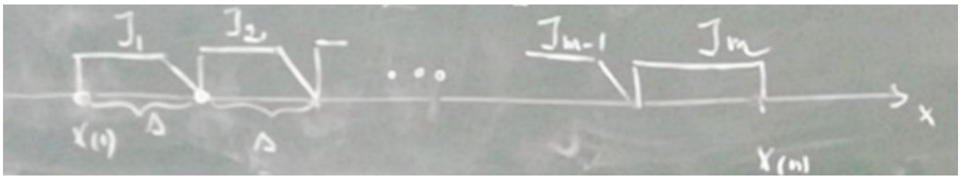
Содержание

1 Лекция 5

1.1 Интервальный статистический ряд

Выше было понятие статистического ряда. Однако, если объем достаточно велик (n > 50), то элементы выборки группируют в так называемый интервальный статистический ряд. Для этого отрезок $J = [x_{(1)}, x_{(n)}]$ разбивают на m равновеликих промежутков. Ширина каждого из них $\Delta = \frac{|J|}{m} = \frac{x_{(1)} - x_{(n)}}{n}$. Данные промежутки строятся по следующему правилу:
$J_i = [x_{(1)} + (i-1)\Delta; x_{(i)} + i\Delta), i = \overline{1, m-1}$
$J_m = [x_{(1)} + (m-1)\Delta; x_{(n)}]$



Определение интервального статистического ряда, отвечающего выборке x называется таблица следующего вида:

J_1	J_2	...	J_m
n_1	n_2	...	n_m

n_i - число элементов выблоки $\overrightarrow{x}$, попавших в промежуток $J_i, i = \overline{1, m}$
Замечание:
1) $\sum\limits_{i=1}^{m} n_i = n$
2) Для выбора m используют формулу:
$m = [log_2 n] + 2$
или
$m = [log_2 n] + 1$

1.2 Эмпирическая плотность

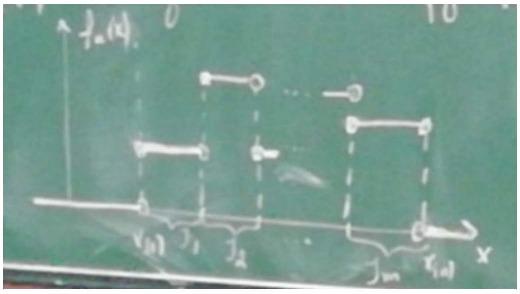
Пусть для данной выборки $\overrightarrow{x}$ построен интервальный статистический ряд $(J_i, n_i), i = \overline{1, m}$
Определение:
Эмпирической плотностью распределения соответствующей выборки $\overrightarrow{x}$ называется функция:

$$f_n(x) = \begin{cases} \frac{n_i}{n \cdot \Delta}, x \in J_i \\ 0 \end{cases} \tag{1.1}$$

Замечание: 1) Очевидно, что $\int\limits_{-\infty}^{+\infty} f_n(x)dx = \int\limits_{x_{(1)}}^{x_{(m)}} f_n(x)dx = \sum\limits_{i=1}^{m} \frac{n_i}{n \cdot \Delta}\Delta = 1$

Таким образом эмпирическая плотность распределения удовлетворяет условию нормировки. Легко показать, что она обладает всеми свойствами функции плотности распределения.

2) $f_n(x)$ является кусочно-постоянной функцией:



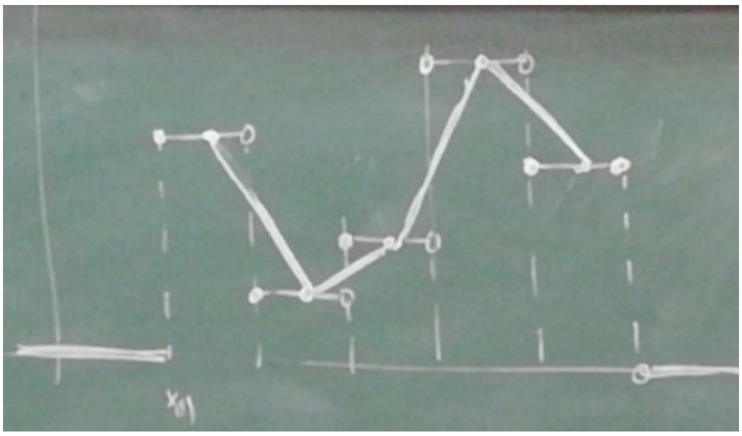
3) Функция $f_n(x)$ вяляется статистическим аналогом функции плотности распределения вероятности. Доказательство - аналогично доказанному выше результату для функции распределения. $\hat{F}_n(x)\overrightarrow{x \to \infty}F(x)$ на P

$f_n(x)$ примерно равна $f(x)$ при n » 1.

Определениие - график эмпирической функции плотности называется гистограммой.

1.3 Полигон частот

Определение полигона частот - пусть для некоторой выборки $\overrightarrow{x}$ построены гистограммы, по определению полигоном частот называется ломаная, звенья которой соединяют середины верхних сторон соседних прямоугольников гистограммы.



1.4 Некоторые распределения, используемые в математической статистике

1.4.1 Гамма-функция Эйлера

По определению гамма-функцией Эйлера называется выражение $\Gamma : R^{+} \to R$, определённое правилом:

$\Gamma(x) = \int\limits_{0}^{+\infty} e^{-t}t^{x-1}dt$

Замечание:

1) Интерграл является несобственным первого рода при $x \geqslant 1$;

при $x \in (0;1)$ этот интеграл является несобственным и имеет следующие особенности: в t = 0 - подинтегральная функция имеет разрыв второго рода, верхний предел равен бесконечности. Легко проверить, что данный интеграл сходится при $x > 0$, при остальных вещественных x он расходится.

Некоторые свойства гамма-функции:

1. $\Gamma(x)$ - является бесконечное число раз дифференцируемой функцией, при этом её к-ая производная задаётся следующей формулой:

$\Gamma^k(x) = \int\limits_{0}^{+\infty} e^{-t}t^{x-1}(lnt)^k dt$

2. $\Gamma(x+1) = x\Gamma(x), x > 0$

3. $\Gamma(1) = 1$

4. $\Gamma(n+1) = n!, n \in N$, по этой причине часто говорят, что гамма-функция является обобщением понятия факториала на вещественные числа.

5. $\Gamma(\frac{1}{2}) = \sqrt{\pi}$, вывод через интеграл Пуассона. 6. $\Gamma(\frac{n+1}{2}) = \left|\text{по второму свойству}\right| = \frac{n-1}{2}\Gamma(\frac{n-1}{2}) = ... = \frac{n-1}{2}\frac{n-2}{2}...\frac{1}{2}\Gamma(\frac{n-1}{2}) = \frac{1\cdot3\cdot5...\cdot(n-1)}{2^n}\sqrt{\pi}$

7. Эскиз графика $\Gamma(x)$

1.4.2 Гамма-распределение

Определение: говорят, что случайная величина ξ имеет гамма-распределение, ели её функция плотности распределения вероятности имеет вид:

$$f_\xi(x) = \begin{cases} \frac{\lambda^\alpha}{\Gamma(\alpha)}x^{\alpha-1}e^{-\lambda x}, x > 0 \end{cases} \tag{1.2}$$

Обозначаеся как ξ $\Gamma(\lambda, \alpha)$

Замечание:

1) Экспоненциальное распределение:

$$f_\xi(x) = \begin{cases} \lambda e^{-\lambda x}, x > 0 \\ 0 \end{cases} \tag{1.3}$$

$Exp(\lambda) = \Gamma(\lambda, 1)$

Теорема:
Пусть случайная величина ξ_1 $\Gamma(\lambda, \alpha_1)$, а ξ_1 $\Gamma(\lambda, \alpha_1)$, ξ_1 и ξ_2 - независимы. Тогда:
$\xi_1 + \xi_2$ $\Gamma(\lambda, \alpha_1 + \alpha_2)$

Следствие:
Если случайные величины $\xi_1, \xi_2, ..., \xi_n$ независимы, причём ξ_i $\Gamma(\lambda, \alpha_i), i = \overline{1, n}$, то:
$\xi_1 + ... + \xi_n$ $\Gamma(\lambda, \alpha_1 + ... + \alpha_n)$

1.4.3 Распределение Релея

Пусть ξ $\mathcal{N}(0, \sigma^2)$
Говорят, что случайная величина ξ имеет распределения Релея с параметром σ.
Замечание:
1) Несложно показать, что:

$$f_y(x) = \begin{cases} \frac{1}{\sigma\sqrt{2\pi x}} e^{\frac{-x}{2b^2}}, x > 0 \\ 0 \end{cases} \tag{1.4}$$

2) Распределение Релея является частным случаем гамма-распределения для $\lambda = \frac{1}{2\sigma^2}$ и $\lambda = \frac{1}{2}$, то есть ν $\Gamma(\frac{1}{2\sigma^2}, \frac{1}{2})$

1.4.4 Распределение хи-квадрат

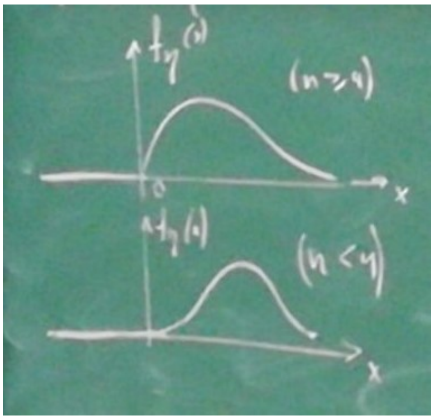
Пусть:
Если случайные величины $\xi_1, \xi_2, ..., \xi_n$ независимы, ξ_i $N(0, 1), i = \overline{1, n}$, $\nu = \xi_1^2 + ... + \xi_n^2$

Определение: в этом случае говорят, что случайная величина ν имеет распределение хи-квадрат с n степенями свободы. Обозначается как ν $X^2(n)$

Замечание:
1) ξ_i $N(0, 1) \Rightarrow \xi_i^2$ имеет распределение Релея с параметром $\sigma = 1$, то есть ξ_i^2 $\Gamma(\frac{1}{2}, \frac{1}{2})$. Так как случайные величины $\xi_1 ... \xi_n$ - независимы с учётом свойства гамма-распределения:
$\nu = \xi_1^2 + ... + \xi_n^2$ $\Gamma(\frac{1}{2}, \frac{n}{2})$, то $X^2 = \Gamma(\frac{1}{2}, \frac{n}{2})$

2) Очевидно, что если независимые случайные величины $\nu_1, ...\nu_m$ имеют распределения $X^2(\nu_i\ X^2(k_i), i$ $\overline{1, m})$, то $\nu_1 + ... + \nu_n\ X^2(k_1 + ...k_m)$

3)График функции плотности $\nu\ X^2(n)$



1.4.5 Распределение Фишера

Пусть:

1) ξ_1, ξ_2 - независимы 2) $\xi_i\ X^2(n_i), i = \overline{1,2}$

3) $\nu = \frac{n_1 \xi_1}{n_2 \xi_2}$

Определение: в этом случае говорят, что случайная величина ν имеет распределение Фишера со степенями свободы $n_1 n_2$, $\nu\ F(n_1, n_2)$

Замечания:

1) Можно показать, что:

$$f_\nu(x) = \begin{cases} C\frac{x^{\frac{n_1}{2}-1}}{(1+\frac{n_1 x}{n_2})^{\frac{n_1+n_2}{2}}}, x > 0 \\ 0 \end{cases} \tag{1.5}$$

$C = \frac{(\frac{n_1}{n_2})^{\frac{n_1}{2}}}{B(\frac{n_1}{2}, \frac{n_2}{2})}$

$B(x, y) = \int\limits_0^1 t^{x-1}(1-t)^{y-1}dt$ - бета-функция Эйлера.

2) Если $\nu\ F(n_1, n_2)$, то $\frac{1}{\nu}\ F(n_2, n_1)$