Министерство науки и высшего образования Российской Федерации
Федеральное государственное бюджетное образовательное учреждение
высшего образования
«Рязанский государственный радиотехнический университет
имени В.Ф.Уткина»
Кафедра «Электронные вычислительные машины»

Отчёт о практическом занятии № 1
«Вероятностно- статистический аппарат эконометрики»
по дисциплине
«Эконометрика»

Выполнил:
Студент группы 1714
Субботин Д.А.
Проверил:
доц. каф. ЭВМ
Хруничев Р. В.

Рязань 2023

Цель работы: оценка свойств случайной величины (СВ) в генеральной совокупности по эмпирическим данным (выборке) путем построения эмпирического распределения, нахождения числовых характеристик СВ по выборке, нахождения точечных и интервальных оценок параметров распределения СВ в генеральной совокупности, проверки гипотезы о нормальном распределении СВ в генеральной совокупности.

Практическая часть

Вариант 8. СВ – время безотказной работы оборудования (тыс.

час.) по результатам 110 наблюдений.

По заданной выборке значений изучаемого признака генеральной совокупности (рисунок 1)

14,64	12,7	14,9	11,33	8,84	12,87	13,78	13,51	11,75	11,2	11,91
12,72	10,26	13,48	12,06	13,64	12,93	11,07	14,21	11,6	11,55	17,22
14,74	12,68	12,72	16,81	14,29	11,47	8,29	12,79	12,29	11,27	8,8
13,22	11,81	10,52	13,27	9,7	12,73	14,68	12,21	12,87	15,89	13,65
13,52	14,15	11,71	12,78	8,39	11,26	12,26	10,91	14,23	12,5	13,7
11,47	13,72	11,85	13,64	15,53	12,77	17,04	9,07	11,83	10	12,62
15,59	10,52	13,31	12,95	12,46	9,49	14,49	14,25	14,01	16,13	13,66
11,6	13,06	12,15	13,48	10,84	14,08	11,08	12,1	13,32	17,71	12,91
17,39	15,55	13,55	15,54	13,67	11,28	14,84	9,22	11,94	13,7	11,7
13,2	12,87	13,56	15,78	11,43	12,75	12,95	12,5	14,02	12,54	16,26

Рисунок 1 – Массив данных

Произведем ранжирование исходных данных в порядке возрастания (рисунок 2).

8,29	8,39	8,8	8,84	9,07	9,22	9,49	9,7	10	10,26	10,52
10,52	10,84	10,91	11,07	11,08	11,2	11,26	11,27	11,28	11,33	11,43
11,47	11,47	11,55	11,6	11,6	11,7	11,71	11,75	11,81	11,83	11,85
11,91	11,94	12,06	12,1	12,15	12,21	12,26	12,29	12,46	12,5	12,5
12,54	12,62	12,68	12,7	12,72	12,72	12,73	12,75	12,77	12,78	12,79
12,87	12,87	12,87	12,91	12,93	12,95	12,95	13,06	13,2	13,22	13,27
13,31	13,32	13,48	13,48	13,51	13,52	13,55	13,56	13,64	13,64	13,65
13,66	13,67	13,7	13,7	13,72	13,78	14,01	14,02	14,08	14,15	14,21
14,23	14,25	14,29	14,49	14,64	14,68	14,74	14,84	14,9	15,53	15,54
15,55	15,59	15,78	15,89	16,13	16,26	16,81	17,04	17,22	17,39	17,71

Рисунок 2 – Проранжированный массив данных во возрастанию

Составить интервальный вариационный ряд по ранжированному дискретному массиву.

Наименьшее значение выборки – 8,29, а наибольшее – 17,71; размах вариации R = 17,71-8,29=9,42. Число частичных интервалов можно оценить по правилу Стерджесса:

$$r = 1 + 3,3\lg n = 1 + 3,3\lg 110 = 7,73 \approx 8.$$

r=8, тогда размер каждого интервала составит: $\Delta = \frac{R}{r} = \frac{17,71}{8} = 2,21$

Находим частоту, соответствующую каждому интервалу, она будет равна сумме частот вариант, попавших в данный интервал, учитывая, что левый конец принадлежит интервалу. Также рассчитываем относительные частоты. Получим интервальный вариационный ряд частот (рисунок 3).

11,81	12,72	16,81	14,29	11,47	8,29	12,79	12,29	11,27	8,8
12,68	10,52	13,27	9,7	12,73	14,68	12,21	12,87	15,89	13,65
12,7	11,71	12,78	8,39	11,26	12,26	10,91	14,23	12,5	13,7
12,87	11,85	13,64	15,53	12,77	17,04	9,07	11,83	10	12,62

Рисунок 3 — Интервальный вариационный ряд частот

Проверяем, что сумма абсолютных частот равна объему выборки:

$$n = \sum ni = 110,$$

а сумма относительных частот – единице:

$$w = \sum wi = 1.$$

Построим гистограмму относительных частот (рисунок 4).

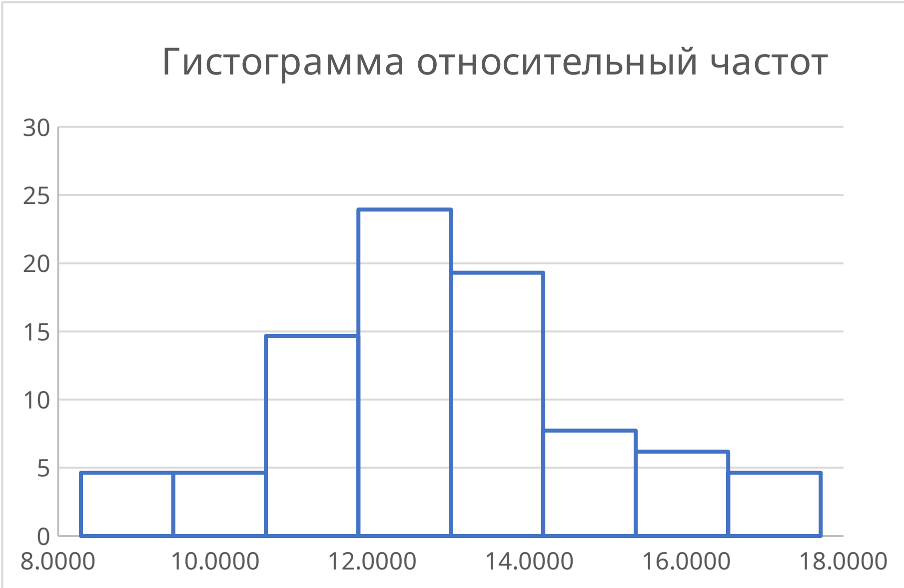

Рисунок 4 – Гистограмма

Чтобы построить кумуляту и эмпирическую функцию распределения, подсчитаем накопленные относительные частоты рисунок 5.

Частичный интервал Δ = 1,1775	Середины интервалов xi	Абсолютные частоты ni	Относительные частоты wi	Накопл. Отн. Частоты wiнак	xiwi	$(xi)^2wi$	$(xi-xcp)^3wi$	(xi-xcp)4wi
8,29-9,4675	8,879	6	0,054545455	0,054545455	0,4843	4,2999	-3,3890	13,4227
9,4675-10,645	10,056	6	0,054545455	0,109090909	0,5485	5,5161	-1,1759	3,2728
10,645-11,8225	11,234	19	0,172727273	0,281818182	1,9404	21,7977	-0,7151	1,1482
11,8225-13	12,411	31	0,281818182	0,563636364	3,4977	43,4110	-0,0221	0,0095
13-14,1775	13,589	25	0,227272727	0,790909091	3,0884	41,9668	0,0956	0,0716
14,1775-15,355	14,766	10	0,090909091	0,881818182	1,3424	19,8220	0,6503	1,2531
15,355-16,5325	15,944	8	0,072727273	0,954545455	1,1595	18,4875	2,1757	6,7540
16,5325-17,71	17,121	5	0,045454545	1	0,7782	13,3244	3,5683	15,2788
Сумма		110	1	-	12,8394	168,6255	1,1878	41,2107

Рисунок 5 – Расчет некоторых вспомогательных величин

Эмпирическая функция распределения (по накопленным частотам рисунок 5):

$$F^{¿}(x)=\begin{cases} 0\,,\,при\,x\leq 20\,,9 \\ 0\,,09\,,\,при\,20\,,9<x\leq 23\,,55 \\ 0\,,19\,,\,при\,23\,,55<x\leq 26\,,2 \\ 0\,,39\,,\,при\,26\,,2<x\leq 28\,,85 \\ 0\,,59\,,\,при\,28\,,85<x\leq 31\,,5 \\ 0\,,79\,,\,при\,31\,,5<x\leq 34\,,15 \\ 0\,,89\,,\,при\,34\,,15<x\leq 36\,,8 \\ 0\,,97\,,\,при\,36\,,8<x\leq 39\,,45 \\ 1\,,\,при\,x>42\,,1 \end{cases}$$

График эмпирической функции распределения $F^*(x)$ – рисунок 6. Более наглядно эмпирическая функция распределения описывается кумулятой (рисунок 7), которая дает представление о виде функции распределения СВ в генеральной совокупности.

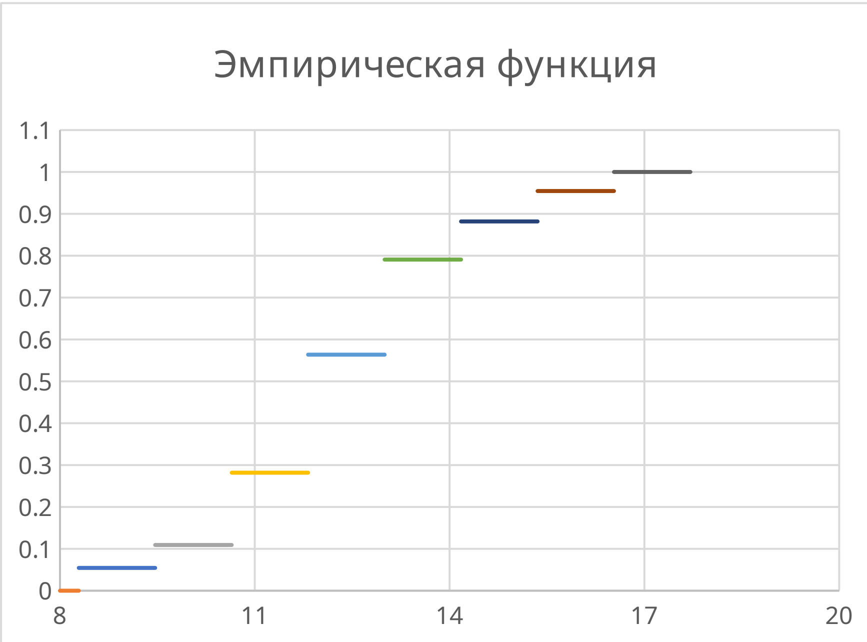

Рисунок 6 – График эмпирической функции

Рисунок 7 – Кумулята

Вычислим методом произведения точечные оценки параметров распределения, используя рисунок 5.

Выборочное среднее:

$$\overline{x_в}=\sum x_i w_i=\text{¿}30{,}4145\text{¿}$$

Выборочная дисперсия:

$$D_в=\overline{x_в^2}-\overline{(x_в)}^2=947{,}466325\text{-}(30{,}4135)^2=22{,}4853428$$

Исправленная выборочная дисперсия:

$$S_в^2=\frac{n}{n-1}D_в=\frac{100}{99}22{,}4853428=22{,}7124675$$

Выборочное среднее квадратическое отклонение:

$$\sigma_в=\sqrt{D_в}=4{,}7418$$

Мода (*Mo*) представляет собой значение изучаемого признака, повторяющееся с наибольшей частотой. Моду для интервального распределения определяем по формуле:

$$Mo=x_{Mo}+i_{Mo}\frac{f_{Mo}-f_{Mo-1}}{(f_{Mo}-f_{Mo-1})+(f_{Mo}-f_{Mo+1})}$$

где x_{Mo} – нижняя граница модального интервала, i_{Mo} – величина модального интервала, f_{Mo}, f_{Mo-1}, f_{Mo+1} – соответственно частоты модального, предмодального (предыдущего) и послемодального (следующего за модальным) интервалов.

Модальный интервал (интервал с наибольшей честотностью) – от 26,2 до 28,85. Тогда мода

$$Mo = 26,2 + 2,65 \frac{20-10}{(20-10)+(20-10)} = 27{,}525$$

Медиана (Ме) – значение признака, приходящееся на середину ранжированной совокупности. Определяется в интервальном вариационном ряду по формуле:

$$Me = x_{Me} + h \frac{0,5 * \sum f - S_{Me-1}}{f_{Me}}$$

где xMe - начальное значение интервала, содержащего медиану; h - величина медианного интервала; ∑f - сумма частот ряда; SMe-1 - сумма накопленных частот, предшествующих медианному интервалу; fMe - частота медианного интервала.

Медианным является тот интервал, накопленная частота которого равна полусумме или превышает полусумму всех частот ряда (fMe = =0,5∑f). У нас это интервал от 26,2 до 28,85. Тогда медиана:

$$Me = 26,2 + 2,65 \frac{0,5 * 100 - 39}{20} = \quad 27{,}6575$$

Коэффициент асимметрии:

$$a_s = \frac{\mu_3}{\sigma^3}$$

где в числителе – центральный момент третьего порядка (рисунок 5).

$$\mu_3 = \sum (x_i - \bar{x}_в)^3 w_i = -\ 59{,}96951656$$

Тогда коэффициент асимметрии:

$$a_S = \frac{-59{,}96951656}{4{,}7418^3} = -0{,}5624$$

Эксцесс - величина, определяемая формулой:

$$e_k = \frac{\mu_4}{\sigma_x^4} - 3$$

где в числителе – центральный момент четвертого порядка (таблица 2):

$$\mu_4 = \sum (x_i - \overline{x}_в)^4 w_i = 1310{,}929445$$

Тогда эксцесс:

$$e_k = \frac{1310{,}929445}{4{,}7418^4} - 3 = -0{,}406977$$

Доверительный интервал для математического ожидания нормально распределенного признака с надежностью γ определяется из условия:

$$p\left(\overline{x}_в - \frac{t\sigma}{\sqrt{n}} < a < \overline{x}_в + \frac{t\sigma}{\sqrt{n}}\right) = 2\Phi(t) = \gamma$$

где значение t определяется из таблиц для функции Лапласа так, чтобы выполнялось равенство $2\Phi(t) = \gamma$.

В условиях примера имеем: $\Phi(t) = \gamma/2 = 0{,}95/2 = 0{,}475$. По таблице для функции Лапласа находим аргумент: $t = 1.96$.

Имеем:

$$\Delta = \frac{t\sigma}{\sqrt{n}} = \frac{1{,}96 * 4{,}7418}{\sqrt{100}} = 0{,}9293928$$

Получаем 95%-ный доверительный интервал:

$(x_{ср}-\Delta;\ x_{ср}+\Delta) = (30{,}4135 - 0{,}9293928;\ 30{,}4135 + 0{,}9293928) = (29{,}4841072$

$;\ 31{,}3428928)$.

По итогам предыдущих пунктов целесообразно выдвинуть гипотезу H_0 – «данный признак X распределен нормально».

Для того, чтобы при заданном уровне значимости проверить гипотезу о нормальном распределении наблюдаемого признака, надо:

а) вычислить теоретические частоты;

б) вычислить наблюдаемое значение критерия χ^2 Пирсона по формуле

$$x^2_{факт} = \sum \frac{\left(n_i^{¿} - n_i\right)^2}{n_i}$$

где n_i* - эмпирические частоты, n_i - теоретические частоты;

в) по таблице критических точек распределения χ^2, по заданному уровню значимости α=0,05 и числу степеней свободы k=s-1-r , где s – число частичных интервалов, r – число параметров предполагаемого распределения, найти критическую точку $\chi^2_{кр}$ (α; k);

г) если $\chi^2_{набл}$ < $\chi^2_{кр}$, то гипотеза о нормальном распределении наблюдаемого признака принимается; если $\chi^2_{набл}$ > $\chi^2_{кр}$, то гипотеза отвергается.

Для нахождения теоретических частот составим расчетную таблицу. Удобнее всего это сделать в Excel, используя функцию НОРМРАСП (рисунок 8).

i	xi	xi+1	zi	zi+1	Φ(zi)	Φ(zi+1)	pi=Φ(zi+1)-Φ(zi)	ni=pi*n
1	8,29	9,4675	-бесконечно	-1,7275	-0,5	-0,457962177	0,042037823	4,624
2	9,4675	10,645	-1,7275	-1,1243	-0,457962177	-0,369547404	0,088414773	9,726
3	10,645	11,823	-1,1243	-0,5210	-0,369547404	-0,198815172	0,170732232	18,78
4	11,823	13	-0,5210	0,0823	-0,198815172	0,032781015	0,231596187	25,48
5	13	14,178	0,0823	0,6855	0,032781015	0,253492351	0,220711337	24,28
6	14,178	15,355	0,6855	1,2888	0,253492351	0,401262576	0,147770224	16,25
7	15,355	16,533	1,2888	1,8920	0,401262576	0,470757014	0,069494438	7,644
8	16,533	17,71	1,8920	+бесконечно	0,470757014	0,5	0,029242986	3,217
						Сумма:	1	110

Рисунок 8 – Таблица для расчета теоретических частот

Для нахождения наблюдаемого значения критерия $x^2_{набл}$

$$x^2_{набл} = \sum \frac{\left(n_i^{¿} - n_i\right)^2}{n_i}$$

	ni*	ni	ni*-ni	(ni*-ni)^2	(ni*-ni)^2/ni
	12	14,35	-2,35	5,52	0,385
	19	18,78	0,22	0,05	0,003
	31	25,48	5,52	30,52	1,198
	25	24,28	0,72	0,52	0,021
	10	16,25	-6,25	39,12	2,407
	13	10,86	2,14	4,57	0,421
Сумма:	110	110			4,435

Рисунок 9 – Расчет наблюдаемого значения критерия хи-квадрат (малые частоты <5 объединены).

Так как уровень значимости α=0,05, число частичных интервалов (после объединения) равно s=7, число параметров нормального распределения r=2, то число степеней свободы распределения χ^2 равно k=7-1-2=4, тогда по таблице критических точек находим критическую точку распределения χ^2, она равна $\chi^2_{крит}$ = 7,8. Поскольку $\chi^2_{набл}<\chi^2_{крит}$ (9,165<9,5), то гипотеза H_0 о нормальном распределении признака X принимается.

Заключение

В ходе работы были оценены свойства случайной величины (СВ) в генеральной совокупности по эмпирическим данным (выборке) путем построения эмпирического распределения, нахождения числовых характеристик СВ по выборке, нахождения точечных и интервальных оценок параметров распределения СВ в генеральной совокупности, проверки гипотезы о нормальном распределении СВ в генеральной совокупности.