

Chapter 2

Essentials of Probability

The study of probability began in the 17th Century as a result of a desire to determine the outcome of games of chance. By definition, a probability is a real number bounded between $[0,1]$. Often, probabilities are expressed as percentages, or positive outcomes in a finite number of trials, e.g. your odds of winning are “1 in 3.” More formally:

Probability

Probability is defined as the mathematical modeling of the phenomenon of *chance* or *randomness*. A probability is always stated in a fraction: a real number from $[0.0,1.0]$.

A complete study in probability begins with a discussion of set theory, elements, subsets, Venn Diagrams, set based proofs, association laws, etc. These are important topics, but are beyond the scope of our discussion here, where we provide a brief discussion of basic probability applications leading up to the likelihood of radioactive decay and probability distributions as they pertain to nuclear processes. It is from this perspective that we begin.

Counting Rules

There are two key rules of “event counting” that enable one to attribute specific events among a possible suite of events, and then determine a likelihood that the specific events occur. The two key rules are the “sum rule” and the “product rule.”

Sum rule of counting

Given: Event A occurs in m ways,
Event B occurs in n ways....

Then Event A **or** B can occur in $(m+n)$ ways.

Product rule of counting

Given: Event A occurs in m ways, and independently,
Event B occurs in n ways....

Then combinations of events A **and** B can occur in $(m \cdot n)$ ways.

Example

Suppose a student at the university has the following choices (assuming no prerequisites):

3 History courses, 4 English courses, 2 Science courses

If a student must choose one of each, the number of ways in which he/she may do this is:

$$(3)(4)(2) = 24$$

If a student must choose only one course, the number of choices he/she has is:

$$(3+4+2) = 9$$

The concept of an individual outcome for a number of trials is frequently encountered. In this context, consider the following:

Tossing a coin once.

Choosing a card from a deck.

Picking winning lotto numbers.

We note that while each individual event among those listed above has a specific outcome not explicitly known, an overall result will be known for “many trials...” Consider x and y as events.

Consider 2 events (or outcomes) x, y in N trials. After each trial $1, 2, \dots, N$, there are *four possibilities* for the outcome:

of Times

x occurred but not y ...	n_1
y occurred but not x ...	n_2
x and y both occurred...	n_3
neither x nor y occurred...	n_4

The number of recorded events are: $(n_1 + n_2 + n_3 + n_4) = N$

Event Probability Metrics

We can define the probability for a number of events, noting that each is a probability that results in a number in the interval $[0,1]$:

Summary of Probability Rules	Probability Set
$\frac{n_1 + n_3}{N} \leftrightarrow P(x) \equiv$ probability x occurred	A
$\frac{n_2 + n_3}{N} \leftrightarrow P(y) \equiv$ probability y occurred	B
$\frac{n_1 + n_2 + n_3}{N} \leftrightarrow P(x + y) \equiv$ prob <u>either</u> x or y occurred	$A \cup B$
$\frac{n_3}{N} \leftrightarrow P(xy) \equiv$ probability <u>both</u> x and y occurred	$A \cap B$

Conditional Probability Rules	Probability Set
-------------------------------	-----------------

$\frac{n_3}{n_2 + n_3} \leftrightarrow P(x y) \equiv \text{probability of event } x, \text{ given}$ <p style="text-align: center;">that y has occurred</p>	$A \text{ given } B$
---	----------------------

$\frac{n_3}{n_1 + n_3} \leftrightarrow P(y x) \equiv \text{probability of event } y, \text{ given}$ <p style="text-align: center;">that x has occurred</p>	$B \text{ given } A$
---	----------------------

Mutual Exclusivity and “Complement”

“ $A \cap B = \phi$ ” is known as “disjoint” or “mutually exclusive”, where A and B cannot occur simultaneously.

The “Complement” involves the probability that the event does not occur, so that, for an event x with a probability $P(x)$, then:

$$P(x^c) = 1 - P(x)$$

Multiplication and Addition Laws

One can often use keywords “either” or “both/all” to distinguish what course of action to follow when determining probabilities. With the probability rules defined above, by inspection, we can write the multiplication and addition laws. Given two events x and y , the probability that *both events will occur* can be gleaned from the multiplication law, where probability sets intersect:

Multiplication Law of probability

$$P(xy) = P(x) P(y|x) = P(y)P(x|y)$$

$A \cap B$ Implies “both” or “all”

If we consider that *either of two events can occur*, the addition law applies, but requires that we account for the exclusion of “both” or “all” as determined from the multiplication law:

Addition Law of probability

$$P(x + y) = P(x) + P(y) - P(xy)$$

$$A \cup B \quad \text{Implies “either”}$$

A number of useful Examples come from a deck of cards. In a standard deck of 52 cards, there are 12 face cards {Jacks – J, Queens – Q, Kings – K} spread among the 4 suits. Suits are {Diamonds – D, Hearts – H, Clubs – C, Spades – S}. There are then 13 cards in each suit, spanning “2” through “Ace.”

Example

What is the probability of drawing one card that is the ace of spades? It is given there are 52 cards in the deck. Also $x =$ drawing ace of spades $= y$, then each deck has only 1 ace of spades.

$$P(x) = P(y) = \frac{1}{52} \approx 0.019$$

Example

What is the probability of choosing the ace of spades twice in two independent draws?

Note that choosing the “ace of spades twice” in two independent draws is as follows: by multiplication law:

$P(xy) = P(x) P(y|x)$ probability that the 1st draw is the ace of spades given that second draw is the ace of spades

$$= \frac{1}{52} \quad (\text{each draw is independent})$$

$$P(y|x) \rightarrow P(y) \quad \therefore P(xy) = \left(\frac{1}{52}\right)\left(\frac{1}{52}\right) \cong 0.00037$$

\therefore it is unlikely two consecutive random draws will yield the ace of spades twice in a row.

The “Complement” of drawing the Ace of Spades 2 times in a row is then

$$P((xy)^c) = 1 - 0.00037 = 0.99963$$

Example: For a standard deck, compute the probability that a

event A = a Heart and

event B = a Face card will be drawn.

Note: “and” implies the *Multiplication Rule*:

$$P(A) \equiv \text{probability we will draw a Heart} = \frac{13}{52} = \frac{1}{4}$$

$$P(B) \equiv \text{probability we will draw a face card} = \frac{12}{52} = \frac{3}{13}$$

$P(AB) = P(A)P(B)$ is also written

$$\therefore P(A \cap B) = P(A)P(B) = \frac{3}{52} = 0.0577$$

Example

Compute the probability that either a *heart* or a *face card* will be drawn from a standard deck.

Note: “Either” implies the *Addition Rule*

$$P(A) \equiv \text{probability we will draw a heart} = \frac{13}{52}$$

$$P(B) \equiv \text{probability we will draw a face card} = \frac{12}{52}$$

$P(A + B) = P(A) + P(B) - P(AB)$ is also written

$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$

$$\therefore P(A \cup B) = \frac{13}{52} + \frac{12}{52} - \left(\frac{13}{52}\right)\left(\frac{12}{52}\right)$$

where $P(A \cap B) \leftrightarrow P(AB) = P(A)P(B|A)$, each is independent

$$P(A \cup B) = 0.42307$$

The Law of Large Numbers

Assume an experiment is repeated N times and “ a ” out of “ n ” times, the result was of type “ x ”.

$P(x) = \lim_{N \rightarrow \infty} \frac{a}{N} \Rightarrow$ in practical limit, an infinite number of trials is not possible

$P(x) = \frac{a}{N}$ where a/N is the frequency of occurrence of x in the first N trials

Example: Consider a single coin toss

100 times \rightarrow 47 heads
53 tails

$$\text{Then } P_{\text{Heads}} = \frac{47}{100} = 0.47$$

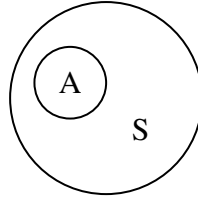
$$P_{\text{Tails}} = \frac{53}{100} = 0.53$$

As $N \rightarrow \infty$, the *law of large numbers* for N says that probabilities will yield their true theoretical outcomes... $P_{\text{Heads}} = P_{\text{Tails}} = 0.50$.

Graphical Definition of Probability

Consider the following graphical definition of probability:

$$P(A) \equiv \frac{\# \text{ Pts in } A}{\# \text{ Pts in } S}$$



$$P(S) = 1$$

$$\frac{f(A)}{n} \equiv \frac{\# \text{ times } A \text{ occurs}}{\# \text{ trials}} \quad \text{where } 0 \leq f(A) \leq 1$$

Note that here a “universe” is defined by the region “S,” where the probability of a point being in “S” is 100% (with a $P((S)^c)=0$). The region “A” is contained inside “S”.

One can consider the number of points sampled randomly in “S,” and with the Law of Large Numbers, ultimately $P(A)$ is the ratio of the area of “A” divided by the area of “S.”

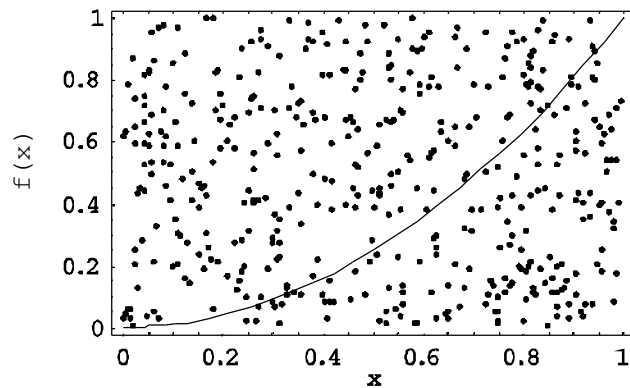
The Monte Carlo Method

The Monte Carlo method was named as such since it involves “playing the probability game”... sampling probabilities for a large number of trials (thus applying the Law of Large Numbers) to statistically arrive at a correct solution. The Monte Carlo method is simple in principal, but becomes fairly tedious without the use of a computer.

A simple application of the Monte Carlo Method can be demonstrated by *Numerical Integration*, where a simple application of Monte Carlo is to “throw darts” in a closed area where a function is plotted, and determine the integral (area under the function) by using a simple ratio of the “darts” striking under the curve to the “darts” hitting the closed area.

Example

Consider estimating $\int_0^1 x^2 dx$ over a 1 x 1 square area. We “throw” random “darts” at the page within the sampling area over which the function is projected:



Representation of Monte Carlo Sampling of an Area.

and 141/450 dots fell under the curve, estimating the integral to be $141/450 * (\text{Area}) = 141/450 (1)(1) = 0.31333$.

Where the actual answer is

$$\int_0^1 x^2 dx = 0.33333$$

... which can be better approximated by throwing more darts!

This points out that the *Monte Carlo method* is ALWAYS subject to *some statistical sampling error*, which in this case, is 6%.

The Monte Carlo method is used a great deal in nuclear engineering due to the complex issues involved in solving radiation transport problems, and it is indeed a valuable method. However, nuclear engineers should always question Monte Carlo results that do not include statistical sampling error, such as the 6% error noted in our simple numerical integration Example. More theory is required in discussing the convergence and variance in Monte Carlo methods. In general applications, mean (μ) and variance (σ) are determined as follows:

$$\mu = E(x) = \frac{\int_{-\infty}^{\infty} x f(x) dx}{\int_{-\infty}^{\infty} f(x) dx} \quad E(x^2) = \frac{\int_{-\infty}^{\infty} x^2 f(x) dx}{\int_{-\infty}^{\infty} f(x) dx} \quad \sigma^2 = E(x^2) - \mu^2$$

Sampling With and Without Replacement

“Replacement,” or a lack of, accounts for a shift in the relative probability as sampling occurs, and this is best discussed using an Example.

Example

Consider that 3 screws are drawn from a lot of 100 that contain 10 defective ones...determine the probability that all 3 are non-defective (with replacement) if we can draw one that is not defective with a probability of 0.9.

With Replacement, 3 sequential draws become

$$(0.9)(0.9)(0.9) = 0.7290 \quad 72.9\%$$

Next, determine the probability that all 3 are non-defective *without replacement*

$$\left(\frac{90}{100}\right)\left(\frac{89}{99}\right)\left(\frac{88}{98}\right) = 0.7265 \quad 72.65\%$$

Permutations [n!]

A number of possible arrangements for a given set of n items is $n!$

Example

Consider a, b, c (a set of 3 letters of the alphabet).

The number of possible arrangements is

abc, acb, bac, bca, cab, cba

the number of which can be determined from

$$3! = 3 \cdot 2 \cdot 1 = 6$$

Combinations

The combination formula can be used to determine the number of possible samples from a defined set. This then can be used to directly determine sampling probabilities.

The Combination formula is:

$$\binom{n}{k} \equiv \left[\frac{n!}{k!(n-k)!} \right] = \frac{n(n-1)\dots(n-k+1)}{1 \cdot 2 \dots k}$$

and yields number of samples or combinations of “ n ” things “ k ” at a time.

Example

What is the probability of winning a “pick 6” lotto where 6 random numbers are chosen from $\{1 \dots 49\}$?

NOTE: let event $A \equiv WIN LOTTO$ from 1 ticket.

First, compute the number of combinations available from choosing 6 random numbers from the set of numbers contained in $\{1..49\}$:

$$\binom{49}{6} = \left[\frac{49!}{6!(49-6)!} \right] = 13.98 \times 10^6$$

Then, the odds of choosing one “winning combination” is

$$P(A) = \frac{1}{\binom{49}{6}} = \frac{1}{13.98 \times 10^6}$$

Example

What if you sold your car and bought \$10k in unique lotto tickets? What is your new probability of choosing a winning ticket?

$$P(A \cup B) = P(x + y) = (10000) \frac{1}{13.98 \times 10^6} = 0.00072$$

(Note the Addition Rule for mutually exclusive events.)

Trying to win the lotto with realistic odds is not an easy (or likely a profitable) task!

Example

Two cards are drawn at random from a standard deck of 52 cards (4 suits of 2... 10, J, Q, K, A). Using the *Combination formula*, find the probability that both cards are Hearts.

Let event $A \equiv$ Both Cards are Hearts.

First, we determine how many ways there are to choose *any* two cards using the combination formula:

$$\binom{52}{2} = 1326 \text{ ways to choose 2 cards from the 52 card deck}$$

Then, we determine how many ways there are to choose two cards that are both Hearts:

$$\binom{13}{2} = 78 \text{ ways to draw 2 Hearts from the 13 available}$$

The solution we seek comes from the **ratio** of these two combination formulas:

$$P(A) = \frac{\binom{13}{2}}{\binom{52}{2}} = \frac{78}{1326} \approx 0.058824$$

Example

Using the *Combination formula*, find the probability that one card is a Heart and one Card is a Spade.

Let $B \equiv$ one card is a Heart, and one card is a Spade.

As before:

$$\binom{52}{2} = 1326 \text{ ways to choose 2 cards from the 52 card deck}$$

Then, there are 13 hearts, and there are 13 spades; herefore, there are $(13)(13) = 169$ ways to draw a single heart and a single spade.

Note this is the same as computing

$$\binom{13}{1} \binom{13}{1} = (13)(13) = 169$$

And the solution is:

$$P(B) = \frac{(13)(13)}{\binom{52}{2}} = \frac{169}{1326} \approx 0.12745$$

Binomial Coefficients

The numbers from the combination formula

e.g. consider the following:

$$(a + b)^3 = a^3 + 3a^2b + 3ab^2 + b^3$$

Nuclear System Reliability

Example:

Consider two redundant, hypothetical nuclear rated valves at a nuclear power plant, where we have valve #1 on cooling circuit #1, and valve #2 on cooling circuit #2. Then consider the following:

event A = valve #1 will operate for 20 years without failure

event B = valve #2 will operate for 20 years without failure

We will assume A and B are independent events, and that

$$P(A) = \frac{1}{4} \quad P(B) = \frac{1}{3}$$

(note these values are for illustrative purposes only!)

The probability that both will operate after 20 years is given by the multiplication rule (note this is indicated by the “both” statement), where:

$P(AB) = P(A) P(B)$ is also written

$$P(A \cap B) = P(A) P(B) = \frac{1}{12} = 0.0833$$

The probability that either A or B will operate after 20 years, as suggested by the wording, is governed by the addition rule, where

$P(A + B) = P(A) + P(B) - P(AB)$ is also written

$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$

$$P(A \cup B) = \frac{1}{4} + \frac{1}{3} - \frac{1}{12} = \frac{1}{2}$$

The probability that neither A or B will operate after 20 years is therefore the complement of $P(A \cup B)$:

$$P((A \cup B)^c) = 1 - P(A \cup B) = \frac{1}{2}$$

Since these events are independent, this could also be determined from a basis using complements, with the probabilities the valves will *fail*:

$$P((A)^c) = \frac{3}{4} \qquad P((B)^c) = \frac{2}{3}$$

Then

$$P((A)^c \cap (B)^c) = \frac{3}{4} \cdot \frac{2}{3} = \frac{1}{2}$$

The probability that only valve B will be operating after 20 years is therefore:

$$P((A)^c \cap B) = \frac{3}{4} \cdot \frac{1}{3} = \frac{1}{4}$$

The probability of survival or failure pertaining to equipment in a nuclear power plant is of course a major issue in licensing the facility for safe operation, and numerous redundancies must be analyzed to assess the risks to the plant.

Note here we have treated the two events considered to be independent. While this is normally the case, it is not always true, and interdependent probabilities can become quite complex to evaluate. Often, it is simpler to have independent redundant systems, since this tends to reduce failure and simplifies the analysis.

Finally, large diagrams that resemble “trees” map out specific chains of events, and this procedure is called “fault tree analysis”, or FTA. FTA is a “top-down” approach for analyzing potential system failures before they occur for systems under development, beginning with the top event (the initiating potential failure). Analysts then determine all pathways

the failure leads to that affect the system, which ultimately determines risk.

This is a routine practice in engineering design, particularly important in nuclear engineering for obvious reasons.

Radioactive Decay

An atom with an unstable nucleus will, at some point in time, rearrange itself to a more stable configuration, although the precise instance of when this will occur is, for an individual atom, *a low probability event*.

For Example, consider the decay probability of a Pu-239 nucleus per second, $9.114\text{E-}13$ 1/s.

If a significant number of like unstable isotopes can be gathered together, then the rearrangements, or “decays” can be made at a predictable rate that is individually characteristic of the isotope.

The “decay” process is traditionally measured in terms of a “half life” where $\frac{1}{2}$ of the nuclides decay after a measure of time :

$$\frac{dN}{dt} = -\lambda N$$

Recognizing this is a separable differential equation describing $N(t)$, then

$$\int \frac{dN}{N} = \int -\lambda dt \quad \text{yields} \quad \ln(N) = -\lambda t + c_1$$

Taking an exponential of both sides...

$$N(t) = \exp(c_1) \exp(-\lambda t)$$

and assuming we start from an initial number N_o , then

$$N(t) = N_o \exp(-\lambda t)$$

The probability of decay of a nuclide λ , in the interval between t and $t + dt$, even if small, can be determined from the half life measured from a very large number of nuclei decaying (applying the Law of Large Numbers):

$$\frac{N(T_{half})}{N_o} = \frac{1}{2} = \exp(-\lambda T_{half})$$

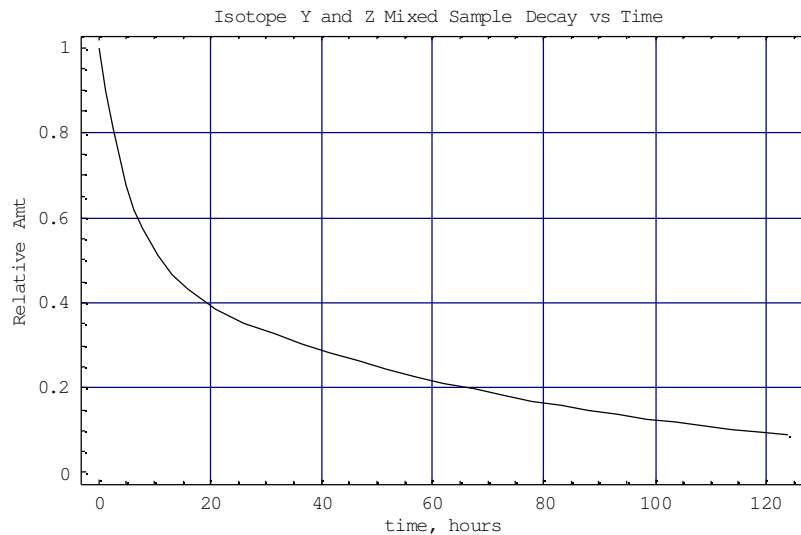
Then, taking the log of both sides, canceling signs, and solving for the probability:

$$\lambda = \frac{\ln(2)}{T_{half}}$$

Often one must consider a practical limit for when an isotope has “disappeared,” even though it is never fully “gone.” Usually, this is after 5 to 10 half lives, depending on limits, where one part in 32 (~3%) or one part in 1024 (~0.1%) remain, respectively.

In the early years of nuclear physics and engineering, radiochemists determined half lives of isotopes graphically by using the fact that two nuclides with different half lives will decay away at different rates.

By plotting the relative activities of the combined “cocktail” of nuclides and iterating on initial guesses for half lives based on the graph, the half lives of the two decaying isotopes can be effectively isolated. An Example of this is shown below.

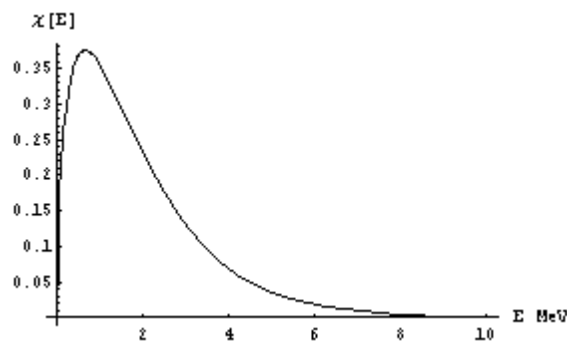


Two isotopes with half lives of 4 and 50 hours, respectively.

Probability Distributions: Fission Spectrum

Fission neutrons are emitted over a continuous probability distribution of energy, known in nuclear engineering as the *Prompt Fission Neutron Spectrum*.... Note this implies the existence of a *Delayed Neutron Spectrum*. A brief word on the subject of *prompt* versus *delayed* neutrons is worth noting. Most neutrons emitted in fission are in fact prompt—they are emitted within $1E-13$ seconds of a fission event. Indeed, a small fraction of the neutrons emitted in fission are *delayed* in their emission following a fission event, with half-lives measured as long as *several minutes*. This small fraction of neutrons, emitted at some protracted time period(s) after a fission event, is known as β – the delayed neutron fraction, and it is 0.65% for U-235, 0.21 % for Pu-239. *Were it not for the delayed neutron fraction*, the only nuclear power available to us would be the *uncontrolled chain reaction of nuclear weapons*, since the delayed neutrons are emitted in times reasonable for mechanical control of the fission chain reaction in a reactor. Needless to say, when a reactor does not rely on the delayed neutrons for a chain reaction, this is described as being in a “*prompt critical*” condition. Unless the reactor is designed for this in a “*pulse*” condition, such a condition would normally result in a reactor accident; many safeguards are in place to prevent “prompt” criticality. We will return to this subject a bit later.

Now, consider that $\chi(E)$ is called the *Prompt Fission Neutron Spectrum*. This spectrum for U-235 is given below:



Prompt fission spectrum for U-235.

For a U^{235} nucleus, the Maxwellian Prompt Fission Neutron Spectrum is:

$$\left[\chi(E) = \frac{77}{100} \sqrt{E} e^{-\left(\frac{0.776}{\text{MeV}} E\right)} \right] \quad \text{with } E \text{ in MeV}$$

$\chi(E) dE \Rightarrow$ is the average number of neutrons emitted with an energy E between E and $E + dE$ per fission neutron,

$$\text{Where } dE = \lim_{\Delta E \rightarrow 0} \Delta E$$

Another way to think of this is:

$\chi(E) dE \Rightarrow$ is the net fraction of fission neutrons emitted with an energy E between E and $E + dE$.

Note that Energy in Nuclear Engineering is given in units of (eV, KeV, or MeV), 1 electron volt (eV) is the energy imparted to a free electron when accelerated by an electric field potential of 1 Volt. Specifically,

$$1 \text{ eV} = 1.602 \times 10^{-19} \text{ J} \quad \text{and} \quad 1 \text{ MeV} = 1.602 \times 10^{-13} \text{ J}$$

$$\text{Noting that by definition, } \int_0^{\infty} \chi(E) dE = 1$$

$$\Rightarrow \chi(E) = C_m \sqrt{E} e^{-\alpha E} \quad \text{where } C_m \equiv 0.77, \quad \alpha \equiv 0.776$$

Now solve for the most probable fission neutron energy:

$$\frac{d\chi}{dE} = \left[\frac{77}{100} \frac{1}{2} E^{-1/2} e^{-0.776E} + \frac{77}{100} E^{1/2} e^{(-0.776E)} (-0.776) \right]$$

$$\text{Set } \frac{d\chi}{dE} = \frac{77}{100} e^{-0.776E} \left(\frac{1}{2\sqrt{E}} + E^{1/2} (-0.776) \right) = 0$$

$$\frac{1}{2\sqrt{E}} + \sqrt{E}(-0.776) = 0 \quad \text{or} \quad \frac{1}{2\sqrt{E}} = \sqrt{E}(0.776)$$

$$\text{Then } E = \frac{1}{2(0.776)} = 0.6443 .$$

Typically, the fission spectrum is given as either a *Maxwellian Fission Spectrum* or a *Watt Fission Spectrum*. The forms are:

$$\text{Maxwellian Spectrum: } \left[\chi(E) = C_m \sqrt{E} e^{-(\alpha E)} \right]$$

$$\text{Watt Spectrum: } \left[\chi(E) = C_m e^{-(\alpha E)} \sinh(\beta E) \right]$$

Note that a unique $\chi(E)$ exists for every fissile nuclide. It is therefore always important to verify that

$$\int_0^{\infty} \chi(E) dE = 1$$

for the spectrum given.

The average fission neutron energy is computed from

$$\bar{E} = \frac{\int_0^{\infty} E \chi(E) dE}{\int_0^{\infty} \chi(E) dE} = \frac{1}{1} \int_0^{\infty} E \chi(E) dE$$

Really, this is an “improper integral”

$$\lim_{b \rightarrow \infty} \int_0^b E (C_m E^{1/2} e^{-\alpha E}) dE = \lim_{b \rightarrow \infty} \int_0^b C_m E^{3/2} e^{-\alpha E} dE$$

This is a special form of the exponential integral

$$\lim_{b \rightarrow \infty} \int_0^b C_m x^{3/2} e^{-\alpha x} dx$$

$$\lim_{b \rightarrow \infty} \left[e^{-\alpha x} C_m \left(-\frac{3\sqrt{x}}{2\alpha^2} - \frac{x^{3/2}}{\alpha} \right) + \frac{3\sqrt{\pi} C_m \operatorname{Erf}[\sqrt{\alpha x}]}{4(\alpha)^{5/2}} \right]_0^b$$

Where $\operatorname{Erf}[x]$ is the Gaussian Error Function, given below.

Gaussian Error Function

$$\operatorname{Erf}[x] = \frac{2}{\sqrt{\pi}} \int_0^x \exp[-u^2] du \quad \text{with}$$

$$\lim_{b \rightarrow 0} \operatorname{Erf}[b] = 0 \quad \text{and} \quad \lim_{b \rightarrow \infty} \operatorname{Erf}[b] = 1$$

Standard Probability Distributions

Binomial Distribution

The *Binomial Distribution* is a probability distribution function (pdf) that can be applied in the following cases:

There are two possible outcomes: {A,B}

P(A) is constant, independent of number of observations and
 $P(B) = 1 - P(A)$

Occurrences of A observed (or B observed) do not affect
 $P(A)$ or $P(B)$.

$$f_B(x) = \binom{n}{x} p^x q^{n-x} = \binom{n}{x} p^x (1-p)^{n-x}$$

So that the Binomial Distribution Formula is

$$f_B(x) = \left[\frac{n!}{x!(n-x)!} p^x q^{n-x} \right]$$

The Binomial Distribution formula applies in cases like a *coin toss* and in the *decay of a few atoms*.

Example

What is the probability of getting heads in a coin toss 3 times in a row?

$A \equiv$ get a heads up coin toss

$B \equiv$ don't get a heads up coin toss (tails)

$n = 3$ total trials

$x =$ getting heads 3 times in a row

$$p = 0.5 \quad q = (1 - 0.5)$$

$$f_B(3) = \left[\frac{3!}{3!(3-3)!} 0.5^3 (1-0.5)^{3-3} \right] = 0.125$$

Two other important Probability Distribution Functions (PDFs) include the *Poisson Distribution* and the *Normal Distribution*

Poisson Distribution

Applies to events with a *small* but *constant* probability of occurrence

Is derived from Binomial Probability Density Function, we apply the

limits as $\left\{ \begin{array}{l} n \rightarrow \infty \\ p \rightarrow 0 \end{array} \right\}$

$$f_p(x) = \frac{\mu^x}{x!} e^{-\mu}$$

where $\mu \equiv$ average (“mean”) for a large number of trials, and $x \equiv$ result of very next trial

$$\sigma_p = \sqrt{\mu}$$

Radiation Detection Statistics

A radiation detector is used to count particles from a radioactive isotope. If the average count rate is 20 counts/min (cpm), what is the probability the next trial will yield 18 cpm?

Then: $\mu = 20$

$x = 18$

$$f_p(x) = \frac{20^{18}}{18!} e^{-20} = 0.0844 \text{ or } \sim 8\%$$

NOTE: As the mean *increases*, the Poisson distribution becomes *symmetric about the mean*.

Other important things to remember:

→ Both the *Binomial* and *Poisson* distributions apply to discrete variables or events.

→ Most variables in experiments are continuous...

Gaussian Distribution

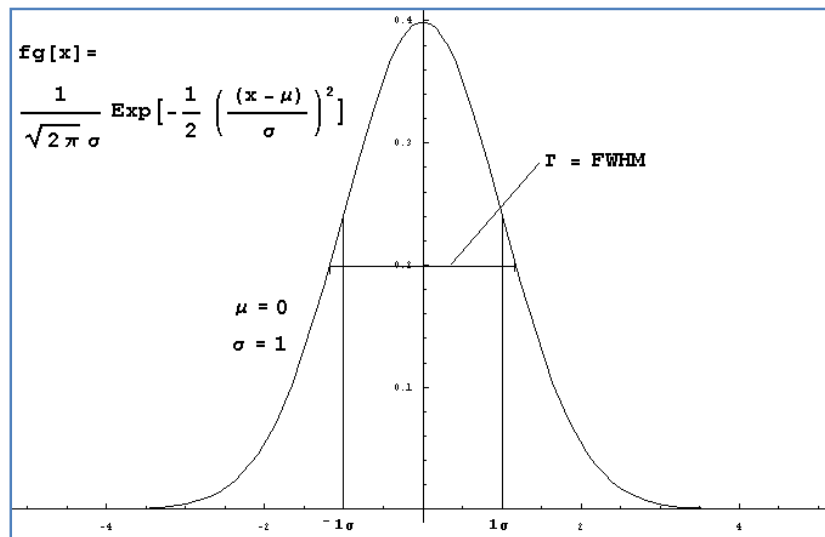
$$f_g(x) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left[-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right]$$

Where “ x ” is for a continuous distribution with a maximum about “ μ ” (the mean), and σ is the standard deviation, where x extends to $-\infty < x < \infty$

The *Full Width Half-Maximum* (FWHM) of the Gaussian function is defined as the peak width at precisely one half of the peak maximum, and is determined mathematically by:

$$f_g\left(\mu - \frac{\Gamma}{2}\right) = f_g\left(\mu + \frac{\Gamma}{2}\right) = \frac{1}{2} f_g(\mu)$$

Solving this yields $\Gamma = 2.35 \sigma$.



A normalized Gaussian distribution function with a zero mean, unit standard deviation.

Central Limit Theorem

The *Central Limit Theorem* states that sampled mean values, e.g. from a collection of Monte Carlo sampling histories, follow a Gaussian distribution for n stochastic samples of a probability density function, and that the uncertainty is proportional to the inverse of the square root of the number of particle histories.

Example

You are asked to run a Monte Carlo simulation to estimate a dose that will be delivered to a patient, and your supervisor has asked for a solution with 1.5% or less uncertainty.

Assume the *Central Limit Theorem* can be applied. The Chief of Physics calls to tell you the solution is needed as quickly as possible for a high risk treatment plan. He tells you he is waiting by the phone for your answer.

You decide to set up an initial one hour run of the Monte Carlo simulation, and after 60 minutes of computing time, the run completed $n = 1,000,000$ particle histories, and the dose tally of interest reported by the Monte Carlo code is 170 Gy/hr +/- 0.4%.

Since you only needed 1.5% accuracy and the Chief of Physics was waiting, estimate how many fewer histories could have been run, and in how much time a shorter run would have required to achieve tally results within +/- 1.5%.

To solve this, we assume that the central limit theorem applies, so that we can assume that the standard deviation is proportional to the inverse of the square root of the number of histories sampled:

$$\sigma = k / \sqrt{N}$$

From this, for the Example given, we determine that $k = 400$. With this knowledge, we can then apply the same principle, assigning $\sigma = 1.5$ and solving for a new N , which at this point yields 71,111 histories required for the standard deviation given.

Then, assuming a linear extrapolation of histories sampled correctly accounts for computation time (a reasonable assumption), we obtain that the actual model should have only required less than 4.3 minutes to yield a 1.5% accuracy goal, and that 60 minutes was quite an overestimate of the time.

Detector Peak Resolution

Detector Peak *Resolution* (R) and the FWHM is an often discussed issue when dealing with radiation detectors. Consider scintillator detectors attached to multi-channel analyzers, which “bin” the energy values initially recorded as light pulses as a result of interactions of the radiation in a gamma radiation scintillator detector. Due to the variation in the signals, these events yield Gaussian peaks that have a mean based on the energy of the incident gamma radiation.

Resolution of the detector is defined as a percentage derived from

$$R(\mu) = \frac{\Gamma}{\mu}$$

so that detectors with better resolution have narrow FWHMs at a given energy. For Example, if the normalized Gaussian in the last figure depicted a counts vs energy over a reasonable number of channels for a detector, and if $\mu = 662$ keV for the gamma ray energy emitted from a Cs-137 gamma ray emitting isotope (Cesium-137 is a fission product), if the resolution of the detector at 662 keV was $R=3\%$, then you would know that the FWHM at 662 keV was 19.9 keV.

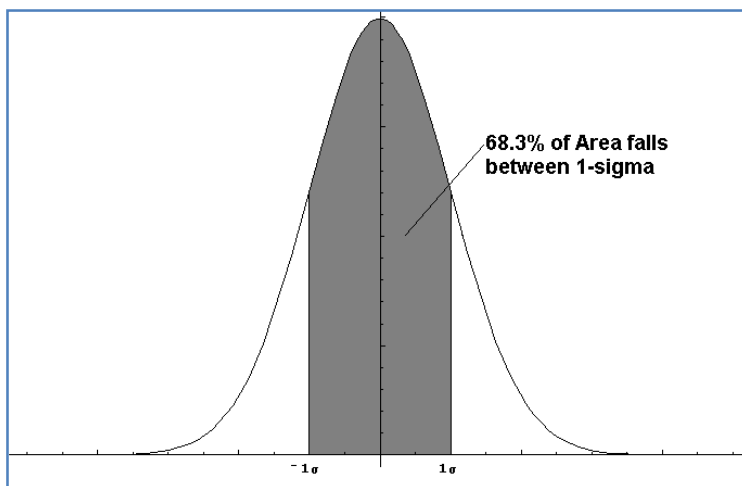
Resolution is an important fact in being able to determine if one can clearly distinguish between two gamma lines in a gamma radiation detector, and it is desirable to have the FWHM as low as possible.

Integration of the Gaussian Function

Integration of the Gaussian yields a special function denoted as $Erf(x)$, the *Gaussian Error Function*, or *ERF*:

$$Erf(x) = \frac{2}{\sqrt{\pi}} \int_0^x \exp(-t^2) dt$$

$$A_\sigma = \int_{\mu-\sigma}^{\mu+\sigma} f_g(x) dx = 2 Erf(\mu + \sigma) = 0.683$$



A graphical depiction of 68% of the area under the Gaussian function.

For a normalized Gaussian, 68.3% of the curve falls under the interval characterized by “1-sigma” spanning the mean. Similar values for “2-sigma” and “3-sigma” are 95.4% and 99.7% of the area, respectively. This is often coined as the “68-95-99+” rule.

This chapter provided a *very* brief look at probability and some issues related to it applied to problems of interest to nuclear engineers and medical physicists. It is recommended that the student further study this interesting and at times complex subject with a dedicated text on the subject.