P1

(1) $x = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \in \mathbb{R}^n$

def^n

L_∞-norm $\|\vec{x}\|_\infty = \max\{|x_1|, |x_2|, \dots |x_n|\}$
$= \max_{1 \le i \le n}\{|x_i|\}$

L^1 norm $\|\vec{x}\|_1 = |x_1| + |x_2| + \dots + |x_n| = \sum_{i=1}^n |x_i|$

L_2 norm $\|\vec{x}\|_2 = \sqrt{|x_1|^2 + |x_2|^2 + \dots |x_n|^2}$
$= \sqrt{\sum_{i=1}^n |x_i|^2}$ [from notes]

o To show, $\|x\|_\infty \geq 0$ for all $x \in \mathbb{R}^n$
& $\|x\|_\infty = 0$ iff $x = 0$

by definition, $\|\vec{x}\|_\infty = \max_{1 \le i \le n} |x_i|$, so for each i, $|x_i| \geq 0$ since absolute value so not negative & max of set of non negative real numbers is also non negative $\therefore \|x\|_\infty \geq 0$

To show, $\|x\|_\infty = 0$ iff $x = 0$, then $x_i = 0$ for all i and so $|x_i| = 0$ and so $\max_{1 \le i \le n} |x_i| = 0$ i.e. $\|x\|_\infty = 0$ iff $x = 0$.

o next, $\|\alpha x\|_\infty = |\alpha| \|x\|_\infty$ for all $x \in \mathbb{R}^n$ & $\alpha \in \mathbb{R}$

since $\alpha \in \mathbb{R}$, $\|\alpha x\|_\infty = \max_{1 \le i \le n} |\alpha x_i| = \max_{1 \le i \le n} (|\alpha| |x_i|)$

$= |\alpha| \max_{1 \le i \le n} |x_i|$

$= |\alpha| \|x\|_\infty$

o Next $\|x_1 + x_2\|_\infty \leq \|x_1\|_\infty + \|x_2\|_\infty$ for all $x_1, x_2 \in \mathbb{R}^n$

let $x_1 = (x_{1,1}, x_{1,2} \dots x_{1,n})$ & $x_2 = (x_{2,1}, x_{2,2}, \dots x_{2,n})$

$\therefore x_1 + x_2 = (x_{1,1} + x_{2,1}+, \dots, x_{1,n} + x_{2,n})$

Since $\|x_1 + x_2\|_\infty = \max_{1 \le i \le n} |x_{1,i} + x_{2,i}|$ & $|x_{1,i} + x_{2,i}| \leq |x_{1,i}| + |x_{2,i}|$ ---- (1)

this comes from triangular inequality $\forall \mathbb{R}$

$\therefore |x_{1,i}| \leq \max_{1 \le j \le n} |x_{1,j}| = \|x_1\|_\infty$

$|x_{2,i}| \leq \max_{1 \le k \le n} |x_{2,k}| = \|x_2\|_\infty$

$\therefore |x_{1,i}| + |x_{2,i}| \leq \|x_1\|_\infty + \|x_2\|_\infty$

$\therefore |x_{1,i} + x_{2,i}| \leq |x_{1,i}| + |x_{2,i}|$
$\leq \|x_1\|_\infty + \|x_2\|_\infty$ ---- (2)

$\therefore \|x_1 + x_2\|_\infty = \max_{1 \le i \le n} |x_{1,i} + x_{2,i}|$.
$\leq \|x_1\|_\infty + \|x_2\|_\infty$ (from (2)) $\forall x_1, x_2 \in \mathbb{R}^n$

$\therefore \|x_1 + x_2\|_\infty \leq \|x_1\|_\infty + \|x_2\|_\infty$

Q.E.D.

P1 (2) Prove $\|x\|_\infty \leq \|x\|_2 \leq \|x\|_1$

$\|x\|_2 = \sqrt{\sum_{i=1}^{n} |x_i|^2}$, $\|x\|_2^2 = \sum_{i=1}^{n} |x_i|^2$ — (1)

$\|x\|_1 = \sum_{i=1}^{n} |x_i|$, $\|x\|_1^2 = \left(\sum_{i=1}^{n} |x_i|\right)^2$

$$= \underbrace{\sum_{i=1}^{n} |x_i|^2}_{>0} + 2 \underbrace{\sum_{1\leq i<j\leq n} |x_i||x_j|}_{>0 \text{ or non negative.}}$$

[e.g $(a_i+a_j)^2 = a_i^2 + a_j^2 + 2a_ia_j$]

$\therefore \|x\|_1^2 \geq \sum_{i=1}^{n} |x_i|^2 = \|x\|_2^2$ from (1)

$\therefore \|x\|_1^2 \geq \|x\|_2^2$

$\therefore \|x\|_2 \leq \|x\|_1$ — (2) since both sides have non negative values $\forall \mathbb{R}^n$

now $\|\vec{x}\|_\infty = \max_{1\leq i\leq n} |x_i| = |x_k|$ say

$\therefore \|x\|_2^2 = \left(\sqrt{\sum_{i=1}^{n} |x_i|^2}\right)^2 = \sum_{i=1}^{n} x_i^2 \geq x_k^2$ at least $= |x_k|^2 = \|x\|_\infty^2$

$\therefore \|x\|_2 = \sqrt{\sum_{i=1}^{n} |x_i|^2} \geq \sqrt{x_k^2} = |x_k| = \|x\|_\infty$

$\therefore \|x\|_\infty \leq \|x\|_2$ — (3) $\forall \; > 0$

Combine 2 & 3 $\|x\|_\infty \leq \|x\|_2 \leq \|x\|_1$ Q.E.D

P1 (3) To prove

$\|x\|_1 \leq \sqrt{n} \|x\|_2$

or, $\|x\|_1 = \sum_{i=1}^{n} |x_i| \leq \sqrt{n} \sqrt{\sum_{i=1}^{n} x_i^2} = \sqrt{n}\|x\|_2$ by def^n.

$\|x\|_1 = \sum_{i=1}^{n} |x_i| . \vec{1}$ → vector product of x_i's & 1's all $\in \mathbb{R}^n$

by Cauchy-Schwarz $u, v \in V$, then

$|\langle u . v\rangle|^2 \leq \langle u, u\rangle . \langle v, v\rangle$ i.e. $\leq \|u\|_2 . \|v\|_2$ ← formula (Source: berkeley-math notes online)

$\therefore \left(\sum_{i=1}^{n} |x_i| . 1\right)^2 \leq \left(\sum_{i=1}^{n} |x_i|^2\right)\left(\sum_{i=1}^{n} 1^2\right)$ | N.B $\sum_{i=1}^{n} 1^2 = n$

or $\|x\|_1^2 . \leq \|x\|_2^2 . n$

$\therefore$ Since $\|x\|_1$ & $\|x\|_2$ are not negative here,

taking sq root yields $\|x\|_1 \leq \sqrt{n}\,\|x\|_2$ Q.E.D

from P1 2 $\quad \|x\|_\infty \leq \|x\|_2 \leq \|x\|_1$

$$\therefore \|x\|_2 \leq \|x\|_1 \leq \sqrt{n}\,\|x\|_2$$

& $$\frac{1}{\sqrt{n}}\|x\|_1 \leq \|x\|_2 \leq \|x\|_1$$

P1 4 Hölder's inequality $x_1, x_2 \in \mathbb{R}^n$

to prove, $|\langle x_1, x_2 \rangle| \leq \|x_1\|_\infty \|x_2\|_1$

$$|\langle x_1, x_2\rangle| = \left|\sum_{i=1}^{n} x_{1,i}\, x_{2,i}\right|$$ (as defined in P1 1) see x_1 & x_2

$$\leq \sum_{i=1}^{n} |x_{1,i}\, x_{2,i}| = \sum_{i=1}^{n} |x_{1,i}|\,|x_{2,i}|$$

$$\therefore \sum_{i=1}^{n} |x_{1,i}|\,|x_{2,i}| \leq \sum_{i=1}^{n} \|x_1\|_\infty \cancel{\sum} \|x_{2i}\|_1 \cancel{= \|x_1\|_\infty \|x_2\|_1}$$

$$\therefore = \|x_1\|_\infty \sum_{i=1}^{n} \|x_{2,i}\|$$

$$= \|x_1\|_\infty \|x_2\|_1$$

$$\therefore |\langle x_1, x_2\rangle| \leq \|x_1\|_\infty \|x_2\|_1$$

P2 $\mathbb{E}_X[X] = \mathbb{E}_Y[\mathbb{E}_X[X|Y]]$, given from tower rule. — ①

① To prove, the tower rule for two discrete random variables X & Y with values in set $\mathcal{X}$ & $\mathcal{Y}$

$$\mathbb{E}_X[X] = \sum_{x\in\mathcal{X}} xP[X=x] \quad — ②$$ → expectation of X

$$\mathbb{E}_X[X|Y=y] = \sum_{x\in\mathcal{X}} x\,P(X=x|Y=y) \quad — ③$$ → Conditional expectation of X given $Y=y$

$$\therefore \mathbb{E}_Y[\mathbb{E}[X|Y]] = \sum_{y\in\mathcal{Y}} P[Y=y]\,\mathbb{E}[X|Y=y] \quad — ④$$ → expectation of conditional expectation

R.h.s of ①

$$\mathbb{E}_Y[\mathbb{E}[X|Y]] = \sum_{y\in\mathcal{Y}} P[Y=y]\,\mathbb{E}[X|Y=y]$$

substitute expression from ③,

$$\mathbb{E}_Y[\mathbb{E}[X|Y]] = \sum_{y\in\mathcal{Y}} P[Y=y]\left(\sum_{x\in\mathcal{X}} x\,P(X=x|Y=y)\right)$$

$$= \sum_{y\in\mathcal{Y}}\sum_{x\in\mathcal{X}} x\,P[Y=y]\,P[X=x|Y=y] \quad — ⑤$$

$$P[X=x|Y=y] = \frac{P[X=x,Y=y]}{P[Y=y]} \to >0 \qquad \therefore P[Y=y]\,.\,P[X=x|Y=y] = P[X=x,Y=y] \quad — ⑥$$

⑥ in ⑤,

$$\mathbb{E}_Y[\mathbb{E}[X|Y]] = \sum_{y\in\mathcal{Y}}\sum_{x\in\mathcal{X}} xP[X=x,Y=y]$$

$$= \sum_{x\in\mathcal{X}}\sum_{y\in\mathcal{Y}} xP[X=x,Y=y]$$

$$= \sum_{x\in\mathcal{X}} x\left(\sum_{y\in\mathcal{Y}} P[X=x,Y=y]\right)$$

N.B. $\to P[X=x]$

$$= \sum_{x\in\mathcal{X}} x\,P[X=x] = \mathbb{E}_X[X]$$

$$\therefore \mathbb{E}_X[X] = \mathbb{E}_Y[\mathbb{E}[X|Y]] \quad \text{Q.E.D}$$

P2 (2) Now do same as before prove tower rule for two continuous real valued random variables X & Y

joint pdf $f_{X|Y}(x|y) = \dfrac{f_{X,Y}(x,y)}{f_Y(y) \to > 0}$ — (1)

individual

$f_X(x) = \int_{-\infty}^{\infty} f_{X,Y}(x,y)\,dy$ — (2)

$f_Y(y) = \int_{-\infty}^{\infty} f_{X,Y}(x,y)\,dx$ — (3)

$\mathbb{E}[X|Y=y] = \int_{-\infty}^{\infty} x f_{X|Y}(x|y)\,dx$ — (5)

$\mathbb{E}_Y[\mathbb{E}[X|Y]] = \int_{-\infty}^{\infty} \mathbb{E}[X|Y=y]\, f_Y(y)\,dy$ — (4)

$\mathbb{E}_Y[\mathbb{E}_X[X|Y]] = \cancel{\mathbb{E}_X[X]} = \cancel{\int_{-\infty}^{\infty} x f_X(x)\,dx}$ ~~from (2)~~

$= \cancel{\int_{-\infty}^{\infty} x\left(\int_{-\infty}^{\infty} f_{X,Y}(x,y)\,dy\right)dx}$

$= \cancel{\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x f_{X,Y}(x,y)\,dy\,dx}$

from (4) & (5)

$= \int_{-\infty}^{\infty}\left(\int_{-\infty}^{\infty} x f_{X|Y}(x|y)\,dx\right) f_Y(y)\,dy$

from (1),

$$= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x f_{X,Y}(x,y)\,dx\,dy \quad \text{— (6)}$$

now, $\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x f_{X,Y}(x,y)\,dx\,dy = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x f_{X,Y}(x,y)\,dy\,dx$

order swapped as we expect values to be finite.

$\therefore$ (6) becomes,

$$\int_{-\infty}^{\infty} x\left(\int_{-\infty}^{\infty} f_{X,Y}(x,y)\,dy\right)dx = \int_{-\infty}^{\infty} x f_X(x)\,dx = \mathbb{E}[X]$$

$$\therefore\ E_Y[\mathbb{E}[X|Y]] = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x f_X(x,y)\,dx\,dy = \mathbb{E}[X]$$

Q.E.D.

P2

3 Extend tower rule $\mathbb{E}_X[X|Z] = \mathbb{E}_Y[\mathbb{E}_X[X|Y,Z][Z]] \longrightarrow 1$

L.H.S

$E_X[X|Z=z] = \sum_{x\in\mathcal{X}} x\, P(X=x \mid Z=z) \longrightarrow 1$

$\mathbb{E}_X[X|Y=y, Z=z]$

$= \sum_{x\in\mathcal{X}} x \cdot P(X=x \mid Y=y, Z=z) \longrightarrow 2$

$P(X = x \mid Z=z)\ Z=z]$

$= \dfrac{P[X=x, Y=y, Z=z]}{P[Y=y, Z=z]} \longrightarrow 3$ $\quad P[Y=y,Z=z] \to >0$

& $P(Y=y \mid Z=z) = \dfrac{P[Y=y, Z=z]}{P[Z=z]} \longrightarrow 4$

given
$\mathbb{P}[X=x]\ x\in\mathcal{X}$
$\mathbb{P}[Y=y]\ y\in\mathcal{Y}$
$\mathbb{P}[Z=z]\ \cancel{z\in z}$
$z\in\mathcal{Z}$

R.H.S of 1

$\mathbb{E}_Y[\mathbb{E}_X[X|Y,Z] \mid Z=z]$

$= \sum_{y\in\mathcal{Y}} E[X|Y=y, Z=z]\, P(Y=y \mid Z=z)$

$= \sum_{y\in\mathcal{Y}} \left(\sum_{x\in\mathcal{X}} x \cdot P(X=x \mid Y=y, Z=z)\right) \cdot P(Y=y \mid Z=z)$

from 2 & 4 substⁿ

$= \sum_{y\in\mathcal{Y}} \left(\sum_{x\in\mathcal{X}} x \cdot \dfrac{P[X=x, Y=y, Z=z]}{P[\cancel{Y=y}, Z=z]} \cdot \dfrac{P[\cancel{Y=y}, Z=z]}{P[Z=z]}\right)$

$= \sum_{y\in\mathcal{Y}} \sum_{x\in\mathcal{X}} x \cdot \dfrac{P[X=x, Y=y, Z=z]}{P(Z=z)}$ $\quad P(Z=z) \to >0$

$= \dfrac{1}{P(Z=z)} \sum_{x\in\mathcal{X}} x \left(\sum_{y\in\mathcal{Y}} P[X=x, Y=y, Z=z]\right)$

$= \dfrac{1}{P(Z=z)} \sum_{x\in\mathcal{X}} x \cdot P[X=x, Z=z]$

$= \sum_{x\in\mathcal{X}} x \dfrac{P[X=x, Z=z]}{P(Z=x)} = \sum_{x\in\mathcal{X}} x\, P[X=x \mid Z=z]$

$= \mathbb{E}_X[X \mid Z=z] = \mathbb{E}_X[X|Z]$

from 1

$\therefore\ \mathbb{E}_X[X|Z] = \mathbb{E}_Y[\mathbb{E}_X[X|Y,Z] \mid Z]$

QED

P2 (4)

Starts at 0

Each step $x_t = 1 \to p$, $x_t = -1 \to 1-p$ $\}$ $p \in (0,1)$ $\quad f_X = 2^{\sqrt{2x}}$

Stops after H steps $H \in \{1, 2, 3, \dots\}$

H is geometric distributed random variable. $q \in (0, 1)$

so $p(H = h) = (1-q)^{h-1} \cdot q$ → this by def^n.

given $L = \sum_{t=1}^{H} x_t = x_1 + x_2 + \dots + x_H$

$\mathbb{E}[x_t] = (+1) \cdot p + (-1) \cdot (1-p) = p - 1 + p = 2p - 1$

Tower rule $\mathbb{E}_X[X] = \mathbb{E}_Y[\mathbb{E}_X[X|Y]]$

so $\mathbb{E}_L[L] = \mathbb{E}_H[\mathbb{E}_L[L|H]]$

now $\mathbb{E}_L[L|H=h] = \mathbb{E}_H\left[\sum_{t=1}^{h} x_t \,\middle|\, H=h\right]$ $\quad$ $L = \sum_{t=1}^{h} x_t$ battery finishes at h steps

$\mathbb{E}_L[L|H=h] = \sum_{t=1}^{h} \mathbb{E}_H[x_t | H=h]$

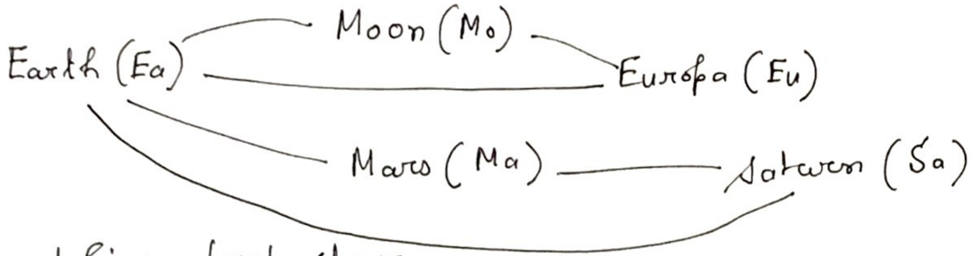
$= \sum_{t=1}^{h} (2p-1) = h \cdot \mu$

$\mathbb{E}_L[L] = \mathbb{E}_L(\mathbb{E}_H[L|H]) = \mathbb{E}_L(H\ (2p-1)) = (2p-1)\,\mathbb{E}_L[H]$

$= \frac{2p-1}{q}$ Ans?

P3

1



Ask is to define prob space defined by $\Omega_v \rightarrow$ sample space, $F_v \rightarrow$ event space & prob measure $\rightarrow P_v$

here sample space $\Omega_v \rightarrow$ all possible voyages based on allowed jumps are finite.

v_1 : $(Ea \rightarrow Mo \rightarrow Ea)$

v_2 : $(Ea \rightarrow Ma \rightarrow Ea)$

v_3 : $(Ea \rightarrow Eu \rightarrow Ea)$

v_4 : $(Ea \rightarrow Sa \rightarrow Ea)$

v_5 : $(Ea \rightarrow Mo \rightarrow Eu \rightarrow Ea)$

v_6 : $(Ea \rightarrow Ma \rightarrow Sa \rightarrow Ea)$

Sigma-algebra $F_v = 2^{\Omega_v}$ for finit Ω_v. It is to measure events related to voyage sequence. It is sigma algebra over voyage space. all possible collection of voyages.

Probability measurement P_v

&

$P_v(v_1) = P_v[Ea \rightarrow Mo \rightarrow Ea] = 0.1 \times 0.6 = 0.06$

$P_v(v_2) = P_v[Ea \rightarrow Ma \rightarrow Ea] = 0.4 \times 0.3 = 0.12$

$P_v(v_3) = P_v[Ea \rightarrow Eu \rightarrow Ea] = 0.25 \times 1.00 = 0.25$ ✓

$P_v(v_4) = P_v[Ea \rightarrow Sa \rightarrow Ea] = 0.25 \times 1.00 = 0.25$

$P_v(v_5) = P_v[Ea \rightarrow Mo \rightarrow Eu \rightarrow Ea] = 0.1 \times 0.4 \times 0.1 = 0.04$ ✓

$P_v(v_6) = P_v[Ea \rightarrow Ma \rightarrow Sa \rightarrow Ea] = 0.4 \times 0.7 \times 1.0 = 0.28$

sum $\rightarrow (0.06 + 0.12 + 0.25 + 0.25 + 0.04 + 0.28) = 1.0$

(3) $R = R(v)$ return voyage v, same way to show prob space $(\Omega_R, F_R, \mathbb{P}_R)$

first sample space $\Omega_R \to$ all possible values of return R.

Sample values: (S)
$Ea=(0)$, $Mo=(5)$, $Ma=(3)$
$Eu=(8)$, $Sa=(9)$
energy cost: (C)
$(Ea \leftrightarrow Mo) = 2$
$(Mo \leftrightarrow Eu) = 15$
$(Ma \leftrightarrow Sa) = 20$
$(Ea \leftrightarrow Eu) = 12$
$(Ea \leftrightarrow Ma) = 10$
$(Ea \leftrightarrow Sa) = 25$

Return of all possible voyeges

Returns $R(v) = \sum$ sample values (S) $- \sum$ cost (C)

✓ $R(Ea \to Mo \to Eo)$
$= (0+5+0) - (2+2) = 1$

✓ $R(Ea \to Ma \to Ea)$
$= (0+3+0) - (10+10) = 3 - 20 = -17$

- $R(Ea \to Eu \to Ea) = (0+8+0) - (12+12)$
$= 8 - 24 = -16$ ✓ ?

- $R(Ea \to Sa \to Ea) = (0+9+0) - (25+25)$
$= 9 - 50 = -41$

- $R(Ea \to Mo \to Eu \to Ea) = (0+5+8+0) - (2+15+12)$
$= 13 - 29$
$= -16$ ✓

- $R(Ea \to Ma \to Sa \to Ea) = (0+3+9+0) - (10+20+25)$
$= (12) - 55 = -43$.

$\therefore$ Return values all possible $\Omega_R = \{-43, -16, -41, \cancel{-16}, -17, 1\}$

Event space F_R is power set of the sample space Ω_R
$F_R = 2^{\Omega_R}$ ie power set of all possible returns.

Next prob measure $\mathbb{P}_R$
$P_R(1) = 0.06$ $\{Ea \to Mo \to Ea\}$
$P_R(-16) \to$ two voyages v_3 & v_5 (same return yield)
$\therefore P_R(-16) = P_v(v_3) + P_v(v_5)$
$= P_v(Ea \to Eu \to Ea) + P_v(Ea \to Mo \to Eu \to Ea)$
$= 0.25 + 0.04 = 0.29$ (Combined)

$P_R(-41) = P_V(V_4) = P_V(Ea \to Sa \to Ea)$

$= 0.25$

$P_R(-17) = P_V(V_2) = 0.12 \; P_V[Ea \to Ma \to Ea]$

$P_R(-43) : P_V(V_6) = P_V[Ea \to Ma \to Sa \to Ea) = 0.28$

$\text{Sum } P_R = 0.06 + 0.29 + 0.25 + 0.12 + 0.28 = 1.00$

Prob space (Ω_R, F_R, P_R)

$= (\{-43, -16, -41, \cancel{-16}, -17, 1\}, 2^{\{-43,-16,-41,-17,1\}}, \mathbb{P}_R\}$

QED. (Ans)

P3
3. Calc true (theoretical) expectation (ave) of return $\mathbb{E}(R)$

$\mathbb{E}(R) = 1.(0.06) + (-16)(0.29) + (-17)(0.12) + (-41)(0.25)$
$+ (-43)(0.28)$ [these already calculated in ① & ②]

$= 0.06 - 4.64 - 2.04 - 10.25 - 12.04$

$= -28.91$

P3
④ (a) There are total 6 possible first-return voyages with the set $\{-43, -16, -41, -17, 1\}$

$\min R_i = -43$
$\max R_i = 1$ $\}$ lower & upper bounds (β) (α)

$\therefore \underset{\alpha}{-43} \leq R_i \leq \underset{\beta}{1}$ for all $i \in 1, 2 \ldots, n$ (Ans)

P3/4 (b) Given $S_n = \sum_{i=1}^{n} R_i$ sum of n returns.

$$S_n = R_1 + R_2 + \cdots + R_n$$

$$\therefore \mathbb{E}[S_n] = \mathbb{E}[R_1 + R_2 + \cdots + R_n]$$

$$= \mathbb{E}[R_1] + \mathbb{E}[R_2] + \cdots + \mathbb{E}[R_n]$$

Using i.i.d
- same distribution so identical
- voyages are independent
- so we can assume iid all have same theoretical average $\mathbb{E}[R]$

$$\therefore \mathbb{E}[S_n] = n.\mathbb{E}[R]$$ [Answer]

P3/4 (c) Hoeffding inequality (wikipedia?)

$\alpha = -43$
$\beta = 1$
$\therefore \beta - \alpha = +44$

$$\mathbb{P}(|S_n - \mathbb{E}[S_n]| \geq \epsilon) \leq 2\exp\left(-\frac{2\epsilon^2}{\sum_{i=1}^{n}(\beta-\alpha)^2}\right) \quad (1)$$

$$\leq 2\exp\left(-\frac{2\epsilon^2}{n(\beta-\alpha)^2}\right) \leq 2\exp\left(\frac{-2\epsilon^2}{n\,44^2}\right)$$

Answer.

P3/4 (d)

$$\bar{S}_n = \frac{1}{n} S_n = \frac{1}{n}\sum_{i=1}^{n} R_i$$

$$\mathbb{E}[\bar{S}_n] = \mathbb{E}\left[\frac{1}{n}\sum_{i=1}^{n} R_i\right] = \frac{1}{n}\sum_{i=1}^{n}\mathbb{E}[R_i]$$

$$= \frac{1}{n} \cdot n \cdot \mathbb{E}[R]$$ we proved earlier iid

$$= \mathbb{E}[R]$$ Answer

P3 (e) From 4(c) $\mathbb{P}(|S_n - \mathbb{E}[S_n]|) \leq 2\exp\left(-\frac{2\epsilon^2}{n(\beta-\alpha)^2}\right)$ 44

we known $\bar{S}_n = S_n$

To derive,

$$\mathbb{P}(\bar{S}_n - \mathbb{E}[\bar{S}_n]) = \mathbb{P}\left(\frac{S_n}{n} - \mathbb{E}\left[\frac{S_n}{n}\right]\right) = \mathbb{P}\left[\frac{1}{n}\left(S_n - \mathbb{E}[S_n]\right) \geq \epsilon'\right.$$

$$\therefore \mathbb{P}(|\bar{S}_n - \mathbb{E}[\bar{S}_n]|) = \mathbb{P}(|S_n - E[S_n]) \geq n\epsilon')$$

from 4(c) eqⁿ ①

$$\mathbb{P}\left(|\bar{S}_n - \mathbb{E}[\bar{S}_n]| \geq \epsilon'\right) = \mathbb{P}\left(|S_n - \mathbb{E}[S_n]| \geq n\epsilon'\right) \leq 2\exp\left(-\frac{2(n\epsilon')^2}{n(\beta-\alpha)^2}\right)$$

$[\beta - \alpha = 44]$

$$\therefore \mathbb{P}\left(|\bar{S}_n - \mathbb{E}[\bar{S}_n]| \geq \epsilon'\right) \leq 2\exp\left(-\frac{2n(\epsilon')^2}{44^2}\right) \quad — ②$$

[Answer I hope!]

4. (f)

Using eqⁿ ② from above, & $\epsilon' = 2$, $\delta = 1 - 0.9 = 0.1$

$$\mathbb{P}\left(|\bar{S}_n - \mathbb{E}[R]| \geq \epsilon'\right) \leq 2\exp\left(-\frac{2n(\epsilon')^2}{44^2}\right)$$

$\therefore \mathbb{P}\left(\bar{S}_n - \mathbb{E}[R] \geq 2\right) \leq 0.1$ also

$2\exp\left(-\frac{2n(2)^2}{44^2}\right) \leq 0.1$ or $2\exp\left(-\frac{8n}{1936}\right) \leq 0.1$

or $2\exp\left(-\frac{n}{242}\right) \leq 0.1$ $\quad \therefore \exp\left(-\frac{n}{242}\right) \leq 0.05$

$$\therefore -\frac{n}{242} \leq \ln(0.05)$$

$$\therefore \frac{n}{242} \geq -\ln(0.05) \qquad \therefore \frac{n}{242} \geq 2.995$$

$$\therefore n \geq 242 \times 2.995$$

$$n \geq 724.96 \sim 724.96$$
$$\sim 725$$

[Answer]

P4

$$\mathbb{P}[X_{t+1} | X_0, X_1, .., X_t] = \mathbb{P}[X_{t+1} | X_t] \quad \forall t \in \mathbb{N}_0$$

①

To prove

LHS

$$\mathbb{P}[X_{t+1}, X_{t+2}, \cdots X_{t+m} | X_0, X_1, \ldots X_t]$$

$$= \mathbb{P}[X_{t+1}, X_{t+2} \cdots X_{t+m} | X_t] \quad \forall t \in \mathbb{N}_0 \; \forall m \in \mathbb{N}$$

RHS.

Use of Chain rule here

$$\mathbb{P}(X_{t+1}, \ldots, X_{t+m} | X_0, \ldots, X_t)$$

$$= \mathbb{P}[X_{t+1} | X_0, \ldots, X_t] \times \mathbb{P}[X_{t+2} | X_{t+1}, X_0 \ldots X_t] \times \ldots \times \mathbb{P}[X_{t+m} | X_{t+m-1} \ldots X_t]$$

1st factor : $\mathbb{P}[X_{t+1} \mid X_0, \dots X_t] = \mathbb{P}[X_{t+1} \mid X_t]$

2nd factor $\mathbb{P}[X_{t+2} \mid X_{t+1}, X_0, \dots X_t] = \mathbb{P}[X_{t+2} \mid X_{t+1}]$

or $\mathbb{P}[X_{t+2} \mid X_0, \dots, X_{t+1}] = \mathbb{P}[X_{t+2} \mid X_{t+1}]$

based on Markov rule simplifies this for each factors

kth factor. ∴ $\mathbb{P}[X_{t+k} \mid X_{t+k-1}, \dots X_{t+1}, X_0, \dots, X_t]$

or $\mathbb{P}[X_{t+k} \mid X_0, \dots, X_{t+k-1}] = \mathbb{P}[X_{t+k} \mid X_{t+k-1}]$

$$\therefore \text{LHS} = \mathbb{P}[X_{t+1} \mid X_t] \times \mathbb{P}[X_{t+2} \mid X_{t+1}] \times \dots \times P[X_{t+m} \mid X_{t+m-1}] \longrightarrow (1)$$

RHS $\mathbb{P}[X_{t+1}, \dots X_{t+m} \mid X_t]$ chain rule

$$= \mathbb{P}[X_{t+1} \mid X_t] \times P[X_{t+2} \mid X_{t+1}, X_t] \times \dots \times [X_{t+m} \mid X_{t+m-1} \dots X_{t+1}, X_t]$$

Using same markov rule, we only care about X_{t+k-1} to eval state of X_{t+k}. $\mathbb{P}[X_{t+k} \mid X_{t+k}, \dots, X_t] = \mathbb{P}[X_{t+k} \mid X_{t+k-1}]$

RHS simplifies as $= \mathbb{P}[X_{t+1} \mid X_t] \times \mathbb{P}[X_{t+2} \mid X_{t+1}] \times \dots \times \mathbb{P}[X_{t+m} \mid X_{t+m-1}] \longrightarrow (2)$

we see from (1) & (2) LHS = RHS

$$\therefore \mathbb{P}[X_{t+1}, \dots, X_{t+m} \mid X_0, \dots, X_t] = \mathbb{P}[X_{t+1}, \dots, X_{t+m} \mid X_t]$$

Q.E.D.

P4 (2) To prove

$$\mathbb{P}[X_{t+d} \mid X_0, X_1, \dots, X_t] = \mathbb{P}[X_{t+d} \mid X_t]$$

also $\mathbb{P}[Y \mid Z] = \sum_{w \in W} \mathbb{P}[Y, W = w \mid Z] \longrightarrow$ Given in problem

$\mathbb{P}(X_{t+1} \mid X_0, \dots X_t) = \mathbb{P}(X_{t+1} \mid X_t)$ when $d = 1$

when $d \geq 1$, we prove holds true for $d+1$ iteratively,

$$\mathbb{P}(X_{t+d} \mid X_0, \dots, X_t) = \mathbb{P}(X_{t+d} \mid X_t)$$

when $W = X_{t+1}$ as missing variable from hint

so next marginalizing over it,

$$\mathbb{P}[X_{t+d+1} \mid X_0, X_1, \dots, X_t] = \sum_{\omega \in W} \mathbb{P}[X_{t+d+1}, W = \omega \mid X_0, X_1, \dots, X_t]$$

Using chain rule r.h.s

$$= \sum_{\omega \in W} \mathbb{P}[X_{t+d+1} \mid W = \omega, X_0, X_1, \dots, X_t] \, . \, \mathbb{P}[W = \omega \mid X_0, X_1, \dots, X_t] \quad \text{——} (1)$$

in eqn (1) r.h.s 2nd factor

$$\mathbb{P}[W = \omega \mid X_0, \dots X_t] = \mathbb{P}[X_{t+1} = \omega \mid X_t]$$ from markov

1st factor in eqn (1)

$$\mathbb{P}[X_{t+d+1} \mid W = \omega, X_0, \dots X_t] = \mathbb{P}[X_{t+d+1} \mid X_{t+1} = w]$$

$$\therefore \mathbb{P}[X_{t+d+1} \mid X_0, X_1, \dots, X_t] = \sum_{\omega \in W} \mathbb{P}[X_{t+d+1} \mid X_{t+1} = w] \, \mathbb{P}[X_{t+1} = W \mid X_t]$$

$$= \mathbb{P}[X_{t+d+1} \mid X_t]$$

$$\therefore \quad \mathbb{P}[X_{t+d+1} \mid X_0, \dots, X_t] = \mathbb{P}[X_{t+d+1} \mid X_t]$$

$d' = d+1$

$$\Rightarrow \mathbb{P}[X_{t+d'} \mid X_0, \dots, X_t] = \mathbb{P}[X_{t+d'} \mid X_t]$$

Generalizing $\mathbb{P}[X_{t+d} \mid X_0, \dots X_t] = \mathbb{P}[X_{t+d} \mid X_t]$

Q.E.D.

P4

(3) To prove,

L.H.S $\qquad$ R.H.S

$$\mathbb{P}[X_{t+k}, X_{t+k+1}, \dots, X_{t+m} \mid X_0, \dots X_t] = \mathbb{P}[X_{t+k}, X_{t+k+1}, \dots X_{t+m} \mid X_t]$$

$$\forall t \in \mathbb{N}_0, \forall k \in \mathbb{N}, m \geq k \quad \text{——} (A)$$

let $H_t = X_0, \dots, X_t$

$$\mathbb{P}[X_{t+k:t+m} \mid H_t] = \prod_{j=k}^{m} \mathbb{P}[X_{t+j} \mid H_t, X_{t+k}, \dots, X_{t+j-1}] \quad \text{——} (1)$$

expanding r.h.s $\longrightarrow$ or, $\mathbb{P}[X_{t+k}, \dots, X_{t+m} \mid H_t] = \mathbb{P}[X_{t+k} \mid H_t] . \mathbb{P}[X_{t+k+1} \mid H_t, X_{t+k}] \dots \mathbb{P}[X_{t+m} \mid H_t, X_{t+k} \dots, X_{t+m-1}$ —— (2)

Using Markov property on ① i.e $\mathbb{P}[X_{t+1} \mid X_0 \dots X_t] = \mathbb{P}[X_{t+1} \mid X_t]$

$$\therefore \mathbb{P}[X_{t+j} \mid H_t, X_{t+k}, \dots X_{t+j-1}] = \mathbb{P}[X_{t+j} \mid X_{t+j-1}]$$

$$\therefore \mathbb{P}[X_{t+k:t+m} \mid H_t] = \prod_{j=k}^{m} \mathbb{P}[X_{t+j} \mid X_{t+j-1}]$$

or eqn ① can be written as

$$\mathbb{P}[X_{t+k}, \dots, X_{t+m} \mid H_t] = \mathbb{P}[X_{t+k} \mid H_t] \prod_{j=k+1}^{m} \mathbb{P}[X_{t+j} \mid X_{t+j-1}, \dots X_{t+k}, H_t] \quad \text{— ③}$$

first factor in r.h.s eqn ③

$$\mathbb{P}[X_{t+k} \mid H_t] = \mathbb{P}[X_{t+k} \mid X_t]$$

next second factor for $j > k$, $\mathbb{P}[X_{t+j} \mid X_{t+j-1}]$ X_{t+j} will depend on prior $t+j-1$ th.

$\therefore$ LHS of eqn Ⓐ

i.e. $\overset{LHS}{\mathbb{P}[X_{t+k}, X_{t+k+1}, \dots X_{t+m} \mid X_0, \dots X_t]}$

$$= \mathbb{P}[X_{t+k} \mid X_t] \cdot \prod_{j=k+1}^{m} \mathbb{P}[X_{t+j} \mid X_{t+j-1}] \quad \text{— ④}$$

Next, R.HS.

$$\mathbb{P}[X_{t+k}, \dots, X_{t+m} \mid X_t] = \mathbb{P}[X_{t+k} \mid X_t] \cdot \prod_{j=k+1}^{m} \mathbb{P}[X_{t+j} \mid X_{t+j-1}, \dots, X_{t+k}, X_t]$$

$$\Downarrow$$

$$\mathbb{P}[X_{t+j} \mid X_{t+j-1}]$$

$$\therefore \text{R.HS} = \mathbb{P}[X_{t+k} \mid X_t] \cdot \prod_{j=k+1}^{m} \mathbb{P}[X_{t+j} \mid X_{t+j-1}] \quad \text{— ⑤}$$

$\therefore$ from ④ & ⑤ LHS = RHS. $\therefore$ QED.

—x—