Stat 100a: Introduction to Probability.

Outline for the day

1. A quick fact about normals.
2. Random walk example.
3. Review list.
4. Example problems.
5. Tournaments.

Hand in HW3.

http://www.stat.ucla.edu/~frederic/100A/S20 .

Thu is the final exam, here in class, 10am to 11:15am.

Again any notes and books are fine. Read the instructions carefully!

1. A fact about normals.

If X and Y are independent and both are normal, then X+Y is normal, and so are –X and –Y.

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, for a random walk starting at (0,0),

by symmetry $P(Y_1 > 0, Y_2 > 0, \ldots, Y_{48} > 0) = ½\ P(Y_1 \neq 0, Y_2 \neq 0, \ldots, Y_{2n} \neq 0)$

$= ½\ P(Y_{48} = 0) = ½$ Choose(48,24)$(½)^{48}$.

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

= 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, multiplying both sides by 2,

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

= 11.46%.

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 and skill.
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) (ch. 7.1).
20) Bayes's rule (ch. 3.4).
20) Continuous RVs, Uniform, Normal, Exponential and Pareto.
21) Probability density function (pdf). Recall $F'(c) = f(c)$, where $F(c)$ = cdf.
22) Moment generating functions
23) Markov and Chebyshev inequalities
24) Law of Large Numbers (LLN) and Fundamental Theorem of Poker.
25) Central Limit Theorem (CLT)
26) Conditional expectation.
27) Confidence intervals for the sample mean and sample size calculations.
28) Random Walks, Reflection Principle, Ballot Theorem, avoiding zero
29) Chip proportions, doubling up, and induction.
30) Bivariate normal distribution and the conditional distribution of Y given X.
31) Covariance and correlation.

Basically, we've done all of ch. 1-7 except 6.7 and the optimal strategy stuff in 6.3.

What integrals do you need to know?

You need to know $\int e^{ax}\,dx$, $\int ax^k\,dx$ for any k, and $\int \log(x)\,dx$,

and basic stuff like $\int [af(x) + g(x)]\,dx = a \int f(x)dx + \int g(x)dx$,

and you need to understand that $\int\int f(x,y)\,dy\,dx = \int [\int f(x,y)\,dy]\,dx$.

Suppose you have 10 players at the table. What is the expected number of players who have 2 face cards? A face card is a J, Q, K.

P(2 face cards) = C(12,2)/C(52,2) = 4.98%.

Let $X1 = 1$ if player 1 has 2 face cards, and $X1 = 0$ otherwise.

$X2 = 1$ if player 2 has 2 face cards, and $X2 = 0$ otherwise. etc.

$X = \sum Xi$ = total number of players with 2 face cards.

$E(X) = \sum E(Xi) = 10 \times 4.98\% = 0.498$.

Let $X = N(0, 0.8^2)$ and $\varepsilon = N(0, 0.1^2)$ and ε is independent of X. Let $Y = 7 + 0.2\,X + \varepsilon$.

Find E(X), E(Y), E(Y|X), var(X), var(Y), cov(X,Y), and ρ = cor(X,Y).

$E(X) = 0$.

$E(Y) = E(7 + 0.2X + \varepsilon) = 7 + 0.2\,E(X) + E(\varepsilon) = 7$.

$E(Y|X) = E(7 + 0.2X + \varepsilon \mid X) = 7 + 0.2X + E(\varepsilon \mid X) = 7 + 0.2X$ since ε and X are ind.

$var(X) = 0.64$.

$var(Y) = var(7 + 0.2\,X + \varepsilon) = var(0.2X + \varepsilon) = 0.2^2\,var(X) + var(\varepsilon) + 2*0.2\,cov(X,\varepsilon)$

$= 0.2^2(.64) + 0.1^2 + 0 = 0.0356$.

$cov(X,Y) = cov(X, 7 + 0.2X + \varepsilon) = 0.2\,var(X) + cov(X, \varepsilon) = 0.2(0.64) + 0 = 0.128$.

$\rho = cov(X,Y)/(sd(X)\,sd(Y)) = 0.128 / (0.8 \times \sqrt{.0356}) = 0.848$.

Suppose (X,Y) are bivariate normal with E(X) = 10, var(X) = 9, E(Y) = 30, var(Y) = 4, $\rho = 0.3$,

What is the distribution of Y given X = 7?

Given X = 7, Y is normal. Write $Y = \beta_1 + \beta_2 X + \varepsilon$ where ε is normal with mean 0, ind. of X.

Recall $\beta_2 = \rho\, \sigma_y/\sigma_x = 0.3 \times 2/3 = 0.2$.

So $Y = \beta_1 + 0.2\,X + \varepsilon$.

To get β_1, note $30 = E(Y) = \beta_1 + 0.2\,E(X) + E(\varepsilon) = \beta_1 + 0.2\,(10) + 0$. So $30 = \beta_1 + 2$. $\beta_1 = 28$.

So $Y = 28 + 0.2\,X + \varepsilon$, where ε is normal with mean 0 and ind. of X.

What is var(ε)?

$4 = \text{var}(Y) = \text{var}(28 + 0.2\,X + \varepsilon) = 0.2^2\,\text{var}(X) + \text{var}(\varepsilon) + 2(0.2)\,\text{cov}(X,\varepsilon)$
$= 0.2^2\,(9) + \text{var}(\varepsilon) + 0$. So $\text{var}(\varepsilon) = 4 - 0.2^2(9) = 3.64$ and $\text{sd}(\varepsilon) = \sqrt{3.64} = 1.91$.

So $Y = 28 + 0.2\,X + \varepsilon$, where ε is $N(0, 1.91^2)$ and ind. of X.

Given X = 7, $Y = 28 + 0.2(7) + \varepsilon = 29.4 + \varepsilon$, so $Y|X{=}7 \sim N(29.4, 1.91^2)$.

Bivariate and marginal density example.

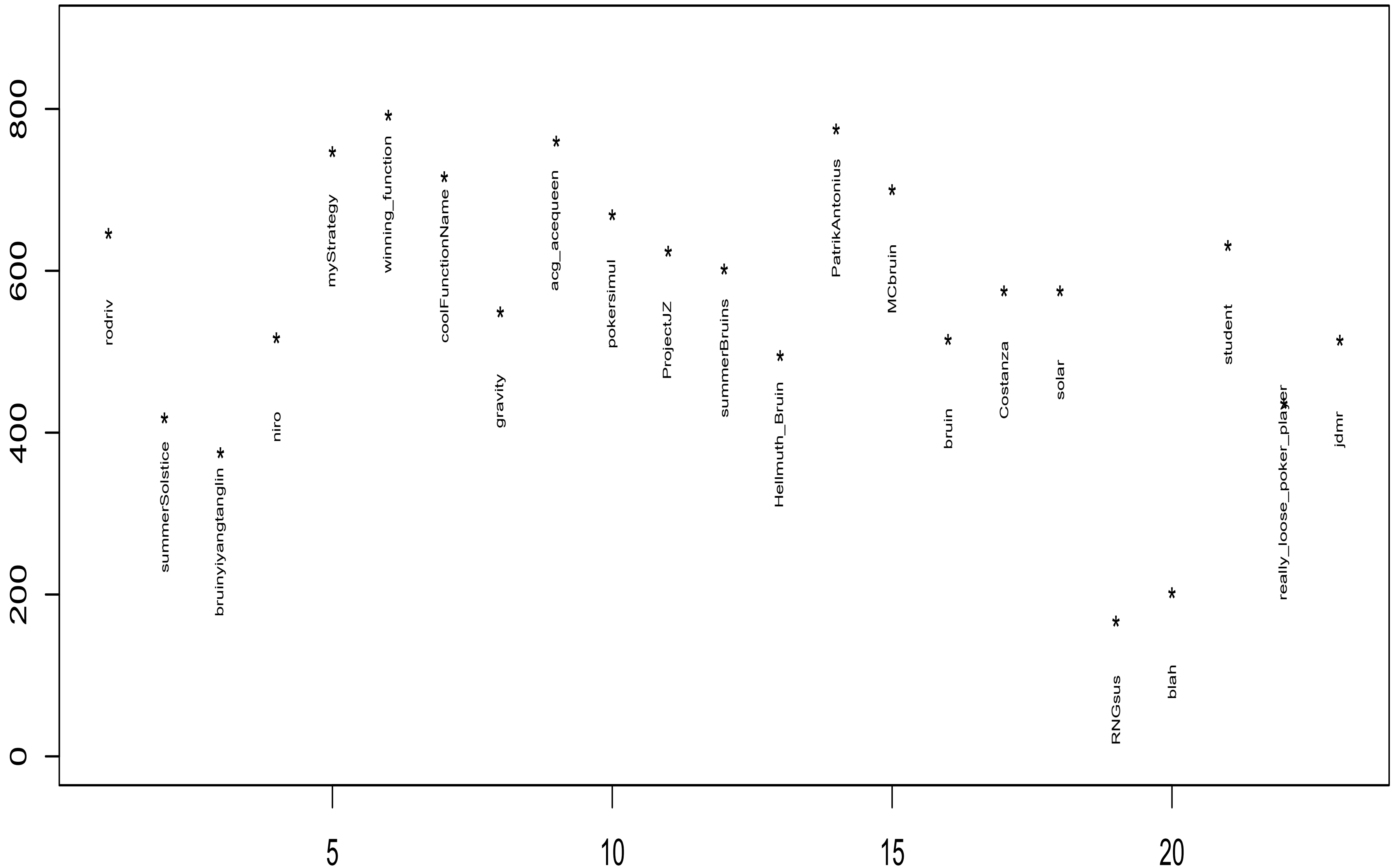
Suppose the joint density of X and Y is f(x,y) = a(xy +x+y), for X and Y in (0,2) x (0,2). What is a? What is the marginal density of Y? What is the density of X conditional on Y? What is E(X|Y)? Are X and Y independent?

$\iint a(xy + x + y)\, dy\, dx = 1 = a\int [xy^2/2 + xy + y^2/2\,]_{y=0}{}^2\, dx = a\int [2x + 2x + 2 -0-0-0]dx$
$= a(x^2 + x^2 + 2x)\,]_{x=0}{}^2 = a(4 + 4 + 4 - 0 - 0 - 0) = 12a$, so $a = 1/12$.
The marginal density of Y is $f(y) = \int_0^2 a(xy+x+y)\, dx$
$= ay \int_0^2 x dx + a\int_0^2 x dx + ay \int_0^2 dx$
$= y/12\ (x^2/2)\,]_{x=0}{}^2 + 1/12\ (x^2/2)]_{x=0}{}^2 + y/12\ x]_{x=0}{}^2$
$= 2y/12 + 2/12 + 2y/12$
$= y/3 + 1/6$.
Check that this is a density. $\int_0^2 (y/3 + 1/6)\, dy = (y^2/6 + y/6)\,]_{y=0}{}^2 = 4/6 + 2/6 - 0 - 0 = 1$.

Conditional on Y, the density of X is $f(x|y) = f(x,y)/f(y) = (xy+x+y) / [12(y/3+1/6)]$
$= (xy+x+y)/(4y+2)$.
$E(X|Y) = \int_0^2 x(xy+x+y)/(4y+2)\, dx = (x^3y/3 + x^3/3 + x^2y/2)/(4y+2)]_{x=0}{}^2$
$= (8y/3+8/3+2y-0-0-0)/(4y+2) = (14y/3 + 8/3)/(4y+2)$.
$f(y) = y/3 + 1/6$ and similarly $f(x) = x/3 + 1/6$,
so $f(x)f(y) = xy/9 + x/18 + y/18 + 1/36 \neq f(x,y)$. So, X and Y are not independent.

Let X = 1 if you are dealt pocket aces and 0 otherwise. Let Y = 1 if you are dealt two black cards and 0 otherwise. What is cov(3X, 7Y)?

cov(3X, 7Y) = 21cov(X,Y).
cov(X,Y) = E(XY) – E(X)E(Y).
E(X) = 1 P(pocket aces) + 0 P(not pocket aces) = C(4,2)/C(52,2) = 0.452%.
E(Y) = 1 P(2 black cards) + 0 P(not 2 black cards) = C(26,2)/C(52,2) = 24.5%.
Here XY = 1 if X and Y are both 1, and XY = 0 otherwise.
So E(XY) = 1 P(X and Y = 1) + 0 P(X or Y does not equal 1)
= P(2 black aces) + 0
= 1 / C(52,2) = 0.0754%.
cov(X,Y) = .000754 - .00452(.245) = -.0003534.
cov(3X, 7Y) = 21 (-.0003534) = -.00742.



0
200
400
600
800
5
10
15
20
rodriv
summerSolstice
bruinyiyangtanglin
niro
myStrategy
winning_function
coolFunctionName
gravity
acg_acequeen
pokersimul
ProjectJZ
summerBruins
Hellmuth_Bruin
PatrikAntonius
MCbruin
bruin
Costanza
solar
RNGsus
blah
student
really_loose_poker_player
jdmr

Suppose X = the number of hands until you get dealt at least one black card. After this, you play 100 more hands and Y = the number of hands where you get dealt pocket aces out of these next 100 hands.

Let Z = 4X + 7Y. What is the SD of X? What is SD(Y)? What is E(Z)? What is SD(Z)?

X is geometric(p), where p = 1 – P(both red) = 1 – C(26,2)/C(52,2) ~ 75.5%. SD(X) = √q/p = 0.656.

Y is binomial(n,p), n = 100 and p = C(4,2)/C(52,2) ~ 0.452%. SD(Y) = √(npq) = 0.671.

E(Z) = 4E(X) + 7E(Y) = 4(1/.755) + 7(100)(.00452) = 8.46.

X and Y are independent so Var(Z) = Var(4X) + Var(7Y) = 16Var(X) + 49Var(Y) = 16(.656^2) + 49(.671^2) = 28.9. So SD(Z) = √28.9 = 5.38.

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 <u>only</u> 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,

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

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

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

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

Some other example problems.

a. Find the probability you are dealt a suited king.

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

b. The typical number of hands until this occurs is ... +/- ....

1/.0362 ~ 27.6.
(√96.38%) / 3.62% ~ 27.1.
So the answer is 27.6 +/- 27.1.

CLT Example

Suppose X1, X2, ..., X100 are 100 iid draws from a population with mean μ=70 and sd σ=10. What is the approximate distribution of the sample mean, $\bar{x}$?

By the CLT, the sample mean is approximately normal with mean μ and sd $\sigma/\sqrt{n}$, i.e. ~ $N(70, 1^2)$.

Now suppose Y1, Y2, ..., Y100 are iid draws, independent of X1, X2, ..., X100, with mean μ=80 and sd σ=25. What is the approximate distribution of $\bar{x} - \bar{y} = Z$?

Now the sample mean of the first sample is approximately $N(70, 1^2)$ and similarly the negative sample mean of the 2nd sample is approximately $N(-80, 2.5^2)$, and the two are independent, so their sum Z is approximately normal.

Its mean is 70-80 = -10,

and var(Z) = $1^2+2.5^2 = 7.25$, so Z ~ $N(-10, 2.69^2)$, because $2.69^2 = 7.25$.

Remember, if X and Y are ind., then var(X+Y) = var(X) + var(Y).

More on luck and skill.

Gain due to skill on a betting round = your expected *profit* during the betting round
= your exp. chips after betting round – your exp. chips before betting round
= (equity after round + leftover chips) –
(equity before round + leftover chips + chips you put in during round)
= equity after round – equity before round – cost during round.

For example, suppose you have A♣ A♠, I have 3♥3♦, the board is
A♥ Q♣ 10♦ and there is $10 in the pot. The turn is 3♣
You go all in for $5 and I call. How much equity due to skill did you gain on the turn?
Your prob. of winning is 43/44.
Your skill gain on turn = your equity after turn bets - equity before turn bets – cost
= ($20)(43/44) - ($10)(43/44) - $5
= $4.77.

Suppose instead you bet $5 on the turn and I folded. How much equity due to skill did you gain on the turn?
$15(100%) - ($10)(43/44) - $5 = $0.23.

♠ ♣ ♥ ♦

Let X = the number of aces you have and Y = the number of kings you have. What is cov(X,Y)?

cov(X,Y) = E(XY) – E(X)E(Y).

$X = X_1 + X_2$, where $X_1 = 1$ if your first card is an ace and $X_2 = 1$ if your 2^{nd} card is an ace,

so $E(X) = E(X_1) + E(X_2) = 1/13 + 1/13 = 2/13$. E(Y) = 2/13.

E(XY) = 1 if you have AK, and 0 otherwise, so E(XY) = 1 x P(AK) = 4x4/C(52,2) = .0121.

So, cov(X,Y) = .0121 – 2/13 x 2/13

= -.0116.