WEBVTT

00:00:00.180 --> 00:00:03.110 - Seminar, so hello everyone.

00:00:03.110 --> 00:00:05.350 My name is Qingyuan Zhao,

00:00:05.350 --> 00:00:10.170 I'm currently a University Lecturer in Statistics

00:00:10.170 --> 00:00:11.853 in University of Cambridge.

00:00:13.020 --> 00:00:15.440 I visited Yale Biostats,

00:00:15.440 --> 00:00:19.153 briefly last year in February.

00:00:21.340 --> 00:00:26.340 And so it's nice to see every guest very shortly this time.

00:00:28.380 --> 00:00:30.280 And today I'll talk

00:00:30.280 --> 00:00:33.939 about sensitivity analysis for observational studies,

00:00:33.939 --> 00:00:37.040 looking back and moving forward.

00:00:37.040 --> 00:00:39.070 So this is based on ongoing work

00:00:39.070 --> 00:00:44.070 with several people Bo Zhang, Ting Ye and Dylan Small

00:00:44.790 --> 00:00:46.480 at University of Pennsylvania,

00:00:46.480 --> 00:00:49.483 and also Joe Hogan at Brown University.

00:00:52.280 --> 00:00:57.280 So sensitivity analysis is really a very broad term

00:00:58.880 --> 00:01:01.890 and you can find in almost any area

00:01:01.890 --> 00:01:04.353 that uses mathematical models.

00:01:05.680 --> 00:01:07.530 So, broadly speaking,

00:01:07.530 --> 00:01:12.380 what it tries to do is it studies how the uncertainty

00:01:12.380 --> 00:01:17.010 in the input of a mathematical model or system,

00:01:17.010 --> 00:01:19.890 numerical or otherwise can be apportioned

00:01:19.890 --> 00:01:23.860 to different sources of uncertainty in it's input.

00:01:23.860 --> 00:01:26.810 So it's an extremely broad concept.

00:01:26.810 --> 00:01:30.380 And you can even fit statistics as part

00:01:30.380 --> 00:01:33.703 of a sensitivity analysis in some sense.

00:01:34.840 --> 00:01:39.533 But here, there can be a lot of kinds of model inputs.

00:01:40.691 --> 00:01:43.400 So, in particular, 00:01:43.400 --> 00:01:47.010 it can be any factor that can be changed in a model

00:01:47.010 --> 00:01:49.183 prior to its execution.

00:01:50.310 --> 00:01:53.320 So one example is structural

00:01:53.320 --> 00:01:57.350 or epistemic sources of uncertainty.

00:01:57.350 --> 00:02:00.680 And this is sort of the things we'll talk about.

00:02:00.680 --> 00:02:03.490 So basically, what our talk about today

00:02:03.490 --> 00:02:06.870 is those things that we don't really know.

00:02:06.870 --> 00:02:08.810 I mean, we made a lot of assumptions

00:02:08.810 --> 00:02:12.970 about when proposing such a model.

00:02:12.970 --> 00:02:16.337 So in the context of observational studies,

00:02:16.337 --> 00:02:20.260 a very common and typical question

00:02:20.260 --> 00:02:23.710 that requires sensitivity analysis is the following.

00:02:23.710 --> 00:02:28.680 How do the qualitative and or the quantitative conclusions

00:02:28.680 --> 00:02:30.652 of the observational study change

00:02:30.652 --> 00:02:34.930 if the no unmeasured confounding assumption is violated?

00:02:34.930 --> 00:02:38.710 So this is really common because essentially,

00:02:38.710 --> 00:02:41.910 in the vast majority of observational studies,

00:02:41.910 --> 00:02:44.610 it's essential to assume this

00:02:44.610 --> 00:02:46.850 no unmeasured confounding assumption,

00:02:46.850 --> 00:02:50.270 and this is an assumption that we cannot test

00:02:50.270 --> 00:02:51.880 with empirical data,

00:02:51.880 --> 00:02:54.213 at least with just observational data.

00:02:55.360 --> 00:02:58.500 So any, if you do any observational studies,

00:02:58.500 --> 00:03:01.750 so you're almost bound to be asked this question

00:03:01.750 --> 00:03:04.023 that, what if this assumption doesn't hold?

00:03:06.051 --> 00:03:08.140 And I'd like to point out that this question

00:03:08.140 --> 00:03:11.650 is fundamentally connected to missing not at random

00:03:11.650 --> 00:03:13.890 in the missing data literature.

00:03:13.890 --> 00:03:16.010 So what I will do today is I'll focus

00:03:16.010 --> 00:03:19.860 on sensitivity analysis for observational studies,

00:03:19.860 --> 00:03:21.860 but a lot of the ideas are drawn

00:03:21.860 --> 00:03:24.380 from the missing data literature.

00:03:24.380 --> 00:03:27.500 And most of the ideas that I'll talk about

00:03:27.500 --> 00:03:30.140 today can be also applied there

00:03:30.140 --> 00:03:32.083 and to related problems as well.

00:03:34.970 --> 00:03:39.970 So, currently, a state of the art of sensitivity analysis

00:03:40.220 --> 00:03:43.400 for observational studies is the following.

00:03:43.400 --> 00:03:47.440 There are many, many masters gazillions of methods

00:03:47.440 --> 00:03:50.490 of exaggeration, but certainly many many methods

00:03:50.490 --> 00:03:54.140 that are specifically designed for different

00:03:54.140 --> 00:03:56.193 kinds of sensitivity analysis.

00:03:57.570 --> 00:04:02.570 It often also depends on how you analyze your data

00:04:02.580 --> 00:04:04.823 under unmeasured confounding assumption.

00:04:06.080 --> 00:04:08.810 There are various forms of statistical guarantees

00:04:08.810 --> 00:04:10.073 that have been proposed.

00:04:11.120 --> 00:04:15.320 And oftentimes, these methods are not always

00:04:15.320 --> 00:04:17.350 straightforward to interpret,

00:04:17.350 --> 00:04:20.470 at least for inexperienced researchers,

00:04:20.470 --> 00:04:23.623 it can be quite complicated and confusing.

00:04:25.950 --> 00:04:29.903 The goal of this talk is to give you a high level overview.

00:04:30.860 --> 00:04:33.860 So this is not a talk where I'm gonna unveil

00:04:33.860 --> 00:04:35.770 a lot of new methods.

00:04:35.770 --> 00:04:39.660 This is more of an overview kind of talk

00:04:39.660 --> 00:04:42.230 that just to try to go through

00:04:42.230 --> 00:04:46.160 some of the main ideas in this area.

00:04:46.160 --> 00:04:47.150 So in particular,

00:04:47.150 --> 00:04:51.880 what I wanted to address is the following two questions.

00:04:51.880 --> 00:04:54.090 What is the common structure behind

00:04:54.090 --> 00:04:57.300 all these sensitivity analysis methods?

00:04:57.300 --> 00:05:01.760 And what are some good principles and ideas we should follow

00:05:01.760 --> 00:05:05.790 and perhaps extend when we have similar problems?

00:05:05.790 --> 00:05:10.230 The perspective of this talk will be global and frequentist.

00:05:10.230 --> 00:05:11.990 By that, I mean,

00:05:11.990 --> 00:05:13.750 there's an area in sensitivity analysis

00:05:13.750 --> 00:05:15.520 called local sensitivity analysis,

00:05:15.520 --> 00:05:18.674 where you're only allowed to move your parameter

00:05:18.674 --> 00:05:23.513 near its maximum likelihood estimate, usually.

00:05:24.500 --> 00:05:29.190 But global sensitivity analysis refer to the method

00:05:29.190 --> 00:05:31.470 that you can model your sensitivity parameter

00:05:31.470 --> 00:05:33.603 freely in a space.

00:05:34.700 --> 00:05:36.913 So that's what we'll focus on today.

00:05:37.900 --> 00:05:40.360 And also, I'll take a frequentist perspective.

00:05:40.360 --> 00:05:43.093 So I won't talk about Bayesian sensitivity analysis,

00:05:43.938 --> 00:05:45.820 which is also a big area.

00:05:45.820 --> 00:05:48.880 And I'll use this portal typical setup

00:05:49.950 --> 00:05:51.950 in observational studies,

00:05:51.950 --> 00:05:55.870 where you have iid copies of these observed data O,

00:05:55.870 --> 00:06:00.250 which has three parts, x is the covariance,

00:06:00.250 --> 00:06:04.300 A the binary treatment, Y is the outcome

00:06:04.300 --> 00:06:06.350 and these observed observed data

00:06:06.350 --> 00:06:10.480 that come from underlying full data, F,

00:06:10.480 --> 00:06:12.770 which includes X and A

00:06:12.770 --> 00:06:15.547 and the potential outcomes, Y(0) and Y(1).

00:06:16.910 --> 00:06:17.973 Okay, so this is,

00:06:19.474 --> 00:06:21.490 if you haven't, if most of you probably have seen this

00:06:21.490 --> 00:06:23.700 many, many times already,

00:06:23.700 --> 00:06:25.481 but if you haven't seen that this

00:06:25.481 --> 00:06:28.521 is the most typical setup in observational studies.

00:06:28.521 --> 00:06:30.383 And it kind of gets a little bit boring

00:06:30.383 --> 00:06:31.610 when you see it so many times.

00:06:31.610 --> 00:06:33.150 But what we're trying to do

00:06:34.204 --> 00:06:37.080 is to use this as the simplest example,

00:06:37.080 --> 00:06:41.010 to demonstrate the structure and ideas.

00:06:41.010 --> 00:06:46.010 And hopefully, if you understand these good ideas,

00:06:46.060 --> 00:06:49.780 you can apply them to your problems

00:06:49.780 --> 00:06:52.793 that are maybe slightly more complicated than this.

00:06:54.930 --> 00:06:56.758 So here's the outline

00:06:56.758 --> 00:06:58.500 and I'll give a motivating example

00:06:58.500 --> 00:07:01.260 then I'll talk about three components

00:07:01.260 --> 00:07:02.850 in the sensitivity analysis.

00:07:02.850 --> 00:07:04.330 There the sensitivity model,

00:07:04.330 --> 00:07:07.633 the statistical inference and the interpretation.

00:07:09.530 --> 00:07:13.330 So the motivating example will sort of demonstrate

00:07:13.330 --> 00:07:16.240 where these three components come from.

00:07:16.240 --> 00:07:20.750 So this example is in the social sciences actually

00:07:20.750 --> 00:07:22.943 it's about child soldiering,

00:07:23.930 --> 00:07:28.930 a paper by Blattman and Annan, 2010.

00:07:29.540 --> 00:07:34.018 On the review of economics and statistics,

00:07:34.018 --> 00:07:39.018 so what they studied is this period of time in Uganda,

00:07:41.320 --> 00:07:43.572 from 1995 to 2004,

00:07:43.572 --> 00:07:45.656 where there was a civil war

00:07:45.656 --> 00:07:49.092 and about 60,000 to 80,000 youth

00:07:49.092 --> 00:07:52.223 were abducted by a rebel force.

00:07:53.120 --> 00:07:54.410 So the question is,

00:07:54.410 --> 00:07:57.980 what is the impact of child soldiering

00:07:57.980 --> 00:08:00.453 sort of this abduction by the rebel force,

00:08:01.380 --> 00:08:04.370 as on various outcomes,

00:08:04.370 --> 00:08:07.820 such as years of education,

00:08:07.820 --> 00:08:11.663 and in this paper to actually study the number of outcomes.

00:08:12.740 --> 00:08:16.590 The authors controlled for a variety of baseline covariates,

00:08:16.590 --> 00:08:19.640 like the children's age, their household size,

00:08:19.640 --> 00:08:22.013 their parental education, et cetera.

00:08:23.210 --> 00:08:25.710 They were quite concerned about

00:08:25.710 --> 00:08:28.480 this possible unmeasured confounder.

00:08:28.480 --> 00:08:32.890 That is the child's ability to hide from the rebel.

00:08:32.890 --> 00:08:37.890 So it's possible that maybe if this child is smart,

00:08:38.620 --> 00:08:41.230 and if he knows that he or she knows

00:08:41.230 --> 00:08:44.010 how to hide from the rebel,

00:08:44.010 --> 00:08:48.610 then he's less likely to be abducted

00:08:48.610 --> 00:08:50.543 to be in this data set.

00:08:51.620 --> 00:08:54.680 And he'll probably also be more likely

00:08:54.680 --> 00:08:58.210 to receive longer education just because maybe

00:09:00.331 --> 00:09:04.023 the skin is a bit more small, let's say.

00:09:05.710 --> 00:09:07.190 So in their analysis,

00:09:07.190 --> 00:09:10.880 they follow the model proposed by Imbens,

00:09:10.880 --> 00:09:12.430 which is the following.

00:09:12.430 --> 00:09:17.430 So basically, they assume this no unmeasured confounding

00:09:18.120 --> 00:09:21.273 after you conditional on this unmeasured confounder U.

00:09:22.480 --> 00:09:24.291 Okay, so X are all covariates

00:09:24.291 --> 00:09:25.233 that U controlled for,

00:09:26.197 --> 00:09:30.410 and U is they assumed is a binary, unmeasured confounder.

00:09:31.840 --> 00:09:34.513 That's just a coin flip.

00:09:35.800 --> 00:09:39.000 And then they assume the logistic model

00:09:39.000 --> 00:09:44.000 for the probability of being abducted

00:09:44.120 --> 00:09:49.120 and the normal linear model for the potential outcomes.

00:09:49.410 --> 00:09:54.410 So notice that here the linear these terms

00:09:55.220 --> 00:09:57.830 depends on not only the observed covariance,

00:09:57.830 --> 00:10:00.920 but also the unmeasured covariates U.

00:10:00.920 --> 00:10:02.440 And of course,

00:10:02.440 --> 00:10:03.910 we don't measure this U.

00:10:03.910 --> 00:10:08.003 So we cannot directly fit these models.

00:10:08.920 --> 00:10:12.040 But what they did is they because they made

00:10:12.040 --> 00:10:16.080 some distribution assumptions on U,

00:10:16.080 --> 00:10:19.100 you can treat U as unmeasured variable.

00:10:19.100 --> 00:10:21.010 And then, for example,

00:10:21.010 --> 00:10:23.873 fit maximum likelihood estimate.

00:10:25.470 --> 00:10:29.397 So they're treated this two parameters lambda and delta,

00:10:29.397 --> 00:10:30.993 as sensitivity parameters.

00:10:31.980 --> 00:10:34.970 So these are the parameters that you vary

00:10:34.970 --> 00:10:37.260 in a sensitivity analysis.

00:10:37.260 --> 00:10:39.220 So when they're both equal to zero,

00:10:39.220 --> 00:10:42.837 that means that there is no unmeasured confounding.

00:10:42.837 --> 00:10:45.810 So you can actually just ignore this confounder U.

00:10:45.810 --> 00:10:48.380 So it corresponds to your primary analysis,

00:10:48.380 --> 00:10:49.810 but in a sensitivity analysis,

00:10:49.810 --> 00:10:52.580 you change the values of lambda and U

00:10:52.580 --> 00:10:55.330 and you see how that changes your result

00:10:55.330 --> 00:10:57.030 above this parameter beta,

00:10:57.030 --> 00:10:59.783 which is interpreted as a causal effect.

00:11:01.540 --> 00:11:06.000 Okay, so the results can be summarized in this one slide.

00:11:06.000 --> 00:11:07.940 I mean they've done a lot more definitely.

00:11:07.940 --> 00:11:11.650 But for the purpose of this talk, basically,

00:11:11.650 --> 00:11:14.760 what they found is that the primary analysis

00:11:14.760 --> 00:11:17.443 found that the average treatment effect is -0.76.

00:11:18.600 --> 00:11:21.270 So remember the outcome was years of education.

00:11:21.270 --> 00:11:23.460 So being abducted,

00:11:23.460 --> 00:11:28.297 has a significant negative effect on education.

00:11:30.150 --> 00:11:32.160 And then it did a sensitivity analysis,

00:11:32.160 --> 00:11:35.740 which can be summarized in this calibration plot.

00:11:35.740 --> 00:11:39.710 What is shown here is that these two axis

00:11:39.710 --> 00:11:43.010 are basically the two sensitivity parameters,

00:11:43.010 --> 00:11:44.890 lambda and delta.

00:11:44.890 --> 00:11:48.350 So what the paper did is they transform it

00:11:48.350 --> 00:11:50.423 to the increase in R-squared.

00:11:51.310 --> 00:11:54.853 But that's that can be mapped to lambda and delta,

00:11:55.720 --> 00:11:57.280 and then they compared

00:11:59.040 --> 00:12:01.780 this curve, so this dashed curve

00:12:03.384 --> 00:12:06.970 is where the values of lambda and delta such that

00:12:06.970 --> 00:12:09.773 the treatment in fact is reduced by half.

00:12:10.700 --> 00:12:12.640 And then they compare this curve

00:12:12.640 --> 00:12:15.080 with all the measured confounders,

00:12:15.080 --> 00:12:16.760 like year and a location,

00:12:16.760 --> 00:12:20.323 location of birth, year of birth, et cetera.

00:12:21.348 --> 00:12:25.020 And then you compare it with the corresponding coefficients

00:12:25.020 --> 00:12:30.020 of those variables in the model

00:12:30.800 --> 00:12:35.800 and then they just plot these in the same figure.

00:12:36.580 --> 00:12:39.260 What is supposed to show is that look,

00:12:39.260 --> 00:12:42.360 this is the point where the treatment effect

00:12:42.360 --> 00:12:44.100 is reduced by half,

00:12:44.100 --> 00:12:47.120 and this is about the same strength

00:12:47.120 --> 00:12:50.093 as location or birth alone.

00:12:50.093 --> 00:12:53.730 So, if you think your unmeasured confounder is in some sense

00:12:53.730 --> 00:12:57.950 as strong as the location or the year of birth,

00:12:57.950 --> 00:13:00.590 then it is possible that the treatment infact,

00:13:00.590 --> 00:13:03.993 is half of what it is estimated to be.

00:13:05.020 --> 00:13:07.760 Okay, so it's a pretty neat way

00:13:07.760 --> 00:13:10.503 to present a sensitivity analysis.

00:13:12.220 --> 00:13:13.550 So in this example, you see,

00:13:13.550 --> 00:13:16.730 there's three components of sensitivity analysis.

00:13:16.730 --> 00:13:19.137 First is model augmentation.

00:13:19.137 --> 00:13:23.510 And you need to expand the model used by primary analysis

00:13:23.510 --> 00:13:26.230 to allow for unmeasured confounding.

00:13:26.230 --> 00:13:29.620 Second, you need to do statistical inference.

00:13:29.620 --> 00:13:31.840 So you vary the sensitivity parameter,

00:13:31.840 --> 00:13:33.340 estimate the effect,

00:13:33.340 --> 00:13:36.490 and then control some statistical errors.

00:13:36.490 --> 00:13:38.116 So what they did

00:13:38.116 --> 00:13:42.260 is, it's they essentially varied lambda and delta,

00:13:42.260 --> 00:13:45.220 and they estimated the average treatment effect

00:13:45.220 --> 00:13:46.813 under that lambda and delta.

00:13:48.510 --> 00:13:52.160 And the third component is to interpret the results.

00:13:52.160 --> 00:13:55.930 So this paper relied on that calibration plot

00:13:55.930 --> 00:13:57.830 for that purpose.

00:13:57.830 --> 00:14:00.630 But this is often quite a tricky

00:14:00.630 --> 00:14:03.800 because the sensitivity analysis is complicated

00:14:04.860 --> 00:14:07.400 as we need to probe different directions

00:14:07.400 --> 00:14:09.090 of unmeasured confounding.

00:14:09.090 --> 00:14:13.540 So the interpretation is actually not always straightforward

00:14:13.540 --> 00:14:17.173 and sometimes can be quite complicated.

00:14:19.230 --> 00:14:23.350 There did you have there do exist two issues

00:14:23.350 --> 00:14:24.813 with this analysis.

00:14:25.790 --> 00:14:29.630 So this is just the model and rewriting it.

00:14:29.630 --> 00:14:33.350 The first issue is that actually the sensitivity parameters

00:14:33.350 --> 00:14:34.490 lambda and Dota,

00:14:34.490 --> 00:14:37.740 where we vary in a sensitivity analysis

00:14:37.740 --> 00:14:40.740 are identifiable from the observed data.

00:14:40.740 --> 00:14:44.320 This is because this is a perfect parametric model.

00:14:44.320 --> 00:14:47.380 And then it's not constructed in any way

00:14:47.380 --> 00:14:51.490 so that these lambda and delta are not identifiable.

00:14:51.490 --> 00:14:52.690 In fact, in the next slide,

00:14:52.690 --> 00:14:55.170 I'm going to show you some empirical evidence

00:14:55.170 --> 00:14:58.890 that you can actually estimate these two parameters.

00:14:58.890 --> 00:15:02.160 So, logically it is inconsistent for us

00:15:02.160 --> 00:15:04.630 to vary the sensitivity parameter.

00:15:04.630 --> 00:15:07.317 Because if we truly believe in this model

00:15:07.317 --> 00:15:08.960 and the data actually tell us what the values

00:15:08.960 --> 00:15:10.110 of lambda and delta is.

00:15:11.010 --> 00:15:12.850 So this is the similar criticism

00:15:12.850 --> 00:15:17.850 that for Hattman selection model, for example.

00:15:20.010 --> 00:15:22.590 The second issue is a bit subtle

00:15:22.590 --> 00:15:24.660 is that in a calibration plot,

00:15:24.660 --> 00:15:27.420 what they did is they use the partial R squared

00:15:27.420 --> 00:15:32.420 as a way to measure lambda and delta

00:15:32.690 --> 00:15:35.780 in a more interpretable way

00:15:35.780 --> 00:15:38.460 But actually the partial R squared for the observed

00:15:38.460 --> 00:15:42.410 and unobserved confounders are not directly comparable.

00:15:42.410 --> 00:15:45.920 This is because they're they use different reference model

00:15:45.920 --> 00:15:47.670 to start with.

00:15:47.670 --> 00:15:50.090 So, actually you need to be quite careful

00:15:50.090 --> 00:15:54.087 about these interpretation this calibration quotes.

00:15:56.150 --> 00:16:00.990 So, here is what I promised that suggests

00:16:00.990 --> 00:16:02.410 that you can actually identify

00:16:02.410 --> 00:16:05.810 these two sensitivity parameters lambda and delta.

00:16:05.810 --> 00:16:07.600 So here the red dots

00:16:07.600 --> 00:16:10.560 are the maximum likelihood estimators.

00:16:10.560 --> 00:16:14.020 And then these solid curves this regions,

00:16:14.020 --> 00:16:15.950 or the rejection,

00:16:15.950 --> 00:16:20.370 or I should say acceptance region

00:16:20.370 --> 00:16:23.080 for the likelihood ratio test.

00:16:23.080 --> 00:16:25.900 So this is at level 0.50,

00:16:25.900 --> 00:16:29.640 this is 0.10, this is 0.05.

00:16:29.640 --> 00:16:34.000 There is a symmetry around the origin that's

00:16:34.000 --> 00:16:37.270 because the U number is symmetric.

00:16:37.270 --> 00:16:40.680 So, lambda like delta is the same

00:16:40.680 --> 00:16:43.024 as minus lambda minus delta.

00:16:43.024 --> 00:16:44.470 But what you see

00:16:44.470 --> 00:16:47.130 is that you can actually estimate lambda and delta

00:16:47.130 --> 00:16:49.780 and you can sort of estimate it

00:16:49.780 --> 00:16:53.050 to be in a certain region.

00:16:53.050 --> 00:16:55.620 So, something a bit interesting here

00:16:55.620 --> 00:17:00.620 is that there's more you can say about Delta,

00:17:01.050 --> 00:17:03.000 which is the parameter for the outcome, 00:17:04.059 --> 00:17:06.827 than the parameter for the treatment lambda.

00:17:09.120 --> 00:17:10.640 But in any case,

00:17:10.640 --> 00:17:12.790 it didn’t look like we can just vary

00:17:12.790 --> 00:17:16.030 this parameter lambda delta freely in this space

00:17:16.030 --> 00:17:18.719 and then expect to get different results

00:17:18.719 --> 00:17:22.510 for each each point.

00:17:22.510 --> 00:17:24.910 What we actually can get is some estimate

00:17:24.910 --> 00:17:27.023 of this sensitivity parameters.

00:17:27.920 --> 00:17:30.100 So the lesson here is that

00:17:30.100 --> 00:17:32.480 if you use a parametric sensitivity models,

00:17:32.480 --> 00:17:34.900 then they need to be carefully constructed

00:17:34.900 --> 00:17:37.143 to avoid these kind of issues.

00:17:40.320 --> 00:17:42.760 So next I’ll talk about the first component

00:17:42.760 --> 00:17:44.430 of the sensitivity analysis,

00:17:44.430 --> 00:17:46.693 which is your sensitivity model.

00:17:47.750 --> 00:17:50.680 So very generally,

00:17:50.680 --> 00:17:53.560 if you think about what is the sensitivity model,

00:17:53.560 --> 00:17:58.560 is essentially it’s a model for the full data F,

00:18:00.270 --> 00:18:03.140 that include some things that are not observed.

00:18:03.140 --> 00:18:04.780 So, what we are trying to do here

00:18:04.780 --> 00:18:07.650 is to infer the full data distribution

00:18:07.650 --> 00:18:11.470 from some observed data, O.

00:18:11.470 --> 00:18:14.220 So a sensitivity model is basically

00:18:14.220 --> 00:18:18.060 a family of distributions of the full data,

00:18:18.060 --> 00:18:22.730 is parameterized by two parameters theta and eta.

00:18:22.730 --> 00:18:26.610 So, I’m using eta to stand for the sensitivity parameters

00:18:26.610 --> 00:18:28.810 and theta is some other parameters

00:18:28.810 --> 00:18:31.293 that parameterize the distribution.

00:18:32.600 --> 00:18:36.090 So the sensitivity model needs to satisfy two properties.

00:18:37.685 --> 00:18:39.701 So first of all,

00:18:39.701 --> 00:18:44.180 if we set the sensitivity parameter eta to be equal to zero,

00:18:44.180 --> 00:18:47.576 then that should correspond to our primary analysis

00:18:47.576 --> 00:18:48.570 assuming no unmeasured confounders.

00:18:48.570 --> 00:18:50.513 So I call this augmentation.

00:18:51.410 --> 00:18:55.960 A second property is that given the value of the

00:18:55.960 --> 00:18:58.740 of this sensitivity prior to eta,

00:18:58.740 --> 00:19:03.410 then we can actually identify this parameters data

00:19:03.410 --> 00:19:05.550 from the observed data.

00:19:05.550 --> 00:19:08.080 So this is sort of a minimal assumption.

00:19:08.080 --> 00:19:10.783 Otherwise, this model is simply too rich,

00:19:12.424 --> 00:19:14.820 and so I call model identifiability.

00:19:14.820 --> 00:19:17.700 So the statistical problem in sensitivity analysis

00:19:17.700 --> 00:19:20.140 is that if I give you the value of eta

00:19:20.140 --> 00:19:22.700 or the range of eta,

00:19:22.700 --> 00:19:25.730 can you use observed data to make inference

00:19:25.730 --> 00:19:28.520 about some causal parameter that is a function

00:19:28.520 --> 00:19:30.383 of the theta and eta.

00:19:31.910 --> 00:19:36.910 Okay, so this is a very general abstraction

00:19:37.310 --> 00:19:40.793 of what we have seen in the previous example.

00:19:42.720 --> 00:19:45.090 But it's a bit too general.

00:19:45.090 --> 00:19:48.600 So let's make it slightly more concrete

00:19:48.600 --> 00:19:53.423 by understanding these observational equivalence causes.

00:19:54.720 --> 00:19:57.500 So essentially, what we're trying to do

00:19:57.500 --> 00:19:59.200 is we observe some data,

00:19:59.200 --> 00:20:01.870 but then we know there's an underlying full data

00:20:01.870 --> 00:20:04.870 some other observe.

00:20:04.870 --> 00:20:07.610 And instead of just modeling the observed data,

00:20:07.610 --> 00:20:10.163 we're modeling the full data set.

00:20:10.163 --> 00:20:13.870 So that makes our model quite rich,

00:20:13.870 --> 00:20:16.743 because we're modeling something that are all observed.

00:20:17.640 --> 00:20:20.560 For that purpose is useful to define this

00:20:20.560 --> 00:20:23.870 observationally equivalence relation

00:20:23.870 --> 00:20:25.983 between two full data distribution,

00:20:26.900 --> 00:20:29.560 which just means that their implied

00:20:29.560 --> 00:20:33.840 observed data distributions are exactly the same.

00:20:33.840 --> 00:20:38.840 So we write this as this approximate equal

00:20:39.050 --> 00:20:43.160 to this equivalence symbol.

00:20:43.160 --> 00:20:45.490 So then we can define the equivalence class

00:20:45.490 --> 00:20:48.100 of a distribution of a full data distribution,

00:20:48.100 --> 00:20:51.380 which are all the other full data distributions

00:20:51.380 --> 00:20:54.530 in this family that are observationally equivalent

00:20:54.530 --> 00:20:56.393 to that distribution.

00:20:57.860 --> 00:21:01.907 Then we can sort of classify these sensitivity models

00:21:01.907 --> 00:21:05.053 based on the behavior of these equivalence classes.

00:21:06.930 --> 00:21:09.540 So, what happened in the last example

00:21:09.540 --> 00:21:14.540 is that the full data distribution full data model

00:21:14.570 --> 00:21:16.180 is not rich enough.

00:21:16.180 --> 00:21:19.683 So these equivalence classes are just singleton's

00:21:19.683 --> 00:21:24.100 so can actually identify the sensitivity parameter eta

00:21:24.100 --> 00:21:25.363 from the observed data.

00:21:26.300 --> 00:21:30.650 So, this makes this model testable in some sense

00:21:30.650 --> 00:21:33.853 with the choice of sensitivity parameter testable,

00:21:34.862 --> 00:21:37.483 and this should generally be avoided in practice.

00:21:39.000 --> 00:21:41.600 Then there are the global sensitivity models

00:21:42.680 --> 00:21:45.650 where you can basically freely vary

00:21:45.650 --> 00:21:48.280 the sensitivity parameter eta.

00:21:48.280 --> 00:21:50.920 And for any eta you can always find the theta

00:21:50.920 --> 00:21:53.960 such that it is observational equivalent

00:21:53.960 --> 00:21:55.443 to where you started from.

00:21:57.140 --> 00:22:01.130 And then even nicer models the separable model

00:22:01.130 --> 00:22:04.090 where basically, this eta,

00:22:04.090 --> 00:22:07.416 the sensitivity parameter doesn't change

00:22:07.416 --> 00:22:11.720 the observation of the observed data distribution.

00:22:11.720 --> 00:22:14.140 So for any theta and eta,

00:22:14.140 --> 00:22:16.883 theta and eta is equivalent to theta and zero.

00:22:17.730 --> 00:22:22.119 So these are really nice models to work with.

00:22:22.119 --> 00:22:25.880 So understand the difference between global models

00:22:25.880 --> 00:22:28.060 and separable models.

00:22:28.060 --> 00:22:32.410 So basically, it's just that they have different shapes

00:22:33.659 --> 00:22:37.480 of the equivalence classes.

00:22:37.480 --> 00:22:39.540 So for separable models,

00:22:39.540 --> 00:22:41.630 these equivalence classes,

00:22:41.630 --> 00:22:45.320 needs to be perpendicular to the theta axis.

00:22:46.350 --> 00:22:50.263 But that's not needed for global sensitivity models.

00:22:53.300 --> 00:22:56.930 So I've talked about what a sensitivity model means

00:22:56.930 --> 00:22:59.970 and some basic properties of it,

00:22:59.970 --> 00:23:02.240 but haven't talked about how to build them.

00:23:02.240 --> 00:23:05.362 So generally, in this setup,

00:23:05.362 --> 00:23:07.590 there's three ways to build a sensitivity model.

00:23:07.590 --> 00:23:09.200 And then they essentially correspond

00:23:09.200 --> 00:23:11.010 with different factorizations

00:23:11.010 --> 00:23:13.420 of the full data distribution.

00:23:13.420 --> 00:23:15.400 So there's a simultaneous model

00:23:15.400 --> 00:23:18.730 that tries to factorize distribution this way.

00:23:18.730 --> 00:23:22.250 So introduces unmeasured confounder, U,

00:23:22.250 --> 00:23:23.920 and then you need to model

00:23:23.920 --> 00:23:26.393 these three conditional probabilities.

00:23:27.495 --> 00:23:30.651 There's also the treatment model

00:23:30.651 --> 00:23:35.450 that doesn't rely on this unmeasured confounder U.

00:23:35.450 --> 00:23:38.550 But whether you need to specify is the distribution

00:23:38.550 --> 00:23:42.373 of the treatment given the unmeasured cofounders and x.

00:23:43.524 --> 00:23:46.350 And once you've specified that you can use Bayes formula

00:23:46.350 --> 00:23:47.913 to get this part.

00:23:49.920 --> 00:23:53.829 And then there's the outcome model that factorizes

00:23:53.829 --> 00:23:56.530 this distribution in the other way.

00:23:56.530 --> 00:24:00.020 So this is basically the propensity score

00:24:00.020 --> 00:24:03.330 and the third turn is what we need to specify

00:24:03.330 --> 00:24:05.830 it's a sensitivity parameter.

00:24:05.830 --> 00:24:08.900 So in the missing data literature,

00:24:08.900 --> 00:24:10.970 second model kind of model

00:24:10.970 --> 00:24:13.137 is usually called selection model.

00:24:13.137 --> 00:24:15.680 And the third kind of models usually called

00:24:15.680 --> 00:24:17.340 pattern mixture model,

00:24:17.340 --> 00:24:19.990 and there are other names that have been given to it.

00:24:22.730 --> 00:24:26.260 And basically different sensitivity models,

00:24:26.260 --> 00:24:29.530 they amount to different ways of specifying these

00:24:30.700 --> 00:24:32.970 either non identifiable distributions,

00:24:32.970 --> 00:24:36.520 which are these ones that are underlined.

00:24:36.520 --> 00:24:41.520 A good review is this report by a committee

00:24:41.580 --> 00:24:44.983 organized by the National Research Council.

00:24:46.043 --> 00:24:49.560 This ongoing review paper that we're writing

00:24:49.560 --> 00:24:54.063 also gives a comprehensive review of many models

00:24:54.063 --> 00:24:58.313 that have been proposed using these factorizations.

00:25:00.169 --> 00:25:03.170 Okay, so that's about the sensitivity model.

00:25:03.170 --> 00:25:06.683 The next component is statistical inference.

00:25:11.480 --> 00:25:14.020 Things get a little bit tricky here,

00:25:14.020 --> 00:25:16.670 because there are two kinds of inference

00:25:16.670 --> 00:25:19.250 or two modes of inference we can talk about

00:25:19.250 --> 00:25:20.880 in this study.

00:25:20.880 --> 00:25:24.490 So, the first mode of inference is point identify inference.

00:25:24.490 --> 00:25:27.200 So you only care about a fixed value

00:25:27.200 --> 00:25:29.187 of the sensitivity parameter eta.

00:25:31.503 --> 00:25:33.620 And the second kind of inference

00:25:33.620 --> 00:25:36.170 is partial identified inference,

00:25:36.170 --> 00:25:40.390 where you perform the statistical inference simultaneously

00:25:40.390 --> 00:25:43.730 for a range of security parameters eta.

00:25:43.730 --> 00:25:45.963 And that range H is given to you.

00:25:50.330 --> 00:25:53.910 And in these different modes of inferences,

00:25:53.910 --> 00:25:56.940 it comes differences to core guarantees.

00:25:56.940 --> 00:26:01.640 So for point identified inference usually let's say

00:26:02.700 --> 00:26:04.080 for interval estimators,

00:26:04.080 --> 00:26:07.840 you want to construct confidence intervals.

00:26:07.840 --> 00:26:12.260 And these confidence intervals depend on the observed theta

00:26:12.260 --> 00:26:15.290 and the sensitivity parameter which

00:26:15.290 --> 00:26:17.390 your last to use

00:26:17.390 --> 00:26:19.760 in a point of identified inference

00:26:19.760 --> 00:26:22.810 and it must cover the true parameter

00:26:22.810 --> 00:26:25.270 with one minus alpha probability

00:26:25.270 --> 00:26:28.130 for all the distributions in your model.

00:26:28.130 --> 00:26:29.410 Okay that's the infimum.

00:26:30.250 --> 00:26:34.630 But for partial identified inference,

00:26:34.630 --> 00:26:37.630 you're only allowed to use an interval

00:26:37.630 --> 00:26:39.723 that depends on the range, H.

00:26:40.880 --> 00:26:43.173 So, it cannot depend on a specific values

00:26:43.173 --> 00:26:45.720 of the sensitivity parameter,

00:26:45.720 --> 00:26:50.480 because you only know eta is in this range H.

00:26:50.480 --> 00:26:55.480 It need to satisfy this very similar criteria.

00:26:55.530 --> 00:26:59.230 So I call this intervals that satisfy this criteria

00:26:59.230 --> 00:27:01.000 in the sensitivity interval.

00:27:01.000 --> 00:27:03.300 But in the literature people have also called this

00:27:03.300 --> 00:27:06.833 uncertainty interval and or just confidence interval.

00:27:07.840 --> 00:27:11.060 But to make it different from the first case,

00:27:11.060 --> 00:27:13.160 we're calling a sensitivity interval here.

00:27:14.610 --> 00:27:19.200 So you can see that these two equations,

00:27:19.200 --> 00:27:21.510 two criterias look very similar,

00:27:21.510 --> 00:27:25.250 besides just that this interval needs to depend on the range

00:27:25.250 --> 00:27:28.970 instead of a particular value of the sensitivity parameter.

00:27:28.970 --> 00:27:31.000 But actually, they're quite different.

00:27:31.000 --> 00:27:32.763 This is usually much wider.

00:27:33.750 --> 00:27:34.603 The reason is,

00:27:35.707 --> 00:27:37.250 you can actually write an equivalent form

00:27:37.250 --> 00:27:38.803 of this equation one,

00:27:39.909 --> 00:27:44.510 because this only depends on the observed data

00:27:44.510 --> 00:27:46.170 and the range H.

00:27:46.170 --> 00:27:48.610 Then for every theta in that,

00:27:48.610 --> 00:27:51.710 sorry for every eta in that range H,

00:27:51.710 --> 00:27:55.607 is missing here, eta in H and also

00:27:55.607 --> 00:27:59.760 that's observationally equivalent to a two distribution.

00:27:59.760 --> 00:28:02.055 This interval also needs to cover

00:28:02.055 --> 00:28:05.823 the corresponding theta parameter.

00:28:07.160 --> 00:28:08.030 So in that sense,

00:28:08.030 --> 00:28:12.240 this is a much stronger guarantee that you have.

00:28:16.112 --> 00:28:20.773 So, in terms of the statistical methods,

00:28:20.773 --> 00:28:25.243 point identified inference is usually quite straightforward.

00:28:26.290 --> 00:28:28.540 It's very similar to our primary analysis.

00:28:28.540 --> 00:28:32.070 So, primary analysis just assumes this eta equals to zero,

00:28:32.070 --> 00:28:36.050 but this sensitivity analysis assumes eta is known.

00:28:36.050 --> 00:28:38.340 So usually you just you can just plug in

00:28:38.340 --> 00:28:41.810 this eta in some way as an offset to your model.

00:28:41.810 --> 00:28:44.680 And then everything works out in almost the same way

00:28:44.680 --> 00:28:46.253 as a primary analysis.

00:28:47.590 --> 00:28:49.930 But for partially identified analysis,

00:28:49.930 --> 00:28:54.510 things become quite more challenging.

00:28:54.510 --> 00:28:57.860 And there are several methods several approaches

00:28:57.860 --> 00:28:59.063 that you can take.

00:29:00.260 --> 00:29:04.960 So, essentially there are two big classes of methods,

00:29:04.960 --> 00:29:07.610 one is bound estimation,

00:29:07.610 --> 00:29:11.010 one is combining point identified inference.

00:29:11.010 --> 00:29:13.660 So, for bond estimation,

00:29:13.660 --> 00:29:17.720 it tries to directly make inference about the two ends

00:29:17.720 --> 00:29:21.060 of this partial identify region.

00:29:21.060 --> 00:29:26.060 So, this set this is the region of the parameter beta

00:29:26.330 --> 00:29:28.840 that are sort of indistinguishable,

00:29:28.840 --> 00:29:33.543 if I only know this sensitivity parameter eta is in H.

00:29:34.912 --> 00:29:39.912 If we can somehow directly estimate the infimum and supremum

00:29:40.060 --> 00:29:43.740 of this in this set, 00:29:43.740 --> 00:29:46.300 but then that gotta get us a way

00:29:46.300 --> 00:29:48.553 to make partial identified inference.

00:29:50.470 --> 00:29:53.170 The second method is basically

00:29:53.170 --> 00:29:58.170 to try to combine the results of point identified inference.

00:29:59.350 --> 00:30:02.410 The main idea is to sort of construct

00:30:02.410 --> 00:30:05.190 let's say interval estimators,

00:30:05.190 --> 00:30:08.090 for each individual sensitivity parameter

00:30:08.090 --> 00:30:10.680 and then take a union of them.

00:30:10.680 --> 00:30:13.630 So, these are the two broad approaches

00:30:13.630 --> 00:30:15.973 to the partially identified inference.

00:30:17.610 --> 00:30:20.150 And so, within the first approach

00:30:20.150 --> 00:30:22.010 the bound estimation approach,

00:30:22.010 --> 00:30:24.730 there are also several variety of,

00:30:24.730 --> 00:30:26.700 there are several possible methods

00:30:26.700 --> 00:30:28.163 depending on your problem.

00:30:29.480 --> 00:30:31.470 So, the first problem,

00:30:31.470 --> 00:30:34.770 the first method is called separable balance.

00:30:34.770 --> 00:30:38.930 But before that, let's just slightly change our notation

00:30:38.930 --> 00:30:43.930 and parameterize this range H by a hyper parameter gamma.

00:30:46.526 --> 00:30:51.526 So, this is useful when we outline these methods.

00:30:51.830 --> 00:30:54.800 And then this beta L of gamma,

00:30:54.800 --> 00:30:59.683 this is the lower end of the partial identify region.

00:31:00.910 --> 00:31:04.913 So the first method is called separable bounds.

00:31:05.937 --> 00:31:10.937 What it tries to do is to write this lower end

00:31:11.418 --> 00:31:15.370 as a function of beta star and gamma,

00:31:15.370 --> 00:31:19.853 where beta star is your primary analysis estimate.

00:31:20.930 --> 00:31:23.650 So let's say theta star zero

00:31:23.650 --> 00:31:26.550 is what you would do in a primary analysis

00:31:26.550 --> 00:31:30.413 that is observationally equivalent to the true distribution.

00:31:31.910 --> 00:31:36.910 And then, if beta star is the corresponding causal effect,

00:31:37.030 --> 00:31:38.563 from that model,

00:31:39.420 --> 00:31:42.380 and if somehow can write this lower end

00:31:42.380 --> 00:31:45.666 as a function of beta star and gamma

00:31:45.666 --> 00:31:47.160 and the function is known,

00:31:47.160 --> 00:31:49.670 then our life is quite easy,

00:31:49.670 --> 00:31:52.540 because we already know how to make inference

00:31:52.540 --> 00:31:55.360 about beta star from the primary analysis.

00:31:55.360 --> 00:31:57.230 And all we need to do is just plug in

00:31:57.230 --> 00:31:59.200 that beta star in this formula,

00:31:59.200 --> 00:32:00.400 and then we're all done.

00:32:01.810 --> 00:32:05.540 And we call this separable because it allows us

00:32:05.540 --> 00:32:09.140 to separate the primary analysis

00:32:09.140 --> 00:32:11.330 from the sensitivity analysis.

00:32:11.330 --> 00:32:14.940 And statistical inference becomes a trivial extension

00:32:14.940 --> 00:32:16.940 of the primary analysis.

00:32:16.940 --> 00:32:20.470 So, some examples of this kind of method

00:32:20.470 --> 00:32:23.650 include the classical cornfields bound

00:32:25.680 --> 00:32:27.150 and the E-value,

00:32:27.150 --> 00:32:29.320 if you have heard about them,

00:32:29.320 --> 00:32:31.340 and E-value seems quite popular

00:32:31.340 --> 00:32:33.980 these days at demonology.

00:32:36.870 --> 00:32:40.949 The second type of bound estimation

00:32:40.949 --> 00:32:44.975 is called tractable bounds.

00:32:44.975 --> 00:32:47.600 So, in these cases,

00:32:47.600 --> 00:32:51.620 we may derive this lower bound as a function

00:32:51.620 --> 00:32:54.300 of theta star and gamma.

00:32:54.300 --> 00:32:58.020 So we are not able to reduce it to just depend

00:32:58.020 --> 00:33:00.360 on beta star the causal effect

00:33:00.360 --> 00:33:04.029 under no unmeasured confounding,

00:33:04.029 --> 00:33:06.660 but we're able to express in terms of theta star.

00:33:06.660 --> 00:33:10.720 And then the function gl is also some practical functions

00:33:10.720 --> 00:33:12.760 that we can compute.

00:33:12.760 --> 00:33:17.170 And then this also makes our lives quite a lot easier,

00:33:17.170 --> 00:33:21.149 because we can just replace this theta star,

00:33:21.149 --> 00:33:24.614 which can be nonparametric can be parametric,

00:33:24.614 --> 00:33:26.973 by its empirical estimate.

00:33:28.110 --> 00:33:31.310 And, often in these cases,

00:33:31.310 --> 00:33:34.670 we can find some central limit theorems

00:33:34.670 --> 00:33:37.930 for the corresponding sample estimator,

00:33:37.930 --> 00:33:41.750 such that the sample estimator of the bounds

00:33:41.750 --> 00:33:46.190 converges to its truth at root and rate

00:33:46.190 --> 00:33:49.453 and it follows the normal limit.

00:33:50.925 --> 00:33:55.240 And then if we can estimate this standard error,

00:33:55.240 --> 00:33:58.482 then we can use this central limit theorem

00:33:58.482 --> 00:34:02.480 to make partial identified inference

00:34:02.480 --> 00:34:04.363 because we can estimate the bounds.

00:34:06.925 --> 00:34:08.630 There's some examples in the literature,

00:34:08.630 --> 00:34:11.093 you're familiar with these papers.

00:34:12.110 --> 00:34:14.230 But one thing to be careful about

00:34:14.230 --> 00:34:16.390 these kind of tractable bounds

00:34:16.390 --> 00:34:20.790 is that things that get a little bit tricky

00:34:20.790 --> 00:34:23.740 with syntactic theory.

00:34:23.740 --> 00:34:26.960 This is because in a syntactic theory,

00:34:26.960 --> 00:34:30.220 the confidence intervals or the sensitivity intervals

00:34:30.220 --> 00:34:31.750 in this case,

00:34:31.750 --> 00:34:36.743 can be point wise or uniform in terms of the sample size.

00:34:38.210 --> 00:34:42.887 So it's possible that if the convergence,

00:34:45.350 --> 00:34:48.925 if there are statistical guarantee is point wise,

00:34:48.925 --> 00:34:53.925 then you sometimes in extreme cases,

00:34:55.725 --> 00:34:58.190 even with very large sample size,

00:34:58.190 --> 00:35:01.160 they're still exist data distributions

00:35:01.160 --> 00:35:03.343 such that your coverage is very poor.

00:35:04.670 --> 00:35:07.770 So this point is discussed very heavily

00:35:07.770 --> 00:35:09.810 in econometrics literature.

00:35:09.810 --> 00:35:13.223 And these are some references.

00:35:15.040 --> 00:35:18.300 So that's the second type of method

00:35:18.300 --> 00:35:20.853 in the first broad approach.

00:35:22.010 --> 00:35:24.727 The third kind of method

00:35:24.727 --> 00:35:28.470 is called stochastic programming.

00:35:28.470 --> 00:35:33.470 And this applies when the model is separable.

00:35:34.338 --> 00:35:39.338 So and we can write this parameter we're interested in

00:35:40.400 --> 00:35:43.460 as some expectation of some function

00:35:43.460 --> 00:35:46.763 of the theta and the sensitivity parameter eta.

00:35:48.140 --> 00:35:49.603 Okay, so in this case,

00:35:50.890 --> 00:35:53.540 the bound becomes the optimal value

00:35:53.540 --> 00:35:56.110 for an optimization problem,

00:35:56.110 --> 00:35:59.753 which you want to minimize expectation of some function.

00:36:00.730 --> 00:36:04.630 And the parameter in this function is in some set

00:36:04.630 --> 00:36:06.113 as defined by U.

00:36:07.660 --> 00:36:10.560 So, this is known as stochastic programming.

00:36:10.560 --> 00:36:14.000 So, this type of problem is known as stochastic programming

00:36:14.000 --> 00:36:15.653 in the optimization literature.

00:36:16.900 --> 00:36:18.980 And what people do there

00:36:18.980 --> 00:36:22.047 is they sample from the distribution, 00:36:22.047 --> 00:36:25.860 and then they try to use it to solve the empirical version

00:36:25.860 --> 00:36:28.900 and try to use that as approximate solution

00:36:28.900 --> 00:36:32.640 to this population optimization problem,

00:36:32.640 --> 00:36:36.100 which we can't directly U value evaluate.

00:36:36.100 --> 00:36:38.950 And the method is called sample average approximation

00:36:38.950 --> 00:36:40.603 in the optimization literature.

00:36:42.470 --> 00:36:44.393 So, what is shown there.

00:36:46.515 --> 00:36:51.260 And Alex Shapiro did a lot of great work on this,

00:36:51.260 --> 00:36:56.260 is that nice problems with compact set age,

00:36:56.540 --> 00:36:58.916 and everything is euclidean.

00:36:58.916 --> 00:37:00.530 So it's finite dimensional.

00:37:00.530 --> 00:37:02.830 Then you actually have a central limit theorem

00:37:03.730 --> 00:37:05.693 for the sample optimal value.

00:37:07.150 --> 00:37:11.820 And this link, is a link between sensitivity analysis

00:37:11.820 --> 00:37:15.753 and stochastic programming is made in this paper

00:37:15.753 --> 00:37:17.263 by Tudball et al.

00:37:20.330 --> 00:37:22.890 Okay, so that's the first broad approach

00:37:22.890 --> 00:37:25.003 with doing bounds estimation.

00:37:26.290 --> 00:37:29.330 The second broad approach is to combine the results

00:37:29.330 --> 00:37:31.423 of points identified inference.

00:37:32.370 --> 00:37:36.930 So, the first possibility is to take a union

00:37:36.930 --> 00:37:40.020 of the individual confidence intervals.

00:37:40.020 --> 00:37:43.332 Suppose these are the confidence intervals

00:37:43.332 --> 00:37:45.282 when the sensitivity from eta is given.

00:37:46.510 --> 00:37:51.134 Then, it is very simple to just apply a union bound

00:37:51.134 --> 00:37:54.060 and to show that if you take a union

00:37:54.060 --> 00:37:57.460 of these individual confidence intervals,

00:37:57.460 --> 00:38:01.100 then they should satisfy the criteria

00:38:01.100 --> 00:38:03.350 for sensitivity interval.

00:38:03.350 --> 00:38:06.994 So now, if you take a union this interval only depends

00:38:06.994 --> 00:38:07.960 on the range H,

00:38:07.960 --> 00:38:11.511 and then you just apply the union bound

00:38:11.511 --> 00:38:13.933 and get this formula from the first.

00:38:17.080 --> 00:38:19.610 And this can be slightly improved

00:38:19.610 --> 00:38:23.270 to cover not just these parameters,

00:38:23.270 --> 00:38:27.210 but also the entire partial identified region

00:38:27.210 --> 00:38:29.910 if the intervals if the confidence intervals

00:38:29.910 --> 00:38:32.653 have the same tail probabilities.

00:38:35.050 --> 00:38:36.923 So we discussed this in our paper.

00:38:38.653 --> 00:38:43.350 And here, so, all we need to do

00:38:43.350 --> 00:38:45.113 is to compute this union.

00:38:45.970 --> 00:38:49.230 So, which essentially is an optimization problem

00:38:49.230 --> 00:38:52.480 we'd like to minimize the lower bound,

00:38:52.480 --> 00:38:57.257 that the lower confidence point Cl of eta over eta in H

00:38:58.988 --> 00:39:00.688 and similarly for the upper bound.

00:39:01.710 --> 00:39:04.550 And usually using of syntactic theory,

00:39:04.550 --> 00:39:09.340 we can get some normal base confidence

00:39:09.340 --> 00:39:12.440 intervals for each fixed eta.

00:39:12.440 --> 00:39:14.430 And then we just need to optimize

00:39:14.430 --> 00:39:19.430 this thing this confidence interval over eta.

00:39:19.940 --> 00:39:21.950 But for many problems this can be

00:39:21.950 --> 00:39:26.440 computationally challenging because the standard errors

00:39:26.440 --> 00:39:29.000 are usually quite complicated

00:39:30.057 --> 00:39:32.370 and it has some very nonlinear dependence

00:39:32.370 --> 00:39:34.010 on the parameter eta.

00:39:34.010 --> 00:39:36.153 So optimizing this can be tricky.

00:39:39.854 --> 00:39:43.840 This is where another method of percentile bootstrap method

00:39:43.840 --> 00:39:46.600 can greatly simplify the problem.

00:39:46.600 --> 00:39:51.600 It’s proposed by this paper that we wrote,

00:39:52.710 --> 00:39:55.920 and what it does is instead of using

00:39:55.920 --> 00:40:00.770 the syntactic confidence interval for fixed eta,

00:40:00.770 --> 00:40:03.790 we use the percentile bootstrap interval.

00:40:03.790 --> 00:40:06.290 Where we take theta samples,

00:40:06.290 --> 00:40:10.850 and then you estimate the causal effect beta

00:40:10.850 --> 00:40:14.057 in each resample and then take quantiles.

00:40:15.230 --> 00:40:19.330 Okay, so if you use this confidence interval,

00:40:19.330 --> 00:40:24.330 then there is a general,

00:40:24.540 --> 00:40:28.700 generalized minimax inequality that allows us to construct

00:40:28.700 --> 00:40:31.873 this percentile bootstrap sensitivity interval.

00:40:32.870 --> 00:40:36.890 So what it does is this thing in the inside

00:40:36.890 --> 00:40:41.010 is just the union of these percentile construct

00:40:41.010 --> 00:40:44.910 intervals for fixed eta,

00:40:44.910 --> 00:40:48.063 taken over eta in H.

00:40:48.910 --> 00:40:51.480 And then this generalized minimax inequality

00:40:51.480 --> 00:40:56.480 allows us to interchange the infimum with quanto

00:40:56.700 --> 00:40:59.940 and the supremum of a quanto.

00:40:59.940 --> 00:41:01.340 Okay, so the infimum of a quanto

00:41:01.340 --> 00:41:04.303 is greater than equal to the quanto of infimum

00:41:05.215 --> 00:41:07.050 and that it’s always true.

00:41:07.050 --> 00:41:08.550 So it’s just a generalization

00:41:08.550 --> 00:41:11.233 of the familia minimax inequality.

00:41:12.560 --> 00:41:15.760 Now, if you look at this order interval,

00:41:15.760 --> 00:41:18.580 this is much easier to compute,

00:41:18.580 --> 00:41:20.210 because all it needs to do

00:41:20.210 --> 00:41:25.098 is you gather data resample,

00:41:25.098 --> 00:41:29.430 then you just need to repeat method 1.3.

00:41:29.430 --> 00:41:33.860 So just get the infimum of the point estimate

00:41:33.860 --> 00:41:37.460 for that resample and the supremum for that resample.

00:41:37.460 --> 00:41:40.150 Then you do this over many, many resamples

00:41:41.215 --> 00:41:43.550 and then you take the quantiles of the infimum,

00:41:43.550 --> 00:41:47.898 lower of the infimum and upper quantile of the supremum,

00:41:47.898 --> 00:41:49.690 and then you're done.

00:41:49.690 --> 00:41:53.370 And because this union sensitivity interval

00:41:53.370 --> 00:41:54.920 is always valid,

00:41:54.920 --> 00:41:58.330 if the individual confidence intervals are valid.

00:41:58.330 --> 00:42:02.370 So you almost got a very you got a free lunch

00:42:02.370 --> 00:42:03.380 in some sense,

00:42:03.380 --> 00:42:06.260 you don't need to show any heavy theory.

00:42:06.260 --> 00:42:07.760 All you need to show is that

00:42:07.760 --> 00:42:10.767 these percentile bootstrap intervals are valid

00:42:10.767 --> 00:42:13.340 for each fixed eta,

00:42:13.340 --> 00:42:18.340 which are much easier to establish in real problems.

00:42:22.630 --> 00:42:24.770 And this is sort of selfish,

00:42:24.770 --> 00:42:27.200 where I'd like to compare this idea

00:42:27.200 --> 00:42:29.140 with Efron's bootstrap,

00:42:29.140 --> 00:42:31.140 where what was found there

00:42:31.140 --> 00:42:33.370 is that you've got a point estimator,

00:42:33.370 --> 00:42:34.990 you resample your data,

00:42:34.990 --> 00:42:38.230 and then many times and then use bootstrap

00:42:38.230 --> 00:42:39.780 to get the confidence interval.

00:42:40.808 --> 00:42:44.030 For partially identified inference,

00:42:44.030 --> 00:42:45.940 you need to do a bit more.

00:42:45.940 --> 00:42:48.140 So for each resample you need

00:42:48.140 --> 00:42:51.550 to get extrema optimal estimator.

00:42:51.550 --> 00:42:55.176 Then the minimax inequality allows you just

00:42:55.176 --> 00:43:00.160 sort of transfer the intuition from the bootstrap, 00:43:00.160 --> 00:43:02.480 for bootstrap from point identification

00:43:02.480 --> 00:43:04.063 to partial identification.

00:43:07.560 --> 00:43:10.553 So the third approach in this,

00:43:11.408 --> 00:43:13.574 is a third method in this general approach

00:43:13.574 --> 00:43:15.010 is to take the supremum of key value.

00:43:15.010 --> 00:43:18.090 And this is used in Rosenbaum sensitivity analysis.

00:43:18.090 --> 00:43:19.823 If you're familiar with that.

00:43:21.680 --> 00:43:24.490 Essentially it's a hypothesis testing analog

00:43:24.490 --> 00:43:27.193 of the Union confidence interval method.

00:43:28.540 --> 00:43:29.880 What it does is that

00:43:29.880 --> 00:43:34.860 if you have individually valid P values for a fixed eta,

00:43:34.860 --> 00:43:37.670 then you just take the supremum of the P values

00:43:37.670 --> 00:43:41.380 over all the etas in this range.

00:43:41.380 --> 00:43:44.693 And that can be used for partially identified inference.

00:43:45.547 --> 00:43:48.680 So what Rosenbaum did,

00:43:48.680 --> 00:43:51.990 and Rosenbaum is really a pioneer in this area

00:43:51.990 --> 00:43:55.620 in the partially identify sensitivity analysis.

00:43:55.620 --> 00:43:59.410 So what he did was use randomization tests

00:43:59.410 --> 00:44:01.073 to construct these key values.

00:44:02.540 --> 00:44:06.570 So, this is usually done for matched observational studies

00:44:06.570 --> 00:44:11.570 and the inside of this line of work

00:44:11.790 --> 00:44:16.044 is that you can use these inequalities

00:44:16.044 --> 00:44:18.940 particularly Holley's inequality

00:44:18.940 --> 00:44:21.500 in probabilistic combinatorics

00:44:21.500 --> 00:44:25.113 to efficiently compute these supremum of the P values.

00:44:26.440 --> 00:44:29.740 So, usually what is done there is that

00:44:29.740 --> 00:44:32.470 the Holley's inequality gives you a way

00:44:32.470 --> 00:44:36.983 to upper bound the distribution of a that, 00:44:38.680 --> 00:44:40.920 to upper bound family of distributions

00:44:42.080 --> 00:44:45.424 in the stochastic dominance sense.

00:44:45.424 --> 00:44:49.793 So, that is used to get these supremum of the P values.

00:44:51.070 --> 00:44:56.070 And so, basically the idea is to use some theoretical tool

00:44:58.520 --> 00:45:02.023 to simplify the computation.

00:45:05.366 --> 00:45:08.140 Okay, so that's the statistical inference.

00:45:08.140 --> 00:45:10.270 The third part, the third component

00:45:10.270 --> 00:45:13.190 is interpretation of sensitivity analysis.

00:45:13.190 --> 00:45:16.950 And this is the area that we actually really need

00:45:16.950 --> 00:45:19.293 a lot of good work at the moment.

00:45:20.460 --> 00:45:25.460 So, overall, there are two good ideas that seem to work,

00:45:25.770 --> 00:45:27.560 that seem to improve the interpretation

00:45:27.560 --> 00:45:28.873 of sensitivity analysis.

00:45:29.990 --> 00:45:31.690 The first is sensitivity value,

00:45:31.690 --> 00:45:35.203 the second is the calibration using measured confounders.

00:45:36.080 --> 00:45:38.460 So the sensitivity value is basically

00:45:38.460 --> 00:45:41.062 the value of the sensitivity parameter

00:45:41.062 --> 00:45:42.170 or the hyper parameter,

00:45:42.170 --> 00:45:46.163 where some qualitative conclusions about your study change.

00:45:47.603 --> 00:45:51.360 And in our motivating example,

00:45:51.360 --> 00:45:54.920 this is where the estimated average treatment effect

00:45:54.920 --> 00:45:58.640 is reduced by half an Rosenbaum sensitivity analysis

00:45:58.640 --> 00:46:00.140 if you are familiar with that.

00:46:01.079 --> 00:46:02.820 This is where, this is the value of the gamma

00:46:02.820 --> 00:46:03.913 in his model, 00:46:04.766 --> 00:46:07.763 where we can no longer reject the causal null hypothesis.

00:46:09.640 --> 00:46:13.610 So, this is can be seen as kind of an extension

00:46:13.610 --> 00:46:15.763 of the idea of a P value.

00:46:16.660 --> 00:46:19.330 So P value is used for primary analysis,

00:46:19.330 --> 00:46:21.680 so assuming no unmeasure confounding,

00:46:21.680 --> 00:46:24.100 and then for sensitivity analysis,

00:46:24.100 --> 00:46:26.953 you can use the sensitivity value to sort of sorry,

00:46:30.142 --> 00:46:32.142 that's the P value it basically measures

00:46:33.293 --> 00:46:36.270 how likely your results,

00:46:36.270 --> 00:46:39.230 your sort of false rejection is due to

00:46:39.230 --> 00:46:43.600 sort of random chance.

00:46:43.600 --> 00:46:45.610 But then what a sensitivity value does

00:46:45.610 --> 00:46:50.590 is measures how much sort of how sensitive your resources is

00:46:50.590 --> 00:46:53.026 in some sense, so, how much deviation

00:46:53.026 --> 00:46:54.940 from the unmeasured confounding it takes

00:46:54.940 --> 00:46:57.113 to alter your conclusion.

00:46:58.350 --> 00:47:00.668 And for sensitivity value,

00:47:00.668 --> 00:47:03.950 there often exists a phase transition phenomenon

00:47:03.950 --> 00:47:05.873 for partially identified inference.

00:47:07.020 --> 00:47:11.290 This is because if you take your hyper parameter gamma

00:47:11.290 --> 00:47:12.850 to be very large,

00:47:12.850 --> 00:47:15.210 then essentially your partially identify region

00:47:15.210 --> 00:47:17.060 already covered in null.

00:47:17.060 --> 00:47:20.330 So, no matter how large your sample size is

00:47:20.330 --> 00:47:21.853 you can never reject null.

00:47:23.240 --> 00:47:26.310 So, this is sort of an interesting phenomenon

00:47:27.632 --> 00:47:32.070 and explained first discovered by Rosenbaum

00:47:32.070 --> 00:47:37.070 in this paper I wrote also clarified some problems

00:47:37.610 --> 00:47:42.153 some issues in both the phase transition.

00:47:44.080 --> 00:47:46.370 So, the second idea is the calibration

00:47:46.370 --> 00:47:48.650 using measured confounders.

00:47:48.650 --> 00:47:50.690 So, you have already seen an example

00:47:50.690 --> 00:47:54.300 in a motivating study.

00:47:54.300 --> 00:47:59.180 It's really a very necessary and practical solution

00:47:59.180 --> 00:48:01.086 to quantify the sensitivity,

00:48:01.086 --> 00:48:05.066 because it's not really very useful if you tell people,

00:48:05.066 --> 00:48:08.230 we are sensitive at gamma equals to two,

00:48:08.230 --> 00:48:09.400 what does that really mean?

00:48:09.400 --> 00:48:12.568 That depends on some mathematical model.

00:48:12.568 --> 00:48:14.950 But if we can somehow compare that

00:48:14.950 --> 00:48:17.483 with what we do observe,

00:48:18.390 --> 00:48:19.510 and we have,

00:48:19.510 --> 00:48:22.950 often the practitioners have some good sense

00:48:22.950 --> 00:48:26.800 about what are the important confounders and what are not.

00:48:26.800 --> 00:48:30.610 Then this really gives us a way to calibrate

00:48:30.610 --> 00:48:35.150 and strengthen the conclusions of a sensitivity analysis.

00:48:35.150 --> 00:48:37.990 But unfortunately, although there are some good heuristics

00:48:37.990 --> 00:48:39.373 about the calibration,

00:48:40.366 --> 00:48:43.890 they're often suffer from some subtle issues,

00:48:43.890 --> 00:48:46.100 like the ones that I described

00:48:46.100 --> 00:48:47.550 in the beginning of the talk.

00:48:48.627 --> 00:48:51.113 If you carefully parameterize your models

00:48:51.113 --> 00:48:52.863 this can become easier.

00:48:53.768 --> 00:48:56.080 And this recent paper sort of explored this

00:48:56.080 --> 00:49:00.540 in terms of linear models.

00:49:00.540 --> 00:49:03.670 But really there's not a unifying framework

00:49:03.670 --> 00:49:07.770 then you can cover more general cases

00:49:07.770 --> 00:49:09.913 and lots of work are needed.

00:49:11.220 --> 00:49:13.200 And when I was writing the slides,

00:49:13.200 --> 00:49:15.450 I thought maybe what we really need

00:49:15.450 --> 00:49:17.530 is to somehow build this calibration

00:49:17.530 --> 00:49:19.570 into the sensitivity model.

00:49:19.570 --> 00:49:21.890 Because currently our workflow is that

00:49:21.890 --> 00:49:23.930 we assume a sensitivity model,

00:49:23.930 --> 00:49:26.380 and we see where things get changed,

00:49:26.380 --> 00:49:28.890 and then we try to interpret those values

00:49:28.890 --> 00:49:30.760 where things get changed.

00:49:30.760 --> 00:49:34.328 But suppose if we somehow build that,

00:49:34.328 --> 00:49:37.750 if we left the range H eta to be defined

00:49:37.750 --> 00:49:40.480 in terms of this calibration.

00:49:40.480 --> 00:49:45.210 Perhaps gamma directly means some kind of comparisons

00:49:45.210 --> 00:49:48.630 that measured confounders this would solve some

00:49:48.630 --> 00:49:50.410 a lot of the issues.

00:49:50.410 --> 00:49:52.930 This is just a thought I came up

00:49:52.930 --> 00:49:54.703 when I was preparing for this talk.

00:49:56.230 --> 00:49:58.536 Okay, so to summarize,

00:49:58.536 --> 00:50:01.460 so there is number of messages,

00:50:01.460 --> 00:50:04.980 which I hope you can take home.

00:50:04.980 --> 00:50:07.860 There are three components of a sensitivity analysis.

00:50:07.860 --> 00:50:10.600 Model augmentations, statistical inference

00:50:10.600 --> 00:50:13.840 and the interpretation of sensitivity analysis.

00:50:13.840 --> 00:50:17.160 So sensitivity model is about parameterizing,

00:50:17.160 --> 00:50:19.210 the full data distribution.

00:50:19.210 --> 00:50:21.780 And that's basically about over parameterizing

00:50:22.660 --> 00:50:24.540 the observed data distribution.

00:50:24.540 --> 00:50:26.150 And you can understand these models

00:50:26.150 --> 00:50:28.603 by the observational equivalence classes.

00:50:29.930 --> 00:50:32.940 You can get different model augmentations

00:50:32.940 --> 00:50:35.427 by factorizing the distribution differently

00:50:35.427 --> 00:50:37.670 and specify different models

00:50:37.670 --> 00:50:39.873 for those that are on identifiable.

00:50:41.392 --> 00:50:45.330 And there's a difference between point identified inference

00:50:45.330 --> 00:50:47.410 and partially identified inference,

00:50:47.410 --> 00:50:50.693 and partially identified inference is usually much harder.

00:50:51.667 --> 00:50:55.090 And there are two general approaches

00:50:55.090 --> 00:50:56.700 for partially identified inference,

00:50:56.700 --> 00:51:01.023 bound estimation and combining point identified inference.

00:51:02.290 --> 00:51:04.970 For interpretation of sensitivity analysis,

00:51:04.970 --> 00:51:07.992 there seem to be two good ideas so far,

00:51:07.992 --> 00:51:10.450 to use the sensitivity value,

00:51:10.450 --> 00:51:12.850 and to calibrate that sensitivity value

00:51:12.850 --> 00:51:14.513 using measured confounders.

00:51:16.040 --> 00:51:17.680 But overall,

00:51:17.680 --> 00:51:22.680 I'd say this is still a very,

00:51:22.712 --> 00:51:25.690 this is still a very open area

00:51:25.690 --> 00:51:28.250 that a lot of work is needed.

00:51:28.250 --> 00:51:30.840 Even for this prototypical example

00:51:30.840 --> 00:51:33.410 that people have studied for decades,

00:51:33.410 --> 00:51:35.910 it seems there's still a lot of questions

00:51:35.910 --> 00:51:37.113 that are unresolved.

00:51:38.070 --> 00:51:41.342 And there are methods that need to be developed

00:51:41.342 --> 00:51:44.860 for this sensitivity analysis

00:51:44.860 --> 00:51:48.030 to be regularly used in practice.

00:51:48.030 --> 00:51:50.850 And then there are many other related problems

00:51:50.850 --> 00:51:52.810 in missing data in causal inference

00:51:53.730 --> 00:51:57.703 that need to see more developments of sensitivity analysis.

00:51:58.810 --> 00:52:00.820 So that's the end of my talk.

00:52:00.820 --> 00:52:03.423 And there are some references that are used.

00:52:04.630 --> 00:52:08.140 I'm happy to take any questions.

00:52:08.140 --> 00:52:11.167 Still have about four minutes left.

00:52:11.167 --> 00:52:12.641 - Thank you.

00:52:12.641 --> 00:52:14.307 That yeah, thank you.

00:52:14.307 --> 00:52:17.180 Thank you, I'm sorry I couldn't introduce you earlier,

00:52:17.180 --> 00:52:20.627 but my connection but it did not to work.

00:52:20.627 --> 00:52:23.153 So we have time for a couple of questions.

00:52:25.560 --> 00:52:28.820 You can write the question in the chat box,

00:52:28.820 --> 00:52:30.663 or just unmute yourselves.

00:52:43.482 --> 00:52:44.315 Any questions?

00:52:53.670 --> 00:52:55.962 I guess I'll start with a question.

00:52:55.962 --> 00:53:00.120 Yeah I guess I'll start with a question.

00:53:00.120 --> 00:53:03.890 This was a great connection between I think,

00:53:03.890 --> 00:53:05.770 sensitivity analysis literature

00:53:05.770 --> 00:53:07.493 and the missing data literature.

00:53:08.600 --> 00:53:11.982 Which I think it's kind of overlooked.

00:53:11.982 --> 00:53:16.982 Even when you when you run a prometric sensitivity analysis,

00:53:17.360 --> 00:53:20.270 it's really something, like most of the times

00:53:20.270 --> 00:53:22.060 people really don't understand

00:53:22.060 --> 00:53:24.023 how much information is given.

00:53:24.860 --> 00:53:28.560 Like, how much information the model actually gives

00:53:29.400 --> 00:53:31.660 on the sensitivity parameters.

00:53:31.660 --> 00:53:34.080 And as you said,

00:53:34.080 --> 00:53:35.530 like it's kind of inconsistent

00:53:35.530 --> 00:53:37.470 to set the sensitivity parameters

00:53:37.470 --> 00:53:40.230 when sensitivity parameters are actually identified

00:53:40.230 --> 00:53:41.193 by the model.

00:53:42.940 --> 00:53:46.190 So I think like my I guess a question of like,

00:53:46.190 --> 00:53:47.603 clarifying question is,

00:53:48.670 --> 00:53:53.660 you mentioned there is this there this testable models,

00:53:53.660 --> 00:53:55.763 this testable models essentially are wherein

00:53:55.763 --> 00:53:59.690 the sensitivity model is such that

00:53:59.690 --> 00:54:03.620 the sensitivity barometer are actually point identified.

00:54:03.620 --> 00:54:04.453 Right?

00:54:04.453 --> 00:54:05.286 - Yes.

00:54:05.286 --> 00:54:07.535 So it re, so you said,

00:54:07.535 --> 00:54:10.850 you reshooting use the sensitivity analysis

00:54:10.850 --> 00:54:13.800 to actually to set the parameters

00:54:13.800 --> 00:54:16.141 if the sensitivity parameters

00:54:16.141 --> 00:54:18.000 are actually identified model.

00:54:18.000 --> 00:54:18.833 - Yeah.

00:54:18.833 --> 00:54:20.690 - Is that what you're trying?

00:54:20.690 --> 00:54:23.480 All right, so and. - Yes, yeah.

00:54:23.480 --> 00:54:27.300 Basically what happened there is the model is too specific,

00:54:27.300 --> 00:54:29.830 and it wasn't constructed carefully.

00:54:29.830 --> 00:54:32.570 So it's possible to construct parametric models

00:54:32.570 --> 00:54:36.770 that are not testable that are perfectly fine.

00:54:36.770 --> 00:54:40.310 But sometimes, if you just sort of

00:54:40.310 --> 00:54:42.170 write down the most natural model,

00:54:42.170 --> 00:54:46.400 if it just extend what the parametric model

00:54:46.400 --> 00:54:50.883 you used for observed data to also model full data,

00:54:52.100 --> 00:54:53.780 then you don't do it carefully,

00:54:53.780 --> 00:54:58.780 then the entire full data distribution becomes identifiable.

00:54:59.530 --> 00:55:02.240 So it does makes sense to treat those parameters

00:55:02.240 --> 00:55:04.580 as sensitivity parameters.

00:55:04.580 --> 00:55:08.190 So this kind of is a reminiscent of the discussion

00:55:08.190 --> 00:55:10.753 in the 80s about the Hackmann selection model.

00:55:11.690 --> 00:55:13.690 Because in that case,

00:55:13.690 --> 00:55:18.291 there was also sir Hackmann has this great selection model

00:55:18.291 --> 00:55:23.200 for reducing or getting rid of selection bias,

00:55:23.200 --> 00:55:26.560 but it's based on very heavy parametric assumptions.

00:55:26.560 --> 00:55:31.560 And you can adapt certainly identify the selection effect

00:55:31.690 --> 00:55:35.040 directly from the model where you actually have no data

00:55:35.890 --> 00:55:38.203 to support that identification.

00:55:39.693 --> 00:55:43.513 Which led to some criticisms in the 80s.

00:55:44.950 --> 00:55:49.950 But I think we are seeing this things repeatedly

00:55:50.600 --> 00:55:53.223 again and again in different areas.

00:55:54.940 --> 00:55:58.910 And it's, I think it's fine

00:55:58.910 --> 00:56:03.910 to use the power metric models that are testable, actually,

00:56:05.090 --> 00:56:07.240 if you really believe in those models,

00:56:07.240 --> 00:56:09.370 but it doesn't seem that they should be used

00:56:09.370 --> 00:56:11.590 this sensitivity analysis,

00:56:11.590 --> 00:56:13.050 because just logically,

00:56:13.050 --> 00:56:14.483 it's a bit strange.

00:56:15.331 --> 00:56:18.483 It's hard to interpret those models.

00:56:20.493 --> 00:56:24.347 And but sometimes I've also seen people

00:56:24.347 --> 00:56:27.650 who use the sort of parameterize the model

00:56:27.650 --> 00:56:30.950 in a way that you include enough terms.

00:56:30.950 --> 00:56:34.510 So the sensitivity parameters are weakly identified

00:56:34.510 --> 00:56:37.650 in a practical example.

00:56:37.650 --> 00:56:42.590 So with a practical data set of maybe the likelihood test,

00:56:43.530 --> 00:56:46.330 Likelihood Ratio Test rejection region,

00:56:46.330 --> 00:56:49.083 that acceptance region is very, very large.

00:56:50.170 --> 00:56:53.420 So there are a suggestions like that,

00:56:53.420 --> 00:56:58.330 that kind of it's a sort of a compromise

00:56:58.330 --> 00:57:01.533 for good practice.

00:57:02.430 --> 00:57:06.330 - Right in that case you gave it either set the parameters

00:57:06.330 --> 00:57:08.520 and drag the causal effects,

00:57:08.520 --> 00:57:13.370 or kind of treat that as a partial identification problem

00:57:13.370 --> 00:57:16.800 and just write use bounds or the methods

00:57:16.800 --> 00:57:18.723 you were mentioning, I guess.

00:57:19.908 --> 00:57:21.369 - Yeah.

00:57:21.369 --> 00:57:22.536 - Yep, thanks.

00:57:25.720 --> 00:57:26.713 Other questions?

00:57:34.367 --> 00:57:36.788 Well I guess you can read the question?

00:57:36.788 --> 00:57:39.763 - It's a question from Kiel Sint.

00:57:40.687 --> 00:57:42.997 Sorry if I didn't pronounce your name correctly.

00:57:42.997 --> 00:57:45.620 ”In the applications of observational studies ideally,

00:57:45.620 --> 00:57:47.440 what confounders should be collected

00:57:47.440 --> 00:57:49.440 for sensitivity analysis,

00:57:49.440 --> 00:57:53.600 power sensitivity analysis for unmeasured confounding?”

00:57:53.600 --> 00:57:54.433 Thank you.

00:57:54.433 --> 00:57:58.453 So if I understand your question correctly,

00:58:01.250 --> 00:58:03.780 basically what sensitivity analysis does

00:58:03.780 --> 00:58:05.910 is you have observational study,

00:58:05.910 --> 00:58:08.870 where you for already collected confounders

00:58:09.730 --> 00:58:12.010 that you believe are important or relevant

00:58:13.167 --> 00:58:16.340 that really that are real confounders,

00:58:16.340 --> 00:58:20.280 that they change the causal unchanged the treatment

00:58:20.280 --> 00:58:22.200 and the outcome.

00:58:22.200 --> 00:58:24.540 But often that's not enough.

00:58:24.540 --> 00:58:29.387 And what sensitivity analysis does is it tries to say,

00:58:29.387 --> 00:58:31.690 "based on what the components already

00:58:32.927 --> 00:58:34.210 you have already collected,

00:58:34.210 --> 00:58:36.840 what if there is still something missing

00:58:36.840 --> 00:58:38.153 that we didn't collect?

00:58:39.480 --> 00:58:43.508 And then if those things behave in a certain way,

00:58:43.508 --> 00:58:46.640 does that change our results?"

00:58:46.640 --> 00:58:51.640 So, I guess sensitivity analysis is always relative

00:58:52.110 --> 00:58:53.930 to a primary analysis.

00:58:53.930 --> 00:58:58.200 So I think you should use the same set of confounders

00:58:58.200 --> 00:59:00.583 that the primary analysis uses.

00:59:02.396 --> 00:59:07.396 I don't see a lot of reasons to vary to say

00:59:09.910 --> 00:59:14.160 use a primary analysis with more confounders,

00:59:14.160 --> 00:59:17.543 but a sensitivity analysis with fewer confounders.

00:59:20.800 --> 00:59:23.340 Sensitivity analysis is really a supplement

00:59:23.340 --> 00:59:26.373 to what you have in the primary analysis.

00:59:35.420 --> 00:59:37.220 - Just one more question if we have?

00:59:39.551 --> 00:59:40.500 There not.

00:59:40.500 --> 00:59:41.333 Yes.

00:59:42.500 --> 00:59:44.027 - So from Ching Hou Soo,

00:59:45.447 --> 00:59:48.740 "How to specify the setup sensitivity parameter gamma

00:59:48.740 --> 00:59:50.193 in the real life question?

00:59:51.220 --> 00:59:53.730 When gamma is too large the inference results

00:59:53.730 --> 00:59:57.000 will always be non informative?"

00:59:57.000 --> 00:59:59.787 Yes, this is always a tricky problem any,

01:00:01.140 --> 01:00:05.930 and essentially the sensitivity values kind of

01:00:05.930 --> 01:00:08.560 trying to get past that.

01:00:08.560 --> 01:00:11.220 So it tries to directly look at the value

01:00:11.220 --> 01:00:15.210 of this sensitivity parameter that changes your conclusion.

01:00:15.210 --> 01:00:18.810 So in some sense, you don't need to specify

01:00:18.810 --> 01:00:20.173 a parameter a priori.

01:00:21.040 --> 01:00:24.760 But obviously, in the end of the day,

01:00:24.760 --> 01:00:29.760 we need some clue about what value of sensitivity parameter

01:00:30.020 --> 01:00:31.360 is considered large.

01:00:31.360 --> 01:00:35.500 In a practical sense, in this application.

01:00:35.500 --> 01:00:39.156 That's something this calibration clause

01:00:39.156 --> 01:00:43.620 this calibration analysis is trying to address.

01:00:43.620 --> 01:00:44.490 But as I said,

01:00:44.490 --> 01:00:47.310 they're not perfect at the moment.

01:00:47.310 --> 01:00:52.310 So for some time, now, at the least,

01:00:52.360 --> 01:00:55.830 we'll have to sort of live through this and

01:00:55.830 --> 01:01:00.600 or will either need to understand really

01:01:00.600 --> 01:01:02.180 what the sensitivity model means,

01:01:02.180 --> 01:01:06.010 and then use your domain knowledge

01:01:06.010 --> 01:01:10.460 to set the sensitivity parameter,

01:01:10.460 --> 01:01:15.460 or we have to use these rely on these

01:01:15.471 --> 01:01:19.323 imperfect visualization tools to calibrate analysis.

01:01:27.689 --> 01:01:28.620 - Yeah, all right.

01:01:28.620 --> 01:01:29.453 Thank you.

01:01:30.372 --> 01:01:33.209 I think we need to wrap up we've run over time.

01:01:33.209 --> 01:01:36.040 So thank you again Qingyuan,

01:01:36.040 --> 01:01:37.850 for sharing your work with us.

01:01:37.850 --> 01:01:40.672 And thank you, everyone for joining.

01:01:40.672 --> 01:01:41.780 Thank you.

01:01:41.780 --> 01:01:42.960 Bye bye.

01:01:42.960 --> 01:01:44.222 See you next week.

01:01:44.222 --> 01:01:45.325 - It's a great pleasure.

01:01:45.325 --> 01:01:46.158 Thank you.