Лекция 5

Цензурирование. Выбор модели надёжности

Пример 1

На испытания поставлены *7 экземпляров* однотипного оборудования.

Поставлена задача найти точечную оценку параметра распределения, считая, что времена отказов подчиняются *экспоненциальному* распределению.

В ходе испытаний были зафиксированы отказы трех экземпляров оборудования с временами 1, 2 и 5 единиц времени.

По прошествии 6 единиц времени было принято решение прекратить испытания.

Пример 1 (продолжение)

Предположим, было принято решение проводить процедуру оценивания, основываясь только на наблюдаемых значениях времени отказа, считая, что количество объектов равно 3.

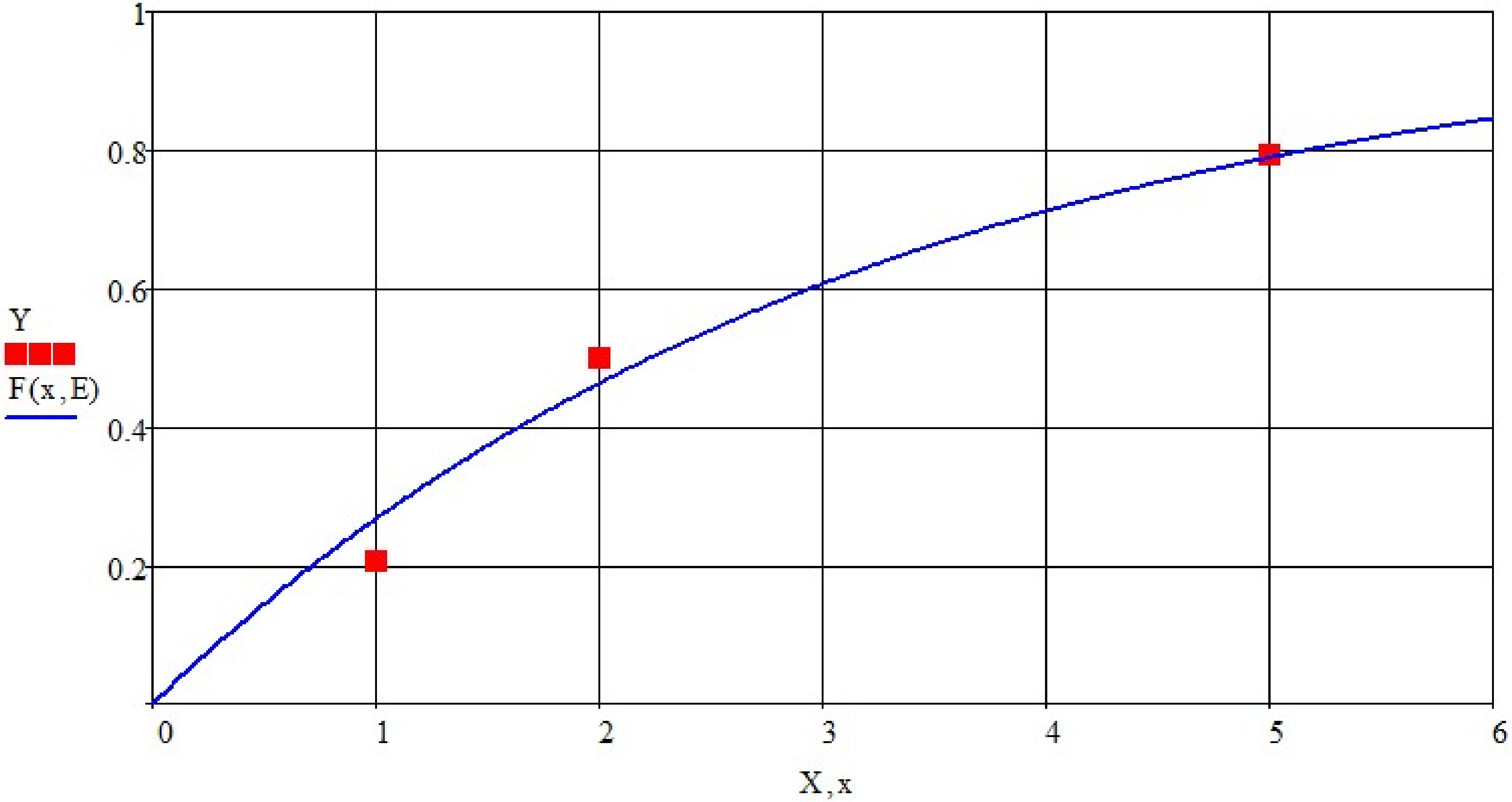
Используем МНК.

$$X = \begin{bmatrix} 1 \\ 2 \\ 5 \end{bmatrix} \quad Y = \begin{bmatrix} 0{,}206 \\ 0{,}5 \\ 0{,}794 \end{bmatrix} \qquad RSS(\lambda) = \sum_{i=1}^{3} [Y_i - F(X_i, \lambda)]^2$$

$$E_{МНК} = 0{,}311$$

$$T_{МНК} = 3{,}215$$

Пример 1 (продолжение)



Пример 1 (продолжение)

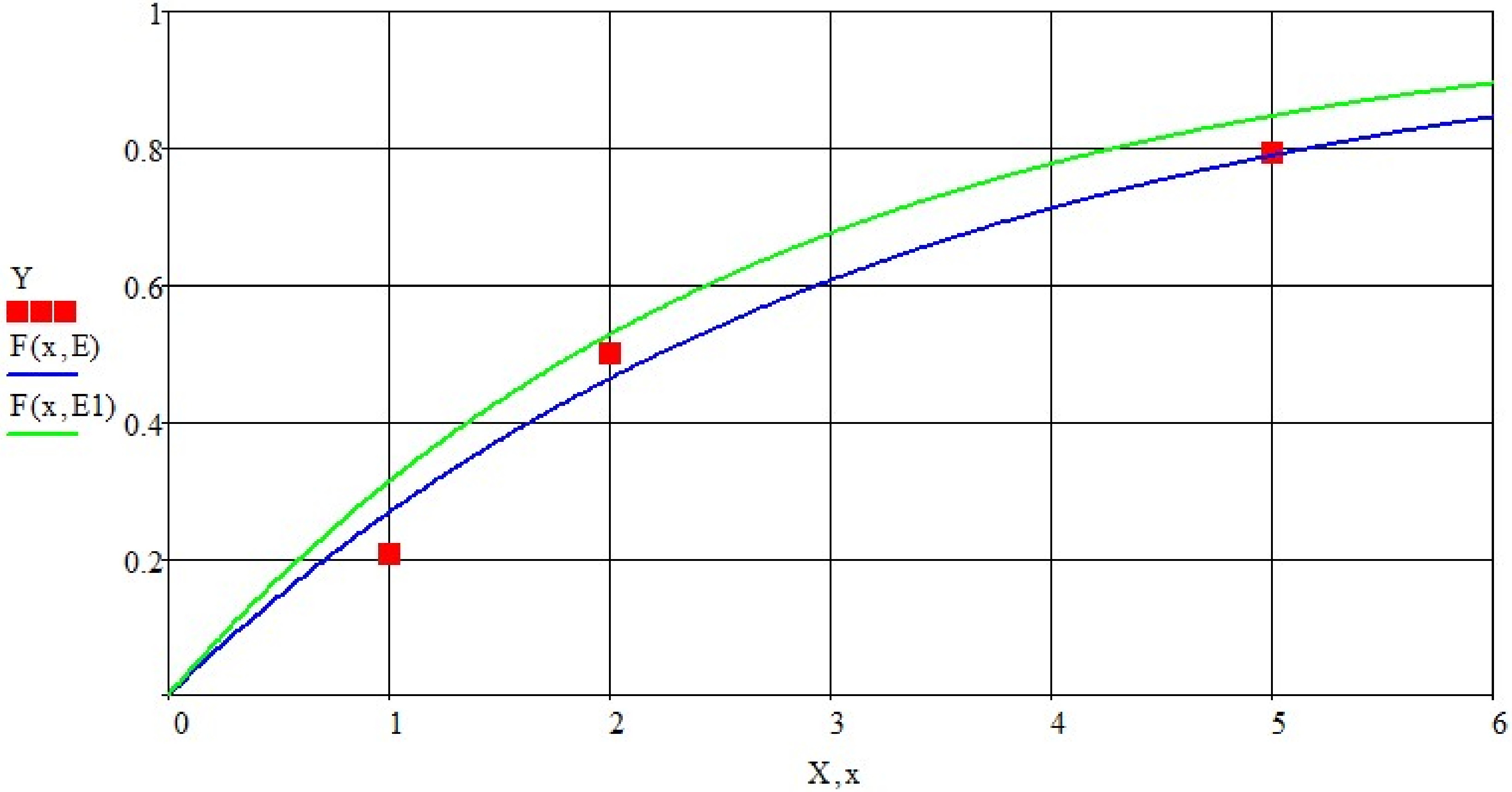
Используем ММП.

$$X = \begin{bmatrix} 1 \\ 2 \\ 5 \end{bmatrix} \qquad \Lambda(\lambda) = \sum_{i=1}^{3} \ln f(X_i, \lambda)$$

$$E_{\text{ММП}} = 0{,}375$$

$$T_{\text{ММП}} = 2{,}667$$

Пример 1 (продолжение)



Пример 1 (продолжение)

Однако, тот факт, что некоторые экземпляры оборудования не отказали к моменту прекращения наблюдения, представляет собой информацию.

Отказ от использования этой информации означает искажение результатов оценивания.

Предположим теперь, что для МНК используются значения Y, рассчитанные, исходя из того, что общее количество экземпляров оборудования равно 7.

Пример 1 (продолжение)

МНК

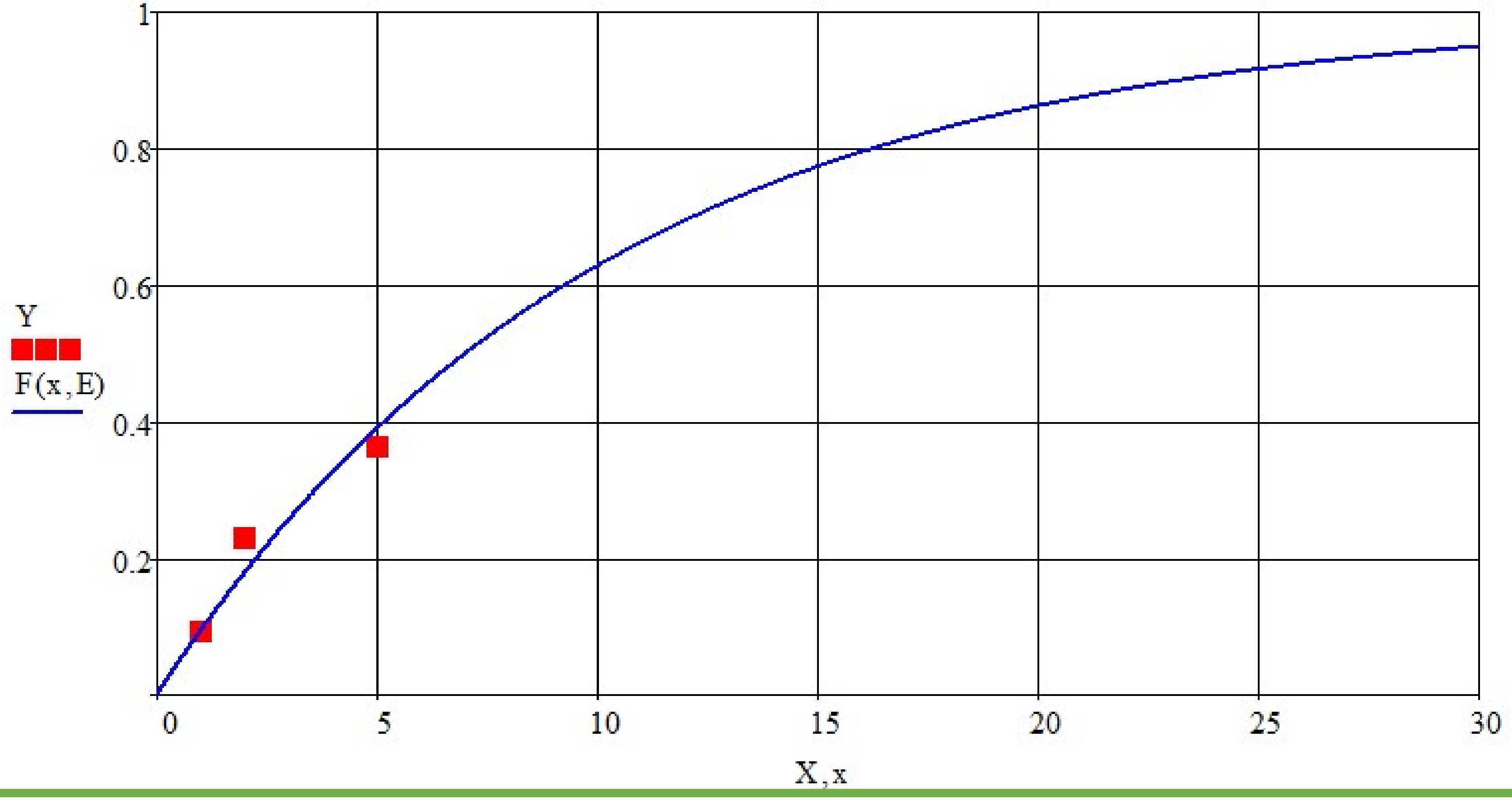
$$X = \begin{bmatrix} 1 \\ 2 \\ 5 \\ ? \\ ? \\ ? \\ ? \end{bmatrix} \qquad Y = \begin{bmatrix} 0{,}094 \\ 0{,}228 \\ 0{,}364 \\ 0{,}5 \\ 0{,}636 \\ 0{,}772 \\ 0{,}906 \end{bmatrix}$$

$$RSS(\lambda) = \sum_{i=1}^{3} [Y_i - F(X_i, \lambda)]^2$$

$$E_{МНК} = 0{,}099$$

$$T_{МНК} = 10{,}099$$

Пример 1 (продолжение)



Пример 1 (продолжение)

Видно, что учет даже части из имеющейся информации существенно изменяет результат оценки параметра.

ММП позволяет учесть больше информации.

Обозначим количество наблюдаемых отказов через $N_F = 3$, а число экземпляров оборудования, которые не отказали к моменту прекращения наблюдения через $N_R = 4$.

Также, обозначим через $T_R = 6$ момент окончания наблюдения.

Пример 1 (продолжение)

Имеющиеся данные можно представить в следующем виде:

1	F	
2	F	отказы
5	F	
6	R	
6	R	прекращение наблюдений
6	R	
6	R	

Пример 1 (продолжение)

Информацию о том, что некоторый экземпляр оборудования не отказал к моменту времени T_{sus} можно учесть через вероятность этого события:

$$\Pr[X \geq T_R] = 1 - F(T_R, \lambda)$$

Для использования этой информации необходимо модифицировать логарифмическую функцию правдоподобия следующим образом:

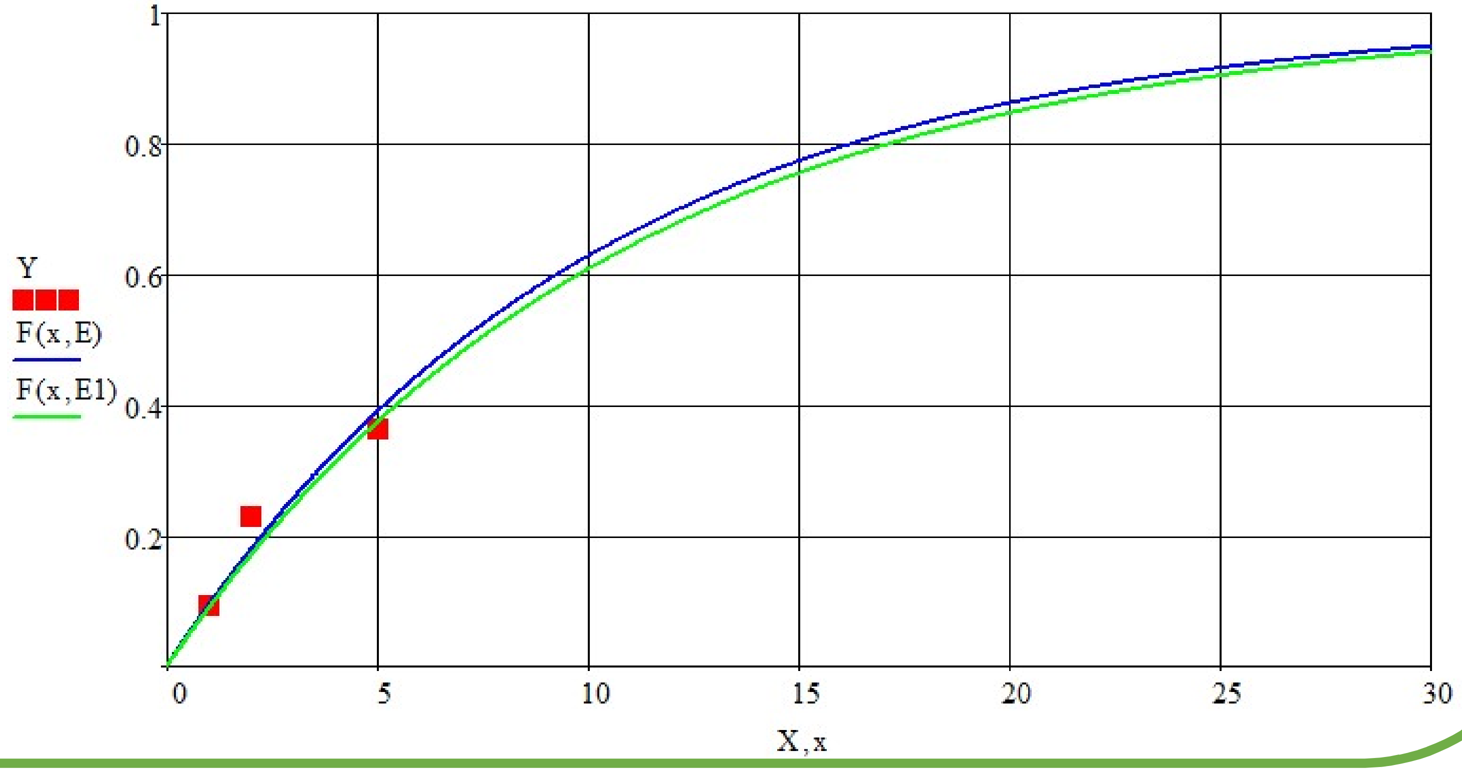
$$\Lambda(\lambda) = \sum_{i=1}^{N_F=3} \ln f(X_i, \lambda) + N_R \cdot \ln[1 - F(T_R, \lambda)]$$

Пример 1 (продолжение)

Далее определяем максимум модифицированной функции.

$E_{\text{ММП}} = 0{,}094$

$T_{\text{ММП}} = 10{,}667$



В общем случае, когда за различными экземплярами оборудования прекращают наблюдать в разное время T_{R_j}, логарифмическая функция правдоподобия принимает вид:

$$\Lambda(\mathbf{\Theta}) = \sum_{i=1}^{N_F} \ln f(X_i, \mathbf{\Theta}) + \sum_{j=1}^{N_R} \ln\left[1 - F\left(T_{R_j}, \mathbf{\Theta}\right)\right]$$

где $\mathbf{\Theta}$ – вектор оцениваемых параметров, X_i - времена наблюдаемых отказов, T_{R_j} - времена прекращения наблюдения.

В процессе анализа надежности приходится сталкиваться с ситуациями, когда определенная часть объектов или систем не отказывает за период наблюдения,

а другая часть отказывает, но моменты отказов точно неизвестны.

В таких ситуациях возникает необходимость проведения статистического анализа надежности на основе специфических выборок, основной особенностью которых является отсутствие сведений о моментах отказов контролируемой части изделий.

Это явление носит название *цензурирование данных*, а получаемые в результате выборки — *цензурированными выборками*.

Под данными, применительно к задачам надежности, понимают фиксированные значения наработок изделий, полученные по результатам испытаний или эксплуатационных наблюдений.

Данными цензурированной выборки являются как наработки отказавших объектов, так и наработки неотказавших объектов, а также могут быть интервалы времени, в течение которых объект отказал, но момент отказа точно неизвестен.

Прежде чем перейти к описанию структуры представления данных цензурированных выборок, приведем некоторые определения.

Событие, приводящее к прекращению испытаний или эксплуатационных наблюдений объекта до наступления отказа (предельного состояния) изучаемого характера, либо к обнаружению отказа изучаемого характера в пределах известного интервала наработки, называется *цензурированием*.

Цензурированной выборкой называется выборка, элементами которой являются значения наработки до отказа и наработки до цензурирования, либо только значения наработки до цензурирования.

Как было отмечено ранее, цензурирование — это процесс возникновения неопределенности момента отказа объекта, причем интервал неопределенности известен аналитику.

Интервал неопределенности — интервал наработки, внутри которого произошел либо произойдет отказ объекта, причем точное значение наработки до отказа неизвестно.

Этот интервал может быть неограниченным справа, тогда говорят о *цензурировании справа*, либо ограниченным справа, тогда говорят о *цензурировании слева*. Если интервал неопределенности момента отказа ограничен слева и справа, то говорят о *цензурировании интервалом*.

Пример 2

На испытания поставлены *7 экземпляров* однотипного оборудования.

Поставлена задача найти точечную оценку параметра распределения, считая, что времена отказов подчиняются *экспоненциальному* распределению.

Непосредственно, наблюдения начались только через 6 единиц времени, при этом 3 экземпляра оборудования уже отказали. В ходе испытаний были зафиксированы отказы четырех экземпляров оборудования с временами 7, 8, 18 и 29 единиц времени.

Пример 2 (продолжение)

МНК

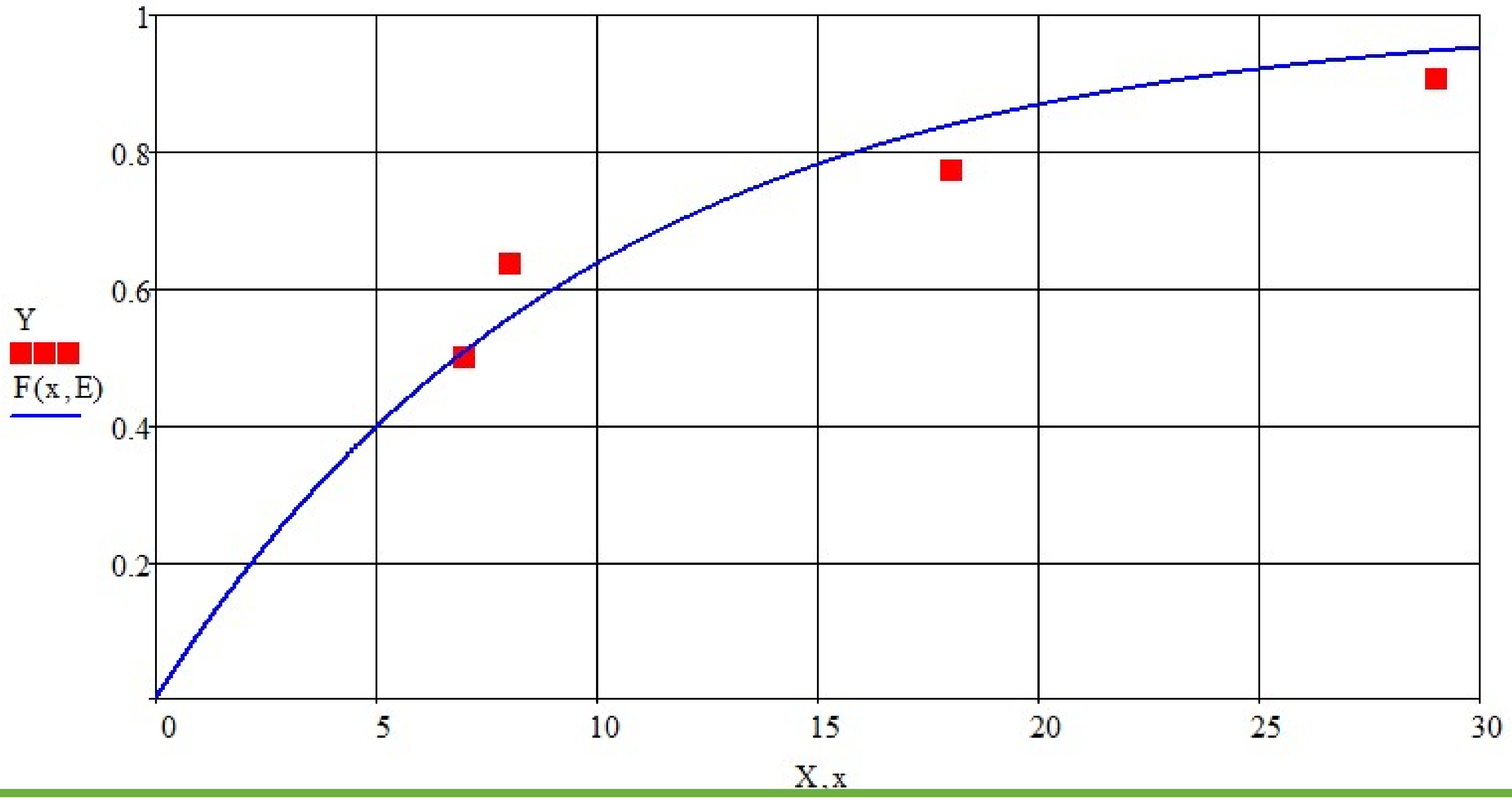
$$X = \begin{bmatrix} ? \\ ? \\ ? \\ 7 \\ 8 \\ 18 \\ 29 \end{bmatrix} \quad Y = \begin{bmatrix} 0{,}094 \\ 0{,}228 \\ 0{,}364 \\ 0{,}5 \\ 0{,}636 \\ 0{,}772 \\ 0{,}906 \end{bmatrix}$$

$$RSS(\lambda) = \sum_{i=4}^{7} [Y_i - F(X_i, \lambda)]^2$$

$$E_{МНК} = 0{,}101$$

$$T_{МНК} = 9{,}869$$

Пример 2 (продолжение)



Пример 2 (продолжение)

ММП

Обозначим количество наблюдаемых отказов через $N_F = 4$, а число экземпляров оборудования, которые отказали к моменту начала наблюдения через $N_{\mathrm{L}} = 3$.

Также, обозначим через $T_{\mathrm{L}} = 6$ момент начала наблюдения.

Пример 2 (продолжение)

Имеющиеся данные можно представить в следующем виде:

6	L	
6	L	отказы до начала наблюдений
6	L	
7	F	
8	F	
18	F	наблюдаемые отказы
29	F	

Пример 2 (продолжение)

Информацию о том, что некоторый экземпляр оборудования отказал к моменту времени T_L можно учесть через вероятность этого события:

$$\Pr[X < T_L] = F(T_L, \lambda)$$

Для использования этой информации необходимо модифицировать логарифмическую функцию правдоподобия следующим образом:

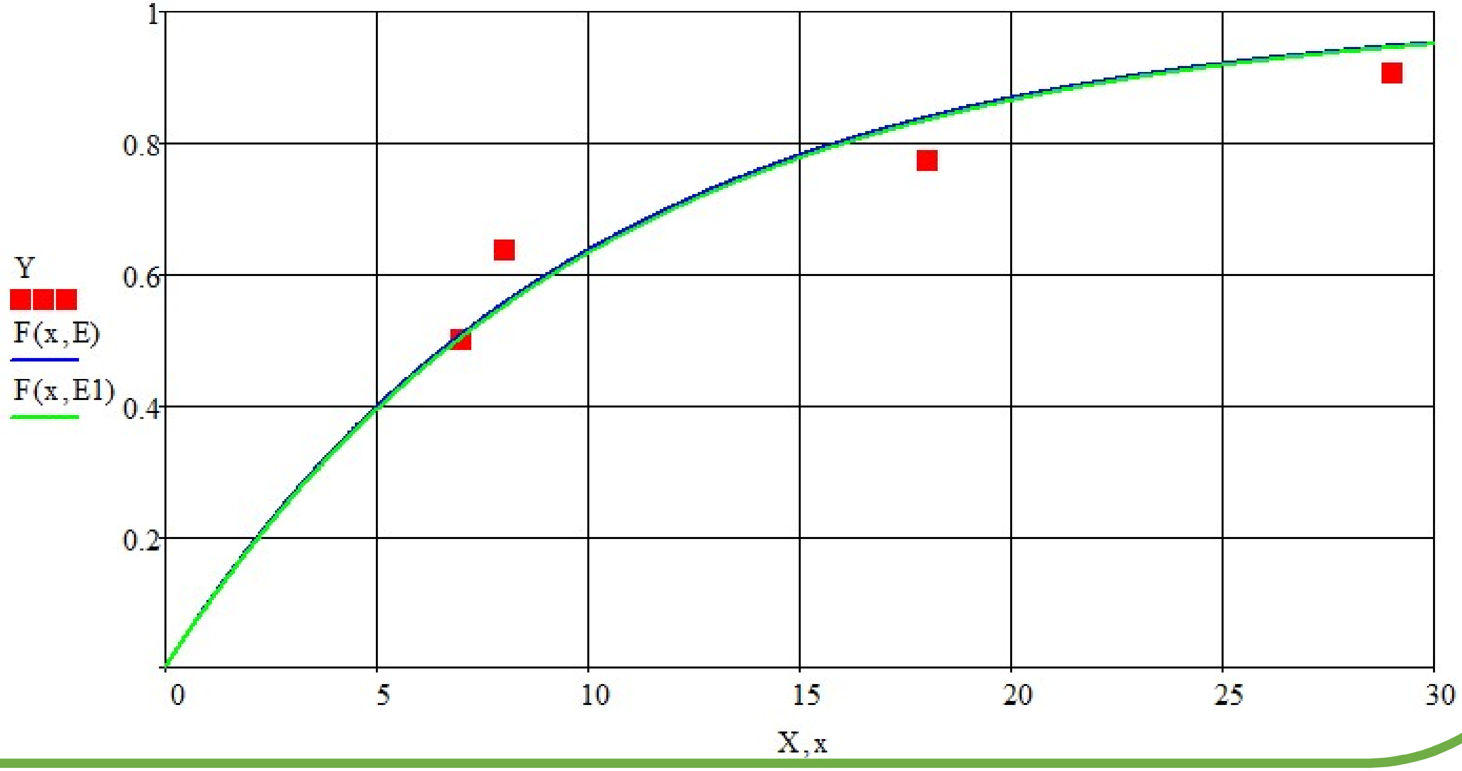
$$\Lambda(\lambda) = \sum_{i=1}^{N_F=4} \ln f(X_i, \lambda) + N_L \cdot \ln[F(T_L, \lambda)]$$

Пример 2 (продолжение)

Далее определяем максимум модифицированной функции.

$$E_{\text{ММП}} = 0{,}0998$$

$$T_{\text{ММП}} = 10{,}015$$



В общем случае, когда за различными экземплярами оборудования начинают наблюдать в разное время T_{L_j}, логарифмическая функция правдоподобия принимает вид:

$$\Lambda(\boldsymbol{\Theta}) = \sum_{i=1}^{N_F} \ln f(X_i, \boldsymbol{\Theta}) + \sum_{j=1}^{N_L} \ln\left[F\left(T_{L_j}, \boldsymbol{\Theta}\right)\right]$$

где $\boldsymbol{\Theta}$ – вектор оцениваемых параметров, X_i - времена наблюдаемых отказов, T_{L_j} - времена начала наблюдения за j-ым объектом.

Пример 3

7 экземпляров однотипного оборудования поставлены на испытания на двух промышленных площадках (3 и 4).
В ходе испытаний на первой площадке зафиксированы времена отказов: 2, 5 и 18.
На второй площадке наблюдения проводились через интервалы в 10 единиц времени. Определено, что в интервале 0-10 произошло 3 отказа, 10-20 – 0 отказов, 20-30 – 1 отказ.

Поставлена задача найти точечную оценку параметра распределения, считая, что времена отказов подчиняются *экспоненциальному* распределению.

Пример 3 (продолжение)

Имеющиеся данные можно представить в следующем виде:

2	F	
5	F	наблюдаемые отказы
18	F	
0-10	3	
10-20	0	интервальные данные
20-30	1	

Пример 3 (продолжение)

Информацию о том, что некоторый экземпляр оборудования отказал на интервале $(a; b]$ можно учесть через вероятность этого события:

$$\Pr[a < X \leq b] = F(b, \lambda) - F(a, \lambda)$$

Для использования этой информации необходимо модифицировать логарифмическую функцию правдоподобия следующим образом:

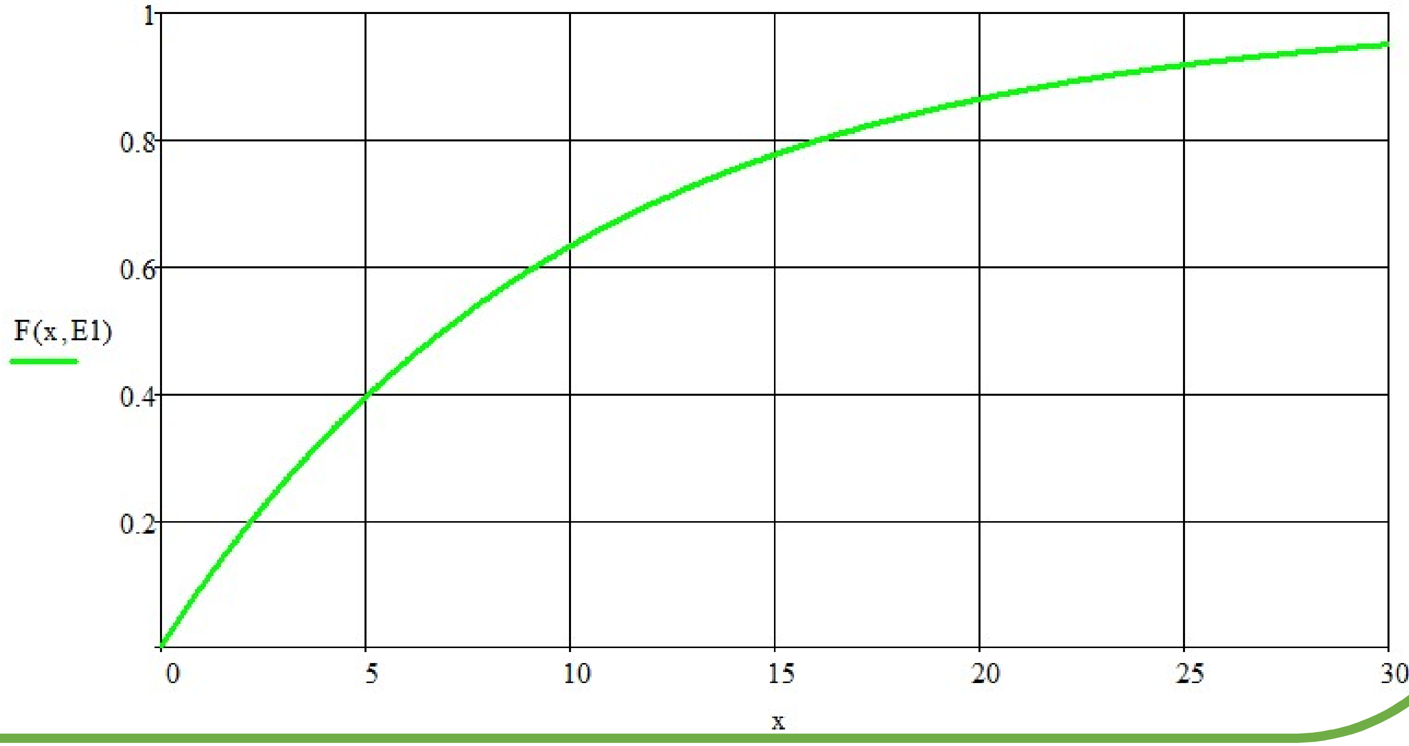
$$\Lambda(\lambda) = \sum_{i=1}^{N_F=3} \ln f(X_i, \lambda) + \sum_{j=1}^{N_I=4} \ln\left[F\left(b_j, \lambda\right) - F\left(a_j, \lambda\right)\right]$$

Пример 3 (продолжение)

Далее определяем максимум модифицированной функции.

$$E_{\text{ММП}} = 0{,}114$$

$$T_{\text{ММП}} = 8{,}753$$



Общий вид модифицированной логарифмической функции правдоподобия, учитывающей правое, левое и интервальное цензурирование имеет вид:

$$\Lambda(\lambda) = \sum_{i=1}^{N_F} \ln f(X_i, \lambda) + \sum_{j=1}^{N_R} \ln\left[1 - F\left(T_{R\,j}, \boldsymbol{\Theta}\right)\right] + \sum_{k=1}^{N_L} \ln\left[F\left(T_{L\,k}, \boldsymbol{\Theta}\right)\right] + \sum_{n=1}^{N_I} \ln\left[F(b_n, \lambda) - F(a_n, \lambda)\right]$$

наблюдаемые отказы

правое цензурирование

левое цензурирование

интервальное цензурирование

Необходимо обратить внимание на то, что **интервальное цензурирование** является наиболее общим видом цензурирования, так как при устремлении правой границы интервала к бесконечности этот вид цензурирования превращается в цензурирование справа, а при устремлении левой границы интервала к нулю — в цензурирование слева.

$$\Lambda(\lambda) = \sum_{i=1}^{N_F} \ln f(X_i, \lambda) + \sum_{j=1}^{N_I} \ln\left[F\left(b_j, \lambda\right) - F\left(a_j, \lambda\right)\right]$$

Правое: $F(\infty, \lambda) - F\left(a_j, \lambda\right) = 1 - F\left(a_j, \lambda\right)$

Левое: $F\left(b_j, \lambda\right) - F(0, \lambda) = F\left(b_j, \lambda\right) - 0$

Так называемое *цензурирование типа I* применяется в ситуациях, когда тестирование заканчивается в определенный заранее момент времени (например, мы начинаем проверять 100 ламп и заканчиваем эксперимент после некоторого фиксированного заранее момента времени).

В этом случае время цензурирования фиксировано, и число отказавших единиц (в примере, ламп) является случайной величиной (см. Пример 1).

При *цензурировании типа II* эксперимент продолжается до того момента, когда отказжет (выйдет из строя) фиксированное заранее число проверяемых единиц (например, эксперимент заканчивается только после выхода из строя 50 ламп).

В этом случае число отказавших элементов фиксировано, а время является случайной величиной.

При оценке параметров статистических моделей надежности с использованием ММП по цензурированным (I и II тип) выборкам логарифмическая функция правдоподобия должна быть изменена следующим образом:

$$\Lambda(\Theta) = \sum_{i=1}^{N} \ln f(T_i, \Theta) + M \cdot \ln\left[1 - F\left(T_{\text{ценз}}, \Theta\right)\right],$$

где N – число наблюдаемых отказов; $N + M$ – объем выборки; T_i – времена наблюдаемых отказов; $T_{\text{ценз}}$ - время цензурирования; Θ – оцениваемый параметр (вектор параметров).

Для цензурирования типа I $T_{\text{ценз}}$ задается заранее.
Для цензурирования типа II $T_{\text{ценз}} = \max T_i$.

Цензурирование в совокупности с ММП предоставляет еще одну возможность анализа.

Пример 4

Пусть имеется система, которая может отказать по двум причинам: A и B. Необходимо по эксплуатационным данным оценить параметры распределения Вейбулла для отказов по каждой причине.

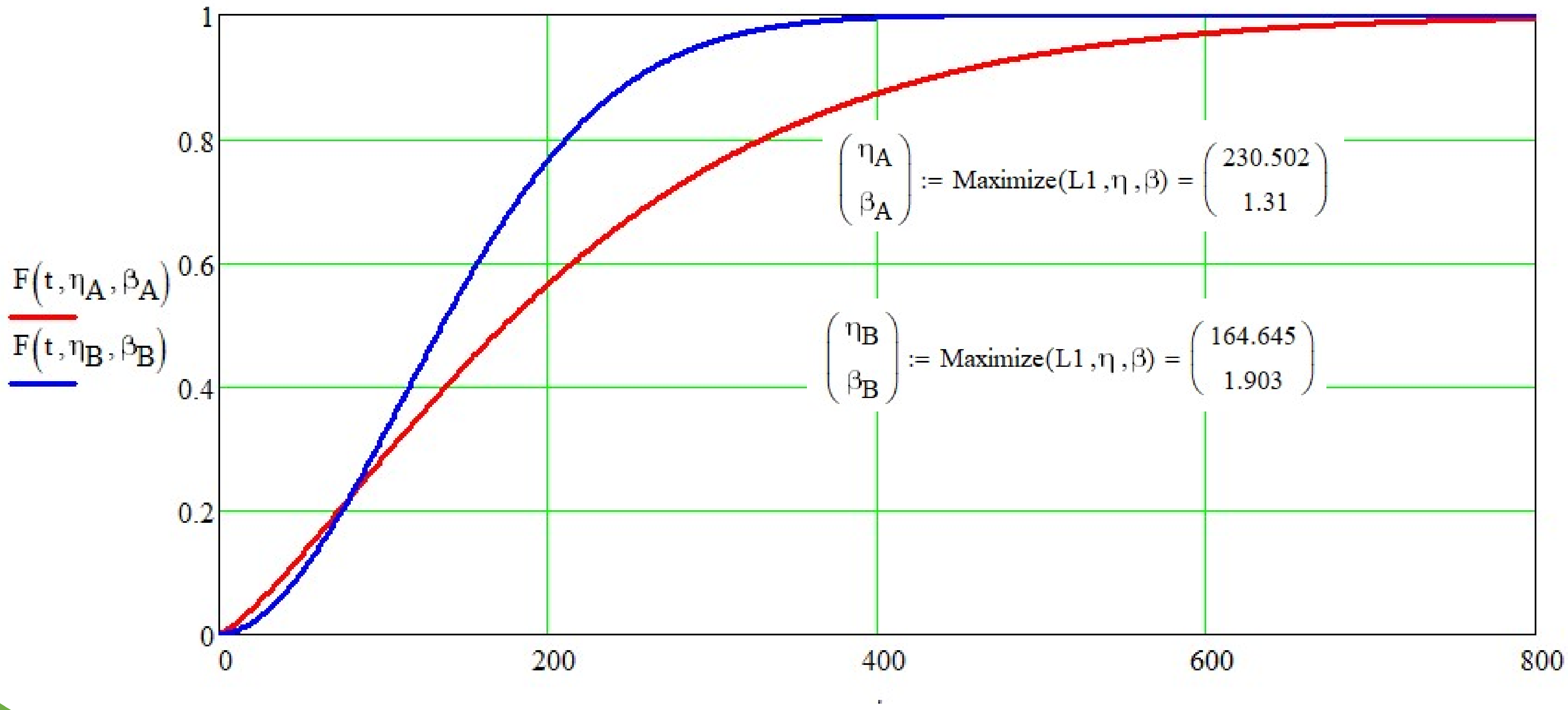
Предположим, имеется выборка из $N = 20$ времен отказов оборудования, причем каждому отказу соответствует одна из причин.

Пример 4 (продолжение)

Время	Причина	Время	Причина	Время	Причина	Время	Причина
8	A	55	B	93	B	167	B
40	B	57	B	114	B	169	B
41	A	65	B	130	B	198	B
42	A	71	A	135	B	226	A
49	A	90	B	148	A	263	A

При определении оценок параметров для причины A будем считать времена отказов по причине B временами правого цензурирования, и наоборот.

Пример 4 (продолжение)



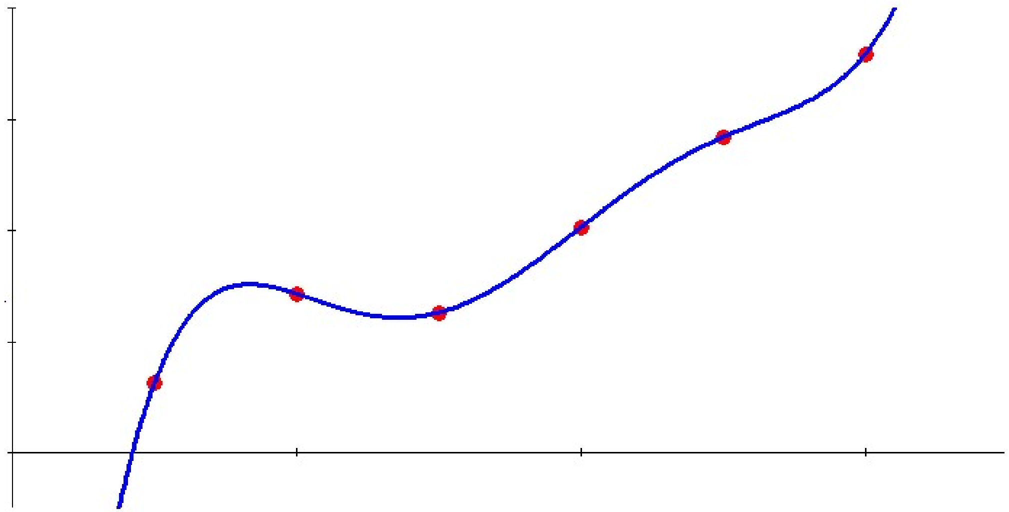
Истинное распределение времени отказа сложной системы с большим количеством компонентов, как правило, неизвестно.

В практических целях выбор статистической модели надежности подразумевает поиск такого закона распределения, который с высокой точностью позволяет описать имеющуюся выборку данных об отказах, одновременно оставаясь относительно простым, т.е. имеющим небольшое количество параметров.

Математический подход к выбору модели позволяет сделать выбор одной из нескольких моделей-кандидатов из заранее определенного множества.

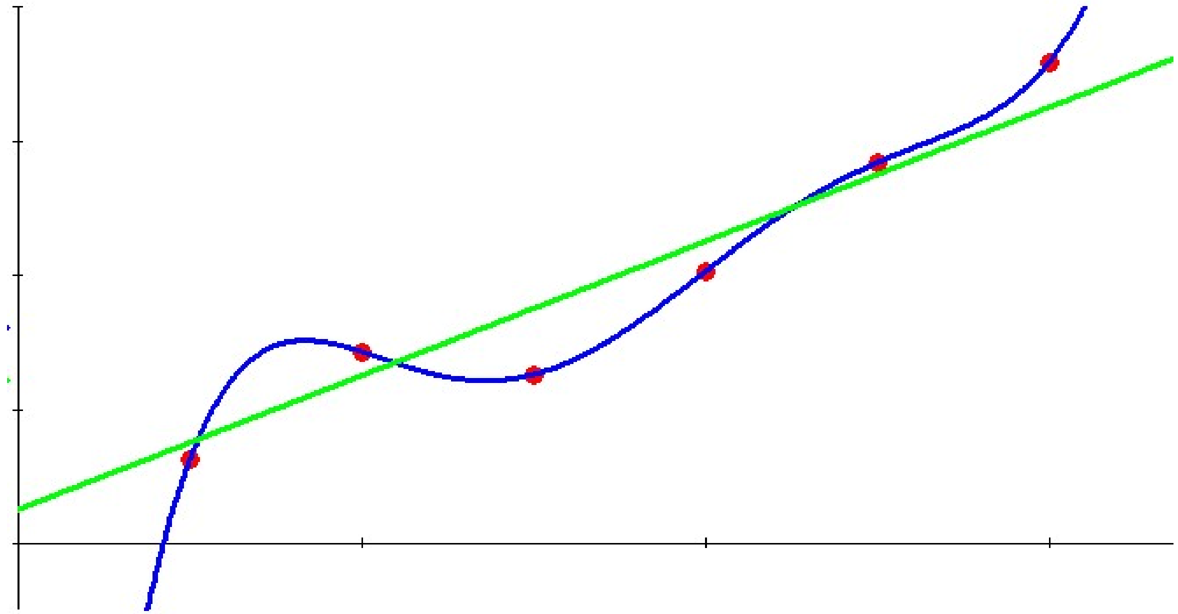
Более сложные модели способны за счет большого количества параметров значительно менять форму функциональных зависимостей.

Например, полином пятого порядка может в точности пройти по шести точкам.

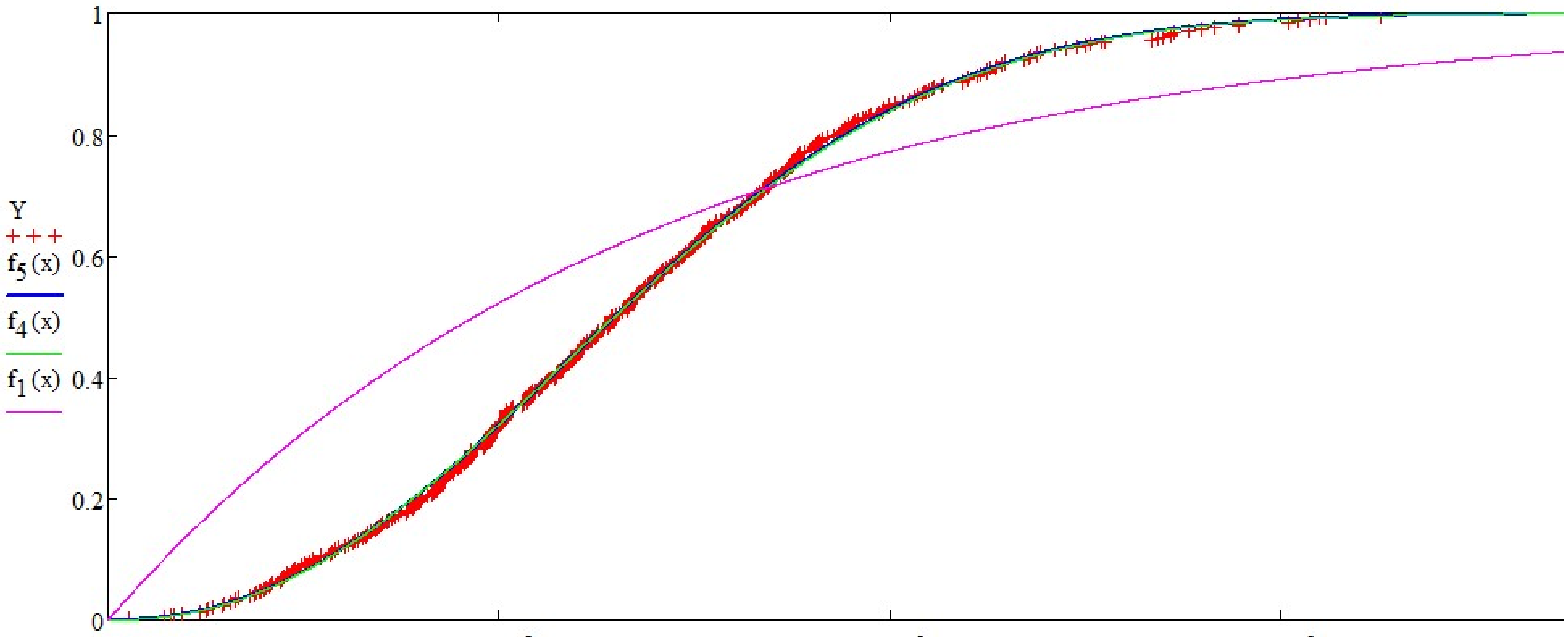


Однако, стремление повысить описательную точность может привести к тому, что модель будет хорошо описывать имеющуюся выборку, и будет непригодна для новых данных.

Модель не должна принимать случайный шум за полезные данные!



Иногда подходящую модель можно определить визуально, сравнивая эмпирическую функцию распределения, полученную по выборке, с функциями распределения моделей-кандидатов. Однако, это не всегда возможно.



Наиболее прямолинейным способом математического выбора модели является выбор модели, обладающей максимальным значением функции правдоподобия.

Пусть $\Lambda_1(\widehat{\Theta}_1)$ и $\Lambda_2(\widehat{\Theta}_2)$ - значения логарифмических функций правдоподобия, вычисленные для двух моделей-кандидатов с оценёнными векторами параметров $\widehat{\Theta}_1$ и $\widehat{\Theta}_2$.

Следует отдать предпочтение модели, имеющей наибольшее значение Λ (или наименьшее значение -2Λ).

Хотя такой подход учитывает согласованность выбранной модели с данными, он не принимает во внимание сложность модели.

Как правило, более сложная модель будет иметь лучшее значение Λ или -2Λ, нежели простая.

Мы можем найти баланс между точностью и простотой модели, используя информационные критерии.

Первым таким критерием был информационный критерий Акаике (*AIC*).

Значение критерия рассчитывается как

$$AIC = 2p - 2\hat{\Lambda},$$

где p – число параметров модели, $\hat{\Lambda}$ - значение логарифмической функции правдоподобия, вычисленное для найденных оценок параметров.

Поскольку p – всегда положительно, большое количество параметров увеличивает (ухудшает) значение $-2\hat{\Lambda}$.

Можно рассматривать величину $2p$ как «штраф» за количество параметров.

Отметим, что значение AIC показывает лишь относительное качество модели в сравнении с другими моделями-кандидатами, и ничего не говорит об абсолютном качестве.

Таким образом, если все модели-кандидаты плохо подходят к имеющимся данным, AIC все равно выберет какую-то модель как лучшую.

Следовательно, после выбора модели по методу AIC, необходимо проводить проверку.

Идея, лежащая в основе метода AIC, была использована другими исследователями, которые предложили другие выражения для вычисления «штрафа».

Байесовский информационный критерий (критерий Шварца):

$$BIC = p \ln N - 2\widehat{\Lambda}$$

Информационный критерий Ханнана-Квина:

$$HQIC = 2p \ln(\ln N) - 2\widehat{\Lambda}$$

где N – объем выборки.

Еще одним подходом к выбору модели является применение различных статистических критериев согласия.

Критерии согласия - это критерии проверки гипотез о соответствии эмпирического распределения теоретическому распределению вероятностей. Такие критерии подразделяются на два класса:

- *Общие критерии согласия* применимы к самой общей формулировке гипотезы, а именно к гипотезе о согласии наблюдаемых результатов с любым априорно предполагаемым распределением вероятностей.
- *Специальные критерии согласия* предполагают специальные нулевые гипотезы, формулирующие согласие с определенной формой распределения вероятностей.

Общие критерии согласия

Простая нулевая гипотеза $H_0 \colon F_n(x) = F_0(x, \boldsymbol{\Theta})$, где $F_n(x)$ - эмпирическая функция распределения вероятностей для выборки объема n; $F_0(x, \boldsymbol{\Theta})$ - известная функция распределения вероятностей, с которой проверяется согласие наблюдаемой выборки, а $\boldsymbol{\Theta}$ - известное значение параметра (скалярного или векторного)

Группы общих критериев согласия:

- критерии, основанные на изучении разницы между теоретической плотностью распределения и эмпирической гистограммой;
- критерии, основанные на расстоянии между теоретической и эмпирической функциями распределения вероятностей.

Критерий согласия Пирсона (критерий согласия χ^2) — непараметрический метод, позволяющий оценить значимость различий между фактическим количеством исходов, попадающих в каждую категорию, и теоретическим количеством, которое можно ожидать в изучаемых группах при справедливости нулевой гипотезы.

Выражаясь проще, метод позволяет оценить статистическую значимость различий двух или нескольких относительных показателей (частот, долей).

Критерий хи-квадрат был разработан и предложен в 1900 году основателем математической статистики английским учёным Карлом Пирсоном.

Процедура проверки гипотез с использованием критерия χ^2 предусматривает группирование наблюдений.

Область определения случайной величины разбивают на k непересекающихся интервалов граничными точками

$$x_{(0)}, x_{(1)}, \ldots, x_{(k-1)}, x_{(k)}$$

где $x_{(0)}$ — нижняя грань области определения случайной величины; $x_{(k)}$ — верхняя грань.

В соответствии с заданным разбиением подсчитывают число n_i выборочных значений, попавших в i-й интервал, и вероятности попадания в интервал

$$p_i = F\left(x_{(i)}, \boldsymbol{\Theta}\right) - F\left(x_{(i-1)}, \boldsymbol{\Theta}\right),$$

соответствующие теоретическому закону с функцией распределения $F(x, \boldsymbol{\Theta})$.

При этом

$$n = \sum_{i=1}^{k} n_i\,, \qquad \sum_{i=1}^{k} p_i = 1.$$

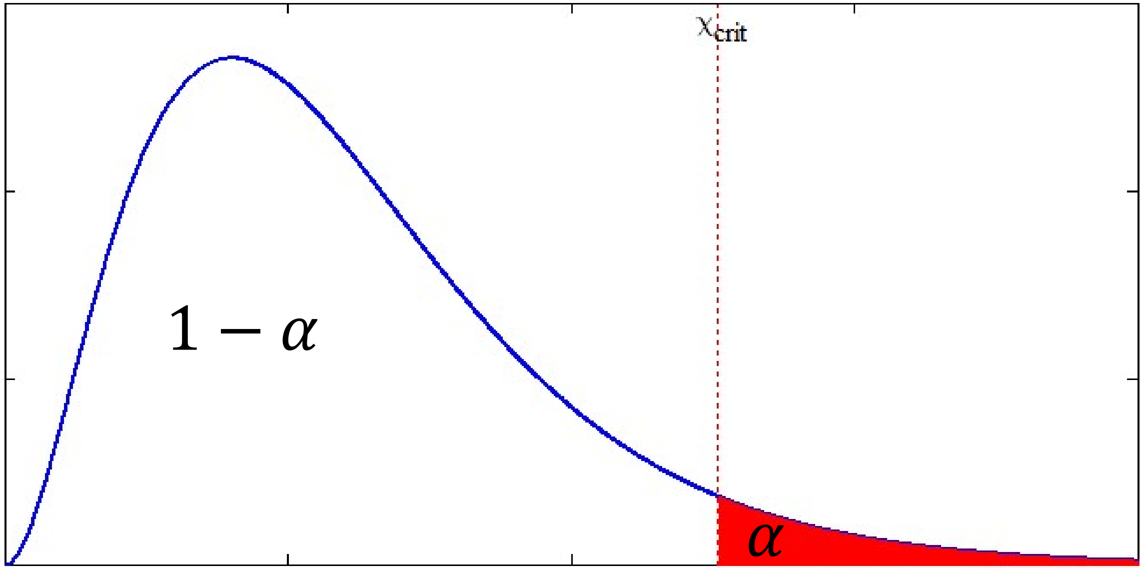
Статистику критерия согласия χ^2 Пирсона вычисляют по формуле

$$\chi^2 = n \cdot \sum_{i=1}^{k} \frac{(n_i/n - p_i)^2}{p_i}.$$

В случае проверки простой гипотезы в пределе при $n \to \infty$ эта статистика подчиняется χ_r^2-распределению с $r = k - 1$ степенями свободы, если верна нулевая гипотеза.

Рекомендуется использовать $k \geq m + 2$ равновероятных интервалов, где m – количество параметров распределения. При этом в каждом интервале должно быть не менее 5-10 выборочных значений.

Для выбранного уровня значимости α, нулевая гипотеза отвергается, если рассчитанное значение статистики превышает критическое значение $\chi^2_{r;1-\alpha}$.



Иначе, нет оснований отклонить нулевую гипотезу.

Классический критерий Колмогорова (иногда говорят Колмогорова-Смирнова) предназначен для проверки простых гипотез о принадлежности анализируемой выборки некоторому полностью известному закону распределения.

В критерии Колмогорова в качестве расстояния между эмпирическим и теоретическим законом используется величина

$$D_n = \sup_{|x|<\infty} |F_n(x) - F_0(x, \boldsymbol{\Theta})|.$$

При проверке гипотез обычно используется статистика вида

$$S_K = \frac{6nD_n + 1}{6\sqrt{n}};$$

где $D_n = \max(D_n^+; D_n^-)$;

$$D_n^+ = \max_{1 \le i \le n} \left\{ \frac{i}{n} - F_0(x_i, \boldsymbol{\Theta}) \right\}; \qquad D_n^- = \max_{1 \le i \le n} \left\{ F_0(x_i, \boldsymbol{\Theta}) - \frac{i-1}{n} \right\};$$

$\boldsymbol{n}$ – объем выборки;

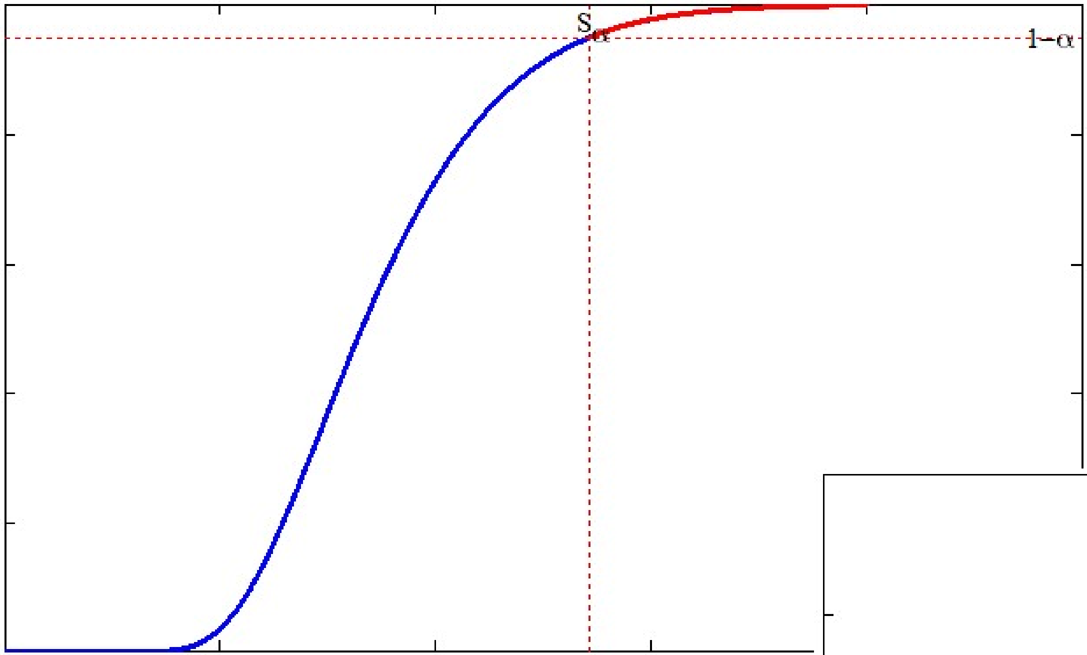
$x_1, x_2, \dots, x_n$ - упорядоченные по возрастанию выборочные значения.

Значения функции распределения Колмогорова можно рассчитать по следующим выражениям:

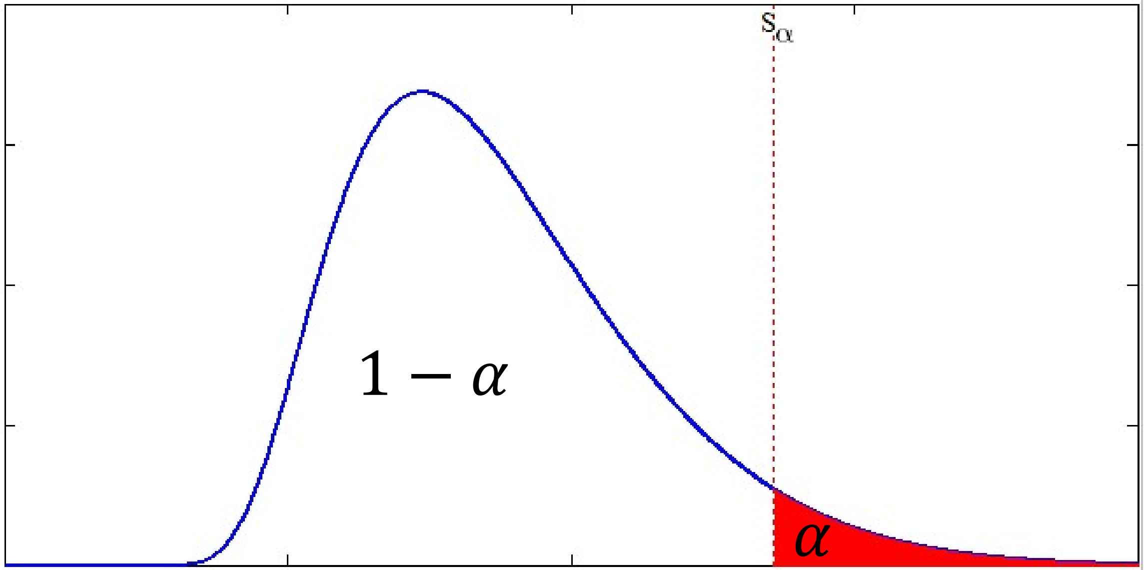
$$F_K(x) = 1 - 2\sum_{k=1}^{\infty}(-1)^{k-1}e^{-2k^2x^2} = \frac{\sqrt{2\pi}}{x}\sum_{k=1}^{\infty}e^{\frac{-(2k-1)^2\pi^2}{8x^2}}$$

Для выбранного уровня значимости α, нулевая гипотеза отвергается, если рассчитанное значение статистики превышает критическое значение S_α, определяемое из соотношения

$$F_K(S_\alpha) = 1 - \alpha.$$



S$_\alpha$
1−α



S$_\alpha$
$1-\alpha$
α

Критерии согласия Кра́мера-фон Ми́зеса-Смирнова и А́ндерсона-Да́рлинга, как и критерий Колмогорова, основаны на расстоянии между теоретической и эмпирической функциями распределения вероятностей.

Статистика Кра́мера-фон Ми́зеса-Смирнова рассчитывается по формуле:

$$S_\omega = \frac{1}{12n} + \sum_{i=1}^{n} \left\{ F_0(x_i, \boldsymbol{\theta}) - \frac{2i-1}{2n} \right\}^2$$

Статистика А́ндерсона-Да́рлинга рассчитывается по формуле:

$$S_\Omega = -n - 2\sum_{i=1}^{n} \left\{ \frac{2i-1}{2n} \ln F_0(x_i, \boldsymbol{\theta}) + \left(1 - \frac{2i-1}{2n}\right) \ln\left(1 - F_0(x_i, \boldsymbol{\theta})\right) \right\}$$

Для выбранного уровня значимости α, нулевая гипотеза отвергается, если рассчитанные значения статистик превышают соответствующие критические значения, которые можно найти в таблицах.

В случае простых гипотез предельные распределения статистик рассматриваемых критериев согласия ***не зависят от вида наблюдаемого закона распределения и от его параметров***.

Говорят, что эти критерии являются **«свободными от распределения»**. Это достоинство предопределило их широкое использование в приложениях.

Исходя из целей конкретной задачи, возможно определить наилучшую модель, используя более простые критерии, как то:

- наименьшее максимальное отклонение $|\Delta_{max}| \to min$;
- наименьшая среднеквадратичная ошибка $\varepsilon^2 = \frac{1}{N}\sum \varepsilon_i^2 \to min$;
- наименьшее отклонение математического ожидания и дисперсии в соответствии с моделью от выборочных моментов.

В последнем случае часто вместо отклонения используют относительную погрешность:

$$\delta = \frac{|x - \hat{x}|}{x} \times 100\%$$

(1 – α)-квантили χ^2-распределения с r степенями свободы

r	α										
	0.5	0.4	0.3	0.2	0.1	0.05	0.025	0.01	0.005	0.001	0.0005
1	0.455	0.708	1.074	1.642	2.706	3.841	5.024	6.635	7.879	10.828	12.116
2	1.386	1.833	2.408	3.219	4.605	5.991	7.378	9.210	10.597	13.816	15.202
3	2.366	2.946	3.665	4.642	6.251	7.815	9.248	11.345	12.838	16.266	17.730
4	3.357	4.045	4.878	5.989	7.779	9.488	11.143	13.277	14.860	18.467	19.997
5	4.351	5.132	6.064	7.289	9.236	11.070	12.832	15.086	16.750	20.515	22.105
6	5.348	6.211	7.231	8.558	10.645	12.592	14.449	16.812	18.548	22.458	24.103
7	6.346	7.283	8.383	9.803	12.017	14.067	16.013	18.475	20.278	24.322	26.018
8	7.344	8.351	9.524	11.030	13.362	15.507	17.535	20.090	21.955	26.125	27.868
9	8.343	9.414	10.656	12.242	14.684	16.919	19.023	21.666	23.589	27.877	29.666
10	9.342	10.473	11.781	13.442	15.987	18.307	20.483	23.209	25.188	29.588	31.420
11	10.342	11.530	12.899	14.631	17.275	19.675	21.920	24.725	26.757	31.264	33.136
12	11.340	12.584	14.011	15.812	18.549	21.026	23.336	26.217	28.300	32.909	34.821
13	12.340	13.636	15.119	16.985	19.812	22.362	24.736	27.688	29.819	34.528	36.478
14	13.339	14.685	16.222	18.151	21.064	23.685	26.119	29.141	31.319	36.123	38.109
15	14.339	15.733	17.322	19.311	22.307	24.996	27.488	30.578	32.801	37.697	39.719
16	15.338	16.780	18.418	20.465	23.542	26.296	28.845	32.000	34.267	39.252	41.308

Функция распределения статистики Колмогорова $K(S)$ при проверке простой гипотезы

S	0	1	2	3	4	5	6	7	8	9
0.2	0.000000	000 000	000 000	000 000	000 000	000 000	000 000	000 000	000 001	000 004
0.3	0.000009	000 021	000 046	000 091	000 171	000 303	000 511	000 826	001 285	001 929
0.4	0.002808	003 972	005 476	007 377	009 730	012 589	016 005	020 022	024 682	030 017
0.5	0.036055	042 814	050 306	058 534	067 497	077 183	087 577	098 656	110 394	122 760
0.6	0.135718	149 229	163 255	177 752	192 677	207 987	223 637	239 582	255 780	272 188
0.7	0.288765	305 471	322 265	339 114	355 981	372 833	389 640	406 372	423 002	439 505
0.8	0.455858	472 039	488 028	503 809	519 365	534 682	549 745	564 545	579 071	593 315
0.9	0.607269	620 928	634 285	647 337	660 081	672 515	684 836	696 445	707 941	719 126
1.0	0.730000	740 566	750 825	760 781	770 436	779 794	788 860	797 637	806 130	814 343
1.1	0.822282	829 951	837 356	844 502	851 395	858 040	864 443	870 610	876 546	882 258
1.2	0.887750	893 030	898 102	903 973	907 648	912 134	916 435	920 557	924 506	928 288
1.3	0.931908	935 371	938 682	941 847	944 871	947 758	950 514	953 144	955 651	958 041
1.4	0.960318	962 487	964 551	966 515	968 383	970 159	971 846	973 448	974 969	976 413
1.5	0.977782	979 080	980 310	981 475	982 579	983 623	984 610	985 544	986 427	987 261
1.6	0.988048	988 791	989 492	990 154	990 777	991 364	991 917	992 438	992 928	993 389
1.7	0.993823	994 230	994 612	994 972	995 309	995 625	995 922	996 200	996 460	996 704
1.8	0.996932	997 146	997 346	997 533	997 707	997 870	998 023	998 165	998 297	998 421
1.9	0.998536	998 644	998 744	998 837	998 924	999 004	999 079	999 149	999 213	999 273
2.0	0.999329	999 381	999 429	999 473	999 514	999 553	999 588	999 620	999 651	999 679
2.1	0.999705	999 728	999 750	999 771	999 790	999 807	999 823	999 837	999 851	999 863
2.2	0.999874	999 886	999 895	999 904	999 912	999 920	999 927	999 933	999 939	999 944
2.3	0.999949	999 954	999 958	999 961	999 965	999 968	999 971	999 974	999 976	999 978
2.4	0.999980	999 982	999 984	999 985	999 987	999 988	999 989	999 990	999 991	999 992

Процентные точки распределения статистики Колмогорова при проверке простой гипотезы

Функция распределения	Верхние процентные точки				
	0.15	0.1	0.05	0.025	0.01
$K(S)$	1.1379	1.2238	1.3581	1.4802	1.6276

Функция распределения статистики ω^2 Крамера–Мизеса–Смирнова $a1(S)$ при проверке простой гипотезы

S	0	1	2	3	4	5	6	7	8	9
0.0	0.00000	00 001	00 300	02 568	06 685	12 372	18 602	24 844	30 815	36 386
0.1	0.41513	46 196	50 457	54 329	57 846	61 042	63 951	66 600	69 019	71 229
0.2	0.73253	75 109	76 814	78 383	79 829	81 163	82 396	83 536	84 593	85 573
0.3	0.86483	87 329	88 115	88 848	89 531	90 167	90 762	91 317	91 836	92 321
0.4	0.92775	93 201	93 599	93 972	94 323	94 651	94 960	95 249	95 521	95 777
0.5	0.96017	96 242	96 455	96 655	96 843	97 020	97 186	97 343	97 491	97 630
0.6	0.97762	97 886	98 002	98 112	98 216	98 314	98 406	98 493	98 575	98 653
0.7	0.98726	98 795	98 861	98 922	98 981	99 036	99 088	99 137	99 183	99 227
0.8	0.99268	99 308	99 345	99 380	99 413	99 444	99 474	99 502	99 528	99 553
0.9	0.99577	99 599	99 621	99 641	99 660	99 678	99 695	99 711	99 726	99 740
1.0	0.99754	99 764	99 776	99 787	99 799	99 812	99 820	99 828	99 837	99 847
1.1	0.99856	99 862	99 869	99 876	99 883	99 890	99 895	99 900	99 905	99 910
1.2	0.99916	99 919	99 923	99 927	99 931	99 935	99 938	99 941	99 944	99 947
1.3	0.99950	99 953	99 955	99 957	99 959	99 962	99 964	99 965	99 967	99 969
1.4	0.99971	99 972	99 973	99 975	99 976	99 978	99 978	99 979	99 980	99 980

Процентные точки распределения статистики ω^2 Крамера–Мизеса–Смирнова при проверке простой гипотезы

Функция распределения	Верхние процентные точки				
	0.15	0.1	0.05	0.025	0.01
$a1(S)$	0.2841	0.3473	0.4614	0.5806	0.7434

Функция распределения статистики Ω^2 Андерсона–Дарлинга $a2(S)$ при проверке простой гипотезы

S	0	1	2	3	4	5	6	7	8	9
0.0	0.00000	00 000	00 000	00 000	00 000	00 000	00 000	00 000	00 000	00 001
0.1	0.00003	00 008	00 020	00 043	00 081	00 141	00 228	00 349	00 508	00 710
0.2	0.00959	01 256	01 605	02 005	02 457	02 961	03 514	04 115	04 762	05 453
0.3	0.06184	06 954	07 759	08 596	09 463	10 356	11 273	12 211	13 168	14 140
0.4	0.15127	16 124	17 132	18 146	19 166	20 190	21 217	22 244	23 271	24 296
0.5	0.25319	26 337	27 351	28 359	29 360	30 355	31 342	32 320	33 290	34 250
0.6	0.35200	36 141	37 071	37 991	38 900	39 798	40 684	41 560	42 424	43 277
0.7	0.44118	44 947	45 765	46 572	47 367	48 150	48 922	49 683	50 432	51 170
0.8	0.51897	52 613	53 318	54 012	54 695	55 368	56 030	56 682	57 324	57 956
0.9	0.58577	59 189	59 791	60 383	60 966	61 540	62 104	62 660	63 206	63 744
1.0	0.64273	64 794	65 306	65 811	66 307	66 795	67 275	67 748	68 213	68 670
1.1	0.69120	69 563	69 999	70 428	70 851	71 266	71 675	72 077	72 473	72 863
1.2	0.73247	73 624	73 996	74 361	74 721	75 075	75 424	75 767	76 105	76 438
1.3	0.76765	77 088	77 405	77 717	78 025	78 328	78 626	78 919	79 209	79 493
1.4	0.79773	80 049	80 321	80 589	80 852	81 112	81 368	81 620	81 868	82 112
1.5	0.82352	82 589	82 823	83 053	83 279	83 503	83 723	83 939	84 153	84 363
1.6	0.84570	84 774	84 975	85 173	85 369	85 561	85 751	85 938	86 122	86 303

S	0	1	2	3	4	5	6	7	8	9
1.7	0.86482	86 659	86 832	87 004	87 173	87 339	87 503	87 665	87 824	87 981
1.8	0.88136	88 289	88 439	88 588	88 734	88 878	89 021	89 161	89 299	89 435
1.9	0.89570	89 703	89 833	89 962	90 089	90 215	90 338	90 460	90 581	90 699
2.0	0.90816	90 932	91 046	91 158	91 269	91 378	91 486	91 592	91 697	91 800
2.1	0.91902	92 003	92 102	92 200	92 297	92 392	92 486	92 579	92 671	92 761
2.2	0.92851	92 939	93 025	93 111	93 196	93 279	93 361	93 443	93 523	93 602
2.3	0.93680	93 757	93 833	93 908	93 983	94 056	94 128	94 199	94 269	94 339
2.4	0.94407	94 475	94 542	94 608	94 673	94 737	94 800	94 863	94 925	94 986
2.5	0.95046	95 105	95 164	95 222	95 279	95 336	95 391	95 446	95 501	95 554
2.6	0.95607	95 660	95 711	95 762	95 813	95 862	95 912	95 960	96 008	96 055
2.7	0.96102	96 148	96 194	96 239	96 283	96 327	96 370	96 413	96 455	96 497
2.8	0.96538	96 579	96 619	96 659	96 698	96 737	96 775	96 813	96 850	96 887
2.9	0.96923	96 959	96 995	97 030	97 064	97 099	97 132	97 166	97 199	97 231
3.0	0.97263	97 295	97 327	97 358	97 388	97 419	97 449	97 478	97 507	97 536
3.1	0.97565	97 593	97 621	97 648	97 675	97 702	97 729	97 755	97 781	97 806
3.2	0.97831	97 856	97 881	97 905	97 929	97 953	97 977	98 000	98 023	98 046
3.3	0.98068	98 090	98 112	98 134	98 155	98 176	98 197	98 217	98 238	98 258
3.4	0.98278	98 297	98 317	98 336	98 355	98 374	98 392	98 410	98 429	98 447
3.5	0.98464	98 482	98 499	98 516	98 533	98 549	98 566	98 582	98 598	98 614
3.6	0.98630	98 645	98 660	98 676	98 691	98 705	98 720	98 734	98 749	98 763
3.7	0.98777	98 791	98 804	98 818	98 831	98 844	98 857	98 870	98 883	98 895
3.8	0.98908	98 920	98 932	98 944	98 956	98 968	98 979	98 991	99 002	99 013
3.9	0.99024	99 035	99 046	99 057	99 067	99 078	99 088	99 098	99 108	99 118
4.0	0.99128	99 221	99 303	99 377	99 442	99 501	99 553	99 600	99 642	99 679
5.0	0.99713	99 742	99 769	99 793	99 814	99 834	99 851	99 866	99 880	99 892
6.0	0.99903	99 913	99 922	99 930	99 937	99 944	99 949	99 954	99 959	99 963
7.0	0.99967	99 970	99 973	99 976	99 978	99 981	99 983	99 984	99 986	99 987
8.0	0.99989	99 990	99 991	99 992	99 993	99 993	99 994	99 995	99 995	99 996
9.0	0.99996	–	–	–	–	–	–	–	–	–

Процентные точки распределения статистики Ω^2 Андерсона–Дарлинга при проверке простой гипотезы

Функция распределения	Верхние процентные точки				
	0.15	0.1	0.05	0.025	0.01
$a2(S)$	1.6212	1.9330	2.4924	3.0775	3.8781

Таблицы приводятся по

Лемешко Б.Ю.
Статистический анализ данных, моделирование и исследование вероятностных закономерностей. Компьютерный подход : монография / Б.Ю. Лемешко, С.Б. Лемешко, С.Н. Постовалов, Е.В. Чимитова. – Новосибирск : Изд-во НГТУ, 2011. – 888 с. (серия «Монографии НГТУ»).