Christopher Lorenz
UD76627SEN85819

STATISTICS

Better Engineering through Statistics

ATLANTIC INTERNATIONAL UNIVERSITY
HONOLULU, HAWAII
SUMMER 2010

DATE 30th /MARCH /2022

TABLE OF CONTENTS

INTRODUCTION-

> "Statistics is, or should be, about scientific investigation and how to do it better, but many statisticians believe it is a branch of mathematics. Now I agree that the physicist, the chemist, the engineer, and the statistician can never know too much mathematics, but their objectives should be better physics, better chemistry, better engineering, and in the case of statistics, better scientific investigation. Whether in any given study this implies mathematics is incidental." George E. P. Box

Although most scientists, engineers and physicists considered nature, at its deepest level, was believed to be based on determinism before quantum theory. After the introduction of quantum theory, most now consider nature, at its deepest level to be random and unpredictable. (Buchanan).

From an electrical engineering viewpoint, nature being random and unpredictable at its fundamental level, causes problematic conditions when trying to formulate engineering solutions. To overcome these problematic conditions engineers, use statistical analysis to help develop their solutions.

Engineers use statistical analysis to help them collect, investigate, and present data to show patterns and trends. This helps engineers to develop models that are used to make decisions, test, and create overall solutions to the engineering problems that they are working on solving. It's easy to capture a quantity of data, but without knowing how the data should be examined and used provides an engineer with no viable solution. Therefore, according to George E. P. Box, the engineer is not providing what he would determine as the concept of better engineering.

By studying the concepts of The Engineering Method and Statistical Thinking, Probability and Probability Models, and the effect that nature's randomness has will provide a better idea on how engineers can provide better engineering solutions.

BETTER ENGINEERING –

Since Engineers are solve problems by the application of scientific principles, solutions are developed through the application of statistical analysis predominately from the viewpoint of mathematics. However, does providing engineering solutions through the window of mathematics provide what George E. P. Box states as "Better Engineering"? To answer that question, the engineer should have a basic knowledge of what better engineering means. The definition of the term better means to improve or surpass. Therefore, when applied to the concepts of engineering, better engineering should be the use engineering practices that surpass or improve on existing engineering practices used to develop an engineer's solution to an engineering problem.

In today's world, there are many different sources that provide information on the practices that and engineer should invent, design, and build to. For electrical engineers, this includes NFPA engineering standards, IEEE engineering standards, and NEC guidelines to name a few. Electrical Engineers also follow a strict code of conduct. These all provide and engineer with a set of engineering practices. With the addition of statistical analysis, electrical engineers can provide better engineering solutions to electrical engineering problems.

According to many of the references about statistics reviewed, they all state the importance that statistics has on engineering. In the journal article *Application of Statistics in Engineering Technology Programs*, the journal writers state that statistics is an important tool for robustness analysis, measurement system error analysis, test data analysis, probabilistic risk assessment, and many other fields in the engineering world (Zahn) According to the International Statistical Engineering Association, Statistical Engineering is a discipline for solving complex problems using data and data analysis. (Association)

NATURE'S RANDOMNESS-

Since the addition of statistical analysis provides an engineer with a tool to perform their work at a better engineering level, applying this tool to Nature's randomness will provide a way for the engineer to develop engineering solutions to problems affected by this concept.

Before the introduction of quantum mechanics, Nature to most science and engineering was considered through the view of Determinism. Determinism is a belief that all events and actions are already determined and are unaffected by the concept of free will. In the case of the Laws of Nature, Causal Determinism is the belief that every current action or event is the result of the actions or events that preceding along with the effects caused by the Laws of Nature. This is what is known as Classical Mechanics.

When science and engineering introduce the concepts of Quantum Mechanics, it was noticed that by most scientists, engineers, and physicist that Nature isn't deterministic. Instead, Nature is basically random at its fundamental levels (Buchanan). In either case, classical or quantum mechanics are both based on the statistical probability of Nature (Schmitz).

So, if Nature is random at its fundamental level, what does that mean, what is its effect on electrical engineering, and how does statistical analysis overcome this? For electrical engineers, developed methodologies captures Nature's randomness as a variable which can then be applied to known and established mathematical equations to help provide solutions to problems. This random variable can be discrete in nature, continuous in nature, or a mixture. All three of these variables can be solved for using the cumulative distribution function.

The cumulative function is a set of events of the form $\{X = b\}$ (Leon-Garcia) where X is a function from the sample space S to R with the property that the set $A_b = \{ \zeta : X(\zeta) \leq X_b\}$ (Leon-Garcia). What the formula states in general is that every set A_b should have a well defined probability in the random experiment. Simply stated the function is a convenient way of specifying the probability of all semi-infinite intervals of the real line of the form (-∞,b].

As stated above Nature's randomness can be one of three different types of variables. Each type of variable will produce a specific type of function output. For Discrete random variables, the continuous discrete function is the sum of probabilities of the outcomes less than x. This type of variable produces a right-continuous staircase function. A continuous random variable is a variable whose function is continuous everywhere. It is smooth and can usually be expressed as an integral of some nonnegative function. The last type of variable is the mixed type. This variable can jump on a countable set of points or increase continuously over at least one interval of values of x. This type of variable is normally shown as a two-step process. Each type of variable outputs can be seen in Fig. 4.1, 4.2 and 4.3 (Leon-Garcia).

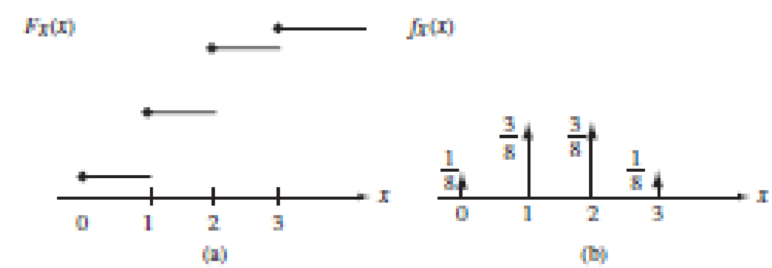

FIGURE 4.1
cdf (a) and pdf (b) of a discrete random variable.

Figure 1:Example Output of a Discrete Random Variable

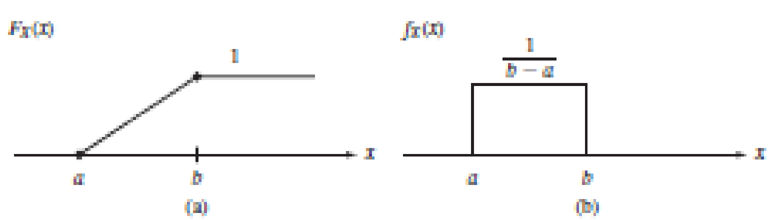

FIGURE 4.2
cdf (a) and pdf (b) of a continuous random variable.

Figure 2: Example Output of a Continuous Random Variable

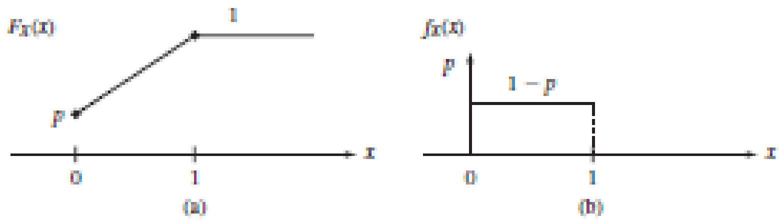

FIGURE 4.3
cdf (a) and pdf (b) of a random variable of mixed type.

Figure 3: Example of Output of a Mixed-Type Random Variable

Understanding that Nature is at its fundamental level random, electrical engineers can therefore create a mathematical variable based on this randomness. In the next section, I will discuss what does probability mean, what are probability models, and how the random variable is used to help solve these models.

PROBABILITY AND PROBABILITY MODELS-

When it comes to statistics, to solve a problem that is most likely complex in nature will require choices to be made. Probability simply stated is the study of randomness. From the view of electrical engineering, randomness is seen as noise on an electrical signal. Probability Theory simply stated is the mathematical study of random phenomena (Handel). Probability Theory consists of several parts. These parts include a Sample Space, Events, Probability Measure, and Probabilistic Modelling to name a few.

The Sample Space is the first ingredient of any probability model. This space simply stated is the specification of all possible outcomes of a random experiment (Handel). Usually notated with the symbol Ω. The next ingredient is known as an Event. Events are statements that can be used to determine whether they are true or false when compared to the Sample Space (Handel). Simply stated an Event is a subset A of the Sample Set Ω (Handel). From a Mathematical perspective and Event is the intersection of subset A and Sample Space B where A and B occur simultaneously.

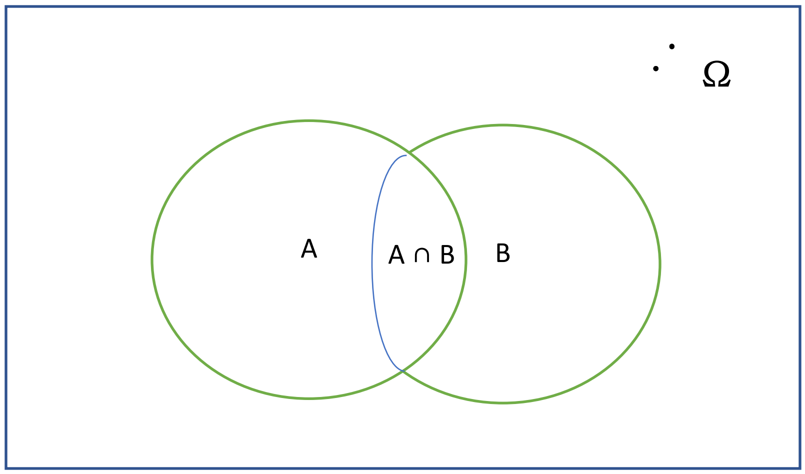

Fig 4: Representation of A ∩ B

The Probability Measure is an assignment of a number P(A) to every Event where $0 \leq P(A) \leq 1$ and $P(\Omega) = 1$, and $A \cap B = 0$ are satisfied (Handel). Finally, there is Probabilistic Modelling. Simply stated Probabilistic Modelling is a mathematical representation of the Sample Space, Events, and the Probability Measure (Handel).

Models can play an important role in analysis of engineering problems and usually come in two forms. Models can be either mechanistic or empirical. Mechanistic models are models built on underlying knowledge of mechanisms relating to variables. An example of a mechanistic model is the following formula: I = E/R or current = voltage/resistance. Empirical models are models built on engineering or scientific knowledge of phenomena which is not directly developed from theoretical of first hand knowledge of an underlying mechanism. An example of this is $M_n = \beta_0 + \beta_1 V + \beta_2 C + \beta_3 T + \varepsilon$ where the β's are unknown variables.

Knowing the randomness found in Nature, and the use of Probability Models, how does an electrical engineer figure out the solution to his/her problem. The solution is developed through a multi-step process called The Engineering Method.

1. Develop a clear and concise description of the problem.
2. Identify, at least tentatively, the important factors that affect this problem or that may play a role in its solution.
3. Propose a model for the problem, using scientific or engineering knowledge of the phenomenon being studied. State any limitations or assumptions of the model.
4. Conduct appropriate experiments and collect data to test or validate the tentative model or conclusions made in steps 2 and 3.
5. Refine the model on the basis of the observed data.
6. Manipulate the model to assist in developing a solution to the problem.
7. Conduct an appropriate experiment to confirm that the proposed solution to the problem is both effective and efficient.
8. Draw conclusions or make recommendations based on the problem solution.

Figure 5: The Steps of The Engineering Method.

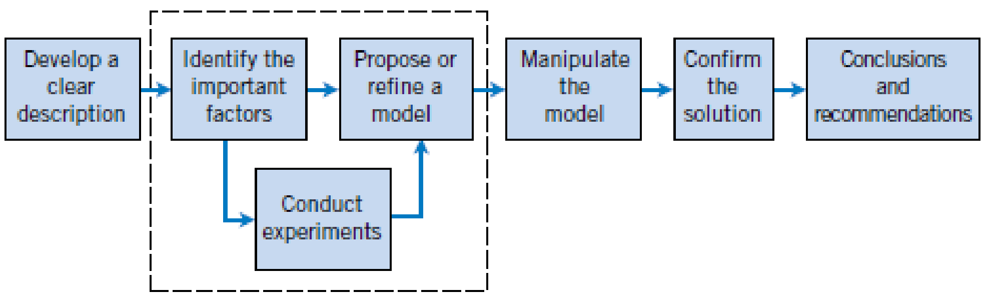

Figure 6: The Engineering Method Process

Statistics is the collection, presentation, analysis, and use of data to make decisions, solve problems, and design products and processes (Douglas C Montgomery). Statistical methods are important for engineers because many problems are complex in nature and statistical methods help describe and understand variability (Douglas C Montgomery). Statistical thinking provides a useful way to incorporate variability into the decision making process (Douglas C Montgomery). Since variability can be found in all aspects of life, it becomes very important for engineers to represent this variability in mathematical terms so they can derive a solution. A good way to represent variability mathematically is to set the variable to the letter X and substitute it into the following formula $X = \mu + \varepsilon$, where μ is a constant component and ε is the random component.

One portion of the Engineering Method is the collection of data. The collection of data can be accomplished in one of three different methods. The first method is thru a retrospective study of historical data. A retrospective study requires engineers to review previous archived data over a period to help in finding relationships in the data. This collection of data may be rather large and could contain a lot of information not relevant to the problem at hand. Along with these issues there may also be times where the data is missing or there may be errors in the information. Using this method may present with problems. These problems include the failure to see relationships between data, a failure to construct an approximate representation, a failure to assess real impact on variables, and a difficulty in separating data information.

The next method to obtain data is thru an observational study of the information. Engineers using this method will observe the process and record any data information that is of interest to the problem being solved. These studies are usually done over short periods of time, and variables not routinely measured can be included in the solution. While this method does help in overcoming the problems of seeing relationships, and developing approximate representations in the data, it still may not solve for the other types of problems that can be presented.

The final method to obtain data is thru the performance of designed experimentation. This method allows the engineer to make deliberate or purposeful changes to variables and observe their effect on the output data. From there the engineer can make inferences, change the experiment, and reobserve the outcome. This method can play a vital role in engineering design, development and ultimately the improvement of different processes. Another benefit of this type of method is the reduction of lead time which is very important to the bottom line of most companies. Simply stated, designed experimentation is a powerful approach to studying and solving complex problems.

Another useful method for working with data is thru the observation of data of a specific period. This allows the engineer to develop data plots based on the time observed. Doing this will allow the engineer to see trends in the data which can help in making inferences and develop better solutions to problems that are being solved.

From the viewpoint of electrical engineering, statistical analysis doesn’t play a big role when it comes to design or basic circuit analysis. It is however a large help when it comes to trying to explain operation of electrical systems, trying to determine correlations between similar products that have been fabricated using the same processes. For example, if a customer purchases a batch of circuit boards for a particular product line. Statistical analysis is normally used in determining the variation in signal to noise ratio between each circuit board. An important area where statistical analysis is needed is in the field of IC research and development. Another field where statistical analysis is of great importance is in the field of signal processing and the effects that noise has on it (Keim).

Noise is unwanted modifications to a signal during capture, storage, transmission, processing, or conversion. It is random in nature and usually carries little or no useful information. Robert Keim discusses noise in a couple of articles that he wrote for All About Circuits. In those articles he discusses electrical noise generated in image processing using charge-coupled device (CCD) image sensors. In those articles the two most important sources of noise in the image sensors are coming from what he called Photon Noise, and Dark Noise.

Free electrons that are generated thermally and collected are known as dark current (Keim). By taking the square root of the combination of this dark current with a time of integration determines the level of dark noise. This formula is shown in Fig. 7. describes the root mean square (RMS) value of dark noise.

$$\text{dark noise} = \sqrt{(\text{dark current}) \times (\text{integration time})}$$

Figure 7: Dark Noise Formula

According to Keim, the RMS value can be determined since the behavior of dark current is discrete in nature and it follows a Poisson relationship. This relationship expresses the probability of a given number of events occurring in a fixed interval of time or space. To meet this fixed interval, the events must occur at a known constant mean rate and they must be independent of time. Alternately known as thermal noise, dark noise usually will produce an electrical offset of the CCD signal. Dark noise compensation is accomplished by measuring the frame to frame rate of the sensor's dark current and since it is thermally based the noise level can be reduced by regulating the temperature of the system. Using thermoelectric cooling or liquid nitrogen cooling can bring this noise level to a negligible level (Keim).

Keim continues his investigation into electrical noise associated with CCD image sensors and states that two other major contributors to circuit noise is Photon Noise and Read Noise. Photon noise, like dark noise, is discrete in nature and it follows Poisson relationships. This type of noise is based the inherent variation in incident photon flux at the detector. The second major source of noise on the image sensors is known as Read Noise. Read noise is the amount of noise generated by electrons as the charge present in the pixels of the camera. Read noise usually will determine the contrast resolution of the camera. The lower the level of read noise, the smaller the contrast difference. The overall result is the lower the read noise the more sensitive the sensor is.

CONCLUSION-

An article published in the American Journal of Engineering Education in 2010, describes the importance of using statistics in engineering applications. According to the article, statistics is a critical tool for robustness analysis, measurement system error analysis, test data analysis, probabilistic risk assessment and many other applications in the engineering world (Zahn). Knowing that statistics supplies all these important applications, the article states that traditionally statistics is not extensively used in undergraduate programs. This is surprising since the American Statistical Association stated in 1980 that the need for engineers with experience and knowledge in statistics necessary for industry (Zahn).

The use of statistical analysis through a step by step process known as The Engineering Method helps engineers understand the importance of using probability and probability models to incorporate the randomness of variables seen in Nature to develop the solutions to many of the problems faced by engineers today. For today's engineers the use of these concepts is what George E. P. Box considers as "better engineering".

Bibliography

Association, International Statistical Engineering. *What is Statistical Engineering*. n.d. Webpage. 19 04 2022.

Buchanan, Mark. "Quantum randomess may not be random." 18 March 2008. *NewScientist.* Web Page. 04 April 2022. <https://www.newscientist.com/article/mg19726485-700-quantum-randomness-may-not-be-random/>.

Douglas C Montgomery, George C Runger. *Applied Statistics and Probability for Engineers 5th edition*. Wiley (3/23/2010), 2011.

Handel, Ramon van. *Probability and Random Processes, ORF309/MAT 380 Lecture Notes Princeton Univeristy*. 2016.

Keim, Robert. *Dark Noise in CCD Image Sensors*. 21 May 2020. <https://www.allaboutcircuits.com/technical-articles/dark-noise-ccd-image-sensors/>.

—. *Descriptive Statistics in Electrical Engineering*. 22 June 2020. <https://www.allaboutcircuits.com/technical-articles/descriptive-statistics-in-electrical-engineering/>.

—. *Introduction to Statistical Analysis in Electrical Engineering*. 09 June 2020. <https://www.allaboutcircuits.com/technical-articles/introduction-to-statistical-analysis-in-electrical-engineering/#:~:text=Statistical%20techniques%20help%20us%20to,signals%20that%20we%20call%20noise.>.

—. *Photon Noise, Read Noise, and Reset Noise in CCD Image Sensors*. 22 May 2020. <https://www.allaboutcircuits.com/technical-articles/photon-noise-read-noise-and-reset-noise-in-ccd-image-sensors/>.

Leon-Garcia, Alberto. *Probability, Statistics, and Random Processes for Electrical Engineering, 3rd Edition*. Upper Saddle River: Pearson Prentice Hall, 2008. pdf.

Schmitz, Kenneth S. "Quantum Statistical Mechanics." *Physical Chemistry* (2017).

Zahn, W., Fink, R., Fang, A. "Application of Statistics in Engineering Technology Programs." *American Journal of Engineering Education* (2010): 65-78. Document.