

АНАЛИЗ СТАТИСТИКИ ЛУЧШИХ ЗООПАРКОВ МИРА И РОССИИ

ВЕРХОГЛЯДОВА Александра Владимировна

магистрант

ФГБОУ ВО «Донской государственный технический университет»

г. Ростов-на-Дону, Россия

Лицо, собирающееся открыть зоопарк, вынуждено решать, какие виды ему стоит включить в список своего зоопарка, а какие придётся оставить без внимания. В этом деле он может опираться на опыт уже устоявшихся больших зоопарков, а может из своих личных приоритетов подобрать животных. При этом необходимо понимать разницу государственных и частных зоопарков, начальный капитал которых может сильно отличаться, так же, как и имеющаяся площадь.

Ключевые слова: математический анализ, статистика, зоопарк, анализ статистики, видовой состав, функции распределения.

Перед началом разработки математической модели видовой состава городского зоопарка мы проводим дополнительные исследования, необходимые для корректного построения модели. Исследования основываются на анализе собранной статистики по лучшим зоопаркам мира. Полученные результаты станут обоснованием выбора параметров и ограничений для оптимизации, примером для сравнения полученных моделей, выявления преимуществ.

Для корректного исследования необходимо

ознакомиться с реальными объектами, то есть провести сбор данных по зоопаркам на предмет видовой состава кошачьих. Для этого была проведена выборка из сорока зоопарков (30 лучших зоопарков мира и 10 лучших зоопарков России) и построение таблицы по полученным данным. Выбор зоопарков осуществлялся из статей, которые на основе ряда параметров составляли топы лучших зоопарков мира, Европы и России. Демонстрация видовой разнообразия находится в открытом доступе, на сайтах зоопарков.

Таблица 1

ВИДОВОЕ РАЗНООБРАЗИЕ ЗООПАРКОВ МИРА

Species/ Name of the zoo	Сингапурский зоопарк	Зоопарк Торонто	Пражский зоопарк	Зоопарк Рануа	Эдинбургский зоопарк	Смитсоновский национальный зоологический парк	Зоопарк Цюриха
Asian Lion					1		1
Amur Tiger		1				1	1
African Lion	1	1	1			1	
The Dune Cat						1	
White Lion							
White Tiger	1						
Bengal Cat							
Bengal Tiger			1				
East Siberian Lynx							
Cheetah	1	1	1			1	
The Far Eastern Leopard							
Far Eastern Forest Cat							
Long - tailed cat							

Clouded leopard		1				1	
Indochinese Leopard							
Indochina Tiger							
Caucasian Leopard							
Jungle cat							
Canadian Lynx		1					
Caracal			1			1	
Temmincka the Cat							
The Angler Cat			1			1	
Geoffrey's Cat							
Leopard	1						
Liligr							
The Malay Tiger							
Manul				1			
Ocelot							
Persian Leopard							
Cougar		1					
Red Lynx							
Lynx				1		1	
Savannah							
Serval							
Siberian Lynx							
Snow Leopard		1					1
Steppe Cat							
Sumatran Cat							
Sumatran Tiger		1			1		
Florida Cougar							
Black - legged cat							
Jaguar		1					
Jaguarundi			1				
The sum of species:	4	9	6	2	2	8	3

После проведения оптимизации необходимо будет проверить модель на корректность по расходам. Для этого имеющиеся данные надо проанализировать, провести исследования на нахождение лучшего графика, описывающего реальные данные.

По собранным данным мы строим таблицу,

по которой будем строить график распределения величины расходов зоопарков на кошачьих. Вместе с этим вычисляем основные показатели, на которые мы будем опираться в дальнейшем, при поиски необходимой функции. Такими показателями являются асимметрия, эксцесс, математическое ожидание, дисперсия, СКО.

Таблица 2

РАСЧЁТ ОСНОВНЫХ ПОКАЗАТЕЛЕЙ

Среднее значение	161 832 325,00
Сигма (СКО)	71 404 874,75
Корень из дисп делим на МО	0,44
Дисперсия	5 098 656 137 858 070
Мат ожидание	161 832 325
Асимметрия	1,03704
Эксцесс исх.дан.	5,10771

Таблица 3

РАСЧЁТ ДАННЫХ ДЛЯ ПОСТРОЕНИЯ ГРАФИКА ПО ИСХОДНЫМ ДАННЫМ И ГРАФИК НОРМАЛЬНОГО РАСПРЕДЕЛЕНИЯ

Расчёты, по исходным данным						
Исходные данные	Вероятность				Kurt - 3	
18 000 000	0,000					
19 729 749	0,000	Норм. Расп	p(x)	F(x)	Асимметрия	1,03704
19 729 750	0,025	0,13525	0,00731	0,00731	Экспесс исх.дан.	5,10771
64 850 500	0,050	0,16087	0,00869	0,01599	Экспесс p(x)	-1,46227
99 894 250	0,075	0,20563	0,01111	0,02710		
100 880 000	0,100	0,25823	0,01395	0,04105		
100 880 000	0,125	0,31860	0,01721	0,05826		
100 900 000	0,150	0,38619	0,02086	0,07912		
104 650 500	0,175	0,45990	0,02484	0,10396		
113 772 000	0,200	0,53806	0,02906	0,13302		
113 772 000	0,225	0,61847	0,03341	0,16643		
113 792 000	0,250	0,69842	0,03772	0,20415		
117 050 000	0,275	0,77487	0,04185	0,24601		
117 802 000	0,300	0,84461	0,04562	0,29163		
117 930 500	0,325	0,90446	0,04885	0,34048		
127 150 000	0,350	0,95157	0,05140	0,39188		
131 856 250	0,375	0,98357	0,05313	0,44501		
134 258 750	0,400	0,99880	0,05395	0,49896		
136 320 500	0,425	0,99647	0,05382	0,55278		
136 822 500	0,450	0,97671	0,05276	0,60554		
137 036 000	0,475	0,94055	0,05080	0,65634		
138 616 700	0,500	0,88983	0,04806	0,70440		
158 403 500	0,525	0,82707	0,04467	0,74908		
160 329 950	0,550	0,75526	0,04079	0,78987		
160 672 250	0,575	0,67758	0,03660	0,82647		
169 210 000	0,600	0,59722	0,03226	0,85873		
170 431 000	0,625	0,51716	0,02793	0,88666		
177 901 300	0,650	0,43997	0,02376	0,91043		
179 348 750	0,675	0,36774	0,01986	0,93029		
179 348 750	0,700	0,30197	0,01631	0,94660		
184 157 000	0,725	0,24362	0,01316	0,95976		
184 261 750	0,750	0,19309	0,01043	0,97019		
198 428 250	0,775	0,15036	0,00812	0,97831		
214 010 500	0,800	0,11503	0,00621	0,98452		
214 704 750	0,825	0,08646	0,00467	0,98919		
215 589 750	0,850	0,06384	0,00345	0,99264		
216 897 500	0,875	0,04631	0,00250	0,99514		
231 178 750	0,900	0,03301	0,00178	0,99693		
233 467 950	0,925	0,02311	0,00125	0,99818		
234 154 250	0,950	0,01590	0,00086	0,99903		
324 376 550	0,975	0,01075	0,00058	0,99961		
438 456 300	1,000	0,00714	0,00039	1,00000		
500 000 000	1,000	18,51365				

Далее мы предполагаем, что исходный график имеет нормальное распределение. Для проверки гипотезы, необходимо сначала

построить таблицу, соответствующую нормальному распределению, а далее по полученным данным строим график.

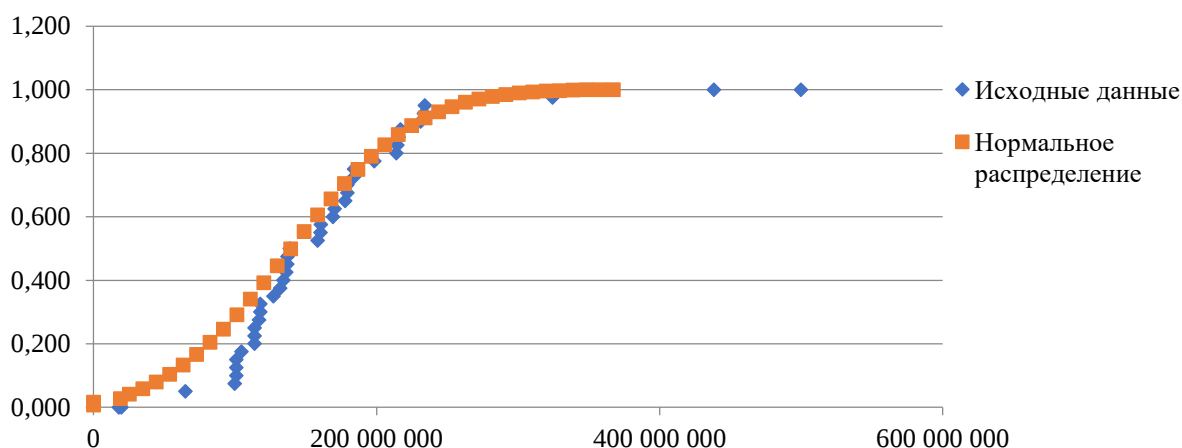


Рисунок 1. Сравнение графика исходных данных с функцией нормального распределения

Очевидно, что для наших данных не подходит нормальное распределение, так же это проверяем по асимметрии и эксцессу.

Таблица 4

РАСЧЁТ ОСНОВНЫХ ПОКАЗАТЕЛЕЙ

Асимметрия	0,338832
Эксцесс	-0,78083

Следовательно, мы должны выдвинуть другую гипотезу, о функции распределения данных. Предположим, что данные имеют распределение Рэля-Райса.

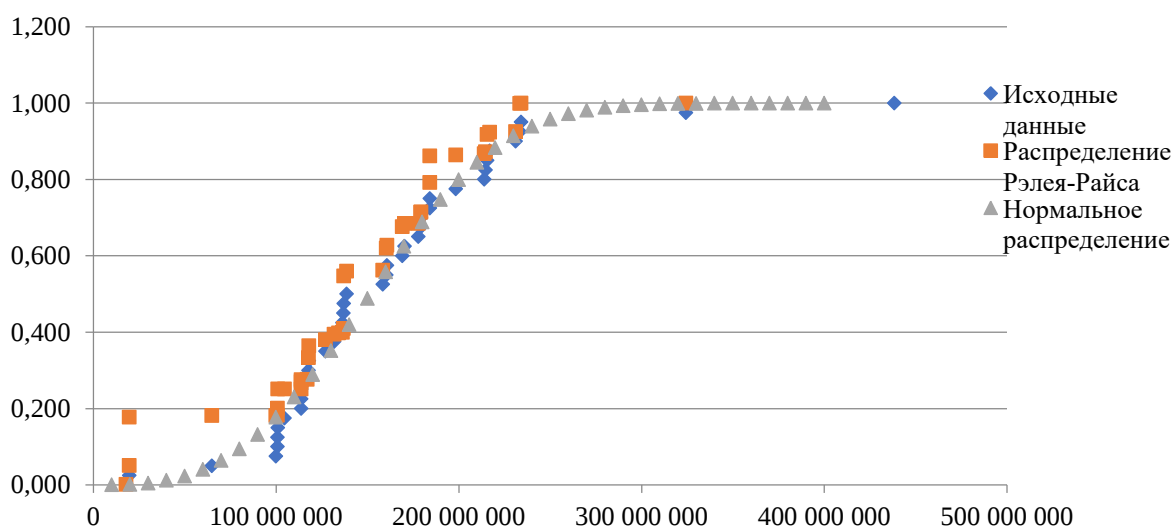


Рисунок 2. Сравнение графика исходных данных с функцией распределения Рэля-Райса

Таблица 5

РАСЧЁТ ДАННЫХ ДЛЯ ПОСТРОЕНИЯ ФУНКЦИИ РАСПРЕДЕЛЕНИЯ РЭЛЕЯ-РАЙСА

Распределение Рэлея-Райса				
n	D	F(x) theor		
3,04	171 068 575			
Невязка	0,06	0,00141	-0,02	0,00056
		0,05109	0,00	0,00000
		0,17712	0,10	0,01043
		0,18197	0,08	0,00672
		0,18197	0,06	0,00325
		0,18207	0,03	0,00103
		0,20113	0,03	0,00068
		0,25135	0,05	0,00264
		0,25135	0,03	0,00069
		0,25147	0,00	0,00000
		0,27064	0,00	0,00002
		0,27515	-0,02	0,00062
		0,27592	-0,05	0,00241
		0,33359	-0,02	0,00027
		0,36444	-0,01	0,00011
		0,38049	-0,02	0,00038
		0,39439	-0,03	0,00094
		0,39780	-0,05	0,00273
		0,39925	-0,08	0,00574
		0,41001	-0,09	0,00810
		0,54685	0,02	0,00048
		0,56008	0,01	0,00010
		0,56242	-0,01	0,00016
		0,61991	0,02	0,00040
		0,62795	0,00	0,00001
		0,67580	0,03	0,00067
		0,68478	0,01	0,00010
		0,68478	-0,02	0,00023
		0,71383	-0,01	0,00012
		0,71445	-0,04	0,00126
		0,79191	0,02	0,00029
		0,86127	0,06	0,00375
		0,86396	0,04	0,00152
		0,86733	0,02	0,00030
		0,87221	0,00	0,00001
		0,91770	0,02	0,00031
		0,92371	0,00	0,00000
		0,92545	-0,02	0,00060
		0,99908	0,02	0,00058
		1,00000	0,00	0,00000
		1,00000	0,00	0,00000

В этом случае распределение также нас не устраивает. Погрешность возникает из-за маленьких зоопарков и больших, которые сильно отклоняются от средних значений. После про-

ведения двух исследований, приходим к выводу: оптимальным выбором будет проведение анализа подгонки по функции распределения и интегральным показателям.

Таблица 6

РАСЧЁТ ОСНОВНЫХ ПОКАЗАТЕЛЕЙ

Выб дисп/Мод дисп	0,87
МО	144 979 774,96
Дисперсия	4441027809817810,00
Асимметрия	1,45
Эксцесс	6,72
ЦФ	0,04

Таблица 7

РАСЧЁТ ДАННЫХ ДЛЯ ПОСТРОЕНИЯ ГРАФИКОВ ПОДГОНКИ ФУНКЦИИ РАСПРЕДЕЛЕНИЯ И ИНТЕГРАЛЬНЫХ ПОКАЗАТЕЛЕЙ

Подгонка функции распределения						
DD	nn					
147 981 500	4,44	Гстеп	Разница	Квадрат разницы		
Невязка	0,05	0,000130	-0,024870	0,000619	10 000 000	0,000179
		0,025019	-0,024981	0,000624	20 000 000	0,001468
		0,148718	0,073718	0,005434	30 000 000	0,005024
		0,154322	0,054322	0,002951	40 000 000	0,012001
		0,154322	0,029322	0,000860	50 000 000	0,023509
		0,154437	0,004437	0,000020	60 000 000	0,040559
		0,176799	0,001799	0,000003	70 000 000	0,064008
		0,237369	0,037369	0,001396	80 000 000	0,094493
		0,237369	0,012369	0,000153	90 000 000	0,132366
		0,237510	-0,012490	0,000156	100 000 000	0,177640
		0,260948	-0,014052	0,000197	110 000 000	0,229943
		0,266468	-0,033532	0,001124	120 000 000	0,288506
		0,267415	-0,057585	0,003316	130 000 000	0,352178
		0,337693	-0,012307	0,000151	140 000 000	0,419469
		0,374671	-0,000329	0,000000	150 000 000	0,488641
		0,393633	-0,006367	0,000041	160 000 000	0,557817
		0,409894	-0,015106	0,000228	170 000 000	0,625115
		0,413847	-0,036153	0,001307	180 000 000	0,688789
		0,415527	-0,059473	0,003537	190 000 000	0,747353
		0,427944	-0,072056	0,005192	200 000 000	0,799685
		0,574970	0,049970	0,002497	210 000 000	0,845085
		0,588030	0,038030	0,001446	220 000 000	0,883295
		0,590322	0,015322	0,000235	230 000 000	0,914466
		0,644548	0,044548	0,001984	240 000 000	0,939092
		0,651827	0,026827	0,000720	250 000 000	0,957914
		0,693713	0,043713	0,001911	260 000 000	0,971819
		0,701303	0,026303	0,000692	270 000 000	0,981738
		0,701303	0,001303	0,000002	280 000 000	0,988561
		0,725313	0,000313	0,000000	290 000 000	0,993084
		0,725816	-0,024184	0,000585	300 000 000	0,995969
		0,786229	0,011229	0,000126	310 000 000	0,997738
		0,837259	0,037259	0,001388	320 000 000	0,998780
		0,839209	0,014209	0,000202	330 000 000	0,999368
		0,841658	-0,008342	0,000070	340 000 000	0,999686
		0,845204	-0,029796	0,000888	350 000 000	0,999851
		0,878742	-0,021258	0,000452	360 000 000	0,999932
		0,883327	-0,041673	0,001737	370 000 000	0,999970
		0,884663	-0,065337	0,004269	380 000 000	0,999988
		0,970239	-0,004761	0,000023	390 000 000	0,999995
		0,992016	-0,007984	0,000064	400 000 000	0,999998

Продолжение Таблицы 7

**РАСЧЁТ ДАННЫХ ДЛЯ ПОСТРОЕНИЯ ГРАФИКОВ ПОДГОНКИ
ФУНКЦИИ РАСПРЕДЕЛЕНИЯ И ИНТЕГРАЛЬНЫХ ПОКАЗАТЕЛЕЙ**

Подгонка интегральных показателей					
Рстеп	Рстеп*dx	МО	Дисп	АС	ЭК
0,00071	0,00268741	53 021,937	42158936750767,00	-0,02	0,033533
0,05935	0,14625724	9 484 855,116	939073961403269,00	-0,254253	0,305714
0,24298	0,242964912	24 270 797,683	493875885174829,00	-0,075237	0,050901
0,24985	0,006895714	695 639,649	13410717079986,60	-0,001998	0,001322
0,24985	0	0,000	0,00	0,000000	0,000000
0,24999	0,000139932	14 119,186	271892508580,77	-0,000040	0,000027
0,27663	0,027002023	2 825 775,214	43917451700682,30	-0,005985	0,003622
0,34376	0,068098801	7 747 736,844	66323140082745,30	-0,006994	0,003275
0,34376	0	0,000	0,00	0,000000	0,000000
0,34391	0,00014932	16 991,369	145239726715,14	-0,000015	0,000007
0,36827	0,024382036	2 853 917,302	19019751534499,60	-0,001795	0,000752
0,37390	0,005626724	662 839,311	4156075133184,94	-0,000382	0,000156
0,37486	0,000961425	113 381,305	703439205277,26	-0,000064	0,000026
0,44324	0,067373826	8 566 581,940	21418198149270,00	-0,001290	0,000345
0,47722	0,033527793	4 420 849,055	5774388095866,16	-0,000256	0,000050
0,49421	0,016842061	2 261 194,045	1935832815593,36	-0,000070	0,000011
0,50855	0,01423248	1 940 178,843	1067194686021,39	-0,000031	0,000004
0,51201	0,003451663	472 265,161	229677572906,83	-0,000006	0,000001
0,51348	0,001465482	200 823,762	92477118427,40	-0,000002	0,000000
0,52426	0,010708865	1 484 427,581	433588283349,41	-0,000009	0,000001
0,64554	0,107619846	17 047 360,247	19392708141303,20	0,000880	0,000177
0,65589	0,010210942	1 637 119,784	2405982500429,80	0,000125	0,000029
0,65770	0,001805906	290 158,973	444711155378,43	0,000024	0,000006
0,70015	0,039884944	6 748 931,444	23416602649881,20	0,001917	0,000697
0,70580	0,00560126	954 628,416	3628299653920,05	0,000312	0,000119
0,73819	0,03055662	5 436 062,503	33118084508813,10	0,003684	0,001820
0,74404	0,005787106	1 037 910,244	6835882757496,69	0,000794	0,000409
0,74404	0	0,000	0,00	0,000000	0,000000
0,76255	0,017800489	3 278 084,633	27321168700425,90	0,003617	0,002126
0,76294	0,000387136	71 334,319	597379012198,07	0,000079	0,000047
0,80965	0,041428575	8 220 599,666	118350646395941,00	0,021374	0,017143
0,84972	0,035008263	7 492 135,891	166822810666717,00	0,038911	0,040306
0,85127	0,001541478	330 962,569	7494004703737,48	0,001766	0,001847
0,85322	0,001935706	417 318,417	9650983165182,65	0,002303	0,002440
0,85605	0,002797566	606 785,066	14469456512794,30	0,003516	0,003795
0,88318	0,023997358	5 547 679,123	178306685115729,00	0,051933	0,067175
0,88696	0,003701623	864 210,218	28984285837066,00	0,008666	0,011507
0,88806	0,001097053	256 879,642	8723861547032,95	0,002629	0,003517
0,96437	0,035984969	11 672 680,198	1158111568622250,00	0,702004	1,889786
0,98826	0,0113661	4 983 538,303	978944841149533,00	0,970748	4,275018

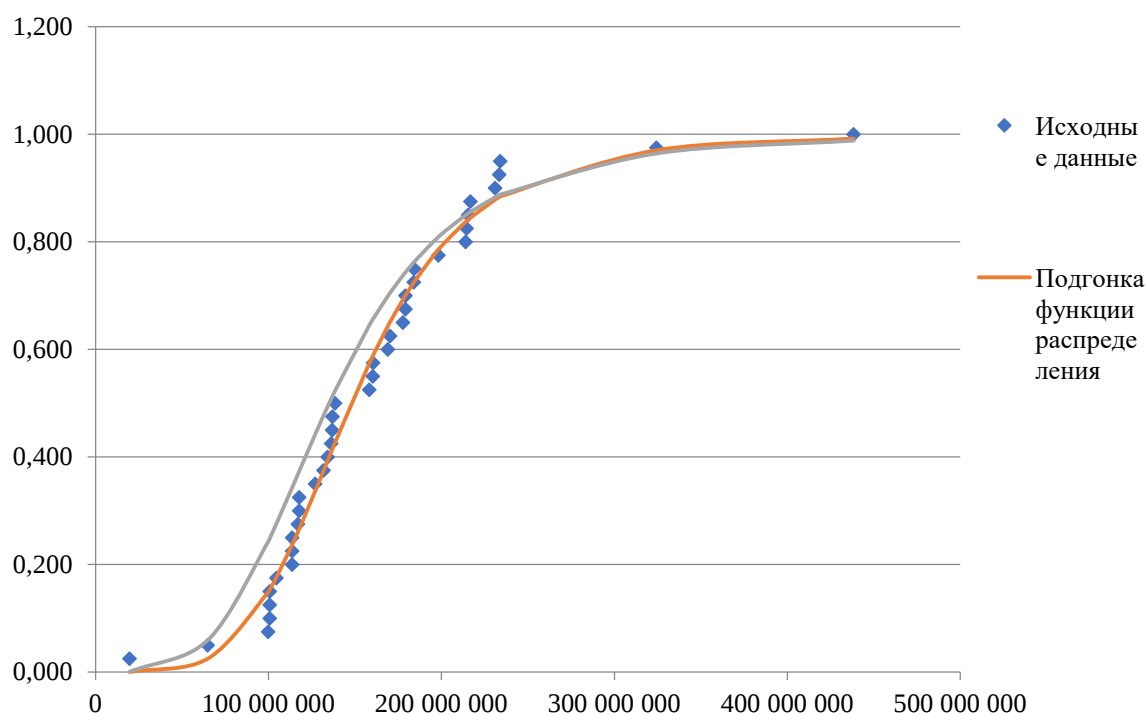


Рисунок 3. Сравнение графика исходных данных с подгонкой функции распределения и интегральных показателей

Сопоставив все три графика, можно прийти к следующему выводу: границы графиков подгонки функции распределения и подгонки интегральных показателей можно считать неким доверительным интервалом для исходных данных. То есть в области, находящейся

между данными графиками, с некоторой долей погрешности, будут лежать другие реальные данные по расходам зоопарков, не вошедших в работу. Вместе с этим, данные выводы будут использоваться для сравнения нашей математической модели с реальными объектами.

СПИСОК ЛИТЕРАТУРЫ

1. Азимова Н.Н., Ладоса Е.Н., Холодова С.Н., Цымбалов Д.С., Яценко О.В. Статистический анализ размерных характеристик пыли, образующейся при механической обработке металлов // Вестник Донского государственного технического университета. – 2020. – № 20(1). – С. 68-78. – URL: <https://doi.org/10.23947/1992-5980-2020-20-1-68-78>.
2. Обзор методов статистического анализа данных. – URL: <http://www.statlab.kubsu.ru/node/4> (дата обращения: 20.04.2022).
3. Практическая статистика для специалистов Data Science: пер. с англ. / П. Брюс, Э. Брюс. – СПб.: БХВ-Петербург, 2018 – 304 с.: ил.
4. Статистические методы анализа: [учеб. пособие] / И.С. Шорохова, Н.В. Кисляк, О.С. Мариев; М-во образования и науки рос. Федерации, Урал. федер. ун-т. – Екатеринбург: изд-во Урал. ун-та, 2015 – 300 с.
5. Статистический анализ данных: учеб. пособие / Е.А. Денискина, П.Э. Коломиец; Самар, гос. аэрокосм. ун-т. – Самара, 2006. – 64 с.