Ravi Dayabhai

Stat 110 Strategic Practice 1, Fall 2011

Prof. Joe Blitzstein (Department of Statistics, Harvard University)

1 Naive Definition of Probability (NDP)

1. For each part, decide whether the blank should be filled in with $=, <,$ or $>$, and give a short but clear explanation.

(a) (probability that the total after rolling 4 fair dice is 21) $>$ (probability that the total after rolling 4 fair dice is 22)

(b) (probability that a random 2 letter word is a palindrome[1]) $=$ (probability that a random 3 letter word is a palindrome)

(SEE BACK) 2. A random 5 card poker hand is dealt from a standard deck of cards. Find the probability of each of the following (in terms of binomial coefficients).

(a) A flush (all 5 cards being of the same suit; do not count a royal flush, which is a flush with an Ace, King, Queen, Jack, and 10)

(b) Two pair (e.g., two 3's, two 7's, and an Ace)

(SEE BACK) 3. (a) How many paths are there from the point $(0, 0)$ to the point $(110, 111)$ in the plane such that each step either consists of going one unit up or one unit to the right?

(b) How many paths are there from $(0, 0)$ to $(210, 211)$, where each step consists of going one unit up or one unit to the right, and the path has to go through $(110, 111)$?

(SEE BACK) 4. A *norepeatword* is a sequence of at least one (and possibly all) of the usual 26 letters a,b,c,...,z, with repetitions not allowed. For example, "course" is a norepeatword, but "statistics" is not. Order matters, e.g., "course" is not the same as "source".

A norepeatword is chosen randomly, with all norepeatwords equally likely. Show that the probability that it uses all 26 letters is very close to $1/e$. (NBD)

[1] A *palindrome* is an expression such as "A man, a plan, a canal: Panama" that reads the same backwards as forwards (ignoring spaces and punctuation). Assume for this problem that all words of the specified length are equally likely, that there are no spaces or punctuation, and that the alphabet consists of the lowercase letters a,b,...,z.

(1A) Probability space for all outcomes of sum of 4 die is symmetric. $\Rightarrow S = \{4, 5, 6, \ldots 23, 24\} = \{x \in \mathbb{Z} \mid 4 \le x \le 24\}$

$P(s \in S \mid s = x)$ increases for $x \to 15$ $(x \ge 4)$ and decreases for $x \to 24$ $(x \ge 15)$.

(2B.) 2-letter case

palindrome $\Leftrightarrow$ same letters

(NDP) $\frac{\cancel{26} \cdot 1}{26^{\cancel{2}}} = \frac{1}{26} = \frac{1}{26} = \frac{\cancel{26} \cdot \cancel{26} \cdot 1}{26^{\cancel{3}}}$

3-letter case

palindrome $\Leftrightarrow$ first letter = last letter

(2A) P(flush, but NOT royal flush) =

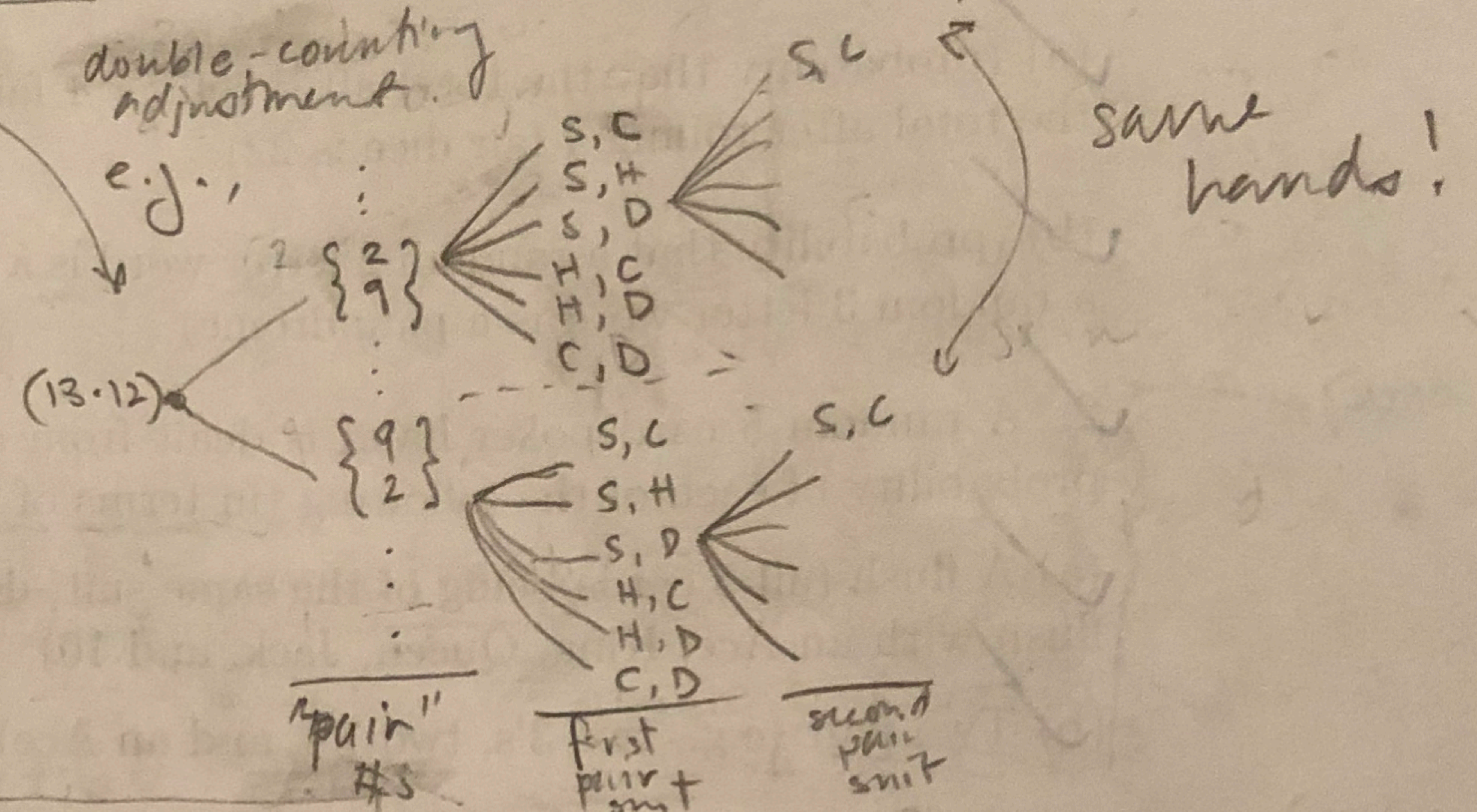
$$\underbrace{\frac{\overset{\text{chooses suit}}{\binom{4}{1}}\overset{\text{chooses number}}{\binom{13}{5}}}{\binom{52}{5}}}_{\text{all flushes}} - \underbrace{\frac{\binom{4}{1}}{\binom{52}{5}}}_{\text{royal flushes}} = \boxed{4\left(\frac{\binom{13}{5}-1}{\binom{52}{5}}\right)}$$

(2B) P(two pair) = $\boxed{\dfrac{13\binom{4}{2}\,12\binom{4}{2}\,11\binom{4}{1}}{(2!)\binom{52}{5}}}$

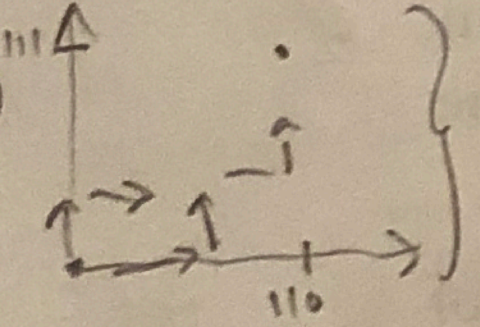
→ CHOOSE TWO VALUES : $\binom{13}{2} = \frac{13\cdot 12}{2!}$ double-counting adjustment e.g.,

→ CHOOSE TWO SUITS PER PAIR : $\binom{4}{2}\binom{4}{2}$

→ CHOOSE REMAINING CARD : $11\cdot\binom{4}{1} = 44$

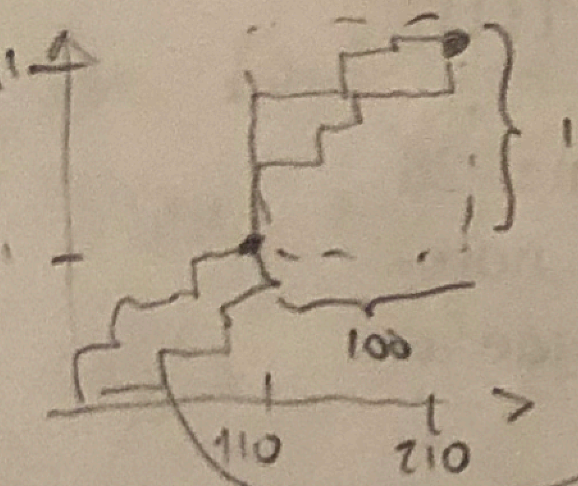


(3A)



how to arrange U's & R's? 110 Rs 111 Us $\boxed{\binom{110+111}{110}}$

(3B)



$\boxed{\binom{110+111}{110}\binom{100+100}{100}}$

(4) P(26 letters | non-repeat word) = $\dfrac{\overset{26!}{\#\text{ of 26-letter non-repeat words}}}{\#\text{ all non-repeat words}} \rightarrow \frac{1}{e}$ } SHOW.

$\sum_{k=1}^{26}\frac{26!}{(26-k)!}$ where $k \equiv$ length of word.

$$\frac{\frac{\cancel{26!}}{25!}+\frac{\cancel{26!}}{24!}+\dots+\frac{\cancel{26!}}{1!}+\frac{\cancel{26!}}{0!}}{\cancel{26!}} = \boxed{1+\frac{1}{1!}+\frac{1}{2!}+\dots+\frac{1}{26!}\approx e}$$

from Taylor series approximation of e^x at $x = 1$.

Imagine we have n switches that can be turned on or off. There are 2^n configurations of these switches (BITS). Alternatively, we can sum the number of ways k switches are turned on ($n-k$ are off) for all choices of k on-switches.

2 Story Proofs

5. Give a story proof that $\sum_{k=0}^{n}\binom{n}{k} = 2^n$.

6. Give a story proof that

$$\frac{(2n)!}{2^n \cdot n!} = (2n-1)(2n-3)\cdots 3 \cdot 1.$$

Partnerships (p.20 of text)

7. Show that for all positive integers n and k with $n \geq k$,

$$\binom{n}{k} + \binom{n}{k-1} = \binom{n+1}{k},$$

doing this in two ways: (a) algebraically and (b) with a "story", giving an interpretation for why both sides count the same thing. (SEE BACK

Hint for the "story" proof: imagine $n+1$ people, with one of them pre-designated as "president".

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

$$\binom{n+1}{k} = \frac{n+1}{k}\binom{n}{k-1} = \frac{n+1}{n+1-k}\binom{n}{k}, \quad \text{let } a = n+1$$

$$\frac{a}{k}\binom{n}{k-1} = \frac{a}{a-k}\binom{n}{k}$$

$$(a-k)(a)\binom{n}{k-1} = ka\binom{n}{k}$$

$$\binom{n}{k-1}\frac{a^{\cancel{2}}}{k} = ka\binom{n}{k} + ka\binom{n}{k-1} = \cancel{ka}\left[\binom{n}{k} + \binom{n}{k-1}\right]$$

$$\frac{a}{k}\binom{n}{k-1} = \binom{n}{k} + \binom{n}{k-1}, \quad \text{substitute } a = n+1$$

$$\frac{n+1}{k}\binom{n}{k-1} = \binom{n}{k} + \binom{n}{k-1}$$

from above

$$\boxed{\binom{n+1}{k} = \binom{n}{k} + \binom{n}{k-1}}$$

(TB) $\underbrace{\binom{n+1}{k}} = \underbrace{\binom{n}{k} + \binom{n}{k-1}}$; story proof.

# of ways to choose k people from group of n+1 people

Decomposes # ways of choosing k people from n+1 people where 1 person is president into 2 cases

- president <u>included</u> in selected group of k
- president <u>excluded</u> in selected group of k

HOMEWORK 1

① all birthorders = $\binom{6}{3} \cdot 3! \cdot 3! = 6! = \frac{6!}{\cancel{3!}\,\cancel{3!}} \cdot \cancel{3!} \cdot \cancel{3!}$ ✓

1 way to choose 3 girls first, 3 boys second (GGGBBB)

$$\frac{\binom{6}{1} \cancel{3!} \cdot \cancel{3!}}{\binom{6}{3} \cancel{3!}\,\cancel{3!}} = \boxed{\frac{3!\,3!}{6!}}$$

Stat 110 Homework 1, Fall 2011

Prof. Joe Blitzstein (Department of Statistics, Harvard University)

(SEE BACK OF PREVIOUS PAGE)

1. A certain family has 6 children, consisting of 3 boys and 3 girls. Assuming that all birth orders are equally likely, what is the probability that the 3 eldest children are the 3 girls?

(SEE BACK)

2. (a) How many ways are there to split a dozen people into 3 teams, where one team has 2 people, and the other two teams have 5 people each?

(b) How many ways are there to split a dozen people into 3 teams, where each team has 4 people?

(SEE BACK)

3. A college has 10 (non-overlapping) time slots for its courses, and blithely assigns courses to time slots randomly and independently. A student randomly chooses 3 of the courses to enroll in (for the PTP, to avoid getting fined). What is the probability that there is a conflict in the student's schedule?

(SEE BACK.)

ALMOST 4. A city with 6 districts has 6 robberies in a particular week. Assume the robberies are located randomly, with all possibilities for which robbery occurred where equally likely. What is the probability that some district had more than 1 robbery?

5. Elk dwell in a certain forest. There are N elk, of which a simple random sample of size n are captured and tagged ("simple random sample" means that all $\binom{N}{n}$ sets of n elk are equally likely). The captured elk are returned to the population, and then a new sample is drawn, this time with size m. This is an important method that is widely-used in ecology, known as *capture-recapture*.

What is the probability that exactly k of the m elk in the new sample were previously tagged? (Assume that an elk that was captured before doesn't become more or less likely to be captured again.)

(SEE BACK OF P. 2)

6. A jar contains r red balls and g green balls, where r and g are fixed positive integers. A ball is drawn from the jar randomly (with all possibilities equally likely), and then a second ball is drawn randomly.

(a) Explain intuitively why the probability of the second ball being green is the same as the probability of the first ball being green.

ALMOST (b) Define notation for the sample space of the problem, and use this to compute the probabilities from (a) and show that they are the same.

(c) Suppose that there are 16 balls in total, and that the probability that the two balls are the same color is the same as the probability that they are different colors. What are r and g (list all possibilities)?

$r+g=16$

(5) $$\frac{\cancel{\binom{N}{n}}\binom{n}{k}\binom{N-n}{m-k}}{\cancel{\binom{N}{n}}\binom{N}{m}}$$

fixes k elk of first sample

choose remaining m-k from untagged elk

# of ways of 1st sample

ways to run 2 samples of size n, m, resp.

(2A) 12 people, 3 teams of 2 : 5 : 5 ; how many ways?

CHOOSE THE TEAM of 2 → $\boxed{\dfrac{\binom{12}{2}\binom{10}{5}}{2}}$ choosing one team of 5 from 10 is the same as choosing the other 5.

(2B) 12 people, 3 evenly divided teams of 4 ; how many ways?

$$\boxed{\frac{12!}{4!\,4!\,4!\,3!}} = \boxed{\frac{\binom{12}{4}\binom{8}{4}\binom{4}{4}}{3!}} = \frac{12!}{\cancel{8}!\,4!}\cdot\frac{\cancel{8!}}{4!\,4!}$$

line of 12 people

order doesn't matter within teams

3! ways to arrange these teams

e.g.; 34 7 8 | 11 12 10 1 | 5 6 9 2

(3A) P(at least 1 conflict) = 1 − P(no conflict) ; ~~combination w/ replacement~~

$\underbrace{\text{P(no conflict)}}_{\#\text{ of ways w/ no conflict}}$

10 time slots

$\cancel{\binom{10+3-1}{3}}$ 10^3

$\binom{10}{3}$; $1 - \boxed{\dfrac{\binom{10}{3}}{\cancel{\binom{10+3-1}{3}}}}$

10^3

SEE WARNING ON p.18 of text (1.4.21)

$\underbrace{\cdot\cdot|\cdot|\cdot|\cdot|\cdot}_{\text{4 slots, 4 classes}}$ $\binom{7}{4}$

10 9 8

(4) 6 districts, 6 robberies; each robbery to have occurred in each district w/ same probability; P(some district had > 1 robbery) = 1 − P(some district had ≤ 1 robbery)

e.g.,

1		2,3		4,5,6		← robberies
A	B	C	D	E	F	← district

equally likely as

4	2	3	6	5	1
A	B	C	D	E	F

6! ways 1 robbery per district;

no way for a district to have 0 w/o another having > 1 !

$$1 - \frac{6!}{\boxed{\text{\# ways robberies could be distributed across districts}}} \quad 6^6$$

$\binom{n}{k}$ k=3, e.g.

youngest · oldest always in $\binom{k+1}{k}$ subgroups. · oldest

"hockey stick" sums

7. (a) Show using a story proof that

$$\binom{k}{k}+\binom{k+1}{k}+\binom{k+2}{k}+\cdots+\binom{n}{k}=\binom{n+1}{k+1},$$

where n and k are positive integers with $n \geq k$.
Hint: imagine arranging a group of people by age, and then think about the oldest person in a chosen subgroup.

(b) Suppose that a large pack of Haribo gummi bears can have anywhere between 30 and 50 gummi bears. There are 5 delicious flavors: pineapple (clear), raspberry (red), orange (orange), strawberry (green, mysteriously), and lemon (yellow). There are 0 non-delicious flavors. How many possibilities are there for the composition of such a pack of gummi bears? You can leave your answer in terms of a couple binomial coefficients, but not a sum of lots of binomial coefficients.

7A e.g. $\binom{3+1}{2+1}=\binom{4}{3}=4=\binom{2}{2}+\binom{2+1}{2}=1+3$

$\binom{4+1}{2+1}=\binom{5}{3}=10=\binom{2}{2}+\binom{2+1}{2}+\binom{2+2}{2}=1+3+6$

1
1 1
1 2 1
1 3 3 1
1 4 6 4 1
1 5 10 10 5 1

diagonal sums of PASCAL'S TRIANGLE!

The RHS counts the number of ways a given person of group size n+1 can be the oldest person in subgroup size k+1. This is the same as the LHS since when choosing k+1 people from n+1 people, someone has to be the oldest in the subgroup!

e.g. n=5, k=3, o's ordered by age, choices of k, • must be oldest

cases

o o o o o • $=\binom{5}{3}=\binom{n}{k}=\binom{k+2}{k}=10$

o o o o • o $=\binom{4}{3}=\binom{n-1}{k}=\binom{k+1}{k}=4$

o o o • o o $=\binom{3}{3}=\binom{n-2}{k}=\binom{k}{k}=1$

$15=\binom{6}{4}=\binom{n+1}{k+1}$

o o • o o o = 0

o • o o o o = 0

• o o o o o = 0

o b/c • can't be oldest in group of k+1 = 4

$\binom{33}{29}$ $\binom{34}{29}$ $\binom{34}{30}$ $\binom{35}{31}$ ⋱ $\binom{54}{50}$ $\binom{55}{50}$ Δ

7B $\sum_{k=30}^{50}\binom{5+k-1}{k}=\sum_{k=30}^{50}\binom{4+k}{k}$ ... but cannot express this way!

diagonal sum of Pascal triangle ⇒ "hockey stick" sums!

$\Delta=\boxed{\binom{55}{50}-\binom{34}{29}}$

$\frac{g(g-1)}{2}$ $\frac{(16-g)(16-g-1)}{2}$ $\frac{g(16-g)}{2}$

$\frac{(16-g)g}{2}$

2 cases: both green, both red

2 cases: red, green; green, red

6c) $r+g=16$; $r=16-g$ — $\underbrace{P(\text{same color, 2 drawn}) = P(\text{different color, 2 drawn})}$ $=\frac{1}{2}$; what are r, g

$$\frac{g(g-1)}{\cancel{2}} + \frac{(16-g)(15-g)}{\cancel{2}} = \frac{2(16-g)g}{\cancel{2}}$$

$$g^2 - 16g + 60 = 0$$

$$\boxed{g = 10 \text{ or } 6} \Leftrightarrow \boxed{r = 6 \text{ or } 10}$$

6A) This is because of symmetry. Say two green balls are labeled 1, 2. We could draw both at same time or 1,2 or 2,1 (sequentially). These are equivalent. The outcome we observe are just two green balls, so in effect, those are all equivalent outcomes.

6B) $S \equiv$ set of all outcomes from drawing 2 balls from jar containing $r+g$ balls.

$$\sum_{j=1}^{2} \binom{r}{j}\binom{g}{2-j} = |S| \text{ where } j = \#\text{ of red balls in outcome} = \binom{r+g}{2}$$ Vandermonde's identity.

OUTCOMES: where 1st ball is green VS OUTCOMES: where 2nd ball is green

$$\frac{\binom{g}{1}\binom{r+g-1}{1}}{|S|\cdot 2!} = \frac{g(r+g-1)}{|S|\cdot 2!} \boxed{=} \frac{\binom{r+g-1}{1}\binom{g}{1}}{|S|\cdot 2!}$$

fix the first ball = green

fix the second ball = green

e.g. G_1, G_2 can be (from A) arranged 2 ways; same for $R_1 G_2$, etc.