Luria-Delbruck Mathematical Supplement

April 20, 2013

0.1 Mathematical Toolbox

0.1.1 The Binomial Distribution

The **binomial distribution** is a discrete probability distribution used for characterizing the probability of binary events, such as a coin toss.

Here's an example:

You are a basketball player and you want to see the probability that you will make 3 out of 5 free throws in your practices. From your previous practices, you know roughly the chance of making each free throw, p. Let's assume that your skills are neither improving nor declining such that p remains a constant and can be obtained from your previous practices.

There are two concepts one must combine to answer this question:**1)** the individual probability of making 3 out of 5 shots, m, and **2)** the number of different ways that you can get 3 successes and 2 failures exactly.

The first point can be calculated by

$$m = p^3(1-p)^2. \tag{1}$$

The second point can be calculated by realizing that this is a combination problem. You have to note that for combination problems, the sequence of events isn't so specific so the number of different sequences is always less than it would for permutation problems. In other words, it doesn't make a difference in our counting if you tell us you have made the 1st, 2nd, and 3rd shots or you tell us that you made the 3rd, 2nd, and 1st shots. In a permutation problem, this subtle difference would be accounted for. In this example, it doesn't even make any physical sense for you to make the 3rd, 2nd, and 1st shots in this order unless you could reverse the passage of time.

Hence to reduce the number of possible outcomes, we have to divide by the redundant counts, 3!, such that

$$\binom{5}{3} = \frac{5!}{(5-3)!(3)!}. \tag{2}$$

More generally, $\binom{N}{n}$ denotes choosing n out of all possible N outcomes, namely,

$$\binom{N}{n} = \frac{N!}{(N-n)!(n)!}. \tag{3}$$

Now we can put the first and second part of our calculations together to arrive at the probability, $B(5,3)$ that you will make 3 out of 5 shots. For this problem, let's consider the probability of making a shot is 0.2 for every shot. Hence, the total probability of you making 3 out of 5 free throws is

$$B(5,3) = \frac{5!}{(5-3)!(3)!}(0.2)^3(0.8)^2. \tag{4}$$

And, if you were curious to get probability values for various numbers of successes out of various number of trials, to help you in betting with your friends, then you would build a probability distribution using this general formula

$$B(N,n) = \frac{N!}{(N-n)!(n)!}p^n(1-p)^{N-n}. \tag{5}$$

0.1.2 Derivation of the Poisson Distribution From the Binomial

The Poisson distribution, $\boldsymbol{P(\lambda)}$ is a great tool for characterizing low probability events that occur in large sample sizes or in large number of trials. And it can be derived as an approximation of the binomial distribution, when N is large and p is small. Let's put p in terms of λ,

$$p = \frac{\lambda}{N}, \tag{6}$$

where λ is the average number of successes out of N trials.

Let's substitute $\frac{\lambda}{N}$ in place of p so that

$$\boldsymbol{P(\lambda)} = \frac{N(N-1)\cdots(N-n+1)}{n!}(\frac{\lambda}{N})^n(1-\frac{\lambda}{N})^{N-n}. \tag{7}$$

This is approximately

$$\boldsymbol{P(\lambda)} \approx \frac{\lambda^n}{n!}(1-\frac{\lambda}{N})^N, \tag{8}$$

if n is small relative to N. When N is large,

$$\boldsymbol{P(\lambda)} \approx \frac{\lambda^n}{n!}e^{-\lambda}. \tag{9}$$

Use

$$\lim_{N\to\infty}(1+\frac{1}{N})^N = e, \tag{10}$$

as the starting point for deriving the $e^{-\lambda}$ statement in the above equation

0.1.3 Expected Value

Let's revisit the basketball example. Based on many rounds of practice, you have determined the probabilities of making 0/5,1/5, 2/5, 3/5, 4/5, and 5/5 shots. The probabilities are 0, 0.15, 0.2, 0.25, 0.3 and 0.1, respectively. What is your expectation of the number of shots you will make on average in a practice session? We multiply each outcome by its probability and sum over all outcomes resulting in

$$0.0(0) + 0.15(1) + 0.2(2) + 0.25(3) + 0.3(4) + 0.1(5) = 3. \tag{11}$$

You should expect to make 3 out of 5 shots on average. Here, we're essentially weighting the contribution of each event by the probability of its occurrence to arrive at the expected value.

More generally, the **Expected Value** can be calculated using:

$$E(X) = \sum_{x \in \omega} x m(x) \tag{12}$$

where X is a discrete random variable found in the sample space ω with a probability distribution function $m(x)$. This equation applies only if the sum is convergent. How is the expected value different from the mean? The limit of the sample mean converges to the expected value as the sample size approaches ∞.

The expected value has a couple of interesting properties. For example, let's reconsider the basketball example, wherein every practice session constitutes 5 free-throws. Let's say that you now toss a die after every practice session and record both the number of shots you made during that practice and the number of points you got in the die toss. You add your score and the die score at the end of every practice session and develop an expected value for the sum of these two variables at the end of many many practice sessions. Alternatively you could simply add the expectation value of the number of basketball shots, which you already calculated to be 3, to the expected value of the die toss which can be calculated as follows: $1/6(1) + 1/6(2) + 1/6(3) + 1/6(4) + 1/6(5) + 1/6(6) = 3.5$. Hence, the expected value of the two scores is $3 + 3.5 = 6$. This is also true if you were to multiply the scores. This can be written more formally as

$$E(X + Y) = E(X) + E(Y), \tag{13}$$

and

$$E(X \cdot Y) = E(X) \cdot E(Y), \tag{14}$$

where Y is another random variable with a separate probability distribution. **Note:** these properties hold true only when X and Y are independent of each other. Another interesting property is that if you multiply your basketball score every time by a constant, c, and calculate the expected value for the multiplied scores, you will see that you have done more work than you had to. You could simply multiply your previous expected value by the constant.

$$E(cX) = cE(X) \tag{15}$$

This

$$E(\phi(X)) = \sum_{x \in \omega} \phi(X) m(X) \tag{16}$$

represents a very useful property of the expected value which we will use in derivations of the variance. Note that $\phi(X)$ represents a function of X.

For proof of these equations and more examples, click here.

0.1.4 Variance

Variance $V(X)$ is a measure of the deviation of a sample or many from a predicted, expected value. Let's say that you had 4 basketball practice sessions and in all practice sessions you only got 1/5 shots. How much does the result of this set of practice sessions deviate from the expected value of 3 shots per session?

$$V(X) = \sum \underbrace{(X-\mu)^2}_{\text{part 1}} \underbrace{m(X)}_{\text{part 2}} \tag{17}$$

Where μ is mean of a population and can be used as a proxy for the expected value when dealing with large sample sizes. Use Figure 1 as a visual aid to understand part 1 of the above equation. Note, the distance of each sample from the mean is squared because we want a positive value, and this is represented as the area of the squares you see in figure 1. Note, we have reduced the number of shots to two in the visual example, for simplicity.

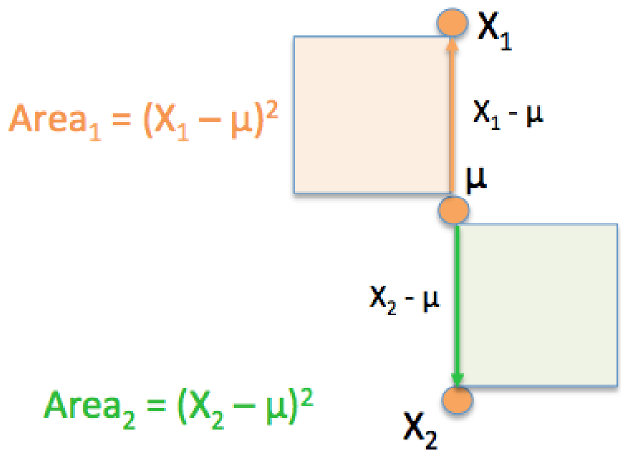

Figure 1: A visual representation of $(X-\mu)^2$. Orange dots X_1 and X_2 represent two samples, and the vectors represent their distances from the mean. The areas of squares represent $(X-\mu)^2$ for each sample.

You can use Figure 2, as a visual aid for the second part of the above equation, where the variance of each event is scaled by the probability of its occurrence. In the event where the probability of the events are equal, we simply divide the sum of the squares by the number of samples included in our variance calculations.

Going back to our basketball player example, here's how we can calculate the variance of the distribution:

$$V(X) = (0-3)^2(0)+(1-3)^2(0.15)+(2-3)^2(0.2)+(3-3)^2(0.25)+(4-3)^2(0.3)+(5-3)^2(0.1) \tag{18}$$

The variance also has interesting properties:

$$V(cX) = c^2V(X) \tag{19}$$

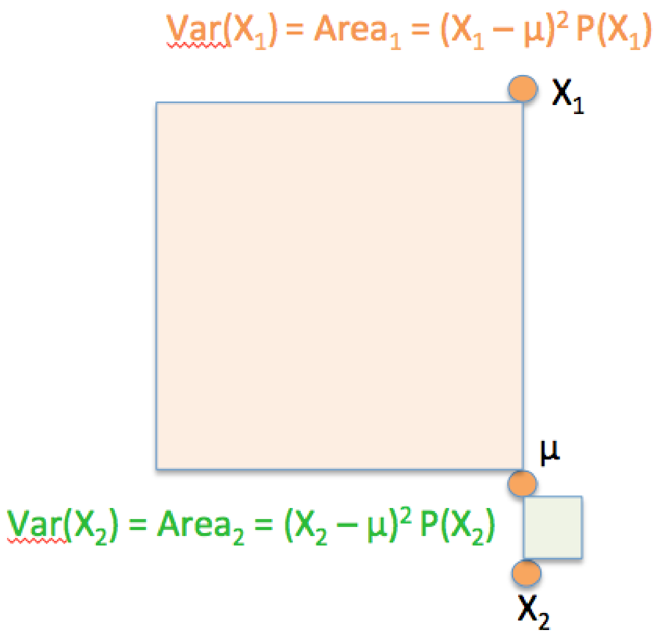

Figure 2: A visual representation of variance for samples X_1 and X_2. The squares from the previous figure have been scaled by the probability of each sample P_1 and P_2 which together comprise the probability distribution $m(x)$ in the variance equation.The collective variance of these two samples is the sum of the scaled areas of these two squares.

$$V(X + c) = V(X) \tag{20}$$

where c, is a constant.

Here is another property of variance,

$$V(X + Y) = V(X) + V(Y) \tag{21}$$

where Y is another random variable, and X and Y are independent. For proof of these properties, you can refer here.

Recall,

$$E(\phi(X)) = \sum_{x \in \omega} \phi(X)m(X) \tag{22}$$

where $\phi(X)$ represents a function of X. By this property of the expected value, we can rewrite our equation for variance, namely,

$$V(X) = \sum (X - \mu)^2 m(X) = E((X - \mu)^2). \tag{23}$$

We can further simplify the expression for variance to

$$V(X) = E((X - \mu)^2) = E(X^2 - 2\mu X + \mu^2). \tag{24}$$

In the case when dealing with large number of samples, E(X) = μ, the mean of the sample as previously discussed. Here you can see the many but same faces of variance:

$$V(X) = E(X^2) - 2\mu E(X) + \mu^2 = E(X^2) - \mu^2 = E(X^2) - [E(X)]^2. \tag{25}$$

Moreover, **Standard Deviation**, D(X), is simply $\sqrt{V(X)}$.

0.1.5 Expected Value and Variance of the Poisson Distribution

Recall the Poisson Distribution is $P(n, \lambda) = \frac{\lambda^n e^{-\lambda}}{n!}$. Now let's calculate it's expectation value according to $E(X) = \sum_{x \in \omega} x m(x)$. Let's get the expression for $E(X)$ in a form in which we can use the properties of e^x when it's expanded out as an infinite sum. Specifically, we have

$$E(X) = \sum_{n=0}^{\infty} nP(X = n, \lambda) = \sum_{n=1}^{\infty} n \frac{\lambda^n e^{-\lambda}}{n!} \tag{26}$$

$$= \sum_{n=1}^{\infty} \lambda \frac{\lambda^{(n-1)} e^{-\lambda}}{(n-1)!} \tag{27}$$

$$= \lambda \sum_{n=0}^{\infty} \frac{\lambda^n e^{-\lambda}}{n!}. \tag{28}$$

Now recall that

$$\sum_{n=0}^{\infty} \frac{\lambda^n}{n!} = e^n. \tag{29}$$

Hence, Equation 28 can be re-written as

$$E(X) = \lambda \sum_{n=0}^{\infty} e^{\lambda} e^{-\lambda}. \tag{30}$$

Finally, the expected value of a Poisson distribution is simply

$$E(X) = \lambda. \tag{31}$$

Now, let's derive the variance for a poisson distribution. Recall variance is

$$V(X) = \underbrace{E(X^2)}_{\text{part 1}} - \underbrace{[E(X)]^2}_{\text{part 2}}. \tag{32}$$

Let's calculate the first portion of this equation as we already know the second part of the equation based on our previous calculation of the expected value,

$$\underbrace{[E(X)]^2 = \lambda^2}_{\text{part 2}}. \tag{33}$$

$$\underbrace{E(X^2)}_{\text{part 1}} = \sum_{n=0}^{\infty} n^2 P(X = n, \lambda) = \sum_{n=1}^{\infty} n^2 \frac{\lambda^n e^{-\lambda}}{n!} \tag{34}$$

$$= \lambda \sum_{n=1}^{\infty} n \frac{\lambda^{n-1} e^{-\lambda}}{(n-1)!} \tag{35}$$

$$= \lambda \sum_{n=0}^{\infty} (n+1) \frac{\lambda^n e^{-\lambda}}{n!} \tag{36}$$

$$= \lambda \sum_{n=0}^{\infty} (n \frac{\lambda^n e^{-\lambda}}{n!} + \frac{\lambda^n e^{-\lambda}}{n!}) \tag{37}$$

$$= \lambda^2 + \lambda. \tag{38}$$

Thus, we have

$$V(X) = \underbrace{\lambda^2 + \lambda}_{\text{part 1}} - \underbrace{\lambda^2}_{\text{part 2}} = \boxed{\lambda}. \tag{39}$$

Hence, in a Poisson distribution, $\frac{\text{variance}}{\text{mean}} = 1$. This will become a critical concept for understanding the Luria-Delbruck paper.