Logistic Regression Analysis of Factors Affecting Robot Arm Movement Testing

Nichanach Katemukda*

Industrial Engineering, Faculty of Engineering, Rajamangala University of Technology Rattanakosin, Salaya, Phuttamonthon, Nakhon Pathom, 73170, Thailand

***Corresponding Author E–mail:** Nichanach.kat@rmutr.ac.th

Received: Dec 08, 2023; **Revised:** May 24, 2024; **Accepted:** May 27, 2024

Abstract

Currently, with the rapid growth of robots in the industrial and service sectors, the robotic arm product is in high demand, and manufacturers need to deliver it on time. The manufacturer has a new product called "robot arm". The issue is a high failure rate at the test station called the robot arm movement test. The manufacturer focused on the test process in order to reduce any variance that may result in a failure rate, with the operator's performance being their primary interest. Since the first group was built and tested, totaling 233 units of the robot arm, the logistic regression was applied with three independent variables. There are operators, working shifts, and product models. The results indicate day or night shifts are not related to the test failing. The operator and product model are important factors in the test failing. The 1.5–meter– long model has a higher chance of passing the test than the 1–meter–long model by about 13.66 times and the 2–meter–long model by 25.25 times. Operator D is the best at performing the robotic arm test and has a better chance of passing than the other operators (2.07 times for Operator A, 6.53 times for Operator B, and 7.01 times for Operator C). The action is that the software test needs to be updated for the 1– and 2–meter–long models. Moreover, Operators B and C need to be retrained as a priority. Then the manufacturer needs to focus more on the assembly process for yield improvement.

Keywords: Improvement, Logistic regression, Operator performance, Robot arm movement test, Root cause analysis

1. Introduction

The use of robotics is increasing rapidly these days. Since 2015, the growth of robotics installation has been approximately four times that of 2021 [1] as shown in **Figure 1**.

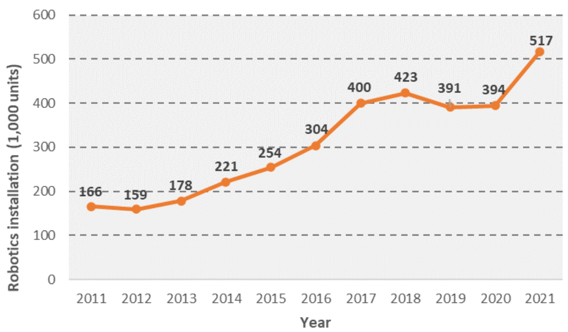

Figure 1 Installation of industrial robots's trend

In the year 2021, China had the most robotic unit installations, Japan was the second, and Thailand was ranked 12th in the world for robotic unit installations [1] as shown in **Figure 2**.

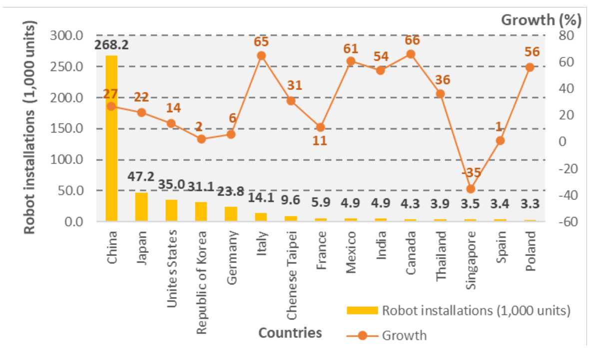

Figure 2 The top 15 largest countries for robot installation in 2021

Robots now play numerous roles in both the industrial and service sectors, particularly the robotic arm. We also see the robot arm pick and place parts, tighten screws, weld, and so on in manufacturing [2]. In the service sector, robots make coffee, food, and so on [3], which is exciting to see. This is the main reason that the manufacturer of the robotic arm needs to introduce a new product and rapidly ramp up volume. The issue based on the yield of the robotic arm movement test has a low yield in the new product introduction phase. The manufacturer introduced the new product of the robotic arm, and they found the output to be low due to the high failure rate of the robotic arm movement test process. Thus, the study aimed to apply a logistic regression statistical analysis to determine the factors that are influenced by the high failure rate of testing. The focus of the manufacturer is mainly on the operator [4]. There are different shifts and different operators. In fact, the researchers required more factors to include in the model for more clarification on the root cause of the test's failure: more interaction between humans and machines. There is the product model, which is the size, which varies for holding, rotating, and testing. The study's benefit was discovering that a logistic regression is appropriate for root cause identification and that interpreting the results leads to action in manufacturing problems, which is practical and can be applied to general problems with similar fundamentals.

2. Objectives

To explore how logistic regression can be applied and identify the root cause (man and machine) of the high failure rate for the robotic arm movement test process. The dependent variables are categorical. There are the working time of the operator (day or night shift), operator (4 operators, represented by operators A, B, C, and D), and the product model, which is represented for the test station (1, 1.5, and 2–meter–long). On the other hand, the independent variable is the test result, which has a qualitative outcome of pass or fail. The research hypothesis as shown in **Figure 3**.

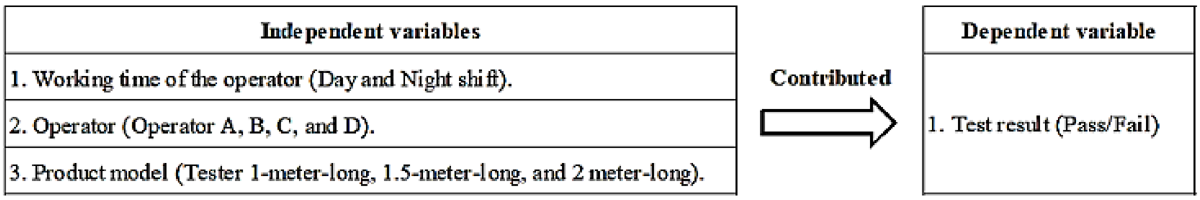

Figure 3 Research hypothesis

3. Materials and methods

3.1 Robotic arms product

A robotic arm is a particular kind of mechanical arm that can be programmed and performs tasks comparable to those of a human arm [5]. It can be a standalone robot or a component of a more sophisticated robot [6]. Such a manipulator has joints that allow for either translational (linear) displacement or rotational motion, similar to an articulated robot. A kinematic chain can be thought of as being formed by the manipulator's linkages. The end effector, which is similar to the human hand, and the function for which it is required, such as a pick– and–place gripper, a welder, and so on [7]. A type of robotic mechanism is collectively referred to as “robotic arms”.

There are many applications for these various kinds of robots. However, each type has unique features that often allow it to be more effective than other robotic arms for particular jobs. 6 different types of robotic in total, there are [8];

Articulated arms; general–purpose robotic arms with five or more joints, or degrees of freedom, are known as articulated arms. Numerous other types of robots are classified as articulated arms. A six–axis robot, for instance, has an articulated arm with six degrees of freedom. The widest variety of industrial robot types are those with articulating arms, which also include six–axis and collaborative robots.

Spherical robot; a spherical (polar) robot served as the first industrial robot in contemporary times. This sort of robot has a straightforward design but is less popular than it previously was. Robots that are spherical resemble cylindrical robots except that they replace the vertical linear axis with a second rotating axis. It has a vertical rotational axis. It was created for straightforward jobs that don't call for rapid or intricate action.

SCARA robot; SCARA robots are Selected Compliance Assembly Robot Arms. They therefore lack the same degree of flexibility that articulated arms have. They are constrained in some ways, but they also have some advantages over articulated arm types as a result.

Cartesian robot; Cartesian robots are unyielding machines that move in a three–dimensional coordinate plane. Three linear actuators are commonly used to build these robots. In the x–axis, one actuator moves left and right. The x–axis actuator is connected to another actuator. In the y–axis plane, this actuator moves upward and downward. The y–axis member is connected to a final actuator, which oscillates in the z–axis plane. Robots designed for little applications are called Cartesian robots.

Cylindrical robot: a single arm that can travel up and down a vertical member is the central component of cylindrical robot arms. The arm is horizontally rotated by this vertical part. In order to complete the work, the arm can extend and retract. These little robots are used for quick, easy jobs and are relatively portable.

Parallel robots: robotic automation options with high speed include parallel robots. These robots can travel at amazing speeds because of their distinctive design.

Almost all the robotic arms of the manufacturer are applied to articulated arms and SCARA robots.

According to the confidential nature of the product, the photo of the actual product does not appear. In any case, the configuration is the same as in **Figure 4**. The manufacturer offers three sizes: 1–, 1.5–, and 2–meter–long products. The only difference between the models is the length (1 meter, 1.5 meters, and 2 meters). The freedom of movement of the robotic arm is used in the functional test [9]. The operator must connect the end of the arm to the points specified by the software test, then rotate the arm anticlockwise or encounter clockwise as specified by the software test. The remainder of the software test will be calculated, with a pass or fail result displayed.

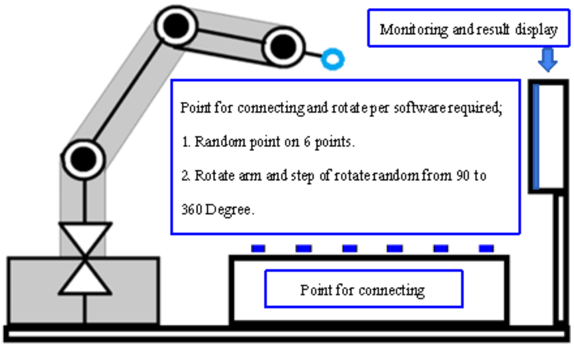

Figure 4 Robotic arm movement test

The interaction between the operator, the robotic arm, and the test software command takes about 15 minutes per unit. This is the main reason that the manufacturer considers the operator to be the main influencer of the failure units of testing. In contrast, the researcher suggests the manufacturer consider how the model may affect the failure units testing as well; then, the product model becomes an independent variable of the research hypothesis.

3.2 Logistic regression

A statistical method for calculating the likelihood of an attribute result given a dependent variable is logistic regression. Binary logistic regression is one kind of logistic regression in which the event has two outcomes, such as "good" or "bad", "pass" or "fail", and so on [10]. Logistic regression is frequently and generally used in many sectors, for example, in the education sector [11], the industrial sector, and so on [12]. Logistic regression's primary theoretical assumption is that it is the same as linear regression. The residual is the difference between the response variables expected and actual values (the error). The residual is typically assumed to have a normal distribution with a mean of zero and constant variance. There is no correlation between the residuals and the variables (both dependent and independent variables) [13]. The dependent variable result will fall between zero and one since the possibility that the event will occur determines the outcome. A probability of zero indicates that the interesting event is not happening, whereas a probability of one indicates that it has a 100% possibility of occurring. Assign the probability of the interesting event, denoted by *P(Y)*, using eq. 1.

$$P(Y) = \frac{e^{b_0+b_1X_i}}{1+\ e^{b_0+b_1X_i}} \quad (1)$$

Eq. 2 can be used to represent the uninteresting event, *Q(Y)*:

$$Q(Y) = 1 - P(Y) \quad (2)$$

where,

b_0 and b_i are the estimated coefficients.

X_i is the independent variable.

e is a Natural Logarithm.

To link the dependent variable to the independent variable, logistic regression can be transformed into a logit equation. Logit link equation form can be written as eq. 3.

$$Logit(P) = Log\left[\frac{P}{Q}\right] \quad (3)$$

The odds ratio is an important number in logistic regression. The odds ratio describes the number of times an event occurs by comparing *P* and *Q*. The odds ratio can be calculated using eq. 4.

$$Odds\ ratio = \frac{P_y}{Q_y} \quad (4)$$

The likelihood value is one of the most important criteria to test for the fit of the logistic regression model. The statistical software will compute the likelihood value and convert it into a Chi–square distribution test with the independent variable's degree of freedom. The Chi–square test's main hypothesis is that "all logistic regression model coefficients are zero." Despite the rejection of the null hypothesis, the conclusion is that "logistic regression has a model." Based on the type I error percentage, the statistical software also provided the P–Value for deciding whether to accept or reject the null hypothesis. The typical type I error rate is 5%.

A rule of ten is a conservative estimate for estimating sample size in terms of sample size. On the other hand, applying type I and type II errors to the calculation can eliminate bias in regression coefficients and improve the predictive accuracy of the regression model [14]. The sample size calculation can be written in terms of the equation as eq. 5 [15]:

$$N = \left[\frac{(Z_{1-\alpha/2} + Z_{1-\beta})}{P_0(1-P_0)b^2}\right]^2 \quad (5)$$

Where, N = Sample size, α = Type I error, β = Type II error, $P_0 = P(y = 1|x = \mu_x)$, $b = ln\ (Odds\ Ratio)$

The formula for more than one variable is as eq. 6.

$$N^* = \frac{N}{1-R^2} \quad (6)$$

Where N is the value obtained from eq. 5, and R^2 (R–Square) is the value obtained by regressing the independent variable of interesting event.

Logistic regression is applied for root cause analysis. The semiconductor industry used regression analysis to define the root cause of the dropping yield and suggest factors to improve wafer processing [16]. Even though the six–sigma project focused on finding out the factors that influence the healthcare topic, logistic regression also applied [17]. Moreover, the microelectronics industry enhances the quality of their products by defining the root cause through a logistic regression analysis [18].

3.3 Independent and dependent variable data

Independent variables: the research has three independent variables. The first independent variable is the operator. Manufacturer has four operators (represented by Operators A, B, C, and D), with two working the day shift (6 a.m.–6 p.m.) and two operators work the night shift (6 p.m.–6 a.m.), rotating every two weeks. Then the type of shift is the second independent variable; there are day and night shifts. The third independent variable was suggested by the researcher, which is the model of product; there are 1, 1.5, and 2–meter–long.

Dependent variable: the research has one dependent variable, which is the test result. The interesting event is a product test pass with a set value of one for successful testing and zero for failed testing. Thus, “1” represented the “pass” of the test result, and “0” represented the “failed” of the test result. The passing test is an interesting event because the manufacturer is more concerned with the output in this situation.

A total of 233 units were built and tested for the first mass production. 128 units passed and 105 units failed, so the yield is 54.94%. After segregating by working shift, they found day shift has a yield of 54.33% (69 units passed from the total of 127 units) and night shift has a yield of 55.66% (59 units passed from the total of 106 units). Even though the difference is only one percent, they have a question, as well as the possibility that the operator's skill also may have contributed. **Table 1** is the data of the research.

Table 1 Raw data of the research

Product model	Operator	Shift	Test result
1Meter	B	NIGHT	1
1Meter	B	NIGHT	0
1Meter	B	NIGHT	1
1Meter	B	NIGHT	1
1Meter	A	DAY	1
1Meter	A	DAY	1
1Meter	A	NIGHT	1
1Meter	C	NIGHT	1
1Meter	B	NIGHT	0
1Meter	B	DAY	0
1Meter	B	DAY	1
1Meter	B	DAY	0
1Meter	A	DAY	1
1Meter	A	DAY	1
1Meter	B	NIGHT	1
1Meter	C	NIGHT	1
1Meter	A	DAY	1
1Meter	B	DAY	0
1Meter	B	DAY	0
1Meter	A	DAY	1
1Meter	B	NIGHT	0
1Meter	A	NIGHT	1
1Meter	C	DAY	1
1Meter	A	DAY	0

Table 1 Raw data of the research *(cont.)*

Product model	Operator	Shift	Test result
1Meter	A	NIGHT	0
1Meter	A	DAY	1
1Meter	B	DAY	0
1Meter	A	DAY	0
1Meter	C	NIGHT	1
1Meter	A	NIGHT	1
1Meter	B	NIGHT	0
1Meter	B	NIGHT	0
1Meter	B	DAY	0
1Meter	A	NIGHT	1
1Meter	B	DAY	1
1Meter	A	NIGHT	1
1Meter	A	DAY	0
1Meter	D	NIGHT	1
1Meter	B	DAY	0
1.5Meter	A	NIGHT	1
1.5Meter	C	NIGHT	1
1.5Meter	C	NIGHT	1
1.5Meter	C	DAY	1
1.5Meter	C	DAY	1
1.5Meter	A	DAY	1
1.5Meter	B	DAY	1
1.5Meter	B	DAY	0
1.5Meter	B	DAY	1
1.5Meter	C	NIGHT	1
1.5Meter	C	DAY	1
1.5Meter	B	DAY	1
1.5Meter	B	DAY	1
1.5Meter	B	DAY	1
1.5Meter	D	DAY	1
2Meters	D	DAY	0
2Meters	B	NIGHT	1

Table 1 Raw data of the research *(cont.)*

Product model	Operator	Shift	Test result
2Meters	B	NIGHT	1
2Meters	A	DAY	1
2Meters	A	DAY	1
2Meters	D	DAY	1
2Meters	A	DAY	1
2Meters	A	DAY	0
2Meters	B	NIGHT	1
2Meters	A	DAY	1
2Meters	B	DAY	0
2Meters	A	DAY	0
2Meters	B	NIGHT	1
2Meters	B	DAY	1
2Meters	B	DAY	0
2Meters	A	DAY	1
2Meters	B	NIGHT	1
2Meters	A	DAY	0
2Meters	B	NIGHT	1
2Meters	A	DAY	1
2Meters	B	DAY	0
2Meters	B	NIGHT	0
2Meters	A	DAY	1
2Meters	B	NIGHT	0
2Meters	B	NIGHT	1
2Meters	A	DAY	1
2Meters	B	DAY	1
2Meters	B	NIGHT	1
2Meters	A	DAY	0
2Meters	C	NIGHT	1
2Meters	B	NIGHT	0
2Meters	A	DAY	1
2Meters	A	NIGHT	1
2Meters	B	NIGHT	1

Table 1 Raw data of the research *(cont.)*

Product model	Operator	Shift	Test result
2Meters	A	NIGHT	0
2Meters	A	DAY	0
2Meters	B	NIGHT	1
2Meters	A	NIGHT	1
2Meters	A	DAY	1
2Meters	B	DAY	0
2Meters	B	NIGHT	1
2Meters	B	NIGHT	0
2Meters	A	NIGHT	1
2Meters	B	NIGHT	0
2Meters	B	NIGHT	1
2Meters	B	NIGHT	0
2Meters	B	NIGHT	0
2Meters	A	NIGHT	0
2Meters	B	NIGHT	0
2Meters	B	NIGHT	1
2Meters	A	DAY	1
2Meters	A	NIGHT	1
2Meters	A	NIGHT	0
2Meters	A	NIGHT	1
2Meters	D	DAY	1
2Meters	A	DAY	0
2Meters	A	NIGHT	1
2Meters	B	NIGHT	1
2Meters	A	DAY	1
2Meters	A	DAY	1
2Meters	B	DAY	0
2Meters	A	DAY	1
2Meters	B	NIGHT	1
2Meters	A	NIGHT	0
2Meters	A	NIGHT	0
2Meters	B	DAY	0

Table 1 Raw data of the research *(cont.)*

Product model	Operator	Shift	Test result
2Meters	B	DAY	0
2Meters	A	NIGHT	1
2Meters	A	NIGHT	1
2Meters	C	DAY	1
2Meters	B	DAY	0
2Meters	A	DAY	1
2Meters	A	NIGHT	0
2Meters	B	DAY	1
2Meters	A	NIGHT	1
2Meters	A	DAY	1
2Meters	B	DAY	0
2Meters	A	NIGHT	0
2Meters	B	DAY	0
2Meters	A	NIGHT	1
2Meters	A	NIGHT	0
2Meters	A	NIGHT	0
2Meters	B	NIGHT	0
2Meters	B	DAY	1
2Meters	A	NIGHT	1
2Meters	A	NIGHT	1
2Meters	B	DAY	0
2Meters	D	DAY	1
2Meters	A	NIGHT	1
2Meters	A	DAY	0
2Meters	A	NIGHT	1
2Meters	D	NIGHT	1
2Meters	B	DAY	1
2Meters	B	NIGHT	0
2Meters	A	NIGHT	1
2Meters	A	NIGHT	0
2Meters	B	DAY	0
2Meters	A	NIGHT	0

Table 1 Raw data of the research *(cont.)*

Product model	Operator	Shift	Test result
2Meters	D	NIGHT	1
2Meters	A	DAY	0
2Meters	A	NIGHT	0
2Meters	A	NIGHT	1
2Meters	D	NIGHT	1
2Meters	A	NIGHT	0
2Meters	A	NIGHT	1
2Meters	A	NIGHT	0
2Meters	A	NIGHT	1
2Meters	B	DAY	0
2Meters	C	DAY	0
2Meters	D	NIGHT	1
2Meters	D	NIGHT	0
2Meters	B	NIGHT	0
2Meters	B	DAY	0
2Meters	B	DAY	0
2Meters	A	NIGHT	1
2Meters	A	DAY	1
2Meters	B	NIGHT	0
2Meters	A	DAY	0
2Meters	A	DAY	1
2Meters	A	DAY	1
2Meters	A	DAY	0
2Meters	C	NIGHT	0
2Meters	A	DAY	1
2Meters	A	DAY	1
2Meters	C	NIGHT	0
2Meters	A	DAY	0
2Meters	A	DAY	0
2Meters	A	DAY	1
2Meters	A	DAY	0
2Meters	A	DAY	0

Table 1 Raw data of the research *(cont.)*

Product model	Operator	Shift	Test result
2Meters	B	NIGHT	0
2Meters	A	DAY	1
2Meters	A	DAY	1
2Meters	B	NIGHT	1
2Meters	A	DAY	0
2Meters	A	DAY	1
2Meters	A	DAY	0
2Meters	B	NIGHT	0
2Meters	A	DAY	0
2Meters	A	DAY	1
2Meters	D	DAY	1
2Meters	C	NIGHT	0
2Meters	A	DAY	0
2Meters	D	DAY	1
2Meters	A	DAY	0
2Meters	A	DAY	0
2Meters	A	DAY	0
2Meters	A	DAY	1
2Meters	A	DAY	1
2Meters	A	DAY	1
2Meters	B	NIGHT	0
2Meters	A	DAY	1
2Meters	A	DAY	1
2Meters	B	NIGHT	0
2Meters	C	NIGHT	0
2Meters	A	DAY	0
2Meters	A	NIGHT	1
2Meters	A	DAY	1
2Meters	D	DAY	1
2Meters	C	DAY	0
2Meters	C	NIGHT	0
2Meters	A	DAY	0

Table 1 Raw data of the research *(cont.)*

Product model	Operator	Shift	Test result
2Meters	A	DAY	1
2Meters	A	DAY	1
2Meters	C	DAY	0
2Meters	A	NIGHT	1
2Meters	D	DAY	1
2Meters	A	DAY	1
2Meters	D	DAY	0
2Meters	B	DAY	0
2Meters	D	DAY	1
2Meters	C	NIGHT	0
2Meters	B	DAY	0
2Meters	B	DAY	0
2Meters	C	NIGHT	0
2Meters	D	DAY	0
2Meters	C	NIGHT	0
2Meters	C	NIGHT	0
2Meters	C	NIGHT	0

4. Results

The research applied the Minitab version 19 for a logistic regression model analysis. Based on the research hypotheses, the result of the research hypothesis is as follow:

Hypothesis 1: Shift of working contributed to the test result. It is linked to the logistic model hypothesis as follows:

H_0: The coefficient of the shift of working is zero.

H_1: The coefficient of the shift of working is not zero.

Hypothesis 2: Operator contributed to the test result. It is linked to the logistic model hypothesis as follows:

H_0: The coefficient of the operator is zero.

H_1: The coefficient of the operator is not zero.

Hypothesis 3: Product model contributed to the test result. It is linked to the logistic model hypothesis as follows:

H_0: The coefficient of the product model is zero.

H_1: The coefficient of the product model is not zero.

The ANOVA table is presented in **Table 2** with the P–Value for testing the hypothesis.

Table 2 ANOVA of the logistic regression model with three independent variables

Source	DF	Chi–Square	P–Value
Regression	6	27.03	0.000
Product model	2	11.87	0.003
Operator	3	21.01	0.000
Shift	1	2.61	0.106

Referring to the P–value, the P–value of the "shift" variable is greater than 0.05, then this variable is not contributing to the test result at the 95% confidence interval. It means that the yield is comparable whether the working shift is during the day or at night. On the other hand, the two independent variables, "product model" and "operator," contributed to the test result. The two independent variables were re–analyzed, and the results are shown in **Table 3**.

Table 3 ANOVA of the logistic regression model with two independent variables

Source	DF	Chi–Square	P–Value
Regression	5	25.37	0.000
Product model	2	10.87	0.004
Operator	3	19.27	0.000

Note: The R–Square (adj.) is 8.86%.

Finally, the regression model has two independent variables that influence the probability of passing the test. The regression equation can be written as eq. 7.

$$Y' = 3.65 + 0.00(1.5\text{–}meter\text{–}long)\text{–}\,2.61\,(1\text{–}meter\text{–}long) - 3.23(2\text{–}meters\text{–}long) + 0.00(OperatorA)\text{–}\,1.15(OperatorB)\text{–}\,1.22\,(OperatorC) + 0.73(OperatorD) \quad (7)$$

It should be noted that the coefficients for 1.5–meter–long and Operator A are extremely small, and with a three–digit number, the coefficient appears to be zero but is not exactly zero. The R–Square Adjust is very low (8.86%), which means that only 8.86% of the two independent variables explain the probability of passing the test. Even though the R–Square Adjust is low, the result led to a conclusion based on the odds ratio in **Table 4**.

Table 4 The Odds Ratio for level A relative to level B of Product model

Level A	Level B	Odds Ratio
1.5–meter–long	1–meter–long	13.66
1.5–meter–long	2–meters–long	25.25
1–meter–long	2–meters–long	1.85

The odds ratio of the product model of **Table 5** indicated that the 1.5–meter–long product model has a higher chance of passing the test than the 1–meter–long product model 13.66 times and the 2–meter–long product model 25.25 times. 1–meter–long is better than 2–meter–Long about 1.85 times. Thus, the manufacturer needs to drill down more details for the different of software test, which is the testing revision of 1.5–meter–long is a new revision and the logic of calculate for the movement free compensation is different.

Table 5 The Odds Ratio for level A relative to level B of Operator

Level A	Level B	Odds Ratio
Operator A	Operator B	3.16
Operator A	Operator C	3.39
Operator D	Operator A	2.07
Operator B	Operator C	1.07
Operator D	Operator B	6.53
Operator D	Operator C	7.01

Operator D is the most skilled at carrying out the test. Operator A is ranked number two. Operators B and C have comparable performances. Operator D has a higher chance of passing the test than Operator A (2.07 times), Operator B (6.53 times), and Operator C (7.01 times). Thus, Operator D is a role model, sharing his practice with the others, and Operators A, B, and C need to be retrained. Operators B and C are a priority of retrained.

According to eq. 5, this is applied for the sample size calculation. Type I error is 5%, and Type II error is 20%. P_0 is 105 divided by 233, then P_0 is equal to 0.45, and the Odds ratio is 1.486 by calculation per Equation 4, then b = ln(1.486), which is 0.396. The sample size is rounded up to 203 units, which is 203 units for the one independent variable. The research had two dependent variables; thus, the sample size was calculated by applying eq. 6, and the R–Square is 0.088. Finally, the sample size with two dependent variables in the logistic regression is rounding up to 204 units. In actuality, the sample size of 233 units is appropriate to do the logistic regression analysis of this research.

5. Discussion

In essence, the research can provide some insight into the root cause of the likelihood of passing the test. On the other hand, other terms need to be verified, such as assembly process, material, and so on. Thus, this research can demonstrate how to apply logistic regression, which is new to this manufacturer as well, and they are aware that this statistical tool is beneficial enough to help them find the root cause. The operator skills as well as the software revision test have a significant impact. This root cause analysis can also be used in any field where there is a lot of interaction between humans and the testing machine, and the outcome is either pass or fail.

6. Conclusion

Based on the research hypothesis, there are three variables to prove that influence the passing test result: operator, working shift (day and night), and product model. The findings show that day or night shifts have no effect on the likelihood of passing the test. The operator and product model are important factors in the test passing. The 1.5–meter–long model has a higher chance of passing the test than the 1–meter–long model by about 13.66 times and the 2–meter–long model by 25.25 times. Operator D is the best at performing the robotic arm test and has a better chance of passing than the other operators. There are 2.07 times for Operator A, 6.53 times for Operator B, and 7.01 times for Operator C, respectively. The action is that the software test needs to be updated for the 1– and 2–meter–long models after comparison of the testing of each software test model. Moreover, Operators B and C need to be retrained as a priority, followed by Operator A (the option for manpower management). Then the manufacturer needs to focus more on other factors in the assembly process for yield improvement in the next step. Moreover, material traceability should be implemented, and data will be available for further analysis of the material portion.

7. Acknowledgements

The researcher is grateful to the Industrial Engineering Department, Faculty of Engineering, Rajamangala University of Technology Rattanakosin, for supporting and funding this research. The researcher also expresses gratitude to the manufacturer for providing the data for analysis and for cooperating until the result was reached.

References

[1] IFR. "World Robotics Report: All–Time High with Half a Million Robots Installed in One Year." IFR.com. https://www.universal–robots.com/in/blog/robotic–arm/ (accessed Nov. 9, 2022)

[2] S. E. Hashemi–Petroodi, S. Thevenin, S. Kovalev and A. Dolgui, "Operations management issues in design and control of hybrid human–robot collaborative manufacturing systems: a survey," *Annual Reviews in Control*, vol. 49, pp. 264–276, 2020, doi: 10.1016/j.arcontrol.2020.04.009.

[3] J. Holland, L. M. Kingston, C. Mccarthy, E. Armstrong, P. O'dwyer, F. Merz and M. McConnell, "Service Robots in the Healthcare Sector," *robotics,* vol. 10, no. 47, 2021, Art. no. 47, doi: 10.3390/robotics10010047.

[4] L. Gazzaneo, A. Padovano and S. Umbrello, "Designing Smart Operator 4.0 for Human Values: A Value Sensitive Design Approach," *Procedia Manufacturing*, vol. 42, pp. 219–226, 2020, doi: 10.1016/j.promfg.2020.02.073.

[5] K. Harish, D. Megha, M. Shuklambari, K. Amit and K. J. Chaitanya, "Pick and place robotic arm using Arduino," *International Journal of Science, Engineering and Technology Research (IJSETR*), vol. 6, no. 12, pp. 1568–1573, 2017.

[6] Y. Chen and F. Dong, "Robot machining: recent development and future research issues," *The International Journal of Advanced Manufacturing Technology*, vol. 66, pp. 1489–1497, 2013, doi: 10.1007/s00170–012–4433–4.

[7] R. Naboni, A. Kunic, A. Kramberger and C. Schlette, "Design, simulation and robotic assembly of reversible timber structures," *Construction Robotics*, vol. 5, pp. 13–22, 2021, doi: 10.1007/s41693–020–00052–7.

[8] C. Panyasai and R. Chockrungissaranukul. "*Industrial Robot Programming*." nectec.or.th. https://www.nectec.or.th/news/news–public–document/industrial-robot-programming.html (accessed Dec. 15, 2022)

[9] P. Michalik, V. Molnár, G. Fedorko, B. Stehlíková, P. Tirpák and J. Macej, “Research on the Influence of Production Technologies on the Positioning Accuracy of a Robotic Arm for Low–Handling Weights.” *applied sciences*, vol. 11, no. 13, 2021, Art. no. 6104, doi: 10.3390/app11136104.

[10] F. E. Harrell, “Binary logistic regression,” in *Regression modeling strategies*, 2nd ed. Aargau, Switzerland: Springer International Publishing, 2015, ch. 10, sec. 10.1–10.13, pp. 219–274.

[11] L. Niu, “A review of the application of logistic regression in educational research: Common issues, implications, and suggestions,” *Educational Review*, vol. 72, no. 1, pp. 41–67, 2020, doi: 10.1080/00131911.2018.1483892.

[12] J. Phillips, E. Cripps, J. Lau and M. Hodkiewicz, “Classifying machinery condition using oil samples and binary logistic regression,” *Mechanical Systems and Signal Processing,* vol. 60–61, pp. 316–325, 2015, doi: 10.1016/j.ymssp.2014.12.020.

[13] D. C. Montgomery, E.A. Peck and G.G. Vining, “Generalized linear model,” in *Introduction to Linear Regression Analysis*, 6th ed. Hoboken: John Wiley & Sons, 2021, ch. 13, sec. 13.2, pp. 441–462

[14] E. O. Ogundimu, D. G. Altman and G. S. Collins, “Adequate sample size for developing prediction models is not simply related to events per variable,” *Journal of Clinical Epidemiology*, vol. 76, pp. 175–182, 2016, doi: 10.1016/j.jclinepi.2016.02.031.

[15] F. Y. Hsieh, D. A. Bloch and M. D. Larsen, “A simple method of sample size calculation for linear and logistic regression,” *Statistics in medicine,* vol. 17*,* no. 14, pp. 1623–1634, 1998, doi: 10.1002/(sici)1097-0258(19980730)17:14<1623::aid-sim871>3.0.co;2-s.

[16] C. F. Chien, C. W. Liu and S. C. Chuang, “Analysing semiconductor manufacturing big data for root cause detection of excursion for yield enhancement,” *International Journal of Production Research*, vol. 55, no. 17, pp. 5095–5107, 2017, doi: 10.1080/002075 43.2015.1109153.

[17] F. H. van Meulen, T. Vermaat and P. Willems, “Case Study: An Application of Logistic Regression in a Six Sigma Project in Health Care,” *Quality Engineering*, vol. 23, no. 2, pp. 113–124, 2011, doi: 10.1080/08982112.2011.553761.

[18] N. Katemukda, “Applying the logistic regression for the quality of microelectronics product,” *Journal of Applied Statistics and Information Technology*, vol. 6, no. 2, pp. 25–35, 2021.