Lecture 6 10/3/16

(start Chap 2)

Counting # of elements in (large!) sets

1. **Multiplication rule**: Let n experiments be performed in order, and let $m_1, m_2, \dots, m_n$ be the # of possible outcomes of exp. $1, 2, \dots, n$ resp. Then there are $m_1 m_2 \cdots m_n$ possible outcomes for the ordered seq of n exp.

Ex1: A has 8 pairs of earrings, 7 dresses and 12 pairs of shoes.
How many possible outfits: $8 \cdot 7 \cdot 12$

Ex2: 30 students in class ⟶ how many seating configurations

$30!$

n Factorial: $n! = n(n-1)(n-2)\cdots 1 \quad (0! = 1)$
$n! = n \cdot (n-1)!$

2. **Permutations**

Ex: Soccer team of 16 ⟶ $\left\{ \begin{array}{l} 1 \text{ Pres} \\ 1 \text{ VP} \\ 1 \text{ Treas} \end{array} \right\}$.

$$16 \cdot 15 \cdot 14 = \frac{16!}{(16-3)!} = P_{16,3}$$

For $k \leq n$: $P_{n,k} = \dfrac{n!}{(n-k)!}$ ⟶ # of permutations of k objects from a set of n (ordered!)

3. **Combinations**

Ex: Soccer team of 16 ⟶ choose admin committee of 3 members

$\dfrac{P_{16,3}}{3!}$ ⟶ # of ways 3 people can be ordered

For $k \leq n$: $C_{n,k} = \dfrac{P_{n,k}}{k!} = \dfrac{n!}{(n-k)!\,k!}$ ⟶ # of combinations of k objects from a set of n (unordered!)

2.7 #5

#10 : how many 4 letter words: 1. no letter repeated twice $\to$ $P_{26,4}$ = # of 4 letter words with no repeated letters
2. must contain at least 1 vowel

Now, subtract those without vowels : $P_{21,4}$

Important properties

- For $k \leq n$: $C_{n,k} = \dfrac{n!}{(n-k)!\,k!} = C_{n,n-k}$ (Symmetry)

- Pascal's triangle

$$
\begin{array}{ccccccccc}
 & & & & 1 & & & & \\
 & & & 1 & & 1 & & & \\
 & & 1 & & 2 & & 1 & & \\
 & 1 & & \underset{C_{n-1,k-1}}{3} & & \underset{C_{n-1,k}}{3} & & 1 & \\
1 & & 4 & & \underset{C_{n,k}}{6} & & 4 & & 1
\end{array}
$$

$$C_{n,k} = C_{n-1,k-1} + C_{n-1,k}$$

- Binomial Thm: $(x+y)^n = \sum_{i=0}^{n} C_{n,i}\, x^i y^{n-i}$

Lecture 7 10/5/16

2.7 #15   10 G (L+R)   # of lineups?
12 C, F(L+R)

$P_{10,2} \cdot P_{12,3}$ (order matters!)

2.7 #13   12 F, 7 invites

(a) A, B feud and can't go both to party : $C_{12,7} - C_{10,5} = 540$

(b) A, B either together, or both don't go : $C_{10,5} + C_{10,7} = 372$

Multinomial

Recall : $C_{n,k}$ = # of ways to pick k objects from n (order does not matter)

Thus : 2 groups, i.e. - those selected (k from n)
- those not selected ((n-k) from n)

Q: What if we have multiple groups?

Ex: 30 students : 5 A's, 10 B's, 15 C's . How many ways

$C_{30,5}$ = # of ways to select the 5 A students from the class

$\to$ 25 left $\to$ $C_{25,10}$ ways to pick

$\Rightarrow$ Multiplication Rule : $C_{30,5} \cdot C_{25,10} = \frac{30!}{\cancel{(30-5)!}\,5!} \cdot \frac{\cancel{(30-5)!}}{(25-10)!\,10!} = \frac{30!}{5!\,10!\,15!}$

Thm: n objects, k groups with $n_1, n_2, \dots, n_k \in \{0, \dots, n\}$ but $n_1 + \dots + n_k = n$

$\Rightarrow \frac{n!}{n_1!\,n_2!\cdots n_k!}$ ways to distribute n objects over these k groups.

$\boxed{2.7 \# 21}$

Lecture 8 10/7/16

2.7 #27 (a) $4! = 24$

(b) From (a) there are 24 different ways to form 4 couples.
Each such group can marry in order in $4!$ ways $\Big\} = (24)\cdot 4! = 57\ldots$

(c) $\dfrac{8!}{2!2!2!2!} \Big/ 4! = 105$

(d) $\dfrac{8!}{2!2!2!2!} = 2520$

(e) $8!$

(f) Note: There a 4 fixed, married couples!
There are 4 M-F couples. These can be lined up in $4!$ ways.
Each couple can line up in 2 ways: (M,F) or (F,M)

$2^4 \cdot 4! = 384$

Binomial and Multinomial distributions

Do n experiments independently
Let p be the prob. of success in 1 experiment

X = RV for the # of successes in n experiments (So $X \in \{0,1,2,\dots,n\}$)

Then X is a Binomial RV (n,p) : $Bin(n,p)$

Distribution f'ion: $P(X=k) = C_{n,k}\, p^k (1-p)^{n-k}$, $k = 0,1,\dots,n$

Ex: Roll a die 12 times.
What is prob. to get 4 twos? At most 4 twos? More than 4 twos?

A: Roll a die once $\to$ prob. to get a two $= \frac{1}{6}$

Let X = RV for the # of twos in 12 rolls

$P(X>4) = 1 - P(X\le 4)$

$P(X=4) = C_{12,2}\left(\frac{1}{6}\right)^4\left(\frac{5}{6}\right)^8$

$P(X\le 4) = C_{12,0}\left(\frac{1}{6}\right)^0\left(\frac{5}{6}\right)^{12} + C_{12,1}\left(\frac{1}{6}\right)^1\left(\frac{5}{6}\right)^{11} + C_{12,2}\left(\frac{1}{6}\right)^2\left(\frac{5}{6}\right)^{10} + C_{12,3}\left(\frac{1}{6}\right)^3\left(\frac{5}{6}\right)^9 + C_{12,4}\left(\frac{1}{6}\right)^1\left(\frac{5}{6}\right)^8$

Thm : If $X \sim Bin(n, p)$, then $E(X) = np$
$Var(X) = np(1-p)$

(typos: $\circledast$ p 43, l 2 no m in Σ
$\circledast$ p 45, l 4 : "38/51" → "38/49 = 0.776")

2.7 # 29

33

Lecture 9 10/10/17

Thm: $X \sim Bin(n,p)$ has $E(X) = np$ and $Var(X) = np(1-p)$

Pf: Proof 1 (Short): Let X_1 be the RV for success in 1 try

Thus $\begin{cases} E(X_1) = 0\cdot(1-p) + 1\cdot(p) = p \\ Var(X_1) = E(X_1^2) - (E(X_1))^2 = 0^2\cdot(1-p) + 1^2\cdot(p) - p^2 = p - p^2 = p(1-p) \end{cases}$

Now: $X = X_1 + X_2 + \dots + X_n$, where the X_i are i.i.d.

$$\rightarrow \begin{cases} E(X) \overset{\text{proved later}}{=} E(X_1) + \dots + E(X_n) = p + \dots + p = np \\ Var(X) = Var(X_1) + \dots + Var(X_n) = p(1-p) + \dots + p(1-p) = np(1-p) \end{cases}$$

because the X_i are independent (proved later)

Proof 2 (longer):

$$E(X) = \sum_{k=0}^{n} k\, C_{n,k}\, p^k (1-p)^{n-k} = \sum_{k=1}^{n} k \frac{n!}{k!(n-k)!} p^k (1-p)^{n-k} = \sum_{k=1}^{n} \frac{n!}{(k-1)!(n-k)!} p^k (1-p)^{n-k}$$

KEY

$$= np \sum_{k=1}^{n} \frac{(n-1)!}{(k-1)!(n-1-(k-1))!} p^{k-1} (1-p)^{n-1-(k-1)} = np \sum_{l=0}^{n} \frac{(n-1)!}{l!(n-l)!} p^l (1-p)^{n-1-l} = np \sum_{l=0}^{n} C_{n-1,l}\, p^l (1-p)^{n-1-l}$$

$$= np \cdot 1 = np$$

$$Var(X) = E(X^2) - (E(X))^2 \overset{\text{trick}}{=} E(X(X-1)) + E(X) - (E(X))^2.$$ Since $E(X)$ known, let's compute $E(X(X-1))$

$$E(X(X-1)) = \sum_{k=0}^{n} k(k-1)\, C_{n,k}\, p^k (1-p)^{n-k} = \sum_{k=2}^{n} k(k-1) \frac{n!}{k!(n-k)!} p^k (1-p)^{n-k}$$

KEY

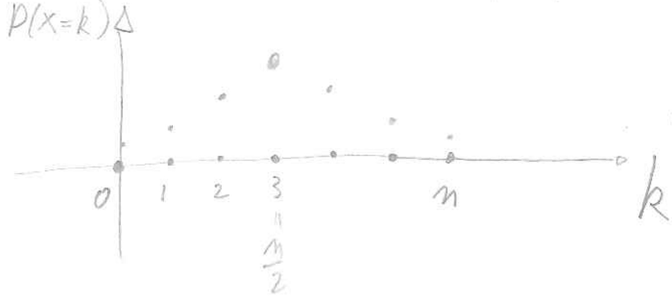
$$= n(n-1)p^2 \sum_{k=2}^{n} \frac{(n-2)!}{(k-2)!(n-2-(k-2))!} p^{k-2} (1-p)^{n-2-(k-2)}$$

$$= n(n-1)p^2 \sum_{l=0}^{n} \frac{(n-2)!}{l!(n-2-l)} p^l (1-p)^{n-2-l} = n(n-1)p^2 \sum_{l=0}^{n} C_{n-2,l}\, p^l (1-p)^{n-2-l}$$

$$= n(n-1)p^2$$

$$\rightarrow Var(X) = n(n-1)p^2 + np - (np)^2 = \cancel{n^2p^2} - np^2 + np - \cancel{(np)^2} = np(1-p)$$

Plot Bin$(n,p) \rightsquigarrow P(X=k) = C_{n,k}\, p^k (1-p)^{n-k}$



<u>Ex</u>: $p=\frac{1}{2} \to np = \frac{n}{2}$. Then $P(X=k) = \left(\frac{1}{2}\right)^n C_{n,k} = \left(\frac{1}{2}\right)^n \frac{n!}{(n-k)!\,k!}$ which is maximized (as a f^{ion} of k, for fixed n) when $k=\frac{n}{2}$, ie when $k=E(X)$.

Furthermore, since $C_{n,k} = C_{n,n-k}$ and $p=\frac{1}{2}$, the plot of the distribution function will be symmetrical around the expected value $\frac{n}{2}$

<u>EX</u>: What if $p < \frac{1}{2}$ or $> \frac{1}{2}$ (experiment at home, or online)

<u>EX 2.16, p44 in text</u>: An example of when <u>not</u> to use Bin!

Pick 13 cards from deck of 52.

Q: $P(1 \text{ ace}), P(2 \text{ aces}), P(3 \text{ aces}), P(4 \text{ aces})$

Tempting, but false A: Use Bin$(4, \frac{1}{4})$.

This is false because the 4 experiments (of success or unsuccessfully drawing aces) are not independent $\to$ Reason: There is no replacement

<u>k</u>	<u>Exact</u>	<u>Bin$(4, \frac{1}{4})$</u>
0	$\frac{C_{4,0}C_{48,13}}{C_{52,13}}$	$C_{4,0}\left(\frac{1}{4}\right)^0\left(\frac{3}{4}\right)^4$
1	$\frac{C_{4,1}C_{48,12}}{C_{52,13}}$	$C_{4,1}\left(\frac{1}{4}\right)^1\left(\frac{3}{4}\right)^3$
2	$\frac{C_{4,2}C_{48,11}}{C_{52,13}} \approx 0.2134$	$C_{4,2}\left(\frac{1}{4}\right)^2\left(\frac{3}{4}\right)^2 \approx 0.2109$
3	$\frac{C_{4,3}C_{48,10}}{C_{52,13}}$	$C_{4,2}\left(\frac{1}{4}\right)^3\left(\frac{3}{4}\right)^1$
4	$\frac{C_{4,4}C_{48,9}}{C_{52,13}}$	$C_{4,4}\left(\frac{1}{4}\right)^4\left(\frac{3}{4}\right)^0$

2.7 (#29)

33: 10 samples of beer
guess at least 8
(a) if guessing
(b) if $p=0.9$

Lecture 10 10/12/16

Ex 2.16 on p44

Draw 13 cards from deck of 52.

$$P(\text{ace of H is drawn}) = \frac{C_{51,12}}{C_{52,13}} = \frac{\frac{51.50\ldots.40}{12!}}{\frac{52.51\ldots.40}{13.12!}} = \frac{13}{52} = \frac{1}{4} = P(\text{ace of D}) = P(\text{ace of C}) = P(\text{ace of S})$$

But: These probabilities are not independent:

ex: $$P(\text{ace of H} \cap \text{ace of D}) = \frac{C_{50,11}}{C_{52,13}} = \frac{\frac{50\ldots40}{11!}}{\frac{52.51.50\ldots40}{13.12.11!}} = \frac{\cancel{13}.\cancel{12}^{3}}{\cancel{52}_{4}.51} = \frac{1}{17} \neq P(\text{ace of H})P(\text{ace of D}) = \frac{1}{16}$$

k: # of aces drawn	Exact	Bin$(4, \frac{1}{4})$
0	$\frac{C_{48,13}}{C_{52,13}} = 0.304$	$C_{4,0}\left(\frac{3}{4}\right)^4 = 0.316$
1	$\frac{C_{4,1}C_{48,12}}{C_{52,13}} = 0.439$	$C_{4,1}\left(\frac{1}{4}\right)\left(\frac{3}{4}\right)^3 = 0.422$
2	$\frac{C_{4,2}C_{48,11}}{C_{52,13}} = 0.213$	$C_{4,2}\left(\frac{1}{4}\right)^2\left(\frac{3}{4}\right)^2 = 0.211$
3	$\frac{C_{4,3}C_{48,10}}{C_{52,13}} = 0.041$	$C_{4,3}\left(\frac{1}{4}\right)^3\left(\frac{3}{4}\right) = 0.047$
4	$\frac{C_{4,4}C_{48,9}}{C_{52,13}} = 0.003$	$C_{4,4}\left(\frac{1}{4}\right)^4 = 0.004$

Compare pretty well!

(Ex 2.7 #33)

Remark: Multinomial distribution

↳ more than 2 categories, compared to the F, S in binomial

Ex: Experiment: height of person:
$P(H \leq 5'5'') = 30\%$
$P(5'5'' < H \leq 6') = 40\%$
$P(H > 6') = 30\%$

Measure 20 people: Prob that 5 people are $\leq 5'5''$, 10 between 5'5'' and 6', and that 5 people are $> 6'$:

$$= \frac{20!}{5!\,10!\,5!}(0.3)^5(0.4)^{10}(0.3)^5$$

Poisson distribution

$X \sim Pois(\lambda)$ if $P(X=k) = e^{-\lambda}\dfrac{\lambda^k}{k!}$, $k=0,1,2,\dots$
($\lambda>0$ is parameter)

Thm: If $X\sim Pois(\lambda)$, then $E(X)=\lambda=Var(X)$

Pf: • $E(X)=\sum\limits_{\substack{k=0\\1}}^{\infty} k\, e^{-\lambda}\dfrac{\lambda^k}{k!} = e^{-\lambda}\sum\limits_{k=1}^{\infty}\dfrac{\lambda^k}{(k-1)!} = e^{-\lambda}\cdot\lambda\sum\limits_{k=1}^{\infty}\dfrac{\lambda^{k-1}}{(k-1)!}$

$$\underset{l=k-1}{=} e^{-\lambda}\lambda\sum_{l=0}^{\infty}\frac{\lambda^l}{l!} = \cancel{e^{-\lambda}}\lambda\cdot\cancel{e^{\lambda}} = \lambda$$

• $Var(X)=E(X^2)-\left(E(X)\right)^2$ (1)

Will calculate $E(X(X-1))$, which is easier than $E(X^2)$.

Then: $E(X(X-1)) = E(X^2)-E(X)$, hence $E(X^2)=E(X(X-1))+E(X)$ (2)

$E(X(X-1))=\sum\limits_{\substack{k=0\\2}}^{\infty} k(k-1)\,e^{-\lambda}\dfrac{\lambda^k}{k!} = e^{-\lambda}\sum\limits_{k=2}^{\infty}\dfrac{\lambda^k}{(k-2)!} = e^{-\lambda}\lambda^2\sum\limits_{k=2}^{\infty}\dfrac{\lambda^{k-2}}{(k-2)!}$

$$\underset{l=k-2}{=} e^{-\lambda}\lambda^2\sum_{l=0}^{\infty}\frac{\lambda^l}{l!} = \cancel{e^{-\lambda}}\lambda^2\cancel{e^{\lambda}} = \lambda^2 \quad (3)$$

$\overset{(1),(2),(3)}{\Longrightarrow}$ $Var(X) = \cancel{\lambda^2}+\lambda-\cancel{\lambda^2}=\lambda$.

An approximation result:

Link between $Bin(n, p)$ and $Pois(\lambda)$

<u>Thm 2.4</u>: If $X_n \sim Bin(n, p_n)$ such that $\left.\begin{array}{l} p_n \to 0 \\ \text{and} \\ np_n \to \lambda \\ (\text{as } n \to \infty) \end{array}\right\}$. Then $P(X_n = k) \to e^{-\lambda}\frac{\lambda^k}{k!}$ for all $k = 0, 1, \ldots, n$, as $n \to \infty$

<u>Ex 2.7 #38</u> $P(\text{at least 1 double 6 in 24}) = 1 - P(\text{no double 6 in 24})$

$$= 1 - \left(\frac{35}{36}\right)^{24} \simeq \underline{0.4914} \text{ (exact)}$$

Poisson approx. $= 1 - P(X=0)$, where $X \sim Pois\left(\lambda = np = \frac{24}{36} = \frac{2}{3}\right)$

$$= 1 - e^{-\frac{2}{3}} \cdot \frac{\left(\frac{2}{3}\right)^0}{0!} = \underline{\underline{0.4866}}$$

pretty close, as expected

Lecture 11 10/14/17

Sec. 2.7 # 52
56
57
61

Capture - recapture experiments (Ex 2.42 on p61)

Q: A population with unknown number of individuals.
Can we estimate this number based on a capture-recapture experiment:

So (k, r, m) are known.

- day 1: capture k individuals and mark them all
- day 2: capture r individuals of which m turn out to be marked

A: Main idea is maximum likelihood estimate:

If there were N individuals, then the probability of capturing m marked individuals on day 2 is:

$$P_N = \frac{C_{k,m}\, C_{N-k,\, r-m}}{C_{N,r}}$$

Think of P_N as a function of N (k, r, m are known), and find that value of N which maximizes P_N.

So let's maximize P_N: First note $C_{p,q} = \frac{p!}{(p-q)!\,q!} = \frac{p \cdot (p-1)!}{(p-q)(p-1-q)!\,q!}$ $= C_{p-1,q}$

$$\text{So } C_{p,q} = \frac{p}{p-q} C_{p-1,q}$$

$$\text{Thus, } P_N = \frac{C_{k,m} \frac{N-k}{N-k-(r-m)} C_{N-1-k,\, r-m}}{\frac{N}{N-r} C_{N-1,r}} = P_{N-1} \frac{N-k}{N-k-(r-m)} \cdot \frac{N-r}{N}$$

$\Rightarrow P_N \geq P_{N-1}$ iff $(N-k)(N-r) \geq (N-k-(r-m))N$ iff $N \leq \frac{kr}{m}$.

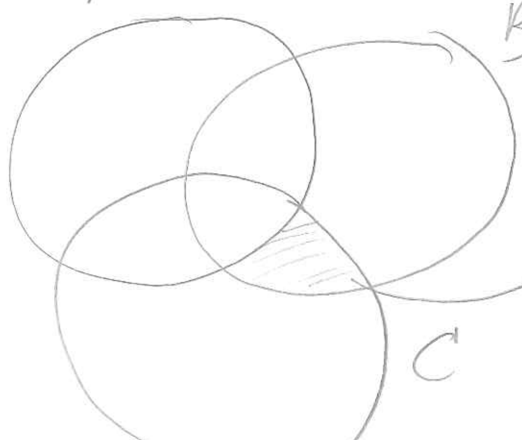
$\Rightarrow P_N$ is maximized for $\boxed{N = \lfloor \frac{kr}{m} \rfloor}$ Do numerical example

Lecture 12   10/17/16

- Finish Capture - Recapture experiment from last class.
- Inclusion - Exclusion generalized
  - ↳ For 2 events: $P(A \cup B) = P(A) + P(B) - P(A \cap B)$

For 3 events: $P(A \cup B \cup C) = P(A) + P(B) + P(C) - (P(A \cap B) + P(A \cap C) + P(B \cap C)) + P(A \cap B \cap C)$

Pf



There are 7 parts to $A \cup B \cup C$. Check how often each of these are counted in expression above.

$A \cup B \cup C$ is disjoint union of:	A	B	C	$A \cap B$	$A \cap C$	$B \cap C$	$A \cap B \cap C$	Net Count
$A \cap B \cap C$	+1	+1	+1	−1	−1	−1	+1	+1
$A \cap B \cap C^c$								
$A \cap C \cap B^c$								
$B \cap C \cap A^c$								
$A \cap (B \cup C)^c$								
$B \cap (A \cap C)^c$								
$C \cap (A \cap B)^c$								

- This counting principle can be extended to n events

$$P\left(\bigcup_{i=1}^{n} A_i\right) = \sum_{i=1}^{n} P(A_i) - \sum_{i<j}^{n} P(A_i \cap A_j) + \sum_{i<j<k}^{n} P(A_i \cap A_j \cap A_k) - \cdots + (-1)^{n+1} P\left(\bigcap_{i=1}^{n} A_i\right)$$

2.7 # 77, 79