ORIE 5355

Lecture 8: Introduction to Algorithmic Pricing

Nikhil Garg

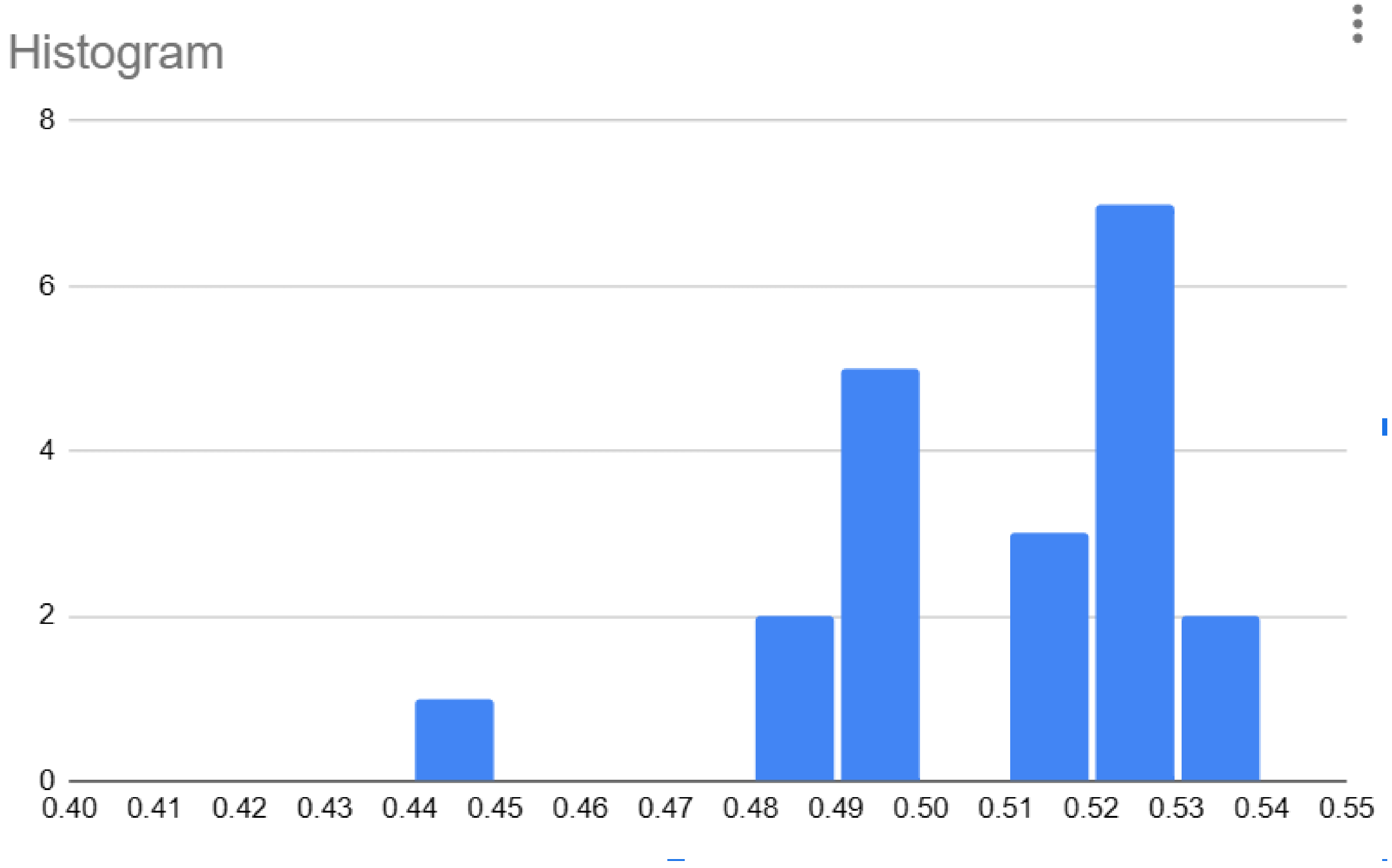
Announcements

HW 2 released, due 10/1

Guest lectures next week – show up in person!

Recommendations and NYC’s 311 system

HW1 Final estimates



Summary of recommendations

There are 3 steps to building a recommendation system:

- Choose the *data* that you will use
  - What does the data imply about people's opinions and future desires?
- Train a model to *predict* ratings between pairs of items and users
  - Different approaches (item- and user similarity, matrix factorization)
  - Can also combine approaches
- *Recommend* items based on predictions and other concerns
  - Capacity constraints, diversity, fairness considerations, long-term objectives

Questions from recommendations?

Algorithmic Pricing

Module Overview

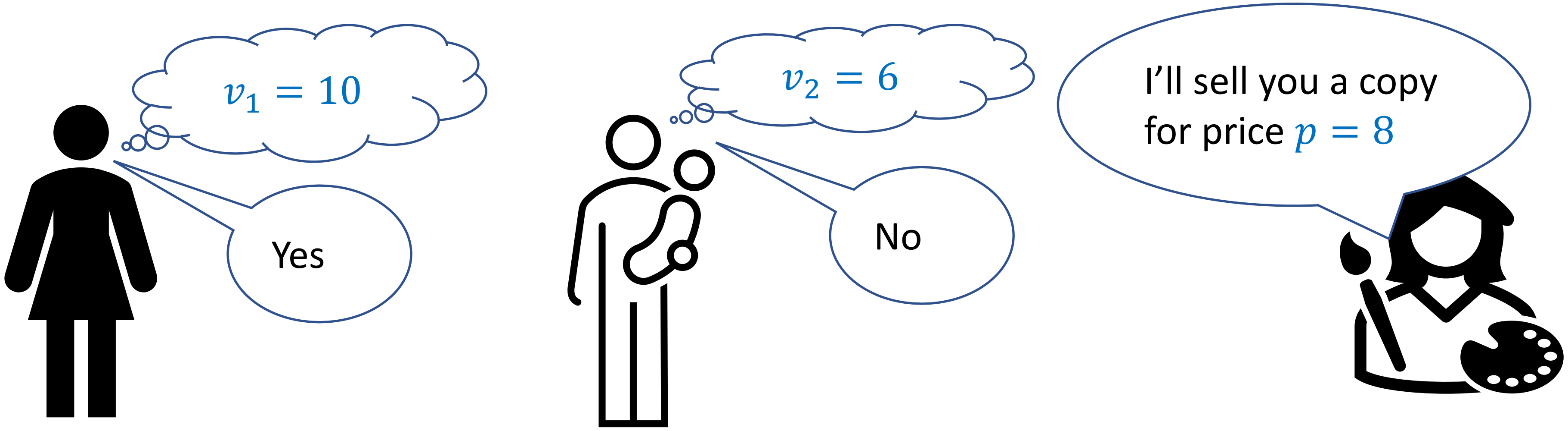
- Basics of pricing and algorithmic pricing
  - Pricing under uncertainty or heterogeneous valuations in population
  - Demand estimation at different prices
- Challenges from practice:

  Capacity constraints, dynamics, competition, selling multiple items (cannibalization)
- Extended case-study: Pricing in online marketplaces [Ride-hailing]
- Ethics, Transparency, and Bias in algorithmic pricing

User model and omniscient pricing

Simple user behavior model

- Suppose you're selling 1 type of item
- Each person i has a *private valuation* v_i for that item
- Suppose you offer the item at price p
- Person i buys the item if $v_i \geq p$
- Omniscient pricing: maximize revenue by setting $p_i = v_i$



Maximizing profit via machine learning

- Omniscient pricing: maximize revenue by setting $p_i = v_i$
- Challenge: we don't know valuation v_i for each person
- Ok, let's just use a machine learning approach!
  - Create an estimate $\hat{v}_i$ for value for person i using historical data
    - KNN, regression, whatever
  - Set price $p_i = \hat{v}_i$
- Problem: the above approach *miserably fails*!

Why does the naïve method fail?

- Your estimated valuation $\widehat{v}_i$ is not perfect
- Example: Suppose the true valuation $v_i = 10$
  - What is your revenue if $\widehat{v}_i = p = 9$?

    Answer: 9
  - What is your revenue if $\widehat{v}_i = p = 11$?

    Answer: 0
- Under the simple behavior model, *small errors* in guessing valuation $\widehat{v}_i$ can have *huge revenue implications*
- Must incorporate *uncertainty* in your pricing decisions!

(You also don't have great data to estimate $\widehat{v}_i$…)

Optimal pricing with uncertainty

"Posted price mechanisms" and personalized pricing

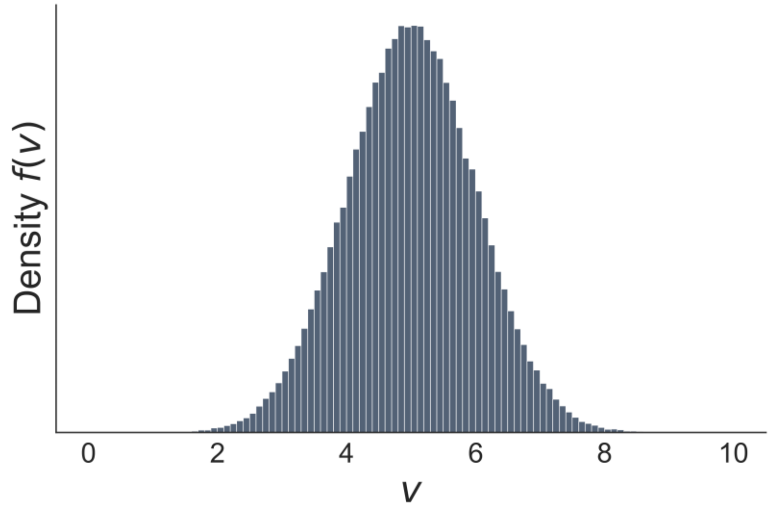
Challenge

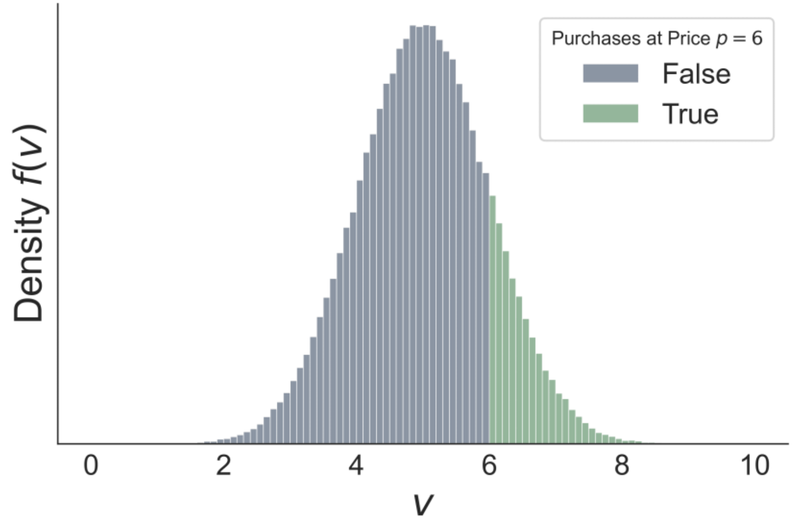
- There is a lot of randomness in whether someone purchases at a given price. Multiple ways to think about it:
  - You have a single price p for the entire population, but people differ in in their valuations v_i (heterogeneity)
  - You do personalized pricing p_i, but your estimate $\widehat{v_i}$ is not perfect (noise)
- Why is this a problem?
  - In recommendations, we ignored noise. Why not ignore it here?
  - Here, dealing with noise is crucial if we want to maximize revenue, even “in expectation”

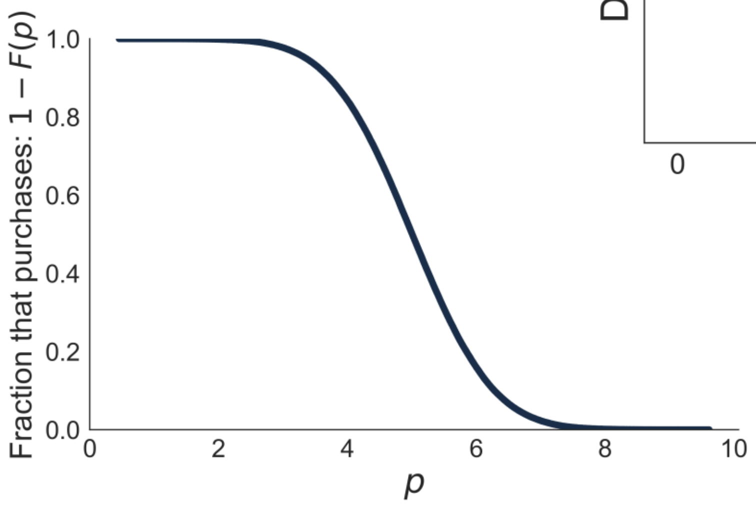
Model

- Here, let's suppose we are posting single price p for entire population
- We have unlimited copies of the item
- Suppose we have a distribution F for the users' valuations: for each user i, valuation $v_i \sim F$
- If we set price p:
  - Each individual with valuation $v_i \geq p$ purchases
  - Overall, fraction $1 - F(p)$ purchases

$\mathrm{d(p)} = 1 - F(p)$ is called the "demand" at price p







Maximizing revenue

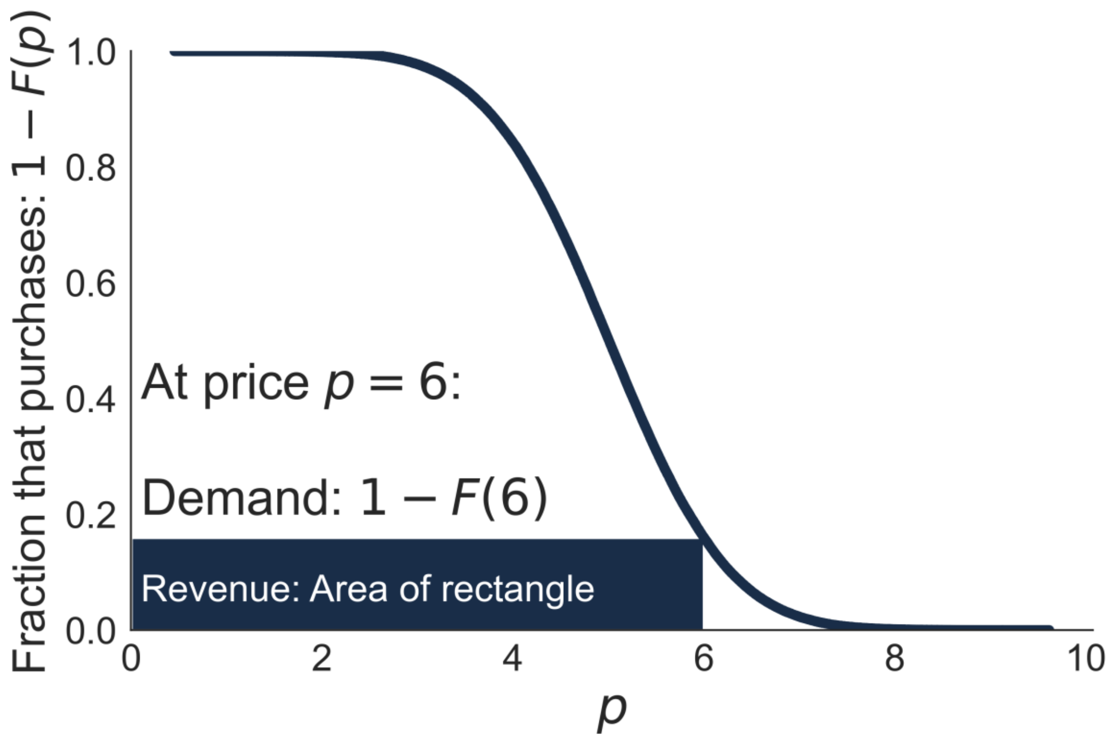
- Expected revenue at price p:

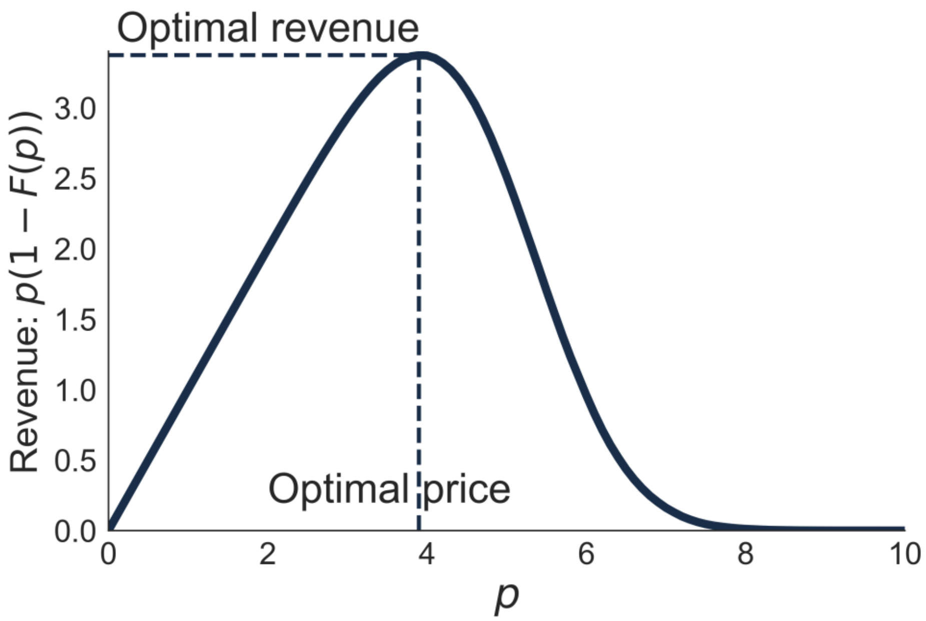
  [Revenue from each sale] x [Demand at price p]

  $$p(d(p))$$

- Revenue maximizing price:

  $$\text{argmax}_p \; p(d(p))$$





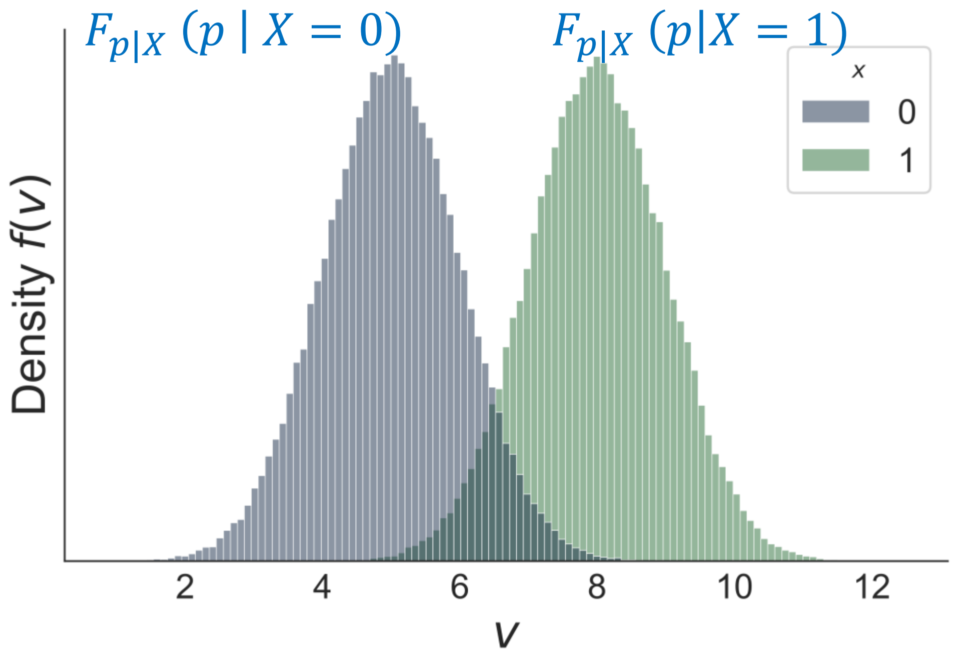
Personalized pricing: Price differentiation via covariates

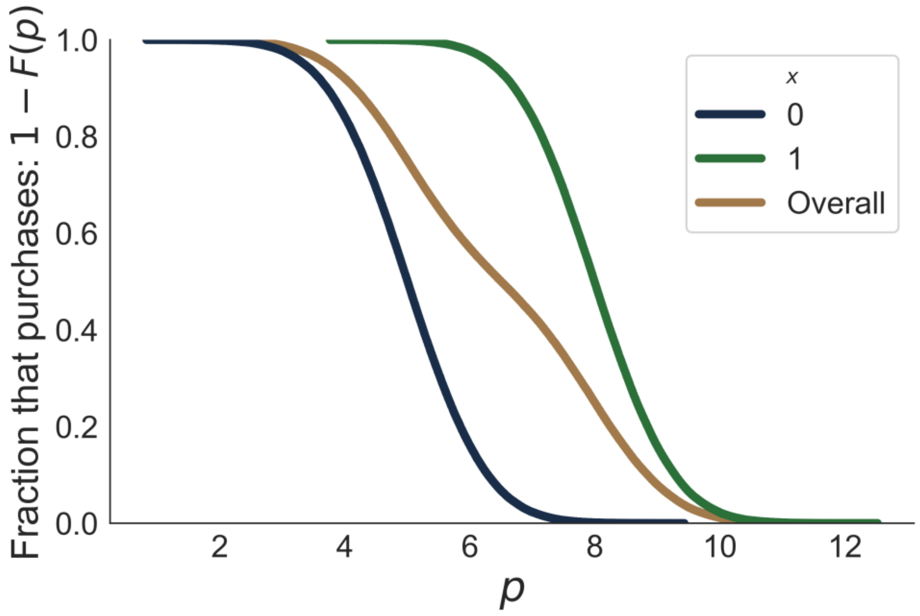
- So far: given the population valuation distribution F, we can find the price p that maximizes revenue: $\text{argmax}_p \, p\big(1 - F(p)\big)$
- Now, suppose we have covariates x_i for each potential customer, and we are allowed to give show different prices to different people
  - Prices by geography (neighborhood)
  - Student or senior citizen discounts
- Now, given the *conditional* distributions $F_{p|X}\,(p \mid X = x)$, simply create a price $p(x)$ that maximizes revenue

$$p(x) \times (1 - F_{p|X}\,(p \mid X = x))$$

Example

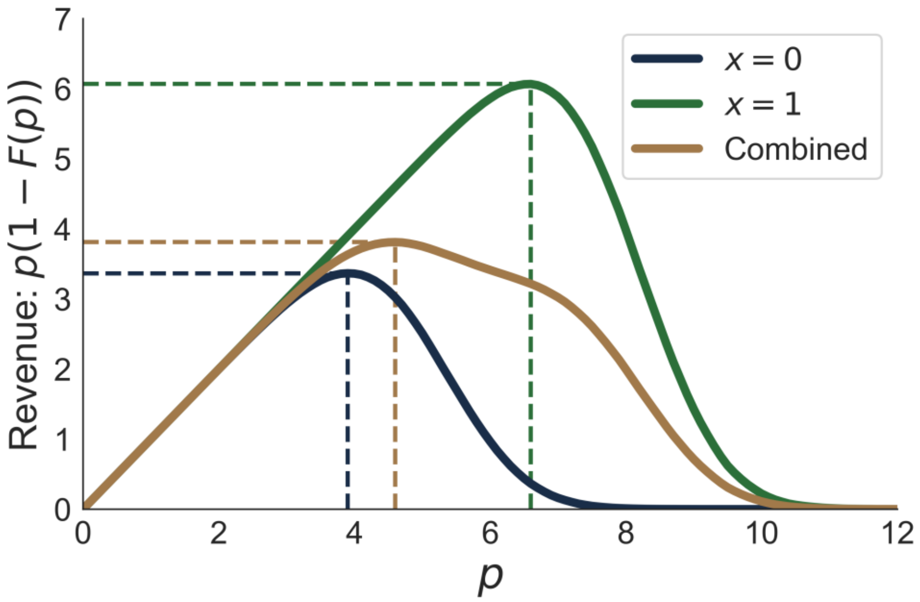
- Suppose we have a binary covariate, $x_i \in \{0, 1\}$. Population evenly split
- Valuation distributions differ
- And then purchase probabilities at each price p also differ





Example cont.

- If we don't have any capacity constraints on the item, we can simply find optimal prices independently for the two customer types
- Value of personalized pricing
  - Revenue from single price: 3.81
  - Revenue from separate prices: 4.72
- Things get more complicated if there are capacity constraints (next time)



Questions?

Demand (distribution) estimation

The challenge

- So far, we've talked about calculating optimal prices if we knew the demand distribution $F(p)$, or the conditional demand distributions $F_{p|X}\,(p \mid X = x)$
- We don't know these distributions! Need to learn them from data
- What does data look like? We never see valuations, just purchase decisions at historical prices p
- Assumption: we see decisions at many prices p

	Location	Income level	Offered price	Purchased
0	Africa	4.40	4.70	False
1	Europe	21.83	0.61	True
2	America	37.60	3.37	True
3	Europe	17.90	1.91	True
4	Africa	9.45	1.57	False
5	Europe	1.45	4.28	False
6	Europe	19.63	3.00	True
7	Europe	15.76	4.44	False
8	Europe	5.87	6.25	False
9	America	30.31	0.51	True

Naïve approach: Empirical Distribution

- Goal: estimate $\mathrm{d(p)} = 1 - F(p)$ for each p in a "reasonable range" of prices
- Naïve approach:
  - Bin the historical prices offered
  - In each bin, construct estimate $\widehat{\mathrm{d(p)}}$ as the fraction of offers in that bin that were accepted

$$\widehat{\mathrm{d(p)}} = \frac{\text{\# offers accepted}}{\text{\# offers}}$$

- When estimating $F_{p|X}\,(p \mid X = x)$, simply do the same thing but for each set of covariates

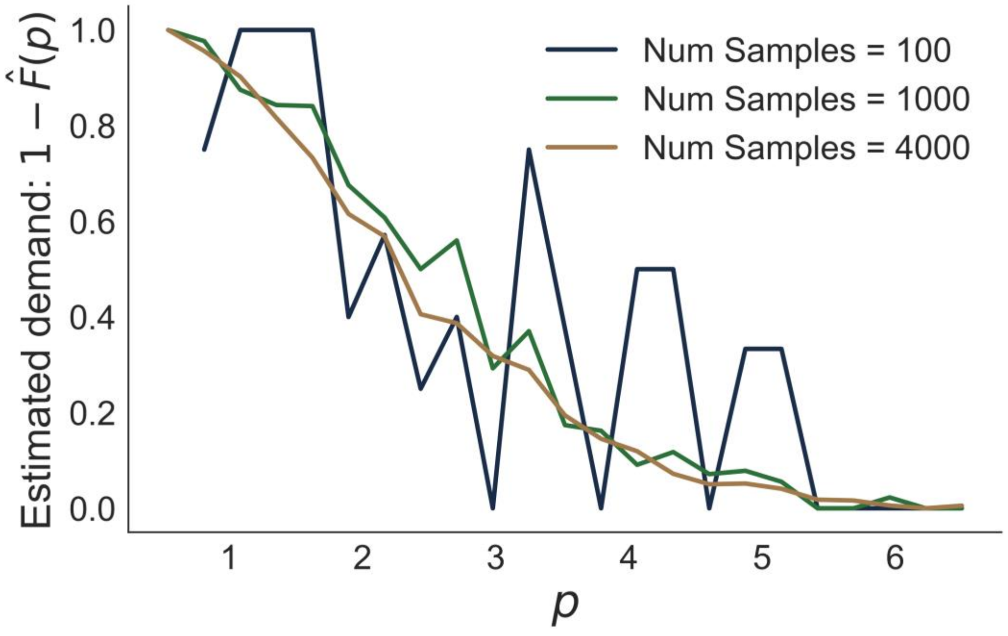
Naïve method pros and cons

Pros:

- Simple to implement
- “Non-parametric” – no assumptions
- As # of historical samples → ∞, converge to truth

Cons:

- Wastes data: only use data for that given price bin and for that given covariate
- Requires many samples



Exactly the same as naïve mean estimation in polling!

Fancier methods: machine learning

- We want to estimate $\mathrm{d}(\mathrm{p}, \mathrm{x}) \stackrel{\text{def}}{=} 1 - F_{p|X}\,(p \mid X = x)$
- In polling module: we replaced mean estimation with “MRP.” More generally, plug in a machine learning model
  - Now, can borrow information across prices and covariates
  - We must make a “parametric” assumption for how prices and covariates relate to purchasing decisions
- One example: Logistic regression
  - Target (Y variable) is purchase decision
  - Covariates are: price offered, user covariates, interactions between price and covariates or between covariates

Demand estimation comments

- Demand estimation and forecasting is probably the *most important and difficult* challenge in revenue management
- Unlike most machine learning challenges, we need to estimate a *function* $F(p)$ [or treat price as a covariate]
- We made a *substantial* assumption that almost never holds in practice: that you have historical data at many different prices p
    - Requires experimentation!

Today’s summary, & complicating factors

Today: We want to sell an item

- Only one item
- No capacity constraints
- No competition from other sellers
- No over-time dynamics
- Allowed to explicitly give different prices to different users

Then: revenue-maximizing price(s) and demand estimation

Next time: Relax (some of) these limiting assumptions

Questions?