Министерство образования и науки Украины
Донбасская государственная машиностроительная академия

МЕТОДИЧЕСКИЕ УКАЗАНИЯ
к лабораторным работам по дисциплине

« Надежность и диагностика »

для студентов специальности
«Электромеханические системы автоматизации и электропривод»
дневной формы обучения

Рекомендовано до перезатвердження
кафедрою ЕСА
(назва кафедри)
Протокол № 34 від 11.06. 2012 р.
(протокол, номер, дата)

В.О Зав. кафедрою ЕСА
(назва кафедри)
_________О.М. Наливайко
(підпис, ініціали, прізвище)

Краматорск 2010

Содержание

ЛАБОРАТОРНАЯ РАБОТА № 1

СТАТИСТИЧЕСКИЙ АНАЛИЗ ПАРАМЕТРОВ И ПОКАЗАТЕЛЕЙ НАДЕЖНОСТИ

1. Цель работы

Ознакомление с методами статистического анализа надежности систем.

2. Теоретическая часть

Основной вероятностной характеристикой надежности является функция распределения параметров и показателей, характеризующих надежность. Исходными данными для ее оценки выступают эмпирические значения $x_1,\ldots,x_N$ параметров (показателей) надежности - выборка случайных величин.
Основой математико-статистического анализа надежности является преобразование вида

$$\{x_1,\ldots,x_N\}\xrightarrow{P}F\left(x,\vec{\theta}\right)$$

где P - алгоритм преобразования выборочных данных в функцию распределения;

F(x, $\vec{\theta}$)- параметрический закон распределения случайной величины;

$\vec{\theta}$ - вектор параметров закона распределения.

В силу того, что на практике может использоваться много различных законов распределения, задача анализа сводится к обоснованному преобразованию выборочных данных в конкретную функцию распределения F(x, $\vec{\theta}$) В общем случае алгоритм преобразования P имеет следующий вид (Рис.1):

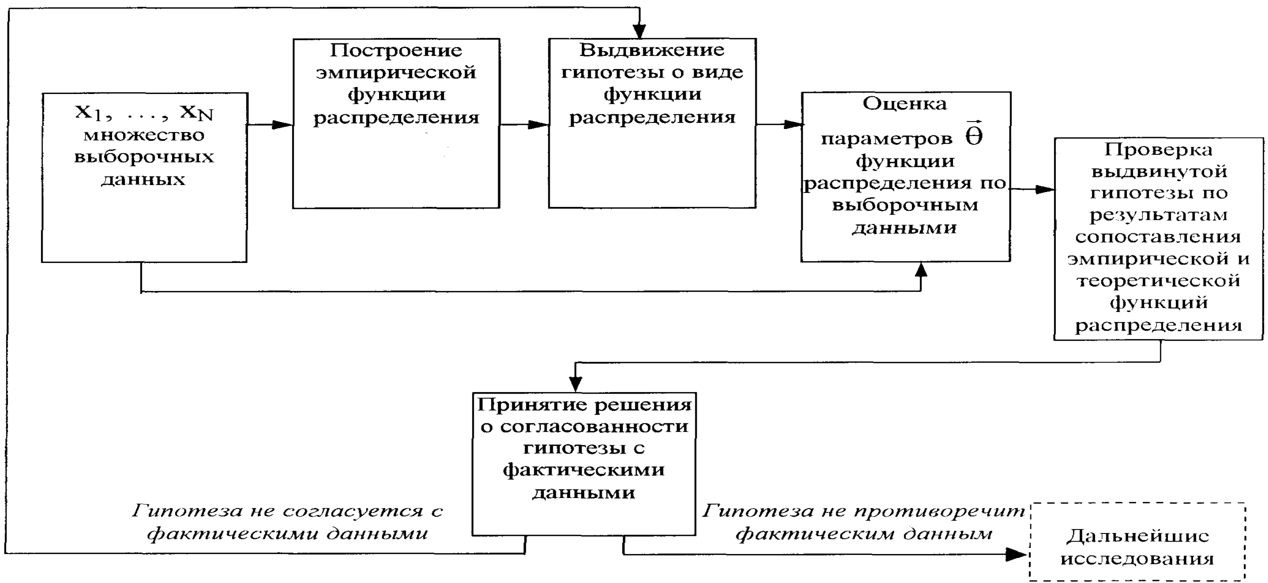

Рис.1. Общий вид алгоритма преобразования .

В зависимости от вида критерия, используемого для проверки выдвинутой гипотезы, построение эмпирической функции распределения осуществляется по группированным, либо не группированным данным.

Обобщенная схема построения эмпирической функции распределения по группированным данным имеет вид (Рис.2)

Рис.2 Обобщенная схема построения эмпирической функции

Здесь x_{min}, x_{max} — минимальные и максимальные элементы выборки.

Построение эмпирической функции распределения по не группированным данным сводится к преобразованию вида

$$\hat{F}(x_i) = i / N \qquad (i=1; N)$$

Выдвижение гипотезы о виде функции распределения не поддается какой-либо формализации. Вид закона (экспоненциальный, Релея, Вейбулла, Гауса и др.) выбирается по результатам визуального анализа эмпирической функции распределения.

Оценка параметров θ функции распределения осуществляется по известным (описанным в литературе) соотношениям. Так, например, параметры σ, ν_1 закона распределения Гаусса

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp \frac{(x-\nu_1)^2}{2\sigma^2}$$

определяются с помощью соотношений

$$\nu_1 = \frac{\sum_{i=1}^{N} x_i}{N}; \; \sigma = \sqrt{\frac{1}{N-1}\left(\frac{\sum_{i=1}^{N} x_i^2}{N} - \left(\frac{\sum_{i-1}^{N} x_i}{N}\right)^2\right)}$$

Проверка выдвинутой гипотезы о виде функции распределения производится с помощью так называемых «критериев согласия». Рассмотрим два наиболее часто используемых критерия - χ^2 и критерий Колмогорова.

Критерий χ^2. Пусть имеется ряд эмпирических значений показателей /параметров $X_1,...,X_N$, характеризующих надежность. Требуется проверить гипотезу о том, что функция распределения F(x,$\vec{\theta}$), вектор параметров $\vec{\theta}$ которой рассчитан по выборочным данным, согласуется со значениями $X_1,...,X_N$.

Разобьем интервал [$X_{\min}$, $X_{\max}$] на n интервалов (правила расчета будут описаны ниже). Рассчитаем теоретическую вероятность попадания случайной величины в I-й интервал

p_i=F(b_i,$\vec{\theta}$)-F(a_i,$\vec{\theta}$).

Подсчитаем число m_i элементов выборки, попавших в тот же интервал [a_i, b_i]. Если выдвинутая гипотеза о виде F(x,$\vec{\theta}$) согласуется с фактическими данными, то по К. Пирсону при больших N величина $\hat{\chi}^2=\sum_{i=1}^{N}\frac{(m_i-np_i)^2}{np_i}$ имеет распределение χ^2 с (N-1) степенями свободы. По заданному уровню значимости (обычно из набора 0,1; 0,05; 0,01; 0,001) по таблице 2 находится такое значение $\chi^2(N-1,\varepsilon)$, что P{$\hat{\chi}^2>\chi^2(N-1,\varepsilon)$)}=$\varepsilon$.

Если $\hat{\chi}^2$ больше $\chi^2(N-1,\varepsilon)$, то считается, что гипотеза не согласуется с фактическими данными. Если $\hat{\chi}^2\le\chi^2(N-1,\varepsilon)$, то считается, что гипотеза не противоречит фактическим данным.

Пример 1. Допустим, имеются 43 выборочных значения параметра:
0; 15; 24; 28; 37; 49; 54; 60; 75; 87; 92; 111; 114; 121; 127; 130; 134; 138;
140; 144; 147; 149; 155; 168; 170; 173; 189; 192; 197; 198; 201; 204; 225;
231;243;248;249;256;265;274;281;300;
Требуется проверить гипотезу о том, что выборочные данные описываются экспоненциальным законом распределения. Уровень значимости критерия принять равным ε = 0,01.

Решение. Разобьем интервал [0; 300] на n=6 равных частей: [0,50), [50,100), [100,150), [150,200), [200,250), [250,300).

Закон распределения имеет вид f(x, λ) =λ exp(-λx), параметр λ которого рассчитывается по формуле

$$\hat{\lambda}=\frac{N}{\sum_{i=1}^{N}x_i}=\frac{43}{6485}=0.00631$$

Рассчитываются значения m_i и p_j. Результаты сводятся в таблицу.

Таблица 1

Номер интервала	m_i	p_i	Np_j	$\frac{[m_i - Np_i]^2}{Np_i}$
1	5	0.27	11.34	3.51
2	6	0.2	8.40	0.62
3	11	0.15	6.30	3.86
4	8	0.11	4.62	2.16
5	7	0.07	2.94	5.41
6	5	0.20	8.40	1.43

Подсчитывается величина

$$\hat{\chi}^2 = \sum_{i=1}^{6} \frac{[m_i - Np_i]^2}{Np_i} = 17{,}04$$

По заданному уровню значимости ε=0,01 по таблице 2 находится $\chi^2(N-1,\varepsilon)=\chi^2$(5,0.01)=15.09. Сопоставление $\chi^2 \subset \chi^2$(5,0.01) позволяет заключить, что гипотеза о том, что выборочные данные соответствуют экспоненциальному закону распределения не согласуется с фактическими данными.

Критерий Колмогорова. По имеющимся выборочным значениям $X_1,...,X_N$ по не сгруппированным данным строится эмпирическая функция распределения $\hat{F}(x)$. Величина

$$\hat{D} = \sup_{0 \le x \le \infty} |\hat{F}(x) - F(x,\theta)|$$

имеет распределение, не зависящее от функции F(x,θ). При N $\le$ 50 по заданному уровню значимости ε с помощью таблицы 3 (ε выбирается из набора 0.1; 0.05; 0.02; 0.01) находится такое значение A(N,χ^2), что

P{D>A(N, ε)}= ε.

Если значение $\hat{D}$=A(N, ε), то считается, что эксплуатационные данные не противоречат гипотезе. При N>50 используется предельное распределение K(y) Колмогорова. Проверка соответствия выдвинутой гипотезы основана на использовании соотношения

$$\lim_{N\to\infty} P\{\sqrt{N}\,\hat{D}>y\}=l-K(y).$$

Пример 2. В условиях примера 1 проверить по критерию Колмогорова гипотезу о том, что распределения показателя опасности является равномерным в интервале [0;300]. Уровень значимости ε принять равным 0.01.

Из таблицы 3 следует, что

A(N, ε)=A(43;0.01)= 0.243

Величина $\hat{D}$ составляет

$$\hat{D} = \max|\hat{F}(x) - \frac{x}{300}| = 0.07 < 0.243$$

$$0 \le x \le 300$$

Таким образом, выдвинутая гипотеза о равномерном распределении показателей опасности на интервале [0;300] не противоречит фактическим данным.

3. Порядок выполнения работы

По выборочным данным проверить их соответствие следующим законам распределения по критериям $\hat{\chi}^2$ и Колмогорова:

- экспоненциальному f(x)= $\lambda e^{-\lambda x}$, 0<x<∞

- усеченному нормальному

$\bar{f}$(x)=cf(x), 0<x<∞ где f(x) - плотность неусеченного распределения; c - нормирующий множитель, определяемый по правилу $c = \dfrac{1}{\int_0^{\infty} f(x)dx}$

- Релея $f(x) = \dfrac{x}{\sigma_t^2}\exp\left(-\dfrac{x^2}{2\sigma_t^2}\right)$, 0<x<∞

(обратите внимание, что для упомянутых законов распределения приведена формальная запись для дифференциальных функций распределения. В случае необходимости следует рассчитать интегральную функцию F(y,θ) по правилу F(y,θ)=$\int_0^y f(x)dx$

Параметры вышеупомянутых законов распределения определяются по правилам:

• экспоненциальное $\lambda = \dfrac{N}{\sum_{i=1}^{N} x_i}$;

• усеченное нормальное - $M = \dfrac{\sum_{i=1}^{N} x_i}{N}; \sigma = \sqrt{\dfrac{1}{N-1}\left(\dfrac{\sum x_i^2}{N} - M^2\right)}; c = \dfrac{1}{\int_0^{\infty} f(x, M, \sigma)dx}$

• Релея $M = 1{,}253\sigma_t$, где $M = \dfrac{\sum_{i=1}^{N} x_i}{N}$.

При использовании критерия χ^2 проверку осуществлять для следующих правил разбиения интервала [x_{min}, x_{max}]:

- n=int(l+3,3lgN)
- n=int(5lgN)
- n = int($\sqrt{N}$)
- n = int($\sqrt[3]{N}$)

Выборочные данные, соответствующие различным вариантам, приведены в таблице 4.

4. Требования к отчету

Отчет по лабораторной работе должен содержать:

1) титульный лист
2) название и цель работы
3) При использовании критерия χ^2:

- таблицу значений $m_i, p_i, np_i, \dfrac{[m_i - np_i]^2}{np_i}$, соответствующих различным правилам расчета n;
- значения $\hat{\chi}^2$, соответствующие различным правилам расчета n;

При использовании критерия Колмогорова:

- эмпирическую функцию $\hat{F}$ (x);
- функцию F(y, θ), параметры которой рассчитаны по выборочным данным;
- значения $\hat{D}$, соответствующие различным F(y, θ).

4) Выводы о соответствии проверяемой гипотезы выборочным данным.

5. Контрольные вопросы

1. Назовите законы распределения случайных величин, используемых в теории надежности.
2. Поясните смысл параметров нормального распределения и влияние их значений на вид кривой распределения.
3. Алгоритм проверки соответствия выборочных данных выбранному закону распределения по критериям Колмогорова и χ^2.

Список литературы

1. Дружинин Г.В. Надёжность автоматизированных производственных систем - 4-е изд., перераб. и доп. - М.: Энергоатомиздат, 1986.
2. Вентцель Е.С., Овчаров Л.А., Теория случайных процессов и ее инженерные приложения. Уч. пособие для ВУЗов. – М.: Высшая школа, 2000.

Таблица 2.

Критические значения $\chi^2(r,q)$ распределения χ^2.

$$q=\frac{1}{2^{\frac{r}{2}}\Gamma\left(\frac{r}{2}\right)}\int\limits_{\chi^2(r,q)}^{\infty}x^{\frac{r}{2}-1}e^{-\frac{x}{2}}dx$$

r	q													
	0,99	0,98	0,95	0,90	0,80	0,70	0,50	0,30	0,20	0,10	0,05	0,02	0,01	0,001
1	0,000	0,001	0,004	0,016	0,064	0,148	0,455	1,074	1,642	2,71	3,84	5,41	6,63	10,83
2	0,020	0,040	0,103	0,211	0,446	0,713	1,386	2,41	3,22	4,60	5,99	7,82	9,21	13,82
3	0,115	0,185	0,352	0,584	1,005	1,424	2,37	3,66	4,64	6,25	7,81	9,84	11,34	16,27
4	0,297	0,429	0,711	1,064	1,649	2,20	3,36	4,88	5,99	7,78	9,49	11,67	13,28	18,46
5	0,554	0,752	1,145	1,610	2,34	3,00	4,35	6,06	7,29	9,24	11,07	13,39	15,09	20,5
6	0,872	1,134	1,635	2,20	3,07	3,83	5,35	7,23	8,55	10,64	12,59	15,03	16,81	22,5
7	1,239	1,564	2,17	2,83	3,82	4,67	6,35	8,38	9,80	12,02	14,07	16,62	18,48	24,3
8	1,646	2,03	2,73	3,49	4,59	5,53	7,34	9,52	11,03	13,36	15,51	18,17	20,1	26,1
9	2,09	2,53	3,33	4,17	5,38	6,39	8,34	10,66	12,24	14,68	16,92	19,68	21,7	27,9
10	2,56	3,06	3,94	4,86	6,18	7,27	9,34	11,78	13,44	15,99	18,31	21,2	23,2	29,6
11	3,05	3,61	4,58	5,58	6,99	8,15	10,34	12,90	14,63	17,28	19,68	22,6	24,7	31,3
12	3057	4,18	5,23	6,30	7,81	9,03	11,34	14,01	15,81	18,55	21,0	24,1	26,2	32,9
13	4,11	4,76	5,81	7,04	8,63	9,93	12,34	15,12	16,98	19,81	22,4	25,5	27,7	34,6
14	4,65	5,37	6,57	7,79	9,47	10,82	13,34	16,22	18,15	21,1	23,7	26,9	29,1	36,1
15	5,23	5,98	7,26	8,55	10,31	11,72	14,34	17,32	19,31	22,3	25,0	28,3	30,6	37,7
16	5,81	6,61	7,96	9,31	11,15	12,62	15,34	18,42	20,5	23,5	26,3	29,6	32,0	39,3
17	6,41	7,26	8,67	10,08	12,00	13,53	16,34	19,51	21,6	24,8	27,6	31,0	33,4	40,8
18	7,01	7,91	9,39	10,86	12,86	14,44	17,34	20,6	22,8	26,0	28,9	32,3	34,8	42,3
19	7,63	8,57	10,11	11,65	13,72	15,35	18,34	21,7	23,9	27,2	30,1	33,7	36,2	43,8
20	8,26	9,24	10,85	12,44	14,58	16,27	19,34	22,8	25,0	28,4	31,4	35,0	37,6	45,3
21	8,00	9,92	11,59	13,24	15,44	17,18	20,3	23,9	26,2	29,6	32,7	36,3	38,9	46,8
22	9,54	10,60	12,34	14,04	16,21	18,10	21,3	24,9	27,3	30,8	33,9	37,7	40,3	48,3
23	10,29	11,29	13,09	14,85	17,19	19,02	22,3	26,0	28,4	32,0	35,2	39,0	41,6	49,7
24	10,86	11,99	13,85	15,66	18,06	19,94	23,3	27,1	29,6	33,2	36,4	40,3	43,0	51,2
25	11,52	12,70	14,61	16,47	18,94	20,9	24,3	28,2	30,7	34,4	37,7	41,7	44,3	52,6
26	12,29	13,41	15,38	17,29	19,82	21,8	25,3	29,2	31,8	35,6	38,9	42,9	45,6	54,1
27	12,58	14,12	16,17	18,11	20,7	22,7	26,3	30,3	32,9	36,7	40,1	44,1	47,0	55,5
28	13,56	14,85	16,93	18,94	21,6	23,6	27,3	31,4	34,0	37,9	41,3	45,4	48,3	56,9
29	14,26	15,57	17,71	19,77	22,5	24,6	28,3	32,5	35,1	39,1	42,6	46,6	49,6	58,3
30	14,95	16,31	18,49	20,6	23,4	25,5	29,3	33,5	36,2	40,3	43,8	48,0	50,9	59,7

Таблица 3

Критические значения *A(n, q)* для наибольшего отклонения эмпирического распределения от теоретического (критерий Колмогорова).

$$P\{\sup|F_n(x)-F(x)|> A(n,q)\}=q$$

n	q			n	q		
	0,1	0.05	0.01		0.1	0.05	0.01
1	0,950	0,975	0,995	26	0.233	0,259	0,311
2	776	842	929	27	229	254	305
3	636	708	829	28	225	250	300
4	565	624	734	29	221	246	295
5	509	563	668	30	218	242	290
6	468	519	617	31	214	238	285
7	436	483	576	32	211	234	281
8	410	454	542	33	208	231	277
9	387	430	513	34	205	227	273
10	369	409	489	35	202	224	269
11	352	391	468	36	199	221	265
12	338	375	449	37	196	218	262
13	325	361	432	38	194	215	258
14	314	349	418	39	191	213	255
15	304	338	404	40	189	210	252
16	295	327	392	41	187	208	249
17	286	318	381	42	185	205	246
18	279	309	371	43	183	203	243
19	271	301	361	44	181	201	241
20	265	294	352	45	179	198	238
21	259	287	344	46	177	196	235
22	253	281	337	47	175	194	233
23	247	275	330	48	173	192	231
24	242	269	323	49	171	190	228
25	238	264	317	50	170	188	226

Таблица 4

ВЫБОРКИ СЛУЧАЙНЫХ ВЕЛИЧИН

значение выборки	№варианта																			
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
1	90	74	87	73	39	90	5	71	64	71	50	84	7	59	6	30	11	47	97	35
2	21	15	6	80	70	9	45	31	46	90	31	6	28	74	82	47	14	9	9	75
3	36	59	80	65	2	96	35	84	56	22	7	7	87	76	40	10	45	79	33	78
4	28	28	70	54	14	88	88	21	81	13	1	76	76	94	4	22	58	82	75	76
5	80	88	52	87	85	13	72	44	71	47	7	27	53	69	93	89	29	94	22	44
6	19	68	43	20	24	77	5	14	78	9	100	68	52	88	16	77	80	49	49	42
7	86	100	21	34	3	67	16	39	50	91	71	55	6	78	84	44	11	25	37	7
8	22	2	45	33	56	35	46	19	74	30	45	19	82	57	95	13	18	7	41	36
9	47	73	95	76	40	30	26	90	85	13	12	65	17	98	67	82	88	47	15	23
10	48	55	40	60	2	74	65	65	47	49	32	15	37	16	70	9	75	85	44	77
11	13	43	10	25	83	53	76	93	39	80	11	95	77	61	65	95	95	54	28	55
12	76	51	94	26	92	24	29	41	16	47	84	53	34	91	67	62	73	45	28	40
13	9	2	30	85	4	48	43	62	83	48	47	46	67	28	8	92	88	60	45	88
14	32	28	65	20	2	11	10	94	62	51	3	50	91	55	1	52	66	8	92	56
15	97	77	9	85	81	50	9	20	95	49	99	56	3	42	30	93	81	92	41	53
16	9	61	86	63	61	90	43	41	95	29	61	70	72	32	24	52	48	99	18	25
17	22	36	21	53	92	47	69	89	41	39	52	44	48	62	91	73	72	91	32	67
18	61	10	1	57	37	65	21	89	64	7	39	5	9	10	64	68	73	33	58	85
19	54	73	92	75	70	37	32	72	28	76	81	51	32	2	79	51	5	48	6	48
20	40	58	15	13	95	9	15	79	77	33	39	64	95	67	9	75	6	12	14	100
21	56	89	53	1	22	15	81	19	52	64	83	67	84	38	78	18	89	43	63	62
22	2	15	58	20	14	47	55	48	66	71	13	10	7	42	84	35	34	91	29	50
23	6	84	17	16	6	31	48	59	17	77	85	31	29	60	20	52	78	14	94	84
24	9	22	94	37	59	46	47	3	93	72	46	10	56	73	81	63	48	68	99	95
25	24	7	25	100	19	71	50	93	72	90	16	86	39	6	97	88	87	30	55	51
26	95	93	85	33	26	33	13	76	69	69	63	7	57	38	41	96	57	19	27	94
27	83	60	75	32	17	74	77	17	70	87	34	99	55	96	37	45	98	18	56	18
28	55	75	37	45	20	99	63	16	64	26	72	93	41	60	19	83	60	69	30	10
29	16	47	75	99	100	66	76	71	80	85	78	91	90	77	59	61	56	42	98	94
30	81	69	20	38	44	34	20	26	29	32	94	17	25	72	46	70	24	59	61	34
31	31	83	29	80	6	97	9	94	88	72	87	87	30	57	47	14	44	63	40	40
32	37	44	43	70	59	96	95	24	50	28	75	89	88	95	40	88	75	72	80	54
33	66	16	8	17	91	18	62	64	0	53	28	97	64	66	78	23	91	69	29	88
34	77	51	51	10	56	64	76	11	98	89	91	31	7	98	44	36	26	()	34	35
35	50	66	12	36	21	87	31	78	90	17	8	77	4	40	22	88	38	19	49	59
36	24	19	21	86	30	5	46	12	100	1	29	34	29	59	65	82	2	72	98	7
37	23	86	65	80	86	76	69	60	86	8	44	65	48	86	12	16	19	47	97	10
38	8	82	35	16	64	24	92	69	7	89	9	92	19	66	82	64	27	92	99	80
39	68	84	27	41	11	64	62	15	25	44	78	38	78	49	71	60	54	69	94	69
40	15	8	74	95	92	51	6	18	5	96	59	13	38	49	29	4	78	23	49	27

ЛАБОРАТОРНАЯ РАБОТА № 2

МАТЕМАТИКО-СТАТИСТИЧЕСКИЕ МЕТОДЫ ОБРАБОТКИ МАЛЫХ ВЫБОРОК

1.Цель работы

Изучение новых статистических методов анализа малых выборок.

2.Теоретическая часть

2.1. Анализ специальных статистических методов обработки малого числа наблюдений

Невозможность применения традиционных методов математической статистики для обработки выборок малого объема стимулировала разработку новых статистических методов, ориентированных специально в обработку малого числа наблюдений.

При анализе специальных статистических методов ограничимся лишь кратким анализом свойств.

Идея о необходимости нового подхода к обработке малых выборок впервые высказана В.В. Чавчанидзе и В.А. Кусишвили где для построения оценки функции распределения предлагалось использовать так называемый метод прямоугольных вкладов (МПВ). Исследование свойств этого метода и полученные при этом результаты стимулировали разработку целой серии методов, основанных на идее использования функций вкладов.

Оценки распределений, получаемые в результате использования этих методов, обобщенно могут быть представлены линейной суммой двух компонент: априорной и эмпирической. При этом эмпирическая компонента $\hat{f}(x)$ строится по данным выборки и представляет собой сумму функций, удовлетворяющих ряду условий:

$$\hat{f}(x) = \alpha_0 f_0(x) + \frac{1-\alpha_0}{N}\sum_{i=1}^{N} p(x-x_i), \qquad (2.1)$$

где $f_0(x)$ - априорная компонента;

$p(x-x_i)$- составляющая эмпирической компоненты, связанная с i *–й* реализацией выборки;

α_0 - вес априорной компоненты.

Анализ «нетрадиционных» методов начнем с МПВ. Названный метод называется на использовании априорной информации о неизвестном распределении и учете случайного характера выбора. Априорная информация о распределении состоит в следующем:

- при известных границах [a, b] интервала, на котором определена случайная величины X, плотность распределения удовлетворяет условию:

$$f(x) \geq 0 \qquad x \in [a, b]$$
$$f(x) = 0 \qquad x \notin [a, b]$$

- плотность распределение непрерывна внутри интервала [a, b] и не имеет очень крутых подъемов и спадов.

Из анализа априорной информации можно заключить, что в качестве априорной компоненты в МПВ наиболее целесообразно использовать равномерное распределение, заданное внутри интервала [a, b].

Учет случайного характера выборки выражается в том, что учитывается возможность появления любых других значений случайной величины из области $[x_i - \frac{d}{2}; x_i + \frac{d}{2}]$, где d - ширина вклада.

Методика построения *f(x)*,основывается на использовании в качестве функции вклада единичного прямоугольника, что и дало название методу. Плотность распределения при этом записывается в виде:

$$f(x) = \frac{1}{N+1}\{f_0(x) + \sum_{i=1}^{N} \Psi_i(x)\}, \qquad (2.2)$$

$\Psi_i(x)$ - функция вклада единичной площади. При этом для некоторых значений x_i функция вклада может выходить за пределы интервала [a, b]. В этом случае часть площади, выходящая за границы интервала, отбрасывается, а над оставшимся основанием прямоугольника равномерно надстраивается площадь, равная отброшенной.

В качестве функции вклада могут использоваться также распре деления иной формы. Так, например, в качестве функции вклада используется распределение Симпсона, потенциальная функция, дельтообразная функция и другие. В известных работах показано, что:

- во-первых, для каждого типа распределения существует оптимальная ширина вклада d_0, при котором эффективность метода максимальна; величина d_0 уменьшается с увеличением объема выборки;
- во-вторых, форма вклада оказывает существенное влияние на точность метода и простоту реализации;
- в-третьих, оптимальная ширина вклада зависит не только от типа, но и от значений параметров распределения; при отсутствии такой информации задача выбора параметров вклада не приводит к определенному решению, поэтому на высокую эффективность методов вкладов при обработке экспериментальных данных в случае, когда отсутствует априорная информация о виде распределения, рассчитывать трудно.

Своеобразный подход к определению формы и параметров вкладов рассматривается в работе В.И. Шаповалова. Функции вкладов предлагается подбирать таким образом, чтобы из имеющейся выборки извлекалось наибольшее количество информации о функции распределения.

Однако, рассматриваемый подход имеет также существенные недостатки. Во-первых, для определения формы вклада используются значения четвертого центрального момента; во-вторых, для определения оптимальных значений параметров требуется априорное значение типа распределения. Если тип распределения заранее неизвестен, обосновать выбор значений параметров не представляется возможным.

Заключая анализ методов вкладов, отметим, что идея учитывать случайный характер каждой отдельной реализации используется ухе при группировании больших выборок. При построении гистограммы каждой реализацией связывается элементарная плотность равномерного распределения на подинтервале, включающем эту реализацию. Суммирование всех элементарных плотностей дает оценку плотности распределения, графическим изображением которой и является гистограмма.

Таким образом, из анализа сущности методов вкладов можно заключить, что принципиально новым элементов, обусловливающим высоки эффективность методов при обработке выборок ограниченного объема, является использование априорной информации, заключенной в границах интервала, на котором определена случайная величина.

И.В. Еременко и А.Н. Свердлик предложили эмпирический метод построения функции распределения, названный метод уменьшения неопределенности (МУН). Отличие этого метода от МПВ заключается в том, что вместо прямоугольного вклада ширины d **,** построенного около реализации x_i**,** используется нормированное равномерное распределение, заданное на интервале $[x_{i-1}, x_{i+1}]$ суть МУН заключается в равномерном распределении скачка вероятности в точке x_i.

Выражение для эмпирической функции распределения, получаемой *с* помощью МУН, записывается в виде

$$F(x) = \frac{x - x_{i-1}}{x_i - x_{i-1}}[F(x_i) - F(x_{i-1})] + F(x_{i-1}), \qquad (2.3)$$

при $x_{i-1} \le x < x_i$ и

$$F(x_i) = \frac{1}{N+1}[\frac{x_i - a}{b - a} + (i-1) + k_i - 1], \qquad (2.4)$$

где k_i - число одинаковых значений x_i **.** МУН является частный случаем МПВ, в котором ширина вклада является случайной величиной, изменяющейся с изменением индекса i.

Метод априорно-эмпирических функций (АЭФ**),** разработанный И.П. Демаковым и В.Е. Потепуном. В случае использования метода АЭФ интегральная функция распределения представляется в виде:

$$F(x) = \omega F_a(x) + (1-\omega) F_э(x) \qquad (2.5)$$

где $F_a(x)$ - априорное распределение, построенное по априорным данным;

$F_э(x)$ - эмпирическое распределение, построенное по данным выборки;

ω - коэффициент достоверности информации об априорном распределении.

Метод АЭФ, так же как и рассмотренные выше методы получения оценок распределений, основан на использовании априорной информации в виде границ интервала $[a, b]$ и на индивидуальном подходе к каждой отдельной реализации случайной величины. Однако при этом априорной информации приписывается некоторый вес ω и полагается, что

$f(x) \geq 0$ при $a - \frac{\Delta}{2} \leq x \leq b + \frac{\Delta}{2}$,

$f(x) = 0$ при $x < a - \frac{\Delta}{2}$, $x > b + \frac{\Delta}{2}$,

где Δ - интервал дискретности, определяемый точностью наблюдения (измерения) случайной величины. В известных работах приводятся сведения о том, что по эффективности метод АЭФ не уступает МПВ, а по простоте реализации сродни МУН, Кроме того, сильной стороной метода является то, что предлагается отказаться от уравнивания значимости априорной информации и информации, получаемой от наблюдений.

Однако указанные преимущества на практике можно реализовать лишь тогда, когда имеется достаточно точная информация о предполагаемой функции распределения. В случае отсутствия такой информации метод АЭФ совпадает с традиционными методами математической статистики.

Метод сжатия области существования интегральных законов распределения (ИЗР), предложен И.В. Еременко. При использовании метода предполагается, что:

- имеется выборка конечного объема, представленная в виде вариационного ряда x_i $(i=1,2, ... ,N)$;
- для каждого элемента выборки x_i существует единственная последовательность чисел $y_i = F(x_i)$ такая, что

$$y_1 \leq y_2 \leq \ldots \leq y_i \leq \ldots \leq y_N; \qquad (2.6)$$

- построение оценки распределения состоит в отыскании приближенных значений $F(x_i)=a_i$:

$$0 \leq a_1 \leq a_2 \leq \ldots \leq a_i \leq \ldots \leq a_N \leq 1, \qquad (2.7)$$

задача оценивания, таким образом, сводится к выбору такой последовательности $a_i(i=1,2,...,N)$, чтобы минимизировать либо математическое ожидание, либо дисперсию погрешности построения ИЗР для каждого i -го члена вариационного ряда $x_i(i=1,2,...,N)$.

Наиболее предпочтительным при инженерных расчетах является алгоритм, основанный на минимизации дисперсии. В этом случае члены ряда (2.7) определяются следующим образом:

$$a_i = \frac{i}{N+1}, \qquad (2.8)$$

Соединяя полученные значения $a_i(i=1,2,...,N)$ отрезками прямых, получаем кусочно-линейную аппроксимацию интегральной функции распределения с узлами в точках $\{x_i, a_i\}$.

Основным достоинством метода сжатия ИЗР является возможность вычислить доверительную вероятность для каждого i-го члена последовательности (2.7). При этом вероятность прохождения ИЗР через заранее выбранный интервал $[a_i-\Delta_I, a_i+\Delta_I]$ определяется как

$$\mathrm{P}_{\Delta\alpha_i,N} = \int_{a_i-\Delta_i}^{a_i+\Delta_i} \mathrm{P}_{y,i,N}\, dy_i \qquad (2.9)$$

где $\mathrm{P}_{y,i,N}$ - плотность вероятности прохождения ИЗР на уровне y_i для i-*го* испытания в серии из N испытаний. Эта величина описывается выражением

$$\mathrm{P}_{y,i,N} = \frac{N!}{(i-1)!(N-i)!} y_i^{i-1}(1-y_i)^{N-i} \qquad (2.10)$$

Кроме того, использование метода сжатия ИЗР обеспечивает высокую точность оценивания функции распределение в окрестности узлов интерполяции.

Недостатком метода является то, что точность воспроизведения остальных участков кривой $y= F(x)$ путем линейной интерполяции при малом числе наблюдений невысока.

Из анализа (2.8) можно заключить, что метод сжатия ИЗР является частным случаем МПВ (в этом можно легко убедиться, продифференцировав интегральную функцию распределения, получаемую с помощью ИЗР). В силу того, что используемые вклады не являются оптимальными, можно заключить, что по точности метод сжатия ИЗР уступает МПВ.

Г.В. Дружининым и О.В. Вороновой разработан эмпирический метод построения интегральной функции распределения, названный авторами «методом последовательных медиан» (МПМ). Суть этого метода заключается в следующем.

Исходные данные $\{x_1, x_2,..., x_N\}$ располагаются в вариационный ряд, находится медиана этого ряда и на графике эмпирической функции распределения ставится точка с координатами $\overset{*}{x}_{\mathrm{ME}}$ и $F(\overset{*}{x}_{\mathrm{ME}})$=0.5. Затем находятся медианы двух половин вариационного ряда и им в соответствие ставятся значения эмпирической функции 0.25 и 0.75 и т.д. Указанная процедура продолжается до тех пор, пока не будут рассмотрены все имеющиеся значения $x_i(i=1,2,...,N)$. В результате получается ряд точек $F_n(x)$.

Нахождение начального и конечного значений функций распределения осуществляется по формулам

$$\overset{*}{x}_н = x_1 - (x_2 - x_1)/k_1 \qquad (2.11)$$

$$\overset{*}{x}_k = x_N + (x_N - x_{N-1})/k_2, \qquad (2.12)$$

где $k_1 = \dfrac{\overset{*}{x}_м - x_1}{x_N - \overset{*}{x}_м}, k_2 = \dfrac{1}{k_1}$.

Здесь x_1, x_2 первые два члена вариационного ряда, $\overset{*}{x}_м$ - оценка моды функции распределения. При малом числе опытных данных считается, что оценка моды совпадает с оценками медианы по выборке.

Из описания МПМ можно сделать вывод, что он также является одной из разновидностей метода вкладов (чтобы убедиться в этом, достаточно продифференцировать получаемую эмпирическую функцию распределения). Весьма интересным моментом в МПМ является то, что границы интервала, на котором определена случайная величина, определяются по данным выборки.

Своеобразный подход к построению оценки распределения предложен в работах Л.Я. Пешеса, М.Д. Степановой, Н.Н. Власовца.

Предлагаемый ими метод основан на выдвижении и проверке гипотез, при чем в качестве критерия согласия предлагается использовать условие совпадения трех-четырех первых моментов распределения. Определение вида аппроксимирующего распределения осуществляется путем оценки попадания расчетных моментов этого распределения в доверительные интервалы для эмпирических моментов. При этом доверительные интервалы для эмпирических моментов определяются путем статистического моделирования эмпирической функции на ЭВМ.

Для построения эмпирической функции распределения по выборке $x_i(i=1,2,...,N)$ авторами предлагается использовать метод прямоугольных вкладов, причем в качестве функции вклада предлагается использовать прямоугольник с переменным основанием $d_i = 0.5(x_{i+1} - x_i)$. Функции вклада строятся относительно $x_i' = 0.5(x_{i+1} + x_i)$ При этом интегральная функция распределения записывается в виде:

$$F(x) = \frac{x - x_{i-1}^{/}}{x_i^{/} - x_{i-1}^{/}}(y_i - y_{i-1}) + y_{i-1}, \; x_{i-1}^{/} < x < x_i^{/}, \qquad (2.13)$$

где y_i – частность i - *ой* реализации случайной величины X.

Алгоритм формирования оценок первых начальных четырех моментов заключается в следующем:

– с помощью программного датчика равномерно - распределенных случайных чисел вырабатываются случайные последовательности равномерно распределенные на интервале (0; 1);
– с помощью обратного преобразования $x=F^{-1}(y)$ получаются случайные последовательности объема N, удовлетворяющие распределению (2.13). Учитывая выражение для $F(x)$, значения x определяются по правилу:

$$x = \frac{y - y_{i-1}}{y_i - y_{i-1}} * \frac{x_{i+1} - x_{i-1}}{2} + \frac{x_{i-1} - x_i}{2}$$, при $y_{i-1} < y < y_i$.

По полученным N реализациям случайных величин X оцениваются значения первых четырех начальных моментов распределения. Указанные операции повторяются k раз, причем k оцениваются при помощи неравенства $k \geq 20(1-\gamma)^{-1}$, где γ - уровень доверия. Для каждого из моментов составлялся вариационный ряд

$$\nu_i^{(1)} \leq \nu_i^{(2)} \leq \ldots \leq \nu_i^{(k)} \quad (i = 1,2,3,4)$$

Границы доверительного интервала определяются величинами $\nu_i^{(r_1)}$ и $\nu_i^{(r_2)}$, где r^1, r^2 целые части чисел $\frac{k(1-\gamma)}{2}$ и $\frac{k(1+\gamma)}{2}$. Методика проверки выдвигаемых гипотез заключается в следующем. Для каждого из рассматриваемых теоретических распределений методом максимального правдоподобия определяются значение параметров, после чего вычисляются первые четыре начальных момента, соответствующие этому распределению.

После этого выясняют, попадают ли эти моменты в доверительные интервалы для эмпирических моментов. При этом переходят последовательно от более высоких уровней к более низким. В качестве подходящего распределения выбирают такое, моменты которого попадают в самый узкий интервал.

Следует отметить, что авторами метода выбран весьма удачный способ идентификации эмпирического и теоретического распределения по значениям начальных моментов, т.к. совокупность моментов образует минимальную систему достаточных статистик, однозначно определяющих распределение.

К недостаткам метода следует отнести:

- во-первых, то, что при малом числе данных значения моментов высших порядков (начиная с третьего) определяются с большой погрешностью;
- во-вторых, вследствие того, что рассматриваемый метод является одной из разновидностей метода вкладов, а ширина вклада является случайной, оценка распределения (2.13) не является оптимальной;
- в-третьих, «разрешающая способность» метода зависит от анализируемых теоретических распределений. Если в их числе не найдется подходящей модели, то перейти к узким доверительным интервалам не удается, что является следствием неконструктивного, проверочного характера метода.

2.2. Информационный подход к построению оценок распределения по ограниченному числу опытных данных

Количество информации (по Шеннону) о функции распределения, содержащееся в выборке малого объема, ограничено, поэтому оценить распределение по опытным данным можно лишь с определенной степенью точности. Целью разработки новых статистических методов является, возможно, более полное использование выборочной информации о функции распределения и, следовательно, получение оценок распределений, как можно более близких к истинным.

На основе проведенного анализа нетрадиционных методов математической статистики, можно заключить, что в случае отсутствия априорных сведений о функции распределения целесообразно строить метод оценивания таким образом, чтобы исключить этапы, требующие использования какой-либо информации, кроме опытной. Использование энтропийного подхода позволяет получать оценку распределения на основе лишь опытной информации. Задача в этом случае формулируется следующим образом.

Пусть случайная величины X (в общем случае векторная) может принимать ряд значений $x_1, x_2,..., x_N$ с вероятностями $p_1, p_2,..., p_N$, которые неизвестны. В результате эксперимента получены средние значения функций $g_j(x)(j=1,2,...,m)$

$$\sum_{i=1}^{n} p_i g_j(x_i) = \nu_j \,, \qquad (2.14)$$

причем $m<<n$.

Требуется на основе имеющейся информации определить значения p_i .

Дополнив исходные данные условием нормировки $\sum_{i=1}^{n} p_i = 1$, получим $m+1$ уравнение с n неизвестными.

Из постановки задачи видно, что однозначно определить значения p_i по имеющейся опытной информации не представляется возможным. Поэтому необходим критерий, который из бесконечного множества распределений позволял бы выбирать такое, которое наиболее точно согласуется с имеющимися экспериментальными данными. В указанных работах в качестве такого критерия используется энтропия распределения. Под энтропией понимается неопределенность относительно истинных значений p_i $(i=1,2,...,N)$. Исходная задача при этом преобразуется к виду: определить значения p_i , доставляющие максимум функционалу

$$S = -\sum_{i=1}^{n} p_i \ln p_i \rightarrow \max \qquad (2.15)$$

при уравнениях связи

$$\left.\begin{aligned} &\sum_{i=1}^{n} p_i = 1 \\ &\sum_{i=1}^{n} p_i g_j(x_i) = \nu_j \end{aligned}\right\} \quad j=1,2,...,m \qquad (2.16)$$

Подчеркнем, что по условию задачи тип распределения p_i полагается неизвестным.

При таком подходе из всех распределений, согласующихся с исходными данными (и представленными в виде уравнений связи (1.16)), следует выбирать наиболее пологое (т.е. наиболее близкое к равномерному). Другими словами следует избегать распределений, имеющих острые пики, при которых выделяется тот или иной результат, за исключением случаев, когда того требует условие задачи.

Таким образом, оценке распределения, полученной на основе энтропийного подхода (апостериорному распределению) соответствует наибольшая неопределенность (согласно (2.15)). Если число возможных значений случайной величины X априорно известно, то количество информации о функции распределения I_{on} , извлекаемое в результате обработки исходных статистических данных, определяется соотношением:

$$I_{on} = S_{anp} - S$$

где через S_{anp} обозначена исходная неопределенность, соответствующая случаю, когда известно лишь число возможных значений случайной величины и определяемая соотношением

$$S_{anp} = -ln(n) \qquad (2.17)$$

Из анализа (2.17) следует, что оценка, доставляющая максимум неопределенному апостериорному распределению, гарантирует нас от использования любой информации, не связанной с данными выборки.

Использование информационного подхода позволяет представить практически все встречающиеся на практике одномерные определения в единой форме.

$$f(x) = \exp(\sum_{j=1}^{m} \mu_j x^j) \qquad (2.18)$$

где μ_j- значения параметров распределения, определяемые по значениям ν_j . Следует отметить, что выбор характеристик распределения в значительной степени произволен. Например, в качестве центральной точки распределения могут выступать среднее, медиана и мода, в качестве характеристик рассеяния - дисперсия, первый абсолютный момент и широта распределения.

В известной литературе показано, что в качестве ν_j целесообразно использовать значения начальных моментов распределения. Во-первых, в силу того, что начальные моменты являются осредненными величинами, то при малом числе данных, в их значениях обнаруживаются более устойчивые закономерности, чем в самих результатах наблюдений. Во-вторых, совокупность начальных моментов образует минимальную систему достаточных статистик, однозначно характеризует функцию распределения и стохастическую зависимость между переменными (в случае системы случайных величин). Поэтому в качестве уравнений связи следует принять значения начальных моментов распределений.

Из теории математической статистики известно, что для описания практически любого реально встречающегося распределения достаточно учитывать три - четыре начальных момента.

В известной литературе показано, что точность получаемых с помощью разработанного метода оценок распределений в целом оказывается выше, нежели у оценок, получаемых с помощью статистических методов аналогичного назначения. Таким образом, для оценивания показателей надежности по выборке малого объема целесообразно использовать информационный метод.

2.3. Методика построения оценок законов распределения случайных величин.

2.3.1. Методика построения оценок неусеченных законов распределения.

В методике использованы следующие обозначения: $\mu_j^{\text{н}}$ - нормированные параметры распределения, соответствующие нормированным значениям математического ожидания $\nu_1^{\text{н}}$ и среднеквадратического отклонения $\sigma^{\text{н}}$, μ_j^* - ненормированные параметры распределения, соответствующие $\overset{*}{X} \in [0,\ b^*]$, μ_j - параметры распределения, соответствующие $X \in [a,\ b]$.

Шаг 1. Исходный интервал [a, b], на котором возможны реализации случайной величины X, преобразуется таким обрезом, чтобы нижняя граница интервала совпала с началом координат $\overset{*}{X} = X - a$ (при этом верхняя граница интервала примет значение $\overset{*}{b} = b - a$). Соответственно преобразуются данные исходной выборки

$$\overset{*}{X}_i = X_i - a \qquad (2.3.1)$$

Шаг 2. Оцениваются значения двух первых начальных моментов распределения

$$\nu_j = \frac{\sum_{i=1}^{N} (\overset{*}{x}_i)^j}{N},\ j = 1,2 \qquad (2.3.2)$$

и с помощью выражения $\sigma = \sqrt{\frac{N}{N-1}(\nu_2 - \nu_1^2)}$ оценивается величина среднеквадратического отклонения распределения.

Шаг 3. Определяется значение первого нормированного начального момента

$$\nu_1^{\text{н}} = \frac{2\nu_1}{\sigma}$$

Шаг 4. Если соблюдается условие $\nu_1^{\text{н}} \geq 7$. то нормированные параметры распределения определяются с помощью соотношений

$$\mu_0^{\text{н}} = -\ln(2\sqrt{2\pi}) - [\nu_1]^2/8$$
$$\mu_1^{\text{н}} = \nu_1^{\text{н}}/4$$
$$\mu_2^{\text{н}} = -1/8 \qquad (2.3.3)$$

и осуществляется переход к шагу 6. Если указанное условие не соблюдается, то переходят к шагу 5.

Шаг 5. Вычисляются нормированные параметры распределения с помощью выражения $$\mu_j^{\text{Н}} = \sum_{i=0}^{5} a_i^{(j)}[\nu_1^{\text{Н}}]^i,\ j = 0,1,2 \qquad (2.3.4)$$

значения коэффициентов полинома $a_i^{(j)}$ приводятся в таблице 2.1.

Таблица 2.1.

j	0	1	2
a_0	*0.78760454*	*-0.23350401*10^{1}*	*0.19935077*
a_1	*-0.25889486*	*0.88266373*	*-0.10402411*
a_2	*-0.31872153*	*0.99215448*10^{-1}*	*-0.7659968*10^{-2}*
a_3	*0.45243271*10^{-1}*	*-0.52450713*10^{-1}*	*0.64909123*10^{-2}*
a_4	*-0.27552103*10^{-2}*	*0.62824786*10^{-2}*	*-0.83410484*10^{-3}*
a_5	*0*	*-0.240757*10^{-3}*	*0.33025782*10^{-4}*

Шаг 6. Определяются ненормированные значения параметров распределения

$$\mu_0^* = \mu_0^{\text{Н}} - \ln(\sigma) + \ln(2)$$

$$\mu_1^* = \frac{2\mu_0^{\text{Н}}}{\sigma}$$

$$\mu_2^* = \frac{4\mu_2^{\text{Н}}}{\sigma^2} \qquad (2.3.5)$$

Шаг 7. Вычисляются значения параметров распределения, соответствующие исходному интервалу [a, b]:

$$\mu_0 = \mu_0^* - \mu_1^* a + \mu_2^* a^2$$

$$\mu_1 = \mu_1^* - 2\mu_2^* a$$

$$\mu_2 = \mu_2^* \qquad (2.3.6)$$

Рассмотрим пример использования методики для оценки параметров неусеченного распределения.

Пример 1. Пусть имеется N=12 реализации случайной величины X:

x_1 =1044.29853	X_5 =836.57545	x_9 =1297.06719
x_2 =1982.45718	X_6 =4750.78613	x_{10} =1658.10047
x_3 =1719.11262	X_7 =1472.99209	x_{11} =2157.29901
x_4 =2817.04650	X_8 =1235.21748	x_{12} =1860.30754

Причем априорно известно, что нижняя граница интервала равна нулю.

Требуется:

а) найти аналитическое выражение плотности распределения случайной величины *f(x)*;

б) построить оценку функции распределения *F(x)*.

Решение. Определим значения двух первых начальных моментов распределения ν_1, ν_2 и величину среднеквадратического отклонения σ:

ν_1 =1900,18835; ν_2 =4698429,18750; σ=1042,93504

Определим значение первого нормированного начального момента $\nu_1^{\text{н}}$ =3,64546.

С помощью выражений (2.3.4) определим значения нормированных параметров распределения:

$\mu_0^{\text{н}} = -2.68633$; $\mu_1^{\text{н}} = 0.61469$; $\mu_2^{\text{н}} = -0.09324$

по которым с помощью (2.3.5), (2.3.6) определим значения ненормированных параметров распределения:

$\mu_0 = -8.916229724$; $\mu_1 = 0.001151079$; $\mu_2 = -0.000000338$.

Оценка плотности распределения запишется в виде:

$$F(x)=\exp(-8.916229724 + 0.001151079x - 0.000000338x^2)$$

На рис. 2.1 сплошной линией показана оценка функции распределения, получаемая с помощью информационного метода. Ломаной линией показана эмпирическая функция распределения, построенная классическим способом.

Исходные данные для примера получены при помощи датчика случайных чисел, подчиняющегося распределению Релея с параметром σ = *1020*. Заданная функция распределения показана на рисунке пунктирной линией.

Как видно из рисунка, наибольшее абсолютное уклонение оценки F*(x)* от заданной $F_3(x)$ составляет 0,14 при x = 1500. Наибольшее абсолютное уклонение эмпирической функции распределения от заданной составляет 0,35 при x = 2157,29901.

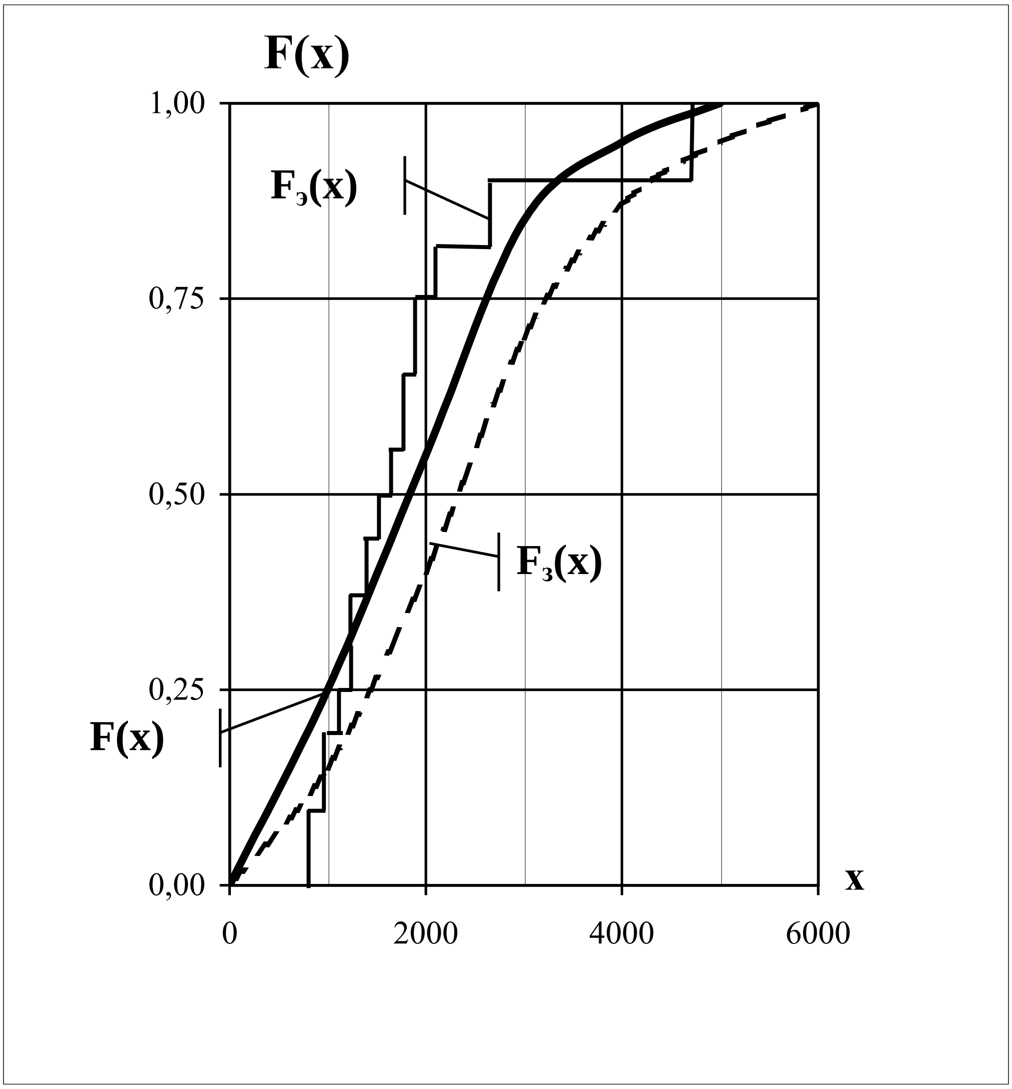

Рис. 2.1

2.3.2. Методика построения оценок усеченных законов распределения

Алгоритм оценки параметров усеченного распределения заключается в следующем.

Первый и второй шаг алгоритма совпадают с первыми двумя шагами алгоритма определения параметров неусеченного распределения.

Шаг 3. Если соблюдаются условия

$$\left.\begin{aligned} &\nu_1 < 0{,}5(b+a) \\ &\nu_2 \geq 3{,}5\sigma + a \end{aligned}\right\},$$

либо

$$\left.\begin{aligned} &\nu_1 \geq 0{,}5(b+a) \\ &(b-\nu_1) \geq 3{,}5\sigma \end{aligned}\right\},$$

то параметры распределения определяются с помощью соотношений

$$\mu_0 = -\ln(\sigma\sqrt{2\pi}) - \nu_1^{\,2} / 2\sigma^2$$

$$\mu_1 = \nu_1 / 2\sigma^2$$

$$\mu_1 = -1/2\sigma^2 \qquad (2.3.7)$$

Если же указанное условие не соблюдается, то значения ν_1 и σ пересчитываются в относительные величины по правилу

$$\nu_1^{\mathrm{H}} = \frac{14\nu_1}{b-a}$$

$$\sigma^{\mathrm{H}} = \frac{14\sigma}{b-a} \qquad (2.3.8)$$

Шаг 4. С помощью таблицы 2.2, а также с помощью интерполяционного полинома Лагранжа четвертой степени

$$\mu_j^{\mathrm{H}}(\varphi) = \sum_{i=0}^{4} [\mu_j^{\mathrm{H}}]^i \frac{(\hat{\varphi} - \varphi_0)\ldots(\hat{\varphi} - \varphi_{i-1})(\hat{\varphi} - \varphi_{i+1})\ldots(\hat{\varphi} - \varphi_n)}{(\varphi - \varphi_0)\ldots(\varphi - \varphi_{i-1})(\varphi - \varphi_{i+1})\ldots(\varphi - \varphi_n)}$$

определяются значения μ_1^{H}, μ_2^{H} соответствующие значениям ν_1^{H}, σ^{H}.

Здесь в качестве $\hat{\varphi}$ выступают либо значения ν_1^{H}, либо σ^{H}, а в качестве φ_i (i=0,1,...) те значения математического ожидания и среднеквадратического отклонения, для которых рассчитана таблица 2.2.

Таблица 2.2(a)

σ^{H}	ν_1^{H}	μ_1^{H}	μ_2^{H}
	0,5	***-2,0275***	***0,01365***
	1	***3,14991***	***-1,65984***
	1,5	***5,90479***	***-1,97136***
	2	***8***	***-2***
	2,5	***10***	***-2***
	3	***12***	***-2***
	3,5	***14***	***-2***
	4	***16***	***-2***
	4,5	***18***	***-2***
	5	***20***	***-2***
	5,5	***22***	***-2***
0,5	***6***	***24***	***-2***
	6,5	***26***	***-2***
	7	***28***	***-2***
	7,5	***30***	***-2***
	8	***32***	***-2***
	8,5	***34***	***-2***
	9	***36***	***-2***
	9,5	***38***	***-2***
	10	***40***	***-2***
	10,5	***42***	***-2***
	11	***44***	***-2***
	11,5	***46***	***-2***
	12	***48***	***-2***
	12,5	***50***	***-2***
0,5	***13***	***43,32561***	***-1,65984***
	13,5	***1,64249***	***0,01374***

Таблица 2.2(b)

σ^{H}	ν_1^{H}	μ_1^{H}	μ_2^{H}
	1	***-1,00879***	***-0,00219***
	1,5	***0,48945***	***-0,2668***
	2	***1,57495***	***-0,41496***
	2,5	***2,34474***	***-0,4732***
	3	***2,95239***	***-0,49284***
	3,5	***3,48819***	***-0,49844***
	4	***4***	***-0,5***
	4,5	***4,5***	***-0,5***
	5	***5***	***-0,5***
1	***5,5***	***5,5***	***-0,5***
	6	***6***	***-0,5***
	6,5	***6,5***	***-0,5***
	7	***7***	***-0,5***
	7,5	***7,5***	***-0,5***
	8	***8***	***-0,5***
	8,5	***8,5***	***-0,5***
	9	***9***	***-0,5***
	9,5	***9,5***	***-0,5***
	10	***10***	***-0,5***
	10,5	***10,5***	***-0,5***
	11	***11***	***-0,5***
	11,5	***10,90483***	***-0,4732***
	12	***10,04393***	***-0,41496***
1	***12,5***	***6,98095***	***-0,2668***
	13	***0,94747***	***0,00218***

Таблица 2.2(c)

$\sigma^{\text{Н}}$	$\nu_1^{\text{Н}}$	$\mu_1^{\text{Н}}$	$\mu_2^{\text{Н}}$
	1,5	***-0,67253***	***0,00097***
	2	***0,01662***	***-0,08248***
	2,5	***0,59821***	***-0,14675***
	3	***1,05013***	***-0,18452***
	3,5	***1,40839***	***-0,20446***
	4	***1,70658***	***-0,21441***
	4,5	***1,96862***	***-0,2191***
	5	***2,22222***	***-0,22222***
	5,5	***2,44444***	***-0,22222***
1,5	***6***	***2,66666***	***-0,22222***
	6,5	***2,88888***	***-0,22222***
	7	***3,11111***	***-0,22222***
	7,5	***3,33333***	***-0,22222***
	8	***3,77777***	***-0,22222***
	8,5	***4,0***	***-0,22222***
	9	***4,22222***	***-0,22222***
	9,5	***4,22222***	***-0,22222***
	10	***4,2969***	***-0,21441***
	10,5	***4,31649***	***-0,20446***
	11	***4,11642***	***-0,18452***
	11,5	***3,151079***	***-0,14675***
1,5	***12***	***2,29282***	***-0,08248***
	12,5	***0,64537***	***0,00096***

Таблица 2.2(d)

$\sigma^{\text{Н}}$	$\nu_1^{\text{Н}}$	$\mu_1^{\text{Н}}$	$\mu_2^{\text{Н}}$
	2	***-0,5***	***0***
	2,5	***-0,11146***	***-0,03518***
	3	***0,24473***	***-0,0667***
	3,5	***0,54152***	***-0,08908***
	4	***0,78748***	***-0,10374***
	4,5	***0,99416***	***-0,11286***
	5	***1,17239***	***-0,1183***
	5,5	***1,33102***	***-0,12146***
	6	***1,4762***	***-0,12321***
	6,5	***1,61281***	***-0,12414***
2	***7***	***1,75***	***-0,125***
	7,5	***1,86311***	***-0,12414***
	8	***1,97368***	***-0,1232***
	8,5	***2,06986***	***-0,12146***
	9,5	***2,16592***	***-0,11286***
	10	***2,11724***	***-0,10374***
	10,5	***1,95272***	***-0,08908***
	11	***1,62286***	***-0,0667***
	11,5	***1,09649***	***-0,03518***
	12	***0,5***	***0***

Таблица 2.2(e)

$\sigma^{н}$	$\nu_1^{н}$	$\mu_1^{н}$	$\mu_2^{н}$
	2,5	***-0,34999***	***-0,00307***
	3	***-0,2249***	***-0,0103***
	3,5	***-0,05***	***-0,03179***
	4	***0,3***	***-0,04991***
	4,5	***0,475***	***-0,0592***
	5	***0,625***	***-0,06601***
	5,5	***0,75***	***-0,06985***
	6	***0,875***	***-0,07381***
	6,5	***0,975***	***-0,07536***
2	***7***	***1,05***	***-0,07491***
	7,5	***1,13508***	***-0,07536***
	8	***1,19167***	***-0,0738***
	8,5	***1,20580***	***-0,06985***
	9	***1,22327***	***-0,06601***
	9,5	***1,1826***	***-0,0592***
	10	***1,09747***	***-0,04991***
	10,5	***0,94012***	***-0,03179***
	11	***0,51339***	***-0,0103***
	11,5	***0,43595***	***-0,00307***

Таблица 2.2(h)

$\sigma^{н}$	$\nu_1^{н}$	$\mu_1^{н}$	$\mu_2^{н}$
	4	***-0,47499***	***0,02282***
	4,5	***-0,39999***	***0,01928***
4	***5***	***-0,24999***	***0,00987***
	5,5	***-0,12499***	***0,00257***
	6	***0***	***-0,00419***
	6,5	***0***	***-0,0021***
	7	***-0,02031***	***0,00139***
	7,5	***-0,0186***	***0,00139***
4	***8***	***0,0588***	***-0,0021***
	8,5	***0,11732***	***-0,00419***
	9	***0,5303***	***0,00257***
	9,5	***-0,2076***	***0,00967***
	10	***-0,13984***	***0,01928***

Таблица 2.2(f)

$\sigma^{н}$	$\nu_1^{н}$	$\mu_1^{н}$	$\mu_2^{н}$
	3	***-0,47499***	***-0,01546***
	3,5	***-0,14999***	***-0,00946***
3	***4***	***-0,01874***	***-0,01725***
	4,5	***0,15***	***-0,02772***
	5	***0,275***	***-0,03423***
	5,5	***0,3875***	***-0,03808***
	6	***0,48125***	***-0,04209***
	6,5	***0,56562***	***-0,04434***
2	***7***	***0,625***	***-0,04459***
	7,5	***0,6759***	***-0,04431***
	8	***0,69727***	***-0,04209***
	8,5	***0,70394***	***-0,03898***
	9	***0,68344***	***-0,03423***
	9,5	***0,62616***	***-0,02772***
	10	***0,50174***	***-0,01725***
	10,5	***0,41487***	***-0,00946***
	11	***0,04211***	***-0,01546***

Таблица 2.2(g)

$\sigma^{н}$	$\nu_1^{н}$	$\mu_1^{н}$	$\mu_2^{н}$
	3,5	***-0,44999***	***0,01763***
	4	***-0,2***	***-0,00015***
	4,5	***-0,09999***	***-0,00597***
	5	***0,0125***	***0,01134***
	5,5	***0,05***	***-0,012***
3,5	***6***	***0,1875***	***-0,0192***
	6,5	***0,25***	***-0,2065***
	7	***0,3***	***-0,214***
	7,5	***0,3282***	***-0,2065***
	8	***0,35009***	***-0,01919***
	8,5	***0,28599***	***-0,01199***
	9	***0,30502***	***0,01134***
	9,5	***0,26714***	***-0,00596***
3,5	***10***	***0,20419***	***-0,00014***
	10,5	***-0,04364***	***0,01762***

Шаг 5. Определяются ненормированные значения μ_1^* и μ_2^* параметров распределения

$$\mu_1^* = \frac{14\mu_1}{b-a}$$

$$\mu_2^* = \frac{196\mu_2}{(b-a)^2} \qquad (2.3.9)$$

Шаг 6. Определяется масштабирующий параметр

$$\mu_0^* = -\ln\left(\int_0^{b^*} \exp\left(\sum_{j=1}^{2} \mu_j^* x^j\right)\, dx\right) \qquad (2.3.10)$$

Шаг 7. С помощью выражений (2.3.6) определяются параметры распределения, соответствующие исходному интервалу *[a, b]*.

Описание методики закончим рассмотрением примера.

Пример 2. Пусть в результате опыта получено 10 реализации случайной величины:

$x_1=0.44856$	$x_2=1.36945$	$x_3=0.80314$	$x_4=1.58958$	$x_5=0.67373$
$x_6=0.58740$	$x_7=1.36845$	$x_8=0.55475$	$x_9=0.51220$	$x_{10}=0.36568$

определенной на интервале [0; 2].

Требуется:

а) найти аналитическое выражение плотности распределения случайной величины *f(x)*;

б) построить оценку функции распределения *F(x)*.

Решение: По данным выборки определим значения двух первых начальных моментов распределения ν_1, ν_2 и величину среднеквадратического отклонения σ:

$\nu_1=0,82739$ $\qquad$ $\nu_2=0,86266$ $\qquad$ $\sigma=0,44481$

По формулам (2.8) определим нормированные значения математического ожидания и среднеквадратического отклонения:

$\nu_1^{н}=5.86172;$ $\qquad$ $\sigma^{н}=3.11366.$

С помощью таблицы 2.2 определим значения нормированных параметров распределения, соответствующих полученному значению при следующих значениях $\sigma^{н}$:

$\sigma_1^{н}=1,5$ $\qquad$ $\sigma_2^{н}=2$ $\qquad$ $\sigma_3^{н}=2,5$ $\qquad$ $\sigma_4^{н}=3$ $\qquad$ $\sigma_5^{н}=3,5$

Для этого, задаваясь поочередно перечисленными значениями $\sigma_i^{\text{н}}$ *(i=1,2,...,5)* с помощью таблицы 2.2 определим значения $\mu_1^{\text{н}}$, $\mu_2^{\text{н}}$, соответствующие следующим значениям $\nu_1^{\text{н}}$:

$\nu_{1.1}^{\text{н}}=4{,}5$ $\nu_{1.2}^{\text{н}}=5$ $\nu_{1.3}^{\text{н}}=5{,}5$ $\nu_{1.4}^{\text{н}}=6$ $\nu_{1.5}^{\text{н}}=6{,}5$

Полученные значения $\mu_1^{\text{н}}$, $\mu_2^{\text{н}}$, занесем во второй, третий, четвертый, пятый и шестой столбцы таблиц 2.3 и 2.4.

С помощью этих таблиц, а также с помощью интерполяционного полинома Лагранжа четвертой степени, определим значения $\mu_1^{\text{н}}$, $\mu_2^{\text{н}}$ при различных значениях $\sigma_i^{\text{н}}$ *(i = 1,2,...,5)* и заданном значении $\nu_1^{\text{н}}$. Полученные значения представлены в седьмых столбцах таблиц 2.3 и 2.4.

Сведем значения $\mu_1^{\text{н}}$ и $\mu_2^{\text{н}}$, соответствующие одному и тому же значению $\nu_1^{\text{н}}$, по различным $\sigma_i^{\text{н}}$, во второй, третьей, четвертой, пятой и шестой столбцы таблицы 2.5.

По этой таблице с помощью интерполяционного полинома Лагранжа четвертой степени определим значения параметров распределения, соответствующие полученному значению $\sigma^{\text{н}}$**.** Полученные значения $\mu_1^{\text{н}}$ и $\mu_2^{\text{н}}$ занесем в седьмой столбец таблицы 2.5.

С помощью формул (2.9) определим ненормированные значения μ_1^* и μ_2^*:

$\mu_1^*=2{,}66133;$ $\mu_2^*=-1{,}72627.$

С помощью формулы (2.10) определим масштабирующий параметр μ_0^*:

$$\mu_2^*=1{,}24388$$

В силу того, что по условию задачи нижняя граница интервала [a, b] совпадает с началом координат, значения ненормированных параметров μ_j^* *(j=0,1,2),* распределения совпадают с искомыми.

Оценка плотности распределения при этом запишется в виде:

$$f(x)=\exp(-1{,}24388+2{,}66133x-1{,}72627x^2).$$

На рисунке 2.2 сплошной линией показана оценка функции распределения, полученная с помощью информационного метода. Ломаной линией показана эмпирическая функция распределения. Исходные данные для примера получены с помощью датчика случайных чисел, подчиняющегося экспоненциальному распределению с параметрами λ=0,5 и интервалом усечения [0; 2] . Заданная функция распределения показана на рисунке пунктирной линией. Как видно из рисунка наибольшее абсолютное уклонение оценки $F(x)$ от заданной функции распределения $F_з\ (x)$ составляет 0,09 при x=0.4. Наивысшее абсолютное уклонение эмпирической функции распределения от заданной составляет 0,22 и достигается при x = 0,80314.

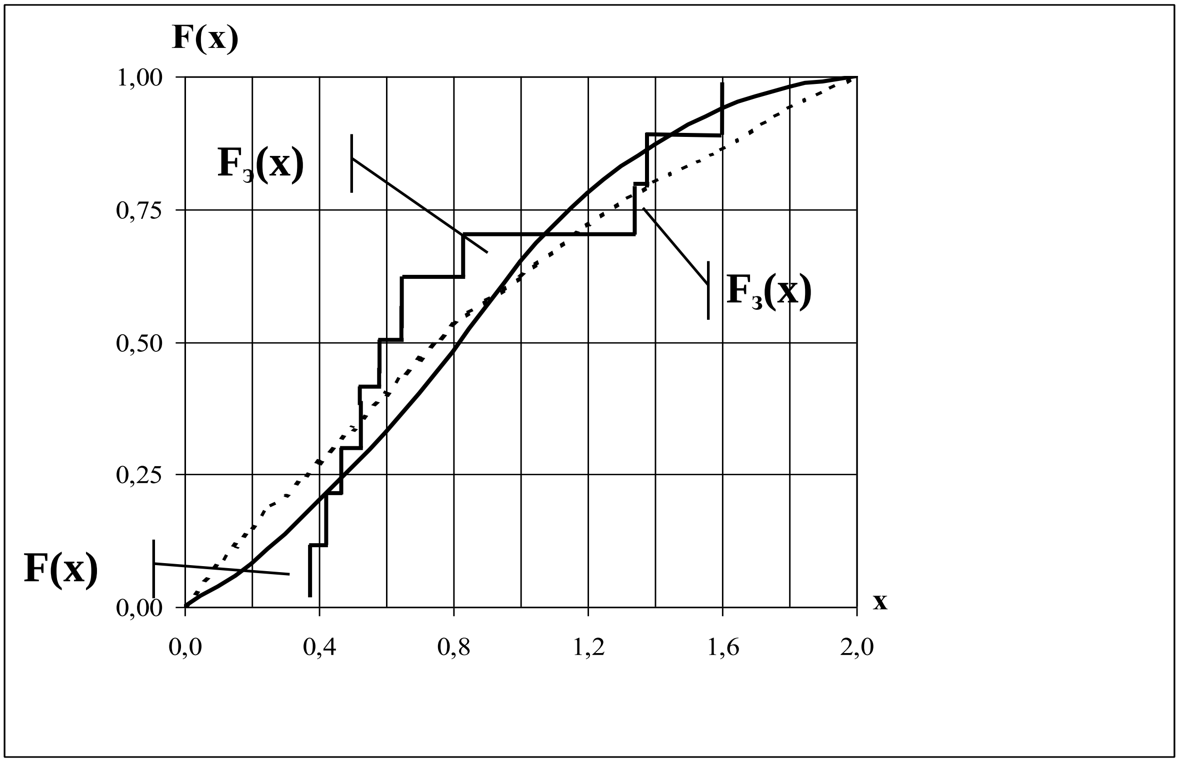

Рис. 2.2

Таблица 2.3.

$\sigma^{\text{н}}$ \ $\nu_1^{\text{н}}$	**4.5**	**5**	**5.5**	**6**	**6.5**	$\nu_1^{\text{н}}$
3.5	-0.0999	0.0125	0.05	0.1875	0.25	0.14310
3	0.15	0.275	0.3875	0.48125	0.56562	0.45694
2.5	0.475	0.625	0.75	0.875	0.975	0.84094
2	0.99416	1.17239	1.33102	1.4762	1.6128	1.43709
1.5	1.968862	2.22222	2.44444	2.66666	2.88888	2.60460

Таблица 2.4.

$\sigma^{\text{н}}$ \ $\nu_1^{\text{н}}$	**4.5**	**5**	**5.5**	**6**	**6.5**	$\nu_1^{\text{н}}$
3.5	-0.00597	-0.01134	-0.012	-0.0192	-0.02065	-0.01682
3	-0.02772	-0.03423	-0.03898	-0.04209	-0.04434	-0.04136
2.5	-0.0592	-0.06601	-0.06985	-0.07381	-0.07536	-0.07274
2	-0.11286	-0.1183	-0.12146	-0.12321	-0.12414	-0.12282
1.5	-0.21910	-0.22222	-0.22222	-0.22222	-0.22222	-0.22214

Таблица 2.5.

$\sigma^{\text{Н}}$	**3.5**	**3**	**2.5**	**2**	**1.5**	$\sigma^{\text{Н}}$
$\mu_1^{\text{Н}}$	0.14310	0.45694	0.84094	1.43709	2.60460	0.38019
$\mu_2^{\text{Н}}$	-0.01682	-0.04136	-0.07274	-0.12282	-0.22214	-0.03523

3.Порядок выполнения работы

3.1 Задание №1

1. На основе выборочных данных (в соответствии с заданием, выданным преподавателем), построить оценки законов распределения наработки до отказа следующими методами:

– последовательных медиан;
– уменьшения неопределенности;
– сжатия ИЗР.

2. Составить полученные оценки на основе следующих метрик:

– $D_1 = \max |F_i(x) - F_j(x)|$

– $D_2 = \int\limits_a^b |F_i(x) - F_j(x)| \quad dx$

– $D_3 = \int\limits_a^b (F_i(x) - F_j(x))^2 dx$

Здесь i, j – идентификаторы используемых методов ($i, j = \overline{1;k}$);

a, b – границы области изменения случайной величины x.

Результаты сопоставления оформить в виде таблицы вида:

Таблица 1

Метод / *метод*	μ_1	**...**	μ_k
μ_1			
...			
μ_k			

3.2 Задание №2

1. На основе выборочных данных (в соответствии с заданием, выданным преподавателем), построить оценку закона распределения на основе информационного метода.

2. Произвести сопоставление полученной оценки с нормальным законом распределения параметрами: матожидание – 8, среднеквадратическое отклонение – 2. Сопоставление произвести на основе метрик, описанных в п.2 задания №1. Результаты сопоставления оформить в виде таблицы вида:

Таблица 2

Наименование метрики	D_1	D_2	D_3
Значение метрики			

3.3 Задание №3

1. На основе выборочных данных (в соответствии с заданием, выданным преподавателем), построить оценку усеченного закона распределения.

2. Произвести сопоставление с равномерным законом распределения, заданным на интервале [0; 10]. Сопоставление произвести на основе метрик, описанных в п.2 задания №1. Результаты сопоставления оформить в виде таблицы 2.

4. Требования к отчету

Отчет по лабораторной работе должен содержать:

1) титульный лист
2) название и цель работы
3) результаты выполнения заданий по п.п. 3.1 – 3.2.
4) выводы.

5. Контрольные вопросы

1. В чем заключается основная идея метода прямоугольных вкладов?
2. Условия применимости метода прямоугольных вкладов.
3. В чем отличие метода уменьшения неопределенности от метода прямоугольных вкладов?
4. Почему в качестве априорной компоненты метода прямоугольных вкладов целесообразно использовать равномерное распределение?
5. Преимущества и недостатки метода сжатия области существования интегральных законов распределения?

Список литературы

1. Дружинин Г.В. Надёжность автоматизированных производственных систем - 4-е изд., перераб. и доп. - М.: Энергоатомиздат, 1986.
2. Вентцель Е.С., Овчаров Л.А., Теория случайных процессов и ее инженерные приложения. Уч. пособие для ВУЗов. – М.: Высшая школа, 2000.
3. Вентцель Е.С.Теория вероятностей :Учебник для ВТУЗов. 6-е изд.,испр.- М.:Высшая школа, 1999.

Составители: Гвоздев Владимир Ефимович
Тагирова Клара Фоатовна

СТАТИСТИЧЕСКИЙ АНАЛИЗ НАДЕЖНОСТИ

МЕТОДИЧЕСКИЕ УКАЗАНИЯ
к лабораторным работам по курсам
«Надежность АСОИиУ», «Надежность и эффективность в технике», «Теория надежности систем» для студентов специальностей 210100 «Управление и информатика в технических системах», 220200 «Автоматизированные системы обработки информации и управления» направлений 552800 «Информатика и ВТ», 550200 «Автоматизация и управление»