Глава 3

СТАТИСТИЧЕСКАЯ МОДЕЛЬ РЕНЬИ

Приводятся основные методы и идеи теории информации, построенной на статистической модели Реньи. Выводятся и анализируются логарифмические меры энтропии и информации различия Реньи и их обобщения для аддитивных случайных объектов.

В отличие от подхода Шеннона–Винера, Альфред Реньи впервые ввел меры, зависящие от некоторого параметра q и играющие важную роль не только в теории информации, но и в теории мультифракталов. Результаты работы [103] и монографии [104] явились значительным шагом вперед и способствовали началу развития обобщенной теории информации. Именем Реньи назван Институт математики Венгерской академии наук, первым директором которого он являлся долгое время.

3.1. Взвешенное среднее Колмогорова–Нагумо с произвольной функцией

Рассмотрим множество всех состояний случайного объекта, описываемых распределением $p=\{p_1,\ldots,p_m\}$ и множество случайных величин $T=\{T_1,\ldots,T_m\}$.

Определение 1. Среднее Колмогорова–Нагумо с весом $p=\{p_1,\ldots,p_m\}$, $p_i>0$ и произвольной функцией $\varphi=\varphi(T)$ имеет вид [81, 94]

$$A_\varphi(T)=\varphi^{-1}\left[\frac{\sum_i^m \varphi(T_i)\,p_i}{\sum_i^m p_i}\right], \qquad (3.1.1)$$

где $\varphi = \varphi(T)$ – непрерывная строго монотонная функция на $\mathbf{R}$, а $\varphi^{-1}(T)$ – функция, обратная $\varphi(T)$.

Если выполняется условие вероятностной нормировки для распределения, то из (3.1.1) следует выражение

$$A_\varphi(T) = \varphi^{-1}\left[\sum_i^m \varphi(T_i)\, p_i\right], \quad \sum_i^m p_i = 1. \tag{3.1.2}$$

Приведем основные свойства взвешенного среднего Колмогорова–Нагумо с произвольной функцией [47, 63, 81, 94].

1. Нормировка. Для неслучайной величины, имеющей постоянное значение C, справедливо равенство

$$A_\varphi(\mathrm{C}) = \mathrm{C}. \tag{3.1.3}$$

Поскольку при $\mathrm{C} = 1$ из (3.1.3) вытекает $A_\varphi(1) = 1$, то имеем свойство нормированности взвешенного среднего на единицу.

2. Эквивалентные средние. Необходимым и достаточным условием равенства

$$A_\varphi(T) = A_\chi(T) \tag{3.1.4}$$

для всех T и p является условие

$$\chi = \alpha\varphi + \beta, \tag{3.1.5}$$

где α и β – постоянные и $\alpha \neq 0$. Средние с функциями φ и χ называются эквивалентными средними.

3. Однородность. Пусть $\varphi(T)$ непрерывна в открытом интервале $(0, \infty)$ и пусть

$$A_\varphi(aT) = aA_\varphi(T) \tag{3.1.6}$$

для всех положительных T, p и a. Очевидно, что (3.1.6) справедливо, когда $\varphi = T^q$ или $\varphi = \log_2 T$. В общем случае, согласно (3.1.5), для φ имеем функциональное уравнение

$$\varphi(T_1 T_2) = \varphi(T_1) + \varphi(T_2) + \varepsilon\varphi(T_1)\varphi(T_2). \tag{3.1.7}$$

При $\varepsilon = 0$ из (3.1.7) вытекает $\varphi(T) = \lambda \log_2 T$. Если $\varepsilon \neq 0$, то общим решением (3.1.7) является

$$\varphi(T) = \varepsilon^{-1}\left(T^q - 1\right). \tag{3.1.8}$$

4. Сравнимость. Пусть ψ и χ являются непрерывными и строго монотонными функциями. Чтобы $A_\psi(T)$ и $A_\chi(T)$ были сравнимы, то есть для них справедливо неравенство

$$A_\psi(T) \le A_\chi(T) \qquad (3.1.9)$$

для всех T и p, необходимо и достаточно, чтобы функция

$$\varphi = \chi\psi^{-1} \qquad (3.1.10)$$

удовлетворяла неравенству

$$\varphi\left(\frac{\sum_i^m T_i p_i}{\sum_i^m p_i}\right) \le \frac{\sum_i^m \varphi(T_i) p_i}{\sum_i^m p_i}, \qquad (3.1.11)$$

если χ – возрастающая функция; если же χ убывает, то в (3.1.11) имеем обратный знак неравенства.

5. Выпуклость. Если для произвольной непрерывной функции φ справедливо строгое неравенство в (3.1.11), кроме тех случаев, когда $T_i = \text{const}$ или $\varphi(T)$ линейная функция, то эта функция обладает свойством выпуклости. Функция $(-\varphi)$ есть вогнутая и в (3.1.11) имеет обратный знак. Необходимым и достаточным условием выпуклости $\varphi(T)$ в рассматриваемом интервале является неравенство $\varphi''(T) \ge 0$. Класс функций φ исследуется в теории выпуклых функций.

6. Сумма с произвольной функцией. Если $p_i = 1$ или $p_i = 1/m$, то с точностью до постоянной $1/m$ имеем сумму с произвольной функцией

$$A_\varphi(T) = \varphi^{-1}\left[\sum_i^m \varphi(T_i)\right], \qquad (3.1.12)$$

которая непрерывна и строго монотонна. При строгом равенстве (3.1.11) с $p_i = 1/m$ выражение (3.1.12) есть среднее Колмогорова–Нагумо с произвольной функцией для равновероятного состояния.

7. Флуктуация. Рассмотрим отклонение произвольной функции случайной величины T_i от среднего по Колмогорову–Нагумо и введем флуктуацию:

$$\Delta\varphi(T_i)=\varphi(T_i)-A_\varphi(T),\quad \sum_i^m\left[\Delta\varphi(T_i)\right]p_i=0, \tag{3.1.13}$$

начальные и центральные моменты n-го порядка

$$\alpha_n=\sum_i^m\varphi^n(T_i)\,p_i, \tag{3.1.14}$$

$$\mu_n=\sum_i^m\left[\varphi(T_i)-A_\varphi(T)\right]^n p_i. \tag{3.1.15}$$

8. f-взвешенное среднее с произвольной функцией. Функциональное обобщение взвешенного среднего Колмогорова–Нагумо запишется так [63]:

$$A_{\varphi,f}(T)=\varphi^{-1}\left[\frac{\sum_i^m\varphi(T_i)\,f(p_i)}{\sum_i^m f(p_i)}\right]. \tag{3.1.16}$$

Если $f(p_i)=p_i$, то f – взвешенное среднее с произвольной функцией совпадает с выражением (3.1.1).

Определение 2. Аналогом взвешенного среднего произвольной функции каждой непрерывной случайной функции $T=T(X)$ в состоянии $p=p(X)$ является выражение

$$A_\varphi(T)=\varphi^{-1}\left[\frac{\int_G\varphi(T)\,pdX}{\int_G pdX}\right], \tag{3.1.17}$$

где

$$0<\int_G pdX<1. \tag{3.1.18}$$

При условии вероятностной нормировки для всего пространства из (3.1.17) вытекает

$$A_\varphi(T)=\varphi^{-1}\left[\int\varphi(T)\,pdX\right],\quad \int pdX=1. \tag{3.1.19}$$

Для квантовой величины T, являющейся эрмитовым оператором, имеем выражение

$$A_\varphi(T)=\varphi^{-1}\left[\mathrm{Sp}\varphi(T)\cdot\rho\right],\ \mathrm{Sp}\rho=1, \tag{3.1.20}$$

где ρ – оператор смешанного состояния случайного объекта.

В заключение рассмотрим различные функции $\varphi(T)$ и получим следующие средние по Колмогорову–Нагумо [47, 63]

$\varphi(T)$: $\qquad\qquad A_{\varphi}(T)$:

$$T \qquad \mathbf{E}(T)=\sum_{i}^{m} T_i p_i\,, \tag{3.1.21}$$

$$T^{\varepsilon} \qquad N_{\varepsilon}(T)=\left(\sum_{i}^{m} T_i^{\varepsilon} p_i\right)^{1/\varepsilon} \tag{3.1.22}$$

$$\log_2 T \qquad N(T)=2^{\sum_{i}^{m}(\log_2 T)p_i}\,, \tag{3.1.23}$$

$$2^{\varepsilon T} \qquad H_{\varepsilon}(T)=\frac{1}{\varepsilon}\log_2\left(\sum_{i}^{m} 2^{\varepsilon T_i} p_i\right), \tag{3.1.24}$$

$$\sin T \qquad S(T)=\arcsin\left[\sum_{i}^{m}\left(\sin T_i\right) p_i\right], \tag{3.1.25}$$

некоторые из которых рассматриваются в статистической теории информации.

3.2. Полунормы

Рассмотрим множество всех состояний случайного объекта, описываемых распределением $p=\{p_1,\ldots,p_m\}$, и множество случайных величин $T=\{T_1,\ldots,T_m\}$. Взвешенное среднее равно

$$\mathbf{E}(T)=\frac{\sum_{i}^{m} T_i p_i}{\sum_{i}^{m} p_i}, \quad 0<\sum_{i}^{m} p_i \le 1. \tag{3.2.1}$$

Определение. Каждой случайной величине T и любому числу q, $-\infty<q<\infty$, соответствует функция [4, 47, 63]:

$$N_q(T)=\left(\frac{\sum_i^m T_i^q p_i}{\sum_i^m p_i}\right)^{1/q}. \qquad (3.2.2)$$

Если веса p_i удовлетворяют условию вероятностной нормировки

$$\sum_i^m p_i = 1, \qquad (3.2.3)$$

то выражения (3.2.1) и (3.2.2) примут следующий вид:

$$\mathbf{E}(T)=\sum_i^m T_i p_i, \qquad (3.2.4)$$

$$N_q(T)=\left(\sum_i^m T_i^q p_i\right)^{1/q}. \qquad (3.2.5)$$

При $p_i = 1$ из (3.2.2) имеем средние арифметическое и геометрическое

$$\mathbf{E}(T)=N_1(T)=\frac{1}{m}\sum_i^m T_i, \quad M(T)=N_{-1}(T)=\left(\frac{1}{m}\sum_i^m T_i^{-1}\right)^{-1}. \qquad (3.2.6)$$

В зависимости от областей изменения числа q: $q<0$, $0<q<1$, $1\le q<\infty$ определяется тот или иной тип вероятностного q-пространства.

Приведем основные свойства функции $N_q(T)$ [4, 47, 63].

1. Однородность. При замене p на ap $(a>0)$ функция (3.2.2) является однородным функционалом нулевой степени относительно p, то есть выполняется свойство однородности.

2. Положительность и нормированность. Для всей области изменения числа $q\in \mathbf{R}$ функция является положительной и неубывающей. Если $q<0$ $(q>0)$ и некоторые T_i равны нулю, то $N_q(T)=0$ $(N_q(T)=\infty)$.

Для неслучайной постоянной величины C имеем равенство:

$$N_q\left(C\right)=\left(\frac{\sum_i^m C^q p_i}{\sum_i^m p_i}\right)^{1/q}=C\,, \qquad (3.2.7)$$

из которого следует свойство нормированности функции на единицу

$$N_q\left(1\right)=1\,. \qquad (3.2.8)$$

3. Полунормы $N_q\left(T\right)$. Для значений $1\le q<\infty$ число $N_q\left(T\right)$ есть полунорма в локально выпуклом пространстве $\mathbf{L}^q$, для которого справедливы соотношения:

$$1)\ N_q\left(aT\right)=\left|a\right|N_q\left(T\right), \qquad (3.2.9)$$

2) $N_q\left(T_1+T_2\right)\le N_q\left(T_1\right)+N_q\left(T_2\right)$ (неравенство Минковского), (3.2.10)

где $a\ne 0$ есть произвольное число. При $0<q<1$ неравенство (3.2.10) не выполняется.

4. Невырожденность. Для полунормы допустимо $N_q\left(T\right)=0$ при $T\ne 0$. Этим свойством полунорма отличается от нормы при $q=2$

$$N_2\left(T\right)=\left(\sum_i^m T_i^2 p_i\right)^{1/2}, \qquad (3.2.11)$$

для которой $N_2\left(T\right)=0$ при $T=0$.

5. Неравенство Гельдера. При $0<q\le\infty$ и $0<q\le\infty$ выполняется неравенство Гельдера

$$N_r\left(T_1T_2\right)\le N_q\left(T_1\right)N_{q'}\left(T_2\right), \qquad (3.2.12)$$

где для произвольных случайных величин T_1 и T_2 имеем

$$N_q\left(T_1\right)=\left(\sum_i^m T_{1i}^q p_i\right)^{1/q},\ N_{q'}\left(T_2\right)=\left(\sum_i^m T_{2i}^{q'} p_i\right)^{1/q'},$$

$$N_r\left(T_1T_2\right)=\left[\sum_i^m\left(T_{1i}T_{2i}\right)^r p_i\right]^{1/r}. \qquad (3.2.13)$$

Конечные числа q и q' есть так называемые сопряженные показатели, которые удовлетворяют равенству

$$\frac{1}{q}+\frac{1}{q'}=\frac{1}{r}, \quad 0<r<q. \tag{3.2.14}$$

Знак равенства в (3.2.12) достигается тогда и только тогда, когда выполняется условие

$$T_{1i}^{q}=bT_{2i}^{q'}, \tag{3.2.15}$$

где положительный коэффициент

$$b=\left[N_q\left(T_1 \right)\right]^q \left[N_{q'}\left(T_2 \right)\right]^{-q'}. \tag{3.2.16}$$

При $1 \le q < \infty$ пространство $\mathbf{L}^{q'}$ есть сопряженное пространству $\mathbf{L}^q$ и величины T_1 и T_2 принадлежат, соответственно, этим пространствам. Для всякого $0<T<1$ имеется сопряженная величина $T'=1-T$.

Пусть $T_1=T$ и $T_2=1$, тогда из (3.2.12) вытекает следствие, что

$$N_r\left(T\right) \le N_q\left(T\right) \tag{3.2.17}$$

при любом $q>0$, где $0<r<q$, то есть $N_q\left(T\right)$ – возрастающая функция при увеличении значения q.

При $r=1$ из (3.2.17) получим важное неравенство

$$\mathbf{E}\left(T\right) \le N_q\left(T\right). \tag{3.2.18}$$

Полунорма больше среднего значения случайной величины.

6. Выпуклость функций $\boldsymbol{N_q\left(T\right)}$ и $\boldsymbol{\log_2 N_q\left(T\right)}$. Рассмотрим неравенство Минковского (3.2.10) при $T=aT_1+\left(1-a\right)T_2$ $\left(0\le a\le 1\right)$. Тогда получим свойство выпуклости для $N_q\left(T\right)$ в виде

$$N_q\left(T\right) \le aN_q\left(T_1\right)+\left(1-a\right)N_q\left(T_2\right). \tag{3.2.19}$$

При значениях $q>0$, для которых $N_q\left(T\right)$ конечно, функция $\ln N_q\left(T\right)$ является строго выпуклой от $1/q$. Для любых $0<a<1$ имеет место неравенство

$$\log N_q\left(T\right) \le a\log_2 N_r\left(T\right)+\left(1-a\right)\log_2 N_\nu\left(T\right), \tag{3.2.20}$$

где $r>0$, $\nu>0$, а числа r, q и ν удовлетворяют равенству:

$$\frac{1}{q}=\frac{a}{r}+\frac{1-a}{\nu}. \qquad (3.2.21)$$

7. Дифференцируемость функций $N_q(T)$ и $\log_2 N_q(T)$. Функции $N_q(T)$ и $\log_2 N_q(T)$ бесконечно дифференцируемы в каждой точке интервала, в котором они конечны. Функция $\left[N_q(T)\right]^q=\sum_i^m T_i^q p_i$ имеет предел

$$\lim_{q\to 0}\sum_i^m T_i^q p_i=1, \qquad (3.2.22)$$

когда q стремится к 0, и имеет в этой точке первую производную,

$$\frac{d}{dq}\sum_i^m T_i^q p_i=\lim_{q\to 0+0}\sum_i^m\left(2^{q\log_2 T_i}\right)p_i=\ln 2\sum_i^m\left(\log_2 T_i\right)p_i. \qquad (3.2.23)$$

При этом $N_q(T)$ и $\ln N_q(T)$ имеют пределы

$$\lim_{q\to 0}N_q(T)=N(T)=2^{\sum_i^m(\log_2 T_i)p_i}, \qquad (3.2.24)$$

$$\lim_{q\to 0}\log_2 N_q(T)=\sum_i^m\left(\log_2 T_i\right)p_i. \qquad (3.2.25)$$

Если T^q и $\ln T$ интегрируемы, то

$$N(T)\le N_q(T), \qquad (3.2.26)$$

причем равенство достигается лишь в том случае, когда T постоянна почти всюду.

Величина $N(T)$ есть среднее геометрическое функции T.

8. Взаимосвязь полунорм. Неравенства. Используем приведенные свойства и дадим дополнительно формулы взаимосвязи полунорм

$$N_q(T)N_{-q}\left(T^{-1}\right)=1, \qquad (3.2.27)$$

$$N_{q\nu}(T)=\left[N_q\left(T^{\nu}\right)\right]^{1/\nu} \qquad (3.2.28)$$

и неравенства

$$N_r\left(T_1T_2\ldots T_n\right)\le N_{q_1}\left(T_1\right)N_{q_2}\left(T_2\right)\ldots N_{q_n}\left(T_n\right), \qquad (3.2.29)$$

$$N_r\left(T_1+T_2+\ldots+T_n\right)\le N_{q_1}\left(T_1\right)+N_{q_2}\left(T_2\right)+\ldots+N_{q_n}\left(T_n\right). \quad (3.2.30)$$

В обобщенном неравенстве Гельдера (3.2.29) имеем

$$\frac{1}{q_1}+\frac{1}{q_2}+\cdots+\frac{1}{q_n}=\frac{1}{r},\ 0<q_i<\infty. \quad (3.2.31)$$

Неравенство (3.2.30) представляет собой обобщенное неравенство Минковского для $1\le q_i<\infty$.

Если $0<r<q<\nu$, то находим неравенство

$$\left[N_q(T)\right]^q<\left\{\left[N_r(T)\right]^r\right\}^{\frac{\nu-q}{\nu-r}}\left\{\left[N_\nu(T)\right]^\nu\right\}^{\frac{q-r}{\nu-r}}, \quad (3.2.32)$$

которое при $q=ra+\nu(1-a)$ и $0<a<1$ запишется так:

$$\sum_i^m T_i^q p_i<\left(\sum_i^m T_i^r p_i\right)^a\left(\sum_i^m T_i^\nu p_i\right)^{1-a}. \quad (3.2.33)$$

9. Взвешенное обратно-гармоническое среднее. Определим для каждой случайной величины T и любого числа q, $-\infty<q<\infty$, взвешенное обратно-гармоническое среднее [63]

$$M_q(T)=\left[\frac{N_q(T)}{N_{q-1}(T)}\right]^{q-1}N_q(T)=\frac{\sum_i^m T_i^q p_i}{\sum_i^m T_i^{q-1}p_i}. \quad (3.2.34)$$

При $q=1$ и $q=0$ из (3.2.24) вытекают взвешенное среднее и взвешенное гармоническое среднее. Если $p_i=1$, то они совпадают с соответствующими значениями (3.2.6).

Функционал (3.2.34) вытекает как частный случай при $s=q-1$ от выражения [63]

$$T_q^s(T)=\left(\frac{\sum_i^m T_i^q p_i}{\sum_i^m T_i^s p_i}\right)^{1/(q-s)}, \quad (3.2.35)$$

имеющего важную роль при определении статистического состояния в обобщенной теории информации [66].

В заключение приведем аналог функции (3.2.2) каждой непрерывной величины $T = T(X)$ в состоянии p

$$N_q(T) = \left(\frac{1}{\Gamma(G)} \int_G T^q d\Gamma \right)^{1/q} = \left(\frac{\int_G T^q p dX}{\int_G p dX} \right)^{1/q}, \qquad (3.2.36)$$

где для весовой функции имеем неравенство

$$0 < \int_G p dX < \infty. \qquad (3.2.37)$$

При условии вероятностной нормировки для всего пространства из (3.2.36) следует

$$N_q(T) = \left(\int T^q d\Gamma \right)^{1/q} = \left(\int T^q p dX \right)^{1/q}, \ \int p dX = 1. \qquad (3.2.38)$$

Для квантовой величины T, представляемой эрмитовым оператором, определяем

$$N_q(T) = \left(\mathrm{Sp} T^q \rho \right)^{1/q}, \ \mathrm{Sp}\,\rho = 1, \qquad (3.2.39)$$

где ρ – оператор смешенного состояния случайного объекта.

3.3. Полунормы распределений

Рассмотрим определение полунормы распределения

$$N_q(p) = \left(\frac{\sum_i^m p_i^{q+1}}{\sum_i^m p_i} \right)^{1/q}, \qquad (3.3.1)$$

которая при условии вероятностной нормировки имеет вид

$$N_q(p) = \left(\sum_i^m p_i^{q+1} \right)^{1/q}. \qquad (3.3.2)$$

Согласно свойствам 3, 4, 6 и 7, рассмотренным в предыдущем разделе, полунорма распределения является невырожденной, выпуклой и дифференцируемой функцией по аргументу q. Причем из (3.2.24) и (3.2.25) вытекают следующие равенства:

$$\lim_{q \to 0} N_q(p) = 2^{\sum_i^m (\log_2 p_i) p_i}, \qquad (3.3.3)$$

$$\lim_{q\to 0}\log_2 N_q\left(p\right)=\sum_i^m\left(\log_2 p_i\right)p_i\ . \tag{3.3.4}$$

Отметим некоторые дополнительные свойства $N_q\left(p\right)$.

1. Мультипликативность. Для совместного распределения p_{12} независимых объектов с распределениями p_1 и p_2 имеем равенство

$$N_q\left(p_{12}\right)=N_q\left(p_1\right)N_q\left(p_2\right), \tag{3.3.5}$$

где полунормы

$$N_q\left(p_1\right)=\left(\sum_i^m p_{1i}^{q+1}\right)^{1/q},\ N_q\left(p_2\right)=\left(\sum_i^m p_{2i}^{q+1}\right)^{1/q}, \tag{3.3.6}$$

$$N_q\left(p_{12}\right)=\left(\sum_i^m\sum_j^n p_{ij}^{q+1}\right)^{1/q} \tag{3.3.7}$$

и $p_{ij}=p_i p_j$. Произведение полунорм в (3.3.5) определяет свойство мультипликативности.

2. Полунормы равновероятного распределения. Подставим в выражение (3.3.2) распределение равновероятного состояния

$$p_i=\frac{1}{m} \tag{3.3.8}$$

и получим, что полунорма

$$N_q\left(p\right)=\frac{1}{m} \tag{3.3.9}$$

не зависит от числа q и совпадает с распределением (3.3.8).

3. Полунорма частного распределения. При $q>0$ справедливо неравенство

$$\left[\sum_i^m\left(\sum_j^n p_{ij}\right)^q\right]^{1/q}\le\sum_j^n\left(\sum_{ij}^m p_{ij}^q\right)^{1/q} \tag{3.3.10}$$

для совместного распределения p_{ij}. Знак равенства достигается при $p_{ij}=p_i p_j$ для независимых объектов с распределениями вероятностей p_i и p_j.

Используем выражение полунормы частного распределения

$$N_{q-1}(p_1)=\left(\sum_i^m p_{1i}^q\right)^{1/(q-1)}, \; p_i=\sum_j^n p_{ij} \qquad (3.3.11)$$

и из (3.3.10) получим

$$\left[N_{q-1}(p_1)\right]^{1/q'} \le \sum_j^n \left[N_{q-1,j}(p_{12})\right]^{1/q'}, \qquad (3.3.12)$$

где

$$N_{q-1,j}(p_{12})=\left(\sum_i^m p_{ij}^q\right)^{1/(q-1)}. \qquad (3.3.13)$$

Далее примем $T=p/u$ и определим, согласно (3.2.5), полунорму

$$N_q\left(\frac{p}{u}\right)=\left[\sum_i^m\left(\frac{p_i}{u_i}\right)^q p_i\right]^{1/q}, \qquad (3.3.14)$$

которая обладает всеми перечисленными свойствами для $N_q(p)$.

При значении $q=0$ из (3.2.24) и (3.2.25) следуют пределы

$$\lim_{q\to 0} N_q\left(\frac{p}{u}\right)=2^{\sum_i^m\left(\log_2\frac{p_i}{u_i}\right)p_i}, \qquad (3.3.15)$$

$$\lim_{q\to 0}\log_2 N_q\left(\frac{p}{u}\right)=\sum_i^m\left(\log_2\frac{p_i}{u_i}\right)p_i. \qquad (3.3.16)$$

Предопределяя дальнейшие результаты, отметим, что функционалы $H_q(p)=-\log_2 N_{q-1}(p)$ и $I_q(p:u)=\log_2 N_{q-1}\left(\frac{p}{u}\right)$ есть аддитивные q-энтропия и q-информация различия Реньи [105, 106].

Приведем наглядную иллюстрацию рассматриваемых свойств функционалов.

На рис.3.1 представлены зависимости полунорм распределений $N_{q-1}(p)$ и $N_{q-1}\left(\frac{p}{u}\right)$ от числа q при значениях $m=2$, $p_1=1/4$, $u_1=1/3$. Видно монотонное возрастание положительных функций при увеличении значения q.

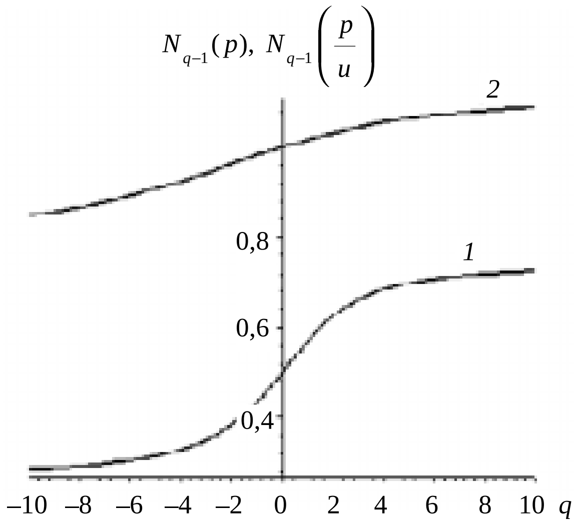


Рис. 3.1. Зависимости полунорм распределений от числа q:

$1 - N_{q-1}(p)$, $2 - N_{q-1}\left(\frac{p}{u}\right)$

Выпуклость полунорм распределений видна из рис.3.2 и рис.3.3, где даются зависимости их от распределения при значениях $m=2$, $p_1=1/4$, $u_1=1/3$ и $q=-3;\ -1;\ 0;\ 1;\ 3$.

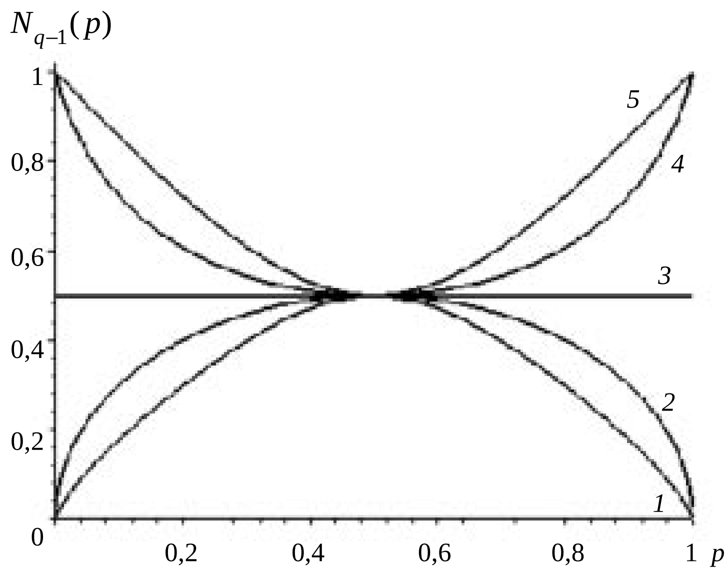


Рис. 3.2. Зависимость полунормы $N_{q-1}(p)$ от распределения:

$1-(q=-3)$, $2-(q=-1)$, $3-(q=0)$, $4-(q=1)$, $5-(q=3)$

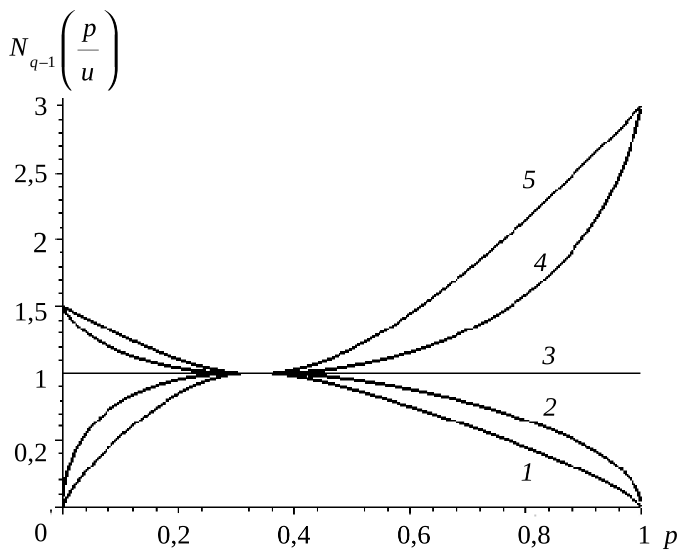


Рис. 3.3. Зависимость полунормы $N_{q-1}\left(\frac{p}{u}\right)$ от распределения:
1 – ($q = –3$), *2* – ($q = –1$), *3* – ($q = 0$), *4* – ($q = 1$), *5* – ($q = 3$).

Для распределения и случайной величины T имеет место функция

$$N_q^*(T) = \left(\frac{\sum_i^m T_i^q u_i}{\sum_i^m u_i} \right)^{1/q}, \qquad (3.3.17)$$

которая при вероятностной нормировке

$$\sum_i^m u_i = 1 \qquad (3.3.18)$$

запишется так

$$N_q^*(T) = \left(\sum_i^m T_i^q u_i \right)^{1/q}. \qquad (3.3.19)$$

При $T = u/p$ из (3.3.19) вытекает выражение полунормы распределения

$$N_q^*\left(\frac{u}{p}\right) = \left(\sum_i^m \left(\frac{u_i}{p_i}\right)^q u_i \right)^{1/q}. \qquad (3.3.20)$$

4. Взаимосвязь полунорм. Подставим в (3.3.14) распределение равновероятного состояния $u_i = 1/m$ и получим следующее равенство

$$N_q\left(\frac{p}{u}\right) = \frac{N_q(p)}{N_q(u)} = m^{\frac{q+1}{q}} N_q(p), \quad q \neq 0. \tag{3.3.21}$$

Для произвольных распределений оно не выполняется, за исключением случая $q = 0$, при котором полунормы становятся средними геометрическими распределений.

Взаимосвязь функционалов (3.3.14) и (3.3.20) определяется соотношениями

$$\begin{gathered} \left[N_q\left(\frac{p}{u}\right)\right]^q = \left[N^*_{-1-q}\left(\frac{u}{p}\right)\right]^{-1-q}, \\ \left[N_{q-1}\left(\frac{p}{u}\right)\right]^{1/q} = \left[N^*_{-q}\left(\frac{u}{p}\right)\right]^{1/(1-q)}. \end{gathered} \tag{3.3.22}$$

Наконец выпишем выражения полунорм в других важных случаях

$$N_q(p) = \left(\frac{\int_G p^{q+1} dX}{\int_G p dX}\right), \; N_q(\rho) = \left(\mathrm{Sp}\rho^{q+1}\right)^{1/q}, \tag{3.3.23}$$

$$N_q\left(\frac{p}{u}\right) = \left(\frac{\int_G \left(\frac{p}{u}\right)^{q+1} dX}{\int_G p dX}\right), \; N_q\left(\rho u^{-1}\right) = \left(\mathrm{Sp}\rho^{q+1} u^{-1}\right)^{1/q}. \tag{3.3.24}$$

3.4. Аксиомы и меры информации Реньи

А. Реньи [103, 104] для вывода энтропий упростил систему аксиом Фадеева (см. главу 1) и предложил следующие аксиомы.

1. $H(p_1, p_2) = H(1-p, p)$ непрерывна при $0 \le p \le 1$ и положительна хотя бы в одной точке.

2. $H(p_1, p_2, \ldots, p_m)$ симметрична относительно $p_1, p_2, \ldots, p_m$.

3. $H_q\left(\frac{1}{2},\frac{1}{2}\right)=1.$ (3.4.1)

4. Для независимых объектов с распределениями $p_1=\{p_{11},\dots,p_{1m}\}$ и $p_2=\{p_{21},\dots,p_{2m}\}$ имеем свойство аддитивности общей энтропии

$$H(p_{12})=H(p_1)+H(p_2). \tag{3.4.2}$$

5. Если $\sum_i^m p_{1i}+\sum_j^n p_{2j}\le 1$, то справедливо равенство

$$H(p_{11},\dots,p_{1m},p_{21},\dots,p_{2n})=$$
$$=\frac{\sum_i^m p_{1i}H(p_{11},\dots,p_{1m})+\sum_j^n p_{2j}H(p_{21},\dots,p_{2n})}{\sum_i^m p_{1i}+\sum_j^n p_{2j}} \tag{3.4.3}$$

для всех p_{1i} и p_{2j}.

Согласно аксиомам, находится выражение энтропии Шеннона–Винера

$$H(p)=-\frac{\sum_i^m(\log_2 p_i)p_i}{\sum_i^m p_i}. \tag{3.4.4}$$

Далее А. Реньи изменил аксиому 5 и сформулировал ее так:

$5'$. Если $\sum_i^m p_{1i}+\sum_j^n p_{2j}\le 1$, то справедливо равенство

$$H(p_{11},\dots,p_{1m},p_{21},\dots,p_{2n})=$$
$$=\varphi^{-1}\left\{\frac{\sum_i^m p_{1i}\varphi\left[H(p_{11},\dots,p_{1m})\right]+\sum_j^n p_{2j}\varphi\left[H(p_{21},\dots,p_{2n})\right]}{\sum_i^m p_{1i}+\sum_j^n p_{2j}}\right\}, \tag{3.4.5}$$

где φ есть функция Колмогорова–Нагумо. Принимая $\varphi(T)=aT+b$ с $a\ne 0$, постулат $5'$ преобразуется в постулат 5. Использовалась функция:

$$\varphi(T) = 2^{(q-1)T}, \quad T = \log_2 p, \tag{3.4.6}$$

что в итоге дает энтропию Реньи

$$H_q(p) = \frac{1}{1-q}\log_2\left(\frac{\sum_i^m p_i^q}{\sum_i^m p_i}\right) \quad (q \neq 1), \tag{3.4.7}$$

зависящую от параметра q.

В пределе $q \to 1$ функционал (3.4.7) равняется энтропии Шеннона–Винера

$$H(p) = \lim_{q\to 1} H_q(p) = -\frac{\sum_i^m (\log_2 p_i) p_i}{\sum_i^m p_i}. \tag{3.4.8}$$

Такой подход к функции $T = \log_2(p/u)$ позволил А. Реньи получить также информацию различия

$$I_q(p:u) = \frac{1}{q-1}\log_2\left(\frac{\sum_i^m p_i^q u_i^{1-q}}{\sum_i^m p_i}\right) \quad (q \neq 1), \tag{3.4.9}$$

которая удовлетворяет условию нормированности $I_q\left(1,0;\frac{1}{2},\frac{1}{2}\right) = 1$ (соответствующая аксиома 3). При $q = 1$ функционал (3.4.9) совпадает с информацией различия Кульбака–Лейблера

$$I(p:u) = \lim_{q\to 1} I_q(p:u) = \frac{\sum_i^m \left(\log_2 \frac{p_i}{u_i}\right) p_i}{\sum_i^m p_i}. \tag{3.4.10}$$

3.5. Параметризованное распределение. Энтропия Реньи

Рассмотрим экстремальные свойства энтропии Шеннона–Винера и информации различия Кульбака–Лейблера:

$$H(f) = -\sum_i^m (\log_2 f_i) f_i, \qquad (3.5.1)$$

$$I(f:p) = -\sum_i^m \left(\log_2 \frac{f_i}{p_i} \right) f_i \qquad (3.5.2)$$

при дополнительных условиях, чтобы экстремум достигался при сохранении меры неточности $H(f:p)$ с фиксированным распределением вероятностей $p = \{p_1, \ldots, p_m\}$ и нормировки для $f = \{f_1, \ldots, f_m\}$

$$H(f:p) = -\sum_i^m (\log_2 p_i) f_i, \quad \sum_i^m f_i = 1. \qquad (3.5.3)$$

Согласно вариационному принципу, находим безусловные экстремумы функционалов

$$L = -\sum_i^m (\log_2 f_i) f_i + q\sum_i^m (\log_2 p_i) f_i - \alpha \sum_i^m f_i, \qquad (3.5.4)$$

$$L = \sum_i^m \left(\log_2 \frac{f_i}{p_i} \right) f_i + (1-q)\sum_i^m (\log_2 p_i) f_i + \alpha \sum_i^m f_i, \qquad (3.5.5)$$

где $(1-q)$ и α есть лагранжевы множители.

Из условия равенства нулю первой вариации

$$\delta L = -\sum_i^m \delta f_i \left(\log_2 f_i + \frac{1}{\ln 2} - q \log_2 p_i + \alpha \right) = 0, \qquad (3.5.6)$$

$$\delta L = \sum_i^m \delta f_i \left(\log_2 \frac{f_i}{p_i} + \frac{1}{\ln 2} + (1-q) \log_2 p_i + \alpha \right) = 0 \qquad (3.5.7)$$

получим параметризованное распределение

$$f_i = p_i^q \Gamma_q^{-1}(p), \ \Gamma_q(p) = \sum p_i^q, \qquad (3.5.8)$$

которое использовалось в теории самоорганизации физических [18–21] и термодинамических хаотических систем [56].

Распределение (3.5.8) дает максимальное значение энтропии Шеннона–Винера и минимальное для информации различия Кульбака–Лейблера:

$$H(f)=qH(f:p)+\log_2 \Gamma_q(p), \tag{3.5.9}$$

$$I(f:p)=(1-q)H(f:p)-\log_2 \Gamma_q(p). \tag{3.5.10}$$

Добавим к (3.5.9) и (3.5.10) информацию различия и расхождение

$$I(p:f)=-(1-q)H(p)+\log_2 \Gamma_q(p), \tag{3.5.11}$$

$$J(f:p)=I(f:p)+I(p:f)=(1-q)\left[H(f:p)-H(p)\right]. \tag{3.5.12}$$

Взаимосвязь рассматриваемых функционалов определяется следующими соотношениями [20, 21]

$$\frac{q}{q-1}I(f:p)=-\left[H(f)-H_q(p)\right],\quad \frac{1}{q-1}I(p:f)=-\left[H_q(p)-H(p)\right], \tag{3.5.13}$$

$$\frac{q}{q-1}I(f:p)+\frac{1}{q-1}I(p:f)=-\left[H(f)-H(p)\right]. \tag{3.5.14}$$

Здесь введена энтропия Реньи

$$H_q(p)=\frac{1}{1-q}\log_2 \Gamma_q(p)=\frac{1}{1-q}\log_2 \sum_i^m p_i^q\,, \tag{3.5.15}$$

которая вместе с энтропией

$$H(f)=-\sum_i^m\left[\log_2\left(p_i^q\Big/\sum_i^m p_i^q\right)\right]p_i^q\Big/\sum_i^m p_i^q \tag{3.5.16}$$

и мерой неточности

$$H(f:p)=H(f)+I(f:p)=-\sum_i^m\left(\log_2 p_i\right)p_i^q\Big/\sum_i^m p_i^q \tag{3.5.17}$$

зависит от действительного параметра q, изменяющегося в области допустимых значений. Если пределы изменения q не обозначены, то полагаем, что $q\in\mathbf{R}$. В пределе $q\to 1$ имеем равенство функционалов (3.5.15)–(3.5.17) с энтропией Шеннона–Винера

$$H(p)=\lim_{q\to 1}\left[H_q(p),H(f),H(f:p)\right]=-\sum_i^m\left(\log_2 p_i\right)p_i\,. \tag{3.5.18}$$

Отметим, что мера неточности (3.5.17) и мера $2^{q-1}\sum_i^m p_i^q H(f:p)$

рассматривались, соответственно, в работах [52] и [112] как выражения, обобщающие меру Шеннона–Винера.

Рассмотрим основные свойства энтропии Реньи.

1. Положительность и выпуклость. Энтропия есть вещественный, неотрицательный и выпуклый функционал с максимумом (минимумом) при $q>0$ $(q<0)$. Справедливы неравенства

$$H_q(p)\geq 0, \tag{3.5.19}$$

$$H_q(a_1p_1+a_2p_2)\leq a_1H_q(p_1)+a_2H_q(p_2), \tag{3.5.20}$$

где $a_1+a_2=1$, $a_1>0$, $a_2>0$ и энтропии

$$H_q(p_1)=\frac{1}{1-q}\log_2\sum_i^m p_{1i}^q,\ H_q(p_2)=\frac{1}{1-q}\log_2\sum_i^m p_{2i}^q \tag{3.5.21}$$

с нормированными распределениями

$$\sum_i^m p_{1i}=\sum_i^m p_{2i}=1. \tag{3.5.22}$$

Неравенство (3.5.20) есть неравенство Иенсена в теории выпуклых функций [47]. При $q=0$ имеем $H_0(p)=\log_2 m$.

На рис. 3.4 представлена зависимость энтропии Реньи $H_q(p)$ от распределения p при значениях $m=2$, p_1p и $q=-3;\ -1;\ 0;\ 1;\ 3$.

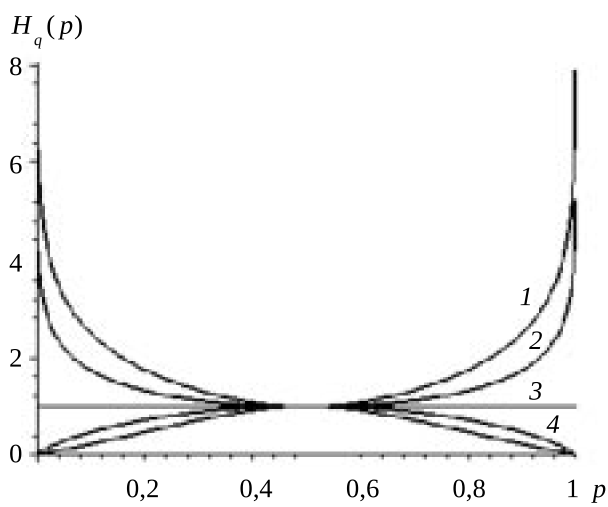


Рис. 3.4. Зависимость энтропии Реньи от распределения:
1 – (q = –3), *2* – (q = –1), *3* – (q = 0), *4* – (q = 1), *5* – (q = 3)

2. Аддитивность для независимых объектов. Пусть состояние случайного объекта описывается совместным мультипликативным распределением вероятностей $p_{ij} = p_i p_j$, p_i и p_j относятся к разным независимым объектам. Запишем общую энтропию

$$H_q\left(p_{12}\right)=\frac{1}{1-q}\log_2 \sum_i^m p_{ij}^q \tag{3.5.23}$$

с условиями нормировки для распределений

$$\sum_i^m \sum_j^n p_{ij} = \sum_i^m p_i = \sum_j^n p_j = 1. \tag{3.5.24}$$

Тогда из (3.5.23) получим свойство аддитивности для энтропий независимых объектов

$$H_q\left(p_{12}\right)=H_q\left(p_1\right)+H_q\left(p_2\right), \tag{3.5.25}$$

где

$$H_q\left(p_1\right)=\frac{1}{1-q}\log_2 \sum_i^m p_i^q,\ H_q\left(p_2\right)=\frac{1}{1-q}\log_2 \sum_j^n p_j^q. \tag{3.5.26}$$

3. Равенство для зависимых объектов. В общем случае зависимых объектов используется теорема разложения $p_{ij} = p_i p_{j|i}$. Рассмотрим два подхода к нахождению условия аддитивности для зависимых объектов. В первом случае исходим из соотношений для параметризованного распределения

$$f_{ij} = f_i f_{j|i}, \quad \sum_i^m \sum_j^n f_{ij} = \sum_i^m f_i = 1, \tag{3.5.27}$$

$$f_{ij} = \frac{p_i^q p_{j|i}^q}{\sum_i^m \sum_j^n p_i^q p_{j|i}^q}, \quad f_i = \frac{p_i^q}{\sum_i^m p_i^q}, \tag{3.5.28}$$

$$f_{j|i} = \frac{\sum_i^m p_i^q}{\sum_i^m \sum_j^n p_i^q p_{j|i}^q} p_{j|i}^q = \frac{p_{j|i}^q}{\sum_i^m f_i \sum_j^n p_{j|i}^q}. \tag{3.5.29}$$

Используем тождество для случайных энтропий $h\left(f_{ij}\right) - h\left(f_i\right) - h\left(f_{j|i}\right) = 0$ и получим равенство для зависимых объектов

$$H_q\left(p_{12}\right) = H_q\left(p_1\right) + H_q\left(p_2 \middle| p_1\right), \tag{3.5.30}$$

где функционал [110]

$$H_q\left(p_2 \middle| p_1\right) = \frac{1}{1-q} \log_2 \sum_i^m f_i \sum_j^n p_{j|i}^q = \frac{1}{1-q} \log_2 \frac{\sum_i^m \sum_j^n p_{ij}^q}{\sum_i^m p_i^q}. \tag{3.5.31}$$

Условие аддитивности (3.5.30) по форме совпадает с соответствующим условием (1.5.17) в статистической модели Шеннона–Винера.

Во втором случае введем параметризованное распределение

$$\varphi_{ij} = \varphi_i \varphi_{j|i}, \quad \sum_i^m \sum_j^n \varphi_{ij} = \sum_i^m \varphi_i = 1, \tag{3.5.32}$$

где

$$\varphi_{ij} = \frac{p_i^q p_{j|i}^q}{\sum_i^m p_i^q \sum_j^n p_{j|i}^q}, \quad \varphi_i = \frac{p_i^q}{\sum_i^m p_i^q}, \tag{3.5.33}$$

$$\varphi_{j|i} = \frac{\varphi_{ij}}{\varphi_i} = \frac{p_{j|i}^q}{\sum_j^n p_{j|i}^q}, \quad \sum_j^n \varphi_{j|i} = 1. \tag{3.5.34}$$

Запишем случайные энтропии

$$h\left(f_{ij}\right) = qh\left(p_{ij}\right) + (1-q) H_q\left(p_{12}\right), \tag{3.5.35}$$

$$h\left(\varphi_i\right) = qh\left(p_i\right) + (1-q) H_q\left(p_1\right), \tag{3.5.36}$$

$$h\left(\varphi_{j|i}\right)=qh\left(p_{j|i}\right)+(1-q)H_{qi}\left(p_{2|1}\right). \qquad (3.5.37)$$

В (3.5.37) имеем условную энтропию

$$H_{qi}\left(p_{2|1}\right)=\frac{1}{1-q}\log_2\sum_{j}^{n}p_{j|i}^{q}\ . \qquad (3.5.38)$$

После усреднения разности между случайными энтропиями $h\left(f_{ij}\right)-h\left(\varphi_i\right)-h\left(\varphi_{j|i}\right)$ получим условие аддитивности

$$H_q\left(p_{12}\right)=H_q\left(p_1\right)+H_q\left(p_2\left|p_1\right.\right)-\frac{1}{1-q}\left[H\left(\psi:f_{12}\right)-H\left(\psi:\varphi_{12}\right)\right]. \qquad (3.5.39)$$

Равенство содержит, в отличие от (3.5.30), дополнительное слагаемое, которое назовем дефектом энтропии. Указанный дефект энтропии зависит от разности мер неточности.

При $\psi_i=f_i$ и $\psi=p_i$ средние условные энтропии имеют следующий вид

$$H_q\left(p_2\left|p_1\right.\right)=\sum_{i}^{m}H_{qi}\left(p_2\left|p_1\right.\right)f_i=$$

$$=\frac{1}{1-q}\sum_{i}^{m}f_i\log_2\sum_{j}^{n}p_{j|i}^{q}=\frac{\sum_{i}^{m}p_i^q\log_2\sum_{j}^{n}p_{j|i}^{q}}{(1-q)\sum_{i}^{m}p_i^q}, \qquad (3.5.40)$$

$$H_q\left(p_2\left|p_1\right.\right)=\sum_{i}^{m}H_{qi}\left(p_2\left|p_1\right.\right)p_i=\frac{1}{1-q}\sum_{i}^{m}p_i\log_2\sum_{j}^{n}p_{j|i}^{q}\ . \qquad (3.5.41)$$

В пределе $q\to 1$ дефект энтропии равен нулю и (3.5.39) совпадает с формулой (1.5.17). Функционалы (3.5.31), (3.5.40) и (3.5.41) принимают значение средней условной энтропии (1.5.10) статистической модели Шеннона–Винера. Выражение (3.5.41) рассматривается в работе [59].

4. Случайное отклонение. Рассмотрим случайные энтропии $h\left(f_i\right)=-\log_2 f_i$ и $h\left(p_i\right)=-\log_2 p_i$. Логарифмируя распределение (3.5.8), имеем

$$h\left(f_i\right)-H_q(p)=q\left[h\left(p_i\right)-H_q(p)\right], \qquad (3.5.42)$$

где ясно проявляется линейная зависимость между случайными отклонениями величин $h\left(f_i\right)$ и $h\left(p_i\right)$ от энтропии Реньи. Энтропия Реньи не имеет аналога случайной энтропии.

В геометрической интерпретации равенство (3.5.42) представляет собой уравнение пучка прямых, проходящих через точку A плоскости $\left\{h\left(f_i\right), h\left(p_i\right)\right\}$. Параметр q есть угловой коэффициент рассматриваемой прямой $\left(q = \operatorname{tg}\alpha\right)$, а точка $A = \left\{H_q\left(p\right), H_q\left(p\right)\right\}$ – центр пучка.

Распределения можно записать в эквивалентных формах

$$p_i = f_i^{1/q} \Big/ \sum_i^m f_i^{1/q}, \quad f_i = p_i^q 2^{(q-1)H_q(p)}. \tag{3.5.43}$$

5. Энтропия равновероятного состояния. Находим экстремум энтропии Реньи при условии сохранения нормировки распределения p. Варьируем функционал

$$L = \frac{1}{1-q}\log_2 \sum_i^m p_i^q - \alpha \sum_i^m p_i \tag{3.5.44}$$

и из равенства $\delta L = 0$ получим $p_i = \mathrm{const}$. Из условия нормировки вытекает равновероятное распределение

$$p_i = \frac{1}{m}. \tag{3.5.45}$$

Экстремальное значение энтропии Реньи

$$H_q\left(p\right) = H\left(p\right) = \log_2 m \tag{3.5.46}$$

не зависит от параметра q и совпадает с соответствующим значением энтропии Шеннона–Винера (1.5.41).

6. Неравенства. Энтропия Реньи удовлетворяет следующим неравенствам:

$$H_q\left(p_{12}\right) \le H\left(p_1\right) + H\left(p_2\right), \tag{3.5.47}$$

$$H_q\left(p\right) \le \log_2 m \ , \quad \left(q > 0\right), \tag{3.5.48}$$

$$H_q\left(p\right) \ge \log_2 m \ , \quad \left(q < 0\right), \tag{3.5.49}$$

$$H(f) > H_q(p) > H(p)\ , \quad (0 < q < 1), \tag{3.5.50}$$

$$H(f) < H_q(p) < H(p)\ , \quad (q > 1), \tag{3.5.51}$$

$$H_q(p) > H(f) \text{ и } H_q(p) > H(p)\ , \quad (q < 0), \tag{3.5.52}$$

$$H_q(p_{12}) \ge H_q(p_1),\ H_q(p_{12}) \ge H_q(p_2), \tag{3.5.53}$$

$$H_q(p_2|p_1) \ge 0\,,\ H_q(p_1|p_2) \ge 0\,. \tag{3.5.54}$$

7. Нормированность и размерность. Выбираем наименьшее число возможных состояний $m = 2$ с равновозможными значениями распределения $p_1 = p_2 = 1/2$. Тогда для энтропии Реньи выполняется свойство нормированности

$$H_q\left(\frac{1}{2},\frac{1}{2}\right) = \frac{1}{1-q}\log_2 \sum_i^2 p_i^q = 1. \tag{3.5.55}$$

Единицей измерения информации в статистической модели Реньи, как следует из (3.5.55), является один бит.

Для физической теории информации имеем энтропию

$$H_q^{phys}(p) = \frac{1}{1-q}\ln \sum_i^m p_i^q\ . \tag{3.5.56}$$

В статистической физике используется энтропия Реньи, имеющая размерность постоянной Больцмана $H_q(p) = kH_q^{phys}(p)$ и совпадающая в пределе $q \to 1$ с энтропией Больцмана–Гиббса [1,6]

$$H(p) = \lim_{q\to 1} H_q^{phys}(p) = -k\sum_i^m (\ln p_i)\, p_i\ . \tag{3.5.57}$$

Энтропия Реньи есть отношение физической безразмерной энтропии (3.5.56) к ее значению при равновероятном состоянии с $m = 2$. Таким образом, имеет место равенство

$$H_q(p) = \frac{H_q^{phys}(p)}{H_q^{phys}\left(\frac{1}{2},\frac{1}{2}\right)}, \quad H_q^{phys}\left(\frac{1}{2},\frac{1}{2}\right) = \ln 2\,. \tag{3.5.58}$$

8. Дифференциальные соотношения. Дифференцируем информацию различия $I(f:p)$ в равенстве (3.5.13) и получим соотношение [20, 21]:

$$dI(f:p) = -\left(\frac{q-1}{q}\right) dH(f), \tag{3.5.59}$$

которое после преобразования Лежандра примет следующий вид

$$dH_q(p) = I(f:p)\, d\left(\frac{q}{q-1}\right). \tag{3.5.60}$$

Дифференциальные соотношения (3.5.59) и (3.5.60) определяют взаимосвязь информации различия Кульбака–Лейблера с энтропией Шеннона–Винера и энтропией Реньи, соответственно, в случае непрерывных значений параметра q.

9. ***f*** **-энтропия.** Функциональным обобщением энтропии Реньи (3.5.15) является единственное выражение

$$H_f(p) = f\left[N_{q-1}(p)\right], \tag{3.5.61}$$

где f – выпуклая функция от полунормы распределения.

При $f = -\log_2 N_{q-1}(p)$ из (3.5.61) вытекает энтропия Реньи

$$H_q(p) = -\log_2 N_{q-1}(p). \tag{3.5.62}$$

Здесь аргументом логарифмической функции является полунорма распределения

$$N_{q-1}(p) = \left(\sum_i^m p_i^q\right)^{1/(q-1)}. \tag{3.5.63}$$

3.6. Информация различия Реньи и мера неточности. Мера Чернова

Рассмотрим минимум информации различия Кульбака–Лейблера

$$I(f:p) = \sum_i^m \left(\log_2 \frac{f_i}{p_i}\right) f_i \tag{3.6.1}$$

при заданности разности мер неточности $H(f:p)$ и $H(f:u)$ с фиксированными распределениями вероятности $p = \{p_1, \ldots, p_m\}$, $u = \{u_1, \ldots, u_m\}$ и сохранении нормировки для $f = \{f_1, \ldots, f_m\}$:

$$H(f:p) = -\sum_i^m (\log_2 p_i) f_i,$$
$$H(f:u) = -\sum_i^m (\log_2 u_i) f_i \ , \quad \sum_i^m f_i = 1. \tag{3.6.2}$$

Согласно вариационному принципу, находим безусловный экстремум функционала

$$L = \sum_i^m \left(\log_2 \frac{f_i}{p_i} \right) f_i + (1-q) \left[\sum_i^m (\log_2 p_i) f_i - \right.$$
$$\left. - \sum_i^m (\log_2 u_i) f_i \right] + \alpha \sum_i^m f_i, \tag{3.6.3}$$

где $(1-q)$ и α есть множители Лагранжа. Используя равенство

$$\delta L = \sum_i^m \delta f_i \left(\log_2 \frac{f_i}{p_i} + \frac{1}{\ln 2} + (1-q) \log_2 \frac{p_i}{u_i} + \alpha \right) = 0, \tag{3.6.4}$$

получим параметризованное распределение

$$f_i = p_i^q u_i^{1-q} \Gamma_q^{-1}(p:u), \ \Gamma(p:u) = \sum p_i^q u_i^{1-q}, \tag{3.6.5}$$

рассматриваемое в теории самоорганизации физических систем [19 – 22] и теории информации [33].

Распределение (3.6.5) максимизирует энтропию Шеннона–Винера при вышеприведенных дополнительных условиях. Подставляя его в (3.5.1) и (3.6.1), имеем экстремальные значения энтропии и информации различия

$$H(f) = qH(f:p) + (1-q)H(f:u) + \log_2 \Gamma_q(p:u), \tag{3.6.6}$$

$$I(f:p) = (1-q)\left[H(f:p) - H(f:u)\right] - \log_2 \Gamma_q(p:u). \tag{3.6.7}$$

Аналогично находим следующие информацию различия и расхождение

$$I(p:f) = -(1-q)\left[H(p) - H(p:u)\right] + \log_2 \Gamma_q(p:u), \tag{3.6.8}$$

$$J(f:p) = I(f:p) + I(p:f) =$$
$$= (1-q)\left[H(f:p) - H(f:u) + H(p:u) - H(p)\right] \tag{3.6.9}$$

и с учетом (3.6.7) получим соотношения [20, 21]

$$\frac{q}{q-1} I(f:p) = I(f:u) - I_q(p:u), \tag{3.6.10}$$

$$\frac{1}{q-1} I(p:f) = I_q(p:u) - I(p:u), \tag{3.6.11}$$

$$\frac{q}{q-1} I(f:p) + \frac{1}{q-1} I(p:f) = I(f:u) - I(p:u) \tag{3.6.12}$$

для взаимосвязи рассматриваемых функционалов.

Здесь введена информация различия Реньи

$$I_q(p:u) = \frac{1}{q-1} \log_2 \Gamma_q(p:u) = \frac{1}{q-1} \log_2 \sum_i^m p_i^q u_i^{1-q} . \tag{3.6.13}$$

Не выписывая функционалы в явном виде, отметим, что в пределе $q \to 1$ имеем информацию различия Кульбака–Лейблера

$$I(p:u) = \lim_{q \to 1} \left[I_q(p:u), I(f:u), H(f:u) \right] = \sum_i^m \left(\log_2 \frac{p_i}{u_i} \right) p_i, \tag{3.6.14}$$

энтропию Шеннона–Винера

$$H(p) = \lim_{q \to 1} H(f:p) = -\sum_i^m \left(\log_2 p_i \right) p_i \tag{3.6.15}$$

и меру неточности Керриджа

$$H(p:u) = \lim_{q \to 1} H(f:u) = -\sum_i^m \left(\log_2 u_i \right) p_i . \tag{3.6.16}$$

Рассмотрим основные свойства информации различия Реньи.

1. Выпуклость. Информация различия есть вещественный, выпуклый и положительный (отрицательный) функционал с минимумом (максимумом) при $q > 0$ $(q < 0)$.

Имеют место следующие неравенства

$$I_q(p:u) > 0, \ (q > 0), \tag{3.6.17}$$

$$I_q(p:u) < 0, \ (q < 0), \tag{3.6.18}$$

$$I_q\left[\left(a_1 p_2 + a_2 p_2\right):u\right] \le a_1 I_q\left(p_1 : u\right) + a_2 I_q\left(p_2 : u\right), \qquad (3.6.19)$$

где $a_1 + a_2 = 1$, $a_1 > 0$, $a_2 > 0$ и информации различия

$$I_q\left(p_1 : u\right) = \frac{1}{q-1}\log_2 \sum_i^m p_{1i}^q u_i^{1-q}, \qquad (3.6.20)$$

$$I_q\left(p_2 : u\right) = \frac{1}{q-1}\log_2 \sum_i^m p_{2i}^q u_i^{1-q}. \qquad (3.6.21)$$

Равенства $q = 0$ или $p = u$ дают значение $I_q\left(p : u\right) = 0$.

На рис.3.5 приводятся зависимости энтропии и информации различия Реньи от числа q при значениях $m = 2$, $p_1 = 1/4$, $u_1 = 1/3$, где видно убывание $H_q\left(p\right)$ и возрастание $I_q\left(p : u\right)$ с увеличением q.

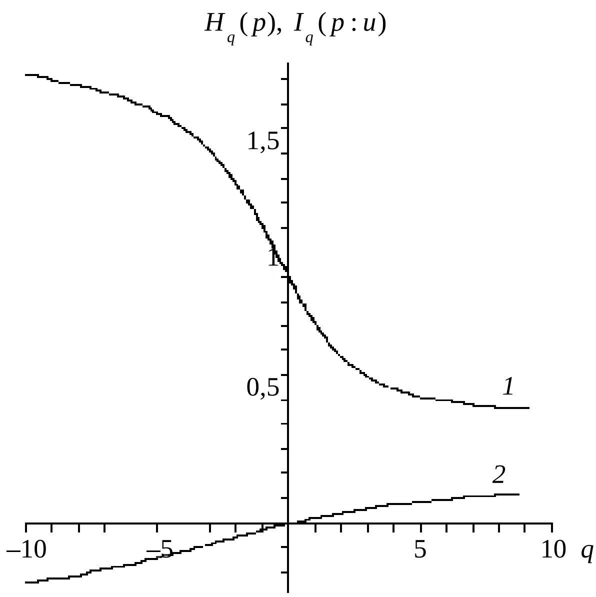


Рис. 3.5. Зависимости функционалов модели Реньи от числа q:
1 – $H_q\left(p\right)$, *2* – $I_q\left(p : u\right)$

На рис.3.6 представлена зависимость информации различия Реньи $I_q\left(p : u\right)$ от распределения при значениях $m = 2$, $p_1 = p$, $u_1 = 1/3$ и $q = -3;\ -1;\ 0;\ 1;\ 3$.

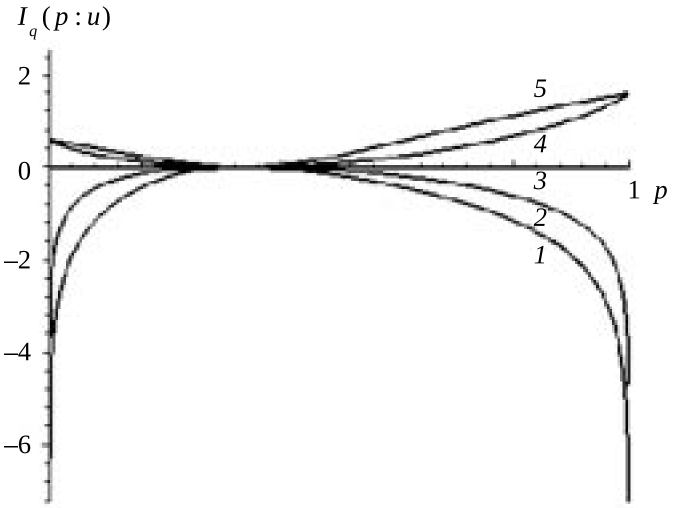


Рис. 3.6. Зависимость информации различия Реньи от распределения:
1 – (q = –3), *2* – (q = –1), *3* – (q = 0), *4* – (q = 1), *5* – (q = 3)

2. Аддитивность для независимых объектов. Пусть состояния случайного объекта описываются нормированными совместными распределениями p_{12} и u_{12}. Информация различия запишется так:

$$I_q\left(p_{12}:u_{12}\right)=\frac{1}{q-1}\log_2\sum_i^m\sum_j^n p_{ij}^q u_{ij}^{1-q}\ . \qquad (3.6.22)$$

В случае статистической независимости состояний имеем равенства $p_{ij}=p_i p_j$ и $u_{ij}=u_i u_j$. Тогда из (3.6.22) вытекает свойство аддитивности

$$I_q\left(p_{12}:u_{12}\right)=I_q\left(p_1:u_1\right)+I_q\left(p_2:u_2\right) \qquad (3.6.23)$$

для информации различия независимых объектов

$$I_q\left(p_1:u_1\right)=\frac{1}{q-1}\log_2\sum_i^m p_i^q u_i^{1-q}\ , \qquad (3.6.24)$$

$$I_q\left(p_2:u_2\right)=\frac{1}{q-1}\log_2\sum_j^n p_j^q u_j^{1-q}\ . \qquad (3.6.25)$$

3. Равенство для зависимых объектов. В общем случае зависимых объектов рассмотрим совместные распределения $p_{ij}=p_i p_{j|i}$ и $u_{ij}=u_i u_{j|i}$ при реализации состояния с p_i и u_i. Приведем два подхода к нахождению условия аддитивности для зависимых объектов. В первом случае запишем параметризованное распределение

$$f_{ij} = f_i f_{j|i}, \quad \sum_i^m \sum_j^n f_{ij} = \sum_i^m f_i = 1, \tag{3.6.26}$$

где

$$f_{ij} = \frac{p_{ij}^q u_{ij}^{1-q}}{\sum_i^m \sum_j^n p_{ij}^q u_{ij}^{1-q}}, \quad f_i = \frac{p_i^q u_i^{1-q}}{\sum_i^m p_i^q u_i^{1-q}}, \tag{3.6.27}$$

$$f_{j|i} = \frac{\sum_i^m \sum_j^n p_i^q u_i^{1-q}}{\sum_i^m \sum_j^n p_{ij}^q u_{ij}^{1-q}} p_{j|i}^q u_{j|i}^{1-q} = \frac{p_{j|i}^q u_{j|i}^{1-q}}{\sum_i^m f_i \sum_j^n p_{j|i}^q u_{j|i}^{1-q}}. \tag{3.6.28}$$

Используем значения случайных информаций различия и после вычислений имеем равенство для зависимых объектов

$$I_q\left(p_{12} : u_{12}\right) = I_q\left(p_1 : u_1\right) + I_q\left(p_{2|1} : u_{2|1}\right), \tag{3.6.29}$$

где функционал

$$I_q\left(p_{2|1} : u_{2|1}\right) = \frac{1}{q-1} \log_2 \sum_i^m f_i \sum_j^n p_{j|i}^q u_{j|i}^{1-q} =$$

$$= \frac{1}{q-1} \log_2 \frac{\sum_i^m \sum_j^n p_{ij}^q u_{ij}^{1-q}}{\sum_i^m p_i^q u_i^{1-q}}. \tag{3.6.30}$$

Условие аддитивности для зависимых объектов (3.6.29) по форме совпадает с соответствующим условием в статистической модели Шеннона–Винера.

Во втором случае введем равенства для другого параметризованного распределения

$$\varphi_{ij} = \frac{p_{ij}^q u_{ij}^{1-q}}{\sum_i^m p_i^q u_i^{1-q} \sum_j^n p_{j|i}^q u_{j|i}^{1-q}}, \quad \varphi_i = \frac{p_i^q u_i^{1-q}}{\sum_i^m p_i^q u_i^{1-q}}, \tag{3.6.31}$$

$$\varphi_{j|i} = \frac{\varphi_{ij}}{\varphi_i} = \frac{p_{j|i}^q u_{j|i}^{1-q}}{\sum_j^n p_{j|i}^q u_{j|i}^{1-q}}, \quad \sum_j^n \varphi_{j|i} = 1, \tag{3.6.32}$$

случайные информации различия

$$I\left(f_{ij} : u_{ij}\right) = qI\left(p_{ij} : u_{ij}\right) - (q-1)I_q\left(p_{12} : u_{12}\right), \tag{3.6.33}$$

$$I\left(\varphi_i : u_i\right) = qI\left(p_i : u_i\right) - (q-1)I_q\left(p_1 : u_1\right), \tag{3.6.34}$$

$$I\left(\varphi_{j|i} : u_{j|i}\right) = qI\left(p_{j|i} : u_{j|i}\right) - (q-1)I_{qi}\left(p_{2|1} \middle| u_{2|1}\right). \tag{3.6.35}$$

и условную информацию различия

$$I_{qi}\left(p_{2|1} \middle| u_{2|1}\right) = \frac{1}{q-1} \log_2 \sum_j^n p_{j|i}^q u_{j|i}^{1-q} . \tag{3.6.36}$$

После усреднения разности между случайными информациями различия $I\left(f_{ij} : u_{ij}\right) - I\left(\varphi_i : u_i\right) - I\left(\varphi_{j|i} : u_{j|i}\right)$ получим условие аддитивности

$$I_q\left(p_{12} : u_{12}\right) = I_q\left(p_1 : u_1\right) + I_q\left(p_{2|1} \middle| u_{2|1}\right) - $$
$$- \frac{1}{q-1}\left[I\left(\psi : f_{12}\right) - I\left(\psi : \varphi_{12}\right)\right] . \tag{3.6.37}$$

Здесь дефект информации различия зависит от разности информации различия.

При $\psi_i = f_i$ и $\psi = p_i$ средние условные информации различия имеют следующий вид

$$I_q\left(p_{2|1} \middle| u_{2|1}\right) = \sum_i^m I_{qi}\left(p_{2|1} \middle| u_{2|1}\right) f_i =$$
$$= \frac{1}{q-1} \sum_i^m f_i \log_2 \sum_j^n p_{j|i}^q u_{j|i}^{1-q} = \frac{\sum_i^m p_i^q u_i^{1-q} \log_2 \sum_j^n p_{j|i}^q u_{j|i}^{1-q}}{(q-1)\sum_i^m p_i^q u_i^{1-q}}, \tag{3.6.38}$$

$$I_q\left(p_{2|1}\middle|u_{2|1}\right)=\sum_i^m I_{qi}\left(p_{2|1}\middle|u_{2|1}\right)p_i=$$

$$=\frac{1}{q-1}\sum_i^m p_i\log_2\sum_j^n p_{j|i}^q u_{j|i}^{1-q}. \tag{3.6.39}$$

В пределе $q\to 1$ дефект информации различия равен нулю и условие аддитивности (3.6.37) совпадает с формулой (1.8.23). Функционалы (3.6.30), (3.6.38) и (3.6.39) принимают значение средней условной информации различия (1.8.25) статистической модели Шеннона–Винера.

4. Случайное отклонение. Рассмотрим случайные информации различия $I(f_i:u_i)=-\left[h(f_i)-h(u_i)\right]=\log_2\left(f_i/u_i\right)$ и $I(p_i:u_i)=-\left[h(p_i)-h(u_i)\right]=$ $=\log_2\left(p_i/u_i\right)$, которые равняются разностям случайных энтропий. После логарифмирования распределения (3.6.5) получим связь случайных отклонений этих величин от информации различия Реньи

$$I(f_i:u_i)-I_q(p:u)=q\left[I(p_i:u_i)-I_q(p:u)\right], \tag{3.6.40}$$

которая не имеет аналога случайной информации различия.

Равенство (3.6.40) есть уравнение пучка прямых в плоскости $\left\{I(f_i:u_i),I(p_i:u_i)\right\}$ с центром в точке $A=\left\{I_q(p:u),I_q(p:u)\right\}$.

Распределение (3.6.5) можно представить в эквивалентной форме

$$f_i=p_i^q u_i^{1-q}2^{-(q-1)I_q(p:u)}. \tag{3.6.41}$$

5. Информация различия с $\boldsymbol{u_i=1/m}$. Подставим равновероятное распределение $u_i=1/m$ в (3.6.13) и получим равенство

$$I_q(p:u)=\frac{1}{q-1}\log_2\sum_i^m p_i^q\left(\frac{1}{m}\right)^{1-q}=-\left[H_q(p)-\log_2 m\right]. \tag{3.6.42}$$

Из условий выпуклости информации различия Реньи, определяемых неравенствами (3.6.17) и (3.6.18), следует, что энтропия меньше (больше) энтропии равновероятного состояния при $q>0$ $(q<0)$.

6. Неравенства. Для информации различия имеем следующие неравенства:

$$I_q\left(p_{12}:u_{12}\right)\le I_q\left(p_1:u_1\right)+I_q\left(p_2:u_2\right), \qquad (3.6.43)$$

$$I\left(f:u\right)>I_q\left(p:u\right)>I\left(p:u\right),\ \left(0<q<1\right), \qquad (3.6.44)$$

$$I\left(f:u\right)<I_q\left(p:u\right)<I\left(p:u\right),\ \left(q>0\right), \qquad (3.6.45)$$

$$I_q\left(p:u\right)>I\left(f:u\right) \text{ и } I_q\left(p:u\right)>I\left(p:u\right),\ \left(q<0\right). \qquad (3.6.46)$$

7. Расхождение. Определим количественную меру информации различия в наблюдениях u относительно p

$$I_q\left(u:p\right)=\frac{1}{q-1}\log_2\sum_i^m u_i^q p_i^{1-q} \qquad (3.6.47)$$

и рассмотрим сумму

$$J_q\left(p:u\right)=I_q\left(p:u\right)+I_q\left(u:p\right)=$$
$$=\frac{1}{q-1}\left(\log_2\sum_i^m p_i^q u_i^{1-q}+\log_2\sum_i^m u_i^q p_i^{1-q}\right), \qquad (3.6.48)$$

которая является расхождением. Расхождение есть симметричная функция $J_q\left(p:u\right)=J_q\left(u:p\right)$ относительно распределений p и u. В пределе $q\to 1$ из (3.6.48) следует известная мера

$$J\left(p:u\right)=\lim_{q\to 1}J_q\left(p:u\right)=\sum_i^m\left(\log_2\frac{p_i}{u_i}\right)\left(p_i-u_i\right). \qquad (3.6.49)$$

Функционал (3.6.48) является выпуклым и аддитивным для независимых объектов.

В работе [117] рассматривается расхождение в следующем виде

$$J_q\left(p:u\right)=\frac{2}{q-1}\left(\log_2\frac{\sum_i^m p_i^q u_i^{1-q}+\sum_i^m u_i^q p_i^{1-q}}{2}\right), \qquad (3.6.50)$$

которое при $q\to 1$ также имеет предел (3.6.49). Для сравнения функционалов (3.6.48) и (3.6.50) учитываем неравенство:

$$\frac{1}{2}\left(\log_2\sum_i^m p_i^q u_i^{1-q}+\log_2\sum_i^m u_i^q p_i^{1-q}\right)\le$$

$$\le\log_2\left[\sum_i^m\frac{p_i^q u_i^{1-q}+u_i^q p_i^{1-q}}{2}\right], \tag{3.6.51}$$

которое вытекает из свойства выпуклости логарифмической функции.

После деления (3.6.51) на $(q-1)$ получим следствие, что функционал (3.6.48) меньше (больше), чем функционал (3.6.32) при $q>0$ $(0<q<1)$.

8. Мера неточности. Мера статистической неточности определяется функционалом при аддитивности мер

$$H_q(p:u)=H_q(p)+I_q(p:u)=$$

$$=-\log_2\frac{N_{q-1}(p)}{N_{q-1}\left(\frac{p}{u}\right)}=\frac{1}{1-q}\log_2\frac{\sum_i^m p_i^q}{\sum_i^m\left(\frac{p_i}{u_i}\right)^{q-1}p_i}, \tag{3.6.52}$$

который был введен в работе [96]. Другие выражения для меры неточности

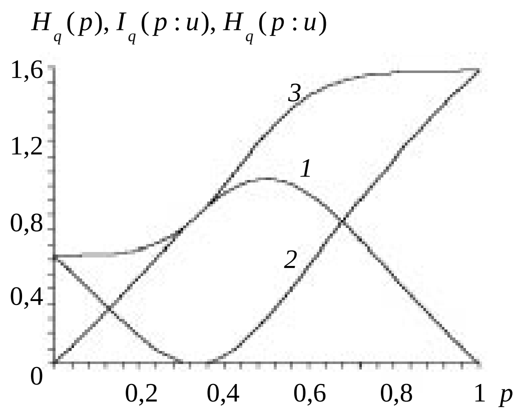
$$H_q(p:u)=-\log_2 N_{q-1}(u)=\frac{1}{1-q}\log_2\sum_i^m u_i^{q-1}p_i, \tag{3.6.53}$$

$$H_q(p:u)=-\log_2\left[N_{q-1}\left(\frac{1}{u^q}\right)\right]^q=\frac{q}{1-q}\log_2\sum_i^m u_i^{(q-1)/q}p_1, \tag{3.6.54}$$

не являющиеся суммой энтропии и информации различия, приводятся, соответственно, в работах [95] и [126]. В пределе $q\to 1$ из (3.6.52)-(3.6.54) вытекает мера неточности Керриджа

$$H(p:u)=\lim_{q\to 1}H_q(p:u)=-\sum_i^m\left(\log_2 u_i\right)p_i\,. \tag{3.6.55}$$

На рис. 3.7 представлены зависимости энтропии $H_q(p)$, информации различия $I_q(p:u)$ и меры неточности $H_q(p:u)$ в виде (3.6.52) от распределения при $m=2$, $q=2$, $p_1=p$ и а) $u_1=1/3$, б) $u_1=1/2$.



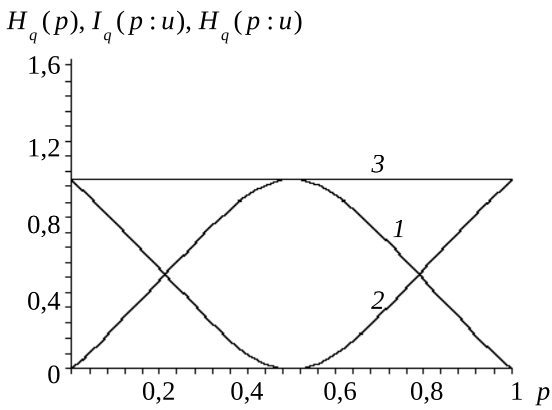


Рис. 3.7. Зависимости функционалов модели Реньи от распределения:
1 – энтропия $H_q(p)$, *2* – информация различия $I_q(p:u)$,
3 – мера неточности $H_q(p:u)$

9. Информационный радиус. Геометрическая интерпретация неравенства (3.5.20) при $a_1 = a_2 = 1/2$ дает формулу для информационного радиуса [99]

$$R_q(p:u) = H_q\left(\frac{p+u}{2}\right) - \frac{1}{2}\left[H_q(p) + H_q(u)\right] =$$

$$= \frac{1}{2}\left[I_q\left(p:\frac{p+u}{2}\right) + I_q\left(u:\frac{p+u}{2}\right)\right] =$$

$$= \frac{1}{1-q}\log_2\left[\sum_i^m\left(\frac{p_i+u_i}{2}\right)^q \Big/ \sqrt{\left(\sum_i^m p_i^q\right)\left(\sum_i^m u_i^q\right)}\right], \tag{3.6.56}$$

которая при $q \to 1$ совпадает с функционалом (1.8.60) статистической теории Шеннона–Винера. Другие выражения для $R_q(p:u)$ приводятся в обзорах [120, 121].

10. Мера Чернова. Информация различия Реньи включает в себя меру Чернова [67]

$$\Gamma_q(p:u) = \sum_i^m p_i^q u_i^{1-q} \tag{3.6.57}$$

для различия распределений p и u. В интервале $0 \le q \le 1$ функция $\Gamma_q(p:u)$ является аналитической, конечной и неотрицательной. На концах интервала имеем равенства $f = u$ $(f = p)$ при $q = 0$ $(q = 1)$ и $\Gamma_0(p:u) = \Gamma_1(p:u) = 1$.

11. Информационное расстояние. Для геометрической интерпретации информации различия в виде половины несимметричного расстояния $\delta(p,u) = \frac{1}{2} I_q^2(p:u)$ от p до u не выполняется неравенство треугольника. Имеет место несимметричный аналог теоремы Пифагора

$$I_q(p:u) = I_q(p:w) + I_q(w:u), \tag{3.6.58}$$

где равенство достигается тогда и только тогда, когда справедливо условие

$$\sum_i^m p_i^q u_i^{1-q} = \sum_i^m p_i^q w_i^{1-q} \sum_i^m w_i^q u_i^{1-q} . \tag{3.6.59}$$

При $q = 1$ из (3.6.58) и (3.6.59) следуют соотношения (1.8.68) и (1.8.69) в статистической модели Шеннона–Винера.

12. Нормированность и размерность. Информация различия Реньи удовлетворяет условию нормированности

$$I_q\left(1,0 : \frac{1}{2}, \frac{1}{2}\right) = \frac{1}{q-1} \log_2 \sum_i^2 p_i^q u_i^{1-q} = 1, \tag{3.6.60}$$

а также равняется отношению физической безразмерной различающей информации на значение физической энтропии $H_q^{phys}(u)$ при равновероятном состоянии с $m = 2$. Выполняется следующее равенство

$$I_q(p:u) = \frac{I_q^{phys}(p:u)}{H_q^{phys}\left(\frac{1}{2}, \frac{1}{2}\right)}, \tag{3.6.61}$$

где

$$I_q^{phys}(p:u) = \frac{1}{q-1} \ln \sum_i^m p_i^q u_i^{1-q} , \; H_q^{phys}(u) = \frac{1}{1-q} \ln \sum_i^2 u_i^q , \tag{3.6.62}$$

$$H_q^{phys}\left(\frac{1}{2},\frac{1}{2}\right)=I_q^{phys}\left(1,0:\frac{1}{2},\frac{1}{2}\right)=\ln 2\,. \qquad (3.6.63)$$

В статистической физике используется физическая размерная информация различия $I_q\left(p:u\right)=kI_q^{phys}\left(p:u\right)$ [23].

13. Дифференциальные соотношения. Дифференцируя информацию различия $I\left(f:p\right)$ в равенстве (3.6.10), получим [20, 21]:

$$dI\left(f:u\right)=\frac{q}{q-1}dI\left(f:p\right). \qquad (3.6.64)$$

Используем преобразование Лежандра и из (3.6.64) имеем дифференциальное соотношение

$$dI_q\left(p:u\right)=-I\left(f:p\right)d\left(\frac{q}{q-1}\right) \qquad (3.6.65)$$

взаимосвязи информации различия Реньи с информацией различия Кульбака–Лейблера для непрерывных значений параметра q.

14. Информация Фишера. Рассмотрим параметрические распределения $p_i\left(\theta\right)$ и $p_i\left(\theta+\delta\theta\right)$, соответствующие малому изменению одномерного параметра $\delta\theta$. Предельные значения меры Чернова и информации различия Реньи имеют вид

$$\Gamma_q\left(\theta:\theta+\delta\theta\right)=\sum_i^m p_i^q\left(\theta\right)p_i^{1-q}\left(\theta+\delta\theta\right)=$$

$$=1+\frac{q\left(q-1\right)}{2\ln 2}\Gamma_{\theta\theta}\left(\delta\theta\right)^2, \qquad (3.6.66)$$

$$I_q\left(\theta:\theta+\delta\theta\right)=\frac{1}{q-1}\log_2\sum_i^m p_i^q\left(\theta\right)p_i^{1-q}\left(\theta+\delta\theta\right)=$$

$$=\frac{q}{2\ln 2}\Gamma_{\theta\theta}\left(\delta\theta\right)^2, \qquad (3.6.67)$$

где величина

$$\Gamma_{\theta\theta}=\sum_i^m\left[\frac{\partial\ln p_i\left(\theta\right)}{\partial\theta}\right]^2 p_i\left(\theta\right) \qquad (3.6.68)$$

есть мера информации Фишера о величине нефлуктуирующего параметра θ в теории оценивания математической статистики [32, 33, 39, 74, 75].

Различные аспекты вопроса об информации Фишера можно найти, например, в работах [106, 119].

15. f-информация различия. Функциональным обобщением информации различия Реньи

$$I_q(p:u) = \log_2 N_{q-1}\left(\frac{p}{u}\right) \tag{3.6.69}$$

определим выражение

$$I_f(p:u) = f\left[N_{q-1}\left(\frac{p}{u}\right)\right], \tag{3.6.70}$$

которое представляет собой функцию полунормы распределения.

В случае трех распределений имеем полунорму

$$N_{q-1}\left(\frac{w}{u}\right) = \left[\sum_i^m \left(\frac{w_i}{u_i}\right)^{q-1} p_i\right]^{1/(q-1)} \tag{3.6.71}$$

и f-информацию различия

$$I_f(p:w:u) = f\left[N_{q-1}\left(\frac{w}{u}\right)\right], \tag{3.6.72}$$

обобщающую функционал (3.6.70). Если f есть логарифмическая функция, то из (3.6.72) получим следующий функционал [95]

$$I_f(p:w:u) = \frac{1}{q-1}\log_2 \sum_i^m \left(\frac{w_i}{u_i}\right)^{q-1} p_i\ . \tag{3.6.73}$$

При $w = p$ из (3.6.72) и (3.6.73) следуют, соответственно, f-информация различия (3.6.70) и информация различия Реньи (3.6.69).

3.7. Обобщенные полунормы и меры

Рассмотрим основополагающие обобщения полунорм произвольной случайной величины

$$N_q(T) = \left(\frac{\sum_i^m T_i^q p_i}{\sum_i^m p_i}\right)^{1/q} \tag{3.7.1}$$

и распределения

$$N_q(p)=\left(\frac{\sum_i^m p_i^{q+1}}{\sum_i^m p_i}\right)^{1/q}, \tag{3.7.2}$$

которые определяются способом усреднения.

Нормированная *f*-полунорма. Функциональным обобщением являются взвешенные нормированные f-полунормы

$$N_{q,f}(T)=\left(\frac{\sum_i^m T_i^q f(p_i)}{\sum_i^m f(p_i)}\right)^{1/q}, \tag{3.7.3}$$

$$N_{q,f}(p)=\left(\frac{\sum_i^m p_i^q f(p_i)}{\sum_i^m f(p_i)}\right)^{1/q}, \tag{3.7.4}$$

где функция от распределения обладает свойством мультипликативности $f(p_i p_j)=f(p_i)f(p_j)$.

Энтропия и информация различия представляются выражениями

$$H_{q,f}(p)=-\log_2 N_{q-1,f}(p)=\frac{1}{1-q}\log_2\left[\frac{\sum_i^m p_i^{q-1} f(p_i)}{\sum_i^m f(p_i)}\right], \tag{3.7.5}$$

$$I_{q,f}(p:u)=\log_2 N_{q-1,f}\left(\frac{p}{u}\right)=\frac{1}{q-1}\log_2\left[\frac{\sum_i^m p_i^{q-1}u_i^{1-q} f(p_i)}{\sum_i^m f(p_i)}\right], \tag{3.7.6}$$

которые имеют предельные значения:

$$H_f(p)=\lim_{q\to 1} H_q(p)=-\frac{\sum_i^m (\log_2 p_i)\, f(p_i)}{\sum_i^m f(p_i)}, \tag{3.7.7}$$

$$I_f(p:u)=\lim_{q\to 1} I_q(p:u)=\frac{\sum_i^m \left(\log_2 \frac{p_i}{u_i}\right) f(p_i)}{\sum_i^m f(p_i)}, \tag{3.7.8}$$

обобщающие энтропию Шеннона–Винера и информацию различия Кульбака–Лейблера.

В качестве примера приведем случай усреднения при помощи распределения

$$f(p_i)=\frac{p_i^{s_i}}{\sum_i^m p_i^{s_i}}, \tag{3.7.9}$$

что дает, согласно (3.7.5), следующую энтропию [100]

$$H_q(p)=\frac{1}{1-q}\log_2\left[\frac{\sum_i^m p_i^{q+s_i-1}}{\sum_i^m p_i^{s_i}}\right], \tag{3.7.10}$$

где $s_i \geq 1$, $q \neq 1$ и $q > 0$.

При $s_i = s$ из (3.7.10) вытекает функционал

$$H_q^s(p)=\frac{1}{1-q}\log_2\left[\frac{\sum_i^m p_i^{q+s-1}}{\sum_i^m p_i^{s}}\right] \tag{3.7.11}$$

с $s \geq 1$, $q \neq 1$ и $q > 0$, который изучался в работах [52, 83].

Мера статистического расстояния. Если в (3.7.11) положить $1-q=s-r$, а при $s=1$ сделать замену q на $q-r+1$, то получим соответствующие энтропии:

$$H_r^s(p) = -\log_2 T_r^s(p) = \frac{1}{s-r}\log_2\left(\frac{\sum_i^m p_i^r}{\sum_i^m p_i^s}\right) \qquad (3.7.12)$$

с $r \neq q$, $r > 0$ и $s > 0$ [52] и

$$H_q^r(p) = \frac{1}{r-q}\log_2 \sum_i^m p_i^{q-r+1} \qquad (3.7.13)$$

с $r-1 < q < r$, $r \geq 1$ [128]. Функционал (3.7.12) является обобщенной мерой статистического расстояния $T_r^s(p)$ между p^r и p^s [63, 66].

$(\boldsymbol{h}, \boldsymbol{\varphi})$**-энтропия.** Определим для каждой случайной величины T и любых чисел q и s функцию

$$G_q^s(T) = \left(\frac{\sum_i^m T_i^q p_i}{\sum_i^m p_i}\right)^s, \qquad (3.7.14)$$

которая при выполнении вероятностной нормировки имеет вид

$$G_q^s(T) = \left(\sum_i^m T_i^q p_i\right)^s. \qquad (3.7.15)$$

Функция (3.7.14) является взвешенным средним по Колмогорову–Нагумо при $s = 1/q$.

Выражение (3.7.15) для распределения

$$G_{q-1}^s(p) = \left(\sum_i^m p_i^q\right)^s \qquad (3.7.16)$$

и его свойства рассматривались в работах [60, 66, 125, 127], а соответствующая энтропия

$$H_q^s(p) = -\log_2 G_{q-1}^s(p) = -s\log_2 \sum_i^m p_i^q \qquad (3.7.17)$$

представляет собой (h, ϕ)-энтропию [105]:

$$H_{\phi}^{h}(p)=h\left[\sum_{i}^{m}\phi(p_i)\right] \tag{3.7.18}$$

при функциях степенной $\phi(p)=p^q$ и логарифмической $h(G)=-\log_2 G$.

Рассмотренные двухпараметрические энтропии и информация различия имеют логарифмическую меру и исследовались в соответствующих работах как обобщающие выражения для энтропии и информации различия Реньи.

3.8. Двух- и *k*-параметрическая информация различия. Мера Хаусдорфа

Получим двухпараметрический аналог информации различия Реньи. Для чего рассмотрим минимум информации различия Кульбака-Лейблера

$$I(f:p)=\sum_{i}^{m}\left(\log_2\frac{f_i}{p_i}\right)f_i \tag{3.8.1}$$

при фиксированных распределениях $p=\{p_1,\ldots,p_m\}$ и $u=\{u_1,\ldots,u_m\}$ с условием сохранения нормировки для $f=\{f_1,\ldots,f_m\}$ и двух мер неточности

$$H(f:p)=-\sum_{i}^{m}(\log_2 p_i)f_i\,,\quad H(f:u)=-\sum_{i}^{m}(\log_2 u_i)f_i\,. \tag{3.8.2}$$

Введем неопределенные множители Лагранжа $(1-q)$, τ, α и, согласно вариационному принципу, исследуем безусловный экстремум функционала [21]

$$L=\sum_{i}^{m}\left(\log_2\frac{f_i}{p_i}\right)p_i+(1-q)\sum_{i}^{m}(\log_2 p_i)f_i+$$
$$+\tau\sum_{i}^{m}(\log_2 u_i)f_i+\alpha\sum_{i}^{m}f_i. \tag{3.8.3}$$

Варьируя функционал и используя равенство

$$\delta L=\sum_{i}^{m}\delta f_i\left[\log_2\frac{f_i}{p_i}+\frac{1}{\ln 2}+(1-q)\log_2 p_i+\right.$$
$$\left.+\tau\log_2 u_i+\alpha\right]=0, \tag{3.8.4}$$

получим нормированное распределение

$$f_i = p_i^q u_i^{-\tau} \Gamma^{-1}(q,\tau). \tag{3.8.5}$$

Здесь определена обобщенная мера

$$\Gamma(q,\tau) = \sum_i^m p_i^q u_i^{-\tau} = \sum_i^m p_i \left(\frac{p_i}{u_i^{D_q}} \right)^{q-1} \tag{3.8.6}$$

и величина

$$D_q = \frac{\tau}{q-1}, \tag{3.8.7}$$

в которых q и τ есть действительные числа, меняющиеся в пределах допустимых значений.

Подставим распределение (3.8.5) в выражения информаций различия и расхождения

$$I(f:p) = \sum_i^m \left(\log_2 \frac{f_i}{p_i} \right) f_i =$$

$$= (q-1) \sum_i^m \left(\log_2 \frac{p_i}{u_i^{D_i}} \right) f_i - \log_2 \Gamma(q,\tau), \tag{3.8.8}$$

$$I(p:f) = \sum_i^m \left(\log_2 \frac{p_i}{f_i} \right) p_i =$$

$$= (q-1) \sum_i^m \left(\log_2 \frac{p_i}{u_i^{D_i}} \right) p_i + \log_2 \Gamma(q,\tau), \tag{3.8.9}$$

$$J(f:p) = I(f:p) + I(p:f) =$$

$$= (q-1) \sum_i^m \left(\log_2 \frac{p_i}{u_i^{D_i}} \right) (f_i - p_i) \tag{3.8.10}$$

и после преобразований имеем следующие соотношения [21]

$$\frac{q}{q-1} I(f:p) = I_{D_q}(f:u) - I_{q,\tau}(p:u), \tag{3.8.11}$$

$$\frac{1}{q-1}I(p:f)=I_{q,\tau}(p:u)-I_{D_q}(p:u), \tag{3.8.12}$$

$$\frac{q}{q-1}I(f:p)+\frac{1}{q-1}I(p:f)=I_{D_q}(f:u)-I_{D_q}(p:u). \tag{3.8.13}$$

В (3.8.11) – (3.8.13) введены нестандартные информации различия

$$I_{D_q}(f:u)=-\sum_i^m\left[h(f_i)-D_qh(u_i)\right]f_i=$$

$$=\sum_i^m\left(\log_2\frac{f_i}{u_i^{D_q}}\right)f_i=-\left[H(f)-D_qH(f:u)\right], \tag{3.8.14}$$

$$I_{D_q}(p:u)=-\sum_i^m\left[h(p_i)-D_qh(u_i)\right]p_i=$$

$$=\sum_i^m\left(\log_2\frac{p_i}{u_i^{D_q}}\right)p_i=-\left[H(p)-D_qH(p:u)\right] \tag{3.8.15}$$

и двухпараметрический аналог информации различия Реньи

$$I_{q,D_q}(p:u)=\frac{1}{q-1}\log_2\sum_i^m p_i^q u_i^{D_q(1-q)}=$$

$$=\frac{1}{q-1}\log_2\sum_i^m\left(\frac{p_i}{u_i^{D_q}}\right)^{q-1}p_i=\log_2 N_{q-1}\left(\frac{p}{u^{D_q}}\right). \tag{3.8.16}$$

При $D_q=1$ из (3.8.14)–(3.8.16) вытекает информация различия Кульбака-Лейблера и Реньи

$$\begin{gathered}I(p:u)=I_1(p:u)=\sum_i^m\left(\log_2\frac{p_i}{u_i}\right)p_i,\\ I_q(p:u)=I_{q,1}(p:u)=\frac{1}{q-1}\log_2\sum_i^m p_i^q u_i^{1-q}.\end{gathered} \tag{3.8.17}$$

Не рассматривая все свойства функционалов, отметим лишь некоторые из них.

1. Выпуклость и аддитивность. Информация различия (3.8.16) есть выпуклый, аддитивный функционал для независимых объектов:

$$I_{q,D_q}(p_{12}:u_{12}) = I_{q,D_q}(p_1:u_1) + I_{q,D_q}(p_2:u_2). \tag{3.8.18}$$

При $D_q > 1$ и $q > 0$ ($D_q < 1$ и $q < 0$) он имеет положительное (отрицательное) значение. Если $q = 0$, то справедливы равенства

$$I_{0,1}(p:u) = 0, \quad I_{0,D_q}(p:u) = -\log_2 \sum_i^m u_i^{D_q}, \tag{3.8.19}$$

а при $p = u$ имеем ненулевые значения

$$I_{q,D_q}(p:p) = \frac{1}{q-1}\log_2 \sum_i^m p_i^{1+(1-D_q)(q-1)} = \\ = -(1-D_q) H_{1+(1-D_q)(q-1)}(p), \tag{3.8.20}$$

$$I_{1,D_q}(p:p) = -(1-D_q) H(p). \tag{3.8.21}$$

2. Информация различия с $u_i = 1/m$. При равновероятном распределении $u_i = 1/m$ из (3.8.16) получим функционал

$$I_{q,D_q}(p:u) = \frac{1}{q-1}\log_2 \sum_i^m p_i^q \left(\frac{1}{m}\right)^{D_q(1-q)} = \\ = -\left[H_q(p) - D_q \log_2 m\right]. \tag{3.8.22}$$

Из условий выпуклости вытекает, что энтропия Реньи меньше (больше) выражения $D_q \log_2 m$ при $q > 0$ и $D_q > 1$ ($q < 0$ и $D_q < 1$).

3. Расхождение. Количественная мера расхождения определяется следующими выражениями

$$J_{D_q}(p:u) = I_{D_q}(p:u) + I_{D_q}(u:p) = \\ = -\left[H(p) + H(u)\right] + D_q\left[H(p:u) + H(u:p)\right], \tag{3.8.23}$$

$$J_{q,D_q}(p:u) = I_{q,D_q}(p:u) + I_{q,D_q}(u:p) = \\ = \frac{1}{q-1}\left[\log_2 \sum_i^m p_i^q u_i^{D_q(1-q)} + \log_2 \sum_i^m u_i^q p_i^{D_q(1-q)}\right]. \tag{3.8.24}$$

4. Мера неточности. Мера статистической неточности определения одного состояния случайного объекта относительно другого определяется функционалами

$$H_{D_q}(p:u) = H(p) + I_{D_q}(p:u) =$$

$$= -D_q H(p:u) = -D_q \sum_i^m (\log_2 u_i) p_i, \tag{3.8.25}$$

$$H_{q,D_q}(p:u) = H_q(p) + I_{q,D_q}(p:u) =$$

$$= -\log_2 \frac{N_{q-1}(p)}{N_{q-1}\left(\frac{p}{u^{D_q}}\right)} = \frac{1}{1-q} \log_2 \frac{\sum_i^m p_i^q}{\sum_i^m p_i^q u_i^{D_q(1-q)}}. \tag{3.8.26}$$

5. Мера Хаусдорфа. Обобщенная мера $\Gamma(q,\tau)$ рассматривалась в теории информации [116]. Отметим взаимосвязь ее с мерой Хаусдорфа в теории мультифракталов. Для этого рассмотрим множество ω, которое разложим на счетное число подмножеств ω_i с диаметрами $\ell_i < \ell$ $(\ell > 0)$. Подмножества имеют размерность $(-\tau)$. Вероятность, что элемент множества ω находится в ω_i есть p_i. Тогда в теории мультифракталов вводится так называемая статистическая сумма [56]

$$\Gamma(q,\tau,\{\omega_i\},\ell) = \sum_i^{m=m(\ell)} p_i^q \ell_i^{-\tau}, \tag{3.8.27}$$

где число состояний $m \approx c\ell^{-r}$.

Обобщение меры Хаусдорфа на случай мультифракталов есть величина [56]

$$\overline{\Gamma}(q,\tau) = \lim_{\ell \to 0} \sum_i^{m(\ell)} p_i^q \ell_i^{-\tau}. \tag{3.8.28}$$

Для каждого q существует одно значение $\tau = \tau(q)$ и мера имеет конечное значение в пределах $0 < \overline{\Gamma}(q,\tau) < \infty$. Тогда вводится обобщенная размерность Реньи

$$D_q = \frac{\tau(q)}{q-1} \ge 0, \tag{3.8.29}$$

которая при $\tau = 0$ имеет значение $D_q = 0$, а при $q = 0$ и $\tau(0) = 1$ совпадает с размерностью Хаусдорфа $D_0 = -r$. В итоге мера $\overline{\Gamma}(0,1)$ представляет собой известную D_0-мерную меру Хаусдорфа множества ω

$$\overline{\Gamma}(q,\tau) = \lim_{\ell \to 0} \sum_{i}^{m(\ell)} p_i^q \ell_i^{-\tau} . \qquad (3.8.30)$$

Эти сведения из теории мультифракталов показывают, что обобщенная мера Хаусдорфа (3.8.28) связана с обобщенной мерой (3.8.6) следующей зависимостью

$$\overline{\Gamma}(q,\tau) = \lim_{\ell \to 0} \ell^{-\tau} \Gamma(q,\tau) \qquad (3.8.31)$$

при $u_i = \ell_i / \ell$.

6. Нормированность и размерность. Двухпараметрический аналог информации различия Реньи (3.8.16) не удовлетворяет условию нормированности на единицу. Справедливо соотношение

$$I_{q,D_q}\left(1,0;\frac{1}{2},\frac{1}{2}\right) = D_q . \qquad (3.8.32)$$

Физическая безразмерная информация различия имеет следующий вид

$$I_{q,D_q}^{phys}(p:u) = \frac{1}{q-1} \ln \sum_{i}^{m} p_i^q u_i^{D_q(1-q)} . \qquad (3.8.33)$$

Функционал (3.8.16) есть отношение выражения (3.8.33) к значению физической энтропии $H_q^{phys}(u)$ при равновероятном состоянии с $m = 2$, то есть выполняется равенство

$$I_{q,D_q}(p:u) = \frac{I_{q,D_q}^{phys}(p:u)}{H_q^{phys}\left(\frac{1}{2},\frac{1}{2}\right)}, \quad H_q^{phys}\left(\frac{1}{2},\frac{1}{2}\right) = \ln 2 . \qquad (3.8.34)$$

Для второго двухпараметрического аналога информации различия Реньи имеем выражение:

$$I_{q,D_q}(p:u)=\frac{I^{phys}_{q,D_q}(p:u)}{I^{phys}_q\left(1,0;\frac{1}{2},\frac{1}{2}\right)}=\frac{1}{D_q(q-1)}\log_2\sum_i^m p_i^q u_i^{D_q(1-q)}, \quad (3.8.35)$$

которое удовлетворяет условию нормированности

$$I_{q,D_q}\left(1,0;\frac{1}{2},\frac{1}{2}\right)=1. \quad (3.8.36)$$

При $D_q=1$ оба аналога совпадают с информацией различия Реньи.

7. *k*-параметрическая информация различия. Находим безусловный экстремум энтропии Шеннона–Винера при заданности k мер неточности $H(f:p_r)$ с фиксированными распределениями $p_r=\{p_{r1},p_{r2},\dots,p_{rm}\}$ $(r=1,\dots,k)$ и сохранении нормировки для $f=\{f_1,\dots,f_m\}$. Из равенства нулю первой вариации функционала

$$L=-\sum_i^m(\log_2 f_i)f_i+\sum_r^k\alpha_r\sum_i^m(\log_2 p_{ri})f_i+\alpha\sum_i^m f_i, \quad (3.8.37)$$

где α_k – множители Лагранжа, получим нормированное распределение

$$f_i=\frac{p_{1i}^{\alpha_1}p_{2i}^{\alpha_2}\dots p_{ki}^{\alpha_k}}{\sum_i^m p_{1i}^{\alpha_1}p_{2i}^{\alpha_2}\dots p_{ki}^{\alpha_k}}, \quad \sum_i^m p_{ri}=1. \quad (3.8.38)$$

После подстановки (3.8.38) в энтропию получим ее максимальное значение

$$H(f)=\sum_r^k\alpha_r H(f:p_r)-\log_2\sum_i^m p_{1i}^{\alpha_1}p_{2i}^{\alpha_2}\dots p_{ri}^{\alpha_k}, \quad (3.8.39)$$

которое после деления на произведение $\prod_r^k\alpha_r$ примет вид:

$$\frac{1}{\prod_r^k \alpha_r} H(f) = \frac{1}{\prod_r^k \alpha_r} \sum_r^k \alpha_r H(f : p_r) +$$

$$+ I_{\alpha_1, \alpha_2, \ldots, \alpha_k}(p_1 : p_2 : \ldots : p_k). \qquad (3.8.40)$$

Здесь введена k-параметрическая информация различия

$$I_{\alpha_1, \alpha_2, \ldots, \alpha_k}(p_1 : p_2 : \ldots : p_k) =$$

$$= -\log_2 \left[\sum_i^m p_{1i}^{\alpha_1} p_{2i}^{\alpha_2} \ldots p_{ki}^{\alpha_k} \right]^{1/(\alpha_1 \alpha_2 \ldots \alpha_k)} =$$

$$= -\frac{1}{\alpha_1 \alpha_2 \ldots \alpha_k} \log_2 \sum_i^m p_{1i}^{\alpha_1} p_{2i}^{\alpha_2} \ldots p_{ki}^{\alpha_k}. \qquad (3.8.41)$$

Рассмотрим пример с двумя распределениями $p_{1i} = p_i$ и $p_{2i} = u_i$. Пусть справедливо равенство $\alpha_1 + \alpha_2 = 1$ с $\alpha_1 = q$ и $\alpha_2 = 1 - q$. Тогда из (3.8.41) вытекает q-информация различия [125]

$$I_{q,1-q}(p : u) = -\frac{1}{q(1-q)} \log_2 \sum_i^m p_i^q u_i^{1-q} =$$

$$= \log_2 \left[N_{q-1}\left(\frac{p}{u}\right) \right]^{1/q} = \log_2 \left[N_{-q}^*\left(\frac{u}{p}\right) \right]^{1/(1-q)}. \qquad (3.8.42)$$

При $q = 0$ и $q = 1$ имеем, соответственно, выражения информаций различия Кульбака–Лейблера

$$I_{0,1}(p : u) = \sum_i^m \left(\log_2 \frac{u_i}{p_i} \right) u_i, \qquad (3.8.43)$$

$$I_{1,0}(p : u) = \sum_i^m \left(\log_2 \frac{p_i}{u_i} \right) p_i. \qquad (3.8.44)$$

3.9. Тип аддитивной *q*-энтропии и *q* -информации различия

Рассмотрим состояние случайного объекта, которое описывается мультипликативным распределением $p_{ij} = p_i p_j$, а p_i и p_j есть распределения двух независимых объектов. Тогда выполняется свойство мультипликативности полунормы

$$N = N_1 N_2, \tag{3.9.1}$$

где $N = N_{q-1}(p_{12})$, $N_1 = N_{q-1}(p_1)$ и $N_2 = N_{q-1}(p_2)$.

Пусть выполняется свойство аддитивности для энтропии в виде равенства

$$H = H_1 + H_2, \tag{3.9.2}$$

в котором $H = H(N)$, $H_1 = H(N_1)$ и $H_2 = H(N_2)$.

В этом случае получим дифференциальное уравнение [21,23]

$$\frac{d\ln N}{dH} = \frac{d\ln N_1}{dH_1} = \frac{d\ln N_2}{dH_2} = -\lambda, \tag{3.9.3}$$

решением которого является физическая безразмерная энтропия

$$H_q^{phys}(p) = -\lambda^{-1} \ln N_{q-1}(p) \tag{3.9.4}$$

с точностью до коэффициента λ^{-1}.

Учитывая условие нормированности энтропии в теории информации, получим $\lambda = \ln 2$ и окончательно имеем энтропию Реньи

$$H_q(p) = -\log_2 N_{q-1}(p) = \frac{1}{1-q} \log_2 \left(\sum_i^m p_i^q \middle/ \sum_i^m p_i \right). \tag{3.9.5}$$

Аналогично используем полунормы

$$N = N_{q-1}\left(\frac{p_{12}}{u_{12}}\right), \; N_1 = N_{q-1}\left(\frac{p_1}{u_1}\right), \; N_2 = N_{q-1}\left(\frac{p_2}{u_2}\right), \tag{3.9.6}$$

удовлетворяющие условию мультипликативности (3.9.1), и закон аддитивности для информаций различия

$$I = I_1 + I_2 \tag{3.9.7}$$

После вычислений имеем уравнение

$$\frac{d\ln N}{dI} = \frac{d\ln N_1}{dI_1} = \frac{d\ln N_2}{dI_2} = \lambda, \tag{3.9.8}$$

выражение физической безразмерной информации различия:

$$I_q^{phys}(p:u)=\lambda^{-1}\ln N_{q-1}\left(\frac{p}{u}\right) \qquad (3.9.9)$$

и функционал Реньи

$$I_q(p:u)=\log_2 N_{q-1}\left(\frac{p}{u}\right)=\frac{1}{q-1}\log_2\left(\frac{\sum_i^m p_i^q u_i^{1-q}}{\sum_i^m p_i}\right). \qquad (3.9.10)$$

Например, при использовании обобщенной полунормы (3.7.5) имеем следующие значения функционалов

$$H_q(p)=\frac{1}{1-q}\log_2\left(\frac{\sum_i^m p_i^{q-1} f(p_i)}{\sum_i^m f(p_i)}\right), \qquad (3.9.11)$$

$$I_q(p:u)=\frac{1}{q-1}\log_2\left(\frac{\sum_i^m p_i^{q-1}u_i^{1-q} f(p_i)}{\sum_i^m f(p_i)}\right). \qquad (3.9.12)$$

При этом должно выполняться свойство мультипликативности функции $f(p_i p_j)=f(p_i)f(p_j)$, с помощью которой производится усреднение в соответствующих полунормах.

Таким образом, имеем один тип логарифмической зависимости от рассматриваемых полунорм, соответствующий свойству аддитивности энтропии и информации различия.

3.10. Экстремум энтропии Реньи и приложения

Исследуем экстремальные свойства мер Реньи, позволяющие находить наиболее вероятное распределение в случае заданности среднего значения произвольной случайной величины.

Приведем основные определения.

Определение 1. Взвешенное среднее каждой случайной величины $T=\{T_1,\dots,T_m\}$ в состоянии с параметризованным распределением $f=\{f_1,\dots,f_m\}$ равно:

$$\mathbf{E}_q(T)=\sum_i^m T_i f_i=\frac{\sum_i^m T_i p_i^q}{\sum_i^m p_i^q}, \quad \sum_i^m p_i=1. \qquad (3.10.1)$$

Если $p_i=1$ или $p_i=1/m$, то выражение (3.10.1) равняется обычному среднему арифметическому

$$\mathbf{E}_q(T)=\frac{1}{m}\sum_i^m T_i . \qquad (3.10.2)$$

Из определения 1 вытекают следующие свойства:

1. Однородность и нормированность. При замене p на λp $(a>0)$, выражение (3.10.1) является однородным функционалом нулевой степени относительно p, что означает его однородность.

Нормированность на единицу означает выполнимость равенства

$$\mathbf{E}_q(1)=1. \qquad (3.10.3)$$

2. Аддитивность и мультипликативность. Пусть случайная величина $T_{12}=T_1+T_2$ или $T_{12}=T_1T_2$ равняется, соответственно, сумме случайных величин или произведению одного или двух независимых объектов. Тогда получим равенства

$$\mathbf{E}_q(T_{12})=\mathbf{E}_q(T_1)+\mathbf{E}_q(T_2), \qquad (3.10.4)$$

$$\mathbf{E}_q(T_{12})=\mathbf{E}_q(T_1)\mathbf{E}_q(T_2), \qquad (3.10.5)$$

означающие аддитивность и мультипликативность наблюдаемых макроскопических величин.

Определение 2. Начальные и центральные моменты n-го порядка случайной величины T равны

$$\alpha_n=\sum_i^m T_i^n f_i , \qquad (3.10.6)$$

$$\mu_n=\sum_i^m\left[T_i-\mathbf{E}_q(T)\right]^n f_i , \qquad (3.10.7)$$

где $\Delta T_i=T_i-\mathbf{E}_q(T)$ есть флуктуация, среднее значение которой равняется нулю.

Вначале рассмотрим экстремум энтропии Реньи

$$H_q(p)=\frac{1}{1-q}\log_2\sum_i^m p_i^q \tag{3.10.8}$$

при дополнительных условиях заданности среднего значения произвольной случайной величины $T=\{T_1,\ldots,T_m\}$ и нормировки распределения

$$\mathbf{E}_q(T)=\sum_i^m T_i f_i\,,\quad \sum_i^m p_i=1. \tag{3.10.9}$$

Варьируем следующий функционал

$$L=\frac{1}{1-q}\log_2\sum_i^m p_i^q+\tau\sum_i^m T_i f_i-\alpha\sum_i^m p_i \tag{3.10.10}$$

и из условия

$$\delta L=\delta H_q(p)-\alpha\sum_i^m\delta p_i+$$
$$+\frac{q\tau}{\sum_i^m p_i^q}\left[\sum_i^m T_i p_i^{q-1}\delta p_i-\frac{\sum_i^m T_i p_i^q}{\sum_i^m p_i^q}\sum_i^m p_i^{q-1}\delta p_i\right]=0, \tag{3.10.11}$$

где первая вариация энтропии Реньи и среднего значения

$$\delta H_q(p)=\frac{q\sum_i^m f_i\delta(\ln p_i)}{(1-q)\ln 2}=\frac{q\sum_i^m p_i^{q-1}\delta p_i}{(1-q)\ln 2\sum_i^m p_i^q}, \tag{3.10.12}$$

$$\delta\mathbf{E}_q(T)=q\sum_i^m f_i(\Delta T_i)\delta(\ln p_i)=\frac{q}{\sum_i^m p_i^q}\sum_i^m p_i^{q-1}(\Delta T_i)\delta p_i\,, \tag{3.10.13}$$

получим равенство

$$\frac{qp_i^{q-1}}{(1-q)\ln 2\sum_i^m p_i^q}\Big\{1+(1-q)\tau\ln 2\big[T_i-\mathbf{E}_q(T)\big]\Big\}-\alpha=0. \tag{3.10.14}$$

Из (3.10.14) следует нормированное распределение

$$p_i = \left\{ 1+(1-q)\tau \ln 2\left[T_i - \mathbf{E}_q(T)\right]\right\}^{1/(1-q)} \Gamma_q^{-1}(\tau). \qquad (3.10.15)$$

Здесь функция

$$\Gamma_q(\tau) = \sum_i^m \left\{ 1+(1-q)\tau \ln 2\left[T_i - \mathbf{E}_q(T)\right]\right\}^{1/(1-q)} \qquad (3.10.16)$$

зависит от q и τ, которые меняются в пределах допустимых значений. При $1+(1-q)\tau \ln 2\left[T_i - \mathbf{E}_q(T)\right] < 0$ имеем $p_i = 0$, а при $q=1$ из (3.10.15) получим распределение (1.9.9) статистической модели Шеннона–Винера

$$p_i = 2^{\tau T_i}\Gamma^{-1}(\tau),\ \Gamma(\tau) = \sum_i^m 2^{\tau T_i}. \qquad (3.10.17)$$

Подставляя (3.10.15) в (3.10.8), получим

$$H_q(p) = \frac{1}{1-q}\log_2 \sum_i^m p_i^q =$$

$$= \frac{1}{1-q}\log_2 \sum_i^m \left\{ \frac{\Gamma_q^{1-q}(\tau)\, p_i}{1+(1-q)\tau \ln 2\left[T_i - \mathbf{E}_q(T)\right]} \right\} =$$

$$= \frac{1}{1-q}\log_2 \left\{ \Gamma_q^{1-q}(\tau) - \sum_i^m \frac{(1-q)\tau \ln 2\,\Gamma_q^{1-q}(\tau)\, p_i\left[T_i - \mathbf{E}_q(T)\right]}{1+(1-q)\tau \ln 2\left[T_i - \mathbf{E}_q(T)\right]} \right\} =$$

$$= \frac{1}{1-q}\log_2 \left\{ \Gamma_q^{1-q}(\tau) - (1-q)\tau \ln 2 \left(\sum_i^m p_i^q \right)\left(\sum_i^m \left[T_i - \mathbf{E}_q(T)\right] f_i \right) \right\} =$$

$$= \frac{1}{1-q}\log_2 \Gamma_q^{1-q}(\tau) = \log_2 \Gamma_q(\tau). \qquad (3.10.18)$$

Согласно (3.10.18), вытекает равенство

$$\Gamma_q(\tau) = \left(\sum_i^m p_i^q \right)^{1/(1-q)} = \left[N_{q-1}(p) \right]^{-1} \qquad (3.10.19)$$

и дифференциальное соотношение

$$dH_q(p) = -\tau d\mathbf{E}_q(T) \qquad (3.10.20)$$

между энтропией Реньи и средним значением случайной величины T.

Далее положим, что распределение p произвольное, а распределение u имеет вид

$$u_i = \left\{ 1 + (1-q)\tau_0 \ln 2\left[T_i - \mathbf{E}_{0q}(T)\right]\right\}^{1/(1-q)} \mathit{\Gamma}_q^{-1}(\tau_0), \qquad (3.10.21)$$

где

$$\mathit{\Gamma}_q(\tau_0) = \sum_i^m \left\{ 1 + (1-q)\tau_0 \ln 2\left[T_i - \mathbf{E}_{0q}(T)\right]\right\}^{1/(1-q)}, \qquad (3.10.22)$$

$$\mathbf{E}_{0q}(T) = \frac{\sum_i^m T_i u_i^q}{\sum_i^m u_i^q}. \qquad (3.10.23)$$

Находим информацию различия Реньи [23]

$$I_q(p:u) = \frac{1}{q-1}\log_2 \sum_i^m p_i^q u_i^{1-q} =$$

$$= \frac{1}{1-q}\log_2\left(\sum_i^m f_i u_i^{1-q} \sum_i^m p_i^q\right) = \frac{1}{q-1}\log_2 \sum_i^m p_i^q +$$

$$+ \frac{1}{q-1}\log_2 \sum_i^m f_i \left\{ 1 + (1-q)\tau_0 \ln 2\left[T_i - \mathbf{E}_{0q}(T)\right]\right\} \mathit{\Gamma}_q^{q-1}(\tau_0) =$$

$$= \frac{1}{q-1}\log_2 \sum_i^m p_i^q - \frac{1}{q-1}\log_2 \sum_i^m u_i^q +$$

$$+ \frac{1}{q-1}\log_2 \left\{ 1 + (1-q)\tau_0 \ln 2\left[\mathbf{E}_q(T) - \mathbf{E}_{0q}(T)\right]\right\} =$$

$$= -\left[H_q(p) - H_q(u)\right] + \frac{1}{q-1}\log_2 \left\{ 1 + (1-q)\tau_0 \ln 2\left[\mathbf{E}_q(T) - \mathbf{E}_{0q}(T)\right]\right\}. \qquad (3.10.24)$$

При $q = 1$ из (3.10.24) следует значение информации различия Кульбака–Лейблера (1.9.24) статистической модели Шеннона–Винера:

$$I(p:u)=\lim_{q\to 1} I_q(p:u)=$$
$$=-\left[H_q(p)-H_q(u)\right]-\tau_0\left[\mathbf{E}_q(T)-\mathbf{E}_{0q}(T)\right] \ . \qquad (3.10.25)$$

Пусть случайными объектами являются частицы, рассматриваемые в статистической физике. Положим, что $\beta=-k\tau\ln 2$ есть обратная температура, H_i – дискретные значения энергии частицы и $E_q=\mathbf{E}_q(H)$ – средняя энергия частицы. Тогда экстремум физической энтропии Реньи

$$H_q(p)=\frac{k}{1-q}\ln\sum_i^m p_i^q=k\ln\Gamma_q(\beta)=\beta\left(E_q-F_q\right) \qquad (3.10.26)$$

достигается, согласно (3.10.15), при равновесном распределении

$$p_i=\left\{\ 1+(1-q)k^{-1}\beta\left(H_i-E_q\right)\right\}^{1/(1-q)}\Gamma_q^{-1}(\beta), \qquad (3.10.27)$$

где F_q – свободная энергия. Дифференциальные соотношения

$$TdH_q(p)=dE_q,\ dF_q=-H_q(p)dT \qquad (3.10.28)$$

соответствуют соотношениям равновесной статистической термодинамики замкнутых систем [6].

Открытые системы, находящиеся в окружении с температурой T_0, характеризуются, согласно (3.10.24), физической информацией различия Реньи [23]

$$I_q(p:p_0)=\frac{1}{q-1}\ln\sum_i^m p_i^q p_{0i}^{1-q}=-\left[H_q(p)-H_q(p_0)\right]+$$
$$+\frac{k}{q-1}\ln\left[1-(1-q)k^{-1}\beta_0\left(E_q-E_{0q}\right)\right] \qquad (3.10.29)$$

с равенством $I_q(p:p_0)=0$ при распределении

$$p_i=p_{0i}=\left\{\ 1+(1-q)k^{-1}\beta_0\left(H_i-E_{0q}\right)\right\}^{1/(1-q)}\Gamma_q^{-1}(\beta_0), \qquad (3.10.30)$$

где $\beta_0=1/T_0$ и $E_{0q}=\mathbf{E}_{0q}(H)$.

Дифференцируя (3.10.29) по времени, получим следующее соотношение для открытых систем:

$$dI_q\left(p:p_0\right) \geq -dH_q\left(p\right) + \frac{dE_q}{T_0\left[\ 1-\left(1-q\right)k^{-1}\beta_0\left(E_q - E_{0q}\right)\right]}, \quad (3.10.31)$$

которое отличается от (1.11.12) наличием разности средних энергий частицы. Знак неравенства достигается для случая необратимых процессов.

При $q=1$ из распределений (3.10.27) и (3.10.30) вытекают канонические распределения Гиббса [6]

$$p_i = \exp\left\{-k^{-1}\beta\left(H_i - F\right)\right\},\ p_{0i} = \exp\left\{-k^{-1}\beta_0\left(H_i - F_0\right)\right\} \quad (3.10.32)$$

а дифференциальные соотношения (3.10.28) и (3.10.31) совпадают с соотношениями статистической теории Гиббса [19]

$$TdH\left(p\right) = dE\ ,\ dF = -H\left(p\right)dT\ , \quad (3.10.33)$$

$$dI\left(p:p_0\right) \geq -dH\left(p\right) + \frac{1}{T_0}dE\ . \quad (3.10.34)$$

В заключение отметим, что в рассматриваемой вариационной задаче остается открытым вопрос о знакоопределенности второй вариации $\delta^2 L$, что осложняет дальнейшую интерпретацию термодинамического соотношения (3.10.28). Другие трудности, связанные с усреднением при помощи распределения f, рассматриваются в монографии [21]. К ним можно отнести нелинейный характер информации различия (3.10.29) относительно изменения средней энергии.