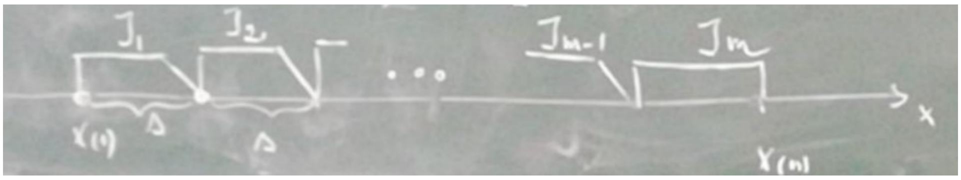
## Содержание

# 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

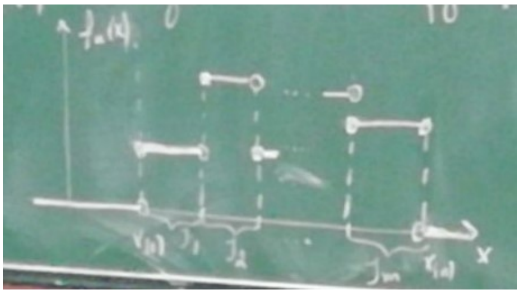
Пусть для данной выборки $\overrightarrow{x}$ построен интервальный статистический ряд $(J_i, n_i), i = \overline{1, m}$
Определение:
Эмпирической плотностью распределения соответствующей выборки $\overrightarrow{x}$ называется функция:

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

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:



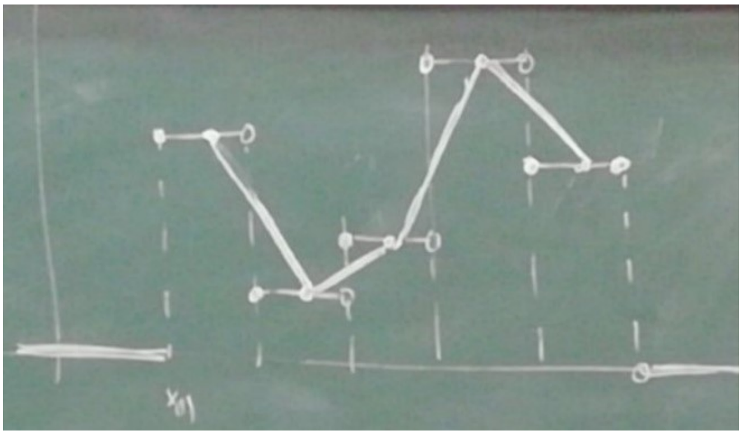
3) Функция $f_n(x)$ вяляется статистическим аналогом функции плотности распределения вероятности. Доказательство - аналогично доказанному выше результату для функции распределения. $\hat{F}_n(x)\overrightarrow{x \to \infty}F(x)$ на P

$f_n(x)$ 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. $\Gamma(\frac{n+1}{2}) = \left|\text{по второму свойству}\right| = \frac{n-1}{2}\Gamma(\frac{n-1}{2}) = ... = \frac{n-1}{2}\frac{n-2}{2}...\frac{1}{2}\Gamma(\frac{n-1}{2}) = \frac{1 \cdot 3 \cdot 5 \cdot ... \cdot (n-1)}{2^n}\sqrt{\pi}$

7. Эскиз графика $\Gamma(x)$

### 1.4.2 Гамма-распределение

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

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

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

Замечание:

1) Экспоненциальное распределение:

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

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

Теорема:

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

Следствие:

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

### 1.4.3 Распределение Релея

Пусть $\xi$ $\mathcal{N}(0, \sigma^2)$

Говорят, что случайная величина $\xi$ имеет распределения Релея с параметром $\sigma$.

Замечание:

1) Несложно показать, что:

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

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

### 1.4.4 Распределение хи-квадрат

Пусть:

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

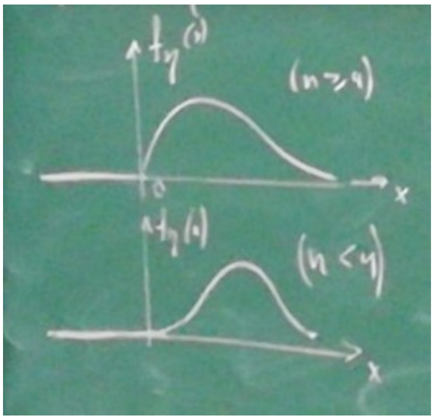
Определение: в этом случае говорят, что случайная величина $\nu$ имеет распределение хи-квадрат с n степенями свободы. Обозначается как $\nu$ $X^2(n)$

Замечание:

1) $\xi_i$ $N(0, 1) \Rightarrow \xi_i^2$ имеет распределение Релея с параметром $\sigma = 1$, то есть $\xi_i^2$ $\Gamma(\frac{1}{2}, \frac{1}{2})$. 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) $\xi_1, \xi_2$ - 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 $\nu$ 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 $\frac{1}{\nu}\ F(n_2, n_1)$

# 2 Лекция 6

## 2.1 Переписать

## 3 Лекция 7

По определению оценка $\hat{\theta}$ называется эффективной оценкой параметра $\theta$, если:

1) $\hat{\theta}$ является наименьшей оценкой для теты

2) оценка $\hat{\theta}$ обладает наименьшей дисперсией среди всех несмещённых $\theta$

Замечание: иногда говорят не об эффективной вообще точечной оценке, а об оценке, эффективной в некотором классе. Пусть $\Theta$ - некоторый класс несмещённых оценок для параметра $\theta$.

По определению оценка $\hat{\theta}$ называется эффективной в классе $\Theta$, если она имеет наименьшую дисперсию среди всех оценок этого класса, т.е. - $(\forall\tilde{\theta})(D\hat{\theta} \leqslant D\tilde{\theta})$.

Пример:

Пусть X - случайная величина, обладающая MX = m и DX = $b^2$. Покажем, что оценка $\hat{m}_1(\overrightarrow{X}) = \overline{X}$ является эффективной оценкой для m и b в классе линейных оценок.

Решение:

1) Линейная оценка имеет вид: $\hat{m}(\overline{X}) = \sum\limits_{i=1}^{n} \lambda_i X_i = \lambda_1 X_1 + ... + \lambda_n X_n (*)$

где $\lambda_i \in R, i = \overline{1,n}$, тогда матожидание линейной оценки (*) : a) $M[\hat{m}] = \lambda_1 M X_1 + ... + \lambda_n M X_n = \left| X_i \ X_j, MX = m \right| = (\lambda_1 + ... + \lambda_n)m$. Так как оценка является несмещённой, то $M[\hat{m}] = m \Rightarrow \sum\limits_{i=1}^{n} = 1$

b) Дисперсия оценки (*):

$D[\hat{m}] = \sum\limits_{i=1}^{n} \lambda_i^2 D X_i = \lambda^2 \sum_{i=1}^{n} \lambda_i^2$ - аналогично матожиданию.

2) Попробуем подобрать коэффициент $\lambda_i, i = \overline{1,n}$, и (*) так, чтобы:

$$\begin{cases} D[\hat{m}] \to min \\ \sum\limits_{i=1}^{n} \lambda_i = 1 \end{cases} \tag{3.1}$$

Для этого нужно решить задачу условной оптимизации:

$$\begin{cases} \phi(\lambda_1 ... \lambda_n) = \lambda_1^2 + ... + \lambda_n^2 \to min \\ \sum \lambda_i = 1 \end{cases} \tag{3.2}$$

Запишем функцию Лагранжа:

$L(\lambda_1 ... \lambda_n, \mu) = \lambda_1^2 + ... + \lambda_n^2 - \mu \sum \lambda_i - \mu$

$$\begin{cases} \frac{dL}{d\lambda_i} = 0 \\ \frac{dL}{d\mu} = 0 \end{cases} \tag{3.3}$$

Следовательно:

$$\begin{cases} \frac{dL}{d\lambda_i} = 2\lambda_i - \mu = 0 \\ \frac{dL}{d\mu} = -(\sum\limits_{i=1}^{n} \lambda_i - 1) = 0 \end{cases} \tag{3.4}$$

Из n первых уравнений - $\lambda_i = \frac{\mu}{2}, i = \overline{1,n}$

Покажем, что найденное решение соответствует точке условного минимума целевой функции, таким образом, подставляя $\lambda_i$ в * получаем искомую оценку с минимальной дисперсией в классе линейных оценок.
$\hat{m}(\overrightarrow{X}) = \frac{1}{n}X_1 + ... + \frac{1}{n}X_n = \overline{X}$
Дисперсия этой оценке:
$D[\hat{m}] = \sigma^2 \sum_{i=1}^{n} \lambda_i^2 = \frac{\sigma^2}{n}$

Теорема:
Теорема о единственности эффективной оценки:
Пусть $\tilde{\theta_1}(\overline{X})$ и $\tilde{\theta_2}(\overline{X})$ - эффективные оценки некой оценки параметра $\theta$, тогда $\tilde{\theta_1}(\overline{X}) = \tilde{\theta_2}(\overline{X})$

### 3.1 Неравенство Рао-Крамера

Пусть X - случайная величина, закон распределения которой зависит от вектора $\overrightarrow{\theta} = (\theta_1, ..., \theta_n)$ параметров.
Пусть $\overrightarrow{X} = (X_1, ..., X_n)$ - случайная выборка из генеральной совокупности X.

Опеределение - функцией правдоподобия, отвечающей случайной выборке $\overrightarrow{X}$ называется функция $L(\overrightarrow{X}, \overrightarrow{\theta}) = p(X_1, \overrightarrow{\theta})...p(X_1, \overrightarrow{\theta})$
где:

$$p(X_i, \overrightarrow{\theta}) = \begin{cases} f(X_i, \overrightarrow{\theta}), \text{если X - непрерывная случайная величина} \\ P(X = X_i), \text{если X - дискретная случайная величина} \end{cases} \tag{3.5}$$

Пусть r = 1, т.е. $\overrightarrow{\theta} = (\theta_1) = (\theta)$

Определение - количество информации по Фишеру, содержащееся в случайной величине $\overrightarrow{X}$, называется число $I(\theta) = M[(\frac{dlnL}{d\theta})^2]$

Теорема:
Неравенство Рао-Крамера:
Пусть рассматриваемая модель является регулярной, $\hat{\theta}(\overrightarrow{X})$ - несмещённая оценка параметра тета. Тогда:

$D[\hat{\theta}] \geqslant \frac{1}{I(\theta)}$ - неравенство Рао-Крамера.

Замечание:

1) При доказательстве теоремы Рао используются дифференциальные параметры под знаком интеграла:

$\frac{d}{d\phi}\int\limits_G \phi(\overrightarrow{X},\theta)dx = \int\limits_G \frac{d\phi(\overrightarrow{X},\theta)}{d\theta}dx$

Т.е. параметрические модели, для которых справедливо это равенство, будем называть регулярными.

2) Неравенство Рао даёт нижнюю границу для дисперсии для всех возможных оценок параметра $\theta$.

3) Величина $e(\hat{\theta}) = \frac{1}{I(\theta)D(\hat{\theta})}$ называется показателем эффективности по Рао точечной оценки $\hat{\theta}$

$0 \leqslant e(\hat{\theta}) \leqslant 1$

Очевидно, что оценка эффективная по Рао будет "просто"эффективной. Вопрос в том, для каких параметричесих моделей существует эффективная по Рао оценка (то есть существует оценка, дисперсия которой равна $\frac{1}{I(\theta)}$) мы оставим без рассмотрения.

4) В некоторых случаях вводят в рассмотрение величину $I_0(\theta) = M[(\frac{dp(X,\theta)}{d\theta})^2]$

где $p(X,\theta)$ имеет тот же смысл, что и функция правдоподобия.

Данную величину можно назваь количеством информации по Фишеру в одном испытании. Для некоторых параметрических моделей справедливо:

$I(\theta) = nI_0(\theta)$,

где n - объём случайной информации.

Пример:

Пусть $X\ N(m,\sigma^2)$, где m - неизвестна, $\sigma$ - известна. Докажем, что оценка $\hat{m}_1(\overrightarrow{X}) = \overline{X}$ для m является эффективной по Рао.

1) Необходимо найти показатель эффективности оценки $\hat{m}_1$:

$e(\hat{m}) = \frac{1}{I(m)D(\hat{m})}$, если данная величина равна 1, то оценка эффективна по Рао, иначе не является эффективной по Рао. 2) $D[\hat{m}] = D[\overline{X}] = ... = \frac{\sigma^2}{n}$

3) $I(\hat{m}) = ?$

$I(\hat{m}) = M[(\frac{dlnL}{dn})^2]$, составим функцию L правдоподобия:

$L(\overline{X},m) = p(X,m)\cdot ... \cdot p(X_n,m) = \frac{1}{(\sqrt{2\pi})^n\sigma^n}e^{-\frac{1}{2\sigma^2}\sum(x_i-m)^2}$

Тогда:

$lnL(\overline{X},n) = -\frac{n}{2}ln2\pi - nln\sigma - \frac{1}{2\sigma^2}\sum_{i=1}^n(m-X_i)^2$

$\frac{dlnL(\overline{X},m)}{dm} = -\frac{2}{2\sigma^2}\sum\limits_{i=1}^n(m-X_i) = \frac{1}{\sigma^2}\sum_{i=1}^n(X_i-m)$

$(\frac{dlnL(\overline{X},m)}{dm})^2 = \frac{1}{\sigma^4}[(X_1-m)+...+(X_n-m)]^2$

Т.о.:

$I(m) = M[(\frac{dlnL(\overline{X},m)}{dm})^2] = \frac{1}{\sigma^4}[\sum\limits_{i=1}^n M[(X_i-m)^2 + 2\sum\limits_{1\leqslant i<j\leqslant n} M[(X_i-m)(X_j-m)]] = \frac{1}{\sigma^4}[\sum\limits_{i=1}^n \sigma^2 + 0] = \frac{1}{\sigma^4}n\sigma^2 = \frac{n}{\sigma^2}$

4) Получаем $e(\hat{m})$