Статистика в микробиологических исследованиях.

1.Погрешности известного происхождения, которые могут быть точно определены.

2.Погрешности, природа которых известна, но точное значение не установлено. К их числу относится погрешность измерительных приборов, когда известно не точное значение погрешности прибора, а максимально возможное.

Типы погрешностей

3.Погрешности, о существовании которых исследователь не подозревает, хотя величина их может быть очень значительна. Ошибка в измерениях может возникать при определении концентрации клеток микроорганизма в растворе по оптической плотности раствора, если этот раствор кроме клеток содержал нерастворимые твердые вещества.

Нормальный закон распределения.

Вероятность того, что случайная величина примет значение в пределах бесконечно малого интервала, определяется как φ(x)dx, *где* φ(x)— *некоторая функция, называемая плотностью вероятности.*

$$\varphi(x)=\frac{1}{\sigma\sqrt{2\pi}}\cdot e^{\frac{x-\alpha*\alpha}{2*\sigma*\sigma}}$$

α и σ^2 – параметры распределения.

α – истинное значение измеряемой величины, он определяет расположение центра распределения на числовой оси

σ^2 - дисперсия, он служит мерой рассеивания случайной величины

Генеральная и выборочная совокупность. Расчет среднего значения и дисперсии по выборочным данным.

Если выполнено бесконечно большое число измерений одной и той же величины, то полученную бесконечную гипотетическую совокупность результатов называют *генеральной совокупностью,* α и σ^2 — параметры генеральной совокупности. Истинное значение измеряемой величины α, называемое *генеральным средним,* равно:

$$\alpha = \sum_{i=1}^{x} XiP(Xi)$$

где $p(x_I)$ — вероятность появления результата $x_{\{}$.

Параметр σ^2 представляет собой *генеральную дисперсию*

$$\sigma^2 = \sum_{i=1}^{x} (Xi - \alpha)^2 \cdot P(Xi)$$

Выборочная совокупность характеризуется значениями выборочных параметров, которые являются функциями величин, составляющих выбор очную совокупность. Для обозначения генеральных параметров используют греческие буквы, а выборочных — латинские. Выборочное среднее определяют как среднее арифметическое величины x из n результатов измерений:

$$\overline{X} = \frac{1}{n}\sum_{i=1}^{n} Xi$$

Выборочную дисперсию вычисляют по формуле:

$$S_{x=}^{2} \frac{\sum_{i=1}^{n}(Xi - X)^{z}}{n-1}$$

Необходимое число измерений.

Для определения необходимо, чтобы доверительный интервал, был существенно меньше систематической погрешности, т.е. $\Delta_\gamma \leqslant \delta.$

В кинетических исследованиях выражению «много больше» соответствует отношение сравниваемых величин 10 и больше, т. е. в том случае, когда $\Delta_\gamma < \delta/10$. условие можно считать выполненным.

$$\varepsilon = \frac{\Delta_\gamma}{S} = \frac{\Delta_{\gamma_{\text{отн}}}}{W}$$

Доверительная вероятность γ практически не превышает 0,95. Величина *w,* выраженная в процентах, носит название *коэффициента вариации*

$$w=S_{x}\text{-}/_{x}\text{-}$$

По значениям γ и w по таблице находят необходимое число измерений

Обнаружение промахов.

Для оценки вероятностей β случайного появления выскакивающих значений в ряду *n* измерений (для *n* < 25) на основании результатов была составлена таблица. Для пользования этой таблицей необходимо вычислить из всех измерений, включая подозреваемое $x_к$, которое кажется недопустимо большим или малым. Далее вычисляем относительное отклонение этого измерения от среднего арифметического, выраженное в долях выборочной дисперсии:

$$V=\left|\frac{\bar{x}-xk}{Sn}\right|$$

По таблице находим, какой вероятности β соответствует полученное значение *v*. Если вероятность β подозреваемого измерения < 0,01, то его следует отбросить; соответственно если β>0,1, то подозреваемое измерение следует непременно оставить. При промежуточных значениях (0,1>β>0,01) представляется одинаково правильным любое из этих действий.

Выравнивание эмпирических рядов регрессии.

1. *Графический способ выравнивания.*

Эмпирический ряд регрессии изображается в виде графика в системе прямоугольных координат. Затем визуально намечают срединные точки регрессии, по которым проводят сплошную линию. Но этот способ не исключает влияние индивидуальных свойств исследователя на результаты выравнивания линий.

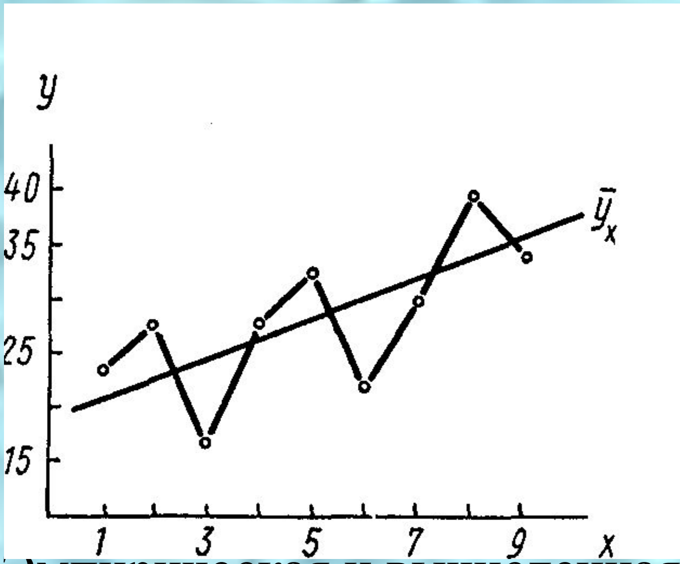

Рисунок 1 Эмпирическая и вычисленная линии регрессии

2. Способ скользящей средней.

Суть этого способа сводится к последовательному вычислению средних арифметических их двух или трех соседних членов эмпирического ряда. Получается выровненный ряд.

3. Метод наименьших квадратов.

Это способ позволяет наиболее точно выравнивать эмпирические ряды. Он основан на предположении, что сумма квадратов отклонений варианта x_i от их средней есть величина ми[illegible] т.е.

$$\sum_{i=1}^{n} (x_i - \bar{x})^2 = \min.$$

Метод наименьших квадратов объективен и универсален

Сопряженность между переменными величинами Y и X можно установить, сопоставляя числовые значения одной из них с соответствующими значениями другой.

Для определения степени сопряженности используют коэффициент корреляции:

$$r_{xy} = \frac{\sum_{i=1}^{n} (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum (x_i - \bar{x})^2 \sum (\bar{y}_i - y)^2}}.$$

Коэффициент корреляции – отвлеченное значение, лежащее в пределах от -1 до +1. При независимом варьировании признаков, когда связь между ними полностью отсутствует, r=0. Чем сильнее сопряженность между признаками, тем выше значение коэффициента корреляции. При положительной связи коэффициент корреляции имеет положительный знак и находится в пределах от 0 до +1. При отрицательной связи в пределах от 0 до -1.