Stat 100a, Introduction to Probability.

Outline for the day:

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

There is no lecture Thu Nov 1!

HW2 is due Nov6. The midterm is Tue Nov 6 in class.

http://www.stat.ucla.edu/~frederic/100a/F18 .

♠ ♣ ♥ ♦

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

e.g. say n=10, k=3: $P(X = 3) = \text{choose}(10,3) * p^3 q^7$.

Why? Could have 1 1 1 0 0 0 0 0 0 0, or 1 0 1 1 0 0 0 0 0 0, etc.

choose(10, 3) choices of places to put the 1's, and for each the prob. is $p^3 q^7$.

Key idea: $X = Y_1 + Y_2 + \ldots + Y_n$, 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 $\sigma = \sqrt{(pq)}$.

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

Binomial Random Variables, continued.

Suppose X = the number of pocket pairs you get in the next 100 hands.

<u>What's P(X = 4)? What's E(X)? σ? Var(X)?</u>   X = Binomial (100, 5.88%).

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

So, $P(X = 4) = \text{choose}(100, 4) * 0.0588^4 * 0.9412^{96} = 13.9\%$, or 1 in **7.2.**

$E(X) = np = 100 * 0.0588 =$ **5.88**. $\sigma = \sqrt{(100 * 0.0588 * 0.9412)} =$ **2.35**.

So, out of 100 hands, you'd *typically* get about 5.88 pocket pairs, +/- around 2.35.

$Var(X) = \sigma^2 = 100 * 0.0588 * 0.9412 = 5.53$.

1. Geometric random variables, ch 5.3.

Suppose now X = # of trials until the *<u>first</u>* 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 $\sigma = (\sqrt{q}) \div p$.

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

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.

2. 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.

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.

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

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

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

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

What is the distribution of XY?

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

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

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

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

$= 0.72 + 0.28e^t$, 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. 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.

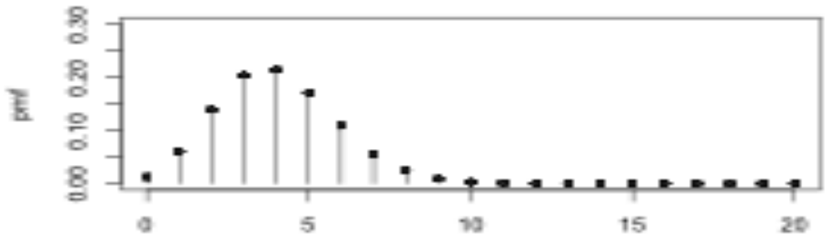
Player 3 plays in a faster game where about 20 hands are dealt per hour, and she bluffs only when the second hand on her watch at the start of the deal is in a 3 second interval, with probability 1/20. Each of the three players will thus average one bluff every hour.

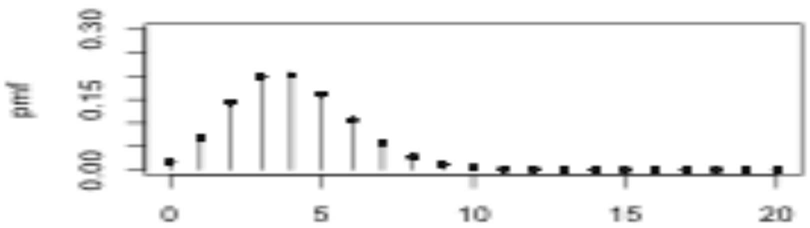
Let X_1, X_2, and X_3 denote the number of big bluffs attempted in a given 4 hour interval by Player 1, Player 2, and Player 3, respectively.

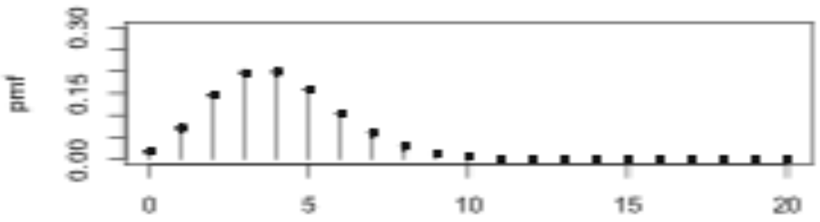
Each of these random variables is binomial with an expected value of 4, and a variance approaching 4.

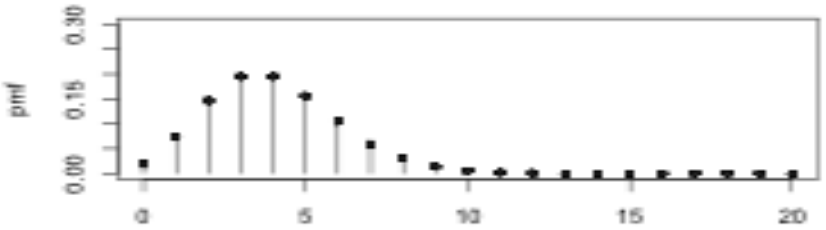
They are converging toward some limiting distribution, and that limiting distribution is called the *Poisson* distribution.

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

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! **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 $V(X_1+X_2) \sim V(X_1) + V(X_2) \sim$ \$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.**