

Ба 14.402

Белорусская Государственная Академия
Сельского Хозяйства

Проф. В. И. ПЕРЕХОД

КОРРЕЛЯЦИЯ (соотношение) МЕЖДУ ЭКОНОМИЧЕСКИМИ ФАКТОРАМИ ЛЕСНОГО ХОЗЯЙСТВА

(Опыт применения вариационно-статистического метода)

Из „Записок Белорусской Государственной Академии
Сельского Хозяйства“, т. IV

Инв. 1083 - БА 14402

Бел. изд-ел
1994 г.

Горки, БССР
Типография Академии
1 9 2 7



II.

Корреляция (соотношение) между экономическими факторами лесного хозяйства.

[Опыт применения вариационно-статистического метода].

1. Предварительные замечания.

Применение вариационно-статистического метода к изучению лесоводственных вопросов в СССР было начато проф. Г. Р. Эйтингеном, который в 1918 году опубликовал¹⁾ свою работу: „Влияние густоты древостоя на рост насаждения“. В этой работе сделаны вычисления следующих вариационно-статистических элементов: 1) средн. арифм. (M), 2) основного отклонения (σ), 3) коэффициента изменчивости (v), 4) коэффициента корреляции (r) и 5) коэффициента регрессии (R). Тот-же метод частично применен проф. Эйтингеном и в другой его работе: „Der Wuchs der Eiche in abhängigkeit von dem Gewicht der Eicheln“, опубликованной за-границей²⁾.

В лесной таксации применение метода вариационной статистики было осуществлено:

1) проф. Д. И. Морохиным³⁾ в работе: „Об'ем 9-ти аршинных еловых бревен по действительному измерению и по разным таблицам“;

2) И. А. Кищенко⁴⁾ в статье: „Опыт применения статистического метода к изучению строения древесно-растительных сообществ“ и

3) В. В. Матренинским в его⁵⁾ статье: „К анализу таблиц о строении насаждений“.

В лесной экономии и статистике названный метод пока не применялся, хотя, казалось-бы, здесь он и должен был укорениться. Объясняется это тем, что у нас до сих пор не была выделена лесная статистика, как самостоятельная дисциплина; что-же касается лесной экономики, долгое время пребывавшей под прессом лесоустройства, то она еще и сейчас некоторыми лесоводами рассматривается как составная часть последнего.

Техник-лесоустроитель, конечно, не мог способствовать выяснению чисто-теоретических вопросов, к каковым, несомненно, относится изучение взаимозависимостей различного рода факторов и явлений статистико-экономического характера.

Теме о „взаимодействии факторов лесного хозяйства“ мы, в свое время⁶⁾, уделяли уже некоторое внимание, но без сопоставления отдельных признаков, их характеризующих.

¹⁾ „Лесной Журнал“. 1918 г. Вып. 6—8;

²⁾ „Forstwissenschaftliches Zentralblatt“. 1926;

³⁾ „Сборник научных статей Казанского Института Сельского Хозяйства и Лесоводства“. 1925 г.

⁴⁾ и ⁵⁾ „Лесоведение и Лесоводство“. Вып. I. 1926 г.

⁶⁾ См. нашу статью „Взаимодействие факторов лесного хозяйства“, помещенную в журнале „Народное Хозяйство Белоруссии“. № 2 1922 г.

В настоящее время, на материале Белоруссии, мы имеем в виду выяснить, насколько это позволяют нам чисто-объективные условия, то соотношение, которое существует между некоторыми экономическими факторами лесного хозяйства, могущих быть выраженными с помощью измерителей.

Из числа этих факторов мы остановились на следующих:

- 1) плотность населения, выраженная количеством жителей, приходящимся на единицу поверхности;
- 2) лесистость, выраженная в процентах;
- 3) степень обеспеченности лесом (на 1 жителя);
- 4) коэффициент доходности (на 1 жителя);
- 5) потребление древесины, как топлива (на 1 хозяйство).

Эти факторы, по прежним округам БССР, выражаются такими цифровыми данными:

НАИМЕНОВАНИЕ ОКРУГОВ.	Экономические факторы лесного хозяйства:				
	1) Плотность населения:	2) Лесистость в %/о	3) Степень обеспечения лесом	4) Коэффициент доходности	5) Потребление древесины:
1. Бобруйский	34,58	28,8	0,92	2,13	2,30
2. Борисовский	31,15	34,7	1,17	2,23	2,26
3. Витебский	49,07	17,6	0,42	1,03	1,90
4. Калининский	47,92	18,3	0,41	0,53	2,03
5. Минский	45,78	22,9	0,65	1,23	1,81
6. Могилевский	43,89	23,2	0,59	1,10	1,77
7. Мозырский	19,62	34,9	1,95	3,30	3,20
8. Оршанский	53,37	19,7	0,39	0,73	1,87
9. Полоцкий	33,85	27,2	0,89	1,65	2,03
10. Слуцкий	32,45	23,2	0,71	0,97	1,60

ПРИМЕЧАНИЕ: Работа делалась до официального присоединения Гомельщины, поэтому новых округов нет.

Обычно интересуются лишь варьированием отдельных признаков или факторов (напр., лесистости, по округам), каждого порознь без сопоставления их с другими. Если же это сопоставление и делается, как, напр., в нашей работе¹⁾ „Плановость и рационализация белорусского лесного хозяйства“, то применяемый „метод сопутствующих изменений“ не дает полной уверенности в существовании определенной связи явлений, или, как теперь принято выражаться, сопряженности.

Корреляция (дословно „соотношение“) и есть, именно, одна из форм взаимозависимости переменных величин.

По словам Кьюве, каждый организм „представляет собою нечто целое, единую и замкнутую систему, части которой взаимно соответствуют; ни одна из этих частей не может изменяться без того, чтобы не изменились другие“...

¹⁾ См. „Савецкае будаўніцтва“ № 8—9 и № 10 за 1926 г.

Лесное хозяйство есть организм, части которого „представляют собою нечто целое“. И мы думаем, что изменение одного фактора должно повлечь за собой изменение другого.

Поэтому изучение коррелятивных или „соотносительных“ изменений факторов лесного хозяйства, безусловно, составляет важнейшую и необходимую задачу современного лесоведа-экономиста.

Разумеется, при изучении коррелятивной изменчивости факторов лесного хозяйства, вопрос о причинной связи не может ставиться. Важно лишь установить: варьируют ли отдельные факторы лесного хозяйства независимо или независимо друг от друга.

Коэффициент корреляции (r) дает лишь ответ на вопрос о том, связаны ли друг с другом вариационные ряды и, если да, то в какой степени.

За основу измерения степени корреляции, как известно, берется сумма произведений ($\sum a_x a_y$), попарно связанных отклонений от „средней величины“ (M).

Известно, что корреляция будет тем выше, чем в большем числе случаев и чем более совершенно будет проявляться взаимозависимость между вариантами двух совокупностей, в отдельных случаях их сочетания.

Поэтому, для выявления корреляции между экономическими факторами лесного хозяйства БССР, мы связывали их попарно, определяя для каждого сочетания, так называемый, „коэффициент корреляции“ (r) по формуле Bravais:

$$r = \frac{\sum a_x a_y}{n \cdot \sigma_x \cdot \sigma_y} \quad (1).$$

В знаменатель этой формулы входит, так наз., „основное отклонение“ (Standardabweichung), обозначаемое малой греческой сигмой (σ) и определяемое по формуле:

$$\sigma = \pm \sqrt{\frac{\sum p a^2}{n}} \quad (2).$$

По этим, именно, формулам и производились вычисления¹⁾, попарно связанных отклонений от среднего арифметического (Mittelwert).

Кроме того, вычислялся коэффициент регрессии, как показатель скорости изменения одного фактора, в случае изменения другого.

Наконец, вычислялась также средняя ошибка коэффициента корреляции по формуле:

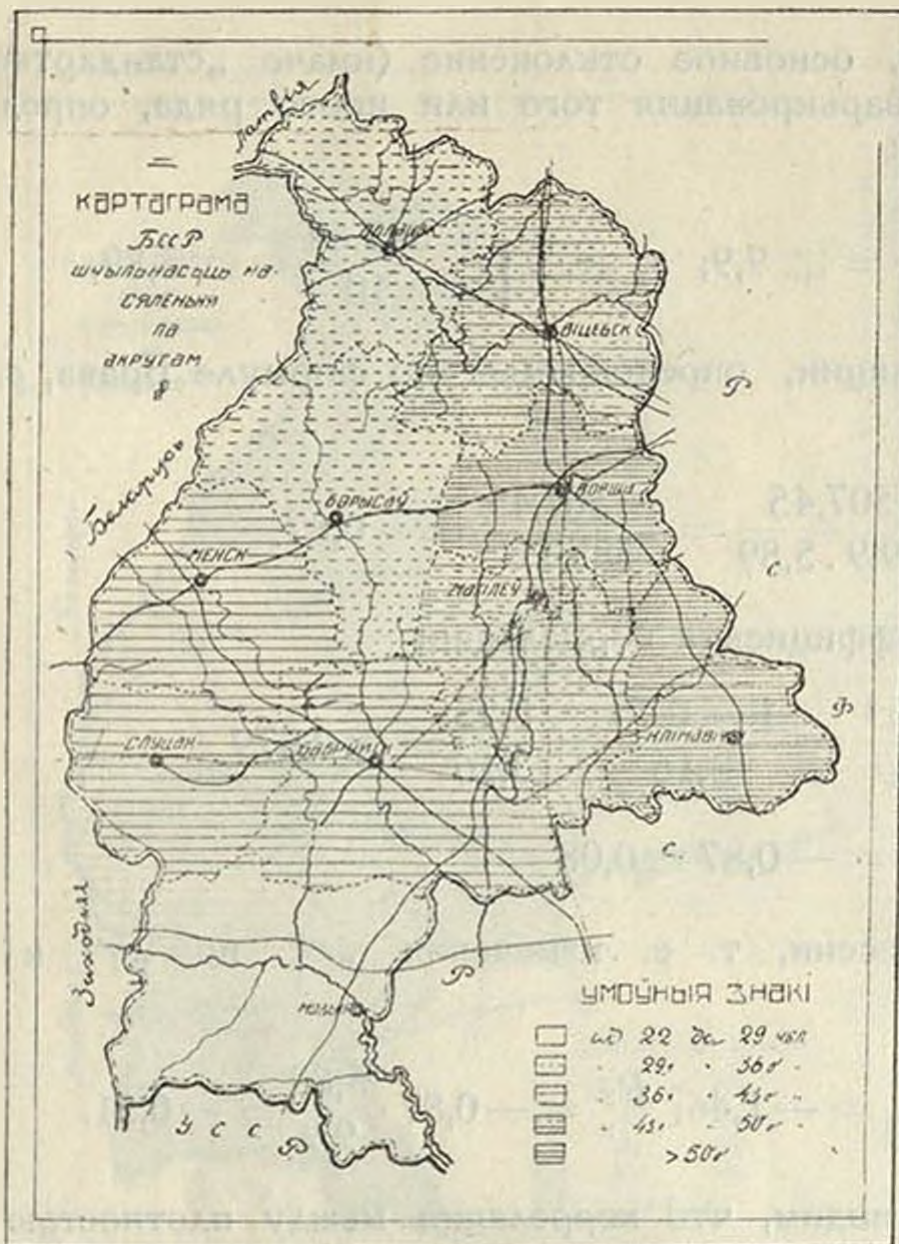
$$m_r = \frac{1 - r^2}{\sqrt{n}} \quad (3).$$

2. Корреляция между плотностью населения и лесистостью.

Соотносительные изменения (корреляция) двух означенных факторов—плотности населения и лесистости—были отмечены давно, пользуясь „методом сопутствующих изменений“. Так, если мы сопоставим, по округам БССР., данные этих двух факторов, то получим следующие ряды:

Номера округов: . . .	1	2	3	4	5	6	7	8	9	10
Плотность населения .	34,58	31,15	49,07	47,92	45,78	43,89	19,62	53,37	33,85	32,45
Лесистость в процентах	28,8	34,7	17,6	18,3	22,9	23,2	34,9	19,7	27,2	23,2

¹⁾ Настоящая работа выполнена лесозакономическим отделом Центральной Лесной Опытной Станции Белоруссии; вычисления, с помощью арифмометра, производил научный сотрудник Л. И. Блюдохо, за что считаю долгом принести ему глубокую благодарность.



1. Картограма плотнoсти населения.

Таблица 1

Номера вариант.	Абсолютные значения отд. вариант.		Отклонения от средней (M)		Произведения отклонений $D_x \cdot D_y$		Квадраты отклонений от средн. вел.	
	Плотн. (v_x)	Лесистость (v_y)	$V_x - M_x$	$V_y - M_y$	Положит. (+)	Отриц. (-)	D_x^2	D_y^2
1	34,58	28,8	- 4,59	+3,7	-	16,72	21,07	13,69
2	31,15	34,7	- 8,02	+9,6	-	76,99	64,32	92,16
3	49,07	17,6	+ 9,9	-7,5	-	64,25	98,01	54,25
4	47,92	18,3	+ 8,75	-6,8	-	59,5	76,56	46,24
5	45,78	22,9	+ 6,61	-2,2	-	14,54	43,69	4,84
6	43,89	23,2	+ 4,72	-1,9	-	8,11	14,28	3,61
7	19,62	34,9	-19,55	+9,8	-	191,59	384,16	96,04
8	53,37	19,7	+14,2	-5,4	-	76,68	201,64	29,16
9	33,85	27,2	- 5,32	+2,1	-	11,17	28,30	4,41
10	32,45	23,2	- 6,72	-1,8	12,2	-	45,16	3,34
M=	39,17	25,1	-	-	12,1	519,55	980,19	347,64
						- 507,45		

Однако, такое сопоставление дает немного; мы не можем судить о степени зависимости, хотя эта зависимость и существует.

Гораздо полнее и рельефнее зависимость выступает, в случае определения „коэффициента корреляции“.

Ниже мы приводим ход наших вычислений.

Среднее квадратное, основное отклонение (иначе „стандартное“), показывающее¹⁾ степень варьирования того или иного ряда, определяется для ряда „x“: и „y“:

$$\sigma_x = \pm \sqrt{\frac{980,19}{10}} = \pm 9,9; \quad \sigma_y = \pm \sqrt{\frac{347,64}{10}} = \pm 5,89.$$

Коэффициент корреляции, определяемый по формуле Бравэ, равен следующему:

$$r = \frac{-507,45}{10 \cdot 9,9 \cdot 5,89} = \frac{-507,45}{583,11} = -0,87.$$

Средняя ошибка коэффициента корреляции:

$$m_r = \frac{1 - r^2}{\sqrt{n}} = \frac{1 - 0,75}{3,16} = \frac{0,25}{3,16} = \pm 0,079$$

В конечном итоге, $r = -0,87 \pm 0,08$ (I)

Коэффициенты регрессии, т. е. изменения „x“ по „y“ и „y“ по „x“, равен:

$$R_{x/y} = -0,87 \cdot \frac{9,9}{5,89} = -1,46; \quad R_{y/x} = -0,87 \cdot \frac{5,89}{9,9} = -0,51.$$

Таким образом, мы видим, что корреляция между плотностью населения и лесистостью не только существует, но и достигает довольно значительного размера, ибо лишь при полной сопряженности коэффициент равен единице ($r=1,0$), у нас он все-же не так далёк от предела ($r=-0,87$).

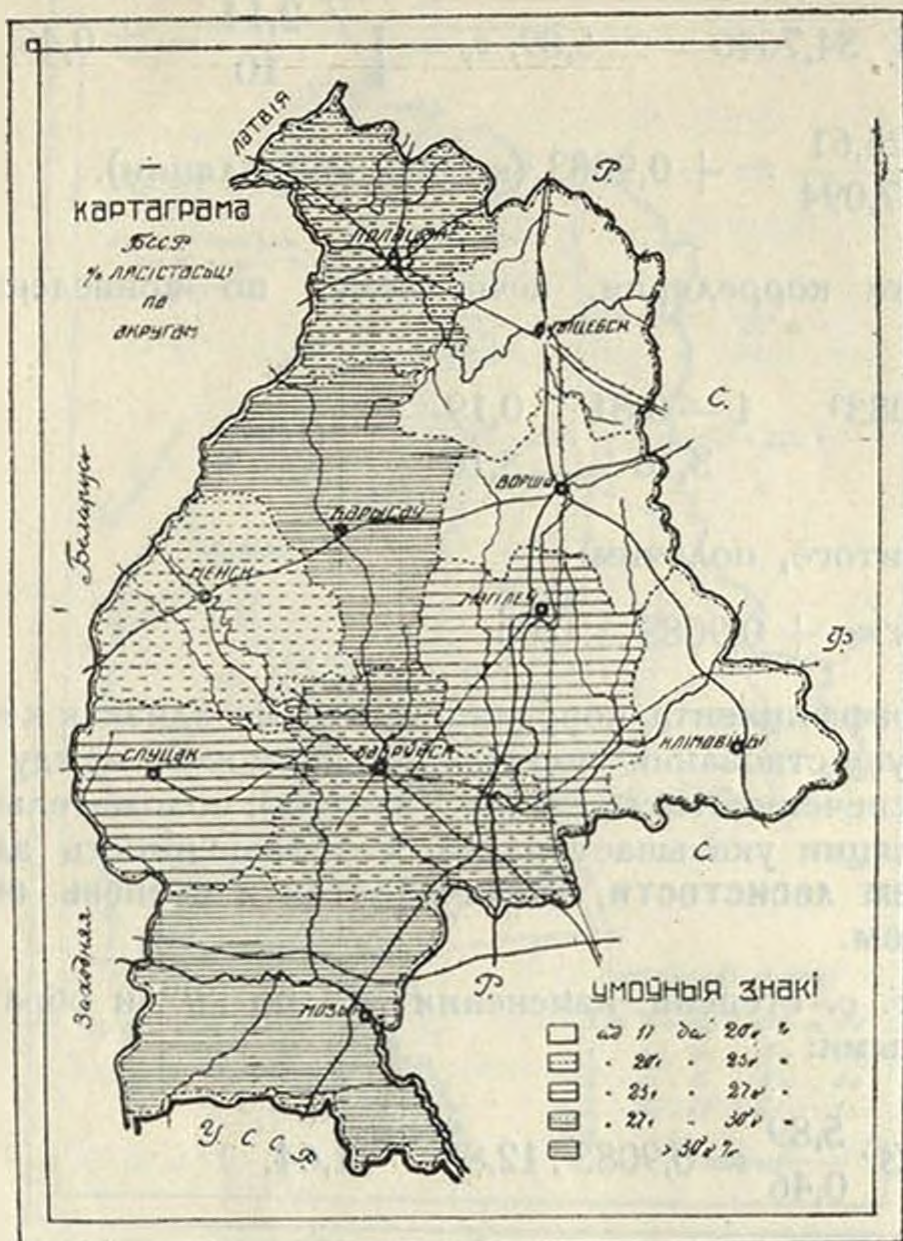
Далее, мы видим, что коэффициент корреляции у нас получился со знаком минус; это значит, что корреляция между плотностью населения и лесистостью—отрицательная, т. е. взаимозависимость здесь обратная, и с увеличением одного фактора (плотность), другой уменьшается (лесистость).

3. Корреляция между лесистостью и степенью обеспеченности населения лесом.

Если взять такую пару вариантов, как лесистость (в процентах) и степень обеспеченности лесом (на 1 жителя), то получим такие ряды:

Номера округов: . . .	1	2	3	4	5	6	7	8	9	10
Лесистость (‰) . . .	28,8	34,7	17,6	18,3	22,9	23,2	34,9	19,7	22,7	23,2
Обеспеченность лесом	9,92	1,17	0,42	0,41	0,65	0,59	1,95	0,39	0,88	0,71

¹⁾ Возможно вычисление и по другой формуле: $\sigma = \pm \sqrt{\frac{\sum pa^2}{n-1}}$, в которой знаменатель уменьшен на единицу; мы, однако, этого уменьшения в нашей работе не делаем.



По приведенным данным еще трудно с достоверностью судить о том, насколько связь между двумя названными экономическими факторами сильна. Поэтому, для определения степени сопряженности, приведем данные по вычислению коэффициента корреляции:

2. Картограмма лесистости.

Таблица 2

Номера вариантов.	Абсолютные значения вариантов		Отклонения от средней вел.		Произведения отклонений (D_x, D_y)		Квадраты отклонений от средней	
	Лесистость	Обеспечен.	D_x	D_y	+	-	D_x^2	D_y^2
1	28,8	0,92	+ 3,7	+ 0,11	0,41	—	13,69	0,12
2	34,7	1,17	+ 9,6	+ 0,36	3,45	—	92,16	0,13
3	17,6	0,42	- 7,5	- 0,39	2,93	—	54,25	0,15
4	18,3	0,41	- 6,8	- 0,40	2,72	—	46,24	0,16
5	22,9	0,65	- 2,2	- 0,16	0,35	—	4,84	0,03
6	23,2	0,59	- 1,9	- 0,22	0,42	—	3,61	0,05
7	34,9	1,95	+ 9,8	+ 1,14	11,72	—	96,04	1,30
8	19,7	0,39	- 5,4	- 0,42	2,27	—	29,16	0,18
9	22,7	0,88	+ 2,1	+ 0,07	0,15	—	4,41	0,01
10	23,2	0,71	- 1,8	- 0,10	0,18	—	3,24	0,01
	25,1	0,81	—	—	24,61	—	347,64	2,14

$$\sigma_x = \pm \sqrt{\frac{347,64}{10}} = \pm \sqrt{34,7640} = \pm 5,89; \sigma_y = \sqrt{\frac{2,14}{10}} = \pm 0,46.$$

$$r = \frac{24,61}{10 \cdot 5,89 \cdot 0,46} = \frac{24,61}{27,094} = + 0,9083 \text{ (коэфф. корреляции).}$$

Ошибка коэффициента корреляции, исчисляемая по приведенной формуле, равна:

$$m_r = \frac{1 - 0,9083^2}{3,16} = \frac{1 - 0,81}{3,16} = \frac{0,19}{3,16} = \pm 0,06$$

Отсюда, в конечном итоге, получим:

$$r = + 0,9083 \pm 0,06 \quad (II)$$

Полученная цифра коэффициента корреляции, весьма близкая к единице, свидетельствует о существовании прочной зависимости между лесистостью и степенью обеспеченности лесом (на 1 жителя); положительный знак коэффициента корреляции указывает на то, что зависимость здесь прямая, т. е. с увеличением лесистости, увеличивается и степень обеспеченности населения лесом.

Быстрота реакции, т. е. степень изменения „x“ по „y“ и обратно, определяется такими данными:

$$R_x = 0,9083 \cdot \frac{5,89}{0,46} = 0,9083 \cdot 12,82 = 11,64.$$

$$R_y = 0,9083 \cdot \frac{0,46}{5,89} = 0,9083 \cdot 0,07809 = 0,0709.$$

Мы не хотим здесь сказать, что приведенная взаимозависимость между лесистостью и степенью обеспеченности лесом населения не была ранее известна.

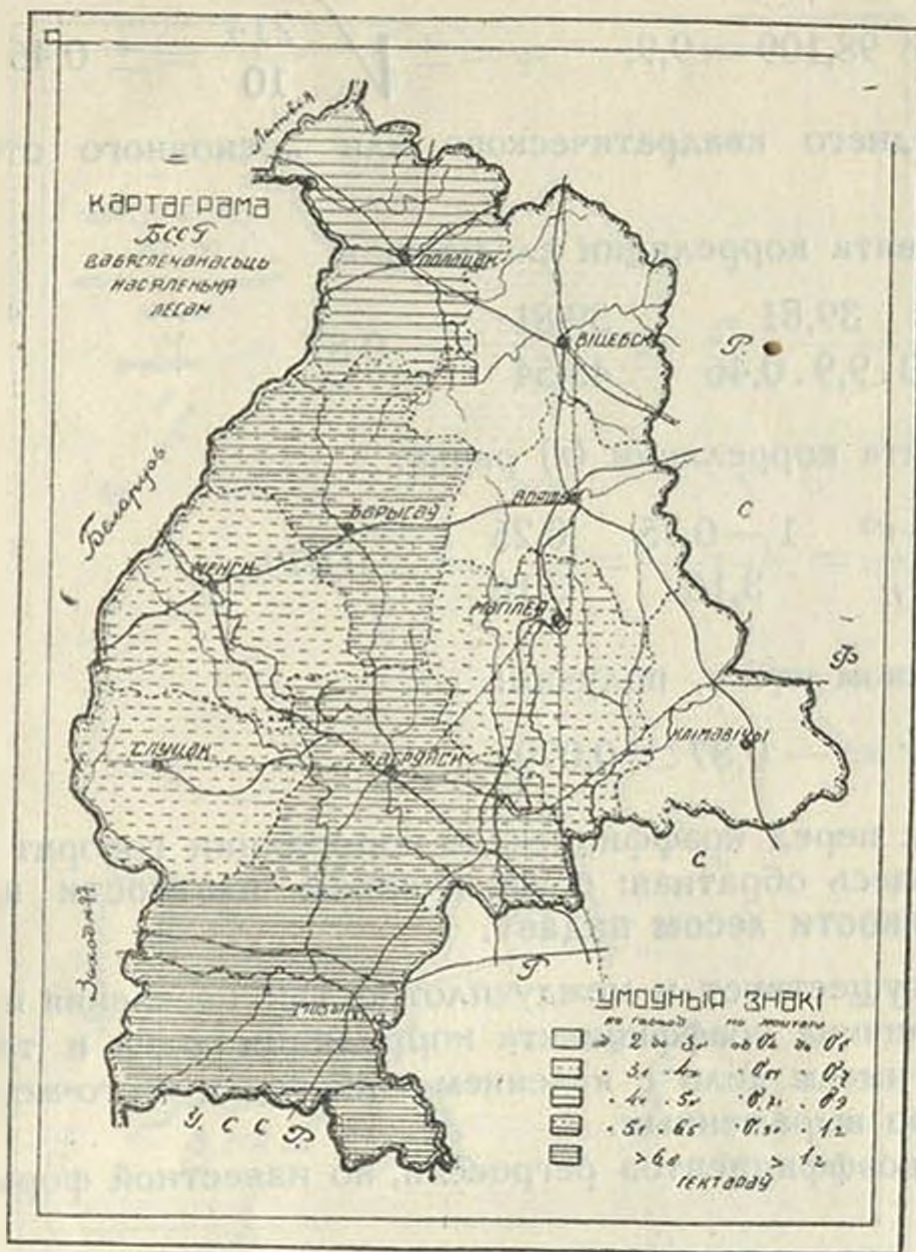
Мы имеем в виду лишь доказать прочность установленной сопряженности двух экономических факторов, с одной старой, а с другой — дать коэффициент указанного соотношения, т. е. численно выразить самый измеритель данного явления.

4. Корреляция между степенью обеспеченности лесом и плотностью населения (по БССР).

Теперь остановимся на выяснении того соотношения (корреляции), которое существует между степенью обеспеченности лесом и плотностью населения, по округам БССР.

Цифровые данные, характеризующие это явление, выражаются следующим образом:

Номера округов	1	2	3	4	5	6	7	8	9	10
Плотность населения	34,58	31,15	49,07	47,92	45,78	43,89	19,62	53,57	33,85	32,42
Обеспеченность лесом: (на 1 жителя)	0,92	1,17	0,42	0,41	0,65	0,59	1,95	0,39	0,88	0,71



Здесь мы имеем, повидимому, обратную зависимость, которую и проверяем, путем вычисления коэффициента корреляции.

3. Картограмма обеспеченности населения лесом.

Таблица 3

Номера округов.	Плотность населения	Обеспеч. лесом	Простые отклонения от средней		Произведения $D_x \cdot D_y$		Квадраты отклонений	
	x		y	D_x	D_y	—	—	D_x^2
1	34,58	0,92	- 4,59	+ 0,11	—	0,51	21,07	0,12
2	31,15	1,17	- 8,02	+ 0,36	—	2,89	64,32	0,13
3	49,07	0,42	+ 9,9	- 0,39	—	3,86	98,01	0,15
4	47,92	0,41	+ 8,75	- 0,40	—	3,5	76,56	0,16
5	45,78	0,65	+ 6,61	- 0,16	—	1,06	43,69	0,03
6	43,89	0,59	+ 4,72	- 0,22	—	1,04	14,28	0,05
7	19,62	1,95	-19,55	- 1,14	—	21,29	384,16	1,30
8	53,37	0,39	+14,2	- 0,42	—	5,96	201,64	0,18
9	33,85	0,88	- 5,32	+ 0,07	—	0,37	28,30	0,01
10	32,42	0,71	- 6,72	- 0,10	0,67	—	46,16	0,01
	39,17	0,81	—	—	0,67	40,48	980,19	2,14

$$\sigma_x = \pm \sqrt{\frac{980,19}{10}} = \pm \sqrt{98,109} = 9,9; \quad \sigma_y = \pm \sqrt{\frac{214}{10}} = \pm 0,46$$

Таковы данные среднего квадратического или „основного отклонения“ (σ).

Величина коэффициента корреляции следующая:

$$r = \frac{39,81}{10 \cdot 9,9 \cdot 0,46} = \frac{39,81}{45,54} = -0,87$$

Ошибка коэффициента корреляции (m_r) равна:

$$m_r = \frac{1 - r^2}{\sqrt{n}} = \frac{1 - 0,75}{3,16} = \frac{0,25}{3,16} = \pm 0,079$$

Стало-быть, в конечном итоге, получим:

$$r = -0,87 \pm 0,079 \quad (III)$$

Отрицательный знак перед коэффициентом корреляции говорит нам о том, что зависимость здесь обратная: с увеличением плотности населения—степень обеспеченности лесом падает.

Та-же зависимость существует и между плотностью населения и лесистостью; численная величина коэффициента корреляции одна и та-же. Это показывает, что мы имеем дело с явлением одного и того-же порядка и притом одинаково выраженным.

Определяя размер коэффициентов регрессии, по известной формуле получим:

$$R_{y/x}^x = -0,87 \cdot \frac{9,9}{0,46} = -0,87 \cdot 21,52 = -18,72.$$

$$R_{x/y}^y = -0,87 \cdot \frac{0,46}{9,9} = -0,87 \cdot 0,046 = -0,04.$$

По этим формулам, квадратические уклонения взяты в их абсолютном виде.

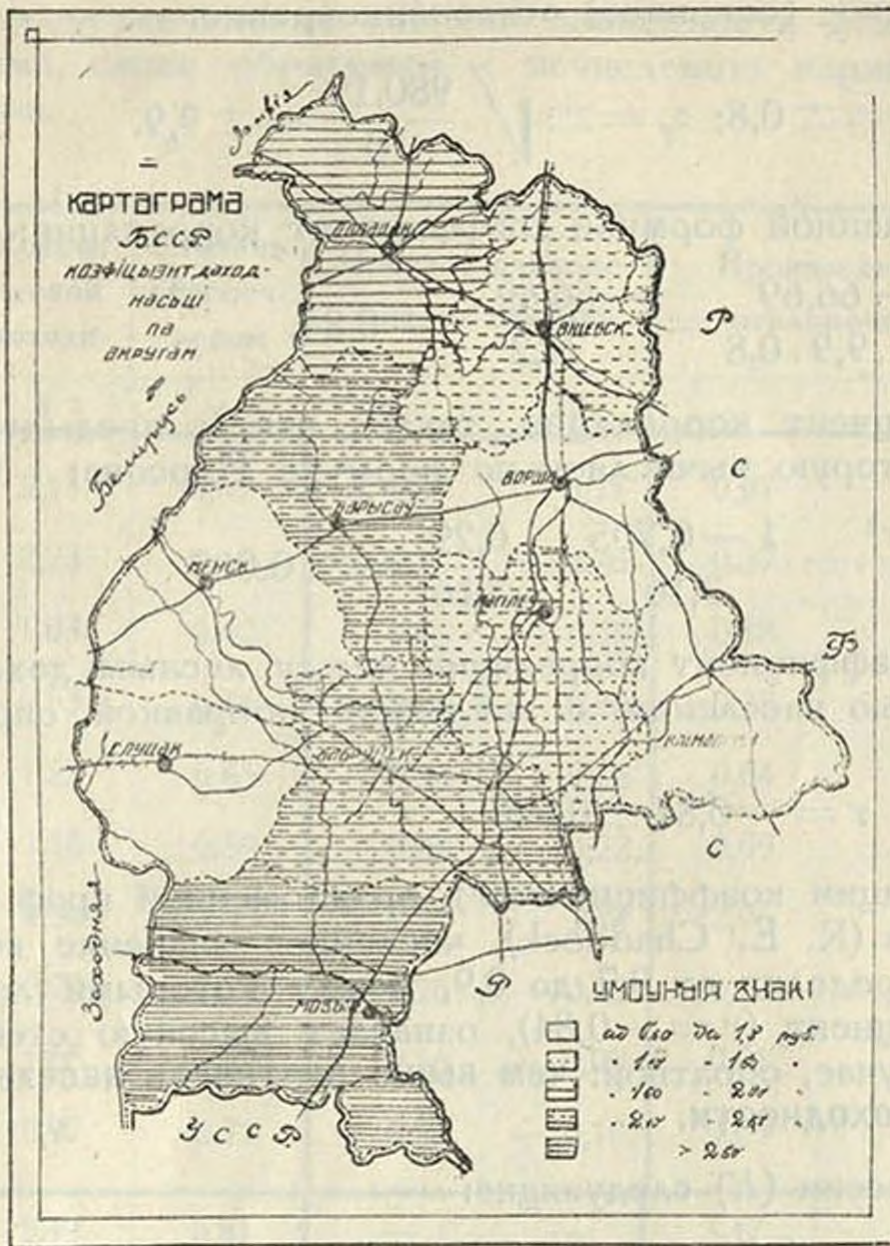
Различные числовые данные коэффициентов регрессии свидетельствуют о неодинаковом соотношении изменения одного свойства (плотности) к изменению другого (степень обеспечения лесом).

5. Корреляция между лесным доходом (на 1 жит.) и плотностью населения.

Мы хотим остановиться на том соотношении, которое существует между размером лесного дохода, приходящегося на одного жителя („коэффициентом доходности“) и плотностью населения.

Эта зависимость может быть обнаружена из сопоставления следующих рядов (по округам):

Номера округов	1	2	3	4	5	6	7	8	9	10
Кoeffиц. дохода	2,13	2,23	1,03	0,53	1,23	1,10	3,30	0,73	1,64	0,97
Плотность населения	34,58	31,15	49,07	47,92	45,78	43,89	19,62	53,37	33,85	32,42



Однако, для числового выражения существующей зависимости, приходится прибегнуть к вычислению тех-же вариационно-статистических элементов.

4. Картограмма коэффиц. доходности

Таблица 4

№№ округов.	Коэффиц. доходн. x	Плотность населения y	Dx	Dy	Dx.Dy		Dx ²	Dy ²
					+	-		
1	2,13	34,58	+ 0,64	- 4,59	-	2,94	0,40	21,07
2	2,23	31,15	+ 0,74	- 8,02	-	5,93	0,55	64,32
3	1,03	49,07	- 0,46	+ 9,9	-	3,55	0,21	98,01
4	0,53	47,92	- 0,96	+ 8,75	-	7,4	0,92	76,56
5	1,23	45,78	- 0,26	+ 6,61	-	1,72	0,07	43,69
6	1,10	43,89	- 0,39	+ 4,72	-	1,67	0,15	14,28
7	3,30	19,62	+ 1,81	- 19,55	-	35,39	3,28	384,16
8	0,73	53,37	- 0,76	+ 14,2	-	10,79	0,58	201,64
9	1,64	33,85	+ 0,15	- 5,32	-	0,79	0,02	28,30
10	0,97	32,42	- 0,52	- 6,72	3,49	-	0,27	45,16
	1,49	39,17	-	-	3,49	70,18	6,45	980,19
					- 66,69			

Среднее квадратическое (основное) отклонение равно:

$$\sigma_x = \pm \sqrt{\frac{6,45}{10}} = \pm 0,8; \quad \sigma_y = \pm \sqrt{\frac{980,19}{10}} = \pm 9,9.$$

Находим по установленной формуле коэффициент корреляции:

$$r = \frac{-66,69}{10 \cdot 9,9 \cdot 0,8} = \frac{-66,69}{79,2} = -0,84.$$

Полученный коэффициент корреляции, также, как и предыдущие, нуждается в поправке, которую вычисляем по формуле Пирсона:

$$m_r = \frac{1 - r^2}{\sqrt{n}} = \frac{1 - 0,705}{3,16} = \frac{0,29}{3,16} = \pm 0,009$$

В конечном итоге коэффициент корреляции между лесным доходом (на 1 жителя) и плотностью населения, с найденной поправкой, определится в следующем виде:

$$r = -0,84 \pm 0,009 \quad (\text{IV.})$$

Согласно классификации коэффициентов¹⁾, предложенной проф. Колумбийского Университета (R. E. Chaddock), численное значение коэффициента корреляции в пределах от 0,7 до 0,9, между которыми лежит найденный выше коэффициент ($r = -0,84$), означает высокую степень зависимости, в данном случае, обратной: чем выше плотность населения, тем ниже коэффициент доходности.

Коэффициенты регрессии (R) следующие:

$$R_{y/x}^x = 0,84 \cdot \frac{0,8}{9,9} = 0,84 \cdot 0,08 = -0,672.$$

$$R_{x/y}^y = 0,84 \cdot \frac{9,9}{0,8} = 0,84 \cdot 12,3 = -10,332.$$

Характер данных коэффициентов регрессии тот-же, что и в предыдущем случае.

6. Корреляция между коэффициентом доходности и степенью обеспечения населения лесом.

Попробуем теперь параллельно сопоставить два таких статистических ряда; как данные о величине коэффициентов лесной доходности и степени обеспечения населения лесом (по округам БССР):

Номера округов: . . .	1	2	3	4	5	6	7	8	9	10
Коэффиц. доходности .	2,13	2,23	1,03	0,53	1,23	1,10	3,30	0,73	1,65	0,97
Степень обеспечения населения лесом: . . .	0,92	1,17	0,42	0,41	0,65	0,59	1,95	0,39	0,89	0,71

¹⁾ Коэффициент ниже 0,3 означает слабую степень зависимости (сомнительного значения); коэффициент 0,3 — 0,5 выражает умеренную степень зависимости (если поправка к коэффициенту невелика); 0,5 — 0,7 отмечает заметную степень зависимости; 0,7 — 0,9 отмечает высокую степень зависимости и, наконец, 0,9 и более — весьма тесную зависимость.

Для установления степени зависимости между двумя названными факторами, снова обратимся к исчислению вариационно-статистических элементов:

Таблица 5

№№ окру- гов.	Коэфф. лесной доходн.	Степень обеспеч. лесом	Простое отклоне- ние от средней		Произведения отклонений		Квадраты отклонений	
	x	y	Dx	Dy	+	-	Dx ²	Dy ²
1	2,13	0,92	+ 0,64	+ 0,11	0,07	—	0,40	0,01
2	2,23	1,17	+ 0,74	+ 0,36	0,27	—	0,55	0,13
3	1,03	0,42	— 0,46	— 0,39	0,18	—	0,21	0,15
4	0,53	0,41	— 0,96	— 0,40	0,38	—	0,92	0,16
5	1,23	0,65	— 0,26	— 0,16	0,04	—	0,07	0,03
6	1,10	0,59	— 0,39	— 0,22	0,09	—	0,15	0,05
7	3,30	1,95	+ 1,81	+ 1,14	2,06	—	3,28	1,30
8	0,73	0,39	— 0,76	— 0,42	0,32	—	0,58	0,18
9	1,65	0,89	+ 0,16	+ 0,08	0,01	—	0,02	0,01
10	0,97	0,71	— 0,52	— 0,10	0,05	—	0,27	0,01
	1,49	0,81	—	—	3,47	—	6,45	2,03

Определим основное или стандартное отклонение:

$$\sigma_x = \pm \sqrt{\frac{6,45}{10}} = \pm 0,8; \quad \sigma_y = \pm \sqrt{\frac{2,03}{10}} = \pm 0,45.$$

По нахождении средних квадратических отклонений, определяем коэффициент корреляции:

$$r = \frac{3,47}{10 \cdot 0,45 \cdot 0,8} = \frac{3,47}{3,60} = +0,96$$

Средняя ошибка коэффициента корреляции следующая:

$$m_r = \frac{1 - r^2}{\sqrt{n}} = \frac{1 - 0,96^2}{\sqrt{10}} = \frac{0,08}{3,1} = \pm 0,025$$

Придерживаясь классификации Chaddock'a, следует сказать, что высокий коэффициент корреляции (+ 0,96), при небольшой к нему поправке (± 0,025), свидетельствует о весьма тесной связи между сопоставляемыми переменными величинами, а именно: коэффициентом доходности и степенью обеспеченности лесом.

В конечном итоге, получим:

$$r = +0,96 \pm 0,025 \quad (V.)$$

Можно было-бы и не вычислять коэффициентов регрессии, но все-же мы приведем и их:

$$R_{y}^x = r \cdot \frac{\sigma_x}{\sigma_y} = 0,96 \cdot \frac{0,8}{0,45} = 0,96 \cdot 1,77 = 1,7.$$

$$R_{x}^y = r \cdot \frac{\sigma_y}{\sigma_x} = 0,96 \cdot \frac{0,45}{0,8} = 0,96 \cdot 0,562 = 0,54.$$

Эти величины, выражающие собою средние изменения одного свойства (зависимого), т. е. в данном случае — коэффициента доходности (от леса на одного жителя), соответственно определенному изменению другого (основного): степени обеспеченности населения лесом.

Найденные численные величины коэффициентов регрессии показывают, что при увеличении степени обеспеченности лесом населения — коэффициент доходности возрастает (на 1,7 р. на 1 гектар).

7. Корреляция между коэффициентом доходности и лесистостью.

По существу, данная зависимость близка к предыдущей, так как „степень обеспеченности лесом“ и „лесистость“ — суть две величины аналогичного порядка.

При параллельном сопоставлении двух вариационных рядов, характеризующих названные факторы, получается следующая картина (по округам):

Номера округов: . . .	1	2	3	4	5	6	7	8	9	10
Коэффиц. доходн. . .	2,13	2,23	1,02	0,53	1,23	1,10	3,30	0,73	1,64	0,97
Лесистость (%) . . .	28,8	34,7	17,6	18,3	22,9	23,2	34,9	19,7	27,2	23,3

Таблица вычислений, для нахождения коэффициента корреляции, приводится ниже:

Таблица 6

№№ округов.	Коэфф. доходн.	Лесистость	Простые отклонения от средней		Dx.Dy		Квадраты отклонений	
	x	y	Dx	Dy	+	—	Dx ²	Dy ²
1	2,13	28,8	+ 0,64	+3,7	2,30	—	0,40	13,69
2	2,23	34,7	+ 0,74	+9,6	7,04	—	0,55	92,16
3	1,03	17,6	— 0,46	—7,5	3,45	—	0,21	54,25
4	0,53	18,3	— 0,96	—6,8	6,53	—	0,92	46,24
5	1,23	22,9	— 0,26	—2,2	0,57	—	0,07	4,84
6	1,10	23,2	— 0,39	—1,9	0,74	—	0,15	3,61
7	3,30	34,9	+ 1,81	+9,8	17,74	—	3,28	96,04
8	0,73	19,7	— 0,76	—5,4	3,10	—	0,58	29,16
9	1,64	27,2	+ 0,16	+2,1	0,32	—	0,02	4,41
10	0,97	23,2	— 0,52	—1,8	0,94	—	0,27	3,24
	1,49	25,1	—	—	42,73	—	6,45	374,64

Стандартное или основное отклонение равно:

$$\sigma_x = \pm \sqrt{\frac{6,45}{10}} = \pm 0,8; \quad \sigma_y = \pm \sqrt{\frac{347,64}{10}} = \pm 5,89.$$

Имея нужные данные, вычисляем искомый коэффициент корреляции:

$$r = \frac{42,73}{10 \cdot 5,89 \cdot 0,8} = \frac{42,73}{47,12} = +0,906.$$

Поправка, вычисляемая на основании формулы Пирсона, выражающая среднюю ошибку коэффициента корреляции, равна:

$$m_r = \frac{1 - r^2}{\sqrt{n}} = \frac{1 - 0,81}{3,16} = \frac{0,19}{3,16} = \pm 0,006.$$

Отсюда, в конечном итоге, коэффициент корреляции, с найденной к нему поправкой, равен:

$$r = +0,906 \pm 0,006 \quad (\text{VI.})$$

Найденный коэффициент говорит о весьма тесной зависимости между сопоставляемыми переменными величинами и притом зависимости прямой: чем выше лесистость, тем больше коэффициент доходности на одного жителя.

Коэффициенты регрессии названных величин следующие:

$$R_{y}^x = 0,91 \cdot \frac{0,8}{5,89} = 0,91 \cdot 0,14 = 0,13.$$

$$R_x^y = 0,91 \cdot \frac{5,89}{0,8} = 0,91 \cdot 0,73 = 0,66.$$

Эти данные показывают, что при увеличении лесистости на 1% происходит повышение коэффициента доходности, в среднем, на 0,13 р.

Таким образом, коэффициент регрессии дает нам представление о соотношении изменения одного свойства, в связи с изменением другого.¹⁾

8. Корреляция между потреблением древесины и лесистостью.

Общеизвестен факт зависимости потребления древесины от лесистости, указываемый хотя-бы в нашем курсе лесной экономики.²⁾

В нашем распоряжении были данные потребления древесины по округам Белоруссии, которые мы и приводим ниже (эти данные освещают только потребление древесины, как топлива) в сопоставлении с лесистостью:

Номера округов: . . .	1	2	3	4	5	6	7	8	9	10
Потребление древесины:	2,30	2,26	1,90	2,03	1,81	1,77	3,20	1,87	2,03	1,60
Лесистость (‰): . . .	28,8	34,7	17,6	18,3	22,9	23,2	34,9	19,7	27,2	23,3

¹⁾ Коэффициент регрессии, также, как и коэффициенты корреляции, как известно, имеют свою „среднюю ошибку“, которую мы, однако, не вычисляем.

Вычисление этой ошибки производится по формуле: $mR_x^y = \pm m_r \cdot \frac{\sigma_y}{\sigma_x}$

²⁾ См. „Теорию лесного хозяйства“. 2-ое изд. 1925 г. Минск. Госиздат Белоруссии.

При подобного рода сопоставлении, зависимость хотя и выявляется, но недостаточно; мы предпочитаем коррелирование этих признаков вычислять по методу вариационной статистики:

Таблица 7

№№ округов.	1-й вариант (потреблен.)	2-й вариант (лесист.)	Простое среднее отклон. ряда „x“	Простое среднее отклон. ряда „y“	Произв. отклон. Dx . Dy		Квадраты отклонений от средн.	
	Vx	Vy	Dx	Dy	Полож. (+)	Отриц. (-)	Dx ²	Dy ²
1	2,30	28,8	+ 0,22	+3,7	0,81	—	0,05	13,69
2	2,26	34,7	+ 0,18	+9,6	1,73	—	0,03	92,16
3	1,90	17,6	— 0,18	—7,5	1,35	—	0,03	54,25
4	2,03	18,3	— 0,05	—6,8	0,34	—	0,003	46,24
5	1,81	22,9	— 0,27	—2,2	0,59	—	0,07	4,84
6	1,77	23,2	— 0,31	—1,9	0,59	—	0,10	3,61
7	3,20	34,9	+ 1,12	+9,8	10,98	—	1,25	96,04
8	1,87	19,7	— 0,21	—5,4	1,13	—	0,04	29,16
9	2,03	27,2	— 0,05	+2,1	—	+ 0,11	0,003	4,41
10	1,60	23,2	— 0,48	—1,8	0,86	—	0,23	32,4
	2,08	25,1	—	—	18,38	0,11	1,81	347,64
					+ 18,27			

$$\sigma_x = + \sqrt{\frac{1,81}{10}} = \pm \sqrt{0,181} = \pm 0,42; \quad \sigma_y = \pm \sqrt{\frac{347,64}{10}} = \pm 5,89.$$

Коэффициент корреляции для изучаемых факторов равен:

$$r = \frac{18,27}{10 \cdot 0,42 \cdot 5,89} = \frac{18,27}{24,738} = + 0,73.$$

Средняя ошибка коэффициента корреляции равна:

$$m_r = \frac{1 - r^2}{\sqrt{n}} = \frac{1 - 0,73^2}{3,16} = \frac{0,47}{3,16} = \pm 0,15.$$

Следовательно, в конечном итоге, искомый коэффициент с поправкой равен:

$$r = + 0,73 \pm 0,15 \quad (\text{VII}).$$

Полученный размер коэффициента корреляции все-же заставляет признать существующую довольно высокую степень зависимости (прямой) между изучаемыми факторами, и чем выше лесистость, тем больше потребление древесины.

Соотношение изменения одного свойства (лесистости) к изменению другого (потребление), выражаемое коэффициентом регрессии, приводится ниже:

$$R_{y}^x = + 0,73 \cdot \frac{0,42}{5,89} = 0,73 \cdot 0,071 = 0,5.$$

$$R_{x}^y = + 0,73 \cdot \frac{5,89}{0,42} = 0,73 \cdot 1,4 = 1,02.$$

Так как корреляция, а, значит, и регрессия—в данном случае—имеют положительное значение (+), то приращение одного из указанных свойств вызывает отмеченное приведенными данными—увеличение другого.

Так, напр., при наших данных, увеличение лесистости на один процент вызывает рост потребления древесины, как топливо на 0,5.

9. Заключение и выводы.

В заключение, мы хотели-бы сказать здесь следующее.

Метод сопутствующих изменений, которым мы обычно пользуемся при статистико-экономических исследованиях того или иного вопроса лесного хозяйства, хотя и открывает некоторую возможность к обнаружению зависимости между отдельными фактами и даже позволяет установить направление этой зависимости (прямое или обратное), но все-же он не достаточно полон, так как не дает количественного (цифрового) измерения обнаруживаемой зависимости. В этом отношении метод корреляции стоит несравненно выше, ибо он не только позволяет констатировать—наличие или отсутствие зависимости между теми или иными факторами, определяя в то же время положительное или отрицательное направление этой зависимости, но и дает возможность—количественно выразить степень связности изучаемых явлений.

Поэтому метод корреляции в экономике и статистике лесного хозяйства не только займет подобающее место, но и обеспечит научное бытие этих дисциплин. так как простое собиране фактов, характеризующее собою первый (описательный) период развития каждой науки, не дает права на признание ее среди своих сверстниц.

Мы продолжаем утверждать, что „лесная экономика стремится овладеть теми соотношениями или взаимной связью, которая существует между лесохозяйственными явлениями“ (см. введение к нашему курсу „Теории лесного хозяйства“). И с помощью метода корреляции лесная экономика—существующими взаимозависимостями между фактами—овладеет.

В частности, настоящая работа, не претендующая, конечно, на характер открытий законов или зависимостей, более или менее известных и до настоящего времени, тем не менее, позволяет сделать следующие выводы:

1) корреляция между плотностью населения и лесистостью—отрицательная, т. е. взаимозависимость здесь обратная, и с увеличением одного фактора (плотность), другой уменьшается (лесистость);

2) направление зависимости между лесистостью и степенью обеспеченности—положительное, т. е. с увеличением лесистости, увеличивается и степень обеспеченности лесом;

3) с увеличением плотности населения, степень обеспеченности лесом падает (корреляция отрицательная);

4) чем выше плотность населения, тем ниже коэффициент лесной доходности на одного жителя (зависимость обратная);

5) при увеличении степени обеспеченности лесом населения, коэффициент доходности возрастает (прямая зависимость);

6) чем выше лесистость, тем больше коэффициент лесной доходности (положительная корреляция); и,

7) наконец, чем выше лесистость, тем больше потребление древесины, как топлива, конечно, при прочих равных условиях; это относится и ко всем предыдущим выводам (что само собой понятно).

Проф. В. И. Переход.

Горки, БССР.
1927.