Chapter 2: Natural distributions in combinational circuit theory

December 19, 2003

Abstract

The goal of this chapter is to introduce the log-normal distribution. Log-normal distribution repeatedly shows up in fault-detection analysis, functional equivalence verification, generalized Cutting alogorithm, and metric theory. Therefore, knowledge and through formulation of log-normal distribution, becomes necessary and crucial.

1 Formulating Distributions in circuit theory

Any circuit function, can be writtern,in sums of products as in,

$$\sum_{k}^{\text{length}} \prod_{i} Y_i \tag{1}$$

where Y_i is either X_i or $1-X_i$. Given that X_i are input probabilities , each term $\prod Y_i$ contributes to the probability. Now, instead of having a large function that blows up exponentially, it is possible,to work with it, statistically. It is important to see it as a statistical representation. This representation will fail, under certain circumstances. we will go into issues, later and preventive measures that can taken, to circumvent to rescue it.

Take $\prod_i Y_i$. Now, if you take the log of it.

$$\log \prod_{i} Y_i = \sum_{i} \log Y_i \tag{2}$$

As said earlier, Y_i will either be X_i or $1 - X_i$. And thus, $\log \prod_i Y_i$ behaves like a gaussian variable. As with most gaussian phenomenon, things of most interest are the mean and the standard deviation.

Let μ_{lt} be the mean of log of the terms. And let σ_{lt} be the standard deviation of log of the terms. Suppose, y is a probability value contributed by a term, then probability of finding a term, that supplies y is given by

$$p(y,\mu_{lt},\sigma_{lt}) = \frac{1}{\sqrt{2\pi}\sigma_{lt}y}\exp\left(-\frac{(\log y-\mu_{lt})^2}{2\sigma_{lt}^2}\right) \tag{3}$$

The problem with this distribution, is that it assumes that the distribution is continous and this assumption has problems, of its own, which we will discuss later. We will fix that assumption, by putting a unit step function in there.

$$p(y,\mu_{lt},\sigma_{lt},y_{max},y_{min}) = \frac{1}{\sqrt{2\pi}\sigma_{lt}y}\exp\left(-\frac{(\log y-\mu_{lt})^2}{2\sigma_{lt}^2}\right)[u(y_{max}-y)-u(y_{min}-y)] \tag{4}$$

y_{max} is the value of the term that supplies the maximum individual probability and y_{min} is the value of the term that supplies the least invidual probability. Suppose, N is the number of terms in the circuit, then,

$$\int_0^1 dyN\cdot y\cdot p(y,\mu_{lt},\sigma_{lt},y_{max},y_{min}) \tag{5}$$

would give the expected value of the output. Now, suppose, we decide to rewrite the above integration. Let $x=-\log y$. Then $y=\exp(-x)$ and $dx=-dy/y$, then $\exp(-x)dx=-dy$. Now,

$$\int_\infty^0 dxN\left[-\exp(-x)\right]\frac{1}{\sqrt{2\pi}\sigma_{lt}}\exp\left(-\frac{(x-\mu_{lt})^2}{2\sigma_{lt}^2}\right)[u(x--\log y_{max})-u(x--\log y_{min})] \tag{6}$$

Now, with a little bit of rearragement, it becomes,

$$\int_0^\infty dxN\exp(-x)\frac{1}{\sqrt{2\pi}\sigma_{lt}}\exp\left(-\frac{(x-\mu_{lt})^2}{2\sigma_{lt}^2}\right)[u(x--\log y_{max})-u(x--\log y_{min})] \tag{7}$$

which is the laplace transform of the gaussian,with s set to 1. We shall generalize the above expression as follows

$$\int_0^\infty dxN\exp(-s\cdot x)\frac{1}{\sqrt{2\pi}\sigma_{lt}}\exp\left(-\frac{(x-\mu_{lt})^2}{2\sigma_{lt}^2}\right)u(x-c) \tag{8}$$

$$\tag{9}$$

for $c\geq 0$, it gives

$$\int_c^\infty dxN\exp(-s\cdot x)\frac{1}{\sqrt{2\pi}\sigma_{lt}}\exp\left(-\frac{(x-\mu_{lt})^2}{2\sigma_{lt}^2}\right) \tag{10}$$

Integrating with Maple or Mathematica,we get

$$\frac{N}{2}\exp\left(-\mu_{lt}s+\frac{\sigma_{lt}^2s^2}{2}\right)\operatorname{erfc}\left(\frac{-\mu_{lt}+\sigma_{lt}^2s+c}{\sqrt{2}\sigma_{lt}}\right) \tag{11}$$

We define

$$\Gamma(s,c,\mu_{lt},\sigma_{lt}) = \frac{1}{2}\exp\left(-\mu_{lt}s+\frac{\sigma_{lt}^2s^2}{2}\right)\operatorname{erfc}\left(\frac{-\mu_{lt}+\sigma_{lt}^2s+c}{\sqrt{2}\sigma_{lt}}\right) \tag{12}$$

2 Properties

Now, $N\Gamma(s, c, \mu_{lt}, \sigma_{lt})$ is the probability of output being a 1. Now, $N^2\Gamma^2(s, c, \mu_{lt}, \sigma_{lt})$ is the probability of the output being a 1, consecutively. Now, $N^k\Gamma^k(s, c, \mu_{lt}, \sigma_{lt})$ is the probability that the function is a 1, in k consecutive runs.
It is interesting to see that in the following light. E[C($\mathbf{X}$)] is the expected value of the function C($\mathbf{X}$) with n independent inputs. Then, E[C($\mathbf{X}$)]E[C($\mathbf{X}$)], is equivalent to the expected value of a function with 2-n independent inputs.To see this explicity, note the following.

$$\log \prod_i Y_i = \sum_i \log Y_i \tag{13}$$

Now, multiplication of the circuit, by itself, constitutes,

$$\log \prod_i^{2n} Y_{\text{new i}} = \sum_i^{2n} \log Y_{\text{new i}} \tag{14}$$

Now, after i goes pass n, the properties of Y_0, Y_1, etc, repeat. So, the sum becomes

$$\log \prod_i^{2n} Y_{\text{new i}} = \sum_i^{n} \log Y_i + \sum_i^{n} \log Y_i' \tag{15}$$

Now, Y_i' is used to denote that, although ,they do have the same properties as Y_i, they are statiscally indepedent.
More importantly, the new addition of n-input, changes the function, in such a way that $\langle lt \rangle$ and c increases two folds, and σ_{lt}^2 also increases two folds.
To see why c increases two folds. Remember that $c = \log(\max_{term})$.

$$\max \log \prod_i^{2n} Y_{\text{new i}} = \max \sum_i^{n} \log Y_i + \max \sum_i^{n} \log Y_i' \tag{16}$$

max of $\log Y_i'$ has to be atleast as maximum as $\max \log Y_i$. Otherwise, it would not be a maximum, since Y_is andY_i' all have the same properties.
Generally, suppose if we take E[C($\mathbf{X}$)]^k, μ_{lt}, c and σ_{lt}^2 increase k-fold.

3 μ_{lt} and σ_{lt}

Now, to compute, μ_{lt} and σ_{lt}, assuming independence of terms involving x_i with terms involving x_j. We use the following lemmas

Lemma 1 *Given, independent variables, y_1, y_2,..., y_n, we have,*

$$\langle \sum_i y_i \rangle = \sum_i \langle y_i \rangle \tag{17}$$

$$\mathrm{Var}(\sum_i y_i) = \sum_i \mathrm{Var}(y_i) \tag{18}$$

Using the condition of independence, and the above lemma, we can prove the following

$$\mu_{lt} = \sum_i -w_i \log(x_i) - (1 - w_i) \log(1 - x_i) \tag{19}$$

where w_i is the probability that a term involve x_i, as opposed to $1 - x_i$. Then, $1 - w_i$, is the probability that the term involves $1 - x_i$. Now, to compute the standard deviation of log of the terms, we just simply take

$$\begin{aligned}
\sigma_{lt}^2 &= \sum_i w_i \log^2(x_i) + (1 - w_i) \log^2(1 - x_i) - \sum_i \left(-w_i \log(x_i) - (1 - w_i) \log(1 - x_i)\right)^2 && (20)\\
&= \sum_i w_i \log^2(x_i) + (1 - w_i) \log^2(1 - x_i) - && (21)\\
&\quad \sum_i w_i^2 \log^2(x_i) + (1 - w_i)^2 \log^2(1 - x_i) - 2w_i(1 - w_i) \log(1 - x_i) \log(x_i) && (22)\\
&= \sum_i (w_i - w_i^2) \log^2(x_i) + \left[(1 - w_i) - (1 - w_i)^2\right] \log^2(1 - x_i) && (23)\\
&\quad -2w_i(1 - w_i) \log(1 - x_i) \log(x_i) && (24)\\
&= \sum_i (w_i - w_i^2) \log^2(x_i) + (-w_i + 2w_i - w_i^2) \log^2(1 - x_i) - 2w_i(1 - w_i) \log(1 - x_i) \log(x_i) && (25)\\
&= \sum_i (w_i - w_i^2) \log^2(x_i) + (w_i - w_i^2) \log^2(1 - x_i) - 2w_i(1 - w_i) \log(1 - x_i) \log(x_i) && (26)\\
&= \sum_i w_i(1 - w_i) \left(\log(x_i) - \log(1 - x_i)\right)^2 && (27)
\end{aligned}$$

For a given circuit, w_i, remain unchanged, even when input probabilities change. So, now, if we reformulate, our fault distribution, in language of w_is. It should be noted that, the above relationship, will still hold, if we replace w_i with an arbitrary a_i and $1 - w_i$ with b_i. To see this, suppose, we multiply the w_i and $1 - w_i$ by c_i, we would have $c_i w_i \log(x_i) + c_i(1 - w_i) \log(1 - x_i)$ and and we woudl have $c^2 w_i(1 - w_i) \left(\log(x_i) - \log(1 - x_i)\right)^2$. So,if $a_i = c w_i$ and $b_i = c(1 - w_i)$, we would have

$$\mu_{lt} = \sum_i a_i \log(x_i) + b_i \log(1 - x_i) \tag{28}$$

and

$$\sigma_{lt}^2 = \sum_i a_i b_i \left(\log(x_i) - \log(1 - x_i)\right)^2 \tag{29}$$

4 Dynamics: Predictables and Unpredictables

The expected value of the log-normal distribution is extremely chaotic. A small percent of the log-normal distribution, contributes to the majority of the expected value.
Take

$$\frac{N}{2}\exp\left(-\mu_{lt}s+\frac{\sigma_{lt}^2 s^2}{2}\right)\operatorname{erfc}\left(\frac{-\mu_{lt}+\sigma_{lt}^2 s+c}{\sqrt{2}\sigma_{lt}}\right) \tag{30}$$

Suppose, the weightset changes from time to time. Knowing w_i, we can successfully estimate σ_{lt}, and μ_{lt}, from the formulas, in the above section. However generally, estimating c is just as bad as finding the actual c. In fact, c, the value of the term, that contributes the highest probability, holds the key to the correct order of the estimate.

4.1 The Unpredictable 'c'

Suppose, we naively decide to estimate c as by taking

$$\prod_i \max\{1-x_i, x_i\} \tag{31}$$

Notice that each time, we take the $\max\{1-x_i, x_i\}$, we would either have $1-x_i$, or x_i. In the end, we would have a product like $(1-x_0)(x_1)x_2(1-x_3)x_4x_5x_5(1-x_7)(1-x_8)$. However, for that term, to be the maximum value supplying term, it must be part of the actual function. It is not required to be part of the function. And when it is not in it, it is merely a worst possible bound on the actual maximum, that is part of the function. In general, estimating c, could be an NP-complete process.

5 Fault Detection Analysis

The following analysis is based on an appoarch by Seth-Agarwal-Farat. We reformulate their method, in terms of vectors that detect those faults , intead of the faults themselves.
Imagine stuck at faults f_1, f_2, f_3, ... f_n in the circuit. Each fault is detected by a set of vectors. Suppose, we build a circuit $\mathrm{F}_n(\mathbf{X})$, which is 1, when vector X, detects the fault f_n, otherwise zero. Now, each such circuit for different faults f_n has σ_n , μ_n, and N_n parameters. Now, the probability that a f_n is detected at time instant k, but not before is.

$$\int_0^1 dy N_n \cdot (1-y)^{k-1} y \cdot p(y, \mu_n, \sigma_n, y_{nmax}) \tag{32}$$

which is same as saying that one of the vectors, that detects f_n, becomes a 1at time instant k.

Now, the probability that it is detected anytime, upto the time instant k is

$$= \int_0^1 dyN_n \left[1 + (1-y) + (1-y)^2 + ... + (1-y)^{k-1}\right] y \cdot p(y, \mu_n, \sigma_n, y_{nmax}) \tag{33}$$

$$= \int_0^1 dyN_n \left[\frac{(1-y)^k - 1}{1-y-1}\right] y \cdot p(y, \mu_n, \sigma_n, y_{nmax}) \tag{34}$$

$$= \int_0^1 dyN_n \left[1 - (1-y)^k\right] p(y, \mu_n, \sigma_n, y_{nmax}) \tag{35}$$

$$= \int_0^1 dyN_n \sum_{i=1}^{k} (-1)^{i+1} \mathrm{C}_i^k y^i p(y, \mu_n, \sigma_n, y_{nmax}) \tag{36}$$

$$\tag{37}$$

which is wonderfully the same as,

$$N_n \sum_{i=1}^{k} (-1)^{i+1} \mathrm{C}_i^k \Gamma(i, c_n, \mu_n, \sigma_n) \tag{38}$$

Now, suppose $q(N, c, \mu, \sigma)$ is the normalized distribution of N, c,μ, and σ, we would have

$$\int\int\int\int dNdcd\mu d\sigma \sum_{i=1}^{k} (-1)^{i+1} Nq(N, c, \mu, \sigma) \mathrm{C}_i^k \Gamma(i, c, \mu, \sigma) \tag{39}$$

would give,the probability of faults detected, by iteration k. We shall come back to this expression later. The above expression, should be seen as a mathematical device,which can be used,to prove several properities of fault distribution, fault entropy, etc, in circuits. In fact, based on aprior model of N, c and μ, and σ , one can modify $q(N, c, \mu, \sigma)$ to prove, properties of such models. We shall, in the following sections, use such models, to prove, why increasing input entropy, would increase fault coverage.
One can reformulate this relationship in terms of **a**'s and **b**'s, and see a direct relationship between input probabilities and fault distribution.

$$\int\int\int\int dNdcdadb \sum_{i=1}^{k} (-1)^{i+1} Nq'(N, c, a, b) \mathrm{C}_i^k \Gamma\left(i, c, \mu(\mathbf{X}, a, b), \sigma(\mathbf{X}, a, b)\right) \tag{40}$$

where

$$\mu(\mathbf{X}, a, b) = \sum_i a \log x_i + b \log(1 - x_i) \tag{41}$$

and

$$\sigma^2(\mathbf{X},a,b)=\sum_i ab\left(\log x_i-\log(1-x_i)\right)^2 \tag{42}$$

Now, suppose, one tries to maximize fault-coverage, with respect to x_i , one needs to differentiate, the above equation with respect to x_i and set it to 0. Doing that would lead to the equation,

$$\int\int\int\int dNdcdadb\sum_{i=1}^{k}(-1)^{i+1}Nq'(N,c,a,b)\mathrm{C}_i^k\left[\Gamma_\mu\frac{\partial\mu}{\partial x_i}-\Gamma_\sigma\frac{\partial\sigma}{\partial x_i}\right]=0 \tag{43}$$

6 A Constrained and Specific Solution

The above equation, has a very interesting solution, for a specific case of as and bs. Suppose, $a=w_i$ and $b=1-w_i$. Now if $q(w_i)$ is a narrow gaussian around $w_i=0.5$. Then $q(w_i)=\delta(w_i-0.5)$ where $\delta(x)$ is the dirac-delta with property,

$$\int_{-\infty}^{\infty}dx\delta(x-a)f(x)=f(a) \tag{44}$$

Given that $w_i=0.5$ is a very ideal assumption. Without the loss of generality, we can assume that the above function is function of σ^2 as apposed to σ. Notice that

$$\frac{\partial\sigma^2}{\partial x_i}=2\frac{w_i(1-w_i)}{x_i(1-x_i)}\log\left[\frac{x_i}{1-x_i}\right] \tag{45}$$

and

$$\frac{\partial\mu}{\partial x_i}=-\left[\frac{w_i-x_i}{x_i(1-x_i)}\right] \tag{46}$$

Notice that for $x_i=0.5$, the fault-coverage is maximum for the ideal case, since $\frac{\partial\sigma^2}{\partial x_i}=\frac{\partial\mu}{\partial x_i}=0$.

7 A Functional Verification Problem

Dr. Agarwal, Dr. Seth and others proved that, given combinational circuits $\mathrm{C}_1(\mathbf{X})$ and $\mathrm{C}_2(\mathbf{X})$, suppose we randomly generate a weightset $\mathbf{X}$, and if $\mathrm{E}[\mathrm{C}_1](\mathbf{X})=\mathrm{E}[\mathrm{C}_2](\mathbf{X})$, it follows that $C_1=C_2$, since the probability that circuits are equal for a random weightset is 0.

7.1 Refutation

Their reasoning is in error, but however, it is still accurate probabilitistically. For a randomly generated weightset x_i, the probability that different circuits have the same probability is finite, but very small.

Proof 1 *Suppose n is the number of inputs of both the functions. Randomly generate a weightset* $\mathbf{X}$. *Now notice that there are only finite number of function that can constructed as sum of products. However, there are* $\mathrm{pow}(2, 2^n)$ *such functions.So, the distribution of probabilities is finite.*

One can compute the distribution of $\mathrm{E}[\mathrm{C}_1](\mathbf{X}) - \mathrm{E}[\mathrm{C}_2](\mathbf{X})$. To get a feel for our formulation, imagine all possible differences between 2-bit circuits, written as.

$$\{0| \pm 1\}(1-x)(1-y) + \{0| \pm 1\}(1-x)y + \{0| \pm 1\}x(1-y) + \{0| \pm 1\}xy \tag{47}$$

the mean μ_Δ is 0. The distribution is gaussian. The variance σ_Δ is given by

$$= \frac{2}{3}\left[(1-x)^2(1-y)^2 + (1-x)^2y^2 + x^2(1-y)^2 + x^2y^2\right] \tag{48}$$

$$= \frac{2^n}{3}\Gamma\left(1, 0, -\sum_i \log\left[x_i(1-x_i)\right], \sqrt{\sum_i \log^2\left[\frac{x_i}{1-x_i}\right]}\right) \tag{49}$$

computed by varying formulas for μ_{lt} and σ_{lt}. Now, worst case probability that difference between 2 circuits are equal is about