Построение теоретического закона распределения по опытным данным. (описать алгоритм для дискретного и интервального вариационных рядов).

I. **Для дискретного вариационного ряда**

(Дано:

- Генеральная совокупность с неизвестным законом распределения.
- Выборка объёма n: x1,x2,…,xn)

1. **Собрать данные:** x1,x2,…,xn
2. **Построить дискретный вариационный ряд:**
    - Упорядочить значения.
    - Подсчитать частоты ni для каждого значения xi.
    - Вычислить относительные частоты (частости) $\hat{p}_i = \frac{n_i}{n}$.
3. **Выдвинуть гипотезу о законе распределения** (например, Пуассона, биномиальное).

- Формулируем **нулевую гипотезу** H_0.
  Например:

$$H_0 : \text{Генеральная совокупность распределена по закону Пуассона } P(X = m) = \frac{\lambda^m e^{-\lambda}}{m!}.$$

Альтернативная гипотеза H_1: закон распределения не является Пуассона.

4. **Оценить параметры** теоретического распределения:
    - Использовать метод моментов или максимального правдоподобия.
    - - Для Пуассона: $M(X) = \lambda$.
        Метод моментов: $\hat{\lambda} = \bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$.
      - Для биномиального: $M(X) = np$, $D(X) = npq$.
        Метод моментов: $\hat{p} = \frac{\bar{x}}{k}$, где k — число испытаний.
5. **Вычислить теоретические вероятности pi** по формуле выбранного закона.
6. **Пересчитать теоретические частоты:** $n_i^{\text{теор}} = n \cdot p_i$.
7. **Сравнить эмпирические и теоретические частоты** с помощью критерия согласия (хи-квадрат Пирсона).

Используем критерий **Пирсона (χ^2)**.

Статистика критерия:

$$K = \chi^2_{\text{набл}} = \sum_{i=1}^{m} \frac{(n_i - n_i^{\text{теор}})^2}{n_i^{\text{теор}}}.$$

При справедливости H_0 величина K имеет распределение χ^2 с числом степеней свободы

$$k = m - 1 - r,$$

где m — число групп (значений), r — число оцененных параметров.

Шаг 7. Определить критическую область

Задаём уровень значимости α (обычно 0,05).

Критическая область — **правосторонняя**:

$$V_{kp} = \{K : K > \chi^2_{1-\alpha;k}\}.$$

8. **Сделать вывод** о принятии/отклонении гипотезы.

Если $K_{\text{набл}} \in V_{kp} \to$ отвергаем H_0.

Если $K_{\text{набл}} \notin V_{kp} \to$ принимаем H_0.

II. <u>Для интервального вариационного ряда</u>

(Дано:

- Непрерывная случайная величина X
- Выборка объёма n**)**

1. **Собрать данные и разбить на интервалы и Построить гистограмму частот.**

- Разбить диапазон значений на m интервалов:

$$[a_0, a_1), [a_1, a_2), \ldots, [a_{m-1}, a_m].$$

- Подсчитать частоты n_i попаданий в интервалы.

Выдвинуть гипотезу о непрерывном распределении (нормальное, показательное и т.д.).

Например:

$$H_0 : X \sim N(a, \sigma^2) \quad (\text{нормальное распределение}).$$

2. **Оценить параметры:**

Для нормального распределения:

- Метод моментов: $\hat{a} = \bar{x}$, $\hat{\sigma}^2 = \hat{D}(X)$ (выборочная дисперсия).
- Метод ММП: $\hat{a} = \bar{x}$, $\hat{\sigma}^2 = \frac{1}{n}\sum(x_i - \bar{x})^2$.

3. **Вычислить теоретические вероятности** попадания в каждый интервал:

$$p_i = \int_{a_{i-1}}^{a_i} f_N(x; \hat{a}, \hat{\sigma}^2)\,dx,$$

где f_N — плотность нормального распределения.

Найти теоретические частоты: $n_i^{\text{теор}} = n \cdot p_i.$

4. **Применить критерий согласия** (хи-квадрат, Колмогорова–Смирнова).

$$K = \sum_{i=1}^{m} \frac{(n_i - n_i^{\text{теор}})^2}{n_i^{\text{теор}}}.$$

Число степеней свободы: $k = m - 1 - r$, где $r = 2$ (оценены a и σ^2).

5. **Сделать вывод.**

Если $K_{\text{набл}} \in V_{kp}$ → отвергаем H_0.
Если $K_{\text{набл}} \notin V_{kp}$ → принимаем H_0.

ДОП ИНФА

Обозначения и термины из ваших файлов

- $\hat{\theta}$ — точечная оценка параметра θ.
- $\bar{x}$ — выборочная средняя.
- $\hat{D}(X)$ — выборочная дисперсия.
- s^2 — исправленная дисперсия:

$$s^2 = \frac{n}{n-1}\hat{D}(X).$$

- H_0 — нулевая гипотеза, H_1 — альтернативная.
- α — уровень значимости (ошибка первого рода).
- V_{kp} — критическая область.
- K — статистика критерия.
- $\chi^2_{1-\alpha;k}$ — квантиль распределения хи-квадрат.