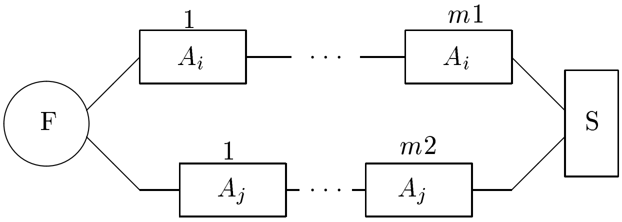

Рис. 1. Двухканальная электрическая цепь

Различие между $m1$ и $m2$ можно объяснить тем, что второй канал попутно выполняет роль резервного на случай обрыва в первом канале. В нашем распоряжении имеется по $m1$ усилителей двух типов с (неизвестными) средними значениями коэффициентов усиления θ_i и θ_j, $i,j = 1,2$ $(i \neq j)$. Нужно $m1 + m2$ из них разместить в каналах с целью максимизации математического ожидания увеличения входного сигнала F, которое равно $\theta_i^{m1} + \theta_j^{m2}$ в предположении, что в каждом канале могут быть усилители только одного типа. Этому допущению отвечает физическая или техническая несовместимость усилителей разных типов.

Имеем

$$\theta_i^{m1} + \theta_j^{m2} \geqslant \theta_i^{m2} + \theta_j^{m1} \quad \text{при} \ \ \theta_i > \theta_j \geqslant 1,$$

поскольку в этом случае выражение (7) удовлетворяет условию (6). Мы хотим добиться максимума (целевой) функции $\theta_i^{m1}+\theta_j^{m2}$ в рамках вальдовской теории принятия решений. Пусть функция потерь имеет вид

$$L_{i,j}(\theta_1,\theta_2) = h_{\{\theta_1,\theta_2\}}\left(\theta_i^{m1} + \theta_j^{m2}\right),$$

где $h_{\{\theta_1,\theta_2\}}(z)$ как функция от z неотрицательная и невозрастающая. (Для того, чтобы соблюсти единообразие по всей статье, потребуем, в отличие от [1,13], чтобы потери при корректном решении непременно равнялись нулю).

Предположим, что коэффициенты усиления устройств, входящих в цепь на рис. 1, независимы и имеют распределение Лапласа с плотностью

$$f_\theta(x) = \frac{1}{2}\, e^{-|x-\theta|}, \quad -\infty < x < \infty, \qquad 1 \leqslant \theta < \infty,$$

которая, очевидно, обладает M.L.R.-свойством. Тогда на основании [5,6] мы должны поставить в первый канал устройства того типа, который при тестировании дал более высокую сумму n реализаций коэффициента усиления. Такой статистический критерий является минимаксным и в то же время при каждом (θ_1,θ_2) имеет наименьший риск среди всех инвариантных статистических критериев. По следствию 1 настоящей работы стремление риска к 0 будет проходить монотонно по n.

Литература

1. M. I. Revyakov, *Schur-convexity of* 2*nd order, certain subclass of multivariate arrangement increasing functions with applications in statistics.* — J. Multivar. Anal. **116** (2013), 25–34.
2. E. L. Lehmann, *Testing Statistical Hypotheses*, Wiley, 1986.
3. C. Ma, *On peakedness of distributions of convex combinations.* — J. Statist. Plann. Inference **70** (1998), 51–56.
4. F. Proschan, *Peakedness of distributions of convex combinations.* — Ann. Math. Statist. **36** (1965), 1703–1706.

5. M. L. Eaton, *Some optimum properties of ranking procedures.* — Ann. Math. Statist. **38** (1967), 124–137.
6. E. L. Lehmann, *On a theorem of Bahadur and Goodman.* — Ann. Math. Statist. **37** (1966), 1–6.
7. M. A. Girshick, *Contributions to the Theory of sequential analysis. I.* — Ann. Math. Statist. **17** (1946), 123–143.
8. R. H. Randles, M. Hollander, *Γ-minimax selection procedures in treatments versus control problems.* — Ann. Math. Statist. **42** (1971), 330–341.
9. A. W. Marshall, I. Olkin, *Inequalities: Theory of Majorization and Its Applications*, Academic Press, 1979.
10. M. Y. An, *Logconcavity versus logcovexity: a complete characterization.* — J. Economic Theory **80** (1998), 350–369.
11. C. R. Rao, *Linear Statistical Inference and its Applications*, Wiley, 1965.
12. G. G. Lorentz, *An inequality for rearrangements.* — Amer. Math. Monthly **60** (1953), 176–179.
13. M. I. Revyakov, *Optimal substitution of arguments of an arrangement increasing function based on sufficient statistics for parameter-arguments.* — J. Math. Sci. **176**, No. 2 (2011), 209–223.

Revyakov M. I. Ranking and selection of populations based on sample means.

A number of directions is indicated in which, for statistical problems of decision making related to ordering the parameters of distributions, it is expedient to lean on comparison of sample means. It is assumed that the corresponding parametric family has no nontrivial sufficient statistics. The key role is played by establishing conditions under which the reliability of inferences increases monotonically with the size of the sampling. Examples of applications are given.

С.-Петербург, Россия
E-mail: revyakov.m@gmail.com

Поступило 10 ноября 2016 г.