IEOR E4650 Business Analytics

Session 18: Personalized MAB and Simulation

Spring 2018

Prof. Adam Elmachtoub

Contextual Multi-armed Bandit

- In the standard multi-armed bandit problem, decision and success probability depend only on the chosen action a_t.
- However, in many settings, decision and probability of success might also depend on a set of observable features about user t, which we call $x_t = (x_{t,1}, ..., x_{t,p})$.
  - Features include demographics, purchase history, click history, OS, etc.
- The Contextual Multi-armed Bandit problem is a generalization of the standard MAB problem where the success probabilities of each action a depend on x_t, i.e., we have $p_a(x_t)$ for each action a
- *There is no optimal action now, it depends on the user.*

Contextual Multi-armed Bandit: Formulation

- At each time step $t = 1, ..., T$
  1. Observe context $x_t = (x_{t,1}, ..., x_{t,p})$
  2. Choose action $a_t \in \{1, ..., K\}$
  3. Observe random, context-dependent reward $Y_t \sim p_{a_t}(x_t)$

Goal: Maximize Expected Total Successes

$$\mathsf{E}[\sum_{t=1}^{T} Y_t] = \sum_{t=1}^{T} p_{a_t}(x_t)$$

Equivalent Goal: Minimize Cumulative Regret

$$\sum_{t=1}^{T} \max_{a=1,...,K} p_a(x_t) - \sum_{t=1}^{T} p_{a_t} x(t)$$

Contextual Multi-armed Bandit: Challenges

- Need to estimate $p_a(x_t)$ for each action.
  - Idea: for each action a at time t, fit a machine learning model $\hat{p}_{t,a}(x_t)$ to estimate $p_a(x_t)$
    - Linear regression

$$\hat{p}_{t,a}(x_t) = \hat{\beta}_{a,0} + \sum_{j=1}^{p} \beta_{a,j} x_{t,j} = \hat{\beta}_a^T x_t$$

    - Logistic regression

$$\hat{p}_{t,a}(x_t) = \frac{1}{1 + e^{-\hat{\beta}_{a,0} - \hat{\beta}_a^T x_t}}$$

    - Other learners: K-NN, Decision Trees, LDA

- As before, need to navigate the exploration-exploitation trade-off
  - The uncertainty of each action's estimated reward now depends on x_t!

MAB Algorithms

1. Pure Exploration
   - Try a random action every period
2. Pure Exploitation
   - Choose the action according to $\text{argmax}_a \, \hat{p}_{t,a}(x_t)$
3. Explore then Exploit
   - Pure Explore for first T_0 periods, then Pure Exploit for $T - T_0$ periods
4. Epsilon-Greedy
   - Pure Explore with probability ϵ_t, pure exploit with probability $1 - \epsilon_t$
5. Upper Confidence Bound (UCB)
   - Choose the action according to $\text{argmax}_a \, \hat{p}_{t,a}(x_t) + \text{``}std.err.(\hat{p}_{t,a})\text{''}$

Contextual Epsilon-Greedy Policy

- Input parameter: $\epsilon \in [0, 1]$
- Contextual Epsilon-Greedy Algorithm: In each time step, choose a random action with probability ϵ (explore). Otherwise, choose the action which maximizes expected reward (exploit).
- Note: any machine-learning method can be used with this policy to model expected reward.

LinUCB Policy

- Input parameter: $\alpha \in [0, \infty)$
- Linear Upper Confidence Bounding (LinUCB): In each time step t, choose the action which maximizes the UCB score:

$$UCB_{t,a} = \hat{\beta}_{t,a}^T x_t + \alpha \text{``}std.err.(\hat{\beta}_{t,a}^T x_t)''$$

  - $\hat{\beta}_{t,a}$: estimated linear regression coefficient for action a
  - The second term represents our *uncertainty* about $\hat{\beta}_{t,a}^T x_t$.
  - Thus, α directly controls the trade-off between maximizing expected reward and maximizing information gain.
- A similar algorithm, GLM-UCB, uses logistic regression to model rewards.

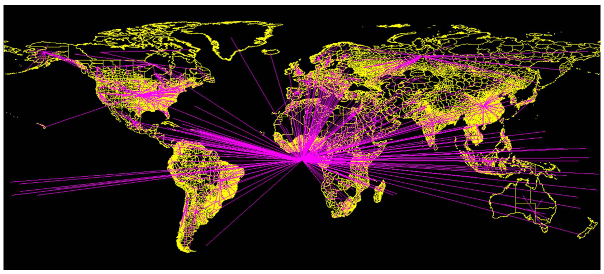
Which one is best?

- Surprisingly, all the algorithms perform decently well except for Pure Exploration
- Even Pure Exploitation is not bad since the randomness in the input (x_t) forces us to do some learning! In other words, we naturally try different actions anyways because the users are random.
- We shall explore more in the homework :)

Simulation Applications – I

Traffic simulation

- Traffic light timing
- One-way versus two-way streets
- Impact of road closures



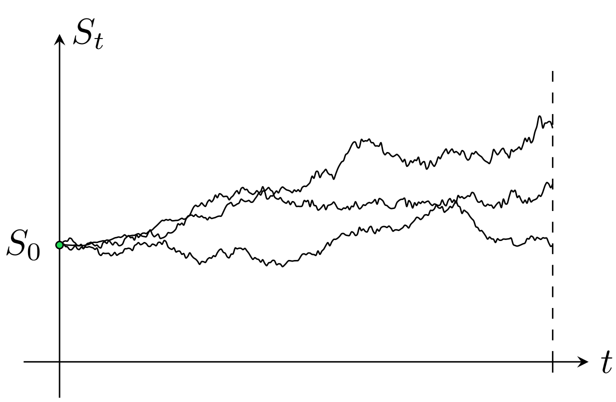
Simulation of epidemics

- Spread through air travel
- Analyze intervention efforts
- Analyze vaccination policies

Simulation Applications – II

Call Center simulation

- Random arrival times of calls
- Analyze impact of staffing plans on key perf. metrics



Financial simulation

- Pricing options and other securities
- Analyze hedging (risk management) strategies
- Capital allocation
- Value-at-risk and other simulation methods mandated by government regulation
- Note: can only simulate the known unknowns

Monte Carlo Simulation Process

Construct a model connecting inputs to outputs

- Output of interest and random inputs that impact the output
  - e.g., daily returns
- How the random inputs impact the outputs
- Nature of random inputs: distribution

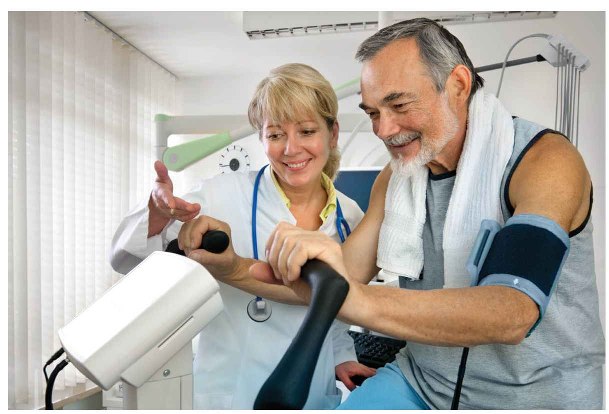
Run the simulation

- Generate many possible values that random inputs may take
  - e.g., positive or negative returns
- For each sequence of events, record the outputs
  - e.g., final wealth

Analyze the output

- Simulation shows how random inputs lead to a range of outcomes for the random outputs
- Distribution of the outputs: average, standard deviation, percentiles, ...

Valuing the Healthcare Pension Liability at GM



- UAW (United Auto Workers) and GM are in negotiations
- The transfer of the healthcare liability from GM to the union for a fixed amount is being discussed

How should this liability be valued?

Pro-Forma Analysis

Year	Age	Healthcare cost	HC borne by employer	Discount factor
2031	63	43.4		0.40
2032	64	47.1		0.38
2033	65	51.1	51.1	0.36
2034	66	55.5	55.5	0.34
2035	67	60.2	60.2	0.33
2036	68	65.3	65.3	0.31
2037	69	70.8	70.8	0.30
2038	70	76.9	76.9	0.28
2039	71	83.4	83.4	0.27
2040	72	90.5	90.5	0.26
2041	73	98.2	98.2	0.24
2042	74	106.5	106.5	0.23
2043	75	115.6	115.6	0.22
2044	76	125.4	125.4	0.21
2045	77	136.1	136.1	0.20
2046	78	147.6		0.19
2047	79	160.2		0.18
2048	80	173.8		0.17
2049	81	188.6		0.16
2050	82	204.6		0.16

Average employee

- Male
- Age: 45 years
- Age at retirement: 65 years
- Age at death: 78 years

Healthcare costs

- Current year: $10,000
- Annual increase in healthcare costs: 8.5%
- Discount rate assumption: 5.0%

How would you value the healthcare pension liability today?

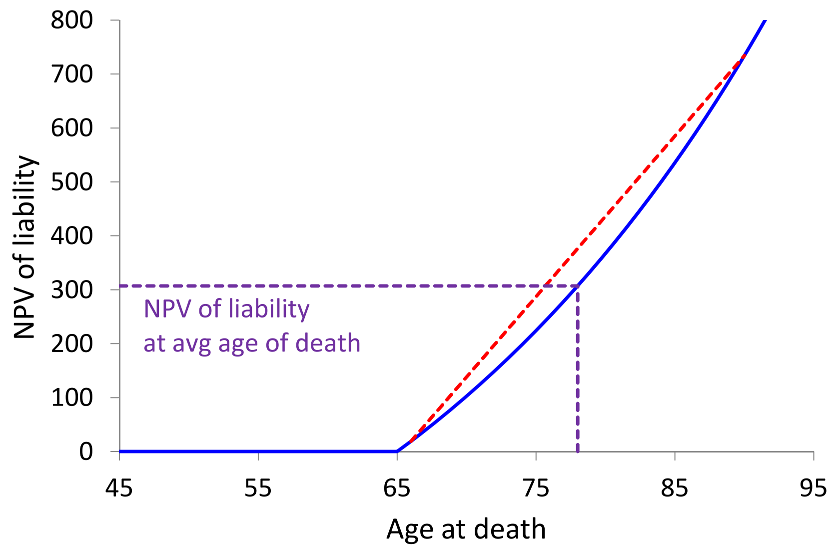
NPV of liability: $307,000

The Impact of Randomness in Age of Death

Age at death	NPV of liability (in $000)
78	307
66	19
90	734

- NPV of liability at average age of death (78): $307,000
- Suppose an employee dies at 66 with probability 1/2 or at 90 with probability 1/2. What is the expected NPV?
- Average NPV of liability with death at 66 or 90: $377,000 ($377 = (19+734)/2$)
- The difference is due to the nonlinear relationship between NPV and age at death

Intuition: Plot of NPV of Liability versus Age at Death



- In this particular case, we could have predicted that working with average death age would lead to underestimating the NPV!
- Effect called Jensen's inequality

Actuarial Life Table: Male

Age	Death probability	Life expectancy	Age	Death probability	Life expectancy	Age	Death probability	Life expectancy	Age	Death probability	Life expectancy	Age	Death probability	Life expectancy
0	0.70%	75.9	20	0.11%	56.8	40	0.22%	38.2	60	1.10%	21.3	80	6.16%	8.1
1	0.04%	75.4	21	0.13%	55.9	41	0.24%	37.3	61	1.18%	20.5	81	6.82%	7.6
2	0.03%	74.5	22	0.13%	55.0	42	0.26%	36.4	62	1.27%	19.7	82	7.53%	7.1
3	0.02%	73.5	23	0.14%	54.0	43	0.29%	35.5	63	1.37%	19.0	83	8.32%	6.7
4	0.02%	72.5	24	0.14%	53.1	44	0.31%	34.6	64	1.49%	18.2	84	9.19%	6.2
5	0.02%	71.5	25	0.14%	52.2	45	0.34%	33.7	65	1.62%	17.5	85	10.16%	5.8
6	0.02%	70.5	26	0.14%	51.3	46	0.37%	32.8	66	1.76%	16.8	86	11.24%	5.4
7	0.01%	69.5	27	0.14%	50.3	47	0.41%	31.9	67	1.91%	16.1	87	12.45%	5.0
8	0.01%	68.6	28	0.14%	49.4	48	0.44%	31.1	68	2.08%	15.4	88	13.78%	4.7
9	0.01%	67.6	29	0.14%	48.5	49	0.49%	30.2	69	2.25%	14.7	89	15.25%	4.3
10	0.01%	66.6	30	0.14%	47.5	50	0.53%	29.4	70	2.45%	14.0	90	16.84%	4.0
11	0.01%	65.6	31	0.14%	46.6	51	0.58%	28.5	71	2.67%	13.4	91	18.55%	3.7
12	0.01%	64.6	32	0.15%	45.7	52	0.63%	27.7	72	2.92%	12.7	92	20.38%	3.5
13	0.02%	63.6	33	0.15%	44.7	53	0.68%	26.8	73	3.19%	12.1	93	22.33%	3.2
14	0.03%	62.6	34	0.16%	43.8	54	0.73%	26.0	74	3.48%	11.5	94	24.39%	3.0
15	0.05%	61.6	35	0.16%	42.9	55	0.79%	25.2	75	3.82%	10.9	95	26.43%	2.8
16	0.06%	60.6	36	0.17%	41.9	56	0.85%	24.4	76	4.21%	10.3	96	28.42%	2.6
17	0.07%	59.7	37	0.18%	41.0	57	0.91%	23.6	77	4.63%	9.7	97	30.32%	2.5
18	0.08%	58.7	38	0.19%	40.1	58	0.97%	22.8	78	5.08%	9.2	98	32.09%	2.4
19	0.10%	57.8	39	0.21%	39.2	59	1.04%	22.0	79	5.59%	8.6	99	33.69%	2.2

Reports the remaining life expectancy

Source: Social Security Administration website:
http://www.ssa.gov/oact/STATS/table4c6.html

Actuarial Life Table: Female

Age	Death probability	Life expectancy	Age	Death probability	Life expectancy	Age	Death probability	Life expectancy	Age	Death probability	Life expectancy	Age	Death probability	Life expectancy
0	0.57%	80.8	20	0.04%	61.5	40	0.13%	42.2	60	0.67%	24.3	80	4.39%	9.7
1	0.04%	80.3	21	0.04%	60.6	41	0.15%	41.3	61	0.73%	23.5	81	4.88%	9.1
2	0.02%	79.3	22	0.05%	59.6	42	0.16%	40.4	62	0.80%	22.6	82	5.44%	8.5
3	0.02%	78.3	23	0.05%	58.6	43	0.18%	39.4	63	0.87%	21.8	83	6.07%	8.0
4	0.02%	77.3	24	0.05%	57.6	44	0.20%	38.5	64	0.94%	21.0	84	6.78%	7.5
5	0.01%	76.4	25	0.05%	56.7	45	0.22%	37.6	65	1.03%	20.2	85	7.57%	7.0
6	0.01%	75.4	26	0.06%	55.7	46	0.24%	36.6	66	1.13%	19.4	86	8.47%	6.5
7	0.01%	74.4	27	0.06%	54.7	47	0.26%	35.7	67	1.24%	18.6	87	9.46%	6.0
8	0.01%	73.4	28	0.06%	53.8	48	0.28%	34.8	68	1.36%	17.8	88	10.57%	5.6
9	0.01%	72.4	29	0.06%	52.8	49	0.30%	33.9	69	1.49%	17.1	89	11.79%	5.2
10	0.01%	71.4	30	0.07%	51.8	50	0.33%	33.0	70	1.64%	16.3	90	13.11%	4.9
11	0.01%	70.4	31	0.07%	50.9	51	0.36%	32.1	71	1.82%	15.6	91	14.56%	4.5
12	0.01%	69.4	32	0.07%	49.9	52	0.38%	31.2	72	2.00%	14.9	92	16.12%	4.2
13	0.01%	68.4	33	0.08%	48.9	53	0.41%	30.4	73	2.20%	14.2	93	17.79%	3.9
14	0.02%	67.4	34	0.08%	48.0	54	0.43%	29.5	74	2.42%	13.5	94	19.58%	3.6
15	0.02%	66.4	35	0.09%	47.0	55	0.46%	28.6	75	2.67%	12.8	95	21.38%	3.4
16	0.03%	65.5	36	0.09%	46.1	56	0.49%	27.7	76	2.96%	12.1	96	23.19%	3.2
17	0.03%	64.5	37	0.10%	45.1	57	0.52%	26.9	77	3.27%	11.5	97	24.95%	3.0
18	0.03%	63.5	38	0.11%	44.1	58	0.56%	26.0	78	3.60%	10.9	98	26.65%	2.8
19	0.04%	62.5	39	0.12%	43.2	59	0.61%	25.2	79	3.97%	10.2	99	28.25%	2.7

Reports the remaining life expectancy

Source: Social Security Administration website:
http://www.ssa.gov/oact/STATS/table4c6.html

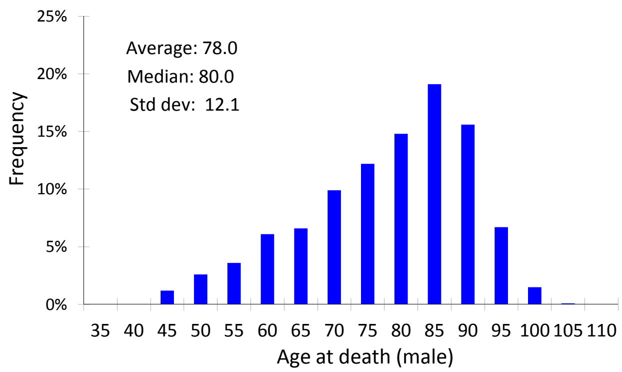
Simulating the Age of Death

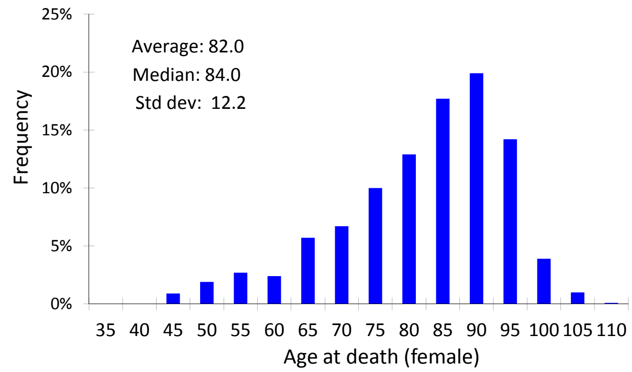
Year	Age	Rand	Prob death at this age	Death	Year	Age	Rand	Prob death at this age	Death	Year	Age	Rand	Prob death at this age	Death
2013	45	0.915	0.3%	0	2033	65	0.554	1.6%	0	2053	85	0.709	10.2%	0
2014	46	0.945	0.4%	0	2034	66	0.003	1.8%	1	2054	86	0.669	11.2%	0
2015	47	0.978	0.4%	0	2035	67	0.380	1.9%	0	2055	87	0.492	12.5%	0
2016	48	0.872	0.4%	0	2036	68	0.467	2.1%	0	2056	88	0.083	13.8%	1
2017	49	0.530	0.5%	0	2037	69	0.311	2.2%	0	2057	89	0.830	15.2%	0
2018	50	0.667	0.5%	0	2038	70	0.792	2.4%	0	2058	90	0.347	16.8%	0
2019	51	0.766	0.6%	0	2039	71	0.467	2.7%	0	2059	91	0.038	18.5%	1
2020	52	0.230	0.6%	0	2040	72	0.513	2.9%	0	2060	92	0.024	20.4%	1
2021	53	0.255	0.7%	0	2041	73	0.966	3.2%	0	2061	93	0.788	22.3%	0
2022	54	0.582	0.7%	0	2042	74	0.814	3.5%	0	2062	94	0.058	24.4%	1
2023	55	0.279	0.8%	0	2043	75	0.150	3.8%	0	2063	95	0.834	26.4%	0
2024	56	0.614	0.9%	0	2044	76	0.988	4.2%	0	2064	96	0.790	28.4%	0
2025	57	0.272	0.9%	0	2045	77	0.237	4.6%	0	2065	97	0.180	30.3%	1
2026	58	0.937	1.0%	0	2046	78	0.722	5.1%	0	2066	98	0.659	32.1%	0
2027	59	0.238	1.0%	0	2047	79	0.302	5.6%	0	2067	99	0.919	33.7%	0
2028	60	0.356	1.1%	0	2048	80	0.065	6.2%	0	2068	100	0.100	35.4%	1
2029	61	0.525	1.2%	0	2049	81	0.497	6.8%	0	2069	101	0.970	37.1%	0
2030	62	0.047	1.3%	0	2050	82	0.889	7.5%	0	2070	102	0.014	39.0%	1
2031	63	0.564	1.4%	0	2051	83	0.200	8.3%	0	2071	103	0.946	41.0%	0
2032	64	0.662	1.5%	0	2052	84	0.818	9.2%	0	2072	104	0.387	43.0%	1

Steps for age of death simulation

- Age k: compare the random number (RAND_k) with the probability of dying (p_k)
- Death indicator: IF($\text{RAND}_k < p_k$, 1, 0)
- Age at death: minimum of the ages (k) with death indicator of 1
- In this example, the age of death is 66

Frequency Histograms: Age of Death

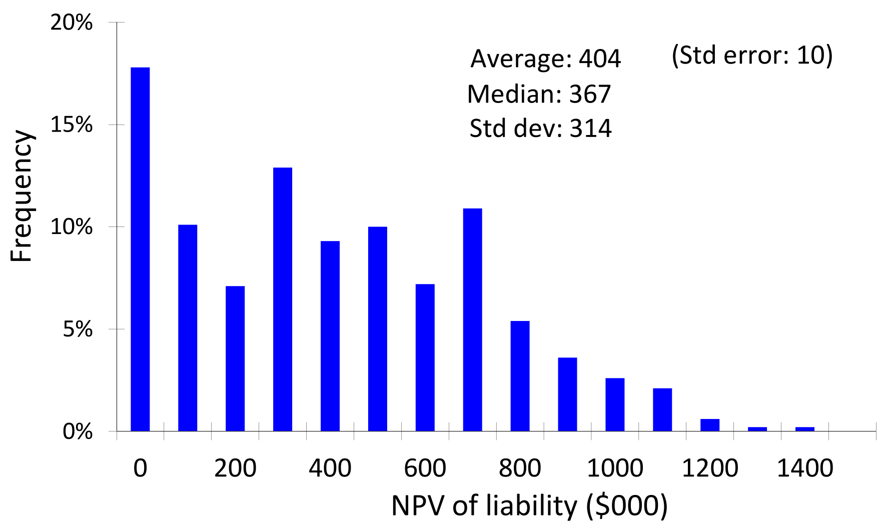




Age of death histograms

- Results are for males (left) and females (right) who are 45 years old today
- Results based on 1000 simulation trials

Frequency Histogram: NPV of Liability



NPV of liability histogram (1000 simulation trials)

- Average NPV of liability: $404,000
- NPV of liability assuming death age equals 78 (using pro-forma analysis): $307,000

Simulation estimate is 30% higher
Reasoning with averages can be misleading

Monte Carlo Simulation: Accuracy

How many paths should one simulate?

- The number of paths one generates will determine the accuracy of the statistics one obtains for the outcome
- The average value of an output variable is approximately normally distributed with a large number (n) of simulation trials
  - Why? Recall the Central Limit Theorem from statistics
- The standard error of the average value is:

$$\sigma/\sqrt{n}$$

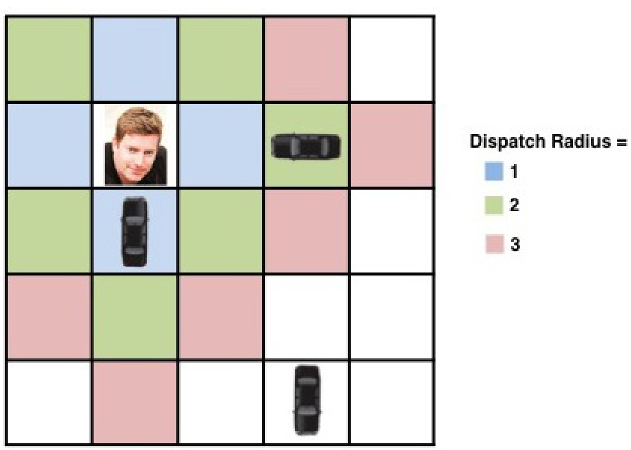
  where σ is the standard deviation of the output variable
- Implication: to cut the standard error in half requires four times as many simulation trials

Additional Simulation Example

Uber: Online Marketplace matching drivers and users

Design question: How to route idle drivers?

Problem can be approached through simulation



http://blog.uber.com/aisimulation

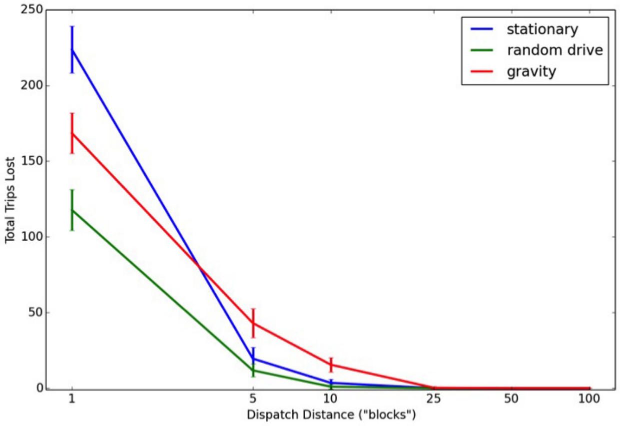
Objectives and metrics

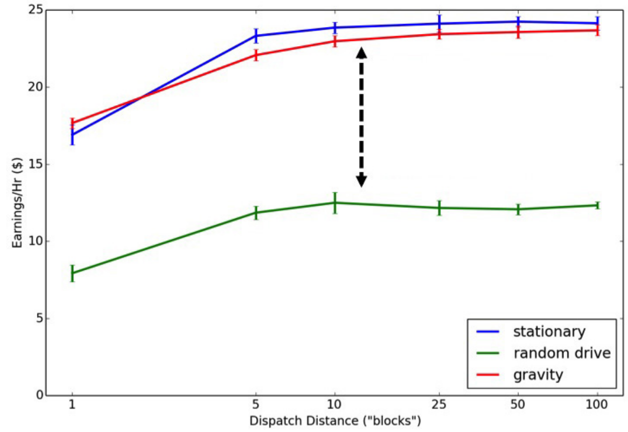
1. On the rider side, getting you a ride when you need it.
   - Total potential trips lost
2. On the driver side, maximize trips taken on the system, which maximizes driver partners' earnings.
   - Average # of trips completed
   - Average distance driven (on and off trips) and gas costs
   - Average driver earnings

Uber Routing Simulation

Three routing policies for idle drivers:

- random drive — drive around while waiting for next pick-up
- stationary — stay put after a drop-off
- gravity — gravitate toward high demand areas





"When dispatch distances are very short drivers should navigate back toward demand density. However when dispatch distances are relatively longer, drivers maximize their earnings by using less gas by remaining stationary between trips."

Wrap Up

- In many business scenarios, reasoning with averages can be misleading
- Simulation is a tool to support decisions
  - Provides range of possible outcomes: applies when it is difficult to assess the impact of decisions due to randomness
  - Sensitivity analysis: simulation can be used to test the sensitivity of an outcome to different assumptions
  - Widely applicable: simulation has been applied in insurance, financial services, healthcare, and many other areas