II-2. Judgement in frequentism (How to use null hypothesis)

The author explained a basic philosophy of frequentism in II-1. In this philosophy, probability of obtaining actual data decrease with distance from the center of parental population assuming that actual data was obtained by random sampling from parental population. This means that possibility that the data obtained is not from the parental population increases with the distance. In this logic, there still remains the possibility that the data come from parental population, even the position of the obtained data is extremely far from the center. Of course, we have to accept possibility of misjudge, as we do not know the truth. However, we can reduce the possibility of misjudging and minimize the impact of misjudge. How we should make judgement? Logically, “different” is contrary concept of “similar”. Differentness and similarity are exclusive events each other. When similarity is proven, differentness is denied and when differentness is proven, similarity is denied. However, this is idealistic theory. This is not directly applicable to the real world. Because we identify phenomena by several limited factors. Judgements in real world is complicated and the allegorical story “how we identify crow” is often use as an example. The author cannot conceive any other example, and the story is as follow.

“Shape of birds are similar among species. We often use body color for identification of bird species particularly when the bird is far from us. When the body color is not black, we can judge the bird is not crow. However, we cannot judge the bird is crow, when the body color is black” That all. The story is finished.”

You may not understand what the allegorical story is saying. However, you can understand we cannot say “the bird is crow because of black body color”. There are many other black birds such as myna bird and jay.

Logical structure used here is as follows.

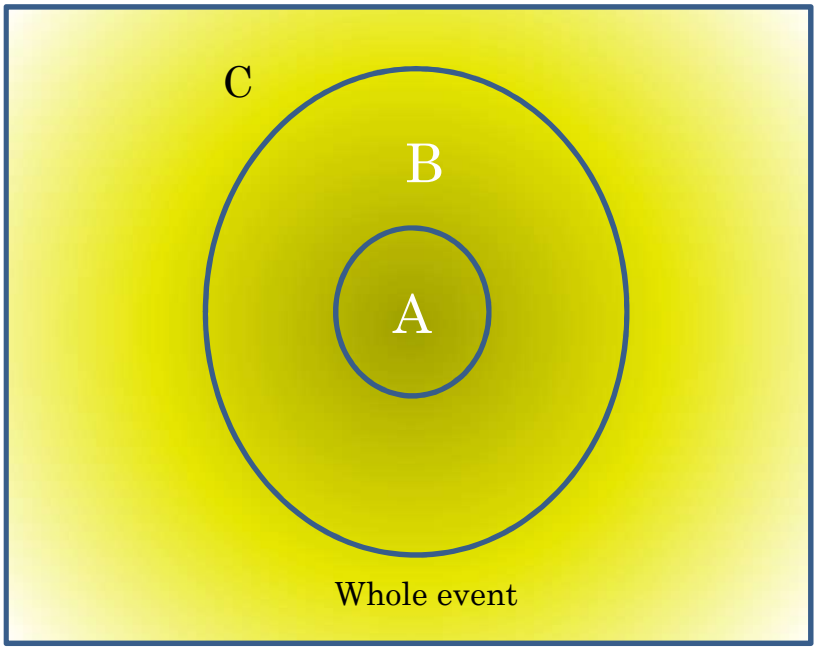

Fig.3. Relationship between necessary condition and sufficient condition (an expression in setting theory, $C \not\supset B, B \not\supset C, B \supseteq A$)

"A is B" is simple sentence. In many cases, those sentence is a description of character of B. We say "Crow is Black" when we want to say as a general description "Crow is generally black". In another case, these sentences are used as judgement or evaluation. In such case, the sentence can be rephrased more logically as "when α is A, then α is B". We also say "Crow is black, when want to say "When a bird is crow, the bier is black". When the sentence "A is B" is used as example of latter case, such statements are called "proposition in logic", and some propositions are true (affirmative preposition) and other preposition are false (negative proposition). Actually, proposition is a sentence that include reservation or condition in many cases. An example of such statement is "If α is B, then α is A". This statement can be rephrased to "Among various α, some particular α which is B in an element of character (factor) is identified as A". In the case of crow, "among all birds, bird which is back in body color is crow" is a proposition. Of course, this proposition is false generally, because there are many other black birds such as myna and jay. However, "proposition "among all birds, bird which is no black in body color is not crow" is correct, generally. In this case we need another reservation. That is "in the case if all craw is black". Actually, there are special crows that is not black. For example, albino or naughty crow which ate lunch of house painter and was

painted white by painter, hasty crow rushed into storehouse of flour shop and so on. Those examples are rare cases, and existence of white crow is a probability event. When we neglect those rare cases, we can say ”among all biers, bird which is no black in body color is not crow” or “If α is not B, then α is not A”. Figure 3 is representing the relation. All A is included in B. or all crow is included in black birds.

When we say that if α is A, α is necessarily B”, A is sufficient condition to B and B is necessary condition to A in logical relationship. Similarly, black is sufficient condition to crow and crow is necessary condition to black, when we say that if a bird is black, the bird is crow. Actually, this proposition is false, because there are many black birds other than crow and the all black bird cannot be included in crow. When we consider that A is “crow” and B is “black” in figure 3, in which A is included in B. This relation is correct. Because, crow is included in black birds in real world. So black is necessary condition to crow, and crow is sufficient condition to black, and proposition that if a bird is crow, the bird is black is correct.

When we consider a proposition “if α is A, α is B, (“if a bird is crow, the bird is black” or more simply “crow is black”)”, the statement “if α is B, α is A (“if a bird is black, the bird is crow” or more simply “black bird is crow”) “ is the converse of proposition. Correct and false are antinomy each other in the classic logic. In other words, all statements are should be correct or false in classic logic. On the contrary, there existence of vague intermediate category between correct and false is allowed in intuitionistic logic. In the classic logic, even if proposition is correct, converse of the preposition is not always correct. For example, even if we accept the statement “a bird is crow, the bird is black” (more simply “Crow is black”) as correct, we cannot accept the statement “a bird is black, the bird is crow” (More simply black bird is crow”) as correct proposition. Because of existence of jay and myna. Only in the caseA ≡ B (there is no black bird other than crow), we can say the converse is correct.

The statement “if α is not A, α is not B, (“if a bird is not crow, the bird is not black” or more simply “the birds other than crow is not black”)” is inverse of the proposition. Obviously, this statement is incorrect generally, because there are many black birds other than crow. Only in the caseA ≡ B (there is no black bird other than crow), we can say the inverse is correct.

Contraposition of proposition is a combination of converse and inverse. The statement “if α is not B, α is not A, (“if a bird is not black, the bird is not crow” or more simply “non-black birds is not crow”)” is contraposition of the proposition. If proposition is

correct the contraposition is correct, Inversely, if contraposition is correct the proposition is correct in classic logic. This is understandable when we look figure 3. A is completely included in B. From this figure, we can understand that "if α is A, α is B (in other words, if the data exists in region A, the data is exist in region B)"and "if α is not B, α is not A(if the data exists in region C, the data does not exist in region A)" are the same meaning. In other word we can understand that "crow is black" and "non-black bird is not crow" is same meaning.

We can use this relation, when we cannot proof correctness of proposition directly. In that case, we can proof the correctness by proofing the correctness of contraposition. However, there is a old theory in logic. We cannot proof correctness by empiric method. We can only proof incorrectness by empiric method. This can be explained as follows. We can deny the correctness by showing an example which does not fit into the proposition. We cannot proof the correctness of the proposition by showing illimitable example fitting in to the proposition, because we cannot deny the possibility of existence of an example which does not fit into the proposition. Proving of absence of possibility is an effort to find the evidence of absence. This is impossible for human being, because human being is finite being. This means that when we cannot confirm "α is A or not A (crow or not crow)" directly and the data exist in the region of "α is not B (not Black) ", we can say "α is not A (not crow).

The author is supposing that most of readers are feeling a frustration reading upper long rephrasing of similar sentences. However, several scientists often make mistakes by neglection of upper logic. They say that α is A, when they only confirm α exists in region B. Or, they say that the bird is crow when they see black bird. These statements are obviously mistakes, because every B is not A, though B includes A and every black bird is not crow though black birds include crow.

In statistics in frequentism, we discuss the similarity of parent population. We can say that two populations are different (more exactly not same) when we observed difference in data obtained from two batches of data population by placing the data in data distribution regions. However, we cannot say that two populations are same when we cannot observe difference in data obtained from two batches of data population even the two data drop same data distribution region.

Classic logic always orients structure of antimony, though real world always includes ambiguity. Introduction of the concept of probability is a possible method for

harmonization of real world and classic logic. When we neglect stochastic variation, average of two batch obtained from same parent population should be the same. And the difference between two averages should be 0. Actual random sampling always makes fluctuation of data, and we need to take account stochastic variation. When we consider the variation, center of the distribution of the difference should be 0 and the probability distribution is maximum at the center theoretically, if there is no difference between two batches of data. Cumulative provability increased with the distance from the center. When we fix a level of probability, we can draw a boundary point or line of the distribution of the data (threshold). If observed difference between the average of then both batch of data exists out of the point or line, we can say two batch of data do not come from the same population, because the possibility to observe such difference is less than the fixed level. In the allegorical story of crow, the probability can be comparable to the strength of color. There are various strengths of body color in crow. The crow standing on the center is real black and the body color becomes weaker depending the distance from the center. In the region where the distance from center is longer than fixed level by allowable error, we can say the body black color is too weak to say the bird is crow. This logic is used in t test in analysis of variance.

The other idea is to use ration of variance (F ratio). When we neglect stochastic variation, variances of two batch obtained from same parent population should be the same. And the ration between two variances should be 1. In this case 1 is not center of distribution of the ratio when we consider stochastic variation, though the probability is highest at 1, because the distribution is asymmetric. However, if we can make the distribution curve from sample size and variances, we can set a point or line for determination whether the distance is too long or not to long. This is the logic of F analysis.