IV Random Walks

- prob. model $(\Omega, \mathcal{F}, P)$

- suppose $S_0, X_1, X_2, \ldots$ are mut. stat. ind. r.v.'s and $X_1, X_2, \ldots$ are identically distributed

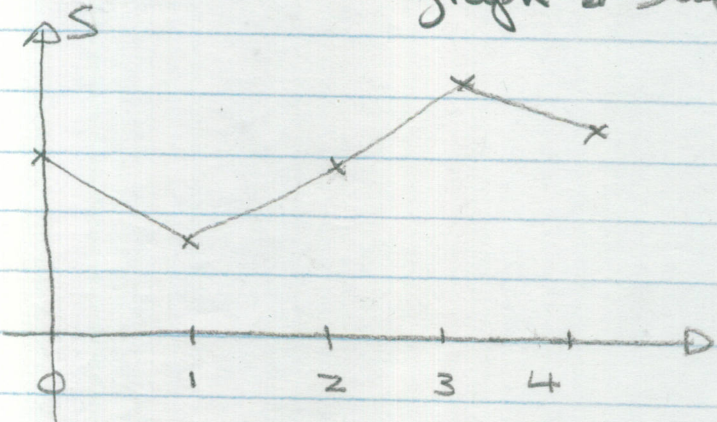
- then we have a s.p. with state space $\mathbb{R}$ and time domain $\mathbb{N}_0$ given by $(S_0, X_1, X_2, \ldots)$ via the product model

- now define a new s.p. with state space $\mathbb{R}$ and time domain $\mathbb{N}_0$ via

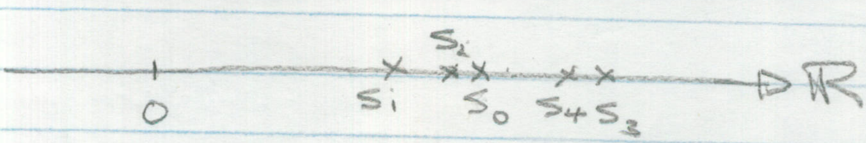
$$S_n = S_0 + \sum_{k=1}^{n} X_k = S_{n-1} + X_n$$

- we call call the s.p. $S = (S_0, S_1, \ldots)$ a <u>random walk</u>

graph of sample function.



note, we join the points by lines but remember time is discrete



Eg - motion of a gas mol

eg stock prices

X_n = price change of stock at time n
$= \prod_{i=1}^{n} (1+r_i) X_0$. $\log X_n = \log X_0 + \sum_{i=1}^{n} \log(1+r_i)$

eg kinetic motion in a gas

X_i = displacement of molecule (along x-axis) due to collision at time i

- a random walk is called a simple random walk (SRW) if $S_0 \in \mathbb{Z}$ is degenerate and $P(X_i = 1) = p$, $P(X_i = -1) = q = 1-p$

- if $p = q = \frac{1}{2}$ it is a symmetric SRW

- note $P(S_n = j \mid S_0 = a)$ refers to the probability SRW is in state j at time n given that it started in state a.

Prop ① For the SRW

(i) (spatial homogeneity) $P(S_n = j \mid S_0 = a) = P(S_n = j+b \mid S_0 = a+b)$

(ii) (temporal homogeneity) $P(S_n = j \mid S_0 = a) = P(S_{n+m} = j \mid S_m = a)$

Proof: (i) $P(S_n = j+b \mid S_0 = a+b) = P(\sum_{k=1}^{n} X_k = j)$

(ii) $P(S_{n+m} = j \mid S_m = a) = P(S_{n+m} = j, S_m = a) / P(S_m = a)$

$= P(S_m = a, \sum_{k=m+1}^{n+m} X_k = j-a) / P(S_m = a)$

independence

$= \cancel{P(S_m = a)}\, P(\sum_{k=m+1}^{n+m} X_k = j-a) / \cancel{P(S_m = a)}$

Prop ③ (Markov Property) For a SRW

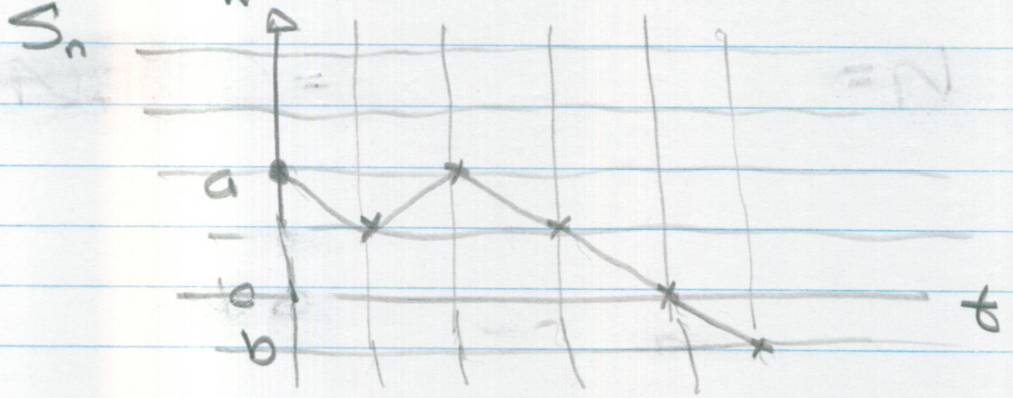
$$P(S_n = i_n \mid S_0 = a, S_1 = i_1, \ldots, S_{n-1} = i_{n-1}) = P(S_n = i_n \mid S_{n-1} = i_{n-1})$$

Proof: $\text{LHS} = \dfrac{P(X_1 = i_1 - a, X_2 = i_2 - i_1, \ldots, X_n = i_n - i_{n-1})}{P(X_1 = i_1 - a, \ldots, X_{n-1} = i_{n-1} - i_{n-2})}$

$\overset{\text{ind. of } X_i}{=} P(X_n = i_n - i_{n-1}) = P(S_n = i_n \mid S_{n-1} = i_{n-1})$.

- so we can ignore past as future only depends on present

- let $N_n(a,b)$ = # of paths from $(0,a)$ to (n,b)



- note each such path has the same values for
$N_+(a,b)$ = # of 1's, $N_-(a,b)$ = # of −1's since

$$N_+(a,b) + N_-(a,b) = n$$
$$N_+(a,b) - N_-(a,b) = b - a$$

which implies $N_+(a,b) = \frac{n+b-a}{2}$, $N_-(a,b) = \frac{n+a-b}{2}$

Lemma $N_n(a,b) = \binom{n}{\frac{n+b-a}{2}}$

Proof: We can choose the $N_+(a,b)$ steps in this many ways.

Prop 2 $P(S_n = b \mid S_0 = a) = N_n(a,b)\, p^{\frac{n+b-a}{2}} q^{\frac{n+a-b}{2}}$.

eg SRW with absorbing barriers (Gambler's ruin)

- suppose $0 \le S_0 = k \le N$ and the SRW stops whenever $S_n = 0$ or $S_n = N$ (k = initial cap of gambler, $N-k$ = initial cap of casino)
- want to compute

$$p_k = P(S_n = 0, \text{ some } n \geq 0 \text{ and } S_1, \ldots, S_{n-1} \neq N \mid S_0 = k)$$

note this prob involves infinitely many r.v.'s

- we have the following difference equation

$$p_k = p_{k+1}\,p + p_{k-1}\,q \qquad (*)$$

condition on what happens at time $n=1$ + temporal homogeneity

with boundary conditions $p_0 = 1$, $p_N = 0$

- note that $p_k \equiv 1$ and $p_k = \left(\frac{q}{p}\right)^k$ both sat $(*)$ and thus so does $p_k = A + B\left(\frac{q}{p}\right)^k \ \forall A, B$
- the bdry conditions imply

$$1 = A + B \quad \text{and} \quad 0 = A + B\left(\frac{q}{p}\right)^N$$

which in turn imply $0 = A\left(1 - \left(\frac{q}{p}\right)^N\right) + \left(\frac{q}{p}\right)^N$

- then provided $p \neq q$ this gives

$$A = -\left(\frac{q}{p}\right)^N \Big/ \left[1 - \left(\frac{q}{p}\right)^N\right], \quad B = 1 - A = 1 \Big/ \left[1 - \left(\frac{q}{p}\right)^N\right]$$

$$p_k = \frac{\left(\frac{q}{p}\right)^k - \left(\frac{q}{p}\right)^N}{1 - \left(\frac{q}{p}\right)^N}$$

- putting $x = \frac{q}{p}$ we have, by L'Hospital, that

$$p_k = \lim_{x \to 1} \frac{x^k - x^N}{1 - x^N} = \lim_{x \to 1} \frac{k x^{k-1} - N x^{N-1}}{-N x^{N-1}} = \frac{k-N}{-N} = 1 - \frac{k}{N}$$

and this solves ⊛ and sats bdry conditions when $p = q$

- note that a solution to ⊛ is determind by p_0 and p_N and so is unique

$p_k = p_{k+2} p^2 + p_k pq + p_k pq + p_{k-2} q^2$

- $$\lim_{N \to \infty} p_k = \begin{cases} \left(\frac{q}{p}\right)^k & q < p \\ 1 & \text{otherwise} \end{cases}$$

so when casino has advantg you go bankrupt with prob. 1

- similarly we can calculate

$$q_k = P(S_n = N \text{ some } n \text{ and } S_1, \dots, S_{n-1} \notin \{0, N\} \mid S_0 = k)$$

$$= \begin{cases} \dfrac{1 - \left(\frac{q}{p}\right)^k}{1 - \left(\frac{q}{p}\right)^N} & p \neq q \\ \dfrac{k}{N} & p = q \end{cases}$$

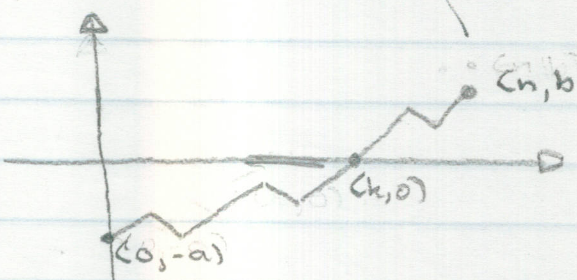
$\therefore \quad p_k + q_k = 1$

so either the gambler or the casino goes bankrupt

- let $N_n^0(a,b)$ = # of paths from $(0,a)$ to (n,b) that contain $(k,0)$ some $k \in \{1,\dots,n\}$

Prop ④ (Reflection Principle)
If $a,b>0$ then $N_n^0(a,b) = N_n(-a,b)$

Proof: Each path from $(0,-a)$ to (n,b) intersects the x-axis at some first point $(k,0)$. Reflect the



path with $0 \le x \le k$ in the x-axis to obtain a path connecting $(0,a)$ to (n,b) and note the 1-1 correspondence.

Corollary (Ballot Theorem)
If $b>0$ then the number of paths from $(0,0)$ to (n,b) which do not revisit the x-axis is $\frac{b}{n} N_n(0,b)$

Proof: This number is $N_{n-1}(1,b) - N_{n-1}^0(1,b)$

$$= N_{n-1}(1,b) - N_{n-1}(-1,b) = \binom{n-1}{\frac{n-2+b}{2}} - \binom{n-1}{\frac{n+b}{2}}$$

$$= \frac{(n-1)!}{\left(\frac{n-2+b}{2}\right)!\left(n-1-\frac{n-2+b}{2}\right)!} - \frac{(n-1)!}{\left(\frac{n+b}{2}\right)!\left(n-1-\frac{n+b}{2}\right)!}$$

$$= \frac{b}{n} N_n(0,b)$$

Why "ballot"? winner receives α votes and loser receives β and voters vote sequentially in random order. What is prob winner is always ahead? There are $\frac{\alpha-\beta}{\alpha+\beta} N_{\alpha+\beta}(0,\alpha-\beta)$ such paths out of $N_{\alpha+\beta}(0,\alpha-\beta)$

Prop ⑤ $P(S_1 S_2 \cdots S_n \neq 0,\ S_n = b \mid S_0 = 0) = \frac{|b|}{n} P(S_n = b \mid S_0 = 0)$

Proof: $P(S_1 S_2 \cdots S_n \neq 0,\ S_n = b \mid S_0 = 0)$

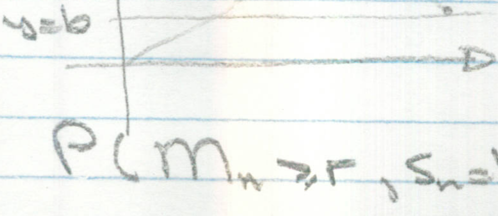
$$\overset{\text{Ballot Thm}}{=} \frac{|b|}{n} N_n(0,b)\ p^{\frac{n+b}{2}} q^{\frac{n-b}{2}} \overset{\text{Prop ⑤}}{=} \frac{|b|}{n} P(S_n = b \mid S_0 = 0)$$

- let $M_n = \max\{S_i \mid 0 \leq i \leq n\}$

Prop ⑥ $P(M_n \geq r,\ S_n = b \mid S_0 = 0) = \begin{cases} P(S_n = b \mid S_0 = 0), & b \geq r \\ \left(\frac{q}{p}\right)^{r-b} P(S_n = 2r-b \mid S_0 = 0) & b < r \end{cases}$

Proof: If $b \geq r$ the result is obvious and so assume $b < r$. Let $N_n^r(0,b)$ = # of paths from $(0,0)$ to (n,b) that include a point (i,r), $0 < i < n$.

Reflect the path from the earliest such point to (n,b) about the line $y=r$ to form a path from $(0,0)$ to $(n, 2r-b)$.



Note the 1-1 correspondence and thus

$$P(M_n \geq r,\ S_n = b \mid S_0 = 0) = N^r(0,b)\ p^{\frac{n+b}{2}} q^{\frac{n-b}{2}}$$

p goes from (0,0) to (n,b) [illegible] , $N_n N_n b$

Prop ③

$$= N^{r-b}_{n}(0, 2r-b)\, p^{\frac{n+b}{2}} q^{\frac{n-b}{2}} = \left(\frac{q}{p}\right)^{r-b} N_n(0, 2r-b)\, p^{\frac{n+2r-b}{2}} q^{\frac{n-2r+b}{2}}$$

$$= \left(\frac{q}{p}\right) P(S_n = 2r-b \mid S_0 = 0)$$

<u>Prop</u> ⑦ <u>(Hitting Time Theorem)</u> For $b \neq 0$

$$f_b(n) = P(\text{ S hits b for first time at } t=n \mid S_0 = 0)$$

$$= \frac{|b|}{n} P(S_n = b \mid S_0 = 0) \text{ if } n \geq 1$$

Proof: For $b > 0$ and writing $P_0 = P(\cdot \mid S_0 = 0)$ then

$$f_b(n) = P_0(M_{n-1} = S_{n-1} = b-1,\ S_n = b)$$

$$= p\, P_0(M_{n-1} = S_{n-1} = b-1)$$

$$= p\left[P_0(M_{n-1} \geq b-1,\ S_n = b-1) - P_0(M_{n-1} \geq b,\ S_{n-1} = b-1)\right]$$

Prop⑥

$$= p\left[P_0(S_{n-1} = b-1) - \left(\frac{q}{p}\right) P_0(S_{n-1} = b+1)\right]$$

$r = b$
$2r - b + 1 = b + 1$
$r - (b-1) = 1$

Prop③

$$= p\left[N_{n-1}(0, b-1)\, p^{\frac{n+b}{2}-1} q^{\frac{n-b}{2}} - \left(\frac{q}{p}\right) N_{n-1}(0, b+1)\, p^{\frac{n+b}{2}} q^{\frac{n-b}{2}-1}\right]$$

$$= p^{\frac{n+b}{2}} q^{\frac{n-b}{2}} \left[\frac{(n-1)!}{\left(\frac{n+b-2}{2}\right)!\left(\frac{n-b}{2}\right)!} - \frac{(n-1)!}{\left(\frac{n+b}{2}\right)!\left(\frac{n-b-1}{2}\right)!}\right]$$

$$= \frac{1}{n} \binom{n}{\frac{n+b}{2}} p^{\frac{n+b}{2}} q^{\frac{n-b}{2}} \left[\frac{n+b}{2} - \frac{n-b}{2}\right]$$

$$= \frac{b}{n} P_0(S_n = b) \text{ and similarly when } b < 0.$$

- put $f_b = \sum_{n=1}^{\infty} f_b(n) = P(S_n = b, \text{ some } n \mid S_0 = 0)$ $b \neq 0$
  and $f_0 = 1$

Prop 8 If $S_0 = 0$ then the mean no. of visits to $b \neq 0$ before returning to 0 is $\mu_b = f_b$.

Proof: $\mu_b = E\left[\sum_{n=1}^{\infty} I_{\{S_1 \cdots S_n \neq 0, S_n = b\}} \mid S_0 = 0\right]$

→ monotone convergence

$\overset{MCT}{=} \sum_{n=1}^{\infty} E\left[I_{\{S_1 \cdots S_n \neq 0, S_n = b\}} \mid S_0 = 0\right]$

$= \sum_{n=1}^{\infty} P(S_1 \cdots S_n \neq 0, S_n = b \mid S_0 = 0)$

$\overset{\text{Prop 6}}{=} \sum_{n=1}^{\infty} \frac{|b|}{n} P(S_n = b \mid S_0 = 0) \overset{\text{Hitting theorem}}{=} \sum_{n=1}^{\infty} f_b(n)$.

Corollary If $p = \frac{1}{2}$, then $\mu_b = 1$.

Proof: We have $f_b = \frac{1}{2}(f_{b-1} + f_{b+1})$ with the boundary conditions $f_0 = 1$, $f_{-b} = f_b$ (by symmetry). Then $f_0 = \frac{1}{2}(f_1 + f_{-1}) = f_1$,
$f_1 = \frac{1}{2}(f_0 + f_2)$ and thus $f_1 = f_2$ etc. which implies $f_b \equiv 1$.

condition on first step

$\Rightarrow$ $pP(S_n = b \text{ some } n \mid S_0 = 0, S_1 = 1)$
$+ qP(S_n = b \text{ some } n \mid S_0 = 0, S_1 = -1)$
$\overset{MP}{=} pP(S_n = b \text{ some } n \mid S_1 = 1) + \cdots$
$\overset{\text{time homog}}{=} pP(S_{n-1} = b \text{ some } n-1 \mid S_0 = 1) +$
$\overset{\text{spat. homog}}{=} pP(S_{n-1} = b-1 \text{ some } n-1 \mid S_0 = 0)$

after an even # of steps walk is in an even state

10.

Prop 9 (Arc sine law for last visit to origin)

For $p=\frac{1}{2}$, $0<k<n$ $P(S_{2k}=0, S_{2k+1} \ldots S_{2n} \neq 0 \mid S_0=0)$ = prob time of last visit to 0 before time 2n is 2k.

$= P(S_{2k}=0 \mid S_0=0)\, P(S_{2n-2k}=0 \mid S_0=0)$

arc-sine distribution at 2k

Proof: $P(S_{2k}=0, S_{2k+1} \ldots S_{2n} \neq 0 \mid S_0=0)$

mp
$= P(S_{2k}=0 \mid S_0=0)\, P(S_{2k+1} \ldots S_{2n} \neq 0 \mid S_{2k}=0)$

time homo
$= P(S_{2k}=0 \mid S_0=0)\, P(S_1 \ldots S_{2n-2k} \neq 0 \mid S_0=0)$

and $P(S_1 \ldots S_{2n-2k} \neq 0 \mid S_0=0)$

$= \sum_{b=-1}^{-\infty} P(S_1 \ldots S_{2n-2k} \neq 0, S_{2n-2k}=2b \mid S_0=0)$

$+ \sum_{b=1}^{\infty} P(S_1 \ldots S_{2n-2k} \neq 0, S_{2n-2k}=2b \mid S_0=0)$

symmetry
$= 2 \sum_{b=1}^{\infty} P(S_1 \ldots S_{2n-2k} \neq 0, S_{2n-2k}=2b \mid S_0=0)$

Prop 5
$= 2 \sum_{b=1}^{\infty} \frac{2b}{2n-2k} P(S_{2n-2k}=2b \mid S_0=0)$

max value for $2b = 2n-2k$ in $2n-2k$ times

Prop 3
$= 2 \sum_{b=1}^{n-k} \frac{2b}{2n-2k} \binom{2n-2k}{\frac{2n-2k+2b}{2}} \left(\frac{1}{2}\right)^{2n-2k}$

$= \binom{2n-2k}{n-k} \left(\frac{1}{2}\right)^{2n-2k} = P(S_{2n-2k}=0 \mid S_0=0)$

since

$$2\sum_{b=1}^{n-k}\frac{2b}{2n-2k}\binom{2n-2k}{\frac{2n-2k+2b}{2}}\left(\frac{1}{2}\right)^{2n-2k}$$

$$=2\left(\frac{1}{2}\right)^{2n-2k}\sum_{b=1}^{n-k}\frac{2b}{2n-2k}\binom{2n-2k}{n-k+b}$$

$$=2\left(\frac{1}{2}\right)^{2n-2k}\sum_{b=1}^{n-k}\left[(n-k+b)-(n-k-b)\right]\frac{(2n-2k-1)!}{(n-k+b)!\,(n-k-b)!}$$

$$=2\left(\frac{1}{2}\right)^{2n-2k}\sum_{b=1}^{n-k}\left[\binom{2n-2k-1}{n+k+b-1}-\binom{2n-2k-1}{n-k+b}\right] \quad \text{=0 when } b=n$$

~~telescope~~

$$=2\left(\frac{1}{2}\right)^{2n-2k}\binom{2n-2k-1}{n+k}=2\left(\frac{1}{2}\right)^{2n-2k}\frac{(2n-2k-1)!}{(n+k)!\,(n+k-1)!}$$

$$=\left(\frac{1}{2}\right)^{2n-2k}\binom{2n-2k}{n-k}$$

note "Why arc sine"?

- $P(S_{2k}=0|S_0=0)\,P(S_{2n-2k}=0|S_0=0)$

$$= \binom{2k}{k}\left(\tfrac{1}{2}\right)^{2k}\binom{2n-2k}{n-k}\left(\tfrac{1}{2}\right)^{2n-2k}$$

$$= \frac{(2k)!}{k!\,k!}\,\frac{(2n-2k)!}{(n-k)!\,(n-k)!}\left(\tfrac{1}{2}\right)^{2n}$$

Stirling's formula $n! \sim n^n e^{-n}(2\pi n)^{\frac{1}{2}}$ as $n\to\infty$

next page for details

$$\sim \frac{1}{\pi}\frac{1}{\sqrt{\frac{k}{n}\left(1-\frac{k}{n}\right)}}\frac{1}{n} \quad \gamma_{n,k} \text{ as } k, n-k \to\infty$$

visit to 0 means equal ~~heads and~~ tails

- if T_{2n} = time of last visit to 0 up to time $2n$

$0<x<1$ then $P(T_{2n}\le 2xn|S_0=0) = \sum_{k=1}^{\lfloor xn\rfloor} P(S_{2k}=0|S_0=0)\,P(S_{2n-2k}=0|S_0=0)$

$$\sim \sum_{k=1}^{\lfloor xn\rfloor}\frac{1}{\pi}\frac{1}{\sqrt{\frac{k}{n}\left(1-\frac{k}{n}\right)}}\frac{1}{n} \qquad \text{Riemann sum}$$

$$\sim \frac{1}{\pi}\int_0^x \frac{1}{\sqrt{z(1-z)}}\,dz = \frac{2}{\pi}\arcsin\sqrt{x}$$

since $\frac{\lfloor xn\rfloor}{n}\sim x \quad \left(\frac{xn-1}{n}\le\frac{\lfloor xn\rfloor}{n}\le\frac{xn+1}{n}\right)$

- e.g. $x=\frac{1}{2}$ $\quad P(T_{2n}\le n)\approx\frac{2}{\pi}\arcsin\sqrt{\tfrac{1}{2}}=\frac{1}{2}$

take n big then walk spends a lot of time either always pos. or always neg. with significant prob.

Subsidiary Calculations

$$\frac{(2n-2k)!}{(n-k)!}\left(\frac{1}{2}\right)^{2n-2k} \overset{n\to\infty}{\sim} \frac{(2n-2k)^{2n-2k}\, e^{-(2n-2k)}\, \left(2\pi(2n-2k)\right)^{1/2}}{(n-k)^{2(n-k)}\, e^{-2(n-k)}\, \left(2\pi(n-k)\right)} \left(\frac{1}{2}\right)^{2n-2k}$$

$$= \frac{2^{1/2}}{\left(\pi(n-k)\right)^{1/2}}$$

so goes to 0 as $n-k \to \infty$, k fixed

$$\frac{(2k)!}{k!\,k!}\left(\frac{1}{2}\right)^{2k} \overset{k\to\infty}{\sim} \frac{2^{1/2}}{(\pi k)^{1/2}}$$

so goes to 0 k large, $n-k$ fixed.

So choose k_0, n_0 s.t. $\forall k > k_0,\ n-k > n_0 - k_0$, $\left(k_0 \le k \le n-(n_0-k_0)\right)$ implies $1-\varepsilon \le P(\ )/\tau_{n,k} \le 1+\varepsilon$ and $\sum_{k \le k_0} + \sum_{k \ge n_0 - k_0} < \varepsilon$

Then $$\left|\frac{\sum P(\ )}{\sum \tau_{n,k}} - 1\right| \le \frac{1}{\left|\sum \tau_{n,k}\right|}\left|\sum_{k\le k_0} + \sum_{k \ge n-n_0+k_0}\right| \text{ finite sum } + \frac{1}{\sum \tau_{n,k}}\left|\sum_{k_0 < k < n-n_0+k_0} P(\ ) - \sum \tau_{n,k}\right| + \frac{1}{\sum \tau_{n,k}}\left|\sum_{k \le k_0} \tau_{n,k} + \sum_{k \ge n-n_0+k_0} \tau_{n,k}\right|$$

$$(1-\varepsilon)\sum_{k_0<k<n-n_0+k_0} \tau_k - \sum \tau_k = -\varepsilon \sum_{k_0<k<n-n_0+k_0} > -\varepsilon \sum_k \tau_k$$

$$(\arcsin x)' = \frac{1}{\sqrt{1-x^2}}$$

$$(\arcsin \sqrt{x})' = \frac{1}{2}\frac{1}{\sqrt{x(1-x)}}$$

$\sin x = \frac{1}{\sqrt{2}}$. $x = \frac{\pi}{4}$

since $\sin x = \cos x$ when $x = \frac{\pi}{4}$

so $2\sin^2 x = 1$