① Генеральная и выборочная совокупности.

Генеральная совокупность – сов-ть объектов, облад. признаками, дост. для наблюдения и количеств. оценки

Элементы – объекты, сост. ген. сов-ть

Объём – число элементов

Примем каждый из N элементов, как эл. исход. нек. вер. пр-ва с вер-ю $1/N$. Тогда соотв. между объектами и изуч. признаками на рассматрив. вер. пр-ве случ. вел. $\xi = \xi(\omega)$ (где распр. случ. величины ξ – распр. ген. сов-ти)

Изучение ген. сов-ти отн. колич. признака: 1) сплошное обследование; 2) обслед-е части ген. сов-ти

Выборка – часть г.с., выбранная для обследов.

Пусть X – ген. сов-ть.

Извлекая эл-ты и фикс. значение признака получаем $x_1, \dots, x_n$ выборку, где n – объём выборки

Виды отбора из генер. сов-ти:
1) простой случ. безповторный (незав. н. при больш. n)
2) простой случ. повторный (незав. наблюдения)

Выборочный метод – метод, по которому делаются выводы о св-вах и распр. ген. сов-ти

Представительная выборка (репрезент.) – выборка, правильно отобрит. пропорции г.с.
↳ в силу закона о больших числах пр-ее случайно

Способы отбора:
① Не требующий разбиение г.с. на части
– простой случайный безповторный / повторный
② требующий разбиение
– серийный (применяется, когда знач. призн. колебл. незначит.)
– типический (отбор из кажд. типической части; при существ. колебаниях знач. признака)
– механический (г.с. разбивается на n частей, из кот. выб. 1 элем.)
+ может быть нарушена представительность

② Эмпирическая функция распределения

Эмпирическая (выборочная) ф-я распр. $F_n(x) = \frac{n_x}{n}$, где n_x - число значений признаков выборки меньших x, n - объем выборки

Теоретическая ф-я распределение $F(x) = p(\xi < x)$

Теор. эмп. функция распр. сход. при $n \to \infty$ по вер. к теор. функции распр.:

$$\lim_{n \to \infty} p(\underbrace{|F_n(x)}_{\text{отн. част.}} - \underbrace{F(x)}_{\text{вер-ть}}| < \varepsilon) = 1, \quad \forall x \in \mathbb{R}, \ \varepsilon > 0$$

▷ след. из т. Бернулли зак. о больших числах $\lim\limits_{n \to \infty} p(|\frac{m}{n} - p| < \varepsilon) = 1$

Св-ва эмп. ф-и (прибл. к теор. ф-и)

1) $0 \leq F_n(x) \leq 1$

2) x_1 - наим. знач. призн. выборки $\Rightarrow$ $x \leq x_1$ $F_n(x) = 0$
x_n - наиб. -//- $\Rightarrow$ $x \geq x_n$ $F_n(x) = 1$

3) график ступенчатый с т. разрыва $x_1, \dots, x_n$

③ Статистическое распределение выборки

Вариационный ряд — ряд, в кот. значения признака расположены в возрастающем (убыв.) порядке с указ. соотв. частот.

Вариант — каждое отдельное значение признака ($x_1, \dots x_n$)

Статистический ряд — таблица знач. признака и соотв. частот

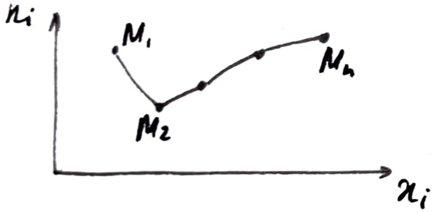
x_1	...	x_n
n_1	...	n_n

1) Дискретный ряд — значения признака удал. друг от друга (разбивается на интервалы при больших объемах выборки)

2) Непрерывный ряд — строятся интервальные ряды (причем в кажд. $[a;b]$ инт. вкл. знач. $a<, \leq b$)

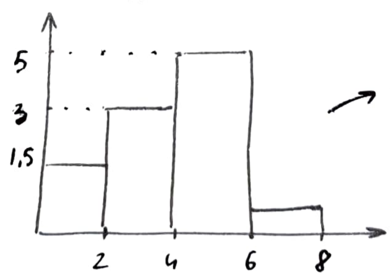
Интервалы одинак. длины $h = \frac{x_{max} - x_{min}}{1 + 3,22 \lg n}$, где $(x_{max} - x_{min})$ — размах

Полигоны частот: $M_k(x_k, n_k)$; аналогично полигон относительных частот $M_k(x_k, \frac{n_k}{n})$



Гистограмма частот — ступенчатая фигура (в основаниях — отн. интервалы, высоты — n_i/h)

h	0-2	2-4	4-6	6-8
n_i	3	6	10	1



→ площадь гистограммы = объему выборки

аналогично гистограмма отн. частот, где высоты — $\frac{n_i}{n \cdot h}$, причем площадь = 1

④ Выборочные числовые характеристики

Выборочные числ. х-ки — ф-и от набл., приближ. оценивающие соотв. числ. х-ки

1) Выборочное среднее — среднее арифм. знач. изучаемого призн. в выборке

$\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$, если все знач. различны

$\bar{x} = \frac{1}{n}\sum^{k} n_i x_i$, если x_i встретилось n_i раз

Св-ва аналогичны мат. ожидани.

$\overline{kx} = k\cdot\bar{x}$

$\overline{x+y} = \bar{x} + \bar{y}$

$\overline{x+c} = \bar{x} + c$

$\frac{1}{n}\sum(x_i - \bar{x}) = \bar{x} - \frac{1}{n}\cdot n\bar{x} = 0$, где $(x_i - \bar{x})$ — откл. знач. призн. от выб. ср-го

групповое среднее — среднее арифм. кажд. из k групп знач. призн.

Св-во (общ. ср. = ср. арифм. групп. ср., взвешенных по объемам групп)

$\bar{x} = \frac{1}{n}(\bar{x}_1\cdot n_1 + \ldots + \bar{x}_k\cdot n_k)$, где $n = \sum n_i$; $\bar{x}_i = \frac{1}{n_i}\sum^{n_i} x_i$

2) Выборочная дисперсия — ср. арифм. квадр. отклон. изуч. призн. от его ср. знач.
(хар-ет меру разброса знач. призн. в выборке отн. средн. знач.)

$\hat{\delta}^2 = \frac{1}{n}\sum^{n}(x_i - \bar{x})^2$, если все знач. различны

$\hat{\delta}^2 = \frac{1}{n}\sum^{k}(x_i - \bar{x})^2\cdot n_i$, если x_i встр. n_i раз

Св-ва $\hat{\delta}^2 \geq 0$

$\hat{\delta}^2(cx) = c^2\hat{\delta}^2(x)$

$\hat{\delta}^2(x+y) = \hat{\delta}^2(x) + \hat{\delta}^2(y)$, x, y — независимые

$\hat{\delta}^2 = \frac{1}{n}\sum x_i^2 - \left(\frac{1}{n}\sum x_i\right)^2 = \overline{x^2} - (\bar{x})^2$

3) Выборочный нач. момент порядка k — ср. арифм. k-х степеней знач. призн. в выборке

$\hat{\alpha}_k = \frac{1}{n}\sum_{i=1}^{n} x_i^k$ (причем $\hat{\alpha}_1 = \frac{1}{n}\sum x_i = \bar{x}$) ; $\hat{\alpha}_k = \frac{1}{n}\sum^{z} n_i x_i^k$, если знач. повтор.

4) Выборочный центр. момент порядка k — ср. арифм. k-х степеней откл. знач. призн. в выборке от его ср. знач.

$\hat{\mu}_k = \frac{1}{n}\sum^{n}(x_i - \bar{x})^k$, $\hat{\mu}_k = \frac{1}{n}\sum^{z}(x_i - \bar{x})^k\cdot n_i$, где $\sum^{z} n_i = n$

⑤ Статистическая оценка

Статистическая оценка $\hat{\theta}_n$ неизвестного пар-ра θ – функция от наблюдений оцениваем. пар-ра θ (при объеме выборки n),

где $\theta_n = f(x_1, \dots, x_n)$ – статистика

1) Несмещенная $\hat{\theta}_n$ – если мат. ожид. оценки при люб. объеме выборки равняется оцениваемому пар-ру

$M(\hat{\theta}_n) = \theta$

2) Состоятельная $\hat{\theta}_n$ – если оценка сходится по вер. к оц. пар-ру

$\lim\limits_{n\to\infty} p(|\hat{\theta}_n - \theta| < \varepsilon) = 1 \quad , \forall \varepsilon > 0$

зам. св-во состоятельности – асимптотическое св-во

3) Эффективная $\hat{\theta}_n$ в нек. классе оценок – если в этом классе дисп. наим.

$\lim\limits_{n\to\infty} \delta^2(\hat{\theta}_n) = \min$

Теор. если $\hat{\theta}_n$ несмещенная и $\delta^2(\hat{\theta}_n) \to 0, \; n \to \infty$,

то $\hat{\theta}_n$ состоятельная

▷ по нер-ву Чебышева $p(|\xi - M(\xi)| < \varepsilon) \geqslant 1 - \dfrac{D(\xi)}{\varepsilon^2}$

$\forall \xi, \varepsilon > 0$

Пусть $\xi = \hat{\theta}_n$: $M(\xi) = M(\hat{\theta}_n) = \theta$

$$p(|\hat{\theta}_n - \theta| < \varepsilon) \geqslant 1 - \frac{D(\hat{\theta}_n)}{\varepsilon^2}$$

$$n\to\infty \quad \lim_{n\to\infty} p(|\hat{\theta}_n - \theta| < \varepsilon) \geqslant 1 - 0, \quad \forall \varepsilon > 0$$

$$\lim_{n\to\infty} p(|\hat{\theta}_n - \theta| < \varepsilon) = 1 \Rightarrow \text{состоятельная}$$

⑥ Несмещенная и состоятельная оценка мат. ожидания

Пусть ГС: $M(\xi) = a$, $D(\xi) = \delta^2$

Теор. $\bar{x}$ – выб. среднее явл. несмещенной и состоятельной оценкой мат. ожидания ГС

▷ для выборки $x_1, \ldots, x_n$ $\quad \bar{x} = \frac{1}{n}\sum^n x_i$

Несмещенность:

$$M(\bar{x}) = \frac{1}{n}\sum^n M(x_i) = \frac{1}{n}\sum^n M(\xi) = \frac{1}{n} \cdot n \cdot a = a \quad \Rightarrow \bar{x} \text{ – несмещ. оценка мат. ож.}$$

Состоятельность:

$$D(\bar{x}) = D\left(\frac{1}{n}\sum^n x_i\right) = \frac{1}{n^2}\sum^n D(x_i) = \frac{1}{n^2}\sum^n D(\xi) = \frac{1}{n^2} \cdot n \cdot \delta^2 = \frac{\delta^2}{n}$$

$$\begin{cases} \lim\limits_{n\to\infty} D(\bar{x}) = \lim\limits_{n\to\infty} \frac{\delta^2}{n} = 0 \\ \text{несмещенность} \end{cases} \Rightarrow \text{состоятельность оценки } \bar{x} \text{ неизв. мат. ож. } a$$

⑦ Несмещенная и состоятельная оценка начальных моментов

<u>Теор.</u> выборочные начальные моменты явл. несмещ. и сост. оценками соответствующих начальных моментов ТС

▷ по выборке $x_1, \dots, x_n$ находим $\hat{\alpha}_k = \frac{1}{n} \sum^n x_i^k$

Несмещенность:

$$M(\hat{\alpha}_k) = \frac{1}{n} \sum^n M(x_i^k) = \frac{1}{n} \sum^n M(\xi^k) = \frac{1}{n} \cdot n \cdot \alpha_k = \alpha_k,$$

где $\alpha_k = M(\xi^k)$

$\Rightarrow$ $\hat{\alpha}_k$ – несмещ. оценка α_k

Состоятельность:

$\forall k \quad D(\xi^k) = M(\xi^{2k}) - M^2(\xi^k)$

$$D(\hat{\alpha}_k) = \frac{1}{n^2} D\left(\sum^n x_i^k\right) = \frac{1}{n^2} \sum^n D(x_i^k) = \frac{1}{n^2} \cdot n \cdot D(\xi^k) = \frac{1}{n}\left(M(\xi^{2k}) - M^2(\xi^k)\right) =$$

$$= \frac{\alpha_{2k} - \alpha_k^2}{n} \longrightarrow 0, \quad n \to \infty$$

несмещенность $\Rightarrow$ состоятельность оценки $\hat{\alpha}_k$

⑦ Несмещенная и состоятельная оценка начальных моментов

Теор. выборочные начальные моменты явл. несмещ. и сост. оценками соответствующих начальных моментов ТС

▷ по выборке $x_1, \dots, x_n$ находим $\hat{\alpha}_k = \frac{1}{n} \sum^n x_i^k$

Несмещенность:

$$M(\hat{\alpha}_k) = \frac{1}{n} \sum^n M(x_i^k) = \frac{1}{n} \sum^n M(\xi^k) = \frac{1}{n} \cdot n \cdot \alpha_k = \alpha_k,$$

где $\alpha_k = M(\xi^k)$

$\Rightarrow$ $\hat{\alpha}_k$ - несмещ. оценка α_k

Состоятельность:

$\forall k \quad D(\xi^k) = M(\xi^{2k}) - M^2(\xi^k)$

$$D(\hat{\alpha}_k) = \frac{1}{n^2} D\left(\sum^n x_i^k\right) = \frac{1}{n^2} \sum^n D(x_i^k) = \frac{1}{n^2} \cdot n \cdot D(\xi^k) = \frac{1}{n}\left(M(\xi^{2k}) - M^2(\xi^k)\right) =$$

$$= \frac{\alpha_{2k} - \alpha_k^2}{n} \longrightarrow 0, \quad n \to \infty$$

несмещенность $\Rightarrow$ состоятельность оценки $\hat{\alpha}_k$

⑨ Теорема Слуцкого. Сходимость выборочных центр. моментов

Теор. (Слуцкого) если $f(x,y)$ – непрерывна в т. $(a;b)$, $x_n \xrightarrow{P} a$, $y_n \xrightarrow{P} b$,
то $f(x_n, y_n) \xrightarrow{P} f(a,b)$

▷ $f(x,y)$ непрерывна в т. (a,b) => $\forall \varepsilon > 0$ $\exists \delta = \delta(\varepsilon)$: $|f(x_n,y_n) - f(a,b)| < \varepsilon$,
если $|x_n - a| < \delta$, $|y_n - b| < \delta$

Тогда если $|f(x_n,y_n) - f(a,b)| \geq \varepsilon$, то верно хотя бы одно из $|x_n - a| \geq \delta$, $|y_n - b| \geq \delta$

Введем в рассмотрение события: $A = \{w : |f(x_n,y_n) - f(a,b)| \geq \varepsilon\}$
$B = \{w : |x_n - a| \geq \delta\}$
$C = \{w : |y_n - b| \geq \delta\}$

$A = B + C$ => $p(A) \leq p(B) + p(C)$
$p(|f(x_n,y_n) - f(a,b)| \geq \varepsilon) \leq p(|x_n - a| \geq \delta) + p(|y_n - b| \geq \delta)$

$x_n \xrightarrow{P} a$ => $\lim\limits_{n\to\infty} p(|x_n - a| \geq \delta) = 0$

$y_n \xrightarrow{P} b$ => $\lim\limits_{n\to\infty} p(|y_n - b| \geq \delta) = 0$

$p(|f(x_n,y_n) - f(a,b)| \geq \varepsilon) = 0$

$p(|f(x_n,y_n) - f(a,b)| < \varepsilon) = 1$ => $f(x_n,y_n) \xrightarrow{P} f(a,b)$

Зам.: по т. Слуцкого, если $f(x_1,\dots,x_n)$ непр., $x_1,\dots,x_n \xrightarrow{P} a_1,\dots,a_n$ соотв.
то $f(x_1,\dots,x_n) \xrightarrow{P} f(a_1,\dots,a_n)$

Зам.: поскольку в. центр. моменты явл. рациональными ф-ями наблюдений, то они сход. по вер-ти к соотв. центр. моментам ГС (т.е. явл. состоятельными оценками)

(10) Оценка мат. ожидания по неравноточным наблюдениям

Равноточные наблюдения — наблюдения с одинак. дисперсией

Пусть $M(x_i) = M(\xi) = a$ и наблюдения неравноточные $D(x_i) = \sigma_i^2$

Класс лин. несмещенных оценок пар-ра a — класс оценок $\hat{a}_n = c_1 x_1 + \ldots + c_n x_n$,

причём $M(\hat{a}_n) = a$; $c_1, \ldots, c_n$ — веса набл.

По св-вам мат. ож. и несмещ. оценок:

$$M(\hat{a}_n) = c_1(\underbrace{M(x_1)}_{=a}) + \ldots + c_n(\underbrace{M(x_n)}_{a}) = a \cdot \sum^n c_i = a \Rightarrow \sum^n c_i = 1$$

$$D(\hat{a}_n) = c_1^2 \sigma_1^2 + \ldots + c_n^2 \sigma_n^2$$

для эффективной оценки $D(\hat{a}_n) \to \min \Rightarrow \begin{cases} \min\ c_1^2 \sigma_1^2 + \ldots + c_n^2 \sigma_n^2 = f(c_1, \ldots, c_n) \\ \sum c_i = 1 \end{cases}$

$$\begin{cases} f = \sum c_i^2 \sigma_i^2 \\ g = \sum c_i - 1 = 0 \end{cases}$$

$$L(c_1, \ldots, c_n, \lambda) = \sum^n c_i^2 \sigma_i^2 - \lambda\left(\sum^n c_i - 1\right)$$

$$\begin{cases} \dfrac{\partial L}{\partial c_1} = 2c_1 \sigma_1^2 - \lambda = 0 \\ \ldots \\ \dfrac{\partial L}{\partial c_n} = 2c_n \cdot \sigma_n^2 - \lambda = 0 \\ \dfrac{\partial L}{\partial \lambda} = 1 - \sum^n c_i = 0 \end{cases} \Rightarrow \frac{\lambda}{2} \cdot \sum^n \frac{1}{\sigma_i^2} = 1$$

$$c_i = \frac{1}{\sigma_i^2} \cdot \frac{1}{\sum^n \frac{1}{\sigma_i^2}}$$

— min (по кр-ую втор пр.)

$$\hat{a}_n = \frac{1}{\sigma_1^2} \cdot \frac{1}{\sum^n \frac{1}{\sigma_i^2}} \cdot x_1 + \ldots + \frac{1}{\sigma_n^2} \cdot \frac{1}{\sum^n \frac{1}{\sigma_i^2}} \cdot x_n$$

(чем больше дисп. набл., тем меньше вес) причём дает эффект. оценку мат. ож.

$$D(\hat{a}_n) = \frac{1}{\left(\sum^n \frac{1}{\sigma_i^2}\right)^2} \left[\frac{1}{\sigma_1^4} \cdot \sigma_1^2 + \ldots + \frac{1}{\sigma_n^4} \cdot \sigma_n^2 \right] = \frac{1}{\left(\sum \frac{1}{\sigma_i^2}\right)^2} \cdot \sum^n \frac{1}{\sigma_i^2} = \frac{1}{\sum^n \frac{1}{\sigma_i^2}}$$

Состоятельность:

пусть $\sigma_i^2 \le \sigma_0^2$ $i = \overline{1, n}$

то $\dfrac{1}{\sigma_i^2} \geqslant \dfrac{1}{\sigma_0^2} \Rightarrow \sum^n \dfrac{1}{\sigma_i^2} \geqslant n \cdot \dfrac{1}{\sigma_0^2} \Rightarrow D(\hat{a}_n) = \dfrac{1}{\sum^n \frac{1}{\sigma_i^2}} \le \dfrac{\sigma_0^2}{n} \to 0,\ n \to \infty$

Зам: при $D(x_i) = \sigma^2$, $i = \overline{1, n}$

$$c_1 = \frac{1}{\sigma^2} \cdot \frac{1}{\sum^n \frac{1}{\sigma^2}} = \frac{1}{n} \Rightarrow \hat{a}_n = \frac{1}{n} \sum^n x_i = \bar{x}$$ — выб. среднее

(11) Асимптотически нормальный характер выборочных х-к

Асимпт. нормальная послед. случ. вел. $\xi_1, \dots, \xi_n$ — если $\exists$ чис. посл. $A_1, \dots, A_n$; $B_1, \dots, B_n$, $B_i > 0$

что $P\left(\frac{\xi_i - A_i}{\sqrt{B_i}} < x\right) \to \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-t^2/2} dt$, $\xi_i \xrightarrow{\epsilon} N(A_i, B_i)$

Зам. если $\xi_1, \dots, \xi_n$ имеют одинак. распр. и независимы; $M(\xi_i) = a$, $D(\xi_i) = \delta^2$

то $\xi = \sum^{n} \xi_i$ явл. асимп. норм. распр., причем $\xi \xrightarrow{\epsilon} N(a \cdot n; \delta^2 \cdot n)$

Теор. $M(\xi) = a$, $D(\xi) = \delta^2$ — конечны => основные выб. хар-ки асимпт. норм.

1) $\bar{x} \xrightarrow{\epsilon} N\left(a, \frac{\delta^2}{n}\right)$

2) $\hat{\delta}^2 \xrightarrow{\epsilon} N\left(\frac{n-1}{n} \delta^2, \frac{2(n-1)}{n^2} \cdot \delta^4\right)$

3) $S^2 \xrightarrow{\epsilon} N\left(\delta^2, \frac{2}{n-1} \cdot \delta^4\right)$

4) $F_n(x) \xrightarrow{\epsilon} N\left(F(x), \frac{1-F(x)}{n} \cdot F(x)\right)$

▷ по центральной предельной теореме

(12) Эффективность оценок по Рао-Фреше-Крамеру

Эффективная оценка – оценка с минимальной дисперсией

Информация Фишера о неизв. параметре θ, сод. в одном из нез. набл. сл. вел. – величина $J(\theta) = M\left[\left(\frac{\partial \ln f(x,\theta)}{\partial\theta}\right)^2\right]$, где $f(x,\theta)$ – значение плотности распр. сл. вел. в т. x (непр.); вер. $P(x,\theta)$, что сл. вел. примет знач. x (дискр.)

<u>Теор.</u>: 1) $G_n = \{x : f(x,\theta) > 0\}$ – не зав. от θ

2) $\int_{-\infty}^{\infty} x \cdot f(x,\theta)\,dx = M(\xi)$

$\int_{-\infty}^{\infty} f(x,\theta)\,dx = 1$; $f(x,\theta)$ диф-ма по θ

3) $J(\theta) > 0$, конечна

то $\forall$ несмещ. $\hat{\theta}_n$ справедливо $D(\hat{\theta}_n) \geqslant \frac{1}{n \cdot J(\theta)}$

$\Delta_n = \frac{1}{n \cdot J(\theta)}$ – нижн. граница дисп. несм. оценок

Эффект-ть несмещ. $\hat{\theta}_n$ по РФК – отношение мин. возм. значения дисперсии в классе всех несмещ. оценок и дисперсии рассм. оценки:

$$e(\hat{\theta}_n) = \frac{\Delta_n}{D(\hat{\theta}_n)} = \frac{1}{n \cdot J(\theta) \cdot D(\hat{\theta}_n)} \in [0; 1]$$

$\hat{\theta}_n$ эффективна по РФК, если $e(\hat{\theta}_n) = 1$

(13) Эффективная оценка мат. ожидания норм. распр.

Теор. $\bar{x}$ выб. среднее – эффективная оценка мат. ожидания норм. распр.

▷ $M(\bar{x}) = a$ – несмещ. $\qquad \bar{x} = \frac{1}{n} \cdot \sum^{n} x_i$

$D(\bar{x}) = \frac{\delta^2}{n}$ – (известн.)

$$f(x, a) = \frac{1}{\delta\sqrt{2\pi}} \cdot e^{-\frac{(x-a)^2}{2\delta^2}}$$

$$\ln f = -\ln\delta - \ln\sqrt{2\pi} - \frac{(x-a)^2}{2\delta^2}$$

$$\frac{\partial \ln f}{\partial a} = \frac{x-a}{\delta^2}$$

$$J(a) = M\left[\left(\frac{\partial \ln f}{\partial a}\right)^2\right] = M\left[\left(\frac{x-a}{\delta^2}\right)^2\right] = \frac{1}{\delta^4} M\left[(x-a)^2\right] = \frac{1}{\delta^4} \cdot D(\xi) = \frac{1}{\delta^4} \cdot \delta^2 = \frac{1}{\delta^2}$$

$$e(\bar{x}) = \frac{1}{n \cdot J(a) \cdot D(\bar{x})} = \frac{1}{n \cdot \frac{1}{\delta^2} \cdot \frac{\delta^2}{n}} = 1$$

$\Rightarrow \bar{x}$ – эффективная оценка мат. ожид. a

(14) Эффективная оценка неизв. вер-ти в схеме Бернулли

Теор. Относительная частота "успеха" явл. эффект. оц. н. вер-ти успеха в схеме нез. исп. Бернулли

▷ $p = \theta$

$$\hat{\theta}_n = \frac{m}{n} = \frac{1}{n}\sum^{n} x_i \text{ , где } x_i = \begin{cases} 1, & \text{если соб. наст.} \\ 0, & \text{если соб. не наст.} \end{cases}$$

$$M(\hat{\theta}_n) = \frac{1}{n}\sum M(x_i) = \frac{1}{n}\cdot n\cdot M(\xi) = \frac{1}{n}\cdot n\cdot p = p = \theta$$ => оценка несмещ.

$$D(\hat{\theta}_n) = \frac{1}{n^2}\sum D(x_i) = \frac{1}{n^2}\cdot n\cdot D(\xi) = \frac{1}{n}\cdot p\cdot q = \frac{1}{n}\cdot\theta(1-\theta)$$

x_i	0	1
	$p(0,x) = 1-\theta$	$p(1,x) = \theta$

$$I(\theta) = \left(\frac{d\ln p(1,\theta)}{d\theta}\right)^2\cdot p(1,\theta) + \left(\frac{d\ln p(0,\theta)}{d\theta}\right)^2\cdot p(0,\theta) =$$

$$= \frac{1}{\theta^2}\cdot\theta + \left(-\frac{1}{1-\theta}\right)^2\cdot(1-\theta) = \frac{1}{\theta} + \frac{1}{1-\theta} = \frac{1}{\theta(1-\theta)}$$

$$e(\hat{\theta}_n) = \frac{1}{n\cdot\frac{1}{\theta(1-\theta)}\cdot\frac{1}{n}\cdot\theta(1-\theta)} = 1$$ => эффективность

15 Асимптотически эффективные оценки

Асимптот. эффективность $e_0(\hat{\theta}_n)$ несмещ. оценки $\hat{\theta}_n$ неизв. пар-ра θ – величина $e_0(\hat{\theta}_n) = \lim\limits_{n\to\infty} e(\hat{\theta}_n)$

$\hat{\theta}_n$ асимпт. эф-на, если $e_0(\hat{\theta}_n) = \lim\limits_{n\to\infty} e(\hat{\theta}_n) = 1$

Теор. Исправленная выборочная дисп. S_n^2 явл. асимп. эф-ной оценкой неизв. дисперсии σ^2 для норм. распр. ГС

$\triangleright$ $M(S_n^2) = \sigma^2$ => S_n^2 – несмещенная оценка σ^2

$D(S^2) = \frac{2\sigma^4}{n-1}$

$f(x,a,\sigma^2) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-a)^2}{2\sigma^2}}$

$\ln f = -\frac{1}{2}\ln\sigma^2 - \ln\sqrt{2\pi} - \frac{(x-a)^2}{2\sigma^2}$

$\frac{\partial \ln f}{\partial \sigma^2} = -\frac{1}{2\sigma^2} + \frac{(x-a)^2}{2\sigma^4} = \frac{(x-a)^2 - \sigma^2}{2\sigma^4}$

$\mathcal{I}(\theta) = M\left[\left(\frac{(x-a)^2-\sigma^2}{2\sigma^4}\right)^2\right] = \frac{1}{4\sigma^8} M\left[\left((x-a)^2-\sigma^2\right)^2\right] = \frac{1}{4\sigma^8}\cdot D\left[(x-a)^2\right] = \frac{1}{2\sigma^4}$

$D\left[(x-a)^2\right] = M\left[(x-a)^4\right] - M^2\left[(x-a)^2\right] = M\left[(x-a)^4\right] - (\sigma^2)^2 = 3\sigma^4 - \sigma^4 = 2\sigma^4$

$M\left[(x-a)^4\right] = \int_{-\infty}^{\infty} (x-a)^4 \frac{1}{\sigma\sqrt{2\pi}} \cdot e^{-\frac{(x-a)^2}{2\sigma^2}} dx = \left\{ \begin{array}{l} t = \frac{x-a}{\sigma} \\ dx = \sigma dt \end{array} \right\} = \frac{\sigma^4}{\sqrt{2\pi}} \int_{-\infty}^{\infty} t^4 \cdot e^{-t^2/2} dt = \ldots$

$= \frac{3\sigma^4}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-t^2/2} dt = \frac{3\sigma^4}{\sqrt{2\pi}} \cdot \sqrt{2\pi} = 3\sigma^4$

$e(\hat{\theta}_n) = e(S^2) = \frac{1}{n\cdot\frac{1}{2\sigma^4}\cdot\frac{2\sigma^4}{n-1}} = \frac{n-1}{n}$ – не явл. эффективной

$e_0(\hat{\theta}_n) = \lim\limits_{n\to\infty} e(S^2) = \lim\limits_{n\to\infty} \frac{n-1}{n} = 1$ – явл. асимп. эффективной

(16) Метод моментов

Метод моментов осн. на том, что выб. нач. и центр. моменты являются состоятельными оценками

суть: приравниваем выб. момент и соотв. моменту ГС, причём число уравнений равно числу оцениваемых пар-ров, из системы получаем оценки неизв. пар-ров согласно мет. мом.

Достоинства: ◦ простая вычислительная реализация
◦ состоятельность оценок (сх-ть по вер-ти выб. мом. к теор.)

Зам. • следует брать моменты более низких порядков (не всегда возможно)
• если полученные ур-я не реш. в элемент. ф-ях, то находится числ. решение
• можно оценивать ф-и, зависящие от пар-ра (оценка эксцесса)
• применяя различные моменты можно получить разные оценки (неопред.)
• оценки часто смещенные

Пример: по выборке $x_1, \dots, x_n$ из ГС имеющей равномерное распр. на $[a; b]$ оценить мет. моментов пар-ры a, b

приравняем нач. моменты 1ого порядка $\hat{\alpha}_1 = \alpha_1$
центр. —//— 2ого порядка $\hat{\mu}_2 = \mu_2$

$$\begin{cases} M(\xi) = \dfrac{a+b}{2} \\ D(\xi) = \dfrac{(b-a)^2}{2} \end{cases}$$

$$\begin{cases} \hat{\alpha}_1 = \bar{x} \\ \hat{\mu}_2 = S^2 \end{cases} \Rightarrow \begin{cases} \bar{x} = \dfrac{a+b}{2} \\ S^2 = \dfrac{(b-a)^2}{2} \end{cases} \Rightarrow \begin{cases} 2\bar{x} = a+b \\ 2S\sqrt{3} = b-a \end{cases} \Rightarrow \begin{cases} \hat{a} = \bar{x} - S\sqrt{3} \\ \hat{b} = \bar{x} + S\sqrt{3} \end{cases}$$

(17) Метод максимального правдоподобия

Метод макс. правдоподобия осн. на введении ф-и правдоподобия, которая представляет собой вероятность совместного появления в выборке значений $x_1, \dots, x_n$ (дискр.) или плотность совместного появления в выборке знач. $x_1, \dots, x_n$ (непр.)

$$L(x_1, \dots, x_n, \theta) = P(x_1, \dots, x_n, \theta) = P(x_1, \theta) \cdot \ldots \cdot P(x_n, \theta) \quad \text{дискр. сл. вел.}$$

или

$$L(x_1, \dots, x_n, \theta) = f(x_1, \theta) \cdot \ldots \cdot f(x_n, \theta) \quad \text{непр. сл. вел.}$$

Суть: 1) $L(x_1, \dots, x_n, \theta) \longrightarrow max$ – отыскание таких оценок пар-ра, кот. макс. ф-ю правдоподобия

2) $\ln L(\theta)$ – логарифмируем ф-ю правдоподобия, т.к. L и $\ln L$ достигают макс. при одинак. знач. аргумента

3) $\frac{\partial \ln L}{\partial \theta} = 0$ – записываем уравнение правдоподобия (решаем и обосновываем max)

Зам. ◦ если у ф-и неск. extr, то выбирается наибольший

◦ если распр. зав. от k пр-ров, то на extr иссл-ся ф-я k переменных

◦ при выполнении усл. регулярности оценки, она является асимпт. несмещ., асимпт. норм. распр., асимпт. эффективной

$$\hat{\theta}_n \rightleftarrows N\left(\theta; \frac{1}{n\, I(\theta)}\right)$$

◦ если $\exists$ эффект. по РФК оценка пар-ра θ, то она будет оц. макс. пр-я

если оц. макс. пр-я не явл. эффективной, то др. эффект. оценок нет

◦ метод хорош при больших объемах выборки

Достоинства: ◦ оценки получ. асимпт. несм., состоятельными, ас. норм., ас. эф.

◦ дост. полно используется инф. о π

17.1. Дискретное распределение

Пусть ГС имеет дискр. распр., закон распр. известен, но неизв. пар-р. Тогда по $x_1, \dots, x_n$ выборке сост. ф-я пр-я L и исслед. на макс.

Пример: мет. макс. пр-я оценить вер-ть успеха в схеме Бернулли, если в n_1 нез. исп. успех наст. m_1 раз, в n_2 - m_2

$$P_n(m) = C_n^m p^m q^{n-m}, \quad \theta = p$$

$$L = P_{n_1}(m_1) \cdot P_{n_2}(m_2) = C_{n_1}^{m_1} C_{n_2}^{m_2} p^{m_1+m_2} (1-p)^{n_1+n_2-m_1-m_2}$$

$$\ln L = \ln C_{n_1}^{m_1} C_{n_2}^{m_2} + (m_1+m_2)\ln p + (n_1+n_2-m_1-m_2)\ln(1-p)$$

$$\frac{\partial \ln L}{\partial p} = \frac{m_1+m_2}{p} - \frac{n_1+n_2-m_1-m_2}{1-p} = 0 \qquad \frac{\partial^2 \ln L}{\partial p^2} < 0, \ \forall p \qquad \Rightarrow p - \max$$

откуда $\hat{p} = \dfrac{m_1+m_2}{n_1+n_2}$ - оценка неизв. пар-ра по мет. макс. пр-я

17.2. Непрерывное распределение

Пусть ГС имеет непрер. распр., кот. опред-ся пар-ром θ.

Тогда $L(x_1, \dots, x_n, \theta) = f(x_1, \theta) \cdot \ldots \cdot f(x_n, \theta)$

<u>Пример</u>: ГС имеет показательное распр., оценить пар-р λ мет. макс. пр-я

$$f(x, \lambda) = \lambda \cdot e^{-\lambda x}, \quad x \geqslant 0, \qquad \theta = \lambda$$

$$L(x_1, \dots, x_n, \lambda) = \lambda \cdot e^{-\lambda x_1} \cdot \ldots \cdot \lambda e^{-\lambda x_n} = \lambda^n \cdot e^{-\lambda \cdot \sum x_i}$$

$$\ln L(\lambda) = n \cdot \ln \lambda - \lambda \cdot \sum x_i$$

$$\frac{\partial \ln L}{\partial \lambda} = \frac{n}{\lambda} - \sum x_i = 0 \quad \Rightarrow \quad \hat{\lambda} = \frac{1}{\bar{x}}$$ – оценка пар-ра λ

$$\frac{\partial^2 \ln L}{\partial \lambda^2} = -\frac{n}{\lambda^2} < 0 \quad \Rightarrow \quad \max$$

17.3. Нерегулярный случай

Общая схема метода не работает в граничных случаях или когда ф-я недиф-ма

В таких случаях исследуется сама ф-я L

<u>Пример</u>: ген. сов. имеет равн. распр. на $[0, \theta]$, мет. макс. пр-я оценить θ

$$f_{\xi}(x) = \begin{cases} 1/\theta, & x \in [0, \theta] \\ 0, & x \notin [0, \theta] \end{cases}$$

$$L = f(x_1, \theta) \cdot \ldots \cdot f(x_n, \theta) = \left(\frac{1}{\theta}\right)^n$$

$$\ln L = -n \ln \theta$$

$$\frac{\partial \ln L}{\partial \theta} = -\frac{n}{\theta} = 0$$ => алгоритм не работает, нерегулярный случай

(не вып-ся 1ое усл. регулярности: область, где плотность распр. > 0 зав. от пар-ра)

Исследуем L непосредственно на максимум

$\theta \geqslant x_{max}$

L убывает по θ => $L \to max$ при наим. θ => $\hat{\theta} = x_{max}$

причем оценка $\hat{\theta}$ - смещенная:

$$F_{x_{max}} = \frac{x^n}{\theta^n}$$

$$f_{x_{max}} = n\frac{x^{n-1}}{\theta^n}$$

$$M(x_{max}) = \int_0^{\theta} x \cdot n \cdot \frac{x^{n-1}}{\theta^n}\, dx = \frac{n}{\theta^n}\int_0^{\theta} x^n\, dx = \frac{n}{\theta^n} \cdot \frac{x^{n+1}}{n+1}\Big|_0^{\theta} = \frac{n\theta}{n+1}$$

$\hat{\theta}_n = \frac{n+1}{n}\theta$ - несмещ. оценка

(18) Метод наименьших квадратов

Метод наим. квадр. осн. на мин. суммы квадр. откл. выч. знач. от наблюдаемых

Применяется для получения оценок, если вид зав-ти известен

Пусть неи. показатель описывается законом $y = \varphi(x, \theta_1, \dots, \theta_k)$, θ_i - пар-ры

В результате набл. получены значения $y_i = \varphi(x_i, \theta_1, \dots, \theta_k) + \varepsilon_i$, $i = \overline{1,n}$,

где ε_i - случайные ошибки

Требуется по набл. (x_i, y_i) оценить пар-ры $\theta_1, \dots, \theta_k$

Можно считать, что распр. явл. нормальным $\varepsilon_i \in N(0, \delta^2)$

в силу зак. о больш. числах

Тогда y_i будут иметь норм. распр. как лин. преобразование ε_i

$y_i \in N(\varphi, \delta^2)$

Составим ф-ю правдоподобия для проведенных наблюдений:

$$L = \left(\frac{1}{\delta\sqrt{2\pi}}\right)^n \cdot e^{-\frac{1}{2\delta^2} \cdot \sum (y_i - \varphi(x_i, \theta_1, \dots, \theta_k))^2} \longrightarrow max$$

причем $L \to max$, когда $T(\theta_1, \dots, \theta_k) = \sum (y_i - \varphi)^2 \longrightarrow min$

далее по обычной схеме находим мин. $\frac{\partial T}{\partial \theta_i}$

<u>Пример</u>: пусть зависимость $y = ax + b$ - линейная

по результатам наблюдений имеем $(x_1, y_1), \dots, (x_n, y_n)$, т.е. $y_i = ax_i + \varepsilon_i + b$,

где ε_i - случ. ош.

$$T(a,b) = \sum (ax_i + b - y_i)^2 \longrightarrow min$$

$$\begin{cases} \frac{\partial T}{\partial a} = (\sum a x_i^2 + b \quad x_i - \sum x_i y_i) \cdot 2 = 0 \\ \frac{\partial T}{\partial b} = (a \sum x_i + nb - \sum y_i) \cdot 2 = 0 \end{cases} \Rightarrow \begin{cases} a \cdot \sum x_i^2 + b \cdot \sum x_i = \sum x_i y_i \\ a \cdot \sum x_i + b \cdot n = \sum y_i \end{cases}$$

$\Delta = n \sum x_i^2 - (\sum x_i)^2 = n^2(\overline{x^2} - (\overline{x})^2) = n^2 \hat{\delta}^2 \neq 0$ (т.к. $\delta^2 = 0$, когда вел. пост.)

$\Rightarrow$ система имеет ед. решение; можно уб., что реш. явл. мин.; $y = \hat{a}x + \hat{b}$ - ур. прям. регрес.

Теор.: пусть зависимость имеет вид $y = \varphi(x, \theta_1, \dots, \theta_n)$

наблюдения $y_i = \varphi(x_i, \theta_1, \dots, \theta_n) + \varepsilon_i$, $i = \overline{1,n}$

1) $M(\varepsilon_i) = 0$, $\delta^2(\varepsilon_i) = \delta^2$, не зав. от i; θ_i

2) ε_i не коррелированы

3) $\varphi \in C$, $\in D$ по параметрам θ_i

тогда оценки наим. мет. наим. квадратов

состоятельные, ас. несмещ., ас. норм., ас. эффективные

▷ след. из того, что оценка такая же, как и в мет. макс. прав-я

Достоинство: не требует знания зак. распр. сл. ошибки и самой сл. вел.

(19) Основные распределения в статистике

1) Распределение χ^2 с n степенями свободы – распределение суммы квадратов n нез. сл. вел., каждая из кот. имеет станд. норм. распр.

$$\chi_n^2 = \sum^n \xi_i^2 , \quad \xi_i \in N(0,1) , \quad i=\overline{1,n}$$

Если $\xi_i \in N(a_i, \sigma_i^2)$, $i=\overline{1,n}$ н. с. в. имеют норм. распр.

то $\dfrac{\xi_i - a_i}{\sigma_i} \in N(0;1)$, $i=\overline{1,n}$ имеют ст. норм. распр.

$$\chi_n^2 = \sum^n \frac{(\xi_i - a_i)^2}{\sigma_i^2}$$ имеет распр. χ_n^2

$$f(x) = \begin{cases} \dfrac{1}{2^{n/2}\,\Gamma(\frac{n}{2})} \cdot x^{n/2-1} \cdot e^{-x/2} , & x>0 \\ 0 , & x \le 0 \end{cases} , \quad \Gamma(\alpha) = \int_0^\infty x^{\alpha-1} e^{-x} dx$$

$M(\chi_n^2) = n$

$D(\chi_n^2) = 2n$

Св-во: χ_n^2, χ_m^2 – н. с. в с распр. хи-квадрат

то $(\chi_n^2 + \chi_m^2)$ имеет распр. хи-квадрат с $(n+m)$ ст. свободы

2) Распределение Стьюдента – распр. сл. вел. (с n степ. свободы)

$$t_n = \frac{\xi_0}{\sqrt{\frac{1}{n} \cdot \sum^n \xi_i^2}}, \quad \xi_i \in N(0;1), \ \xi_0, \dots, \xi_n \text{ – н.с.в.}, \ i = \overline{0,n}$$

$$f(x) = \frac{\Gamma(\frac{n+1}{2})}{\Gamma(\frac{n}{2}) \cdot \sqrt{n\pi}} \left(1 + \frac{x^2}{n}\right)^{-\frac{n+1}{2}}, \quad x \in (-\infty; +\infty)$$

$M(t_n) = 0$

$$D(t_n) = \frac{n}{n-2}, \quad n > 2$$

Распр. Стьюдента при $n > 30$ сход. к норм. распр.

3) Распределение Фишера – Снедекора $F(n,m)$ с n и m. ст. свободы – распределение $F(n,m) = \frac{\chi^2_n}{n} : \frac{\chi^2_m}{m}$, где χ^2_n, χ^2_m – нез. сл. в.

$$f(x) = \frac{\Gamma(\frac{n+m}{2})}{\Gamma(\frac{n}{2}) \cdot \Gamma(\frac{m}{2})} \cdot \frac{n^{n/2} \cdot m^{m/2} \cdot x^{n/2 - 1}}{(nx+m)^{n+m/2}}, \quad x > 0$$

$$M(F(n,m)) = \frac{m}{m-2}, \quad m > 2$$

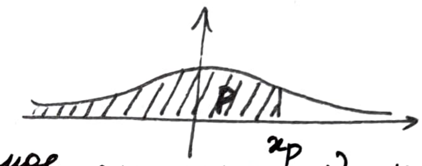
$$D(F(n,m)) = \frac{2m^2(n+m-2)}{n \cdot (m-2)^2 (m-4)}, \quad m > 4$$

(20) Квантили, процентные и критические точки

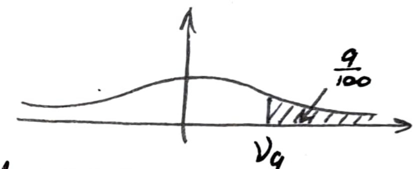
Квантиль уровня p н.с.в. ξ с ф-я распр. $F_\xi(x)$ – такое значение x_p этой сл. вел., что $p(\xi < x_p) = p$, где $p \in (0;1)$,

т.е. x_p – решение ур-я $F_\xi(x_p) = p$

геометрически – это такое значение сл. величины x_p : $\int_{-\infty}^{x_p} f(x)dx = p$



Процентная точка уровня $q \in [0:100]$ – такое значение ν_q н.с.в. ξ, что $p(\xi \geq \nu_q) = \frac{q}{100}$

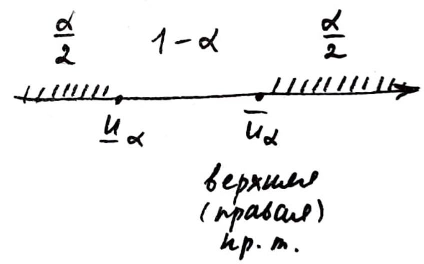


Нижняя (левая) критическая точка $\underline{u}_\alpha$, соотв. уровню значимости α – такие знач. сл. в. ξ, что вып. след. условия:

$p(\xi < \underline{u}_\alpha) = \frac{\alpha}{2}$

$p(\xi \geq \overline{u}_\alpha) = \frac{\alpha}{2}$

$p(\underline{u}_\alpha \leq \xi < \overline{u}_\alpha) = 1 - \alpha$



α – уровень значимости – дост. малая вер-ть (обычно зад-ся заранее)

Крит. точки опр. границы изм. значений сл. вел. ξ (приближенно)

(21) Доверительный интервал. Точность и надёжность оценки

Точечные оценки – оценки пар-ра, определяемые одним числом

Интервальные оценки – оценки пар-ра, опр. двумя чис. – границами инт.

Доверительная вероятность (надежность оценки) $\hat{\theta}_n$ неизв. пар-ра θ – вероятность γ: $p(|\hat{\theta}_n - \theta| < \varepsilon) = \gamma$ (надежность γ опр. заранее, $\to 1$)

- ясно, что чем меньше ε, тем точнее оценка (ε – точность инт. оценки)
- $p(|\hat{\theta}_n - \theta| < \varepsilon) = p(\hat{\theta}_n - \varepsilon < \theta < \hat{\theta}_n + \varepsilon) = \gamma$

Доверительный интервал, покрывающий неизв. пар-р θ с надежностью γ – интервал $(\hat{\theta}_n - \varepsilon; \hat{\theta}_n + \varepsilon)$, причем границы инт. не зав. от пар-ра θ

- пар-р θ – пост. величина, а границы инт. – случ. вел.

Построение д.и. начинается с отыскания "хорошей" точ. оценки (несмещ.)

Методы построения: 1) классический:

по выб. $x_1, \dots, x_n$ находятся ф-и $\hat{\theta}_1 = \theta_1(x_1, \dots, x_n)$

$\hat{\theta}_2 = \theta_2(x_1, \dots, x_n)$

$p(\hat{\theta}_1 < \theta < \hat{\theta}_2) = \gamma$, $\hat{\theta}_1 < \hat{\theta}_2$, $\forall x_1, \dots, x_n$

2) осн. на асимпт. норм-ти оценок.

(21) Доверительный интервал. Точность и надёжность оценки

Точечные оценки – оценки пар-ра, определяемые одним числом

Интервальные оценки – оценки пар-ра, опр. двумя чис. – границами инт.

Доверительная вероятность (надежность оценки) $\hat{\theta}_n$ неизв. пар-ра θ – вероятность γ: $p(|\hat{\theta}_n - \theta| < \varepsilon) = \gamma$ (надежность γ опр. заранее, $\to 1$)

- ясно, что чем меньше ε, тем точнее оценка (ε – точность инт. оценки)
- $p(|\hat{\theta}_n - \theta| < \varepsilon) = p(\hat{\theta}_n - \varepsilon < \theta < \hat{\theta}_n + \varepsilon) = \gamma$

Доверительный интервал, покрывающий неизв. пар-р θ с надежностью γ – интервал $(\hat{\theta}_n - \varepsilon; \hat{\theta}_n + \varepsilon)$, причем границы инт. не зав. от пар-ра θ

- пар-р θ – пост. величина, а границы инт. – случ. вел.

Построение д.и. начинается с отыскания "хорошей" точ. оценки (несмещ.)

Методы построения: 1) классический:

по выб. $x_1, \dots, x_n$ находятся ф-и $\hat{\theta}_1 = \theta_1(x_1, \dots, x_n)$
$\hat{\theta}_2 = \theta_2(x_1, \dots, x_n)$

$p(\hat{\theta}_1 < \theta < \hat{\theta}_2) = \gamma$, $\hat{\theta}_1 < \hat{\theta}_2$, $\forall x_1, \dots, x_n$

2) осн. на асимпт. норм-ти оценок.

(22) Доверительный интервал для мат. ожидания

1) <u>при известной дисперсии</u>:

Пусть $\xi \in N(a, \sigma^2)$, требуется с зад. над. γ построить д.и. для a:

произведем выборку $x_1, \dots, x_n$; $\bar{x}$ – точечная оценка a

$p(|\bar{x} - a| < \varepsilon) = \gamma$, отсюда $A: \bar{x} - \varepsilon < a < \bar{x} + \varepsilon$
$B: a - \varepsilon < \bar{x} < a + \varepsilon$ – равносильные соб. $P(A) = P(B)$

Учитывая, что $\bar{x} \in N(a, \sigma^2/n)$ найдем вер-ть соб. B:

$$p(a - \varepsilon < \bar{x} < a + \varepsilon) = \Phi\left(\frac{(a + \varepsilon - a)\sqrt{n}}{\sigma}\right) - \Phi\left(\frac{(a - \varepsilon - a)\sqrt{n}}{\sigma}\right) = 2\Phi\left(\frac{\varepsilon\sqrt{n}}{\sigma}\right) = \gamma$$

$$\Rightarrow \Phi\left(\frac{\varepsilon\sqrt{n}}{\sigma}\right) = \frac{\gamma}{2}$$

по табл. знач. ф-и Лапласа $\left(= \frac{\gamma}{2}\right)$ находим знач. арг. $U_\gamma = \frac{\varepsilon\sqrt{n}}{\sigma}$

отсюда $\varepsilon = \frac{\sigma}{\sqrt{n}} \cdot U_\gamma$

$$p\left(\bar{x} - \frac{\sigma}{\sqrt{n}} \cdot U_\gamma < a < \bar{x} + \frac{\sigma}{\sqrt{n}} \cdot U_\gamma\right) = \gamma$$

Д.и. $(\bar{x} - \varepsilon < a < \bar{x} + \varepsilon)$ покр. неизв. пар-р a, причем $\varepsilon = \frac{\sigma}{\sqrt{n}} U_\gamma$, U_γ – арг. ф. Лапл. при кот. ф-ие знач-те $\gamma/2$

<u>Зам.</u>: с ростом n точность оценки растет $(\varepsilon\downarrow)$

с ростом γ точность оценки умен. $(\varepsilon\uparrow)$

Если ди. строится с заранее зад. надежностью и точностью, то мин. объем выборки $n \geq \frac{\sigma^2 \cdot U_\gamma^2}{\varepsilon^2}$

2) при неизвестной дисперсии:

$\xi \in N(a, \sigma^2)$; построить д.и. для a

$\bar{x} = \frac{1}{n}\sum x_i$; $S^2 = \frac{1}{n-1}\sum^n (x_i - \bar{x})^2$

по т. Фишера

$$\bar{x} \in N(a, \frac{\sigma^2}{n}) \Rightarrow \frac{\bar{x}-a}{\sigma}\sqrt{n} \in N(0;1)$$

$$\frac{n\hat{\sigma}^2}{\sigma^2} \in \chi^2_{n-1}; \quad \frac{(n-1)S^2}{\sigma^2} \in \chi^2_{n-1}$$

$$\frac{n}{\sigma^2} \cdot \frac{1}{n} \cdot \sum^n (x_i - x)^2 = \frac{n-1}{\sigma^2} \cdot \frac{1}{n-1} \sum^n (x_i - \bar{x})^2$$

учитывая это рассмотрим сл. величину:

$$t_{n-1} = \frac{\bar{x}-a}{\sigma}\sqrt{n} : \sqrt{\frac{1}{n-1} \cdot \frac{(n-1)S^2}{\sigma^2}}$$ $\in$ распр. Стьюдента с $n-1$ ст. своб.

$$t_{n-1} = \frac{\bar{x}-a}{S}\sqrt{n}$$

Введем в рассмотр. крит. точки распр. Стьюдента $p(|t_n| < t_\gamma) = \gamma = 1-\alpha$

$$p\left(\left|\frac{\bar{x}-a}{S}\sqrt{n}\right| < t_\gamma\right) = \gamma = 1-\alpha$$

$$\bar{x} - \frac{S}{\sqrt{n}} \cdot t_\gamma < a < \bar{x} + \frac{S}{\sqrt{n}} \cdot t_\gamma$$ – искомый д.и.

где t_γ – крит. т. распр. Стьюдента с $n-1$ ст. своб для двусторонней крит. обл-ти с ур. зн. $\alpha = 1-\gamma$

Зам.: для получ. инт. оценок для истин. знач крит. случ. вел. можно применять результаты пред. п.

(23) Доверительный интервал для дисперсии случ. величины

1) при известном мат. ожидании:

Пусть $\xi \in N(a, \delta^2)$; построим д.и. для δ^2 с зад. надежностью γ

$$S^2 = \frac{1}{n}\sum^{n}(x_i - a)^2$$ – несмещ. оценка δ^2

$$\frac{n S^2}{\delta^2} = \sum^{n}\left(\frac{x_i - a}{\delta}\right)^2 = \chi^2_n$$, т.к. $\frac{x_i - a}{\delta} \in N(0,1)$

Введем в рассмотрение крит. т. распр. χ^2_n

$$p(\chi^2_n < \chi^2_{\text{левая крит. т.}}) = \frac{\alpha}{2}$$

$$p(\chi^2_n > \chi^2_{\text{правая крит. т.}}) = \frac{\alpha}{2}$$

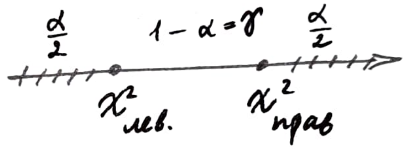


таблица крит. т. распр χ^2_n содержит только правые крит. т.

для левых: $p(\chi^2_n < \chi^2_{\text{лев.}}) + p(\chi^2_n > \chi^2_{\text{лев.}}) = 1$

$$p(\chi^2_n > \chi^2_{\text{лев.}}) = 1 - p(\chi^2_n < \chi^2_{\text{лев.}}) = 1 - \frac{\alpha}{2}$$

$\Rightarrow$ левая кр. т. ищется при ур. зн-ти $1 - \frac{\alpha}{2}$

$$p\left(\chi^2_{1-\frac{\alpha}{2}, n} < \frac{n S^2}{\delta^2} < \chi^2_{\frac{\alpha}{2}, n}\right) = \gamma = 1 - \alpha$$

отсюда д.и. с надежностью $\gamma = 1 - \alpha$, где $\chi^2_{\frac{\alpha}{n}, n}, \chi^2_{1-\frac{\alpha}{n}, n}$ – крит. т. распр. χ^2_n с ур. зн-ти $\frac{\alpha}{n}$; $1 - \frac{\alpha}{n}$.

$$\frac{n S^2}{\chi^2_{\frac{\alpha}{2}, n}} < \delta^2 < \frac{n S^2}{\chi^2_{1-\frac{\alpha}{2}, n}}$$

2) при неизвестной дисперсии:

$\xi \in N(a, \sigma^2)$; построить д.и. для σ^2 с зад. надежностью $\gamma = 1 - \alpha$

$\bar{x} = \frac{1}{n}\sum^n x_i$; $S^2 = \frac{1}{n-1}\sum^n (x_i - \bar{x})^2$

$$\frac{(n-1)S^2}{\sigma^2} \in \chi^2_{n-1} \quad \text{по т. Фишера}$$

Аналогично пред. случаю

$\chi^2_{n-1,\text{прав.}} = \chi^2_{\frac{\alpha}{2}, n-1}$ $\qquad p(\chi^2_{n-1} < \chi^2_{\text{лев.}}) = \frac{\alpha}{2}$

$\chi^2_{n-1,\text{лев.}} = \chi^2_{1-\frac{\alpha}{2}, n-1}$ $\qquad p(\chi^2_{n-1} > \chi^2_{\text{пр.}}) = \frac{\alpha}{2}$

Отсюда искомый д.и. с над. $\gamma = 1 - \alpha$

$$\frac{(n-1)S^2}{\chi^2_{\frac{\alpha}{2}, n-1}} < \sigma^2 < \frac{(n-1)S^2}{\chi^2_{1-\frac{\alpha}{2}, n-1}}$$

Зам.: отл. от пред. случая только числом степ. свободы

2) при неизвестной дисперсии:

$\xi \in N(a, \sigma^2)$; построить д.и. для σ^2 с зад. надежностью $\gamma = 1 - \alpha$

$\bar{x} = \frac{1}{n}\sum^n x_i$; $S^2 = \frac{1}{n-1}\sum^n (x_i - \bar{x})^2$

$$\frac{(n-1)S^2}{\sigma^2} \in \chi^2_{n-1} \quad \text{по т. Фишера}$$

Аналогично пред. случаю

$\chi^2_{n-1,\,прав.} = \chi^2_{\frac{\alpha}{2},\,n-1}$ $\qquad p(\chi^2_{n-1} < \chi^2_{лев.}) = \frac{\alpha}{2}$

$\chi^2_{n-1,\,лев.} = \chi^2_{1-\frac{\alpha}{2},\,n-1}$ $\qquad p(\chi^2_{n-1} > \chi^2_{пр.}) = \frac{\alpha}{2}$

Отсюда искомый д.и. с над. $\gamma = 1 - \alpha$

$$\frac{(n-1)S^2}{\chi^2_{\frac{\alpha}{2},\,n-1}} < \sigma^2 < \frac{(n-1)S^2}{\chi^2_{1-\frac{\alpha}{2},\,n-1}}$$

Зам.: отл. от пред. случая только числом степ. свободы

(24) Асимптотические доверительные интервалы

Асимпт. дов. инт. строятся для пар-ров, им. распр., отл. от норм.

Асимпт. дов. инт. $(\hat{\theta}_1, \hat{\theta}_2)$ при оц. пар-ре θ — интервал $(\theta_1; \theta_2)$:

$$p(\hat{\theta}_1 < \theta < \hat{\theta}_2) \xrightarrow[n \to \infty]{} \gamma$$

Методы построения ас. инт.:

1. В ф-лу дисперсии вместо пар-ра подставляется его оценка:

$$\hat{\theta}_n - \frac{\sigma(\hat{\theta}_n)}{\sqrt{n}} u_\gamma < \theta < \hat{\theta}_n + \frac{\sigma(\hat{\theta}_n)}{\sqrt{n}} \cdot u_\gamma ,$$

где u_γ - знач. арг. ф-и Лапласа при ур. знач. $\gamma/2$

$$\Phi(u_\gamma) = \frac{\gamma}{2}$$

2. Метод функционального преобразования

   1) сначала находится ф-я $g(\theta) = \int \frac{d\theta}{\sigma(\theta)}$

   2) записывается интервал для $g(\theta)$:

$$g(\hat{\theta}_n) - \frac{1}{\sqrt{n}} u_\gamma < g(\theta) < g(\hat{\theta}_n) + \frac{1}{\sqrt{n}} u_\gamma$$

   3) используя обратное преобразование строится ас. д.и. пар-ра θ

(25) Асимптотический д.и. в схеме Бернулли

Пусть для n независ. исп. вер. 'успеха' $- p = \Theta$

$\hat{\Theta}_n = W = \frac{1}{n} \cdot \sum^n x_i$; $x_i = 1$ - успех; $x_i = 0$ - неуспех

$M(\hat{\Theta}_n) = p = \Theta$ => несмещенность

$D(\hat{\Theta}_n) = \frac{1}{n^2} \cdot n \cdot \Theta(1-\Theta) = \frac{\Theta(1-\Theta)}{n}$

Построим д.и. <u>1ым методом</u> (подстановки)

$\Theta = p$

$\hat{\Theta}_n = W$ => $W - \sqrt{\frac{W(1-W)}{n}} \cdot U_\gamma < p < W + \sqrt{\frac{W(1-W)}{n}} \cdot U_\gamma$ - д.и.

-//- <u>2ым методом</u> (функц. преобр-е)

$\frac{\delta^2(\Theta)}{n} = \frac{\Theta(1-\Theta)}{n}$ => $\delta(\Theta) = \sqrt{\Theta(1-\Theta)}$; $\quad \hat{\Theta} \xrightarrow{\in} N\left(\Theta; \frac{\delta^2(\Theta)}{n}\right)$

1) $g(\Theta) = \int \frac{d\Theta}{\delta(\Theta)} = \int \frac{d\Theta}{\sqrt{\Theta(1-\Theta)}} = \int \frac{d\Theta}{\sqrt{\Theta} \cdot \sqrt{1-\cancel{\Theta}^2}} = 2\arcsin\sqrt{\Theta}$

2) $2\arcsin\sqrt{W} - \frac{1}{\sqrt{n}} \cdot U_\gamma < 2\arcsin\sqrt{\Theta} < 2\arcsin\sqrt{W} + \frac{1}{\sqrt{n}} \cdot U_\gamma$

3) $\sin^2\left(\arcsin\sqrt{W} - \frac{1}{2\sqrt{n}} U_\gamma\right) < \underset{(=p)}{\Theta} < \sin^2\left(\arcsin\sqrt{W} + \frac{1}{2\sqrt{n}} \cdot U_\gamma\right)$

(26) Доверительный интервал для пар-ра распр-я Пуассона

Пусть г.с. распр. по закону Пуассона; построим ас. д.и. для λ с зад. надежностью γ

$M(\xi) = \lambda$
$D(\xi) = \lambda$

В кач-ве точечной оценки: $\bar{x} = \hat{\theta}_n$

$$M(\hat{\theta}_n) = M(\bar{x}) = \frac{1}{n} \cdot n \cdot M(\xi) = \frac{1}{n} \cdot n \cdot \theta = \theta = \lambda$$

$$D(\hat{\theta}_n) = \frac{1}{n^2} \cdot n \cdot \theta = \frac{\theta}{n} = \frac{\lambda}{n}$$

Построим д.и. методом функц. преобр-я: $\hat{\theta} \xrightarrow{\epsilon} N(\theta, \frac{\sigma^2(\theta)}{n})$

1) $g(\theta) = \int \frac{d\theta}{\sigma(\theta)} = \int \frac{d\,\theta}{\sqrt{\theta}} = 2\sqrt{\theta}$

2) $g(\hat{\theta}_n) - \frac{1}{\sqrt{n}} \cdot u_\gamma < g(\theta) < g(\hat{\theta}_n) + \frac{1}{\sqrt{n}} \cdot u_\gamma$

$$2\sqrt{\bar{x}} - \frac{u_\gamma}{\sqrt{n}} < 2\sqrt{\theta} < 2\sqrt{\bar{x}} + \frac{u_\gamma}{\sqrt{n}}$$

3) $\left(\sqrt{\bar{x}} - \frac{u_\gamma}{2\sqrt{n}}\right)^2 < \underset{(=\lambda)}{\theta} < \left(\sqrt{\bar{x}} + \frac{u_\gamma}{2\sqrt{n}}\right)^2$

(27) Доверительный интервал для пар-ра показательного распр.

Пусть $f_\xi(x) = \begin{cases} \lambda e^{-\lambda x} & , x \geq 0 \\ 0 & , x < 0 \end{cases}$; построим д.и. для λ с зад. над. γ

$M(\xi) = \frac{1}{\lambda}$

$D(\xi) = \frac{1}{\lambda^2}$

$\theta = \frac{1}{\lambda}$

$\hat{\theta}_n = \bar{x}$ – несмещ. оценка

$D(\hat{\theta}_n) = \frac{1}{n} \cdot \sum D(x_i) = \frac{1}{n^2} \cdot n \cdot \theta^2 = \frac{\theta^2}{n}$

Построим д.и. методом подстановки:

$$\bar{x} - \frac{\bar{x}}{\sqrt{n}} \cdot u_\gamma < \theta < \bar{x} + \frac{\bar{x}}{\sqrt{n}} \cdot u_\gamma \quad , \theta = \frac{1}{\lambda}$$

$$\frac{1}{\bar{x} + \frac{\bar{x}}{\sqrt{n}} \cdot u_\gamma} < \lambda < \frac{1}{\bar{x} - \frac{\bar{x}}{\sqrt{n}} \cdot u_\gamma}$$

Построим д.и. методом функц. пр-я:

1) $g(\theta) = \int \frac{d\theta}{\delta(\theta)} = \int \frac{d\theta}{\theta} = \ln \theta$

2) $\ln \bar{x} - \frac{1}{\sqrt{n}} \cdot u_\gamma < \ln \theta < \ln \bar{x} + \frac{1}{\sqrt{n}} \cdot u_\gamma$

3) $\bar{x} \cdot e^{-\frac{u_\gamma}{\sqrt{n}}} < \theta < \bar{x}\, e^{\frac{u_\gamma}{\sqrt{n}}}$

$$\frac{1}{\bar{x} e^{\frac{u_\gamma}{\sqrt{n}}}} < \lambda < \frac{1}{\bar{x} \cdot e^{-\frac{u_\gamma}{\sqrt{n}}}}$$

(28) Доверительный интервал для разности мат. ожиданий

Пусть $\xi_1 \in N(a_1, \sigma^2)$ причем a_1, a_2 – неизвестны

$\xi_2 \in N(a_2, \sigma^2)$ σ^2 – неизв., но равны

Построить дов. инт. для $(a_1 - a_2)$ при зад. над. $\gamma = 1 - \alpha$

Произв. нез. выборки $x_1, \dots, x_n$; $y_1, \dots, y_n$

Точечные оценки $\hat{a}_1 = \bar{x}$, $\hat{a}_2 = \bar{y}$

$$S_x^2 = \frac{1}{n-1} \cdot \sum^n (x_i - \bar{x})^2$$

$$S_y^2 = \frac{1}{n-1} \cdot \sum^n (x_i - \bar{y})^2$$

$M(\bar{x} - \bar{y}) = M(\bar{x}) - M(\bar{y}) = a_1 - a_2$

$D(\bar{x} - \bar{y}) = D(\bar{x}) + D(\bar{y}) = \frac{\sigma^2}{n} + \frac{\sigma^2}{m} = \frac{\sigma^2(m+n)}{mn} \Rightarrow \frac{(\bar{x} - \bar{y}) - (a_1 - a_2)}{\sigma} \cdot \sqrt{\frac{nm}{m+n}} \in N(0,1)$

по т. Фишера $\frac{(n-1) S_x^2}{\sigma^2} \in \chi^2_{n-1}$; $\frac{(m-1) S_y^2}{\sigma^2} \in \chi^2_{m-1}$

$$\Rightarrow \frac{(n-1) S_x^2}{\sigma^2} + \frac{(m-1) S_y^2}{\sigma^2} \in \chi^2_{n+m-2}$$

тогда t_n имеет распр. Стьюдента с $n+m-2$ ст. свободы

$$t_n = \frac{(\bar{x} - \bar{y}) - (a_1 - a_2)}{\sigma} \sqrt{\frac{mn}{m+n}} : \sqrt{\frac{1}{n+m-2} \cdot \frac{(n-1) S_x^2 + (m-1) S_y^2}{\sigma^2}}$$

при $\alpha = 1 - \gamma$ опр. крит. точки распр. Стьюдента с $m+n-2$ ст. свободы для двусторонней крит. обл-ти – $t_{\alpha, m+n-2}$

$$P(|t_n| < t_{\alpha, n+m-2}) = \gamma = 1 - \alpha$$

$$P\left(-t_{\alpha, n+m-2} < \frac{(\bar{x} - \bar{y}) - (a_1 - a_2)}{\sqrt{(n-1) S_x^2 + (m-1) S_y^2}} \sqrt{\frac{mn}{m+n}} \sqrt{n+m-2} < t_{\alpha, n+m-2}\right) = 1 - \alpha$$

$$(\bar{x} - \bar{y}) - \sqrt{\frac{(n-1) S_x^2 + (m-1) S_y^2}{n+m-2}} \cdot \sqrt{\frac{m+n}{mn}} \cdot t_{\alpha, m+n-2} < a_1 - a_2 < (\bar{x} - \bar{y}) + \sqrt{\frac{(n-1) S_x^2 + (m-1) S_y^2}{n+m-2}} \cdot \sqrt{\frac{m+n}{n}} t_{\alpha, m+n-2}$$

(29) Статистическая гипотеза

Статистическая гипотеза – любое предположение о неизв. законе распр. или н. параметре изв. закона распр.

Виды:
1) простые – одно предположение
2) сложные – > 1 предположений

1) параметрические – речь идет о пар-ре
2) непараметрические – остальные

1) основная (нулевая) – гипотеза H_0, выдв. для проверки
2) конкурирующая (альтерн.) – гипотеза H_1, противореч. нулевой

Ошибки:

H_0	верна	верна	неверна	неверна
принимаемое знач. по рез. проверки	прин.	отв.	отв.	прин.
		ошибка $1^{ого}$ рода (α)		ошибка $2^{ого}$ рода (β)

Ошибка $1^{ого}$ рода сост. в том, что H_0 верна, но она отвергается
– вероятность совершить ошибку $1^{ого}$ рода – наз-ся ур. знач-ти α

Ошибка $2^{ого}$ рода сост. в том, что принимается неверная гипотеза
– вероятность сов. ошибку $2^{ого}$ рода – β,
причем $(1-\beta)$ – мощность критерия, т.е. вер-ть отвергнуть неверную гипотезу

Задача – при зад. α построить наиб. мощ-ти критерий

(30) Общая схема проверки стат. гипотезы

Статистический критерий K – сл. вел., подобранная для проверки стат. гип., точное или прибл. распр. кот. известно

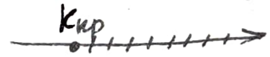
Наблюдаемое значение критерия $K_{набл.}$ – знач., наход. использ. данные выборки

- **Допустимая область** (область принятия нул. гип.) – часть обл-ти всех знач. K, при попадании в кот. гип. H_0 не отвергается (осн. отвергнуть H_0 – нет)
- **Критическая область** – часть об-ти всех знач. K, при попад. в кот. гип. H_0 отвергается

Критические точки критерия K – точки, отд. крит. обл. от обл-ти доп. знач.

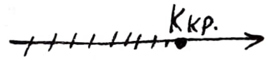
В зависимости от H_0 различают односторонние и двусторонние кр. обл-ти

1) правостор-ие $K > K_{крит.}$



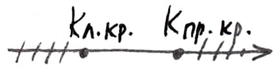
$P(K > K_{крит.}) = \alpha$

2) левостор-ие $K < K_{крит.}$



$P(K < K_{крит.}) = \alpha$

3) двустор-ие $\begin{cases} K > K_{пр.\,крит.} \\ K < K_{л.\,крит.} \end{cases}$



$(K_{л.кр.} < K_{пр.кр})$

$P(K<K_л) + P(K>K_п) = \alpha$

$P(K<K_л) = P(K>K_п) = \frac{\alpha}{2}$

Схема проверки стат. гипотезы (построения крит. наиб. мощн-ти при зад α)

1) Формулируется гипотеза H_0, зад. ур. зн-ти α
2) Формул-ся конкурирующая гипотеза H_1
3) Выбирается крит. проверки H_0
4) В зав-ти от вида H_1 исп. табл. крит. т., опр. крит. обл.
5) По данным выборки вычисл. $K_{набл.}$ знач. крит.
6) Если $K_{набл.}$ попадает в доп. область, то осн. отвергнуть гип. H_0 – нет, если $K_{набл.}$ попадает в крит. область, то H_0 отвергается

(31) Критерий отношения правдоподобия

Критерий наибольшей мощности – критерий отношения правдоподобия

Отношение правдоподобия $\frac{L_1}{L_0}$ – отношение, в пот. ξ – непр. сл. вел., пот. при спр-ти H_0 имеет распр. $f_0(x)$,

а при спр-ти H_1 имеет распр. $f_1(x)$, по пот. стр-ся ф-и правд-я L_0, L_1

Теорема (Неймана–Пирсона): из всех крит., пров. простую гипотезу H_0 при зад. ур. зн-ти α, наиб. мощностью облад. крит. отн. правдоподобия

Откуда след., что $\exists C$ – const : крит. обл-ть наиболее мощного крит. имеет вид $S = \{x : \frac{L_1}{L_0} > c\}$, где постоянная c находится из условия $P_{H_0}(\frac{L_1}{L_0} > C) = \alpha$

Пример: Пусть $\xi \in N(a, \sigma^2)$, σ^2 – изв., a – неизв.

на ур-не зн-ти α проверить г-зу $H_0 : a = a_0$

$H_1 : a = a_1 > a_0$

$H_0: f_0(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-a_0)^2}{2\sigma^2}}$ $\qquad L_0 = \left(\frac{1}{\sigma\sqrt{2\pi}}\right)^n e^{-\frac{1}{2\sigma^2}\sum^n (x_i - a_0)^2}$

$H_1: f_1(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-a_1)^2}{2\sigma^2}}$ $\qquad L_1 = \left(\frac{1}{\sigma\sqrt{2\pi}}\right)^n e^{-\frac{1}{2\sigma^2}\sum^n (x_i - a_1)^2}$

$$\frac{L_1}{L_0} = e^{-\frac{1}{2\sigma^2}\sum^n (x_i-a_1)^2 + \frac{1}{2\sigma^2}\sum^n (x_i - a_0)^2} = \ldots = e^{-\frac{1}{2\sigma^2}(a_1 - a_0)(2\bar{x} - a_1 - a_0)n}$$

по т. Н-П $\exists C(\alpha)$: $S = \{x : \frac{L_1}{L_0} > c\}$, где c опр-ся из $P_{H_0}(\frac{L_1}{L_0} > c) = \alpha$

$\frac{L_1}{L_0} > c \iff \bar{x} > c_1 \Rightarrow S = \{x : \bar{x} > c_1\}$, где $P_{H_0}(\bar{x} > c_1) = \alpha$.

$P_{H_0}(c_1 < \bar{x} < +\infty) = \Phi(\infty) - \Phi\left(\frac{c_1 - a_0}{\sigma}\cdot\sqrt{n}\right) = \alpha \Rightarrow \Phi\left(\frac{c_1 - a_0}{\sigma}\sqrt{n}\right) = \frac{1}{2} - \alpha$

Тогда: если $\bar{x} > a_0 + \frac{\sigma}{\sqrt{n}}\cdot u_\alpha$, то H_0 отв.

если $\bar{x} < a_0 + \frac{\sigma}{\sqrt{n}}\cdot u_\alpha$, то осн. отв. H_0 нет

$\frac{c_1 - a_0}{\sigma}\sqrt{n} = u_\alpha \Rightarrow c_1 = \frac{\sigma}{\sqrt{n}}u_\alpha + a_0$

$P_{H_1}(\bar{x} < c_1) = \beta$ – вер. ош. 2 рода $\Rightarrow \Phi\left(\frac{c_1 - a_1}{\sigma}\sqrt{n}\right) + \Phi(\infty) = \beta \Rightarrow 1 - \beta = \frac{1}{2} + \Phi\left(\frac{a_1 - a_0}{\sigma}\sqrt{n} - u_\alpha\right)$ – мощн. кр.

(32) Гипотеза о неизвестном среднем

(1) при известной дисперсии

$\xi \in N(a, \sigma^2)$, σ^2 – изв.

по $x_1, \dots, x_n$ выборке на ур. знач. α $H_0: a = a_0$

$H_1: \left\{ \begin{array}{l} a > a_0 \\ a < a_0 \\ a \neq a_0 \end{array} \right.$

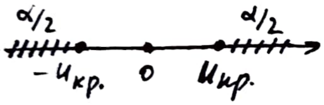
во всех случаях в кач-ве крит. пров. H_0 берем $U = \frac{\bar{x} - a_0}{\sigma}\sqrt{n} \in N(0,1)$ при спр-сти гип. H_0

$\bar{x} \in N(a, \frac{\sigma^2}{n})$

Кр. обл. стр-ся в зав-ти от вида H_1

1) при $H_1: a \neq a_0$ кр. обл. двусторонняя

наиболее мощн. критерий будет, если $u_{кр.л.}$, $u_{кр.пр.}$: $\underbrace{p(u < u_{кр.л.})}_{\frac{\alpha}{2}} + \underbrace{p(u > u_{кр.пр})}_{\frac{\alpha}{2}} = \alpha$



тогда $p(u > u_{кр}) + p(0 < u < u_{кр}) = \frac{1}{2}$

$p(u > u_{кр.}) + \Phi\left(\frac{u_{кр} - 0}{1}\right) - \underbrace{\Phi\left(\frac{0-0}{1}\right)}_{=0} = \frac{1}{2} \Rightarrow \Phi(u_{кр.}) = \frac{1}{2} - \frac{\alpha}{2}$

значит кр. т. нах-ся из $\Phi(u_{кр.}) = \frac{1}{2} - \frac{\alpha}{2}$

если $|u_{набл.}| > u_{кр}$, то H_0 – отв.

$|u_{набл.}| < u_{кр.}$, то нет осн. отв. H_0

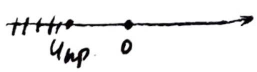
2) при $H_1: a > a_0$ кр. обл. правосторонняя

значит кр. т. нах-ся из $\Phi(u_{кр.}) = \frac{1}{2} - \alpha$

если $u_{набл.} > u_{кр}$, то H_0 – отв

$u_{набл.} < u_{кр.}$, то H_0 – не отв.

3) при $H_1: a < a_0$ кр. обл. левосторонняя



– // – если $u_{набл.} < -u_{кр}$, то H_0 – отв.

$u_{набл.} > -u_{кр}$, то H_0 – не отв.

② при неизв. дисперсии:

$\xi \in N(a, \sigma^2)$, σ^2 - неизв. ; при α

$H_0 : a = a_0$

$H_1 : \begin{cases} a > a_0 \\ a < a_0 \\ a \neq a_0 \end{cases}$

крит. проверки H_0

$t = \frac{\bar{x} - a_0}{s}\sqrt{n}$,

кот. при спр. H_0 имеет распр. Стьюдента с $n-1$ ст. св.

1) при $H_1 : a \neq a_0$ кр. обл. двусторонняя

кр. точка нах. по табл. кр. т. распр. Стьюдента с $(n-1)$ ст. свободы и α для двуст. кр. обл.

если $|t_{набл.}| > t_{кр.дв.}(\alpha, n-1) \Rightarrow H_0$ - отв.

$|t_{набл.}| < t_{кр.дв.}(\alpha, n-1) \Rightarrow H_0$ - не отв.

2) при $H_1 : a > a_0$ кр. обл. правосторонняя

кр. т. нах по -//- для одност. кр. обл.

если $t_{набл.} > t_{кр.одн.}(\alpha, n-1) \Rightarrow H_0$ - отв

$t_{набл.} < t_{кр.одн.}(\alpha, n-1) \Rightarrow H_0$ - не отв.

3) при $H_1 : a < a_0$ кр. обл. левосторонняя

-//-

если $t_{набл.} < -t_{кр.одн.} \Rightarrow H_0$ - отв.

$t_{набл.} > -t_{кр.одн.} \Rightarrow H_0$ - не отв.

(33) Гипотеза о неизв. дисперсии для одной выборки

Пусть $\xi \in N(a, \delta^2)$

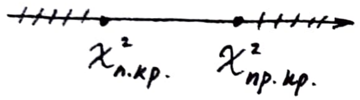
по $x_1, \dots, x_n$ при α проверить $H_0 : \delta^2 = \delta_0^2$

$$H_1 : \begin{cases} \delta^2 \neq \delta_0^2 \\ \delta^2 > \delta_0^2 \\ \delta^2 < \delta_0^2 \end{cases}$$

в кач-ве крит. проверки H_0 исп. случ. вел.

$$\chi^2_{n-1} = \frac{(n-1)S^2}{\delta^2}$$

1) при $H_1 : \delta^2 \neq \delta_0^2$ крит. обл. двусторонняя



крит. т. выбираются так: $\underbrace{P(\chi^2 < \chi^2_{л.кр.})}_{\alpha/2} + \underbrace{P(\chi^2 > \chi^2_{пр.кр.})}_{\alpha/2} = \alpha$

что мощность крит. наиб.

при построении д.и. для неизв. дисп. при неизв. a было получено:

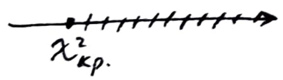
$$\chi^2_{пр.кр.} = \chi^2_{\frac{\alpha}{2}, n-1}$$

$$\chi^2_{л.кр.} = \chi^2_{1-\frac{\alpha}{2}, n-1}$$

если $\left\{ \begin{array}{l} \chi^2_{набл.} > \chi^2_{\alpha/2, n-1} \\ \chi^2_{набл.} < \chi^2_{1-\alpha/2, n-1} \end{array} \right.$ или $\Rightarrow H_0$ - отв.

$\chi^2_{1-\alpha/2, n-1} < \chi^2_{набл.} < \chi^2_{\alpha/2, n-1} \Rightarrow H_0$ - не отв.

2) при $H_1 : \delta^2 > \delta_0^2$ крит. обл. правосторонняя

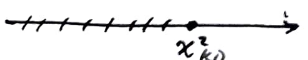


$\chi^2_{кр.}$ нах. по табл. кр. т. распр. χ^2_{n-1} и α : $\chi^2_{кр.} = \chi^2_{\alpha, n-1}$

если $\chi^2_{набл.} > \chi^2_{\alpha, n-1} \Rightarrow H_0$ - отв.

$\chi^2_{набл.} < \chi^2_{\alpha, n-1} \Rightarrow H_0$ - не отв.

3) при $H_1 : \delta^2 < \delta_0^2$ крит. обл. левосторонняя



-//- $\chi^2_{кр.} = \chi^2_{1-\alpha, n-1}$

если $\chi^2_{набл.} < \chi^2_{1-\alpha, n-1} \Rightarrow H_0$ - отв.

$\chi^2_{набл.} > \chi^2_{1-\alpha, n-1} \Rightarrow H_0$ - не отв.

(34) Гипотеза о неизв. вероятности успеха в схеме Бернулли (одна серия)

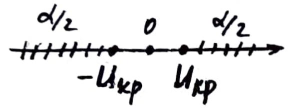
Проверим при α $\quad H_0: p = p_0$

$$H_1: \begin{cases} p \neq p_0 \\ p > p_0 \\ p < p_0 \end{cases}$$

гипотеза провер. на осн. ас. норм. част. события $\Rightarrow$ объем выборки долж. быть дост. больш.

$$w \xrightarrow{\varepsilon} N\left(p, \frac{p(1-p)}{n}\right), \quad u = \frac{w - p_0}{\sqrt{p_0(1-p_0)}} \cdot \sqrt{n} \xrightarrow{\varepsilon} N(0;1) \text{ при спр-ти } H_0$$

1) при $p \neq p_0$

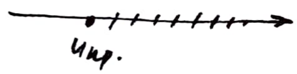


кр. т. $u_{кр.}$ нах-ся из рав-ва $\Phi(u_{кр.}) = 1/2 - \alpha/2$

если $|u_{набл.}| > u_{кр.} \Rightarrow H_0$ – отв.

$|u_{набл.}| < u_{кр.} \Rightarrow H_0$ – не отв.

2) при $p > p_0$

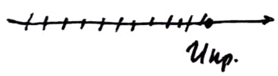


кр. т. $u_{кр.}$ –//– $\Phi(u_{кр}) = 1/2 - \alpha$

если $u_{набл.} > u_{кр.} \Rightarrow H_0$ – отв.

$u_{набл.} < u_{кр.} \Rightarrow H_0$ – не отв.

3) при $p < p_0$



–//–

если $u_{набл.} < -u_{кр.} \Rightarrow H_0$ – отв.

$u_{набл.} > -u_{кр.} \Rightarrow H_0$ – не отв.

(35) Гипотеза о значимости выборочного коэф. корреляции

Пусть (ξ_1, ξ_2) имеют норм. распр. (известно),

для характ. лин. зав-ти между ξ_1 и ξ_2 исп. коэф. корреляции

По выборке $(x_1, y_1), \dots, (x_n, y_n)$ найдем выб. коэф. корреляции

$$\rho_{в.} = \frac{\sum^{n} (x_i - \bar{x}) \cdot (y_i - \bar{y})}{\sqrt{\sum^{n} (x_i - \bar{x})^2 (y_i - \bar{y})^2}}$$

Пусть $\rho_{г} = 0$

Поскольку выборка произв-ся случайно нельзя утв., что $\rho_{\xi_1 \xi_2} = 0$

$H_0 : \rho_{\xi_1 \xi_2} = 0$

$H_1 : \rho_{\xi_1 \xi_2} \neq 0$

Если в рез. проверки H_0 не отв., то ξ_1, ξ_2 – некоррелированные => независимые

иначе, H_0 отв, т.е. ξ_1, ξ_2 – кор-ны => зависимые

Пусть при α, в кач-ве крит. проверки H_0 берется $t = \frac{\rho_{в.} \sqrt{n-2}}{\sqrt{1 - \rho_{в.}^2}}$,

кот. при спр-ти H_0 имеет распр. Стьюдента с $(n-2)$ ст.св.

Крит. обл. двусторонняя

$t_{кр.}$ находится по табл. кр. т. распр. Стьюдента для зад. α

если $|t_{табл.}| > t_{кр. дв. \alpha, n-2}$ => H_0 – отв.

$|t_{табл.}| < t_{кр. дв. \alpha, n-2}$ => H_0 – не отв.

(36) Гипотеза о равенстве средних для двух зав. выборок

Пусть $x_1, \dots, x_n$; $y_1, \dots, y_n$ — выборки до и после возд. на призн.

Проверим гипотезу о рав-ве средних до и после возд. на признак

$H_0: a_x = a_y$

$$H_1 \begin{cases} a_x \neq a_y \\ a_x > a_y \\ a_x < a_y \end{cases}$$

Введем $d_i = y_i - x_i$, $i = \overline{1,n}$

Тогда $H_0' : a_d = 0$

$$H_1' \begin{cases} a_d \neq 0 \\ a_d < 0 \\ a_d > 0 \end{cases}$$

критерий проверки H_0' : $t = \dfrac{\bar{d} - a_d}{S_d} \cdot \sqrt{n} = \dfrac{\bar{d}\sqrt{n}}{S_d}$

при спр-ти H_0 имеет распр. Стьюдента с $(n-1)$ ст. св.

1) при $H_1' : a_d \neq 0$ крит. обл. двусторонняя

если $|t_{набл.}| > t_{кр. дв., \alpha, n-1} \Rightarrow H_0$ — отв.

$|t_{набл.}| < t_{кр. дв., \alpha, n-1} \Rightarrow H_0$ — не отв.

2) при $H_1' : a_d < 0$ крит. обл. левосторонняя

если $t_{набл.} < -t_{кр. одн., \alpha, n-1} \Rightarrow H_0$ — отв.

3) при $H_1' : a_d > 0$ крит. обл. правосторонняя

если $t_{набл.} > t_{кр. одн., \alpha, n-1} \Rightarrow H_0$ — отв.

(37) Гипотеза о равенстве дисперсий двух нез. выборок

$\xi_1 \in N(a_1, \sigma_1^2)$; $\xi_2 \in N(a_2, \sigma_2^2)$

$x_1, \dots, x_n$; $y_1, \dots, y_m$ выборки: $S_x^2 = \frac{1}{n-1}\sum^{n}(x_i - \bar{x})^2$

$$S_y^2 = \frac{1}{m-1}\sum^{m}(y_i - \bar{y})^2$$

при α $H_0: \sigma_1^2 = \sigma_2^2$

$$H_1: \begin{cases} \sigma_1^2 \neq \sigma_2^2 \\ \sigma_1^2 > \sigma_2^2 \\ \sigma_1^2 < \sigma_2^2 \end{cases}$$

по т. Фишера S_x^2, S_y^2 – независимы $\frac{(n-1)S_x^2}{\sigma_1^2} \in \chi^2_{n-1}$; $\frac{(m-1)S_y^2}{\sigma_2^2} \in \chi^2_{m-1}$

если H_0 справедлива, то $\sigma_1^2 = \sigma_2^2 = \sigma^2$

$$F_{n-1,m-1} = \frac{(n-1)S_x^2}{\sigma^2(n-1)} : \frac{(m-1)S_y^2}{\sigma^2(m-1)} = \frac{S_x^2}{S_y^2}$$

имеет распр. Фишера с $n-1$, $m-1$ ст. свободы

в кач-ве крит. проверки H_0 берем отн. большей дисп. и меньшей

$$F = \frac{S^2_{б.}}{S^2_{м.}}$$

1) при $H_1: \sigma_1^2 \neq \sigma_2^2$ крит. обл. двусторонняя

$P(F < F_{1кр.}) = P(F > F_{2кр.}) = \alpha/2$

если выбрать $F_{2кр}$ так же, как $F_{1кр.}$ на ур. знач $\frac{\alpha}{2}$, то будет вып-ся и $P(F < F_{1кр.})$

=> если $F_{набл.} > F_{кр.}(\frac{\alpha}{2}, k_1, k_2)$ => H_0 – отв.

2) при $H_1: \sigma_1^2 > \sigma_2^2$ крит. обл. правосторонняя

$F_{кр.}$ нах-ся по табл. кр. т. распр. Фишера – Снедекора при α

$F_{кр.}(\alpha, k_1, k_2)$, где k_1 – число ст. св. большей дисп.; k_2 – // – меньш.

=> если $F_{набл.} > F_{кр.}(\alpha, k_1, k_2)$ => H_0 – отв.

3) при $H_1: \sigma_1^2 < \sigma_2^2$ сводится к $\sigma_1^2 > \sigma_2^2$ перестановкой σ_1^2 и σ_2^2

(38) Гипотеза о равенстве средних

1) при известных дисперсиях

$\xi_1 \in N(a_1, \delta_1^2)$, $\xi_2 \in N(a_2, \delta_1^2)$, δ_1^2, δ_2^2 – известные

при α $H_0: a_1 = a_2$

$$H_1: \begin{cases} a_1 \neq a_2 \\ a_1 > a_2 \\ a_1 < a_2 \end{cases}$$

выборки $x_1, \dots, x_n$; $y_1, \dots, y_n$: $\hat{a}_1 = \bar{x}$; $\hat{a}_2 = \bar{y}$

В кач-ве критерия проверки H_0 бер. случ. вел. $u = \dfrac{\bar{x} - \bar{y}}{\sqrt{\dfrac{\delta_1^2}{n} + \dfrac{\delta_2^2}{m}}}$

при спр-ти H_0 и нез-ти выборок $u \in N(0; 1)$

1) при $H_1: a_1 > a_2$

$u_{кр.}$ находится из рав-ва $\Phi(u_{кр.}) = \frac{1}{2} - \alpha$

если $u_{набл.} > u_{кр} \Rightarrow H_0$ – отв.

2) при $H_1: a_1 < a_2$

$u_{кр.}$ находится из рав-ва $\Phi(u_{кр.}) = \frac{1}{2} - \alpha$

если $u_{набл.} < -u_{кр.} \Rightarrow H_0$ – отв.

3) при $H_1: a_1 \neq a_2$ кр. обл. двусторонняя

$u_{кр.}$ находится из рав-ва $\Phi(u_{кр.}) = \frac{1}{2} - \frac{\alpha}{2}$

если $|u_{набл.}| > u_{кр.} \Rightarrow H_0$ – отв.

2) <u>при неизв. дисперсиях</u> (равных)

$\varsigma_1 \in N(a_1, \delta^2)$; $\varsigma_2 \in N(a_2, \delta^2)$

при выборках $x_1, \dots, x_n$; $y_1, \dots, y_n$

отсюда $M(\varsigma_1) = \bar{x}$; $M(\varsigma_2) = \bar{y}$; $D(\varsigma_1) = S_x^2$; $D(\varsigma_2) = S_y^2$

при α $\quad H_0: a_1 = a_2$

$$H_1: \begin{cases} a_1 \neq a_2 \\ a_1 > a_2 \\ a_1 < a_2 \end{cases}$$

В качестве крит. проверки H_0 бер. сл. вел. $t = \dfrac{\bar{x} - \bar{y}}{S\sqrt{\frac{1}{n} + \frac{1}{m}}}$,

где $S = \sqrt{\dfrac{(n-1)S_x^2 + (m-1)S_y^2}{n+m-2}}$ — несмещ. оценка неизв. дисп. δ^2

при спр-ти H_0 t имеет распр. Стьюдента с $(n+m-2)$ ст. св.

1) при $H_1: a_1 > a_2$

$t_{кр.} = t_{кр.\,одн.,\, n+m-2,\, \alpha}$

если $t_{набл.} > t_{кр.\,одн.,\, n+m-2,\, \alpha}$ $\Rightarrow H_0$ — отв

2) при $H_1: a_1 < a_2$

$t_{кр.} = t_{кр.\,одн.,\, n+m-2,\, \alpha}$

если $t_{набл.} < -t_{кр.\,одн.,\, n+m-2,\, \alpha}$ $\Rightarrow H_0$ — отв.

3) при $H_1: a_1 \neq a_2$

$t_{кр.} = t_{кр.\,дв.,\, n+m-2,\, \alpha}$

если $|t_{набл.}| > t_{кр.\,дв.,\, n+m-2,\, \alpha}$ $\Rightarrow H_0$ — отв

<u>Зам.</u>: если нет уверенности о рав-ве дисп., то прежде чем проверять гип. о рав-ве средних надо проверить гип. о рав-ве дисп. по кр. Фишера

(39) Гипотеза о рав-ве вероятностей успеха в схеме Бернулли (две серии исп.)

Две серии испытаний Бернулли $n_1 - m_1 \quad w_1 = m_1/n_1$

$n_2 - m_2 \quad w_2 = m_2/n_2$

при α $H_0: p_1 = p_2$

$$H_1: \begin{cases} p_1 \neq p_2 \\ p_1 > p_2 \\ p_1 < p_2 \end{cases}$$

в кач-ве крит. проверки H_0 берется сл. вел.

$$u = \frac{w_1 - w_2}{\sqrt{\hat{p}(1-\hat{p})\left(\frac{1}{n_1}+\frac{1}{n_2}\right)}}$$

, т.к. при спр-ти H_0: $u \in N(0;1)$

$$w_1 - w_2 \xrightarrow{\in} N \Leftarrow \begin{matrix} w_1 \xrightarrow{\in} N\left(p_1, \frac{p_1(1-p_1)}{n_1}\right) \\ w_2 \xrightarrow{\in} N\left(p_2, \frac{p_2(1-p_2)}{n_2}\right) \end{matrix}$$

причем $\frac{p_1(1-p_1)}{n_1} + \frac{p_2(1-p_2)}{n_2} = \frac{p(1-p)(n_1+n_2)}{n_1 n_2}$

$\hat{p}$ – оценка неизв. вер-ти: $\hat{p} = \frac{m_1+m_2}{n_1+n_2}$

1) при $H_1: p_1 > p_2$

$u_{кр.}$ нах. из рав-ва $\Phi(u_{кр.}) = \frac{1}{2} - \alpha$

если $u_{набл.} > u_{кр.} \Rightarrow H_0$ – отв.

2) при $H_1: p_1 < p_2$

– // –

если $u_{набл.} < -u_{кр.} \Rightarrow H_0$ – отв.

3) при $H_1: p_1 \neq p_2$

$u_{кр.}$ нах. из рав-ва $\Phi(u_{кр.}) = \frac{1}{2} - \frac{\alpha}{2}$

если $|u_{набл.}| > u_{кр.} \Rightarrow H_0$ – отв.

Зам: указ. мет. пров. на осн. ас. норм. отн. част. w_1, w_2

$\Rightarrow n_1, n_2$ должно быть дост. большими

(40) Критерий Бартлетта и Кохрена

$\xi_1, \dots, \xi_m \in N(a_m, \sigma_m^2)$

из них извл. выборки объемов $n_1, \dots, n_m$, по кот. найд. $S_1^2, \dots, S_m^2$

при α $H_0: \sigma_1^2 = \dots = \sigma_m^2$ (гип. об. однородности дисперсий)

$H_1: \; i,j: \; \sigma_i^2 \neq \sigma_j^2$

Для проверки H_0 можно использовать <u>критерий Бартлетта</u>

$k_i = n_i - 1, \; i = \overline{1,m}$ – число ст. свободы дисперсии S_i^2

$K = \sum k_i$

$S^2 = \frac{1}{K} \sum^m k_i S_i^2$ – ср. арифм. дисп., взвеш. по числ. ст. свободы

В кач-ве крит. исп-ся сл. величина $B = \frac{V}{C}$,

где $V = K \ln S^2 - \sum^m k_i \ln S_i^2$

$C = 1 + \frac{1}{3(m-1)} \left(\sum^m \frac{1}{k_i} - \frac{1}{K} \right)$

Бартлетт доказал, что при спр-ти H_0 $B \xrightarrow{\epsilon} \chi^2_{m-1}$,

если объем кажд. из выборок ≥ 4

кр. обл. правосторонняя

кр. т. $\chi^2_{кр.}$ нах. по табл. кр. точек χ^2_{m-1} при α

если $B_{набл.} > \chi^2_{кр.}(\alpha, m-1) \Rightarrow H_0$ отв.

<u>Зам.</u> если $V_{набл.} < \chi^2_{кр}$, то $B_{набл.} << \chi^2_{кр.}$, т.к. $C > 1$

<u>Зам.</u> если объемы выборок одинаковы, то исп. др. критерий (Кохрена)

$\xi_1, \dots, \xi_n \in N(a_m, \delta_m^2)$

из ген. извл. выборки объема n, по ген. нах-ся $S_1^2, \dots, S_m^2$

при α $\quad H_0: \delta_1^2 = \dots = \delta_m^2$

$H_1: \exists\, i,j: \delta_i^2 \neq \delta_j^2$

Гипотеза H_0 при одинак. объемах выборки пров. по крит. Кохрена

В кач-ве крит. исп. сл. вел. $G = \dfrac{S_{max}^2}{S_1^2 + \dots + S_m^2}$

распр. этой величины зависит от числа ст. свободы $k = n - 1$
и кол-ва выборок m

Крит. обл. правосторонняя

кр. т. нах. по табл. распр. Кохрена в з-ти от числа ст. св, размера выб.
и ур. знач. α

$G_{кр.} = G_{кр.}(k, m, \alpha)$

если $G_{набл.} > G_{кр.} \Rightarrow H_0$ - отв.

Зам. если H_0 - не отв., то в кач-ве оценки неизв. дисп.
можно взять ср. ар. испр. выбор. дисп.

(41) Дисперсионный анализ.

$\xi_1, \dots, \xi_k \in N(a_k, \delta^2)$ δ^2 – неизв., но равные

при α $H_0 : a_1 = \dots = a_k$

$H_1 : \exists\, i, j : a_i \neq a_j$

для проверки H_0 извл. выборки объемов $n_1, \dots, n_k$

гипотеза H_0 проверяется методом дисперсионного анализа, осн. на сравнении дисперсий (исп. для изуч. влияния какого-либо фактора, имеющего неск. уровней)

<u>Идея:</u> сравнение факт. и ост. дисперсии

$$S^2_{фак} = \frac{1}{k-1}\left((\bar{x}_{гр.1} - \bar{x})^2 \cdot n_1 + \dots + (\bar{x}_{гр.k} - \bar{x})^2 \cdot n_k\right)$$

– отр. возд-е факторов

$$S^2_{ост.} = \frac{1}{N-K}\left((x_{11} - \bar{x}_{гр.1})^2 + (x_{21} - \bar{x}_{гр.1})^2 + \dots + (x_{nk} - \bar{x}_{гр.k})^2\right)$$

– отр. возд-е случ. величин, где x_{ij} j – номер группы, i – номер наблюдения

$$N = \sum^{k} n_i$$

при спр-ти H_0 $S^2_{фак.}$, $S^2_{ост.}$ явл. несмещ. оц. дисп. ГС

Проверка H_0 о рав-ве мат. ожид. св-ся к проверке гип. о рав-ве $S^2_{фак.}$ и $S^2_{ост}$ по крит. Фишера

▷ $H_0 : a_1 = \dots = a_k$; если H_0 – не отв., то $S^2_{ф.}$ и $S^2_{ост.}$ явл. несм. оценками и по кр. Фишера $H_0' : S^2_{ф.} = S^2_{ост.}$ – не отв.

◁ H_0 – отв. => фактор оказ. существ. влияние и нек. гр. средние будут расти, тогда $F_{набл.} = S^2_{фак}/S^2_{ост.} > F_{кр.}$ => H_0' – отв.

Кр. обл. правосторонняя; $F_{кр.}$ нах. по табл. кр.т. расп. Фишера с $N-k, k-1$ ст.св. при α

если $F_{набл.} > F_{кр.}(N-k, k-1, \alpha)$ => H_0' – отв. => H_0 – отв.

<u>Зам.</u> если нет уверенности, что дисп. равны, сначала надо пров. гип. об их рав-ве

(42) Гипотеза о рав-ве вероятностей "успеха" в схеме Бернулли
(l серий)

Пусть $n_1 - m_1$; ... ; $n_l - m_l$ оценки $w_1 = m_1/n_1$, ..., $w_l = m_l/n_l$

при α $H_0 : p_1 = \ldots = p_l$

$H_1 : \exists i, j : p_i \neq p_j$

В кач-ве крит. проверки H_0 бер. сл. вел. $u = \frac{1}{\hat{p}(1-\hat{p})} \sum^{l} (w_i - \hat{p})^2 \cdot n_i$

кот. при спр-ти H_0 и дост. объеме выб. $u \xrightarrow{\epsilon} \chi^2_{l-1}$

крит. обл. правосторонняя

кр. т. $\chi^2_{кр.} = \chi^2_{кр.}(\alpha, l-1)$ нах. по табл.

если $\chi^2_{набл.} > \chi^2_{кр.} \Rightarrow H_0$ - отв.

причем $\hat{p} = \frac{m_1 + \ldots + m_l}{n_1 + \ldots + n_l}$ - оценка неизв. вер-ти

(43) Критерий согласия Пирсона

Критерий согласия — крит. проверки гипотез о виде неизв. зак. распр.
- простая гипотеза (о зак. распр. с изв. пар-рами)
- сложная гипотеза (неизв. зак. распр. с неизв. пар-рами)

Критерий согласия Пирсона (χ^2) — самый распространённый, поскольку применим для непр. и дискр. сл. вел.; простых и сложных гипотез

Суть: сравн. эмпир. частоты (получ. по выборке) и теор. частоты. (получ. из спр-ти нул. гип.)

Пусть из ГС по $x_1, \dots, x_n$ получ. стат. распр.

знач. пр	x_1	...	x_s
эмп. част.	n_1		n_s

$n = \sum^s n_i$

предположим, что найдены теор. частоты $n_i' = p_i \cdot n$

рассмотрим сл. вел. $\sum^s \frac{(n_i - n_i')^2}{n_i'}$

Независимо от зак. распр. ГС эта вел. имеет распр. χ^2_{s-r-1},
где s — кол-во групп (после объед.
r — число пар-ров, нот. оц. по выборке

по выборке нах. $\chi^2_{набл.}$, по табл. $\chi^2_{кр.}(\alpha, s-r-1)$

если $\chi^2_{набл.} > \chi^2_{кр.} \Rightarrow H_0$ — отв.

Схема:
1) $n \geq 50$ дост. больш. объем выборки
2) диапазон знач. призн. в выб. разб. на S инт. один. длины, причем в каждом не менее 5-8 значений призн. (иначе инт. объед-ся в один; если инт. неогр слева/справа, то крайние инт. увелич. до бескон-ти)
3) сост. статист. ряд и оценки пар-ров
4) исходя из H_0 опр. вер-ти попад. сл. вел. в получ. инт (непр.); прин. пол. зн. (дискр.)
5) прим. критерий, опис. выше

(44) Критерий согласия Колмогорова

Крит. согл. Колмогорова применяется для проверки гипотез о виде неизв. зак. распр. для непр. сл. вел. $H_0: F(x) = F_0(x)$

крит. Кол-ва осн. на том, что распр-е $\sup |\underset{\text{эмп.}}{F_n(x)} - \underset{\text{теор}}{F(x)}|$ одно и то же, незав. от закона распр.

Статистикой явл сл. вел $D_n = \sup|F_n(x) - F(x)|$

(причем при малых n есть табл. кр. т. $D_n = D(\alpha, n)$)

при $n > 20$ исп. крит. т. распр Колмогорова

(пот. явл. пред. для статистики)

$$\lambda = \sqrt{n} \cdot D_n$$

$P(\sqrt{n} D_n < x) \longrightarrow$ распр. К-ва $\lambda_{кр.} = \lambda(\alpha)$

В обоих случаях если набл. знач. больше крит.

$$\left.\begin{array}{l} D_{набл.} > D_{кр.} \\ \lambda_{набл.} > \lambda_{кр.} \end{array}\right. \Rightarrow H_0 - отв.$$

а практике D_n нах. по ф-ле

$$D_n = \max \left| F_0(x_i) - \underbrace{\frac{2i-1}{2n}}_{F_n(x)} \right| + \frac{1}{2n}$$

<u>Зам.</u> для применения кр. Кол-ва знач. призн. в выборке упор. по возр.

<u>Зам</u>. не исп., если неизв. пар-ры распр.

не исп., если значения призн. сгруппированы

(45) Критерий Колмогорова – Смирнова

Гип. об одн-сти выборок – гип. о том, что выборки пр-ся из одной ГС

$H_0: F_1(x) = F_2(x) \quad \forall x$

$H_1: \exists x: \quad F_1(x) \neq F_2(x)$, где $F(x)$ – ф-я распр. ГС

Критерий Колм-ва – Смирнова исп. идею крит Кол-ва
(о том, что распр. $\sup|F_n(x) - F(x)|$ одно и то же)
но теперь ср-ся эмп. ф-и двух выборок

=> статистикой яв. $\lambda' = \sqrt{\dfrac{n_1 \cdot n_2}{n_1 + n_2}} \cdot \max\left|F_{n_1}(x) - F_{n_2}(x)\right|$,

где n_1, n_2 – объемы выборок (≥ 50)

кр. т. $\lambda'_{кр.}$ нах. по табл. кр. т. распр. Колмогорова при α

если $\lambda'_{набл.} > \lambda'_{кр.}$ => H_0 – отв.

Зам. $F_{n_1}(x)$, $F_{n_2}(x)$: ф-и распр. (теор.) непрерывны

(46) Критерий χ^2

Гипотезы об одн-ти выборк — гипотезы о том, что выборки произв. из одной и той же ГС

H_0 : $F_1(x) = F_2(x)$ $\forall x$, где $F_i(x)$ - ф-и распр. i-ой ГС

H_1 : $\exists x$: $F_1(x) \neq F_2(x)$

Пусть $x_1, \ldots, x_{n_1}$; $y_1, \ldots, y_{n_2}$ (набл. независимы)

для проверки H_0 при α весь диапазон знач. призн. в этих выб. разб-ся на S групп и подсч. число знач. призн., попавш. в кажд. группу в кажд. выб.

$$\begin{array}{ll} n_1 & m_1, \ldots, m_s \\ n_2 & l_1, \ldots, l_s \end{array}$$

, где m_i, l_i - число знач. призн.

$n_1 = \sum m_i$; $n_2 = \sum l_i$

В кач-ве крит. пров. H_0 исп. сл. вел.

$$\chi^2_{s-1} = n_1 \cdot n_2 \cdot \sum^{s} \frac{\left(\frac{m_i}{n_1} - \frac{l_i}{n_2}\right)^2}{m_i + l_i}$$

при спр-ти H_0 и дост. бол. n_1, n_2 имеет распр. χ^2_{s-1}

крит. обл. правосторонняя

если $\chi^2_{\text{набл.}} > \chi^2_{\text{кр.}}(\alpha, s-1)$ $\Rightarrow$ H_0 - отв.

(47) Критерий Вилкоксона

Ранговые критерии одн-ти выб. — осн. на поредк. ном. (рангах) знач. из призн., получ. в рез. объед. рассм. выборок и расположения в объед. выборке знач. призн. по возр.

$H_0: F_1(x) = F_2(x) \quad \forall x$

$H_1 \begin{cases} F_1(x) \neq F_2(x) \\ F_1(x) < F_2(x) \Rightarrow P(\xi_1 < x) < P(\xi_2 < x) \Rightarrow P(\xi_1 \geq x) > P(\xi_2 \geq x) \Rightarrow \xi_1 > \xi_2 \\ F_1(x) > F_2(x) \Rightarrow \text{—//—} \Rightarrow \xi_1 < \xi_2 \end{cases}$

Крит. Вилкоксона прим-ся если ф-я распр. непрерывна

Случай 1 объемы выборок: $n_1 \leq n_2 \leq 25$

Для проверки H_0 две выборки объед. в одну и расп. по возр. И найдем набл. знач. крит. $W_{набл.}$ — сумма номеров знач. призн. для знач. из одной выборки (одинак. для $\forall H_1$)

① при $H_1: F_1(x) \neq F_2(x)$ при α из табл. кр.т. кр. Вилкоксона нах. кр. т.

$W_{нижн. кр.т.} = w\left(\frac{\alpha}{2}, n_1, n_2\right)$

$W_{верх. кр.т.} = (n_1 + n_2 + 1)n_1 - W_{нижн.кр.}$

если $W_{н.кр.} < W_{набл.} < W_{в.кр.} \Rightarrow H_0$ — не отв.

② при $H_1: F_1(x) < F_2(x)$ если $W_{набл.} > W_{в.кр} \Rightarrow H_0$ — отв.

③ при $H_1: F_1(x) > F_2(x)$ если $W_{набл.} < W_{н.кр.} \Rightarrow H_0$ — отв.

Случай 2 $n_1 > 25$ или $n_2 > 25$

при $H_1: F_1(x) \neq F_2(x)$ $\quad W_{н.кр.} = \left[\frac{(n_1+n_2+1)n_1 - 1}{2} - u_{кр} \cdot \sqrt{\frac{(n_1+n_2+1)n_1 n_2}{12}}\right]$ ← цел. часть

$u_{кр.}$ нах. из рав-ва $\Phi(u_{кр.}) = 1/2 - \alpha/2$, далее —//—

при $H_1: F_1(x) > (<) F_2(x)$ $\quad u_{кр.}$ нах. из рав-ва $\Phi(u_{кр.}) = 1/2 - \alpha$, далее —//—

(48) Критерий Вилкоксона – Манна – Уитни

$H_0: F_1(x) = F_2(x) \quad \forall x$

$H_1: F_1(x) \neq F_2(x) \quad \exists x$

В этом крит. исп-ся статистики

$$V = S_1 - \frac{1}{2} n_1 (n_1 + 1)$$
$$V = S_2 - \frac{1}{2} n_2 (n_2 + 1)$$

, где n_1, n_2 – объемы выборок

S_i – сумма рангов знач. i-ой выборки в общ. выборке

$$S = S_1 + S_2 = \frac{1}{2}(n_1 + n_2)(n_1 + n_2 + 1)$$

В кач-ве крит. пр-ки H_0 исп. сл. вел.

$$u = \frac{V - 1/2\, n_1 \cdot n_2}{\sqrt{\frac{1}{12}(n_1 + n_2) n_1 \cdot n_2}} \xrightarrow{\epsilon} N(0,1)$$

при спр-сти H_0 и дост. больших n_1, n_2

крит. т. $u_{кр.}$ нах. из рав-ва $\Phi(u_{кр.}) = \frac{1}{2} - \frac{\alpha}{2}$

если $|u_{набл.}| > u_{кр.} \Rightarrow H_0$ – отв.

Сверхэффективная оценка

Если дисперсия оценки пропорциональна $\frac{1}{n^{\alpha}}$, $\alpha > 1$, то оценка лучше, чем эффективная, т.е. сверхэффективная

например: $\hat{\theta}_n = \frac{n+1}{n} \cdot x_{max}$

$$D(\hat{\theta}_n) = M(\hat{\theta}_n^2) - M^2(\hat{\theta}_n) = \int_0^{\theta} \left(\frac{n+1}{n}\right)^2 \cdot x^2 \cdot \frac{n}{\theta^n} \cdot x^{n-1} dx - \theta^2 =$$

$$= \frac{(n+1)^2 x^{n+2}}{\theta^n \cdot n \cdot (n+2)} \Big|_0^{\theta} - \theta^2 = \frac{\theta^2}{n(n+2)}$$

Теорема Фишера

Теор.: пусть ГС имеет норм. распр. (a, σ^2), если $\hat{\alpha}_n$, $\hat{\sigma}_n^2$, то:
(выб. средн., выб. дисп.)

1) $\hat{\alpha}_n \in N(a, \frac{\sigma^2}{n})$

2) $\frac{n \hat{\sigma}_n^2}{\sigma^2} \in \chi^2_{n-1}$

3) $\hat{\alpha}_n$, $\hat{\sigma}_n^2$ – независимы

След.:
- $M(\hat{\sigma}_n^2) = \frac{n-1}{n} \sigma^2$
- $D\left(\frac{n\hat{\sigma}_n^2}{\sigma^2}\right) = 2(n-1) \Rightarrow D(\hat{\sigma}_n^2) = \frac{2n-1}{n^2} \sigma^4$
- $S^2 = \frac{n}{n-1} \hat{\sigma}_n^2 \Rightarrow D(S^2) = \frac{2}{n-1} \sigma^4$