DS102 - Homework 5

If you are submitting a handwritten version please make sure your answers are legible as you may lose points otherwise.
Data science is a collaborative activity. While you may talk with others about the homeworks, we ask that you write your solutions individually. If you do discuss the homework with others please include their name in your submission.
Due by: 1:59pm, Tuesday 11th November, 2019

1. (15 points) **Simpson's Paradox**

   In this question we discuss an (in)famous paradox in causal inference called Simpson's paradox.

   This paradox refers to the phenomenon where an event D increases the probability of an event R in a given population P, and, at the same time, decreases the probability of R in every subpopulation of P.

   For a concrete example, suppose D is the event of having taken some drug, and R is the event of recovery from a disease. Suppose we divide all people that have taken the drug into two subpopulations, A - adults (18 years old or above) and K - kids (under 18 years old). Then, Simpson's paradox says it's possible to observe the following:

$$\mathbb{P}(R|D) > \mathbb{P}(R|D^c), \quad \mathbb{P}(R|D,A) < \mathbb{P}(R|D^c,A), \quad \mathbb{P}(R|D,K) < \mathbb{P}(R|D^c,K), \tag{1}$$

   where by $(\cdot)^c$ we denote the complement of an event $(\cdot)$, implying $\mathbb{P}(F) + \mathbb{P}(F^c) = 1$, for any event F. Note also that $K = A^c$ and $A = K^c$. In words, the above example says it's possible for a drug to be harmful for every subpopulation, but beneficial for the population given by the union of these subpopulations!

   The goal of this question is to understand when such a phenomenon can happen.

   (a) (5 points) Suppose that we treat 80 individuals, and get the following table of outcomes.

Combined (adults + kids)	R	R^c	$R + R^c$	Recovery rate
Drug (D)	20	20	40	0.5
No drug (D^c)	16	24	40	0.4
	36	44	80	

Table 1: Table of outcomes.

   We observe a higher recovery rate for people who have taken the drug.

   Suppose we estimate all probabilities in equation (1) via empirical quantities, for example:

$$\mathbb{P}(R|D) = \frac{\#\text{recoveries among drug-takers}}{\#\text{drug-takers}},$$

or

$$\mathbb{P}(R|D, A) = \frac{\#\text{recoveries among adult drug-takers}}{\#\text{adult drug-takers}},$$

and so on. Fill out the following two tables, *in a way that is consistent with Table 1*, such that you observe Simpson's paradox. Consistency implies, for example, that you should have 80 people in total, 36 recoveries in total, etc. You can distribute them across the two subpopulations, adults and kids, as you choose. Note that there are many possible solutions to this problem.

Can you also explain in your own words (without any math) why Simpson's paradox might happen? A short answer is fine, just give an intuitive explanation of how you went about designing the two tables.

Adults	R	R^c	$R + R^c$	Recovery rate
Drug (D)	?	?	?	?
No drug (D^c)	?	?	?	?
	?	?	?	

Kids	R	R^c	$R + R^c$	Recovery rate
Drug (D)	?	?	?	?
No drug (D^c)	?	?	?	?
	?	?	?	

(b) (3 points) Prove that, if D and A are independent events, then D and K are independent, D^c and A are independent, and D^c and K are independent.

(c) (7 points) Prove that, if D and A are independent events, then Simpson's paradox cannot happen. More precisely, if D and A are independent, knowing $\mathbb{P}(R|D, A) < \mathbb{P}(R|D^c, A)$ and $\mathbb{P}(R|D, K) < \mathbb{P}(R|D^c, K)$ implies $\mathbb{P}(R|D) < \mathbb{P}(R|D^c)$.
Hint: You might want to use the law of total probability.

2. (20 points) **Experiment Design for Linear Models**

In lecture we discussed three different design methodologies for collecting data to estimate the parameters of a linear trend. In this problem we'll investigate these methodologies more closely through analysis and experiments.

Specifically, we'll first assume that we have a true underlying linear model that governs our data, for all $x_i \in [0, 1]$:

$$y_i = ax_i + b + \epsilon_i$$

Where $\epsilon_i \sim \mathcal{N}(0, \sigma^2)$ is the observation noise, and a, b are the true parameters of the model. We are interested in designing n query points x_i, such that the (x_i, y_i) pairs we receive will result in the best estimates of a and b. We'll use least squares to do this.

We'll compare three different designs for x:

- Evenly spaced: query n points evenly spaced within $[0, 1]$ (inclusive).
- Dumbbell: query half of the allotted points at $x_i = 0$, the other half at $x_i = 1$.
- Quadratic: query one third of the allotted points at $x_i = 0$, the one third at $x_i = 1$, and one third at the halfway point: $x_i = 0.5$.

Once we've designed $\{x_i\}_{i=1}^n$ and conducted our experiment, we have pairs $\{(x_i, y_i)\}_{i=1}^n$. We will use OLS to estimate a. Recall that for the univariate regression above, the OLS estimators of a is

$$\hat{a}_{OLS} = \frac{\sum_{i=1}^n (x_i - \overline{x})(y_i - \overline{y})}{\sum_{i=1}^n (x_i - \overline{x})^2}$$

and has variance

$$Var(\hat{a}_{OLS}) = \frac{\sigma^2}{\sum_{i=1}^n (x_i - \overline{x})^2}$$

Where $\overline{x} := \frac{1}{n} \sum_{i=1}^n x_i$.

(a) (8 points) Prove the following properties of the variance of the estimators.

(i) Show that the variance of the dumbbell estimator for this problem is

$$Var(\hat{a}_{DB}) = \frac{4\sigma^2}{n} \ .$$

.

(ii) Show that the variance of the estimator resulting from a quadratic design on this problem setting is

$$Var(\hat{a}_{quad}) = \frac{3}{2} Var(\hat{a}_{DB}) \ .$$

(iii) Show that the variance of the estimator resulting from a evenly spaced design on this problem setting is

$$Var(\hat{a}_{even}) = \frac{3(n-1)}{n+1} Var(\hat{a}_{db}) \ .$$

(b) (6 points) Implement the following experiment to assess these three design methodologies;

1. For $n = 100$, instantiate three design vectors $x = x_1, \ldots, x_n$ corresponding to the designs above.
2. For $x_1, \ldots, x_n$ in each of the three samples, sample from the "true" model:

$$y_i = a \cdot x_i + \epsilon_i$$

where $a = 5$ and $\epsilon_i = \mathcal{N}(0, 0.5)$.

3. Use these three sets of 100 (x, y) pairs to compute three OLS estimates of α.
4. Make three scatter plots of y vs. x, one for each experiment. On these plots overlay the fitted trendline using the OLS estimate of α (*note: there's no intercept in our model*).

(c) (2 points) Repeat the experiment above 1000 times (keep n fixed at 100), and record the distribution of OLS estimates of α for each of the three experiment designs. Calculate the empirical average and standard deviations of the estimators for α over these 1000 trials.

(d) (2 points) Now follow the same procedure as in part (b) (you need only run each experiment once), but for a model which is not what you expected, but rather is

$$y_i = b(x_i - 0.5)^2 + \epsilon_i$$

where $b = 8$ and $\epsilon_i = \mathcal{N}(0, 0.5)$. Run the same procedure to get the three OLS estimates of a, and plot the scatter plots with trend line overlaid.

(e) (2 points) Using the results above, compare and contrast the three sampling strategies. In particular, address which strategy you would want to use if (a) you were absolutely certain the relationship between x and y is linear for $x \in [0, 1]$, (b) you were fairly certain the relationship is linear for x in this range, but it's possible that it is also quadratic, and (c) you don't really know the true relationship between x and y, but as a first exploratory step you'd model it as linear.

Submit code and plots for the coding parts of this question as either a pdf printed python notebook (as you do with labs), or by submitting a python script (as pdf printout) with plots separately.

3. (20 points) **Regret of the Explore-then-Commit Algorithm** In this problem, we will analyze the regret of the Explore-then-Commit algorithm in multi-armed-bandit problems.

Suppose that you have a stochastic multi-armed bandit environment where there are K arms, each of which has a reward distribution P_i for $i = 1, ..., K$, with mean $\mu_i = \mathbb{E}_{P_i}[X_i]$. In this problem, we will assume that the rewards are bounded in $[a, b]$ (i.e $P(a \leq X_i \leq b) = 1$ for all arms $i = 1, ..., K$, where $a < b$).

An Explore-then-Commit algorithm pulls each of the K arms c times. After cK rounds, it commits to the arm with the highest sample mean $\hat{\mu}_i$:

$$\hat{\mu}_i = \frac{1}{c}\sum_{s=1}^{c} X_s,$$

and always chooses that one for all time afterwards. The algorithm is described below in Algorithm 1. By A_t we denote the choice of arm to pull.

Algorithm 1 Explore-then-Commit Algorithm

input: Number of exploratory pulls: c

For $t = 1, 2, ...$:

$$A_t = \begin{cases} (t \bmod k) + 1 & : \quad t \leq cK \\ \arg\max_{i \in 1,...,K} \hat{\mu}_i & : \quad t > cK \end{cases}$$

Define the mean of the optimal arm as:

$$\mu^* = \max_{i \in \{1,...,K\}} \mu_i,$$

and the sub-optimality gap of a sub-optimal arm i as:

$$\Delta_i = \mu^* - \mu_i,$$

We would like to analyze the pseudo-regret of this algorithm for various choices of c. Recall that the pseudo-regret of an algorithm is given by:

$$R(n) = n\mu^* - \mathbb{E}\left[\sum_{t=1}^{n} X_{A_t}\right]$$

(a) (5 points) Let us define the random variable $T_a(t)$ as the number of times arm $a = 1, ..., K$ has been pulled up to and including time t:

$$T_a(t) = \sum_{s=1}^{t} \mathbb{I}\{A_s = a\}$$

Show that we can decompose the regret as:

$$R(n) = \sum_{i=1}^{K} \Delta_i \mathbb{E}[T_i(n)]$$

(Hint: recall that the choice of action, and the reward are independent)

(b) (3 points) Show that for an Explore-then-Commit algorithm, if $n > cK$:

$$\mathbb{E}[T_i(n)] = c + (n - Kc)\mathbb{P}\left(\hat{\mu}_i > \max_{j=1,...,K; j \neq i} \hat{\mu}_j\right)$$

(c) (2 points) Suppose, without loss of generality, that the optimal arm (the arm with the highest mean) is 1. Show that for any sub-optimal arm i:

$$\mathbb{P}\left(\hat{\mu}_i > \max_{j=1,...,K; j \neq i} \mu_j\right) \leq \mathbb{P}\left(\hat{\mu}_i > \hat{\mu}_1\right)$$

(d) (5 points) In the last few parts, we have shown that:

$$\mathbb{E}[T_i(n)] \leq c + (n - Kc)\mathbb{P}\left(\hat{\mu}_i > \hat{\mu}_1\right)$$

Using the Hoeffding Bound, show that

$$\mathbb{P}\left(\hat{\mu}_i > \hat{\mu}_1\right) \leq exp\left(-\frac{c\Delta_i^2}{(b-a)^2}\right)$$

Hint: Recall that the Hoeffding bound applies to random variables $Y_1, ..., Y_d$ where each random variable Y_j is bounded between a_j and b_j for $j = 1, ..., d$. It is given by:

$$\mathbb{P}\left(\sum_{j=1}^{d} Y_j - \mathbb{E}\left[\sum_{j=1}^{d} Y_j\right] > t\right) \leq exp\left(-\frac{2t^2}{\sum_{j=1}^{d}(b_j - a_j)^2}\right)$$

(e) (5 points) Putting it all together, we have that:

$$\mathbb{E}[T_i(n)] \leq c + (n - Kc)exp\left(-\frac{c\Delta_i^2}{(b-a)^2}\right)$$

Suppose that you knew the minimum sub-optimality gap $\Delta = \min_{i>1} \Delta_i$. Then, for each sub-optimal arm $i = 2, ..., K$:

$$\mathbb{E}[T_i(n)] \leq c + n\ exp\left(-\frac{c\Delta^2}{(b-a)^2}\right),$$

where we trivially upper bound $n - Kc$ by n.

Find the value of c which guarantees that:

$$exp\left(-\frac{c\Delta^2}{(b-a)^2}\right) \leq \frac{1}{n}$$

What is the resulting upper bound on the pseudo-regret with this value of c? Is the pseudo-regret sublinear?