# **Лекция 10. Определение объема выборки при данных ошибках.**

**Пример.** Найдем объем выборки, обеспечивающий при решении задачи о различении гипотез о параметре показательного закона ошибки и .

Согласно критерию Неймана-Пирсона . Тогда при больших

Решаем систему:

Графическая иллюстрация критерия отношения правдоподобия

Вывод может, таким образом, зависеть от

# **Критерий Вальда.**

Вальд предложил последовательный критерий отношения правдоподобия:

Рассмотрим задачу проверки гипотезы

выборка из известного закона распределения с параметром

: - простая альтернатива

Идея: найти такие границы и , чтобы

Положим , то есть статистикой критерия будет последовательность

Критерий: если , то принимается

Если , то принимается

Тогда - ошибка первого рода,

- ошибка второго рода

# **Геометрическая иллюстрация критерия Вальда**

Как найти и ?

**Теорема 1**. (При любой гипотезе момент конечен!)

**Теорема 2.** Для критерия Вальда имеют место неравенства:

**Замечание**. Система неравенств при фиксированных значениях изображается заштрихованным множеством. Следовательно, выбирая для заданных величин и граничные значения и , заключаем, что сумма ошибок

и удовлетворяет неравенству

(см. рисунок)

# **Средний объем испытаний**

Найдем математическое ожидание , имеющее смысл среднего числа испытаний до принятия решения.

Рассмотрим функцию и статистики

Тогда имеет место

**Теорема 3**. Если

то

Найдем и .

Приближенно случайная величина принимает значения и

Если истинна , то

Если истинна , то

Следовательно,

**Пример**. Построим критерий Вальда для различения гипотез

Для параметра показательного закона с ошибками и .

По формулам Вальда

Найдем среднее число испытаний.

Среднее число испытаний при условии :

Среднее число испытаний при условии :

**Замечание**. Таким образом, среднее число испытаний в любом случае примерно в два раза меньше, чем для критерия Неймана-Пирсона.

# **Критерий согласия Пирсона. Проверка гипотезы о виде распределения**.

Рассмотрим выборку из некоторого закона распределения и гипотезы

выборка из распределения с законом распределения (c плотностью )

выборка не подчиняется этому закону

Переходим к группированной выборке.

В дискретном случае:

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Значения |  |  |  |  |
| Эмпирические  частоты |  |  |  |  |
| Теоретические  частоты |  |  |  |  |

Во второй строке представлены эмпирические частоты выпадения значений из верхней строки; в третьей теоретические частоты при гипотезе

В непрерывном случае:

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Интервалы |  |  |  |  |
| Эмпирические  частоты |  |  |  |  |
| Теоретические  частоты |  |  |  |  |

Во второй строке представлены частоты попадания значений в интервалы из верхней строки; в третьей теоретические частоты при гипотезе

Вычисляется статистика . Ее распределение при больших приближенно равно , где – число параметров распределения, оцениваемых по выборке.

**Теорема Пирсона для простой гипотезы**. При справедливости основной гипотезы статистика , .

**Пример 1**. Имеется выборка объема

: выборка взята из равномерного распределения

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  |  | 2 |  | 4 | 5 | 6 | 7 | 8 | 9 |  |
|  | 11 | 12 |  | 13 | 7 | 10 | 11 | 9 | 14 | 8 |
|  | 10,4 | 10,4 | 10,4 | 10,4 | 10,4 | 10,4 | 10,4 | 10,4 | 10,4 | 10,4 |

Вычисляем

По таблице находим

Поскольку гипотеза принимается

**Пример 2**. Опыт Менделя 556 горошин, полученных при скрещивании во втором поколении

: Гипотеза Менделя о доминантных признаках согласуется с практикой

|  |  |  |  |
| --- | --- | --- | --- |
| Признаки | Частоты | Вероятность |  |
| Круглые и желтые | 315 | 9/16 | 312,75 |
| Круглые и зеленые | 108 | 3/16 | 104,25 |
| Морщинистые и желтые | 101 | 3/16 | 104,25 |
| Морщинистые и зеленые | 32 | 1/16 | 34,75 |

Гипотеза Менделя подтверждается

**Замечание**. Если , то этот столбец объединяется с соседним.

# **Лекция 11**. Непараметрические гипотезы. Критерий согласия Пирсона. Проверка гипотезы о виде распределения.

Рассмотрим выборку из некоторого закона распределения и гипотезы

выборка из распределения с законом распределения (c плотностью )

выборка не подчиняется этому закону

Переходим к группированной выборке.

В дискретном случае:

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Значения |  |  |  |  |
| Эмпирические  частоты |  |  |  |  |
| Теоретические  частоты |  |  |  |  |

Во второй строке представлены эмпирические частоты выпадения значений из верхней строки; в третьей теоретические частоты при гипотезе

В непрерывном случае:

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Интервалы |  |  |  |  |
| Эмпирические  частоты |  |  |  |  |
| Теоретические  частоты |  |  |  |  |

Во второй строке представлены частоты попадания значений в интервалы из верхней строки; в третьей теоретические частоты при гипотезе

## Теорема Пирсона для простой гипотезы.

Рассмотрим некоторую выборку и проверим гипотезу

выборка подчиняется распределения закону распределения с функцией . Тогда при справедливости основной гипотезы статистика в пределе распределена по закону :

**Доказательство**.

При справедливости основной гипотезы вектор имеет полиномиальное распределение: где

Заметим, что

Найдем характеристическую функцию вектора : (столбец)

Далее определим случайный вектор , положив .

Его характеристическая функция может быть выражена через следующим образом:

Исследуем асимптотическое поведение при больших значениях

где -ортогональная матрица,

-первый элемент в столбце .

Тогда для случайного вектора получим:(– столбцы)

Следовательно,

Таким образом, характеристическая функция вектора имеет вид:

, что означает, что в пределе вектор нормален с нулевым математическим ожиданием и матрицей ковариаций , то есть .

Следовательно, в пределе подчиняется распределению . Теорема доказана.

## Теорема Пирсона для сложной гипотезы

Пусть закон распределения генеральной совокупности зависит от параметра , пусть

Тогда если существуют непрерывные производные и , , а матрица имеет ранг , то статистика где , - ОМП, в пределе распределена по закону :

**Выводы**:

Большие значения статистики говорят о расхождении практики с предположением . Критическое множество имеет вид: .

**Критерий**

гипотеза отклоняется

гипотеза принимается

**Пример**. Бомбардировки Лондона, на 576 участков по 0,25 упало 537 снарядов. Проверить гипотезу о пуассоновском распределении числа снарядов, упавших на участок.

**Решение**. Найдем параметр пуассоновского закона: , вычисляем теоретические вероятности и статистику (два последних интервала при этом объединяются, чтобы , что влечет уменьшение значения )

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  |  | 1 |  | 3 |  | 7 |
|  | 229 | 211 |  | 35 | 7+1 | |
|  | 226,74 | 211,39 | 98,54 | 30,62 | 8,71 | |

Находим

Поскольку – основная гипотеза принимается.