**Порівняння дисперсій двох нормальних розподілів**

Нехай .– вибірка з нормально розподіденої статистичної змінної  , а  – незалежна від неї вибірка з нормально розподіленої статистичної змінної .

Потрібно перевірити гіпотезу про те, що дисперсії генеральних сукупностей, з яких взяті ці вибірки, однакові:  

На основі даних вибірок знаходимо їх вибіркові середні та варіанси відповідно:

, , , .

Виберемо рівень значущості α та розглянемо статистику

. (\*)

Р. Фішер (англійський статистик) у 1926 р. довів, що у випадку істинності висунутої нульової гіпотези статистика  має розподіл із густиною розподілу:

, (\*\*)

де .

Вираз (\*) називають дисперсійним відношенням Фішера, а (\*\*) – щільністю Фішера з  ступенями вільності:

- число ступенів вільності чисельника

 - число ступенів вільності знаменника дисперсійного відношення (\*). Зауважимо, що при  статистика  збігається до статистики . п.4 ().

На основі статистики  визначаємо критичну область для гіпотези. Очевидно, що для гіпотези сприятливими будуть ті випадки коли  близьке до 1. Тому критична область для гіпотези складається з двох частин: із дуже малих і дуже великих значень статистики .

Довірча і критична область критерію Фішера є наступні: критична зліва на величину α/2, довірча посередині на величину 1-α і критична справа на величину α/2.

Тут α – заданий рівень значущості і, якщо , або , то гіпотезу відкидаємо. В протилежному випадку кажемо, що вона не суперечить експериментальним даним.

**Зауваження 1.**  – табульована для різних пар чисел ступенів вільності = і рівнів значущості α. Нижнє критичне значення дістаємо як обернену величину верхнього критичного значення статистики  при зміненій черговості ступенів вільності, тобто

.

Якщо в чисельнику дисперсійного відношення  ставити більшу варіансу, то це відношення буде більше одиниці, і тоді достатньо розглядати лише верхні критичні значення. У цьому випадку, якщо емпіричне значення статистики  попаде у верхню критичну область (), то гіпотезу відхиляємо як хибне твердження.

**Приклад.** На основі вибірок з двох нормально розподілених генеральних сукупностей обсягів  обчислено ,  . Чи можна вважати при десяти процентному рівні значущості, що дисперсії обох популяцій однакові?

У цьому випадку . Згідно означення дисперсійного відношення Фішера . У цьому відношенні у чисельнику маємо більшу за величиною варіансу. Тому з таблиці (додаток 8), при  та кількості ступенів вільності , знаходимо лише верхнє критичне значення статистики Фішера: . Оскільки , то немає підстав для сумнівів, що дисперсії генеральних сукупностей рівні між собою.

**Приклади застосування критерію**

При вивченні міцності 2-х типів бетону одержано





* Гіпотеза H1: міцність бетону не залежить від типу.
* Шукаємо 
* Вибираємо α=0,05 d.f.=14 |-> tкр=2,14
* Гіпотезу відкидаємо, бо 
* Гіпотеза H2: Ті, що виготовляють бетон 2-х типів однаково кваліфіковані.
* Шукаємо 
* Вибираємо α=0,05 d.f.=14 |-> Fкр=3,79
* Гіпотезу приймаємо, бо 
* На певному підприємстві розробили 2 методи виготовлення означеного виробу. Витрати сировини при роботі 2-ма методами є наступні:
* I: 2.0, 2.7, 2.5, 2.9, 2.3, 2.6
* II: 2.5, 3.2, 3.5, 3.5, 3.5
* Що можна про це сказати?
* Гіпотеза H1: методи однаково матеріалоємкі.
* Визначаємо величини:

 

* Шукаємо 
* Вибираємо α=0,05 d.f.=9 |-> tкр=2,26
* Гіпотезу приймаємо, бо 
* Гіпотеза H2: Ті, що виготовляють продукцію 2-х видів однаково кваліфіковані.
* Шукаємо 
* Вибираємо α=0,05 d.f.=(4,5) |-> Fкр=5,13
* Гіпотезу приймаємо, бо 

**Одновибірковий критерій погодженості(Крит.Колмогорова)**

* Нехай буде ряд незалежних спостережень над неперервною статистичою змінною ξ. Потрібно перевірити H про те, що популяція, з якої взята вибірка має функцію розподілу F(x), де F(x) – вповні означена. H: F(x)
* На основі вибірки знаходимо емпіричну функцію розподілу 

Одновибірковий критерій погодженості

* Розглянемо статистику 
* А. М. Колмогоров (1933 ) довів, що статистика типу  має розподіл незалежний від неперервної гіпотетичної функції розподілу F(x) і при збігається до розподілу



* Ця функція K(x) табульована в Tаблиці 7 при різних рівнях значущості α.
* Приклад. Дано вибірку з n=25 незалежних спостережень xi над незалежною змінною ξ. Потрібно перевірити H про те, що вибірка взята з нормально (розподіленої) популяції з середнім α=5 і стандартним відхиленням (стандартом) σ=10.

-1,5 4,8 2,0 8,7 4,3 -10,4 -10,3 2,4 9,4 7,1 7,7 7,8 15,0 11,0 17,5 13,8 19,3 4,4 -3,1 -5,2 -3,6 12,7 3,3 21,8 -6,7

* Якщо гіпотеза вірна, то лінійно перетворена за формулою  переведе нашу вибірку у вибірку з нормальною популяцією із сподіванням 0 і стандартом 1
* Запишемо лінійно перетворену вибірку:

-0,65 -0,02 -0,80 0,37 -0,01 -1,54 -1,53 -0,26 0,44 0,21 0,27 0,28 1,00 0,60 1,25 0,88 1,43 -0,26 -0,81 -1,02 -0,80 0,77 -0,17 1,68 -1,17

* Запишемо для останніх даних варіаційний ряд y(i), значення емпіричного та нормального розподілів у пунктах варіаційного ряду, а також абсолютні різниці в тих пунктах і зліва в них між обома розподілами  і 
* Результати запишемо далі.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  |  |  |  |  |
| -1,54 | 0,04 | 0,0618 | 0,0218 | 0,0618 |
| -1,53 | 0,08 | 0,0630 | 0,0171 | 0,0230 |
| -1,17 | 0,12 | 0,1210 | 0,0010 | 0,0410 |
| -1,02 | 0,16 | 0,1539 | 0,0061 | 0,0339 |
| -0,86 | 0,2 | 0,1949 | 0,0051 | 0,0349 |
| -0,81 | 0,24 | 0,2090 | 0,0310 | 0,0090 |
| -0,65 | 0,28 | 0,2578 | 0,0222 | 0,0178 |
| -0,3 | 0,32 | 0,3821 | 0,0621 | 0,1021 max |
| -0,26 | 0,36 | 0,3974 | 0,0374 | 0,0774 |
| -0,17 | 0,4 | 0,4325 | 0,0325 | 0,0725 |
| -0,06 | 0,44 | 0,4761 | 0,0361 | 0,0761 |
| -0,02 | 0,48 | 0,4920 | 0,0120 | 0,0520 |
| -0,01 | 0,52 | 0,4960 | 0,0240 | 0,0160 |
| 0,21 | 0,56 | 0,5832 | 0,0232 | 0,0632 |
| 0,27 | 0,6 | 0,6064 | 0,0064 | 0,0464 |
| 0,28 | 0,64 | 0,6103 | 0,2944 | 0,0103 |
| 0,37 | 0,68 | 0,6443 | 0,357 | 0,0043 |
| 0,44 | 0,72 | 0,6700 | 0,0500 | 0,0100 |
| 0,60 | 0,76 | 0,7257 | 0,0343 | 0,0057 |
| 0,77 | 0,80 | 0,7294 | 0,0206 | 0,0194 |
| 0,88 | 0,84 | 0,8106 | 0,0294 | 0,0106 |
| 1,00 | 0,88 | 0,8413 | 0,0387 | 0,0013 |
| 1,25 | 0,92 | 0,8944 | 0,0256 | 0,0144 |
| 1,43 | 0,96 | 0,9236 | 0,0304 | 0,0036 |
| 1,68 | 1,00 | 0,9535 | 0,0465 | 0,0065 |

* З цього видно, що максимальне відхилення між емпіричним та гіпотетичним розподілами є 0,1021, тобто D25= 0,1021
* Воно значно менше від критичного значення при рівні значущості α=0,05, тому гіпотезу приймаємо. Ми довели гіпотезу для yi, але вона – лінійне перетворення. Тому ми довели нашу початкову гіпотезу.

Двовибірковий критерій погодженості (Крит.Смірнова)

* Нехай буде ряд незалежних спостережень над неперервною статистичною змінною ξ1, а - над неперервною статистичною змінною ξ2. Потрібно перевірити,гіпотезу про те, що популяції з двох взятих вибірок є однаково розподілені:
* 
* На основі вибірки знаходимо емпіричну форму розподілу
* 
* Розглянемо статистику



* М.В Смірнов (1933) довів, що статистика  має розподіл незалежний від неперервних гіпотетичних розподілів F і G, який при  збігається з розподілом Колмогорова K(x)
* Критерій Смірнова стосується випадку великих вибірок. Для малих m,n (2<= m,n <=50) існують окремі таблиці розподілу Смірнова
* Приклад: Дано 2 незалежні вибірки незалежних спостережень над абсолютно неперервною випадковою змінною.
* (x): 1,9 3,2 0,4 1,2 1,1 1,0 1,5 2,7 0,6 2,0
* (y): 0,6 3,0 2,3 2,1 1,2 2,6 1,9 1,0 2,4 2,7
* H: Обидві множини спостережень взято у одинаково розподілених абсолютно неперервних популяцій 
* Запишемо для обох рядів спостережень спільний варіаційний ряд значень обох емпіричних функцій розподілу в точках спільного варіаційного ряду та абсолютну різницю в цих точках між розподілами.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  |  |  |  |  |
| 0,4 |  | 0,1 | 0 | 0,1 |
| 0,6 | 0,6 | 0,2 | 0,1 | 0,1 |
| 1,0 | 1,0 | 0,3 | 0,2 | 0,1 |
| 1,1 |  | 0,4 | 0,2 | 0,2 |
| 1,2 | 1,2 | 0,5 | 0,3 | 0,2 |
| 1,5 |  | 0,6 | 0,3 | 0,3 |
| 1,9 | 1,9 | 0,7 | 0,4 | 0,3 |
| 2,0 |  | 0,8 | 0,4 | 0,4 max |
|  | 2,1 | 0,8 | 0,5 | 0,3 |
|  | 2,2 | 0,8 | 0,6 | 0,2 |
|  | 2,4 | 0,8 | 0,7 | 0,1 |
|  | 2,6 | 0,8 | 0,8 | 0 |
| 2,7 | 2,7 | 0,9 | 0,9 | 0,1 |
|  | 3,0 | 0,9 | 1 | 0,1 |
| 3,2 |  | 1 | 1 | 0 |



* Для рівня значущості α=0,05, Sкр=1,36



* Гіпотезу приймаємо
* Приклад: Дано 2 незалежні вибірки незалежних спостережень над абсолютно неперервною випадковою змінною.
* (x): 0,4 -0,5 1,7 0,0 -1,1 1,2 -0,3 -0,9 -0,4 0,5
* (y): -0,9 1,1 1,5 0,4 0,8 -0,5 0,6 0,9 -0,3 1,2
* H: Обидві множини спостережень взято у однаково розподіленої абсолютно неперервної генеральної сукупності

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  |  |  |  |  |
| * -1,1 |  | * 0,1 | * 0 | * 0,1 |
| * -0,9 | * -0,9 | * 0,2 | * 0,1 | * 0,1 |
| * -0,5 | * -0,5 | * 0,3 | * 0,2 | * 0,1 |
| * -0,4 |  | * 0,4 | * 0,2 | * 0,2 |
| * -0,3 | * -0,3 | * 0,5 | * 0,3 | * 0,2 |
| * 0,0 |  | * 0,6 | * 0,3 | * 0,3 |
| * 0,4 | * 0,4 | * 0,7 | * 0,4 | * 0,3 |
| * 0,5 |  | * 0,8 | * 0,4 | * 0,4 max |
|  | * 0,6 | * 0,8 | * 0,5 | * 0,3 |
|  | * 0,8 | * 0,8 | * 0,6 | * 0,2 |
|  | * 0,9 | * 0,8 | * 0,7 | * 0,1 |
|  | * 1,1 | * 0,8 | * 0,8 | * 0 |
| * 1,2 | * 1,2 | * 0,9 | * 0,9 | * 0 |
|  | * 1,5 | * 0,9 | * 1 | * 0,1 |
| * 1,7 |  | * 1 | * 1 | * 0 |



* Для рівня значущості α=0,05, Sкр=1,36



* Гіпотезу приймаємо