

## ВЕРОЯТНОСТНЫЕ МОДЕЛИ

УДК 519.2

### О ЧИСЛЕННОЙ АППРОКСИМАЦИИ МОЩНОСТИ КРИТЕРИЕВ В СЛУЧАЕ РАСПРЕДЕЛЕНИЯ ЛАПЛАСА

**Королев Р.А.**

кафедра математической статистики,  
факультет вычислительной математики и кибернетики МГУ, г. Москва  
(*stochastique@gmail.com*)

---

*Поступила в редакцию 29.01.2009, после переработки 17.02.2009.*

---

В работе приводятся численные результаты аппроксимации для мощностей асимптотически оптимального и наилучшего критериев в случае распределения Лапласа (см. [1], [2]). Для обоих критериев исследуется точность аппроксимации функций мощности их асимптотическими разложениями до порядка  $n^{-1}$  (см. [2], Теорема 2.1, Теорема 3.3). Численно исследуется поведение дефекта.

In the paper we obtain numerical results for the power of the asymptotically optimal test and for the power of the asymptotically most powerful test for the case of Laplace distribution (see [1], [2]). For both tests we investigate the numerical accuracy of approximations to power functions using their asymptotic expansions of the order  $n^{-1}$  (see [2], Theorem 2.1, Theorem 3.3). We numerically research deficiency of the asymptotically optimal test.

**Ключевые слова:** асимптотическое разложение, численная аппроксимация, функция мощности, распределение Лапласа, дефект.

**Keywords:** asymptotic expansion, numerical approximation, power function, Laplace or double exponential distribution, deficiency.

#### 1. Введение

Следуя работам [1], [2], рассмотрим задачу проверки гипотезы

$$H_0 : \theta = 0 \tag{1.1}$$

против последовательности простых близких альтернатив вида

$$H_{n,1} : \theta = \frac{t}{\sqrt{n}}, \quad 0 < t \leq C, \quad C > 0, \tag{1.2}$$

на основе выборки  $(X_1, \dots, X_n)$  – независимых одинаково распределенных наблюдений, имеющих распределение Лапласа с параметром сдвига вида

$$p(x, \theta) = \frac{1}{2} e^{-|x - \theta|}, \quad x, \theta \in \mathbb{R}^1. \tag{1.3}$$

Для каждого фиксированного  $t \in (0, C]$  обозначим через  $\beta_n^*(t)$  мощность наилучшего критерия размера  $\alpha \in (0, 1)$  против простой альтернативы  $H_{n,t} : \theta = t/\sqrt{n}$ . Такой критерий всегда существует согласно фундаментальной лемме Неймана-Пирсона (см., например, [3], стр. 94) и основан на логарифме отношения правдоподобия  $\Lambda_n(t)$  (см. формулу (1.4) ниже).

В работе [2] были получены асимптотические разложения для мощностей  $\beta_n^*(t)$  и  $\beta_n(t)$ , соответственно, наилучшего и асимптотически эффективного критериев уровней  $\alpha \in (0, 1)$ , основанных, соответственно, на статистиках

$$\Lambda_n(t) = \sum_{i=1}^n (|X_i| - |X_i - tn^{-1/2}|), \quad (1.4)$$

$$T_n = \frac{1}{\sqrt{n}} \sum_{i=1}^n \text{sign}(X_i). \quad (1.5)$$

Обозначим через  $\Phi(x)$ ,  $\varphi(x)$ , соответственно, функцию распределения и плотность стандартного нормального закона,  $\Phi(u_\alpha) = 1 - \alpha$ ,  $\alpha \in (0, 1)$ .

Асимптотические разложения для мощностей  $\beta_n^*(t)$  и  $\beta_n(t)$  представлены следующими формулами (см. [2], Теорема 2.1 и Теорема 3.3)

$$\begin{aligned} \beta_n^*(t) = & \Phi(t - u_\alpha) - \frac{t^2}{6\sqrt{n}}\varphi(t - u_\alpha) - \\ & - \frac{t}{12n} \left( \frac{t^4}{6} - \frac{t^2}{3} - 1 - u_\alpha \left\{ \frac{t^3}{6} - u_\alpha + t \right\} \right) \varphi(t - u_\alpha) + \mathcal{O}(n^{-3/2}), \end{aligned} \quad (1.6)$$

$$\begin{aligned} \beta_n(t) = & \Phi(t - u_\alpha) - \frac{t^2}{2\sqrt{n}}\varphi(t - u_\alpha) + \\ & + \frac{t}{2n} \varphi(t - u_\alpha) \left( \frac{t^2}{3} - \frac{t^3}{4}(t - u_\alpha) - \right. \\ & \left. - \frac{1}{6}(u_\alpha^2 + u_\alpha t - 3t^2 + 24\varepsilon_n(1 - \varepsilon_n) - 3) \right) + \mathcal{O}(n^{-3/2}), \end{aligned} \quad (1.7)$$

где

$$\varepsilon_n = \frac{n+1}{2} + \frac{\sqrt{nu_\alpha}}{2} - \left[ \frac{n+1}{2} + \frac{\sqrt{nu_\alpha}}{2} \right] + \mathcal{O}(n^{-1/2}),$$

и  $[y]$  – целая часть числа  $y \in \mathbb{R}^1$ .

Формулы (1.6) и (1.7) полезны при вычислении мощностей обоих критериев в случае, если размер выборки  $n$  велик. Для малых  $n$  мощности могут быть вычислены с помощью явных формул, приведенных ниже (см. (2.4) и (3.5) с использованием формул (3.1), (3.4)). Мы используем явные формулы при построении алгоритмов вычислений. При этом можно отметить, что для  $n > 50$  вычислительный процесс становится ресурсозатратным. С другой стороны, асимптотические разложения до порядка  $n^{-1}$  для функции распределения знаковой статистики и, соответственно, мощности критерия дают неудовлетворительный результат при

малых  $n$  и больших  $t$  (см. Таблицу 2.1). Объясняется это простым типом аппроксимации, который мы молчаливо выбрали среди всех представленных в работе [10] альтернатив (см. [10], стр. 72, формулы (2.1), (2.2)). Заметим, что разложение (1.7) дает хорошее приближение для малых  $n$ , но только для небольших значений параметра  $t$  (см. Таблицу 2.1). С ростом  $t$  при фиксированном  $n$  симметричный случай  $(k + \frac{1}{2} - np)(npq)^{-1/2}$  (см. (2.2) в работе [10]) становится всё менее пригодным, поскольку  $p$  значительно больше  $q$ , и ошибка фактически имеет порядок  $n^{-1/2}$ . Здесь можно рассматривать приближение более общего вида

$$(k + c - np)\{(n + d)pq + \delta\}^{-1/2},$$

где параметры  $c, d, \delta$  выбираются соответствующим образом (см. [10], стр. 75-79). Эти параметры можно выбрать так, что для заданного уровня значимости  $\alpha$  члены порядков  $n^{-1/2}$  и  $n^{-1}$  будут обращены в ноль (см. [10], стр. 88-90), но тогда формулу (1.7) нужно продолжить до порядка  $n^{-3/2}$  включительно (см. формулу (2.1) работы [10]).

С учетом приведенных замечаний в разделах 2 и 3 мы проводим сравнение численных оценок для мощностей критерия, основанного на знаковой статистике, и критерия, основанного на логарифме отношения правдоподобия, соответственно,  $\beta_n(t)$  и  $\beta_n^*(t)$ , с их асимптотическими приближениями до порядка  $n^{-1}$ , которые мы будем обозначать, соответственно,  $\bar{\beta}_n(t)$  и  $\bar{\beta}_n^*(t)$ , указывая на то, что член  $\mathcal{O}(n^{-3/2})$  опускается. В разделе 4 мы вычисляем дефект критерия знаковой статистики по отношению к наиболее мощному критерию в случае простой альтернативы  $H_{n,t}$  с использованием явных формул и через их асимптотические аналоги. Вопросы численной аппроксимации дефектов критериев рассматривались также в работах [11], [12]. Однако наш подход основан на явных формулах, возможных в случае распределения Лапласа.

## 2. Мощность критерия в случае знаковой статистики

В этом разделе будет использоваться явная формула для функции распределения знаковой статистики

$$\bar{T}_n = \sum_{j: X_j > 0} 1 \tag{2.1}$$

как при гипотезе  $H_0$  (см. (1.1)), так и при простой альтернативе  $H_{n,t}$  при фиксированном размере выборки  $n$ . Заметим, что преобразование статистики  $T_n$  (см. 1.5) к виду (2.1) не влияет на мощность  $\beta_n(t)$ . Обозначим через  $\bar{\beta}_n(t)$  правую часть формулы (1.7) без остаточного члена  $\mathcal{O}(n^{-3/2})$ .

Для каждого фиксированного  $n$  при гипотезе  $H_0$  знаковая статистика  $\bar{T}_n$  принимает значение  $k$  ( $k = 0, 1, \dots, n$ ) с вероятностью  $\binom{n}{k} \cdot 2^{-n}$ , так как в этом случае случайная величина  $X_j$  симметрично распределена. Ввиду того, что распределение знаковой статистики дискретно, будем использовать процедуру рандомизации: при заданном уровне значимости  $\alpha$ , если знаковая статистика  $\bar{T}_n$  больше или равна целого числа  $k_\alpha + 1$ , то критерий равен единице, а если  $\bar{T}_n = k_\alpha$ , то критерий равен  $\gamma_\alpha \in (0, 1)$ . Причем числа  $k_\alpha$  и  $\gamma_\alpha$  выбираются из условия

$$P_{n,0}(\bar{T}_n \geq k_\alpha + 1) + \gamma_\alpha P_{n,0}(\bar{T}_n = k_\alpha) = \alpha. \tag{2.2}$$

При гипотезе  $H_0$  выражение (2.2) примет вид

$$\sum_{k=k_\alpha+1}^n \binom{n}{k} \cdot 2^{-n} + \gamma_\alpha \binom{n}{k_\alpha} \cdot 2^{-n} = \alpha. \quad (2.3)$$

Функция мощности критерия, основанного на  $\bar{T}_n$ , может быть записана в виде

$$\beta_n(t) = \sum_{k=k_\alpha+1}^n \binom{n}{k} p^k q^{n-k} + \gamma_\alpha \binom{n}{k_\alpha} p^{k_\alpha} q^{n-k_\alpha}, \quad (2.4)$$

где (см., например, [4], стр. 102 – 103)

$$p = P_{n,tn-1/2}(X_1 > 0) = \int_0^\infty \frac{1}{2} e^{-|x-t/\sqrt{n}|} dx = 1 - \frac{1}{2} e^{-t/\sqrt{n}},$$

$q = 1 - p$ ,  $k_\alpha$  и  $\gamma_\alpha$  находятся из (2.3).

Рассмотрим аппроксимацию функции мощности  $\beta_n(t)$  (2.4) с помощью выражения  $\bar{\beta}_n(t)$ , определенного выше (см. (1.7)). Были проведены вычисления для трёх значений уровня значимости  $\alpha = 0.01; 0.05; 0.10$ , поскольку они часто используются в прикладных задачах.

Как упоминалось в разделе 1, при альтернативе  $\theta = t/\sqrt{n}$  при малых  $t$  приближение  $\bar{\beta}_n(t)$  будет давать хорошие результаты для мощности критерия знаковой статистики даже при малых  $n$ . Результаты вычислений для  $t = 0.25(0.25)1.50(0.50)4.50$  показаны в таблице 2.1, где значение  $\beta_n(t)$  расположено сверху, а  $\bar{\beta}_n(t)$  – снизу под ним. Обратим внимание, что для случаев  $n = 100; 1000$  вычисления по формуле (2.4) проводились с точностью до 15 знаков с использованием встроённых функций программного комплекса Matlab. В таблице 2.1 результаты округлены до пятого разряда.

Из таблицы 2.1 видно, что даже при небольших  $t$  для  $n = 10$  ошибка заметнее остальных случаев. Из этой же таблицы видно, что для больших  $t$  и малых  $n$  в случае  $\alpha = 0.01$  аппроксимация дает ошибку уже в первом знаке, и лишь при  $n = 1000$  приближение можно считать качественным. На графике 2.1 представлена функция мощности  $\beta_n(t)$  для  $n = 10$ ,  $\alpha = 0.01$  и  $t$  пробегает полуинтервал  $(0, 10]$ , причем приближение  $\bar{\beta}_n(t)$  дает неудовлетворительный результат уже для  $t > 2$ . Аналогично, на графике 3.1 (см. раздел 3) представлена похожая ситуация, где приближение для мощности  $\beta_n^*(t)$  показывает хороший результат уже при  $n = 3$ .

С ростом  $n$  ошибка

$$\delta_n(t, \alpha) \equiv \bar{\beta}_n(t) - \beta_n(t), \quad (2.5)$$

уменьшается, и для наглядности результатов (см. идею работы [5], разделы 1,2) мы рассмотрим график 2.2

$$\left\{ (\beta_n(t), \delta_n(t, \alpha)) : t > 0 \right\},$$

на котором представлена ошибка  $\delta_n(t, \alpha)$  в зависимости от значений мощности  $\beta_n(t)$  для  $n = 50$ ,  $\alpha = 0.01; 0.05; 0.10$  и  $t$  пробегает полуинтервал  $(0, 10]$ . График 2.3 показывает, что ошибка  $\delta_n$  «имеет порядок  $n^{-3/2}$ », т.е. величина  $n^{3/2}\delta_n(\cdot, \alpha)$  мало зависит от  $n$ .

Сравнение результатов этого раздела и результатов раздела 3 наглядно демонстрирует как поведение распределения статистики влияет на качество аппроксимации. В силу решетчатости распределения знаковой статистики точность аппроксимации нарушается при малых  $n$ , больших  $t$  и  $\alpha$  близких к единице. Во введении мы указали, что в этих случаях аппроксимацию можно улучшить специальным подбором параметров асимптотического разложения функции распределения знаковой статистики, а также повышением порядка аппроксимации (см. раздел 1).

$t =$	$\alpha =$	$n =$	10	20	30	40	50	100	1000
.25	.01		.01815	.01820	.01845	.01854	.01857	.01874	.01888
			.01788	.01824	.01850	.01857	.01858	.01873	.01888
	.05		.07971	.08006	.08065	.08039	.08057	.08091	.08136
			.07943	.07996	.08068	.08041	.08055	.08092	.08136
	.10		.14659	.14853	.15001	.14943	.15018	.15037	.15089
			.14657	.14848	.14999	.14943	.15017	.15038	.15090
.50	.01		.03014	.03064	.03158	.03195	.03212	.03284	.03358
			.02920	.03070	.03169	.03202	.03214	.03281	.03358
	.05		.11758	.11950	.12155	.12115	.12180	.12321	.12527
			.11688	.11923	.12164	.12116	.12174	.12321	.12527
	.10		.20107	.20705	.21131	.21030	.21239	.21357	.21605
			.20088	.20691	.21130	.21026	.21238	.21358	.21605
.75	.01		.04655	.04827	.05061	.05163	.05214	.05408	.05641
			.04445	.04830	.05083	.05176	.05215	.05401	.05641
	.05		.16294	.16808	.17271	.17245	.17398	.17751	.18300
			.16210	.16773	.17300	.17247	.17388	.17752	.18300
	.10		.26147	.27361	.28198	.28085	.28490	.28807	.29442
			.26143	.27355	.28212	.28082	.28496	.28811	.29442
1.00	.01		.06772	.07183	.07657	.07875	.07991	.08413	.08971
			.06444	.07194	.07705	.07903	.07996	.08402	.08971
	.05		.21455	.22473	.23317	.23345	.23631	.24321	.25435
			.21487	.22477	.23399	.23362	.23629	.24326	.25435
	.10		.32560	.34563	.35913	.35833	.36479	.37097	.38312
			.32697	.34616	.35975	.35848	.36508	.37109	.38312
1.25	.01		.09376	.10172	.11008	.11409	.11633	.12420	.13538
			.09062	.10256	.11129	.11484	.11659	.12410	.13538
	.05		.27083	.28774	.30104	.30235	.30694	.31840	.33740
			.27466	.28911	.30293	.30299	.30726	.31857	.33741
	.10		.39126	.42021	.43939	.43935	.44843	.45836	.47770
			.39587	.42211	.44083	.43995	.44912	.45861	.47770
1.50	.01		.12447	.13799	.15130	.15788	.16168	.17475	.19438
			.12475	.14097	.15414	.15970	.16255	.17481	.19438
	.05		.33005	.35495	.37386	.37663	.38322	.40008	.42855
			.33965	.35865	.37728	.37809	.38417	.40048	.42856
	.10		.45648	.49455	.51933	.52039	.53199	.54592	.57285
			.46539	.49824	.52163	.52161	.53311	.54632	.57286

Таблица 2.1 (начало). Мощность  $\beta_n(t)$  (сверху) и величина  $\bar{\beta}_n(t)$  (снизу) при различных  $t$ ,  $\alpha$  и  $n$ .

$t =$	$\alpha =$	$n =$	10	20	30	40	50	100	1000
2.0	.01		.19806	.22779	.25461	.26841	.27672	.28365	.29146
			.21692	.24021	.26371	.27438	.28031	.28582	.29299
	.05		.45053	.49287	.52323	.52963	.54023	.55078	.55430
			.46784	.50007	.52828	.53236	.54226	.55249	.55554
	.10		.57936	.63318	.66675	.67033	.68551	.68693	.69497
			.59036	.63776	.66825	.67174	.68630	.68778	.69545
2.5	.01		.28337	.33444	.37745	.39992	.41372	.42504	.43729
			.30697	.35136	.38908	.40752	.41845	.42807	.43932
	.05		.56455	.62207	.66098	.67068	.68395	.69700	.70179
			.56033	.62155	.65985	.67042	.68400	.69661	.70181
	.10		.68552	.74843	.78538	.79085	.80670	.80912	.81753
			.68049	.74644	.78094	.78961	.80465	.80848	.81640
3.0	.01		.37415	.44816	.50628	.53650	.55514	.57011	.58577
			.33551	.44293	.50342	.53419	.55302	.56810	.58332
	.05		.66480	.73175	.77398	.78540	.79914	.81240	.81754
			.61918	.71615	.76204	.77953	.79526	.80835	.81517
	.10		.77164	.83592	.87079	.87684	.89075	.89345	.90069
			.75213	.82825	.86153	.87350	.88648	.89165	.89838
3.5	.01		.46456	.55917	.62786	.66277	.68410	.70081	.71769
			.32629	.51824	.60215	.64518	.67162	.69133	.70880
	.05		.74810	.81770	.85800	.86923	.88144	.89294	.89752
			.69318	.80067	.84423	.86341	.87777	.88879	.89538
	.10		.83818	.89759	.92679	.93226	.94285	.94522	.95055
			.83084	.89533	.92058	.93128	.94030	.94482	.94940
4.0	.01		.55002	.65989	.73275	.76836	.78966	.80585	.82161
			.38984	.61739	.70541	.75042	.77796	.79761	.81363
	.05		.81437	.88076	.91547	.92509	.93466	.94339	.94689
			.79936	.88144	.91214	.92655	.93638	.94354	.94817
	.10		.88765	.93849	.96072	.96497	.97215	.97389	.97732
			.90862	.94755	.96187	.96843	.97325	.97598	.97839
4.5	.01		.62753	.74584	.81656	.84929	.86825	.88216	.89516
			.56570	.74944	.81635	.85040	.87115	.88558	.89677
	.05		.86532	.92457	.95206	.95944	.96619	.97211	.97446
			.90041	.94447	.96025	.96789	.97285	.97635	.97870
	.10		.92332	.96427	.97991	.98287	.98728	.98839	.99036
			.96275	.97958	.98556	.98841	.99034	.99152	.99247

Таблица 2.1 (окончание).

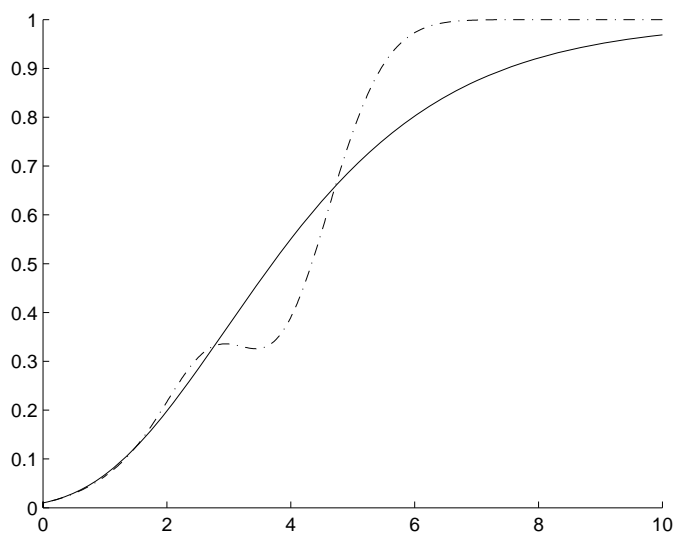


График 2.1. Случай  $\alpha = 0.01$ ,  $n = 10$ . Сплошной линией обозначена функция мощности  $\beta_n(t)$ , штрих-пунктирной — приближение  $\bar{\beta}_n(t)$ .

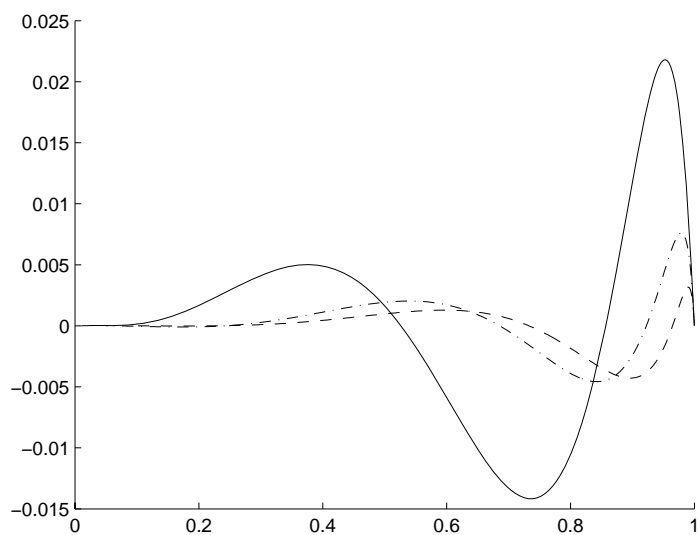


График 2.2. Ошибка аппроксимации  $\delta_n(\cdot, \alpha)$  в зависимости от значений мощности  $\beta_n(t)$ . Сплошной линией обозначена  $\delta_n(\cdot, \alpha)$ ,  $\alpha = 0.01$ ,  $n = 50$ , штрих-пунктирной —  $\delta_n(\cdot, \alpha)$ ,  $\alpha = 0.05$ ,  $n = 50$ , штриховой —  $\delta_n(\cdot, \alpha)$ ,  $\alpha = 0.10$ ,  $n = 50$ .

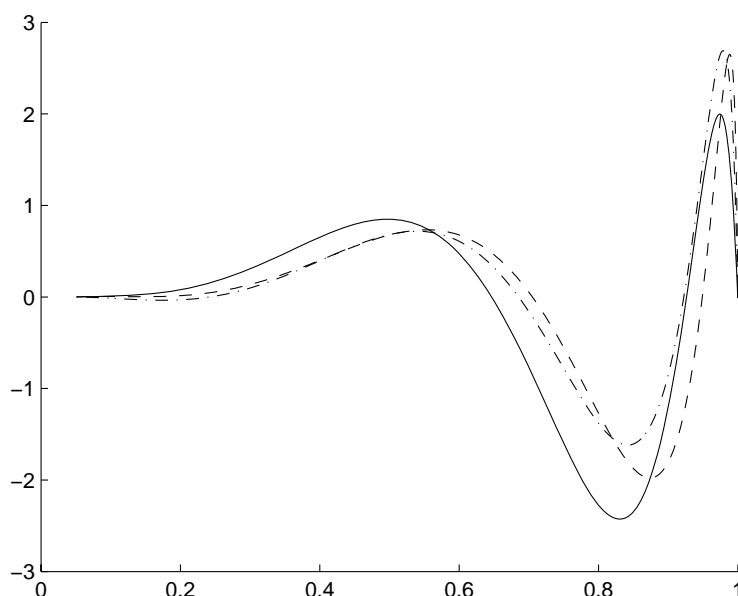


График 2.3. График величины  $n^{3/2}\delta_n(\cdot, \alpha)$  в зависимости от значений мощности  $\beta_n(t)$  при  $\alpha = 0.05$ . Сплошной линией обозначена  $n^{3/2}\delta_n(\cdot, \alpha)$ ,  $n = 30$ ,  $n^{3/2}\delta_n(\cdot, \alpha)$ ,  $n = 50$ , штриховой —  $n^{3/2}\delta_n(\cdot, \alpha)$ ,  $n = 100$ .

### 3. Аппроксимация мощности критерия, основанного на логарифме отношения правдоподобия

В этом разделе рассматривается задача численной аппроксимации распределения логарифма отношения правдоподобия  $\Lambda_n(t)$  (см. (1.4)). Проблема аппроксимации мощности критерия, основанного на  $\Lambda_n(t)$ , осложняется тем, что распределение этой статистики представляет собой смесь непрерывного и дискретного распределений. Явный вид функции распределения для  $\Lambda_n(t)$  получен в работе [6]. В работе [7] отмечается, что аппроксимация нормальным распределением не всегда достаточно точна для малых и средних размеров выборки (о сходимости распределения  $\Lambda_n(t)$  к нормальному закону см. также работу [1], Лемма 2.1). Ниже мы приводим результат из работы [6] (стр. 867-869, формулы 7,8) в виде Теоремы 3.1 (см. также [8], стр. 108-110).

**ТЕОРЕМА 3.1.** *Функция распределения логарифма отношения правдоподобия  $\Lambda_n(t)$  при гипотезе  $H_0$  имеет вид*

$$P_{n,0}(\Lambda_n(t) \leq x) = 2^{-n} \sum_{k=1}^n \binom{n}{k} \sum_{r=0}^k (-1)^r \binom{k}{r} \sum_{s=0}^{n-k} \binom{n-k}{s} \times \\ \times \left\{ e^{-(r+s)tn^{-1/2}} - e^{-\frac{1}{2}(x+ntn^{-1/2})} \cdot e_{k-1} \left( \frac{1}{2}(x + (n-2r-2s)tn^{-1/2}) \right) \right\} \times$$



$$\begin{aligned} & \times \mathbf{1}_{[0, \infty)}(x + (n - 2r - 2s)tn^{-1/2}) + \\ & + 2^{-n} \sum_{l=0}^n \binom{n}{l} e^{-ltn^{-1/2}} \mathbf{1}_{[0, \infty)}(x + (n - 2l)tn^{-1/2}), \end{aligned} \quad (3.1)$$

где  $e_{k-1}(\cdot)$  – неполная экспоненциальная функция

$$e_{k-1}(z) = \sum_{i=0}^{k-1} \frac{z^i}{i!}.$$

Функция распределения логарифма отношения правдоподобия  $\Lambda_n(t)$  при альтернативе  $H_{n,t}$  имеет вид

$$P_{n,t}(\Lambda_n(t) \leq x) = 1 - P_{n,0}(\Lambda_n(t) < -x). \quad (3.2)$$

ДОКАЗАТЕЛЬСТВО.

Доказательство первого утверждения теоремы приведено в работах [6] (стр. 867-869) и [8] (стр. 108-110). Доказательство (3.2) проведем самостоятельно, заметив, что в формуле (3.2) в правой части нужно рассматривать событие  $\{\Lambda_n(t) < -x\}$  вместо события  $\{\Lambda_n(t) \leq -x\}$ , которое ошибочно следует из формальной записи в работах [6], [8].

Рассмотрим случайную величину  $\xi_i$  с распределением  $H_0$ , тогда случайная величина  $X_i = \xi_i + tn^{-1/2}$  имеет распределение  $H_{n,t}$ . Так как случайные величины  $\xi_i$  и  $-\xi_i$  ввиду симметричности закона Лапласа одинаково распределены, по определению  $\Lambda_n(t)$  (см. (1.4)) имеем

$$\begin{aligned} P_{n,t}(\Lambda_n(t) \leq x) &= P_{n,0}\left(\sum_{i=1}^n (|\xi_i + tn^{-1/2}| - |\xi_i|) \leq x\right) = \\ &= P_{n,0}\left(\sum_{i=1}^n (|-\xi_i + tn^{-1/2}| - |-\xi_i|) \leq x\right) = P_{n,0}\left(\sum_{i=1}^n (|\xi_i - tn^{-1/2}| - |\xi_i|) \leq x\right) = \\ &= P_{n,0}\left(\sum_{i=1}^n (|\xi_i| - |\xi_i - tn^{-1/2}|) \geq -x\right) = 1 - P_{n,0}(\Lambda_n(t) < -x). \end{aligned}$$

□

Критическое значение  $c_{n,t}$  находится из условия

$$P_{n,0}(\Lambda_n(t) \leq c_{n,t}) - \gamma_\alpha^* P_{n,0}(\Lambda_n(t) = c_{n,t}) = 1 - \alpha. \quad (3.3)$$

Однако прямое применение формулы (3.1) приводит к громоздким вычислениям, что, в свою очередь, сказывается на точности результатов. Следуя работе [7] (стр. 580) мы несколько изменим формулу (3.1), заметив, что  $1 - \alpha$  находится вблизи единицы, поэтому запишем формулу (3.1) в виде (см. 7, формула A1, стр. 580)

$$P_{n,0}(\Lambda_n(t) \leq x) = 1 - 2^{-n} \sum_{k=1}^n \binom{n}{k} \sum_{r=0}^k (-1)^r \binom{k}{r} \sum_{s=0}^{n-k} \binom{n-k}{s} \times$$

$$\begin{aligned}
& \times \left\{ e^{-(r+s)tn^{-1/2}} - e^{-\frac{1}{2}(x+ntn^{-1/2})} \cdot e_{k-1} \left( \frac{1}{2}(x + (n-2r-2s)tn^{-1/2}) \right) \right\} \times \\
& \quad \times \mathbf{1}_{(-\infty, 0)}(x + (n-2r-2s)tn^{-1/2}) + \\
& \quad + 2^{-n} \sum_{l=0}^n \binom{n}{l} e^{-ltn^{-1/2}} \mathbf{1}_{(-\infty, 0)}(x + (n-2l)tn^{-1/2}). \quad (3.4)
\end{aligned}$$

Как видно из структуры функции распределения  $\Lambda_n(t)$ , критическое значение  $c_{n,t}$  может оказаться как точкой разрыва, так и точкой непрерывности. В случае, если  $c_{n,t}$  точка разрыва функции распределения, мы будем применять процедуру рандомизации (см. (3.3)), в случае, если  $c_{n,t}$  – точка непрерывности, то уравнение (3.3) приобретает вид

$$\mathbf{P}_{n,0}(\Lambda_n(t) \leq c_{n,t}) = 1 - \alpha.$$

Поскольку в явном виде решить уравнение (3.3) нельзя, то для второго случая мы используем численный метод решения уравнений, а именно, метод деления отрезка пополам (см., например, [9], стр. 190-191). Точность численного метода в наших вычислениях составляет  $10^{-7}$ .

Выражение

$$1 - \beta_n^*(t) = \mathbf{P}_{n,t}(\Lambda_n(t) \leq c_{n,t}) - \gamma_\alpha^* \mathbf{P}_{n,t}(\Lambda_n(t) = c_{n,t})$$

с учетом (3.2) и равенства

$$\mathbf{P}_{n,t}(\Lambda_n(t) = c_{n,t}) = \mathbf{P}_{n,0}(\Lambda_n(t) = -c_{n,t})$$

дает формулу для вычисления мощности

$$\beta_n^*(t) = \mathbf{P}_{n,0}(\Lambda_n(t) \leq -c_{n,t}) + (\gamma_\alpha^* - 1) \mathbf{P}_{n,0}(\Lambda_n(t) = -c_{n,t}). \quad (3.5)$$

Таблица 3.1 содержит результаты расчетов мощности, полученных с помощью явной формулы (3.1) и по формуле (1.6) без остаточного члена, соответственно,  $\beta_n^*(t)$  расположено сверху, а  $\bar{\beta}_n^*(t)$  - снизу для  $\alpha = 0.01; 0.05; 0.10$ ,  $t = 0.25(0.25)1.50(0.50)4.50$ , и  $n = 10(10)50$ .

Аналогично разделу 2 здесь также рассмотрим ошибку приближения  $\bar{\beta}_n^*(t)$  к функции мощности  $\beta_n^*(t)$  наилучшего критерия. График 3.1 показывает функцию мощности для  $n = 3$ ,  $\alpha = 0.01$ , где  $t$  пробегает полуинтервал  $(0, 10]$ . Для больших  $n$  ошибка мала, поэтому на графике 3.2 аналогично разделу 2 представлена ошибка  $\delta_n^*(t, \alpha)$  в зависимости от значений мощности  $\beta_n^*(t)$  для  $n = 10$ ,  $\alpha = 0.01; 0.05; 0.10$  и  $t$  пробегает полуинтервал  $(0, 10]$ . График 3.3 показывает, что ошибка  $\delta_n^*$  «имеет порядок  $n^{-3/2}$ », то есть величина  $n^{3/2}\delta_n^*(\cdot, \alpha)$  мало зависит от  $n$ .

Как отмечалось выше, логарифм отношения правдоподобия содержит участки с непрерывности, кроме того, он ограничен, что объясняет достаточно хорошее приближение асимптотическими разложениями типа Эджворта (см. [2], Лемма 3.2). В итоге наблюдается вполне удовлетворительное приближение функции мощности  $\beta_n^*(t)$  даже при малых  $n$ , что позволяет использовать величину  $\bar{\beta}_n^*(t)$  вместо громоздких вычислений с использованием формул (3.1), (3.4).

$t =$	$\alpha =$	$n =$	10	20	30	40	50	
.25	.01		.01837	.01856	.01865	.01873	.01879	
			.01841	.01864	.01872	.01876	.01879	
	.05		.08072	.08106	.08093	.08113	.08125	
			.08063	.08098	.08111	.08118	.08123	
		.10		.14933	.15077	.15040	.15078	.15061
				.15022	.15052	.15065	.15072	.15077
.50	.01		.03138	.03283	.03306	.03320	.03326	
			.03193	.03271	.03300	.03316	.03326	
	.05		.12153	.12332	.12430	.12467	.12470	
			.12274	.12387	.12434	.12460	.12478	
		.10		.21434	.21461	.21493	.21533	.21560
				.21349	.21456	.21504	.21533	.21553
.75	.01		.05131	.05424	.05496	.05529	.05561	
			.05235	.05421	.05493	.05534	.05560	
	.05		.17482	.17951	.18087	.18121	.18186	
			.17714	.17964	.18073	.18138	.18181	
		.10		.28954	.28992	.29215	.29259	.29313
				.28843	.29085	.29198	.29268	.29316
1.0	.01		.08140	.08482	.08657	.08751	.08807	
			.08153	.08514	.08662	.08747	.08804	
	.05		.24290	.24818	.24980	.25125	.25198	
			.24346	.24795	.24999	.25123	.25208	
		.10		.37160	.37641	.37836	.37987	.38072
				.37219	.37648	.37856	.37985	.38076
1.25	.01		.12164	.12718	.12982	.13133	.13234	
			.12100	.12713	.12977	.13131	.13236	
	.05		.32030	.32673	.33006	.33214	.33360	
			.31993	.32691	.33021	.33224	.33365	
		.10		.45744	.46729	.47033	.47238	.47390
				.46077	.46729	.47052	.47254	.47397
1.50	.01		.17194	.18099	.18513	.18765	.18943	
			.17171	.18105	.18524	.18775	.18947	
	.05		.40315	.41224	.41778	.42076	.42290	
			.40346	.41324	.41801	.42097	.42305	
		.10		.54727	.55799	.56270	.56553	.56752
				.54955	.55844	.56289	.56568	.56766

Таблица 3.1 (начало). Мощность  $\beta_n^*(t)$  (сверху) и величина  $\bar{\beta}_n^*(t)$  (снизу) при различных  $t$ ,  $\alpha$  и  $n$ .

$t =$	$\alpha =$	$n =$	10	20	30	40	50
2.0	.01		.30024	.32056	.32956	.33490	.33861
			.30500	.32203	.33026	.33538	.33895
	.05		.57151	.58977	.59774	.60275	.60626
			.57505	.59055	.59826	.60309	.60650
			.71023	.72337	.72995	.73407	.73696
			.71058	.72371	.73018	.73422	.73706
2.50	.01		.45859	.48775	.50126	.50964	.51544
			.46433	.48982	.50248	.51042	.51601
	.05		.72517	.74559	.75541	.76141	.76558
			.72564	.74582	.75548	.76144	.76560
			.83294	.84763	.85457	.85884	.86180
			.83174	.84715	.85430	.85865	.86166
3.0	.01		.61770	.65256	.66911	.67920	.68620
			.61907	.65317	.66938	.67936	.68632
	.05		.84110	.86058	.86968	.87521	.87903
			.83729	.85915	.86887	.87467	.87863
			.91339	.92603	.9317	.93511	.93744
			.91043	.92483	.93102	.93466	.93711
3.5	.01		.75397	.78946	.80559	.81534	.82204
			.74703	.78659	.80393	.81422	.82121
	.05		.91669	.93225	.93914	.94322	.94599
			.91146	.93029	.93805	.94250	.94547
			.95962	.96822	.97194	.97412	.97557
			.95688	.96725	.97142	.97378	.97533
4.0	.01		.85616	.88607	.89911	.90679	.91198
			.84500	.88178	.89670	.90520	.91082
	.05		.96030	.97071	.97502	.97748	.97911
			.95735	.96969	.97448	.97714	.97887
			.98276	.98783	.98987	.99100	.99174
			.98202	.98764	.98978	.99096	.99171
4.5	.01		.92336	.94485	.95371	.95874	.96205
			.91655	.94254	.95249	.95796	.96150
	.05		.98272	.98866	.99094	.99218	.99297
			.98277	.98881	.99106	.99227	.99305
			.99322	.99582	.99676	.99726	.99757
			.99384	.99610	.99693	.99738	.99766

Таблица 3.1 (окончание).

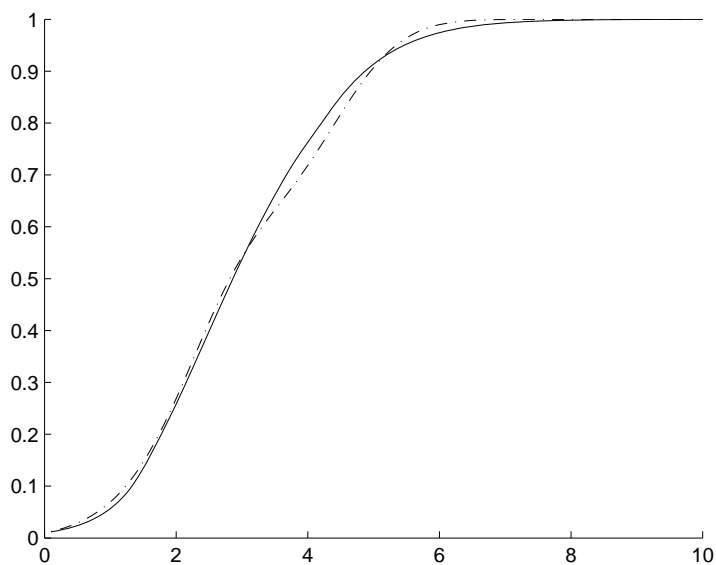


График 3.1. Случай  $\alpha = 0.01$  и  $n = 3$ . Сплошной линией обозначена функция мощности  $\beta_n^*(t)$ , штрих-пунктирной — приближение  $\bar{\beta}_n^*(t)$ .

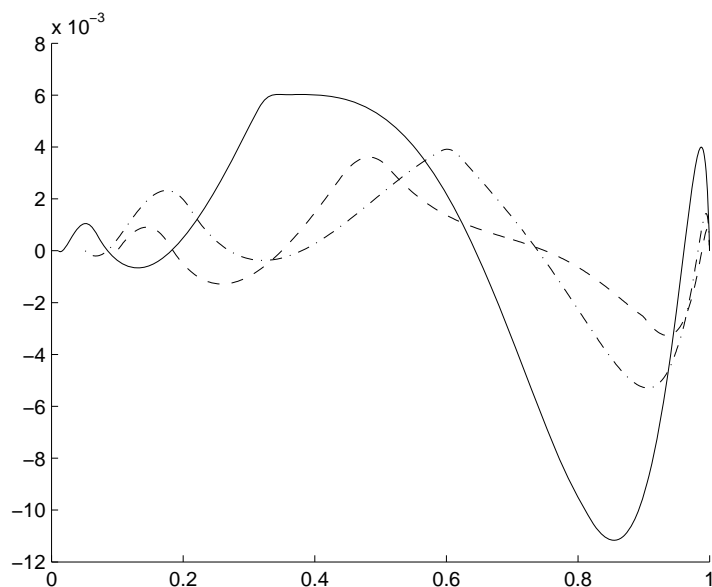


График 3.2. Ошибка аппроксимации  $\delta_n^*(\cdot, \alpha)$  в зависимости от значений мощности  $\beta_n^*(t)$ . Сплошной линией обозначена  $\delta_n^*(\cdot, \alpha)$ ,  $\alpha = 0.01$ ,  $n = 10$ , штрих-пунктирной —  $\delta_n^*(\cdot, \alpha)$ ,  $\alpha = 0.05$ ,  $n = 10$ , пунктирной —  $\delta_n^*(\cdot, \alpha)$ ,  $\alpha = 0.10$ ,  $n = 10$ .

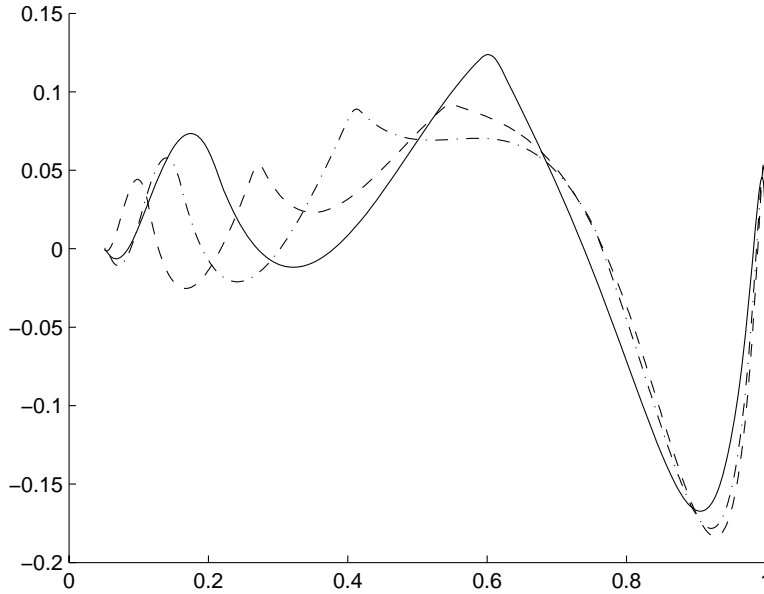


График 3.3. График величины  $n^{3/2}\delta_n^*(\cdot, \alpha)$  в зависимости от значений мощности  $\beta_n^*(t)$  при  $\alpha = 0.05$ . Сплошной линией обозначена  $n^{3/2}\delta_n^*(\cdot, \alpha)$ ,  $n = 10$ , штрихпунктирной —  $n^{3/2}\delta_n^*(\cdot, \alpha)$ ,  $n = 20$ , пунктирной —  $n^{3/2}\delta_n^*(\cdot, \alpha)$ ,  $n = 30$ .

#### 4. Аппроксимация для дефекта критерия, основанного на знаковой статистике

В этом разделе получены приближенные значения для дефекта критерия, основанного на знаковой статистике  $T_n$ .

Для любых фиксированных  $t$  и  $n$  наилучший критерий для проверки гипотезы (1.1) против простой альтернативы

$$H_{n,t} : \theta = \frac{t}{\sqrt{n}} \quad (4.1)$$

основан на логарифме отношения правдоподобия  $\Lambda_n(t)$  (см. работу [1]). Таким образом, функция мощности такого критерия является огибающей функцией мощности, поэтому для фиксированных  $\alpha$  и  $n$  выполняется неравенство

$$\beta_n^*(t) \geq \beta_n(t) \quad (4.2)$$

для всех  $t \in (0, C]$ ,  $C > 0$ , см. (1.2).

Пусть  $m_n(\alpha, t)$  обозначает наименьшее число  $m$  такое, что

$$\beta_m\left(\left(\frac{m}{n}\right)^{1/2}t\right) \geq \beta_n^*(t).$$

Пусть  $p_n(\alpha, t)$  будет решение относительно  $p$  уравнения

$$p\beta_m\left(\left(\frac{m}{n}\right)^{1/2}t\right) + (1-p)\beta_{m-1}\left(\left(\frac{m-1}{n}\right)^{1/2}t\right) = \beta_n^*(t)$$

где  $m = m_n(\alpha, t)$ .

Тогда

$$\bar{m}_n(\alpha, t) = p_n(\alpha, t) m_n(\alpha, t) + (1 - p_n(\alpha, t))(m_n(\alpha, t) - 1)$$

будет размер выборки  $m$ , для которого асимптотически оптимальный критерий эквивалентен наилучшему критерию. Дефект тогда определяется как

$$d_n(\alpha, t) = \bar{m}_n(\alpha, t) - n. \quad (4.3)$$

Используя асимптотические разложения для мощностей  $\beta_n^*(t)$  и  $\beta_n(t)$  (см. (1.6), (1.7)), получаем соотношение

$$\begin{aligned} \hat{d}_n(\alpha, t) \equiv \hat{d}_n = \sqrt{n}\frac{2t}{3} + \frac{2t^4}{9} - \frac{2u_\alpha t^3}{9} - \frac{7t^2}{9} + \frac{u_\alpha t}{3} - \frac{1}{3} + \\ + 4 \cdot \varepsilon_{n+\hat{d}_n}(1 - \varepsilon_{n+\hat{d}_n}) + \mathcal{O}(n^{-1/2}), \end{aligned} \quad (4.4)$$

где

$$\varepsilon_{n+\hat{d}_n} = \frac{n + \hat{d}_n + 1}{2} + \frac{u_\alpha \sqrt{n + \hat{d}_n}}{2} - \left[ \frac{n + \hat{d}_n + 1}{2} + \frac{u_\alpha \sqrt{n + \hat{d}_n}}{2} \right] + \mathcal{O}(n^{-1/2}).$$

При получении (4.4) применяется формула Тейлора к функции  $\beta_m((m/n)^{1/2}t)$  (см. (1.7)) при  $m = n + \hat{d}_n$ , при этом  $\hat{d}_n$  интерпретируется как непрерывная переменная (в отличие от  $d_n(\alpha, t)$ ) и используется формула (1.6). Заметим, что с ростом  $n$  величина  $\hat{d}_n(\alpha, t)$  растет со скоростью  $\sqrt{n}$ , что объясняется нерегулярностью распределения Лапласа (см. [1], стр. 18), выражение (4.4) содержит члены, не зависящие от  $n$ , и  $0 \leq 4 \cdot \varepsilon_{n+\hat{d}_n}(1 - \varepsilon_{n+\hat{d}_n}) \leq 1$ .

Для наглядности численных результатов определим  $t_n(\alpha, \beta)$  так, что

$$\beta_n^*(t_n(\alpha, \beta)) = \beta \quad (4.5)$$

и рассмотрим

$$D_n(\alpha, \beta) = d_n(\alpha, t_n(\alpha, \beta)),$$

соответственно,

$$\hat{D}_n(\alpha, \beta) = \hat{d}_n(\alpha, t_n(\alpha, \beta)).$$

Из (1.6) следует, что

$$t_n(\alpha, \beta) = u_\alpha + u_\beta + \frac{(u_\alpha + u_\beta)^2}{6\sqrt{n}} + \mathcal{O}(n^{-1}), \quad (4.6)$$

где  $\Phi(u_\beta) = \beta$ . Для получения  $D_n(\alpha, \beta)$  мы решаем уравнение (4.5) относительно  $t_n(\alpha, \beta)$  численным методом (методом деления отрезка пополам) с точностью  $10^{-7}$  и далее используем явные формулы для мощностей критериев. Для получения

асимптотического дефекта выражение (4.6) подставляется в (4.4) и опускаются члены порядка  $\mathcal{O}(n^{-1/2})$ . Обозначим

$$\hat{D}_n(\alpha, \beta) = \bar{D}_n(\alpha, \beta) + \mathcal{O}(n^{-1/2}),$$

тогда, решая относительно  $\bar{D}_n(\alpha, \beta)$  получившееся уравнение, имеем

$$\begin{aligned} \bar{D}_n(\alpha, \beta) &= \sqrt{n} \frac{2}{3}(u_\alpha + u_\beta) + \frac{2}{9}(u_\alpha + u_\beta)^4 - \frac{2u_\alpha}{9}(u_\alpha + u_\beta)^3 - \frac{2}{3}(u_\alpha + u_\beta)^2 + \\ &+ \frac{u_\alpha}{3}(u_\alpha + u_\beta) - \frac{1}{3} + 4 \cdot \varepsilon_{n+\bar{D}_n(\alpha, \beta)}(1 - \varepsilon_{n+\bar{D}_n(\alpha, \beta)}) \equiv \\ &\equiv C_n(\alpha, \beta) + 4 \cdot \varepsilon_{n+\bar{D}_n(\alpha, \beta)}(1 - \varepsilon_{n+\bar{D}_n(\alpha, \beta)}), \end{aligned} \quad (4.7)$$

где  $C_n(\alpha, \beta)$  – константа при фиксированных  $n$ ,  $\alpha$  и  $\beta$ . Уравнение (4.7) решается численным методом (методом деления отрезка пополам) с точностью  $10^{-7}$ . Из всех возможных решений уравнения (4.7) мы выбираем точку, попадающую в полуинтервал (см. График 4.1)

$$\left[ 2M - 1 + \frac{u_\alpha^2}{2} - u_\alpha \sqrt{2M - 1 + \frac{u_\alpha^2}{4}} - n, 2(M+1) - 1 + \frac{u_\alpha^2}{2} - u_\alpha \sqrt{2(M+1) - 1 + \frac{u_\alpha^2}{4}} - n \right),$$

где

$$M = \left\lceil \frac{n + C_n(\alpha, \beta) + 1}{2} + \frac{u_\alpha \sqrt{n + C_n(\alpha, \beta)}}{2} \right\rceil.$$

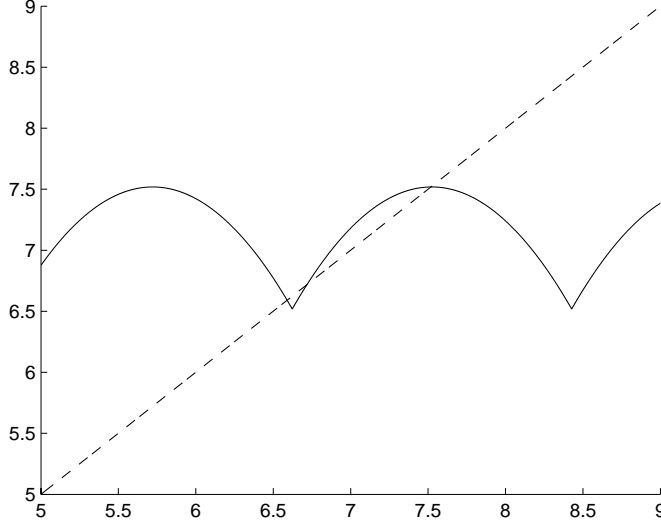


График 4.1. Решение уравнения (4.7) для случая  $n = 50$ ,  $\alpha = 0.05$ ,  $\beta = 0.5$ . Сплошной линией обозначена функция  $g(\bar{D}_n(\alpha, \beta)) \equiv C_n(\alpha, \beta) + 4 \cdot \varepsilon_{n+\bar{D}_n(\alpha, \beta)}(1 - \varepsilon_{n+\bar{D}_n(\alpha, \beta)})$ , пунктирной – функция  $f(\bar{D}_n(\alpha, \beta)) \equiv \bar{D}_n(\alpha, \beta)$ .



$\alpha = 0.05$						
$\beta =$	$n =$	10	20	30	40	50
0.5		4.43	5.67	6.54	7.70	8.60
		2.79	4.40	5.09	5.81	6.59
		3.47	4.90	6.01	6.94	7.75
0.9		7.10	9.30	11.21	12.55	13.85
		9.42	11.64	13.84	15.97	17.13
		6.17	8.72	10.69	12.34	13.80

Таблица 4.1. Случаи  $\alpha = 0.05$ ,  $\beta = 0.5$  и  $0.9$   
 дефект  $D_n(\alpha, \beta)$  представлен сверху  
 аппроксимация  $\bar{D}_n(\alpha, \beta)$  - в середине  
 аппроксимация  $\tilde{D}_n(\alpha, \beta)$  - снизу

Результаты вычислений представлены в таблице 4.1 для  $\alpha = 0.05$  и двух значений мощности  $\beta = 0.5$  и  $\beta = 0.9$ ,  $n = 10(10)50$ . Как видно из Таблицы 4.1, для данных случаев величина  $\bar{D}_n(\alpha, \beta)$  дает ошибку больше единицы, и может рассматриваться в качестве аналога дефекта весьма условно. Заметим, что величина

$$\tilde{D}_n(\alpha, \beta) = \sqrt{n} \frac{2}{3} (u_\alpha + u_\beta)$$

дает лучшее приближение к дефекту по сравнению с величиной  $\bar{D}_n(\alpha, \beta)$  для этих случаев (см. в Таблице 4.1 нижние значения). Учитывая тот факт, что в выражении (4.4) присутствуют члены, не зависящие от  $n$ , то, по-видимому, аппроксимация более высокого порядка для функций мощности не улучшит ситуацию. По нашему мнению, повышению порядка аппроксимации должна сопутствовать настройка приближения для мощности критерия знаковой статистики такая, как указано во введении. Исследованию этого вопроса будет посвящена одна из следующих работ, до появления которой мы рекомендуем использовать в качестве аппроксимации дефекта величину  $\tilde{D}_n(\alpha, \beta)$  вместо  $\bar{D}_n(\alpha, \beta)$ .

Автор выражает благодарность профессору Бенингу Владимиру Евгеньевичу за внимание и рекомендации, используемые при подготовке данной работы. Все вычисления проводились с использованием программного комплекса Matlab.

### Список литературы

- [1] Королев Р. А., Тестова А. В., Бенинг В. Е. О мощности асимптотически оптимального критерия в случае распределения Лапласа // Вестник Тверского государственного университета, серия Прикладная математика, 2008, выпуск 8, номер 4(64), стр. 5 – 23.

- 
- [2] Королев Р. А., Бенинг В. Е. Асимптотические разложения для мощностей критериев в случае распределения Лапласа // Вестник Тверского государственного университета, серия Прикладная математика, 2008, выпуск 3(10), номер 26(86), стр. 97 – 107.
- [3] Леман Э. Проверка статистических гипотез. – М.: Наука, 1964, 498 стр.
- [4] Gibbons J. D. Nonparametric Statistical Inference. – McGraw-Hill, USA, 1971, 306 p.
- [5] Pfaff T., Pfanzagl J. On the Accuracy of Asymptotic Expansions for Power Functions. – Preprints in Statistics, University of Cologne, 1983, No. 81, 64 p.
- [6] Marks R. J., Wise G. L., Haldeman D. G., Whited J. L. Detection in Laplace noise // IEEE Trans. Aerospace Electron. Systems, 1978, AES-14(6), p. 866-871.
- [7] Dadi M. I., Marks R. J. II. Detector relative efficiencies in the presence of Laplace noise // IEEE Trans. Aerospace Electron. Systems, 1987, AES-23(4), p. 568-582.
- [8] Kotz S., Kozubowski T. J., Podgórski K. The Laplace distribution and generalizations: a revisit with applications to communications, economics, engineering, and finance. – Birkhäuser Boston, USA, 2001, 349 p.
- [9] Волков Е. А. Численные методы. – 2-е изд., испр. – М.: Наука, 1987, 248 p.
- [10] Molenaar W. Approximations to the Poisson, binomial and hypergeometric distribution functions. – Mathematical Centre, Amsterdam, 1970, 160 p.
- [11] Albers W. Asymptotic Expansions and the Deficiency Concept in Statistics.– Mathematisch Centrum, Amsterdam, 1974, 145 p.
- [12] Pfanzagl J. Asymptotic Expansions in Parametric Statistical Theory.– in: Developments in Statistics, Ed. by Krishnaiah P. R., Academic Press, New York–London, 1980, p. 1–97.