Stat 100a, Introduction to Probability.

Outline for the day:

1. Binomial random variables, continued, and variance of sum.
2. Poisson random variables.
3. Moment generating functions.
4. Examples and review.
5. Harman/Negreanu and running it twice.

There is no lecture or OH Thu Oct24.

Exam1 is Tue Oct29.

There is no lecture Tue Nov5 because of the faculty meeting.

Read through chapter 5.

Homework 2 is on the course website. It is due Thu Nov7.

Your emails are in the bottom of teams.txt.

http://www.stat.ucla.edu/~frederic/100a/F19 ♠ ♣ ♥ ♦

1. 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 $\mu = p$, and var(X) = pq, so $\sigma = \sqrt{(pq)}$.

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

2. 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.

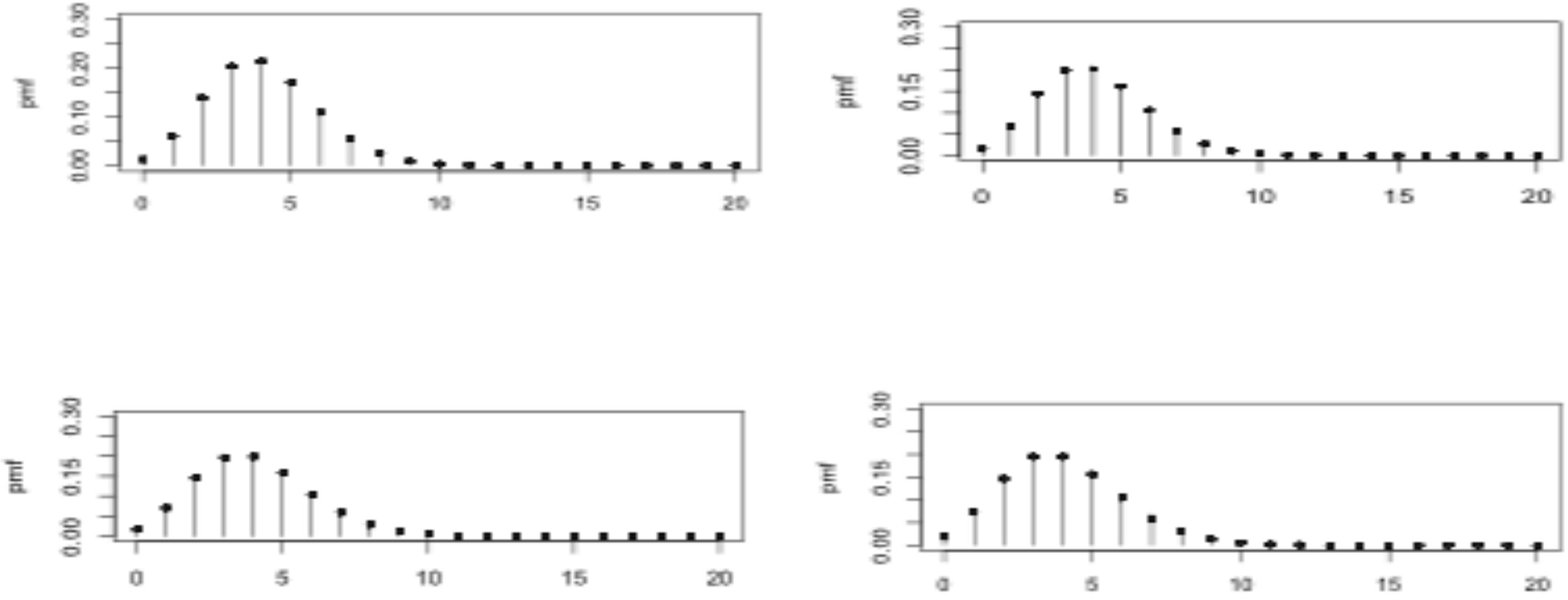
Player 3 plays in a faster game where about 20 hands are dealt per hour, and she bluffs only when the second hand on her watch at the start of the deal is in a 3 second interval, with probability 1/20. Each of the three players will thus average one bluff every hour.

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 $e = 2.71828...$.
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, 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) $P(X = 5 \mid X > 1) = P(X = 5 \text{ and } X > 1) \div P(X > 1) = P(X = 5) \div P(X>1) = [e^{-1.4}\ 1.4^5/5!] \div [1 - e^{-1.4}\ 1.4^0/0! - e^{-1.4}\, 1.4^1/1!] \sim 2.71\%$.

3. 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.

Moment generating functions, continued.

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

Suppose X is Bernoulli (0.4). What is $\text{ø}_X(t)$?

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

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

What is the distribution of XY?

$\text{ø}_{XY}(t) = E(e^{tXY}) = P(XY=0)\,(e^{t(0)}) + P(XY=1)(e^{t(1)})$

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

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

$= [1 – 0.4 \text{ x } 0.7] + 0.4\text{x}0.7e^t$

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

Therefore XY is Bernoulli (0.28).

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

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

4. Review.

1) Basic principles of counting.
2) Axioms of probability, and addition rule.
3) Permutations & combinations.
4) Conditional probability.
5) Independence.
6) Multiplication rules. $P(AB) = P(A)\,P(B|A)$ [= $P(A)P(B)$ if ind.]
7) Odds ratios.
8) Discrete RVs and probability mass function (pmf).
9) Expected value.
10) Pot odds calculations.
11) Luck and skill.
12) Variance and SD.
13) $E(aX+b)$, $E(X+Y)$, $V(X+Y)$.
14) Bayes's rule.
15) Markov and Chebyshev inequalities.
16) Bernoulli, binomial, geometric, negative binomial, and Poisson rvs.
17) Moment generating functions.

We have basically done all of chapters 1-5.

on the midterm, FOR EACH QUESTION, WRITE THE LETTER OF YOUR ANSWER TO THE LEFT OF THE QUESTION.

Example problems.

___ 1. What is the probability that you will be dealt a king and another card of the same suit as the king?

a. 1.69%. b. 3.62%. c. 4.89%. d. 5.02%. e. None of the above.

4 * 12 / C(52,2) = 3.62%.

P(flop an ace high flush)? [where the ace might be on the board]

-- 4 suits

-- one of the cards must be an ace. C(12,4) possibilities for the others.

So P(flop ace high flush) = 4 * C(12,4) / C(52,5)

= 0.0762%, or 1 in **1313**.

P(flop a straight | 87 of different suits in your hand)?

It could be 456, 569, 6910, or 910J. Each has 4*4*4 = 64 suit combinations.

So P(flop a straight | 87) = 64 * 4 / C(50,3)

= 1.31%.

P(flop a straight | 86 of different suits in your hand)?

Now it could be 457, 579, or 7910.

P(flop a straight | 86) = 64 * 3 / C(50,3)

= 0.980%.

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 = ($\sqrt{q}$) / 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%.

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 $ø_X(t)$?

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) = $E(X^2) - \mu^2 = 1.875^2 – 0.875^2 = 2.75$.

SD(X) = $\sqrt{2.75}$ = 1.66.

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

P(flop two pairs).

If you're sure to be all-in next hand, what is P(you will flop two pairs)?

This is a tricky one. Don't double-count (4♠ 4♦ 9♠ 9♦ Q♦) and (9♠ 9♦ 4♠ 4♦ Q♦).

There are choose(13,2) possibilities for the NUMBERS of the two pairs.

For each such choice (such as 4 & 9),

there are choose(4,2) choices for the suits of the lower pair,

and the same for the suits of the higher pair.

So, choose(13,2) * choose(4,2) * choose(4,2) different possibilities for the two pairs.

For each such choice, there are 44 [52 - 8 = 44] different possibilities for your fifth card, so that it's not a full house but simply two pairs. So,

P(flop two pairs) = C(13,2) * C(4,2) * C(4,2) * 44 / C(52,5)

~ 4.75%, or 1 in **21**.

P(flop two pairs).

Here is another way to do it. Find the mistake.

P(flop 2 pairs) = P(pocket pair and flop 2 pairs) + P(no pocket pair and flop 2 pairs)

= P(pocket pair) P(flop 2 pairs | pocket pair) + P(no pocket pair) P(flop 2 pairs | no pocket pair)

= P(pocket pair aa) * P(bbc | aa) + P(ab)*P(abc | ab)

= 13*C(4,2)/C(52,2) * 12*C(4,2)*44/C(50,3) + C(13,2)*4*4/C(52,2) * 3*3*44/C(50,3)

= 2.85%.

What is the problem here?

P(flop two pairs).

Here is another way to do it. Find the mistake.

P(flop 2 pairs) = P(pocket pair and flop 2 pairs) + P(no pocket pair and flop 2 pairs)

= P(pocket pair) P(flop 2 pairs | pocket pair) + P(no pocket pair) P(flop 2 pairs | no pocket pair)

= P(pocket pair aa) * P(bbc | aa) + P(ab)*P(abc | ab)

= 13*C(4,2)/C(52,2) * 12*C(4,2)*44/C(50,3) + C(13,2)*4*4/C(52,2) * **3*3*44**/C(50,3)

= 2.85%.

What is the problem here?

P(flop 2 pairs | no pocket pair) ≠ P(ab)*P(abc | ab). If you have ab, it could come acc or bcc on the flop.

13*C(4,2)/C(52,2) * 12*C(4,2)*44/C(50,3) + C(13,2)*4*4/C(52,2) * **(3*3*44 + 6*11*C(4,2))** /C(50,3)

= 4.75%.

5. Harman / 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.**

<u>If they run it twice, what is E(X)?</u>

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! **<u>Say they run it once</u>.** (see p127.)

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

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

$E(X^2)$ = ($$156,100^2$)(71.31%) + ($0^2$)(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**.

<u>If they run it twice?</u> 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.**