

## ДИФЛЕКС-АНАЛИЗ ТЕХНОЦЕНОЗА ПО ЭЛЕКТРОПОТРЕБЛЕНИЮ

В.И. Гнатюк

Еще раз обратимся к ранговому анализу как основному методу исследования техноценозов. По определению – это метод исследования техноценозов, имеющий целью их статистический анализ, а также оптимизацию, и полагающий в качестве критерия форму ранговых параметрических распределений. Включает стандартные процедуры параметрического нормирования, интервального оценивания, прогнозирования, нормирования и потенцирования. Более тонкий анализ рангового параметрического распределения позволяет существенно повысить эффективность рангового анализа. Он осуществляется в рамках следующих (так называемых «тонких») процедур: дифлекс-анализа (на этапе интервального оценивания), GZ-анализа (на этапе прогнозирования), ASR-анализа (на этапе нормирования) и ZP-анализа (на этапе потенцирования).

Известно, что основным инструментом рангового анализа является ранговое параметрическое распределение. Вообще под ранговым распределением понимают полученное в результате процедуры ранжирования видов или особей техноценоза по какому-либо параметру распределение Ципфа в ранговой дифференциальной форме, по сути, являющееся невозрастающей последовательностью значений самих параметров, поставленных в соответствие рангу. Различают ранговые распределения, в которых ранжируются виды по количеству особей, которым они представлены в техноценозе (ранговые видовые), или объекты по значению параметра (ранговые параметрические). Применительно к параметру электропотребления нас будут интересовать ранговые параметрические распределения. И здесь необходимо вспомнить ряд базовых понятий.

Начнем с понятия ранга, под которым понимается номер по порядку при расположении объектов техноценоза в порядке снижения их электропотребления. При параметрическом описании техноценоза изначально мы всегда имеем дело с множеством эмпирических значений электропотребления объектов в фиксированный момент времени:

$$\{W_k\}_{k=1}^n, \quad (1)$$

где  $W_k$  – значение электропотребления  $k$ -ого объекта техноценоза;  
 $n$  – общее количество объектов техноценоза.

После процедуры ранжирования появляется возможность установить взаимно-однозначное соответствие между множествами:

$$\{W_k\}_{k=1}^n \xrightarrow{f:W \rightarrow R} \{R_k\}_{k=1}^n, \quad (2)$$

где  $\{R_k\}_{k=1}^n$  – множество возможных рангов объектов техноценоза в фиксированный момент времени;  
 $f : W \rightarrow R$  – числовая функция, устанавливающая соответствие между элементами множеств.

Следует отметить, что числовая функция соответствия между множествами значений электропотребления и рангов может быть определена на основе процедуры установления однозначного функционального соответствия (очевидно, что данная процедура определяется самой сутью процедуры параметрического ранжирования в техноценозе).

Технически функциональное соответствие между множествами может быть установлено посредством процедуры аппроксимации, под которой в общем случае понимается научный метод, состоящий в замене одних объектов другими, в каком-то смысле близкими к исходным, но более простыми. Аппроксимация позволяет исследовать числовые характеристики и качественные свойства объекта, сводя задачу к изучению более простых или более удобных объектов (например, таких, характеристики которых легко вычисляются или свойства которых уже известны). В теории чисел изучаются диофантовы приближения, в частности, приближения иррациональных чисел рациональными. В геометрии рассматриваются аппроксимации кривых ломаными. Некоторые разделы математики, в сущности, целиком посвящены аппроксимации, например, теория приближения функций, а также численные методы анализа. Теория приближений – раздел математики, изучающий вопрос о возможности приближенного представления одних математических объектов другими, как правило, более простой природы, а также вопросы об оценках вносимой при этом погрешности. Значительная часть теории приближений относится к приближению одних функций другими, однако есть и важные результаты, относящиеся к абстрактным векторным или топологическим пространствам.

В теории рангового анализа аппроксимация носит фундаментальный характер и связана с явлением случайности. В философии случайность – категория, выражающая отношение к основанию (сущности) процесса его отдельных форм (проявлений). При этом полагается, что случайность имеет свое основание не в сущности явления, а в воздействии на него других явлений; что это то, что может быть, а может и не быть, может произойти так, а может и иначе. В ряде концепций рассматривается как форма, за которой скрывается непознанная закономерность. В математике рассматривается как определение класса событий, которые при осуществлении некоторого комплекса условий иногда происходят, а иногда не происходят. В

алеатике (науке о случайности) рассматривается как важный атрибут объектов материального мира, отражающий континуальность параметров и фрактальность систем отсчета, а также имеющий следующие возможные причины: 1) непонятая закономерность; 2) скрещение несогласованных процессов; 3) уникальность; 4) неустойчивость движения; 5) относительность знания; 6) имманентная случайность; 7) произвольный выбор. При исследовании объектов техноценологического типа мы, в той или иной степени, имеем дело с причинами пятого и седьмого типов. Случайным в широком смысле является сочетание (именно фиксированное сочетание) видов технических изделий, составляющих техноценоз, если мы его рассматриваем среди большого количества других подобных техноценозов. Судить о статистическом (и далее – вероятностном) распределении данных сочетаний можно лишь полномасштабно исследовав поведение техноценозов в более общем таксономическом образовании – метаценозе (доступной для исследования в данный момент времени совокупности техноценозов). В узком смысле случайной является форма видового распределения, описывающего номенклатуру техноценоза, что делает случайной величиной значение соответствующего формального параметра. С другой стороны, если рассматривать совокупность одноименных параметров технических изделий техноценоза как выборку из параметрического пространства, то значение фиксированного параметра конкретного технического изделия может рассматриваться как случайная величина, а саму выборку в этом случае можно описать как статистическое распределение.

Формально в ранговом анализе аппроксимация реализуется после операции ранжирования для получения числовой функции соответствия множества значений параметра техноценоза множеству определения рангов. В результате получается числовая функция рангового параметрического распределения. Как сказано выше, аппроксимация эмпирических распределений в ранговом анализе обладает существенной спецификой. Например, если рассматривать совокупность значений электропотребления объектов техноценоза как выборку из параметрического пространства, то значение электропотребления конкретного объекта может рассматриваться как случайная величина, а саму выборку в этом случае можно описать как статистическое распределение. Учитывая, что в процессе аппроксимации мы фактически без изменения формы обобщаем конечную выборку значений электропотребления техноценоза до континуума генеральной совокупности, можно заключить, что аппроксимационная форма – это и есть соответствующее вероятностное распределение. Таким образом, в данном случае аппроксимация позволяет перейти от частной эмпирической выборки к вероятностному распределению генеральной совокупности. Однако, для того чтобы определить функцию рангового параметрического распределения по электропотреблению, прежде необходимо рассмотреть понятия ранговой топологии и ранговой топологической меры.

Топология (от древнегреческих «топос» – место и «логос» – слово, учение) – раздел математики, который изучает: в самом общем виде – явление непрерывности; в частности – свойства пространств, которые остаются неизменными при непрерывных деформациях (например – связность, ориентируемость). В частности, ранговая топология – раздел рангового анализа техноценозов, в котором изучаются свойства ранговых параметрических пространств (множеств параметров с дополнительной структурой рангового типа). Двумерным примером рангового параметрического пространства является множество значений одного отдельного параметра, заданное на множестве определения ранговой топологической меры (оба данных множества имеют мощность «алеф 1»). В простейшем случае – это числовая функция рангового параметрического распределения, определенная на множестве ранговой топологической меры, полученная в результате аппроксимации отранжированного множества значений параметра. Ранговая топологическая мера – количественная форма, отражающая качественное свойство объекта обладать большим или меньшим значением параметра. В конкретном техноценозе ранговая топологическая мера численно определяется как помноженная на количество объектов вероятность того, что в техноценозе будет превышено значение параметра, соотносимое с данной ранговой топологической мерой (при условии, что количество объектов стремится к бесконечности). Она дает континуальное обобщение понятия ранга как целочисленной меры близости объектов по значению параметра в упорядоченной последовательности, построенной по убыванию данного параметра. При этом ранги соответствуют целочисленным значениям ранговой топологической меры и задают на ранговом параметрическом распределении граничные значения параметра, близость к которым, в конечном итоге, и ранжирует объекты. Принципиально важным видится то, что континуальная ранговая топологическая мера позволяет достаточно точно определить место произвольного значения параметра на ранговом параметрическом распределении устоявшегося техноценоза.

В качестве исследуемого параметра здесь рассмотрим электропотребление – количественную форму одноименного показателя, фиксируемую счетчиками электроэнергии за интервал времени и определяемую как разность между значениями электропотребления в конце и начале рассматриваемого интервала. В случае стандартизации интервала времени (час, сутки, месяц, квартал, год и т.д.) значение электропотребления конкретного приемника или потребителя электроэнергии в базе данных будет фиксироваться в кВт·ч за принятый промежуток времени.

Итак, аналитически ранговое параметрическое распределение по электропотреблению представляет собой числовую функцию, определенную на множестве ранговой топологической меры, полученную в результате аппроксимации отранжированного множества значений электропотребления объектов (приемников, потребителей) техноценоза:

$$[\{W_k\}_{k=1}^n \xrightarrow{f:W \rightarrow R} \{R_k\}_{k=1}^n] \xrightarrow{\text{Approx}} W = f(x), \quad (3)$$

- где  $\{W_k\}_{k=1}^n$  – множество значений электропотребления;  
 $\{R_k\}_{k=1}^n$  – множество ранговой топологической меры;  
 $W(x)$  – функция рангового параметрического распределения техноценоза по электропотреблению (рис. 1);  
 $x$  – ранговая топологическая мера.

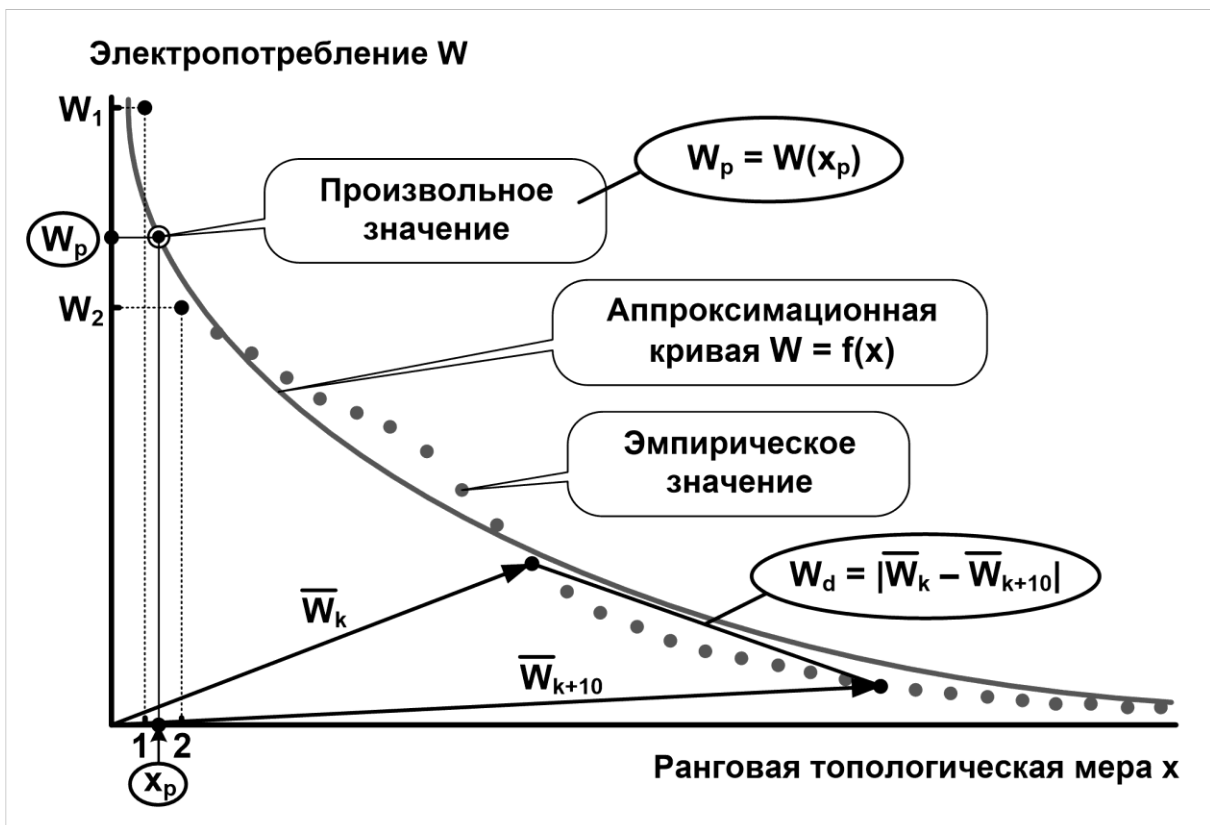


Рис. 1. К понятию рангового параметрического распределения техноценоза по электропотреблению

Рассмотрим подробнее рисунок 1, на котором построен график аппроксимационной функции рангового параметрического распределения техноценоза по электропотреблению  $W(x)$ . По оси ординат здесь откладываются значения параметра дифференциального электропотребления  $W$  (кВт·ч/Т), а по оси абсцисс – ранговая топологическая мера  $x$ . Кроме того, на этом же рисунке изображены точки, соответствующие отранжированным эмпирическим значениям электропотребления объектов техноценоза  $W_k$  (кВт·ч/Т). При этом точки расположены так, что их ранги соответствуют целочисленным значениям ранговой топологической меры:

$$\left\{ \begin{array}{l} (W_1; R_1) \rightarrow (W = W_1; x = 1); \\ (W_2; R_2) \rightarrow (W = W_2; x = 2); \\ \dots \\ (W_k; R_k) \rightarrow (W = W_k; x = k); \\ \dots \\ (W_n; R_n) \rightarrow (W = W_n; x = n). \end{array} \right. \quad (4)$$

Таким образом, следует различать два ранговых параметрических распределения одного и того же техноценоза: первое – эмпирическое, построенное на основе обработки данных по электропотреблению, полученных со счетчиков электроэнергии, установленных на объектах; второе – теоретическое, построенное после аппроксимации эмпирических данных. Повторимся и еще подчеркнем, что в процессе аппроксимации мы фактически без изменения формы обобщаем конечную выборку эмпирических точек техноценоза до континуума всей генеральной совокупности. Следовательно, аппроксимационная форма – это и есть соответствующее вероятностное распределение, которое мы должны рассматривать как единственный корректный инструмент рангового анализа техноценозов.

Какие задачи позволяет решать подобная аппроксимационная форма рангового параметрического распределения по электропотреблению? Прежде всего, она позволяет определять дробное (межранговое) значение ранговой топологической меры, соответствующее произвольному значению электропотребления. Это позволяет точно позиционировать (с точки зрения ранговой динамики) значение электропотребления, извне привносимое в устоявшийся техноценоз. На рисунке 1 наглядно проиллюстрировано, сколь значительную погрешность можно получить, если оперировать только эмпирическим ранговым параметрическим распределением. Заметим, что это имеет большое значение в процедурах параметрического нормирования и синтеза, а также в бифуркационных методах прогнозирования электропотребления. Кроме того, аппроксимационная форма рангового параметрического распределения позволяет корректно вычислять совокупное значение электропотребления техноценоза в целом или отдельного параметрического кластера (например, если его зафиксировать в пределах от  $x_1$  до  $x_2$ ). Для этого берутся определенные интегралы:

$$\int_0^{\infty} w(x) dx \quad \text{или, соответственно,} \quad \int_{x_1}^{x_2} w(x) dx. \quad (5)$$

Следует отметить, что в первом случае мы имеем дело с несобственным интегралом, который, в общем случае, может быть сходящимся или расходящимся. Однако в теории рангового анализа показано, что несобственные интегралы ранговых параметрических распределений всегда являются сходящимися, т.е. позволяют получать корректные конечные результаты. Более того, сам факт расходимости подобного интеграла является свидетельством, что исследуемое ранговое распределение, по какой-либо причине, не может быть отнесено к области негауссовых. Интегрирование ранговых параметрических распределений применяется в процедурах потенцирования и ZP-анализа, а также при определении интегральных показателей качества и затрат по электропотреблению.

Ранее нами в рамках процедуры дифлекс-анализа был предложен параметр, описывающий качество процесса электропотребления. Дифлекс-анализ – тонкая процедура рангового анализа техноценоза, осуществляемая на этапе интервального оценивания с целью разработки оптимального плана углубленных обследований «аномальных» объектов на среднесрочную перспективу (до 5 – 7 лет). При этом предполагается, что основным параметром процедуры дифлекс-анализа является дифлекс-параметр, под которым понимается отклонение (абсолютное или относительное) эмпирического значения электропотребления объекта техноценоза от нижней границы области допустимых значений. Учитывая сугубо эмпирический характер данного параметра, а также ряд важных теоретических обобщений, касающихся ранговой топологии, которые мы сделаем ниже, предлагается его отныне называть ранговым дифлекс-параметром.

Как известно, дифлекс-анализ является тонким дополнением к процедуре интервального оценивания. По определению интервальное оценивание – процедура оптимального управления электропотреблением техноценоза, заключающаяся в определении точек эмпирического рангового параметрического распределения по электропотреблению, выходящих за пределы области допустимых значений, построенной относительно аппроксимационной кривой распределения. Точки, выходящие за пределы области допустимых значений, фиксируют объекты, аномально потребляющие электроэнергию. При этом если точка находится ниже области допустимых значений, то считается, что объект потребляет электроэнергию аномально мало, а если выше области, то аномально много. В обоих случаях объект нуждается в углубленном энергетическом обследовании с целью выявления причин его подобного аномального состояния.

Предполагается, что в качестве источника данных используется база данных по электропотреблению за 10 – 15 лет предыстории. Интервальное оценивание проводится с целью определения границ параметрической области допустимых значений. Нижняя граница области допустимых значений – гиперболическая кривая, полученная в результате аппроксимации

нижних границ 95 %-ых доверительных интервалов, рассчитанных для каждого из рангов рангового параметрического распределения. Построение области допустимых значений на основе значений электропотребления рангов позволяет учесть системное влияние техноценоза на объекты, объектов на техноценоз, а также множественное взаимное влияние объектов друг на друга. Анализ, выполненный для большого числа разнородных объектов различных техноценозов, позволил подтвердить предположение о нормальном распределении значений электропотребления внутри рангов, что дает возможность на основе эмпирических данных за ряд временных интервалов построить для каждого ранга доверительный интервал (рис. 2).

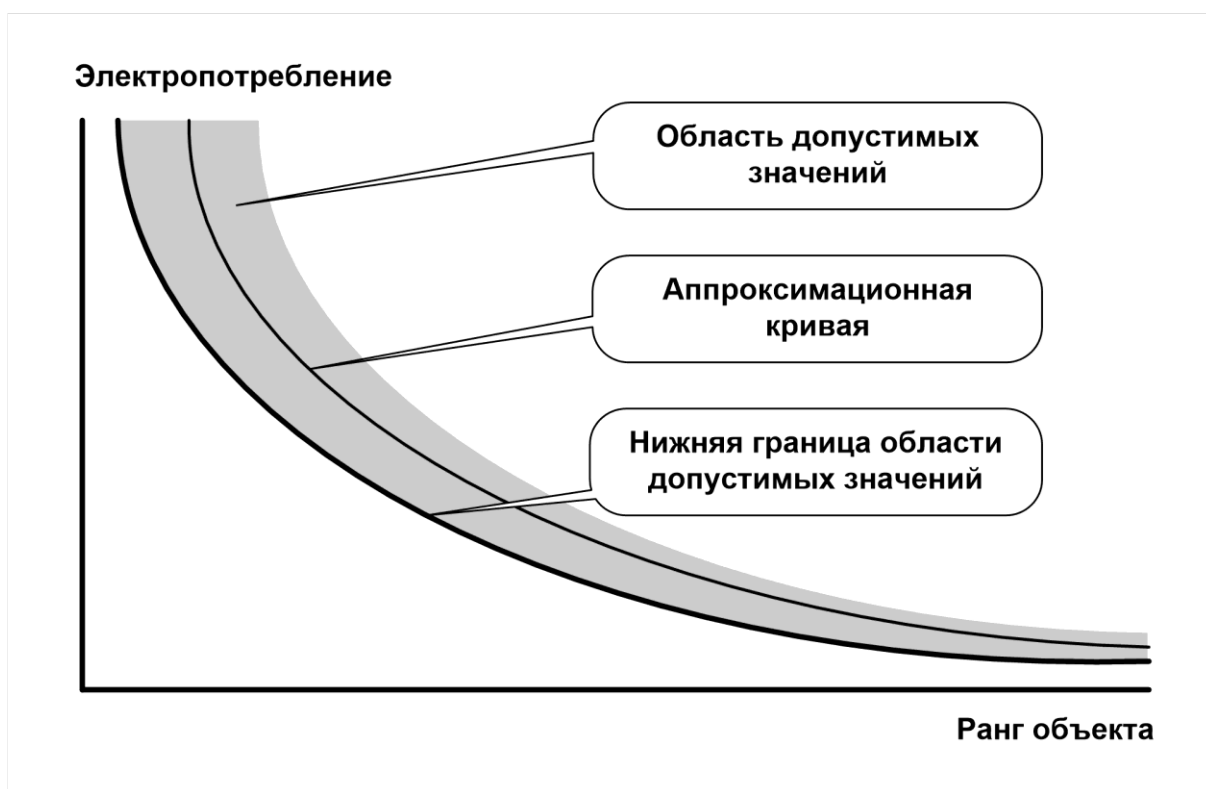


Рис. 2. Область допустимых значений электропотребления

Как известно, если  $\Theta^*$  служит оценкой неизвестного параметра  $\Theta$ , то доверительным называется интервал  $[\Theta^* - \delta, \Theta^* + \delta]$ , который покрывает неизвестный параметр с надежностью  $\gamma$  и точностью  $\delta$ :

$$P [\Theta^* - \delta < \Theta < \Theta^* + \delta] = \gamma. \quad (6)$$

Если случайная величина  $W$  распределена нормально, то по данным выборки объемом  $n$  можно ввести случайную величину  $T$ , которая имеет распределение Стьюдента с  $k = n - 1$  степенями свободы:



$$T = \frac{\bar{W} - m}{S / \sqrt{n}}, \quad (7)$$

где  $\bar{W}$  – выборочная средняя;  
 $m$  – неизвестное математическое ожидание;  
 $S$  – исправленное среднее квадратичное отклонение.

Плотность распределения Стьюдента определяется выражением:

$$S(t, n) = \frac{\Gamma(n/2)}{\sqrt{\pi(n-1)} \cdot \Gamma((n-1)/2)} \left(1 + \frac{t^2}{n-1}\right)^{-n/2}, \quad (8)$$

где  $\Gamma(y) = \int_0^{\infty} u^{y-1} e^{-u} du$  – гамма-функция ( $t$  – табличный).

Как видно из (8), распределение Стьюдента определяется одним параметром – объемом выборки  $n$  и не зависит от неизвестных величин. Так как  $S(t, n)$  – четная функция от  $t$ , то вероятность неравенства

$$\frac{\bar{W} - m}{S / \sqrt{n}} < t_{\gamma} \text{ определяется следующим условием:} \quad (9)$$

$$P \left( \left| \frac{\bar{W} - m}{S / \sqrt{n}} \right| < t_{\gamma} \right) = 2 \int_0^{t_{\gamma}} S(t, n) dt = \gamma. \quad (10)$$

При замене строгого неравенства в выражении (9) двойным неравенством, а также с учетом уравнения (10) получаем вероятностную оценку неизвестного математического ожидания  $m$  с надежностью  $\gamma$  :

$$P (\bar{W} - t_{\gamma} S / \sqrt{n} < m < \bar{W} + t_{\gamma} S / \sqrt{n}) = \gamma. \quad (11)$$

При замене случайных величин  $\bar{W}$  и  $S$  неслучайными величинами  $\bar{w}$  и  $s$ , найденными по выборке, получается доверительный интервал, покрывающий неизвестный параметр  $m$  с надежностью  $\gamma$  (рис. 3).

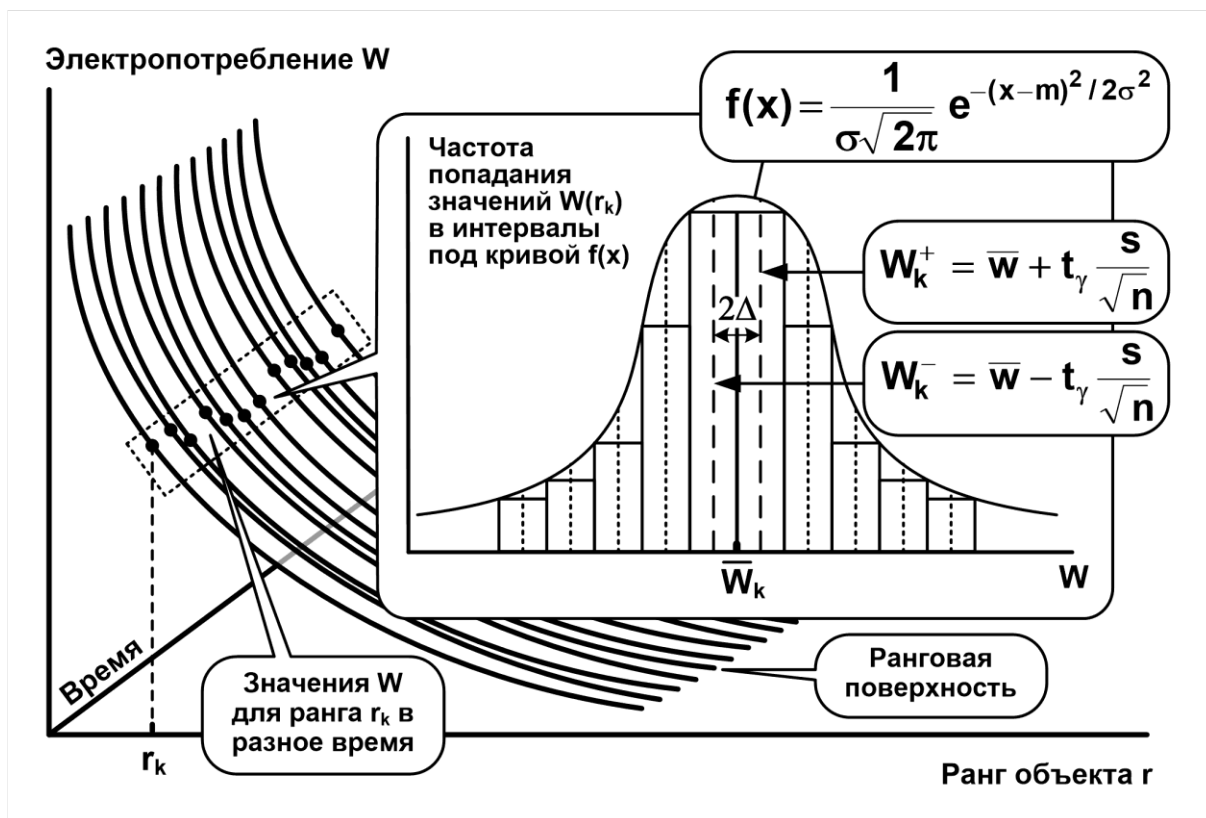


Рис. 3. Границы доверительных интервалов рангов техноценоза

Как показано на рисунке 3, нижняя и верхняя границы доверительных интервалов рангов могут быть вычислены следующим образом:

$$((\bar{w} - t_\gamma s / \sqrt{n}), (\bar{w} + t_\gamma s / \sqrt{n})), \quad (12)$$

где  $s$  – выборочное среднее квадратичное отклонение;  
 $\bar{w}$  – выборочное среднее (находится по выборке);  
 $t_\gamma$  – аргумент (находится таблично по заданным  $n$  и  $\gamma$ ).

Данный подход для построения доверительных интервалов имеет следующие преимущества: возможность применения для выборок с малым объемом ( $n < 30$ ), а также отсутствие неизвестных параметров распределения. В наших исследованиях неизвестным параметром для фиксированного ранга является истинное электропотребление  $W$ , а его оценкой выступает выборочное среднее значение электропотребления  $\bar{w}$ .

После расчета для каждого из рангов границ доверительных интервалов осуществляется аппроксимация нижней и верхней границ области допустимых значений электропотребления техноценоза (рис. 2). Как мы полагаем, область допустимых значений включает в себе, своего рода, зону «физически нормального» разброса значений электропотребления для данного техноценоза на данном интервале времени (с учетом возможных

труднопредсказуемых колебаний значительного количества объективных факторов, влияющих на процесс электропотребления). При этом, учитывая, что расчеты осуществляются на эмпирических данных самого техноценоза, можно сделать вывод, что подобный уровень электропотребления одновременно является и достаточным для объектов (с технологической точки зрения). Если исходить из цели минимизации расходования энергетических ресурсов, то логично предположить, что нижняя граница области допустимых значений показывает уровень наилучшего электропотребления. Отклонение же от нижней границы (как в большую, так и в меньшую стороны) можно считать показателем качества процесса электропотребления, который предлагается описывать дифлекс-показателем.

Итак, дифлекс-показатель – это мера, отражающая свойство объектов техноценоза (приемников или потребителей) осуществлять процесс электропотребления с большей или меньшей степенью энергоэффективности. В данном случае под энергоэффективностью понимается показатель, отражающий уровень минимизации количества электроэнергии для полного обеспечения питаемого технологического процесса.

Дифлекс-показатель  $k$ -го объекта техноценоза количественно может характеризоваться абсолютным ранговым дифлекс-параметром (рис. 4):

$$\Delta W(r_k) = |W(r_k) - W^H(r_k)|, \quad (13)$$

где  $W(r_k)$  – эмпирическое значение электропотребления  $k$ -го объекта в рассматриваемый момент времени;  
 $W(r)$  – ранговое параметрическое распределение объектов техноценоза по электропотреблению;  
 $r_k$  – ранг  $k$ -го объекта на распределении;  
 $W^H(r_k)$  – значение электропотребления, соответствующее  $k$ -му рангу на нижней границе области допустимых значений рангового распределения.

В ряде случаев (при рассмотрении техноценоза в динамике) можно вести речь также и об относительном ранговом дифлекс-параметре:

$$\Delta W^o(r_k) = \frac{|W(r_k) - W^H(r_k)|}{W(r_k)}. \quad (14)$$

В заключение здесь отметим, что изложенный выше метод получения области допустимых значений не является единственным. Для этого также может применяться и комбинаторный ранговый анализ.

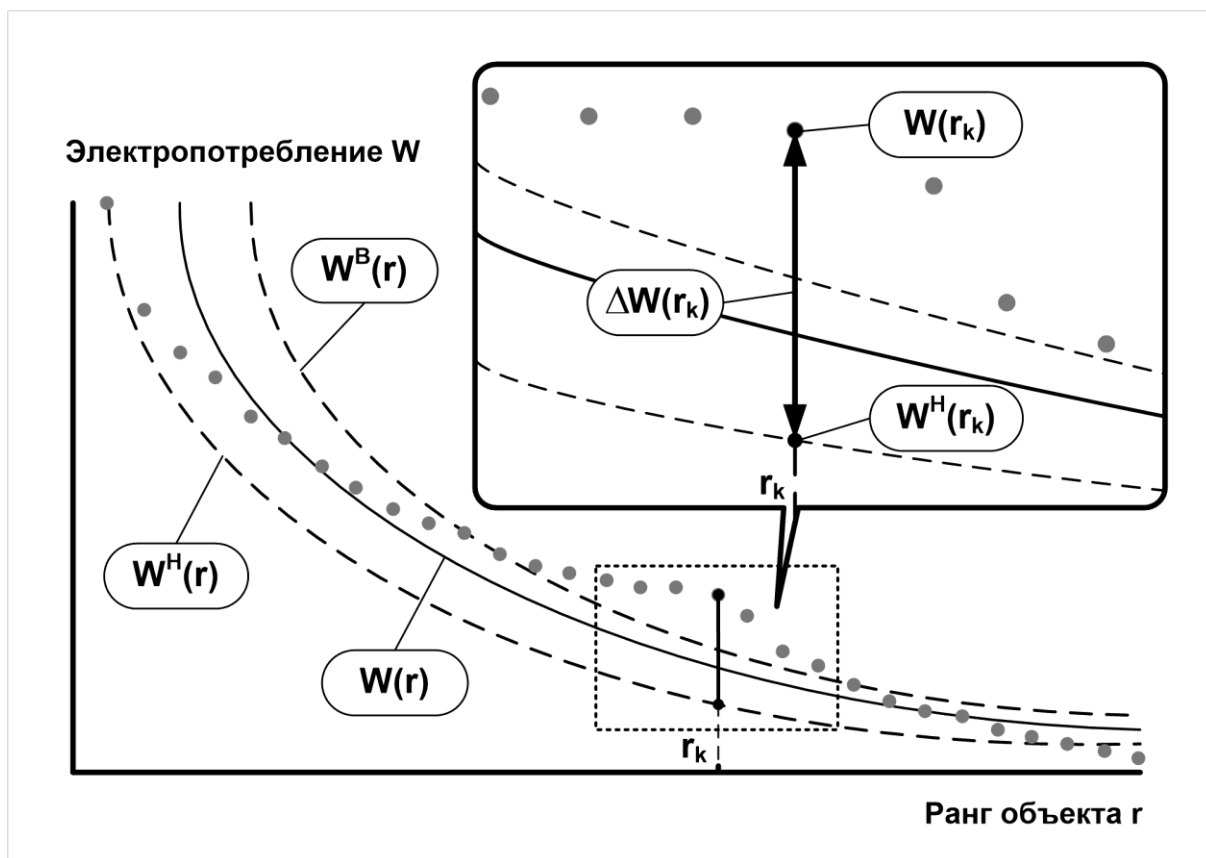


Рис. 4. К понятию рангового дифлекс-параметра

С целью теоретического обобщения рассмотрим процедуру дифлекс анализа в области ранговой топологии – раздела рангового анализа техноценозов, в котором изучаются свойства ранговых параметрических пространств (множеств параметров с дополнительной структурой рангового типа). Как было уже сказано выше, аппроксимация рангового параметрического распределения по электропотреблению позволяет перейти от частной эмпирической выборки к вероятностному распределению генеральной совокупности. Это позволяет ввести понятие ранговой топологической меры и определить топологический дифлекс-параметр (для простоты далее мы будем его именовать просто дифлекс-параметром) (рис. 5).

Подобное континуальное представление позволяет впервые установить однозначное функциональное соответствие между областью значений дифлекс-параметра и множеством значений ранговой топологической меры. При этом дифлекс-параметр для любого расчетного значения ранговой топологической меры может быть определен как разность значений электропотребления, полученная после аппроксимации распределений:

$$\Delta W(x_p) = W(x_p) - W^H(x_p), \quad (15)$$

где  $W(x_p)$  – значение электропотребления, соответствующее расчетному значению на аппроксимационной кривой;  
 $W^H(x_p)$  – значение электропотребления на нижней границе области допустимых значений.

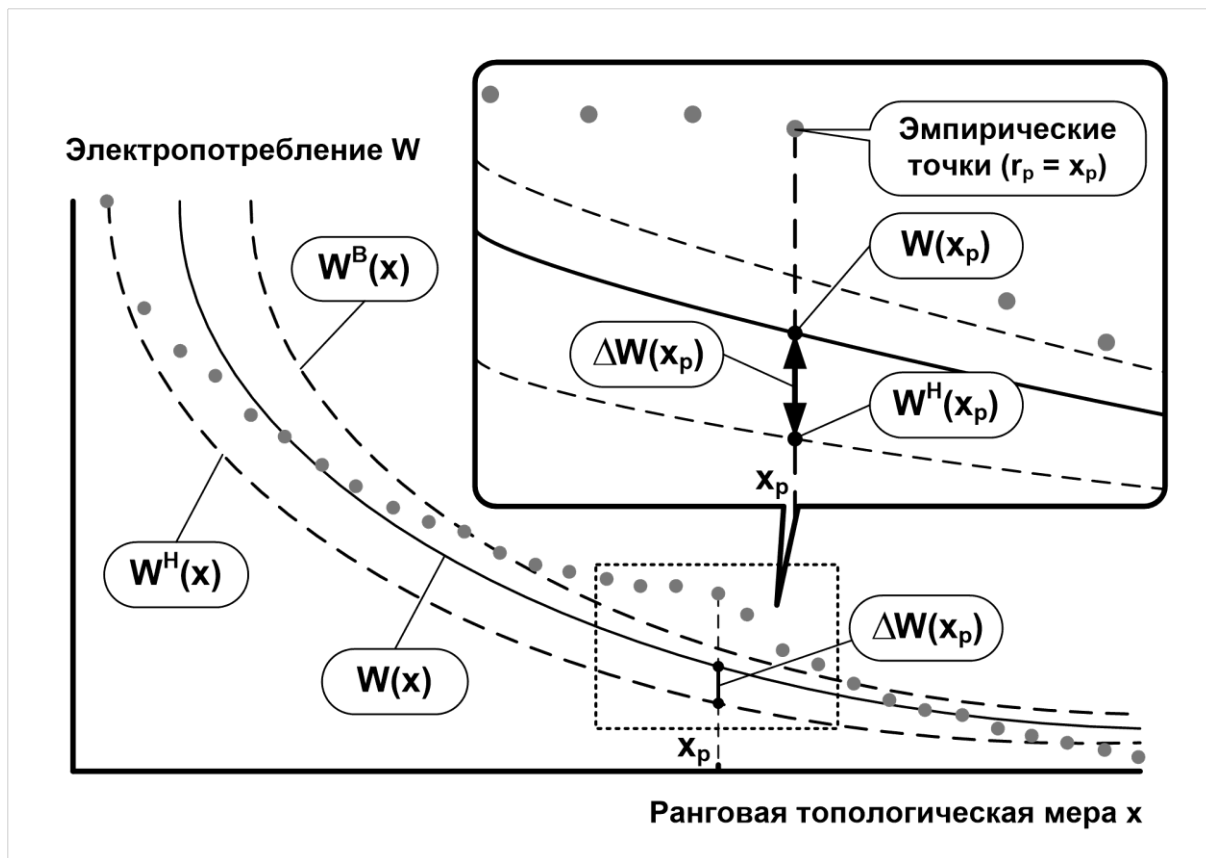


Рис. 5. Топологический дифлекс-параметр по электропотреблению

Таким образом, на этапе статистического анализа и построения эмпирической модели процесса электропотребления осуществляется глубокая обработка данных по электропотреблению объектов, которая включает интервальное оценивание, прогнозирование и нормирование. Интервальное оценивание выявляет в динамике и наглядно представляет объекты с аномальным электропотреблением. Прогнозирование электропотребления отдельными объектами и техноценозом в целом осуществляется с использованием гауссовых и ципфовых методов. Кластерный анализ позволяет разбить объекты по группам и осуществить нормирование электропотребления объектов в каждой группе с подробным статистическим описанием норм. Более тонкий анализ рангового параметрического распределения по электропотреблению позволяет существенно повысить эффективность рангового анализа. Он осуществляется в рамках следующих процедур: дифлекс-анализа (на этапе интервального оценивания), GZ-анализа (на этапе прогнозирования) и ASR-анализа (на этапе нормирования).