УДК 621.382

ОЦЕНКА ПАРАМЕТРОВ РАСПРЕДЕЛЕНИЯ ВРЕМЕНИ ЗАПАЗДЫВАНИЯ ЗАЖИГАНИЯ ГАЗОРАЗРЯДНЫХ ИНДИКАТОРОВ

А. Н. Шестеркин, д.т.н., профессор кафедры ВПМ РГРТУ, Рязань, Россия;
orcid.org/0000-0002-0633-0591, neon60@inbox.ru

Рассмотрены способы вычисления оценок математического ожидания и смещения времени запаздывания зажигания при использовании для его описания экспоненциального распределения. ***Цель работы*** *– анализ способов вычисления статистических характеристик времени запаздывания зажигания. Рассмотрены классические методы вычисления параметров экспоненциального распределения – метод моментов и максимального правдоподобия. При измерении времени запаздывания в условиях сильной ионизации предложено результаты измерений (выборку) подвергать цензурированию слева, а оценку параметров распределения проводить на основе остальных, не цензурированных измерений. При исследовании времени запаздывания в режимах, когда статистическое время запаздывания может быть очень большим, среднее время запаздывания предложено вычислять на основе выборки, цензурированной справа. Проанализированы различные способы оценки математического ожидания времени запаздывания на основе порядковых статистик. Рассмотрена возможность ориентировочной оценки среднего значения на основе гистограммы. Приведены сведения об устройствах, реализующих предложенные методы оценки параметров. Достоверность рассмотренных способов вычисления оценок параметров экспоненциального распределения подтверждена путем статистического моделирования.*

Ключевые слова: *время запаздывания разряда, экспоненциальное распределение, оценка параметров распределения, объем выборки, порядковые статистики, устройства оценки параметров.*

DOI: 10.21667/1995-4565-2022-79-141-153

Введение

В работе [1] сформулированы рекомендации по определению числа измерений, обеспечивающих вычисление с заданной точностью и достоверностью оценок математического ожидания и смещения экспоненциально распределенных значений времени запаздывания зажигания, вычислены доверительные интервалы этих статистик. В зависимости от метода проведения измерений (плана испытаний), достоверности зарегистрированных значений, задач, стоящих перед исследователем и др., перечисленные статистики могут быть вычислены различными способами. Поэтому при проведении исследований не менее важно рационально выбрать метод и устройство, его реализующее, которые обеспечат вычисление статистик, наиболее объективно характеризующих исследуемый процесс.

Цель работы: анализ способов (устройств) вычисления оценок смещения и среднего значения времени запаздывания зажигания элементов отображения газоразрядных матричных индикаторов.

Теоретические исследования

В большинстве случаев вычисление параметров распределения времени запаздывания τ на основе совокупности измеренных значений $t_i (i = 1,\ 2,\ \ldots,N)$ осуществляют методом моментов или максимального правдоподобия.

Для несмещенного экспоненциального распределения с плотностью распределения $f(\tau) = \frac{1}{m} \cdot \exp\left(-\frac{\tau}{m}\right)$ оценка математического ожидания $\hat{m}$, равная выборочному среднему значению, вычисляется для обоих методов по классической формуле:

$$\hat{m} = \frac{\sum_{i=1}^{N} t_i}{N}. \qquad (1)$$

Эта оценка состоятельная, несмещенная и эффективная (имеет наименьшую дисперсию). Так как для экспоненциального распределения математическое ожидание и среднее квадратическое (стандартное) отклонение равны, то оценку стандартного отклонения, равного корню квадратному из несмещенной оценки дисперсии $\hat{S}^2 = \frac{1}{N-1}\sum_{i=1}^{N}\left(t_i - \hat{m}\right)^2$, вычисляют лишь в некоторых случаях для косвенного подтверждения экспоненциальности распределения совокупности результатов измерений.

Если время формирования разряда сопоставимо со статистическим временем задержки, то используют смещенное показательное распределение с плотностью распределения $f(\tau) = \frac{1}{m}\cdot\exp\left(-\frac{(\tau - c)}{m}\right)$, $\tau > c$, c – смещение. Для смещенного распределения оценки параметров указанными методами вычисляются по различным формулам [2]. При использовании метода максимального правдоподобия оценка смещения равна:

$$\hat{c}_0 = min(t_{1,}\ t_2, \ldots, t_N).$$

Эта оценка $\hat{c}_0$ является смещенной. Несмещенная оценка вычисляется по формуле [3]:

$$\hat{c}_1 = \hat{c}_1 - 1/N. \qquad (2)$$

Оценка математического ожидания, вычисленная методом максимального правдоподобия, равна:

$$\hat{m}_1 = \hat{m} - \hat{c}_1. \qquad (3)$$

Здесь $\hat{m}$ – выборочное среднее, определяемое по формуле (1). При использовании метода моментов оценка математического ожидания $\hat{m}_2$ равна квадратному корню из несмещенной оценки дисперсии $\hat{S}^2$, а оценка смещения $\hat{c}_2$ – разности между выборочным средним и оценкой математического ожидания $\hat{m}_2$:

$$\hat{m}_2 = \sqrt{\hat{S}^2}, \quad \hat{S}^2 = \frac{1}{N-1}\sum_{i=1}^{N}\left(t_i - \hat{m}\right)^2, \quad \hat{c}_2 = \hat{m} - \hat{m}_2. \qquad (4)$$

Для вычисления оценки дисперсии $\hat{S}^2$ часто используют более удобное выражение:

$$\hat{S}^2 = \frac{N}{N-1}\left(\frac{1}{N}\sum_{i=1}^{N} t_i^2 - \hat{m}^2\right).$$

Вычисленными оценками следует пользоваться, если известна их точность. Значения верхней m_H и нижней m_L границ доверительного интервала для математического ожидания m при экспоненциальном распределении можно определить по формулам [4]:

$$m_H = \frac{2\hat{m}N}{\chi^2_{/2}(2N+1)}, \quad m_L = \frac{2\hat{m}N}{\chi^2_{1-/2}(2N+1)}.$$

Здесь $\chi^2_s(v)$ – квантиль χ^2 – распределения, s – уровень доверия, для которого вычисляются границы. При «инженерной» достоверности 0,95 и объеме выборки, состоящей из 300÷350 элементов, границы доверительных интервалов отличаются от оцениваемого математического ожидания примерно на 10 %.

Длина доверительного интервала для оценки сдвига, найденной методом максимального правдоподобия, равна $-\ln(\alpha)/N$ [5]. Этот интервал откладывают влево от минимального значения, а истинным значением сдвига считают середину доверительного интервала. Длина этого интервала при $\alpha = 0{,}05$ для объемов выборки ≅300 элементов меньше одного процента от оценки смещения, т.е. оценка вычисляется достаточно точно.

Рассмотренные оценки и многочисленные устройства, позволяющие их вычислить, используют, если достоверны все зарегистрированные значения времени запаздывания. При исследовании временных характеристик газоразрядных индикаторов в ряде случаев это обеспечить трудно. В частности, достаточно трудно достоверно измерить малые длительности статистического времени запаздывания зажигания в условиях сильной ионизации элементов, например при «подсвете» исследуемого элемента отображения горящим или ранее горевшим, вблизи расположенным элементом.

Действительно, для измерения времени запаздывания зажигания на исследуемый элемент отображения подают стимулирующее напряжение – сотни вольт. Если определяют время запаздывания при «подсвете» горящими или горевшими элементами, то в это же время аналогичные высоковольтные сигналы подают и на «подсвечивающие» элементы. При зажигании элемента отображения в нем возникает ток разряда, который регистрируется для измерения времени запаздывания. Из-за переходных процессов, возникающих при коммутации высоковольтных напряжений, значительных емкостных и индуктивных связей между электродами матричного индикатора на электродах исследуемого элемента непосредственно после формирования стимулирующих сигналов действуют помехи. Амплитуда этих помех по величине сравнима, а в ряде случаев больше, чем ток разряда, протекающий через элемент отображения. В таких режимах сравнимы также длительность существования указанных помех (переходных процессов) и значительное число возможных случайных значений времени запаздывания. Например, если среднее время запаздывания составляет 3 мкс (режим «подсвета» близко расположенным элементом отображения), то в течение первой микросекунды, когда действуют значительные помехи, будет происходить в среднем ≈28 % зажиганий. Аналогичные проблемы возникают и при регистрации момента зажигания фотоэлектрическими методами. Так как измерение времени запаздывания с такими длительностями при действии значительных помех трудно провести с высокой достоверностью, то в дальнейшем на их основе с недостаточной достоверностью будут определены и параметры распределения.

Если самые малые результаты измерений недостаточно достоверны, то при вычислении оценок параметров распределения их целесообразно исключить из рассмотрения, т.е. результаты измерений (выборку) подвергнуть цензурированию слева, а оценку параметров распределения проводить на основе остальных, не цензурированных измерений. Для выборки, элементы которой имеют экспоненциальное распределение и цензурированы слева (исключены малые значения времени запаздывания), оценки смещения $\hat{\tau}_{zl}$ и среднего значения $\hat{m}_{zl}$ можно вычислить по формулам [6]:

$$\hat{\tau}_{zl} = C\left\{\left[\frac{1}{C} + (N-r)\cdot a_{r+1}\right] t_{(r+1)} - a_{r+1}\cdot\sum_{i=r+1}^{N} t_{(i)}\right\}, \tag{5}$$

$$\hat{m}_{zl} = C\left[\sum_{i=r+1}^{N} t_{(i)} - (N-r)\cdot t_{(r+1)}\right]. \tag{6}$$

Здесь $C = \dfrac{1}{N-r-1}$, r – число цензурированных измерений, $t_{(j)}$ – j-я порядковая статистика, $a_{r+1} = \sum\limits_{i=1}^{r+1} \dfrac{1}{N-i+1}$. Оценки параметров, вычисленные по формулам (5) и (6), являются оптимальными [6].

Из формул (5) и (6) следует, что для вычисления первого слагаемого оценки смещения и второго слагаемого оценки среднего значения необходимо определить $t_{(r+1)}$ порядковую статистику, т.е. в достоверных результатах измерений выделить минимальное значение. Эта процедура, особенно при реализации её аппаратным способом, достаточно громоздкая и требует дополнительного времени. Оценим возможность ее исключения.

Вычислим средний временной интервал между соседними порядковыми статистиками (спейсинг). Вероятность зажигания элемента отображения при экспоненциальном распределении за время t равна $P(t) = 1 - \exp(-t/m)$, m – среднее время запаздывания. При расчетах удобно использовать относительное время подключения $t_o = t/m$. Если относительное время подключения элемента $t_o =$ 0,200, то вероятность его зажигания равна 0,1813. Для «инженерного» объема выборки из 300 элементов за это время произойдет в среднем 54,38 зажиганий. Большее на единицу среднее число зажиганий, т.е. в среднем следующее зажигание, происходит при увеличении относительной длительности на 0,004 (до 0,204, среднее число зажиганий 55,36). Если $t_o =$ 0,400, то увеличение числа зажиганий на 1 (от 98,90 до 99,91) происходит при увеличении относительной длительности подключения до 0,405.

Для выборки, содержащей 100 элементов, увеличение среднего числа зажиганий на единицу происходит при увеличении относительного времени подключения соответственно от 0,200 до 0,212 или от 0,400 до 0,415. Таким образом, средний временной интервал между соседними начальными порядковыми статистиками не превышает нескольких процентов и вместо $\tau_{(r+1)}$ можно принять максимальное значение цензурированной статистики τ_m, практически это время цензурирования. Более того, если известен временной интервал, в течение которого происходит определенное число зажиганий, то из формулы $N[\exp(-\tau_m) - -\exp(-\tau_{r+1})] = 1$ можно приближенно определить значение следующей порядковой статистики τ_{r+1}.

Для вычисления длительности τ_{r+1} можно использовать простую и достаточно точную формулу $\tau_{r+1} \cong \tau_m + 1/[N \cdot \exp(-\tau_m)]$. Нетрудно убедиться, что τ_{r+1} незначительно превышает τ_m. Например, для объема выборки 300 элементов, эти длительности отличаются на 2,0 % при $t_o =$ 0,200 и на 1,2 % при $t_o =$ 0,400. Отметим также, что возможная погрешность оценки среднего значения из-за замены в формуле (6) $\tau_{(r+1)}$ на τ_m или τ_{r+1} будет существенно меньше, чем разность $\tau_{(r+1)}$ и τ_m или τ_1. Это обусловлено тем, что второе слагаемое в формуле (6) в несколько раз меньше, чем первое. Таким образом, для вычисления параметров, особенно при объемах выборки, обеспечивающих исследования с «инженерной» точностью, в формулах (5) и (6) $\tau_{(r+1)}$ можно заменить на τ_m или τ_{r+1}. Это позволяет исключить процедуру нахождения в достоверных результатах измерений минимального значения $-\tau_{(r+1)}$ порядковую статистику, т.е. не ранжировать достоверные результаты измерений, следовательно, построить более простое устройство оценки параметров распределения.

Практически цензурирование недостаточно достоверных результатов измерений времени запаздывания, имеющих длительность, сравнимую с длительностью помех, можно обеспечить, например, блокировкой измерителя времени запаздывания в течение некоторого временного интервала после подачи стимулирующих сигналов. Если в момент окончания этого интервала (помехи к этому времени значительно уменьшатся) элемент отображения горит, то фиксируется лишь факт зажигания. В противоположной ситуации время запаздывания, отсчитываемое от момента подачи стимулирующих сигналов, измеряется и регистрируется традиционным способом. Таким образом, будут зарегистрированы достоверно измеренные значения времени запаздывания и число недостоверных измерений. Максимальное значение цензурированной статистики τ_m в этом случае равно длительности блокировки записи недостоверно измеренных значений. Такой способ оценки параметров распределения времени запаздывания реализован в устройстве [7]. Цензурирование наименьших элементов выборки позволяет также снизить требования к устройству измерения времени, в частности к узлу, который обеспечивает определение момента возникновения разряда.

При исследовании времени запаздывания в «критических» режимах возбуждения, например с малыми перенапряжениями, статистическое время запаздывания может быть очень большим (сотни микросекунд) и поэтому зажигание исследуемого элемента может произойти не во всех циклах формирования стимулирующих сигналов. Например, если максимальное время возбуждения элемента установить равным утроенному ожидаемому среднему времени запаздывания, то за это время в среднем элемент отображения может не зажечься примерно в 5 % случаев. В таких ситуациях исследователи либо повторяют эксперимент, как правило, с лучшими условиями зажигания элементов, либо используют результаты лишь тех циклов, в которых произошло зажигание элемента отображения. Игнорирование таких незажиганий для рассмотренного примера может привести к занижению среднего значения времени запаздывания ≅ на 18 %! Отметим, что в некоторых случаях максимальное время возбуждения элемента не может быть выбрано сколь угодно большим, т.е. принципиально ограничено, например, при исследовании времени запаздывания при различных частотах возбуждения элементов отображения. Кроме того, увеличение максимального времени возбуждения элемента в «критических» режимах приводит к значительному увеличению общего времени исследований.

Ситуация, при которой в некоторых циклах формирования стимулирующих сигналов элемент отображения не загорается, т.е. неизвестны большие (максимальные) значения времени запаздывания зажигания, эквивалентна цензурированию выборки справа с несмещенным экспоненциальным распределением (временем формирования разряда в этом случае можно пренебречь). Оценка среднего значения времени запаздывания для такой цензурированной справа выборки, когда неизвестны r_b значений времени запаздывания зажигания, больших, чем максимальное время возбуждения *T,* можно вычислить по формуле [8]:

$$\widehat{m}_{zp} = \frac{\sum_{i=1}^{N-r_b} t_i + r_b \cdot T}{N - r_b}. \tag{7}$$

Эта оценка наилучшая: несмещенная, эффективная и достаточная (содержит информацию, достаточную для статистических выводов о распределении) [8]. Очевидно при зажигании элемента отображения во всех циклах формирования стимулирующих сигналов ($r_b = 0$) получим классическую формулу для оценки среднего значения.

Практическая реализация такого способа оценки среднего времени запаздывания сводится к регистрации всех длительностей зажиганий и числа случаев незажигания исследуемого элемента при максимальном времени формирования стимулирующих сигналов *T.* Такой способ и устройство оценки среднего времени запаздывания защищены патентом [9].

Оценка среднего значения может быть вычислена на основе порядковых статистик. Для случайной величины, имеющей экспоненциальное распределение, оценку среднего значения можем вычислить всего лишь по одной $t_{(n)}$ – *n*-й порядковой статистике ряда, используя формулу [6,8]:

$$\widehat{m}_p = t_{(n)} \cdot K_n, \quad K_n = \left(\sum_{i=1}^{n} \frac{1}{N - i + 1} \right)^{-1}. \tag{8}$$

Эффективность этой оценки по сравнению выборочным средним значением больше 0,98 при $n/N \le 1/2$ и 0,96 при $n/N \le 2/3$ [6], т.е. при использовании порядковых статистик с большими номерами эффективность оценки уменьшается. Если для оценки среднего значения использовать начальные порядковые статистики, которые измерены с некоторой погрешностью, то применение формулы (8) может привести к существенной ошибке из-за больших значений коэффициента K_n. Действительно, для первой порядковой статистики $K_1 = N$, для второй $K_2 = \dfrac{N(N-1)}{2N-1} \cong \dfrac{N}{2}$, для третьей $K_3 \cong \dfrac{N}{3}$ и т.д. Таким образом, при выборе порядковой статистики, на основе которой будет оцениваться среднее значение, нецелесообразно брать как начальные, так и последние статистики вариационного ряда.

Возможно более простое вычисление оценки среднего значения на основе ранжированного ряда. Для этого следует выбрать статистику с номером *n*, для которой коэффициент K_n

равен единице. Определим такое значение *n*. Элементы знаменателя коэффициента K_n представляют собой расходящийся гармонический ряд. Для вычисления суммы первых *k* членов этого ряда воспользуемся асимптотической формулой Эйлера

$$\sum_{i=1}^{k}\frac{1}{i}=Ln\left(k\right)+C+\varepsilon. \qquad (9)$$

Здесь *C* – постоянная Эйлера, равная 0,577215…, ε – погрешность определения суммы. Значение погрешности стремится к нулю при больших значениях *k*. Например, уже при *k* = 100 относительная погрешность составляет меньше 0,1 %.

Запишем гармонический ряд: $1,\ \frac{1}{2},\ \frac{1}{3},\ldots,\frac{1}{N-n},\underbrace{\frac{1}{N-n+1},\frac{1}{N-n+2},\ldots,\frac{1}{N-1},\frac{1}{N}}.$

Для вычисления суммы элементов этого ряда от элемента $\frac{1}{N-n+1}$ до $\frac{1}{N}$, т.е. коэффициента K_n, вычислим сумму всех членов гармонического ряда от 1 до $\frac{1}{N}$, из которой вычтем сумму членов гармонического ряда от 1 до $\frac{1}{N-n}$. Порядковый номер n должен быть меньше *N*, в противном случае последний элемент гармонического ряда от 1 до $\frac{1}{N-n}$. будет равен бесконечности. Воспользовавшись формулой Эйлера, получим $\sum_{i=1}^{n}\frac{1}{N-i+1}=$ $=Ln\left(N\right)-Ln\left(N-n\right)+\varepsilon_1$ при $n<N$, ε_1 – погрешность вычисления разности последовательностей, равная разности погрешностей ε. При $n=N$ сумма элементов знаменателя коэффициента K_n определяется по формуле (9).

Пренебрегая значением погрешности ε_1, номер порядковой статистики, для которой коэффициент K_n равен единице, можем легко найти из выражения $Ln(N)-Ln(N-n)=1$. Очевидно $n=N(1-e^{-1})$. Так как порядковый номер элемента ряда может быть только целым числом, то $n\ [N(1-e^{-1})]$, где $[\ldots]$ – целая часть числа, ближайшая к целому. Следовательно, для оценки среднего значения экспоненциального распределения достаточно взять порядковую статистику с номером $n[N(1-e^{-1})]\quad[0,632N]$, т.е. $\widehat{m}_{p1}=t_{([0,632N])}$.

Таким образом, для оценки среднего значения случайной величины, имеющей экспоненциальное распределение, из вариационного ряда можно выбрать порядковую статистику $t_{(n)}$, для неё вычислить коэффициент K_n и на величину этого коэффициента скорректировать значение выбранной порядковой статистики. Более простой вариант оценки среднего значения случайной величины предполагает выбор статистики с номером $[0,632\cdot N]$, значение которой и является искомой оценкой. Относительная эффективность этих оценок среднего значения экспоненциального распределения по сравнению с оптимальной линейной оценкой превосходит 0,90 [9]. Оптимальная оценка среднего значения величины, имеющей экспоненциальное распределение, по одной порядковой статистике обеспечивается при использовании из вариационного ряда измерения с номером $0,7968\cdot N+1$, … – наибольшее целое [6]. Оценка математического ожидания в этом случае $\widehat{m}_{p2}=t_{(0,7968\cdot N+1)}{*}0,6275$. Устройства, реализующие оценку среднего значения времени запаздывания на основе одной порядковой статистики, рассмотрены в патенте РФ [10].

Статистические оценки, вычисленные на основе порядковых статистик, можно характеризовать относительной эффективностью – отношением количества информации, извлекаемой из выборки *KI*, к количеству информации в полной выборке объемом N [6]. Количество информации, извлекаемой из выборки *KI* при вычислении $\widehat{m}_{p1}$, составляет 58,20 %, а при вычислении $\widehat{m}_{p2}$ – 64,76 %. Традиционным способом увеличения информации, извлекаемой из выборки, является увеличение числа элементов, на основе которых вычисляется эта оценка. В работе [6] показано, что количество информации, извлекаемой из выборки, составляет 99,18 % уже при использовании пятнадцати её элементов с оптимальными номерами и в дальнейшем практически не увеличивается. При инженерных расчетах для оценки среднего значения достаточно использовать пять порядковых статистик с оптимальными номерами, которые позволяют из всей выборки извлечь 94,76 % информации (шестая статистика добавляет 1,30 %, седьмая – еще 0,87 %).

Для вычисления оценки среднего значения $\widehat{m}_{pn}$ по n порядковым статистикам с оптимальными номерами следует воспользоваться формулой [6]:

$$\widehat{m}_{pn} = \sum_{i=1}^{n}(s_i \cdot N+1)\cdot b_i. \tag{10}$$

Значения коэффициентов s_i, позволяющих определить оптимальные номера порядковых статистик и их вес b_i, приведены в таблице 1 [6].

Таблица 1 – Коэффициенты для вычисления оценки среднего значения экспоненциального распределения на основе порядковых статистик с максимальной относительной эффективностью

Table 1 – Factors for calculating an estimate of average value for exponential distribution based on order statistics with maximum relative efficiency

Коэффициенты	Число порядковых статистик, n				
	1	2	3	4	5
s_1	0,7968	0,6386	0,5296	0,4514	0,3931
b_1	0,6275	0,5232	0,4477	0,3907	0,3463
s_2		0,9266	0,8300	0,7419	0,6670
b_2		0,1790	0,2266	0,2361	0,2320
s_3			0,9655	0,9067	0,8434
b_3			0,0775	0,1195	0,1402
s_4				0,9810	0,9434
b_4				0,0409	0,0709
s_5					0,9885
b_5					0,0243
KI, %	64,76	82,03	89,10	92,69	94,76

Анализ коэффициентов таблицы подтверждает, что наибольший вклад в значение оценки среднего значения вносят порядковые статистики со «средними» номерами. Вычисление оценок среднего значения на основе порядковых статистик наиболее целесообразно при проверке сложных гипотез о возможном законе распределения с помощью непараметрических критериев типа Колмогорова, которые предполагают ранжирование результатов измерений. Отметим, оценка математического ожидания на основе порядковых статистик может быть вычислена и для цензурированных выборок. Если цензурирование проведено для малых значений (слева), то у оставшихся элементов выборки индексы следует увеличить на число цензурированных элементов. При цензурировании только справа такая оценка проводится обычным способом. Очевидно, для такой оценки среди оставшихся элементов должны быть соответствующие порядковые статистики.

Значение плотности экспоненциального распределения при $t = 0$ $f(0) = \frac{1}{m}$, т.е. для оценки среднего значения времени запаздывания достаточно определить значение плотности распределения при $t = 0$. При экспериментальной оценке плотности распределения – построении гистограммы – начальное значение соответствует относительной частоте попадания случайной величины в первый дифференциальный коридор. Так как это значение зависит от ширины (числа) дифференциальных коридоров, то найденную относительную частоту следует скорректировать на ширину дифференциального коридора. Ориентировочную оценку среднего времени запаздывания $\widehat{m}_g$ зажигания можно вычислить на основе первого значения гистограммы, и она равна:

$$\widehat{m}_g = \frac{N}{d_1}\Delta. \tag{11}$$

Здесь d_1 – число случайных значений, попавших в первый дифференциальный коридор; Δ – ширина дифференциального коридора, равная (при одинаковых интервалах) разности между максимальным и минимальным значениями измеряемого времени запаздывания, деленной на число дифференциальных коридоров k.

При проведении исследований число дифференциальных коридоров и их ширину определяют на основе предполагаемого диапазона изменения и требуемой точности построения гистограммы (объема выборки). Эти параметры выбирают перед началом исследований, следовательно, в формуле (11) неизвестно лишь d_1. Такой способ оценки среднего значения реализован в устройстве [11]. Формирование гистограммы в этом устройстве осуществляется в процессе измерений (поступления данных от измерительного устройства) за счет использования значения случайных величин в качестве адресов ячеек блока памяти, содержимое которых увеличивается на единицу при поступлении очередной случайной величины. Так как в процессе построения гистограммы осуществляется группирование результатов измерений, т.е. теряется часть информации, то оценка среднего значения экспоненциального распределения, вычисленная по формуле (10), является менее точной по сравнению с оценкой выборочного среднего. Таким способом нетрудно вычислить экспресс-оценку интенсивности зажигания в процессе исследований для «круглых» значений длительности интервала и числа проведенных измерений. Для увеличения точности оценки математического ожидания времени запаздывания можно использовать все значения гистограмм. В этом случае вычисления будут более громоздкими, а оценка математического ожидания, как правило, хуже (из-за округления при группировке результатов измерений), чем оценка выборочного среднего.

Установление соответствия характеристик серийно изготавливаемых газоразрядных индикаторов заданным требованиям целесообразно проводить с помощью статистического контроля характеристик индикатора – контрольных испытаний. Такие испытания, проводимые последовательно, позволяют принять решение о соответствии (или несоответствии) характеристик объекта заданным требованиям по небольшому числу измерений. Методика проведения таких испытаний и её эффективность рассмотрены в работе [1]. Устройство последовательного статистического контроля времени запаздывания для принятия с заданным риском решения о соответствии исследуемых индикаторов категории годных либо категории негодных защищено патентом РФ [12].

Достоверность предложенных методов оценки параметров распределения обеспечивается лишь в случае, если подтверждена экспоненциальность распределения. Так как время запаздывания возникновения разряда в некоторых случаях может характеризоваться суперпозицией экспоненциальных распределений или гамма распределением, то необходимо убедиться в экспоненциальности совокупности результатов измерений. Кроме анализа физических процессов для этого следует использовать многочисленные критерии согласия, разработанные в том числе и для цензурированных выборок. При объемах выборки, состоящей из более 50 элементов, наибольшей мощностью для экспоненциального распределения обладает критерий Фроцини [8], который используется в устройстве [13].

Результаты имитационного моделирования

Достоверность вычисления оценок параметров экспоненциального распределения рассмотренными способами проверялась путем моделирования в среде MATLAB. Для этого при некоторых исходных параметрах распределения: среднем значении (интенсивности зажигания) m (λ) и смещении τ генерировался массив, содержащий N случайных значений времени запаздывания, элементы которого соответствуют экспоненциальному распределению.

Оценки среднего значения и смещения методом моментов и максимального правдоподобия вычислялись по формулам (1)-(4). Для вычисления оценок по формулам (5) и (6) сформированный массив сортировался по возрастанию. Устанавливалось некоторое пороговое значение времени τ_m как часть от исходного среднего значения времени запаздывания m. Из ранжированного массива исключались элементы (случайные значения) $t_{(1)}$, $t_{(2)},...,t_{(r-1)}$, $t_{(r)}$, которые меньше установленного порогового значения τ_m, фиксировалось число удаленных элементов r, т.е. проводилось цензурирование массива слева. Среди оставшихся после удаления «недостоверных» элементов массива выделялся наименьший элемент $t_{(r+1)}$. Затем на основе цензурированной выборки по формулам (5) и (6) вычислялись оценки r_b среднего значения $\hat{m}_{zl}$ и смещения $\hat{\tau}_{zl}$ с использованием наименьший статистики $t_{(r+1)}$ и при ее замене на время τ_m и $t_{(r+1)}$.

Для моделирования ситуации, соответствующей не зажиганию элемента отображения в некоторых циклах формирования стимулирующих сигналов, т.е. когда неизвестны максимальные значения времени запаздывания зажигания, устанавливалось некоторое пороговое значение времени T также как часть от исходного среднего значения времени запаздывания m. Далее из ранжированного по возрастанию массива исключались элементы (случайные значения) $t_{(b)}$, $t_{(b+1)},...,t_{(N)}$, которые больше, чем T, фиксировалось число удаленных элементов r_b. Оценка математического ожидания вычислялась по формуле (7), а также на основе элементов массива, цензурированного справа по формуле (1) для объема выборки N-r_b.

Оценки математического ожидания на основе порядковых статистик вычислялись по формуле (8) для $K_n = 1$, а с максимальной относительной эффективностью – по формуле (10). Номера этих порядковых статистик и их вклад приведены в таблице 1. Для ориентировочной оценки среднего времени запаздывания на основе первого значения гистограммы определялся диапазон изменения сформированных случайных значений, устанавливалось число дифференциальных коридоров, их ширина, определялось число случайных значений, попадающих в первый коридор. Найденные значения подставлялись в формулу (11).

Все оценки вычислялись для одного и того же массива случайных значений. Элементы этого массива и, следовательно, вычисленные оценки являются случайными, поэтому для каждых исходных значений моделирование повторялось многократно (M раз), а затем вычислялись минимальное, среднее и максимальное значения оценок. Моделирование проводилось многократно при различных условиях. В таблицах 2 и 3 приведены результаты моделирования для выборки, содержащей 300 экспоненциально распределенных элементов, которая позволяет получить оценки с «инженерной» точностью и достоверностью [1]. Число повторений $M = 100$. Исходное относительное среднее время запаздывания зажигания равно единице, смещение – ноль. Пороговое значение времени для цензурирования слева 0,25, а для цензурирования справа 3,00 от относительного среднего значения времени запаздывания. Оценка среднего значения на основе гистограммы проводилась по первому из 25 равных дифференциальных коридоров.

Анализ результатов моделирования показывает, что и часто используемые для вычисления оценок параметров экспоненциального распределения методы моментов и максимального правдоподобия и рассмотренные методы на основе цензурированных выборок дают практически одинаковые значения (таблица 2). Эти оценки минимального, среднего и максимального значения математического ожидания незначительно отличаются от соответствующих оценок выборочного среднего, которые равны соответственно 0,8685; 1,0058 и 1,1170 и практически укладываются в доверительные интервалы для «инженерной» достоверности 0,95. Оценка показателей распределения при замене в формулах (5) и (6) $t_{(r+1)}$ на τ_m или τ_{r+1} приводила к изменению оценки математического ожидания не более чем на единицу в третьем знаке после запятой, т.е. не более чем на 1 %. Оценка смещения при этом не изменялась.

Таблица 2 – Оценки параметров экспоненциального распределения на основе различных методов

Table 2 – Parameter estimates of exponential distribution based on various methods

Оцениваемый параметр и его значение		Метод максимального правдоподобия	Метод моментов	Оценка на основе цензурированной выборки		Оценка на основе гистограммы
				слева (-22 %)	справа (-5 %)	
Смещение	Мин. знач.	-0,0033	-0,1786	-0,0846	-	-
	Сред. знач.	0,0004	0,0003	$6{,}061\cdot10^{-4}$	-	-
	Макс. знач.	0,0142	0,1217	0,0846	-	-
Математическое ожидание	Мин. знач.	0,8718	1,0471	0,8788	0,8354 (0,7320)	0,5942 (1,0204)
	Сред. знач.	1,0054	1,0055	1,0044	1,0050 (0,8477)	0,9594 (1,1072)
	Макс. знач.	1,1028	0,9953	1,1389	1,1440 (0,9678)	1,5586 (1,2048)

Несколько больше отличаются от идеальных значений оценки, вычисленные на основе выборки, цензурированной справа, тем не менее, они существенно точнее оценок, найденных на основе лишь тех случайных значений, которые меньше T, т.е. зарегистрированных значений времени запаздывания зажигания элемента отображения (данные приведены в скобках).

Оценка на основе первого столбца гистограммы имеет большую погрешность, однако следует помнить, что для построения гистограммы, близкой к плотности распределения, требуется на порядок большее число наблюдений [1]. При увеличении объема выборки до 3000 элементов и диапазоне изменения случайных значений 5m, в который попадает 99,33 % элементов выборки, оценки математического ожидания становятся приемлемыми (указаны в скобках в последнем столбце таблицы 2).

Оценки математического ожидания, вычисленные путем статистического моделирования на основе порядковых статистик (таблица 3), имеют несколько больший разброс по сравнению с соответствующими оценками выборочного среднего.

Минимальные и максимальные оценки, вычисленные на основе порядковых статистик, незначительно выходят из границ доверительных интервалов для «инженерной» достоверности. Вычисленные оценки порядковых статистик практически не зависят от номеров и количества порядковых статистик, на основе которых проведены вычисления. Незначительное уменьшение разброса оценок среднего значения при увеличении числа порядковых статистик выявлено для выборки из 50 элементов, однако эти оценки были существенно менее точными. Так как оценки математического ожидания на основе порядковых статистик вычисляются более просто и достоверно отражают исследуемые случайные процессы, то их целесообразно использовать наряду с оценками, вычисленными другими способами.

Таблица 3 – Оценки математического ожидания экспоненциального распределения на основе порядковых статистик

Table 3 – Estimates of mathematical expectation of exponential distribution based on order statistics

Значение параметра	Оценка на основе одной статистики ($K_n = 1$)	Число статистик, используемых для вычисления оценки с максимальной относительной эффективностью				
		1	2	3	4	5
Мин. знач.	0,8053	0,8324	0,8687	0,8774	0,7942	0,8000
Сред. знач.	1,0069	1,0233	1,0160	1,0200	0,9542	1,0281
Макс. знач.	1,1928	1,2606	1,1485	1,1683	1,0906	1,1615

В соответствии с «инженерной» достоверностью исследований каждый из сформированных массивов содержал достаточно небольшое число элементов (300). Естественно, оценки, вычисленные для каждого из массивов, отличались. В процессе исследований установлено, что все оценки, вычисляемые для различных массивов, изменялись «синхронно», принимая бо̀льшие, средние или меньшие значения. Многократное повторение процесса моделирования сглаживает эти случайности, обеспечивая точное вычисление средних значений оценок.

Выводы

Для вычисления оценок математического ожидания и смещения времени запаздывания зажигания газоразрядных матричных индикаторов целесообразно использовать не только широко используемые методы моментов или максимального правдоподобия. В условиях сильной взаимной ионизации элементов отображения цензурирование выборки слева может быть единственным способом получения достоверных результатов при одновременном снижении требований к узлу определения состояния элемента отображения. При больших задержках времени запаздывания достоверность результатов исследований существенно увеличивается (в среднем на 15 %), а время исследований может быть уменьшено, если оценка показателей распределения вычисляется на основе выборок, цензурированных справа. Достоверное и более простое вычисление оценки математического ожидания обеспечивается при использовании лишь одной порядковой статистики.

Достоинства рассмотренных методов и устройств реализуются для экспоненциального распределения. Возможность его использования для описания времени запаздывания зажигания должна быть подтверждена, прежде всего, на основе исследования соответствующих физических процессов либо на основе наиболее мощных критериев согласия, которые учитывают особенности экспоненциального распределения.

Библиографический список

1. **Шестеркин А. Н.** Определение объема выборки для исследования времени запаздывания зажигания газового разряда // Вестник Рязанского государственного радиотехнического университета. 2020. № 71. С. 220-233.

2. **Вадзинский Р. Н.** Справочник по вероятностным распределениям. СПб.: Наука, 2001. 295 с.

3. **Горянов Б. В., Павлов И. В., Цветкова Г. М. и др.** Математическая статистика: учеб. для вузов. Вып. XVII. М.: Издательство МГТУ им. Н.Э. Баумана, 2001. 424 с.

4. ГОСТ Р 50779.26-2007. Статистические методы. Точечные оценки, доверительные, предикционные и толерантные интервалы для экспоненциального распределения. М.: Стандартинформ, 2008. 27 с.

5. **Лагутин М. Б.** Наглядная математическая статистика. М.: Бином. Лаборатория знаний, 2007. 472 с.

6. **Сархан А. Е., Гринберг Б. Г.** Ведение в теорию порядковых статистик. М.: Статистика, 1970. 416 с.

7. **Шестеркин А. Н.** Устройство для оценки параметров экспоненциального распределения. Патент РФ 208739.

8. **Кобзарь А. И.** Прикладная математическая статистика. Для инженеров и научных работников. М.: ФИЗМАТЛИТ, 2006. 816 с.

9. **Шестеркин А. Н.** Способ оценки параметров распределения времени запаздывания зажигания и устройство для его осуществления. Патент РФ 2646897.

10. **Шестеркин А. Н.** Устройство для оценки среднего времени запаздывания возникновения разряда (его варианты). Патент РФ 2678646.

11. **Шестеркин А. Н.** Устройство для оценки параметров распределения времени запаздывания возникновения разряда. Патент РФ 181880.

12. **Шестеркин А. Н.** Устройство для статистического приемочного контроля газоразрядных индикаторов. Патент РФ 2714382.

13. **Шестеркин А. Н.** Устройство для определения статистических характеристик времени запаздывания зажигания элементов матричного индикатора. Патент РФ 2767598.

UDC 621.382

PARAMETER ESTIMATION OF GAS DISCHARGE INDICATORS IGNITION DELAY TIME DISTRIBUTION

A. N. Shesterkin, Dr. Sc. (Tech), full professor, Department of computing and applied mathematics, RSREU, Ryazan, Russia;
orcid.org/0000-0002-0633-0591, e-mail: neon60@inbox.ru

Methods to compute the estimates of mathematical expectation and time displacement of ignition delay time when describing it by exponential distribution are considered. ***The aim of the work*** *is to analyze methods for calculating statistical estimates of ignition delay time. Classical methods of calculating exponential distribution parameters – method of moments and maximum likelihood method- are analyzed. It is proposed to censor measurement results (sampling) of delay time in conditions of strong ionization on the left and to evaluate distribution parameters based on other, uncensored measurements. The article offers to calculate average delay time based on right-censored sample while investigating delay time in the modes when statistical delay time can be time-consuming. Various ways of estimating mathematical expectation of delay time based on order statistics are analyzed. The possibility of approximate estimate of average value based on a histogram is considered. The information about the devices implementing proposed methods of parameter estimation is provided. The accuracy of considered methods of computing exponential distribution parameter estimates is confirmed by statistical modeling.*

Key words*: discharge delay time, exponential distribution, distribution parameter estimation, sample size, order statistics, parameter estimation devices.*

DOI: 10.21667/1995-4565-2022-79-141-153

References

1. **Shesterkin A. N.** Opredelenie ob#ema vyborki dlja issledovanija vremeni zapazdyvanija zazhiganija gazovogo razrjada. *Vestnik Rjazanskogo gosudarstvennogo radiotehnicheskogo universiteta.* 2020, no. 71, pp. 220-233 (in Russian).

2. **Vadzinskij R. N.** *Spravochnik po verojatnostnym raspredelenijam*. SPb.: Nauka. 2001, 295 p. (in Russian).

3. **Gorjanov B. V., Pavlov I. V., Cvetkova G. M. i dr.** *Matematicheskaja statistika: ucheb. dlja vuzov*. Vyp. XVП. M.: Izdatel'stvo MGTU im. N.Je. Baumana. 2001, 424 p. (in Russian).

4. GOST R 50779.26-2007. *Statisticheskie metody. Tochechnye ocenki, doveritel'nye, predikcionnye i tolerantnye intervaly dlja jeksponencial'nogo raspredelenija*. Moscow: Standartinform. 2008, 27 p. (in Russian).

5. **Lagutin M. B.** *Nagljadnaja matematicheskaja statistika.* M.: Binom. Laboratorija znanij. 2007, 472 p. (in Russian).

6. **Sarhan A. E., Grinberg B. G.** *Vedenie v teoriju porjadkovyh statistik*. Moscow: Statistika. 1970, 416 p. (in Russian).

7. **Shesterkin A. N.** *Ustrojstvo dlja ocenki parametrov jeksponencial'nogo raspredelenija*. Patent RF 208739. (in Russian).

8. **Kobzar' A. I.** *Prikladnaja matematicheskaja statistika. Dlja inzhenerov i nauchnyh rabotnikov*. Moscow: FIZMATLIT, 2006. 816 p. (in Russian).

9. **Shesterkin A. N.** *Sposob ocenki parametrov raspredelenija vremeni zapazdyvanija zazhiganija i ustrojstvo dlja ego osushhestvlenija*. Patent RF 2646897. (in Russian).

10. **Shesterkin A. N.** *Ustrojstvo dlja ocenki srednego vremeni zapazdyvanija vozniknovenija razrjada (ego varianty)*. Patent RF 2678646. (in Russian).

11. **Shesterkin A. N.** *Ustrojstvo dlja ocenki parametrov raspredelenija vremeni zapazdyvanija vozniknovenija razrjada*. Patent RF 181880. (in Russian).

12. **Shesterkin A. N.** *Ustrojstvo dlja statistichesogo priemochnogo kontrolja gazorazrjadnyh indikatorov*. Patent RF 2714382. (in Russian).

13. **Shesterkin A. N.** *Ustrojstvo dlja opredelenija statisticheskih harakteristik vremeni zapazdyvanija zazhiganija elementov matrichnogo indikatora*. Patent RF 2767598 (in Russian).