

О КОРРЕЛЯЦИИ ПЛОТНОСТИ ИСТИННЫХ ПРЕДСКАЗАНИЙ В ЧИСЛОВЫХ ЛОТЕРЕЯХ С СОЛНЕЧНОЙ АКТИВНОСТЬЮ И ТЕСТОМ ПИККАРДИ*

В 1960–1980 гг. астрофизиком Н.А. Козыревым (ГАО АН СССР) было высказано предположение о том, что процессы, в которых существенно меняется организация, происходят по-разному (ускоряются или тормозятся), если рядом с рассматриваемым процессом осуществляется другой, в котором также существенно изменяется энтропия [1]. Это предположение было автором формализовано в терминах энтропийных потенциалов.

Определение. Энтропийный потенциал $Z(A)$ события A , происходящего в момент T_0 в системе R , есть разность математических ожиданий энтропии системы R , вычисленных после того, как событие A произошло (т.е. в момент $T_0 + dT$), и до того как оно произошло (т.е. в момент $T_0 - dT$):

$$Z(A) = \hat{S}_T(T_0 + dT) - \hat{S}_T(T_0 - dT)$$

Уточним, что математическое ожидание \hat{S}_T вычисляется для некоторого момента $T > T_0$, тем самым $\hat{S}_T(T_0 + dT)$ – оценка энтропии системы R на момент T , сделанная в момент $T_0 + dT$ (см. рис. 1).

При этом T – конечно; что происходит с $Z(A)$ при $T \rightarrow \infty$ вопрос особый и в данной статье не обсуждается. Более подробное определение энтропийного потенциала и пример его использования (см. в Приложении А).

Поскольку в реальном мире практически для любого события A может осуществиться и событие \bar{A} («не A »), с вероятностью $P(\bar{A}) \neq 0$, то обычно $Z(A) \neq 0$. Понятно, что тем самым $Z(A)$ говорит о влиянии события A на развитие или угасание системы R . Если до того, как событие A произошло (т.е. пока неизвестно произойдет событие A или \bar{A}) математическое ожидание энтропии $\hat{S}_T(T_0 - dT)$, а в результате события A математическое ожидание оказывается ниже: $\hat{S}_T(T_0 + dT) < \hat{S}_T(T_0 - dT)$, то по сути это означает, что событие A препятствует росту энтропии в системе R , защищает ее от деградации и развала. Иными словами, энтропийный потенциал событий «полезных» для системы R отрицателен, а «вредных» – «положителен». (Замечание. Если под S понимать только «информационную» энтропию, то $Z(A)$ становится мерой ценности сообщения A , поскольку $Z(A)$ – показатель вклада сообщения в развитие или деградацию системы R , принявшей сообщения. В самом деле, ведь «ценность» сообщения A для системы R проявляется именно во вкладе A в дальнейшее развитие или угасание системы R , а мерой этого вклада как раз и является энтропийный потенциал $Z(A)$).

В чем эвристическая ценность энтропийного потенциала события?

Бытовые формулировки «везет – не везет», «удача – неудача» легко и

* Печатается по изданию: Зильберман М.Ш. О корреляции плотности истинных предсказаний в числовых лотереях с солнечной активностью и тестом Пиккарди. – Л., 1989. – Деп. в ВИНТИ 22.12.89, № 3168-В89.

естественно формализуются в терминах энтропийных потенциалов, поскольку «везение», как правило, ведет к упорядочению или сохранению структуры: производственной («лишь чудом избежали аварии...»), межличностной («удачная встреча, знакомство»), в экстремальных случаях даже биологической («спасение», пройди пуля чуть ниже...» и т.д.). Бытовое же «невезение», «неудача», как правило, такую структуру разрушает. Поэтому энтропийный потенциал, будучи мерой потенции события A изменит в будущем энтропию системы R , хорошо подходит для формализации терминов типа «везение» и т.д. Однако важнее следующее.

Гипотеза (*). «Процессы с энтропийными потенциалами одного знака способствуют протеканию друг друга».

Нетривиальность этой гипотезы заставляет чрезвычайно тщательно выбирать материал для ее проверки. Понятно, что материал такого рода должен быть, с одной стороны, статистически весьма надежным, а с другой, полностью чистым от различных неучтенных физических, социальных и прочих воздействий. Это не могут быть различные дорожно-транспортные происшествия, эпидемии, несчастные случаи и т.д., поскольку все эти процессы сильно коррелируют с солнечной активностью [2], [3]. Это не могут быть и чисто стохастические явления (типа отклонения употребимости данной буквы на данной газетной странице), т.к. к ним неприменима высказанная гипотеза.

В результате для проверки гипотезы был выбран материал массовых числовых лотерей. В самом деле такой материал, по сути, представляет собой длинный ряд однородных наблюдений над социальными следствиями, порождаемыми физическим датчиком случайных чисел. Этот материал весьма чист от тривиальных физических воздействий и вместе с тем статистически чрезвычайно надежен. Согласно гипотезе, плотность истинных предсказаний в лотереях должна возрастать весной и в первой половине лета (в природе идет активный биосинтез, передается генетическая информация) и падать в августе, сентябре (биосинтез угнетается) и в конце зимы (метеофакторы не дают начаться биосинтезу, растения находятся в состоянии вынужденного покоя). Поясним высказанное утверждение. Плотность истинных предсказаний – отношение числа верно заполненных (т.е. выигравших) комбинаций к общему числу заполненных комбинаций. Чем она больше, тем больше число выигрышей, тем меньше число отрицательных стрессов, связанных с неоправдавшимися надеждами. Поэтому, если гипотеза (*) верна, то в периоды, когда в природе осуществляются процессы с большим отрицательным потенциалом Z (прекратись биосинтез или образование плодов, через 100 лет планета будет голой), число отрицательных стрессов должно падать, т.е. плотность истинных предсказаний должна возрастать. Каким образом? Например, посредством роста частоты выбрасывания шаров с теми номерами, которые чаще заполняют участники. Впрочем, это совсем другой вопрос.

Исследовались результаты двух советских лотерей («5 из 36» и «6 из 49») и французской числовой лотереи «ЛОТО» на протяжении 8 лет – с 1980 по 1987 гг. Было рассмотрено 938 тиражей советских и 517 тиражей французской (из-за пропусков в публикациях и отсутствия материала в библиотеках автор располагал 84 % материала всех прошедших в «ЛОТО» тиражей, для лотерей СССР – 100 %). Плотность истинных предсказаний в тираже i вычислялось как $\rho_i = N_{3i} / N_i \times 2000$, где N_{3i} – число комбинаций с верно предсказанными 3 цифрами, N_i – общее число заполненных комбинаций, множитель 2000 введен для удобства записи. Величины N_{3i} (число верно предсказанных «троек») на 1,5 порядка больше, чем N_{4i} (число верно предсказанных «четверок») и более чем на 3 порядка превосходят числа N_{5i} . Поэтому в статье для статистической надежности использовались именно числа N_{3i} . Но т.к. величины N_{3i} , N_{4i} , N_{5i} не независимы и сильно коррелируют, то найденные эффекты проявляются и в них.

В дальнейшем плотности ρ («5 / 36») обозначаются ρ_3^5 , т.е. плотность верно предсказанных 3 цифр при заполненных 5 (по условиям лотереи «5 из 36» на поле из 36 чисел требуется отметить 5, которые сравниваются с 5, выпавшими в тираже. Выигравшими считаются билеты с верно отмеченными 3, 4, 5 цифрами). Аналогично в лотерее «6 / 49» плотности ρ («6 / 49») обозначаются ρ_3^6 . Во французской лотерее «ЛОТО» (по правилам аналогичной «6/49») – $\rho(F)$. В принятом масштабе математическое ожидание $\hat{\rho}_3^5 \approx 24,669$, $\hat{\rho}_3^6 \approx 35,301$. (0 величинах $\rho(F)$ французской лотереи «ЛОТО» см. в п. 7–2). В 1986г. лотерея «6/49» была преобразована в «6/45». Поэтому для однородности величины ρ_3^6 1986–1987 гг. домножались на отношение математических ожиданий:

$$\hat{\rho}_3^6 (\text{«6/49»}) / \hat{\rho}_3^6 (\text{«6/45»}) \approx 0,78654$$

Числа комбинаций, заполняемых в данном тираже в среднем по 1980–1987 гг. составляли N («ЛОТО») > 12 млн, N («5/36») ~ 8 млн, N («6/49») ~ 1,4 млн.

Сформулируем «нулевую» гипотезу: все отклонения выборочных средних являются чисто случайными.

В отличие от нее, гипотеза (х) приводит к наличию сезонного хода плотности ρ (по причинам, изложенным выше), возможному наличию многолетнего хода ρ (который может быть связан, например, с активными и менее активными диссипативными процессами на Солнце) и коррелированности хода плотности истинных предсказаний в различных странах (как порождаемых одними факторами) и т.д. В «нулевой» гипотезе ни значимого хода плотности, ни тем более коррелированности этого хода наблюдаться не должно.

Разобьем тиражи каждого года $J = 1980, \dots, 1987$ случайным образом на

2 выборки равного объема. Если нет факторов, систематически смещающих среднегодовые плотности, то при этом для каждого года образуется пара независимых, непересекающихся, одинаково распределенных выборок $\{\rho_i\}_J^I$, $\{\rho_i\}_J^{II}$, чьи выборочные средние положительно коррелировать не должны. Если же существуют «удачные» и «неудачные» годы (см. выше), то образующиеся пары выборок зависимы от внешних факторов и выборочные средние $\bar{\rho}_J^I$, $\bar{\rho}_J^{II}$ должны коррелировать положительно.

Разбиение на случайные выборки производилось датчиком случайных чисел RND(O) компьютера «ATARI 130 XE». Для надежности получаемых выводов разбиение и проверка коррелированности получаемых пар выборочных средних производилась неоднократно (10 раз). Результаты приведены в таблице 1.

Видно, что *все* 10 коэффициентов корреляции как для «Loto France», так и для «5/36» положительны, причем надежность коррелированности для 2 коэффициентов «Loto France» и 5 коэффициентов «5/36» превосходит 95%-ный уровень (для 2 коэффициентов «5/36» даже 99%-ный уровень). Разумеется, т.к. данные получены на едином материале (хотя и независимыми случайными извлечениями), то нельзя использовать биномиальное распределение для оценки вероятности, случайно получить в обеих лотереях все 10 коэффициентов положительными, однако факт систематического смещения среднегодовых плотностей, по-видимому, можно считать доказанным. Заметим при этом, что на материале «6/49» установленные закономерности не проявляются (см. подробнее п. 4).

Для контроля полученных результатов та же операция на том же компьютере производилась над «перетасованным» массивом $\{\rho_i\}^{RND}$, который был сформирован из реального массива $\{\rho_i\}$ с помощью 10 000 случайных обменов местами величин ρ_i . Полученный в результате массив $\{\rho_i\}^{RND}$ обладал теми же распределением и средними параметрами, что и $\{\rho_i\}$, и отличался лишь «замытостью» реального порядка расположения ρ_i . Результаты приведены в таблице 1. Видно, что положительная корреляция выборочных средних исчезла. Это говорит как о существенности номера i , характеризующем положение плотности ρ_i в общем массиве данных $\{\rho_i\}$, так и о наличии факторов, систематически смещающих среднегодовые плотности.

1. Сезонный ход плотности истинных предсказаний приведен на рисунке 2. График советской лотереи $\rho(\Sigma)$ получен по объединенному материалу «5/36» и «6/49» (ρ («6/49») домножены на масштабный коэффициент, соответствующий отношению математических ожиданий). Видно сильное сходство полученного хода с графиком плотности истинных

предсказаний во Франции (коэффициент корреляции +61 %). Практически единственное отличие состоит в месячном опережении весеннего пика во Франции, но это, скорее, не противоречит, а подтверждает гипотезу (х), т.к. во Франции раньше наступает весна, и процессы биосинтеза активизируются раньше. Значимо также непостоянство среднемесячной плотности истинных предсказаний, которого не должно быть в «нулевой» гипотезе. Сходен и многолетний ход ρ в лотереях СССР и Франции: коэффициент корреляции среднегодовых значений плотности истинных предсказаний $\rho(\Sigma)$ и $\rho(F)$ составляет + 86,6 % с надежностью коррелированности 99,7 % [4] (см. рис. 3).

2. Соотнесение сезонного хода ρ с тестом Пиккарди. Тест Пиккарди [5] – химическая реакция гидролиза $BiCl_3$, проводимая ежедневно в стандартных условиях на протяжении многих лет. Пиккарди [5] и более поздними исследованиями [6] был показан сезонный ход ее протекания, коррелированность с солнечной активностью. Объяснение этому обычно ищут в области солнечно-земных связей, осуществляемых электромагнитным механизмом. Однако на механическую стохастическую систему – лотерейный барабан – электромагнитный механизм не воздействует, а сезонный ход плотности в советской лотерее «5/36» и «Loto France» близок к сезонному ходу теста Пиккарди (коэффициенты корреляции составляют – 62 % для советской лотереи «5/36» и – 45 % для «Loto France») (см. рис. 2).

3. Близость графиков сезонного хода плотности истинных предсказаний и величин T_p теста Пиккарди, при доказанном влиянии солнечной активности (С.А.) на ход теста [5], приводит к мысли о возможном влиянии С.А. на плотность истинных предсказаний. Поскольку увеличение С.А. приводит к росту числа дорожно-транспортных происшествий [2], числа заболеваний [7] и т.д., т.е. приводит к росту энтропии (правда обычно утверждается достаточность электромагнитного механизма для объяснения всех этих явлений), то логично ожидать, что в годы роста С.А. среднегодовые плотности истинных предсказаний будут ниже, а в годы спада – выше. Это в самом деле имеет место, причем для всех трех типов рассмотренных лотерей. Коэффициенты корреляции среднегодовых чисел Вольфа [8] и среднегодовых плотностей в «Loto France», «5/36», «6/49» составляют: -78,0, -67,4, -58,8 % соответственно (надежность коррелированности по одностороннему критерию 99, 96,5 и 94 % соответственно [4]) (см. рис. 3).

4. Установленные эффекты ранжируются в порядке числа участников лотерей (табл. 2).

К этим эффектам можно отнести:

- Выраженность максимума графика сезонного хода ρ тах и возможность получить установленное превышение в рамках «нулевой» гипотезы.

- Вариабельность сезонного хода: чем сильнее установленные эффекты, т.е. чем сильнее влияние процессов в природе на ход плотности ρ , тем выше вариабельность.
- Коррелированность с числами Вольфа (показателем солнечной активности).
- Корреляция с тестом Пиккарди.
- Пусть $\{\rho_i\}^+$ множество тех ρ_i , каждое из которых превосходит математическое ожидание $\bar{\rho}$. Аналогично сформируем множество $\{\rho_i\}^-$ ($\rho_i^- < \bar{\rho}$). Обозначим: N^+ – число ρ_i^+ в некотором фиксированном интервале времени (например, 1 месяц), а $\bar{\rho}^+$ – средняя плотность по ρ_i^+ из этого интервала. Тогда, с точки зрения «нулевой» гипотезы величины $(\bar{\rho}^+, \bar{\rho}^-)$, а также $(\bar{\rho}^+, N^+)$ значимо коррелировать не должны, с точки зрения «альтернативы Б» (см. п. 5б) коэффициент корреляции должен быть отрицательным, а с точки зрения гипотезы (х) коэффициент корреляции должен быть положительным и тем больше, чем сильнее в ней проявляются найденные эффекты.

В таблице 2 приведены соответствующие величины. Видно, что почти все они выстраиваются в цепочку «Loto France», «5/36», «6/49», как это и должно быть, если установленные эффекты пропорциональны количеству участников данной лотереи.

5. Необычность гипотезы (*) приводит к необходимости опровержения весьма необычных альтернативных моделей.

- а) Модель «Мафиози» – предположение о том, что сезонные подъемы плотности над математическим ожиданием объясняются сезонным подбрасыванием выигранных карточек в общий массив. Опровергается, поскольку средняя плотность для каждого из рассмотренных годов (кроме 1987 г.) не уклоняется значимо от математического ожидания, т.е. весенне-летний подъем компенсируется осенним и зимним спадами. Кроме того, если изъять из рассмотрения майский и июльский пики плотности, как обусловленные «мафиозными» причинами, то средняя плотность по всем остальным тиражам окажется аномально низкой ($P_{сл.} < 2,2\%$ для материала 1980–1986 гг., на котором пики выявляются с надежностью 99 %). Следовательно, эти пики не привнесены, а органичны для рассмотренного материала.
- б) Массовая сезонная смена тактики. Занумеруем все возможные комбинации каким-либо образом. Например, для лотереи «5/36»: $1, 2, \dots, C_{36}^5$. Обозначим число дублей данной комбинации m , представленной в тираже i – ν_m^i . Множество $\pi_i = \{ \nu_1^i, \nu_2^i, \dots, \nu_m^i \}$.

$\nu_{C_{36}^5}$ } есть тем самым, групповая тактика заполнения карточек в «i»-м тираже.

Может ли сезонная смена тактики породить наблюдаемые эффекты? В принципе такой механизм мог бы при конечном времени усреднения приводить к уменьшению \bar{p} в данном сезоне (например, при вырожденной тактике $\pi_d = \{N_i, O, O, \dots, O\}$). Однако такой механизм мог бы объяснить только сезонные спады плотности p , значимые же подъемы \bar{p} выше математического ожидания он объяснить не может (см. рис. 2). Выраженность максимума и вероятность его появления в рамках «нулевой» гипотезы приведены в таблице 2. Заметим при этом, что июльский пик превышает мат. ожидание не только интегрально – при вычислении \bar{p} [VII] по всем июльским тиражам. Величина \bar{p} ([VII], j) превышает \hat{p} и при рассмотрении каждого года j независимо. Так, в лотерее «5/36» из 8 рассмотренных годов в 7 \bar{p} ([VII], j) > \hat{p} (j = 1980–1982, 1984–87). С учетом того, что в 1980 г. существовало два независимых варианта лотереи «5/36» и в обоих \bar{p} [VII] > \hat{p} , вероятность такого ежегодного превышения:

$$P \leq 1 - B(8,9,0,448) \approx 0,88 \%,$$

где $B(x, y, z)$ – дается биномиальным распределением, 0,448 – отношение числа тиражей с $p_i > \hat{p}$ к общему числу тиражей. Та же картина наблюдается и в «Loto France». В этой лотерее во все 6 лет, по которым имелась 100%-ная информация об июльских тиражах, \bar{p} [VII] превосходит мат.ожидание (точнее \bar{p} [VII] > $\bar{p} + 3S/\sqrt{517}$, где \bar{p} – выборочное среднее, S – выборочный стандарт по всей совокупности тиражей «Loto France», см. п. 7-2). Для оставшихся 2 годов (1983, 1984 гг.) \bar{p} [VII] тоже превосходит $\bar{p} + 3S/\sqrt{517}$, однако автор располагает лишь 50 % июльского материала для них, и утверждать, что \bar{p} ([VII], 1983) > \hat{p} , \bar{p} ([VII], 1984) > \hat{p} нельзя. Как бы там ни было:

$$P \leq 1 - B(6,6,0.41) \approx 0,48 \%.$$

Следовательно, значимость июльского превышения плотности над мат. ожиданием оказывается более 99 % и при независимом рассмотрении годов. Кроме того, если альтернатива 5б верна, то величины (\bar{p}^+ , \bar{p}^-), введенные в п. 4, должны антикоррелировать. В самом деле, если в сезон [T] множество π менее однородно, чем в другие сезоны (т.е. некоторые комбинации дублируются сильнее, чем обычно, что и объясняет падение \bar{p} [T]), то рост \bar{p}^+ [T] с точки зрения альтернативы 5б означает увеличение неоднородности распределения π [T] в сезон [T]. Но тогда из-за роста неоднородности π [T] величины \bar{p}^- [T] падают, поскольку играющие заполняют меньшее количество комбинаций и чаще «попадают впросак». И наоборот, уменьшение \bar{p}^+ [T] означает рост однородности с соответствующим приближением и \bar{p}^+ и \bar{p}^- к \hat{p} (очевидно, что при полностью однородном распределении $\bar{p}^+ = \bar{p}^- = \hat{p}$), но только \bar{p}^+ приближается к \hat{p} сверху, а \bar{p}^- – снизу, что приводит к их антикорреляции. По аналогичным причинам

должны антикоррелировать и величины $(\bar{\rho}^+, N^+)$.

В реальности же величины $(\bar{\rho}^+, \bar{\rho}^-)$ и $(\bar{\rho}^+, N^+)$ в лотереях «Loto France» и «5/36» коррелируют положительно (см. табл. 2, рис. 4), что говорит о неприменимости альтернативы 5б к этим лотереям (но не к «6/49»!). Так, для «5/36» в периоды невезения (т.е. в марте, феврале, декабре – ранги 12, 11, 10 при помесечном ранжировании ρ) падение средней ρ определяется одновременным уменьшением числа тиражей с $\rho_i > \hat{\rho}$ (ранги $R(N^+_{\text{фев}}) = 12$, $R(N^+_{\text{март}}) = 11$, $R(N^+_{\text{дек}}) = 6.5$) и падением средней ρ^+ по этим тиражам ($R(\bar{\rho}^+_{\text{фев}}) = 7$, $R(\bar{\rho}^+_{\text{март}}) = 12$, $R(\bar{\rho}^+_{\text{дек}}) = 11$).

В случае применимости альтернативы 5б – это невозможно. Аналогично и в «Loto France» для минимальных среднемесячных значений $\bar{\rho}$ в феврале, январе и декабре (ранги 12, 11, 10) минимизируются как N^+ (ранги 12, 11, 8), так и $\bar{\rho}^+$ (ранги 10, 8, 12). Альтернатива 5б предсказывает обратное.

6. Весьма интересен следующий неожиданный факт, также не соответствующий «нулевой» гипотезе (и альтернативе 5б). Средняя плотность $\bar{\rho}$ [Т] за некоторый сезон [Т] может превзойти мат. ожидание $\hat{\rho}$ только случайно в рамках «нулевой» гипотезы и альтернативы 5б. Как видно из графика рисунка 3, для каждого из годов $J=1984-1987$ гг. среднегодовая ρ_i («5/36») $> \hat{\rho}$ («5/36»). Следовательно, эти величины «чисто» случайны. При рассмотрении $\bar{\rho}$, усредненной в 1984–1987 гг. по полугодиям, в 7 из 8 полугодий $\bar{\rho}$ [пол.] $> \hat{\rho}$. Следовательно, с точки зрения «нулевой» гипотезы они также «чисто» случайны. Однако плотность истинных предсказаний во французской лотерее, усредненная по тем же полугодиям, дает 96 %-ный (!) коэффициент корреляции с плотностью $\bar{\rho}$ [пол.] советской лотереи (рис. 5а).

Это невозможно с точки зрения «нулевой» гипотезы, однако факт. То же наблюдается и при рассмотрении всех полугодий 1980–1987 гг., в которых $\bar{\rho}$ [пол.] $> \hat{\rho}$ (рис. 5б, коэффициент корреляции $R = 76\%$ с надежностью коррелированности 99 %).

Коррелированность ρ в советской и французской лотереях сохраняется и при 3-месячном интервале усреднения в период спада С.А. Промежуток 1984–1987 гг. содержит 16 трёхмесячных интервалов, из которых в 12 $\bar{\rho}$ [3 мес] $> \hat{\rho}$. При этом коэффициент корреляции $R = 76,8\%$ с надежностью коррелированности 99,8 % (см. рис. 5в).

Замечания к статье

7-1. Критерий Стьюдента, критерий равенства средних и другие критерии однородности затабулированы в таблицах для нормального распределения. Однако распределение $\{\rho_i\}$ и в советской и во французской лотереях не является нормальным, в частности по критериям асимметрии и эксцесса (см. рис. 6). Вместе с тем оно не слишком «не-нормально»: так

отличия распределений $\{\rho_i\}$ от нормального по Д-критерию [4] не слишком сильны (вероятность применимости нормального распределения – 5 % для «5/36» и 30 % для «Loto France»). Критерий χ^2 тоже оставляет 5%-ную вероятность несущественности отличия $\{\rho_i\}$ от нормального. Поэтому оценки вероятности отклонения выборочных средних и критерии однородности вычислялись стандартными методами [4], но следует быть осторожным, полагаясь на конкретные величины этих вероятностей.

Это, разумеется, относится только к критериям, в которых важен вид распределения случайной величины. На соображения о многолетнем ходе, коррелированности ρ с внешними факторами W , T_p и лотерей между собой «не-нормальность» распределения существенно не влияет, так же как и на вывод о значимости июльского пика, полученного с помощью формул биномиального распределения в п. 5б.

7-2. В сообщениях о тиражах «Loto France» приводится только размер среднего выигрыша, приходящегося на комбинацию с 3, 4, 5, 6 верно угаданными числами. Поэтому в качестве индикатора плотности ρ_i использовалась величина $100/S_{3i}$, где S_{3i} – размер выигрыша, приходящегося на комбинацию с верно угаданными тремя номерами. Понятно, что чем выше плотность истинных предсказаний ρ_i , тем ниже размер среднего выигрыша S_{3i} . Для советской лотереи «5/36», в которой независимо известны ρ_i и S_{3i} , коэффициент корреляции $R(\rho_i, 100/S_{3i})$ весьма высок 96,2 %, что дает нам право, за неимением лучшего использовать $100/S_{3i}$ в качестве заменителя ρ_i . Таким же «косвенным» образом была получена и оценка числа заполняемых комбинаций N в «Loto France». Поскольку среди 517 рассмотренных тиражей лишь в 16 не было угадано ни одной комбинации с шестью верными числами, то, зная вероятность угадывания такой комбинации в одной попытке (дается гипергеометрическим распределением и составляет $P = 1/13983816$), нижнюю границу N получаем из условия $B(501, N, P) < 0,99$, где вероятность $B(x, y, z)$ дается биномиальным распределением. Отсюда $N > 12$ млн.

Заключение

Вывод. Совокупность рассмотренных массивов данных лотерей Франции («Loto France») и СССР («5/36») не удовлетворяет «нулевой» гипотезе.

Предложение. Сформулированная в начале статьи гипотеза (х) верна, и именно ею объясняются установленные нетривиальные факты. Если это так, то благодаря своей универсальности и универсальности физического параметра энтропии она может иметь применения в очень широком круге проблем.

Благодарности. В заключение хочу искренне поблагодарить В. Каймановича за чрезвычайно ценную конструктивную критику (им, в частности, была сформулирована альтернатива 5б, уточнены многие математические положения и формулировки), а также В. Орлова за ценные замечания и советы.

Автор был бы весьма благодарен в случае получения материала по национальным лотереям других стран, особенно южного полушария.

ПРИЛОЖЕНИЕ А

Энтропийный потенциал

Пусть в момент времени T_0 в точке (X_0, Y_0, Z_0) физического пространства осуществляется процесс..., например, выбор элемента X_i из конечного множества, содержащего m различных элементов $\{X_j\}_{j=1}^m$. Назовем выбор элемента x_i событием A_i . Таким образом, имеем несовместную систему событий $\{A_j\}_{j=1}^m$. Обозначим априорные вероятности осуществления каждого из событий A_i - P_i . Понятно, что в такой постановке задачи $\sum P_i = 1$.

Зафиксируем некоторый момент времени $T > T_0$. Обозначим энтропию системы R , в которой происходит процесс λ в момент времени T_0 - $S[R(T_0)]$, а энтропию системы R в момент времени T - $S[R(T)]$. Величина $S[R(T_0)]$ - фиксирована, а величина $S[R(T)]$, вообще говоря, априори не детерминирована и зависит от того, какое из событий A_i произойдет.

Из несовместности событий A_i не следует несовместность их следствий. Кроме того, поскольку мы рассматриваем систему R , находящуюся в условиях реального физического пространства, то из-за отсутствия абсолютной детерминированности событию A_i в момент времени T соответствует не одно, а множество состояний системы R (T, A_i, W) с определенными вероятностями $P(W)$ (W - многомерный параметр состояния системы R). Тогда, если известно, что событие A_i произошло, мат. ожидание энтропии системы R в момент T дастся интегралом (1)

$$\hat{S}[R(A_i, T)] = \int S[R(T, A_i, W)] \rho(W) d(W). \quad (1)$$

Если же неизвестно, какое из событий A_i произойдет, мат. ожидание энтропии системы R в момент T дастся интегралом (2)

$$\hat{S}[R(T)] = \sum \{ \int S[R(T, A_i, W)] \rho(W) d(W) \} \rho(A_i). \quad (2)$$

Вообще говоря, мат. ожидания $\hat{S}[R(T, A_i)]$ и $\hat{S}[R(T)]$ различны. В результате осуществления события A_i мат. ожидание энтропии системы R отклоняется от мат. ожидания энтропии $\hat{S}[R(T)]$ в ту или иную сторону.

Определение. Энтропийным потенциалом события A_i , определенным на момент T , называется разность мат. ожидания энтропии системы R после того как произошло событие A_i - $\hat{S}[R(A_i, T)]$, и до того, как оно произошло $\hat{S}[R(T)]$. Обозначая энтропийный потенциал $Z(A_i, T)$, имеем:

$$Z(A_i, T) = \int S[R(T, A_i, W)] \rho(W) dW - \sum_{i=1}^m \{ \int S[R(T, A_i, W)] \rho(W) dW \} \rho(A_i). \quad (3)$$

Замечания.

1. Понятно, что вырожденной системы, т.е. такой, что $1, W = W_0$
 $P(W) = \{ 0, W \neq W_0$
 $Z(A_i, T) = 0$ для любого i, T .
2. Очевидно, что энтропийный потенциал положителен, когда в результате события A_i мат. ожидание энтропии повышается

относительно мат. ожидания, и отрицателен в противном случае.

ПРИЛОЖЕНИЕ Б

Статья была написана в 1988 г. по массиву данных 1980–1987 гг. Данные 1988 г. также подтверждают выводы статьи.

- Июльские плотности ρ_3^5 и ρ (F) в 1988 г. снова превзошли мат. ожидания и еще более увеличили значимость июльского пика плотности (ср. п. 5б).
- Среднегодовые величины ρ_3^5 (1988), ρ_3^6 (1988 г.) (полных данных по «Loto France» 1988г. у автора пока нет), опустились ниже среднегодовых ρ_3^5 (1987 г.), ρ_3^6 (1987 г.), как это и должно быть при антикорреляции ρ с солнечной активностью (см. п. 3, рис. 3). С.А. в 1988 г. возросла по сравнению с 1987 г., т.к. Солнце находится на ветви роста 11-летнего цикла С.А.
- Поскольку, согласно гипотезе (х), плотность ρ можно рассматривать как индикатор энтропийного потенциала процессов, происходящих в данный момент в природе, ставился вопрос о моментах достижения величинами ρ своих экстремальных значений. Оказалось, что максимальная плотность

$\rho_3^5 = 53,14$ была зарегистрирована 5 мая 1985 г., а самая низкая в 1980–88 гг. $\rho_3^5 = 12,78$ пришлась на 4 декабря 1988 г., т.е. была зарегистрирована *за три дня до* трагического Армянского землетрясения. Автор уверен, что систематическая индикация на лабораторном стохастическом материале причин, формирующих ρ , могла бы позволить оперативно оценивать энтропийный потенциал процессов, происходящих в данный момент в природе, и принимать в случае необходимости соответствующие меры. Во Франции экстремальные значения $\rho(F) = 20$ и $\rho(F) = 5$ достигались 8 июля 1987 г. и 14 мая 1986 г.

БИБЛИОГРАФИЧЕСКИЙ СПИСОК

1. Козырев, Н.А. О воздействии времени на вещество / Н.А. Козырев; АН СССР // Физические аспекты современной астрономии. – Л.: Б.И., 1985. – С. 82–91. – Проблемы исследования Вселенной; Вып. 2.
2. Загускин, Ю.С. Исследование связи солнечной активности и тяжести последствий дорожно-транспортных происшествий в Москве / Ю.С. Загускин, В.Н. Иванов // Проблемы космической биологии. Т.43. – М.: Наука, 1982. – С. 59–63.
3. Соколовский, В.В. О биохимическом механизме реакции живых организмов на изменение солнечной активности / В.В. Соколовский // Проблемы космической биологии. Т. 43. – М.: Наука, 1982. – С. 180–193.
4. Большев, Л.Н. Таблицы математической статистики / Л.Н. Большев, Н.В.

- Смирнов. – М.: Наука, 1983. – 416 с.
5. Пиккарди, Дж. Химические основы медицинской климатологии / Дж. Пиккарди; пер. с англ. – Л.: Гидрометеоиздат, 1967. – 96 с.
 6. Опалинская, А.М. Влияние естественных и искусственных электромагнитных полей на физико-химические и элементарную биологическую системы / А.М. Опалинская, Л.П. Агулова. – Томск: Изд-во Томского ун-та, 1984. – 192 с.
 7. Ягодинский, В.Н. О зависимости эпидемического процесса от солнечной деятельности / В.Н. Ягодинский, З.П. Коноваленко, И.Л. Дружинин // Влияние солнечной активности на атмосферу и биосферу Земли. – М.: Наука, 1971. – С. 81–91.
 8. Solar-Geophysical Data. 1980-1987. NOAA. Boulder. P. 1.

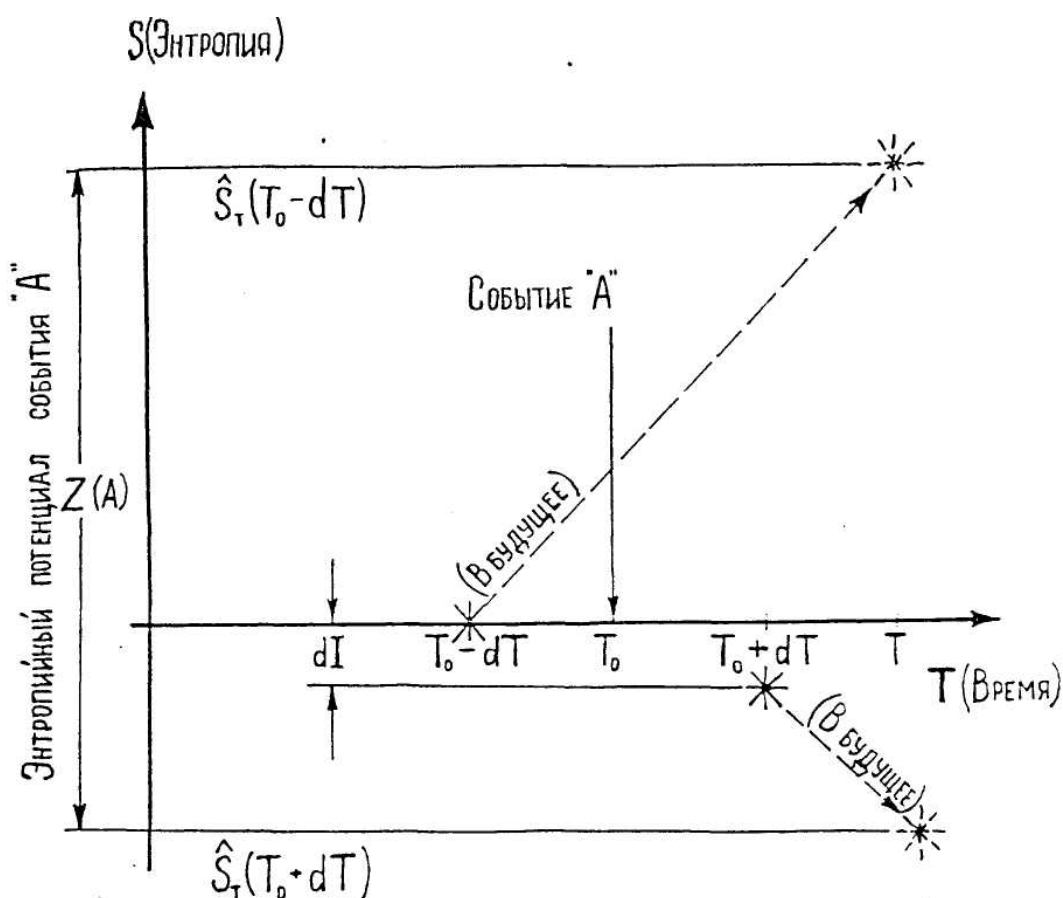


Рис. 1. Иллюстрация к определению энтропийного потенциала Z :
 $\hat{S}_T(T_0 - dT)$ – оценка энтропии на момент T , сделанная в момент $T_0 - dT$. $\hat{S}_T(T_0 + dT)$ – оценка энтропии на момент T , сделанная в момент $T_0 + dT$. Разность $\hat{S}_T(T_0 + dT) - \hat{S}_T(T_0 - dT) = Z(A)$ – энтропийный потенциал. Если событие «А» – приход сообщения А, то $Z(A)$ – ценность информации А, а $I = S(T_0 - dT) - S(T_0 + dT)$ – количество информации сообщения А

Таблица 1

Номер серии	«Loto France»			«5/36»			«6/49»	
	Реальный массив		Контр. массив	Реальный массив		Контр. массив	Реальный массив	Контр. массив
№	R	Q	R _{конт}	R	Q	R _{конт}	R	R _{конт}
1	57,3	93	-59,3	39,3	-	-17,0	-48,8	-55,6
2	50,1	-	-39,4	79,8	99	35,7	14,7	-43,3
3	63,3	95	-53,9	7,7	-	-21,0	-56,0	-18,2
4	56,4	92	-37,1	28,4	-	-39,1	28,7	-18,5
5	47,7	-	-1,5	63,6	95	19,6	-1,0	-46,2
6	18,9	-	-19,2	68,3	96	-38,6	-28,0	-55,3
7	76,1	98	-6,5	81,7	99	-32,6	-17,4	-42,5
8	31,8	-	-16,9	64,9	95	-23,9	-3,7	-47,5
9	43,2	-	-5,8	30,2	-	-25,8	-60,7	-58,5
10	36,3	-	-61,5	50,6	-	21,8	-18,1	-47,9

Коэффициенты корреляции R, надежность коррелированности Q пар $\bar{p}_J^I, \bar{p}_J^{II}$ – выборочных средних плотностей, полученных при случайном разбиении множества тиражей года J на подвыборки равного объема, при J=1980-87. № – номер серии («серия»: = «разбиение + вычисление коэффициента корреляции по 8 парам $\bar{p}_J^I, \bar{p}_J^{II}$ »). R_{конт.} – коэффициент корреляции, полученный описанной операцией на контрольном («перетасованном») массиве

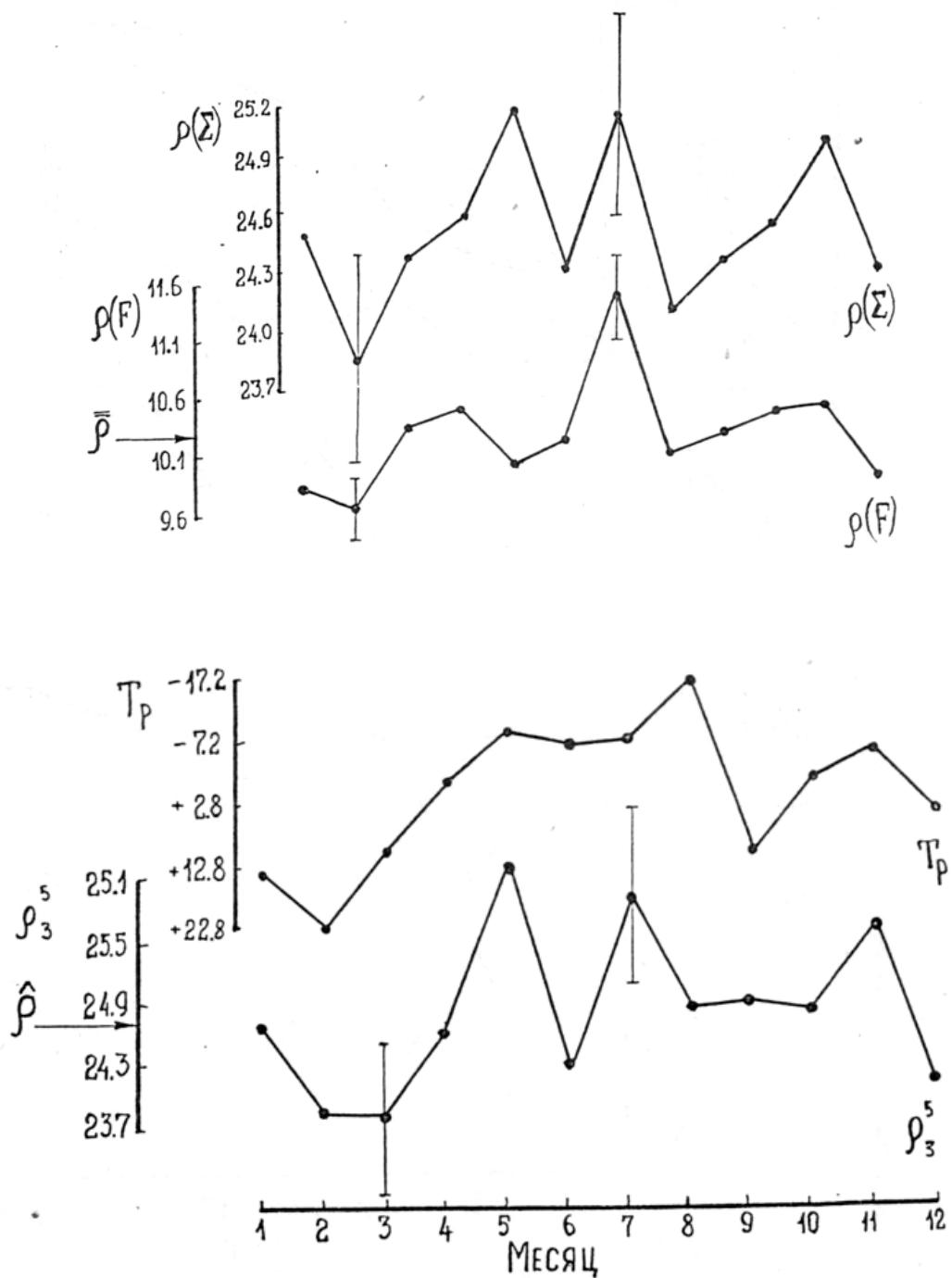


Рис. 2. Сезонный ход плотности истинных предсказаний в лотереях СССР: ρ_3^5 – «5/36», $\rho(\Sigma)$ – «5/36» + «6/49», Франции – $\rho(F)$ и сезонный ход теста Пиккарди T_p в 1971–1973 гг. по [6]

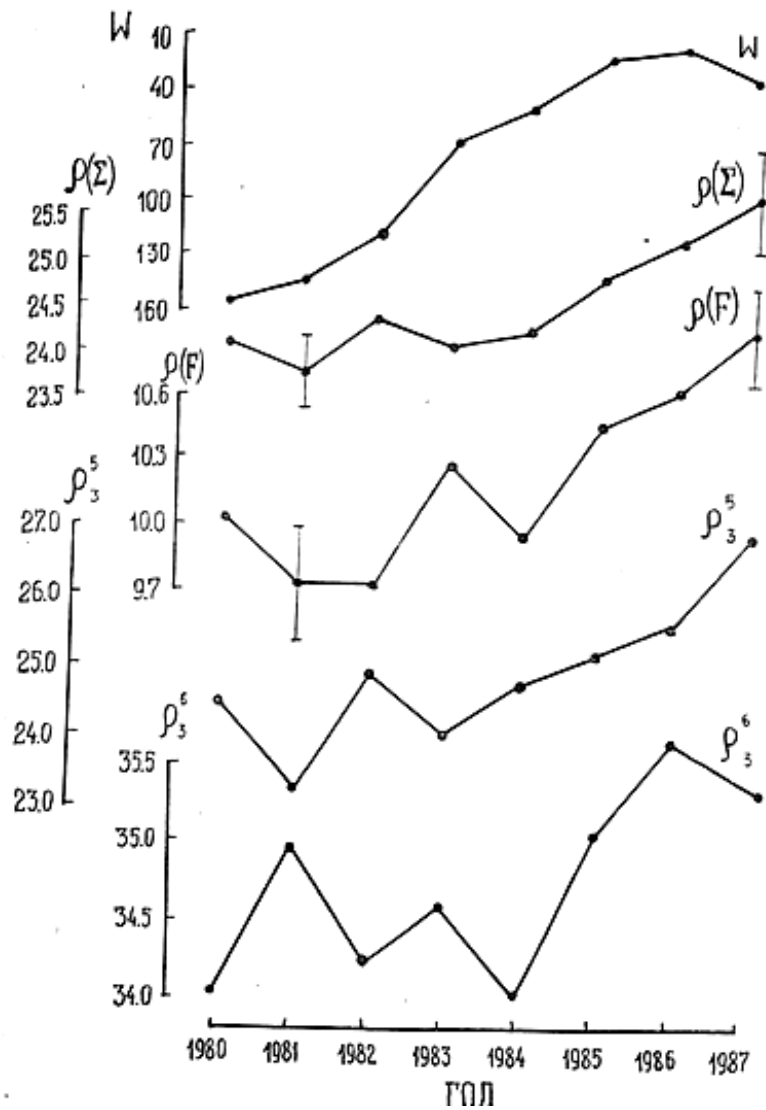


Рис. 3. Среднегодовые числа Вольфа W и среднегодовые плотности истинных предсказаний в лотереях:

$\rho(\Sigma)$ – «5/36» + «6/49»; $\rho(F)$ – «Loto France», ρ_3^5 – «5/36» ρ_3^6 – «6/49»

Таблица 2

Ранжировка эффектов в зависимости от числа участников лотереи

	L.F.	5/36	6/49
N (млн)	>12	8	1,4
$\rho_{\max} / \rho - 1$ %	11,5	4,4	3,2
$R_{сл}$ (%)	0,14	7,1	>20
V(%)	4,38	2,96	1,71
$R(\rho_j, W)$ %	-78,0	-67,4	-58,8
$R(\rho_k, T)$ %	-45,3	-61,7	+30,8
$R(N^+, \rho^+)$ %	66,7	22,4	-58,6
$R(\rho^+, \rho^{+-})$ %	67,2	39,8	-35,4

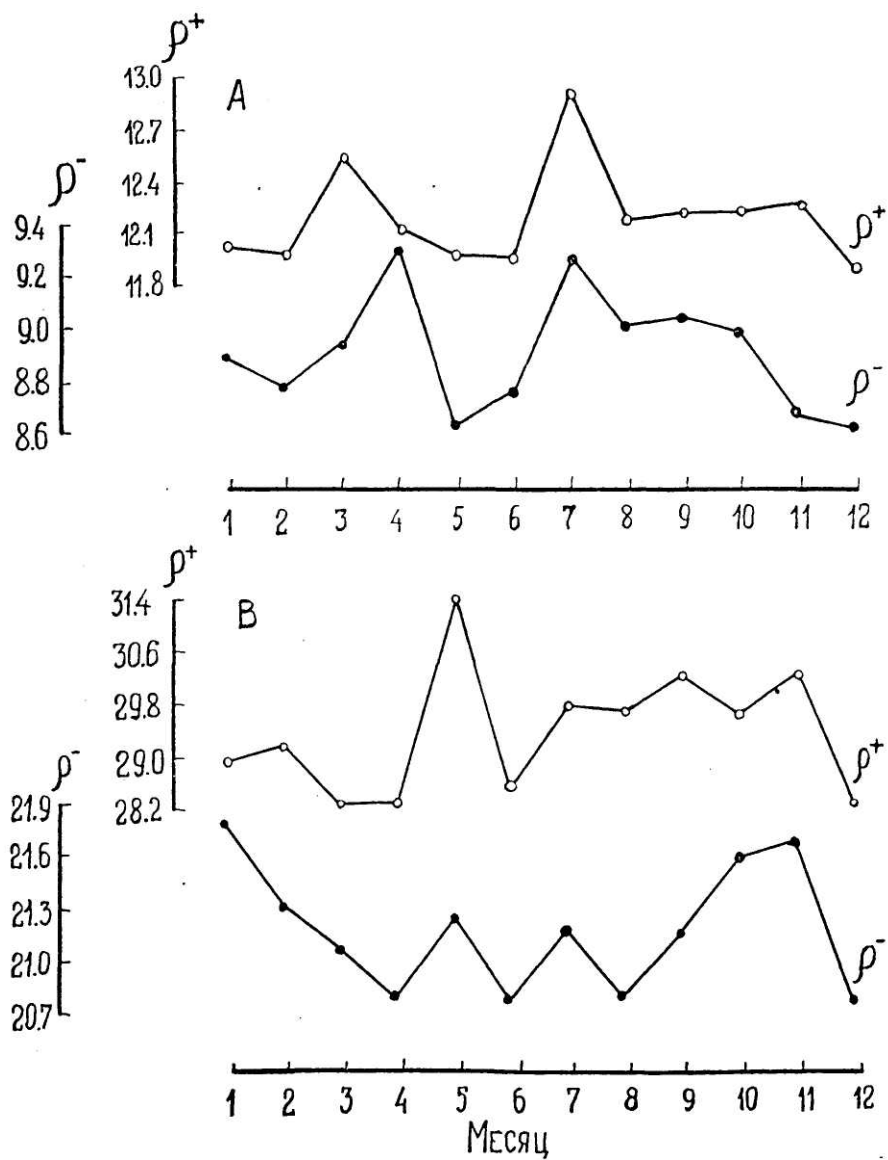


Рис. 4. Корреляция сезонного хода плотности для непересекающихся выборок $\{\rho\}^+ \{\rho\}^-$. «Loto France» – (A), «5/36» – (B)

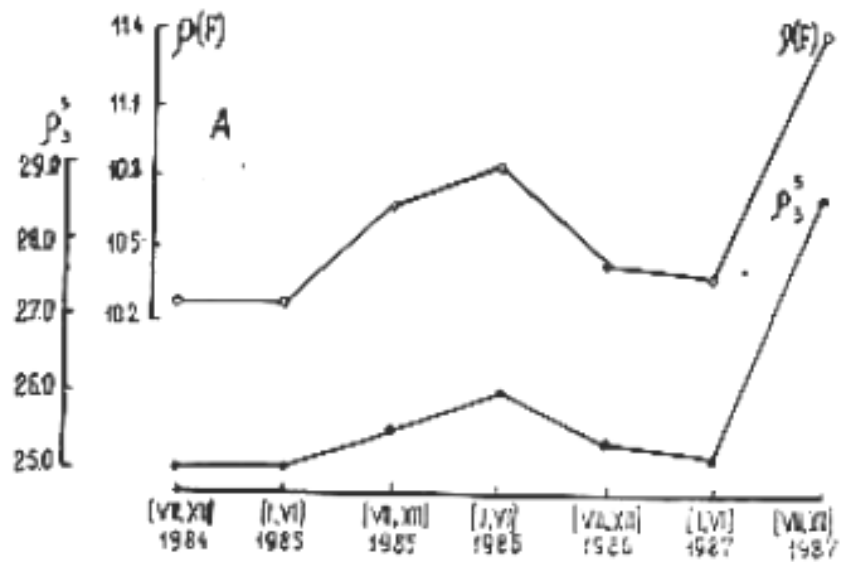


Рис. 5а. Корреляция средних плотностей плотностей ρ_3^5 и $\rho(F)$ в лотереях «5/36», «Loto France», для тех полугодий 1984–1987 гг., в которых $\bar{\rho}_3^5$ [пол] превосходит мат. ожидание $\hat{\rho}_3^5$

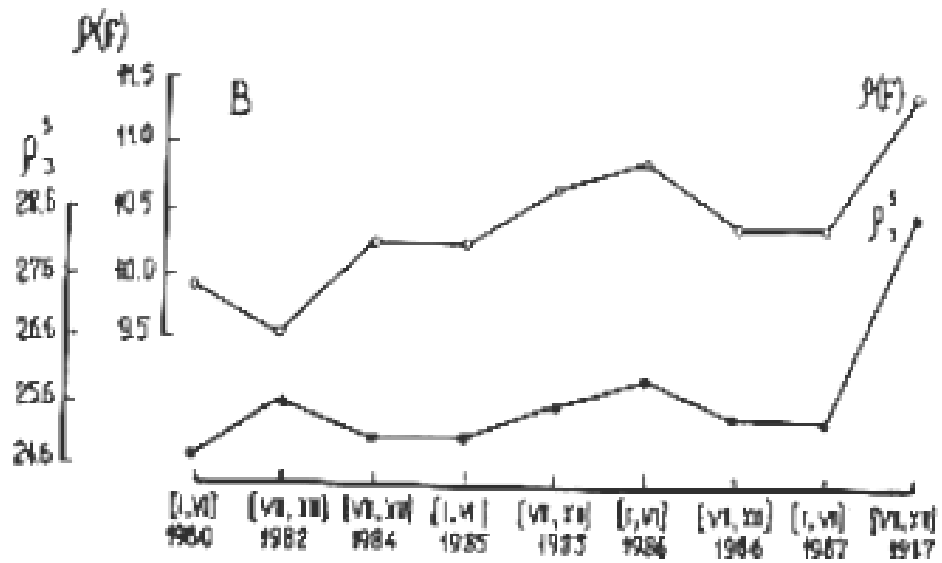


Рис. 5б. То же, что и 5а, но по всем полугодиям 1980–1987 гг., в которых $\bar{\rho}_3^5$ [пол] превосходит мат. ожидание $\hat{\rho}_3^5$

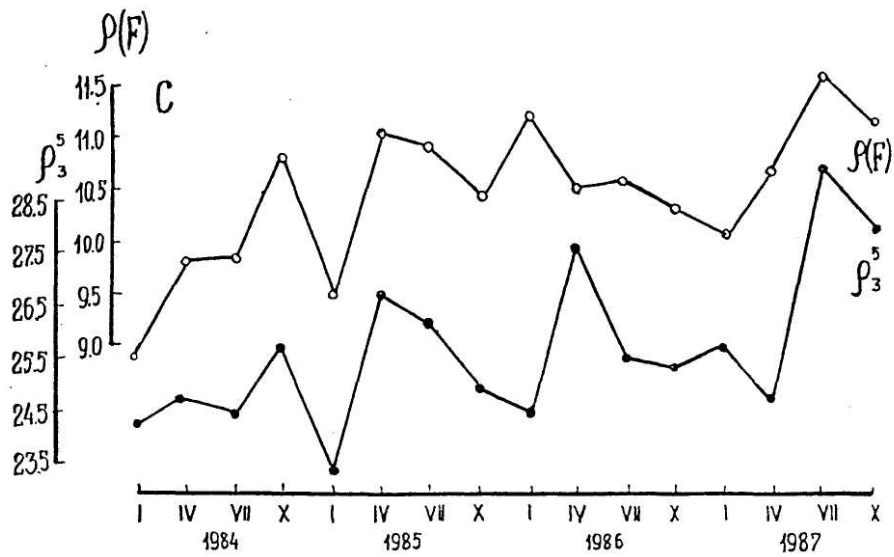


Рис. 5в. Корреляция средних в 3-месячных интервалах плотностей $P(F)$ и ρ_3^5 в лотереях «Loto France», «5/36», 1984–1987 гг.

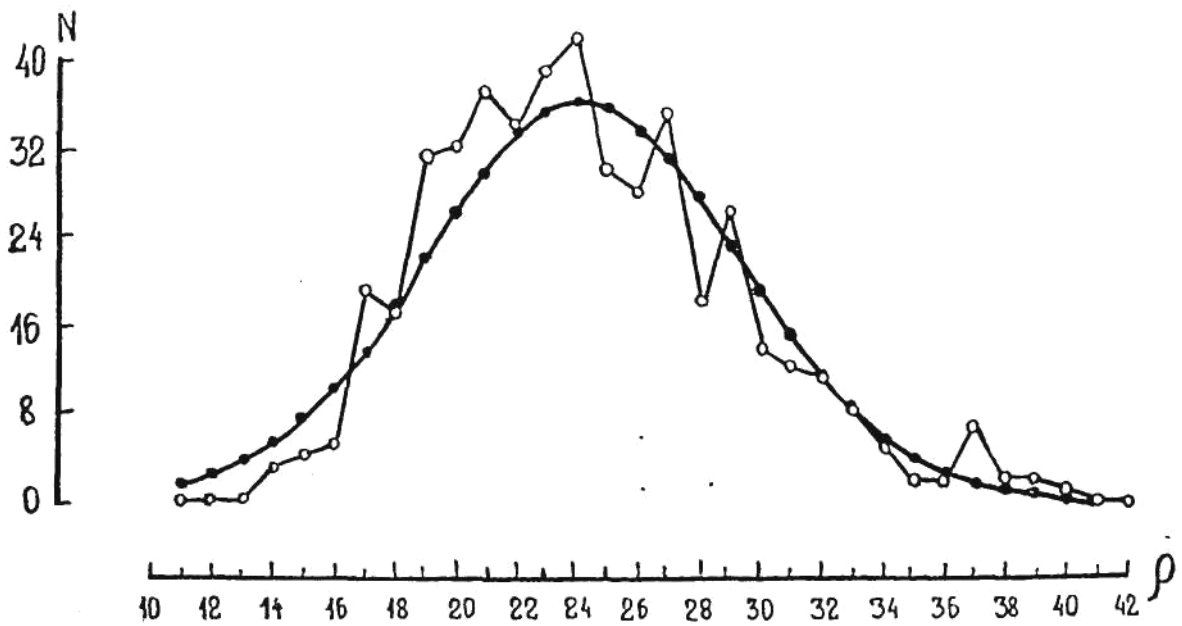


Рис. 6. Гистограмма плотности ρ_3^5 лотереи «5/36» и гауссиана с мат. ожиданием, равным мат. ожиданию ρ и нормированная к общему числу тиражей ($N = 469$)