Критерии проверки гипотез об однородности

Лемешко Борис Юрьевич

E-mail: Lemeshko@fpm.ami.nstu.ru
http://www.ami.nstu.ru/~headrd/
Тел. 346-06-00

Среди критериев однородности выделяют **3 группы критериев**:

– Критерии проверки гипотез об однородности законов распределения;

– Критерии проверки гипотез об однородности средних (о равенстве математических ожиданий);

– Критерии проверки гипотез об однородности дисперсий (о равенстве дисперсий).

Критерии однородности законов

Задача проверки гипотезы об однородности законов, соответствующих k выборкам, формулируется следующим образом. Имеется k выборок

$$x_{11}, x_{12}, ..., x_{1n_1}, \quad x_{21}, x_{22}, ..., x_{2n_2}, \quad \ldots, \quad x_{k1}, x_{k2}, ..., x_{kn_k},$$

где n_i – объём i -й выборки, $i = \overline{1,k}$. Проверяется гипотеза о том, что все выборки извлечены из одной и той же генеральной совокупности, т. е.

$$H_0: \ F_1(x) = F_2(x) = \ldots = F_k(x) = F(x)$$

при любом x . Конкурирующая гипотеза может иметь вид

$$H_1: \ F_i(x) \neq F_j(x)$$

для некоторых $i \neq j, \ i, j \leq k$.

Как правило, критерии проверки гипотезы об однородности законов непараметрические.

Как правило, на практике используется либо критерий **Смирнова (1939 г.)**, либо критерий **Лемана–Розенблатта (1951 г.)**.

В русскоязычной литературе практически не упоминается о применении двухвыборочного критерия **Андерсона–Дарлинга** (**Андерсона–Дарлинга–Петита, 1976 г.**) или, тем более, об использовании **многовыборочного варианта критерия Андерсона–Дарлинга (1987 г.)**.

Пожалуй, нельзя найти и примеров использования **критериев Жанга (1999 г.)**, которые являются развитием вышеупомянутых критериев Смирнова, Лемана–Розенблатта и Андерсона–Дарлинга.

Модификация критерия однородности Смирнова

$F(x)$ и $G(x)$ - непрерывны. Статистика критерия Смирнова

$$S_C = \sqrt{\frac{mn}{m+n}} D_{m,n},$$

$$D_{m,n} = \sup_x \left| G_m(x) - F_n(x) \right|.$$

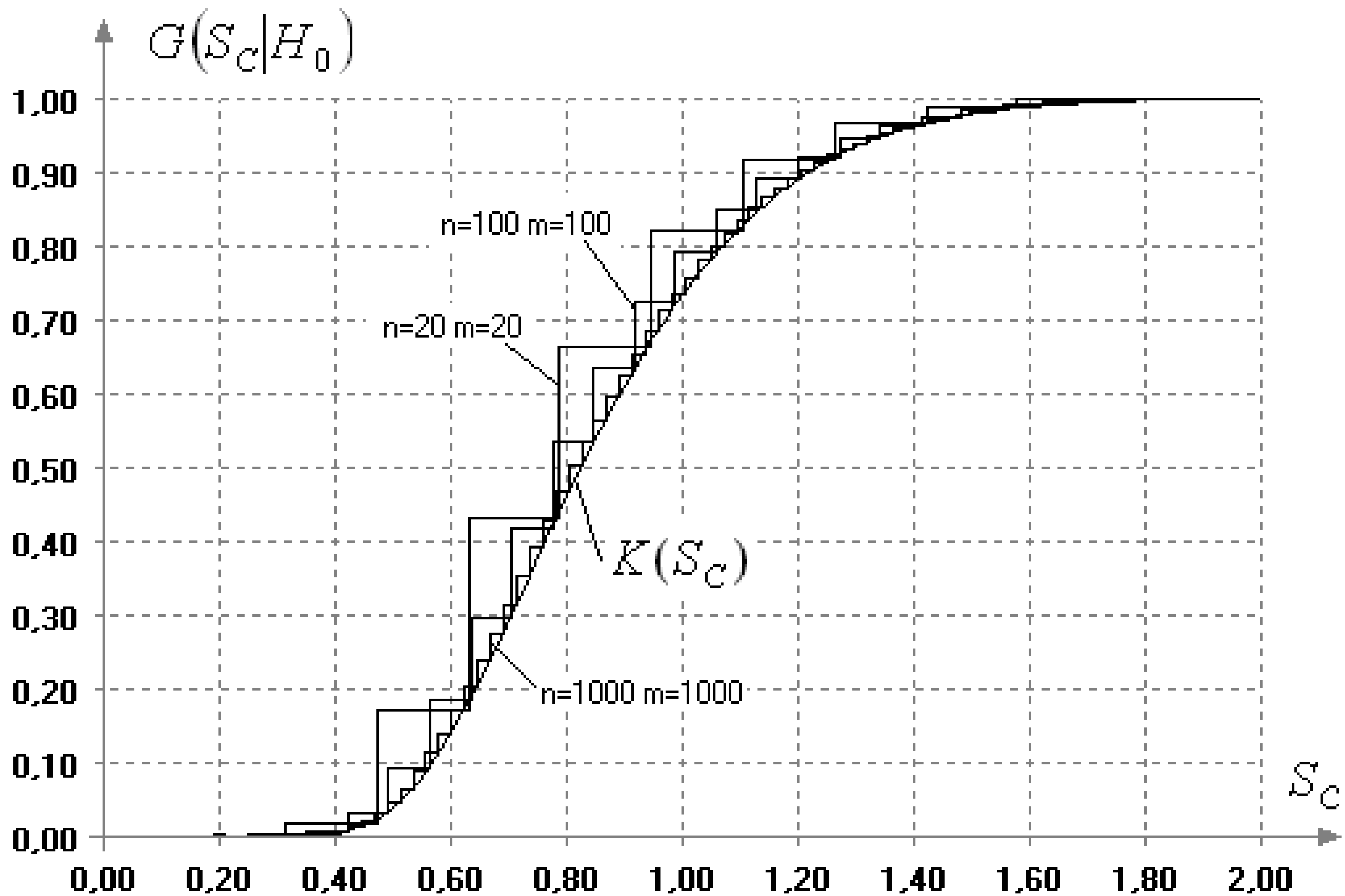
При практическом использовании критерия значение статистики $D_{m,n}$ рекомендуется вычислять в соответствии с соотношениями

$$D_{m,n}^{+} = \max_{1 \le r \le m} \left[\frac{r}{m} - F_n(x_r) \right] = \max_{1 \le s \le n} \left[G_m(y_s) - \frac{s-1}{n} \right],$$

$$D_{m,n}^{-} = \max_{1 \le r \le m} \left[F_n(x_r) - \frac{r-1}{m} \right] = \max_{1 \le s \le n} \left[\frac{s}{n} - G_m(y_s) \right],$$

$$D_{m,n} = \max(D_{m,n}^{+}, D_{m,n}^{-}).$$

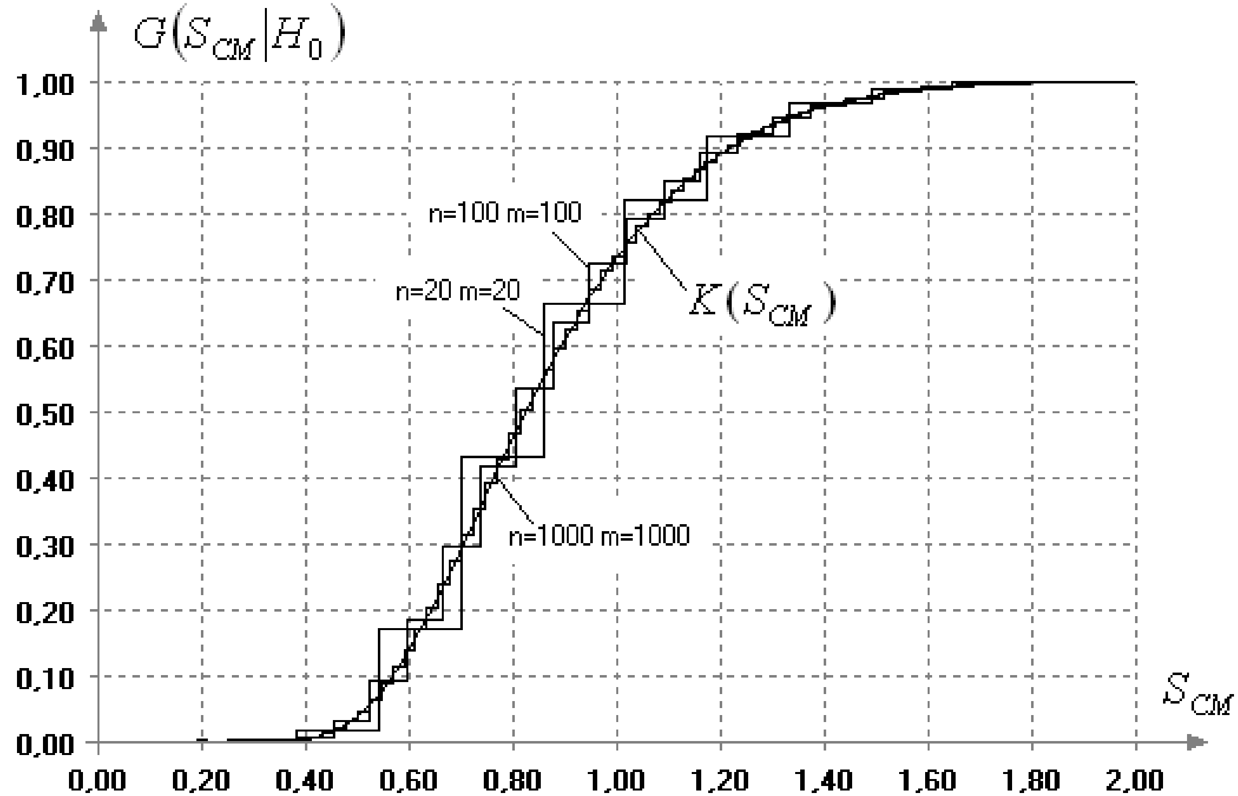
Статистика в пределе подчиняется распределению Колмогорова $K(s)$. Однако при ограниченных значениях m и n случайные величины $D_{m,n}^{+}$ и $D_{m,n}$ являются дискретными, и множество их возможных значений представляет собой решетку с шагом $1/k$, где k наименьшее общее кратное m и n.



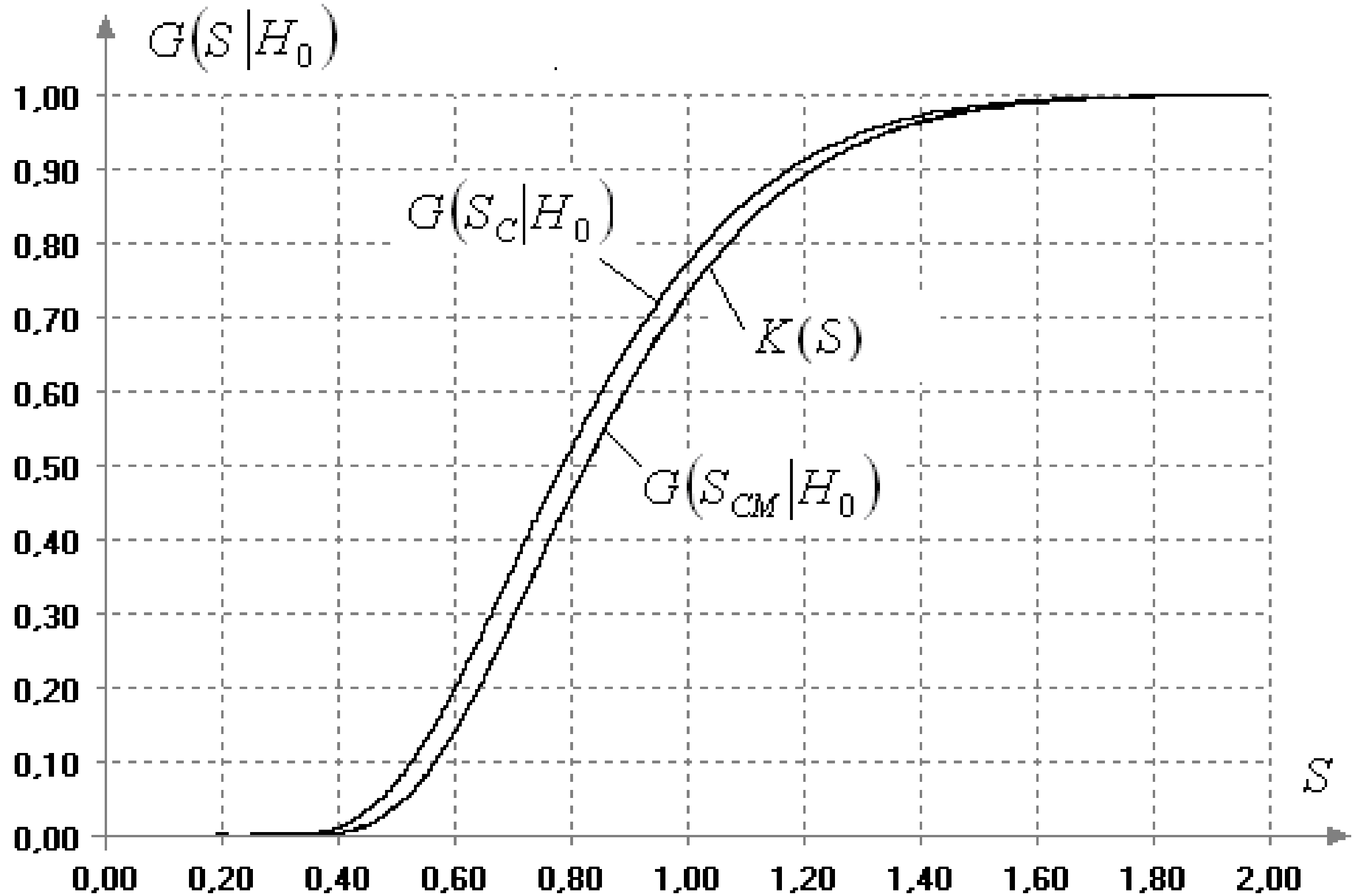
Распределения статистики критерия Смирнова при справедливости H_0 в зависимости от *m* и *n*

Распределение статистики Смирнова заметно сдвинуто влево от $K(s)$. В этой связи **предложена простая модификация статистики**,

$$S_{CM} = \sqrt{\frac{mn}{m+n}}\left(D_{m,n} + \frac{m+n}{4.6mn}\right),$$



Распределения модифицированной статистики при справедливости H_0 в зависимости от m и n



Распределения статистик Смирнова и модифицированнойпри справедливости H_0, и взаимно простых $m = 61$ и $n = 53$

Чтобы распределение статистики критерия Смирнова было более гладким, объемы выборок m и n надо выбирать в виде взаимно простых чисел.

Исследование мощности критериев однородности

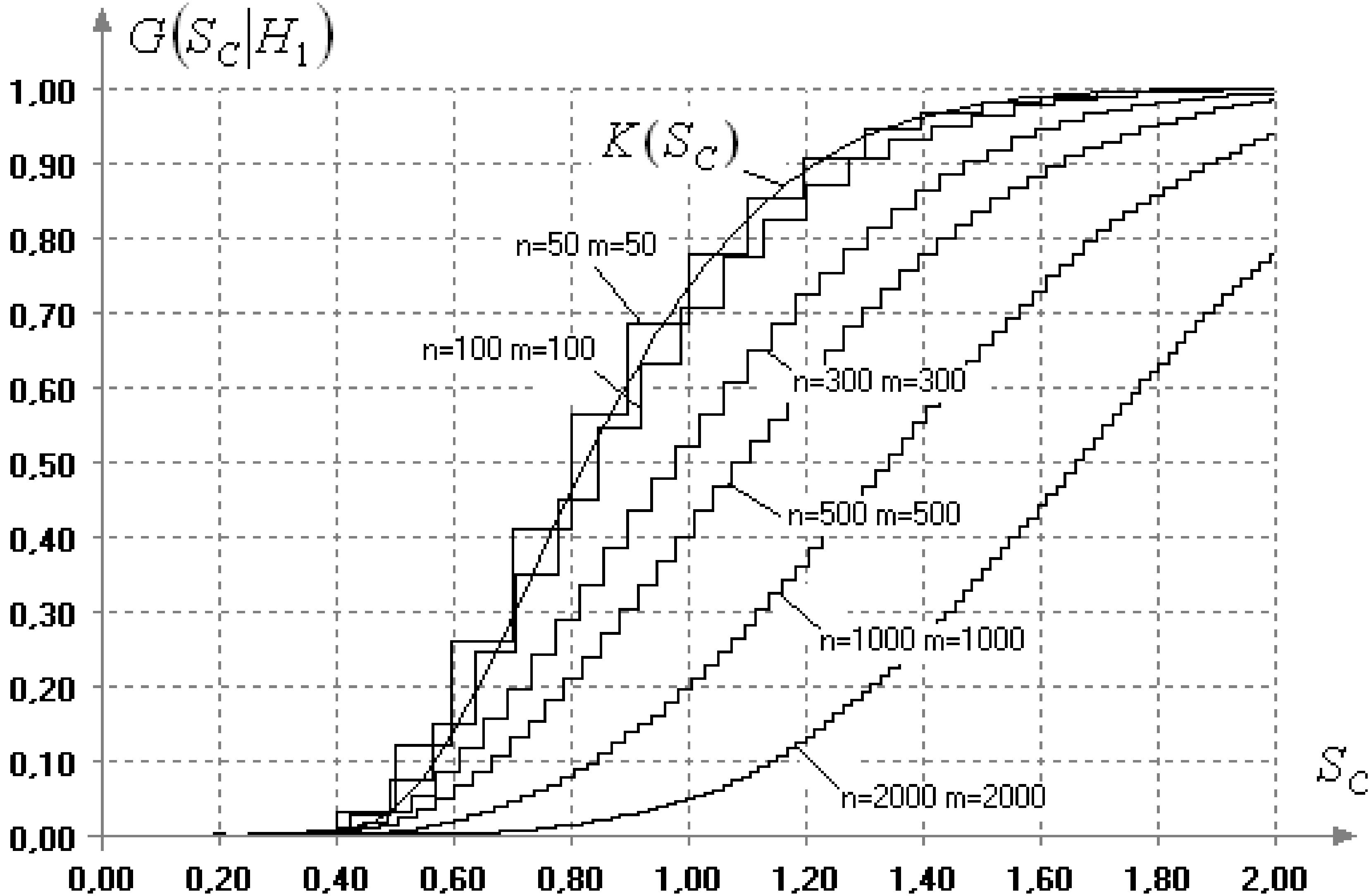
Для определенности гипотезе H_0 соответствовала принадлежность выборок одному и тому же стандартному нормальному закону распределения с плотностью

$$f(x) = \frac{1}{\theta_1\sqrt{2\pi}} \exp\left\{-\frac{(x-\theta_0)^2}{2\theta_1^2}\right\}$$

с параметрами сдвига $\theta_0 = 0$ и масштаба $\theta_1 = 1$. При всех альтернативах первая выборка всегда соответствовала стандартному нормальному закону, а вторая – некоторому другому. В частности, в случае гипотезы H_1 вторая выборка соответствовала нормальному закону с параметром сдвига $\theta_0 = 0.1$ и параметром масштаба $\theta_1 = 1$. В случае гипотезы H_2 – нормальному закону с параметрами $\theta_0 = 0.5$ и $\theta_1 = 1$. В случае гипотезы H_3 – нормальному закону с параметрами $\theta_0 = 0$ и $\theta_1 = 1.1$. В случае гипотезы H_4 – нормальному закону с параметрами $\theta_0 = 0$ и $\theta_1 = 1.5$. В случае гипотезы H_5 – вторая выборка соответствовала логистическому закону с плотностью

$$f(x) = \frac{\pi}{\theta_1\sqrt{3}} \exp\left\{-\frac{\pi(x-\theta_0)}{\theta_1\sqrt{3}}\right\} \Bigg/ \left[1 + \exp\left\{-\frac{\pi(x-\theta_0)}{\theta_1\sqrt{3}}\right\}\right]^2$$

и параметрами $\theta_0 = 0$ и $\theta_1 = 1$.



Распределения статистики Смирнова при справедливости H_1 $N(0.1;1)$

Критерий однородности Лемана-Розенблатта

Критерий однородности представляет собой критерий типа ω^2. Статистика критерия имеет вид

$$T = \frac{mn}{m+n} \int_{-\infty}^{\infty} \left[G_m(x) - F_n(x) \right]^2 dH_{m+n}(x),$$

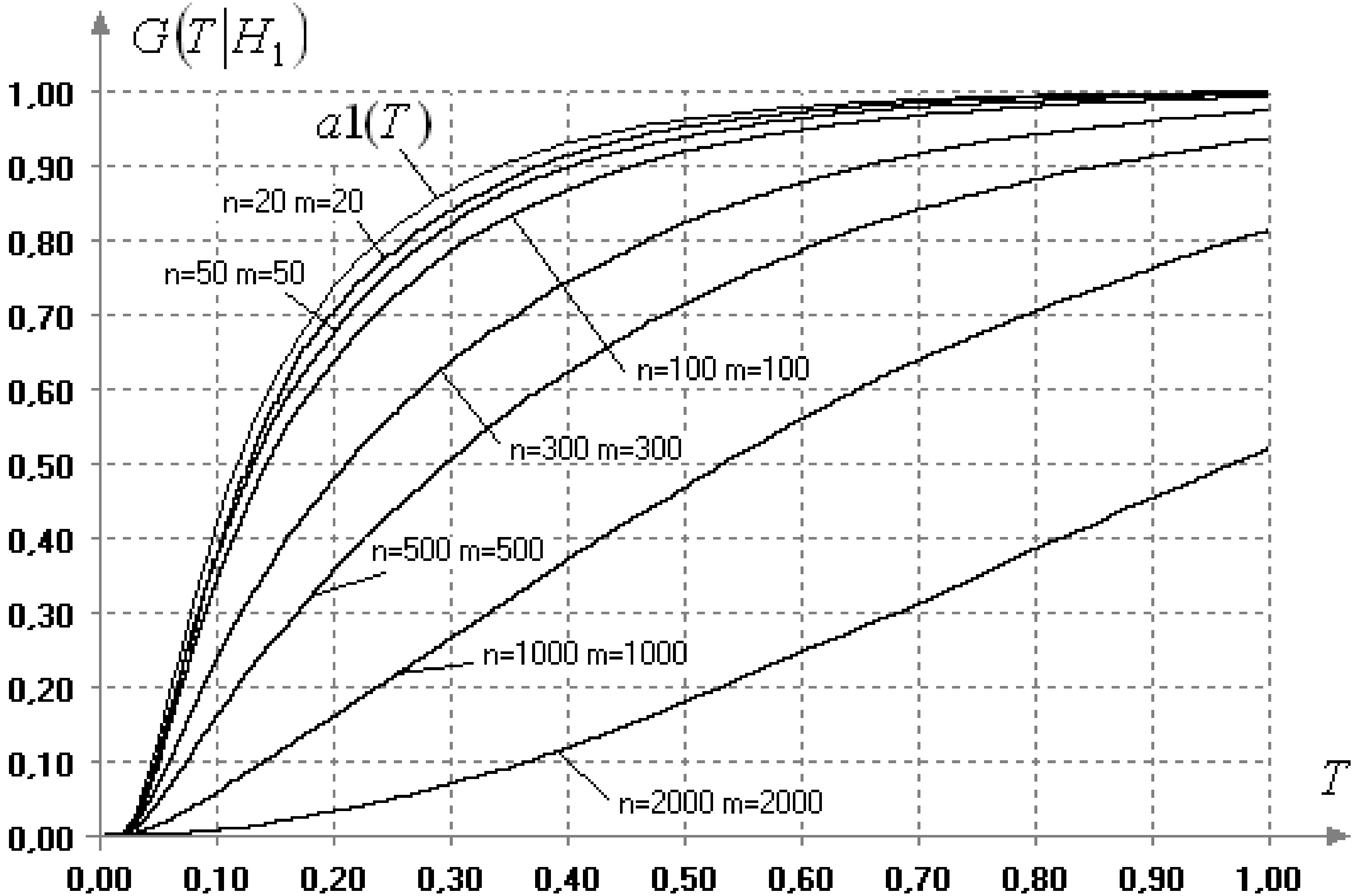
где $H_{m+n}(x) = \frac{m}{m+n} G_m(x) + \frac{n}{m+n} F_n(x)$ – эмпирическая функция распределения, построенная по вариационному ряду объединения двух выборок.

Статистика T используется в форме

$$T = \frac{1}{mn(m+n)} \left[n \sum_{i=1}^{n} (r_i - i)^2 + m \sum_{i=1}^{m} (s_i - j)^2 \right] - \frac{mn-1}{6(m+n)}, \tag{4}$$

где r_i – порядковый номер (ранг) y_i, s_j – порядковый номер (ранг) x_j в объединенном вариационном ряде.

Статистика (4) в пределе распределена как $a1(t)$ и, в отличие от статистики критерия Смирнова, быстро сходится к предельному.



Распределения статистики Лемана-Розенблатта при справедливости H_1

Критерий Андерсона–Дарлинга

Двухвыборочный критерий Андерсона–Дарлинга (критерий однородности) рассмотрен в работе [47]. Статистика критерии определяется выражением

$$A^2 = \frac{mn}{m+n} \int_{-\infty}^{\infty} \frac{[G_m(x) - F_n(x)]^2}{(1 - H_{m+n}(x))H_{m+n}(x)} dH_{m+n}(x). \tag{2.6}$$

Для выборок непрерывных случайных величин выражение для этой статистики принимает простой вид [47]

$$A^2 = \frac{1}{mn} \sum_{i=1}^{m+n-1} \frac{\left(M_i(m+n) - mi\right)^2}{i(m+n-i)}, \tag{2.7}$$

где M_i – число элементов первой выборки, меньших или равных i-му элементу вариационного ряда объединенной выборки.

При объёмах выборок $m,n \geq 45$ отклонение функции распределения $G\left(A^2 \middle| H_0\right)$ от предельного распределения $a2(A^2)$ не превышает 0.01.

Многовыборочный критерий Андерсона–Дарлинга

Необходимо проверить гипотезу вида

$$H_0 : F_1(x) = F_2(x) = \ldots = F_k(x)$$

без указания общего для них закона распределения.

В предположении о непрерывности $F_i(x)$, **упорядочив объединённую выборку** $Z_1 \le Z_2 \le \ldots \le Z_n$, $n = \sum_{i=1}^{k} n_i$, можно получить простое выражение для вычисления статистики :

$$A_{kn}^2 = \frac{1}{n} \sum_{i=1}^{k} \frac{1}{n_i} \sum_{j=1}^{n-1} \frac{\left(nM_{ij} - jn_i\right)^2}{j(n-j)}, \tag{2.10}$$

где M_{ij} – число элементов в i-й выборке, которые не больше чем Z_j.

Проверяемая гипотеза H_0 отклоняется при больших значениях статистики (2.10).

Но в критерии используется статистика вида:

$$T_{kn} = \frac{A_{kn}^2 - (k-1)}{\sqrt{D[A_{kn}^2]}}, \tag{2.11}$$

где дисперсия статистики A_{kn}^2 определяется выражением

$$D\left[A_{kn}^2\right] = \frac{an^3 + bn^2 + cn + d}{(n-1)(n-2)(n-3)} \tag{2.12}$$

при

$$a = (4g-6)(k-1) + (10-6g)H,$$

$$b = (2g-4)k^2 + 8hk + (2g-14h-4)H - 8h + 4g - 6,$$

$$c = (6h+2g-2)k^2 + (4h-4g+6)k + (2h-6)H + 4h,$$

$$d = (2h+6)k^2 - 4hk,$$

где

$$H = \sum_{i=1}^{k} \frac{1}{n_i}, \quad h = \sum_{i=1}^{n-1} \frac{1}{i}, \quad g = \sum_{i=1}^{n-2} \sum_{j=i+1}^{n-1} \frac{1}{(n-i)j}.$$

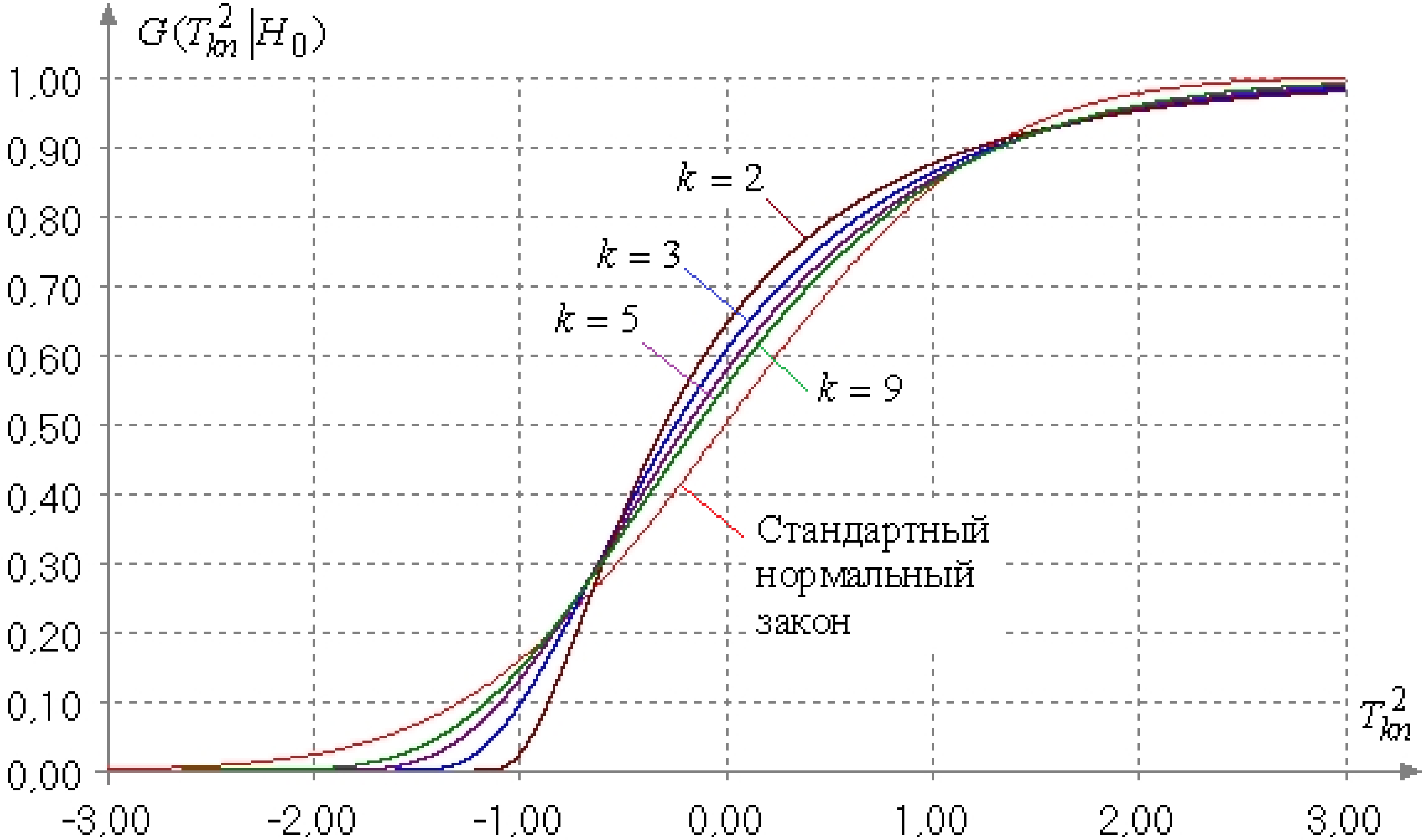


Рис. 2.20. Зависимость предельных распределений статистики (2.11) от числа сравниваемых выборок

Для предельных распределений статистик (для $k = 2 \div 11$) были построены приближенные модели законов для распределений статистики (2.11). Хорошими моделями оказались законы семейства бета-распределений III рода с плотностью

$$f(x) = \frac{\theta_2^{\theta_0}}{\theta_3 \mathrm{B}(\theta_0, \theta_1)} \frac{\left(\frac{x-\theta_4}{\theta_3}\right)^{\theta_0 - 1} \left(1 - \frac{x-\theta_4}{\theta_3}\right)^{\theta_1 - 1}}{\left[1 + (\theta_2 - 1)\frac{x-\theta_4}{\theta_3}\right]^{\theta_0 + \theta_1}}$$

при конкретных значениях параметров этого закона $\mathrm{B}_{\mathrm{III}}(\theta_0, \theta_1, \theta_2, \theta_3, \theta_4)$, найденными по выборкам статистик объёмом $N = 10^6$, полученным в результате моделирования.

Представленные в таблице 2.8 модели $\mathrm{B}_{\mathrm{III}}(\theta_0, \theta_1, \theta_2, \theta_3, \theta_4)$, с приведенными значениями параметров, позволяют по значениям статистики, вычисленным в соответствии с соотношением (2.11), находить оценки p_{value} при соответствующем числе k сравниваемых выборок.

Таблица 2.8 **Модели предельных распределений статистики (2.11)**

k	Модель
2	B_{III} (3.1575, 2.8730, 18.1238, 15.0000, –1.1600)
3	B_{III} (3.5907, 4.5984, 7.8040, 14.1310, –1.5000)
4	B_{III} (4.2657, 5.7035, 5.3533, 12.8243, –1.7500)
5	B_{III} (6.2992, 6.5558, 5.6833, 13.010, –2.0640)
6	B_{III} (6.7446, 7.1047, 5.0450, 12.8562, –2.2000)
7	B_{III} (6.7615, 7.4823, 4.0083, 11.800, –2.3150)
8	B_{III} (5.8057, 7.8755, 2.9244, 10.900, –2.3100)
9	B_{III} (9.0736, 7.4112, 4.1072, 10.800, –2.6310)
10	B_{III} (10.2571, 7.9758, 4.1383, 11.186, –2.7988)
11	B_{III} (10.6848, 7.5950, 4.2041, 10.734, –2.8400)
∞	N(0.0, 1.0)

Критерии однородности Жанга

Пусть $x_{i1}, x_{i2}, \ldots, x_{in_i}$ упорядоченные выборки непрерывных случайных величин с функциями распределения $F_i(x)$, $(i = \overline{1,k})$ и пусть $X_1 < X_2 < \ldots < X_n$, где $n = \sum_{i=1}^{k} n_i$, объединённая упорядоченная выборка.

Обозначим R_{ij} ранг j-го упорядоченного наблюдения x_{ij} i-й выборки в объединённой выборке. Пусть $X_0 = -\infty$, $X_{n+1} = +\infty$, а ранги $R_{i,0} = 1$, $R_{i,n_i+1} = n+1$.

В критериях используется модификация эмпирической функции распределения $\hat{F}(t)$, принимающая в точках разрыва X_m, $m = \overline{1,n}$, значения $\hat{F}(X_m) = (m - 0.5) / n$.

Статистика Z_{K} критерия однородности Жанга имеет вид:

$$Z_{\text{K}} = \max_{1 \le m \le n} \left\{ \sum_{i=1}^{k} n_i \left[F_{i,m} \ln \frac{F_{i,m}}{F_m} + \left(1 - F_{i,m}\right) \ln \frac{1 - F_{i,m}}{1 - F_m} \right] \right\}, \tag{2.13}$$

где $F_m = \hat{F}(X_m)$, так что $F_m = (m - 0.5)/n$, а вычисление $F_{i,m} = \hat{F}_i(X_m)$ осуществляется следующим образом. В начальный момент значения $j_i = 0$, $i = \overline{1,k}$. Если $R_{i,j_i+1} = m$, то $j_i := j_i + 1$ и $F_{i,m} = (j_i - 0.5)/n_i$, в противном случае если $R_{i,j_i} < m < R_{i,j_i+1}$, то $F_{i,m} = j_i / n_i$.

Критерий правосторонний: проверяемая гипотеза H_0 отклоняется при **больших** значениях статистики (2.13).

Зависимость распределений $G(Z_{\text{K}} | H_0)$ статистики от объёмов выборок (при равных n_i и $k = 2$) иллюстрирует рис. 2.22, а зависимость от числа сравниваемых выборок k при $n_i = 20$, $i = \overline{2,k}$ – рис. 2.23.

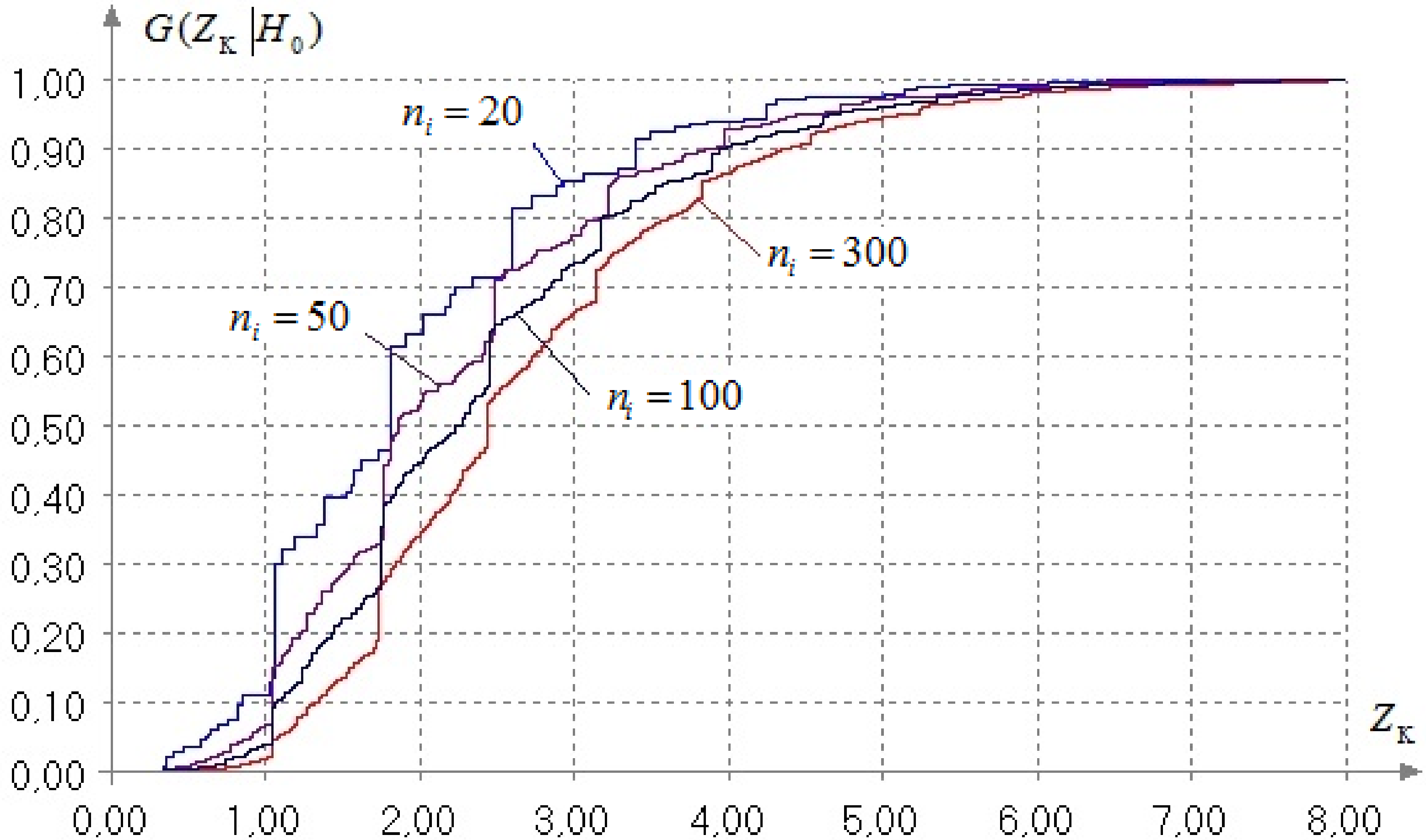


Рис. 2.22. Зависимость распределений статистики (2.13) от объёмов выборок

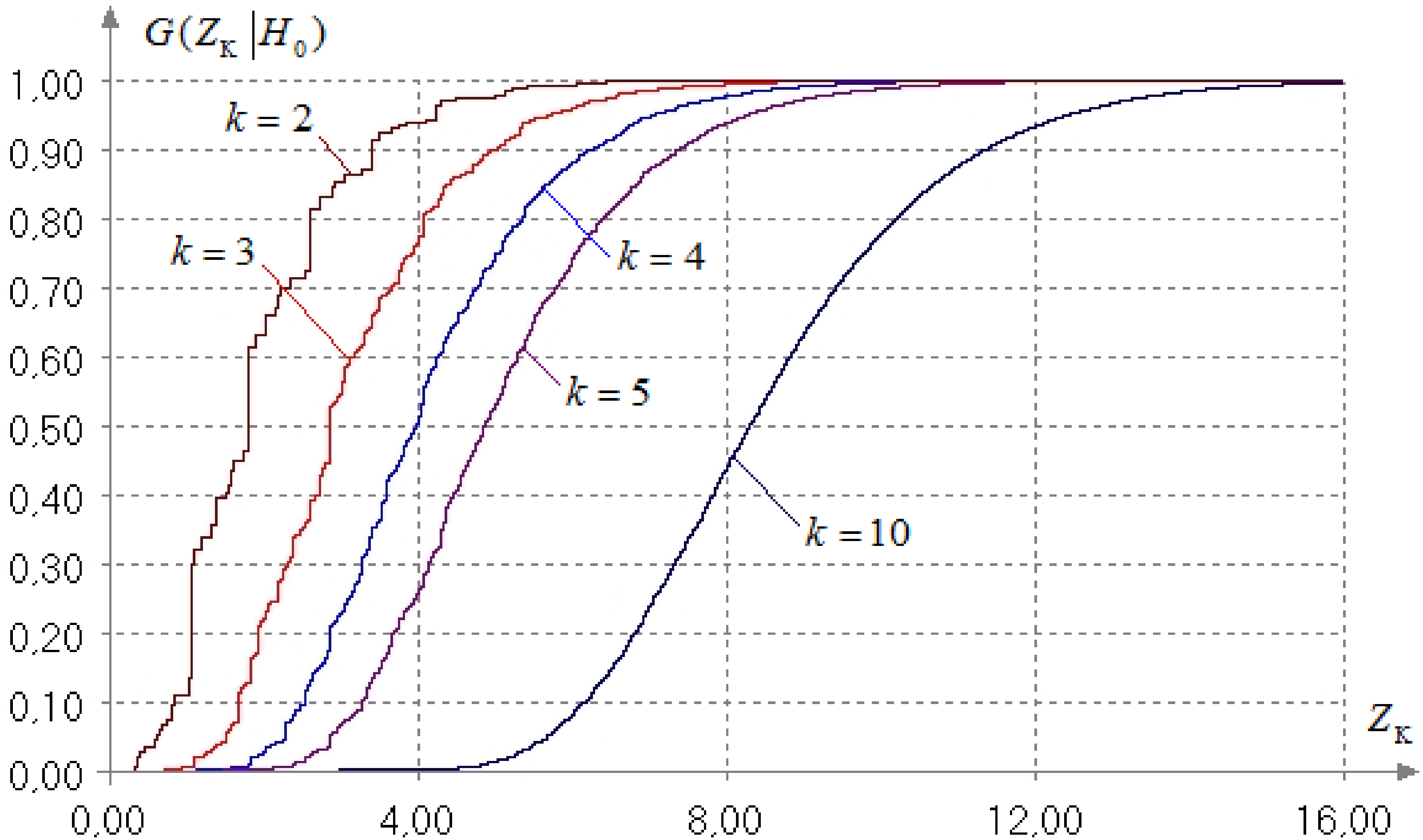


Рис. 2.23. Зависимость распределений статистики (2.13) от k при $n_i = 20$

Статистика Z_{A} **критерия однородности** k **выборок** определяется выражением:

$$Z_{\mathrm{A}} = -\sum_{m=1}^{n}\sum_{i=1}^{k} n_i \frac{F_{i,m}\ln F_{i,m} + \left(1-F_{i,m}\right)\ln\left(1-F_{i,m}\right)}{(m-0.5)(n-m+0.5)}, \qquad (2.14)$$

где F_m и $F_{i,m}$ вычисляются, как определено выше.

Критерий левосторонний: проверяемая гипотеза H_0 отклоняется при **малых** значениях статистики (2.14).

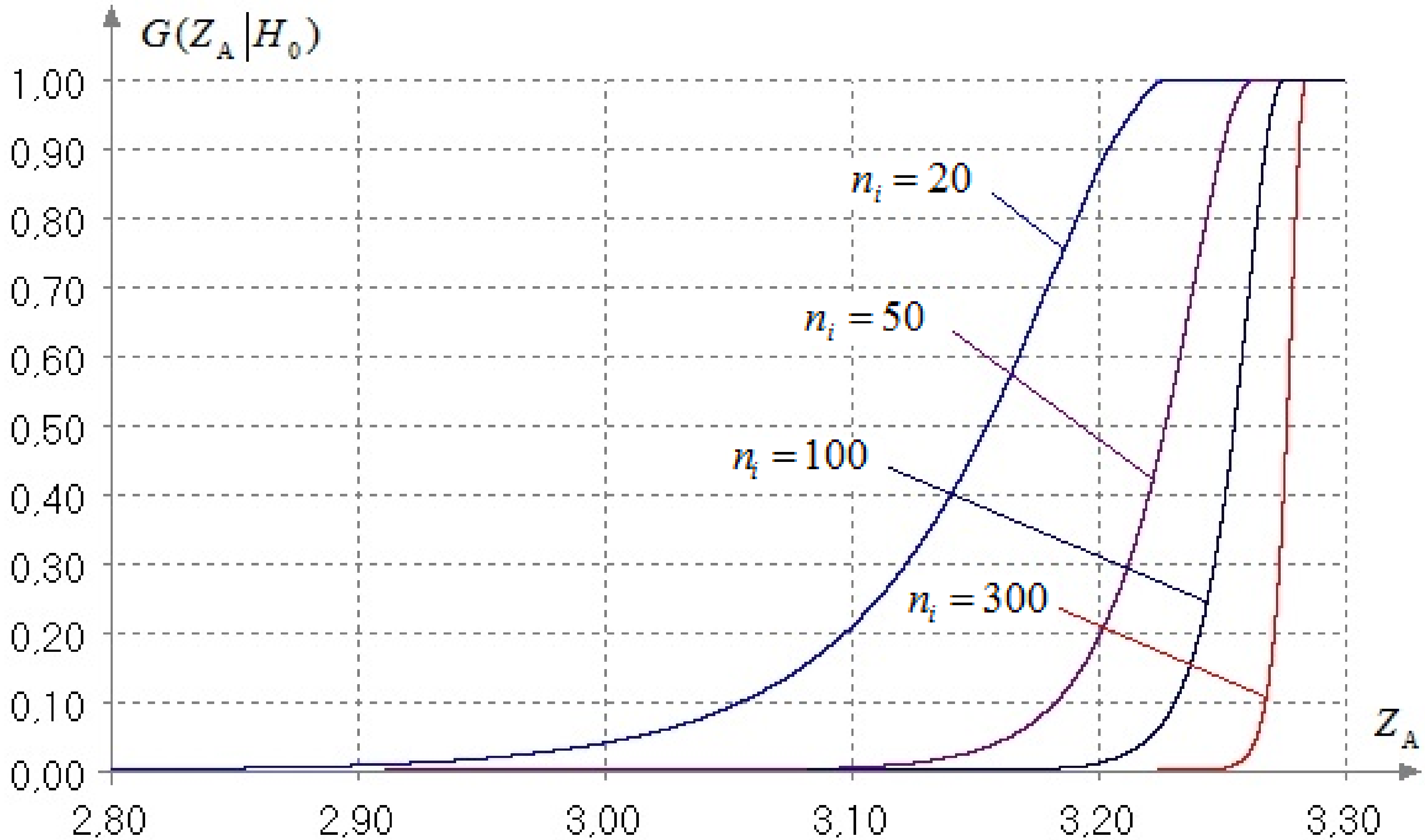


Рис. 2.24. Зависимость распределений статистики (2.14) от объёмов выборок при равных n_i и $k = 2$

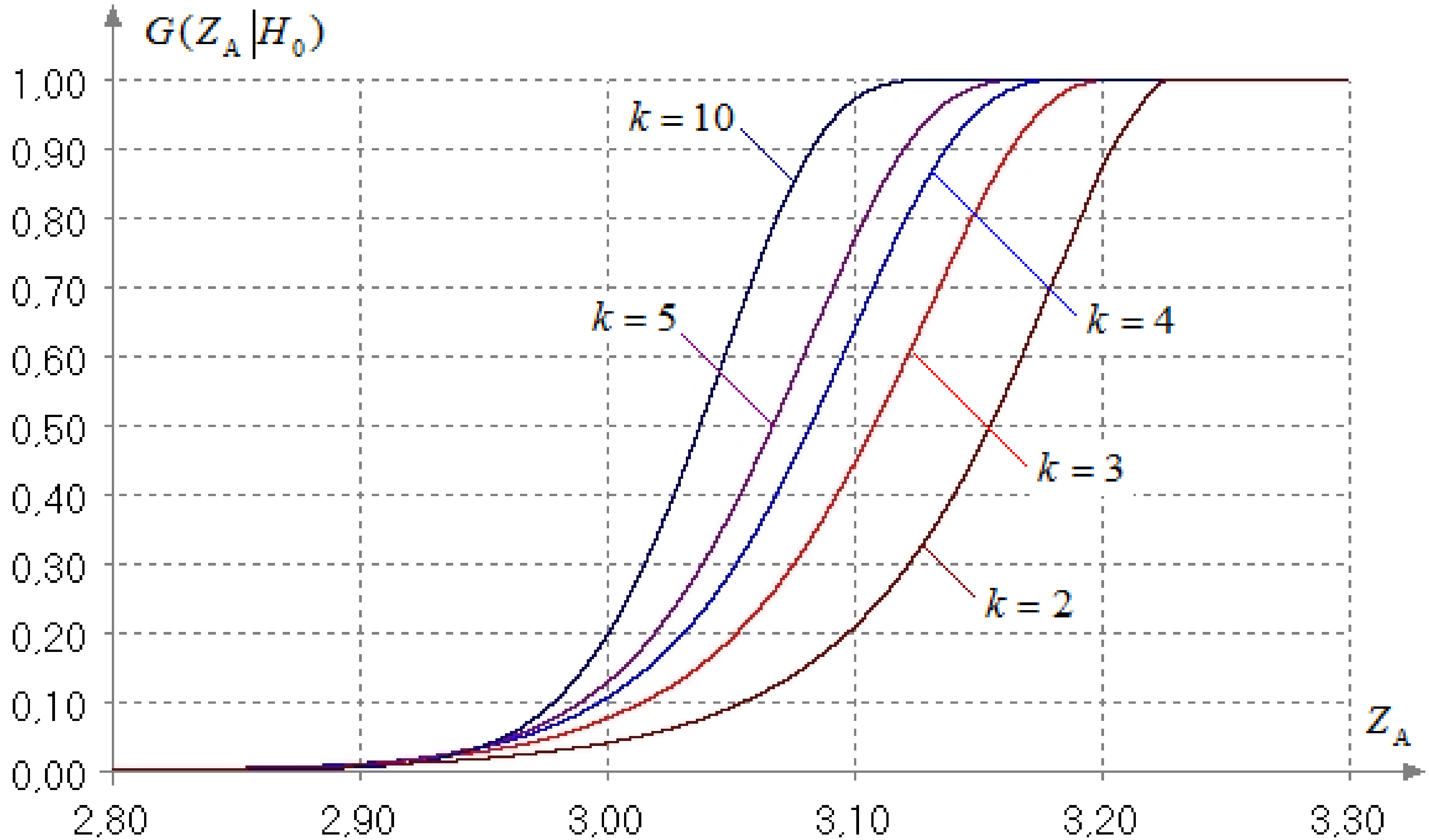


Рис. 2.25. Зависимость распределений статистики (2.14) от k при $n_i = 20$, $i = \overline{2,k}$

Статистика Z_C **критерия однородности** k **выборок** определяется выражением:

$$Z_C = \frac{1}{n}\sum_{i=1}^{k}\sum_{j=1}^{n_i}\ln\left(\frac{n_i}{j-0.5}-1\right)\ln\left(\frac{n}{R_{i,j}-0.5}-1\right). \qquad (2.15)$$

Критерий левосторонний: проверяемая гипотеза H_0 отклоняется при **малых** значениях статистики (2.15).

Изменение распределений $G(Z_C | H_0)$ статистики от объёмов выборок при равных n_i и $k=2$ демонстрируется на рис. 2.26, а зависимость от числа сравниваемых выборок k при $n_i = 20,\ i=\overline{2,k}$ – на рис. 2.27.

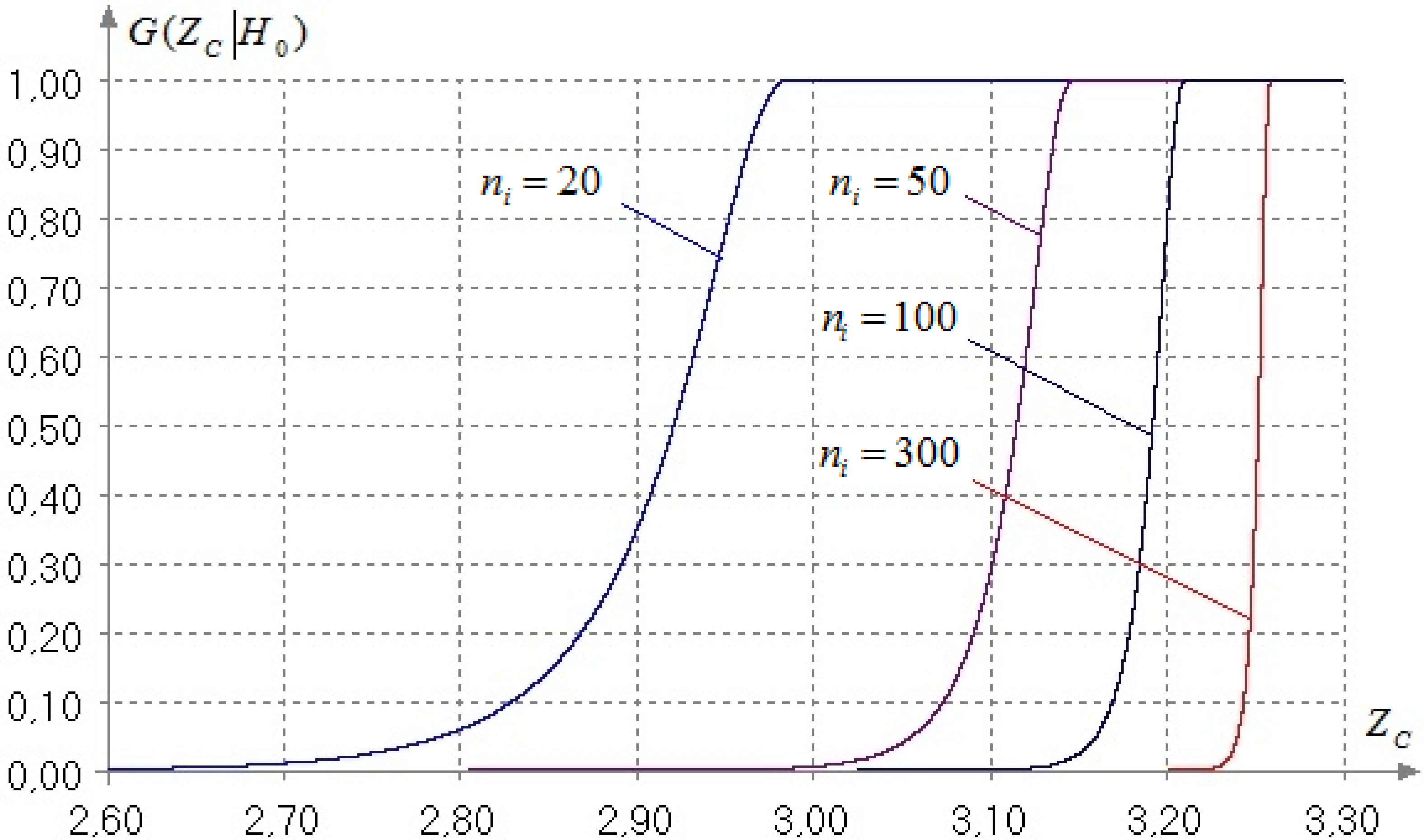


Рис. 2.26. Зависимость распределений статистики (2.15) от объёмов выборок при равных n_i и $k = 2$

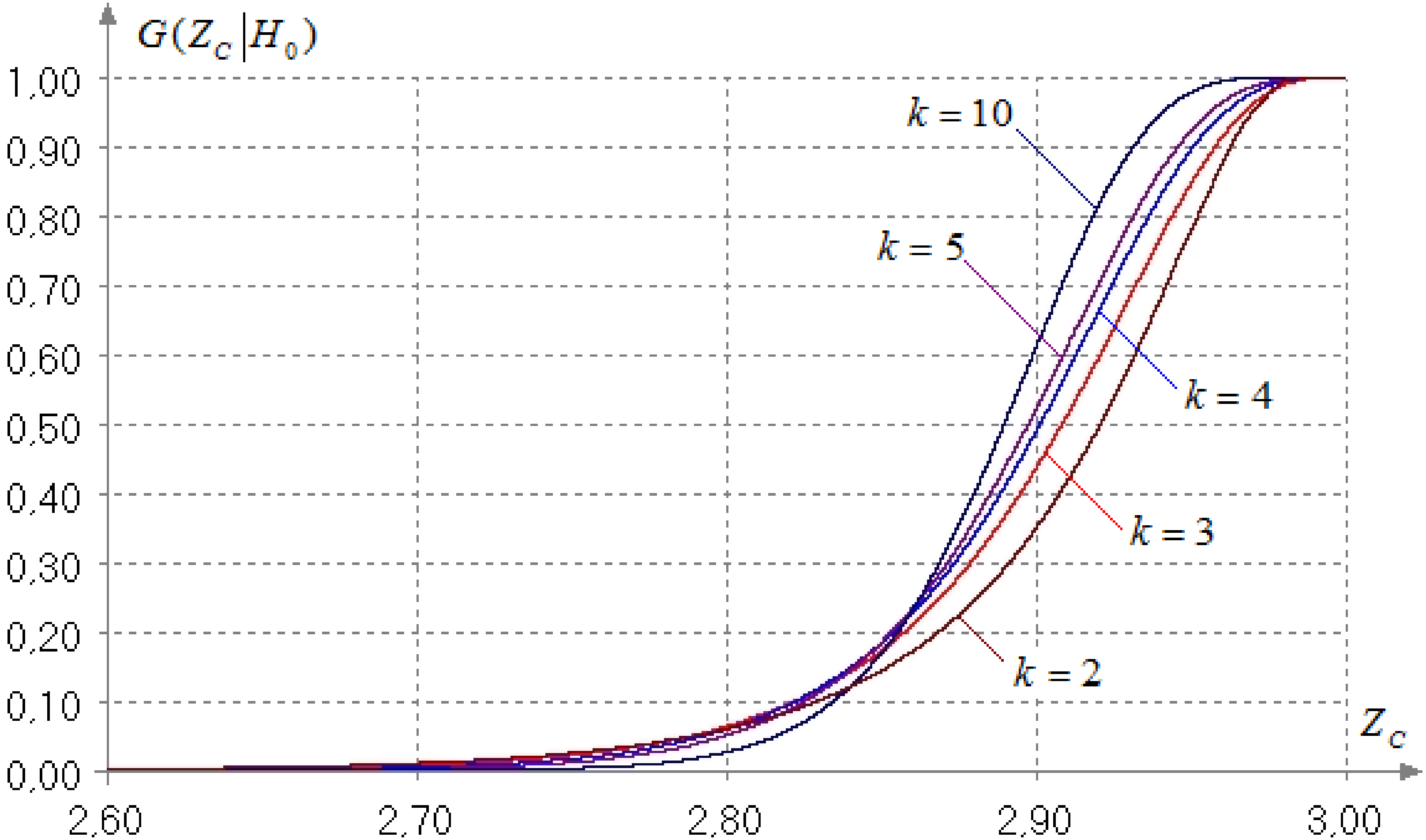


Рис. 2.27. Зависимость распределений статистики (2.15) от k при $n_i = 20$

Table 1: **Estimates of the power of tests to alternatives $H_1 - H_5$ where $k = 2$ with equal n_i, $\alpha = 0.1$**

Test	n_i =20	n_i =50	n_i =100	n_i =300	n_i =500	n_i =1000	n_i = 2000
Concerning the alternative H_1							
AD	0.114	0.137	0.175	0.319	0.447	0.691	0.919
LR	0.115	0.136	0.173	0.313	0.438	0.678	0.910
Z_C	0.114	0.134	0.164	0.278	0.382	0.600	0.869
Sm	0.111	0.132	0.164	0.280	0.381	0.617	0.859
Z_A	0.113	0.133	0.162	0.272	0.374	0.583	0.851
Z_K	0.111	0.126	0.152	0.238	0.333	0.526	0.798
Concerning the alternative H_2							
AD	0.435	0.768	0.959	1	1	1	1
LR	0.430	0.757	0.954	1	1	1	1
Z_C	0.425	0.743	0.946	1	1	1	1
Z_A	0.419	0.733	0.941	1	1	1	1
Sm	0.365	0.703	0.910	1	1	1	1
Z_K	0.344	0.650	0.906	1	1	1	1

Test	n_i =20	n_i =50	n_i =100	n_i =300	n_i =500	n_i =1000	n_i = 2000
Concerning the alternative H_3							
Z_C	0.107	0.127	0.163	0.320	0.468	0.748	0.961
Z_A	0.108	0.128	0.164	0.318	0.464	0.745	0.958
Z_K	0.107	0.127	0.154	0.268	0.390	0.624	0.892
AD	0.104	0.112	0.128	0.202	0.290	0.528	0.861
Sm	0.105	0.108	0.120	0.150	0.186	0.297	0.551
LR	0.103	0.107	0.114	0.149	0.1908	0.324	0.624
Concerning the alternative H_4							
Z_A	0.267	0.651	0.937	1	1	1	1
Z_C	0.256	0.640	0.936	1	1	1	1
Z_K	0.248	0.552	0.849	1	1	1	1
AD	0.185	0.424	0.777	1	1	1	1
LR	0.154	0.280	0.548	0.989	1	1	1
Sm	0.152	0.288	0.510	0.964	0.999	1	1

Test	n_i =20	n_i =50	n_i =100	n_i =300	n_i =500	n_i =1000	n_i = 2000
Concerning the alternative H_5							
Z_A	0.104	0.108	0.115	0.177	0.275	0.563	0.916
Z_C	0.104	0.108	0.116	0.1721	0.265	0.556	0.913
Z_K	0.105	0.110	0.122	0.179	0.266	0.429	0.759
AD	0.103	0.108	0.117	0.156	0.203	0.343	0.640
Sm	0.104	0.110	0.121	0.159	0.198	0.319	0.564
LR	0.103	0.106	0.113	0.142	0.178	0.288	0.547

Table.2: Estimates of power of the k-sampling homogeneity test to alternatives H_1, H_3, H_5 where $k=4$ with equal n_i

Test	n_i =20	n_i =50	n_i =100	n_i =300	n_i =500	n_i =1000
Concerning the alternative H_1						
AD	0.112	0.131	0.164	0.301	0.433	0.701
Z_C	0.111	0.126	0.155	0.260	0.368	0.595
Z_{A}	0.111	0.127	0.153	0.255	0.360	0.579
Z_{K}	0.109	0.121	0.141	0.219	0.300	0.502
Concerning the alternative H_3						
Z_C	0.106	0.122	0.158	0.306	0.468	0.761
Z_{A}	0.107	0.124	0.158	0.305	0.463	0.745
Z_{K}	0.106	0.120	0.145	0.249	0.367	0.606
AD	0.104	0.110	0.123	0.180	0.254	0.474
Concerning the alternative H_5						
Z_{A}	0.103	0.107	0.116	0.179	0.274	0.566
Z_C	0.103	0.107	0.115	0.173	0.257	0.555
Z_{K}	0.103	0.107	0.114	0.161	0.222	0.410
AD	0.102	0.106	0.113	0.143	0.179	0.291

Относительно конкурирующих гипотез, соответствующих **изменению параметра сдвига**, двухвыборочные критерии однородности по убыванию мощности располагаются в следующем порядке:

$$\text{Андерсона–Дарлинга} \succ \text{Лемана–Розенблатта} \succ \text{Жанга } Z_C \succ \text{Жанга } Z_{\text{A}} \succ \text{Смирнова} \succ$$
$$\succ \text{Жанга } Z_{\text{K}}\,.$$

Относительно конкурирующих гипотез, соответствующих **изменению параметра масштаба**, критерии располагаются уже в другом порядке:

$$\text{Жанга } Z_{\text{A}} \succ \text{Жанга } Z_C \succ \text{Жанга } Z_{\text{K}} \succ \text{Андерсона–Дарлинга} \succ \text{Лемана–Розенблатта} \succ$$
$$\succ \text{Смирнова.}$$

В ситуации, когда при конкурирующей гипотезе **одна выборка принадлежит нормальному закону, а вторая – логистическому**, критерии упорядочиваются по мощности следующим образом:

$$\text{Жанга } Z_{\text{K}} \succ \text{Жанга } Z_{\text{A}} \succ \text{Жанга } Z_C \succ \text{Смирнова} \succ$$
$$\succ \text{Андерсона–Дарлинга} \succ \text{Лемана–Розенблатта.}$$

Использование двухвыборочных критериев для анализа k выборок

Для анализа k выборок можно к каждой паре выборок применить двухвыборочный критерий со статистикой S (всего $(k-1)k/2$ вариантов), а решение принимать по совокупности результатов. В качестве статистики такого k-выборочного критерия (в случае использования правостороннего двухвыборочного критерия) можно рассмотреть, например, статистику вида **(2018 г.)**

$$S_{\max} = \max_{\substack{1 \le i \le k \\ i < j \le k}} \{S_{i,j}\},$$

где $S_{i,j}$ – значения статистик используемого двухвыборочного критерия, полученные при анализе i-й и j-й выборок.

Проверяемая гипотеза H_0 будет отклоняться при **больших** значениях статистики $S_{\max}$.

Преимуществом такого рода критерия является и то, что в результате будет определена пара выборок, различие между которыми оказывается наиболее значимым с позиций используемого двухвыборочного критерия.

В качестве $S_{i,j}$ можно использовать статистики двухвыборочных критериев Смирнова (лучше в модифицированном виде), Лемана–Розенблата, Андерсона–Дарлинга. В этом случае распределения соответствующих статистик $S_{\max}$ сходятся к некоторым предельным, модели которых могут быть найдены по результатам статистического моделирования.

2.6.1. k –выборочный критерий Смирнова (max)

В качестве $S_{i,j}$ в (2.17) в этом случае будет рассматриваться модификация статистики Смирнова (2.3), распределение которой всегда ближе к предельному распределению Колмогорова $K(S)$. Статистику $S_{\max}$ в этом случае будем обозначать как $S_{\max}^{Sm}$.

При равных объёмах сравниваемых выборок распределения статистики $S_{\max}^{Sm}$ (как и в двухвыборочном варианте) обладают существенной дискретностью (см. рис. 2.28) и отличаются от асимптотических (предельных) распределений (см. рис. 2.29). Если есть такая возможность, то предпочтительней в качестве n_i выбирать взаимно простые числа, тогда распределения $G(S|H_0)$ статистики $S_{\max}^{Sm}$ практически не будут отличаться от асимптотических.

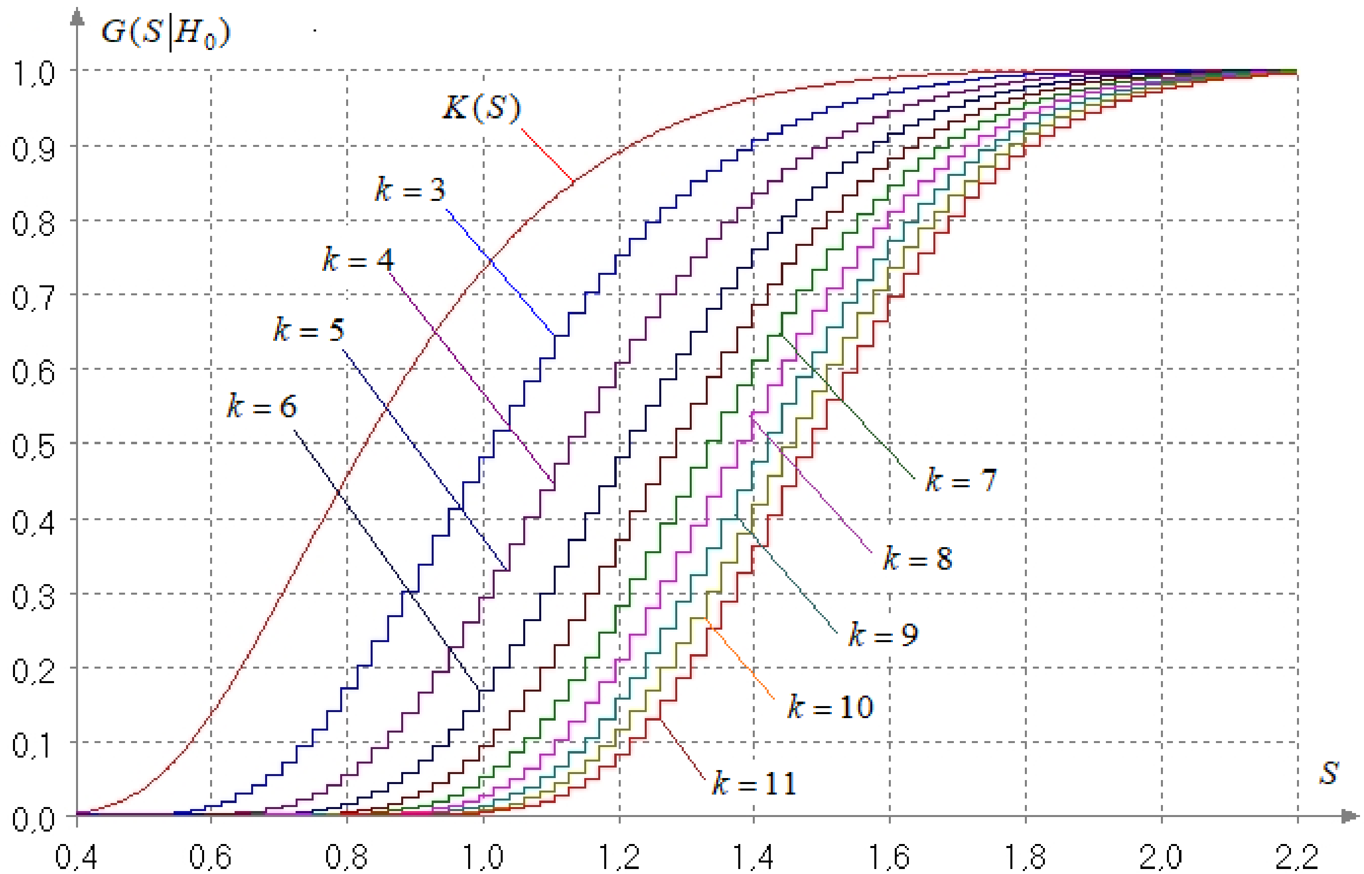


Рис. 2.28. Распределения статистики $S_{\max}^{Sm}$ k-выборочного критерия Смирнова при $n_i = 1000$, $i = \overline{1, k}$

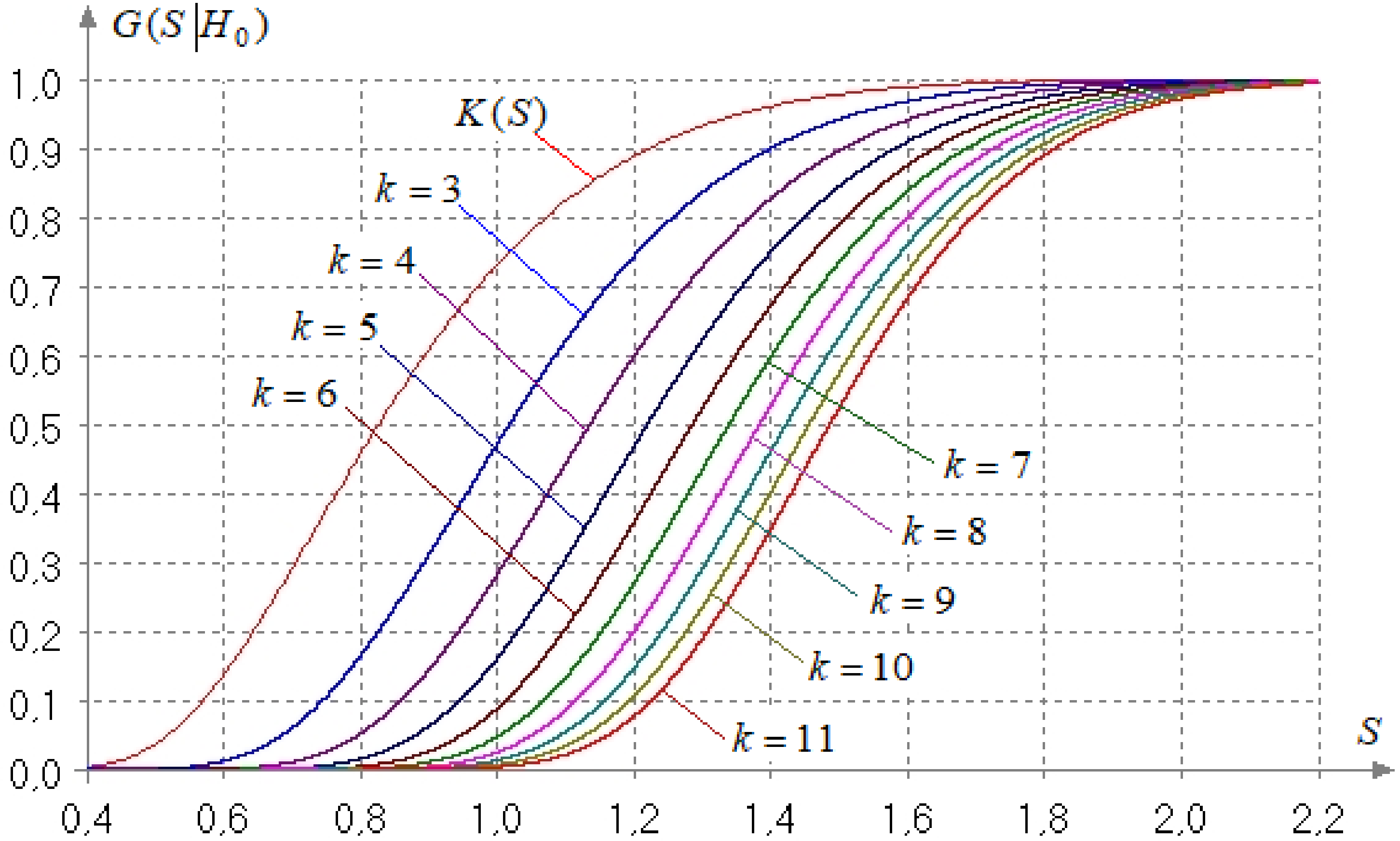


Рис. 2.29. Асимптотические распределения статистики $S_{\max}^{Sm}$ k -выборочного критерия Смирнова

Таблица 2.20

Верхние критические значения статистики $S_{\max}^{Sm}$ Смирнова

k	$1-\alpha$				
	0.75	0.90	0.95	0.975	0.99
2	1.0192	1.2238	1.3581	1.4802	1.6276
3	1.2059	1.4006	1.5266	1.6396	1.7782
4	1.3194	1.5070	1.6278	1.7372	1.8701
5	1.4002	1.5822	1.6997	1.8057	1.9329
6	1.4620	1.6396	1.7545	1.8573	1.9833
7	1.5127	1.6867	1.7989	1.9009	2.0245
8	1.5546	1.7264	1.8362	1.9362	2.0596
9	1.5911	1.7598	1.8682	1.9672	2.0879
10	1.6224	1.7893	1.8969	1.9940	2.1148
11	1.6506	1.8156	1.9217	2.0182	2.1371

Таблица 2.21

Модели предельных распределений статистики $S_{\max}^{Sm}$

k	Модель
2	$K(S)$
3	B_{III} (6.3274, 6.6162, 2.8238, 2.4073, 0.4100)
4	B_{III} (7.2729, 7.2061, 2.6170, 2.3775, 0.4740)
5	B_{III} (7.1318, 7.3365, 2.4813, 2.3353, 0.5630)
6	B_{III} (7.0755, 8.0449, 2.3163, 2.3818, 0.6320)
7	B_{III} (7.7347, 8.6845, 2.3492, 2.4479, 0.6675)
8	B_{III} (7.8162, 8.9073, 2.2688, 2.4161, 0.7120)
9	B_{III} (7.8436, 8.8805, 2.1696, 2.3309, 0.7500)
10	B_{III} (7.8756, 8.9051, 2.1977, 2.3280, 0.7900)
11	B_{III} (7.9122, 9.0411, 2.1173, 2.2860, 0.8200)

2.6.2. k –выборочный критерий Лемана–Розенблатта (max)

В качестве $S_{i,j}$ в статистике $S_{\max}^{LR}$ вида (2.17) в этом случае используется статистика (2.4) Лемана–Розенблатта. Зависимость распределений статистики при справедливости H_0 от числа выборок иллюстрирует рис. 2.30.

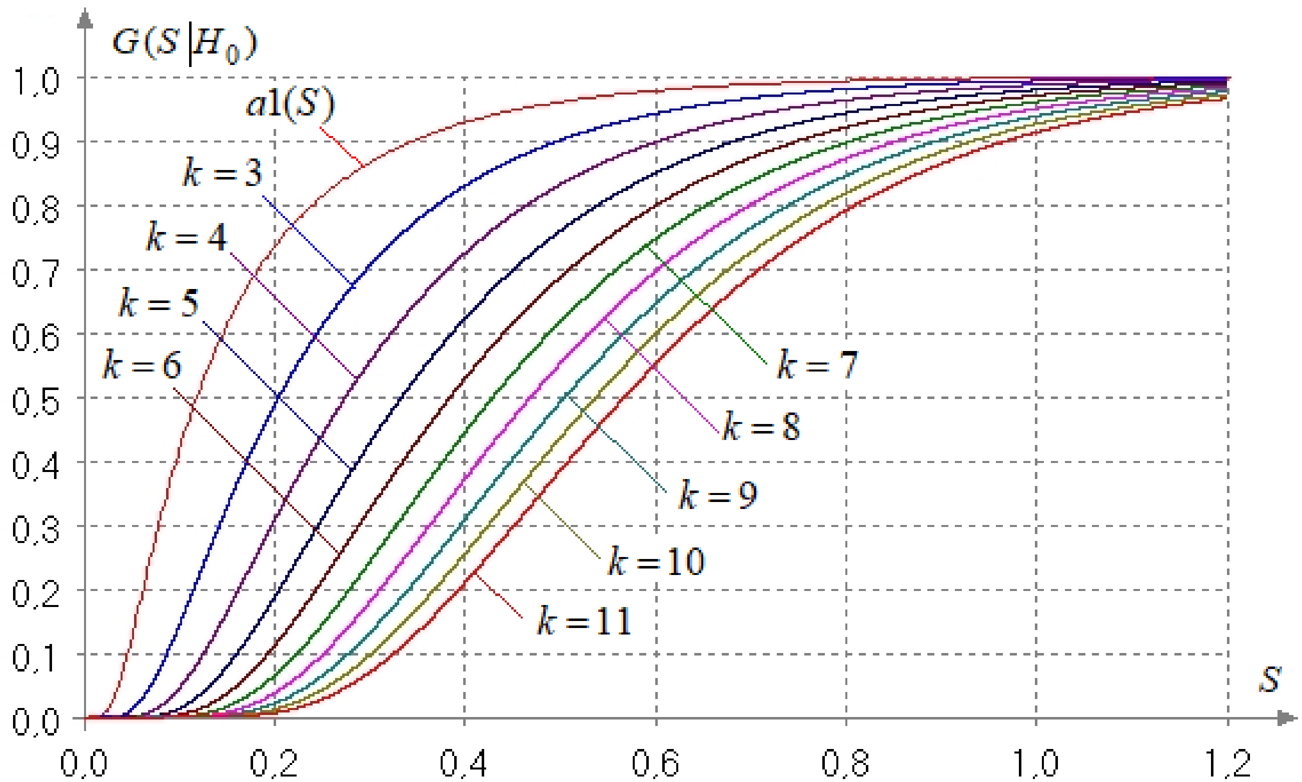


Рис. 2.30. Распределения статистики $S_{\max}^{LR}$ k-выборочного критерия Лемана–Розенблатта

Таблица 2.24

Верхние критические значения статистики $S_{\max}^{LR}$ Лемана–Розенблатта

k	$1-\alpha$				
	0.75	0.90	0.95	0.975	0.99
2	0.2094	0.3473	0.4614	0.5806	0.7435
3	0.3306	0.4995	0.6283	0.7581	0.9308
4	0.4206	0.6050	0.7413	0.8770	1.0550
5	0.4924	0.6856	0.8267	0.9653	1.1429
6	0.5521	0.7512	0.8959	1.0365	1.2175
7	0.6037	0.8072	0.9524	1.0947	1.2774
8	0.6481	0.8550	1.0015	1.1444	1.3298
9	0.6876	0.8976	1.0457	1.1902	1.3781
10	0.7237	0.9355	1.0858	1.2303	1.4171
11	0.7563	0.9703	1.1214	1.2655	1.4535

Т а б л и ц а 2.25

Модели предельных распределений статистики $S_{\max}^{LR}$

k	Модель
2	$a1(t)$
3	Sb(3.2854, 1.2036, 3.0000, 0.0215)
4	Sb(2.5801, 1.2167, 2.2367, 0.0356)
5	Sb(3.1719, 1.4134, 3.1500, 0.0320)
6	Sb(2.9979, 1.4768, 2.9850, 0.0380)
7	Sb(3.2030, 1.5526, 3.4050, 0.0450)
8	Sb(3.2671, 1.6302, 3.5522, 0.0470)
9	Sb(3.4548, 1.7127, 3.8800, 0.0490)
10	Sb(3.4887, 1.7729, 3.9680, 0.0510)
11	Sb(3.4627, 1.8168, 3.9680, 0.0544)

2.6.3. k –выборочный критерий Андерсона–Дарлинга (max)

В статистике $S_{\max}^{AD}$ вида (2.17) качестве $S_{i,j}$ используется статистика (2.7) Андерсона–Дарлинга. Зависимость распределений статистики $S_{\max}^{AD}$ при справедливости H_0 от числа выборок иллюстрирует рис. 2.31.

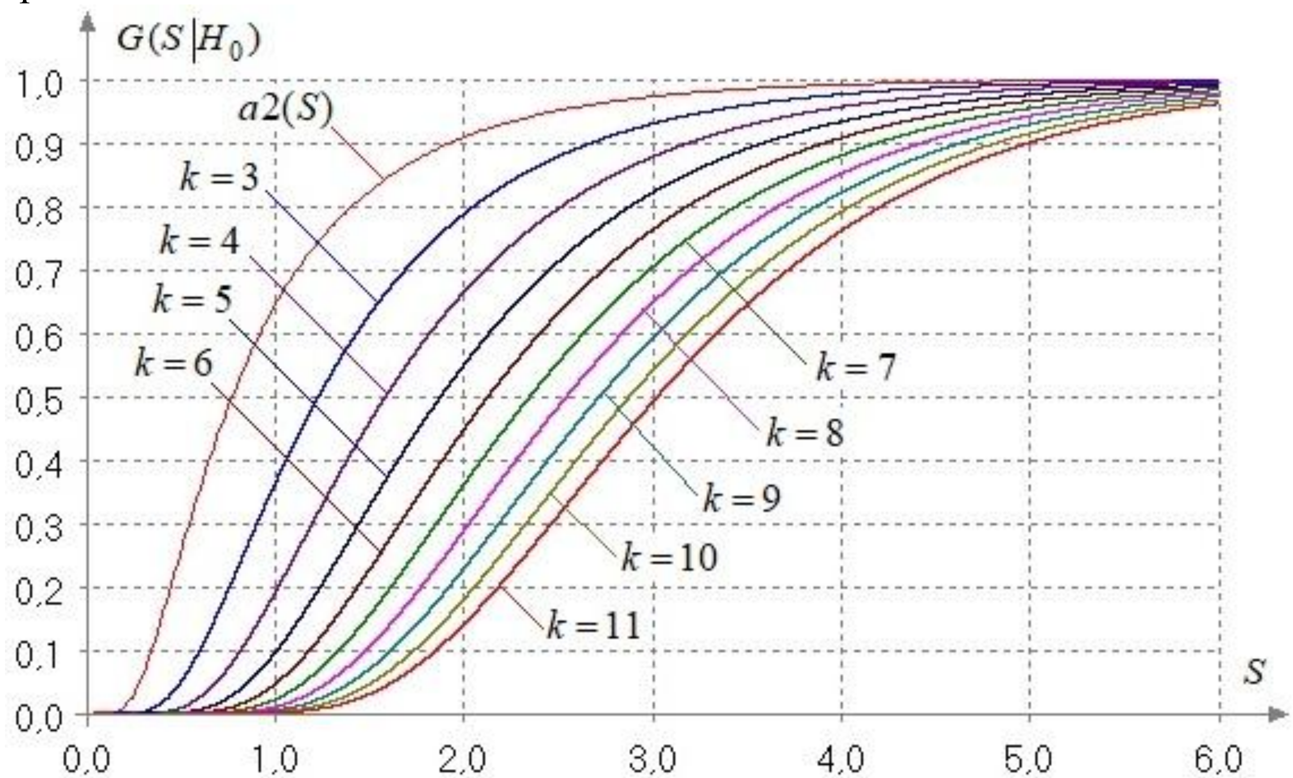


Рис. 2.31. Распределения статистики $S_{\max}^{AD}$ k -выборочного критерия Андерсона–Дарлинга

Таблица 2.28

Верхние критические значения статистики $S_{\max}^{AD}$ Андерсона–Дарлинга

k	$1-\alpha$				
	0.75	0.90	0.95	0.975	0.99
2	1.2479	1.9330	2.4924	3.0775	3.8781
3	1.8535	2.6796	3.31215	3.95176	4.7924
4	2.2990	3.1966	3.8682	4.5368	5.4076
5	2.6514	3.5948	4.2877	4.9686	5.8472
6	2.9431	3.9187	4.6292	5.3175	6.2089
7	3.1943	4.1950	4.9097	5.6063	6.5118
8	3.4135	4.4292	5.1501	5.8531	6.7710
9	3.6094	4.6395	5.3660	6.0733	7.0076
10	3.7867	4.8270	5.5616	6.2720	7.2042
11	3.9466	4.9974	5.7384	6.4512	7.3837

Таблица 2.29

Модели предельных распределений статистики $S_{\max}^{AD}$

k	Модель
2	$a2(t)$
3	B_{III} (4.4325, 2.7425, 12.1134, 8.500, 0.1850)
4	B_{III} (5.2036, 3.2160, 10.7792, 10.000, 0.2320)
5	B_{III} (5.7527, 3.3017, 9.7365, 10.000, 0.3000)
6	B_{III} (5.5739, 3.4939, 7.7710, 10.000, 0.3750)
7	B_{III} (6.4892, 3.6656, 8.0529, 10.500, 0.3920)
8	B_{III} (6.3877, 3.8143, 7.3602, 10.800, 0.4800)
9	B_{III} (6.7910, 3.9858, 7.1280, 11.100, 0.5150)
10	B_{III} (6.7533, 4.2779, 6.6457, 11.700, 0.5800)
11	B_{III} (7.1745, 4.3469, 6.6161, 11.800, 0.6100)

2.7. Критерий однородности χ^2

Пусть имеется k выборок $x_{i1}, x_{i2}, \ldots, x_{in_i}$, $i = \overline{1,k}$, непрерывных случайных величин и $n = \sum_{i=1}^{k} n_i$. Общая область, которой принадлежат выборки, разбивается на r интервалов (групп). Пусть η_{ij} – количество элементов i-й выборки, попавших в j-й интервал, тогда $n_i = \sum_{j=1}^{r} \eta_{ij}$.

Статистика критерия однородности χ^2 имеет вид

$$\chi^2 = n\sum_{i=1}^{k}\sum_{j=1}^{r}\frac{(\eta_{ij} - \nu_j n_i / n)^2}{\nu_j n_i} = n\left(\sum_{i=1}^{k}\sum_{j=1}^{r}\frac{\eta_{ij}^{2}}{\nu_j n_i} - 1\right),$$

где $\nu_j = \sum_{l=1}^{k} \eta_{lj}$ – общее число элементов всех выборок, попавших j-й интервал.

Асимптотическим распределением статистики (2.19) является χ^2-распределение с числом степеней свободы $(k-1)(r-1)$.

В случае k выборок порядок предпочтения **по мощности** для k-выборочных критериев **относительно изменения параметра сдвига** порядок предпочтения имеет вид:

$$S_{\max}^{AD} \succ \text{Андерсона–Дарлинга} \succ S_{\max}^{LR} \succ S_{\max}^{Sm} \succ \text{Жанга } Z_C \succ \text{Жанга } Z_{\mathrm{A}} \succ \text{Жанга } Z_{\mathrm{K}} \succ \chi^2 .$$

Относительно **изменения параметра масштаба** –

$$\text{Жанга } Z_C \succ \text{Жанга } Z_{\mathrm{A}} \succ \text{Жанга } Z_{\mathrm{K}} \succ \text{Андерсона–Дарлинга} \succ \chi^2 \succ S_{\max}^{AD} \succ S_{\max}^{Sm} \succ S_{\max}^{LR} .$$

При этом критерии Жанга со статистиками Z_{A} и Z_C практически эквивалентны по мощности, а критерий Андерсона–Дарлинга заметно уступает критериям Жанга.

Относительно ситуации, когда, например, три выборки принадлежат **нормальному закон**у, а четвёртая – **логистическому**, критерии располагаются по мощности в следующем порядке:

$$\text{Жанга } Z_{\mathrm{A}} \succ \text{Жанга } Z_C \succ \text{Жанга } Z_{\mathrm{K}} \succ \chi^2 \succ \text{Андерсона–Дарлинга} \succ S_{\max}^{Sm} \succ S_{\max}^{AD} \succ S_{\max}^{LR} .$$

Критерии проверки гипотез об однородности средних

при нарушении классических предположений о принадлежности наблюдений (ошибок измерений) нормальному закону

Проверяемая гипотеза о равенстве математических ожиданий задается в виде

$$H_0 : \mu_1 = \mu_2,$$

конкурирующая –

$$H_1 : \mu_1 \neq \mu_2.$$

Для проверки гипотезы H_0 **могут использоваться *параметрические критерии*, предназначенные для проверки гипотезы в случае принадлежности наблюдений нормальному закону:**

- **критерий сравнения двух выборочных средних при известных дисперсиях;**
- **критерий сравнения двух выборочных средних при неизвестных, но равных дисперсиях (критерий Стьюдента);**
- **критерий сравнения двух выборочных средних при неизвестных и неравных дисперсиях (проблема Беренса-Фишера);**
- **F-критерий однородности средних (k-выборочный).**

***Непараметрические критерии*:**

- ***U*-критерий Уилкоксона, Манна–Уитни;**
- ***H*-критерий Краскела–Уаллиса;**
- **критерий Ван дер Вардена;**
- **Критерий Фишера–Йэйтса–Терри–Гёфдинг;**
- **многовыборочный критерий Ван дер Вардена.**

Статистики непараметрических критериев

№	Статистика	Условия	Распределение статистики
1	$z=\dfrac{\bar{x}_1-\bar{x}_2-\mu_1+\mu_2}{\sqrt{\dfrac{\sigma_1^2}{n_1}+\dfrac{\sigma_2^2}{n_2}}}$, $\bar{x}_i=\dfrac{1}{n_i}\sum_{j=1}^{n_i}x_j$	σ_1^2, σ_2^2 - известны	Стандартное нормальное $N(0,1)$
2	$t=\dfrac{(\bar{x}_1-\bar{x}_2)}{\sqrt{\left[\dfrac{n_1+n_2}{n_1 n_2}\right]\left[\dfrac{Q_1+Q_2}{n_1+n_2-2}\right]}}$, $Q_i=\sum_{j=1}^{n_i}(x_{ij}-\bar{x}_i)^2$	$\sigma_1^2=\sigma_2^2$ - неизвестны	t_ν-распределение $\nu=n_1+n_2-2$
3	$t=\dfrac{(\bar{x}_1-\bar{x}_2)}{\sqrt{\left[\dfrac{s_1^2}{n_1}+\dfrac{s_2^2}{n_2}\right]}}$, $s_i=\dfrac{1}{n_i-1}\sum_{j=1}^{n_i}(x_{ij}-\bar{x}_i)^2$	$\sigma_1^2\neq\sigma_2^2$ - неизвестны	t_ν-распределение $\nu=\dfrac{\left(\dfrac{s_1^2}{n_1}+\dfrac{s_2^2}{n_2}\right)^2}{\dfrac{\left(\dfrac{s_1^2}{n_1}\right)^2}{n_1-1}+\dfrac{\left(\dfrac{s_2^2}{n_2}\right)^2}{n_2-1}}$

***F*-критерий**

В случае справедливости гипотезы о постоянстве (о равенстве) дисперсий с помощью F-критерия можно проверять гипотезу об однородности математических ожиданий по k выборкам.

Пусть у нас имеется k выборок объема n. Общая сумма квадратов отклонений по всем выборкам

$$Q_{kn} = \sum_{i=1}^{k}\sum_{j=1}^{n}\left(x_{ij} - \bar{x}_{kn}\right)^2, \text{ где } \bar{x}_{kn} = \frac{1}{kn}\sum_{i=1}^{k}\sum_{j=1}^{n} x_{ij} = \bar{\bar{x}}_k,$$

раскладывается на два компонента $Q_{kn} = Q_1 + Q_2$,

$$Q_1 = n\sum_{i=1}^{k}\left(\bar{x}_{in} - \bar{\bar{x}}_k\right)^2 = n\sum_{i=1}^{k}\left(\bar{x}_{in}^{2} - k\bar{\bar{x}}_k\right)^2,$$

$$Q_2 = \sum_{i=1}^{k}\sum_{j=1}^{n}\left(x_{ij} - \bar{x}_{in}\right)^2 = (n-1)\sum_{i=1}^{k} s_{in}^2 .$$

Для проверки гипотезы используется критерий со статистикой

$$F = \frac{Q_1/(k-1)}{Q_2/[k(n-1)]}.$$

Если все выборки извлекаются из нормальной генеральной совокупности, то при справедливости гипотезы H_0 статистка подчиняется F_{ν_1,ν_2}-распределению Фишера со степенями свободы $\nu_1 = k-1$ и $\nu_2 = k(n-1)$.

Об устойчивости параметрических критериев

Давно известно, что параметрические критерии, предназначенные для проверки гипотез об однородности средних анализируемых выборок, достаточно устойчивы к нарушению стандартного предположения о нормальности.

При исследовании степени устойчивости параметрических критериев в данном случае использовалась та же методика компьютерного моделирования и анализа статистических закономерностей, что и в других наших работах, например. Распределения статистик (3.1)–(3.3) и (3.5) при справедливой проверяемой гипотезе H_0 исследовались для различных законов, в частности, в случае принадлежности наблюдений обобщённому нормальному закону с плотностью

$$f(x) = \frac{\theta_2}{2\theta_1 \Gamma(1/\theta_2)} \exp\left\{ -\left(\frac{|x-\theta_0|}{\theta_1} \right)^{\theta_2} \right\} \qquad (3.6)$$

с различными значениями параметра формы θ_2. При $\theta_2 = 2$ выражение (3.6) дает плотность нормального закона распределения. При больших значениях θ_2 распределение (3.6) стремится к равномерному, при малых θ_2 получаем симметричные законы с «тяжелыми хвостами».

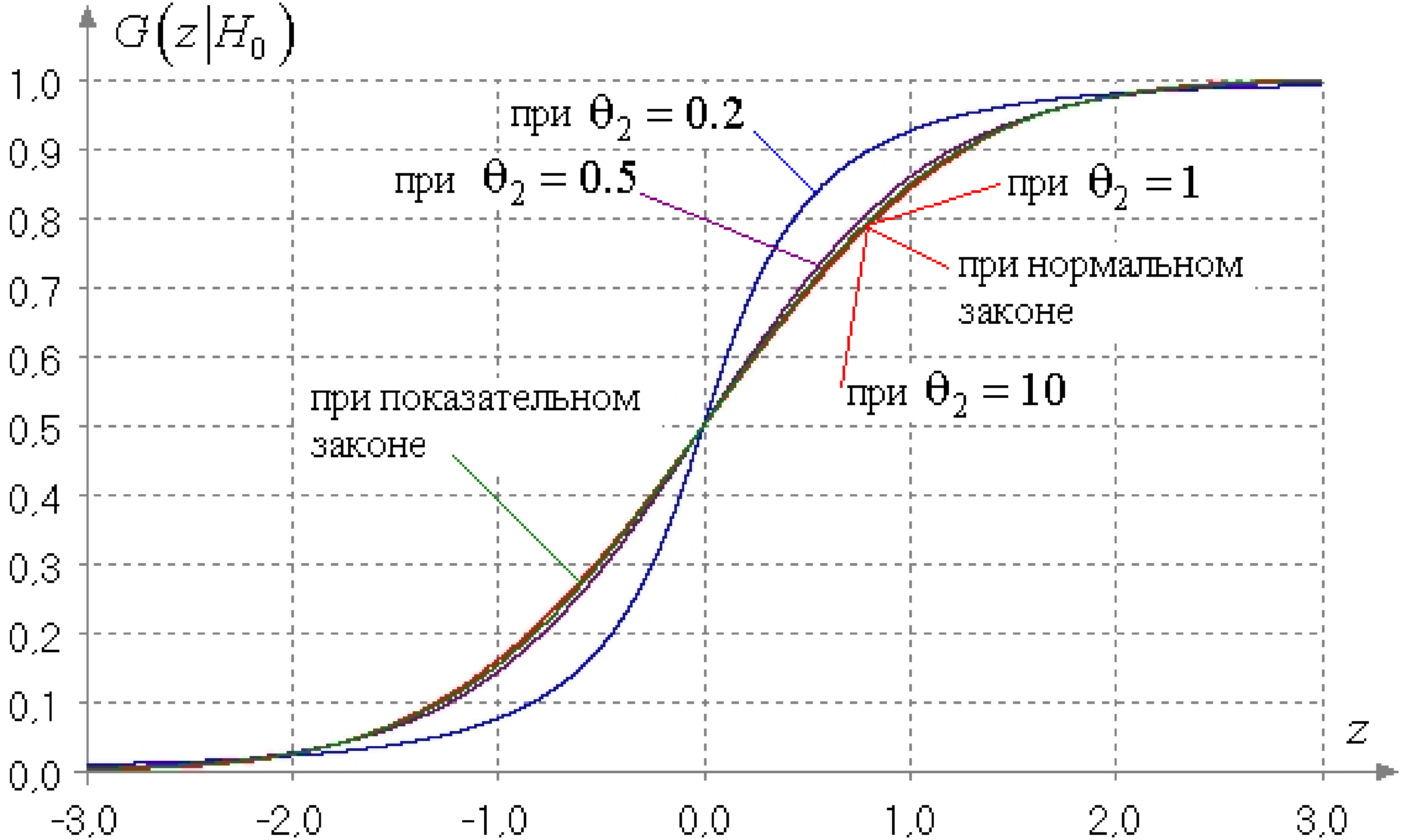


Рис. 3.2. Эмпирические распределения статистики (3.1) при различных законах распределения наблюдаемых величин и справедливости гипотезы H_0 при объемах выборок $n_1 = n_2 = 10$

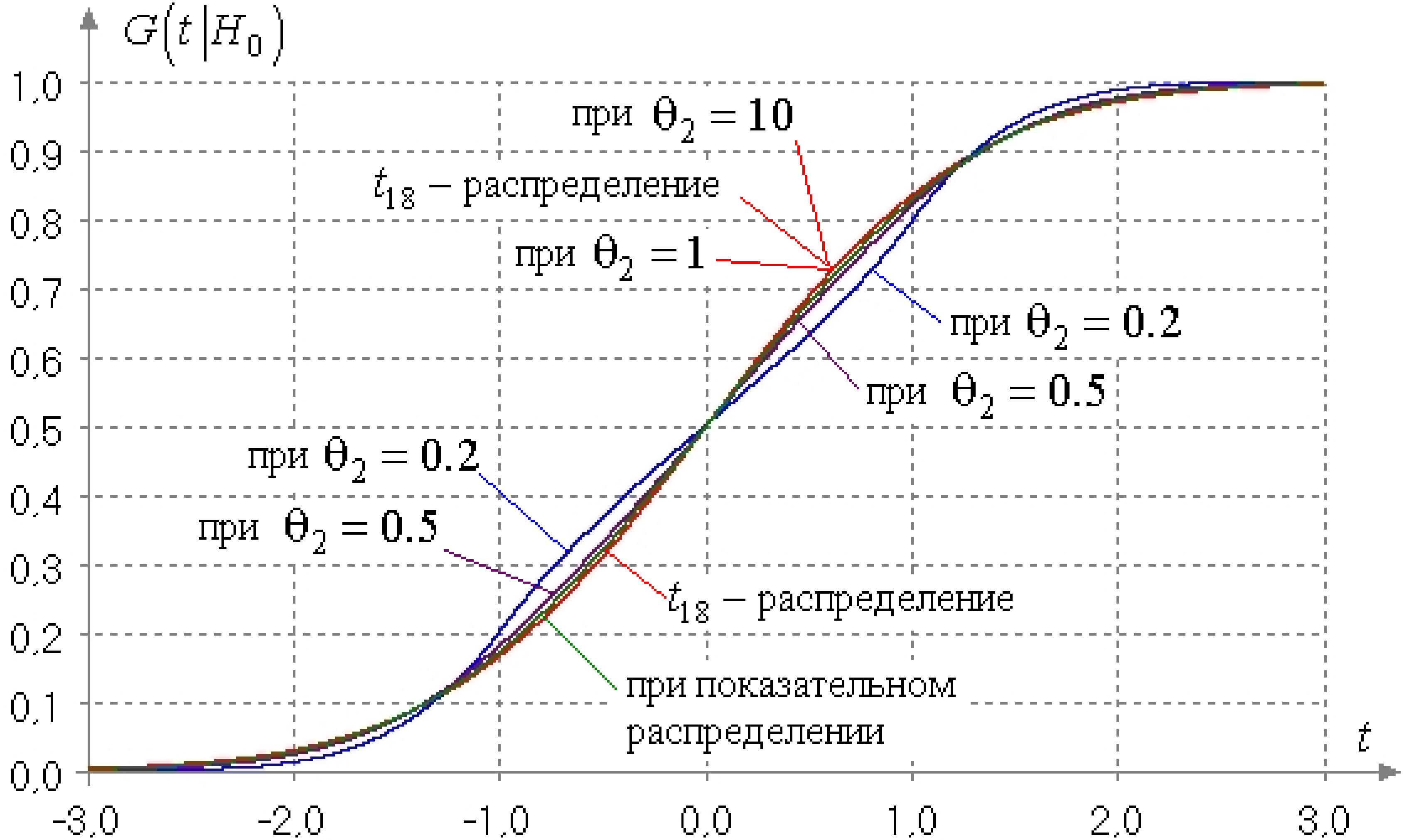


Рис. 3.3. Эмпирические распределения статистики (3.2) при различных законах распределения наблюдаемых величин и справедливости гипотезы H_0 при объемах выборок $n_1 = n_2 = 10$

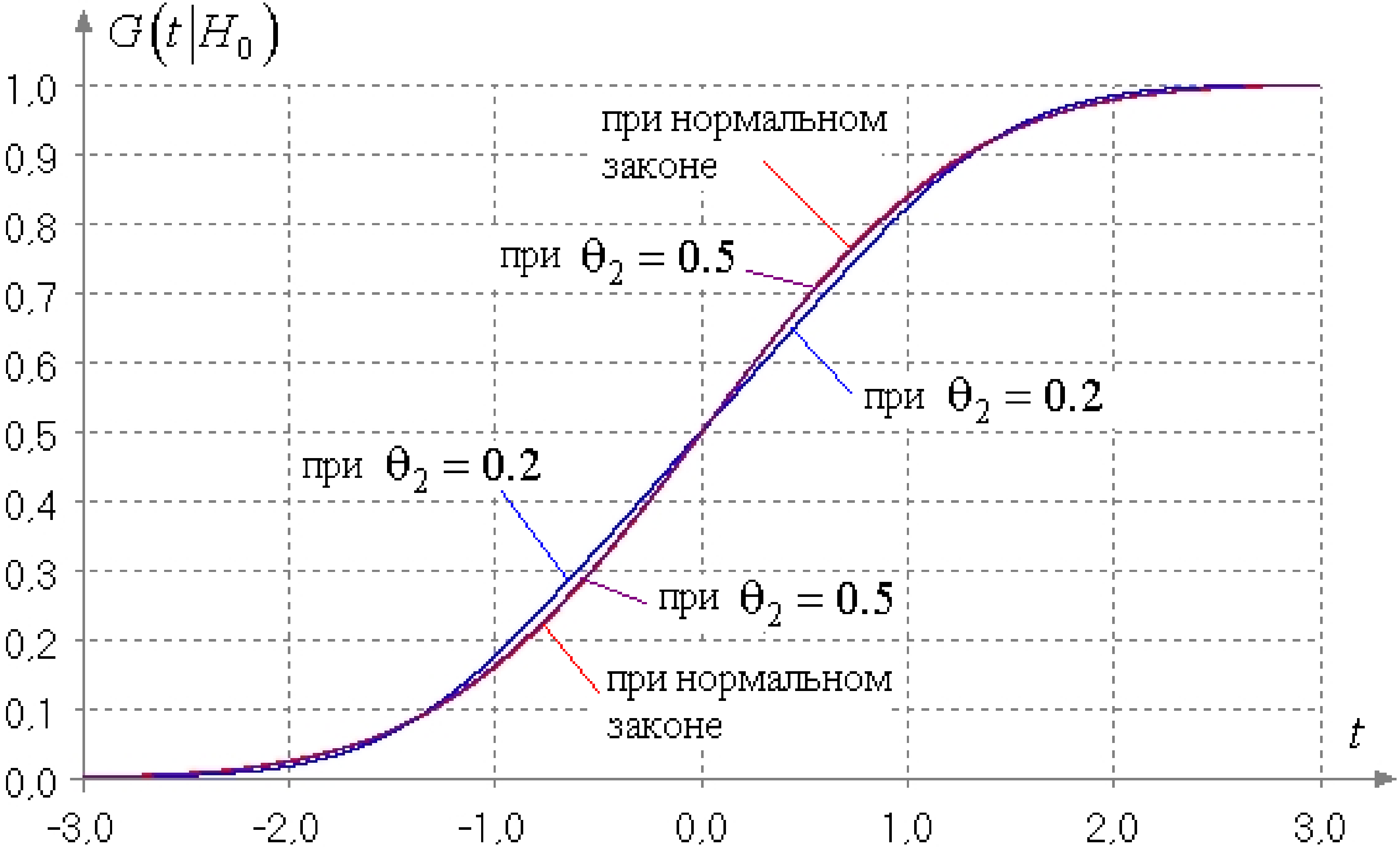


Рис. 3.4. Эмпирические распределения статистики (3.2) при различных законах распределения наблюдаемых величии и справедливости гипотезы H_0 при объемах выборок $n_1 = n_2 = 100$

Критерий Манна и Уитни

Ранговый критерий Манна и Уитни основан на критерии Уилкоксона для независимых выборок. Он является непараметрическим аналогом *t*-критерия для сравнения двух средних значений непрерывных распределений. Для вычисления статистики упорядочивают n_1+n_2 значений объединенной выборки, определяют сумму рангов R_1, соответствующую элементам первой выборки, и сумму рангов второй R_2. Вычисляются

$$U_1 = n_1 n_2 + \frac{n_1(n_1-1)}{2} - R_1,$$

$$U_2 = n_1 n_2 + \frac{n_2(n_2-1)}{2} - R_2.$$

Статистика критерия имеет вид

$$U = \min\{U_1, U_2\}.$$

Вместо *U*-статистики удобнее использовать статистику

$$\tilde{z} = \frac{U - \frac{n_1 n_2}{2}}{\sqrt{\frac{n_1 n_2 (n_1+n_2+1)}{12}}},$$

дискретное распределение которой в случае справедливости проверяемой гипотезы H_0 при $n_1+n_2>60$ хорошо приближается стандартным нормальным законом, когда объем каждой из выборок не слишком мал $n_1 \geq 8$, $n_2 \geq 8$. При меньших объемах выборок следует учитывать, что достигаемый уровень значимости (*p*-значение), вычисляемый по значению статистики в соответствии с функцией распределения стандартного нормального закона, может заметно отличаться от истинного.

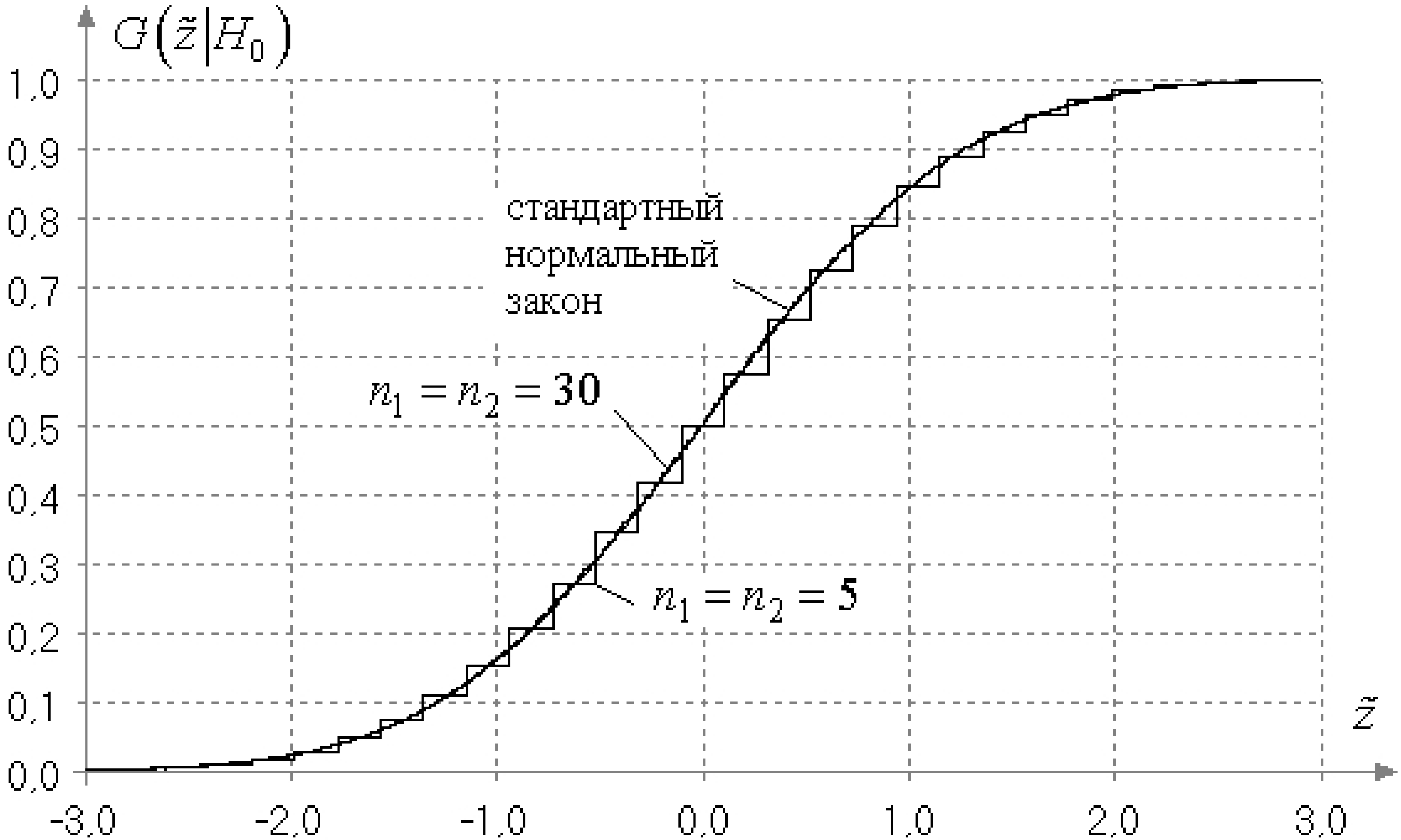


Рис. 3.6. Сходимость распределения статистики Манна–Уитни к стандартному нормальному закону

Критерий Краскела–Уаллиса

H –критерий Краскела–Уаллиса [10,11] является развитием U –критерия для проверки гипотезы о равенстве средних по k выборкам. Объединенную выборку $n=\sum_{i=1}^{k} n_i$ упорядочивают и вычисляют суммы рангов R_i для i-й выборки, $i=\overline{1,k}$. Статистика для проверки гипотезы H_0 имеет вид

$$H=\left[\frac{12}{n(n+1)}\right]\left[\sum_{i=1}^{k}\frac{R_i^2}{n_i}\right]-3(n+1).$$

H представляет собой дисперсию ранговых сумм. При больших n_i и k в случае справедливости проверяемой гипотезы H_0 статистика подчиняется χ^2_{k-1}-распределению. В описаниях критерия говорится, что χ^2_{k-1}-распределением практически можно пользоваться при $n_i \geq 5,\ k \geq 4$.

В действительности же при $k=2$ дискретностью можно практически пренебречь при $n_i \geq 30$. С ростом числа выборок влияние дискретности быстро убывает. При $k=3$ распределение статистики достаточно хорошо приближается χ^2_{k-1}-распределением, начиная с $n_i \geq 20$, а при $n_i \geq 30$ согласие распределения статистики с χ^2_{k-1}-распределением не отклоняется по всем применяемым критериям согласия. При $k \geq 5$ согласие распределения статистики с χ^2_{k-1}-распределением не отклоняется при $n_i \geq 20$.

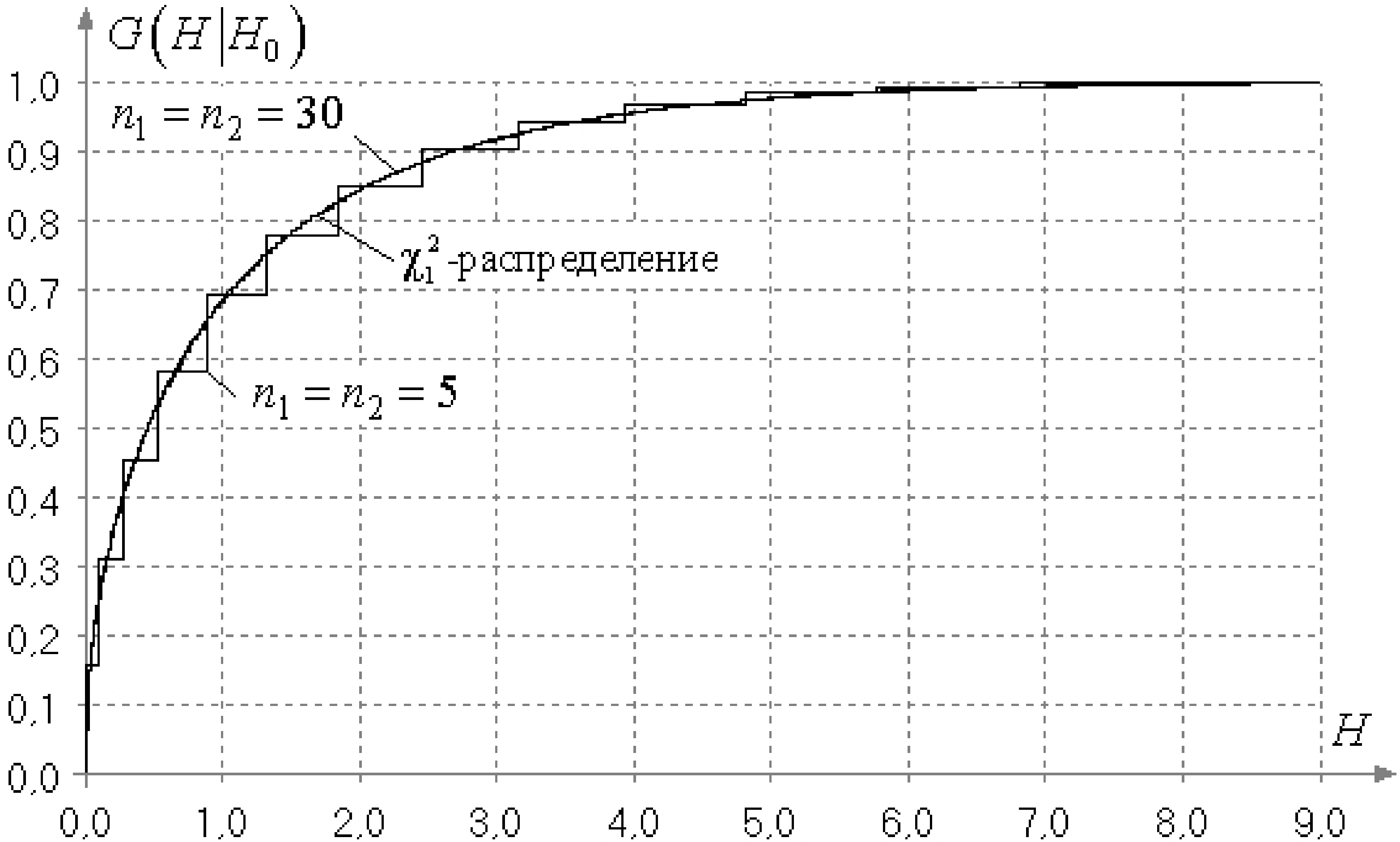


Рис. 3.7. Сходимость распределения статистики (3.9) Краскела–Уаллиса к χ^2_{k-1}-распределению при $k = 2$

Критерий Ван дер Вардена

Критерий предназначен для анализа 2-х выборок. Статистика непараметрического критерия Ван дер Вардена вычисляется в соответствии с выражением:

$$V = \sum_{i=1}^{n_2} u_{R_i/(n_1+n_2+1)}, \qquad (3.10)$$

где u_γ – γ-квантиль стандартного нормального закона, R_i, $i = \overline{1, n_2}$ – ранг i-го наблюдения, например, как в (3.10), второй выборки в общем вариационном ряду объединенной выборки из $n_1 + n_2$ наблюдений.

Считается [**Ошибка! Источник ссылки не найден.**], что при $n_1, n_2 \geq 20$ распределение статистики (3.10) удовлетворительно описывается нормальным законом с нулевым математическим ожиданием и дисперсией

$$D[V] = \frac{n_1 n_2}{(n_1+n_2)(n_1+n_2-1)} \sum_{i=1}^{n_1+n_2} u_{i/(n_1+n_2+1)}^2 .$$

Нормализованная статистика

$$V^* = \frac{V}{\sqrt{D[V]}} \qquad (3.11)$$

должна подчиняться стандартному нормальному закону.

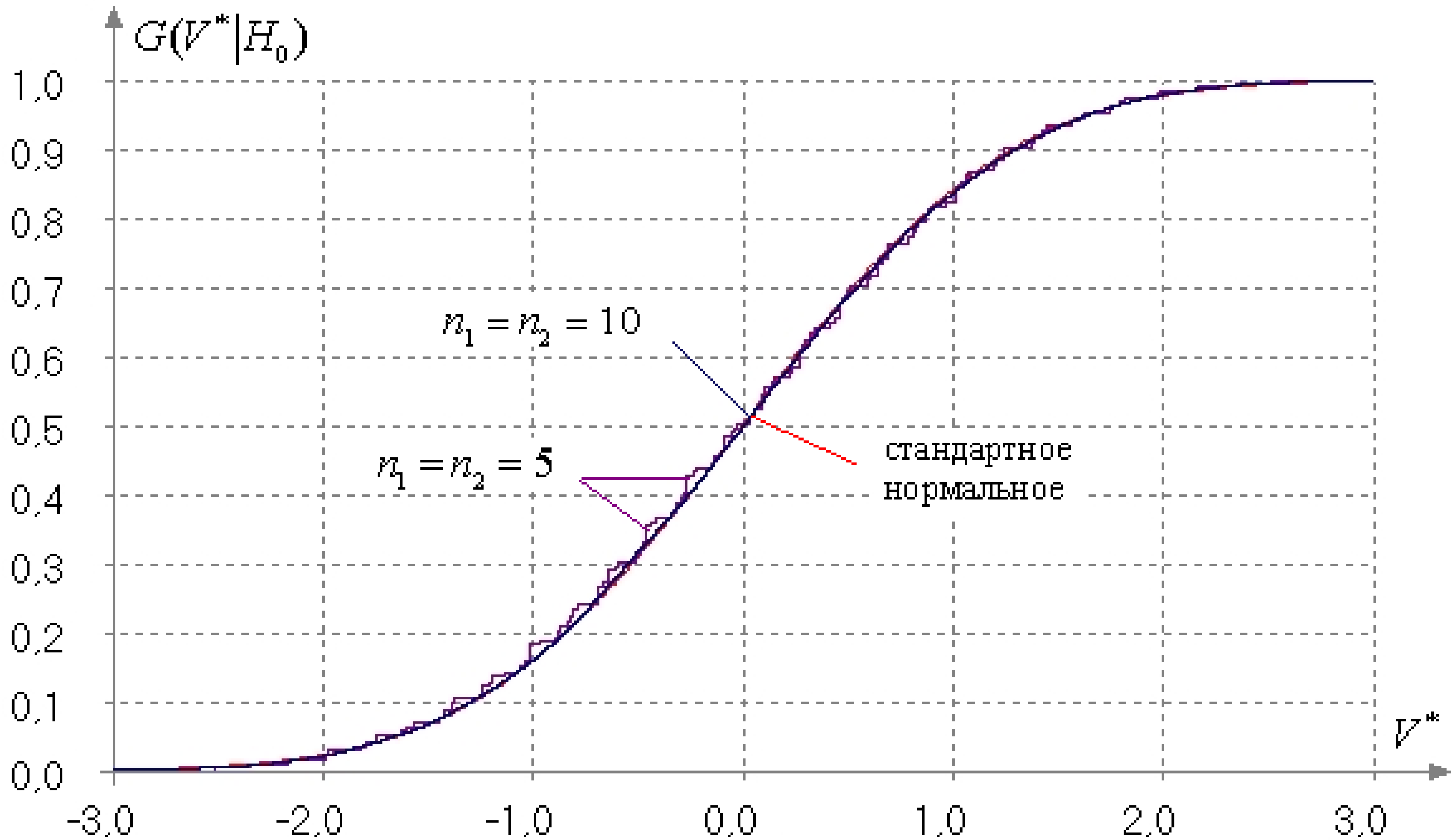


Рис. 3.8. Сходимость распределения нормализованной статистики (3.11) критерия Вандер Вардена к стандартному нормальному закону

Критерий Фишера–Йэйтса–Терри–Гёфдинга

тот критерий очень близок к критерию Ван дер Вардена. В качестве меток в критерии выбраны математические ожидания соответствующих порядковых статистик в выборке объёмом $n = n_1 + n_2$ из стандартного нормального закона. Статистика критерия имеет вид:

$$S = \sum_{i=1}^{n_2} u_{(R_i - 3/8)/(n+1/4)}, \tag{3.12}$$

где u_γ – γ-квантиль стандартного нормального закона, R_i, $i = \overline{1, n_2}$ – ранг i-го наблюдения, например, второй выборки в общем вариационном ряду объединенной выборки из $n_1 + n_2$ наблюдений.

Так же как и в случае критерия Ван дер Вардена, статистика (3.12) достаточно хорошо описывается нормальным законом с нулевым математическим ожиданием и дисперсией

$$D[S] = \frac{n_1 n_2}{(n_1 + n_2)(n_1 + n_2 - 1)} \sum_{i=1}^{n_1+n_2} u^2_{(i-3/8)/(n_1+n_2+1/4)},$$

а нормализованная статистика

$$S^* = \frac{S}{\sqrt{D[S]}} \tag{3.13}$$

– стандартным нормальным законом.

Сравнительный анализ мощности критериев

Мощность исследуемых критериев сравнивалась при одинаковых дисперсиях анализируемых выборок относительно следующих конкурирующих гипотез:

$$H_1^1:\ \mu_2 = \mu_1 + 0.1\sigma;$$

$$H_1^2:\ \mu_2 = \mu_1 + 0.2\sigma;$$

$$H_1^3:\ \mu_2 = \mu_1 + 0.3\sigma;$$

$$H_1^4:\ \mu_2 = \mu_1 + 0.4\sigma;$$

$$H_1^5:\ \mu_2 = \mu_1 + 0.5\sigma;$$

$$H_1^6:\ \mu_2 = \mu_1 + \sigma.$$

Выборки распределений статистик моделировались объемом $N = 10^6$, что позволило оценивать значения мощности с погрешностью в пределах $\pm 10^{-3}$.

Table 1. **Power of Tests Relative to the Alternative $H_1^1: \mu_2 = \mu_1 + 0.1\sigma$**

α	n = 10	n = 20	n = 30	n = 50	n = 100
z - test with known variances					
0.1	0.108	0.117	0.126	0.142	0.184
Student's t test with unknown and equal variances					
0.1	0.108	0.117	0.126	0.141	0.183
Student's t test with unknown and unequal variances					
0.1	0.108	0.117	0.126	0.141	0.183
F- test					
0.1	0.108	0.116	0.125	0.141	0.183
Van der Waerden test					
0.1	0.108	0.116	0.125	0.141	0.183
k-sampling Van der Waerden test					
0.1	0.108	0.116	0.125	0.141	0.182
Kruskal–Wallis H-test					
0.1	0.113	0.118	0.126	0.140	0.179
Mann–Whitney $\tilde{z}$ -test					
0.1	0.102	0.113	0.124	0.138	0.179

Table 4. **Power of Tests Relative to the Alternative $H_1^3:\ \mu_2 = \mu_1 + 0.4\sigma$**

α	$n = 10$	$n = 20$	$n = 30$	$n = 50$	$n = 100$
z - test with known variances					
0.1	0.232	0.355	0.462	0.638	0.881
Student's *t* test with unknown and equal variances					
0.1	0.222	0.348	0.456	0.633	0.879
Student's *t* test with unknown and unequal variances					
0.1	0.222	0.348	0.456	0.633	0.879
F- test					
0.1	0.222	0.346	0.455	0.633	0.879
Van der Waerden test					
0.1	0.218	0.339	0.449	0.628	0.877
k-sampling Van der Waerden test					
0.1	0.217	0.340	0.449	0.628	0.877
Kruskal–Wallis *H*-test					
0.1	0.224	0.338	0.445	0.616	0.865
Mann–Whitney $\tilde{z}$ -test					
0.1	0.201	0.327	0.439	0.613	0.865

Table 6. **Power of Tests Relative to the Alternative $H_1^6: \mu_2 = \mu_1 + \sigma$**

α	n = 10	n = 20	n = 30	n = 50	n = 100
z - test with known variances					
0.1	0.722	0.935	0.987	1.000	1.000
Student's t test with unknown and equal variances					
0.1	0.693	0.928	0.985	1.000	1.000
Student's t test with unknown and unequal variances					
0.1	0.693	0.928	0.985	1.000	1.000
F-test					
0.1	0.692	0.927	0.985	1.000	1.000
Van der Waerden test					
0.1	0.677	0.922	0.984	0.999	1.000
k-sampling Van der Waerden test					
0.1	0.676	0.921	0.984	0.999	1.000
Kruskal–Wallis H-test					
0.1	0.680	0.917	0.982	0.999	1.000
Mann–Whitney $\tilde{z}$ -test					
0.1	0.645	0.912	0.981	0.999	1.000

Оценки мощности k -выборочных критериев

Table 7. **Power of Tests Relative to the Alternative** $H_1^2: \mu_k = \mu_1 + 0.2\sigma$

Test	α	*k* = 3		*k* = 5	
		n = 30	*n* = 50	*n* = 30	*n* = 50
F-test	0.1	0.191	0.254	0.175	0.232
k-sampling Van der Waerden test	0.1	0.190	0.252	0.174	0.231
Kruskal–Wallis *H*-test	0.1	0.187	0.247	0.172	0.225

Table 8. **Power of Tests Relative to the Alternative** $H_1^5: \mu_k = \mu_1 + 0.5\sigma$

Test	α	*k* = 3		*k* = 5	
		n = 30	*n* = 50	*n* = 30	*n* = 50
F- test	0.1	0.618	0.823	0.585	0.807
k-sampling Van der Waerden test	0.1	0.612	0.819	0.581	0.804
Kruskal–Wallis *H*-test	0.1	0.598	0.806	0.564	0.786

Выводы

1. Параметрические критерии проверки однородности средних устойчивы к отклонениям законов распределения наблюдаемых случайных величин от нормального.
2. Отличия распределений статистик от “классических” становятся заметными в случае асимметричности наблюдаемых законов, но эти отклонения не очень значительны.
3. Существенные отклонения распределений статистик от “классических” наблюдаются в случае законов с тяжелыми хвостами.
4. Мощность параметрических критериев выше непараметрических. Следовательно, их применение в свете устойчивости к отклонениям от классических предположений о нормальности является предпочтительным.

О требуемых объемах выборок

На практике, применяя рассматриваемые критерии для проверки гипотезы об однородности математических ожиданий, **как правило, задаются только вероятностью** α **ошибки 1-го рода**. Процедуры контроля предусматривают, чаще всего, небольшие объемы выборок.

Как правило, **не заходит речи о вероятности** β **ошибки 2-го рода**: не отклонить проверяемую гипотезу при справедливости конкурирующей. В то же время в процедуре контроля **при задании** α **желательно гарантировать величину** $\beta \le \alpha$.

В данном случае относительно конкурирующей гипотезы $H_1^1 : \mu_2 = \mu_1 + 0{,}1\sigma$ для $\alpha = 0{,}1$ и объемах выборок $n = 100$ **вероятность ошибки 2-го рода составит величину** $\beta = 1 - 0{,}283 = 0{,}717$ **(!)** для z-критерия.

При $\alpha = 0{,}1$ и объемах выборок $n = 100$ данный критерий обеспечит величину $\beta = 0{,}061 < 0{,}1$ только для более далекой конкурирующей гипотезы $H_1^4 : \mu_2 = \mu_1 + 0{,}4\sigma$.

А **чтобы с заданным качеством различать гипотезы** H_0 **и** H_1^1 **необходимо иметь выборки объемом порядка 1350 наблюдений!**

Что можем различать при заданных *n*?

А какие пары конкурирующих гипотез можно с таким же качеством (при $\alpha \le 0{,}1$ и $\beta \le 0{,}1$) различать по выборкам объемом $n = 10$?

При $n = 10$ – конкурирующие гипотезы, когда μ_2 отличается от μ_1 на величину не менее, чем $1{,}15\sigma$!

При $n = 20$ – когда μ_2 отличается от μ_1 на величину порядка $0{,}82\sigma$!

При $n = 30$ – на величину порядка $0{,}67\sigma$!

При $n = 50$ – на величину порядка $0{,}51\sigma$!

При $n = 100$ – на величину порядка $0{,}364\sigma$!

Критерии проверки гипотез об однородности дисперсий

при нарушении классических предположений о принадлежности наблюдений (ошибок измерений) нормальному закону

В критериях проверки однородности дисперсий проверяемая гипотеза о постоянстве дисперсий k выборок имеет вид

$$H_0: \sigma_1^2 = \sigma_2^2 = \ldots = \sigma_k^2, \tag{4.1}$$

а конкурирующая с ней гипотеза

$$H_1: \sigma_{i_1}^2 \neq \sigma_{i_2}^2, \tag{4.2}$$

где неравенство выполняется, по крайней мере, для одной пары индексов i_1, i_2.

Параметрические критерии:

Бартлетта, Кокрена, Хартли, Фишера, Неймана–Пирсона, О'Брайена, Z–критерий Оверолла–Вудворда, модифицированный Z–критерий, Левене, Блисса–Кокрена–Тьюки, Кадуэлла–Лесли–Брауна, Линка, Ньюмана.

Непараметрические критерии:

Ансари–Бредли, Сижела–Тьюки, Муда, Клотца, Кейпена, Флайне–Киллина

Критерий Бартлетта

Статистика критерия Бартлетта вычисляется в соответствии с соотношением [2]:

$$\chi^2 = M\left[1 + \frac{1}{3(m-1)}\left(\sum_{i=1}^{m}\frac{1}{\nu_i} - \frac{1}{N}\right)\right]^{-1} \qquad (5)$$

где n_i – объемы выборок, $\nu_i = n_i$, если математическое ожидание известно, и $\nu_i = n_i - 1$, если неизвестно, $N = \sum_{i=1}^{m}\nu_i$,

$$M = N\ln\left(\frac{1}{N}\sum_{i=1}^{m}\nu_i S_i^2\right) - \sum_{i=1}^{m}\nu_i \ln S_i^2,$$

S_i^2 – оценки выборочных дисперсий. При неизвестном математическом ожидании оценки $S_i^2 = \frac{1}{n_i - 1}\sum_{j=1}^{n_i}\left(X_{ji} - \bar{X}_i\right)^2$, где $\bar{X}_i = \frac{1}{n_i}\sum_{j=1}^{n_i}X_{ji}$ и $\nu_i = n_i - 1$. Если гипотеза H_0 верна, все $\nu_i > 3$ и выборки извлекаются из нормальной генеральной совокупности, то статистика (5) приближенно подчиняется χ^2_{m-1}-распределению.

При нормально распределенных наблюдаемых (измеряемых) величинах распределение статистики (5) практически не зависит от изменения объема выборки. Например, на рис. 1 приведены практически совпадающие функции распределения статистики критерия Бартлетта (5) при различных объемах выборок ($n = 10, 50, 100$). Это означает, что в случае нормальности наблюдаемых величин выводы остаются корректными и при очень малых объемах анализируемых выборок.

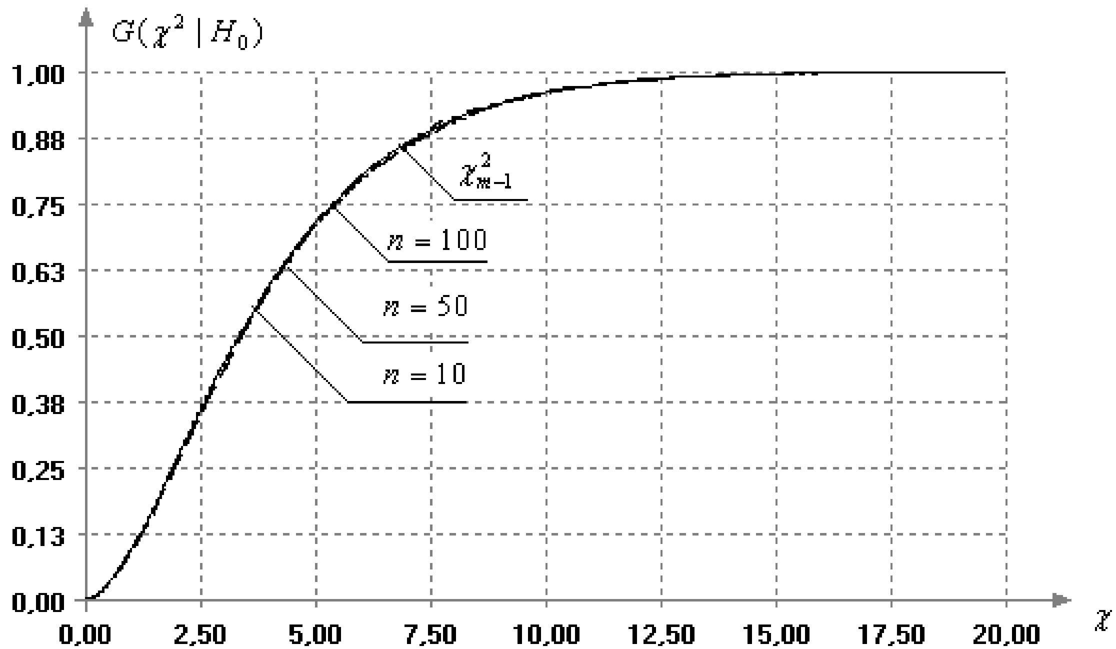


Рис. 1. Функции распределения статистики классического критерия Бартлетта при различных объемах выборок при $m = 5$

В то же время, распределения статистики (5) очень чувствительны к отклонениям наблюдаемого закона от нормального. Вид распределения статистики (5) исследовался при различных наблюдаемых законах, в частности, в случае принадлежности моделируемых выборок законам логистическому с плотностью

$$f(x) = \frac{\pi}{\theta_0\sqrt{3}}\exp\left\{-\frac{\pi(x-\theta_1)}{\theta_0\sqrt{3}}\right\}\Bigg/\left[1+\exp\left\{-\frac{\pi(x-\theta_1)}{\theta_0\sqrt{3}}\right\}\right]^2 ,$$

Лапласа с плотностью

$$f(x) = \frac{\theta_0}{2}\exp\left\{-\theta_0\left|x-\theta_1\right|\right\} ,$$

экспоненциальному семейству распределений с различными параметрами формы с плотностью

$$De(\lambda) = f(x) = \frac{\lambda}{2\sqrt{2}\theta_1\Gamma\left(1/\lambda\right)}\exp\left(-\left(\frac{\left|x-\theta_0\right|}{\sqrt{2}\theta_1}\right)^{\lambda}\right), \qquad (6)$$

где λ – параметр формы. Законы нормальный и Лапласа являются частными случаями данного семейства распределений при значениях параметра формы 2 и 1 соответственно. Семейство (6) может быть хорошей моделью для законов распределения ошибок различных измерительных систем.

Рис. 2 отражает зависимость распределений статистики (5) от вида наблюдаемого закона при различных объемах выборок. Видно, что при отклонении закона распределения наблюдаемого показателя от нормального закона распределение статистики критерия Бартлетта (5) существенно отличается от χ^2_{m-1}-распределения. При этом распределения статистики становятся более зависимыми от объема выборки, чем в случае нормального закона.

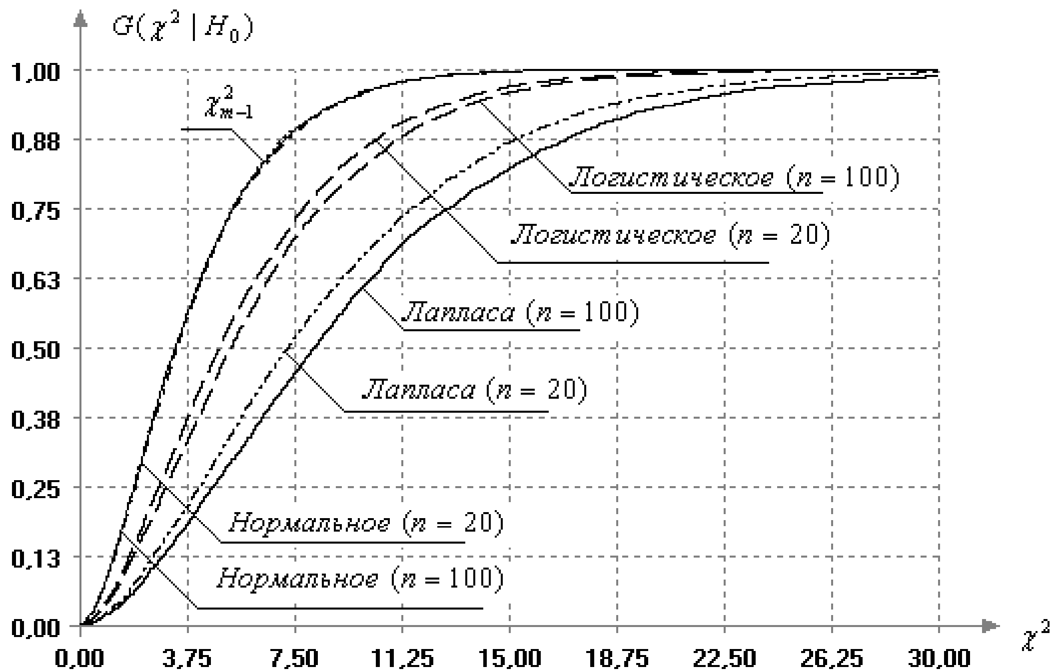


Рис. 2. Функции распределения статистики Бартлетта при отклонении закона распределения наблюдаемого показателя от нормального при различных объемах выборки и $m = 5$

На рис. 3 показано, как меняется распределение статистики Бартлетта, если наблюдаемая случайная величина подчиняется экспоненциальному семейству распределений с различными значениями параметра формы.

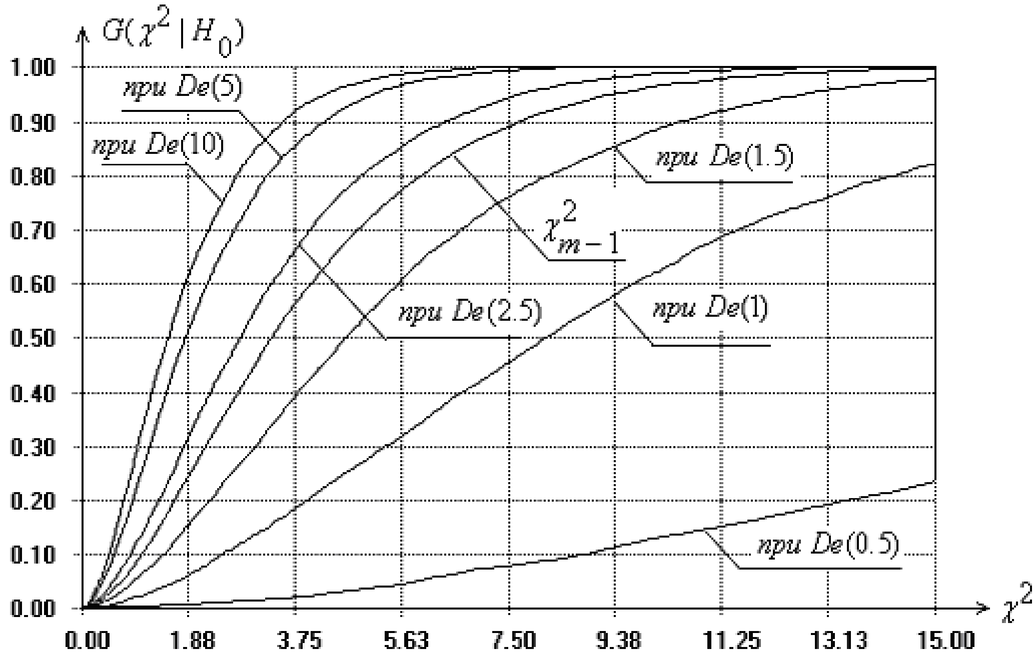


Рис. 3. Функции распределения статистики критерия Бартлетта в случае распределений экспоненциального семейства $De(\lambda)$ с различными значениями параметра формы при $n = 100$ и $m = 5$

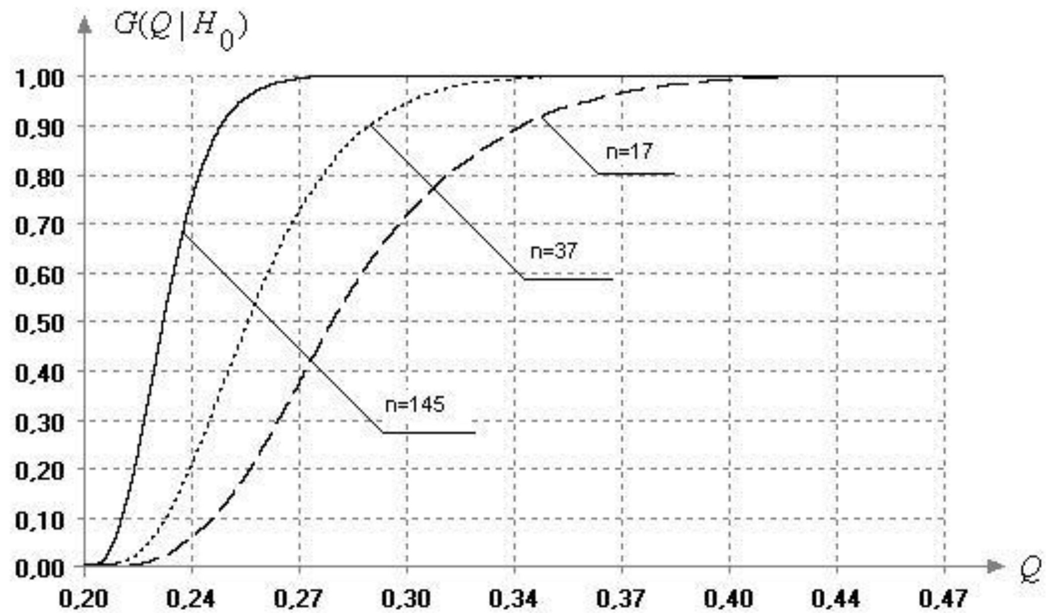
Критерий Кокрена

В том случае, когда все n_i одинаковы, $n_1 = n_2 = \ldots = n_m = n$, возможно использование более простого критерия Кохрена. Статистика Q критерия Кокрена выражается формулой [2]

$$Q = \frac{S_{\max}^2}{S_1^2 + S_2^2 + \cdots + S_m^2}, \qquad (7)$$

где $S_{\max}^2 = \max\left(S_1^2, S_2^2, \ldots, S_m^2\right)$, где m – число независимых оценок дисперсий (число выборок).

Распределения статистики Кокрена сильно зависят от объема наблюдаемых выборок. Поэтому в справочной литературе приводятся только таблицы процентных точек.



Распределения статистики критерия Кокрена при различных объемах выборок при $m = 5$

Как и критерий Бартлетта, критерий Кокрена используется в предположении, что наблюдаемая случайная величина принадлежит нормальному закону. Насколько сильно меняется распределение статистики Кокрена (7) в случае определенных отклонений закона распределения наблюдаемой случайной величины (контролируемого показателя) от нормального показывает следующий рисунок.

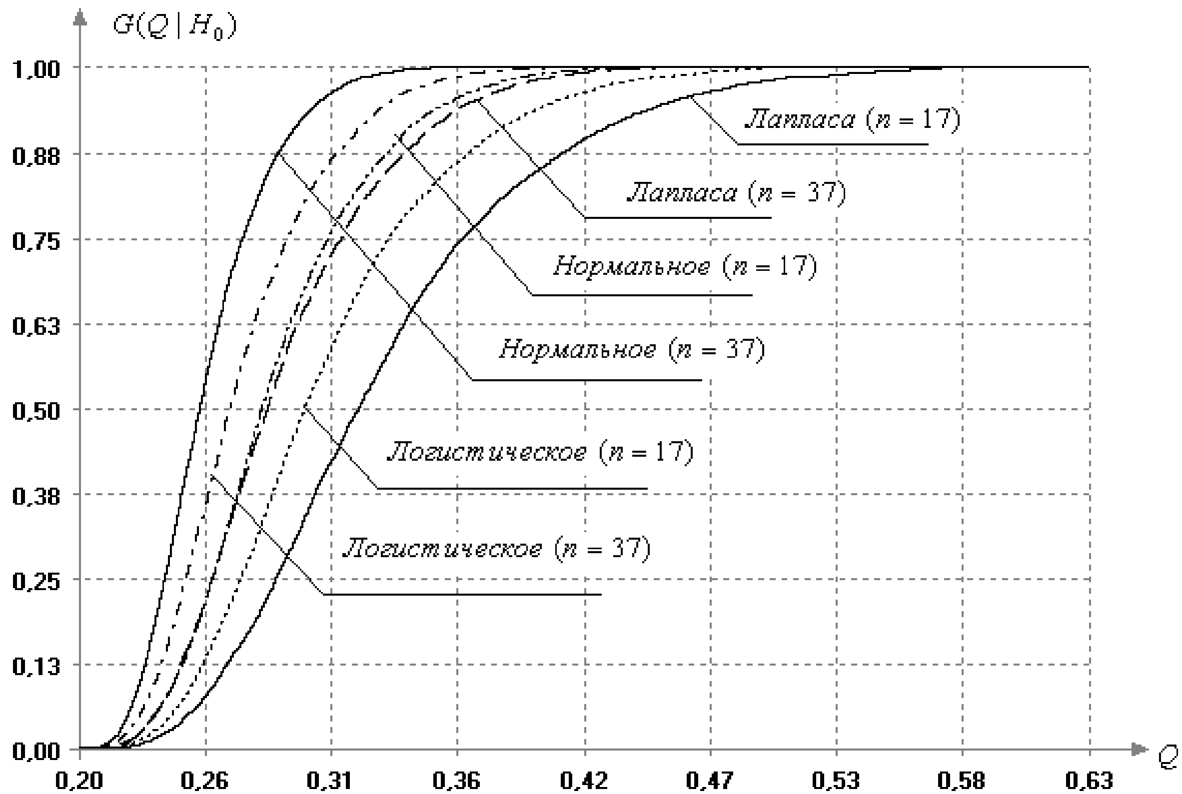
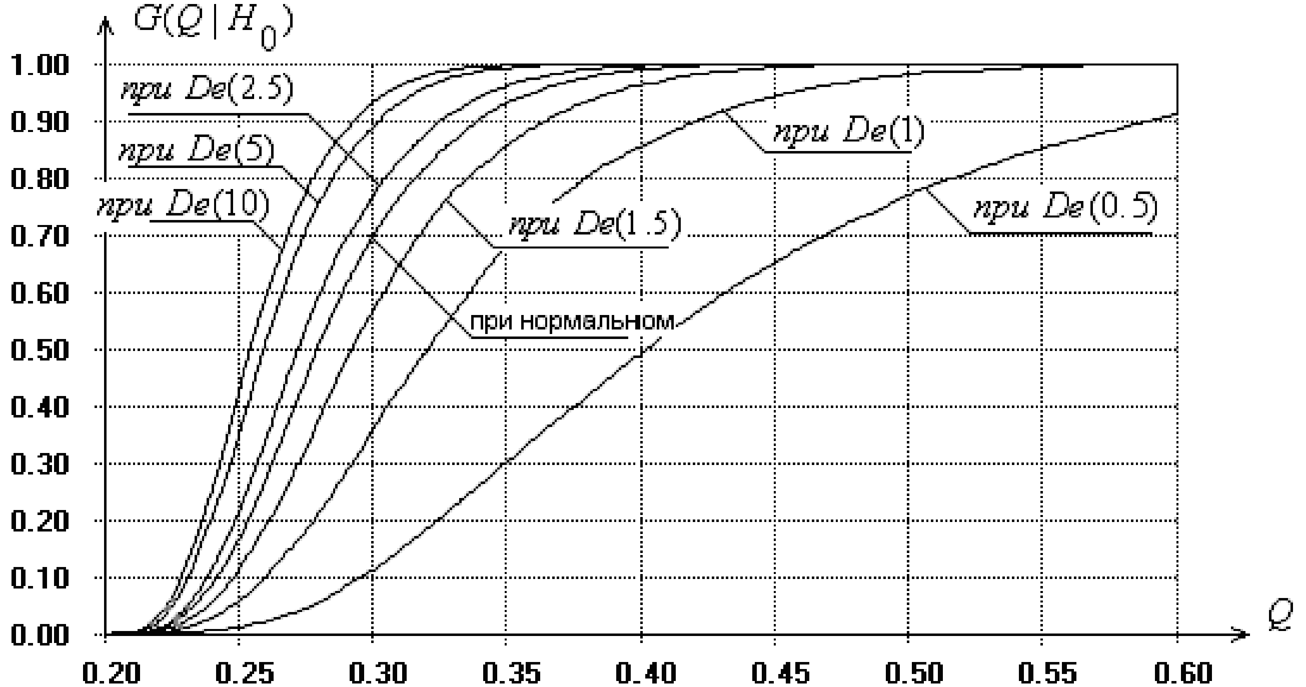


Рис. 5. Функции распределения статистики критерия Кокрена при отклонении закона распределения наблюдаемого показателя от нормального при различных объемах выборки при $m = 5$



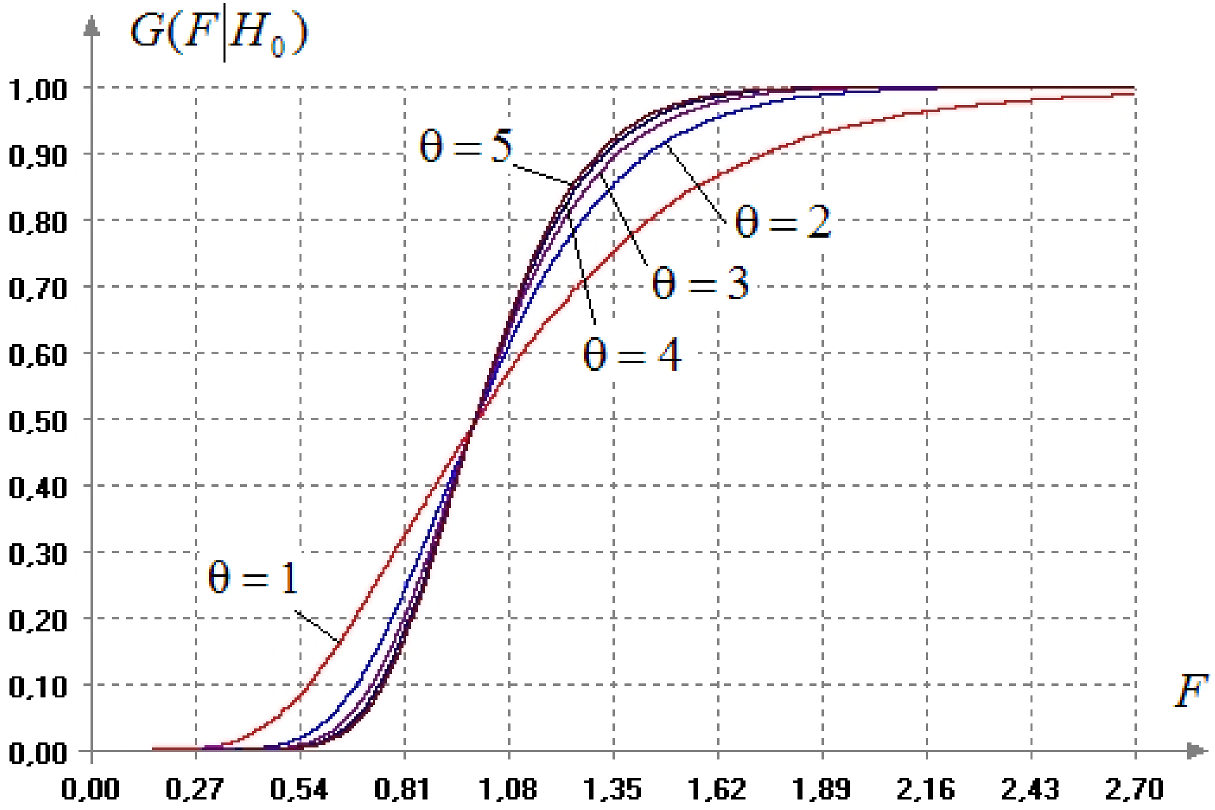
Функции распределения статистики критерия Кокрена в случае распределений экспоненциального семейства с различными значениями параметра формы при $n = 17$ и $m = 5$

Критерий Фишера

Для сравнения двух выборочных дисперсий из нормальных совокупностей. Для определения того, относятся ли две выборки к одной и той же генеральной совокупности, проверяется гипотеза вида $H_0 : \sigma_1^2 = \sigma_2^2$. Статистика для проверки гипотезы имеет вид

$$F = \frac{s_1^2}{s_2^2}.$$

В случае принадлежности выборок нормальному закону и справедливости H_0 эта статистика подчинятся F_{ν_1,ν_2} -распределению Фишера с числом степеней свободы $\nu_1 = n_1 - 1$ и $\nu_2 = n_2 - 1$.



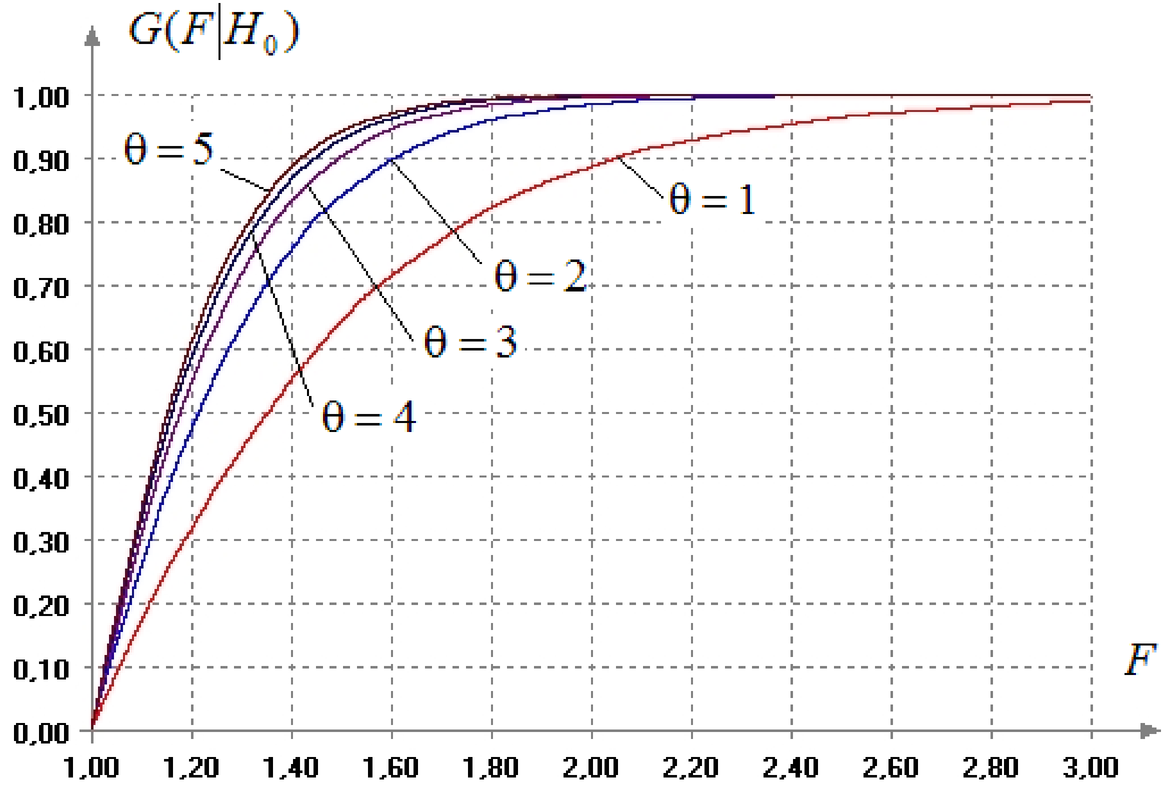
Распределения статистики Фишера при справедливой гипотезе H_0 объеме 2-х выборок по $n = 50$ в случае принадлежности выборок различным законам из экспоненциального семейства: $\theta = 1$, $\theta = 2$ (нормальный), $\theta = 3$, $\theta = 4$, $\theta = 5$.

Критерий Хартли

Проверка равенства нескольких дисперсий для выборок равного объема по Хартли. Статистика для проверки гипотезы имеет вид

$$F = \frac{s_{\max}^2}{s_{\min}^2}.$$

Степенями свободы для распределения статистики являются число выборок $\nu_1 = m$ и $\nu_2 = n - 1$. В литературе для статистики приводятся лишь таблицы процентных точек.



Распределения статистики Хартли при справедливой гипотезе H_0 объеме 2-х выборок по $n = 50$ в случае принадлежности выборок различным законам из экспоненциального семейства: $\theta = 1$, $\theta = 2$ (нормальный), $\theta = 3$, $\theta = 4$, $\theta = 5$.

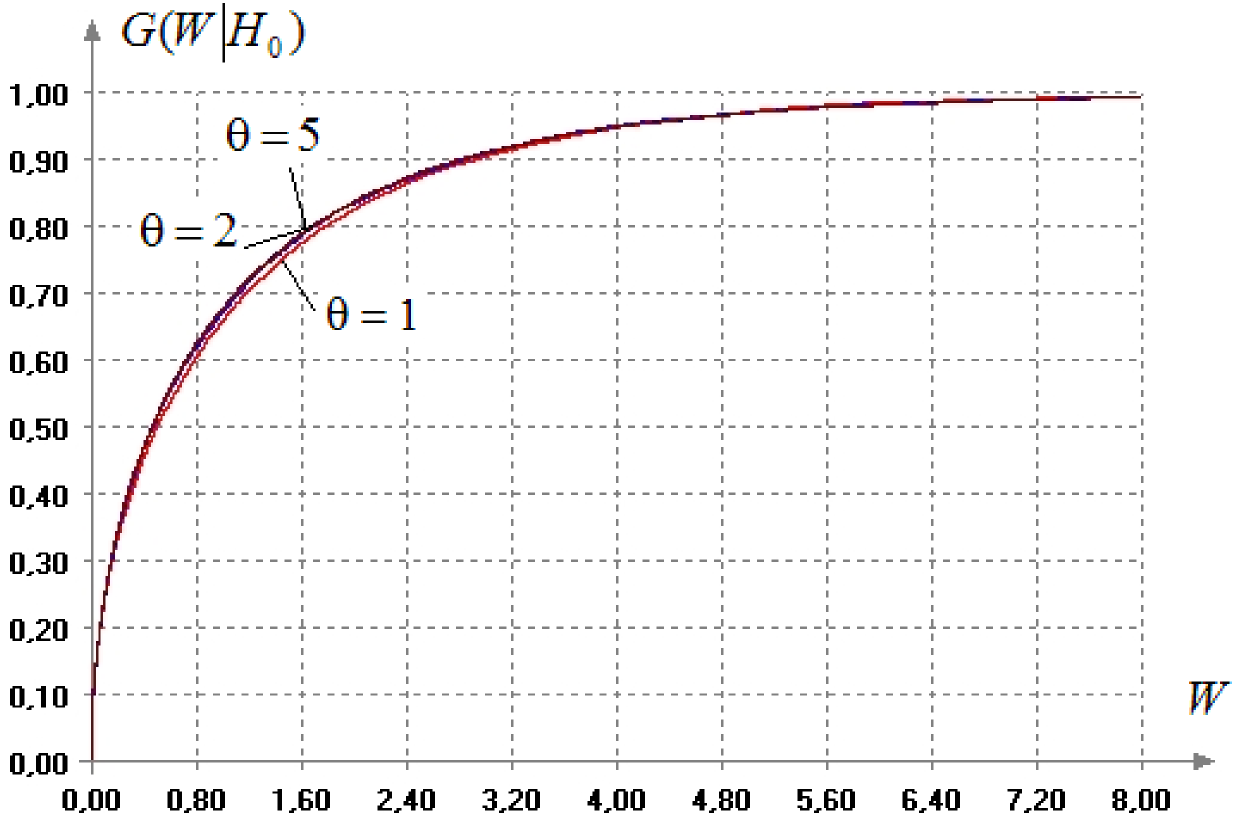
Критерий Левене

Считается, что критерий Левене менее чувствителен к отклонениям от нормальности, чем критерий Бартлетта.

Пусть n_i − объем i-й выборки, $N=\sum_{i=1}^{m} n_i$, X_{ij} − j-е наблюдение в i-й выборке. Статистика критерия Левене имеет вид:

$$W=\frac{N-m}{m-1}\frac{\sum_{i=1}^{m} n_i\left(\bar{Z}_{i\bullet}-\bar{Z}_{\bullet\bullet}\right)^2}{\sum_{i=1}^{m}\sum_{j=1}^{n_i}\left(Z_{ij}-\bar{Z}_{i\bullet}\right)^2},$$

где Z_{ij} определяется выражением $Z_{ij}=\left|X_{ij}-\bar{X}_{i\bullet}\right|$, в котором $\bar{X}_{i\bullet}$ − среднее в i-й выборке. $\bar{Z}_{i\bullet}$ − среднее Z_{ij} по i-й выборке, $\bar{Z}_{\bullet\bullet}$ − среднее Z_{ij} по всем выборкам. В случае принадлежности выборок нормальному закону и справедливости H_0 эта статистика подчинятся F_{ν_1,ν_2} - распределению Фишера с числом степеней свободы $\nu_1=m-1$ и $\nu_2=N-m$.



Распределения статистики Левене при справедливой гипотезе H_0 объеме 2-х выборок по $n = 50$ в случае принадлежности выборок различным законам из экспоненциального семейства: $\theta = 1$, $\theta = 2$ (нормальный), $\theta = 3$, $\theta = 4$, $\theta = 5$.

Как видим, критерий Левене оказывается устойчивым к нарушению предположения нормальности.

В оригинальном критерии Левене предусмотрено использование только выборочных средних. Brown и Forsythe расширили критерий Левене на случай использования выборочных медиан и усеченного среднего ($Z_{ij} = \left| X_{ij} - \tilde{X}_{i\bullet} \right|$, где $\tilde{X}_{i\bullet}$ − медиана в i-й выборке; $Z_{ij} = \left| X_{ij} - \bar{X}'_{i\bullet} \right|$, где $\bar{X}'_{i\bullet}$ − усеченное среднее в i -й выборке). В этой работе говорится, что в этих случаях критерий становится устойчивей к нарушению предположений о нормальности. Однако возникает опасение, не приводит ли это к тому, что распределение статистики будет несколько отличаться от F_{ν_1,ν_2} -распределения? Проверка показала, что существенных отличий от F_{ν_1,ν_2} -распределения при использовании в качестве оценок среднего медианы и усеченного среднего не происходит.

Выводы

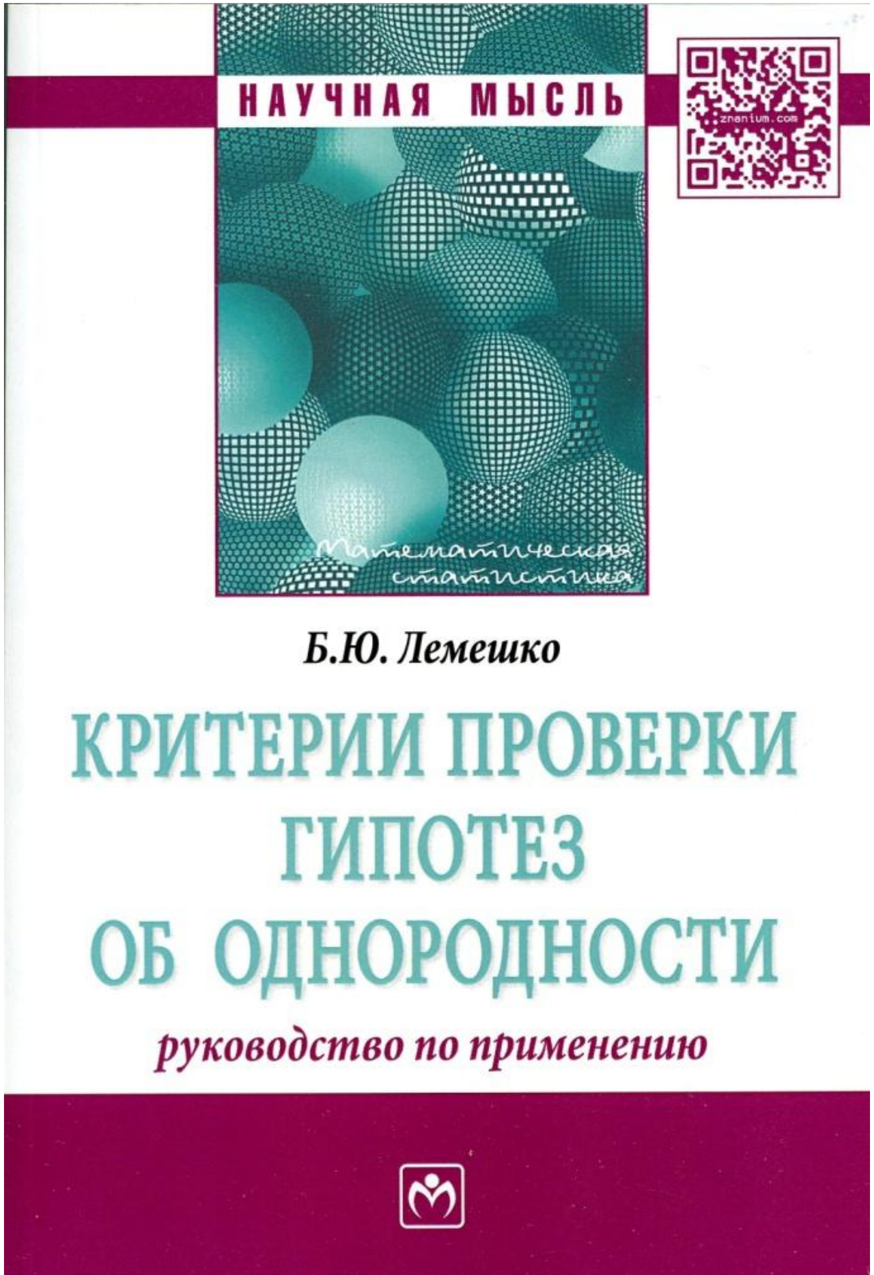
Критерии проверки однородности дисперсий, как правило, очень чувствительны к нарушению предположения о принадлежности наблюдений нормальному закону.

Исключение составляет критерий Левене. Этому есть объяснение. В критерии Левене проверка однородности дисперсий сводится к проверке «однородности средних».

Аналогичная ситуация характерна для F-критерия, применяемого для проверки гипотез об однородности средних.

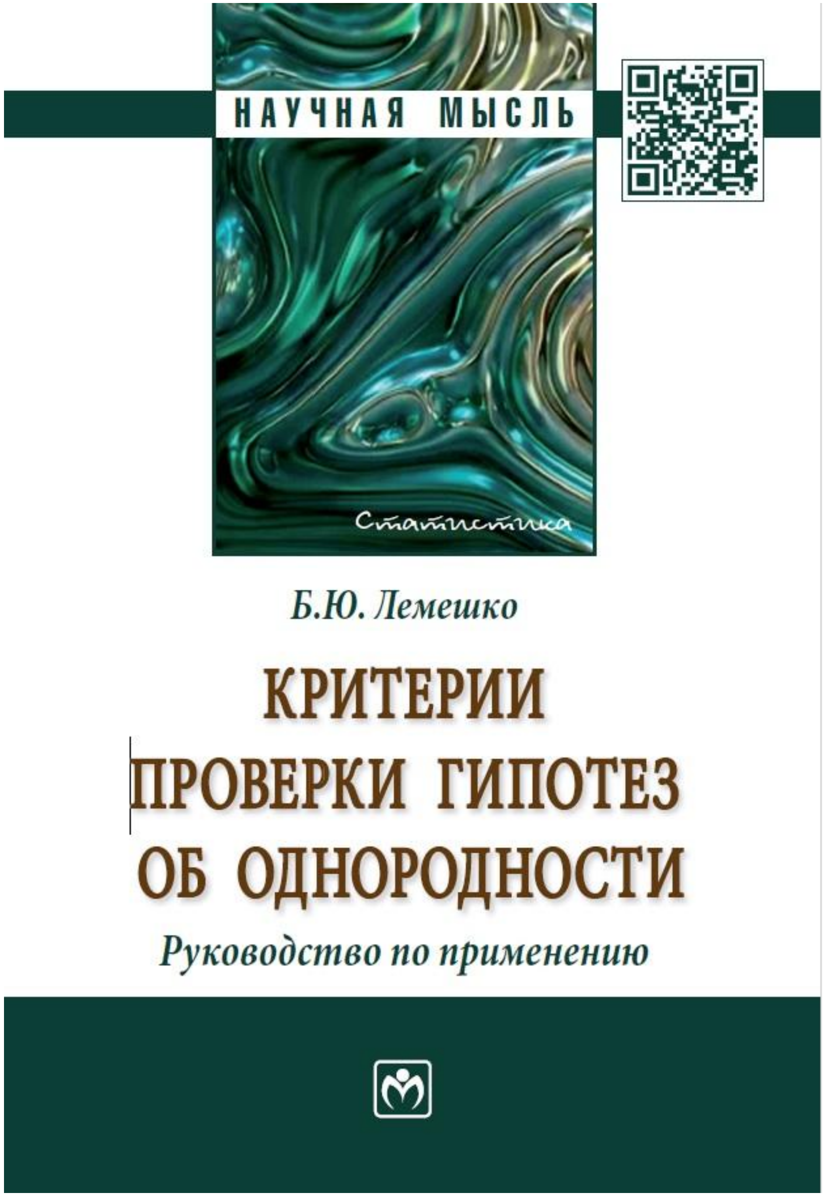
Критерий Шеффе, используемый в дисперсионном анализе в аналогичной ситуации проверки однородности средних, также обладает известной устойчивостью.

При необходимости можно исследовать распределения статистики любого критерия однородности дисперсий в случае принадлежности наблюдений любому закону распределения.



1. **Лемешко Б. Ю. Критерии проверки гипотез об однородности. Руководство по применению: Монография. – М.: ИНФРА-М, 2017. – 207 с**.

DOI: 10.12737/22368



2. **Лемешко Б. Ю. Критерии проверки гипотез об однородности. Руководство по применению: Монография. – 2-е изд., перераб. и доп. – М.: ИНФРА-М, 2021. – 248 с**.

DOI: 10.12737/986695