

תרגיל 6 - יותם גרדוש - 208541334

שאלות הקודמה

1. לפי ההגדרות שראינו בכיתה, מה נכון לומר על הקשר בין $V_\pi(s)$ לבין $Q_\pi(s, a)$?

(א) כדי להגיע מ- s ל- $V_\pi(s, a)$, יש לפחות פניו כל הפעולות האפשריות, a , בהתאם למדייניות $\pi(a|s)$

(ב) כדי להגיע מ- s ל- $V_\pi(s, a)$, יש לפחות פניו כל המ מצבים הבאים האפשריים $'s$, בהתאם למעברי העולם $P(s'|s, a)$

(ג) כדי להגיע מ- s ל- $Q_\pi(s, a)$, יש לפחות פניו כל הפעולות האפשריות, a , בהתאם למדייניות $\pi(a|s)$

(ד) כדי להגיע מ- s ל- $Q_\pi(s, a)$, יש לפחות פניו כל המ מצבים הבאים האפשריים $'s$, בהתאם למעברי העולם $P(s'|s, a)$

השאלה היא שאלת סטטיסטיקה אוניברסלית. נסמן a בפ' $Q_\pi(s, a)$: $|Q_\pi(s, a)|$ נסמן בפ' (א).

השאלה היא שאלת סטטיסטיקה אוניברסלית. נסמן $V_\pi(s)$ בפ' (ב). $V_\pi(s)$ נסמן בפ' (ב). $Q_\pi(s, a)$ נסמן בפ' (ב).

השאלה היא שאלת סטטיסטיקה אוניברסלית. נסמן $V_\pi(s)$ בפ' (ב). $V_\pi(s)$ נסמן בפ' (ב). $Q_\pi(s, a)$ נסמן בפ' (ב).

השאלה היא שאלת סטטיסטיקה אוניברסלית. נסמן $Q_\pi(s, a)$ בפ' (ב). $Q_\pi(s, a)$ נסמן בפ' (ב).

2. כאשר אנו מדברים על "מודל" ומפרידים בין Model-Free ו-Model-Based, מה נכון בתחום הגדרת המודל? בחרו בכל התשובות הנכונות

(א) מודל - נסמן: אין לנו גישה דיאלקטית כלשהי זיהוי גזורה.

(ב) מודל - נסמן: גזורה אטומית הסמן גזורה גזורה.

(ג) מודל - נסמן: גזורה אטומית גזורה.

(ד) מודל - נסמן: גזורה אטומית גזורה.

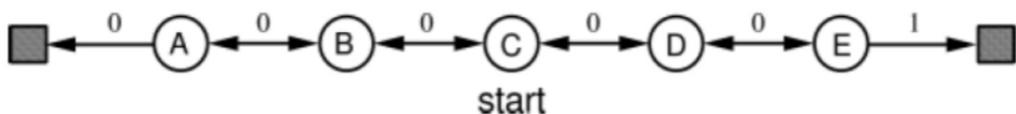
(ה) מודל - נסמן: גזורה אטומית גזורה גזורה גזורה.

(ו) מודל - נסמן: גזורה אטומית גזורה גזורה גזורה.

(ז) מודל - נסמן: גזורה אטומית גזורה גזורה גזורה.

1. למידת חיזוק - ייצור עולם חדש מימידי

יצור חי בעולם חד-ממדדי המתואר באירוע:



היצור מתחילה בעיגול C ובכל צעד, נע שמאלה או ימינה בהסתברות $0.5 = p$. כאשר הוא מגיע לאחד מהריבועים הקיימים נגמר ה-*trial*, הוא מוחזר אוטומטית למצב C ומתחילה trial חדש. הגמול היחיד שיש במשחק ניתן רק על מעבר מעיגול E ימינה, לכיוון הריבוע האפור, וערכו 1. ערך הדיסקונטינג בשאלתנו זו $= 1 - \gamma$.

1. תארו את המשחק במושגים של *MDP*. מהי הסביבה? מהם המצבים, הפעולות, סיכויי המעבר, והגמולים המתארים אותה? מיהו הסוכן? מהי המדיניות שלו?

.1

אנו נאבקים באלימות כלכלית, מינית ופיזית.

{A, B, C, D, E} $\vdash_{\text{PA}} S \rightarrow \neg \neg A \rightarrow A$

רְקָדֶה וְבָדָק כִּי מֵעַתָּה תְּלַבֵּשׂ נְסָמָךְ וְכֵן : נְסָמָךְ

$$P = 0.5 \quad (4\pi \cdot \sigma^2_N \cdot \int_{\Delta} \text{d}\Omega) \approx 0.51$$

תְּמִימָנָה וְבָרֶךְ נַעֲמָן

2. מצאו את ערכי המצבים בהנטן המדיניות של הסוכן (s). $V_{\pi}(s)$

*רמז: התחילו מלחשב את הערך של המצב C , $(C)_\pi$ ולאחר מכן חשבו את הערכים האחרים באמצעות משוואות בלמן. הסבירו כיצד ניתן את התוצאה שקיבלתם.

$$V_{\pi}(s) = \mathbb{E}_{\pi} \left[\sum_{t=0}^{\infty} \gamma^t r_t | s_0 = s \right] = \sum_a \pi(a|s) \left[r(s,a) + \gamma \sum_{s'} P(s'|s,a) V_{\pi}(s') \right]$$

• **נקודות מילוי** fill_N left right $A = \{\text{left}, \text{right}\}$, $S = \{\text{A,B,C,D,E}\}$ יי'ס

• גורנשטיין מושגון פוליאי ר' = 1, $\pi(a|s) = 0.5$ $s \in S, a \in A$ סט $\{N\}$

$$V_{f_k}(C) = 0.5 \left(r(C, \text{left}) + 1 \cdot V_{f_k}(B) \right) + 0.5 \left(r(C, \text{right}) + 1 \cdot V_{f_k}(D) \right) = \frac{V_{f_k}(B) + V_{f_k}(D)}{2}$$

NP |alka

$$V_{f_k}(B) = 0.5 \left(r(\overset{O}{B}, \text{left}) + 1 \cdot V_{f_k}(A) \right) + 0.5 \left(r(\overset{O}{B}, \text{right}) + 1 \cdot V_{f_k}(C) \right) = \frac{V_{f_k}(A) + V_{f_k}(C)}{2}$$

$$V_{R^*}(D) = 0.5 \left(r(D_{\text{left}}) + 1 \cdot V_{R^*}(C) \right) + 0.5 \left(r(D_{\text{right}}) + 1 \cdot V_{R^*}(E) \right) = \frac{V_{R^*}(C) + V_{R^*}(E)}{2}$$

$$V_R(A) = 0.5 \left(r(A, \text{left}) \right) + 0.5 \left(r(A, \text{right}) + 1 \cdot V_R(B) \right) = \frac{V_R(B)}{2}$$

$$\text{V}_{\text{rc}}(E) = 0.5 \left(\overbrace{r(E,\text{left}) + 1}^0 \cdot \text{V}_{\text{rc}}(\text{D}) + \overbrace{r(c,\text{right}) + 1}^1 \right) = \frac{\text{V}_{\text{rc}}(\text{D}) + 1}{2}$$

trial prison

$$\left\{ \begin{array}{l} f V_R(C) = V_R(B) + V_R(D) \\ f V_R(B) = V_R(A) + V_R(C) \\ f V_R(D) = V_R(C) + V_R(E) \\ f V_R(A) = V_R(B) \\ f V_R(E) = V_R(D) + 1 \end{array} \right.$$

: $P(N \neq 0) = S_p$ ≈ 0.141 $S_{\text{infty}} = \sqrt{R_1}$

$$V_R(A) = \frac{1}{\delta} V_R(B) \quad \downarrow \quad \delta V_R(B) = \frac{1}{\delta} V_R(B) + V_R(C) \quad \Rightarrow \quad \delta V_R(B) = \delta V_R(C)$$

$$V_{R^+}(E) = \frac{1}{\delta} (V_R(D) + 1) \quad \Rightarrow \quad \delta V_R(D) = V_R(C) + \frac{1}{\delta} (V_R(D) + 1) \quad \Rightarrow \quad \delta V_R(D) = \delta V_R(C) + 1$$

$$\Rightarrow V_{rc}(B) = \frac{2}{3} V_{rc}(C) \stackrel{(1)}{\Rightarrow} \quad \frac{2}{3} V_{rc}(C) = \frac{2}{3} V_{rc}(C) + \frac{1}{3}(2V_{rc}(C) + 1) \Rightarrow V_{rc}(C) = \frac{1}{2}$$

WIR SIND DAS PUNKT V_R(C) PSS NUR IN N

$$V_{1,t}(B) = \frac{2}{6} = \frac{1}{3} \quad \Rightarrow \quad V_{\pi}(A) = \frac{1}{6}$$

$$V_{\pi}(D) = \frac{2}{3} \quad \Rightarrow \quad V_{\pi}(E) = \frac{5}{6}$$

‘I like you too John John’

$$\left\{ \begin{array}{l} V_R(A) = \frac{1}{6} \\ V_R(B) = \frac{1}{3} \\ V_R(C) = \frac{1}{2} \\ V_R(D) = \frac{2}{3} \\ V_R(E) = \frac{5}{6} \end{array} \right.$$

ל'אך כי לא מתקיים מושג של נס

הנ' הילא הילא יולן.

3. לצורך החישובים שביצעתם בסעיף 2, השתמשתם במידע שיש לכם על הסביבה (סיכום המעבר והגמולים). נניח שתליזור אין את הידע זהה והוא מנסה למדוד את פונקציית הערך שחייבת. לצורך כך, הוא עושה שני trial-ים, ומשתמש בכל העדכון TD (הכל מופיע בסוף התרגול). הניחו שקצב הלימוד $\eta = 0.1$. בשני ה trial-ים, הבחירה שעה היצור הם: ימינה, ימינה, ימינה.

נניח שהשערוק ההתחלתי של ערכי המוצבים הינו: $V_\pi(s) = 0$. השתמשו בכל העדכון TD על מנת למצוא את השערוק של ערכי המוצבים לאחר כל צעד. ככלمر השלימו את הטבלה הבאה:

trial	step	מיקום היצור	$V_\pi(A)$	$V_\pi(B)$	$V_\pi(C)$	$V_\pi(D)$	$V_\pi(E)$
1	0	C	0	0	0	0	0
	1	D	0	0	0	0	0
	2	E	0	0	0	0	0
	3	█	0	0	0	0	0.1
2	0	C	0	0	0	0	0.1
	1	D	0	0	0	0	0.1
	2	E	0	0	0	0.01	0.1
	3	█	0	0	0	0.01	0.19

*שים לב - היצור "זוכר" את ערכי המוצבים בין ה trial-ים, ככלמר, לאחר שמצאתם את השורה של step 3 ב trial הראשון, שכפלו אותה לשורה של 0 ב step 0 ב trial השני.

$$V(s_t) = V(s_{t-1}) + \eta (r_t + \gamma V(s_{t+1}) - V(s_t)) \quad \text{לפניהם הנקודותTD: C, D, E}$$

$$V(c) = 0 + 0.1(0 + 1 \cdot 0 - 0) = 0 \quad : D - \int c - N : 1 \text{ גז}$$

$$V(D) = 0 + 0.1(0 + 1 \cdot 0 - 0) = 0 \quad : E - \int D - N : 2 \text{ גז}$$

$$V(E) = 0 + 0.1(1 + 1 \cdot 0 - 0) = 0.1 \quad : \square - \int E - N : 3 \text{ גז}$$

$$V(c) = 0 + 0.1(0 + 1 \cdot 0 - 0) = 0 \quad : D - \int c - N : 1 \text{ גז}$$

$$V(D) = 0 + 0.1(0 + 1 \cdot 0.1 - 0) = 0.01 \quad : E - \int D - N : 2 \text{ גז}$$

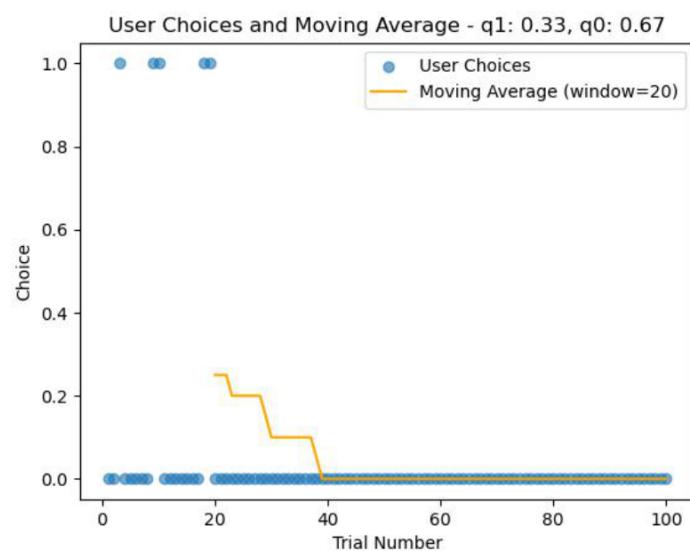
$$V(E) = 0.1 + 0.1(1 + 1 \cdot 0 - 0.01) = 0.19 \quad : \square - \int E - N : 3 \text{ גז}$$

- מה היו העריכים ההתחלתיים של כל אחד ממחצביים?
 - הערך של איזה מחצבים השתנה לאחר trial הראשון? איפה הסתיים ה trial הראשון - מימין או משמאלי? נזכיר.

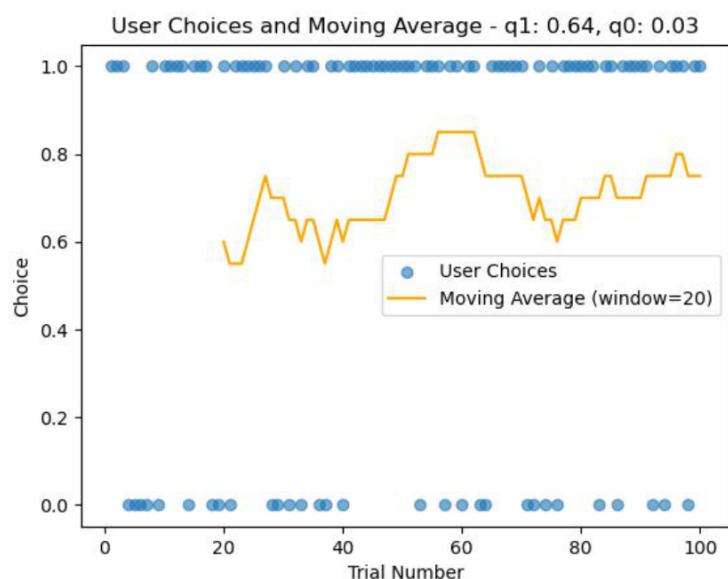
0.5 -> 10% error in λ

לעומת הדוגמה הקודמת, מטרת הבדיקה היא לא בדיקת השוני בין אמצעים, אלא בדיקת השוני בין אמצעים אחדים (בנוסף ל- μ_0) ביחס למשתנה אחד בלבד (בנוסף ל- σ^2). נניח ש- μ_1 ו- μ_2 הם אמצעי ה- μ עבור קבוצות נבדיקות נפרדיות, ו- μ_0 הוא אמצעי ה- μ עבור קבוצה שלרשות. מטרת הבדיקה היא בדיקת השוני בין אמצעי ה- μ של קבוצות נבדיקות נפרדיות (בנוסף ל- μ_0).

2.5 הצגת הנתונים



• f_1, f_0 מוגדרים על \mathbb{R}



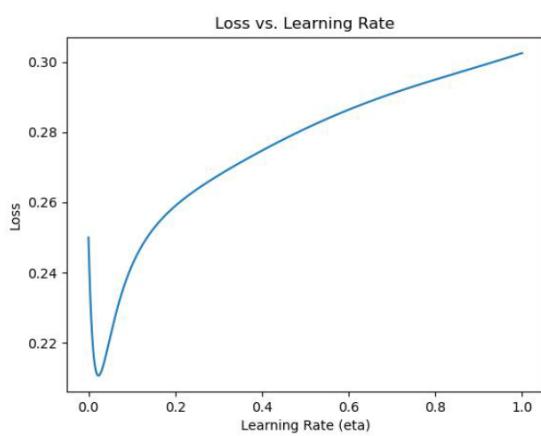
(10) 2

הוּא מִבְּרָכָה לְפָנֶיךָ יְהוָה אֱלֹהֵינוּ וְעַמּוּדֵינוּ בְּבָרְכָתֶךָ
בְּבָרְכָתֶךָ יְהוָה אֱלֹהֵינוּ וְעַמּוּדֵינוּ בְּבָרְכָתֶךָ
בְּבָרְכָתֶךָ יְהוָה אֱלֹהֵינוּ וְעַמּוּדֵינוּ בְּבָרְכָתֶךָ

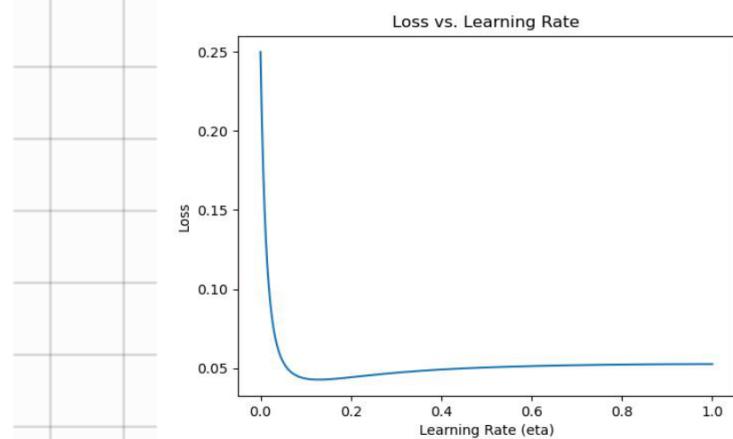
2.6 ניתוח והתאמת מודל

הסבירו במילים מודיע *loss* יתיר את תהליכי הלמידה של הנבדק “טייצר” רצף הסתברויות שעבورو ה-*יהיה קטן*.

```
experiment 2:  
Best eta: 0.023, Lowest loss: 0.2106701019479911
```



```
experiment 1:  
Best eta: 0.129. Lowest loss: 0.042681233100134074
```



הנ'ת עלי זען וו' ואל.

דיאן 2.7

כתבו בקורס ברורה מהו הפרמטר שהתקבל עבור הנבדק שעובדו $q_0 = \frac{1}{3}$, $q_1 = \frac{2}{3}$, ומהו הפרמטר שהתקבל עבור הנבדק שעובדו q_0 , q_1 .

בסיום התרגיל, אנחנו נאסף את קצבי הלימוד שוכלתם קיבלתם ונشرط שתี้ היסטוגרמות: אחת של קצבי הלימוד שהתקבלו מהנדקרים שעברו את הניסוי עם $q_0 = \frac{1}{3}$, ואחת נוספת שהתקבלו מהנדקרים שעבורם הגרלות את q_1 באקראי.

1. נניח שמודל *LRIA* לא מתואר התנהוגות אנושית בצורה טובה. במלils אחרות, הפרמטר שקיבלו לא מאפיין רק את התנהוגות הנבדק אלא מושפע גם ממערך הניסוי. במקרה זה, מה אפשר לצפות מההסתטוגרמות? האם יהיה דומות זו לזו? שונות? באיזה צורה?
 2. אם מודל *LRIA* כן מתואר התנהוגות אנושית בצורה טובה, מה נצפה מההסתטוגרמות?

$$\eta_2 = 0.023 \cdot ! \quad \eta_1 = 0.129 \quad : 10^3 \quad | " 10^3 \ell \quad \text{!} \circ$$

לעתה נזכיר את הדרישות שקיים בפונקציית f על מנת שפונקציית $\int f$ תהיה מוגדרת:

чисוביות וקוגניציה - תרגיל 6

להגשה עד: 07/03/2024

שימו לב: בתחילת התרגיל מופיעות כמה שאלות הקדמה אמריקאיות. יש להגיש את התשובות אליהן, עם משפט נימוק קצר לכל שאלה. שאלה 1 היא שאלה אנלידית ושאלה 2 כוללת ביצוע ניסוי התחנוגטי וחומר מודול בצורה נומריאת.

שאלות הקדמה

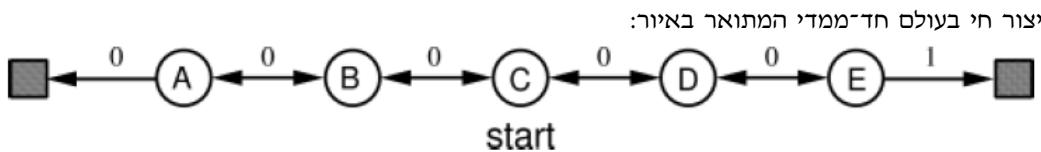
1. לפי ההגדרות שראינו בכיתה, מה נכון לומר על הקשר בין $V_\pi(s)$ לבין $Q_\pi(s, a)$?

- (א) כדי להגיע מ- $V_\pi(s)$ ל- $Q_\pi(s, a)$, יש למצוע על פני כל הפעולות האפשריות a , בהתאם למדייניות $\pi(a|s)$
- (ב) כדי להגיע מ- $V_\pi(s)$ ל- $Q_\pi(s, a)$, יש למצוע על פני כל המצביעים הבאים האפשריים s' , בהתאם למעברי העולם $P(s'|s, a)$
- (ג) כדי להגיע מ- $Q_\pi(s, a)$ ל- $V_\pi(s)$, יש למצוע על פני כל הפעולות האפשריות a , בהתאם למדייניות $\pi(a|s)$
- (ד) כדי להגיע מ- $Q_\pi(s, a)$ ל- $V_\pi(s)$, יש למצוע על פני כל המצביעים הבאים האפשריים s' , בהתאם למעברי העולם $P(s'|s, a)$

2. כאשר אנו מדברים על "מודל" ומפרידים בין Model-Free ו-Model-Based, מה כלל בתחום הגדרת המודל?
בחרו מכל התשובות הנכונות

- (א) $P(s'|s, a)$
(ב) $\pi(a|s)$
(ג) $P(R|s, a)$
(ד) $V_\pi(s)$
(ה) $V^*(s)$
(ו) $Q_\pi(s, a)$
(ז) $Q^*(s, a)$

1. למידת חיזוק - ייצור בעולם חד מימדי



היצור מתחילה בעיגול C ובכל צעד, נע שמאלה או ימינה בהתאם להסתברות $0.5 = p$. כאשר הוא מגע לאחד מהריבועים הכהים נגמר ה trial, הוא מוחזר אוטומטית במצב C ומתחילה trial חדש. הגמול היחיד שיש במשחק ניתן רק על מעבר מעיגול E ימינה, לכיוון הריבוע האפור, וערךו 1. ערך הדיסקונטינג בשאלתנו זו $\gamma = 0.9$.

1. תארו את המשחק במושגים של MDP . מהי הסביבה? מהם המצבים, הפעולות, סיכויי המעבר, והגמולים המתארים אותה? מיהו הסוכן? מהי המדיניות שלו?

2. מצאו את ערכי המצבים בהנתן המדיניות של הסוכן (s). V_π .

*רמז: התחלו מלהշב את הערך של המצב C ($V_\pi(C)$) ולאחר מכן חשבו את הערכים האחרים בעזרת משוואות בלבדן. הסבירו בצורה איקוונית את התוצאה שקיבתם.

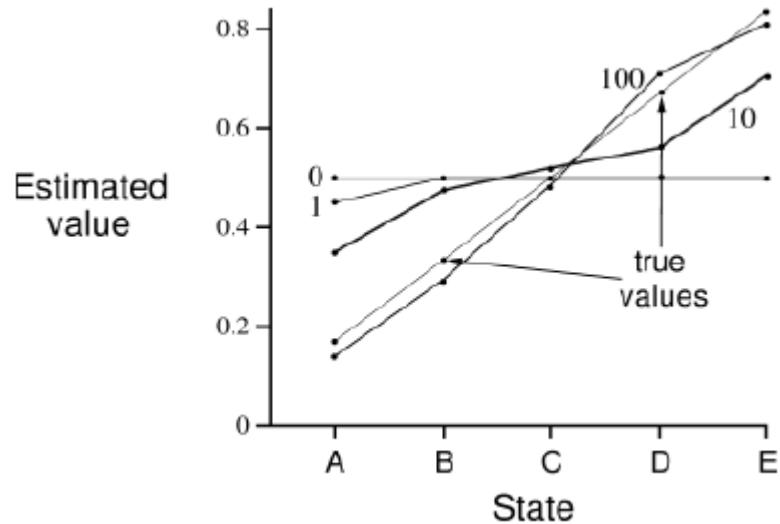
3. לצורך החישובים שביצעתם בסעיף 2, השתמשתם בידע שיש לכם על הסביבה (סיכויי המעבר והגמולים). נניח שתליצור אין את הידע הזה והוא מנסה ללמידה את פונקציית הערך שחייבת. לצורך כך, הוא עושה שני trial-ים, ומשתמש בכלל העדכון TD (הכלל מופיע בסוף התרגול). הניחו שקצב הלמידה $\alpha = 0.1$. בשני ה trial-ים, הבחירה שעשאה הייצור הם: ימינה, ימינה, ימינה.

נניח שהשערון ההתחלתי של ערכי המצבים הינו: $0 = V_\pi(s)$. השתמשו בכלל העדכון TD על מנת למצוא את השערון של ערכי המצבים לאחר כל צעד. כולם השלימו את הtablלה הבאה:

trial	step	מקום הייצור	$V_\pi(A)$	$V_\pi(B)$	$V_\pi(C)$	$V_\pi(D)$	$V_\pi(E)$
1	0	C	0	0	0	0	0
	1	D					
	2	E					
	3	■					
2	0	C					
	1	D					
	2	E					
	3	■					

*שיםו לב - הייצור "זוכר" את ערכי המצבים בין ה trial-ים, כלומר, לאחר שמצאתם את השורה של step 3 ב trial הראשון, שכפלו אותה לשורה של 0 ב step 3 ב trial השני.

4. הגרף הבא מתאר את למידת ערכי המצבים בהנתן המדיניות של הייצור, באמצעות אלגוריתם TD, על פני 100 trial-ים. המספר ליד כל אחד מהעוקמות מציין את מספר ה trial שבסופו שורטטו ערכי המצבים המשוערכים.



- מה היו הערכים ההתחלתיים של כל אחד מהמצבים?
- הערך של איזה מהמצבים השתנה לאחר trial הראשון? איפה הסטיים ה trial הראשון - מימין או משמאל? נמקו.

2 התאמת מודל לתנהגות

בכיתה דיברתם על כללי למידה שנייתן למدل בעורכם התנהגות אנושית. בתרגיל זה נבדוק עד כמה המודלים האלה אכן מתארים בצורה טובה התנהגות אנושית. אנא קראו את כל השאלה לפני שתתחלו.

2.1 פרדיגמת הניסוי

בשאלה זו אתם תרצו ניסוי התנהוגות על שני משתתפים בפרדיגמה ניסויית פשוטה שנקראת two alternative forced choice. בניסוי זה:

- המשתתף מבצע רצף של trials כאשר בכל trial הוא צריך לבחור בין שתי אופציות: $y = 0$ או $y = 1$. לאחר הבחירה, הוא מקבל גמול ($r = 1$) או לא מקבל גמול ($r = 0$). המשתתף לא מודע לחוקיות שועמדת בבסיס קבלת הגמול.
- החוקיות לפיה מתќבל לא מתќבל גמול היא כזו: עבור כל אופציה y , ישנה הסתברות q_y שמתחילה את הסיכוי לקבל גמול בהינתן שהאופציה נבחרה:

$$P(r = 1|y = 0) = q_0$$

$$P(r = 1|y = 1) = q_1$$

כאשר ההסתברויות לא לקבל גמול הן ההסתברויות המשלימות:

$$P(r=0|y=0) = 1 - q_0$$

$$P(r=0|y=1) = 1 - q_1$$

בצורה פורמלית אפשר לכתוב שהגמול r בהינתן הבחירה y מתפלג בהתאם לתרגול עם פרמטר q_y

$$r|y \sim \text{Ber}(q_y)$$

(a) מסמן התפלגות ברנולי עם פרמטר a , כלומר משתנה מקרי שמקבל 1 בסיכוי a ו 0 בסיכוי $a-1$

- מטרת המשתנה היא לקבל כמה שיותר גמולים. נגיד שוב – הוא לא מודע לסיכוי קבלת הגמול – כל המידע שיש לו זה האם הוא קיבל או לא קיבל גמול בעבר הבחירה שלו.

2.2 המודל שתבחן

בכיתה, דיברתם על כלל למידה $LRIA$. זהו מודל שיכול לתאר את תהליך הלמידה והבחירה של בני אדם בפרדיגמה לעיל. במודל זה, אנחנו מגדירים שהמשתנה בניסוי בוחר בין שתי האופציות בצורה הסתברותית: הוא בוחר באופציה y בהסתברות p ובהתאם, הוא בוחר באופציה $0 = y$ בהסתברות $p-1$. כלומר, לפי המודל, הבחירה של המשתנה ב trial ה – t מתפלגת כך:

$$y_t \sim \text{Ber}(p_t)$$

בנוסף, המודל מתאר איך משתנה ההסתברות p בעקבות קבלת גמול ב trial ה – t :

$$p_{t+1} = p_t + \eta \cdot r_t \cdot (y_t - p_t) \iff \Delta p_t = \eta \cdot r_t \cdot (y_t - p_t)$$

כאשר η קצב הלימוד, r_t הוא הפרמטר היחיד במודל. בהמשך, אתם תשתמשו במודל זה על מנת לתאר את ההתנהגות של המשתתפים בניסוי שלהם ותבדקו האם הוא טוב.

2.3 בניית הניסוי

כתבו פונקציה ש谋יצה ניסוי ושומרת את התוצאות. הניסוי מאופיין על ידי שני הפרמטרים q_1, q_0 (מתוארים לעיל). בכלל צעד, על הנבדק לבחור פעולה דרך המקלדת עי' הקשה של 0 או 1. השתמשו בפונקציה `input(str)` שמצוינה טקסט ומחייב קלט מהמשתמש (וודאו שהקלט הוא אחד מהאפשרות החוקיות). לאחר הבחירה, התוכנה מגירלה את הגמול (0 או 1) מההתפלגות המתאימה ומציג אותה לנבדק למשך s שניות – השתמשו בפונקציה `time.sleep(s)` (בפייטון) או ב `pause(s)` (במטלב) שעוצרת את התוכנה ל s שניות. לאחר מכן, יש 'לנקות' את המסך והנבדק עבר ל trial הבא. ע"מ 'לנקות' את המסך השתמשו בפקודה `clc` במטלב. בפייטון, הדרך הקללה ביותר היא `('cls')os.system('cls')` לוינדוס ו `('clear')os.system('clear')` בלינוקס.

2.4 הרצת הניסוי

בחרו שני נבדקים.ות (חברים שאינם תלמידים בקורס, בני משפחה, וכו') שעלייהם תריצו את הניסוי.

כל נבדק יעבור את הניסוי פעם אחת, והניסוי אורך $T = 100$ trials. הפרמטרים q_0, q_1 שמאפיינים את הניסוי הינם:

- עברו נבדק א': $q_1 = \frac{2}{3}, q_0 = \frac{1}{3}$.
- עברו נבדק ב': הגרילו באופן אקרי q_0 מתוך התפלגות איחידה בין 0 ל 1 (כלומר הגרילו את שני הפרמטרים האלה פעם אחת לפני תחילת הניסוי).

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

2.5 הצגת הנתונים

עבור כל נבדק:

1. הציגו גרף שמתאר את התנהוגות שלו: שרטטו את הבחירה שלו (כלומר $y = 0$ או $y = 1$) כפונקציה של מספר ה trial. אל תחויבו בקו את הנקודות בgap. כתבו בכותרת מהם ההסתברויות q_0, q_1 .
2. לפני שנתאים מודל לתנהוגות, נרצה לראות אם אנחנו רואים שינויים בהסתברויות הבחירה של הנבדק. בשביל לעשות זאת, שרטטו על גבי אותו הגרף ממוצע נע (moving average) של בחירות הנבדק כפונקציה של מספר ה trial (השתמשו בחלון של 20).
3. האם אתם רואים שינויים בגרף לאורך זמן? האם הם מתאימים (בצורה איקוונית) לפרמטרים q_0, q_1 שמאפיינים את הניסוי שעבר הנבדק?

2.6 ניתוח והתאמת מודל

מודל LRIA מאופיין על ידי פרמטר יחיד: קצב הלימוד η . ככלומר, אם המודל מותאר נכון לתנהוגות אנושית, אז כל אדם מאופיין ע"י פרמטר η בלבד.
Cut, אנחנו נניח שהנבדקים מתחילה לפיה כל הלמידה $LRIA$, ותחת ההנחה הזאת נתאים להם את הפרמטר שמאפיין אותם בצורה הטובה ביותר. בשביל לעשות זאת:

1. כתבו פונקציה שמקבלת רצף של בחירות, רצף של גמולים ופרמטר קצב לימוד η , ומחזירה את רצף ההסתברויות p_t שאוthon חוזה המודל. הניתו שההסתברות ההתחלתית היא $p_1 = \frac{1}{2}$ (כלומר - אנחנו מניחים שב trial הראשון, לפני שקיבלו איזשהו גמול, הנבדקים בוחרים באקרי).
2. Cut, נשאל מהי ה η שעובדת רצף ההסתברויות שהוזה המודל מותאים בצורה הטובה ביותר לתחנהוגות הנבדקים. בשביל לעשות זאת צריך להגיד מה הכוונה ב"מתאים בצורה הטובה ביותר". בתרגיל זה אנחנו נניח שהמטרה היא למצוא את ה η שמשמעותו את הביטוי הבא:

$$Loss = \frac{1}{T} \sum_{t=1}^T (p_t - y_t)^2$$

כאשר p_t הוא ההסתברות שחוזה המודל ב trial ה - t ו y_t הוא בחירת הנבדק ב trial ה - t .
הסבירו במילים מדויקים שמתארת טוב יותר את תהליך הלמידה של הנבדק “תיצר” רצף הסתברויות שעבורו הLoss יהיה קטן.

3. עבור כל נבדק, מצאו את הערך של η שמצויר את ה Loss. כדי לעשות זאת, בחרו אחת משלוש הדרכים הבאות (אם תרצו אתם מוזמנים לנסוט בכמה דרכים ולודא שיוצאה לכם תוצאות דומות):

דרך א

- כתבו פונקציה שמקבלת רצף הסתברויות p_t ורץ של בחירות y_t ומחזירה את ה Loss.
- הריצו את הפונקציה שכתבتم בסעיף 1 עבור כל ערכי η בתחום בין 0 ל 1 בקפיצות של 0.001. עבור כל הרצתה, השתמשו בפונקציה שכתבتم בסעיף הנוכחי על מנת לחשב את ה Loss.
- שרטטו גרף של ה Loss כפונקציה של η . בחרו את קצב הלימוד שעבורו ה Loss הוא המינימי.

דרך ב

- השתמשו בgradient descent לשם כך, ראשית נכתוב את ה Loss כפונקציה של הפרמטר η .

$$\begin{aligned} Loss(\eta) &= \frac{1}{T} \sum_{t=1}^T (p_t(\eta) - y_t)^2 \\ &= \frac{1}{T} \left[(p_1(\eta) - y_1)^2 + \sum_{t=2}^T (p_t(\eta) - y_t)^2 \right] \\ &= \frac{1}{T} \left[\left(\frac{1}{2} - y_1 \right)^2 + \sum_{t=2}^T (p_{t-1} + \eta \cdot r_t \cdot (y_{t-1} - p_{t-1}) - y_t)^2 \right] \end{aligned}$$

- לאחר מכן, נחשב את הגרדיינט (במקרה זה, יש לנו רק פרמטר אחד η ולכן הגרדיינט הוא פשוט הנגזרת לפיה).

$$\begin{aligned} \frac{d}{d\eta} Loss(\eta) &= \frac{d}{d\eta} \left(\frac{1}{T} \left[\left(\frac{1}{2} - y_1 \right)^2 + \sum_{t=2}^T (p_{t-1} + \eta \cdot r_t \cdot (y_{t-1} - p_{t-1}) - y_t)^2 \right] \right) \\ &= \frac{1}{T} \sum_{t=2}^T (p_{t-1} + \eta \cdot r_t \cdot (y_{t-1} - p_{t-1}) - y_t) \cdot r_t \cdot (y_{t-1} - p_{t-1}) \end{aligned}$$

- כתבו פונקציה שמקבלת רצף בחירות y_t , רצף גמולים r_t , רצף הסתברויות p_t , וקצב לימוד הנוכחי η ומחזירה את הגרדיינט של ה Loss.
- אתחלו את הפרמטר $\eta_0 = \eta$ באופן שרירותי בין 0 ל 1. הריצו לולאה שככל ריצת:

- מיצרת, ע"פ ערך ה η הנוכחי, רצף הסטבריות p_t (ע"י שימוש בפונקציה שכתבתם בסעיף 1).
- לוקחת את רצף הסתבריות p_t , רצף הבחרות y_t ורצף גמולים r_t ומוחירה את הגרדיינט של ה $Loss$ (ע"י שימוש בפונקציה שכתבתם בסעיף הקודם)
- מעדכנת את הפרמטר η ע"פ :

$$\eta_{t+1} = \eta_t - \epsilon \frac{d}{d\eta} Loss(\eta)$$

שימו לב, אנחנו משתמשים בשיטת gradient descent על מנת למצוא את פרמטר קצב הלמידה η המזער את ה $Loss$. כחלק מהשיטה, אנחנו צריכים לבחור פרמטר קצב לימוד. זהו פרמטר אחר - קראנו לו ϵ כדי למנוע בלבול. בחרו אותו מספיק קטן על מנת שתקבלו התוצאות.

- ערכו כאשר השינוי ב η יהיה מאד קטן (כלומר, מעד על כך שהאלגוריתם הטענן).

- שרטטו גרף של η כפונקציה של הזמן (כלומר צעדי עדכון). מהו ה η שקיבלתם?

דרך 2

- השתמשו בפונקציות מובנות של אופטימיזציה על מנת למצוא את η שמייער את ה $Loss$. מהו ה η שקיבלתם?

2.7 דיוון

כתבו בפורה מהו הפרמטר שהתקבל עבור הנבדק שעבורו $q_0 = \frac{1}{3}$, $q_1 = \frac{2}{3}$, ומהו הפרמטר שהתקבל עבור הנבדק שעבורו הגרלתם את q_0, q_1 . בסיום התרגיל, אנחנו נאוסף את קבועי הלמידה ש כולכם קיבלתם ונشرط שתי היסטוגרמות: אחת של קבועי הלמידה שהתקבלו מהנבדקים שעברו את הניסוי עם $q_0 = \frac{1}{3}, q_1 = \frac{2}{3}$, ואחת נוספת שהתקבלו מהנבדקים שעבורם הגרלתם את q_0, q_1 באקראי.

1. נניח שמודל $LRIA$ לא מותאר התנהגות אנושית בפורה טובה. במקרים אחרים, הפרמטר שקיבלו לא מאפיין רק את התנהגות הנבדק אלא מושפע גם ממערך הניסוי. במקרה זה, מה אפשר לצפות מההיסטוגרמות? האם היה ייחו דומות זו לזו? שונות? באיזה צורה?
2. אם מודל $LRIA$ כן מותאר התנהגות אנושית בפורה טובה, מה נצפה מההיסטוגרמות?