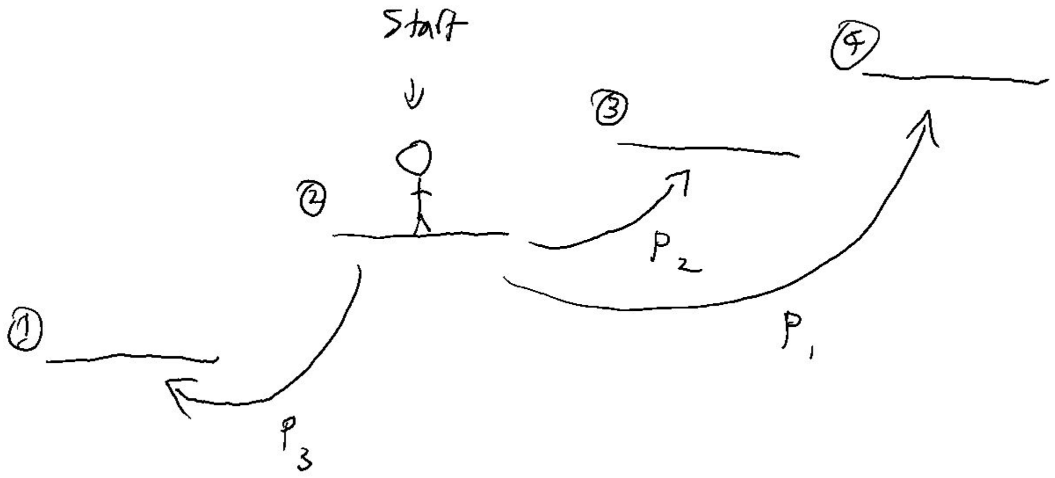
Problem 1

Context (1a)

New workout!

- Start at floor **2**
  - p_1 : +2
  - p_2 : +1
  - p_3 : -1
- Do this **twice**
  - No resetting to floor 2



Questions

- After **just 1 iteration**, what is your expected change in floor #? (E[Y])
- **After 2 iterations, starting on floor 2**, what is the expected floor you end on? (E[X])

<u>Bonus:</u> Does E[X] change if you pick a p_1, p_2, p_3, and do the exercise twice without re-selecting?

Y = Change in floor, ignore start

$$E[Y] = \overset{+2}{2p_1} + \overset{+1}{p_2} - \overset{-1}{p_3}$$

X = Absolute ending floor, Start at 2

↳ 1 iteration Change = $E[Y]$

$$\therefore E[X] = 2 + E[Y] + E[Y]$$

$$= 2 + 4p_1 + 2p_2 - 2p_3$$

No Change, because:

$$\left.\begin{array}{l} p_1 \to +4 \\ p_2 \to +2 \\ p_3 \to -2 \end{array}\right] \to = 2 + 4p_1 + 2p_2 - 2p_3 = E[X]$$

Context (1b)

Fact 1. *Let $X_{(j)}$ denote the jth order statistic in a sample of i.i.d. random variables; that is, the jth element when the items are sorted in increasing order $X_{(1)} \leq X_{(2)} \leq \ldots \leq X_{(n)}$.*

The PDF of $X_{(j)}$ is given by:

$$f_{X_{(j)}}(x) = \frac{n!}{(n-j)!(j-1)!}[F(x)]^{j-1}[1-F(x)]^{n-j} f(x). \tag{1}$$

Questions

(b) When a sample of $2N+1$ i.i.d. random variables is observed, the $(N+1)^{\text{st}}$ smallest is called the sample median. If a sample of size 3 from a uniform distribution over $[0,1]$ is observed, find the probability that the sample median is between $\frac{1}{4}$ and $\frac{3}{4}$. *Hint: use Fact* $\boxed{1}$.

<u>More confusing than it needs to be…</u>

First sentence: Defining what a median is. You already know this, don't let it confuse you

Second sentence: The actual question

Find the following, then plug & chug:

- *n* (not *N*)
- *j*
- *f(x)* (PDF)
- *F(x)* (CDF)

Hint: Both the PDF and CDF are piecewise

$\underline{n = 3}$: given

Sample median would be the "middle" sample when sorted

↳ $2N+1 = 3 \quad \therefore N = 1$

$N+1 = 2 \quad \therefore \underline{j = 2}$

$$P\left(\frac{1}{4} \le X_{(2)} \le \frac{3}{4}\right) = \int_{\frac{1}{4}}^{\frac{3}{4}} f_{X_{(2)}}(x)\, dx$$

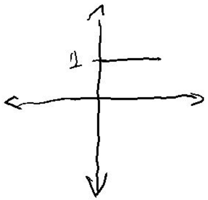
$$= \int_{\frac{1}{4}}^{\frac{3}{4}} \frac{3!}{(3-2)!\,(2-1)!} (x)^{2-1} \cdot (1-x)^{3-2} \cdot \underbrace{1}_{f(x)} \, dx$$

$$= 6 \int_{\frac{1}{4}}^{\frac{3}{4}} x(1-x)\, dx$$

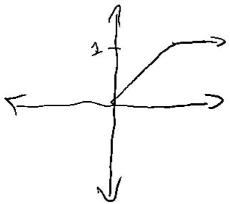
$$\downarrow$$

$$\frac{11}{16}$$

$$f(x) = \begin{cases} 1 & \text{if } 0 \le x \le 1 \\ 0 & \text{otherwise} \end{cases}$$



$$F(x) = \begin{cases} 0 & \text{if } x < 0 \\ x & \text{if } 0 \le x \le 1 \\ 1 & \text{if } x \ge 1 \end{cases}$$



(comes from Unif. dist PDF\CDF)

Problem 2

Quick Matrix Algebra/Calculus Refresher

Unsure if this has been taught before to you all (when I took 208, I definitely did not learn it)

Rule	Comments
$(\mathbf{AB})^T = \mathbf{B}^T\mathbf{A}^T$	order is reversed, everything is transposed
$(\mathbf{a}^T\mathbf{Bc})^T = \mathbf{c}^T\mathbf{B}^T\mathbf{a}$	as above
$\mathbf{a}^T\mathbf{b} = \mathbf{b}^T\mathbf{a}$	(the result is a scalar, and the transpose of a scalar is itself)
$(\mathbf{A}+\mathbf{B})\mathbf{C} = \mathbf{AC}+\mathbf{BC}$	multiplication is distributive
$(\mathbf{a}+\mathbf{b})^T\mathbf{C} = \mathbf{a}^T\mathbf{C}+\mathbf{b}^T\mathbf{C}$	as above, with vectors
$\mathbf{AB} \neq \mathbf{BA}$	multiplication is **not** commutative

Note: For the matrix calculus section to the right, **B** is a constant matrix.

Scalar derivative			Vector derivative		
$f(x)$	$\rightarrow$	$\frac{\mathrm{d}f}{\mathrm{d}x}$	$f(\mathbf{x})$	$\rightarrow$	$\frac{\mathrm{d}f}{\mathrm{d}\mathbf{x}}$
bx	$\rightarrow$	b	$\mathbf{x}^T\mathbf{B}$	$\rightarrow$	$\mathbf{B}$
bx	$\rightarrow$	b	$\mathbf{x}^T\mathbf{b}$	$\rightarrow$	$\mathbf{b}$
x^2	$\rightarrow$	$2x$	$\mathbf{x}^T\mathbf{x}$	$\rightarrow$	$2\mathbf{x}$
bx^2	$\rightarrow$	$2bx$	$\mathbf{x}^T\mathbf{Bx}$	$\rightarrow$	$2\mathbf{Bx}$

Thank you: Kirsty McNaught

Context & Questions (2a-b)

You are given some matrices, their shapes, and are asked to manipulate them!

Hint 1: The "2" superscript means square the whole norm. This can be used on the norm equivalencies at the top of the section handout

$\|Xw - Y\|_2^2$

$X_{m \times n}$; $w_{n \times 1}$; $Y_{m \times 1}$

Show that $\|Xw - Y\|_2^2 = \sum_{i=1}^{m} (x_i^T w - y_i)^2$

what is $\nabla_w \|Xw - Y\|_2^2$?

Hint 2: Verify your matrix calculus is correct by ensuring the shapes of the matrices allow for valid matrix multiplication.

$\|Xw - Y\|_2^2$; $Xw - Y$ has Shape $m \times 1$ $\begin{array}{cc} & \\ m & \begin{bmatrix} 0 \\ \\ \\ \end{bmatrix} \\ & 1 \end{array}$ A single element is: $(X_i^T w - Y_i)$

Think of $[Xw - Y]$ as a vector $(m \times 1)$, call it "v"

$$\|v\|_2^2 = \sqrt{v^T v}^{\,2} = \sqrt{\sum_{i=1}^{m} (v_i)^2}^{\,2}$$

$$\downarrow$$

$$\sum_{i=1}^{m} (v_i)^2$$

substitute v_i with $X_i^T w - Y_i$

$$\|Xw - Y\|_2^2 = \sum_{i=1}^{m} \left[(X_i^T w - Y_i)^2 \right]$$

$\nabla_w \|Xw - Y\|_2^2$: $(Xw - Y)$ is a vector, let's call it V

$=$

$\nabla_w \|V\|_2^2 = \nabla_w V^T V$

Take the gradient, apply chain rule $(\nabla_x X^T X = 2X)$

$= \nabla_w V \cdot 2V$; Now lets bring back in $(Xw - Y)$

$= \underbrace{\nabla_w (Xw - Y)} \cdot 2(Xw - Y)$

cancelled!

$\nabla_w Xw - \cancel{\nabla_w Y}$

$\downarrow$

① X^T

① $\nabla_w (Xw)$

$=$

$\nabla_w (w^T X^T)$

$=$

X^T

Rule:

$\nabla_z (z^T B) = B$

$\nabla_w \|Xw - Y\|_2^2 = 2X^T(Xw - Y)$