v. 2, n. 1, 2022

Journal of
Engineering
Research

VARIABLE-RATE PROBABILITY DISTRIBUTION THEOREM

Alejandro Aristizábal Correa
ORCID: 0000-0001-5449-979X

Dedicated to my Grandfather, Alejandrino (Nino)

INTRODUCTION

The following theorem consists of a distribution function for any variable-rate process or system with a given lifecycle. It was the result of an effort for improving the manner how the probability of disruption is frequently estimated in organizational processes during a risk analysis.

The result of the research was a probability distribution function for any stochastic process with a variable-rate function.

Reliability engineering, organizational processes and IT systems are direct applications of the Variable-Rate Probability Distribution (VRPD) theorem.

For example, the probability of failure of an electro-mechanical machine component in any time interval, that sometimes, requires complicated approaches through probability distributions that not always match the variable distribution, is easily performed through the VRPD theorem by merely considering the failure rate function.

Organizational processes and IT systems evolve with time, which means that, optimization and patches respectively, the probability of interruption decreases along the process/system (p/s) lifecycle.

Keywords: Evolutionary Process, Failure rate, Variable-rate probability distribution.

THEOREM

The probability of one or more interruptions (probability of success: $P(X=1)$) of a p/s in any interval $t_\alpha \leq t \leq t_b$ along its lifecycle with a known interruption rate function $\lambda(t)$ is given by:

$$P(X=1)_{t_a \leq t < t_b} = \frac{1}{K(t)} \int_{t_a}^{t_b} \lambda(t)dt$$

With K, the subspace of events of the p/s along its lifecycle

POSTULATE

There will be at least one interruption along the process/system lifecycle.

PROOF

Suppose an experiment with m number of processes or systems (p/s) of the *same kind*, starting their lifecycle at the same time ($t=0$).

Let us take a timeframe at t_i out of the p/s lifecycle of size $\Delta T = t_b - t_\alpha$.

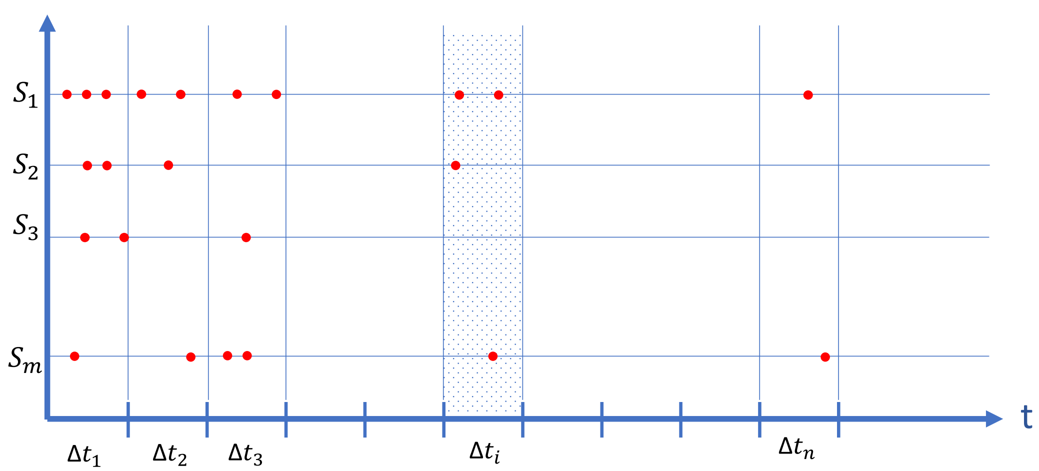


Figure 1 – failures of m cloned systems running

The mean count of systems interrupted c_i out of m systems on Δt_i is given by[1]: $\frac{c_i}{m}$, which is the *empirical probability of interruption* of the p/s type, as it is stablished by the theorem of large numbers for Bernoulli trials:

$$\lim_{m\to\infty} P\left(\left|\frac{c_i}{m} - p_i\right| \geq \epsilon\right) = 0 \ (eq1)$$

With p_i: Probability of a Bernoulli event at Δt_i, $c_i \leq m$

An equivalent identity is given by the general low of large numbers which states that:

$$\lim_{m\to\infty} \frac{c_i}{m} = p_i \ (eq2)$$

Making ΔT $[t_0, t_n)$ the time frame of analysis, e.g., the p/s lifecycle.

With $|\Delta t| = \frac{\Delta T}{n}$

We can say that the interruption rate of a p/s in any time interval Δt in $[t_0, t_n)$ is given by:

$$\lim_{m\to\infty} \left(\frac{c_i}{m * \Delta t}\right) = \frac{p_i}{\Delta t} = \lambda_i \ (eq3)$$

With X: the random variable.

$X = 1$: sucessful event: one or more interruptions.

Now, if we divide the time interval ΔT in two, we will have that the probability of an event, resulting from any two independent and non-exclusive events, is given by:

$$P(X=1)_{t_0 \leq t < t_n} = p_1 + p_2 - p_1 p_2 =$$

$$\lambda_1 \frac{\Delta T}{2} + \lambda_2 \frac{\Delta T}{2} - \left(\frac{\Delta T}{2}\right)^2 \lambda_1 \lambda_2$$

Whereby, for n ≥ 2, the probability of one or more interruptions, i.e., ($P(X = 1)$), taking place along $t_0 \leq t \leq t_n$, can be expressed as follows:

$$P(X=1)_{t_0 \leq t < t_n}$$

$$= \sum_{i=1}^{n} p_i - \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} p_i p_j + \sum_{i=1}^{n-2} \sum_{j=i+1}^{n-1} \sum_{k=j+1}^{n} p_i p_j p_k - \cdots + (-1)^n \sum_{i=1}^{2} \sum_{j=i+1}^{3} \sum_{k=j+1}^{4} \cdots \sum_{r=n-1}^{n} p_i p_j p_k \dots p_r + (-1)^{n-1} \prod_{i=1}^{n} p_i \ (eq4)$$

With:

$[t_0, t_n)$, any interval in $[0, \infty)$

$i \neq j \neq k \neq \dots r$ $\forall i, j, k, \dots, r = 1, 2, 3, \dots, n$.

pi, pj, pk, ..., pr: the probability of the event in the time interval i, j, k, ..., r respectively. i.e., pi: Probability of the event in $[t_0, t_n)$.

n: a possitive interger.

Note:

The first big operation of eq4 has $\binom{n}{1}$ terms, i.e., n terms; the second, $\binom{n}{2}$ terms; the third, $\binom{n}{3}$ terms, and son on. The big operation before the last one has $\binom{n}{n-1}$ terms, i.e., n terms.

The whole equation has a symmetric pattern, alternating between positive and negative. For instance, if we divide the subspace of events in ten intervals Δt, the big operations that compound the eq4 will have the following number of terms plus the last term; corresponding to $(-1)^{n-1} \prod_{i=1}^{10} p_i$:

$$\binom{10}{1} = 10; \binom{10}{2} = 45; \binom{10}{3} = 120; \binom{10}{4} 210; \binom{10}{5} = 252;$$

$$\binom{10}{6} 210; \binom{10}{7} = 120; \binom{10}{8} = 45; \binom{10}{9} = 10$$

When n is even, it appears a mid-term (252 in the example above). When n is odd, the

1. The count of systems interrupted disregards the number of interruptions per system in the given interval

pattern is a perfect mirror. The total number of terms of the eq is given by:

$$2\left[n+\binom{n}{2}+\binom{n}{3}+\cdots+\binom{n}{\frac{n-2}{2}}\right]+\binom{n}{\frac{n}{2}}, if\ n\ is\ even$$

$$2\left[n+\binom{n}{2}+\binom{n}{3}+\cdots+\binom{n}{\frac{n-1}{2}}\right], if\ n\ is\ odd$$

END OF NOTE

Now, to calculate the probability of an event (one or more interruptions or failures) in any variable rate processes using the eq5, we must consider that if n increases, $\frac{\Delta T}{n}$ decreases. Therefore, if ΔT is divided in n intervals with n very large, such that $\frac{\Delta T}{n} \leq 10^{-1}$, eq4 converges on the second big operation and can be written as follows:

$$P(X=1)_{t_0 \leq t < t_n} = \sum_{i=1}^{n} p_i - \sum_{i=1}^{n-1}\sum_{j=i+1}^{n} p_i p_j$$
$$= \frac{\Delta T}{n}\sum_{i=1}^{n}\lambda_i - \left(\frac{\Delta T}{n}\right)^2 \sum_{i=1}^{n-1}\sum_{j=i+1}^{n}\lambda_i\lambda_j \quad (eq5)$$

Making $\frac{\Delta T}{n} = \Delta t$, from eq3 we have:

$$\sum_{i=1}^{n-1}\sum_{j=i+1}^{n} p_i p_j = (\Delta t)^2 \sum_{i=1}^{n-1}\sum_{j=i+1}^{n}\lambda_i\lambda_j \quad (eq6)$$

But,

$$\sum_{i=1}^{n-1}\sum_{j=i+1}^{n}\lambda_i\lambda_j = \frac{1}{2}\left[\sum_{i=1}^{n}\sum_{j=1}^{n}\lambda_i\lambda_j - \sum_{i=1}^{n}\lambda_i^2\right] \quad (eq7)$$

Then, eq5 becomes:

$$P(X=1)_{t_0 \leq t < t_n} = \sum_{i=1}^{n} p_i - \frac{(\Delta t)^2}{2}\left[\sum_{i=1}^{n}\sum_{j=1}^{n}\lambda_i\lambda_j - \sum_{i=1}^{n}\lambda_i^2\right] \quad (eq8)$$

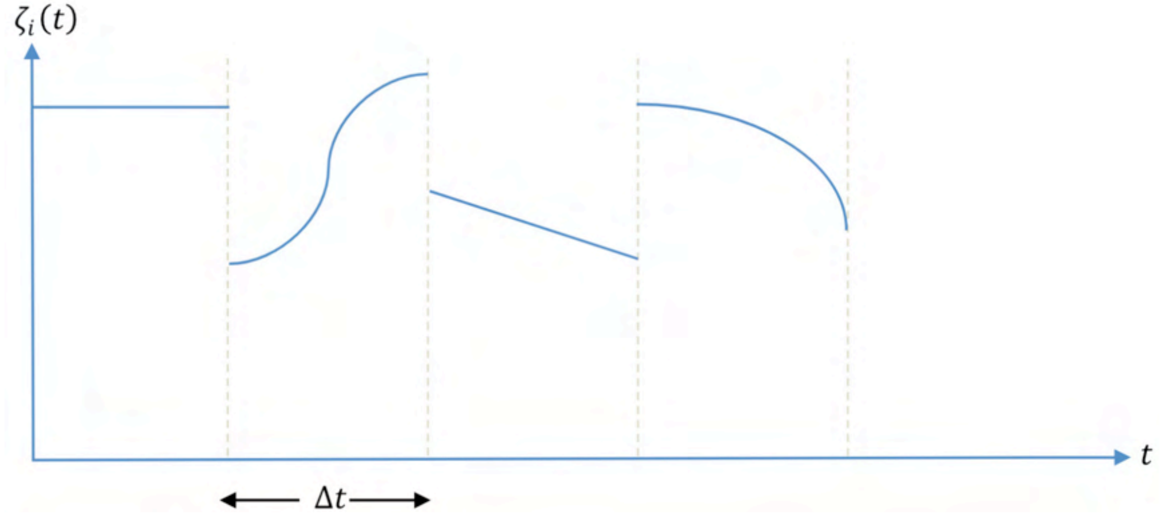
Let us suppose that for $t_0 \leq t < t_n$, there is a continuous function ζ_i (t) that represents the process/system (p/s) interruption rate in every Δt such that:

$$\zeta_i(t) = \begin{cases} \zeta_i(t) > 0, & t_{i-1} \leq t < t_i \\ 0, & otherwise \end{cases}$$

And $\zeta_i(u = v) = \lambda_i$, $t_{i-1} \leq v < t_i$, *with* $i = 1, 2, 3, \ldots, n$

According to the Mean Value Theorem we can find a value λ_i such that:

$$\lambda_i = \frac{\int_{t_{i-1}}^{t_i}\zeta_i(u)du}{\Delta t} \quad (eq9)$$



Graph 1 – Random interruption rate function $\zeta_i(t)$ of any p/s

Journal of Engineering Research ISSN 2764-1317 DOI 10.22533/at.ed.317212210014

or, what is the same: $\Delta t * \lambda_i = p_i =$

$$\int_{t_{i-1}}^{t_i} \zeta_i(u)du \ (eq10)$$

Now, with:

$$S = \sum_{i=1}^{n} p_i = \int_{t_0}^{t_1} \zeta_1(u)du + \int_{t_1}^{t_2} \zeta_2(u)du + \cdots + \int_{t_{n-1}}^{t_n} \zeta_n(u)du \ (eq11)$$

Eq8 becomes:

$$P(x=1)_{t_0 \le t < t_n} = S - \frac{1}{2}\left[S^2 - \sum_{i=1}^{n}\left(\int_{t_{i-1}}^{t_i} \zeta_i(u)du\right)^2\right] \ (eq12)$$

Let us remember that $\zeta_i(u)$ is a continuous function by parts, i.e., continuous in every interval $t_{i-1} \le u < {}_i t$, such that $\frac{d\Lambda_i(u)}{du} = \zeta_i(u)$.

Now, suppose that instead of having a continuous function by parts, there exists a continuous function $\lambda(u)$ in $t_0 \le u < t_n$, such that:

$$\lambda(u) = \sum_{i=1}^{n} \zeta_i \ (eq13)$$

Besides:

$$\frac{d\Lambda(u)}{du} = \lambda(u)$$

Thus:

$$\sum_{i=1}^{n}\left(\int_{t_{i-1}}^{t_i} \zeta_i(u)du\right)^2 = \sum_{i=1}^{n}\left(\int_{t_{i-1}}^{t_i} \lambda(u)du\right)^2$$

Therefore:

$$\sum_{i=1}^{n}\left(\int_{t_{i-1}}^{t_i} \lambda(u)du\right)^2$$

$$= [\Lambda(t_1)-\Lambda(t_0)]^2 + [\Lambda(t_2)-\Lambda(t_1)]^2 + [\Lambda(t_3)-\Lambda(t_2)]^2 + \cdots + [\Lambda(t_n)-\Lambda(t_{n-1})]^2$$

$$= \Lambda^2(t_0) + \Lambda^2(t_n) + 2[\Lambda^2(t_1) + \Lambda^2(t_2) + \Lambda^2(t_3) + \cdots + \Lambda^2(t_{n-1})] - 2[\Lambda(t_1)\Lambda(t_0) + \Lambda(t_2)\Lambda(t_1) + \cdots + \Lambda(t_{n-1})\Lambda(t_{n-2}) + \Lambda(t_n)\Lambda(t_{n-1})] \ (eq15)$$

As Δt is very small (Δt<<1), we have that: $\Lambda(t_{i+1})\Lambda(t_i) \approx \Lambda^2(t_i)$, therefore the eq14 becomes:

$$(\Delta t)^2 \sum_{i=1}^{n} \lambda_i^{\,2} = \sum_{i=1}^{n}\left(\int_{t_{i-1}}^{t_i} \lambda(u)du\right)^2 = \Lambda^2(t_n) - \Lambda^2(t_0) \ (eq15)$$

From eq15 into eq12:

$$P(x=1)_{t_0 \le t < t_n} = S - \frac{1}{2}[S^2 - S(S + 2\Lambda(t_0))]$$

Now, factorizing we have:

$$P(x=1)_{t_0 \le t < t_n} = S(1 + \Lambda(t_0)) \ (eq16)$$

From eq11 we know that K is the subspace of all the events (interruptions or failures) of a p/s. K will depend on the p/s under analysis, for instance, if we take the subspace of all events along the p/s lifecycle, i.e., $t_0 = 0$, and $t_n \to \infty$, and if the limit exists, we have:

$$\lim_{t_n \to \infty} S = \lim_{t_n \to \infty} \int_0^{t_n} \lambda(t)dt = K$$

As we stated, an Evolutionary Process will have at least one interruption event along its lifecycle. Thus, from eq16:

$$\lim_{t_n \to \infty} P(x=1)_{0 \le t < t_n} = K(1 + \Lambda(0)) = 1 \Leftrightarrow (1 + \Lambda(0)) = \frac{1}{K} \ (eq17)$$

Factorizing, we have:

$$\Lambda(t_n) - \Lambda(0) = K$$

Then, with eq17 into eq16, without losing generality, the probability of interruption of an EP in any time interval $t: [t_a, t_b)$ is given by:

$$P(x = 1)_{t_a \le t < t_b} = \frac{S}{K}$$

Note:

In Conditional Probability of Failure, the subspace of events K, goes from t_a to $t \to \infty$, which makes a function of time.

Therefore, the general case is given by:

$$P(X = 1)_{t_a \le t < t_b} = \frac{1}{K(t)} \int_{t_a}^{t_b} \lambda(t)dt \ (eq18)$$

CONCLUSION

The probability of interruption of an EP in an interval is given by eq18.

The probability density function of an EP with continuous interruption rate function $\lambda(t)$, along EP lifecycle, is given by:

$$f(t) = \frac{1}{K(t)} \lambda(t)$$

Q.E.D.

REFERENCES

1. R. Murray., J. Schiller., R. Srinivasan "Special Probability Distributions" in Probability and Statistics, 4th, USA: Mc Graw Hill, 2013, pp. 108–109.

2. J. A. Aristizábal, "Annex-4. Variable-Rate processes Probability Distribution (VRPD) applied to Conditional Probability of Failure (CPF)" in Probability of Interruption and Business Impact, 1ed. Menlo Park, CA, USA: IBUKKU, 2020, pp. 133–138.

3. L. Leithold, "Integral Definida e Integración" in El Cálculo, 7th ed. MX: Oxford University Press, 1998.

4. A. K. Verma., S. Ajit., D.R. Karanki. "Basic Reliability Mathematics" in Safety Engineering, England: SpringerLink, 2016, pp. 28–36.