https://doi.org/10.25130/tjaes.17.54.3.31

Estimation and Prediction the Probability of Failure with Application

Assist. Prof. Dr. Bekhal Samad Sedeeq
College of Administration and Economics
Salahaddin University/Erbil

Researcher: Zhiyan Muhsin Jalal
College of Administration and Economics
Salahaddin University/Erbil

Abstract:

One of the reasons that reduces the production of factories is the failure of these machines that are used for production. One of the reasons for these failures is due to failures of the components of the machines, which is due to all the stress that these components face while working. That is why trying to make the life of these machines longer we need to know the (Probability of Failure) and (Cycle of Failure) for the components of these machines.

The Researcher, to achieve and obtain the main goal of this study, after very hard work started to collect data on five components (time of failure and repair) of these components in a steel company, the company that is in Erbil City Kurdistan Region of Iraq. The Researcher reached a conclusion finding the probability of failure and cycle to failure for each component in these machines, also estimating the value of cycle to failure for nine specific values of the probability of failure. We also designed a model for predicting the probability of failure for each machine including all the components together through using Factor Analysis and Multiple Regression Analysis.

STAT Graphics, SPSS and Microsoft Excel have been used for all the calculations, graphics, and analysis.

Keywords: Reliability, Probability of failure, Cycle to failure, Regression analyze.

التقدير والتنبؤ باحتمالية الفشل مع التطبيق

أ.م.د. بيخال صمد صديق
كلية الإدارة والاقتصاد
جامعة صلاح الدين/أربيل
Bikhal.sedeeq@su.edu.krd

الباحثة: ژيان محسن جلال
كلية الإدارة والاقتصاد
جامعة صلاح الدين/أربيل
zhianjalal782@gmail.com

المستخلص:

أحد الاسباب تخفيض نسبة الانتاج في المصانع هو تعطيل المكائن الموجودة والمستخدمة في عملية الانتاج، والسبب التي يؤدي الي تعطيل هذه المكائن يعود الي الاعطال/عطل المكونات الموجودة داخل المكائن، والذي يرجع كل أنواع الضغط التي تتعرض لها هذه المكونات أثناء العمل، لذلك السعي من اجل إطالة عمر هذه المكائن يكون بمعرفة (احتمالية الفشل) و(دورة الفشل) لمكونات هذه المكائن.

للوصول الي الهدف الاساسي للدراسة قام الباحث بعد عمل شاق جداً بدأ بجمع البيانات عن خمسة مكونات متعلقة بـ (وقت تعطيل واصلاح) هذه المكونات في معمل صناعة الحديد والتي تقع في محافظة اربيل اقليم كردستان العراق. توصل الباحث الي استنتاج أيجاد احتمالية الفشل ودورة الفشل لكل مكون في هذه المكائن، وتم التقدير قيمة دورة الفشل لتسعة قيم احتمالية للفشل، وبعدها

قمنا بتكوين نموذج لغرض التنبؤ احتمالية الفشل لكل ماكنة بمكوناته وذلك بواسطة استخدام تحليل العاملي و تحليل الانحدار المتعددة.

تم استخدام (Statgraphic, SPSS, Microsoft Excel) لكل الإجراءات الحسابية والرسومات والتحاليل التي توصلنا اليها في هذه الدراسة.

الكلمات المفتاحية: المعولية، احتمالية الفشل، دورة الفشل، تحليل الانحدار.

1. introduction:

The reliability discipline started in the early 1930s, when probability ideas were applied to problems involving electric power generation. Germans used simple reliability principles to increase reliability during World War II. Many factors have influenced the consideration of reliability in product design, including product complexity, the use of reliability-related clauses in design requirements, competition, cost-effectiveness perception, public demand, and so on. As well as previous device failures (dhillon, 1999: 16). The probability of applying a service or product for a given time period under specified conditions without failure is defined as reliability. Clearly, reliability is an indicator of a product's potential dependability and trustworthiness based on its physical consistency over time. (Ohring, 1998: 3).

2. Function of Reliability: Reliability is the capability of the item to perform the required function for a specified period of time under given operating conditions. Reliability function, is defined as the probability that the system, devise or components will not failure during the specified period time (t), under specified operating conditions. If TTF indicates the time -to-failure random variable with failure function (cumulative distribution function) F(t), then the function of reliability is given by:(Kumar, 2000: 61)

$R(t) = P(T > t), \qquad t \geq 0$ ………. (1)

R (t) = P {the system does not failure during [0, t]} = I - F(t)

The Reliability function has the following properties:

A. Reliability is a reducing function with time t. That is, for: $t_1 < t_2; R(t_1) \geq R(t_2)$.

B. It is mostly assumed that R (0) = 1. As t becomes bigger and bigger R (t) approaches zero, that is, $R(\infty)$.

3. Failure Density Function: Represents the failure distribution over the entire period. The greater the value of f (t), the more failures occur in a small period of time around t. It is the essential for deriving other metrics and performing in-depth analyses. (Meeker and Escobar, 1998: 28).

$f(t) = p_r(t < T \leq t + \Delta t)$ ………………… (2)

4. **Mean Time to Failure:** A crucial reliability measures is the mean time to failure, which is the average time to the first failure. (Todinov, 2005: 33) and (Yang-2007: 16)

   The MTTF can be written with integration as the following:

   $$MTTF = E(t) = \int_0^\infty t.f(t)dt = \int_0^\infty R(t)dt \qquad (3)$$

5. **Mean Time between Failure (MTBF):** Is its expected time between two successive failures (repairable component) (Bertsche, 2008: 31).
6. **Failure:** A failure happens when the any component or devise stops performing its required function, applying it to a complex item can be hard. On the other hand, is described as when a structure can no longer perform its function safely and reliably, whether as a result of deterioration, deformation, wear, or breakage; the conditions would have breached the limit and serviceability states used in production. (Qua et al., 2015: 1).
7. **Prediction:** The aim of reliability prediction is to solve problems that have been found prior to the actual product or service, resulting in highly reliable product. To this end, while the prediction of reliability result is significant, it is also critical to forecast reliability early on and provide feedback based on the prediction results. That is why predicting reliability is so important. (lee and lee, 2008: 22).
8. **Stress:** The stress methods used in acceleration lifetime tests are constant stress and step stress. The constant stress method is a lifetime experiment where stress, such as temperature or voltage, is held constant and the level of degradation of properties and time-to-failure lifetime classification are evaluated. In the step stress method, contrary to the constant stress method, the time is kept constant and the stress is dramatically increased in steps and the level of stress causing failure is observed. (Bernstein, 2014: 26).
9. **S-N Diagram:** Data are charted as stress (s) versus the logarithm of the number (N) of cycle to failure for each of the specimens in a curve called S-N curve, also called the Wohler curve. A sequence of the experiment is begun by subjecting a specimen to the stress cycling at a relatively large maximum stress amplitude (σ_{max}), usually on the older of two-thirds of the static tensile strength; the number of cycles to failure is counted. This

technique is a replication on other specimens at progressively declining maximum stress amplitudes. (Callister and Rethwisch, 2018: 231)

10. Cyclic Stresses: Some practical applications, involve cycling between high and low stress levels that are steady. This is called constant amplitude stressing and as shown in Fig. (1.1). The stress range, $\Delta\sigma = \sigma\max - \sigma\min$, is the difference between the maximum and the minimum values. Averaging the maximum and minimum values gives the mean stress, σ_m. The mean stress may be zero, as in Fig. (1), (a), but often it is not, as in (b). Half the range is called the stress. Amplitude, σ_m, which is the variation about the mean. The mathematical expressions for these basic definitions are as follows: (Dowling, 2013: 418).

$$\sigma_a = \frac{\Delta\sigma}{2} = \frac{\sigma_{max} + \sigma_{min}}{2}, \quad \sigma_m = \frac{\sigma_{max} + \sigma_{min}}{2}$$

The signs of σ_a and $\_\sigma_m$ are all the time positive, since $\sigma_{max} > \sigma_{min}$, where tension is considered positive. The quantities σ_{max}, σ_{min}, and σ_m, it's possible to be either positive or negative. Sometimes, the following ratios of two of these variables are used:

R = $\frac{\sigma_{min}}{\sigma_{max}}$, $A = \frac{\sigma_a}{\sigma_m}$; where R is called the stress ratio and A the amplitude ratio. Some additional relationships derived from the preceding equations are also useful:

$$\sigma_a = \frac{\Delta_\sigma}{2} = \frac{\sigma_{max}}{2}(1-R) \quad , \sigma_m = \frac{\sigma_{max}}{2}(1+R) \qquad (4)$$

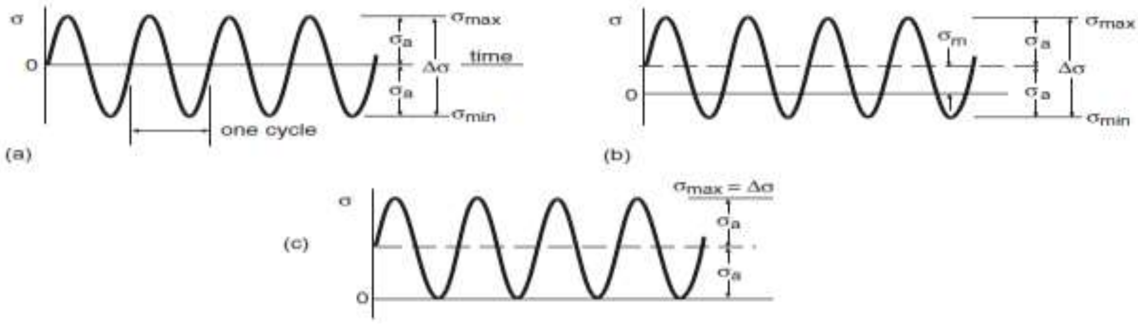

Figure (1): illustrate unchanging amplitude cycling and its related classification. Case-(a) is absolutely reversible when it comes to stressing, $\sigma_m = 0$, (b) has a nonzero mean stress , σ_m, and (c) is zero-to-tension stressing , $\sigma_m = 0$.

11. **Cycle to Failure and Stress:** Cycle to failure and Stress could be calculated in the following way from the original data: (Joshi; Dhaug and Pandey, 2004: 85, 163-168)

Cycle to failure=N = $t_{fi} - t_{ci} + 1$; i = 1, 2, …., n …….. (5)

Where:

t_f : failure's date

t_c : reparation's date

Stress = S = $\sum_{i=1}^{k} P_i$; i = 1, 2, 3, k ……. (6)

Where:

P_i: Production on time in the minute i.

k: minute of failure.

12. **Probability of Failure:** The values within each machine number should be ranked in the dictate of the number of cycles to failure, and the probability of failure is counted from the equation: (Joshi; Dhaug and Pandey, 2004: 85, 163-168)

$P_f = \frac{m}{u+1}$ ………. (7)

Where:

m: is the rank of the specimen.

u: is the total number of specimens within one stress level.

13. **Factor analysis:** Factor analysis is a method of multivariate statistical analysis that is used in the analysis of the correlation matrix and the covariance matrix to determine the factors of the hypothesis behind the nature of the relationships (Internal relationships) through the common factors causing these relationships to discover a modern set of variables less in number than the original set of variables (Hypothetical variables). The factor is his ability to summarize the many variables and arrange them into a small number of hypothesized variables.

These factors are not related (orthogonal) to each other and reflect the common variance between the variables, and thus we get rid of the problem of Multicollinearity. (Ali and Muhammad, 2019: 23, 14-29).

14. **Linear regression:** Is a model that is linear in the parameters, neglect of the shape of the predictors in the model. (Ryan, 2007: 233)

14.1.Simple linear regression:

It is the most basic type of relationship that studies the relationship between two variables, the first of which is known as the independent (explanatory) variable and the second of which is known as the dependent (response) variable. (Hadar and Mahadevan, 2000: 157)

$Y_i = \beta o + \beta_1 X_i + \varepsilon_i$ ………. (8)

14.2.Multiple linear regression:

A regression model that contains more than one predictor variable. The ordinary least square method is used to estimate the coefficient of the multiple linear regression model. (Ibrahim and Mahdi, 2011: 59).

The following is the model for the i^{th} observation:

$$y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \beta_3 x_{i3} + \cdots + \beta_k x_{ik} + \varepsilon_{i,} \quad i = 1,2,3,\dots n \qquad (9)$$

$$\begin{bmatrix} y_1 \\ y_2 \\ \cdot\cdot \\ \cdot\cdot \\ y_n \end{bmatrix}_{n\times 1} = \begin{bmatrix} \beta_o & \beta_1 x_{11} & \cdots & \beta_k x_{1k} \\ \beta_o & \beta_1 x_{21} & \cdots & \beta_k x_{2k} \\ \cdot\cdot & \cdot\cdot & \cdots & \cdots \\ \cdot\cdot & \cdot\cdot & \cdots & \cdots \\ \beta_o & \beta_1 x_{n1} & \cdots & \beta_k x_{nk} \end{bmatrix}_{n\times(k+1)} \begin{bmatrix} \beta_o \\ \beta_1 \\ \cdot\cdot \\ \cdot\cdot \\ \beta_k \end{bmatrix}_{(k+1)\times 1} + \begin{bmatrix} U_1 \\ U_2 \\ \cdot\cdot \\ \cdot\cdot \\ U_n \end{bmatrix}_{n\times 1}$$

$$\underline{Y} = X\underline{\beta} + \underline{U}$$

Where:

Y: The vector for observations values of predict variable of order (n × 1).

X: The matrix of explanatory variables of order [n × (k + 1)].

β : The vector of unknown coefficients for model of order [(k + 1) × 1].

U: The vector of random errors of order (n × 1).

14.2.1. F-test: Testing coefficients significance by using F-test, this test aims to measure the significantly of the linear relationship between the predict variable and the explanatory variables through the hypothesis: (Ibrahim and Mahdi, 2011: 67).

Source of variation	Degrees of freedom	Sum of squares	Mean square	F_0
Regression	K	$\hat{\beta}'x'y - n\bar{y}^2$	$\hat{\beta}'x'y - n\bar{y}^2$ /k	$= \frac{MSR}{MSE}$
Error or Residual	n-k-1	$y'y - \hat{\beta}x'y$	$y'y - \hat{\beta}x'y$ /n-k-1	
Total	n-1	$y'y - n\bar{y}^2$	$y'y - n\bar{y}^2$ /n-1	

14.2.2. Coefficient of determination (R^2): To assess how good, the overall model fits the data, the coefficient of determination (R) can be consulted. This statistic represents the percentage of variation explained in the dependent variable by the independent variables in the model. R-squared is always between 0 and 1. In general, the higher the R-squared, the better the model fits your data. (Samprit Chatterjee and Jeffrey s. simonoff-2013: 2I)

15. **Application:** Machine malfunction is one of the problems that the factories face, and that will negatively impact the production of these factories. The researcher has tried very hard to collect data from different factories for these machines, but due to different reason she was not able to collect data from these different factories. However, she was able to collect data for the year of (2019) from a steel company, which is in Erbil city.

    This factory is consisted of five machines, for producing and manufacturing steel. Each of one of them is consisted of different components. After paying several visits to the factory, the outcome was collecting and studying data on all the components, in terms of (the time of machine failure and the time of repair) for each machine. This was done to build a model for predicting the probability of failure for each machine including all the components together. This is helpful for the factories to know or estimate the time of these failures and to be more alert, that process will ultimately benefit the companies and factories to increase production.

16. **Delineation of the Data:** In this study a data was taken from five machines, each machine contains five components. Each component has its own history of failure. This history contains certain steps like the appearance of a complete failure, the performance of cleaning or changing a part or all the parts, and the date of failure and reparations. Thus, data about 5 related components to failure are collected.

    Unfortunately, sometimes the data was incomplete, because of the laziness or any other reason. Hereby, it was not possible to deal with all the components, so we concentrated on the components which had enough data to find out a statistical model. the following components were analyzed, and it was known that they have a deferent number of failures:

A. Electric Arc Furnace had (232) failures.
B. Continues Casting Machine had (290) failures.
C. Reheat Furnace had (261) failures.
D. Rolling Mill Line had (263) failures.
E. Cooling Bed had (229) failures.

17. **Calculation of Cycle to Failure and Stress:** The first step we should calculated the Cycle to Failure and Stress by using equations (5) and (6) from the original data.

The calculated *N* and *S* for Electric Arc Furnace appear particularly in Table (1) below.

Table (1): Calculated cycle to failure and stress for Electric Arc Furnace

Failure	Cycle to failure (N)	Stress (S)
1	3	2555
2	8	6240
3	9	5783
…	…	…
…	…	…
…	…	…
230	4	4221
231	8	7790
232	5	4730

18. **Calculation of Probability of Failure:** is calculated for each machine, using via equation (7). Table (2) below shows the calculated value of probability of failure for each machine appearing particularly for Electric Arc Furnace, In this table, cycle to failure is arranged in an ascending order and ranked.

Table (2): Calculation of Probability of failure for (Electric Arc Furnace)

S (Stress)	MN (Machine number)	N	Rank	P_f	log N
850	1	1	1	0.017241	0
642	1	1	2	0.034483	0
855	1	2	3	0.051724	0.30103
..	..	..	..	..	..
..	..	..	..	..	..
23548	5	18	47	0.94	1.255273
14488	5	18	48	0.96	1.255273
17002	5	18	49	0.98	1.255273

We can plot the $P_f - N$ Diagram by using the calculated P_f versus the logarithm of cycle to failure, for the (Electric Arc Furnace) as it is shown in figure (2).

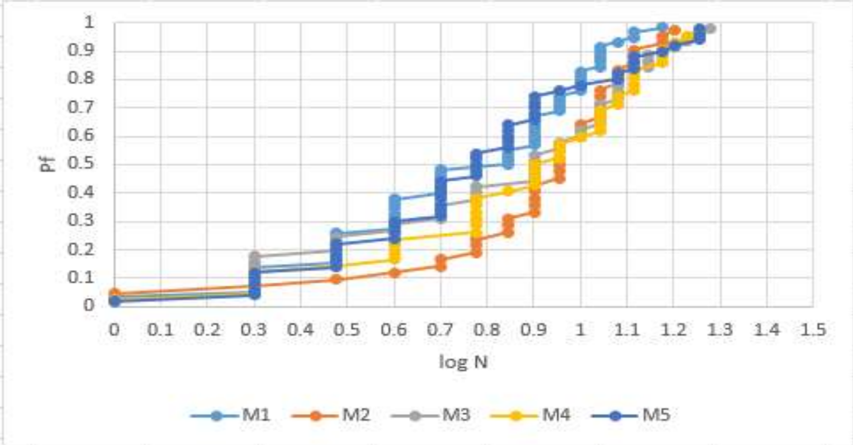

Figure (2): P_f *-N Diagram* for (Electric Arc Furnace)

The Figure above illustrates the relation between P_f and log N, and it can be used to estimate log N for different machines. (See Table 3).

19. Estimation Cycle to failure: From the Figures of the above Figure (2) we select Nine probability values from (0.1 to 0.9) Using the intersection points between the selected probability value (P_f) and the curve, Log(N) for each Machine is found out Table (3) will show the results regarding the **Electric Arc Furnace** .

Table (3): Log (N) for each Machine at different probability of failures for (Electric Arc Furnace)

P_f	M 1	M 2	M 3	M 4	M 5
0.1	0.301029	0.602059	0.301029	0.301029	0.301029
0.2	0.477121	0.778151	0.477121	0.602059	0.477121
0.3	0.602059	0.845098	0.698970	0.778151	0.602059
0.4	0.698970	0.903089	0.778151	0.845098	0.698970
0.5	0.845098	0.954242	0.903089	0.903089	0.778151
0.6	0.903089	1	1	1.041392	0.845098
0.7	0.954242	1.041392	1.041392	1.079181	0.903089
0.8	1	1.079181	1.079181	1.113943	1.079181
0.9	1.041392	1.113943	1.204119	1.176091	1.176091

20. Factor Analysis Procedure: After estimation Log N at nine probability values from (0.1 to 0.9) for each component in all the machines, we can use the factor analysis procedure to reduce the huge date for each component by finding MN, P_f, Log N as it is shown in table (3). The table that follows displays the number of factors and their corresponding eigenvalues, the percentage of variance describing and the cumulative percentage.

Table (4): Factor Analysis

Factor Number	Eigenvalue	Percent of Variance	Cumulative Percentage
1	1.9227	64.090	64.090
2	1.0	33.333	97.423
3	0.0772995	2.577	100.000

The following scree plot graphically shows the eigenvalues for each factor and suggests that there are two predominant factors.

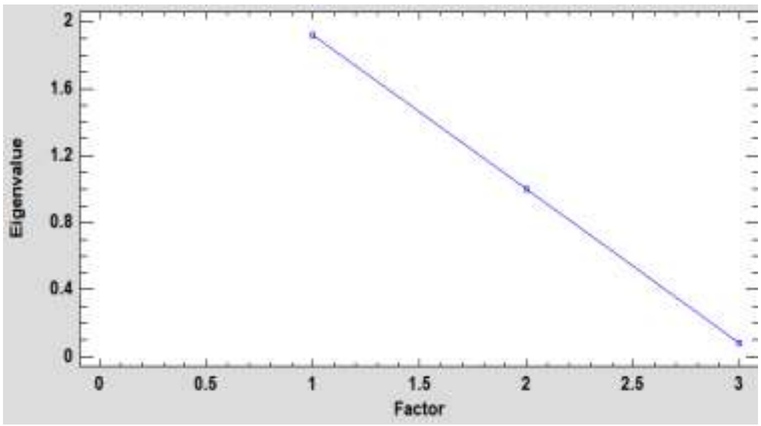

Figure (3): Scree plot

Table (5): Factor Loading Matrix (before and after Rotation)

	Factors Loading Matrix Before Rotation			Factors Loading Matrix After Varimax Rotation		
	Factor 1	Factor 2	Factor 3	Factor 1	Factor 2	Factor 3
P_f	0.98012	0.0270495	0.196523	0.98051	0.00564798	0.196376
MN	-0.02652	0.999634	-0.0053178	-0.00623	0.999978	0.00236461
Log N	0.98048	2.71993E-17	-0.196595	0.98024	-0.0183828	-0.196939

Factor scores were calculated for each component, using the following equations:

Score 1 = S_{1i} = 0.980512* $\mathbf{P_f}$ -0.00623368***MN** + 0.980243***Log N**

Score 2 = S_{2i} = 0.00564798* $\mathbf{P_f}$ + 0.999978***MN** + -0.0183828***Log N**

Score 3 = S_{3i} = 0.196376* $\mathbf{P_f}$ + 0.00236461***MN** -0.196939***Log N**

Where

i = 1,2,3, , 45.

The results above show that we need only two factors to describe the data, but for the purposes of model building, we are going to keep the three factors. The scores for **Electric Arc Furnace** after rotation are appearing particularly in Table (6) below.

Table (6): Factor Scores for (Electric Arc Furnace)

P_f	Score1	Score2	Score3
0.1	-3.51925	-1.36904	0.102896
0.1	-2.36942	-0.691494	-0.127337
0.1	-3.52797	0.029341	0.106203
0.1	-3.53233	0.728531	0.107856
0.1	-3.53669	1.42772	0.10951
0.2	-2.46858	-1.37954	0.042457
.	…	…	…
.	…	…	…
0.8	2.07542	1.38691	0.036534
0.9	2.32347	-1.40497	0.134236
0.9	2.59728	-0.710996	0.0800025
0.9	2.93867	-0.018289	0.0121922
0.9	2.82685	0.682916	0.0354358
0.9	2.82249	1.38211	0.0370892

The approach of this work here is to find out a general model, which could be used not only for one component, but for two or more than two components, we can be using the scores of the five components together and they are used considered as a basis for forecasting probability of failure for all machines.

The table below (7), shows the scores all components for all machines at any level of probability of failure and the whole Table we put in the appendix.

Table (7): Factor Scores for (all components)

P_f	Component Name	Score_1	Score _2	Score _3
0.1	Electric Arc Furnace	-3.51925	-1.36904	0.102896
0.1	Electric Arc Furnace	-2.36942	-0.691494	-0.127337
0.1	Electric Arc Furnace	-3.52797	0.029341	0.106203
0.1	Electric Arc Furnace	-3.53233	0.728531	0.107856
0.1	Electric Arc Furnace	-3.53669	1.42772	0.10951
.	…	…	…	…
.	…	…	…	…
0.9	Cooling Bed	2.94815	-1.34151	0.00997616
0.9	Cooling Bed	2.66155	-0.658584	-0.0595029
0.9	Cooling Bed	2.0631	0.00749048	-0.207411
0.9	Cooling Bed	2.74239	0.7426	-0.0339684
0.9	Cooling Bed	2.36038	1.42037	-0.127441

21. **Model Forecasting:** A Multiple Linear Regression Analysis was used taking the scores as independent variables and the variable P_f as dependent variable see Table (8, 9) separately. The results show that this model is adequate, and we are able to predict the above variable P_f with the help of the overall scores.

Table (8): Multiple Regression Analysis
(Dependent variable: P_f, machine1)

Parameter	Estimate	Standard Error	t Statistic	P-Value
Constant	0.440898	0.207221	2.12767	0.0394
Score_1	0.130793	0.00421775	31.0101	0.0000
Score_2	-0.0411656	0.147832	-0.278461	0.7821
Score_3	-0.212942	0.0924031	-2.30449	0.0263

Table (9): ANOVA table

Source	Sum of Squares	Df	Mean Square	F-Ratio	P-Value	R-squared (Adj) %
Model	2.89126	3	0.963752	363.36	0.0000	96.11
Residual	0.108744	41	0.0026523			
Total (Corr.)	3.0	44				

The table (8) shows that only score 1 and 3 are significant but score 2 is not statistically significant this model because of its P-value <0.05. So we can predict the $P_{f\ for}$ machine 1 as following:

$$\hat{P}_f = 0.440898 + 0.130793 score1 - 0.232942 score3 \ \ldots\ldots\ldots \ (10)$$

Table (9) shows that Analysis of variance, the results that represent the mean square for model and residual is equal to (0.963752, 0.0026523) respectively, the calculated value of F is 363.36 and P-value is less than signification level, The R-Squared adjusted statistic indicates that the model as fitted explains 96.11% of the variability in P_f ,which is suitable for forecasting. Likewise, Pf for machines 2, 3, 4, and 5 can be predicted via using the result in tables (10-17).

Table (10): Multiple Regression Analysis
(Dependent variable: P_f, machine 2)

Parameter	Estimate	Standard Error	t Statistic	P-Value
Constant	0.138564	0.107663	1.28701	0.2053
Score-1	0.144555	0.00455728	31.7195	0.0000
Score-2	-0.503466	0.152335	-3.30498	0.0020
Score-3	-0.0777442	0.0926609	-0.83901	0.4063

Table (11): ANOVA table

Source	Sum of Squares	Df	Mean Square	F-Ratio	P-Value	R-squared (Adj) %
Model	2.89781	3	0.965936	387.54	0.0000	96.3443
Residual	0.102193	41	0.00249251			
Total (Corr.)	3.0	44				

In the table (10) illustrates, that it is only by score1 and score2 we can predict the P_f for machine2 but we cannot predict P_f for machine 2 by score3 through the following equation:

$$\hat{P}_f = 0.138564 + 0.144544\,score1 - 0.503466\,score2 \quad \dots\dots\dots (11)$$

As far as table (11) is concerned, the results show that the model in equation (11) is adequate.

Table (12): Multiple Regression Analysis
(Dependent variable: P_f, machine 3)

Parameter	Estimate	Standard Error	t Statistic	P-Value
Constant	0.510449	0.00570721	89.4394	0.0000
Score_1	0.137119	0.00301407	45.493	0.0000
Score_2	0.296969	0.131096	2.26528	0.0288
Score_3	-0.350063	0.0656151	-5.33509	0.0000

Table (13): ANOVA table

Source	Sum of Squares	Df	Mean Square	F-Ratio	P-Value	R-squared (Adj)
Model	2.94299	3	0.980996	705.46	0.0000	97.9605
Residual	0.0570134	41	0.00139057			
Total (Corr.)	3.0	44				

Table (12), explains that the equation (12) can be used for predicting probability of failure for machine3.

$$\hat{P}_f = 0.510449 + 0.137119\,score1 + 0.296969\,score2 - 0.350063\,score3 \quad \dots\dots\dots (12)$$

By comparing p-value with level of significance in the table (13), then we realized the model has significance, and the outcome R-square adjust (97.9605) hence appropriate for forecasting.

Table (14): Multiple Regression Analysis
(Dependent variable: P_f, machine 4)

Parameter	Estimate	Standard Error	t Statistic	P-Value
Constant	0.565276	0.11701	4.83101	0.0000
Score-1	0.139539	0.00357901	38.9882	0.0000
Score-2	-0.117723	0.167143	-0.704325	0.4852
Score-3	-0.164921	0.0949362	-1.73718	0.0899

Table (15): ANOVA table

Source	Sum of Squares	Df	Mean Square	F-Ratio	P-Value	R-squared (Adj) %
Model	2.92379	3	0.974597	524.33	0.0000	97.2738
Residual	0.0762084	41	0.00185874			
Total (Corr.)	3.0	44				

Table of (14) points out that the P_f for machine 4 can be predicted by using the equation number (13)

$$\hat{P}_f = 0.565276 + 0.139539 score1 \text{ ………. (13)}$$

The results of table (15) show that the model is fitting.

Table (16): Multiple Regression Analysis

(Dependent variable: P_f, machine 5)

Parameter	Estimate	Standard Error	t Statistic	P-Value
Constant	-0.410077	0.303877	-1.34948	0.1846
Score_1	0.128225	0.00539603	23.7627	0.0000
Score_2	0.662745	0.215936	3.06917	0.0038
Score_3	-0.510939	0.112934	-4.52422	0.0001

Table (17): ANOVA table

Source	Sum of Squares	Df	Mean Square	F-Ratio	P-Value	R-squared (Adj) %
Model	2.92379	3	0.974597	524.33	0.0000	95.9102
Residual	0.0762084	41	0.00185874			
Total (Corr.)	3.0	44				

The table (16) shows that all score 1, 2 and score 3 are significant, and the following equation can be predicted:

$$\hat{P}_f = -0.410077 + 0.128225 score1 + 0.662745 score2 - 0.510939 score3 \text{ ……. (14)}$$

The ANOVA table (17) displays that the mean square for model is equal to (0.96189, 0.0027885) respectively, and the calculated value of F is 344.95 and P-value is less than the signification level. The value of R-squared adjust indicates that the model which is suitable for forecasting is fitting.

22. Conclusion: According to the analysis of the results from the practical part, the following concluded points have been drawn:

A. continues casting machine is the component with the highest number of failure whereas cooling bed is the lowest number of failure.

B. We could find the value of probability of failure and cycle to failure through knowing the time of failure and repair of the components in each machine.

C. Estimating cycle to failure for nine specific value of probability of failure for all components in each machine by $P_f - N$ using Diagram.

D. Regression analysis and factor analysis are suitable for this kind of data.

E. We can predict probability of failure for (machine1 by using equation (10), machine2 by using equation (11), machine3 by using equation (12), machine4 by using equation (13), machine5 by using equation (14).

F. According the results in the table (8, 9, 10, 11, 12, 13, 14, 15, 16 and 17) the model are adequate in all machine.

References:

1. Ali, I. M. & Muhammad, Z. A., (2019), A statistical study on the most important factors that lead to the increase of car accidents in Erbil governorate using the analysis and cluster analysis. J Zanco Journal of Humanity Sciences, 23**,** 14-29.
2. Bertche, B., (2008), Reliability in Automotive and Mechanical Engineering. © Springer-Verlag Berlin Heidelberg 2008.
3. Bernstein, J. B., (2014), Reliability Prediction from Burn-In Data Fit to Reliability Models. British Library Cataloguing-in-Publication Data.
4. Chatterjee, S. & Simonoff, J.S., (2013), Handbook of Regression Analysis. Published by John Wiley & Sons, Inc., Hoboken, New Jersey published simultaneously in Canada.
5. Callister, JR. WD & Rethwisch, D.G, (2018), Materials Science and Engineering. An introduction. John Wiley & Sons, Inc., 111 River Street, Hoboken,NJ 07030-5774.
6. Dowling, NE., (2013), Mechanical Behavior of Materials.4th edition. Published by Pearson Education. The United States of America.
7. Dhillon, B. S., (1999), Design Reliability (Fundamentals and Application) CRC Press LLC.
8. Hadar, A. & Mahadevan, S., (2000), Probability, Reliability, and Statistical Methods in Engineering Design. John Wiley & Sons, Inc. USA.
9. Ibrahim, Nazir Abbas & Mahdi, Degla Ibrahim, (2011), Econometrics and Its Applications, Al-Jazeera Printing and Publishing Office, Iraq-Baghdad.
10. Joshi, Y. P, Dhaug, N and Pandey, B. B., (2004), Probability Analysis of Fatigue of Paving Concrete. Institution of Engineering (India) Journal, IE (I), Vol, 85, PP163-168.
11. Kumar, U., Crocker, J., Knezevic, J. & El-Haram, M., (2000), Reliability,Maintenance and Logistic Support: a life cycle approach. kluwer Academic.
12. Lee, S. W & Lee, H., (2008), Reliability prediction system based on the failure rate model for electronic components. Journal of mechanical science and technology, 22 (2008) 957-964.
13. Meeker, W., Q., & Escobar, L. A., (1998), Statistical Methods for Reliability data. John Wiley & Sons, Canada. Library of congress cataloging-in-publication data.

14. Ohring, M., (1998), Reliability and Failure of Electronic Materials and Devices. United Kingdom Edition published by academic press limited an imprint of Elsevier.
15. Qua, H. Q; Tan, C, S; Wong, K; Ho, J; Wang, X; Yap, E; Ooi, I & Wong, Y., (2015), Applied Engineering Failure Analysis; Theory and Practice. Taylor & Francis Group, LLC.
16. Ryan, T. P., (2007), Modern Engineering Statistics. John Wiley & Sons, Inc., Hoboken, New Jersey, Canada.
17. Todinov, M.T., (2005), Reliability and Risk Models: Setting Reliability Requirements. John Wiley & Sons Ltd, ISBN 0-470-09488-5 (HB).
18. Yang, G., (2007), Life cycle reliability engineering. John Wiley & Sons, Inc.