А. О. САПІГИН
Професор генетики та селекції
Одеського С.-Г. Інституту

ВАРІЯЦІЙНА СТАТИСТИКА

ЕЛЕМЕНТАРНИЙ ПІДРУЧНИК ДЛЯ АГРОНОМІВ

КНИГОСПІЛКА

А. О. САПІГИН
Професор генетики та селекції
Одеського С.-Г. Інституту

ВАРІЯЦІЙНА СТАТИСТИКА

ЕЛЕМЕНТАРНИЙ ПІДРУЧНИК
ДЛЯ АГРОНОМІВ

Державний Науково-Методологічний Комітет Наркомосвіти УСРР вживати як підручника по вищих сіл.-госп. школах **дозволив.**

КНИГОСПІЛКА

Київськ. Окрліт № 17184.
Держтрест „Київ-Друк“,
1-ша фот.-літ.-цин.-друк.
Зам. № 845—2500.

ПЕРЕДМОВА.

Моя „Варіяційна Статистика“ українською мовою виходить після двох попередніх російських видань її, що вийшли 1922 р. з піврічною проміжкою. З того часу розуміння варіяційної статистики в агрономічних колах стало дуже поширене і застосування її є річ звичайна.

Завданням мого підручника було і є допомогти агрономові та студентові, що не беруть всілякі „тези“ та висліди з числових даних „на віру“, але хотять підійти до них критично, перевірюючи їхню статистичну надійність.

В українському виданні я взяв на увагу всі корисні вказівки критики, також збільшений власний досвід із педагогічної практики, і це видання виходить виправлене, доповнене й поширене.

Автор.

5 липня 1926 р.

I. ВСТУП.

Величини, що на них завжди доводиться агрономові спиратися у своїй роботі, за дуже малими винятками—варіюють, тоб-то змінюються в певних межах; тому їх виражають не одним якимось числом, але рядом чисел, що більш-менш відрізняються одно від одного.

Приклад. Вияснюючи для себе, як упливають на врожай різні способи обробки грунту, терміни засіву, його густота, способи загортання насіння, угноєння, відміни пару то-що, агроном звертається до числових даних досвідних установ того району, що його цікавить, і відразу-ж натрапляє на ряди, що варіюють: числа врожаїв одного й того самого досліджуваного питання не залишаються однаковими, але змінюються з року в рік, стаючи раз більшими, раз меншими; ту саму картину дають і числа різниць поміж урожаями всякої пари способів або питань, як-що простежити за цим протягом цілої низки років.

Через це доводиться вдаватися до середніх величин, до середніх аритметичних; їх виводять вивчаючи ряди чисел, що варіюють. Найближче розглядання їх, проте, виявляє, що ряди чисел, які варіюють, инакше кажучи, варіяційні ряди являються далеко не однохарактерними і далеко не рівноцінними що до надійности середніх аритметичних, які з них одержують. Так, напр., середнє, рівне 10, можна одержати із ряду 1, 4, 8, 12, 16 і 19, але також і з ряду 8,5; 9,0; 9,5; 10,5; 11,0 і 11,5. Границі (розмах) варіювання чисел першого ряду є рівні 1—19, отже 18 яких-небудь одиниць, а для другого ряду 8,5—11,5, себ-то 3 одиниці. Тут, в оцих двох рядах і характер варіювання, як бачимо, є різний і надійність середніх аритметичних є неоднакова, бо матеріял, що дав числа першого ряду, є без порівнання різноманітніший і менш надійний за матеріял, що з нього втворено другий ряд. Уявім собі, що наведені ряди чисел виражають вагу коренів двох сортів буряка (у фунтах) і нам треба схарактеризувати їх середніми вагами коренів. В обох випадках ми мали доціру в середньому 10 фун., але беручи для визначення инші групи коренів, ми дістали-б і инші середні, бо в одну групу попало-б більше дрібних коренів, в другу-ж більше середніх аритметичних. В крайніх випадках могли-б вийти такі групи: для першого сорту—всі 6 коренів по 1 ф. або всі 6 по 19 ф., а для другого сорту—всі 6 по 8,5 ф. або всі 6 по 11,5 ф. В такому випадку середні аритметичні вийшли-б: для першого сорту в 1 ф. і в 19 ф., а для другого сорту в 8,5 ф. і в 11,5 ф. Звідци видко, що на такім різноманітнім матеріялі, що так широко варіює, як у першого сорту, можна зробити більшу похибку при визначенні середнього аритметичного, ніж на матеріялі другого сорту, що мало варіює. Щоб збільшити надійність, точність середнього аритметичного, для першого сорту довелося-б виважити більшу кількість коренів, ніж для другого сорту.

Подібні міркування стосуються і до всяких инших варіяційних рядів. Через те, агрономові потрібно вміти просто і ясно схарактеризувати даний варіяційний ряд і визначити величину точности, або надійности його середнього аритметичного, а також число членів досліджуваного ряду, потрібних, щоб одержати бажаний степінь точности цього середнього.

Часто буває потрібно виразити простим способом степінь взаємної залежности або супряжености двох величин, що варіюють. Так, напр., кількість азоту в зерні дуже часто буває звязана з його величиною; врожай озимого залежить дуже в значній мірі від часу засіву; вишина врожаю в тій чи иншій мірі є супряжена з густотою травостою і т. и.

Агрономові, таким чином, конче треба вміти виражати просто і ясно степінь супряжености, або звязку поміж двома величинами, що варіюють.

Розробку поставлених вище питань та їх розвязання дає нам статистика величин, що варіюють, коротче сказати—варіяційна статистика. Над її розробленням працювало і працює багато дослідувачів, особливо—математиків, бо в багатьох своїх основних положеннях варіяційна статистика спирається на теорію ймовірностей. Найважливіші для нас є праці Кетле, Браве, Гальтона та Пірсона.

В дальших розділах читач познайомиться в елементарній формі з найважнішими способами і формулами варіяційної статистики в застосуванні до потреб агрономії.

II. Варіяційні ряди та криві.

§ 1. Перші відомості в явищах варіювання і в характері варіяційних рядів ми дістанемо, коли звернемо нашу увагу на те, яким способом визначають певні окремості, особи, величини, що варіюють. Одні з них виражають цілими числами, напр., число зернин у колосі, число пелюсток у квітці, число плугів, коней, корів або иншої худоби в господарстві, число рослин на одиниці площі то-що. Таке варіювання звуть цілим або дискретним (переривним). Инші окремості визначають відношенням двох мір, беручи одну з них за одиницю (міри довжини, ваги то-що), напр., довжина і грубина зернин або вага їх, величина врожаю з одиниці площі, довжина колосу то-що. Вартості цих відношень можуть лежати в кожнім місці поміж двома цілими числами, бо це залежить від точности вимірювання, а також від того, яку саме міру взято за одиницю. Через те тут доводиться задовольнятися більш-менш наближеними вартостями тих окремостей, що варіюють, і вдаватися до штучного розподілу одержаних чисел по групах або класах, звідки й назва такого варіювання—класове.

§ 2. Умовмося тепер що до змісту понять, які ми звязуємо з термінами: варіювання, варіянт, варіяція. Варіюванням ми звемо саме явище мінливости речей, окремостей, величин. Окремі речі, величини і т. и., що варіюють, ми зватимемо варіянтами. Варіяціями-ж зватимемо ступені, градації варіювання. Напр., явище мінливости числа зернин у різних колосках (навіть одного й того самого найчистішого сорту) ми звемо варіюванням числа зернин в колосі; окремі-ж колоски з тим чи иншим числом зернин будуть варіянтами, кожне-ж число зернин, що може повторюватися в багатьох колосках, явить собою окрему варіяцію.

§ 3. Приклади. Взято (під ряд, без вибору) 100 колосів ярої пшениці „ульки" і підраховано число зернин в кожного з цих 100 колосів. Вислід одержано такий: дане варіювання дало 6 варіяцій—в 15, 16, 17, 18, 19 і 20 зернин, при чому окремі варіяції виявлено тут такими числами (кількостями) варіянтів:

варіяцію „15 зернин“	виявлено	4	колосами,
„ 16 „	„	19	„
„ 17 „	„	34	„
„ 18 „	„	28	„
„ 19 „	„	11	„
„ 20 „	„	4	„

Це можна висловити ще так: **ч и с е л ь н і с т ь** варіяції „15 зернин“ = 4, чисельність варіяції „16 зернин“ = 19 і т. д.

На підставі цих даних можна написати такий варіяційний ряд:

варіяції	число зернин	15	16	17	18	19	20
варіянти	Чисельність відповідних варіяцій	4	19	34	28	11	4

Тут ми мали діло з цілими, дискретними варіяціями.

Візьмім тепер приклад класових варіяцій. Зважено 386 зернин „ульки“, взятих під ряд, без вибору. Виявилось, що

з вагою більшою за	10 mgr., але не вищою за	15 mgr. було	5 зернин
„ „ „ „	15 „ „ „ „ „	20 „ „	22 „
„ „ „ „	20 „ „ „ „ „	25 „ „	101 „
„ „ „ „	25 „ „ „ „ „	30 „ „	190 „
„ „ „ „	30 „ „ „ „ „	35 „ „	63 „
„ „ „ „	35 „ „ „ „ „	40 „ „	4 „
„ „ „ „	40 „ „ „ „ „	45 „ „	1 „

Ці дані можна розкласти в такий ряд:

варіяції	класи	10	15	20	25	30	35	40	45 mgr
варіянти	Чисельність відповідних клас		5	22	101	190	63	4	1

Коли порівняти цей класовий ряд з вищенаведеним рядом цілих варіяцій, то різниця між ними покажеться цілком виразно: числа варіянтів містяться тут поміж межами клас, яких величина є довільна, бо її можна взяти то ширшу, то вужчу.

Иноді буває потрібно представити цілі варіяції у вигляді класового ряду. Тоді вдаються до інтерполяції, намічаючи межі клас поміж даними цілими величинами.

Так, у нашім прикладі варіювання числа колосків в ульки можна взяти за межі клас 14,5—15,5—16,5—17,5—18,5—19,5—20,5 колосків і тоді вийде такий ряд:

число колосків	14,5	15,5	16,5	17,5	18,5	19,5	20,5
число колосів	4	19	34	28	11	4	

Другим прикладом цілих варіяцій, упорядкованих за класами, буде такий ряд:

число зернин у колосі	0	3	6	9	12	15	18	21	24	27	30
число відповід. колосів	1	6	14	27	37	77	115	79	25	6	

Цей ряд треба розуміти так: 6 колосів мало понад 3, але не більше як 6 зернин, тоб-то 4, 5 або 6 зернин, 14 колосів—7, 8 або 9 зернин і т. д.

§ 4. Розглянуті вище варіяційні ряди, цілі та класові, виражають варіювання *кількосних* ознак. Як що ми звернемося до ознак *якосних*, то вислід спостережень являтиметься звичайно в иншім вигляді: кожний даний варіянт або матиме досліджувану ознаку або не матиме її. Напр., досліджуючи купу колосів якогось звичайного сорту пшениці, ми відзначимо, що вона складається із стількох, скажім, червоних колосів і стількох нечервоних (напр., білих); або—із стількох остюватих і стількох без остюків. Таке варіювання зветься *альтернативним*.

Прикладами альтернативних варіяцій можуть служити такі ряди:

В однім племінній господарстві народилося 1914 року:

бичків	теличок	% бичк.	% телич.
92	108	46	54

При схрещуванні білокольорового жовтонасінного сорту квасолі з фіялковокольоровим чорнонасінним одержано фіялковокольоровий брудночорнонасінний мішанець; нащадки його, числом 558 особин, дали таке варіювання:

<table>
<tr><td colspan="2">з білим віночком
160 (28,7%)
з насінням</td><td colspan="2">з фіялковим віночком
398 (71,3%)
з насінням</td></tr>
<tr><td>жовтим
39 (7%)</td><td>бронзовим
121 (21,7%)</td><td>фіялковим
105 (18,8%)</td><td>чорним
293 (52,5%)</td></tr>
</table>

В останнім прикладі ми маємо 2 альтернативи, що до барви віночка, і 4 альтернативи, що до барви насіння.

§ 5. Значно наочніше, як числовими рядами, можна ілюструвати варіювання графічними схемами, так званими варіяційними кривими.

Криві цілих варіяцій роблять так: на поземній простій (на осі абсцис) відкладають відповідні цілі числа за допомогою точок, рівновіддалених одна від одної; в цих точках проводять прямовисні прості (ординати); довжини цих простих (ординат) і виражають собою знайдені кількості варіянтів, або чисельність відповідних варіяцій. Далі лишається тільки сполучити верхні кінці ординат лінією, і вийде крива цілих варіяцій. Треба, проте, завважити, що справжню криву ми одержимо, власне, тільки тоді, коли ми нарисовану, як це звичайно роблять, ламану лінію, перетворимо в криву, заокругливши відповідним способом її злами; однак робити це доводиться дуже рідко, при чому треба вживати особливих правил, щоб часом не понівечити характеру кривої.

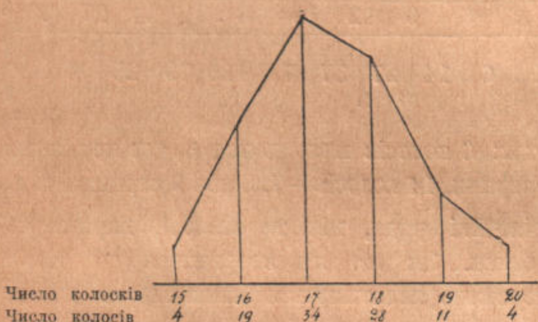


Рис. 1.

Криву наведеного вище варіювання число колосків в ульки дає нам рис. 1.

При рисуванні кривої класових варіяцій на поземній осі відкладають рівні відтинки, що виражають собою класи, і на цих відтинках, як на основах, будують прямокутники з таким розрахунком, щоб поверхня кожного з них виражала собою чисельність (число варіянтів) відповідної класи; основи прямокутників тут всі однакові, тому можна взяти, як міру, висоту прямокутника, беручи її відповідно до числа варіянтів класи: поверхні прямокутників будуть тоді пропорціональні до цих чисел. Верхні боки всіх прямокутників дадуть тоді криву, що звуть «ступеньковою» або «східчастою». Відзначмо тепер у всіх верхніх боків їх середини точками і сполучімо їх лініями: тоді одержимо відшукувану варіяційну криву, як, напр., на рис. 2.

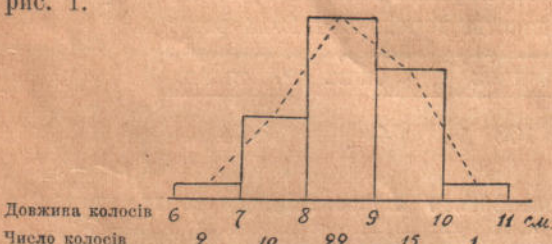


Рис. 2.

§ 6. Вивчаючи варіяційні ряди, Кетле постеріг, що вони доволі добре відповідають рядам сучинників в розкладі Ньютонового біному. Справді, розклад біному $(a+b)^n$ дає таке:

$$(a+b)^1 = a+b$$
$$(a+b)^2 = a^2+2ab+b^2$$
$$(a+b)^3 = a^3+3a^2b+3ab^2+b^3$$
$$(a+b)^4 = a^4+4a^3b+6a^2b^2+4ab^3+b^4$$
$$\vdots$$
$$(a+b)^{10} = a^{10}+10a^9b+45a^8b^2+120a^7b^3+210a^6b^4+252a^5b^5+$$
$$+210a^4b^6+120a^3b^7+45a^2b^8+10ab^9+b^{10}$$

і т. д.

Коли прийняти $a=b=1$, то з правих частин попередніх формул вийдуть ряди, схожі з варіяційними:

1,1
1,2,1
1,3,3,1
1,4,6,4,1
⋮
1,10,45,120,210,252,210,120,45,10,1
і т. д.

Коли візьмемо $n=\infty$, то вийде біноміяльна крива, як на рис. 4.

Це явище відповідности в розподілі чисельностей варіяцій до розподілу сучинників в формулі бінома зветься кетлетовим законом. Вивчання властивостей біномінальної кривої дало початок варіяційній статистиці. До цього ми ще вернемося в одному з дальших розділів, тут-же тільки те завважмо, що в таких „нормальних“ (правильніше, симетричних) кривих найбільшу чисельність має середня варіяція і що далі від цієї середини (вправо та вліво), то все менші будуть і числа варіянтів.

III. Середнє аритметичне (M).

§ 1. Найвідоміший в щоденній практиці варіяційностатистичний елемент, що характеризує варіяційний ряд, є його середнє аритметичне. Обчисляють його звичайно так, що помножають числа варіянтів на відповідні їм числові вартості варіянтів. Для прикладу візьмім знайомий уже нам ряд, що виражає варіювання числа колосків в ульки:

Число колосків.	15	16	17	18	19	20
Число колосів.	4	19	34	28	11	4

Середнє аритметичне тут дорівнює:

$$(4.15 + 19.16 + 34.17 + 28.18 + 11.19 + 4.20) : 100 = 17{,}35 \text{ колоска.}$$

Як що середнє аритметичне зазначити літерою M, число всіх варіянтів літерою n, числову вартість варіяції літерою W, а відповідне до неї число варіянтів літерою p, то для нашого прикладу матимемо:

$$\begin{array}{c|c|c|c|c|c} W_1 = 15 & W_2 = 16 & W_3 = 17 & W_4 = 18 & W_5 = 19 & W_6 = 20 \\ p_1 = \ \ 4 & p_2 = 19 & p_3 = 34 & p_4 = 28 & p_5 = 11 & p_6 = \ \ 4, \end{array}$$

звідки

$$M = (p_1 W_1 + p_2 W_2 + p_3 W_3 + p_4 W_4 + p_5 W_5 + p_6 W_6) : 100 = \\ = 17{,}35 \text{ колоска;}$$

взагалі-ж

$$M = (p_1 W_1 + p_2 W_2 + \ldots + p_k W_k) : n$$

Уживаючи знака суми, в вигляді великої грецької літери сигми (Σ), ту саму формулу для середнього аритметичного можна написати так:

$$M = \frac{\Sigma p W}{n}.$$

Середнє аритметичне являє собою ніби точку рівноваги варіяційного ряду, а через те сума додатніх відхилів від M дорівнює сумі від'ємних, с у м а-ж у с і х в і д х и л і в в і д M, т а к и м ч и н о м, д о р і в н ю є н у л е в і. Справді, в нашому прикладі $M = 17{,}35$; сума додатніх відхилів дорівнює $(18 - 17{,}35) \,.\, 28 + (19 - 17{,}35) \,.\, 11 + (20 - 17{,}35) \,.\, 4 = 46{,}95$; сума-ж від'ємних відхилів дорівнює $(17 - 17{,}35) \,.\, 34 + (16 - 17{,}35) \,.\, 19 + (15 - 17{,}35) \,.\, 4 = - 46{,}95$; сума-ж усіх відхилів $= 46{,}95 - 46{,}95 = 0$.

Як що відхил варіяції від M зазначити літерою a, число-ж відповідних варіянтів — p, то сума всіх pa дорівнюватиме нулеві, або символічно $\Sigma pa = 0$.

§ 2. Нам потрібно тепер познайомитися з простішим, ніж показаний вище, способом обчисляти середнє аритметичне, що його широко вживають у варіяційній статистиці.

Коли ми обчисляємо M звичайним способом, то беремо абсолютні числові вартості варіяцій, тоб-то беремо їх, як відповідні відхили в і д нуля. Так, в нашому прикладі, варіяції відхиляються (віддалені) від нуля, відповідно на 15, 16, 17, 18, 19 та 20. З цього погляду M є середній відхил від нуля. Але, обчислюючи відхили варіяцій, взяти не нуль, а якесь инше число, напр. 15; тоді наші варіяції будуть відхилятися відповідно на 0, 1, 2, 3, 4 та 5, а це значно спростить нам обчислення M, бо числа варіянтів доведеться помножати не на двоцифрові, а на одноцифрові числа. Для нашого прикладу ця операція матиме такий вигляд:

$$(0.4 + 1.19 + 2.34 + 3.28 + 4.11 + 5.4) : 100 = 2{,}35.$$

Таким чином, середній відхил від нашого умовного нуля (від вихідної точки) дорівнює 2,35; але сам умовний нуль дорівнює 15, отже $M = 15 + 2{,}35 = 17{,}35$.

Як що умовний нуль (вихідну точку) зазначити літерою А, а відхил варіяції від вихідної точки—літерою a, то середній відхил від умовного нуля (зазначмо його літерою b) можна виразити формулою

$$b = \frac{\Sigma pa}{n},$$

середнє-ж аритметичне

$$M = A + \frac{\Sigma pa}{n} = A + b.$$

Це обчислення середнього аритметичного можна спростити ще більше. Для цього беруть за вихідну точку звичайно варіяцію з найбільшим числом варіянтів, отже, десь біля середини варіяційного ряду. Тоді праворуч від цього A будуть додатні a, ліворуч-же—від'ємні, при чому всім або майже всім додатнім a відповідатимуть рівновеликі від'ємні. Так, як що ми візьмемо в нашім прикладі за A варіяцію 17, то праворуч матимемо відхили $(a) + 1, + 2, + 3$, а ліворуч $- 1, - 2$. При обчисленні b матимемо, таким чином, ряд:

$$+ 1.28 - 1.19 + 2.11 - 2.4 + 3.4,$$

або взагалі:

$$a_1 p_1 - a_1 p^{(1)} + a_2 p_2 - a_2 p^{(2)} + \ldots + a_k p_k - a_k p^{(k)},$$

як що знаками $p_1, p_2, \ldots, p_k$ зазначити числа варіянтів відповідних додатніх a, а знаками $p^{(1)}, p^{(2)} \ldots, p^{(k)}$ — від'ємних a.

Виносячи однакові чинники за дужки, одержимо:

$$1.(28 - 19) + 2.(11 - 4) + 3.4,$$

або в загальній формі:

$$a_1 (p_1 - p^{(1)}) + a_2 (p_2 - p^{(2)}) + \ldots + a_k (p_k - p^{(k)})$$

Цей вираз можна записати ще й так:

$$\underbrace{\overbrace{\begin{matrix}p_1\\-p^{(1)}\end{matrix}}^{a_1}} + \underbrace{\overbrace{\begin{matrix}p_2\\-p^{(2)}\end{matrix}}^{a_2}} + \cdots + \underbrace{\overbrace{\begin{matrix}p_k\\-p^{(k)}\end{matrix}}^{a_k}}\text{, або }\left|\begin{array}{c|c|c|c|c}a_1 & & a_2 & \ldots\ldots & a_k\\ \hline \begin{matrix}p_1\\-p^{(1)}\end{matrix} & + & \begin{matrix}p_2\\-p^{(2)}\end{matrix} & +\cdots+ & \begin{matrix}p_k\\-p^{(k)}\end{matrix}\\ \hline \end{array}\right|$$

Зручність такого способу записування виявиться на прикладах, даних в § 3.

Цей, викладений допіру, спосіб скорочує число потрібних перемножень і приводить до зменшення числових вартостей других чинників; крім того, беручи за нуль варіяцію з найбільшим числом варіянтів, ми тим самим виключаємо з операції множення найбільшого множача, бо $p_0 . 0 = 0$.

§ 3. Приклад. Підрахунок скрайніх квіток у 1000 головках Chrysanthemum segetum, зроблений Людвігом, дав такий варіяційний ряд:

$\pm a$	-6	-5	-4	-3	-2	-1	A	$+1$	$+2$	$+3$	$+4$	$+5$	$+6$	$+7$	$+8$
Число скрайніх квіток.	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21
Число відпов. головок.	1	6	3	25	46	141	**529**	129	47	30	15	12	8	6	2

Визначмо середнє аритметичне цього ряду. $n = 1000$; візьмім за A варіяцію 13, тоб-то $A = 13$. Обчислення ведемо так:

Відхили від A . . , . ($\pm a$)	1	2	3	4	5	6	7	8
Число варіянтів (p) при $+a$	129	47	30	15	12	8	6	2
„ „ „ „ $-a$	141	46	25	3	6	1	0	0
$a_1(p_1-p^{(1)})+a_2(p_2-p^{(2)})+\cdots+a_k(p_k-p^{(k)})$	$-12.1+1.2+5.3+12.4+$ $+6.5+7.6+6.7+2.8=$							

$= \Sigma pa = -12 + 2 + 15 + 48 + 30 + 42 + 42 + 16 = -12 + 195 =$ $= +183; b = \frac{\Sigma pa}{n} = \frac{+183}{1000} = +0{,}183$; отже, $M = A + b = 13 +$ $+ 0{,}183 = 13{,}183$.

При обчисленні M класового ряду за A беруть середину класи з найбільшим числом варіянтів.

Приклад. Визначення ваги 226 зерен пшениці „єжовки" дало такий ряд:

Вага mgr	5–10	10–15	15–20	20–25	25–30	30–35	35–40
p	3	21	48	86	56	9	3

$n = 226$; візьмім $A = 22{,}5$ mgr; тоді

$\pm a$	5	10	15
$+p$	56	9	3
$-p$	48	21	3

$\Sigma pa = +8.5 - 12.10 \pm 0.15 =$
$= +40 - 120 \quad \pm 0 = -80$

$$b = \frac{\Sigma pa}{n} = \frac{-80}{226} = -0{,}35$$

$$M = A + b = 22{,}5 - 0{,}35 = = 22{,}15 \text{ mgr.}$$

У всіх схожих з даним прикладом випадках можна запровадити ще одне спрощення. А саме, можна взяти спершу проміжку [1]) класи за одиницю і, обчисливши b при класі $= 1$ (зазначмо літерою β), внести потім поправку, помножаючи β на числову вартість класової проміжки. В останнім прикладі обчислення піде тоді так ($a' = a$ при класовій проміжці, взятій за 1):

$\pm a'$	1	2	3
$+p$	56	9	3
$-p$	48	21	3

$\Sigma pa' = +8.1 - 12.2 \pm 0 =$
$= +8 - 24 = -16$

$$\beta = \frac{\Sigma pa'}{n} = \frac{-16}{226} = -0{,}07$$

$$b = \beta \,.\, 5 = -0{,}07 \,.\, 5 = -0{,}35$$
$$M = A + b = 22{,}5 - 0{,}35 = 22{,}15 \text{ mgr.}$$

Запис можна вести й трохи инакше. Напр., визначити M такого ряду:

Довжина колосів	50	60	70	80	90	100	110 mm.
Число відп. „	1	4	27	69	43	6	

$n = 150$; візьмімо $A = 85$ mm., тоді матимемо:

$\pm a'$	1	2	3
$+p$	43	6	.
$-p$	27	4	1

$\Sigma(+pa') = 16.1 + 2.2 \quad . \quad = 16 + 4 = 20$
$\Sigma(-pa') = \quad . \quad . - 1.3 = \quad -3$
$\Sigma pa' = +17$

$$\beta = \frac{\Sigma pa'}{n} = \frac{+17}{150} = +0{,}113$$
$$b = \beta \,.\, 10 = +0{,}113.10 = +1{,}13$$
$$M = A + b = 85 + 1{,}13 = = 86{,}13 \text{ mm.}$$

[1]) Цим ми при знаходженні суми виносимо її за дужки: $b = \frac{\Sigma pa}{n} \times$ проміжку класи.

Спрощеного способу обчислення M (з умовним нулем) можна вживати і при малочисленних даних, коли числовий матеріял не можна розташувати в повний варіяційний ряд. Нехай, напр., нам треба обчислити M таких чисел: 235, 239, 242, 243, 243, 247, 248, 256, 270. Візьмім $A = 230$; $\Sigma a = 5 + 9 + 12 + 13 + 13 + 17 + 18 + 26 + 40 = 153$, але тут $n = 9$, отже $b = 153 : 9 = 17$; звідси $M = A + b = 230 + 17 = 247$. Як що A взяти рівним 235, то $\Sigma a = 0 + 4 + 7 + 8 + 8 + 12 + 13 + 21 + 35 = 108$; $b = 108 : 9 = 12$, $M = 235 + 12 = 247$. Вигоди цього способу є очевидні.

IV. Основний відхил (σ).

§ 1. В розділі I було вже показано, що середнє аритметичне дає тільки значіння середньої варіяції, при чому, по обидва боки цієї середньої варіяції, неначе від точки рівноваги, відхиляються вправо і вліво, збільшуючись і зменшуючись, варіянти ряду. Але того, як розподіляються і як далеко відхиляються ці варіянти, инакше кажучи, самого характеру і розмаху варіювання середнє аритметичне не показує. Для цього потрібний другий варіяційно-статистичний елемент.

Найпростіше, здавалося-б, скористуватися з самого розмаху варіювання, тоб-то з тої проміжки, що в ній укладаються всі дані варіяції. Так, в наведенім вище прикладі варіювання довжини колосів варіяційний розмах буде 110 mm — 60 mm = 50 mm.

Однак ця міра є занадто неточна, як це легко бачити на такім прикладі, взятім у Йогансена: розмах варіювання у бобів, змірених під ряд, без вибору, був:

у	120	бобин	15,50—10,75 mm	= 4,75 mm
„	2500	„	16,25— 8,25 „	= 8,00 „
„	5000	„	17,00— 8,25 „	= 8,75 „
„	10000	„	17,25— 8,25 „	= 9,00 „

Тоб-то, розмах варіювання із збільшенням числа варіянтів раз-у-раз зростає, дарма, що чисельність варіянтів є дуже велика. Це так вийшло тому, що при збільшенні числа варіянтів, збільшується також і ймовірність потрапити найбільш крайні з них. Цею-ж випадковістю потрапити при мірянні такі варіянти, що більш-менш далеко відхиляються, можна пояснити і результат вимірювання трьох рівночисельних груп тих самих бобів, а саме: розмах варіювання вийшов для—

1-ої	групи	в	2500	бобин	$16{,}25 - 8{,}25 = 8{,}00$	mm
2 „	„	„	„	„	$17{,}00 - 8{,}25 = 8{,}75$	„
3 „	„	„	„	„	$17{,}00 - 9{,}75 = 7{,}25$	„

Таким чином, варіяційний розмах так сильно змінюється, що його не можна вживати, як міру варіювання.

§ 2. Щоб запобігти цьому, инколи користуються з „середнього відхилу“ від M. Для цього беруть абсолютні вартості відхилів варіяцій від M, незалежно від їх знаку (це зазначаємо: $|\alpha|$), бо в противному разі вийде, як це ми бачили в розділі III, нуль ($\Sigma p\alpha = 0$),— і суму всіх цих відхилів ділять на число варіянтів; отже, середній відхил (зазначмо його B) можна виразити формулою:

$$B = \frac{\Sigma p|\alpha|}{n}.$$

Приклад. Візьмім знайомий нам варіяційний ряд:

15	16	17	18	19	20	колосків
4	19	34	28	11	4	колосів.

$M = 17{,}35$ колосків: $\Sigma p|\alpha| = 4.2{,}35 + 19.1{,}35 + 34.0{,}35 + 28.0{,}65 + 11.1{,}65 + 4.2{,}65 = 9{,}40 + 25{,}65 + 11{,}90 + 18{,}20 + 18{,}15 + 10{,}60 = 93{,}9$. Звідси $B = \frac{\Sigma p|\alpha|}{n} = \frac{93{,}9}{100} = 0{,}94$ колоска.

§ 3. Проте і середній відхил є не досить добра варіяційна міра, бо величина його порівнюючи мало залежить від характеру розподілу варіянтів в ряді, від характеру варіяційної кривої. Справді, як що ми переробимо наведений вище ряд, перекинувши частину варіянтів з крайніх варіяцій на сусідні з ними, то дістанемо, напр., такий ряд з крутими краями кривої:

15	16	17	18	19	20
1	24	34	29	14	2

Тут $\frac{\Sigma p|\alpha|}{n} = 0{,}90$, тоб-то дарма, що крива дістала тепер значні зміни, середній відхил відбив їх на собі так мало, що перший десятковий знак його навіть не змінився. Для того, щоб зміна характеру кривої відбилася на розглядуваній варіяційній мірі в достатній мірі, треба щоб у сумі відхилів, з яких визначають відшукувану міру, крайні варіяції порівнюючи переважали середні, бо характер кривої залежить від крайніх варіяцій відносно більше, як від середніх. Цього досягають піднесенням числових вартостей варіяцій до степеня: до другого (квадрата), третього (куба), четвертого то-що. Так роблять тому, що при степенюванні відносна числова вартість відхилу наростає тим швидче, що є далі вона від *M*. Напр., відхил 3 дає в квадраті 9, тоб-то буде в 3 рази більший, а відхил 6 дасть в квадраті 36, тоб-то буде в 6 разів більший.

Виходячи з оцих міркувань, а також із цілої низки положень, що сюди стосуються і які можна одержати за допомогою метод математичної аналізи, за відшукувану нами міру варіювання беруть середній квадратовий відхил, рівний $\pm\sqrt{\frac{\Sigma p\alpha^2}{n}}$; його ми зватимем „основним відхилом“ [1]) і зазначатимем малою грецькою сигмою σ; отже:

$$\sigma = \pm\sqrt{\frac{\Sigma p\alpha^2}{n}}.$$

З вищих степенів відхилів користуються для числового схарактеризовування инших рис варіяційного ряду. Так, напр., середній кубічний відхил є дуже придатний для визначування степеня асиметричности ряду, а середній відхил четвертого степеня дає можливість добре схарактеризувати відносне виступання або пониження вершка варіяційної кривої. З цими мірами ми познайомимося згодом, тепер-же перейдімо до способів обчислення основного відхилу.

§ 4. Щоб визначити основний відхил (σ), потрібно: 1) відшукати відхили варіяцій (W) від M (тоб-то α), 2) піднести знайдені відхили до квадрату (α^2), 3) помножити знайдені квадрати на відповідні числа варіянтів ($p\alpha^2$), 4) додати знайдені добутки ($\Sigma p\alpha^2$), 5) поділити цю суму на число всіх варіянтів $\left(\frac{\Sigma p\alpha^2}{n}\right)$ і 6) з одержаної частки добути квадратовий корінь $\left(\sigma = \pm\sqrt{\frac{\Sigma p\alpha^2}{n}}\right)$.

Приклад: Обчислім основний відхил відомого вже нам варіяційного ряду, з $M = 17{,}35$:

15	16	17	18	19	20 колосків
4	10	34	28	11	4 колоси

[1]) Звуть його ще „стандардною девіяцією“, передаючи майже літера в літеру англійський термін „standard deviation“.

Обчислення ведемо так:

варіяції W	відхили від M α	квадрат відхилів α^2	число варіянтів p	добутки $p\alpha^2$
15	$-2{,}35$	5,5225	4	22,0900
16	$-1{,}35$	1,8225	19	38,6275
17	$-0{,}35$	0,1225	34	4,1650
18	$+0{,}65$	0,4225	28	11,8300
19	$+1{,}65$	2,6225	11	28,8475
20	$+2{,}65$	7,0225	4	28,0900

$$\Sigma p\alpha^2 = 129{,}65$$

$$n = 100$$

$$\frac{\Sigma p\alpha^2}{n} = 1{,}2965$$

$$\sigma = \pm\sqrt{\frac{\Sigma p\alpha^2}{n}} = \pm\sqrt{1{,}2965} = \pm 1{,}14 \text{ кол.}$$

Отже, основний відхил для нашого ряду дорівнює $\pm 1{,}14$ колоска. Для його обчислення довелося проробити цілу низку досить складних степенювань, множень, додавань, що вимагає багато часу і терпіння. Тому цілком натурально було намагатися заступити в формулі основного відхилу α, що являє собою в переважній більшості випадків ціле число з дробом, відхилом варіяцій від основного нуля, тоб-то заступити числом a, як це робиться за спрощеного способу обчислення середнього аритметичного.

§ 5. Цю формулу для спрощеного способу обчисляти основний відхил виводять таким способом. За нашим зазначенням $\alpha = W - M$, $W - A = a$. Але $M = A + b$, значить $\alpha = W - M = W - (A + b) = (W - A) - b = a - b$; звідки $\alpha + b = a$. Таким чином, кожне a є завжди більше (або менше, якщо „b“ є від'ємна величина) за відповідне α на b, і його можна заступити сумою $\alpha + b$.

Тому ми можемо написати:

$$\frac{\Sigma p\,(\alpha + b)^2}{n} = \frac{\Sigma pa^2}{n}, \text{ або } \frac{\Sigma p\,(\alpha^2 + 2\,\alpha\, b + b^2)}{n} = \frac{\Sigma pa^2}{n}.$$

Розкриваючи в останній рівності дужки, можемо написати її так:

$$\frac{\Sigma p\alpha^2}{n} + \frac{\Sigma p\alpha . 2b}{n} + \frac{\Sigma pb^2}{n} = \frac{\Sigma pa^2}{n}.$$

Розгляньмо три доданки лівої частини цеї рівности. $\frac{\Sigma pa^2}{n}$ є σ^2, тоб-то квадрат основного відхилу $\left(\sigma^2 = \frac{\Sigma pa^2}{n}\right)$. У другім доданкові Σpa є сума всіх відхилів від M; вона, як це ми знаємо з розділу III, дорівнює нулеві; отже й весь доданок $\frac{\Sigma pa \,.\, 2b}{n} = 0$. В третім доданкові Σp є сума всіх варіянтів, тоб-то n; отже $\frac{\Sigma p}{n} . b^2 = b^2$. Таким чином, наша рівність переходить на таку:

$$\sigma^2 + b^2 = \frac{\Sigma pa^2}{n}, \text{ або } \sigma^2 = \frac{\Sigma pa^2}{n} - b^2,$$

звідки й маємо відшукувану формулу: $\sigma = \pm\sqrt{\frac{\Sigma pa^2}{n} - b^2}$.

Щоб мати нагоду обчислити σ із цеї формули, візьмімо для прикладу знайоме нам варіювання числа колосків в ульки. При $A = 17$ колосків b дорівнює 0,35, отже $b^2 = 0,1225$. Обчислення ведемо так (порівняйте з обчисленням M!):

	15	16	17	18	19	20	колосків
	4	19	34	28	11	4	колосів
$\pm a$	1	2	3				
p при $(+a)^2$	28	11	4				
p при $(-a)^2$	19	4					

pa^2 47.1^2 15.2^2 4.3^2 [1])

$$\Sigma pa^2 = 47.1 + 15.4 + 4.9 =$$
$$= 47 + 60 + 36 = 143$$

$$\frac{\Sigma pa^2}{n} = \frac{143}{100} = 1,43$$
$$b^2 = 0,1225$$
$$\frac{\Sigma pa^2}{n} - b^2 = 1,3075$$
$$\sigma = \pm\sqrt{\frac{\Sigma pa^2}{n} - b^2} = \mp\sqrt{1,3075} =$$
$$= \pm 1,14 \text{ колоска.}$$

Для даного ряду ми маємо, таким чином, такі варіяційні міри: $M = 17,35$ колосків; $\sigma = \pm 1,14$ колоска.

Обчисляючи основний відхил класового ряду, беруть за A середину підхожої класи і проводять обчислення так, як в попередньому прикладі.

П р и к л а д. Обчислімо σ для знайомого нам варіювання ваги зерна „єжовки“. При $A = 22,5$ mgr, $b = 0,35$, отже $b^2 = 0,1225$.

5	10	15	20	25	30	35	40 mgr
3	21	48	86	56	9	3	зернин

---

[1]) Числа варіянтів ми тут додаємо, бо $p\,(+a)^2$ і $p\,(-a)^2$ будуть обоє додатні; $p_n\,(+a_n)^2 + p^{(n)}\,(-a_n)^2 = a_n^{\,2}\,(p_n + p^{n})$.

$\pm a$	5	10	15
p при $(+a)^2$	56	9	3
p при $(-a)^2$	48	21	3
pa^2	104.5^2	30.10^2	6.15^2

$$\Sigma pa^2 = 104.25 + 30.100 + 6.225 =$$
$$= 2600 \quad + 3000 \quad + 1350 = 6950$$

$$\frac{\Sigma pa^2}{n} = \frac{6950}{226} = 30{,}75$$
$$b^2 = \qquad 0{,}1225$$
$$\frac{\Sigma pa^2}{n} - b^2 = 30{,}6275$$

$$\sigma = \pm \sqrt{30{,}6275} = \pm 5{,}55 \text{ mgr.}$$

В останньому прикладі і в усіх схожих випадках значно вигідніше, як і при обчисленні M, взяти спершу проміжку класи за одиницю і, обчисливши таке σ „одиничне“ (зазначмо його літерою s), внести потім відповідну поправку, помножаючи s на класову проміжку; при цьому треба, звичайно, брати для обчислення не „b“, а „β“; $s = \pm \sqrt{\frac{\Sigma pa'^2}{n} - \beta^2}$.

Останній наш приклад матиме такий вигляд:

$\pm a'$	1	2	3
$+p$	56	9	3
$-p$	48	21	3

$$\Sigma pa'^2 = 104.1 + 30.4 + 6.9 =$$
$$= 104 \quad + 120 \quad + 54 = 278$$

$$\frac{\Sigma pa'^2}{n} = \frac{278}{226} = 1{,}23$$
$$\beta^2 = \qquad 0{,}0049$$
$$s = \mp \sqrt{1{,}2251} = \pm 1{,}11$$
$$\sigma = s.5 = \pm 1{,}11.5 = \pm 5{,}55 \text{ mgr.}$$

§ 6. Коли число варіянтів є дуже мале (менше за 20), то треба користуватися з иншої формули для σ, а саме:

$$\sigma = \pm \sqrt{\frac{\Sigma D^2}{n-1}},$$

де літерою D зазначено відхил варіянта від M. Напр., дано 8 варіянтів (V): 50, 48, 41, 54, 49, 43, 49 і 50; визначити σ.

Обчислення ведемо так:

V	D	D^2
50	+2	4
48	0	0
41	−7	49
54	+6	36
49	+1	1
43	−5	25
49	+1	1
50	+2	4
$M = 384 : 8 = 48$	$\Sigma D^2 = 120$	

$$\sigma = \pm \sqrt{\frac{\Sigma D^2}{n-1}} = \pm \sqrt{\frac{120}{8-1}} =$$
$$= \pm \sqrt{\frac{120}{7}} = \pm \sqrt{17{,}14} = \pm 4{,}14.$$

При цьому, в тих випадках, коли M є рівне цілому з дробом, вигідніше користуватися з формули

$$\sigma = \pm\sqrt{\frac{\Sigma a^2}{n} - b^2} \cdot \sqrt{\frac{n}{n-1}}\ ^{1)}$$

розглядаючи кожний варіянт (V), як варіяцію з одним варіянтом.

Приклад. Нехай дано ряд: 51, 48, 41, 54, 49, 43, 49, 50; $\sigma = ?$ Візьмім $A = 50$ і проведімо обчислення так:

V	a	a^2
51	$+1$	1
48	-2	4
41	-9	81
54	$+4$	16
49	-1	1
43	-7	49
49	-1	1
50	0	
	$\Sigma a = -15$	$\Sigma a^2 = 153$

$$b = \frac{\Sigma a}{n} = \frac{-15}{8} = -1{,}875;$$

$$M = A + b = 50 - 1{,}875 = 48{,}125;$$

$$\frac{\Sigma a^2}{n} = \frac{153}{8} = 19{,}125;\ b^2 = 3{,}516;$$

$$\sigma = \pm\sqrt{\frac{\Sigma a^2}{n} - b^2} . \sqrt{\frac{n}{n-1}} =$$

$$= \pm\sqrt{19{,}125 - 3{,}516} . \sqrt{\frac{8}{7}} =$$

$$= \pm\sqrt{15{,}609 . 1{,}143} = \pm\sqrt{17{,}84} =$$

$$= \pm 4{,}22.$$

Нарешті, коли ряд складається з малої кількости невеликих чисел, то є рація виходити від справжнього нуля, взяти $A = 0$; тоді $a = V$, $b = \frac{\Sigma a}{n} = \frac{\Sigma V}{n} = M$, -- і формула матиме такий вигляд:

$$\sigma = \pm\sqrt{\frac{\Sigma V^2}{n} - M^2} . \sqrt{\frac{n}{n-1}}.$$

Приклад. Дано ряд: 2, 4, 3, 7, 5, 5, 6, 4; $\sigma = ?$ Обчислімо:

V	V^2
2	4
4	16
3	9
7	49
5	25
5	25
6	36
4	16
$\Sigma V = 36$	$\Sigma V^2 = 180$

$$M = \frac{36}{8} = 4{,}5;\ M^2 = 20{,}25;\ \frac{\Sigma V^2}{n} = \frac{180}{8} = 22{,}5;$$

$$\frac{\Sigma V^2}{n} - M^2 = 22{,}5 - 20{,}25 = 2{,}25;$$

$$\sigma = \pm\sqrt{\frac{\Sigma V^2}{n} - M^2} . \sqrt{\frac{n}{n-1}} =$$

$$= \pm\sqrt{2{,}25 . 1{,}14} = \pm\sqrt{2{,}57} = \pm 1{,}6.$$

---

1) Бо $\sqrt{\frac{\Sigma D^2}{n-1}} = \sqrt{\frac{\Sigma D^2}{n}} . \sqrt{\frac{n}{n-1}}$, а $\sqrt{\frac{\Sigma D^2}{n}} = \sqrt{\frac{\Sigma a^2}{n} - b^2}$.

§ 7. Нам залишається ще познайомитися з визначенням основного відхилу при альтернативнім варіюванні. Спинімося спершу на випадкові 2 альтернатив; напр., білокольорові і фіялкові (§ 4, роз. II).

Як що одну альтернативу взяти за одиницю, то числова вартість другої дорівнюватиме нулеві, і ми матимемо діло з двома варіяціями: нулевою та одиничною. В нашому прикладі, коли за одиницю взяти фіялкову варіяцію, то всіх білокольорових рослин треба віднести до не фіялкової, тоб-то до нулевої варіяції.

W	0	1	n
p	160	398	558

Маючи за підставу цей ряд, можна обчислити σ одним з вищепоказаних шляхів, але краще скористуватися з особливої, дуже простої формули, яку можна вивести таким способом.

Зазначмо, відповідно знаками p_0 і p_1 чисельність варіяцій 0 і 1; $n = p_0 + p_1$. Тоді можна зробити такий запис:

W	0	1	n
p	p_0	p_1	p_0+p_1

Візьмім за A варіяцію 0; тоді $b = \frac{p_1}{n}$ і $b^2 = \frac{p_1^2}{n^2}$.

Тут $a = 1$, отже, $\Sigma pa^2 = p_1$ і $\frac{\Sigma pa^2}{n} = \frac{p_1}{n}$. Тому:

$$\sigma = \pm\sqrt{\frac{\Sigma pa^2}{n} - b^2} = \pm\sqrt{\frac{p_1}{n} - \frac{p_1^2}{n^2}} = \pm\sqrt{\frac{np_1 - p_1^2}{n_2}} =$$

$$= \pm\sqrt{\frac{(p_0 + p_1)\,p_1 - p_1^2}{n^2}} = \pm\sqrt{\frac{p_0 p_1 + p_1^2 - p_1^2}{n^2}} = \pm\sqrt{\frac{p_0 p_1}{n^2}} =$$

$$= \pm\sqrt{\frac{p_0 \cdot p_1}{n}}.$$

Для нашого прикладу—

$$\sigma = \pm\frac{\sqrt{160.398}}{558} = \pm 0{,}452.$$

Справа ще простішає, як що числа варіянтів дано у відсотках (тоб-то в сотих частинах всього числа варіянтів). В цім випадку можемо написати:

$$\sigma = \pm 100 \frac{\sqrt{p_0 p_1}}{n} \% = \pm \sqrt{\frac{100 p_0}{n} \cdot \frac{100 p_1}{n}} \%$$ або, через те, що $\frac{100 p_0}{n}$ і $\frac{100 p_1}{n}$ є p_0 і p_1 у відсотках від n, инакше кажучи, відсоткове p_0 ($\% p_0$) і відсоткове p_1 ($\% p_1$) —

$$\sigma = \pm \sqrt{\% p_0 \cdot \% p_1} \% .$$

В нашім прикладі—

W	0	1
$\% p$	28,7	71,3

$$\sigma = \pm \sqrt{28,7 \cdot 71,3} \% = \pm 45,2\%.$$

Цієї-ж самої формули вживають і при обчисленні σ для більшого числа альтернатив. У використованім допіру прикладі що до барви насіння було 4 альтернативи: було —

39 жовтих і 519 нежовтих, 105 фіялкових і 453 нефіялкових,
121 брондзових і 437 неброндзових, 293 чорних і 265 нечорних.

Виразивши ці гасла у відсотках, одержимо:

жовті 7,0 проти 93,0; $\sigma = \pm \sqrt{7.93} = \pm 25,5\%$
брондзові 21,7 « 78,3; $\sigma = \pm \sqrt{21,7.78,3} = \pm 41,2\%$
фіялкові 18,8 « 81,2; $\sigma = \pm \sqrt{18,8.81,2} = \pm 39,1\%$
чорні 52,5 « 47,5; $\sigma = \pm \sqrt{52,5.47,5} = \pm 49,6\%$

V. Значіння основного відхилу.

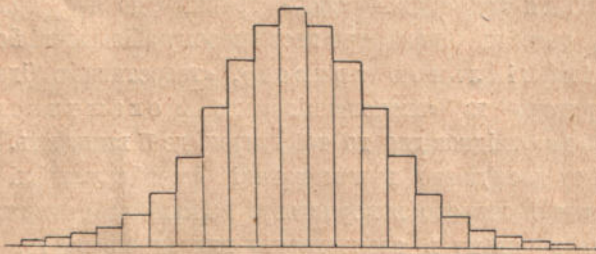
§ 1. Ми вже знаємо, що коефіцієнти розкладу Ньютонового біному дають ідеально-симетричну варіяційну криву. Так, при $a = b = 1$ ми маємо (як що ряд представити в вигляді класового, а кожного коефіцієнта виразити у відсотках від усього їх числа, від їх суми):

	−8	−7	−6	−5	−4	−3	−2	−1	0	+1	+2	+3	+4	+5	+6	+7	+8
$(a+b)^2 =$								25,00	50,00	25,00							
$(a+b)^4 =$							6.25	25,00	37,50	25,00	6,25						
$(a+b)^{10} =$				0,10	0,98	4,39	11,71	20,52	24,60	20,52	11,71	4,39	0,98	0,10			
[1] $(a+b)^{20} =$	0,02	0,11	0,46	1,48	3,70	7,39	12,01	16,02	17,62	16,02	12,01	7,39	3,70	1,48	0,46	0,11	0,02

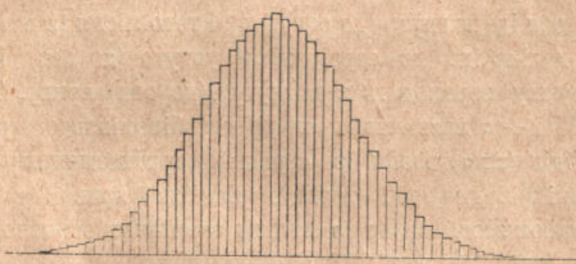
[1]) 1-ша, 2-га, 20-та і 21-ша класи є остільки малі численно, що ми їх не виписуємо.

Як що ці ряди зобразити графічно, то вийдуть такі східчасті криві (рис. 3).

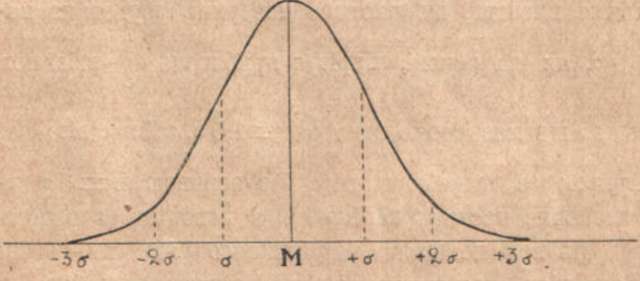
На таких кривих ми бачимо, що із збільшенням степеня бінома збільшується і число клас (рівне, взагалі, покажчикові степеня плюс



Крива $(a+b)^{20}$
Рис. 3-а.



Крива $(a+b)^{60}$
Рис. 3-б.



Біноміяльна крива.
Рис. 4.

одиниця), поверхні, що виявляють собою кількості відповідних варіянтів, стають порівнюючи що-раз важчими і ширина-ж східців що-раз меншою. При дуже високому степені бінома ширина кожної класової поверхні буде остільки малою, що поверхня практично зведеться на лінію, ширина-ж кожного східця на точку, і замість східчастої кривої ми дістанемо тоді практично (на рисунку) справжню криву лінію особливої форми, подану на рис. 4.

Теоретично ця крива являє собою границю, до якої прямують ряди сучинників у розкладі бінома при нескінченому збільшенні його степеня, і зветься «біноміяльною кривою», «теоретичною (ідеальною) варіяційною кривою» та «експоненціяльною (покажчиковою) кривою похибок», відповідно до тої галузи, де її вживають. Біноміяльну криву детально досліджено в математиці, і найважливіші результати цього дослідження широко використовує варіяційна статистика.

Спинятися на цьому в елементарнім викладі не варто, ми відзначимо тільки, як стосується до біноміяльної кривої основний відхил.

Розглянувши біноміяльну криву, рис. 4, де відзначено також віддалі від середини до σ, 2σ, и 3σ, ми легко помітимо, що крива, починаючи від поземної оси іде вигнуто (до оси) аж до того місця її, що лежить понад точкою $-\sigma$, далі-ж іде до середини M угнуто (до оси) і в точці, що понад $\pm\sigma$, знову переходить в вигнуту (до оси) криву. Таким чином, точки $-\sigma$, M і $+\sigma$ визначають характер біноміяльної кривої, а тому основний відділ ($\pm\sigma$) разом із середнім аритметичним (M) являють собою основні (кординальні) точки варіяційної кривої.

§ 2. Познайомимося тепер із иншим значінням основного відхилу, що має особливу вагу для варіяційної статистики.

Вище ми бачили (на рис. 4), що площа, обмежена варіяційною кривою та поземною віссю, визначає собою всю кількість (n) варіянтів даного ряду, всі 100% їх. З того-ж рисунку можна вивести, що частина поверхні, збудована на відтинкові від $-\sigma$ до $+\sigma$, як на основі, вирізує приблизно $\frac{2}{3}$ всієї поверхні кривої; поверхня, що має своєю основою $\pm 1{,}5\,\sigma$, займає вже коло $\frac{6}{7}$ всієї поверхні кривої; за межами поверхні з основою в $\pm 3\sigma$ залишається вже дуже незначна частина варіянтів, а поверхня з основою $\pm 3{,}5\sigma$ практично дасть всю поверхню варіяційного ряду, тоб-то число варіянтів, визначене цією проміжкою, можна взяти за n, або за 100%. При точнім обчисленні виявляється, що числа варіянтів розподіляються таким способом (див. табл. на стор. 25):

На підставі цих залежностів, маючи який-будь варіянт ряду і знаючи його σ, можемо з певною ймовірністю робити висновок за величину середнього аритметичного (M)

При проміжці	В межах проміжки	Поза проміжкою	Відношення 1-ої до 2-ої (кругло)
$M \pm \sigma$	68,3 %	31,7 %	2 : 1
$M \pm 1{,}50\,\sigma$	86,6 —	13,4 —	6 : 1
$M \pm 2{,}00\,\sigma$	95,5 —	4,5 —	21 : 1
$M \pm 2{,}50\,\sigma$	98,8 —	1,2 —	82 : 1
$M \pm 2{,}75\,\sigma$	99,4 —	0,6 —	116 : 1
$M \pm 3{,}00\,\sigma$	99,7 —	0,3 —	332 : 1
$M \pm 3{,}25\,\sigma$	99,9 —	0,1 —	999 : 1
$M \pm 3{,}50\,\sigma$	99,96—	0,04—	2500 : 1

цього ряду. Ми маємо 21 шанс проти 1 за те, що від узятого навгад варіянту M лежить не далі, як $\pm 2\,\sigma$; можемо з досить великою імовірністю (82 проти 1) сказати, що M лежить в межах, віддалених від усякого варіянта не далі, як на $\pm 2{,}5\,\sigma$; нарешті, практично без ризикування, тоб-то безпомилково (332 проти 1) можемо твердити, що кожний, взятий навгад, варіянт виражає собою середнє аритметичне з похибкою, яка не перевищує $\pm 3\,\sigma$, бо число варіянтів, віддалених від M на величину, більшу за $3\,\sigma$, надзвичайно мале (менше як 0,3%).

З огляду на такі залежності основний відхил називають ще „середньою похибкою окремих варіянтів“.

Через це саме, маючи варіянт, якого величина відрізняється від M відомого нам ряду, що дає варіювання тої самої ознаки, більше, як на $3\,\sigma$, ми можемо з практично цілковитою певністю твердити, що цей варіянт належить до иншого ряду, з иншим M.

Напр., коли для відомого нам варіяційного ряду, що виражає варіювання числа колосків в колосі, $M = 17{,}35$ кол. і $\sigma = \pm 1{,}14$ кол., то натрапивши на колос з 13 колосками, отже з числом їх, меншим за відоме нам M на 4,35 колосків, ми можемо цілком упевнено, практично безпомилково твердити, що цей колос належить до иншого варіяційного ряду, напр., є представник иншого сорту, або, хоч і того самого сорту, але виріс в инших життьових обставинах, бо відрізняється від нашого M на $-3{,}8\,\sigma$.

Або, напр., коли середня вага (M) зернини дослідженого нами насінневого зразку дорівнює 30 mgr і її $\sigma = 5$ mrg, то натрапивши на зернину, яка важить 43 mgr, ми можемо з великою імовірністю твердити, що ця зернина належить до насінневого матеріялу иншого характеру, але устоювати за це з цілковитою певністю, категоричністю не будемо, бо вага нової зернини відрізняється від M відомого нам зразку тільки на 13 mgr, тоб-то на 2,6 σ, отже, може належати до того самого ряду.

VI. Квартиль, або ймовірний відхил (Q).

§ 1. З метою, зазначеною наприкінці попереднього розділу часто користуються ще й тепер з так зван. „квартилю“, або „ймовірного відхилу“, що його звуть також „ймовірною похибкою окремого варіянту“. Цею назвою називають половину проміжки між двома внутрішніми (середніми) чвертями всього числа варіянтів ряду, як що весь даний матеріял поділити на 4 рівні групи.

Сказане стане зрозумілішим, коли перевести визначення квартиля на якімось прикладі. Ми спинимося на знайомому нам варіюванні ваги зерна в „єжовки“:

	5	10	15	20	25	30	35	40 mgr
	3	21	48	86	56	9	3	
Налічуваний ряд		3	24	72	158	214	223	226

Вміщений внизу „налічуваний“ ряд утворюється через додавання чисел варіянтів від класи до класи, а саме так, що до числа варіянтів кожної класи додається число варіянтів всіх попередніх клас; так, у нашому прикладі маємо: для 1-ої класи 3, для 2-ої $21 + 3 = 24$, для 3-ої $48 + 24 = 72$ і т. д. Числа налічуваного ряду показують, скільки варіянтів не переступають верхньої межі відповідної класи. Отже, налічуваний ряд у нашому прикладі показує, що 3 варіянти були не важчі за 10 mgr, 72—не важчі за 20 mgr, 223—не важчі за 35 mgr то що. Инакше кажучи, числа налічуваного ряду стосуються до верхньої межі класи, але не до його середини.

Нам треба відшукати межі 1-ої і 3-ої чверти. Межею 1-ої чверти (назвім її q_1) буде на класова міра, яку не переступає $^1/_4$ всіх варіянтів, тоб-то 56,5 $(= 226:4)$; вона повинна лежати поміж 15 і 20 mgr. Припустивши (без великої похибки) рівномірний розподіл варіянтів у класі, ми міркуємо так: при 15 mgr є 24 варіянти, отже, до q_1 не вистачає 32,5; 48 варіянтів, що лежать у класі 15—20 mgr, зсовують межу на 5 mgr, а 32,5 пересунуть її на $\frac{5}{48} \cdot 32{,}5 = 3{,}39$ mgr. Тому $q_1 = 15 + 3{,}39 = 18{,}39$ mgr.

Межа 3-ої чверти, q_3, виділяє 56,5 найважчих варіянтів (226—160,5). Вона лежить поміж 25 і 30 mgr. У 25 mgr є 158 варіянтів; 11,5 (169,5—158) варіянтів, що їх не вистачає, зсунуть межу на $\frac{5}{56} \cdot 11{,}5 = 1{,}03$ mgr. Отже, $q_3 = 25 + 1{,}03 = 26{,}03$ mgr.

Таким чином, межа 1-ої чверти $q_1 = 19{,}39$ mgr, межа 3-ої чверти $q_3 = 26{,}03$ mgr; проміжка між ними (зайнята двома внутрішніми чвертями всього числа варіянтів), поділена на 2, і дасть, як це було сказано, квартиль (Q).

$$Q = \frac{q_3 - q_1}{2} = \frac{26{,}03 - 18{,}39}{2} = \frac{7{,}64}{2} = 3{,}82 \text{ mgr.}$$

§ 2. Значіння квартиля зразуміти не важко. Проміжка в $2Q (\pm Q)$ містить, як ми бачили, половину всіх варіянтів, друга-ж половина їх лежить у крайніх чвертях ряду, себ-то за межами $q_1 - q_3$. Через те, коли ми беремо навгад який-небудь варіянт з нашого матеріялу, буде однаково ймовірне (50% проти 50%), як те, що цей варіянт попаде в серединну половину нашого ряду, так і те, що він опиниться поза її межами. Обидві можливості є однаково ймовірні, а тому квартиль називають також „імовірним відхилом“.

При инших проміжках варіянти розташовуються так:

При проміжці	В межах проміжки	Поза проміжкою	Відношення 1-ої до 2-ої (в цілих числах)
$M \pm Q$	50,0 %	50,0 %	1 : 1
$M \pm 2Q$	82,3 —	17,7 —	5 : 1
$M \pm 3Q$	95,7 —	4,3 —	22 : 1
$M \pm 4Q$	99,3 —	0,7 —	142 : 1
$M \pm 5Q$	99,93—	0,07—	1428 : 1

§ 3. Обчислений таким способом квартиль дуже мало відбиває на собі розподіл варіянтів по окремих класах, особливо крайніх, які сильніше за инші впливають на характер варіяційного ряду. Тому квартиль являє собою недосить добру міру варіювання, в порівнанні з основним відхилом значно нижче стоїть за нього.

§ 4. Для симетричної біноміяльної кривої Q і σ звязує така стала залежність: $Q = 0{,}6745\ \sigma$. Тому квартиль (імовірний відхил) обчислять звичайно з формули для σ:

$$Q = \pm 0{,}6745 \sqrt{\frac{\Sigma p a^2}{n}}.$$

Однак, не кажучи вже про те, що нема ніякої рації користуватися з Q, як з частини σ, і що значно простіше, в такому разі, брати σ безпосередньо, слід указати ще на те, що $Q = 0{,}6745\ \sigma$ тільки для ідеальної біноміяльної кривої. Коли-ж розподіл варіянтів іде за иншою кривою, за иншим законом, то Q буде більше або менше за вказане відношення. Напр., для кривої, даної на рис. 6, $Q = 0{,}294\ \sigma$. Беручи в таких випадках Q рівним $0{,}6745\ \sigma$, ми зробимо найгрубішу помилку.

Третя хиба квартиля це — неможливість уживати безпосереднього обчислення його для альтернативного варіювання.

З усіх указаних причин квартиля або ймовірного відхилу вживають тепер у варіяційно-статистичних дослідах у дуже рідких випадках із специяльною метою, і в багатьох агрономічних працях він є лише з непорозуміння.

VII. Варіяційний коефіціент (v).

§ 1. Розгляньмо такі ряди:

$$M_1 = 2{,}465\ \mathrm{m},\ \sigma_1 = \pm 0{,}005\ \mathrm{m};\ M_2 = 246{,}5\ \mathrm{cm},\ \sigma_2 = \pm 0{,}5\ \mathrm{cm};$$
$$M_3 = 2465\ \mathrm{mm},\ \sigma_3 = \pm 5\ \mathrm{mm}.$$

Абсолютна вартість основного відхилу для цих трьох рядів дорівнює відповідно 0,005 m, 0,5 cm. і 5 mm. Чи це значить, що останній ряд, для якого $\sigma = 5$, варіює більше за перші два ряди? Звичайно, ні; бо σ, так само, як і M, є число йменоване і, як таке, не дає безпосередніх вказівок що до відносного степеня варіювання, не говорить безпосередньо, який саме ряд варіює сильніше і який менше. Так, у нашому прикладі всі три σ, як і M, можна виразити в одній якійсь назві (мірі), напр., в „mm“, і тоді виявиться, що всі вони є однакові, всі дорівнюють ± 5 mm, і що, таким чином, всі три виражають один і той самий ряд, тільки в різних назвах (в метрах, в центиметрах і міліметрах).

Візьмім тепер другий приклад:

$$M_1 = 48\ \text{гр.},\ \sigma_1 = \pm 4\ \text{гр.};\ M_2 = 64\ \text{гр.},\ \sigma_2 = \pm 4\ \text{гр.}$$

Основний відхил в обох рядів виражено одним і тим самим числом (± 4 гр.). Чи це значить, що обидва ряди варіюють однаково? Ні, не значить, бо M_1 значно різниться від M_2 і порівнювати варіювання цих рядів безпосередньо за σ ми не маємо права.

Ще приклад:

$M_1 = 17{,}35$ колосків, $\sigma_1 = \pm 1{,}14$ колоска; $M_2 = 86{,}1$ mm, $\sigma_2 = \pm 8{,}8$ mm.

Перші два числа характеризують варіювання числа колосків у колосі, а других два—варіювання довжини колоса. Чи можна безпосередньо, беручи на увагу σ, сказати, що варіює більше: число колосків в колосі, чи довжина колосу? Звичайно, не можна, бо і середні аритметичні тут є неоднакові і назви основного відхилу є різні.

Для наших цілей треба виразити σ відносним числом, а саме, взявши M за 1, виразити числом $\frac{\sigma}{M}$, або-ж, взявши $M = 100$, тоб-то у відсотках. Цей приведений основний відхил (зазначмо його літерою v) звуть „варіяційним коефіцієнтом“ і виражають звичайно у відсотках, тоб-то:

$$v = \frac{100\sigma}{M} \%.$$

Для наших прикладів ми одержимо:

для першого—$M_1 = M_2 = M_3 = 2465$ mm, $v_1 = v_2 = v_3 = \pm 0{,}2\%$;
для другого—$M_1 = 48$ гр., $v_1 = \pm 8{,}3\%$; $M_2 = 64$ гр., $v_2 = \pm 6{,}3\%$;
для третього—$M_1 = 17{,}35$ кол., $v_1 = \pm 6{,}6\%$; $M_2 = 86{,}1$ mm, $v_2 = \pm 10.2\%$.

Величини v вказують, таким чином, уже безпосередньо на степінь варіювання ряду. Так, перші три варіюють однаково ($v_1 = v_2 = v_3$) і до того дуже мало ($\pm 0{,}2\%$), у других двох перший ряд варіює сильніше за другого ($8{,}4 > 6{,}3$); у третіх двох, навпаки, варіювання першого ряду є значно менше, ніж другого ($6{,}6 < 10{,}2$), який, взагалі, варіює досить сильно ($\pm 10{,}2\%$!).

§ 2. Для альтернативного варіювання відносною мірою його буде вже самий основний відхил, бо тут його виражають відсотками, а не іменованим числом ($\sigma = \pm \sqrt{\%p_1 . \%p_2}\, \%$). Тому, щоб мати певну думку про степінь альтернативного варіювання, користуються безпосередньо з основного відхилу. Легко пересвідчитися, що альтернативне варіювання буде найбільшим, коли обидві альтернативи представлено однаковими кількостями, тоб-то по 50%: тоді $\sigma = \pm \sqrt{50{,}50}\,\% = \pm 50\%$. Це варіювання буде тим менше, що менша є чисельність одної з альтернатив; так, для чисел 30% і 70% $\sigma = \pm \sqrt{30.70}\,\% = \pm 45{,}83\%$, для альтернатив 20% і 80% $\sigma = \pm \sqrt{20.80}\,\% = \pm 40\%$, для 10% і 90% $\sigma = \pm \sqrt{10.90}\,\% = \pm 30\%$ для 2% і 98% $\sigma = \pm \sqrt{2{,}98}\,\% = \pm 14\%$; нарешті, як що всі 100%

даних для вивчання об'єктів представлено тільки одною якістю, то й не можна говорити про варіювання: $\sigma = \pm\sqrt{100.0}\,\% = 0$.

§ 3. Варіяційний коефіцієнт показує, наскільки сильно або наскільки мало варіює величина, виражена даним варіяційним рядом, инакше кажучи, показує, наскільки різноманітним або, навпаки, вирівненим являється досліджуваний матеріял що до даної ознаки або властивости. Тому варіяційний коефіцієнт ми можемо назвати також покажчиком різноманітности або одноманітности (вирівнености) тої ознаки або властивости, що варіює. Що є більше v, то більша є різноманітність і менша вирівненість даного варіяційного ряду.

У цьому своєму значінні варіяційний коефіцієнт може найти різноманітне застосування в агрономії.

1-ий приклад. Однаково оброблене і однаково-ж засіяне буряками поле розбито на 256 рівних ділянок, кожна завбільшки 3×6 саж. Урожай з кожної такої ділянки зібрано і зважено окремо. Виявилось, що врожайність окремих ділянок є не однакова, але варіює. Як-би врожаї з цих дрібних ділянок перерахувати (множенням на $\frac{2400}{3.6} = 133{,}3$) на десятину, то вони виразяться числами від 1156 п. до 1972 п. Розташувавши числа врожаїв по класах, одержимо такий ряд:

1100—	1200—	1300—	1400—	1500—	1600—	1700—	1800—	1900—	2000 пуд.
6	25	61	48	50	34	22	6	4	

σ цього ряду $= \pm 172$ п., що при M, рівному тут 1495 п., дає $v = \pm 11{,}5\%$. Різноманітність цього бурякового поля, або варіювання його родючости, вираженої врожайністю ділянок у 18 кв. с., можна визнати досить значною.

Коли перевести облік урожаю цього-ж самого поля, беручи ділянки вчетверо більші, 6 с. $\times$ 12 с., то вийде такий ряд:

1250—	1350—	1450—	1550—	1650—	1750—	1850 пуд.
6	14	30	8	5	1	

Тут $\sigma = \pm 122$ п., а $v = \pm 8{,}2\%$. Таким чином, при ділянках у 72 кв. с. різноманітність поля помітно знизилась.

Дальше збільшення поверхні облікових ділянок дасть чим раз більше зниження різноманітности, инакше, — чим раз більшу вирівненість показів урожайности по окремих ділянках. Що є менше v, то з меншою похибкою, то точніше ми можемо визначити M, тоб-то справжню врожайність даного поля (або дослідної станції). Отже, збільшення поверхні облікової (або дослідної) ділянки призводить до збільшення точности визначування справжнього врожаю (або точности спроби). Отже покажчик вирівнености дослідної ділянки є також і покажчик точности спроби, як що оцінювати точність її з окремих ділянок, тоб-то без повторення спроби.

2-й приклад. Оцінюючи різні якості зерна, часто надають певної ваги його вирівненості. З'ясовують степінь вирівнености звичайно так: беруть до уваги відносну кількість зерна, що проходить і що не проходить через 1—2 сита з різними завбільшки дірочками, або, як то кажуть, беруть на увагу відносну кількість зерна 1-го, 2-го і 3-го «сорту».

Однак, цей спосіб визначати вирівненість зерна часто дає дуже значні помилки, бо у різних сортів середня величина зернини буває звичайно не зовсім однаковою, а частенько й зовсім різною (напр., зернина арнаутки, пересічно, є вдвоє більша за зернину гирки). Вживаючи сит із тими самими дірочками, для дрібнозерного сорту, ми одержимо більше зерна друго- і третьосортного, а для дорідноверного, навпаки, більше першо- і другосортного зерна, хоча-б обидва сорти, що їх порівнюємо, мали зерно однакової вирівнености.

Варіяційний коефіціент, як покажчик вирівнености тої ознаки або властивости, що варіює, може й тут дуже нам стати в пригоді. Змірявши ширину (а ще краще й довжину) 200—500 зернин (що більше, то краще), ми розташовуємо одержані дані за класами варіяційного ряду; або-ж—за допомогою кількох сит з такими дірочками, що відповідають класовим проміжкам (напр., кожне на 0,25 mm. ширше за попереднє), просто „розсіваємо" наш зразок за класами, і чисельності цих клас дадуть нам відповідний варіяційний ряд. Обчисливши тоді v цього ряду, ми дістанемо число, що точно характеризуватиме вирівненість зерна досліджуваного сорту. Так, для зерна лінійних і „господарських" сортів озимої пшениці, що виросли в однакових умовах, ми одержали 1917 року такі M і v.

	Банатка місцева	№ 4	№ 6	№ 7	Ячмінь № 3	Ячмінь № 4
v ширини	10,2	10,8	11,4	9,8	7,4	6,8
v довжини	8,2	7,7	8,2	5,9	7,9	6,6
M ширини зерна в *mm* .	26,8	27,5	28,0	29,3	33,7	31,3
M довжини зерна в *mm* .	61,3	65,1	62,0	62,8	86,8	93,1

З цієї таблиці виразно помітно, що у пшениць довжина зерна була більш вирівненою, як ширина, в ячмінів-же обидві ці величини варіювали однаково; помітно також, що в окремих сортів вирівненість зерна дуже не однакова: від $\pm 5{,}9\%$ у № 7 до $\pm 8{,}2\%$ у № 6 та банатки місцевої, це що до ширини; що-ж до довжини, то найменшу вирівненість виявив № 6.

3-ій приклад. При роботі на малих ділянках сівалка кидає на кожну з них не однакову кількість зерна, тоб-то ці кількості зерна варіюють; треба визначити, які завеликі можуть бути ці коливання для окремих ділянок різного розміру, напр., в 1, 3 і 10 кв. с. Підвязавши до отворів висівального апарату мішечки, даємо сівалці пройти ту путь, що потрібна для засіву 1 кв. с. Викинуте при цьому в мішечки зерно зважуємо. Повторивши цю операцію 2—3 десятки разів, одержимо 20—30 чисел (що більше, то краще); з найдених чисел обчисляємо v, яке й покаже нам степінь вирівнености кількости насіння, що висіває дана сівалка на 1 кв. с.

Так само визначаємо v і для засіву на ділянках в 3 і 10 кв. с.

При одному такому досліджуванні селекційної сівалки Ельворті ми одержали такий результат:

при засіві ділянки в	1 кв. с.	$v = \pm 5{,}7\%$
„ „ „ „	3 „ „	$v = \pm 3{,}5\%$
„ „ „ „	10 „ „	$v = \pm 1{,}2\%$

Через те, що коливання для окремих ділянок, згідно з таблицею, наведеною в § 2 розділу V, можуть досягати величини в 3 рази більшої, то зрозуміло, що провадити, напр., звичайного типу спробу на вплив густоти засіву, беручи ділянки в 1 і в 3 кв. с., не можна; тільки величина ділянки в 10 кв. с. забезпечує достатню вирівненість кількости зерна, що засівають на ділянці.

Аналогічно з цим, установляючи висівальний апарат сівалки на потрібну кількість засіву, можна визначати, скільки оборотів повинно зробити колесо, щоб вийшла бажана точність установлення. Так, напр., для тої самої селекційної сівалки при 10 оборотах колеса $v = \pm 5{,}5\%$, точність, видима річ, недостатня; при 30 оборотах $v = \pm 1{,}1\%$, добра точність.

VIII. Серединна похибка (m).

§ 1. З розділу V ми знаємо, що основний відхил зветься також середньою похибкою окремого варіянту, бо дає можливість визначити з одного варіянту даного ряду його середнє аритметичне з певною надійністю або, инакше кажучи, з певної величини похибкою; бо на практиці вона не перевищує $\pm 3\sigma$, коли розтошування варіянтів у ряду буде приблизно біноміяльним. Можна висловитись і так: кожний даний варіянт являє собою приблизне середнє аритметичне ряду, що відхиляється від його справжнього середнього аритметичного практично не далі, як $\pm 3\sigma$; величина σ визначає, таким чином, степінь цього наближення, цієї похибки.

Що буде з цією середньою похибкою, з цим степенем наближення до справжнього середнього аритметичного, коли його визначити не з одного варіянту, а з середнього аритметичного із 2... 5... 20... n варіянтів, взятих навгад із даного матеріялу, що варіює? Щоденний досвід учить нас, та цього й наперед можна чекати, що надійність визначення M буде при цьому зростати: що більшу групу варіянтів ми візьмемо для визначення M, то ближче до справжнього середнього всього матеріялу буде M цієї групи.

Спробуймо тепер визначити, чому дорівнює такий середній основний відхил, така середня похибка середнього аритметичного групи, що складається з n варіянтів.

Для цього розгляньмо спочатку, чому дорівнюватиме середня похибка суми двох довільно взятих варіянтів V_x і V_y (із двох варіяційних рядів), яких середні похибки зазначмо відповідно σ_x і σ_y.

Величина V_x є величина мінлива, буває вона то більшою, то меншою, з огляду на те, на який саме варіянт ми потрапимо. Зазначмо відхили V_x від справжнього середнього знаками $x_1, x_2, x_3, \ldots x_n$. Тоді $\sigma_x = \pm \sqrt{\frac{x_1^2 + x_2^2 + x_3^2 + \cdots + x_n^2}{n}}$, або $\sigma_x^2 = (x_1^2 + x_2^2 + x_3^2 + + \cdots x_n^2) : n$.

Відхили V_y від його справжнього середнього зазначмо літерами $y_1, y_2, y_3, \ldots y_N$. Тоді $\sigma_y^2 = (y_1^2 + y_2^2 + y_3^2 + \cdots + y_N^2) : N$.

Нам треба відшукати основний відхил, що характеризує варіювання суми V_x і V_y або — в знаках σ, для $S = V_x + V_y$. Утворюючи пари із V_x і V_y, ми повинні комбінувати кожне V_x з кожним V_y, бо ми вважаємо їх тут за незалежні одне від одного, отже кожний варіянт x-го ряду можна комбінувати з кожним варіянтом y-го ряду. Всіх V_x буде n і всіх V_y буде N, тому вийде $n.N$ комбінацій $V_x + V_y$. Відхил кожної з цих nN комбінацій від справжнього середнього буде виражений своєю особливою величиною, напр., $x_1 + y_1$, $x_4 + y_{10}, \ldots$, а всіх їх можна записати так:

$$\begin{array}{c} \\ n \text{ рядків} \end{array} \left\{ \overbrace{\begin{array}{lllll} x_1 + y_1; & x_1 + y_2; & x_1 + y_3; & \ldots\ldots\ldots & x_1 + y_N \\ x_2 + y_1; & x_2 + y_2; & x_2 + y_3; & \ldots\ldots\ldots & x_2 + y_N \\ x_3 + y_1; & x_3 + y_2; & x_3 + y_3; & \ldots\ldots\ldots & x_3 + y_N \\ \vdots & \vdots & \vdots & & \vdots \\ x_n + y_1; & x_n + y_2; & x_n + y_3; & \ldots\ldots\ldots & x_n + y_N \end{array}}^{N \text{ стовпців.}} \right.$$

Таким чином, ми маємо тут nN відхилів суми $V_x + V_y$ від справжнього середнього суми (M_s). Як що суму квадратів цих відхилів поділити на число їх (nN), то вийде квадрат відшукуваного основного відхилу суми $V_x + V_y$, тоб-то σ_s^2. Визначмо його.

Зауваживши, що $(x+y)^2 = x^2 + y^2 + 2xy$, можемо написати:

$$(x_1^2 + y_1^2 + 2x_1y_1) + (x_1^2 + y_2^2 + 2x_1y_2) +$$
$$+ (x_1^2 + y_3^2 + 2x_1y_3) + \ldots + (x_1^2 + y_N^2 + 2x_1y_N)$$
$$(x_2^2 + y_1^2 + 2x_2y_1) + (x_2^2 + y_2^2 + 2x_2y_2) +$$
$$+ (x_2^2 + y_3^2 + 2x_2y_3) + \ldots + (x_2^2 y_N^2 + 2x_2y_N)$$
$$(x_3^2 + y_1^2 + 2x_3y_1) + (x_3^2 + y_2^2 + 2x_3y_2) +$$
$$+ (x_3 2 + y_3^2 + 2x_3y_3) + \ldots + (x_3^2 + y_N^2 + 2x_3y_N)$$
$$\vdots \qquad\qquad\qquad\qquad \vdots$$
$$(x_n^2 + y_1^2 + 2x_ny_1) + (x_n^2 + y_2^2 + 2x_ny_2) +$$
$$+ (x_n^2 + y_3^2 + 2x_ny_3) + \ldots + (x_n^2 + y_N^2 + 2x_ny_N).$$

Кожний x^2 повторюється тут N разів і кожний y^2 — n разів, тому ми можемо перетворити наш запис на такий:

$$(Nx_1^2 + Nx_2^2 + Nx_3^2 + \ldots + Nx_n^2) + (ny_1^2 + ny_2^2 + ny_3^2 + \ldots + ny_N^2) +$$
$$+ (2x_1y_1 + 2x_1y_2 + \ldots + 2x_1y_N) + (2x_2y_1 + 2x_2y_2 + \ldots + 2x_2y_N) + \ldots +$$
$$+ (2x_ny_1 + 2x_ny_2 + \ldots + 2x_ny_N).$$

Або, коротче:

$$N(x_1^2 + x_2^2 + \ldots + x_n^2) + n(y_1^2 + y_2^2 + \ldots + y_N^2) +$$
$$+ 2(x_1 + x_2 + \ldots + x_n)(y_1 + y_2 + \ldots + y_N),$$

що, нарешті, дає

$$N(x_1^2 + x_2^2 + \ldots + x_n^2) + n(y_1^2 + y_2^2 + \ldots + y_N^2),$$

бо $(x_1 + x_2 + \ldots + x_n) = 0$ і $(y_1 + y_2 \ldots + y_N) = 0$ (одже сума всіх відхилів від M дорівнює нулеві). Таким чином квадрат відшукуваного основного відхилу буде:

$$\sigma_s^2 = \frac{N(x_1^2 + x_2^2 + \ldots + x_n^2) + n(y_1^2 + y_2^2 + \ldots + y_N^2)}{nN} =$$
$$\frac{x_1^2 + x_2^2 + \ldots + x_n^2}{n} + \frac{y_1^2 + y_2^2 + \ldots + y_N^2}{N} = \sigma_x^2 + \sigma_y^2, \text{ звідки}$$
$$\sigma_s = \pm\sqrt{\sigma^2_x + \sigma_y^2}.$$

Коли нам дано три варіянти: V_1, V_2 і V_3 з основними відхилами: σ_1, σ_2 і σ_3, то основний відхил суми $V_1 + V_2 + V_3$ можна одержати так, що спершу визначити σ_{1+2}^2 для $S_{1+2} = V_1 + V_2$ (воно $= \sigma_1^2 + \sigma_2^2$), а тоді вже σ_{1+2+3}^2 для $S_{1+2+3} = (V_1 + V_2) + V_3$ (воно $= \sigma_{1+2}^2 + \sigma_3^2$); остаточно матимемо $\sigma_{1+2+3}^2 = \sigma_1^2 + \sigma_2^2 + \sigma_3^2$, звідки $\sigma_{1+2+3} = \pm\sqrt{\sigma_1^2 + \sigma_2^2 + \sigma_3^2}$.

Отже для суми n варіянтів: V_1, V_2, V_3, ..., V_n з основними відхилами: σ_1, σ_2, σ_3, ... σ_n можна написати в загальнім вигляді:

$$S = V_1 + V_2 + V_3 + \ldots + V_n \pm \sqrt{\sigma_1^2 + \sigma_2^2 + \sigma_3^2 + \ldots + \sigma_n^2}.$$

Як що суму всіх V поділити на число їх (n), то вийде їх середнє аритметичне (M), а як що вираз для суми всіх основних відхилів (середніх похибок) цих варіянтів поділити на їх число (теж n), то вийде середня похибка цього середнього аритметичного:

$$M = \frac{V_1 + V_2 + V_3 + \ldots + V_n}{n} \pm \frac{\sqrt{\sigma_1^2 + \sigma_2^2 + \sigma_3^2 + \ldots + \sigma_n^2}}{n}.$$

Коли всі n варіянтів належать до того самого варіяційного ряду, то $\sigma_1 = \sigma_2 = \sigma_3 = \ldots \pm \sigma_n$, бо у того самого варіяційного ряду може бути тільки один основний відхил. Тому середня похибка середнього аритметичного одного варіяційного ряду (зазначмо її m) дорівнює $\frac{\sqrt{n\sigma^2}}{n}$. Спростивши цей вираз $\left(\frac{\sqrt{n\sigma^2}}{n} = \frac{\sigma\sqrt{n}}{\sqrt{n}\cdot\sqrt{n}} = \right.$ $\left. = \frac{\sigma}{\sqrt{n}}\right)$, дістанемо надзвичайно важливу варіяційно-статистичну формулу: середня похибка середнього аритметичного або — назвім її коротче — серединна похибка

$$m = \frac{\sigma}{\sqrt{n}}.$$

Ця формула показує, що надійність визначення середнього аритметичного (M) є пропорціональна до кореня квадратового з числа (n) варіянтів, взятих для визначення цього середнього аритметичного.

§ 2. Як що потрібно є визначити m безпосередньо, минаючи обчислення σ, то слід тільки підставити у формулу для m підхожий вираз для σ (розділ IV). Тоді матимемо:

$$m + \pm \sqrt{\frac{\Sigma pa^2}{n^2}}, \text{ або } m = \pm \sqrt{\left(\frac{\Sigma pa^2}{n} - b^2\right) : n}, \text{ або } m =$$
$$= \pm \sqrt{\frac{\Sigma D^2}{(n-1)n}}, \text{ або } m = \pm \sqrt{\frac{\Sigma V^2}{n} - M^2} \cdot \sqrt{\frac{1}{n-1}} =$$
$$= \pm \sqrt{\left(\frac{\Sigma V^2}{n} - M^2\right) : (n-1)}.$$

Для альтернативного варіювання $m = \pm \sqrt{\frac{\%p_1 \cdot \%p_2 \%}{n}}.$

§ 3. Значіння і суть основного відхилу не міняється від того, чи збільшимо ми або зменшимо числовий його вираз, тому $\sigma : \sqrt{n}$ можна записати, як якесь σ_k; значить, серединна похибка суми двох середніх аритметичних (m_s) із серединними похибками: $m_1 (= \sigma_{k_1})$ і m_2 (σ_{k_2}) дорівнюватиме $\pm\sqrt{\sigma_{k_1}^2 + \sigma_{k_2}^2} = \pm\sqrt{m_1^2 + m_2^2}$, а в загальнім вигляді:

$$\mathbf{m}_s = \pm\sqrt{\mathbf{m}_1^2 + \mathbf{m}_2^2 + \ldots + \mathbf{m}_n^2}.$$

Тому:

$$(M_1 \pm m_1) + (M_2 \pm m_2) + \ldots + (M_n \pm m_n) = M_1 + M_2 + \ldots + M_n \pm \\ \pm\sqrt{m_1^2 + m_2^2 + \ldots + m_n^2}.$$

§ 4. Серединна похибка різниці двох середніх аритметичних (M_1 і M_2) з серединними похибками: m_1 і m_2 дорівнює також $\pm\sqrt{m_1^2 + m_2^2}$. Справді, для виводу основного відхилу різниці ($\sigma_{\text{різн.}}$) двох варіянтів ($V_x - V_y$) ми можемо всюди вважати $x - y = x + (-y)$; в остаточних фазах, в цьому виводі, ми одержимо $(-y)^2 = y^2$, отже, кінець-кінцем

$$\sigma_{\text{різн.}} = \pm\sqrt{\sigma_x^2 + \sigma_y^2},$$

звідки витікає друга дуже важлива формула:

$$\mathbf{m}_{\text{різн.}} = \pm\sqrt{\mathbf{m}_1^2 + \mathbf{m}_2^2}.$$

Таким чином:

$$(M_1 \pm m_1) - (M_2 \pm m_2) = M_1 - M_2 \pm\sqrt{m_1^2 + m_2^2}.$$

§ 5. Занотуймо ще такі вирази:

$$(M \pm m) \,.\, k = kM \pm km; \; (M \pm m) : k = \frac{M}{k} \pm \frac{m}{k};$$

$$\frac{(M_1 \pm m_1) + (M_2 \pm m_2) + \ldots + (M_n \pm m_n)}{n} = \frac{M_1 + M_2 + \ldots + M_n}{n} \pm \\ \pm \frac{\sqrt{m_1^2 + m_2^2 + \ldots + m_n^2}}{n};$$

$$(M_1 \pm m_1) \,.\, (M_2 \pm m_2) = M_1 \,.\, M_2 \pm\sqrt{(M_1 m_2)^2 + (M_2 m_1)^2};$$

$$(M_1 \pm m_1) : (M_2 \pm m_2) = \frac{M_1}{M_2} \pm \frac{\sqrt{(M_1 m_2)^2 + (M_2 m_1)^2}}{M_2^2}.$$

Перші три є безпосередньо очевидні, два-ж останніх можна вивести тим самим шляхом, яким ми вище вивели формулу для σ суми.

IX. Значіння срединної похибки.

§ 1. Як що з великої кількости матеріялу, що варіює, напр., із купи зерна набрати велике число N груп по n варіянтів у кожній і визначати середнє аритметичне цих груп: M_1, M_2, $M_3, \ldots$, M_k, то виявляється, що всі ці середні аритметичні окремих груп варіюють, розташовуючись навколо „справжнього“ M всього матеріялу подібно до того, як це роблять окремі варіянти всякої групи навколо її середнього аритметичного. Можна розташувати ці M_1, $M_2, \ldots$, M_n по класах, як це роблять з V_1, $V_2, \ldots$, V_n і одержати відповідний варіяційний ряд середніх аритметичних. Серединна похибка всякої з груп варіянтів дає бачити границі відхилів цих M_1, $M_2, \ldots$ від „справжнього“ M, так само, як основний відхил указує на границі відхилу окремих V від їх M; адже серединна похибка є те саме σ, але зменшене в $k = \sqrt{n}$ разів, суть-же її залишається незмінною. Тому звязок поміж проміжкою, яка залежить від тої чи иншої величини m та кількістю варіянтів, що лежать у середині і поза цією проміжкою (ці варіянти тут представлено середніми аритметичними), залишається тим самим, як і для основного відхилу (§ 2 розділу V), а це дає нам можливість із середнього аритметичного якої-будь групи варіянтів і його серединної похибки визначити справжнє M з певним степенем надійности.

Міркуючи, як у розділі V, ми можемо твердити з досить великою імовірністю (82 проти 1), що справжнє середнє лежить не далі, як $\pm 2{,}5m$, від середнього даної групи варіянтів, і з практично цілковитою певністю (332 проти 1), що справжнє середнє не більше, як на $\pm 3m$ відрізняється від середнього даної групи.

Приклад. Вимірювання 100 зернин одної лінії ульки дало для їх ваги $M \pm m = 32 \pm 0{,}8$ мгр. Через те можна було сподіватися, що з дальшим вимірюванням цього зерна по сотнях середні аритметичні їх ваги коливатимуться (на практиці) в межах від $32 - 3 \,.\, 0{,}8$ до $32 + {} + 3 \,.\, 0{,}8$ мгр., тоб-то від 29,6 до 34,4 мгр. Фактична перевірка, вимірювання великого числа груп по 100 зернин, дала середні для них від 29,5 до 34 мгр.

§ 2. Із тих самих (розд. V) залежностей витікає, що два середніх аритметичних $M_1 \pm m_1$ і $M_2 \pm m_2$ ми можемо вважати не за варіянти одного й того самого справжнього M, але за приналежні до різних варіяційних рядів, лише в тому разі, коли різниця між цими середніми, принаймні, втроє більша за свою серединну похибку, тоб-то коли

$(M_1 - M_2)$ не менша за $3m_{\text{різн.}}$ [1]). Число, що показує, у скільки разів різниця між середніми є більша за її серединну похибку, можна назвати покажчиком певности цієї різниці.

1-ий приклад. $M_1 \pm m_1 = 32{,}0 \pm 0{,}8$ мгр.

$$\underline{M_2 \pm m_2 = 29{,}6 \pm 0{,}7 \quad „}$$

Різниця $= M_1 - M_2 \pm \sqrt{m_1{}^2 \pm m^2{}_2} = 2{,}4 \pm \sqrt{0{,}8^2 \pm 0{,}7^2} = 2{,}4 \pm 1{,}1$, тоб-то 2,2 : 1. Покажчик певности різниці дорівнює 2,2.

Вивід: M_1 і M_2 треба вважати за варіянти одного й того самого M, бо не доведено, щоб вони відрізнялися.

2-ий приклад. $M_1 = 32{,}0 \pm 0{,}8$ мгр.

$$\underline{M_2 = 27{,}5 \pm 0{,}7 \quad „}$$

Різниця $= 4{,}5 \pm 1{,}1$, тоб-то 4,1 : 1.

Вивід: середнє M_1 і M_2 належать до різних рядів.

3-й приклад. Із одного сорту ячменю Йоганнсен відібрав рослини 1) з найменшою і 2) з найбільшою черезернецею і засіяв кожну групу окремо. Наступними роками він повторив добір у тому-ж напрямі, засіваючи у І-ої групи зерно від рослин з найменшою черезернецею, а у ІІ-ої з найбільшою. За кілька років ці обидві групи виявили такі середні:

група І, що добирали її за найменшою черезернецею: 35,63% черезер.
„ ІІ, „ „ „ „ найбільшою „ 37,99% „

Отже, вийшла різниця в 2,36%. Чим її з'ясувати: впливом добору чи випадковістю? Розвязати це можна тільки за допомогою серединних похибок; вони дали таке:

для гр. І $M_1 = 35{,}63 \pm 1{,}08$

„ „ ІІ $M_2 = 37{,}99 \pm 1{,}21$

Різниця $M_2 - M_1 = 2{,}36 \pm 1{,}62$, тоб-то 1,5 : 1.

Таким чином, всі підстави були міркувати, що цю різницю викликала звичайна випадковість.

В таких разах слід повторити спробу і в ширшому маштабі. Спробу було повторено, і, справді, наступними роками цієї різниці більше не помічали.

4-й приклад. Дослідження впливу видів пару, зроблене на Плотянській дослідній станції, дало такий результат: урожай зерна банатки в пудах на десятину був—

---

1) Як що користуються з „імовірної похибки" $= \frac{Q}{\sqrt{n}}$, то ця різниця повинна бути в $4^1/_2$ — 5 разів більша за свою „ймовірну похибку".

	1. По чорному пару	2. По квітневому пару	Різниця 1-й — 2-й
1898 року	127,0	99,5	+27,5
1899 „	122,6	118,0	+ 4,6
1900 „	102,0	111,4	— 9,4
1901 „	172,7	179,6	— 6,9
1902 „	215,8	191,2	+24,6
1903 „	182,0	173,7	+ 8,3
1904 „	169,4	164,6	+ 4,8
Пересічно	155,9	148,3	+ 7,6

Вийшла невеличка надбавка (коло 5%) на користь чорного пару, але за окремі роки різниця врожаїв виявляє значне варіювання, як в той, так і у другий бік, і цілком природньо повстає питання, чи є надійні різниці середніх.

В таких випадках для вияснення надійности M неможна безпосередньо користуватися з чисел урожаїв, бо ці числа одержано в різні роки, отже, в обставі неоднаковій, навіть напевне різній що до умов виростання. В цих випадках треба брати для відповідних обчислень ту величину, що її має з'ясувати дана спроба, отже, в нашому прикладі— різницю врожаїв по чорному та квітневому пару.

Але, як-би ми хотіли дізнатися, напр., з метою економічного розрахунку, оскільки точне число 155,9 пуда, одержане в середньому за 7 літ, виражає величину середнього а б с о л ю т н о г о врожаю [1]) банатки по чорному пару, тоді нам потрібно було-б визначити m безпосередньо з тих числових даних, що з них виведено $M = 155{,}9$ п.

Отже, надійність середньої різниці + 7,6 п. ми обчисляємо так:

Роки	Різниця V	Відхил від середн. різн. D	Квадрат відхилів D^2
1898	27,5	+ 19,9	396,01
1899	4,6	— 3,0	9,00
1900	—9,4	—17,0	289,00
1901	—6,9	—14,5	210,25
1902	24,6	+17,0	289,00
1903	8,3	+ 0,7	0,49
1904	4,8	— 2,8	7,84
	$M = 7{,}6$		$\Sigma D^2 = 1201{,}59$

[1]) Для даної місцевости.

$$m = \pm\sqrt{\frac{\Sigma D^2}{(n-1)\,n}} = \pm\sqrt{\frac{1201{,}59}{42}} = \pm\sqrt{28{,}61} = \pm 5{,}3 \text{ п.}$$

$$M = 7{,}6 \pm 5{,}3 \text{ п.}$$

Таким чином, цікава для нас середня різниця тільки в (7,6 : 5,3) 1,4 рази є більша за свою серединну похибку, отже, її треба визнати недоведеною. Щоб осягнути потрібну надійність виводів, треба мати численніші дані, треба впровадити достатню повторність спроби.

5-й приклад. В тій самій спробі для абсолютної ваги житнього зерна (вага 1000 зернин у грамах) одержано було такі числа:

	1898	1899	1901	1902	1903	1904	Середнє
1. По чорному пару	31,1	24,0	24,6	28,6	29,1	30,1	27,9
2. „ квітнев. „	31,6	24,2	24,8	29,1	29,9	31,0	28,4
Різниця 2.—1. (V)	0,5	0,2	0,2	0,5	0,8	0,9	0,5 $= M_{\text{різн.}}$
Відхили від $M_{\text{різниці}}$ (D)	0	—0,3	—0,3	0	+0,3	+0,4	
D^2	0	0.09	0,09	0	0,09	0,16	$=0{,}43 = \Sigma D^2$

$$m = \pm\sqrt{\frac{\Sigma D^2}{(n-1)\,n}} = \pm\sqrt{\frac{0{,}43}{30}} = \pm 0{,}12 \text{ гр.}; \quad M = 0{,}5 \pm 0{,}12 \text{ гр.}$$

Отже, різниця, що розглядаємо її, в (0,5 : 0,12 =) 4,2 рази є більша за свою серединну похибку, тоб-то її треба визнати цілком доведеною хоч абсолютна величина її є дуже мала (всього тільки ½ гр).

6-й приклад. 11-літнє дослідження на Одеському дослідному полі того, яку вагу має час засіву озимої пшениці для її врожаю, виявило, що серпневий засів, порівнюючи з вересневим, дає лишку 10,4 п. Перевірьмо надійність цього виводу.

Спроби провадили при двох повтореннях (зазначмо їх цифрами I і II), при чому одержано такі числа:

Роки		Серпень (VIII)	Вересень (IX)	Різниця VIII—IX
1897	I	40,8	18,3	22,5
	II	31,5	20,4	11,1
1898	I	130,5	115,8	14,7
	II	87,9	98,7	—10,8
1899	I	61,8	73,2	—11,4
	II	63,6	81,0	—17,4
1900	I	85,9	89,2	— 3,3
	II	66,3	80,8	—14,5
1901	I	126,0	102,3	23,7
	II	133,8	97,2	36,6
1902	I	198,9	171,3	27,6
	II	180,0	162,9	17,1
1903	I	141,6	134,7	6,9
	II	129,9	124,5	5,4
1904	I	40,8	57,0	—16,2
	II	9,3	60,0	—50,7
1905	I	86,4	57,0	29,4
	II	93,0	57,6	35,4
1906	I	166,8	142,8	24,0
	II	138,9	134,4	4,5
1907	I	105,0	54,0	51,0
	II	96,6	46,2	44,4
Середнє		100,4	90,0	10,4

З опису методики цієї спроби видно, що в окремих повтореннях ділянка із серпневим засівом містилася поруч із ділянкою вересневого засіву, але між самими повтореннями була досить значна проміжка ~~часу,~~ поля через те одно повторення, порівнюючи з другим, могло проходити в помітно неоднакових, однобічно змінених умовах виростання [1]). В таких випадках було-б ризиковано користуватися з середніх із кожної пари повторень, значно є надійніше вжити й тут способу, що з його ми користувалися в 2 попередніх прикладах. Заокругливши різниці до цілих чисел, бо це заокруглення на остаточний результат особливого впливу не зробить, а все-ж обчислення значно спростить, ми поведемо це обчислення так:

[1]) Цю низку спроб роблено на незначному схилі, одно повторення вище від одного, при чому віддаль між ними була у кілька десятків сажнів.

V	V^2
23	529
11	121
15	225
—11	121
—11	121
—17	289
— 3	9
—15	225
24	574
37	1369
28	784
17	289
7	49
5	25
—16	256
—51	2601
29	841
35	1225
24	576
5	25
51	2601
44	1936
$\Sigma V = 231$	$\Sigma V^2 = 14793$

$$M = \frac{\Sigma V}{n} = \frac{231}{22} = \pm 10{,}5 \text{ п.}$$

$$\sigma = \pm\sqrt{\frac{\Sigma V^2}{n} - M^2} = \pm\sqrt{\frac{14793}{22} - 10{,}5^2} =$$

$$= \pm\sqrt{672{,}41 - 110{,}25} = \pm\sqrt{562{,}16} =$$

$$= \pm 23{,}7 \text{ п.}$$

$$m = \pm\frac{\sigma}{\sqrt{n}} = \pm\frac{23{,}7}{\sqrt{22}} = \pm\frac{23{,}7}{4{,}7} = \pm 5 \text{ п.}$$

$$M = 10{,}5 \pm 5 \text{ п., тоб-то } 2{,}1:1.$$

Вивід цієї многолітньої спроби залишається, таким чином, недоведенним. Проте, иншого результату при такому широкому варіюванні врожаїв ($v = \pm 24\%$!) ледві чи можна було й сподіватися. З цього становища є тільки один вихід: значне збільшення повторности спроби, як що тільки даний розмах варіювання не походить від якихось хиб у техніці переведення спроби.

7-й приклад. Підрахунок рослин гречки з короткостовпчатими квітками та з довгостовпчатими, який зробив Лебедянцев (на Шатилівській досвідній станції) 1910 року, дав (при числі рослин = 364) 45,1% довгостовпчатих і 54,9% короткостовпчатих, тоб-то різницю в 9,8% на користь короткостовпчатих. Яка є надійність цього результату?

Розвязуємо:

$$m = \pm\sqrt{\frac{\%p_1 \cdot \%p_2}{n}} = \pm\sqrt{\frac{45{,}1 \cdot 54{,}9}{364}} = \pm 2{,}7\%$$

$$\begin{array}{r} 45{,}1 \pm 2{,}7\% \\ -54{,}9 \pm 2{,}7 \text{ „} \\ \hline -\ 9{,}8 \pm 3{,}8\%, \text{ тоб-то } 2{,}6:1. \end{array}$$

Різницю, таким чином, не доведено, особливо, коли взяти на увагу величину m. Справді, межі коливань, одержаних в цьому дослідженні даних, можуть бути такими:

для довгостовпчатих — $45{,}1 \pm 3 \,.\, 2{,}7\%$, тоб-то від 37% до $53{,}2\%$
„ короткостовпчатих — $54{,}9 \pm 3 \,.\, 2{,}7\%$, „ „ 63% „ $46{,}8\%$,
инакше кажучи, можна чекати, що „справжні“ цих груп рослин можуть бути як-раз оберненими до результату даної спроби. І дійсно, як що ми звернемося до такого-ж самого підрахунку 1908 року, зробленого на матеріялі із значно більшою чисельністю ($n = 2953$), то ми одержимо ось що: було—

дов.-ст.	$52{,}7\%$		$52{,}7 \pm 0{,}9\%$
кор.-ст.	$47{,}3\%$	$m = \pm\sqrt{\dfrac{52{,}7 \,.\, 47{,}4}{2953}} = \pm\, 0{,}9\%$	$47{,}3 \pm 0{,}9\%$
різниця	$5{,}4\%$		$5{,}4 \pm 1{,}3\%$.
			тоб-то $4{,}2 : 1$.

На цей раз ми маємо доведену різницю на користь довгостовпчатих. Числова вартість різниці тепер є мало не вдоє менша ($5{,}4\%$), ніж у першому випадку ($9{,}8\%$), але через те, що число варіянтів є незрівняно більше ($2953 : 364$), то й надійність виводу є значно вища, вона майже втроє ($\sqrt{2953 : 364}$) більша і, для даної різниці, є цілком достатня.

8-й приклад. Коли нащадки мішанця, гетерозиготного в 2 факторах, менделюють, то виходить 4 групи зовні різних особин, тоб-то 4 альтернативи, в таких числових відношеннях:

$$94 : 33 : 29 : 12;\ \text{всіх особин}\ n = 168.$$

Теоретично тут треба було сподіватися 16 комбінацій (зазначмо $K = 16$) з такими відношеннями тих самих 4 груп:

$$9 : 3 : 3 : 1;\ \text{всіх комбінацій}\ K = 16.$$

Треба перевірити, чи сходяться дослідні дані з теоретичними.

Для цього відносять спершу, множачи на $\dfrac{n}{K}$, теоретичні числа до одержаного із спроби числа всіх варіянтів, щоб довідатися, яких чисельностей (зазначмо їх літерою q) окремих груп можна чекати: в нашому прикладі

спостережені числа (p): $94 : 33 : 29 : 12;\ n = 168$
теоретичні числа: $9 : 3 : 3 : 1;\ K = 16$

$$\text{сподівані числа } (q)\text{: } \left(9 \,.\, \frac{168}{16} =\right) 94{,}5 : \left(3 \,.\, \frac{168}{16} =\right) 31{,}5 : 31{,}5 :$$

$$: \left(1 \,.\, \frac{168}{16} =\right) 10{,}5;\ n = 168.$$

Далі, як що одну альтернативу прийняти за q, то другу (або суму всіх инших) можна виразити числом $(n-q)$, а тому можна написати (див. § 7, розд. IV):

$\sigma = \pm \frac{\sqrt{q(n-q)}}{n}$. Для середнього числа сподівань $\Big($тоб-то для $\frac{q}{n}$ і $\frac{n-q}{n}\Big)$ $m = \sigma : \sqrt{n}$, отже, для абсолютного сподіваного числа (тоб-то для q і $n-q$) буде $m_{абс.} = n.(\sigma : \sqrt{n}) = n.\frac{\sqrt{q(n-q)}}{n} : \sqrt{n}$ або, по де-яких спрощеннях,

$$m_{абс.} = \pm \sqrt{\frac{q(n-q)}{n}}.$$

Користуючися з цеї формули, одержимо: для групи з теоретичним числом варіянтів

9 сподіване $q = 94{,}5$; $n - q = 73{,}5$; $m_{абс.} = \pm \sqrt{\frac{94{,}5 \,.\, 73{,}5}{168}} = \pm 6{,}4$

3 „ $q = 31{,}5$; $n - q = 136{,}5$; $m_{абс.} = \pm \sqrt{\frac{31{,}5 . 136{,}5}{168}} = \pm 5{,}1$

1 „ $q = 10{,}5$; $n - q = 157{,}5$; $m_{абс.} = \pm \sqrt{\frac{10{,}5 . 157{,}5}{168}} = \pm 3{,}1$

Остаточний підсумок дає нам:

1. Одержано	2. Сподіваємося	Різниця (1—2)
94	$94{,}5 \pm 6{,}4$	$-0{,}5$
33	$31{,}5 \pm 5{,}1$	$+1{,}5$
29	$31{,}5 \pm 5{,}1$	$-2{,}5$
12	$10{,}5 \pm 3{,}1$	$+1{,}5$

Спостережені числа, таким чином, сходяться з теоретичними як найкраще.

Для розвязання тої самої задачі можна піти й навпаки, тоб-то віднести спостережені числа до теоретичних, помноживши спостережені на $\frac{K}{n}$; тоді матимемо:

абсолютно	94	33	29	12
відносно (до 16) . . .	8,952	3,143	2,763	1,142
теоретично	9	3	3	1

Далі, аналогічно з формулою $\sigma = \pm 100 \sqrt{\frac{p_0 p_1}{n}}\,\%$ (тоб-то до 100, або сотих) ми можемо написати

$$\sigma = K\frac{\sqrt{p_0 p_1}}{n} \text{ до } K, \text{ або } K\text{-тих.}$$

Зазначивши теоретичну чисельність одної альтернативи літерою N, матимемо для другої (або суми всіх инших) $K-N$ і тому σ можна виразити в K-тих частках (бо тут $p_0 = N$; $p_1 = K-N$; $n = K$) такою формулою:

$$\sigma_k = K\frac{\sqrt{N(K-N)}}{K} = \pm\sqrt{N(K-N)}.$$

Звідси відшукуване $m_k = \frac{\sigma_k}{\sqrt{n}} = \pm\sqrt{\frac{N(K-N)}{n}}$.

В нашому прикладі з цієї формули одержимо: для теоретичного числа

$$9 \quad m_k = \pm\sqrt{\frac{9\,.\,7}{168}} = \pm 0{,}612$$

$$3 \quad m_k = \pm\sqrt{\frac{3\,.\,13}{168}} = \pm 0{,}482$$

$$1 \quad m_k = \pm\sqrt{\frac{1\,.\,15}{168}} = \pm 0{,}299.$$

Остаточний підсумок:

1. Одержано		2. Сподіваємося	Різниця
абсолютно	відносно (до 16)	(до 16)	1 — 2.
94	8,952	$9 \pm 0{,}612$	$-0{,}048$
33	3,143	$3 \pm 0{,}482$	$+0{,}143$
29	2,763	$3 \pm 0{,}482$	$-0{,}237$
12	1,142	$1 \pm 0{,}299$	$+0{,}142$

Отже, різниця всюди є менша, ніж відповідна серединна похибка.

X. Покажчик точности досліду або спроби (P).

§ 1. Так, як і σ, m є число йменоване. Лише для альтернативного варіювання m є число відносне (приведене), бо обчисляють його у відсотках. Тому лише для альтернативного варіювання різні m можна порівнювати безпосередньо одне з одним; так у 7-ому прикладі попереднього розділу ми мали в одному досліді $m = \pm 2{,}7\%$, а в другому $m = \pm 0{,}9\%$ і можемо сказати, що у другому випадку точність досліду була втроє вища, як у першому. Отже, для альтернативного варіювання серединна похибка являє собою одночасно і покажчика точности досліду (або спроби).

Инакше стоїть справа при других рядах варіювання. Ми не можемо, напр., сказати безпосередньо, що дослід варіювання довжини колосу, який дав $M_1 \pm m_1 = 86,1 \pm 0,7$ *mm*, буде не такий точний, як дослід варіювання числа колосків, який дав $M_2 \pm m_2 = 17,4 \pm 0,2$ кол.

Так, як і для σ, тут треба брати приведене m, взявши M за 1 або, що звичайно є зручніше, за 100, тоб-то, виразивши m у відсотках од M. Таку відсоткову серединну похибку назвімо покажчиком точности досліду або спроби і зазначмо літерою P; отже, формула покажчика точности досліду буде така:

$$P = \frac{100m}{M}\ \%.$$

Для допіру наведених дослідів матимемо:

$$P_1 = \frac{100m_1}{M_1}\ \% = \pm\frac{100\ .\ 0,7}{86,1}\ \% = \pm 0,81\%$$

$$P_2 = \frac{100m_2}{M_2}\ \% = \pm\frac{100\ .\ 0,2}{17,4}\ \% = \pm 1,15\%$$

Отже, M_1 визначено з більшою точністю (з меншою відносною похибкою), ніж M_2.

Знаючи v, можна обчислити P з формули

$$P = \frac{v}{\sqrt{n}}\ \%,$$

$$\text{бо } m = \frac{\sigma}{\sqrt{n}};\ \text{ а }\ \frac{100\sigma}{M} = v.$$

Можемо сказати, таким чином, що варіяційний коефіцієнт (v) показує точність визначення середнього аритметичного (M) з одного варіянту і що із збільшенням їх числа (або із збільшенням повторности спроби) точність визначення M зростає пропорціонально до кореня квадратового з цього числа (або з повторности спроби). Напр., як що точність спроби без повторности (покажчик різноманітности дослідної ділянки) $v = \pm 8,2\%$, то при 4-ній повторності $P_{(4)} = \frac{v}{\sqrt{n}} =$
$= \pm \frac{8,2}{\sqrt{4}}\% = \pm 4,1\%$, тоб-то точність є вдвоє більша.

§ 2. Визначення точности спроби безпосередньо з даної σ (або v) є можливе тільки при великім n, бо сама σ завжди веде за собою похибку (m_σ), рівну:

$$m_\sigma = \frac{\sigma}{\sqrt{2n}}.$$

При великім n ця похибка основного відхилу є порівнюючи невелика, напр., для $n = 100$ $m_\sigma = \frac{1}{14}\sigma$, але при малій повторності спроби m_σ досягає досить значних розмірів. Так, при 2-ній повторності, що найчастіше буває в дослідній ділі, теоретично $m_\sigma = \frac{1}{2}\sigma$! Справді, через дуже малу кількість варіянтів, відхили σ від справжньої величини бувають тут значно більшими, а иноді й надмірними. Визначування P безпосередньо з даної σ в цих випадках є занадто неточні. З тих-же причин треба відкинути також і вживання звичайного способу порівнювати середні аритметичні з їх „m", одержаних безпосередньо із спроби.

Крім того, розвідочні засіви (часткові обліки врожаю) показали, чого можна було сподіватися й заздалегідь, що точність спроби не залишається постійною з року на рік, але міняє свою величину, зважаючи на життьові умови того чи иншого року. Отже, розвідочний засів, скільки-б разів його не повторювали, дає тільки загальні вказівки що до характеру варіювання родючости дослідного поля, робити-ж висновки з даних часткового обліку про точність наступних спроб ніяк не можна. Тому конче потрібно визначати точність спроби що-року, користуючись тільки з даних самої-ж спроби. Як-би ми мали діло з великим числом повторних спроб, з десятками повторень, то могли-б вказаного вище способу визначувати точність спроби, а також ужити звичайних правил для порівнювання середніх аритметичних за допомогою їх серединних похибок [1]). Але така велика повторність є виняткова в дослідній справі, і сподіватися її в майбутньому, як правила, теж не доводиться, бо переведення дуже великої повторности в польовім досліджуванні зустрічає значні технічні перешкоди. Отже, для дослідної справи склалося дуже важке становище, і треба було найти якийсь досить надійний вихід із нього, бо працювати, не знаючи точности спроби є, звичайно, річ неможлива.

[1]) Тільки в дуже рідких випадках, при надзвичайно вирівненій родючості поля, вже незначна повторність дає P, рівне 1—2%, і дозволяє, таким чином, безпосереднє користуватися з m.

Я найшов досить добрий спосіб визначати точність спроби при малій повторності, який можна назвати методою зведення. Вона полягає ось у чому:

Уявім собі поле, поділене на рівновеликі ділянки одної й тої самої форми. Всіх ділянок нехай буде N. Ми провадимо спробу, повторюючи її n разів; отже, досліджуємо польовою методою N/n питань (сортів, способів обробки то-що).

Урожай окремої ділянки, що вийшов-би на ній при розвідочнім засіві, зазначмо літерою V. При n повтореннях середнє аритметичне M окремого питання було-б тоді $M = \frac{\Sigma V}{n}$. Але, вносячи для даних n ділянок инший технічний засіб або инший сорт, ми, взагалі змінюємо врожай цих n ділянок; припустімо, що цю зміну врожаю виражає певний сучинник k, однаковий для кожної з n даних ділянок; тоді цей урожай можна записати як kV. Середнє аритметичне для цих ділянок також зміниться (збільшиться або зменшиться) в k разів:

$$\frac{\Sigma kV}{n} = kM.$$

В спробі ми завжди маємо kV і kM. Але з даної допіру формули видко, як позбутися k, що нам заважає: досить лише взяти kM за 1 або за 100, тоб-то одержати зведене kV. Як що такі зведені відсоткові kV зазначити літерою w, то

$$w = 100 . \frac{kV}{kM} \% = 100 . \frac{V}{M} \%.$$

Ми одержуємо зведені відсоткові варіянти такі-ж самі, які вийшли-б при розвідочнім засіві і, таким чином, можемо тепер визначати покажчика точности нашої спроби із значно більшою надійністю.

Проробивши аналогічні обчислення для всіх N/n питань, матимемо для визначення v і P вже не по n варіянтів, а по числу в N/n разів більшому, тоб-то по N.

Одержавши всі w, обчисляємо відшукуване $P_{(n)}$ (тоб-то P при n повтореннях) так:

Спершу обчислім різниці між w і відсотковим M ($=100\%$); зазначивши їх літерою d ($d = w - 100$), матимемо $v = \pm\sqrt{\frac{\Sigma d^2}{N}} \%$, звідки

$$P_{(n)} = \frac{v}{\sqrt{n}} = \pm\sqrt{\frac{\Sigma d^2}{Nn}} \%.$$

Як що повторність у окремих питань спроби є неоднакова і дорівнює, напр., в одних n_1, а в инших n_2, то P для них обчисляють або через v:

$$P_{(n_1)} = \frac{v}{\sqrt{n_1}}; \quad P_{(n_2)} = \frac{v}{\sqrt{n_2}},$$

або безпосередньо:

$$P_{(n_1)} = \pm \sqrt{\frac{\Sigma d^2}{Nn_1}}\,\%; \quad P_{(n_2)} = \pm \sqrt{\frac{\Sigma d^2}{Nn_2}}\,\%.$$

Знаючи $P_{(n)}$, легко обчислити для кожного одержаного із спроби середнього аритметичного (kM) його значно точнішу серединну похибку (km), бо [1])

$$km = \frac{kMP_{(n)}}{100}$$

1-ий приклад. В спробі на порівняння було 17 сортів озимої пшениці; спробу повторювано 4 рази. Отже $N/n = 17$; $n = 4$; звідки $N = N/n \,.\, n = 17.4 = 68$. Ці сорти при повтореннях (kV) виявили такий урожай у фунтах (див. табл. на стор. 50):

Звідси $\Sigma d^2 = 5879$ і $P_{(4)} = \pm \sqrt{\frac{\Sigma d^2}{nN}} = \pm \sqrt{\frac{5879}{4.68}} = \pm 4{,}7\%$.

Тепер можемо порівняти врожаї окремих сортів, обчисляючи „точні“ серединні похибки їх дослідних середніх аритметичних; а саме, для останнього сорту одержимо: $km = \frac{kMP}{100} = \pm \frac{27{,}06.4{,}7}{100} = \pm 1{,}27$ ф.; для сорту, що дав у спробі середній урожай $kM = 30{,}5$ ф., серединна похибка $km = \frac{30{,}5.4{,}7}{100} = \pm 1{,}43$ ф. Взагалі-ж, в даній спробі всі сорти з урожайністю від 30,5 ф. до 25,4 ф. треба віднести до одної й тої самої найвищої своєю врожайністю класи.

$$\begin{array}{r} 30{,}5 \pm 1{,}43 \text{ ф.} \\ -25{,}4 \pm 1{,}19 \text{ ,,} \\ \hline 5{,}1 \pm 1{,}86 \text{ ф.,} \end{array} \text{ тоб-то } 2{,}7:1.$$

Нижче іде сорт з урожаєм 24,6 ф.:

$$\begin{array}{r} 30{,}5 \pm 1{,}43 \text{ ф.} \\ 24{,}6 \pm 1{,}16 \text{ ,,} \\ \hline 5{,}9 \pm 1{,}66 \text{ ф.,} \end{array} \text{ тоб-то } 3{,}6:1,$$

отже, за межами найвищої класи.

[1]) $P = \frac{100m}{M}$; звідси $m = P : \frac{100}{M} = \frac{MP}{100}$.

kV	*w*	*d*	*d*²	*kV*	*w*	*d*	*d*²
28	103	+ 3	9	25	109	+ 9	81
24	88	— 12	144	20	86	— 14	196
29	107	— 7	49	26	113	+ 13	169
28	103	+ 3	9	21	91	— 9	81
kM = 27,25				*kM* = 23			
20,50	85	— 15	225	29	107	+ 7	49
21,25	88	— 12	144	21	77	— 23	529
29,75	123	+ 23	529	31,5	116	+ 10	256
25,50	105	+ 5	25	27	100	0	0
kM = 24,25				*kM* = 27,13			
27,5	102	+ 2	4	27,5	97	— 3	9
20,5	76	— 24	576	28	99	— 1	1
31,5	117	+ 17	289	25,5	90	— 10	100
28.5	106	+ 6	36	32,5	114	+ 14	196
kM = 27				*kM* = 28,4			
29	104	+ 4	16	21,5	97	— 3	9
26,5	95	— 5	25	23	104	+ 4	16
27,5	99	— 1	1	24.5	111	+ 11	121
28,5	102	+ 2	4	19,5	88	— 12	144
kM = 27,9				*kM* = 22.1			
26	102	+ 2	4	23,75	99	— 1	1
27	106	+ 6	36	24,75	103	+ 3	9
23,5	95	— 5	25	25	104	+ 4	16
25	97	— 3	9	23	95	— 5	25
kM = 25,4				*kM* = 24,1			
23,5	100	0	0	24	98	— 2	4
23	97	— 3	9	29	118	+ 18	324
24	101	+ 1	1	23	93	— 7	49
24	101	+ 1	1	22,5	91	— 8	64
kM = 23,6				*kM* = 24.6			
21	97	— 3	9	30,5	100	0	0
20,5	94	— 6	36	27,5	90	— 10	100
24,5	113	+ 13	169	32,5	107	+ 7	49
21	97	— 3	9	31,5	103	+ 3	9
kM = 21.75				*kM* = 30,5			
26,75	94	— 6	36	25,5	94	— 6	36
31,25	110	+ 10	100	27,5	101	+ 1	1
27	95	— 5	25	31,5	116	+ 16	256
28.75	101	+ 1	1	23.75	88	— 12	144
kM = 28,44				*kM* = 27.06			

kV	*w*	*d*	*d*²
26	96	— 4	16
24	89	— 11	121
30,25	112	+ 12	144
28	103	+ 3	9
kM = 27,06			

Взагалі-ж, при наявній точності $P_{(4)} = \pm 4{,}7\%$, сорти даної спроби можна за врожайністю поділити лише на дві класи:

$$\begin{array}{rl} \text{I.} & 30{,}5 \pm 1{,}43 \text{ ф.} \\ & -25{,}0 \pm 1{,}17 \text{ ,,} \\ \hline & 5{,}5 \pm 1.85 \text{ ф.} \\ & 3 : 1 \end{array} \qquad \begin{array}{rl} \text{II.} & 25{,}0 \pm 1{,}17 \text{ ф.} \\ & -20{,}5 \pm 0{,}96 \text{ ,,} \\ \hline & 4{,}5 \pm 1{,}5 \text{ ф.} \\ & 3 : 1 \end{array}$$

Цей розподіл зроблено так. За початок узято сорт з найвищою врожайністю (30,5 ф.) і для визначення нижньої межі цієї найвищої врожайністю класи д о б р а н о число (25 ф.) так, щоб різниця між ним і найвищим урожаєм була втроє більша за її серединну похибку. За верхню межу другої класи взято нижню межу першої класи (25 ф.) і до неї д о б р а н о число (20,5 ф.) так, щоб різниця між цими двома числами знову була втроє більша за свою серединну похибку.

Відшукуване M_2 можна обчислити і з формули. Ми шукаємо:

$$\begin{array}{r} M_1 \pm m_1 \\ -\ M_2 \pm m_2 \\ \hline M_1 - M_2 = 3\sqrt{m_1^2 + \left(\frac{M_2 P}{100}\right)^2}, \text{ звідки} \end{array}$$

$$M_2 = M_1 \pm 3m\underline{\sqrt{2 - 0{,}0009\ P^2}};\ \text{і}\ m_2 = 0{,}01\ M_2\ P.$$

$1 - 0{,}0009 P^2$

2-й п р и к л а д. У звіті Одеського дослідного поля за 1914 р. ми знайдемо з групи B (сорти кукурудзи) такі дані:

Весняна оранка		1-ше повтор.	2-ге повтор.	M
1) 2 в., рослини 1 арш. × 10 в.	Чинквантино	144,0	124,8	134,4
2) 4 » » » » » »	Чинквантино	125,0	133,0	129,0
3) 6 » » » » » »	Чинквантино	110,0	115,5	112,8
4) 2 » » 1 » × 1 арш.	Чинквантино	102,5	116,0	109,3
5) » » » 1½ » × 1 »	Чинквантино	128,5	119,0	123,8
6) » » » » » » »	Чинквантино	161,0	137,5	149,6
7) » » » » » » »	Лімінг . . .	149,8	159,8	154,8
8) » » » » » » »	Кавказька .	144,7	156,6	150,6
9) » » » » » » »	Зуб Рейдів .	131,8	152,8	142,3
10) » » » » » » »	Король Філіп біл. . . .	149,3	161,5	155,4
11) » » » » » » »	Грушівська .	217,1	207,8	212,5
12) » » » » » » »	Функ 90-ден.	187,0	237,1	212,1
13) » » » » » » »	Угорська . .	196,4	198,4	197,4

Завваживши, що в данім випадку, тоб-то при подвійній повторності обидва варіянти того самого питання матимуть однакові d, обчислення $P_{(2)}$ записуємо так:

	d	d^2	$2d^2$
1)	7,1	50,41	100,82
2)	3,1	9,61	19,22
3)	2,4	5,76	11,52
4)	6,1	37,21	74,42
5)	3,8	14,44	28,88
6)	7,6	57,76	115,52
7)	3,2	10,24	20,48
8)	4,0	16,00	32,00
9)	7,4	54,76	109,52
10)	3,9	15,21	30,42
11)	2,2	4,84	9,68
12)	11,8	139,24	278,48
13)	0,5	0,25	0,50
			$\Sigma d^2 = 831,46$

$$P_{(2)} = \pm\sqrt{\frac{\Sigma d^2}{nN}} = \pm\sqrt{\frac{831,46}{2.26}} = \pm 4\%.$$

Перше ніж робити висновки, звернім увагу на те, що повторність даної спроби є найменша, яка може бути, і число всіх варіянтів є також невелике, тому є можливі великі коливання серединних похибок окремих M і, щоб довести різниці між двома M, є потрібно, щоб m було хоча-б у 3,5 рази менше за саму різницю. Розглядувана спроба є з цієї сторони особливо повчаюча. А саме, в ній питання 5-е і 6-е є зовсім ідентичні своїми умовами: той самий сорт, та сама глибина оранки, та сама рясність засіву,—проте середні вийшли різко несхожі: 123,8 п. і 149,6 п., тоб-то дослідна різниця між середніми досягла 25,8 п., дарма, що сама постава спроби мусіла-б дати однакові середні. Виходячи з обчисленого показчика точности даної спроби, одержимо для цих двох середніх таке:

$$\begin{array}{r} 149,6 \pm 6 \text{ п.} \\ -\,123,8 \pm 5 \;\text{„} \\ \hline 25,8 \pm 7,8 \text{ п., тоб-то } 3,3:1. \end{array}$$

Отже, з огляду на зроблене допіру завваження, різницю не можна визнати доведеною (як воно й є справді).

Після всього сказаного, можна зробити із даної спроби такі основні виводи.

По-перше. До найменш урожайних сортів в данім році треба віднести: чинквантино, кавказьку і Рейдів зуб, бо $m_{\text{різн.}}$ навіть поміж крайніми з цих середніх тільки в 3,4 рази менше за різницю між ними, що при 2-кратній повторності неможна вважати за цілком достатнє:

$$\begin{array}{r} 150,6 \pm 6,2 \text{ п.} \\ -\,123,8 \pm 5,0 \;\text{„} \\ \hline 26,8 \pm 8,0 \text{ п., тоб-то } 3,4:1. \end{array}$$

Далі. Найбільш урожайними були сорти: грушівська, Функ 90-денний і угорська, бо найвищий урожай, при даній точності спроби, треба рахувати від 212,5 п. до 174 п.:

$$\begin{array}{r} 212{,}5 \pm 8{,}5 \text{ п.} \\ -174{,}0 \pm 7{,}0 \text{ „} \\ \hline 38{,}5 \pm 11{,}0 \text{ п., тоб-то } 3{,}5:1. \end{array}$$

Останні два сорти: Лимінг і Король Філіп білий становлять проміжну группу.

По-друге. Вплив заглиблення оранки на пониження врожайности, як це було-б виведено на підставі дослідної різниці між відповідними середніми (питання 1, 2, 3-є), в даній поставі спроби неможна вважати за доведений, бо різниця навіть між крайніми з цих середніх дорівнює:

$$\begin{array}{r} 134{,}4 \pm 5{,}4 \text{ п.} \\ -112{,}8 \pm 4{,}5 \text{ „} \\ \hline 21{,}6 \pm 7{,}0 \text{ п., тоб-то } 3{,}1:1. \end{array}$$

Отже, двократна повторність в умовах даного поля є замала, щоб виявити, як впливає глибина оранки на врожайність.

XI. Форми варіяційних кривих.

Покажчик асиметрії *(S)* і покажчик ексцесу *(E)*.

§ 1. Розглянуті в попередніх розділах варіяційні заходи характеризують в достатній мірі симетричну біноміяльну криву, схожу з „кривою похибок“, що служить її границею. Варіяція з найбільшим числом варіянтів збігається тут із середнім аритметичним, звідки в обидві сторони крива спадає цілком аналогічно, так що права її половина неначе являє собою дзеркальне відбиття лівої. Такий характер варіювання буває, напр., в тому разі, коли величина, що варіює, може змінятися в обидві сторони від *M* безмежно або, принаймні, дуже далеко, коли, напр., нуль не являє собою границю відхилу варіяцій від *M,* але вони можуть відходити і лівіше за його, тоб-то ставати від'ємними. Другою умовою являється тут незалежність факторів, що викликають варіювання, відсутність однобічної дії де-яких з них, то-що.

Але як-раз ті величини, що з ними доводиться мати діло біологові, отже й агрономові, з самої природи своєї дуже рідко справджують указані умови. Так, напр., здебільшого варіювання реальних величин має можливість управо від *M* сягати дуже далеко (теоретично безмежно), вліво-ж лише на обмежену віддаль, що її теоретично визначають положенням *M* відносно нуля, практично-ж вона обмежена в більшій мірі, ніж управо від *M*. Иноді така границя варіювання міститься вправо від *M*. Багато факторів росту й розвивання діють з неоднаковою силою на особини різної величини, то-що.

Тому криві варіювання реальних величин здебільшого повинні мати иншу форму, ніж „криві похибки“. Математичне дослідження показує, що характерною особливістю таких кривих буде їх асиметричність, або перекошеність.

Рис. 5 дає нам приклад асиметричної (перекошеної) кривої. Точка *M* визначає положення середнього аритметичного (середньої варіяції), а точка M_0—положення найчастішої варіяції, так званої моди (тоб-то: «наймоднішої» варіяції). У даної кривої *M* лежить правіше, як мода; такі криві ми називатимемо додатньо-асиметричними.

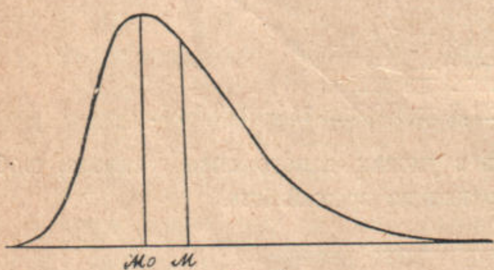


Рис. 5.

Криві з розподілом варіянтів, противним указаному, називатимемо від'ємно-асиметричними.

Однак, часто буває більш-менш ризиковано робити висновки безпосередньо з варіяційного ряду про наявність асиметричности у кривої, бо з огляду на довільне вибирання меж для клас, нам може часом здатися, що крива є перекошена, чого справді нема. Так, напр., коли матеріял, що з нього вийшов був ряд:

9	10	11	12	13	14	15	16	17	*mm*
	5	67	466	761	201	15	5	1	

очевидячки асиметричний, знову перерозташовано було по класах з иншими межами, то вийшов без сумніву симетричний ряд:

8,75	9,75	10,75	11,75	12,75	13,75	14,75	15,75	16,75	*mm*
	2	43	314	**809**	316	30	6	2	

Так чи инакше, треба мати певну варіяційну міру, що давала-б змогу визначати степінь перекошености варіяційного ряду незалежно від того, які варіяції взято за основу для будування розглядуваного ряду, тоб-то обчисляти безпосередньо з його данних.

Для відшукуваної міри в основу взято *середній кубічний відхил*, тоб-то третій степінь середнього із суми відхилів [1]. З цього ми маємо подвійну користь: по-перше, крайні варіяції, що мають особливу вагу відносно асиметричности кривої, дістають порівнюючи ще більшого значіння; по-друге, при наявності перекошености в розподілі варіянтів, одна галузь кривої дає суму третіх степенів більшу, ніж друга, а тому, що третій степінь зберегає знак відхилу, то різниця цих сум прямо вказує на те, з якою саме асиметрією, додатньою чи від'ємною, ми маємо діло. Як-що крива є симетрична, то сума додатніх відхилів дорівнює сумі від'ємних, але третій степінь зберегає ці знаки,— тому сума всіх відхилів дорівнює нулеві. Так, напр., для симетричного ряду:

1	2	3	4	5	6	7
1	6	15	20	15	6	1

$$\text{при } M = 4, \text{ маємо } \left.\begin{array}{l} +1^3 . 15 + 2^3 . 6 + 3^3 . 1 \\ -1^3 . 15 - 2^3 . 6 - 3^3 . 1 \end{array}\right\} \Sigma p\alpha^3 = 0$$

Для аналогічного, але перекошеного ряду

2	3	4	5	6	7
5	18	21	14	4	2

$$\text{при } M = 4, \text{ маємо } \left.\begin{array}{ll} +1^3 . 14 + 2^3 . 4 + 3^3 . 2 & = +100 \\ -1^3 . 18 - 2^3 . 5 & = -58 \end{array}\right| \Sigma p\alpha^3 = +42,$$

тоб-то додатню перекошеність кривої.

Виявилось, що найзручніше за *покажчика асиметричности* варіяційного ряду взяти *відношення середнього кубічного відхилу до кубу основного відхилу*, тоб-то, як-що покажчика перекошености зазначити літерою S, то

$$S = \frac{\Sigma p\alpha^3}{n} : \sigma^3.$$

Для практичних потреб цю формулу є вигідніше перетворити в иншу, замінивши α на a.

Для цього нагадаймо, що $\alpha + b = a$, і запишім

$$\frac{\Sigma p(\alpha + b)^3}{n} = \frac{\Sigma pa^3}{n}, \text{ звідки}$$

$$\frac{\Sigma p\alpha^3}{n} + \frac{3b\Sigma p\alpha^2}{n} + \frac{3b^2\Sigma p\alpha}{n} + b^3 = \frac{\Sigma pa^3}{n}.$$

---

[1]) При дуже точних дослідженнях, напр., в астрономії, вдаються і до вищих степенів (звичайно, непаристих).

Тут третій член зникає, бо $\Sigma p\alpha = 0$. Далі, $\frac{\Sigma p\alpha^2}{n} = \frac{\Sigma pa^2}{n} - b^2$; тому $\frac{3b\Sigma p\alpha^2}{n} = \frac{3b\Sigma pa^2}{n} - 3b^3$. Отже весь наш вираз перетворюється в такий:

$$\frac{\Sigma p\alpha^3}{n} + \frac{3b\Sigma pa^2}{n} - 3b^3 + b^3 = \frac{\Sigma pa^3}{n},$$

звідки $\frac{\Sigma p\alpha^3}{n} = \frac{\Sigma pa^3}{n} - \frac{3b\Sigma pa^2}{n} + 2b^3$. Отже,

$$S = \left(\frac{\Sigma pa^3}{n} - \frac{3b\Sigma pa^2}{n} + 2b^3\right) : \sigma^3.$$

Для класового ряду з проміжками, не рівними 1, зручніше прийняти їх за одиницю і заступити в такому разі у формулі для S числа a, b і σ відповідно числами; a', β і s; тоді вона матиме вигляд:

$$S = \left(\frac{\Sigma pa'^3}{n} - \frac{3\beta\Sigma pa'^2}{n} + 2\beta^3\right) : s^3.$$

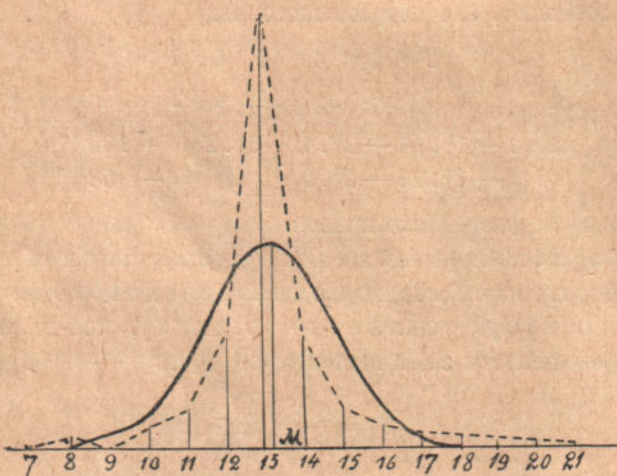


Рис. 6.

§ 2. Приклад на обчислення з цієї формули ми знайдемо трохи нижче, після того, як познайомимося з покажчиком ексцесу кривої. Так називають явище висування вершка кривої поза межі «нормального» положення, як це видно, напр., на рис. 6, де дано криву знайомого вже нам ворівування числа скрайніх квіток у цвітостанах Chrysanthemum segetum (§ 3, розділ III). Для таких кривих є характерне переступання варіяцій за межі $\pm 3\sigma$, хоча m і в цих випадках являє собою досить добру міру для визначення надійности M.

Для визначення ексцесу кривої користуються з середнього відхилу четвертого степеня $\left(\frac{\Sigma p\alpha^4}{n}\right)$ і саме з того факту, що поміж цим і квадратовим відхилами для ідеальної біноміяльної кривої існує певна залежність:

$$\frac{\Sigma p\alpha^4}{n} : \frac{\Sigma p\alpha^2}{n} = 3\sigma^2,$$

Звідки

$$\frac{\Sigma p\alpha^4}{n} : \left(\frac{\Sigma p\alpha^2}{n} \cdot \sigma^2\right) = \frac{\Sigma p\alpha^4}{n} : \sigma^4 = 3.$$

В разі ексцесу ця рівність порушується, і сама різниця між правою та лівою частиною наведеного рівнання дасть, таким чином, величину і знак покажчика ексцесу. Зазначивши його літерою E, матимемо:

$$E = \frac{\Sigma p\alpha^4}{n} : \sigma^4 - 3.$$

Заступім і тут для зручности обчислення число α числом a. Але $\alpha + b = a$, отже

$$\frac{\Sigma p(\alpha + b)^4}{n} = \frac{\Sigma pa^4}{n},$$

або

$$\frac{\Sigma p\alpha^4}{n} + \frac{4b\Sigma p\alpha^3}{n} + \frac{6b^2\Sigma p\alpha^2}{n} + \frac{4b^3\Sigma p\alpha}{n} + b^4 = \frac{\Sigma pa^4}{n}.$$

Далі:

$$\frac{4b\Sigma p\alpha^3}{n} = 4b\left(\frac{\Sigma pa^3}{n} - \frac{3b\Sigma pa^2}{n} + 2b^3\right);$$

$$\frac{6b^2\Sigma p\alpha^2}{n} = 6b^2\left(\frac{\Sigma pa^2}{n} - b^2\right); \qquad \frac{4b^3\Sigma p\alpha}{n} = 0.$$

Тому:

$$\frac{\Sigma p\alpha^4}{n} + 4b\left(\frac{\Sigma pa^3}{n} - \frac{3b\Sigma pa^2}{n} + 2b^3\right) + 6b^2\left(\frac{\Sigma pa^2}{n} - b^2\right) + b^4 = \frac{\Sigma pa^4}{n},$$

або

$$\frac{\Sigma p\alpha^4}{n} = \frac{\Sigma pa^4}{n} - \frac{4b\Sigma pa^3}{n} + \frac{6b^2\Sigma pa^2}{n} - 3b^4.$$

Отже,

$$E = \left(\frac{\Sigma pa^4}{n} - \frac{4b\Sigma pa^3}{n} + \frac{6b^2\Sigma pa^2}{n} - 3b^4\right) : \sigma^4 - 3.$$

В разі варіяцій з класовою проміжкою, нерівною 1, слід брати її за 1 і користуватися з формули такої:

$$E = \left(\frac{\Sigma pa'^4}{n} - \frac{4\beta\Sigma pa'^3}{n} + \frac{6\beta^2\Sigma pa'^2}{n} - 3\beta^4\right) : s^4 - 3.$$

Приклад. Частковий облік пшеничного поля в Офросимівці, на Київщині, який пробив Юровський, дав для 480 ділянок, у 2×6 кв. с. кожна, наступний ряд (розд. XIII, 1-ий приклад):

235	255	275	295	315	335	355	375	395	415	435 пуд.
1	6	55	183	156	62	11	4	0	2	

Визначити S і E.

Прийнявши $A = 325$ і проміжку класи за 1, обчисляємо:

	$\pm a'$	1	2	3	4	5
	$+p$	62	11	4	0	2
	$-p$	183	55	6	1	
4. різниці $+p$ і $-p$		-121	-44	-2	-1	$+2$
5.	$\pm a'$	1	2	3	4	5
6.	$\pm a'^3$	1	8	27	64	125
7. суми $+p$ і $-p$		245	66	10	1	2
8.	a'^2	1	4	9	16	25
9.	a'^4	1	16	81	256	625

З рядків 4. і 5. маємо:

$$\Sigma pa' = -121.1 - 44.2 - 2.3 - 1.4 + 2.5 = -209.$$

З рядків 4. і 6. маємо:

$$\Sigma pa'^3 = -121.1 - 44.8 - 2.27 - 1.64 + 2.125 = -341.$$

З рядків 7. і 8. маємо:

$$\Sigma pa'^2 = 245.1 + 66.4 + 10.9 + 1.16 + 2.25 = 665.$$

З рядків 7. і 9. маємо:

$$\Sigma pa'^4 = 245.1 + 66.16 + 10.81 + 1.256 + 2.625 = 3617.$$

Поділяючи на $n = 480$, одержимо:

$$\frac{\Sigma pa'}{n} = \frac{-209}{480} = -0,4354 = \beta; \quad \frac{\Sigma pa'^2}{n} = \frac{665}{480} = 1,3854$$

$$\frac{\Sigma pa'^3}{n} = \frac{-341}{480} = -0,7104; \quad \frac{\Sigma pa'^4}{n} = \frac{3617}{480} = 7,5354.$$

Звідки:

$$\beta = -0,4354; \quad \beta^2 = 0,1896; \quad \beta^3 = -0,0825; \quad \beta^4 = 0,0359$$

$$s = \sqrt{\frac{\Sigma pa'^2}{n} - \beta^2} = \sqrt{1,3854 - 0,1896} = 1,0935; \quad s^2 = 1,1957;$$

$$s^3 = 1,3075; \quad s^4 = 1,4298.$$

Отже:

$$S = \left(\frac{\Sigma pa'^3}{n} - \frac{3\beta\Sigma pa'^2}{n} + 2\beta^3\right) : s^3 =$$

$$= [-0{,}7104 - 3.(-0{,}4354).1{,}3854 + 2.(-0{,}0825)] : 1{,}3075 = +0{,}71.$$

$$E = \left(\frac{\Sigma pa'^4}{n} - \frac{4\beta\Sigma pa'^3}{n} + \frac{6\beta^2\Sigma pa'^2}{n} - 3\beta^4\right) : s^4 - 3 =$$

$$= [7{,}5354 - 4.(-0{,}4354).(-0{,}7104) + 6.0{,}1896.1{,}3854 - 3.0{,}0359] :$$
$$: 1{,}4298 - 3 = +2{,}43.$$

Криві з S від 0,25 до 0,50 вважається за помірно-перекошені і з E до 0,4 — за мало-ексцесивні, отже в нашому прикладі ми маємо доволі перекошений і доволі ексцесивний варіяційний ряд.

§ 3. Із збільшенням перекошености відповідна сторона кривої стає чим раз крутішою і в крайніх випадках може вийти одостороння крива. Напр.,

0	5	10	15 % череззерниці
261	39	1	

Тут $M = 3{,}18\%$; $\sigma = \pm 1{,}76\%$ і $S = +2{,}34$.

Коли ексцес є відʼємний, то вершок кривої неначе притиснено. Із збільшенням E цей притиснений вершок стає втисненим (угнутим), і ми одержуємо 2-верхову криву. Так, напр., варіювання задніх „кліщів“ щипавки дало за Бейтсоновими вимірюваннями такий ряд:

2,75	3,25	3,75	4,25	4,75	5,25	5,75	6,25	6,75	7,25	7,75	8,25	8,75	9,25
64	**125**	52	7	12	24	42	42	**90**	68	44	8	6	

Тут $E = 1{,}56$.

При $E = -2$ двоверхова крива розпадається на дві одноверхових.

Трапляються, правда дуже рідко, криві, що значно відрізняються від наведених вище, напр., такі, що своєю формою нагадують літеру V, тоб-то з верхами по обох кінцях кривої, та инші. Щоб точно їх характеризувати, треба вживати особливих способів, що їх розробив Пірсон з його школою, але це вже виходить поза рямці нашого „елементарного підручника“.

XII. Коефіцієнт кореляції, або супряжености (r).

§ 1. Досі ми цікавилися варіюванням окремих ознак будь-якого матеріялу, напр., довжини колосу, числа колосків у колосі, ваги зернин в колосі то-що. Ми вивчали варіювання ознак, кожної нарізно, не зіставляючи їх одна з одною. Проте часто з'являється потреба такого зіставлення. Так, напр., вивчаючи варіюваня ваги зернин вівса і кількости товщу в них, ми помічаємо, що в даного сорту дорідніше зерно, загалом, є бідніше на товщ; однак, безпосередньо виходячи із спостережень, дати певну відповідь на питання не можна, бо і вага зернин і кількість жирів у них варіюють; отже при звичайному перегляді матеріялу може повстати фальшиве враження, ніби між вагою зернин і кількістю товщу є певний звязок і вони являють собою ознаки супряжені (корелятивні), тоб-то певній зміні одної величини відповідає певна зміна і другої.

В цих випадках повстає потреба вміти підпорядковувати окремі ознаки одну одній так, щоб ясно було видно, чи є між ними якийсь звязок, инакше — супряженість, або кореляція, та щоб величину цеї супряжености можна було виразити означеним і зручним у користуванні числом.

Припустім, що ми зіставляємо дві ознаки якогось матеріялу, що варіює: одну з них зазначмо X, а другу — Y. Зміривши ці ознаки по окремих варіянтах, ми дістали, напр., такі дані:

Варіянт №	1	2	3	4	n	
Величина X	20	15	30	25	10	mm.
„ Y	4	3	8	3	4	гр.

Відзначмо тепер класи для X, напр., 0 — 5 — 10 — 15 . . . . і для Y, напр., 0 — 3 — 6 — 9 . . . . і, вибравши всі варіянти, що стосуються до одної й тої самої класи групи X, розподілім їх у варіяційний ряд за класами Y-ка. Напр., у варіянтів, що стосуються до класи 5—10 мм групи X, Y дає, припустімо, такий ряд:

класи Y . . . . (гр.)	0—3	3—6	6—9	9—12	12—15	15—18	18—21	21—24	24—27	27—30	30—33
p варіянтів з X від 5 до 10 мм.				1	2	3	2	1			

Ми чинимо так з усіма X, від класи до класи, і одержуємо стільки рядів для Y, скільки клас має X. Можна звести ці ряди в одну загальну таблицю, або „мережу“, як що ряди Y-ів написати один під одним за тою чергою, як ідуть класи X-а; напр., в нашім теоретичнім випадку може вийти така мережа (див. табл. на стор. 61):

З цеї мережі відразу видко, як ідуть ряди Y що до X, і навпаки, прямовисні ряди в той самий час показують, як розташовано ряди X-ів що до клас Y-а. Взявши суми чисел (p) варіянтів по відповідних рядах

Класи Y

Класи X	0–3	3–6	6–9	9–12	12–15	15–18	18–21	21–24	24–27	27–30	30–33	Σp_y	M_y
5						1						1	16,5
10				1	2	3	2	1				9	16.5
15			1	3	6	9	6	3	1			29	16,5
20		1	3	6	13	14	13	6	3	1		60	16,5
25		2	6	13	17	19	17	13	6	2		95	16,5
30	1	3	9	14	19	20	19	14	9	3	1	112	16,5
35		2	6	13	17	19	17	13	6	2		95	16,5
40		1	3	6	13	14	13	6	3	1		60	16,5
45			1	3	6	9	6	3	1			29	16,5
50				1	2	3	2	1				9	16,5
55						1						1	16,5
Σp_x	1	9	29	60	95	112	95	60	29	9	1	$n=500$	
M_x	27,5	27,5	27,5	27.5	27,5	27,5	27,5	27,5	27,5	27,5	27,5		

одержмио Σp для окремих рядів; нарешті, сума всіх окремих Σp дасть число всіх варіянтів $= n$.

Уже безпосередній розгляд цеї мережі досить яскраво говорить за те, що в рядах Y варіянти розташовано по класах зовсім аналогічно, без ніякої залежности від того, до якої саме класи X стосується кожний даний ряд. Але особливо це помітно, коли розглядати середні аритметичні окремих рядів: для всіх клас X ряди Y мають одне й те саме $M_y = 16{,}5$ і, навпаки, для всіх клас Y ряди X мають одне й те саме $M_x = 27{,}5$. Отже, дарма, що X наростає, M_y залишається незмінним: між X і Y, таким чином, нема жадної кореляції.

§ 2. Дані розглянутої кореляційної мережі з'образімо тепер графічно, бо це полегчить нам дальшу роботу.

Візьмімо прямокутну систему координат. Поземна вісь представить нам класи X, а прямовисна — класи Y. Відзначмо точками положення M_x і M_y; тоді ми дістанемо для розглянутого прикладу, що всі M_x лежать на осі X, всі M_y на осі Y (рис. 7), як-що точку перетину простих для M_x і M_y взяти за початок координат.

Коли нема супряжености і залежности між ознаками, що ми їх зіставляємо *одну з одною*, то середні аритметичні окремих рядів Y, як це видко з графіку, розміщуються на осі Y-ів, а середні величини рядів X — на осі X-ів, так що кут між простими M_x і M_y є рівним 90°.

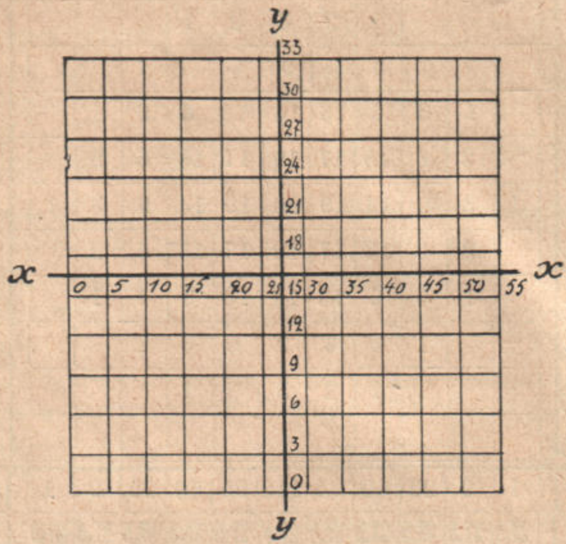


Рис. 7.

§ 3. Уявім собі тепер такий випадок, коли ознаки X і Y так звязані одна з одною, що зміна одної з них іде поруч з цілком аналогічною зміною другої: зміна X на 1 класову проміжку веде за собою змінку Y також на 1 класову проміжку. Знаючи в цім випадку X якогось варіянту, ми можемо відразу-ж, без міряння, сказати, яка буде величина Y цього варіянту. Такий випадок був-би, напр., з наступних даних:

Варіянт №	1	2	3	4	5	6 . . . .
Величина X	2,5	7,5	12,5	17,5	7,5	12,5 мм. і т. д.
„ Y	1,5	4,5	7,5	10,5	4,5	7,5 гр. і т. д.

Тут кожна зміна X на 5 мм. іде поруч із зміною Y на 3 гр. Це буде випадок п о в н о ї супряжености, і кореляційна мережа матиме такий вигляд (див. табл. на стор. 63):

Як що M_y, що виражають зміну, чи то р е г р е с і ю Y по X, з'образити графічно (роблячи так, як у § 2), то вийде такий графік (рис. 8).

Проста CC, поведена через точки M_y, показує, як міняється M_y, переходячи з одної X-ої класи у другу: це буде проста регресії Y по X.

	0	3	6	9	12	15	18	21	24	27	30 33	Σ_p	M_y
0	1											1	1,5
5		9										9	4,5
10			29									29	7,5
15				60								60	10,5
20					95							95	13,5
25						112						112	16,5
30							95					95	19,5
35								60				60	22,5
40									29			29	25,5
45										9		9	28,5
50 55											1	1	31,5
Σ_p	1	9	29	60	95	112	95	60	29	9	1	$n=500$	
M_x	2,5	7,5	12,5	17,5	22,5	27,5	32,5	37,5	42,5	47,5	52,5		

Коли таким самим способом збудувати просту регресії X по Y

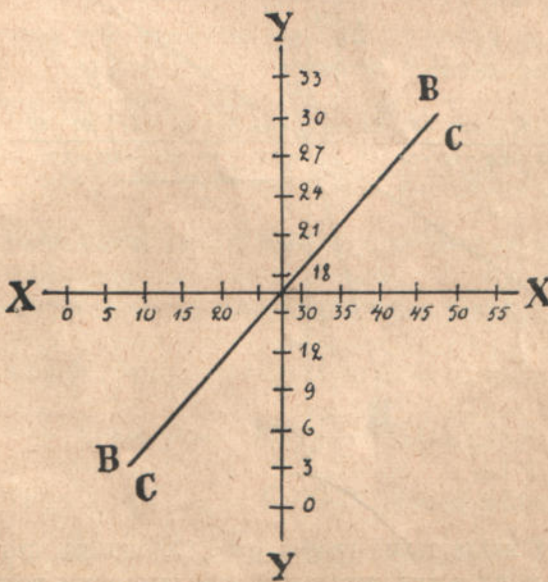


Рис. 8.

(BB), то виявиться, що в розглядуванім випадку ця проста перейде через ті самі точки, як і проста CC; отже, в разі повної кореляції проста регресії X по Y зливається з простою регресії Y по X.

§ 4. Уявім собі тепер, що M_y дано такі вартості:

відхили (x)		—25	—20	—15	—10	—5	M	+5	+10	+15	+20	+25
класи X	0	5	10	15	20	25	30	35	40	45	50	55 мм.
відповід. M_y	9	10,5	12	13,5	15	16,5	18	19,5	21	22,5	24	гр.,

а M_x такі:

відхили (y)		—15	—12	—9	—6	—3	M	+3	+6	+9	+12	+15
класи Y	0	3	6	9	12	15	18	21	24	27	30	33 гр.
відповідн. M_x	22,5	23,5	24,5	25,5	26,5	27,5	28,5	29,5	30,5	31,5	32,5	мм.

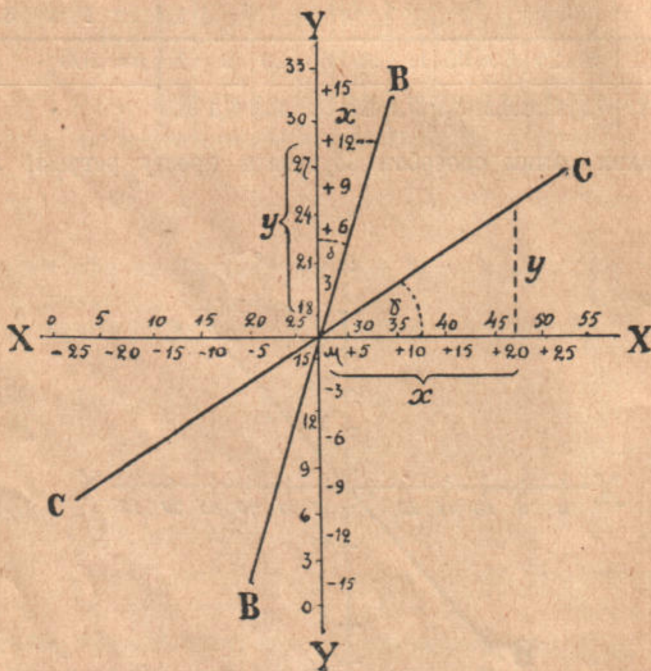


Рис. 9.

Із цих даних вийде такий графік (рис. 9), що на ньому точка M з'ображає положення точки загального середнього для всіх X і, разом з тим, точки загального середнього для всіх Y, бо ці обидві точки зливаються; від M, як від початку координат, відкладено відхили x

клас X праворуч (+) і ліворуч (—), а також відхили y клас Y вгору (+) і вниз (—).

M_y розташуються на простій CC (це є проста регресії Y по X, або, що є те саме, y по x), а M_x—на BB (це є проста регресії x по y).

З числових даних цього прикладу видко, що, при зміні x на класову проміжку, y зміняється, пересічно, лише на 0,5 своєї класової проміжки, і що, при зміні y на класову проміжку, x зміняється лише на 0,2 своєї класової проміжки.

Величину зміни y по x, вимірювану відношенням $y:x$[1]), звуть коефіцієнтом регресії y по x; зазначивши його $R_{\frac{y}{x}}$, матимемо в нашому прикладі: $R_{\frac{y}{x}} = \frac{y}{x} = \frac{1{,}5 \text{ гр.}}{5 \text{ мм.}} = \frac{3 \text{ гр.}}{10 \text{ мм.}} = 0{,}3$ гр. на кожний мм.

Аналогічно, величину зміни x по y, вимірювану відношенням $x:y$, звуть коефіцієнтом регресії x по y; зазначивши його $R_{\frac{x}{y}}$, одержимо для нашого прикладу: $R_{\frac{x}{y}} = \frac{x}{y} = \frac{1 \text{ мм.}}{3 \text{ гр.}} = \frac{2 \text{ мм.}}{6 \text{ гр.}} = \ldots . = 0{,}33$ мм. на кожний гр.

§ 5. Зазначмо кут, утворений простою регресії y по x (CC) з поземною віссю, літерою γ, а кут, утворений простою регресії x по y (BB) з прямовисною віссю, літерою δ. Тоді:

$$R_{\frac{y}{x}} = \frac{y}{x} = tg\gamma \text{ і } \quad R_{\frac{x}{y}} = \frac{x}{y} = tg\delta.$$

Ці залежності можна використати, щоб добрати дуже зручну міру величини супряжености двох ознак. А саме, таку міру являє собою середнє геометричне із коефіцієнтів регресії, тоб-то корінь квадратовий з їх добутка. Його звуть коефіцієнтом кореляції і зазначають літерою r. Отже,

$$r = \sqrt{R_{\frac{y}{x}} \cdot R_{\frac{x}{y}}} = \sqrt{tg\gamma \cdot tg\delta.}$$

Коефіцієнт кореляції, як видко з попереднього, є величина абстрактна (напр., в наведенім вище випадку $R_{\frac{y}{x}} \cdot R_{\frac{x}{y}} = \frac{1{,}5 \text{ гр.}}{5 \text{ мм.}}$.

[1]) Замісць абсолютних вартостей Y і X тут і далі взято їх відхили від середнього аритметичного, тоб-то y і x, що є значно зручніше.

$.\ \frac{1\ \text{мм.}}{3\ \text{гр.}} = \frac{1\ \text{мм.}}{5\ \text{мм.}} . \frac{1{,}5\ \text{гр.}}{3\ \text{гр.}}$, тоб-то виходить пара таких відношень, що в кожному з них обидва члени є одної назви).

Ми вже знаємо, що, коли супряжености нема, то прості (CC і BB) регресії зливаються з відповідними осями координат. Тоді $tg\gamma = 0$ $tg\delta = 0$, а тому і $r = \sqrt{tg\gamma \,.\, tg\delta} = \sqrt{0\,.\,0} = 0$.

З другого боку, за повної кореляції прості (CC і BB) регресії зливаються одна з одною, так що $\gamma + \delta = 90°$, отже $\delta = (90 - \gamma)^0$. Тому тут $r = \sqrt{tg\gamma \,.\, tg\delta} = \sqrt{tg\gamma \,.\, tg(90^0 - \gamma)} = \sqrt{tg\gamma \,.\, ctg\gamma} = 1$.

Таким чином, коли нема супряжености, то $r = 0$, а коли є повна супряженість, то $r = 1$, взагалі-ж r може коливатися від 0 до 1: від 0 до $+1$ і від 0 до -1, бо кореляція може бути й від'ємна, обернена, коли із збільшенням одної ознаки (напр. y) друга (x) зменшується; в цьому разі лінії регресії підносилися-б не з лівого боку у правий, але з правого боку в лівий.

§ 6. З причин, за які мова буде нижче (§ 10), на практиці дуже клопотливо користуватися з безпосередніх даних кореляційної мережі для обчислення R з tg, щоб потім з них визначити r; доводиться, навпаки, спершу обчисляти r з формули, куди входять відхили варіяцій (α) і основні відхили, як x-ого ряду (σ_x), так і y-ого (σ_y); тоді вже з r легко обчислити R.

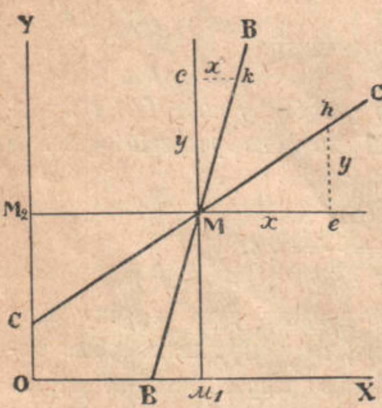


Рис. 10.

Перейдімо тепер до виводу цеї формули.

Ми поки що спинимося на теоретичнім випадку, коли всі M_x і M_y лежать на відповідних простих лініях.

Припустімо (рис. 10), що M_2 є середнє аритметичне для всіх Y, і нехай проста регресії x по y (BB) перетинає поземну просту M_2x у точці M.

Тоді прямовисна проста через M перетне поземну вісь OX в точці M_1, що являє собою середнє аритметичне всіх X.

Справді. $R_{\frac{x}{y}} = \frac{kc}{cM} = \frac{x}{y}$; звідци $x = R_{\frac{x}{y}} \,.\, y$. Для всякого y-ого ряду, з числом варіянтів p, сума x-ів буде в p разів більша, тоб-то $pR_{\frac{x}{y}} \,.\, y$, а для всіх рядів (для всеї мережі) $\Sigma x = R_{\frac{x}{y}} \,.\, \Sigma py$. Але Σpy, тоб-то сума всіх відхилів від середнього аритметичного дорівнює 0, тому й $\Sigma x = 0$, а це можливо лише тоді, як M_1 є середнє аритметичне всіх x-ів.

Далі. Ми можемо написати Σxy у вигляді $y\Sigma x$, а через те, що для всякого ряду $\Sigma x = R_{\frac{x}{y}} py$, буде $\Sigma xy = y\Sigma x = y R_{\frac{x}{y}} py = R_{\frac{x}{y}} py^2$, а для всіх рядів

$$\Sigma xy = R_{\frac{x}{y}} \Sigma py^2 = R_{\frac{x}{y}} n\sigma_y^2,$$

бо для всіх рядів Σpy^2 є сума всіх y^2, а $\frac{\Sigma py^2}{n} = \sigma^2$, отже $\Sigma py^2 = n\sigma^2$.

Із останньої рівности витікає, що

$$R_{\frac{x}{y}} = \frac{\Sigma xy}{n\sigma_y^2}.$$

Так само, розглядаючи просту регресії y по x (CC) знайдемо, що

$$R_{\frac{y}{x}} = \frac{\Sigma xy}{n\sigma_x^2}.$$

Тепер ми можемо одержати r.

$$r = \sqrt{R_{\frac{x}{y}} \cdot R_{\frac{y}{x}}} = \sqrt{\frac{\Sigma xy}{n\sigma_y^2} \cdot \frac{\Sigma xy}{n\sigma_x^2}} = \frac{\Sigma xy}{n\sigma_y\sigma_x} = \frac{\Sigma D_x D_y}{n\sigma_x\sigma_y},$$

бо відхили (від середнього) окремих X ми можемо зазначити також відомим D_x, а відхили Y зазначити D_y.

§ 7. З формули для r маємо $\Sigma xy = r \cdot n\sigma_x\sigma_y$.

Вносячи цей вираз для Σxy у формули для R, одержимо для них загальновживані формули:

$$R_{\frac{x}{y}} = \frac{r \cdot n\sigma_x\sigma_y}{n\sigma_y^2} = r\,\frac{\sigma_x}{\sigma_y} \quad \text{і} \quad R_{\frac{y}{x}} = r\,\frac{\sigma_y}{\sigma_x}.$$

П р и к л а д. Супряженість між середньою довжиною меживузля колосу і врожаєм його зерна виражається: $r = +0{,}56$; $\sigma_x = 0{,}54$ мм., $\sigma_y = 0{,}16$ гр.

Звідци, $R_{\frac{y}{x}} = +0{,}56 \cdot \frac{0{,}16 \text{ гр.}}{0{,}54 \text{ мм.}} = +0{,}166$ гр. на 1 мм., тоб-то подовження меживузля колосу пересічно на 1 мм. супряжене із збільшенням урожаю його зерна на 0,166 гр.

З другого боку, $R_{\frac{x}{y}} = +0.56 \cdot \frac{0{,}54 \text{ мм.}}{0{,}16 \text{ гр.}} = +1{,}89$ мм. на 1 гр., тоб-то збільшення врожаю зерна колосу на 1 гр. іде поруч із подовженням його (середнього) меживузля на 1,89 мм.

§ 8. Формули $r = \frac{\Sigma D_x D_y}{n\sigma_x\sigma_y}$ уживають лише в разі н е в е л и к о г о n, коли матеріял не можна розташувати по класах і збудувати кореляційну мережу. Як що-ж ми маємо таку мережу, то в окремих комірках її (див., напр., табл. в § 1, розд. XII) добутки $D_x . D_y$ повторюються кілька (p) разів; тому ми можемо заступити їх (аналогічно із знайомим нам виразом $p\alpha$) добутками $p\alpha_x\alpha_y$. Тоді

$$r = \frac{\Sigma p\alpha_x\alpha_y}{n\sigma_x\sigma_y}.$$

§ 9. Щоб полегчити обчислювання, ми всюди заступали α величинами a і b; зробімо це й тут.

Як що, взагалі, $a = \alpha + b$, то

$$a_x = \alpha_x + b_x \quad \text{і} \quad a_y = \alpha_y + b_y.$$

Тому:

$$a_xa_y = (\alpha_x + b_x)(\alpha_y + b_y) = \alpha_x\alpha_y + \alpha_xb_y + \alpha_yb_x + b_xb_y.$$

Звідци

$$\Sigma pa_xa_y = \Sigma p\alpha_x\alpha_y + \Sigma p\alpha_xb_y + \Sigma p\alpha_yb_x + \Sigma pb_xb_y.$$

Але $\Sigma p\alpha = 0$, тому $\Sigma p\alpha_xb_y = 0$ і $\Sigma p\alpha_yb_x = 0$;

крім того $\Sigma pb_xb_y = (\Sigma p)\, b_xb_y = nb_xb_y$, бо b_x і b_y є величини сталі. Отже:

$$\Sigma pa_xa_y = \Sigma p\alpha_x\alpha_y + nb_xb_y,$$

$$\text{звідки } \Sigma p\alpha_x\alpha_y = \Sigma pa_xa_y - nb_xb_y.$$

Впроваджуючи в формулу для r, одержимо:

$$r = \frac{\Sigma pa_xa_y - nb_xb_y}{n\sigma_x\sigma_y}.$$

Як-що класова проміжка не є рівна 1, то слід узяти її за 1, виносячи за дужки і скорочуючи квадрат класової проміжки в чисельнику і знаменику, і обчисляти тоді з як-найбільше спрощеної формули

$$r = \frac{\Sigma pa'_xa'_y - n\beta_x\beta_y}{ns_xs_y}.$$

§ 10. Формулу для r виведено в тім припущенні, що всі M_x і M_y лежать на відповідних простих лініях.

На практиці, через варіювання, M_x і M_y трохи відхиляються то в той, то в той бік від теоретичних простих BB і CC, але і в такому разі вказана формула є найкраща для визначення сучинника кореляції, певну-ж похибку, що з нею, через варіювання, є звязане r (виведене вказаним способом), визначають з формули

$$m_r = \frac{1 - r^2}{\sqrt{n}},$$

досить точної для нашої мети.

Відповідно матимемо:

$$\text{для } R_{\frac{y}{x}} \qquad m_R = m_r \cdot \frac{\sigma_y}{\sigma_x}$$

$$\text{і для } R_{\frac{x}{y}} \qquad m = m_r \cdot \frac{\sigma_x}{\sigma_y}.$$

§ 11. Наведену формулу для r можна вживати лише в разі прямолінійної кореляції, або близької до неї. В разі-ж супряжености криволінійної ця формула може дати грубо-помилковий вислід, і вживати її в таких разах не можна. Так, напр., як-що матеріял розташовано в таку (теоретичну) мережу (за Йогансеном):

	0	3	6	9	12	15	18	21	24	27	30 33	P_y	M_y
0									1			1	25,5
5						1	1	3	3	1		9	23,2
10				1	2	7	7	7	4	1		29	19,9
15			1	3	13	17	15	8	2	1		60	17,5
20		1	6	16	21	21	17	9	3	1		95	15,7
25	1	7	15	20	23	20	15	6	3	1	1	112	12,9
30		1	6	16	21	21	17	9	3	1		95	15,7
35			1	3	13	17	15	8	2	1		60	17,5
40				1	2	7	7	7	4	1		29	19,9
45						1	1	3	3	1		9	23,2
50 55									1			1	25,5
P_x	1	9	29	60	95	112	95	60	29	9	1	500	
M_x	27,5	27,5	27,5	27,5	27,5	27,5	27,5	27,5	27,5	27,5	27,5		

то обчислене з указаної формули r дорівнює нулеві. Проте, тут виразно помітно кореляцію: із збільшенням x, y в середньому спочатку зменшується, а далі, від середини, знову збільшується; кореляція тут, як показують M_y, не простолінійна, а криволінійна. З огляду на це, рекомендується в усіх сумнівних прикладах обчисляти M_x і M_y і в разі їх безперечної криволінійної зміни братися до визначування супряжености иншими способами, але спинятися на цих способах ми тут не будемо.

§ 12. Як саме проводити обчислення r, нам покажуть наступні приклади.

1-й приклад (з малим числом варіянтів, взято у Йогансена). Зважено 25 зернин ячменю і визначено в них кількість азоту. Треба визначити, чи є і яка на величину супряженість між цими двома ознаками. Тут n є невелике, тому користуємося з формули $r = \frac{\Sigma D_x D_y}{n \sigma_x \sigma_y}$ і провадимо обчислення, як на таблиці, що її подано нижче.

Користуючися з ΣX, ΣY, ΣD_x^2, ΣD_y^2 і $n = 25$, знайдемо: $\sigma_x = 4{,}258$ і $\sigma_y = 0{,}154$. Отже,

$$r = \frac{\Sigma D_x D_y}{n \sigma_x \sigma_y} = \frac{+12{,}628}{25.4{,}258.0{,}154} = +0{,}77$$

$$m_r = \frac{1 - r^2}{\sqrt{n}} = \frac{1 - 0{,}77^2}{\sqrt{25}} = \pm 0{,}08.$$

Остаточно, $r = +0{,}77 \pm 0{,}08$ (див. табл. на стор. 71).

2-й приклад (з варіянтами, розташованими в кореляційній мережі). В 365 колосів визначено: щільність кожного колосу і врожай його зерна; одержані числа дали таку мережу (поземні ряди — вага у гр., прямовисні — щільність у мм.):

mm. \ gr.	0–0,1	0,1–0,2	0,2–0,3	0,3–0,4	0,4–0,5	0,5–0,6	0,6–0,7	0,7–0,8	0,8–0,9	0,9–1,0	
3,7–4,2	1	6	4	2	1						14
4,2–4,7		1	15	9	11	6					42
4,7–5,2		1	7	20	28	36	26	6	1		125
5,2–5,7			6	4	20	40	39	13	2		124
5,7–6.2					6	12	21	4	4	2	50
6.2–6,7					3	3	2	1			9
6,7–7,2									1		1
	1	9	32	35	69	97	88	24	8	2	365

В таких випадках користуємося з формули

$$r = \frac{\Sigma p a'_x a'_y - n \beta_x \beta_y}{n s_x s_y}.$$

№ варі-янту	Вага зерна X	Азот у % Y	Відхили від середнього		Добутки $D_x D_y$		Квадрати відхилів	
			D_x	D_y	+	—	D_x^2	D_y^2
1	66,0	1,71	+ 11,52	+ 0.236	2,719		132,7102	0,0557
2	62,4	1,57	+ 7,92	+ 0,096	0,760		62,7264	0,0092
3	58,8	1,66	+ 4,32	+ 0,186	0,804		18,6624	0,0346
4	53,4	1,52	— 1,08	+ 0,046		0,050	1,1664	0,0021
5	51,1	1,36	— 3,38	— 0,114	0,385		11,4244	0,0130
6	51,2	1,41	— 3,28	— 0,064	0,210		10,7584	0,0041
7	49,0	1,29	— 5,48	— 0,184	1,008		30,0304	0,0339
8	51,2	1,31	— 3,28	— 0,164	0,538		10,7584	0,0269
9	55,2	1,45	+ 0,72	— 0,024		0,017	0,5184	0,0006
10	55,3	1,42	+ 0,82	— 0,054		0,044	0,6724	0,0029
11	48,5	1,31	— 5,98	— 0,164	0,981		35,7604	0,0269
12	52,4	1,44	— 2,08	— 0,034	0,071		4,3264	0,0012
13	54,8	1,31	+ 0,32	— 0,164		0,052	0,1024	0,0269
14	51,8	1,33	— 2,68	— 0,144	0,386		7,1824	0,0207
15	59,6	1,74	+ 5,12	+ 0,266	1,362		26,2144	0,0708
16	56,8	1,51	+ 2,32	+ 0,036	0,084		5,3824	0,0013
17	53,4	1,67	— 1,08	+ 0,196		0,212	1,1664	0,0384
18	54,8	1,39	+ 0,32	— 0,084		0,027	0,1024	0,0071
19	51,8	1,49	— 2,68	+ 0,016		0,043	7,1824	0,0003
20	51,8	1,45	— 2,68	— 0,024	0,064		7,1824	0,0006
21	55,4	1,53	+ 0,92	+ 0,056	0,052		0,8464	0,0031
22	51,0	1,24	— 3,48	— 0,234	0,814		12.1104	0,0548
23	54,6	1,41	+ 0,12	— 0,064		0,008	0,0144	0,0041
24	50,2	1,45	— 4,28	— 0,024	0,103		18,3184	0,0006
25	61,4	1,87	+ 6,92	+ 0,396	2,740		47,8864	0,1568
Суми	1361,9 ΣX	36, 84 ΣY			$\underbrace{+13,081 \quad -0,453}_{\Sigma D_x D_y = +12,628}$		453,2060 ΣD_x^2	0,5966 ΣD_y^2
M	54,48	1,474						

Вибравши тут $A_y = 0{,}55$ гр., $A_x = 5{,}45$ мм. і запровадивши в мережу (у відповідні комірки) добутки $a'_x \cdot a'_y$, надамо їй такого вигляду:

$A_y = 0{,}55$

	−5	−4	−3	−2	−1	0	+1	+2	+3	+4	
−3	$+1._{15}$	$6._{12}$	$4._{9}$	$2._{6}$	$1._{3}$					—	14
−2		$2._{8}$	$15._{6}$	$9._{4}$	$11._{2}$	$6._{0}$					42
−1		$1._{4}$	$7._{3}$	$20._{2}$	$28._{1}$	$36._{0}$	$26._{1}$	$6._{2}$	$1._{3}$		125
$A_x = 5{,}45$ 0			$6._{0}$	$4._{0}$	$20._{0}$	$40._{0}$	$39._{0}$	$13._{0}$	$2._{0}$		124
+1					$6._{1}$	$12._{0}$	$21._{1}$	$4._{2}$	$4._{3}$	$2._{4}$	50
+2					$3._{2}$	$3._{0}$	$2._{2}$	$1._{4}$			9
+3	—								$1._{9}$	+	1
	1	9	32	35	69	97	88	24	8	2	365

X-й ряд

Y-ий ряд.

Добутки $a'_x a'_y$ у лівім верхнім і правім нижнім квадрантах є додатні, а в правім верхнім і лівім нижнім є від'ємні, що зазначено в кутах таблиці. Дальший запис (а саме, $pa'_x a'_y$) можемо вести так:

+	−
15+72+36+12+3+8+90+36+ +22+4+21+40+28=+387	−26−12−3=−41
−6−6=−12	+21+8+12+8+4+4+9=+66
−	+

Звідци $\Sigma pa'_x a'_y = +(387+66)-(41+12) = +400$.

Тепер обчислім β_x, β_y, s_x, s_y.

$\pm a'_x$	1	2	3
$+p$	50	9	1
$-p$	125	42	14

$\Sigma pa'_x$ −75.1−33.2−13.3=−180

$\Sigma pa'^2_x$ 175.1+51.4+15.9=514

$$\beta_x = \frac{-180}{365} = -0{,}49; \beta_x^2 = 0{,}24$$

$$s_x = \pm\sqrt{\frac{514}{365} - 0{,}24} = \pm 1{,}08$$

$\pm a'_y$	1	2	3	4	5
$+p$	88	24	8	2	
$-p$	69	35	32	9	1

$$\beta_y = \frac{-108}{365} = -0{,}3; \quad \beta_y^2 = 0{,}09$$

$$s_y = \pm\sqrt{\frac{954}{365} - 0{,}09} = \pm 1{,}58$$

$\Sigma pa'_y$ $+19.1-11.2-24.3-7.4-1.5=-108$

$\Sigma pa'^2_y$ $157.1+59.4+40.9+11.16+1.25=954$

З цих даних маємо:

$$r = \frac{\Sigma pa'_x a'_y - n\beta_x\beta_y}{ns_x s_y} = \frac{+400-365(-0.49)(-0.3)}{365 \,.\, 1{,}08 \,.\, 1{,}58} = \frac{+400-53.66}{622{,}84} = +0{,}56$$

$$m_r = \frac{1-r^2}{\sqrt{n}} = \frac{1-0{,}3125}{19} = \pm 0{,}04.$$

Одержаний $r = +0{,}56 \mp 0{,}04$ говорить, що між щільністю (що її визначає середня довжина меживузля) колоса та врожаєм його зерна є велика додатня супряженість: що є довше середнє колосове меживузля, тоб-то що рідший є колос, то вищий є врожай його (в зерні).

XIII. Значіння коефіціента кореляції.

§ 1. Вказуючи на супряженість двох ознак, коефіціент кореляції, все-ж таки, не визначає сам по собі причинової залежности між ними. Так, напр., у нащадках гібрида рідкоколосої пшениці з щільноколосою, що складаються з рослин з колосами різної щільности, ми знайдемо, що для довжини колоскових і квіткових лусок $r = 0{,}94 \pm 0{,}006$, тоб-то надзвичайно високу супряженість. Але з цього зробити висновок, що довжина колоскових лусок обумовлює довжину квіткових (або навпаки), ми не маємо права, бо тільки відповідне біологічне дослідження може відшукати причиновий звязок у цих явищах. В нашому прикладі, як показує гібридологічна аналіза, звязок між довжиною тих і тих лусок є другоразовий: особливий спадковий фактор, що своєю дією вкорочує колоскові і квіткові луски, спричиняє також ущільненість колоса, обумовлюючи сам безпосередньо кореляцію, що ми спостерігаємо між ними.

§ 2. Коефіціент кореляції дуже допомагає при розвязуванні різноманітних агрономічних питань. Раніше ми познайомилися вже з 3 прикладами таких агрономічних питань, тут ми спинимося ще на трьох.

1-й приклад. Користуючися з числових даних урожаїв 480 ділянок, кожна 2 с. $\times$ 6 с., пшеничного поля (див. нижче табл. врожайности), що на ньому переведено частковий облік, дослідити, чи є супряжені один з одним урожаї сусідніх, а також і більш-менш віддалених одна від одної ділянок, чи вони що до цього є незалежні.

Умовмося зазначати двома числами f і g таке взаємне розміщення ділянок, коли якась ділянка є віддалена від иншої на f прямовисних рядів (тоб-то в поземному напрямі) і на g поземних рядів (тоб-то в прямовисному напрямі). Так, напр., для наведених тут 24 ділянок.

1	2	3	4	5	6
7	8	9	10	11	12
13	14	15	16	17	18
19	20	21	22	23	24

положення 3 відносно 1 ми зазначаємо символом (2,0); 9 відносно 1 символом (2,1); 7 відносно 1—(0,1); 23 відносно 8—(3,2) і т. д.

Отже, нашу задачу можна звести на визначення r_{fg}.

Для обчислення $r_{0,1}$, зіставляємо кожну пару ділянок з відносним розміщенням (0,1) і на підставі цих пар чисел складаємо кореляційну мережу. Для даного розміщення ділянок нашого поля вона матиме такий вигляд:

	225	245	265	285	305	325	345	365	385	405	425	445
225				1								1
245						2						2
265		1	5	5	7	1	1					20
285		1	6	42	43	10	3					105
305	1		7	44	97	45	6		1			201
325			1	12	46	28	13	2		1		103
345				5	8	11	3	2				29
365						2	2					4
385						1						1
405											1	1
425									1			1
445	1	2	19	109	201	100	29	4	2	1	1	468

Ярова пшениця.

Врожай у пудах, облік переведено ділянками у 2 × 6 кв. с.

(за Юровським).

359	305	313	323	328	325	313	313	281	328	309	315
366	318	346	321	308	304	291	239	308	345	316	324
344	326	326	304	290	301	305	298	290	350	389	341
349	321	328	336	293	309	315	283	288	294	343	324
303	327	338	321	290	336	332	298	289	311	361	318
320	321	321	290	303	301	294	280	290	330	348	320
306	298	308	304	316	314	296	311	308	316	329	316
293	302	315	314	304	299	325	311	305	303	294	291
288	321	325	318	311	324	313	284	283	283	301	325
284	320	314	298	315	327	313	298	284	283	330	305
284	320	320	321	321	303	305	308	319	316	363	320
288	307	329	331	313	315	300	309	316	306	310	311
299	324	306	326	299	289	296	304	284	294	303	328
314	342	308	324	304	304	306	311	316	351	333	307
317	339	329	327	305	298	301	294	283	310	306	290
301	321	303	298	298	286	306	323	283	316	314	307
297	320	329	351	292	298	320	311	336	336	321	336
309	338	304	341	302	305	353	305	333	356	338	334
320	348	344	326	296	299	303	304	298	326	337	309
309	331	354	320	290	288	311	302	311	316	357	315
336	339	353	313	293	286	323	305	331	333	296	309
349	343	335	308	291	300	334	308	323	385	324	357
354	314	304	299	306	309	302	296	269	328	308	335
319	328	326	322	314	304	315	315	314	308	318	321
306	328	306	302	318	288	324	332	303	398	303	339
311	337	332	311	316	299	327	314	281	319	330	343
323	341	309	322	311	304	311	308	301	320	334	357
333	321	306	321	323	288	347	329	293	308	321	339
328	308	336	351	327	309	319	314	274	313	316	307
308	337	319	322	316	303	330	318	285	313	325	318
341	340	313	328	293	329	318	320	303	319	311	293
350	336	330	332	331	319	325	319	331	335	302	290
323	320	314	341	311	301	322	313	318	328	313	289
313	337	338	346	314	323	334	329	306	268	313	295
307	306	353	339	289	298	309	309	270	313	313	299
324	334	343	383	319	333	315	295	281	354	311	311
333	325	328	356	310	308	278	263	263	316	309	300
421	334	326	373	313	301	361	334	326	339	339	350
429	336	344	349	298	304	342	306	323	313	327	302
386	337	316	325	319	319	341	319	318	291	308	293

Із цеї мережі одержимо $r_{0,1} = +0,42 \pm 0,04$, тоб-то досить велику супряженість.

Зіставляючи ділянки з відносним розміщенням (0,2), одержимо $r_{0,2} = +0,27 \pm 0,04$. Далі матимемо: $r_{0,3} = +21 \pm 0,05$; $r_{0,4} = +0,22 \pm$ $\pm 0,05$; $r_{0,5} = \mp 0,17 \pm 0,05$; $r_{0,6} = +0,14 \pm 0,05$; $r_{0,7} = +0,11 \pm 0,05$ (тоб-то $= 0$, як що взяти на увагу m_r). Ідучи в инших напрямах, одержимо: $r_{1,0} = +0,28 \pm 0,04$; $r_{2,0} = 0$; $r_{1,1} = 0$; $r_{1,2} = 0$.

Отже, між урожаями сусідніх ділянок є певний звязок, або супряженість, але що далі лежать ділянки одна від одної, то звязок цей стає все слабкішим, аж поки не зникає зовсім. Ділянки з розміщенням (0,6) лежать одна від одної на 12 саж. і з розміщенням (1,0) теж на 12 саж.: це є віддалі, на які сягає звязаність поміж ділянками, далі вона зникає.

2-й приклад. Покажчик різноманітности поля (v) являє собою, як ми знаємо, також покажчика точности спроби, коли нема повторности, Коли-ж увести повторність, розміщуючи окремі повторення поза межами сягання звязанности, що про неї була мова в першому прикладі. то точність спроби (P) наростає пропорціонально до кореня квадратового з числа (n) повторень: $P = \frac{v}{\sqrt{n}}$. Як буде наростати точність спроби без повторности, але при збільшенні поверхні ділянки?

Умовмося зазначати збільшену ділянку, втворену з t послідовних прямовисних рядів вихідних ділянок по u послідовних ділянок в кожнім, символом ($t.\ u$); напр., коли вихідна ділянка є рівна 2 кв. с., то ділянку із 3 взятих поряд прямовисних рядів по 4 взятих поряд ділянки в кожнім зазначаємо (3.4), тоб-то 12,2 кв. с. $= 24$ кв. с. Щоб зазначати відносне розміщення ділянок, будемо брати знайомий уже нам символ (f,g).

Як-що основний відхил для вихідних ділянок зазначити $\sigma_{1\cdot1}$, то основний відхил для збільшених ділянок, із ($t.\ u$) відхідних можна виразити наступною формулою, що її візьмемо тут в готовому вигляді:

$$\sigma_{t.u} = \frac{\sigma_{1\cdot1}}{\sqrt{t.u}}\sqrt{1 + 2\Sigma\left(1 - \frac{f}{t}\right)\left(1 - \frac{g}{u}\right) r_{fg}}\,.$$

В разі повторности $r = 0$, а $t.u = n$, тоді $\sigma_{t.u} = m_{(n)} = \frac{\sigma}{\sqrt{n}}$, знайомий нам вираз. Як видко з наведеної формули, при збільшенні поверхні $\frac{\sigma}{\sqrt{n}}$ множиться на число, більше за 1; отже, точність спроби від збільшення поверхні ділянки наростає повільніше, як від збільшення повторности.

Користуються з указаної формули так. Спершу визначають $\sigma_{1\cdot 1}$ і r_{fg}; припустімо, що для вихідних ділянок, кожна в 1 квадр. саж., $\sigma_{1\cdot 1} = \pm 67$ пуд. і

$$r_{1,0} = +0{,}42 \qquad r_{0,1} = +0{,}33$$
$$r_{2,0} = +0{,}28 \qquad r_{1,1} = +0{,}30$$
$$r_{3,0} = +0{,}23 \qquad r_{2,1} = +0{,}19$$
$$r_{4,0} = +0{,}23$$

Инші r_{fg}, коли взяти на облік m_r, є рівні 0.

Обчислім $\sigma_{(48)} = \sigma_{3.16}$, тоб-то основний відхил для збільшених ділянок, складених кожна з 48 вихідних ділянок (3 ряди по 16 вихідних ділянок в кожному). З даної доціру формули:

$$\sigma_{(48)} = \sigma_{3.16} = \pm \frac{67}{\sqrt{3.16}} \sqrt{1 + 2\left[\left(1-\frac{1}{3}\right)\left(1-\frac{0}{16}\right).0{,}42 + \left(1-\frac{2}{3}\right)\left(1-\frac{0}{16}\right).0{,}28 + \left(1-\frac{0}{3}\right)\left(1-\frac{1}{16}\right).0{,}33 + \left(1-\frac{0}{3}\right)\left(1-\frac{2}{16}\right).0 + \left(1-\frac{1}{3}\right)\left(1-\frac{1}{16}\right).0{,}3 + \left(1-\frac{2}{3}\right)\left(1-\frac{1}{16}\right).0{,}19\right]} = \pm 15 \text{ пуд.}$$

Звичайно, такі вирази, як $\left(1-\frac{0}{16}\right)$, $\left(1-\frac{0}{3}\right)$, а також і всі ті, де $r_{fg} = 0$, не треба навіть і виписувати; їх наведено тут лише для того, щоб полегшити розуміння порядку обчислення $\sigma_{t.u}$.

3-й приклад. Коефіцієнта регресії можна використати в польовій спробі для обчислення виправлених урожаїв з дослідних ділянок в тих випадках, коли вони в різній мірі порідшали від якогось шкідника або з иншої причини. Так, напр., одна спроба на сорт, де було 50 ліній баклажанів при 4 повтореннях, дуже порідшала від дротяного хробака: замісць насаджених 16 кущів на ділянці залишилося здебільшого 3—6, загалом-же від 1 до 15. Тому рясність травостою на окремих ділянках, як одної й тої самої лінії, так і різних, виявилася занадто неоднаковою, і числа, одержані у спробі, безпосередньо непридатними для порівняльного визначення врожайности ліній. В таких випадках треба вносити де-яку поправку. Часто цю поправку роблять так, що поділюють урожай з ділянки (зазначмо його V) на число рослин, що залишилися на на цій ділянці (зазначмо n), і одержану частку помножають на число насаджених (або посіяних) на цій ділянці особин (зазначмо його N); як-що такий виправлений урожай зазначити літерою U, то тоді

$$U = \frac{V}{n} . N.$$

Проте, цей спосіб є занадто неточний. Річ у тому, що окремі рослини розвиваються, взагалі, тим буйніше, що менше їх є на ділянці, тоб-то що є рідший травостій. Тому частка $\frac{V}{n}$ буде тим хибніша, що рідшою буде дана ділянка. Середній урожай одної рослини на розрідженій ділянці буде, таким чином, відносно вищий, ніж був-би він за нормальної або принятої густоти. Для обчислення точної поправки на розрідженість треба з дослідного середнього врожаю одної рослини відняти вказаний відносний приріст урожаю.

Часто цей приріст можна визначити досить добре за допомогою коефіціента регресії (R). Розмістивши у кореляційну мережу числа врожаїв на 1 рослину, одержані для окремих ділянок, обчисляємо r, а тоді R. Величина його показує, наскільки, пересічно, збільшується врожай одної рослини, коли густота травостою зменшується на 1 рослину, тоб-то—дає відшукувану поправку. Але користуватися з одержаної абсолютної вартости R також не можна, бо на врожай упливає не тільки густота травостою, але також і якість грунту, природа сортів то-що. Вони вносять, у свою чергу, ту чи иншу зміну врожаю з ділянки, тим часом як R показує лише с е р е д н ю зміну сумарно взятих урожаїв. Тому треба користуватися із зведеного R, прийнявши один із середніх (в кореляційній мережі) урожаїв за одиницю.

Візьмімо за вихідне те середнє, що виражає врожай на 1 рослину при найбільшім числі (N) рослин на ділянці, і зазначмо це середнє M_N. Тоді середній урожай 1 рослини при рідшім травостою, з n рослинами на ділянці, виразимо так: $M_N + (N - n) . |R|$, або $M_N - (N - n)R$. Припустімо, що, напр., природа сорту змінює «справжній» урожай на даній ділянці в k разів. Тоді врожай 1 рослини при N їх на ділянці дорівнював-би kM_N, а врожай 1 рослини при n їх на ділянці, тоб-то одержане із спроби $\frac{V}{n}$ дорівнюватиме $kM_N + (N - n) |R| k$.

Другий член цеї суми помножмо на $\frac{M_N}{M_N}$. Тоді

$$\frac{V}{n} = kM_N + (N - n) |R| . \frac{kM_N}{M_N}, \text{ або}$$

$$\frac{V}{n} = kM_N \left[1 + (N - n) . \frac{|R|}{M_N}\right].$$

Невідоме тут є kM_N. Розвязуючи відносно цього невідомого наше рівнання, одержимо

$$kM_N = \frac{V:n}{1+(N-n)\,|\,R\,|:M_N}$$

звідци виправлений урожай на дану ділянку—

$$U = N.\,kM_N = \frac{N.\,V:n}{1+(N-n)\,|\,R\,|:M_N}$$

У поданому прикладі числа врожаїв на одну рослину вляглися в кореляційній мережі так:

0 1 2 3 4 5 6 7 8 9 10 фунт. на 1 рослину

рослин на ділянці

1		2	1	3	1	3		1		1	12 (4,5)
2			3	7	5		2				17 (4)
3		3	6	11	2	2		1	1		26 (3,6)
4		4	9	3	2	1					19 (2,8)
5		8	12	8	1						29 (2,6)
6	1	6	15	3	1						26 (2,4)
7		7	11	1							19 (2,2)
8	1	10	7								18 (1,8)
9	1	6	2	1							10 (1,8)
10		5	4								9 (1,9)
11	1	3	2								6 (1,7)
12	1	1									2 (1)
13		2									2 (1,5)
14		1									1 (1,5)
15		1									1 (1,5)
	5	59	72	37	12	6	2	2	1	1	197

Середні врожаї на 1 рослину для різних густот, наведені праворуч у дужках, спершу падають круто (це буде ясніше, як що нанести їх на графік), потім (від 4 до 11 рослин) що-раз рівніше, а далі йдуть поземно. Пам'ятаючи про сказане в розд. XII § 11, при обчисленні r відкидаємо ці крайні ряди, обмежуючись середніми рядами (від 4 до 11 рослин).

Із цеї вилученої кореляційної мережі одержимо: $r = -0{,}44$; $R_{\frac{y}{x}} =$ $= -0{,}17$ ф., або наше $|R| = 0{,}17$ ф.; $\frac{|R|}{M} = \frac{0{,}17}{1{,}7} = 0{,}1$, $N = 11$.

Звідци в нашій спробі

$$U = \frac{11 \,.\, V:n}{1 + (11 - n) \,.\, 0{,}1}.$$

Запроваджуючи для кожної з ділянок її $\frac{V}{n}$ і n, одержимо виправлені врожаї їх, причому вичеркнуті з мережі ділянки вичеркуємо, звичайно, і з цих перечислень.

Відношення $\frac{R}{M_N}$ має тут другорядне значіння, тому буде найкраще, коли взяти „справжнє" M_N, тоб-то таке, що лежить на теоретичній простій регресії. Обчислити його є дуже легко. Як відомо, $R_{\frac{y}{x}} = \frac{y}{x}$; звідци $y = R_{\frac{y}{x}} \,.\, x$. В нашім випадку y є відхил „справжнього" M_N від середнього аритметичного всіх $\frac{V}{n}$ нашої вилученої кореляційної мережі; x є відхил N від середнього числа рослин для всіх ділянок тої самої мережі. Як-що середнє для всіх y зазначити M_y, а середнє для всіх x зазначити M_x, то виправлене $M_N = M_y - |R|\,(N - M_x)$.

В нашому прикладі:

$$M_{11} = 2{,}28 - 0{,}17 \,.\, 4{,}4 = 1{,}53 \text{ ф., бо } M_y = 2{,}28 \text{ і } M_x = 6{,}6.$$

Примітка. Обчислена цим способом поправка на розрідженість є вірна лише в першім наближенні і на порівнюючи невеликих інтервалах, бо в дійсності із зміною густоти травостою врожай змінюється не простолінійно, а криволінійно.

XIV. Практичні вказівки, різні уваги та приклади до вправ.

До розд. II.

До § 3. Розміщаючи варіяційний матеріял за класами, не варто брати їх занадто багато; звичайно досить брати від 5 до 10, в залежності від числа варіянтів: що їх є менше, то й клас треба брати менше, инакше вийде фальшива многоверхова крива.

Розміщення варіянтів за класами найзручніше вести так. Відзначають найменший і найбільший варіянт і проміжку між ними поділяють на бажане число клас; таким шляхом одержують величину класової проміжки. Зробивши схему варіяційного ряду з цими класовими проміжками, прочитують одне по одному числові вартості всіх варіянтів і відзначають при цьому точками у відповідних класах положення кожного з варіянтів у варіяційнім ряді. Тоді, підрахувавши числа точок в кожнім класі, знатимемо чисельність варіяцій. Приклади до вправ дано (в таблицях) нижче (стор. 82 і 83).

До розд. III.

До § 3. При обчисленні M не забувати про знак у b або β; що до самої β, не забути наприкінці помножити її на класову проміжку.

В разі однобокої кривої, щоб точніше обчислити M, треба брати вузькі класи.

До розд. IV.

До § 6. Для роботи з аритмометром найкраща формула є

$$\sigma = \pm \sqrt{\frac{\Sigma V^2}{n} - M^2}.$$

Приклади до вправ.

1.	7,3	7,8	8,3	8,8	9,3	9,8	10,3	10,8	11,3	11,8	12,3	мм.
		3	17	55	82	90	103	153	77	20	1	насінин.
2.	76	79	82	85	88	91	94	97	100	103	106	пуд.
		5	13	24	29	46	36	27	22	13	3	ділянок.
3.	4	5	6	7	8	9	10	пелюсток.				
	3	22	86	152	140	34	2	квіток.				

Що до σ — не забувати за класову проміжку.

4. 13, 18, 14, 14, 17, 15, 16, 17, 16, 19, 12 пуд.
 $M = ?$ $\sigma = ?$

5. 130,9; 130,9; 139,7; 135,3; 144,2; 148,6; 133,1; 137,5; 141,9; 148,6; 161,9; 150,8; 130,9; 130,8. $M = ?$ $\sigma = ?$

До § 3, розд. II. Приклади до вправ.

1. Довжина зернин банатки (в мм.)

Рознести по класах, нарисувати криву то-що.

6,14	5,69	5,92	5,67	6,32	5,88	6,94
5,85	7,05	5,78	6,65	6,01	6,47	5,69
6,48	7,41	6,64	6,42	5,98	6,17	5,70
5,81	6,37	5,20	6,05	6,51	6,40	6,32
6,56	5,70	7,23	5,03	5,34	5,87	6,08
6,22	6,35	5,59	6,06	6,39	6,06	6,36
5,18	5,76	6,45	6,02	6,32	6,15	5,24
6,99	6,30	6,02	6,18	5,50	6,16	6,67
5,68	6,12	5,89	6,31	5,59	6,35	5,66
6,58	5,98	6,79	6,02	6,21	6,74	6,01
5,60	5,92	6,57	6,42	6,52	6,67	6,52
6,79	6,68	6,50	6,00	5,32	6,09	6,11
6,28	5,52	5,81	6,12	5,68	5,50	6,31
6,10	6,89	6,25	5,07	6,28	6,25	6,58
5,64	6,04	5,79	6,45	6,60	6,20	6,00
6,58	5,92	5,78	4,99	6,32	6,51	6,28
5,91	6,94	6,58	5,32	6,35	6,83	5,95
6,59	5,82	6,38	6,59	6,11	6,98	6,06
5,95	5,90	5,23	5,71	5,09	6,15	6,65
5,93	6,19	6,37	6,67	5,72	6,30	6,93
6,01	6,38	6,26	5,58	5,69	6,21	5,91
6,56	5,78	6,81	5,37	6,95	5,95	5,40
6,50	6,40	6,12	4,00	6,73	6,22	6,25
6,25	6,75	6,08	6,17	6,78	5,97	6,50
6,24	6,74	5,86	7,05	5,99	6,58	6,62
6,22	6,39	5,90	5,49	5,50	6,36	5,71
6,14	5,93	6,58	6,26	5,94	6,23	6,15
5,87	5,98	5,26	6,08	6,43	6,60	5,55
6,03	5,42	5,83	5,71	6,13	6,69	5,65
6,14	6,22	6,08	6,26	5,92	6,36	6,30
5,95	6,24	5,12	5,72	5,84	6,95	6,15
5,83	5,87	4,99	5,87	5,92	5,13	6,89
6,06	5,17	6,92	6,27	6,21	5,50	6,39

2. Ширина зернин банатки (в сотих частинах мм.)

Розмістити по класах то-що.

299	270	266	210	259	254	262
250	265	231	214	249	281	305
298	299	255	277	270	266	271
273	300	258	256	242	297	290
278	312	251	292	286	274	219
271	239	241	285	282	252	202
224	280	278	287	201	270	308
258	259	240	238	295	254	229
292	287	310	310	228	268	251
271	272	252	245	304	270	298
248	281	251	287	288	279	337
303	265	269	299	251	288	268
288	289	309	281	268	229	274
280	221	289	298	289	295	268
262	261	265	276	261	308	292
248	300	279	279	278	275	325
264	291	252	230	249	270	232
264	259	259	272	263	245	274
292	247	214	220	298	281	286
271	271	253	222	260	299	284
233	306	227	321	277	335	335
301	291	250	243	268	288	281
292	247	265	283	240	211	265
290	292	269	290	235	256	288
285	278	240	211	281	245	325
306	310	269	208	324	279	289
249	248	272	239	292	290	281
240	289	261	311	239	269	306
316	251	314	252	241	282	280
249	263	252	237	276	280	275
305	257	278	275	262	297	231
257	308	287	278	285	274	294

До розд. VII.

До § 3. Матеріял для прикладу 1-го взято у Юровського.

До розд. VIII.

Вправа: визначити m для вищенаведених прикладів.

До розд. IX.

До § 2. Порівняти $M_1 \pm m_1$ з $M_2 \pm m_2$ в наступних прикладах:

1\. 5,5 6,0 6,5 7,0 7,5 8,0 8,5 9,0 мм.
1 9 55 84 48 22 3 і
6,5 7,0 7,5 8,0 8,5 9,0 9,5 10,0 мм.
1 10 38 130 137 62 5

2\. 13 18 23 28 33 38 43 48 пуд.
6 47 98 70 27 3 1 і
18 23 28 33 38 43 48 53 58 63 пуд.
1 6 30 115 146 65 14 12 2

3\. 0 3 6 9 12 15 18 21 24 27 30 зернин
2 6 19 20 36 40 59 35 7 1 і
0 3 6 9 12 15 18 21 24 27 30 зернин
1 6 14 27 37 77 115 79 25 6

4\. Порівняти дані про число народжень бичків та теличок:

	бичків	теличок	n
р. 1890	52,2%	47,8%	1507
„ 1891	49,3%	50,7%	1571

Приклад 8-й взято у Йогансена.

До розд. X.

Із усього сказаного в розд. X та із звітів досвідних станцій видко, оскільки є „хистка" методика польових спроб у більшості наших досвідних установ; на жаль, поки що лише небагато з них зважають на вимоги варіяційної статистики. Вказівок на звичайну „середню“ точність, що їх зустрічаємо у працях багатьох дослідувачів, як тепер ми бачимо, є надто замало, а такі твердження, як, напр.: „обробка пара чотирьохлемішником збільшує врожай на 1 пуд., порівнюючи з обробкою плоскорізом“, є зовсім нікчемні, коли мати на оці, що спроби ведено в найкращому разі при подвійній повторності.

Взагалі, що менша є сподівана різниця в урожаях з тих чи инших технічних засобів або сортів, то більшою повинна бути повторність спроби: як що при значних різницях (напр., серпневий та жовтневий засів озимини) звичайно доволі двох повторень, то при малих різницях (напр., густота засіву в 4 і 5 пуд. то-що) повторність треба збільшити в багато разів, бо, инакше, точність спроби буде остільки мала, що спроба не дасть нічого, і даремно тільки пропаде час, праця і кошти, витрачені на цю спробу. Користуватися з таких виводів, без перевірки їх надійности, є ризиковано, радити-ж селянству вживати цих виводів це часто є те саме, що ошукувати його і дискредитувати досвідну справу.

Що до методу зведення (§ 2, розд. X), то треба завважити таке: умови його вживання випливають із самої сути нашого основного припущення, а саме, що врожаї всіх повторень даного питання змінюються в одне й те саме число k разів, що, таким чином, v залишається незмінним для всіх питань; неточність буде тоді, коли це основне припущення не має місця. Однак у польових дослідах звичайно межі коливань v остільки є малі порівнюючи з загальною різноманітністю поля, що на практиці дуже рідко доводиться на це зважати.

Напр. З даних часткового обліку Безенчуцької досвідної станції v для поля озимої пшениці дорівнювало $\pm 10{,}3 \pm 0{,}5\%$, для ярої пшениці $\pm 12{,}6 \mp 0{,}6\%$. Різниця є невелика, хоч озима пшениця йшла по угноєному пару, а яра—в значно гірших умовах, так що різниця в середніх урожаю першої і другої пшениці досягала 75% (!), тоб-то була такою, яку ми зустрічаємо в спробах одної й тої самої групи питань дуже рідко. Аналогічні результати бувають і тоді, коли дуже велика повторність стандардної ділянки дозволяє обчислити σ для стандарду, як з v для всієї дільниці (способом зведення), так і безпосередньо.

Тому запропонований спосіб дає значно точніші σ і m, ніж обчислені звичайним шляхом, і при малій повторності спроби дозволяє точніше оцінювати числові результати її, ніж це поки що є можливо якимось иншим способом.

До розд. XI.

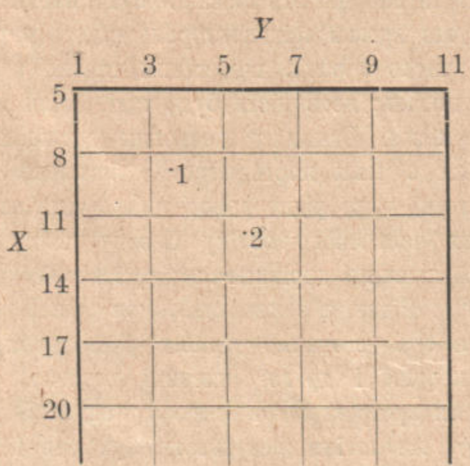
До § 1. Для правильного розуміння асиметричних кривих є дуже важливі Каптейнові праці.

До § 3. Для детальних зіставлень емпіричних кривих з теоретичними можна користуватися з „елементарного пособия к применению методов Пирсона к биологической статистике“ А. Леонтовича (Київ, 1911) або з підручника Лахтина: „Кривые распределения“, 1922.

До розд. XII.

До § 1. Розміщення варіантів у кореляційну мережу можна провадити так. Визначивши, як указано в § 3 розд. II, межі клас для обох зіставлюваних ознак, рисуємо відповідну мережу.

Тоді, спинившися на першому варіанті, відзначаємо згідно з величиною його X і Y точкою його місце в відповідній комірці мережі. Проробивши так з усіма варіантами, підраховуємо чисельності точок (варіантів) в кожній комірці, і мережа є готова. Напр., як-що схема мережі є така:



то варіант з $У = 4$ і $X = 11$ треба відзначити, як показано на схемі, точкою із знаком „1“, а варіант з $У = 7$ і $X = 13$ — точкою із знаком „2“.

До § 11. Про засоби, що їх уживають в цих випадках, можна довідатися в указаній книжці Леонтовича, а також у Е. Е. Слуцького: „Теория корреляции“ (Київ, 1915).

До § 12. Приклади до вправ.

1.

	55–60	60–65	65–70	70–75	75–80	80–85	85–90
73–78	1	2					
78–83		7	10				
83–88		2	34	18	1		
88–93			19	53	13	1	
93–98			2	5	35	6	
98–103					1	9	3

2.

	60–65	65–70	70–75	75–80	80–85	85–90	90–95
78–83	7	8					
83–88	1	21	17	1			
88–93		1	56	13			
93–98			12	36	5		
98–103				12	13		
103–108				2	14	7	
108–113						2	1

До розд. XIII.

До § 2. Наведену в 2-му прикладі формулу вивів професор математики Ю. Г. Рабинович, на підставі звязаности сусідніх ділянок поля (див. приклад 1-ий), яку встановив Сапігин. Подробиці можна найти в статті Сапігина: „Определение точности полевого опыта помощью элементов вариационной статистики“ (Одеса, „Изв. Обл. Упр. по опытному делу“, 1922).

При обчисленні r у випадку прикладів 1-го і 2-го, знаменик $\sigma_x\sigma_y = \sigma^2$, бо для варіювання родючости одного й того самого поля $\sigma_x = \sigma_y$.

Докладніше про обчислення поправки на розрідженість див. в статті тої-ж назви (за підписом А. А. Сапєгін і А. М. Левшин) в „Труд. Одесск. Селикц. Ст.", вип. VII, 1922; див. вип. X, 1925 і А. Сапігин: „Методика селекції", 1925.

Для дальшого теоретичного поглиблення у варіяційну статистику можна порадити в найпершу чергу „элементы биометрики" Левитського, згадану книгу Слуцького, підручники Лахтіна, Орженцкого та инш.

Спис символів та формул.

A—довільний нуль, вихідна точка при обчисленні M, σ та инш.

a—відхил варіяції від A.

a'—відхил варіяції від A при класовій проміжці, взятій за 1.

α—відхил варіяції від M.

b—середній відхил A від M.

β—середній відхил A від M при класовій проміжці, взятій за 1.

D—відхил варіянту від M.

d—відхил зведеного відсоткового варіянту від $M = 100\%$.

E—покажчик ексцесу кривої.

M—середнє аритметичне.

m—серединна похибка.

n—число всіх варіянтів ряду.

P—покажчик точности дослідження.

p—чисельність варіяції або альтернативи.

Q—квартиль, або ймовірний відхил.

R—сучинник регресії.

r—сучинник кореляції.

Σ—знак суми.

σ—основний відхил.

s—основний відхил при класовій проміжці, взятій за 1.

S—покажчик асиметричности (перекошености) кривої.

V—числова вартість варіянту.

v—варіяційний сучинник.

W—числова вартість варіяції.

w—зведена відсоткова вартість варіянту.

X і Y числові вартості 2 зіставлених ознак варіянту.

x і y відхили 2 зіставлених ознак від своїх M.

$$A = M - b;\ b = \frac{\Sigma pa}{n};\ \beta = \frac{\Sigma pa'}{n}.$$

$$E = \frac{\Sigma p\alpha^4}{n} : \sigma^4 - 3 = \left(\frac{\Sigma pa'^4}{n} - \frac{4\beta\Sigma pa'^3}{n} + \frac{6\beta^2\Sigma pa'^2}{n} - 3\beta^4\right) : s^4 - 3.$$

$$M = \frac{\Sigma V}{n} = \frac{\Sigma pW}{n} = A + b.$$

$$m = \frac{\sigma}{\sqrt{n}} = \sqrt{\frac{\Sigma p\alpha^2}{n^2}};\ m_r = \frac{1 - r^2}{\sqrt{n}};\ m_\sigma = \frac{\sigma}{\sqrt{2n}}.$$

$$P = \frac{v}{\sqrt{n}} = \frac{100m}{M} = \frac{100\sigma}{M\sqrt{n}}.$$

$$R_{\frac{x}{y}} = r\frac{\sigma_x}{\sigma_y};\ R_{\frac{y}{x}} = r\frac{\sigma_y}{\sigma_x}.$$

$$r = \frac{\Sigma D_x D_y}{n\sigma_x\sigma_y} = \frac{\Sigma p\alpha_x\alpha_y}{n\sigma_x\sigma_y} = \frac{\Sigma pa'_x a'_y - n\beta_x\beta_y}{ns_x s_y}.$$

$$\sigma = \pm\sqrt{\frac{\Sigma D^2}{n}} = \pm\sqrt{\frac{\Sigma p\alpha^2}{n}} = \pm\sqrt{\frac{\Sigma p\alpha^2}{n} - b^2} = \pm\sqrt{\frac{\Sigma V^2}{n} - M^2}.$$

і для альтернативного варіювання: $\sigma = \pm\sqrt{{}^0\!/\!_0 p_1 \cdot {}^0\!/\!_0 p_2}\ {}^0\!/\!_0$.

$$s = \pm\sqrt{\frac{\Sigma pa'^2}{n} - \beta^2}.$$

$$S = \frac{\Sigma p\alpha^3}{n} : \sigma^3 = \left(\frac{\Sigma pa'^3}{n} - \frac{3\beta\Sigma pa'^2}{n} + 2\beta^3\right) : s^3.$$

$$v = \frac{100\sigma}{M}.$$

ЗМІСТ

Помилки та поправки.

Стор.	Рядок	Надруковано	Треба
20	1 знизу	$\sqrt{\frac{\Sigma a^2}{n}} - b^2.$	$\sqrt{\frac{\Sigma a^2}{n} - b^2}.$
20	6 знизу	$\pm\sqrt{\frac{\Sigma V^2}{n}} - M^2$	$\pm\sqrt{\frac{\Sigma V^2}{n} - M^2}$
21	13 зверху	$n = p^0 + p_1.$	$n = p_0 + p_1.$
35	11 зверху	$+\sigma_n.$	$=\sigma_n.$
41	12 знизу	часу	поля
51	20 зверху	$M_2 = M_1 \pm 3m\sqrt{2 - 0{,}0009P^2}$	$M_2 = \frac{M_1 \pm 3m_1\sqrt{2 - 0{,}0009P^2}}{1 - 0{,}0009P^2}$
59	8 знизу	$V.$	$U.$
61	15 зверху	$n - 500.$	$n = 500.$

Підручна література для Соцвиху.

Бернашевський і Звягинцев. Віки та людська праця. . . . . Ц. 70 коп.
Горовий О. Працею та знанням. Нариси й оповідання для дітей з історії матеріяльної культури й техніки. . . . . . . Ц. 75 коп.
Григор'їв М. Найголовніші речі щоденного вжитку. Вироб та властивості продуктів широкого споживання . . Ц. 1 крб. 50 коп.
Клодницький І. Тварина в господарстві людини. . . . . . . Ц. 80 коп.
Соколовський-Москвичів. Молодий хлібороб. Робоча книга з сільського господарства . . . . Ц. —

ПІДРУЧНИКИ ДЛЯ ПРОФШКІЛ.

Мова. Література.

Горецький і Шаля. Курс української мови. . . . . . Ц. 1 крб. 50 коп.
Дорошкевич О. Підручник української літератури. . . Ц. 1 крб. 80 коп.
Гаєвський С. Теорія поезії. . . . . Ц. 85 коп.

Сільське господарство. Економіка.

Колкунов. Загальне рослинознавство . . . . Ц. —
Кельнер. Основи годівлі. . . . . Ц. 2 крб. 20 коп.
Вовк П. Сільсько-господарське машинознавство. Ч. І. . . . . Ц. 75 коп.
„ „ „ „ Ч. ІІ. . . . . Ц. —

Підручна література для профшкіл.

Кисіль О. Український театр. . . . . Ц. 1 крб. 50 коп.
Навроцький Б. Мова та поезія . . . . Ц. 2 крб. —

Педагогіка. Методика.

Ананьїн С. Трудове виховання. . . . . Ц. 30 коп.
Чепіга Я. Практична трудова педагогіка. . . . . Ц. 25 коп.
Дога В. Навчання грамоти. . . . . Ц. 90 коп.

С К Л А Д   В И Д А Н Н Я.

Харків, Горяїнівський пр. 2; Київ, ул. Короленка 46; Одеса, ул. Лассаля 12.

Ціна 1 крб. 25 коп.

СКЛАД ВИДАННЯ:
КНИГОСПІЛКА
Київ, вул. Короленка 46
Харків, Горяїнівск, пр. 2