**Лекция 6.**

**Интервальные оценки параметров выборки:** *Понятие доверительного интервала, уровень значимости, надежность оценки, доверительный интервал для математического ожидания нормально распределенной случайной величины при известной и при неизвестной дисперсии, доверительный интервал для вероятности случайного события в схеме Бернулли при большом объеме выборки.*

Доверительный интервал для а при неизвестном . Прежде всего докажем важное свойство выборочного среднего, дано .



и дисперсии



для выборки  (1) из нормального распределения .

Теорема 1. Статистики  и  для выборки (1) из нормального распределения независимы. Случайная величина  имеет распределение с (n-1)-й степенью свободы.

Доказательство. Случайные величины  независимы и имеют нормальное распределение с параметрами (0,1). Обозначим

, .

Тогда  , . Доказать, что  и  независимы и что  имеет распределение с (n-1)-й степенью свободы . Случайный вектор  имеет сферическое нормальное распределение с плотностью

. (2)

Пусть -ортогональное преобразование , заданное соотношением

 (3)

 k=1.,2,…,n .

Всегда можно подобрать коэффициенты  так, чтобы равенства (3) задавали ортогональное преобразование. Тогда  также будут иметь сферическое нормальное распределение с плотностью (2) . Так как  и  (из-за ортогональности преобразования C ), то

,

поэтому  имеет распределение  с (n-1)-й степенью свободы . Теорема доказана .

Теорема 2. Пусть (1)-выборка из нормального распределения . Статистика

, (4)

называемая отношением Стьюдента , имеет распределение Стьюдента с

(n-1)-степенью свободы.

Распределение Стьюдента. Доверительный интервал для математического ожидания при неизвестных а и σ2.

Пример распределения Стьюдента. Найти функцию распределения частного , где  и  – независимые величины, причем  распределено по нормальному закону

, а  ,

.

Согласно формуле



 .



Сделав замену



Находим

.

Плотность распределения вероятностей



носит название закона Стьюдента. (Стьюдент – псевдоним английского статистика Госсета, впервые нашедшего этот закон эмпирическим путем.)

При n=1 закон Стьюдента превращается в закон Коши.

Пример 7. Поворот осей координат. По функции распределения двумерной случайной величины  найти функцию распределения величин

, .

Доверительный интервал для математического ожидания при неизвестных а и σ2.

В случае, когда параметры *а* и  неизвестны, для построения доверительных интервалов используется величина , где  определено формулой . Распределение величины  не зависит от параметров *а* и  нормального распределения, соответствующего выборке; плотность распределения  определяется формулой , где нормирующая постоянная .Распределение величины  называется *распределением Стьюдента с п-1 степенями свободы*. Определим величину как решение уравнения  Отсюда, положив , получим доверительный интервал для *а* :

. (9)

Приведем небольшую табличку значений :

  5 10 20 30 

0,05 2,571 2,228 2,086 2,042 1,960

0,01 4,032 3,169 2,845 2,750 2,576

При  плотность  сходится кплотности нормально распределенияпараметрами (0,1).Так как при любом конечном n плотность  при 

убывает медленнее плотности нормального распределения, то . Это приводит к тому, что средняя длина доверительного интервала (9) больше длины доверительного интервала (6). При построении доверительных интервалов (6), (9), предполагалось, что элементы выборок имеют нормальное распределение. Пусть теперь -произвольная выборка. При больших n можно построить простые приближенные доверительные интервалы без предположения нормальности . Найдем, например, доверительный интервал для параметра .Пусть.Тогда по центральной предельной теореме распределение величины  близко к нормальному распределению с параметрами (0,1). Отсюда, используя состоятельность оценки параметра, можно показать, что



или

 (10)

при .Доказательство этого утверждения приводить не будем.

Доверительный интервал для центра распределения при неизвестном σ .

Перейдем теперь к случаю, когда оба параметра , ν и σ , неизвестны и должны оцениваться по выборке . Как мы видели величина критерия  распределяется по закону Стьюдента с (n-1) степенями свободы . Поэтому для вероятности  мы можем , пользуясь таблицей ( Значения q-процентных пределов tq,k в зависимости от k степеней свободы и от вероятности  для распределения Стьюдента ) , найти q% пределы ±tq,n-1 , такие что

 (1) или



при любом ν . Следовательно , интервал  будет доверительным интервалом , отвечающий доверительной вероятности  .

Пусть требуется построить 99%-ый доверительный интервал для оценки генерального среднего диаметра ν валика по «пробе» из 10 деталей , обработанных на токарном автомате , если отклонения размеров этих деталей от середины поля допуска оказались следующими ( см. таблицу 1)

# Таблица 1

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| № деталей по порядку | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
| Отклонения размеров в *мк* | +2 | +1 | -2 | +3 | +2 | +4 | -2 | +5 | +3 | +4 |

Находим , что  *мк , s=2.3 мк .* Из (1) имеем :

,

откуда 1- P = Q = 1-0.99 = 0.01 = 1% .

По таблице приложений V при q=1% и k=n-1=10-1=9 находим t1;9=3.25 . Далее *мк* ,

## откуда 2-2,49<ν<2+2,49

## или 0,49<ν<4,49 (2)

## Таким образом , согласующиеся с нашими опытными данными или , иными словами , «допустимые» (с надежностью в 99%) значения параметра ν лежат в интервале (-0,49 *мк* ; 4,49 *мк*) .

Заметим, что если бы мы приняли найденное значение , s*=2.3 мк* за значение параметра σ и исходили бы из нормального закона при оценке , то «классические» 99%-ые доверительные границы были бы значительно уже . В самом деле , вместо t1;9=3,25 мы имели бы по таблице II (Нормирования функции Лапласа  ) Z1=2,576 и  и вместо неравенства (2) мы получили бы 0,12<ν<3,88 .

Таким образом, мы значительно преувеличили бы действительную точность нашей оценки.

Следует подчеркнуть , что получение при малом числе наблюдений с помощью закона Стьюдента более широкого доверительного интервала (или , иначе говоря , более широкой зоны неопределенности ) при оценке центра рассеивания , нежели при использовании нормального закона , вовсе не является недостатком метода Стьюдента , а связано с существом дела : ведь мы предположили , что никакой дополнительной информации относительно величины параметра σ , кроме той , которую дает выборка , мы не имеем . В этом случае мы не должны для построения доверительного интервала для центра рассеивания применять нормальный закон , так как тем самым мы преувеличим действительную точность нашей оценки , как это видно из приведенного примера .

Заметим, что в некоторых случаях экспериментатор хотя и не располагает точным значением σ , однако может с уверенностью гарантировать верхнюю границу для этого параметра . Если ее принять за σ и определить по нормальному закону доверительный интервал для центра рассеивания , то при малом числе наблюдений этот интервал иногда оказывается все же более узким , чем по методу Стьюдента , и тем самым несколько уточняется оценка центра рассеивания.

Доверительный интервал для σ. Предполагая выборку произведенной из нормальной совокупности N(x;ν;σ) , построим теперь доверительный интервал для параметра σ и σ2 . Отправным пунктом будет при этом положение о том , что величина  распределена по законуχ2 с (n-1) степенями свободы . Заметим , что для данного уровня доверительной вероятности  мы можем вообще построить бесконечным числом способов интервал , которому отвечала бы вероятность , в точности равная  . Целесообразно выбрать в данном случае два предела χ12  иχ22 так , чтобы



и, следовательно ,

 (3)

Из (3) следует 

Или , что равносильно 

Таким образом , интервал  есть доверительный интервал для оценки σ2 , а интервал  - для оценки σ с доверительной вероятностью