**Тема 1. Нахождение оценок параметров распределения. Построение доверительных интервалов**

***Задание 1.1***

При оценке свойств сахарной свеклы было обследовано *n* проб и получены следующие значения содержания сахара *X* %.

Требуется:

1. определить выборочную среднюю *Х*в, выборочную и исправленную дисперсии.
2. полагая, что распределение признака X описывается законом нормального распределения, найдите доверительный интервал для среднего содержания сахара в обследуемой партии свеклы на уровне заданной надежности *γ*.

Исходная информация в таблице.

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **№ варианта** | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|  | 15,5 | 12,4 | 12,3 | 14,8 | 11,6 | 15,4 | 15,1 | 16,1 | 14,1 | 14,2 |
|  | 16,0 | 14,1 | 11,0 | 11,5 | 14,5 | 15,8 | 15,5 | 17,1 | 15,1 | 13,6 |
|  | 14,6 | 13,2 | 11,8 | 12,1 | 13,8 | 15,1 | 143 | 17,2 | 12,2 | 15,2 |
| **Значения** | 14,9 | 14,0 | 13,2 | 13,4 | 12,0 | 14,9 | 16,6 | 16,9 | 13,8 | 13,0 |
| ***X*i** | 15,3 | 14,1 | 14,0 | 13,1 | 13,8 | 14,5 | 16,2 | 15,2 | 14,1 | 15,9 |
|  | 14,0 | 14,0 | 14,1 | 11,9 | 14,2 | 13,5 | 16,1 | 14,3 | 15,9 | 15,2 |
|  | 14,3 | 14,6 | 12,0 | 12,8 | 13,4 | 13,0 | 15,4 | 15,5 | 16,8 | 16,5 |
|  | 15,1 | 14,0 | 13,6 | 13,5 | 14,2 | 13,2 | 14,9 | 16,2 | 12,5 | 13,9 |
|  | 15,4 | 13,4 | 11,3 | 14,2 | 14,5 | 13,3 | 15,5 | 14,1 | 13,4 | 14,2 |
|  | 16,0 | 12,2 |  | 12,9 | 11,0 | 14,0 |  | 15,0 | 16,0 | 15,3 |
|  | 15,7 | 12,3 |  |  |  | 14,1 |  |  |  | 16,4 |
|  |  |  |  |  |  |  |  |  |  | 15,0 |
| ***n*** | 11 | 11 | 9 | 10 | 10 | 11 | 9 | 10 | 10 | 12 |
| ***γ*** | 0,95 | 0,99 | 0,999 | 0,95 | 0,99 | 0,999 | 0,95 | 0,99 | 0,999 | 0,99 |

***Задание 1.2***

Для оценки средней загрязненности сточных вод было взято 24 пробы по 1л и получены следующие статистические данные о содержании в них вредных примесей (в мг):

|  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| № | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 |
| масса | 40+N | 28 | 62 | 54+N | 46 | 52 | 72 | 53 | 60+N | 56 | 34 | 54 |
| № | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 | 21 | 22 | 23 | 24 |
| масса | 57+N | 46 | 71 | 50+N | 41 | 38 | 55 | 42 | 60+N | 49 | 44 | 31 |

а) Представить эти данные в виде интервального ряда распределения с шагом 10мг и построить гистограмму относительных частот.

б) На основании этих данных дать интервальную оценку средней массы примесей с доверительной вероятностью 0,95 (считать, что примеси распределены практически нормально).

***Задание 1.3***

1. Найти минимальный объем выборки, при котором с надежностью *γ* = 0,95 точность *ε* оценки математического ожидания нормально распределенной генеральной совокупности по выборочной средней равна 0,N, если известно, что среднее квадратическое отклонение *σ* = N+1,5 генеральной совокупности.

2. Выборка из большой партии электроламп содержит 200+N дамп. Средняя продолжительность горения лампы выборки оказалась равной 1000+N часам. Найти с надежностью 0,99 доверительный интервал для средней продолжительности горения лампы всей партии, если известно, что среднее квадратическое отклонение продолжительности горения 50+N часов. Предполагается, что продолжительность горения ламп распределена нормально.

***Образец выполнения задания 1.1***

При оценке свойств винограда было обследовано 10 проб и получены следующие значения содержания углеводов *Y* (г/100 г): 17,1; 16,2; 18,3; 16,6; 15,3; 18,6; 16,4; 16,5; 17,5; 16,9. Требуется: 1) определить выборочную среднюю , выборочную *D*в и исправленную *S*2 дисперсии; 2) полагая, что распределение признака *Y* описывается нормальным законом распределения, найдите доверительный интервал для среднего содержания *а* углеводов в обследуемой партии винограда на уровне заданной надежности γ = 0,999.

***Решение.***

1. Определим выборочную среднюю , выборочную *D*в и исправленную *S*2 дисперсии. Вспомогательные расчеты проведем в таблице.

Выборочная средняя .

Выборочная дисперсия

*D*в.

Исправленная дисперсия

*s*2.



1. Полагая, что распределение признака *Y* описывается нормальным законом распределения, найдём доверительный интервал для среднего содержания *а* углеводов в обследуемой партии винограда на уровне заданной надежности γ = 0,999. Доверительный интервал найдём по формуле:

, где . по таблице по заданным *n*=10 γ = 0,999 находим *t*γ=4,78.

Итак,  и доверительный интервал



Ответ: Выборочная средняя ; выборочная дисперсия *D*в; исправленная дисперсия *s*2.

С надёжностью 0,999 среднее содержание *а* углеводов в обследуемой партии винограда заключено в доверительном интервале 15,45 < *а* < 18,43.

***Образец выполнения задания*** *1*.***2***

Для оценки содержания минеральных веществ в артезианском источнике было взято 16 проб воды по 1 л и получены следующие статистические данные (в мг/л):

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| № | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 |
| масса | 41 | 35 | 61 | 53 | 44 | 51 | 70 | 52 |
| № | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 |
| масса | 43 | 57 | 32 | 55 | 58 | 47 | 72 | 48 |

а) Представить эти данные в виде интервального ряда распределения с шагом 10мг и построить гистограмму частот.

б) На основании этих данных дать интервальную оценку средней массы примесей с доверительной вероятностью 0,95.

***Решение.***

а) Представим эти данные в виде статистического ряда распределения с шагом 10 мг. Для этого возьмем в качестве середин интервалов значения 30, 40, 50, 60, 70.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Частичные интервалы | 25-35 | 35-45 | 45-55 | 55-65 | 65-75 |
| Середина интервала (мг/л) | 30 | 40 | 50 | 60 | 70 |
| Попавшие варианты | 32 | 41, 35, 44, 43 | 53, 51, 52, 47,48 | 61, 55, 57,58 | 70, 72 |
| Колич. набл. (*ni*) | 1 | 4 | 5 | 4 | 2 |
| *ni* /Δ*xi* | 0,1 | 0,4 | 0,5 | 0,4 | 0,2 |

Найдём высоты столбцов гистограммы по формулам *ni*/Δ*xi*, где *ni* частота попадания в интервал длиной Δ*xi* (у нас все интервалы имеют длину 10). Строим гистограмму частот

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| 0,5 |  |  |  |  |  |
| 0,4 |  |  |  |  |  |
| 0,3 |  |  |  |  |  |
| 0,2 |  |  |  |  |  |
| 0,1 | 30 | 40 | 50 | 60 | 70 |

Найдем среднее значение выборки:

(30+160+250+240+140)/16=820/16=51,25.

Среднее значение квадрата:

(900+6400+12500+14400+9800)/16=44000/16=2750

Дисперсия: σ2=–()2=2750–(51,25)2=123,438.

Среднее квадратическое отклонение: σ===11,11.

Исправленное среднее квадратическое отклонение: *s*=σ=11,11=11,474.

Исправленное среднее квадратическое отклонение средней выборочной: =2,868.

Коэффициент Стьюдента: =2,131.

Полуширина доверительного интервала: Δ*x*= ⋅ =2,131⋅2,868=6,113≈6

**Ответ***:* доверительный интервал для среднего значения 51,25±6 (мг/л).

***Образец выполнения задания 1.3***

1. Найти минимальный объем выборки, при котором с надежностью *γ* = 0,925 точность *ε* оценки математического ожидания нормально распределенной генеральной совокупности по выборочной средней равна 0,2, если известно, что среднее квадратическое отклонение *σ* = 1,5 генеральной совокупности.

***Решение.***

Точность оценки , если дано *σ* = 1,5, поэтому , *ε* = 0,2, , *Ф*(*t*) = 0,4625 –> *t* = 1,78



2. Выборка из большой партии электроламп содержит 100 дамп. Средняя продолжительность горения лампы выборки оказалась равной 1000 часам. Найти с надежностью 0,95 доверительный интервал для средней продолжительности горения лампы всей партии, если известно, что среднее квадратическое отклонение продолжительности горения 40 часов. Предполагается, что продолжительность горения ламп распределена нормально.

***Решение.***

Запишем данные задачи в принятых обозначениях. *Х* – продолжительность горения ламп, *Х**N*(*a*,40), *γ* = 0,95, *n* = 100,  = 1000,

*а* – средняя продолжительность горения ламп всей партии (генеральной совокупности).



По таблице функции  находим *t = 1,96*



*lγ* = (1000 – 7,84; 1000 + 7,84)

*lγ* = (992,16; 1001,84)

Итак, с надежностью 95% можно утверждать, что неизвестный параметр а покрывается интервалом *lγ*, т.е.

992,16 < *а* < 1007,84.

**Тема 2. Методы расчёта сводных характеристик выборки**

***Задание 2.1***

**1.** Случайная величина *Х* – отклонение контролируемого размера изделия от номинала *Х**N*(*a*,*σ*). Приведено эмпирическое распределение отклонений 200 изделий: в первой строке указано отклонение *xi* (мм), во второй строке приведена частота *ni*– количество изделий, имеющих отклонение:

|  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| *xi* | 3+N | 5+N | 7+N | 9+N | 11+N | 13+N | 15+N | 17+N | 19+N | 21+N | 23+N |
| *ni* | 6 | 9 | 26 | 25 | 30 | 26 | 21 | 24 | 20 | 8 | 5 |

а) Найти методом моментов точечные оценки неизвестных параметров *а* и *σ* нормального распределения.

б) Найти: , *Dв*, *σв*, используя условные варианты.

***Задание 2.2***

С целью определения рациональной структуры размерного ассортимента детской одежды проведено выборочное обследование определённых половозрастных групп детского населения и получено следующее распределение количества детей по величине обхвата груди *Х*.

Требуется:

1. Построить гистограмму относительных частот для наблюдаемых значений признака *Х*.
2. Определить выборочную среднюю, выборочную и исправленные дисперсии.
3. Используя полученные результаты и полагая, что распределение признака *Х* подчинено нормальному закону, найдите:
4. доверительный интервал для ожидаемого среднего значения обхвата груди на уровне надёжности *γ*;
5. вероятность того, что величина признака *Х* у выбранного наугад из данной группы ребёнка окажется в пределах от *α* до *β* см.

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Обхват груди  *Х*, см | Количество детей для варианта № | | | | | | | | | |
| 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
| 56-58 | 13 |  |  |  | 28 |  |  |  | 35 | 5 |
| 59-61 | 28 |  | 27 |  | 48 | 16 | 33 |  | 52 | 18 |
| 62-64 | 40 | 11 | 40 | 42 | 70 | 47 | 54 | 21 | 71 | 57 |
| 65-67 | 47 | 37 | 44 | 51 | 78 | 64 | 66 | 36 | 67 | 49 |
| 68-70 | 41 | 64 | 48 | 67 | 36 | 54 | 55 | 78 | 44 | 30 |
| 71-73 | 21 | 74 | 34 | 74 | 20 | 37 | 30 | 69 | 31 | 21 |
| 74-76 |  | 30 | 17 | 54 |  | 22 | 12 | 49 |  |  |
| 77-79 |  | 14 |  | 32 |  |  |  | 27 |  |  |
| Итого | 190 | 230 | 210 | 320 | 280 | 240 | 250 | 280 | 300 | 180 |

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Номер варианта | **1** | **2** | **3** | **4** | **5** | **6** | **7** | **8** | **9** | **10** |
| *γ* | 0,95 | 0,99 | 0,95 | 0,99 | 0,95 | 0,99 | 0,999 | 0,999 | 0,99 | 0,95 |
| *α*, см | 66 | 62 | 72 | 69 | 58 | 64 | 63 | 67 | 60 | 65 |
| *β*, см | 70 | 66 | 78 | 77 | 66 | 68 | 69 | 72 | 65 | 71 |

***Образец выполнения задания 2.1***

1. Случайная величина *Х* – отклонение контролируемого размера изделия от номинала *Х**N*(*a*,*σ*). Приведено эмпирическое распределение отклонений 200 изделий: в первой строке указано отклонение *xi* (мм), во второй строке приведена частота *ni*– количество изделий, имеющих отклонение:

|  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| *xi* | 0,3 | 0,5 | 0,7 | 0,9 | 1,1 | 1,3 | 1,5 | 1,7 | 1,9 | 2,1 | 2,3 |
| *ni* | 6 | 9 | 26 | 25 | 30 | 26 | 21 | 24 | 20 | 8 | 5 |

а) Найти методом моментов точечные оценки неизвестных параметров *а* и *σ* нормального распределения.

б) Найти: , *Dв*, *σв*., используя условные варианты.

***Решение.***

**а)** Т.к. требуется оценить два параметра. То можно приравнять моменты первого и второго порядков. Приравняем начальные моменты первого порядка и центральные моменты второго порядка:



Итак, 

Таким образом, нужно вычислить среднюю выборочную и выборочную дисперсию. Составим расчетную таблицу.

|  |  |  |  |
| --- | --- | --- | --- |
| *xi* | *ni* | *xini* |  |
| 0,3 | 6 | 1,8 | 0,54 |
| 0,5 | 9 | 4,5 | 2,25 |
| 0,7 | 26 | 18,2 | 12,74 |
| 0,9 | 26 | 22,5 | 20,25 |
| 1,1 | 30 | 33,0 | 36,30 |
| 1,3 | 26 | 33,8 | 43,94 |
| 1,5 | 21 | 31,5 | 47,25 |
| 1,7 | 24 | 40,8 | 69,36 |
| 1,9 | 20 | 38,0 | 72,20 |
| 2,1 | 8 | 16,8 | 35,28 |
| 2,3 | 5 | 11,5 | 26,45 |
|  | *n* = 200 | 252,4 | 366,56 |

 Dв =,  *а* = 1,262.

**б)** Выберем в качество ложного нуля *С* = *х*5 = 1,1 (с наибольшей частотой), *h* = 0,5 – 0,3 = 0,2, тогда . Составим расчетную таблицу:

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| *xi* | *ni* | *Ui* | *niUi* |  | *ni*(*Ui* + 1)2 |
| 0,3 | 6 | – 4 | – 24 | 96 | 54 |
| 0,5 | 9 | – 3 | – 27 | 81 | 36 |
| 0,7 | 26 | – 2 | – 52 | 104 | 26 |
| 0,9 | 25 | – 1 | – 25 | 25 | 0 |
| 1,1 | 30 | 0 | 0 | 0 | 30 |
| 1,3 | 26 | 1 | 26 | 26 | 104 |
| 1,5 | 21 | 2 | 42 | 84 | 189 |
| 1,7 | 24 | 3 | 72 | 216 | 384 |
| 1,9 | 20 | 4 | 80 | 320 | 500 |
| 2,1 | 8 | 5 | 40 | 200 | 288 |
| 2,3 | 5 | 6 | 30 | 180 | 245 |
|  | *n* = 200 | 162 | = 1332 |  |  |

Контроль: 

1865 = 1332 + 2·162 + 200; 1856 = 1856

Вычисляем искомые величины:



 = 0,81·0,2 + 1,11 = 1,262

Dв = (6,662 – 0,812)·0,22 = 0,238, .

***Образец выполнения задания 2.2***

С целью определения месячного дохода работников банка, была обследована группа специалистов и получено следующее распределение количества работников по величине заработной платы *X* (у.е.).

Требуется: 1) построить гистограмму относительных частот для наблюдаемых значений признака *X*; 2) определить выборочную среднюю, выборочную и исправленную дисперсии; 3) используя полученные результаты и полагая, что распределение признака *X* подчинено нормальному закону, найдите: а) доверительный интервал для ожидаемого среднего значения заработной платы на уровне надежности γ= 0,99;

б) вероятность того, что величина признака *X* у выбранного наугад из данной группы работника окажется в пределах от 196 у.е. до 202 у.е.

|  |  |
| --- | --- |
| Зар. плата  *X*, у.е. | Количество работников |
| 187-189 | 1 |
| 190-192 | 4 |
| 193-195 | 7 |
| 196-198 | 8 |
| 199-201 | 25 |
| 202-204 | 29 |
| 205-207 | 19 |
| 208-210 | 8 |
| Итого | 101 |

***Решение.***

1. Построим гистограмму относительных частот для наблюдаемых значений признака *X*. Для этого перейдём к частичным интервалам, разбив весь промежуток (187,210) на 8 промежутков равной длины . Для каждого из них найдём значения плотности относительных частот  (*i* =1,2,…,*n*). Расчеты проводим в следующей таблице:



Построим гистограмму относительных частот:

187

189,875

192,75

195,625

198,5

201,375

204,25

207,125

210

2. Определим выборочную среднюю, выборочную и исправленную дисперсии. Для упрощения расчетов перейдём к условным вариантам , где *x*i – середина *i*-го интервала, *C* = 203 – ложный нуль (варианта с наибольшей частотой), *h* = 3 – шаг, т.е. разность между любыми

двумя соседними первоначальными вариантами. Расчеты проводим в следующей таблице:



Проверим правильность выполненных расчетов: .

260+2∙(-48)+101=265; 265=265. Вычисления произведены правильно.

Вычислим условные моменты первого и второго порядков: 

Вычислим искомые выборочную среднюю и выборочную дисперсию:





Исправленная дисперсия *s*2.

3. Используя полученные результаты и полагая, что распределение признака *X* подчинено нормальному закону, найдём:

а) доверительный интервал для ожидаемого среднего значения заработной платы на уровне надежности γ = 0,99. Так как среднее квадратическое отклонение не известно, то используем формулу

, где . по таблице по заданным *n*=101 γ = 0,99 находим =2,627.

Итак,  и доверительный интервал



б) вероятность того, что величина признака *X* у выбранного наугад из данной группы работника окажется в пределах от 196 у.е. до 202 у.е., найдем по формуле:

, где *а* ≈ , *α* = 196, *β* = 202, *σ*≈*s*=.4,6202738.



Ответ: Выборочная средняя; выборочная дисперсия ; исправленная дисперсия *s*2=..

С надёжностью 0,99 среднее значение *а* заработной платы работников банказаключено в доверительном интервале 200,37 < *а* < 202,78.

Вероятность того, что величина признака *X* у выбранного наугад из данной группы работника окажется в пределах от 196 у.е. до 202 у.е, равна 0,42.