Appendix

Ing. Viktor Matovič

July 30, 2023

Contents

0.1 Derivation of the Maximum Likelihood Function

The probability that the organizational pattern referenced in the text description of another organizational pattern would be applied as second in pattern sequence can be calculated with Bernoulli distribution. The maximum likelihood function, used to verify the organizational pattern's usefulness, is if based on the Bernoulli probability distribution, derived from the multiplication of partial functions modeling the probability of the application of the organizational pattern using the Maximum likelihood estimation method discussed in the work of Varga [Ste12] or Gareth [JWHT13]. Partial likelihood functions in Figure 1 form a chain by which it is possible to calculate the likelihood of the usefulness of the organizational pattern. The maximum likelihood function is a string of partial functions which are probability mass functions of the Bernoulli distribution. The decision to apply an organizational pattern can be modeled as a binary random variable *X:* $\{0, 1\}$ which turns value 1 when an organizational pattern is applied or 0 otherwise.

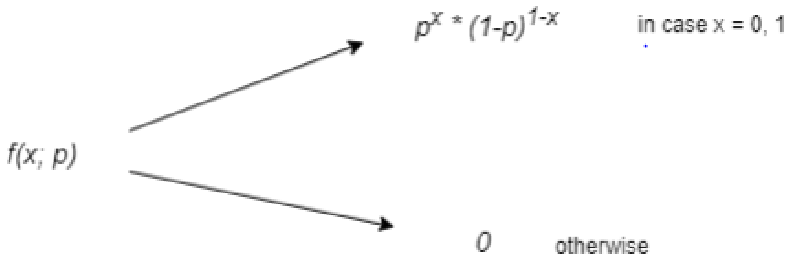

Figure 1: Maximum likelihood function is a string of the partial Bernoulli probability mass functions

X in Figure 1 is a random variable, whether the assumed domain expert chooses to apply the organizational pattern in the pattern sequence. The random variable X is any force conditioning application of the organizational pattern. Variable p is an unknown parameter that influences the application of the organizational pattern. The probability of the application of the organizational pattern is affected by the parameter p. The parameter p is a set of numerically represented context forces that the pattern is meant to resolve. The maximum likelihood function is, in general, for N organizational patterns documented in the pattern language in the form as can be seen in Figure 2:

$$L(p) = p^{X_1} * (1-p)^{1-X_1} \quad \ldots \quad p^{X_n} * (1-p)^{1-X_n} = \prod_{i=1}^{n} p^{X_i} * (1-p)^{1-X_i}$$

Figure 2: General form of the maximum likelihood function to estimate the likelihood of the application of pattern sequence

As in Figure 2, maximum likelihood functions are complex if the pattern tries to resolve multiple context forces. Derivation of the complex maximum likelihood function with numerous parameters and multiple functions is complicated. Because of this, approximation techniques are employed to calculate its first derivation. To calculate the maximum of the function of multiple variables as in Figure 2, the function from Figure 2 must be derived. This is problematic because the more products of partial functions need to be calculated, the more difficult derivation must be performed. In order to solve this problem, the function from Figure 2 must be transformed using the natural logarithm as in Figure 3 to derive the sum instead of the product of partial functions. This is possible because it can be shown that if an unbiased estimator *h(x)* maximizes likelihood *L(p)*, then it also maximizes the logarithm of the *ln L(p)* function.

$$\ln L(p) = \sum_{i=1}^{n} xi * \ln(p) + (n - \sum_{i=1}^{n} xi * \ln(1-p)$$

Figure 3: It is much easier to transform maximum likelihood function by logarithm to find its maximum

Equating this to 0 and solving for p results in derivation in Figure 4:

$$d \ln L(p) / dp = \sum_{i=1}^{n} xi / p - \sum_{i=1}^{n} xi / (1-p) = 0$$

Figure 4: Derivation of the maximum likelihood function to find its maximum

And after solving it, the best-unbiased estimator is the function *h(x)* in Figure 5:

$$h(x) = (1/n) \sum_{i=1}^{n} xi$$

Figure 5: Unbiased estimator of maximum likelihood

Using the unbiased estimator from Figure 5 as the Maximum likelihood function, the probability of the event the assumed domain expert would apply two organizational patterns can be calculated. This probability can be formally written as $P(chosen, chosen|h(x)) = h(x)^2$, which is also a maximum likelihood of such an event. This means the maximum likelihood that two organizational patterns would be applied in the expected pattern sequence is probability $P(applied, applied, unused\ |h(x)) = h(x)^2 * (1 - h(x))$. The value $h(x)^2 * (1 - h(x))$ is a likelihood that two organizational patterns out of 3 candidates would be applied in pattern sequence, each with a probability $h(x)$ with a value between 0.0 and 1.0.

The method to compute likelihood can be generalized and described by following two steps:

1. First step is to form a function that depends on a parameter Θ constraining the application of the patterns. The probability of the application of each pattern can be modeled with the Bernoulli distribution. The application of each organizational pattern or each pattern language can be modeled using different probability mass functions.

2. Second step dictates the derivation of the maximum likelihood function. A natural logarithm of the maximum likelihood function whose general form is $L(\theta) = f(x_1, \theta) * \ldots * f(x_n, \theta)$ is calculated as an approximation of the derivation of the maximum likelihood function. Function in this form must be derived, put equal to 0, and solved for the parameter to get an unbiased and most efficient estimator of the maximum likelihood of the application of the organizational pattern in the pattern sequence.

Derivation of the maximum likelihood, which maximizes the probability of application of the pattern in pattern sequence, is according to Kruschke [Kru15] frequentist approach. The Bayesian approach is used to convert the Maximum likelihood function to a valid probability distribution.

Bayes' rule is used to convert likelihood into a probability distribution of a parameter Θ. Probability that organizational pattern X would be applied is:

$$Pr(X = 1|\theta) = \theta \tag{1}$$

or

$$Pr(X = 0|\theta) = (1 - \theta) \tag{2}$$

otherwise. The value of the parameter Θ must be taken from a valid probability distribution. Random variable equals value α with a probability of $Pr(X = \alpha \ |\theta) = \theta^{\alpha} * (1-\theta)^{(1-\alpha)}$. Probability of application of organizational pattern is $Pr(X = 1|\theta) = \theta^1 * (1 - \theta)^{1-1} = \theta$ and probability of opposite is $Pr(X = 0|\theta) = \theta^0 * (1 - \theta)^{1-0} = (1 - \theta)$.

0.2 Probability of the Application of Pattern Sequence

The previous section 0.1 discussed only one situation that only one organizational pattern would be applied. Before the probability that multiple patterns would be applied in the expected pattern sequence can be calculated, according to Kruschke [Kru15] the following conditions must be satisfied:

- *Assumption of statistical independence* between applications of organizational patterns, sometimes in mathematical literature denoted as i.i.d.
- *Assumption of sample from one population*, which means the sample of descriptions of the organizational patterns must represent one complete set. This condition is satisfied if pattern language is applied.

The probability two organizational patterns would be applied in the pattern sequence is $Pr(X_1 = \alpha_1, X_2 = \alpha_2 \mid \Theta_1, \Theta_2) = Pr(X_1 = \alpha_1 \mid \Theta_1) * Pr(X_2 = \alpha_2 \mid \Theta_2) = (\Theta_1 * (1-\Theta_1)^{1-\alpha_1}) * (\Theta_2 * (1-\Theta_2)^{1-\alpha_2})$ in case that a different set of factors influences the applicability of each organizational pattern in a sequence, Θ_1 and Θ_2. In order to find the actual value of this probability, parameter Θ needs to be replaced with a function from Figure 5.

This probability $Pr(X_1 = \alpha_1, X_2 = \alpha_2 \mid \Theta_1, \Theta_2)$ can also be formally expressed for case that two patterns resolve the same contradicting forces. If contradicting forces to be resolved are the same $\Theta_1 = \Theta_2 = \Theta$ probability of application of two patterns in pattern sequence is $Pr(X_1 = \alpha_1, X_2 = \alpha_2 \mid \Theta) = \Theta^{\alpha_1+\alpha_2} * (1 - \Theta)^{2-\alpha_1-\alpha_2}$. The probability that two organizational patterns which resolve the same context forces would be applied in the expected pattern sequence is $Pr(X_1 = 1, X_2 = 1|\Theta) = \Theta^{1+1} * (1 - \Theta)^{2-1-1} = \Theta^2$.

Considering any pattern language P and any sequence of 2 organizational patterns probability that random variable X representing the number of patterns from language P applied in expected pattern sequence turns zero value decreases as probability of the existence of condition under which patterns from P could be applied increases, because $Pr(X = 0 \mid \Theta) = Pr(\neg P \mid \Theta) = Pr(pattern_1,\ pattern_2 \mid \Theta) = Pr(pattern_1|\Theta) * Pr(pattern_2\ |\Theta) = (1 - \Theta)^2$.

0.3 Usefulness of the Pattern Sequence

Using information from the previous two sections, Section 0.1 and Section 0.2, the likelihood of the usefulness of the organizational pattern can be calculated. Calculation of the likelihood of the usefulness of organizational pattern with maximum likelihood function can be presented using the simple example. Verification of the usefulness of the organizational pattern X_0, which mentions 4 other patterns in its textual description, can be calculated in a way that only the pattern being verified X_0 is useful and the patterns it links to $X_1..X_4$ describe its usefulness. It is probability of the random variable *"X = number of applications of organizational patterns"* turning zero value formally written as $Pr(X = 0 \mid \Theta_1, \Theta_2, \Theta_3, \Theta_4) = Pr(X_1, X_2.X_3, X_4 \mid \Theta_1, \Theta_2, \Theta_3, \Theta_4) = Pr(\neg X_1|\Theta_1) * Pr(\neg X_2|\Theta_2) * Pr(\neg X_3|\Theta_3) * Pr(\neg X_4|\Theta_4) = (1-\Theta_1) * (1-\Theta_2) * (1-\Theta_3) * (1-\Theta_4)$. The same applies if the maximum likelihood of the event 2 or 3 patterns are referenced by the pattern expected to be verified. The resulting output value is a probability of the usefulness of the organizational pattern.

Bibliography

[JWHT13] Gareth James, Daniela Witten, Trevor Hastie, and Robert Tibshirani. *An introduction to statistical learning*, volume 112. Springer, 2013.

[Kru15] John Kruschke. *Doing Bayesian Data Analysis (Second Edition)*. Academic Press, Boston, 2015.

[Ste12] Varga Stefan. *Matematická Štatistika*. Slovak Technical University, 2012.