09/14/26: Metrics lunch, Zhiheng You

Ask Prof Diebold about high-dim'l asymptotics for generalized GMM / sieve GMM? Also ask about his Sargent stories.

"How to think about 'restrictiveness measure' for blackbox models?" [SJ: For what purpose?] GP model?

Restrictiveness measure: given a model $f : \mathcal{X} \to \mathcal{Y}$, some parametric family $\mathcal{F}_\theta$, some eligible set $\mathcal{F}$, test distribution $\lambda_\mathcal{F}$ over $\mathcal{F}$, and define $r(\mathcal{F}_\theta; \mathcal{F}) = \mathbb{E}[d(\mathcal{F}_\Theta, f)]/\mathbb{E}[d(\mathcal{F}, f)]$

[SJ: Very intuitive. Think back to math-stat.]

[SJ: Diebold question: where are dof's? Schorfheide: this is all population level, bad.]

[SJ: The sampling $f \sim \lambda_\mathcal{F}$? Rewrite as a ratio of distances between $\lambda_\mathcal{F}$ and $\text{Unif}(\mathcal{F}_\Theta)$ and similarly for $\text{Unif}(\mathcal{F}_{\text{base}})$?]

Completeness measure (complementary to the above): basically generalization of R^2, given by $\frac{\mathcal{E}(f_n)-\mathcal{E}(f^*_\Theta)}{\mathcal{E}(f_n)-\mathcal{E}(f^*)}$ where f_n is a "naive" benchmark.

[SJ: In what sense are these complementary? They seem to be very correlated. Speaker clarifies that in former, we are sampling *predictors* from $\lambda_\mathcal{F}$, our *belief* on what the predictor may be; whereas in the latter, we are computing generalization error with respect to the test data distribution.]

[SJ: We distinguish $\lambda_\mathcal{F}$ from P. Schorfheide: gives concrete example with AR models.]

[SJ: Xu: what is a J-statistic?]

How to draw $f \sim \lambda_\mathcal{F}$ when $\mathcal{F}$ is infinite-dim'l? Use a GP, BNP prior, Matérn 3/2 covariance kernel. We can impose shape constraints like boundedness, monotonicity, convexity. [SJ: Cool.]

[SJ: Wayne: alternative is tabular foundation model as generative prior, somehow trained on the internet.]

[SJ: Do you prove a theorem that bounds the effect of perturbing your prior on your results? What is the standard in BNP? "Sensitivity analysis".]

Ex: predict the certainty equivalent: $f(x) = z_l + (z_u - z_l) \cdot \sigma(z_u(x))$.

Ex: predict demand and supply via reduced form model $Q = a + bP + cX + \varepsilon$ and $P = a + bQ + cX + \varepsilon$. Rewrite this as $Y = AX + \varepsilon$ and compute all the aforementioned quantities. [SJ: Here, we need to impose slope constraints due to the natural economic relation between demand and supply. Need to understand map from structural to reduced.] [SJ: Schorfheide: structural AR literature have a lot of debates about this about priors. What's a "reasonable" distribution? The implications can vary dramatically on demand and supply. At what point do we impose the natural economic constraints on these two variables?]

Ex: entry with strategic substitutes, via best equilibrium selection. [SJ: ?] $y_1 = \mathbf{1}\{a + by_2 + cx_1 \geqslant \varepsilon_1\}$ and $y_2 = \mathbf{1}\{a + by_1 + cx_2 \geqslant \varepsilon_2\}$. [SJ: So y_i indicates whether participant i enters the market, based on what the other participant does?] Both $(1, 0)$ and $(0, 1)$ can be equilibria. An eligible prediction is the four-vector of probability predictions.

Ex: Berry–Levinsohn–Pakes (BLP) demand.

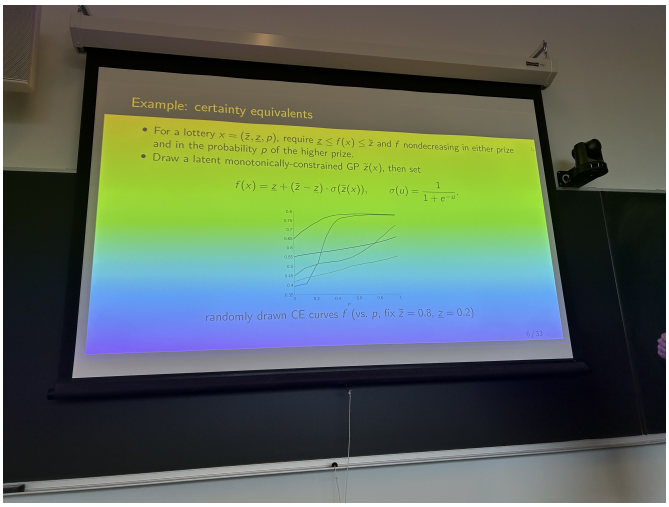


Figure 1: Lottery, with (implicitly defined) GP on models f.

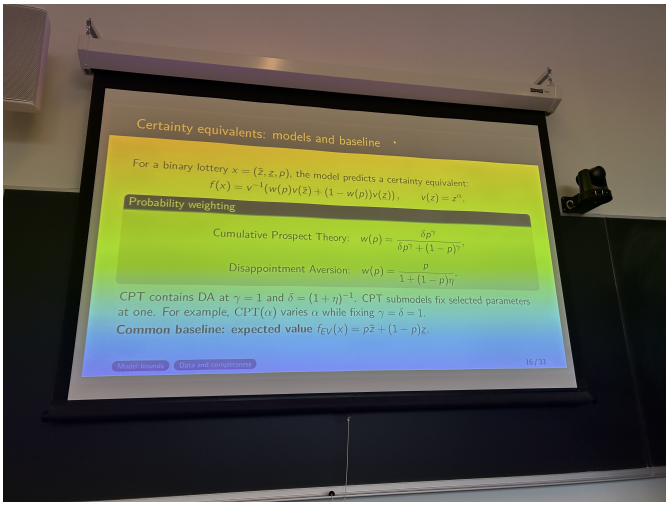


Figure 2: Lottery, results.