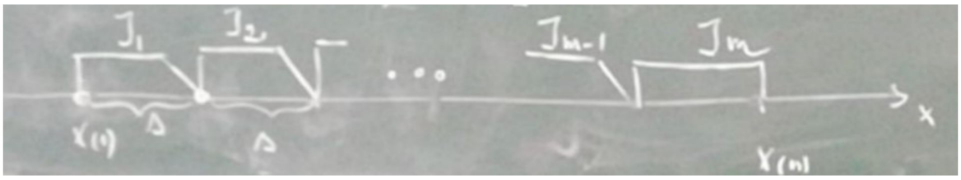
Содержание

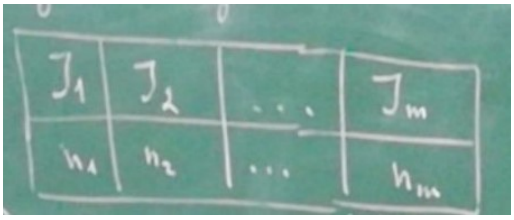
1 Лекция 5

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

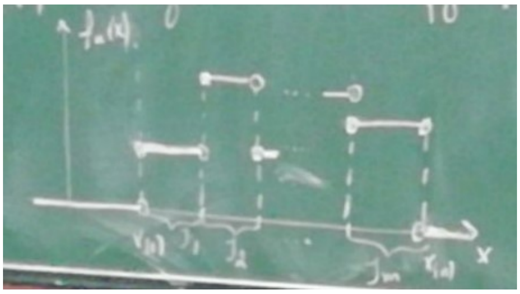
Пусть для данной выборки $\overrightarrow{x}$ построен интервальный статистический ряд $(J_i, n_i), i = \overline{1, m}$
Определение:
Эмпирической плотностью распределения соответствующей выборки $\overrightarrow{x}$ называется функция:

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

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:



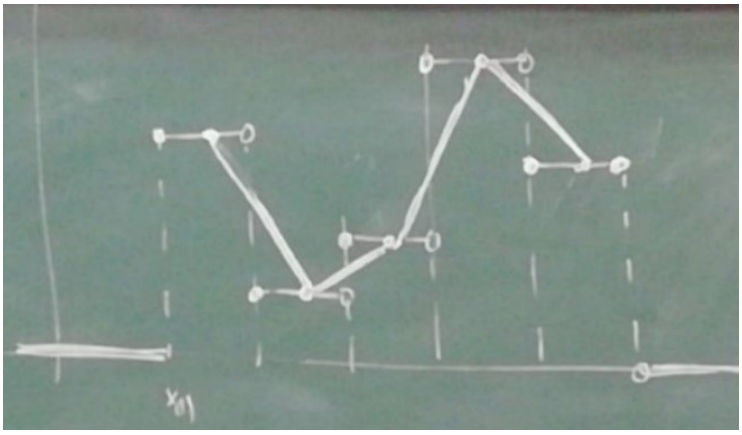
3) Функция $f_n(x)$ вяляется статистическим аналогом функции плотности распределения вероятности. Доказательство - аналогично доказанному выше результату для функции распределения. $\hat{F}_n(x)\overrightarrow{x \to \infty}F(x)$ на P

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

Опредениение - 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 верних 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. $\Gamma(\frac{n+1}{2}) = \left|\text{по второму свойству}\right| = \frac{n-1}{2}\Gamma(\frac{n-1}{2}) = \dots = \frac{n-1}{2}\frac{n-2}{2}\dots\frac{1}{2}\Gamma(\frac{n-1}{2}) = \frac{1\cdot3\cdot5\dots\cdot(n-1)}{2^n}\sqrt{\pi}$

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 $X^2(n)$

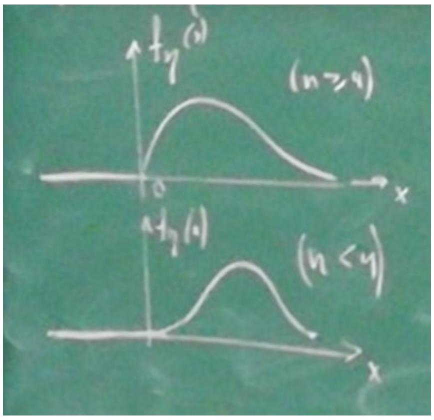
Замечание:

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

2 Лекция 6

2.1 Переписать

3 Лекция 7

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

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

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

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

По определению оценка $\hat{\theta}$ называется эффективной в классе Θ, 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) $M[\hat{m}] = \lambda_1 MX_1 + ... + \lambda_n MX_n = \left| X_i \ X_j, MX = m \right| = (\lambda_1 + ... + \lambda_n)m$. Так как оценка является несмещённой, то $M[\hat{m}] = m \Rightarrow \sum\limits_{i=1}^{n} = 1$

b) Дисперсия оценки (*):

$D[\hat{m}] = \sum\limits_{i=1}^{n} \lambda_i^2 DX_i = \lambda^2 \sum_{i=1}^{n} \lambda_i^2$ - 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 λ_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 θ, 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 θ.

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 назвать 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, σ - 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 $e(\hat{m})$

4 Лекция 8

4.1 Методы построения точечных оценок

1. Метод моментов
2. Метод максимального правдоподобия.

4.2 Метод моментов

Пусть:

1) X - некий случайный вектор, распередление которого зависит от вектора $\overrightarrow{\theta} = (\theta_1, ..., \theta_n)$ неизвестных параметров.

2) $\exists r$ первых моментов случайного вектора X, то есть $\exists M[X^k], k = 1,..r$

Тогда в методе моментов:

1) Вычисляются теоретические моменты 1-го, 2-го, ..., r-го порядков, зависящих от неизвестных параметров:

$m_k(\theta_1, ..., \theta_r = M[X^k], k = \overline{1, r}$ - теоретические моменты порядка k

2) Теоретические моменты приравниваются к выборочным аналогам:

$$\begin{cases} m_1(\theta_1, ..., \theta_r) = \hat{m}_1(\overrightarrow{X}) \\ ... \\ m_r(\theta_1, ..., \theta_r) = \hat{m}_r(\overrightarrow{X}) \end{cases} \tag{4.1}$$

Система уравнений (возможно, нелинейных), относительно неизвестных параметров тета.

3) Решаем получившуюся систему:

$$\begin{cases} \theta_1 = \hat{\theta}_1(\overrightarrow{X}) \\ ... \\ \theta_r = \hat{\theta}_r(\overrightarrow{X}) \end{cases} \tag{4.2}$$

Полученные зависимости используются в качестве точечных оценок для полуинтервалов.

Замечание:

Иногда некоторые уравнения системы из подпункта (2) удобнее записывать относительно центральных, а не начальных моментов. В этом случает k-е уравнение будет иметь вид:

$\mathring{m}_k(\theta_1, ..., \theta_r) = \circ\hat{m}_k(\overrightarrow{X})$, где:

$\mathring{m}_k(\theta_1, ..., \theta_r) = M[(X - MX)^k]$

$\mathring{\hat{m}}_k(\overrightarrow{X}) = \frac{1}{n}\sum\limits_{i=1}^{n}(X_i - \overline{X})^k$

Пример: Пусть $X\ \ R[a,b]$, гже a, b - произвольные параметры. С помощью метода моментов построить точечные оценки для a, b.

1)

$$X\ \ R[a,b] \Rightarrow f(x) = \begin{cases} \frac{1}{b-a}, x \in [a,b] \\ 0, \text{иначе} \end{cases} \tag{4.3}$$

Неизвестные параметры (a и b), следовательно требуется два уравнения:

$$\begin{cases} m_1(a,b) = \hat{m}_1(\overrightarrow{X}) \text{ - относительно начального момента 1-го порядка} \\ m_2(a,b) = \hat{m}_2(\overrightarrow{X}) \text{ - относительно центрального момента второго порядка} \end{cases} \tag{4.4}$$

2) Найдём теоретические моменты:

$m_1(a,b) = MX = \frac{a+b}{2}$

$\circ m_2(a,b) = DX = \frac{(b-a)^2}{12}$

Запишем выборочные моменты:

$\hat{m}_1(a,b) = \frac{1}{n}\sum_{i=1} nX_i = \overline{X}$

$\mathring{m}_2(a,b) = \frac{1}{n}\sum_{i=1} n(X-\overline{X})^2 = \hat{\sigma}^2(\overrightarrow{X})$

или

$\mathring{m}_2(a,b) = \frac{1}{n-1}\sum_{i=1} n(X-\overline{X})^2 = S^2(\overrightarrow{X})$

Используем S^2

3) Приходим к следующей системе уравнений:

$$\begin{cases} \overline{X} = \frac{a+b}{2} \\ S^2 = \frac{(b-a)^2}{12} \end{cases} \tag{4.5}$$

a, b = ?

Далее система решается стандартно

Замечание:

1) Поскольку выбранные моменты $\hat{m}_k(\overrightarrow{X})$ и $\mathring{\hat{m}}_k(\overrightarrow{X})$ являются состоятельными оценками соотвествтующих теоретических моментов, то можно показать, что в случае непрерывной зависимости решения системы из пункта (2) от $\hat{m}_k$ (или от $\mathring{m}_k$) оценки параметров, полученные с использованием этого метода также являются состоятельными.

2) Так как выборочные моменты степени k при $k \geqslant 2$ являются смещёнными оценками сво-

их теоретических аналогов, то и оценки параметров, полученные с помощью метода моментов также могут быть смещёнными.

4.3 Метод максимального правдоподобия

Пусть:
1) X - случайная величина, зависящая от вектора параметров $\overrightarrow{\theta} = (\theta_1, ..., \theta_r)$

Ранее было введено понятие функции правдоподобия случайной выборки X:
$L(\overrightarrow{X}, \overrightarrow{\theta}) = p(X_1, \overrightarrow{\theta}) \cdot ... \cdot p(X_n, \overrightarrow{\theta})$
, где:

$$p(X_i, \overrightarrow{\theta}) = \begin{cases} f(X_i, \overrightarrow{\theta}), \text{если X - непрерывная случайная величина} \\ P(X = X_i), \text{если X - дискретная случайная величина} \end{cases} \tag{4.6}$$

Можно показать, что чем ближе значение вектора $\overrightarrow{\tilde{\theta}}$ к теоретическому значению вектора тета, тем большие значения принимает функция правдоподобия $L(\overrightarrow{x}, \overrightarrow{\theta})$

В методе максимального правдоподобия в качестве точечных оценок неизвестного параметра выступают значения, доставляющие максимальное значение функции правдоподобия. Для реализации метода необходимо решить задачу
$L(\overrightarrow{X}, \overrightarrow{\theta}) \to max_{\overrightarrow{\theta}}$ (*)
тогда:
$\overrightarrow{\hat{\theta}}(\overrightarrow{X}) = argmax_{\overrightarrow{\theta}} L(\overrightarrow{X}, \overrightarrow{\theta})$

Замечание:
1) Для решения задачи (*) можно использовать неоходимые условия дифференциальной функции нескольких переменных:

$$\begin{cases} \frac{\delta L(\overrightarrow{X}, \overrightarrow{\theta})}{\delta \theta_1} = 0 \\ \frac{\delta L(\overrightarrow{X}, \overrightarrow{\theta})}{\delta \theta_2} = 0 \end{cases} \tag{4.7}$$

Данные уравнения называются уравнениями правдоподобия
2) Функция L является произведением n сомножетелей и работать с ней не всегда удобно. Поэтому часто вместо задачи (*) решют эквивалентную задачу (**):
$lnL(\overrightarrow{X}, \overrightarrow{\theta}) \to max_{\overrightarrow{\theta}}$
т.е.:
$\overrightarrow{\hat{\theta}}(\overrightarrow{X}) = argmax_{\overrightarrow{\theta}} lnL(\overrightarrow{X}, \overrightarrow{\theta})$
Данная замена эквивалентна, так как ln - монотонная возрастающая функция.

Пример:

$X\ R[a,b]$, a,b - неизвестные параметры, необходимо с помощью метода максимального правдоподобия построить точеные оценки параметров a и b.

Решение:

1)

$$X\tilde{R}[a,b] \Rightarrow f(x) = \begin{cases} \frac{1}{b-a}, x \in [a,b] \\ 0, \text{иначе} \end{cases} \tag{4.8}$$

Выпишем функцию правдоподобия $L(\overrightarrow{X}, a, b) =$ так как X - непрерывная случайная величина $=$ $f(X_1, a, b) \cdot ... \cdot f(X_n, a, b) = \frac{1}{b-a} \cdot \frac{1}{b-a} \cdot ... \cdot \frac{1}{b-a} = \frac{1}{(b-a)^n}$

2) $lnL = ln\frac{1}{(b-a)^n} = -nln(b-a) \rightarrow max_{a,b}$

Уравнения правдоподобия:

$$\begin{cases} \frac{\delta lnL}{\delta a} = \frac{n}{b-a} = 0 \\ \frac{\delta lnL}{\delta b} = -\frac{n}{b-a} = 0 \end{cases} \tag{4.9}$$

n $= 0$ - противоречие, возникло из-за того, что неверно записана формула для L.

Правильная запись формулы:

$$L(X, a, b) = \begin{cases} \frac{1}{(b-a)^n}, if X(1) \geqslant a, X(n) \leqslant b \\ 0, \text{иначе} \end{cases} \tag{4.10}$$

Вычисление $\frac{\delta L(\overrightarrow{X},a,b)}{\delta a}$, $\frac{\delta L(\overrightarrow{X},a,b)}{\delta b}$ затруднительно, так как a и b также зависят от области, в которой L > 0

Попробуем в "в лоб" найти $max_{a,b}L(\overrightarrow{X},a,b)$

a) если

$$\begin{cases} a \leqslant X_{(1)} \\ b \geqslant X_{(n)} \end{cases} \tag{4.11}$$

то L > 0, в противном случае L $= 0$. Так как $L \rightarrow max$, то условия на a, b в любом случае должны быть верны

b) при вычислении условий на a, b

$L = \frac{1}{(b-a)^n}$

Так как $L \rightarrow max$, то $b - a \rightarrow min$

Тогда приближаем к подзадаче:

$$\begin{cases} b-a \to min \\ a \leqslant X_{(1)} \\ b \geqslant X_{(n)} \end{cases} \tag{4.12}$$

Таким образом:

$$\begin{cases} \hat{a}(\overrightarrow{X}) = X_{(1)} \\ \hat{b}(\overrightarrow{X}) = X_{(n)} \end{cases} \tag{4.13}$$

4.4 Интервальные оценки

Основные понятия:

1) Рассматривается вторая задача математической статистики.

Ранее для решения этой задачи использовались точечные оценки. Тогда принимались равенства $\theta_j = \hat{\theta}_j(\overrightarrow{x}), j = \overline{1,k}$

Для некоторых статистик $\hat{\theta}_1, ..., \hat{\theta}_r$. Недостатком данного подходя является то, что он выдаёт информацию о вероятностных характеритиках точности оценивания неизвестных параметров. Кроме точечных оценок для решения второй задачи математической статистики используется другой подход. Для простоты будем считать, что у нас только один неизвестный параметр $r = 1, \overrightarrow{\theta} = (\theta_1) = (\theta) = \theta$.

Опеределение:

Интервальной оценкой параметра θ уровня γ (γ-интервальной оценкой), $\gamma \in (0;1)$ называют пару интервальных статистик $\underline{\theta}(\overrightarrow{X}), \overline{\theta}(\overrightarrow{X})$ таких, что:

$P\{\theta \in (\underline{\theta}(\overrightarrow{X}), \overline{\theta}(\overrightarrow{X})\} = \gamma$

5 Лекция 9

Замечание:

1) Интервальная оценка является интервалом со случайными границами $\underline{\theta}(\overrightarrow{X}), \overline{\theta}(\overrightarrow{X})$, который накрывает инизвестно теоретическое значение параметра с вероятностью γ. 2) Вероятность совершить ошибку при построении интервальной оценки уровня γ:
$1-\gamma = P\{\theta \notin (\underline{\theta}(\overrightarrow{X}), \overline{\theta}(\overrightarrow{X}))\}$
3) Вероятностной оценкой интервальной оценкой уровня гамма является случайная величина $l(\overrightarrow{X}) = \overline{\theta}(\overrightarrow{X}) - \underline{\theta}(\overrightarrow{X})$
4) Иногда удобно строить односторонние интервальные оценки.
Определение: нижней односторонней интервальной оценкой гамма-доверительной границей для параметра θ называют статистику $\underline{\theta}$ такую, что $P\{\theta > \underline{\theta}\} = \gamma$, для верхней оценкой аналогично.

Определение: гамма-доверительным интервалом (доверительным интервалом уровня гамма) для параметра тэта называют реализацию (выборочное значение) интервальной оценки уровня гамма для этого параметра, т.е. интервал $(\underline{\theta}(\overrightarrow{x}), \overline{\theta}(\overrightarrow{x}))$ с детерминированными границами.

Замечание: иногда там, гем это не будет приводить к путанцие, мы будем допускать вольность речи, не разделяя строго понятия интервала и интервальной оценки доверительного интервала.

5.1 Построение интервальных оценок

Пусть:
1) X - случайная величина, закон распределения которой известен с точностью до значения параметра $\theta \in \mathcal{R}$
Требуется построить интервальную оценку уровня гамма для параметра θ.

Определение: по определению статистика $g(\overrightarrow{X}, \theta)$ называется центральной, если закон её распределения не зависит от неизвестного параметра θ.

Пример:
$X \ N(\theta, \sigma^2)$, где θ - неизвестно, σ^2 - известно.

Рассмотрим статистику $g(\overrightarrow{X}, \theta) = \frac{\theta - \overline{X}}{\sigma}\sqrt{n}$.

Покажем, что g - центральная статистика.
а) $X_i \ N(\theta, \sigma^2), i = \underline{1, n} \Rightarrow \overline{X} \ N(?, ?)$
Следовательно: $g(\overrightarrow{X}, \theta) = \frac{\theta\sqrt{n}}{\sigma} - \frac{\sqrt{n}}{\sigma}\overline{X} \ N(m_g, \sigma_g^2)$
б) найдём:
$m_y = M[g(\overrightarrow{X}, \theta)] = M[\frac{\theta\sqrt{n}}{\sigma} - \frac{\sqrt{n}\overline{X}}{\sigma}] = 0$

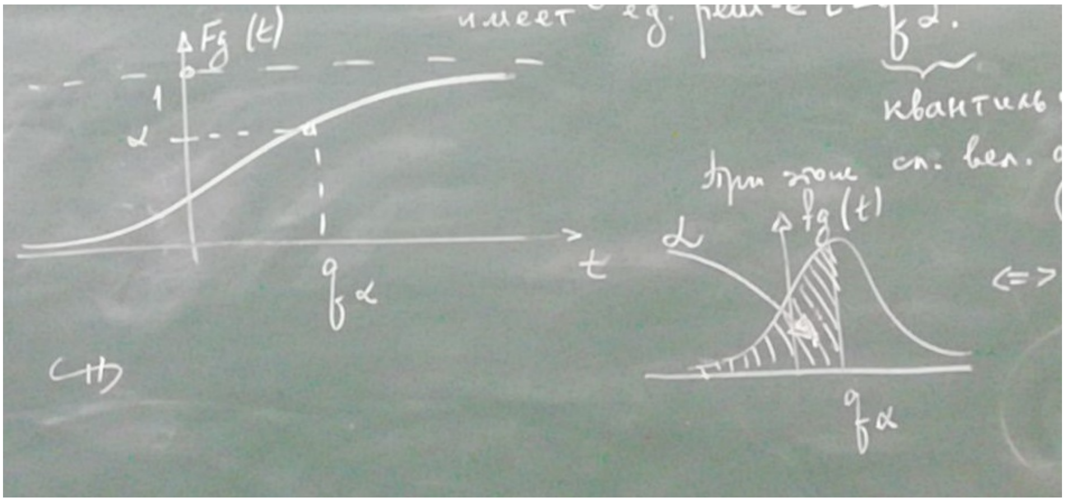
$\sigma_g^2 = D[g(\overrightarrow{X},\theta)] = D[\frac{\theta\sqrt{n}}{\sigma} - \frac{\sqrt{n}\overline{X}}{\sigma}] = \frac{n}{\sigma^2}D\overline{X} = 1$

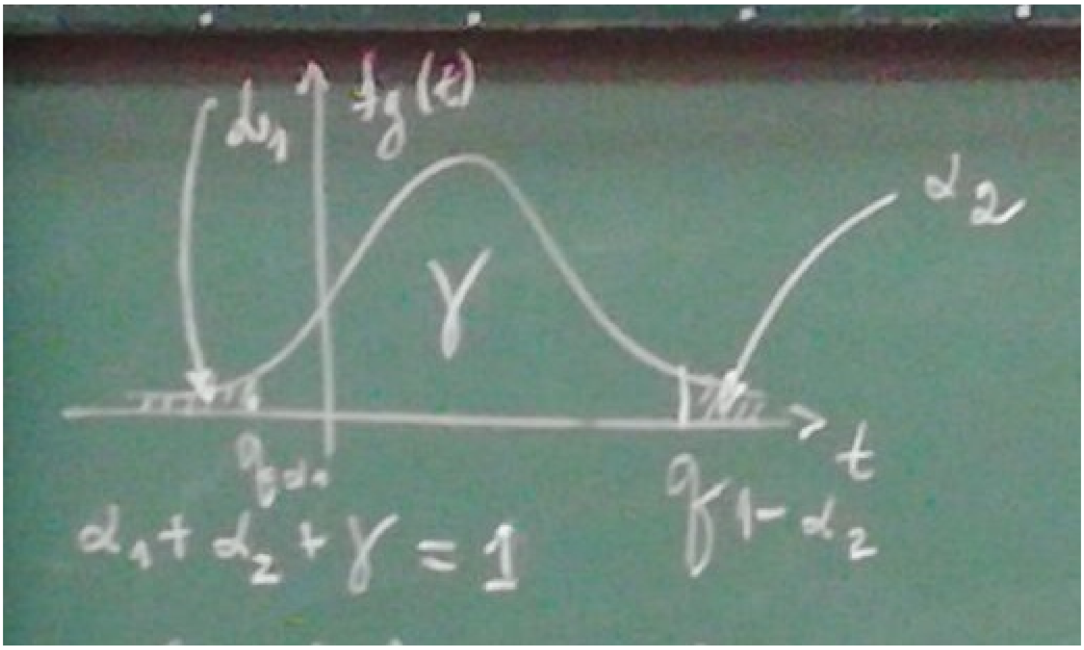
т.е. $g(\overrightarrow{X},\theta)$ $N(0,1)$ не зависит от θ
Следовательно g - центральная статистика.

5.2 Общий алгоритм построения интервальной оценки с использованием центральной статистики

Пусть:
1) $g(\overrightarrow{X},\theta)$ - центральная статистика
2) $g(\overrightarrow{X},\theta)$ - как функция параметра тета также является монотонно возрастающей
3) $F_g(y)$ - является монотонно возрастающей
4) Выбраны некоторые значения $\alpha_1, \alpha_2 > 0$ такие что $\alpha_1 + \alpha_2 = 1 - \gamma$, где гамма - заданный уровень доверия.

Замечание:
из условния (3) сделаем вывод, что уравнени $F_g(t) = \alpha$ для любого $\alpha \in (0,1)$ имеет единственное решение $t = q_\alpha$ - квантиль уровня α случайной величины g $P\{X < q_\alpha\} = \alpha$.





Из свойств непрерывных случайных величин следует: $P\{q_{\alpha_1} < g(\overrightarrow{X}, \theta) < q_{1-\alpha_2}\} = \gamma$

Запишем событие, стоящее под вероятностью в эквивалентном виде:
$\{q_{\alpha_1} < g(\overrightarrow{X}, \theta) < q_{1-\alpha_2}\} \Leftrightarrow \{g^{-1}(\overrightarrow{X}, q_{\alpha_1}) < \theta < g^{-1}(\overrightarrow{X}, q_{1-\alpha_2})\}$
Выражаем из двойного неравенства θ , с учётом условия (2) знаки неравенств не изменияются.

Обозначим:
$\underline{\theta}(\overrightarrow{X}) = g^{-1}(\overrightarrow{X}, g_{\alpha_1})$
$\overline{\theta}(\overrightarrow{X}) = g^{-1}(\overrightarrow{X}, g_{1-\alpha_2})$
Так как события эквивалентны, то:
$P\{\underline{\theta}(\overrightarrow{X}) < \theta < \overline{\theta}(\overrightarrow{X})\} = \gamma$

Замечения:
1) Построенная оценка зависят от выбранных α_1, α_2, обычно используют $\alpha_1 = \alpha_2 = \frac{1-\gamma}{2}$. Однако это не всегда так.
2) Если выбрать:
$\alpha_1 = 1 - \gamma, \alpha_2 = 0$, то $q_{1-\alpha_2} = +\infty \Rightarrow \underline{\theta}(\overrightarrow{X}) = g^{-1}(\overrightarrow{X}, q_{1-\gamma}), \overline{\theta}(\overrightarrow{X}) = +\infty$
Т.е. $\underline{\theta}\overrightarrow{X}$ - нижнаяя гамма-доверительная граница для тэты.
3) Аналогично при $\alpha_1 = 0, \alpha_2 = 1 - \gamma$
$\underline{\theta}(\overrightarrow{X}) = -\infty$
$\overline{\theta}(\overrightarrow{X}) = g^{-1}(\overrightarrow{X}, q_\gamma)$

Пример:
Пусть $X\ N(m, \sigma^2)$
где m - неизвестно, σ - неизвестно.

Построить гамма-доверительную интервальную оценку для m

Решение:

1) Рассмотрим статистику $g(\overrightarrow{X}, m) = \frac{m-\overline{X}}{\sigma}\sqrt{n}\ N(0,1)$

$\alpha_1 = \alpha_2 = \frac{1-\gamma}{2}$

Тогда:

$P\{-u_{\frac{1+\gamma}{2}} < \frac{m-\overline{X}}{\sigma}\sqrt{n} < u_{\frac{1+\gamma}{2}}\} = \gamma$

$P\{\overline{X} - \frac{\sigma u_{\frac{1+\gamma}{2}}}{\sqrt{n}} < m < \overline{X} + \frac{\sigma u_{\frac{1+\gamma}{2}}}{\sqrt{n}}\} = \gamma$

Замечание:

1) Мы использовали $\alpha_1 = \alpha_2 = \frac{1-\gamma}{2}$. Однако решая эту систему для произвольных альфа больших нуля мы получили бы результаты:

$\underline{m}\overrightarrow{X} = \overline{X} - \frac{\sigma u_{\alpha_1}}{\sqrt{n}}$

$\overline{m}\overrightarrow{X} = \overline{X} + \frac{\sigma u_{1-\alpha_2}}{\sqrt{n}}$

В этом случае различные доверительные интервалы:

$l(\overrightarrow{X}) = \overline{m}(\overrightarrow{X}) - \underline{m}(\overrightarrow{X}) = \frac{\sigma}{\sqrt{n}}(u_{1-\alpha_2} - u_{\alpha_1})$

Очевидно, что $l(\overrightarrow{X}) \to min, \alpha_1 = \alpha_2$, в любом случае в рассматривемом примере размах является не случаной, а дискретной величиной.

2) Зависимость размаха доверительного интервала от параметров:

Чем больше n, тем меньше размах.

При $\gamma \to 1, l \to +\infty$

Пример:

Пусть $X\tilde{N}(m, \sigma^2)$

где m - неизвестно, σ - неизвестно.

Построить гамма-доверительную интервальную оценку для m

Решение:

1) Рассмотрим статистику $g(\overrightarrow{X}, m) = \frac{m-\overline{X}}{S(\overrightarrow{X})}\sqrt{n}$?

2) Представим в виде:

$g(\overrightarrow{X}, m) = \frac{\frac{m-\overline{X}}{\sigma}\sqrt{n}}{fracS(\overrightarrow{X})\sigma} = \frac{\frac{m-\overline{X}}{\sigma}\sqrt{n}}{\sqrt{\frac{(n-1)S^2(\overrightarrow{X})}{\sigma^2}}}\sqrt{n-1}$

Пусть $\xi = \frac{m-\overline{X}}{\sigma}\sqrt{n}, \nu = \frac{(n-1)S^2(\overrightarrow{X})}{\sigma^2}$

Тогда $g(\overrightarrow{X}, m) = \frac{\xi}{\nu}\sqrt{n-1}$

$\xi\ N(0,1), \nu\ X^2(n-1)$, ξ, ν - независимы. 3) Таким образом g зависит от распределения Стьюдента от (n - 1)

4) Выражаем интервал через m и S аналогично предыдущему примеру и таким образом получаем интервальные оценки.

6 Лекция 10

Пример:

Пусть $X\tilde{N}(m,\sigma^2)$, где σ^2 - неизвестная, m - неизвестная. Построить интервальные оценки.

Рассмотрим статистику:

$g(\overrightarrow{X},\sigma^2) = \frac{(n-1)\sigma^2(\overrightarrow{X})}{\sigma^2}$

Можно показать, что $g\tilde{X}^2(n-1)$, т.о. принимаем $\alpha_1 = \alpha_2 = \frac{1-\gamma}{2}$, получаем:

$P\{h_{\frac{1-\gamma}{2}} < g(\overrightarrow{X},\sigma^2) < L_{\frac{1+\gamma}{2}}\} = \gamma$

$P\{h_{\frac{1-\gamma}{2}} < \frac{(n-1)\sigma^2(\overrightarrow{X})}{\sigma^2} < L_{\frac{1+\gamma}{2}}\} = \gamma$

$P\{\frac{1}{h_{\frac{1-\gamma}{2}}} > \frac{\sigma^2}{(n-1)\sigma^2(\overrightarrow{X})} > \frac{1}{h_{\frac{1+\gamma}{2}}}\} = \gamma$

$P\{\frac{(n-1)\sigma^2(\overrightarrow{X})}{h_{\frac{1+\gamma}{2}}} < \sigma^2 < \frac{(n-1)\sigma^2(\overrightarrow{X})}{h_{\frac{1-\gamma}{2}}}\} = \gamma$

т.е.:

$\underline{\sigma}^2(\overrightarrow{X}) = \frac{(n-1)\sigma^2(\overrightarrow{X})}{h_{\frac{1+\gamma}{2}}}$

$\overline{\sigma}^2(\overrightarrow{X}) = \frac{(n-1)\sigma^2(\overrightarrow{X})}{h_{\frac{1-\gamma}{2}}}$

6.1 Модуль 2. Проверка параметрических гипотез. Основные понятия

Пусть X - случайная величина, закон распределения которой неизвестен, или известен не полностью. По определению статистической гипотезой называется любое утверждение о законе распределения случайной величины X.

Определение - статистическая гипотеза называется простой, если она однозначно определяет закон распределения случайной величины X, т.е. однозначно определяет функцию распределения случайной величины X.

Определение - статистическая гипотеза называется параметрической, если она является утверждением относительно значений неизвестного параметра, общий вид которого известен.

В противном случае гипотеза называется непараметрической.

Примеры:

1) Пусть известно, что случайная величина X имеет нормальное распределение. где параметры m и σ^2 - неизвестны.

Рассмотрим гипотезу:

$H_1 = \{m = 0, \sigma^2 = 1\}$ - простая параметрическая гипотеза

$H_2 = \{m > 0, \sigma^2 = 1\}$ - сложная параметрическая

$H_3 = \{m = 1\}$ - сложная параметрическая

2) Пусть X - случайная величина, закон распределения которой неизвестен. Рассмотрим гипотезы:
$H_1 = \{X\tilde{N}(0{,}1)\}$ - простая непараметрическая гипотеза
$H_2 = \{X\tilde{N}(m, \sigma^2)\}$ - сложная непараметрическая
$H_3 = \{N(1, \sigma^2)\}$ - сложная непараметрическая

Задачу проверки гипотез обычно ставят следующим образом:
1) Выдвигают гипотезу H_0, которую называют основной
2) Затем выдвигают H_1, которую называют альтернативной или конкурирующей, при этом гипотезы H_0 и H_1 не пересекаются, но возможно их сумма не исчерпывает все возможные случаи
3) На основании выборки $\overrightarrow{x} \in X$ принимают решение об истинности либо основной гипотезы, либо конкурирующей.

Определение - правило, в соответствии с которым принимается решение об истинности гипотез H_0, H_1, называется критерием проверки гипотез. Как правило критерий задают с использованием критического множества. $W \subseteq X_n$, при этом соответствующее правило имеет вид:
если $\overrightarrow{x} \in W \Rightarrow$ отклонить H_0, принять H_1
если $\overrightarrow{x} \in X_n \setminus W \Rightarrow$ принять H_0, отклонить H_1

Замечание:
1) Критерий полностью задаётся критическим множеством W
2) Множество $X_n \setminus W$ называется доверительным

При использовании любого критеря возмодны ошибки двух видов:
1) Принять конкурирующую гипотезу при условии, что истинная основная:
$\alpha = P\{\overrightarrow{X} \in W | H_0\}$
2) Ошибка второго рода - принять основную гипотезу при истинности конкурирующей. Вероятность совершения этой ошибки:
$\beta = P\{\overrightarrow{X} \in X_n \setminus W | H_1\}$
Величину $1 - \beta$ называют мощностью критерия.

Замечание:
Конечно, хотелось бы построить критерий так, чтобы вероятности совершения обеих этих ошибок были минимальны, однако это невозможно, поэтому при построении критерия обычно фиксируют уровень значимости (т.е. α) и максимизируют мощность. То есть минимизируют β при фиксированном α.

6.2 Критерий Неймана-Пирсона приверки двух простых гипотез

Пусть X - случайная величина, общий вид закона распределения которой известен с точностью до неизвестного параметра θ.
Рассмотрим задачу проверки двух простых гипотез:
$H_0 = \{\theta = \theta_0\}$ против $H_1 = \{\theta = \theta_1\}$, где $\theta_0 \neq \theta_1$.
Введём статистику:
$\phi(\overrightarrow{X}) = \frac{L(\overrightarrow{X}, \theta_1)}{L(\overrightarrow{X}, \theta_0)}$
где L - функция правдоподобия случайной величины X.

Определение - статистику ϕ называют отношением правдоподобия.

Понятно, что чем больше занчение $\phi(\overrightarrow{x})$, тем более предпочтительной будет гипотеза H_1, поэтому критическое множество в рассматриваемой задаче будет иметь вид:
$W = \{\overrightarrow{x} \in X_n : \phi(\overrightarrow{x}) \geqslant C_\phi\}$
для некоторой подхощядей константы C_ϕ, которая получается из уравнения:
$P\{\phi(\overrightarrow{X}) \geqslant C_\phi | H_0\} = \alpha$
Так как проверяется две простые гипотезы, то при выполнении условия C_ϕ значение β вероятности совершит ошибку второго рода и будет отвергнуто однозначно.

Замечание - построенный критерий называется критерием Неймана-Пирсона для простых гипотез.

Замечание:
Если X - непрерывная случайная величина, то:
$P = \{\overrightarrow{X} \in W | H_0\} = \int ... \int\limits_W f_{\overrightarrow{X}}(t_1, ..., t_n, \theta_0) dt_1, ..., dt_n$, $\theta = \theta_0$, так как $H = H_0$
Так как $f_{\overrightarrow{X}}(t_1, ..., t_n, \theta) = f_{X_1}(t_1, ..., t_n, \theta) \cdot ... \cdot f_{X_n}(t_1, ..., t_n, \theta) = f(t_1, \theta) \cdot ... \cdot f(t_n, \theta) =$ по определению функции правдоподобия $= L(t_1, ..., t_n, \theta)$
То:
$P = \{\overrightarrow{X} \in W | H_0\} = \int ... \int\limits_W f_{\overrightarrow{X}}(t_1, ..., t_n, \theta_0) dt_1, ..., dt_n = \int ... \int\limits_W L(t_1, ..., t_n) dt_1, ..., dt_n = \alpha, \{(t_1, t_n) : \phi(t_1, ..., t_n) \geqslant C_\phi\}$

Пример:
Пусть $X\tilde{N}(m, \sigma^2)$, где сигма - известна, m - неизвестна. Рассмотрим задачу проверки гипотезы: $H_0 = \{m = m_0\}$ против $H_1 = \{m = m_1\}$, где $m_1 > m_0$ (2 простые параметрические гипотезы), используем критерий Неймана-Пирсона. Составим отношение правдоподобия:
$\phi(\overrightarrow{X}) = \frac{L(\overrightarrow{X}, m_1)}{L(\overrightarrow{X}, m_0})$

Запишем $L(\overrightarrow{X}, m) = \frac{1}{\sqrt{2\pi}\sigma^n} e^{-\frac{1}{2\sigma^2}\sum\limits_{i=1}^{n}(X_i - n_i)^2}$

Тогда $\phi(\overrightarrow{X}) = \frac{L(\overrightarrow{X}, m_1)}{L(\overrightarrow{X}, m_0})} = e^{-\frac{1}{\sigma^2}[\sum\limits_{i=1}^{n}(X_i - m_1)^2 - \sum\limits_{i=1}^{n}(X_i - m_0)^2]} = ... = e^{\frac{m_1 - m_0}{2\sigma^2}\sum 2X_i} \cdot e^{-\frac{m_1 - m_0}{2\sigma^2}\sum(m_0 + m_1)}$