

УДК 004.8:623

А.С. ГОРСКИЙ, кандидат технических наук

## МЕТОДИЧЕСКИЙ ПОДХОД К ОЦЕНИВАНИЮ ХАРАКТЕРИСТИК ИНТЕЛЛЕКТУАЛЬНЫХ РОБОТОТЕХНИЧЕСКИХ СРЕДСТВ

*Статья посвящена исследованию вопросов методологии проведения государственных испытаний и оценки качества функционирования интеллектуальных робототехнических средств. Предлагаемый подход заключается в применении методов теории вероятностей, которые позволяют проводить измерение степени возможности случайных результатов испытаний в различных условиях на основе уже выявленных закономерностей массовых событий, что дает возможность прогнозировать результаты испытаний и обосновать требования к условиям их проведения.*

**Ключевые слова:** искусственный интеллект; испытания; технология; теория вероятностей; нормальное распределение; уровень доверия.

### Введение

Активное внедрение технологий искусственного интеллекта (ИИ) в различные сферы человеческой деятельности вызывает необходимость переосмысления многих аспектов создания новых технических систем и комплексов [1-6]. В этой связи одним из основных проблемных вопросов является оценка качества функционирования интеллектуальных робототехнических средств (ИРТС). Под ИРТС будем понимать робототехнические средства (РТС), в которых реализованы технологии ИИ. Объективные сложности в решении данной проблемы возникают ввиду отсутствия на сегодняшний день научно-методических подходов, определяющих единые требования к методам проведения испытаний ИРТС.

В работе [7] рассматривается подход, который предполагает проведение испытаний ИРТС в два этапа:

1. Испытания программных (программно-аппаратных) средств, реализующих технологии ИИ и входящих в состав образца, с помощью программно-аппаратного комплекса имитационных моделей (виртуальные испытания).
2. Испытания опытного образца ИРТС (натурные испытания).

Необходимо отметить, что любая технология ИИ реализуется в виде программного обеспечения (ПО), а потом уже это ПО устанавливается на соответствующую программно-аппаратную платформу и применяется в составе образца. Это говорит о том, что качество внедрения технологии ИИ в образец РТС будет зависеть, прежде всего, от качества функционирования разработанного ПО. Принимая во внимание данный факт, можно сказать, что вышеуказанный подход вполне обоснован. Поэтому в данной статье предлагается ограничить исследования поиском и разработкой научно-методического подхода к решению задачи проведения виртуальных испытаний ИРТС. Под виртуальными (цифровыми) испытаниями будем понимать определение количественных и (или) качественных характеристик свойств объекта испытаний как результата исследования свойств цифровой модели (или цифрового двойника) этого объекта<sup>1</sup>.

<sup>1</sup> ГОСТ Р 57700.37-2021. Компьютерные модели и моделирование. Цифровые двойники изделий. Общие положения. М., 2021.

## Анализ и выбор методологического аппарата оценки характеристик интеллектуальных РТС

Отличительной особенностью проведения испытаний ИРТС является поиск и выбор методов оценки, как правило, при отсутствии какой-либо имеющейся статистики. В условиях такой статистической неопределенности наиболее подходящей методологией для обработки полученных результатов является теория вероятностей.

Методы теории вероятностей позволяют проводить измерение степени возможности случайных результатов испытаний в различных условиях на основе уже выявленных закономерностей массовых событий, что дает возможность прогнозировать результаты испытаний и обосновать требования к условиям их проведения.

С ростом количества повторений одного и того же испытания усложняется математический анализ его результатов. Если это количество небольшое, например, в пределах двух-пяти, то достаточно использования теорем сложения и умножения вероятностей; для десятка можно применить формулу биномиального распределения, а для сотни, тысячи и более – нормальное распределение [8; 9].

Для обеспечения уровня доверия к итоговому результату испытаний применяются два основных способа. Первый способ заключается в формировании статистической выборки (вариационного ряда) с дальнейшей обработкой и определением ее параметров. Второй – в определении априорной вероятности случайного события и других его параметров до испытания, приводящего к определенному исходу. Такой способ применяется тогда, когда формирование репрезентативной выборки до проведения испытаний невозможно либо является сложным или ресурсоемким процессом. В случае с виртуальными испытаниями получение статистических данных возможно только непосредственно в ходе их проведения. Поэтому для обоснования требований к этому виду испытаний необходимо использовать последний способ.

Рассмотрим более подробно порядок применения этого способа.

Существуют два пути повышения уверенности в результатах испытаний:

1. Расширение доверительного интервала значений случайной величины.
2. Увеличение выборки, то есть количества повторений одного и того же испытания.

Если пойти по первому пути, то это означает, что при фиксированном количестве повторений опыта необходимо будет повышать погрешность успешного исхода события. Такой подход применяется, когда нет возможности увеличить объем выборки по каким-либо причинам (ограниченность ресурсов, времени проведения испытаний, сложность организации и т.д.). В этом случае нужно понимать, насколько мы готовы ухудшить результат, то есть увеличить интервал ошибки каждого испытания, чтобы повысить уровень доверия к конечному результату всех испытаний.

С учетом постоянного роста высокой производительности и ресурсоемкости современных средств вычислительной техники мы можем принять допущение, что количество виртуальных испытаний ИРТС должно быть таким, чтобы обеспечивался требуемый уровень доверия. Это допущение позволяет увеличить число степеней свободы выборки по количеству испытаний до нужного значения. Другими словами, путь увеличения выборки для повышения уверенности в результатах виртуальных испытаний ИРТС является вполне приемлемым.

Виртуальные испытания проводятся по схеме Бернулли и их результат может быть одним из двух – успех или неуспех. То есть мы имеем дело с дискретной случайной величиной, которая всегда может быть описана биномиальным распределением. Теоретически вероятность того, что  $k$  испытаний в одной выборке из  $n$  испытаний завершиться успехом равна:

$$\hat{P}(k) = \frac{n!}{k!(n-k)!} \cdot p^k (1-p)^{n-k},$$

где  $\frac{n!}{k!(n-k)!}$  – биномиальный коэффициент;

$n$  – число испытаний;

$p$  – априорная вероятность успеха при однократном испытании, как правило, принимается  $p = 0,5$ .

В общем случае для любого  $\hat{P}$  биномиальное распределение для  $n$  испытаний с вероятностью успеха  $p$  описывается нормальной кривой с ожидаемым числом успехов, равным математическому ожиданию:

$$M(\hat{P}) = n \cdot p, \quad (1)$$

и среднеквадратическим отклонением:

$$\sigma = \sqrt{n \cdot p(1-p)}. \quad (2)$$

Доказано, что с увеличением  $n$  все биномиальные распределения в итоге становятся нормальными [8; 9]. В соответствии с центральной предельной теоремой существует такое преобразование  $z$ :

$$z = \frac{\hat{P} - M}{\sigma}, \quad (3)$$

превращающее любую нормальную случайную величину в случайную величину со стандартным нормальным распределением, при котором  $M = 0$  и  $\sigma = 1$ . Теперь найти вероятность того, что  $z$  попадет в произвольный интервал  $a \leq z \leq b$ , можно по общей таблице для стандартного нормального распределения  $f(z)$ , используя выражение:

$$P(a \leq z \leq b) = \int_a^b f(z) dz. \quad (4)$$

Чтобы получить наиболее точное непрерывное приближение дискретной биномиальной величины, необходимо учитывать поправку на непрерывность, равную 0,5. С учетом этой поправки, подставив выражения (1) и (2) в (3), а формулу (4) вместо  $a$  и  $b$ , запишем общее правило вычисления вероятностей для нормального распределения:

$$P(a \leq z \leq b) \approx P\left(\frac{a-0,5-n \cdot p}{\sqrt{n \cdot p(1-p)}} \leq z \leq \frac{b+0,5-n \cdot p}{\sqrt{n \cdot p(1-p)}}\right). \quad (5)$$

В теории вероятностей доказано, что среднеквадратическое отклонение выборочных средних значений для нормального распределения равно:

$$\sigma = \sqrt{\frac{p(1-p)}{n}}. \quad (6)$$

Как уже было отмечено ранее, основной задачей перед проведением виртуальных испытаний является определение количества повторений одного и того же испытания, которое бы обеспечило требуемый уровень доверия к его результату.

Чтобы рассчитать это число, прежде всего необходимо определить требуемый уровень доверия. Тогда можно будет найти разницу  $\alpha$  между желаемым уровнем доверия и абсолютной уверенностью, а уровень доверия будет равен  $(1 - \alpha) \cdot 100\%$ . Чтобы рассчитать доверительный интервал вероятности успешных испытаний  $M(\hat{P})$  для требуемого уровня доверия, нужно найти по справочной таблице стандартного распределения<sup>2</sup> значение  $z_{\alpha/2}$ , после чего с учетом (3) выражение (5) для  $M(x)$  можно переписать в следующем виде:

$$P(M(\hat{P}) - z_{\alpha/2} \cdot \sigma \leq M(\hat{P}) \leq M(\hat{P}) + z_{\alpha/2} \cdot \sigma) = P(M(\hat{P}) \pm z_{\alpha/2} \cdot \sigma). \quad (7)$$

Теперь  $z_{\alpha/2} \cdot \sigma$  обозначим через  $E$ , то есть:

$$E = z_{\alpha/2} \cdot \sigma, \quad (8)$$

и преобразуем выражение (7) к следующему виду:

$$P(M(\hat{P}) - E \leq M(\hat{P}) \leq M(\hat{P}) + E) = P(M(\hat{P}) \pm E). \quad (9)$$

Учитывая, что границы интервала  $M(\hat{P})$  задаются выражением  $M(\hat{P}) \pm E$ , где  $E$  по сути есть ошибка, с учетом (6)  $E$  имеет вид:

$$E = z_{\alpha/2} \cdot \sqrt{\frac{p(1-p)}{n}}.$$

Принимая априорную вероятность успеха  $p = 0,5$  в единичном испытании, взяв  $z_{\alpha/2}$  из справочной таблицы<sup>3</sup> и зная  $E = 1 - P$ , объем выборки можно определить по формуле:

$$n = \frac{(z_{\alpha/2})^2 \cdot p(1-p)}{(1-P)^2}. \quad (10)$$

После проведения виртуальных испытаний ИРТС, имея количество  $k$  успехов в одной выборке из  $n$  испытаний, можно судить о степени соответствия полученной статистической вероятности  $\hat{P}(k) = \frac{k}{n}$  успешных испытаний требованию тактико-технического задания (ТТЗ), например, вероятности обнаружения и распознавания объектов системой технического зрения ИРТС. Теперь рассмотрим порядок применения выбранного математического аппарата на конкретных примерах.

## Обработка и анализ результатов проведения виртуальных испытаний ИРТС

Большинство задач теории вероятностей решаются либо дедуктивным методом, то есть от гипотезы к выводу, либо индуктивным – от выборочных наблюдений к разумной гипотезе. Также возможен комбинированный метод, представляющий собой сочетание первых двух. В случае с виртуальными испытаниями целесообразным представляется применять первый метод. Далее рассмотрим возможный порядок применения метода дедукции.

<sup>2</sup> Вентцель Е.С. Теория вероятностей: учебник. 4-е изд. М.: Наука, 1969. – 576 с.

<sup>3</sup> Там же.

На первом этапе при разработке программ и методик государственных испытаний выдвигается гипотеза о нормальном законе распределения успешных исходов каждого виртуального испытания. Затем определяем требуемый уровень доверия к результату каждого испытания. К примеру, заказчик предъявляет уровень доверия 95% или  $P = 0,95$ , при котором  $\alpha = 1 - 0,95 = 0,05$ . Чтобы рассчитать доверительный интервал для данного уровня доверия, необходимо, используя плотность функции нормального распределения, найти точки  $z$  и  $-z$  такие, что  $P(M(\hat{P}) - z \leq M(\hat{P}) \leq M(\hat{P}) + z) = 1 - \alpha = 0,95$ . Тогда, учитывая симметрию графика плотности нормального распределения, можно утверждать, что  $z = |-z| = z_{\alpha/2}$  и  $P(z \geq z_{\alpha/2}) = 1 - P(z \leq z_{\alpha/2}) = \frac{\alpha}{2} = 0,025$ . Затем по таблице нормального распределения (приложение Г) [10] при  $P(z \leq z_{\alpha/2}) = 1 - 0,025 = 0,975$  находим  $z_{\alpha/2} = 1,96$ .

Теперь, используя выражение (10), можно определить объем необходимой выборки для уровня доверия 95%:

$$n = \frac{(z_{\alpha/2})^2 \cdot p(1-p)}{(1-P)^2} = \frac{1,96^2 \cdot 0,5^2}{0,05^2} = 384.$$

Таким образом, чтобы достичь уровня доверия 95% с погрешностью 5%, количество испытаний должно быть 384.

После завершения виртуальных испытаний ИРТС можно судить о степени соответствия результатов испытаний установленному заказчиком требованию. Допустим, в ТТЗ на разработку ИРТС задано следующее требование:

использование технологий ИИ должно обеспечить решение задачи по обнаружению и распознаванию объектов на маршруте движения (автомобиль, мост, дорога, овраг и т.д.) в светлое время суток с вероятностью не менее 0,9.

Вместе с тем по результатам проведенных виртуальных испытаний искусственной нейронной сети (ИНС), которая должна обеспечить выполнение указанного требования, была установлена доля успешных испытаний  $\frac{k}{n} = \frac{334}{384} = 0,87$ . Это говорит о том, что данное испытание не пройдено, потому что не выполнено требование ТТЗ, так как  $0,87 < 0,9$ . В этом случае может быть принято решение увеличить количество испытаний для того, чтобы повысить уровень доверия к результату, например, до 96%. Для этого нужно повторить вышеуказанный порядок действий и рассчитать  $n$  для уровня 96%. В соответствии с проведенными расчетами, указанными в таблице 1, это значение будет равно 659. Затем на основе повторно полученного результата испытаний принимать окончательное решение: либо продолжать повышать уровень доверия до достижения приемлемого результата, либо принять решение о доработке ИНС (изменение архитектуры, обучающей выборки и других параметров ИНС) [11].

Таблица 1 – Расчетные значения выборки для различных уровней доверия

№ п/п	Уровень доверия, $P$	Объем выборки, $n$
1	0,999	2387334
2	0,99	16641
3	0,98	3382
4	0,97	1308
5	0,96	659
6	0,95	384
7	0,90	67
8	0,85	23
9	0,80	10

Однако, исходя из интересов потребителя ИРТС, уровень доверия рекомендуется выбирать наибольшим с учетом вычислительной мощности испытательного стенда.

На основе полученных расчетных значений построен график зависимости выборки  $n$  от уровня доверия  $P$  (рисунок 1).

Форма представленной зависимости в большей степени напоминает экспоненциальное распределение. Из таблицы 1 видно, что при последовательном изменении уровня доверия на 0,01 на интервале  $[0,95; 0,98]$  размер выборки изменяется примерно в 2 раза, при значении 0,99 возрастает примерно в 5 раз, а при 0,999 – в 144 раза!

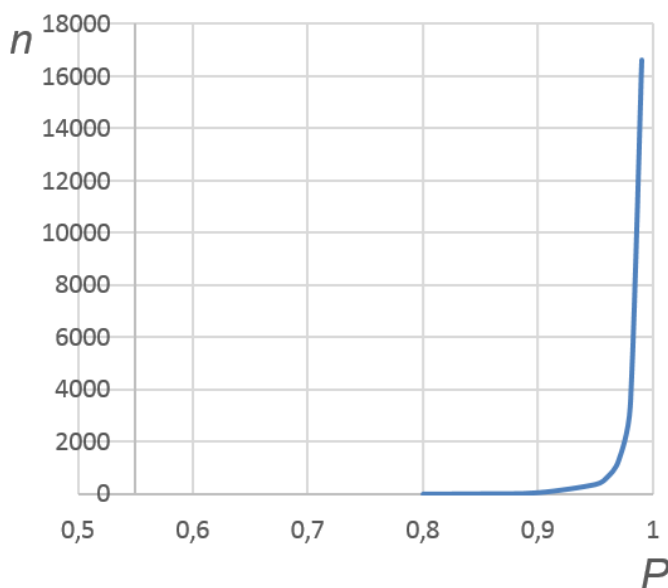


Рисунок 1 – График зависимости  $n$  от  $P$

Анализ полученных расчетных данных показал, что с ростом уровня доверия количество испытаний в выборке увеличивается по экспоненциальному закону. Также необходимо отметить, что при значительных ограничениях в ресурсах, например, при проведении натурных испытаний, приходится решать обратную задачу. То есть исходить из того количества испытаний, которое возможно обеспечить имеющимися средствами. В таком случае нужно четко понимать, что и уровень доверия к результатам испытаний будет существенно снижаться. Допустим, при количестве испытаний, равном 10, доверие к полученному результату будет всего 80%. С другой стороны, этот уровень доверия может быть приемлем, если в ходе 10 испытаний будет получен результат, который удовлетворяет соответствующему требованию заказчика.

Таким образом, можно сделать вывод о том, что количество виртуальных испытаний ИРТС должно быть таким, чтобы обеспечивался требуемый уровень доверия, а количество натурных испытаний определяется, исходя из имеющихся материальных и временных ресурсов и минимального допустимого уровня доверия.

## Заключение

В ходе исследований проведен анализ и обоснован выбор методологического аппарата для оценки характеристик интеллектуальных РТС в ходе государственных испытаний, который представляет собой методы теории вероятностей. Данные методы позволяют прогнозировать результаты испытаний и обосновать требования к условиям их проведения, а также определить степень соответствия результатов испытаний установленному заказчиком требованию. Порядок применения указанных методов рассмотрен на конкретных примерах. Предложенный методический аппарат может быть использован как при разработке методик испытаний ИРТС, так и других технических средств с различными уровнями автономности систем управления.

Основная суть применения предложенного подхода в том, чтобы понять, как повлияет снижение или увеличение количества испытаний на уровень доверия, и насколько можно поступиться доверием, чтобы получить приемлемый результат. Это своего рода методология выстраивания системы сдержек и противовесов перед принятием ответственного решения по итогам испытаний.

## Список использованных источников

1. Корнилов В.И., Маев С.А., Машков К.Ю., Пантелеев А.Л., Соколенко В.Н. Зарубежные робототехнические комплексы военного назначения и требования, предъявляемые к ним // Труды НАМИ. 2015. №263. – С. 65-85.
2. Веревкин А.С., Горский А.С. Актуальные вопросы разработки отечественного связующего программного обеспечения для управления робототехнических комплексов военного назначения // Стратегическая стабильность. 2022. №1(98). – С. 64-66.

3. Кудряшов Д.В., Липатников А.Д., Щербинин Е.В. Подходы к испытаниям искусственного интеллекта в части автономности мобильных роботов // Состояние и перспективы развития современной науки по направлению «Робототехника»: сб. статей IV Всероссийской науч.-технич. конф. (г. Анапа, 20-21 июля 2022 г.). Анапа: ВИТ «ЭРА», 2022. – С. 84-89.
4. Щербинин Е.В., Липатников А.Д., Кудряшов Д.В. К вопросу об испытании адаптивности в системе управления сервисных мобильных роботов как показателя технологии искусственного интеллекта // Состояние и перспективы развития современной науки по направлению «Робототехника»: сб. статей IV Всероссийской науч.-технич. конф. (г. Анапа, 20-21 июля 2022 г.). Анапа: ВИТ «ЭРА», 2022. – С. 204-211.
5. Малов А.В., Лапа М.В., Гончаренко А.Ю. Методы испытаний систем технического зрения для морских роботов // Качество продукции: контроль, управление, повышение, планирование: сб. науч. трудов 8-й Международной молодежной науч.-практич. конф. (г. Курск, 21 ноября 2021 г.). Курск: ЮЗГУ, 2021. – С. 205-207.
6. Громов С.П., Корнилов В.И., Соколенко В.Н., Пантелеев А.Л., Наумов В.Н. Испытания наземных робототехнических комплексов военного назначения за рубежом // Труды НАМИ. 2016. №264. – С. 79-91.
7. Горский А.С. Проблемные вопросы испытаний робототехнических комплексов военного назначения с элементами искусственного интеллекта и пути их решения // Состояние и перспективы развития современной науки по направлению «Робототехника»: сб. статей III Всероссийской науч.-технич. конф. (г. Анапа, 27-28 июля 2021 г.). Анапа: ВИТ «ЭРА», 2021. – С. 58-63.
8. Венецкий И.Г., Венецкая В.И. Основные математико-статистические понятия и формулы в экономическом анализе. М.: Статистика, 1974. – 279 с.
9. Рассел С., Норвиг П. Искусственный интеллект. Современный подход. 4-е изд. Т.2. Знания и рассуждения в условиях неопределенности. СПб.: Диалектика, 2021. – 480 с.
10. Джонсон Н., Лион Ф. Статистика и планирование эксперимента в технике и науке. М.: Мир, 1980. – 611 с.
11. Частиков А.П., Глушко С.П., Тотухов К.Е. Компьютерная симуляция программного управления виртуальным роботом. Исследование, разработка, испытание. L.: LAP Lambert Academic Publishing, 2012. – 144 с.