Stat 100a: Introduction to Probability.

<u>Outline for the day:</u>

1. Hand in HW3.
2. Avoiding zero.
3. Chip proportions and induction.
4. Doubling up.
5. Bivariate densities.
6. Review list.
7. Examples.

No class Mon. Final exam is Wed from 10am to 11:30am. It's open book and open note. Bring a calculator and a pen or pencil.

1. Hand in HW3. **2. Avoiding zero.**

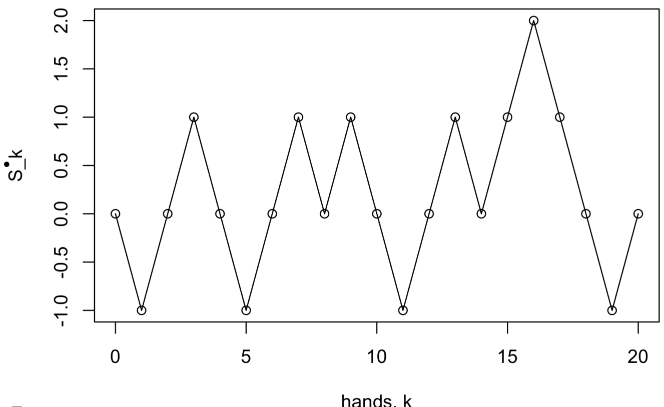
For a simple random walk, for any even # n, $P(S_1 \neq 0, S_2 \neq 0, \ldots, S_n \neq 0) = P(S_n = 0)$.

Proof. The number of paths from (0,0) to (n, j) that don't touch the x-axis at positive times

= the number of paths from (1,1) to (n,j) that don't touch the x-axis at positive times

= paths from (1,1) to (n,j) - paths from (1,-1) to (n,j) by the *reflection principle*

$= N_{n-1,j-1} - N_{n-1,j+1}$



Let $Q_{n,j} = P(S_n = j)$.

$P(S_1 > 0, S_2 > 0, \ldots, S_{n-1} > 0, S_n = j) = ½[Q_{n-1,j-1} - Q_{n-1,j+1}]$.

Summing from j = 2 to ∞,

$P(S_1 > 0, S_2 > 0, \ldots, S_{n-1} > 0, S_n > 0)$

$= ½[Q_{n-1,1} - Q_{n-1,3}] + ½[Q_{n-1,3} - Q_{n-1,5}] + ½[Q_{n-1,5} - Q_{n-1,7}] + \ldots$

$= (1/2)\ Q_{n-1,1}$

$= (1/2)\ P(S_n = 0)$, because to end up at (n, 0), you have to be at (n-1,+/-1),

so $P(S_n = 0) = (1/2)\ Q_{n-1,1} + (1/2)\ Q_{n-1,-1} = Q_{n-1,1}$.

By the same argument, $P(S_1 < 0, S_2 < 0, \ldots, S_{n-1} < 0, S_n < 0) = (1/2)\ P(S_n = 0)$.

So, $P(S_1 \neq 0, S_2 \neq 0, \ldots, S_n \neq 0) = P(S_n = 0)$.

3. Chip proportions and induction, Theorem 7.6.6.

P(win a tournament) is proportional to your number of chips.

Simplified scenario. Suppose you either go up or down 1 each hand, with prob. 1/2.

Suppose there are n chips, and you have k of them.

Let p_k = P(win tournament given k chips) = P(random walk goes k -> n before hitting 0).

Now, clearly $p_0 = 0$. Consider p_1. From 1, you will either go to 0 or 2.

So, $p_1 = 1/2\ p_0 + 1/2\ p_2 = 1/2\ p_2$. That is, $p_2 = 2\ p_1$.

We have shown that $p_j = j\ p_1$, for j = 0, 1, and 2.

***(induction:)* Suppose that, for j = 0, 1, 2, …, m, $p_j = j\ p_1$.**

We will show that $p_{m+1} = (m+1)\ p_1$.

Therefore, $p_j = j\ p_1$ for all j.

That is, P(win the tournament) is prop. to your number of chips.

$p_m = 1/2\ p_{m-1} + 1/2\ p_{m+1}$. If $p_j = j\ p_1$ for $j \leq m$, then we have

$mp_1 = 1/2\ (m-1)p_1 + 1/2\ p_{m+1}$,

so $p_{m+1} = 2mp_1 - (m-1)\ p_1 = (m+1)p_1$.

4. Doubling up. Again, P(winning) = your proportion of chips.
Theorem 7.6.7, p152, describes another simplified scenario.
Suppose you either double each hand you play, or go to zero, each with probability 1/2.

Again, P(win a tournament) is prop. to your number of chips.

Again, $p_0 = 0$, and $p_1 = 1/2\ p_2 = 1/2\ p_2$, so again, $p_2 = 2\ p_1$.

We have shown that, for j = 0, 1, and 2, $p_j = j\ p_1$.

(induction:) **Suppose that, for $j \leq m$, $p_j = j\ p_1$.**

We will show that $p_{2m} = (2m)\ p_1$.

Therefore, $p_j = j\ p_1$ for all $j = 2^k$. That is, P(win the tournament) is prop. to # of chips.

This time, $p_m = 1/2\ p_0 + 1/2\ p_{2m}$. If $p_j = j\ p_1$ for $j \leq m$, then we have

$mp_1 = 0 + 1/2\ p_{2m}$, so $p_{2m} = 2mp_1$. Done.

In Theorem 7.6.8, p152, you have k of the n chips in play. Each hand, you gain 1 with prob. p, or lose 1 with prob. q=1-p.

Suppose $0<p<1$ and $p \neq 0.5$. Let $r = q/p$. Then P(you win the tournament) = $\mathbf{(1-r^k)/(1-r^n)}$.

The proof is again by induction, and is similar to the proof we did of Theorem 7.6.6.

5. Bivariate density.

Suppose X and Y are random variables.

If X and Y are discrete, we can define the joint pmf $f(x,y) = P(X = x \text{ and } Y = y)$.

Suppose X and Y are continuous for the rest of this page.

Define the bivariate or joint pdf $f(x,y)$ as a function with the properties that $f(x,y) \geq 0$, and for any a,b,c,d,

$$P(a \leq X \leq b \text{ and } c \leq Y \leq d) = \int_a^b \int_c^d f(x,y)\, dx\, dy.$$

The integral $\int_{-\infty}^{\infty} f(x,y)\, dy = f(x)$, the pdf of X, and this is sometimes called the *marginal* density of X.

$E(X) = \int_{-\infty}^{\infty} x\, f(x)\, dx = \int_{-\infty}^{\infty} x\, [\int_{-\infty}^{\infty} f(x,y)\, dy]\, dx = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x\, f(x,y)\, dx\, dy.$

If X and Y are independent, then $f(x,y) = f_x(x) f_y(y)$.

Now $E(XY) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} xy\, f(x,y)\, dx\, dy$, and

$$\begin{aligned} \text{cov}(X,Y) &= E(XY) - E(X)E(Y) \\ &= E[(X - \mu_x)(Y - \mu_y)] \\ &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} [x - E(X)][y - E(Y)]\, f(x,y)\, dx\, dy. \end{aligned}$$

6. Review list.

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) Random variables (RVs).
9) Discrete RVs, and probability mass function (pmf).
10) Expected value.
11) Pot odds calculations.
12) Luck, skill, and deal-making.
13) Variance and SD.
14) Bernoulli RV. [0-1. $\mu = p$, $\sigma = \sqrt{(pq)}$.]
15) Binomial RV. [# of successes, out of n tries. $\mu = np$, $\sigma = \sqrt{(npq)}$.]
16) Geometric RV. [# of tries til 1st success. $\mu = 1/p$, $\sigma = (\sqrt{q}) / p$.]
17) Negative binomial RV. [# of tries til rth success. $\mu = r/p$, $\sigma = (\sqrt{rq}) / p$.]
18) Poisson RV [# of successes in some time interval. [$\mu = \lambda$, $\sigma = \sqrt{\lambda}$.]
19) E(X+Y), V(X+Y), covariance, and correlation. (ch. 7.1).
20) Bayes's rule (ch. 3.4).
20) Continuous RVs
21) Probability density function (pdf)
22) Uniform, normal, exponential, and Pareto RVs.
23) Moment generating functions
24) Markov and Chebyshev inequalities
25) Law of Large Numbers (LLN)
26) Central Limit Theorem (CLT)
27) Conditional expectation.
28) Confidence intervals for the sample mean.
29) Fundamental theorem of poker
30) Random Walks, Reflection Principle, Ballot Theorem, avoiding zero
31) Chip proportions and induction.
32) Bivariate densities.

We've done all of ch. 1-7 except 6.7.

7. Examples.

(Chen and Ankenman, 2006). Suppose that a $100 winner-take-all tournament has $1024 = 2^{10}$ players. So, you need to double up 10 times to win. Winner gets $102,400.

Suppose you have probability $p = 0.54$ to double up, instead of 0.5.

What is your expected profit in the tournament? (Assume only doubling up.)

Answer. P(winning) = 0.54^{10}, so exp. return = 0.54^{10} ($102,400) = $215.89. So exp. profit = $115.89.

What if each player starts with 10 chips, and you gain a chip with $p = 54\%$ and lose a chip with $p = 46\%$? What is your expected profit?

Answer. $r = q/p = .46/.54 = .852$. P(you win) = $(1\text{-}r^{10}/1\text{-}r^{10240}) = 79.9\%$.

So exp. profit = .799($102400) - $100 ~ $81700.

Random Walk example.

Suppose you start with 1 chip at time 0 and that your tournament is like a simple random walk, but if you hit 0 you are done. P(you have not hit zero by time 47)?

We know that starting at 0, $P(Y_1 \neq 0, Y_2 \neq 0, \ldots, Y_{2n} \neq 0) = P(Y_{2n} = 0)$.

So, $P(Y_1 > 0, Y_2 > 0, \ldots, Y_{48} > 0) = ½ P(Y_{48} = 0) = ½ \text{Choose}(48,24)(½)^{48}$

$= P(Y_1 = 1, Y_2 > 0, \ldots, Y_{48} > 0)$

= P(start at 0 and win your first hand, and then stay above 0 for at least 47 more hands)

= P(start at 0 and win your first hand) x P(from (1,1), stay above 0 for ≥ 47 more hands)

= 1/2 P(starting with 1 chip, stay above 0 for at least 47 more hands).

So, P(starting with 1 chip, stay above 0 for at least 47 hands) = $\text{Choose}(48,24)(½)^{48}$

= 11.46%.

Bivariate pdf examples.

Suppose $f(x,y) = ax^2 \div y^5$, for x and y between 1 and 3, and $f(x,y) = 0$ otherwise, where $a = 243/520$.

a. What is the marginal density of Y?
b. Are X and Y independent?

a. $f_y(y) = \int f(x,y)\,dx = ay^{-5} \int_1^3 x^2\,dx = ay^{-5}\,[x^3/3]_1^3 = ay^{-5}\,[3^3/3 - 1^3/3]$
$= 26/3\ ay^{-5}$, for y between 1 and 3.

b. $f_x(x) = \int f(x,y)\,dy = ax^2 \int_1^3 y^{-5}\,dy = ax^2\,[-y^{-4}/4]_1^3 = ax^2\,[-1/324+1/4] = 80/324\ ax^2$, for x between 1 and 3.
$f_x(x)\,f_y(y) = 26/3\ ay^{-5} * 80/324\ ax^2 = 520/243\ a^2\,x^2y^{-5} = a\,x^2y^{-5} = f(x,y)$,
so X and Y are independent.

Bayes' rule example.

Your opponent raises all-in before the flop. Suppose you think she would do that *80%* of the time with AA, KK, or QQ, and she would do that *30%* of the time with AK or AQ, and 1% of the time with anything else.

Given only this, and not even your cards, what's P(she has AK)?

Given nothing, P(AK) = 16/C(52,2) = 16/1326. P(AA) = C(4,2)/C(52,2) = 6/1326.

Using Bayes' rule,

$$P(AK \mid \text{all-in}) = \frac{P(\text{all-in} \mid AK) * P(AK)}{P(\text{all-in}|AK)P(AK) + P(\text{all-in}|AA)P(AA) + P(\text{all-in}|KK)P(KK) + \ldots},$$

$$= \frac{30\% \times 16/1326}{[30\%\times16/1326] + [80\%\times6/1326] + [80\%\times6/1326] + [80\%\times6/1326] + [30\%\times16/1326] + [1\%\ (1326-16-6-6-6-16)/1326)]}.$$

(AK) (AA) (KK) (QQ) (AQ) (anything else)

= **13.06%.** Compare with 16/1326 ~ 1.21%.

Conditional probability examples.

Approximate P(SOMEONE has AA, given you have KK)? Out of your 8 opponents?

Note that given that you have KK,

P(player 2 has AA & player 3 has AA)

= P(player 2 has AA) x P(player 3 has AA | player 2 has AA)

= choose(4,2) / choose(50,2) x 1/choose(48,2)

= 0.0000043, or 1 in 230,000.

So, very little overlap! Given you have KK,

P(someone has AA) = P(player2 has AA or player3 has AA or … or pl.9 has AA)

~ P(player2 has AA) + P(player3 has AA) + … + P(player9 has AA)

= 8 x choose(4,2) / choose(50,2) = 3.9%, or 1 in **26.**

What is **exactly** P(someone has an Ace | you have KK)? (8 opponents)

or more than one ace

Given that you have KK, P(someone has an Ace) = 100% - P(nobody has an Ace).

And P(nobody has an Ace) = choose(46,16)/choose(50,16)

= **20.1%.**

So P(someone has an Ace) = **79.9%.**

Suppose $f(x,y) = c\ y\ e^{3x}$, for x and y between 0 and 1.
a) Find the marginal densities of X and Y,
b) find c,
c) find E(X) and E(Y), and
d) find cov(x,y).

a) $f_x(x) = \int f(x,y)\ dy = \int_0^1 c\ y\ e^{3x}\ dy = c\ e^{3x}/2.$
$f_y(y) = \int f(x,y)\ dx = \int_0^1 c\ y\ e^{3x}\ dx = c\ y\ (e^3 - 1)/3.$

b) Since $\int f_x(x)\ dx$ must = 1, we have $1 = c/2 \int_0^1 e^{3x}\ dx = c(e^3 - 1)/6$, so $c = 6/(e^3 - 1)$.
Note that this means $f_y(y) = 2y$.

c) $E(X) = \int_{-\infty}^{\infty} x\ f_x(x)\ dx = c/2 \int_0^1 x\ e^{3x}\ dx$ (integrating by parts)
$= c/2[xe^{3x}/3 - \int e^{3x}/3dx] = c/2\ [e^3/3 - 1/9(e^3-1)] = c/2(e^3/3 - e^3/9 + 1/9) = c(2e^3 + 1)/18.$
$E(Y) = \int_{-\infty}^{\infty} y\ f_y(y)\ dx = c\ (e^3 - 1)/3 \int_0^1 y^2\ dy = 2 \int_0^1 y^2\ dy = 2/3.$

d) cov(x,y) = E(XY) – E(X) E(Y).
$E(XY) = \int_0^1 \int_0^1 (xy)\ c\ y\ e^{3x}\ dx\ dy = [\int_0^1 x\ ce^{3x}/2\ dx\ ][\int_0^1 y\ (2y)\ dy\ ] = E(X)\ E(Y).$
Thus cov(x,y) = 0.

Suppose $f(x,y) = c(x+y)$, for x and y between 0 and 1.
a) Find the marginal densities of X and Y,
b) find c,
c) find E(X) and E(Y), and
d) find cov(x,y).

a) $f_x(x) = \int f(x,y)\, dy = c\int_0^1 (x+y)\, dy = cx + c/2.$
$f_y(y) = \int f(x,y)\, dx = c \int_0^1 (x+y)\, dx = cy + c/2.$

b) Since $\int f_x(x)\, dx$ must $= 1$, we have $1 = c/2 + c/2 = c$, so $c = 1$.
Note that this means $f_y(y) = x + ½$ and $f_y(y) = y + ½$.

c) $E(X) = \int_{-\infty}^{\infty} x\, f_x(x)\, dx = \int_0^1 x\, (x+1/2)\, dx = \int_0^1 x^2\, dx + \int_0^1 x/2\, dx = 1/3 + ¼ = 7/12.$
$E(Y) = \int_0^1 y\, (y+1/2)\, dy = 7/12.$

d) $cov(x,y) = E(XY) – E(X)E(Y) = E(XY) - (7/12)^2.$
$E(XY) = \int_0^1 \int_0^1 xy\, (x+y)\, dx\, dy = \int_0^1 \int_0^1 x^2y + y^2x\, dxdy$
$= (\int_0^1 x^2\, dx)\, (\int_0^1 ydy) + (\int_0^1 y^2\, dy)\, (\int_0^1 xdx)$
$= (1/3)(½) + (1/3)(½) = 1/3.$
Thus $cov(x,y) = 1/3 - (7/12)^2 = 41/144 \sim 0.285.$