

2.8 Тема. Ошибки средних величин и выборочной разности

Цель. Знакомство с методами вычисления основных биометрических показателей количественных признаков.

При обработке данных, полученных из документации или из специальных опытов, возможны три типа ошибок: технические (просчеты, описки); ошибки, обусловленные неточностью используемого прибора; статистические. Первые две ошибки называются систематическими и не имеют отношения к биометрии.

Статистические ошибки обусловлены самим статистическим методом, при котором из генеральной совокупности отбирают часть объектов. Ошибки, возникающие при характеристике генеральной совокупности показателями, полученными при изучении выборки, называются ошибками репрезентативности. Величина статистической ошибки зависит от степени изменчивости признака и от числа членов, вошедших в выборку. Чем сильнее изменчивость признака, тем больше статистическая ошибка; чем больше объем выборки, тем меньше ошибка.

Данные ошибки используются для установления доверительных границ в генеральной совокупности, достоверности выборочных показателей и разности, установления объема выборок при научно-исследовательских работах. Статистическую ошибку принято обозначать буквой m с подстрочным значком того параметра, для которого она вычисляется.

Доверительными интервалами называют крайние значения, в пределах которых может находиться параметр генеральной совокупности.

Вариационной статистикой установлено, что средняя арифметическая генеральной совокупности лежит в границах $\pm 3m$ от средней арифметической (тоже σ и S) выборочного исследования; в большинстве же случаев она близка к ней.

Ошибка средней арифметической ($m_{\bar{x}}$) вычисляется по формуле:

$$m_{\bar{x}} = \pm \frac{\sigma}{\sqrt{n}} \quad (23)$$

Пример. Средний суточный удой 100 коров составляет $\bar{x} = 21,26$ кг, а $\sigma = \pm 3,65$. Ошибка средней арифметической в данном случае составит $m_{\bar{x}} = \frac{3,68}{\sqrt{100}} = 0,368$ кг.

Это означает, что средняя ошибка на 100 голов оставляет 0,368 кг. Следовательно, среднесуточные удои изучаемой выборки характеризуются $\bar{X} \pm m = 21,26 \pm 0,368$.

Ошибка среднего квадратического отклонения вычисляется по формуле:

$$m_{\sigma} = \pm \frac{\sigma}{\sqrt{2n}}. \quad (24)$$

Ошибка коэффициента вариации вычисляется по формуле:

$$m_c = \pm \frac{C}{\sqrt{2n}}. \quad (25)$$

В биометрических исследованиях исключительное значение имеет разность – результат вычитания одной величины из другой. В конце всех исследований или наблюдений именно разность и дает ответ на тот вопрос, ради которого производилось исследование.

По разности производится сравнение отдельных животных (растений) или групп животных (растений) между собой и намечается дальнейшее их использование. По разности между признаками потомков и признаками матерей (или других групп животных) определяют качество производителей. По разности между опытной и контрольной группами судят об эффективности опыта и т.д. Поэтому понятно значение достоверности разности, как основной величины, на основании которой выводится окончательное суждение об исследовании.

Вопрос достоверности разности не возникает там, где сравниваются две генеральные совокупности.

Пример. Сравнивается средний удой стада коров в двух хозяйствах за какой-то год: в одном из них он был 3525 кг, в другом 3500 кг, разность (d) = 3500-3525=25 кг. В данном случае не вызывает никакого сомнения, что средний удой коров в первом хозяйстве выше, чем во втором на 25 кг, так как удой подсчитан по всем коровам, какие были в этом году в одном и другом хозяйстве.

Совершенно другое отношение к себе вызывает разность между двумя выборками.

Пример. При установлении разности в удое коров двух пород средний удой породы устанавливается выборочным методом: исследуется часть животных какой-то породы и какое-то определенное время. Каждое среднее, по отношению к животным всей породы, здесь определено с какой-то ошибкой, следовательно, разность между ними может быть ошибочной, и если не учесть этой ошибки, то и вывод будет ошибочный.

Ошибка выборочной разности вычисляется по следующей формуле:

$$m_d = \pm \sqrt{m^2_1 + m^2_2}, \quad (26)$$

где m_d – ошибка выборочной разности;

m_1 – ошибка средней арифметической признака одной группы;

m_2 – ошибка средней арифметической признака другой группы.

В тех случаях, когда между сравниваемыми признаками существует взаимосвязь, то ошибка выборочной разности вычисляется по формуле:

$$m_d = \pm \sqrt{m_1^2 + m_2^2 - 2rm_1m_2},$$

где r – коэффициент корреляции (о взаимосвязях и о вычислении коэффициента корреляции будет описано ниже).

Задание 1. Средняя живая масса дочерей быка в возрасте одного года: $\bar{X}_{\pm m} = 320 \pm 4$ кг. Средняя живая масса их матерей в возрасте одного года $\bar{X}_{\pm m} = 300 \pm 3$ кг. Все ли дочери (а не только исследованные в данный момент) будут иметь живую массу больше своих матерей?

Задание 2. В отобранных случайным способом 50 колосьях двухрядного ячменя были подсчитаны зерна, содержащиеся в каждом колосе. Результаты оказались следующие:

21 17 27 20 22 12 24 13 20 19 22 16 22
 9 21 16 23 16 21 24 18 11 22 15 23 21
 10 15 18 15 21 14 15 18 22 15 17 19 17
 18 17 24 18 19 16 17 15 17 25 16

Распределите эти данные в вариационный ряд, постройте его линейный график и определите среднюю арифметическую, среднее квадратическое отклонение, коэффициент вариации, их ошибки.

Задание 3. Вычислить среднюю арифметическую показателей изменчивости и ошибку средней арифметической по плодовитости свиноматок: 10, 11, 13, 10, 9, 14, 12, 12, 13.

Контрольные вопросы.

1. Какие ошибки называются ошибками выборочного метода?
2. Как вычисляются ошибка средней арифметической, ошибка среднего квадратического, ошибка коэффициента вариации?
3. Приведите формулы вычисления ошибок σ, C_v, r, R .
4. Как изменяется величина $m_{\bar{x}}$ при изменении объема выборки и величины сигму?