УДК 519.23

Мощность критерия однородности как функция полезности в задачах принятия решения в условиях риска и неопределенности[1]

П. А. Филоненко, С. Н. Постовалов

В задачах проверки статистических гипотез существует большое количество статистических методов. Некоторые статистические критерии являются предпочтительными при определенных альтернативных гипотезах. Необходим некоторый надежный способ выбора предпочтительного критерия. Мы формируем типы альтернативных гипотез (с различным количеством точек пересечений функций надежности), для каждого типа строим несколько альтернативных гипотез (с разными законами распределения моментов отказа), а затем в соответствие с теорией принятия решений в условиях риска и неопределенности (критерий Вальда) мы получаем, какой критерий предпочтительнее при определенном типе альтернативной гипотезы, используя результаты компьютерного моделирования мощности критериев.

Ключевые слова: проверка статистических гипотез, гипотеза однородности распределений, данные типа времени жизни, критерий Вальда, метод Монте-Карло.

1. Введение

Мощность критерия [1] – это важная характеристика статистического критерия, потому что это вероятность отклонить проверяемую статистическую гипотезу, когда верна альтернативная гипотеза. Каждый статистический критерий имеет определенную статистическую мощность на заданной альтернативной гипотезе. Чем выше мощность статистического критерия, тем выше надежность статистических выводов. Наиболее популярным способом исследования мощности критериев является метод имитационного моделирования Монте-Карло [2]. В отличие от других статей [1, 2] мы представляем несколько иной способ исследования мощности статистического критерия: разбиваем множество альтернативных гипотез на определенные типы, при которых функции надежности (следовательно, и мощности критериев) близки. Затем, применяя критерий Вальда [3] для принятия решений в условиях риска и неопределенности [4], мы можем сделать вывод, какой критерий предпочтителен при определенном типе альтернативных гипотез. Критерий Вальда – это критерий выбора оптимальной стратегии в том случае, когда среда находится в наихудшем состоянии (значения мощности критериев минимальны). Это значит, что мощность критерия будет гарантированно наибольшей в ситуации, когда нет априорных предположений о виде альтернативной гипотезы.

[1] Данная работа выполнена в рамках проектной части государственного задания Министерства образования и науки Российской Федерации (номер заявки 1.1009.2017/ПЧ) и в рамках реализуемой программы стратегического развития НГТУ по итогам конкурса проектов среди молодых учёных (проект С-15, 2017).

Таким образом, мы определяем типы альтернативных гипотез, затем для каждого типа мы описываем несколько функций надежности, используя различные семейства распределений.

В разделе 2 мы рассмотрим постановку задачи. Кроме того, в разделе 2 мы представим использованные статистические критерии, а также представим модификацию критерия Крамера–Уэлча для данных типа времени жизни. В разделе 3 мы представим критерий Вальда и определение функции полезности [4]. Функция полезности – это основная характеристика в теории принятия решений в условиях риска и неопределенности. В разделе 4 – типы альтернативных гипотез, а в разделе 5 – результаты компьютерного моделирования.

2. Статистические критерии

Предположим, что имеются две выборки $X_1 = \left\{t_{11},...,t_{1n_1}\right\}$ и $X_2 = \left\{t_{21},...,t_{2n_2}\right\}$ из двух случайных величин ξ_1 и ξ_2 соответственно, которым соответствуют функции надежности $S_1(t)$ и $S_2(t)$. Каждый элемент выборки t_{ij} представим в виде $t_{ij} = \min\left(T_{ij}, C_{ij}\right)$, где T_{ij} и C_{ij} – моменты времени отказа и цензурирования *j*-го объекта в *i*-ой группе соответственно. Моменты времени T_{ij} и C_{ij} являются независимыми и одинаково распределенными случайными величинами с функциями распределения вероятностей $F_i(t)$ и $F_i^C(t)$.

Значение функции надежности $S(t)$ – это вероятность события, что не произойдет отказ объекта до момента времени t:

$$S_i(t) = P\left\{\xi_i > t\right\} - 1 - F_i(t).$$

Таким образом, нулевая гипотеза может быть записана в следующем виде:

$$H_0 : S_1(t) = S_2(t).$$

В дальнейшем мы будем предполагать, что элементы выборок упорядочены $t_{i,j} \le t_{i,j+1} \forall i, j$. Кроме того, объединим выборки $X = X_1 \cup X_2 = \left\{t_1,...,t_n\right\}$, $n = n_1 + n_2$, $t_j \le t_{j+1} \forall j$.

Введем в рассмотрение следующие индикаторы:

$$v_i = \begin{cases} 0, t_i \in X_1, \\ 1, t_i \in X_2 \end{cases}; \quad c_{ij} = \begin{cases} 0, t_{ij} \text{ – отказ}, \\ 1, t_{ij} \text{ – цензурирование} \end{cases}; \quad c_i = \begin{cases} 0, t_i \text{ – отказ}, \\ 1, t_i \text{ – цензурирование} \end{cases}.$$

К данному моменту мы ввели все необходимые обозначения и можем рассмотреть статистические критерии для проверки гипотезы однородности. В данной работе мы используем следующие критерии: обобщенные критерии Уилкоксона, логарифмический ранговый критерий, критерий Кокса–Мантела, Q-критерий, критерий максимального значения, взвешенный критерий Каплана–Мейера, обобщенный критерий Крамера–Уэлча и критерии Багдонавичуса–Никулина.

2.1. Обобщенный Гехановом критерий Уилкоксона

Статистики обобщенного Гехановом критерия Уилкоксона (в дальнейшем будем называть критерий Гехана) – это классический критерий Уилкоксона [5, 6, 7], обобщенный на случай данных типа времени жизни [2]. Статистика критерия может быть вычислена следующим образом:

$$S_G = \frac{\sum_{i=1}^{n}(1-v_i)h_i}{\sqrt{\frac{n_1 n_2}{n(n-1)}\sum_{i=1}^{n}(1-v_i)h_i^2}}, \text{ где } h_i = \sum_{i=1}^{n} v_j h_{ij}, h_{ij} = \begin{cases} +1, & t_i > t_j \,\&\, c_j = 0 \,\&\, v_i = 0 \,\&\, v_j = 1 \\ -1, & t_i < t_j \,\&\, c_j = 0 \,\&\, v_i = 0 \,\&\, v_j = 1. \\ 0, & \text{иначе} \end{cases}$$

Нулевая гипотеза H_0 отклоняется с достигнутым уровнем значимости α, если $|S_G| > z_{1-\alpha/2}$, где $z_{1-\alpha/2}$ – это $(1-\alpha/2)$-квантиль стандартного нормального распределения.

2.2. Обобщенный Пето критерий Уилкоксона

Статистики обобщенного Пето критерия Уилкоксона (в дальнейшем будем называть критерий Пето) – это классический критерий Уилкоксона [5, 6, 8], обобщенный на случай данных типа времени жизни [2]. Статистика критерия может быть вычислена следующим образом:

$$S_P \frac{n(n-1)\sum_{i=1}^{n} u_i(1-v_i)}{n_1 n_2 \sum_{i=1}^{n} u_i^2},$$

где

$$c_0 = 0,\ u_i = \begin{cases} \hat{s}(t_i) + \hat{s}(t_{k_i}) - 1, & c_i = 0 \\ \hat{s}(t_{k_i}) - 1, & c_i = 1 \end{cases},\quad S_P \frac{n(n-1)\sum_{i=1}^{n} u_i(1-v_i)}{n_1 n_2 \sum_{i=1}^{n} u_i^2},\ k_i = \max\left\{ j \mid j \in \{0,\dots,i-1\}, c_j = 0 \right\},$$

и $\hat{s}(t) = \prod_{i:t_i<t, c_i=0} \frac{n-i}{n-i+1}$ – непараметрическая оценка функции надежности Каплана–Мейера [9] функции $S(t)$ по объединенной выборке X.

Нулевая гипотеза H_0 отклоняется с достигнутым уровнем значимости α, если $|S_P| > z_{1-\alpha/2}$, где $z_{1-\alpha/2}$ – это $(1-\alpha/2)$-квантиль стандартного нормального распределения.

2.3. Логарифмический ранговый критерий

Статистика логарифмического рангового критерия может быть вычислена следующим образом [6, 10]:

$$S_L = \frac{\sum_{i=1}^{n}(1-c_i) - \sum_{j=1}^{i}\frac{1}{n-j+1}}{\sqrt{\left(\sum_{i=1}^{n}(1-c_i)\frac{n-i}{n-i+1}\right)\frac{n_1 n_2}{n(n-1)}}}.$$

Нулевая гипотеза H_0 отклоняется с достигнутым уровнем значимости α, если $|S_L| > z_{1-\alpha/2}$, где $z_{1-\alpha/2}$ – это $(1-\alpha/2)$-квантиль стандартного нормального распределения.

2.4. Критерий Кокса–Мантела

Критерий Кокса–Мантела [11] – это статистический критерий логарифмически рангового типа. Статистика этого критерия может быть вычислена следующим образом:

$$S_{CM} = \frac{r_2 - \sum_{i=1}^{n} (1-c_i) A_{(i)}}{\sqrt{\sum_{i=1}^{n} (1-c_i) A_{(i)} (1-A_{(i)})}}, \text{ где } r_2 = \sum_{i=1}^{n} v_i (1-c_i),\ A_{(i)} = \frac{1}{n-i} \sum_{j=i}^{n} v_j .$$

Нулевая гипотеза H_0 отклоняется с достигнутым уровнем значимости α, если $|S_{CM}| > z_{1-\alpha/2}$, где $z_{1-\alpha/2}$ – это $(1-\alpha/2)$-квантиль стандартного нормального распределения.

2.5. Взвешенный критерий Каплана–Мейера

Взвешенный критерий Каплана–Мейера [12] – это дистанционный статистический критерий. Основная идея заключается в измерении расстояния между оценками функций надежности. В данном критерии в качестве оценок используются оценки Каплана–Мейера [9]. Статистика критерия может быть вычислена следующим способом:

$$S_{WKM} = \frac{U}{\sqrt{\sigma}}, \text{ где } U = \sqrt{\frac{n_1 n_2}{n_1 + n_2}} \int_0^{t_n} w(t) \left(\hat{S}_1(t) - \hat{S}_2(t) \right) dt, \quad w(t) = \frac{(n_1 + n_2) \hat{F}_1^C(t-) \hat{F}_2^C(t-)}{n_1 \hat{F}_1^C(t-) + n_2 \hat{F}_2^C(t-)},$$

$$\sigma = -\int_0^{t_n} \left\{ \int_t^{t_n} w(\tau) \hat{S}(\tau) d\tau \right\}^2 \frac{n_1 \hat{F}_1^C(t-) + n_2 \hat{F}_2^C(t-)}{(n_1 + n_2) \hat{F}_1^C(t-) \hat{F}_2^C(t-)} \frac{d\hat{S}(t)}{\hat{S}(t) \hat{S}(t-)},$$

а $\hat{F}_i^C(t)$ и $\hat{S}_i(t)$ – оценки Каплана–Мейера соответствующих функций, описанных в разделе 2, а $\hat{S}(t)$ – оценка Каплана–Мейера по объединенной выборке X.

Нулевая гипотеза H_0 отклоняется с достигнутым уровнем значимости α, если $|S_{WKM}| > z_{1-\alpha/2}$, где $z_{1-\alpha/2}$ – это $(1-\alpha/2)$-квантиль стандартного нормального распределения.

2.6. *Q*-критерий

Мартинез и Норанхо предложили предварительную проверку [13], какой критерий предпочтительнее использовать: критерий Уилкоксона или логарифмический ранговый критерий. Мы используем этот критерий в качестве двухэтапного критерия для решения задачи однородности распределений. Статистика *Q*-критерия может быть вычислена следующим образом:

$$S_Q = \begin{cases} S_L, Q < 0, \\ S_P, Q \geq 0 \end{cases} \text{ где } Q = \left| \hat{S}_2(q_{0.6}) - \hat{S}_1(q_{0.6}) \right| - \left| \hat{S}_2(q_{0.2}) - \hat{S}_1(q_{0.2}) \right|, \quad q_p = \hat{S}_1^{-1}(p),$$

а $\hat{S}_i(t)$ – оценка Каплана–Мейера для $S_i(t)$.

Данный критерий имеет двустороннюю критическую область, следовательно, нулевая гипотеза H_0 отклоняется, если $|S_Q| > M$ при больших значениях M.

2.7. Критерий максимального значения

Как и *Q*-критерий, критерий максимального значения является двухэтапным критерием. Статистика критерия максимального значения может быть вычислена с помощью значений статистик S_G и S_L. Этот критерий был предложен нами [14, 15]. Статистика критерия может быть вычислена следующим образом:

$$S_{MAX} = \max\left(|S_G|, |S_L|\right),$$

где S_G – значение статистики критерия Гехана, S_L – значение статистики логарифмического рангового критерия.

Нулевая гипотеза H_0 отклоняется с достигнутым уровнем значимости α, если $|S_{MAX}| > G\left(S_{MAX} \mid H_0\right)_{1-\alpha}$, где $G\left(S_{MAX} \mid H_0\right)_{1-\alpha}$ – это $(1-\alpha)$-квантиль распределения статистик критерия при справедливости гипотезы H_0. Данное распределение описывается следующей функцией плотности:

$$f(t;r) = 4\varphi(t)\left(\Phi_0\left(t\sqrt{\frac{1-r}{1+r}}\right) + \Phi_0\left(t\sqrt{\frac{1+r}{1-r}}\right)\right), t \ge 0, \text{ где } \Phi_0(t) = \int_0^t \varphi(x)dx = \int_0^t \frac{\exp\left(-x^2/2\right)}{\sqrt{2\pi}}dx.$$

В соответствие с [15] функция плотности $f(t;r)$ может применяться в качестве наилучшей оценки предельного распределения статистик критерия максимального значения с параметром $r \in [0.84; 0.92]$, например, $r = 0.90$, если объемы выборок $n_1, n_2 > 50$ и $\alpha < 0.15$. В противном случае $G\left(S_{MAX} \mid H_0\right)$ следует получить с помощью методов компьютерного моделирования, например, метода Монте-Карло.

2.8. Критерий Крамера–Уэлча

Критерий Крамера–Уэлча [16] – это критерий для проверки гипотезы однородности средних. Кроме того, это непараметрический критерий и альтернатива параметрическому критерию Стьюдента [17]. В общем случае проверка гипотезы однородности распределений может осуществляться через проверку однородности характеристик распределений, например, математических ожиданий или дисперсий. Данная модификация критерия может использоваться для данных типа времени жизни. Статистика критерия может быть вычислена следующим образом:

$$S_{CW} = \frac{\sqrt{n_1 n_2}\left(\hat{\mu}_1 - \hat{\mu}_2\right)}{\sqrt{n_2\hat{\sigma}_1^2 + n_1\hat{\sigma}_2^2}},$$

где

$$\hat{\mu}_i = \sum_{j=1}^{n_i} t_{ij}\left(\hat{S}_i(t-) - \hat{S}_i(t+)\right), \quad \hat{\sigma}_i^2 = \frac{n_i}{n_i - 1}\left(\sum_{j=1}^{n_i} t_{ij}\left(\hat{S}_i(t-) - \hat{S}_i(t+)\right) - \hat{\mu}_i^2\right),$$

а $\hat{\mu}_i$, $\hat{\sigma}_i^2$, $\hat{S}_i(t)$ – оценки математического ожидания дисперсии и функции надежности соответственно для данных типа времени жизни.

Нулевая гипотеза H_0 отклоняется с достигнутым уровнем значимости α, если $|S_{CW}| > z_{1-\alpha/2}$, где $z_{1-\alpha/2}$ – это $(1-\alpha/2)$-квантиль стандартного нормального распределения.

2.9. Критерий Багдонавичуса–Никулина для однократных пересечений

Данный критерий основан на простой модели с пересечением (SCE-модель, [18]), которая предполагает, что функции надежности могут пересекаться один раз. Статистика критерия имеет следующий вид [18]:

$$S_{BN1} = \left(U_1, U_2\right)\Sigma^{-1}\left(U_1, U_2\right)^T,$$

где

$$U_1=\sum_{j:c_{1j}=0}\frac{Y_2(t_{1j})}{Y(t_{1j})}-\sum_{j:c_{2j}=0}\frac{Y_1(t_{2j})}{Y(t_{2j})},$$

$$U_2=-\sum_{j:c_{1j}=0}\frac{Y_2(t_{1j})}{Y(t_{1j})}\ln\left(1+\Lambda(t_{1j})\right)+\sum_{j:c_{2j}=0}\frac{Y_1(t_{2j})}{Y(t_{2j})}\ln\left(1+\Lambda(t_{2j})\right),$$

$$Y(t)=Y_1(t)+Y_2(t),\ Y_i(t)=\sum_{j=1}^{n_i}Y_{ij}(t),\ Y_{ij}(t)=1_{\left\{t_{ij}\ge t\right\}},\ \ \Lambda(t)=\sum_{i=1}^{2}\sum_{j:c_{ij}-0,t_{ij}\le t}\frac{1}{Y(t_{ij})}.$$

Элементы матрицы Σ следующие:

$$\sigma_{11}=\sum_{i=1}^{2}\sum_{j:c_{ij}=0}\frac{Y_1(t_{ij})Y_2(t_{ij})}{Y^2(t_{ij})},\ \sigma_{22}=\sum_{i=1}^{2}\sum_{j:c_{ij}=0}\frac{Y_1(t_{ij})Y_2(t_{ij})}{Y^2(t_{ij})}\ln^2\left(1+\Lambda(t_{ij})\right),$$

$$\sigma_{12}=\sigma_{21}=\sum_{i=1}^{2}\sum_{j:c_{ij}=0}\frac{Y_1(t_{ij})Y_2(t_{ij})}{Y^2(t_{ij})}\ln\left(1+\Lambda(t_{ij})\right).$$

Нулевая гипотеза H_0 отклоняется с достигнутым уровнем значимости α, если $S_{BN1}>\chi^2_{1-\alpha}(2)$, где $\chi^2_{1-\alpha}(2)$ – это $\left(1-\alpha\right)$-квантиль распределения хи-квадрат со степенью свободы 2.

2.10. Критерий Багдонавичуса–Никулина для многократных пересечений

Этот критерий основан на модели множественных пересечений (MCE-модель, [19]), которая предполагает, что функции надежности могут иметь одно пересечение или два пересечения. Статистика критерия имеет следующий вид [19]:

$$S_{BN2}=\left(U_1,U_2,U_3\right)\Sigma^{-1}\left(U_1,U_2,U_3\right)^T,$$

где

$$U_1=\sum_{j:c_{1j}=0}\frac{Y_2(t_{1j})}{Y(t_{1j})}-\sum_{j:c_{2j}=0}\frac{Y_1(t_{2j})}{Y(t_{2j})},\ U_2=-\sum_{j:c_{1j}=0}\frac{Y_2(t_{1j})}{Y(t_{1j})}\Lambda(t_{1j})+\sum_{j:c_{2j}=0}\frac{Y_1(t_{2j})}{Y(t_{2j})}\Lambda(t_{2j}),$$

$$U_3=-\sum_{j:c_{1j}=0}\frac{Y_2(t_{1j})}{Y(t_{1j})}\Lambda^2(t_{1j})+\sum_{j:c_{2j}=0}\frac{Y_1(t_{2j})}{Y(t_{2j})}\Lambda^2(t_{2j}).$$

Элементы матрицы Σ следующие:

$$\sigma_{11}=\sum_{i=1}^{2}\sum_{j:c_{ij}=0}\frac{Y_1(t_{ij})Y_2(t_{ij})}{Y^2(t_{ij})},\ \sigma_{12}=\sigma_{21}=\sum_{i=1}^{2}\sum_{j:c_{ij}=0}\frac{Y_1(t_{ij})Y_2(t_{ij})}{Y^2(t_{ij})}\Lambda(t_{ij}),$$

$$\sigma_{13}=\sigma_{31}=\sigma_{22}=\sum_{i=1}^{2}\sum_{j:c_{ij}=0}\frac{Y_1(t_{ij})Y_2(t_{ij})}{Y^2(t_{ij})}\Lambda^2(t_{ij}),$$

$$\sigma_{23}=\sigma_{32}=\sum_{i=1}^{2}\sum_{j:c_{ij}=0}\frac{Y_1(t_{ij})Y_2(t_{ij})}{Y^2(t_{ij})}\Lambda^3(t_{ij}),\ \sigma_{33}=\sum_{i=1}^{2}\sum_{j:c_{ij}=0}\frac{Y_1(t_{ij})Y_2(t_{ij})}{Y^2(t_{ij})}\Lambda^4(t_{ij}).$$

Нулевая гипотеза H_0 отклоняется с достигнутым уровнем значимости α, если $S_{BN1}>\chi^2_{1-\alpha}(3)$, где $\chi^2_{1-\alpha}(3)$ – это $\left(1-\alpha\right)$-квантиль распределения хи-квадрат со степенью свободы 3.

2.11. Критерий Багдонавичуса–Никулина для постоянного отношения

Данный статистический критерий проверки гипотезы однородности для модели, предполагающей, что отношение соответствующих функций интенсивности [20], является постоянным. Статистика критерия может быть вычислена следующим способом [21]:

$$S_{BN3}=\left(U_1,U_2\right)\Sigma^{-1}\left(U_1,U_2\right)^T,$$

где

$$U_i = \sum_{j=1}^{n_1}\left(1-c_{1j}\right)\frac{K_i\left(t_{1j}\right)}{Y\left(t_{1j}\right)}Y_2\left(t_{1j}\right) - \sum_{j=1}^{n_2}\left(1-c_{2j}\right)\frac{K_i\left(t_{2j}\right)}{Y\left(t_{2j}\right)}Y_1\left(t_{2j}\right), i=1,2.$$

Элементы матрицы Σ следующие:

$$\sigma_{ij} = \sum_{r=1}^{2}\sum_{s=1}^{n_r}\left(1-c_{rs}\right)K_i\left(t_{rs}\right)K_j\left(t_{rs}\right)\frac{Y_1\left(t_{rs}\right)Y_2\left(t_{rs}\right)}{Y^2\left(t_{rs}\right)}.$$

Авторами предлагается использовать следующие функции:

$$K_1(t)=\exp\left(-\Lambda(t)\right)\Big/\sqrt{n},\ \ K_2(t)=\left\{\exp\left(-\Lambda(t)\right)-1\right\}\Big/\sqrt{n},\ \ \Lambda(t)=\sum_{i=1}^{2}\sum_{j:c_{ij}=0,t_{ij}\le t} 1\Big/Y\left(t_{ij}\right),\ \ Y(t)=Y_1(t)+Y_2(t),$$

$$Y_i(t)=\sum_{j=1}^{n_i}Y_{ij}(t),\ \ Y_{ij}(t)=\begin{cases}1, t_{ij}\ge t,\\ 0, t_{ij}<t.\end{cases}$$

Нулевая гипотеза H_0 отклоняется с достигнутым уровнем значимости α, если $S_{BN3} > \chi^2_{1-\alpha}(2)$, где $\chi^2_{1-\alpha}(2)$ – это $\left(1-\alpha\right)$-квантиль распределения хи-квадрат со степенью свободы 2.

В следующем разделе мы рассмотрим критерий Вальда и свойства функции полезности в задачах принятия решений в условиях риска и неопределенности.

3. Критерий Вальда для принятия решений в условиях риска и неопределенности

Пусть r – это стратегия применения статистического критерия со статистикой S при альтернативной гипотезе H. Тогда каждой стратегии r может быть присвоена некоторая полезность, используя функцию полезности $U(r\,|\,H)$ [4] (или сокращенно $U(r)$), для которой будут справедливыми следующие утверждения:

1) стратегия r_i предпочтительнее стратегии r_j (обозначим как $r_i \succ r_j$), если $U\left(r_i\right) \ge U\left(r_j\right)$, где $U\left(r_i\right), U\left(r_j\right)$ – значения полезности стратегий r_i и r_j;
2) если $r_i \succ r_j$ и $r_j \succ r_k$, тогда $U\left(r_i\right) \ge U\left(r_j\right)$;
3) если r представима в виде $r=\left(1-\alpha\right)r_1+\alpha r_2$, тогда $U\left(r\right)=\left(1-\alpha\right)U\left(r_1\right)+\alpha U\left(r_2\right)$;
4) если $U\left(r_1,r_2\right)$ – это объединенная полезность r_1, r_2, тогда $U\left(r_1,r_2\right)=U\left(r_1\right)+U\left(r_2\right)$.

В данной работе мы используем мощность статистического критерия [1] со статистикой S при альтернативной гипотезе H в качестве функции полезности $U\left(r\right)$.

Критерий Вальда – это критерий для принятия решений в условиях риска и неопределенности [4]. Это критерий «осторожного наблюдателя», т.к. критерий оптимизирует полезность (мощность критерия) при выполнении предположения, что для принятия решения среда находится в самом неблагоприятном состоянии, т.е. значения мощности всех критериев минимальны. Правило может быть записано как:

$$W = \max_{i=\overline{1,k}}\min_{j=\overline{1,m}} U\left(r_i\,|\,H_j\right),$$

где k – это число рассматриваемых статистических критериев, m – число рассматриваемых альтернативных гипотез.

Фактически наилучший критерий по правилу Вальда – это статистический критерий, чья минимальная мощность среди всех альтернативных гипотез максимальна среди всех критериев.

4. Типы альтернативных гипотез

Мы делим множество близких альтернативных гипотез на 9 типов, как показано в табл. 1 и на рис. 1. Рассмотрим 3 типа альтернатив без пересечений функций надежности $S_1(t)$ и $S_2(t)$: с наибольшими отклонениями в ранние ($H_{01}-H_{03}$), средние ($H_{04}-H_{06}$) и поздние ($H_{07}-H_{09}$) моменты времени. Также рассмотрим 3 типа альтернатив с одной точкой пересечений функций надежности $S_1(t)$ и $S_2(t)$: точка пересечения в ранние ($H_{11}-H_{13}$), средние ($H_{14}-H_{16}$) и поздние ($H_{17}-H_{19}$) моменты времени. Кроме того, рассмотрим еще 3 типа альтернатив с двумя точек пересечений функций надежности $S_1(t)$ и $S_2(t)$: в ранние и средние ($H_{21}-H_{23}$), в ранние и поздние ($H_{24}-H_{26}$) и в средние и поздние ($H_{27}-H_{29}$) моменты времени.

Для каждого типа альтернатив рассмотрим по 3 альтернативных гипотезы со схожим поведением функций надежности, но на разных семействах распределений. В табл. 1 обозначения *Exp, We, LgN* и *Γ* обозначают семейства распределений (экспоненциальное, Вейбулла–Гнеденко, лог-нормальное и гамма соответственно), а μ_i и σ_i – это математическое ожидание и стандартное отклонение случайной величины с функцией надежности $S_i(t)$ соответственно. Функции плотности вероятностей для указанных случайных величин приведем ниже:

Вейбулла–Гнеденко, *We* $$f_{We}\left(t;\theta_1,\theta_2,\theta_3\right)=\theta_3\left(t-\theta_1\right)^{\theta_3-1}\exp\left(-\left(\frac{t-\theta_1}{\theta_2}\right)^{\theta_3}\right)\Big/\theta_2^{\theta_3},\, t\ge\theta_1;$$

экспоненциальное, *Exp* $$f_{Exp}\left(t;\theta_1,\theta_2\right)=\theta_1\exp\left(-\theta_1\left(t-\theta_2\right)\right),\, t\ge\theta_1;$$

гамма, *Γ* $$f_{\Gamma}\left(t;\theta_1,\theta_2\right)=\frac{\left(t/\theta_1\right)^{\theta_2-1}\exp\left(-t/\theta_1\right)}{\theta_1\Gamma\left(\theta_2\right)},\, t\ge 0;$$

лог-нормальное, *LgN* $$f_{LgN}\left(t;\theta_1,\theta_2\right)=\frac{\exp\left(-\left(\ln t-\theta_1\right)^2\Big/2\theta_2^2\right)}{t\sqrt{2\pi\theta_2^2}},\, t\ge 0.$$

Ранние моменты времени – это область функции надежности $S(t)$ между квантилями 1.0 и 0.67. Средние моменты времени – это область функции надежности $S(t)$ между квантилями 0.67 и 0.33. Поздние моменты времени – это область функции надежности $S(t)$ между квантилями 0.33 и 0.0.

Для того чтобы измерить отклонения между функциями надежности $S_1(t)$ и $S_2(t)$, мы используем норму L^p в пространстве Лебега [22] при $p=1$. В общем случае норма L^1 для функций надежности $S_1(t)$ и $S_2(t)$ может быть вычислена следующим образом:

$$L^1\left(S_1,S_2\right)=\int_{-\infty}^{+\infty}\left|S_1(t)-S_2(t)\right|dt.$$

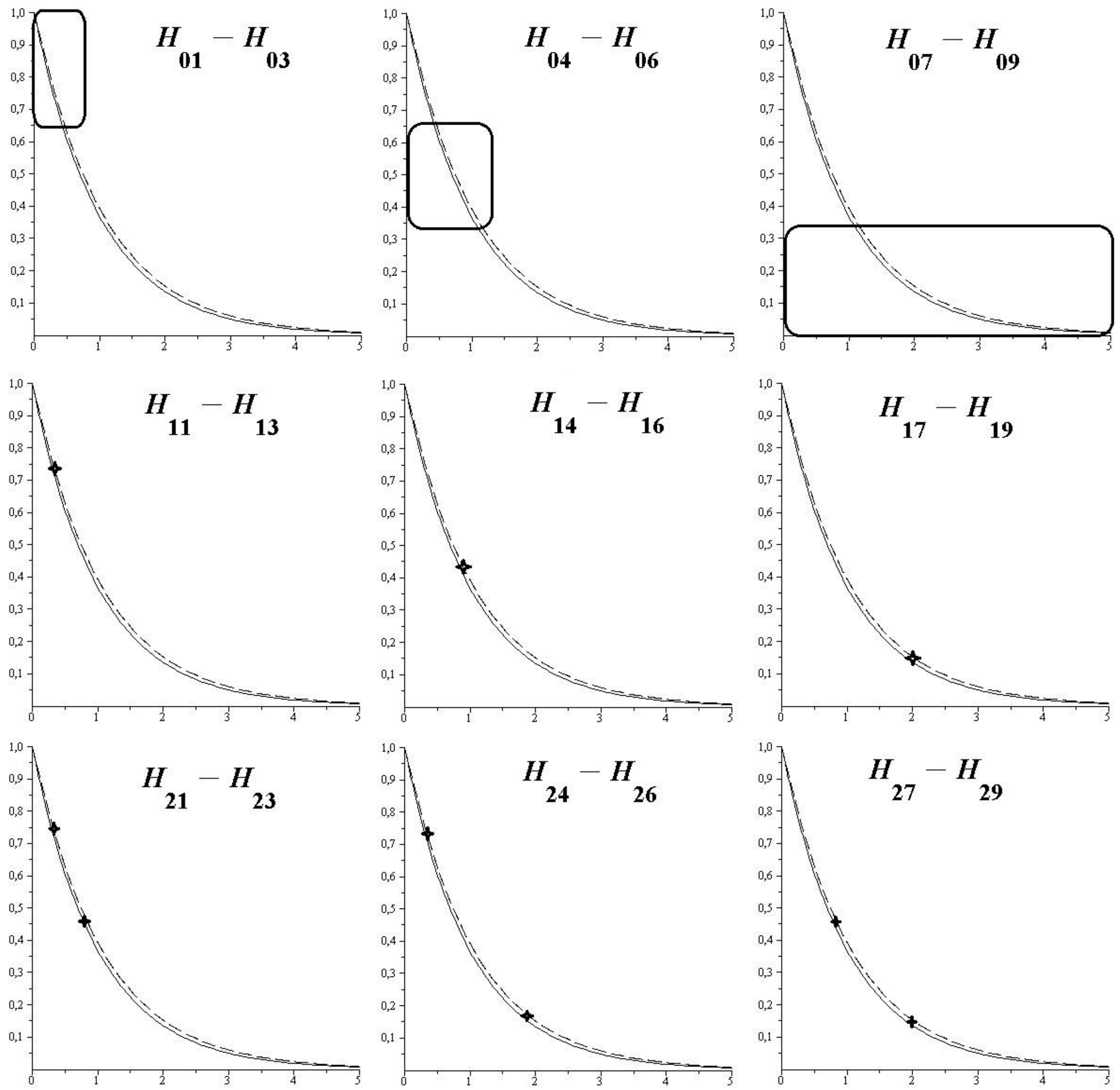

Рис. 1. Типы альтернативных гипотез

В данной работе для каждой альтернативной гипотезы общее отклонение между функциями надежности $S_1(t)$ и $S_2(t)$ в норме L^1 составляет приблизительно 0.10, т.е. $L^1(S_1, S_2) \approx 0.10$ (приблизительно, потому что параметры распределений были получены из решения нелинейного уравнения, которое имеет некоторую погрешность). Отклонение $L^1(S_1, S_2) \approx 0.10$ зафиксировано для того, чтобы сравнить, как тип альтернативы влияет на функцию полезности (мощность критерия). Значения L^1 для каждой альтернативной гипотезы представлены в табл. 1.

Таблица 1. Типы альтернативных гипотез

H_i	Тип альтернативы	$S_1(t)$	$S_2(t)$	Пересечения	$L^1(S_1,S_2)$	$\lvert\mu_1-\mu_2\rvert$	$\lvert\sigma_1-\sigma_2\rvert$
H_{01}	Различия в ранние моменты времени	*Exp*	*Exp*	–	0.099	0.10	0.10
H_{02}		*We*	*LgN*	–	0.096	0.10	0.66
H_{03}		*LgN*	*Exp*	–	0.109	0.18	0.00
H_{04}	Различия в средние моменты времени	Γ	*Exp*	–	0.075	0.07	0.00
H_{05}		*Exp*	*We*	–	0.100	0.10	0.08
H_{06}		*Exp*	*We*	–	0.162	0.16	0.16
H_{07}	Различия в поздние моменты времени	*Exp*	Γ	–	0.089	0.09	0.03
H_{08}		*We*	*Exp*	–	0.116	0.13	0.57
H_{09}		*We*	*LgN*	–	0.107	0.11	0.41
H_{11}	Точка пересечения в ранние моменты времени	*Exp*	*Exp*	0.363	0.100	0.09	0.05
H_{12}		Γ	Γ	0.763	0.078	0.06	0.35
H_{13}		*We*	*Exp*	0.571	0.125	0.12	0.00
H_{14}	Точка пересечения в средние моменты времени	*We*	Γ	0.611	0.081	0.08	0.01
H_{15}		*Exp*	*We*	0.843	0.097	0.04	0.12
H_{16}		Γ	*Exp*	1.040	0.068	0.04	0.00
H_{17}	Точка пересечения в поздние моменты времени	*We*	*Exp*	1.346	0.099	0.01	0.00
H_{18}		Γ	Γ	1.878	0.132	0.03	0.15
H_{19}		*We*	*Exp*	3.626	0.097	0.09	0.47
H_{21}	Точки пересечения в ранние и средние моменты времени	*LgN*	*We*	0.243, 0.655	0.071	0.01	0.04
H_{22}		*Exp*	*LgN*	0.814, 1.038	0.079	0.08	0.48
H_{23}		*We*	*Exp*	0.577, 1.327	0.105	0.09	0.53
H_{24}	Точки пересечения в ранние и поздние моменты времени	*LgN*	Γ	0.683, 3.074	0.095	0.06	0.13
H_{25}		*LgN*	*Exp*	0.232, 0.831	0.068	0.01	0.19
H_{26}		*We*	*LgN*	1.018, 2.381	0.131	0.06	0.60
H_{27}	Точки пересечения в средние и поздние моменты времени	*We*	*LgN*	1.254, 3.265	0.099	0.11	0.09
H_{28}		Γ	*LgN*	0.793, 2.994	0.088	0.07	0.11
H_{29}		*Exp*	*LgN*	1.321, 3.463	0.104	0.11	0.08

5. Результаты компьютерного моделирования

В нашей работе компьютерное моделирование происходило для одинаковых объемов выборок $n_1 = n_2 = 200$ и равной степени цензурирования в диапазоне 0 % – 50 %. Число повторений метода Монте-Карло составляет 100 000, т.к. по теореме Колмогорова [23] это обеспечивает точность моделирования распределения статистик критерия приблизительно 0.0051. Для вычисления мощности критерия ошибка первого рода $\alpha = 0.05$. В процессе моделирования использовались 3 семейства распределений моментов цензурирования $F^C(t)$: Вейбулла–Гнеденко, экспоненциальное и гамма.

Результаты представлены в виде таблиц, содержащих минимальные значения мощности критериев среди всех $F^C(t)$ с соответствующим значением критерия Вальда для степеней цензурирования 0 % и 50 %. Жирным шрифтом выделены значения, близкие (вычисление мощности критериев осуществляется со статистической погрешностью) к значению статистики критерия Вальда *W*.

Выбирая лучшую стратегию по критерию Вальда, исследователь выбирает критерий, который будет иметь наибольшую мощность в заведомо наихудшей ситуации, т.е. при выполнении предположения, что мощности всех критериев минимальны. Однако в общем случае, когда это предположение не выполнено, наилучшая стратегия по Вальду не обязательно соответствует критерию с наибольшей мощностью.

5.1. Функции надежности не пересекаются

Группа альтернатив, не имеющих точек пересечения функций надежности $S_1(t)$ и $S_2(t)$, содержит 3 типа: наибольшие отклонения в ранние ($H_{01} - H_{03}$), средние ($H_{04} - H_{06}$) и поздние ($H_{07} - H_{09}$) моменты времени. Результаты представим в табл. 2–4.

Таблица 2. Мощность критериев для типа альтернатив без пересечений и с наибольшим отклонением в ранние моменты времени

Статистика критерия	**Степень цензурирования 0 %**				**Степень цензурирования 50 %**			
	H_{01}	H_{02}	H_{03}	min	H_{01}	H_{02}	H_{03}	min
S_G	0.374	0.526	0.307	**0.307**	0.527	0.739	0.569	**0.527**
S_P	0.376	0.524	0.304	**0.304**	0.442	0.548	0.431	0.431
S_{CM}	0.162	0.306	0.176	0.162	0.263	0.296	0.204	0.204
S_L	0.169	0.305	0.175	0.169	0.263	0.299	0.209	0.209
S_Q	0.343	0.497	0.258	0.258	0.336	0.525	0.300	0.300
S_{MAX}	0.325	0.476	0.266	0.266	0.481	0.680	0.519	0.481
S_{BN1}	0.293	0.309	0.155	0.155	0.442	0.433	0.503	0.433
S_{BN2}	0.276	0.601	0.470	0.276	0.494	0.598	0.662	0.494
S_{BN3}	0.366	0.444	0.247	0.247	0.489	0.536	0.568	0.489
S_{WKM}	0.374	0.529	0.310	**0.310**	0.538	0.749	0.565	**0.538**
S_{CW}	0.170	0.621	0.206	0.170	0.115	0.398	0.089	0.089
				***W*=0.310**				***W*=0.538**

Как очевидно из полученных результатов, в случае наибольших отклонений в ранние моменты времени оптимально использовать критерий Гехана S_G и взвешенный критерий Каплана–Мейера S_{WKM}. Аналогичный результат получен и в случае наибольших отклонений в средние моменты времени. Однако в случае наибольших отклонений в поздние моменты времени оптимальным выбором является использование критерия Крамера–Уэлча S_{CW}.

Таблица 3. Мощность критериев для типа альтернатив без пересечений и с наибольшим отклонением в средние моменты времени

Статистика критерия	Степень цензурирования 0 %				Степень цензурирования 50 %			
	H_{04}	H_{05}	H_{06}	min	H_{04}	H_{05}	H_{06}	min
S_G	0.172	0.376	0.671	**0.172**	0.206	0.338	0.790	**0.206**
S_P	0.175	0.372	0.678	**0.175**	0.183	0.328	0.712	0.183
S_{CM}	0.094	0.239	0.346	0.094	0.132	0.265	0.505	0.132
S_L	0.094	0.240	0.346	0.094	0.139	0.268	0.496	0.139
S_Q	0.146	0.341	0.634	0.146	0.153	0.319	0.597	0.153
S_{MAX}	0.148	0.333	0.620	0.148	0.182	0.321	0.747	0.182
S_{BN1}	0.146	0.284	0.576	0.146	0.150	0.257	0.667	0.150
S_{BN2}	0.121	0.232	0.512	0.121	0.134	0.220	0.679	0.134
S_{BN3}	0.161	0.298	0.651	0.161	0.158	0.259	0.704	0.158
S_{WKM}	0.173	0.371	0.676	**0.173**	0.202	0.351	0.799	**0.202**
S_{CW}	0.098	0.247	0.363	0.098	0.088	0.156	0.223	0.088

W=0.175 **W =0.206**

Таблица 4. Мощность критериев для типа альтернатив без пересечений и с наибольшим отклонением в поздние моменты времени

Статистика критерия	Степень цензурирования 0 %				Степень цензурирования 50 %			
	H_{07}	H_{08}	H_{09}	min	H_{07}	H_{08}	H_{09}	min
S_G	0.213	0.108	0.100	0.100	0.171	0.065	0.080	0.065
S_P	0.219	0.108	0.099	0.099	0.174	0.072	0.071	0.071
S_{CM}	0.197	0.282	0.347	0.197	0.161	0.084	0.076	0.076
S_L	0.198	0.277	0.334	0.198	0.158	0.084	0.077	0.077
S_Q	0.212	0.219	0.251	**0.212**	0.171	0.075	0.073	0.073
S_{MAX}	0.215	0.228	0.276	**0.215**	0.169	0.075	0.079	0.075
S_{BN1}	0.161	0.342	0.560	0.161	0.130	0.082	0.067	0.067
S_{BN2}	0.133	0.285	0.608	0.133	0.118	0.094	0.119	0.094
S_{BN3}	0.166	0.329	0.482	0.166	0.132	0.081	0.065	0.065
S_{WKM}	0.215	0.107	0.102	0.102	0.172	0.065	0.079	0.065
S_{CW}	0.195	0.283	0.342	**0.195**	0.113	0.119	0.129	**0.113**

W=0.215 **W=0.113**

5.2. Функции надежности с одной точкой пересечения

Группа альтернатив, имеющих точку пересечения функций надежности $S_1(t)$ и $S_2(t)$, содержит 3 типа: пересечение в ранние ($H_{11}-H_{13}$), средние ($H_{14}-H_{16}$) и поздние ($H_{17}-H_{19}$) моменты времени. Результаты представим в табл. 5–7.

Таблица 5. Мощность критериев для типа альтернатив с одним пересечением в ранние моменты времени

Статистика критерия	**Степень цензурирования 0 %**				**Степень цензурирования 50 %**			
	H_{11}	H_{12}	H_{13}	min	H_{11}	H_{12}	H_{13}	min
S_G	0.065	0.057	0.062	0.057	0.053	0.050	0.059	0.050
S_P	0.065	0.057	0.062	0.057	0.050	0.049	0.051	0.049
S_{CM}	0.150	0.071	0.210	**0.071**	0.056	0.055	0.052	0.052
S_L	0.151	0.075	0.204	**0.075**	0.055	0.052	0.051	0.051
S_Q	0.119	0.065	0.162	**0.065**	0.053	0.053	0.051	0.051
S_{MAX}	0.125	0.067	0.169	**0.067**	0.062	0.056	0.064	0.056
S_{BN1}	0.177	0.069	0.339	**0.069**	0.133	0.071	0.143	**0.071**
S_{BN2}	0.139	0.065	0.257	**0.065**	0.137	0.077	0.132	**0.077**
S_{BN3}	0.194	0.073	0.359	**0.073**	0.146	0.079	0.150	**0.079**
S_{WKM}	0.066	0.057	0.061	0.057	0.058	0.052	0.058	0.052
S_{CW}	0.149	0.068	0.204	**0.068**	0.077	0.057	0.072	0.057

W=0.075 **W=0.079**

Таблица 6. Мощность критериев для типа альтернатив с одним пересечением в средние моменты времени

Статистика критерия	**Степень цензурирования 0 %**				**Степень цензурирования 50 %**			
	H_{14}	H_{15}	H_{16}	min	H_{14}	H_{15}	H_{16}	min
S_G	0.051	0.080	0.055	0.051	0.056	0.233	0.090	0.056
S_P	0.052	0.078	0.054	0.052	0.051	0.162	0.073	0.051
S_{CM}	0.105	0.071	0.063	0.063	0.049	0.068	0.053	0.049
S_L	0.105	0.070	0.062	0.062	0.050	0.067	0.053	0.050
S_Q	0.093	0.116	0.062	0.062	0.054	0.120	0.061	0.054
S_{MAX}	0.089	0.099	0.066	0.066	0.056	0.194	0.081	0.056
S_{BN1}	0.169	0.323	0.138	**0.138**	0.076	0.312	0.107	**0.076**
S_{BN2}	0.131	0.257	0.110	0.110	0.072	0.326	0.099	**0.072**
S_{BN3}	0.174	0.391	0.151	**0.151**	0.076	0.348	0.111	**0.076**
S_{WKM}	0.052	0.078	0.054	0.052	0.056	0.233	0.089	0.056
S_{CW}	0.111	0.069	0.063	0.063	0.052	0.045	0.052	0.045

W=0.151 **W=0.076**

Таблица 7. Мощность критериев для типа альтернатив с одним пересечением в поздние моменты времени

Статистика критерия	Степень цензурирования 0 %				Степень цензурирования 50 %			
	H_{17}	H_{18}	H_{19}	min	H_{17}	H_{18}	H_{19}	min
S_G	0.201	0.151	0.342	0.151	0.363	0.241	0.330	0.241
S_P	0.203	0.149	0.339	0.149	0.276	0.186	0.288	0.186
S_{CM}	0.049	0.054	0.142	0.049	0.147	0.113	0.180	0.113
S_L	0.047	0.054	0.144	0.047	0.149	0.116	0.175	0.116
S_Q	0.170	0.140	0.283	0.140	0.226	0.170	0.242	0.170
S_{MAX}	0.173	0.126	0.290	0.126	0.303	0.204	0.274	0.204
S_{BN1}	0.480	0.207	0.302	**0.207**	0.319	0.197	0.246	**0.197**
S_{BN2}	0.378	0.167	0.238	0.167	0.268	0.178	0.217	0.178
S_{BN3}	0.518	0.239	0.336	**0.239**	0.340	0.206	0.264	**0.206**
S_{WKM}	0.197	0.147	0.338	0.147	0.350	0.244	0.323	**0.244**
S_{CW}	0.051	0.059	0.154	0.051	0.077	0.057	0.118	0.057
				***W*=0.239**				***W*=0.244**

Это наиболее сложный тип альтернативных гипотез для критериев однородности, имеет точку пересечения в ранние моменты времени, т.к. значение статистики Вальда W минимальное. Можно использовать большинство критериев (например, критерии Багдонавичуса–Никулина или критерий максимального значения), но мощность может быть небольшой. В случае наличия пересечения функций надежности в средние или поздние моменты времени оптимальной стратегией является использование критериев Багдонавичуса–Никулина со статистиками S_{BN1} и S_{BN3}.

5.3. Функции надежности с двумя точками пересечений

Группа альтернатив, имеющих две точки пересечений функций надежности $S_1(t)$ и $S_2(t)$, содержит 3 типа: пересечения в ранние и средние ($H_{21}-H_{23}$), ранние и поздние ($H_{24}-H_{26}$) и средние и поздние ($H_{27}-H_{29}$) моменты времени. Результаты представим в табл. 8–10.

В том случае, если конкурирующие гипотезы имеют два пересечения, то наиболее оптимальной стратегией является использование критерия Багдонавичуса–Никулина со статистикой S_{BN2}, т.к. значения мощности были близки к значению статистики Вальда на всех типах альтернатив с двумя пересечениями функций надежности. Кроме того, в случае пересечений в ранние и средние моменты времени может также использоваться критерий Багдонавичуса–Никулина со статистикой S_{BN1}.

Таблица 8. Мощность критериев для типа альтернатив с двумя пересечениями в ранние и средние моменты времени

Статистика критерия	Степень цензурирования 0 %				Степень цензурирования 50 %			
	H_{21}	H_{22}	H_{23}	min	H_{21}	H_{22}	H_{23}	min
S_G	0.065	0.099	0.050	0.050	0.051	0.109	0.055	0.051
S_P	0.065	0.097	0.051	0.051	0.056	0.088	0.053	0.053
S_{CM}	0.052	0.121	0.169	0.052	0.064	0.077	0.050	0.050
S_L	0.054	0.120	0.163	0.054	0.063	0.080	0.050	0.050
S_Q	0.055	0.120	0.135	0.055	0.059	0.085	0.050	0.050
S_{MAX}	0.061	0.110	0.140	0.061	0.060	0.096	0.060	**0.060**
S_{BN1}	0.082	0.100	0.393	**0.082**	0.064	0.071	0.113	**0.064**
S_{BN2}	0.147	0.091	0.303	**0.091**	0.070	0.086	0.099	**0.070**
S_{BN3}	0.065	0.099	0.389	0.065	0.067	0.078	0.115	**0.067**
S_{WKM}	0.064	0.097	0.051	0.051	0.050	0.101	0.055	0.050
S_{CW}	0.052	0.137	0.162	0.052	0.055	0.086	0.073	0.055
				***W*=0.091**				***W*=0.070**

Таблица 9. Мощность критериев для типа альтернатив с двумя пересечениями в ранние и поздние моменты времени

Статистика критерия	Степень цензурирования 0 %				Степень цензурирования 50 %			
	H_{24}	H_{25}	H_{26}	min	H_{24}	H_{25}	H_{26}	min
S_G	0.051	0.052	0.049	0.049	0.068	0.055	0.090	0.055
S_P	0.051	0.052	0.050	0.050	0.053	0.048	0.058	0.048
S_{CM}	0.048	0.056	0.068	0.048	0.050	0.052	0.048	0.048
S_L	0.047	0.055	0.066	0.047	0.051	0.052	0.050	0.050
S_Q	0.046	0.050	0.057	0.046	0.052	0.051	0.055	0.051
S_{MAX}	0.050	0.053	0.061	0.050	0.073	0.056	0.081	0.056
S_{BN1}	0.085	0.054	0.154	0.054	0.096	0.079	0.053	0.053
S_{BN2}	0.272	0.140	0.376	**0.140**	0.207	0.098	0.180	**0.098**
S_{BN3}	0.056	0.054	0.098	0.054	0.120	0.086	0.063	0.063
S_{WKM}	0.051	0.051	0.051	0.051	0.067	0.055	0.091	0.055
S_{CW}	0.061	0.053	0.089	0.053	0.043	0.053	0.061	0.043
				***W*=0.140**				***W*=0.098**

Таблица 10. Мощность критериев для типа альтернатив с двумя пересечениями в средние и поздние моменты времени

Статистика критерия	**Степень цензурирования 0 %**				**Степень цензурирования 50 %**			
	H_{27}	H_{28}	H_{29}	min	H_{27}	H_{28}	H_{29}	min
S_G	0.239	0.058	0.283	0.058	0.568	0.125	0.627	0.125
S_P	0.238	0.058	0.276	0.058	0.400	0.086	0.456	0.086
S_{CM}	0.093	0.054	0.098	0.054	0.164	0.052	0.190	0.052
S_L	0.095	0.054	0.098	0.054	0.162	0.051	0.189	0.051
S_Q	0.230	0.059	0.272	0.059	0.293	0.070	0.342	0.070
S_{MAX}	0.197	0.054	0.234	0.054	0.504	0.114	0.585	0.114
S_{BN1}	0.135	0.055	0.179	0.055	0.535	0.144	0.611	0.144
S_{BN2}	0.446	0.263	0.488	**0.263**	0.692	0.272	0.749	**0.272**
S_{BN3}	0.258	0.052	0.327	0.052	0.608	0.184	0.686	0.184
S_{WKM}	0.241	0.057	0.284	0.057	0.571	0.131	0.645	0.131
S_{CW}	0.111	0.066	0.111	0.066	0.070	0.055	0.072	0.055
				***W*=0.263**				***W*=0.272**

6. Заключение

В итоге нами было получено, что в случае отсутствия пересечений функций надежности оптимальными стратегиями является использование критериев Гехана, взвешенного критерия Каплана–Мейера, и критерий Крамера–Уэлча. В случае наличия одной точки пересечения в разные моменты времени оптимальной стратегией является использование критериев Багдонавичуса–Никулина со статистиками S_{BN1} и S_{BN3}, в случае двух точек пересечений функций надежности – S_{BN2}. Данные результаты могут использоваться в качестве готовой инструкции по применению критериев на практике.

Литература

1. *Philonenko P., Postovalov S.* A power comparison of homogeneity tests for randomly censored data / P. Philonenko, S. Postovalov // Applied methods of statistical analysis. Applications in survival analysis, reliability and quality control AMSA2013, Novosibirsk, 25-27 Sept. 2013: proc. of the intern. workshop. Novosibirsk: NSTU publ., 2013. P. 227–237.
2. *Philonenko P., Postovalov S.* A Comparison of Homogeneity Tests for Different Alternative Hypotheses // Statistical Models and Methods for Reliability and Survival Analysis: monograph. London: Wiley-ISTE, 2013. Chap. 12. P. 177–194.
3. *Wald A.* Statistical decision functions which minimize the maximum risk. The Annals of Mathematics. 1945. V. 46, № 2. P. 265–280.
4. *Johnson E. J., Payne J. W.* Effort and accuracy in choice. Management Science. 1985. V. 31, № 4. P. 395–414.
5. *Wilcoxon F.* Individual comparisons by ranking methods. Biometrics Bulletin. 1945. V. 1, № 6. P. 80–83.
6. *Lee E. T., Wang J. W.* Statistical Methods for Survival Data Analysis. N. Y: Wiley, 2003.

7. *Gehan E. A.* A generalized Wilcoxon test for comparing arbitrarily singly-censored samples // Biometrika. 1965. V. 52, № 1/2. P. 203–223.
8. *Peto R., Peto J.* Asymptotically efficient rank invariant test procedures // J. Royal Statist. Soc. Ser. A (General). 1972. V. 135, № 2. P. 185–207.
9. *Kaplan E. L. and Meier P.* Nonparametric estimator from incomplete observation // J. Amer. Statist. Assoc. 1958. V. 53. P. 457–481.
10. *Savage I. R.* Contributions to the Theory of Rank Order Statistics: The Two Sample Case // Annals of Mathematical Statistics. 1956. V. 27. P. 590–615.
11. *Mantel N.* Evaluation of Survival Data and Two New Rank Order Statistics Arising in Its Consideration // Cancer Chemotherapy Reports. 1966. V. 50. P. 163–170.
12. *Pepe M. S., Fleming T. R.* Weighted Kaplan-Meier statistics: A class of distance tests for censored survival data // Biometrics. 1989. V. 45. P. 497–507.
13. *Ruvie Lou Maria C. Martinez, Joshua D. Naranjo.* A pretest for choosing between logrank and wilcoxon tests in the two-sample problem // International Journal of Statistics. 2010. V. LXVIII, № 2, P. 111–125.
14. *Philonenko P., Postovalov S.* A new two-sample test for choosing between log-rank and Wilcoxon tests with right-censored data // Journal of Statistical Computation and Simulation. 2015. V. 85, № 14. P. 2761–2770.
15. *Philonenko P., Postovalov S., Kovalevskii A.* The limit test statistic distribution of the maximum value test for right-censored data // Journal of Statistical Computation and Simulation. 2016. V. 86, № 17. P. 3482–3494.
16. *Welch B. L.* The generalization of Student`s problem when several different population variances are involved // Biometrika. 1947. V. 34. P. 29–35.
17. *Student.* The probable error of a mean. // Biometrika. 1908. V. 6, № 1. P. 1–25.
18. *Bagdonavičus V. B., Levuliene R. J., Nikulin M. S.* Zdorova-Cheminadeo “Tests for equality of survival distributions against non-location alternatives” // Lifetime Data Analysis. 2004. V. 10, № 4. P. 445–460.
19. *Bagdonavičus V. B., Nikulin M.* On goodness-of-fit tests for homogeneity and proportional hazards // Applied Stochastic Models in Business and Industry. 2006. V. 22, № 1. P. 607–619.
20. *Bagdonavičus V. B., Kruopis J., Nikulin M. S.* Nonparametric tests for censored data. John Wiley & Sons, Inc., New York, 2010, 233 p.
21. *Bagdonavičus V. B., Kruopis J., Nikulin M. S.* Censored and Truncated Data, in Non-parametric Tests for Censored Data. John Wiley & Sons, Inc, Hoboken, NJ, USA. 2013.
22. *Titchmarsh E.C.* The theory of functions. Oxford University Press, 1976.
23. *Kolmogorov A.* Some Works of Recent Years in the Field of Limit Theorems in the Theory of Probability // Vestnik Moskov, Univ. Ser. Fiz.-Mat. Estest. Nauk. 1953. V. 8. P. 29–38.

Статья поступила в редакцию 14.09.2016

Филоненко Петр Александрович

аспирант кафедры теоретической и прикладной информатики НГТУ (630073, Новосибирск, пр. Карла Маркса, 20), тел. (383) 346-06-00, e-mail: petr-filonenko@mail.ru.

Постовалов Сергей Николаевич

д.т.н., доцент кафедры теоретической и прикладной информатики НГТУ (630073, Новосибирск, пр. Карла Маркса, 20), тел. (383) 346-06-00, e-mail: postovalov@ngs.ru.

Homogeneity test power as utility function in the theory of decision making under risk and uncertainty

P. Philonenko, S. Postovalov

There are a lot of statistical tests for hypothesis testing and some statistical tests are preferable than others in a certain alternative hypothesis. It needs some reliable method for selecting the powerful statistical test. To solve the problem, we create types of alternative hypotheses (with different number of intersection points of reliability function). For every type we create some alternative hypotheses (with various distribution of survival function) with similar behavior of survival functions (the behavior of test power is similar too) to simulate the power of statistical tests and then we apply the Wald test to conclude what test is the most preferable in a certain type of alternative hypotheses.

Keywords: two-sample problem, lifetime data, Wald test, Monte-Carlo method.