
Clase 190 — Uplift modeling, DiD (difference-in-differences)

Parte: 3 — Estadística Inferencial y Causal · Fuente: Gutierrez & Gerardy (2017), Causal Inference and Uplift Modeling + Abadie, Diamond & Hainmueller (2010), Synthetic Control Methods. Duración estimada: 90 min.

Clase 190 — Uplift modeling, DiD (difference-in-differences)

Parte: 3 — Estadística Inferencial y Causal · Fuente: Gutierrez & Gerardy (2017), Causal Inference and Uplift Modeling + Abadie, Diamond & Hainmueller (2010), Synthetic Control Methods. Duración estimada: 90 min.

Objetivo

Dominar las dos técnicas causales más usadas en industria cuando hay datos panel/observacionales: DiD (Difference-in-Differences) —comparar la evolución antes/después en grupo tratado vs control— y uplift modeling —predecir a quién conviene tratar (heterogeneidad del efecto causal a nivel individuo). Conocer el complemento moderno Synthetic Control Method para cuando no hay grupo de control natural y solo se trata una unidad (una ciudad, un país).

Resultados de aprendizaje

Al finalizar, el estudiante podrá:

- Aplicar DiD clásico con OLS: $Y = \beta + \beta \cdot \text{tratado} + \beta \cdot \text{post} + \beta \cdot (\text{tratado} \times \text{post}) + \epsilon$. El coeficiente β es el efecto causal bajo parallel trends.
- Diagnosticar la asunción de parallel trends con un gráfico antes/después y un placebo test.
- Construir modelos de uplift: T-learner, S-learner, X-learner (Künzel et al. 2019), Causal Forest (econml).
- Evaluar uplift con Qini curve y uplift@k (no con AUC clásico — uplift es individual, no global).
- Aplicar Synthetic Control con pysyncon o SparseSC cuando una sola unidad recibe tratamiento (estudio de caso).

Temas

- DiD: comparación dos-por-dos en panel (2 grupos × 2 tiempos).
- Asunción crítica: parallel trends — sin tratamiento, ambos grupos hubieran evolucionado paralelos.
- Generalización: DiD con muchos tiempos, two-way fixed effects (TWFE), event study designs.
- Uplift = CATE individual = $E[Y(1) - Y(0) | X=x]$.
- 4 cuadrantes de uplift: persuadables, sure things, lost causes, do-not-disturb (no tocarlos).
- Métricas: Qini, uplift@k, AUUC (area under uplift curve).
- Complemento moderno: Synthetic Control Method (Abadie et al.) — construye un "país sintético" como combinación convexa de unidades no tratadas que replica la trayectoria pre-tratamiento.

Versión profundizada — 2026

El tema moderno que antes vivía como complemento dentro de esta clase ahora tiene su(s) clase(s) propia(s) con patrón completo, ejercicios y homework:

- Clase 157a — Synthetic Control Method dedicado (pysyncon, SparseSC)

Definiciones y características

- DiD: estima $(Y_{\text{tratado,post}} - Y_{\text{tratado,pre}}) - (Y_{\text{control,post}} - Y_{\text{control,pre}})$. Equivale al coeficiente de la interacción $\text{tratado} \times \text{post}$ en OLS.
- Parallel trends: sin el tratamiento, ambos grupos hubieran tenido la misma trayectoria. NO testeable directamente; sí en el pre.
- TWFE (Two-Way Fixed Effects): regresión con dummies de unidad + dummies de tiempo + tratamiento. Generaliza DiD a panel.
- Event study: grafica los coeficientes de $\text{dummy} \times (t - t_{\text{tratamiento}})$ para cada t . Permite ver dinámica del efecto y testear pre-tendencias.
- Uplift / CATE: $\tau(x) = E[Y(1) - Y(0) | X=x]$. Diferencia entre lo que pasa con tratamiento vs sin.
- T-learner: ajustar dos modelos separados, uno por grupo ($\mu(x) = E[Y|X, T=1]$, $\mu(x) = E[Y|X, T=0]$); restar.
- S-learner: un solo modelo con T como feature; predecir con $T=1$ y $T=0$ y restar.
- X-learner (Künzel et al. 2019): combina T y S , usando propensity para ponderar; mejor con grupos desbalanceados.
- Causal Forest (econml.CausalForestDML): random forest entrenado para minimizar heterogeneidad del efecto, con IC válidos.
- Qini curve: ranking individuos por uplift predicho; eje x = % tratado; eje y = ganancia incremental acumulada vs random.

Dataset / recursos

- DiD clásico: Card & Krueger 1994 (mínimum wage en NJ vs PA).
- Uplift: Hillstrom email dataset (criteo), Lenta uplift dataset.
- Synthetic Control: California Prop 99 (smoking) — clásico de Abadie.
- Librerías: linearmodels (PanelOLS), econml, causalml (Uber), pysyncon, SparseSC.

Ejercicios

1. DiD ingenuo: simulá panel con 2 grupos y 2 períodos. Aplicá DiD con OLS y verificá que β recupera el efecto verdadero. Probá violar parallel trends y ver el sesgo.
2. Event study: con panel 10 períodos (5 pre, 5 post), graficá coeficientes por período. Si los pre son ≈ 0 → parallel trends plausible.
3. T-learner: en Hillstrom (binario tratamiento email), entrená dos RandomForestClassifier y predecí uplift = $p(x) - p(x)$.
4. Qini curve: con las predicciones del ej. 3, calculá Qini con `sklift.metrics.qini_score` o a mano. Compará contra "tratar al azar".
5. Synthetic Control: con un dataset panel simulado (10 estados \times 20 años, tratamiento en California año 11), ajustá pesos con `pysyncon` y graficá `path_plot` + `gaps_plot`. Aplicá `placebo_test`.

Homework verificable

Sobre el dataset Card & Krueger (mínimum wage NJ vs PA, 1992):

1. Cargar datos, calcular promedio de empleo por estado en pre (Feb 1992) y post (Nov 1992).
2. Aplicar DiD con OLS y reportar β con IC95 %.
3. Hacer un event study si hay más de 2 períodos disponibles, o el gráfico 2×2 si no.
4. Discutir en 4 líneas: ¿se cumple parallel trends? ¿Cómo se relaciona el β con el debate clásico (Card vs Neumark)?

Criterio de aceptación: el coeficiente DiD debe ser positivo y $\approx +2.7$ empleados (consistente con el paper original); la discusión menciona parallel trends y un riesgo concreto (sesgo de selección de stores, attrition).

Errores comunes

Síntoma / mensaje	Causa y cómo arreglar	
Aplico DiD sin verificar parallel trends	Si los grupos tenían trayectorias divergen	
Uplift evaluado con AUC	AUC mide clasificación, no uplift. Un mode	X) puede ser pésimo para uplift. Fix: Qin
T-learner con un grupo mucho más chico que	El modelo del grupo chico overfittea. Fix:	
Synthetic control con pre-período de 2 año	Pesos no convergen a algo confiable. Fix:	
Interpreto el sintético como "lo que hubie	Sin placebo, no podés saber si tu "efecto"	

Preguntas frecuentes

¿DiD funciona con un solo período pre y un solo post?

Sí (es el "2x2 DiD"), pero no podés testear parallel trends — solo asumirlo. Con más períodos, podés hacer event study y verificarlo.

¿Uplift modeling vs causal forest?

Causal forest es uplift modeling — produce CATE por instancia con IC. Las otras (T/S/X-learner) son métodos clásicos. En la práctica, X-learner y causal forest tienen el mejor performance en benchmarks.

¿Cuántas unidades de control mínimas para synthetic control?

10-20 idealmente; con menos los pesos saturan en 1 o 2 unidades y pierde robustez. Si pocas unidades pero muchos períodos: synthetic DiD o factor models.

¿Negative weights en synthetic control?

El método clásico fuerza $w \geq 0$ (combinación convexa). Variantes modernas (Doudchenko & Imbens 2016, SparseSC) relajan esto y permiten negativos con regularización — más flexible pero menos interpretable.

¿DiD con TWFE es siempre correcto?

No con tratamientos escalonados (unidades tratadas en momentos distintos). TWFE puede pesar con signo negativo períodos donde unidades "ya tratadas" sirven de control de "recién tratadas". Fix: Callaway & Sant'Anna 2021, did package en R o differences en Python.

Referencias

- Angrist & Pischke (2009), Mostly Harmless Econometrics, cap. 5 (DiD).
- Künzel, Sekhon, Bickel & Yu (2019), Metalearners for estimating heterogeneous treatment effects using ML, PNAS.
- Abadie, Diamond & Hainmueller (2010), Synthetic Control Methods, JASA.
- Callaway & Sant'Anna (2021), Difference-in-Differences with Multiple Time Periods, J. of Econometrics.
- Gutierrez & Gerardy (2017), Causal Inference and Uplift Modeling, JMLR Workshop.
- econml, causalml (Uber), pysyncon, SparseSC.

Siguiente clase

Clase 191 — Synthetic Control Method dedicado (pysyncon, SparseSC)

Apéndice: notebook (primer bloque)

Primera celda ejecutable del notebook de la clase.

```
# Imports y configuración inicial
```

Archivos complementarios

- notebook.ipynb