Mosteller Cheat sheet by Alain Chenier, page 1 of 2, 13th March 2017

1 Sock Drawer

Question

drawer contains r red sock and b blue socks. 2 socks are drawn. **force probability to be 1/2.**
a) how small num of red and black socks
b) how small num of socks if num black socks is even.

Answer

$$\text{prob first is red } \frac{r}{r+b} \text{ then prob second is red given first is red } \frac{r-1}{r+b-1}$$

$$\text{so need to find r b such that} \left(\frac{r}{r+b}\right)\left(\frac{r-1}{r+b-1}\right) = \frac{1}{2}$$

$$\frac{r}{r+b} > \frac{r-1}{r+b-1} \text{so with 1/2 condition above}$$

$$\frac{r}{r+b} > \frac{1}{\sqrt{2}} > \frac{r-1}{r+b-1} \quad \text{so } (\sqrt{2}+1)b+1 > r > (\sqrt{2}+1)b$$

2 Win 2 tennis matches in a row

Question

win at least 2 tennis matches in a row in a 3 set with Champion > Father
Q: Choose Father-Champion-Father or Champion - Father - champion ?

Answer

Compare both sequences FCF CFC, add up the probs of winning:

F	F	F	Prob
W	W	W	fcf
W	W	L	fc(1-f)
L	W	W	(1-f)cf
Total			**fc (2-f)**

F	F	F	Prob
W	W	W	cfc
W	W	L	cf(1-c)
L	W	W	(1-c)fc
Total			**fc(2-c)**

3 Jury: 2 with p + random or 1 with p ?

Answer

prepare table of all correct decisions:

Serious agree	p	p	not matter as overruled	p^2
Serious disagree 1	p	1-p	chooses right with 1/2	$p(1-p)\frac{1}{2}$
Serious disagree 2	1-p	p	chooses right with 1/2	$(1-p)p\frac{1}{2}$
Total				$p^2+p(1-p)=p$

4 How many throws of dice to get 6 ?

Answer 1 - force

let p = prob of 6 = 1/6 then prob of success p (len=1). qp (len2), qqp (len 3) ...

$$m = 1(p)+2(qp)+3(qqp)... = p(1+2q+3qq+...) = p\left(\frac{1}{1-q}\right)^2 = p\frac{1}{p^2} = \frac{1}{p}$$

Answer 2 - First step analysis = Markov FSA

If first is failure, then m->m+1, if first is good then m=1 so

$$m = q(m+1)+p(1) \Leftrightarrow m(1-q) = q+p = 1 \Leftrightarrow m = \frac{1}{p}$$

Answer 3 - Martingale

...put martingale here...

5 throw 3/4cm diameter coin on 1cm x 1cm grided table

Question

Get 5c (in all) if whole inside a square , lose 1c otherwise. CHANCE TO WIN and How MUCH on AVG?

Answer

Centre of coin falls in one square, so consider one square, and divide allowable area for centre by area of square

 | $p = \frac{1}{4}\text{x}\frac{1}{4}$and $m = (+5)\frac{1}{16}+(-1)\frac{15}{16}$

6 Dice Game

Question

Bet on 1 or 2 or ... or 6 and throw 3 dice. Receive 1x-2x-3x stake + original stake. WHAT IS the EXPECTED LOSS ?

Answer

when you roll either all 3 diff, or 2 same, or 3 same. Bet 6 on all 6. If all 3 different then house wins 3, you lose 3 (tot 0), if 2 same h wins 1(for die=3)+1 (for die=4)+1(for 5)+1(for 6), you win 2(for 1)+1(for 2) (tot you=3-4=**-1**). if 3 same you win 3 house wins 5 (tot you = 3-5=-2). Then count how many ways to get 1 alike, 2 alike, 3 alike and multiply by loss.
Total ways = $6x6x6 = 216$, all 3 different = $6x5x4 = 120$, all 3 same = $6x1x1 = 6$, therefore 2 same $216-120-6 = 216-126 = 90$ [Note: 2 same (alt way) : die 1=2 and 3 dif). or (die 1=3 and 2 dif) or (die 2=3 and 1 dif)so $6x5+6x5+6x5 = 30x3 = 90$] So loss for betting on all 6 = all diff + 2 same + 3 same = $\frac{0}{6}\frac{120}{216}+\frac{(-1)}{6}\frac{90}{216}+\frac{(-2)}{6}\frac{6}{216} = \frac{-102}{216} = \frac{-17}{36}$

7 Follow-up Bet on outcome of initial roulette bet

Question

Bet 1 number out of 38 then bet 20 ahead after 36 games. HOW MUCH IS THE WIN or LOSS ?

Answer

(expected **gain** one turn) $= (+35)\frac{1}{38} + (-1)\frac{37}{38} = \frac{-2}{38}$, and in 36 turns (expected **gain** 36 turns) $= m_1 = (36)\frac{-2}{38} = -\frac{72}{38}$
Now for the 20 bet: 'ahead in 36 games' = 'win once or more in 36 games' since **gain** (for win 0 times in 36) = $36(-1) = -36 < 0$, and **gain** (win 1 in 36) = $1(+35)+35(-1) = 0 >= 0$ so relevant prob for the 20 bet = P(win one or more in 36)= 1- P(lose all 36) = $1-\left(\frac{37}{38}\right)^{36} = 0.617$ so for the 20 bet expected **gain** = $m_2 = (+20)(+0.617)+(-20)(1-0.617)$ and so total **gain** $m = m_1+m_2 = ...$

8 All cards same suit in 52 card pack

Question

Probability of all 13 cards of same suit?

Answer 1 - straight

$P = \frac{\text{good hands}}{\text{all hands}}$. for good: pick any card , after must be of same suit = $52*12*11*...*1 = 52*12!$, for all hands: any card then, any card left, ... (13 times)= $52*51*...*40 = \frac{52!}{39!}$ so Prob = $\frac{52*12!}{\frac{52!}{39!}} = \frac{39!12!}{51!}$

Answer 2 - binomials

$P = \frac{\text{good hands}}{\text{all hands}}$. for good: pick 1 of 4 suits, then pick 13 cards in the 13 card suits , for all, pick 13 cards in 52, so $P = \frac{\binom{4}{1}\binom{13}{13}}{\binom{52}{13}} = \frac{(4)(1)}{\frac{52!}{39!13!}} = \frac{4*13!*39!}{52!} = \frac{4*13*12!*39!}{52*51!} = \frac{12!*39!}{51!}$ = same as above!!

9 Craps game

Game description and question

Only total of sum of 2 dice counts. Lose if sum=2,3,12. Win if sum=7,11. record sum otherwise, then keep throwing until either 'same sum again' (win) or 7 (lose). **Question:** odds of winning ?

Answer 1 - Conditional probability

get probs of totals by counting in the matrix:

	1	2	3	4	5	6
1	2	3	4	5	6	7
2	3	4	5	6	7	8
3	4	5	6	7	8	9
4	5	6	7	8	9	10
5	6	7	8	9	10	11
6	7	8	9	10	11	12

so

total	2	3	4	5	6	7	8	9	10	11	12
p(total)	$\frac{1}{36}$	$\frac{2}{36}$	$\frac{3}{36}$	$\frac{4}{36}$	$\frac{5}{36}$	$\frac{6}{36}$	$\frac{5}{36}$	$\frac{4}{36}$	$\frac{3}{36}$	$\frac{2}{36}$	$\frac{1}{36}$

so $P(\text{win 1st throw}) = P(7)+P(11) = \frac{6}{36}+\frac{2}{36}$ and $P(\text{loss 1st throw}) = P(2)+P(3)+P(12) = \frac{1}{36}+\frac{2}{36}+\frac{1}{36}$ then after first throw only: throwing either 'initial sum again' (which was either=4,5,6,8,9 or 10) (win) or 7 (lose) matters, so we have a conditional prob on first throw. Say the 'initial sum was 4', then win if 4, or lose if 7, so P(win if init sum=4)=P(win|4)=$\frac{\text{num ways of 4}}{\text{num ways of 4 or 7}} = \frac{3}{3+6}$ so P(win if 4)=P(4)*P(win|4), so in total P(win)=P(7)+P(11)+P(4)*P(win|4)+P(5)*P(win|5)+P(6)*P(win|6)+P(7)*P(win|7) +P(8)*P(win|8)+P(9)*P(win|9)+P(10)*P(win|10) and expected gain = (+1)*P(win)+(-1)*P(loss)

note: Answer 1' - infinite summation method untill you get the init sum or 7

if into 2nd or more throw , you have the possibilities P(init sum again), P(ignore ie not count), or P(7), and win if 'init again', or, 'not count'+'init again, or, 'not count' 'not count' then 'init again' etc. So say init sum was 4 (same as above), then P(win|init sum=4) = P+RP+RRP+... with P=prob(init sum of 4 again) and R=prob(not count), which is P(1+R+RR+..)=P/(1-R) , so in summary to win with a 4 then P(4)*P/(1-R) with P=3/36 and R=1-P(4)-P(7)=1-3/36-6/36 – note P/(1-R) = P/(P+Q) <1 if Q=prob(7)

10 Which game do you prefer

Question: prefer known contents or unknown contents?

game1:urn1:10 black balls,urn2: 10 white balls, pick a colour
game2:urn1:unknown,urn2: urn2, pick a colour

Answer: same

game1: P(win)=P(urn1)P(B|urn1)+P(urn2)P(B|urn1)=1/2 1/2 + 1/2 1/2 = 1/2 for B=black, same for W
game2: P(win)=P(urn1)P(B|urn1)+P(urn2)P(B|urn1)=1/2 x + 1/2 (1-x) = 1/2 for B=black, same for W because $x = \blacksquare/total$ so in urn2: $(total-\blacksquare)/total = 1-x$

11 Prisoner's dilemma

Question - find the flaw in this argument

3 prisoners A,B,C and 2 will be released. **W**arden knows. if A asks W, W will name 1 other than A. A thinks his chance before asking =2/3, and if **W**arden say 'B' then chance now = 1/2 < 2/3, hence should not ask. Where the flaw ?

Answer: list ALL events, must condition on the Warden

Num	Release Outcome	Warden	Prob outcome AND W	example W says B
1	**A,B**	C must	(1/3)(1)=1/3	
2	**A,C**	B must	(1/3)(1)=1/3	ok
3	B,C	B	(1/3)(1/2) = 1/6	ok
4	B,C	C	(1/3)(1/2) = 1/6	

so P(A released | W says B) = P(A released and W says B) /P(W says B)= row2/(row2+row3) = (1/3)/(1/6+1/3) = (1/3)/(3/6) = 6/9 = 2/3 = same as without asking similarly P(A released | W says C) = P(A released and W says C) /P(W says C)= row1/(row1+row3) = same since same probs = same as without asking

12 Coupon Collector

Question: How many boxes required to collect all 5 coupons

5 possible coupons 1..5, 1 per box, how many boxes on average to get all 5?

Answer: See FSA problem 4 - num throws to get a 6 = 1/(1/6)=6

1st box always good. second box:new number with prob (4/5) $\to$ from FSA prob 4 must wait 1/(4/5)=5/4, third box: new number with prob (3/5) $\to$ must wait 1/(3/5)=5/3 etc so answer $5/4+5/3+5/2+5/1 = 5(1/4+1/3+1/2+1)$. Note $1+1/2+..+1/n = \ln n + \frac{1}{2n} + euler$ so $n\ln n + 1/2 + n\gamma \equiv n\ln n + n\gamma$for big n

13 Theater Row - pairs of adjacent but different

Question: How many pairs of adjacent but different

8 B and 7 M in a row, how many pairs of **BM** or **MB** on average? For example **BMB**MMMMMM**MB**BBB is 3 good pairs. **Then same question with 6B 5M 4W**

Answer1: prob 2 balls are adjacent x num of good pairs

prob 2 balls in the 15 are adjacent = 14 slots possible / (15 choose 2) , then num pairs=8x7=56 so $\frac{14}{\binom{15}{2}}(56) = \frac{14}{\frac{15x14}{2}}(56) = (2)\frac{56}{15}$. For next one first part is exact same !, just more pairs=(6x5+6x4+5x4)

Answer2: indicator variable + linearity of expectation even though the X_i not independent!!

score 1 if slot 1,2 = good, 0 otherwise, and same for slot 2,3 etc ($\leftarrow$ the indicator variable). First slot 1,2 is good if (BM or MB) ie $\frac{8}{15}\frac{7}{14}+\frac{7}{15}\frac{8}{14}$, same for slot 2,3 ... slot 14,15 so $14(\frac{8}{15}\frac{7}{14}+\frac{7}{15}\frac{8}{14}) = (2)\frac{56}{15}$ ($\leftarrow$ linearity of expectation even if not independent!!)

14 2^n Tennis tournament with elimination

Question: Prob second-best makes it to runner-up?

$2^3 = 8$ players, better one always defeats next best. **Question**: prob second best is in final ?

Answer: Second best must be in lower half to meet the best in the final (where he will be defeated of course)

1 2 **3** 4 5 6 7 8 ($\leftarrow$ second-best must occupy a lot in lower half to win it)= $\frac{\text{good slots for 2nd best}}{\text{possible slots for 2nd-best – keep 1 for best}} = \frac{2^{n-1}}{2^n-1}$

15 2^n tournament - prob of meeting at all

Question: Chance of meeting in tournament

Answer: enumerate

1 2 3 4 5 6 7 8 <= $\frac{1}{7}$ chance, and they meet with prob 1,**1 3 or 1,4** the rest <= $\frac{1+1}{7}$ chance, and they meet with prob $\frac{1}{2}, \frac{1}{2}$ as they both have to win, finally bottom half: **1 5,6,7,8** with prob $\frac{1+1+1+1}{7}$ and they meet with prob $\frac{1}{2x2x2x2}$ as they both have to win their matches and meet in final, so $\frac{1}{7}(1)+\frac{2}{7}\frac{1}{4}+\frac{4}{7}\frac{1}{16} = \frac{1}{4}$. Can prove general case by induction.

16 Bernoulli trial - prob of given number

Question: throw 100 fair coins, prob of 50 heads exactly?

Answer: $\binom{100}{50}\left(\frac{1}{2}\right)^{100} = \frac{100!}{50!50!}\left(\frac{1}{2}\right)^{100}$

use sterling approximation to evaluate $n! = \sqrt{2\pi}n^{n+1/2}e^{-n}$

17 1 6 with 1 dice, 2 6s with 2x6 dice, 3 6s with 3x6 dice?

Answer: enumerate the probs for 1 6, 2 6s,

prob of 1 6 exactly with 6 dice: (6 choose 1)(1 6 exactly) = $\binom{6}{1}\frac{1}{6}\frac{5}{6}^5$, prob of x 6s exactly with n dice: (n choose x)(x 6s exactly in n trials) = $\binom{n}{x}\frac{1}{6}^x\left(\frac{5}{6}\right)^{n-x}$

18 3 way duel

Question

A,B,C fight duel. A shoots and wins with prob 0.3, B with prob 1, C with prob 0.5. Best strategy for A ?

Answer: enumerate the options

if A shoots C , B will shoot A so no good - if A shoots B and misses then B will choose to shoot C since C better than A, then A gets to shoot B with his prob =0.3 - if A shoots B and wins, then C <-> A <-> C <-> A ... untill one dies, and the prob sequence is P(A eventual wins)= P(A terminates the sequence) = (C miss, A wins) + (C miss, A miss, C miss, A wins) + ... = $(.5)(.3)+(.5)^2(.7)(.3)+(.5)^3(.7)^2(.3)+... = (.5)(.3)((\text{geometric of }.5\ .7) < 0.3$ so A should deliberately miss B and let B take care of C and then try hitting B with his prob 0.3

19 Sample with and without replacement

Urn A has 2R 1B, Urn B has 101 R,100 B. try and guess the right urn based on 2 draws from same urn: 1st draw, see the colour, choose to replace it or not replace it before drawign again form same urn. which is best ?

Answer: enumerate

if 1st is red and urn 1, then 1B-1R and if urn2 100B -100R so 2nd draw from same urn is the same, so cannto tell which urn <– put it back if red

if 1st is black, then if no replace, then the probs depends on the sequence of colours observed as follows:

	prob it is urn A	prob it is urn B	biggest
R*,R	(choose A)(R)(R) = $\frac{1}{2}\frac{2}{3}\frac{2}{3}$	$\frac{1}{2}\frac{101}{201}\frac{101}{201}$	urn A
R*,B	$\frac{1}{2}\frac{2}{3}\frac{1}{3}$	$\frac{1}{2}\frac{101}{201}\frac{100}{201}$	B
B,R	$\frac{1}{2}\frac{1}{3}\frac{1}{1}$	$\frac{100}{201}\frac{101}{200}\frac{1}{1}$	A - note the no rep
B,B	$\frac{1}{2}\frac{1}{3}0$		B - note the 0 in A

20 Ballot theorem - part 1 - vote counting

Question suppose candidate A has more votes than B: a>b. Prob of having same number at least once at some point during the vote counting?

Answer

We know A does eventually win so ALL sequences with B starts first are good sequences, since there is eventual hit-and-overtake by A. All sequences with B first=$\frac{b}{a+b}$. Now each sequence has a mirror : B <-> A, so there is same number of sequence starting with A with eventual equal tally (then A wins). **There are no other good sequences with A start else they would have a mirror and we have already counted all the mirrors.** so answer = $\frac{2b}{a+b}$

21 Ballot theorem - part2 - Bernoulli sums

Question: Bernoulli sums of 1,-1 - prob of no tie

Answer: for even N=2n ..., prob a string x prob of a tie as per above

Prob of x wins: $\binom{N=2n}{x}$ each string prob $\frac{1}{2^N}$ so $\binom{N=2n}{x}\frac{1}{2^N}$, and to use part 1 we need a win, so if x<=n -> loss -> use part 1 direct -> prob of a tie = $\frac{2x}{N}$, if x<=n -> loss -> convert to loss to use part 1 -> -> prob of a tie = $\frac{2(N-x)}{N}$ then add up the possibilities x =0..n $\binom{N=2n}{x}\frac{1}{2^N}\frac{2x}{N}$ and for x =n+1..N $\binom{N=2n}{x}\frac{1}{2^N}\frac{2(N-x)}{N}$ then expand out and remember fact that $(1+1)^N = 2^N = \sum\binom{N}{x}$

22 Subway arrival

Answer: train1 arrives 3:00 3:10.., train 2 arrives 3:02 3:12... - must arrive in good interval to catch1 before 2

23 prob(length Random Chord) > radius

Answer: several possibilities

(1) chord distance / radius = $\frac{\sqrt{r^2-(r/2)^2}}{r} = \sqrt{3}/2$

(2) midpoint of chord less than altitude of regular hexagon = $\frac{\pi d^2}{\pi r^2} = (\sqrt{3}/2)^2 = 3/4$

(3) 2 random points on circle : given 1 point , second must fall either way of corresponding arc subtended by hexagon = 2 x 1/6 = 1/3 so bigger than radius = 1-1/3=2/3

24 probability of meeting if each arrive between 6/7am and stay 5mns

Answer: stretch of time between $y = x+1/12$ and $y = x-1/12$ with x=0..1

because y can arrive either way of $[x-1/12, x+1/12]$ with x 0..1 and constraint 0<y<1

25 fixed number of urns with sampling - part1

100 coins per urn, 1 fake coin per urn, draw 1 coin per urn, prob of detecting one false coin?

Answers

probability of no false coins in all 100 urns with 1 false coins per urn $(1-\frac{1}{100})^{100}$, for n: $(1-\frac{1}{n})^n$, probability of no false coins in all n urns with m false coins $(1-\frac{m}{n})^n \to e^{-m}$

26 multiple urns with sampling - part 2 – POISSON DISTRIBUTION INTRODUCTION

n coins per urn, m fake coins per urn, draw 1 coin from the n urns, prob of drawing exactly r false coins?

Answers

r events each with prob $p = m/n$ so P= (n choose r) (prob of a string of r good and n-r bad) = $\boxed{\binom{n}{r}\frac{m}{n}^r(1-\frac{m}{n}))^{n-r}}$, then with n large and m and r fixed $\boxed{\binom{n}{r}\frac{m}{n})^r(1-\frac{m}{n}))^{n-r} = \frac{1}{r!}\frac{n(n-1)(n-r+1)}{n^r}m^r(1-\frac{m}{n})^n(1-\frac{m}{n})^{-r} = \frac{1}{r!}m^re^{-m}}$ which is the POISSON distribution $\boxed{P(r) = e^{-m}\frac{m^r}{r!}\ \ r=0,1,2..}$ – NOTE: large n -> means r can be very large,=$P(0) = e^{-m}$

27 POISSON: Fixed number of colonies per plate with plates divided into small

Question1: large number of plates, fixed average m=3 number of colonies per plate, how many plates with exactly m=3 colonies ? Question 2: let m=3 be large, what happens ?

divide each plates into large number of small subplates, let p = prob = num colonies per sub-plate, then there on a plate $m = 3 = \text{(num sub plates)(prob per sub plate)} = np$ ie np stays constant. Then if prob of r colonies in 1 plate = n sub-plates, with each colony = prob p per sub-plate is Bernoulli experiment !

So $P(r) = \binom{n}{r}p^r(1-p)^{n-r}$ with $np = 3 = m \to p = \frac{m=3}{n}$ So $P(r) = \binom{n}{r}\frac{3}{n}^r(1-\frac{3}{n})^{n-r}$ with n $\to\infty$

See Question 26 ! POISSON distribution $\boxed{P(r) = e^{-m}\frac{m^r}{r!}\ \ m=3 \text{ and } r=0,1,2..}$, Now set m=3 , and then r=3=exact number to answer the question.

Actually m=3 is the mean of the POISSON distribution: $mean = \sum_{r=0}^{\infty} xP(x) = \sum_{r=0}^{r=\infty} rP(r) = \sum_{r=0}^{r=\infty} re^{-m}\frac{m^r}{r!} = \sum_{r=1}^{r=\infty} e^{-m}\frac{m^r}{r-1!} = me^{-m}\sum_{r=1}^{r=\infty}\frac{m^{r-1}}{r-1!} = me^{-m}\sum_0^{\infty}\frac{m^r}{r!} = me^{-m}e^m = m$

Again : POISSON distribution $\boxed{P(r) = e^{-m}\frac{m^r}{r!}}$ r=0,1,2.. . |Check out what happens when mean m $\to\infty$ :: Use sterling formula for r! with r=m $\to\infty$ ie $r! = \sqrt{2\pi}r^{r+\frac{1}{2}}e^{-m}$ so $P(r=m) = e^{-m}\frac{m^m}{m!} = e^{-m}\frac{m^m}{\sqrt{2\pi}m^{m+\frac{1}{2}}e^{-m}} = \frac{1}{\sqrt{2\pi m}}$

28 POISSON: Sell m=20 cakes per round - Prob of selling even number of cakes

Number of cakes sold per round = POISSON distribution = $\boxed{P(r) = e^{-m}\frac{m^r}{r!}}$ and 'selling evens' means r=0,or 2 or 4 etc so of course answer = $\sum_{r=even}^{\infty} P(r)$ now use expansions of $e^m = \sum\frac{m^r}{r!}$ and e^{-m} to eliminate odd terms and divide by 2, and with the e^{-m} in front of each expansion you get $\frac{1}{2}(1+e^{-2m})$ and set m=20 in there for the answer.

29 BIRTHDAY problem: smallest num. of people required in group for at least 2 of them in the group to have same birthdate with prob > 1/2

find prob of no person having same birthdate in group, take complement.

If N=365 birthdates, first person in group has N dates to choose from, second has $(N-1)$ dates to choose from ... r th person has $(N-r+1)$ dates to choose from so $N(N-1)..(N-r+1)$, to be divided by each person having N dates to choose from = N^r so $\frac{N(N-1)..(N-r+1)}{N^r}$ with N=365 and try rs such that the prob < 1/2 so that 1-prob > 1/2

approximation: $e^{-x} = 1-x$ x small and $\frac{N(N-1)..(N-r+1)}{N^r} = 1(1-\frac{1}{N})(1-\frac{2}{N})..(1-\frac{r-1}{N})$ so $e^Oe^{-\frac{1}{N}}e^{-\frac{2}{N}}..e^{-\frac{r-1}{N}} = e^{-\frac{\sum_0^{r-1} i}{N}} = e^{-\frac{r(r-1)}{2N}}$

30 PERSONAL BIRTHDAY problem: how many people needed in group for one in group to have same birthday as you with prob > 1/2

Note **difference**: need person in the group to have same birthday as you - NOT same birthday as each other (and different from you). this is the PERSONAL birthday problem.

say N=365 then Each person has $\frac{N-1}{N}$ chance NOT having your birthday, so in group of people : (not having yoru birthdays) (not having yoru birthday) .. (not having your birthday)=$(\frac{N-1}{N})^r$ so 1 of them has your birthday = $1-(\frac{N-1}{N})^r$, now find r by trial and error such that this > 1/2 for N=365

approximation : same thing, $e^{-x} = 1-x$ for small x so $1-(\frac{N-1}{N})^r = 1-(1-\frac{1}{N})^r = 1-e^{-\frac{r}{N}}$ then set = 1/2, use logs.

31 BIRTHDAY problem: link the 2 above

- each person in personal birthday problems gives 1 chance to match your personal birthday
- each person in group shared birthday problem has r-1 chances to have same birthdate as someone else -> $r(r-1)/2$ pairs
- with the approximation set $n = r(r-1)/2$ so that both probs are equal

32 BIRTHDAY problem: whole group gets a holiday if one of them has a birthday. how large the group to have max number of working days = num people x num days worked ?

- Note: group of 1 : only gets 1 holiday per year -> 1 person works 364 days -> 1x364 days
- Note: group of 2 : each gets 2 holiday per year -> 2 people work 2 x 363 days
- Note: group massive : every one gets holidays whole time-> massive x 0 days = 0 days worked
- so max is in between somewhere [2 .. massive] - but where ?
- Answer : everyone works same number days (days when there are no birthday), so E(num days worked) = (num people) x E(num days worked by 1 person) = (num people) (num days in year) prob(day is working day) = nN [(not a birthday for person1) (not a birthday for person2)...(not a birthday for person n)]= $nN(1-\frac{1}{N})^n$ and find which n gives max for N=365
- note more rigorous : expected (num days offered by first working day for n people) = n x [(+1)xProb(+1)+(+0)xProb(+0)] = $n\left\{(+1)\left(1-\frac{1}{N}\right)^n+(+0)\left(1-(1-\frac{1}{N})^n\right)\right\}$
- so max xp^x, turns out x=N so $N^2(1-1/N)^N == N^2e^{-N/N} = N^2e^{-1}$ so just e^{-1} per worker per day with N workers and N days

33 RANDOM WALK ON INTEGER LINE - CLIFFHANGER: 1 step away from falling, prob p of stepping right=away

Note: possible paths: 1->0, 1->2->1->0, 1->2->1->2->1->0 etc

FSA Answer: FSA: prob (x=0|start=1) = prob (left)+prob (step right) prob (x=0|start=2)=$(1-p)P_1+pP_2$ but P_2=prob (x=0|start=2)=P[all paths that go from 2 to 0 somehow]=P[all paths that go from 2 to 1 somehow AND all paths that go from 1 to 0 somehow] = P[all paths that go from 2 to 1 - somehow] x P [all paths that go from 1 to 0 - somehow] = P_1P_1 (because markov and going form 1 to 0 somehow = going from 2 to 1 somehow) [Note: somehow means : in any way, maybe numerous steps like above] , anyway we get $P_1 = (1-p)+pP_1{}^2$ so P_1 is either 1 or $\frac{1-p}{p}$ and P must be <1 hence =1 for $p < 1/2$ and = $\frac{1-p}{p}$ for p>1/2

Note: for $p = 2/3$ this is $1/2$ and $p = 1/2$ this is 1 - implication for casino : $P_1 = 1/2$ to avoid bankruptcy requires you to have 2/3 chance of winning... tons of paths can lead you to x=0 / bankruptcy

FSA Answer general: $P_2 = P_1{}^2$ so $P_n = P_1{}^n$ so P(x=0|start m)= 1 or $(\frac{1-p}{p})^m$ **Martingale**

Answer: use de Moivre martinagle (p/q) ....

34 RANDOM WALK ON INTEGER LINE - GAMBLER RUIN: 2 players M,N play each other, winning odds: p(M win)=1/3<p(N win)=2/3, and initial fortune: M=1<N=2. they play +/-1 untill one bankrupt. what prob (M=ultimate winner) ? Answer: DOUBLE USAGE of the above

- M starts with m=1 and N starts with n=2. Random walk on integer, and When x=0 , then M bankrupt, when x=m+n then M win
- from above with unlimited resource ie $n = \infty$ then x=0 ie bankrupt with prob $\frac{q}{p}^m$
- but 2 ways to reach x= 0 = go bankrupt , (1) paths that always stay < m+n and (2) paths that reach m+n and then go back to 0. If let Q = paths that always stay < m+n, then if do reach m+n it is like a new start at x=m+n and we know already that these will go back to x=0 with prob $\frac{q}{p}^{m+n}$ ← double usage of the above
- so basically $\text{prob x= 0 with n=}\infty = \frac{q}{p}^m = Q+(1-Q)\frac{q}{p}^{m+n} \leftarrow \text{solve for Q}$ and P=prob of reaching m+n ever = 1 - prob stay below m+n = 1-Q
- answer = $P = \frac{1-\frac{q}{p}^m}{1-\frac{q}{p}^{m+n}}$ and with p=q=1/2 we get 0/0 and with l Hospital rule = $\frac{m}{m+n}$

35 start with 20 $ and want to reach $40: 20$ on evens in one go or bet 1$ on even each time ?

-Answer: bet 1$ each time = random walk on integer line same as above m=start=20, want extra n=20, prob p = 18/38 and so q=1-p=20/38 - so use formula above for 1$ at a time and compare with all in one go = p = 18/38

36 PRINCIPLE - drop n points on a circle - then all symmetric ie length of (n+1) segments have identical distributions

37 Mark sticks of length L with blue and red dots at the end , break sticks in 3 places, average length of blue-dot bit ?

Answer : equally distributed so L/3

38 52 cards, how many cards till first ace ?

4 aces -> 5 segments of 0–48 -> 48/5 cards + ace

39 German tanks 1..N, see n=60 so what is N?

Answer 1: 1 point dropped in 1..N -> your n creates 2 segments -> symmetry -> 2 segments of 1..59 + your obs -> 2x59+ the one you see = 2x59+1=119// Answer 2: prob of sample containing the max = (1 x (N-1)) / $\binom{N}{2} = 2/N$, max likelihood = make that prob biggest -> N=60 = as small as possible = what you see = n// Answer 3: confidence interval = P(n you see>N/x)=some value) eg P(n>2/3 N)=1/3 -> ???

40 German tanks 1..N, see 5 numbers n1...n5=60 so what is N?

Answer 1: 5 points -> 6 segments -> average length of (60-5)/5=11 -> N=60+11=71

41 break stick at random in 2 places, average length of smallest stick ?

Answer 1 break in 2 -> each segment average length L/2 by PRINCIPLE, now smallest = if break point left -> 1/2 x length of left = 1/2 x 1/2 = 1/4
Answer 2: Smaller stick has length x which is uniformly distributed in the range [0,0.5]. The probability distribution of x is p(x) = 1/(b-a) = 1 / [0.5-0] = 2. Yes it is a pdf, $\int_0^{0.5} 2dx = 1$. So E[x] = $\int xp(x)dx = \int_0^{0.5} x(2)dx = 2\left[\frac{x^2}{2}\right]_0^{0.5} = 0.25 = 1/4$.

42 break stick at random in 2 places, average length of ratio of lengths ?

assume fell on right then x [1/2 .. 1] uniform dist 1/(b-a)=2 so $2\int_{1/2}^{1}\frac{1-x}{x}dx = 2ln(2)-1$

43 break stick at random in 2 places, average length of smallest middle, largest ?

- from PRINCIPLE : break in 2 places -> 3 segments with uniform distribution -> average length of each segment = L/3
- however: so what ? we want SMALLEST and SMALLEST (3 uniforms) NOT uniform !! DIFFERENT from case above where the min was directly the distribution so was uniform
- Answer 0: complicated geometry methods with centroids
- Answer 1 FOR MIN:
• let z = min(x,y-x,1-y) = minimum length of the stick and P(z>a) = prob z > a
• P(z>a) = 2 * P(x>a, y-x>a, 1-y>a) = 2 P(x>a, a+x < y < 1-a) ⟵ 2 because have to consider case x < y and y < x – note MAX bit NOT entering yet thisis a general relation
• FIRST: a+x < y < 1-a → x < 1-2a so a < x < 1-2a → must have a < 1/3 ($Leftarrow$ obviously right , the min length will be less than 1/3 with 3 segments)
• SECOND: P(z>a) = 2 * 2 P(x>a, a+x < y < 1-a) → P(z>a) = 2 $\int_a^{1-2a}\int_{a+x}^{1-a} dydx$ (← y first, then x) ⇒ P(z>a) = 2 $\int_a^{1-2a}(1-2a-x)dx = 2((1-2a)(1-2a-a)-1/2((1-2a)^2-a^2) = 2(1-2a)(1-3a)-(1-3a)(1-a) = (1-3a)(2(1-2a)-(1-a) = (1-3a)(1-3a) = (1-3a)^2$
• CONCLUSION: P(z>a) = $(1-3a)^2$ so P(z<a) = 1 - P(z>a) = $1-(1-3a)^2$ so p(a) = derivative of $1-(1-3a)^2 = 6-18a = 6(1-3a)$ ← btw : g'∘f x f' = (-2)f x f' = (-2) (1-3a)(-3) = 6(1-3a) and E(a) = $\int ap(a)da = \int_0^{1/3} 6(1-3a)da$ ← again remember 0 < a < 1/3 from above – integration LIMITS !!
- Now Answer 1 (ie calculus) FOR MAX: similar but stick to P(z<a)
• let z = max(x,y-x,1-y)
• Pr(z < a) = Pr(x < a, y-x < a, 1-y < a) = Pr(1-2a < x < a, 1-a < y < x + a) ← EXACT SAME AS ABOVE – the MAX bit has not entered yet • max bit now : ???

44 your (P win) < P(lose) - How many $2n$ games to play to max your chances of winning?

• n+1 or more games in 2n to win so $P(\text{win in 2n}) = P_{2n} = \sum\binom{2n}{k>=n+1}p^k q^{n-k}$
• what 2n that maximises that sum?
• if go on to n+2 and not stop at n, it must be because it is better for you, ie P(win in 2n+2) > P(win in 2n)
•but 2n+2 games = 2n games + another 2
• in these 2 extra games your status changes in 2 cases only
• case 1: won n+1 in 2n but lose both 2 : then from winner -> loser and you have reduced the prob by (lose x lose x n+1 wins in
• case 2: won only n in 2n but win both 2 : then now from loser -> winner
• and explicitly $P_{2n+2} = P_{2n}$ - P(2 losses) P(won n+1 | 2 losses) + P(2 wins)P(won n | 2 wins) and note P(won n+1 | 2 losses) = P(won n+1) becasue it is independent and same for the other and note ht - becasue that possibility REDUCES your prob of win and the other INCREASES it
• and so explicitly $P_{2n+2} = P_{2n} - q^2\binom{n+1}{2n}p^{n+1}q^{n-1} + p^2\binom{2n}{n}p^n q^n$ ⟵ note the MINUS to REDUCE the prob of winning and the PLUS
• and for optimal 2n we want P(2n) >= P(2n-2) and also P(2n) >= P(2n+2)
• and for this to happen the added bits need to be positive/negative
etc...

45 MATCHING PROBLEM: 52 cards laid out, then another 52 cards laid out beneath - AVERAGE NUMBER of matches ?

• COUNTING problem -> INDICATOR variables
• each card: $1xP(match) + 0xP(not\ match) = 1x(1/n) + 0xwhatever = 1/n$ so n cards = n x 1/n = 1 (yes - don't care what comes before - E(sum of X) = sum of E(X) even if the Xs are dependent - so here what turns up IS dependent but the average is NOT dependent)

46 MATCHING PROBLEM: secretary puts letters in envelopes at random : SAME PROBLEM - SAME AVERAGE

47 MATCHING PROBLEM: 52 cards laid out, then another 52 cards laid out beneath - PROBABILITY of num matches=r (not JUST the AVERAGE over all the probs as per above) ? this is a DERANGEMENT problem

• P(r matches in n cards) = (prob 1st card matches) x (prob 2nd card matches) .. (prob rth card matches) x (choose r cards in n) x PROB(no match in n-r) $\to P(r|n) = \frac{1}{n}\frac{1}{n-1}\frac{1}{n-r+1}\binom{n}{r}P(0|n-r) = \frac{1}{r!}P(0|n-r)$
• P(0 matches in N cards) = num derangement of n cards / num card arrangements = num derangement of n cards / n!
• DERANGEMENT = (1st card not in place) AND (2nd card not in place) AND .. (nth card not in place) = $NumT1^c \cap T2^c \cap ... Tn^c = Num(\text{everything}) - Num(T1 \cup T2 \cup ... \cup Tn) = |S| - \sum_i |T_i| + \sum_{i<j}|T_i \cap T_j| - \sum_{i<j<k}|T_i \cap T_j \cap T_k| + \cdots + (-1)^n|T_1 \cap \cdots \cap T_n| = n! - \binom{n}{1}(n-1)! + \binom{n}{2}(n-2)! - \binom{n}{3}(n-3)! + \cdots + (-1)^n\binom{n}{n}(n-n)! == n! - \frac{n!}{1!} + \frac{n!}{2!} - \frac{n!}{3!} + \cdots + (-1)^n\frac{n!}{n!} = n!\left[1 - \frac{1}{1!} + \frac{1}{2!} - \frac{1}{3!} + \cdots + (-1)^n\frac{1}{n!}\right]$.
• so prob = $\frac{1}{r!}\text{Derangements}\frac{1}{n!} = \frac{1}{r!}\left[1 - \frac{1}{1!} + \frac{1}{2!} - \frac{1}{3!} + \cdots + (-1)^{n-r}\frac{1}{n-r!}\right]$ ← goes to 1/e 1/r!

48 DECIDE IF MAXIMUM part 1: review N slips 1 by 1, decide if biggest or move on, if decision = actual biggest, then get it else get 0 - BEST STRATEGY ?

• OPTIMAL POLICY : skip some , then select best so far : REASON: prob =k/n of having seen the max increases from 1/n to 1 but prob of winning decreases as we get near the end (and at the end it is just 1/n) so must be in the middle and any subsequent choice after your pick must be fails so whatever criterion you choose must be 'pick best so far and statement[k..n] is true if the true max is later' , so best policy is 'skip then pick'
• given optimal policy, the question is HOW MANY TO SKIP (s-1) before picking the sunsequent max so far (at k) :
• we want P(win with draw k = k is max so far) > P (win pick something after)
• P (win at k = k is the max) = 1/n
• P (max of the k-1 numbers seen so far is in the s-1 skipped – else we have lost) = (s-1)(k-1)
• hence P (picking the maximum with the skipping strategy) = $\frac{1}{n}\sum_{k=s}^{n-1}\frac{s-1}{k-1}$
• hence to choose best we want P(win with draw s = s is max so far) > P (win by picking something after s)
• ie $s/n > \frac{1}{n}\sum_{k=s}^{n-1}\frac{s-1}{k-1} = \frac{s}{n}\left[\frac{1}{s} + \frac{1}{s+1} + ... + \frac{1}{n-1}\right]$
• note limit: $\frac{s}{n}\left[\frac{1}{s} + \frac{1}{s+1} + ... + \frac{1}{n-1}\right] \equiv \frac{s}{n}ln(\frac{n}{s}) \to 1/e = s/n$ so cutoff $s \equiv n/e$

49 DECIDE IF MAXIMUM part 2: same as before : review N slips 1 by 1, decide if biggest or move on, if decision = actual biggest, then get it else get 0 BUT the number on the slip is uniformly distributed !- BEST STRATEGY NOW ?

• QUICK NOTE: before, did not know the dist of the dowry amounts. now we know the dist of the numbers. So say first number you get is 0.999, then prob (that this is > than next 99 numbers) = complement of (all number < 0.999) = 1 -{ (2nd num < .999) and (3rd num < .999) .. } = $1-(0.999)^{99}$
• on this problem, work backward:

- last draw : if biggest so far, then choose it (deterministic)
- one draw before last : only one number can beat it (the last one), so if > 1/2 choose it, if not leave it, because if < 1/2 then prob (last one > 1/2) is > 1/2
- 2 draws before last: then only next 2 draws can beat it. say current value is x then
  - both bigger : $(1-x)^2$
  - one bigger, one smaller : $x(1-x) + (1-x)x = 2x(1-x)$
  - both smaller: x^2
  - so prob of winning later is (CAREFUL) : (both bigger AND WE CHOOSE THE RIGHT ONE ← ie prob $\frac{1}{2}$) + (ONE BIGGER ← so we are sure to choose it) = $\frac{1}{2}(1-x)^2 + 2x(1-x)$
  - so chance of winning later = $\frac{1}{2}(1-x)^2 + 2x(1-x)$
  - chance of winning now = now is bigger than later 2 = (bigger than next one) and (bigger than last one) = $x.x = x^2$ ← uniform dist everywhere in this problem !
  - FINALLY : solve $x^2 = \frac{1}{2}(1-x)^2 + 2x(1-x)$ and choose candidate ie if bigger than solution
- n draws before last: same reasoning , end up solving $x^r = \binom{r}{1}x^{r-1}(1-x) + \frac{1}{2}\binom{r}{2}x^{r-2}(1-x)^2 + ... + \frac{1}{r}\binom{r}{r}x^{r-r}(1-x)^r$
- approx: r big and x > 1/2 (always wait for this) then 1-x gets small so x^r = first terma $= \binom{r}{1}x^{r-1}(1-x) \to x^r = rx^{r-1}(1-x) \to x = \frac{r}{r+1}$

50 DOUBLE ACCURACY - measurement with noise

- $D = A - B + d \quad S = A + B + s \quad A = \frac{1}{2}(D+S) - \frac{1}{2}(d+s)$ etc
- average of $\frac{1}{2}(d+s)$, variance of $\frac{1}{2}(d+s) = \frac{1}{4}(\sigma_s^2 + \sigma_d^2) = \frac{1}{2}\sigma^2$

51 Probability that $X^2 + 2bX + c$ has real roots with b,c uniform over whole area

• here the $b^2 - 4ac$ is $4(b^2 - c)$ so need $b^2 - c > 0$ so $b^2 > c$
so choose b so that $b^2 > c$
so choose b so that b outside of $b^2 = c$
• restrict b to [-B ... B] and find area outside $b^2 = c$
• then let B → ∞ • area inside $b = \sqrt{c}$ integrate $\int b = \int\sqrt{c} = \frac{x^{\frac{3}{2}}}{\frac{3}{2}}$ with the +/- = $2\frac{2}{3}x^{\frac{3}{2}} = \frac{4}{3}x^{\frac{3}{2}} = \frac{4}{3}B^{\frac{3}{2}}$
• whole area $4B^2$ so divide to get $\frac{1}{3\sqrt{B}}$
• NOTE : NOT SAME problem as $aX^2 + 2bX + c$ with a.b.c uniform because need divide by a and ratio of 2 uniforms not uniform !!

52 2D random walk

• in 2n moves, , just need n moves up (← and consequently n moves down) to return to origin, and this in both X and Y dimension • so in 2n moves, P (X=0) = $\binom{2n}{n}\frac{1}{2}^n\frac{1}{2}^n = \binom{2n}{n}\left(\frac{1}{2}\right)^{2n}$
• so in 2n x 2n moves, P(return at 0) = P (X=0)P (Y=0) = $\left[\binom{2n}{n}\left(\frac{1}{2}\right)^{2n}\right]^2$
• sterling approximation $n! = \sqrt{2\pi n}n^{n+1/2}e^{-n}$ so in 2n x 2n moves, P(return at 0) = $\frac{1}{\pi n}$ ← straight calc
• so number of returns at origin = P(return after 2) (counts for 1 return) + P(return after 4) (counts for 1 return) + ... (← for ever) = $\frac{1}{\pi}\sum\frac{1}{n} = \infty$
• so Q=0, P=1 since Q = prob of termination in $\text{mean} = \mu = \sum xP^xQ$ and we know form 1D case $\mu = 1/Q$

53 3D random walk

• same principle P (X=0)P (Y=0) = $\left[\binom{2n}{n}\left(\frac{1}{2}\right)^{2n}\right]^3$
• same thing in 2n x 2n x2n moves, P(return at 0) = $\frac{1}{(\pi n)^{3/2}}$ ← straight calc ← sum of this converges
• so this sum = mean number of returns = $\sum\frac{1}{(\pi n)^{3/2}} = \mu$
• same as above with terminating sequence $\text{mean} = \mu = \sum xP^xQ$ • num of succsses + 1 failure = 1/Q -> Q = 1/(num success + 1) (← always < 1 !!) and $P = 1 - Q$

54 Buffon SHORT needle of lengh $2l$ tossed over vertical lines $2a$ apart

• compute distance of centre of needle to nearest parallel given angle with parallel
• $lcos\theta$, assume θ uniform over $0..\frac{\pi}{2}$
• so average distance is $\frac{l\cos\theta}{a}$ with θ uniform over $0..\frac{\pi}{2}$
• so $\int_0^{\frac{\pi}{2}}\left(\frac{1}{\frac{\pi}{2}}\frac{l\cos\theta}{a}d\theta\right) = \frac{2l}{\pi a}\int_0^{\frac{\pi}{2}}\cos\theta d\theta = \frac{2l}{\pi a}[sin\theta]_0^{\frac{\pi}{2}} = \frac{2l}{\pi a}$

55 Buffon SHORT needle of lengh $2l$ tossed over vertical and horizontal lines $2a$ apart

56 Buffon LONG needle of lengh $2l$ tossed over vertical lines $2a$ apart