

Notes on Qualitative Inference

CU Senior Hons

September 9, 2016

1 What is qualitative analysis?

It's surprisingly hard to pinpoint what the analytic difference between qualitative and quantitative approaches is.

- Number of cases?
- Use of numbers?
- Exploratory/inductive versus deductive?
- Interest in case effects versus general effects?
- Interest in what caused Y versus what's the effect of X?
- Subjective versus Objective?
- Descriptive versus Causal?
- Different notions of causality – INUS, equifinality, necessity and sufficiency? Deterministic versus stochastic?
- Difference in data sources, eg direct observation, interviews, archives?
- A difference in general orientation?
- Something else? What?

2 Causal inference for qualitative studies

2.1 The problem of causal inference

Consider, to begin, a situation in which some individuals in a diseased population are observed to have received a treatment while others have not (X). Assume that, subsequently, a researcher observes which individuals become healthy and which do not (Y). Let us further assume that each individual belongs to one of four unobserved “types,” defined by the potential effect of treatment on the individual:

- **adverse**: Those who would get better if and only if they do not receive the treatment
- **beneficial**: Those who would get better if and only if they do receive the treatment
- **chronic**: Those who will remain sick whether or not they receive treatment
- **destined**: Those who will get better whether or not they receive treatment

Throughout, we will use the letters a, b, c , and d to denote these causal types and the terms, $\lambda_a, \lambda_b, \lambda_c, \lambda_d$, to denote the relative share of these types in the population.

These types differ in their “potential outcomes” — that is on what outcomes, Y , they *would* take on under alternative treatment conditions, X . More formally, we let $Y(x)$ denote a case or type’s potential outcome when $X = x$. Thus, the potential outcomes are $Y(0) = 1, Y(1) = 0$ for type a ; $Y(0) = 0, Y(1) = 1$ for type b ; $Y(0) = 0, Y(1) = 0$ for type c ; and $Y(0) = 1, Y(1) = 1$ for type d . These potential outcomes are illustrated in Table 1.

	Type a adverse effects	Type b beneficial Effects	Type c chronic cases	Type d destined cases
Not treated	Healthy	Sick	Sick	Healthy
Treated	Sick	Healthy	Sick	Healthy

Table 1: Potential Outcomes: What would happen to each of four possible types of case if they were or were not treated.

2.2 What we know and what we do not know

The treatment effect for a case is defined as the difference in potential outcomes for that case between the treatment and control conditions, $Y(1) - Y(0)$. The *fundamental problem of causal inference* is that, for any given case, it is only possible to observe $Y(1)$ or $Y(0)$. Thus, for no case is it possible to observe the difference between these two quantities and, hence, the treatment effect. Put differently, it is impossible to directly observe the *type* of an individual case.

Table 2 displays the ambiguity that we face about the type of a case given an observation of X and Y for that case.

The key thing to note is that we know (or can measure) the value of X and the value of Y . That means that we know which cell we are in. But we still do not know the *type*.

	$Y = 0$	$Y = 1$
$X = 0$	$b \text{ or } c$	$a \text{ or } d$
$X = 1$	$a \text{ or } c$	$b \text{ or } d$

Table 2: The fundamental problem of type ambiguity

3 A Qualitative Strategy for Causal Inference

3.1 Strategy

The key strategy in process tracing is to look for “clues” that are indicative of the causal type. These clues are sometimes just called “observable implications” (especially when they come from a theory) or **CPOs**: Causal Process Observations (as distinct from DSOs – Data Set Observations). Note that clues may be indicative of the causal type but you may also have clues that are indicative of more complex propositions, such as the presence of a given mechanism etc.

A clue has *probative power* to distinguish between two hypotheses (say) when the probability of observing the clue is different under the two hypotheses.

Formally (or intuitively) you update given the clue using Bayes’s rule which says:

$$\Pr(H|\mathcal{D}) = \frac{\Pr(\mathcal{D}|H) \Pr(H)}{\Pr(\mathcal{D})} \quad (1)$$

This means that the probability of the Hypothesis given the data is a function of: the prior probability of the hypothesis times the probability of seeing the data if the hypothesis is true, divided by the total probability of seeing the data.

An example of Bayes rule: I am interested in whether a coin is a fair coin (which comes up heads or tails with equal probability) or a biased coin (which comes up heads with a 2/3 probability). I think these two are equally likely. I toss the coin twice and it comes up heads twice! What should I infer about the probability that it is a fair coin?¹

In this case observing the two heads is a clue (really just data) and you want to use it to draw inferences.

The inferences you can draw from any clue depend hugely on how likely you are to observe the clue under different conditions. Different combinations of probabilities give rise to four different types of qualitative “tests” (tests in quotes because this logic is useful even if you are not actually testing): the *hoop test*, the *smoking gun test*, the *straw in the wind test*, and the *doubly decisive test*.

These tests and the kinds of inferences you can make from them are illustrated in the next two figures.

3.2 Questions

Make sure you can answer all these questions:

1. Say you look at a case with $X = 1$ and $Y = 1$: what is this unit’s causal type?

1

$$93.0 = 57/6 = \frac{\frac{5}{1} \times \frac{6}{1} + \frac{2}{1} \times \frac{1}{1}}{\frac{5}{1} \times \frac{1}{1}} \quad \text{Answer:}$$

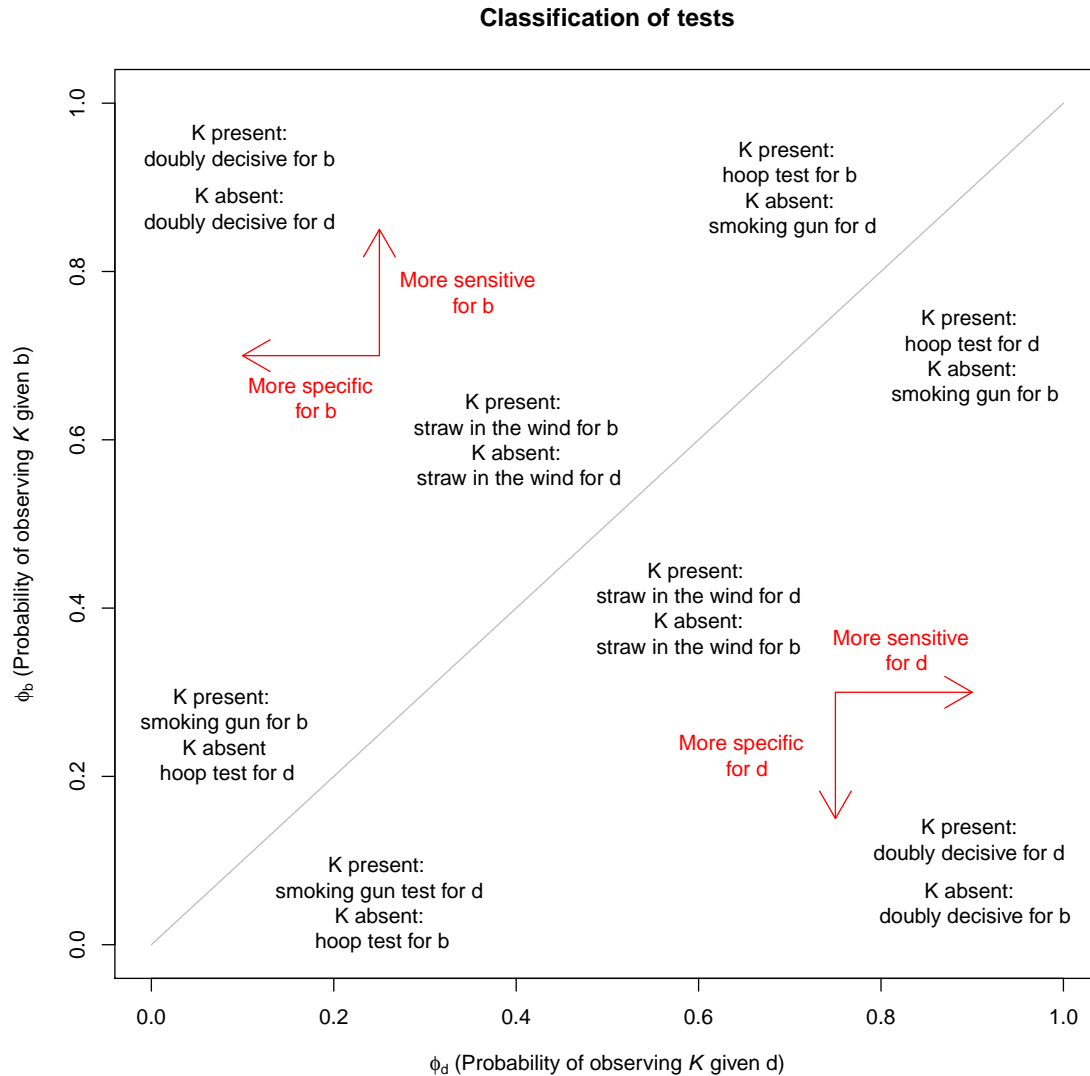


Figure 1: A mapping from the probability of observing a clue if a proposition " b " to a generalization of the classic qualitative tests.

2. Say you look at a case with $X = 1$ and $Y = 0$: what is this unit's causal type?
3. Give an example of a smoking gun test
4. Explain how you know that this is a smoking gun test
5. Give an example of a hoop test
6. Explain how you know that this is a hoop test

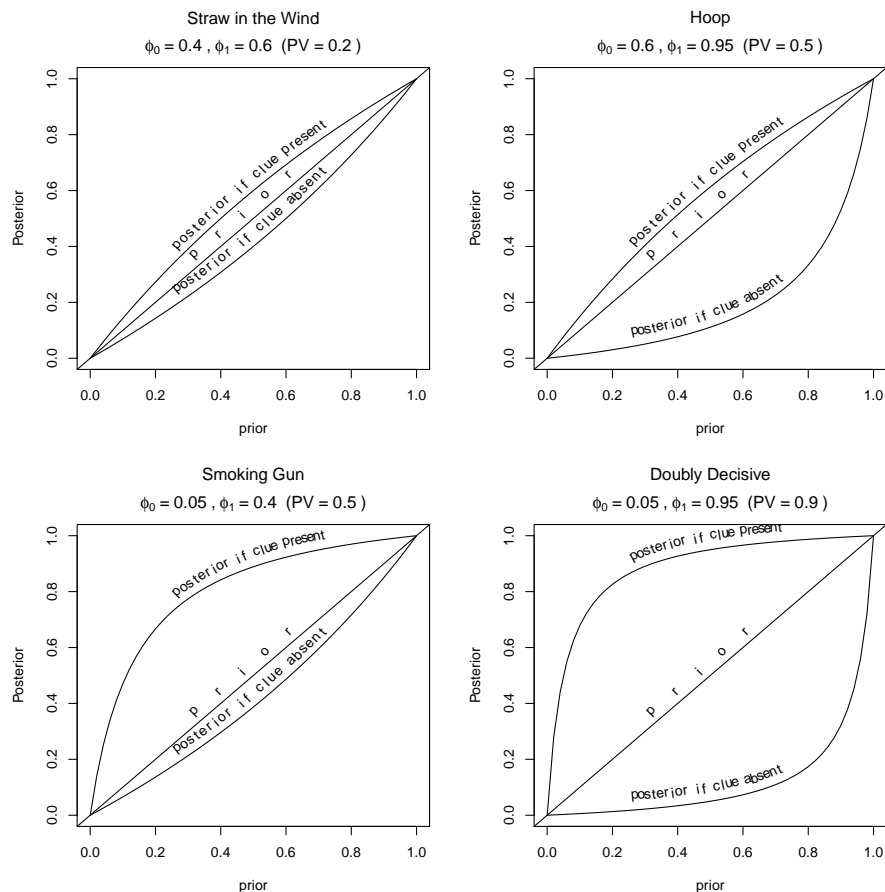


Figure 2: Figure shows how the learning from different types of tests depends on priors regarding the proposition. A “smoking gun” test has the greatest impact on beliefs when priors are middling low and the clue is observed; a “hoop test” has the greatest effect when priors are middling high and the clue is not observed. Here “PV” denotes the probative value of the test.

4 In-class Exercise

Here are the same four arguments from yesterday. For each argument you should:

1. identify the **overall causal relationship** of interest
2. propose a smoking gun OR a hoop test for the relationship
3. give an argument for why it is a smoking gun (or hoop) test
4. identify a key **mechanism** of interest (or, better, pair of rival mechanisms)
5. propose a smoking gun OR a hoop test for the mechanism
6. give an argument for why it is a smoking gun (or hoop) test

7. describe what sort of case would be the best one to choose in order to assess the overall causal relationship (eg $X = 1, Y = 1$)?

4.1 Natural resources and conflict

In developing countries that discover natural resources, such as oil, the ruling elite can extract wealth without needing to tax citizens and develop the state apparatus. Because the state does not rely on taxation for government revenue, it does not need to set up accountability structures or extend its reach and citizens do not feel that they have ownership over the state. The state therefore becomes both less democratic and weaker than if it had not discovered the resources.

4.2 Democracy and growth

Rich countries are more likely to be democratic for the simple reason that when people become wealthier they refuse to be dictated to by others and they demand a role in government. The marginal effects of income increases are greater for poorer countries because the impacts on education are greatest at these levels. You can test this proposition by exploiting natural variation in commodity prices which provide shocks to national income, especially for countries dependent on primary commodity exports.

4.3 Factor Endowments and Coalitions

When countries increase trade (imports and exports), the returns to economic factors (such as labor, land and capital) are affected differently. Specifically, the returns to factors that are the most *abundant* are positive, while the returns to factors that are the most *scarce* are negative. Therefore, the relative factor endowments of a country will predict what sort of political coalitions will form (eg Land versus Labor + Capital) and which groups will favor free trade policies.

4.4 Democratic peace

In democratic states, leaders are accountable for any losses incurred as a result of the wars that they enter into. Two states with democratic leaders are also more likely to share a common set of norms, and to engage in trade with one another. Therefore, two democracies are far less likely to enter into war with one another than a democracy and a non-democracy, or two non-democracies.

5 Case selection

5.1 Selection on the dependent variable

eg to understand why there are revolutions you have to look at those cases with revolutions. Concern though is that you may be wrong thinking that X is a cause of Y if X is also present in those cases where Y is not. More generally censoring on Y can lead to bias. When might it make sense?

5.2 Random case selection

What advantages and disadvantages?

5.3 Typical cases?

Typical of model or typical of data? Say $X = (1, 2, 3, 100)$, $Y = c(1,2,3,0)$. What case is typical?

5.4 Diverse cases?

To represent the distribution of the support of the data?

5.5 Extreme cases?

Extreme in terms of data, Hard cases and bad laws?

5.6 Deviant cases?

Extreme in terms of model error. Hard cases and bad laws?

5.7 Influential?

Is this really deviant?

5.8 Most similar? Most Different

Similar **except** on var of interest. Different **except** on vars of interest.

Question: What is the implicit inference strategy? What do you *do* with the cases?

5.9 Easy cases? Hard cases?

Sometimes people describe choosing “most likely” and “least likely” cases. For the intuition consider the claim that a drug is effective.

A most likely case might examine a case in which the subject is very diligent about taking the drug and does not have unusual health complications: if it works anywhere it should work here; if it does not work here, that is bad news for everywhere...

A least likely case might examine a case in which the subject is not very diligent about taking the drug and has many unusual health complications: if it works here it will work anywhere; if it does not work here, we don't learn too much.

What's a political science example of a most likely or least likely case?

Note that “least likely” should be interpreted not as the least likely case in which the basic theory is correct but the hardest case to find evidence for it. In a sense if the theory holds when it should not, that is bad news for the theory! If the theory explains even though it's set up to fail, then good for it.

6 Recap on Identification Strategies

There are ten strategies for figuring out causal effects. Each require different assumptions to make them fly.

1. Randomization
2. Experimental Control (induced unit homogeneity)
3. Natural experiments (as-if randomization)
4. Before / after comparisons
5. Ex Post Controlling: Regression, Matching and Weighting
6. Synthetic Matching
7. Instrumental variables (IV)
8. Regression discontinuity designs (RDD)
9. Process tracing
10. Front Door Strategies (Argument from mechanisms)

More on each of these here:

<http://egap.org/methods-guides/10-strategies-figuring-out-if-x-caused-y>