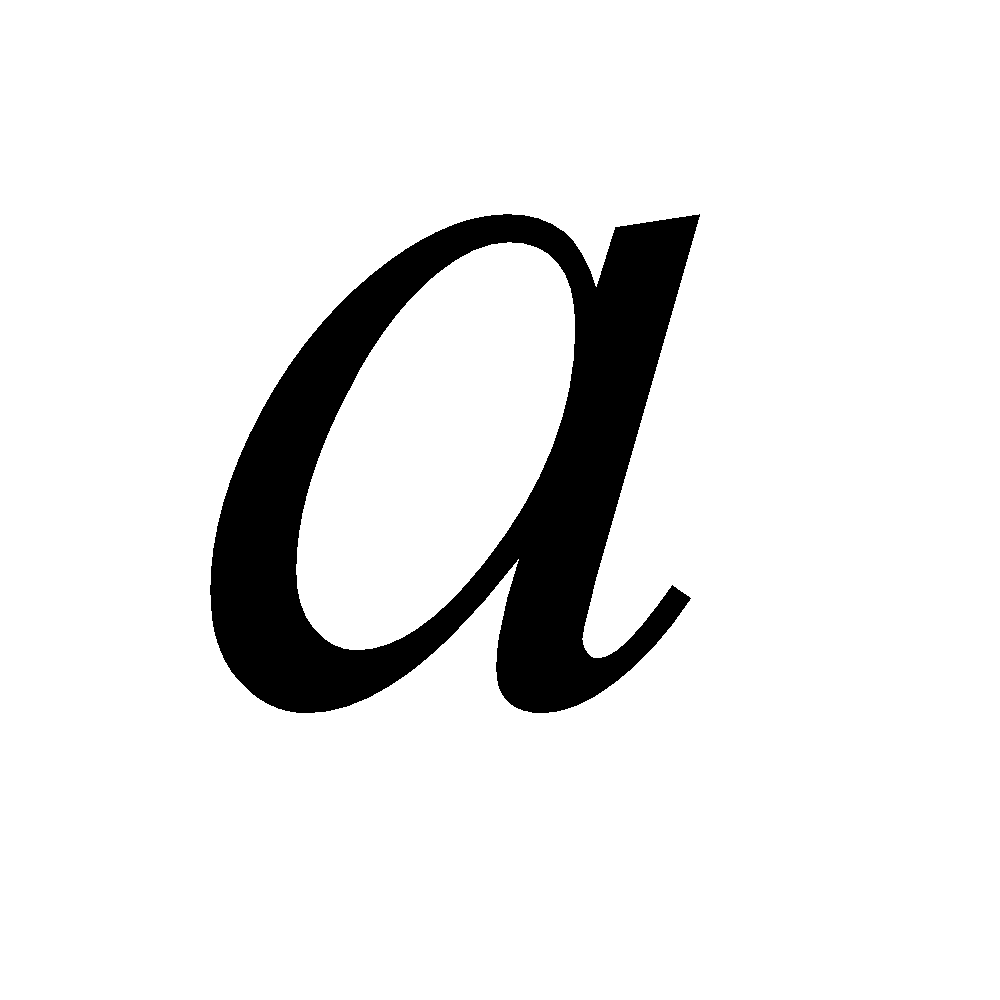
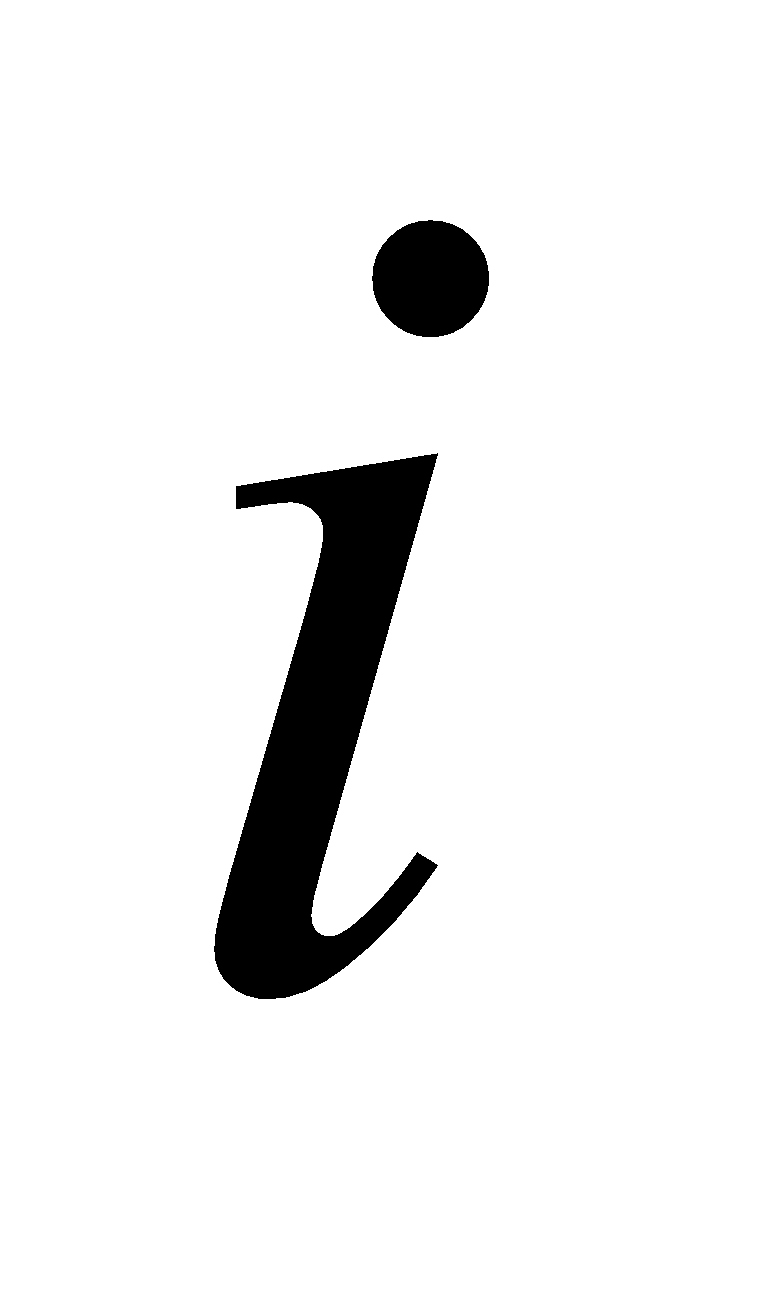
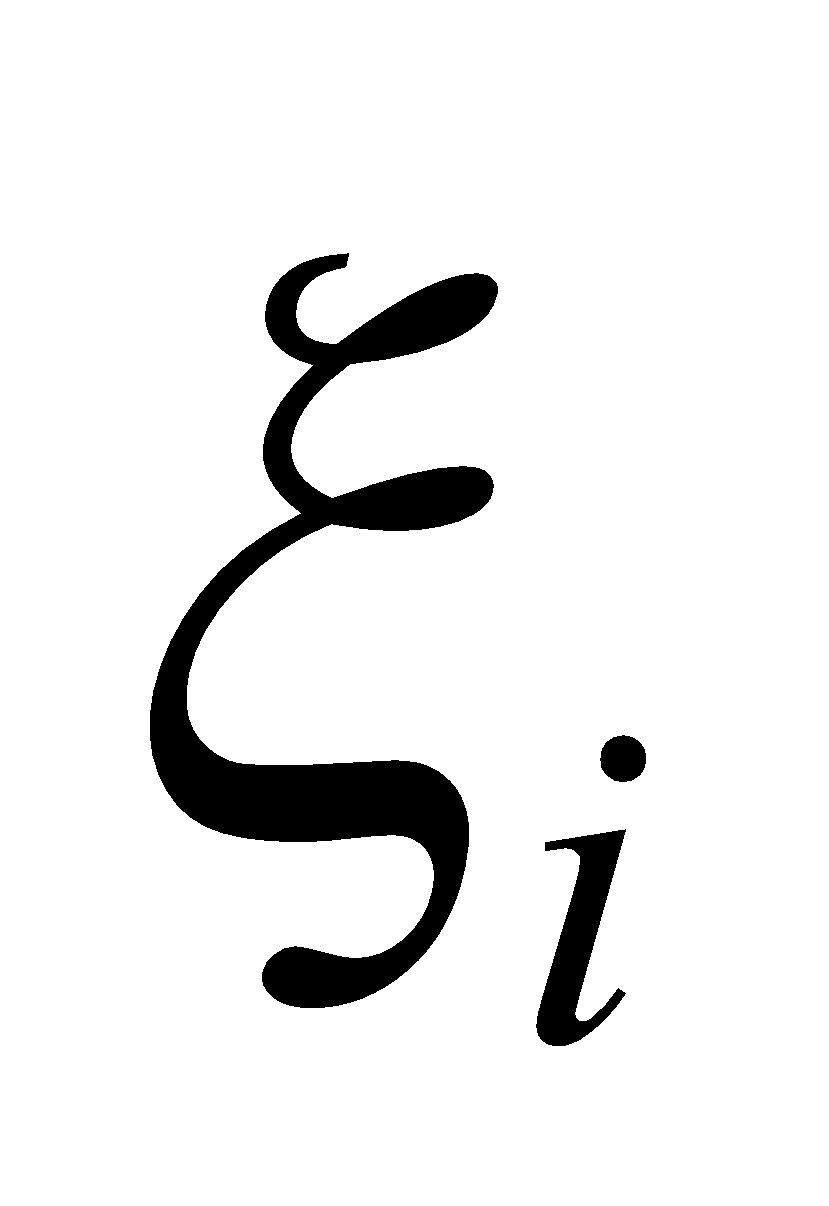
**Розділ 4**

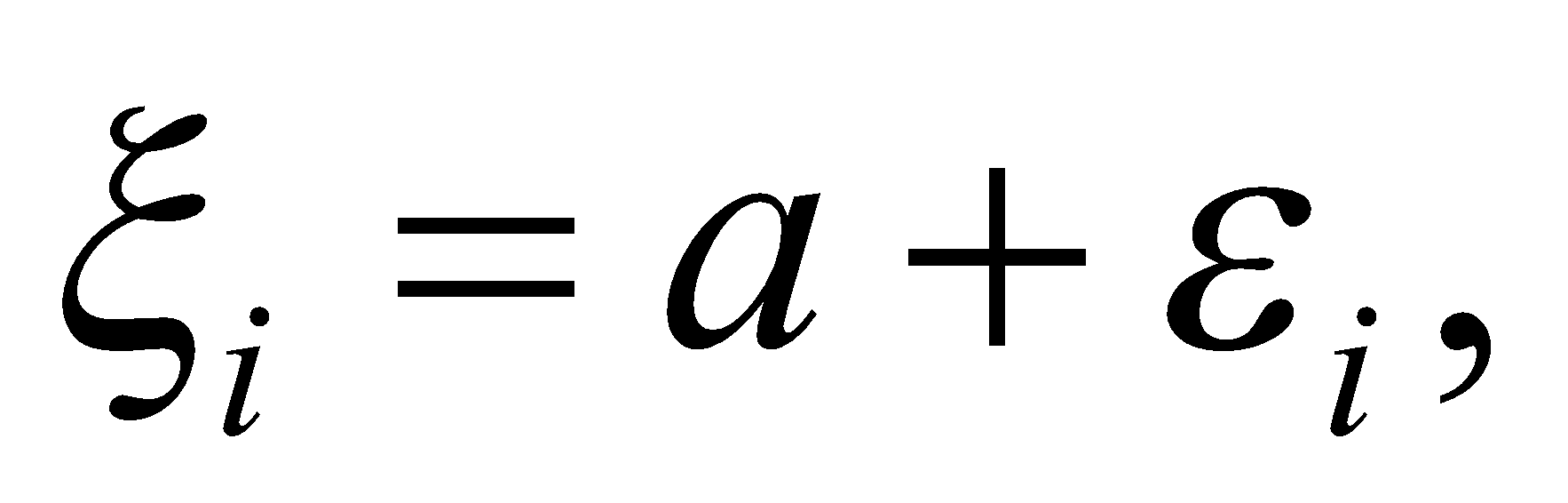
**ПЕРЕВІРКА СТАТИСТИЧНИХ ГІПОТЕЗ**

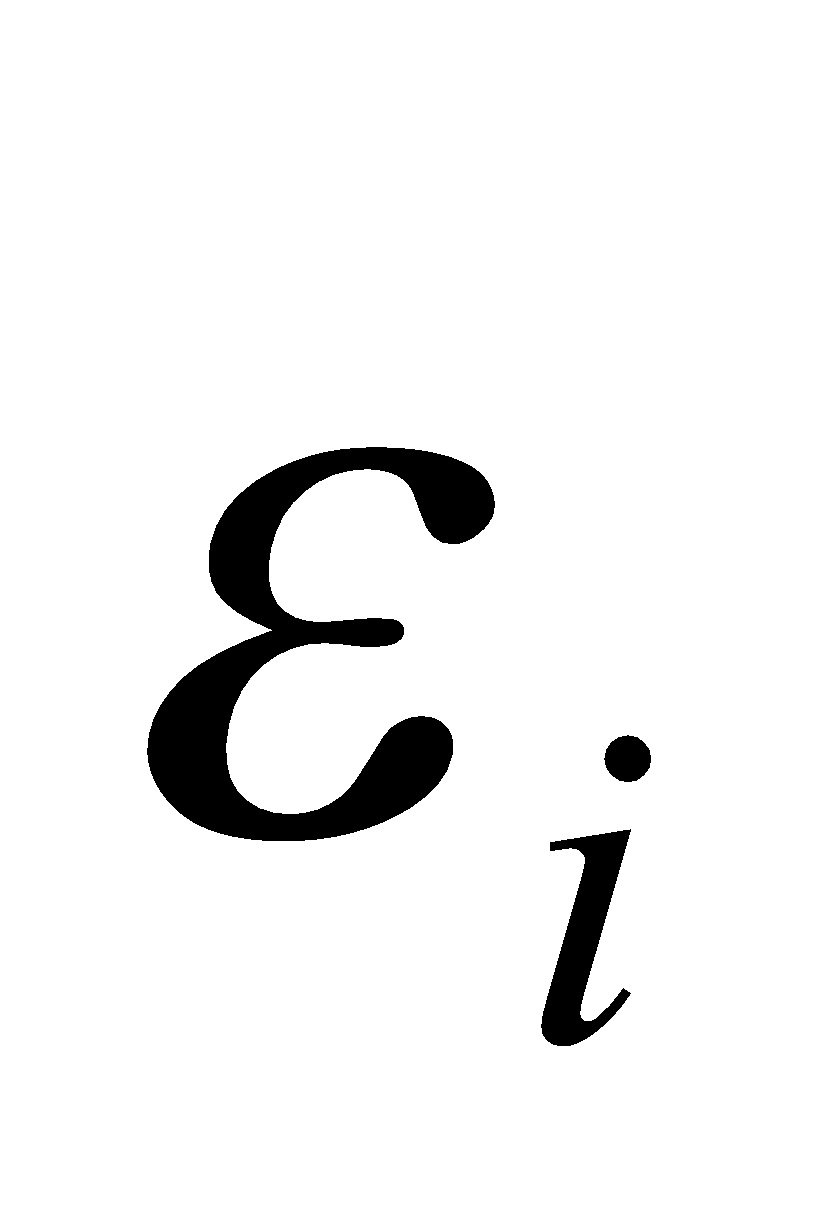
**4.1. Поняття статистичної гіпотези і статистичного критерія**

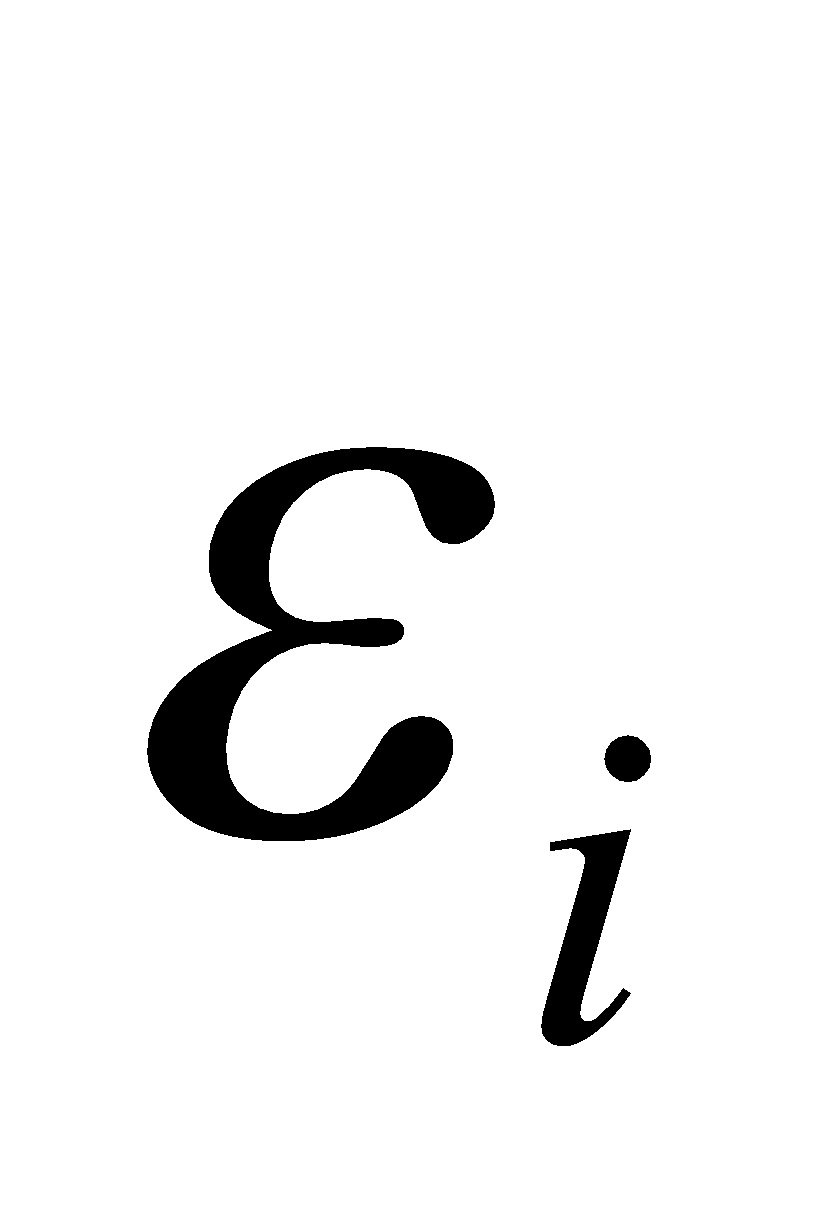
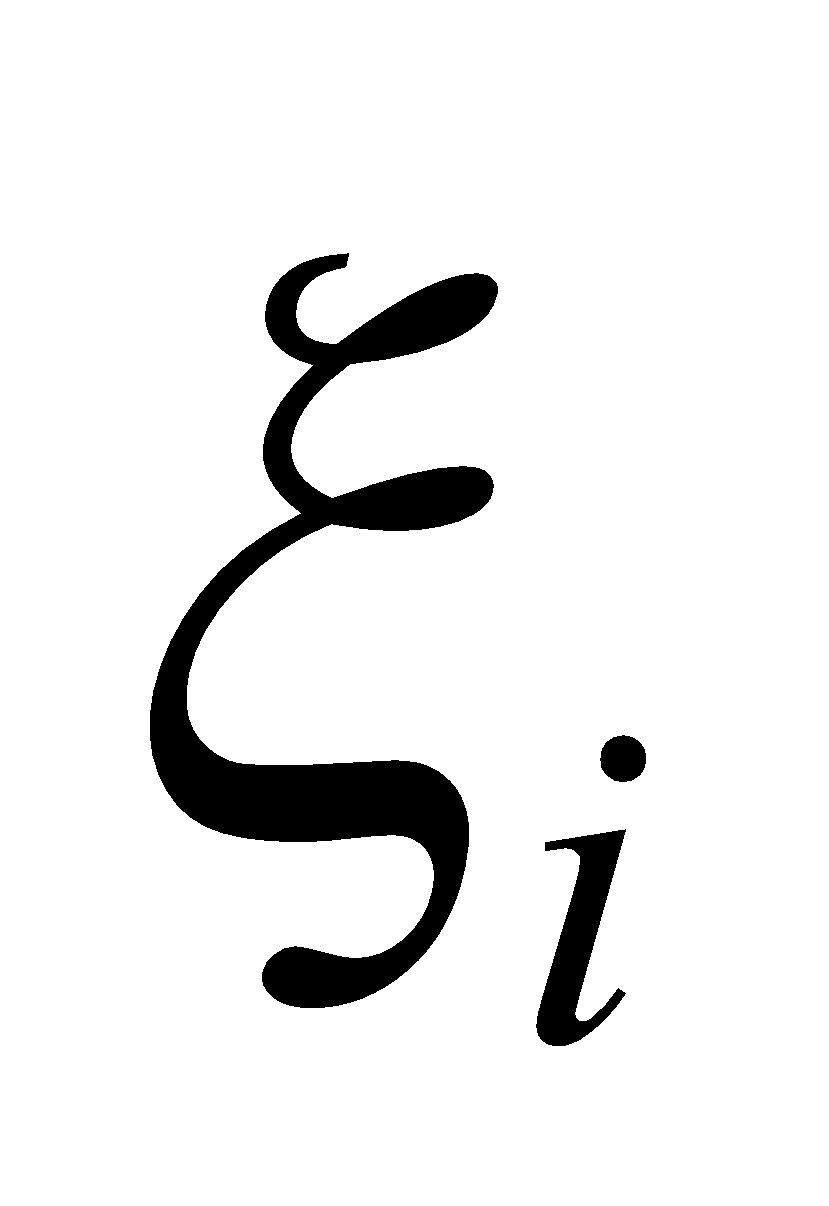
***Статистична гіпотеза*** – це довільне твердження про вид або властивості розподілів випадкових величин, що спостерігаються в експерименті.

**Приклад 4.1.** Нехай експеримент полягає у багаторазовому вимірюванні деякої фізичної величини, точне значення якої “” невідоме і в процесі вимірювань не змінюється.

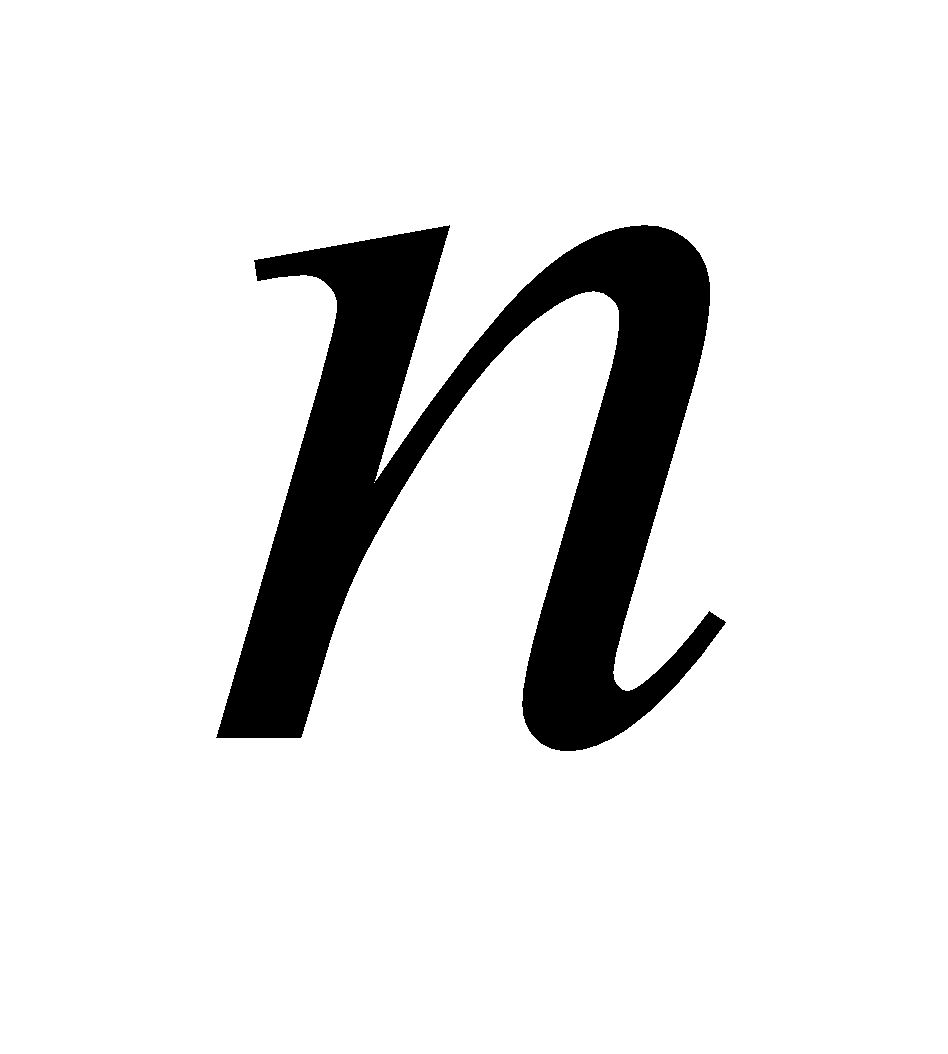
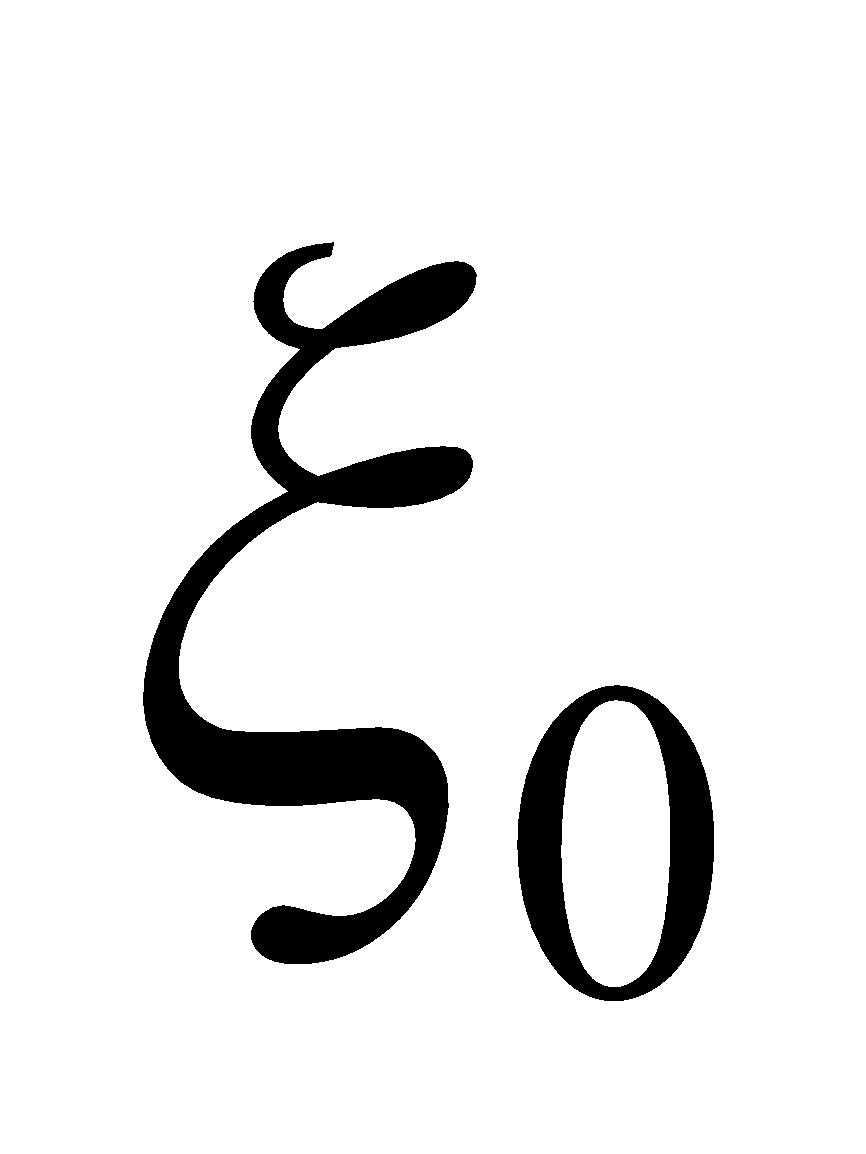
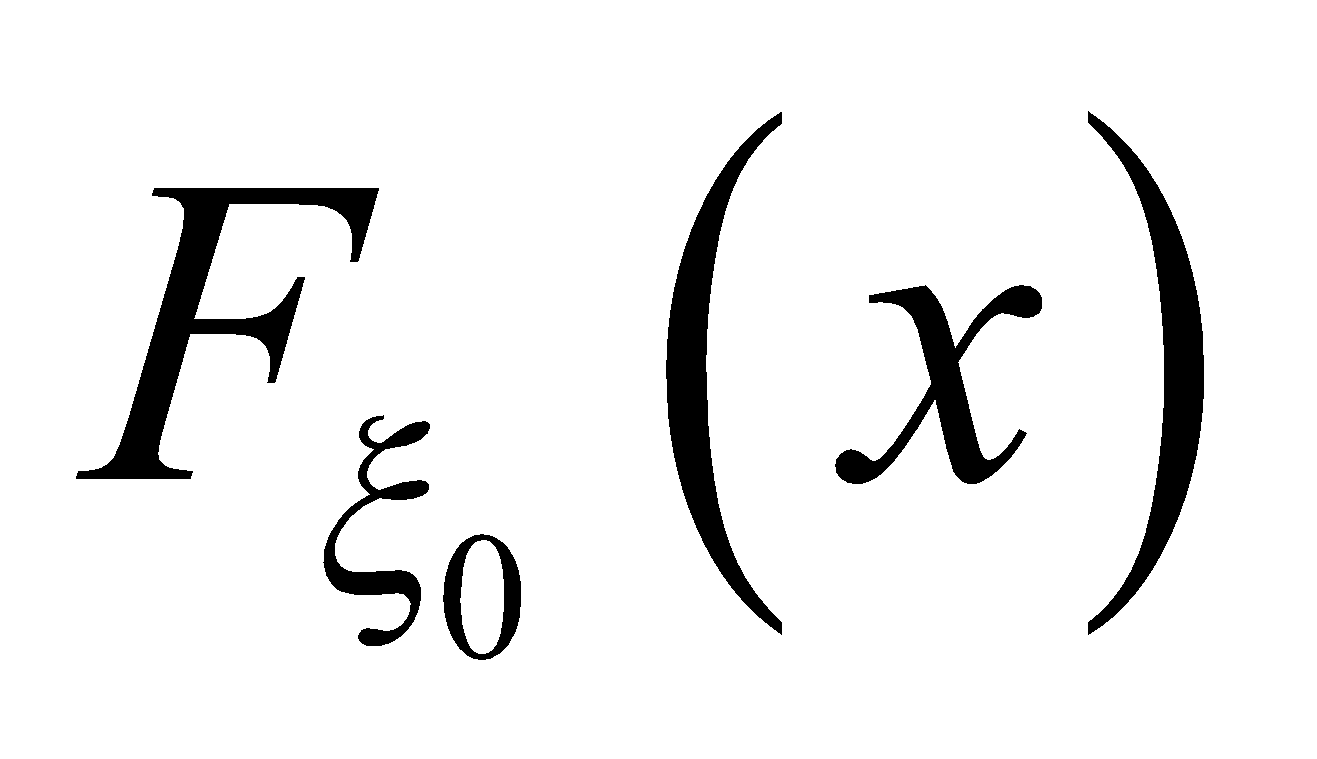
На результати вимірювань впливають багато факторів: точність налагодження приладу, похибка заокруглення і т. д., тому результат - го вимірювання  можна записати у вигляді

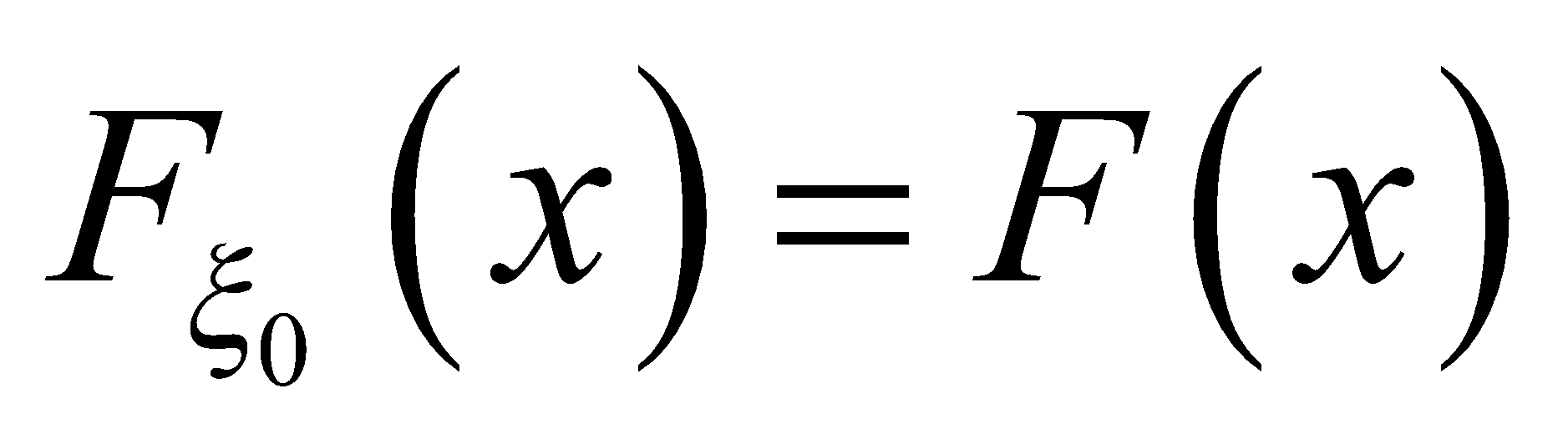
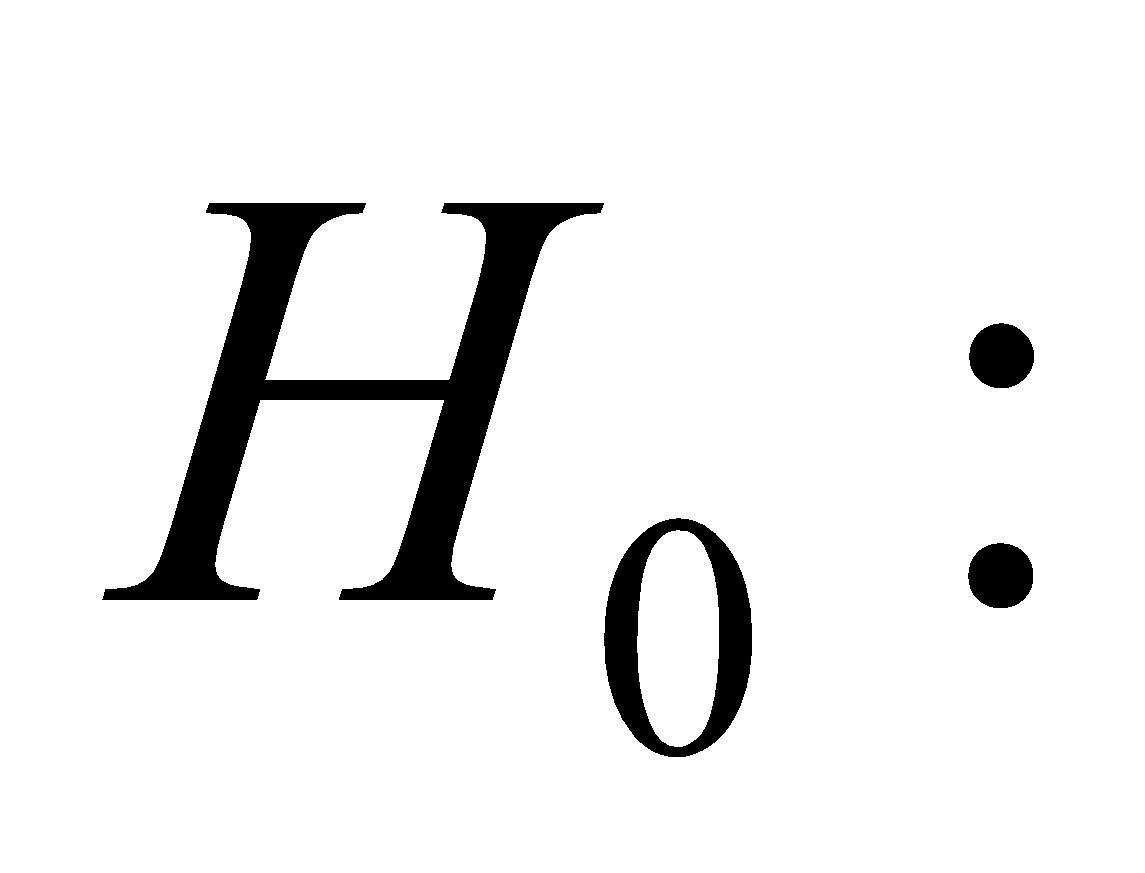


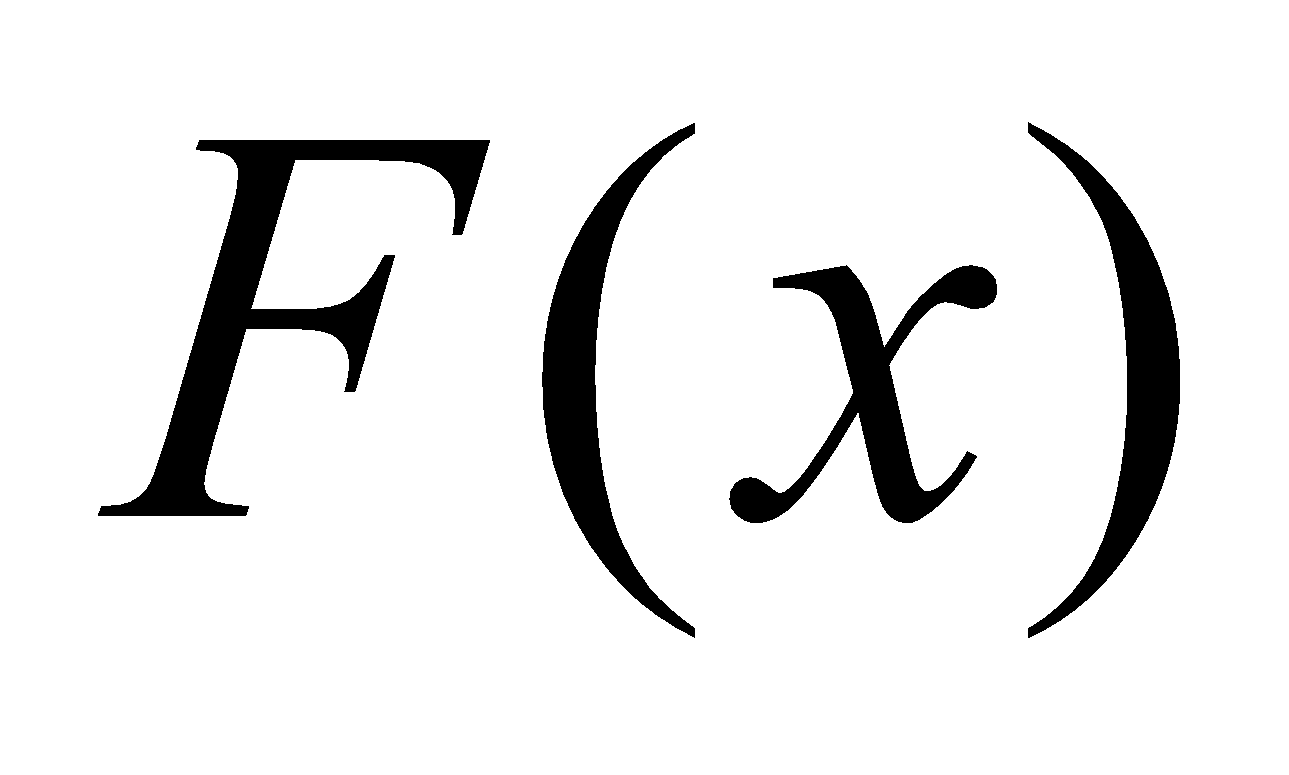
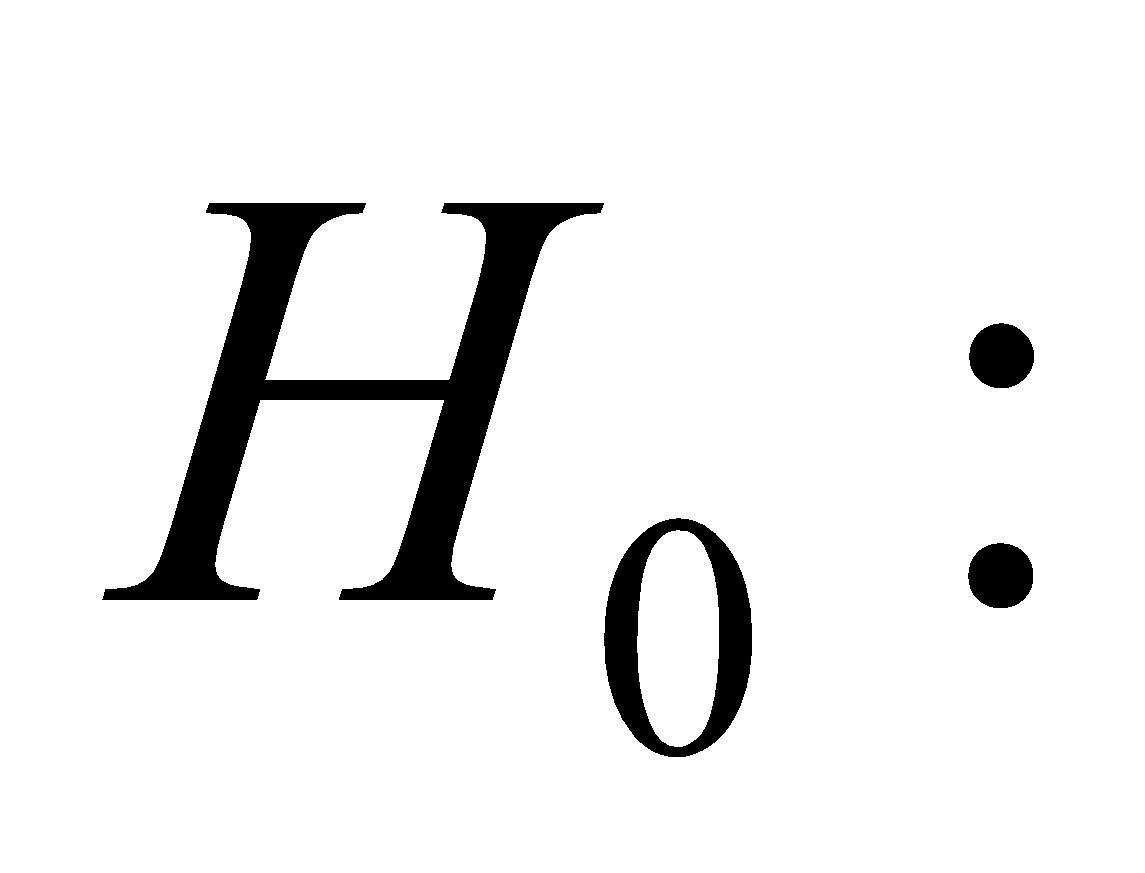
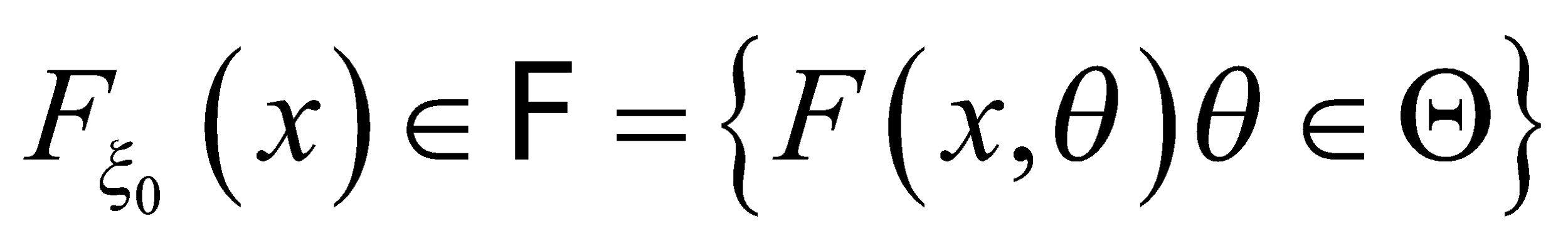
де  – випадкова похибка вимірювання.

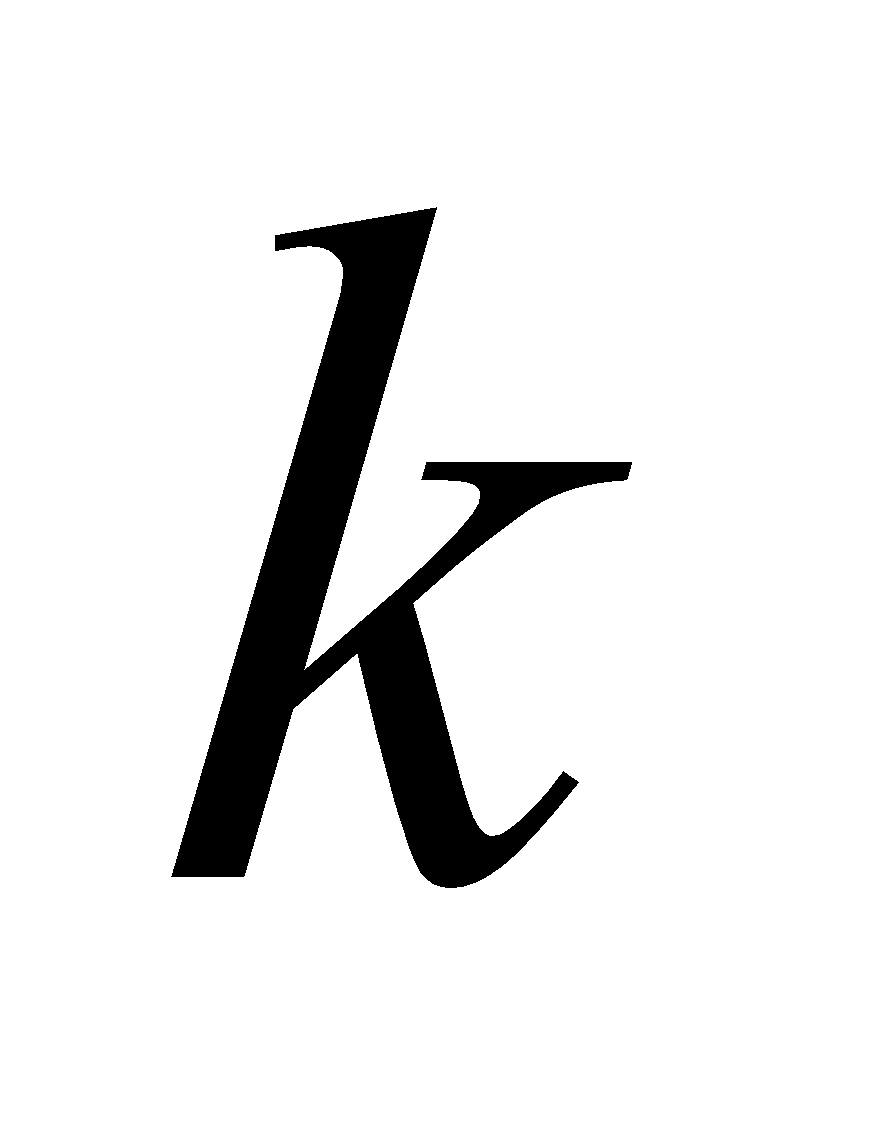
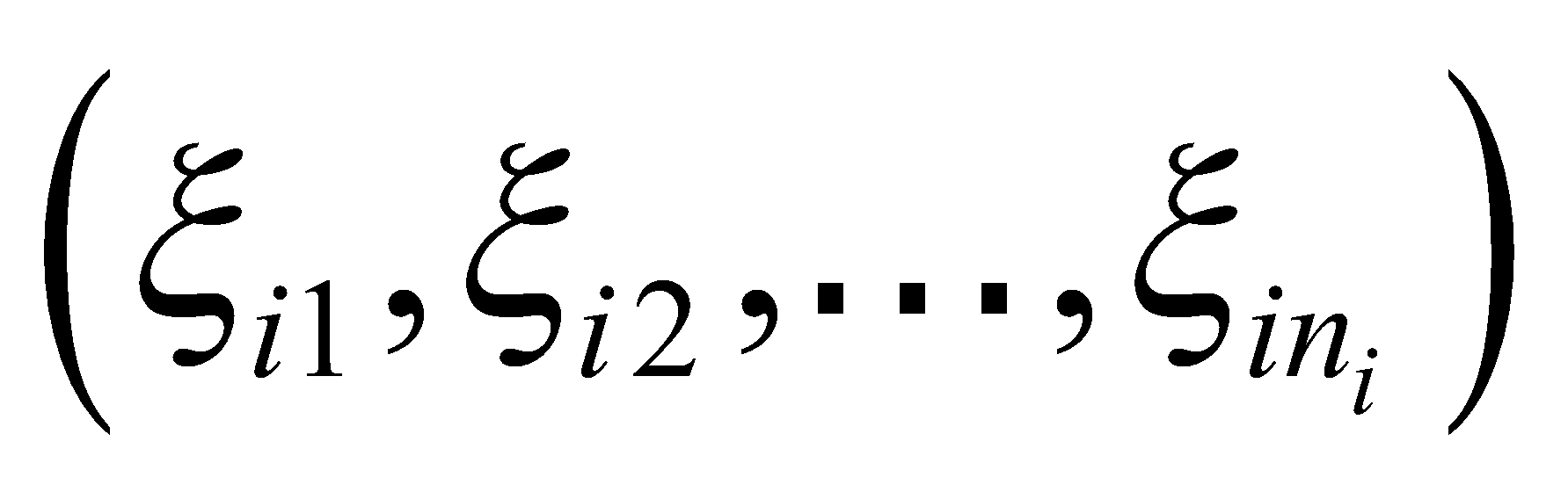
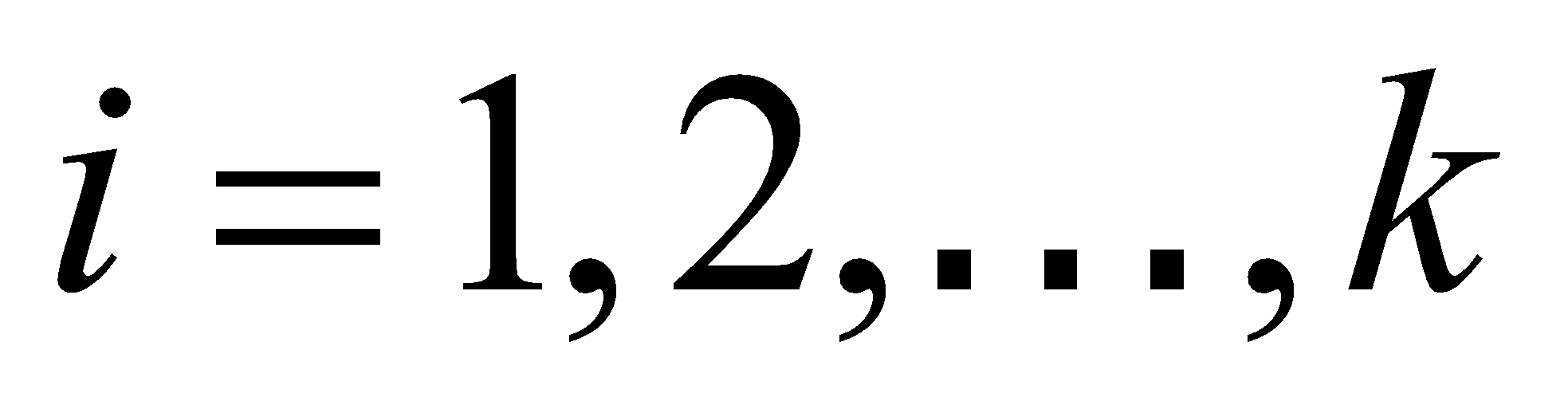
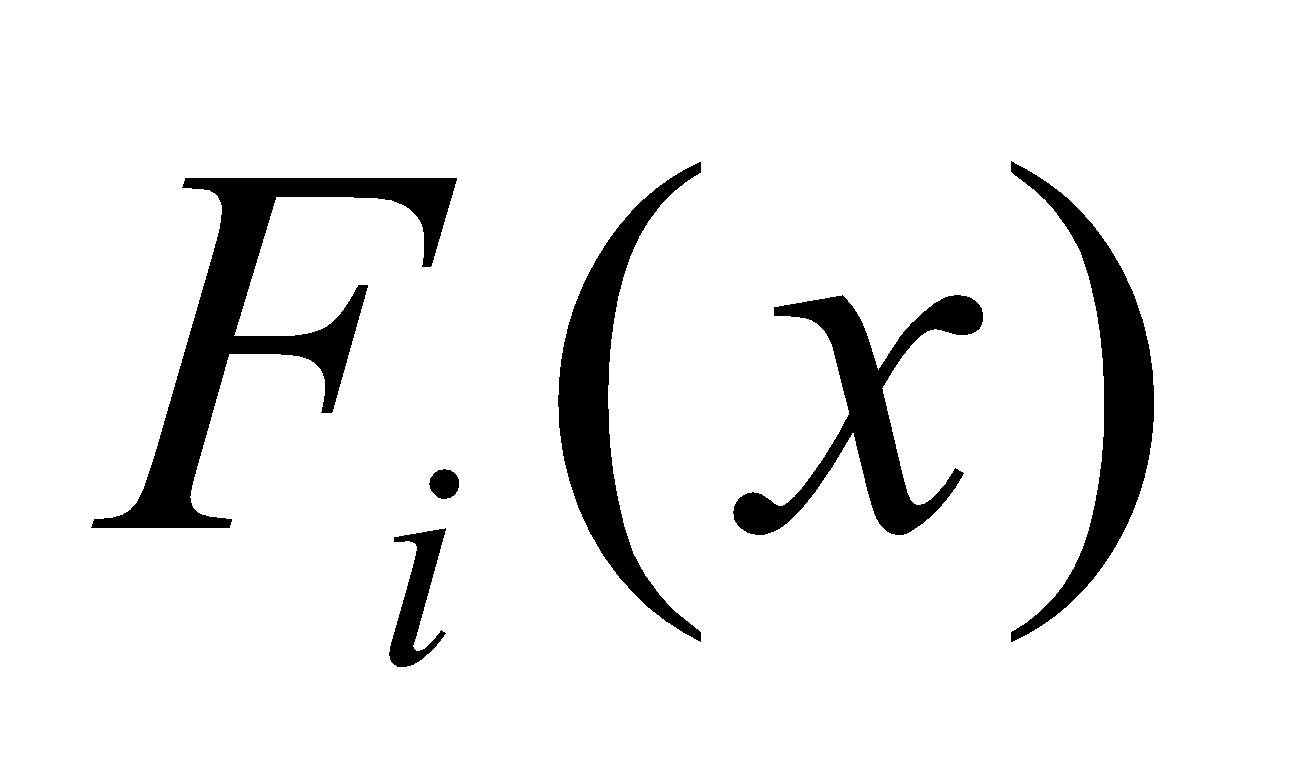
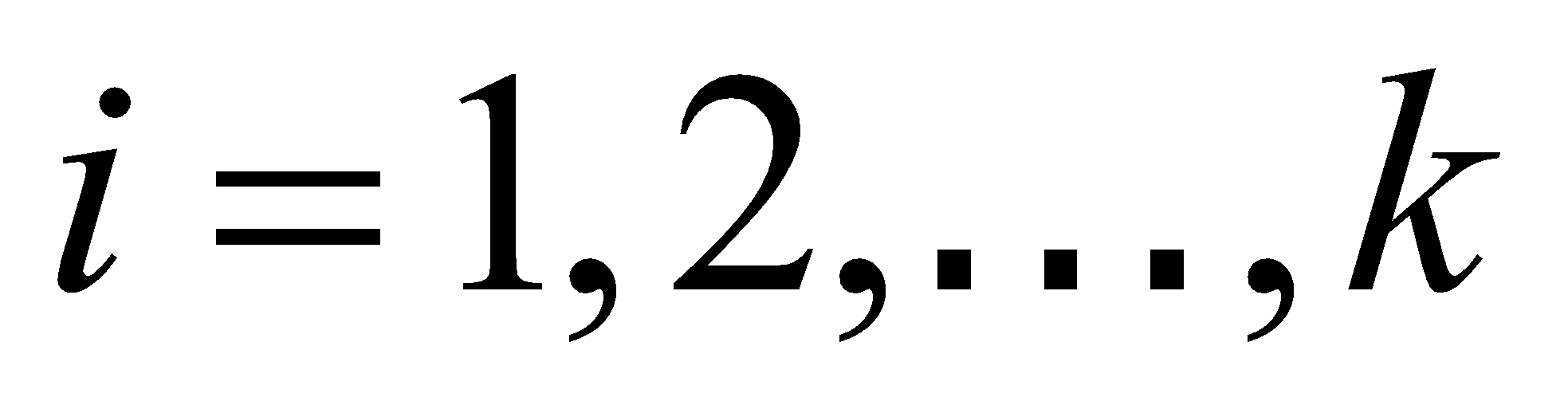
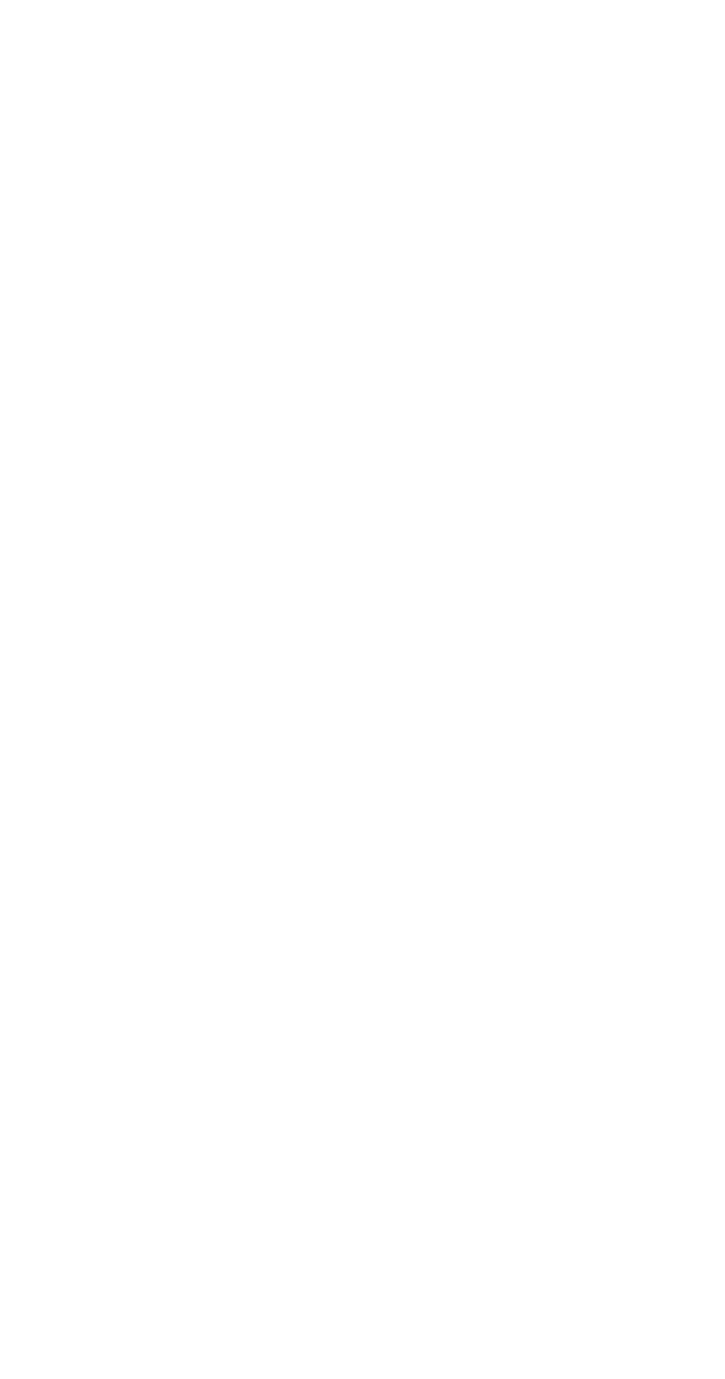
Будемо вважати, що загальна похибка  складається із великого числа похибок, кожна з яких невелика. На основі центральної граничної теореми припустимо, що випадкові величини  мають нормальний розподіл. Таке припущення є статистичною гіпотезою про вид розподілу випадкових величин, що спостерігаються.

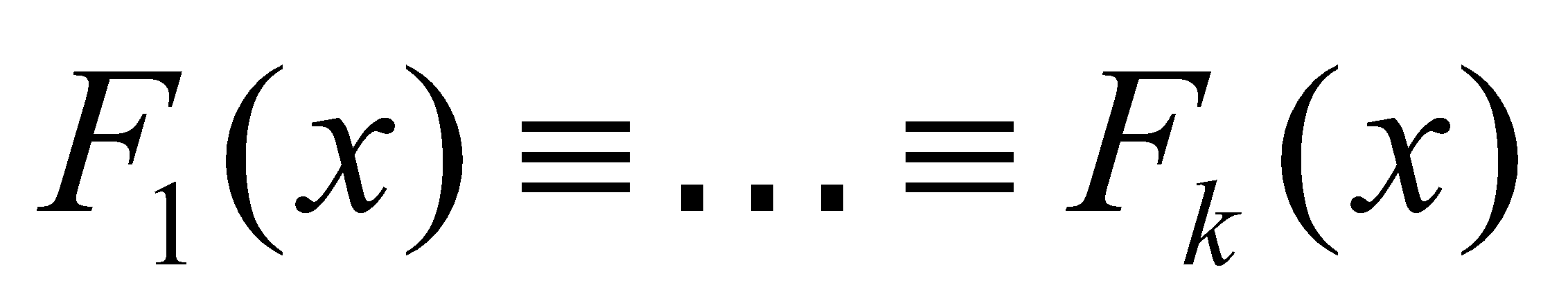
Наведемо декілька типів статистичних гіпотез.

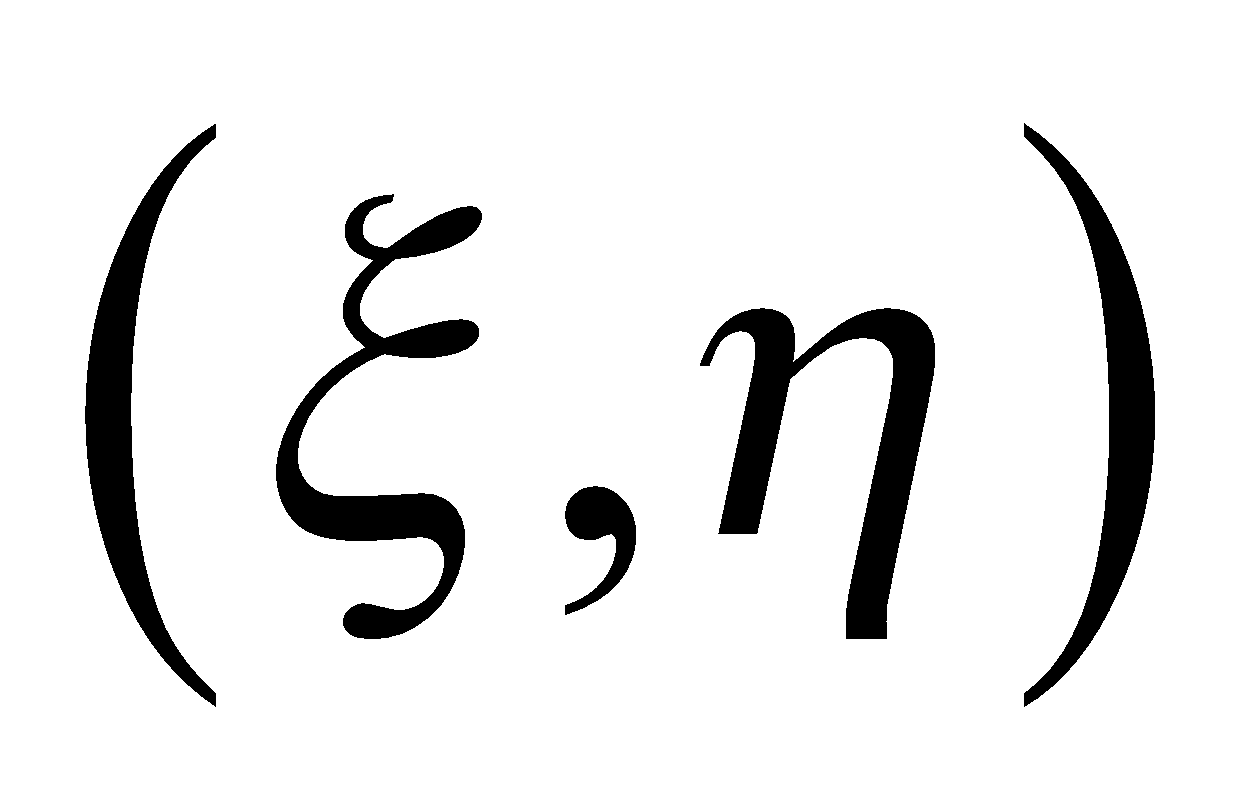
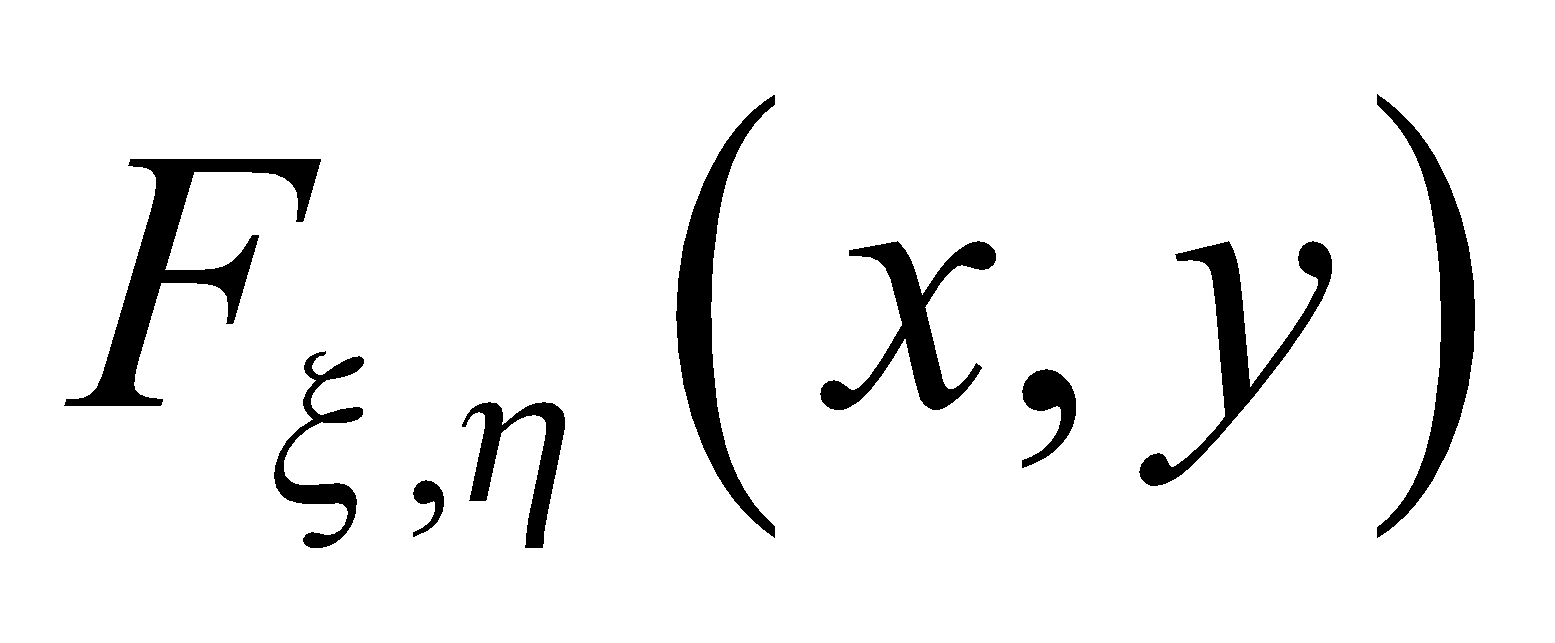
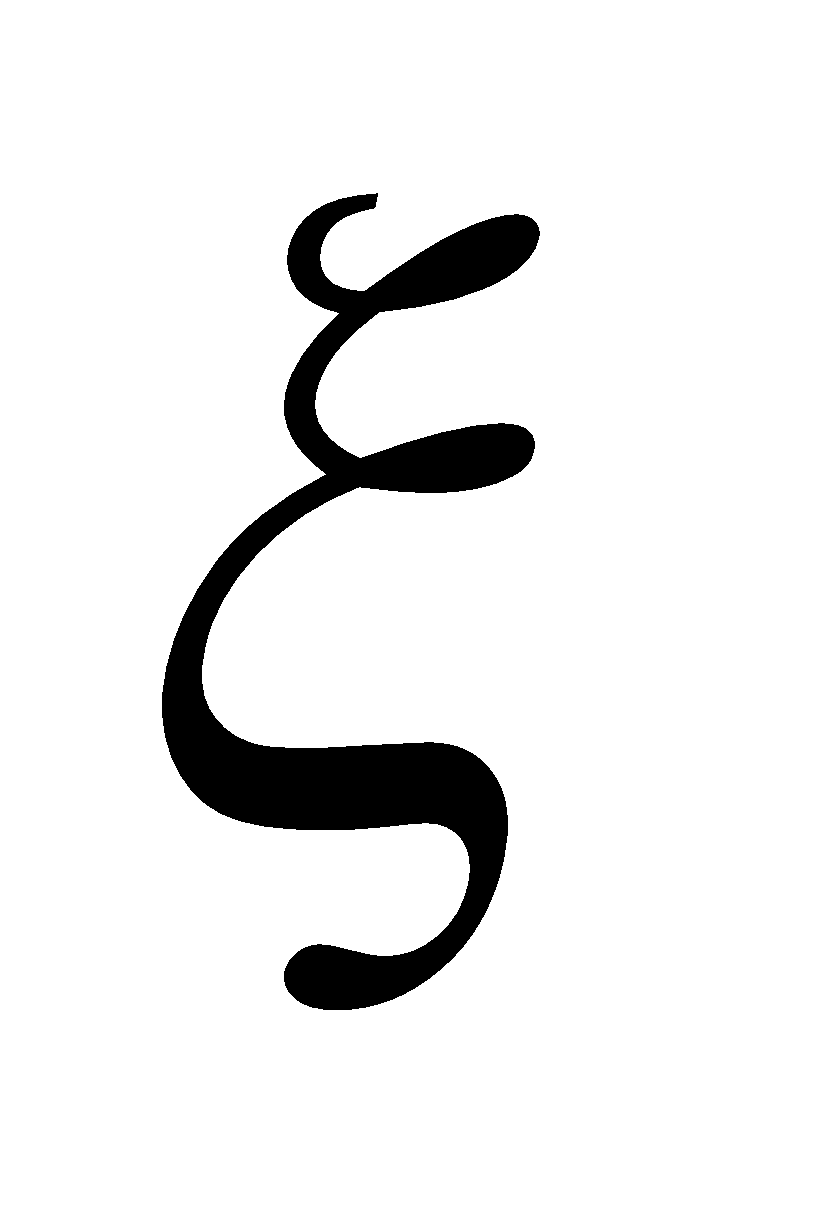
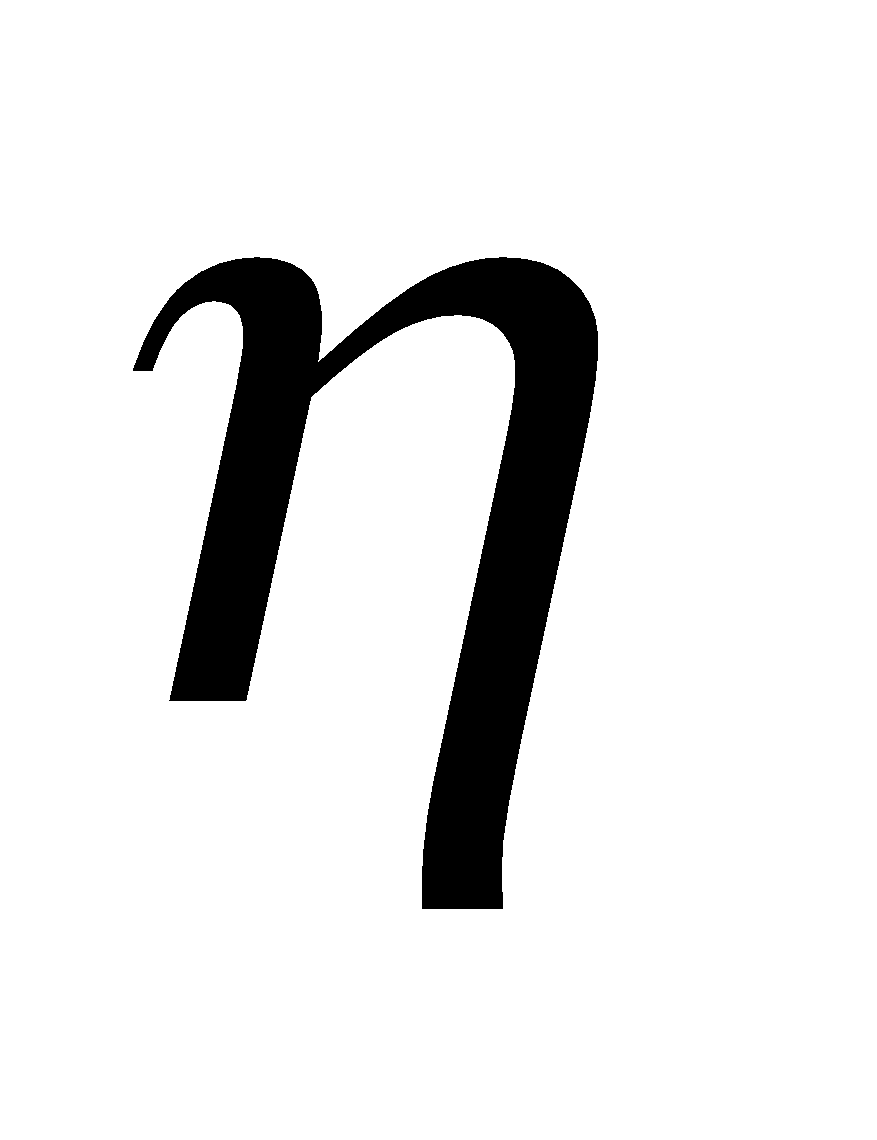
1. Гіпотеза про вид розподілу. Нехай проводиться “” незалежних спостережень над деякою випадковою величиною  з невідомою функцією розподілу . Гіпотеза, яка підлягає перевірці:

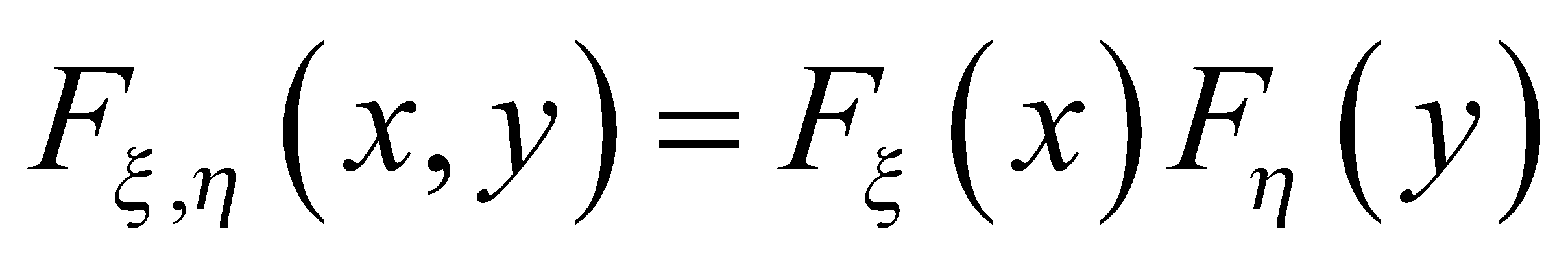
,

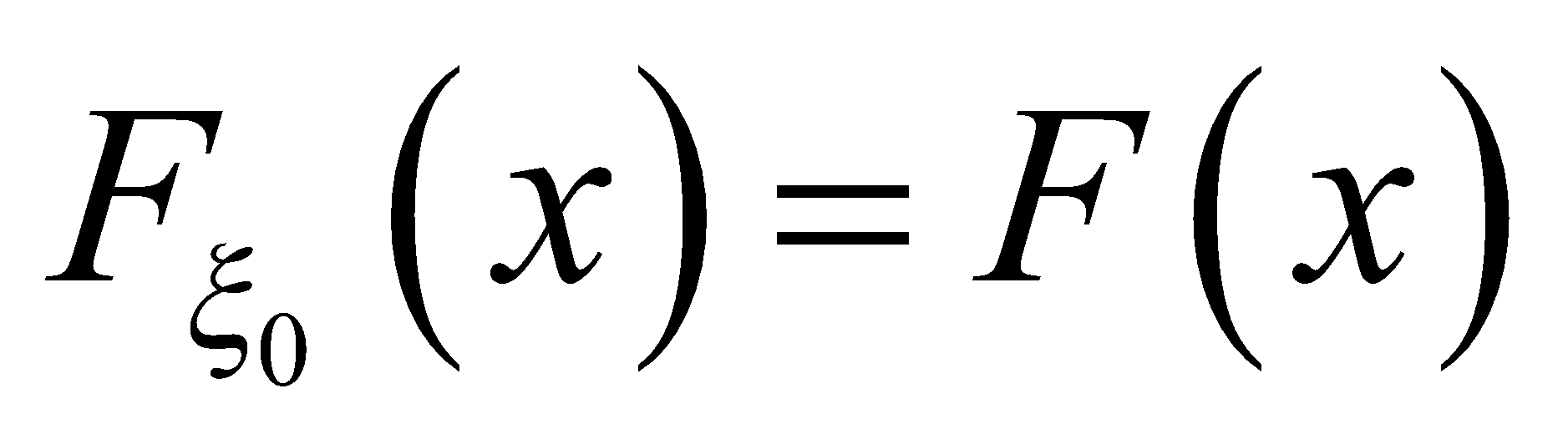
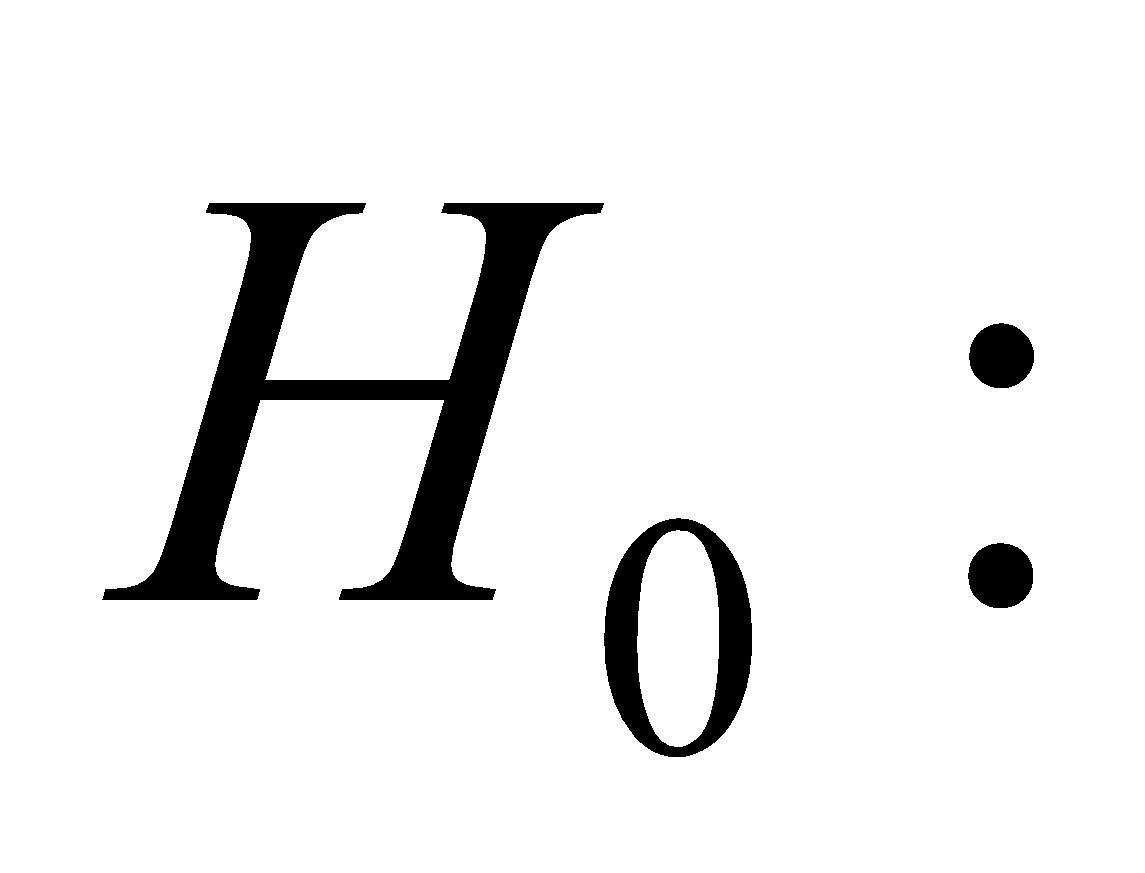
де функція  повністю задана, або   – задане сімейство функцій розподілу.

2. Гіпотеза однорідності. Нехай проведено  серій незалежних спостережень ,  з генеральних сукупностей з функціями розподілу ,  (взагалі кажучи невідомими). Чи є підстава розглядати ці дані як результати спостережень над одною і тою ж самою випадковою величиною? Якщо це так, то говорять, що статистичні дані однорідні. Відповідно перевіряється гіпотеза однорідності

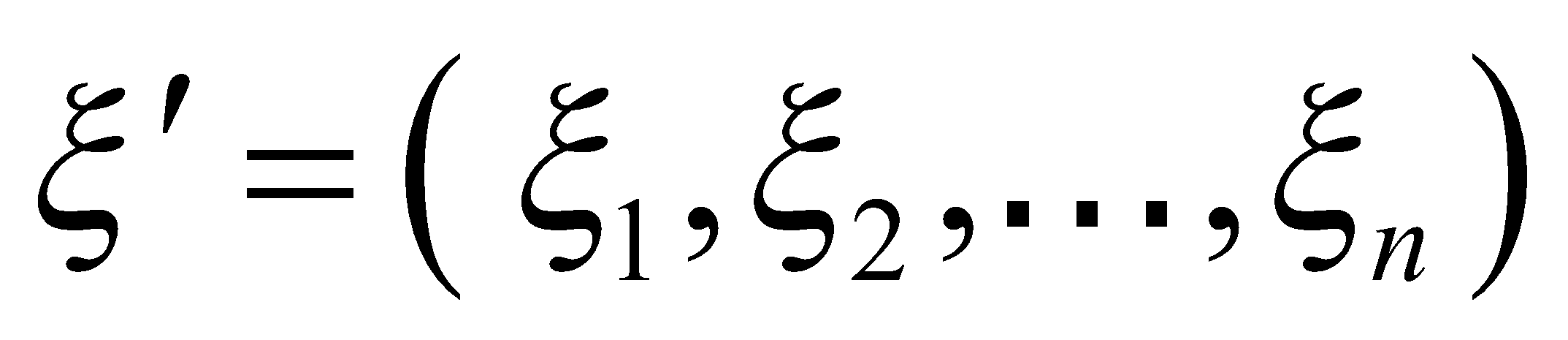
: .

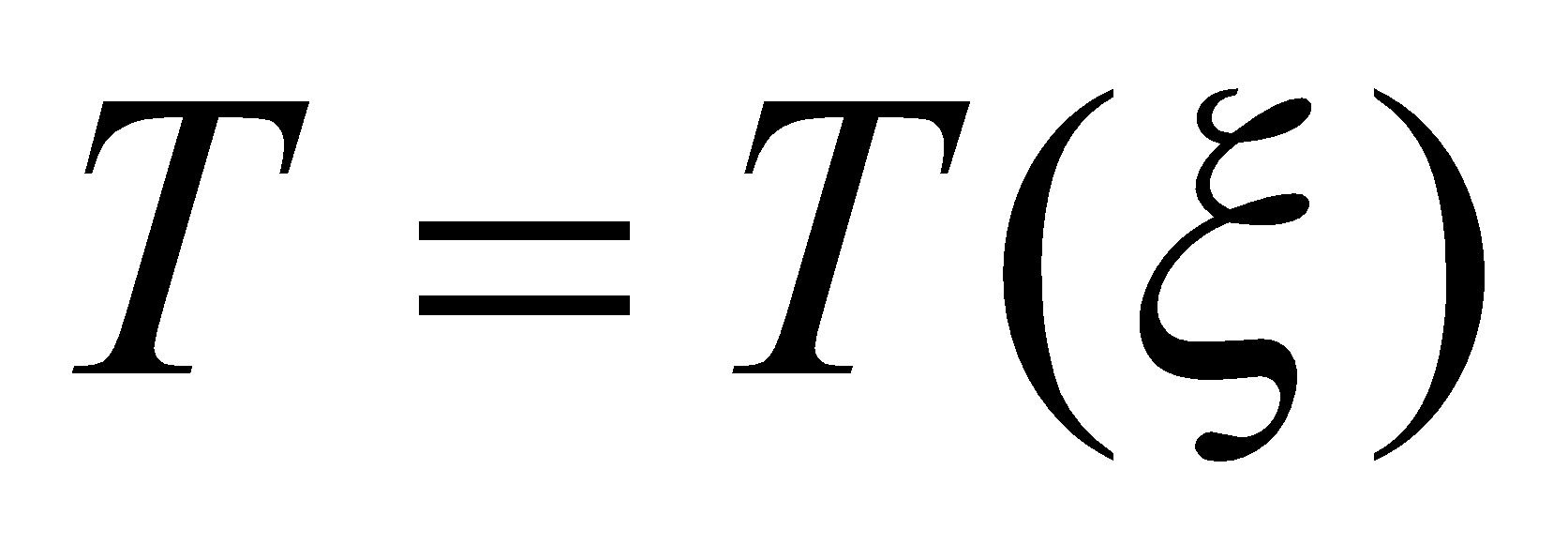
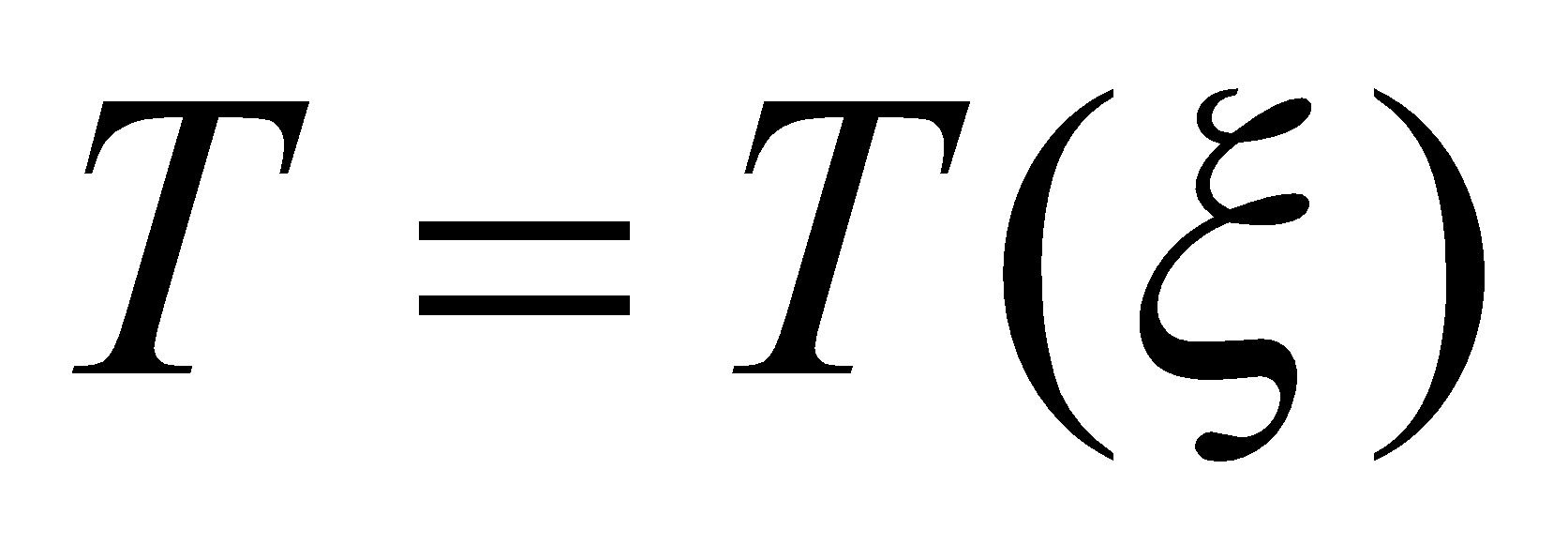
3. Гіпотеза незалежності. У експерименті спостерігається двовимірна випадкова величина  з невідомою сумісною функцією розподілу  і є підстава вважати, що компоненти  і  незалежні. В цьому випадку треба перевірити гіпотезу незалежності, тобто

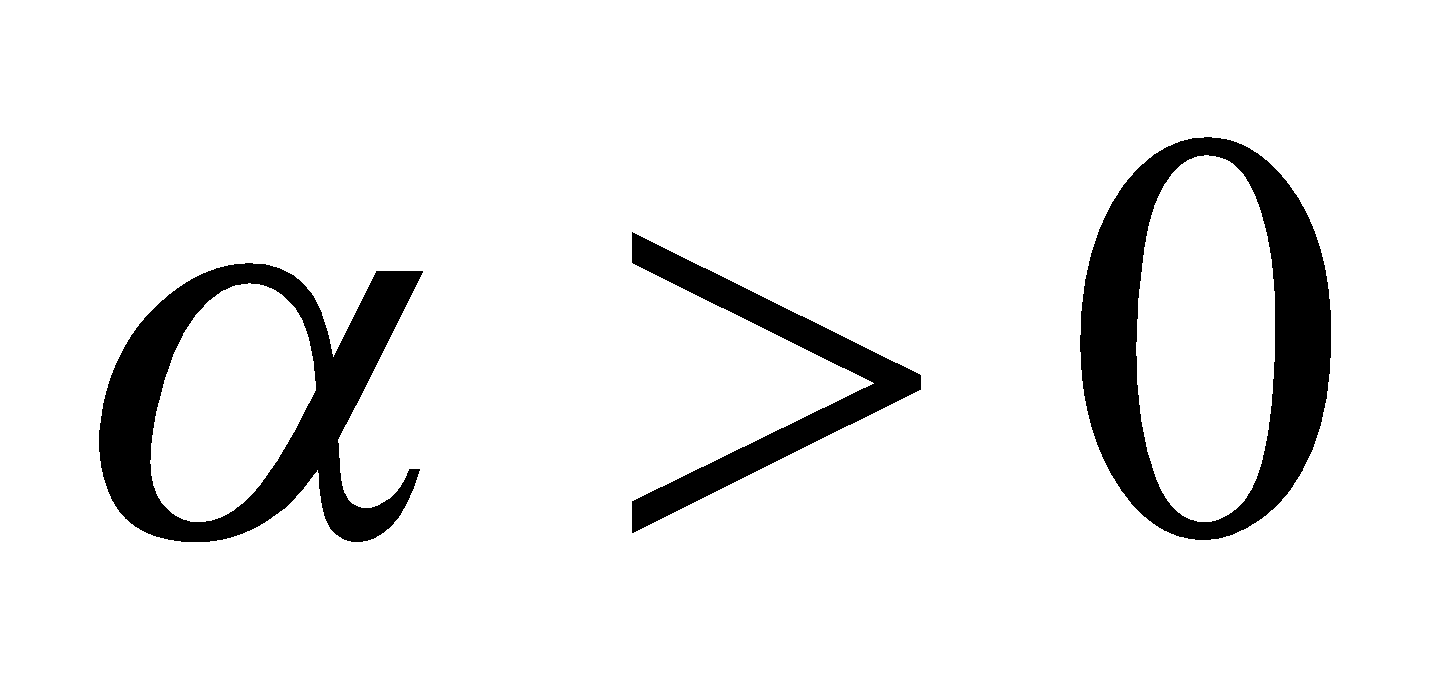
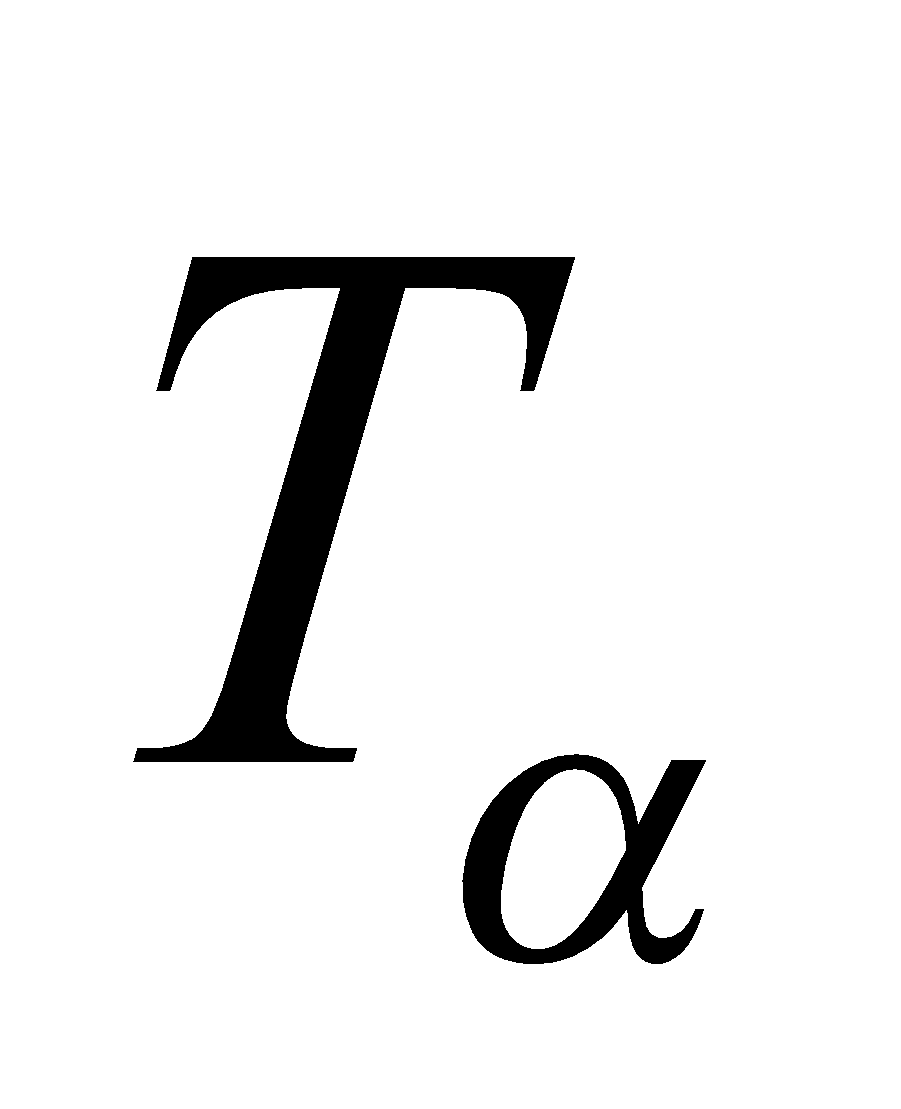
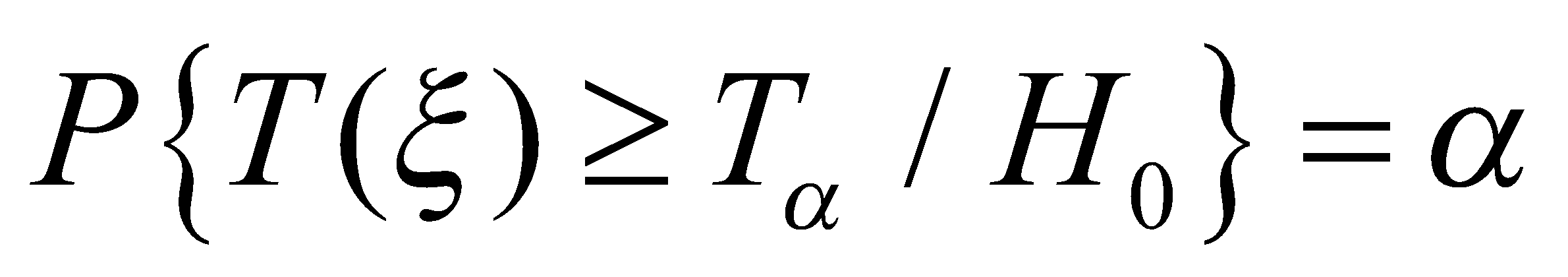
: .

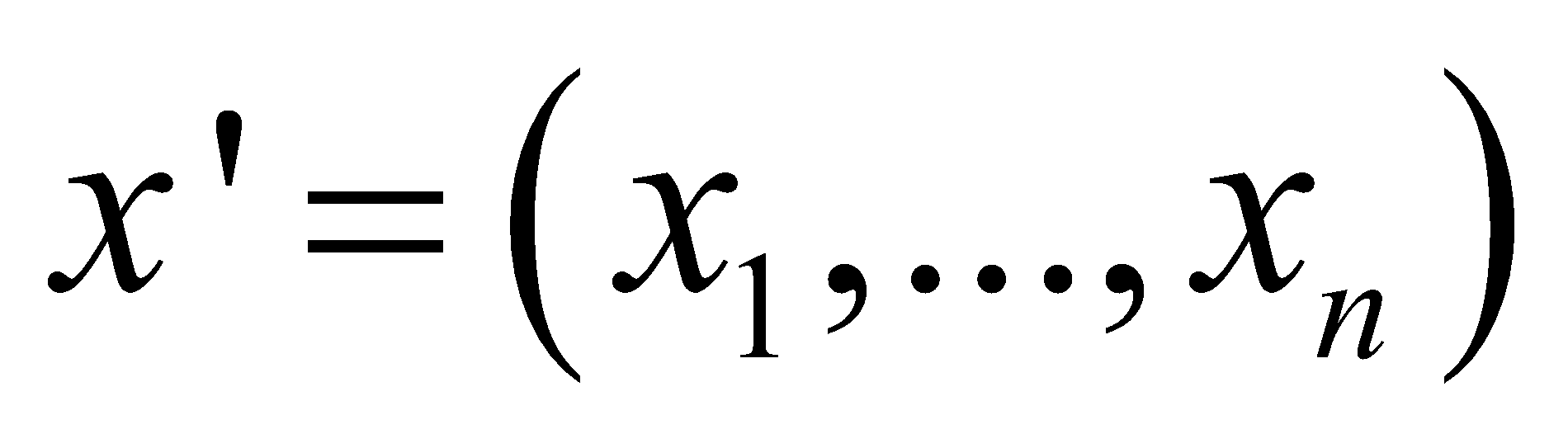
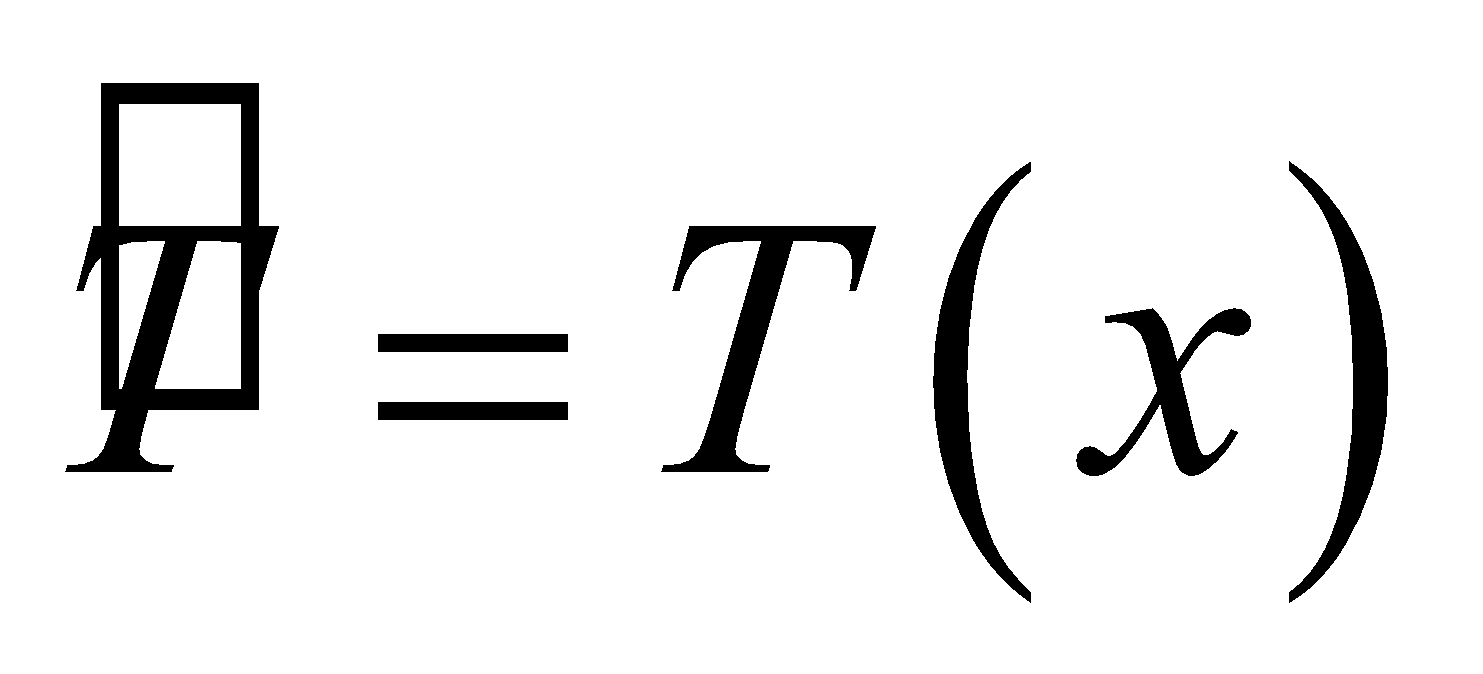
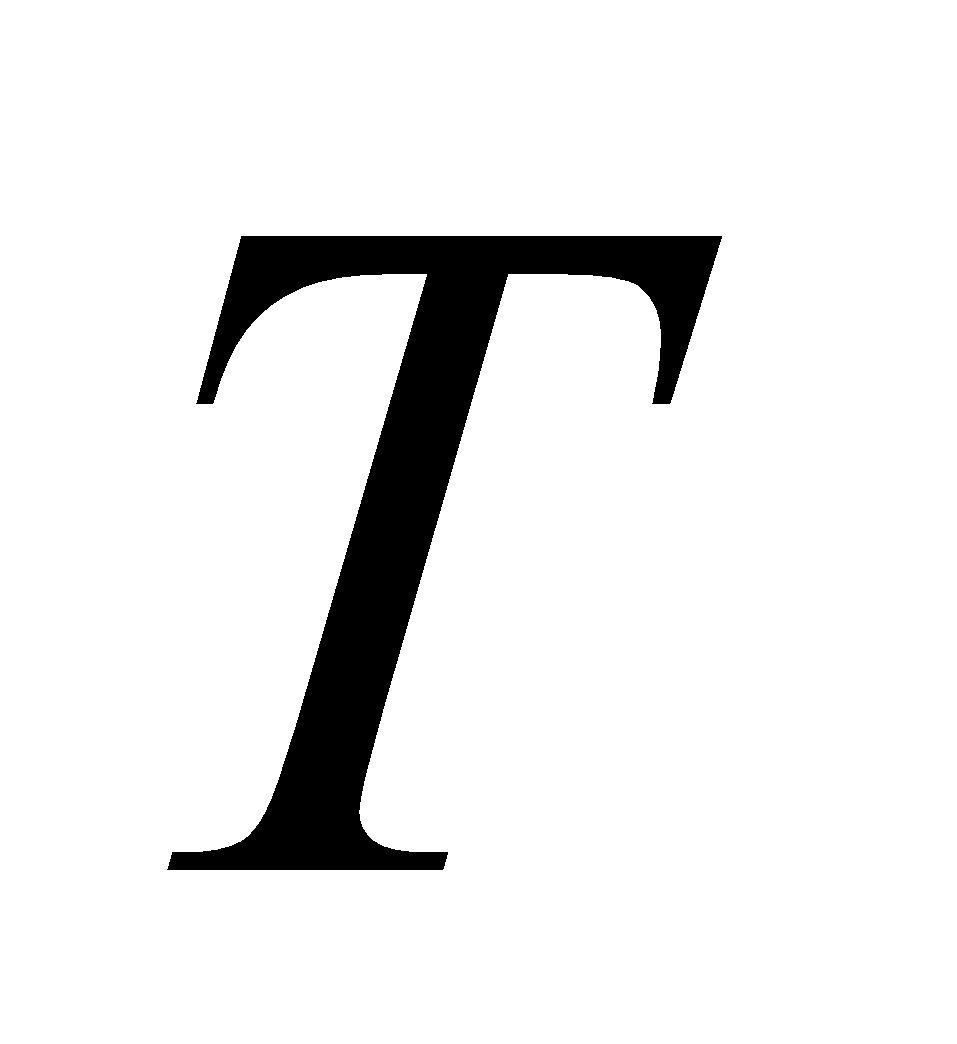
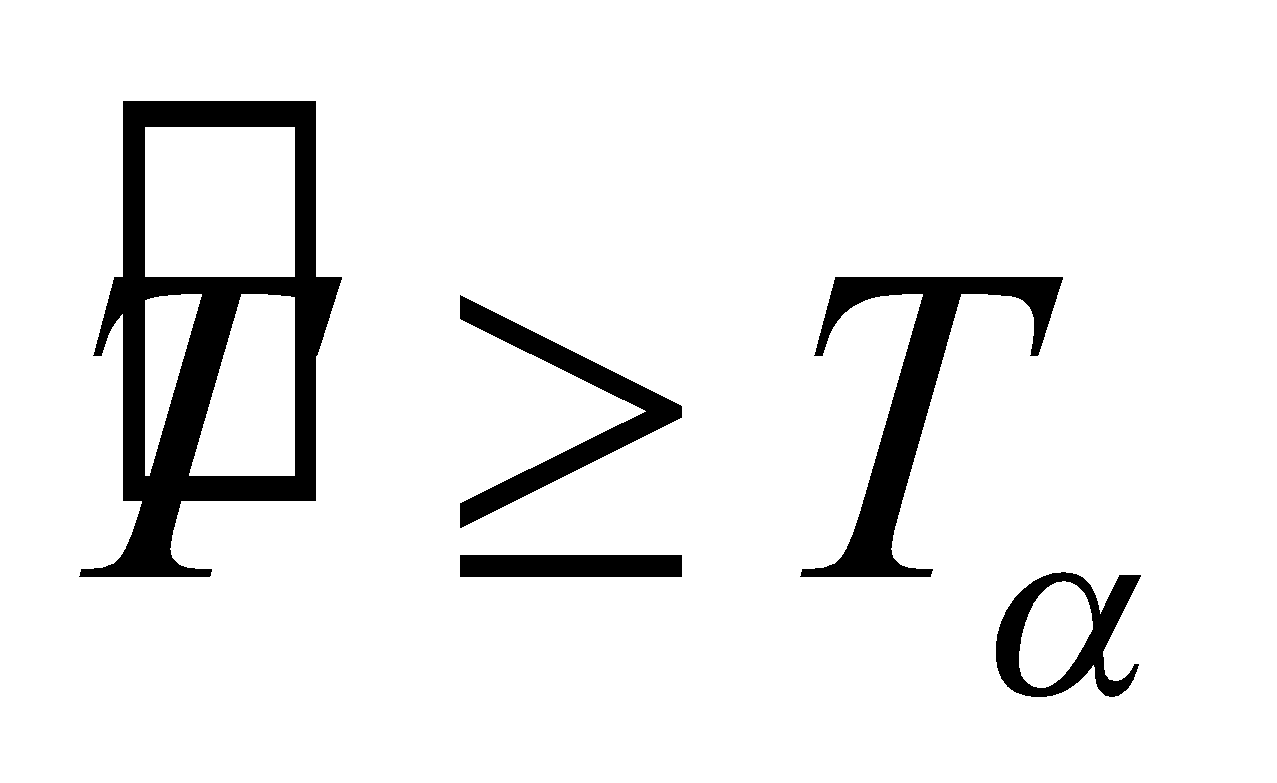
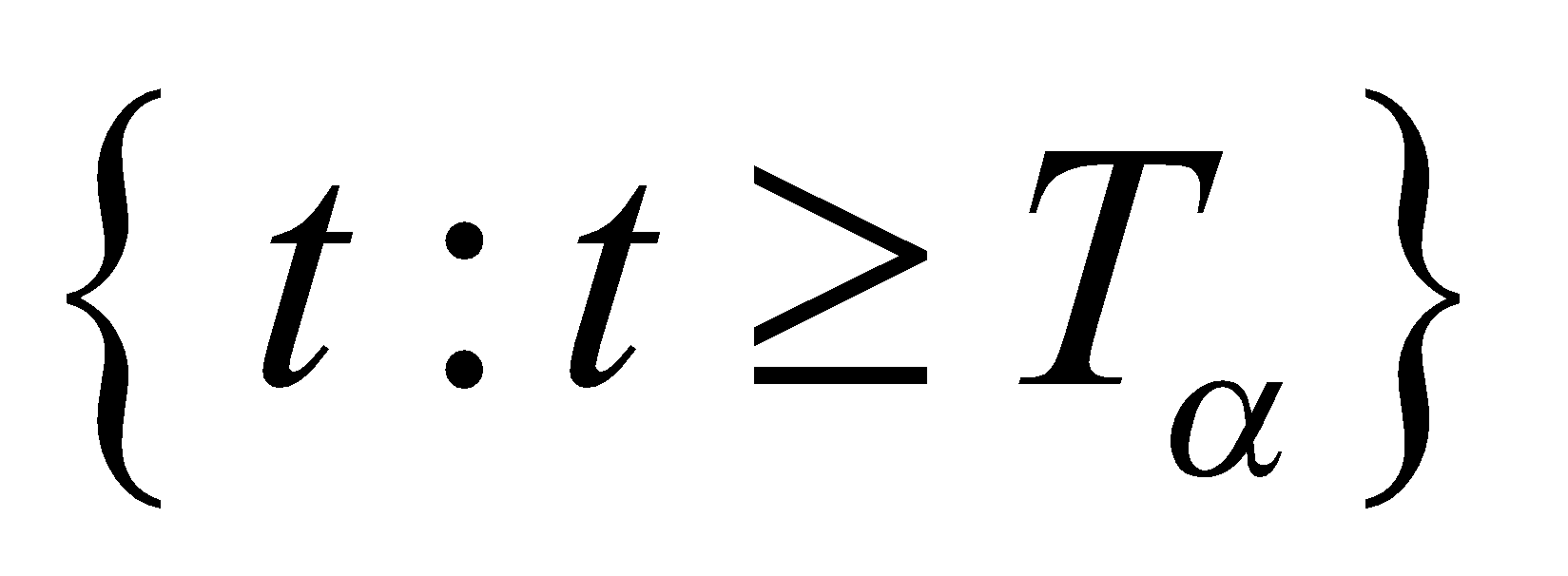
Якщо гіпотеза однозначно фіксує розподіл спостережень, то її називають ***простою***, у протилежному випадку – ***складною***. У наведених вище прикладах лише гіпотеза про вигляд розподілу  є простою.

***Статистичний критерій*** – це правило, згідно якому гіпотеза, що перевіряється, приймається або відкидається.

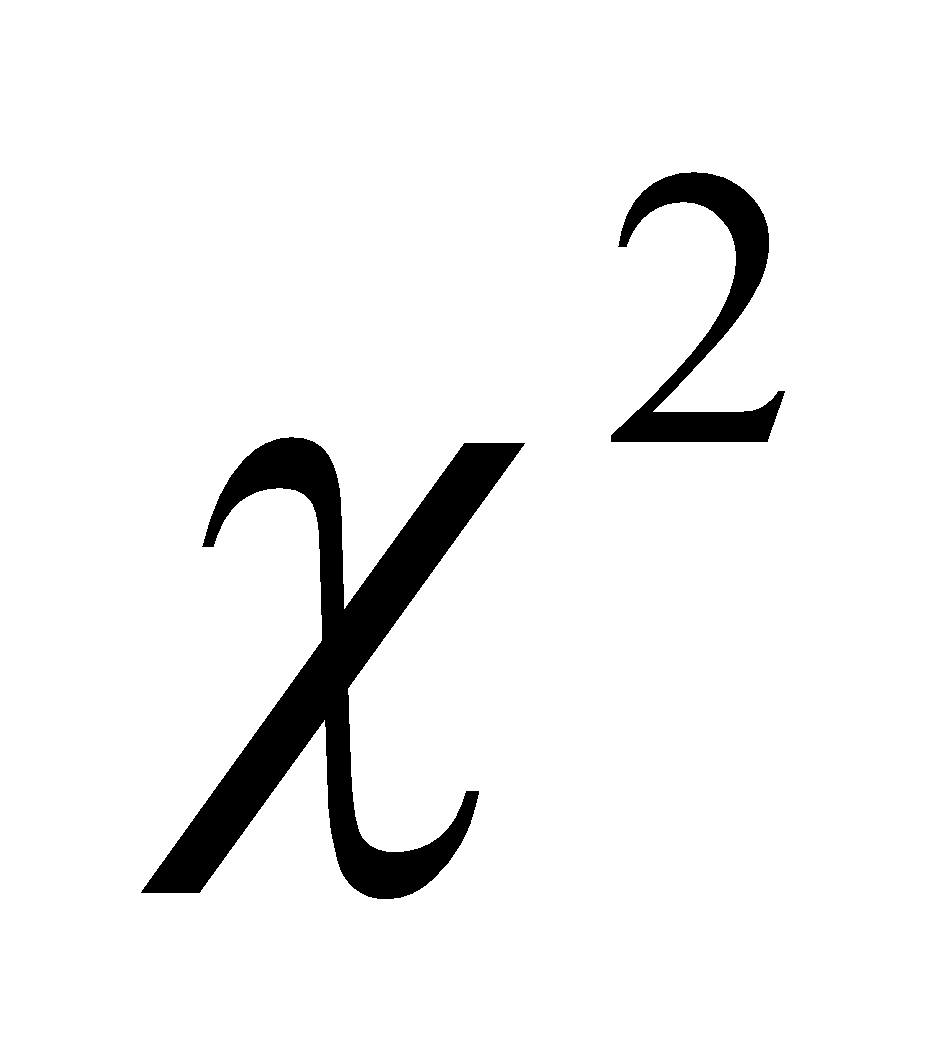
Розглянемо методи перевірки гіпотез описаних вище типів. Нехай про розподіл вибірки , що описує результати експерименту, сформульована гіпотеза . Необхідно перевірити, чи узгоджуються статистичні дані з цією гіпотезою, чи ні. Відповідні критерії називаються критеріями згоди. Наведемо методику побудови критеріїв згоди.

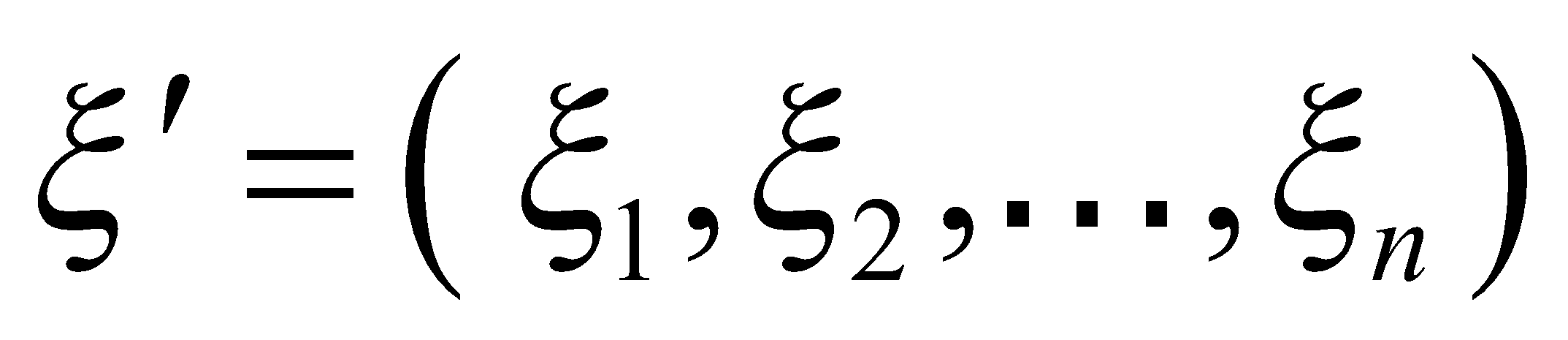
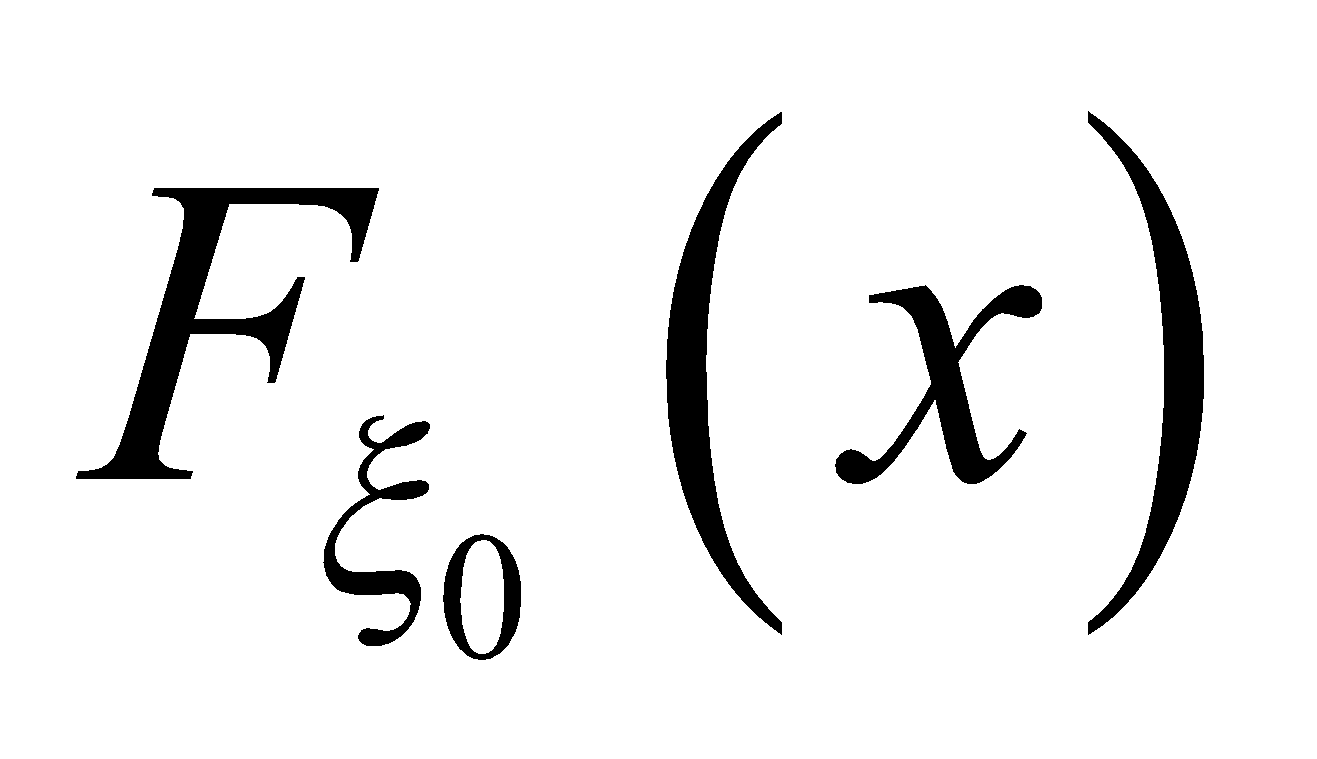
Вибирається статистика , яка характеризує відхилення емпіричних даних від гіпотетичних значень, що відповідають гіпотезі  (є мірою розбіжності статистичного та гіпотетичного законів розподілу), та називається статистикою критерію. Розподіл статистики  треба визначити точно або наближено в припущенні, що розподіл спостережень співпадає з гіпотетичним.

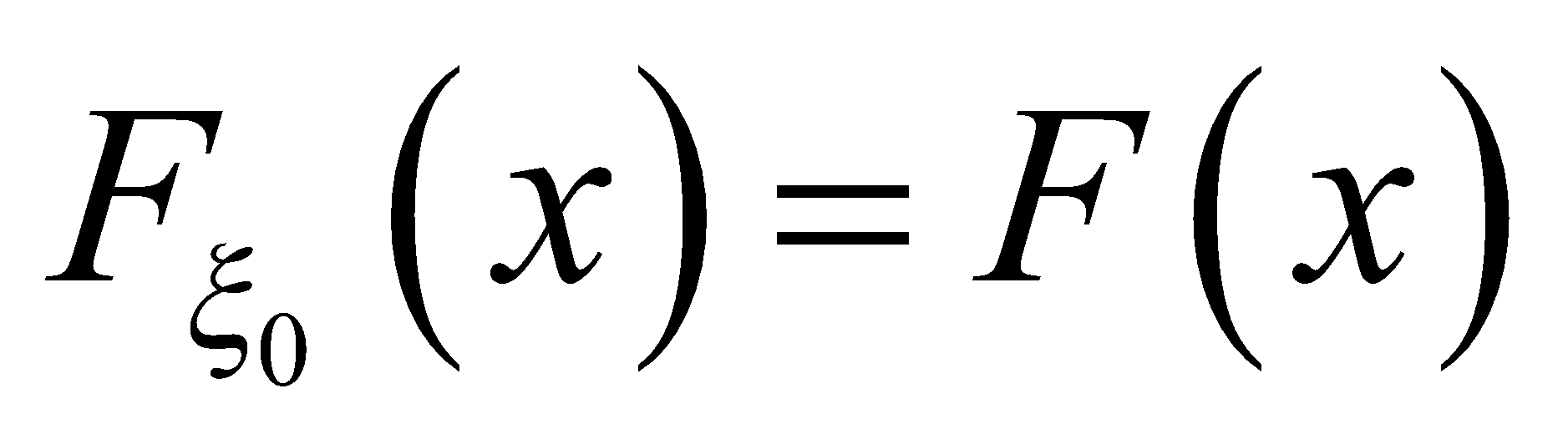
Нехай таку статистику знайдено. Визначимо для фіксованого достатньо малого числа  число  так, щоб у випадку справедливості гіпотези  імовірність настання події . Число  називається рівнем значущості критерію.

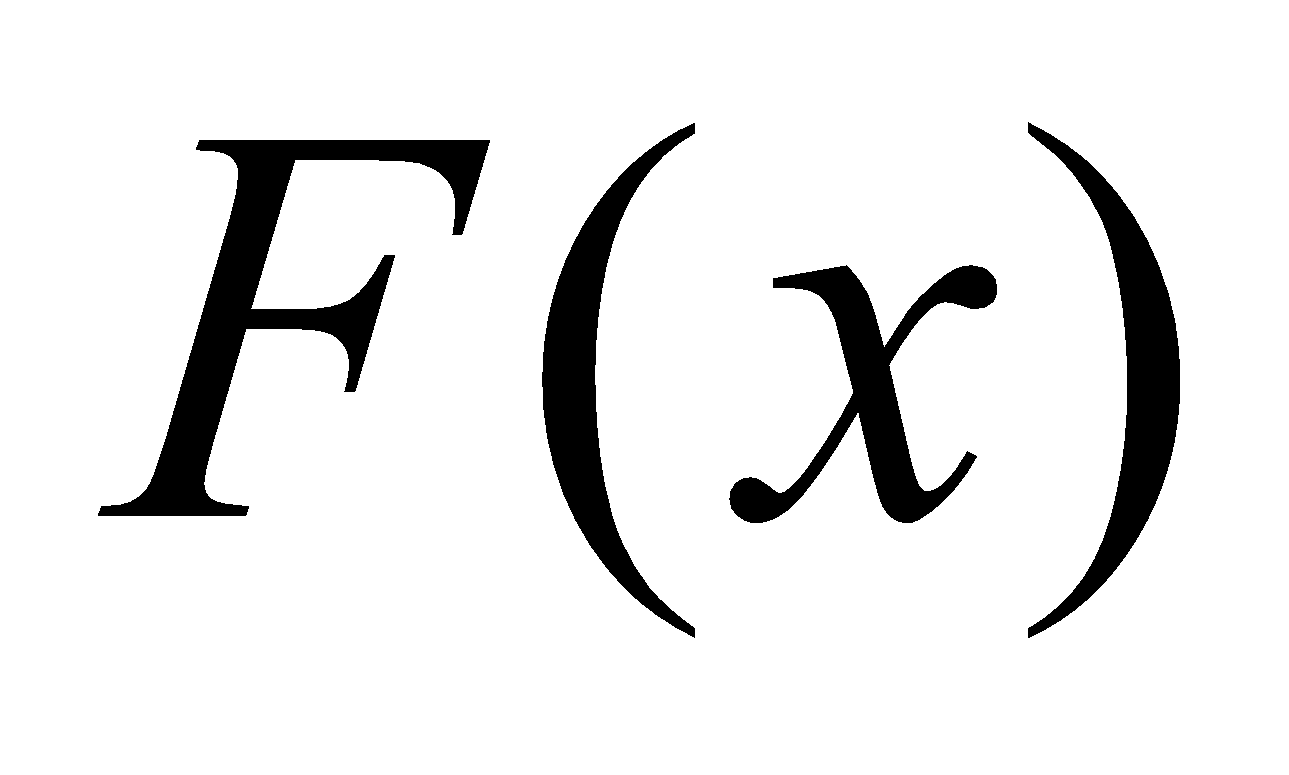
Нехай  – реалізація вибірки, і  – відповідне значення статистики , обчислене за статистичними даними. Якщо виявиться, що , то відхилення від гіпотетичного закону розподілу вважається значущим і гіпотеза відхиляється. У протилежному випадку немає підстав відмовлятися від висунутої гіпотези і слід вважати, що спостереження не протирічать гіпотезі (на рівні ). Область  називається критичною областю для гіпотези .

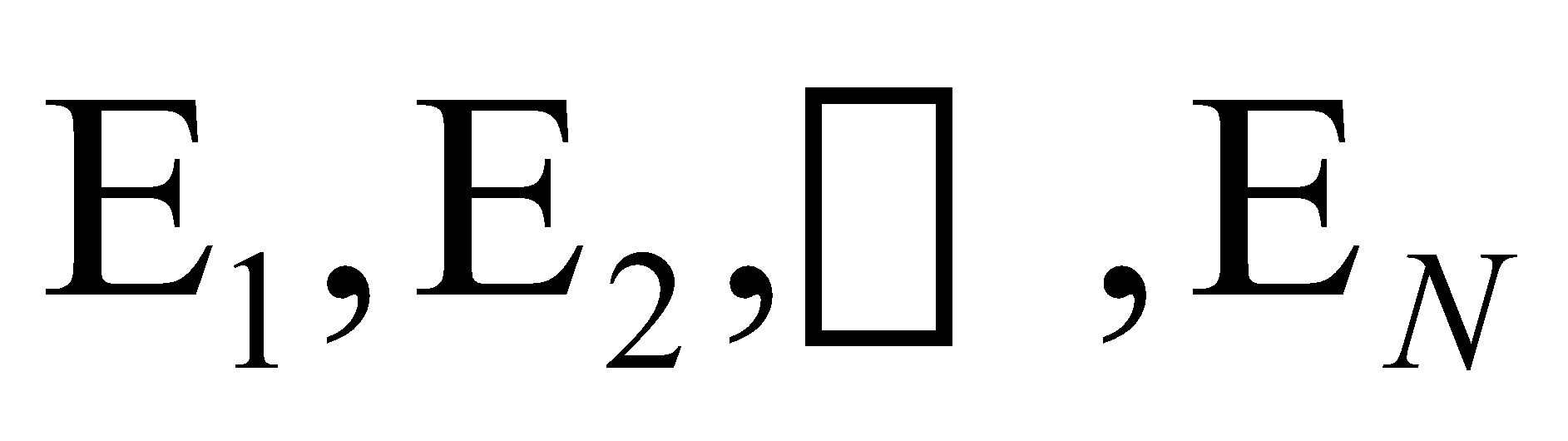
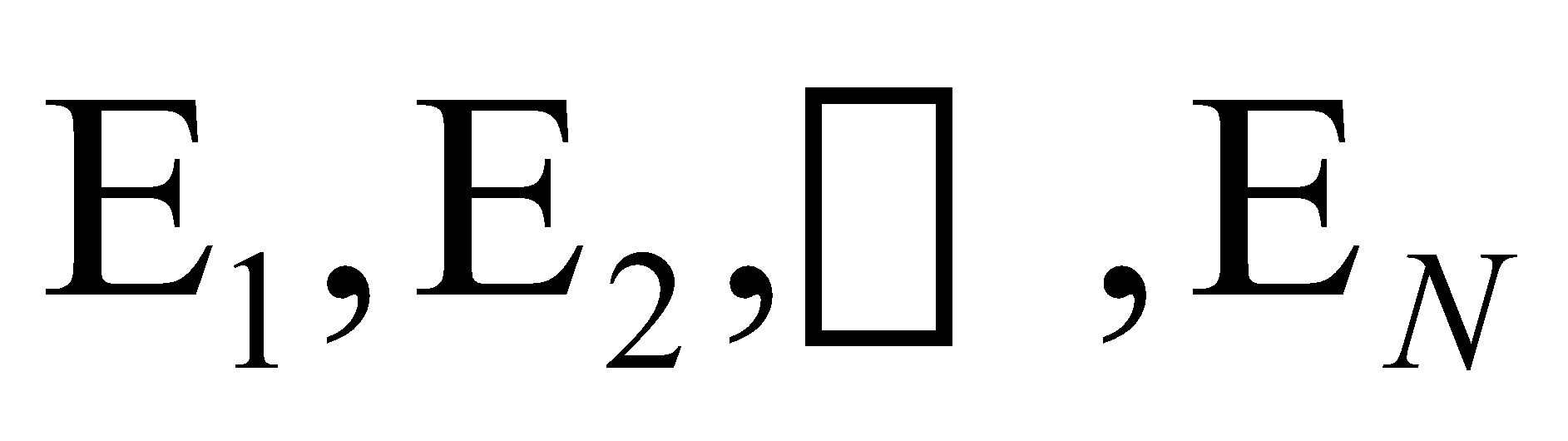
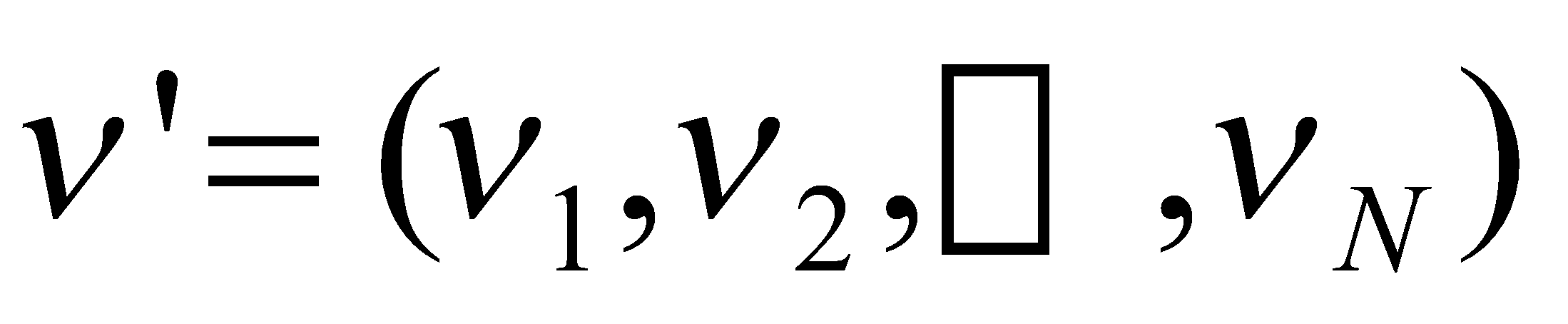
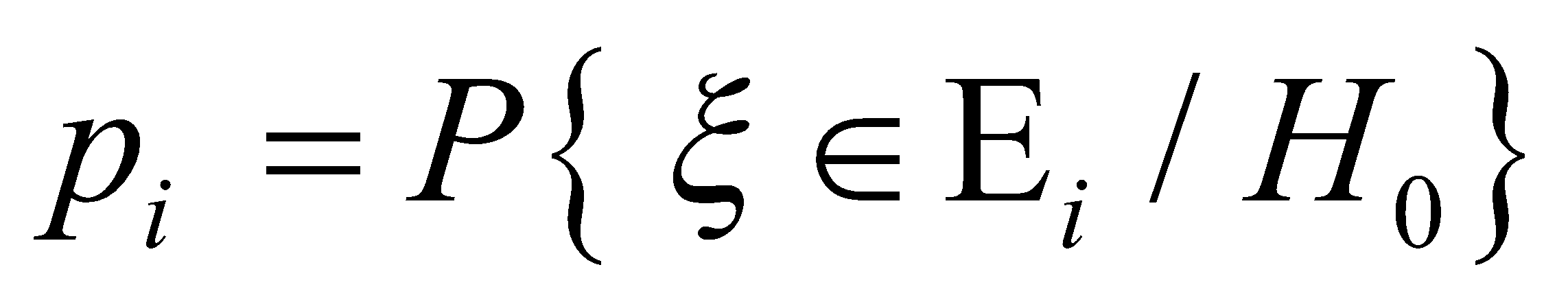
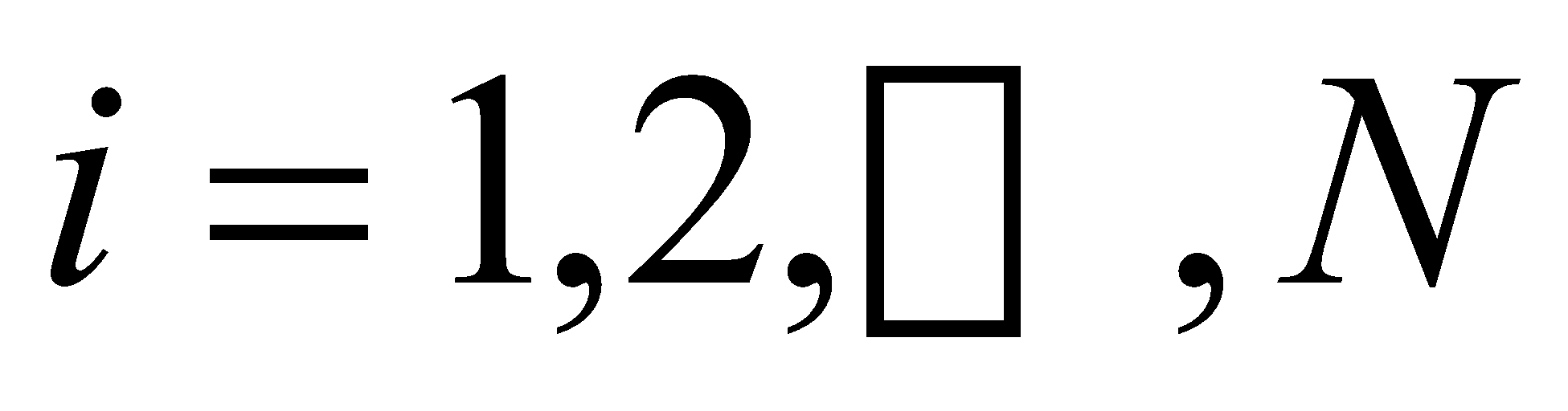
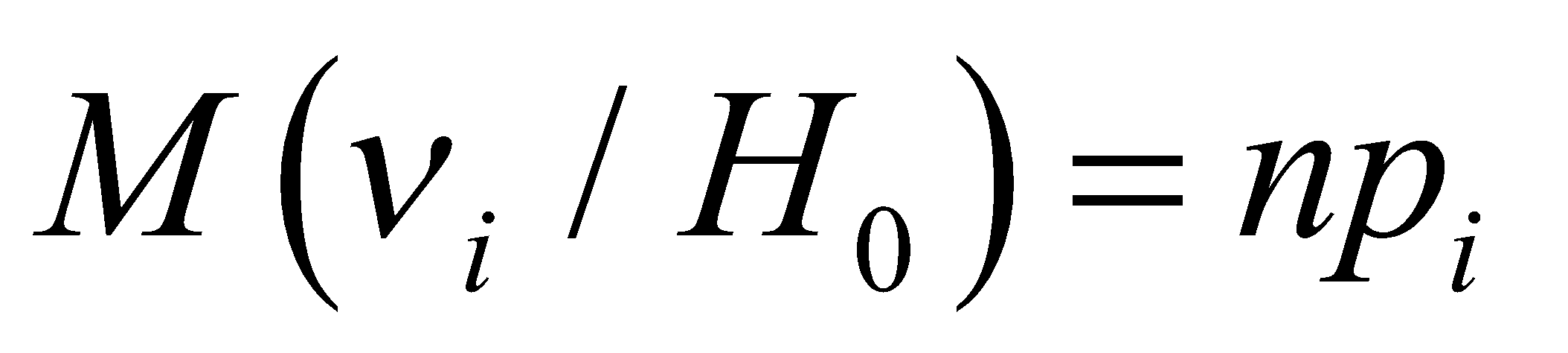
**4.2. Гіпотези про вид розподілу**

**4.2.2. Критерій**  **К. Пірсона**

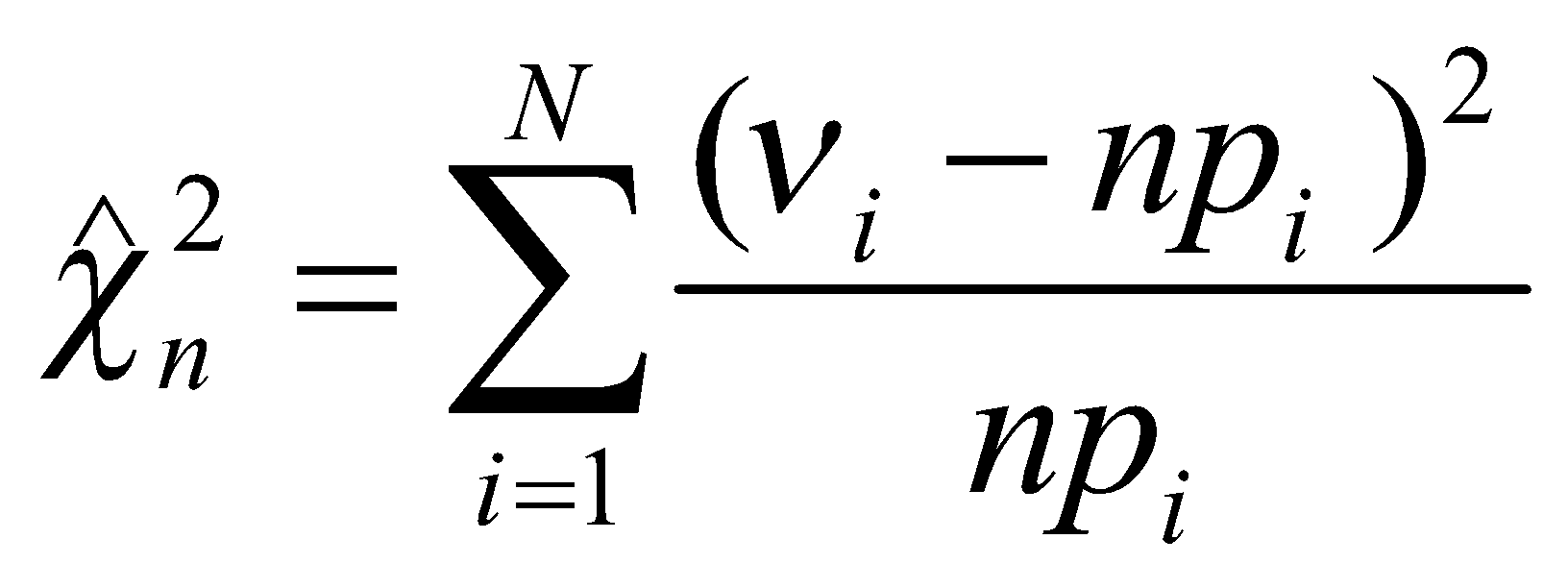
Нехай  – вибірка з невідомою функцією розподілу , про яку висунута проста гіпотеза

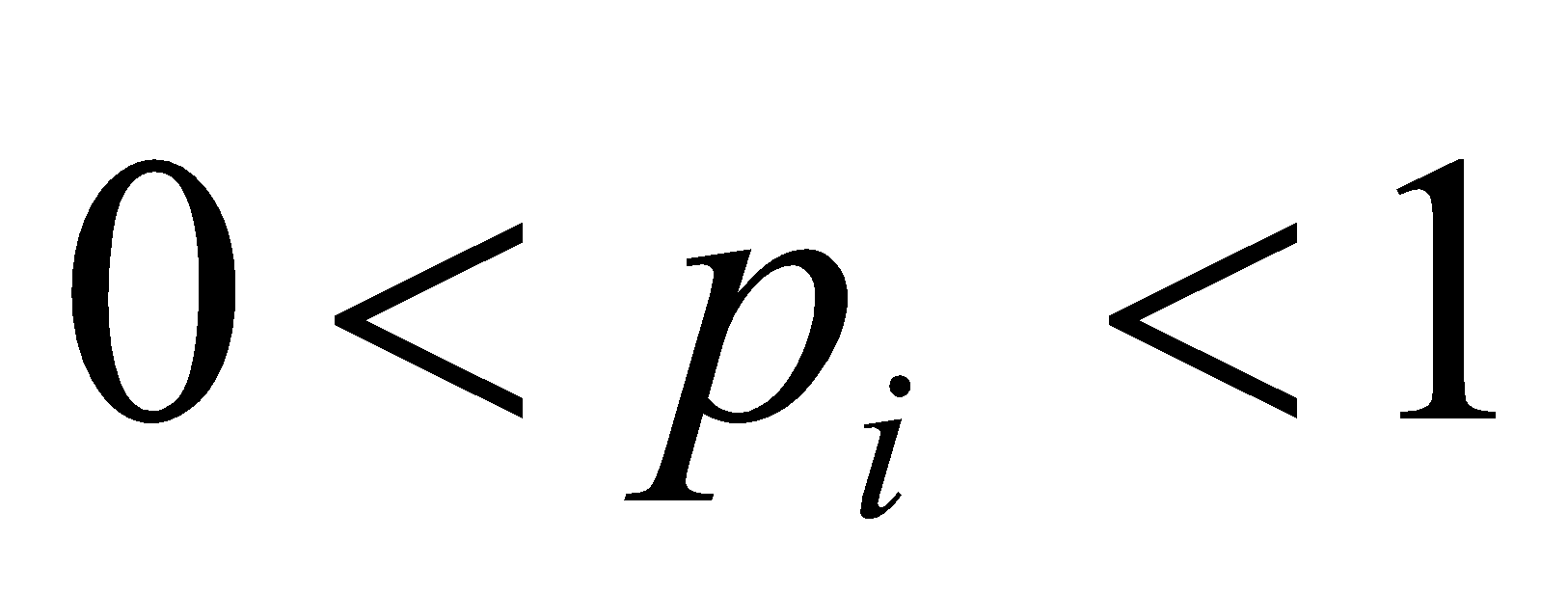
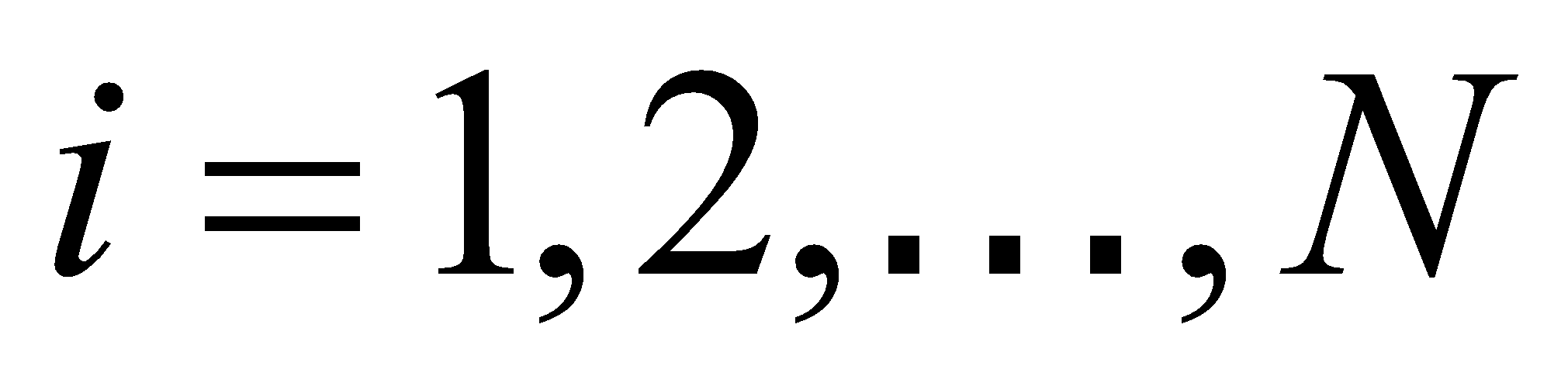
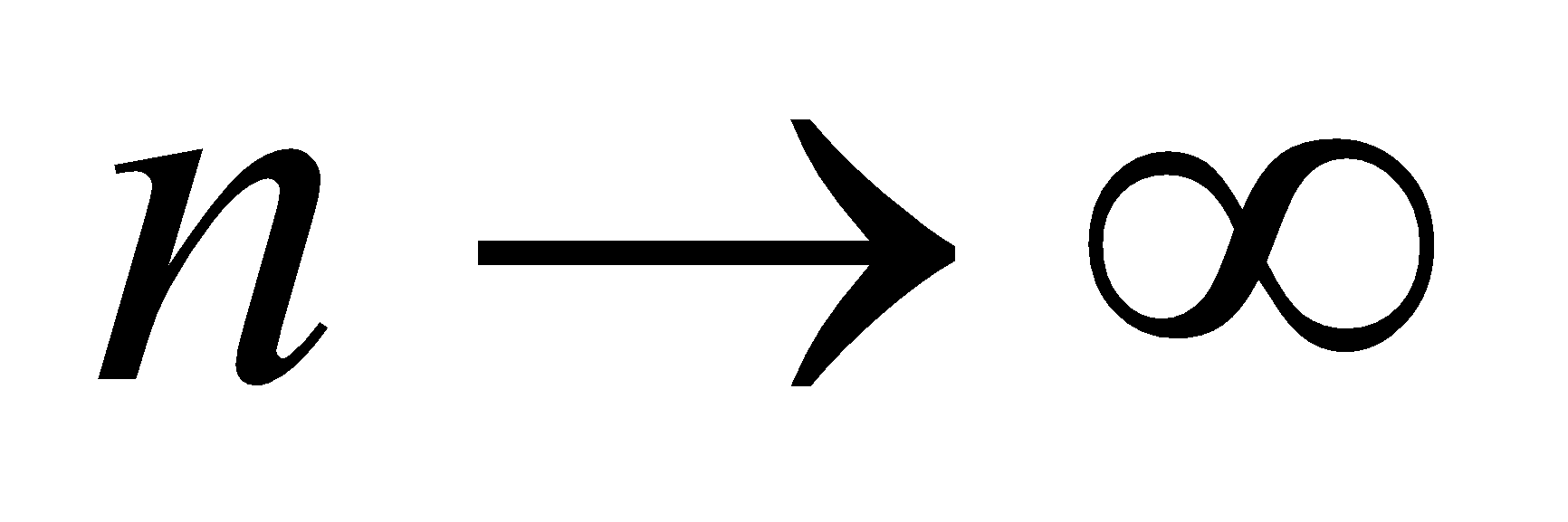
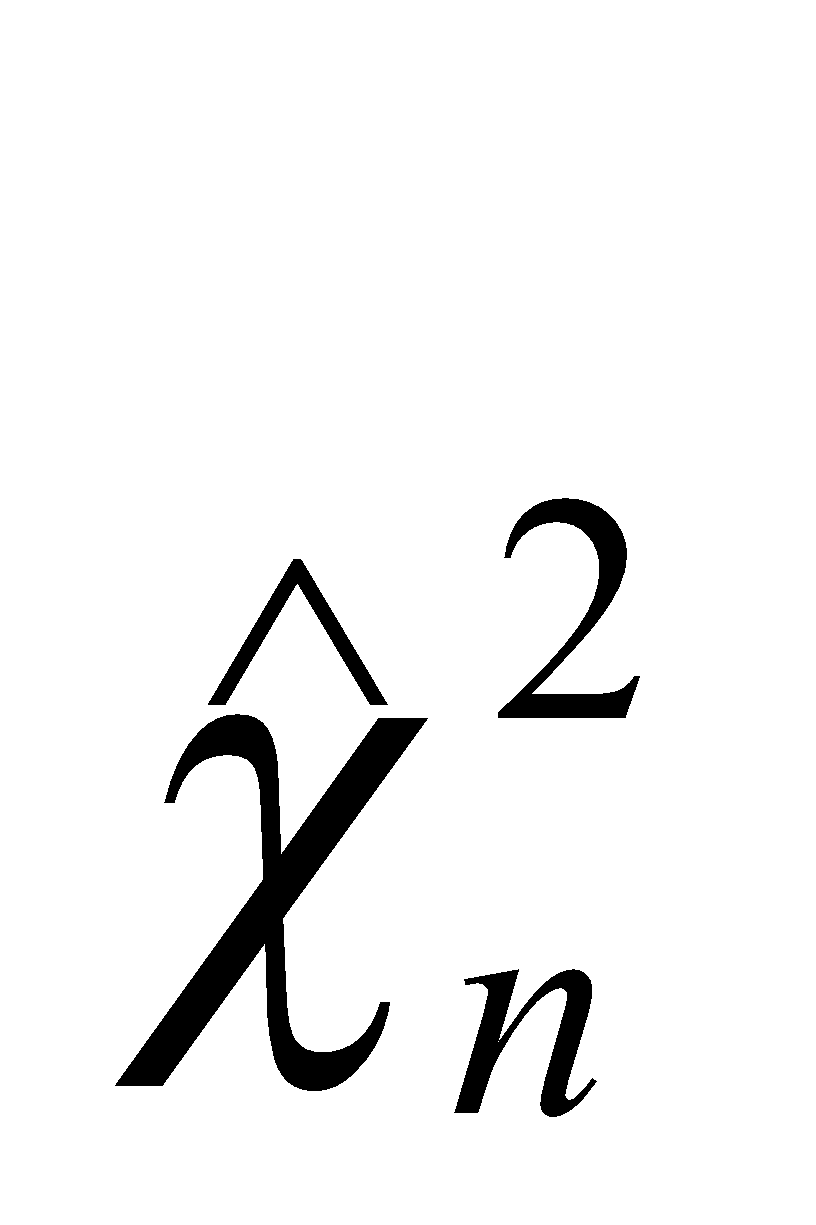
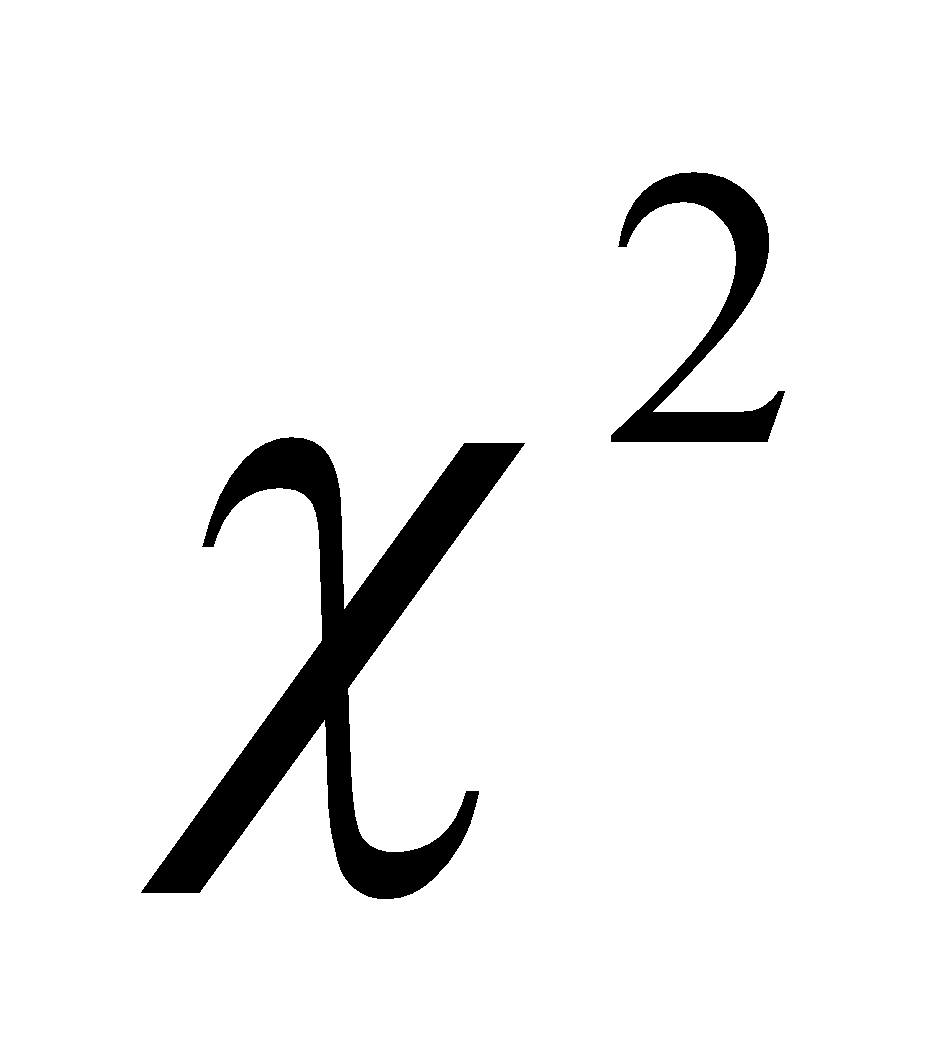
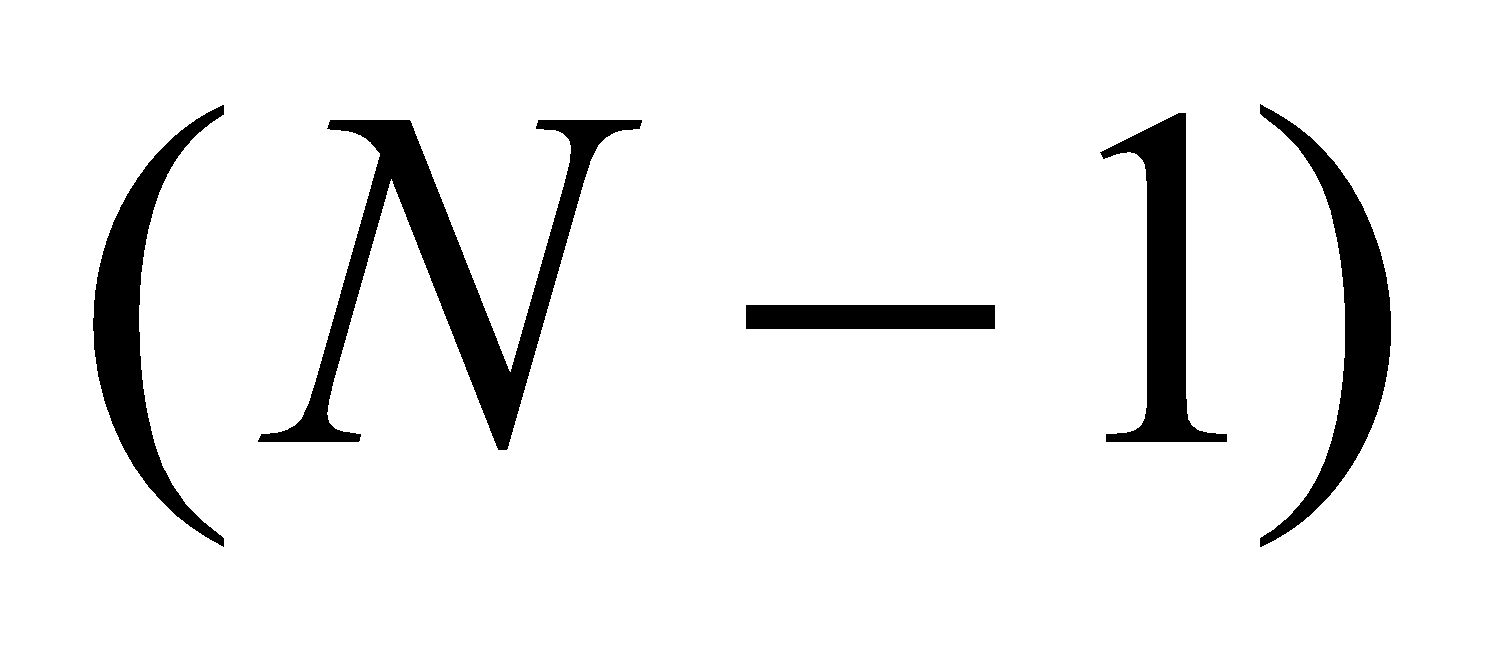
: .

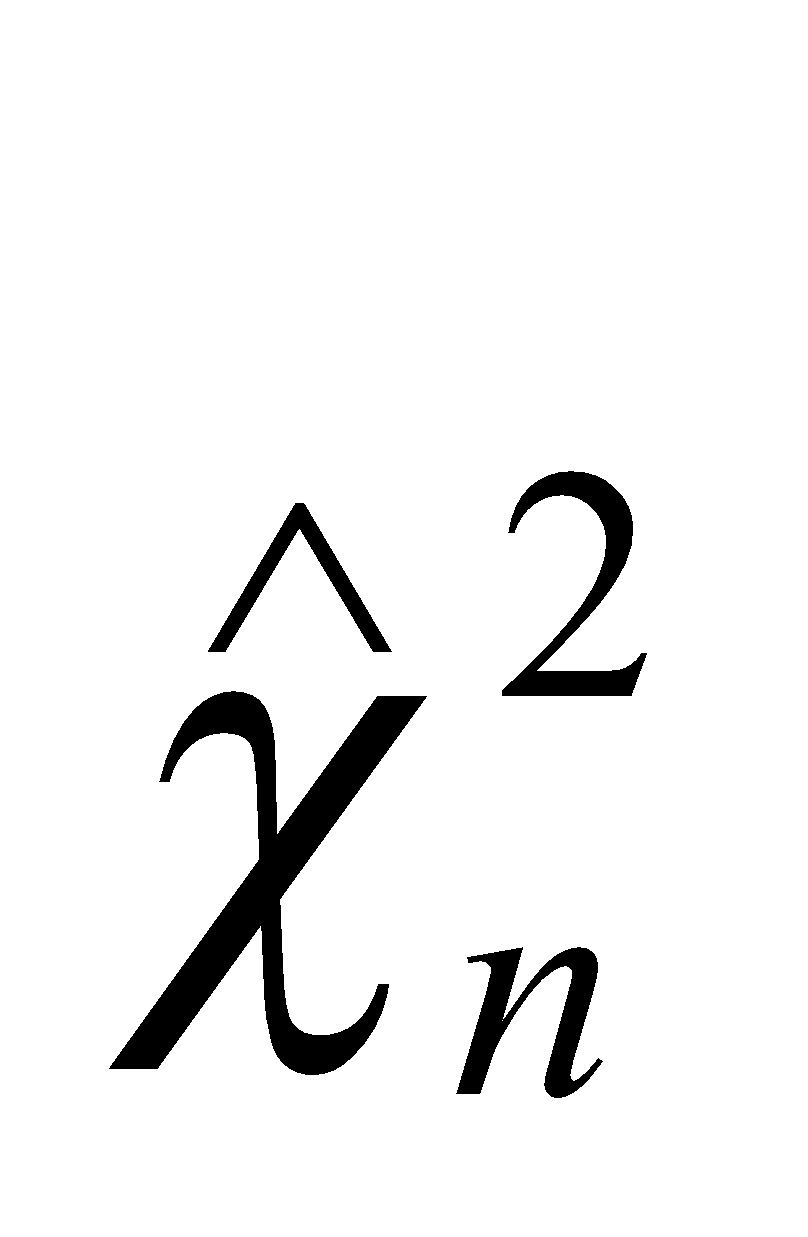
Про властивості гіпотетичної  в даному випадку нічого не відомо, тобто цей критерій можна використовувати як для неперервних, так і для дискретних розподілів.

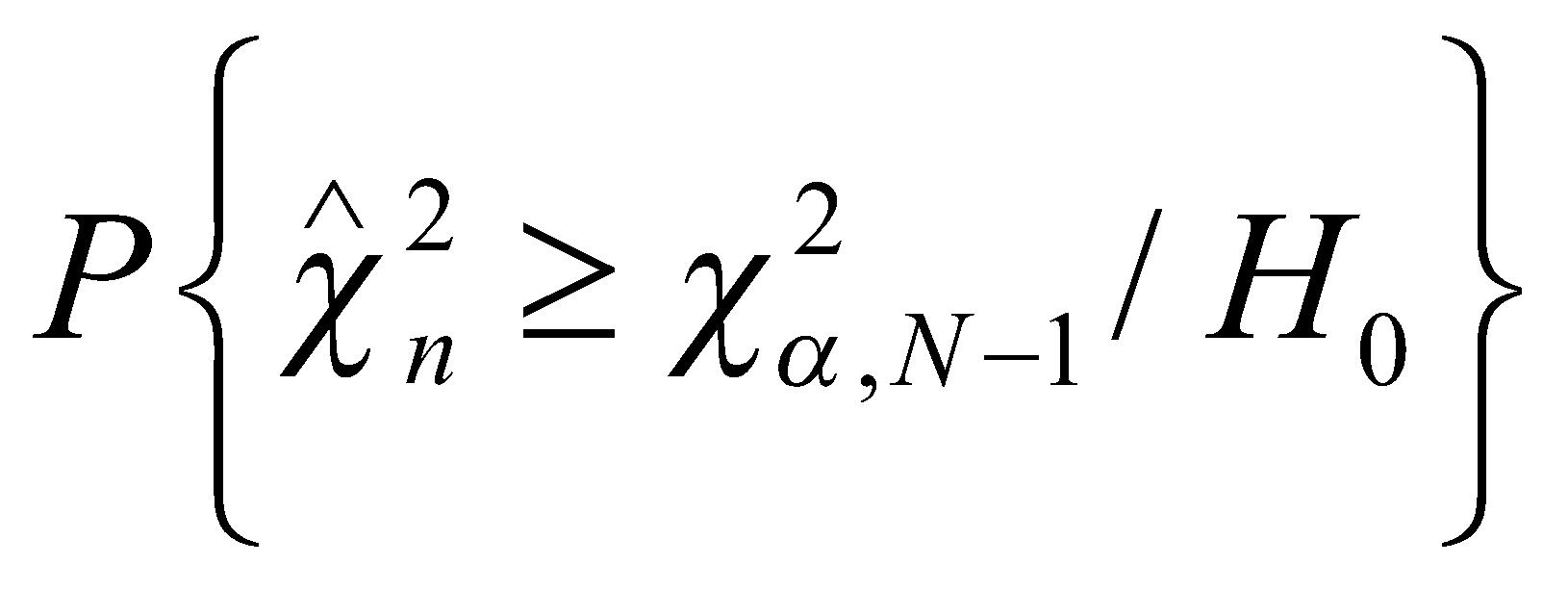
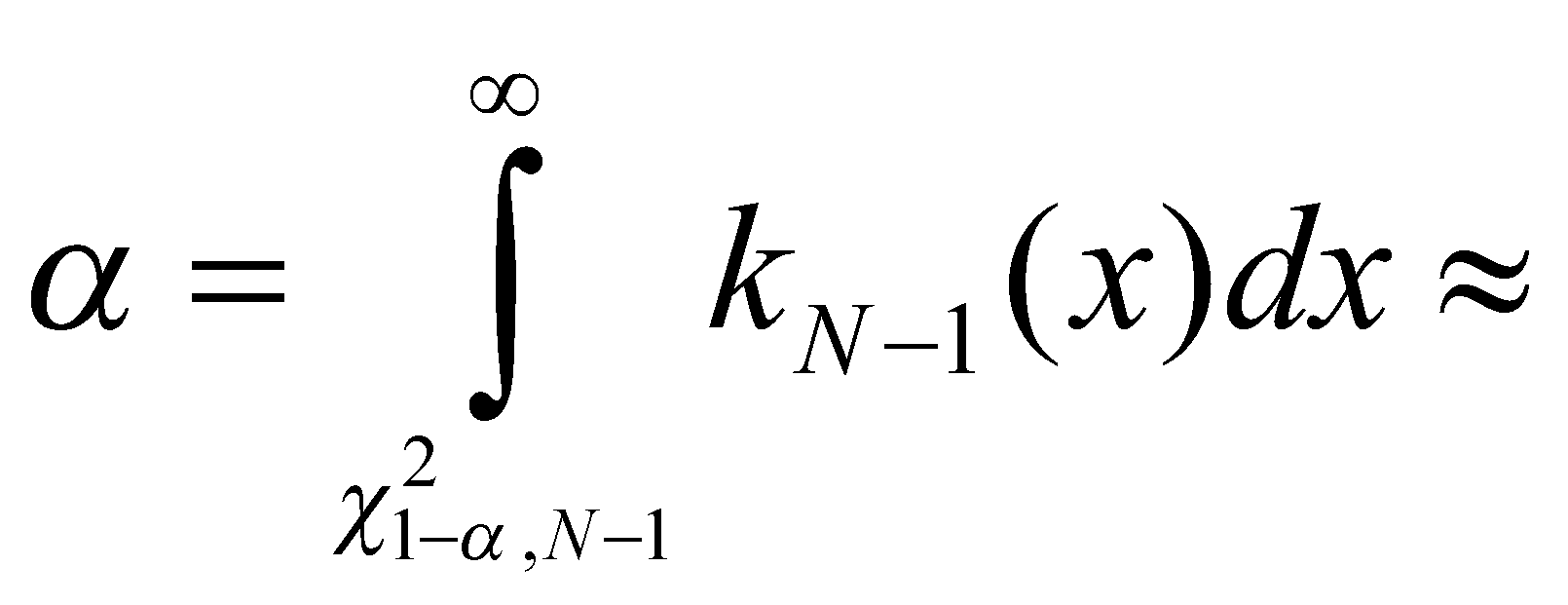
Задамо  – інтервали групування даних, що не перетинаються. Якщо спостерігається дискретна випадкова величина, то  – це різні значення цієї величини. Нехай  – вектор частот влучення елементів вибірки у відповідні інтервали групування. Позначимо , . Очевидно, що .

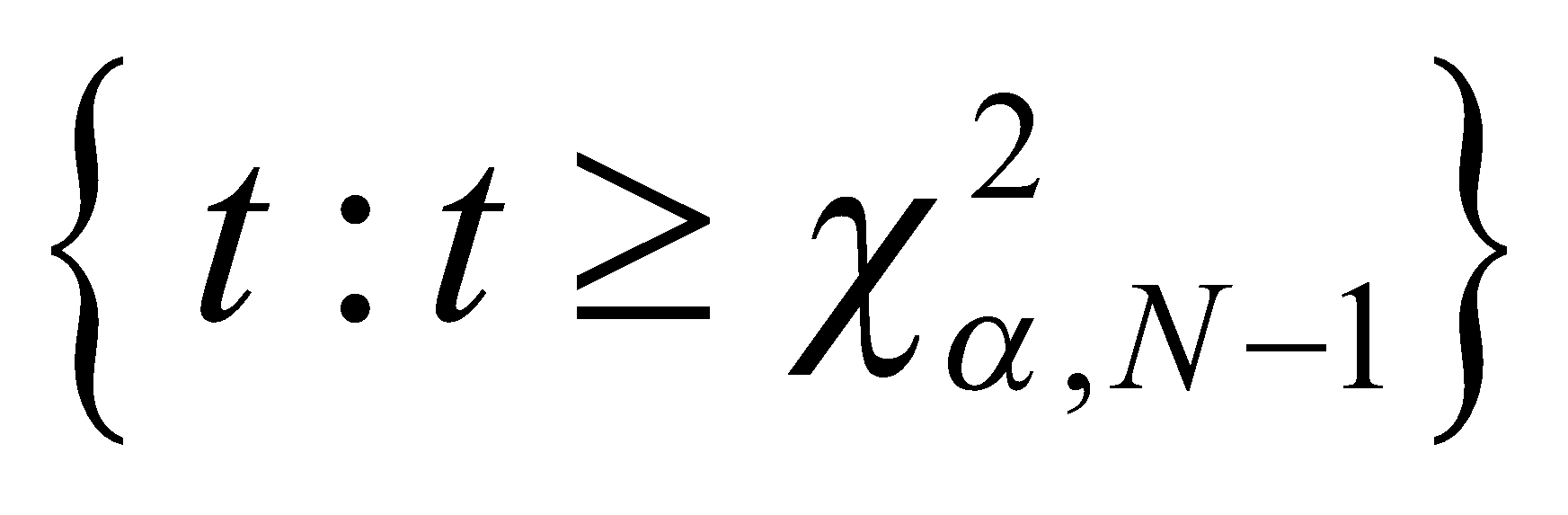
Як міру відхилення емпіричних даних від їх гіпотетичних значень візьмемо статистику

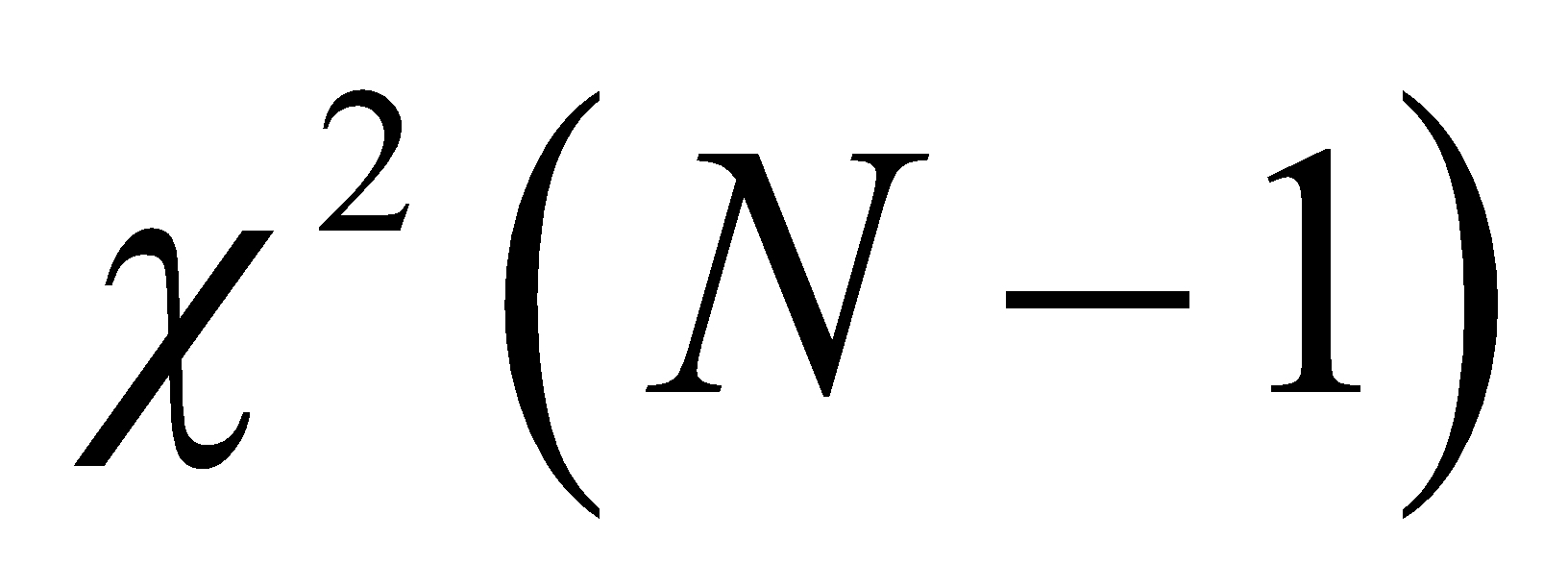
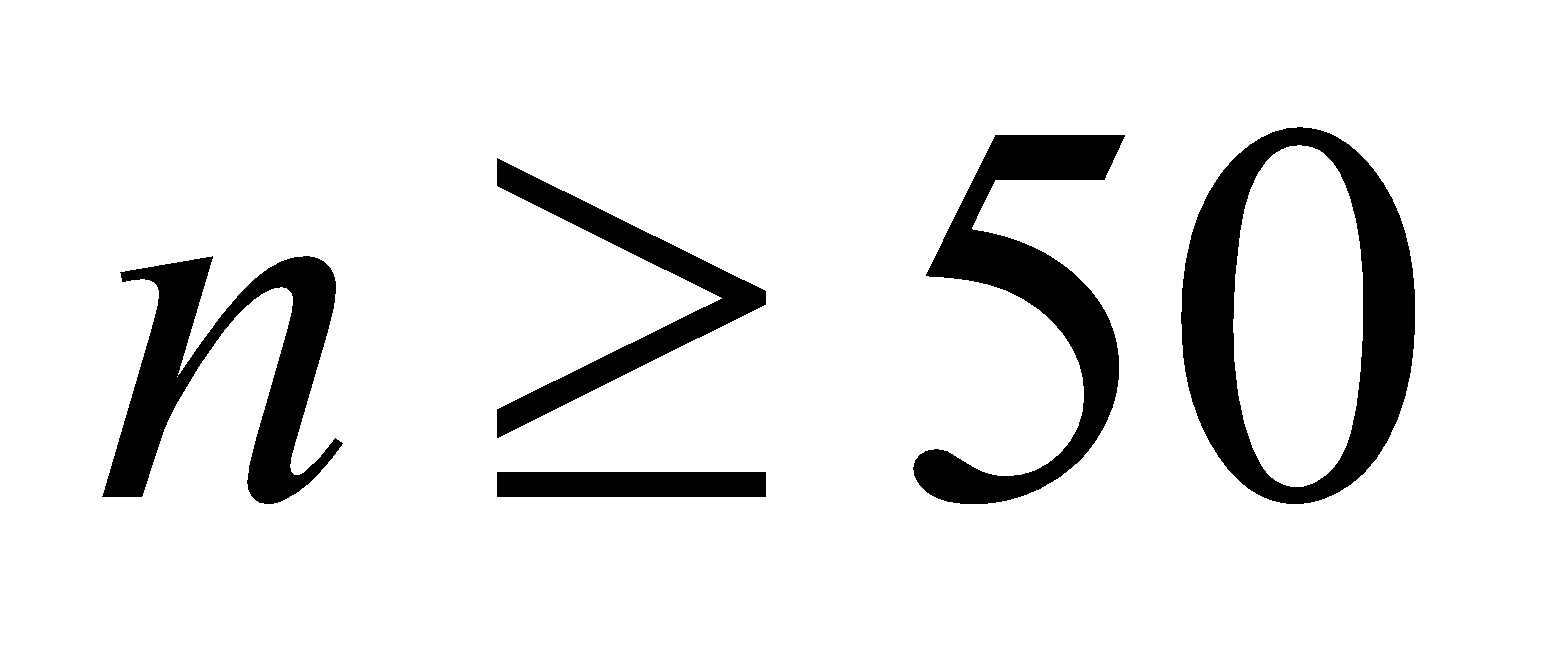
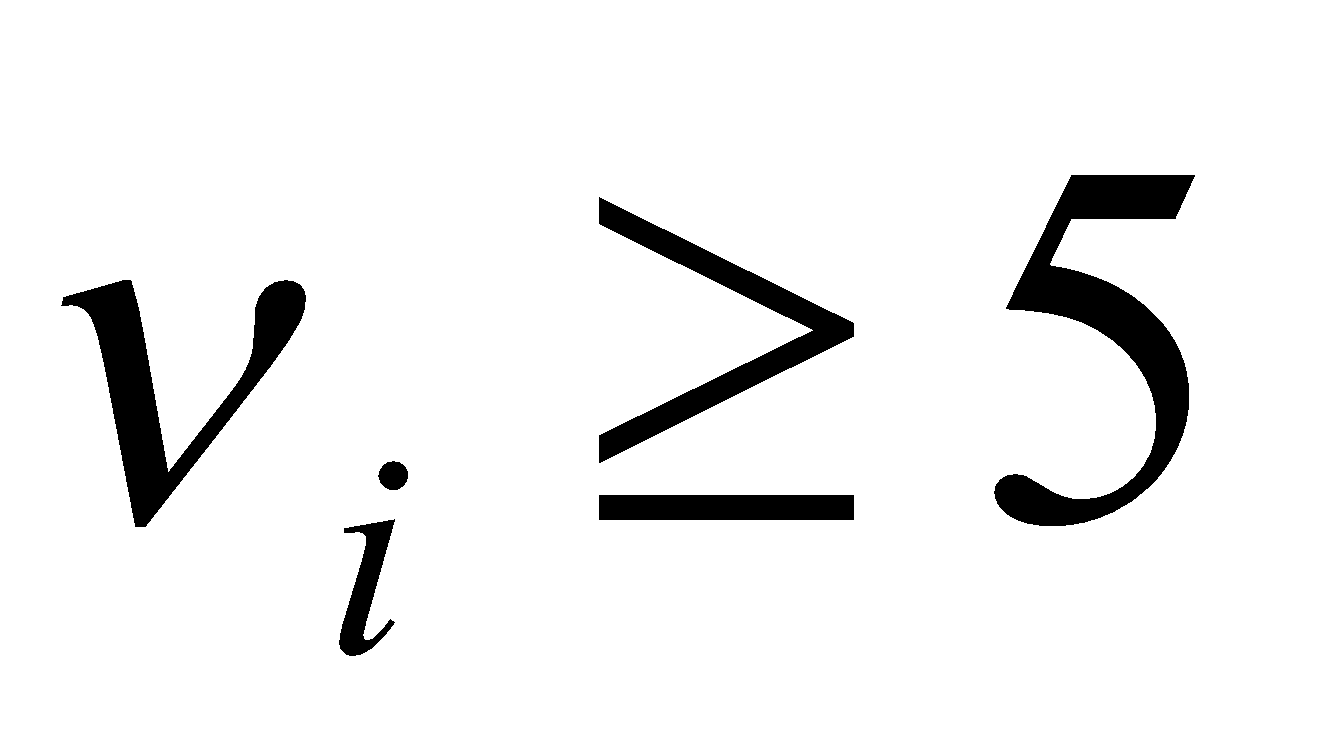
****. (4.2)

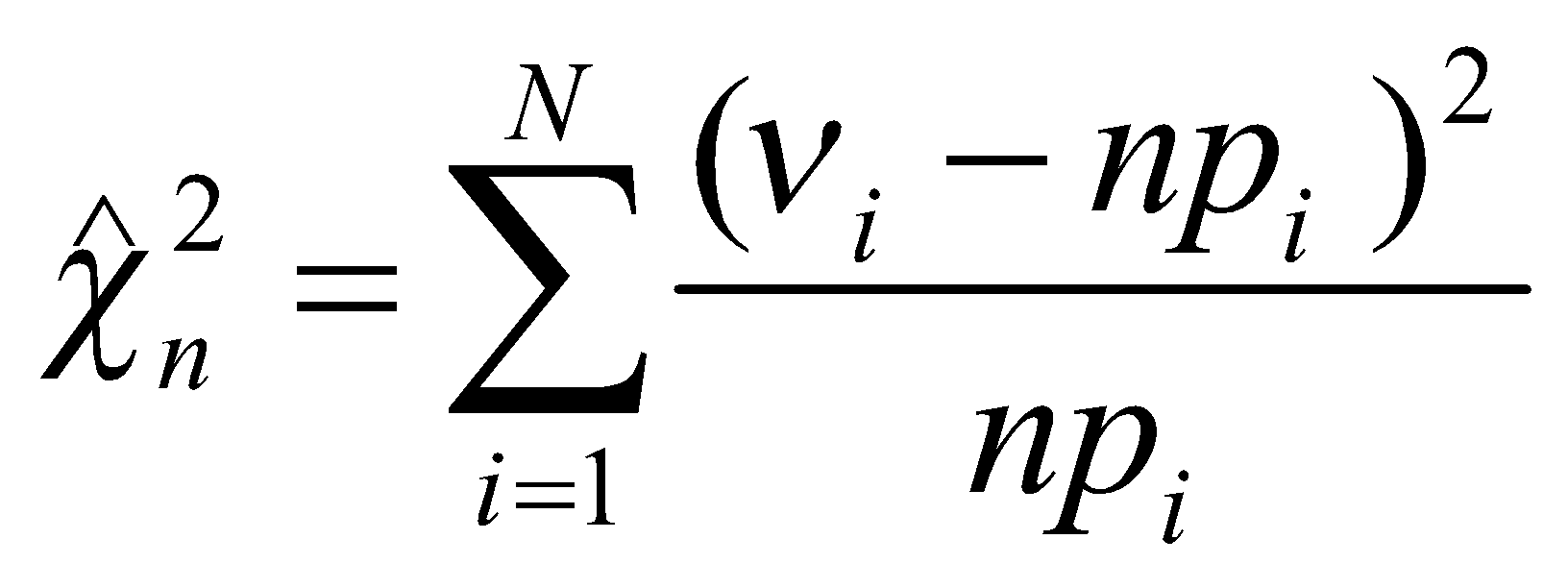
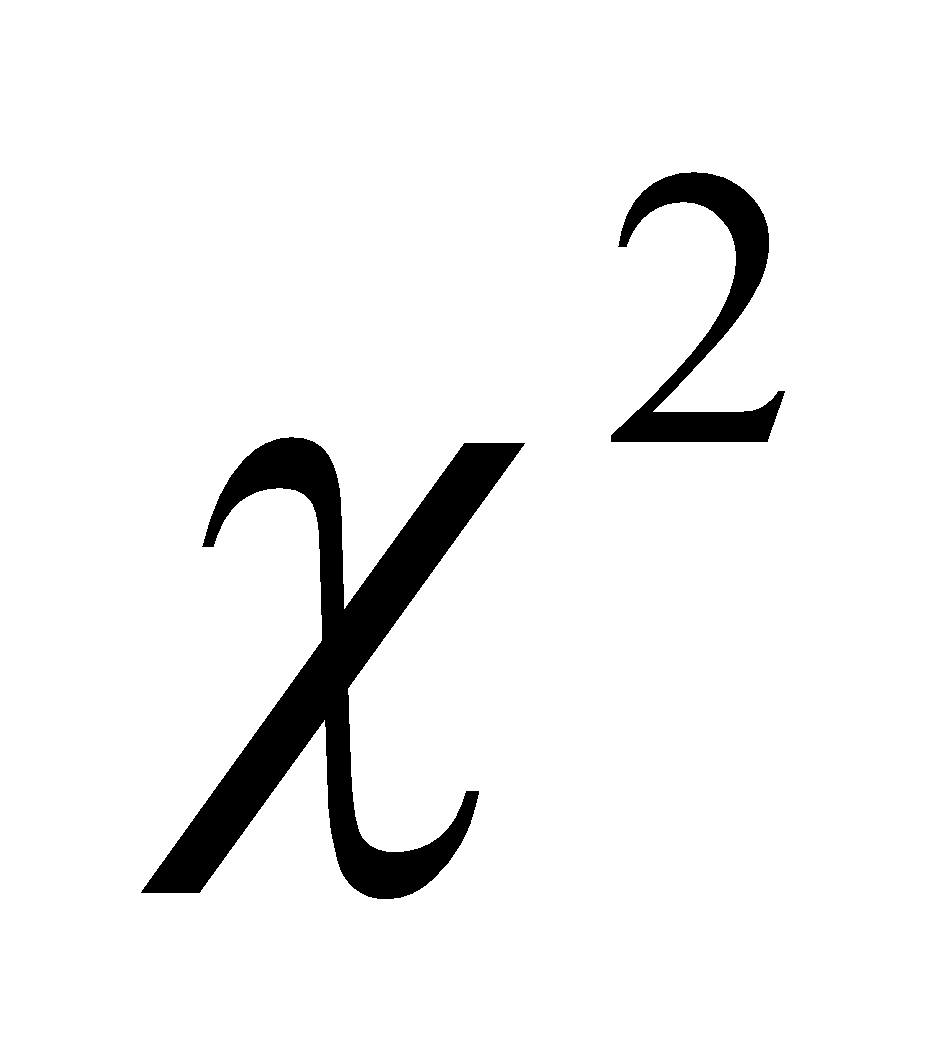
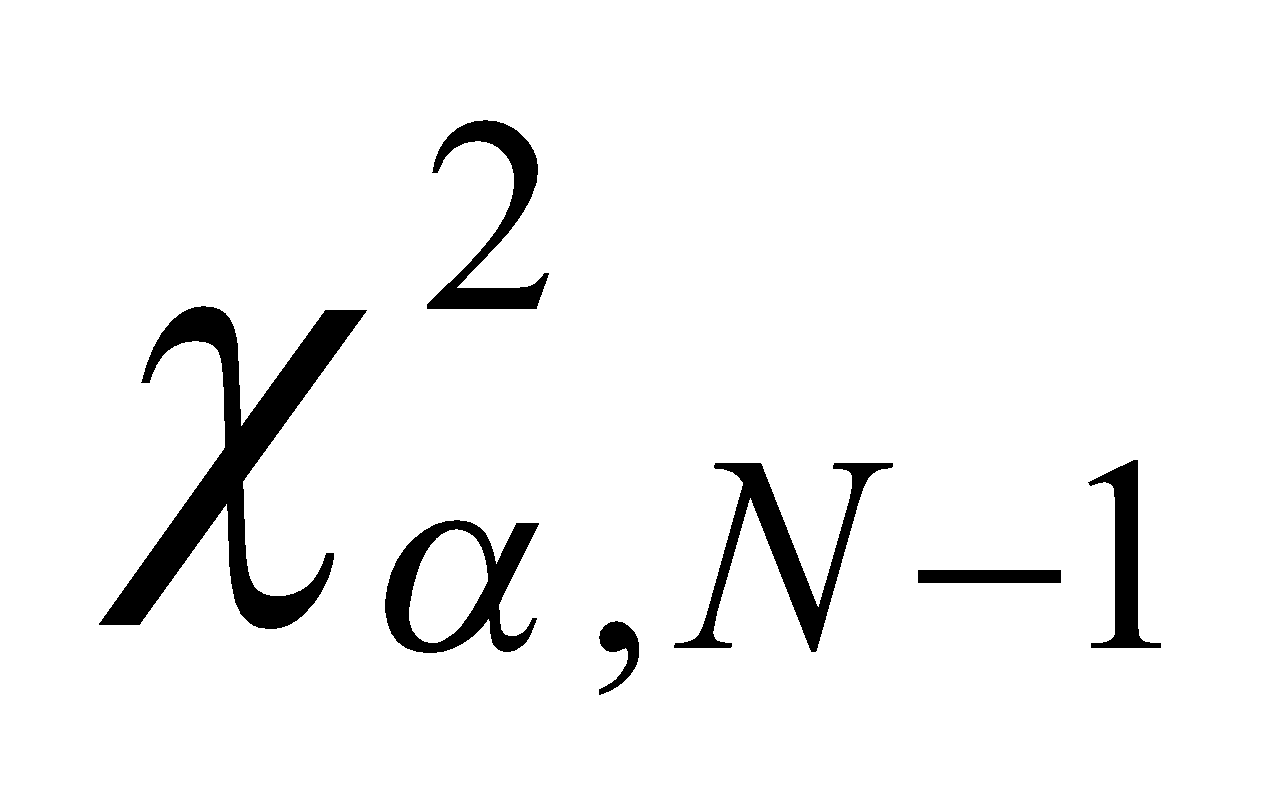
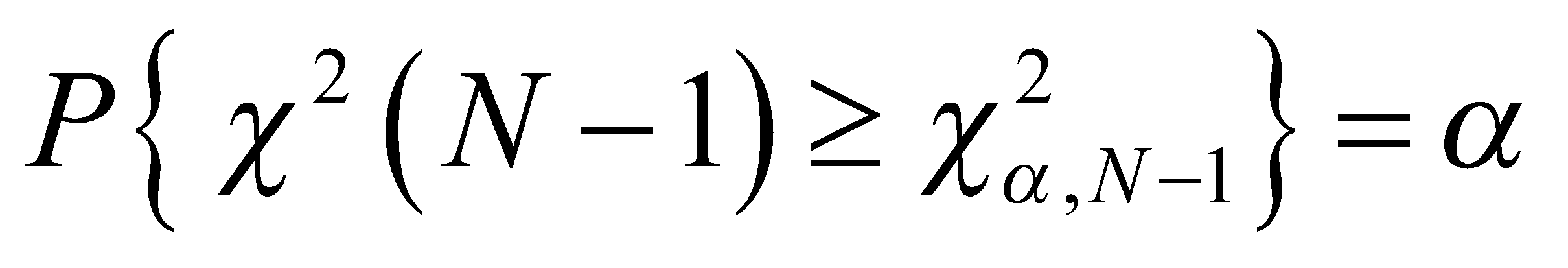
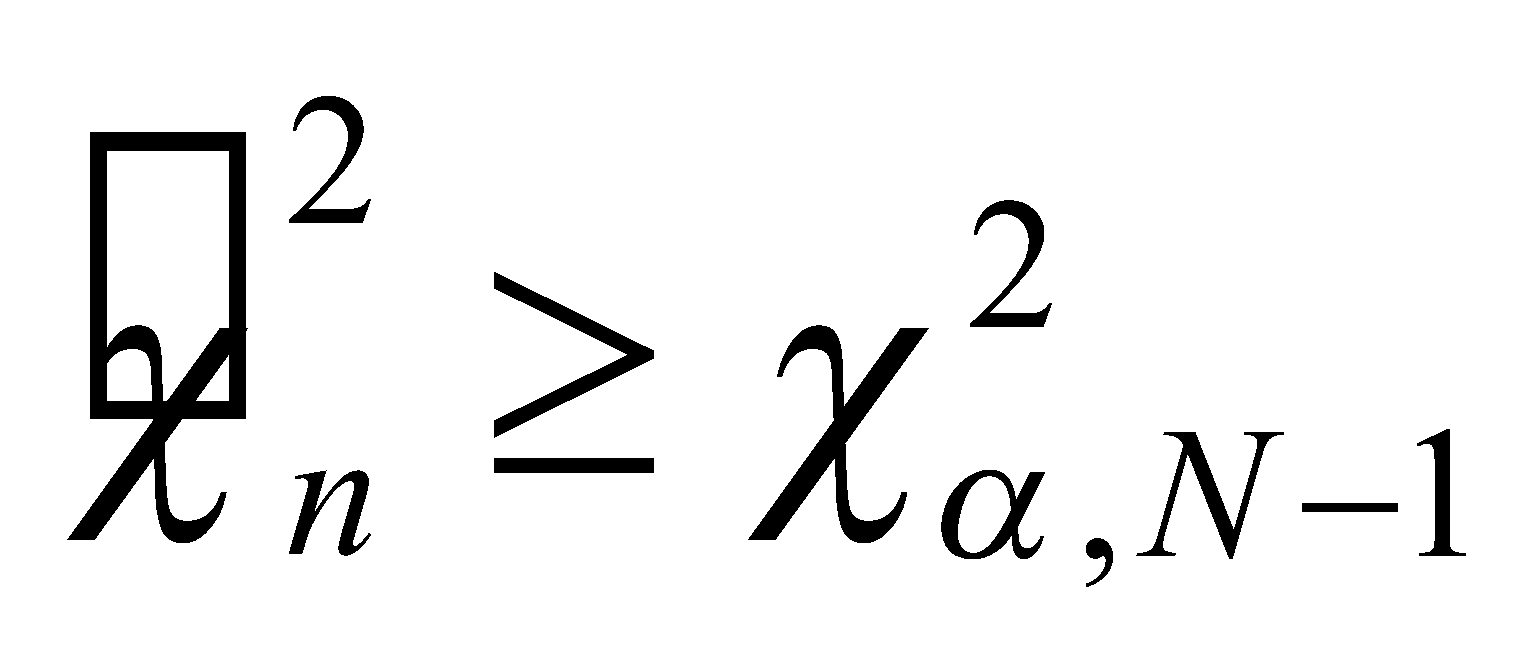
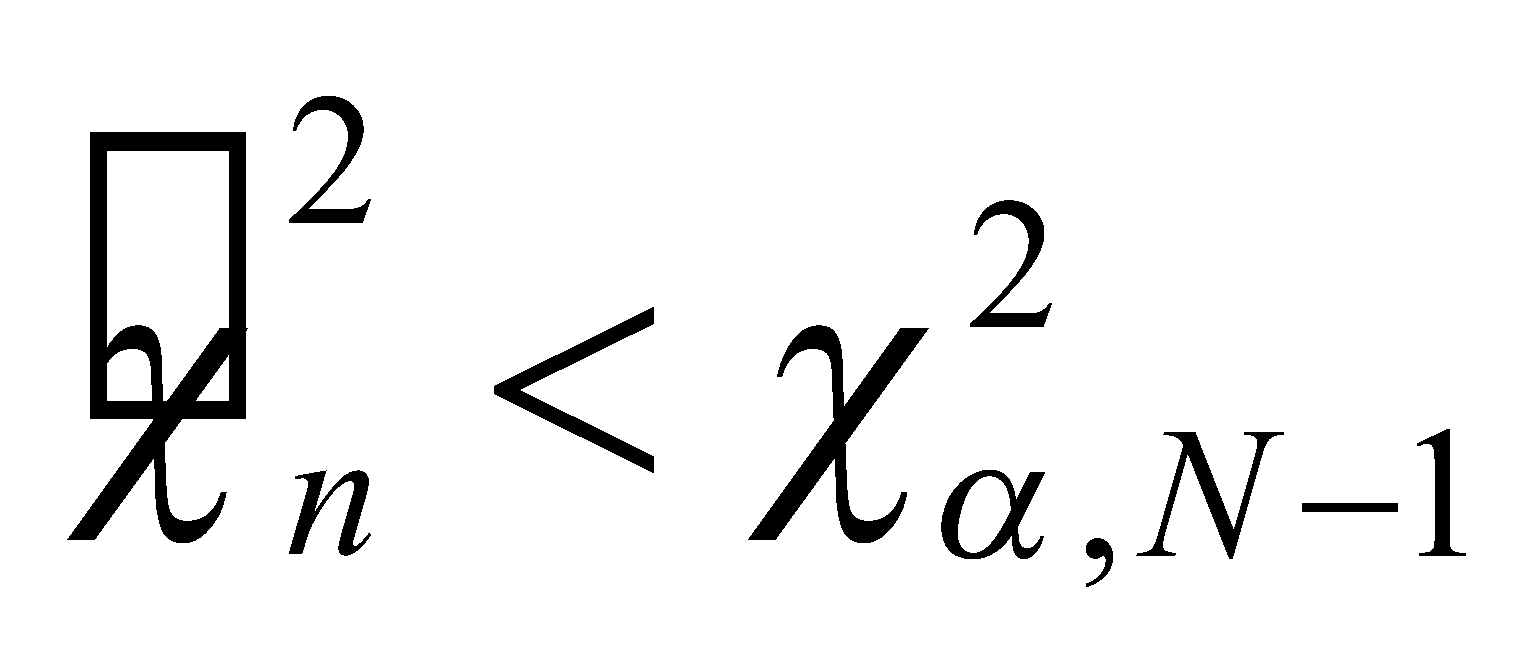
**Теорема 4.1.** *Якщо , , то при  розподіл величини  слабко збігається до - розподілу з* ****** *ступенем свободи.*

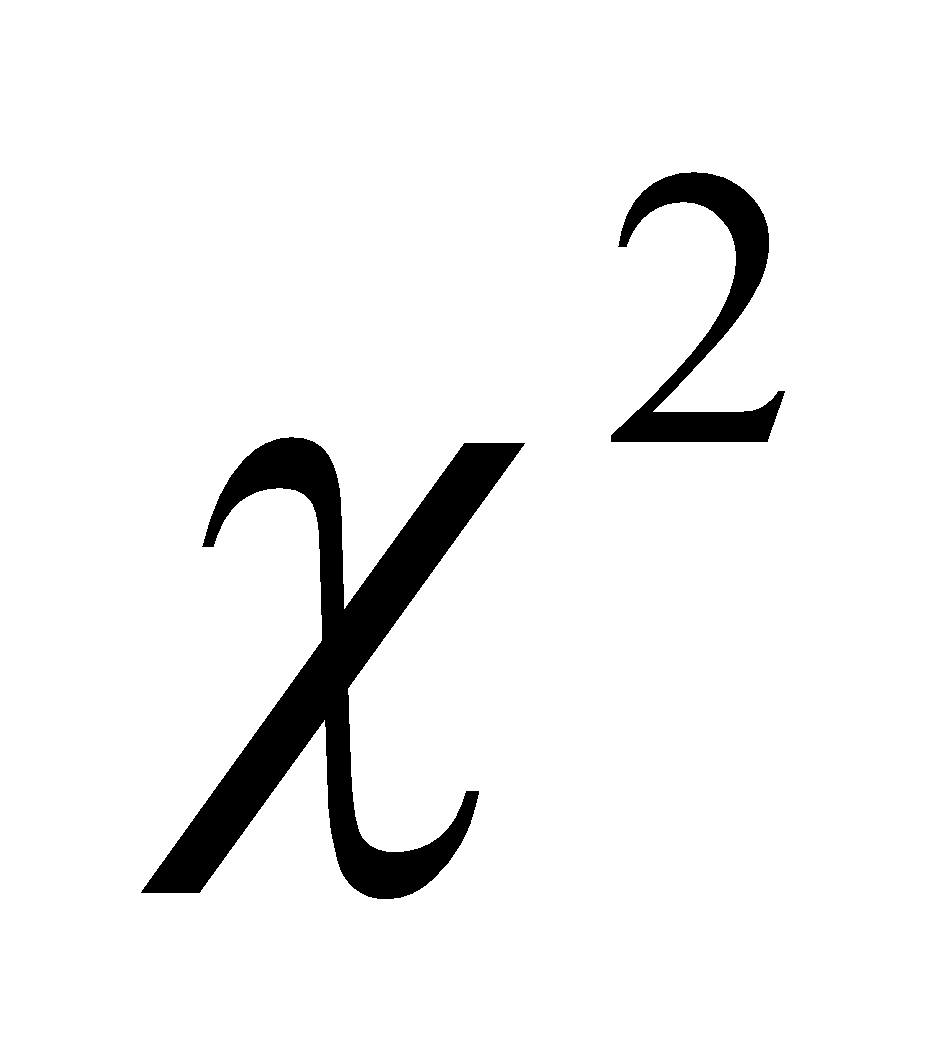
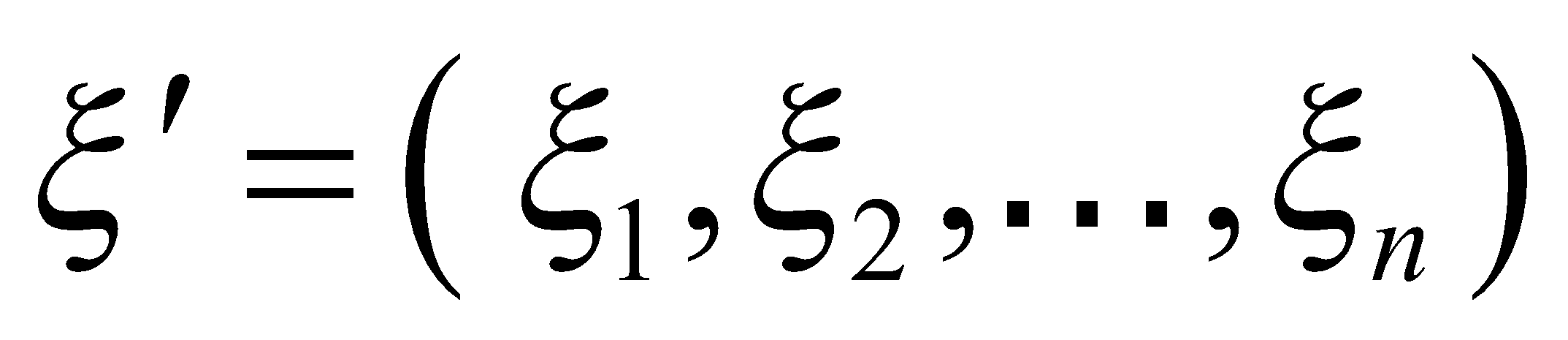
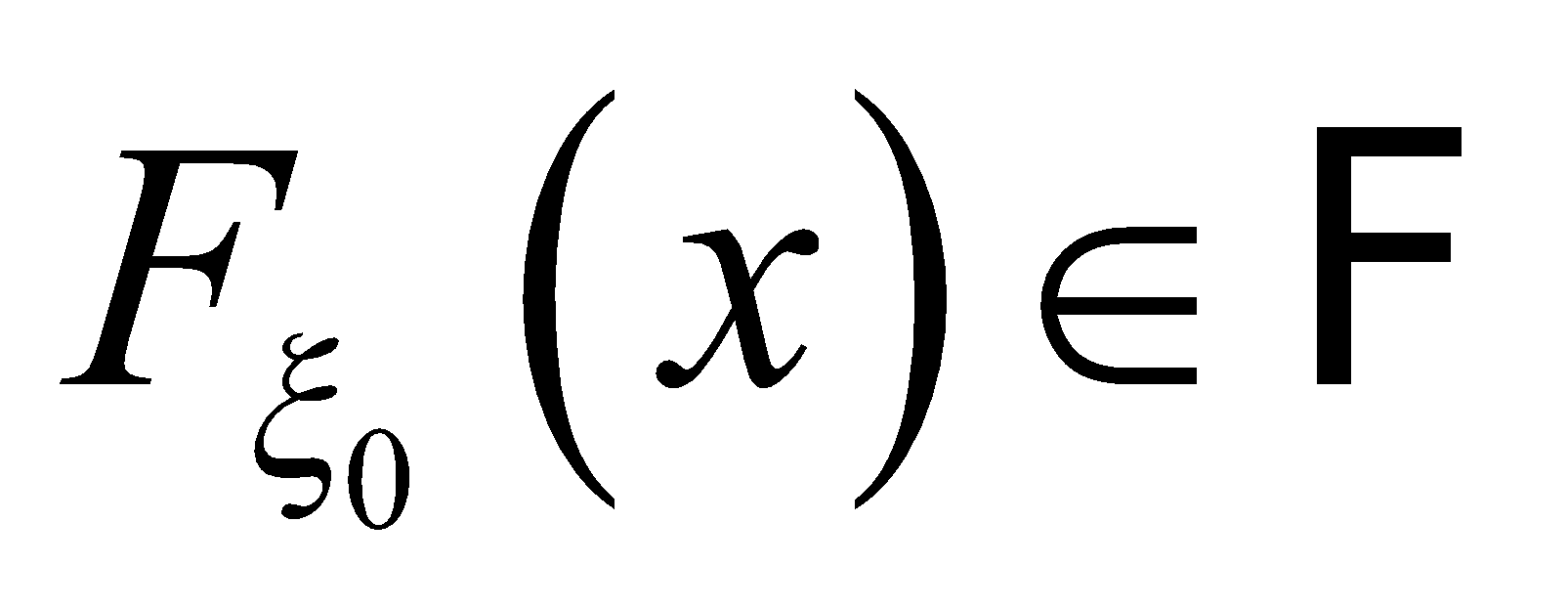
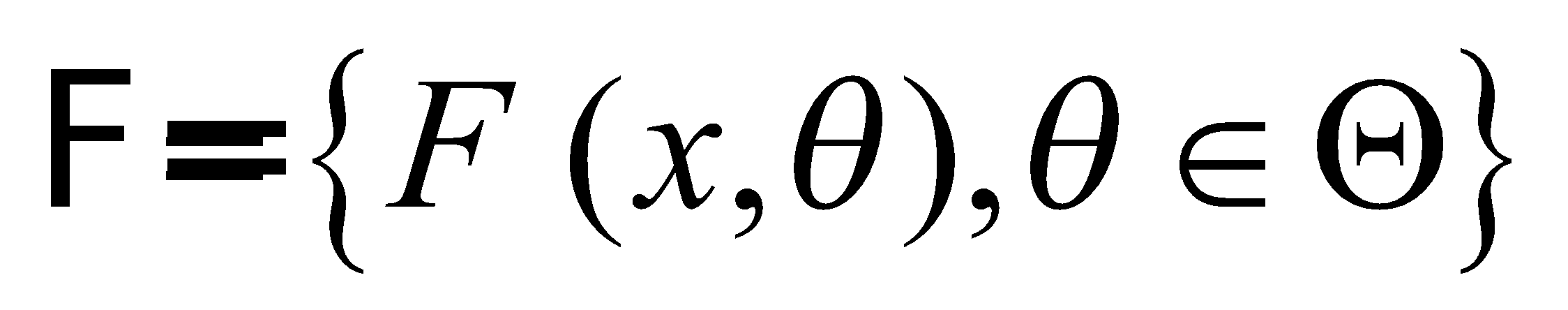
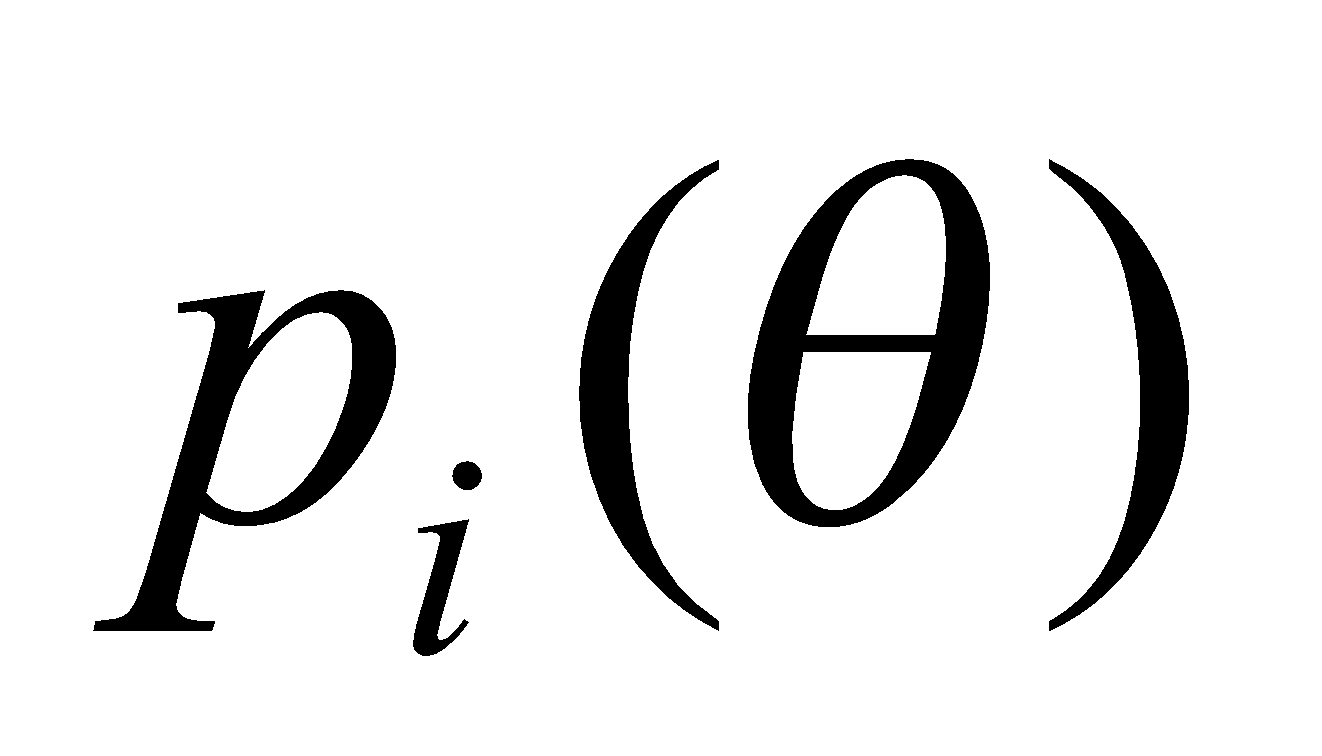
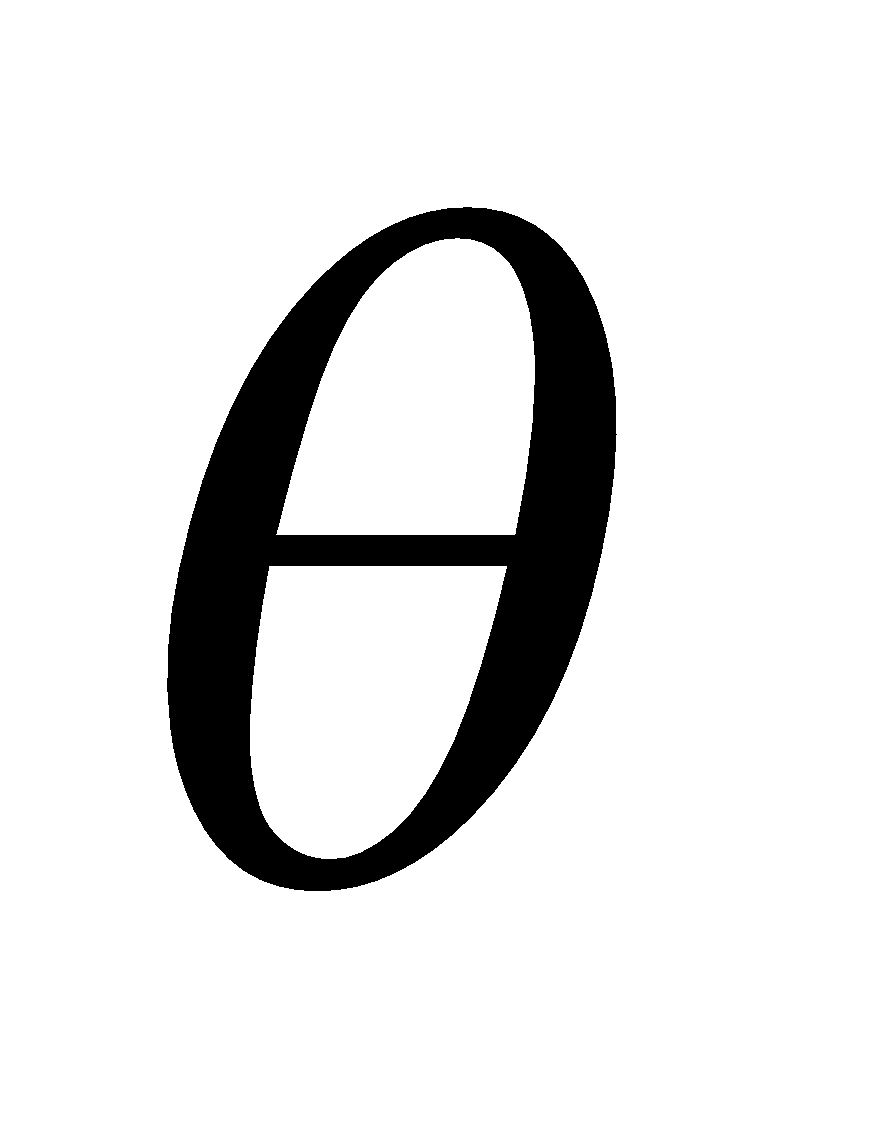
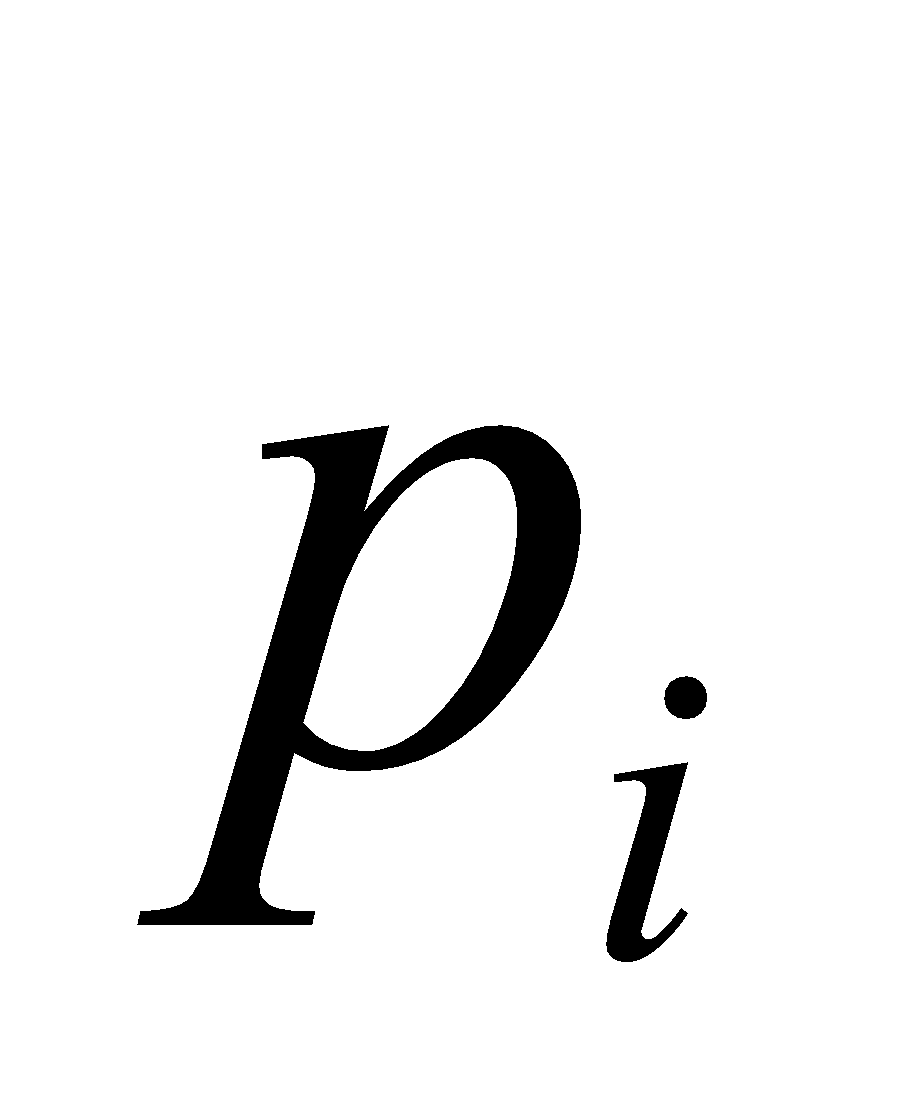
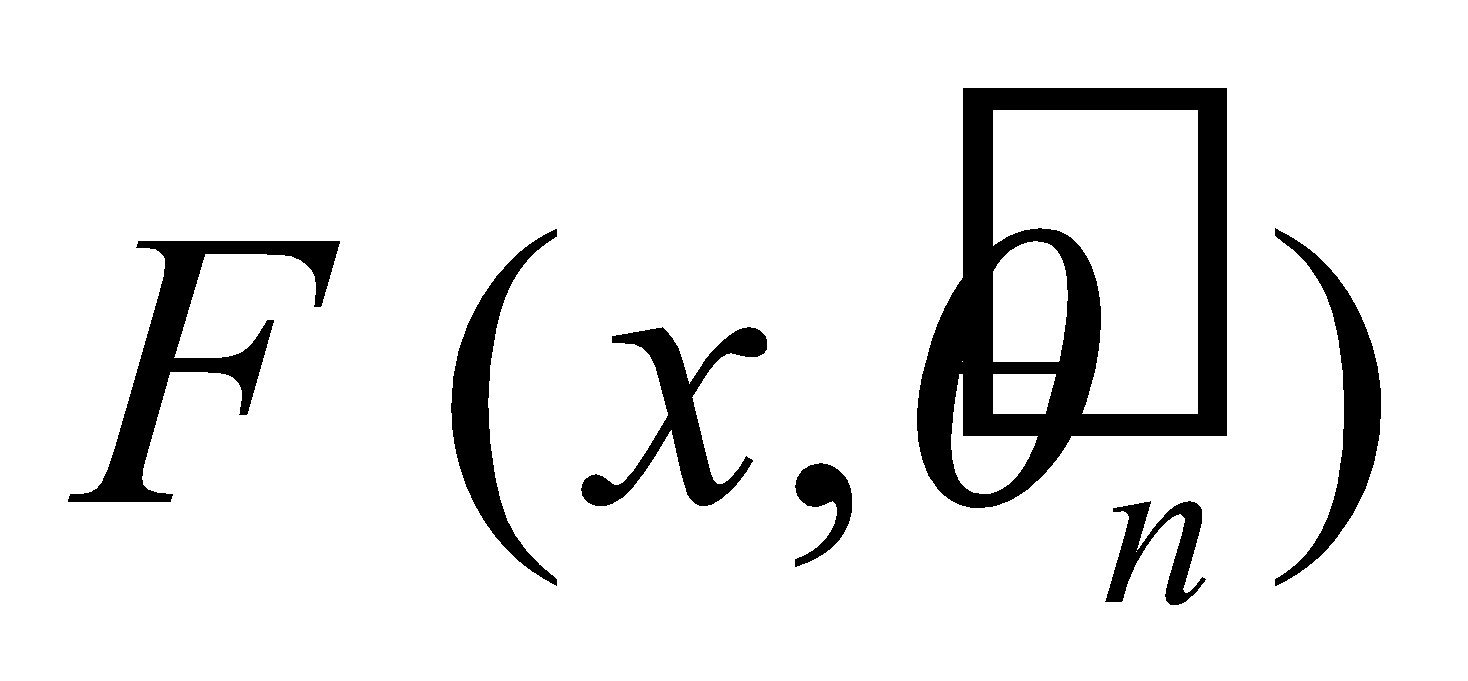
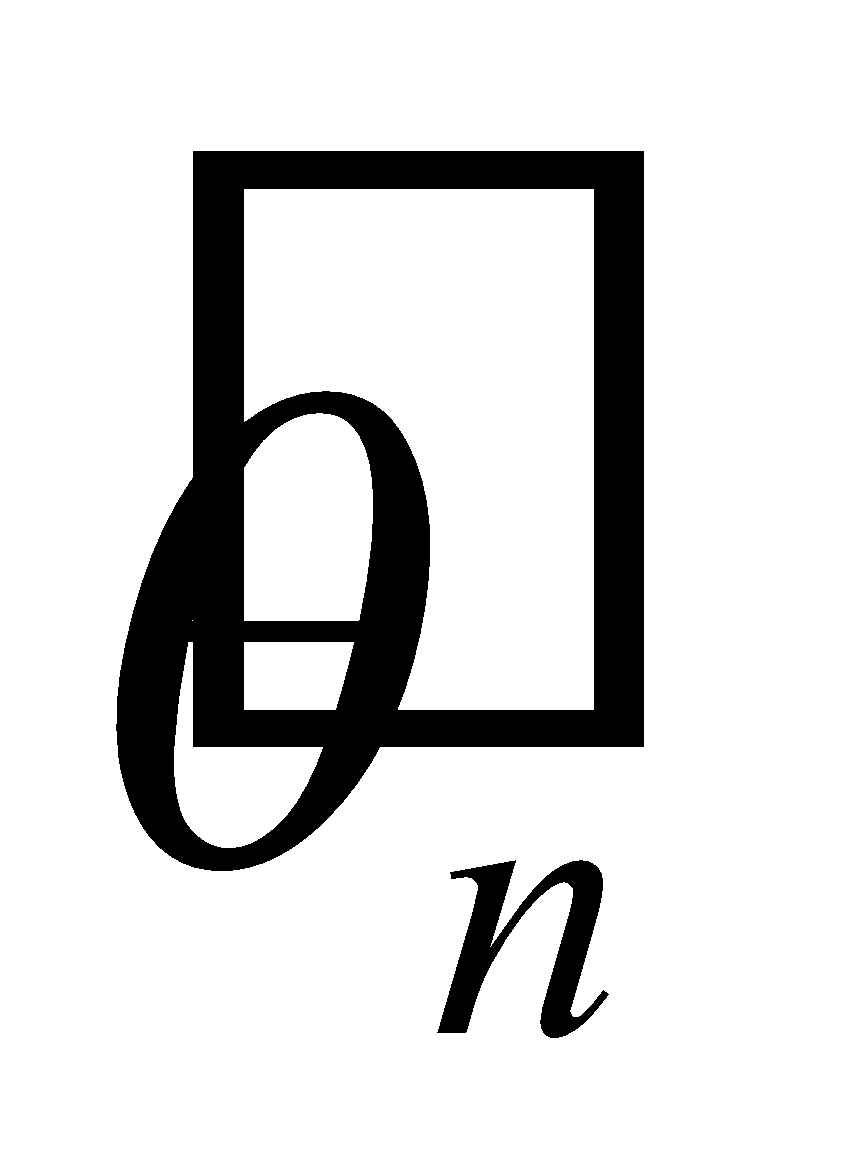
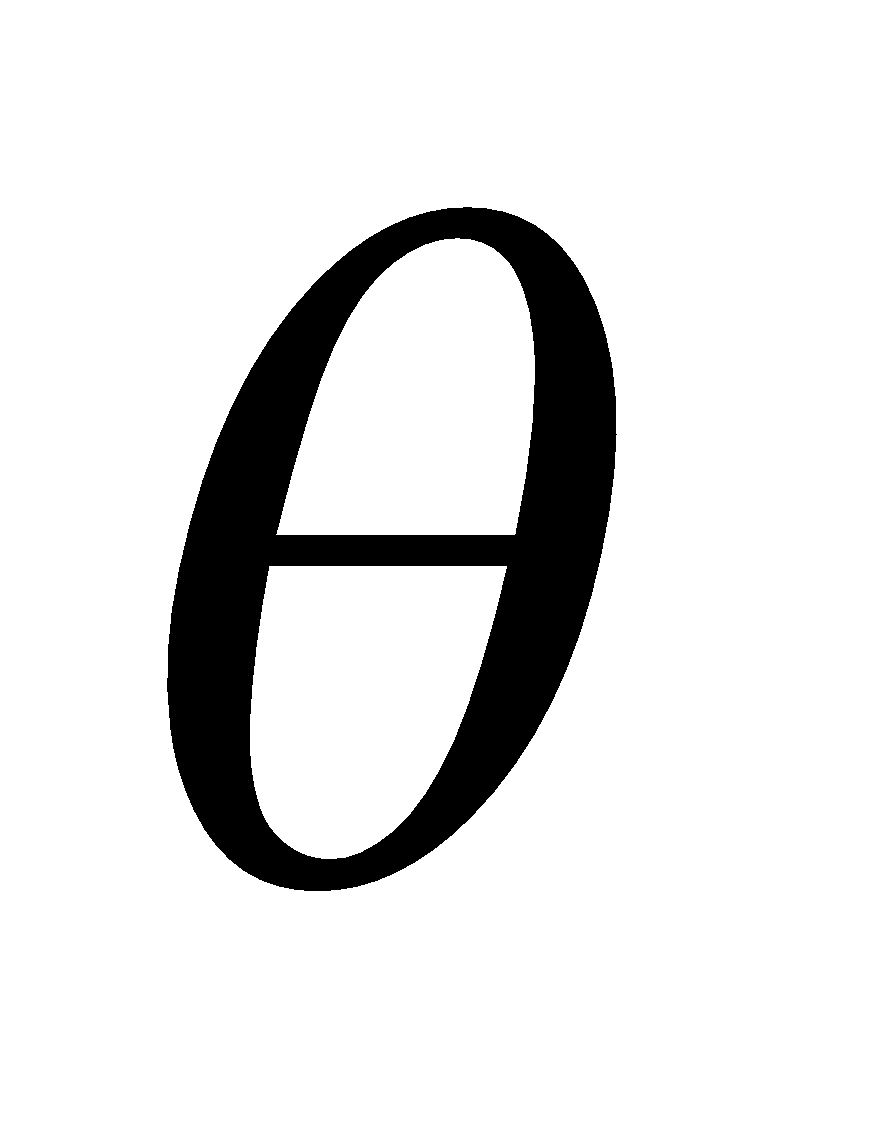
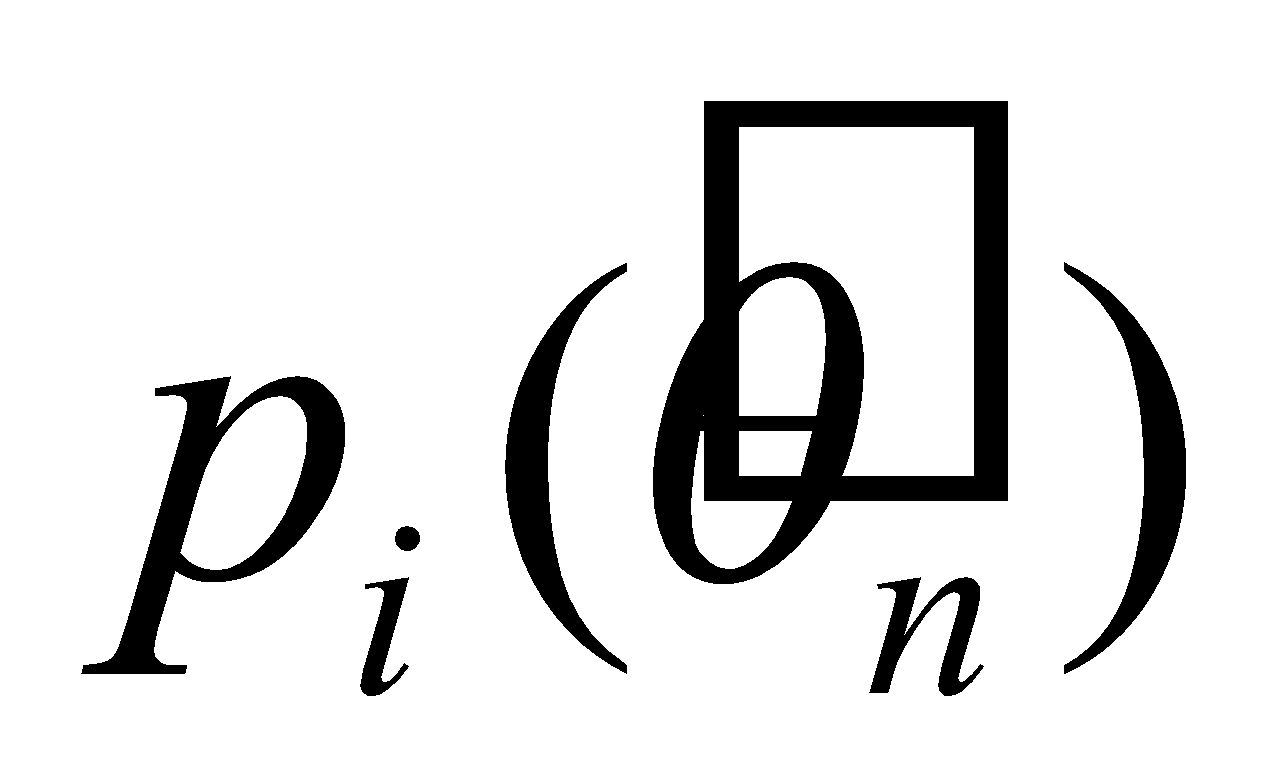
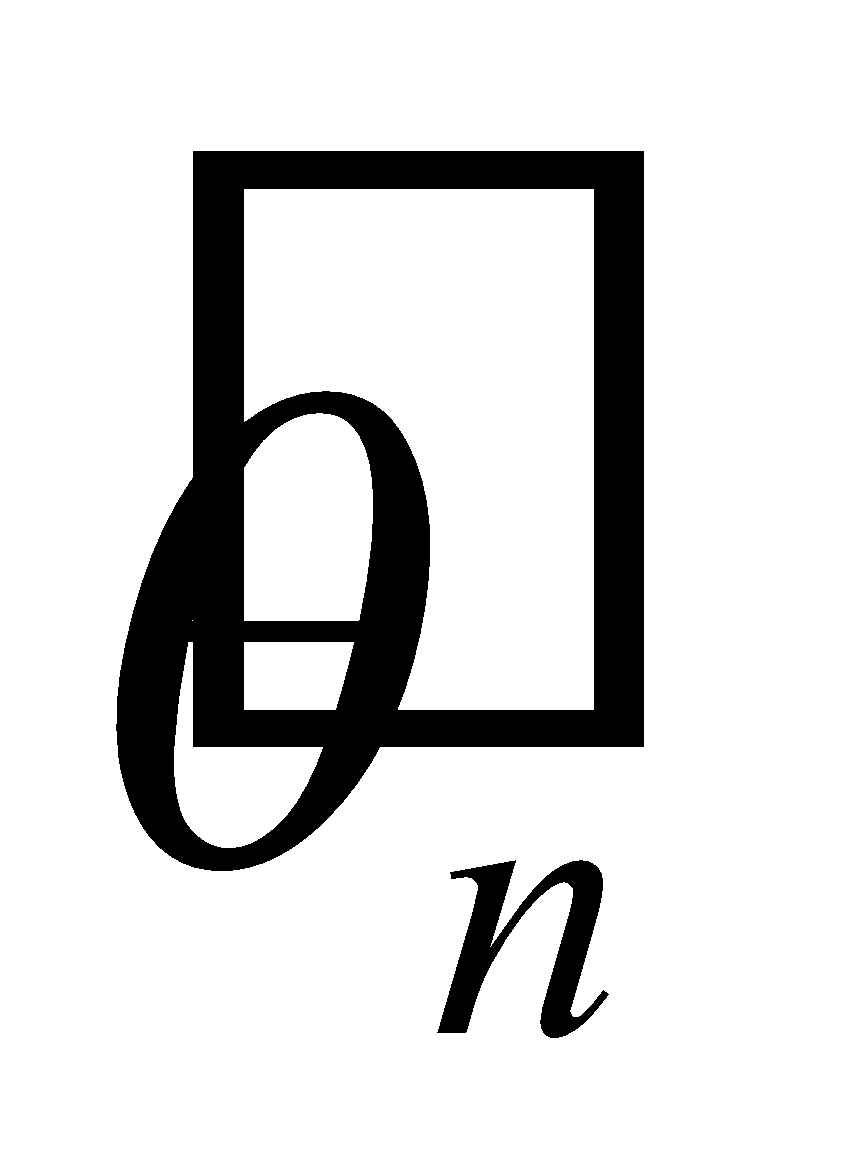
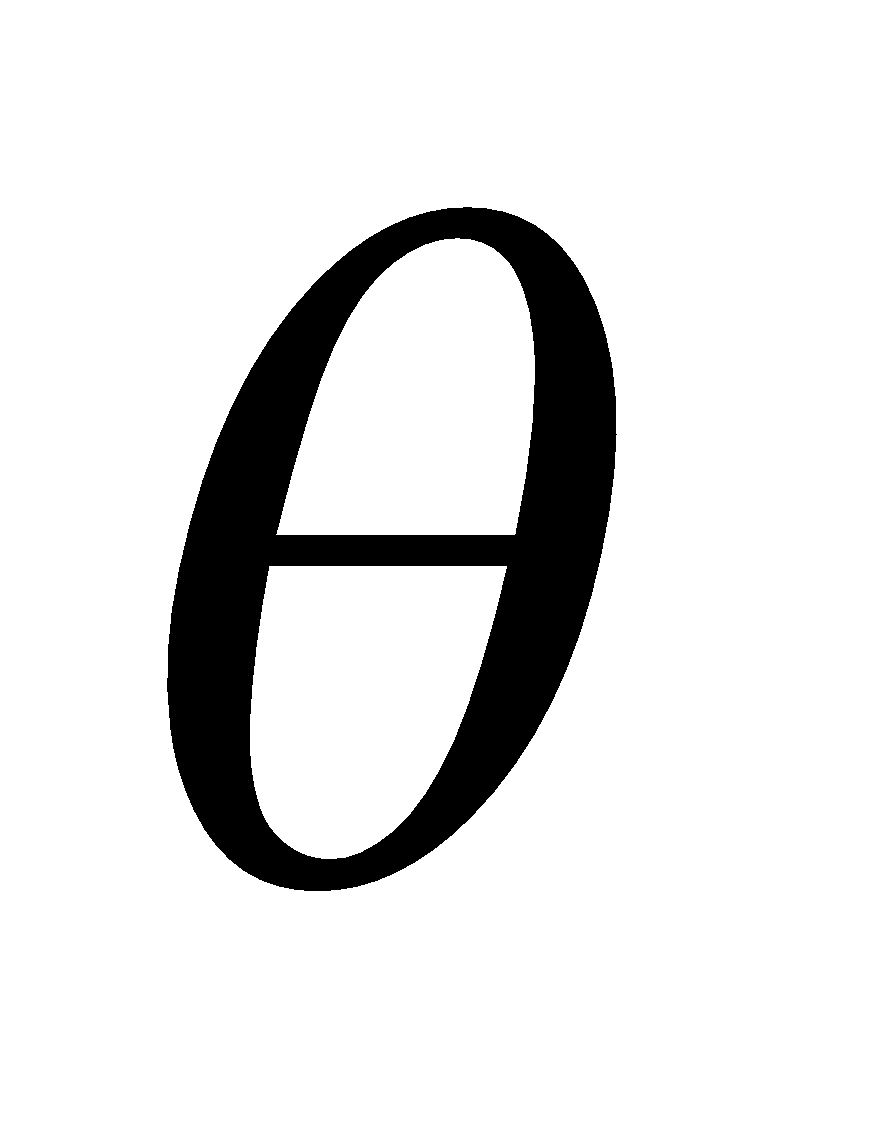
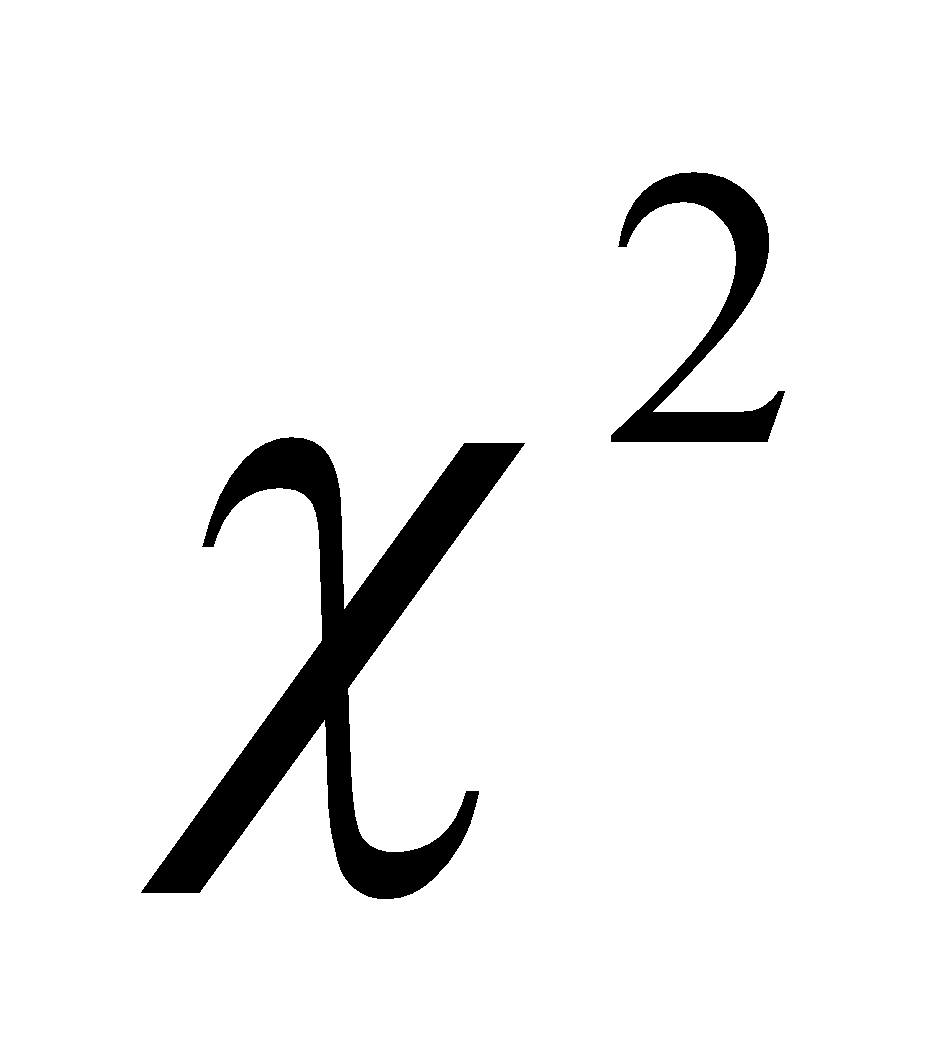
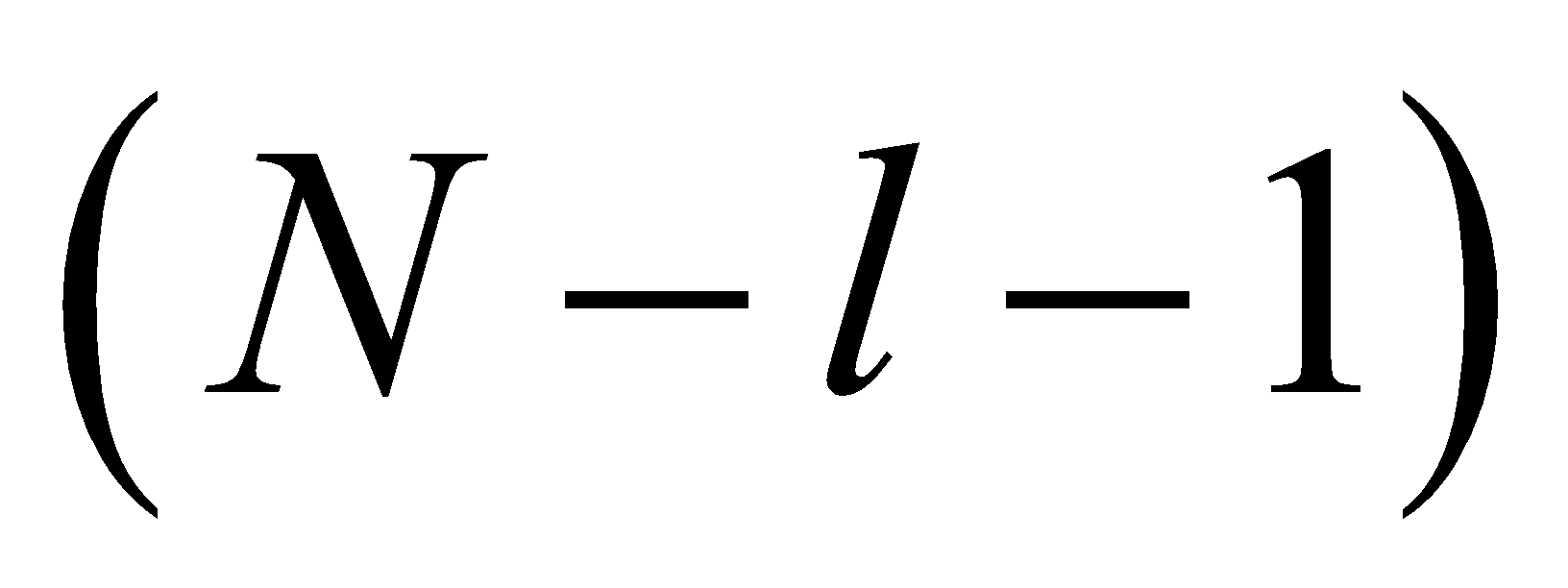
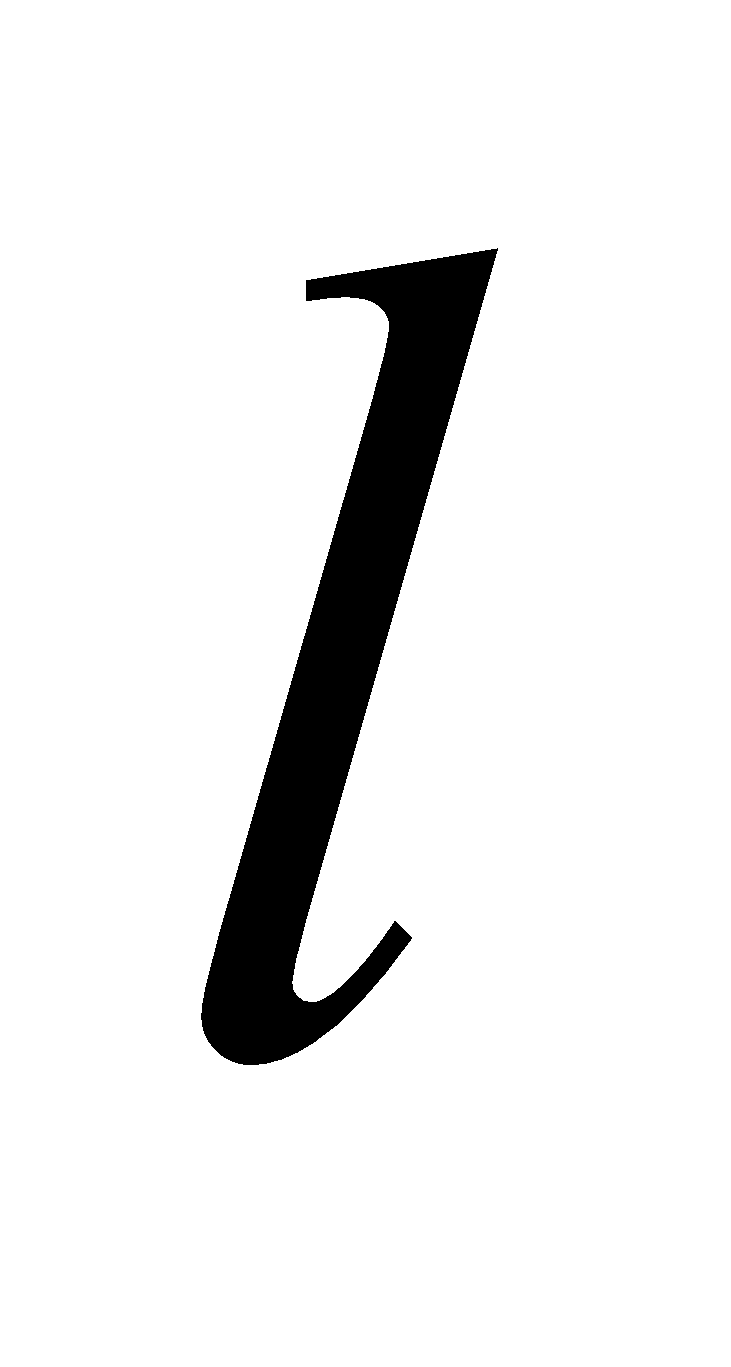
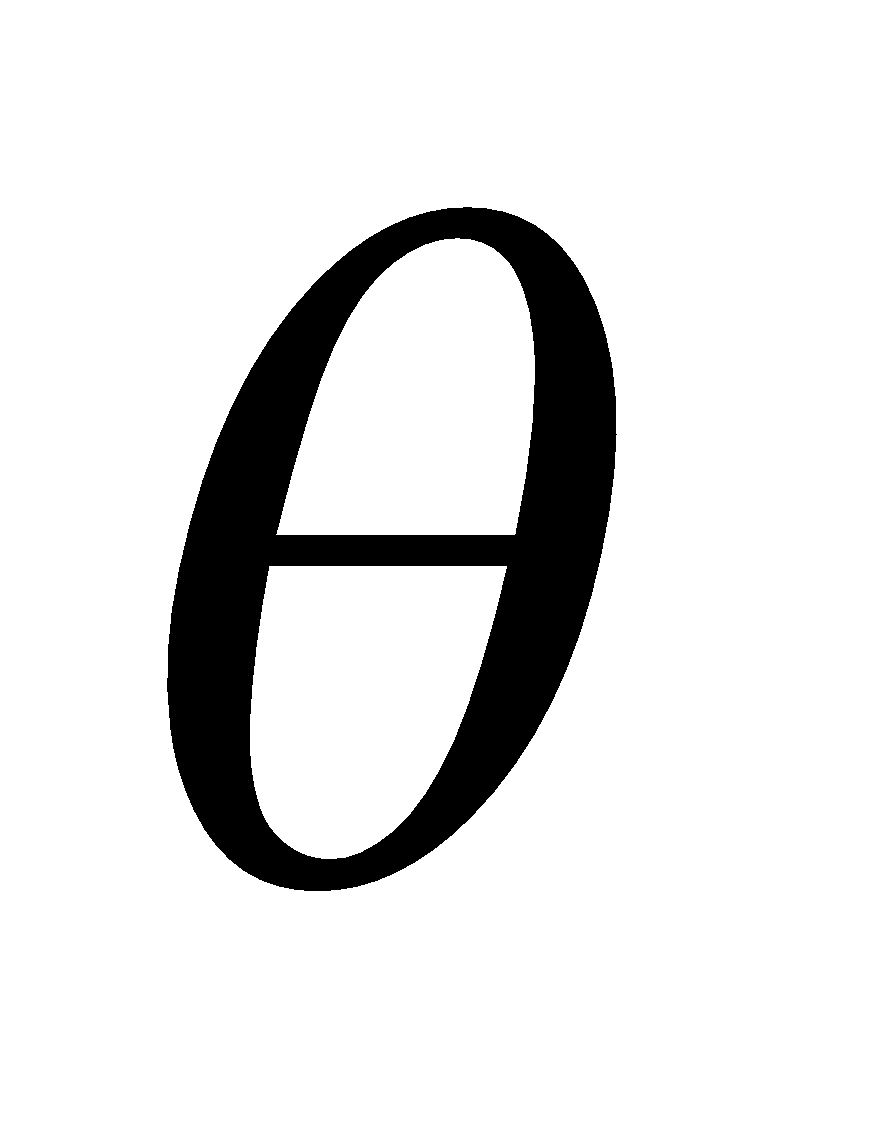
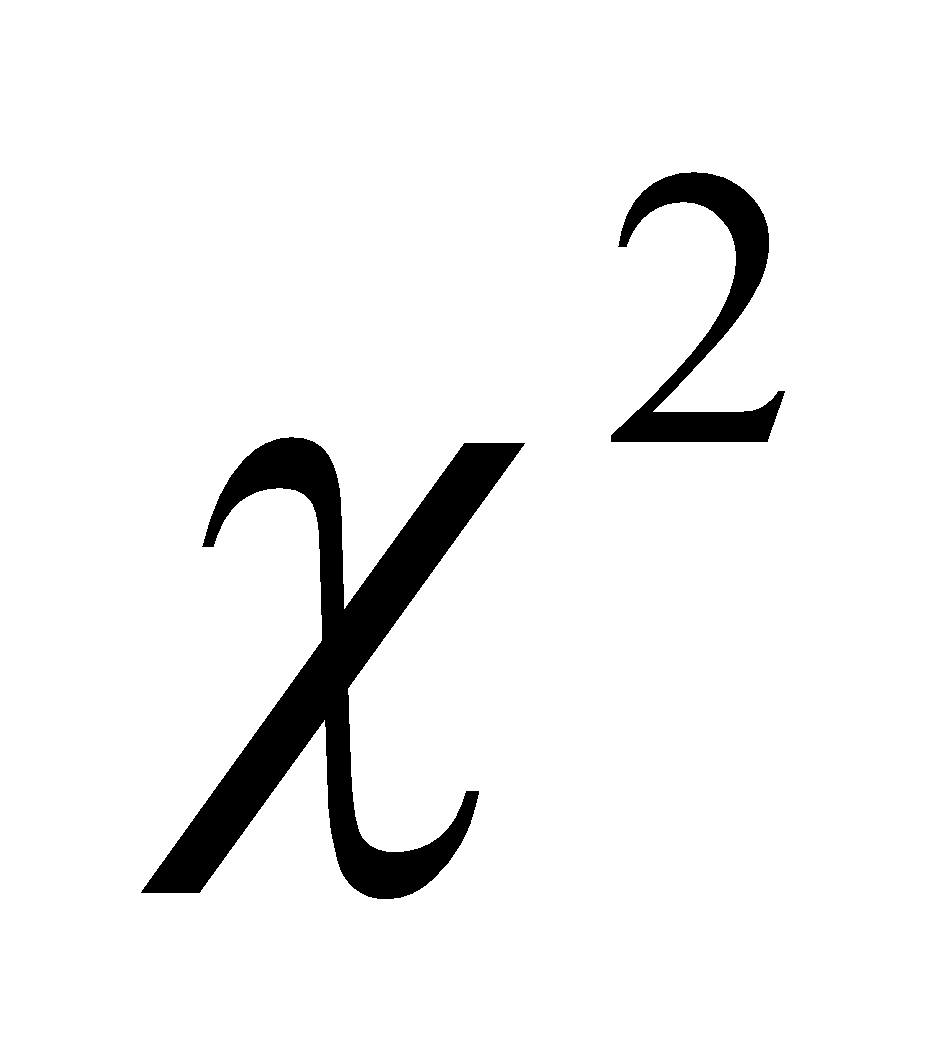
Використовуючи апроксимацію розподілу статистики **** розподілом хі-квадрат, маємо

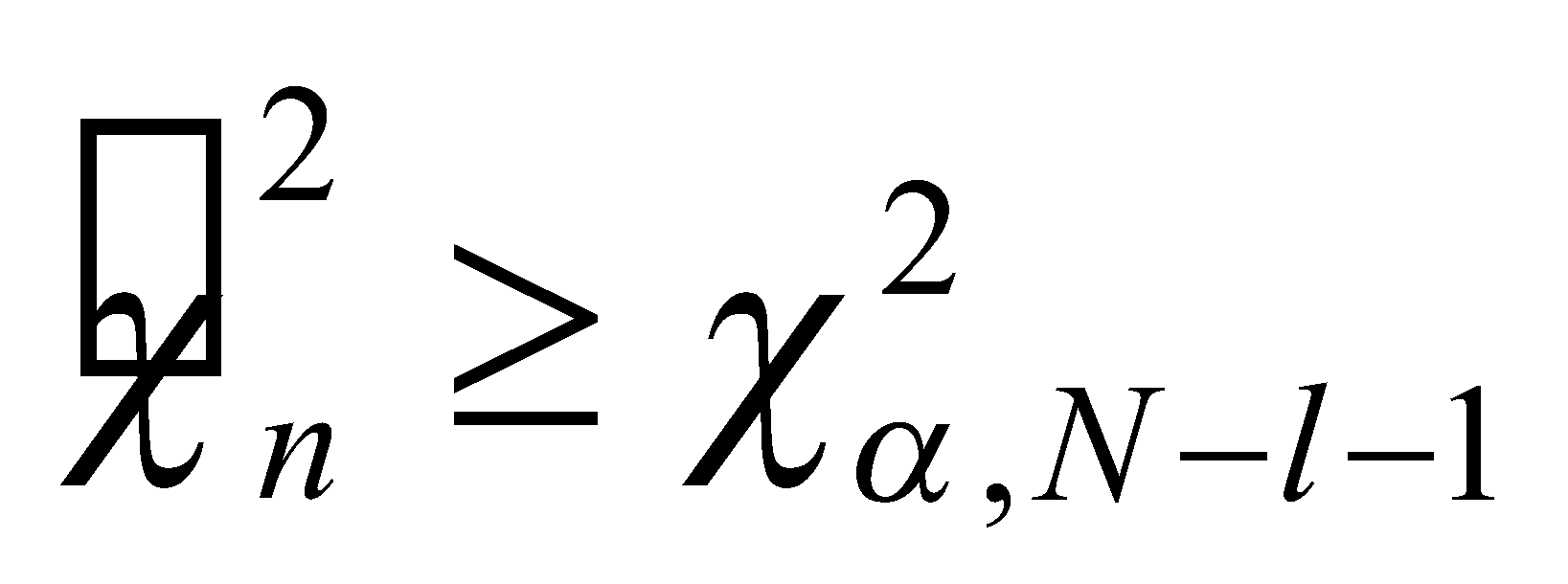
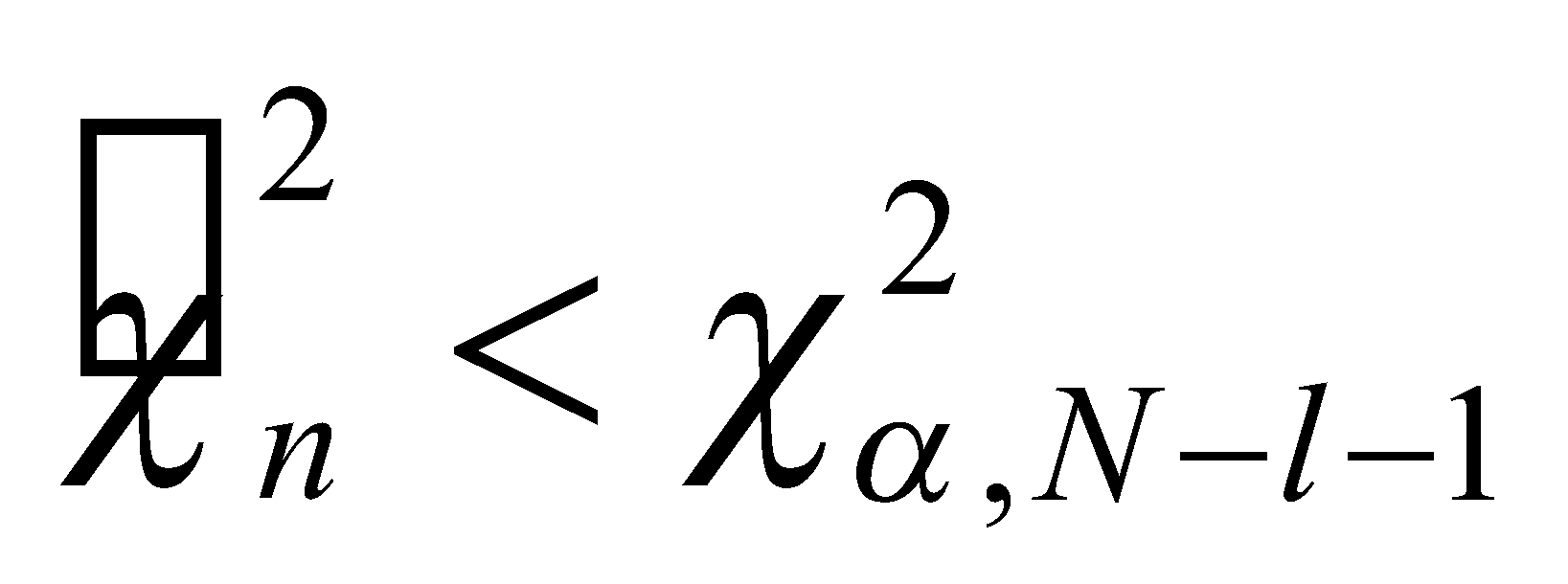
.

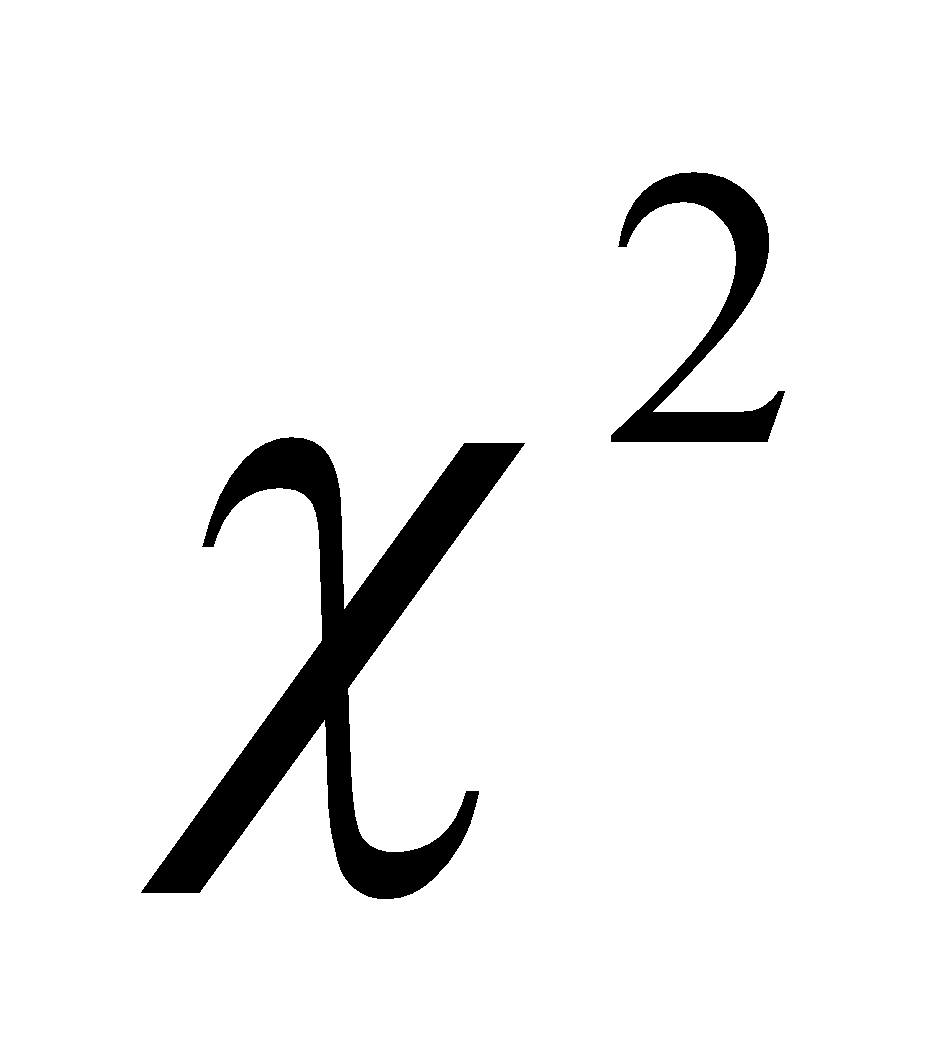
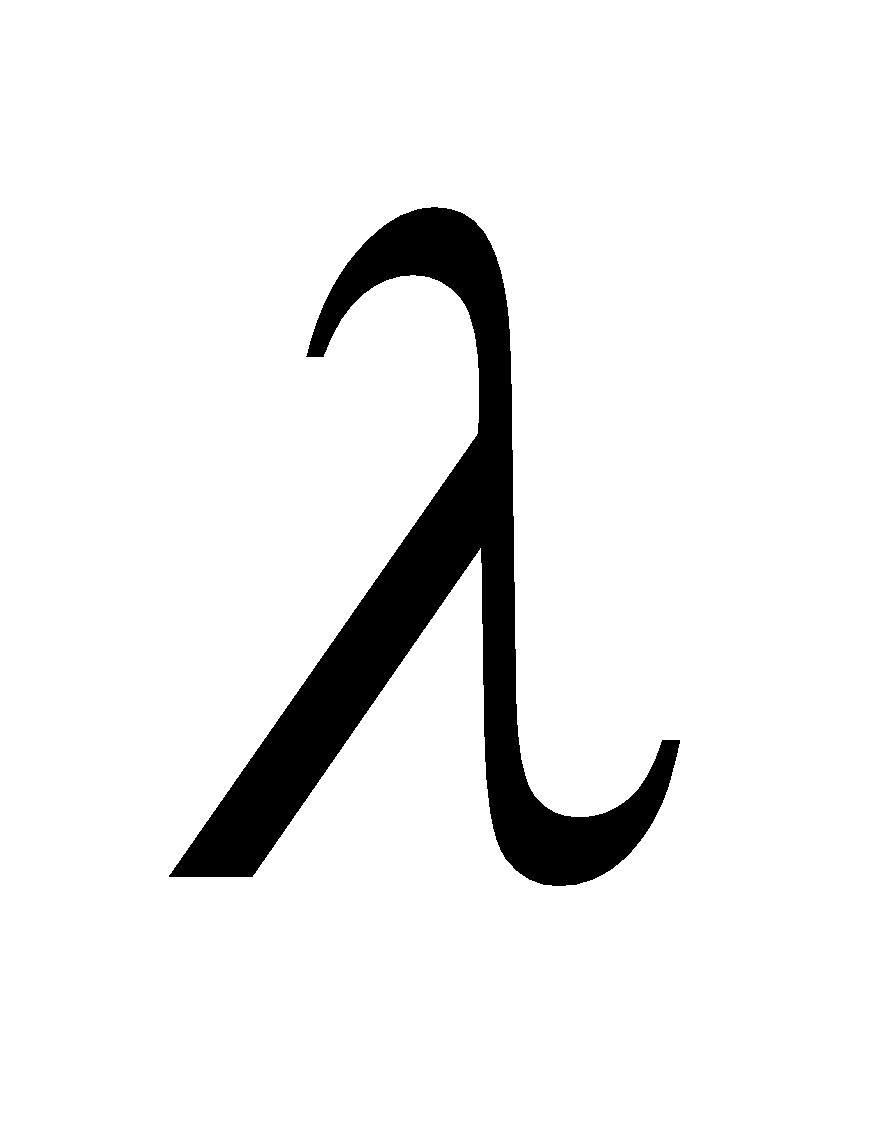
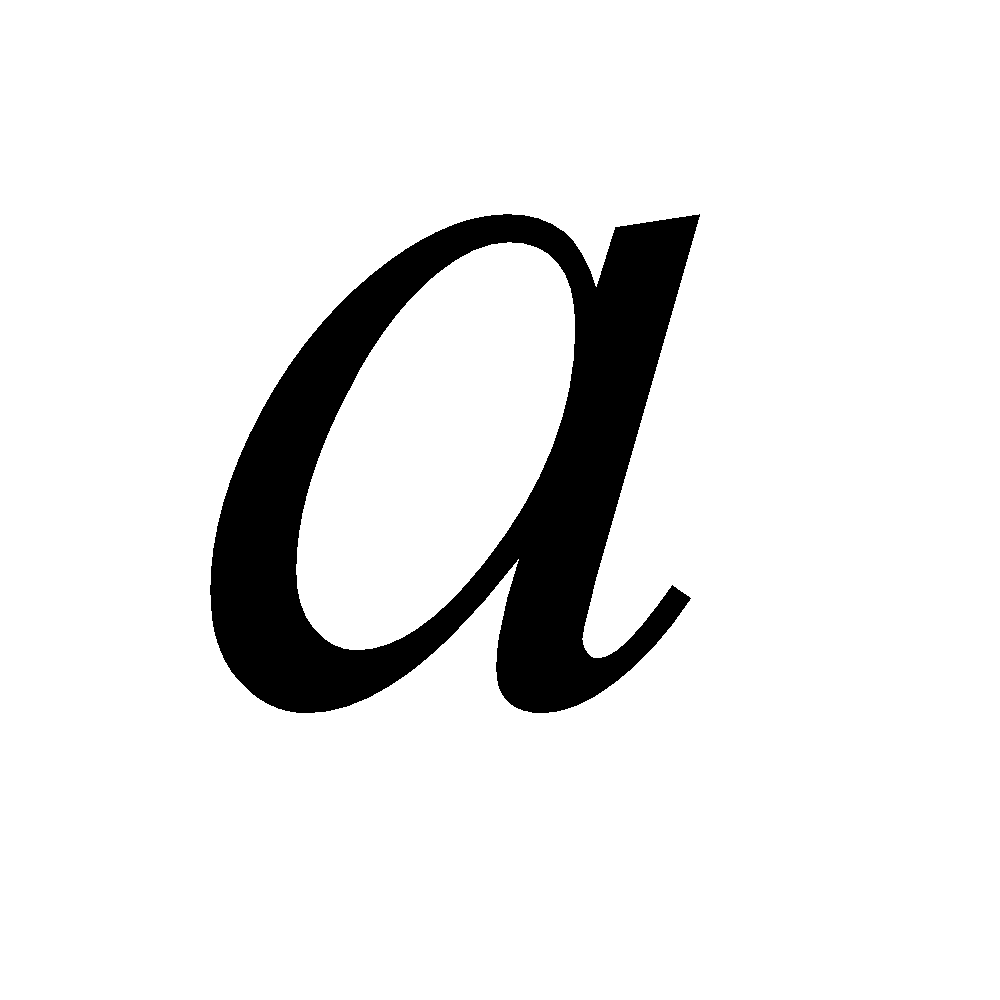
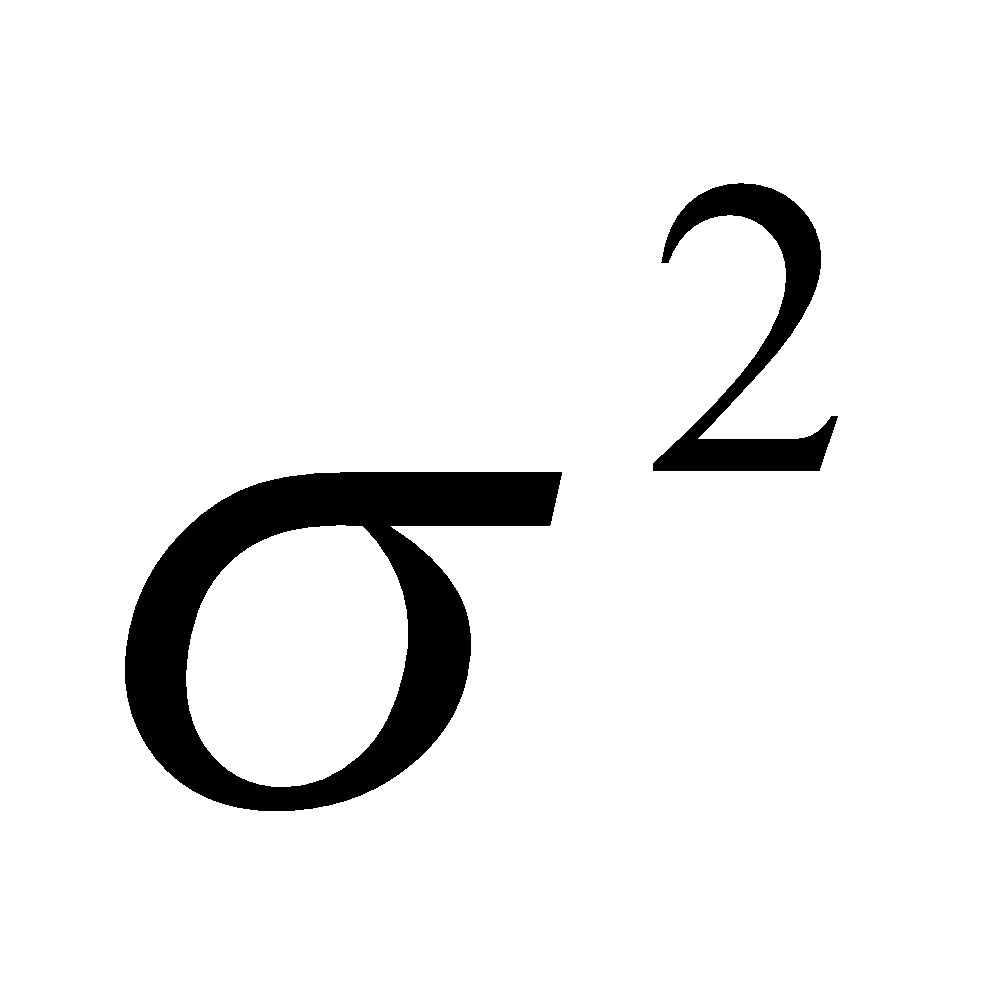
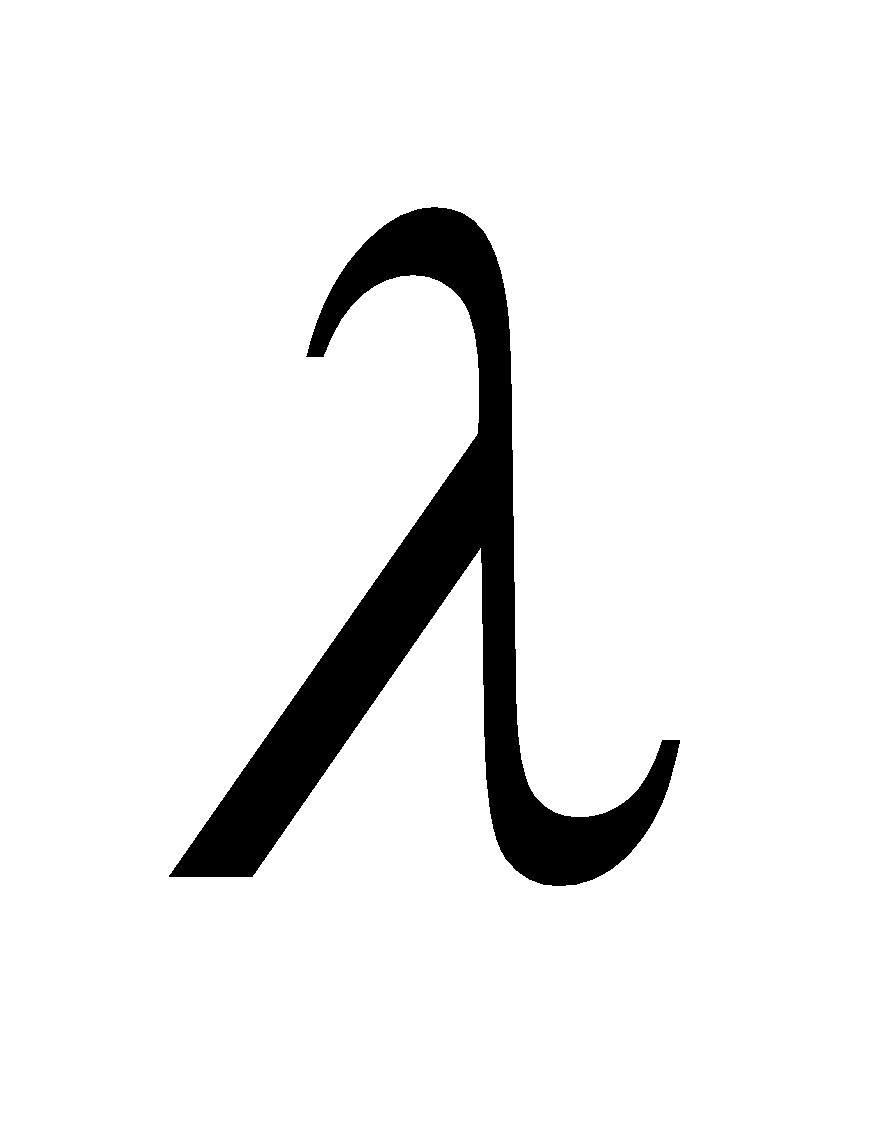
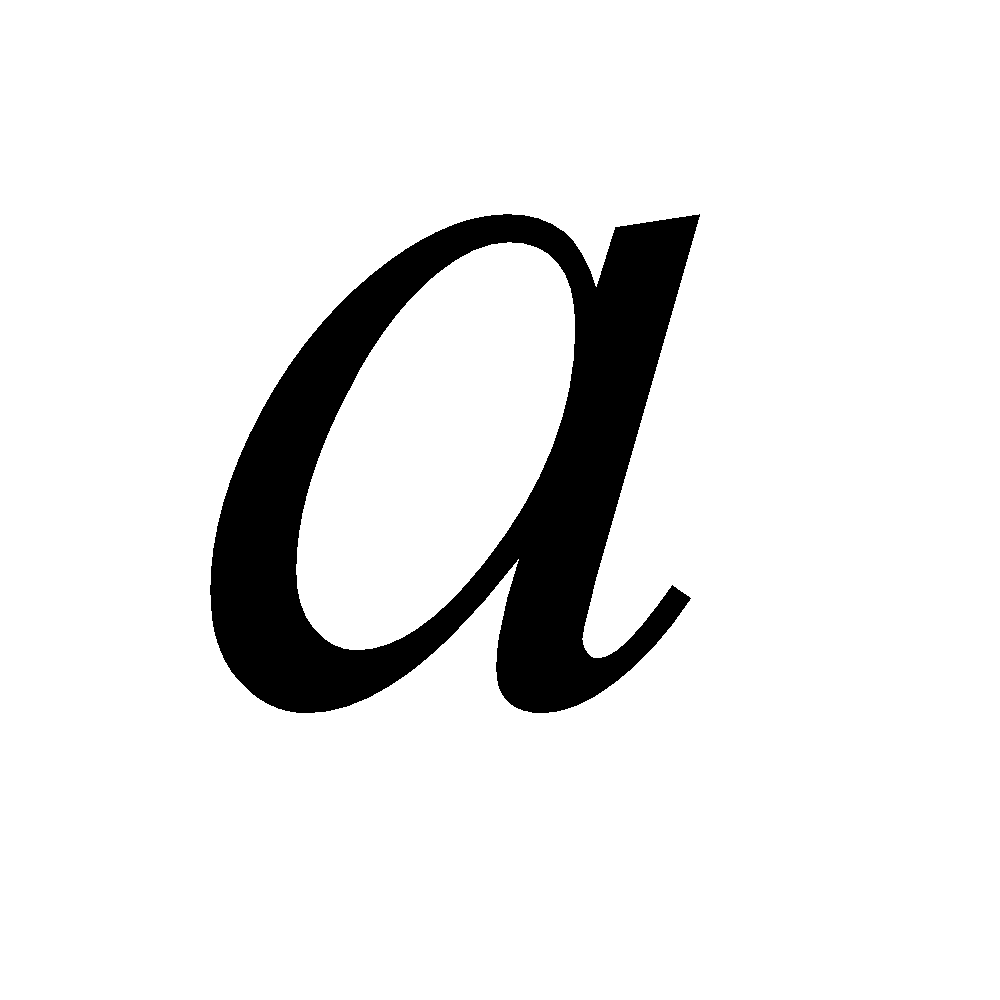
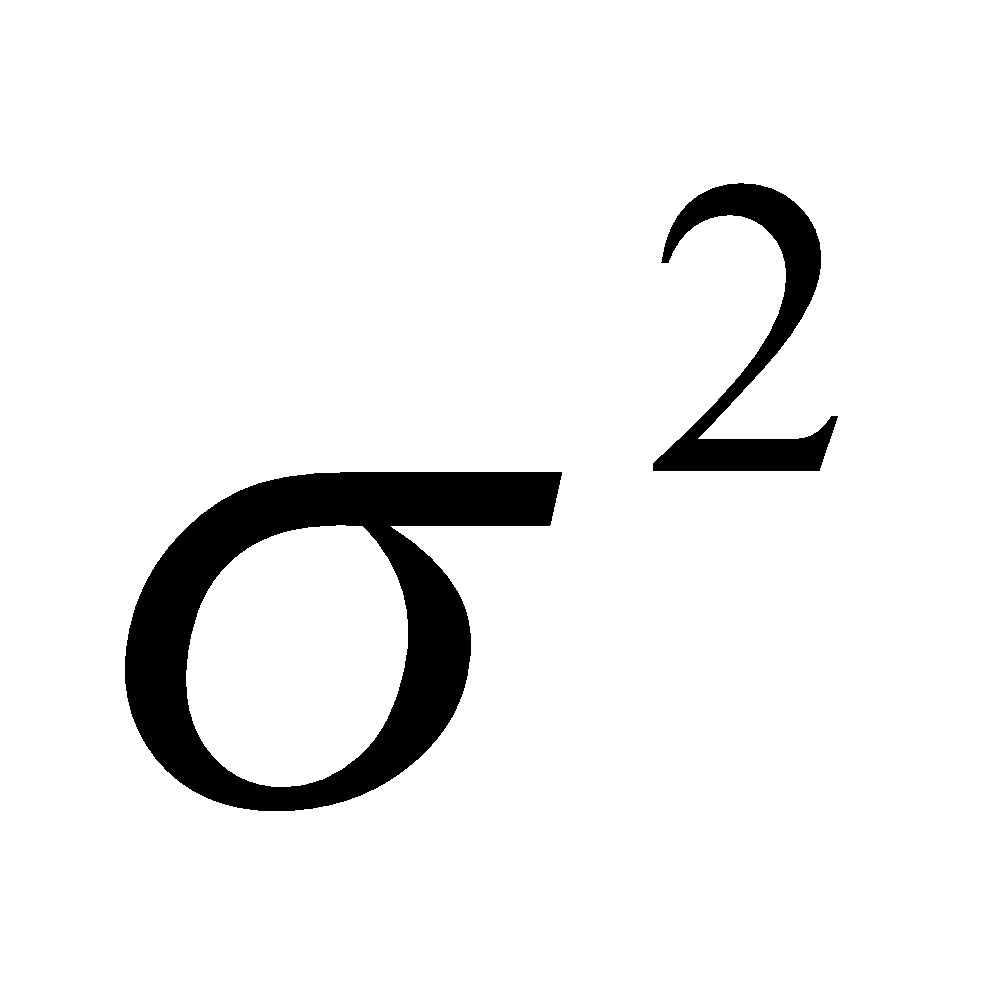
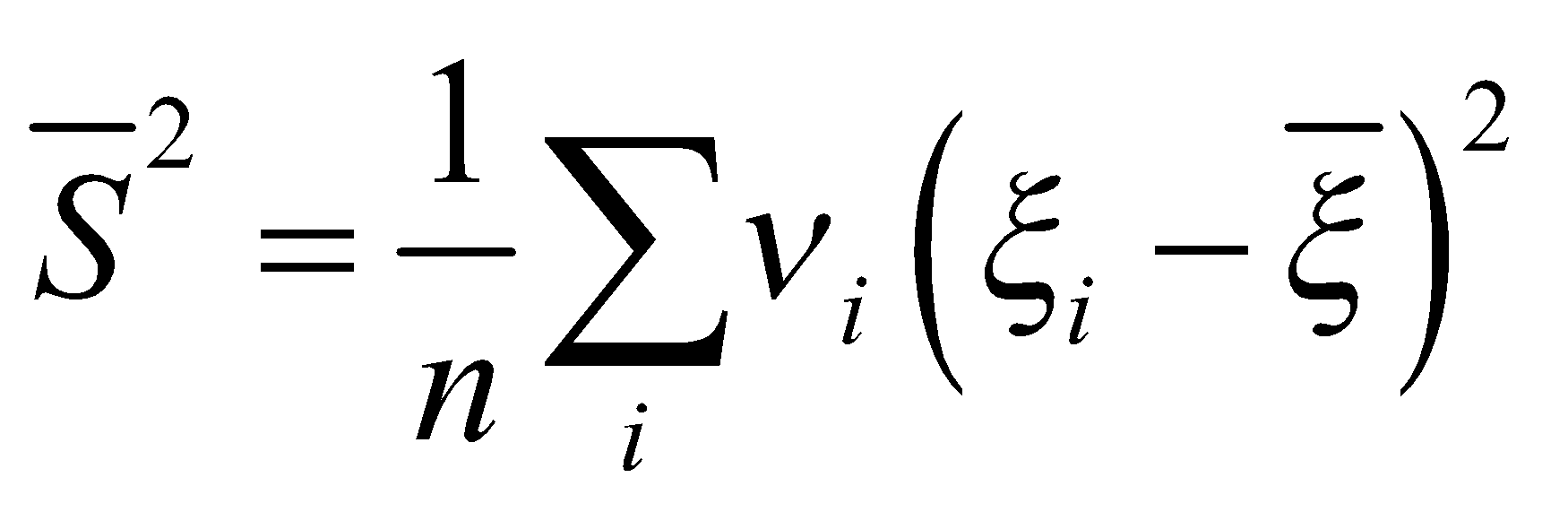
Таким чином критична область представляє собою множину .

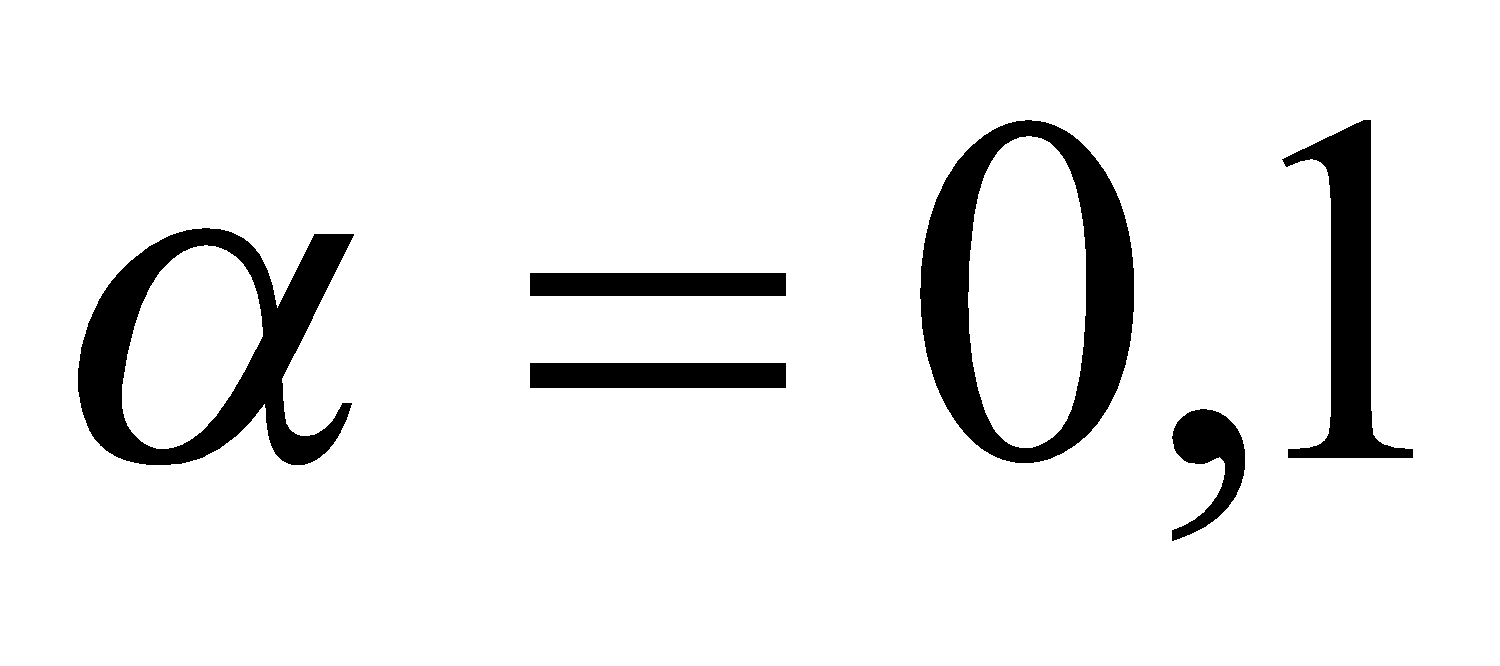
На практиці граничний розподіл  можна використовувати з непоганим наближенням вже при  і .

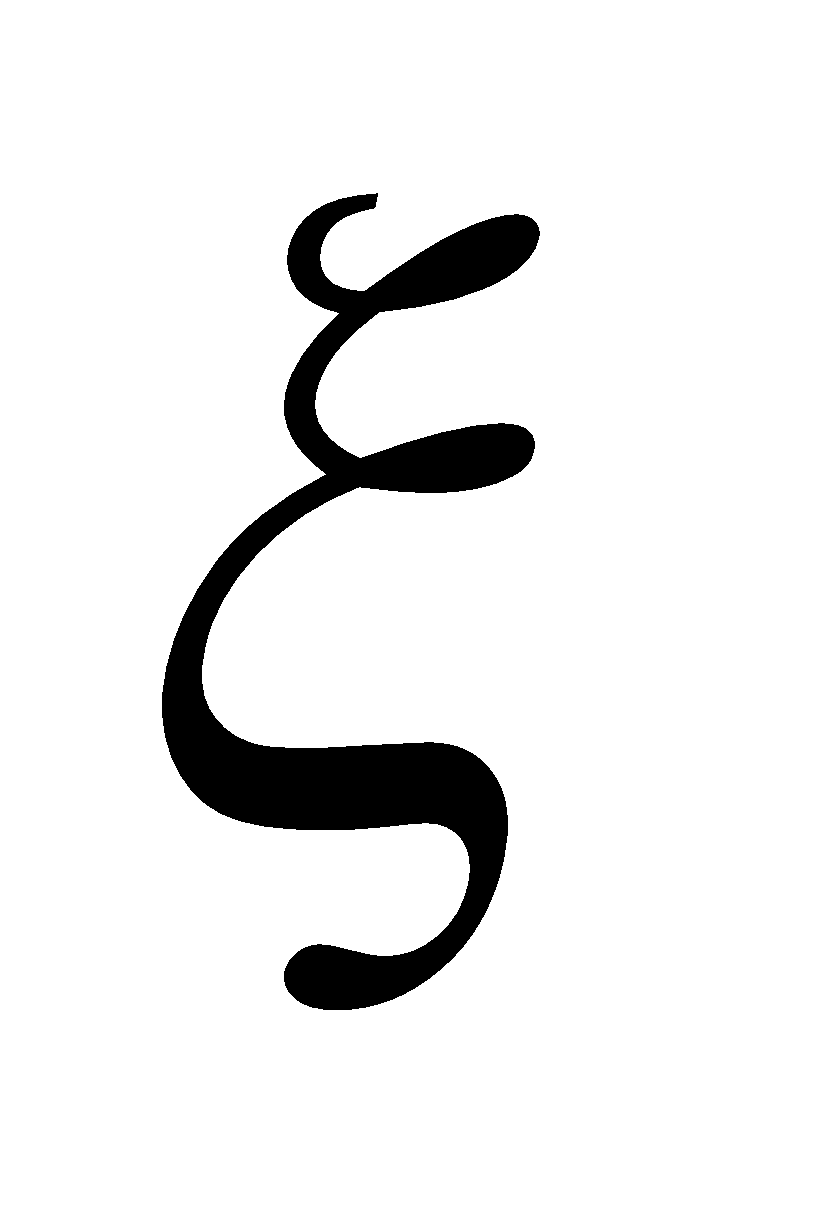
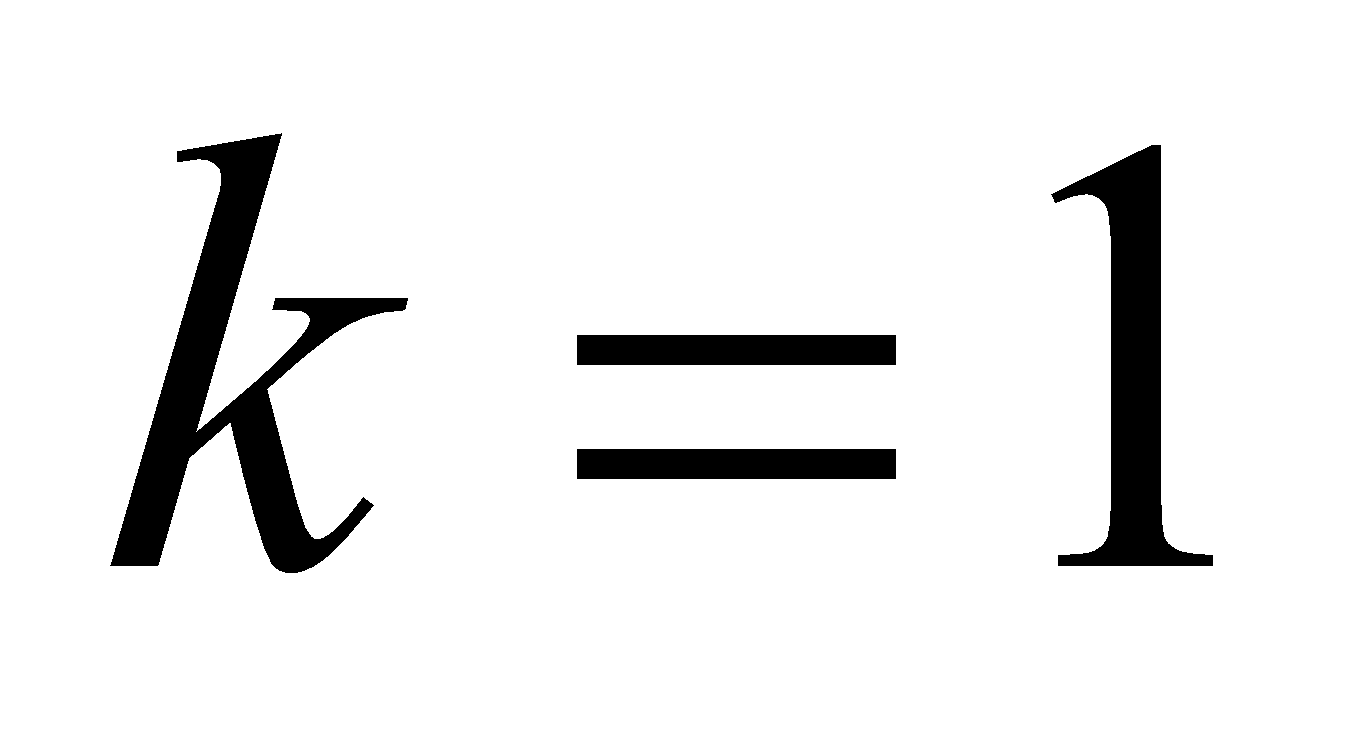
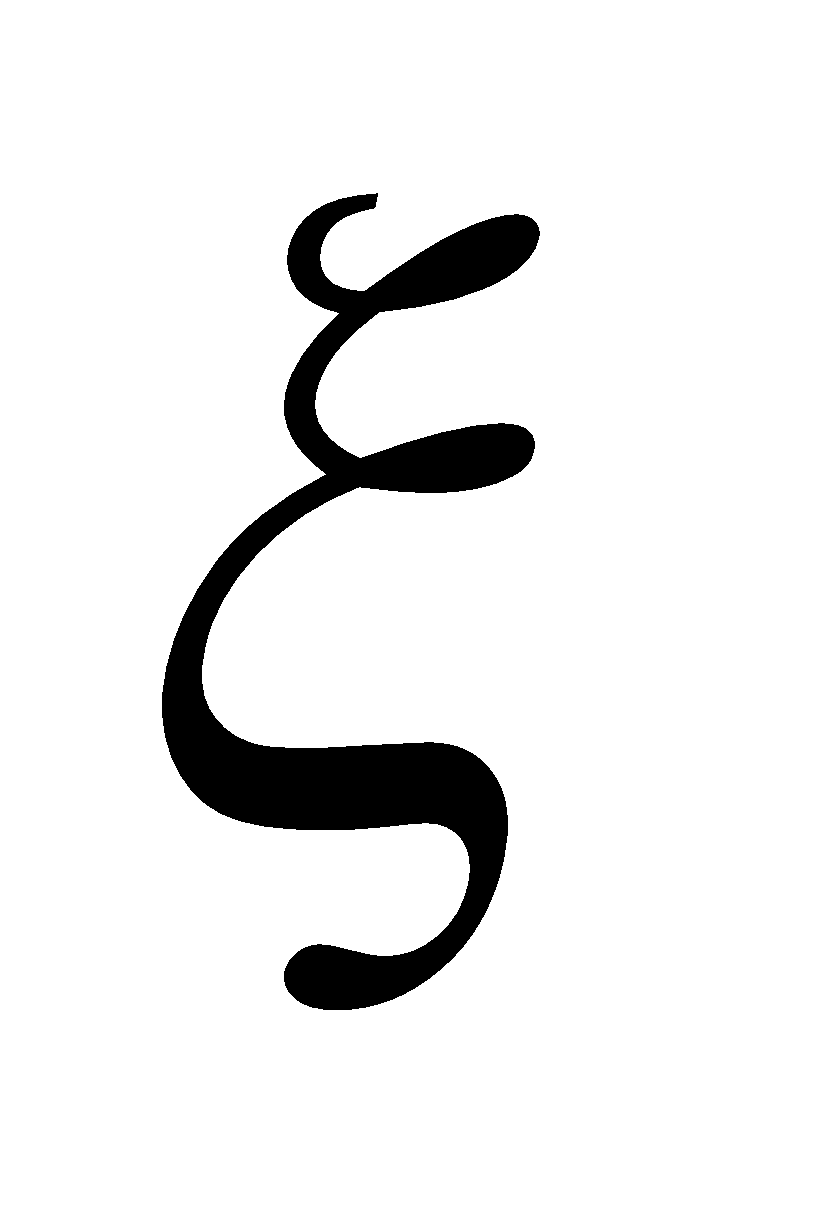
Критерій перевірки гіпотези  будується таким чином. Обчисливши значення статистики критерію ****і вибравши рівень значущості , по таблиці значень - розподілу (таблиця 3 додатку) визначимо величину  таку, що . Якщо , то гіпотеза  відхиляється, якщо ж , то гіпотеза приймається.

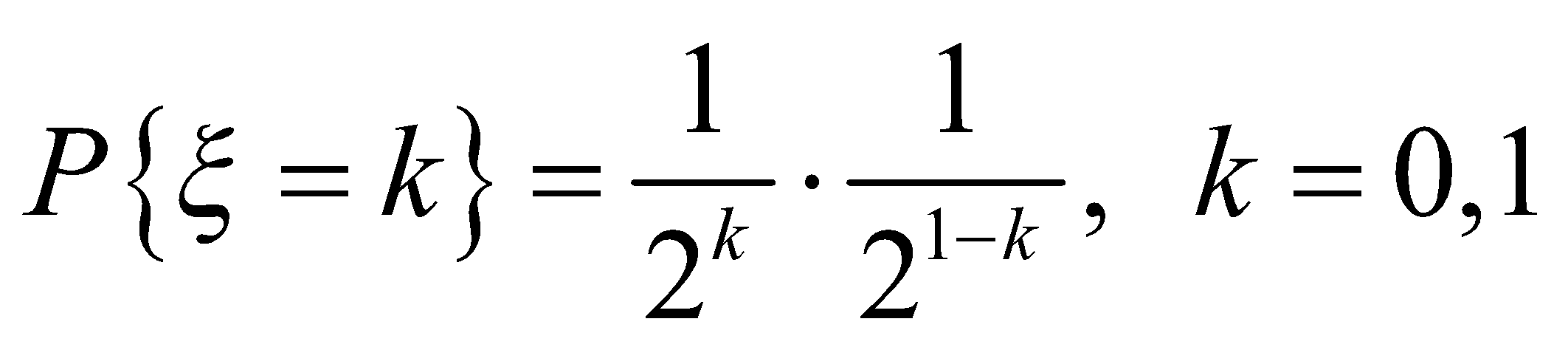
**Зауваження 1.** Гіпотетичний розподіл, що не залежить від параметрів, на практиці зустрічається рідко. Як правило гіпотетичний розподіл залежить від невідомих параметрів, відносно значень яких ми маємо лише ту інформацію, яка міститься у виборці. В цій ситуації ми маємо задачу перевірки складних гіпотез, для яких також можна використовувати критерій . Нехай по вибірці  треба перевірити гіпотезу : , де  – задане сімейство функцій розподілу. Значення параметрів, а отже і ймовірностей  невідомі. Тому природньо оцінити невідомий параметр  за вибіркою і в статистику (4.2) підставити ймовірності , підраховані через , де  – оцінка  методом максимальної вірогідності. В даному випадку величини  вже не сталі, а представляють собою функції від вибірки, а отже є випадковими величинами. Тому, взагалі кажучи, теорема 4.1 не може бути застосована. Окрім того, розподіл цієї статистики буде залежати від методу побудови оцінки . Р. Фішер показав, що існують методи оцінювання параметра , при яких граничним розподілом для статистики критерію буде - розподіл з  ступенями свободи, де  – розмірність параметра . Одним з таких методів є метод мінімума .

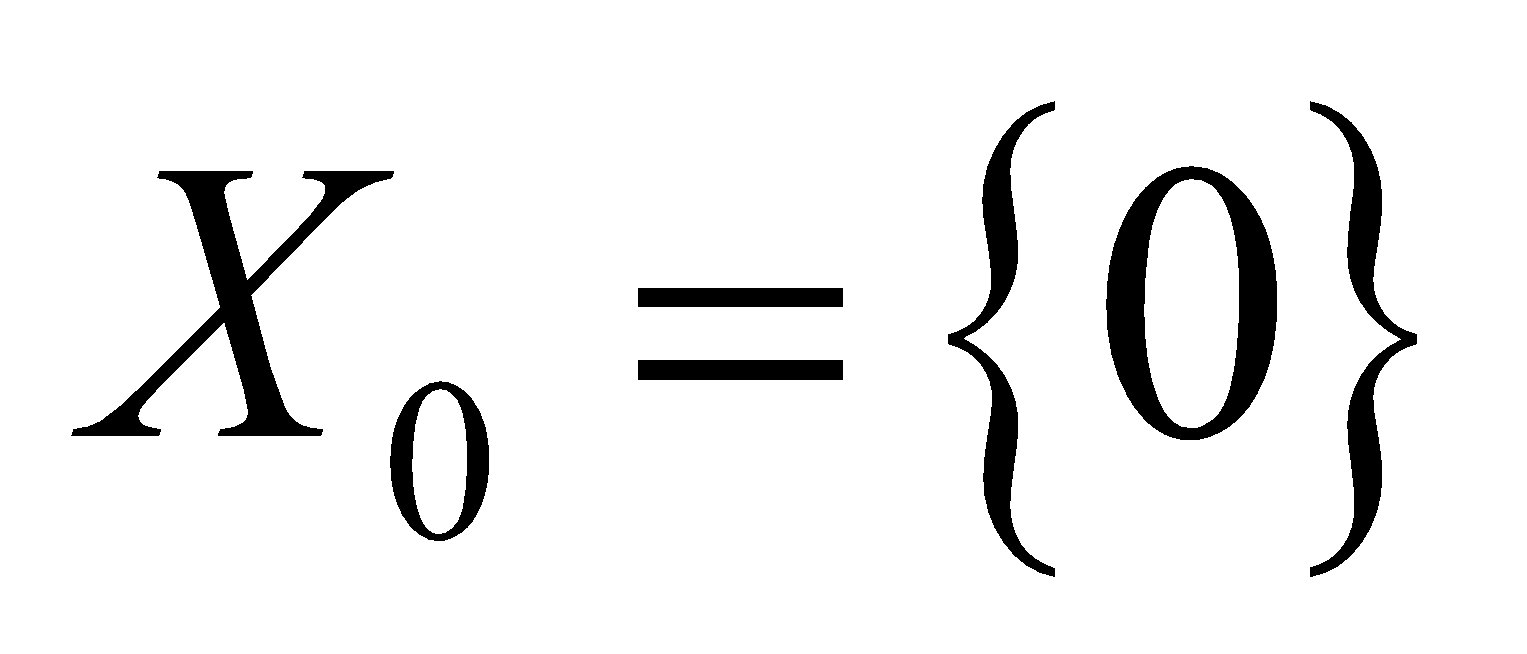
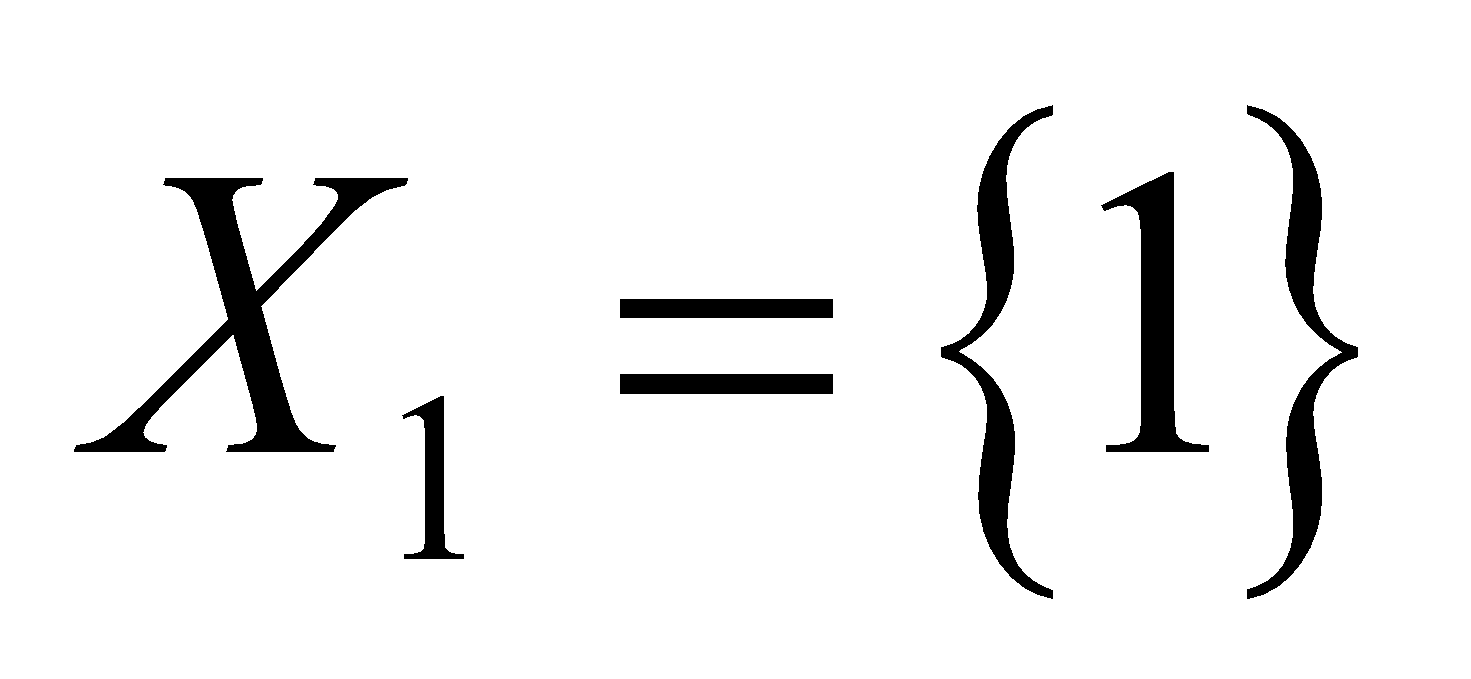
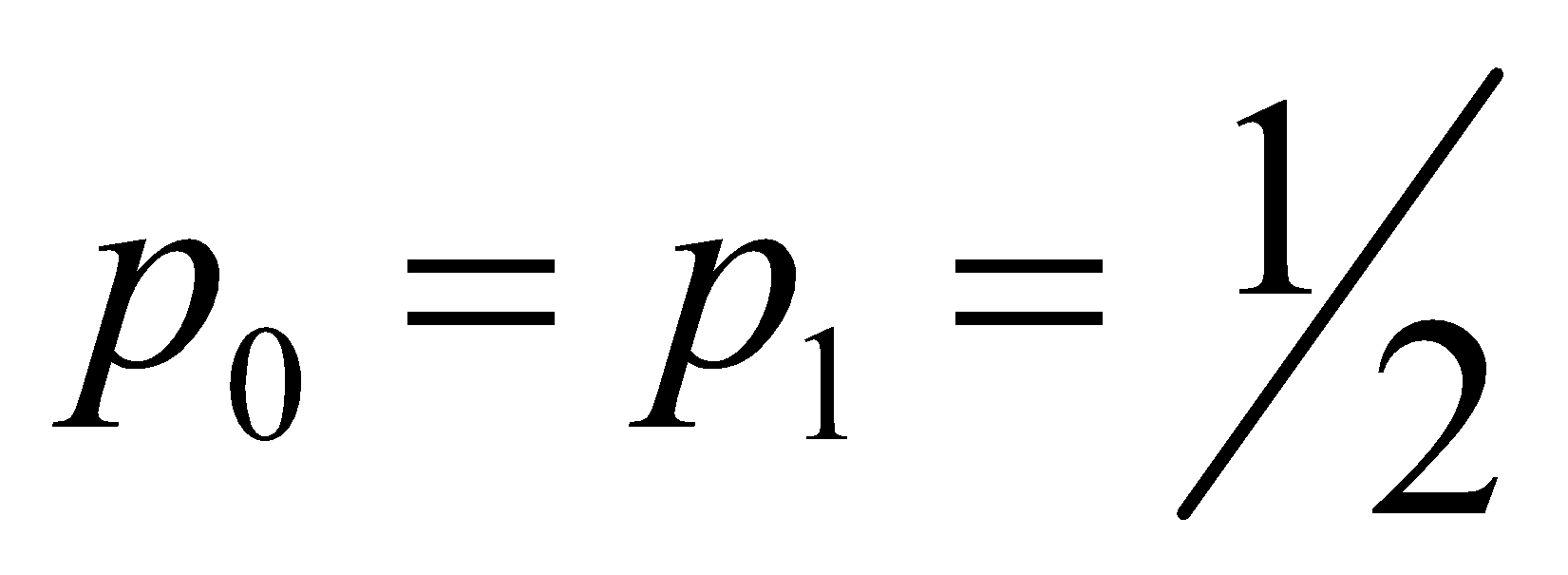
Далі критерій перевірки гіпотези будується аналогічно наведеному вище. Якщо , то гіпотеза  відхиляється, якщо ж , то гіпотеза приймається.

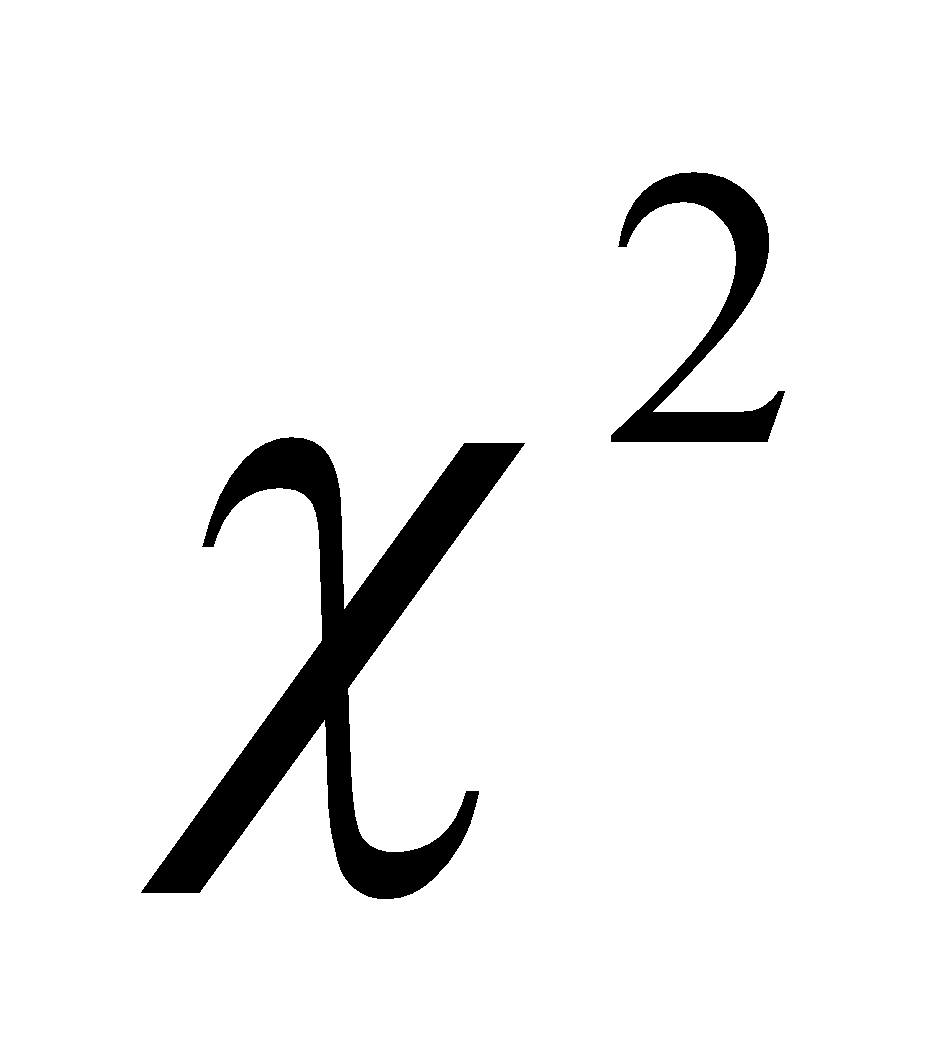
**Зауваження 2.** За допомогою методу мінімума  знайдено оцінки для параметру  розподілу Пуассона та параметрів  та  нормального розподілу. Доведено, що оцінки параметрів  та  близькі до вибіркового середнього, а оцінкою  є вибіркова дисперсія 

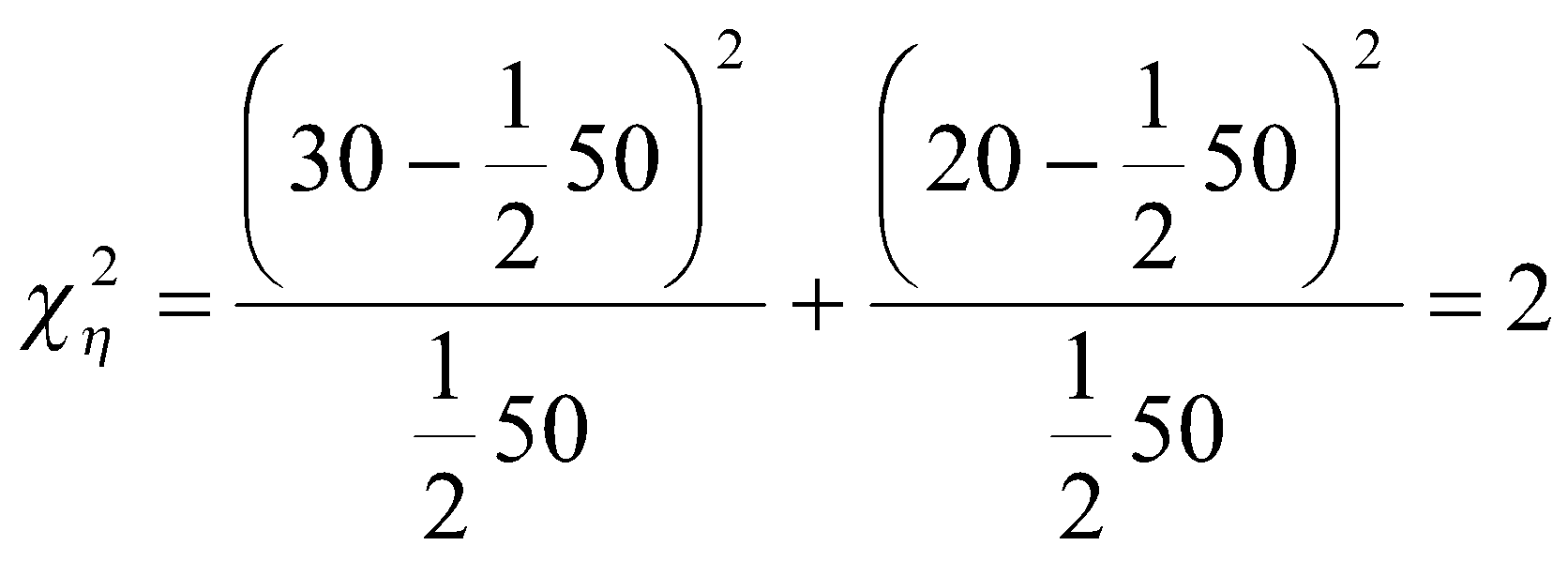
**Приклад 4.3.** При 50 підкиданнях монети герб з’явився 20 раз. Чи можна вважати, що монета симетрична? Прийняти .

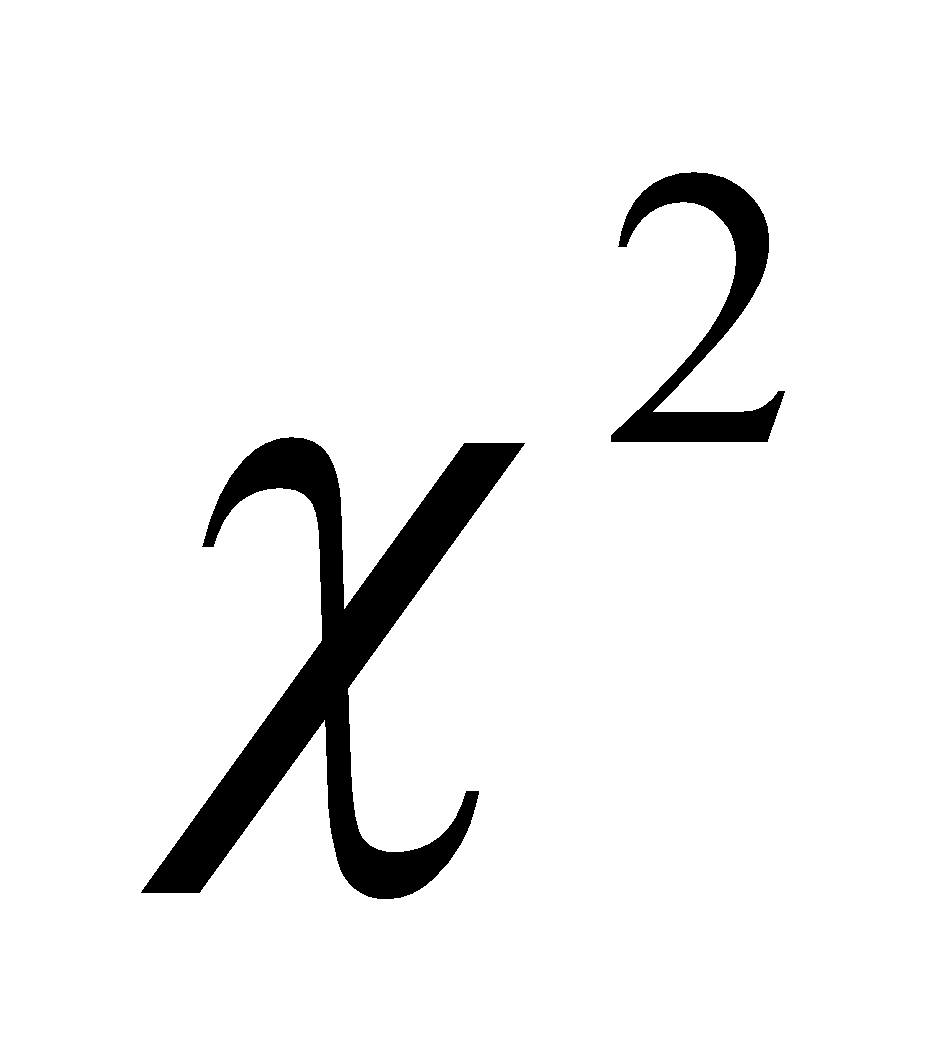
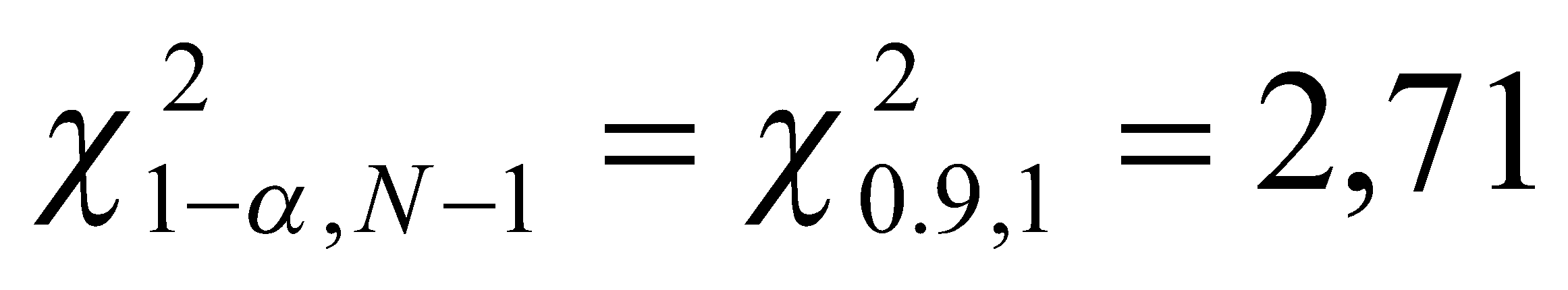
Експеримент з підкиданням монети можна описати в термінах незалежних спостережень випадкової величини , яка приймає два значення: , якщо випав герб, та 0, якщо випала решка. Гіпотеза про симетричність монети в термінах розподілу випадкової величини формулюється так: розподілом випадкової величини  є

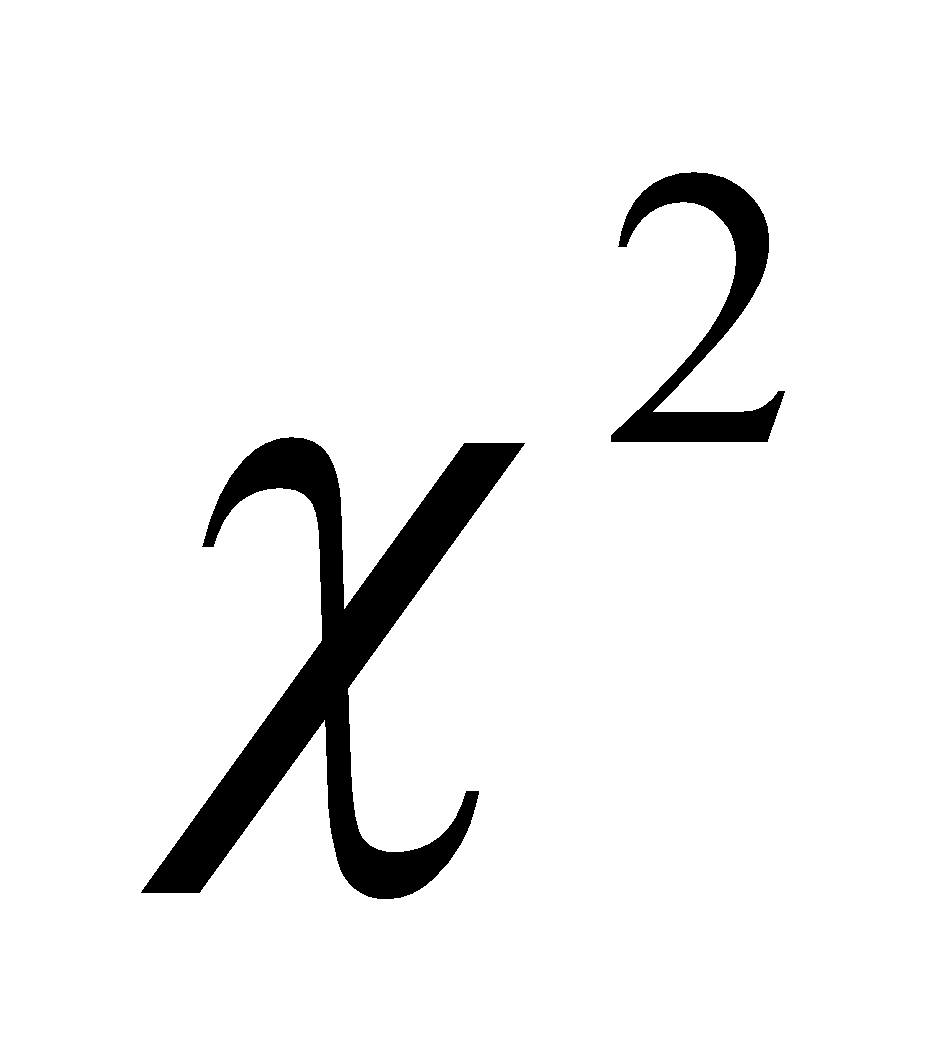
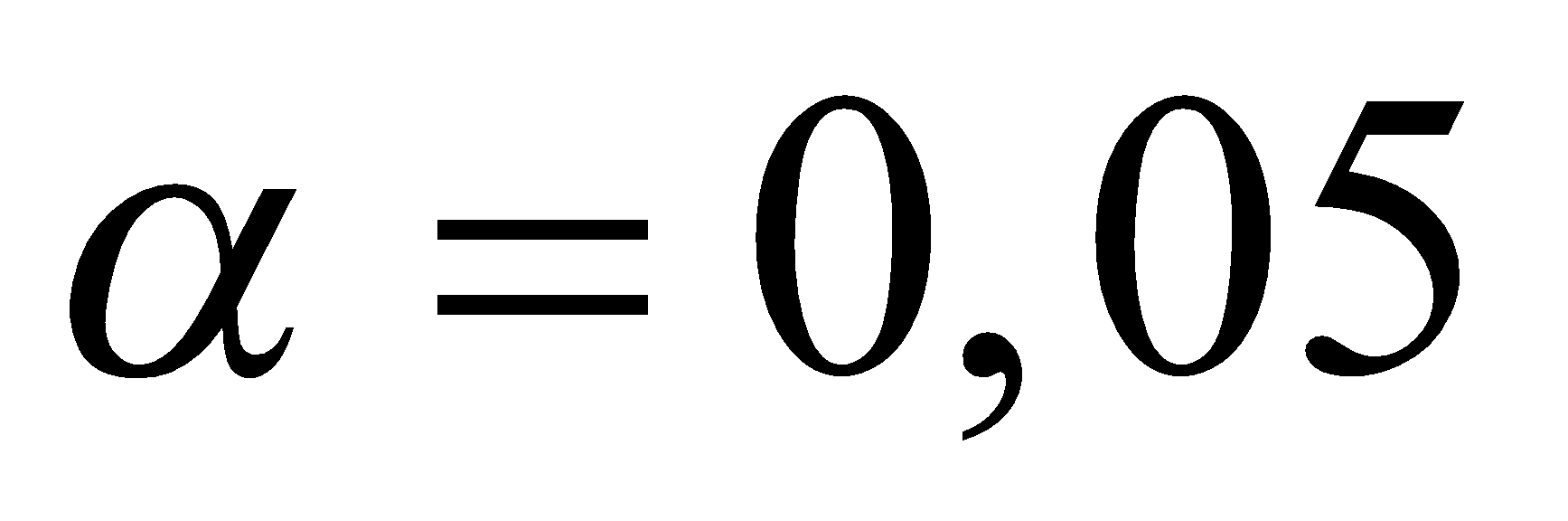
.

Ймовірності попадання вибіркових значень в підмножини вибіркових значень, якими є дві підмножини  та , підраховані за гіпотетичним розподілом, дорівнюють .

Підрахуємо значення статистики критерію :

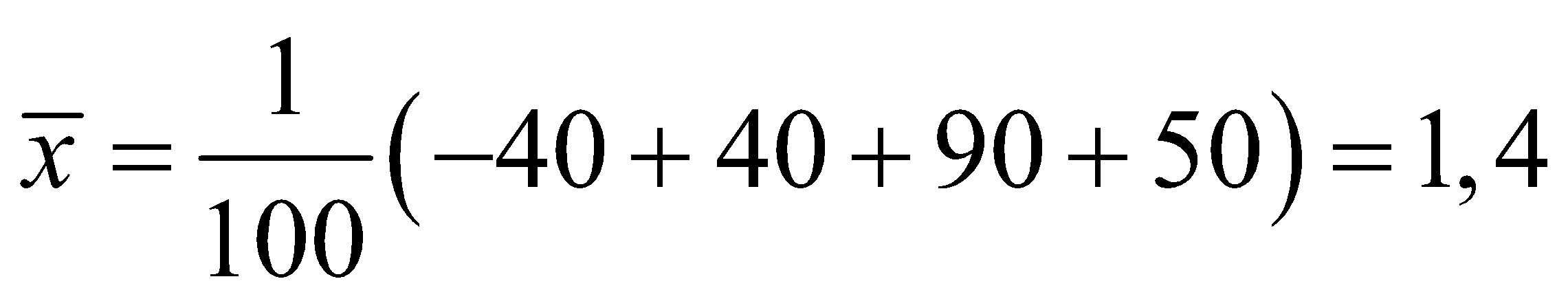
.

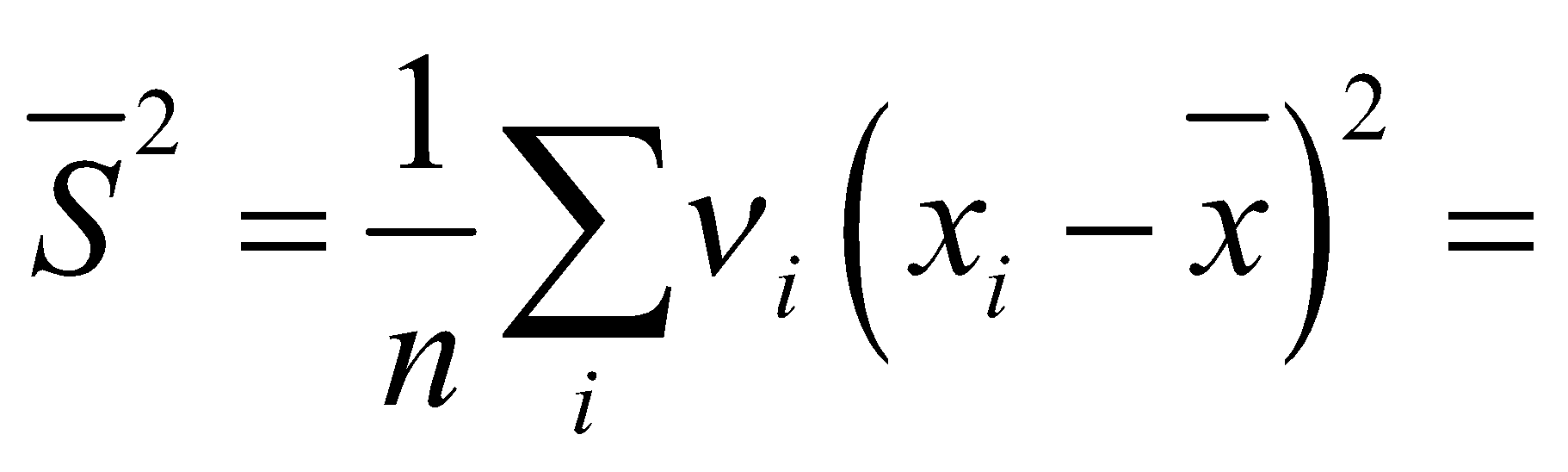
В таблиці 3 значень -розподілу шукаємо значення . Оскільки значення статистики критерію не перевищує табличне значення, гіпотеза не відхиляється. Таким чином гіпотеза узгоджується з даними експерименту, або, іншими словами, гіпотеза про симетричність монети не протирічить експериментальним даним.

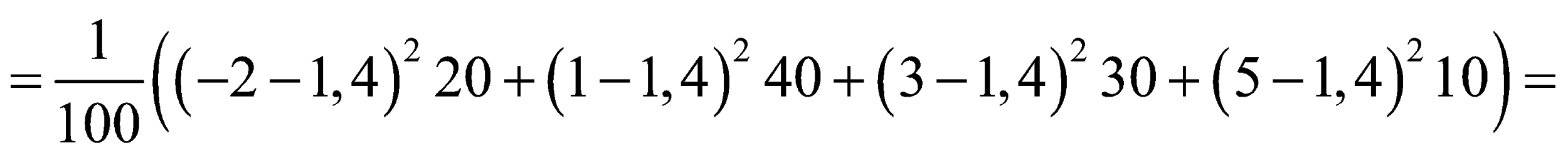
**Приклад 4.4.** По спостереженням, наведеним у таблиці, за допомогою критерію  з рівнем значущості  перевірити гіпотезу, що випадкова величина має нормальний розподіл.

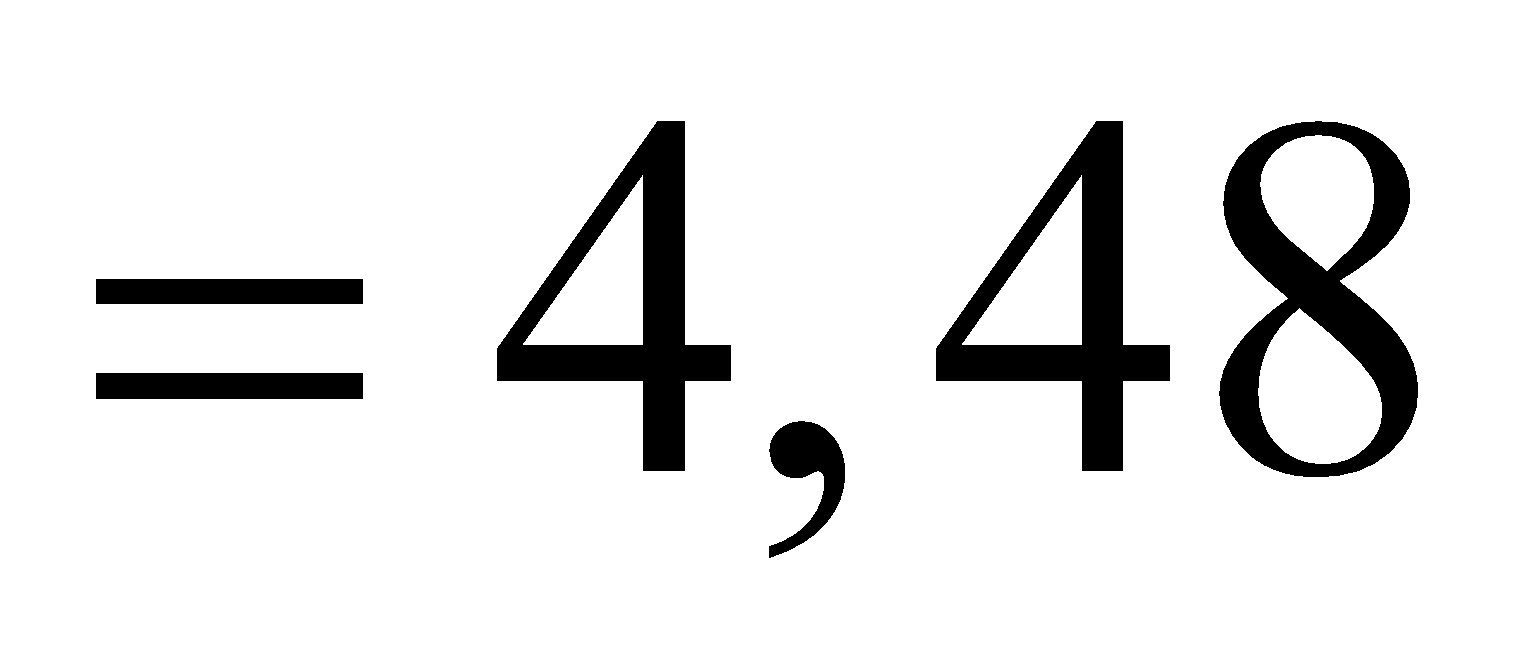
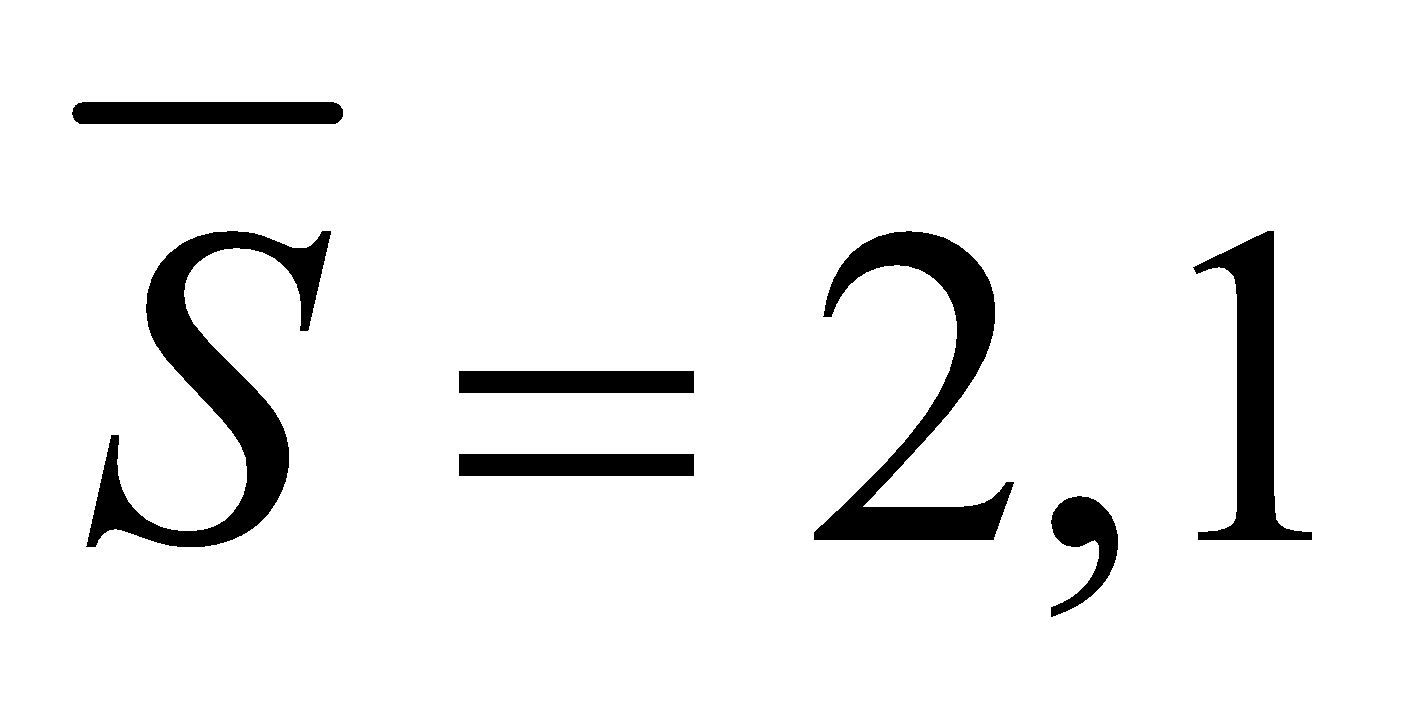
|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Інтервал |  |  |  |  |
|  | 20 | 40 | 30 | 10 |

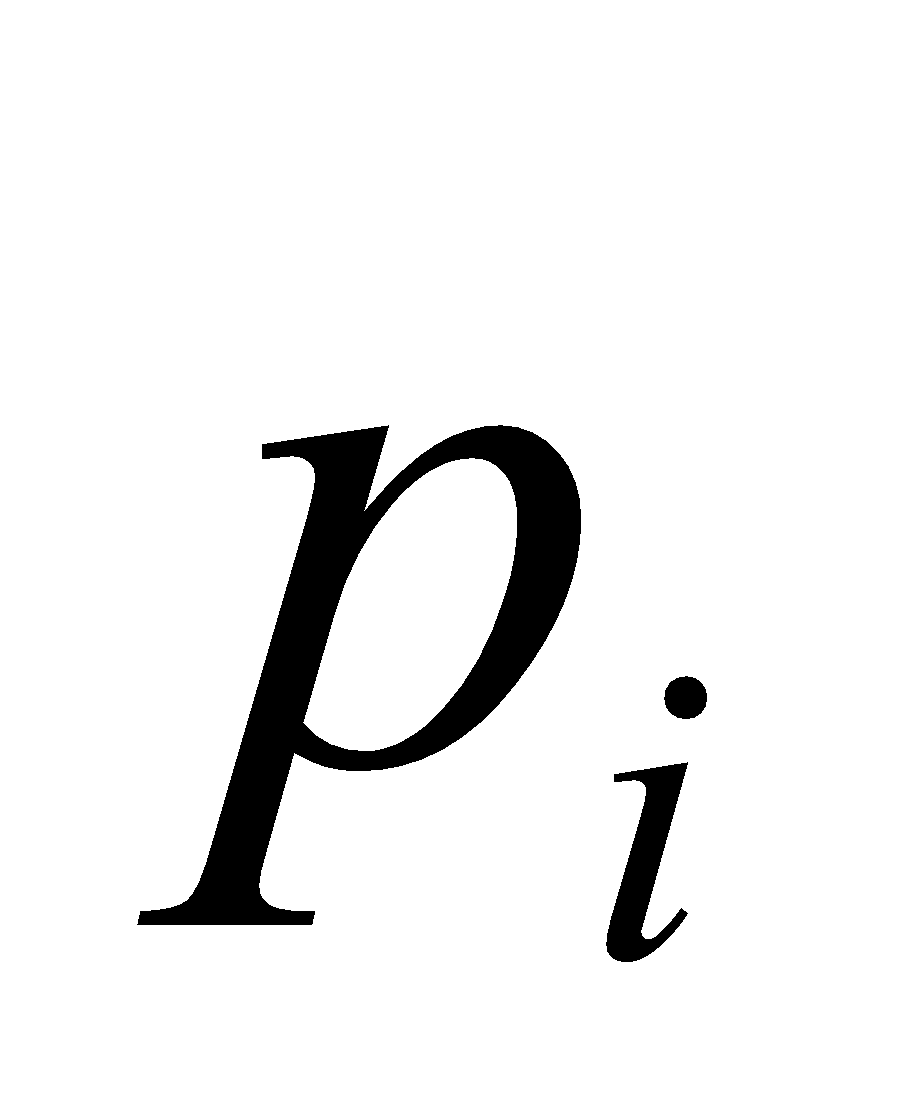
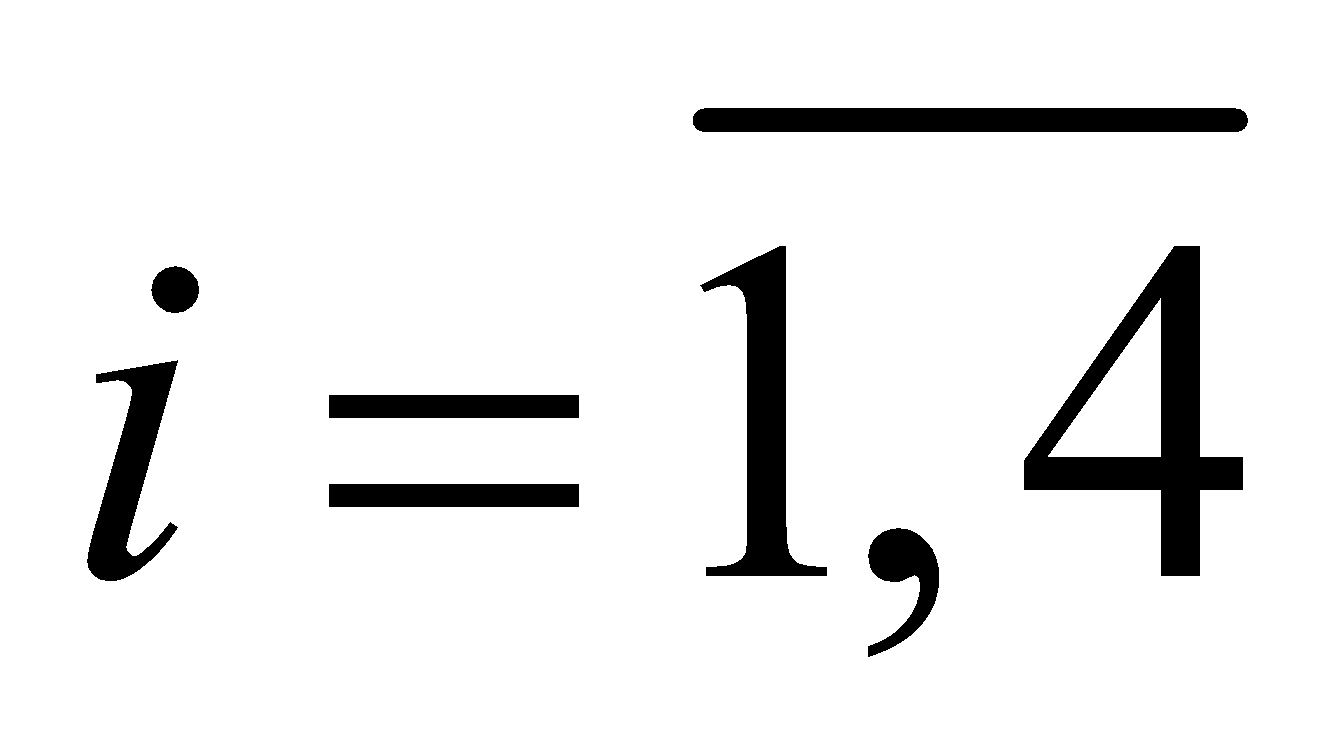
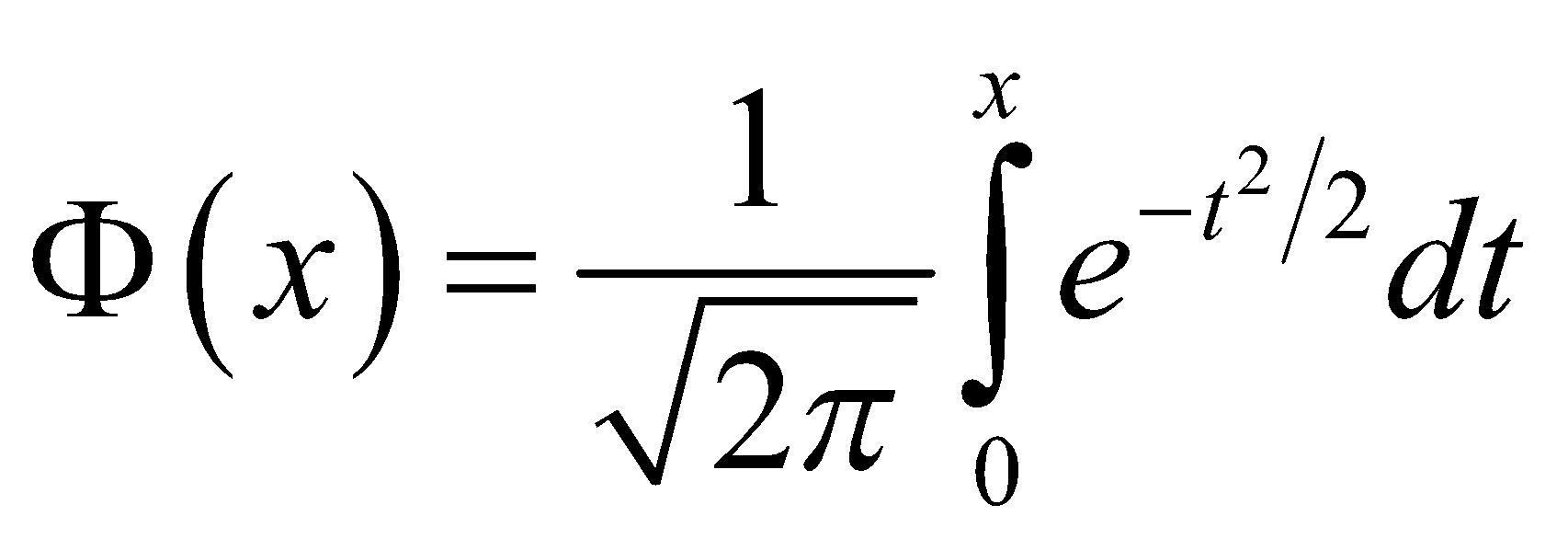
Обчислимо оцінки параметрів нормального розподілу за вибіркою. Згідно зауваженню 2:

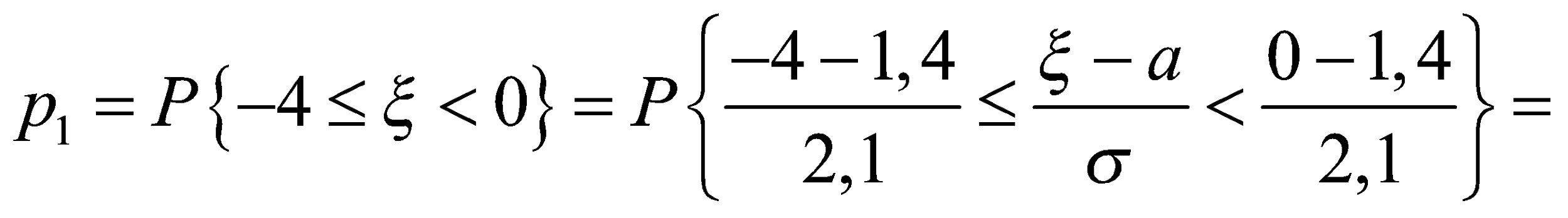
;

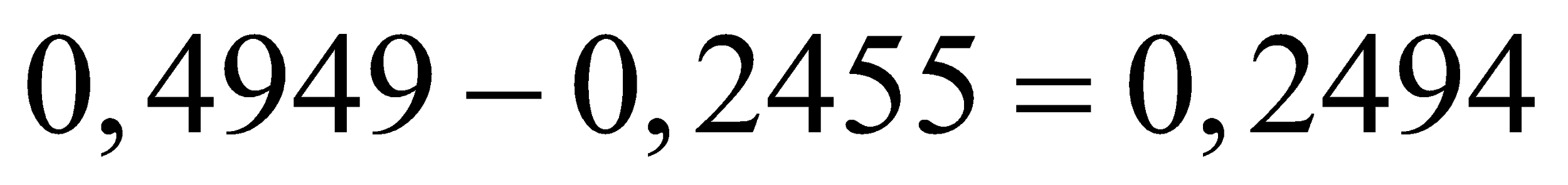
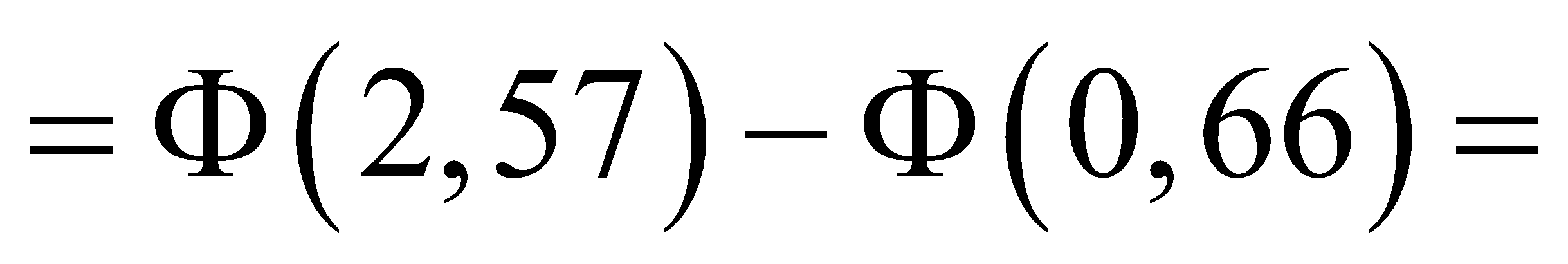


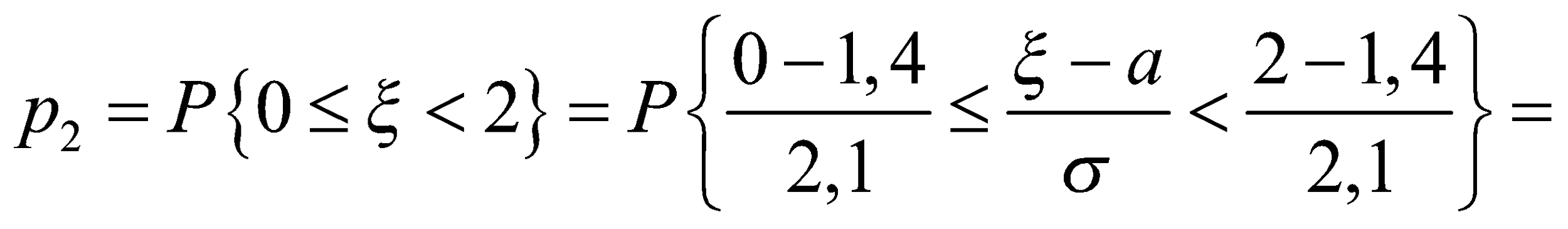


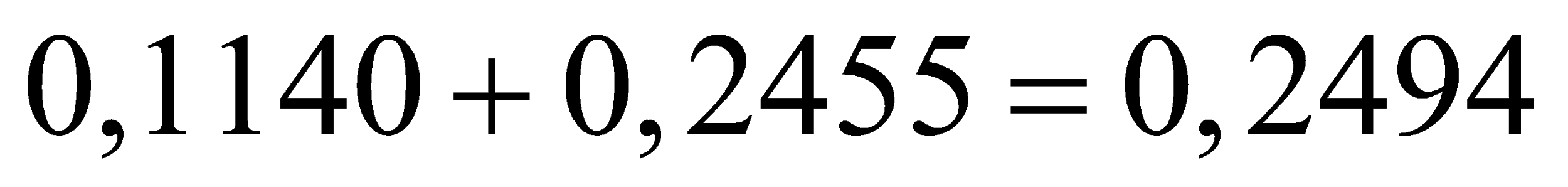
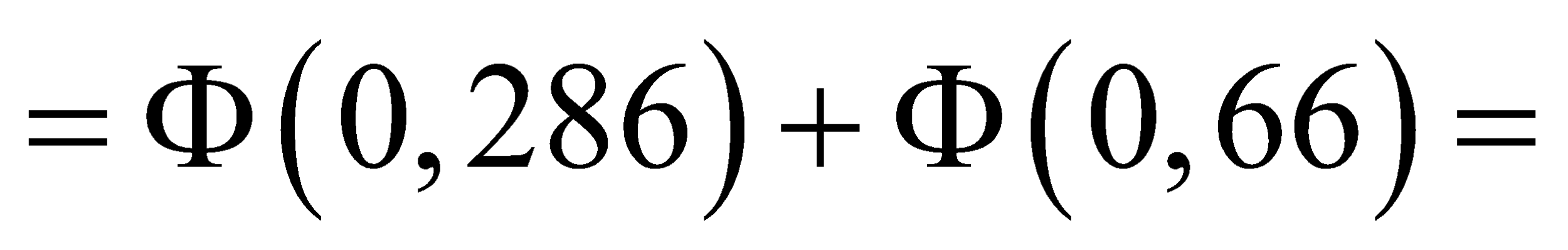
; .

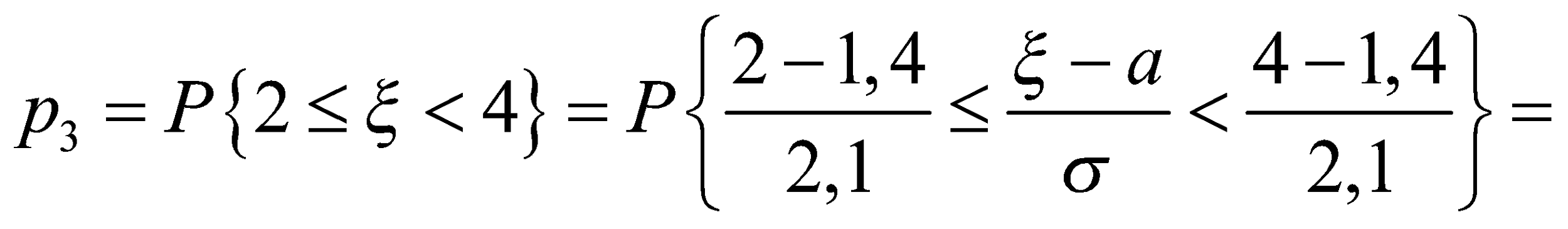
Тепер підрахуємо ймовірності , , використовуючи функцію Лапласа  та таблицю 1 додатка з її значеннями.

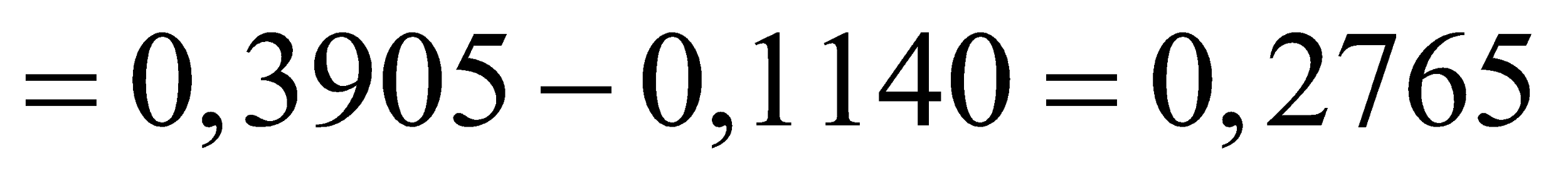
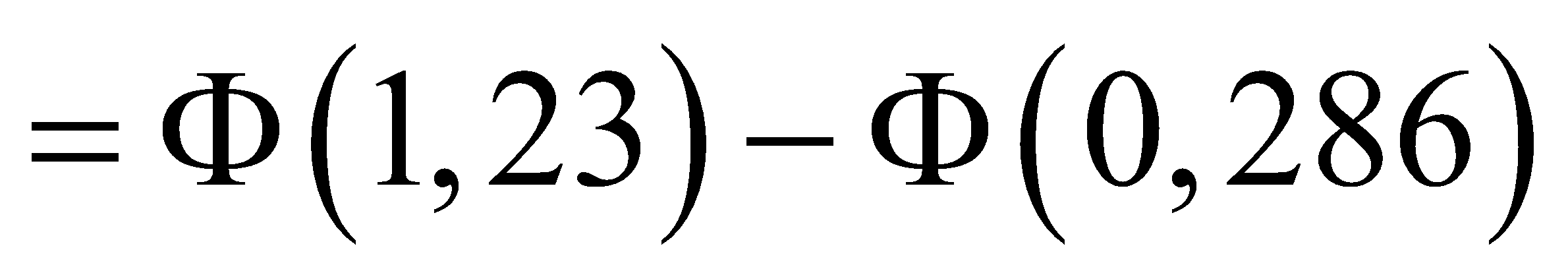


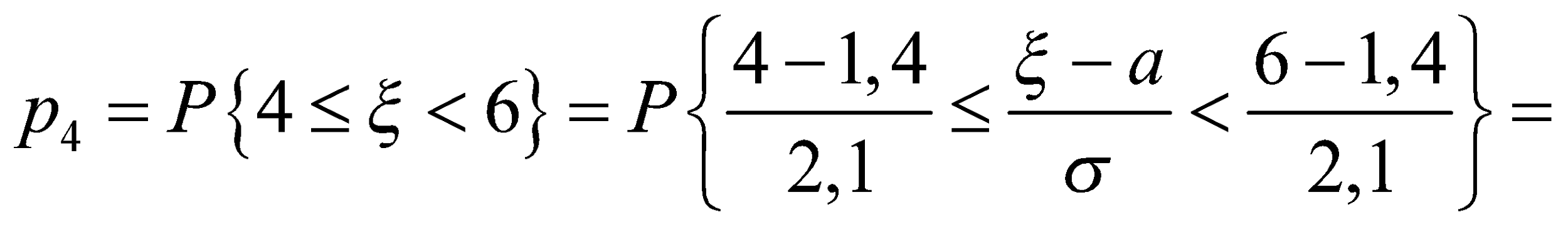
;

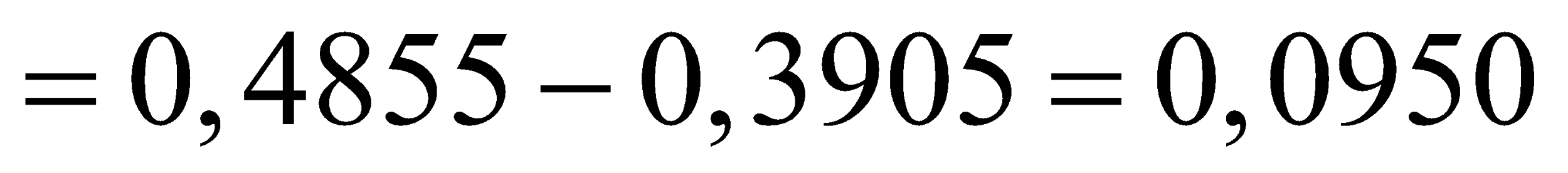
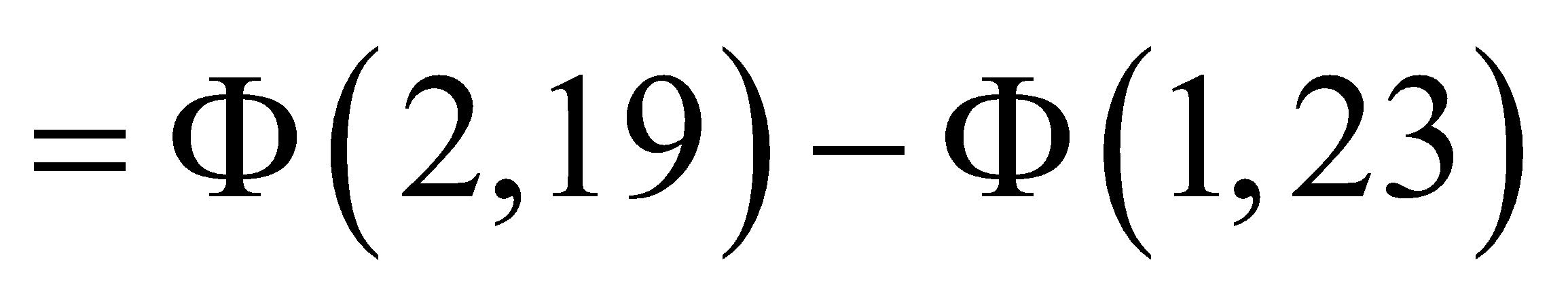


;



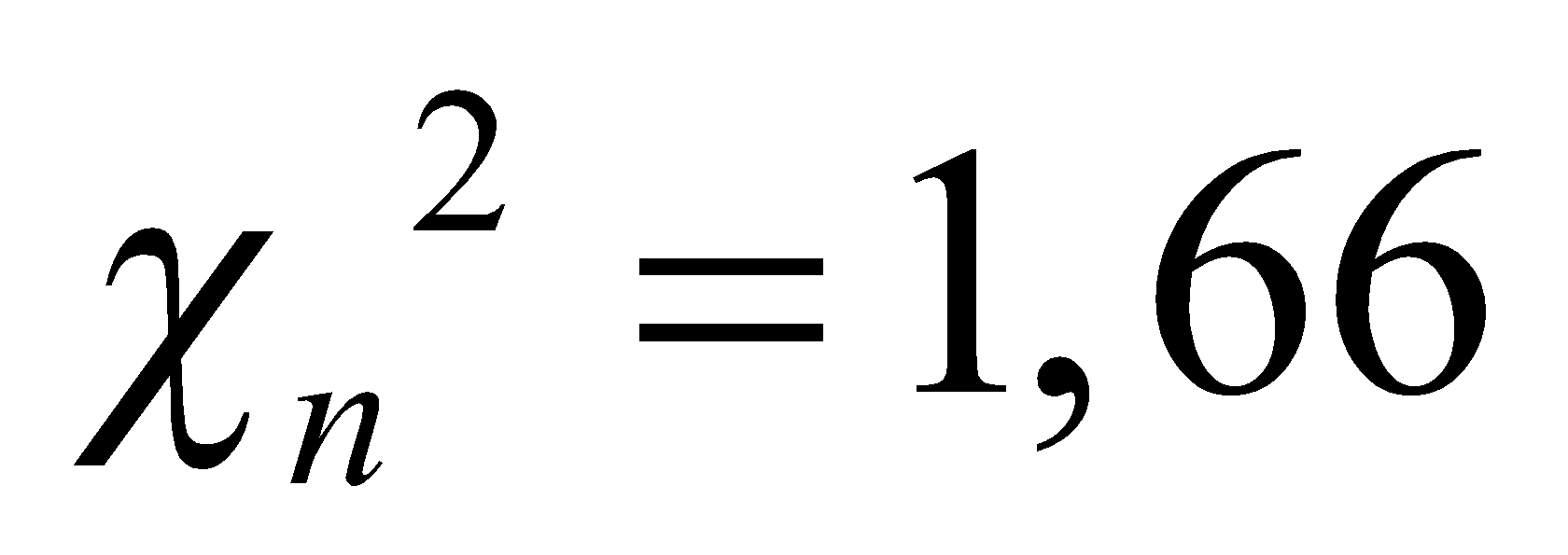
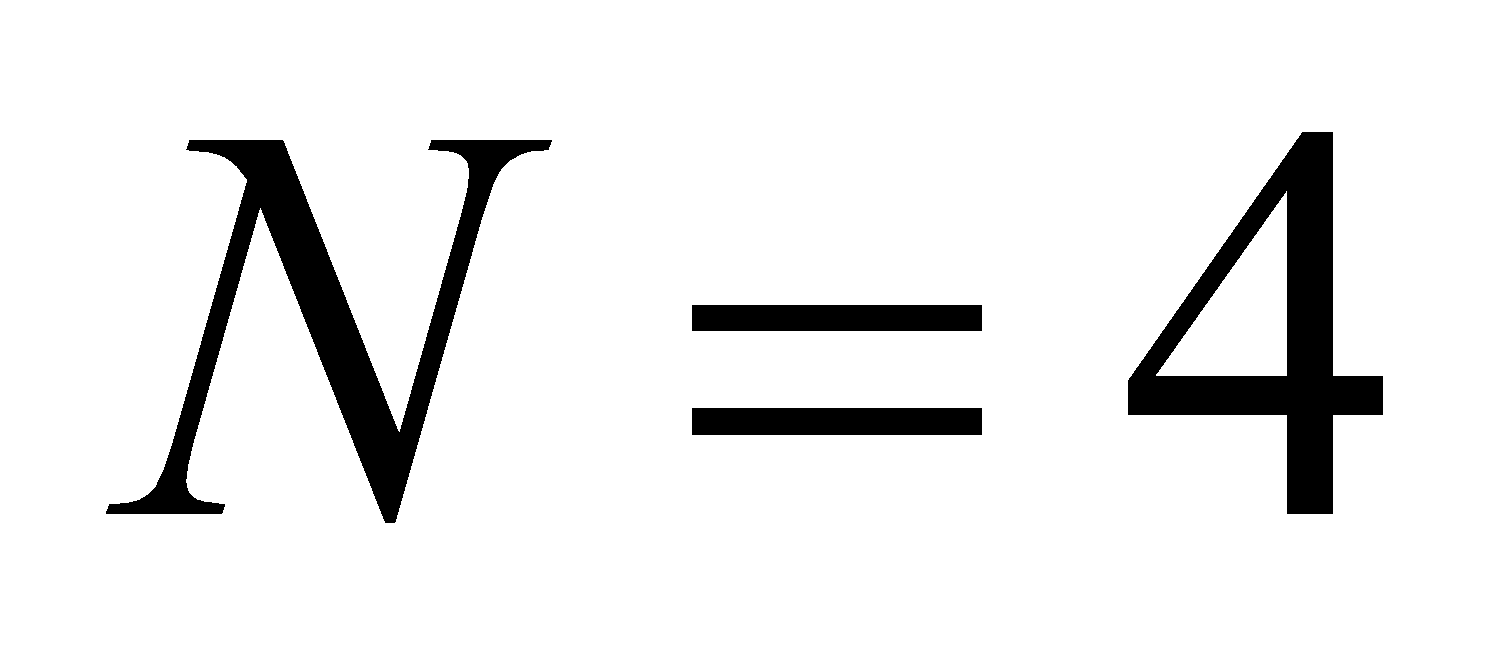
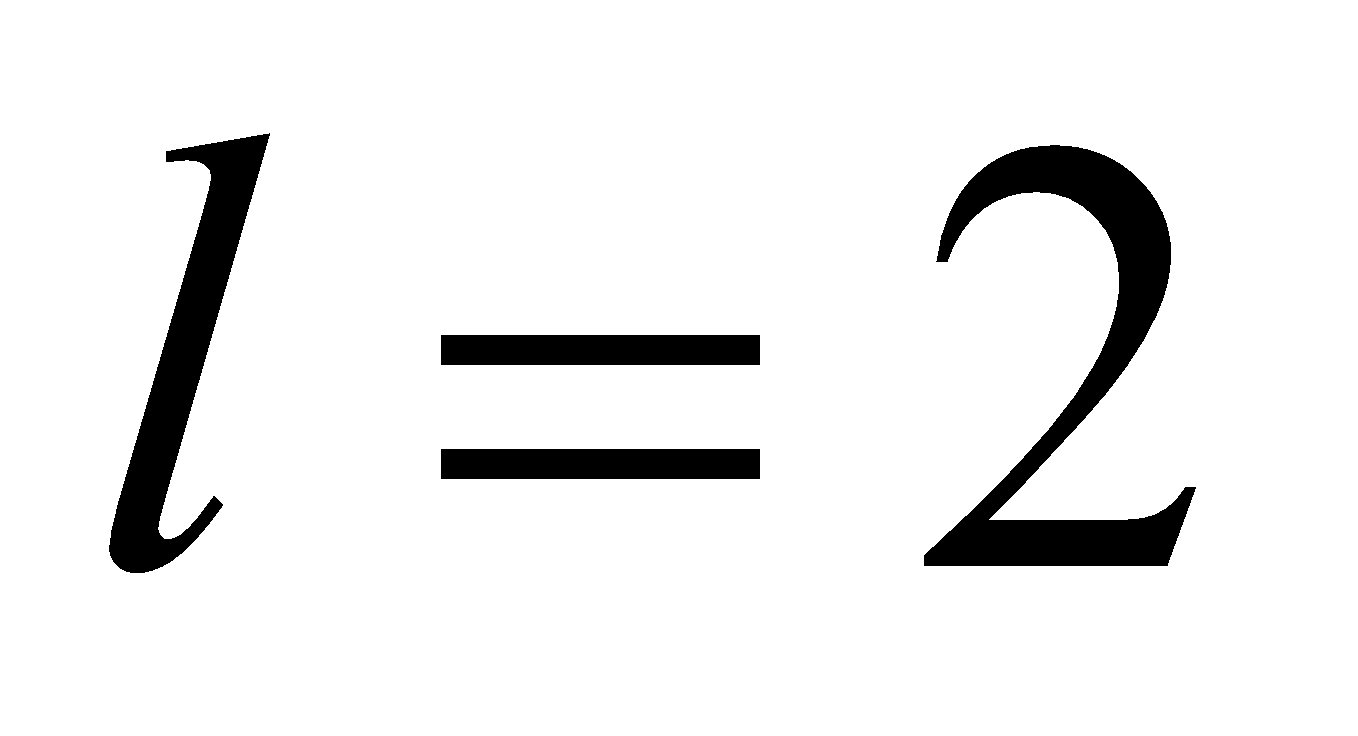
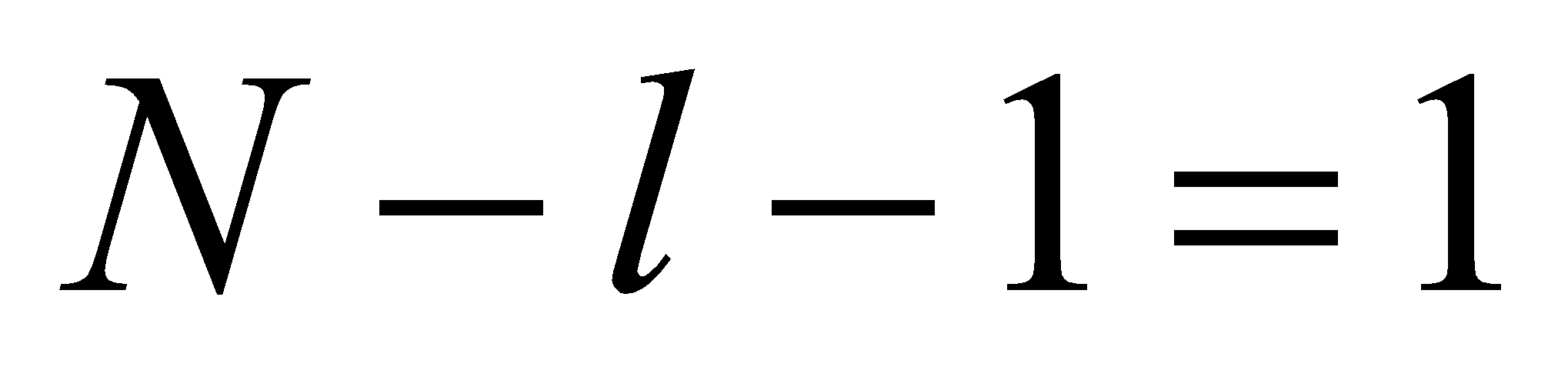
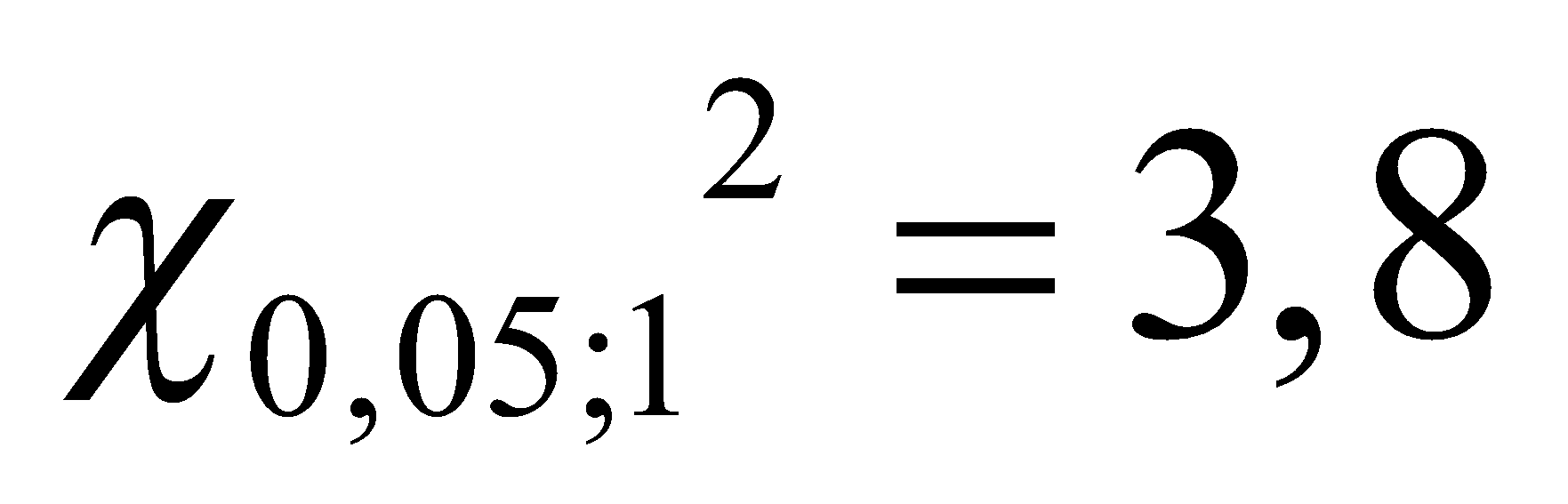
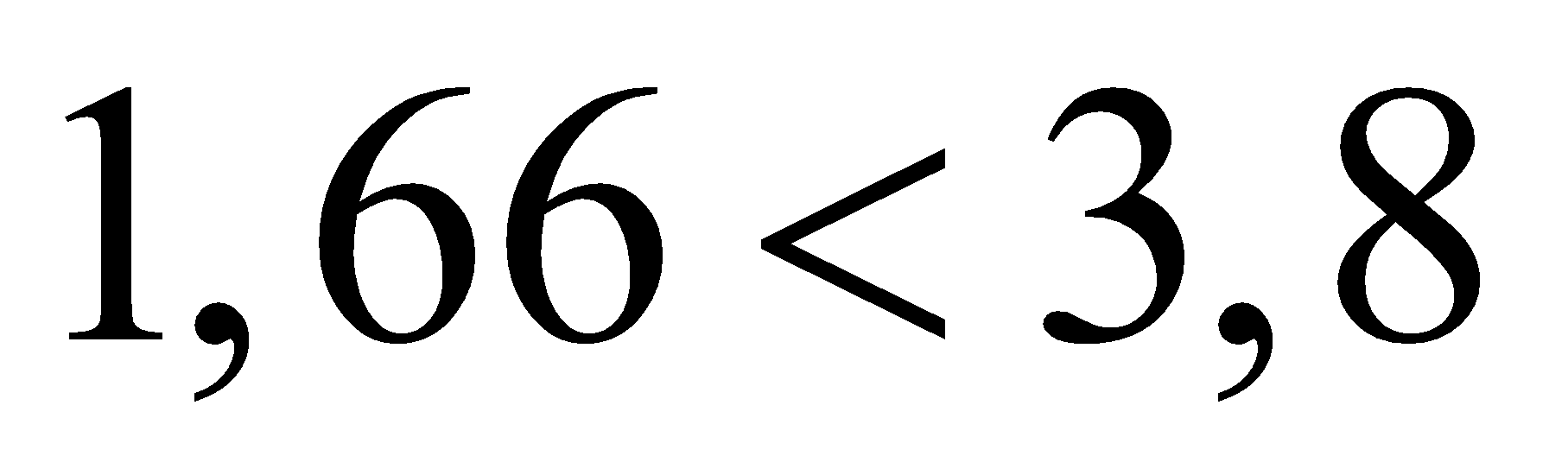
;

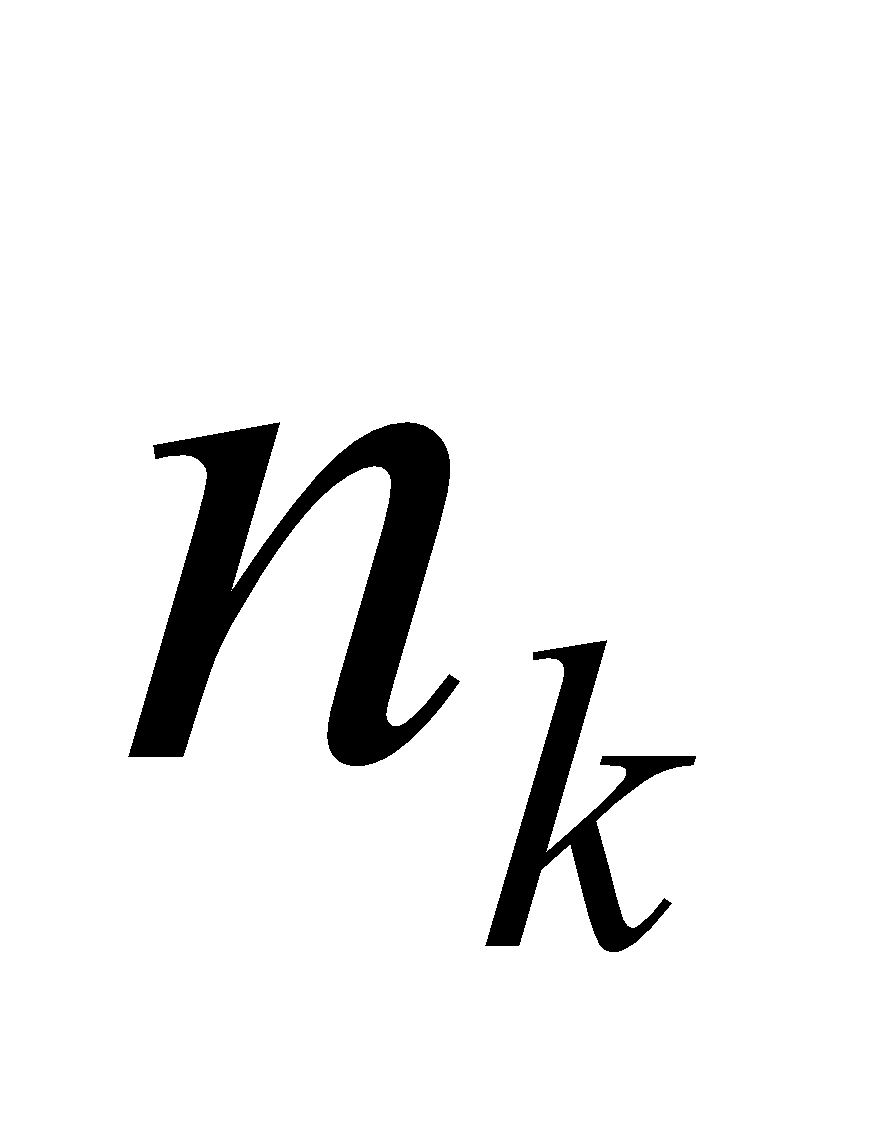
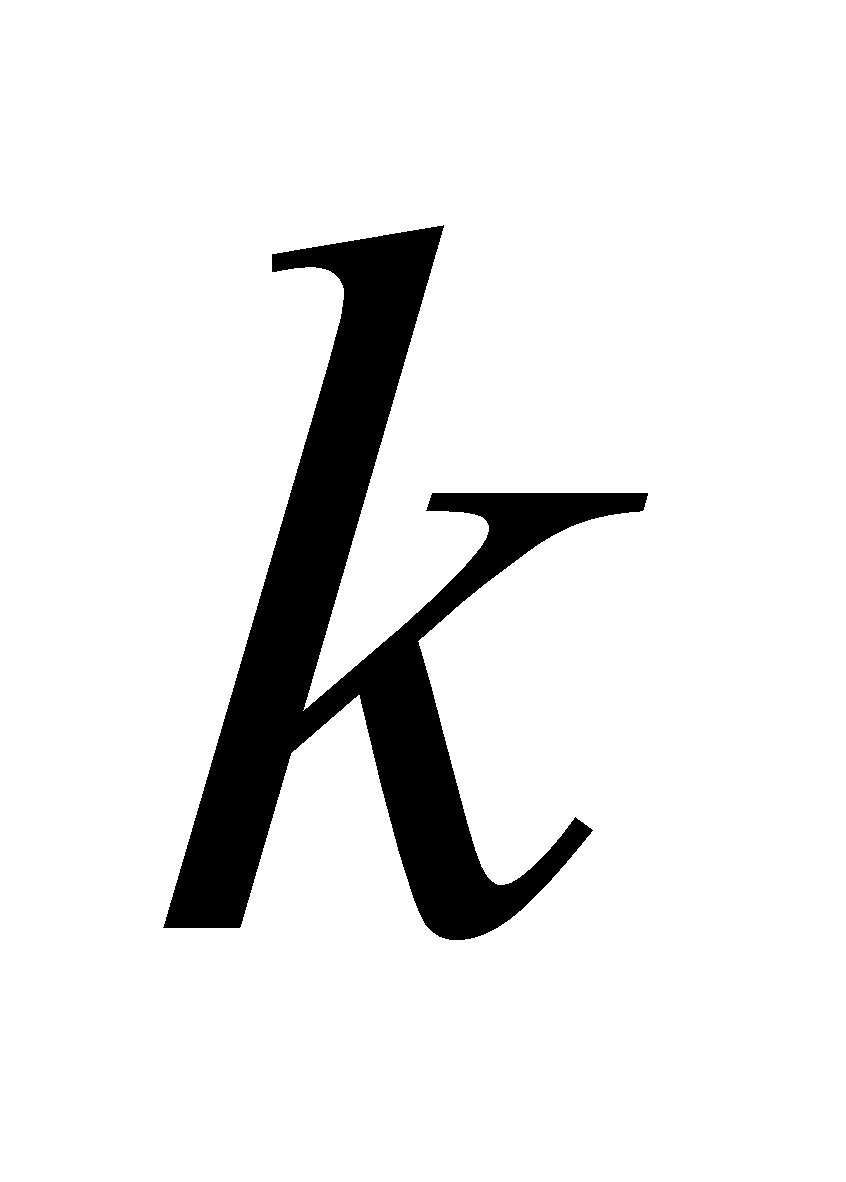


.

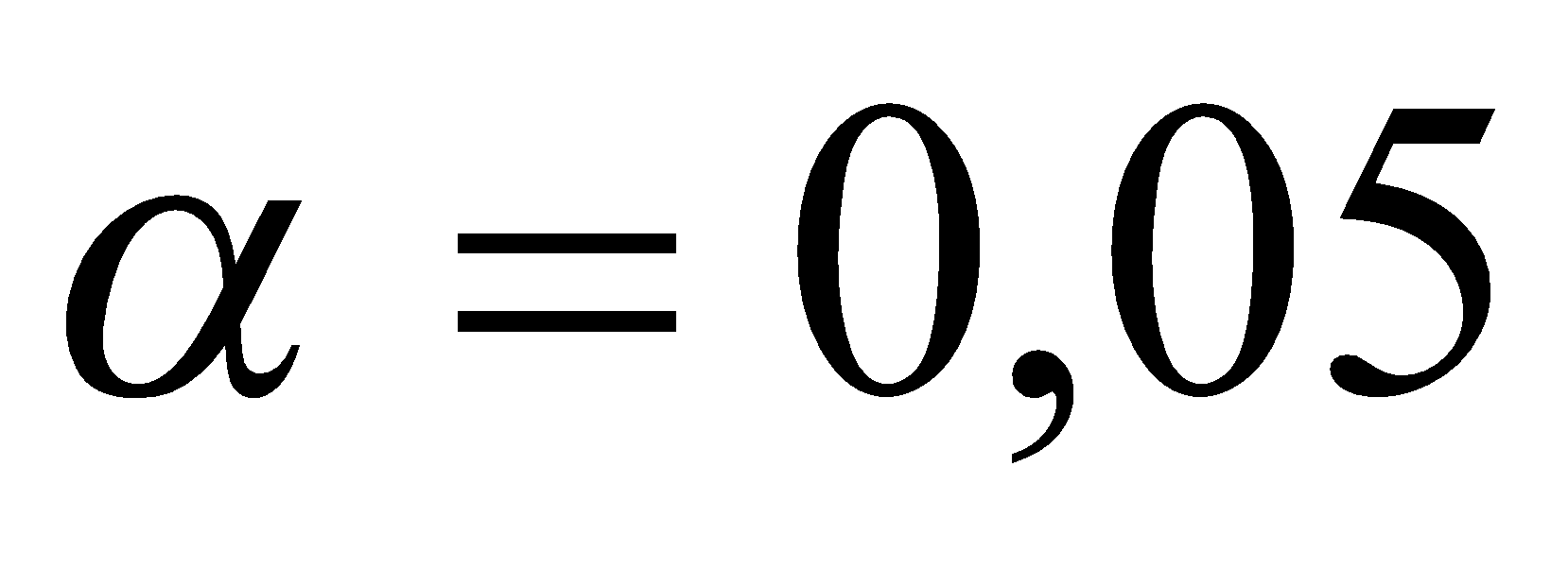
Обчислені результати заносимо в таблицю:

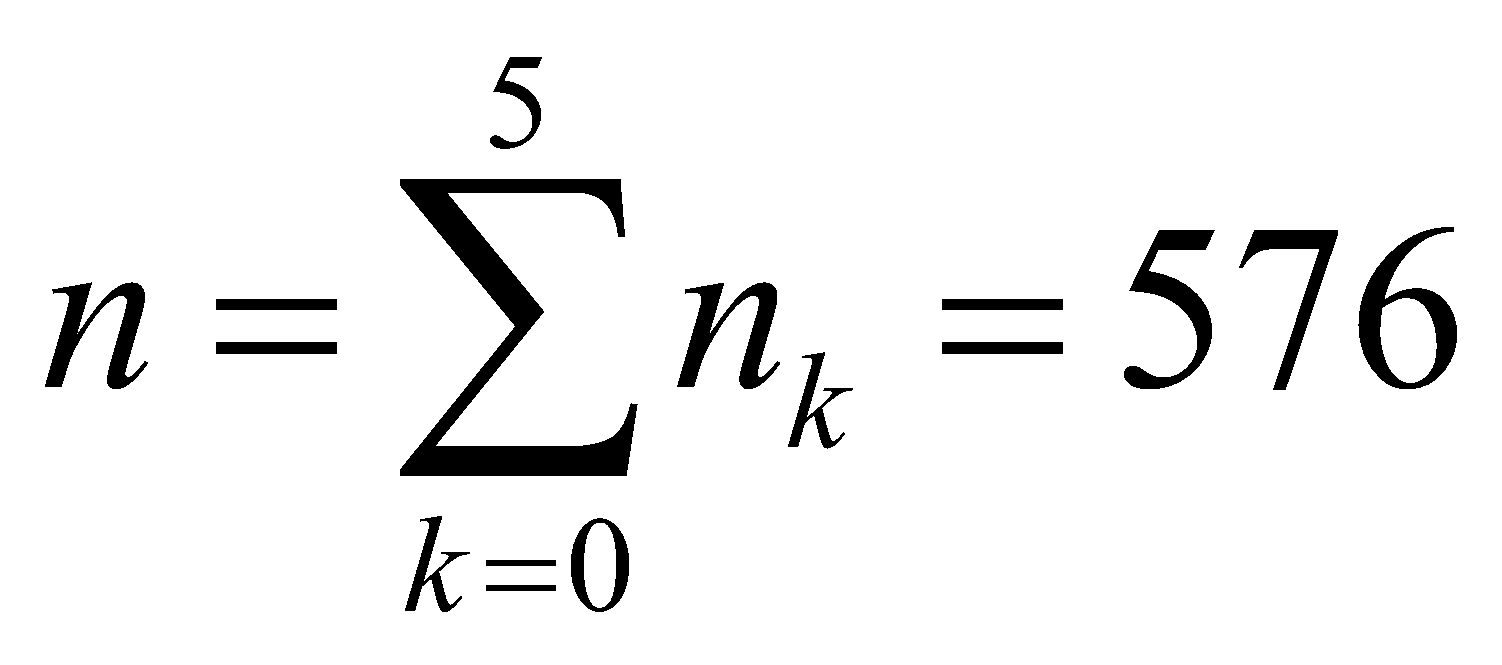
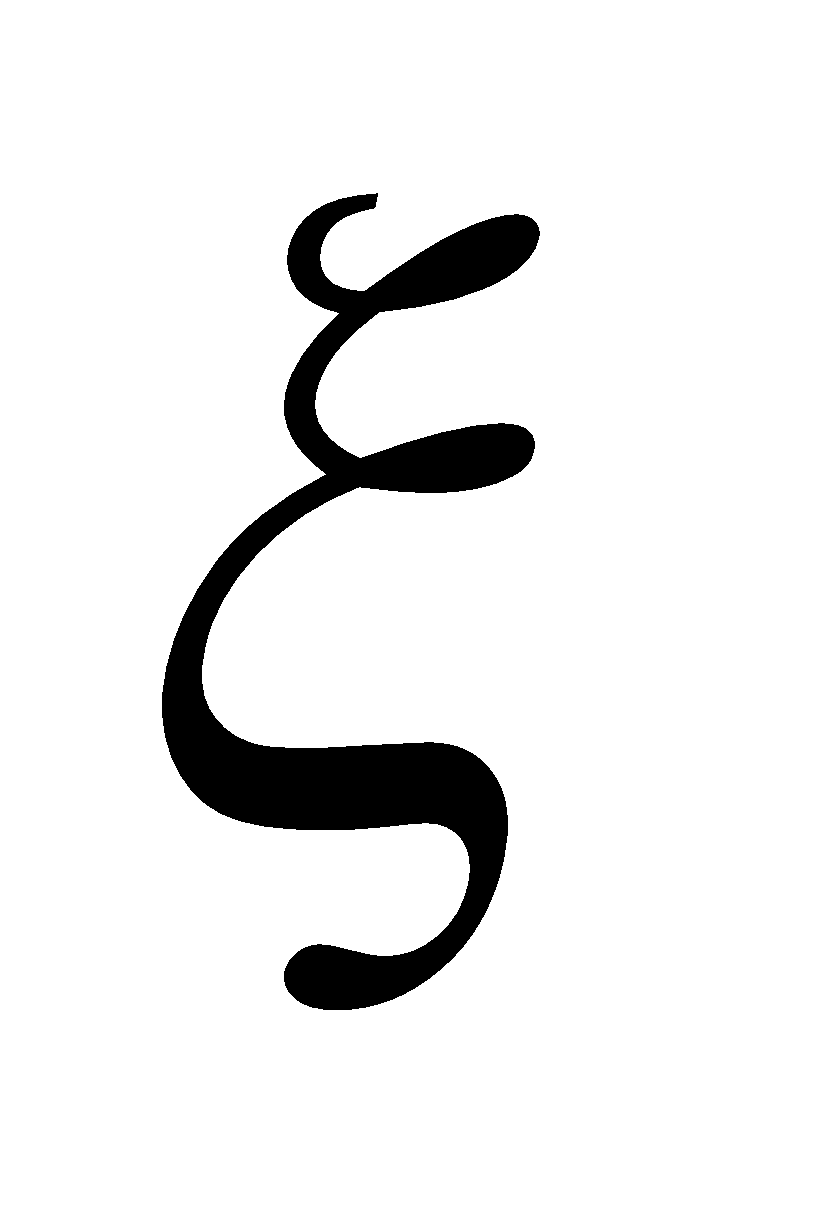
|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Інтервал |  |  |  |  |
|  | 20 | 40 | 30 | 10 |
|  | 0,2494 | 0,3595 | 0,2765 | 0,095 |
|  | 24,94 | 35,95 | 27,65 | 9,5 |
|  | 24,40 | 16,40 | 5,52 | 0,25 |
|  | 0,98 | 0,46 | 0,2 | 0,02 |

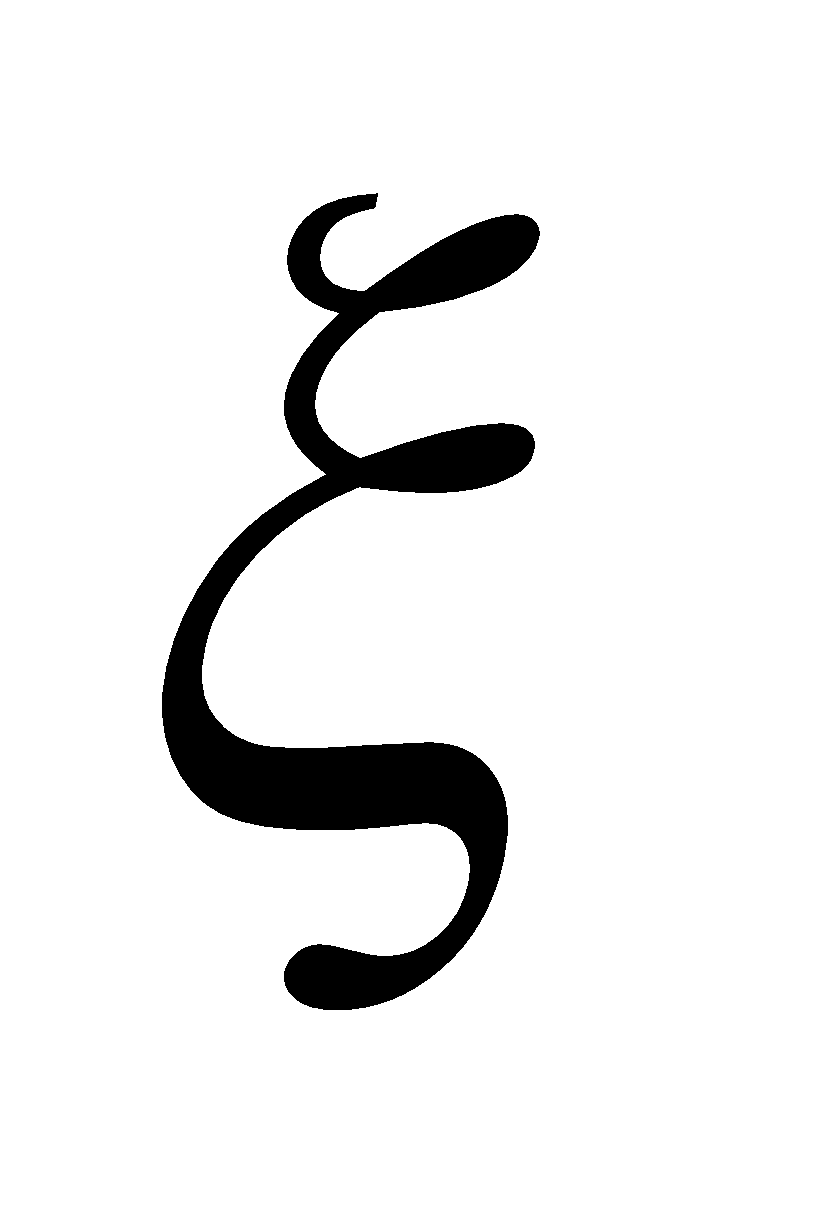
Значення статистики критерію . Кількість інтервалів , кількість невідомих параметрів . Число ступенів свободи . З таблиці беремо критичне значення . Оскільки , то гіпотеза приймається.

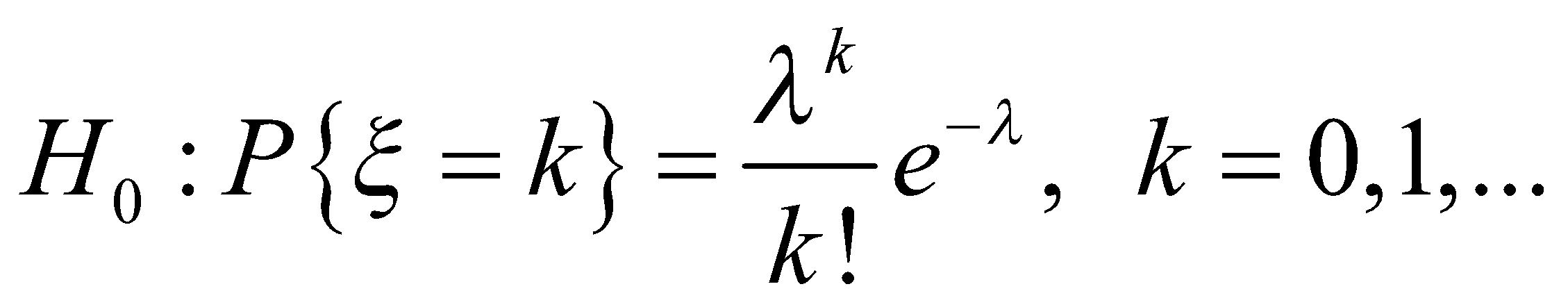
**Приклад 4.5.** За час другої світової війни на Лондон впало 535 літаків-снарядів. Уся територія Лондону була розділена на 576 ділянок площею по 0,25 км². Нижче наведені числа ділянок  на які впало  снарядів

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | 0 | 1 | 2 | 3 | 4 | 5 |
|  | 229 | 211 | 93 | 35 | 7 | 1 |

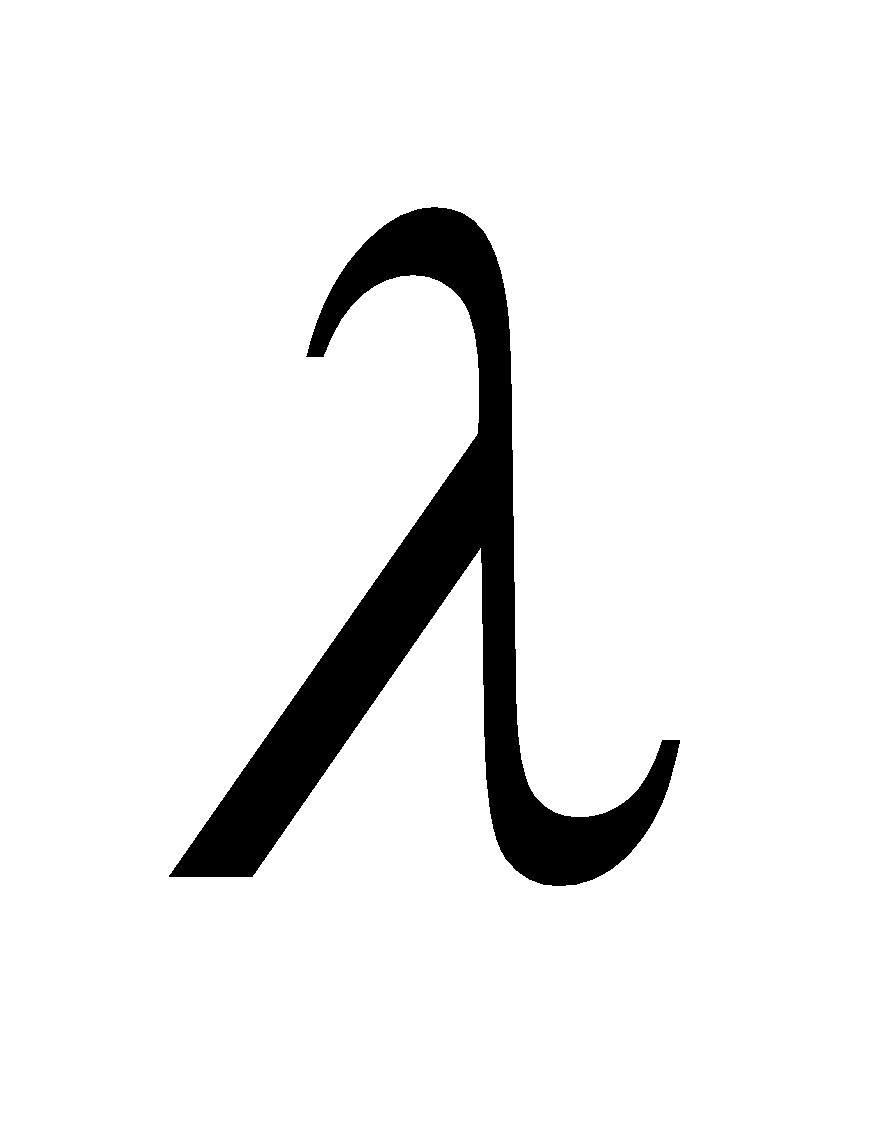
Чи узгоджуються ці дані з гіпотезою про те, що число снарядів, які впали на ділянку, мають розподіл Пуассона? Прийняти .

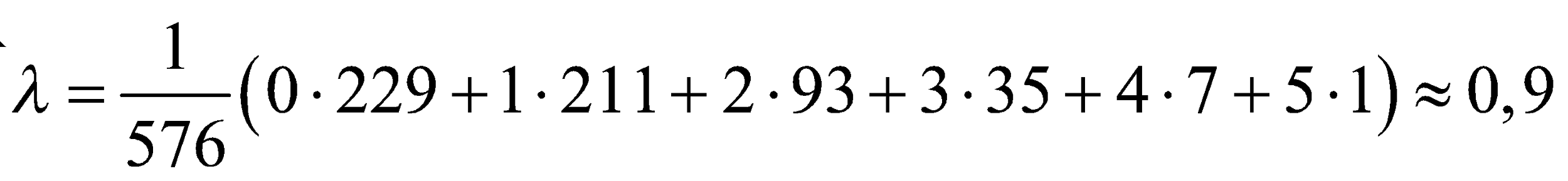
Маємо  незалежних спостережень випадкової величини  – кількості літаків-снарядів, що впали на ділянку.

Відносно невідомого розподілу випадкової величини  висувається гіпотеза

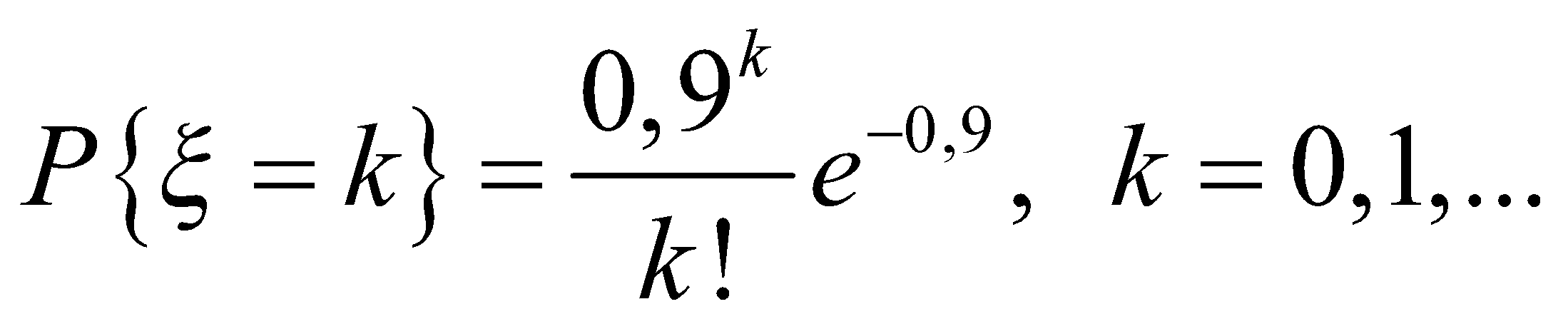
,

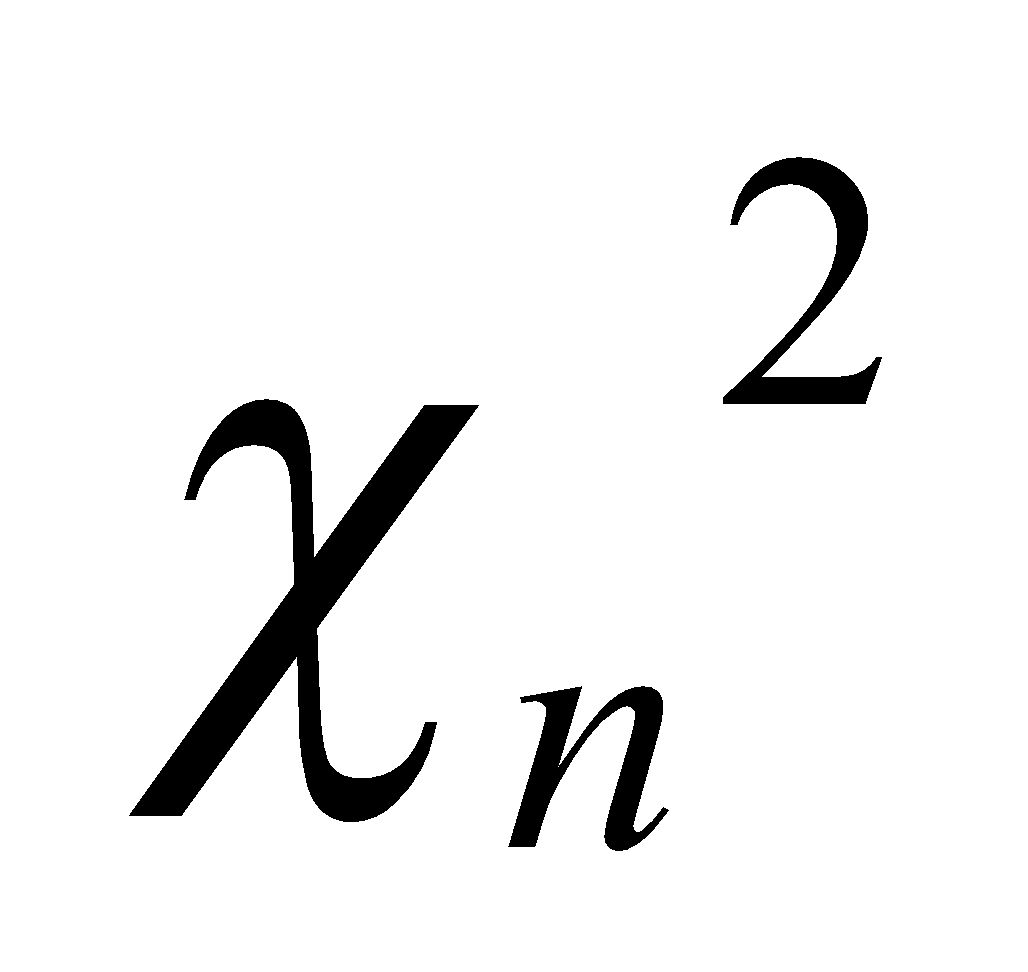
яку треба перевірити.

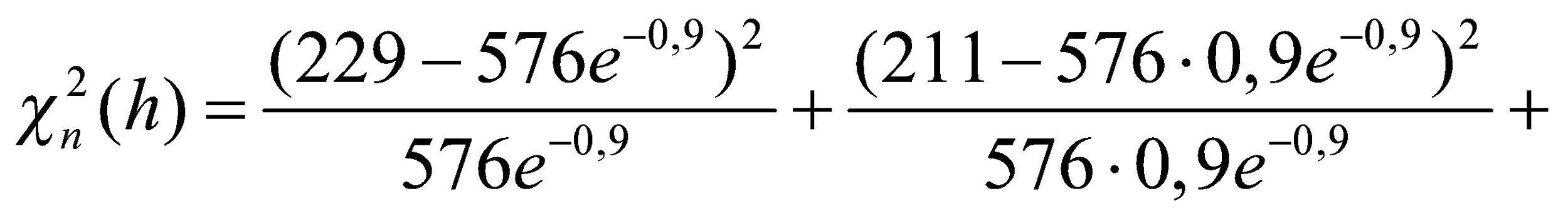
Параметр  гіпотетичного розподілу невідомий. Згідно зауваженню 2, в якості його оцінки ми можемо взяти вибіркове середнє:

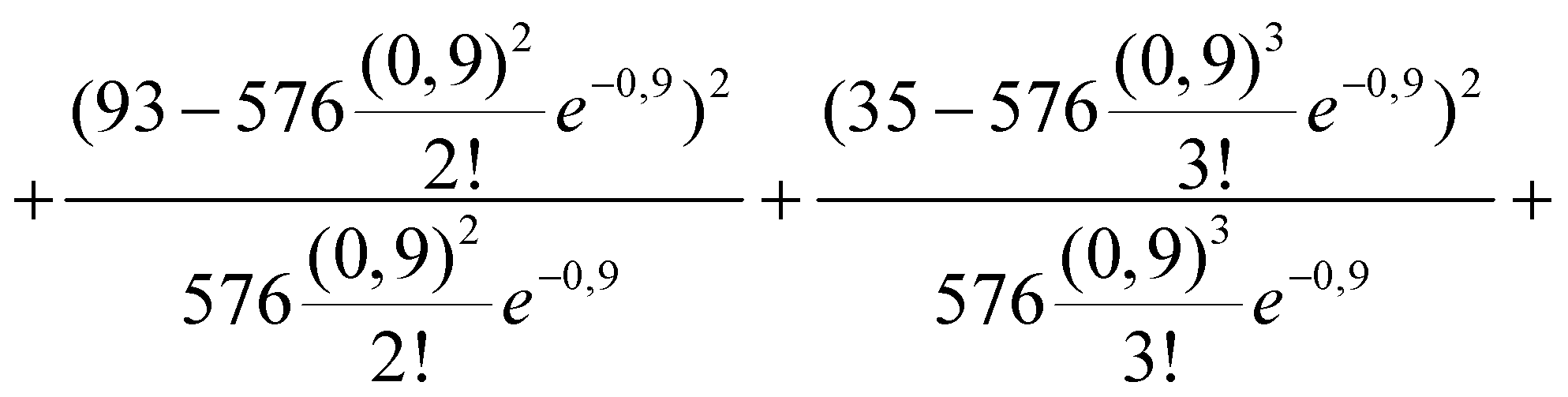
,

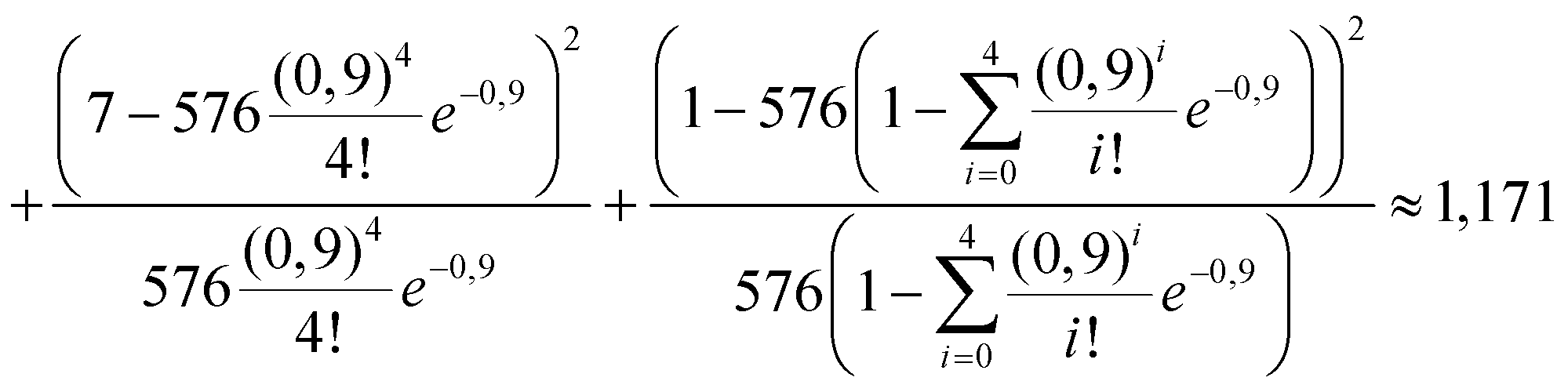
отже гіпотетичний розподіл має вигляд

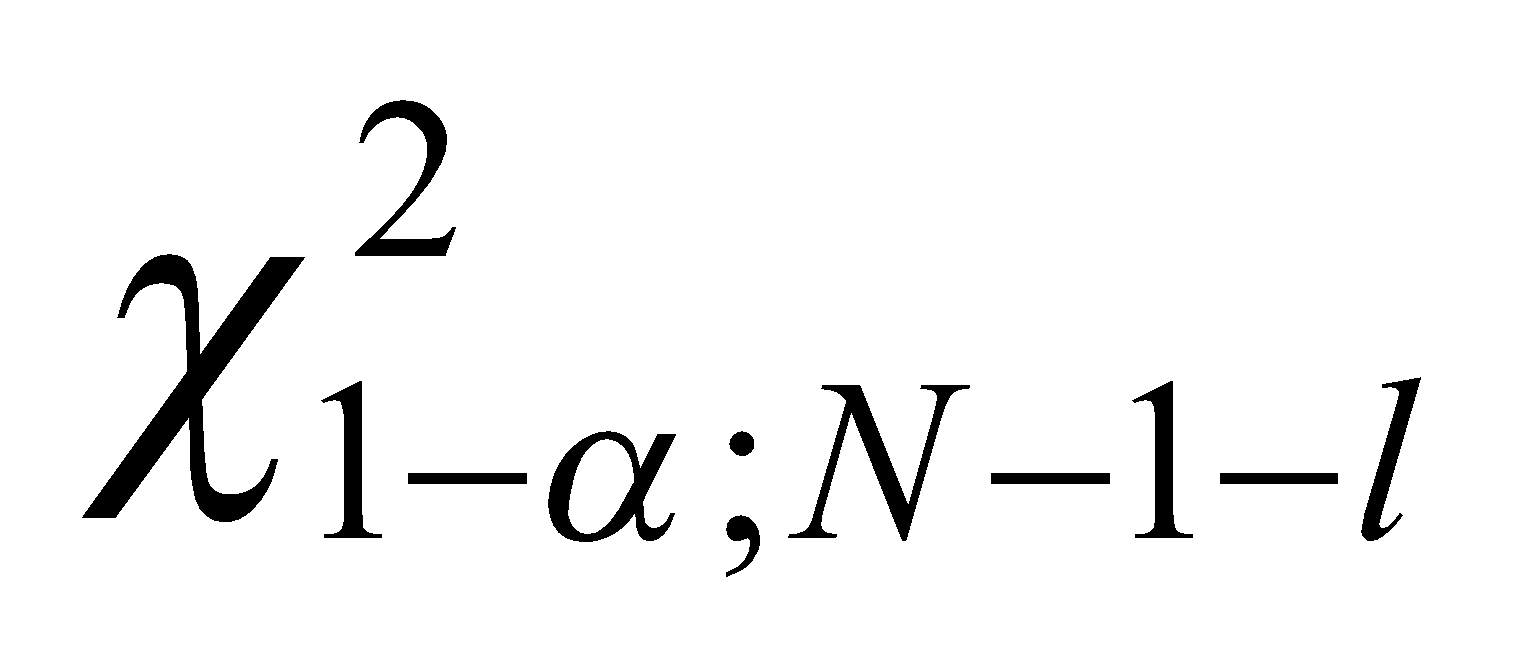
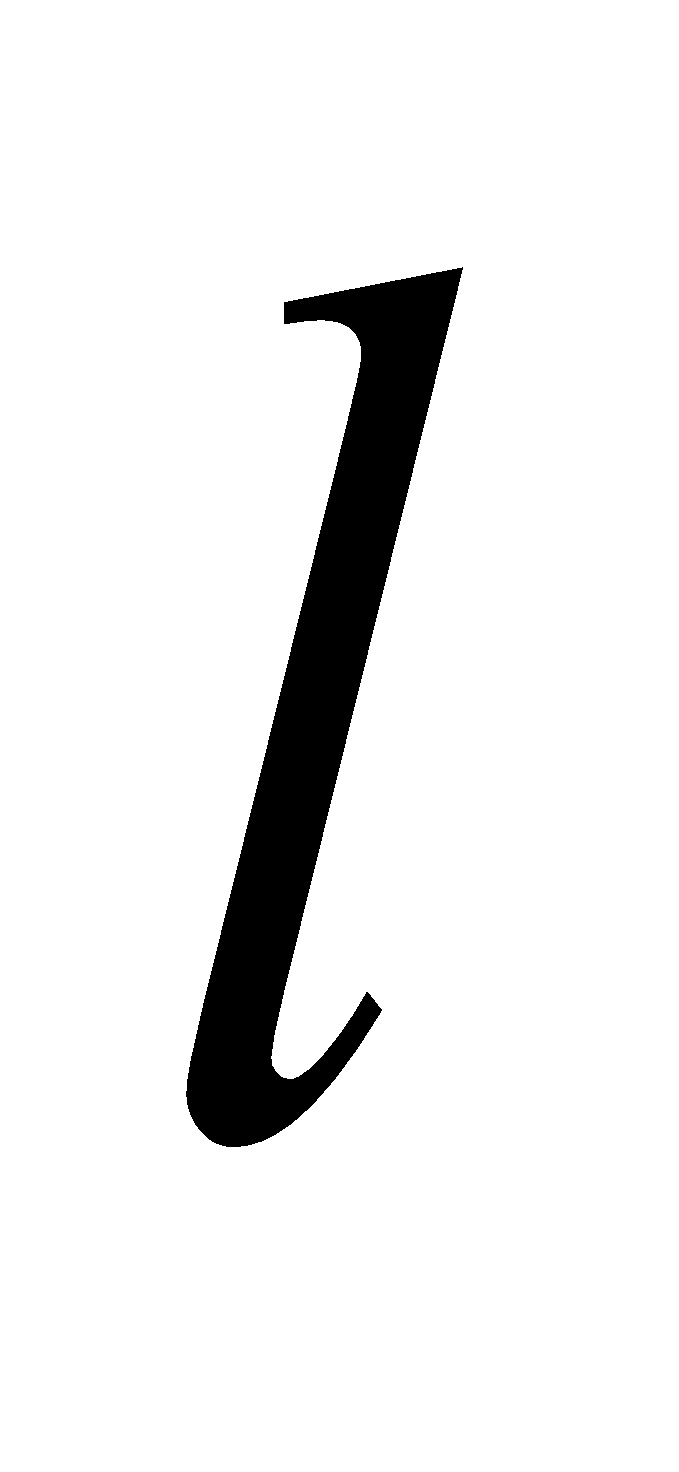
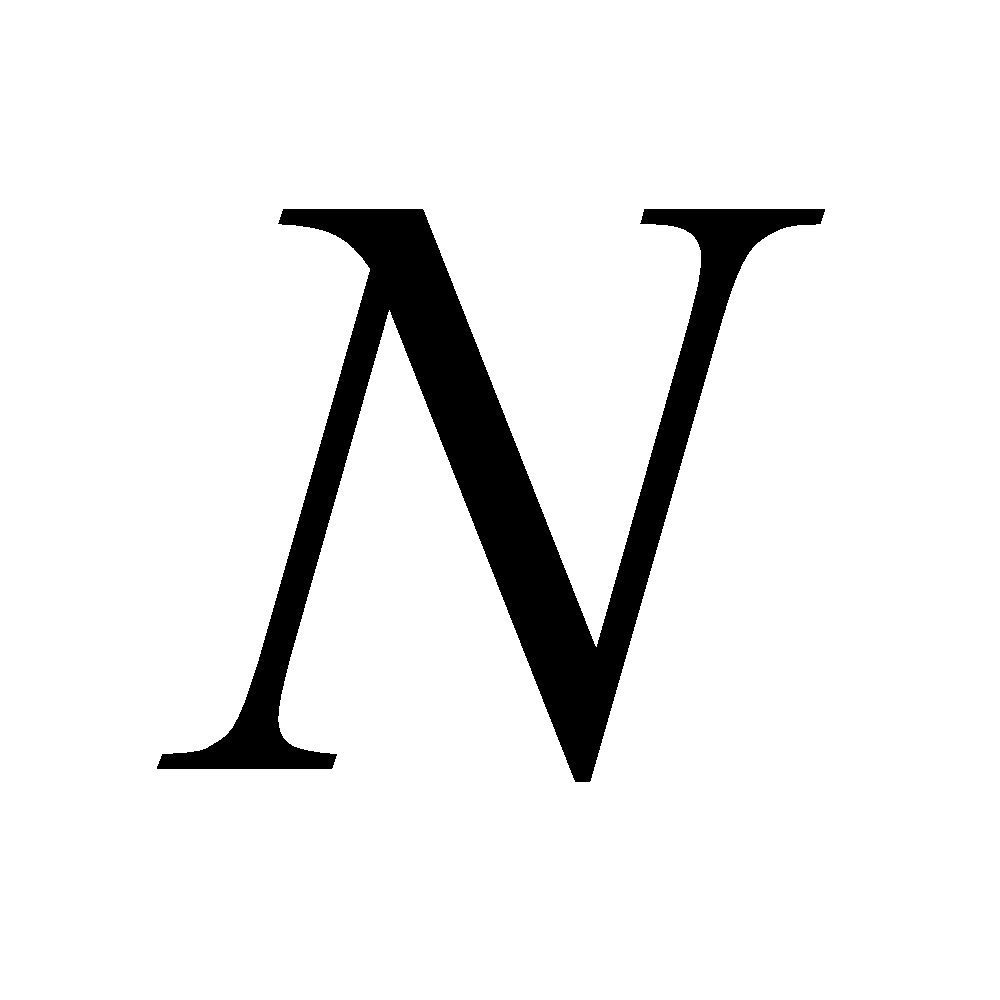
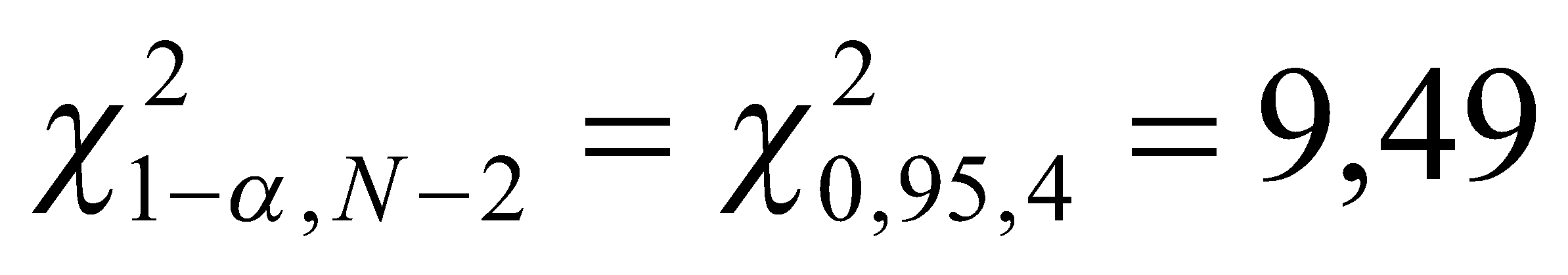
.

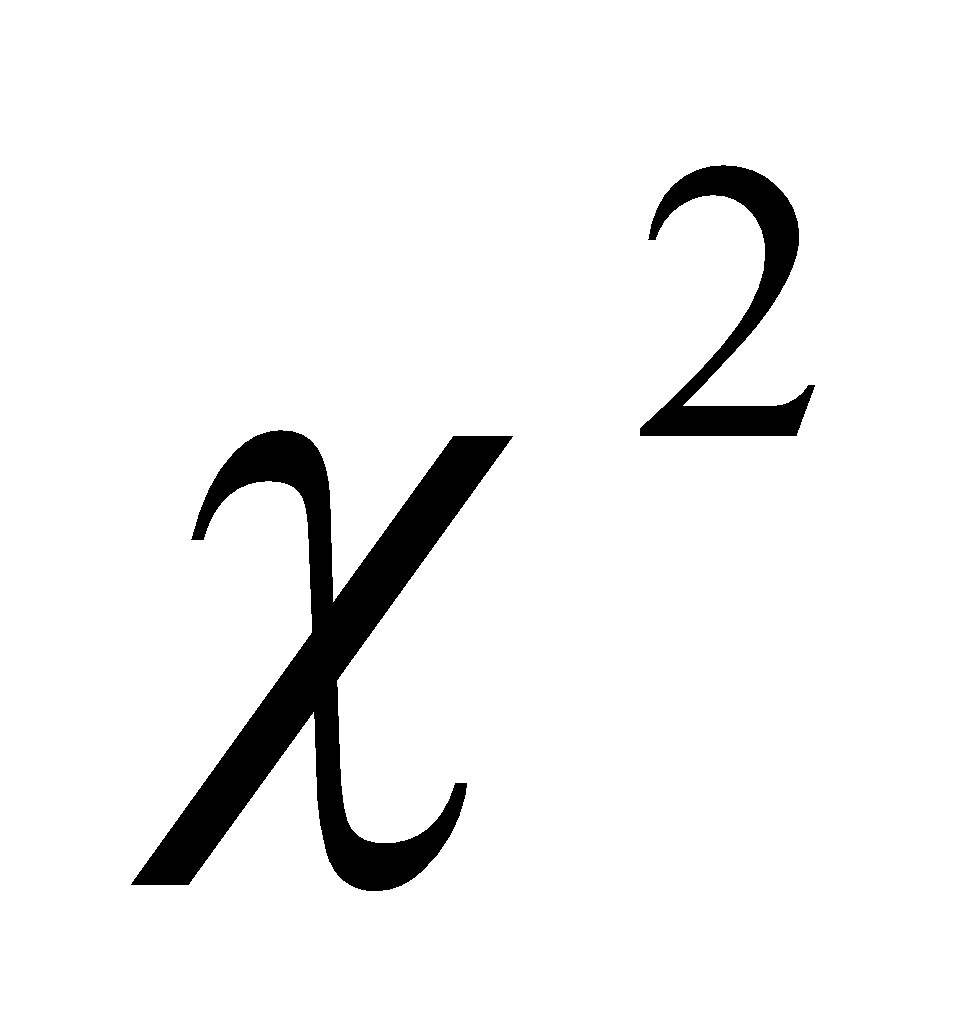
Користуючись наведеною вище методикою, шукаємо значення статистики критерію :

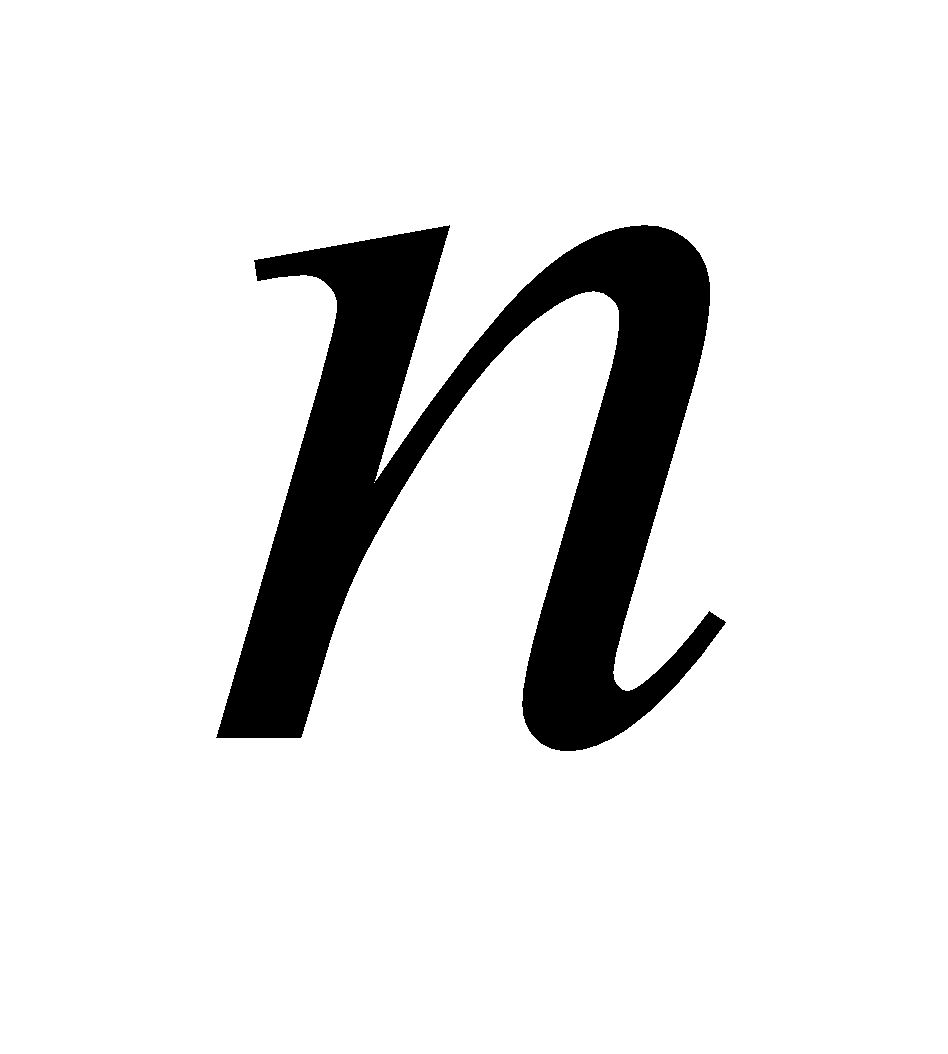
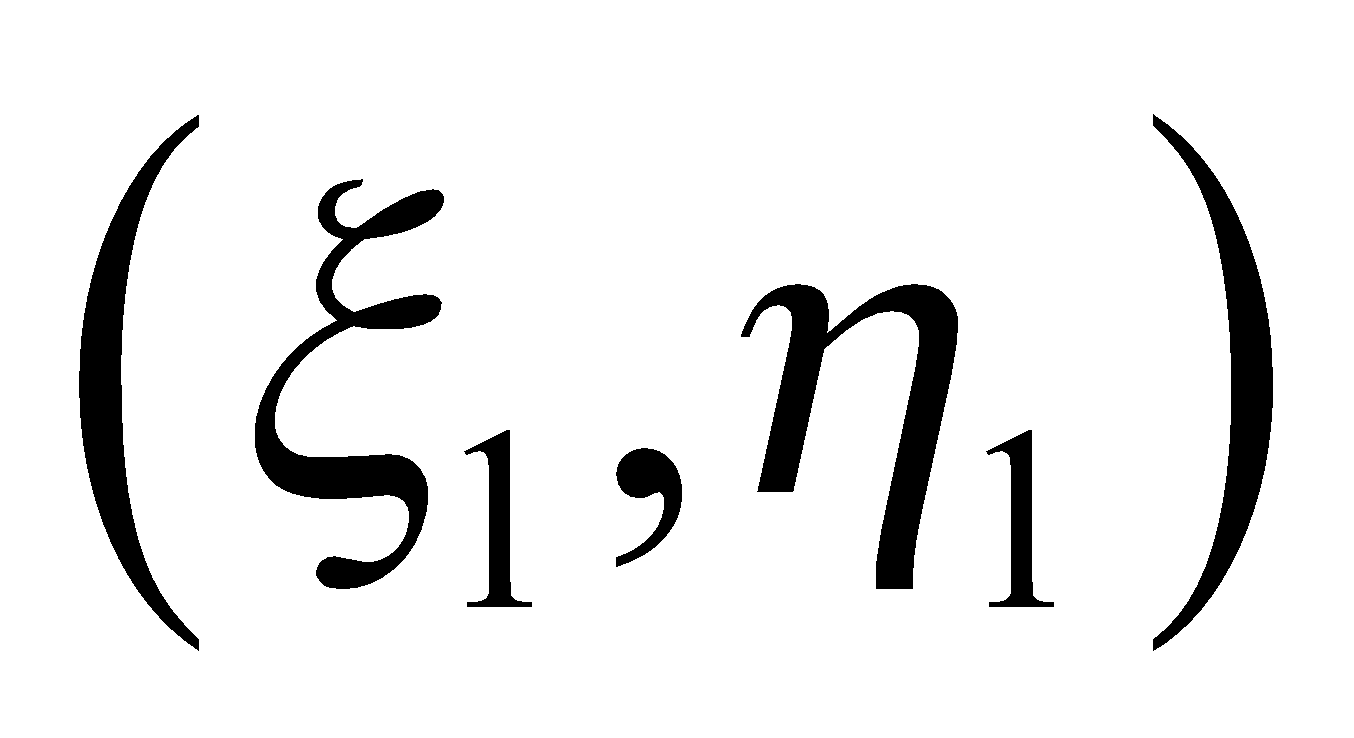
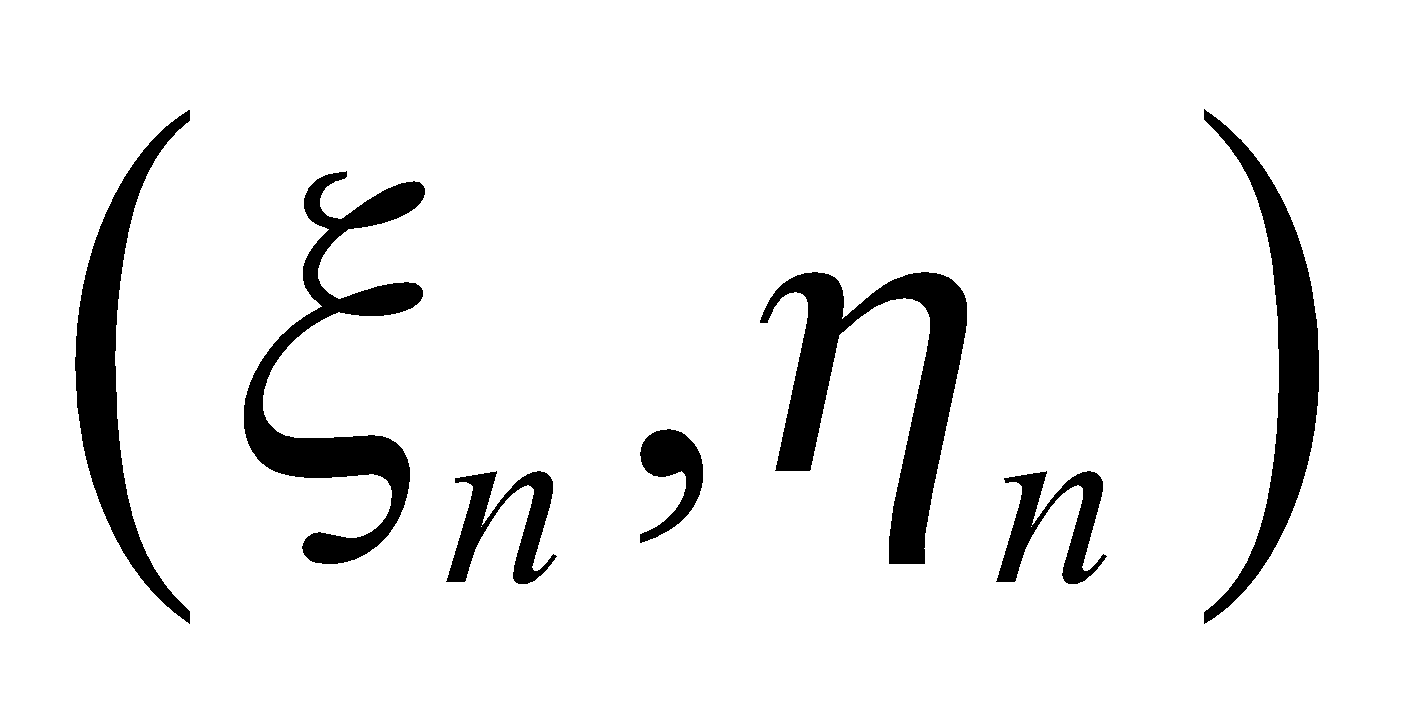
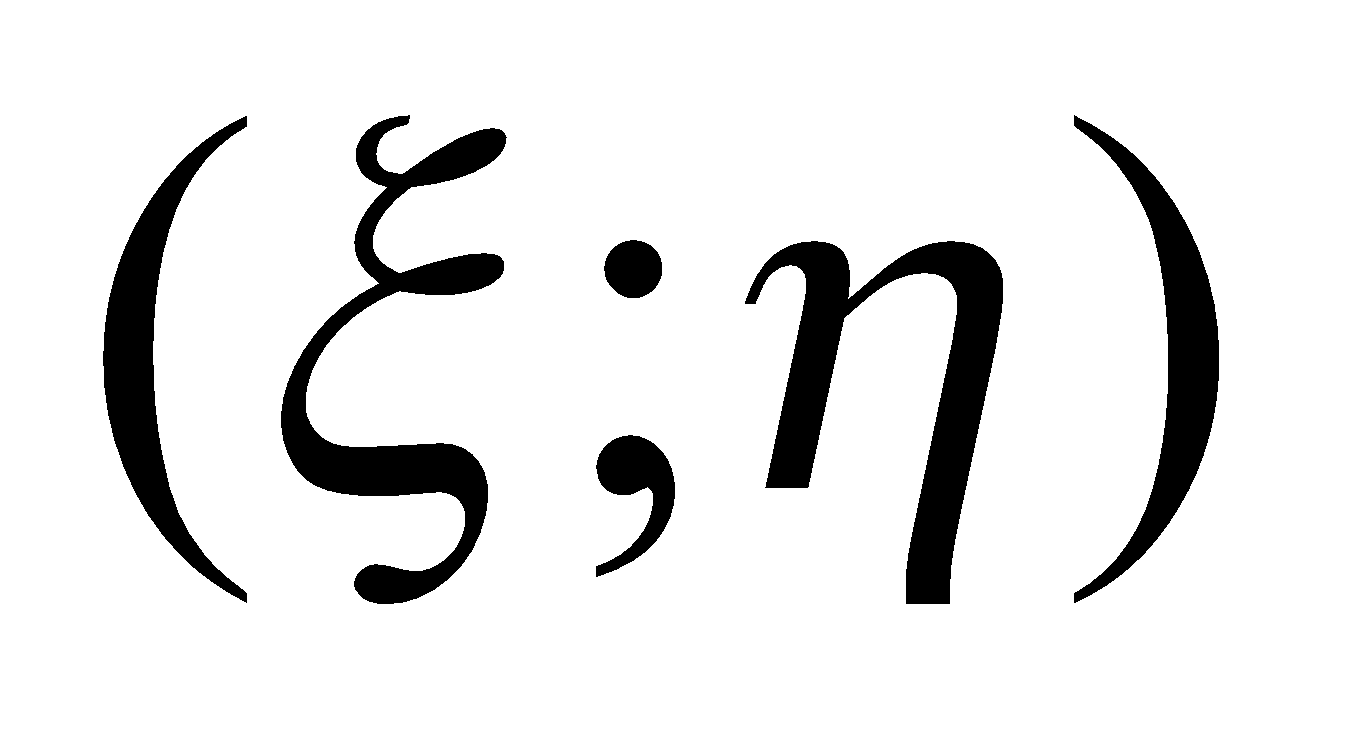
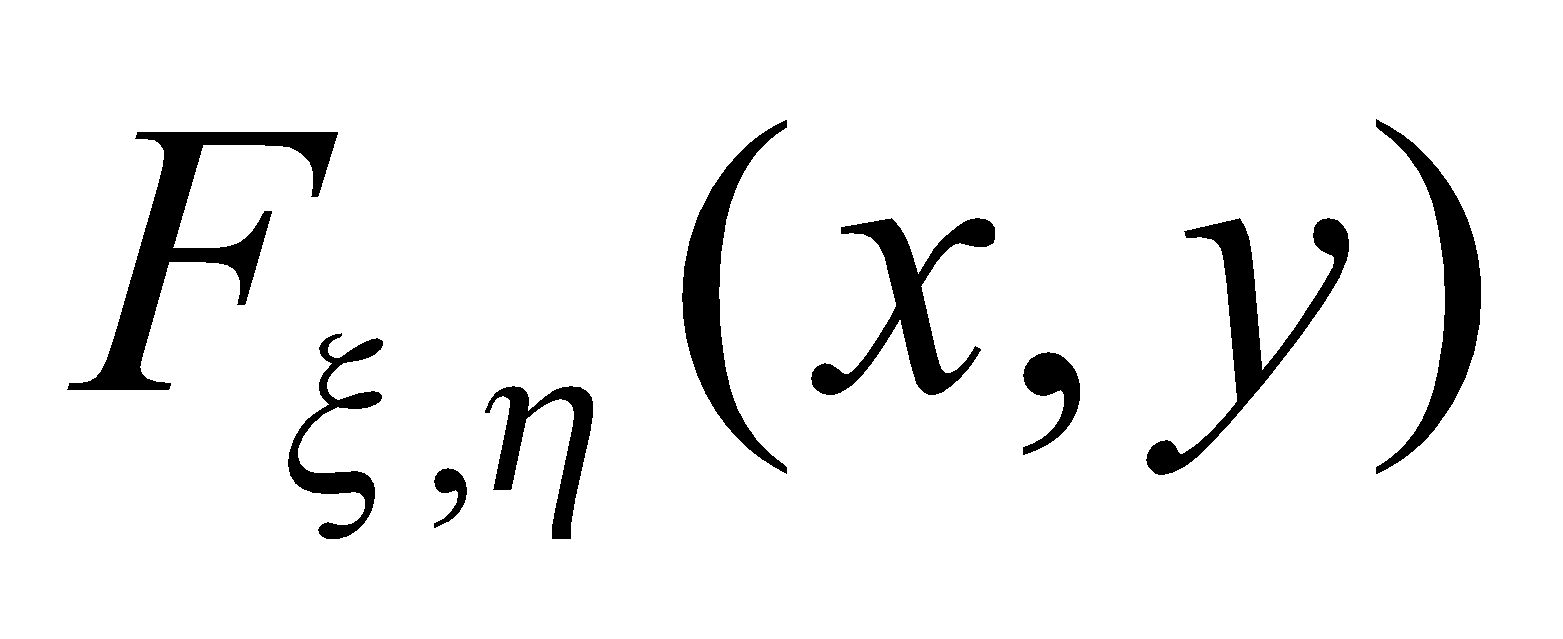


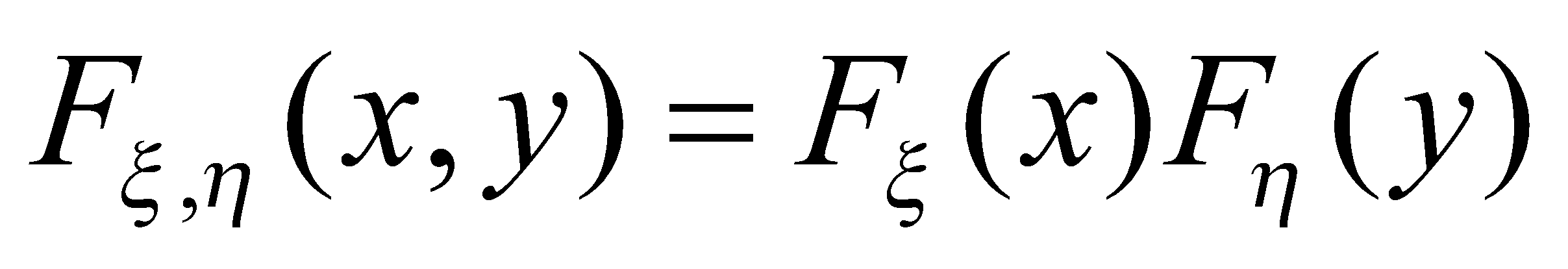


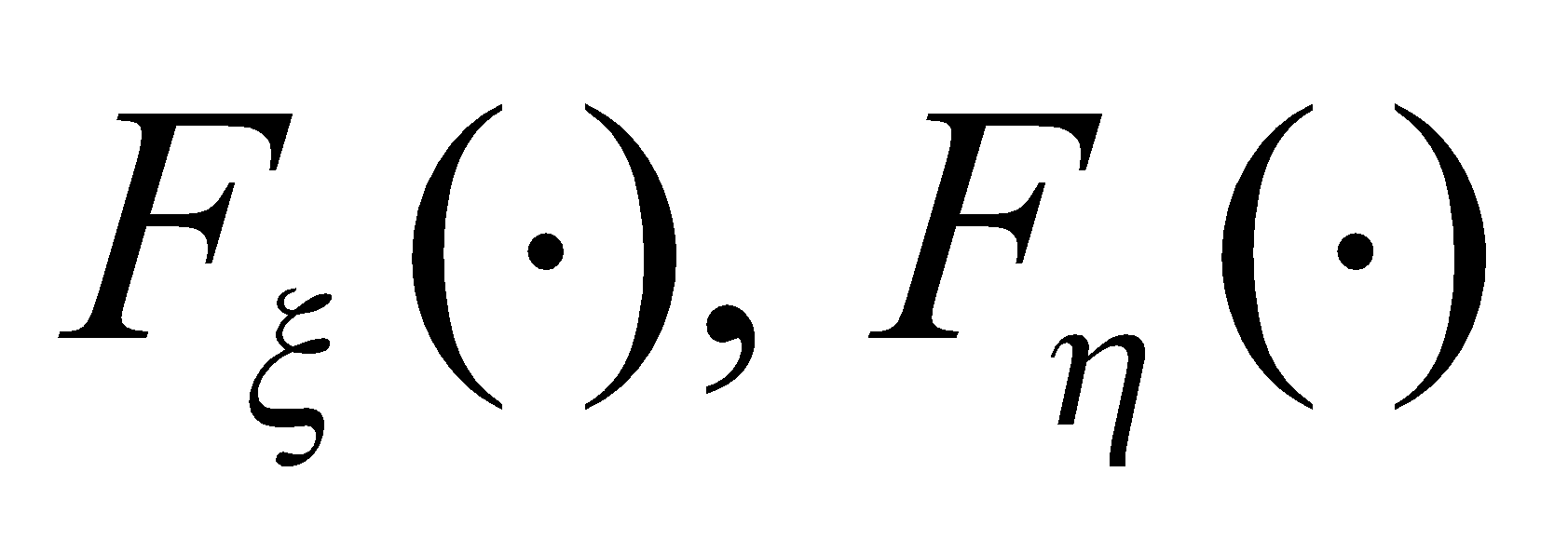


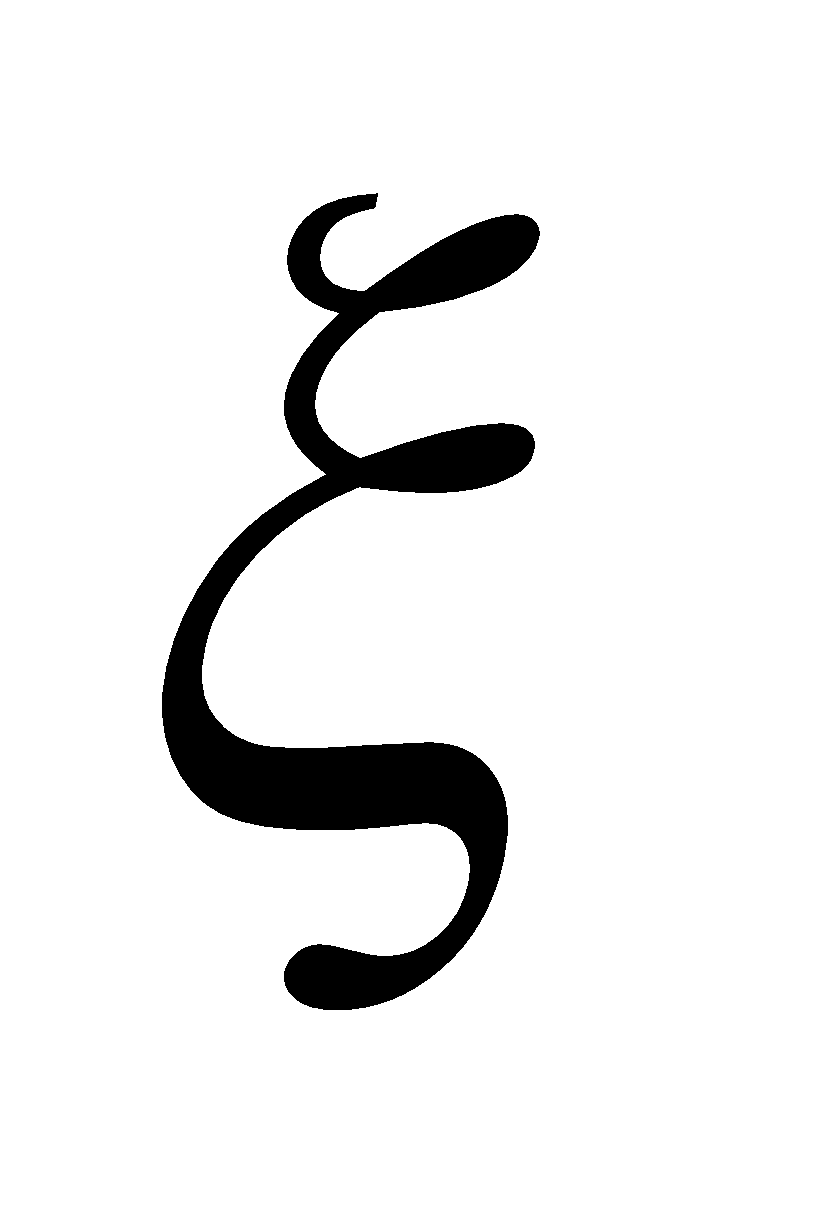
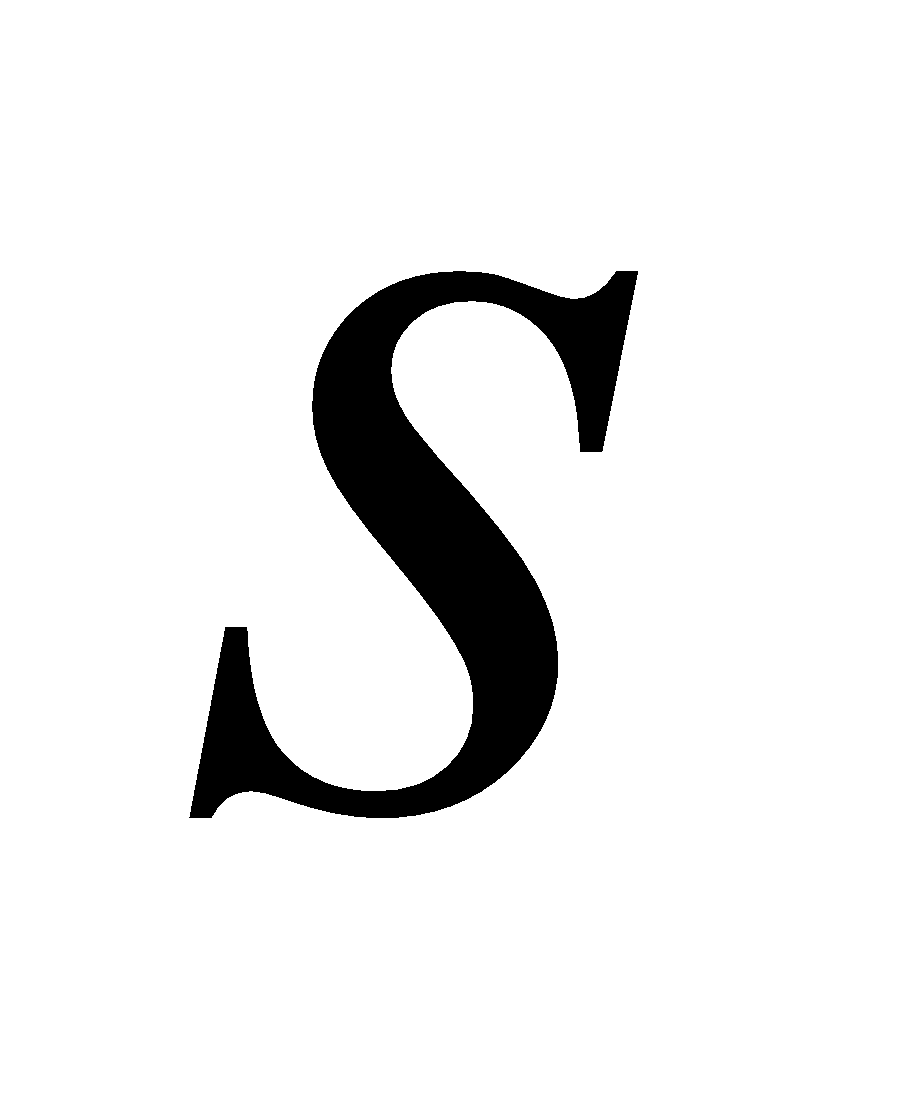
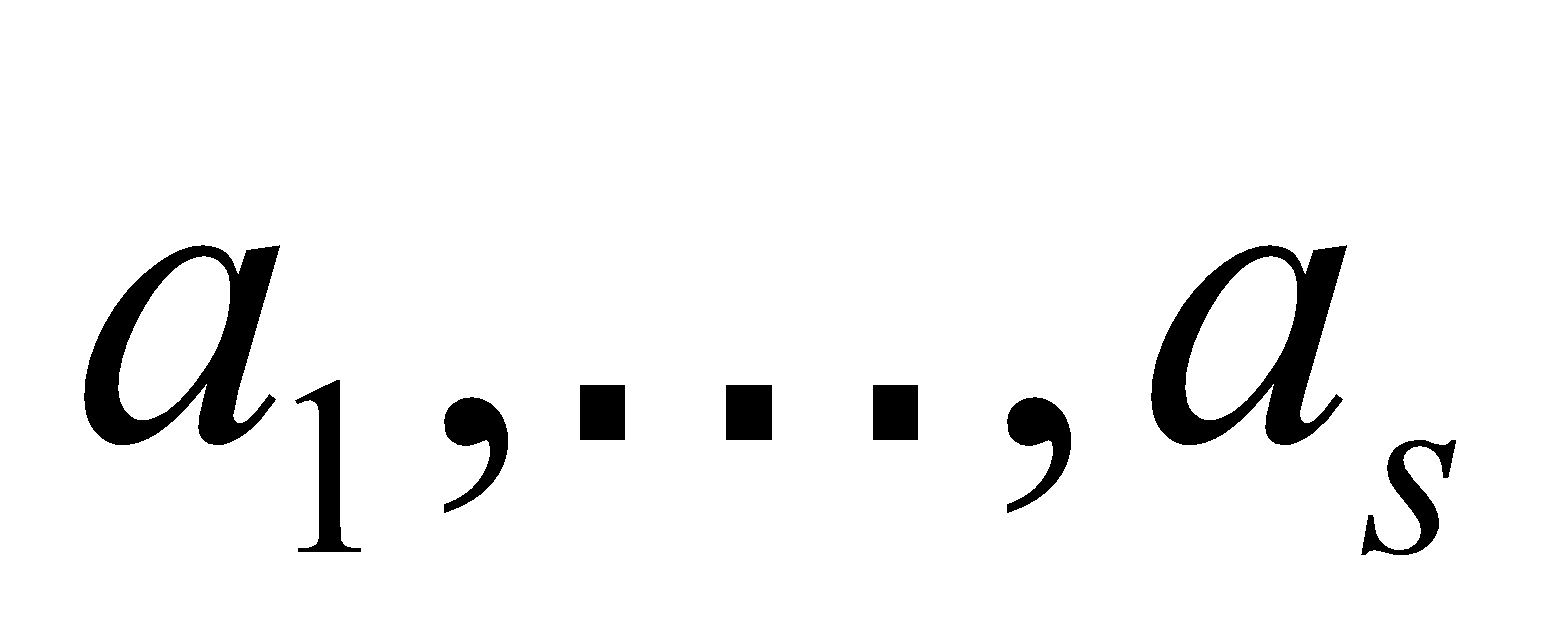
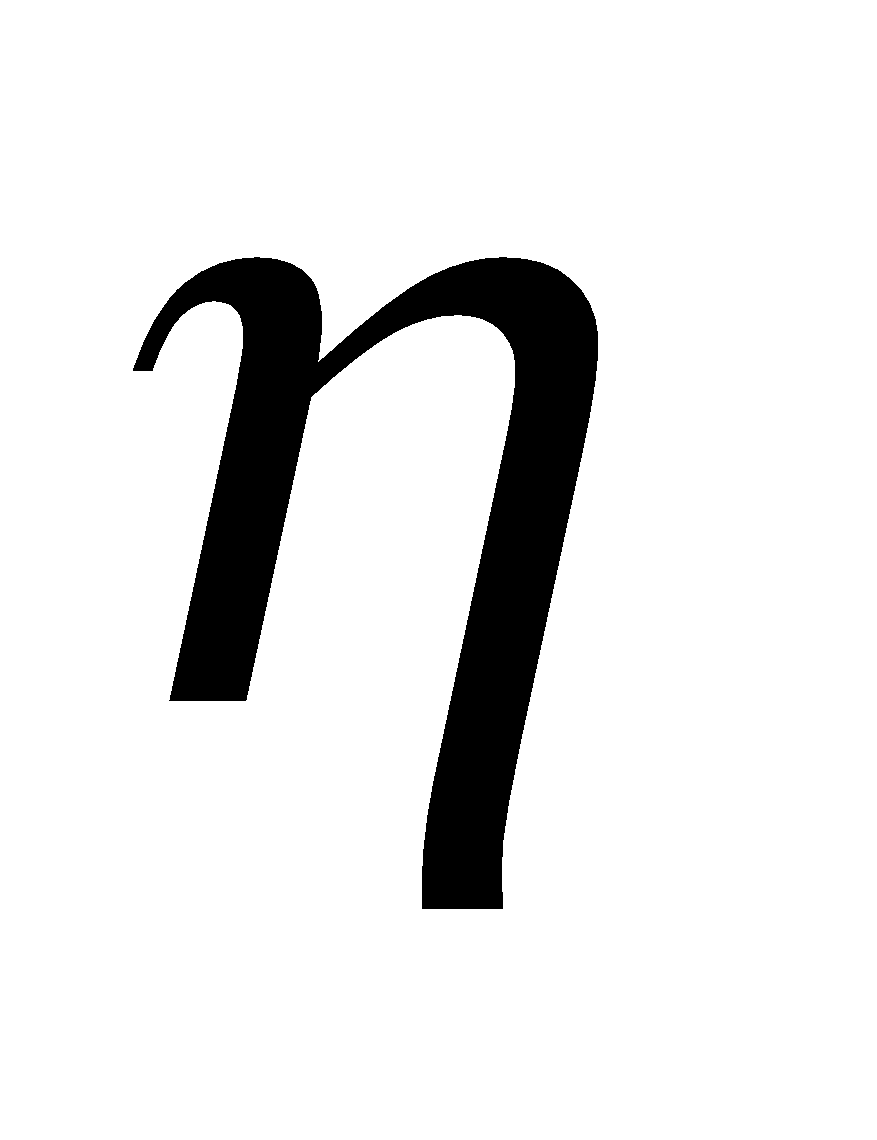
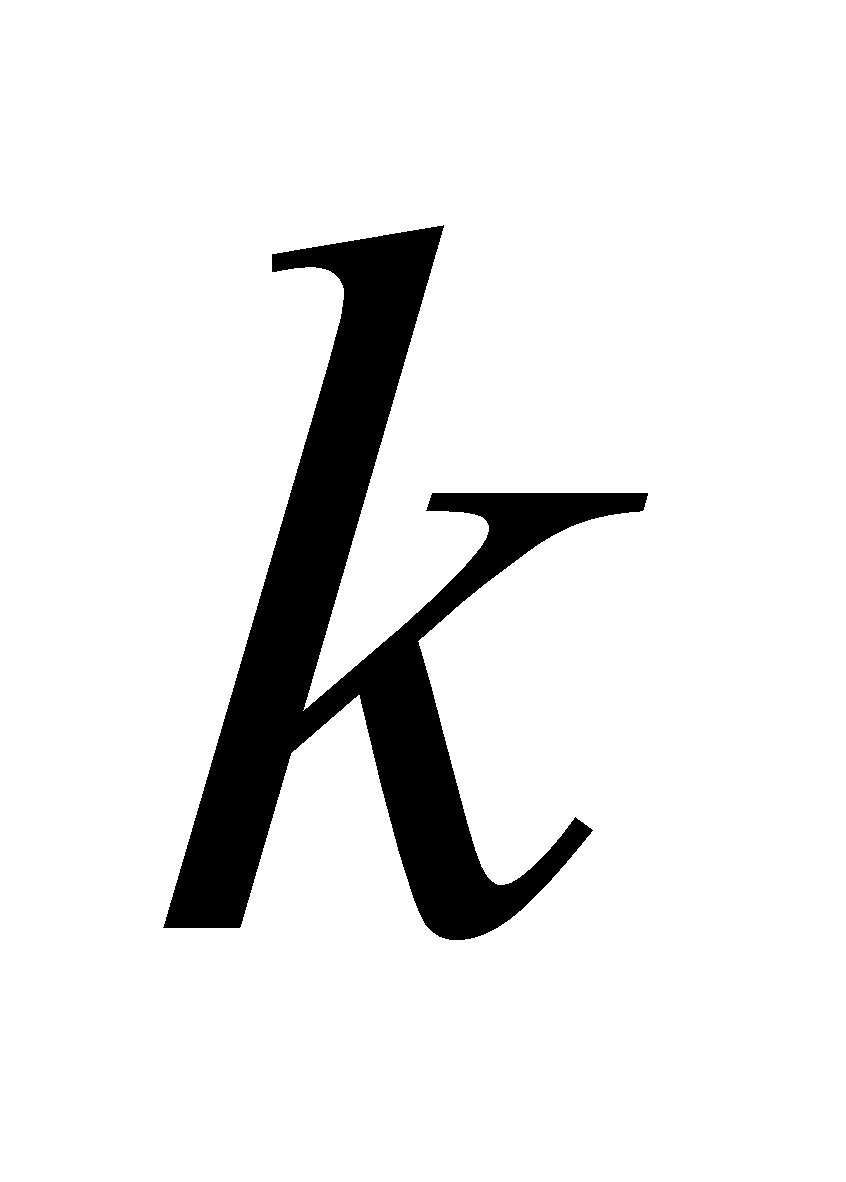
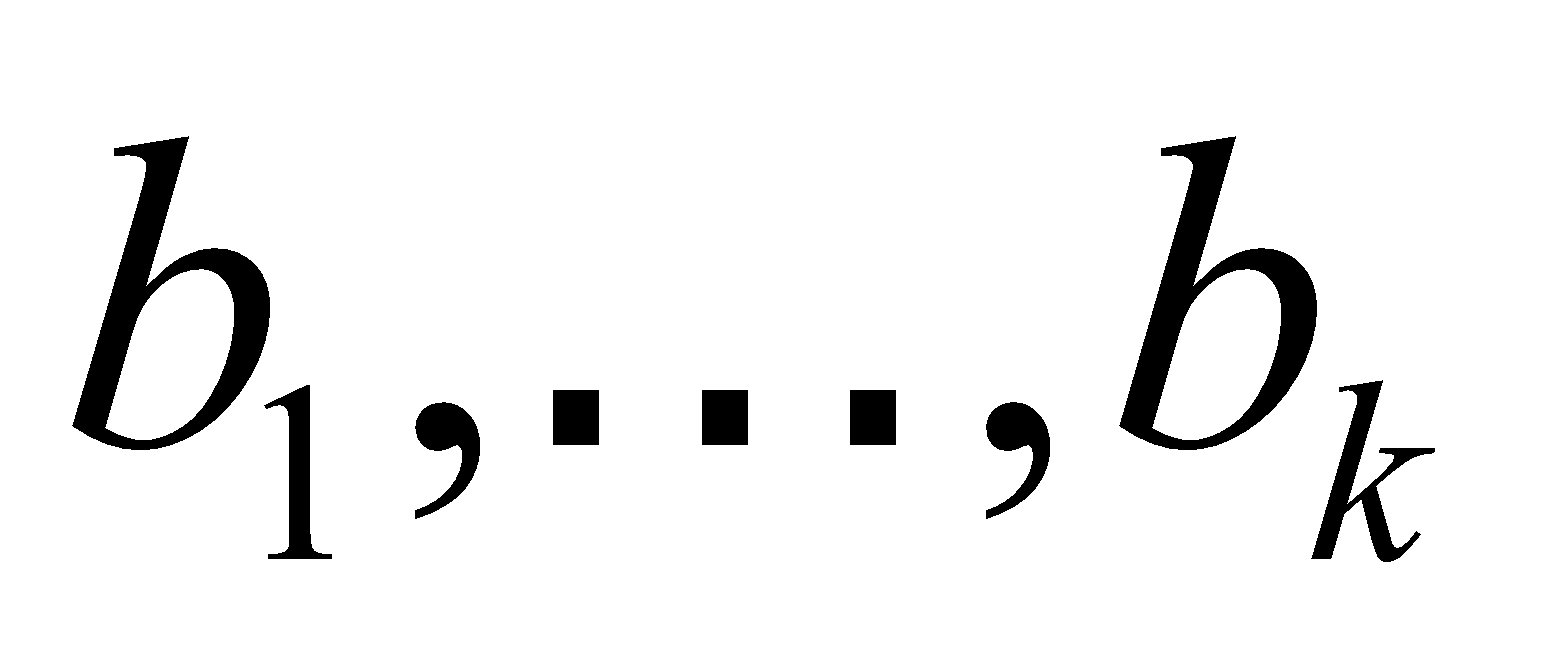
Отримане значення порівнюємо з табличним значенням , де  – кількість параметрів, оцінених по вибірці,  – кількість підмножин, на які розбито вибірковий простір. Маємо . Таким чином експериментальні дані узгоджуються з гіпотезою.

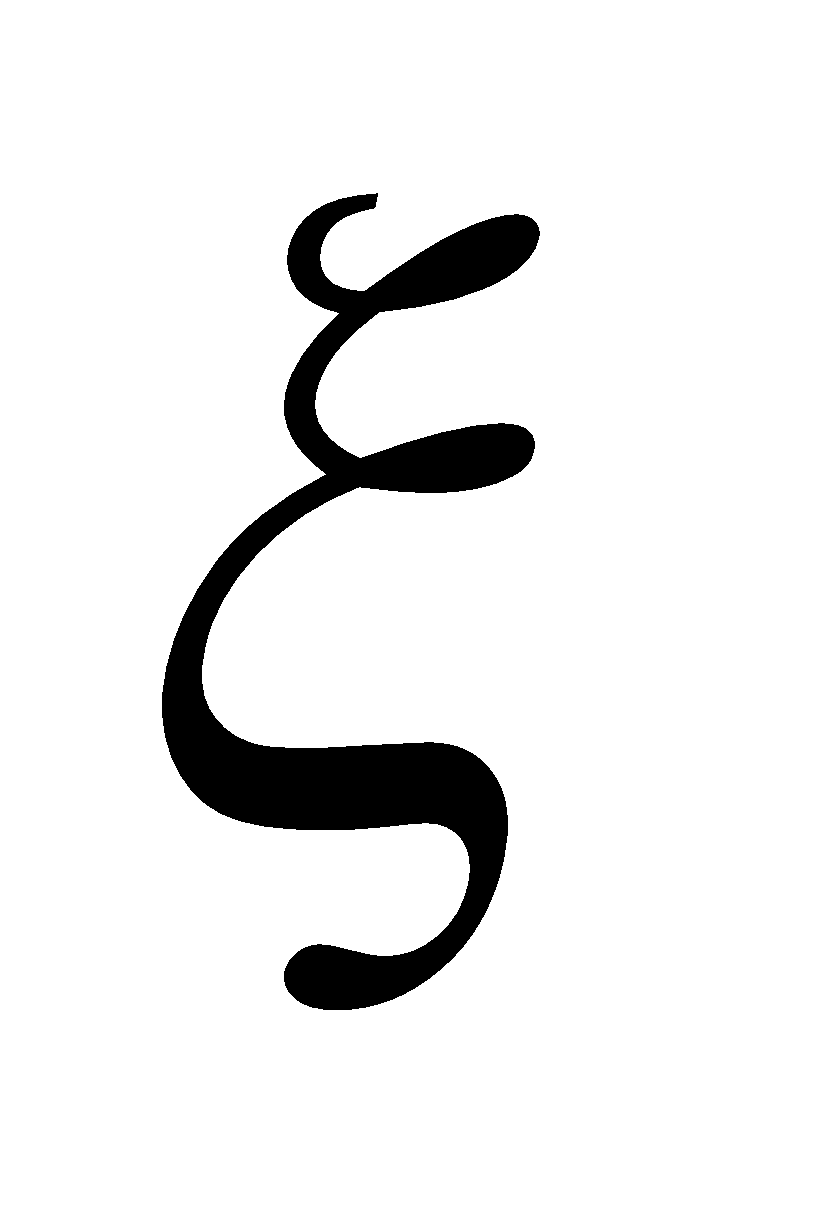
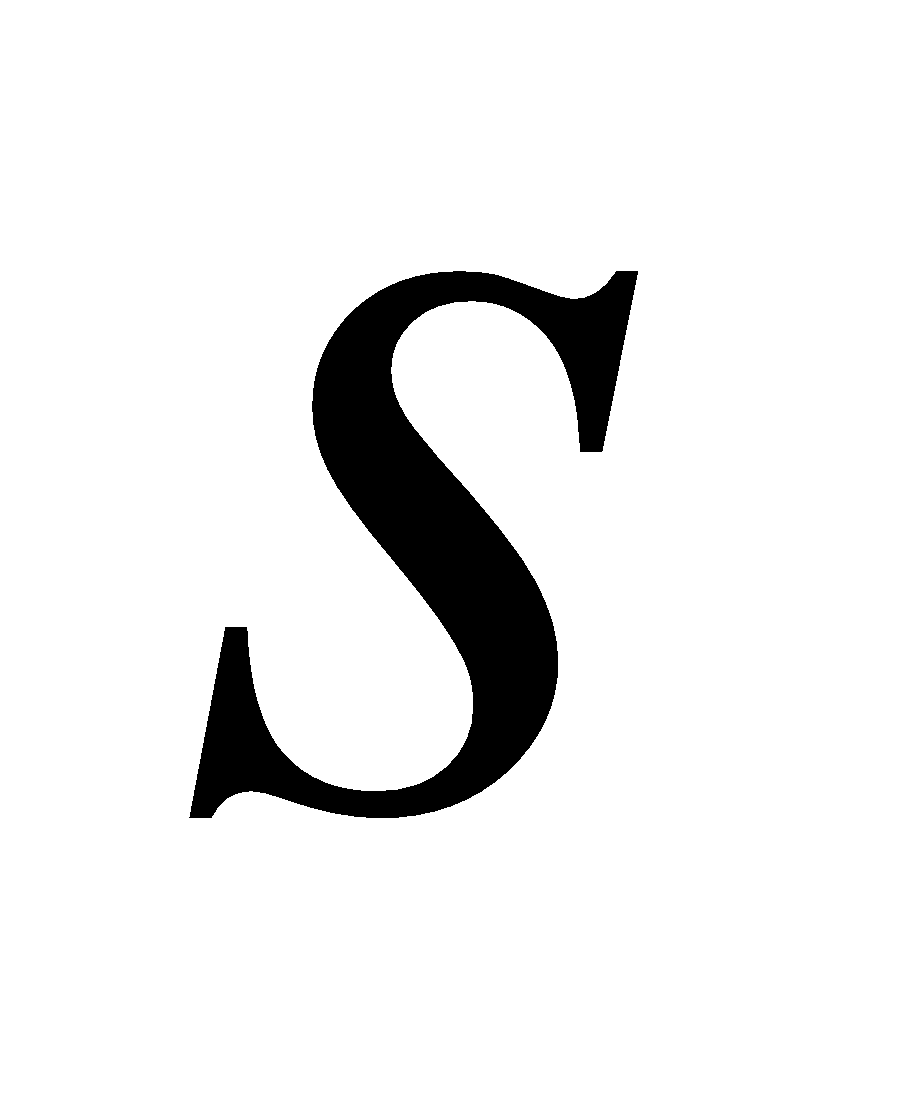
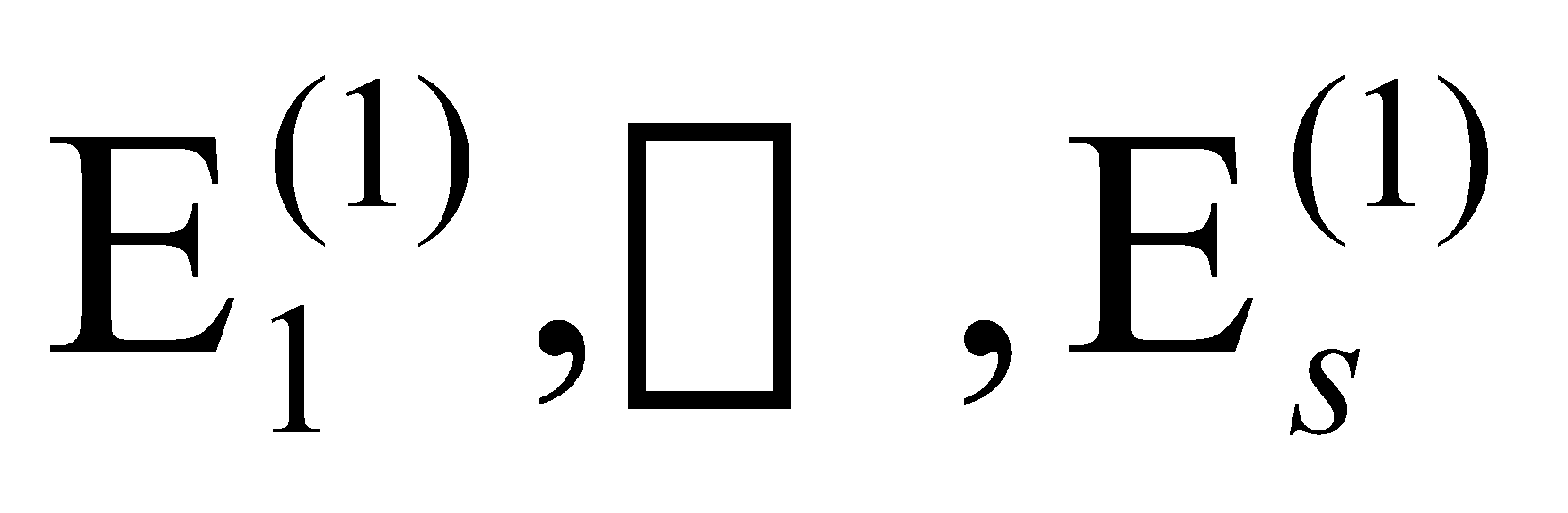
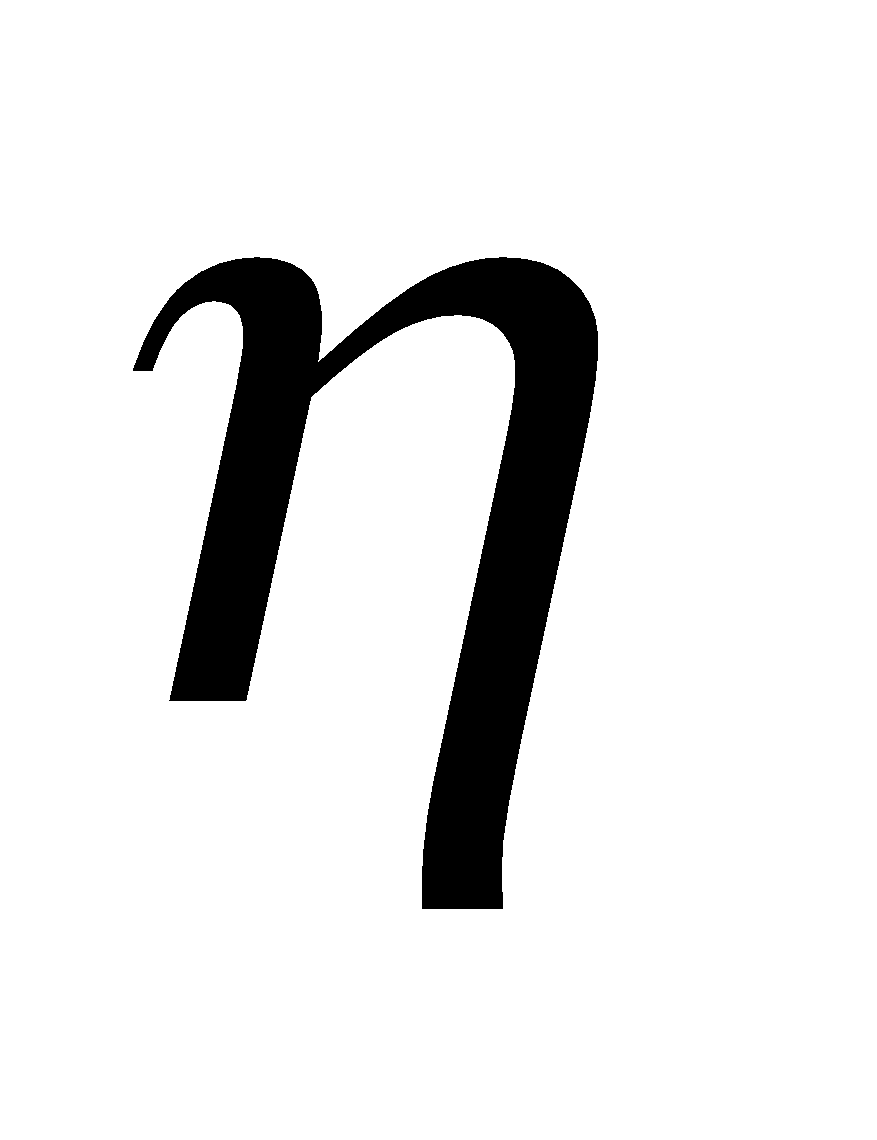
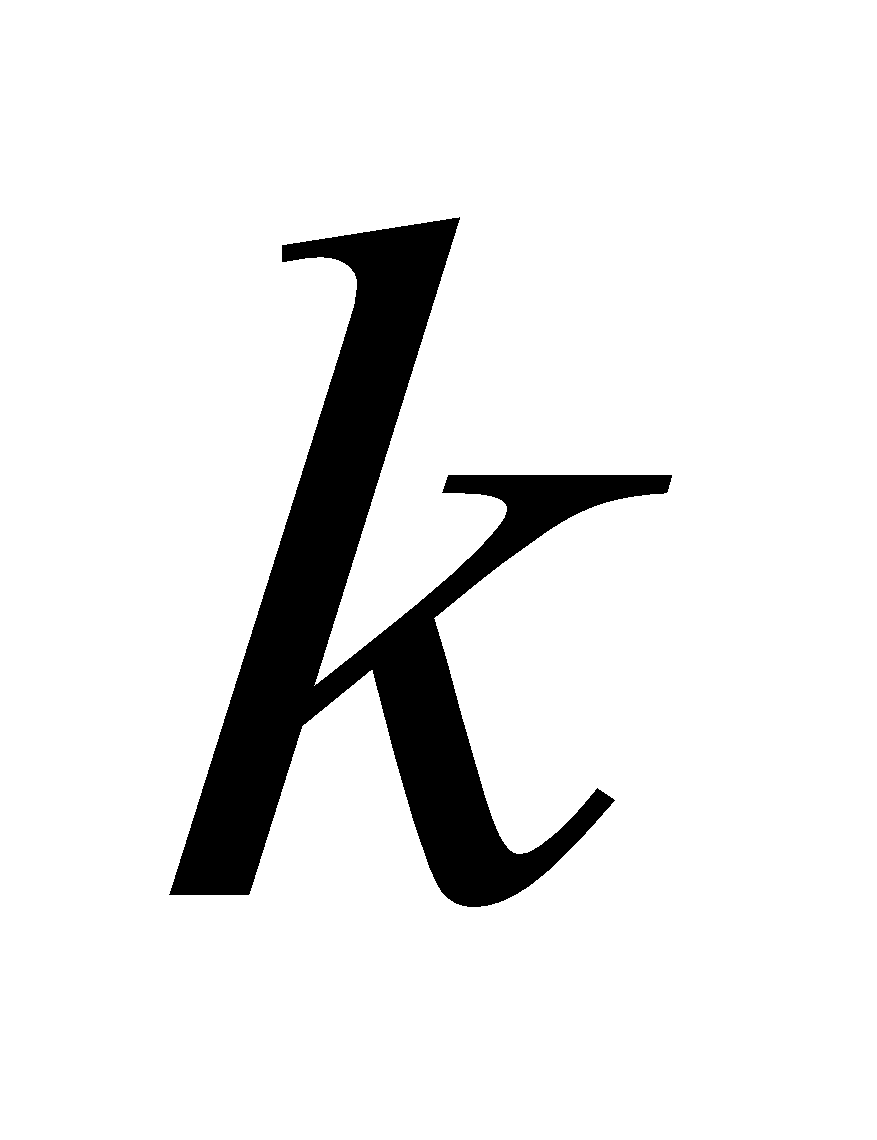
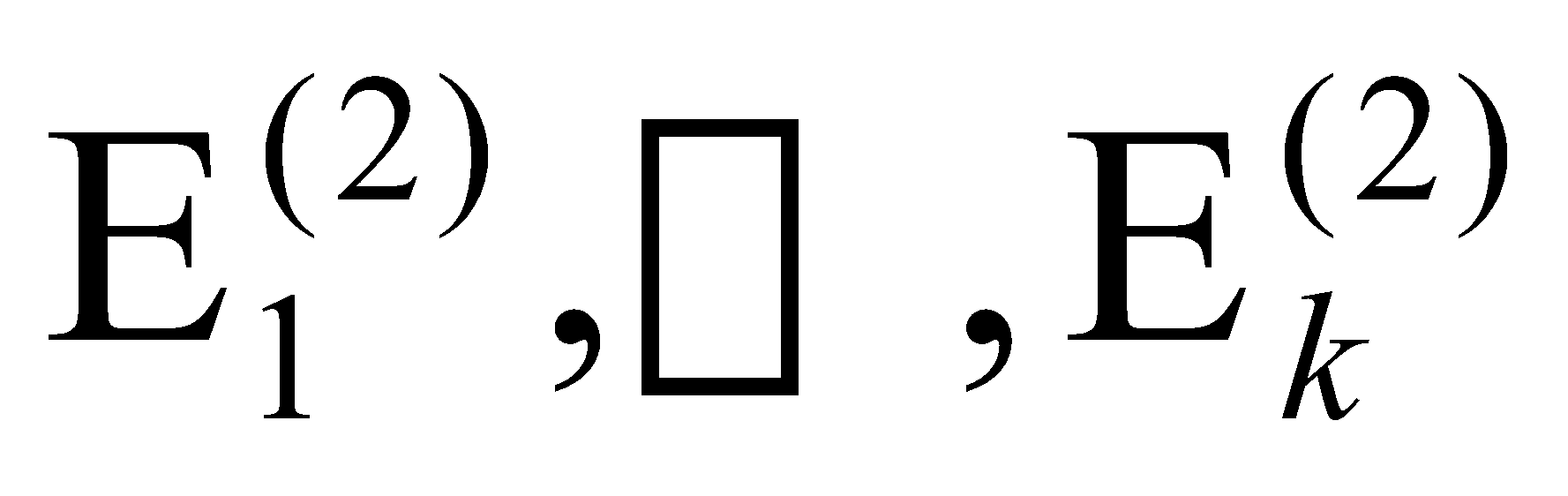
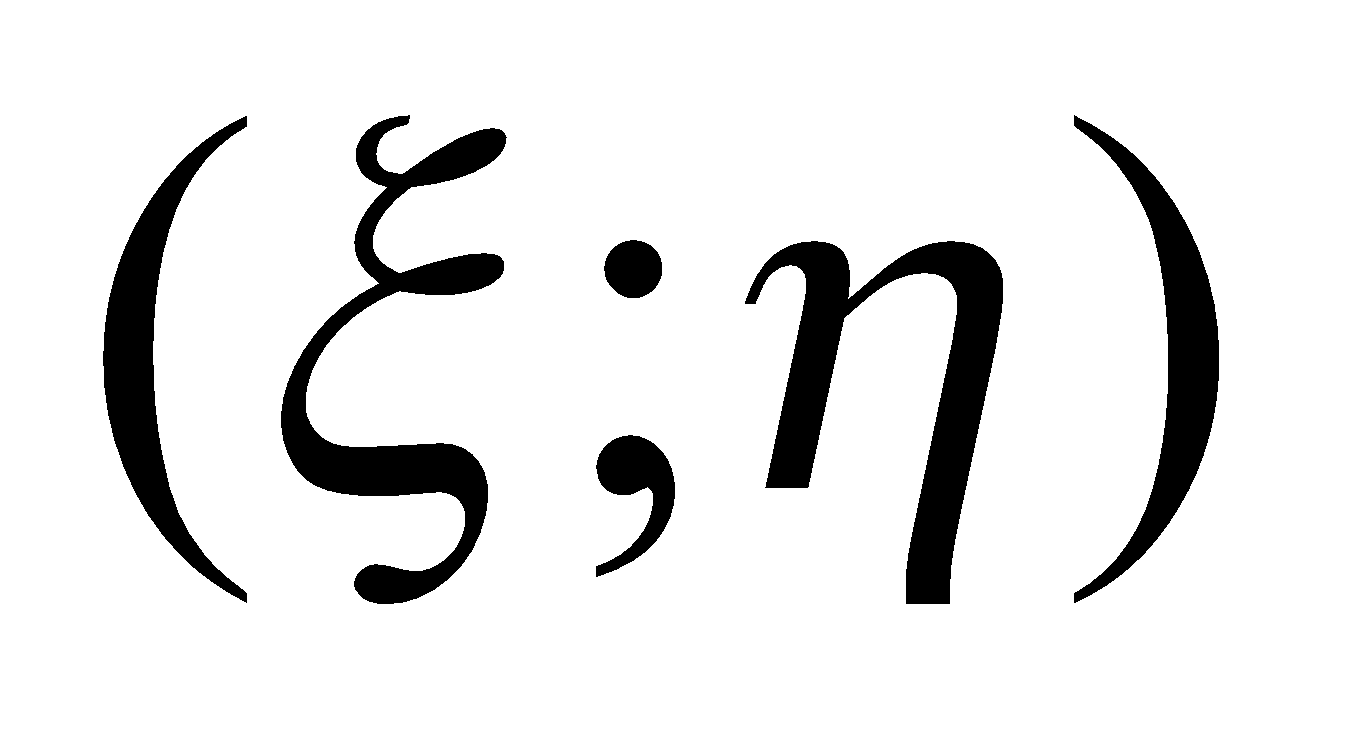
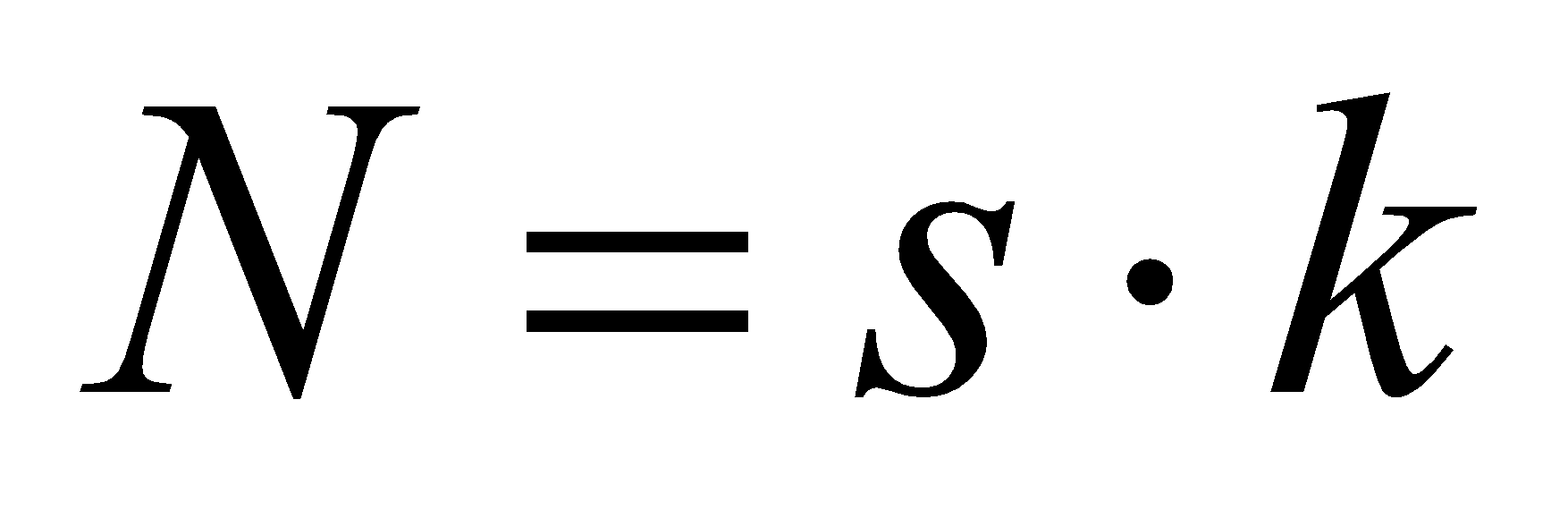
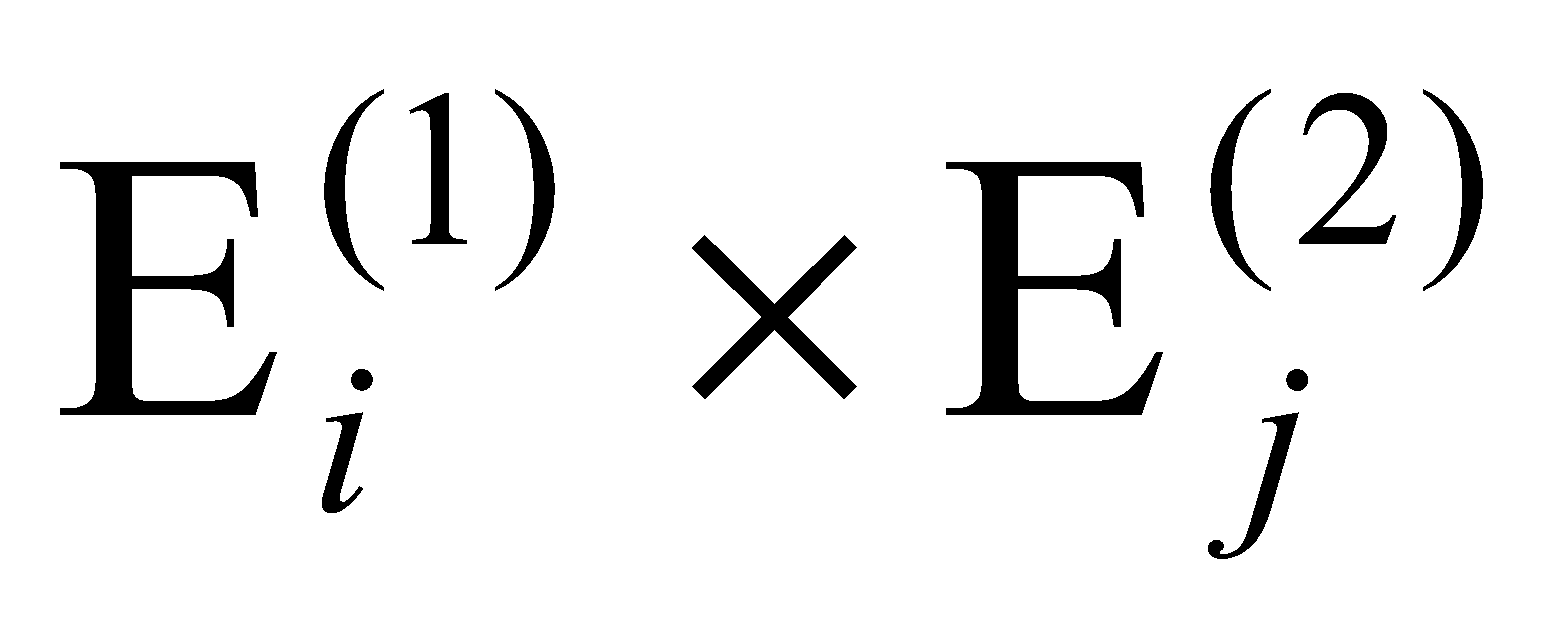
**4.4. Гіпотези незалежності. Критерій незалежності **

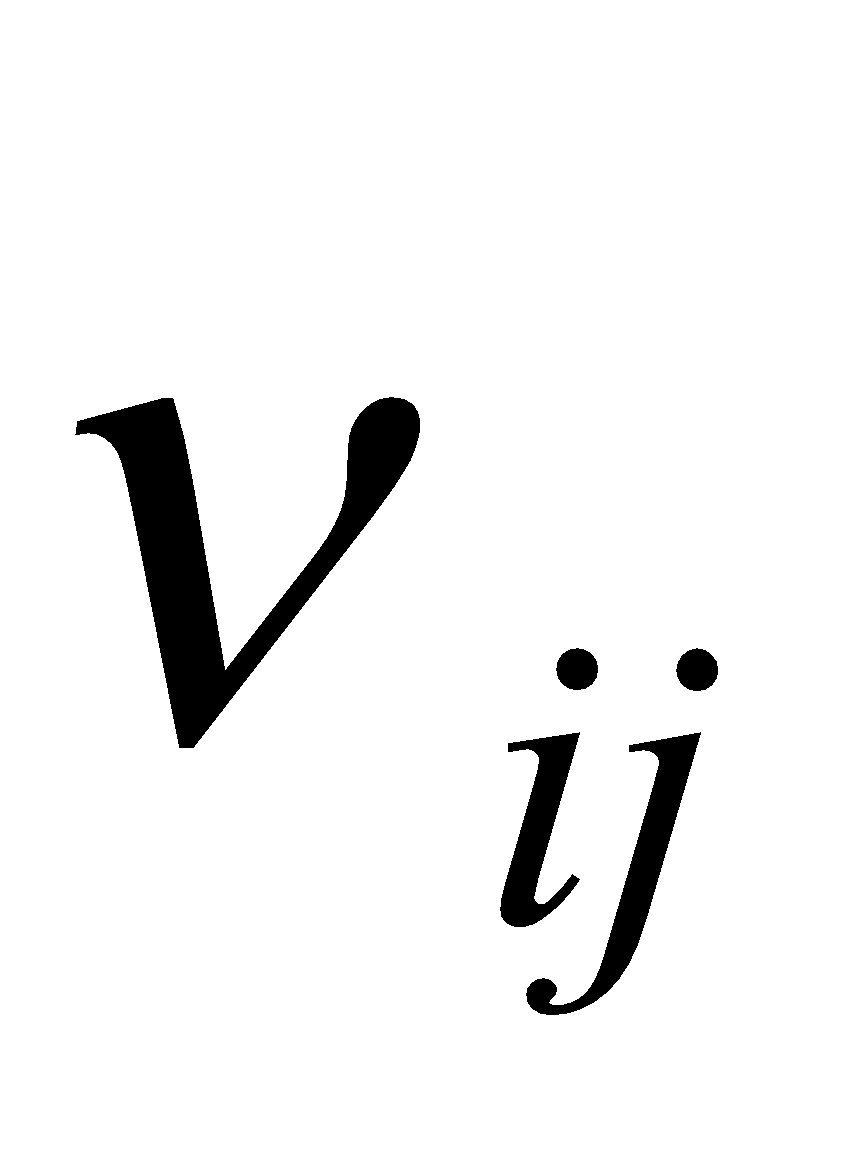
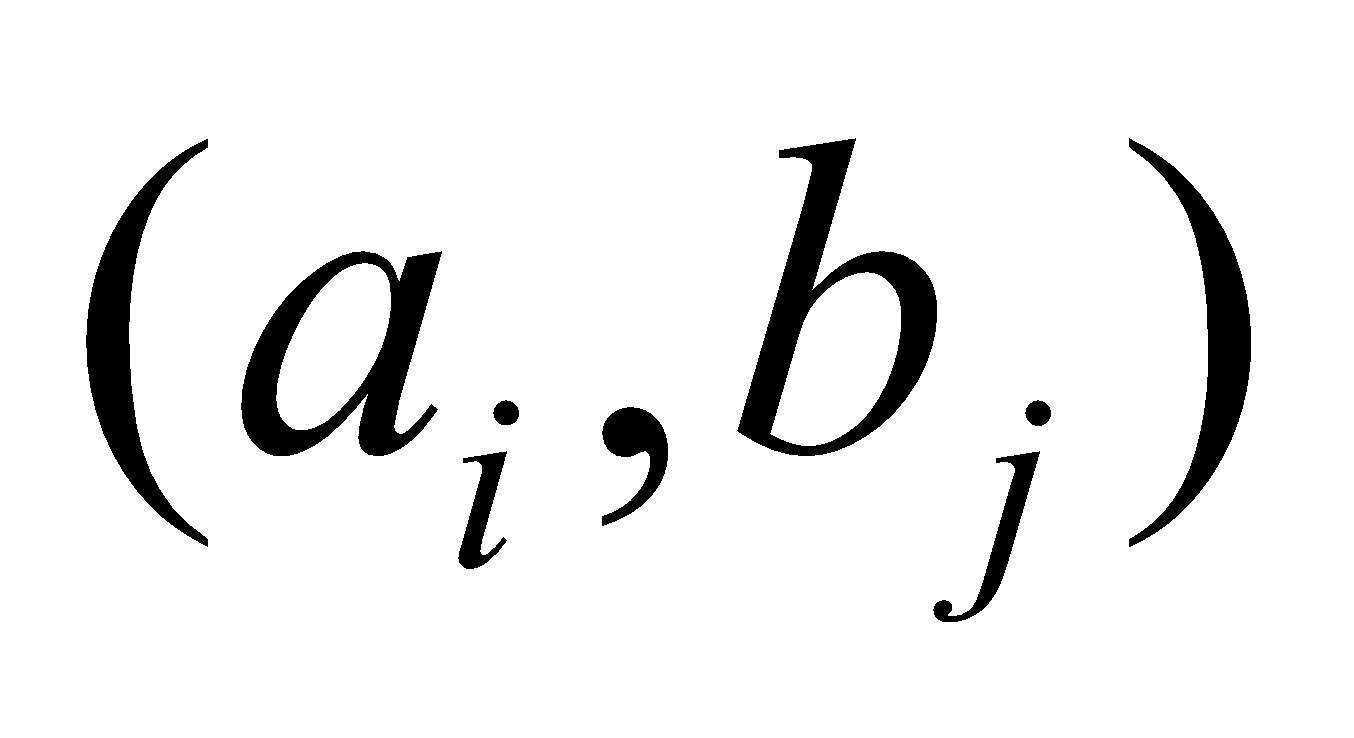
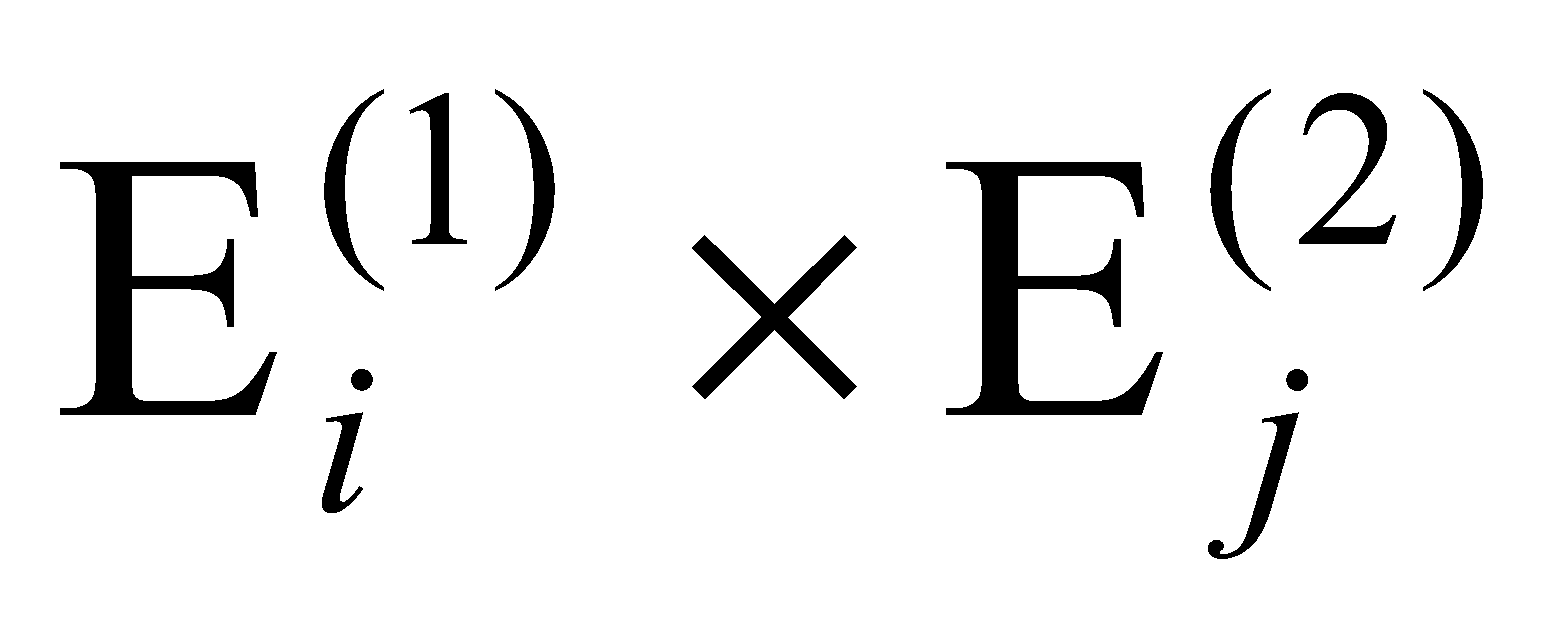
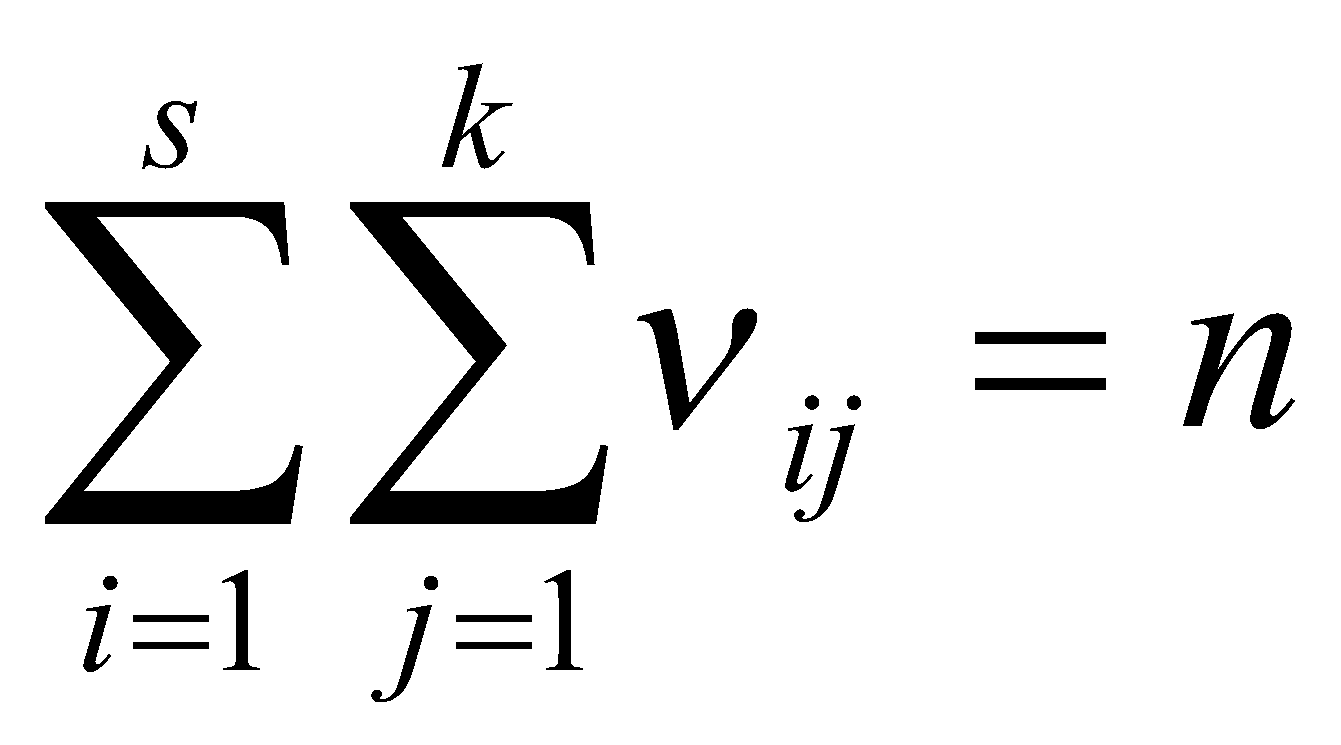
Є  пар незалежних спостережень  ,…,  випадкових величин  з невідомою функцією розподілу , для якої треба перевірити гіпотезу

: ,

де  – одновимірні функції розподілу.

Будемо припускати, що випадкова величина  приймає скінченне число значень , які будемо позначати літерами , а друга компонента  –  значень .

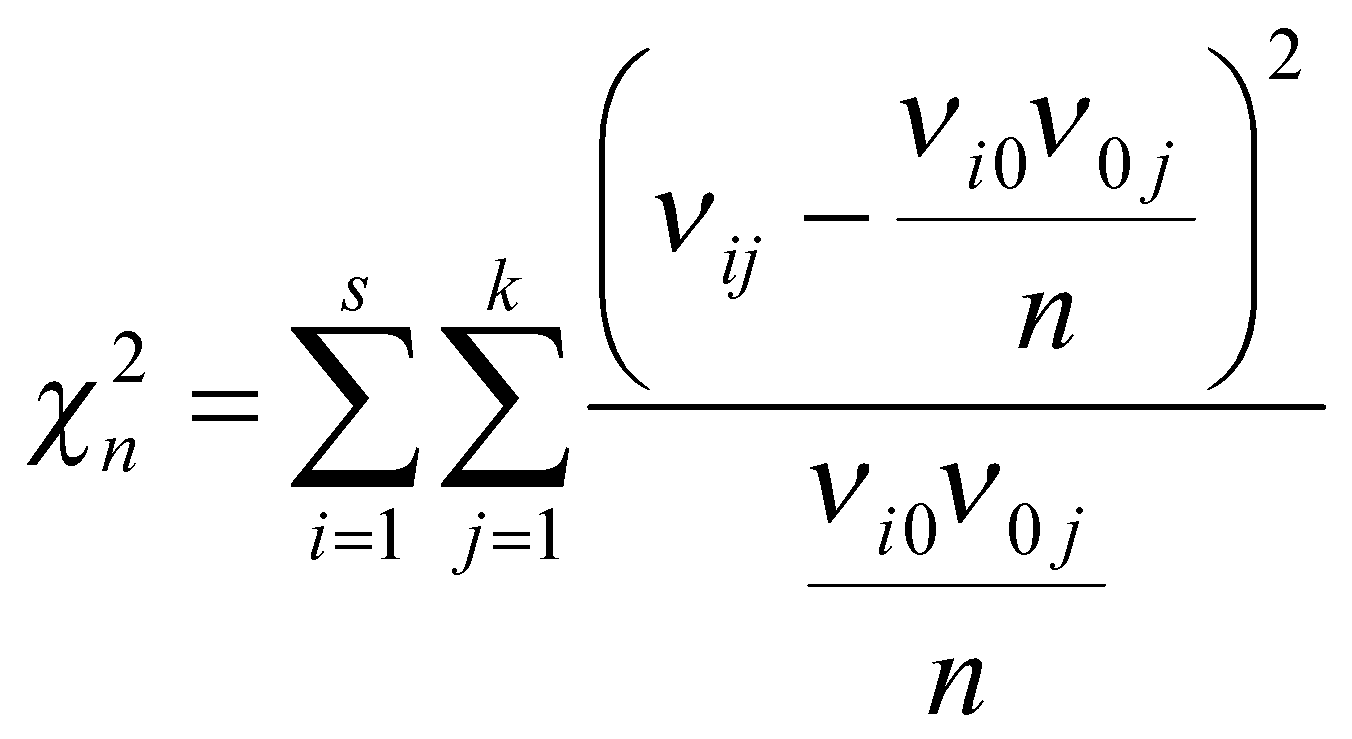
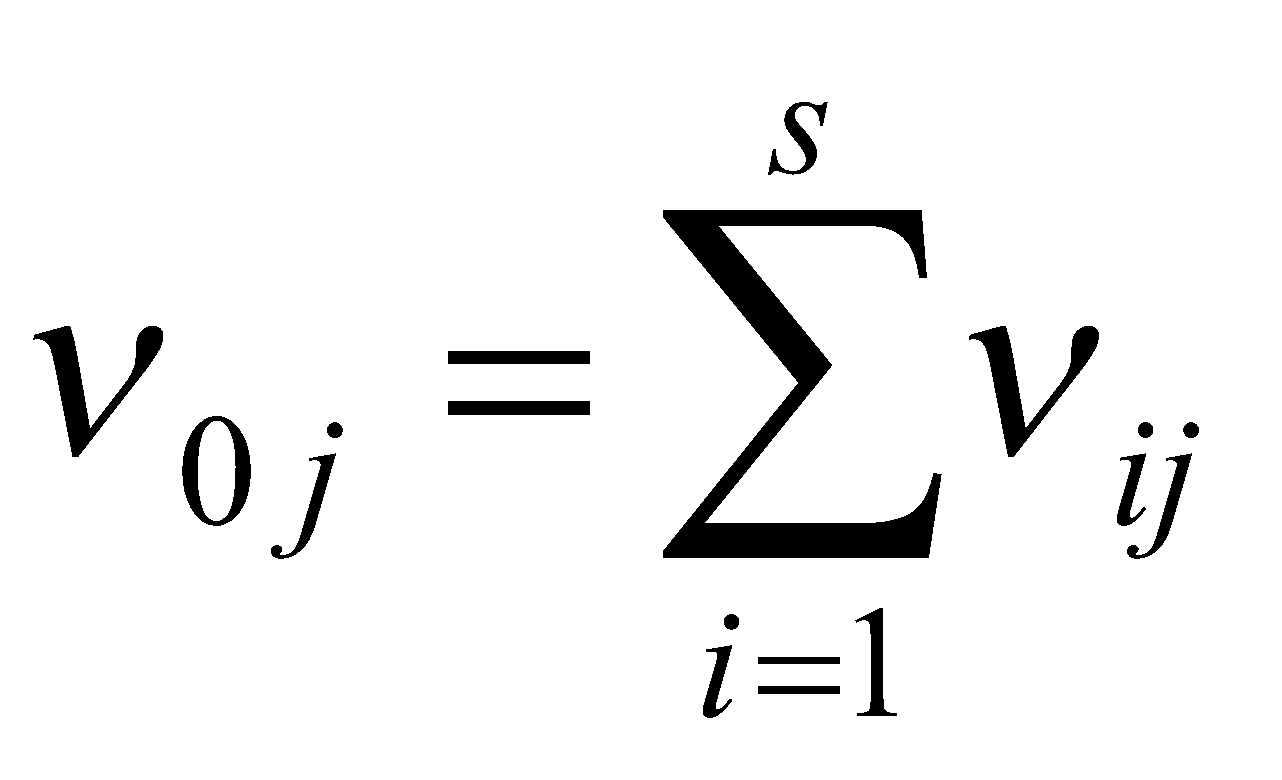
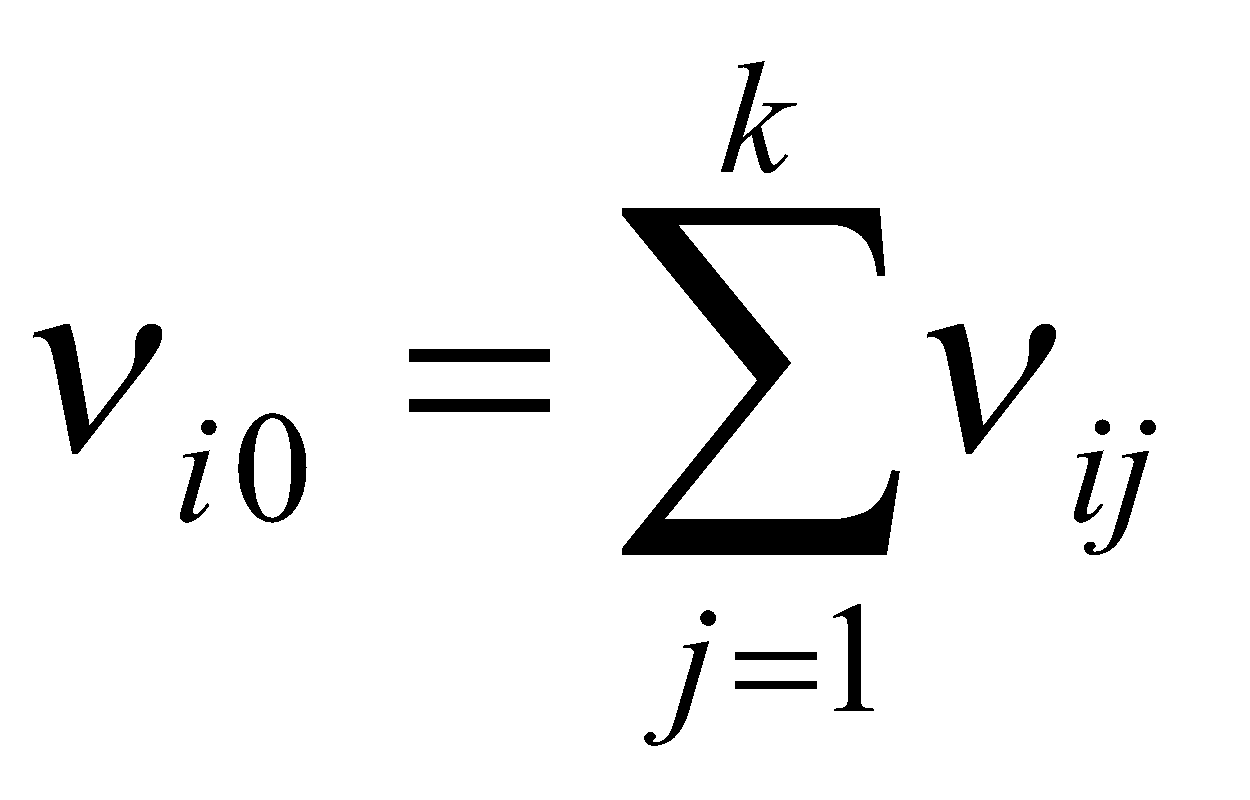
Якщо модель має іншу структуру, то групують усі можливі значення випадкових величин окремо по першій і другій компонентам. В цьому випадку множина значень  розбивається на  інтервалів , а множина значень  – на  інтервалів , множина значень вектора  – на  прямокутників .

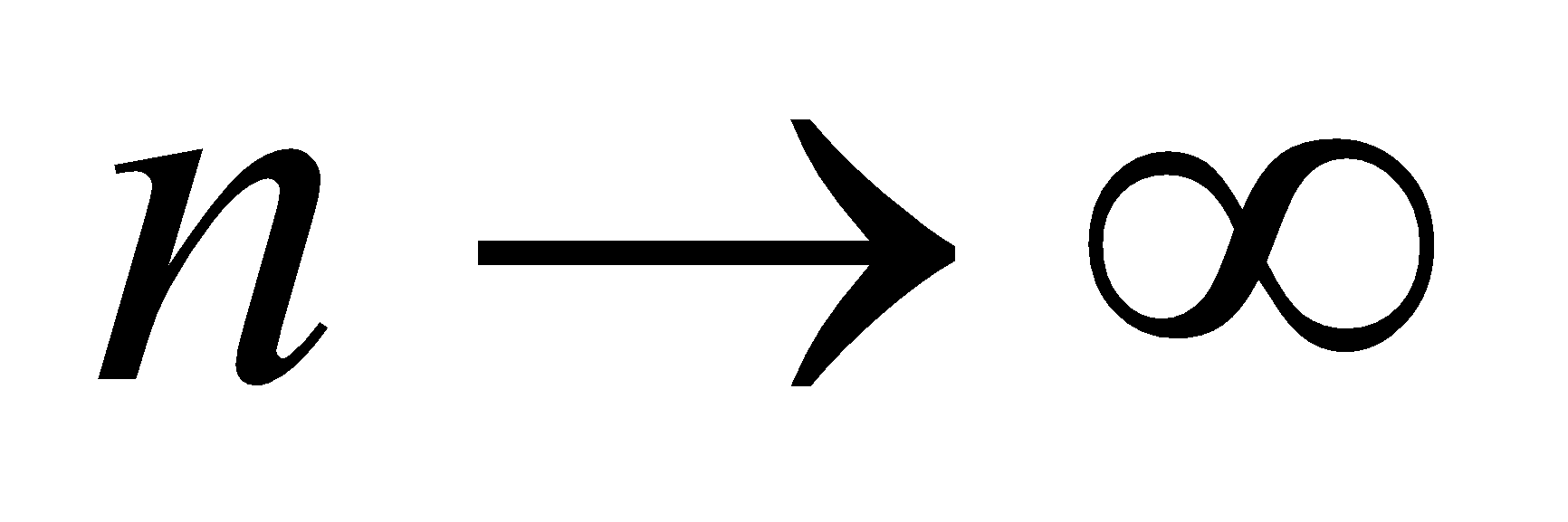
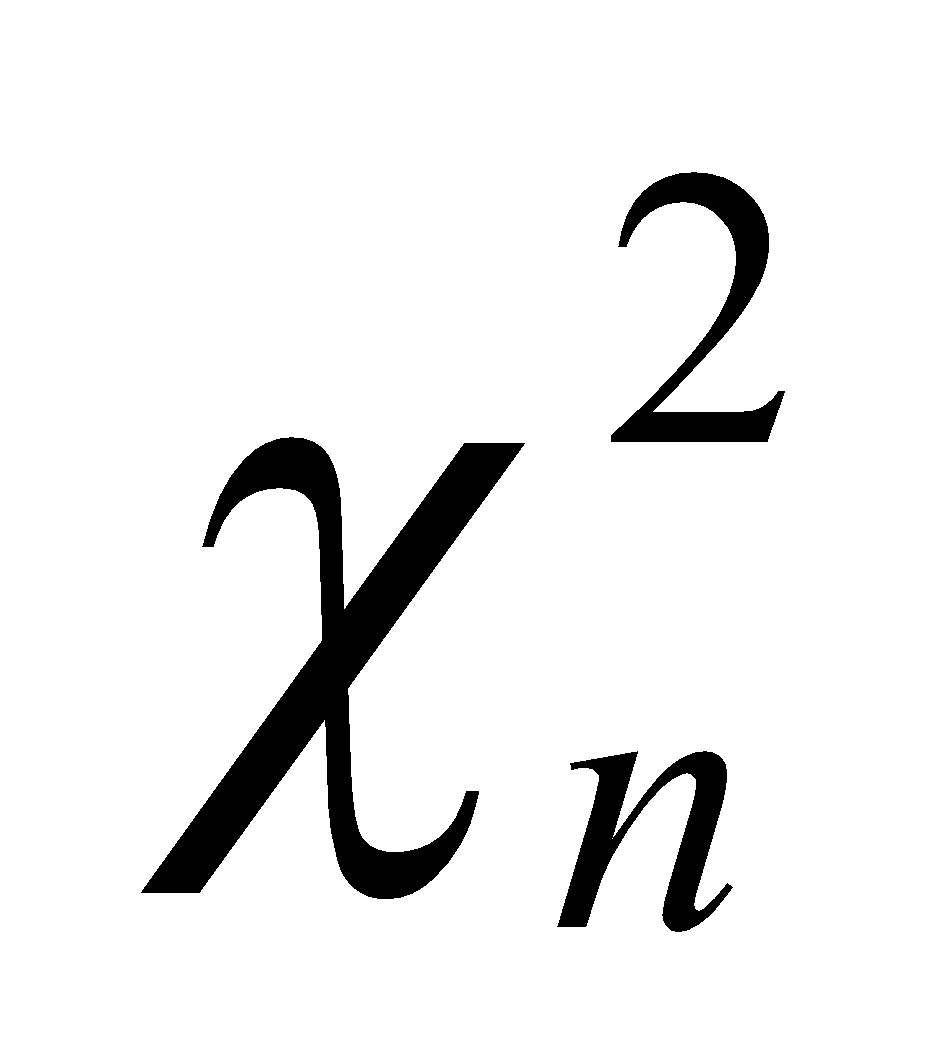
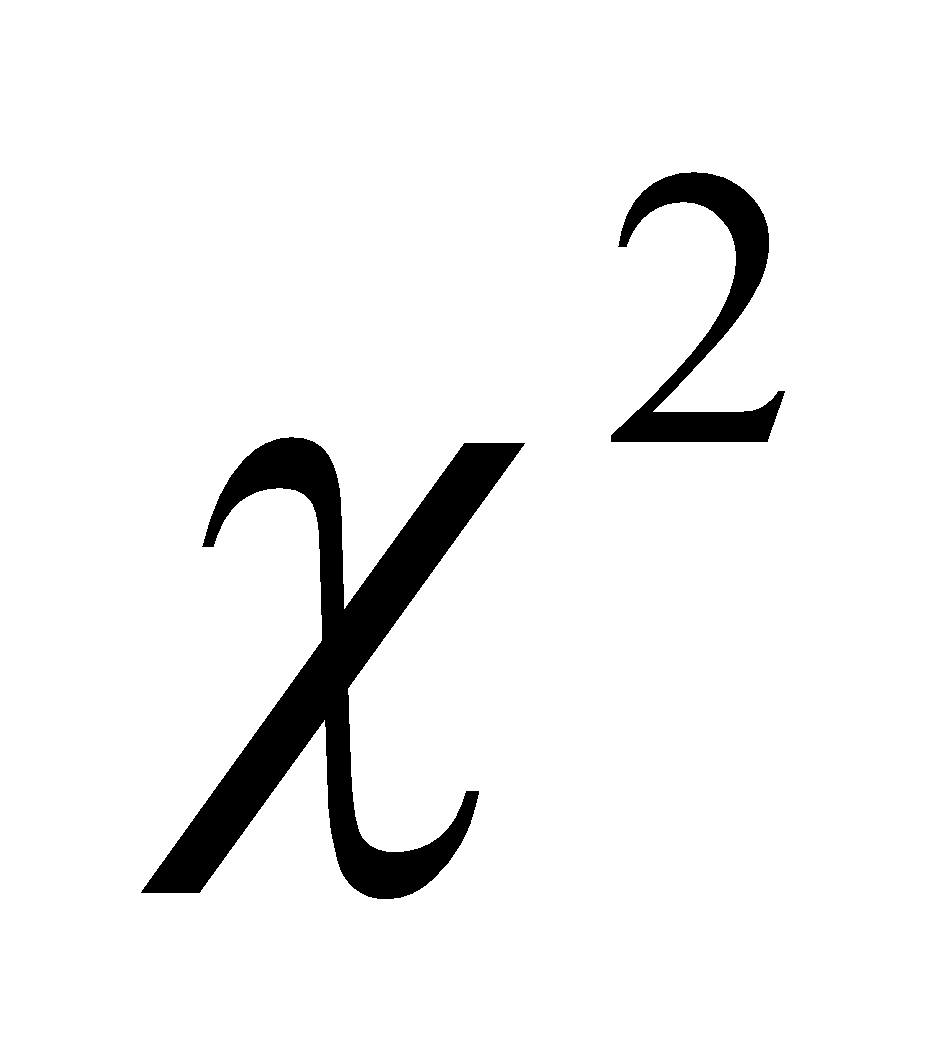
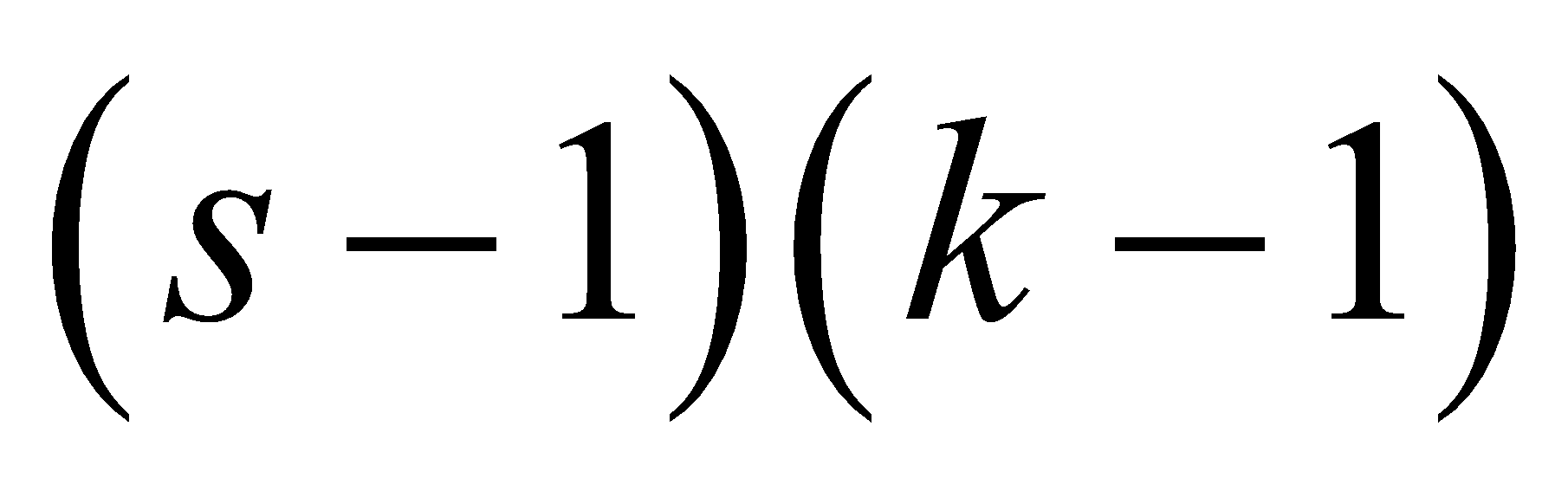
Позначимо через  число спостережень пари  (або число елементів вибірки, які належать прямокутнику , якщо дані групуються), .

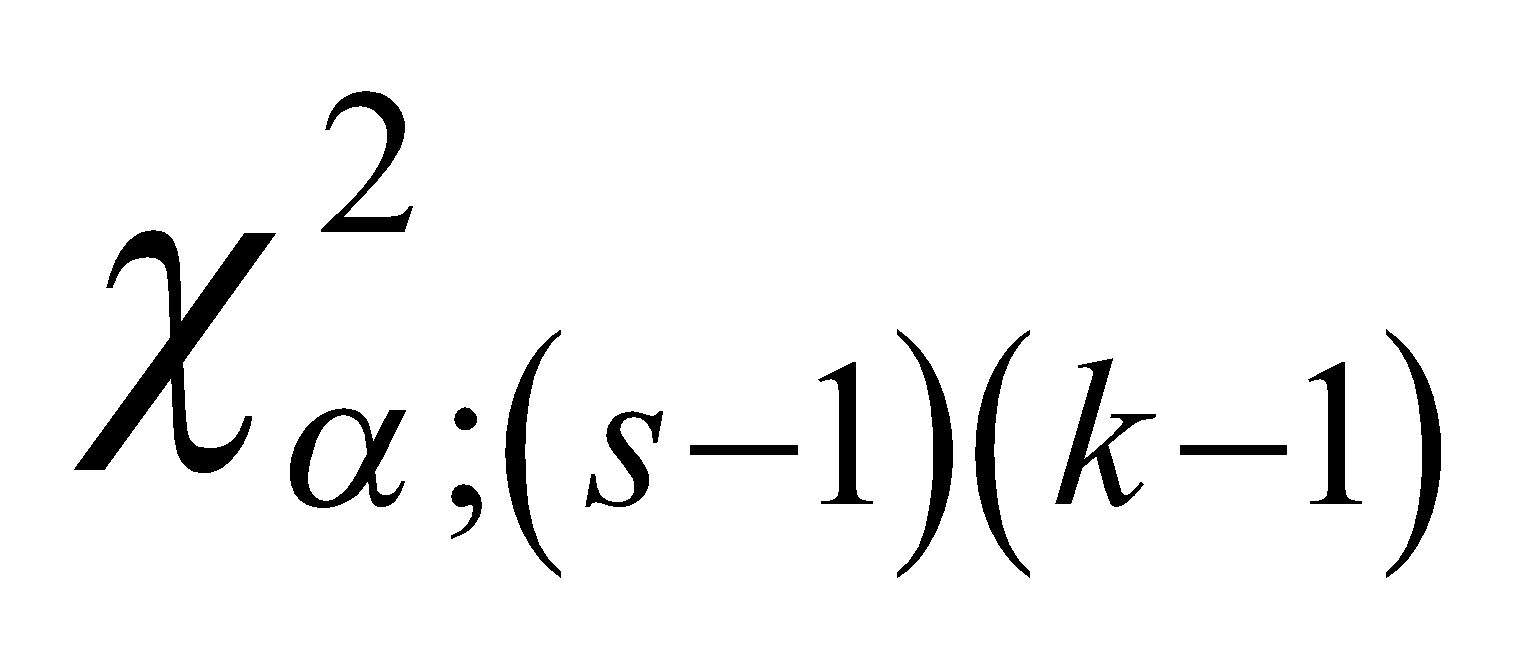
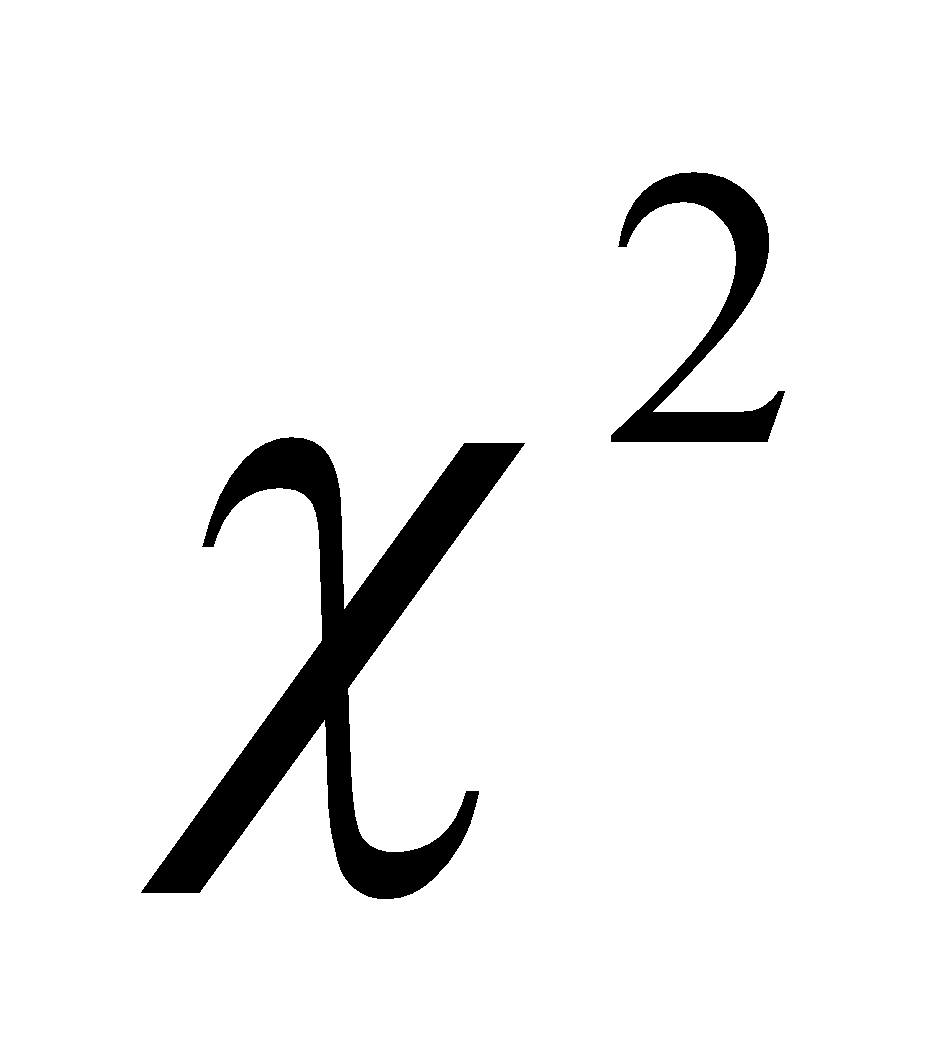
Результати спостережень зручно розташовувати у вигляді таблиці спряженості двох ознак

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  |  | | | | Сума |
|  |  |  |  |
|  |  |  |  |  |  |
|  |  |  |  |  |  |
|  |  |  |  |  |  |
|  |  |  |  |  |  |
| Сума |  |  |  |  |  |

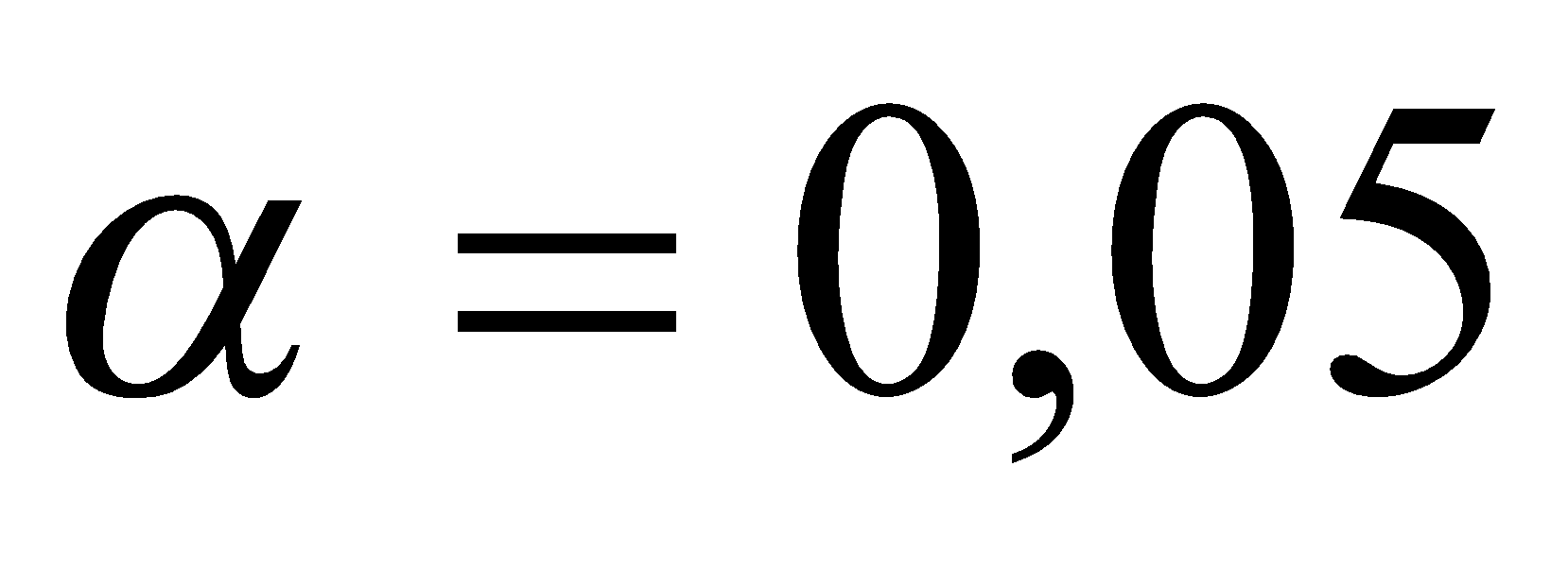
Відстань між емпіричними даними і їх гіпотетичними значеннями має вигляд

, , .

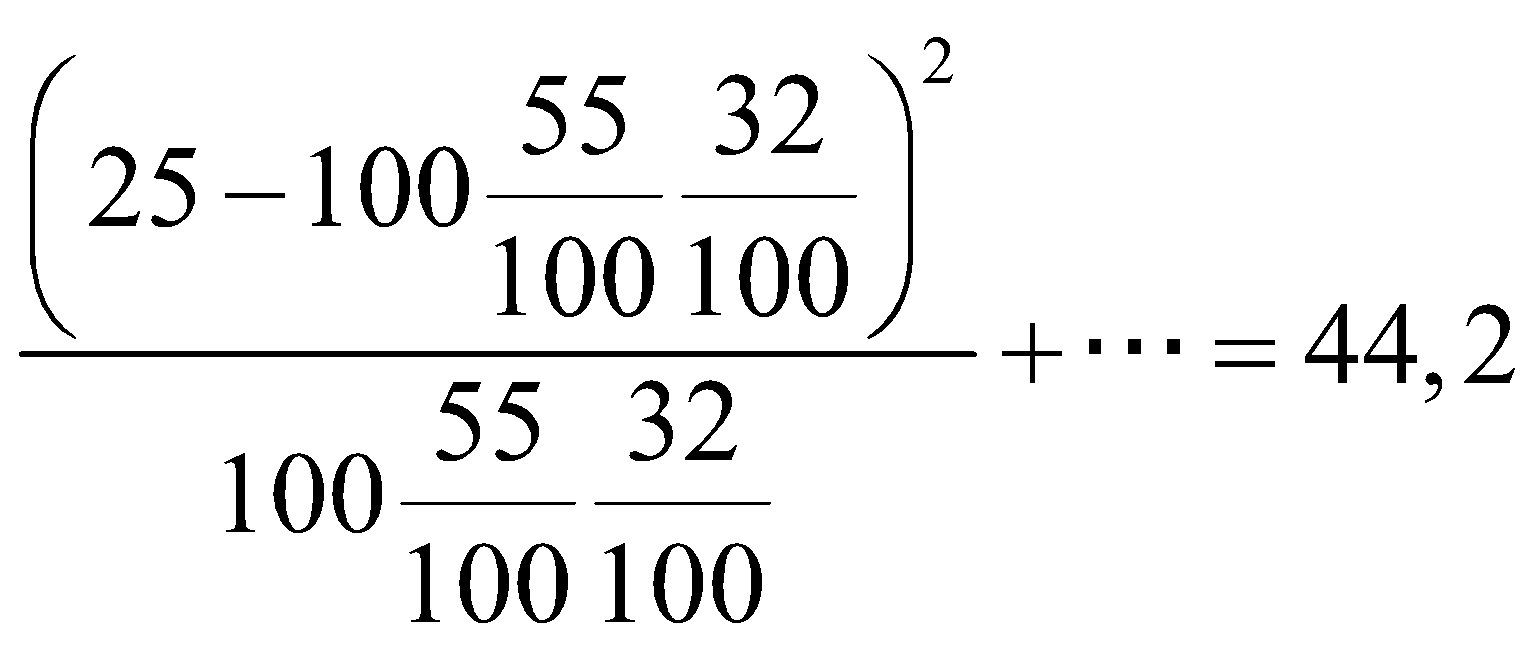
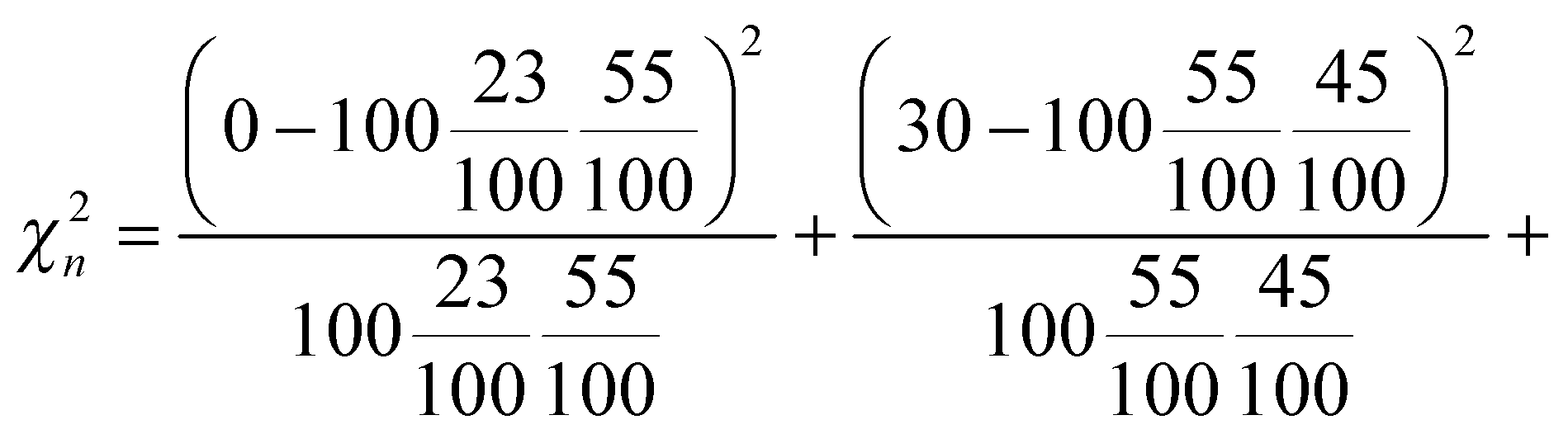
При  розподіл відхилення **** збігається до ****– розподілу з  степенями свободи.

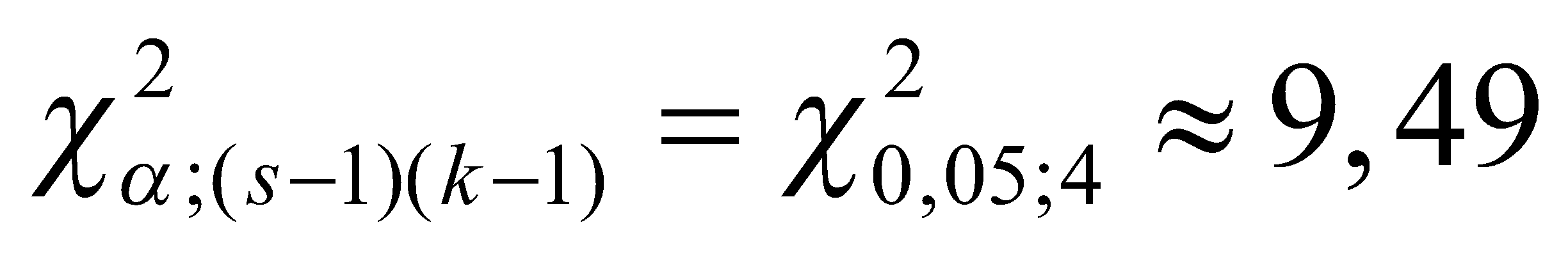
Вибір табличного значення і прийняття рішення про допустимість гіпотези проводиться аналогічно описаній вище процедурі для критерію однорідності ****. При цьому з імовірністю **** гіпотеза **** буде відхилятися, коли вона вірна.

**Приклад 4.7.** Нижче наведені результати опитування 100 студентів перших трьох курсів, яким ставилося одне питання: “Чи вважаєте ви, що куріння заважає навчанню?”

З’ясувати, чи підтведжують ці дані припущення про те, що відношення до куріння студентів на різних курсах різне? Прийняти .

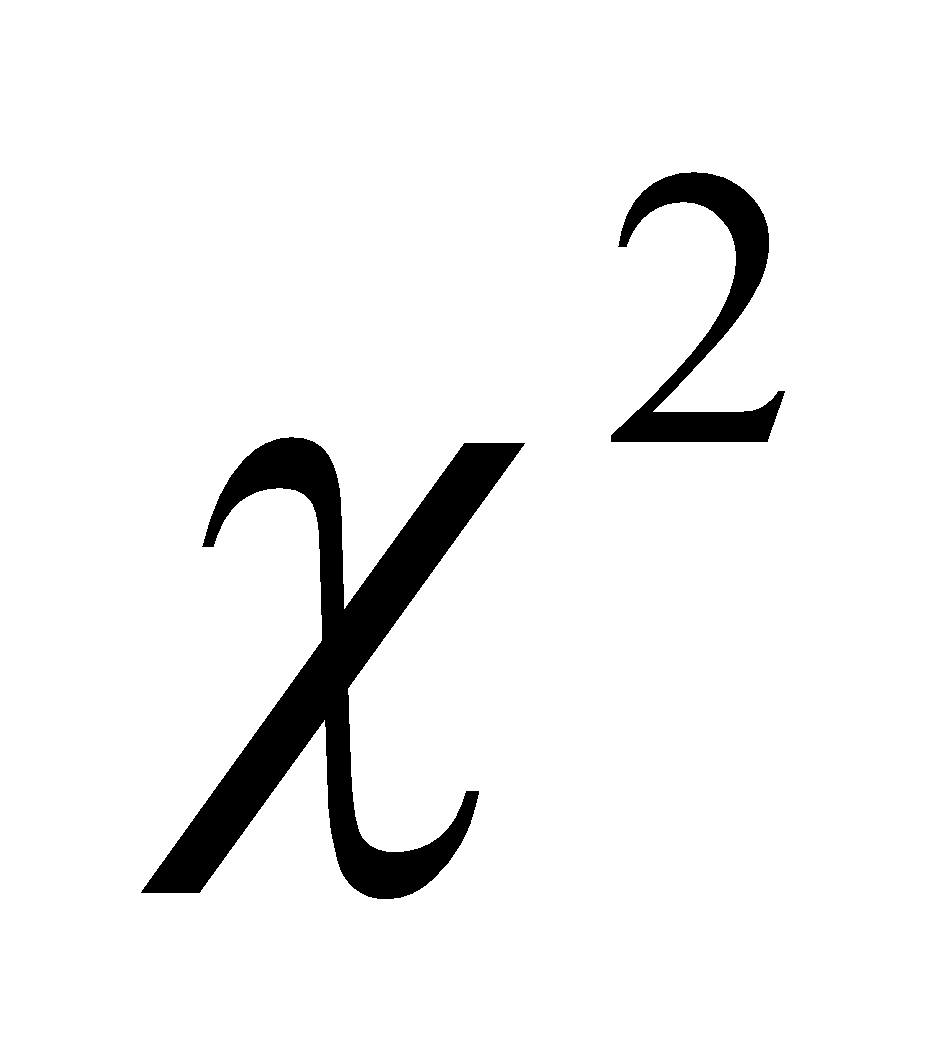
|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Відповідь | Курс | | |  | |
| 1 | 2 | 3 |
| Так | - | 30 | 25 | 55 |  |
| Не знаю | 8 | 5 | 7 | 20 |  |
| Ні | 15 | 10 | - | 25 |  |
|  | 23 | 45 | 32 | 100 | |
|  |  |  |

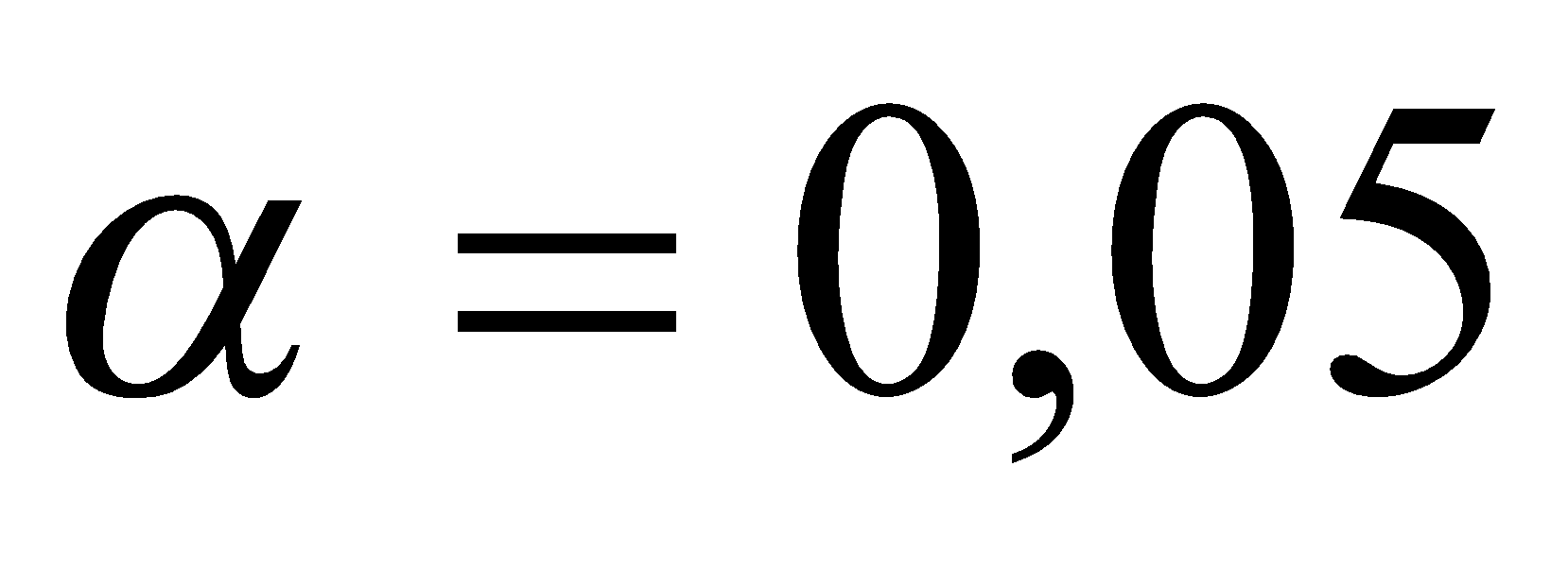


.

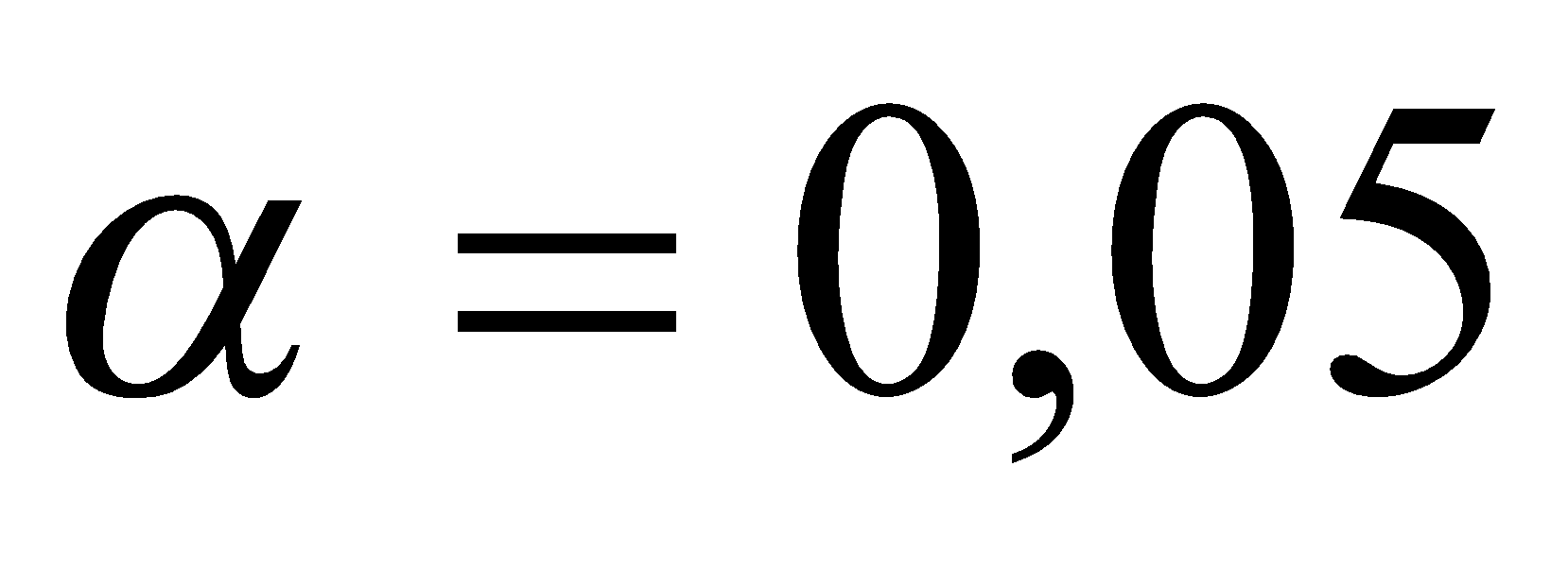
Гіпотеза про незалежність відкидається і сформульована теза приймається.

**ЗАДАЧІ**

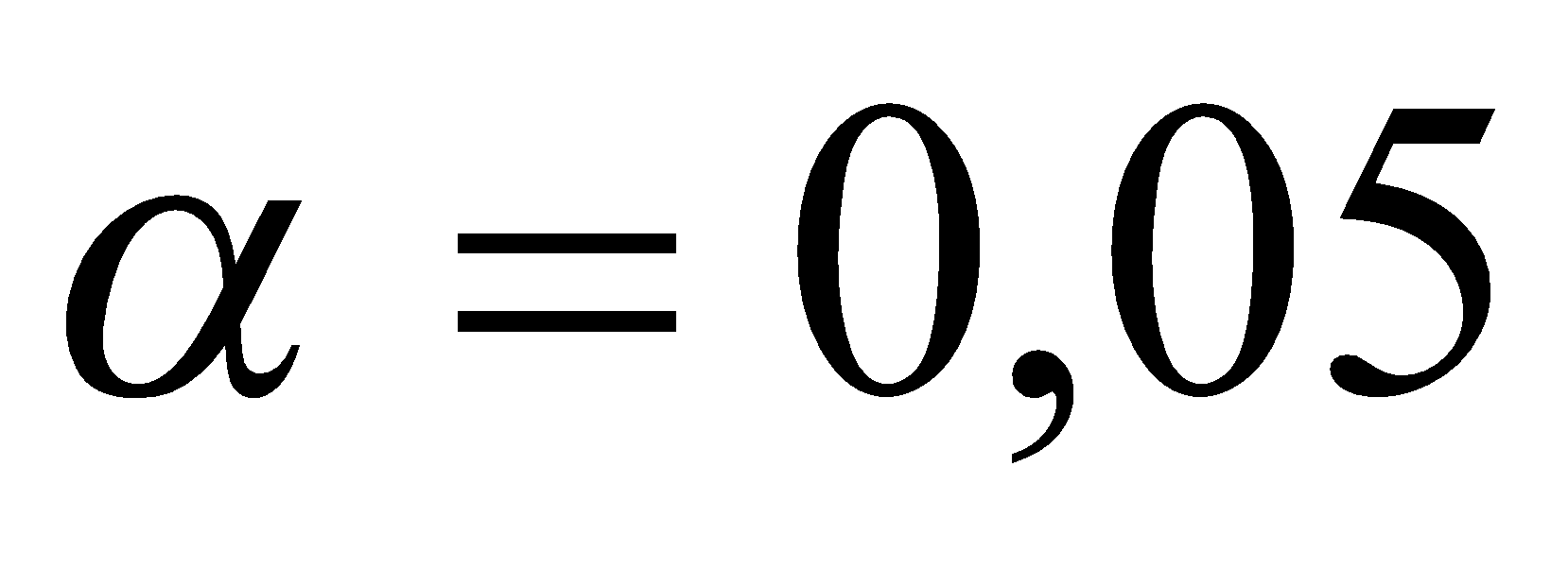
**4.1.** За спостереженнями, наведеними у таблиці, за допомогою критерію **** перевірити гіпотезу, що випадкова величина має пуассонівський розподіл.

а) ,

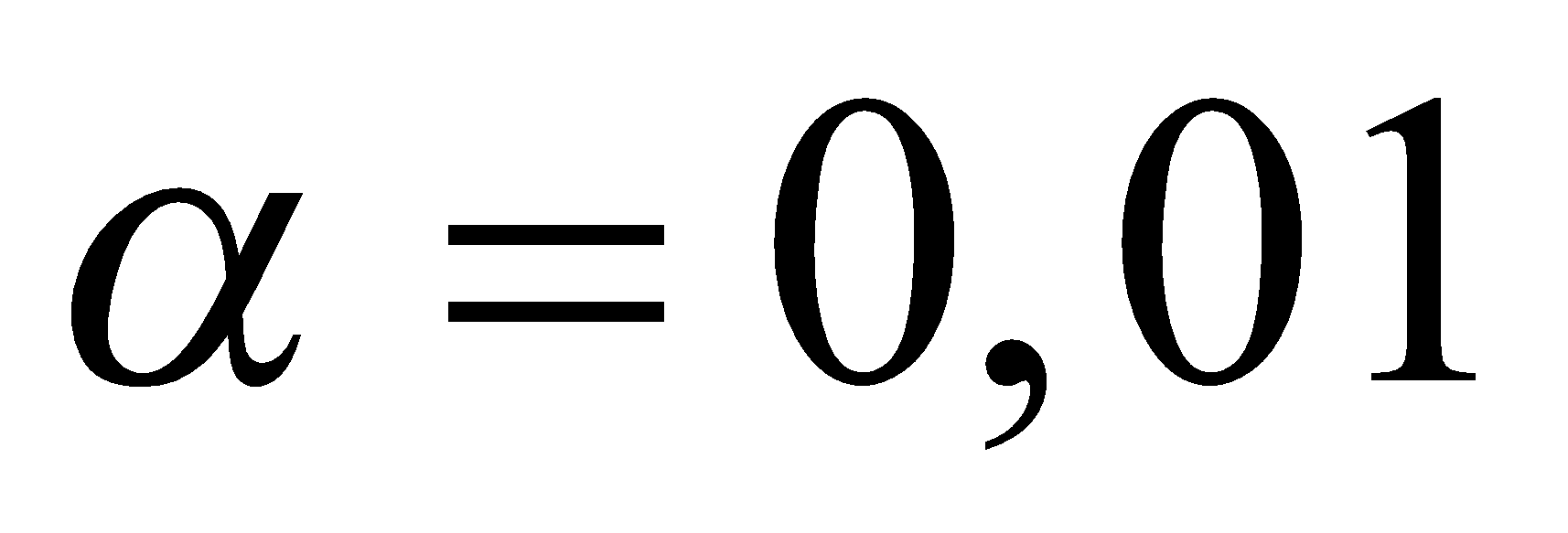
|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | 0 | 1 | 2 | 3 | 4 |
|  | 109 | 65 | 22 | 3 | 1 |

б) ,

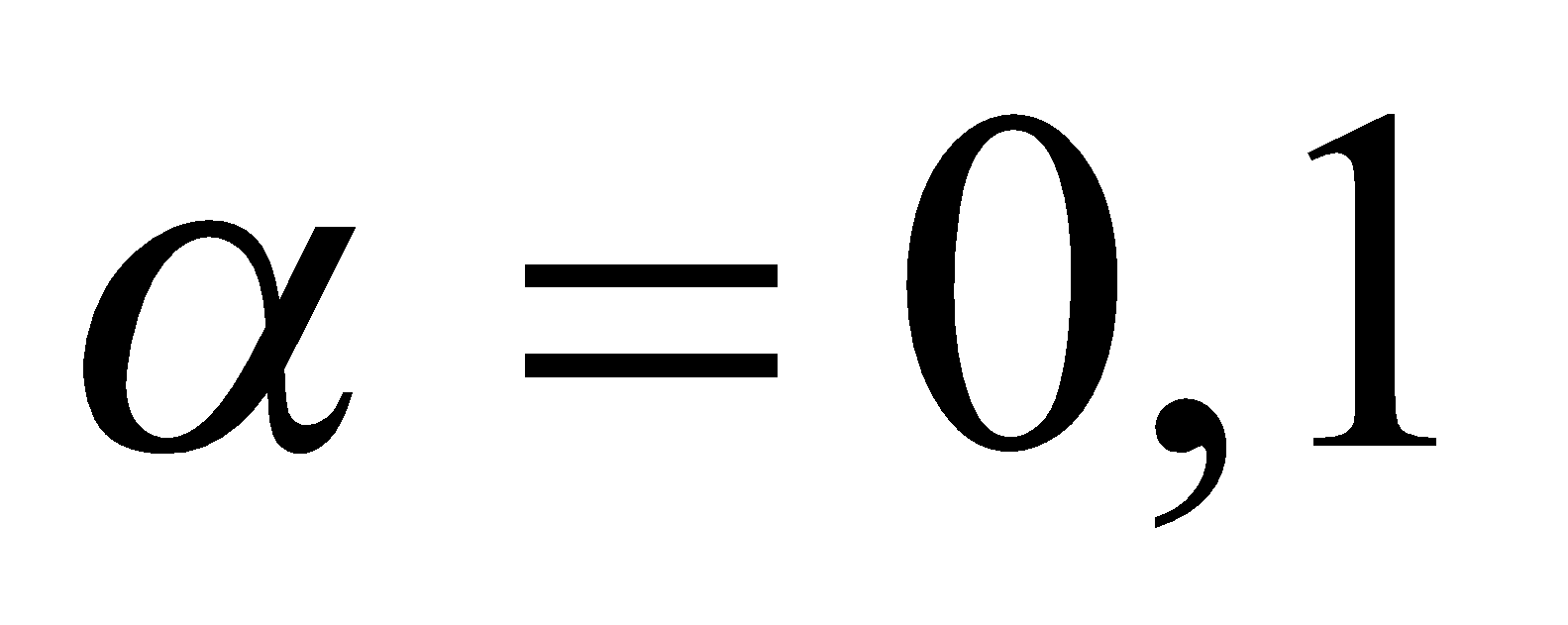
|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
|  | 112 | 168 | 130 | 68 | 32 | 5 | 1 | 1 |

в) ,

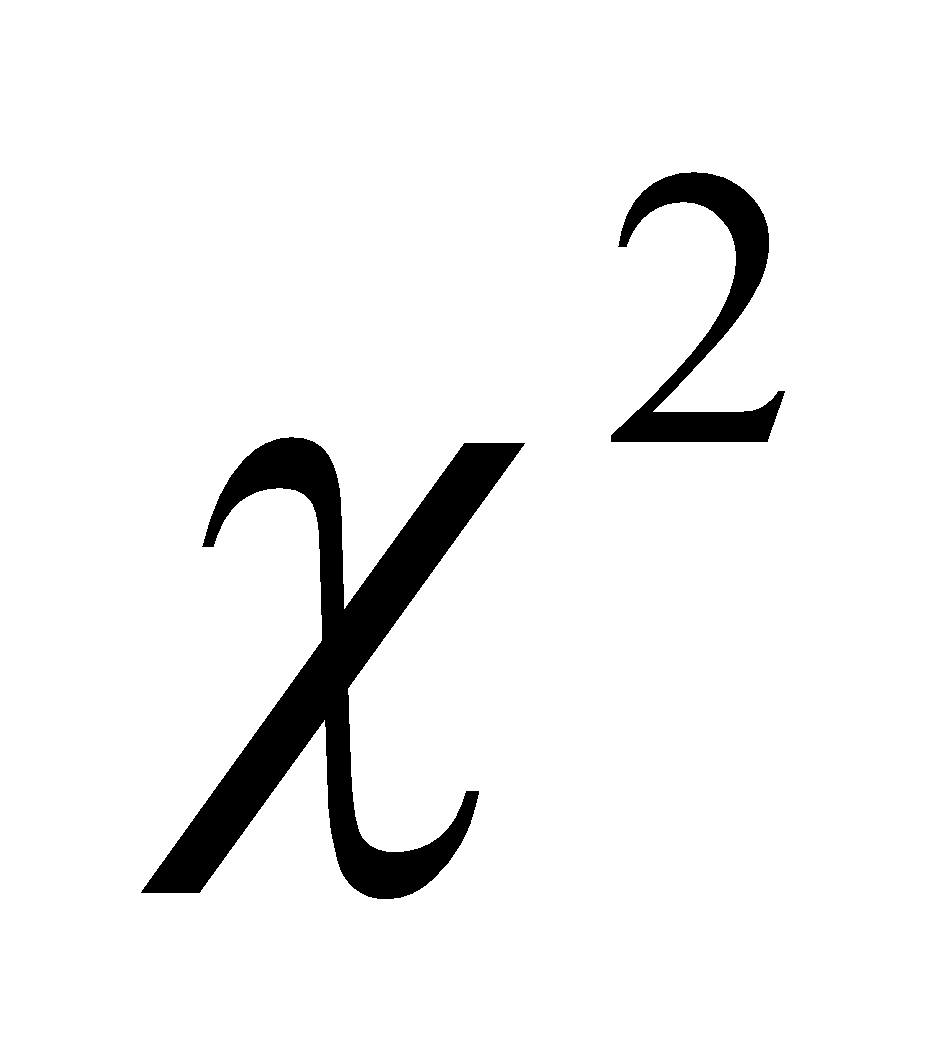
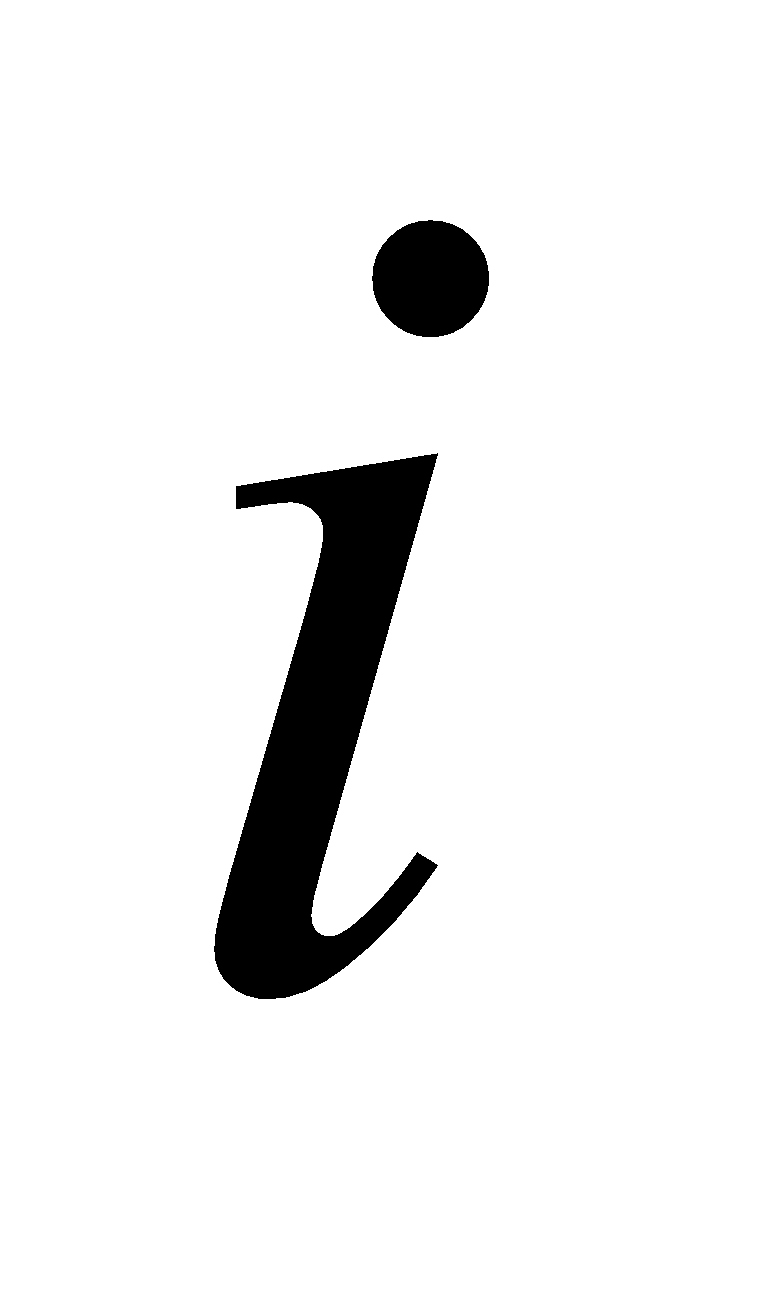
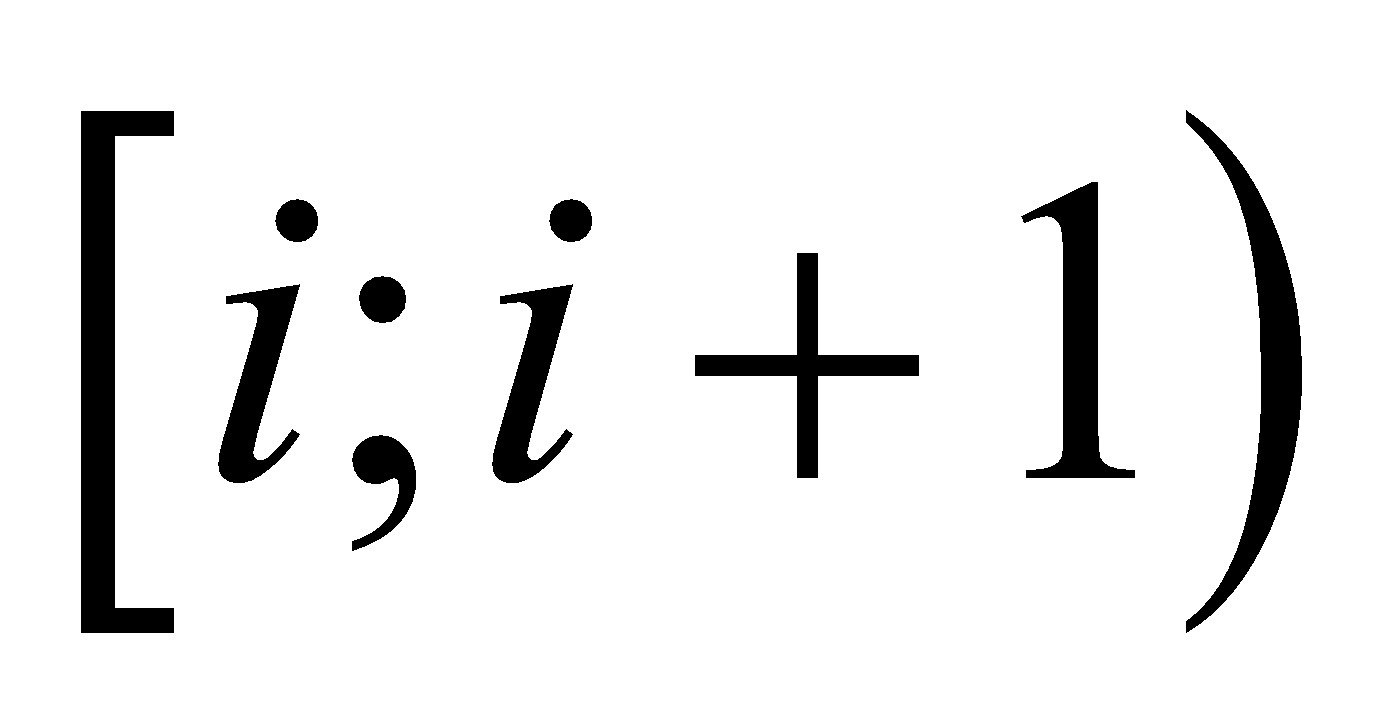
|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | 0 | 1 | 2 | 3 | 4 | 5 |
|  | 229 | 211 | 93 | 35 | 7 | 1 |

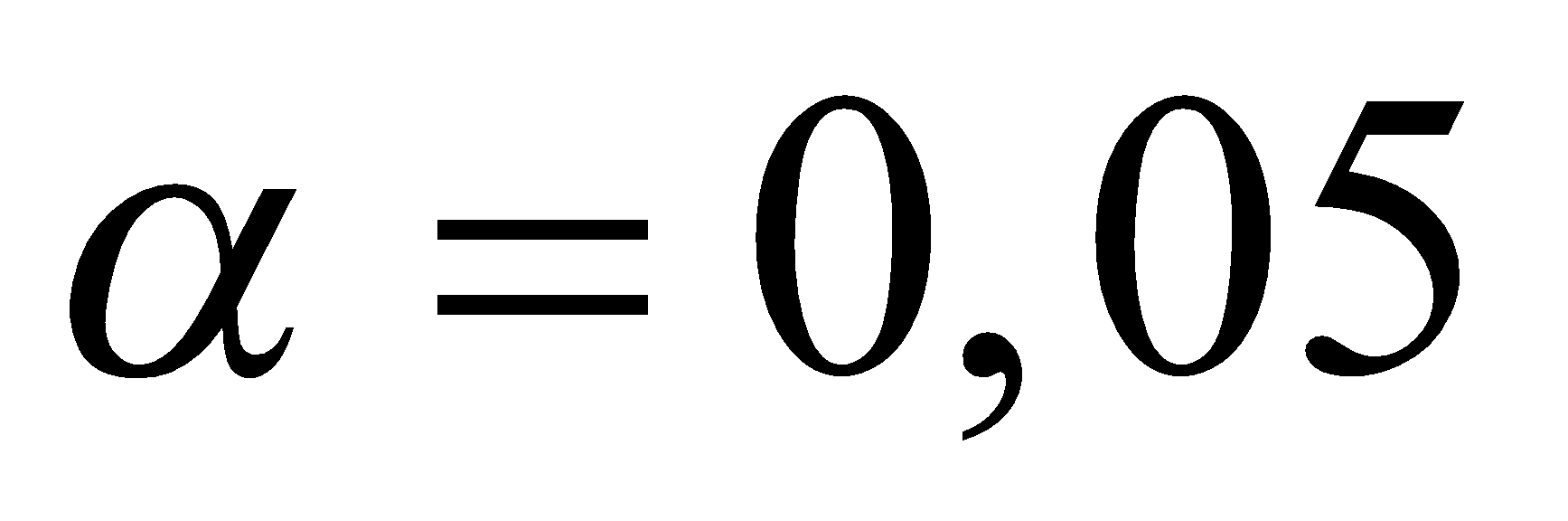
г) ,

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
|  | 8 | 17 | 16 | 10 | 6 | 2 | 0 | 1 |

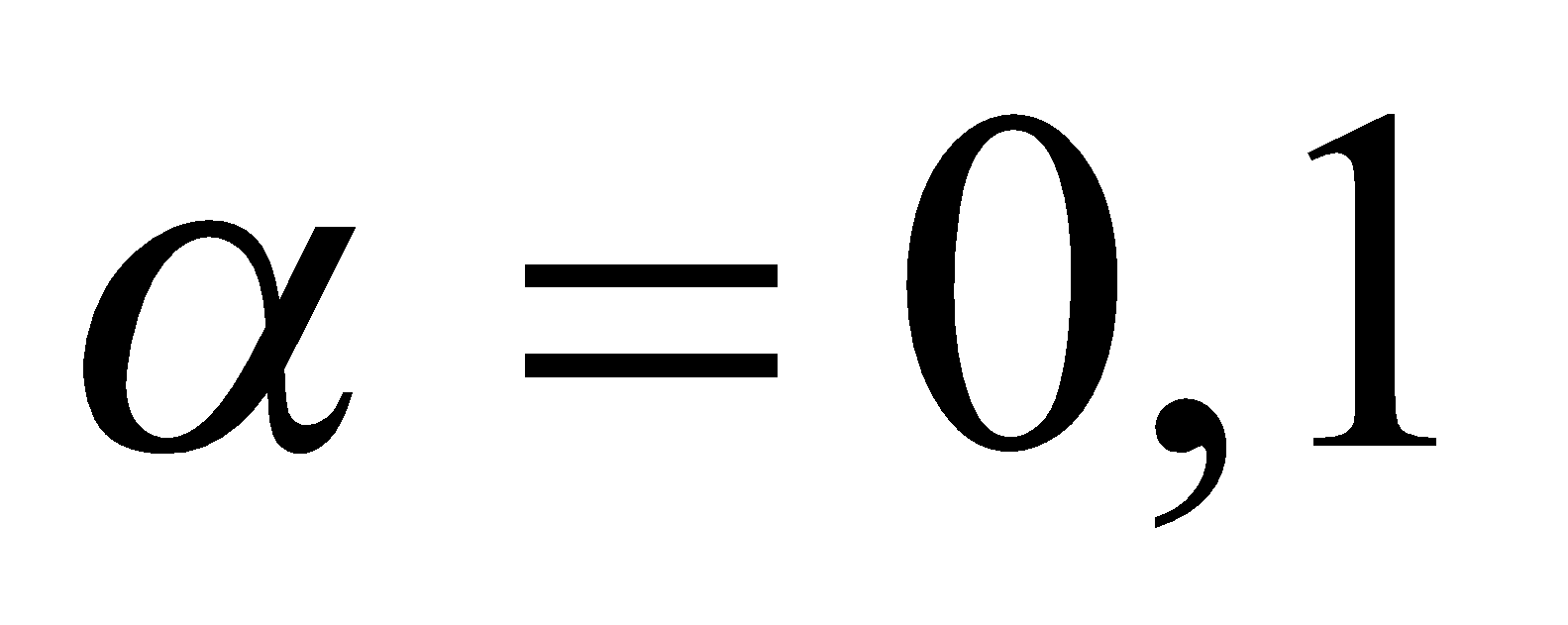
д) ,

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | 0 | 1 | 2 | 3 | 4 | 5 |
|  | 376 | 100 | 81 | 35 | 7 | 1 |

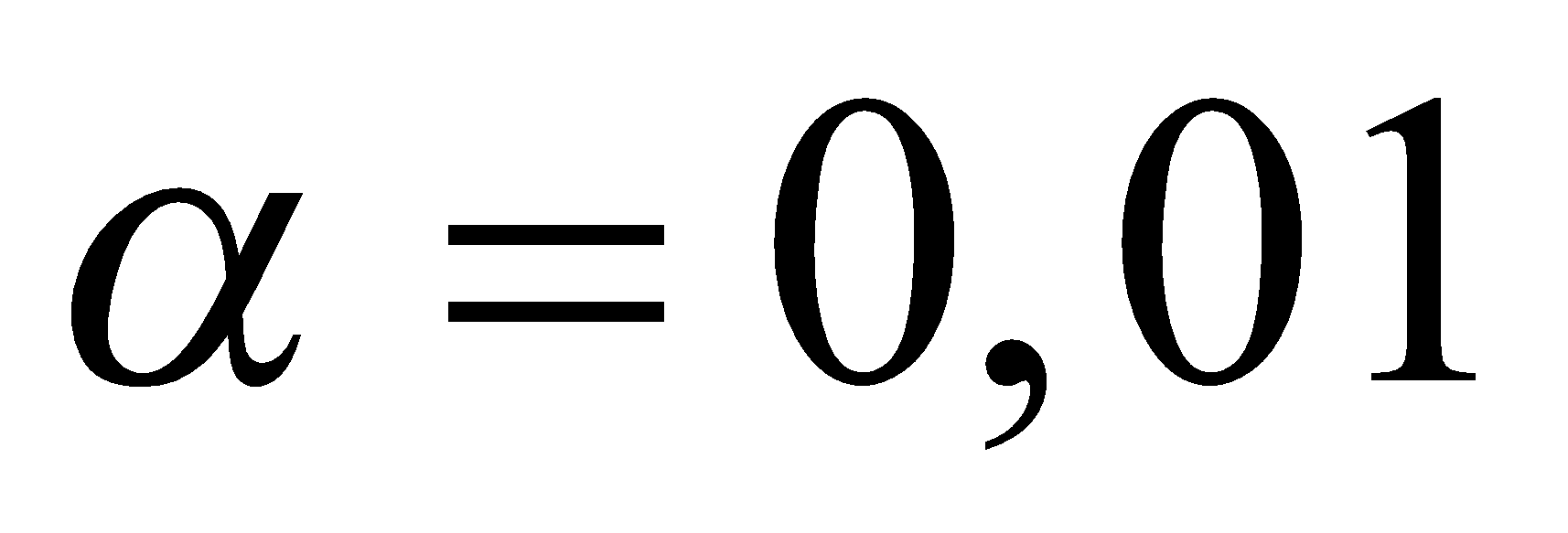
**4.2.** За спостереженнями, наведеними у таблиці, за допомогою критерію **** перевірити згоду з рівномірним розподілом. У першому рядку таблиці вказана ліва границя інтервалу ( – номер інтервалу ).

а) ,

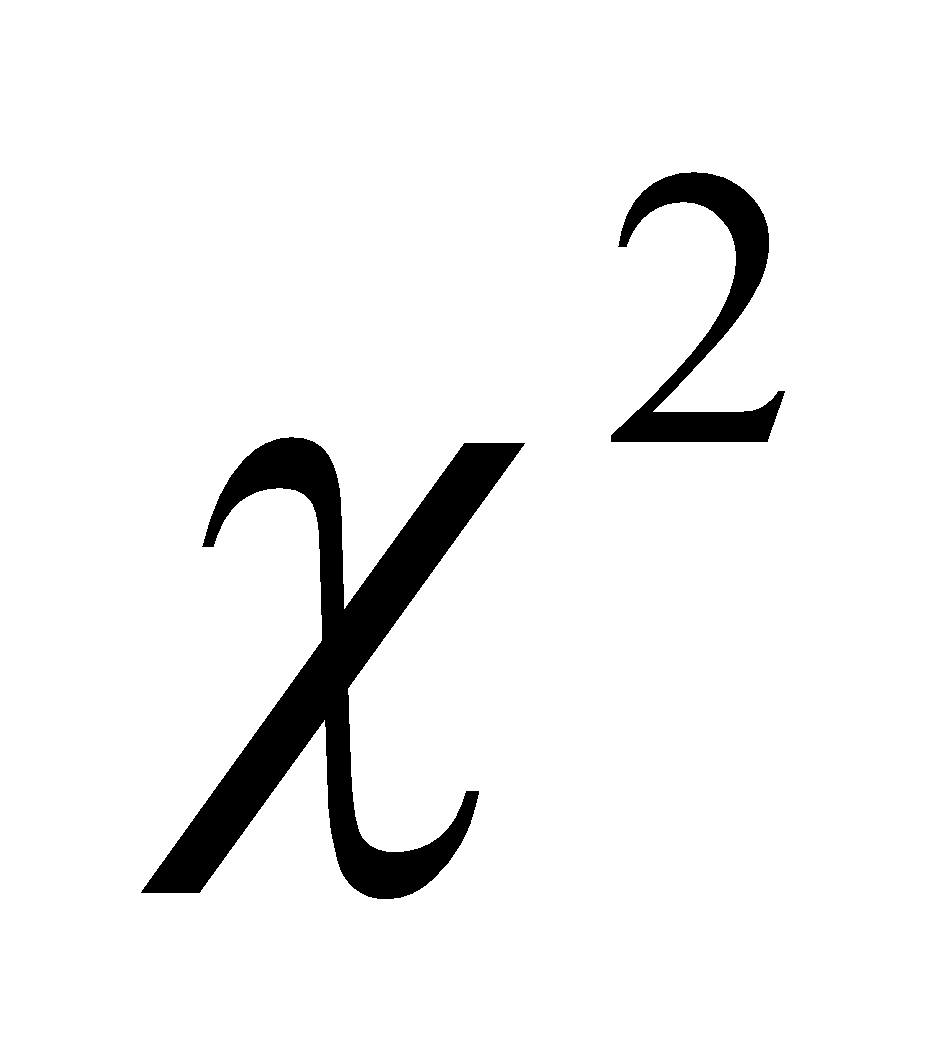
|  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 |
|  | 48 | 42 | 36 | 54 | 39 | 43 | 41 | 33 | 37 | 41 | 47 | 39 |

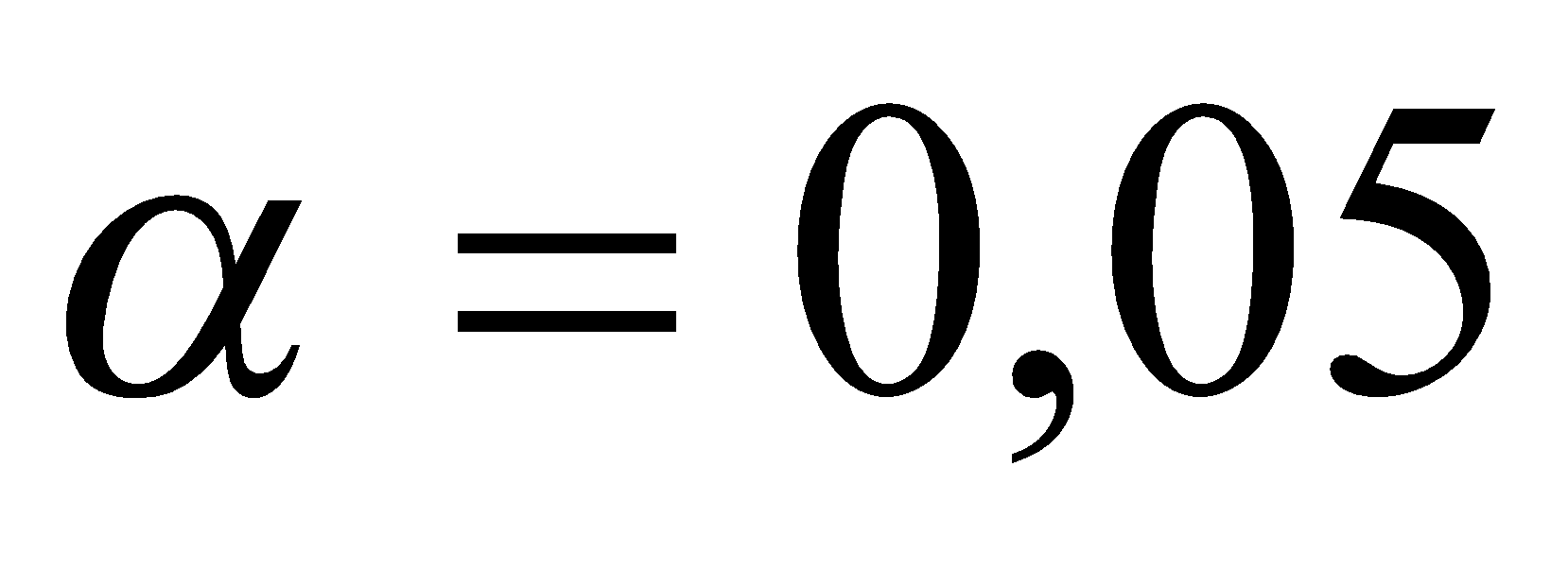
б) ,

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |
|  | 69 | 89 | 83 | 79 | 80 | 73 | 77 | 75 | 76 | 91 |

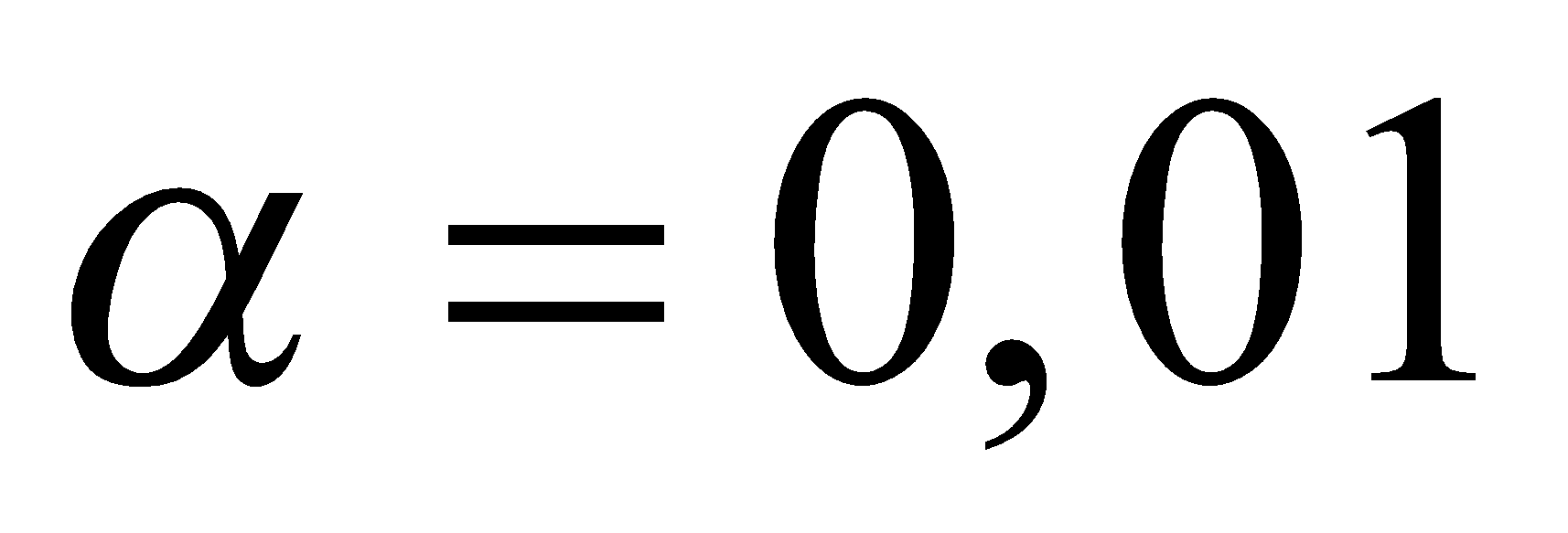
в) ,

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |
|  | 16 | 15 | 19 | 13 | 14 | 19 | 14 | 12 | 17 | 13 |

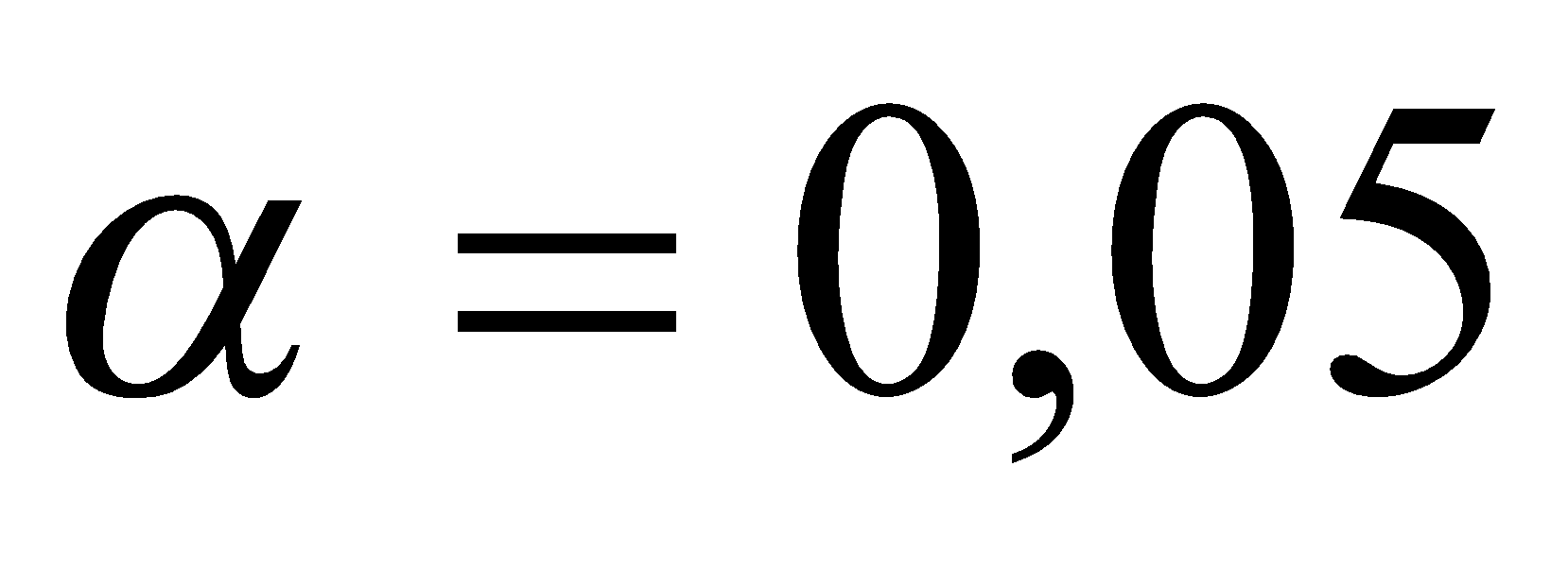
**4.3.** По спостереженням, наведеним у таблиці, за допомогою критерію **** перевірити згоду з нормальним розподілом.

а) ,

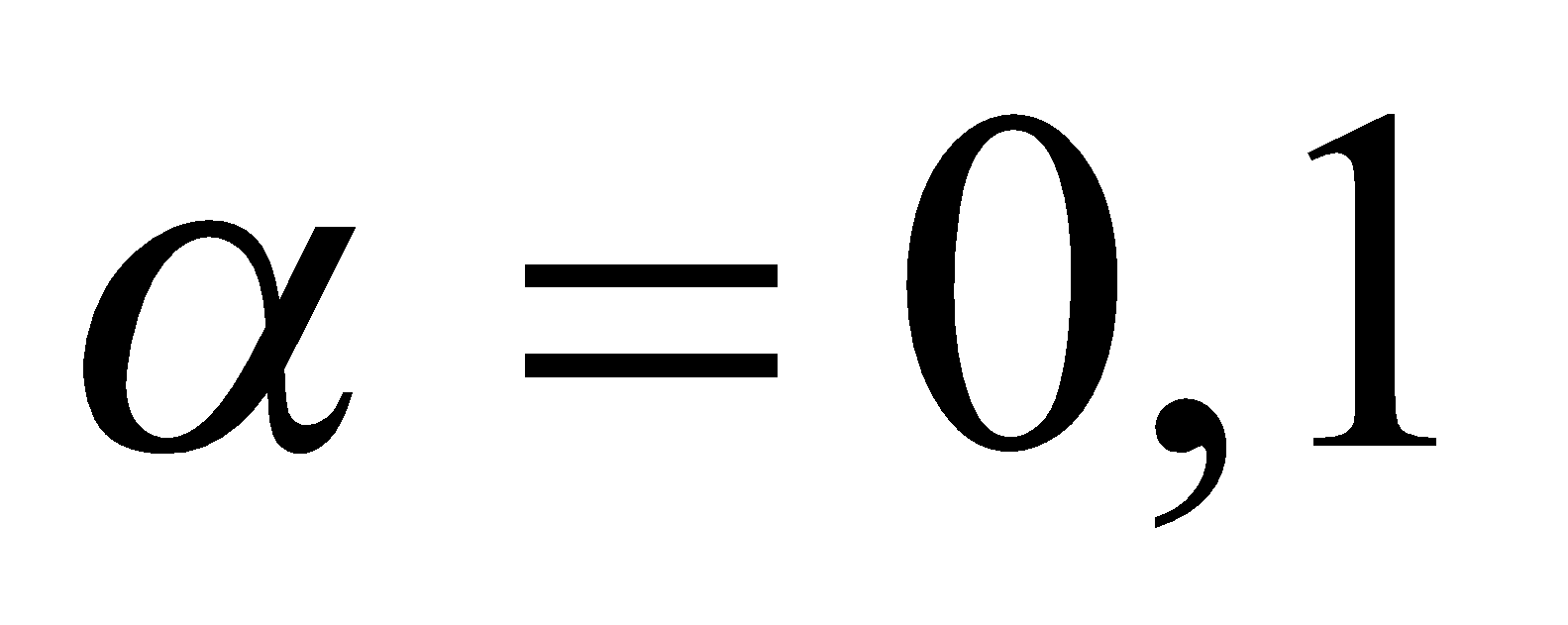
|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Інтервал | [0;5) | [5;10) | [10;15) | [15;20) | [20;25) |
|  | 15 | 75 | 100 | 50 | 10 |

б) ,

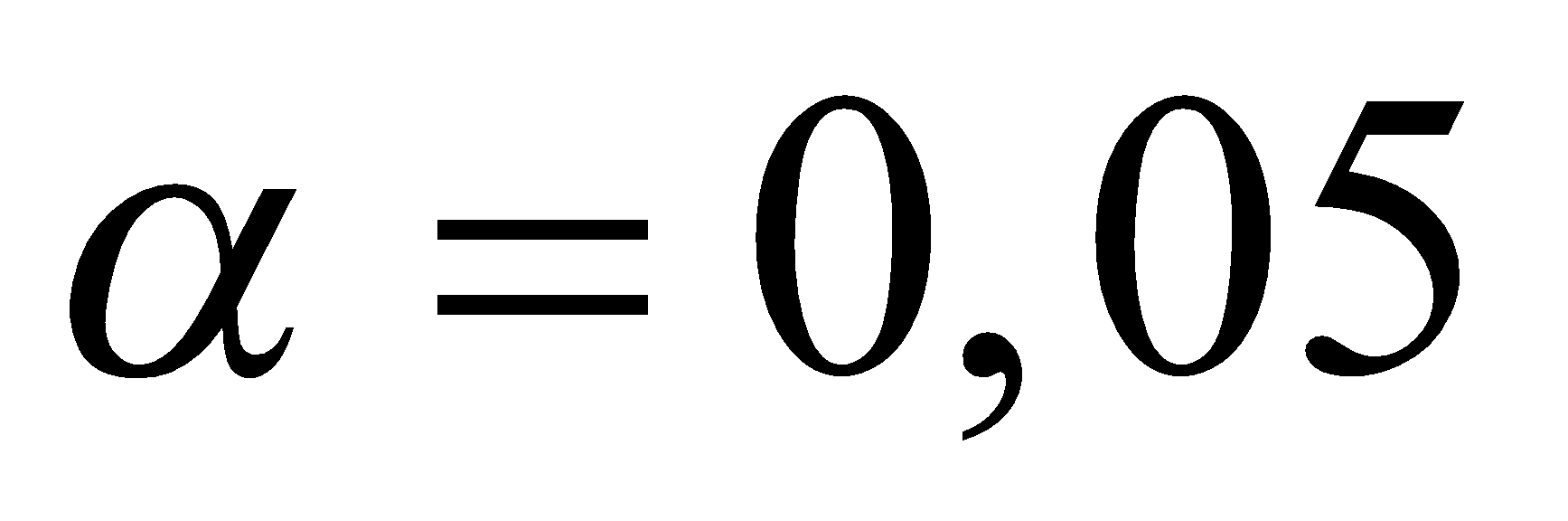
|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Інтервал | [3,0;3,6) | [3,6;4,2) | [4,2;4,8) | [4,8;5,4) | [5,4;6,0) | [6,0;6,6) |
|  | 2 | 8 | 35 | 43 | 22 | 10 |

в) ,

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Інтервал | [-3;-1) | [-1;0) | [0;1) | [1;2) | [2;3) | [3;5) |
|  | 13 | 15 | 24 | 25 | 13 | 10 |

г) ,

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Інтервал | [-8;-2) | [-2;4) | [4;10) | [10;16) |
|  | 10 | 50 | 30 | 10 |

г) ,

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Інтервал | [-4;0) | [0;2) | [2;4) | [4; 6) |
|  | 25 | 40 | 25 | 10 |

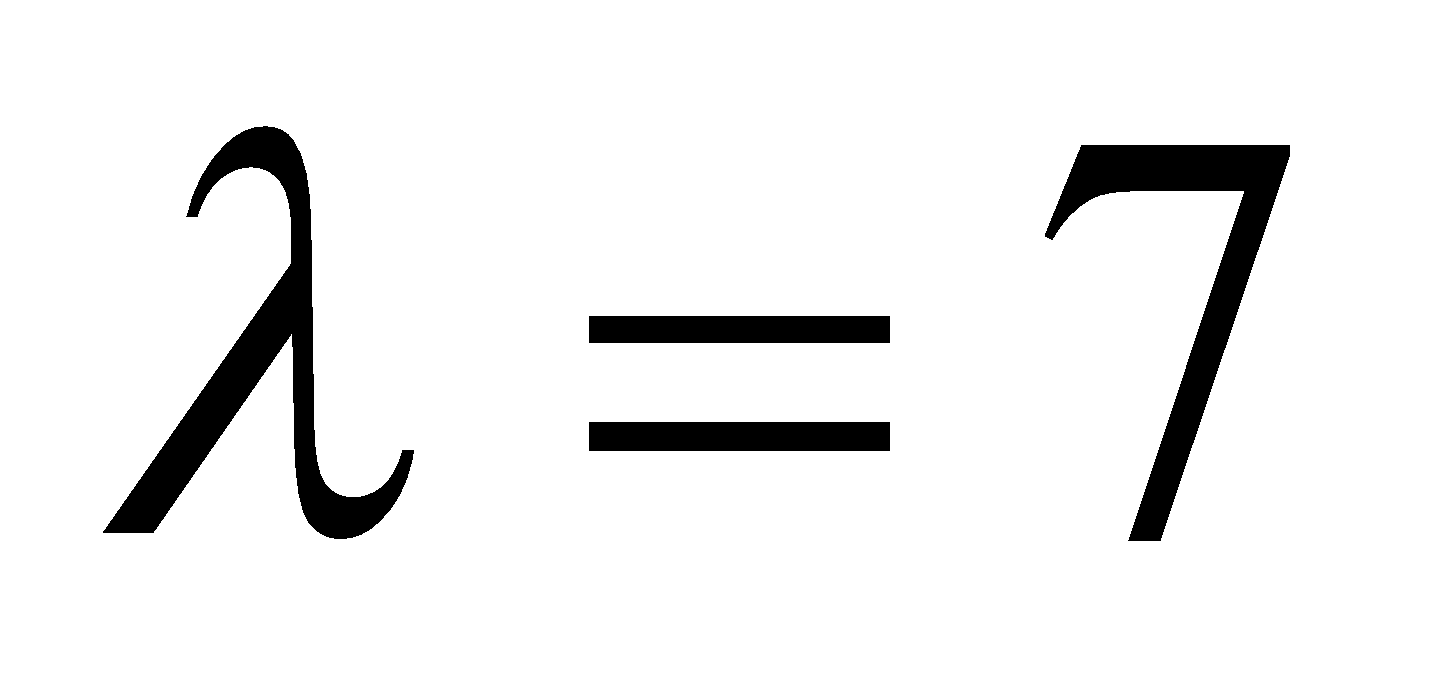
**4.4.** Для наступних 50 реалізацій випадкової величини

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 2,03 | 2,25 | 2,94 | 2,30 | 1,00 | 2,18 | 1,93 | 1,60 | 1,52 | 2,42 |
| 2,32 | 1,43 | 1,79 | 2,07 | 1,89 | 1,49 | 1,31 | 2,58 | 2,17 | 1,53 |
| 2,55 | 2,46 | 2,65 | 1,68 | 1,81 | 1,21 | 2,34 | 2,00 | 1,35 | 2,53 |
| 2,49 | 1,30 | 2,79 | 2,76 | 2,60 | 1,25 | 1,71 | 2,57 | 1,70 | 1,65 |
| 1,58 | 1,93 | 2,84 | 1,03 | 2,85 | 2,25 | 2,85 | 2,45 | 1,37 | 1,90 |

за критерієм Колмогорова-Смірнова або χ2 перевірити гіпотезу про рівномірний розподіл на рівні значущості 0,1.

**4.5.** Для наступних 50 реалізацій випадкової величини

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 9 | 8 | 8 | 6 | 10 | 6 | 4 | 7 | 6 | 10 |
| 6 | 5 | 7 | 6 | 6 | 4 | 12 | 2 | 12 | 4 |
| 5 | 6 | 10 | 7 | 8 | 3 | 9 | 6 | 4 | 8 |
| 7 | 4 | 8 | 4 | 6 | 2 | 10 | 5 | 6 | 3 |
| 9 | 8 | 5 | 4 | 7 | 8 | 7 | 4 | 3 | 7 |

за критерієм χ2 перевірити а) гіпотезу про пуассонівський розподіл, б) гіпотезу про пуассонівський розподіл з параметром  з надійністю 0,9.

**4.6.** Для наступних 50 реалізацій випадкової величини

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 2,30 | 2,04 | 3,62 | 2,21 | 1,82 | 2,82 | 2,77 | 1,44 | 1,72 | 2,69 |
| 1,72 | 4,08 | 2,11 | 0,91 | 2,82 | 1,82 | 2,91 | 0,14 | 0,76 | 1,45 |
| 1,59 | 1,80 | 2,33 | 2,25 | 1,76 | 3,31 | 1,92 | 3,32 | 0,60 | 1,81 |
| 2,96 | 1,33 | 3,01 | -1,23 | 0,76 | 2,19 | 2,82 | 2,95 | 3,22 | 2,48 |
| 1,56 | 5,06 | 0,61 | 1,83 | 2,04 | 1,57 | 1,49 | 1,07 | 3,06 | 1,77 |

за критерієм χ2 перевірити а) гіпотезу про гауссівський розподіл, б) гіпотезу про гауссівський розподіл N(2;1) з надійністю 0,95.

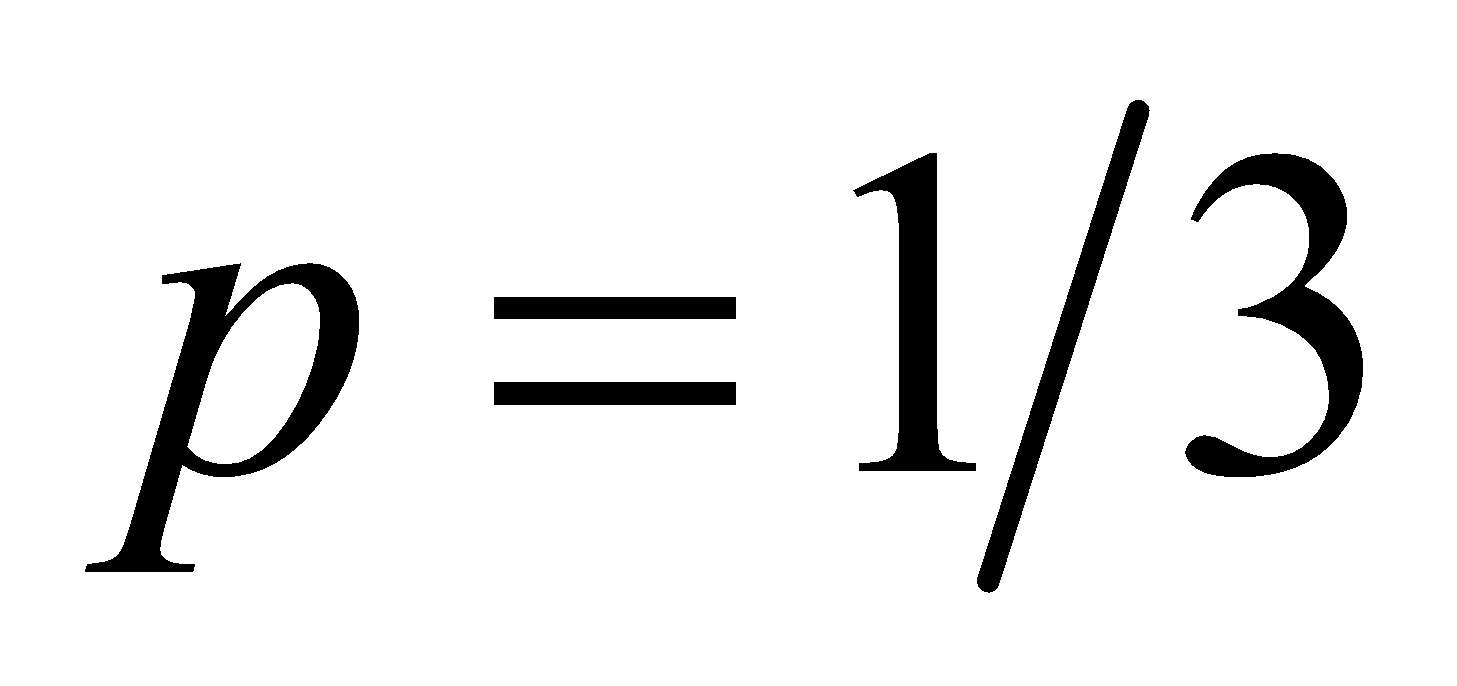
**4.7.** Для наступних 50 реалізацій випадкової величини

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 9,68 | 4,31 | 27,89 | 5,67 | 12,62 | 0,68 | 16,79 | 7,05 | 9,10 | 6,59 |
| 10,87 | 2,07 | 4,84 | 6,65 | 3,42 | 25,28 | 3,40 | 61,42 | 3,51 | 5,53 |
| 26,68 | 9,85 | 13,96 | 1,34 | 5,72 | 3,12 | 17,00 | 17,03 | 2,79 | 8,36 |
| 33,51 | 7,21 | 129,97 | 47,03 | 3,22 | 15,63 | 7,78 | 22,26 | 5,26 | 8,26 |
| 1,53 | 33,77 | 0,87 | 2,91 | 19,83 | 36,54 | 4,93 | 5,64 | 3,44 | 2,86 |

за критерієм χ2 перевірити: а) гіпотезу про логнормальний розподіл, б) гіпотезу про логнормальний розподіл logN(2; 1) з надійністю 0,9.

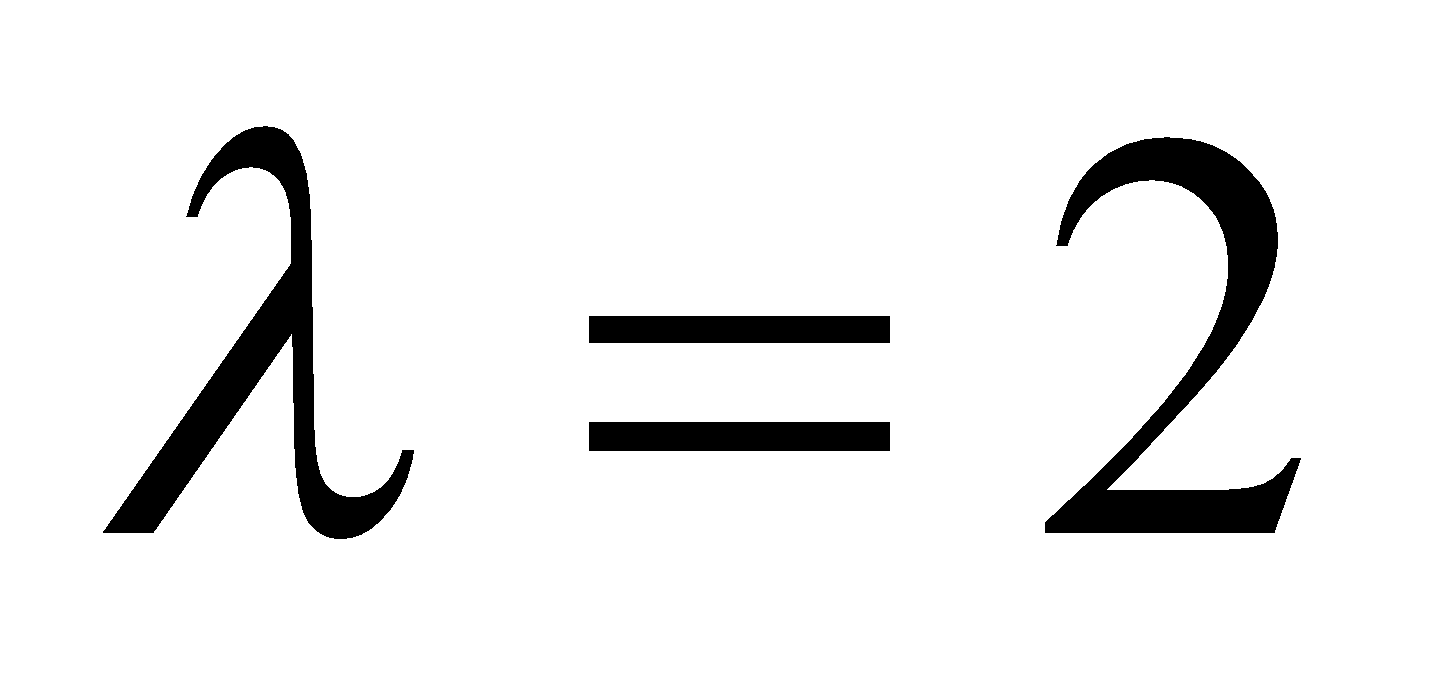
**4.8.** Для наступних 50 реалізацій випадкової величини

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 0 | 7 | 0 | 8 | 0 | 4 | 0 | 0 | 0 | 2 |
| 1 | 0 | 0 | 2 | 1 | 0 | 2 | 2 | 0 | 0 |
| 1 | 0 | 3 | 10 | 1 | 0 | 4 | 3 | 1 | 0 |
| 1 | 3 | 0 | 1 | 2 | 0 | 0 | 0 | 0 | 3 |
| 0 | 1 | 6 | 0 | 4 | 6 | 4 | 3 | 2 | 0 |

за критерієм χ2 перевірити: а) гіпотезу про геометричний розподіл, б) гіпотезу про геометричний розподіл з параметром  з надійністю 0,95.

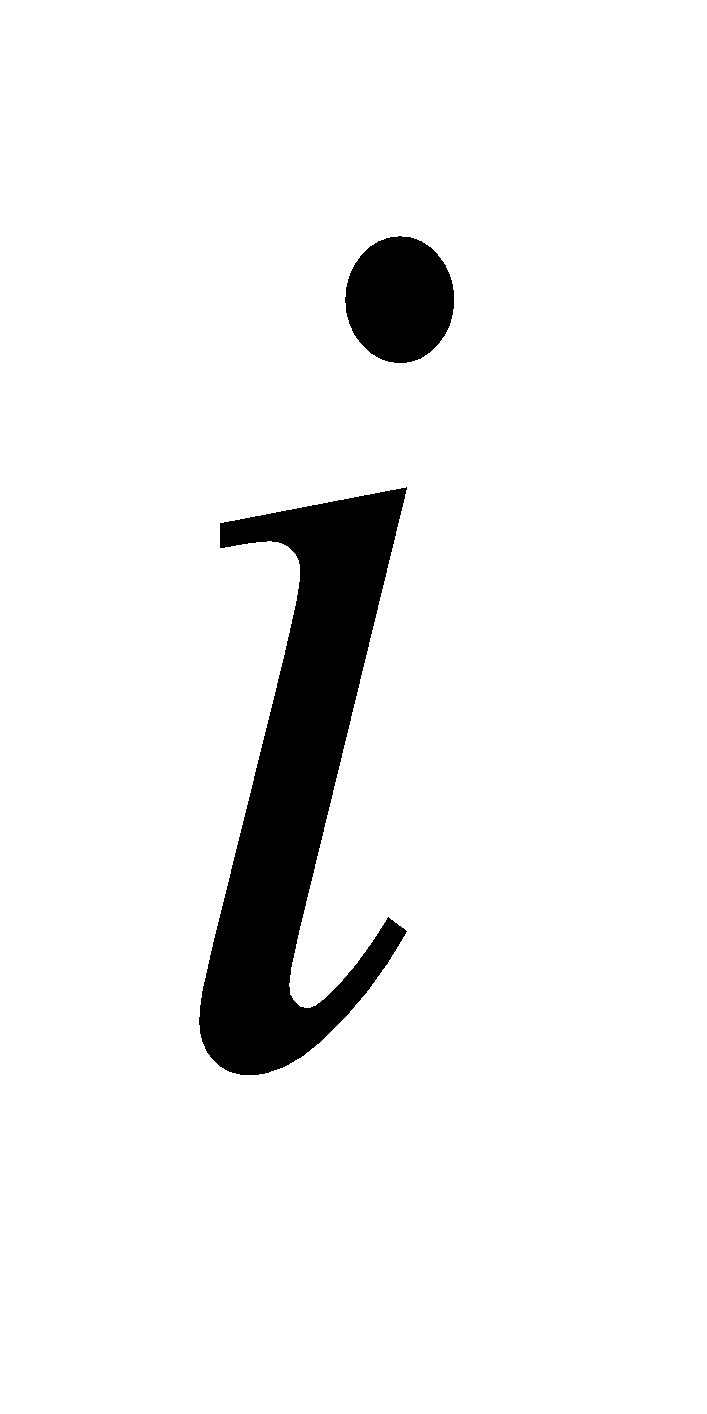
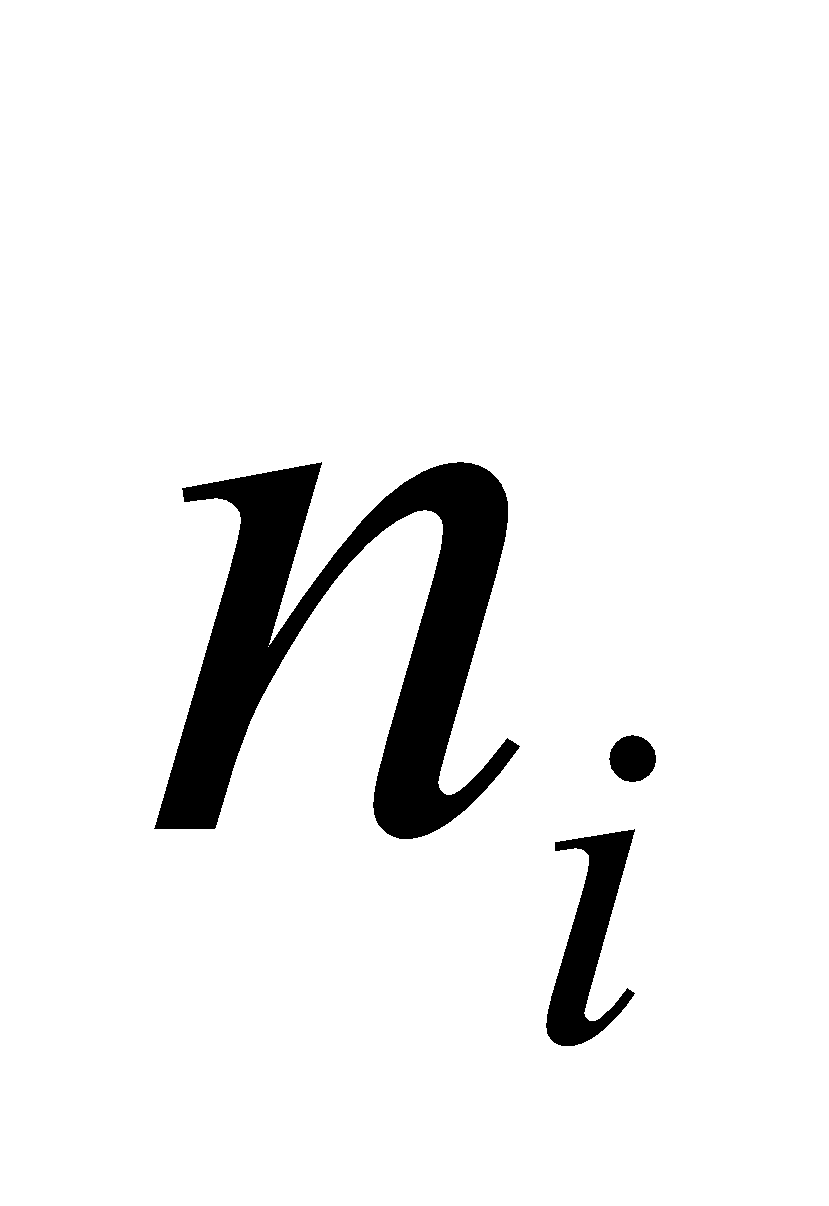
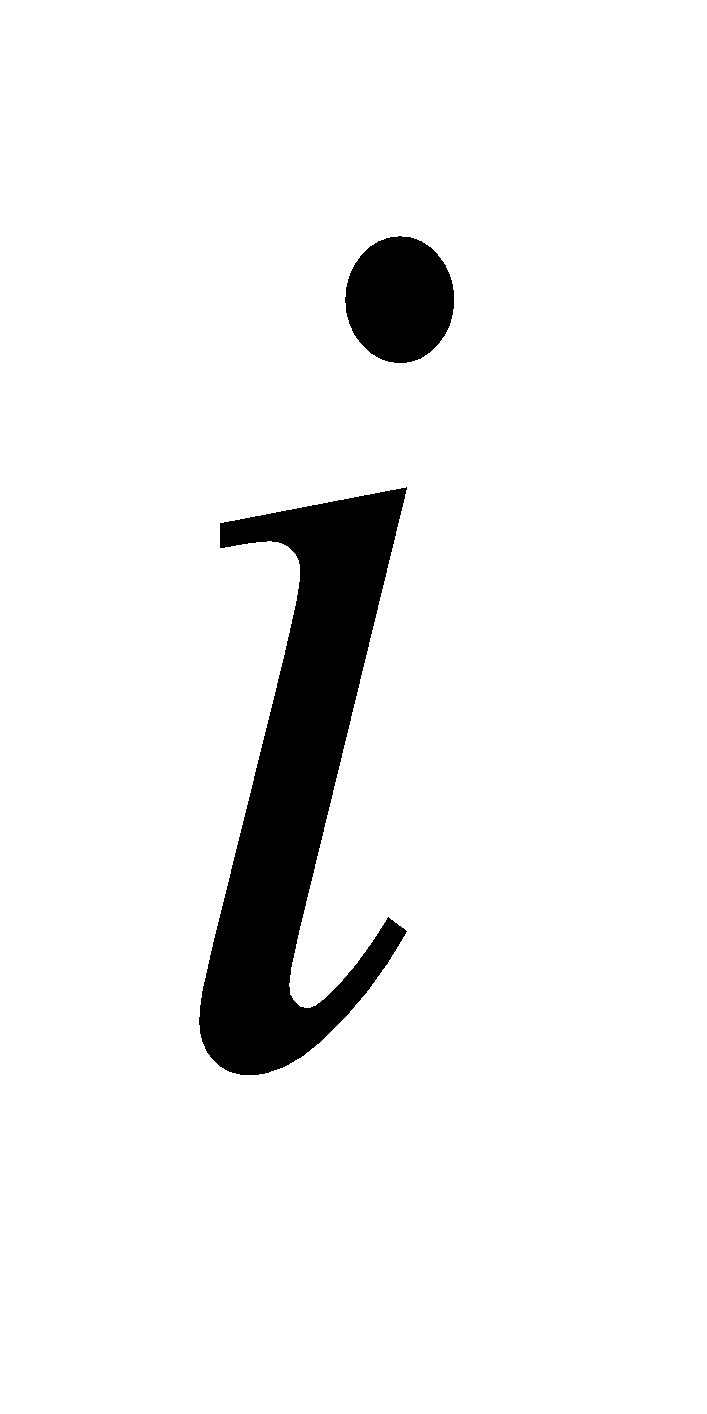
**4.9.** Для наступних 50 реалізацій випадкової величини

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 0,57 | 0,95 | 0,15 | 0,40 | 0,56 | 0,29 | 1,64 | 0,18 | 0,62 | 0,00 |
| 0,92 | 0,43 | 0,16 | 1,22 | 0,56 | 0,21 | 0,02 | 0,22 | 0,15 | 0,49 |
| 0,27 | 0,38 | 0,05 | 0,31 | 0,20 | 0,09 | 0,28 | 0,05 | 0,30 | 0,32 |
| 0,41 | 0,49 | 0,81 | 0,33 | 0,71 | 1,59 | 0,58 | 0,59 | 0,18 | 0,14 |
| 0,04 | 0,07 | 0,03 | 0,45 | 0,16 | 0,78 | 0,25 | 0,08 | 0,02 | 0,31 |

за критерієм χ2 перевірити а) гіпотезу про експоненціальний розподіл, б) гіпотезу про експоненціальний розподіл з параметром  з надійністю 0,95.

**4.10.**Рентгенівське випромінення викликає в органічних клітинах певну перебудову хромосом. В таблиці наведено результати експеримента (підраховувалась кількість перебудов хромосом під впливом рентгенівських променів).

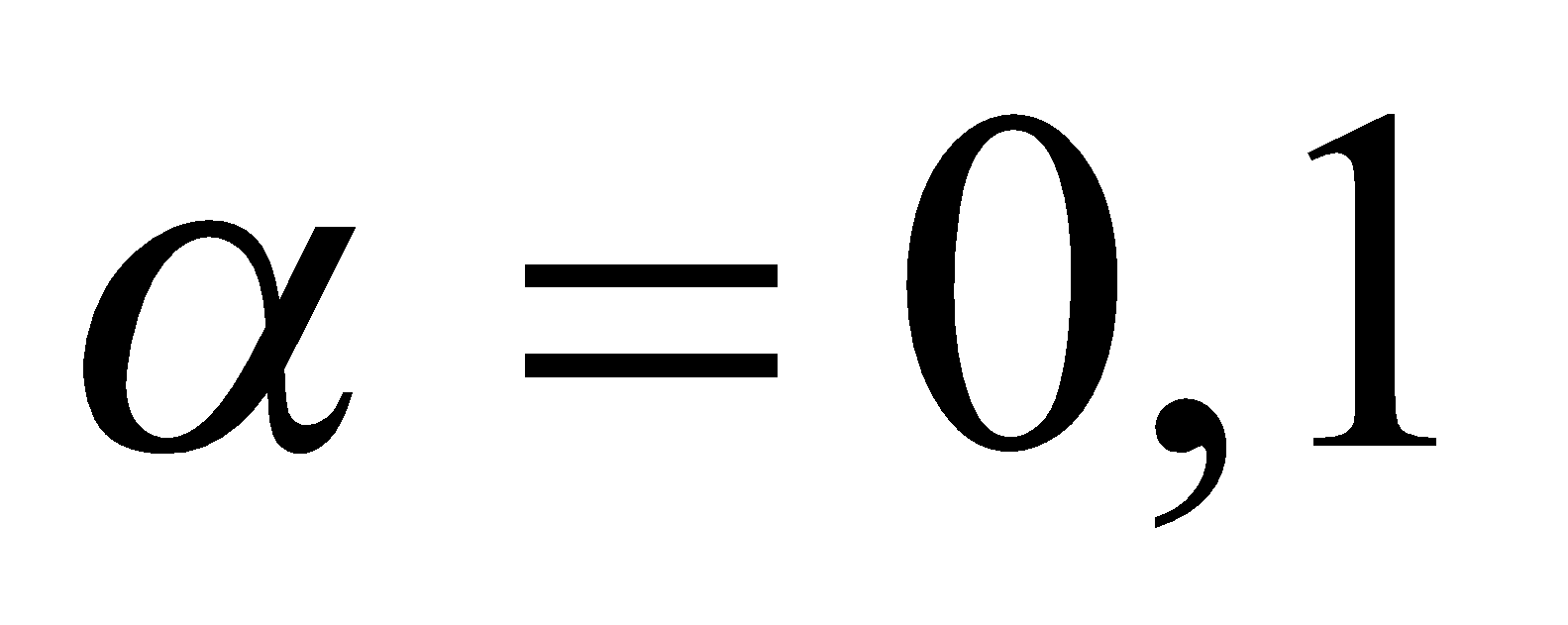
|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | 0 | 1 | 2 | 3 | 4 і більше | Всього |
|  | 434 | 195 | 44 | 9 | 0 | 682 |

Тут  – кількість змін в клітині,  – кількість клітин з  змінами.

Чи узгоджується з наведеними даними гіпотеза про пуассонівський розподіл кількості перебудов у клітині?

**4.11.** З продукції двох верстатів зробили дві вибірки по 38 виробів:

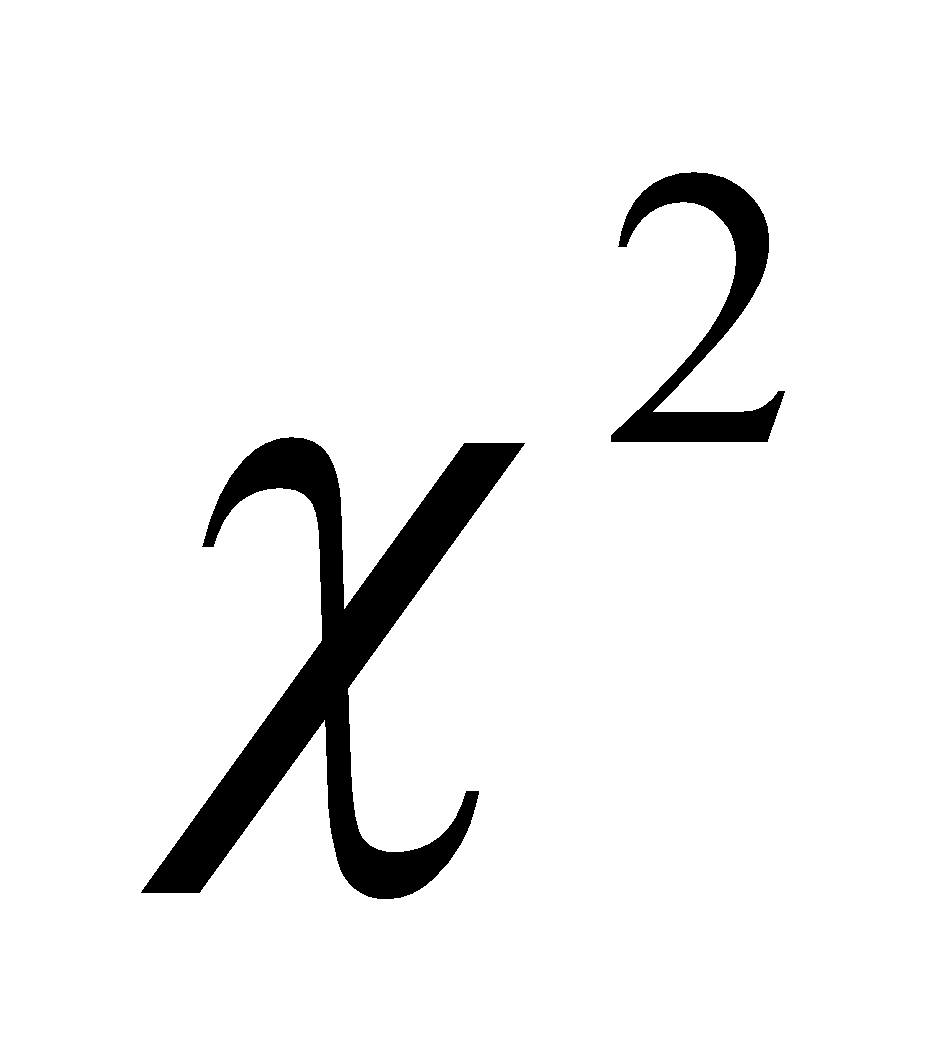
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Розмір  деталі | 24 | 26 | 28 | 30 | 32 | 34 | 36 | 38 | 40 | 42 | 44 | 46 | 48 | 50 | 52 | 54 | 56 |
|  | 2 | 1 | 2 | 2 | 1 | 1 | 4 | 1 | 1 | 0 | 5 | 0 | 0 | 6 | 4 | 3 | 5 |
|  | 2 | 0 | 0 | 2 | 3 | 0 | 1 | 6 | 0 | 5 | 3 | 1 | 5 | 3 | 3 | 2 | 4 |

Перевірити, використовуючи критерій Смірнова-Колмогорова, гіпотезу про те, що ці виборки належать одній і тій самій генеральній сукупності при рівні значущості .

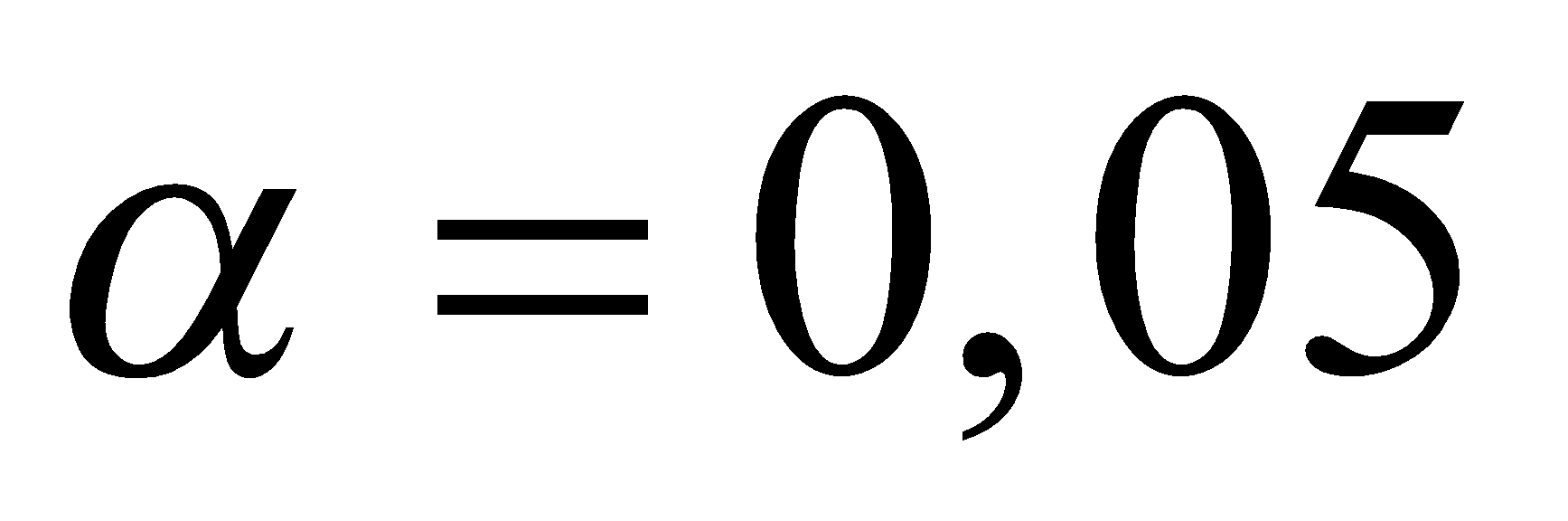
**4.12.** В першому потоці з 300 абітурієнтів оцінку «2» отримало 33 особи, «3» – 43 особи, «4» – 80 осіб, «5» – 144, а у другому потоці інші 300 абітурієнтів мали такий результат: «2» – 39 осіб, «3» – 35, «4» – 72, «5» – 154. Чи можна вважати обидва потоки однорідними при рівні значущості 0,05?

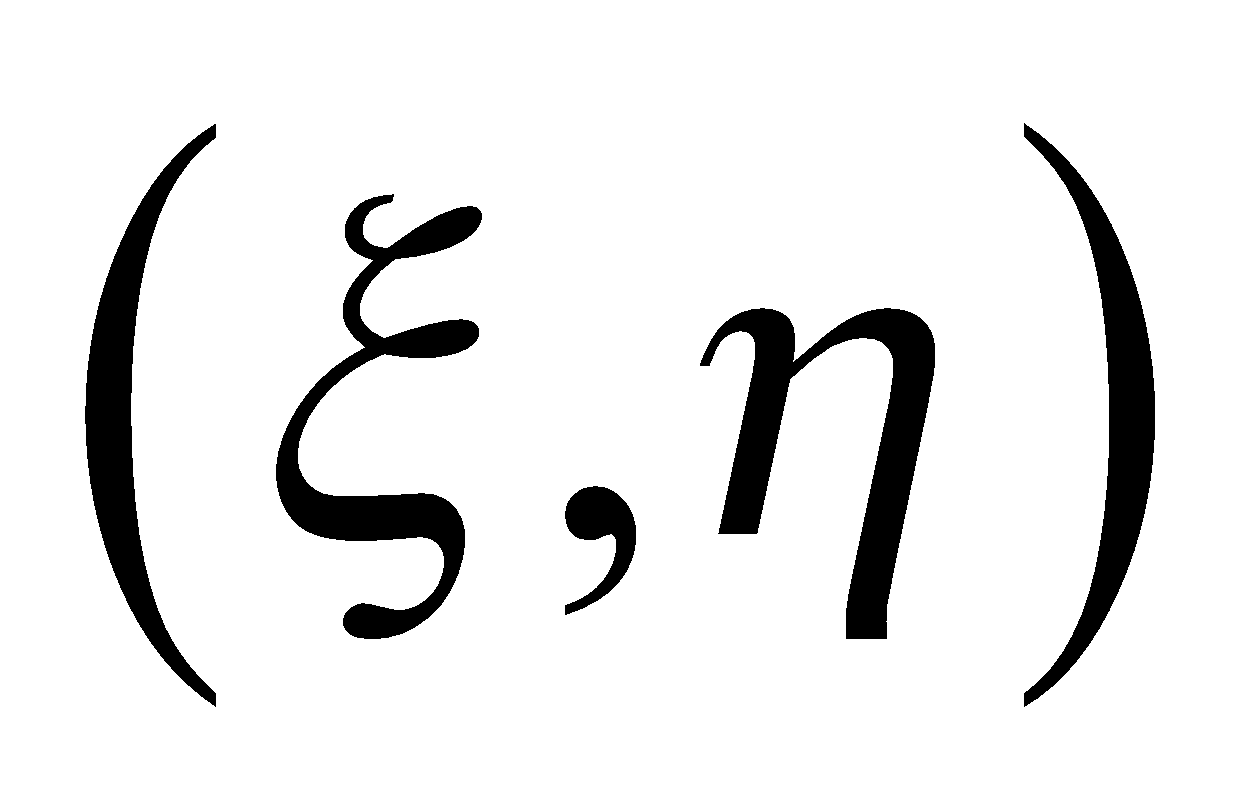
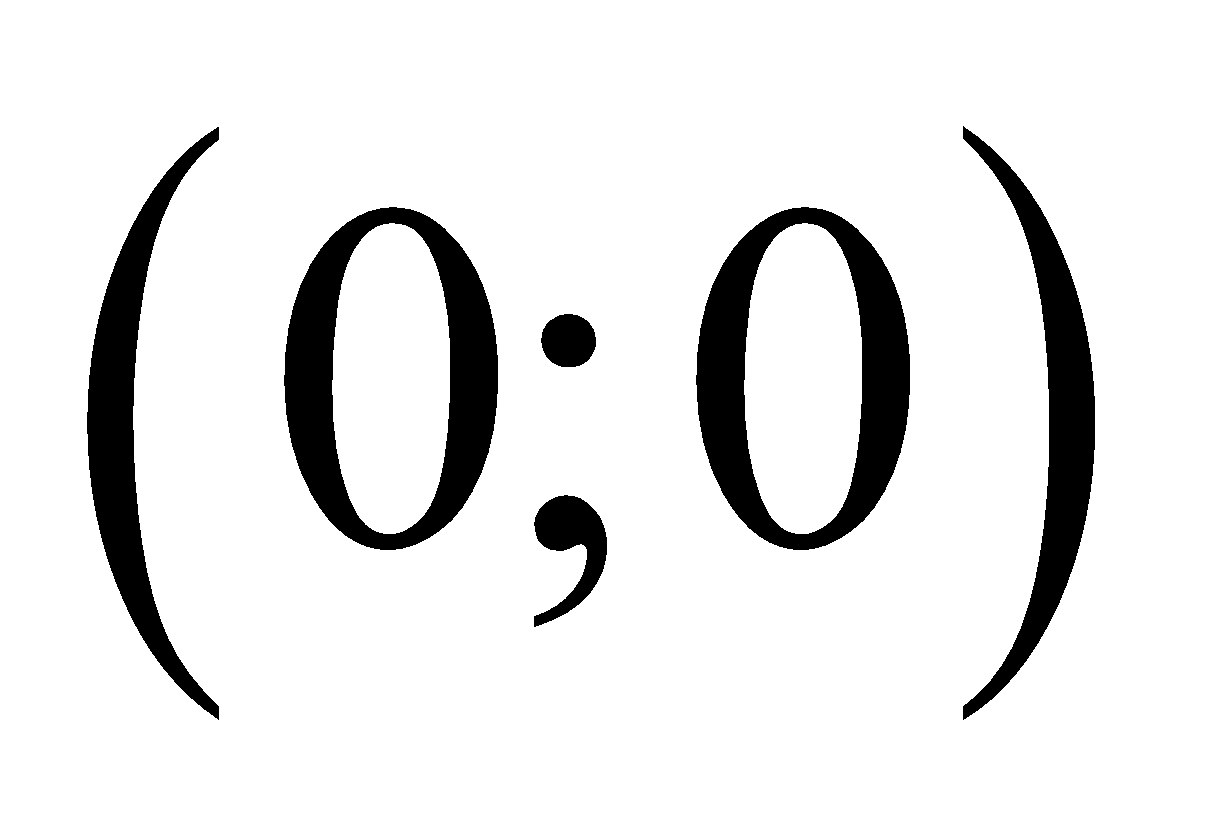
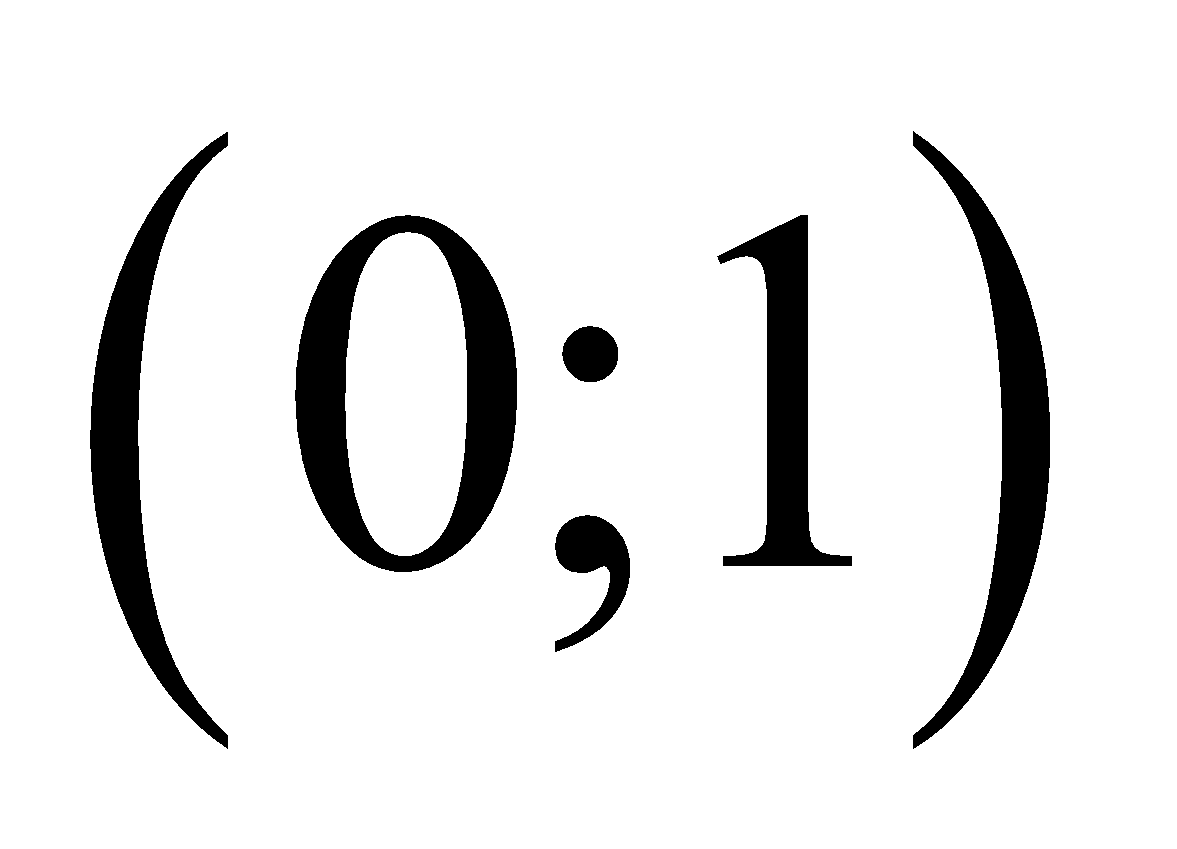
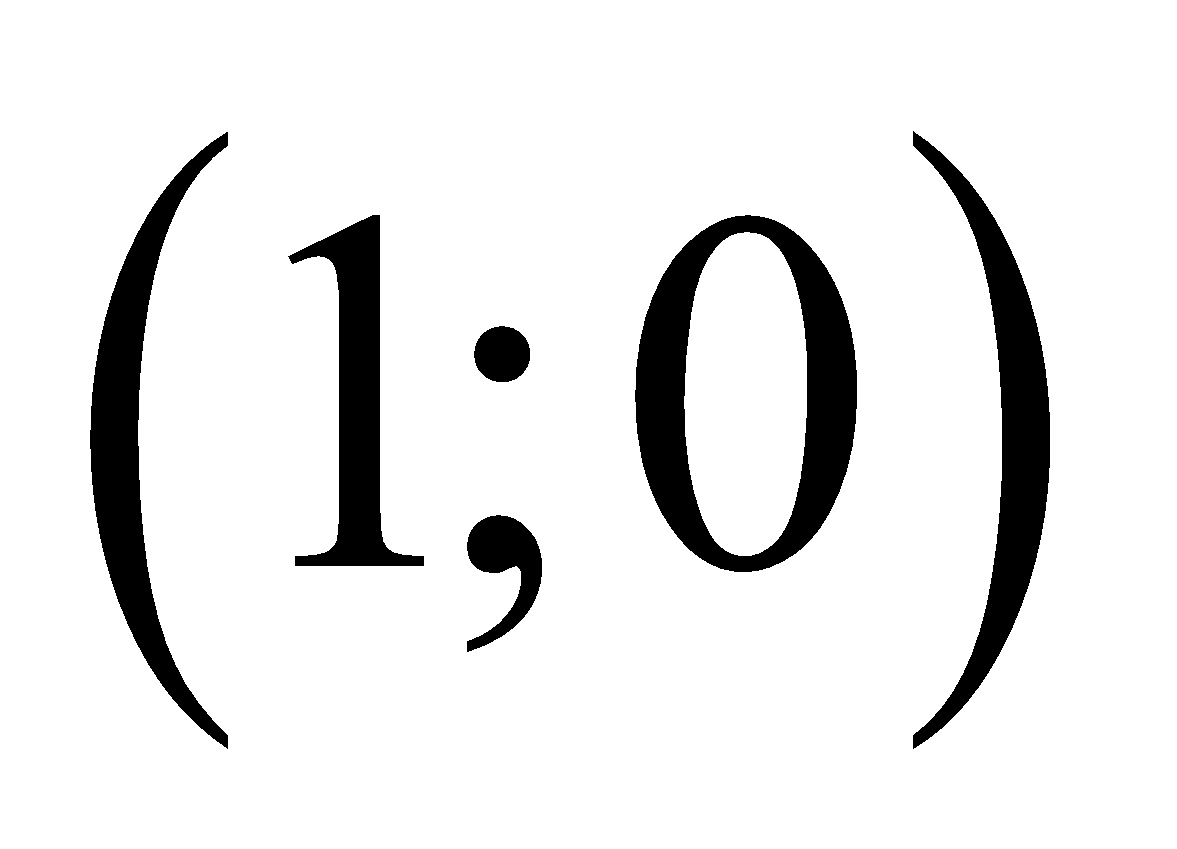
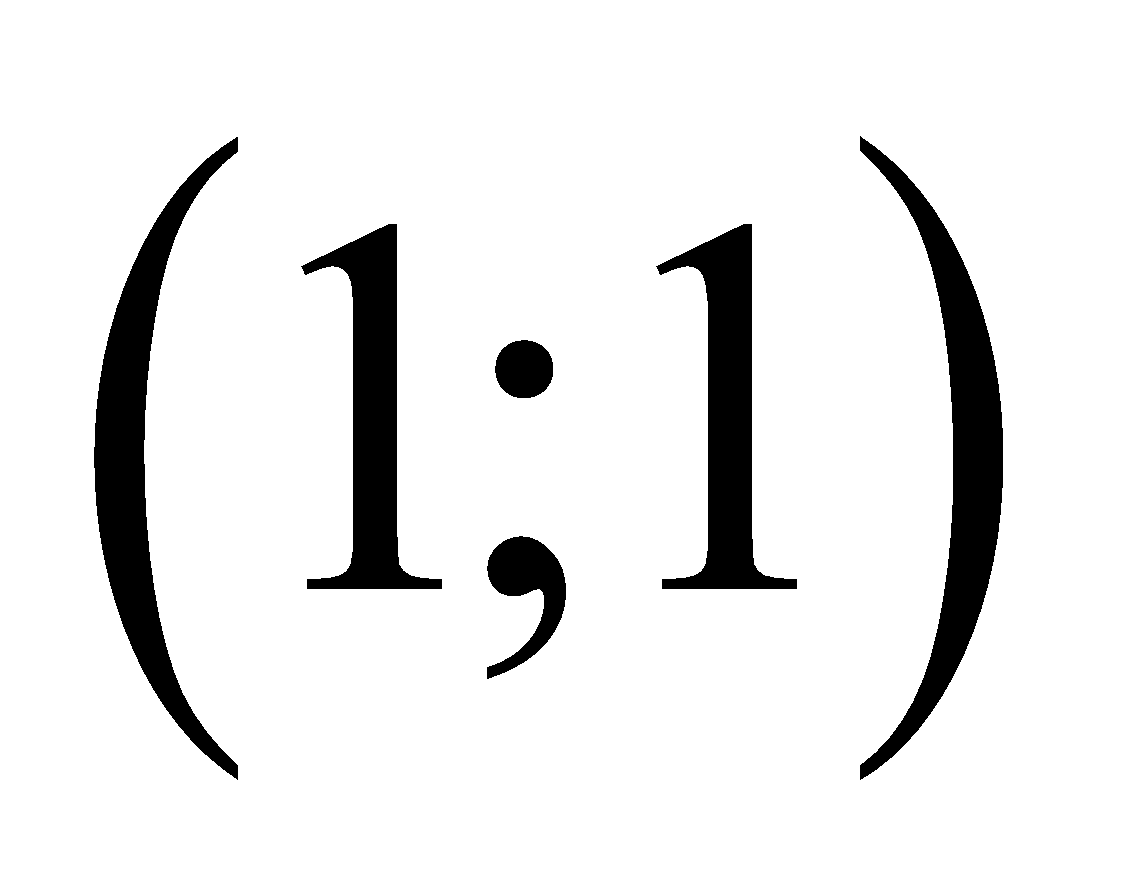
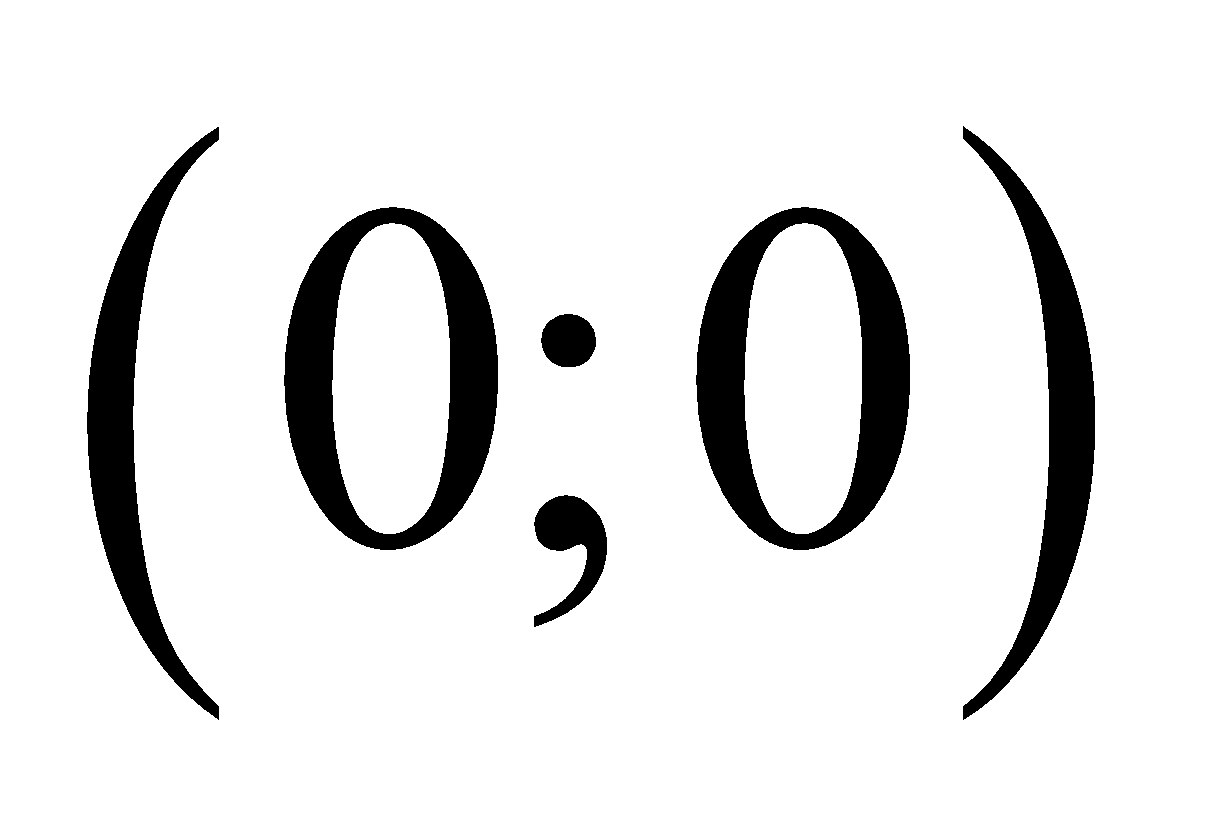
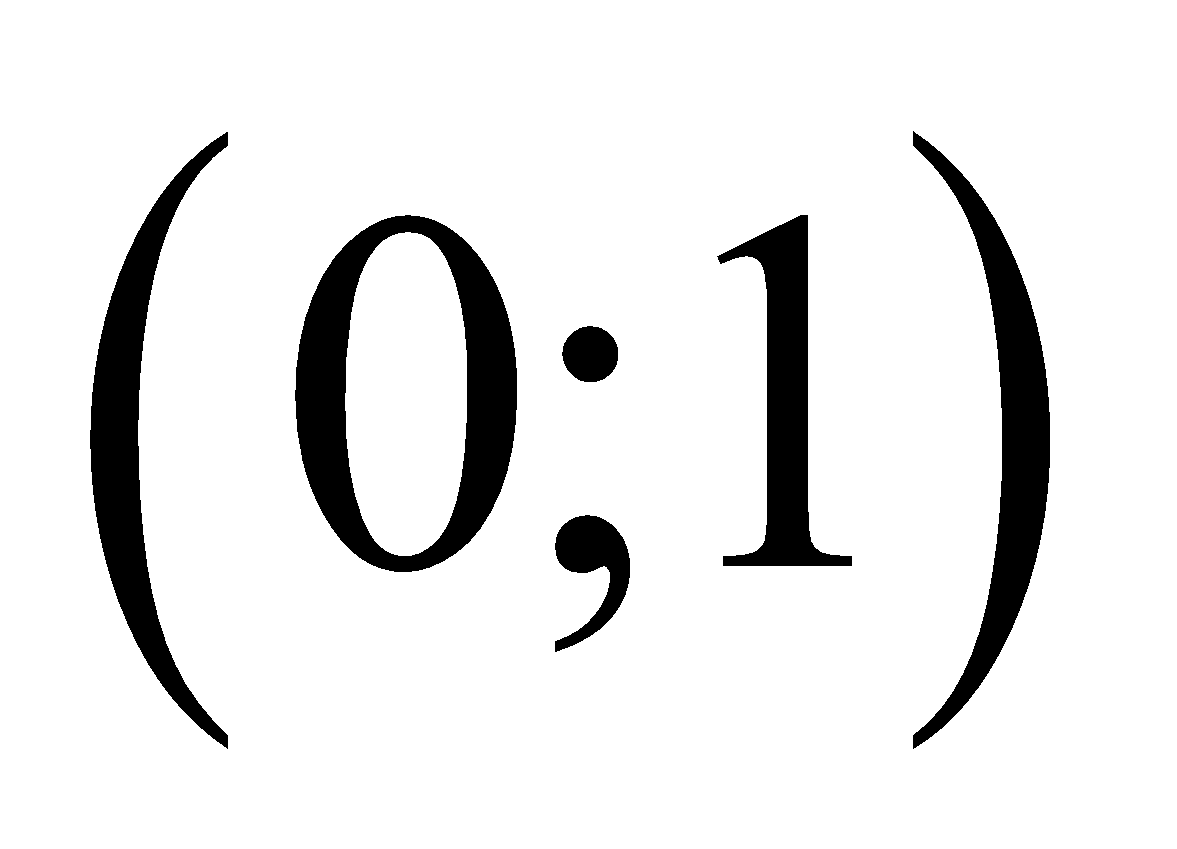
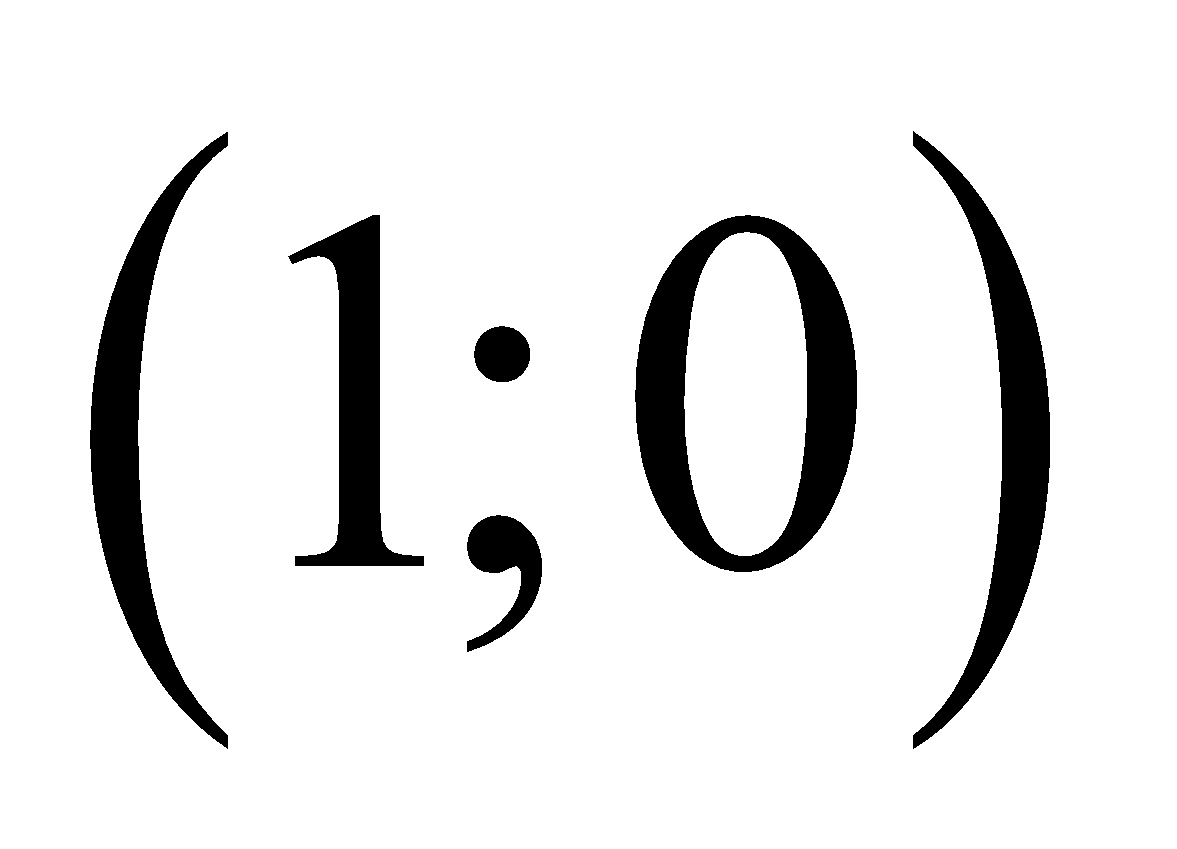
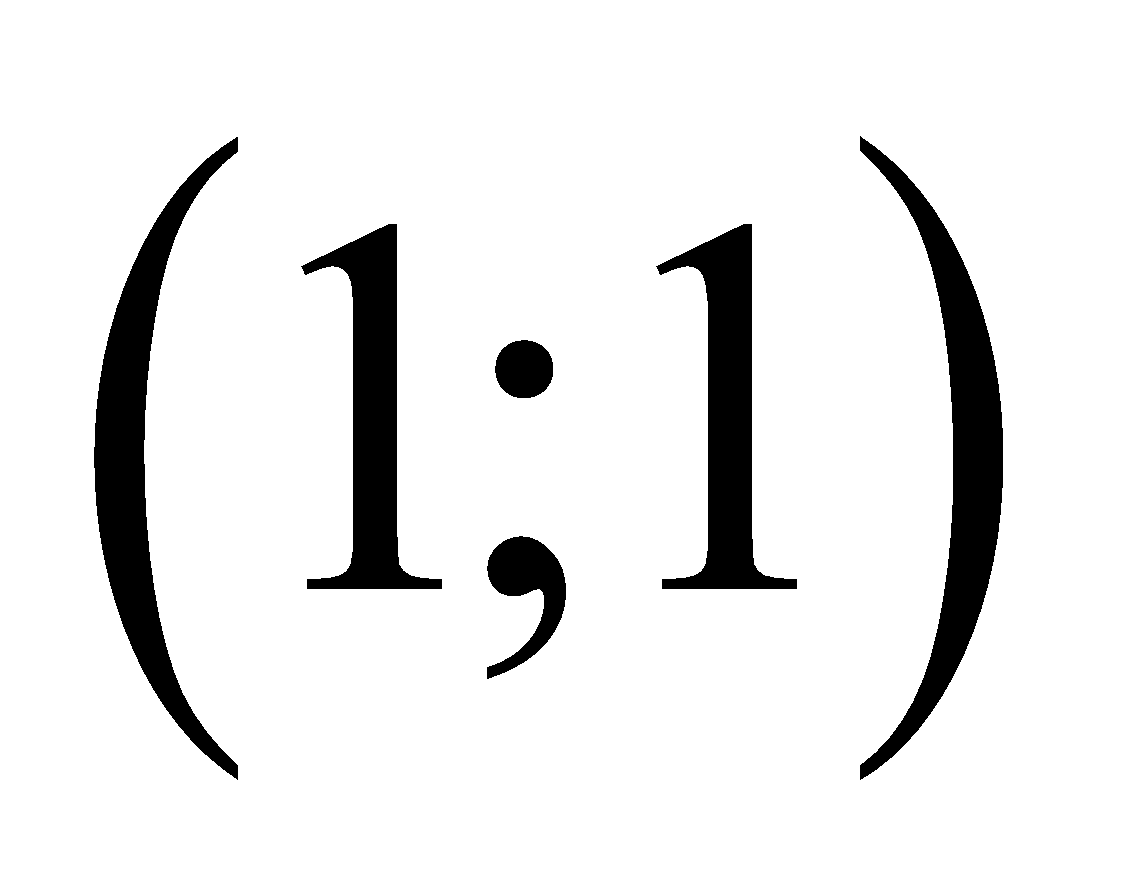
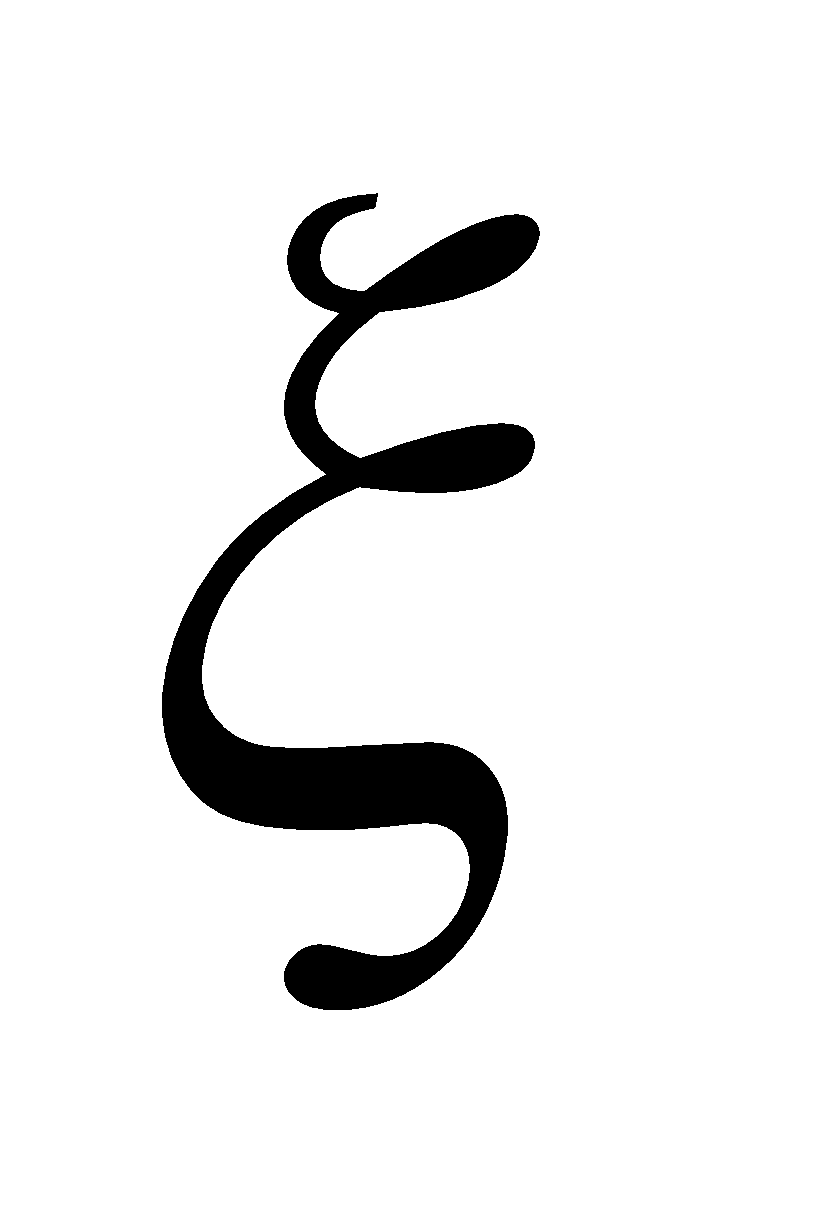
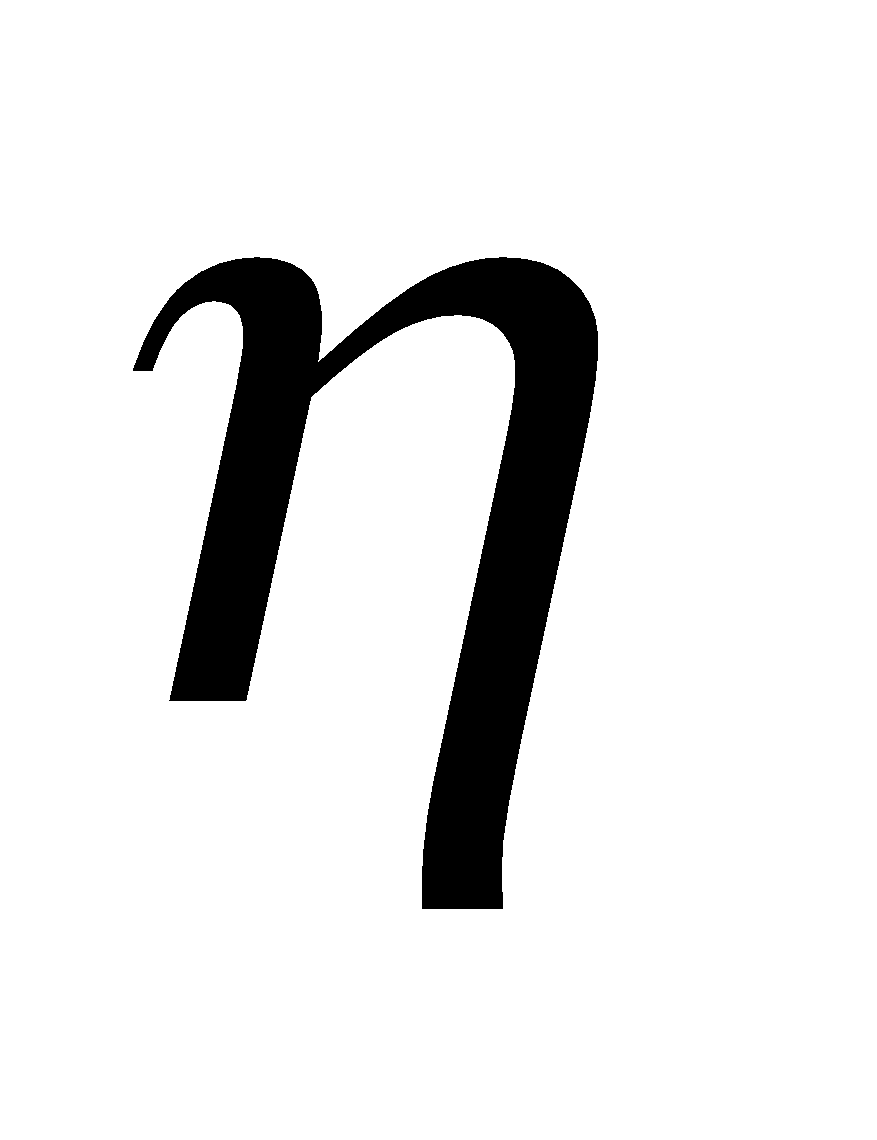
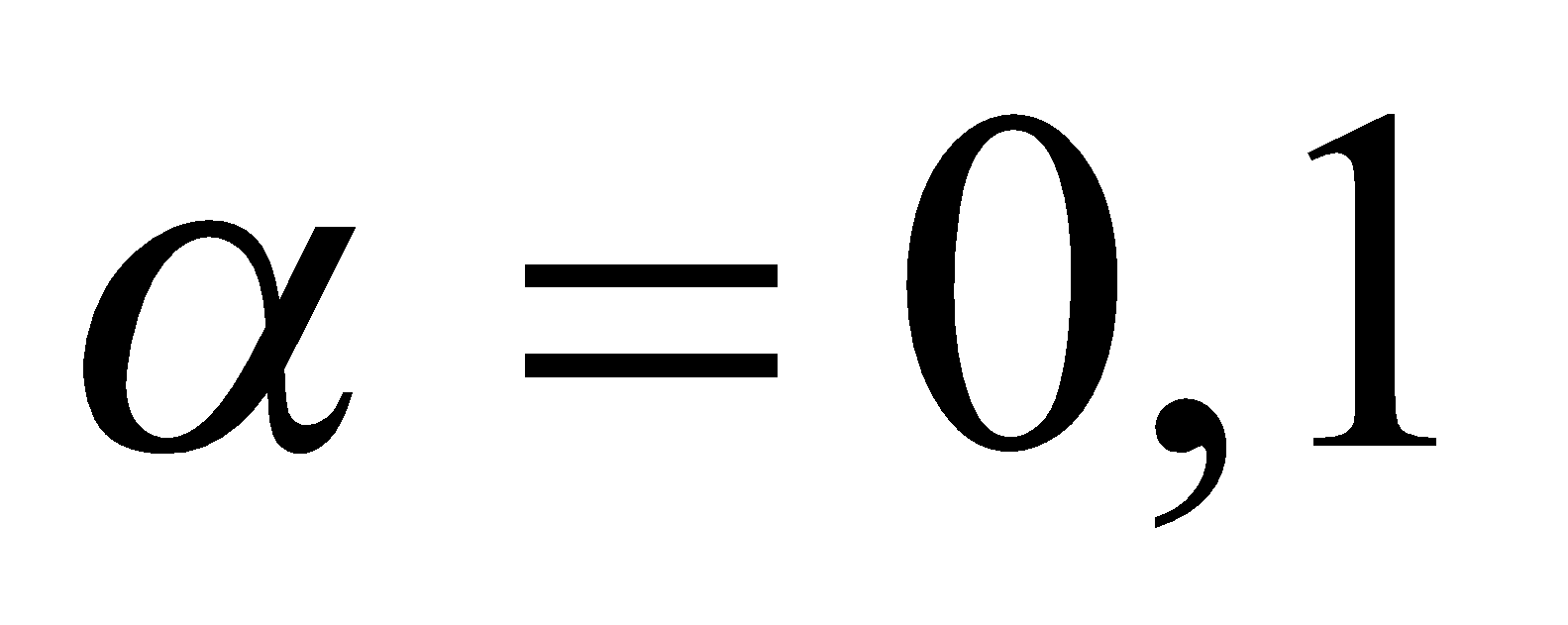
**4.13.** В таблиці наведено результати обстеження 697 школярів. Хлопці були впорядковані за рівнем IQ та відповідно до умов їх проживання вдома. При цьому використано позначення: А – дуже здібний, B – досить здібний, C – має середні здібності, D – недостатньо розвинутий, E – розумово відсталий. Чи можна вважати, що умови життя (забезпеченість) дітей впливають на їх здібність?

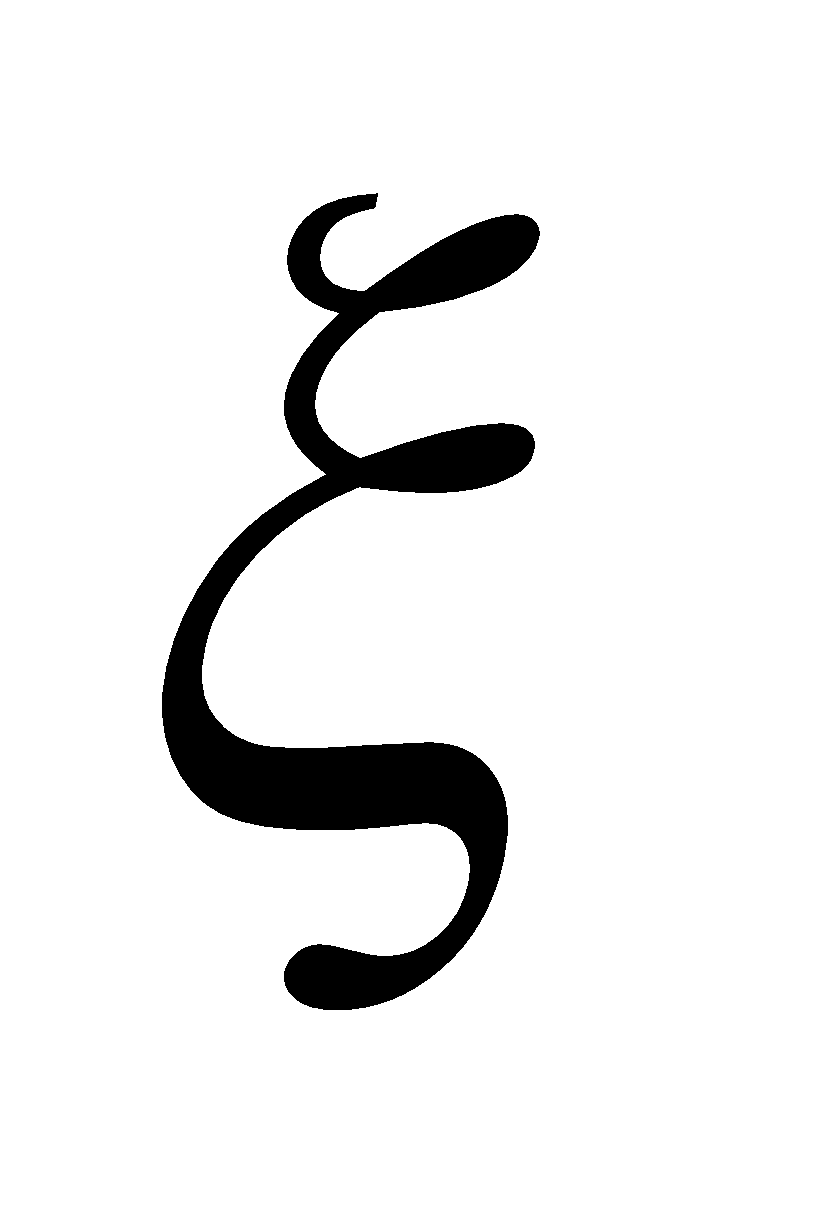
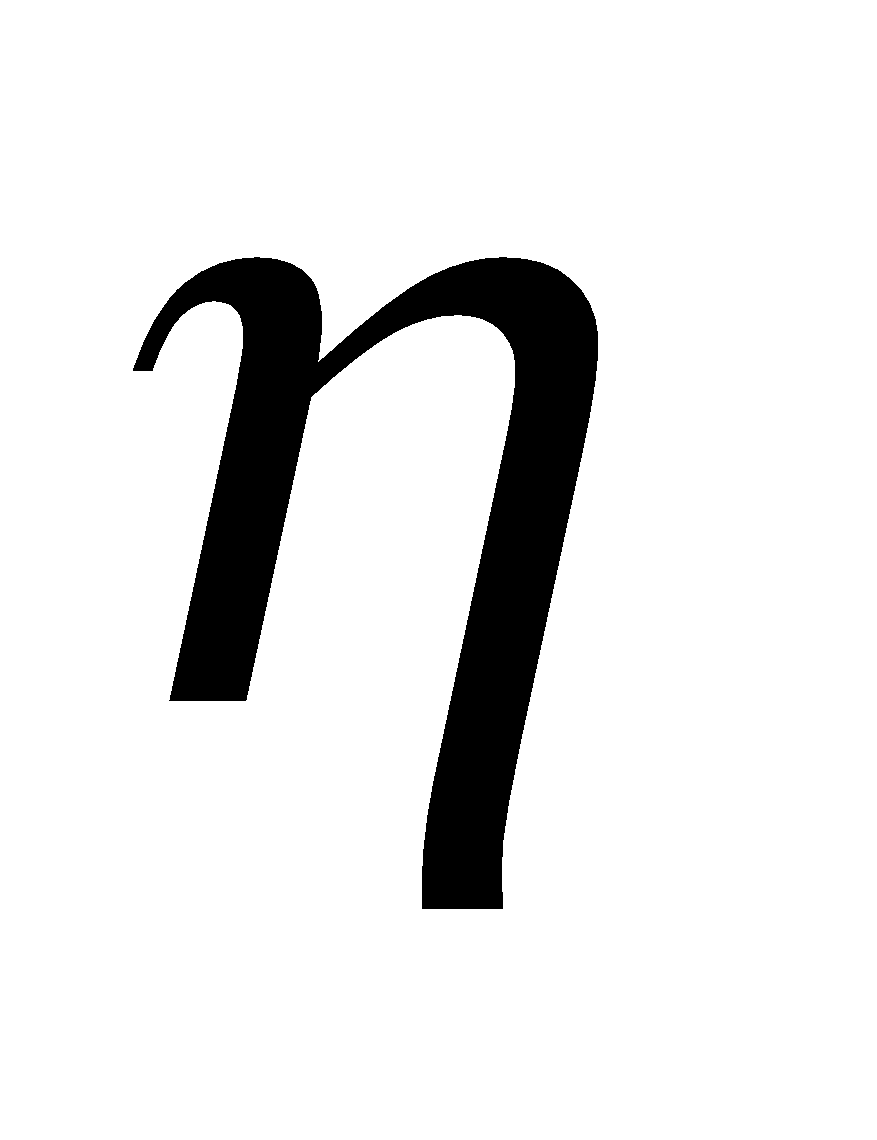
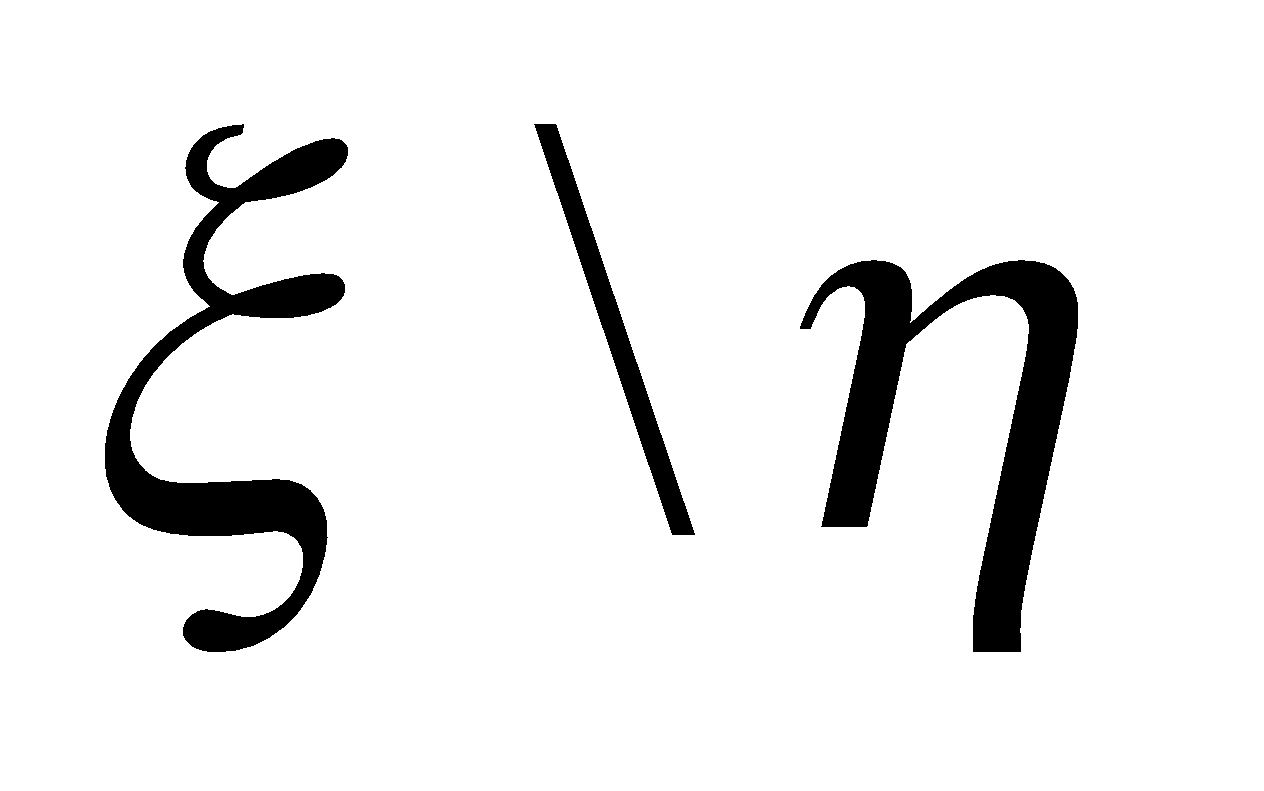
|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Забезпеченість | Здібність хлопців | | | | | Всього |
| A | B | C | D | E |
| Хороша | 33 | 137 | 125 | 47 | 8 | 350 |
| Погана | 21 | 127 | 129 | 61 | 9 | 347 |
| Всього | 54 | 264 | 254 | 108 | 17 | 697 |

**4.14.** За допомогою критерію **** перевірити гіпотезу однорідності двох вибірок, наведених у таблиці.

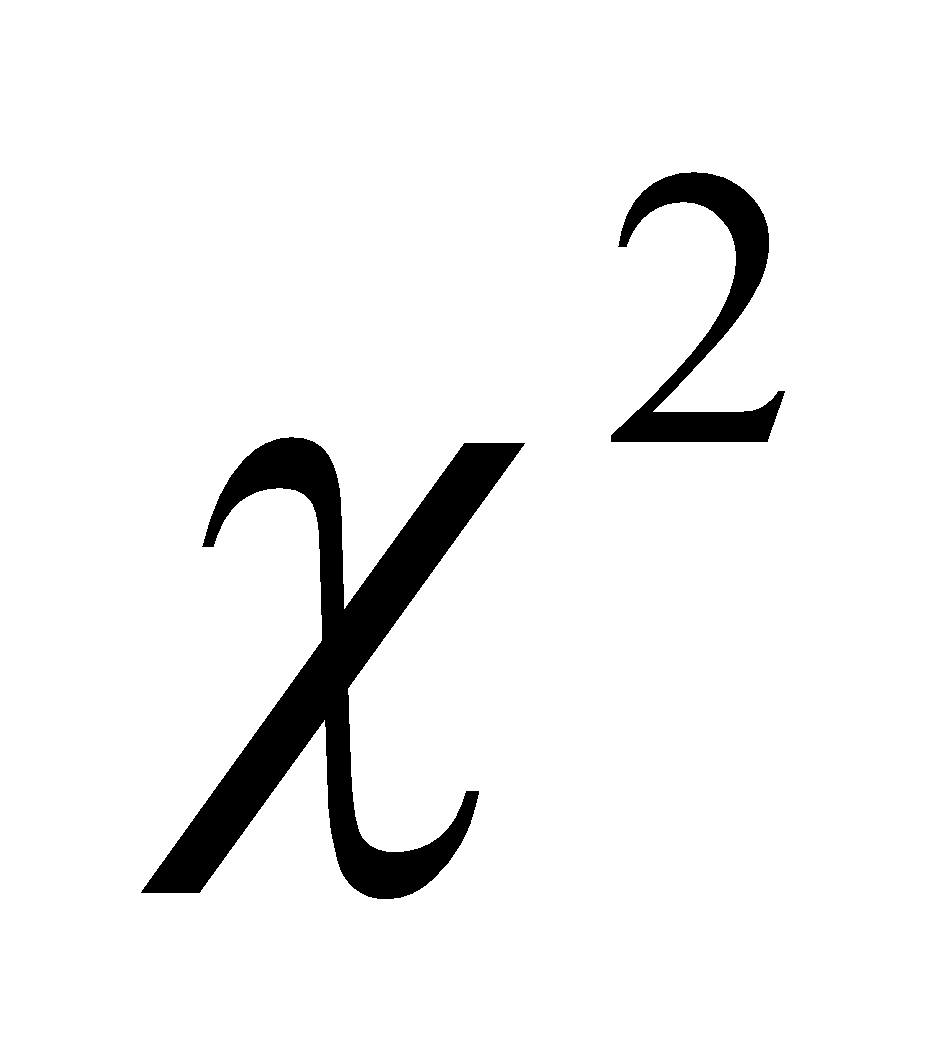
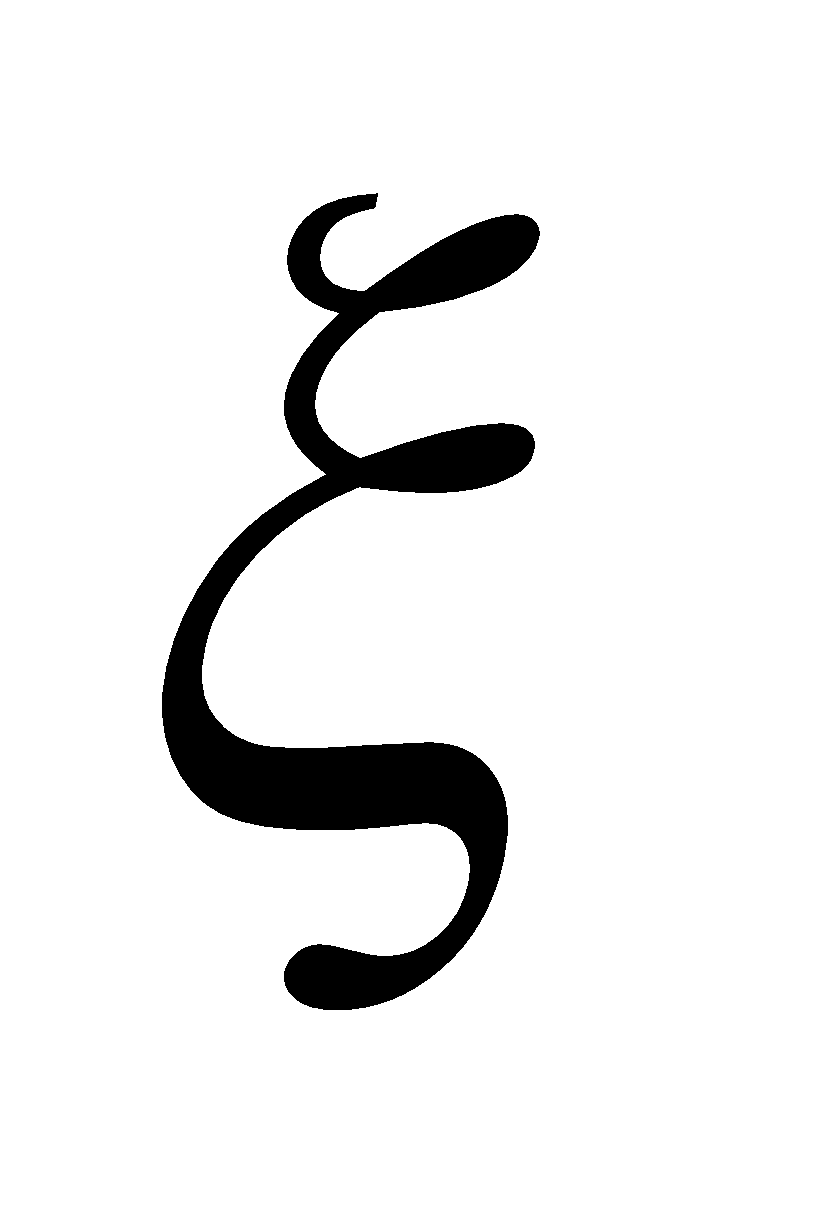
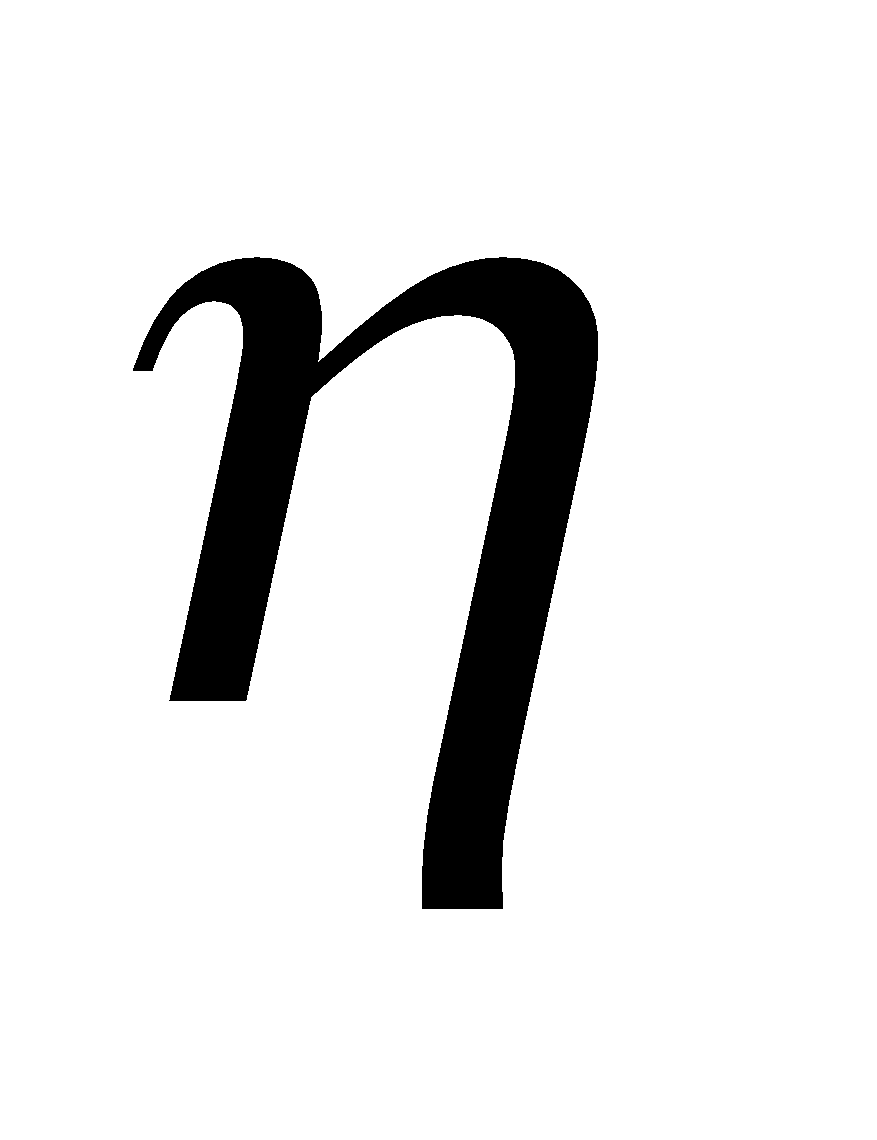
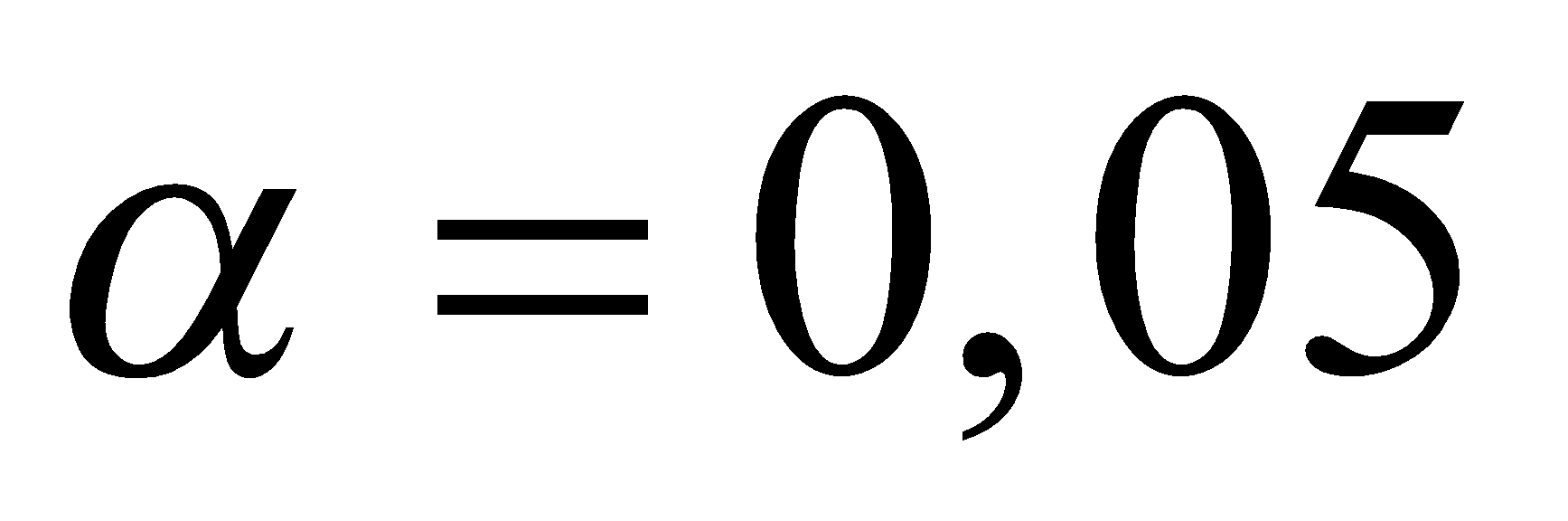
|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 |
|  | 4 | 4 | 15 | 51 | 22 | 3 | 1 | 0 |
|  | 1 | 1 | 8 | 43 | 34 | 7 | 3 | 3 |

.

**4.15.** Двовимірна випадкова величина  може приймати 4 значення: , , , . 180 незалежних спостережень дали такі результати: значення  з’явилося 39 разів,  – 50,  – 53,  – 38. Чи можна вважати, що  і  – незалежні? Рівень значущості прийняти .

**4.16.** Проведено 200 спостережень над випадковими величинами  і ., які приймають значення 1, 2 та 1, 2, 3 відповідно. Результати спостережень наведені у таблиці: 

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | 1 | 2 | 3 |  |
| 1 | 25 | 50 | 25 | 100 |
| 2 | 52 | 41 | 7 | 100 |
|  | 77 | 91 | 32 | 200 |

Перевірити за допомогою критерію **,** чи будуть незалежними випадкові величини  і  при .

**4.17.** Серед 300 осіб, які поступали в університет, 97 мали оцінку «5» в школі, 48 отримали «5» на вступних іспитах по тому ж самому предмету, причому лише 18 осіб мали «5» і в школі, і на вступних іспитах. На рівні значущості 0,1 перевірити гіпотезу незалежності оцінок «5» в школі і на вступних іспитах.

**4.18.** В таблиці (Грінвуд, Юл, 1915р.) наведено дані про 818 випадків, класифікованих по двох ознаках: наявність щеплення проти холери та відсутність захворювання.

Чи можна на основі цих даних прийти до висновку про залежність між відсутністю захворювання та наявністю щеплення?

|  |  |  |  |
| --- | --- | --- | --- |
| Наявність щеплення | Наявність захворювання | | Всього |
| Не захворіли | Захворіли |
| Щеплені | 276 | 3 | 279 |
| Не щеплені | 473 | 66 | 539 |
| Всього | 749 | 69 | 818 |