Stat 100a, Introduction to Probability.

Outline for the day:

1. Binomial random variables.
2. Variance of sums.
3. Geometric random variables.
4. Negative binomial random variables.
5. Moment generating functions.
6. Poisson random variables.
7. Sample ch5 problems.
8. Harman and Negreanu.
9. Continuous random variables.
10. Uniform random variables.
11. Bruin, teams, and exam 1 questions.

Read through chapter 6.

HW2 is due tomorrow morning. Exam 2 on Mon.

http://www.stat.ucla.edu/~frederic/100A/sum21 ♠ ♣ ♥ ♦

1. Binomial Random Variables, ch. 5.2.

Suppose now X = # of times something with prob. p occurs, out of n independent trials

Then X = *Binomial (n.p)*.

e.g. the number of pocket pairs, out of 10 hands.

Now X could = 0, 1, 2, 3, …, or n.

pmf: P(X = k) = choose(n, k) * $p^k q^{n-k}$.

Key idea: X = $Y_1 + Y_2 + \ldots + Y_n$, where the Y_i are independent and *Bernoulli* (p).

If X is Bernoulli (p), then μ = p, and $\sigma = \sqrt{(pq)}$.

If X is Binomial (n,p), then μ = np, and $\sigma = \sqrt{(npq)}$.

2. Variance of sums and binomial random variables, ch. 5.2.

If X = # of times something with prob. p occurs, out of n independent trials, then X = *Binomial (n.p)*.

For example, the number of pocket pairs out of 10 hands is binomial(10, 5.88%).

When X is binomial(n,p), $X = Y_1 + Y_2 + \ldots + Y_n$, where the Y_i are independent and *Bernoulli* (p).

Fact about variance. If X_i are independent, then $Var(X_1+...+X_n) = Var(X_i) + ... + Var(X_n)$.

If X is Bernoulli (p), then μ = p, and var(X) = pq, so $\sigma = \sqrt{(pq)}$.

If X is Binomial (n,p), then μ = np, and var(X) = npq, so $\sigma = \sqrt{(npq)}$.

3. Geometric random variables, ch 5.3.

Suppose now X = # of trials until the *first* occurrence.

(Again, each trial is independent, and each time the probability of an occurrence is p.)

Then X = *Geometric (p)*.

e.g. the number of hands til you get your next pocket pair.

[*Including the hand where you get the pocket pair. If you get it right away, then X = 1.*]

Now X could be 1, 2, 3, …, up to ∞.

pmf: $P(X = k) = p^1 q^{k-1}$.

e.g. say k=5: $P(X = 5) = p^1 q^4$. Why? Must be 0 0 0 0 1. Prob. = q * q * q * q * p.

If X is Geometric (p), then $\mu = 1/p$, and $\sigma = (\sqrt{q}) \div p$.

e.g. Suppose X = the number of hands til your next pocket pair. P(X = 12)? E(X)? σ?

X = Geometric (5.88%).

$P(X = 12) = p^1 q^{11}$ = 0.0588 * 0.9412 ^ 11 = **3.02%.**

E(X) = 1/p = **17.0**. σ = sqrt(0.9412) / 0.0588 = **16.5**.

So, you'd typically *expect* it to take 17 hands til your next pair, +/- around 16.5 hands.

4. Negative binomial random variables, ch5.4.

Recall: if each trial is independent, and each time the probability of an occurrence is p, and X = # of trials until the *first* occurrence, then:

X is Geometric (p), $P(X = k) = p^1 q^{k-1}$, $\mu = 1/p$, $\sigma = (\sqrt{q}) \div p$.

Suppose now X = # of trials until the *rth* occurrence.

Then X = *negative binomial (r,p)*.

e.g. the number of hands you have to play til you've gotten r=3 pocket pairs.

Now X could be 3, 4, 5, …, up to ∞.

pmf: $P(X = k) = \text{choose}(k-1, r-1)\, p^r q^{k-r}$, for k = r, r+1, ….

e.g. say r=3 & k=7: $P(X = 7) = \text{choose}(6,2)\, p^3 q^4$.

Why? Out of the first 6 hands, there must be exactly r-1 = 2 pairs. Then pair on 7th.

P(exactly 2 pairs on first 6 hands) = $\text{choose}(6,2)\, p^2 q^4$. P(pair on 7th) = p.

If X is negative binomial (r,p), then $\mu = r/p$, and $\sigma = [\sqrt{(rq)}] \div p$.

e.g. Suppose X = the number of hands til your 12th pocket pair. P(X = 100)? E(X)? σ?

X = Neg. binomial (12, 5.88%).

$P(X = 100) = \text{choose}(99,11)\, p^{12} q^{88}$

= choose(99,11) * 0.0588 ^ 12 * 0.9412 ^ 88 = **0.104%.**

E(X) = r/p = 12/0.0588 ~ **204**. σ = sqrt(12*0.9412) / 0.0588 = **57.2**.

So, you'd typically *expect* it to take 204 hands til your 12th pair, +/- around 57.2 hands.

5. Moment generating functions, ch. 4.7

Suppose X is a random variable. $E(X)$, $E(X^2)$, $E(X^3)$, etc. are the *moments* of X.

$ø_X(t) = E(e^{tX})$ is called the *moment generating function* of X.

Take derivatives with respect to t of $ø_X(t)$ and evaluate at t=0 to get moments of X.

1st derivative $(d/dt)\, e^{tX} = X\, e^{tX}$, $(d/dt)^2\, e^{tX} = X^2\, e^{tX}$, etc.

$(d/dt)^k\, E(e^{tX}) = E[(d/dt)^k\, e^{tX}] = E[X^k\, e^{tX}]$, (see p.84)

so $ø'_X(0) = E[X^1\, e^{0X}] = E(X)$,

$ø''_X(0) = E[X^2\, e^{0X}] = E(X^2)$, etc.

The moment gen. function $ø_X(t)$ uniquely characterizes the distribution of X.

So to show that X is, say, Poisson, you just need to show that it has the moment generating function of a Poisson random variable.

Also, if X_i are random variables with cdfs F_i, and $ø_{X_i}(t) \to ø(t)$, where $ø_X(t)$ is the moment generating function of X which has cdf F, then $X_i \to X$ in distribution, i.e. $F_i(y) \to F(y)$ for all y where F(y) is continuous, see p85.

Moment generating functions, continued.

$ø_X(t) = E(e^{tX})$ is called the *moment generating function* of X. $e \sim 2.72$.

Suppose X is Bernoulli (0.4). What is $ø_X(t)$?

$E(e^{tX}) = (0.6)(e^{t(0)}) + (0.4)(e^{t(1)}) = 0.6 + 0.4 e^t$.

Suppose X is Bernoulli (0.4) and Y is Bernoulli (0.7) and X and Y are independent.

What is the distribution of XY?

$ø_{XY}(t) = E(e^{tXY}) = P(XY=0)(e^{t(0)}) + P(XY=1)(e^{t(1)})$

$= P(X=0 \text{ or } Y=0)(1) + P(X=1 \text{ and } Y=1)e^t$

$= [1 - P(X=1)P(Y=1)] + P(X=1)P(Y=1)e^t$

$= [1 - 0.4 \times 0.7] + 0.4 \times 0.7 e^t$

$= 0.72 + 0.28e^t$, which is the moment generating function of a Bernoulli (0.28) random variable. Therefore XY is Bernoulli (0.28).

What about $Z = \min\{X,Y\}$?

Z = XY in this case, since X and Y are 0 or 1, so the answer is the same.

Suppose X = 0 with probability ½, 1 with probability ¼, 2 with probability 1/8, and 3 with probability 1/8.
What is E(X)? What is $E(X^2)$? What is Var(X)? What is SD(X)? What is $\phi_X(t)$?

$E(X) = 0(1/2) + 1(1/4) + 2(1/8) + 3(1/8) = 0.875.$

$E(X^2) = 0(1/2) + 1(1/4) + 4(1/8) + 9(1/8) = 1.875.$

$Var(X) = E(X^2) - \mu^2 = 1.875 – 0.875^2 = 1.11.$

$SD(X) = \sqrt{1.11} = 1.05.$

$\phi_X(t) = E(e^{tX}) = ½ (1) + ¼ (e^t) + 1/8 (e^{2t}) + 1/8 (e^{3t}).$

6. Poisson random variables, ch 5.5.

Player 1 plays in a very slow game, 4 hands an hour, and she decides to do a big bluff whenever the second hand on her watch, at the start of the deal, is in some predetermined 15 second interval, i.e. with probability ¼.

Now suppose Player 2 plays in a game where about 10 hands are dealt per hour, so he similarly looks at his watch at the beginning of each poker hand, but only does a big bluff if the second hand is in a 6 second interval, i.e. with probability 1/10.

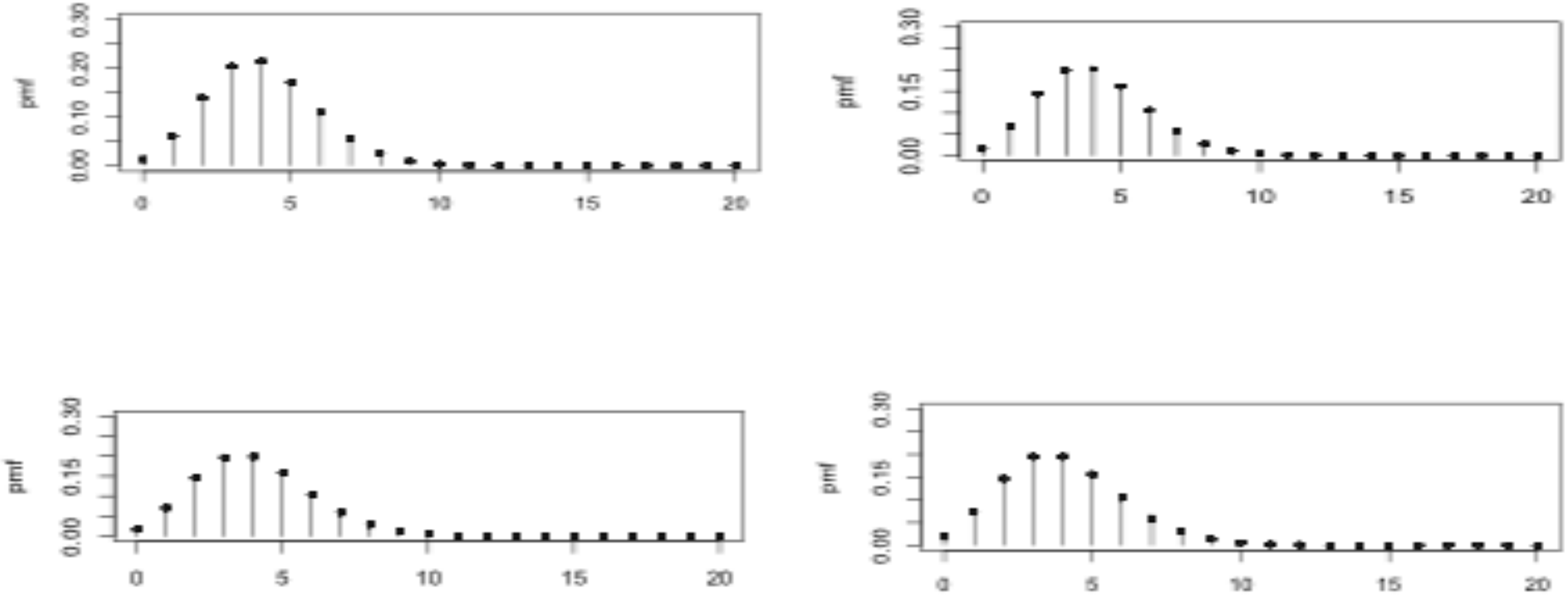
Player 3 plays in a faster game where about 20 hands are dealt per hour, and she bluffs only when the second hand on her watch at the start of the deal is in a 3 second interval, with probability 1/20. Each of the three players will thus average one bluff every hour.

Let X_1, X_2, and X_3 denote the number of big bluffs attempted in a given 4 hour interval by Player 1, Player 2, and Player 3, respectively.

Each of these random variables is binomial with an expected value of 4, and a variance approaching 4.

They are converging toward some limiting distribution, and that limiting distribution is called the *Poisson* distribution.

They are converging toward some limiting distribution, and that limiting distribution is called the *Poisson* distribution. Unlike the binomial distribution which depends on two parameters, n and p, the Poisson distribution depends only on one parameter, λ, which is called the *rate*. In this example, $\lambda = 4$.



The pmf of the Poisson random variable is $f(k) = e^{-\lambda}\lambda^k/k!$, for $k=0,1,2,...$, and for $\lambda > 0$, with the convention that 0!=1, and where $e = 2.71828\ldots$.
The Poisson random variable is the limit in distribution of the binomial distribution as $n \to \infty$ while np is held constant.

For a Poisson(λ) random variable X, $E(X) = \lambda$, and $Var(X) = \lambda$ also. $\lambda = rate$.

Example. Suppose in a certain casino jackpot hands are defined so that they tend to occur about once every 50,000 hands on average. If the casino deals approximately 10,000 hands per day, **a)** what are the expected value and standard deviation of the number of jackpot hands dealt in a 7 day period? **b)** How close are the answers using the binomial distribution and the Poisson approximation? Using the Poisson model, if X represents the number of jackpot hands dealt over this week, what are **c)** $P(X = 5)$ and **d)** $P(X = 5 \mid X > 1)$?

Answer. It is reasonable to assume that the outcomes on different hands are iid,
iid = independent and identically distributed,
and this applies to jackpot hands as well. In a 7 day period, approximately 70,000 hands are dealt, so X = the number of occurrences of jackpot hands is binomial$(n=70,000, p=1/50,000)$. Thus **a)** $E(X) = np = 1.4$, and $SD(X) = \sqrt{(npq)} = \sqrt{(70,000 \times 1/50,000 \times 49,999/50,000)} \sim 1.183204$. **b)** Using the Poisson approximation, $E(X) = \lambda = np = 1.4$, and $SD(X) = \sqrt{\lambda} \sim 1.183216$. The Poisson model is a very close approximation in this case. Using the Poisson model with rate $\lambda = 1.4$,

c) $P(X=5) = e^{-1.4}\, 1.4^5/5! \sim 1.105\%$.

d) $P(X = 5 \mid X > 1) = P(X = 5 \text{ and } X > 1) \div P(X > 1) = P(X = 5) \div P(X>1) = [e^{-1.4}\ 1.4^5/5!] \div [1 - e^{-1.4}\ 1.4^0/0! - e^{-1.4}\, 1.4^1/1!] \sim 2.71\%$.

7. Sample ch5 problems.

Let X = the # of hands until your 1st pair of black aces. What are E(X) and SD(X)?

X is geometric(p), where p = 1/C(52,2) = 1/1326.

E(X) = 1/p = 1326.

SD = (√q) / p, where q = 1325/1326. SD = 1325.5.

What is P(X = 12)?

$q^{11}p$ = 0.0748%.

You play 100 hands. Let X = the # of hands where you have 2 black aces. What is E(X)? What is P(X = 4)?

X is binomial(100,p), where p = 1/1326.

E(X) = np = .0754.

P(X = 4) = C(100,4) p^4 q^{96} = .000118%.

8. Harman and Negreanu, and running it twice.

Harman has 10♠ 7♠ . Negreanu has K♥ Q♥ . The flop is 10◆ 7♣ K◆ .

Harman's all-in. $156,100 pot. P(Negreanu wins) = 28.69%. P(Harman wins) = 71.31%.

Let X = amount Harman has after the hand.

If they run it once, E(X) = $0 x 29% + $156,100 x 71.31% = **$111,314.90.**

If they run it twice, what is E(X)?

There's some probability p_1 that Harman wins both times ==> X = $156,100.

There's some probability p_2 that they each win one ==> X = $78,050.

There's some probability p_3 that Negreanu wins both ==> X = $0.

E(X) = $156,100 x p_1 + $78,050 x p_2 + $0 x p_3.

If the different runs were *independent,* then p_1 = P(Harman wins 1st run & 2nd run) would = P(Harman wins 1st run) x P(Harman wins 2nd run) = 71.31% x 71.31% ~ 50.85%.

But, they're not quite independent! Very hard to compute p_1 and p_2.

However, you don't need p_1 and p_2 !

X = the amount Harman gets from the 1st run + amount she gets from 2nd run, so

E(X) = E(amount Harman gets from 1st run) + E(amount she gets from 2nd run)

= $78,050 x P(Harman wins 1st run) + $0 x P(Harman loses first run)

+ $78,050 x P(Harman wins 2nd run) + $0 x P(Harman loses 2nd run)

= $78,050 x 71.31% + $0 x 28.69% + $78,050 x 71.31% + $0 x 28.69% = **$111,314.90.**

HAND RECAP Harman 10♠ 7♠ Negreanu K♥ Q♥ The flop is 10♦ 7♣ K♦ .

Harman's all-in. $156,100 pot.P(Negreanu wins) = 28.69%. P(Harman wins) = 71.31%.

The standard deviation (SD) changes a lot! **Say they run it once.** (see p127.)

$V(X) = E(X^2) - \mu^2$.

μ = $111,314.9, so μ^2 ~ $12.3 billion.

$E(X^2)$ = ($156,100²)(71.31%) + (0²)(28.69%) = $17.3 billion.

V(X) = $17.3 billion - $12.3 bill. = $5.09 billion. SD σ = sqrt($5.09 billion)~$71,400.

So if they run it once, Harman expects to get back about $111,314.9 **+/- $71,400**.

If they run it twice? Hard to compute, but approximately, if each run were independent, then $V(X_1+X_2) = V(X_1) + V(X_2)$,

so if X_1 = amount she gets back on 1st run, and X_2 = amount she gets from 2nd run,

then $V(X_1+X_2) \sim V(X_1) + V(X_2)$ ~ $1.25 billion + $1.25 billion = $2.5 billion,

The standard deviation σ = sqrt($2.5 billion) ~ $50,000.

So if they run it twice, Harman expects to get back about $111,314.9 **+/- $50,000.**

9. Continuous random variables and their densities, ch6.1.

Density (or pdf = Probability Density Function) f(y):

$\int_B f(y)\, dy = P(X \text{ in } B)$.

If F(c) is the cumulative distribution function, i.e. $F(c) = P(X \leq c)$,

then $f(c) = F'(c)$.

The survivor function is $S(c) = P(X > c) = 1 - F(c)$.

If X is a continuous rv, then $P(X \leq a) = P(X < a)$, because $P(X = a) = \int_a^a f(y)dy = 0$.

Expected value, $\mu = E(X) = \int y\, f(y)\, dy$. (= $\sum y\, P(y)$ for discrete X.)

For any function g, $E(g(X)) = \int g(y)\, f(y)\, dy$. For instance $E(X^2) = \int y^2 f(y)dy$.

Variance, $\sigma^2 = V(X) = Var(X) = E(X-\mu)^2 = E(X^2) - \mu^2$.

$SD(X) = \sqrt{V(X)}$.

For examples of pictures of pdfs, see p104, 106, and 107.

10. Uniform example.

Recall for a continuous random variable X,
the pdf f(y) is a function where $\int_a^b f(y)dy = P\{X \text{ is in } (a,b)\}$,
$E(X) = \mu = \int_{-\infty}^{\infty} y\, f(y)dy$,
and $\sigma^2 = Var(X) = E(X^2) - \mu^2$. $sd(X) = \sigma$.
If X is uniform(a,b), then f(y) = 1/(b-a) for y in (a,b), and y = 0 otherwise.

For example, suppose X and Y are independent uniform random variables on (0,1), and Z = *min*(X,Y). **a)** Find the pdf of Z. **b)** Find E(Z). **c)** Find SD(Z).

a. For c in (0,1), $P(Z > c) = P(X > c \ \& \ Y > c) = P(X > c)\, P(Y > c) = (1-c)^2 = 1 - 2c + c^2$.
So, $P(Z \le c) = 1 - (1 - 2c + c^2) = 2c - c^2$.
Thus, $\int_0^c f(c)dc = 2c - c^2$. So f(c) = the derivative of $2c - c^2 = 2 - 2c$, for c in (0,1).
Obviously, f(c) = 0 for all other c.
b. $E(Z) = \int_{-\infty}^{\infty} y\, f(y)dy = \int_0^1 c\,(2-2c)\, dc = \int_0^1 2c - 2c^2\, dc = c^2 - 2c^3/3]_{c=0}^{1}$
$= 1 - 2/3 - (0 - 0) = 1/3$.
c. $E(Z^2) = \int_{-\infty}^{\infty} y^2 f(y)dy = \int_0^1 c^2\,(2-2c)\, dc = \int_0^1 2c^2 - 2c^3\, dc = 2c^3/3 - 2c^4/4]_{c=0}^{1}$
$= 2/3 - 1/2 - (0 - 0) = 1/6$.
So, $\sigma^2 = Var(Z) = E(Z^2) - [E(Z)]^2 = 1/6 - (1/3)^2 = 1/18$.
$SD(Z) = \sigma = \sqrt{(1/18)} \sim 0.2357$.

11. Teams, emails, and bruin.
The project is problem 8.2, page 249.
You need to write code to go all in or fold.
For instance, if your function is named "bruin", you might do:

```
bruin = function (numattable, cards, board, round, currentbet, mychips, pot, roundbets,
      blinds, chips, ind, dealer, tablesleft) {
## all in with any pair higher than 7s, or if lower card is J or higher,
## or if you have less than 3 times the big blind
a = 0
if ((cards[1, 1] == cards[2, 1]) && (cards[1, 1] > 6.5)) a = mychips
if (cards[2,1] > 10.5) a = mychips
if(mychips < 3*blinds) a = mychips
a
} ## end of bruin
```

cards[1,1] is your higher card (2-14).
cards[2,1] is your lower card (2-14).
cards[1,2] and cards[2,2] are suits of your higher card & lower card.

Here is bruin again.

```
bruin = function (numattable, cards, board, round, currentbet, mychips, pot, roundbets, blinds, chips, ind, dealer,
        tablesleft) {
## all in with any pair higher than 7s, or if lower card is J or higher, or if you have less than 3 times the big blind
a = 0
if ((cards[1, 1] == cards[2, 1]) && (cards[1, 1] > 6.5)) a = mychips
if (cards[2,1] > 10.5) a = mychips
if(mychips < 3*blinds) a = mychips
a
} ## end of bruin
```

The top 3 finishers get 13,8,5 points respectively. 1000 chips. Blinds start at 10/20 and increase by 50% every 10 hands.
numattable = number of players at your table.
cards[1,1] = number of higher card, cards[2,1] = lower card. cards[1,2] and cards[2,2] are their suits.
currentbet = the maximum bet so far this betting round.
mychips is how many chips you have left.
pot = the number of chips in the pot.
blinds = the amount of the big blind. The small blind will be half as much. Both are rounded to integers.
chips is a vector indicating how many chips everyone at your table has left.
ind is your seat number at the table.
dealer is the seat number of the dealer at your table.
tablesleft is how many tables of up to 10 players each are left in the tournament.
Ignore board, round, and roundbets because you have to be all in or fold before the flop.

```
Iveybruin = function (numattable, cards, board, round, currentbet,mychips,
      pot, roundbets, blinds, chips, ind, dealer,tablesleft) {
## all in with any 9 pair or higher, or if lower card is 10 or higher,
## or if I have less than 3 times the big blind
a = 0
if ((cards[1, 1] == cards[2, 1]) && (cards[1, 1] > 8.5)) a = mychips
if (cards[2,1] > 9.5) a = mychips
if(mychips < 3*blinds) a = mychips
a
} ## end of Iveybruin
```

```
acj = function(numattable, cards, board, round, currentbet, mychips, pot, roundbets, blinds, chips, ind, dealer,
tablesleft){
        # all in with pocket pair of Queens, Kings, or Aces, or AK of any suits
        # all in with AQ, AJ, A10, KQ if same suit and no one else is all in yet
        # all in with any pocket pair if only 1-2 players left to play and nobody is all in yet
        # all in if chip count less than 3 times the big blind with any cards
        a = 0
        if((cards[1,1] == cards[2,1]) && (cards[1,1] >= 12)) a = mychips
        if((cards[1,1] == 14) && (cards[2,1] == 13)) a = mychips
        if(currentbet <= blinds){
                if((cards[1,1] == 14) && (cards[2,1] >= 10) && (cards[1,2] == cards[2,2])) a = mychips
                if((cards[1,1] == 13) && (cards[2,1] == 12) && (cards[1,2] == cards[2,2]))a = mychips
                }
        big.blind = dealer + 2
        if(big.blind > numattable) big.blind <- big.blind - numattable
        players.left = big.blind - ind
        if(players.left < 0) players.left = players.left + numattable
```

```
acj = function(numattable, cards, board, round, currentbet, mychips, pot, roundbets, blinds, chips, ind, dealer, tablesleft){
# all in with any pocket pair if only 1-2 players left to play and nobody is all in yet
# all in if chip count less than 3 times the big blind with any cards
                    if(currentbet <= blinds){
                              if((players.left <= 2) && (cards[1,1] == cards[2,1])) a = mychips
                    }
                    if(mychips < 3 * blinds) a = mychips
          a
}
```

I randomly assigned people to teams Wed, for students who were here. For the others, I randomly assigned you to teams over the weekend.

The teams are

1. Shally Li. Andra Velea. Kevin Wang.
2. Adreama Islam. Ilayda Karsidag.
3. Omer Demirkan. Haley Hallman.
4. Kaili Nguyen. Nuo Chen.
5. Daniel Cruz. Tushar Roy.
6. Yuxi Chang. Yuetong Xu.
7. Christopher Clark. Xingbo Zhou.
8. Laiyin Dai. Joy Richardson.
9. Joseph Shen. Bakur Madini. Ruiqi Xin.
10. Hong Chung. Jiaqi Li.
11. Jason Sung. Oscar Yen. Tiffany Zhou.
12. William Fann. Christy Hui.
13. Najia Pan. Sahen Rai. Dominick Won.
14. Jaime Perez. Daniel Zhou.

Maybe you can exchange emails now in the chat if your partner is here, or if not, email me and I will give you the email address of your partner.