

חישוביות וקוגניציה – תרגיל 8

להגשה עד: 09/01/2020

שאלה 1

בשאלה זו תשמשו במודל של למידת חיזוק שנלמד בכיתה ע"מ לתאר התנהגות של נבדקים במשימה פשוטה של בחירה בין 2 אפשרויות.

תכנון הניסוי

בכל Trial הנבדק\ת בוחר\ת אחת מבין 2 אפשרויות $y_t \in \{0, 1\}$, ומקבל\ת גמול שנדגם מההתפלגות $p(r_t|y_t)$. עליכם לבחור מראש את הסטטיסטיקה של הגמול, כאשר מבנה הניסוי הוא גמול בינארי (1 או 0 "נקודות") עם סיכוי הצלחה שונה עבור כל בחירה. פורמלית, $r_t|y \sim \text{Ber}(q_y)$, ועליכם לבחור את q_0 ו q_1 (סיכויי ה"הצלחה" בכל פעולה). ציינו במפורש בקובץ שאתם מגישים מהם הפרמטרים שבחרתם. פה ובהמשך התרגיל $\text{Ber}(a)$ מסמן התפלגות ברנולי עם פרמטר a , כלומר מ"מ שמקבל 1 בסיכוי a ו 0 בסיכוי $1 - a$.

ההוראות שיש לתת לנבדקים הן שבכל צעד עליהם לבחור בין שתי אפשרויות, והמטרה היא לקבל כמה שיותר נקודות – אך אין להסביר להם על המבנה הסטטיסטי של הניסוי.

חלק א' – איסוף נתונים

כתבו פונקציה שמריצה ניסוי ושומרת את התוצאות.

בכל צעד, על הנבדק לבחור פעולה דרך המקלדת ע"י הקשה של 0 או 1. השתמשו בפונקציה `input(str)` שמציגה טקסט ומחכה לקלט מהמשתמש (וודאו שהקלט שהתקבל הוא אחד מהספרות החוקיות). לאחר הבחירה, התוכנה מגרילה את הגמול מההתפלגות המתאימה ומציגה אותו לנבדק למשך שניה – השתמשו בפונקציה `time.sleep(s)` (בפייתון) או ב `pause(s)` (במטלב) שעוצרת את התוכנה ל s שניות. לאחר מכן, יש "לנקות" את המסך והנבדק עובר ל trial הבא. ע"מ "לנקות" את המסך השתמשו בפקודה `clc` במטלב. בפייתון, הדרך הקלה ביותר היא `os.system('cls')` לווינדוס ו `os.system('clear')` בלינוקס.

הריצו את הניסוי למשך $T = 100$ trials פעמיים, על שני נבדקים שונים (פעם אחת לכל נבדק) – חברים (שאינם תלמידים בקורס) או בני משפחה.

חלק ב' – הצגת הנתונים

הציגו גרף שמתאר את הלמידה של כל אחד משני הנבדקים: חשבו את סיכויי הבחירה של הנבדק בפעולה העדיפה במהלך 20 ה trials הראשונים ו 20 ה trials האחרונים של הניסוי.

חלק ג' – ניתוח והתאמת מודל

נרצה להתאים פרמטרים שמסבירים את ההתנהגות של הנבדקים לפי מודל חישובי פשוט. נבדוק שני מודלים אפשריים – אחד שמניח שהנבדקים מתנהגים לפי אלגוריתם REINFORCE והשני שמניח שהנבדקים מתנהגים לפי TD-Learning (מודל פשוט של Q -Learning). הפרמטר שנרצה להתאים כדי להסביר את התנהגות הנבדק הוא קצב הלימוד, בכל אחד מהמודלים. המשוואות שמגדירות את המודל ההסתברותי מפורטות להלן:

REINFORCE מודל

$$\begin{aligned}y_t &\sim \text{Ber}(p(t)) \\ r_t &\sim \text{Ber}(q_{y_t}) \\ p(t+1) &= p(t) + \eta \cdot r_t \cdot p(t) (1 - p(t)) \cdot (y_t - p(t)) \\ p(1) &= \frac{1}{2}\end{aligned}$$

TD-Learning מודל

$$\begin{aligned}p(t) &= \frac{e^{\beta V_1(t)}}{e^{\beta V_1(t)} + e^{\beta V_0(t)}} \\ y_t &\sim \text{Ber}(p(t)) \\ r_t &\sim \text{Ber}(q_{y_t}) \\ V_i(t+1) &= V_i(t) + \eta \cdot \delta(i, y_t) \cdot (r_t - V_i(t)) \\ V_0(1) &= V_1(1) = 0 \\ \beta &= 2\end{aligned}$$

כאשר $\delta(i, y_t) = \begin{cases} 1 & y_t = i \\ 0 & y_t \neq i \end{cases}$ (כלומר, בכל trial מעדכנים רק את V המשווייך לפעולה שנבחרה באותו Trial)
עבור כל אחד מהמודלים:

1. כתבו פונקציה שמקבלת רצף של בחירות, רצף של גמולים ופרמטר קצב לימוד, ומחזירה את לוג הנראות (log-likelihood) של הנתונים (כלומר, הבחירות של הנבדק) בהינתן קצב הלימוד.
2. כתבו פונקציה שמקבלת רצף של בחירות ושל גמולים, ומחזירה את קצב הלימוד שממקסם את לוג הנראות (log-likelihood). לצורך כך, הריצו את הפונקציה שכתבתם בסעיף 1 עבור כל ערכי η בתחום בין 0 ל 1 בקפיצות של 0.005, ומצאו את הערך הטוב ביותר. האם לשני הנבדקים מתאים קצב לימוד שונה?
3. נרצה לבדוק עד כמה ניתן לסמוך על ההבדלים בקצב הלימוד שמצאנו לשני הנבדקים השונים. לצורך כך, נרצה להעריך את השונות באמידת הפרמטרים.
 - (א) כתבו פונקציה שמקבלת קצב לימוד, ומבצעת סימולציה של למידה לפי המודל הרלבנטי. הפונקציה תחזיר תוצאה של "ניסוי", כלומר את רצף הבחירות של המודל ורצף הגמולים שנצפו.
 - (ב) הריצו את הפונקציה מהסעיף הקודם 100 פעמים, ולכל "ניסוי" בצעו את הפרוצדורה מסעיף 2, כלומר – שערכו מחדש את קצב הלימוד בהנתן תוצאות הסימולציה. בסופו של דבר, תקבלו 100 ערכים של קצב לימוד משוער.
 - (ג) ציירו על אותו גרף 2 היסטוגרמות של קצבי הלימוד המשוערים, אחת עבור כל נבדק, בצבעים שונים. לאור התוצאה שקיבלתם, מה ניתן לומר על רמת הבטחון בנוגע להבדל בין הנבדקים?
4. דווחו את התוצאות של הניסוי בלינק הבא: <https://forms.gle/7o8ujuDfAgoBJTud9>

חלק ד'

איזה מודל לדעתכם מתאר טוב יותר את התנהגות הנבדקים? הסבירו.