**БИОМЕТРИЧЕСКАЯ ОБРАБОТКА МАЛЫХ ВЫБОРОК (N˂30)**

**Цель занятия:** приобрести навык обработки малых выборок и расчета основных показателей.

Основными статистическими величинами, определяемыми при обработке данных являются: средняя арифметическая (), показатели изменчивости изучаемых признаков: лимиты (lim), среднее квадратическое отклонение (*δ*), коэффициент изменчивости (Cv), ошибка средней арифметической () и др.

*Средняя арифметическая ()*

Средняя арифметическая () – наиболее употребляемая и распространенная характеристика выборочной совокупности по средней величине признака. Она бывает простой и взвешенной.

В малых выборках, т.е. при n ≤ 30, определение простой средней арифметической величины заключается в суммировании всех значений варьирующего признака и делении полученной суммы на число животных, составляющих эту выборку, т.е.:



где  – средняя арифметическая;– значения признака для каждого члена совокупности (варианты); *n* – общее число членов совокупности (группы); (сигма – греческая прописная буква) – знак суммирования.

Этот простой способ подсчета средней арифметической применяется во всех случаях, когда каждое значение признака входит в сумму одинаковым образом, увеличивая ее на полную величину.

Определение средневзвешенной арифметической производится для характеристики признаков, представляющих собой отношение двух варьирующих величин:



где – значение признака (варианта); – математический вес признака.

Средняя арифметическая и другие средние величины являются очень важными, но явно недостаточными в статистической характеристике выборочной совокупности (группы животных, например), так как любая группа неоднородна и особи, составляющие ее, имеют ряд отличий друг от друга. Эти различия иногда небольшие и малозаметные, а в других случаях они очень велики. Но средняя величина этого разнообразия особей в группе по изучаемому признаку не отражает совершенно.

*Показатели разнообразия (вариации изменчивости)*

В биометрии степень разнообразия в основном принято выражать тремя показателями: лимитами, средним квадратическим отклонением и коэффициентом вариации.

*Лимиты (lim)* – это крайние варианты в группе, т.е. максимальное и минимальное значения признака.

Например, в двух бассейнах имеется по 5 осетров, живой вес которых составляет (кг):

1-й бассейн – 8,40; 8,45; 8,50; 8,55; 8,60; х = 8,50 кг;

2-й бассейн – 8,10; 8,30; 8,50; 8,70; 8,90; х = 8,50 кг.

Средний живой вес осетров в обоих бассейнах одинаков: по 8,50 кг, но разнообразие их по живому весу в первом бассейне значительно меньше, чем во втором. Об этом свидетельствуют лимиты:

lim 1 = 8,40 – 8,60 (разность 0,20 кг);

lim2 = 8,10 – 8,90 (разность 0,80 кг).

Лимиты показывают: а) размах разнообразия; б) максимальное значение признака (живой вес, убойный выход, содержание жира в молоке и т.д.); в) минимальное значение признака. Но лимиты не точно отражают степень разнообразия признака.

К примеру, взяты две подопытные группы рыбы по 10 особей в каждой.

Длина тела рыбы первой группы (см):

40,0; 41,2; 41,8; 43,0; 43,5; 45,5; 45,7; 47,3; 49,0; 48,0.

 = 44,5 см; lim 1 = 40,0 – 49,0 см.

Длина тела рыбы второй группы (см):

40,0; 44,5; 44,5; 44,5; 44,5; 44,5; 44,5; 44,5; 44,5; 49,0.

 = 44,5 см; lim2 = 40,0 – 49,0 см.

Средние арифметические и лимиты в обеих группах одинаковы. Но в 1-ой группе все рыбы имеют различную длину, а во 2-ой – из 10 экземпляров у 8 одинаковая длина. Следовательно, степень разнообразия в 1-ой группе намного больше, чем во 2-ой, но с помощью лимитов выразить это невозможно. Поэтому, помимо лимитов, вычисляют основной показатель разнообразия – среднее квадратическое или стандартное отклонение от средней арифметической величины.

*Среднее квадратическое отклонение* обозначается буквой *δ* (сигма).

При расчете среднего квадратического отклонения небольшого числа наблюдений пользуются формулой:



где– сумма квадратов отклонений, т.е. сумма квадратов разностей между каждой величиной признака и средней арифметической; – значение признака у каждой особи в группе; – средняя арифметическая; n – 1 – число степеней свободы, равное числу признаков без одной.

Среднее квадратическое отклонение высчитывается с точностью на один знак большей, чем средняя арифметическая, и выражается в тех же, что и она, единицах.

Среднее квадратическое отклонение является основным показателем степени разнообразия значений признака в группе. Кроме того, оно используется для вычисления многих других показателей: коэффициента вариации, ошибки средней арифметической, коэффициентов корреляции и регрессии.

Использование среднего квадратического отклонения дает возможность судить о том, сколько и в каких пределах вокруг средней арифметической величины признака размещается особей по величине своего признака. При любом нормальном распределении число особей с величиной признака, отличающегося не более чем на одну сигму ( ± 1*δ*), составляет 68,3 % всех рыб, не более, чем на две сигмы ( ± 2*δ*), - 95,5 % и не более, чем на три сигмы (± 3*δ*), - 99,7 %, т.е. практически весь исследуемый материал.

Знание среднего квадратического отклонения позволяет также установить, относится та или иная особь к данному ряду или нет. Если по изучаемому признаку особь отличается от средней арифметической более, чем на 3 сигмы, то это значит, что она попала в совокупность случайно, т.е. выращена в других условиях или относится к другой породе, подвергалась воздействию тех или иных специфических факторов и т.д. Эти закономерности нормального распределения имеют большое значение в биометрии, так как на их основе построены приемы определения достоверности, полученных в экспериментах, данных.

*Коэффициент изменчивости (разнообразия, вариации)* представляет собой не что иное, как среднее квадратическое отклонение, выраженное в процентах от средней арифметической величины, и рассчитывается по выражению:



Коэффициент изменчивости служит для оценки уравненности вариант в обрабатываемых совокупностях и позволяет сравнивать относительное варьирование признаков, выражаемых в любых единицах измерения.

Если Сv ˂ 10 %, то считают, что изменчивость изучаемого признака в выборочной совокупности незначительная; при Сv = 10 – 20 % изменчивость признака характеризуется как средняя; при Сv ˃ 20 % изменчивость признака значительная, а при Сv ˃ 33 % она настолько велика, что ставит под сомнение достоверность действия изучаемого фактора на формирование анализируемого признака в опытной выборочной совокупности.

*Ошибка средней статистической и достоверность разницы между средними величинами*

При проведении выборочных наблюдений возникает два рода ошибок:

а) организационные (ошибки методики, точности, типичности, ошибки внимания). Этого рода ошибки не могут быть устранены никакими методами математической обработки, и их стремятся свести к минимуму путем тщательного проведения исследований. В частности, чтобы выборка в наибольшей степени представляла генеральную совокупность, т.е. была типичной, в нее отбираются особи из всех частей этой генеральной совокупности;

б) ошибки репрезентативности, т.е. степени соответствия выборки генеральной совокупности. Этот род ошибок не связан с организацией и тщательностью проведения наблюдений. Источником их является сам метод выборок: целое (генеральная совокупность) характеризуется по одной части этого целого – выборке.

*Ошибка средней арифметической* () рассчитывается по формуле:

 (если *n*<30)

Она высчитывается с точностью на один знак большей, чем среднее квадратическое отклонение, и на два знака большей, чем средняя арифметическая; выражается в тех же, что и они, единицах.

Среднюю арифметическую принято записывать вместе с ее ошибкой следующим образом:  ± .При этом ошибка показывает, в каких пределах находится истинная средняя величина.

Выраженная в процентах от средней величины  % = ( / ) · 100 %, она показывает точность, с которой определена средняя величина, и таким образом характеризует точность самого опыта:

при  % ˂ 2 % - точность отменная (отличная);

при  % = 2 – 3 % - точность хорошая;

при  % = 3 – 5 % - точность вполне удовлетворительная;

при  % = 5 – 7 % - точность удовлетворительная;

при  % ˃ 7 % - точность неудовлетворительная.

*Достоверность разницы между средними величинами* **(****).** При исследовании необходимо определить, достоверной ли является полученная разница между двумя средними показателями выборок. Достоверность выборочной разницы измеряется критерием достоверности разницы - . Последний рассчитывается по формуле:



где – критерий достоверности;  – средние арифметические для первой и второй группы;  – ошибки средних арифметических для первой и второй группы.

***Примечание:*** *при определении разности (**) из большей величины вычитают меньшую, т.к. разница всегда положительная.*

– стандартное значение критерия достоверности, определяемое по таблице Стьюдента для каждого порога надежности в зависимости от числа степеней свободы.

Отношение числа благоприятных случаев к числу возможных в биометрии называют вероятностью. Максимально возможная степень вероятности – полное совпадение числа благоприятных и возможных случаев. Вероятность при этом принимается за 1 и составляет 100 %. В биологических исследованиях принято 4 порога вероятности (надежности).

1.Нулевой порог – пониженные требования к надежности, допускаемые в грубоориентировочных исследованиях. Вероятность по этому порогу составляет 0,90 (90 %).

2.Обычные требования к вероятности в большинстве биологических исследований (первый порог вероятности) – Р = 0,95, т.е. надежность прогноза составляет 95 %.

3.Повышенные требования к надежности (второй порог вероятности) при проверочных опытах – Р = 0,99, т.е. надежность прогноза 99 %.

4.Высокие требования к надежности (третий порог вероятности) при проверочных опытах – Р = 0,999, надежность прогноза при такой степени вероятности 99,9 %.

Разница между средними арифметическими двух сравниваемых выборок в том случае достоверна, если критерий достоверности не меньше стандартного значения критерия, найденного по таблице Стьюдента, для первого порога вероятности = 0,95.

В научных публикациях, как правило, указывают не вероятность благоприятных случаев, а так называемый уровень существенности полученных результатов.

Величина Р (уровень существенности) показывает вероятность, с которой проверяемая гипотеза может дать отрицательный результат.

Уровень существенности, равный 0,1, соответствует вероятности 0,90, а уровень существенности 0,05 – вероятности 0,95 и т.д.

Стандартное значение критерия Стьюдента () можно посмотреть в таблице.

Таблица 1 – Стандартные значения критерия достоверности (по Стьюденту) при трех уровнях вероятности

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Число степеней свободы (v) | Уровень вероятности (Р) | | | Число степеней свободы (v) | Уровень вероятности (Р) | | |
| 0,95 | 0,99 | 0,999 | 0,95 | 0,99 | 0,999 |
| 1 | 12,71 | 63,66 | 637,00 | 21 | 2,08 | 2,83 | 2,82 |
| 2 | 4,30 | 9,93 | 31,60 | 22 | 2,07 | 2,82 | 3,79 |
| 3 | 3,18 | 5,84 | 12,94 | 23 | 2,07 | 2,81 | 3,77 |
| 4 | 2,78 | 4,60 | 8,61 | 24 | 2,06 | 2,80 | 3,75 |
| 5 | 2,57 | 4,03 | 6,86 | 25 | 2,06 | 2,76 | 3,73 |
| 6 | 2,45 | 3,71 | 5,96 | 26 | 2,06 | 2,78 | 3,71 |
| 7 | 2,37 | 3,50 | 5,41 | 27 | 2,05 | 2,77 | 3,69 |
| 8 | 2,31 | 3,36 | 5,04 | 28 | 2,05 | 2,76 | 3,67 |
| 9 | 2,26 | 3,25 | 4,78 | 29 | 2,05 | 2,76 | 3,66 |
| 10 | 2,23 | 3,17 | 4,59 | 30 | 2,04 | 2,75 | 3,65 |
| 11 | 2,20 | 3,11 | 4,44 | 35 | 2,03 | 2,72 | 3,59 |
| 12 | 2,18 | 3,06 | 4,32 | 40 | 2,02 | 2,70 | 3,55 |
| 13 | 2,16 | 3,01 | 4,22 | 45 | 2,01 | 2,69 | 3,52 |
| 14 | 2,15 | 2,98 | 4,14 | 50 | 2,01 | 2,68 | 3,50 |
| 15 | 2,13 | 2,93 | 4,07 | 60 | 2,00 | 4,66 | 3,46 |
| 16 | 2,12 | 2,92 | 4,02 | 70 | 1,99 | 2,65 | 3,43 |
| 17 | 2,11 | 2,90 | 3,97 | 80 | 1,99 | 2,64 | 3,42 |
| 18 | 2,10 | 2,88 | 3,92 | 90 | 1,98 | 2,63 | 3,40 |
| 19 | 2,09 | 2,86 | 3,88 | 100 | 1,98 | 2,62 | 3,37 |
| 20 | 2,09 | 2,85 | 3,85 | 120 и выше | 1,96 | 2,56 | 3,29 |

**Задание 1.** Определить среднее квадратическое отклонение по живой массе трехлеток форели (г):

3590 3540 2730 2850 3140 3170 3190 2950

3130 3150 2880 2650 3600 3510 3320 3280

**Задание 2.** Определить достоверность разницы по длине тела молоди рыб двух опытных групп, обработав данные методом малых выборок.

1 группа:

7, 10, 12, 11, 8, 9, 11, 10, 8, 9, 11, 8, 7, 9, 8, 6; *n* = 16.

2 группа:

10, 11, 11, 12, 13, 14, 9, 11, 10, 10, 11, 11, 12, 10, 11, 10; *n* = 16.

При обработке вначале в каждой группе определяете; *δ* ; .

Затем рассчитываете критерий достоверности  разницы между этими двумя группами по длине тела рыб и делаете заключение по вычисленным статистическим величинам.

**Задание 3.** Определить достоверность разницы по длине тушки опытных осетров (см).

1 группа:

15, 20, 26, 18, 22, 24, 17, 19, 16, 23; *n* = 10.

2 группа:

20, 18, 19, 17, 14, 23, 16, 40, 22, 21; *n* = 10.

При обработке вначале в каждой группе определяете; *δ* ; .

Затем рассчитайте критерий достоверности  разницы между этими двумя группами осетров по длине тушки и сделайте заключение по всем вычисленным показателям.

**Задание 4.** Определить достоверность разницы по ихтиомассе рыб контрольной и опытной групп, используя следующие данные:

Таблица 2 – Ихтиомасса рыб, кг

|  |  |
| --- | --- |
| Контрольная группа | Опытная группа |
| 9,5 | 3,4 |
| 12,7 | 1,8 |
| 8,4 | 4,5 |
| 12,0 | 5,6 |
| 10,8 | 6,5 |
| 7,9 | 6,3 |
| 2,8 | 7,6 |
| 10,3 | 8,0 |
| 11,5 | 7,9 |
| 8,2 | 5,0 |
| 7,4 | 4,2 |
| 9,8 | 6,5 |
| 10,9 | 7,3 |
| 9,7 | 6,9 |
| 3,2 | 4,6 |
| 7,6 | 4,5 |
| 3,9 | 5,1 |
| 10,8 | 7,4 |
| 10,8 | 4,7 |
| 14,3 | 6,0 |

Определите вначале  ±  для групп, затем рассчитайте  и сделайте заключение по всем вычисленным показателям.