

МАТЕМАТИЧЕСКАЯ СТАТИСТИКА 2

Основные темы

- Проверка статистических гипотез
- Критерии согласия:
 - Критерий Пирсона (хи-квадрат),
 - Критерий Колмогорова.
- Параметрические критерии:
 - Критерий Фишера (сравнение дисперсий);
 - Сравнение математических ожиданий
 - при больших и малых объёмах выборок.
- Ранговые критерии
- Однофакторный дисперсионный анализ
- Элементы теории корреляции

Проверка статистических гипотез

§ Основные сведения

Статистической гипотезой называется любое предположение о виде или параметрах неизвестного закона распределения.

Проверяемую гипотезу называют *нулевой (основной)*, обозначают её H_0 .

Конкурирующей (альтернативной) называют гипотезу, которая противоречит нулевой, обозначают её H_1 .

Задача: проверить, верна ли нулевая гипотеза H_0 при альтернативной гипотезе H_1 ?

Гипотеза H_0	Принимается	Отвергается
Верна	Правильное решение	Ошибка 1-го рода
Неверна	Ошибка 2-го рода	Правильное решение

Обозначим через α – вероятность допустить ошибку 1-го рода, через β – вероятность ошибки 2-го рода.

Вероятность α допустить ошибку 1-го рода, то есть отвергнуть верную гипотезу H_0 , называют ***уровнем значимости***.

1. Задаём уровень значимости α .
2. Строим случайную величину K , называемую *статистическим критерием*, для которой выполняются следующие условия:
 - 1) она является функцией от выборочных данных:
$$K = K(x_1, x_2, \dots, x_n);$$
 - 2) её значения позволяют судить о «расхождении выборки с гипотезой H_0 », то есть о том, надо принимать или отвергать гипотезу H_0 ;
 - 3) распределение этой величины известно.

3. Вычисляем значения критерия, подставляя в него выборочные данные. Это число называют *наблюдаемым значением критерия* и обозначают $K_{набл}$.

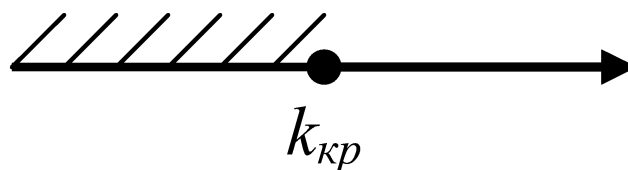
4. Находим *критическую область* данного критерия, то есть совокупность значений критерия, при которых нулевую гипотезу отвергают. Все остальные значения критерия образуют область, называемую *областью принятия нулевой гипотезы*.

5. Если наблюдаемое значение критерия попадает в критическую область, то нулевую гипотезу отвергаем, в противном случае нулевую гипотезу принимаем.

Точки, которые отделяют критическую область от области принятия гипотезы, называют **критическими точками**.

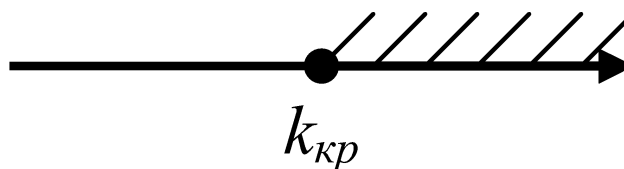
Чаще всего встречаются следующие виды критических областей:

а) левосторонняя



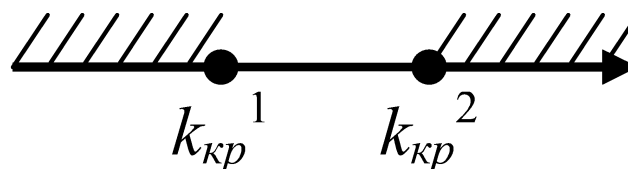
$$K < k_{кр}$$

б) правосторонняя



$$K > k_{кр}$$

в) двусторонняя



$$K < k_{кр}^1$$
$$K > k_{кр}^2$$

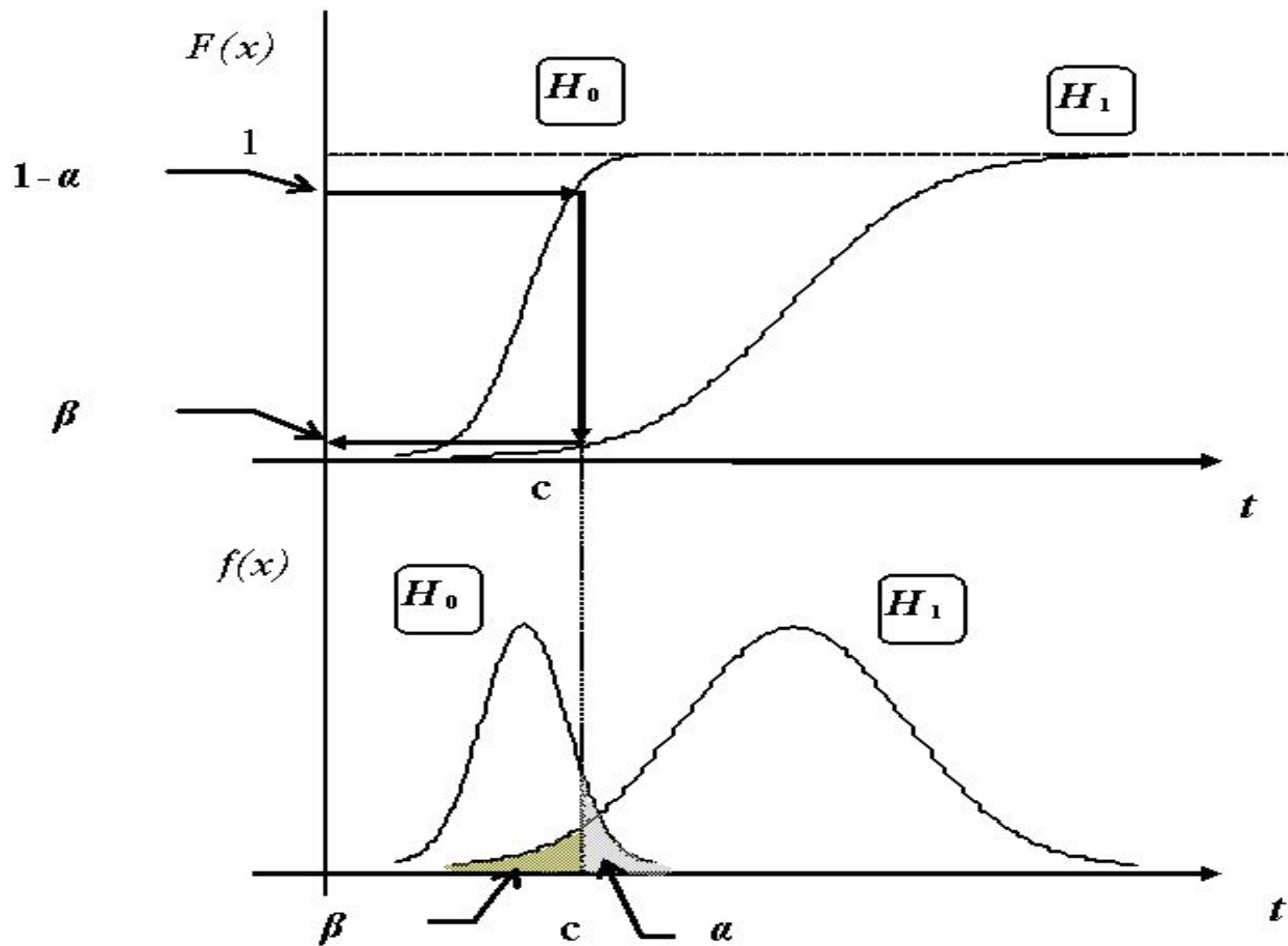
Критическую область W целесообразно находить согласно следующим требованиям:

1. $p(K \in W) = \alpha$
2. вероятность β ошибки 2-го рода – минимальная, то есть вероятность $(1 - \beta)$ – максимальная

Вероятность $(1 - \beta)$ не допустить ошибку 2-го рода, то есть отвергнуть гипотезу H_0 , когда она неверна, называется **мощностью** критерия.

Мощность критерия – максимальная.

- При разработке статистического критерия невозможно одновременно минимизировать обе ошибки. Поэтому поступают следующим образом: при заданном числе испытаний n устанавливается верхняя граница для ошибки первого рода.
- Выбирается тот критерий, у которого наименьшая ошибка второго рода.



Пять шагов проверки

ГИПОТЕЗЫ

1. Сформулировать нулевую H_0 и альтернативную H_1 гипотезы.
2. Выбрать статистику критерия $T(X)$ и уяснить её закон распределения.
3. Задать уровень значимости критерия. По таблицам квантилей распределения статистики найти **критические точки** и указать критическую область.
4. Подсчитать **наблюдаемое значение статистики критерия** и проверить условие его попадания в критическую область.
5. Сделать вывод о принятии нулевой или альтернативной гипотезы.

Критерии, с помощью которых проверяется гипотеза о теоретическом законе распределения, называются *критериями согласия*.

H_0 : генеральная совокупность имеет некоторое определённое распределение

Параметрические критерии тестируют гипотезы о параметрах некоторого распределения :

1. Генеральная совокупность имеет биномиальное распределение с параметрами $m = 10$ и $p = 0.4$.
2. Генеральная совокупность распределена нормально с математическим ожиданием, равным 5 и дисперсией, равной 4.

§ Критерий согласия Колмогорова

Нулевая гипотеза: исследуемая случайная величина имеет заданный закон распределения.

$F(x)$ – теоретическая функция распределения

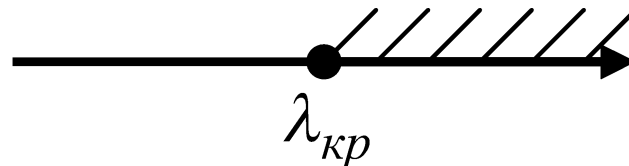
$F_n(x)$ – эмпирическая функция распределения

Обозначим $D = \max |F_n(x) - F(x)|$

– *статистика* критерия Колмогорова

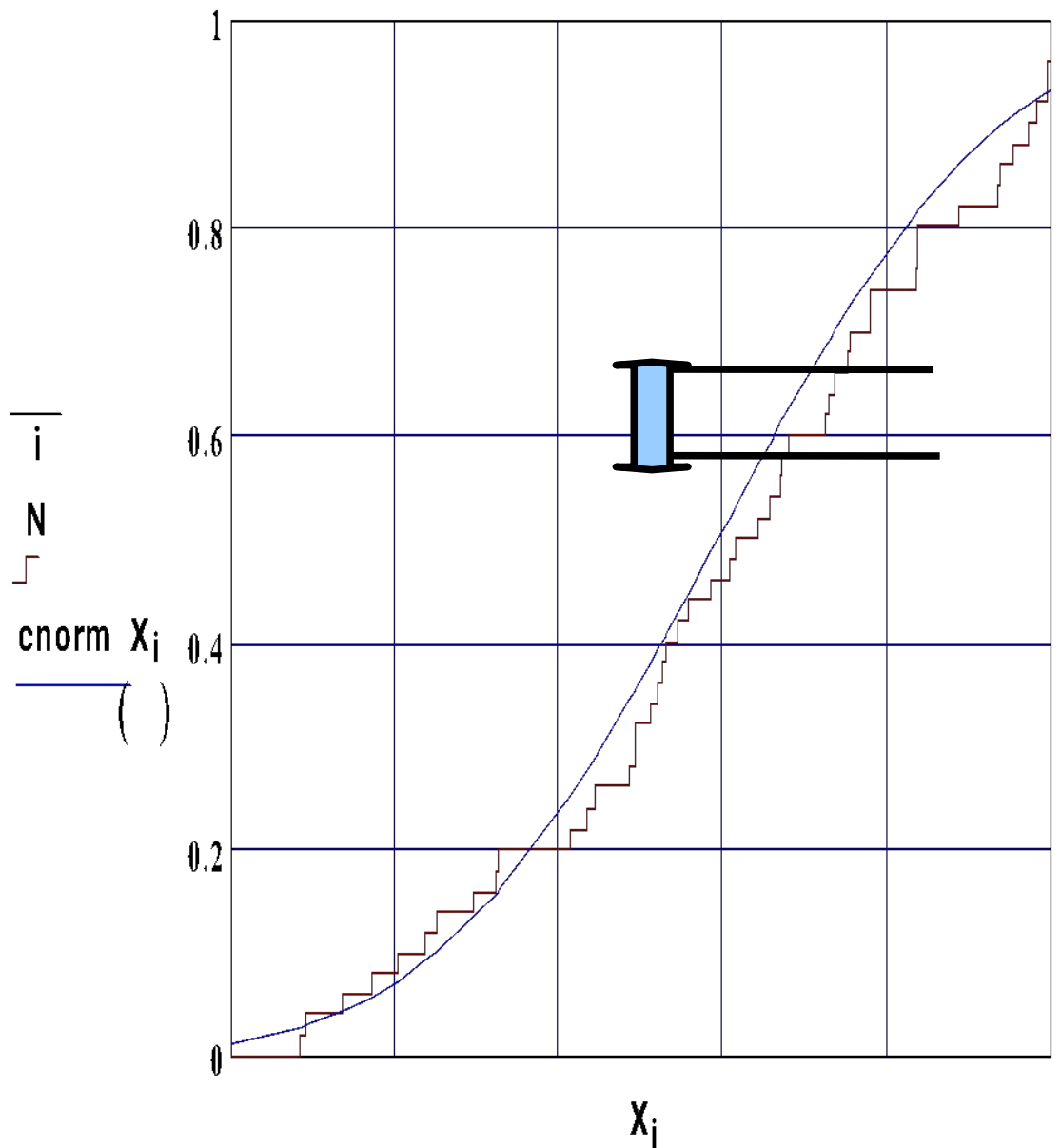
Критерий: $\lambda = D\sqrt{n}$

Критическая область W – правосторонняя:



Из требования для критической области:

$$p(\lambda \in W) = p(\lambda > \lambda_{кр}) = p(D\sqrt{n} > \lambda_{кр}) = \alpha$$



Можно доказать, что при $n \rightarrow \infty$

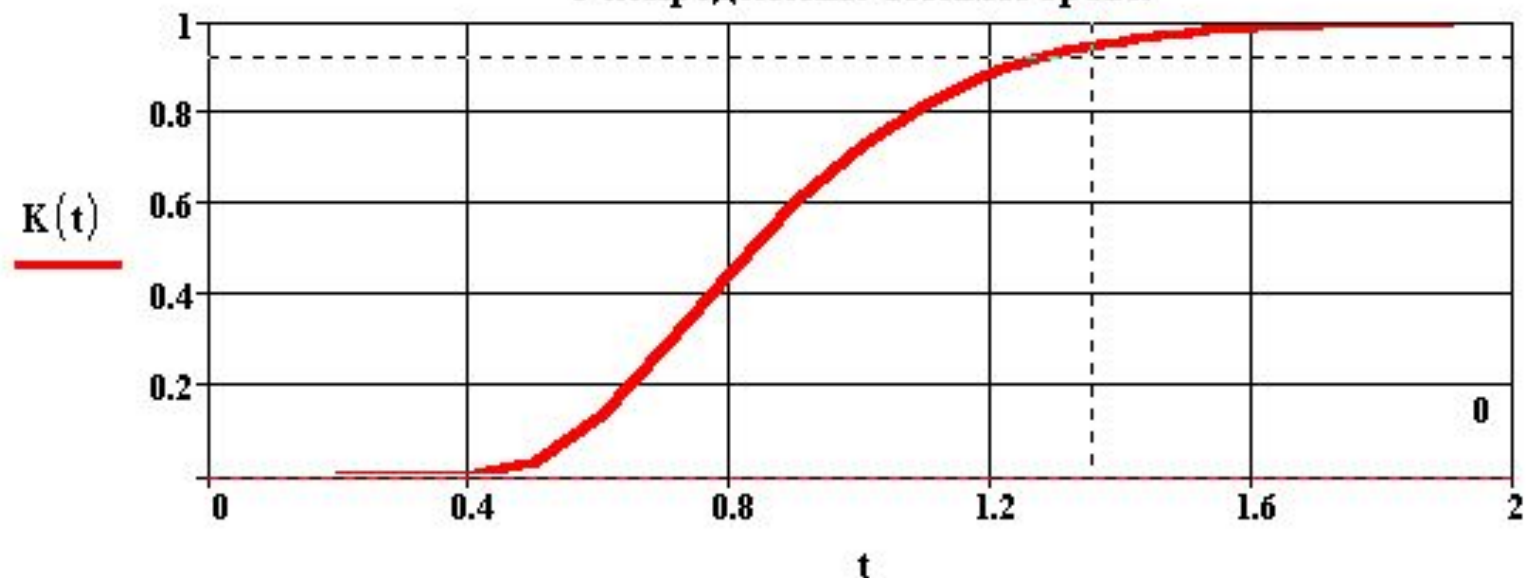
$$p(D\sqrt{n} > \lambda_{kp}) \rightarrow p(\lambda_{kp}) = 1 - \sum_{k=-\infty}^{+\infty} (-1)^k e^{-2k^2 \lambda_{kp}^2}$$

$$\Rightarrow 1 - \sum_{k=-\infty}^{+\infty} (-1)^k e^{-2k^2 \lambda_{kp}^2} = \alpha \quad \Rightarrow \quad \lambda_{kp}$$

α	0.4	0.3	0.2	0.1	0.05	0.025	0.01	0.005	0.001	0.0005
λ_{kp}	0.89	0.97	1.07	1.22	1.36	1.48	1.63	1.73	1.95	2.03

$$K(t) := \sum_{i = -2001}^{2001} \left[[(-1)^{-i}] \cdot e^{-\left(2 \cdot i^2 \cdot t^2\right)} \right] \quad t := 0, 0.1 \dots 2$$

Распределение Колмогорова



Расчет квантилей распределения Колмогорова

$$qK(t, \text{alfa}) := \text{root}(K(t) - \text{alfa}, t)$$

$$qK(1, 0.95) = 1.358099$$

§ Критерий согласия Пирсона (хи-квадрат) χ^2

Найдём *теоретические частоты* вариант.

1. Распределение дискретное $\Rightarrow p(x)$.

x_i	x_1	x_2	\dots	x_{l-1}	x_l
p_i	$p_1=p(x_1)$	$p_2=p(x_2)$	\dots	$p_{l-1}=p(x_{l-1})$	$p_l=1-p_1-p_2-\dots-p_{l-1}$

Теоретическая частота появления варианты x_i – это np_i .

2. Распределение непрерывное $\Rightarrow F(x)$.

x_i	(x_1, x_2)	(x_2, x_3)	\dots	(x_{l-1}, x_l)	(x_l, x_{l+1})
p_i	$p_1=p(X < x_2)$ $=F(x_2)$	$p_2=p(x_2 < X < x_3)$ $=F(x_3) - F(x_2)$	\dots	$p_{l-1}=p(x_{l-1} < X < x_l)$ $=F(x_l) - F(x_{l-1})$	$p_l=1-p_1-p_2-\dots-p_{l-1}$

Теоретическая частота попадания в интервал (x_i, x_{i+1}) – это np_i .

Критерий:
$$\chi^2 = \sum_{i=1}^l \frac{(n_i - np_i)^2}{np_i}$$

n_i – эмпирические частоты

np_i – теоретические частоты

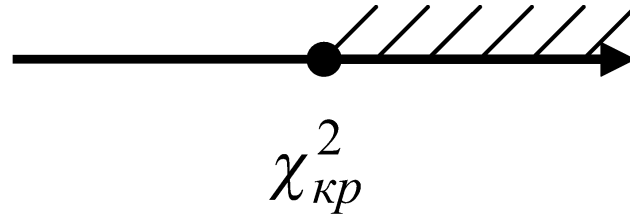
При $n \rightarrow \infty$ случайная величина χ^2 имеет распределение Пирсона с k степенями свободы, где

$$k = l - r - 1,$$

l – число вариантов (интервалов),

r – число параметров предполагаемого распределения, оцениваемых по выборке.

Критическая область W – правосторонняя:



Из требования для критической области:

$$p(\chi^2 \in W) = \alpha \quad \Rightarrow \quad p(\chi^2 > \chi_{кр}^2) = \alpha$$

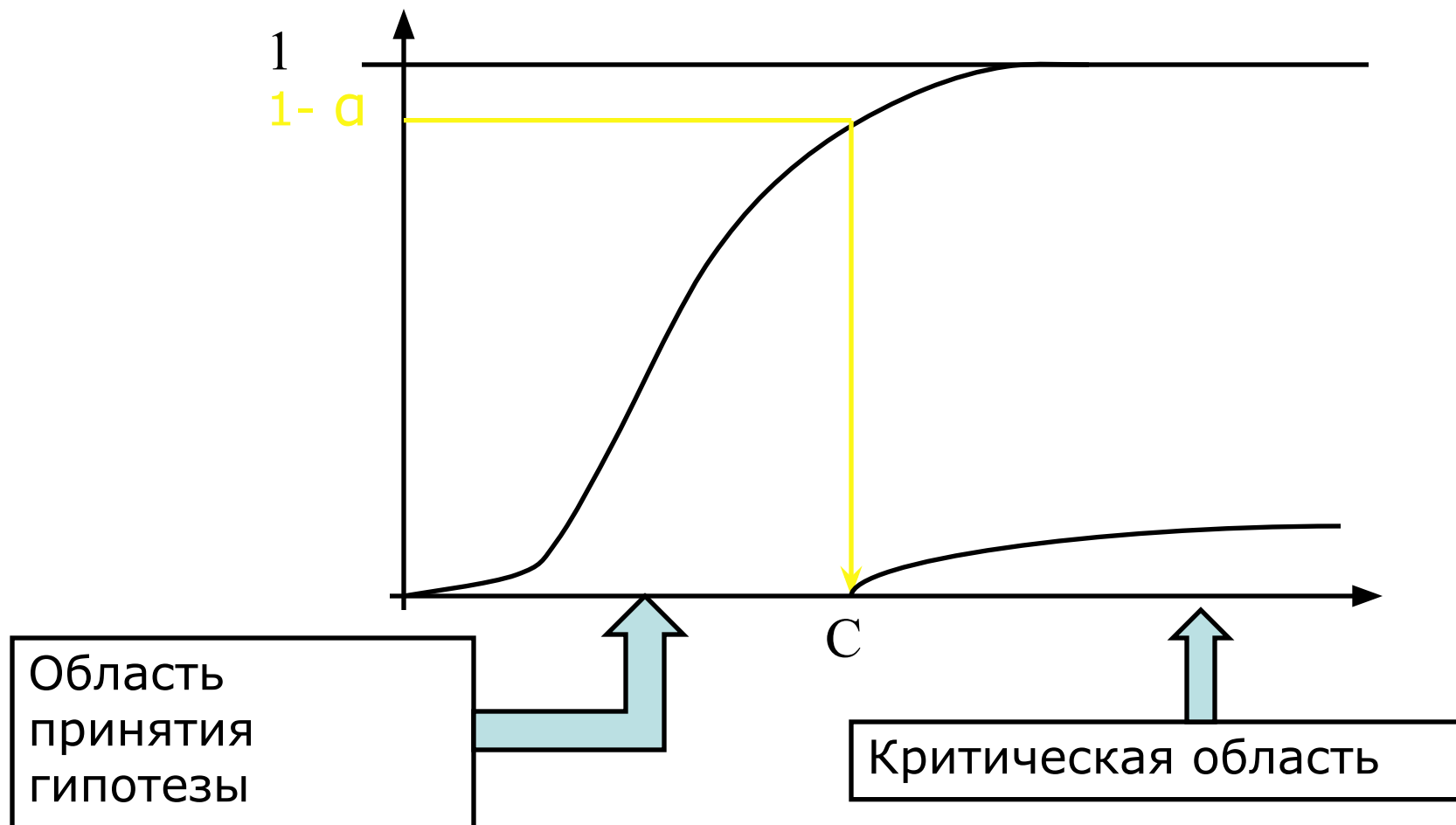
$$p(\chi^2 < \chi_{кр}^2) = 1 - p(\chi^2 > \chi_{кр}^2) = 1 - \alpha$$

$$p(\chi^2 < \chi_{кр}^2) = F(\chi_{кр}^2), \quad F(x) - \text{функция распределения } \chi^2$$

$$F(\chi_{кр}^2) = 1 - \alpha \quad \Rightarrow \quad \chi_{кр}^2 = F^{-1}(1 - \alpha)$$

$F(x)$ – функция распределения Пирсона с $k = l - r - 1$ степенями свободы, l – число вариантов (интервалов), r – число параметров, оцениваемых по выборке.

Для нахождения критической области необходимо по заданной вероятности ошибки первого рода (уровню значимости критерия) α найти квантиль хи-квадрат распределения на уровне $1 - \alpha$.



Алгоритм применения критерия согласия Пирсона

Подсчитываем значение статистики критерия и сравниваем его с критической точкой. Если статистика критерия попадает в критическую область,

$$\chi^2 > C$$

то нулевая гипотеза: исследуемая случайная величина имеет заданный закон распределения отвергается.

В противном случае она принимается на уровне значимости α

Критерий легко приспособливается и для непрерывных распределений путем их *дискретизации*.

Проверку гипотезы удобно совмещать с построением гистограмм.

§ Критерий Фишера

Две генеральные совокупности X и Y распределены нормально.

Проверить гипотезу: $H_0 : D(X) = D(Y)$

Обозначим n_X – объём выборки из совокупности X ,
 n_Y – объём выборки из совокупности Y ,
 s^2_X и s^2_Y – исправленные выборочные дисперсии.

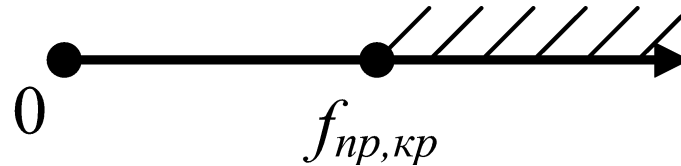
Критерий:
$$F = \frac{s^2_X}{s^2_Y}$$

F имеет распределение Фишера с $(n_X - 1)$ и $(n_Y - 1)$ степенями свободы

Критическая область строится в зависимости от вида конкурирующей гипотезы.

1. $H_1 : D(X) > D(Y)$

Критическая область W – правосторонняя:



Так как $s^2_X > 0$ и $s^2_Y > 0$, то $F > 0 \Rightarrow$ положительная часть

Из требования 1 для критической области:

$$p(F \in W) = \alpha \quad \Rightarrow \quad p(F > f_{np,kr}) = \alpha$$

$$p(F < f_{np,kr}) = 1 - p(F > f_{np,kr}) = 1 - \alpha$$

$p(F < f_{np,kr}) = F(f_{np,kr})$, $F(x)$ – функция распределения F

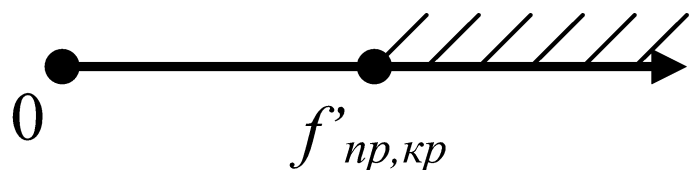
$$F(f_{np,kr}) = 1 - \alpha \quad \Rightarrow \quad f_{np,kr} = F^{-1}(1 - \alpha)$$

$F(x)$ – функция распределения Фишера с $(n_X - 1)$ и $(n_Y - 1)$ степенями свободы

2. $H_1 : D(X) < D(Y)$

Обозначим $F' = \frac{1}{F} = \frac{s_Y^2}{s_X^2}$, F' имеет распределение Фишера с $(n_Y - 1)$ и $(n_X - 1)$ степенями свободы

$H_1 : D(Y) > D(X) \Rightarrow$ предыдущий случай:



$f'_{np,kr} = F^{-1}(1 - \alpha)$, где $F(x)$ — функция распределения F'

$$p(F' > f'_{np,kr}) = \alpha$$

$$p(F' > f'_{np,kr}) = p(1/F > f'_{np,kr}) = p(F < 1/f'_{np,kr})$$

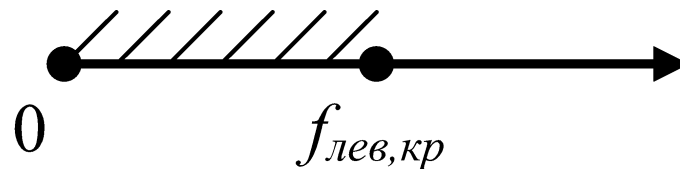
$$\Rightarrow p(F < 1/f'_{np,kr}) = \alpha$$

$$p(F < 1/f'_{np,kr}) = \alpha$$

Обозначим $f'_{лев,kr} = 1/f'_{np,kr}$, тогда

$$p(F < f'_{лев,kr}) = \alpha$$

Таким образом, критическая область для критерия F имеет вид:

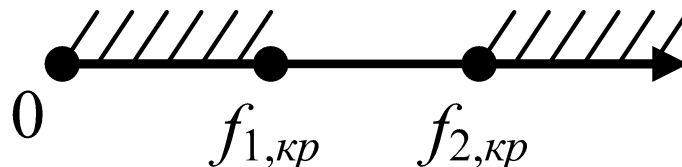


$$f'_{лев,kr} = \frac{1}{f'_{np,kr}} = \frac{1}{F^{-1}(1-\alpha)}, \text{ где } F(x) \text{ — функция}$$

распределения Фишера
с $(n_Y - 1)$ и $(n_X - 1)$
степенями свободы

3. $H_1 : D(X) \neq D(Y)$

Критическая область W – двусторонняя:



Пусть $p(F < f_{1,кр}) = p(F > f_{2,кр}) = \alpha / 2$

Аналогично пунктам **1** и **2** получаем:

$$f_{2,кр} = F_1^{-1}(1 - \alpha / 2)$$

где $F_1(x)$ – функция распределения Фишера с $(n_X - 1)$ и $(n_Y - 1)$ степенями свободы

$$f_{1,кр} = \frac{1}{f'_{2,кр}} = \frac{1}{F_2^{-1}(1 - \alpha / 2)}$$

где $F_2(x)$ – функция распределения Фишера с $(n_Y - 1)$ и $(n_X - 1)$ степенями свободы

§ Критерий Стьюдента (t -критерий)

Генеральная совокупность распределена нормально.

Проверить гипотезу: $H_0 : a = a_0$

a_0 – некоторое число

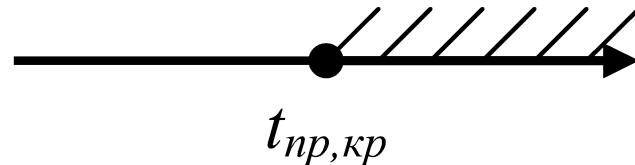
Критерий:
$$T = \frac{\bar{x}_v - a}{s} \cdot \sqrt{n}$$

T имеет распределение Стьюдента с $(n-1)$ степенями свободы

Критическая область строится в зависимости от вида конкурирующей гипотезы.

$$1. \quad H_1 : a > a_0$$

Критическая область W – правосторонняя:



Из требования 1 для критической области:

$$p(T \in W) = \alpha \quad \Rightarrow \quad p(T > t_{np,kr}) = \alpha$$

$$p(T < t_{np,kr}) = 1 - p(T > t_{np,kr}) = 1 - \alpha$$

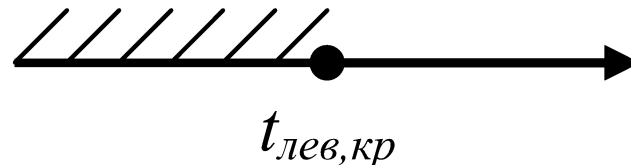
$$p(T < t_{np,kr}) = F(t_{np,kr}), \quad F(x) - \text{функция распределения } T$$

$$F(t_{np,kr}) = 1 - \alpha \quad \Rightarrow \quad t_{np,kr} = F^{-1}(1 - \alpha)$$

$F(x)$ – функция распределения Стьюдента с $(n-1)$ степенями свободы

2. $H_1 : a < a_0$

Критическая область W – левосторонняя:



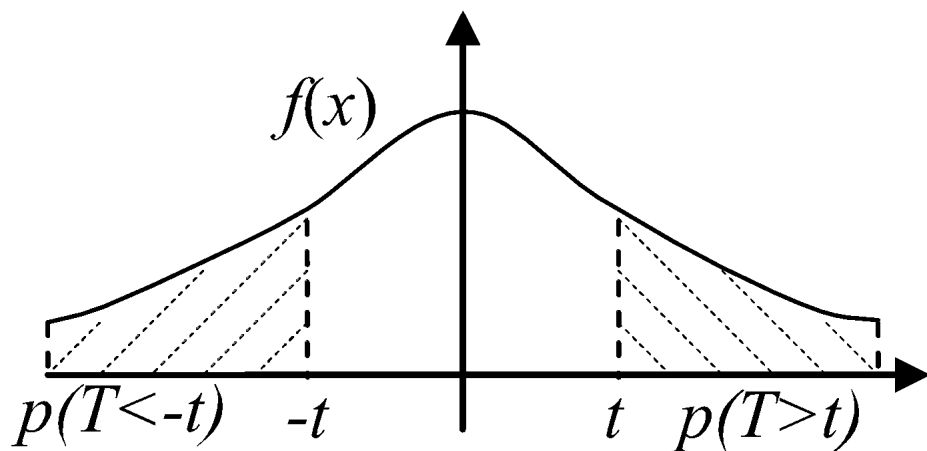
Из требования 1 для критической области:

$$p(T \in W) = \alpha \quad \Rightarrow \quad p(T < t_{лев,кр}) = \alpha$$

$p(T < t_{лев,кр}) = F(t_{лев,кр}) = \alpha$, $F(x)$ – функция
распределения T

$\Rightarrow t_{лев,кр} = F^{-1}(\alpha)$, $F(x)$ – функция распределения
Стьюдента с $(n-1)$ степенями
свободы

Плотность распределения Стьюдента – чётная функция



$$\Rightarrow p(T > t) = p(T < -t)$$

Критическая точка $t_{np,kr}$ находится из требования:

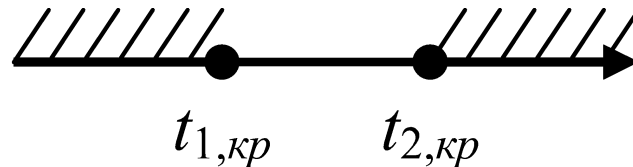
$$p(T > t_{np,kr}) = \alpha \quad \Rightarrow \quad p(T < -t_{np,kr}) = \alpha \quad \Rightarrow$$

$-t_{np,kr}$ является критической точкой для левосторонней области:

$$t_{лев,kr} = -t_{np,kr}$$

3. $H_1 : a \neq a_0$

Критическая область W – двусторонняя:



Пусть $p(T < t_{1,кр}) = p(T > t_{2,кр}) = \alpha / 2$

В силу чётности плотности распределения Стьюдента:

$$t_{1,кр} = -t_{2,кр}$$

Аналогично пунктам **1** и **2** получаем:

$$t_{2,кр} = F^{-1}(1 - \alpha / 2), \quad t_{1,кр} = -t_{2,кр}$$

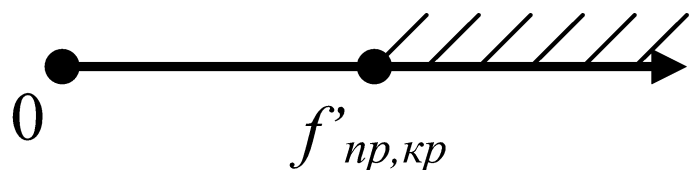
ИЛИ

$$t_{1,кр} = F^{-1}(\alpha / 2), \quad t_{2,кр} = -t_{1,кр}$$

2. $H_1 : D(X) < D(Y)$

Обозначим $F' = \frac{1}{F} = \frac{s_Y^2}{s_X^2}$, F' имеет распределение Фишера с $(n_Y - 1)$ и $(n_X - 1)$ степенями свободы

$H_1 : D(Y) > D(X) \Rightarrow$ предыдущий случай:



$f'_{np,kr} = F^{-1}(1 - \alpha)$, где $F(x)$ — функция распределения F'

$$p(F' > f'_{np,kr}) = \alpha$$

$$p(F' > f'_{np,kr}) = p(1/F > f'_{np,kr}) = p(F < 1/f'_{np,kr})$$

$$\Rightarrow p(F < 1/f'_{np,kr}) = \alpha$$

Однофакторный дисперсионный анализ

Пример: выявить зависимость объёма выполненных на стройке работ за смену от работающей бригады.

Номер бригады \ Номер наблюдения	1	2	3	4
1	20	27	18	23
2	25	31	19	23
3	22	22	29	21
4	24	32	26	20
5	30		28	26
6	23			25
Средний объём	24	28	24	23

X – случайная величина

F – фактор, воздействующий на случайную величину X

F_1, F_2, \dots, F_p – уровни фактора

a_1, a_2, \dots, a_p – математические ожидания на уровнях F_1, F_2, \dots, F_p соответственно

$$H_0: a_1 = a_2 = \dots = a_p$$

Дисперсионным анализом называется статистический метод, предназначенный для выявления влияния отдельных факторов на результат эксперимента, а также для последующего планирования эксперимента.

Критерий Бартлетта

$$H_0: D_1(X) = D_2(X) = \dots = D_p(X)$$

гипотеза о равенстве дисперсий на каждом уровне

q_1, q_2, \dots, q_p – количество наблюдений на уровнях
 F_1, F_2, \dots, F_p соответственно

$s_1^2, s_2^2, \dots, s_p^2$ – исправленные выборочные дисперсии
на уровнях F_1, F_2, \dots, F_p соответственно

$$s_0^2 = \frac{\sum_{i=1}^p (q_i - 1) s_i^2}{\sum_{i=1}^p (q_i - 1)}, \quad Q = \left(1 + \frac{1}{3(p-1)} \left[\sum_{i=1}^p \frac{1}{q_i - 1} - \frac{1}{\sum_{i=1}^p (q_i - 1)} \right] \right)^{-1}$$

Критерий:
$$\varphi = Q \cdot \sum_{i=1}^p \left((q_i - 1) \cdot \ln \frac{s_0^2}{s_i^2} \right)$$

Если $q_1, q_2, \dots, q_p > 3$, то критерий имеет распределение, близкое к распределению Пирсона с $(p-1)$ степенями свободы.

Критическая область – правосторонняя.

$$p(\varphi \in W) = p(\varphi > \varphi_{кр}) = \alpha \quad \Rightarrow \quad p(\varphi < \varphi_{кр}) = 1 - \alpha$$
$$p(\varphi < \varphi_{кр}) = F(\varphi_{кр})$$

$$F(\varphi_{кр}) = 1 - \alpha \quad \Rightarrow \quad \varphi_{кр} = F^{-1}(1 - \alpha),$$

где $F(x)$ – функция распределения Пирсона с $(p-1)$ степенями свободы.

Уровень фактора F	F_1	F_2	...	F_p
Номер наблюдения				
1	x_{11}	x_{12}	...	x_{1p}
2	x_{21}	x_{22}	...	x_{2p}
...				
Число наблюдений	q_1	q_2	...	q_p
Среднее значение	y_1	y_2	...	y_p

$$H_0: a_1 = a_2 = \dots = a_p$$

Объём выборки: $n = q_1 + q_2 + \dots + q_p$

Уровень фактора F	F_1	F_2	\dots	F_p
Номер наблюдения				
1	x_{11}	x_{12}	\dots	x_{1p}
2	x_{21}	x_{22}	\dots	x_{2p}
\dots				
Число наблюдений	q_1	q_2	\dots	q_p
Среднее значение	y_1	y_2	\dots	y_p

1-ая группа – уровень F_1 : $x_{11}, x_{21}, \dots, x_{q_1 1}$

2-ая группа – уровень F_2 : $x_{21}, x_{22}, \dots, x_{q_2 2}$

\dots

p -ая группа – уровень F_p : $x_{1p}, x_{2p}, \dots, x_{q_p p}$

$$D_v = D_{\text{межгр}} + D_{\text{внгр}}$$

1-ая группа – уровень F_1 : $x_{11}, x_{21}, \dots, x_{q_1 1}$

2-ая группа – уровень F_2 : $x_{21}, x_{22}, \dots, x_{q_2 2}$

...

p -ая группа – уровень F_p : $x_{1p}, x_{2p}, \dots, x_{q_p p}$

$$1. \quad D_{\text{межгр}} = \frac{\sum_{i=1}^p q_i (y_i - \bar{x}_v)^2}{n}$$

Факторная сумма:

$$S_{\text{факт}} = \sum_{i=1}^p q_i (y_i - \bar{x}_v)^2$$

$$2. \quad D_{\text{внгр}} = \frac{\sum_{i=1}^p q_i D_{igr}}{n}, \text{ где } D_{igr} \text{ – дисперсия } i\text{-той группы}$$

i -тая группа: $x_{1i}, x_{2i}, \dots, x_{q_i i}$, групповая средняя: y_i

$$D_{igr} = \frac{\sum_{j=1}^{q_i} (x_{ji} - y_i)^2}{q_i}$$

$$D_{внгр} = \frac{\sum_{i=1}^p q_i D_{igr}}{n} = \frac{\sum_{i=1}^p q_i \left(\sum_{j=1}^{q_i} (x_{ji} - y_i)^2 / q_i \right)}{n} = \frac{\sum_{i=1}^p \sum_{j=1}^{q_i} (x_{ji} - y_i)^2}{n}$$

Остаточная сумма:

$$S_{ост} = \sum_{i=1}^p \sum_{j=1}^{q_i} (x_{ji} - y_i)^2$$

Факторная дисперсия: $s_{\text{факт}}^2 = \frac{S_{\text{факт}}}{p - 1}$

Остаточная дисперсия: $s_{\text{ост}}^2 = \frac{S_{\text{ост}}}{n - p}$

$D(x) \approx s_{\text{ост}}^2$ – всегда

$D(x) \approx s_{\text{факт}}^2$ – если несущественно влияние фактора

$$H_0: a_1 = a_2 = \dots = a_p$$

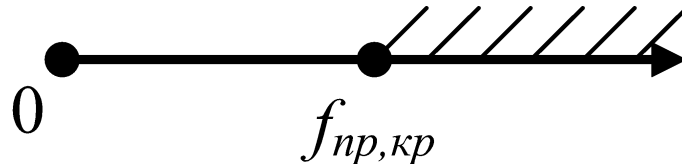
$$H_0: S_{\text{ост}}^2 = S_{\text{факт}}^2$$

$$H_0 : S_{ост}^2 = S_{факт}^2$$

Критерий: $F = \frac{S_{факт}^2}{S_{ост}^2}$ имеет распределение Фишера с $(p-1)$ и $(n-p)$ степенями свободы

$$H_1 : S_{факт}^2 > S_{ост}^2$$

Критическая область W – правосторонняя:



Из требования 1 для критической области:

$$p(F \in W) = \alpha \Rightarrow p(F > f_{np,kr}) = \alpha \Rightarrow f_{np,kr} = F^{-1}(1 - \alpha)$$

$F(x)$ – функция распределения Фишера с $(p-1)$ и $(n-p)$ степенями свободы

Элементы теории корреляции

Зависимость величины Y от X называется *функциональной*, если каждому значению величины X соответствует единственное значение величины Y .

Зависимость величины Y от X называется *статистической (вероятностной, стохастической)*, если каждому значению величины X соответствует не одно, а множество значений величины Y , причём сказать заранее, какое именно значение примет величина Y невозможно.

Среднее значение, которое принимает величина Y при $X=x$, называется математическим ожиданием случайной величины Y , вычисленным при условии, что $X=x$, или *условным математическим ожиданием*:

$$M(Y|X=x)$$

Если при изменении x условные математические ожидания $M(Y|X=x)$ изменяются, то говорят, что имеет место *корреляционная зависимость* величины Y от X .

При этом функцию $f(x)=M(Y|X=x)$ называют *функцией регрессии*.

$$f(x)=M(Y|X=x) - ?$$

$$f(x) = M(Y|X=x) - ?$$

Условным средним \bar{y}_x называют среднее арифметическое наблюдавшихся значений Y , соответствующих $X=x$.

Условное среднее является оценкой условного математического ожидания: $M(Y|X=x) \approx \bar{y}_x$

Каждому x соответствует своё значение \bar{y}_x , следовательно, \bar{y}_x – есть функция от x :

$$\bar{y}_x = f^*(x)$$

это уравнение называется **выборочным уравнением регрессии**, а функция $f^*(x)$ – **выборочной функцией регрессии**.

$$f(x) \approx f^*(x)$$

$$f(x) = M(Y|X=x) - ?$$

Если функция регрессии – линейная:

$$f(x) = M(Y|X=x) = ax + b,$$

то выборочное уравнение регрессии имеет вид:

$$\bar{y}_x = \bar{y} + r_{xy} \frac{\sigma_Y}{\sigma_X} (x - \bar{x}), \quad \text{где } r_{xy} = \frac{\sum_{x,y} n_{xy} xy - n \bar{x} \bar{y}}{n \sigma_X \sigma_Y} - \text{выборочный коэффициент корреляции}$$

\bar{x}, \bar{y} – выборочные средние

σ_X, σ_Y – выборочные средние квадратические отклонения

n_{xy} – частота пары вариантов (x, y)

Корреляционная таблица

$Y \backslash X$	10	20	30	40	n_Y
0.4	5	—	7	14	26
0.6	—	2	6	4	12
0.8	3	19	—	—	22
n_X	8	21	13	18	$n=60$