CSCI 5512 • HW 4

Nikki Kyllonen
kyllo089 • 4/8/19

① Stochastic Dominance

(1) A : N(0, 1) B : N(0.5, 0.5) → Does B stochastically dominate?

* First-order stochastic dominance (FSD) : $\forall_x\, P(X \geq x) \geq P(Y \geq x)$ and $\exists_x\, P(X \geq x) > P(Y \geq x)$

X gives at least as high a probability of receiving at least x as does Y and for some x values, X gives a strictly higher probability

$P(a \leq X \leq b) = \int_a^b f_x(x)\,dx \longrightarrow f_x(x) = N(\mu_x, \sigma_x^2)$

$F_x(x) = \int_{-\infty}^{x} f_x(u)\,du$ = cumulative distribution function $= P(X \leq x)$

⇒ FSD using CDFs : $\forall_x\, F_x(x) \leq F_y(x)$ and $\exists_x\, F_x(x) < F_y(x)$

→ if B stochastically dominated over A

$\forall_x\, F_B(x) \leq F_A(x)$ and $\exists_x\, F_B(x) < F_A(x)$

$\boxed{\text{FALSE}}$ @ $x > 1$, $F_B(x) \geq F_A(x)$

* insert CDF graph/plot

(2) $X : N(\mu_x, \delta_x^2)$ $Y : N(\mu_y, \delta_y^2) \rightarrow$ when does Y stochastically dominate X?

$\rightarrow \forall_x F_y(x) \leq F_x(x)$ and $\exists_x F_y(x) < F_x(x)$

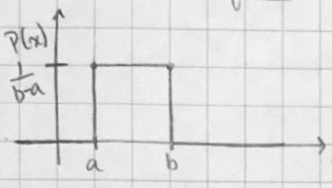
$\boxed{@\ \delta_x^2 = \delta_y^2 : \quad \mu_y < \mu_x} \longrightarrow$ $F_y(x)$ would be the same form as $F_x(x)$, but translated to the left $\therefore$ being less than $F_x(x)$

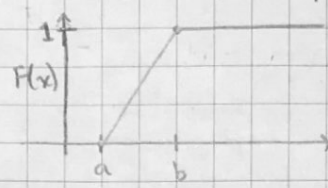
@ $\mu_y = -0.5$, $\mu_x = 0$

@ $\mu_y = -0.8$, $\mu_x = 0$

MIQUELRIUS

(3) * can a <u>uniform</u> distribution stochastically dominate a <u>normal</u> dist?





UNIFORM DISTRIBUTIONS

$\Rightarrow \forall_x \; F_U(x) \leq F_N(x)$ and $\exists_x \; F_U(x) < F_N(x)$

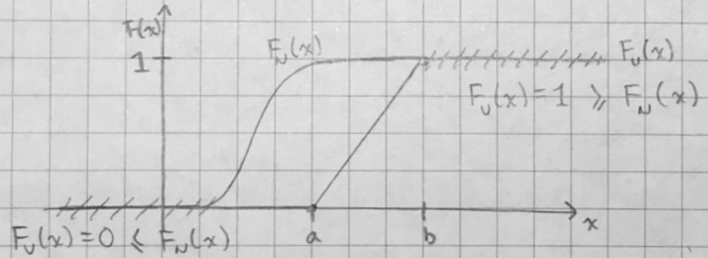
$\boxed{NO}$: $F_U(x) = 1$ @ $x > b$ and $F_N(x) = 1$ @ $x = +\infty$

$\therefore$ @ $x > b$, $F_U(x) \geq F_N(x)$

* can a <u>normal</u> distribution stochastically dominate over a <u>uniform</u> dist?

$\Rightarrow \forall_x \; F_N(x) \leq F_U(x)$ and $\exists_x \; F_N(x) < F_U(x)$

$\boxed{NO}$: @ $x < a$, $F_U(x) = 0$ $\therefore$ it will be <u>less</u> than $F_N(x)$



② Slot Machines

X:

P(X)	rewards
0.1	$20
0.3	$5
0.4	$1
0.2	$0

Y:

P(Y)	rewards
0.05	$40
0.25	$4
0.30	$2
0.40	$0

Z:

P(Z)	rewards
0.25	$10
0.25	$5
0.25	$2
0.25	$0

* $4 to play
* single turn (X ‖ Y ‖ Z)

$E[X] = (0.1)U(\$20) + (0.3)U(\$5) + (0.4)U(\$1) = \boxed{3.9}$

$E[Y] = (0.05)U(\$40) + (0.25)U(\$4) + (0.30)U(\$2) = \boxed{3.6}$

$E[Z] = (0.25)U(\$10) + (0.25)U(\$5) + (0.25)U(\$2) = \boxed{4.25}$ **

*let the Utility function be

$U(x) = x$

x "utility worth of money"

* without paying to know, 1/3 chance of X, Y, Z machines

→ E[randomly choosing X] = (1/3)(3.9) = $\boxed{1.3}$

→ E[" Y] = (1/3)(3.6) = $\boxed{1.2}$

→ E[" Z] = (1/3)(4.25) = $\boxed{1.417}$ ** greatest amount of utility money w/o knowing

MIQUELRIUS

* with paying to know about one machine

→ choose and identify X :

→ E[choosing X] = (1.0)(3.9) = $\boxed{3.9}$

→ E[" " Y] = (1/2)(3.6) = 1.8

→ E[" " Z] = (1/2)(4.25) = 2.125

→ choose and identify Y :

→ E[choosing X] = (1/2)(3.9) = 1.95

→ E[" " Y] = (1.0)(3.6) = $\boxed{3.6}$

→ E[" " Z] = (1/2)(4.25) = 2.125

→ choose and identify Z :

→ E[choosing X] = (1/2)(3.9) = 1.95

→ E[" " Y] = (1/2)(3.6) = 1.8

→ E[" " Z] = (1.0)(4.25) = $\boxed{4.25}$ ** greatest amount of utility money w/ knowing one machine

* already have to pay $4
therefore

4.25 − 4 = 0.25 ⟶ most possible earning w/ knowing
1.417 − 4 = −2.583 ⟶ max possible "earning" (least possible loss) w/o knowing

0.25 + (+2.583) = $\boxed{2.833}$ worth / gain of knowing

→ pay more upfront for the higher chance of <u>NOT</u> <u>losing</u> money?

∴ pay an extra $\boxed{2.833}$ "utility worth of money" on top of the original $4

3 Value Iteration

(1) $ python3 value_iteration.py ⟶ 5 iterations

X	50	X
X	34.05	17.71
-50	11.70	0.52
X	2.98	-1.70

X	50	X
X	↑	←
-50	↑	↑
X	↑	←

* with starting utilities :

end state

X	50	X
X	0	0
-50	0	0
X	0	0

end state

(2) $ python3 value_iteration.py zeros → 3 iterations

X	0	X
X	-0.23	-4.95
0	-1.51	-11.76
X	-4.71	-5.03

X	O	X
X	←	←
O	←	←
X	↑	↓

* with starting utilities:

X	O	X
X	O	O
O	O	O
X	O	O

end

end

(3) largest difference = 34.05 + (+0.23) = $\boxed{34.28}$

theoretical band for utility: $U(s_0, s_1, \ldots) \leq \frac{R_{max}}{1-\gamma} = \frac{50}{1-0.8} = \frac{50}{2/10} = \boxed{250}$

$\boxed{34.28 < 250}$ * less than the band

④ Policy Iteration

$ python3 policy_iteration.py → 3 iterations

X	50	X
X	34.38	18.78
-50	12.53	2.30
X	4.70	1.03

X	50	X
X	↑	←
-50	↑	↑
X	↑	←

⑤ POMDP

→ actions : { left , down } → P(going in intended dir) = 70%
→ P(going 90° off) = 15% left / 15% right

→ all possible belief states that result from 2 actions ?

* Rewards:

-1	-4
2	-2

* Initial guesses:

20	30
0	50

%

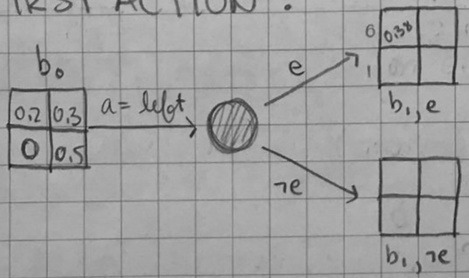
* starting belief state

* P(e|s):

0.3	0
0.9	0.2

likelihood of evidence e being True given state s

* FIRST ACTION :



$$b'(s') = \alpha\, P(e|s') \cdot \sum_s P(s'|s,a) \cdot b(s)$$

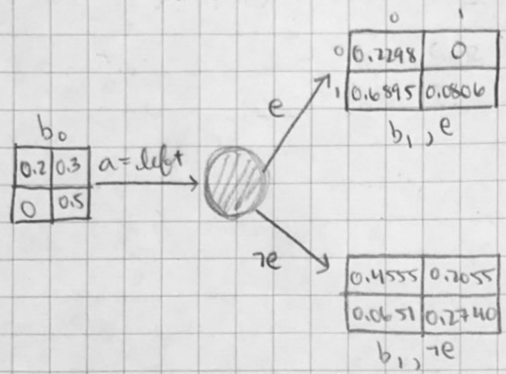
* let s' = (0,0) and action = left : e

s	actual move	$P(s'\|s,a)$	
(0,0)	left	0.7	→ (0.7)(0.2)
(0,0)	up	0.15	→ (0.15)(0.2)
(0,1)	left	0.7	→ (0.7)(0.3)
(1,0)	up	0.15	→ (0.15)(0)

$\boxed{b'(0,0) = 0.38}$

MIQUELRIUS

* FIRST ACTION - continued



* let $s' = (0,1)$: e

s	actual move	$P(s' \mid s,a)$
(0,1)	up	0.15
(1,1)	up	0.15

$b'(0,1) = (0.3)(0.15) + (0.5)(0.15) = \boxed{0.12}$

* let $s' = (1,0)$: e

(1,1)	left	0.7
(0,0)	down	0.15
(1,0)	left	0.7
(1,0)	down	0.15

$b'(1,0) = (0.5)(0.7) + (0.2)(0.15) + (0) + (0) = \boxed{0.38}$

* let $s' = (1,1)$: e

s	actual	P
(1,1)	down	0.15
(0,1)	down	0.15

$b(1,1) = (0.5)(0.15) + (0.3)(0.15) = \boxed{0.12}$

* apply $P(e \mid s)$:

$$\begin{array}{c|c} 0.38 & 0.12 \\ \hline 0.38 & 0.2 \end{array} \Rightarrow \begin{array}{c|c} (0.38)(0.3) & (0) \\ \hline (0.38)(0.9) & (0.2)(0.2) \end{array} \Rightarrow \begin{array}{c|c} 0.114 & 0 \\ \hline 0.342 & 0.04 \end{array}$$

* apply α :

$$\alpha = \frac{1}{0.114 + 0.342 + 0.04} = \underline{2.016} \longrightarrow \underset{b_1,\ e}{\begin{array}{|c|c|} \hline 0.2298 & 0 \\ \hline 0.6895 & 0.0806 \\ \hline \end{array}} \longrightarrow R(b_1) = \sum_s b(s) \cdot R(s)$$

$= (0.23)(-1) + (0.69)(2) + (0.08)(-2)$

$= \boxed{1.1089}$

$\boxed{R(b_1) = 1.1089}$

* apply $P(\neg e \mid s)$:

$$\begin{array}{c|c} 0.38 & 0.12 \\ \hline 0.38 & 0.2 \end{array} \Rightarrow \begin{array}{c|c} (0.38)(1-0.3) & (0.12)(1-0) \\ \hline (0.38)(1-0.9) & (0.2)(1-0.2) \end{array} \Rightarrow \begin{array}{c|c} 0.266 & 0.12 \\ \hline 0.038 & 0.16 \end{array}$$

* apply α :

$$\alpha = \frac{1}{0.266 + 0.12 + 0.038 + 0.16} = \underline{1.712} \Rightarrow \underset{}{\overset{b_1,\ \neg e}{\begin{array}{|c|c|} \hline 0.4555 & 0.2055 \\ \hline 0.06507 & 0.2740 \\ \hline \end{array}}} \quad \boxed{R(b_1) = -1.695}$$

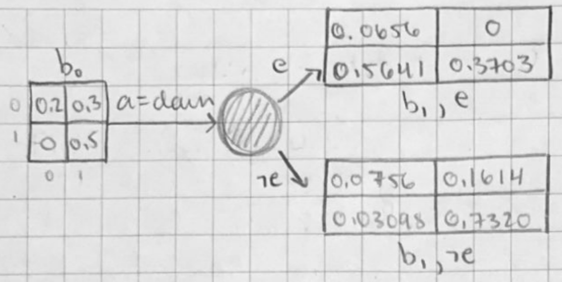
$$R(b_1) = \sum_s b(s) \cdot R(s) = (0.456)(-1) + (0.2055)(-4) + (0.06507)(2) + (0.2740)(-2)$$

$= \boxed{-1.695}$

$$P(b_1 \mid b_0, a = \text{left}) = \sum_e P(b' \mid b, a, e) \sum_{s'} P(e \mid s') \sum_s P(s' \mid a, s)\, b(s)$$

$\rightarrow e : 1(0.114 + 0.342 + 0.04) = \boxed{0.496}$ $\quad\rightarrow \neg e : 1(0.266 + 0.12 + 0.038 + 0.16) = \boxed{0.584}$

?? don't add up to 1...



* let $s' = (0,0)$: e

(0,1)	left	0.15
(0,0)	left	0.15

$b'(0,0) = (0.3)(0.15) + (0.2)(0.15)$
$= \boxed{0.075}$

* let $s' = (0,1)$: e

(0,1)	right	0.15
(0,0)	right	0.15

$b'(0,1) = \boxed{0.075}$

* let $s' = (1,0)$: e

(0,0)	down	0.7
(1,0)	left	0.15
(1,0)	down	0.7
(1,1)	left	0.15

$b'(1,0) = (0.2)(0.7) + (0.5)(0.15) = \boxed{0.215}$

* let $s' = (1,1)$: e

(1,1)	down	0.7
(0,1)	down	0.7
× (1,0)	right	0.15
(1,1)	right	0.15

$b'(1,1) = (0.5)(0.7) + (0.3)(0.7) + (0.5)(0.15)$
$= \boxed{0.635}$

* apply $P(e|s)$:

$\begin{array}{c|c} 0.075 & 0.112 \\ \hline 0.215 & 0.635 \end{array} \Rightarrow \begin{array}{c|c} (0.075)(0.3) & 0 \\ \hline (0.215)(0.9) & (0.635)(0.2) \end{array} \Rightarrow \begin{array}{c|c} 0.0225 & 0 \\ \hline 0.1935 & 0.127 \end{array}$

* apply α :

$\alpha = \dfrac{1}{0.0225 + 0.1935 + 0.127} = 2.915 \longrightarrow$

b_1, e

0.0656	0
0.5641	0.3703

$R(b_1, e) = \sum_s b(s) \cdot R(s) = (0.0656)(-1) + (0.5641)(2) + (0.3703)(-2)$
$\boxed{= 0.322}$

$P(b_1 | b_0, a = down) = 1(1/\alpha) = \boxed{0.343}$ @ e

* apply $P(\neg e|s)$:

$\begin{array}{c|c} 0.075 & 0.112 \\ \hline 0.215 & 0.635 \end{array} \Rightarrow \begin{array}{c|c} (0.075)(0.7) & (0.112)(1) \\ \hline (0.215)(0.1) & (0.635)(0.8) \end{array} \Rightarrow \begin{array}{c|c} 0.0525 & 0.112 \\ \hline 0.0215 & 0.508 \end{array}$

* apply α :

$\alpha = 1/\Sigma = 1.441 \longrightarrow$

$b_1, \neg e$

0.0756	0.1614
0.03098	0.7320

$\longrightarrow P(b_1 | b_0, a = down) = 1/\alpha$
$= \boxed{0.694}$

$R(b_1, \neg e) = (0.0756)(-1) + (0.1614)(-4) +$
$(0.03098)(2) + (0.7320)(-2)$
$= \boxed{-1.5422}$

* doesn't add up to 1...

* SECOND ACTION :

left, b_1, e

0.23	0
0.69	0.08

a = left

e →

0.335	0
0.999	0.0039

b_2, e

¬e →

0.842	0.0213
0.120	0.0170

b_2, ¬e

* let s' = (0,0) :

(0,0)	left	0.7
× (0,1)	left	0.7
(1,0)	up	0.15
(0,0)	up	0.15

$b'(0,0) = (0.23)(0.7) + (0.23)(0.15) + (0.69)(0.7) = \boxed{0.6785}$

* let s' = (1,0) :

(1,0)	left	0.7
(1,0)	down	0.15
(0,0)	down	0.15
(1,1)	left	0.7

$b'(1,0) = (0.23)(0.15) + (0.69)(0.7+0.15) + (0.08)(0.7) = \boxed{0.677}$

* let s' = (0,1) :

× (0,1)	up	0.15
(1,1)	up	0.15

$b'(0,1) = (0.08)(0.15) = \boxed{0.012}$

* let s' = (1,1) :

(1,1)	down	0.15
× (0,1)	down	0.15

$b'(1,1) = (0.08)(0.15) = \boxed{0.012}$

* apply P(e|s) :

0.6785	0.012
0.677	0.012

⟹

0.204	0
0.609	0.0024

* apply α :

$\alpha = \frac{1}{\Sigma} = 1.64$ ⟹

b_2, e

0.335	0
0.999	0.0039

$R(b_2, e) = \sum_s b(s) R(s) = \boxed{1.655}$

$P(\ ) = 1/\alpha = \boxed{0.610}$

* apply P(¬e|s) :

0.6785	0.012
0.677	0.012

⟹

×(0.7)	×1
×(0.1)	×(0.8)

=

0.475	0.012
0.0677	0.0096

* apply $\alpha = 1.77$

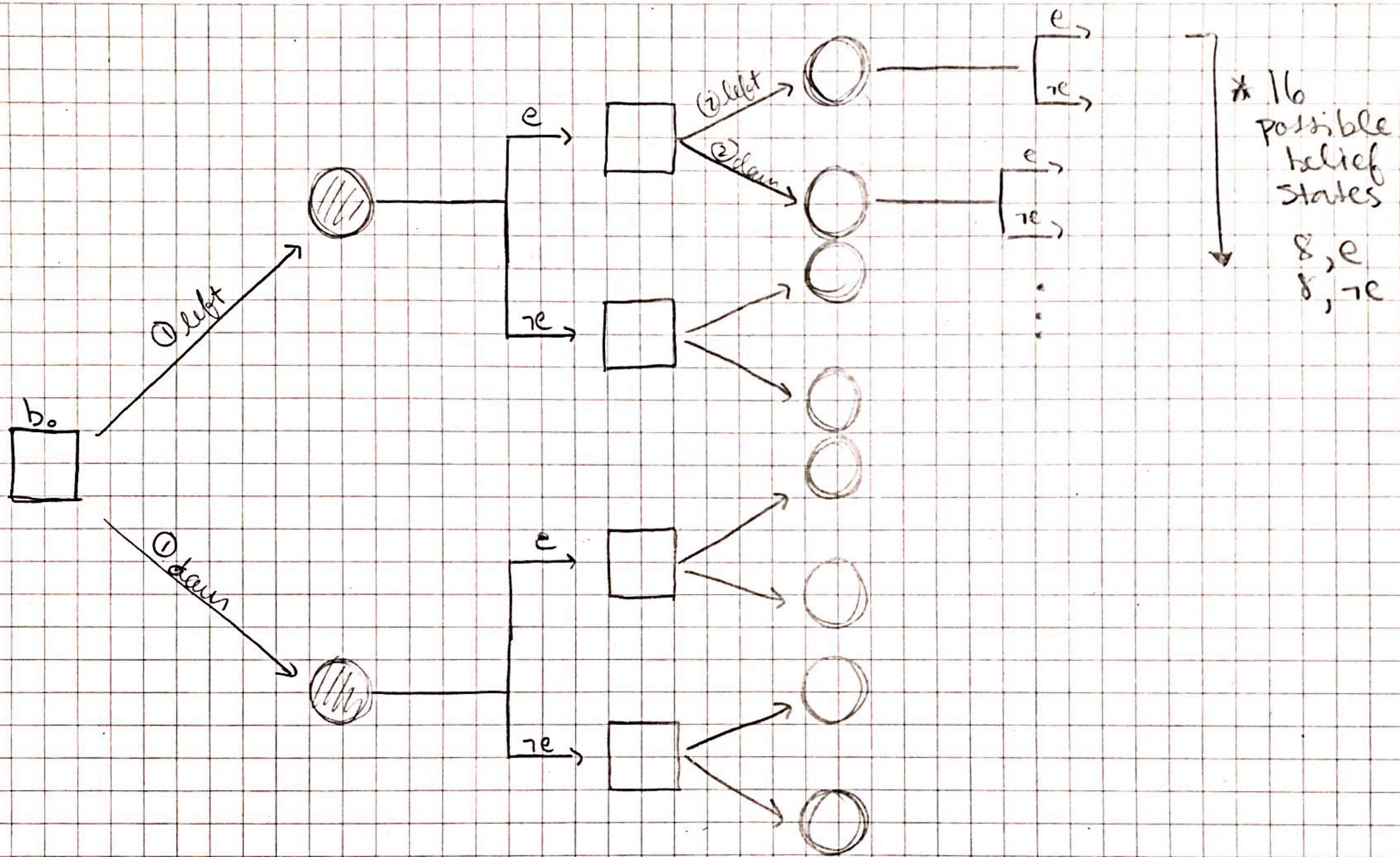
0.842	0.0213
0.120	0.0170

b_2, ¬e

$R(b_2, \neg e) = \boxed{-0.7212}$

$P(\ ) = 1/\alpha = \boxed{0.564}$

* my probabilities do not add up to 1 even though I think they should be


b₀
① left
① down
e
¬e
e
¬e
② left
② down
e
¬e
e
¬e
* 16 possible belief states
8, e
8, ¬e