Детали мехатронных модулей и роботов, их конструирование, диагностика и надежность

Лабораторная работа №5

Оценки параметров моделей надежности систем. Выбор моделей надежности

Разработал: А.А. Ефремов

Томский политехнический университет, 2021

Теоретические сведения

Теория оценивания – это раздел статистики, занимающийся оценкой параметров на основе эмпирических данных.

Параметры отражают основополагающие физические свойства объектов таким образом, что их значения влияют на распределение измеряемых данных.

Оценка (результат процедуры оценивания) – это попытка аппроксимации неизвестных параметров на основе измеренных данных.

Теоретические сведения

Различают *точечные* и *интервальные* оценки.

Точечные оценки – это числовые значения, вычисленные по результатам измерений (наблюдений).

Интервальные оценки ставят в соответствие оцениваемому параметру некоторый интервал, в котором с некоторой, заранее заданной вероятностью, находится истинное значение параметра.

Теоретические сведения

Основными методами точечного оценивания параметров являются

- метод наименьших квадратов;
- метод максимального правдоподобия.

Менее популярным является метод моментов.

Метод наименьших квадратов (МНК) является стандартным подходом к аппроксимации решения переопределенных систем, т.е. систем уравнений, в которых число неизвестных меньше числа уравнений.

Общее решение минимизирует сумму квадратов отклонений наблюдаемых значений от теоретических.

Задача МНК подразделяется на две категории:

- линейный МНК;
- нелинейный МНК,

в зависимости от того, являются ли отклонения линейными по всем неизвестным.

Линейная задача МНК имеет решение в явном виде. Нелинейные задачи зачастую решаются с помощью итеративного приближения. На каждом шаге система линеаризуется, что позволяет использовать те же основные вычисления, что и для линейного случая.

Цель МНК состоит в настройке параметров модели таким образом, что моделируемая функция наилучшим образом описывает имеющиеся данные.

Пусть имеется набор данных, состоящий из n пар (x_i, y_i), $i = 1, \dots, n$.

Моделируемая функция распределения имеет вид $F(x, \Theta)$, где m неизвестных параметров составляют вектор Θ.

Необходимо найти значения оценок $\hat{\theta}_j$, $j = 1, \dots, m$.

Близость модели к точке данных определяется отклонением, т.е. разностью между измеренным значением и значением, предсказанным согласно модели:

$$r\left(\widehat{\Theta}\right)_i = y_i - F\left(x_i, \widehat{\Theta}\right)$$

Процедура МНК минимизирует функцию суммы квадратов отклонений:

$$S\left(\widehat{\Theta}\right) = \sum_{i=1}^{n} r\left(\widehat{\Theta}\right)_i^2 \quad \rightarrow \quad \frac{\partial S}{\partial \hat{\theta}_j} = 0, \qquad j = 1,2,\ldots,m$$

<u>Теоретические сведения</u>

При анализе данных об отказах координаты x представляют собой времена отказов оборудования. Однако, для выполнения МНК нам необходимы и координаты y – вероятности отказов.

В качестве координат y можно использовать значения эмпирической функции распределения, однако это может исказить результат.

Основным методом является использование <u>медианных рангов</u>.

Теоретические сведения

Медианный ранг – это значение, которое имела бы истинная функция вероятности отказа $F\left(x_j\right)$ в момент j-го отказа из выборки n элементов с 50% доверительным уровнем.

Точные значения медианных рангов $MR_j \equiv y_j$ можно получить, решив уравнения

$$\sum_{k=j}^{n} \binom{n}{k} MR_j^k \left(1 - MR_j\right)^{n-k} = \frac{1}{2}$$

для всех j.

Теоретические сведения

Более простым способом получения значений медианных рангов является использование следующей формулы:

$$MR_j = \frac{1}{1 + \frac{n - j + 1}{j} \cdot F^{-1}(0.5, m, n)}$$

где $m = 2 \cdot (n - j + 1)$, $n = 2j$, и $F^{-1}(0.5, m, n)$ - значение обратной функции *F*-распределения Фишера с m и n степенями свободы, вычисленное в точке 0.5.

Еще одним быстрым, но менее точным способом является использование оценки Филлибена:

$$MR_j = \begin{cases} 1 - 0.5^{1/n} & if\ j = 1; \\ \dfrac{j - 0.3175}{n + 0.365} & if\ j = 2,3, \ldots, n - 1; \\ 0.5^{1/n} & if\ j = n. \end{cases}$$

или аппроксимации Бенарда:

$$MR_j = \frac{j - 0.3}{n + 0.4}$$

Пример:

Пусть известны времена отказов шести идентичных объектов: 93, 34, 16, 120, 53 и 75 часов. Необходимо вычислить медианные ранги различными методами.

Упорядочим времена отказов в восходящем порядке:

Время до отказа t_j, час	Порядковый номер отказа
16	1
34	2
53	3
75	4
93	5
120	6

Пример:

Начнем с простейшего метода – аппроксимации Бенарда.

$MR_j = \frac{j-0.3}{n+0.4}$, где $n = 6$:

Время до отказа t_j, час	Порядковый номер отказа	Медианный ранг
16	1	0,10937
34	2	0,26563
53	3	0,42188
75	4	0,57813
93	5	0,73438
120	6	0,89063

Пример:

Оценка Филлибена:

$$MR_j = \begin{cases} 1 - 0.5^{1/n} & if\ j = 1; \\ \dfrac{j - 0.3175}{n + 0.365} & if\ j = 2,3,\dots,n-1; \\ 0.5^{1/n} & if\ j = n. \end{cases}$$

Время до отказа t_j, час	Порядковый номер отказа	Медианный ранг
16	1	0,1091
34	2	0,26434
53	3	0,42145
75	4	0,57855
93	5	0,73566
120	6	0,8909

Пример:

Далее, вычислим медианные ранги с использованием формулы с F-распределением:

$$MR_j = \frac{1}{1 + \frac{n-j+1}{j} \cdot F^{-1}(0.5, m, n)} \qquad \begin{array}{c} m = 2(n-j+1) \\ n = 2j \end{array}$$

№	m	n	qF(0.5,m,n)
1	12	2	1.36097
2	10	4	1.11257
3	8	6	1.02975
4	6	8	0.97111
5	4	10	0.89882
6	2	12	0.73477

Значения обратной (квантильной) функции вычислим с использованием функции qF из пакета Mathcad.

Пример:

Тогда медианные ранги равны:

$$MR_j = \frac{1}{1 + \frac{n-j+1}{j} \cdot F^{-1}(0.5, m, n)}$$

$$m = 2(n-j+1)$$
$$n = 2j$$

Время до отказа t_j, час	Порядковый номер отказа	Медианный ранг
16	1	0.1091
34	2	0.26445
53	3	0.42141
75	4	0.57859
93	5	0.73555
120	6	0.8909

Пример:

Значения, полученные с помощью F-распределения, совпадают с точными значениями, получаемыми путем решения уравнений с биномиальными коэффициентами.

Бенард	Филлибен	F-распределение
0.10937	0.1091	0.1091
0.26563	0.26434	0.26445
0.42188	0.42145	0.42141
0.57813	0.57855	0.57859
0.73438	0.73566	0.73555
0.89063	0.8909	0.8909

Использование аппроксимаций вносит небольшую ошибку, которой во многих случаях можно пренебречь.

Теоретические сведения

Пусть X – непрерывная случайная величина, имеющая функцию плотности распределения

$$f(x, \Theta) \equiv f(x, \theta_1, \theta_2, \dots, \theta_k)$$

где $\theta_1, \theta_2, \dots, \theta_k$ - это k неизвестных параметров, которые необходимо оценить на основе N независимых наблюдений $x_1, x_2, \dots, x_N$, ($N > k$), которые в случае теории надежности представляют собой времена отказов.

Функция правдоподобия:

$$\mathcal{L}(\widehat{\Theta}) = \prod_{i=1}^{N} f(x_i, \widehat{\Theta})$$

На практике удобнее работать с логарифмической функцией правдоподобия:

$$\Lambda(\widehat{\Theta}) = \ln \mathcal{L}(\widehat{\Theta}) = \sum_{i=1}^{N} \ln f(x_i, \widehat{\Theta})$$

Оценки максимального правдоподобия получаем, находя максимум значения $\mathcal{L}(\widehat{\Theta})$ или $\Lambda(\widehat{\Theta})$:

$$\frac{\partial \Lambda}{\partial \hat{\theta}_j} = 0, \qquad j = 1,2, \dots, k$$

Теоретические сведения

Оценки параметров, полученные разными способами для одной и той же выборки, будут различаться.

Поскольку на практике истинные значения параметров неизвестны, мы не можем с уверенностью предпочесть один метод другому.

Большинство исследователей в области надежности используют метод максимального правдоподобия.

Теоретические сведения

Оценки параметров, полученные методом МНК для одной выборки, но разных моделей, можно сравнить между собой, используя сумму квадратов отклонений (RSS) или среднеквадратичную ошибку (MSE):

$$RSS(\widehat{\Theta}) = \sum_{i=1}^{N}\left(Y_i - F(X_i, \widehat{\Theta})\right)^2$$

$$MSE(\widehat{\Theta}) = \frac{RSS(\widehat{\Theta})}{N}$$

Меньшее значение указывает на лучшую модель.

Теоретические сведения

Оценки параметров, полученные методом ММП для одной выборки, но разных моделей, можно сравнить между собой, используя удвоенное значение логарифмической функции правдоподобия, взятое со знаком «минус»: $-2\Lambda(\widehat{\Theta})$.

Меньшее значение указывает на лучшую модель.

ЗАДАНИЕ

Цель работы:

ознакомиться с основными методами получения точечных и интервальных оценок параметров моделей надежности.

Ход работы:

1. Для выборок случайных величин, полученных в ходе выполнения лабораторной работы №4, сгенерированных из:
   - экспоненциального распределения;
   - нормального распределения;
   - распределения F_3, заданного по варианту;

   найти оценки параметров распределений с использованием МНК, считая, что случайные величины распределены:

ЗАДАНИЕ

- в соответствии со своим истинным распределением;
- в соответствии с двухпараметрическим распределением Вейбулла.

3. Повторить п.1 с использованием ММП.
4. Для результатов, полученных в п.1 и п.2 :
   - определить максимальную ошибку;
   - определить среднеквадратическую ошибку;
   - определить значение -2Λ;
   - построить графики функций вероятности отказа в соответствии с полученными моделями надежности.
5. Заполнить Таблицу 1.

ЗАДАНИЕ